СЕКЦИЯ ПО КОМПЮТЪРНА ЛИНГВИСТИКА

ИНСТИТУТ ЗА БЪЛГАРСКИ ЕЗИК „ПРОФ. ЛЮБОМИР АНДРЕЙЧИН“

БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ

БЪЛГАРСКО-АНГЛИЙСКИ ГРАМАТИЧНИ ПАРАЛЕЛИ С ОГЛЕД НА МАШИННИЯ ПРЕВОД

Дисертация за присъждане на образователната и научната степен „доктор“

Професионално направление
2.1. Филология

Научна специалност
Общо и сравнително езикознание
(Компютърна лингвистика)

Автор:
Тодор Лазаров

Научен ръководител:
проф. д-р Светла Коева

ДЕКЛАРАЦИЯ ЗА ОРИГИНАЛНОСТ И ДОСТОВЕРНОСТ

от Тодор Димитров Лазаров

Във връзка с провеждането на процедура за получаване на образователната и научната степен „доктор“ към Института за български език, Българска академия на науките,

декларирам, че:

1. Резултатите и приносите в научната разработка на тема „***Българско-английски граматични паралели с оглед на машинния превод***“, които представям във връзка с процедурата, са оригинални и не са заимствани от изследвания и публикации, в които нямам участия.

2. Всички чужди публикации, използвани за разработване на дисертацията, са надлежно цитирани и описани в библиографията, приложена към разработката.

3. Научната разработка в цялост, както и нейните отделни части, не са представяни в подкрепа на друга процедура за придобиване на образователна и/или научна степен в образователна или научна организация в България и чужбина.

4. Представената информация във вид на документи, публикации, езикови ресурси и други съответства на обективната истина.

25.02.2020 г. ДЕКЛАРАТОР:

гр. София

(Тодор Димитров Лазаров)

СЪДЪРЖАНИЕ

I. Увод 1

I.1. Актуалност и мотивация на изследването 1

I.2. Цели и задачи на изследването 2

I.3. Обект и предмет на изследването 7

I.4. Методи на изследване 8

I.5. План на изследването 9

II. Българско-английски паралели при граматичната категория време 11

II.1. Глаголът като част на речта 11

II.2. Времето като глаголна категория – концепции и дискусионни въпроси . 13

II.3. Характеристика на глаголите в български 19

II.3.1. Морфологичните категории лице, число и род при глаголите в български 22

II.3.2. Особености на българската темпорална система 24

II.3.3. Семантична характеристика на българската темпорална система 31

II.3.4. Образуване на темпоралните форми в български език 41

II.4. Характеристика на глаголите в английски 43

II.4.1. Категориите лице и число на глаголите в английски 46

II.4.2. Особености на английската темпорална система 46

II.4.3. Семантична характеристика на английската темпорална система 52

II.4.4. Особености на формообразуването на глаголните форми в английски език 60

II.5. Съпоставка на характеристиките на темпоралните системи в български и английски 64

III. Създаване на „Българско-английски паралелен корпус със съотнесени глаголни форми“ 85

III.1. Машинен превод – цели и историческо развитие 85

III.2. Езиковите корпуси като ресурс за машинен превод 97

III.3. Мотивация за избора на корпусния подход при изследването на българско-английските граматични паралели 102

III.4. „Българско-английски паралелен корпус със съотнесени глаголни форми“ 107

III.4.1. Подбор на езиковите ресурси, включени в Корпуса 108

III.4.2. Структура на „Българско-английския паралелен корпус със съотнесени глаголни форми“ 111

III.4.3. Анотационни конвенции при създаването на „Българско-английския паралелен корпус със съотнесени глаголни форми“ 114

III.5. Характеристика на „Българско-английския паралелен корпус със съотнесени глаголни форми“ 122

IV. Анализ на граматичните паралели в „Българско-английския паралелен корпус със съотнесени глаголни форми“ 127

IV.1. „Българско-английският паралелен корпус със съотнесени глаголни форми“ в контекста на изследването 127

IV.2. Статистически обусловен анализ на данните за българските темпорални форми в Корпуса 129

IV.3. Статистически обусловен анализ на данните за английските темпорални форми в Корпуса 135

IV.4. Статистически обусловен анализ на превода на темпоралните форми от български на английски 140

IV.4.1. Сегашно време 140

IV.4.2. Минало свършено време 143

IV.4.3. Минало несвършено време 146

IV.4.4. Минало неопределено време 149

IV.4.5. Бъдеще време 152

IV.4.6. Минало предварително време 154

IV.4.7. Бъдеще време в миналото 157

IV.4.8. Бъдеще предварително време и бъдеще предварително време в миналото 159

IV.5. Обобщен анализ на превода на темпоралните морфологични категории от български на английски 161

IV.5.1. Морфологичната категория време 161

IV.5.2. Морфологичната категория вид на действието 163

IV.5.3. Морфологичната категория таксис 165

IV.5.4. Морфологичните категории лице, число и род 166

IV.5.5. Лексикално-граматичната категория вид на глагола в български 169

IV.6. Изводи за българско-английските граматични паралели на темпорални форми 171

V. Практическо приложение и оценка на „Българско-английския паралелен корпус със съотнесени глаголни форми“ 176

VI. Заключение 181

VI.1. Обобщение 181

VI.2. Приноси на изследването 185

Библиография 187

Приложения 197

I. Увод

I.1. Актуалност и мотивация на изследването

Дигиталната обработка на езикови данни през последните години става все по-популярна и се използва все по-широко не само в областта на лингвистиката, но и в редица свързани области. Една от причините за нарастващата ѝ популярност е, че позволява на изследователите да обработват много големи количества текстови данни и да установяват различни текстови характеристики, общи черти, модели, които може да са трудни или дори невъзможни за установяване чрез класическите методи на анализ. Свързана тенденция, също доста силна и все по-популярна, е създаването на различни езикови корпуси, включващи различни видове езикови данни, за целите на компютърната обработка на езика и създаването на функционални езикови модели.

Актуалността на машинния превод през последните десетилетия е продиктувана от нуждата в глобален аспект от бърз достъп до различна по вид, обем и език информация. В съвременното общество процесите на глобализация водят до нарастване на обмена на информация в реално време. По данни на Европейската комисия в генерална дирекция „Писмени преводи“ работят 1 650 преводачи, които през 2012 г. са превели над 1 760 000 страници, като услугите по писмен и устен превод струват по 2 евро годишно на всеки гражданин от ЕС[1]. Разширяването на географския обхват на страните, активно участващи в интернационални и мултинационални икономически структури, изисква активна работна комуникация на различни езици. Това предполага превод в много кратки срокове на големи обеми информация и е сериозно предизвикателство за професионалните преводачи. Нуждата от качествен и бърз превод налага използването на автоматизирани системи, които да подкрепят и подобряват работата на преводача. В търсене на адекватно решение вниманието се насочва към възможностите, които предлага машинният превод. Основните преимущества на системите за машинен превод са гарантирането на непротиворечивост на превода, подобрена производителност и намалени разходи. Разбира се, системите за машинен превод предполагат непрекъснато подобряване на качеството на превода, като по този начин се намалява необходимостта от редакция, както и възможностите за погрешна интерпретация на текста или неадекватен превод. Съществен недостатък на системите за

[1] https://ec.europa.eu/info/publications/annual-activity-report-2018-translation_en

машинен превод е, че с малки изключения са разработени и достъпни предимно за по-разпространени езици като английски, френски, немски, испански, руски, арабски и други.

Дисертацията е насочена към изследване на българско-английските граматични паралели с оглед на машинния превод. Обектът на изследване е ограничен до темпоралните системи на двата езика: изразяването на темпоралните отношения в български и английски и начините за трансфер на граматичното значение на темпоралните форми от български към английски. От една страна, изследването на тази проблематика във връзка с подходите и методите, използвани при машинния превод, е мотивирано от възможностите, които научното поле предлага за систематичен и формален анализ на комплексни езикови явления, каквито са темпоралните системи на двата езика. От друга страна, липсата на паралелни езикови ресурси, които съдържат информация за съотнасянето на глаголните форми при превод от български на английски език, предполага теоретико-приложния аспект на изследването.

I.2. Цели и задачи на изследването

Основната цел на дисертацията е да се представи единно теоретично описание на темпоралните системи в български и английски и да се проверят теоретично изведените хипотези за преводните еквиваленти посредством статистически анализ на данните от паралелен корпус със съотнесени темпорални форми между български и английски: „Българско-английския паралелен корпус със съотнесени глаголни форми“. Резултатите от изследването могат да бъдат приложени при усъвършенстването на работата на системи за компютърно подпомогнат превод, както и при подбора на едноезикови и паралелни данни за обучение на системи за машинен превод.

За целите на дисертацията се въвеждат следните терминологични уточнения: с термина граматична категория време се означава традиционната деветчленна система на времената в български и шестнадесетчленната система на времената в английски (и инфинитив). Деветте времена в български са: сегашно време, минало свършено време, минало несвършено време, минало неопределено време, минало предварително време, бъдеще време, бъдеще предварително време, бъдеще време в миналото и бъдеще предварително време в миналото, а шестнадесетте времена в английски са: сегашно просто време (Present Simple Tense), сегашно продължително време (Present Continuous

Tense), сегашно перфектно време (Present Perfect Tense), сегашно перфектно продължително време (Present Perfect Continuous Tense), минало просто време (Past Simple Tense), минало продължително време (Past Continuous Tense), минало перфектно време (Past Perfect Tense), минало перфектно продължително време (Past Perfect Continuous Tense), бъдеще просто време (Future Simple Tense), бъдеще продължително време (Future Continuous Tense), бъдеще перфектно време (Future Perfect Tense), бъдеще перфектно продължително време (Future Perfect Continuous Tense), бъдеще просто време в миналото (Future Simple Tense in the Past), бъдеще продължително време в миналото (Future Continuous Tense in the Past), бъдеще перфектно време в миналото (Future Perfect Tense in the Past) и бъдеще перфектно продължително време в миналото (Future Perfect Continuous Tense in the Past).

Приема се, че посредством граматичната категория време (още хиперкатегорията време) се изразяват различни (темпорални) отношения и тя може да се разглежда като съвкупност от три морфологични категории в български: време, вид на действието и таксис, и две морфологични категории в английски: време и вид на действието. Морфологичната категория време (наричана още същинска категория време) в български „се основава на отношението на действието към изказването“ (Куцаров 2007: 244), което отношение може да изразява „едновременност или разновременност“ (неедновременност). Морфологичната категория време (наричана още същинска категория време) в английски език „изразява едновременност или неедновременност на действието спрямо момента на говоренето“ (Кабакчиев 2003: 132). Морфологичната категория вид на действието (резултативност) в български изразява „отношение между действие и резултат от действие в един и същ ориентационен момент“ (Куцаров 2007: 249), който може да бъде сегашен, минал, бъдещ или бъдещ спрямо минал. Морфологичната категорията вид на действието (вид) в английски изразява отношението между „продължителността или завършеността на действието към [...] ориентационен момент“ (Валейка и Буиткийн 2003: 84), сегашен, минал, бъдещ или бъдещ спрямо минал. Необходимо е да бъде уточнено, че категорията вид на действието в български и категорията вид на действието в английски представляват две различни категории с различна семантика и категориална структура, въпреки съвпадението на термините, които са възприети. Морфологичната категория таксис в български изразява „отношенията на действията или резултатите от действия, ориентирани към миналия

ориентационен момент (минало време или друг сигнализатор), към действията, неориентирани към миналия ориентационен момент“ (Куцаров 2007: 255).

Представените дефиниции са възприети като най-ясно дефиниращи значението на съответните морфологични категории за български и английски. Тъй като нито един от авторите не дава дефинициите във връзка със съпоставката на темпоралните системи на двата езика, не може да се говори за обобщения, които имат типологичен характер.

С термина темпорални отношения се означават отношенията, изразявани от трите морфологични категории: време, вид на действието и таксис в български, и двете морфологични категории: време и вид на действието в английски. С термина темпорална система се означава системата от морфологичните категории време, вид на действието и таксис в български и време и вид на действието в английски. С термина темпорални форми се означават глаголните форми в двата езика, които са част от съответните темпорални системи.

Основната цел на дисертацията се постига последователно посредством целите и задачите, чието изпълнение е описано в отделните глави на дисертацията.

Целта на Втора глава „Българско-английски паралели при граматичната категория време“ е да се изгради единна концепция за изследване на темпоралните форми в двата езика, да се представи формално-семантично описание на темпоралните значения както поотделно в двата езика, така и съпоставителен план и да се създаде хипотеза за възможните съответствия на семантично равнище. Формално-семантично описание на глаголните форми, натоварени с темпорално значение, ще бъде използвано при изграждането на хипотеза за българско-английските граматични паралели на равнище темпорални форми.

Конкретните задачи за постигането на поставената цел във втората глава от дисертацията са:

1. Да се разгледат и съпоставят различни схващания за морфологичните категории време, вид на действието и таксис в български и време и вид на действието в английски.

2. Да се създаде единна за целите на настоящото изследване концепция за описание на темпоралните системи в български и английски език, морфологичните категории, които ги изграждат, и отношенията между членовете им.

3. Да се направи формално-семантично описание на темпоралните значения в двата езика.

4. Да се обособят граматичните паралели между формално-семантичните описания на темпоралните значения в двата езика, на основата на които ще бъде създадена хипотеза за преводните съответствия и семантичния трансфер на граматична информация при превод.

Целта, към която е ориентирано изследването в Трета глава „Създаване на „Българско-английски паралелен корпус със съотнесени глаголни форми“, е създаването на лингвистично анотиран паралелен корпус, съдържащ информация за преводното съотнасяне на българските темпорални форми с английските им съответствия. За осъществяването на поставената цел е необходимо да бъдат изпълнени следните задачи:

1. Анализ на основните характеристики на машинния превод и изискванията към езиковите ресурси, които се използват за машинен превод.

2. Преглед на съществуващите паралелни ресурси за български и английски с цел подбор на подходящите за целите на изследването ресурси.

3. Формулиране на критерии за изграждането на „Българско-английския паралелен корпус със съотнесени глаголни форми“.

4. Формулиране на анотационни конвенции за обогатяването на „Българско-английския паралелен корпус със съотнесени глаголни форми“ с лингвистична информация за съотнесените глаголни форми.

5. Практическа работа за създаване и анотиране на „Българско-английския паралелен корпус със съотнесени глаголни форми“.

6. Описание на основните характеристики на „Българско-английския паралелен корпус със съотнесени глаголни форми“.

Основната цел на изследването в Четвърта глава е да се представи статистически обусловен анализ на темпоралните форми в български и английски както поотделно, така и в съпоставителен план. Анализът е в няколко направления: по отношение на данните за българските и английските темпорални форми поотделно; по отношение на трансфера на информация за морфологичните категории време, вид на действието и таксис в български към английските морфологични категории време и вид на действието, както и по отношение на трансфера на информацията за морфологичните категории лице, число

и род и лексикално-граматичната категория вид на глагола в български при превод на английски език. Цели се проверка на теоретично изведената хипотеза за съответствието при формално-семантичния анализ на двата езика и наблюдаваните езикови факти при превода.

Конкретните задачи за постигането на поставената цел в четвъртата глава от дисертацията са:

1. Извличане на данни и статистически обусловен анализ на българските темпорални форми, разделени по темпорални категории и тематични области в Корпуса.

2. Извличане на данни и статистически обусловен анализ на английските темпорални форми, разделени по темпорални категории и тематични области в Корпуса.

3. Статистически обусловен анализ на превода на темпоралните форми от български на английски.

4. Формулиране на изводи за българско-английските граматични паралели на темпоралните форми.

5. Представяне на основните резултати от анализа на данните в „Българско-английския паралелен корпус със съотнесени глаголни форми“ в контекста на теоретичното описание на българската и английската темпорална система.

Друга основна цел на изследването е приложение на данните от Корпуса за създаването на преводни модели, обогатени с лингвистична информация за съотнасянето на български и английски темпорални форми. По-конкретно, практическо приложение на „Българско-английския паралелен корпус със съотнесени глаголни форми“ в система за компютърно подпомогнат превод. За изпълнението на тази цел се реализират следните задачи, представени в Пета глава „Практическо приложение и оценка на „Българско-английския паралелен корпус със съотнесени глаголни форми“:

1. Създаване на „терминологична“ база от данни, съдържаща информация от „Българско-английския паралелен корпус със съотнесени глаголни форми“.

2. Провеждане на експеримент с терминологичната база от данни при компютърно подпомогнат превод.

3. Оценка на резултатите от експеримента.

В заключение, целта на дисертацията може да бъде представена като съвкупност от поставените цели по отношение на няколко концептуално отделни части. От една страна, дефинирането на единна концепция за описание на темпоралните системи и формообразуването в български и английски е необходима предпоставка за съпоставянето на двата езика на теоретично равнище. От друга страна, създаването на хипотеза, базирана на теоретичното описание на темпоралните системи на двата изследвани езика, позволява да се направи оценка на извлечените статистически езикови данни от „Българско-английския паралелен корпус със съотнесени глаголни форми“.

По-дългосрочна цел (извън рамките на дисертацията) е на базата на дефинираните характеристики на изследваната езикова двойка да се потърсят езиково зависими начини за оптимизиране на статистически преводен модел за превод на българските темпорални форми на английски.

I.3. Обект и предмет на изследването

Обект на изследването са съвременните темпорални системи в български и английски, начините на изразяване на хиперкатегорията време в двата езика, семантиката и трансфера на морфологичните категории време, вид на действие и таксис в български към морфологичните категории време и вид на действието в английски. Предмет на изследване са конкретните граматични паралели между темпоралните отношения в двата езика, засвидетелствани в паралелен корпус, създаден за нуждите на проучването.

Предназначението на „Българско-английския паралелен корпус със съотнесени глаголни форми“ е да представя изчерпателна и достоверна езикова информация, приложима за създаването на статистически модел за превод на глаголните форми. Необходимо е да се направи уточнението, че при създаването на „Българско-английския паралелен корпус със съотнесени глаголни форми“ не са отразени евиденциалността, наклонението, залогът, положителният или отрицателният статус. Дефинирането на езикова посока (от български към английски) е продиктувано от комплексността на българското глаголно формообразуване в сравнение с английското.

I.4. Методи на изследване

В последните години се откроява подчертана тенденция към формално-семантично описание на свойствата на езика и към приложение на това описание за практически цели. Разглеждането на хиперкатегорията време в двата езика предоставя възможността за паралелен формално-семантичен анализ на значенията на темпоралните категории.

Трябва да се отбележи, че изследването във формален план на особеностите на формообразуването в български и английски (както отделно за двата езика, така и в съпоставителен план) не е новост; съществуват изследвания, акцентиращи върху по-общи или върху по-специфични проблеми; известни са и опити за представянето на формални граматични паралели за превод между български и английски – както за глаголите, така и за останалите части на речта (Атанасова и др. 1963, Стамболиева 2015). В дисертация проблемът за българско-английските граматични паралели между темпоралните системи в двата езика се разглежда чрез методите на формално-семантичния анализ, корпусния анализ и статистическото моделиране.

Избраните методи за анализ на българско-английските граматични паралели на равнището на темпоралните системи на двата езика са тясно обвързани с компютърната лингвистика и по-специално с машинния превод. Изследването е повлияно от съвременните тенденции за систематичен анализ на езикови феномени чрез прилагане на методи от компютърната лингвистика и статистическото моделиране (Паскалева 1987: 32). Актуалността на статистическото моделиране предполага интегриране на съвременните тенденции за изследване на езика и класическите теоретични изследвания. Към съвременните тенденции за изследване на езиковите феномени трябва да бъде причислено разработването и създаването на езикови ресурси, което се подчинява на определени критерии, целящи предоставянето на надеждни данни за научен анализ.

Граматичните паралели между български и английски, и по-специално формално-семантичното описание на различията и сходствата на темпоралните системи в двата езика, се подлага на емпирична верификация на базата на статистически данни, извлечени от „Българско-английския паралелен корпус със съотнесени глаголни форми“. За целта корпусният анализ се съчетава със статистическото моделиране.

При машинния превод се разчита на представителни паралелни езикови корпуси с достатъчен обем, които представят лингвистичните феномени в техните различни проявления. За създаването на статистически езиков модел за превод е необходимо да се знае каква е вероятността дадена езикова единица (дума, фраза) е в целевия език да бъде превод на езикова единица f от изходния – p(e|f). За целта са необходими алгоритми, които да обработват информацията, и езикови ресурси, от които да се извлекат лингвистичните данни. За български и английски са налични различни по вид свободно достъпни езикови корпуси. Както е известно, „корпусите [...] представляват надежден източник за наблюдение, анализ и изводи (подкрепени от обективни количествени и дистрибутивни данни) за [...] различни езикови явления (за употребата им в различни стилове, жанрове и тематични области), както и за автоматично извличане на езикови данни, езикови отношения и модели“ (Коева 2014: 49). Граматичните паралели между български и английски и по-специално формалното описание на различията и сходствата на темпоралните системи в двата езика се основава на анализа на данните от „Българско-английския паралелен корпус със съотнесени глаголни форми“, създаден за нуждите на изследването.

Достоверността на резултатите е проверена експериментално с помощта на софтуер за компютърно подпомогнат превод.

I.5. План на изследването

Дисертацията се състои от шест глави, библиография, едно приложение, за което е предоставен хиперлинк, и е в обем от 197 страници.

Във Втора глава на дисертацията, озаглавена „Българско-английски паралели при граматичната категория време“, ще бъдат разгледани различни схващания за характеристиките на граматичната категория (хиперкатегорията) време в български и английски език. Ще бъде направен анализ на теоретичните постановки, имащи пряко отношение към изследването. Ще бъдат представени семантичните характеристики на морфологичните категории, влизащи в състава на хиперкатегорията време, както и формообразуването в български и английски. Като основен резултат ще бъде представена съпоставката на формално-семантичното представяне на темпоралните категории в двата езика.

В Трета глава, озаглавена „Създаване на „Българско-английски паралелен корпус със съотнесени глаголни форми“, ще бъде направен кратък преглед на развитието на машинния превод и съвременните практики при създаването на лингвистични ресурси за целите на статистическото езиково моделиране и статистическия машинен превод. В Трета глава ще бъде представен и основният резултат от изследването – „Българско-английският паралелен корпус със съотнесени глаголни форми“: включените в него ресурси, структурата на Корпуса, включените метаданните, анотационните конвенции, работата по конструирането на Корпуса, както и актуалното му състояние.

В Четвърта глава, озаглавена „Анализ на граматичните паралели в „Българско-английския паралелен корпус със съотнесени глаголни форми“, последователно се представят и анализират подробни статистически данни за темпоралните форми в български и английски, разпределени в 4 тематични области, както и реализацията на българските темпорални категории при превод на английски. Изградената хипотеза за трансфера на граматична информация при превода на темпорални форми на базата на формално-семантичен анализ ще бъде апробирана посредством извлечените статистически данни от Корпуса.

В Пета глава, озаглавена „Практическо приложение и оценка на „Българско-английския паралелен корпус със съотнесени глаголни форми“ ще бъде представен експериментът, направен със софтуер за компютърно подпомогнат превод, базиран на данните, извлечени от Корпуса, за предписване на най-вероятна форма при превод на темпорални форми от български на английски. В същата глава ще бъде представена и оценката, направена от професионални преводачи, по отношение на проведения експеримент.

II. Българско-английски паралели при граматичната категория време

Граматичната категория време е представена първо поотделно за български и английски, а след това е направена съпоставка. Описва се съотнасянето между средствата за изразяване на морфологично и семантично ниво на темпоралните отношения в български и английски език. Представени са съответствията между морфологичните категории при превод на глаголните форми от български на английски и корелацията между отделните морфологични категории в рамките на хиперкатегорията време.

II.1. Глаголът като част на речта

Един от основните въпроси при описанието на езиковата система е свързан със спецификата и мястото на глагола в тази система. В това отношение предмет на многобройни изследвания е думата, която Сосюр (цит. по Станков 2009: 11) определя като „централен фактор в механизма на езика“. Принадлежността на определена лексикална единица към определена част на речта е свързана, от една страна, с нейната семантика, а от друга страна, с функционирането ѝ в езика.

Опитите за дефинирането на частите на речта могат да бъдат формално разделени на класически, свързани с абстракция по отношение на семантиката на определени думи, които влизат в съответния клас, и модернистични, свързани с опитите за по-задълбочено формално описание на езиковата система и вътрешните зависимости между градивните ѝ части.

Глаголът като част на речта присъства в теоретичните описания на езика още от дълбока древност, като мястото му е неизменно, независимо от концептуалните различия (Йесперсен 1965, Комри 1981). Подходите за дефинирането на глагола като част на речта могат да бъдат представени обобщено в две основни направления – семантично-описателно и функционално-структурно. От гледна точка на семантичната класификация на класовете думи глаголът се описва като част на речта, „която означава действие или състояние, свързано с глаголно лице“ (ГСБКЕ II: 209). Могат да се извлекат следните заключения – първо, семантиката на глагола е свързана с означаването (но не назоваването) на действие или състояние; второ, тази семантика трябва да бъде съотнесена със субект, за да бъде реализирана; трето, съвкупността от назованото

действие или състояние и неговия субект трябва да бъде разположена във времето. Подобен вид описание има своите преимущества, но не винаги е подходящо за формален систематичен анализ на особеностите на глаголите.

Съвременните концепции за описание на значението се опитват да избегнат посочения от Ф. дьо Сосюр феномен в езикознанието, при който обектът на изследване (езикът) бива изследван чрез самия себе си (Сосюр 1992: 16). По-новите подходи за класификация на глаголите като клас думи съвместяват класическите виждания за семантичните особености на думите и съвременните опити за формално описание на функционалните им характеристики. Езиците, в които частите на речта са различими, могат да бъдат разделени на три типа според наличния инвентар от класове думи (Крофт 1991: 58): езици с пълен и ясно изразен инвентар – NAV езици, които имат ясно изградени принципи за разграничаване на имена (N), прилагателни (A) и глаголи (V) (такива са повечето индоевропейски езици, заедно с български и английски); езици с инвентар, при който се противопоставят имена и неимена – N[AV]; езици, които имат структурно изразени имена, а характеристиките на останалите части на речта не са ясно разграничими (Бхат 1994) (такива са езиците кора и салишан); езици с инвентар, при който се противопоставят глаголи и неглаголи – V[AN] (например кечуа). Също така съществуват езици, за които класификацията по части на речта не е приложима.

От гледна точка на съвременното формално описание на естествения език класификационните характеристики на глаголите могат да бъдат разглеждани като взаимозависими и йерархично подредени отношения. При тези класификации съществува строго определена подредба между значенията на частите на речта, при която елементите са подредени йерархично според изграждащите ги семантични части.

Ако се приеме, че глаголът означава определен вид признаци – признаци от първи род, както ги нарича О. Йесперсен (Йесперсен 1965: 34), то разликата между тази част на речта и другите, означаващи също признаци (прилагателно, числително, наречие), е в това, че признаците, изразявани от глаголите, са динамични, локализирани по някакъв начин на определено място по темпоралната ос. От това следва, че глаголът може да бъде определен като „част на речта, която означава (но не именува) действия или състояния като динамични признаци на явления, назовани със съществителни имена“ (Куцаров 2007: 103).

II.2. Времето като глаголна категория – концепции и дискусионни въпроси

Описанието на времето като феномен отдавна е обект на изследване както във философията, така и в лингвистиката. Признавайки необходимостта от наблюдател за съществуването на времето, Аристотеловата традиция представя времето като „непрекъсната линия, която тече от ляво надясно, т.е. от миналото към бъдещето“ (Манинг и др. 2013: 233). Това линейно представяне на времето се свързва с обективното схващане за времето като способ за измерване на естествените процеси. Посочва се (Манинг и др. 2013: 235), че схващанията на св. Августин за времето и паметта бележат фундаментална промяна в представянето на времето като център на човешкото съзнание и познание чрез въвеждането на понятието за субективно време. По този начин при субективно представяне на реалността миналото и бъдещето съществуват само ако присъстват в човешкото съзнание. „Миналото време се възстановява чрез паметта, а бъдещото време се очаква от емпиричното настояще“ (Аюн и др. 2018: 19).

От гледна точка на лингвистиката „настоящето на предстоящите събития съответства на бъдещето, внедрено в настоящето“ (Флашман 1991: 298). Лингвистичната традиция е силно повлияна от Аристотеловата концепция за времето. Х. Путсма (Путсма 1922: 7) и О. Йесперсен (Йесперсен 1924: 257) застъпват становището, че времето е линейно. В по-ново време при невролингвистичните подходи се подчертава, че обективното време не може да бъде съпоставено със субективното време, тъй като „човешкото възприемане на времето зависи от динамиката на когнитивните процеси“ (Васенхов 2017: 167).

С оглед на граматичното време се наблюдава съгласие между лингвистите за неговата семантична природа. Времето е граматична категория, която „граматикализира протичането на събития или състояния във времето чрез локализиране на възможните събития“ (Комри 1985: 9). По-точно, граматичното време локализира възможни събития по отношение на момента на изказване, т.е. то е деиктическа категория (Аюн и др. 2018: 27). Граматичното време обикновено се разглежда като „флективна категория“ (Комри 1985: 14). Граматичното време се различава от наративното (текстовото) време, което цялостните езикови продукти (текстовете) моделират като неотделим атрибут на представяния в тях свят (Хъдълстън и Пулъм 2002: 116). Наративното време не следва да се отъждествява със съвкупността на глаголните форми за време в текста (Аюн и др. 2018: 29). Необходимо е ясно да се разграничават граматичната категория време и

времето като категория на семантиката на текста. Граматичното време е начин за изразяване на съдържателната категория текстово време, но далеч не е единственият начин. Текстовото време има, на първо място, онтологичен аспект (представлява цялостен смислов конструкт от времеви параметри и връзки на обектите и фактите от текстовия свят) и на второ място, езиков аспект (представя се в езиковата форма на речевия продукт чрез цяла редица от изразни средства) (Савова 2007). В дисертацията обект на изследване е граматичното време, като е необходимо да се направи уговорката, че в повечето случаи трансферът на граматикализираната информация при превод зависи от по-широкия текстов контекст.

Както в българското езикознание, така и в изследванията за английски език категорията време винаги е била дискусионна по отношение на обхвата на категорията, инвентара от граматични значения и кодирането на граматична информация. Към категорията време са били причислявани други граматични категории, които изразяват самостойни значения и изцяло модифицират значението на категорията време, но не могат да бъдат включени в нейната парадигма. В наши дни е възприето схващането за т. нар. „хиперкатегории“, съставени от взаимосвързани категории със собствени граматични парадигми, които са с кумулативно значение по отношение на формообразуването (Чергова 2014: 195). Съвременната езиковедска наука дава достатъчно основания за интерпретиране на класическите глаголни категории не като една цялостна граматична стойност, а като йерархични структури от множество значения, изразявани от различни езикови форми на различни структурни равнища на езика (Бондарко и др. 1987: 84). Все по-продуктивни от функционална гледна точка се оказват похватите, които разглеждат глаголните категории като „сложни функционално-семантични структури с разнообразни езикови инструменти“ (Вейга 1991: 56) или хиперкатегории, при които „между крайните елементи на отношението, изразявано от съответната проста категория, се вмъкват един или повече елементи посредници“ (Герджиков 2000). Дж. Грийнбърг (Грийнбърг 1971: 306–310) определя йерархичността на граматичните категории като една от универсалиите на езика и извежда постулата, че граматична категория, изградена върху маркирания елемент на друга граматична категория, разполага с равен или по-малък брой форми от главната категория. Може да се твърди (Чергова 2014: 200), че „категориалните отношения се организират около даден семантичен континуум: отношение на говорещия към изказаното действие (модалност); отношение на изказаното действие към момента на

говорене (темпоралност); отношение на глаголното действие към собственото му вътрешно протичане (аспектуалност) и т. н., но този семантичен континуум не изчерпва всички възможни подзначения, които може да се включат в него, нито се ограничава до дадена форма“. По този начин само част от категориалното съдържание се систематизира под формата на специфични грамеми, а „възможността да се изразяват допълнителни стойности в рамките на този семантичен континуум чрез различни средства остава отворена“ (Чергова 2014: 202).

Глаголните категории в езиковата система може да не се идентифицират с един-единствен категориален признак. Категорията време се отъждествява със семантичния континуум „отношение на изказаното действие към момента на говорене“ (Чергова 2014: 203), но в рамките на това значение може да се обособят подзначения за пряко и непряко отношение на действието към момента на говорене, което формира две подкатегории в темпоралността: глаголно време и таксис (Чергова 2012: 18 – 19).

В това отношение се намесва и разликата, т. е. йерархичността, зададена от синтетичната и от аналитичната реализация на категориалните значения. Аналитичните категориални форми внасят допълнително значение, което се наслагва върху синтетично изразените категориални признаци, като задават ново категориално отношение в рамките на общия семантичен континуум. Това оформя темпоралността като сложна и многопластова категория, в която основните значения се реализират от синтетични изразни средства, а надстроените – от аналитични изразни средства (Чергова 2012: 19 – 21). Категорията време продължава да се изгражда на синтагматично, синтактично и дискурсно ниво чрез други езикови средства (наречия за време, синтактични конструкции, дискурсно време), чийто обект надхвърля морфологичното равнище (Чергова 2014: 190). Допълнителна възможност за изразяване на темпорални отношения е използването на други езикови средства, например, наречия за време като „вчера“ или „утре“ или изрази, съдържащи предлози като „преди“ или „след“. Тези изразни средства нямат същия семантичен статус като граматичното време: те са лексикални, а не граматични маркери за темпоралност.

Йерархичността като организация на глаголните категории е убедително представена от Г. Герджиков (Герджиков 1984: 191 – 193). Хиперкатегориите се отличават от простите категории по това, че не изразяват само отношения между едни и същи крайни елементи. В значенията на отделните членове на хиперкатегорията освен отношенията между крайните елементи може да се включват и елементи, които

опосредстват съответните отношения между крайните елементи. „Сложните категории (хиперкатегориите) се състоят от синкатегории, които нямат равноправно положение в системата поради елементите, които изграждат значенията им (само крайни елементи, крайни елементи и елементи посредници и т.н.)“ (Алексова 2003).

Терминът „граматично време“ може да бъде дефиниран като вторична граматична категория, която служи за „локализиране на събитие или ситуация във времето спрямо момента на говорене“ (Комри 1985: 32). Значението на термина обхваща два аспекта: морфологичен аспект – системата от граматични времена, изразявани от морфологичните характеристики на глагола, и семантичен аспект, отнасящ се до времевото разполагане на събитието или събитията, за които се говори в речевия акт – т.е. „значението“ на различните времена (Чергова 2014: 203). Важен въпрос е дали различните времена в тази система изразяват само отношението между момента на говорене и събитието, като по този начин локализират събития във времето, или изпълняват и други функции като съотнасяне на действията едно с друго и изразяване на завършеността на действията.

За описанието на семантичния аспект на граматичното време ще бъде представена системата на Х. Райхенбах (Райхенбах 1947). Ще бъдат анализирани характеристиките на граматичното време и ще бъде представено формалното съотнасяне между морфологичните и семантичните особености на отделните времена.

Ако речевият акт бъде означен като **момент на говорене**, от гледна точка на темпоралната логика са възможни три времеви отношения на момента на събитието спрямо момента на говорене: „преди момента на говорене“, „едновременно с момента на говорене“ и „след момента на говорене“ (Райхенбах 1947: 287). Трите темпорални отношения не могат да бъдат съотнесени еднозначно към морфологично изразените времена, тъй като броят на глаголните времена в езиците е различен. Нужна е интерпретация с по-комплексни ориентации. Ако се вземе за пример изречение като „Иван гледаше телевизия“, се вижда, че граматичното време засяга събитие, което се определя по отношение на момент, различен от момента на говоренето. Момент, който „не съвпада с момента на говорене и служи за ситуиране на момента на събитието спрямо него, се нарича **отправна точка**“ (Райхенбах 1947: 288). В примерното изречение моментът на събитието е времето, когато се случва „гледането“, и той съвпада с отправната точка, която е преди момента на говорене. При примери като дадения по-горе не винаги е ясно дали моментът на събитието е в отношение с момента на говорене, или

с друга отправна точка и може да се конкретизира в зависимост от контекста на речевия акт. Граматичното време в много строг смисъл съдържа само две семантични стойности: едновременност и неедновременност (Райхенбах 1947: 291). Други темпорални отношения се формират с допълнителни категории като аспект или относителност. Що се отнася до аспекта, „трудно е да се направи разграничение между граматичното време, от една страна, и аспекта, от друга“ (Хорнщайн 1990: 7). Аспектът е граматична категория на глагола, която също има отношение към темпоралността, но за разлика от граматичното време, тази категория има повече общо с темпоралната структура на самото събитие или ситуация, като по този начин липсва релационният характер на граматичното време. Разграничаването на аспекта и граматичното време не е еднотипно реализирано при различните езици, макар че те могат да се разглеждат като взаимозависими и обвързани (Хорнщайн 1990: 9). В най-общ смисъл аспектът представлява „допълнителен маркер на глаголното събитие по отношение на момента на референция (говорене)“ (Накимовски 1988: 12). Аспектът в български е лексикално-граматична категория и формално остава извън рамките на хиперкатегорията време, като обаче си взаимодейства с нея (посредством морфологичната категория вид на действието), а в английски се изразява от категорията вид на действието, а в някои случаи и на лексикално равнище (подробният анализ на начините на изразяване на аспектуалността в български и английски е извън обхвата на дисертацията).

Както бе споменато по-горе, граматичните времената изпълняват функция за локализиране на събития във времето. Те могат да бъдат представени като „семантични оператори или функции на пропозиции, означаващи определен момент или период време, на които може да се определи една истинна стойност“ (Ватер 1994: 17). Тъй като точката във времето или времевият период, които характеризират действията или състоянията, са в отношение с времето на говорене, граматичното време може да бъде разглеждано и като деиктична категория (Райхенбах 1947: 291). Под деиксис се разбира местоположението и идентификацията на лица, предмети, събития, процеси и дейности, за които се говори или се споменава във връзка с пространствено-временния контекст, създаден и поддържан от речевия акт и участието в него (Лионс 1977: 637). Граматичните времена са деиктически, тъй като определят темпоралните отношения спрямо времето на говорене (Райхенбах 1947: 291).

За нуждите на формално представяне на темпоралната система на Х. Райхенбах ще се използва съкращението S, за да се означи моментът (времето) на говорене. Времето на

събитието, изразявано от глагола, ще бъде означено с E. Има две възможни отношения между S и E: отношение на последователност с две подразделения – E предхожда S и обратно; и отношение на едновременност – E и S се случват по едно и също време и S се включва в E. По този начин може да бъде описана следната опростена темпорална система (Райхенбах 1947), представена в Таблица 1:

Отношение	**Значение**	**Време**
E – S	Моментът на събитието е преди момента на говорене	Минало време
S – E	Моментът на събитието е след момента на говорене	Бъдеще време
E, S	Моментът на събитието съвпада с момента на говорене	Сегашно време

Таблица 1. ***Описание на значенията в темпорална система на Х. Райхенбах.***

Много важна характеристика на темпоралната система на Х. Райхенбах е, че за разлика от бинарната структура, представена по-горе, тя се характеризира с три темпорални точки: вече дефинираните точки E и S и така наречената „отправна точка“ (R) (Райхенбах 1947: 293). Освен три темпорални точки (S, E и R) системата включва и два вида отношения между тях (едновременност и последователност). От решаващо значение е, че тези типове отношения могат да бъдат само между точките R и S или R и E, отношението между S и E е непряко и е изградено на базата на другите отношенията в системата (Клайн 1994: 54). Следователно отправната точка R може да се разглежда като свързващ възел в цялата система. Отношенията между S и R, от една страна, и E и R, от друга страна, според системата на Х. Райхенбах (Райхенбах 1947: 295) са представени в Таблица 2:

Отношение	***Значение***	***Наименование***	***Отношение***	***Значение***	***Наименование***
S, R	S е едновременно с R	Настояще	E, R	E е едновременно с R	Едновременност
S – R	S предхожда R	Бъдеще	E – R	E предхожда R	Предходност
R – S	R предхожда S	Минало	R – E	R предхожда E	Следходност

Таблица 2. ***Описание на отношенията между елементите в системата на Х. Райхенбах.***

Важна особеност на тази система е фактът, че всички възможни времена са дефинирани с помощта на отправна точка – дори тези с просто съотнасяне, които могат да бъдат дефинирани само с E и S.

Първоначално приложена за английски език, системата на Х. Райхенбах позволява конструирането на 13 времена (Моллов 2014: 73). За разлика от системите,

получени въз основна на класическите описания на морфологичните времена на даден език, системата на Х. Райхенбах е по-независима, тъй като „функционира като вид автономна логическа конструкция, лежаща в основата на морфологичните времена“ (Ницолова 2008: 264). Тази независимост може да се види във факта, че системата на Х. Райхенбах съдържа конструкции, които не са морфологично реализирани на английски език, но могат да бъдат реализирани на други езици. Под една или друга форма характеристиките на системата на Х. Райхенбах днес са включени в много теории, занимаващи се с описанието на морфологичната и логическата структура на темпоралните системи. Направени са и редица предложения относно изменението и подобряването на тази система. Б. Комри (Комри 1985: 113) дефинира разграничаването на „абсолютни времена“ (прости времена) и „относителни времена“ (сложни времена) и твърди, че само последните трябва да бъдат представени чрез използване на референтна точка. Друг подход е представен от А. Приор (Приор 1967), който въвежда концепцията за множество референтни точки и предлага структури като S – R1 – E – R2.

С оглед на поставените цели в първата глава на дисертацията за формално описание и съпоставителен анализ в следващите части на Втора глава ще бъдат представени темпоралните системи на български и английски според системата на Х. Райхенбах. В резултат от съпоставителния анализ ще бъде дефинирана хипотеза за формалните граматични паралели между българските и английските темпорални глаголни форми. Ще бъдат представени и способите за формообразуване в двата езика с оглед на повърхнината реализация на морфологичните значения на различните категории.

II.3. Характеристика на глаголите в български

Както е известно, в българския език глаголът е най-развитата част на речта – „слон в царството на българската граматика“ (цит. на Ал. Теодоров-Балан по Ницолова 1976: 7). Единствено в български в сравнение с останалите славянски езици глаголът запазва повечето най-важни особености от старобългарския език, но и претърпява допълнително развитие (Ницолова 2008: 224). Подобно на схващанията в класическото езикознание в българските граматики характеризирането на глагола винаги се е свързвало с „действие или състояние“. Тази традиция води началото си още от зараждането на съвременното българско езикознание по време на българското просветно движение. Пръв Н. Рилски характеризира глагола като част на речта, която

„изявлява действие или страдание, или сродно нещо“ (Рилски 1835: 121). Като се изключат най-ранните определения, в българското езикознание действието или състоянието обикновено се свързват (по един или друг начин) с някоя от глаголните морфологични категории – време, залог, лице и т.н. „Глагол е дума, която означава действие, което се върши от лице (предмет), или пък състояние, в което се намира лицето или предметът“ (Младенов 1939: 273). „Глаголът е дума, която означава действие или състояние, свързано с глаголно лице“ (Стоянов 1964: 320), „Глаголът като част на речта означава действие (или състояние) като процес във времето. Действието винаги е отнесено към лице или предмет“ (Пашов 1999: 120). Л. Андрейчин подчертава синтактичната натовареност на глагола: „Глаголът изразява движение или състояние в предикативна връзка с даден субект“ (Андрейчин 1938). Вижда се, че основната връзка при тези дефиниции е с морфологичната категория лице, което влиза в противоречие с дефинирането на безличните глаголи. От казаното следва, че или тези глаголи не могат да бъдат причислени към глаголната парадигма, или че дефиницията трябва да бъде изменена.

Освен описателното дефиниране на глагола чрез неговите семантични характеристики, което може да е или твърде абстрактно, или твърде специфично, е възможно да се вземат предвид и функционалните характеристики на глагола и неговите отношения с други единици от различни езикови равнища.

Според морфологичните си характеристики глаголът попада в групата на изменяемите части на речта, а в глаголната парадигма се включват голям брой словоформи с множество граматични значения.

Според синтактичната си функция и характеристики глаголът попада в групата на самостойните части, тъй като той е тази част от изречението, която изразява отношението между другите съставни части на изречението – той свързва аргументите в рамките на пропозицията (Ницолова 2008: 224). Основното лексикално деление при глаголите, което се основава на техните функционални особености, е на пълнозначни и спомагателни глаголи (Куцаров 2007: 108). Пълнозначните глаголи имат свое лексикално значение и могат да образуват словоформи за цялата парадигма от граматични значения, докато спомагателните глаголи са загубили основното си лексикално значение и се използват при образуването на аналитичните глаголни форми (Куцаров 2007: 108). Друго основно лексикално-семантично деление на глаголите е според това дали изразяваното от тях действие може да засегне пряко друг обект. Според

това деление глаголите биват преходни, при които има възможност действието да засегне друг обект и чийто формален показател е възможността за съчетаване с пряко допълнение в синтактичен план, и непреходни, при които е невъзможно действието да засегне друг обект, а в синтактичен план формалният показател е невъзможността да имат преки допълнения (Куцаров 2007: 106 – 107). Според възможността глаголите да изразяват граматичното значение лице те се делят на лични и безлични (Куцаров 2007: 105). Безличните глаголи, които се ограничават до няколко семантични области, са представени само от формата за среден род, трето лице, единствено число на глагола. Посочвано е (Коева 2008: 59), че осен като лични и безлични глаголите могат да се разделят на третолични, множественолични и множественотретолични.

Характерна черта на българския глагол е възможността да изразява на равнището на лексикалната единица отношението на действието към неговия предел. Граматическата категория вид на глагола „дава възможност протичането на едно действие да се представи по два начина: като завършено действие и като неприключил процес. Следователно видът е морфологическа категория, която показва начина на протичане на глаголното действие по отношение на признака цялостност (изчерпаност, завършеност, комплексност)“ (Гаравалова 2003: 186). Тази отличителна характеристика на глаголите в българския език е обект на множество научни дискусии и различни концепции. Възприема се, че делението на глаголите в българския език на глаголи от свършен вид и на глаголи от несвършен вид е в рамките на лексикално-граматичната категория вид на глагола (Куцаров 2007: 551). Според делението в тази категория глаголите се разпределят в зависимост от това дали действието е насочено към някакъв предел, или не и съответно се означават като пределни и непределни. „Само пределните глаголи са представени в свършен и несвършен вид […], а непределните глаголи са само несвършени“ (Ницолова 2008: 248). Също така трябва да се има предвид, че „глаголните категории вид на глагола и време не са абсолютно независими една от друга и вследствие от това тяхното взаимодействие налага някои ограничения в употребите на свършен и несвършен вид както в различните времена, така и в частните значения на едно и също време, някои от които не ограничават функционирането на двата вида, други обаче изключват функционирането на свършен или несвършен вид (Гаравалова 2003: 187).

От гледна точка на основната морфологична класификация на частите на речта в български език глаголът попада в групата на изменяемите части на речта. В глаголната парадигма се включват голям брой словоформи с множеството значения, които са

граматикализирани. Посочвани са различни обхвати за глаголната парадигма – както по отношение на броя на граматичните форми, така и по отношение на категориите, които трябва да бъдат включени в нея (Андрейчин 1938: 17, Маслов 1956: 85). Броят на глаголните форми варира, в зависимост от това дали се включват (или не се включват) отрицателните форми, евиденциалните форми и т.н.

II.3.1. Морфологичните категории лице, число и род при глаголите в български

Глаголните граматични категории могат да бъдат разделени според обобщеното семантично значение, което изразяват в дадена глаголна форма. Граматичните значения, които характеризират предмета, на който се приписва динамичният признак, са лице, число и род (Ницолова 2008: 224). Значенията, които са свързани с динамичния процесуален признак, който стои в основата на лексикалното значение на глагола, са време и вид (Ницолова 2008: 224). Отношението на глаголния субект (глаголното лице) към действието се изразява от категорията залог (Куцаров 2007: 341, Ницолова 2008: 224). Дискусионни са категориите, които изразяват отношението на говорещия към действието, както и структурирането им в отделна група – наклонение, евиденциалност, преизказност и т.н.

Морфологичните категории, които притежават глаголите в български език и са обект на анализ в това изследване, са: лице, число, в някои случаи род, време, вид на действието и таксис. Акцентът е върху категориите, които изграждат темпоралните отношения в граматичната категория време, и техните формообразователни характеристики с оглед на преноса на граматикализираната информация при превод. Накратко ще бъдат представени категориите лице, число и род.

Морфологичната категория лице е една от най-безспорните глаголни категории. Семантичното съдържание на тази категория „се основава на отношението на глаголния субект към изказването“ (Куцаров 2007: 201). Лицето при глагола „характеризира предмета, на който се приписва динамичният процесуален признак, от гледна точка на отношението му към речевия акт“ (Ницолова 2008: 225). Ще бъде представено накратко значението на категорията лице. От прагматична гледна точка, морфологичната категория лице при глаголите има деиктична същност. Категорията представя персоналната референция на участниците в комуникативния акт чрез краен набор от езикови средства. Редно е да се отбележи посочваната от редица автори двояка същност

на тази категория или нейното разделение на глаголна (морфологична) и местоименна (лексикално-граматична), което е изградено на базата на функционално-семантичния подход на А. Бондарко (Бондарко 1976).

Както беше посочено, семантиката на морфологичната категория лице е свързана с изразяване на отношението на глаголния субект към изказването. Л. Андрейчин (Андрейчин и др. 1997: 202) изразява схващането, че значението на категорията означава „някои граматически разновидности на глаголното лице в зависимост от неговото отношение към говорещото лице“. В българското езикознание се е наложило схващането, че категорията лице е тричленна и е изградена върху опозицията комуникатор – некомуникатор (Герджиков 1984: 222). Признакът комуникатор обединява първо и второ лице, а трето лице притежава признака некомуникатор.

При синтетичните и аналитичните форми на личните глаголи за изразяване на граматичната категория време в български има ясно изразени показатели за различаване на 1., 2. и 3. лице. Ограничената парадигма на останалите видове глаголи по отношение на категорията лице (безлични, третолични и др.) не променя включването им в модела за изразяване на категорията.

Морфологичната категория число обхваща всички изменяеми части на речта в това число и глаголите. Категорията число в традиционното българско езикознание е била изследвана в контекста на имената (най-вече съществителните), това е и причината за нейното сравнително слабо проучване в контекста на глаголите и глаголните форми. Още първите изследователи говорят за категорията число като категория, която изразява „липса или наличие на информация за количество на предметите“ (Теодоров-Балан 1958: 46). В по-ново време Св. Иванчев отбелязва, че към значението на морфологичната категория число трябва да се причисли и „способността на […] глаголните лица да се броят и измерват“ (Иванчев 1978: 102). Семантичното съдържание на категорията число се основава на „информация за разделно множество или единственост на явленията“ (Куцаров 2007: 179). Значението число при глаголите представлява „характеристика на обективните количествени признаци на явленията, на които се приписва динамичният признак“ (Ницолова 2008: 225). Формалните маркери за категорията число при глаголите са свързани с формалните маркери за категорията лице. Определени морфеми, които могат да бъдат и нулеви, носят значенията на категориите лице и число. Този факт е бил причина за редица опити двете категории да бъдат слети в една обща категория по отношение на глаголните форми. При неличните глаголни

форми обаче проявата на тези категории е следната: причастията не притежават категорията лице, но притежават категорията число; при деепричастията и двете категории не са изразени.

При синтетичните и аналитичните форми на личните глаголи, изразяващи граматичната категория време, има ясни морфологични показатели за различаване на единствено и множествено число. Показателите са налице и при глаголите с ограничена парадигма по отношение на категорията число.

Категория род, подобно на категорията число, е една от категориите, която по традиция се представя в контекста на имената като части на речта. Характеристиките на категорията род са различни при различните части на речта – при едни тя е лексикално-граматична, при други – морфологична (Куцаров 2007: 184, 556). В контекста на глаголите голяма част от аналитичните глаголни форми за единствено число, които съдържат в състава си минало свършено деятелно причастие или минало страдателно причастие, са натоварени с допълнително граматично значение – род. Морфологичната категория род при глаголите дава информация за граматичния род на глаголното лице.

Категорията род по отношение на темпоралните форми се изразява от аналитичните глаголни форми, образувани с причастия (минало неопределено време, минало предварително време, бъдеще предварително време и бъдеще предварително време в миналото).

Представените три морфологични категории – лице, число и род при глаголите – попадат към групата на глаголните морфологични категории, които характеризират явлението, на което се приписва динамичният признак, изразяван от глагола, и тяхната морфологична реализация позволява импликация на субекта в изречението.

II.3.2. Особености на българската темпорална система

Морфологичните категории, които ще бъдат разгледани по-подробно, са време, вид на действието и таксис. Тук е мястото да се спомене, че в българската езиковедска традиция има доста разногласия относно статута и инвентара на глаголните морфологични категории. В първите сериозни езиковедски проучвания липсват строги дефиниции за морфологични категории и категориални отношения, характеризирането на категориите е по-скоро описателно.

Теоретичното описание на граматичната категория време в български език е засвидетелствано още в първите български граматики, като за наличието на такава категория спорове в езиковедската литература няма. Споровете са относно обхвата и количеството на грамемите и значенията им. В първите български граматики броят на глаголните времена е силно колеблив – в граматиката на Н. Рилски броят на глаголните времена е 12 (Рилски 1835 :126), докато в граматиката на Н. Бозвели от същото време броят им е 5, като тук за пръв път се споменава, че „глаголните времена свойствено са три – настоящее, прешедшее и будеще“ (Бозвели 1835: 35). Първите опити за дефинирането на категорията време и представянето на пълния инвентар от глаголни времена в българския език са продиктувани от чужди езиковедски традиции. Именно под тяхно влияние Ив. Момчилов, имайки предвид църковнославянския език, дава за пръв път формална семантична характеристика на категорията време, също така описателно характеризира и значенията за резултативност и относителност (Момчилов 1847: 52). По това време започва теоретичното осмисляне на таксиса като категория – за пръв път се употребява терминът „относително преминало време“ от Й. Груев (Груев 1858: 42), а първото описание на тази категория дава Г. Миркович (Миркович 1860: 23). Започва и теоретичното представяне на категорията вид на действието – описателна дефиниция дава в осемчлената си темпорална система С. Радулов (Радулов 1863: 122), който характеризира „определено преминало време“ като действие с резултат. Ю. Трифонов приема употребата на названието минало неопределено време, но само при условие, че в системата съществуват и други определени времена (Трифонов 1905: 163). При всички изброени автори броят на глаголните времена се различава, но също така липсва и единна концепция за отношенията в системата – глаголните времена по-скоро се изброяват, вместо да се представят в категориални отношения помежду си. Поставено е обаче началото на разбирането за темпорална система, изградена от множество граматични категории с определени зависимости помежду си, чиито значения могат да се комбинират.

„Зараждането“ на описанието на съвременната българска темпорална система може да се отнесе към докторската дисертация на Л. Андрейчин (Андрейчин 1938) и по-късната му „Основна българска граматика“ (Андрейчин 1944). В тези свои трудове Л. Андрейчин ясно описва значението резултативност при перфектните времена и значението относителност, както и отношенията между тях и морфологичната категория време, като разделя времената на абсолютни и относителни и на времена с просто и

сложно хронологическо отношение (Андрейчин 1944: 208). В по-късни изследвания характеристиките на така описаната темпорална система са подложени на повторно дефиниране и осмисляне.

Теоретичното описание на категорията вид на действието не е еднотипно. Докато всички изследователи след средата на XX век говорят за времена с абсолютна или двойна ориентация, времена с проста или сложна ориентация и т.н., осъзнаването на значението резултативност не е прокарано последователно. П. Пашов категорично разграничава резултативни и нерезултативни глаголни форми (Пашов 1965: 55). Той приема деветчлената темпорална система на Л. Андрейчин и засяга въпроса за систематизирането на отношенията между членовете ѝ: „Българските глаголни времена се подреждат по отношение на това дали показват отношение на глаголно действие към даден ориентационен момент (акционност), или показват резултат от действие към ориентационен момент (резултативност) в стройна система“ (Пашов 1965: 58). По този начин П. Пашов представя единна система, изградена от няколко категории, чиито значения могат да имат кумулативен характер и строго определени отношения както вътре в системата, така и с други морфологични категории. В по-ново време Г. Герджиков изгражда концепцията за хиперкатегорията време, която се състои от три синкатегории – базисно, надстроечно и нешифтърно време (Герджиков 2000), и „при която между крайните елементи на отношението, изразявано от съответната проста категория, се вмъкват един или повече елементи посредници“ (Герджиков 2000). Преимуществото на теорията за хиперкатегорията време на Г. Герджиков е, че се дефинира ясна йерархия между отношенията на отделните членове, които я изграждат. В последните години концепцията за йерархичност и отношения между категориите е възприета и доразвита от Ив. Куцаров (Куцаров 2007), който приема с някои забележки системата на Г. Герджиков и използвайки принципа на И. Смирницки (Смирницки 1955: 21), разлага отделните опозиции от системата на Г. Герджиков в отделни морфологични категории със собствени значения и инвентар от грамеми.

За целите на дисертацията и изследването на преноса на граматикализираната информация при превод от български на английски ще бъдат представени отношенията, чрез които е изградена българската темпорална система, като се използва концепцията за системност и кумулативност на морфологичните категории, които я изграждат, с някои терминологични коментари.

За представянето на системните отношения и формалното конструиране на значенията на членовете на различните морфологични категории ще бъде използван семантичният модел на Х. Райхенбах (Райхенбах 1947). Представянето на темпоралните отношения в български чрез модела на Х. Райхенбах не е нов подход – редица автори, сред които Св. Иванчев (Иванчев 1971: 68), М. Янакиев (Янакиев 1976: 230) и Р. Ницолова (Ницолова 2008: 265), са представяли темпоралните отношенията чрез този формално-семантичен модел. Формално-семантичното описание на темпоралните значения в двата езика ще бъде използвано при изграждането на хипотеза за българско-английските граматични паралели на равнище темпорални форми. Ще бъдат представени накратко трите морфологични категории, които влизат в състава на граматичната категория време. Формалната репрезентация на значението на всяко граматично време в български ще бъде представена поотделно.

Авторите на различните концепции от по-ново време са единодушни относно същинското значение на морфологичната категория време и разграничаването ѝ от други значения, носени от българските глаголни форми. Същинското значение на морфологичната категория време може да бъде конструирано само по отношение на определен ориентационен момент (момента на речевия акт или друг момент), който е формалният „темпорален център“, спрямо който отношенията могат да бъдат в привативната опозиция разновременност – неразновременност (Куцаров 2007: 244). В рамките на разновременността допълнително се изгражда опозицията следходност и предходност (Куцаров 2007: 244). Или иначе казано, морфологичната категория време в българския език изразява отношението на динамичния признак към момента на изказването – съвпадение (едновременност) или несъвпадение (разновременност). Трябва да се отбележи „по повод на едновременността на действието с ориентационния момент (независимо дали става дума за момента на говорене или за миналия ориентационен момент), че тук не става дума за пълно съвпадение, а за съвпадение само на една от фазите на действието с акта на комуникацията“ (Гаравалова 2003: 191). Морфологичната категория време в български съдържа следните грамеми (Куцаров 2007: 244):

- грамемата **сегашно време** изразява едновременност на динамичния признак с изказването или липса на информация за разновременност;
- грамемата **минало време** изразява предходност на динамичния признак по отношение на изказването;

- грамемата **бъдеще време** изразява следходност на динамичния признак по отношение на изказването.

Според представената от Ив. Куцаров система (Куцаров 2007: 244) в български език информация за едновременност (или липса на информация за разновременност) на действието и изказването носят формите на сегашно време, минало несвършено време (заедно с информация за относителност), минало неопределено време (заедно с информация за резултативност) и минало предварително време (заедно с информация за относителност и резултативност). Информация за следходност на действието спрямо изказването носят формите на бъдеще време, бъдеще предварително време (заедно с информация за резултативност), бъдеще време в миналото (заедно с информация за относителност) и бъдеще предварително време в миналото (заедно с информация за относителност и резултативност). Информация за предходност на действието спрямо изказването носят формите на минало свършено време.

Втората категория, която влиза в състава на хиперкатегорията време в български език, е категорията **вид на действието** (резултативност) (Куцаров 2007: 249). Това е една от глаголните морфологични категории, чието значение е представено колебливо през историческото развитие на теоретичното описание на българската глаголна система. Съществуват различни виждания за вида на опозицията: еквиполентна (резултативност – акционност) (Герджиков 1984: 115) или привативна (резултативност – нерезултативност) (Куцаров 2007: 250). Семантиката на морфологичната категория вид на действието се изгражда върху „отношението между действие и резултат от действие към един и същ ориентационен момент“ (Куцаров 2007: 249).

Вижданията по отношение на ориентационните моменти също не са единни. Някои автори (Андрейчин 1944, Пашов 1999) причисляват ориентационните моменти към морфологичната категория време, а други (Куцаров 2007) – към категорията вид на действието. Приема се схващането, че резултативните моменти (Куцаров 2007: 250) могат да бъдат четири – сегашен, минал, бъдещ и бъдещ спрямо минал. Информация за вида на действието в български език могат да изразяват следните грамеми:

- грамемата **резултативност** изразява резултат от действие, актуализиране на действието в най-близкия следващ ориентационен момент (Куцаров 2007: 250);

- грамемата **нерезултативност** изразява липса на данни за резултат от действие, а в главното си значение сигнализира за действие (акционност) (Куцаров 2007: 250).

Според представената от Ив. Куцаров система (Куцаров 2007: 250 – 253) в български резултативност се изразява от формите на минало неопределено време, бъдеще предварително време, минало предварително време (заедно с информация за относителност към минал ориентационен момент) и бъдеще предварително време в миналото (заедно с информация за следходност на резултата и относителност към минал ориентационен момент). Нерезултативност се изразява от сегашно време, минало свършено време, минало несвършено време, бъдеще време и бъдеще време в миналото.

Третата категория, която влиза в състава на граматичната категория (хиперкатегорията) време в българския език, е категорията **таксис** (относителност) и изразява „отношение на действията или резултатите от действия, ориентирани към миналия ориентационен момент (минало време или друг сигнализатор), към действията, неориентирани към миналия ориентационен момент“ (Куцаров 2007: 255). Подробно описание на различните мнения за семантичното съдържание на категорията е представено от Куцаров (Куцаров 2007: 221 – 226). Макар семантичното съдържание на категорията да е теоретично обосновано като отделно от морфологичната категория време и морфологичната категория вид на действието, съществуват доста разногласия по отношение на опозициите, които изграждат категорията, нейното семантично съдържание, както и за името на категорията. Л. Андрейчин пръв прокарва разделението между „времена с просто хронологическо отношение и със сложно хронологическо отношение“ (Андрейчин 1944: 206). Това разделение е прието в българското езикознание и относителността се разбира като отнасяне на действие към друг ориентационен момент, различен от момента на говорене. П. Пашов по-късно използва термина на Р. Якобсон (Якобсон 1972: 101) таксис, за да именува тази категория, и дефинира опозицията, изграждаща категорията, като еквиполентната опозиция акционност – абсолютност (Пашов 1965: 58). По отношение на ориентационния момент Г. Герджиков противопоставя времената, ориентирани към момента на говорене, срещу времената, ориентирани към минал ориентационен момент (Герджиков 2000). В по-ново време Ив. Куцаров посочва, че относителните форми са ориентирани към минало свършено време или друг негов сигнализатор, докато неотносителните или не са ориентирани, или са ориентирани към момента на изказването, базирайки се на мнението на Г. Герджиков,

че „минало свършено време е единственото минало време в българския език“ (Куцаров 2007: 229). Докато морфологичната категория време се изгражда върху отношението на действие към момента на говорене, категорията таксис се основава на отношението между действие към друго действие (ориентационен момент). Или с други думи, категорията таксис се основава на привативната опозиция относителност – неотносителност (зависим – независим таксис) (Куцаров 2007: 255). На морфологично ниво информация за съотнасяне към определен минал момент или липса на такава могат да носят следните грамеми:

- грамемата **относителност** изразява действия (или резултати от действия), ориентирани към единственото минало време в българския език (минало свършено време) или към друг сигнализатор на миналия ориентационен момент (Куцаров 2007: 255);
- грамемата **неотносителност** изразява действия (или резултати от действия), при които липсва информация за отнасяне към някакъв ориентационен момент или (при изразяване на главното значение) са ориентирани към момента на говоренето (абсолютност) (Куцаров 2007: 255).

Според представената от Ив. Куцаров система (Куцаров 2007) в български език относителност се изразява от формите на минало несвършено време, бъдеще време в миналото, минало предварително време (заедно с информация за резултативност) и бъдеще предварително време в миналото (заедно с информация за следходна резултативност спрямо минало изказване – минал ориентационен момент). Липса на данни за относителност носят формите на сегашно време, минало свършено време, бъдеще време, бъдеще време в миналото и минало неопределено време.

Българската темпорална система се характеризира със стройна йерархичност – тя започва с максимално немаркираното сегашно време, продължава с единично маркираните минало свършено време (по признака предходност), бъдеще време (по признака следходност), минало несвършено време (по признака относителност) и минало предварително време (по признака резултативност), следват двойно маркираните бъдеще предварително време (по признаците следходност и резултативност), минало предварително време (по признаците относителност и резултативност) и бъдеще време в миналото (по признаците следходност и относителност) и завършва с тройно маркираното бъдеще предварително време в миналото (маркирано по признаците следходност, резултативност и относителност (Гаравалова 2003: 195).

Представена по този начин, граматичната категория време в българския език има структурен вид, който поставя различните морфологични категории и техните значения в зависима йерархия. Необходимо да се отбележи, че представената система е една от концепциите, които могат да получат вярна системно-структурна интерпретация само ако се прокара последователно разграничение между опозициите, изграждащи различните категории.

В следващата част ще бъдат представени поотделно българските глаголни времена, заедно с формално-семантично описание, базирано на представената по-горе система на Х. Райхенбах.

II.3.3. Семантична характеристика на българската темпорална система

Както вече беше посочено, един от успешните опити за формално-семантичен анализ на особеностите на темпоралните системи, е моделът на Х. Райхенбах (Райхенбах 1947). За разлика от моделите, получени въз основна на класическите описания на морфологичните времена на един език, моделът на Х. Райхенбах е по-независим, тъй като функционира като автономна логическа конструкция, подходяща за представянето на граматичните времена. Моделът на Х. Райхенбах може да бъде прилаган при изследването на типологично различни темпорални системи, както и да бъде средство за сравнително формално описание.

За български език подобен подход е системно приложен през 70-те години на XX век като опит за абстрахиране от формалното морфологично описание на езиковата система и прилагане на инвариантно семантично тълкуване. Св. Иванчев (Иванчев 1971: 68) е един от първите учени, който изгражда модел на българската темпорална система, подобен на модела на Х. Райхенбах. Неговата система използва три ориентационни момента – момент на говоренето; момент, за който се говори, и момент, в който се върши действието (Иванчев 1971: 75). Р. Ницолова също представя концепция за българската темпорална система, изградена върху идеите на Х. Райхенбах (Ницолова 2008: 265). Преимуществото на нейната концепция е, че не използва момента на говорене като основен ориентационен момент за описание на категорията време, а предлага структура, изградена върху три ориентационни момента – момент на референцията, момент на говоренето и момент на действието. Моментът на референцията е т. нар. „момент, за който се говори“ от традиционното българско

езикознание, като се отчита и тясната връзка между него и момента на действието (Ницолова 2008: 262).

За да има основание да се твърди, че определено време в български език отговаря по формално-семантичните признаци на определено време в английски език, е необходимо да се изгради формално-семантичен модел на българската и английската темпорална система. Ще бъде използвана системата, предложена от Р. Ницолова, основана на темпоралния модел на Х. Райхенбах (Ницолова 2008: 264). При тази система събитието, означено от глагола, се локализира върху темпоралната ос като взаимоотношение между няколко момента. Ще бъдат използвани следните означения, представени от Р. Ницолова (Ницолова 2008: 264 – 265):

T_0 – моментът на говорене (още акт на комуникацията);

R – интервалът на референтност, като е възможно да има повече от един интервал на референтност (R_1) (ще се изобразява с ( ));

E – интервалът на действието (ще се изобразява с < >);

t_{res} – моментът, в който е наличен резултат от действието.

Релациите между изброените моменти ще се означават по следния начин:

$X < Y$ – X предхожда Y, X е преди Y;

$X \in Y$ – X е елемент от Y;

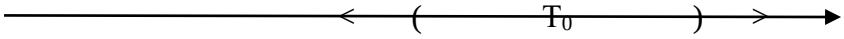
$X \notin Y$ – X не е елемент от Y;

$X \supset Y$ – X включва Y;

$X \supseteq Y$ – X включва или съвпада (едновременно е) с Y.

Използвайки тази система, може да бъде изграден следният формален модел за семантиката на времената в български език:

II.3.3.1. Сегашно време



(R); < E >

$T_0 \in R$; $E \supseteq R$[2]

„При сегашно време в български език интервалът на действието включва в себе си интервала на референтност, който пък включва в себе си момента на говоренето.

[2] За български представените формули са по Р. Ницолова (2008: 269).

Сегашно време в български език изразява действие, което е континуативно“ (Ницолова 2008: 272). Подобно на английски език, а и на други езици, сегашно време в български изразява действие, чийто интервал може да бъде неограничен по отношение на интервала на референтност, като диференциалният признак при сегашно време е включването на момента на говоренето в интервала на референтност. От това произлизат и различните употреби на сегашно време, при които интервалът на действие реферира към събития, предходни и/или следходни, спрямо момента на говорене, но включени в интервала на референтност. Така нареченото сегашно историческо време например има стилистична употреба, която се „характеризира с преместване както на интервала на референтност, така и на интервала на действие в плана на миналото“ (Ницолова 2008: 275).

(1) Пример за сегашно време[3]:

Македонският президент ***обявява*** *нова охранителна мисия* (*в този момент*).

Сегашно време може да изразява и действия, които са актуални по отношение на цялото протежение на темпоралната ос – например при сегашно гномическо време интервалът на референтност обхваща цялата темпорална ос. „То означава действия, които са валидни във всеки интервал от темпоралната ос“ (Ницолова 2008: 277). „Сегашно повторително време изразява действия, при които моментът на говорене не е елемент от интервала на референтност, а поредицата интервали от повторителни действия включва в себе си поредица от интервали на референтност“ (Ницолова 2008: 265). В сегашно гномическо време се „употребяват само глаголи от несвършен вид, при които действието е разчленимо на фази, а глаголите от свършен вид, при които действието е неразчленимо, не се употребяват в сегашно гномическо време, но могат да се употребяват в сегашно повторително време, защото в този случай едно от действията от веригата, които поотделно са неразчленими на фази, съвпада с момента на говорене“ (Гаравалова 2003: 191). Формално представени, тези разновидности на сегашно време изглеждат така:

- сегашно историческо време – $R < T_0$; $E \supseteq R$;
- сегашно гномическо време – $T_0 \in E$; $E \supseteq (R_1, R_2, R_3 \ldots R_n)$;
- сегашно повторително време – $T_0 \notin R$; $(E_1, E_2, E_3 \ldots E_n) \supseteq R(R_1, R_2, R_3 \ldots R_n)$.

[3] Примерите за глаголните форми в тази глава са от Българско-английския паралелен корпус със съотнесени глаголни форми.

Р. Ницолова предлага интересно наблюдение по отношение на локализацията на действията при сегашно време. „Интервалът на действието се движи заедно с интервала на референтност, който променя мястото си спрямо момента на говорене T_0 по темпоралната ос, но във всички случаи темпоралното отношение между интервала на действие E и интервала на референтност R при едно и също време остава постоянно“ (Ницолова 2008: 265-266).

II.3.3.2. Минало несвършено време

(R); < E >

$R < T_0$; $E \supset R$

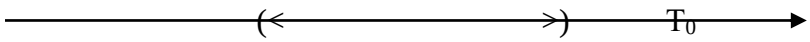
Минало несвършено време в български означава действие, чийто интервал включва в себе си миналия интервал на референтност (Ницолова 2008: 282). Интервалът на действието е „отворен“ по отношение на интервал на референтност, т.е. действието се означава като континуативно (Ницолова 2008: 282). В рамките на противопоставянето по признака континуативност/неконтинуативност минало несвършено време се противопоставя на минало свършено време (Ницолова 2008: 282). Минало несвършено време е относително време. Някои учени го наричат сегашно относително време (Куцаров 2007: 256).

(2) Пример за минало несвършено време:

*Иван беше смачкал гайдата си и я **прибираше** във вулията си.*

Подобно на сегашното време миналото несвършено време може да има различни стилистични употреби, произтичащи от означеното действие, включващо минал интервал на референтност. Основните транспозиционни употреби на минало несвършено време са актуално в миналото минало несвършено време; хабитуално минало несвършено време; повторително минало несвършено време (Ницолова 2008: 283). При повторителното минало несвършено време веригата от действия […] е незавършена, отворена в плана на миналия интервал на референтност (Ницолова 2008: 283).

II.3.3.3. Минало свършено време



(R); < E >

$R < T_0$; $R \supseteq E$

Семантиката на минало свършено време изразява действие, чийто интервал е едновременен с минал интервал на референтност и е затворен в рамките на този интервал (Ницолова 2008: 290). Затвореното действие в рамките на интервала на референтност се означава като неконтинутивно.

Въпреки факта, че има възгледи (Маслов 1982: 254 – 255), които причисляват миналото свършено и миналото несвършено време към една обща морфологична категория на базата на сходства във формообразуването и семантиката, както беше споменато по-горе, миналото свършено време се приема от редица учени за единственото същинско минало време в българския език (Герджиков 1973: 142, Куцаров 2007: 229).

Миналото свършено време в някои видове сложни съставни изречения се явява и време „опора“, към което могат да се ориентират другите действия в изречението и което сигнализира миналия интервал на референтност.

(3) Пример за минало свършено време*:*

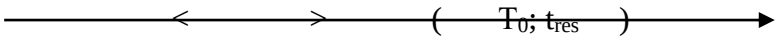
*Някой **почука** на вратата.*

Подобно на другите времена миналото свършено време може да има различни стилистични употреби. Например цялостността на действието, означена от глаголи от свършен вид, се съчетава много добре с минал интервал на референтност (Ницолова 2008: 290), а формите на глаголите от несвършен вид могат да означат и определена по-голяма продължителност на действието (Ницолова 2008: 290).

Също така формите на глаголите от свършен вид в минало свършено време представят основните действия в разказа, които „придвижват“ напред повествованието. Минало свършено време е основното разказвателно време, а минало несвършено време, минало неопределено време и минало предварително време представят допълнителни действия, свързани с основните, предадени с минало свършено време (Ницолова 2008: 290). Този факт прави формите за минало свършено време основни разказвателни форми в текста, „придвижващи“ наратива темпорално, а формите на другите времена, ориентирани спрямо миналото свършено време, изразяват

допълнителни съпътстващи действия. Миналото свършено време в български език може да има транспозиционна употреба, която изразява бъдещи действия – с тази употреба се изразява толкова голяма увереност на говорещия, че действието ще се осъществи, че то може да се представи в момента на говорене като вече реализирано (Ницолова 2008: 291). Тази употреба изразява близки или предстоящи действия, които са представени като логически и следствено неизбежни (и силно емоционално маркирани).

II.3.3.4. Минало неопределено време



(R); < E >

$T_0 \in R$; $E < R$, T_0; $t_{res} \in R$

При минало неопределено време се изразява значение, което представя действие, чийто интервал е в плана на миналото, предходен е спрямо интервала на референтност, който включва в себе си момента на говорене (Ницолова 2008: 294). Значението на глаголните форми в минало неопределено време представя резултативно действие, като резултатът (в материален или абстрактен вариант) е наличен в момента на говоренето (Андрейчин 1976: 279). Поради факта, че „при минало неопределено време интервалът на референтност съвпада с *интервала на речта*[4], когато е налице и резултатът от действието, самото действие не се свързва с определен минал интервал за разлика от минало свършено или минало несвършено време. Това позволява за действието да се мисли като за факт, като за нещо, което някога е станало, без да се конкретизира кога“ (Ницолова 2008: 295). Формите на минало неопределено време в общото си значение изразяват липса на данни за разновременност или едновременност на резултат от действие с изказването, а в главното си значение – едновременност на резултат от действие с изказването (Куцаров 2007: 252).

(4) Пример за минало неопределено време:

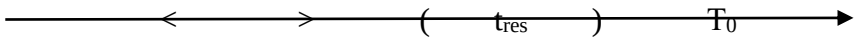
*Кумът **е пристигнал** снощи!*

Транспозиционните употреби на минало неопределено време акцентират върху резултативността на семантиката му. Например при повторителното минало неопределено време освен еднократно действие в много случаи глаголните форми от

[4] Терминът в курсив е използван от Р. Ницолова (Ницолова 2008: 295) и се различава от възприетата терминология в този текст.

несвършен вид, а понякога и формите от свършен вид, могат да означават повтарящо се действие, за което сигнализират обстоятелствени пояснения за време или квантификатори в именните групи, означаващи субекта или обекта (Ницолова 2008: 297). „В някои случаи при глаголи от свършен вид с определени начини на действие в минало неопределено време се обръща по-голямо внимание на резултата от действието, отколкото на самото действие. Тогава се говори за статално минало неопределено време“ (Ницолова 2008: 296). Формите за минало неопределено време могат да имат употреба като форми за предположение. „Поради наличието в момента на говорене на резултат […], говорещият вследствие на абдукция предполага, че действието е било извършено“ (Ницолова 2008: 297). Другите транспозиционни употреби на минало неопределено време са: акционно минало неопределено време, минало неопределено време за означаване на бъдещо предварително действие и минало неопределено време за означаване на гледната точка на участник – наблюдател в събитието, в плана на миналото и на бъдещето (Ницолова 2008: 295 – 298).

II.3.3.5. Минало предварително време



(R); < E >

$R < T_0$; $E < R < T_0$; $t_{res} \in R$

Семантиката на минало предварително време в български език „означава действие, което се извършва в плана на миналото преди даден минал интервал на референтността“ (Ницолова 2008: 302). Минало предварително време е резултативно време – резултатът му е наличен в миналия интервал на референтност (Ницолова 2008: 302). „В общото си темпорално значение формите на минало неопределено време изразяват липса на данни за разновременност или едновременност на резултат от действие с изказване в минало време (минал ориентационен момент), а в главното си значение – едновременност на резултат на действие с изказване в минало време (минал ориентационен момент)“ (Куцаров 2007: 260).

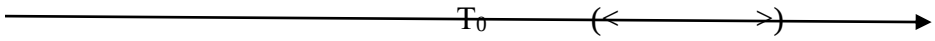
(5) Пример за минало предварително време:

Покрай реката имаше пейки, край които ***бяха насядали*** *младите двойки.*

Тъй като значението на минало неопределено и минало предварително време се различава основно по интервалите, в които е наличен резултатът от действието, минало

неопределено време и минало предварително време имат сходни стилистични употреби. Както минало неопределено време, така и минало предварително време при глаголи от свършен вид, които „означават достигане на някакъв видим резултат“ (Андрейчин 1944: 246 – цит. по Ницолова 2008: 302) може да има статално значение, т.е. подчертава се резултатът от действието, наличен в миналия момент на референтността (Ницолова 2008: 302). Глаголите от несвършен вид в минало предварително време се употребяват с акционно значение, при което се подчертава процесността на действието (Ницолова 2008: 303) – акционно минало предварително време. Също така акционното минало предварително време на глаголите от несвършен вид може да означава и повторителни действия, като в изречението и в контекста има лексикални маркери за повторителност (Ницолова 2008: 303). Формите за минало предварително време могат да изграждат второстепенна, допълнителна линия на повествованието. Тя е съставена от поредица последователни действия, станали преди основната линия на повествованието, изградено с минало свършено време (Ницолова 2008: 303).

II.3.3.6. Бъдеще време



(R); < E >

$T_0 < R; R \supseteq E$

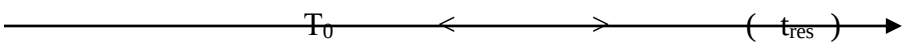
Значението на бъдеще време в български език „означава действие, което се извършва в бъдещ интервал на референтност, разположен след момента на говорене“ (Ницолова 2008: 306). Бъдеще време изразява „следходност на действие по отношение на изказването“ (Куцаров 2007: 244). Бъдеще време е третият член наред със сегашно време и минало свършено време на същинската категория време в български език. Както беше посочено, при морфологичната категорията време отношението между интервала на референтност, в който се проявява интервалът на действието, и момента на говорене е пряко. При бъдеще време началото на интервала на референтност може да бъде непосредствено след момента на говорене, но не и да съвпада с него. Поради семантичната структура на бъдеще време в български език, представяща нереални към момента на говорене действия, редица автори (Янакиев 1976: 230) смятат, че неговите форми нямат място при парадигмата на индикативните глаголни форми.

(6) Пример за бъдеще време:

*Късно е, утре вечер **ще замина**!*

Подобно на останалите времена от същинската категория време, бъдеще време може да има стилистични употреби, с които да означи възможни действия в миналото и сегашни действия. Формите за бъдеще време могат да изразяват повторителни действия. „Интервалът на референтност при бъдеще време може да се мултиплицира по темпоралната ос и заедно с него се мултиплицират и действията“ (Ницолова 2008: 308). Другите транспозиционни употреби на бъдеще време са: таксисна употреба в подчинени изречения, употреба в подчинени изречения за време и условие, бъдеще време за подкана, бъдеще време за необходимо действие, бъдеще време за възможно действие в ехо-въпроси, бъдеще време за възможно действие вследствие на сложно условно изречение и бъдеще време за означаване на реални минали и сегашни действия (Ницолова 2008: 308 – 310).

II.3.3.7. Бъдеще предварително време



(R); < E >

$T_0 < E < R$; $t_{res} \in R$

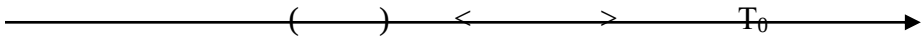
Бъдеще предварително време означава действие, което се е извършило преди бъдещ интервал на референтност, в който ще бъде налице резултатът от действието (Ницолова 2008: 315) Иначе казано, подобно на минало предварително време, при което резултатът от действието е наличен спрямо минал момент, при бъдеще предварително време резултатът от действието е следходен спрямо момента на говорене. Формите на бъдеще предварително време изразяват „следходност на резултат от действие по отношение на изказването“ (Куцаров 2007: 253).

(7) Пример за бъдеще предварително време:

След два дена ***ще сме се пренесли***.

Формите на това време в български език имат типична статална употреба, при която на преден план излиза резултатът от действието, а не неговото протичане (Ницолова 2008: 315). Също така формите на бъдеще предварително време се използват „в следствието на сложните изречения за условие и в главните изречения на други сложни изречения (например за време с *когато, като,* като в подчиненото изречение се използва сегашно време от свършен вид), за да означат възможно действие“ (Ницолова 2008: 315).

II.3.3.8. Бъдеще време в миналото



(R); < E >

$R < E < T_0$

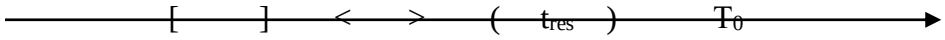
Бъдеще време в миналото означава действие, което е предстоящо по отношение на минал интервал на референтност (Ницолова 2008: 311 – 312). Това е едно от относителните времена, т.е. формите му изразяват „следходност на действие по отношение на изказване в минало време (минал ориентационен момент)“ (Куцаров 2007: 259). Този факт дава основание, подобно на формите за бъдеще време, формите на глаголите за бъдеще време в миналото да бъдат причислявани към наклоненията (Пантелеева 1997: 80).

(8) Пример за бъдеще време в миналото:

***Щях да му** я **дам** вчера.*

Основната употреба на бъдеще време в миналото е „за означаване на възможни действия, при които значението показва, че възможният свят, с който се свързва глаголното действие, е достижим до реалния свят“ (Ницолова 2008: 312). Употреба на бъдеще време в миналото е възможна и за „представяне на контрафактивно действие, а също и за означаване на припомняне“ (Ницолова 312 – 313).

II.3.3.9. Бъдеще предварително време в миналото



[R]; (R_1); < E >

$R < E < R_1 < T_0$; $t_{res} \in R_1$

При формирането на значението на бъдеще предварително време в миналото се използват два интервала на референтност. Единият от тези интервали R е в плана на миналото и действието е предстоящо по отношение на него. Тъй като бъдеще предварително време в миналото е резултативно време, резултатът от действието е наличен във втори интервал на референтност R_1 (Ницолова 2008: 316). Бъдеще предварително време в миналото е относително и резултативно време, чиито форми „изразяват следходност на резултат от действие по отношение на изказване в минало време (минал ориентационен момент)“ (Куцаров 2007: 261).

(9) Пример за бъдеще предварително време в миналото:

*До утре те **щяха да са изчезнали.***

Подобно на бъдеще време в миналото бъдеще предварително време в миналото може да означава възможно, контрафактивно и дори реално действие (Ницолова 2008: 316). Възможна е и употребата на бъдеще предварително време в миналото, при която вторият интервал R_1 е след момента на говорене, т.е. $R < E < T_0 < R_1$ (Ницолова 2008: 316). Поради сложността на своето значение бъдеще предварително време в миналото се среща твърде рядко (Ницолова 2008: 316).

II.3.4. Образуване на темпоралните форми в български език

Образуването на темпоралните форми на българските глаголи е доста комплексно и обвързано с различни лингвистични фактори – произход на думата, лексикално-граматичната категория, към която принадлежи глаголът, фонетични особености на думата и т.н. За целите на труда ще бъдат представени основните типове формообразуване при темпоралната система в български език.

Глаголите в български език притежават сегашна основа, която се образува с тематичните гласни, които определят и спрежението им: за I спрежение – е, за II второ спрежение – и, за III спрежение – а (Куцаров 2007: 110 – 111). Глаголите от I и II спрежение притежават и имперфектна и аористна основа (ГСБКЕ II: 304 – 314). Според формообразуването словоформите на българския глагол се делят на синтетични и аналитични. Накратко представено, формообразуването на българските темпорални форми изглежда така:

Синтетични форми:

– Сегашно време се образува от сегашната основа на глаголите и окончанията за лице и число. Подробно са представени от Р. Ницолова (Ницолова 2008: 271).

– Минало свършено време се образува от аористната основа (при глаголите от I и II спрежение) или сегашната основа (при глаголите от III спрежение) и окончанията за лице и число. Подробно са представени от Р. Ницолова (Ницолова 2008: 286).

– Минало несвършено време се образува от имперфектната основа (при глаголите от I и II спрежение) или от сегашната основа (при глаголите от III спрежение) и окончанията за лице и число. Подробно са представени от Р. Ницолова (Ницолова 2008: 281).

Формалните показатели на категорията лице зависят от спрежението на глагола и биват изразявани чрез набор от морфеми. При сегашно време и времената, при чието формообразуване участват формите за сегашно време, те са: за първо лице – първо и второ спрежение, единствено число – краесловна морфема А [ъ]; за първо лице – трето спрежение, единствено число – краесловна морфема М; за второ лице, единствено число – краесловна морфема Ш; трето лице, единствено число формално остава немаркирано; при множествено число морфемите сигнализират лице и число – при първо лице, множествено число – краесловна морфема М; при второ лице, множествено число – краесловна морфема ТЕ, при трето лице множествено число – краесловна морфема АТ [ът], а за глаголите от трето спрежение – Т. При минало свършено време формалните показатели за лице и число са: за първо лице, единствено число – краесловна морфема Х; за второ и трето лице, единствено число – нулева морфема (при минало свършено време има формално съвпадение при формите за второ и трето лице, единствено число); за първо лице, множествено число – краесловна морфема ХМЕ; за второ лице, множествено число – краесловна морфема ХТЕ; за трето лице, множествено число – краесловна морфема ХА. Показателите за лице при формите за минало несвършено време са: за първо лице, единствено число – краесловна морфема Х; за второ и трето лице, единствено число – краесловна морфема ШЕ (при минало несвършено време има формално съвпадение при формите за второ и трето лице, единствено число); за първо лице, множествено число – краесловна морфема ХМЕ; за второ лице, множествено число – краесловна морфема – ХТЕ; за трето лице, множествено число – краесловна морфема – ХА. При аналитичните глаголни форми, образувани с причастия, информацията за лице е кодирана от спомагателните глаголи.

Аналитични форми:

– Бъдеще време се образува от два глагола: „вкаменена“ форма на глагола *ща* (ще) и спрегнатите форми за сегашно време на глагола (Ницолова 2008: 304).

– Минало неопределено време се образува от формите за сегашно време на спомагателния глагол съм и миналото свършено деятелно причастие на глагола (Ницолова 2008: 292).

– Минало предварително време се образува от формите за минало несвършено време на спомагателния глагол съм и миналото свършено деятелно причастие на глагола (Ницолова 2008: 301).

– Бъдеще време в миналото се образува от формите за минало несвършено време на спомагателния глагол *ща*, частицата *да* и формите за сегашно време на глагола (Ницолова 2008: 311).

– Бъдеще предварително време се образува от формите за бъдеще време на спомагателни глагол *съм* (или *бъда*) и миналото свършено деятелно причастие на глагола (Ницолова 2008: 313).

– Бъдеще предварително време в миналото се образува от формите за бъдеще време в миналото на спомагателния глагол *съм* (или *бъда*) и миналото свършено деятелно причастие на глагола (Ницолова 2008: 315).

Вижда се, че при аналитичните форми на времената в български език формообразувателните елементи са 5: формите за сегашно време на спрегаемия глагол; формите на спомагателния глагол *съм* в сегашно време, минало несвършено време, бъдеще време и бъдеще време в миналото; миналото свършено деятелно причастие на спрегаемия глагол; неизменяема форма *ще* или формите за минало несвършено време на глагола *ща* и частицата *да*, а основните формообразувателни типове при аналитичните времена в български език са два:

– форма на глагола *ща* (неизменяема или за минало несвършено време) и два подтипа:

- форми на спрегаемия глагол в сегашно време – *ще пиша*;
- частицата *да* и форми на спрегаемия глагол в сегашно време – *щях да пиша;*

– форма на спомагателния глагол *съм* (*бъда*) и минало свършено деятелно причастие на спрегаемия глагол – *писал съм, бях писал, ще съм писал, щях да съм писал.*

II.4. Характеристика на глаголите в английски

Английският език принадлежи към групата на германските езици и въпреки че част от характеристиките му, които са общи с останалите езици от тази група, са нарушени от по-късни интензивни езикови влияния, особено от нормандски, много от общите черти са останали и в съвременния английски език. За разлика от българския

език, английският език днес се говори от над 335 000 000 души по целия свят и е официален в 47 страни. Също така за разлика от българския език, при който теоретичното осъзнаване на езиковите категории започва сравнително късно, в английския език се наследява вековната традиция на класическите граматики, изградени върху класическите средновековни езици – латински и гръцки.

В исторически план теоретичното анализиране на английския език започва през XVI век с „Кратка граматика на английския език“ на У. Булокар (Brief Grammar for English) (по Клеър 1966: 19). В този първи съпоставителен труд е представено влиянието на латинските граматики върху английския език. Най-ранните английски граматики представляват прескриптивни трудове, които трудно могат да бъдат приети като представящи актуалното развитие на езика по онова време. Глаголът е описан като една от осемте части на речта, която „изразява състояния на битието, които могат да бъдат начални, продължаващи, приключващи, подновявани, прекратявани и повтаряни” (Уард 1767: 41). В граматиката на Б. Джонсън глаголът се характеризира като броима част на речта, която притежава лице и време (Джонсън 1640: 45). Вижда се, че теоретичното представяне на глаголите в английския език на този етап, подобно на представянето им в българския език, е описателно. На по-късен етап теоретичното описание на глагола като част на речта в английски започва да включва основната му синтактична функция и придобива функционален характер – „Глаголът е част на речта, с която правим твърдения или задаваме въпроси“ (Курме 1930: 22). В по-ново време е възприета практиката характеризирането на глагола в английските граматики да става и чрез изброяване на формите му (Куърк и др. 1985: 93). Развитието на езикознанието от края на XIX и през целия XX век е белязано от множество научни направления и школи, които дават траен отпечатък върху формалното и функционално представяне на езика, а английският език обикновено се явява основен обект на разглеждане при представянето на различните научни схващания.

Описанието на характеристиките на глаголите в английски може да бъде представено чрез класифицирането им според техните лексикални и граматични особености.

В английски език глаголите също се делят във формални лексико-семантични класове и притежават определени морфологични категории (Ранкова и Иванова 1980: 88). Основните лексикални разреди, в които попадат английските глаголи по отношение на своите функционални особености, са спомагателни и пълнозначни

глаголи. Както и в български, пълнозначните глаголи в английски имат самостоятелно лексикално значение и пълна парадигма от възможните граматични значения (Филипова 2002: 38). Спомагателните глаголи са загубили своето лексикално значение и са натоварени с част от граматичното значение при формообразуването на различните времена. Глаголите в английски език се делят на преходни и непреходни според семантиката на изразяваното действие (Валейка и Буиткийн 2003: 89). В случаите, когато действието, изразявано от преходните глаголи, може да засегне пряко друг обект, който да бъде синтактично реализиран посредством пряко допълнение, глаголите попадат в класа на преходните глаголи. При непреходните глаголи действието, което изразяват, не може да засегне друг обект и съответно не е възможна синтактичната реализация на пряко допълнение (Валейка и Буиткийн 2003: 91).

Глаголите в английски, макар и само в ограничени случаи, могат да изразяват на лексикално ниво отношението на действието към някакъв предел, те могат да бъдат модифицирани с пояснения в рамките на изречението и да бъдат както пределни, така и непределни. Също така само ограничен брой глаголи, изразяващи статични състояния, постоянни признаци на предметите, психологически състояния и дейности, не могат да бъдат употребявани в продължителни времена (Филипова 2002: 82). Съществено деление при английските глаголи е групирането им в два основни граматични класа, според начините, по които се образуват формите им за минало време и минало причастие (Мърфи 1991: 23). Според това деление глаголите в английски език се делят на правилни и неправилни. Формите за минало време и минало причастие на правилните глаголи се образуват чрез прибавяне на окончание –(e)d към инфинитива. Неправилните глаголи образуват тези форми по различни формообразователни типове, които могат да бъдат разделени на отделни подтипове (Мърфи 1991: 26).

Морфологичните категории при глаголите в английски, които ще бъдат разгледани, са категориите лице, число, време и вид на действието. Граматична категория род не е глаголна категория в английски и значение за „родовост“ не се изразява посредством глаголната форма.

Спорни са различните виждания за броя на формите в глаголната парадигма в английския език – както по отношение на категориите, така и по отношение на формите им, които трябва да бъдат включени в нея. Класическите граматики и наследената традиция подкрепят виждането, че изискването за експлицитен подлог в английския език (пълнозначна именна група или местоимение) определя около сто възможни глаголни

форми в английски език (Кабакчиев 2003: 132). Структуралистките схващания от втората половина на XX век пък минимализират броя на възможните форми до пет – основна форма, третолична сегашна форма, сегашно причастие, минало причастие и инфинитив (Деклерк и др. 2008: 87). Фактически различимите формите са пет, но тези форми функционират за изразяването на различни граматични значения.

II.4.1. Категориите лице и число на глаголите в английски

По отношение на глаголните категории лице и число английският език е загубил голяма част от окончанията си, които изразяват отношението на глаголното лице към изказването. В този смисъл не може да се говори за флективни морфологични категории лице и число. Тези категории съществуват в английския език дотолкова, доколкото глаголните форми задължително се съчетават с подлог, изразен чрез именна фраза или местоимение. С формален маркер за лице и число във формите на синтетичните времена е натоварено само 3. лице, единствено число в сегашно просто време (Филипова 2002: 27). При формите на аналитичните времена със значение за лице и число са натоварени спомагателните глаголи във формите за 1. и 3. лице, единствено число, сегашно продължително и минало продължително време; за 3. лице, единствено число, сегашно перфектно, сегашно перфектно продължително време и минало перфектно време (Филипова 2002: 67).

II.4.2. Особености на английската темпорална система

Теоретичното описание на темпоралните характеристики на английските глаголи може да бъде разделено на два отделни исторически етапа – дефиниране на броя на глаголните форми и определяне на морфологичните категории.

Сред първите граматики на съвременния английски език наред с „Кратка граматика на английския език“ на У. Булокар е и „Английската граматика“ на Б. Джонсън, публикувана през 1640 г. (Джонсън 1640). В първите прескриптивни граматики глаголът е дефиниран чрез връзката му с категориите лице и време (Джонсън използва думата *time* вместо *tense*) (Джонсън 1640: 21). Б. Джонсън представя значенията на глаголните форми за трите времена (сегашно, минало и бъдеще), като ги съпоставя със спрежението на латинския глагол *обичам* (Джонсън 1640: 22). В своята граматика Б. Джонсън твърди, че другите глаголни форми се образуват посредством синтаксиса и нямат пряко отношение към времето като морфологична категория (Джонсън 1640: 23). Важно е да се спомене, че през XVII век английският език притежава пълна парадигма

на глаголните форми по отношение на категориите лице и число (Палмър 2014: 89). Разбира се, по времето, когато се появяват първите английски граматики, няма единно схващане за естеството на категорията време в английски език и различните възгледи получават различна приемственост сред научното общество. Друг възглед от този период за категорията време в английски език принадлежи на известния енциклопедист Дж. Пристли. В своите „Основи на английската граматика“ той излага възгледите си за две времена в английски, като разграничава термините за граматично време и физическо време (tense и time) и дефинира граматичното време чрез физическото: „Глаголите имат две времена; сегашно време (Present Tense), означаващо настоящето, и минало време (Preter Tense), изразяващо миналото време“ (Пристли 1761: 54). Вижданията на Ф. Палмър и Дж. Пристли за членовете на категорията време в английски език остават популярни и през XX век. Останалите глаголни форми Дж. Пристли разделя в типове и подтипове според формообразувателния им тип (Пристли 1761: 55). В инвентара на Дж. Пристли, включващ две времена, три типа и три съставни подтипа, влизат всички форми на класическата английска темпорална система. По същото време за пръв път се споменава и за вида на протичащото време – Р. Уолт в своето „Кратко въведение в английската граматика с критични бележки“ говори за тричленна темпорална система, но споменава за свършени (Perfect) и несвършени (Imperfect) форми, различавайки 9 темпорални форми – 3 неопределени, без допълнително значение за вида на действието и 6 определени, които могат да бъдат свършени и несвършени (Уолт 1762: 34). Вижда се, че още на този етап граматичното значение на категориите на английската темпорална система се тълкува описателно чрез наименованията на видовете глаголни форми. През XIX век се появяват множество дескриптивистки граматики за целите на училищното образование. В граматиката на Смит и Хал от 1889 е представена тричленната система на английските времена с подвидове от форми – неопределен вид (indefinite forms), незавършен вид (incomplete forms), завършен вид (complete forms) и перфектен незавършен вид (perfect-incomplete forms) (Смит и Хал 1889: 50). По това време започва да се усеща влиянието на класическата латинска традиция и съвременния дескриптивизъм. Това влияние се проявява по два начина. В латински език има 6 времена, като в него липсват прогресивни (или продължителни) времена, а също така имперфектът в латински има различна семантика спрямо миналото време в английски. Като се има предвид статутът на латинския език през XVIII век, не е изненадващо, че много от ранните граматици са се опитвали да дефинират шестчленна система на времената в английски с известни колебания в терминологията. В тези системи

продължителните форми не се приемат за времена и обикновено се разглеждат като варианти на тези шест времена (Смит и Хал 1889: 50). Вторият начин, по който тогавашните граматици се опитват да се доближат до латинската традиция, е чрез изкуственото приписване на класическите имена на времената върху английските глаголни форми (Джонсън 1766: 66). В училищната граматика на Л. Мъри от 1795 г. несистемно са използвани класическите наименования на латинските времена, като към системата са прибавени първо и второ бъдеще време (Куакенбос 1869: 90).

Схващането за тричленна или двучленна категория време с 2 или 3 члена на категорията вид на действието се утвърждава през втората половина на XIX и началото на XX век. Виждания за дванадесетчленна темпорална система се появява в началото на XX век с ясно дефиниран инвентар, но с различни наименования. В граматиката на У. Баскервил и Дж. Сеуел за пръв път са изброени 12 темпорални форми – 3 времена с по 4 вида на действието (Баскервил и Сеуел 1895: 152). Тази система се запазва в училищните граматики през XX век. Разбира се, необходимо е да се споменат и алтернативните възгледи за броя на „същинските“ времена. Подобно на Дж. Пристли, модерните лингвисти, базирайки се на морфологичните особености на глаголните форми в английски, различават минало и неминало време. Това се дължи на факта, че те вземат предвид единствено формите на пълнозначните глаголи, без да отчитат употребата на модални и спомагателни глаголи (Грийнбаум и Нелсън 1991: 77). В някои от съвременните теоретични английски граматики се говори за две морфологични времена в английски език, като се отчита спрежението – минало и бъдеще (Хъдълстън и Пулъм 2002: 250). Съществуват и по-усложнени съвременни системи, в които са описани основни и вторични времена, аспектуални форми и модални форми. За целите на тази работа се възприема схващането за граматичната категория време в английски като хиперкатегория, изградена от морфологични категории, които добавят своите значения при формообразуването.

Следвайки представянето на българската темпорална система, ще бъдат представени характеристиките на морфологичните категории, изграждащи английската темпорална система, както и формално-семантичното представяне на всяко от времената поотделно.

Вижданията за 2 същински морфологични времена в английски език са доста популярни и могат да бъдат подкрепени с различни лингвистични доказателства. Тези доказателства могат да бъдат разделени основно в две групи – формообразователни и

семантични (Деклерк и др. 2008: 100). Формообразувателните доводи са свързани с виждането, че същинските морфологични категории са тези, които имат формообразувателна парадигма, представена само чрез синтетични словоформи – единствените синтетични флективни форми на английските глаголи, които изразяват темпорални признаци, са сегашното причастие, миналото причастие и третоличната сегашна форма (Деклерк и др. 2008: 101). Семантичните доводи са тясно свързани с формообразувателните и изхождат от виждането, че семантичната опозиция, която може да бъде изразена флективно от глаголни форми в английски, е опозицията минало – неминало (Деклерк и др. 2008: 101). Както формообразувателните, така и семантичните виждания приемат аналитичните глаголни форми в английски за синтактични конструкции, при които комбинаторното структурно изразяване на морфологичната опозиция минало – неминало формира темпоралното значение (Хъдълстън и Пулъм 2002: 251). Представените доводи не могат еднозначно да бъдат приети като изчерпателни по отношение на броя на „същинските“ времена. Аналогично, за български може да се твърди, че има само 3 морфологично маркирани времена – сегашно, минало несвършено и минало свършено време. Прилагането на този критерий за български би показал, че морфологично изразените времена в рамките на една синтетична словоформа отново образуват опозицията минало – неминало, а в рамките на маркираните с минало време форми образуват подвидове: свършено – несвършено време. Може да се твърди, че при прилагането на критерия за морфологично маркирани словоформи в английски език не може да се говори са 2 времена, а за 2 вида времена, образуващи се по различни формообразувателни модели (Деклерк и др. 2008: 105).

Спорен е статусът и на бъдещите времена в английски език. Твърди се (Куърк и др. 1985: 213), че в английски език няма бъдещи времена и че глаголът will е модален глагол. Посочвано е също (Хъдълстън и Пулъм 2002: 188), че бъдещите времена изразяват епистемична модалност. Въпреки тези схващания, съществуват и мнения, които обуславят наличието на бъдещи времена в английски. Тези виждания (Деклерк 1991: 8), които опирайки се на дефинициите за време на Дж. Лионс (1977: 68) и Б. Комри (1985: 9), че времето е граматическо средство за локализиране на събития във времето, обясняват темпоралната принадлежност на формите за бъдеще време с характеристиката да локализират събития във времето, без оглед на възможността за тяхното реализиране.

Повечето съвременни модели на глаголните морфологични категории в английски приемат виждането за 3 същински времена – минало, сегашно и бъдеще (Хаус и Стивънс 2000; Луис и Мол 2013). Съществуват също така мнения за 4 същински времена – минало, сегашно, бъдеще и бъдеще в миналото или минало бъдеще време (Ранкова и Иванова 1980; Филипова 2002; Кабакчиев 2003). Наличието на бъдещи времена в миналото в инвентара на същинските времена е оправдано от гледна точка на категорията таксис и споровете около нейното съществуване в английския език (Деклерк и др. 2008: 97). От гледна точка на твърде ограничената употреба на формите на бъдеще време в миналото – при формите за условно наклонение, редица автори причисляват тези глаголни форми към морфологичната категория наклонение в английски език. Причисляването му към категорията на времена е мотивирано от редица прагматични съображения.

Диференциалните признаци на морфологичната категория време в английски език са два – предходност и следходност (Кабакчиев 2003: 132). С липсата на информация за предходност или следходност на действието е натоварено сегашно време. С информация за предходност е натоварено минало време, а за следходност – бъдеще време и бъдеще време в миналото. Бъдеще време в миналото носи информация за следходност на действието спрямо минал ориентационен момент.

В тази дисертация се приема общоприетото виждане за четири същински времена и четири члена на категорията вид на действието (четири вида) (Кабакчиев 2003: 131).

В рамките на същинските времена се обособяват абсолютни и относителни времена. В рамките на групата на абсолютните времена попадат минало, сегашно и бъдеще време, а бъдеще време в миналото е единственото относително време в английски език, чийто ориентационен момент е сигнализиран от миналото време в главното изречение (Кабакчиев 2003: 136). В английски език категорията вид на действието е сложна морфологична категория, която съотнася продължителността или завършеността на действието към външен (спрямо самото действие) ориентационен момент (Валейка и Буиткийн 2003: 84). Действието може да бъде продължително, завършено, да липсва информация за това или да бъде продължително, но завършено спрямо всеки един от ориентационните моменти на четирите времена, като тези отношения са независими от лексикалното значение на глагола (Филипова 2002: 42). Някои автори разлагат морфологичната категория вид на действието в английски на две самостоятелни категории – категорията вид (aspect) (Валейка и Буиткийн 2003: 84), която

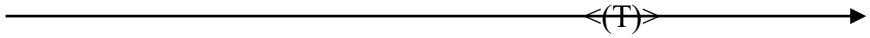
изразява продължителността или липсата на информация за продължителност на действието спрямо определен ориентационен момент, и категорията подредба (order), която отнася завършеността на действието към определен ориентационен момент – подрежда момента на действието преди ориентационния момент (Валейка и Буиткийн 2003: 97). Категорията вид на действието в английски език може да се разглежда като сложна категория, която включва в себе си синкатегориални значения. Категорията вид на действието в английски дава информация за завършеността и продължителността на действието или липсата на такава информация спрямо момента на референция на четирите времена. Разлагането на тази категория на две подкатегории е оправдано от прилагането на принципа на Смирницки, при който две последователно маркирани грамеми в една и съща словоформа не могат да бъдат членове на една морфологична категория (Смирницки 1955). Въпреки това в английския език е прието разглеждането на категорията вид на действието като една категория, изградена от еквиполентната опозиция продължителност – резултативност (Ранкова и Иванова 1980; Филипова 2002; Кабакчиев 2003). Диференциалните признаци на категорията вид на действието в английски език са 2: продължителност и перфективност (завършеност/резултативност) (Кабакчиев 2003: 134). С липсата на информация за продължителност или завършеност са натоварени т.нар. прости времена (simple tenses) (Филипова 2002: 69). Продължителните времена (continuous tenses) са носители на признака продължителност (Филипова 2002: 82), а перфектните времена (perfect tenses) – на признака завършеност (Филипова 2002: 90). Наличието на двата признака се наблюдава в сложните форми на перфектните продължителни времена (Филипова 2002: 101). В английския език различните видове времена се наименуват по следния начин: първо се назовава времето, а след това – видът на действието. Така, за разлика от български, където в голяма степен се пази традиционното българско или латинско наименование на времената, което по-скоро ситуира действието във времето, без да дава информация за неговата завършеност, в английски директно се назовава и част от значението на съответното време. Представеното описание по-горе дава основание да се твърди, че инвентарът на английските времена придобива йерархично организирана структура. За целите на тази дисертация – с оглед на съпоставителния анализ на българско-английските граматични паралели, ще се приложи същият формален семантичен анализ и за английската темпорална система, подобно на българската.

II.4.3. Семантична характеристика на английската темпорална система

Както беше посочено, системата на Х. Райхенбах може да функционира като вид автономна логическа конструкция, лежаща в основата на формално-семантичното описание на времената. Както за български, така и за английски е валидно, че глаголните форми изразяват темпорални, таксисни и аспектуални значения, именно защото става дума за различни категории с различен инструментариум и място в глаголната система (Вейга 1991: 21). Съвместното присъствие на тези значения в текста изисква формален анализ на инвариантното значение, което обосновава разглеждането на формите в системата на езика (Чергова 2014: 187). Системата на Х. Райхенбах (Райхенбах 1947) е приложена първоначално за английски език, а впоследствие е детайлно анализирана и приложена и за други езици.

Преди да бъдат представени формално-семантичните описания на времената в английски, е необходимо да бъде уточнено, че в своята система Х. Райхенбах извежда отношения между изграждащите я елементи, които не се реализират в английски език, но са възможни в други езици. Също така при системата, описана от Х. Райхенбах, няма интервали, а само моменти. Всъщност Х. Райхенбах нарича продължителните времена в английски разширени (extended) (Райхенбах 1947: 290), като по този начин представя тяхното протичане като момент с определена продължителност. В системата също липсва и момент, в който е наличен резултат от действие. За английски език ще бъдат използвани означенията и отношенията от системата на Х. Райхенбах, но модифицирани, за да изразят по-точно значенията на отделните английски времена.

II.4.3.1. Сегашно просто време



(R); < E >

E = R = T[5]

Сегашно просто време в английски език изразява обичайно, повтарящо се действие в настоящето (Филипова 2002: 72). При изграждането на семантиката на това време моментът на говорене, моментът на референтност и моментът на действие съвпадат напълно, без да има информация за завършеност или повторяемост на действието. Може да означава завършено цялостно действие в настоящето, когато се

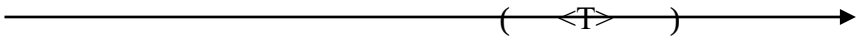
[5] Представените за английски формули са по Х. Райхенбах 1947 и А. Приор 1967

изброяват няколко последователни действия, и общоизвестен факт, научна или неопровержима истина (Ранкова и Иванова 1980: 97). Сегашно просто време може да се използва и за да означи бъдещи и минали действия, като в тези случаи употребата е стилистична (Ранкова и Иванова 1980: 98).

(10) Пример за сегашно просто време:

*The new initiative **aims** to contribute to the development, enhancement and promotion of Balkan cinematography.*

II.4.3.2. Сегашно продължително време



(R); < E >

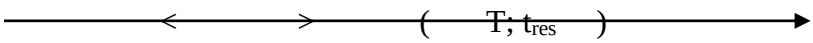
T = R; E ⊇ R

Семантиката на сегашно продължително време в английски език изразява действие, при което моментът на действие обхваща момента на референтност, който съвпада с момента на говорене. Това време може да изразява действие, което се извършва в настоящия момент, или действие, което обхваща период в настоящето, като е започнало преди момента на говорене и вероятно ще продължи и след него (Филипова 2002: 84 – 58). Сегашно продължително време може да се употребява и за бъдещи действия, които се представят като сигурни (Ранкова и Иванова 1980: 104); обикновено означава фонови действия или състояния за други действия или състояния, които се предават с глаголи в сегашно просто време.

(11) Пример за сегашно продължително време:

*Balkan countries **are making** progress.*

II.4.3.3. Сегашно перфектно време



(R); < E >

E < T = R; t_{res} ∈ R

Сегашно перфектно време изразява действие в миналото, което е преди момента на говорене. Сегашно перфектно време е резултативно време, като резултатът е наличен към момента на говоренето, който съвпада с момента на референтност.

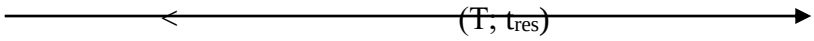
Формите за това време в английски език изразяват действия, извършени в неопределен момент (период) в миналото (Филипова 2002: 92), а също и действие, което

е извършено в миналото, но има някаква връзка с настоящето или резултатът от действието е налице в настоящето (Ранкова и Иванова 1980: 109). Сегашно перфектно време се употребява, когато времето на действието не е означено точно, действието е току-що завършено или действието е извършено в период, който още не е завършил (Ранкова и Иванова 1980: 109 – 111).

(12) Пример за сегашно перфектно време:

*I know that you **have cut** the furrow!*

II.4.3.4. Сегашно перфектно продължително време



(R); < E >

$T \in R$; $E < R$; $t_{res} \in R$

Семантиката на сегашно перфектно продължително време в английски език се изгражда върху момент на действие, който е преди моментите на референтност и говорене, които съвпадат, но резултатът от действието е наличен в рамките на момента на референтност, без информация за приключване на действието.

При употреба на сегашно перфектно продължително време се изразяват продължителни действия, които са започнали в неопределен период в миналото, но са свързани с настоящия момент (Филипова 2002: 103). Сегашно перфектно продължително време означава също и действия, които са започнали неотдавна в миналото, но продължават и в настоящето, или действия, които са се повтаряли в близкото минало (Ранкова и Иванова 1980: 116 – 117).

(13) Пример за сегашно перфектно продължително време:

*They say she **has been fasting** for months.*

II.4.3.5. Минало просто време

(R); < E >

$E = R < T$

Семантиката на минало просто време изразява действие, чийто момент съвпада с минал момент на референтност.

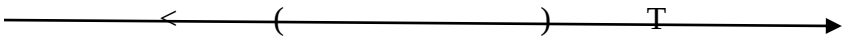
Минало просто време може да изразява еднократно действие, извършено в даден момент или период в миналото, без връзка с настоящето (Филипова 2002: 77). Може да

изразява и обичайни повтарящи се действия в миналото, действие или състояние, което е било характерно за подлога в миналото, общоизвестни или научни факти, които се отнасят за миналото (Ранкова и Иванова 1980: 99 – 100)

(14) Пример за минало просто време:

*I **gave** it to her.*

II.4.3.6. Минало продължително време



(R); < E >

R < T; E ⊃ R

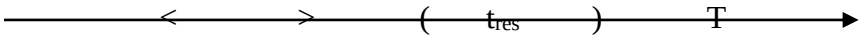
При минало продължително време в английски език моментът на референтност е преди момента на говоренето, а моментът на действието преминава през него – началото на действието е преди момента на референтност и липсва информация за прекратяване на действието в рамките му, както и към момента на говорене. Минало продължително време е едно от „разширените“ времена, при които събитието обхваща определен период от време (Райхенбах 1947: 290)

Минало продължително време изразява действие, което се извършва в определен минал момент (момента, за който се говори), или действие, което обхваща период от миналото, като е започнало преди него и продължава и след него (Филипова 2002: 87). Понякога минало продължително време изразява действие или състояние, което служи за фон на други действия, извършвани в миналото (Ранкова и Иванова 1980: 106)

(15) Пример за минало продължително време:

*He **was making** noise.*

II.4.3.7. Минало перфектно време



(R); < E >

E < R < T; t_{res} ∈ R

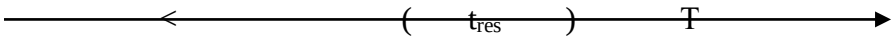
Семантиката на минало перфектно време в английски език се изразява чрез момента на действието, който е предхождащ спрямо момент на референтност в миналото, но резултатът от действието е наличен в момента на референтност.

Минало перфектно време изразява действие, завършено преди даден момент, период или друго действие в миналото (Филипова 2002: 98). Изразява също и действие, което е започнало преди даден момент, но продължава и през този момент (Ранкова и Иванова 1980: 113).

(16) Пример за минало перфектно време:

*My grandfather **had achieved** his glory during the war.*

II.4.3.8. Минало перфектно продължително време



(R); < E >

$R \supseteq E$; $R<T$; $t_{res} \in R$

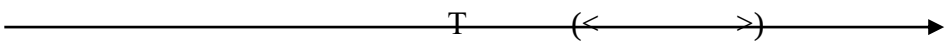
Минало перфектно продължително време изразява действие, чийто момент е преди минал момент на референтност, а резултатът от действието е наличен в момента на референтност. При глаголните форми на това време липсва информация за прекратяване на момента на действието в рамките на момента на референтност или преди него.

При типичната си употреба минало перфектно продължително време изразява действия или състояния, които са започнали в неопределен момент или период и са минали спрямо други действия или период, които от своя страна също са минали спрямо момента на говорене; означените действия или състояния продължават в момента, спрямо който се определят като минали и резултатът от тях е наличен в този момент или период (Филипова 2002: 106 – 107).

(17) Пример за минало перфектно продължително време:

*He **had been dreaming** for this ever since.*

II.4.3.9. Бъдеще просто време



(R); < E >

$T < R = E$

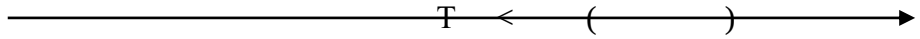
Значението на бъдеще просто време изразява момент на референтност, в който се съдържа моментът на действието и който е след момента на говорене.

Бъдеще просто време изразява действия или състояния, бъдещи спрямо момента на говорене; може да означава свършени и несвършени действия, краткотрайни и дълготрайни действия, еднократни действия, обичайни и общовалидни действия, повтарящи се действия, повтарящи се ограничен или неограничен брой пъти (Филипова 2002: 81 – 82).

(18) Пример за бъдеще просто време:

Macedonia ***will involve*** *itself in new politics.*

II.4.3.10. Бъдеще продължително време



(R); < E >

T < R; R ⊇ E

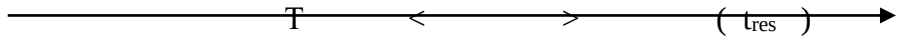
При бъдеще продължително време в английски език моментът на референтност е следхождащ по отношение на момента на говоренето, а моментът на действието е включен в момента на референтност.

Формите на глаголите за бъдеще продължително време изразяват действия, които ще се извършват в определен бъдещ момент, като може да са започнали преди него и да продължават в него, а също и действия, които обхващат определен бъдещ период (Филипова 2002: 89 – 90).

(19) Пример за бъдеще продължително време:

I am sure that the poor man ***will be killing*** *himself.*

II.4.3.11. Бъдеще перфектно време



(R); < E >

T < E < R; $t_{res} \in R$

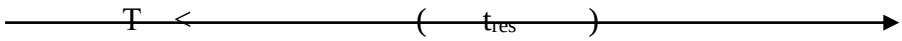
При бъдеще перфектно време моментът на действието е след момента на говорене, а резултатът от действието е наличен в бъдещ момент на референтност, който е след момента на говорене и момента на действието.

Бъдеще перфектно време изразява действие, завършено до (преди) даден момент или период в бъдещето (Филипова 2002: 100). Също така може да изразява действие, което с някаква сигурност ще бъде завършено спрямо определен бъдещ момент (Ранкова и Иванова 1980: 114)

(20) Пример за бъдеще перфектно време:

*In two days we **will have moved** and settled down.*

II.4.3.12. Бъдеще перфектно продължително време



(R); < E >

$R \supseteq E$; $T < E$; $t_{res} \in R$

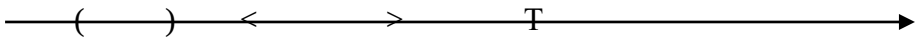
Семантиката на бъдеще перфектно продължително време се изгражда от момент на референтност, който е бъдещ спрямо момента на говоренето, и резултат от действие, който е наличен в момента на референтност. Моментът на действието е бъдещ спрямо момента на говорене.

Глаголните форми за това време в английски език изразяват действия или състояния, бъдещи спрямо ориентационния момент или период, който от своя страна също е бъдещ спрямо момента на говорене; характерното е, че действието е само несвършено и се включва в ориентационния момент или период (Филипова 2002: 109).

(21) Пример за бъдеще перфектно продължително време:

*I **will have been working** there.* (от Ранкова и Иванова 1980: 120)

II.4.3.13. Бъдеще просто време в миналото



(R); < E >

$R < E < T$

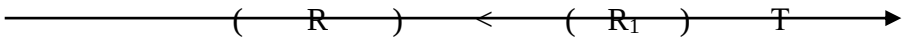
При бъдеще просто време в миналото моментът на действието е следходен спрямо минал момент на референтност. Семантиката на това време в английски език дава основание то да бъде причислено към относителните времена, т.е. формите му изразяват предстоящо действие спрямо ориентационен момент.

Формите на глаголите за това време изразяват действия или състояния от различен тип – свършени и несвършени действия, краткотрайни и дълготрайни действия, еднократни действия, обичайни и общовалидни действия, повтарящи се действия, повтарящи се ограничен и неограничен брой пъти, които са бъдещи спрямо ориентационния момент, който от своя страна е минал спрямо момента на говорене (Филипова 2002: 111).

(22) Пример за бъдеще просто време в миналото*:*

*We **would complain** about this.*

II.4.3.14. Бъдеще продължително време в миналото



(R) (R_1); < E >

$T > R > R_1$; $E \supset R_1$

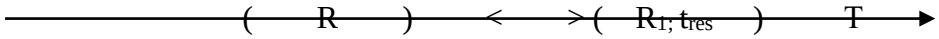
При бъдеще продължително време в миналото в английски език и двата момента на референтност са преди момента на говоренето, единият момент предхожда другия, а момента на действието е бъдещ спрямо първия момент на референтност и достига до втория – началото на действието е преди втория момент на референтност и липсва информация за прекратяване на действието в рамките му.

Бъдеще продължително време в миналото изразява действия, които са били извършени в даден момент или период от време, който е бил бъдещ по отношение на даден момент в миналото (Ранкова и Иванова 1980: 107).

(23) Пример за бъдеще продължително време в миналото*:*

*They **would be moving** faster, if it wasn`t the accident.*

II.4.3.15. Бъдеще перфектно време в миналото



(R) (R_1); < E >

$R < E < R_1 < T$; $t_{res} \in R_1$

Семантиката на бъдеще перфектно време в миналото се изгражда от два последователни момента на референтност, предхождащи момента на говорене. Моментът на действието се намира между двата момента на референтност, като спрямо

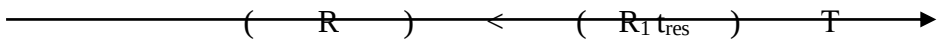
първия е бъдещ, а резултатът от действието е наличен във втория момент на референтност.

Това време в английски език изразява действие, което е било свършено в даден момент, който е бил бъдещ по отношение на момент в миналото (Ранкова и Иванова 1980: 115).

(24) Пример за бъдеще перфектно време в миналото:

*I **would not have had** problems, if I was not there.*

II.4.3.16. Бъдеще перфектно продължително време в миналото



(R) (R_1); < E >

$R < E < R_1 < T$; $E \supset R_1\ t_{res} \in R_1$

Значението на бъдеще перфектно продължително време в миналото се изгражда от два последователни момента на референтност, минали спрямо момента на говорене. Моментът на действието е бъдещ спрямо първия момент на референтност, като достига и/или обхваща изцяло втория, като резултатът от действието е наличен във втория момент на референтност и в него липсва информация за прекратяване на действието.

Бъдеще перфектно продължително време в миналото означава действие, което е започнало преди даден момент, който е бил бъдещ по отношение на минал момент и е завършено преди него или продължава и след него (Ранкова и Иванова 1980: 119)

(25) Пример за бъдеще перфектно продължително време в миналото:

*He **would have been working*** (от Ранкова и Иванова 1980: 118)

II.4.4. Особености на формообразуването на глаголните форми в английски език

Основна характеристика на английското формообразуване при глаголите е силно развитата аналитичност при глаголните форми. Морфологично маркираните синтетични форми, които изразяват темпорални значения, са 2, а комбинаторното им прилагане при спомагателните и пълнозначните глаголи в рамките на темпоралните форми формира граматичното им значение. Глаголите в английски имат 3 нелични глаголни форми, които влизат в употреба при формообразуването на времената – инфинитив, минало

причастие и сегашно причастие (Филипова 2002: 28, Ранкова и Иванова 1980: 88). По начина, по който глаголите образуват формите за минало време и миналото си причастие, те се делят на правилни и неправилни. Също така синтетичните глаголни форми в английски не носят информация за глаголното лице и число, освен в 3. лице единствено число, сегашно просто време (Ранкова и Иванова 1980: 90).

Тук е мястото да се отбележи друга основна характеристика на английския език, която е по-скоро синтактична, но оказва влияние на формообразуването – правилото за съгласуване на времената. Тази характеристика предполага, че времето на глаголите в подчинените изречения в английски зависи от времето на глагола в главното (Филипова 2002: 250). При това правило, когато глаголът в главното изречение е в някое минало време, глаголът в подчиненото изречение трябва да бъде също в минало време. Когато действието в подчиненото изречение става едновременно с действието в главното, употребява се минало просто или минало продължително време. Когато действието на подчиненото изречение предхожда това в главното, употребява се минало перфектно време. Когато действието в подчиненото изречение е бъдещо по отношение на действието в главното, употребява се бъдеще време в миналото (Ранкова и Иванова 1980: 226). Това правило действа при структуриране на сложни изречения и последователни изказвания и е по-скоро синтактична особеност на изграждане на текст. Принципът, който задължително се следва, е базиран на последователното съотнасяне на действията едно спрямо друго.

Темпоралните глаголни форми в английския език се формират по следния начин:

Синтетични форми:

– Сегашно просто време се образува от инфинитива на глагола, като при 3. лице, единствено число се добавя окончание –(e)s (Ранкова и Иванова 1980: 96);

– Минало просто време се изразява от единствената форма за минало време на глагола (Ранкова и Иванова 1980: 98);

Аналитични форми:

– Бъдеще просто време се образува със спомагателните глаголи *shall* (will) за 1. лице, единствено и множествено число и *will* за останалите лица и инфинитива на спрегаемия глагол (Ранкова и Иванова 1980: 101);

– Бъдеще просто време в миналото се образува от миналото време спомагателните глаголи *shall* (will) за 1. лице, единствено и множествено число и *will* за останалите лица и инфинитива на спрегаемия глагол (Ранкова и Иванова 1980: 102);

– Сегашно продължително време се образува от формите на спомагателния глагол *be* в сегашно време и сегашното причастие на спрегаемия глагол (Ранкова и Иванова 1980: 103);

– Минало продължително време се образува от формите на спомагателния глагол *be* в минало време и сегашното причастие на спрегаемия глагол (Ранкова и Иванова 1980: 105);

– Бъдеще продължително време се образува от спомагателните глаголи *shall* (will) за 1. лице, единствено и множествено число и *will* за останалите лица и продължителния инфинитив на спрегаемия глагол (*be* и сегашно причастие) (Ранкова и Иванова 1980: 106);

– Бъдеще продължително време в миналото се образува от миналото време на спомагателните глаголи *shall* (will) за 1. лице, единствено и множествено число и *will* за останалите лица и продължителния инфинитив на смисловия глагол (Ранкова и Иванова 1980: 107);

– Сегашно перфектно време се образува от спомагателния глагол *have* в сегашно време и миналото причастие на спрегаемия глагол (Ранкова и Иванова 1980: 109);

– Минало перфектно време се образува от миналото време на глагола *have* и миналото причастие на спрегаемия глагол (Ранкова и Иванова 1980: 112);

– Бъдеще перфектно време се образува от спомагателните глаголи *shall* (will) за 1. лице, единствено и множествено число и *will* за останалите лица и перфектния инфинитив на спрегаемия глагол (*have* и минало причастие) (Ранкова и Иванова 1980: 113);

– Бъдеще перфектно време в миналото се образува от миналото време на спомагателните глаголи *shall* (will) за 1. лице, единствено и множествено число и *will* за останалите лица и перфектния инфинитив на спрегаемия глагол (Ранкова и Иванова 1980: 114);

– Сегашно перфектно продължително време се образува от спомагателния глагол *have* в сегашно време, миналото причастие на спомагателния глагол *be* и сегашното причастие на спрегаемия глагол (Ранкова и Иванова 1980: 116);

– Минало перфектно продължително време се образува от формите на спомагателния глагол *have* в минало време, миналото причастие на спомагателния глагол *be* и сегашното причастие на спрегаемия глагол (Ранкова и Иванова 1980: 117);

– Бъдеще перфектно продължително време се образува от формите за сегашно време на спомагателните глаголи *shall* (will) за 1. лице, единствено и множествено число и *will* за останалите лица и перфектния продължителен инфинитив на спрегаемия глагол (*have* и *been* и сегашно причастие на глагола) (Ранкова и Иванова 1980: 118);

– Бъдеще перфектно продължително време в миналото се образува от миналото време на спомагателните глаголи *shall* (will) за 1. лице, единствено и множествено число и *will* за останалите лица и перфектния продължителен инфинитив на спрегаемия глагол (Ранкова и Иванова 1980: 119).

При аналитичните форми на глаголите в английски формообразувателните елементи са 6 – инфинитивът, сегашното и миналото причастие на спрегаемия глагол, форми на спомагателния глагол *be*, форми на спомагателния глагол *have* и форми на спомагателните глаголи *will*/*shall*. Формообразувателните типове са 7:

– форми на *will*/*shall* и инфинитив на глагола – *will write, would write*;

– форми на *be* и сегашното причастие на глагола – *am writing, was writing*;

– форми на *have* и миналото причастие на глагола – *have written, had written*;

– форми на *will*/*shall, be* и сегашното причастие на глагола – *will be writing, would be writing;*

– форми на *have* и продължителен инфинитив (миналото причастие на *be* и сегашното причастие на глагола) – *have been writing, had been writing*;

– форми на *will*/*shall* и перфектен инфинитив (форми на *have* и миналото причастие на глагола) – *will have written, would have written*;

– форми на *will/shall* и перфектен продължителен инфинитив (*have been* и сегашното причастие на глагола) – *will have been writing, would have been writing.*

II.5. Съпоставка на характеристиките на темпоралните системи в български и английски

Може да бъде направен съпоставителен анализ в плана на историческото развитие на схващанията за същността на темпоралните категории в двата езика и в плана на съвременното състояние на теоретичните схващания.

В исторически план се вижда, че теоретичното описание на характеристиките на темпоралните системи в български и английски преминава през сходни етапи. Теоретичните лингвистични школи и на двата езика са подложени на чужди влияния, които в ранните етапи налагат изкуственото прокарване на граматични правила и морфологични категории, които са несвойствени за езиците към момента на текущото им историческо развитие. Ще бъдат представени обобщено етапите, през които се преминава, и влиянието, което тези етапи оказват на съвременните теории, като историческият съпоставителен анализ на развитието на теоретичните виждания за характеристиките на глаголните системи в двата езика е извън целите на настоящата работа.

Първият етап, през който преминава теоретичното осъзнаване на темпоралните системи в български и английски, е етапът на дефинирането на същинското съдържание и броя на темпорални форми в двата езика. През този етап споровете между изследователите на езика са свързани основно с това кои са глаголните форми, които принадлежат към категорията време, и с какво граматично значение са натоварени те. Вторият сходен етап от историческото развитие на описанието на темпоралните системи в български и английски е „разлагането“ на традиционните темпорални системи и дефинирането на същинските отношения между изграждащите ги категории. И в двата езика от категорията време първоначално е отделена категорията вид на действието, като изграждащите я значения и опозиции остават спорни през по-нататъшното развитие на лингвистичните теории. На трето място след категорията време и категорията вид на действието е отделена категорията таксис, която в българското езикознание може да се определи като по-скоро стабилна по отношение на описанието си, докато в английски език и до днес се спори за мястото ѝ в темпоралната система на езика.

От гледна точка на съвременното състояние на лингвистичните схващания за български и английски език съпоставителният анализ може да бъде представен в две направления – във връзка с особеностите на глаголите и във връзка с особеностите на темпоралните системи в двата езика.

От представените характеристики на глаголите в български и английски се вижда, че те споделят еднакви категориални особености при класификацията в граматични класове. И в двата езика има пълнозначни и спомагателни глаголи, като спомагателните са натоварени с част от граматичното значение при различните форми на времената. Съществена разлика между български и английски е, че българските глаголи притежават лексикално-граматичната категория вид на глагола. Категорията вид на глагола в български изразява отношението на действието към неговия предел (Куцаров 2007: 551). Това отношение не се отнася до определен момент от времето, а се определя като вътрешноприсъщо за семантиката на самия глагол – завършеността или незавършеността на действието се изразяват на лексикално ниво във всички форми на дадения глагол. В английски такова отношение се изразява от ограничен брой глаголи на лексикално ниво (Филипова 2002: 82), а при по-голяма част от глаголите завършеността или незавършеността на действието се изяснява в контекста от допълнителни лексикални елементи (Валейка и Буиткийн 2003: 84). Морфологичната категория вид на действието в английски определя действието като завършено или продължително спрямо определен деиктичен момент, външен за самото действие. Така действието може да бъде завършено или незавършено спрямо всеки момент от темпоралната ос, но това значение не е вътрешноприсъщо на семантиката на изразяваното действие. Иначе казано, видът на глагола в български, като лексикално-граматична категория, изразява завършеността/незавършеността на действието спрямо границата на самия процес (вътрешна (не)завършеност, цялостност), докато в английски категорията вид на действието изразява завършеността на действието спрямо някаква друга граница по темпоралната ос (външна (не)завършеност, продължителност). Приема се, че категорията вид на глагола не е граматическа за английски език за разлика например от славянските езици, в които формата на глагола показва и вида му (Ницолова 2008: 248). В английските граматики видът на глагола се разглежда в по-голяма степен като синтактична категория, независимо дали рефлектира с перфектността или перфектността и прогресивността (Грийнбаум и Нелсън 1991: 97). В тази дисертация се следва граматическото представяне на К. Кабакчиев (Кабакчиев 1992), в чийто труд

видът на глагола се определя като характеристика, присъща по-скоро на изказването или на изречението, отколкото само на глагола (Кабакчиев 1992: 42). Пак там авторът обобщава факторите, които оказват влияние върху видовата характеристика на изречението. Отношение към проблема за вида в английски имат формалната структура и граматическата оформеност на подлога и допълнението, лексикалното значение на подлога, допълнението и сказуемото, противопоставянето на лични и нелични местоимения, различаването на пределни и непределни глаголи, членуването (Кабакчиев 1992: 56). Обстоятелственото пояснение е най-мощният показател за вида. Може да се направи изводът, че английският глагол има маркирани откъм вид форми само за изразяване на несвършени действия и това са прогресивните форми, образувани със сегашното причастие на глагола. А видовото значение в изречение с немаркирани откъм вид глаголни форми зависи от редица фактори (Кабакчиев 1992: 48).

В темпоралната система на българския език и в английската темпорална система има наличие на различни по вид морфологични категории, които изразяват завършеност или резултат от действие спрямо определен ориентационен момент – категорията вид на действието в български и категорията вид на действието в английски. За разлика от английски, българският език не притежава морфологична глаголна категория, която изразява продължителността на действието, но притежава лексикално-граматична категория вид на глагола, която изразява незавършеността на действието.

Синтетичните форми в български език са натоварени с информация за глаголното лице и число, докато в английски с подобна информация е натоварена само формата за 3. лице, единствено число в сегашно просто време. При аналитичните форми се забелязват доста близки особености. И двата езика образуват голяма част от тези форми с помощта на глагола *съм*/*be* в различните му форми. Докато в български този глагол има пълна парадигма в сегашно време, то в английски само формите за сегашно време 1. и 3. лице, единствено число се различават от формите за 2. лице, единствено число и 1., 2. и 3. лице, множествено число. Така в сегашно време глаголът *съм* в български има 6 форми, а в английски глаголът *be* – 3. В минало свършено време глаголът *съм* притежава 5 форми (формите за 2. и 3. лице, единствено число не се различават), докато в английски формите за минало време на *be* са 2 – *was* за 1. и 3. лице, единствено число и *were* за 2. лице, единствено число и 1., 2. и 3. лице, множествено число. По отношение на формите за бъдеще време – и двата езика образуват бъдеще време с неизменяема форма на спомагателен глагол (за български това е „вкаменената“ частица *ще*, а за

английски – *will*). И в двата езика тези спомагателни глаголи могат да бъдат спрягани в други времена за целите на формообразуването. Освен формите на спомагателните глаголи *съм*/*be* и *ща*/*will*, английският език използва и още един спомагателен глагол – формите на глагола *have*, спрегнати в различни времена. В сегашно време този глагол има 2 форми – различава се само 3. лице, единствено число, а в минало време формата е една за всички лица. Така при формообразуването в български език има 2 спомагателни глагола във формите за изразяване на различните времена, а в английски – 3. По отношение на неличните глаголни форми отново се забелязва по-големият им брой в английски. В английски се използват формите на сегашното причастие, миналото причастие (включително и на глагола *be*) и инфинитивът, а в български се използва само една нелична глаголна форма – миналото свършено деятелно причастие, което е натоварено с допълнителна информация относно граматичния род на глаголното лице. Така по отношение на формообразувателния инвентар английският използва 3 нелични глаголни форми, а българският – само 1. Българският език също така използва и частицата *да* при образуването на формите за бъдещи времена в миналото.

Макар да са налични и в български, и в английски, категориите лице и число са застъпени с много по-малка честота в английски. В български език формален маркер за лице и число имат всички форми на глаголите, като само в изолирани случаи маркерите съвпадат – само формите за 2. и 3. лице, единствено число за минало несвършено време, минало свършено време и минало предварително време. Съществена разлика между двата езика е наличието на морфологичната категория род при глаголните форми в български и отсъствието ѝ в английски. Така формите на българските глаголи, които са натоварени в максимална степен с информация за глаголното лице, позволяват подлогът да не бъде изразен експлицитно, докато в английски глаголните форми, които носят подобна информация само в ограничен брой от случаите, изискват глаголното лице задължително да бъде изразено чрез подлог.

При разглеждането на характеристиките на глаголите и по-специално на структурата на глаголното сказуемо в английски и български език би следвало да се вземат под внимание някои основни синтактични особености. В сравнение с английските съгласувателни принципи в рамките на сложното изречение, където действа правилото за съгласуване на времената, в българските сложни изречения се наблюдава т. нар. неутрализация на темпоралните опозиции. Според Ив. Куцаров (Куцаров 1998: 302) „възможността на немаркирания член на привативната опозиция в общото си значение

да се употребява във функциите на маркирания член“ се нарича неутрализация. Темпоралните неутрализации в български език задават „както съотнасянето към плана на неминалото или миналото (оформени спрямо акта на комуникация), така и отношенията на предходност – едновременност – следходност между действието в подчиненото изречение спрямо предиката в главното“ (Алексова 2008б: 14). „При логическото съгласуване глаголът в главното изречение получава своята темпорална оформеност чрез съотнасяне към акта на комуникацията, а от своя страна задава основната темпорална ориентация на събитията в подчиненото изречение, доколкото ги ориентира към плана на настоящето, на миналото или на бъдещето“ (Алексова 2008а: 174). „Тази неутрализация може да бъде обяснена с факта, че главното изречение в български език доминира не само в синтактично отношение в условията на субординационни отношения. Предикатът в него се отличава с темпорална доминация, която се разпростра върху подчинените изречения“ (Алексова 2008б: 16).

„Тази доминация на темпоралните признаци на предиката в главното изречение […] е необходимото условие за настъпването на неутрализация в подчиненото изречение“ (Алексова 2008б: 16). Иначе казано, ако в главното изречение има наличие на глагол в минало време, това е условие глаголите в подчинените изречения да бъдат заменени от неминалите си корелати (да бъдат неутрализирани по признака разновременност).

Посочвано е (Алексова 2008б: 16), че най-честите неутрализации настъпват в подчинени обстоятелствени изречения за време и условие. Също така неутрализации могат да настъпят по отношение на други темпорални признаци. „Ако е контекстуално изразено наличието на […] резултати от действието, то тогава в подчиненото изречение е възможна неутрализация на признака перфектност, а изразяването от контекста/ситуацията на следходност може да доведе до замяна на бъдещите времена от техните небъдещи корелати“ (Алексова 2008б: 16).

Необходимо е да се изясни, че за разлика от английски, където правилото за съгласуване на времената е задължително в рамките на сложното изречение, темпоралните неутрализации в български могат да бъдат както задължителни, така и факултативни. Задължителните неутрализации настъпват винаги, когато са налице определени контекстуално-ситуативни условия. Например неутрализацията по признака предходност в подчинени изречения е посочена като задължителна в подчинени изречения за време и условие (Герджиков 1984: 228). Незадължителните неутрализации

може да бъдат възможни в определен контекст, като ненастъпването им в този контекст може да се окаже натоварено с определен прагматичен, модален или друг стилистичен нюанс (Алексова 2008б: 18).

Тъй като в български език има 9 глаголни времена, а в английски те са 16 на брой, не може да се търси (пълно) формално съответствие между българските и английските времена. Различните значения на българските времена могат да бъдат преведени с форми на различни времена в английски.

Относно разположението на действието във времето, сравнено спрямо сегашния момент, деветте времена в български език включват едно сегашно време, четири минали, две бъдещи и две бъдещи времена в миналото, докато в английски език темпоралната система е симетрична и е изградена от 4 сегашни времена, 4 минали времена, 4 бъдещи времена и 4 бъдещи времена в миналото. Въпреки тези диспропорции, може да се твърди, че членовете на двете темпорални системи изграждат симетрични отношения помежду си. Четирите същински английски времена могат да изразяват четири различни вида на действието, като по този начин се създава симетричност и при значението, и при именуването на видовете времена. Така в английската темпорална система се образува отношението 4:4:4:4. В български език с оглед на опозицията неразновременност – следходност има 4 „сегашни“ времена, 4 „бъдещи“ и едно „същинско“ минало време. Опозицията резултативност – нерезултативност дава съотношение от 4 резултативни срещу 5 нерезултативни времена (Куцаров 2007: 249). Опозицията относителност – неотносителност дава 4 относителни срещу 4 неотносителни времена и едно време – ориентационен момент (Куцаров 2007: 255), което по същество е неотносително време.

Съществена разлика в особеностите на глаголите в двата езика е видът на глагола. Докато в български видът на глагола е лексикално-граматична категория, която е постоянна за дадената лексема и не се изменя при различните форми на времената, то английските глаголи не притежават еквивалентна категория (Валейка и Буиткийн 2003: 84). Така видовите двойки глаголи в български език имат един лексикален еквивалент в английски. Например: *сядам/седна – sit; видя/виждам – see.* Морфологичната категория вид на действието в български може да изразява само завършеността (резултативността) на действието спрямо определен ориентационен момент, докато в английски категорията вид на действието може да изразява и продължителност на действието спрямо ориентационния момент. Така има разминаване

при изразяването на значението за продължителност/незавършеност на действието – в български това значение е присъщо на самата лексема, докато в английски това значение се носи от формите на продължителните времена и продължителните перфектни времена. Тези две близки по характеристики значения се различават не само по отношение на категориите, които ги изразяват, но и по самата си същност. Както беше казано по-горе, видът на глагола в български определя действието като завършено или не спрямо целостта на самото действие, докато формите н продължителните времена в английски определят действието като продължително или не спрямо външен за действието деиктичен момент. В определени случаи значението на несвършения вид на глаголите в български може да се пренесе със съответното продължително време в английски. Например:

(26) Ще пиша писмото утре. – I will be writing the letter tomorrow.

(27) Ще напиша книгата догодина. – I will write the book next year.

Но това не може да се отнесе до всички случаи. Например в примери (28) и (29) лексикално-граматичното значение на свършения глагол се неутрализира от резултативността на глаголната форма:

(28) Чел съм романа. – I have read the novel.

(29) Прочел съм романа. – I have read the novel.

Представените прилики и разлики между темпоралните системи в български и английски и значенията на категориите, които ги изграждат, не предоставят достатъчно средства за еднозначно формално съотнасяне на значенията на отделните глаголни времена.

С оглед на целите на дисертацията, за да бъде създадена хипотеза за преноса на граматичните значения при превод на глаголните форми от български на английски, може да се подходи към съпоставянето на характеристиките на темпоралните системи в двата езика по два начина – ще бъдат представени формалните съответствия между значенията на темпоралните форми на семантично ниво и между темпоралните форми на морфологично ниво.

Първият подход, който може да се приложи, за да бъдат представени формалните съответствия между значенията на отделните времена, е съпоставката им на семантично ниво според представените семантични конструкции по модела на Х. Райхенбах. Съответствията са изведени на базата на основната употреба на конкретното време,

която се определя от позициите на момента на говорене (T), момента, в който има наличен резултат от действие (T_{res}), интервалите/моментите на референция ($R/R_1/R_2$) и интервала/момента на действието (E) върху темпоралната ос. Сравнения между българските и английските глаголни времена са представени като сходства на отношения между представените по-горе градивни елементи. Сравненията имат за цел да представят основната употреба на времената. Възможни транспозиции, както и модалните употреби, не се разглеждат поради факта, че единствено формата е обект на изследване, а тя се запазва в такива случаи. Също така формалният анализ ще бъде ограничен до представяне на отношения между елементите T, T_{res}, R и E. Изчерпателното описание на всички възможни транспозиционни значения и формалните съвпадения между значенията им е извън целите на настоящото изследване. Отношенията между момента на говорене, момента/интервала на действието и момента/интервала на референтност могат да бъдат представени така:

– отношение на съвпадение между момента на говорене, интервала на действието и интервала на референтност (E=R=T) в двата езика се изразява от семантиката на формите за сегашно време в български език и формите за сегашно просто и сегашно продължително време в английския език. Основната разлика е в рамките на интервала на действието спрямо интервала на референтност. В български език при значението на сегашно време интервалът на действието е включен в интервала на референтност, докато при значението на сегашно просто време в английски има пълно съвпадение между интервала на действието и интервала на референтност;

– отношение на съвпадение на интервала на действието с интервала на референтност и отношение на предхождане на интервала на референтност спрямо момента на говорене (E=R<T) изразяват формите на минало свършено и минало несвършено време в български език и формите на минало просто и минало продължително време в английски език. При значението на минало несвършено време в български интервалът на действието включва в себе си миналия интервал на референтност, а при минало свършено време – интервалът на действието съвпада с минал интервал на референтност. При значението на минало просто време в английски интервалът на действието съвпада изцяло с минал интервал на референтност, докато при минало продължително време моментът на референтност е преди момента на говоренето, а моментът на действието преминава през него;

– отношение на съвпадение между интервалите на действие и референтност и отношение на следхождане на интервала на референтност спрямо момента на говорене (E=R>T) изразяват бъдеще време в български език и бъдеще просто и бъдеще продължително време в английски. При бъдеще продължително време интервалът на действието преминава през интервала на референтност като липсва информация за прекратяване на действието в рамките на интервала на референтност.

Отношенията между момента на говорене, момента/интервала на действието, момента/интервала на референтност и момента, в който е наличен резултат от действието, могат да бъдат обобщени по следния начин:

– отношение на съвпадение между интервала на референтност, момента на говорене и момента на наличие на резултат от действие и отношение на предхождане на интервала на действие ($E<R=T=T_{res}$) в български език изразяват формите на минало неопределено време, а в английски – формите на сегашно перфектно и сегашно перфектно продължително време. При сегашно перфектно продължително време има наличие на информация за предхождане на началото на интервала на действие спрямо интервала на референтност, но липсва информация за край;

– отношение на предходност на интервала на действието спрямо интервала на референтност и отношение на предходност на интервала на референтност спрямо момента на говорене и на съвпадение спрямо момента на наличие на резултат от действие ($E<R=T_{res}<T$) в български език изразява семантиката на минало предварително време, а в английски – семантиката на минало перфектно и минало перфектно продължително време. При минало перфектно продължително време има наличие на информация за предхождане на началото на интервала на действие спрямо интервала на референтност, но липсва информация за край;

– отношение на предходност на интервала на действието спрямо интервала на референтност и отношение на следходност на интервала на референтност спрямо момента на говорене и на съвпадение спрямо момента на наличие на резултат от действие ($E<R=T_{res}>T$) в български език изразява семантиката на бъдеще предварително време, а в английски – семантиката на бъдеще перфектно и бъдеще перфектно продължително време. При бъдеще перфектно

продължително време има наличие на информация за предхождане на началото на интервала на действие спрямо интервала на референтност, но липсва информация за край.

Отношенията между представените по-горе елементи – момента на говорене, момента/интервала на действието, момента/интервала на референтност, момента, в който е наличен резултат от действието, и втори интервал на референтност – в двата езика могат да бъдат представени така:

– отношение на следходност на интервала на действие спрямо първи интервал на референтност, който се намира в отношение на предходност спрямо втори интервал на референтност, който от своя страна е в отношение на предходност спрямо момента на говорене ($R>E>R_1>T$), изразяват значенията на формите за бъдеще време в миналото в български език и бъдеще просто време в миналото и бъдеще продължително време в миналото в английски език. При бъдеще продължително време в миналото има наличие на информация за предхождане на началото на интервала на действие спрямо втория интервал на референтност, но липсва информация за край;

– отношение на следходност на интервала на действие спрямо първи интервал на референтност, който се намира в отношение на предходност спрямо втори интервал на референтност, който от своя страна е в отношение на предходност спрямо момента на говорене и в отношение на съвпадение с момента на наличие на резултат от действие ($R>E>R_1=T_{res}>T$), в български език изразяват формите на бъдеще предварително време в миналото, а в английски – формите на бъдеще перфектно време в миналото и бъдеще перфектно продължително време в миналото. При бъдеще перфектно продължително време в миналото има наличие на информация за предхождане на началото на интервала на действие спрямо втория интервал на референтност, но липсва информация за край.

Така описани, семантичните отношения по модела на Х. Райхенбах са представени в Таблица 3.:

Български	Английски
Сегашно време $T \in R; E \supseteq R$ При сегашно време в български език интервалът на действието включва в себе си интервала на референтност, който пък включва в себе си момента на говоренето.	Сегашно просто време $E=R=T$ При изграждането на семантиката на това време моментът на говорене, моментът на референтност и моментът на действието съвпадат напълно, без да има информация за

Български	Английски
	завършеност или повторяемост на действието. Сегашно продължително време $T = R; E \supseteq R$ Семантиката на сегашно продължително време в английски език изразява действие, чийто момент обхваща момента на референтност, който съвпада с момента на говорене.
Минало свършено време $R < T; R \supseteq E$ Семантиката на миналото свършено време изразява действие, чийто интервал е едновременен с минал интервал на референтност и е затворен в рамките на този интервал.	Минало просто време $E = R < T$ Семантиката на минало просто време изразява момента на действието, който съвпада с минал момент на референтност, който е преди момента на говорене..
Минало несвършено време $R < T; E \supset R$ Минало несвършено време в български означава действие, чийто интервал включва в себе си миналия интервал на референтност.	Минало продължително време $R < T; E \supset R$ При минало продължително време в английски език моментът на действието преминава през момента на референтност, който е в миналото.
Бъдеще време $T < R; R \supseteq E$ Значението на бъдеще време в български език означава действие, което се извършва в бъдещ интервал на референтност, разположен след момента на говорене	Бъдеще просто време $T < R = E$ Семантиката на бъдеще просто време е изградена от момент на референтност, в който се съдържа интервалът на действието и който е след момента на говорене. Бъдеще продължително време $T < R; R \supset E$ При бъдеще продължително време в английски език моментът на референтност е следхождащ по отношение на момента на говоренето, а моментът на действието е включен в него.
Минало неопределено време $T \in R; E < R; T_{res} \in R$ При минало неопределено време се изразява значение, което представя действие, чийто интервал е в плана на миналото, предходен е спрямо интервала на референтност, който включва в себе си момента на говорене.	Сегашно перфектно време $E < T = R; T_{res} \in R$ Сегашно перфектно време изразява действие в миналото, което е преди момента на говорене. Сегашно перфектно време е резултативно време, като резултатът е наличен към момента на говоренето, който съвпада с момента на референтност. Сегашно перфектно продължително време $T \in R; E < R; T_{res} \in R$ Семантиката на сегашно перфектно продължително време в английски език се

Български	Английски
	изгражда върху момент на действие, който е преди моментите на референтност и говорене, които съвпадат, но резултатът от действието е наличен в рамките на момента на референтност, без информация за приключване на действието.
Минало предварително време $R < T; E < R; T_{res} \in R$ Семантиката на минало предварително време в български език означава действие, което се извършва в плана на миналото преди даден минал интервал на референтността.	Минало перфектно време $E < R < T; T_{res} \in R$ Семантиката на минало перфектно време в английски език се изразява чрез момента на действието, който е предхождащ спрямо момент на референтност в миналото, но резултатът от действието е наличен в момента на референтност. Минало перфектно продължително време $R \supseteq E; R<T; T_{res} \in R$ Минало перфектно продължително време изразява действие, чийто момент е преди минал интервал на референтност, а резултатът от действието е наличен в интервала на референтност. При глаголните форми на това време липсва информация за прекратяване на интервала на действието в рамките на интервала на референтност или преди него.
Бъдеще предварително време $T < E < R; T_{res} \in R$ Бъдеще предварително време означава действие, което се е извършило преди бъдещ интервал на референтност, в който ще бъде налице резултатът от действието.	Бъдеще перфектно време $T < E < R; T_{res} \in R$ При бъдеще перфектно време моментът на действието е след момента на говорене, а резултатът от действието е наличен в един бъдещ интервал на референтност, който е след момента на говорене и интервала на действието. Бъдеще перфектно продължително време $R \supseteq E ; T < E; T_{res} \in R$ Семантиката на бъдеще перфектно продължително време се изгражда от момент на референтност, който е бъдещ спрямо момента на говоренето, и резултат от действие, който е наличен в момента на референтност. Моментът на действието е бъдещ спрямо момента на говорене.
Бъдеще време в миналото $R < E < T$ Бъдеще време в миналото означава действие, което е предстоящо по отношение на минал интервал на референтност.	Бъдеще просто време в миналото $R < E < T$ При бъдеще просто време в миналото моментът на действието е следходен спрямо един минал момент на референтност.

Български	Английски
	Бъдеще продължително време в миналото $T > R > R_1$; $E \supset R_1$ При бъдеще продължително време в миналото в английския език и двата момента на референтност са преди момента на говоренето, единият момент предхожда другия, а моментът на действието е бъдещ спрямо първия интервал на референтност и достига до втория – началото на действието е преди втория интервал на референтност и липсва информация за прекратяване на действието в рамките му.
Бъдеще предварително време в миналото $R < E < R_1 < T$; $t_{res} \in R_1$ При формирането на значението на бъдеще предварително време в миналото се използват два интервала на референтност. Единият от тези интервали R е в плана на миналото и действието е предстоящо по отношение на него. Тъй като бъдеще предварително време в миналото е резултативно време, резултатът от действието е наличен във втори интервал на референтност R_1.	Бъдеще перфектно време в миналото $R < E < R_1 < T$; $t_{res} \in R_1$ Семантиката на бъдеще перфектно време в миналото се изгражда от два последователни момента на референтност, предхождащи момента на говорене. Моментът на действието се намира между двата момента на референтност, като спрямо първия е бъдещ, а резултатът от действието е наличен във втория момент на референтност. Бъдеще перфектно продължително време в миналото $R < E < R_1 < T$; $E \supset R_1$ $t_{res} \in R_1$ Значението на бъдеще перфектно продължително време в миналото се изгражда от два последователни момента на референтност, минали спрямо момента на говорене. Моментът на действието е бъдещ спрямо първия момент на референтност, като достига и/или обхваща изцяло втория, като резултатът от действието е наличен във втория момент на референтност и в него липсва информация за прекратяване на действието.

Таблица 3. ***Отношение между семантиката на времената в български и английски.***

Вторият подход, който може да се приложи, е формално съпоставяне на граматичните значения на категориите време, вид на действието и таксис в български и време и вид на действието в английски. На първо място е нужно да се отбележи, че тъй като не е възможно на формално ниво да се съотнесе с пълна еквивалентност значението „вид на глагола“ в български към „продължително време“ в английски, а е необходим и лексикален контекст, за целите на настоящата работа е прието, че определени глаголни форми, в зависимост от лексикално-граматичната категория вид на глагола в български

могат да се превеждат с две времена в английски (продължителното и съответното просто време или перфектното време и съответното перфектно продължително време).

От гледна точка на граматичните значения на категориите, които изграждат темпоралните отношения в български и английски могат да бъдат направени следните заключения:

- глаголни форми, които изразяват едновременност или липса на информация за разновременност на действието с момента на говорене, ненатоварени с други значения, притежават и двата езика. За български това са глаголните форми за сегашно време, за английски – формите за сегашно просто и сегашно продължително време (носещи информация за продължителност на действието);
- глаголни форми, които изразяват информация за следходност на действието спрямо момента на говорене, ненатоварени с други значения, притежават и двата езика. За български това са формите за бъдеще време, за английски – формите за бъдеще просто и бъдеще продължително време (носещи информация за продължителност на действието);
- глаголни форми, които изразяват информация за следходност на резултат от действие спрямо момента на говорене, притежават и двата езика. За български това са формите на бъдеще предварително време, за английски – формите на бъдеще перфектно и бъдеще перфектно продължително време (носещи информация за продължителност на действието);
- глаголни форми, които изразяват предходност на действие спрямо момента на говорене, притежават и двата езика. В български това са формите на минало свършено време, в английски – формите на минало просто време.
- глаголни форми, които изразяват едновременност на резултат от действие с момента на говорене, притежават и двата езика. В български това са формите на минало неопределено време, в английски – формите на сегашно перфектно и сегашно перфектно продължително време;
- глаголни форми, които изразяват наличие на резултат от действие спрямо минал момент, притежават и двата езика. В български това са формите на минало предварително време, в английски – формите за минало перфектно и минало перфектно продължително време;
- глаголни форми, които изразяват следходност на действие спрямо минал ориентационен момент, притежават и двата езика. В български това са формите

за бъдеще време в миналото, в английски – формите за бъдеще просто време в миналото и бъдеще продължително време в миналото;

- глаголни форми, които изразяват следходност на резултат от действие спрямо минал ориентационен момент, притежават и двата езика. В български това са формите за бъдеще предварително време в миналото, в английски – формите за бъдеще перфектно време в миналото и бъдеще перфектно продължително време в миналото.

Представени по този начин, отношенията между двете темпорални системи са отразени в Таблица 4.:

Отношение на:	Предходност		Едновременност		Следходност		Спрямо
Действие	Минало свършено време	Минало просто време *Минало продължително време*	Сегашно време	Сегашно просто време Сегашно продължително време	Бъдеще време	Бъдеще просто време Бъдеще продължително време	Момент на говорене
Резултат от действие		*Минало перфектно време* *Минало перфектно продължително време*	Минало неопределено време	Сегашно перфектно време Сегашно перфектно продължително време	Бъдеще предварително време	Бъдеще перфектно време Бъдеще перфектно продължително време	
Действие			*Минало несвършено време*		Бъдеще време в миналото	Бъдеще просто време в миналото Бъдеще продължително време в миналото	Минал момент
Резултат от действие			*Минало предварително време*		Бъдеще предварително време в миналото	Бъдеще перфектно време в миналото Бъдеще перфектно продължително време в миналото	
	Български	Английски	Български	Английски	Български	Английски	

Таблица 4. ***Отношения между значенията на членовете на темпоралните системи в български и английски.***

Поради липсата на категория, която изразява значение на предходност на действие или резултат от действие спрямо минал ориентационен момент, в български за целите на съпоставителния анализ и на базата на семантичните съответствия, представени в Таблица 3., е прието, че предходността на действие или резултат от действие спрямо минал ориентационен момент, изразявана в английския език от формите на минало перфектно време и минало перфектно продължително време, съответства на ориентацията на действие или резултат от действие към минал момент, изразявана от формите на минало предварително време в български. Формите, които

изразяват съответните отношения в български и английски, са изобразени в курсив в Таблица 4.

Прилагайки отношенията, представени в Таблица 4., се получава следното съотношение между глаголните форми за лице и число в български и английски, представено в Таблица 5.

Български език	**Английски език**	Брой форми
Сегашно време	**Сегашно просто/Сегашно продължително**	
1. лице, ед. число; 2. лице, ед. число; 3. лице, ед. число; 1. лице, мн. число; 2. лице, мн. число; 3. лице, мн. число;	3. лице, ед. число; 1/2. лице, ед. число/1/2/3 лице, мн.число/ 1. и 3. лице, ед. число; 2 лице, ед. число/1/2/3 лице, мн.число	6:2/6:3
Минало свършено време	**Минало просто време**	
1. лице, ед. число; 2/3. лице, ед. число; 1. лице, мн. число; 2. лице, мн. число; 3. лице, мн. число;	1 форма за лице и число	5:1
Минало несвършено време	**Минало продължително време**	
1. лице, ед. число; 2/3. лице, ед. число; 1. лице, мн. число; 2. лице, мн. число; 3. лице, мн. число;	1. и 3. лице, ед. число; 2 лице, ед. число/1/2/3 лице, мн. число	5:2
Минало неопределено време	**Сегашно перфектно време/Сегашно перфектно продължително време**	
1. лице, ед. число, м.р.; 1. лице, ед. число, ж.р.; 1. лице, ед. число, ср.р.; 2. лице, ед. число, м.р.; 2. лице, ед. число, ж.р.; 2. лице, ед. число, ср.р.; 3. лице, ед. число, м.р.; 3. лице, ед. число, ж.р.; 3. лице, ед. число, ср.р.; 1. лице, мн. число; 2. лице, мн. число; 3. лице, мн. число;	3. лице, ед. число; 1/2 лице, ед. число/1/2/3 лице, мн. число	12:2
Минало предварително време	**Минало перфектно/Минало перфектно продължително**	
1. лице, ед. число, м.р.; 1. лице, ед. число, ж.р.; 1. лице, ед. число, ср.р.; 2/3. лице, ед. число, м.р.; 2/3. лице, ед. число, ж.р.; 2/3. лице, ед. число, ср.р.; 1. лице, мн. число; 2. лице, мн. число; 3. лице, мн. число;	1 форма за лице и число	9:1
Бъдеще време	**Бъдеще просто време/Бъдеще продължително време**	
1. лице, ед. число; 2. лице, ед. число; 3. лице, ед. число; 1. лице, мн. число; 2. лице, мн. число; 3. лице, мн. число;	1 форма за лице и число	6:1
Бъдеще предварително време	**Бъдеще перфектно време/Бъдеще перфектно продължително време**	
1. лице, ед. число, м.р.; 1. лице, ед. число, ж.р.; 1. лице, ед. число, ср.р.; 2. лице, ед. число, м.р.; 2. лице, ед. число, ж.р.; 2. лице, ед. число, ср.р.; 3. лице, ед. число, м.р.; 3. лице, ед. число, ж.р.; 3. лице, ед. число, ср.р.; 1. лице, мн. число; 2. лице, мн. число; 3. лице, мн. число;	1 форма за лице и число	12:1
Бъдеще време в миналото	**Бъдеще просто време в миналото**	
1. лице, ед. число; 2. лице, ед. число; 3. лице, ед. число; 1. лице, мн. число; 2. лице, мн. число; 3. лице, мн. число;	1 форма за лице и число	6:1
Бъдеще предварително време в миналото	**Бъдеще перфектно време в миналото/Бъдеще перфектно продължително време в миналото**	
1. лице, ед. число, м.р.; 1. лице, ед. число, ж.р.; 1. лице, ед. число, ср.р.; 2. лице, ед. число, м.р.; 2. лице, ед. число, ж.р.; 2. лице, ед. число, ср.р.; 3. лице, ед. число, м.р.; 3. лице, ед. число, ж.р.; 3. лице, ед. число, ср.р.; 1. лице, мн. число; 2. лице, мн. число; 3. лице, мн. число;	1 форма за лице и число	12:1

Таблица 5. ***Отношение между формите, изразяващи граматичната категория време в български и английски.***

Представените отношения между съответните темпорални системи в български и английски са изградени изцяло върху морфологичните категории, чиито значения изграждат съответните темпорални форми. За да има основание да се твърди, че определено време в български език отговаря на определено време в английски език, и при превод да се прави избор между съответните корелативни времена, е нужно тълкуването на времената на базата на тяхната семантика.

Въпреки очевидната диспропорция между броя на времената в български и английски, на формално ниво се откриват редица сходства. На първо място, от представения сравнителен анализ може да се очаква, че граматикализираната информация, носена от категорията вид на действието в български език, се пренася в английски чрез глаголните форми, маркирани със значението завършеност (перфективност). И двете морфологични категории изразяват близки значения – на отношение на наличие на резултат от действие или завършено действие към определен ориентационен момент (Куцаров 2007: 249; Кабакчиев 2003: 145). Граматикализираната информация, носена от формите на резултативните времена в български език, може да се пренесе успешно при превод чрез формите на перфектните времена в английски. Изградените по-горе семантични модели за двата езика показват, че на семантично ниво отношенията между елементите, изграждащи значението на темпоралните форми, са сходни и може да се очаква максимален семантичен пренос при превод. От гледна точка на формообразуването информацията за наличие на резултативност в български език се носи от миналото свършено деятелно причастие, което влиза в състава на аналитичните глаголни форми, и по-точно от резултативната морфема –Л (Куцаров 2007: 251), което означава, че всички резултативни форми в български се аналитични. В английски със значение за завършеност (перфективност) на действието е натоварен спомагателният глагол have и формите му, които влизат в състава на аналитичните глаголни форми (Филипова 2002: 92). При превод на глаголните форми от български на английски се очаква граматикализираната информация за наличие на резултативност, носена при темпоралните форми в български от миналото свършено деятелно причастие, да се пренася в английски чрез формите на спомагателния глагол have в синтетичните глаголни форми.

Предвид поставените цели за тази глава от дисертацията и на базата на формалните съответствия между граматикализираните значения на глаголните форми в български и английски, които изведохме, ще бъде поставена хипотеза за поведението на информацията при превод от български на английски.

От представения съпоставителен анализ се вижда, че повечето български глаголни времена могат да имат две съответствия в английски – съответното просто време или съответното продължително време, както и съответното перфектно време или съответното продължително перфектно време. Както беше посочено, видът на глагола в български е лексикално-граматична категория и изразява завършеността/ незавършеността на действието спрямо границата на самия процес („вътрешна (не)завършеност“, цялостност), докато в английски видът на действието изразява завършеността на действието спрямо някаква друга граница по темпоралната ос (външна (не)завършеност, продължителност). Така има разминаване при изразяването на значението за продължителност/незавършеност на действието – в български това значение е присъщо на самата лексема, докато в английски това значение се носи от формите на продължителните времена и продължителните перфектни времена. Тези две значения се различават не само по отношението между категориите в езиците, които са техни носители, но и по своите характеристики. В български език значението на лексикално-граматичната категория вид се носи от съответната глаголна словоформа и се запазва във всичките ѝ (възможни) форми (Куцаров 2007: 551). В английски език морфологичната информацията за продължителност на действието се носи от морфемата –ing, която е част от една словоформа в рамките на аналитичната глаголна форма (Ранкова и Иванова 1980: 103). Тук е мястото да бъде представен един от основните въпроси, който трябва да бъде разрешен от извлечената информация от „Българско-английския паралелен корпус със съотнесени глаголни форми“: за поведението на лексикално-граматичната категория вид в български език при превод на съответните глаголните форми на английски – как тази информация се пренася при превод и може ли да се очаква някакъв вид съответствие между категорията вид на глагола в български и продължителните глаголни форми в английски.

Друг въпрос, на който ще се търси отговор чрез извлечената статистическа информация от „Българско-английския паралелен корпус със съотнесени глаголни форми“, е свързан с поведението на глаголните форми при превод на минало свършено и минало несвършено време. От представения по-горе анализ има основание да се очаква, че единственото „същинско минало време“ в български се превежда с формите на минало просто време в английски, а информацията за едновременност на действието спрямо минал ориентационен момент се предава чрез формите на минало продължително време. И тук в извлечената статистическа информация трябва да се търси отговор за поведението на глаголите от свършен и несвършен вид.

Също така в английски език действа правилото за съгласуване на времената и минало продължително време, минало перфектно време и минало перфектно продължително време обикновено изразяват вторични действия в подчинени изречения, които са фонови за действието, изразено от сказуемото в главното изречение. Близките способи за ориентация на действията, изразени чрез глаголните форми за различните времена в двата езика, предполагат близко семантично значение и еднотипно поведение при конструирането на сложни изречения. Тези особености дават основание да се очаква, че подобно на български, където минало предварително и минало несвършено време са ориентирани към минало свършено време или друг негов сигнализатор (Куцаров 2007: 258), в английски формите на минало предварително и минало несвършено време ще отговарят съответно на минало перфектно или минало перфектно продължително време и минало продължително време. Нужно е да се отбележи и фактът, че реализацията на българските времена в свързан текст е подчертано контекстуално зависима, особено в сложното изречение, за разлика от английския, където действа правилото за съгласуването на времената. Причините за това се крият в силното влияние на контекста (в тесен и широк смисъл) върху функционирането на граматичната категория време в българския език и възможността семантично по-натоварените варианти, които са и формално по-сложни, да бъдат заменени с по-ненатоварените, ако контекстът (лексикални показатели или други граматични форми) е изразил същия семантичен признак (или признаци), който съответната грамема носи (Алексова 2018: 39).

Характерна особеност на българския език за разлика от английски е, че при формите на глаголите има формални маркери за лице и число, а при по-голямата част от аналитичните форми на времената има формален маркер и за род. С тази информация е натоварено миналото свършено деятелно причастие при формите за минало неопределено време, минало предварително време, бъдеще предварително време и бъдеще предварително време в миналото. От това следва, че при превод ще има загуба на морфологично ниво на информацията за рода на глаголното лице. Тъй като българският глагол изразява по-голям брой граматични значения, то при превод на английски множество български форми ще имат едно съответствие (например една форма обслужва цялата парадигма за минало просто време в английски език), а това води до редукция на информацията, която българските форми кодират на морфологично равнище.

Отличителна черта на формите на българския глагол е, че те съдържат граматична информация, която позволява подлогът да остане имплицитен, докато в английски език подлогът задължително трябва да бъде изразен експлицитно чрез местоимение или номинална фраза. Това в дадена степен улеснява автоматичния превод на глаголните форми в посока от български към английски и го затруднява в обратната посока. При случай на имплицитно зададен подлог в български (*чета, четеш, четем* ...) изниква проблемът за дефинирането на рамката на глаголната форма в английски – дали трябва експлицитен подлог с лично местоимение (I read, you read, we read …) да бъде включен към глаголната форма. За целите на настоящата работа е прието мнението, че няма основание да се смята, че задължителният подлог е част от темпоралните глаголни форми в английски. Приема се, че тази особеност на английския език е част от синтактичните особености на езика, при които подложната позиция трябва да бъде запълнена с лексикален материал, и не може да се търси морфологично основание за причисляването на местоименията към рамките на глаголните форми.

От направените заключения по-горе ще бъде представена хипотеза за статистически обусловения анализ на информацията, извлечена от „Българско-английския паралелен корпус със съотнесени глаголни форми“. Заключенията са на базата на направеното теоретично описание на темпоралните системи на български и английски и характеристиките на глаголите. Хипотезата представлява рамка за дефинирането на целите и задачите при създаването на „Българско-английския паралелен корпус със съотнесени глаголни форми“ и съответно е предпоставка за процеса по подбор на ресурсите, предварителната обработка, анотацията и използваните софтуерни приложения за работата по Корпуса. Основният въпрос, на който ще се търси отговор в Четвърта глава „Анализ на граматичните паралели в „Българско-английския паралелен корпус със съотнесени глаголни форми“, е относно трансфера на граматикализираната информация от темпоралните глаголни форми – начините, по които тази информация се предава при превод, и степента на трансфер между двата езика. Конкретният въпрос е може ли да се очаква значенията на граматичните категории, изразяващи еднотипни значения в двата езика, да съответстват при превод и в каква степен.

Въз основа на направените сравнения могат да се формулират следните предпоставки, въпроси и очаквани резултати, представени в Таблица 6.

Предпоставка	Въпрос	Очаквани резултати
Наличие на граматични категории, изразяващи близки значения в български и английски – наличието на категориите вид на действието в български и вид на действието в английски (по-конкретно перфектните времена).	В каква степен си съответстват значенията на резултативните времена в български и перфектните времена в английски?	Резултативните форми да се превеждат с перфектни глаголни форми в английски.
Наличие на лексико-граматичната категория вид на глагола в български и морфологично изразен продължителен вид на действието в английски.	Какво е поведението при превод на значението на лексикално-граматичната категория вид на глагола в английски? Дали и в каква степен може да се търсят съответствия на значенията, изразявани от продължителните времена в английски и глаголите от несвършен вид в български.	Темпоралните форми на глаголите от свършен и несвършен вид да се превеждат с глаголни форми съответно ненатоварени със значение продължителност или натоварени със значение продължителност в английски.

Таблица 6. ***Въпроси, на които се търси отговор чрез статистически обусловен анализ на данните от „Българско-английския паралелен корпус със съотнесени глаголни форми“.***

III. Създаване на „Българско-английския паралелен корпус със съотнесени глаголни форми“

В тази глава ще бъде направен кратък преглед на развитието на концепциите на машинния превод като дисциплина и съвременните теории и практики при създаването на лингвистични ресурси за целите на статистическото езиково моделиране и статистическия машинен превод. Ще бъде представен и основният резултат от изследването – „Българско-английският паралелен корпус със съотнесени глаголни форми“.

Обект на разглеждане в третата глава на дисертацията е машинният превод, осъществяван по метода, който доминира съвременните изследвания в областта – статистическият машинен превод. Поради големия интерес сред научната общност, от една страна, технологията е добре документирана, а от друга – има разработени и се предлагат свободно редица програмни инструменти с отворен код за създаване на езикови ресурси и извличане на лингвистични данни от тях. Статистическият метод за машинен превод се появява късно в кратката история на научното поле на компютърната лингвистика и представлява една от най-значимите и популярни тенденции в тази научна област. Статистическото езиково моделиране предполага изучаването на езика в неговите естествени проявления чрез реалната употреба и анализ на езикови феномени извън рамките на традиционната граматика.

Задача в третата глава на дисертацията е да бъдат разгледани различните методи за машинен превод и да бъде мотивиран изборът на статистическия метод за целите на настоящото изследване. Като основна част за изграждането на статистически езиков модел и извличането на надеждни лингвистични данни се разглежда конструирането на качествени езикови ресурси, подборът, изборът и анализът на налични езикови данни. Ще бъде представен „Българско-английският паралелен корпус със съотнесени глаголни форми“: създаване, анотация и средство за извличане на статистически данни.

III.1. Машинен превод – цели и историческо развитие

За да бъде поставен машинният превод във функционален за целите на изследването контекст, е подходящо да се даде формално определение на превода като процес и да се изяснят някои негови особености. Едно от най-разпространените определения е, че преводът представлява комуникация между езиците

(Маккартър 1992: 1051), следователно преводът може да бъде определен като процес на пренасяне на различна по тип информация между езиците. Информацията, която думата кодира със своето номинативно значение, се пренася без значителни затруднения. Разбира се, тук трябва да се отбележи, че многозначността, която се проявява на различни равнища, представлява сериозна пречка при превод. Полисемията е свойство, което се проявява в различна степен, но все пак думата може да има краен брой значения, които могат да бъдат описани. Информацията за денотата, се означава в различните езици с различни езикови знаци и с различни по тип лексикални единици. Например в български за роднинските названия за по-голям брат или сестра съществува отделна лексикална единица (*батко* и *кака*), докато в английски се използват словосъчетания от две езикови единици (*older brother* и *older sister*).

Освен преносът на информацията, закодирана в лексикалното значение на думите, при превод се извършва и пренос на друг тип информация, която зависи от особеностите на определените езици. Информацията, която се съдържа в граматичното значение на думите, се превежда трудно, тъй като всеки език се различава със своите особености – първо, при инвентара на граматичните категории и техните функции и второ, по типа и обема на информацията, която се кодира чрез тях в дадения език. При превод преносът на информацията за граматичното значение на езиковия знак от изходния език може да бъде предадена в значителна степен в целевия език, ако и двата езика притежават еднакви или сходни граматични характеристики. Възможно е и голяма част от тази информация да бъде загубена или да изисква допълнителен езиков материал, за да бъде предадена в достатъчна степен. Например граматичното значение на *щял съм бил да съм пишел* не може да бъде предадено на английски без използването на допълнителни лексикални средства, а значението на *пиввам* (‘пия, отпивам от нещо малко, от време на време; сръбвам’) може да се прояви в най-близка степен в по-голям езиков контекст – *I'm itching a little bit*.

При преноса на информация от подобен тип се предполага, че преводачът добре познава в дълбочина различните равнища (не само граматичните, но и етимологията, и историята) на изходния и целевия език. При добрия превод в единство трябва да се пренасят формата и смисълът с минимални изменения и загуби. Следователно, работата на добрия преводач се изразява в пренасяне (*translatio*) на различна по вид информация от една езикова система в друга.

Машинният превод (Machine Translation или MT) е раздел от компютърната обработка на езика, който се занимава с автоматичното превеждане на текстове. Под термина машинен превод може да се разбира практическото приложение на компютрите за превод на текстове от един естествен език на друг (Хътчинс 1995: 431). През историческото развитие на тази научна област могат да бъдат посочени различни подходи при изграждането на абстрактни или действащи системи за машинен превод. Основната причина за появата на машинния превод като научно поле е практическото приложение: на първо място, няма човешка сфера, в която да няма нужда от комуникация на различни езици и обмен на информация от различно естество, чийто превод изисква значителни финансови и човешки ресурси, и на второ място, съвременната глобализация е неосъществима без насърчаването на международното сътрудничество чрез премахване на езиковите бариери, обмен на техническа, икономическа и друга информация. В тези насоки приложението на системите за машинен превод има естествен положителен резултат от гледна точка на обема и скоростта на обмена на информацията, както и от гледна точка на достъпността ѝ. Не на последно място трябва да се споменат и практическите търговски и икономически мотиви за използването на системи, осигуряващи достъп до различна по вид и обем информация.

Зараждането на интереса към системите за машинен превод датира от 40-те години на XX век и възниква още с появата на първите компютри (Наков 2009). Това е периодът непосредствено след края на Втората световна война, по време на която в САЩ, във връзка с необходимостта от декодиране на съдържанието на прихванатите немски съобщения, силно развитие получават теорията на информацията и криптографията. Създаденият в резултат мощен математическо-статистически апарат съвсем естествено се разглежда като средство за постигане на машинен превод. През 1954 г. в САЩ се създава първият руско-английски прототип на система за машинен превод, което подсилва очакванията за бърз напредък, включително и за други двойки езици. Изследванията в България също не закъсняват и през 1964 г. се създава специална група за машинен превод между руски и български език под ръководството на проф. Александър Людсканов в Института по математика на БАН (Наков 2009).

Машинният превод се появява като подразделение на лингвистиката и приложната математика в средата на XX век, но днес е преди всичко област на изкуствения интелект и абстрактното проучване и описание на естествения език, както и научно поле, занимаващо се с практическото приложение на технологиите и

съвременните информационни теории (Хътчинс 1995: 435). В съвременния свят терминът „машинен превод“ се отнася до компютризирани системи, предназначени за създаването на преводни текстове със или без човешка помощ (Боукър 2002: 87). Появяват се и системите за компютърно подпомогнат превод (Computer-Assisted Translation Tools), които на базата на предимно статистически принципи осигуряват интегрирана среда за превод чрез предоставяне на достъп до онлайн речници, преводна памет (Translation Memory), терминологични бази от данни (Terminology Base), автоматична проверка и корекция на текстове и др. (Каланци 2002: 34). Тук е мястото да се отбележи, че границите между двете основни подразделения на приложния машинен превод – компютърно подпомогнат превод, който се извършва от хора (Machine-Aided Human Translation (MAHT)), и машинен превод, подпомогнат от хора (Human-Aided Machine Translation (HAMT)), не са ясно дефинирани и терминът компютърно подпомогнат превод (CAT) може да обхване и двете, а като централно ядро на машинния превод може да се изведе процесът по автоматизацията на превода (Корнацки 2018: 20). Въпреки че идеалната цел на системите за машинен превод е да произвеждат висококачествен превод, на практика изходният материал (преводът) обикновено се подлага на човешка редакция. Редакцията на машинно генерирани преводи е разпространена практика, като при определени обстоятелства крайният резултат от системите за машинен превод може да не бъде редактиран, а само преработен, например, ако е предназначен за специалисти, които са запознати с текста. Целевият текст може да се използва и като работен вариант за последващ човешки превод, от който да бъдат извлечени термини, да бъде автоматично разпозната сферата на текста, текстовите особености и т.н. Работата на системите за машинен превод и качеството на изходния материал може да бъдат подобрени чрез разработване на повече езиково зависими методи, които да изпълняват функцията на специфични ограничители за алгоритмите, които изграждат ядрото на системите за автоматичен превод (Хътчинс 1995:434), или чрез налагане на определени изисквания към входния текст. Системите за автоматичен превод могат да бъдат проектирани например за работа с текстове от ограничена тематична област (нормативни актове, медицински документи и т.н.), при които входните езикови вариации са крайно ограничени и предвидими. Възможна предварителна стъпка и изискване към входните текстове е да бъдат написани на определен контролиран език, който ограничава обхвата на лексиката и избягва омонимията, полисемията и сложни изреченски структури – този метод се прилага при текстове, които са изградени от затворено множество от възможни конструкции –

изречения, фрази, устойчиви словосъчетания и т.н., предназначени например за описания, упътвания и инструкции на продукти (Боукър 2002: 85). Друга възможност е предварителното анотиране на входните текстове с лингвистични маркери от различен вид – за част на речта, за снемане на многозначност, за граници на фрази и изречения. Системите за автоматичен превод може да бъдат конструирани така, че да подпомагат преводача по време на процеса на превод чрез интерактивно предсказващо въвеждане, снемане на многозначност, локализиране на непреводими единици и т.н. (Родриго 2008: 16).

Съвременните системи за машинен превод обикновено са проектирани за работа с два конкретни езика (двуезикови системи), но съществуват и системи, които работят с повече от една двойка езици (многоезикови системи). Двуезиковите системи могат да бъдат проектирани както като еднопосочни, така и като двупосочни, практическото приложение на многоезиковите системи е доста по-голямо. Многоезиковите системи за машинен превод позволяват обработката на повече от една двойка езици, като използват унифициран модел за превод. Използването на многоезикови системи за превод е доказано като по-успешно в сравнение с двуезиковите системи (Донг и др. 2015: 1729). Използвайки методите на невронния машинен превод, многоезиковите системи се различават от традиционните двуезикови системи за превод. Невронните системи кодират променливите по дължина изречения от изходния език във фиксирани по дължина векторни репрезентации. Векторните репрезентации след това се декодират с друга невронна мрежа от фиксирани вектори към променливи по дължина целеви изречения (Донг и др. 2015: 1724). Успехът на тези системи зависи преди всичко от наличието на достатъчни по обем езикови корпуси.

Развитието на съвременните системи за машинен превод следва историческото развитие на теориите в тази научна област. Характеристиките на дизайна на системите за машинен превод може да бъдат разделени на три основни типа.

Първият (и исторически най-старият) тип е т. нар. подход на „директен превод“ (direct translation). При този подход системите за машинен превод са специално създадени за определена двойка и посока езици и са проектирани във всички детайли на езиково зависимо ниво (Хътчинс 1995: 433). При разработването на подобен тип системи е необходимо теоретичното съпоставително изследване и практическото имплементиране на междуезиковите зависимости в системата. Системите за машинен превод от типа „директен превод“ са наречени така, защото преводът се осъществява

директно от текста на изходния език към текста на целевия език. Основна теоретична предпоставка при този подход е схващането, че лексиката, морфологията и синтаксисът на изходния език не трябва да бъдат анализирани изолирано, а само в съпоставителен план с целевия език. Системите „за директен превод“ обикновено са изградени от ръчно написани правила за анализ, генериране и преобразуване на текстове и достатъчно голям двуезиков речник.

Вторият тип дизайн на системи за машинен превод исторически се заражда от системите „за директен превод“ и представлява своеобразно продължение на теоретичните постановки и практическите методи от предходния етап (Черагю 2012: 161). Такъв е подходът с „трансферни правила“, при който се осъществява последователно надграждане на различните равнища на изходния език, за да се представи междинно абстрактно работно състояние между изходния и целевия език. Системите от този тип са разработени основно за работа с конкретна двойка езици и една посока за превод. Работата на системите с трансферни правила може да бъде разделена на три работни етапа – анализ на изходния текст, трансфер и синтез на целевия текст. Първият работен етап превръща текстовете на изходния език в краен набор от абстрактни елементи, които отразяват спецификата на езика; вторият етап преобразува (трансферира) тези елементи от специфични за изходния език към еквивалентни и специфични за целевия език елементи; третият етап генерира (синтезира) целевия текст на базата на абстрактните елементи, които отразяват спецификата на целевия език. Системите за машинен превод, изградени на базата на трансферни правила, се състоят обикновено от три вида речници и граматика – речник за анализ на изходния език, съдържащ подробна морфологична, синтактична и семантична информация за езика, от който ще се превежда; трансферен двуезиков речник между двата езика; граматика, която съдържа трансферните правила между структурите на изходния и целевия език и речник за синтез на целевия език, който съдържа морфологична, синтактична и семантична информация за езика, на който ще се извършва преводът (Хътчинс 1995: 436). В рамките на етапите за анализ, трансфер и синтез много от системите с трансферни правила показват ясно разделени компоненти, предназначени за различни нива на езиково описание: морфология, синтаксис, семантика. Следователно, анализът, който тези системи извършват върху изходния език, може да бъде на различни езикови равнища – морфологичен анализ (токънизиране, приписване на маркер за част на речта), синтактичен анализ (идентификация на фразови структури, синтактични

зависимости и връзки), семантичен анализ (разрешаване на лексикална и структурна многозначност). Синтезът на целевия език може също така да премине през морфологичен синтез (генериране на правилни словоформи), синтактичен синтез (генериране на структурите на фрази и изречения) и семантичен синтез (подбор на подходящи и съвместими лексикални и структурни форми). В системите, базирани на трансферни правила, компонентът за трансфер между изходния и целевия език може да бъде изграден от отделни приложения, предназначени за лексикален трансфер (подбор на речникови еквиваленти) и за структурен трансфер (трансформация на единици от изходния език към единици от целевия език).

Като естествено развитие на системите, базирани на трансферни правила, се появяват и системите, изградени на основата на единен абстрактен и независим език – език посредник (интерлингва) (Джурафски и Мартин 2006: 1005). Системите, изградени на базата на език посредник, по същество са системи, базирани на трансферни правила, при които етапът на трансфер е интегриран в етапите на анализ и синтез между изходния и целевия език и езика посредник. Етапите на анализ на изходния език и синтез на целевия език се осъществяват съответно между двата работни езика и езика посредник. Езикът посредник представлява абстрактно логическо и езиково независимо представяне на езиковите единици и техните значения. Различните подходи за създаване на език посредник се основават на специално конструирани изкуствени метаезици (логически системи, изградени от семантични примитиви, които са общи за работните или за множество езици) или на универсален езиково независим речник. Като свидетелство за по-съвършения дизайн на системите, базирани на език посредник, се посочва сложността за осъществяване на превод между n на брой работни езици – при системите за „директен превод“ и системите, базирани на трансферни правила, се изискват n(n-1) на брой системи, докато при системи, базирани на език посредник, са необходими само 2n на брой системи за машинен превод (Наков 2009). От друга страна, сложността на езика посредник и представянето на езиковите реалии на абстрактен и езиково независим междинен език значително усложнява създаването на подобни системи. Използването на език посредник – както езиково зависимо, така и езиково независимо – засега остава извън възможностите на съвременните технологии за машинен превод, макар и да са правени някои опити в това отношение върху силно ограничени по обем данни. Досегашното развитие на машинния превод недвусмислено показва, че всяко успешно приближаване до език посредник води до по-добри резултати

и до малка „революция“ в разбирането за езика като цяло. За момента са добре разработени с всички техни предимства и недостатъци едва първите две нива на системите за машинен превод, базирани на трансферния подход – системите за директен превод и системите с трансферни правила при това не напълно (Хътчинс 1995: 441).

С натрупването на изчислителна мощ на компютрите и развитието на съвременните схващания за същността на езиковите проявления се появяват и теории за усъвършенстването на машинния превод чрез прилагане на нелингвистични подходи. И трите представени типа на системи за машинен превод имат за цел да постигнат машинен превод чрез детайлно анализиране на езиковите зависимости и естествените езикови взаимовръзки между отделните езикови равнища. Въпреки множеството преимущества на подобен подход, при него възможностите са ограничени, тъй като е необходимо предварителното познание и описание на лингвистичните феномени, за да е възможно изследването и приложението им. Статистическото езиково моделиране се съсредоточава над крайните продукти на езика, за да изследва естествените процеси в него и проявленията им (Коен 2010: 74). Като преимущества на този подход се посочват възможностите за изследване на езикови феномени – като hapax legomenon (рядка употреба) и правилото на Зипф (Зипф 1935) за честотата на срещане, чийто анализ е извън обхвата на традиционните системи за моделиране на езика и машинния превод. Статистическото моделиране на езика възниква в края на 80-те години на XX век, като основната цел е разпознаването на реч, но също така играе важна роля и в различни други приложения за обработка на естествения език като машинен превод, автоматично приписване на част от речта, интелигентен метод на въвеждане на текст и др. Научното поле е преминало през два периода в своето развитие – статистическо моделиране на езика, базирано на думи, и статистическо моделиране на езика, базирано на фрази. Основният принцип при статистическото моделиране на езика е „повечето данни са по-добри данни“ (more data is better data) (Хътчинс 1995: 438).

Статистическият машинен превод има за цел да представи вероятността за съответствие на две езикови единици в два или повече различни езика. Подобно на превода, извършван от хора, при който целевият текст е максимално близък по смисъл до текста на изходния език и максимално естествен за целевия език, при статистическия машинен превод се разчита на статистически модел, който представя чрез стойностна функция съответствието между два текста и принадлежността на целевия текст към определен език. В най-общ план тази функция може да бъде представена чрез

Формула 1., където T' е търсеното изречение на целевия език, S е изречение на изходния език, а T е изречение на целевия език:

$$T' = argmax_T \text{ съответствие}(S,T) \text{ принадлежност}(T)$$

Формула 1. ***(Джурафски и Мартин 2006: 1007)***

Иначе казано, функцията T' максимализира продукта на логическото и семантично съответствие на две изречения от изходния и целевия език (S и T) и граматичността (принадлежността) на T към целевия език. Може да се представи вероятността за съответствие на двете изречения чрез вероятностен модел, който дава вероятността изречение T да бъде превод на изречение S:

$$T' = _{argmax_T} P(T|S)$$

Формула 2. ***(Джурафски и Мартин 2006: 1007)***

Прилагайки правилото за условна вероятност на Бейс, се получава Формула 3.:

$$\mathrm{T}' = argmax_T \frac{P(S|T)P(T)}{P(S)}$$

Формула 3. ***(Джурафски и Мартин 2006: 1006)***

Поради факта, че функцията търси определено преводно изречение T' за фиксирано изречение на изходния език S, може да се игнорира знаменателят във Формула 3 и да се получи Формула 4.:

$$\mathrm{T}' = argmax_T P(S|T)P(T)$$

Формула 4. ***(Джурафски и Мартин 2006: 1007)***

Във Формула 4. могат да бъдат откроени трите основни компонента на една система за статистически превод, която търси най-добро преводно изречение T' за изходно изречение S. Първият компонент е преводният модел P(S|T) или вероятността две изречения S и T от различни езици да си съответстват на логическо и семантично ниво. Преводният модел P(S|T) следва да дава висока вероятност за двойката (S,T), ако S може да бъде превод на T, и ниска вероятност – в противен случай (Коен 2010: 104). Тук трябва да се отбележи, че преводният модел отчита единствено това дали е вероятно двете изречения да са превод едно на друго, без значение дали резултатното T е граматично правилно построено изречение от целевия език. Последното е задача на езиковия модел P(T) или вероятността дадено изречение да принадлежи към целевия език – да бъде граматично и разбираемо за съответния език (Джурафски и

Мартин 2006: 1008). Езиковият модел следва да дава висока вероятност на граматично правилни изречения за целевия език T и ниска вероятност на неправилни, както и на граматично правилни, но малко вероятни изречения (например в семантично отношение: „Безцветната зелена идея спи яростно" (Чомски 1957: 15)). В генеративната граматика понятието граматичност се отнася до съответствието на дадено изречение с правилата, определени от конкретната граматика на даден език. Граматичността е теоретичен термин и определя едно изречение като граматично, ако то се генерира от граматиката, и за неграматично, ако не се генерира от същата граматика (Нюмайър 1983: 8). Ако се представи определено изречение, което езиковият модел верифицира, като крайна поредица от думи от целевия език w_i, вероятността за изречение T от целевия език може да се представи като:

$$P(\mathrm{T}) = P(w_1 w_2 w_3 \ldots w_n)$$

Формула 5. ***(Джурафски и Мартин 2006: 98)***

Поради факта, че подобна вероятност би приписвала твърде малки стойности на изречения, които са граматични, но не се срещат в наличните езикови ресурси, може да се приложи верижното правило на Марков (Фрейзър 2008: 54) и да се опрости вероятността за даден низ от думи в рамките на контекста от предходните n думи – марковски модел от N-ти ред. Марковският модел от N-ти ред изчислява вероятността на дадена поредица от думи като произведение от условните вероятности на всяка дума в контекста на предходните n-1 думи. Във Формула 6. е представен обобщен марковски езиков модел от N-ти ред.

$$P(w_1^n) \approx \prod_{k=1}^{n} P(w_k | w_{k-1})$$

Формула 6. ***(Джурафски и Мартин 2006: 98)***

Разделянето на двете задачи в два отделни модела позволява, от една страна, езиковият модел да се опрости значително, а от друга страна, създава възможност двата модела да се обучават поотделно, включително на базата на различни езикови ресурси (Наков 2009). Разбира се, в тези случаи преводният модел изисква паралелни двуезикови текстове, които са по-трудни за извличане и създаване, докато езиковият модел използва едноезикови текстове, които могат да бъдат леснодостъпни в големи обеми (Наков 2009).

Третият компонент на системите за статистически машинен превод е декодерът, който по зададено изречение S търси най-вероятното изречение T (Джурафски и

Мартин 2006: 1009), т.е. това, което максимизира произведението P(S|T)P(T). За да се разреши целта, която седи в основата на тези системи, е необходимо да се подходи в обратен ред – чрез прилагане на метода на шумовия канал (Джурафски и Мартин 2006: 1008). При шумовия канал се приема, че дадено изречение на целевия език T е „замърсена“ версия на изречение от изходния език, преминало през канал с шум. Така целта на статистическия машинен превод може да бъде представена като търсене на изречение S от изходния език, което е породило изречение T от целевия (Джурафски и Мартин 2006: 1008).

Добрата взаимообвързаност на трите основни описани компонента на системите за статистически машинен превод и обучението на преводния модел и езиковия модел върху достатъчно представителни и обемни езикови ресурси гарантират качеството на работа на системите за статистически машинен превод и адекватността на резултатните преводи.

Основната цел за преводния модел като компонент на системите за статистически машинен превод е съотнасянето на единиците от различни езикови равнища между различните езици (Наков 2009). При представения по-горе модел на работа на системите за статистически машинен превод основната единица, с която беше илюстрирана работата им, е изречението. Преводните модели се обучават над достатъчно обемни двуезикови корпуси, в които въпреки количеството съдържащи се изречения вероятността да се срещне точно определено изречение S е твърде малка. Това налага сегментирането на езиковите единици на по-малки функционални фрагменти, чрез които се моделират по-прецизни и по-лесно приложими езикови системи (Джурафски и Мартин 2006: 1008). Един от възможните подходи, подобно на подходите, базирани на трансферни правила, е преводът дума по дума. Моделите на IBM (1, 2, 3, 4 и 5) представляват поредица от последователно нарастващи по сложност модели, предназначени за съотнасяне на езикови единици и обучение на преводни модели. Моделите на IBM интегрират различни възможни езикови проявления – от вероятност за директен лексикален превод до дублиране и изпадане на думи и пренареждане. Предимство на моделите на IBM е, че предлагат вероятностно съотнасяне на езиковите единици и проследяемост на процесите (Браун и др. 1993: 283). От друга страна, важно ограничение на моделите на IBM е, че могат да бъдат обучавани и да моделират само отношения от тип 1:M (едно към много), но не и M:1 (много към едно) или M:M (много към много), т.е. една дума от изходния език може да поражда 0, 1 или повече думи от

целевия, но една дума от целевия език може да се поражда само от една дума от изходния език (Наков 2009). Този подход създава трудности при превод от български на английски, тъй като се губи връзката между членната морфема в български и определителния член в английски. Друг проблем на моделите на IBM е невъзможността за отчитане на контекста. Неотчитането на контекста силно затруднява и съгласуването по род, число, членуване и др. и може да доведе до генериране на фрази, които не винаги могат да се коригират успешно от езиковия модел (Наков 2009). Като естествено продължение на системите за статистически машинен превод, базирани на метода за превод дума по дума, се появява и преводът на цели фрази (Коен и др. 2003). Методът за превод на цели фрази дава решение на повечето от проблемите, свързани с моделите на IBM. Този метод е близък до моделите на IBM, като при него работните единици не са токъни, а суперсегментни единици, образувани от тях – фразови модели (Наков 2009). Фразите се конструират на базата на последователни стъпки от възможни пермутации (Имамура и др. 2004: 100). С всяка от стъпките е асоциирана съответна вероятност, която отново се извлича от паралелен двуезичен корпус от двойки изречения на изходния и целевия език. Този метод води до съществено подобрение на качеството на машинния превод, тъй като при него е възможно представянето на съответствия от типа M:M (много към много) и дава възможност за по-добро отчитане на локалния контекст (Имамура и др. 2004: 102). Някои от типовете модели за превод на цели фрази включват и препинателни знаци, което позволява да се моделират и разлики в правилата за пунктуация между два езика.

През XXI век успешно се развива машинният превод, базиран на невронни мрежи. Невронните мрежи фундаментално се различават по начина, по който се извършва преводът в сравнение с традиционните статистически модели и моделите, базирани на правила. Невронните мрежи до голяма степен „копират“ устройството на биоелектричните мрежи в мозъка на човека, образувани от неврони и техните синапси. В невронните мрежи съществуват входни и изходни слоеве от неврони и взаимозависими слоеве скрити неврони. Математическият аналог на невронните мрежи са йерархично подредените изчислителни мрежи, при които всеки изчислителен елемент има тежест (значимост). Комуникацията между отделните неврони се осъществява на различни нива. При машинния превод, изграден на базата на невронни мрежи, входната информация представляват единиците от изходния език, които могат да бъдат думи, фрази или изречения. Основната структура на невронните мрежи за машинен превод се

състои от два повтарящи се слоя от неврони. Входният слой (анализаторът) анализира входните текстови единици с променлива дължина и ги кодира във вектор с фиксирана дължина. Втората невронна мрежа (декодерът) генерира целевите текстови единици на базата на вектора, генериран от първата мрежа (Фират и др. 2016: 866).

Както традиционните статистически системи за машинен превод, така и невронните системи разчитат преди всичко на достатъчно обемни и надеждни езикови корпуси, върху които да се обучават и оценяват. Съществено важна роля имат ръчно създадените и анотирани корпуси при дизайна на подобен вид системи. Ръчно създадените корпуси имат преимущество пред автоматично компилираните корпуси, защото данните в тях (които могат да бъдат автоматично събрани) са обработени от хора и вероятността за грешки е малка. Тази тяхна характеристика предполага по-голямата им надеждност за изучаване на езика и извличане на коректни данни, отколкото автоматичното събиране на текстове.

III.2. Езиковите корпуси като ресурс за машинен превод

При статистическия (а в по-ново време и при невронния) машинен превод се разчита на представителни паралелни езикови корпуси с достатъчен обем, които да представят езиковите феномени в техните различни проявления. За създаването на статистически езиков модел за превод е необходимо да се знае каква е вероятността дадена дума t в целевия език да бъде превод на дума s от изходния – p(t|s). За целта са необходими алгоритми и системи, които да обработват информацията, и езикови ресурси, от които да се извлекат лингвистичните данни. Създаването, структурирането и анотирането на езикови корпуси е основен предмет на изследване на корпусната лингвистика. Корпусната лингвистика е научно направление в лингвистиката, което се занимава с разработването и създаването на достатъчно представителни като обем езикови корпуси, съставени от реални езикови примери или текстове. В световен мащаб корпусната лингвистика има сериозни постижения и традиционно приложение основно при създаването на различни типове речници за определени лингвистични цели, както и в обучението по чужд език. В корпусната лингвистика се използват големи колекции от анотирани и неанотирани текстове, за да се извличат статистически закономерности за употребата на дадено лингвистично явление. Корпусната лингвистика може да бъде в помощ и при ефективното използване на различен тип корпусни данни при формулирането на правописни и правоговорни правила, различен тип граматики и т.н.

Корпусно-базираният подход може да бъде полезен както при изследването на множество различни аспекти и области на лингвистичното познание на различни езикови нива (фонетика, морфология, синтаксис и т. н.), така и при различни практически и теоретични цели и задачи (Тишева 2014). Корпусният подход е в помощ на езиковедите при изследването на голям обем езиков материал, съставен от различни текстове при изучаването и описанието на естествения език и формулирането на теоретични постановки относно функционирането на езика. Съвременните корпуси се използват като емпиричен езиков материал, служещ като база за обективността на съответните изследвания и изводи, които не разчитат само на индивидуалния езиков усет на езиковеда.

Според определението, дадено от Европейската асоциация на езиковите ресурси, с термина езиков ресурс се означава системата от езикови данни или описанията на такива данни, които са достъпни за машинна обработка с различна цел. „Езиковите ресурси варират от обикновени списъци до сложни ресурси, в които са организирани различни типове езикова информация: списъци от думи, колекция от текстове, корпуси, електронни речници, тезауруси, онтологии, тематични речници, бази от данни, звукови записи, видеоматериали и др.“ (Коева 2009: 59 – 60). В най-общ план езиковите ресурси може да се класифицират според структурата, в която са представени езиковите данни, или според конкретното съдържание на включените данни. От гледна точка на начина, по който са структурирани, езиковите ресурси могат да представляват списъци от думи, корпуси, електронни речници, тезауруси, тематични речници или бази от данни (Тишева 2014). Ресурсите може да съдържат само писмени текстове или да имат комплексно съдържание, ако включват и изображения, аудио- или видеоматериали. При по-детайлна класификация на езиковите ресурси се отчитат форматът, езикът и тематичната област на включените данни. Според формата, в която са представени данните, се определят три вида ресурси: речеви – включват записи на диалектна или разговорна реч, радио- или телевизионни предавания; писмени и многомодални (мултимодални, мултимедийни) (Тишева 2014). Според представените езици ресурсите се делят на едноезикови, двуезикови или многоезикови, а според тематичната област, с която са свързани езиковите данни – на общоезикови и специализирани.

Съществува известна неяснота при определянето на относително сходни понятия като текстов архив, библиотека от текстове, съвкупност от текстове и корпуси (Коева 2009: 60). Корпусите се отличават по отношение на своето структуриране и

използване. Те представляват колекция от текстове в електронна форма, подбрани чрез външни критерии, така че да представят, доколкото е възможно, даден език или негова разновидност и да предоставят данни за лингвистични изследвания (Синклер 2005 – цит. по Коева 2014: 29). Репрезентативността е признак, определящ обективността на корпусите. При ексцерпирането им се избират максимални по обхват и разнообразие материали, с цел да се даде възможност за най-пълно описание на обективната картина на речевата практика. Именно тази характеристика различава лингвистичния корпус от случайния набор от текстове (Баева 2011: 49). Предварителната обработка на текста обхваща различни нива от заложената в текста езикова информация: токънизиране (разделяне на текста на последователно от символи), анализиране (разделяне на текста на определени единици – фонеми, морфеми, лексеми, словосъчетания, изречения и др.); анотиране (приписване на лингвистична информация за характеристики на всяка отделна единица – морфологични, синтактични, морфосинтактични, семантични и др. признаци); разрешаване на различни езиково специфични явления като местоименни и неместоименни анафори, елипси и др. и на различните типове езикова многозначност. Различните нива на анализ използват различни програми: морфологични анализатори (тагери), синтактични анализатори (парсери), разширители на анафори, анализатори на текстовата структура и др. От съвременно гледище теоретичното определение за езиков корпус е по-скоро описателно, отколкото дефинитивно. Като основни критерии, за да бъде определен даден набор от текстове като корпус, обикновено се посочват компютърно достъпната, електронна форма, подчинена на даден формат (например txt, XML); целенасочената и добре документирана структурна организация; представителността за езиковата вариативност за предварително определени, конкретни езикови явления; балансираността при разпределението на текстовете в общата структура спрямо употребата в изследваната езикова общност или времеви отрязък. В определенията за езиков корпус „обикновено се подчертават критериите при подбора на корпусните единици, значението на корпусите за различни лингвистични изследвания, както и възможността, която корпусите предоставят за компютърна обработка“ (Коева 2014: 29). Анализът и обработката на различните типове корпуси са обект на компютърната обработка на естествения език. Корпусът представлява множество от текстове или фрагменти с езиков материал в писмен или речев формат, главно в електронна форма, организирани по систематичен и структуриран начин, така че да могат да бъдат използвани за изследването на различни езикови употреби. Определението, възприето в този текст, е „колекция от езикови примери, които са

избрани и подредени според експлицитни лингвистични критерии, за да се използват като модел на езика“ (ЕГПКОЕ 1996: 4). Текстовите корпуси обикновено се използват за статистически анализ, за проверка на честотата на определени употреби или лингвистични правила, специфични за дадена езикова или литературна сфера, използват се и за изучаване на исторически документи, антични ръкописи и др.

Характеристики и критерии при определянето на езикови корпуси, подобно на дефинирането на определение за корпус, също са обект на дискусии. Характеристиките на даден езиков корпус могат да бъдат неговата големина, текстова типология, реализацията на езиковите феномени (общи или специализирани), вътрешна структура, анотационна структура и др. Трябва да бъде направено уточнението, че и тези характеристики нямат единно тълкуване и принципи за дефиниране, използването им е продиктувано по-скоро от стремеж към концептуализирано прилагане на различен набор от методи при изграждането на езикови корпуси.

В този текст ще бъде направен опит за характеризиране на корпусите чрез традиционно възприетите характеристики, които до голяма степен представляват обединение на по-широкия спектър от характеристики за езиков корпус – балансираност, представителност и илюстративност (Лийч 1991: 10). Представителността на даден езиков корпус се определя със степента, в която той може да обхваща пълния диапазон от проявления на даден езиков феномен. Смята се, че корпусът е представителен, ако констатациите за езика въз основа на неговото съдържание могат да бъдат възприети като обобщения, характеризиращи този език (Лийч 1991: 27). Стремежът е всеки езиков корпус да осигурява представителна извадка за изследвания езиков феномен. Подборът на представителни текстове е основна част от съставянето на даден корпус от езикови данни. В този смисъл представителността на повечето корпуси до голяма степен се определя от два фактора: обхват на жанрове, включени в корпуса (балансираност), и критерии за подбор на текстове от всеки жанр (илюстративност) (Синклер 2005: 12). „Представителността на даден корпус се свързва с възможността му да предоставя достоверни езикови факти за определено състояние на езика, специализирана област на употреба на езика и/или множество от езикови явления. [...] Представителността се определя от разнообразните видове текстове, които илюстрират дадено състояние на езика или негова специализирана област на употреба, по-точно зависи от дефинирането на множеството от представителни текстове, йерархичната организация на техните типове и съотношението между тях (балансираност) и техниките, които се използват, за

да се извлекат представителни извадки (илюстративност)“ (Коева 2014: 30). Балансираността на корпуса зависи от текстовата типология – класификацията и характеристиките на включените текстови типове. Балансираният корпус трябва да покрива в някаква пропорция колкото е възможно повече типове текстове, за да бъде представителен за езика като цяло или за даден езиков феномен (Байбър 1993: 243). Известни са различни начини за постигане на балансираност, основани както на външни, така и на вътрешни критерии. Критерии за подбор на текстове, „които са получени въз основа на изследване на комуникативната функция на текста, се наричат външни критерии, а тези, които отразяват езиковите характеристики на текста, се наричат вътрешни критерии. Корпусите следва да бъдат проектирани и изграждани изключително въз основа на външни критерии“ (Синклер 2005: 6). „Представителността и балансираността са тясно свързани с илюстративността“ (Коева 2014: 33). Определянето на текстовите типове, които да бъдат включени при изграждането на корпус, предполага използваните корпусни единици да демонстрират в максимална степен всички езикови свойства за текстовия тип, към който принадлежат, тъй като „различните езиковите особености са разпределени по различен начин (в текстове, между текстове, между различни типове текстове) и представителният корпус трябва да даде възможност за анализ на тези различните срещания“ (Байбър 1993: 243). Постигането на илюстративност е зависимо от избора на представителни текстове, чийто брой за всяка текстова категория трябва да бъде съобразен с честотата или значимостта на тази категория за езиковия феномен, който се илюстрира.

В много отношения напредъкът в изследванията на естествения език зависи пряко от наличието на езикови данни. Това важи с пълна сила за областта на статистическия и невронния машинен превод, които разчитат на големи количества паралелни текстове (Коен 2004: 79). Тъй като и двата подхода се нуждаят от достатъчно представителни данни, качество на паралелните корпуси, използвани за обучение на подобен тип системи, е важен фактор за приложимостта и надеждността на системите. Поради тази причина подборът и подготовката на паралелните корпуси обуславят качеството на получените резултати. Това може да се демонстрира с широко достъпните системи за машинен превод, които предлагат добър превод за езици с достатъчно налични ресурси. От това следва, че съвременните системи за машинен превод зависят пряко от качеството и количеството на наличните езикови ресурси. Езици или определени езикови феномени, за които няма достатъчно представителни корпуси, са слабо проучени и възможностите,

които съвременните системи за машинен превод предоставят, в повечето случаи са неприложими за тях. Не за всички езици има лесно достъпни и достатъчно големи и представителни езикови корпуси. Поради липсата на такива корпуси е трудно подходите на статистическия и невронния машинен превод да бъдат използвани за създаването на системи за автоматичен превод или за емпирични изследвания. Ресурсите, които интернет предлага, не могат да бъдат използвани пряко за обучението на системи за машинен превод, поради факта, че всеки свободно достъпен текст трябва да бъде обработен преди да се превърне в надежден езиков корпус. Стъпките, които тази обработка включва са: събиране на текстове; съотнасяне на текстове; подготовка за статистическа обработка (нормализация, токънизация); съотнасяне на изреченско ниво на оригинали с техните преводи (Коен 2004: 79). Представените характеристики на езиковите ресурси показват, че възможностите, които статистическите и невронните подходи предлагат за изследване на езиците, зависят преди всичко от наличието на достатъчно представителни данни за тях.

III.3. Мотивация за избора на корпусния подход при изследването на българско-английските граматични паралели

Темпоралните системи в български и английски споделят основна характеристика – състоят се от категории, които са йерархично и структурно организирани и могат да се разглеждат като членове на хиперкатегория. Както в български, така и в английски съществува категория, която изразява завършеността (резултативността) на действието спрямо даден ориентационен момент – видът на действието (резултативността) в български и съответно формите, натоварени със значение перфективност (завършеност) при категорията вид на действието в английски. От представения сравнителен анализ се вижда разликата при категорията вид на действието в английски и категорията вид на действието в български. При категорията вид на действието в английски има продължителни времена, които изразяват незавършено действие спрямо определен ориентационен момент, докато при категорията вид на действието в български такива форми не съществуват. Друга ясно отличима разлика е наличието на лексикално-граматичната категория вид на глагола в български, която е част от семантиката на лексикалната единица като цяло и дава информация за „отношението на действието към неговия предел“ (Куцаров 2007: 551).

Това допринася за една от основните трудности при превод между български и английски – предаването на значението на лексикално-граматична категория вид на глагола от български на английски.

По-големият брой възможни морфологични категории, следователно и по-големият брой на възможните граматикализирани значения в български език допринася за по-голямата многозначност по време на превода поради факта, че граматикализираната информация от изходния език трябва да бъде намалена или неравномерно разпределена между различни граматични категории на целевия език. Независимо от това характеристиките на граматикализираната информация в българските и английските глаголни форми имат много съпоставими характеристики (Лазаров 2016а): и двата езика притежават морфологична категория време, която изразява отношението на действието спрямо момента на говорене; в български и английски има наличие на морфологична категория, която изразява отношението на резултат от действие спрямо ориентационен момент – в български това е категорията вид на действието, а в английски това е категорията вид на действието, която е сложна морфологична категория и може да изразява и незавършеността/продължителността на действието спрямо ориентационен момент; във формообразуването на двата езика се използват спомагателни глаголи, които носят част от граматичното значение на глаголната форма; морфологичните категории, изграждащи темпоралните значения, в двата езика могат да бъдат разглеждани като хиперкатегория. Оттук следва и очакването, че при превод от български на английски глаголните форми споделят близки граматични значения и изразяват близки темпорални отношения. Поради характеристиките на двата езика може да се очаква, че част от граматикализираната информация се предава чрез различни езикови средства и на различни езикови равнища при семантичния трансфер при превод между двата езика – например значението на категорията вид на глагола в български няма как да се предаде точно в английски чрез идентична категория. Граматичното значение за лице и число се носи от всяка лична глаголна форма в български език, а аналитичните глаголни форми, образувани с минало свършено деятелно причастие, носят информация и за граматичния род на вършителя на действието, докато в английски език има само една глаголна форма, която обслужва цялата парадигма, с изключение на формите за сегашно просто време, сегашно продължително време, сегашно перфектно продължително време и минало продължително време.

Голямото разнообразие от български глаголни форми води до редица затруднения при превод към английски. В предишни разработки по темата е демонстриран методът за машинен превод, базиран на трансферните правила, като възможно решение за проблема при превод на глаголните форми от български на английски (Лазаров 2016а). От една страна, по-малкият брой глаголни форми в английски може да бъде от помощ, тъй като при трансфера на по-голям брой български глаголни форми към по-малък брой форми в английски броят на възможните трансферни правила е равен на броя на възможните форми в целевия език (Лазаров 2016а). От друга страна, повечето български глаголни форми, които граматикализират информация за наклонение, евиденциалност и залог, имат твърде рядка употреба в съвременния език. Разбира се, трябва да се спомене, че при съставянето на изчерпателни трансферни правила за превод тези проблеми могат да бъдат избегнати чрез разработването на по-детайлни и контекстно зависими правила. В настоящата работа се застъпва тезата, че тези проблеми могат да получат по-изчерпателно и прецизно тълкуване чрез прилагането на методите на статистическото езиково моделиране и машинния превод.

Беше посочено, че трансферните правила могат да съдържат информация за точния семантичен и лексикален трансфер при превод между двата езика, въпреки това не е възможно приписването на трансферни правила със 100% вероятност поради големия обем граматична информация, която се предава чрез допълнителни езикови средства при превод. Независимо от множеството си недостатъци трансферните правила позволяват да бъде представено (граматичното) съотношение на избраните езици и илюстрират подробно лексикалния и семантичния трансфер, който се осъществява при превода. Тъй като по същество трансферните правила представляват вид граматика, която отразява зависимости между двата езика, е възможно такава граматика да бъде допълнена с информация за актуалната употреба на глаголните форми – да бъде създаден преводен модел за съотнасяне на глаголните форми между двата езика. Статистическият преводен модел може да бъде от помощ за съставяне и прилагане на трансферни правила при превод. Стъпка към създаването на преводен модел и/или извличането на надеждни трансферни правила за превод е конструирането на достатъчно представителен и обемен паралелен двуезиков корпус със съотнесени глаголни форми и налична информация за частите на речта.

Най-често споменаваните предимства на статистическия машинен превод са: в общия случай не е необходима адаптация за конкретни езикови двойки; изисква се по-

малко време за разработване и е приложим за различни езикови двойки (за сравнение подходът с правила изисква ръчната разработка на правила, което предполага много време, а правилата обикновено не могат да се използват за други езици, при това подходите с правила често създават буквални преводи) (Хътчинс 1995: 436). Преводите, генерирани от системи за статистически машинен превод, са по близки до естествения език, но за сметка на това използването на този подход често създава преводи, които включват очевидни грешки (ако няма достатъчно количество информация в преводните и езиковите модели). При подхода с правила моделирането на езика се опира на възприети теоретични концепции за езика, докато резултатите от статистическото моделиране се опират на реални езикови проявления, извлечени от актуалната употреба на езика.

Статистическото езиково моделиране може да предостави възможност за сравнително изследване на особеностите на темпоралните системи в двата езика, извън априорното познание за езика и граматичните категории. Също така чрез статистическото езиково моделиране може да бъде даден отговор на въпросите, поставени при сравнителния теоретичен анализ на темпоралните системи на двата езика. Друга гледна точка при използването на статистическа информация, извлечена от паралелен двуезиков корпус, е изследването на особеностите на двата езика поотделно – честотата, разпространението и употребата на глаголните форми в езика – както е посочвано „някои от типовете форми имат крайно ограничена (почти нулева) фреквентност. Този факт се обуславя от голямата им натовареност с граматични значения. А както е известно, много сложните неща се срещат рядко“ (Куцаров и др. 1999: 448)[6]. В този смисъл е очаквано, че по-комплексните глаголни форми в български език, натоварени със сложни значения, са рядко използваеми в съвременния език и биват измествани от форми, които носят по-малко граматикализирана информация в по-широк лексикален контекст.

Също така е възможно и представянето на честотата на употреба на различните глаголни форми при различните типове текстове. В зависимост от обема и качеството на текстовете в корпуса може да се направи заключение за количеството и вида на информацията, която се „губи“ при превод, и за начините за нейното пренасяне в целевия език.

[6] Коментарът на Ив. Куцаров е по отношение на пасивните форми на *преизказно конклузивно бъдеще предварително време в миналото* (в терминологията на автора). Както се вижда, твърдението може да бъде приложено и по отношение на други глаголни форми с комплексно значение.

Усилията за подобряването на машинния превод са насочени в няколко посоки. Една от тях е интегрирането на по-богата синтактична, семантична, прагматична информация (включително използването на дискурсивни маркери за оптимизиране на превода, усъвършенстване на превода на определени класове думи или парадигми, какъвто е случаят със системата на глаголните форми в български език) и друга езикова информация. Основното приложение на системите за компютърна обработка на езика, включително за машинен превод – бързият достъп до различна информация на друг език, не е единственото преимущество. Съществуват други области на приложение на тези системи, в които надеждният машинен превод е задължителен компонент, като автоматичен превод на стандартизирани текстове с ограничена езикова употреба (например новини, прогноза за времето, нормативни актове); извличане на информация за тема и жанр на текстове; езиково обучение и др. Системите за компютърна обработка на езика имат съществена приложимост за изучаване на езикови феномени – особености на граматиката на отделен език или съпоставително изследване на граматиката на различни езици, прагматични зависимости при употребата и др.

Една от основните цели на тази дисертация е създаването и описанието на преводен модел, съдържащ информация за преводното съотнасяне на българските глаголни форми с английските им съответствия. За осъществяването на поставената цел е необходимо: да бъде създаден достатъчно голям паралелен корпус, анотиран с глаголните форми в български и техните преводни съответствия на английски; да бъдат извлечени и анализирани статистически данни от анотирания корпус и да се направят експерименти с преводни модели, които са обогатени с данните от анотирания корпус, за да се провери практическото му приложение.

Резултатите от работата по създаването на лингвистично обогатен статистически модел за автоматичен превод ще бъдат основно в две насоки: теоретична – изясняване на зависимостите при превод на български глаголни форми на английски, и практическа – обогатяване на статистически преводни модели за автоматичен превод с лингвистична информация за преводното съотнасяне на глаголни форми, изразяващи различни категории. Анализът на честотата на глаголните форми в български и преводните им съответствия на английски предоставя възможност за изследване на спецификите на темпоралните категории и преноса на информацията, която те изразяват при превод. Изследването на особеностите на глаголните форми и граматичните категории чрез методите на статистическия анализ е продиктувано от доказания успех на този метод.

Трябва да се посочи, че статистическите данни не могат да бъдат тълкувани изолирано от теоретичния анализ на изследвания езиков феномен, а в съпоставителен план.

III.4. „Българско-английски паралелен корпус със съотнесени глаголни форми“

„Българско-английският паралелен корпус със съотнесени глаголни форми“ е създаден през периода 2017 – 2018 година, като целта му е да бъде средство за изследване на граматичните паралели между глаголните системи на български и английски: реализациите на морфологичните категории, влизащи в състава на темпоралните системи на двата езика. Корпусът е стъпка към формалното описание на граматичните паралели между темпоралните системи на български и английски и извличането на статистически данни за отношенията на преводните реализации на глаголните форми в двата езика. В контекста на изследванията по темата за граматичните сходства и различия между българската и английската глаголни системи предназначението на „Българско-английския паралелен корпус със съотнесени глаголни форми“ е: първо, да представя изчерпателна и надеждна езикова информация и второ, да бъде практически приложим ресурс за създаването на статистически модел за превод на глаголните форми.

Целите при създаването на Корпуса са основно в две насоки – теоретични и практически. Теоретичните цели са да се подберат и съотнесат обективни количествени и дистрибутивни езикови данни за темпоралните системи на български и английски и изчерпателни (за наличния обем) качествени и количествени езикови проявления за надежден анализ. Практическите цели се свеждат до предоставяне на достоверни лингвистични данни, подпомагащи създаването на различни езикови приложения за компютърна обработка на естествения език като статистически модел за анотиране и статистически модел за превод на глаголните форми.

Работният процес по „Българско-английския паралелен корпус със съотнесени глаголни форми“ може да бъде формално разделен на три основни етапа: подбор, анализ и оценка на наличния езиков материал; ръчна анотация на събраните данни; последваща полуавтоматична анотация и ръчна корекция на данните.

III.4.1. Подбор на езиковите ресурси, включени в Корпуса

„Българско-английският паралелен корпус със съотнесени глаголни форми“ съдържа извадки от други езикови ресурси, подбрани по предварително разработени критерии. Стъпката, която предхожда създаването на Корпуса, е подборът и оценката на подходящи и надеждни езикови ресурси, които отговарят на целите на създаването на Корпуса (Лазаров 2016б).

Първоначалният критерий, по който са избрани езикови ресурси за включване в „Българско-английския паралелен корпус със съотнесени глаголни форми“, е наличието на качествена морфологична анотация, тъй като тя представлява съществена част от обработката на лингвистичните ресурси за целите на по-късното извличане на статистически данни и конструирането на езиков модел. Необходимо е да бъде направено уточнението, че трите основни характеристики за качествен езиков корпус – илюстративност, представителност и балансираност, са присъщи за избраните от нас езикови ресурси в различна степен.

Основен анализ на езикови ресурси, които могат да бъдат интегрирани в „Българско-английския паралелен корпус със съотнесени глаголни форми“, е направен за „Българския национален корпус“ (заедно с българските паралелни корпуси) и „Българския корпус, анотиран с части на речта и граматични характеристики“. Основание за това е преди всичко наличната комплексна анотация на различни езикови равнища.

„Българският национален корпус“ е най-големият представителен корпус за български език. „Българският национален корпус“ се състои от ядрен корпус, включващ текстове само на български език, и 47 различни по големина паралелни корпуса на други езици (Коева 2014: 38). Общият обем на Корпуса е около 5,4 милиарда думи. Различните нива на анотация в „Българския национален корпус“ – подробни метаданни, едноезикова и частична многоезикова анотация, позволяват работата с него за различни цели. Характеристиките на „Българския национален корпус“, които го правят подходящ за изпълнението на целите на дисертацията, са:

- подробна морфологична анотация на българските текстове, направена с помощта на Българската многокомпонентна система за първична обработка и

лингвистична анотация на текстове[7], и последваща ръчна проверка на автоматичната анотация (за част от Корпуса);

- наличието на паралелни корпуси на 47 езика (в това число и английски), което позволява използването му за конструирането на статистически модели за превод.

Част от 47-те паралелни корпуса на „Българския национален корпус“ е и „Българско-английският паралелен корпус“ (Коева и др. 2012а). С големината си от около 260 милиона думи за език той представлява и най-големият паралелен корпус в състава на „Българския национален корпус“. Подобно на останалите паралелни корпуси и „Българско-английският паралелен корпус“ се състои от текстове, които имат българско съответствие – като българският текст може да бъде оригинал или превод, както и превод от трети език. Съставна част от „Българско-английския паралелен корпус“ е „Българско-английският паралелен корпус със съотнесени изречения (и прости изречения в състава на сложното)“[8] (Коева и др. 2012б). Неговият обем е от около 367 000 думи, разпределени неравномерно между български и английски. Корпусът включва различни нива на едноезикова и многоезикова анотация. Ключовите характеристики на „Българско-английския паралелен корпус“ заедно с „Българско-английския паралелен корпус със съотнесени изречения“, които го правят подходящ за поставената цел, са:

- подробна морфологична едноезикова и многоезикова анотация, направена с помощта на Българската многокомпонентна система за първична обработка и лингвистична анотация на текстове;

- многоезикова анотация, която включва съотнасяне на изреченията в двата езика (и на простите изречения в състава на сложното), като съотнасянето е проверено или направено от човек, а проверката и корекцията на автоматичното съотнасяне са извършени със специално разработена програма.

Българско-английският паралелен корпус и „Българско-английският паралелен корпус със съотнесени изречения“ са особено ценни езикови ресурси за създаването на статистически езиков модел, тъй като са снабдени с достатъчно метаданни и лингвистична информация за осъществяване на по-нататъшна статистическа обработка.

[7] https://dcl.bas.bg/webservices/
[8] https://dcl.bas.bg/bulenac/

Наличието на подробна морфологична анотация на езиковия материал в корпусите е предпоставка за по-нататъшната им обработка за различни цели. Една от основните характеристики при избора на ресурси е именно тази – подробна едноезикова анотация за части на речта и граматични характеристики, на която да бъдат базирани статистическите езикови модели. Разбира се, при създаване на статистически модел от съществено значение е количеството на данните.

Изборът на ресурси за създаване на „Българско-английския паралелен корпус със съотнесени глаголни форми“ е продиктуван най-вече от наличието на подробна морфологична анотация в описаните корпуси, тъй като целта е да се проследи какво се случва при превод с граматикализираната в глаголните форми информация за темпоралните значения. Всички изброени по-горе ресурси притежават анотация за части на речта и граматични характеристики. Разбира се, представените корпуси се характеризират с различни нива и различни типове анотация, което възпрепятства директната работа с тях и изисква унифициране на анотационните модели на различните корпуси. От представените корпуси единствено „Българско-английският паралелен корпус със съотнесени изречения“ съдържа съотнесени езикови единици на равнище изречения и прости изречения в състава на сложното, което улеснява до известна степен работата с него за целите на изследването.

Оценката на всеки от подходящите ресурси е направена според няколко критерия. Критериите за подбор на езиковите ресурси могат да бъдат разделени според няколко характеристики:

- според характеристиките, които са възприети за езиков корпус: да бъде представителен по отношение на езиковия феномен, който представя (спецификите на глаголната употреба предполагат небалансираността на текстовите типове, които ще бъдат включени в Корпуса);
- според съвременното състояние на ресурсите: друго основно изискване е ресурсът да представлява актуалната употреба на езика;
- според допълнителната лингвистична информация в наличните ресурси: подходящите ресурси трябва да представляват паралелни българско-английски сдвоени текстове на определено ниво – изречения или параграфи, и да имат анотационен слой с морфологична информация – маркери за част на речта и граматични характеристики;

- според екстралингвистичните характеристики на ресурсите: в метаинформацията на ресурсите трябва да бъде посочено кой е езикът на оригиналния и езикът на преводния текст (посоката на превод в ресурса);

- според авторските права: естеството на подбора предопределя и статута на авторските права на ресурсите: те трябва да бъдат освободени от авторски права или да бъдат предоставени такива за целите на работата.

В хода на подбора и оценката на езиковите ресурси някои от дефинираните критерии са преработени заради особеностите на самите ресурси и целите на създаване на Корпуса. Степента на промяна на критериите ще бъде представена по-подробно в „Характеристика на „Българско-английския паралелен корпус със съотнесени глаголни форми“.

С оглед на определените критерии изборът на ресурси е ограничен до „Българско-английския паралелен корпус със съотнесени изречения“, събиране на освободени от авторски права двуезикови документи от интернет и извлечени за целта примери от „Българска граматика“ (Bulgarian Grammar) (Ницолова 2017). Събирането на езикови данни от интернет е предприето с цел да се осигури наличието на срещания на глаголни форми, които не се срещат в инкорпорираните части от „Българско-английския паралелен корпус със съотнесени изречения“. Изборът е продиктуван най-вече от прагматични съображения и от факта, че инкорпорирането на твърде голям обем ресурси е извън практическите възможности на изследването.

III.4.2. Структура на „Българско-английския паралелен корпус със съотнесени глаголни форми“

„Българско-английският паралелен корпус със съотнесени изречения“ представлява надежден паралелен българско-английски корпус, съдържащ информация за морфологичните характеристики на думите, информация за източника и целевия език за всеки от съставящите го текстове. От друга страна, събраните документи от интернет, са подбрани така, че да съдържат информация за жанра, посоката на превод и други специфики на текста, но събирането им не е подчинено на строго определена таксономия, дефинирана върху текстовите типове и техните характеристики. Трябва да се посочи все пак, че и трите избрани източника („Българско-английският паралелен корпус със съотнесени изречения“, текстовете от интернет и примерите от „Българска граматика“) не отговарят в пълна степен на предварително зададените критерии. За

целите на Корпуса се налага унификация и всички предимства и недостатъци на описаните езикови ресурси са взети предвид при изграждането на структурата на „Българско-английския паралелен корпус със съотнесени глаголни форми“. Корпусът се състои от избрани части от „Българско-английския паралелен корпус със съотнесени изречения“ и текстовете от интернет, като за минимална единица, представляваща всеки отделен текст, е възприет параграфът. Инкорпорираните работни единици представляват както специално подбрани параграфи от отделни текстове, така и цели кохерентни текстове от различни източници – новини, художествена литература, филмови субтитри и други произведения (Лазаров 2019а: 109). Метаинформацията на всеки отделен документ в Корпуса включва данни за източника (име на файла или URL), изходния и целевия език, датата на събиране и датата на вписване в Корпуса (ноември 2018 – януари 2019), наличието на маркери за част на речта и граматични характеристики и информация за тематичната област (художествени текстове, публицистични текстове, научни текстове, субтитри и примери от „Българска граматика“). В крайния си вид „Българско-английският паралелен корпус със съотнесени глаголни форми“ се състои от 84% текстове от „Българско-английския паралелен корпус със съотнесени изречения“, 15,5% текстове от интернет и 0,5% от изречения, извлечени от примери от „Българска граматика“. За да бъдат упоменати авторските права на примерите от „Българска граматика“, информацията за източника е включена в метаданните на съответните файлове.

При работата върху дизайна на Корпуса е възприет един от основните методи за постигане на балансираност при изграждането на езиков ресурс – базирането му на вече възприета структура и извличането на стратифицирани случайни извадки от по-голям корпус. При първоначалната оценка на подходящите ресурси е избран Българско-английският паралелен корпус със съотнесени изречения (Коева и др., 2012а). По външни критерии са подбрани ограничен брой текстове от интернет, както и примери от „Българска граматика“ с цел изчерпателност на данните в Корпуса. Първоначалният критерий за предварително дефинирана посока на превода на текстовете, включени в Корпуса, е изоставен поради липсата на достатъчно ресурси. Подбраните текстове от интернет са разделени по област и жанр (в „Българско-английския паралелен корпус със съотнесени изречения“ те вече са разделени), след което поетапно са включени в състава на „Българско-английския паралелен корпус със съотнесени глаголни форми“. За всяка една част от подбраните ресурси са извлечени статистически данни за количеството и

разнообразието на глаголните форми за дадения текстов тип. След всяко ново добавяне на езиков материал в Корпуса се проверява съотношението между общия брой на българските глаголните форми и разнообразието от темпорални форми с цел да се избегне нарастването на еднообразни примери. В този смисъл таксономията на текстовите типове в Корпуса няма дефинирана структура, а е основана на периодична проверка за езиково насищане на Корпуса – при внасянето на поредния обработен текст към него се проверява дали съотношението на изследваните езикови единици (глаголните форми) ще се промени значително, или броят им ще остане сравнително постоянен. Както е известно (Байбър 1993: 252 – цит. по Коева 2014: 34), езиковите характеристики с висока честота на употреба са сравнително стабилни при разпространението си и затова за тяхната илюстрация са достатъчни относително кратки извадки, докато за по-редките езикови характеристики са необходими по-големи текстове. Ниската фреквентност на някои темпорални форми в български предполага, че при фазите на добавяне на езиков материал към Корпуса и достигането на езиково насищане определените предполагаеми стойности за честотата на тези темпорални форми няма да се променят значително и те ще запазят ниските си стойности през целия процес на работата. С други думи – при включването на специално подбран по определени критерии текстов тип, който внася нови срещания на темпорални форми към масива, не може да се очаква, че броят на новите срещания ще продължи да се изменя пропорционално при следващите етапи на работа. В този смисъл „Българско-английският паралелен корпус със съотнесени глаголни форми“ представлява небалансиран езиков корпус.

„Българско-английският паралелен корпус със съотнесени глаголни форми“ е предназначен за изследване на граматичните паралели между темпоралните системи на български и английски (както теоретично, така и с практическа насоченост). Представителността като характеристика на Корпуса е предварително изискване. Основната цел, поставена при създаването на Корпуса, е да бъде надежден езиков ресурс, който представя изследвания феномен: трансфера на темпоралните форми при превод от български на английски. Тази характеристика на Корпуса оказва влияние върху структурата и дизайна му. Представителността по отношение на преводните реализации на глаголните форми между български и английски поражда конфликт между балансираността и илюстративността на Корпуса – от една страна, ако основен критерий е само постигането на илюстративност, балансираността на текстовите типове ще бъде

нарушена поради факта, че глаголните форми с ниска честота на срещане не са характерни за типовете текстове, включени в Корпуса, а от друга страна, балансираността на текстовете ще окаже влияние на илюстративността по отношение на нискочестотните глаголни форми.

III.4.3. Анотационни конвенции при създаването на „Българско-английския паралелен корпус със съотнесени глаголни форми“

Лингвистичните данни в Корпуса имат два слоя анотация. При работата върху Корпуса са възприети няколко анотационни конвенции (Лазаров 2019б: 7). При първия слой са използвани вече съществуващи инвентари от маркери за морфологичните характеристики на думите, които са уеднаквени в процеса на работа. Първият слой е слоят на морфологичните характеристики на думите и се състои от маркери, които ги илюстрират. Текстовете в „Българско-английския паралелен корпус със съотнесени изречения“ са анотирани с помощта на Българската многокомпонентна система за първична обработка и лингвистична анотация на текстове (Коева и Генов 2011). Тя включва програми за обработка – токънизатор и разделител на изречения, основани на регулярни изрази, тагер за определяне на частите на речта, лематизатор на основата на речник за определяне на основната форма и граматичните характеристики на словоформата, които са адаптирани да работят в единна система, при което се осигурява тяхната свързаност, ефективност и висока точност (Коева и Генов 2011). Апаче ОпънЕнЕлПи (Apache OpenNLP) с предварително обучени модели и Станфорд КорЕнЕлПи (Stanford CoreNLP) са използвани за анотация на английските текстове в Българско-английския паралелен корпус със съотнесени изречения – сегментиране на изреченията, токънизация и приписване на маркери за морфологичните характеристики на думите. Лематизирането на английските текстове се извършва с помощта на Станфорд КорЕнЕлПи (Stanford CoreNLP) и РАСП (RASP) (Брискоу и др. 2006).

Прибавянето на анотация за морфологичните характеристики на думите от ресурсите, за които няма налична анотация, и уеднаквяването на анотационните модели между тях и ресурсите с налична анотация („Българско-английския паралелен корпус със съотнесени изречения“) представлява важна стъпка от работата поради факта, че от първия слой анотация зависят следващите етапи от работата – приписването на втория слой анотация и извличането на представителни статистически данни от корпуса.

Първоначално събраните езикови данни са разделени на малки работни файлове, в които са съотнесени на изреченско ниво. Всеки файл получава идентификационен пореден номер, а всяка двойка изречения получава пореден номер в рамките на общия брой на съотнесени изречения в корпуса. Езиковите ресурси, събрани от интернет, и примерите от „Българска граматика“ са анотирани с ТриТагер (TreeTagger). ТриТагер е софтуерно приложение за анотиране на текстове с информация за частите на речта и лематизиране. Приложението е създадено от Х. Шмит (Шмид 1995) като част от проекта Техническа комуникация (Technical Communictaion) към Института за компютърна лингвистика в Университета в Щутгард. Изборът на ТриТагер е продиктуван от факта, че приложението е приспособимо към различни езици и нива на анотация, ако се представят налични речници и ръчно анотирани ресурси. Трябва да бъде отбелязано, че инвентарът от маркери за морфологичните характеристики на думите в ТриТагер (Санторини 1991) се различава от наличния инвентар от маркери за морфологичните характеристики на думите в „Българско-английския паралелен корпус със съотнесени изречения“. След анотирането на ресурсите без анотация двете анотационни схеми са уеднаквени с помощта на ТриТагер, който позволява конвертиране между различни анотационни конвенции. Уеднаквяването на анотационните маркери между Българско-английския паралелен корпус със съотнесени изречения и ресурсите, анотирани с ТриТагер, е наложено поради спецификите на софтуерното приложение, използвано за втория анотационен слой. Използваният инвентар от маркери в „Българско-английския паралелен корпус със съотнесени изречения“ е описан за български (Коева 1998)[9], а за английски инвентарът от маркери на Апаче ОпънЕнЕлПи е съотнесен с българския. Използваният от ТриТагер инвентар от маркери за морфологичните характеристики на думите е описан за английски (Санторини 1991) и за български (Симов и др. 2004). Наличната анотация от „Българско-английския паралелен корпус със съотнесени изречения“ и анотираните чрез ТриТагер ресурси от интернет са уеднаквени и приведени към морфо-синтактичните спецификации на МУЛТИТЕКСТ-Ист (MULTEXT-East)[10]. След приключването на работата по втория анотационен слой маркерите за анотация на първия анотационен слой на целия Корпус са реконвертирани към инвентара от маркери за морфологични характеристики, използван в „Българско-английския паралелен корпус със съотнесени изречения“. По този начин вторият анотационен слой за съотнасянето на темпоралните категории при

[9] http://dcl.bas.bg/english-bulgarian-pos-tagset/
[10] http://nl.ijs.si/ME/

превод от български към английски надгражда морфологичната анотация на Българско-английския паралелен корпус със съотнесени изречения.

В следващата стъпка от хода на работата езиковите ресурси с унифициран първи слой анотация са разделени в 22 работни файла, съдържащи неравномерно разпределени сдвоени текстове и морфологичните характеристики на думите в тях, а където липсва сдвояване (за текстовете, събрани от интернет) са сдвоени ръчно на ниво изречение. Всяка двойка изречения на български и английски получава пореден номер в рамките на работните файлове. След процеса на анотиране, конвертиране и сдвояване, данните са проверени ръчно и са направени корекции, където е необходимо. Анотираните работни файлове са разделени за двата езика и към тях е добавена метаинформация за източника на текста; посоката на превод (където е налична); датата на присъединяване на текста в корпуса; дали съответният текст представлява извадка, или не от съответния файл на Българско-английския паралелен корпус със съотнесени изречения или е от интернет; жанра на текста (за текстовете от интернет). Всеки от работните файлове получава идентификационен номер и е конвертиран в триколонен TSV (tab separated values) файл. Всеки ред от файла се отнася за една дума/токън. Първата колона от всеки ред от файла съдържа думата/токъна, втората колона представя лемата на думата, а третата колона е съответният маркер за морфологичните характеристики на думите. Празният ред сигнализира изреченска граница (Шмид 1995).

Вторият анотационен слой съдържа информация за аналитичните глаголни форми и тяхното съотнасяне в двата езика. Лингвистичната анотация на втория анотационен слой в „Българско-английския паралелен корпус със съотнесени глаголни форми“ е разработена, за да бъдат изпълнени целите на създаването на Корпуса (Лазаров 2018). Инвентарът на втория анотационен слой се състои от 26 стойности, които представят информация за граматичната хиперкатегория време, разглеждана според класическите наименования на членовете ѝ. При разработването и прилагането на маркерите от този слой са възприети следните принципи:

- лингвистичните маркери за български и английски представят членове на хиперкатегорията време в двата езика – 9 за български и 17 за английски (16 за темпоралните форми и 1 за инфинитива);
- лингвистичните маркери не съдържат информация за лице, число или род на глаголните форми, тъй като тази информация се формира от стойностите на слоя с маркери за морфологичните характеристики на думите;

- допустимо е съотнесените форми да не бъдат поравно разпределени в двата езика – т.е. възможно е глаголна форма в български да няма съответния еквивалент в английски (преводът може да бъде осъществен с други части на речта) и обратно;

- маркерите отразяват само темпоралните характеристики на формите, без да отразяват другите граматични характеристики като наклонение, залог, евиденциалност, положителна или отрицателна форма. Отрицателни форми, форми, които са в пасивен залог, или форми, които не са в индикатив, също са маркирани, като са отразени само темпоралните им характеристики.

Вторият анотационен слой е добавен ръчно и/или полуавтоматично чрез софтуерно приложение, което позволява цялостно използване на функция за приписване на лингвистична информация на базата на вероятностен модел и последваща корекция. При работата върху него е спазен принципът на надграждането между равнищата – „различните равнища на анотация приписват нова информация към езиковите единици, без да променят анотациите от предходните равнища“ (Коева и др. 2010: 3681). За втория слой на анотация е използван свободно достъпният инструмент за работа с езикови ресурси УебАно (WebAnno) (Имам и др. 2014). УебАно е универсален инструмент за анотиране за различни равнища, включващи различни слоеве от морфологични, синтактични и семантични характеристики. Освен предварително наличните инвентари от маркери, могат да бъдат дефинирани допълнителни равнища от езикови характеристики, което позволява УебАно да се използва и за структурно недетерминирани цели като приписване на характеристики върху езиковите данни, без да е налична предварителна информация за морфологичните характеристики на думите. Приложението поддържа различни режими на анотация, включително режим на ръчно приписване на характеристики, корекция и самообучение, въз основа на предварително дефинирана таксономия и структурни правила, което позволява приписване на лингвистична информация на базата на вероятностен модел. УебАно позволява работа с множество файлови формати, но за целите на Корпуса е избран форматът CoNLL[11]. УебАно използва преработена версия на файловия формат CoNLL-X[12]. Работните файлове се кодират в нормализирани текстови файлове (UTF-8, с използване на LF (line feed) символ за прекъсване на реда, включващ LF символ в края на файла/текста) с три

[11] https://universaldependencies.org/format.html
[12] https://ilk.uvt.nl/conll/#dataformat

типа редове: редове, съдържащи анотацията на думата/токъна с 5 колони, разделени с табулация; празни редове, маркиращи края на изреченията; редове с коментари (Букхолз и Марси 2006). Изреченията се състоят от един или повече редове с думи, а редовете с думи съдържат следните колони:

- пореден номер на токъна;
- токън;
- маркер за морфологичната характеристика на думата от първия анотационен слой;
- маркер за съставната глаголна форма, ако токънът е част от такава;
- идентификационен номер на глаголната форма, който може да бъде в три различни отношения между двата езика:
- идентификационният номер на глаголната форма в български да отговаря на идентификационния номер на глаголната форма в английски;
- идентификационният номер на глаголната форма в български да няма съответния еквивалент в английски, поради превод в английски с други лексикални средства, които не са глаголна форма;
- идентификационният номер на глаголната форма в английски да няма съответния еквивалент в български, поради превод на части от българското изречение, които не са глаголни форми, с глаголни форми в английски.

Стойностите на езиковите маркери от втория анотационен слой представляват краен списък от по-малко на брой възможни маркери, отколкото в първия слой на анотацията. Те се приписват върху първия анотационен слой, съдържащ маркери за морфологичните характеристики на думите. Инструментът УебАно има функция да разглежда глаголните форми (в този случай) като компоненти, съставени от определен брой структурни части, която позволява един маркер от втория слой да може да бъде приписан на повече от една единица от първия слой (Имам и др. 2014). Използваният инвентар от маркери за втория анотационен слой са:

За български:

– Vpraesens глаголна форма в сегашно време;
– Vaor глаголна форма в минало свършено време;

– Vimperf глаголна форма в минало несвършено време;

– Vfutur глаголна форма в бъдеще време;

– Vperfect глаголна форма в минало неопределено време;

– Vplusqperf глаголна форма в минало предварително време;

– Vfutexact глаголна форма в бъдеще предварително време;

– Vfutexpraet глаголна форма в бъдеще предварително време в миналото;

– Vfutpraet глаголна форма в бъдеще време в миналото.

За английски:

– Vprs глаголна форма в сегашно просто време;

– Vps глаголна форма в минало просто време;

– Vfs глаголна форма в бъдеще просто време;

– Vprp глаголна форма в сегашно перфектно време;

– Vpp глаголна форма в минало перфектно време;

– Vfp глаголна форма в бъдеще перфектно време;

– Vprc глаголна форма в сегашно продължително време;

– Vpc глаголна форма в минало продължително време;

– Vfc глаголна форма в бъдеще продължително време;

– Vprpc глаголна форма в сегашно перфектно продължително време;

– Vppc глаголна форма в минало перфектно продължително време;

– Vfpc глаголна форма в бъдеще перфектно продължително време;

– Vfsp глаголна форма в бъдеще просто време в миналото;

– Vfpp глаголна форма в бъдеще перфектно време в миналото;

– Vfcp глаголна форма в бъдеще продължително време в миналото;

– Vfpcp глаголна форма в бъдеще перфектно продължително време в миналото;

– Vinf инфинитив.

Работата по събирането, анализирането и инкорпорирането на езиковите ресурси в „Българско-английския паралелен корпус със съотнесени глаголни форми“, както и първичното анотиране на текстовете, събрани от интернет, и примерите от „Българска граматика“, синхронизирането на анотационните маркери между Българско-английския паралелен корпус със съотнесени изречения и анотираните за целта ресурси и работата по вторичното анотиране, може да бъде обобщена в по следния начин:

1. Съществено преимущество е, че за част от работния процес са използваните софтуерни приложения – ТриТагер и УебАно. Двете приложения са предопределящи за начина на работа при анализирането, анотацията и хода на работа по корпуса. Фактът, че ТриТагер предлага вече възприети и унифицирани анотационни конвенции за български и английски, както и интегрирани преобразуватели за налични анотационни конвенции, даде възможността да се използват големи масиви от езикови данни с налична анотация, като по този начин се намали времето, необходимо за унифициране на анотационните конвенции и ръчната корекция на слоя с маркери за морфологичните характеристики на езиковите данни в Корпуса. Ръчната корекция е направена в хода на приписване на втория анотационен слой в УебАно, където регистрираните грешки в първия слой с морфологичните характеристики са ръчно коригирани. Друго основно преимущество на ТриТагер е, че може да бъде обучен върху налични предефинирани набори от анотационни маркери, което позволява промяната на възприетите анотационни конвенции в хода на работата и уеднаквяването на различни набори от анотационни маркери. Тази възможност на ТриТагер е използвана за уеднаквяване на конвенциите от Българско-английския паралелен корпус със съотнесени изречения спрямо анотационните модели на ТриТагер.

2. Другото основно софтуерно приложение, което е използвано при конструирането и създаването на „Българско-английския паралелен корпус със съотнесени глаголни форми“ – УебАно, предлага множество възможности за работа с различни по вид езикови данни и анотационни модели, както и разработването на правила за автоматична анотация в хода на работата. Основното преимущество на приложението е, че работи с около 12 различни файлови формата, които са подходящи за различните стъпки от създаването на Корпуса и за целите на извличане на представителни статистически данни от

него. При първоначалната обработка на данните е използван изходният файлов формат на ТриТагер, който е съвместим с входните файлови формати на УебАно. При добавянето на втория слой са използвани посочените CoNLL-X файлови формати. Софтуерното приложение УебАно също така предлага работа с различни конвенции и нива на анотация, които могат да бъдат независими или логически обвързани по между си. В контекста на работата по втория анотационен слой на „Българско-английския паралелен корпус със съотнесени глаголни форми“ най-полезна се оказа функцията за предсказващо приписване на анотационни маркери (predictive annotation). Предсказващото приписване на анотационни маркери предлага определени маркери на следходните езикови данни в хода на работата, базирайки се на модел, който е изграден на информация от вече наличните данни и срещанията на определени лингвистични маркери в тях. Настройките на функцията позволяват да бъде моделирана по два начина: чрез контекстно зависими правила или чрез логически правила за съставните части на по-голямата анотационна единица. И при двете възможности се приписват езикови маркери върху наличните езикови данни със съответната процентна достоверност, като се позволява ръчна корекция на правилата или приписаните маркери. Тази функция на УебАно позволява анотирането на големи по обем езикови масиви да бъде по-ефективно от гледна точка на времето за работа и с по-малък праг на допустими грешки от гледна точка на консистентността на резултата. Функцията подпомага и последващата ръчна корекция чрез изчисляването на несъответствието между автоматично и ръчно приписаните или коригирани маркери.

3. Изборът на представените по-горе езикови ресурси също се оказа съществен при работата върху конструирането и изграждането на „Българско-английския паралелен корпус със съотнесени глаголни форми“. Основното изискване по отношение на езиковите данни в Корпуса е да бъдат представителни. Това означава, че данните трябва да представляват всички възможни проявления на изследвания езиков феномен в различни контексти. В обема на Корпуса тази характеристика не е реализирана в по-голяма част от текстовете. Наличната вътрешна структура на Българско-английския паралелен корпус със съотнесени изречения предостави възможността за работа върху различни по жанр и тип езикови данни, съобразно целите на Корпуса – област на

употреба, жанр, източник на текста, година на публикуване и т.н., които са използвани при предварителния анализ на ресурсите и при критериите за включване на текстовете в Корпуса. Тази характеристика на Българско-английския паралелен корпус със съотнесени изречения допринесе за илюстративността на „Българско-английския паралелен корпус със съотнесени глаголни форми“. Съобразяването с езиковото насищане при добавянето на езиковите данни от интернет и етапната работа по Корпуса също са дефинирани като необходими за работата, за да се избегне пренасищането с еднотипни езикови проявления.

4. Основният проблем при хода на работата върху „Българско-английския паралелен корпус със съотнесени глаголни форми“ е недостатъчното разнообразие от различни по вид глаголни форми както в български, така и в английски. Въпреки че Българско-английския паралелен корпус със съотнесени изречения предоставя представителни и илюстративни езикови данни, той не може да подсигури граматическо разнообразие от глаголни форми. В предишни изследвания, предхождащи конструирането на Корпуса (Лазаров 2017), е показано, че разпространението на глаголните форми в български като изходен език е неравномерно и надеждни статистически данни могат да бъдат извлечени само от небалансирани, но достатъчно големи по обем корпуси.

III.5. Характеристика на „Българско-английския паралелен корпус със съотнесени глаголни форми“

В настоящия си вид „Българско-английският паралелен корпус със съотнесени глаголни форми“ представлява неголям, небалансиран, илюстративен (за поставената си цел), двуезиков паралелен корпус с общ обем от 5 070 анотирани и съотнесени глаголни форми и техните преводни съответствия. Българските текстове съставляват 45 843 думи (48,12% от общия обем), а английските – 49 432 думи (51,88% от общия обем) (Лазаров 2019б). Първоначално поставената цел за детерминирана посока на превод – Корпусът да представлява изцяло български текстове с техните английски преводи – е изоставена по време на подбора на текстовете, поради недостатъчно свободно достъпни и представителни оригинални български текстове в интернет, които да обхващат разглеждания езиков феномен във всичките му проявления. Поради тази причина в Корпуса са включени както преводи от английски на български, така и преводи на

български и английски от текстове на трети език. Таксономията на текстовите типове няма единна структура, а е основана на периодичната проверка за езиково насищане на Корпуса – при внасянето на поредния обработен текст към него, се проверява дали съотношението на изследваните езикови единици (глаголните форми) ще се промени значително, или броят им ще остане сравнително постоянен.

Работата по „Българско-английския паралелен корпус със съотнесени глаголни форми“ е продиктувана от неговите цели и причината за конструирането му. Поставените и изпълнени задачи при разработването на Корпуса могат да бъдат разделени в две направления – теоретични и приложни (Лазаров 2019б: 13). Основната теоретично-практическа задачи е да бъде конструиран статистически преводен модел за глаголните форми между български и английски, базиран на информацията, извлечена от Корпуса – тази задача включва статистически обусловен анализ, създаване на езиков модел и преводен модел и верификацията им. Друга приложна задача е публикуването на Корпуса в интернет и осигуряването на свободен достъп до него и създаването или модифицирането на съществуващ инструмент за търсене в Корпуса и извличане на информация. Осъществяването на свободен достъп до Корпуса ще предостави възможност за бъдещи изследвания върху езиковия материал и информацията, представени в него, които са извън обхвата на първоначално дефинираните задачи за настоящето проучване.

Схематично представени, текстовете в „Българско-английския паралелен корпус със съотнесени глаголни форми“ изглеждат така (Лазаров 2019а: 111).

- По област:
    - 65,5% художествени текстове – разкази, романи, приказки и др., от които 60% представляват текстове от „Българско-английския паралелен корпус със съотнесени изречения“, а 40% – текстове от интернет;
    - 18% публицистични текстове – новини, изявления, съобщения, от които 100% представляват текстове от „Българско-английския паралелен корпус със съотнесени изречения“;
    - 3% научни текстове – изследвания, статии, от които 100% представляват текстове от „Българско-английския паралелен корпус със съотнесени изречения“;

- 13% филмови субтитри, от които 100% представляват текстове от „Българско-английския паралелен корпус със съотнесени изречения“;
- 0,5% други текстове, които са примери от „Българска граматика“;

- По посока на превода:
  - 40,1% от текстовете представляват български оригинал с английски превод;
  - 35,4% от текстовете представляват английски оригинал с български превод;
  - 24,5% от текстовете представляват преводи на български и английски от оригинал на трети език;
- По произход:
  - 84% от текстовете представляват текстове от Българско-английския паралелен корпус със съотнесени изречения;
  - 15,5% от текстовете са текстове, извлечени от интернет;
  - 0,5% от текстовете са конструирани изречения, извлечени от примери от „Българска граматика“.

В настоящия си вид „Българско-английският паралелен корпус със съотнесени глаголни форми“ представлява езиков ресурс с ограничен достъп под авторски права. Корпусът е създаден като част от проекта „Българо-английски граматични паралели с оглед на машинния превод. Обогатяване на статистически модел за превод от български на английски с лингвистична информация“ по Програма за подпомагане на младите учени в БАН.

Езиковите данни в Корпуса са структурирани в 22 двойки от работни файлове, именувани с поредни номера и съответния езиков код. Всеки отделен файл представлява tsv (tab separated values) файл, който съдържа две части.

Първата част представлява метаинформацията на езиковите данни, съдържащи се в този файл, и се състои от следните полета:

- <origin = „X“> – където X може да представлява:
  - името на корпуса/субкорпуса и името на файла, от който са извлечени езиковите данни;

- хиперлинк към интернет ресурса, от който са извлечени езиковите данни;

- – където X е булева стойност, която дава информация за това дали извлечените данни представляват някаква част от упоменатия в предишното поле ресурс, или представляват целия ресурс;

- <date = „X“> – където X е числова стойност, която представлява датата на инкорпориране на езиковите данни;

- <direction = „X“> – където X може да има следните стойности и значения:

  - bg>en – посоката на превода на съответния езиков ресурс е от български на английски;

  - en>bg – посоката на превода на съответния езиков ресурс е от английски на български;

  - bg = en – посоката на превода на съответния езиков ресурс е от трети език съответно на български и английски;

- <genre = „X“> – където X може да има следните стойности и значения:

  - literature – езиковите данни представляват художествена литература;

  - news – езиковите данни представляват публицистични текстове;

  - science – езиковите данни представляват научни текстове;

  - sub – езиковите данни представляват филмови субтитри;

  - control – езиковите данни представляват извадка от примери от „Българска граматика“.

Втората част от всеки файл съдържа езиковите данни. Те са структурирани в петколонен TSV формат, където:

- първата колона съдържа поредния номер на токъна;

- втората колона съдържа токъна;

- третата колона съдържа маркер за морфологичните характеристики на токъна;

- четвъртата колона съдържа приписан маркер на аналитичната глаголна форма, ако токънът е част от такава;
- петата колона съдържа идентификационния номер на глаголната форма, чрез който съответните глаголни форми в български и английски са съотнесени.

Файловата структура и файловият формат на работните файлове от „Българско-английския паралелен корпус със съотнесени глаголни форми“ позволяват реимпортирането им в УебАно за целите на допълнителна анотация, корекция или анотация на нови езикови ресурси чрез функцията за предсказващо приписване на лингвистична информация.

За целите на извличането на статистическа информация от „Българско-английския паралелен корпус със съотнесени глаголни форми“ и изследването на граматичните паралели при превод на глаголните форми е създадено работно приложение за работа с данните от Корпуса. Приложението е базирано на структурата и файловия формат на Корпуса. Чрез него могат да бъдат извлечени данни за съответствието между българските и английските глаголни форми, за броя на токъните при съставните глаголни форми, за вида на граматикализираната информация от отделните глаголни форми и информация за пренасянето на граматикализираната информация при превод.

IV. Анализ на граматичните паралели в „Българско-английския паралелен корпус със съотнесени глаголни форми“

Целта на описанието в четвърта глава е да се представи съпоставителен анализ между емпиричните статистически данни, извлечени от „Българско-английския паралелен корпус със съотнесени глаголни форми“, представен в третата част, и изградената формулировка за поведението на граматикализираната от темпоралните глаголни форми информация при превод от български на английски език.

IV.1. „Българско-английският паралелен корпус със съотнесени глаголни форми“ в контекста на изследването

Преди да бъдат представени статистическите данни, извлечени от „Българско-английския паралелен корпус със съотнесени глаголни форми“, ще бъдат представени основните му характеристики в контекста на изследването. Предварително беше създаден по-малък експериментален корпус, върху който бяха направени опити, свързани с бъдещата работа по Корпуса (Лазаров 2017). Това предварително изследване показва, че глаголните форми в български и английски имат неравномерно разпределение в различните текстове и не може да се очаква да демонстрират всички възможни проявления в различен контекст. Също така предварително разработеният корпус показа, че корпусно базираният подход е подходящ за извличане на статистическа информация за преводните реализации на глаголните форми от български на английски, както и че в ограничена степен наличието на определени граматични категории в български като изходен език предопределя наличието на определени граматични категории в английски като целеви език.

Голяма част от преимуществата и недостатъците на „Българско-английския паралелен корпус със съотнесени глаголни форми“ се дължат на неговата анотационна структура. Анотационните маркери приписват морфологични характеристики на синтетичните глаголни форми в първия анотационен слой, а във втория се означават аналитичните форми, съобразени с класическите схващания за категорията време в български и английски (разглеждана в дисертацията като хиперкатегория, състояща се от отделни темпорални морфологични категории).

Основната информация, която може да се извлече от Корпуса, е за употребата на темпоралните форми в двата езика поотделно и в съпоставителен план. В първия анотационен слой се анотират само синтетичните глаголни форми, включително изразяващи морфологичната категория време (сегашно време, минало свършено време, минало несвършено време за български и сегашно просто време и минало просто време за английски език), наклонение (изявително за двата езика), залог (активен залог за двата езика). Характеристиките, които вторият анотационен слой приписва, се отнасят до аналитичните темпорални форми в български и английски и не съдържат информация за други морфологични категории като наклонение, модус на изказването (Герджиков 1984: 189), залог, положителна или отрицателна форма[13]. Въпреки това върху тези форми в Корпуса е приписана информация за темпоралните характеристики на формата. Например:

(30) Щели сте да сте пътували вчера. – 17V532 - Vfutexact (бъдеще предварително време)

Глаголната форма в (30) е омонимна: едновременно умозаключителна (конклузивна) и преизказна (инверитативна, дубитативна) форма на бъдеще предварително време и бъдеще предварително време в миналото, за която в Корпуса има приписани маркери за формите на морфологичните категории време, вид на действието и таксис. При случаи на омонимия между дадени форми е приписан маркер за по-малко натоварената форма, в случая за бъдеще предварително време в индикатив. В рамките на бъдещи изследвания ще бъде изведена статистика за омонимните форми. Тази особеност на Корпуса може да бъде преодоляна при бъдеща изследвания и работа върху Корпуса и използването му като ресурс за обучение на система за граматична анотация по-голям обем от езикови данни.

Друга характеристика на втория анотационен слой е, че липсва съотнасяне между глаголните форми в единия език и други части на речта в другия език в случаите на превод на дадена глаголна форма с части на речта, които не са глаголи. В тези примери идентификационният номер на анотираната глаголна форма в единия език няма съответствие в другия език. Тази характеристика на Корпуса позволява да се извлекат статистически данни за преводните реализации между двата езика, но не позволява автоматичното извличане на данни за преводните съответствия в случаите на превод на

[13] В дисертацията се изпозлват традиционно възприети термини и не се навлиза в дискусия за същността и състава на изброените категории в български.

глаголни форми с части на речта, които не са глаголи. Преимуществото на тази характеристика, от друга страна, е, че темпоралните категории са представени в количествен и качествен аспект по отношение на категориите време, вид на действието и таксис в български и категориите време и вид на действието в английски.

В тази глава ще бъдат анализирани статистическите данни както за български и английски поотделно, така и в съпоставителен план между двата езика. Ще бъде представен и анализ на поведението на отделните темпорални морфологични категории при превод между български и английски.

В следващата глава ще бъде представено и вграждането на данните от Корпуса в софтуер за компютърно-подпомогнат превод и оценката върху предсказващото приписване на определени темпорални форми на целевия език при превод, базирано на изходния език, и извършената оценка върху предсказващото приписване от хора. Внедряването има за цел да оцени на практика статистическите данни от корпуса и надеждността им като практически приложим езиков ресурс.

IV.2. Статистически обусловен анализ на данните за българските темпорални форми в Корпуса

Разпределението на темпоралните форми за български език е представено в Таблица 7:

Време	Брой форми	Процентно съотношение
Сегашно време	2 437	47,39%
Минало свършено време	1 311	25,50%
Минало несвършено време	710	13,81%
Минало неопределено време	269	5,23%
Бъдеще време	189	3,68%
Минало предварително време	164	3,19%
Бъдеще време в миналото	47	0,91%
Бъдеще предварително време	10	0,19%
Бъдеще предварително време в миналото	5	0,10%

Таблица 7. ***Разпределение на темпоралните форми за български език.***

От Таблица 7. се вижда, че в български език формите за сегашно време съставляват почти половината от анотираните глаголни форми в Корпуса. Най-рядко

срещани са формите на бъдеше предварително време в миналото, които съставляват 0,1% от всички глаголни форми в Корпуса. Трябва да се посочи, че форми за бъдеще предварително време и бъдеще предварително време в миналото не се срещат в текстовете от „Българско-английския паралелен корпус със съотнесени изречения“ и интернет, а само в примерите от книгата „Българска граматика“. Данните доказват представеното по-горе твърдение, че по-комплексните глаголни форми имат по-ниска честота на срещане, като биват заменяни в езика с опростени форми и повече лексикален контекст, който компенсира граматичното значение. Също така се вижда, че 86,7% от всички български глаголни форми в Корпуса са синтетични, а останалите 13,3% са аналитични. При аналитичните глаголни форми съотношението токън : глаголна форма е средно 2,28:1, като пет пъти се срещат съставни глаголни форми, изградени от 4 токъна (всички от примерите от „Българска граматика“).

(31) Утре сутринта ***щях да съм пристигнал.***

Разпространението в Корпуса на глаголните форми за български език по област на употреба е представено в Таблица 8.

	Художествена литература		**Публицистични текстове**		**Научни текстове**		**Филмови субтитри**	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Сегашно време	1201	36,0	771	83%	94	94%	361	53,2%
Минало свършено време	1114	33,4	57	6,1%	1	1%	133	19,6%
Минало несвършено време	578	17,3	12	1,3%	–	–	107	15,8%
Минало неопределено време	210	6,3	34	3,7%	2	2%	22	3,2%
Минало предварително време	118	3,5	4	0,4%	–	–	14	2,1%
Бъдеще време	95	2,8	51	5,5%	3	3%	38	5,6%
Бъдеще време в миналото	23	0,7	–	–	–	–	3	0,4
Бъдеще предварително време	–	–	–	–	–	–	–	–
Бъдеще предварително време в миналото	–	–	–	–	–	–	–	–
Общо	**3339**	**65%**	**929**	**18%**	**100**	**2%**	**678**	**13%**

Таблица 8. ***Разпространение на глаголните форми за български език по област на употреба.***

От Таблица 8. се вижда, че според разпределението на темпоралните глаголни форми различните области на употреба на езика ясно се диференцират една от друга. В представените статистически данни се показва, че в целия обем на Корпуса форми за бъдеще предварително време и бъдеще предварително време в миналото не се

реализират в нито една област (причината за това е, че всички форми за тези времена са от примерите от „Българска граматика“, които не са представени в Таблица 8.). Ясно изразено е и диференцирането при глаголните темпорални форми с най-висока честота на употреба за всяка от областите. Докато при публицистичните и научните текстове има доминиране на глаголни форми в определено време, при филмовите субтитри и художествените текстове честотата на употреба на определени темпорални форми не може еднозначно да бъде изведена като диференциращ признак. Въз основа на представените статистически данни в Таблица 8. може да се направи заключение, подкрепящо тезата, че различните области на употреба на езика се характеризират с определени доминиращи части на речта (Димитрова 2009: 418), но също така и че за дадени области на употреба може да се очакват характерни темпорални отношения, изразени чрез съответните глаголни форми. Данните от Корпуса за разпространението на темпоралните форми в различни области на употреба, разбира се, не могат да бъдат тълкувани абсолютно еднозначно. Анализът зависи преди всичко от обема и обхвата на езиковите данни. Според различни изследвания (Димитрова 2009, Тодорова 1993) за честотата на частите на речта в различни типове текстове, глаголите (обобщено като част на речта, без оглед на морфологичните им характеристики) по честота на срещанията си обикновено се появяват след съществителните имена, прилагателните имена, наречията и предлозите. Спецификите на отделните типове текстове също така предполагат и разлики в употребата на предпочитаните темпорални форми. Съответно представеният анализ за характеристиките на областите на употреба на езика с оглед на разпространението на темпоралните форми трябва да бъде тълкуван в дескриптивен план и като епифеномен за особеностите на съответната област на употреба на езика.

От друга страна, обхватът на разпространение на различните темпорални форми за български показва, че най-често срещаните форми имат и най-голям обхват на разпространение в различните типове текстове, докато темпоралните форми с най-ниска честота на срещане имат сравнително ограничено разпространение – например минало предварително време се среща в Корпуса само в художествени текстове и филмовите субтитри, докато сегашно време, макар с различна честота, не е ограничено само за даден тип текстове.

При художествените текстове почти поравно са разпространени сегашно време, минало несвършено време и формите за останалите времена – като сегашно време и минало несвършено време заемат по около 1/3 от всички срещания, а формите за всички

останалите времена заемат също 1/3 от броя на всички глаголни форми. Данните за честотата на употреба на темпоралните форми при художествените текстове от „Българско-английския паралелен корпус със съотнесени глаголни форми“ потвърждават виждането за характеристиката на художествените текстове (Васева 1989: 87) като текстове, при които естетическото въздействие се постига чрез динамично съотнасяне на действията във времето. Значително по-вариативното разпространение на формите на тези времена се дължи на факта, че имат по-голям спектър от транспозиционни употреби, които могат да имат по-широк семантичен обхват, докато по-рядко разпространените глаголни времена (бъдеще време в миналото бъдеще предварително време и бъдеще предварително време в миналото) имат строго определено значение и при тях транспозиционни употреби не са характерни в същата степен както при най-разпространените времена (сегашно време, минало свършено време, минало несвършено време и минало неопределено време). При филмовите субтитри се вижда, че обхватът на разпространение на темпоралните форми е еднакъв с този в художествените текстове, като разпределението се различава съществено при съотношението между формите за сегашно време и минало несвършено време. Този факт се дължи на характеристиката на субтитрите като тип текст, при който действията и състоянията се представят като протичащи в настоящето и по-рядко е необходимо темпорално съотнасяне. Съответно и комплексността на темпоралните отношения е по-опростена в сравнение с художествените текстове.

При публицистичните и научните текстове от представените данни в Таблица 8. се вижда, че приоритетно разпределение имат темпоралните форми в сегашно време и с малки колебания бъдеще време при публицистичните текстове и минало несвършено време при научните текстове. Това може да се обясни с типологичните характеристики на двата вида текстове – докато при публицистичните текстове се съобщават текущи събития и динамиката на темпоралните отношения е ориентирана предимно към настоящето, то при научните текстове действията се представят като актуални към момента на говорене или към определена референтна точка. При научните текстове рядко има съпоставително ситуиране на действията и състоянията по темпоралната ос едно спрямо друго. При публицистичните текстове е характерно представянето на текущо протичащи събития, като връзките с предходни събития се изграждат с допълнителни лексикални елементи и чрез употреба на съюзи.

Данните от гледна точка на разпределението на морфологичните категории,

влизащи в състава на хиперкатегорията време в български език, са представени в Таблица 9.

Морфологична категория		**Брой форми**	**%**
Време	Неразновременност	3580	69,62%
	Предходност	1311	25,50%
	Следходност	251	4,88%
Вид на действието	Нерезултативни времена	4694	91,29%
	Резултативни времена	448	8,71%
Таксис	Неотносителни времена	4216	82%
	Относителни времена	926	18%

Таблица 9. ***Разпределение на темпоралните морфологични категории в български език.***

От представените данни в Таблица 9. се вижда, че приоритетно разпределение имат немаркираните членове на съответните темпорални морфологични категории: неразновременност, нерезултативност и неотносителност. Тази особеност може да се обясни с факта, че немаркираните членове могат да изразяват както главното си значение – противоположно на значението на маркирания член, така и общото значение – когато няма информация за наличие или отсъствие на основното значение на маркирания член. Именно формалната и семантична немаркираност в тези случаи дава възможност в общото си значение да „не дава информация за наличие или отсъствие на положителната характеристика на маркирания член, а в определени условия може да изразява имплицитно характеристиката на маркирания член“ (Куцаров 2007: 35).

Обобщено разпределение на темпоралните морфологични категории, реализирани в глаголните форми в Корпуса по типове текстове, е представено в Таблица 10.

Морфологична категория		Художествена литература		Публицистични текстове		Научни текстове		Филмови субтитри	
		Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Време	Неразновременност	2107	63,1	821	88,4	96	96	504	74,3
	Предходност	1114	33,4	57	6,1	1	1,0	133	19,6
	Следходност	118	3,5	51	5,5	3	3	41	6
Вид на действието	Нерезултативни времена	3011	90,2	891	95,9	98	98	642	94,7
	Резултативни времена	328	9,8	38	4,1	2	2	36	5,3
Таксис	Неотносителни времена	2620	78,5	913	98,3	100	100	554	81,7
	Относителни времена	719	21,5	16	1,7	–	–	124	18,3

Таблица 10. ***Обобщено разпределение на темпоралните морфологични категории по типове текстове***

От Таблица 10. се вижда, че приоритетното разпределение на немаркираните членове на съответните морфологични категории се запазва в различните видове текстове. Най-голямо разнообразие от реализации на различни морфологични категории се наблюдава в художествените текстове, където типовите характеристики на съответните текстове предполагат по-голяма вариативност на изразните средства (Станева 2001: 158). В художествените текстове формите, изразяващи морфологичните категории вид на действието и таксис, са разпространени в значително по-голяма степен в сравнение с научните и публицистичните текстове. Това се дължи на факта, че в художествените текстове ситуирането на действията по темпоралната ос и подреждането им едно спрямо друго изисква прецизно изразяване на взаимоотношенията между значенията на различните морфологични категории. Както беше посочено, по-голямата част от глаголните форми, натоварени със значенията резултативност и относителност, служат за темпорален контекст в художествените текстове, за да изразят отношенията между отделните времена – те служат за времена ориентири, докато транспозиционните употреби на най-често срещаните глаголни форми (сегашно време, минало свършено време, минало несвършено време и бъдеще време) са с по-голяма вариативност и могат да обхванат по-голям спектър от значения. При научните и публицистичните текстове разнообразието от глаголни форми е силно редуцирано (при научните текстове в Корпуса няма относителни времена) в полза на немаркираните членове на морфологичните категории. При публицистичните текстове предходността и следходността като грамеми на морфологичната категория време са приблизително поравно разпределени, което се дължи на факта, че при този тип текстове е характерно опростеното съотнасяне на действията и състоянията чрез некомплексни изразни средства (Станева 2001: 132).

От гледна точка на посоката на превода на текстовете в Корпуса и корелацията между съотнесените глаголни форми в българските текстове има 5 срещания на други части на речта (без глаголи), които в английските текстове са преведени с глаголни форми. В 4 от случаите това са атрибутивно употребени причастия, като 1 от срещанията е деепричастие.

(32) Той не я видя, докато вратата, отваряйки се, привлече вниманието му….
While the door was opening, he noticed her.

При всички случаи на несъответствие текстовете представляват английски оригинал с български превод.

IV.3. Статистически обусловен анализ на данните за английските темпорални форми в Корпуса

Разпределението на глаголните форми за английски език е представено в Таблица 11.:

Време	Брой форми	%
Минало просто време	2008	43,71%
Сегашно просто време	1430	31,13%
Инфинитив	294	6,40%
Бъдеще просто време	264	5,75%
Минало перфектно време	159	3,46%
Минало продължително време	117	2,55%
Бъдеще просто време в миналото	103	2,24%
Сегашно перфектно време	96	2,09%
Сегашно продължително време	73	1,59%
Минало перфектно продължително време	15	0,33%
Бъдеще перфектно време в миналото	12	0,26%
Бъдеще перфектно време	11	0,24%
Сегашно перфектно продължително време	7	0,15%
Бъдеще продължително време в миналото	3	0,07%
Бъдеще продължително време	2	0,04%
Бъдеще перфектно продължително време	0	0%
Бъдеще перфектно продължително време в миналото	0	0%

Таблица 11. ***Разпределение на глаголните форми за английски език.***

От представените дистрибутивни данни за английски език в Таблица 11. се вижда, че приоритетното разпределение на определени глаголни форми се запазва и в английски. Изолираното разглеждане на разпределението на темпоралните форми както в български, така и в английски показва, че дадени темпорални форми имат значително по-висока фреквентност и по-голям обхват на разпространение в двата езика. Това наблюдение може да бъде тълкувано в две направления. Първо, този факт се появява като потвърждение на тезата (Маккарти и Уарингтън 1988: 429) за определени темпорални домейни на възприятие при разбирането и генерирането на езика и неговите

равница. Тази особеност на функционирането на езика в човешкото съзнание предполага предсказуемия набор от езикови средства (в това число и темпорални отношения) при конструирането на изказванията, а оттам и възприемането им в съзнанието. И второ, това наблюдение може да бъде интерпретирано за български и английски като проявление на тезата на Е. Бенвенист (Бенвенист 1971: 210), който основно за френски твърди, че темпоралните форми не могат да бъдат разглеждани нито изолирано в лексикален план, нито изолирано в граматически план. Ако се разглеждат сигнализаторите за темпорална ориентираност в текста, ще се види, че някои от тях (като дати, адвербиални фрази и уточняващи думи) са твърде редки. Техните проявления все пак не са случайни, а се подчиняват на определени контекстни условия. В подкрепа на тезата на Е. Бенвенист за същността на темпоралните форми (Бенвенист 1971) Х. Уайнрих (Уайнрих 1973: 62) изследва т. нар. „темпорални групи“ – тенденцията определени текстове да имат преобладаващи темпорални форми, а самите темпорални форми – „валентност“ помежду си. Това може да бъде обяснено с факта, че граматическото време има локативна служба в текста и няма директна връзка с реалното време извън рамките на даден текст. По този начин естеството на самия текст може да направи нефункционално използването на широк диапазон от темпорални форми (Фабиан 2014: 83). И двете представени схващания могат да бъдат подкрепени чрез дистрибутивните данни, представени за български и английски поотделно. И в английски, подобно на български, се вижда, че синтетичните глаголни форми заемат значителна част от броя на срещанията на глаголни форми в езиковите данни, включени в Корпуса, като комплексните глаголни форми имат твърде рядка честота на срещане или не се реализират. Това отново може да бъде обяснено с по-широкия обхват от значения при транспозиционната употреба на синтетичните глаголни форми и със стремежа към ограничаване на комплексните темпорални форми за сметка на лексикалния контекст. При аналитичните глаголни форми в английски съотношението токън : глаголна форма е средно 2,76.

Разпространението на глаголните форми за английски език по област на употреба е представено в Таблица 12.

	Художествена литература		Публицистични текстове		Научни текстове		Филмови субтитри	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Минало просто време	1640	56,32	91	10,68	19	20,88	244	37,60
Сегашно просто време	701	24,07	415	48,71	64	70,33	241	37,13
Инфинитив	136	4,67	97	11,38	5	5,49	56	8,63
Минало перфектно време	111	3,81	3	0,35	0	–	17	2,62
Минало продължително време	86	2,95	0	–	0	–	30	4,62
Бъдеще просто време	67	2,30	176	20,66	0	–	20	3,08
Сегашно продължително време	56	1,92	12	1,41	0	–	5	0,77
Бъдеще просто време в миналото	50	1,72	15	1,76	0	–	21	3,24
Сегашно перфектно време	45	1,55	40	4,69	3	3,30	8	1,23
Минало перфектно продължително време	12	0,41	0	–	0	–	3	0,46
Сегашно перфектно продължително време	4	0,14	3	0,35	0	–	0	–
Бъдеще продължително време	2	0,07	0	–	0	–	0	–
Бъдеще перфектно време в миналото	1	0,03	0	–	0	–	2	0,31
Бъдеще продължително време в миналото	1	0,03	0	–	0	–	1	0,15
Бъдеще перфектно време	0	–	0	–	0	–	1	0,15
Бъдеще перфектно продължително време	0	–	0	–	0	–	0	–
Бъдеще перфектно продължително време в миналото	0	–	0	–	0	–	0	–
Общо	***2912***	***63,39***	***852***	***18,55***	***91***	***1,98***	***649***	***14,13***

Таблица 12. ***Разпространение на глаголните форми за английски език по област на употреба.***

От представените статистически данни в Таблица 12. за темпоралните форми в английски личи, че разпространението им не е еднотипно за различните области на употреба. И в английски се наблюдава по-голямо разнообразие от темпорални форми в художествените текстове и филмовите субтитри, докато публицистичните и научните текстове се характеризират със значително по-малко разнообразие от възможни проявления на темпорални форми. При художествените текстове се реализира най-голям брой от различни глаголни форми като в английски двете най-разпространени времена при художествените текстове и филмовите субтитри заемат общо около 3/4 от общия брой на глаголните форми. И в английски при публицистичните и научните текстове се наблюдават еднотипни видове темпорални глаголни форми, което служи за доказателство, че различните типове текстове в различните езици имат съответните

„темпорални групи“, които могат да имат по-разнообразен обхват и взаимоотношения при художествените текстове и по-ограничен – при научни и публицистични текстове. Това се дължи на характеристиките на текстовете – тъй като публицистичните и научните текстове актуализират действията и състоянията основно като случващи се в настоящето и общовалидни, при тях не се проявява използването на широк спектър от темпорални глаголни форми, темпоралните отношения се предават чрез други езикови елементи (Фабиан 2014: 79), и обратно – при художествените текстове (и филмовите субтитри) темпоралните отношения са по-комплексни и се налага използването на времена ориентири, „спомагателни времена“ и „фонови времена“ за по-динамично представяне на взаимовръзките между отделните действия и състояния.

Направените наблюдения за употребата на темпоралните форми в български и английски поотделно са в голяма степен предопределени от факта, че анотираните текстове в Корпуса представляват преводни еквиваленти. От друга страна обаче, наблюденията потвърждават надеждността на двуезиковите корпуси при създаването на модели за превод от един език на друг.

Данните за разпределението на различните видове времена, влизащи в състава на хиперкатегорията време в английски език, са представени в Таблица 13.

Вид времена	***Брой форми***	***%***
Прости времена	3805	82,83%
Перфектни времена	278	6,05%
Продължителни времена	195	4,24%
Перфектни продължителни времена	22	0,48%
Инфинитив	294	6,40%

Таблица 13. ***Разпределение на темпоралните морфологични категории в английски език.***

От данните в Таблица 13. се вижда, че приоритетно разпределение в английски език имат формите на простите (неперфектни, непродължителни) времена, докато най-комплексните форми на перфектните продължителни времена имат най-ниска честота на разпространение. В английски език на второ място по разпространение са инфинитивните форми на глаголите. От семантична гледна точка инфинитивните форми изразяват действието само по себе си, без връзка с глаголното формоизменение (Ранкова и Иванова 1980: 154), от което следва, че в английски език необходимата граматична

информация за пълноценен превод е налична в лексикалния контекст, а инфинитивът представя самото действие.

Обобщено разпределението на видовете времена, реализирани в глаголните форми в английската част на Корпуса по типове текстове, е представено в Таблица 14.

	Художествена литература		**Публицистични текстове**		**Научни текстове**		**Филмови субтитри**	
Вид времена	***Брой форми***	***%***	***Брой форми***	***%***	***Брой форми***	***%***	***Брой форми***	***%***
Прости времена	2458	84,41	697	81,81	83	91,21	526	81,05
Перфектни времена	157	5,39	43	5,05	3	3,30	28	4,31
Продължителни времена	145	4,98	12	1,41	0	–	36	5,55
Перфектни продължителни времена	16	0,55	3	0,35	0	–	3	0,46
Инфинитив	136	4,67	97	11,38	5	5,49	56	8,63

Таблица 14. ***Обобщено разпределение на темпоралните морфологични категории в английски по типове текстове.***

От представените статистически данни в Таблица 14. се вижда, че различните видове темпорални категории имат значителен обхват на разпространение в различните типове текстове, като простите времена имат приоритетно разпределение и обхват във всички видове текстове. Изключение правят научните текстове, където не се срещат продължителни и перфектни продължителни времена, а перфектните времена имат ниска честота. Ако се приемат глаголните форми на простите (неперфектни, непродължителни) времена за немаркирани в английски, спрямо маркираните продължителни и перфектни времена, значително по-голямото им разпространение също може да бъде обяснено с факта, че немаркираните членове на граматическите категории могат да изразяват както главното си значение – противоположно на значението на маркирания член, така и общото значение – когато няма информация за наличие или отсъствие на основното значение на маркирания член. По този начин транспозиционните употреби на немаркираните членове са съществено по-разнообразни и по-разпространени от транспозиционните употреби на по-комплексните маркирани членове и имат по-висока честота.

От гледна точка на посоката на превода на текстовете в Корпуса и корелацията между съотнесените глаголни форми в английските текстове има 530 срещания на други части на речта (без глаголи), които в българските текстове са реализирани с глаголни форми. Броят на тези срещания е 372 при текстове с оригинал на български и превод на

английски, 96 при текстове с оригинал на английски и превод на български и 61 при текстове с оригинал на трети език и преводи на български и английски. Броят им е съществено по-голям от същите срещания за български, което предполага, че голяма част от граматикализираната информация в българските глаголни форми не се реализира на същото ниво при превод или се губи.

(33) Къде ли е? – помисли си първо той. – The first thought was about her: Where was she?

В следващата част ще бъде представен подробен съпоставителен анализ за поведението на морфологичните категории и граматикализираната информация при превод.

IV.4. Статистически обусловен анализ на превода на темпоралните форми от български на английски

В тази част ще бъдат представени поотделно преводните реализации на българските темпорални форми на английски език. Необходимо е да бъде направено уточнението, че при някои от времена, за които не са намерени примери за български и техните преводи на английски в „Българско-английския паралелен корпус със съотнесени изречения“ и интернет, единствените преводни реализации са само в примерите от „Българска граматика“. Също така поради ниската честота на срещане на тези глаголни форми в български език статистическите данни за преводните им реализации на английски, извлечени от Корпуса, не могат да бъдат еднозначно приети като представителни примери за поведението на българските форми и трансфера на граматичната информация при превод.

IV.4.1. Сегашно време

Глаголните форми в сегашно време от българските текстове в Корпуса са 2 437 или 47,39% от състава на всички глаголни форми. Обемът на формите за сегашно време и обхватът им спрямо различните типове текстове предполага адекватно представяне на особеностите на преводните реализации на тези форми. Формите на сегашно време по отношение на морфологичната категория време са неутрални и немаркирани и възможните транспозиционни употреби са най-много. На това се дължи и сравнително високата честота на срещане на тези форми в Корпуса. Разпределението на преводните реализации на формите за сегашно време в българските текстове на английски език, подредено по честота на срещане, е представено в Таблица 15.

	Общо		Художествена литература		Публицистични текстове		Научни текстове		Филмови субтитри	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Сегашно просто време	1346	55,41	643	53,41	412	53,23	65	67,01	226	64,02
Минало просто време	289	11,90	186	15,45	36	4,65	17	17,53	50	14,16
Други части на речта	274	11,28	177	14,70	70	9,04	7	7,22	20	5,67
Инфинитив	210	8,65	92	7,64	79	10,21	5	5,15	34	9,63
Бъдеще просто време	161	6,63	14	1,16	146	18,86	0	–	1	0,28
Сегашно продължително време	59	2,43	45	3,74	11	1,42	0	–	3	0,85
Сегашно перфектно време	29	1,19	9	0,75	14	1,81	3	3,09	3	0,85
Бъдеще просто време в миналото	23	0,95	10	0,83	5	0,65	0	–	7	1,98
Минало продължително време	21	0,86	13	1,08	0	–	0	–	8	2,27
Минало перфектно време	9	0,37	9	0,75	0	–	0	–	0	–
Сегашно перфектно продължително време	4	0,16	3	0,25	1	0,13	0	–	0	–
Минало перфектно продължително време	3	0,12	3	0,25	0	–	0	–	0	–
Бъдеще продължително време в миналото	1	0,04	0	–	0	–	0	–	1	0,28

Таблица 15. ***Разпределение на преводните реализации на английски език на формите за сегашно време в българските текстове.***

От представените статистически данни в Таблица 15. се вижда, че формите за сегашно време в български език имат разнообразни преводни реализации в английски – 11 темпорални глаголни форми, инфинитивни форми и превод с други части на речта. Въпреки разнообразието от възможни форми в английски фактически разнообразните реализации се наблюдават в художествените текстове и филмовите субтитри. В научните и публицистичните текстове процентното разпределение на формите за сегашно просто време се запазва сравнително постоянно спрямо художествените текстове и филмовите субтитри, но разнообразието от възможните реализации е съществено ограничено. Значително големият брой от възможни преводни реализации на формите за сегашно време в английски може да бъде обяснен с факта, че в български сегашно време притежава най-голям брой възможни транспозиционни употреби (Ницолова 2008: 274 – 276), някои от които е невъзможно да бъдат преведени с форми за сегашни времена в английски.

Употребата на сегашно време също така актуализира настоящето в контекста на текстовия дискурс, който може да бъде разновременен спрямо реалния момент на говорене – сегашно време може да се използва при описание на минали или бъдещи събития и събития с по-комплексна темпорална ориентираност. Също така в български

език граматическите категории се отличават с повишена степен на контекстна зависимост – ако в контекста е изразено съответното значение, в случая – темпорален признак, то маркираното с този признак време може да бъде заменено с неговия немаркиран корелат. Вижда се, че въпреки разнообразието от преводни реализации, доминираща честота имат формите на сегашно просто време в английски. В представения теоретичен анализ във Втора глава от дисертацията двете времена бяха разгледани като семантично най-близки в двата езика.

В рамките на Корпуса не се наблюдават преводни реализации на сегашно време в български език с форми за бъдеще продължително време, бъдеще перфектно време, бъдеще перфектно продължително време, бъдеще перфектно време в миналото и бъдеще перфектно продължително време в миналото в английски език. Разнообразието от преводни реализации на формите на сегашно време в английски предполага, че е възможно да се очаква в други случаи преводът на формите за сегашно време в български да се реализира чрез форми, които не са открити в рамките на Корпуса. Семантиката на формите за бъдеще продължително време, бъдеще перфектно време, бъдеще перфектно продължително време, бъдеще перфектно време в миналото и бъдеще перфектно продължително време в миналото в английски език се обединява около значението следходност на действие или резултат от действие спрямо определен ориентационен момент. Въпреки това от представените данни в Таблица 15. се вижда, че преводни еквиваленти с подобно значение са възможни в английски.

От гледна точка на морфологичните категории лице, число и род, характеризиращи субекта на глаголното действие или състояние, формите на сегашно време в български носят информация за граматичното лице и число на „вършителя“ на действието, а информация за граматичен род не се съдържа в глаголната форма. Възможните форми за сегашно време в български са 6. В рамките на Корпуса се срещат всички възможни глаголни форми за единствено и множествено число за първо, второ и трето лице. Преводните английски еквиваленти на формите на сегашно време в български са разнообразни по отношение на възможни форми, носещи информация за граматическо лице и число. Информацията за лице и число в български се реализира в съществено намален обем в английски на морфологично равнище, тъй като възможните форми на преводните еквиваленти носят информация за лице и число в ограничена степен, а информация за граматичен род не може да се пренесе в рамките на глаголната форма. Както беше посочено в представеното описание на особеностите на

формообразуването, английският език изисква експлицитен подлог чрез местоимение или именна фраза, посредством който се реализира информацията за граматическото лице и число. Възможните глаголни форми на преводните английски еквиваленти на формите на сегашно време и глаголните форми, които се срещат в рамките на Корпуса, са представени в Таблица 16.

	Брой възможни графични форми в английски	Граматична информация за лице, число	Брой реализирани форми
Сегашно просто време	2	3. лице, ед. число	2
Минало просто време	1	–	1
Бъдеще просто време	1	–	1
Сегашно продължително време	3	1. и 3. лице, ед. число	3
Сегашно перфектно време	2	3. лице, ед. число	2
Бъдеще просто време в миналото	1	–	1
Минало продължително време	2	1. и 3. лице, ед. число	2
Минало перфектно време	1	–	1
Сегашно перфектно продължително време	2	3. лице, ед. число	2
Минало перфектно продължително време	1	–	1
Бъдеще продължително време в миналото	1	–	1

Таблица 16. ***Съотношение между възможни и реализирани глаголни форми при превода на форми за сегашно време, носещи информация за лице и число в английски.***

Представените данни в Таблица 16. показват, че в рамките на преводните английски реализации на формите на сегашно време се срещат всички възможни форми, носещи граматична информация за лице и число в английски, въпреки ограничената глаголна парадигма в английски.

IV.4.2. Минало свършено време

Глаголните форми в минало свършено време от българските текстове в Корпуса са 1 311 или 25,50 % от състава на всички български форми. Броят на формите за минало свършено време съставлява около 1/4 от корпуса и 1/2 от броя на формите за сегашно време. Минало свършено време е немаркирано по отношение на категориите вид на действието и таксис и е маркирано с предходност в рамките на морфологичната категория време. Разпределението на преводните реализации на английски език на

формите за минало свършено време в българските текстове, подредено по честота на срещане, е представено в Таблица 17.

	Общо		Художествена литература		Публицистични текстове		Научни текстове		Филмови субтитри	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Минало просто време	1084	82,87	931	83,57	36	64,29	1	100	111	84,09
Други части на речта	90	6,88	87	7,81	1	1,79	0	–	2	1,52
Сегашно просто време	35	2,68	28	2,51	1	1,79	0	–	6	4,55
Инфинитив	30	2,29	18	1,62	5	8,93	0	–	7	5,30
Минало перфектно време	21	1,61	18	1,62	1	1,79	0	–	2	1,52
Сегашно перфектно време	21	1,61	10	0,90	10	17,86	0	–	1	0,76
Минало продължително време	15	1,15	13	1,17	0	–	0	–	2	1,52
Бъдеще просто време в миналото	6	0,46	6	0,54	0	–	0	–	0	–
Сегашно продължително време	3	0,23	2	0,18	1	1,79	0	–	0	–
Бъдеще просто време	1	0,08	1	0,09	0	–	0	–	0	–
Минало перфектно продължително време	1	0,08	0	–	0	–	0	–	1	0,76
Сегашно перфектно продължително време	1	0,08	0	–	1	1,79	0	–	0	–

Таблица 17. ***Разпределение на преводните реализации на английски език на формите за минало свършено време в българските текстове.***

От представените данни в Таблица 17. се вижда, че преводните реализации на формите на минало свършено време в английски също са разнообразни. С най-висока честота са формите за минало просто време в английски, като след тях се нареждат преводните реализации с други части на речта. И тук при художествените текстове и филмовите субтитри разнообразието е по-голямо в сравнение с научните и публицистичните текстове. При научните текстове в целия Корпус има само една форма в минало свършено време в български, която на английски е преведена с форма за минало просто време. Поради малкия обем на срещанията на подобни форми в научните текстове, не може еднозначно да се заключи, че в този тип текстове преводните реализации на формите за минало свършено време ще бъдат еднообразни. Формите за минало свършено време имат приблизително същото разнообразие от преводни реализации в английски както формите за сегашно време, но очевидна разлика е преимуществено високата честота на формите за сегашно просто време в английски, използвани при превод на формите за минало свършено време. С факта, че минало свършено време изразява значенията на маркиран член единствено на морфологичната категория време (по отношение на категорията вид на действието минало свършено

време е немаркирано със значение резултативност, а по отношение на категорията таксис минало време е немаркирано със значение относителност) може да се обяснят сравнително разнообразните преводни реализации в английски. Като едно от „същинските“ времена минало свършено време представя събитията като минали спрямо настоящето, описвано в текста, и основната му роля е на „време-ориентир“ за относителните времена. Както беше посочено, правилото за съгласуване на времената в английски предполага, че ако глаголната форма в главното изречение е в минало просто време, другите глаголни форми в подчинените изречения се съгласуват по определен начин с глаголната форма в главното изречение, докато в български е възможно глаголните форми в подчинените изречения да бъдат в същото или в различно време спрямо глаголната форма в главното изречение.

В рамките на Корпуса не се срещат преводни английски реализации на формите на минало свършено време с форми на бъдеще продължително време, бъдеще перфектно време, бъдеще перфектно продължително време, бъдеще продължително време в миналото, бъдеще перфектно време в миналото и бъдеще перфектно продължително време в миналото. Семантиката на изброените глаголни форми се обединява около значението следходност на резултат от действие или действие спрямо определен ориентационен момент. Значението на изброените форми и данните от Корпуса могат да послужат за заключението, че преводни реализации на форми за минало свършено време с форми на изброените по-горе времена са малко вероятни.

Възможните форми за минало свършено време в български език, носещи граматична информация за лице и число, са 5. При тези форми има съвпадение на формите за второ и трето лице единствено число, а информация за граматичен род не се носи от глаголните форми. Възможните глаголни форми на преводните английски еквиваленти на формите на минало свършено време и глаголните форми, които се срещат в рамките на Корпуса, са представени в Таблица 18.

	Брой възможни графични форми в английски	Граматична информация за лице, число	Брой реализирани форми
Минало просто време	1	–	1
Сегашно просто време	2	3. лице, ед. число	2
Минало перфектно време	1	–	1
Сегашно перфектно време	2	3. лице, ед. число	2
Минало продължително време	2	1. и 3. лице, ед. число	2
Бъдеще просто време в миналото	1	–	1

	Брой възможни графични форми в английски	Граматична информация за лице, число	Брой реализирани форми
Сегашно продължително време	3	1. и 3. лице, ед. число	3
Бъдеще просто време	1	–	1
Минало перфектно продължително време	1	–	1
Сегашно перфектно продължително време	2	3. лице, ед. число	2

Таблица 18. ***Съотношение между възможни и реализирани глаголни форми при минало свършено време, носещи информация за лице и число в английски.***

Данните, представени в Таблица 18, показват, че всички възможни форми от парадигмите за съответните времена се срещат в рамките на преводните английски реализации на формите на минало свършено време в българските текстове.

IV.4.3. Минало несвършено време

Глаголните форми в минало несвършено време в българските текстове в Корпуса са 710 или 13,81% от състава на всички глаголни форми. Формите за минало несвършено време съставляват приблизително 1/3 от броя на формите за сегашно време и приблизително 1/2 от броя на формите за минало свършено време. По отношение на мястото на минало несвършено време в хиперкатегорията време в българския език, то може да бъде описано като относително сегашно време (Куцаров 2007: 256) – в общото си темпорално значение изразява липса на данни за разновременност – неразновременност на действие с изказване в минало време, а в главното – едновременност на действие с изказване в минало време (Куцаров 2007: 258). Разпределението на преводните реализации на формите за минало несвършено време в българските текстове на английски език, подредено по честота на срещане, е представено в Таблица 19.

	Общо		Художествена литература		Публицистични текстове		Филмови субтитри	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Минало просто време	444	62,62	362	62,63	9	75,00	66	61,68
Други части на речта	120	16,93	119	20,59	0	–	1	0,93
Минало продължително време	64	9,03	46	7,96	0	–	17	15,89
Инфинитив	34	4,80	19	3,29	1	8,33	14	13,08
Сегашно просто време	14	1,97	11	1,90	1	8,33	1	0,93
Минало перфектно време	10	1,41	6	1,04	0	–	3	2,80

	Общо		Художествена литература		Публицистични текстове		Филмови субтитри	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Бъдеще просто време в миналото	8	1,13	3	0,52	0	–	3	2,80
Минало перфектно продължително време	7	0,99	5	0,87	0	–	2	1,87
Сегашно продължително време	4	0,56	4	0,69	0	–	0	–
Сегашно перфектно време	3	0,42	2	0,35	1	8,33	0	–
Бъдеще просто време	1	0,14	1	0,17	0	–	0	–

Таблица 19. ***Разпределение на преводните реализации на английски език на формите за минало несвършено време в българските текстове.***

От Таблица 19. се вижда, че форми за минало несвършено време не се срещат в научни текстове в целия Корпус. И при формите на минало несвършено време най-разпространените реализации в английски език – формите на минало просто време, съставляват над половината от общия брой на реализациите. Възможните форми за преводни съответствия при минало несвършено време също са разнообразни в художествените текстове и филмовите субтитри, докато в публицистичните текстове, поради сравнително малкия брой примери не може еднозначно да се твърди, че възможното разнообразие от преводни съответствия е застъпено в Корпуса. Семантиката на минало несвършено време в български език има същата формално-семантична интерпретация като минало продължително време в английски език – в български минало несвършено време изразява действието като протичащо преди момента на говорене, а интервалът на действието включва в себе си миналия интервал на референтност, такава е семантичната интерпретация и за минало продължително време в английски.

От представените данни в Таблица 19. се вижда обаче, че не може да се намери еднозначно съответствие между употребата на минало несвършено време в български и минало продължително време в английски. Най-висока честота при превода на форми на минало несвършено време в български имат формите за минало просто време в английски, докато реализациите с форми за минало продължително време се нареждат на трето място след превода с други части на речта.

В рамките на Корпуса не се срещат преводни реализации на формите за минало несвършено време в български с форми за бъдеще продължително време, бъдеще перфектно време, сегашно перфектно продължително време, бъдеще перфектно продължително време, бъдеще продължително време в миналото, бъдеще перфектно

време в миналото и бъдеще перфектно продължително време в миналото в английски език. Глаголните форми в английски, които не се срещат като преводни реализации на формите за минало несвършено време, съвпадат с изброените глаголни форми, които не се срещат като преводни реализации на формите за минало свършено време, с изключение на формите за сегашно перфектно продължително време. При формите за минало несвършено време семантиката на изброените по-горе глаголни форми заедно с представените статистически данни дават основание да се направи заключението, че преводни реализации с изброените по-горе глаголни форми в английски на формите за минало несвършено време в български не са вероятни.

Възможните глаголни форми при формите за минало несвършено време са 5. Подобно на формите за минало свършено време и при минало несвършено време има съвпадение на формите за второ и трето лице, единствено число, а информация за граматичен род не се носи от глаголните форми. Възможните глаголни форми на преводните английски еквиваленти на формите на минало несвършено време и глаголните форми, които се срещат в рамките на Корпуса, са представени в Таблица 20.

	Брой възможни графични форми в английски	Граматична информация за лице, число	Брой реализирани форми
Минало просто време	1	–	1
Минало продължително време	2	1. и 3. лице, ед. число	2
Сегашно просто време	2	3. лице, ед. число	2
Минало перфектно време	1	–	1
Бъдеще просто време в миналото	1	–	1
Минало перфектно продължително време	1	–	1
Сегашно продължително време	3	1. и 3. лице, ед. число	1
Сегашно перфектно време	2	3. лице, ед. число	1
Бъдеще просто време	1	–	1

Таблица 20. ***Съотношение между възможни и реализирани глаголни форми при минало несвършено време, носещи информация за лице и число в английски.***

Данните, представени в Таблица 20., показват, че в Корпуса преводните английски еквиваленти на минало несвършено време в български имат по-ограничено разпространение в сравнение с формите за минало свършено време и сегашно време. Разбира се, в тези случаи преводът с глаголни форми, които не се срещат в Корпуса, е възможен. Като преводни еквиваленти на форми за минало несвършено време се срещат форми за минало просто време, минало продължително време, сегашно просто време, минало перфектно време, бъдеще просто време в миналото, минало перфектно

продължително време, сегашно продължително време, сегашно перфектно време и бъдеще просто време, които не носят информация за лице и число.

IV.4.4. Минало неопределено време

Глаголните форми в минало неопределено време от българските текстове в Корпуса са 269 или 5,23% от състава на всички български форми. Формите за минало неопределено време съставляват приблизително 1/3 от броя на формите за минало несвършено време. Мястото на минало неопределено време в хиперкатегорията време в българския език е сред резултативните времена. Минало неопределено време означава в общото си значение липса на данни за разновременност или едновременност на резултат от действие с изказването, а в главното значение – едновременност на резултат от действието с изказването (Куцаров 2007: 252). По отношение на морфологичната категория време минало неопределено време е маркирано с признака неразновременност, а по отношение на категорията вид на действието – с признака резултативност. Разпределението на преводните реализации на формите за минало неопределено време в българските текстове на английски език, подредено по честота на срещане, е представено в Таблица 21.

	Общо		Художествена литература		Публицистични текстове		Научни текстове		Филмови субтитри	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Минало просто време	134	50	113	53,81	8	23,53	1	50	12	54,55
Сегашно перфектно време	42	15,67	23	10,95	15	44,12	0	–	4	18,18
Минало перфектно време	29	10,82	26	12,38	2	5,88	0	–	1	4,55
Други части на речта	17	6,34	16	7,62	1	2,94	0	–	0	–
Сегашно просто време	14	5,22	11	5,24	1	2,94	1	50	1	4,55
Инфинитив	8	2,99	5	2,38	3	8,82	0	–	0	–
Минало продължително време	8	2,99	8	3,81	0	–	0	–	0	–
Бъдеще просто време в миналото	11	4,10	5	2,38	3	8,82	0	–	3	13,64
Сегашно перфектно продължително време	2	0,75	1	0,48	1	2,94	0	–	0	–
Бъдеще перфектно време в миналото	1	0,37	0	–	0	–	0	–	1	4,55
Минало перфектно продължително време	1	0,37	1	0,48	0	–	0	–	0	–
Сегашно продължително време	1	0,37	1	0,48	0	–	0	–	0	–

Таблица 21. ***Разпределение на преводните реализации на формите за минало неопределено време в българските текстове на английски език.***

От представените данни се вижда, че формите за минало неопределено време са с по-ниска честота в български език спрямо сегашно време, минало свършено време и минало несвършено време. Статистическите данни, извлечени от Корпуса, показват нееднозначно поведение при превод на английски език. Таблица 21. показва, че с най-голямо разпределение са преводните реализации с минало просто време в английски език. На второ място по честота са формите за сегашно перфектно време, които бяха представени във Втора глава като формален семантичен корелат на формите за минало неопределено време в английски. В научните текстове формите за минало неопределено форми са много редки, а в публицистичните текстове най-разпространен е преводът с форми за сегашно перфектно време. Колебанието при най-високо честотните реализации на формите за минало неопределено време между формите на минало просто и сегашно перфектно време в английски могат да се обяснят с това, че в езиковата практика формите за минало неопределено време в български могат да служат за темпорален контекст – изразяват действия или резултати от действия, чийто интервал е предходен спрямо интервала на референтност, който включва в себе си момента на говорене. Изразяването на значение за вече приключило действие в английски се предава чрез подчинени изречения, които са зависими от правилото за съгласуване на времената в английски – приключилите действия в подчинените изречения се предават чрез глаголни форми в минало перфектно време (Филипова 2002: 250).

Формите на времената в английски, които не се срещат в Корпуса като преводни съответствия на формите на минало неопределено в български, са бъдеще просто време, бъдеще продължително време, бъдеще перфектно време, бъдеще перфектно продължително време, бъдеще продължително време в миналото и бъдеще перфектно продължително време в миналото. Представените данни показват, че не е вероятно форми на бъдещи времена в английски да се използват като преводни съответствия на форми на минало неопределено време в български. Изключение правят бъдеще просто време в миналото и бъдеще перфектно време в миналото поради факта, че те се използват в подчинени условни изречения, където сигнализират за възможно приключило действие.

Формите на минало неопределено време в български са аналитични глаголни форми, образувани с минало свършено деятелно причастие. Аналитичните глаголни форми от този тип носят информация за граматичния род на вършителя на глаголното действие посредством формите на причастието. При минало неопределено време

парадигмата от възможни глаголни форми включва 12 члена. В рамките на езиковите данни от „Българско-английския паралелен корпус със съотнесени глаголни форми“ се срещат 10 глаголни форми на минало неопределено време, носещи информация за лице, число и род, съответно за 1. лице, единствено число, мъжки и женски род, 2. лице, единствено число, мъжки и женски род, 3. лице, единствено число, мъжки, женски и среден род, 1. лице, множествено число, 2. лице, множествено число и 3. лице, множествено число. Формите за 1. и 2. лице, единствено число, среден род са редки, но потенциално възможни, тъй като посредством 1. и 2. лице, единствено число се сигнализира за участници в комуникативния акт, които са одушевени, а среден род в множество от случаите се свързва с неодушевен денотат. Възможните глаголни форми на преводните английски еквиваленти на формите на минало неопределено време и глаголните форми, които се срещат в рамките на Корпуса, са представени в Таблица 22.

	Брой възможни графични форми в английски	Граматична информация за лице, число	Брой реализирани форми
Минало просто време	1	–	1
Сегашно перфектно време	2	3. лице, ед. число	2
Минало перфектно време	1	–	1
Сегашно просто време	2	3. лице, ед. число	2
Минало продължително време	2	1. и 3. лице, ед. число	2
Бъдеще просто време в миналото	1	–	1
Сегашно перфектно продължително време	2	3. лице, ед. число	2
Бъдеще перфектно време в миналото	1	–	1
Минало перфектно продължително време	1	–	1
Сегашно продължително време	3	1. и 3. лице, ед. число	1

Таблица 22. ***Съотношение между възможни и реализирани глаголни форми при минало неопределено време, носещи информация за лице и число в английски.***

От представените данни в Таблица 22. се вижда, че формите за съответните времена като преводни съответствия в английски на минало неопределено време са ограничени в сравнение с формите за сегашно време, минало свършено време и минало несвършено време. При времената с ниска честота на срещане, които имат няколко възможни глаголни форми, маркирани с информация за лице и число (сегашно продължително време) се срещат само немаркираните с информация за лице и число форми. При сегашно перфектно продължително време се срещат и двете възможни глаголни форми (3. лице, единствено число и немаркираната форма). Граматичната информация за род, която се носи от формите за минало неопределено време в български,

не се реализира в рамките на глаголните форми в английски. Както беше посочено, английският изисква задължително запълнена подложна позиция с местоимение или именна фраза, в рамките на които информацията за граматичния род може да се реализира, ако е изразена с местоимение в 3. лице, единствено число.

IV.4.5. Бъдеще време

Глаголните форми в бъдеще време от българските текстове в Корпуса са 189 или 3,68% от състава на всички български форми. Формите за бъдеще време съставляват приблизително 1/2 от броя на формите за минало неопределено време. Мястото на бъдеще време в хиперкатегорията време в български език е при нерезултативните и неотносителните времена, а от гледна точка на морфологичната категория време бъдеще време е маркирано с признака разновременност и изразява следходност на действие или резултат от действие спрямо момента на говоренето. Разпределението на преводните реализации на формите за бъдеще време от българските текстове на английски език, подредено по честота на срещане, е представено в Таблица 23.

	Общо		Художествена литература		Публицистични текстове		Научни текстове		Филмови субтитри	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Бъдеще просто време	102	54,26	52	54,74	29	56,86	1	33,33	19	50
Бъдеще просто време в миналото	23	12,23	10	10,53	5	9,80	0	–	8	21,05
Сегашно просто време	17	9,04	9	9,47	2	3,92	1	33,33	5	13,16
Други части на речта	15	7,98	8	8,42	5	9,80	1	33,33	1	2,63
Инфинитив	12	6,38	2	2,11	9	17,65	0	–	1	2,63
Сегашно продължително време	6	3,19	4	4,21	0	–	0	–	2	5,26
Минало просто време	6	3,19	5	5,26	1	1,96	0	–	0	–
Минало продължително време	5	2,66	3	3,16	0	–	0	–	2	5,26
Бъдеще продължително време	2	1,06	2	2,11	0	–	0	–	0	–

Таблица 23. ***Разпределение на преводните реализации на формите за бъдеще време в българските текстове на английски език.***

От представените данни за разпределението на преводните реализации на формите на бъдеще време в Таблица 23. се вижда, че приоритетно разпределение имат формите за бъдеще просто време. Разнообразието от преводни реализации е сравнително по-ограничено от реализациите на глаголните форми за сегашно време, минало свършено време, минало несвършено време и минало неопределено време. Преводните

реализации са сравнително ограничени при научните текстове, докато максимално разнообразие от форми се среща при художествените текстове. Формалният семантичен корелат на българското бъдеще време в английски език е бъдеще просто време, чиито форми са с най-висока честота на срещане при преводните реализации в английски. Въпреки сравнително малкия брой срещания на форми за бъдеще време в „Българско-английския паралелен корпус със съотнесени глаголни форми“, данните, представени в Таблица 23., удостоверяват, че при разпределението на преводните реализации се показва еднообразно предпочитание при превода на българското бъдеще време с английското бъдеще просто време.

В „Българско-английския паралелен корпус със съотнесени глаголни форми“ не се срещат преводни реализации на форми за бъдеще време в български с форми за сегашно перфектно време, минало перфектно време, бъдеще перфектно време, сегашно перфектно продължително време, минало перфектно продължително време, бъдеще перфектно продължително време, бъдеще продължително време в миналото, бъдеще перфектно време в миналото и бъдеще перфектно продължително време в миналото на английски. Изброените глаголни форми включват всички перфектни и перфектни продължителни времена, както и бъдеще продължително време в миналото. Семантиката на бъдеще време в български изразява следходност на действие или резултат от действие спрямо момента на говорене. Семантиката на изброените по-горе форми на времена, които не се срещат като преводни еквиваленти в английски на формите за бъдеще време в български, изразява наличие на резултат от приключило или протичащо действие към определен ориентационен момент. Може да се направи заключението, че преводни еквиваленти с форми на други времена, които не са регистрирани в Корпуса, са възможни, но малко вероятни.

Формите за бъдеще време в български са 6. аналитични форми, които се образуват със спомагателния глагол *ща*. Формите за бъдеще време носят информация за граматичното лице и число на глаголния субект, но не носят информация за граматичен род. В езиковите данни от Корпуса се срещат всички възможни форми за бъдеще време в български. Възможните глаголни форми на преводните английски еквиваленти на формите на минало неопределено време и глаголните форми, които се срещат в рамките на Корпуса, са представени в Таблица 24.

	Брой възможни графични форми в английски	Граматична информация за лице, число	Брой реализирани графични форми
Бъдеще просто време	1	–	1
Бъдеще просто време в миналото	1	–	1
Сегашно просто време	2	3. лице, ед. число	2
Сегашно продължително време	3	1. и 3. лице, ед. число	2
Минало просто време	1	–	1
Минало продължително време	2	1. и 3. лице, ед. число	1
Бъдеще продължително време	1	–	1

Таблица 24. ***Съотношение между възможни и реализирани глаголни форми при бъдеще време, носещи информация за лице и число в английски.***

Данните от Таблица 24. показват ограниченото разнообразие от възможни форми при преводните съответствия в английски на формите на бъдеще време в български. При времената, при които са възможни повече от 1 форма в парадигмата, в Корпуса се наблюдава следната особеност – при сегашно продължително време се срещат само форми за 3. лице, единствено число и немаркираната с информация за лице и число форма, а при минало продължително време се среща само немаркираната за лице и число форма.

IV.4.6. Минало предварително време

Глаголните форми в минало предварително време от българските текстове в Корпуса са 164 или 3,19% от състава на всички български форми. Формите за минало предварително време почти съвпадат по брой с формите за бъдеще време. Мястото на минало предварително време в хиперкатегорията време в българския език е при резултативните и относителни времена. Общото значение на минало предварително време означава липса на данни за едновременност или разновременност на резултат от действие с изказване в минало време, а в главното – едновременност на резултат от действие с изказване в минало време (Куцаров 2007: 260). По отношение на морфологичната категория време то е маркирано с признака неразновременност (едновременност), а по отношение на категорията таксис – с признака относителност. Разпределението на преводните реализации на формите за минало предварително време от българските текстове на английски език, подредено по честота на срещане, е представено в Таблица 25.

	Общо		Художествена литература		Публицистични текстове		Филмови субтитри	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Минало перфектно време	90	54,88	52	44,07	0	0	11	78,57
Минало просто време	44	26,83	39	33,05	4	100	1	7,14
Други части на речта	13	7,93	12	10,17	0	0	1	7,14
Минало продължително време	4	2,44	4	3,39	0	0	0	–
Бъдеще перфектно време в миналото	3	1,83	1	0,85	0	0	1	7,14
Бъдеще просто време в миналото	3	1,83	3	2,54	0	0	0	–
Минало перфектно продължително време	3	1,83	3	2,54	0	0	0	–
Сегашно просто време	3	1,83	3	2,54	0	0	0	–
Сегашно перфектно време	1	0,61	1	0,85	0	0	0	–

Таблица 25. ***Разпределение на преводните реализации на английски език на формите за минало предварително време в българските текстове.***

От статистическите данни за разпределението на преводните реализации на формите за минало предварително време в български се вижда, че с най-висока честота са формите на минало перфектно време в английски. Изключение правят преводните реализации в публицистичните текстове, при които всички 4 срещания на форми за минало предварително време са реализирани с форми за минало просто време в английски. В научните текстове не се срещат форми за минало предварително време. Формалният семантичен корелат на формите за минало предварително време в български език според формално-семантичното описание във Втора глава е минало перфектно време в английски език. Както се вижда от Таблица 25., преимуществено разпределение имат именно преводните реализации с форми за минало перфектно време, на второ място по честота се нареждат формите за минало просто време, като в публицистичните текстове те представляват единствените преводни реализации. Минало предварително време изразява приключило действие или наличен резултат от действие преди определен минал интервал на референтност, а семантиката на двете най-разпространени преводни реализации в английски е следната: формите за минало перфектно време означават резултат от действие спрямо минал момент (моментът на действието е предхождащ спрямо минал момент на референтност), докато формите за минало просто време означават приключило действие в миналото (моментът на действието съвпада с минал момент на референтност).

В езиковите данни от „Българско-английския паралелен корпус със съотнесени глаголни форми“ не се срещат преводни реализации на форми за минало предварително

време със следните времена в английски: бъдеще просто време, сегашно продължително време, бъдеще продължително време, бъдеще перфектно време, сегашно перфектно продължително време, бъдеще перфектно продължително време, бъдеще продължително време в миналото и бъдеще перфектно продължително време в миналото. Изброените времена включват формите на всички бъдещи времена в английски и формите на две бъдещи времена в миналото. Предвид семантиката на минало предварително време и изброените по-горе времена, може да се твърди, че преводни реализации с форми за бъдещи времена в английски на форми за минало предварително време в български не са възможни.

Както при минало неопределено време, формите на минало предварително време в български спадат към аналитичните глаголни форми, които се образуват с минало свършено деятелно причастие. При тези форми причастието носи допълнителна информация за граматичния род на глаголния субект. Възможните форми при минало предварително време в български са 9, като при тях формите за второ и трето лице единствено число съвпадат. В Корпуса се срещат 6 форми за минало предварително време – 1. лице, единствено число, мъжки и женски род, 2./3. лице, единствено число, мъжки и женски род, 1. лице, множествено число и 3. лице множествено число. Тук отново трябва да бъде обърнато внимание, че форми за среден род при 1. и 2. лице, единствено число са възможни, но малко вероятни, поради факта, че 1. и 2. лице, единствено число означават участниците в комуникативния акт, които са одушевени. Възможните глаголни форми на преводните английски еквиваленти на формите на минало неопределено време и глаголните форми, които се срещат в рамките на Корпуса, са представени в Таблица 26.

	Брой възможни графични форми в английски	Граматична информация за лице, число	Брой реализирани форми
Минало перфектно време	1	–	1
Минало просто време	1	–	1
Минало продължително време	2	1. и 3. лице, ед. число	1 (маркирана)
Бъдеще перфектно време в миналото	1	–	1
Бъдеще просто време в миналото	1	–	1
Минало перфектно продължително време	1	–	1
Сегашно просто време	2	3. лице, ед. число	1 (немаркирана)
Сегашно перфектно време	2	3. лице, ед. число	1 (немаркирана)

Таблица 26. ***Съотношение между възможни и реализирани глаголни форми при минало предварително време, носещи информация за лице и число в английски.***

От Таблица 26. се вижда, че граматичната информация за лице, род и число на формите за минало предварително време в български не се пренася в рамките на съответните глаголни форми в английски. В английските преводни съответствия на минало предварително време, при чиято парадигма са възможни повече от 1 форма – минало продължително време, сегашно просто време и сегашно перфектно време, се среща само немаркираната форма. Може да се твърди, че ниската честота на срещане на формите за минало предварително време и възможните преводни съответствия в английски силно редуцират трансфера на граматичните значения за лице и число на глаголния субект. Значението на категориите лице и число в български се изразява лексикално и синтактично от местоименния подлог в английски.

IV.4.7. Бъдеще време в миналото

Глаголните форми в бъдеще време в миналото от българските текстове в Корпуса са 47 или 0,91% от състава на всички български форми. Броят на формите за бъдеще време в миналото представлява 1/3 от броя на формите за минало предварително време. Бъдеще време в миналото в рамките на хиперкатегорията време в български език изразява разновременност, нерезултативност и относителност – формите му предават значение за следходност на действието спрямо минал ориентационен момент (изказване в минало време).

Разпределението на преводните реализации на формите за бъдеще време в миналото в българските текстове на английски език, подредено по честота на срещане, е представено в Таблица 27.

	Общо		Художествена литература		Филмови субтитри	
	Брой форми	%	Брой форми	%	Брой форми	%
Бъдеще просто време в миналото	29	63,04	15	62,50	0	–
Бъдеще продължително време в миналото	4	8,70	1	4,17	0	–
Бъдеще перфектно време в миналото	3	6,52	0	–	0	–
Минало просто време	3	6,52	3	12,50	0	–
Други части на речта	3	6,52	3	12,50	0	–
Сегашно просто време	2	4,35	1	4,17	1	50
Минало продължително време	1	2,17	1	4,17	0	–

Таблица 27. ***Разпределение на преводните реализации на формите бъдеще време в миналото в българските текстове на английски език.***

От статистическите данни в Таблица 27. се вижда, че въпреки малкия брой срещания на форми за бъдеще време в миналото с най-голяма честота са преводните реализации с форми за бъдеще просто време в миналото в английски език. Необходимо е да направим уточнението, че около 50% от формите за бъдеще време в миналото за български се срещат само в примерните текстове от специално конструирани изречения, събрани от „Българска граматика“. Тези данни не са представени в таблицата по-горе. При тях формите на бъдеще време в миналото в български са реализирани изцяло чрез форми за бъдеще просто време в миналото в английски. Формите на бъдеще просто време в български са сред групата на формите на времената с много ниска честота на срещане в „Българско-английския паралелен корпус със съотнесени глаголни форми“ – под 1%. От данните за приоритетното разпределение на преводни реализации с определени темпорални форми при времената с ниска честота на срещане в Корпуса може да се направи заключение, че те имат твърде специфична семантика, която се предава чрез ограничен брой форми при превод. От данните в Корпуса може да се твърди, че на значението на формите за бъдеще време в миналото в български отговарят формите за бъдеще просто време в миналото в английски.

В езиковите данни, включени в „Българско-английския паралелен корпус със съотнесени глаголни форми“, не се срещат преводни реализации на английски на формите на бъдеще време в миналото в български с форми на следните времена: бъдеще просто време, сегашно продължително време, бъдеще продължително време, сегашно перфектно време, минало перфектно време, сегашно перфектно продължително време, минало перфектно продължително време, бъдеще перфектно продължително време и бъдеще перфектно продължително време в миналото. Изброените времена, които не се срещат като преводни съответствия на бъдеще време в миналото, в по-голямата си част представляват минали и сегашни времена. От представените статистически данни и на базата на формалното описание на семантиката на бъдеще време в миналото може да се направи заключение, че преводните му реализации в английски са предимно с форми за бъдещи времена в миналото.

Формите за бъдеще време в миналото на български са аналитични и носят информация за лице и число на глаголния субект. Възможните форми за бъдеще време в миналото са 6, а в рамките на Корпуса се срещат 5 от тях – форми за 1., 2. и 3. лице, единствено число и форми за 1. и 3. лице, множествено число. Възможните глаголни

форми на преводните английски еквиваленти на формите на минало неопределено време и глаголните форми, които се срещат в рамките на Корпуса, са представени в Таблица 28.

	Брой възможни графични форми в английски	Граматична информация за лице, число	Брой реализирани форми
Бъдеще просто време в миналото	1	–	1
Бъдеще продължително време в миналото	1	–	1
Бъдеще перфектно време в миналото	1	–	1
Минало просто време	1	–	1
Сегашно просто време	2	3. лице, ед. число	1 (маркирана)
Минало продължително време	2	1. и 3. лице, ед. число	1 (немаркирана)
Бъдеще перфектно време	1	–	1

Таблица 28. ***Съотношение между възможни и реализирани глаголни форми при бъдеще време в миналото, носещи информация за лице и число в английски.***

Данните от Таблица 28. показват, че възможните форми за преводните реализации в ограничена степен могат да носят информация за граматично лице и число, а тези времена, които имат такива форми в парадигмата си, са сред преводните реализации с ниска честота. От реализираните форми се среща само форма (при сегашно време), която носи информация за 3. лице, единствено число, останалите форми не изразяват лице и число. От формите за минало продължително време се срещат само немаркираните форми.. Информацията за лице и число се реализира на лексикално и синтактично равнище в рамките на изречението – чрез изразяване на субекта.

IV.4.8. Бъдеще предварително време и бъдеще предварително време в миналото

Причината да се представят статистическите данни за преводните реализации на формите за бъдеще предварително време и бъдеще предварително време в миналото заедно е, че формите за тези две времена практически се срещат много рядко в „Българско-английския паралелен корпус със съотнесени глаголни форми“. Всички срещания на форми за тези две времена са единствено от „Българска граматика“. Глаголните форми в бъдеще предварително време от българските текстове в Корпуса са 10 или 0,19% от състава на всички български форми, а формите за бъдеще предварително време в миналото са 5 или 0,1% от състава на всички български форми. Форми за тези две времена не се срещат в езиковите данни от Българско-английския паралелен корпус

със съотнесени изречения и в езиковите данни, събрани от интернет. Преводните реализации, извлечени от примери, са представени в Таблица 29.

Темпорални форми в български	Темпорални форми в английски	Брой срещания в примерите
Бъдеще предварително време	Бъдеще перфектно време	10
Бъдеще предварително време в миналото	Бъдеще перфектно време в миналото	5

Таблица 29. ***Разпределение на преводните реализации на формите бъдеще предварително време и бъдеще предварително време в миналото в българските текстове на английски език.***

Преводните реализации на формите за бъдеще предварително време и бъдеще предварително време в миналото в примерите от „Българска граматика“ са представени само чрез формите съответно на бъдеще перфектно време и бъдеще перфектно време в миналото. Въпреки това, поради нулевата честота на срещане в останалите текстове не може еднозначно да се определят преводните реализации на тези две времена от български на английски език. Честотата на срещане на формите на бъдеще предварително време и бъдеще предварително време в миналото в български представлява доказателство на тезата, че поради сложността на своето значение бъдеще предварително време в миналото се среща в текстовете твърде рядко (Ницолова 2008: 316). Данните от Корпуса могат да служат за същото доказателство и по отношение на бъдеще предварително време.

От гледна точка на морфологичните категории лице, род и число бъдеще предварително време и бъдеще предварително време в миналото притежават по 12 възможни форми, които носят информация за лице, число и род на глаголния субект. В езиковите данни в Корпуса се срещат следните форми за бъдеще предварително време: 1. лице, единствено число, мъжки род; 3. лице, единствено число, мъжки род; 3. лице, единствено число, женски род; 1. лице, множествено число и 3. лице, множествено число. Формите за среден род, 1. и 2. лице, единствено число са възможни, но малко вероятни, тъй като 1. и 2. лице, единствено число означават комуникаторите, които са одушевени предмети. Формите за бъдеще предварително време в миналото, които се срещат в Корпуса, са: 1. лице, единствено число, мъжки род; 3. лице, единствено число, мъжки род; 1. лице, множествено число и 3. лице, множествено число. От гледна точка на трансфера на значението на категориите лице и число формите за двете времена се реализират в английски чрез времена, които имат само една форма, която е за цялата парадигма и не е маркирана с информация за лице и число. В тези случаи граматичната

информация не се реализира в рамките на глаголните форми при превод на английски, а посредством допълнителни лексикални и синтактични средства – чрез субекта в изречението.

IV.5. Обобщен анализ на превода на темпоралните морфологични категории от български на английски

В тази част ще бъдат анализирани статистическите данни, извлечени от „Българско-английския паралелен корпус със съотнесени глаголни форми“, за трансфера на темпоралните морфологични категории, влизащи в състава на хиперкатегорията време в български език, при превод на английски. Ще бъдат представени обобщено преводните реализации на значенията, изграждащи морфологичните категории време, вид на действието и таксис в български. В допълнение ще бъдат анализирани обобщените данни за категориите лице, число и род, с които се характеризират темпоралните форми в български. Накратко ще бъде засегнат и въпросът, свързан с лексикално-граматичната категория вид в български. Анализираните данни са поставени извън лексикалния, граматическия и прагматичния контекст, в който се срещат.

IV.5.1. Морфологичната категория време

Както беше посочено, морфологичната категория време в български език е тричленна категория с два маркирани и един немаркиран член. Основание за обединение на двата маркирани члена (предходност и следходност) при сравнението е значението разновременност. Основната формално-семантична опозиция в категорията е разновременност – неразновременност (с главно значение на немаркирания член – едновременност), според която минало и бъдеще време се противопоставят на немаркираното сегашно време. В рамките на разновременността се противопоставят предходност – следходност (Куцаров 2007: 244). С признака неразновременност са маркирани формите на сегашно време, минало несвършено време, минало неопределено време и минало предварително време. С признака следходност са маркирани формите на бъдеще време, бъдеще предварително време, бъдеще време в миналото и бъдеще предварително време в миналото. Информация за предходност на действието носят формите на минало свършено време. Статистическите данни за преводните реализации на трите грамеми от български на английски са представени в Таблица 30.

	Сегашни времена в английски		Минали времена в английски		Бъдещи времена в английски		Инфинитив в английски		Други части на речта в английски	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Неразновременност в български език	1532	42,79	1160	32,40	212	5,92	252	7,04	424	11,84
Следходност в български език	28	11,24	12	4,82	179	71,89	12	4,82	18	7,23
Предходност в български език	60	4,59	1121	85,70	7	0,54	30	2,29	90	6,88

Таблица 30. ***Данни за преводните реализации в английски на грамемите, изграждащи морфологичната категория време в български.***

Данните в Таблица 30. показват, че форми, изразяващи неразновременност, се срещат 3 580 пъти в българските текстове в „Българско-английския паралелен корпус със съотнесени глаголни форми“. Реализациите на неразновременността спрямо морфологичната категория време в английския език са нееднотипни – преобладаващо разпределение имат реализациите с форми за сегашни времена и форми за минали времена в английски. Глаголни форми, маркирани със значение следходност, се срещат 250 пъти в българските текстове, като при тях преводните реализации са сравнително еднотипни – преимуществено разпределение имат реализациите с форми за бъдещи времена в английски. Глаголни форми, маркирани със значение предходност, се срещат 1 310 пъти, като при тях реализациите също са сравнително еднотипни – с най-висока честота са преводните реализации с форми за минали времена в английски език. От трите значения, изграждащи морфологичната категория време в българския език, с най-стабилно поведение при превод се отличават формите, носещи значението разновременност; в английски следходността се реализира най-често с форми за бъдещи времена, а предходността – с форми за минали времена. Разнообразно е поведението на преводните реализации на формите, маркирани със значение неразновременност. Това разнообразие може да се обясни с факта, че в български със значение неразновременност са маркирани формите освен на същинското сегашно време, но и на голяма част от „миналите“ времена според класическото схващане за категорията време – минало несвършено време, минало неопределено време и минало предварително време. Семантиката на главните значения на „миналите“ времена може да бъде разделена на два подвида – отнасяща действия или резултат от действия като едновременни спрямо минал ориентационен момент (семантиката на минало несвършено време и минало предварително време) или отнасяща резултати от действия (минали приключили действия) към момента на говорене (семантиката на минало неопределено време) (в

рамките на термините от формално-семантичното представяне във Втора глава). Значението на формите за „миналите“ времена се пренася в английски с форми за минали времена (с най-висока честота при минало несвършено и минало неопределено са преводните реализации с минало просто време, а при минало предварително – преводните реализации са с минало перфектно време). От представените статистически данни може да се направи заключението, че макар че са близки по семантика, значенията на морфологичната категория време в български и английски имат различни проявления в темпоралните форми на двата езика и по различен начин маркират глаголните форми със значенията, които ги изграждат.

Различните значения, изграждащи морфологичната категория време в български и английски, може да споделят множество особености, но не може да се търсят формални съответствия при преводните им реализации – форма, маркирана с признака неразновременност в български език, може да се реализира като темпорална форма, маркирана с признака предходност в английски. Пренасянето на граматикализираната информация в подобни случаи се осъществява чрез допълнителни лексикални и синтактични елементи, както и чрез други морфологични категории, при чиито комбинации значението от български се пренася на английски в максимална степен. От представените статистически данни може да се изведе зависимостта, че признаците следходност и предходност се пренасят при превод, докато признакът неразновременност в български може да се реализира с темпорални форми, изразяващи едновременност или предходност в английски. Това се дължи най-вече на преводните съответствия на формите за сегашно време и минало неопределено време, от една страна, и съответствията на формите за минало несвършено време и минало предварително време, от друга страна.

IV.5.2. Морфологичната категория вид на действието

В представянето на хиперкатегорията време в български език морфологичната категория вид на действието беше представена като категория, изградена от привативната опозиция резултативност – нерезултативност. Темпоралните глаголни форми, маркирани с признака резултативност, представят приключили действия или резултати от действия спрямо определен ориентационен момент (Куцаров 2007: 250), докато формите, маркирани с признака нерезултативност, не съдържат такава информация и представят действията като актуализирани действие към даден ориентационен момент

като също така може да сигнализират и самото действие (акционност) (Куцаров 2007: 255). С резултативност в български език са маркирани формите на минало неопределено време, бъдеще предварително време, минало предварително време и бъдеще предварително време в миналото. Значението нерезултативност се носи от формите на сегашно време, минало свършено време, минало несвършено време, бъдеще време и бъдеще време в миналото. Статистическите данни за преводните реализации на резултативните и нерезултативните форми от български на английски са представени в Таблица 31.

	Прости времена в английски		Продължителни времена в английски		Перфектни времена в английски		Перфектни продължителни времена в английски		Инфинитив в английски		Други части на речта в английски	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Резултативност в български	175	37,72	13	2,90	198	49,33	8	1,56	8	1,79	30	6,70
Нерезултативност в български	3630	76,80	182	3,83	80	2,05	14	0,51	287	6,11	502	10,69

Таблица 31. ***Данни за преводните реализации на резултативните и нерезултативните форми от български на английски.***

В данните, представени в Таблица 31., с резултативност са маркирани 448 български темпорални форми от „Българско-английския паралелен корпус със съотнесени глаголни форми“, а с нерезултативност – 4 694 форми. Данните показват, че нерезултативните форми в български се реализират с най-висока честота в английски чрез форми за прости времена, докато резултативните форми имат нееднотипна преводна реализация, като с най-висока честота са формите на перфектните времена в английски. Глаголните форми, маркирани с признака резултативност в български, могат да бъдат разделени в две подгрупи според разпространението си: времена с относително висока честота на срещане – формите за минало неопределено време и минало предварително време, и времена с много ниска честота на срещане – формите за бъдеще предварително време и бъдеще предварително време в миналото. Докато формите на времената с ниска честота са реализирани единствено чрез форми за перфектни времена в „Българска граматика“, то формите за минало неопределено време и минало предварително време имат преводни реализации с висока честота с форми както за прости времена, така и за перфектни времена. От представените данни може да се изведе зависимостта, че значението нерезултативност в български език в значителна степен се пренася чрез форми на прости времена в английски, докато значението резултативност

има разнородни преводни реализации, като приоритет имат формите за перфектните времена. Тези разнородни реализации се дължат на различната семантика на минало неопределено време (едновременност на резултат с изказването); минало предварително време (едновременност на резултат с минал ориентационен момент); бъдеще предварително време (следходност на резултат към изказването) и бъдеще предварително време в миналото (следходност на резултат към минал ориентационен момент).

IV.5.3. Морфологичната категория таксис

Морфологичната категория таксис в български език е изградена от привативната опозиция между грамемите зависим и независим таксис, а маркираните значения на категорията разделят времената на относителни времена, ориентирани към ориентационен момент, различен от момента на говорене, и неотносителни времена, ориентирани към момента на говоренето. Относителните времена в български език са минало несвършено време, бъдеще време в миналото, минало предварително време и бъдеще предварително време в миналото. Липса на данни за относителност носят формите на сегашно време, минало свършено време, бъдеще време, бъдеще време в миналото и минало неопределено време. Статистическите данни за преводните реализации на формите, чрез които функционират двете грамеми, от български на английски са представени в Таблица 32.

	Прости времена в английски		Продължителни времена в английски		Перфектни времена в английски		Перфектни продължителни времена в английски		Инфинитив в английски		Други части на речта в английски	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Относителни времена в български	544	59,72	73	8,32	116	12,42	7	1,19	34	3,67	136	14,69
Неотносителни времена в български	3261	77,40	122	2,94	162	3,68	15	0,43	260	6,17	396	9,39

Таблица 32. ***Съотношение на преводните реализации на относителните и неотносителни времена в български на английски.***

В представените в Таблица 32. данни глаголните форми на относителните времена в българския език са общо 926, а на неотносителните – 4 216. И при двата вида времена преводните реализации на английски са сравнително еднотипни – осъществяват се с форми за прости времена. В Таблица 32. не са представени данните за преводни

реализации чрез бъдещи времена в миналото на английски – тези преводни реализации представляват около 1% от преводните еквиваленти на относителните времена и 0,5% от преводните еквивалентни на неотносителните времена. Статистическите данни показват, че при категорията таксис няма приоритетно разпределение на преводните реализации на английски – относителните и неотносителните времена се реализират чрез форми на прости времена в английски. От данните може да се изведе зависимостта, че категорията таксис в български няма строго определени преводни еквиваленти в английски. Значенията на относителните и неотносителните времена се трансферират при превод чрез значения на различни морфологични категории в английски език.

IV.5.4. Морфологичните категории лице, число и род

Морфологичните категории лице, число и род при глаголите са категориите, които характеризират глаголния субект. Както беше посочено, съществена разлика между български и английски е възможността на глаголните форми да бъдат натоварени с граматичните значения на категориите лице, число и род. В български език граматична информация за лице и число носят всички глаголни форми, като в някои ограничени случаи има изключения и съвпадение на форми (формите за 2. и 3. лице, единствено число при минало свършено време, минало несвършено време и минало предварително време). Информация за граматичен род носят формите на минало неопределено време, минало предварително време, бъдеще предварително време и бъдеще предварително време в миналото. Тези характеристики правят възможно подлогът в български да не бъде експлицитно изразен и глаголната форма да носи в пълна степен самостоятелно граматична информация за тези категории. В английски при по-голяма част от времената всички членове на парадигмата се изразяват от една немаркирана по отношение на лице и число форма (минало просто време, минало перфектно време, минало перфектно продължително време, бъдеще просто време, бъдеще продължително време, бъдеще перфектно време, бъдеще перфектно продължително време, бъдеще просто време в миналото, бъдеще продължително време в миналото, бъдеще перфектно време в миналото и бъдеще перфектно продължително време в миналото). В други случаи информация за граматично лице и число се носи само от определени форми, а за останалите членове на парадигмата се използва една немаркираната форма (3. лице, единствено число при сегашно просто време, 1. и 3. лице, единствено число при сегашно продължително време, 3. лице, единствено число при сегашно перфектно време и 3. лице,

единствено число при сегашно перфектно продължително време). В английски подлогът задължително се изразява експлицитно чрез именна фраза или местоимение, които носят граматичната информация за лице и число на глаголния субект. При безличните глаголи подлогът се реализира с третоличното местоимение it или показателното there. Род по отношение на граматичните характеристики на глаголния субект носят само местоименията в ролята на подлог в английски. Информация в синтезиран вид за трансфера на граматичните значения лице, число и род и съотношението между броя форми при превод от български на английски е представена в Таблица 35.

	Възможни форми в английски	Сегашно време в български	Минало свършено време в български	Минало несвършено време в български	Минало неопределено време в български	Бъдеще време в български	Минало предварително време в български	Бъдеще време в миналото в български	Бъдеще предварително време в български	Бъдеще предварително време в миналото в български
Възможни форми в български		6	5	5	12	6	9	6	12	12
Форми в Корпуса		6	5	5	10	6	6	5	5	4
Сегашно просто време в английски	2	2	2	2	2	2	1	1	–	–
Минало просто време в английски	1	1	1	1	1	1	1	1	–	–
Бъдеще просто време в английски	1	1	1	1	–	1	–	–	–	–
Сегашно продължително време в английски	3	3	3	1	1	2	–	–	–	–
Минало продължително време в английски	2	2	2	2	2	1	1	1	–	–
Бъдеще продължително време в английски	1	–	–	–	–	1	–	–	–	–
Сегашно перфектно време в английски	2	2	2	1	2	–	1	–	–	–
Минало перфектно време в английски	1	1	1	1	1	–	1	–	–	–
Бъдеще перфектно време в английски	1	–	–	–	–	–	–	1	1	–
Сегашно перфектно продължително време в английски	2	2	2	–	1	–	–	–	–	–
Минало перфектно продължително време в английски	1	1	1	1	1	–	1	–	–	–
Бъдеще перфектно продължително	1	–	–	–	–	–	–	–	–	–
Бъдеще просто време в миналото в английски	1	1	1	1	1	1	1	1	–	–
Бъдеще продължително време в миналото в английски	1	1	–	–	–	–	–	1	–	–
Бъдеще перфектно време в миналото в английски	1	–	–	–	1	–	1	1	–	1
Бъдеще перфектно продължително време в миналото в английски	1	–	–	–	–	–	–	–	–	–

Таблица 35. ***Съотношение между броя форми при превод от български на английски.***

Дистрибутивните данни в Таблица 35. показват, че в по-голямата част от случаите граматическата информация за лице и число не се пренася при превод от български на английски в рамките на съответните глаголни форми. Докато граматическата информация за род в български няма пряк корелат в английски и преносът е невъзможен на равнището на глаголните форми, то при категориите лице и число граматическата информация се пренася само в 8% от случаите (при 3. лице, единствено число, сегашно време в български към 3. лице, единствено число, сегашно просто време в английски; 1. лице, единствено число, сегашно време в български към 1. лице, единствено число, сегашно продължително време в английски; 3. лице, единствено число, сегашно време в български към 3. лице, единствено число, сегашно продължително време в английски; 1. или 3. лице, единствено число, сегашно време в български към 1. или 3. лице, единствено число, минало продължително време в английски; 3. лице, единствено число, сегашно време в български към 3. лице, единствено число, сегашно перфектно време в английски; 3. лице, единствено число, сегашно време в български към 3. лице, единствено число, сегашно перфектно продължително време в английски; 3. лице, единствено число, минало свършено време в български към 3. лице, единствено число, сегашно просто време в английски; 1. лице, единствено число, минало свършено време в български към 1. лице, единствено число, сегашно продължително време в английски; 3. лице, единствено число, минало свършено време в български към 3. лице, единствено число, сегашно продължително време в английски; 1. или 3. лице, единствено число, минало свършено време в български към 1. или 3. лице, единствено число, минало продължително време в английски; 3. лице, единствено число, минало свършено време в български към 3. лице, единствено число, сегашно перфектно време в английски; 3. лице, единствено число, минало свършено време в български към 3. лице, единствено число, сегашно перфектно продължително време в английски; 3. лице, единствено число, минало несвършено време в български към 3. лице, единствено число, сегашно просто време в английски; 1. или 3. лице, единствено число, минало несвършено време в български към 1. или 3. лице, единствено число, минало продължително време в английски; 3. лице, единствено число, минало неопределено време в български към 3. лице, единствено число, сегашно просто време в английски; 1. или 3. лице, единствено число, минало неопределено време в български към 1. или 3. лице, единствено число, минало продължително време в английски; 3. лице, единствено число, минало неопределено време в български към 3. лице, единствено число, сегашно перфектно време в английски; 3. лице, единствено число, бъдеще време в български към 3. лице,

единствено число, сегашно просто време в английски и 1. или 3. лице, единствено число, бъдеще време в български към 1. или 3. лице, единствено число, минало продължително време в английски). Във всички останали случаи формите в български, носещи граматичната информация за лице и число, се превеждат чрез немаркирана форма със значение за лице и число в английски. Информацията може да се реализира на синтактично или лексикално равнище чрез лексикалното значение на субекта или местоименен подлог.

IV.5.5. Лексикално-граматичната категория вид на глагола в български

Във втората част от дисертацията категорията вид на глагола беше представена като лексикално-граматична категория, която характеризира вътрешното развитие на действието от гледна точка на това, дали действието се развива до своята цялостност, представено е като комплексно – свършен вид, или е представено като процес без достигане на неговата цялостност и завършеност – несвършен вид (Пашов 1999: 136). Статистическите данни за преводните реализации на глаголните форми, образувани с глаголи от свършен и несвършен вид, са представени в Таблица 36.

	Форми на прости времена в английски		Форми на продължителни времена в английски		Форми на перфектни времена в английски		Форми на перфектни продължителни времена в английски	
	Брой форми	%	Брой форми	%	Брой форми	%	Брой форми	%
Глаголни форми с глаголи от свършен вид в български	2674	90,61%	64	2,17%	204	6,91%	9	0,30%
Глаголни форми с глаголи от несвършен вид в български	1131	83,84%	131	9,71%	74	5,49%	13	0,96%

Таблица 36. ***Съотношение на преводните реализации на глаголите от свършен и несвършен вид в български на английски.***

В данните в Таблица 36. се показва, че при глаголните форми на глаголите от свършен и несвършен вид в български се наблюдава приоритетно разпределение на преводни реализации на английски с форми за прости времена. Ако се изключат преводните съответствия с форми за простите времена, се получава следното разпределение на преводните реализации, представено в Таблица 37.

	Форми на продължителни времена в английски		Форми на перфектни времена в английски		Форми на перфектни продължителни времена в английски	
	Брой форми	%	Брой форми	%	Брой форми	%
Глаголни форми с глаголи от свършен вид в български	64	23,10%	204	73,65%	9	3,25%
Глаголни форми с глаголи от несвършен вид в български	131	60,09%	74	33,94%	13	5,96%

Таблица 37. ***Съотношение на преводните реализации на глаголите от свършен и несвършен вид в български с форми за продължителни и перфектни времена на английски.***

Данните в Таблица 37. показват, че при глаголите от свършен вид в български приоритетно разпределение с най-висока честота след формите за прости времена имат формите за перфектните времена в английски, а при глаголите от несвършен вид приоритетно разпределение с най-висока честота след формите за простите времена имат формите за продължителните времена в английски. Разглеждането на преводните реализации на формите на глаголите от свършен и несвършен вид по този начин предоставя данни в подкрепа на твърдението, формулирано във Втора глава, за еднакво значение и поведение значението на несвършения вид на глагола в български и значението на продължителните времена при категорията вид на действието в английски. Разбира се, трябва да се спомене и фактът, че е възможна и реализация на значението на глаголите от двата вида с допълнителни лексикални елементи – например глаголите от свършен вид се превеждат понякога с фразеологични глаголи, т.е. преносът на граматичното значение става на лексикално ниво. Въпреки експерименталните данни, представени в Таблица 37., данните в Таблица 36. показват, че честотата на разпределение на глаголи от несвършен вид съответно с форми на продължителни или перфектни времена, не може да бъде достатъчна за извеждане на зависимост между двете категории в български и английски. Данните показват, че не може да се търси пряка формална връзка между лексикално-граматичната категория вид на глагола в български и продължителните и перфектните времена в английски, но се наблюдава тенденция за превод на глаголи от несвършен вид в български с форми на продължителни времена в английски. Представените данни в Таблица 36. показват, че независимо от вида на глагола в български, най-голяма честота на разпространение имат формите на простите времена в английски, което доказва тезата за разликата в семантиката на категорията, която се дефинира като вид на глагола в български, и категорията, която се дефинира като вид на действието в английски, граматикализиран чрез продължителни глаголни и

перфектни глаголни форми. Категорията вид на глагола в български се различава спрямо вида на действие в английски както на семантично, така и на граматично ниво – видът на глагола в български определя действието като завършено или не спрямо целостта на самия изразяван процес, докато видът на действие в английски определя действието като продължително или завършено спрямо външен за действието момент.

IV.6. Изводи за българско-английските граматични паралели на темпорални форми

Представеният анализ на статистическите данни, извлечени от „Българско-английския паралелен корпус със съотнесени глаголни форми“, може да бъде интерпретиран в две направления – от гледна точка на избрания изследователски подход и от гледна точка на теоретичното семантично съотнасяне на темпоралните форми в български и английски.

Създаването на езикови корпуси и извличането на статистическа информация от тях е обвързано с граматическите особености на съответните езици и наличието на достатъчно надеждни и представителни езикови ресурси. Наличието на разработен корпус (Българско-английския паралелен корпус със съотнесени изречения), избран за съставянето на „Българско-английския паралелен корпус със съотнесени глаголни форми“, е значително преимущество за избрания корпусно базиран подход в изследването. Наличните анотационни слоеве на Българско-английския паралелен корпус със съотнесени изречения и разнообразието от жанрове на текстовете в него, както и текстовете, събрани от интернет, предоставиха възможността за изграждането на представителен лингвистичен ресурс. Една от основните пречки при съпоставителното статистическо изследване на езикови явления са различните морфологични категории и различната им проява на морфологично равнище. Броят на граматичните характеристики при глаголите и съответните им анотационни маркери за английски е в значителна степен по-малък от този за български. Това наложи опростяването на втория анотационен слой с цел унифициране на разликите между граматичните характеристики на двата езика за целите на изследването. Първоначално дефинираните цели за Корпуса са изпълнени, а именно: извлечени са статистически данни и е направен анализ за преводните реализации на деветте български глаголни времена към английски; извлечени са статистически данни и е направен анализ за поведението на хиперкатегорията време и морфологичните категории, влизащи в нейния състав

изолирано от контекста. Характеристиките на Корпуса не предвиждат възможността за изследване на други глаголни морфологични категории като залог, наклонение, свидетелственост и т.н. Също така извън обхвата на изследването са неличните глаголни форми (извън съставните глаголни форми) и техните проявления и поведение при превод между двата езика.

Формулираната хипотеза, че темпоралните морфологични категории в български език системно предопределят морфологичните категории на съответните преводни реализации в английски се потвърждава според представените статистически данни: в 70% от случаите, които са представени, темпоралната форма в български еднозначно предопределя темпоралната форма в английски.

Формално-семантичните описания за български и за английски, представени във Втора глава според системата на Х. Райхенбах, представят корелационни отношения на различните темпорални форми, изразявани в български и английски. Според представените статистически данни теоретично предвидените корелационни отношения между темпоралните форми не се реализират консистентно при езиковата употреба. Корелацията между формално-семантичния анализ и данните от Корпуса е следната: в 65% от случаите преводният английски еквивалент е предсказуем (на базата на формално-семантичния анализ) с 40% точност.

Причината за несъответствието между формално-семантичния анализ и емпиричните данни от Корпуса е комплексна и може да бъде разгледана в две направления – особеностите на двата езика и контекста на употреба. Както беше споменато, в английски език действа правило за съгласуване на времената, което предполага йерархична структура на темпоралните форми в изреченията и текста (Ранкова и Иванова 1980: 226). В български език голяма част от времената с по-ниска честота на разпространение изразяват фонови действия, които служат за темпорална ориентация на основното действие в изречението. Тези особености на двата езика предопределят темпоралните отношения в тях – докато в английския език те са силно зависими, в българския са сравнително свободни. Съпоставката между разпределението на формите на българските и на английските времена в двете части на Корпуса (българска и английска) дава основание за подкрепа на виждането, че честотата на употребата следва семантичната натовареност на членовете на една категория, а разпространението им е зависимо и от стиловите особености на текста. Друга характерна особеност е, че лексикално-граматичната категория вид на глагола и морфологичната

категория вид на действието в български имат значителна роля при представянето на действията като приключили, докато в английския език формално маркираното минало време е функционален сигнализатор за приключило събитие (Ранкова и Иванова 1980: 113). От гледна точка на прагматичния контекст, в който се появяват темпоралните глаголни форми и означаваните от тях действия, в български език разпространението на глаголни форми в различните видове текстове е значително по-ограничено от тези в английски. Например при художествените текстове преобладаващата част от темпоралните форми е заета от форми за минали времена, докато в английски разнообразието от темпорални форми при тези текстове е сравнително по-голямо (виж Таблици 12. и 14.).

Представянето на преводните реализации на морфологичните категории, влизащи в състава на хиперкатегорията време, изолирано от контекста, показва, че макар близки в семантичен и структурен план, категориите в български и английски имат различно поведение при превод. Както е показано, при превод на българските глаголни форми на английски информацията за отношенията на действието спрямо ориентационния момент или момента на говорене и отношенията на резултата от действието спрямо ориентационния момент или момента на говорене се пренася в значително голяма степен. Представеният анализ на особеностите на категорията вид на действието в български и вид на действието в английски показа, че значенията резултативност – нерезултативност в български и значенията перфективност – неперфективност при перфектните времена в английски имат предвидимо поведение на преводните им реализации. Нерезултативните времена в български език имат своите корелати в английски и в по-голяма част от случаите при превода запазват значението си. Нерезултативността като част от значението на глаголните форми в български има значително по-консистентни преводни еквиваленти в английски в съответните прости времена, докато резултативността като част от значението на глаголните форми има разнородни преводни реализации – около 50% от резултативните форми на български се реализират с перфектни или перфектни продължителни времена в английски. Това се дължи на факта, че докато в български резултативността „изразява резултат от действие, актуализиране на действието в най-близкия следващ във времето ориентационен момент“ (Куцаров 2007: 250), в английски език перфектните времена изразяват само наличен резултат от приключило действие или наличен резултат от продължаващо действие (Ранкова и

Иванова 1980: 115), а приключило действие може да бъде изразено и чрез форми на минали времена.

Статистическите данни за категорията таксис в българския език представят сравнително еднотипно поведение при превод от гледна точка на морфологичните категории в английски. Както относителните, така и неотносителните глаголни форми се реализират в значителна част от случаите с форми за прости времена в английски. Представените данни не могат да дадат еднозначен отговор за естеството на морфологичните корелати на българската категория таксис в английски. Голямото разнообразие от преводни реализации на относителните глаголни форми и малката честота на разпространение на бъдещи времена в миналото в английски доказват, че между категорията таксис в български и бъдещите времена в миналото в английски не могат да се търсят съответствия на формално ниво.

Представените статистически данни за лексикално-граматичната категория вид също не предоставят солидно доказателство за формалната обвързаност между несвършения вид на глагола в български и продължителните времена в английски. Преводните реализации на български глаголи от свършен вид в повечето случаи са непродължителни форми в английски, докато реализациите на глаголи от несвършен вид са почти поравно разпределени между продължителни и непродължителни глаголни форми в английски. Свършеният вид може да се изразява на английски и лексикално – с глаголи с предлози или наречия.

Представените данни за поведението на граматичната информация за лице, число и род при превод от български на английски показват, че в по-голяма част от случаите тази информация не се пренася при превод в рамките на съответстващата глаголна форма в английски. Информацията за лице, число и род на глаголния субект в английски се носи от експлицитно изразен подлог, докато в български може да остане неизразена на лексикално ниво. Тези особености определят трансфера на граматичната информация за лице и число в български извън глаголната форма в английски. Информация за граматичен род се пренася само в случаите, когато експлицитният подлог в английски е изразен чрез третолично местоимение в единствено число и в по-редки случаи лексикално, когато има отделни думи за индивиди от различен пол.

В голяма степен темпоралните системи в български и английски и морфологичните категории, които ги изграждат, имат формално-семантични сходства помежду си. Структурата и йерархията на категориите, които изграждат

хиперкатегорията време в двата езика, подлежи на съпоставителен анализ от гледна точка на естеството на граматикализираната информация и начините на изразяване на темпоралните отношения. Данните от Корпуса показват, че в значителна степен лексикалният контекст и синтактичните особености на езиците, както и прагматичната употреба предопределят използването на темпоралните форми. Граматикализираната информация се запазва в значителна степен, но се пренася също така и чрез допълнителни езикови елементи – лексикален материал или контекстна съгласуваност между отделните темпорални форми в рамките на текста. Статистическите данни също показват, че изследването на морфологичните кореспонденции между темпоралните системи в български и английски и преводните реализации чрез „Българско-английския паралелен корпус със съотнесени глаголни форми“ е възможно и ползотворно.

V. Практическо приложение и оценка на „Българско-английския паралелен корпус със съотнесени глаголни форми“

След създаването на Корпуса са извлечени данни за преводните реализации на глаголните форми в Корпуса и е направен сравнителен анализ. Теоретичният анализ на статистическите данни е първичният етап за оценка на новосъздадени езикови ресурси (Родриго 2008: 124). За да бъде „Българско-английският паралелен корпус със съотнесени глаголни форми“ пълноценен и функционален лингвистичен ресурс, след извличането и анализа на статистическите езикови данни част от него е преработена за използване в софтуер за компютърно подпомогнат превод. Внедряването и използването на Корпуса като практически приложим езиков ресурс е втората основна поставена цел след статистическия обоснования теоретичен анализ. Обогатяването на преводни модели с лингвистична информация за съотнасянето на български и английски глаголни форми е приложено към преводна памет с цел използване на езиковите данни за компютърно подпомогнат превод. Преводната памет е най-подходящият ресурс за създаване на статистически или невронни преводни модели поради факта, че предлага сдвоени на изреченско (или фразово) ниво преводни еквиваленти на два или повече езика, преведени (и редактирани/проверени) от хора. Преводната памет е „многоезиков текстов архив, съдържащ сегментирани, сдвоени, анотирани като части на речта или синтактично и класифицирани многоезикови текстове, позволяващ съхранение, търсене и предоставяне на подравнени многоезикови текстови сегменти, отговарящи на различни критерии за търсене“ (ЕГПКОЕ 1996: 6). Преводната памет съдържа сегментирани по определени критерии текстове. Сегментирането най-често е на изреченско ниво, което позволява бързо и лесно търсене, превод на вече повтарящи се части от изреченията и еднозначно съотнасяне на преводните единици между различните езици. В определени случаи е възможна работата с по-малки езикови единици – фрази или думи. Преводната памет съхранява преводни единици – сегменти, изречения или фрази в зависимост от избраното ниво на сегментация. Системите за компютърно подпомогнат превод предлагат търсене в преводната памет чрез два основни алгоритъма (Лагудаки 2006: 19). Най-разпространените системи за компютърно подпомогнат превод използват алгоритъм, който оценява разстоянието между сравняваните низове при брой операции: изтривания, замествания или вмъквания (edit distance). Най-често това е разстояние по Левенщайн. Недостатъкът на този алгоритъм е, че преводни съответствия се търсят само

в един възможен сегмент – системите за компютърно подпомогнат превод приемат, че даден изходен сегмент има само един възможен превод. Резултатът може да е пълно съвпадение или частично съвпадение между сегментите. При такъв тип търсене единственият низ за търсене може да бъде цялата преводна единица. Съществуват системи за компютърно подпомогнат превод, които използват алгоритъм, който търси в несегментиран текст най-дългите съвпадащи низове (Лагудаки 2006: 21). При този метод на търсене съвпаденията могат да са само пълни и не са ограничени по дължина – могат да търсят извън рамките на изречение, параграф, документ. Алгоритъмът търси в двойки съответстващи документи всички съвпадения с произволна дължина. При този метод се използват двуезикови корпуси, данните от които се обобщават и предоставят на преводача.

Направени са експерименти и анализ със софтуера за компютърно подпомогнат превод – memoQ[14]. memoQ е един от най-използваните системи за компютърно подпомогнат превод. Създаден е в Будапеща през 2004 година като софтуерна среда за преводи, която интегрира множество приложения не само за превод, но и за компютърна обработка на входния/целевия език, създаване на преводна памет, езикови корпуси, терминологични бази от данни и др. Към момента разработването на memoQ е съсредоточено предимно върху приложенията за компютърна обработка на езика. Съществено преимущество на memoQ е възможността да работи с регулярни изрази и съставянето и внедряването на детерминирани и недетерминирани крайни автомати, както и преводаческата работа върху предварително анотирани текстове с различни нива на анотация.

За да бъде експериментално доказана надеждността на извлечените статистически данни и приложимостта на „Българско-английския паралелен корпус със съотнесени глаголни форми“, информацията от анотационните слоеве на Корпуса е преработена в терминологична база от данни с възможност за предсказващо приписване на вероятностна метаинформация върху целевия език на базата на входния език при работа с memoQ. По този начин е направена възможна работата върху предварително анотирани с информация за морфологичните характеристики на думите входни български текстове. Терминологичната база от данни работи върху предварително анотиран текст с маркери за морфологичните характеристики на думите, който се задава като изходен език за превод в програмата. Терминологичните бази от данни използват

[14] https://www.memoq.com/

унифициран SRX стандарт, по който системите или други инструменти за езикова обработка сегментират текст. SRX стандартът представлява специализиран файл с регулярни изрази, чрез които е възможно деленето на преводните единици на по-малки (фразови) структури, термини, специфициране на граматични форми и т.н. С негова помощ системите за компютърно подпомогнат превод и други лингвистични инструменти могат да описват езиково специфичния процес, при който текстът се разделя на сегменти с цел по-нататъшната им обработка (XLIFF 1.2 2007: 14).

На базата на маркерите с морфологични характеристики, които присъстват в отделните сегменти (изречения) от изходния език (български), терминологичната база предлага най-вероятния маркер за преводната реализация на глаголната форма от целевия език (английски). След зареждането на такъв текст в софтуера за компютърно подпомогнат превод memoQ, на базата на маркерите за морфологичните характеристики на думите на входния текст софтуерът сравнява наличието на определена съвкупност от маркери за морфологичните характеристики на думите в изходния текст спрямо наличните статистически данни в базата данни. Въз основа на тези сравнения предоставя възможните маркери (с наличие на информация за глаголните форми от втория анотационен слой на Корпуса) за преводния еквивалент на глагола в изречението на целевия език (английски). Имплементирани по този начин, данните от Корпуса и възможностите на memoQ не предлагат превод на глагола, а само съответната му форма за целевия език. По този начин данните от Корпуса могат да бъдат използвани като помощно средство при превод на текстове с български за изходен език и английски като целеви език.

За да бъде оценен „Българско-английският паралелен корпус със съотнесени глаголни форми“ и неговото практическо приложение, резултатите от експеримента с данните от Корпуса и memoQ са оценени от хора – преводачи, работили върху различни текстове, съвпадащи по област с областите на текстовете от Корпуса. Преводачите са двама и са предварително запознати с тестовата постановка, естеството на текстовете и особеностите на работата със софтуера за компютърно подпомогнат превод. Тестването е извършено с помощта на независими преводачи и софтуера за компютърно подпомогнат превод memoQ върху извадки от текстове, подбрани на случаен принцип. Подборът на текстовете е направен така, че да осигури максимално разнообразие от възможни глаголни форми в изходния език. Тестовите текстове представляват извадки от текстове на български, съвпадащи по област на употреба с текстовете, включени в

Корпуса. Общият обем на тестовите текстове е 1 580 думи, като в тях се съдържат 200 глаголни форми на български. Текстовете са разделени в две групи – за тестване и за оценка, като за тестване за използвани текстове, съдържащи 150 глаголни форми, а за оценка – текстове, съдържащи 50 глаголни форми. Текстовете за тестване са предварително анотирани с маркери за морфологичните характеристики на думите и използвани като входни текстове в memoQ. След приписването на най-вероятен маркер на целевия текст на базата на базата данни, извлечена от Корпуса, и входния анотиран текст, е проследено в колко от случаите преводачите ще приемат приписаната автоматично глаголна форма за глагола в текущото изречение или ще я отхвърлят. Тестовете за оценка са предварително преведени, след което ръчно е приписана глаголна форма в главното изречение според анотационните конвенции на Корпуса. Анализирано е в каква степен статистическите данни за разпределението на глаголните форми от тези текстове се доближават до статистическите данни, извлечени от Корпуса.

Резултатът от проведения експеримент за практическото приложение на Корпуса е, че на базата на моделите за съотнасяне, съдържащи се в Корпуса, могат да се съотнесат автоматично глаголните форми на двата езика, за да се създадат обогатени преводни памети и модели с лингвистична информация, които могат да се използват както за подпомагане на преводния процес, така и за автоматична редакция на вече преведени текстове. Проведеният експеримент с терминологични бази от данни и софтуер за компютърно подпомогнат превод има за цел да се провери практическото приложение на лингвистичните данни и статистическата информация от „Българско-английския паралелен корпус със съотнесени глаголни форми“. В хода на тестването е установено, че в 85% от случаите, в които терминологичната база, изградена върху данните от Корпуса, предлага маркер за глаголната форма в целевия език на базата на езиковите маркери от изходния език, преводачът се съгласява с приписаната глаголна форма. В останалите случаи автоматично предложената глаголна форма не съвпада с избора на преводача, като в 9% от тези случаи избраната от преводача глаголна форма съвпада с втората по вероятност предложена форма, а в останалите 6% – с друга поредна предложена форма. Не е регистриран случай, в който избраната от преводача глаголна форма да е извън предложения списък с възможни глаголни форми. Данните от тестването са представени в Таблица 38.

Глаголни форми в български	Брой форми в тестовите текстове	Случаи с избор на първата предложена глаголна форма	Случаи с избор на втората предложена глаголна форма	Случаи с избор на друга предложена глаголна форма
Сегашно време	98	85	9	4
Минало свършено време	47	39	3	5
Минало несвършено време	36	28	5	3
Бъдеще време	15	14	1	0
Минало неопределено време	4	4	0	0

Таблица 38. ***Данни от тестването на практическото приложение на Корпуса.***

В резултат от тестването са планирани подобрения, които имат за цел да коригират терминологичната база в съответствие с редакциите на текста, извършвани от хора, както и последващо обогатяване на Корпуса с примери и редактираните текстове. Имплементирани са корекции по терминологичната база, които правят възможно присъединяването на новопреведени текстове от хора към Корпуса. В тези случаи новите текстове, които се добавят, съдържат информация само за глаголните форми и анотационни маркери от втория анотационен слой. По този начин Корпусът може да бъде разширен с примери, които произтичат от непосредствените решения на преводачите, като примерите може да са подбрани така, че да илюстрират явленията с най-ниска честота в Корпуса.

VI. Заключение

VI.1. Обобщение

Основната цел на настоящата дисертация, дефинирана в Увода, е изследването на граматичните паралели при темпоралните системи на българския и английския език.

Дисертацията предлага теоретичен подход за анализ на свойствата и организацията на темпоралните глаголните системи в български и английски. Избран е изследователски подход, който е доказан като успешен при изследването на съпоставителни проблеми в езици с различна структура. Изследователският подход е базиран на съпоставителното разглеждане на формално-семантичното представяне на темпоралните категории в двата езика и статистически модел за превод на темпорални форми, извлечени от специално разработен за целта паралелен двуезиков корпус: „Българско-английския паралелен корпус със съотнесени глаголни форми“.

От теоретична гледна точка изследването е мотивирано от хипотезата, че темпорални системи на двата езика предоставят възможност за създаване на преводен модел, изграден на базата на формално-семантичната съпоставка на темпоралните категории, които изграждат тези темпорални системи. Емпиричната верификация на този модел, осъществена на базата на статистически данни, извлечени от „Българско-английския паралелен корпус със съотнесени глаголни форми“, показва достоверността на теоретичния преводен модел. От приложна гледна точка, приложимостта на преводния модел е показана в среда за компютърно подпомогнат превод. Приложният аспект на изследването е свързан и със създаването на нов езиков ресурс („Българско-английския паралелен корпус със съотнесени глаголни форми“), чрез който обектът на изследването да бъде представен чрез естествените му езикови проявления.

Представеното изследване е актуално на фона на засиления през последните години интерес към различни научно-приложни разработки за описание и анализ на езиковите явления чрез методите на компютърната лингвистика и в частност чрез възможностите на корпусно базирания подход. Работата е насочена към комплексно и непротиворечиво описание на механизмите за пренос на граматикализираната информация от глаголните форми при превод от български на английски. Изучаването на механизма на взаимодействие между темпоралните системи в български и английски

и морфологичните категории, които ги изграждат, има теоретична стойност за изясняване на закономерностите при съотнасянето между граматикализираната информация, семантичната структура и начините за изразяване на темпоралните отношения в двата езика.

За постигането на основната цел на дисертацията, както и на съвкупността от цели, които я изграждат, са поставени няколко взаимосвързани задачи, изпълнението на които е довело да следните резултати:

Изработена е единна концепция за изследване и описание на темпоралните форми в двата езика и тяхната семантика, която се основава на концепцията за граматичната категория (хиперкатегорията) време в български и английски. Приема се, че за български хиперкатегорията включва морфологичните категории време, вид на действието и таксис, а за английски: време и вид на действието. Единната концепция за изследване и описание на темпоралните форми в български и английски се основава на формално-семантично описание на темпоралните морфологични категории в български и английски в рамките на модела на Х. Райхенбах (Райхенбах1947). За български се следва формално-семантичният модел, представен от Р. Ницолова (Ницолова 2008), а за английски моделът е надграден, така че да бъде съпоставим с формално-семантичната репрезентация за български. Обособени са граматичните паралели между формално-семантичните описания на темпоралните отношения в двата езика, представящи сами по себе си хипотеза за преводните съответствия и семантичния трансфер на граматична информация при превод.

Създаденият за целите на дисертацията „Българско-английски паралелен корпус със съотнесени глаголни форми“ демонстрира на базата на реално съществуващи примери и техните преводи създаденото формално-семантично описание на темпоралните системи на двата езика. Дизайнът на „Българско-английския паралелен корпус със съотнесени глаголни форми“ е обвързан с теоретичната хипотеза, а извлечените от него статистически данни служат за нейно доказателство или отхвърляне. Анализът на изискванията към езиковите ресурси, които се използват за машинен превод, и на съществуващите паралелни ресурси за български и английски са в основата на избора на критерии за изграждането на „Българско-английския паралелен корпус със съотнесени глаголни форми“. Паралелните глаголни форми в „Българско-английския паралелен корпус със съотнесени глаголни форми“ са анотирани ръчно с помощта на системата за анотация Уебано. За целта са разработени и последователно спазени

анотационни конвенции, които осигуряват консистентната анотация на глаголните форми в Корпуса.

В резултат е създаден паралелен лингвистично анотиран корпус, който съдържа 45 843 думи за български и 49 432 думи за английски, в който са анотирани 4 458 синтетични и 684 аналитични темпорални форми за български, както и 3 438 синтетични и 862 аналитични форми за английски. Текстовете в Корпуса са извлечени от „Българско-английския паралелен корпус със съотнесени изречения (и прости изречения в състава на сложното)“ (Коева и др. 2012б) (84%), от интернет (15,5%) и от „Българска граматика“ (0,5%). Корпусът съдържа 65,5% текстове художествена литература, 18% публицистични текстове, 13% филмови субтитри, 3% научни текстове и 0,5% илюстративни примери от „Българска граматика“.

Анализът на граматичните паралели в „Българско-английския паралелен корпус със съотнесени глаголни форми“ представя: данните за българските и английските темпорални форми поотделно, данните за трансфера на информация за морфологичните категории време, вид на действието и таксис в български към английските морфологични категории време и вид на действието. Предимство на анализа е, че се прави в рамките на различните тематични области в Корпуса.

Представените данни доказват, че теоретичната хипотеза за съответствието при формално-семантичния анализ на двата езика е доказуема чрез избрания изследователски подход, и показват, че в съществена част от случаите темпоралната форма в английски като целеви език зависи от комбинаторното наличие на определени морфологични категории в български като изходен език. Значението на контекста за интерпретацията на глаголните форми в български и английски представлява обективна пречка за пълно потвърждение на хипотезата: при формално-семантичния анализ контекстът не се взема под внимание, докато при анализа на реалните случаи на преводни еквиваленти, регистрирани в „Българско-английския паралелен корпус със съотнесени глаголни форми“, контекстът играе съществена роля за изразяване на темпоралните значения.

От представените статистически данни, разглеждани независимо от теоретичната хипотеза, се вижда, че глаголните форми с най-висока честота на разпространение имат сравнително разнообразни преводни реализации, докато глаголните форми с по-ниска честота на разпространение имат сравнително малък брой възможни преводни реализации в английски. Също така определени глаголни форми не се срещат в рамките

на езиковите данни от Корпуса, което доказва тезата за намаляващата им употреба. Могат да бъдат направени следните изводи:

1. Темпоралните системи в българския и английския език имат достатъчно сходства във формално-семантичен план, за да бъдат разглеждани паралелно и да е възможно изграждането на надежден преводен модел, базиран на семантичните свойства на морфологичните категории, които изграждат двете системи.

2. Морфологичните категории, които съставят темпоралните системи в български и английски, имат сходни формално-семантични характеристики и предполагат сходно моделиране на граматикализираните значения в двата езика.

3. Статистическите данни, извлечени от Корпуса, доказват, че комбинаторното наличие на определени морфологични категории в български език, които изграждат темпоралното значение, предопределя темпоралната форма при превод на английски. Различават се начините за изразяване на граматикализираната информация между двата езика. Статистически данни също така показват, че най-често срещаните глаголни форми в български имат сравнително разнообразни преводни реализации в английски, докато по-комплексните и съответно по-рядко разпространени глаголни форми имат сравнително еднотипни преводни реализации.

Изводите спрямо изследователския подход, избран за анализ на обекта на изследването, са:

1. Корпусно базираният подход се явява функционален и успешен за изследването на граматичните паралели между темпоралните системи в български и английски.

2. Обемът и представителността на „Българско-английския паралелен корпус със съотнесени глаголни форми“ са достатъчни, за да дадат отговор на поставените въпроси в това изследване, но не са достатъчни, за да представят пълна картина на различните особености на темпоралните системи в български и английски. Ниската честота на разпространение и твърде специфичните срещания на някои темпорални форми са причина да няма достатъчно надеждни лингвистични данни за тяхното изследване.

3. Провеждането на експеримент с използването на терминологична база от данни, съдържаща преводната граматична информация от Корпуса, за

компютърно подпомогнат превод показва успешното приложение на паралелните езикови модели за подобряване на превода. Данните от Корпуса могат да бъдат приложени като помощно средство при работа със системите за компютърно подпомогнат превод, както и при подбора на едноезикови и паралелни данни за обучение на системите за машинен превод.

Изследването представя актуален проблем, свързан с комплексен езиков феномен. Проблемът е представен чрез съвременен и научно доказан изследователски подход, чрез който са реализирани целите на проучването и са предоставени емпирични и количествени данни за обекта на изследване: българско-английските граматични паралели при превода на глаголните форми от български на английски.

VI.2. Приноси на изследването

Научните приноси на изследването може да се обобщят в няколко основни направления, които отговарят на организацията на частите на дисертацията, и могат да бъдат формулирани, както следва:

Представяне на теоретично описание на темпоралните системи в български и английски, както поотделно, така и в съпоставка. Проследено е развитието на теоретичните виждания за категориите, изграждащи темпоралните значения при глаголните форми. В съпоставителен план е изследван механизмът на конструиране на граматичното темпорално значение в българските и английските глаголни форми и са формулирани основните закономерности при взаимодействието между темпоралните морфологични категории в двата езика. Представено е формално-семантично съпоставяне между значенията на темпоралните категории в български и английски.

Конструиране и анотиране на „Българско-английския паралелен корпус със съотнесени глаголни форми“. При работата по дисертацията не беше намерена информация за наличието на друг такъв паралелен ресурс за български и втори език. Представени са процесът по подбор на ресурсите, предварителната обработка, анотацията и използваните софтуерни приложения за работата по Корпуса.

„Българско-английският паралелен корпус със съотнесени глаголни форми“ предоставя възможности за изследването на граматическите паралели между български и английски език. Корпусът може да се разглежда като средство за формалното описание на граматичните паралели между темпоралните системи на

български и английски и източник за извличането на статистически модели за отношенията на преводните реализации на глаголните форми в двата езика. Езиковите данни от Корпуса могат да бъдат използвани за обогатяване на преводни модели за нуждата на машинния превод, както и за теоретични изследвания. Представеният работен процес по структурирането и анотирането на Корпуса може да послужи при разработването на подобни лингвистични ресурси за изследване на сходни езикови явления.

Представена е обективна количествена и дистрибутивна информация за глаголните системи на български и английски, подходяща за различен тип научен анализ извън обхвата на проучването. Осигурени са надеждни лингвистични данни за създаването на различни езикови приложения за компютърна обработка на естествения език като статистически модел за анотация и статистически модел за превод на глаголните форми.

Допълнителен принос е интегрирането на лингвистичните данни от „Българско-английския паралелен корпус със съотнесени глаголни форми“ в софтуер за компютърно подпомогнат превод с цел практическото му приложение като помощно средство при превод. По този начин статистическите данни могат както да бъдат използвани при процеса на превод, така и да бъдат коригирани и обогатени с допълнителен езиков материал.

Библиография

Алексова 2003: Алексова, Кр. Синтагматичната неутрализация на граматическите опозиции като причина за различните състояния, които морфологичната система може да заема (върху материал от темпоралната система на българския език). – В: *Материали от Международния славистичен симпозиум в рамките на Юбилейните Паисиеви четения*. Пловдив: Пловдивски университет [http://belb.info/personal/aleksova/neutralizacii_1.pdf] (посетен на 30.01.2020)

Алексова 2008а: Алексова, Кр. Проблемът за неутрализациите на темпоралните опозиции и смяната на гледната точка в българското сложно изречение в преподаването на български език на чужденци, говорещи френски език. – В: *Доклади от Петата международна конференция „Езикът – феномен без граници”*, 12-14 юни 2008 г., Варна, 173– 181.

Алексова 2008б: Алексова, Кр. За два феномена при съвместната употреба навремената в българското сложно изречение: емпатията и темпоралните неутрализации. – В: *сп. Езиков свят, кн. 6*, 7 – 22.

Алексова 2018: Алексова, Кр. Облигаторност и контекстова зависимост на морфологичните категории (върху материал от съвременния български език). – В: *Научни трудове, т. 55, кн. 1, сб. А, 2017. Филология. Паисиеви четения, Българистика*. Пловдив: УИ „Паисий Хилендарски“, с. 38 – 51.

Андрейчин 1938: Андрейчин, Л. Обща характеристика на българската глаголна система (превод на Ogólna charakterystyka bułgarskiego system koniugacyinego, 1938). – В: Пашов, П. Ницолова, Р. (съст.) *Помагало по българска морфология. Глагол.* София: Наука и изкуство.

Андрейчин 1944: Андрейчин, Л. *Основна българска граматика*. София: Наука и изкуство.

Андрейчин и др. 1997: Андрейчин, Л. Попов, К. Стоянов, Ст. Граматика на българския език. София: Наука и изкуство.

Атанасова и др. 1963: Атанасова, Т. Ранкова, М. Русев, Р. Радулова, Н. *Английска граматика в сравнение с български език*. София: Наука и изкуство.

Аюн и др. 2018: Ayoun, D. Celle, A. Lansari, L. (eds.), *Tense, Aspect, Modality, and Evidentiality: Crosslinguistic perspectives*. Amsterdam: John Benjamins Publishing Company.

Баева 2011: Баева, Д. За корпусите като езикови ресурси и тяхното приложение. – В: *Научни трудове на Русенския университет - 2011*, том 50, 48-53.

Байбър 1993: Biber, D. Representativeness in Corpus Design. – In: *Literary and Linguistic Computing, 8*, 243-257.

Баскервил и Сеуел 1895: Baskervill , W. Sewell, J. *An English Grammar: For the Use of High School, Academy, and College Classes.* New York: American Book Compnay [https://archive.org/stream/anenglishgramma00sewegoog#page/n156/mode/2up] (посетен на 20.02.2020).

Бенвенист 1971: Benveniste, E. *Problems in General Linguistics*. Miami: University of Miami Press.

Бозвели 1835: Бозвели, Н. *Славеноболгарское детеводство за малките деца. Част III*. Крагуевац.

Бондарко 1976: Бондарко, А. *Теория морфологических категорий*. Ленинград: Наука.

Бондарко и др. 1987: Бондарко, А. Маслов, Ю. Шелякин, М. *Теорияфункциональной грамматики: введение, аспектуальность, временная локализованность, таксис.* Ленинград: Наука.

Боукър 2002: Bowker, L. *Computer-aided translation technology: A practical introduction.* Ottawa: University of Ottawa Press.

Браун и др. 1993: Brown, P. Pietra V. Pietra, S. Mercer, R. The Mathematics of Statistical Machine Translation: Parameter Estimation. – In: *Computational Linguistics. Volume 19, Issue 2. Special issue on using large corpora: II*, 1993, 263 – 311.

Букхолз и Марси 2006: Buccholz, S. Marsi, E. CoNLL-X shared task on multilingual dependency parsing. – In: *Proceedings of the Tenth Conference on Computational Natural Language Learning*, 2006, 149 – 164.

Бхат 1994: Bhat, D.N.S. *The Adjectival Category: Criteria for differentiation and identification.* Amsterdam: John Benjamins Publishing Company.

Валейка и Буиткийн 2003: Valeika, L. Buitkiene, J. *An introductory course in theoretical English grammar*. Vilnius: Vilnius Pedagogical University.

Васева 1989: Васева, И. *Стилистика на превода: специфика на превода на различни видове и жанрове текст*. София: Наука и изкуство.

Васенхов 2017: Wassenhove, V. Defining moments for conscious time and content. – In: PsyCh Journal, 10, 2017, 166 – 169.

Ватер 1994: Vater, H. *Einführung in die Zeit-Linguistik*. Frankfurt am Main: Gabel.

Вейга 1991: Veiga, A. *Condicionales, concesivas y modo verbal en español*. Santiago de Compostela: Universidade de Santiago de Compostela.

Гаравалова 2003: Гаравалова, Ил. *Школата на натуралната морфология и българската морфологическа система*. София: СемаРШ

Герджиков 1984: Герджиков, Г. Преизказването на глаголното действие в българския език. София: Наука и изкуство.

Герджиков 2000: Герджиков, Г. Категорията време като хиперкатегория – В: сп. *Български език*, бр. 1, 2000 [https://liternet.bg/publish/ggerdzhikov/hyper.htm] (посетен на 20.02.2020).

Грийнбаум и Нелсън 1991: Greenbaum, S. Nelson, G. *An Introduction To English Grammar*. Oxford: Oxford University Press.

Грийнбърг 1971: Greenberg, J. Language Universals: A Reaserch Frontier. – In: *Language, Culture and Communication.* California: Stanford University Press, 295-313.

Груев 1858: Груев, Й. *Основа за българска граматика*. (1987), София: Наука и изкуство.

ГСБКЕ II: *Граматика на съвременния български книжовен език. Том 2: Морфология.* (отг. ред. Ст. Стоянов), София: Институт за български език, БАН.

Деклерк 1991: Declerck, R. *Tense in English: Its Structure and Use in Discourse.* Abingdon-on-Thames: Routledge

Деклерк и др. 2008: Declerck, R. Reed, S. Cappelle, B. *The grammar of the English verb phrase, volume 1: The grammar of the English tense system: A comprehensive analysis*. New York: Mouton de Gruyter.

Джонсън 1640: Johnson, B. *The English Grammar*. 1640, London [https://archive.org/details/englishgrammar00jons/page/n8] (посетен на 20.02.2020).

Джонсън 1766: Johnson, S. *A Grammar of the English Tongue*. 1766, [https://archive.org/stream/dictionaryofengl01johnuoft#page/n63/mode/2up] (посетен на 20.02.2020).

Джурафски и Мартин 2006: Jurafsky, D. Martin, J. *Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River: Prentice Hall PTR.

Димитрова 2009: Димитрова, М. Езикови особености на текста на енциклопедията – В: Викторова, К. (съст.) *Проблеми на граматичната система на българския език – глагол*. София: Академично издателство „Проф. Марин Дринов“ 419-488.

ЕГПКОЕ 1996: Expert Advisory Group for Language Engineering Standards, *Preliminary recommendations on corpus typology*. Instituto di Linguistica Computazionale, 1996 [http://www.ilc.cnr.it/EAGLES/corpustyp/corpustyp.html] (посетен на 20.02.2020).

Зипф 1935: Zipf, G. *The Psychobiology of Language: An Introduction to Dynamic Philology*. Camridge: M.I.T. Press.

Иванчев 1971: Иванчев, Св. *Проблеми на аспектуалността в славянските езици.* София: БАН.

Иванчев 1978: Иванчев, Св. *Приноси в българското и славянското езикознание*. София: Наука и изкуство.

Йесперсен 1965: Йесперсен, О. *The Philosophy of Grammar*. (1992), Chicago: The Universityof Chicago Press.

Имам и др. 2014: Yimam, S. Castilho, E. Gurevych, I. Automatic Annotation Suggestions and Custom Annotation Layers in WebAnno. – In: *Proceedings of ACL-2014, demo session*. Baltimore: ACL.

Имамура и др. 2004: Imamura, K. Okuma, H. Watanabe, T. Sumita, E. Example-based Machine Translation Based on Syntactic Transfer with Statistical Models. – In: *Proceedings of the 20th International Conference on Computational Linguistics*, 2004, 99 – 105.

Кабакчиев 1992 : Кабакчиев, К. *Видът в английския език*. София: Албо.

Кабакчиев 2003: Кабакчиев, К. *Английска граматика*. София: Пенсофт.

Каланци 2002: Kalantzi, D. Teaching MT/CAT tools in Greece: The State of the Art. – In: *Proceedings of the 2002 EAMT Conference*, 2002 33-43.

Клайн 1994: Klein, W. *Time in Language*. London: Routledge.

Клеър 1966: Clair, C. *History of Printing in Britain*. Oxford: Oxford University Press.

Коева 1998: Коева, Св. *Граматичен речник на българския език. Описание на концепцията за организацията на лингвистичните данни.* [http://lml.bas.bg/~stoyan/gws/dokument1.html] (посетен на 20.02.2020).

Коева 2008: Коева, Св. Българският ФреймНет. Семантико-синтактичен речник на българския език – концептуален модел. – В: *Българският ФреймНет. Семантико-синтактичен речник на българския език*. Коева, Св. (съст.). София: Институт за български език, 5 - 57.

Коева 2009: Коева, Св. Езикови ресурси и компютърни програми с приложение в лингвистичните изследвания – В: *Приложение на информационните технологии в работата на филолога и при изграждането на езикови ресурси*. Бъркалова, П. (съст.) София: Архимед, 57-74.

Коева 2014: Коева, Св. Българският национален корпус в контекста на световната теория и практика – В: *Езикови ресурси и технологии за български език*. София: Академично издателство „Проф. Марин Дринов“, 29-53.

Коева и Генов 2011: Koeva, Sv. Genov, A. Bulgarian language processing chain. – In: *Proceedings of Integration of multilingual resources and tools in Web applications. Workshop in conjunction with GSCL 2011*, Hamburg: University of Hamburg.

Коева и др. 2010: Koeva Sv. Blagoeva, D. Kolkovska, S. Bulgarian National Corpus Project. – In: *Proceedings of LREC-2010, Valletta*, ELRA, 3678-3684.

Коева и др. 2012а: Koeva, Sv. Stoyanova, Iv. Leseva, Sv. Dimitrova, Tsv. Dekova, R. Tarpomanova, E. The Bulgarian National Corpus: Theory and Practice. – In: *Corpus Design". Journal of Language Modelling*, 2012, 65–110.

Коева и др. 2012б: Koeva, Sv. Rizov, B. Tarpomanova, E. Dimitrova, Tsv. Dekova, R. Stoyanova, I. Leseva, Sv. Kukova, Hr. Genov, A. Application of Clause Alignment for Statistical Machine Translation. – In: *Proceedings of SSST-6, Sixth Workshop on Syntax, Semantics and Structure in Statistical Translation*, Jeju, Republic of Korea, The Association for Computational Linguistics, 2012, 102-110.

Коен 2010: Koehn, P. *Statistical Machine Translation*. Cambridge: Cambridge University Press.

Коен и др. 2003: Koehn, P. Och, F. Marcu, D. Statistical Phrase-Based Translation. – In: *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, 2003, 127-133.

Комри 1981: Comrie, B. *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago: University of Chicago Press.

Комри 1985: Comrie, B. *Tense*. Cambridge: Cambridge University Press.

Корнацки 2018: Kornacki, M. *Computer-Assisted Translation (CAT) Tools in the Translator Training Process.* Lodz: University of Lodz.

Крофт 1991: Croft, W. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press.

Куакенбос 1869: Quackenbos, G, *An English Grammar*. [https://archive.org/details/englishgrammar00quac] (посетен на 20.02.2020).

Курме 1930: Curme, G. *A Grammar of the English Language*. (1978) Michigan: Gale.

Куцаров 2007: Куцаров, И. *Теоретична граматика на българския език: морфология*. Пловдив: Университетско издателство „Паисий Хилендарски".

Куцаров и др. 1999: Куцаров, Ив. Бояджиев, Т. Пенчев, Й. *Съвременен български език (фонетика, лексикология, морфология, синтаксис)*. София: Изток – Запад

Куърк и др. 1985: Quirk, R. Greenbaum, S. Leech, G. Svartvik, J. *A Comprehensive Grammar of the English Language*. London: Longman.

Лагудаки 2006: Lagoudaki, E. *Translation Memory Systems: Enlightening users' perspective*. London: Imperial College London.

Лазаров 2016а: Лазаров, Т. Особености на глаголните системи и начините за изразяване на времето в български и английски. Семантичен трансфер при превод на глаголните форми от български на английски – В: *Littera et Lingua*, 2, 2016 [https://naum.slav.uni-sofia.bg/lilijournal/2015/12/1-2/tlazarov] (посетен на 20.02.2020).

Лазаров 2016б: Лазаров, Т. Анализ на ресурсите за статистически модел за превод на глаголните форми от български на английски“ – В: *Български език*, бр. 4, 2016, 96 – 102 .

Лазаров 2017: Лазаров, Т. Функционални граматични паралели при превода на глаголните форми от български на английски – В: *Доклади от Международната юбилейна конференция на Института за български език „Проф. Любомир Андрейчин (15/16 май 2017)*“, 321-327.

Лазаров 2018: Lazarov, T. Bulgarian-English Parallel Corpus for the Purposes of Creating Statistical Translation Model of the Verb Forms. General Conception, Structure, Resources and Annotation. – In: *Proceedings of the Third International Conference Computational Linguistics in Bulgaria*, 2018, 192-201.

Лазаров 2019а: Лазаров, Т. Съвременно състояние на Българско-английския паралелен корпус със съотнесени глаголни форми – В: *Български език,* бр. 1, 2019, 103 – 113.

Лазаров 2019б: Лазаров, Т. Българо-английски английския паралелен корпус със съотнесени глаголни форми – В: *Съпоставително езикознание*, бр. 2, 2019, 5 – 15.

Лийч 1991: Leech, G. The state of the art in corpus linguistics. – In: *English Corpus Linguistics*, London: Longman, 8 – 29.

Лионс 1977: Lyons, J. *Semantics*. Camridge: Cambridge University Press.

Луис и Мол 2013: Lewis, G. Mol, H. *Grammar for Young Learners*. Ofxord: Ofxord University Press .

Маккарти и Уарингтън 1988: McCarthy, R. Warrington, E. Evidence for modality-specific meaning systems in the brain. – In: *Nature*, 334, 428 – 434.

Манинг и др. 2013: Manning, L. Cassel, D. Cassel, J. C. St. Augustine’s Reflections on Memory and Time and the Current Concept of Subjective Time in Mental Time Travel. – In: *Behavioral Sciences*, 2013, 3, 232 – 242.

Маслов 1956: Маслов, Ю. *Очерк болгарской грамматики*. Москва: Издательство литературы на иностранных языках.

Маслов 1982: Маслов, Ю. *Граматика на българския език*. София: Наука и изкуство.

Миркович 1860: Миркович, Д. *Кратка и методическа българска граматика*. Цариград: Галата.

Младенов 1939: Младенов, Ст. *Граматика на българския език.* София: Казанлъшка долина.

Моллов 2014: Моллов, Б. *Лекции по философия на езика*. София: Проектория.

Момчилов 1847: Момчилов, Ив. *Писменица на славянский язик.* Белград: Правителствен книгопечат.

Мърфи 1991: Murphy, R. *English grammar in use*. Cambridge: Cambridge University Press.

Накимовски 1988: Nakhimovsky, A. Aspect, aspectual class, and the temporal structure of narrative. – In: *Computational Linguistics*, 14(2), 1988. 29 - 43.

Наков 2009: Наков, Пр. Съвременен статистически машинен превод: кратък обзор. – В: сп. *Автоматика и информатика*, 2009, бр. 2 [http://people.ischool.berkeley.edu/~nakov/selected_papers_list/nakov_prevod_sp_Avtomatika_Informatika.pdf] (посетен на 20.02.2020).

Ницолова 1976: Ницолова. Р. Българския глагол. Обзор. – В: Пашов, П. Ницолова, Р. (съст.) *Помагало по българска морфология. Глагол.* София: Наука и изкуство.

Ницолова 2008: Ницолова, Р. *Българска граматика: Морфология*. София: Университетско издателство „Св. Климент Охридски“.

Ницолова 2017: Nicolova, R. *Bulgarian Grammar*. Berlin: Frank & Timme GmbH.

Нюмайър 1983: Newmeyer, F. *Grammatical Theory: Its Limits and Its Possibilities*. Chicago: University of Chicago Press.

Палмър 2014: Palmer, F. *The English Verb*. London: Routledge.

Пантелеева 1997: Пантелеева, Хр. Бъдеще в миналото – т.нар. кондиционал от балкански *тип. Наблюдения над темпоралното му значение в съпоставка с полски.* – В *Общност и многообразие на славянските езици. Сборник в чест на професор Иван Леков.* София: Академично славистично дружество, 1997, 79–85.

Паскалева 1987: Паскалева, Е. *Компютър и превод*. София: Наука и изкуство.

Пашов 1965: Пашов, П. Българските глаголни времена (за основните им значения и онагледяването им със схеми). – в *Народна просвета*, кн. 3, 52 – 63.

Пашов 1999: Пашов, П. *Българска граматика*, София: Издателска къща „Хермес“.

Приор 1967: Prior, A. *Past, Present and Future.* Oxford: Oxford University Press.

Пристли 1761: Priestly, J. The Rudiments of English Grammar, 1761, London: Griffiths.

Путсма 1922: Poutsma, H. *Mood and tense of the English verb.* Groningen: P. Noordhoff.

Радулов 1863: Радулов, С. *Учебник за болгарский язик, преведен и преработен от подобен руски учебник*, 1863, Болград: А. Дамянов.

Райхенбах 1947: Reichenbach, H. *Elements of Symbolic Logic*. New York: The Macmillan Company, 1947, 287-298.

Ранкова и Иванова 1980: Ранкова, М. Иванова, Цв. *Английска граматика*. София: Наука и изксутво.

Рилски 1835: Рилски, Н. *Болгарска грамматика.* 1835, Крагуевац: Братя Мустакови.

Родриго 2008: Rodrigo, E. *Topics in Language Resources for Translation and Localisation*. Amsterdam: John Benjamins Publishing Company.

Савова 2007: Савова, Ив. Категорията текстово време. – В: сп. *LiterNet*, 3 (88) [https://liternet.bg/publish7/isavova/tekstovo.htm] (посетен на 20.02.2020).

Санторини 1991: Santorini, *B. Part-of-Speech Tagging Guidelines for the Penn Treebank Project, Penn Treebank Project Technical Report № 3*, University of Pennsylvania [https://repository.upenn.edu/cgi/viewcontent.cgi?article=1603&context=cis_reports] (посетен на 20.02.2020).

Симов и др. 2004: Simov, K. Osenova, P. Slavcheva, M. *BTB-TR03: BulTreeBank Morphosyntactic Tagset, BulTreeBank Project Technical Report № 03*, [http://bultreebank.org/wp-content/uploads/2017/04/BTB-TR03.pdf] (посетен на 20.02.2020).

Синклер 2005: Sinclair, J. Corpus and Text: Basic Principles. – In: *Developing Linguistic Corpora: A guide to Good Practice*, 2005 [https://pdfs.semanticscholar.org/07fe/f4d946b51c7045c6968066f222f31192a389.pdf] (посетен на 20.02.2020).

Смирницки 1955: Смирницки, Ив. Логическое и грамматическое в слове – В: *Вопросы грамматического строя*, 1955, 11 – 53.

Смит и Хал 1889: Smith, W. Hall, T. *A School Manual of English Grammar with Copious Exercises*. 1889 [https://archive.org/stream/cihm_35024#page/n75/mode/2up] (посетен на 20.02.2020).

Сосюр 1992: Сосюр, Ф. *Курс по обща лингвистика*. 1992, София: Наука и изкуство.

Стамболиева 2015: Стамболиева, М. Класове думи: английско-български паралели – В: *Годишник на департамент „Англицистика”*. София: НБУ.

Станева 2001: Станева, Хр. *Стилистика на българския книжовен език*. София: Абагар.

Станков 2009: Станков, В. Граматично значение и контекст. – В: Станков, В. (съст.) *Проблеми на граматичната система на българския език - глагол.* София: Академично издателство „Проф. Марин Дринов“

Стоянов 1964: Стоянов, Ст. *Граматика на българския книжовен език*. София: Наука и изкуство.

Теодоров-Балан 1958: Теодоров-Балан, Ал. *Нова българска граматика за всякого. За думите. Част III. Глагол*. София: БАН.

Тишева 2014: Тишева, Й. Езикови бази данни, корпуси и електронни ресурси за българската устна реч. – В сп. *Littera et Lingua*, Том: 11 (1-2), 2014 [https://naum.slav.uni-sofia.bg/en/lilijournal/2014/11/1-2/ytisheva] (посетен на 20.02.2020).

Тодорова 1993: Тодорова, Е. *Езикът на българсакта публицистика (1944-1989). Лингвостатистическо изследване*. София: Евтимов-Илинда.

Трифонов 1905: Трифонов, Ю. Синтактични бележки за съединението на минало действително причастие с глагола съм в новобългарския език. – В: *Периодическо издание на българското книжовно дружество в София* (66), 3 – 4, 155 – 192 .

Уайнрих 1973: Weinrich, H. *Le Temps.* Paris: Seuil.

Уард 1767: Ward, W. *An Essay on English Grammar*. London: Bookfeller [https://archive.org/details/agrammarenglish00wardgoog/page/n10] (посетен на 20.02.2020).

Уолт 1762: Lowth, R. *A Short Introduction to English Grammar with Critical Notes*. Philadelphia: Aitken [https://archive.org/stream/shortintroductio00lowtrich#page/34/mode/2up] (посетен на 20.02.2020).

Фабиан 2014: Fabian, J. *Time and the other*. New York: Columbia University Press.

Филипова 2002: Филипова, М. *Практическа английска граматика*. София: ЕМАС.

Фират и др. 2016: Firat, O. Cho, K. Bengio, Y. Multi-Way, Multilingual Neural Machine Translation with a Shared Attention Mechanism. – In: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp 866–875*

Флашман 1991: Fleischman, S. Discourse as space/discourse as time: Reflections on the metalanguage of spoken and written discourse. – In: *Journal of Pragmatics*, (16), 4, 291 – 306 .

Фрейзър 2008: Fraser, A. *Hidden Markov models and dynamic systems*, Philadelphia: Society for Industrial and Applied Mathematics.

Хаус и Стивънс 2000: House, C.; Stevens, J. *Grammar no problem*, 2000, Berlin: Cornelsen.

Хорнщайн 1990: Hornstein, N. *As Time Goes By: Tense and Universal Grammar*, Cambridge: MIT Press.

Хъдълстън и Пулъм 2002: Huddleston, R. Pullum, G. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Хътчинс 1995: Hutchins, W. Machine Translation: A Brief History. – In: Koerner, E. Asher, R. (eds.) *Concise history of the language sciences: from the Sumerians to the cognitivists*. Oxford: Pergamon Press, 431-445.

Черагю 2012: Chéragui, M. Theoretical overview of machine translation. – In: *Proceedings of 4th International Conference on Web and Information Technologies*. 160 – 169.

Чергова 2012: Чергова, В. *Конюнктивният имперфект в съвременния португалски език*. София: Университетско издателство „Св. Климент Охридски“

Чергова 2014: Чергова, В. Йерархичност на категориалните значения и форми – В: Кръстев, Н. (съст.) *Мислене, език, реч.* София: Университетско издателство „Свети Климент Охридски“.

Чомски 1957: Chomsky, N. *Syntactic Structures*. The Hague/Paris: Mouton.

Шмид 1995: Schmid, H. Improvements in Part-of-Speech Tagging with an Application to German. – In: *Natural Language Processing Using Very Large Corpora*. (11), Dordrecht: Springer, 13–22.

Якобсон 1972: Якобсон, Р. Шифтеры, глаголньые категории и руский глагол – В: *Принципы типологического анализа языков различного строя*. Москва: Наука.

Янакиев 1976: Янакиев, М. За грамемите, наричани в българската граматика „сегашно време“ и „бъдеще време“ – В: Пашов, П. Ницолова, Р. (съст.) *Помагало по българска морфология. Глагол.* София: Наука и изкуство.

XLIFF 1.2 2007: *A white paper on version 1.2 of the XML Localisation Interchange File Format (XLIFF)* [https://www.oasis-open.org/committees/download.php/26817/xliff-core-whitepa-per-1.2-cs.pdf] (посетен на 20.02.2020).

Приложения

Описание на представените приложения в електронна форма.

Хиперлинк	Описание
http://www.dcl.bas.bg/bepcavf/	Българско-английски паралелен корпус със съотнесени глаголни форми