

КОРПУСЕН ПОДХОД ПРИ ИЗСЛЕДВАНЕТО НА ЕЗИКОВИТЕ ОТКЛОНЕНИЯ В БЪЛГАРСКИЯ ЕЗИК КАТО ЧУЖД

Даниела Иванова

Шуменски университет „Епископ Константин Преславски“

CORPUS APPROACH IN THE STUDY OF THE LANGUAGE DEVIATIONS IN BULGARIAN AS A FOREIGN LANGUAGE

Daniela Ivanova

Shumen University “Episkop Konstantin Preslavski”

This paper is an attempt to present a comprehensive overview of the learner corpora as a significant tool in the research on foreign language acquisition. It aims to show the possibilities of the corpus approach in the study of language deviations in Bulgarian as a foreign language. Some of the applications and ways to extract information from error-tagged electronic corpora are also provided. The paper contains the initial steps of the project for a learner corpus of Bulgarian as a foreign language.

Keywords: Bulgarian as a foreign language, corpus linguistic, learner corpus

Увод

Изследванията на процесите, свързани с овладяването на чужд език, са много важна част от съвременната приложна лингвистика. Проучванията в тази изследователска област помагат за усъвършенстване на методите на обучение, за оформянето на различни стратегии при преподаването и усвояването на езика, за систематизиране на взаимодействието и влиянието между различните езици. Търсенето на адекватни решения на тези предизвикателства се осъществява главно чрез подход, който събира в едно постиженията на корпусната лингвистика, модерните компютърни технологии, усвояването на чужд език и лингвистичните теории за развитието и промените на езика. В центъра на това интердисциплинарно поле се оказват учебните корпуси – електронни колекции с езикова продукция на изучаващи чужд език. Именно тези специализирани корпуси, в които отрицателният езиков материал е систематизиран, а различни софтуерни програми

позволяват той да бъде анализиран според целите на изследователите, са във фокуса и на настоящата работа.

Авторските усилия са насочени по-конкретно към открояването на необходимостта от прилагането на корпусния подход при изследването на езиковите отклонения в процеса на изучаването на българския език като чужд, като акцентът е поставен върху възможностите, които той предоставя. В следващото изложение е направен опит в контекста на панорамното представяне на постигнатото досега в подобни изследвания да се покажат пилотните стъпки в изграждането на електронен учебен корпус с езикова продукция на български език като чужд. Корпусът се разработва в рамките на Националната интердисциплинарна изследователска Е-инфраструктура (КЛАДА-БГ), интегрирана в проекта CLARIN, с участието на Лабораторията по приложна лингвистика към Шуменския университет „Епископ Константин Преславски“. Изследователската работа в тази специфична област неизменно е свързана с теоретичните основи на анализа на грешките и междинния език, които ще бъдат представени накратко. След преглед на българския опит и традиции в това интердисциплинарно поле, както и на съществуващите учебни корпуси на различни езици, ще бъдат разгледани въпроси, свързани с процеса на съставяне, обработване и използване на такива платформи с лингвистично анотирани бази данни.

На границата между знание и незнание.

Грешката и междинният език

В областта на лингвистиката въпросът за отклоненията от правилната употреба на езика и за кодифицирането на неговите норми има дълга история. Изследването на този проблем често изисква интердисциплинарен подход заради сложността и комплексния му характер. Необходимостта от изучаването на *грешките*, или т.нар. „отрицателен езиков материал“, е изтъквана още от руския лингвист Л. В. Шчерба (Шчерба/Shcherba 1974: 24). С термина „грешки“ той обозначава различни отклонения от езиковите норми, подчертавайки важното му значение при изследването и моделирането на процеса на овладяването на родния език от децата, както и при усвояването на чужд език.

Отрицателният материал се използва за изграждането на адекватни модели на усвояването на езика. В тези модели грешките имат много важна роля, защото анализът им дава ценни сведения не само за етапите на усвояването, а и за йерархическата организация от правила, чрез които функционира езикът.

От изследванията на езиковото развитие в детската възраст, които приемат грешките като част от временна и динамична езикова система, е повлиян и подходът на П. Кордър. Той ги разглежда като източник на много важна информация и за процеса на овладяване на втори език (Corder/Кордър 1967: 161).

Според възприетата сред българските лингвисти терминология понятието „междинен език“ обхваща специфична самостоятелна езикова система, която се разглежда като трети език, съществуващ наред с родния и чуждия език. Междинният език може да бъде определен като динамична картина на процеса на овладяване на новия език. Тя е съставена от правила и стратегии на изучаващите езика, които Л. Селинкър обобщава като опростяване, трансфер и свръхгенерализация (Selinker/Селинкър 1972: 210). Специфична черта на междинния език е, че в речевата продукция се наблюдава смесване както на грешни (с оглед на езика цел), така и на нормативни (цели) единици. Той се характеризира и със специфична динамика, тъй като придобиването на езиковата компетентност не става изведнъж, а преминава през няколко последователни етапа, в хода на които самият междинен език търпи постоянни промени. Това дава основание той да се разглежда и като своеобразен феномен, представляващ динамично единство между родния език и езика цел.

С коментираните дотук идеи се свързват задачите на проекта за изграждането на учебен корпус с речеви данни на лица, за които българският е чужд. Този електронен езиков масив с анотирани данни ще бъде основен инструмент в изследването на определени специфични особености на междинния български език. Очакваният резултат е открояването на морфологичните трудности в процеса на усвояването му, както и адекватната оценка на тяхното значение в процеса на комуникацията.

Българският опит и традиции в корпусната парадигма

Съвременният български език е представен в няколко големи корпуса, сред които е важно да се спомене Българският национален корпус, създаден от екип на Института за български език. Това е най-големият, корпус с фокус върху български език, в който корпусните единици са снабдени с подробни метаданни и разнообразна лингвистична анотация (Български национален корпус – <http://dcl.bas.bg/bulnc/>). Представителни данни за развитието на езика и неговите варианти съдържат и редица езикови ресурси, които оформят един своеобразен дял в духа на българската корпусна тради-

ция – корпусите с устна българска реч. Тези бази данни, част от които вече са и в електронен вид, представят езика в неговата устна форма с цялото му разнообразие и трансформации. В тях разговорната реч присъства в автентичния си вид, често много отдалечена от книжовните норми. Тук могат да бъдат открити данните, публикувани в рамките на инициативата BgSpeech на www.bgspeech.net с материали, събирани и транскрибирани от различни екипи на Факултета по славянски филологии на Софийския университет. В рамките на проекта е публикуван и първият мултимедиен корпус на българската устна реч, обединяващ аудио- и видеозаписи с писмени записи (транскрипции) (Тишева/Tisheva 2014).

Специфично поле в тази област представляват корпусите, които са свързани с представянето на масиви от данни с т.нар. отрицателен езиков материал. Корпусната перспектива при организирането и обработката на голям обем емпиричен материал е във фокуса на изследването на българската детска реч (Попова/Popova 2017). Първият български електронен корпус с детска реч, разработен в термините на системата CHILDES, вече е част от базата данни на едноименната платформа (Bulgarian LabLing Corpus – <https://childes.talkbank.org/access/Slavic/Bulgarian/LabLing.html>). Той е в основата на обширно изследване на ранната езикова онтогенеза. С широката приложимост на корпусния подход В. Попова разкрива тайните на детската граматика, като успява да присъедини българската детска реч към многоезичната мозайка, събрана в проекта CHILDES.

Изследванията, свързани с усвояването на българския като чужд, също имат своите добри традиции у нас. Корпусният подход, обработката и анализът на отрицателен езиков материал са сред основните в това изследователско поле. С езиковата асиметрия и с анализа на грешки, събирани в различни по обем езикови масиви, се занимават редица учени. В свое изследване Г. Куртева прави опит за типология на грешките и очертаването на ареала на трудностите и проблемните области при усвояването на някои граматични категории на българския език от чуждестранни обучаеми. Във фокуса на вниманието попада именната система на българския език (Куртева/Kurteva 2016). В монографията си, посветена на усвояването на българския език като чужд, С. Банова изследва употребата на притежателните местоимения (клитични и пълни форми) и на приглаголните клитики (Банова/Banova 2018). Неслучайно в психолингвистиката също се отделя достойно внимание на усвояването на българския език като втори. Сред учените, които се занимават с този проблем, е Ю. Стоянова.

Тя представя проблематиката на научното поле, познато като „овладяване на втори език“ (Second Language Acquisition), и я свързва с усвояването на българския език като втори или чужд (Стоянова/Stoyanova 2014).

Въпреки наличието на редица изследвания в тази област, на систематизирани масиви от данни с емпиричен материал и на корпуси с писмена и разговорна реч, до този момент българският език не е представен в електронен учебен корпус, за разлика от повечето западноевропейски и много от славянските езици.

В следващия параграф ще се потърси отговор на въпроса необходимо ли е създаването на учебен корпус на българския език, като при това във фокус ще се окажат досегашни приноси в научната традиция и тяхното изключително значение като инструменти за лингвистичен анализ.

Учебният корпус като инструмент за лингвистичен анализ

Белгийската лингвистка С. Грейнджър от Университета в Льовен определя появата на учебните корпуси като революция в приложената лингвистика. Те свързват в едно общо поле корпусната лингвистика и изследванията за овладяването на втори (чужд) език, като предоставят модерен и работещ ресурс за различни по рода си проучвания (Granger, Hung, Petch/Грейнджър, Хънг, Печ 2002: 3).

Учебните корпуси започват да се появяват в базите данни на корпусната лингвистика през 90-те години на миналия век. Това са големи електронни колекции с текстове, автори на които са хора, изучаващи чужди езици. В самото начало езикът най-често е английски, след време езиците, които са обект на изследване, стават все повече. Така се изработват корпуси с немски, френски, чешки, унгарски, руски, арабски, китайски и други езици, изучавани като чужди. Корпусите обикновено съдържат документирани писмени текстове (есета, диалози, статии, граматически задачи и упражнения) или устна реч, записана с помощта на аудиотехника. Участниците са предимно млади хора, ученици или студенти. Причините за доминирането на английския като втори език са разбираеми – нарастващата необходимост от владееене на този международен език в съвременния свят, както и нуждата да се усъвършенства процесът на неговото усвояване. Изследователите откриват в обширните бази данни богат ресурс за аотирание, систематизиране и анализ на грешките, допускани по време на обучението от носителите на различни изходни езици. Модерните компютърни технологии помагат събраният емпиричен материал да

бъде обработен, анализиран и използван за различни цели. Променя се подходът при създаването на учебни материали, речници, помагала, както и цялостната методика на преподаването на чужд език.

В подкрепа на казаното дотук вниманието ще бъде насочено към няколко от най-значимите функциониращи учебни корпуси в света. Към момента, по информация на Международната асоциация на учебните корпуси на различни езици (Learner Corpus Association), броят им е над 100. Повече от половината са англоезични, като изходните езици на обучаемите варират. Броят на проектите непрекъснато се увеличава и сред езиците са френски, шведски, норвежки, немски, полски, чешки, руски. В процентно съотношение разликите изглеждат така:

- Английски език – 57%
- Френски език – 17%
- Испански език – 13%
- Немски език – 10%
- Италиански език – 3%

Първият и най-известен учебен корпус е The International Corpus of Learner English (ICLE), създаден от група лингвисти в Университета в Льовен, Белгия. Той е съставен от есета, написани от изучаващи английски език от различни националности със средно до високо ниво на владение на изучавания език (upper intermediate, advanced). Версия 1.1 ICLE, издадена на CD-ROM през 2002 г., съдържа над 2,5 милиона думи в 3640 текста, написани от студенти от 11 страни (България, Чехия, Холандия, Финландия, Русия, Франция, Германия, Италия, Полша, Испания, Швеция). Корпусът продължава да се допълва с подкорпуси и в него се включват участници от още страни. Посочената версия не е морфологично аотирана.

Другият не по-малко известен мащабен учебен корпус на английски език е The Cambridge Learner Corpus (<http://www.cambridge.org/corpus>). Той съдържа над 30 млн. думи и непрекъснато се допълва с примери от анонимни изпитни тестове (Cambridge ESOL English) на участници от 190 страни по света. Корпусът е аотиран, грешките са анализирани, съдържа информация за родния език (E1) на участниците, националност, възраст, ниво на владение на английски език. Той може да бъде използван само от автори и изследователи, работещи в Кеймбриджкия университет.

Първият учебен корпус с устна реч е LINDSEI и е копие на The International Corpus of Learner English (ICLE). Първоначалната му версия съдържа стенограми на петдесет 15-минутни интервюта със сту-

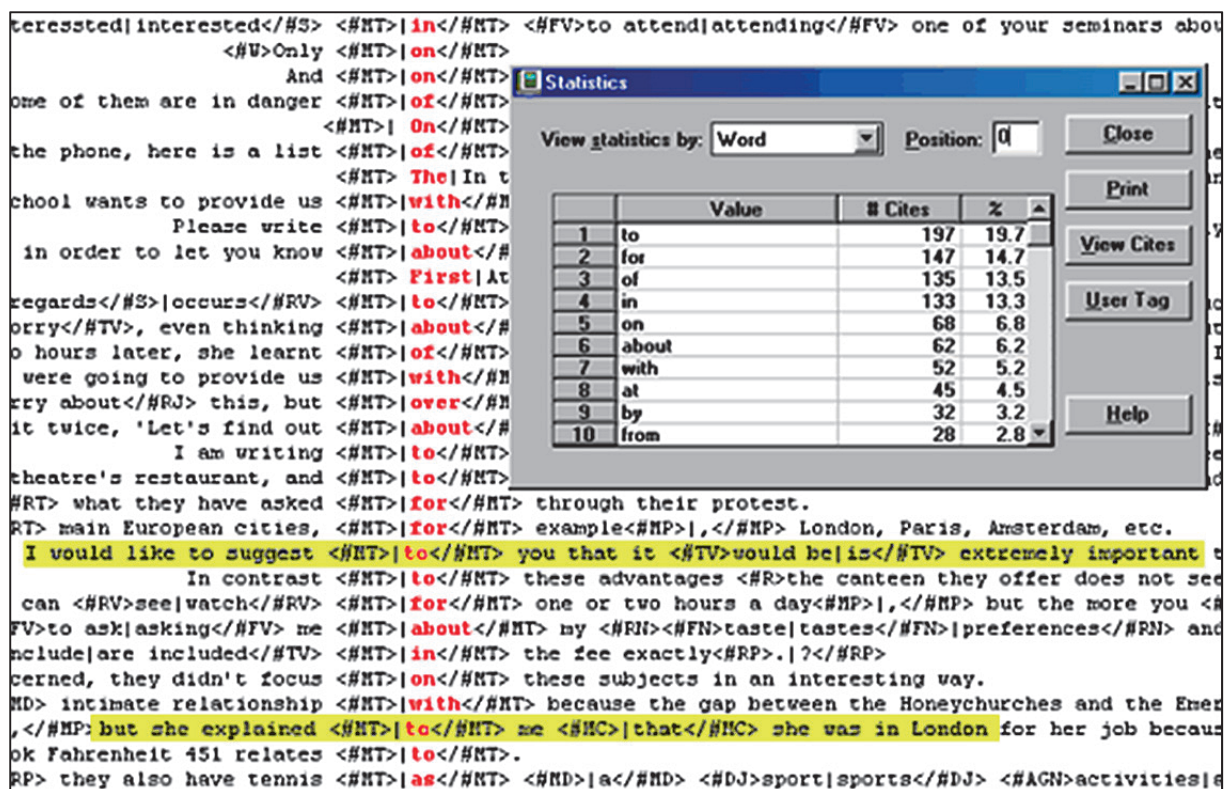
денти с високо ниво на владеене на английски. Базата данни непрекъснато се разширява. Доколкото учебните корпуси използват предимно писмена реч, колекцията в LINDSEI дава нов, по-широк диапазон от възможности за изследвания на особеностите на устната реч в условията на интернационална комуникация.

Руският учебен корпус (RLC, <http://web-corpora.net/RLC/>) е създаден през 2013 г. Основа на корпуса са устни и писмени текстове на две категории носители на руския език – чужденци, изучаващи руски като чужд, и билингви с роден език руски, който обаче не се използва като водещ език. Метаданните включват информация за изследваните лица (пол, възраст, първи език, ниво на владеене на руски) и за текстовете, включени в колекцията (писмен, устен, жанр). Корпусът е анотиран за различните типове отклонения от стандартния руски език. Данните и системата за търсене са основа за съпоставителни изследвания на текстовете на билингвите и чуждестранните носители на езика. Корпусът дава възможност да се проучат такива сложни явления и категории като вида на глагола, падежите, предложните конструкции в контекста на усвояването на руския език като чужд. Анализират се неправилното използване на думите или погрешното разбиране на смисъла в словосъчетанията и по-големите конструкции в рамките на изречението.

Има още много други корпуси, използвани за различни цели, които могат да дадат представа за същността на подобен тип масиви и за техните възможности. Инфраструктурата на проекта CLARIN например осигурява достъп до 74 учебни корпуса (<https://www.clarin.eu/resource-families/L2-corpora>). Сред тях 11 са многоезични, а останалите предоставят писмени, устни и дори видео данни в едноезични корпуси на арабски, чешки, английски, финландски, френски, немски, унгарски, норвежки и шведски. Много от корпусите са публично достъпни в интернет, а част от тях са интегрирани в CLARIN. Те са представени като важна част от теорията и практиката на изследванията и методиката на обучение по чужд език, защото позволяват систематично проучване на усвояването на новия език на лексикално и синтактично ниво, както и проследяването на влиянието на родните езици при тези процеси. Според участниците в проекта едно от основните предимства на тези бази данни е маркирането на грешките, което позволява различна по характер и обем работа със събрания материал.

Възможности на учебните корпусни платформи

С данните от подобни електронни масиви могат да се систематизират етапите на усвояване на езика. Изследователите от Университета в Кеймбридж например използват корпуса, за да опишат трудностите при изучаване на английски език и да проверяват коректното оценяване по време на изпитите. Системата за кодиране на грешки позволява да бъдат проследени думи, конструкции, граматични категории с най-висока честотност на отклонения от нормата на езика и да се търсят конкретни грешки. Ето един пример: думите в червено са предлозите, които учениците трябва да използват, но са изпуснали, а <#MT> е кодът за липсващ предлог (вж. Фиг. 1).



Фигура 1. The Cambridge learner corpus

В сивия прозорец е предложена статистическата информация, която еднозначно демонстрира кои от предлозите се пропускат най-често. Тук “to” и “for” липсват съответно 197 и 147 пъти в извадка от 1000 цитата. Програмата може да посочи и думите, след които най-често се пропускат предлози (в тази извадка това е глаголят “explain”). Анотираният за грешки корпус дава важно предимство и при търсенето на конкретни типове отклонения от нормата или на групи грешки, като това се случва по същия начин, по който търсим отделни думи.

Друго полезно приложение на данните от корпуса е изследването на фреквентността на грешките при различни нива на усвояване, както и при различни изходни езици. Пример за търсене на проблеми при съгласуването на глагола по лице, число и род е представен в извадката от руския учебен корпус (RLC – Russian Learner Corpus). В изречението „Все хвалили её за доброту, мать только **сожалел** о том, что девочка была также близорука“ е маркирана неправилната употреба на глагола при съгласуването със съществително име от женски род. В две допълнителни полета са показани правилната форма на глагола в дадената конструкция (**сожалела**) и инфинитива му (**сожалеть**). Метаданните маркират основния език и нивото на владеене на руския на всеки от авторите, което позволява систематизиране и групиране на типовете грешки (вж. Фиг. 2).

мне тепло и удобно.

2. summary (text for oral presentation) (non-timed, paragraph, Russian in the Major), Anna (F, HL, AL) | eng | HL | 2009
Самая главное эта специальность актуальна почти во всех **рабочих сферах**. <...>
 Самое главное, что эта специальность актуальна почти во всех **профессиональных сферах**.

3. Антон и его бабушка, Manon | fr | FL | None
 Но время **шла**, он **возра** m / f, Verb да мог поговорить ее погугль. <...>
 Но время **шло** он **рос** и всегда

4. Антон и его бабушка, Ma Syntax AgrGender
 Но время **шла** и ему уже **сожалела** садиться чтобы поговорить с бабушкой. <...>
 Но время **шло** и ему уже надо было садиться, чтобы поговорить с бабушкой.

5. О Красной Шапочке, Cosme | fr | FL | None
 Все хвалили её за **доброту**, мать только **сожалел** о том что девочка была **также** близорука. <...>
 Все хвалили её за **доброту** мать только **сожалела** о том что девочка была **также** близорука.

6. О себе, Manon | fr | FL | None
Команды этой спорта в **высшей** школе э **амические** – и мужская и женская. <...>
 Команды **этого** спорта в **Высшей** школе экономики очень **динамичные** и мужская и женская.

7. О себе, Manon | fr | FL | None
 У них турнир каждый **выходни** в разных **местах** и это мне **помогал** открыть Россию **на первом семестре**. <...>
 У них турнир каждый **выходные** в разных **местах** и это мне **помогло** открыть Россию **в первом семестре**.

8. О себе, Manon | fr | FL | None
 У меня еще **два** **старые** **сестры**, и все мы очень **близкие**. <...>
 У меня еще **две** **старшие** **сестры** и все мы **близкие**

Фигура 2. Russian Learner Corpus

Разнообразните инструменти за работа в езиковите корпуси (напр. честотни списъци, конкорданси, парсери) не само помагат за анализа и типологизацията на грешките в събраните текстове, те могат да формират и актуални речници на междинния език на обучаваните. Лесно е да се установи и актуалната компетентност в различни етапи от усвояването на чуждия език. Автори, редактори и лексикографи използват международния корпус ICLE, като търсят примери за

употребата на езика от обучаващите се, както и кои думи, модели и граматически структури се прилагат успешно. Така те могат лесно да установят кои категории предизвикват затруднения.

Първи стъпки по изграждането на български учебен корпус в системата на Кларин

В настоящата работа беше направен опит да се представят функциониращите в момента учебни корпуси на различни езици и тяхното значение в съвременната лингвистика и в обучението по чужд език. В този план бе открито необходимостта от създаването на български учебен корпус, който да запълни липсващото звено в българското езиковознание. Подобен корпус би могъл да обогати палитрата от разноезични учебни корпуси в динамичното поле на световната лингвистика.

Процесът на работа по създаване на учебния корпус на български език е в своята начална фаза. При тези първи стъпки е събрана писмена продукция на чуждестранни студенти с нива на владеење на български от А1 до В1 (според Европейската езикова рамка) Източници на ексцерпирания материал са писмени работи, есета, тестови задачи, диалози, презентации, създавани по време на обучението. Участници са студенти от англоезични програми по медицина в България. Езикът посредник в обучението по български е английският, а изходните езици на изследваните лица са различни – английски, немски, италиански, испански, турски, гръцки, фарси, руски, украински, индийски, полски и др.

Акумулирането на данни все още продължава, тъй като създаването на адекватни модели на човешкото речево поведение и на процеса на езикоусвояването би могло да се осъществи единствено върху солидна и надеждна емпирична база.

Следващ етап от проекта е изграждането на многофункционална платформа, която ще осигури интерфейс за систематизиране на езиковата продукция, както и възможности за многопластово аотиране и търсене. За да бъде използван успешно, всеки корпус трябва да предоставя информация за произхода и специфичните характеристики на продукцията, която е представена в него. Това са метаданните, с които обикновено са аотирани учебните корпуси. Корпусът с български език на чуждестранни студенти ще бъде анонимизиран, но ще предоставя информация за пола на изследваните лица, тяхната националност, първи език, ниво на владеење на български език и др. Добре организирани метаданни ще позволят различни опции за търсене и ще дадат широки възможности за анализ на данните. Учебният корпус на

българския език ще бъде и морфологично аотиран, което ще даде възможност писмените текстове да бъдат сегментирани на единици, означени с граматическите характеристики за всяка част на речта. Освен с класическия за стандартните корпуси морфологичен анализ, корпусът ще бъде обработен и чрез анализ на грешките. Това ще позволи извличането на разнообразна информация и ще предложи данни за количествени изследвания върху усвояването на българския език като чужд.

Заклучение

В предложената работа беше направен опит за панорамно представяне на част от учебните корпуси по света. Сред тях бяха открити няколко значими проекта и ползите им като важен инструмент за лингвистичен анализ, който повишава качеството на изследванията и улеснява работата на изследователите. Като илюстрация на възможностите, предоставяни от електронните масивиот данни, бяха представени и част от начините за работа и извличане на информация от тези инструменти. Изграждането на български корпус от този тип с аотиран данни и инструменти за търсене ще очертае нови възможности за изследвания в областта на езиковото овладяване.

БИБЛИОГРАФИЯ

- Банова/Ванова 2018:** Банова, С. *Езиково усвояване. Вариации в параметрите при морфосинтактични реализации в междинния български език.* София: УИ „Св. Климент Охридски“, 2018. [Banova, S. *Ezikovo usvoyavane. Variatsii v parametrite pri morfosintaktichni realizatsii v mezhdinniya balgarski ezik.* Sofia: UI „Sv. Kliment Ohridski“, 2018.]
- Granger, Hung, Petch/Грейнджър, Хънг, Печ-Тисон 2002:** Granger, S., J. Hung, S. Petch-Tyson. A Bird's-eye view of learner corpus research. // *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching.* Amsterdam & Philadelphia: Benjamins, 3 – 33.
- Corder/Кордър 1967:** Corder, S. P. The significance of learner's errors. // *International Review of Applied Linguistics*, Vol. 5, 1967, 161 – 170. <<https://eric.ed.gov/?id=ED019903>> 20 may 2019.
- Куртева/Kurteva 2016:** Куртева, Г. *Реализация на категориите род, число и определеност в номиналната система на българския език като чужд през призмата на отклоненията от нормативната граматика.* София: Нов български университет, 2016. [Kurteva, G. *Realizatsiya na kategoriite rod, chislo i opredelenost v nominalnatasistema na balgarskiya ezik kato chuzhd prez prizmata na otkloneniyata ot normativnata gramatika.* Sofia: Nov balgarski universitet, 2016.]

- Попова/Ророва 2017:** Попова, В. *Ранна онтогенеза на събитийната модалност (върху материал от българска детска реч)*. Шумен: УИ „Епископ Константин Преславски“, 2017. [Popova, V. *Ranna ontogeneza na sabitiynata modalnost (varhu material ot balgarska detska rech)*. Shumen: UI „Episkop Konstantin Preslavski“, 2017.]
- Selinker/Селинкър 1972:** Selinker, L. Interlanguage. // *International Review of Applied Linguistics*, Vol.10, 1972. <https://www.academia.edu/21533333/Selinker_Interlanguage>, 20 may 2019.
- Стоянова/Stoyanova 2014:** Стоянова, Ю. Овладяване на втори език: проблеми на теорията и практиката. // *Littera et Lingua – Series, Dissertations*. София, 2014 <<http://slav.uni-sofia.bg/naum/liliserries/diss/2014/6>>, 20 march 2019 [Stoyanova, Yu. Ovladyavane na vtori ezik: problem na teoriyata i praktikata. // *Littera et Lingua – Series, Dissertations*. Sofia, 2014 <<http://slav.uni-sofia.bg/naum/liliserries/diss/2014/6>>, 20 march 2019]
- Тишева/Tisheva 2014:** Тишева, Й. Езикови бази данни, корпуси и електронни ресурси за българската устна реч. // *Littera et Lingua*. София, 1 – 2, 2014 < <https://naum.slav.uni-sofia.bg/lilijournal/2014/11/1-2/ytisheva> >, 12 may 2020. [Tisheva, Y. Ezikovi bazi danni, korpusi i elektronni resursi za balgarskata ustna rech // *Littera et Lingua*. Sofia, 1 – 2, 2014 < <https://naum.slav.uni-sofia.bg/lilijournal/2014/11/1-2/ytisheva> >, 12 may 2020.]
- Шчерба/Shcherba 1974:** Щерба, Л. В. *Языковая система и речевая деятельность*. Москва: Наука, 1974. [Shcherba, L. V. *Yazykovaya sistema i rechevaya deyatel'nost'*. Moskva: Nauka, 1974.]