

СОФИЙСКИ УНИВЕРСИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“
ФАКУЛТЕТ ПО КЛАСИЧЕСКИ И НОВИ ФИЛОЛОГИИ
КАТЕДРА „АНГЛИЦИСТИКА И АМЕРИКАНИСТИКА“

Румяна Любомирова Благоева

**СВЪРЗВАЩИ ДУМИ И ИЗРАЗИ В ПИСМЕНИЯ
БЪЛГАРСКИ-АНГЛИЙСКИ МЕЖДИНЕН ЕЗИК
НА НИВО НАПРЕДНАЛИ – КОНТРАСТИВНО
КОРПУСНО ИЗСЛЕДВАНЕ**

Автореферат
на дисертация за присъждане
на образователната и научна степен
„Доктор“
Научна специалност: 2.1. Филология
(Германски езици - английски език)

Рецензенти:

Проф. д-р Мария Георгиева
Проф. д-н Лиляна Грозданова

София, 2016

Дисертационният труд е обсъден и предложен за публична защита на заседание на катедра „Англицистика и американистика“ при ФКНФ на СУ „Св. Климент Охридски“ от 27.10.2015 г. Дисертационният труд е написан на английски език и се състои от увод, четири глави, заключения, шест приложения и списък на цитираната литература.

Защитата на дисертационния труд ще се състои на 05.04.2016 г.
от 14:00 ч. в Зала 2 на СУ „Св. Климент Охридски“,
бул. „Цар Освободител“ №15, гр. София.

Съдържание

Автореферат.....	3
Библиография на цитираната в автороферата литература.....	31
Справка за приносните моменти в дисертацията.....	33
Публикации по темата на дисертацията.....	34
Доклади по темата на дисертацията, изнесени на конференции и симпозиуми	35
Цитирания на някои от публикациите по темата на дисертацията от други автори	36

АВТОРЕФЕРАТ

Представената тук дисертация „Свързващи думи и изрази в писмения български-английски междинен език на ниво напреднали – контрастивно корпусно изследване” е с обем от 286 страници и се състои от увод, четири глави, заключения, шест приложения и списък на цитираната литература, включващ 158 източника. Основната част съдържа 47 таблици и 16 графики.

В увода (1-13 стр.) е очертан основният проблем, възникнал от практиката, към чието разрешаване са насочени целите и задачите, поставени в дисертацията. А именно, защо писмената продукция на английски език, създавана от българи, достигнали високо ниво на владенето му, която не съдържа съществени грешки на лексикално, колокативно и синтактично ниво, се различава от начина, по който биха се изразили естествените носители на езика в подобна ситуация.

Уводът също така разглежда състоянието на изследвания проблем. Направен е кратък преглед на предходни изследвания на българския английски междинен език (БАМЕ) от края на 70-те години на 20-ти до първото десетилетие на 21-ви век. Отбелязва се фактът, че след бума на изследванията на БАМЕ през 80-те години, се забелязва рязък спад на публикациите в тази област. Направените през споменатия период изследвания поставят стабилна основа за развитието на анализа на грешки в нашата страна и насочват вниманието на теоретици и практики към реалните нужди на българите, изучаващи английски език. За съжаление обаче липсата на технически средства за събиране, съхраняване и автоматизирана обработка на данни за БАМЕ през онези години ги прави твърде трудоемки и фокусирани основно върху „грешни“ лексикални и граматически структури, при което се маргинализират проблемите, свързани с организацията на цялостни текстове.

С напредъка на технологиите и развитието на корпусната лингвистика през 90-те години на 20-ти и първото десетилетие на 21-ви век интересът към изследванията на английските междинни езици на учещи с различен роден език се възроди в световен мащаб, докато България остана встрани от тези процеси. С настоящата разработка се поставя началото на нов етап в изучаването на БАМЕ в съответствие със съвременните тенденции в контрастивния анализ на междинните езици (Contrastive Interlanguage Analysis), който вече е един от най-широкопризнатите методи в корпусната лингвистика.

Основна **цел** и **предмет** на изследването са някои текстуални характеристики на БАМЕ и, по-специално, поведението на свързващите думи и изрази като най-експлицитни езикови маркери на текстова организация, които са устойчиви по отношение на тематичните вариации на текстовете. Обективното изучаване на употребата на свързващи думи и изрази в междинния език хвърля светлина върху идеите, които изучаващите имат за постигане на кохезия и единство на текста, и върху стратегиите, които те използват при писане на чуждия език.

Изборът именно на текстуални, а не на други характеристики на междинния английски език на напреднали учещи е продиктуван не само от липсата на подобни изследвания в България, но също и от факта, че в чужбина автори като Грейнджър (1994), Грейнджър и Тисон (1996), Алтенберг и Тапер (1998), Дамаскели (1999), Кашубски (2000) и др., на базата на своите количествени анализи доказват, че неестествените носители на английски език са склонни да разчитат предимно на употребата на свързващи думи и изрази за постигането на логическа последователност в своите текстове. Настоящата разработка е част от тези изследвания и резултатите от нея са съпоставими с международните стандарти и критерии в тази област.

В увода са представени и основните **изследователски въпроси** и **задачи** на дисертацията.

Изследователските въпроси са:

- В каква степен употребата на логически свързващи думи и изрази в писмената продукция на българи, учещи английски език на ниво напреднали, се отклонява от тази на естествените носители на английския език?
- Какви са възможните причини за наблюдаваните отклонения, ако има такива, и дали те са свързани с трансфер от родния език, или със стратегии на преподаване, учене и комуникация?
- Какви коригиращи действия може да се използват в процеса на чуждозиково обучение за подобряване на писмената продукция на българите, учещи английски?

Конкретните задачи, които са изпълнени, за да се постигне целта на изследването и да се отговори на тези въпроси, са:

1. Разглеждане на ролята на големите електронни корпуси за лингвистичните изследвания и обосноваване на необходимостта от тяхното използване с цел получаване на валидни, надеждни и обективни данни за предмета на изследване.

2. Описване на процеса на събиране на първия (и все още единствен) електронен корпус от аргументативни есета на български учещи на ниво напреднали, както и на проблемите, съпътстващи тази дейност.
3. Осъществяване на количествен сравнителен анализ на 50 високочестотни свързващи думи и изрази от писмения БАМЕ на напреднали и от данни за употребата им в писмената продукция на естествените носители на английски език, за да се установят разликите в честотата на употреба и да се набележат евентуалните области, изискващи по-сериозно внимание при преподаването на английски език.
4. Извършване на качествен анализ на няколко свързващи думи и изрази, избрани според тяхната честотност и степен на отклонение от езика-цел, и изследване на тяхното поведение в съответните контексти на употреба, за да се разкрият и сравнят техните текстуални колигации в писмения БАМЕ на напреднали с писмената продукция на естествените носители на английски език.
5. Предлагане на насоки за подобряване на преподаването на академично писане на българи, изучаващи английски език.

В края на уводната част е описана структурата и обхвата на изследването и е резюмирано съдържанието на отделните глави.

Глава първа на дисертацията (14 - 35 стр.) очертава **теоретичната рамка** на изследването, която има три основни аспекта, обусловени от предмета и целите на изследването: първият аспект е свързан с приложната лингвистика, чуждоезиковото обучение и изучаването на характера на междинния език; вторият аспект засяга общата теория за текста и очертава модела за изучаване на характеристики на текстуалността, който се използва в изследването; а третият аспект се отнася конкретно до свързващите думи и изрази в системата от средства за постигане на кохезия в писмения английски език.

Като се изхожда от теорията на Селинкър (1972, 1992) за междинния език се отбелязва, че голямата степен на подобие между езика-цел и междинния език на най-високите нива на чуждоезиково усвояване води до измамното впечатление сред учещи и преподаващи, че след достигане на определена комуникативна компетентност няма място за развитие на езиковите умения. От една страна, учещите нямат усещане за напредък и това е доста обезкуражаващо; от друга страна, преподаващите срещат затруднения при разработването на адекватни

програми за по-нататъшно чуждоезиково обучение, особено по отношение на продуктивните езикови умения на учещите.

Всъщност най-големите предизвикателства пред изследователя на междинния език са именно тези високи нива на владеене на чуждия език, където „грешките“ не са непосредствено очевидни и въпреки това разликите между междинния език и езика-цел все още съществуват. Двойно по-голямо предизвикателство е да се анализират несъответствия от текстуален характер, където понятия като „правилно“ и „грешно“, изглежда, са нерелевантни. В тази връзка, на базата на „психологически релевантни данни от изучаването на чужд език“, т.е. „изказвания на езика-цел от естествените му носители и на родния и междинния език на учещите“ (Селинкър, 1972: 229), получени чрез съвременните методи на корпусната лингвистика, може да се открият обективни и надеждни доказателства за такива несъответствия, което да хвърли светлина върху проблемите на писмената продукция на българи, учещи английски език на най-високо ниво, с цел подобряване и усъвършенстване на учебните програми и материали по английски език.

По отношение на аспектите, свързани с текстуалността, дисертационният труд приема позицията на Богранд и Дреслер (1981) за текста като процес, а не като краен продукт. Само при такъв подход качествата на текстовете на учещите може да се преценяват от гледна точка на това дали изпълняват своята комуникативна функция и максимално се приближават до текстовете, създавани от естествените носители на езика в подобни ситуации, а не дали организацията им е „грешна“ или „вярна“. В такъв смисъл изследването е насочено към търсенето на прекомерното или недостатъчно използване на определени езикови единици, които маркират организацията на текста. И пак в този смисъл изследването не е връщане към „анализа на грешки“, а е стъпка към „контрастивния анализ на междинния език“.

Избраният конкретен теоретичен модел, на който се основава процедурата на изследването, се определя от спецификата на обекта на изследване – свързващите думи и изрази и мястото им в системата на езика като текстоизграждащи елементи. За разлика от другите средства за постигане на кохезия, които са предимно граматически или предимно лексикални, свързващите думи и изрази са на границата между двете: предимно граматически, но с лексикален компонент, изразяващ логическите връзки в текста, и следователно, на пръв поглед, най-експлицитните маркери на неговата организация.

Теорията, която дава възможност да се разглеждат едновременно лексикалният, граматическият и текстуалният аспект на всяка една езикова единица, е развита от Хоуи (2004, 2005). Според него всяка езикова единица има определен „прайминг“, т.е. граматични и лексикални предпочитания за поява в определено езиково и текстуално обкръжение. Тази съвкупност от „прайминги“ той нарича „текстуална колигация“. Текстуалната колигация отразява положителните или отрицателни комбинаторни предпочитания на всяка лексикална единица, някои от които са:

- да се явява като част от конкретен тип семантична връзка;
- да се явява в началото, средата или края на определен отрязък от текст, например параграф;
- да участва в кохезионни вериги;
- да се явява в определен тип или жанр текстове, насочени към определена група потребители.

Според Хоуи нашето знание за всяка лексикална единица е съвкупността от всички или от част от тези прайминги. В процеса на обучение е възможно да се усвоят само някои от тези прайминги и това да доведе до неестествената употреба на съответната езикова единица.

Подобна гледна точка предоставя многоаспектна перспектива за проучване и описание на езиковите средства, като обръща внимание на различни техни черти на поведението им в текста. Затова комбинацията от теоретичните постановки на Селинкър (1972), Богранд и Дреслер (1981) и Хоуи (2004, 2005) осигуряват подходяща основа за теоретичния модел, използван в дисертацията, за да се постигне обективно и цялостно изследване на свързващите думи и изрази в БАМЕ.

При описанието на свързващите думи и изрази е използвана класификацията на Халидей и Хасан (1976), разделяща тези лексикални елементи на четири вида: присъединителни, противопоставителни, причинно-следствени и темпорални, както и една допълнителна категория на вметнати оценъчни фрази, наречени според тази класификация „континуативи“.

Във **Втора глава** (35-56 стр.) е представена **методологията** на изследването, но се засягат и някои основни теоретични въпроси на корпусната лингвистика. Тя е сравнително нов подход към лингвистичните данни, развиващ се от началото на 90-те години на 20-век, и може да се опише по-скоро като методология, отколкото като отделен аспект на езика, изискващ обяснение и описание. Според Макенери и Уилсън (2001) основно предимство на корпусната лингвистика е, че се фокусира върху обективното изучаване на естествено възникващи

езикови данни от реалния живот, а не върху интроспективни умозакljučения, извлечени от субективното интернализирано когнитивно възприятие на езика и интуитивните предположения на отделния лингвист. Това става възможно благодарение на набирането на огромни бази данни от цялостни текстове, наречени корпуси, и разработването на инструменти за съхранение, анотация и обработка на тези данни.

По дефиниция корпус в съвременната корпусна лингвистика представлява единствено и само колекция с краен обем, състояща се от автоматично обработваеми текстове в електронна форма, подбрани с цел да бъдат максимално репрезентативни по отношение на даден език или негова разновидност и достъпни за широк кръг лингвисти. Само при работа с такъв тип бази данни в лингвистичните изследвания се смята, че е възможно получаването на обективни, валидни и надеждни резултати, които са сравними и могат да бъдат допълнително тествани от други лингвисти в същата област (Лийч 1995: 3, Макенери и Уилсън, 2001: 32).

Основен въпрос, който се дискутира от възникването на корпусната лингвистика в началото на 90-те години до днес, е връзката между данни и теория и ролята на теоретичния модел в процеса на изследването. Най-общо казано, към корпуса може да се подходи по два начина: първо, като се изходи от емпиричния материал, без предварително зададен теоретичен модел, да се разгледат, опишат и класифицират наблюдаваните единици и явления в езика, което в крайна сметка да доведе до създаване на теоретичен модел, т.е. да се използва “corpus-driven” анализ (анализ, воден от корпусните данни); или второ, като се вземат предвид теоретичните модели за езика, въз основа на тях да се потърсят структури и явления, които да докажат или отхвърлят теоретични или приложни модели, т.е. да се използва “corpus-based” анализ (анализ, базиран на корпусните данни).

В настоящата разработка е избрана комбинация от двата подхода. Към предмета на изследване най-напред се прилага corpus-driven процедура, за да се установят аспектите на междинния език, в който има проблем. Това се налага, тъй като в България до този момент настоящото изследване е все още първо и единствено от този вид, анализиращо със средствата на корпусната лингвистика големи бази данни от езика на учещи английски, и поради това липсват сравними анализи, които да послужат като отправна точка.

Corpus-based подходът се оказва по-уместен за

анализа на конкретните езикови единици, защото corpus-driven методът предполага обстойно разглеждане на цялата система на даден език и установяване на закономерности, които може да не съвпадат с нито един от съществуващите теоретични модели. Но целта на нито едно изследване на междинен език не може да бъде подробното описание и утвърждаване на проявените в него девиации и включването им в някакъв нов теоретичен модел за самия него. Обикновено към езика на учещите се подхожда от гледна точка на явленията и закономерностите в езика-цел, които се търсят, описват и преподават, за да се постигне максимално доближаване на междинния език към езика-цел в процеса на обучение.

За да се разбере мястото и значението на описания и използван за настоящото изследване корпус на междинен език, тук е направен също така и **кратък преглед на различните видове електронни корпуси** и на тяхното приложение в съвременната лингвистична наука. Представени са някои от най-значимите корпуси на английски език, събрани до този момент.

Втората част на тази глава е посветена на корпусната лингвистика и изследванията на междинните езици. Дава се дефиниция на „корпус на учещи“ (learner corpus) и се описват неговите характеристики. Според Грейнджър (1998) „корпусът на учещите представлява колекция от автентични текстове на учещи чужд език, събрана и структурирана според ясно дефинирани критерии за постигането на определена цел на чуждоезиковото обучение. Тези данни са кодирани по стандартен и хомогенен начин и са документирани по отношение на техния произход.“ Това определение предполага наличието на база данни от цялостни текстове, които съдържат както неправилна, така и правилна употреба на езика, и би било погрешно да се говори за корпус, ако дадено изследване на междинен език се базира само на сбор от отделни думи, изречения или грешни структури, извадени от общия контекст. Следователно едно от големите предимства на работата с електронни корпуси е контекстуализирането на изследваните структури и разглеждането им на фона на цялостната употреба на езика-цел от страна на учещите. Подобно третиране на изследвания материал не води нито до отхвърляне, нито до подценяване на предходни изследвания на междинните езици, а по-скоро продължава традициите, установени през 70-те и 80-години на миналия век, като ги премества в друга плоскост и насочва изследванията към разглеждане на цялостния междинен

език, а анализът на грешките се явява маргинален проблем.

Трета глава (57-79 стр.) е посветена на описание на използваните в дисертацията **електронни корпуси и технически средства**.

Както отбелязва Синклер (1991), началото на всяко корпусно изследване е създаването на самия корпус, а валидността на резултатите се определя от валидността на материала, с който се работи. Поради това тази глава е по-обширна и подробно описва процеса и принципите на създаването на Международния корпус на английския език на учещи (ICLE) и най-вече на българския му компонент (BICLE), както и мястото му сред други подобни корпуси, набрани по света. Като първи и единствен засега структуриран корпус на БАМЕ от този вид, до който имат достъп широк кръг изследователи, той дава възможност да се използва за проверка и сравняване на резултатите и това прави изследванията, базирани на него, съпоставими с други подобни изследвания на междинни езици.

На описанието на този корпус е отделено по-голямо и по-специално място и поради факта, че неговото набиране се инициира още през втората половина на 90-те години на 20-ти век, когато самата корпусна лингвистика прави своите първи стъпки, но вече е набрала достатъчно подкрепа сред изследователите-лингвисти по света. В тези условия голяма част от усилията на създателите на този корпус бяха насочени към изработването на критериите за неговото събиране, структуриране и дизайн, за изработване на софтуер и тагери за аотиране и обработка на данните и, не на последно място, за убеждаване на лингвистичната общност, че този вид корпуси имат своето място и значение в лингвистичните изследвания. Затова и голяма част от теоретичните и приложнопрактически публикации до излизането на първата му версия на електронен носител през 2002 година са насочени главно към описание на тези процеси и трудностите по пътя на разработването му.

В първата версия на ICLE (2002) участват екипи от 11 държави и са набрани 3640 есета, наброяващи 2 500 350 думи, като всеки национален подкорпус съдържа около 200 000 думи. Втората му версия (2009) е обогатена с над 1 200 000 думи, събрани от още 5 национални подкорпуса. От самото начало на процеса са въведени стриктни критерии за събиране на данните, така че крайният резултат да има „статут на корпус“ (Аткинс и Клиър, 1992). Основно изискване към всяка национална разновидност е събирането на данни от хомогенно население. Въведени са контролирани променливи като: среда на обучение, възраст, роден

език, ниво на владеене на езика и характер на задачата. Релевантната биографична информация за всеки учещ, като години на изучаване на чужд език, престой в англоговоряща страна, знание на други чужди езици, използване на речници и граматика при изпълнение на задачата, и съгласие за участие в корпуса, е събрана в профил на учещия. Цялата тази информация е налична в крайния софтуерен продукт и дава възможност да се съставят подкорпуси на базата на избрани от потребителя критерии.

Всички учещи са студенти на възраст между 22 и 27 години, изучаващи английския по-скоро като чужд, отколкото като втори език, при което най-малко 60% от тях са достигнали ниво на владеене C1-C2. Текстовете са основно аргументативни академични есета от 500 до 1000 думи. Темите са избрани от предварително предложен списък, еднакъв за всички национални екипи. Част от корпуса (25%) е аотиран за части на речта (POS tagged) и за грешки (error tagged).

Българският компонент на ICLE е събран от писмената продукция на студенти от втори курс английска филология на СУ „Климент Охридски“ в периода 1997 – 2002 г. Данните напълно отговарят на зададените за този корпус критерии. Повечето есета първоначално са написани от студентите на ръка и след това са превърнати в електронен формат от координатора на проекта, при стриктно съблюдаване на „политиката на чистия текст“ (Синклер, 1991: 21), т.е. възпроизвеждане на текста точно без поправяне на грешките и без въвеждане на нови.

Предимствата на съвместен проект като ICLE са очевидни. Вместо да се анализират изолирани измислени изречения и да се правят изследвания в ограничен мащаб, което е трудоемко и нерентабилно, този корпус дава възможност да се осъществяват мащабни проучвания на различни варианти на междинни езици и по този начин да се вадят обективни изводи. Освен това изследователите имат достъп до огромни бази данни, които могат да се обработват автоматично или полуавтоматично, което ускорява работата с английския език на учещите и улеснява развитието на методологията и средствата за чуждоезиково обучение.

В тази глава се отбелязват и някои общи характеристики на корпусите на английския език на учещите, които правят работата с тях зависима от наличието на други корпуси. Всеки корпус на естествен език може да бъде използван самостоятелно за изследване на характерни черти на съответния естествен език, от който са събрани данните. За разлика от него, корпусът

от езика на учещи представя данни за междинен език. А, както вече беше споменато, междинният език, независимо от етапа на своето развитие, е само едно приближение и подобие на естествения език-цел, към който се стремят учещите, и може да носи елементи от родния им език. Поради тази причина всеки корпус на междинен език е необходимо да се сравнява с най-малко два други корпуса, съставени от данни за езика цел и за родния език на учещите.

Освен това, трябва да се има предвид, че естествените носители на езика, както и учещите, преминават през серия от етапи, преди да достигнат високо ниво на владение на своя роден език, особено по отношение на писмената си продукция. Както е посочено от Лоренц (1999), съществува „скала на владение на езика“ (proficiency cline) не само при неестествените носители на езика. Това налага всеки корпус на учещи английски език да се сравнява и с корпус от писмена продукция на учещи, които са естествени носители на английски език.

Всичко казано дотук обосновава необходимостта при анализа на данните в тази дисертация, наред с корпуса от **писмения език на българските учещи** (BICLE - Bulgarian International Corpus of Learner English, съдържащ 200 704 думи), да се използват и следните допълнителни корпуси:

- **LOCNESS** (Louvain Corpus of Native English Essays) Лувенски корпус, съдържащ 207095 думи от аргументативни есета, написани от британски и американски ученици и студенти, представящ **писмения език на естествените носители на английски, които все още са учещи**.
- **BNC-Sub** – подкорпус на Британския национален корпус (BNC – British National Corpus), съдържащ 207 000 думи от случайно подбрани нехудожествени текстове от области, подобни на тези, разработени в темите на студентските есета. Този корпус представя **нормата на езика-цел** за изследваните тематични области и типове текст.
- **BGC** (Bulgarian Language Corpus) – корпус, наброяващ 200 507 думи от публикувани текстове на български език от области и типове, подобни на тези, разработени в темите на студентските есета. Този корпус представя **нормата на родния език на български учещи** по отношение на изследваните тематични области и типове текст.

Данните от ICLE/BICLE, LOCNESS и BGC са обработени с WordSmith Tools (Скот, 1997) – основен софтуер при лингвистични корпусни изследвания. А данните от BNC-Sub са обработени със собственото приложение на BNC – SARA.

Проверката на значимостта и валидността на резултатите от количествения анализ е направена с онлайн хи-квадрат тест, широко прилаган към лингвистични статистически данни, което дава основание резултатите да бъдат съотнесени не само към текстовете, включени в конкретния корпус, а и към всички текстове от този вид.

В Четвърта глава (79-240 стр.) е представено **емпиричното изследване**. То се състои от две основни части – количествен и качествен анализ на данните.

Както във всички изследвания, базирани на корпус, като водещ фактор при избора на конкретните единици на езика се използва честотата на употребата им, което според Макенери и Уилсън (2001) е единственият надежден начин да се открие кои явления са наистина отражение на поведението на даден език и разновидност и кои от тях се появяват случайно. Честотата на употреба на дадена единица е един от аспектите на езика, за които имаме много малко интуитивно знание, но тя е аспект, който играе централна роля при изясняване на въпроси, свързани не само с това какво е възможно, но и какво е вероятно да се появи в езика. Затова първоначално към данните се прилага corpus-driven подходът, чрез което се установят свързващите думи и изрази, показващи най-големи различия между езика-цел и езика на учещите, и чието поведение в текста следва да бъде подложено на по-задълбочен качествен анализ. И така, **количественият анализ търси отговор на първия от изследователските въпроси**, поставени в началото на дисертацията, а именно, степента на несъответствие между писмената реч на българските учещи и тази на естествените носители на английски език – учещи и норма.

От българския корпус на учещите – BICLE, с помощта на WordSmith Tools, са извлечени конкорданси и честотни списъци на всички свързващи думи и изрази, включени в класификацията на Халидей и Хасан (1976: 242-243). Установява се, че 50 от тях се явяват поне един път в данните. За тези 50 свързващи думи и изрази са зададени същите търсения и в другите два корпуса на английски език – BNC-Sub и LOCNESS – и са направени сравнения на честотата на тяхната поява, отразени в Таблица 3 на дисертацията (стр. 82 – 83). Числата в нея, както и във всички останали таблици до края на дисертацията,

са нормализирани, за да са напълно сравними. Сравненията показват сериозна свръхупотреба от страна на българските учещи както във всяка една от категориите като цяло, така и при индивидуалните единици. Спрямо LOCNESS 35 показват свръхупотреба, 13 показват недостатъчна употреба и само 2 са еднакви, а спрямо BNC-Sub съотношението е съответно 37:12:1.

Сравненията на данните по категории свързващи думи и изрази са представени графично в таблици и фигури и са проверени за статистическа значимост с хи-квадрат теста, за да се установи доколко предполагаемата зависимост съществува обективно и неслучайно (стр. 84-98).

Този **количествен анализ** води до следните изводи:

1. **БАМЕ** на това ниво демонстрира ясно изразено различие по отношение на употребата на свързващи думи и изрази в сравнение с естествените носители на езика – учещи и норма, което подкрепя това предположение, направено в началото на дисертацията.
2. Отделни единици може и да се явяват по-рядко, отколкото в данните от естествените носители на езика, но в голямото си мнозинство свързващите думи и изрази показват стабилна тенденция за свръхупотреба, което потвърждава второто интуитивно убеждение, изказано в началото на дисертацията, че българските учещи като цяло разчитат повече на тези езикови единици при организацията на своите аргументативни текстове, отколкото на други средства за постигане на кохезия.
3. Отделните категории свързващи думи и изрази показват различни тенденции на употреба от страна на българските учещи английски език. Присъединителните, противопоставителните и континуативите демонстрират значителна свръхупотреба в сравнение с нормата на езика-цел и с подобен тип писмена продукция от учещи, които са естествени носители на езика. Причинно-следствените и темпоралните единици показват значителна свръхупотреба, както при българските учещи, така и при учещите, които са естествени носители на езика, като и двете групи учещи показват отклонение от нормата на езика-цел. Следователно може да се каже, че употребата на причинно-следствените и темпоралните свързващи думи и изрази е типична за академичното аргументативно писане.

4. Степента на отклонение при високочестотните свързващи думи и изрази (*and, or, but, because, so* и др.) в писмената реч на българските учещи е значителна, но относително ниска в сравнение писмената реч на естествените носители на езика – учещи и норма.
5. Нискочестотните свързващи думи и изрази (*in/by contrast, likewise, in spite of, hence, for one thing, surely* и др.) също демонстрират свръхупотреба от страна на българските учещи, но броят им е недостатъчен, за да даде статистически значими резултати и основания за сравнение.
6. Най-драматичните и значими отклонения се явяват при средночестотните (*for example, thus, in fact, actually, although, therefore, first/finally* и др.) свързващи думи и изрази, особено от присъединителен, противопоставителен и континуативен тип, което показва, че при напредналите учещи английски българи едни от най-големите несъответствия в праймингите са свързани с тези езикови единици.

На базата на тези изводи са отделени четири средночестотни свързващи думи и изрази, които пропорционално показват най-големи девиации от езика-цел и на които се прави качествен анализ във втората част на тази глава (стр. 99-239). Те са: *in fact, actually, of course* и *thus*. Изборът им се оправдава от факта, че такова значително несъответствие твърде вероятно допринася за неестественото звучене на писмената продукция на българите, учещи английски език, дори и когато са достигнали високи нива на владеење на чуждия език. Изследването само на тези единици не изчерпва въпроса изцяло, но задава модел и очертава насоките за по-нататъшни изследвания на проблематични области в БАМЕ.

Процедурата на качествения анализ се основава на теорията на Хоуи, обсъдена в Първа глава на дисертацията и на някои от критериите, предложени от него за определяне на текстуалната колигация на езикови единици (виж стр. 8 по-горе). Всяка една от изследваните езикови единици е разгледана по отношение на семантични връзки, позиция в определен отрязък от текст, участие в кохезионни вериги и предпочитания за употреба в определен тип текст, като това е направено най-напред в корпусите от данни за езика-цел (BNC-Sub и LOCNESS) поради липса на подобни изследвания, които да послужат за сравнение. По същите критерии е разгледана проявата им в изследвания вариант на БАМЕ (BICLE) и е направен сравнителен профил на текстуалната колигация на анализираната езикова единица в трите корпуса на английски език.

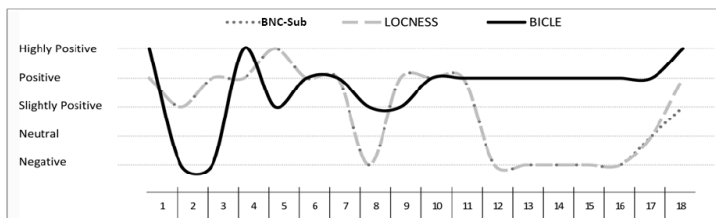
За да се установи дали родният език на учещите влияе върху междинния им английски, такива наблюдения са направени и по отношение на българските съответствия на изследваните езикови единици в корпуса на български език (BGC).

Сравнения са направени и с останалите подкорпуси на ICLE, за да се анализира доколко изведените характеристики на съответните свързващи думи и изрази се повтарят при учещи с друг роден език и по този начин да се прецени влиянието на други фактори при усвояването на техните прайминги, като например стратегии на преподаване, обучение и комуникация.

* * *

От сравняването на **текстколигационната прозодия на *in fact*** в данните за естествения английски език и за БАМЕ (стр. 120-121) се установява следното:

	BNC-Sub	LOCNESS	BICLE
⇒ Preference for occurring as part of specific semantic relations			
Semantic relation			
1 • Avowal	Positive	Positive	Highly positive
2 • Contrast	Slightly positive	Slightly positive	Negative
Environment of occurrence			
3 • Negative	Positive	Positive	Negative
4 • Positive	Positive	Positive	Highly positive
⇒ Preference for occurring at the beginning or end of a recognisable chunk			
5 • Paragraph-initial position	Highly positive	Highly positive	Slightly positive
6 • Sentence-initial position	Positive	Positive	Positive
7 • Mid-sentence position	Positive	Positive	Positive
8 • End-position	Negative	Negative	Slightly positive
Patterns:			
9 • <i>In fact</i> – S – V – O	Positive	Positive	Slightly positive
10 • S – (aux) – <i>in fact</i> – V – O	Positive	Positive	Positive
11 • S – be – <i>in fact</i> – C	Positive	Positive	Positive
12 • S – <i>in fact</i> – aux (or modal) – V	Negative	Negative	Negative
13 • Relative pronoun – <i>in fact</i> – clause	Negative	Negative	Positive
14 • Conjunction (but) – <i>in fact</i> – clause	Negative	Negative	Positive
15 • In wh-interrogative clause	Negative	Negative	Positive
16 • S – V – O – <i>in fact</i>	Negative	Negative	Positive
⇒ Preference for participating in cohesive chains			
17	Neutral	Neutral	Positive
⇒ Constraints on operation of preferences			
18 • Argumentative writing, Applied Science, Social Science and World Affairs.	Slightly positive	Positive	Highly positive



1. При естествените носители на езика – учещи и норма – *in fact* има силно позитивно предпочитание да обозначава логическата връзка утвърждаване (на предходно казаното) и

малко по-слабо, но все пак положително предпочитание за изразяване на контраст (с предходно казаното). В изследвания вариант на БАМЕ тази езикова единица има силно позитивно предпочитание към първата семантична връзка и ясно изразено негативно предпочитание към втората. Освен това, при естествените носители на езика-цел *in fact* се асоциира с положителни и отрицателни контексти в съотношение съответно 1:2, докато при българските учещи данните се свързват почти изключително само с положителна текстуална среда в съотношение 5:1.

2. При материала от естествените носители на езика се забелязва, че *in fact* има положителен прайминг за употреба в начална и средна позиция на параграфа или изречението. Данните за българските учещи показват леко положителна тенденция за появата на тази езикова единица и в крайна позиция. При естествените носители на езика се забелязва сложна текстуална колигация, показваща зависимост между изразената семантична връзка и позицията в изречението (параграфа), при което утвърждаването се употребява предимно в средна позиция и по-често в обкръжение на негативни думи, а контрастът се асоциира предимно с начална позиция и положителни контексти. В БАМЕ *in fact* не показва такива взаимозависимости между семантична връзка, позиция в изречението и лексикално обкръжение, тъй като се явява предимно в смисъл на утвърждаване във всички позиции и предимно в компанията на положителна или неутрална лексика.
3. При естествените носители на езика се явяват по-малко синтактични модели, отколкото при българските учещи. Проверката на различните модели спрямо по-голяма база данни от Британския национален корпус (BNC) показва, че те не са непременно грешни, но са нискочестотни и пропорционално представляват незначителна част от всички примери за употреба на *in fact*, което предполага че те не са типични за изследвания тип текст. Оттук следва изводът, че на това ниво българите владеят разнообразие от модели, но не винаги ги употребяват уместно.
4. Чрез задаване на по-широк контекст на конкордансите се установява, че трите типа английски данни показват равномерно разпределение на *in fact* в отделни текстове от корпусите и много рядка употреба в един и същ текст, в съседни изречения или дори параграфи. Но веригите от свързващи думи и изрази, наблюдавани в примерите

от езика-цел, са много по-кратки от тези при българските учеци, което показва силна зависимост на аргументацията им от този тип езикови средства.

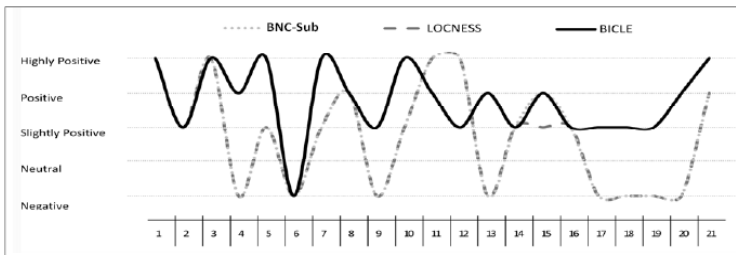
5. И в трите типа данни *in fact* показва положително предпочитание за употреба в аргументативни текстове в областите: приложни науки социални науки, глобални проблеми, като при българските учеци това предпочитание е изразено в по-голяма степен.

* * *

Преди да се пристъпи към сравнение с българските еквиваленти на *in fact* за установяване на евентуален трансфер от роден език, в трите корпуса на английски език е анализирана и другата средночестотна лексикална единица *actually*, тъй като съществуващите двуезични речници дават еднакви еквиваленти на *in fact* и *actually* („всъщност, действително, в действителност, наистина, фактически“). Игнорирането на този факт би могло да създаде „шум“ в статистическите данни, защото е възможно свръхупотребата на всяка от двете лексикални единици да се дължи на активиране на другата в съзнанието на учеция при възникване на комуникативна нужда от изразяване на съответната логическа връзка.

Изводите от анализа на текстколигационната прозодия на *actually* са следните:

		BNC-Sub	LOCNESS	BICLE
	⇒ Preference for occurring as part of specific semantic relations			
1	Semantic relation			
2	• Avowal	Highly Positive	Highly Positive	Highly positive
3	• Contrast	Slightly positive	Slightly positive	Slightly positive
	Environment of occurrence			
4	• Negative	Highly Positive	Highly Positive	Highly positive
5	• Positive	Negative	Negative	Positive
5	• Neutral	Slightly positive	Slightly positive	Highly positive
	⇒ Preference for occurring at the beginning or end of a recognisable chunk			
6	• Paragraph-initial position	Negative	Negative	Negative
7	• Sentence-initial position	Slightly positive	Slightly positive	Highly positive
8	• Mid-sentence position	Positive	Positive	Positive
9	• End-position	Negative	Negative	Slightly positive
	Patterns:			
	Initial position			
10	<i>Actually</i> – S – V – O	Slightly positive	Slightly positive	Highly positive
	Mid-sentence position			
11	(55) <i>S(noun phrase) – actually – V – O</i>	Highly positive	Highly positive	Positive
12	(56) <i>S(noun phrase) – (aux) – actually – V – O</i>	Highly positive	Highly positive	Slightly positive
13	(59) <i>S(clause = what actually V) – V – O</i>	Negative	Negative	Positive
14	(57) <i>S(noun phrase) – (modal aux) – actually – (aux) – V – O</i>	Slightly positive	Slightly positive	Slightly positive
15	(58) <i>S(noun phrase) – be – actually – C</i>	Slightly positive	Positive	Positive
16	(60) / (61) <i>actually + ing/ inf</i>	Slightly positive	Slightly positive	Slightly positive
17	(62) cleft sentences “it + be + actually + S	Negative	Negative	Slightly positive
18	(63) appositive “... actually, ...”	Negative	Negative	Slightly positive
	End position			
19	(64) <i>S- V- O - actually</i>	Negative	Negative	Slightly positive
	⇒ Preference for participating in cohesive chains			
20		Neutral	Neutral	Positive
	⇒ Constraints on operation of preferences			
21	• Argumentative writing, Applied Science, Social Science and World Affairs.	Positive	Positive	Highly positive



1. Логическите връзки утвърждаване и контраст при естествените носители на езика (учещи и норма) са в съотношение 4:1. И при българските учещи това съотношение не е много по-различно, но все пак те още по-категорично предпочитат да използват *actually* за утвърждаване на току-що казаното. Но докато естествените носители използват *actually* почти винаги само в негативни или неутрални контексти, при българските учещи трите типа обкръжение – позитивно, неутрално и отрицателно – показват равномерни дялове. Тази асоциация на *actually* с позитивни идеи води до дисбаланс в семантичните прайминги на думата и обяснява отчасти нейната свръхупотреба.
2. По отношение на позицията в изречение/параграф/текст *actually*, както и *in fact*, не показва драстични разминавания, като се среща предимно в средна позиция и в трите вида данни. И отново както при *in fact*, българските учещи показват по-голямо разнообразие на синтактични модели, които макар и да не са изцяло грешни, са много нискофrequentни или изобщо не се срещат в изследваните типове текст.
3. Що се отнася до кохезионните вериги, трите английски корпуса леко се различават, като отново българските учещи показват много по-дълги кохезионни вериги, които обхващат отделните текстове изцяло.
4. Предпочитанията за участие в определени типове текст са същите както и при *in fact*.

Въпреки почти еднаквите речникови еквиваленти на *in fact* и *actually*, този анализ, освен всичко друго, доказва, че те не са взаимозаменяеми в който и да е контекст на естествения английски

и че всяка от тях има своето специфично място и функция, тъй като показват различни колокационни и колигационни, но не и текстуални модели.

От сравнението с другите междинни езици в ICLE става ясно, че съществува обща тенденция за свръхупотреба при учещите английски език, които не са естествени негови носители. Оттук следва заключението, че поради стратегии на преподаване, които са присъщи на повечето учебни материали по академично писане, учещите английски, независимо от техния роден език, обикновено асоциират тези езикови единици с аргументативни типове тест и изказване на категорично лично мнение.

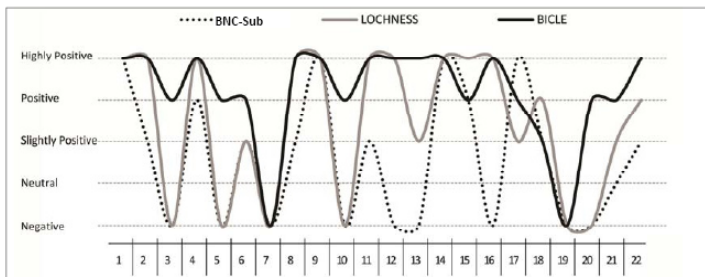
От друга страна, с интернационализирането на английския език е възможно различните варианти на междинен език да са развили черти, присъщи на повечето неестествени носители на езика под влияние на тяхната комуникация както с естествени, така и с други неестествени носители на езика. На този етап данните представят обективни доказателства за такива процеси, но това не изключва възможността за влияние на съответните родни езици. Доколко всеки друг роден език, различен от български, влияе върху съответната група учещи не е предмет на настоящата дисертация. Но една от задачите, поставени в началото ѝ е да се изследва влиянието на българския език върху БАМЕ и това е следващата стъпка тук.

Кумулативните резултати за българските еквиваленти потвърждават интуитивните оценки за влияние на роден език, но отиват една стъпка по-нататък, като разкриват конкретния характер на девиациите. Наблюдават се най-малко три характеристики, които се коренят в трансфер от роден език: а) рядка употреба за въвеждане на контраст, б) използване в крайна позиция или като вметната част на изречението и в) предпочитание за употреба в положителен контекст.

* * *

Анализът на текстколигационната прозодия на *of course* води до следните изводи:

		BNC-Sub	LOCNESS	BICLE
	⇒ Preference for occurring as part of specific semantic relations			
	Semantic relation			
1	• Acceptance	Highly Positive	Highly Positive	Highly positive
2	• Rejection	Slightly positive	Highly Positive	Highly positive
3	• Digression	Negative	Negative	Positive
	Environment of occurrence			
4	• Negative	Positive	Highly Positive	Highly positive
5	• Positive	Negative	Negative	Positive
6	• Neutral	Slightly positive	Slightly positive	Positive
	⇒ Preference for occurring at the beginning or end of a recognisable chunk			
7	• Paragraph-initial position	Negative	Negative	Negative
8	• Sentence-initial position	Slightly positive	Highly positive	Highly positive
9	• Mid-sentence position	Highly positive	Highly positive	Highly positive
10	• End-position	Negative	Negative	Positive
	Patterns:			
	• Initial position			
11	– <i>Of course</i> – S – V – O	Slightly positive	Highly positive	Highly positive
12	– <i>Of course</i> (adverbial /parenthetical phrase /clause)– S – V – O	Negative	Highly positive	Highly positive
13	– <i>Of course</i> – S (there) – V (be) – O	Negative	Slightly positive	Highly positive
	• Mid-sentence position			
14	connector/ reaction signal/ initiator – <i>of course</i> – clause	Highly positive	Highly positive	Highly positive
15	S – <i>of course</i> – V – ...	Positive	Highly positive	Positive
16	S – V – <i>of course</i> – ...	Slightly negative	Highly positive	Highly positive
17	S – V – NP(C/O/A) – <i>of course</i> – NP(C/O/A)	Highly positive	Slightly positive	Positive
18	within a non-finite clause	Slightly positive	Positive	Slightly positive
19	within the verb phrase	Slightly negative	Negative	Slightly Negative
	• End position			
20	S – V – O. – <i>of course</i> .	Negative	Negative	Positive
	⇒ Preference for participating in cohesive chains			
21		Slightly positive	Slightly positive	Positive
	⇒ Constraints on operation of preferences			
22	• Argumentative writing, Applied Science, Social Science and World Affairs.	Slightly positive	Positive	Highly positive



1. По отношение на семантичните връзки, изразявани чрез *of course*, двата корпуса на учещите – BICLE и LOCNESS – са много подобни и значенията на „приемане“ и „отхвърляне“ имат почти еднакви дялове. Причината

за това може да са приликите между типовете текст, събрани от учещите, и изискванията на поставената им писмена задача да се съгласят или да отхвърлят дадени твърдения. Във всички случаи, в съответните пасажи има лексикални вериги, които допълват свързващата роля на *of course*, но 1/5 от данните на българските учещи не показва такава характеристика и *of course* поема цялата кохезионна сила. Непосредствената лексикална среда потвърждава, че пишещият се отклонява от главната посока на мисълта и свързващата дума придобива „вътрешна“ функция за кохезията в текста. Това е първата сериозна разлика между данните на българските учещи и двата типа данни от естествените носители на езика. Другото наблюдавано отклонение е свързано с факта, че в корпусите от естествени носители на английския *of course* избягва контексти с положително маркирана лексика, докато при българите положителната среда се явява най-често.

2. Резултатите за позициите на *of course* показват, че като цяло данните при българските учещи се доближават до общите честоти на начална и средна употреба, наблюдавани при естествените носители на езика, но освен това демонстрират и малко по-силно изразена тенденция към употреба в крайна позиция. Също така може да се твърди, че с почти равните дялове на начална и средна употреба в BICLE и LOCNESS българските учещи отново се доближават повече до данните от английските учещи, отколкото до данните за нормата на езика-цел – BNC. Това отново потвърждава факта, че някои такива употреби са обусловени от типа задача и от навиците за учене, свързани с нея.

Освен тази употреба в крайна позиция се забелязват и други различия при предпочитаните синтактични конструкции, като например употреба на *of course* след глаголната фраза, както и честа употреба в *there*-конструкции, които не се откриват при естествените носители на езика.

3. Разлики между продукцията на българските учещи и двата типа данни от естествените носители на английски се наблюдават и във връзка с участието на *of course* в кохезионни вериги, които при българските учещи са два пъти по-дълги.
4. Както и при другите изследвани езикови единици,

сравнението с останалите подкорпуси на ICLE доказва, че употребата на *of course* е типична за академичното писане на напреднали учещи английски език. Това се наблюдава и при учещи с родни езици, принадлежащи дори към различни езикови семейства, което отново потвърждава заключенията, направени по-горе, че влиянието на родния език намалява за сметка на други фактори, като например, стратегиите на учене и комуникация, както и че стратегиите на преподаване, налагащи определени текстуални модели при академичното писане допринасят за свъхупотребата на *of course*. Това обаче не изключва обективна проверка чрез корпусни данни на възможното влияние на родния език върху БАМЕ.

На базата на анализа на честотността и семантиката на *of course* и на неговия българския еквивалент „разбира се“ може да се заключи, че родният език не е основен фактор при формирането на праймингите на българските учещи за употребата на *of course* в академичното писане. Родният език може да има известен принос при формиране на мнението, че английският еквивалент на „разбира се“ е полезен, когато се дава оценка на текстовата ситуация и се акцентира върху личното мнение за постигане на ефект на по-голяма компетентност и увереност при изказаните твърдения, но в изследваните данни няма достатъчно доказателства за такава тенденция.

Вариантите на синтактични модели и сравненията на синтактичната и текстова среда на „разбира се“ и *of course* също не предлагат достатъчно доказателства за трансфер от роден език, което отново води до извода, че праймингите на *of course* при българските учещи са в голяма степен резултат от стратегиите, използвани в процеса на обучение и комуникация.

По отношение на участие в кохезионни вериги, българското „разбира се“ не се различава особено много от английското *of course* в корпуса на нормата (BNC-Sub). Този факт не е изненадващо неочакван, след като корпусът на български език е съставен от текстове, написани и редактирани от добре образовани хора, така както и BNC-Sub, и поради това и двата корпуса не показват необоснована и редундантна употреба на ресурси за организация на текста. Тези наблюдения отново отделят писмената продукция на българските учещи на английски от нормата за писане на български език и потвърждават предположението, че на това ниво на чуждоезиково усвояване и във връзка с употребата на някои

свързващи думи и изрази влиянието на родния език отслабва за сметка на други фактори в процеса на обучение.

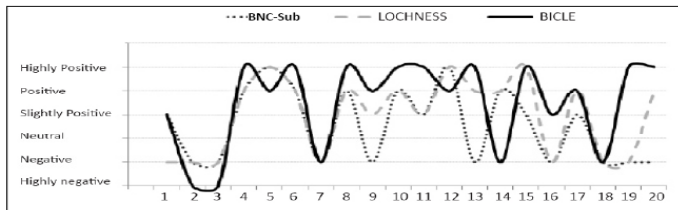
Всички тези твърдения за поведението на „разбира се“, както и за останалите лексикални единици, изследвани тук, се отнасят до конкретните области и жанрове на приложните науки, социалните науки и глобалните проблеми, застъпени в корпусите.

* * *

Thus се оказва изненадващо предпочитана свързваща дума за българските учещи не само при сравнение с LOCNESS и BNC-Sub, но и в рамките на собствената си категория, където е четвъртата по честота сред останалите присъединителни свързващи думи в BICLE.

Текстколигационната прозодия на *thus* показва следното:

		BNC-Sub	LOCNESS	BICLE
	⇒ Preference for occurring as part of specific semantic relations			
	Semantic relation			
1	• addition	Slightly positive	Negative	Slightly positive
2	• manner	Negative	Negative	Highly negative
3	• time	Negative	Negative	Highly negative
4	• cause	Positive	Positive	Highly positive
5	- result	Highly positive	Highly positive	Positive
6	- summary	Positive	Positive	Highly positive
	⇒ Preference for occurring at the beginning or end of a recognisable chunk			
7	• Paragraph-initial position	Negative	Negative	Negative
8	• Sentence-initial position	Positive	Positive	Highly positive
9	- First sentence in paragraph	Negative	Slightly positive	Positive
10	- Mid-sentence in paragraph	Positive	Positive	Highly positive
11	- Last sentence in paragraph	Slightly positive	Slightly positive	Highly positive
12	• Mid-sentence position	Highly positive	Highly positive	Positive
	Patterns:			
	• Mid-sentence position			
13	Clause – (and) thus – clause	Negative	Positive	Highly positive
14	S – (aux) thus – V – O – ...; or S – verb/be – thus – O/C ...	Positive	Positive	Negative
15	Clause – (and) thus – non-finite (-ing) ...	Slightly positive	Highly positive	Highly positive
16	Clause – (and) thus – non-finite (to inf)	Negative	Negative	Slightly positive
17	Clause – (and) thus – verb (elliptical clause)	Slightly positive	Positive	Positive
18	• End-position	Negative	Negative	Negative
	⇒ Preference for participating in cohesive chains			
19		Negative	Negative	Highly positive
	⇒ Constraints on operation of preferences			
20	• Argumentative writing, Applied Science, Social Science and World Affairs.	Negative	Positive	Highly positive



1. По отношение на семантичните връзки, изразявани от *thus* в различните контексти, данните за българските учещи не са много по-различни от тези за двата типа естествени носители на езика. Дори изглежда, че българските учещи не се отклоняват много от нормата, след като всички семантични характеристики на *thus* са представени по подобен начин и в трите корпуса. Това се отнася за смисъла *manner* (начин) и *time* (време) и отчасти за *addition* (добавяне). В този последен смисъл обаче средата на употреба при българските учещи показва редуваност на значения, изразени чрез употребата на други свързващи думи в непосредствена близост. Има също така сериозни различия в най-голямата група примери – *causals* (причинно-следствени), където българските учещи демонстрират точно обратните тенденции за предпочитания при подгрупите *resultative* (резултативност) и *summative* (обобщение) в сравнение и с двата английски корпуса – на учещите и на нормата. Това е индикация, че българските учещи нямат ясно разбиране за всички семантични характеристики на тази езикова единица.
2. Синтактично българските данни показват голямо приближение до данните на естествените носители на езика и това е друг положителен аспект на употребите, който показва, че на лексикално ниво българският учещ е наясно с основните прайминги на тази дума и рядко произвежда очевидно грешни синтактични конструкции. Но тук също се забелязват някои интересни характерни черти, свързани не толкова с позициите на *thus* в рамките на изречението, колкото с позицията на изречението, съдържащо *thus*, в рамките на параграфа и текста. При българските учещи тези изречения се явяват предимно като последно изречение на параграфа, което драстично се различава от предпочитанията на естествените носители на езика, при които *thus* предпочита по-скоро средно

положение в параграфа. Това е сигнал, че текстуалната дистрибуция на лексикалните единици има значение за естественото звучене на текстовете толкова, колкото и правилната им граматическа употреба в изречението, и това е още един аспект, на който трябва да се обръща внимание в процеса на преподаване и учене.

3. Другата наблюдавана черта – участие в кохезионни вериги, отново потвърждава по-рано направените заключения. Струпването на други думи-връзки около *thus* в сравнително кратки отрязъци от текст се проявява ясно при българските учещи, но не и при естествените носители на езика – учещи и норма. Като изключим отделни кохезионни връзки и още по-редки вериги, основани на повторения на самата лексикална единица или на употреба на други свързващи думи и изрази, изразяващи логически връзки, различни от тези на *thus*, при естествените носители на английски език *thus* почти винаги се явява самостоятелно на такива места в текста, които изискват да се покаже обрат на мисълта или отношението на автора към развитието на аргумента. За разлика от това, българското *thus* участва в по-дълги вериги от свързващи думи и изрази, често с подобни значения, което привидно прави текстовете „прекомерно логични“, но всъщност те се оказват по-скоро нелогични поради липсата на други средства за постигане на кохезия.
4. Дали наблюдаваните характеристики на *thus* са специфични само за българските учещи или се проявяват при учещи с друг роден език е проверено отново посредством сравнения с подкорпусите на ICLE. Тези сравнения доказват, че от 16-те групи учещи само четири се приближават до честотите, получени от нормата на езика-цел, а всички останали, включително и учещите естествени носители на английски в различна степен показват свръхупотреба. Очевидно тази характеристика е също специфична за изследваните области и жанра на аргументативните писмени текстове на учещите.

Данните от английските корпуси показват, че проблемите на свръхупотреба при българските учещи се явяват главно при причинно-следствия – резултативен и обобщителен смисъл. Но сравненията с българските еквиваленти на *thus* – „така, по този/такъв начин, така например, и така, следователно“ не потвърждават влияние на роден език, тъй като данните за българските му еквиваленти се доближават до данните за нормата на езика-цел, така както еквивалентите на *of course*. Както беше отбелязано,

това не са изненадващи факти, като се има предвид, че речниците и граматиките обикновено отразяват значенията и употребите на всяка езикова единица достатъчно точно. По-интересен тук е фактът, че данните и този път отхвърлят вероятността за силно негативно влияние на родния език при българите и предполагат, че свръхупотребата от страна на учещите е резултат от процеси, базирани на съществуващото знание на учещите за езика-цел.

* * *

В заключението на дисертацията (стр. 246-254) се дава резюме на теоретичната основа, методологията и процедурата на изследването и се обобщават резултатите от поставените в началото изследователски въпроси, цели и задачи.

Наблюденията от количествения анализ дават отговор на първия изследователски въпрос. Данните за българските учещи показват сериозна свръхупотреба на свързващи думи и изрази в сравнение с естествените носители на езика-цел – учещи и норма, което показва, че дори и при високите нива на усвояване на чуждия език българите се отклоняват от текстуалните модели, типични за естествените носители на езика, и това е проблем, на който трябва да се обърне внимание при чуждоезиковото обучение.

Наблюдаваната свръхупотреба е разпределена неравномерно между отделните свързващи думи и изрази, както и между отделните техни категории.

Отклонението на българските учещи от английските учещи и нормата на езика-цел е различна за високочестотните, средночестотните и нискочестотните свързващи думи и изрази, независимо от типа логическа връзка, която те изразяват. Високочестотните и нискочестотните единици показват значителна свръхупотреба, но пропорционално най-съществените несъответствия се явяват при средночестотните.

Наблюденията върху четири от тях дават отговор на втория изследователски въпрос, свързан с характеристиките на тяхното поведение в текста и с възможните причини за свръхупотреба. Корпусните данни доказват, че изследваните езикови единици са типични за аргументативни писмени текстове на учещи, независимо от техния роден език, и това може да се разглежда като общо явление при междинните езици. Това е разбираемо, като се има предвид огромното количество учебници, учебни помагала, курсове и онлайн ресурси, които предлагат списъци от свързващи думи и убеждават учещите, че тяхната употреба е задължителна, ако желаят текстовете им да бъдат добре

структурирани и логични. Много малко от тези ресурси обясняват, че логичните текстове се базират на логично мислене, а мисленето се изразява чрез цялата гама от лексикални и граматически езикови средства вместо да се „залепва“ върху несвързани изречения като допълнителна рамка. Освен това те, в повечето случаи, не посочват, че текстуалната прозодия на всяка свързваща дума представлява сложно взаимодействие от колокационни, колигационни и семантични асоциации и учещите често придобиват знание за тези черти изолирано поради непълната или понякога противоречива информация в тези източници. Така учещите формират фрагментирани и непълни прайминги за думите и техните комбинации, а дори, когато само една черта на прайминга липсва или е неточна, това се отразява върху цялостното качество на техните текстове. Независимо дали това е неестествена синтактична структура, неуместна лексикална среда или струпане на ненужно голям брой други логически маркери, текстовете се изместват от типичните текстуални модели на естествените носители на езика.

Много важно наблюдение, извлечено от качествения анализ, е, че писмената продукция на българските напреднали учещи има много допирни точки с продукцията на учещите, които са естествени носители на езика, и тези допирни точки маркират редица позитивни характеристики и постижения на българските учещи, които ги нареждат на четвърто място сред 16-те други групи в международния корпус ICLE. Също така корпусите дават достатъчно доказателства да се твърди, че колкото по-напреднали са учещите, толкова повече отслабва влиянието на родния език и се засилва ефектът от другите психолингвистични фактори върху тяхното знание на чуждия език. Въпреки това, и тук все още има следи от негативен трансфер на прайминги, останал от писмените навици, създавани в часовете по български език, като например стремежа да се използват дълги вериги от свързващи думи и изрази при липса на друга лексикална кохезия в тяхното обкръжение.

Игнорирането на положително и отрицателно натоварените контексти е друг аспект на употребата, на който никога не се обръща внимание в часовете по български език и това може би е още един важен момент от негативния трансфер на прайминги.

Отделни синтактични модели също могат да се причислят към трансфера от роден език, но, трябва да се признае, че българският учещ на това ниво знае къде да постави свързващата дума в изречението, макар да не знае къде тя стои в параграфа или в цялостния текст.

По отношение на третия изследователски въпрос и последната от поставените в дисертацията задачи, в края на заключението се дават предложения във връзка с преподаването на английски език – още в началните етапи по-експлицитно и целенасочено да се обръща внимание на всички тези колокационни, колигационни и текстуални характеристики на езиковите единици, за да се избегнат съществуващите проблеми и по-нататък в горните нива на обучение. Лексикалната кохезия при предаването на идеи, съобразяването с положителното или отрицателното текстуално обкръжение при избора на думите, правилното им позициониране в текста са част от истинските „липсващи връзки“ в академичното писане на българите на английски и тези характеристики на езика трябва нарочно да се показват на учещите чрез добре подбрани примери и упражнения.

Като разкрива тези страни на БАМЕ, настоящото изследване може да допринесе за обективизиране на критериите за оценка на писмената продукция на учещите и за вграждане на тези критерии в софтуерни продукти за проверка на текстове, а също и за изготвяне на учебни материали, съобразени с нуждите на българите, учещи английски език, което да доведе до приближаването на езика на учещите към езика-цел.

Изследването може да се използва за усъвършенстване също и на процеса на аотиране на корпуси от текстове на учещи и подобряване на съществуващите тагери, като се вземат предвид не само граматическите и лексикалните аспекти на езика, но се отчитат и някои от текстуалните им характеристики. Така усъвършенстваните корпуси биха дали възможност за още по-подробни и обективни изследвания на междинните езици.

Освен това, заключението очертава някои допълнителни приложения на изследването и изводите от него, които не са свързани пряко със сферата на образованието. Разбирането на начина, по който писмено общуват различни групи потребители, разграничени по възраст, роден език, страна на произход, ниво на владеене на английски, придобива все по-голямо значение в днешния свят на виртуални общности и комуникации. Описаните характеристики на езика в резултат на подобни изследвания могат да намерят приложение при разработване на мобилни приложения, TTS софтуер, и други инструменти, пригодени за различни потребители, а дори и в съдебната лингвистика за установяване на авторство, виртуални престъпления и др.

БИБЛИОГРАФИЯ НА ЦИТИРАНАТА В АВТОРОФЕРАТА ЛИТЕРАТУРА

- Алтенберг и Тапер (1998)** : Altenberg, B. and Tapper, M. (1998) The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (ed.), *Learner English on Computer*. (pp.80-93). London and New York: Longman
- Аткинс и Клиър (1992)**: Atkins, S. and J. Clear (1992) Corpus design criteria. *Literary and Linguistic Computing* 7(1), 1-16
- де Богранд и Дреслер (1981)**: Beaugrande, R. de. and W. Dressler (1981) *Introduction to text linguistics*. Longman: London and New York
- Грейнджър (1994)**: Granger, S. (1994) The Learner Corpus: a revolution in applied linguistics. In *English Today* 39, Vol.10, No. 3 (July 1994), Cambridge University Press, pp. 25-32.
- Грейнджър (1998)**: Granger S. (1998a) The computerized learner corpus: a versatile new source of data for SLA research. In Granger S. (ed.) *Learner English on Computer*. London & New York: Addison Wesley Longman, 3-18
- Грейнджър и Тисон (1996)**: Granger, S. and Petch-Tyson S., (1996) Connector usage in the English essay writing of native and non-native EFL speakers of English, *World Englishes*, 15, p. 19-29.
- Дамаскели (1999)**: Damascelli, A. (1999) Argumentative Essays Written by English and Italian University Students. Aspects of Contrastive Rhetoric: the Use of Logical Connectors. Unpublished MA dissertation, University of Torino, Italy (personal communication)
- Кашубски (2000)**: Kaszubski, P. (2000) *Selected aspects of lexicon, phraseology and style in the writing of Polish advanced learners of English: a contrastive, corpus-based approach*. School of English, Adam Mickiewicz, Poznań
- Лийч (1995)**: Leech, G., Mayers, G. and Thomas, J (eds.) (1995) *Spoken English on Computer*, London: Longman
- Макенери и Уилсън (2001)**: McEnery, T and A. Wilson (2001) *Corpus Linguistics*, Edinburgh: Edinburgh University Press.
- Селинкър (1972)**: Selinker, L. (1972) Interlanguage. *IRAL; International Review of Applied Linguistics in Language Teaching*, 10:3, pp. 209-31.
- Селинкър (1992)**: Selinker, L. (1992) *Rediscovering Interlanguage*. Longman Group UK Ltd.
- Синклер (1991)**: Sinclair, J (1991) *Corpus, concordance, collocation*. Oxford University Press.

- Скот (1997):** Scott, M. (1997) *Wordsmith Tools. Version 2.* Oxford: Oxford University Press
- Халидей и Хасан (1976):** Halliday, M.A.K and R. Hasan (1976) *Cohesion in English*, (English Language Series 9.) London: Longman
- Хоуи (2004):** Hoey, M (2004) Lexical priming and the properties of text in *Corpora and discourse*, A. Partington, J. Morley, and L. Haarmann, eds., Peter Lang, 385-412.
- Хоуи (2005):** Hoey, M. (2005) *Lexical priming A new theory of words and language.* Routledge.

СПРАВКА ЗА ПРИНОСНИТЕ МОМЕНТИ В ДИСЕРТАЦИЯТА

1. Създаден е първият, структуриран международно признат електронен корпус на БАМЕ, който е включен в ICLE (Международен корпус на езика на учещите английски) и е сравним с електронни корпуси на учещи английски език от 16 други страни от целия свят.
2. Участие в изработването на дизайна и критериите за събирането на данните за езика на учещите от страните, включени в ICLE, както и осъществяване на самото събиране, обработка и електронизиране на данните.
3. За първи в България е направено международно съпоставимо изследване на големи бази данни от езика на учещи английски като чужд със средствата на корпусната лингвистика и по този начин е поставено началото на нов етап в изучаването на БАМЕ в съответствие със съвременните тенденции в контрастивния анализ на междинните езици (Contrastive Interlanguage Analysis).
4. Принос към методологията на корпусните изследвания на междинните езици е изводът, че работата с корпуси от учещи изисква наличие и сравнение не само с корпус на езика-цел, но и с корпус на родния език на учещите.
5. Извършено е първото корпусно изследване и описание на текстуални характеристики на писмената реч на българи, учещи английски език.
6. Направено е първото цялостно количествено изследване на свързващите думи и изрази в БАМЕ на ниво напреднали чрез използване на corpus-driven метод (метод за анализ, воден от корпусните данни).
7. Модел за изследване на естествен език е адаптиран и успешно приложен за изследване на междинен език.
8. Зададена е процедура за качествен анализ на лексикални единици, която може да се използва за по-нататъшни изследвания на проблематични области в БАМЕ.

ПУБЛИКАЦИИ ПО ТЕМАТА НА ДИСЕРТАЦИЯТА

I. Статии по темата на дисертацията

1. **Благоева, Р. (1998):** Blagoeva, R. (1998) Some Considerations Concerning Interlanguage and Second Language Curriculum for a High Level of Proficiency in English. In *Foreign Language Teaching*, No 2. 1998, pp. 61-65, Sofia, ISSN 0205-1834
2. **Благоева, Р. (2001a):** Blagoeva, R. (2001a) Comparing Cohesive Devices: Conjunctions and Other Cohesive Relations and Their Place in the Bulgarian-English Interlanguage. In *Foreign Language Teaching*, No 1. 2001, pp. 11-19, Sofia, ISSN 0205-1834
3. **Благоева, Р. (2001b):** Blagoeva, R. (2001b) Comparing Cohesive Devices: a Corpus-based Analysis of Conjunctions in Written and Spoken Learner Discourse. In *Proceedings of the Corpus Linguistics 2001 Conference*, pp. 59-63, Lancaster, ISBN 1 86220 107 2.
4. **Благоева, Р. (2001c):** Blagoeva, R. (2001c) Personal Reference and Other Cohesive Relations in Learner Writing: a Corpus-based Analysis. In *Proceedings of the 22nd International Computer Archive of Modern and Medieval English Conference*, 14-16, Louvain-la-Neuve
5. **Благоева, Р. (2002):** Blagoeva, R. (2002) The Status of English in Bulgaria. In Granger, S., Dagneaux E. and Meunier F. (2002) *The International Corpus of Learner English. Handbook and CD-ROM*, pp. 81-86, Louvain-la-Neuve: UCL Presses Universitaires de Louvain
6. **Благоева, Р. (2004):** Blagoeva, R. (2004) Demonstrative reference as a cohesive device in advanced learner writing: a corpus-based study. In Aijmer, K., B. Altenberg, *Advances in Corpus Linguistics*, pp. 297-307, Rodopi B.V., Amsterdam – New York, NY 2004, ISBN: 90-420-1741-4
7. **Благоева, Р. (2005):** Blagoeva, R. (2005) Electronic Corpora and the Study of Discourse Features in Bulgarian-English Interlanguage. In M. Georgieva (ed) *Spaces, Gaps, Borders*, pp. 153-165, Sofia: St. Kliment Ohridski University Press
8. **Благоева, Р. (2009):** Blagoeva, R. (2009) The Status of English in Bulgaria. In Granger, S., Dagneaux E., F. Meunier and M. Paquot. *The International Corpus of Learner English Version 2. Handbook and CD-ROM*. pp 114-120, Louvain-la-Neuve: UCL Presses Universitaires de Louvain , ISBN: 978-2-87463-143-6

II. Доклади по темата на дисертацията, изнесени на конференции и симпозиуми

- 1. Благоева, Р. (1998):** Blagoeva, R. (1998) Comparing Grammatical Categories: Evidentials, paper given at the First International Symposium on Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching, The Chinese University of Hong Kong, 14-16 December.
- 2. Благоева, Р. (2000):** Blagoeva, R. (2000) Comparing Cohesive Devices: Conjunctions and Other Cohesive Relations and Their Place in the Bulgarian-English Interlanguage, paper given at the Third Contemporary European Studies Conference, Voneshta Voda, Bulgaria, May.
- 3. Благоева, Р. (2001a):** Blagoeva, R. (2001a) Comparing Cohesive Devices: a Corpus-based Analysis of Conjunctions in Written and Spoken Learner Discourse, poster presentation at Corpus Linguistics 2001 Conference, Lancaster University, UK, 29 March – 2 April.
- 4. Благоева, Р. (2001b):** Blagoeva, R. (2001b) Personal Reference and Other Cohesive Relations in Learner Writing: a Corpus-based Analysis, paper given at 22nd International Computer Archive of Modern and Medieval English Conference, Louvain-la-Neuve, Belgium, 16-20 May.
- 5. Благоева, Р. (2002):** Blagoeva, R. (2002) Demonstrative reference as a cohesive device in advanced learner writing: a corpus-based study, poster presentation at 23rd International Computer Archive of Modern and Medieval English Conference 'The Theory and Use of English Language Corpora', Gothenburg, Sweden, May 22-26.
- 6. Благоева, Р. (2003):** Blagoeva, R. (2003) Electronic Corpora and the Study of Discourse Features in Bulgarian-English Interlanguage, paper given at 8th International Conference of the Bulgarian Society for British Studies, Sofia, 24-26 October.
- 7. Благоева, Р. (2006):** Blagoeva, R. (2006) A Corpus-based Study of the Lexical Primings of Conjunctions in Bulgarian English Interlanguage, paper given at ICAME-27 Variation, Contacts and Change, University of Helsinki, Finland, 24-28 May.

III. Цитирания от други автори на някои от публикациите по темата на дисертацията

1. **Гилкин, Г. (2003):** Gilquin, G (2003) Automatic Retrieval of Syntactic Structures: The Quest for the Holy Grail. In *International Journal of Corpus Linguistics* Volume 7, Number 2, 2003, pp. 183-214(32) John Benjamins Publishing Company
2. **Диез-Бедмар, М. Б. (2005):** Díez Bedmar M. B (2005) Struggling with English at University level: error patterns and problematic areas of first-year students' interlanguage. In *Proceedings from The Corpus Linguistics Conference Series*. Available at <http://www.academia.edu/7799133>
3. **Котос, Е. (2014):** Cotos, E (2014) Enhancing writing pedagogy with learner corpus data. In *ReCALL / Volume 26 / Special Issue 02 / May 2014* Cambridge University Press
4. **Кушафар (2012):** Kooshafar, M (2012) The Effect of Dictogloss Technique on Learners' Writing Improvement in Terms of Writing Coherent Texts. In *Journal of Language Teaching and Research*, Vol. 3, No. 4, pp. 716-721, July 2012 ACAD-EMY PUBLISHER
5. **Ларсон, Т. (2012):** Larsson, T (2012) On spelling behavior: A corpus-based study of advanced EFL learners' preferred variety of English. In: *Nordic Journal of English Studies*, ISSN 1654-6970, E-ISSN 1654-6970, Vol. 11, no 3, 127-154
6. **Ли Юн Ёо (2007):** Lee, Eun Joo. (2007). A corpus based analysis of the Korean EFL learners' use of evaluative adjectives. In *Foreign Languages Education*, 14 (4), 73-92
7. **О, С. (2009):** Oh, Sun-Young (2009) Functions of Because in Native English and Korean Learner Corpora. In *Korean Journal of Applied Linguistics*, Vol.25 No.3, 2009.12, pp. 443-470.
8. **Пайва К. С. Ф. М. и Е. Лима (2013):** Paiva, K. S. F. M. and E. Lima (2013). Systemic functional linguistics as a desirable approach to L2 writing evaluation: a case study. In: *Seminário de Línguas Estrangeiras : a formação e a prática de professores de línguas estrangeiras, 2013, Goiânia. Anais do VIII Seminário de Línguas Estrangeiras: a formação e a prática de*

professores de linguas estrangeiras. Goiânia: FUNAPE: FACULDADE DE LETRAS UFG, 2013. pp. 207-21

9. **Шифтнер, Б. и Т. Ранкин (2012):** Shiftner, B and T. Rankin (2012) The use of Demonstrative reference in English texts by Austrian School-Age Learners, in *Developmental and Cross-linguistic Perspectives*. In *Learner Corpus Research*, edited by Yukio Tono, Yuji, John Benjamins Publishing Company,
10. **Юксел, Х. Г. и С. Каваноз (2010):** Yüksel, H. G., S. Kavanoz (2010) Bir Çapraz Çözümleme: Türkçe ve İngilizce Yazılı Metinlerde Bağlaç Kullanımı, In Çiğdem Sağın-Şimşek ve Çiler Hatipoğlu (eds) *24. Ulusal dilbilim kurtultayi Bildiri kitabı*, 17-18 Mayıs 2010, Ankara: ORTA DOĞU TEKNİK ÜNİVERSİTESİ, Eğitim Fakültesi, Yabancı Diller Eğitimi Bölümü. pp. 354-362.
11. **Ян и Сун (2012):** Wenxing Yang and Ying Sun (2012) The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. In *Linguistics and Education*, Volume 23, Issue 1, March 2012, pp.31–48
12. **<http://www.dirk-siepmann.de/Links/LearnerLanguage/learnerlanguage.htm>** (In: Course outline and Bibliography of Prof. Dirk Siepmann, the Institute of English and American Studies at Osnabrueck University, Germany)