

Илина ДОЙКОВА

Медицински университет „Проф. д-р Параскев Стоянов“

Варна, България

ilina.doykova@mu-varna.bg

ЛИНГВИСТИЧНА КОРЕКТНОСТ И ВАЛИДИРАНЕ НА МЕДИЦИНСКА ЛЕКСИКА

Iliana DOYKOVA

Prof. Dr. Paraskev. Stoyanov Medical University of Varna

Varna, Bulgaria

ilina.doykova@mu-varna.bg

LINGUISTIC CORRECTNESS AND VALIDATION OF MEDICAL LEXIS

Translation of specialized medical textbooks is investigated as a source of bilingual terminology and professional lexis in the language pair English–Bulgarian. The loan vocabulary used by the Bulgarian medical community in written scientific discourse is verified in a parallel corpus for the purpose of building specialized e-resources for training purposes. The collocational potential of the key terms is highlighted with linguistic samples of contextual usage.

The aim of the corpus-based study is to compile English–Bulgarian specialized medical glossaries where key words and specific language patterns are identified in an authentic source text (English) and their translation equivalents in Bulgarian are checked for consistency, appropriateness, and linguistic correctness.

Keywords: bilingual medical texts, key vocabulary, terminology, translation equivalents.

Увод

Обект на настоящото проучване е българската професионална лексика и терминология в областта на медицината и здравеопазването. Чрез корпусен лингвистичен анализ на паралелни текстове се проследява функционирането на терминологичните единици и високочестотната лексика в езиковата двойка английски – български. В съответствие с наблюдаваните процеси на интернационализация и национализация (Колковска 2006) се проверява съответстви-

ето и употребата на изходните единици от английски език спрямо преводните им еквиваленти в езика приемник (български) и описанието им в специализираната литература.

Интересът към тясноспециализираната лексика и новите терминосъчетания в писмената медицинска комуникация е провокиран от факта, че въпреки установените критерии по отношение на адаптацията и интеграцията на заемките (Попова 2003; Крумова-Цветкова 2013) в българския медицински език липсва системност от гледна точка на прагматиката и езиковата употреба. Много новообразувани лексеми и словосъчетания функционират като условни терминологични единици, като за условни приемаме тези, които не са описани в стандартните речници, но имат активна употреба в специализирания речников фонд на езика през последните 10 години.

Целта на изследването е описание и систематизиране на неологизмите в специализирания медицински език, навлизащи активно през английски език, както и образуването на хибридни думи, т.е. състоящи се от интернационален и домашен компонент (Смирнов 1989, цит. по Колковска 2006) при създаване на двуезични терминологични ресурси за специализиран превод (области медицина и здравеопазване). Традиционните медицински речници най-често са едноезични, а съществуващите двуезични електронни ресурси за професионални цели (български-английски) имат ограничен обем. През последните години нараства броят на електронните справочници и речници (напр. <https://www.puls.bg/reference/dictionary/>, <https://medpedia.framar.bg/>), където могат да се открият някои описания на английски заемки с честотна употреба в професионалния език на медицината и здравеопазването.

Методология

Изследването е базирано на английско-български паралелен корпус за извличане и проверка на двуезична специализирана лексика и терминология при създаването на електронни ресурси за целите на езиково обучение и превод, които предоставят възможност за осъвременяване на съдържанието (Bowker, Pearson 2002: 21) и проверка на специфична терминология в автентичен контекст. Корпусът на изследването (MEDENCA) съдържа авторски базирани данни (оригинални учебници по медицина и техните официални преводи), от които с помощта на лицензирани онлайн инструменти са ексцерпирани данни от двата езика. За изграждането на паралелен корпус, извличане и систематизация на професионална лексика са използвани програмите за текстов анализ Sketch Engine (Kilgariff 2014) и WordSmith (Scott 2012), които позволяват обработката на корпусни данни по отношение на българския език.

Основната задача при опита за валидиране на двуезична професионална лексика е изграждането на представителен паралелен корпус, който да позволи откриването на възможно най-голям брой специализирана лексика, представена от различни преводи/преводачи и установяването на доминантен екви-

валент (напр. абдоминален vs. коремен). Представителността на изследването е гарантирана от обема на корпуса и от факта, че текстове на оригиналните академични учебници са успоредени от официалните преводи на дванадесет преводачи на медицинска литература (фиг. 1.).

MEDENCA		Surgery Book - Translation	
user/IDoykova/medenca • created: 8/29/2019, 9:35:44 AM		user/IDoykova/surgery_book	
Research articles (BMC)		Bulgarian translation	
GENERAL INFO		GENERAL INFO	
Language	English	Language	English
Tagset	DESCRIPTION	Tagset	DESCRIPTION
Word sketch grammar	SHOW	Word sketch grammar	SHOW
COUNTS		COUNTS	
Tokens	2,323,273	Tokens	2,366,058
words	1,767,365	words	1,925,968
Sentences	77,470	Sentences	65,686
Documents	51	Documents	48

Фиг. 1. Специализиран чужд език за професионални цели – паралелен корпус

Корпусът позволява търсене по лема или части на речта, установяване на значението в контекст, специфична употреба и комбинативни възможности (Powers 1998; Selinker 2009). Езиковите ресурси, включени в корпуса на изследването са подбрани в съответствие със следните принципи:

- съдържат актуална, тясноспециализирана лексика като индикатор на дискурса (писмена професионална комуникация/written expert-to-expert discourse);
- съдържат терминология и метатерминология в съответната научна област.

Използваните конкорданс програми идентифицират ключовите думи и понятия в съответната област, комбинативните възможности и преводните еквиваленти (фиг. 2.). Статистическите данни подпомагат системността в описанието на професионалната лексика, както в понятийно-семантичен, така и в лексикално-граматичен план.

<s> Дневна скорост на освобождаване за всеки микроелемент и/или витамин, ако се използва болус .</s>	<s> Daily release rate for each trace element and/or vitamin if a bolus is used.</s>
<s> Максимален период на непрекъснатото освобождаване на микроелемент или витамин, ако се използва болус .</s>	<s> Maximum period of continuously release of the trace element or vitamin if a bolus is used.</s>
<s> Разрешава се прилагане под формата на болус .</s>	<s> Application in the form of bolus is allowed.</s>

Фиг. 2. Паралелна употреба на лема болус/bolus (двyezичен конкорданс)

Подборът на лексикални единици за включване в двyezични тематични глосари за целите на обучението (и специализирания превод) се основава на автоматично генериран списък с термини и високочестотна специализирана

Astrocytoma is the most common **primary** central nervous system (CNS) neoplasm. The term glioma often is used to refer to astrocytomas **specifically, excluding** other glial tumors. Astrocytomas are **graded** from I to IV. Prognosis varies **significantly** between **grades** I/II, III, and IV, but not between I and II. Median **survival** is 8 years after diagnosis with a low- **grade** tumor, 2 to 3 years with an anaplastic astrocytoma, and roughly 1 year with a GBM. rGBMs account for almost two-thirds of all astrocytomas, anaplastic astrocytomas account for two-thirds of the rest, and low- **grade** astrocytomas the remainder.

Фиг. 7. Вариант А: Общи термини, включени в списъка с академични думи (Wang et al. 2008)

Astrocytoma is the most common **primary central nervous system (CNS) neoplasm**. The **term glioma** often is used to **refer to astrocytomas** specifically, excluding other **glial tumors**. **Astrocytomas** are graded from I to IV. **Prognosis** varies significantly between grades I/II, III, and IV, but not between I and II. **Median survival** is 8 years after **diagnosis** with a **low-grade tumor**, 2 to 3 years with an **anaplastic astrocytoma**, and roughly 1 year with a **GBM**. **GBMs** account for almost two-thirds of all **astrocytomas**, **anaplastic astrocytomas** account for two-thirds of the rest, and **low-grade** astrocytomas the remainder.

Фиг. 8. Вариант Б: Тясноспециализирани термини, включени в глосар Хирургично лечение

Общите термини и тези в съответната професионална област (хирургия) са представени двуезично в корпуса на изследването (Таблица 1.). Примерите са еквивалентни на ниво изречение (брой думи съответно в лява и дясна колона) и онагледяват основната преводаческа стратегия – пряко (материално) заемане от английски език. Валидирането на професионалната лексика на български език (или проверка на употребата на термина в специализиран двуезичен контекст) е постигнато чрез ръчно аотиран корпус на основни номинационни единици в езика на медицинските специалисти (MEDENCA).

Таблица 1. Модул Неврохирургия:
паралелен двуезичен корпус MEDENCA

Източник: Английски език	Превод на български език
<i>Posterior fossa lesions such as tumors, hemorrhage, or stroke can cause mass effect that can rapidly kill the patient in two ways.</i>	<i>Лезии в задната ямка като хеморагия или инсулт, които предизвикват мас-ефект, може бързо да убият пациента по два начина.</i>

<i>Occlusion of the fourth ventricle can lead to acute obstructive hydrocephalus, raised intracranial pressure (ICP), herniation, and eventually death.</i>	<i>Оклузията на четвърти вентрикул предизвиква остра обструктивна хидроцефалия, водеща до повишено ИКН, херниране и смърт.</i>
<i>Symptoms of brain stem compression include hypertension, agitation, and progressive obtundation, followed rapidly by brain death.</i>	<i>Симптомите на мозъчностволова компресия включват хипертония, възбуда, прогресивно помрачаване на съзнанието и хипертензия, последвана бързо от мозъчна смърт.</i>
<i>A patient exhibiting any of these symptoms needs an emergent neurosurgical evaluation for possible ventriculostomy or suboccipital craniectomy (removal of the bone covering the cerebellum).</i>	<i>Пациент с някой от тези симптоми се нуждае от спешна неврохирургична оценка за възможна вентрикулостомия или субоципитална краниектомия (отстраняване на костта, покриваща церебелума).</i>
<i>This situation is especially critical, as expeditious decompression can lead to significant functional recovery.</i>	<i>Тази ситуация е особено важна, тъй като експедитивната декомпресия може да доведе до значително функционално възстановяване на пациентите.</i>
<i>Astrocytoma is the most common primary central nervous system (CNS) neoplasm.</i>	<i>Астроцитомът е най-честият първичен тумор на ЦНС.</i>
<i>The term glioma often is used to refer to astrocytomas specifically, excluding other glial tumors.</i>	<i>Терминът глиома (glioma) често се използва за обозначаване на астроцитомы, изключвайки други глиални тумори.</i>
<i>Grades I and II are referred to as low-grade astrocytoma, grade III as anaplastic astrocytoma, and grade IV as glioblastoma multiforme (GBM).</i>	<i>Степени I и II са известни като нискостепенен астроцитом, степен III като анапластичен астроцитом, и степен IV като мултиформен глиобластом (GBM).</i>
<i>Median survival is 8 years after diagnosis with a low-grade tumor, 2 to 3 years with an anaplastic astrocytoma, and roughly 1 year with a GBM.</i>	<i>Средната преживяемост е 8 години след поставяне на диагнозата нискостепенен тумор, 2 до 3 години с анапластичен астроцитом и приблизително една година с мултиформен глиобластом.</i>
<i>Figure 42-2 demonstrates the typical appearance of a GBM.</i>	<i>Фиг. 42-20 демонстрира типичен вид на глиобластом (GBM).</i>

Дискусия

Основно предизвикателство в медицинския език е липсата на кратко и икономично съответствие (преводен еквивалент) на преките заемки, които са значителна част от професионалната комуникация на здравните специалисти. Създаването на нови лексикални единици чрез пряко заемане, което не е продиктувано от необходимост, а от фактори като небрежност, за удобство или престиж и активното добавяне на лексикални неологизми в медицинския научен език е практически обусловен и наложил се лесен начин за справяне с непълнотите в специализираната литература (Благоева 2008). Употребата на лексикални неологизми обаче често нарушава два от стилите маркери на научния текст – точността и яснотата на научното изложение.

Процесът на установяване на английските заемки и интегрирането им в лексикалната система на езика-приемник (български) се подпомага от двуезичните специалисти в съответната научна област и преводачите, които в съответствие със своята родна и чуждоезикова компетентност формално възпроизвеждат и „въвеждат“ заетите единици като интуитивно се съобразяват с нормите на родния си език (Данчев 1981: 190–204).

При терминологичните единици от нормативна гледна точка обаче е важна прозрачността (в морфологично отношение), функционирането им в кохерентна система и участието им в словообразователни гнезда. Лингвистичната коректност на термините изисква изграждането им по установени модели и в съответствие с фонологичните, морфологични и морфосинтактични норми на езика-цел¹. Тясноспециализираните термини често нямат установен еквивалент в българския език, където образуването им е мотивирано от различни съображения и се допускат компромисни решения. Наблюдават се няколко вида авторски или преводачески стратегии: превод със съществуваща дума или израз (*survival* – *преживяемост*), описателен превод (*off-label medication* – *лекарства извън лекарското предписание/лекарствени продукти за показання, които не са официално одобрени/лекарства off-label*), пряко заемане чрез транскрибиране (*target* – *таргет*, *grading* – *грейдинг*) или транслитерация (*pattern* – *патерн*).

Изследването на двуезични медицински текстове показва съзнателната ориентацията на преводния текст към английския източник по отношение избора на лексикални средства, т.е. прагматичността в научната литература е изразена ясно в стремежа да се запази прототипът на формално равнище чрез директно заемане от английски език и графична адаптация. Пример за трансфер на дума от общия речников фонд на езика-източник, която активно се използва в българския медицински език за изразяване на съответствие е *къ/омплайънс* (*камерен, белодробен, сърдечен, миокарден, висок, нисък, повишен, редуциран* и др.). В медицинската секция на речника Merriam-Webster² съществителното

¹ ISO 704:2009: Terminology work – Principles and methods.

² <https://www.merriam-webster.com/>

име *compliance* има две значения, които частично съответстват на употребата в българския медицински език:

1. the ability or process of yielding to changes in pressure without disruption of the structure or function (еластична деформация)

2. the process of complying with a regimen of treatment (спазване, съответствие, съобразяване).

Българската лексикално-семантична база данни Bulnet³, създадена по аналогия с WordNet⁴, българският национален корпус⁵ и стандартните речници не предлагат определение или примери за употреба на тази езикова единица, използвана в официалните медицински публикации и преводи. Единствено в електронните научни публикации на български език откриваме следните три дефиниции:

1. Къмплайънсът е величина, противоположна на еластичността, показваща съотношението между променения обем на белите дробове (ΔV) и промененото налягане във въздухоносните пътища (ΔP)⁶.

2. ...осигуряване на съгласието и сътрудничеството на пациента за спазване на предписаното лечение и препоръки, т.нар. къмплайънс⁷.

3. ...степената, в която пациентът следва правилно указанията на лекаря, като има малко или е без участие в изграждането на терапевтичната стратегия (пасивна роля)⁸.

От направените наблюдения в корпуса се открива семантична адаптация (Бояджиев 1986: 208) и стесняване на значението на заемката *compliance* в българския медицински език (фиг. 3), за да означа степената на съпричастност на пациента към собственото му лечение като част от придържането към лекарското предписание. Отношението на пациента обаче включва освен съпричастност, още постоянство и съгласие с предписаната терапия. С други думи в българския език *комплайънс* изразява само колко точно пациентът следва предписания режим, но не включва останалите компоненти като постоянство и съгласуване, за да може да се приеме, че се придържа към терапията⁹.

³ BulNet <<http://dcl.bas.bg/bulnet/>>

⁴ WordNet <<https://wordnet.princeton.edu/>>

⁵ Български национален корпус < <http://search.dcl.bas.bg/>>

⁶ <https://medpedia.framar.bg/0>

⁷ Дякова, М. Къмплайънс при лица, преживели остър миокарден инфаркт – реалности и препоръки. // *Обща медицина*, 11, 2009, № 2, с. 10 – 14.

⁸ <https://inspiro-bg.com/komplayans-i-bridarzhane-kam-terapiyata-pri-bronhialna-astma/>

⁹ <https://www.medinfo.bg/>

1	frammar.bg	: контактът с пациента. </s><s> Основната цел е постигането на висок	комплайънс	от негова страна. </s><s> Някои стъпки в тази насока са: </s><s> - под
2	frammar.bg	път лекарствата... </s><s> Фрамар: Самолечение? </s><s> Не, това е	комплайънс	съучастие на пациента към терапията. </s><s> Фрамар: В лекциите д
3	lev.bg	о Холандия, като Франция и други, те прилагат тези елементи на крос	комплайънс	</s><s> Те ги прилагат, но от 2012 г. ще бъдат задължителни за всичк
4	frammar.bg	: леха, най-вече на медикаментозната терапия, е постигането на висок	комплайънс	от страна на пациента. </s><s> Колкото до Снежанка, тя е част именн
5	vestnikataka.bg	каганат. </s><s> Трябва ли да се чудим и да се държим за нивото на "	комплайънс	" между Хартата за правата на човека, Шенгенското споразумение и Н
6	frammar.bg	/хо око" и повишаване на качеството на живот на пациентите и техния	комплайънс	</s><s> Основни принципи на медикаментозната терапия </s><s> Гл
7	lev.bg	А това е важно, защото 2012 година ние влизаме в системата на крос	комплайънс	(cross compliance), на кръстосаните съответствия. </s><s> Но вижте, т
8	mysurgery.bg	ане, туризъм по равна до леко хълмиста местност, и – при съответния	комплайънс	– бягане в ограничени размери. </s><s> Трябва да се избягват всякак
9	frammar.bg	иката, вероятно удължаващи ефекта от терапията </s><s> - подобрен	комплайънс	на пациента чрез използването на комбиниран медикамент с фиксира
10	frammar.bg	лация с друг вид лечение са: непоносимост към медикаментите, нисък	комплайънс	високата цена на полимедикаментозната терапия. </s><s> За успеха
11	frammar.bg	: рецептори (за разлика от TCA, mirtazapine и SNRI) </s><s> По-добър	комплайънс	на пациентите </s><s> Добрият профил на поносимост, в комбинация
12	cio.bg	ния риск и отговаряне на регулаторните изисквания, както и на тези за	комплайънс	</s><s> Процесът за прогнозиране предлага очаквания за проходи и

Фиг. 3. Лематизирано търсене за същ. комплайънс в корпус Bulgarian Web 2012/bgTenTen12, Sketch engine

Други примери за преводни еквиваленти, предложени от медицинските специалисти (и преводачите) са синонимните прилагателни имена (*абдоминален в м. корем, пулмонарен в м. белодробен*), производните съществителни (*болус доза, болус преливане, малигненост, сензитивност*), причастни форми (*дискутабилен, нерезектабилен*), отглаголни съществителни имена за процедури и методи, оформени чрез добавяне на суфикси-аналози в българския език (напр. *-ация: пролиферация, дерегулация, екстравазация, ирадикация, инициация, инактивация, експлантация, агрегация; -иране: минимизиране, репарирание, таргетиране, антагонизиране, типизиране, шънтиране, клампиране, анастомизиране, експресиране, приоритизиране, катетеризиране, тромбозирание*), глаголи с английска основа (*потенцирам, инактивирам, метастазирам, персистирам, експресирам, експлатирам, резицирам, индикирам*). При наличието на синонимни изрази (напр. *surgical procedure/ хирургична намеса/ операция/манипулация/процедура; repair/ възстановяване/реконструкция*), установяването на доминантен преводен еквивалент е въз основа на високочестотната употреба на лексикалната единица или брой употреби в оригинален контекст. Тъй като процесът на интегриране на английските заемки в българския език е стремителен и некоординиран (Крумова-Цветкова 2013), в тези случаи не се регистрира сериозно активизиране на словообразователния капацитет на езика, а се включва в употреба чуждоезиковият прототип. Многобройни случаи на междуезикова интерференция и отрицателен пренос могат да се наблюдават в българския медицински език в следствие на прякото заемане от английски (фиг. 4.).

При необходимост се поставя гипсова имобилизация и ръката се **елевира**.

Типът на инцизията е **дискутабилна**.

Лонгитудинална инцизия върху максималната зона на **флукутация** (виж фиг. 44–22) е сигурна и избягва ятрогенни **компликации**, посочени при други типове инцизия – трансверзална и латерална.

Фиг. 4. Примери за пряко заемане и адаптация от английски език

Очевидно термините в съответната област се осигуряват (Nyland, Tse 2004; Мавродиева, Тишева 2016) чрез пряко заемане и морфологична и словообразователна адаптация към нормите на езика-приемник като най-продуктивен начин за обогатяване на специалния език (Колковска, Благоева 2016). Наблюдава се висока степен на графично-фонетична, морфологична и словообразователна адаптация, но и стремеж към запазване на оригиналната ортографична форма (напр. glioblastoma (GBM)/глиобластома, glioma/глиома). В редки случаи има налице формална вариантност (Попова 2012: 571), като според броя употреби в корпуса можем да твърдим, че се предпочита домашната форма (напр. gerair: *възстановяване (413)/реконструкция (345)*).

Двуезичният корпусен анализ на новосъздадените термини допринася за идентифициране на нови терминологични словосъчетания и за установяването на техните доминантни преводни еквиваленти. Въпреки ясно изразената тенденция към интернационализация (Попова 2012: 419) при заемането на специализирана лексика се налага валидиране поради неустановеността на много от новите заемки и възможните варианти на словообразователно адаптиране в българския език.

Специализираната лексика и комбинативните възможности на неологизмите (напр. комплайънс) представляват едно голямо предизвикателство в преподаването на чужд език за специални цели. Извличането на преводни еквиваленти според честотността на употреба в двуезичен контекст е опит да бъдат преодолените непълнотите в обучителната и лексикографската литература, както и да се осигури точност, прецизност и еднозначност при употребата на лексикални неологизми в професионалната комуникация на специалистите.

Заклучение

Систематизирането и изграждането на електронни двуезични тематични ресурси по специализиран чужд език, базирани на корпусни данни позволява описание на съвременното състояние на специализираната лексика и бързите промени, настъпващи в нея. Необходимостта от точна и еднозначна информация за значението и преводните еквиваленти на тясноспециализираната лексика налага използването на актуални специализирани речникови ресурси, съдържащи и лексикални неологизми (описание на лексикосемантичните отношения, дефиниции, употреба в контекст), тъй като българският език се

попълва от все по-синтетични (кратки и имплицитни), но и универсално разпознаваеми езикови единици.

ЛИТЕРАТУРА

Благоева 2008: Благоева, Д. *Неологизмите в съвременния български език*. Liternet: https://liternet.bg/publish13/d_blagoeva/neologizmite.htm (10 септември 2019).

Blagoeva 2008: *Neologizmite v savremenniya balgarski ezik*. LiterNet: https://liternet.bg/publish13/d_blagoeva/neologizmite.htm (10 September 2019).

Бояджиев 1986: Бояджиев, Т. *Българска лексикология*. София: Наука и изкуство. **Boyadzhiev 1986:** Boyadzhiev, T. *Balgarska leksikologia*. Sofia: Nauka i izkustvo.

Данчев 1981: Данчев, А. Англицизмите в българския език. // *Съпоставително езикознание*, кн. V, 190-204. **Danchev 1981:** Danchev, A. *Anglitsizmite v balgarskia ezik*. // *Sapostavitelno ezikoznanie*, vol. V, 190–204.

Колковска, 2006: Колковска, С. Тенденциите към интернационализация и национализация в новата българска икономическа терминология. // *Националният език в условията на чужди влияния и глобализация*. София: СУБ, 190–212. **Kolkovska 2006:** Kolkovska, S. *Tendentsiite kam internatsionalizatsiya I natsionalizatsiya v novata balgarska ikonomicheska terminologiaya*. Sofia: SUB, 2006, 190–212.

Колковска, Благоева 2016: Колковска, С., Благоева, Д. *Лексикалната норма и неографската практика*. София: АИ „Проф. Марин Дринов“. **Kolkovska, Blagoeva 2016:** Kolkovska, S., Blagoeva, D. *Leksikalna norma i neografskata praktika*. Sofia: AI “Prof. Marin Drinov”.

Мавродиева, Тишева 2016: Мавродиева, И., Тишева, Й. *От реферата до магистърската теза. Академично писане за студенти*. София: „БГ Учебник“. **Mavrodiieva, Tisheva 2016:** Mavrodiieva, I., Tisheva, Y. *Ot referata do magistarskata teza. Akademichno pisane za studenti*. Sofia: BG Uchebnik, 2016.

Biber 2012: Biber, D. Corpus-based and corpus-driven analyses of language variation and use. // *The Oxford handbook of linguistic analysis*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199544004.013.0008>

Bowker, Pearson 2002: Bowker, L., Pearson, J. *Working with specialized language: a practical guide to using corpora*. Routledge.

Hartley 2012: Hartley, J., New ways of making academic articles easier to read. // *International Journal of Clinical and Health Psychology*, 12(1), 143.

Hyland, Tse 2004: Hyland, K., Tse, P. Metadiscourse in academic writing: A reappraisal. *Applied linguistics*, 25(2), 156–177.

Kilgarriff et al. 2014: Kilgarriff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., Rychlý, P., Suchomel, V. The Sketch Engine: ten years on. // *Lexicography*, 1, 2014, 7–36.

Powers 1998: Powers, D. Applications and Explanations of Zipfs Law. // *New Methods in Language Processing and Computational Natural Language Learning*, *ACL*, 151–160. <https://www.aclweb.org/anthology/W98-1218.pdf> (30 август 2019)

Scott 2012: Scott, M. *WordSmith Tools version 6, Stroud: Lexical Analysis Software*. <https://www.puls.bg/reference/dictionary/>

Selinker 2009: Selinker, L., 2009. Interlanguage. // *IRAL - International Review of Applied Linguistics in Language Teaching*, 10(1-4), 209–232.

Simpson-Vlach, Ellis 2010: Simpson-Vlach, R., Ellis, N. An academic formulas list: New methods in phraseology research. // *Applied linguistics*, 31(4), 487–512.

Tribble 2002: Tribble, C. Corpora and Corpus Analysis: New windows on academic writing. // *Academic Discourse*, J. Flowerdew (ed.). London: Longman, 131–149.

Wang et al. 2008: Wang, J., Liang, S., Ge, G. Establishment of a medical academic word list. // *English for Specific Purposes*, 27(4), 442–458.