

УНИВЕРЗИТЕТ У БЕОГРАДУ
ФИЛОЛОШКИ ФАКУЛТЕТ

ИГОР Д. ИВАНОВИЋ

Контрастивна анализа стручне
терминологије ЕУ кроз перспективу
корпусне лингвистике

- ДОКТОРСКА ДИСЕРТАЦИЈА -

Београд, 2013.

UNIVERSITY IN BELGRADE
FACULTY OF PHYLOLOGY

IGOR D. IVANOVI

Contrastive Analysis of EU Terminology
Through the Perspective of Corpus
Linguistics

- DOCTORAL THESIS -

Belgrade, 2013

Ментор: др Весна Половина, редовни професор, Филолошки факултет,
Универзитет у Београду;

Чланови комисије:

др Весна Половина, редовни професор,

др Ивана Трбојевић, доцент,

др Игор Лакић, ванредни професор,

Датум одбране:

Захвалност

Овај рад настао је као плод трогодишњег истраживања и жеље да се кроз призму контрастивне и корпусно-рачунарске лингвистике обради група текстова који су настали у протеклих пар година. Аутор овог раде се нада да је одговорио захтјевима који су пред њега постављени и који морају бити испуњени како би настао квалитетан докторски рад.

Наравно, највећу захвалност дугујем својој менторки проф. др Весни Половини као и члановима комисије чија је професионална помоћ била од немјерљивог значаја приликом израде овог рада.

Аутор

Посвета

Овај рад посвећујем мојој мајци Бранки, моме оцу Драгољубу и мојој супрузи Весни Ивановић.

Наслов рада: Контрастивна анализа стучне терминологије ЕУ кроз перспективу корпусне лингвистике

Резиме: Овај рад је замишљен као рад у којем се прожимају контрастивна лингвистика са корпусном методологијом рада. Основна поставка јесте да се контрастирање разлика и сличности два језика (српског и енглеског) може урадити на основу корпусне методологије која, заједно са рачунарском лингвистиком, представља незаобилазно средство у модерној анализи разноврсних текстова у оквиру лингвистике. Корпус одабраних текстова се састоји од докумената ЕУ (закони, декларације, повеље, прописи, директиве, итд.) који дефинишу различите аспекте живота у смислу економије, права, пољопривреде, културе, итд. Поред овог, у раду смо, кроз многобројне промјере, показали сличности и разлике у вези са појединим граматичким и лексичким елементима српског и енглеског језика. Такође смо покушали да одговоримо на питање: Да ли је корпусна лингвистика метод рада или самостална научна дисциплина? Приказали смо предности и мане коришћења корпуса и на тај начин анализирали наш корпус који се доминантно састоји од административне врсте текста. На крају смо дали резултате истраживања као и одређене закључке.

Кључне ријечи: контрастирање, стручна терминологија, корпус, рачунарска лингвистика, документа ЕУ

Научна област: Лингвистика

Ужа научне област: Контрастивна и рачунарско/корпусна лингвистика

Thesis title: Contrastive Linguistics of EU Terminology through the Perspective of Corpus Linguistics

Abstract: This paper is envisaged as the paper where contrastive linguistics is interwoven with corpus methodology. The basic premise is that contrasting of differences and similarities between the two languages (Serbian & English) can be done through corpus linguistics methodology which, together with computational linguistics, is an essential tool in modern analysis of different texts within the scope of linguistics. Our corpus is composed of EU documents (laws, declarations, charters, regulations, directives, etc.) which define different aspects of life in the sense of economics, law, agriculture, culture, etc. Apart from this, in our paper, we have, through numerous examples, shown similarities and differences regarding particular grammatical and lexical elements of Serbian and English. Furthermore, we tried to answer the questions: If Corpus Linguistics is a method of an independent scientific field? We have shown advantages and disadvantages of corpuse use and in such way analysed our corpus which is primarily composed od administrative texts. In the end we have provided some findings and results.

Keywords: contrasting linguistics, terminology, corpus, computational linguistics, EU documents

Scientific field: Linguistics

Narrow scientific field: Contrastive and Computational/Corpus Linguistics

САДРЖАЈ

1	Увод	12
1.1	Подручје рада.....	13
1.2	Почетне хипотезе	13
1.3	Циљеви/проблеми истраживања	15
1.4	Очекивани научни и/или практични допринос.....	16
1.5	Дефиниције и релевантне информације.....	18
1.5.1	Стручни термини.....	19
1.5.2	Електронски корпус	21
1.5.3	Предности и проблеми у вези употребе корпуса	24
1.5.4	Корпусна лингвистика.....	26
2	Контрастивна анализа корпуса (1. дио).....	28
2.1.1	Контрастирање на основу конкорданце	28
2.1.2	Учесталост ријечи	48
2.1.3	Колокације	55
2.1.4	Техничке карактеристике корпуса и програма за његову обраду.....	57
2.1.5	Превођење стручних термина <i>Acquis Communautaire-a</i>	58
2.1.6	Превођење стручних термина помоћу рачунара	63
2.1.7	Машинско превођење стручних термина.....	66
3	Корпусна лингвистика; метод или наука?.....	68
3.1	Дефинисање корпусне лингвистике.....	68
3.1.1	Корпусни приступ наспрам интуитивног.....	73

3.2	Испитивање помоћу WebCorp-а.....	75
3.3	Корпус истраживачких чланака из лингвистике.....	77
4	Контрастивна анализа корпуса (2. дио).....	81
4.1	Увод	81
4.2	Резултати истраживања корпуса (карактеристике нашег корпуса)	82
4.2.1	Административни стил	82
4.2.2	Импersonализовани регистар	83
4.2.3	Декомпоновани предикати	88
4.2.4	Колико можемо вјеровати интуицији?	94
4.2.5	Изражавање забране.....	96
4.2.6	Грешке које смо уочили током анализе корпуса.....	97
4.2.7	Анализа употребе модалних глагола у корпусу	103
4.2.8	Зипфов закон.....	109
4.2.9	Хипов закон.....	111
4.2.10	Карактеристике превођења текстова административног регистра	112
4.3	Контрастивна анализа граматичких елемената у корпусу	113
4.3.1	Одређени члан.....	113
4.3.2	Колокације придјева и именица.....	116
4.3.3	Глаголски облици.....	119
4.3.4	Копуле остајања.....	130
4.3.5	Прошли партицип.....	132
4.3.6	Can, Could.....	137

4.3.7	Пасив	138
5	Закључак.....	141
6	Литература.....	146
7	Прилог 1.	164
7.1	Глосар економских термина.....	164
8	Прилог 2.	165
8.1	Глосар правних термина	165
9	Прилог 3.	166
9.1	Лиценце и слике коришћених програма	166
10	Прилог 4.	173
10.1	Састав корпуса	173
11	Прилог 5.	181
11.1	Интернет странице.....	181
11.2	Прегледи корпусне лингвистике	181
11.3	Референце о корпусној лингвистици	181
11.4	Корпуси	181
11.5	Часописи и базе података	182
11.6	Остали релевантни пројекти.....	183
11.7	Странице за преузимање програма Р и података	183
11.8	Други корисни рачунарски програми	184
11.9	Изрази	184
12	Биографија аутора	185

13	Прилог 6.	186
14	Прилог 7.	187
15	Прилог 8.	188

1 Увод

Овај рад је замишљен као комбинација практичне примјене контрастивне анализе и корпусне лингвистике да бисмо, како нам се чини, на најбољи начин анализирали корпус који се састоји од разноврсних докумената ЕУ. Под документима подразумевамо законе, декларације, уговоре, споразуме итд... Но овом мјесту жељели бисмо да истакнемо посебно један документ који ће чинити окосницу нашег корпуса а то је: Упитник Европске комисије - Информација коју од Владе Црне Горе захтијева Европска комисија у циљу припреме Мишљења о захтјеву Црне Горе за чланство у Европској унији. С обзиром на то да је ово нови документ и као такав није био детаљније анализиран, сматрамо да ће то унијети и неопходну свјезину у наше истраживање.

Овакву врсту корпуса смо одабрали, јер се надамо да укључује неке од најбитнијих критеријума које квалитетан корпус мора да посједује. У првом плану корпус је референтан јер представља пресјек цјелокупног модерног језика, са још једном битном карактеристиком, а то је да је сачињен од различитих стручних тематских цјелина које кроз језик представљају свако модерно друштво (економија, политика, пољопривреда, здравство, наука, спорт, образовање итд.). Друга битна карактеристика ове групе текстова јесте да је састављен радом више десетина људи из различитих дјелова Црне Горе што, надамо се, доприноси лингвистичком тј. лексичком богатству прикупљених текстова који ће се користити у овом раду. Одабиром оваквог корпуса избјегли смо проблем циркуларности, тј. одабир такве врсте текстова чија би анализа показала управо онакве резултате какве смо и прижељкивали. То смо учинили тако што смо одабрали преводе који су превели други колеге

преводиоци, и на тај начин субјективни утицај аутора овог рада је сведен на најмању могућу мјеру. И на крају, документа ЕУ, представљају веома битна документа за све земље бивше СФРЈ и, што је још битније за овај рад, одсликавају актуелни лингвистички тренутак и као таква заслужују одговарајућу пажњу. Све текстове које смо користили у оквиру нашег корпуса налазе се на DVD-у, који се налази на крају нашег рада, како би се могла провјерити научна утемељеност и методолошка ваљаност истог. На крају рада у Прилогу 5 навели смо све интернет изворе као и рачунарске програме који се нам помогли приликом израде овог рада. На овај начин смо на једном мјесту окупили релевантне интернет странице које могу бити од помоћи и остаим колегама у неким наредним њиховим лингвистичким и другим истраживањима.

1.1 Подручје рада

Лингвистика → Корпусна лингвистика → Паралелни корпус оригиналног и преведеног текста докумената ЕУ

1.2 Почетне хипотезе

Разлике у административном дискурсу текстова на српском и енглеском језику ће бити условљене не само формалним граматичко-структуралним разликама између ова два језика, већ и културолошким и системско-правним разликама које се пресликавају и утичу на њих,

Корпус може бити веома користан приликом контрастирања лингвистичких елемената два језика, у овом случају између српског и енглеског језика.

Корпус је много поузданије оруђе у лингвистичким анализама него што је то интуиција изворног говорника, па тако се неке ријечи понављају свега пар пута или се уопште не налазе у корпусу иако вјероватно већина извоних говорника не би тако оцијенила на основу своје интуиције,

Анализа корпуса уз помоћ рачунара може веома брзо и прецизно да нам пружи податке о колокацијама и учесталости лексичких елемената. Које су најчешће ријечи које колоцирају са рјечју нпр. ИЗВЈЕШТАЈ? Веома брзо, уз посједовање одређене количине знања рада на рачунару и са рачунарским програмима, можемо видјети да су то ГОДИШЊИ, ПЕРИОДИЧНИ и ФИНАНСИЈСКИ, што за посљедицу има да се „најочекиванија“ колокација са ИЗВЈЕШТАЈ (конструкција главни глагол + директни објекат) ПОДНИЈЕТИ ИЗВЈЕШТАЈ налази „тек“ на четвртом мјесту,

Корпус нам може рећи да ли се нешто јавља у језичком окружењу или не, али нам не може рећи да ли је то граматички правилно или није правилно у оквиру одређеног језика, (Видјети за примјер *EXPIRE OF* под насловом Проблеми у вези коришћења корпуса)

Недостатак корпуса јесте то што представља језик који не садржи вањезичке елементе,

Учесталост ријечи:

- На основу програма који су намјењени обради корпуса можемо да сазнамо учесталост ријечи које се налазе у том корпусу. На основу ове анализе научници су у стању да анализирају које и какве ријечи се налазе у корпусима који су тематски различити. Тако, на примјер, стручњаци који се баве овом облашћу нашли су да у корпусима који као главну тему имају политику, најчешће ријечи ће бити граматичке ријечи (*the, of, and...*), док ће лексичке ријечи бити чешће у корпусима који су састављени од научних текстова

(Haliday and Martin: 1993: 76-77). Покушаћемо да потврдимо ову хипотезу кроз анализу нашег корпуса.

1.3 Циљеви/проблеми истраживања

Кроз овај рад покушаћемо да:

- Контрастирамо лексичке и реченичне елементе из српског језика и њима одговарајуће преводачке еквиваленте у енглеском језику и обратно,
- Приликом контрастирања докажемо предности и мане корпусно-контрастивног приступа,
- На основу контрастирања анализираћемо преводне еквиваленте и граматичке јединице којима они припадају,
- Анализирамо врсту дискурса који се користи у оваквим типовима административних докумената (дискурсна обиљежја) и да одговоримо на питања која се тичу сличности или разлика између ова два типа регистра у српском и енглеском језику,
- Докажемо да је ова врста регистра веома осјетљива на „слободнији“ превод, јер обично подразумијева текстове чија је примарна сврха да буду информативног типа, а та информација мора бити прецизно пренесена у циљни језик. Ова „осјетљивост“ ће нарочито бити видљива у текстовима облигационог типа гдје једна ријеч може промијенити структуру облигационог односа,
- Покажемо који то лексички и надлексички елементи чине административне текстове различитим од осталих врста дискурса,
- статистички прикажемо и анализирамо учесталост ријечи,

- статистички прикажемо и анализирамо конкорданцу најчесталијих ријечи
- одговоримо на питање да ли је корпусна лингвистика научна методологија рада или је издиференцирана наука у оквиру лингвистике
- направимо глосаре економских и правних термина

1.4 Очекивани научни и/или практични допринос

Највећи научни и/или практични допринос овог рада треба тражити у:

- промовисању контрастивне лингвистике као дијела лингвистике која, по мишљењу аутора овог текста, има највећи потенцијал када се ради о откривању сличности и разлика међу језицима; у нашем конкретном случају говоримо о српском и енглеском,
- дубљем разумијевању преводних еквивалената у оквиру административног дискурса. Преводаца мора да потпуно разумије дубље значење административне лексике која потиче са простора који имају другачије културолошке и законодавне норме. Боље разумијевање ове врсте дискурса води ка бољим преводима, који доприносе бољем разумијевању међу стручњацима који имају додирних тачака са овом врстом докумената.
- изради глосара који ће помоћи свим будућим преводиоцима који ће на основу нашег глосара и анализе боље моћи да разумију начин на који функционише скуп текстова који имају разнородну тематику,
- промовисању корпусне лингвистике и корпусне методологије истраживања,

- наглашавњу корисности употребе рачунара и специјализованих рачунарских програма за потребе разноврсних лингвистичких анализа јер сматрамо да је ово будућност свих таквих анализа,
- анализи актуелних докумената ЕУ који, на овај начин, нијесу били раније анализирани,
- глосарима који су већином тематски везани за области: банкарства, права и економије, а који се налазе при крају нашег рада у виду додатака. У њима са налазе 10.442 одреднице од којих су већина стручни термини заједно са њиховим преводним еквивалентима.

Наш корпус се састоји из текстова на српском и енглеском језику, па тако говоримо о двојезичном корпусу и из те чињенице можемо извући и следеће доприносе:

- Паралелни корпуси омогућавају бољи увид у језике који се проучавају,
- Могу се користити за сврхе упоредних анализа, као и да се боље проуче типолошке и културолошке разлике, али и да се нађу универзални елементи,
- Омогућавају да се уоче разлике између оригинала и превода, тј. између изворног и преведеног текста,
- Могу се користити за друге практичне примјене као што су: лексикографија, учење језика и превођења.

Међутим, као да је и овај дио корпусне лингвистике осуђена на терминолошка неслагања међу научницима. Наиме, Алтенберг (1996) и Грангер (1996: 38), нашу врсту корпуса би назвали „*translation corpus*“, дакле, корпус преведених текстова, док би паралелни корпус био сваки онај корпус који је направљен коришћењем исте методе узимања узорка

текстова. За потребе овог рада, сложићемо се са групом научника¹ који под паралелним корпусом сматрају онај који је састављен од изворног текста и његовог превода на један или више језика.

У раду ћемо користити контрастивну лингвистику, прецизније контрастивну анализу, која се дефинише као методолошки приступ који помаже лингвисти-аналитичару да утврди у којим аспектима су испитивани језици слични, а у којима се разликују. Такође, поред знања из контрастивне лингвистике, користићемо и знања других лингвистичких дисциплина које су се нашле у контрастивном теоријско-методолошком опусу. Ту, прије свега, мислимо на когнитивну лингвистику, прагматику, лингвистику корпуса итд. Путем контрастивне лингвистике у стању смо да дубље анализирамо ова два језика што може бити од користи будућим преводиоцима и методичарима наставе.

1.5 Дефиниције и релевантне информације

У нашем раду веома често ћемо користити одговарајућу стручну терминологију који је неопходно ближе дефинисати како би се на прави начин могао пратити ток овог рада. Такође ћемо се кратко осврнути и на поља истраживања и рада која су релевантна за наш рад и боље разумијевање истог.

¹ McEneyr и Wilson (1997: 57), Baker (1993: 248, 1995, 1999) и Hunston

1.5.1 Стручни термини

Стручна терминологија представља специјализовани вокабулар који се користи у оквиру одређеног струковног поља. У оквиру тих струковних поља, термини могу имати једно или више значења која се могу разликовати од њиховог стандардног значења у оквиру уобичајене употребе. Дакле, у већини случајева говоримо о проширењу семантичког поља. Жаргон је слична појава, али се сматра да је мањег степена формалности од употребе специјализованог вокабулара:

„Жаргоном се може назвати сваки неформални и претежно говорни варијетет неког језика који служи за идентификацију и комуникацију унутар неке друштвено одређене групе – по професији, социјалном статусу, узрасту и слично – чије чланове повезује заједнички интерес или начин живота, а која уз то може бити и територијално омеђена“. (Р. Бугарски: 2003: 9)

Стручна терминологија се често доживљава као прва-очигледна демаркациона линија између стручњака за одређену област и лаика. Та линија може бити толико широка да представља непремостиву лингвистичку препреку у комуникацији стручњака и оних који то нијесу. Да би неки термин био препознат као стручни мора да буде формализован. То значи да су прецизни стручни термини, као и њихове дефиниције формално препознати, документовани, и представљају дио обуке за нове стручњаке из дате области. Међутим, граница између стручних термина који се користе искључиво у оквиру специјализованих области и термина који су у општој употреби није константна. Наиме, неки стручни термини током време постају дио општег рјечника. Ово је

веома изражена појава у свијету рачунара и интернета. Па су тако нпр. стручни термини порт (енг. *port*), фајервол (енг. *firewall*) постали дио општег језичког блага које учестало користе (са већим или мањим степеном разумијевања) и особе са скромним познавањем рачунара. Овакав прелаз из ексклузивне у општу употребу се десио јер су и рачунарски елементи коју су везани за ове термине доспијели у општу употребу:

„Компјутерски жаргон је, на пример, нов. Ако ослушнете разговор двоје младих људи који много времена проводе за компјутером, који тече нормално, он је прожет компјутерским жаргоном толико да је потпуно неразумљив некое ко не ради са компјутерима“.²

Стручна терминологија је настала из потребе да се идеје пренесу на сажет и прецизан начин, што представља очигледан примјер лексичке економичности. Други разлог јесте и то, што се на овај начин из струковних група „искључују“ сви они који њима не припадају, било да припадају другој струци (дифернецијација) или су лаици (ексклузивитет). Наравно, ово може да проузрокује потешкоће у комуникацији, па тако можемо имати проблема са разумијевањем следеће реченице:

Тросистемски бас рефлекс звучници са хибридном дуал wooфер-ом и веома ефикасним рожнатим твитером...³

Изразићемо сумњу да би велики број потенцијалних купаца имао идеју који конкретно производ купују, ако би им се тај производ представио на начин који је изложен у овом примјеру. Међутим, управо на овај начин се дјелује на психологију купца којима се ове ствари приказују и дефинишу на овакав начин како би добиле макар на звучности.

² Бугарски, Ранко: <http://arhiva.glas-javnosti.co.yu/arhiva/2003/12/28/srpski/I03122701.shtml>
28.02.2011. у 14.55.

³ Овај примјер је преузет из Сонијевог каталога производа; Јесен/Зима 2007 – 2008.

1.5.2 Електронски корпус

Корпус се може дефинисати поштујући различите критеријуме, али се стручњаци из области корпусне лингвистике слажу да корпус представља скуп текстова који се природно појављују у језику и као такви варирају од неколико реченица до колекције дисукрса и транскрипта, а све у сврху лингвистичке анализе. У нашем случају, конкретно се ради о електронском корпусу који, у односу на већ поменути дефиницију, има још један детаљ. Карактеристика електронских корпуса јесте да се могу сачувати и да им се може приступити електронским путем. На тај начин у стању смо да обрадимо много већу количину података, дакле електронски корпус има своју квантитативну вриједност. Квалитативна вриједност корпуса лежи у комбинацији прецизности анализе специјализованих рачунарских програма са анализом особе која се бави корпусном лингвистиком. Корпус је плански скуп текстова који су прикушљени према одређеним критеријумима (референтност, уравнотеженост, организованост) који су унапријед одређени и који се морају поштовати како би корпус испунио намјену за коју је и предвиђен. Ово је елеменат који корпус разликује од неке електронске архиве или библиотеке који чувају поменуте текстове реди неке њихове унутрашње вриједности. Један од критеријума које корпус мора да задовољава, јесте да се састоји од језика који је произведен у природном комуникативном окружењу тј. да није произведен само у сврху укључивања у корпус. Тако нпр. корпус може да се састоји од текстова који се налазе у новинским чланцима. Њихова главна функција јесте да пренесу одређену поруку читаоцима, а не да их користе лингвисти приликом својих истраживања. У нашем раду користећемо паралелни корпус који се састоји од изворног језика (српски)

и превода (енглески). На основу превода можемо да направимо упоредну анализу елемената који су нам од значаја.

1.5.2.1 Врсте корпуса

Основна подјела корпуса јесте на двије врсте: општи и специјализовани корпус.

- **Општи корпуси** садрже и писани и говорни језик, а такође могу укључивати текстове који воде поријекло из различитих земаља. Дакле, њихова основна поставка јесте да укључе што шири спектар текстова. Због ове чињенице, овакви корпус су обично знатно већи од осталих врста корпуса, па тако корпус *Bank of English* садржи преко 400 милиона ријечи и састоји се од многих мањих корпуса који обухватају различите подобласти.
- **Специјализовани корпуси** садрже текстове који су ограничени на одређени тип текста, нпр. академске чланке о одређеним темама и на тај начин представљају ту врсту текста. Корпус може бити ограничен и у смислу временског оквира којим се бави, па тако имамо корпусе језика који се говорио током неких од прошлих вјекова. Корпус може бити ограничен и друштвеним окружењем из којег одабрани текстови потичу (корпус разговора у парламенту ЕУ). Вјероватно два најпознатија специјализована корпуса су: *Cambridge and Nottingham Corpus of Discourse in English* (CANCODE; који садржи неформални регистар британског енглеског) и *Michigan Corpus of Academic Spoken English* (MICASE; који садржи усмени академски регистар).

- **Упоредни корпуси** садрже текстове из истих области различитих језика (нпр. новинарски текстови на енглеском и француском), или садрже варијетете једног језика (британски, амерички и аустралијски енглески). На основу ових корпуса се могу радити контрастивне анализе елемената који интересују поједине научнике.
- **Паралелни корпуси** садрже колекцију преведених текстова са једног на други језик, и као такви могу да имају изузетну вриједност за преводиоце и контрастивне лингвисте. Овакав тип корпуса ћемо користити у овом раду, гдје се конкретно ради о паралелном корпусу докумената европске уније и њихових превода на наш језик. На основу ових корпуса могу се уочити сличности или разлике које постоје међу посматраним језицима.
- **Педагошки корпуси** се састоје од текстова који се презентују особи која учи неки језик. На основу њега ученик може лакше научит одређену употребу неке ријечи или пак научити разлику између двије ријечи које се чине сининимичним.
- Подјелу даље можемо наставити са **дијахроним и синхроним корпусом**. Дијахрони корпус презентује језик кроз његове промјене током времена, док синхрони корпус даје изглед језика у одређеном времену. Затим можемо да говоримо о једнојезичним и паралелним корпусима и на основу њихових назива лако можемо да закључимо да се први бави једним језиком, док је други окренут паралелним текстовима који потичу из различитих језика. Примјере оваквог корпуса представљају преводи дебата у парламенту ЕУ (које су преведене на 23 језика), или Канадски Хансард корпус, који садржи дебате канадског парламента које су преведене на енглески и француски.

За све врсте корпуса битно је нагласити да се морају непрекидно ажурирати како би на прави начин одсликавали актуелни језички тренутак.

1.5.3 Предности и проблеми у вези употребе корпуса

1.5.3.1 Предности коришћења корпуса

- Зашто је корпус битан, иако има поједине недостатке; један од главних адута добро компонованог корпуса јесте то што он даје вјернију слику о језику него што то може да нам пружи говорникова интуиција, на основу корпуса можемо лакше да додјемо до колокација и фраза него што то може да уради изворни говорник „из главе“. Готово је немогуће одредити колико је нека ријеч фреквентна у одређеним жанровима, осим паушалних изјава веома/мало учестала ријеч. Па тако имамо код Халидеја (1993: 3) да су изворни говорници свјеснији да је глагол *GO* чешћи од глагола *STROLL*, али никако не могу да знају које граматичке категорије су учесталије од других. Наравно, из ове приче не треба закључити да интуиција није битна. Понекад је интуиција незамјенљиво средство када се праве неке лингвистичке генерализације.
- Корпуси се користе приликом учења страног језика, јер могу да пруже детаљнији увид у језик као што је фразеологија. Такав увид говорнику који учи неки страни језик може бити недоступан на интуитивном нивоу. Тако на основу истраживања Миндта из 2000. знамо да у колоквијаном енглеском, око 80% фраза које имају будућу референцу гради се помоћу модалног глагола *will*, *be going to*

конструкција чини 10% фразе, док садашње трајно вријеме чини око 5% фразе.

- Упоредни корпуси (корпуси изворног језика и превода) су веома корисни за преводиоце јер на тај начин могу да дођу до превода одговарајућих термина, односно могу да сазнају да ли преводни еквиваленти постоје (и које су има варијанте) или не постоје. Тако смо на основу нашег контрастивног истраживања на магистарском студијама током 2009. дошли до закључка да се одређени члан *the*, приликом превода на наш језик, најчешће не преводи (у око 80% случајева), или се преводи неком од демонстративних замјеница (тај/та/то; овај/ова/ово), у око 15% случајева.

1.5.3.2 Проблеми у вези коришћења корпуса

- Свакако, морамо бити свјесни и недостатака корпуса: корпус нам не може дати одговор да ли је нешто могуће или не, већ само колико је учестало. Па тако код Синклера (Синцлаир 1991: 17); Бибера (*Biber et al* 1998: 3) налазимо да се конструкција *EXPIRE OF* и *Bank of English* налази у пет линија, тј. 0,1% од цјелокупног корпуса. Иако ова информација може бити мање или више корисна, не говори нам да ли је конструкција *EXPIRE OF* граматички тачна.
- Корпус нам не може рећи ништа ван самог себе. Колико год корпус тежио да буде репрезентативан, увије ће остати неки дио језика који ће бити „не покривен“. Иако тај дио језика може бити мањи дио укупног система језика, опет остаје чињеница да је апсолутна репрезентативност немогућа.
- Корпус нуди информације, али их не може објаснити. Ово је идеално мјести за интуицију да се укључи у процес анализе.

- Корпус не садржи вањезичке елементе (говор тијела, интонацију и остале паралингвистичке информације) и као такав може се сматрати као скупом текстова који су ван контекста. Наравно, ово добија на значају када поменемо чињеницу да иста ријеч има различиту конотацију у зависности од тога како је емоционално обојена. Управо из овог разлога, интуиција је назамјенљиво „оруђе“ приликом анализе било које врсте корпуса.

1.5.4 Корпусна лингвистика

Основна корпусна истраживања налазимо на почетку 20. вијека када су стручњаци корпистили веома скромне корпусе текстова који су се налазили у кутијама од ципела и који су били састављени од цедуљица на којим је текст био обично ручно унесен. Овакви корпуси су се обично користи за истраживања везана за фонетику (Линг: 1999: 240). Средином прошлог вијека Џесперсон (*Jasperson* 1949) и Фрајс (*Fries* 1952) користе „папирни“ корпус како би спроводили истраживања на пољу граматике. Током касних педесетих прошлог вијека, корпусна методологија је била изложена оштрој критици (видјети Чомски 1962) због наводне „пристрасности“ корпуса, тј. могућности да се узме корпус који би лако доказао било какве почетне претпоставке. Велики дио тих и таквих критика је био оправдан јер су коришћени корпуси били исувише мали да би се из њихове анализе могли извући било какви озбиљнији закључци. Овакав начин обраде података, је готово закономјерно био пристрасан. Међутим, протоком времена, квантитативни елементи корпуса су превазишли могућности које посједује човјек или група људи, па је увођење рачунара у ову проблематику био сљедећи логичан корак. Веома брзо рачунари су понудили идеалне опције за манипулацију корпусима.

- **Омогућили су брзу и лаку обраду велике количине текстова**
- **Тачну и конзистентну обраду тих текстова**
- **Избјегнута је евентуална пристрасност приликом лингвистичких анализа, самим тим је повећана поузданост добијених података**
- **Чување огромних количина података које неизбежно прате такве врсте анализа. Од тада постоји нераскидива веза између рачунара и корпуса.**

Први модерни корпуси састављени су почетком шездесетих година и то је био *Brown* корпус. Од осамдесетих година па на даље, величина и број корпуса, као и анализа који су засновани на истим се веома повећао (*Johansson: 1998: стр. 3*)

2 Контрастивна анализа корпуса (1. дио)

2.1.1 Контрастирање на основу конкорданце

Најчешће коришћено оруђе у корпусно-лингвистичким анализама јесте тражење конкорданци. Конкорданца представља шири контекст у којем се нека ријеч налази, односно тражење кључне ријечи у контексту (енгл. *KWIC = Key Word In Context*). Приликом прављења табела и тражења конкорданци, обично говоримо о ријечи која се налази у оквиру великог дијела реченице или пак цијеле реченице. Та ријеч се може означити како би се лакше уочила. Такође, многи рачунарски програми одзвољавају да се тачно дефинише број ријечи које ће „окруживати“ кључну ријеч (нпр. двадесет са лијеве и десне стране), или се пак дата ријеч може ставити на почетну или крајњу позицију. Конкорданца омогућава да се уоче правилности које би нам иначе промакле у свакодневним текстовима.

Тако нпр. можемо дати следећу табелу у којој је енглеска ријеч *HEALTH* кључна:

Табела 1. Конкорданца енглеске ријечи *HEALTH* на основу резултата из нашег корпуса

entities and entrepreneurs that perform HEALTH-care activities (Official Gazette of
and practical vocational work in HEALTH-care institutions which meet the
on conditions that have to be met by HEALTH-care institutions which perform
spreading of infectious diseases within HEALTH-care institutions, other

legal entities
and, therefore, patients do not go to a HEALTH-care unit, which in turn make it
of sanitary-epidemiology department in HEALTH-care units and the Institute for Public
Having organized implementation of HEALTH-promotional and preventive measures for
on the prevention and the reduction of HEALTH-related harm associated with drug
further development and improvement of HEALTH-related statistic and information
the competent institution, is monitoring HEALTH-related statistics data for all three
the Public Health Institute manages HEALTH-related statistics data in the public
Ministry of Health i.e. Department of HEALTH-sanitary inspection affairs. Competent
requirements (the Institute for Public HEALTH - Podgorica). List of skills Doctor
on 28th April 2009. The Ministry of HEALTH - the IHR team located in the country,
and organizations such as Market, HEALTH - Sanitary, Phytosanitary, Tourist and o

Или можемо посматрати конкорданцу наше ријечи ЗДРАВЉЕ:

Табела 2. Конкорданца ријечи ЗДРАВЉЕ на основу резултата из нашег корпуса

верзији на сајту Института за јавно ЗДРАВЉЕ - Подгорица: www.ијзцг.ме). б) Сваке
у Подгорици лоциран Институт за јавно ЗДРАВЉЕ (установа која се у домену свог рада
доставља пријаве Институту за јавно ЗДРАВЉЕ (ИЈЗ) о поремећају здравља од јавно
институционалне бриге везане за ментално ЗДРАВЉЕ (тренутно практично једини постојећи
здравља на предлог Института за јавно ЗДРАВЉЕ а након разматрања Комисије за заразне
У савјетовалишту за репродуктивно ЗДРАВЉЕ ће се у 2009.години реализовати Програм
и обрађују у Институту за јавно ЗДРАВЉЕ Црне Горе објављују се на националном
Према истраживањима Института за јавно ЗДРАВЉЕ Црне Горе у 2008. години, свега 11,5%
јавне водоводе врши Институт за јавно ЗДРАВЉЕ Црне Горе, у сарадњи са хигијенско
у стратешком плану Института за јавно ЗДРАВЉЕ Црне Горе. Активности промовисања
поступци процјене ризика по живот и ЗДРАВЉЕ даваоца и

примаоца. Пресађивање дјелова
критеријумима прихватљив ризик по ЗДРАВЉЕ даваоца, те да постоји вјероватноћа за
стилова живота. Институт за јавно ЗДРАВЉЕ дефинисао је програме промоције здравља
играчке – због потенцијалног ризика по ЗДРАВЉЕ дјеце и информација о небезбједности из
док је преосталих 57,8% сматрало своје ЗДРАВЉЕ добрим или веома добрим. Значајно је

2.1.1.1 Шта можемо да закључимо на примјеру конкорданце?

2.1.1.1.1 *Контрастирање на основу конкорданце*

Као што смо већ нагласили корпус нам не може дати одговор на питање да ли је нешто могуће или није. Оно што је сасвим могуће јесте да се у корпусу нађу ријечи или реченичне конструкције које би други изворни говорници српског или енглеског одбацили као нетачне. То је оно што Свон (1994) назива „демаркација“ и наводи да нам корпус не може понудити демаркацију између тачног и нетачног. Умјесто демаркације, корпус нам може понудити „типичну или централну“ употребу посматране ријечи (Хенкс 1987: 124-125; Синклер 1987а: 108, 114; Синклер 1991: 17). Ови аутори често типичну и централну употребу изједначавају и дефинишу је када желе описати најчешће значење, колокације или фразе појединачне ријечи. Да бисмо илустровали овакву типичну употребу ријечи, испитаћемо шта нам корпус каже о томе које ријечи најчешће фроммирају колокације са енглеском ријечи *law/s* и нашом ријечи закон/а.

Овај пар ријечи одабрали смо јер је ово најчешће понављана лексичка ријеч (именица) у нашем корпусу.

Резултати контрастирања показују сљедеће:

2.1.1.2 Контрастирање кроз конкорданцу

2.1.1.2.1 Контрастирање именица Закон и Law

Да је именица ЗАКОН веома често праћена будућим обликом глагола бити или хтјети:

1. не за уређење банкарског сектора. Нови ЗАКОН биће у високом степену усаглашен са Дир
2. тјевима у И кварталу 2008. године. Овај ЗАКОН биће усклађен са такозваном нотификацио
3. итошћу рада спроводилаца поступка. Овај ЗАКОН биће усвојен у трећем кварталу 2008. го
4. донијет у И кварталу 2008. године. Тај ЗАКОН ће бити главни правни основ за доношењ
5. дња система акредитације у Црној Гори. ЗАКОН ће бити усклађен са захтјевима члана 77
6. ени у земљама ЕУ од јула 2003. године, ЗАКОН ће бити ускоро измјењен у циљу усагла
7. и њиховом међународном регистровању. ЗАКОН ће бити усвојен у другом кварталу 2008.

8. декларацијом о правима човјека; Такође, ЗАКОН ће обезбиједити досљедну примјену рати
9. степена развијености региона и општина. ЗАКОН ће припремити Секретаријат за развој, у
10. донијет у првом кварталу 2008. године. ЗАКОН ће прописати кретање слијепог лица уз

Да је именица ЗАКОН најчешће праћена глаголом бити (у свом облику који је прилагођен броју и гласи ЈЕ):

1. Са једним дијелом европских прописа ЗАКОН је дјелимично усклађен. Институционални
2. њавања међународно преузетих обавеза. ЗАКОН је добрим дијелом усаглашен са правом Е
3. редставља Закон о управљању отпадом. ЗАКОН је донесен 2005. године (Сл. лист РЦГ б
4. не фармацеуте (Сл. лист РЦГ бр. 26/06). ЗАКОН је донијет на основу препорука и деклар
5. нама Закона о спрјечавању прања новца ЗАКОН је допуњен одредбама о спрјечавању фи
6. дних права и судска заштита тих права). ЗАКОН је хармонизован са сљедећим међународн
7. ама и обавјештењима ЕУ из ове области. ЗАКОН је измијењен у мају 2007. године (Сл.
8. овим тржишним и конкурентским условима. ЗАКОН је конципиран тако да је пружао основу

9. ијских услуга и производа у Црној Гори. ЗАКОН је конципиран тако да је пружао основу
10. ални промет наркотичких дрога. Такодје, ЗАКОН је неопходно ускладити са Стратегијом и
11. бољи услови за развој и конкурентност. ЗАКОН је обезбиједио оквир да се у поступку п
12. ор и инострана информативна дјелатност. ЗАКОН је омогућио да се слободним протоком ин
13. зрада новог Закона о културним добрима. ЗАКОН је планиран да буде усвојен 2008. годин
14. и други пут до 31. марта 2003. године². ЗАКОН је почео да се примјењује од 1. априла
15. гулативе које уређују ову област. Нови ЗАКОН је потпуно усаглашен са Директивом 2005

Да је именица ЗАКОН праћена препозицијом О (то је уједно и најчешћа конструкција која се јавља у нашем корпусу):

1. лан за њену имплементацију. Усвојен је ЗАКОН о надзору државне границе, спроведена ј
2. и сарадња служби на државној граници. ЗАКОН о надзору државне је дјелимично усаглаш
3. ржавању квалитета њихових дјелатности ЗАКОН о научно - истраживачкој дјелатности пр
4. , од стране државе врши се, у складу са ЗАКОН о невладиним организацијама и Законом о

5. дио чини и Програм правосудног испита. ЗАКОН о нотарима (“Сл. лист РЦГ бр. 68/05) ув
6. за организацију спровођења едукације. ЗАКОН о нотарима (Сл. лист РЦГ бр. 68/05) уст
7. их функција имају право по основу рада. ЗАКОН о нотарима (Сл. лист РЦГ бр. 68/05) уво
8. адвокатури (Сл. лист РЦГ бр. 79/06), ЗАКОН о нотарима (Сл. лист РЦГ бр.68/05), Зак
9. , Закон о извршењу кривичних санкција, ЗАКОН о нотарима и подзаконски акти: Судски п
10. дини утврдити Предлог измијена и допуна ЗАКОН о нотарима којим ће се створити правне
11. Средоземља (ИВ квартал 2008. године): ЗАКОН о ратификацији Конвенције о заштити и
12. урању (Сл. лист РЦГ бр. 78/06 и19/07), ЗАКОН о обавезном осигурању у саобраћају (Сл
13. кид трудноће (Сл. лист СРЦГ бр. 16/80). ЗАКОН о обезбјеђивању крви (Сл.лист ЦГ бр. 1
14. горице. 3.6.1.2. Краткорочни приоритети ЗАКОН о објављивању прописа и других аката
15. на ЦГ је у априлу 2005. године усвојила ЗАКОН о обједињеној регистрацији и систему и
16. н, то су Закон о ваздушном саобраћају и ЗАКОН о облигационим односима и основама свој
17. ом саобраћају (Сл. лист СРЈ бр.12/98) и ЗАКОН о облигационим односима и основама свој

18. првом квартал 2009. године усвојиће се ЗАКОН о облигационим односима и основама свој
19. ом испиту, Законик о кривичном поступку, ЗАКОН о облигационим односима, Закон о рјешав
20. Г бр. 49/07). У 2008.години донијеће се ЗАКОН о облигационим односима, Закон о својин
21. н о архивској дјелатности (1991, 1994), ЗАКОН о обнови и ревитализацији старих градов
22. тресу од 15. априла 1979. (1984, 1986), ЗАКОН о обнови споменичког фонда Котора (1991
23. заштити (Сл. лист РЦГ бр. 78/05); Општи ЗАКОН о образовању и васпитању (Сл. лист РЦ
24. и у складу са посебним прописима. Општи ЗАКОН о образовању и васпитању (Сл. лист РЦ
25. едине нивое и области образовања Општи ЗАКОН о образовању и васпитању (Сл. лист РЦ
26. а различитим нивоима образовања. Општи ЗАКОН о образовању и васпитању дефинише пит
27. ојена је Стратегија одбране и донесени: ЗАКОН о одбрани (Сл. лист РЦГ бр. 47/07) и За
28. ити свједока (Сл. лист РЦГ бр. 65/04) и ЗАКОН о одговорности правних лица за кривична
29. та података, евиденције и финансирање. ЗАКОН о одговорности правних лица за кривична
30. екцијском надзору, и то Закон о дувану, ЗАКОН о ограничавању употребе дуванских прои

У енглеском језику именица *LAW* је много рјеђе праћена модалним глаголом *WILL* + инфинитив без *TO*. Овакав резултат се вјероватно најбоље може објаснити чињеницом да оваква конструкција није довољно „јака“ за будуће вријеме би се на одговарајући начин изразила облигација која је неопходна у текстовима правне природе. У дијелу корпуса на енглеском језику именица *LAW* је најчешће градила колокације са:

***LAW* + садашње просто вријеме (уобличено кроз именицу која је у трећем лицу са обавезним наставком ES/S)**

1. opted in the I quarter of 2008. The new *LAW* aims at establishing state control for
2. 56 National Program for Integration The *LAW* contains measures for the preventi
3. property law in one place. To wit, the *LAW* covers issues currently regulated
4. of importance for waste management. The *LAW* defines packaging waste as packagi
5. nt. 93 Government of Montenegro e/ The *LAW* eliminates from the legal system s
6. man and Minority Rights Protection. The *LAW* emphasizes the role of civil secto
7. ction of Human and Minority Rights. The *LAW* emphasizes the role of the civil s

8. t (Public Procurement Directorate), the LAW ensures greater transparency, whic
9. rly transnational organized crime. This LAW ensures establishment of a security str
10. passed within 90 days from the day this LAW enters into force. Legislation By the e
11. ace the previous progressive rates. The LAW envisages a transitional period for app
12. d by national and foreign persons. This LAW envisages adoption of the National Prog
13. the issue of property in general. This LAW envisages the establishment of town as
14. and guarantee) are also envisaged. The LAW envisages the obligation for excise tax
15. ducer and importer of excise goods. The LAW envisages the possibility for excise li
16. men in all spheres of social life. The LAW establishes a public administration bod
17. to court (justice accessibility), this LAW establishes some institutes such as ass
18. itoring of transition and reforms. This LAW establishes the basic principles underl
19. 1/88); 167 Government of Montenegro The LAW establishes the harmonization of legisl

20. e implementation of these measures. The LAW establishes the harmonization of legisl

LAW + глагол TO BE (како у садашњем тако и у прошлом облику, најчешће, као помоћни глагол у функцији грађења активних или пасивних форми простог садашњег и простог прошлог времена).

1. ts of children in family relations. The LAW is approximated with basic principles o
2. to viticulture and wine production. The LAW is based on the EU Acquis and is partly
3. Government of Montenegro Concept of the LAW is based on the modern solutions for th
4. on and trade in safe food and feed. The LAW is based on the principle of reliable,
5. ments on human rights and freedoms. The LAW is based on the principles: freedom of
6. ich the implementation of international LAW is carried out. Organization and manner
7. omic reforms. Global concept of the new LAW is change of public administration role
8. specified within EC Treaty, so that the LAW is completely harmonized with the prima
9. tion at the Parliament session, a draft LAW is considered by the competent committe

10. he constitutional basis for passing the LAW is contained in Article 68 of the Const
11. of legality and application of lenient LAW is contained in articles 33 and 34, whi
12. vice provided. The fourth change of the LAW is currently being conducted for the pu
13. eded. The drafting and adopting of this LAW is envisaged within two years following
14. ne 2009. The aim of the adoption of the LAW is establishment and development of tra
15. ction of hydrocarbons. Adoption of this LAW is expected in 2008. Montenegro will st
16. prohibited by the Constitution and the LAW is free in Montenegro. Restriction of h
17. ational Program for Integration The new LAW is fully harmonized with Directive 2005
18. nagement in bankruptcy procedures. This LAW is harmonized with the Council of Minis
19. rate Celex No. 32002L0087. The proposed LAW is highly in compliance with the Direct
20. ermine that procedure for adoption of a LAW is initiated by submitting a draft law.
21. tion, the Law on the Amendments to this LAW is intended to harmonize the text of th

22. the termination of the proceedings. This LAW is not completely harmonised with the D
23. framework of their sphere of work. The LAW is partially aligned with the Directive
24. Mainly as a result of the fact that the LAW was not adopted at the time, the harmon
25. exempt excise goods users. The mentioned LAW was partially harmonized with the EU di
26. Republic of Montenegro No. 26/06). The LAW was passed based on recommendations
27. transport of passengers and cargo. The LAW was passed in 1995 and it is harmonized
28. market and competitive conditions. The LAW was prepared in a way that provided the
29. , no 25/06) from April 2006, the former LAW was revised in segments on freedom of a
30. work With regards to the fact that this LAW was taken over from the federal level,

Међутим именица LAW је најчешће колоцирала са препозицијом ON што представља фразу која је у овом контексту анализе формалних докумената сасвим очекивана:

1. plication ART. Deadline: IV Q 2008. The LAW on Abortion Conditions and Procedure

2. framework Montenegro still applies the LAW on Air transport (Official Gazette of F
3. improving freedoms of information. The LAW on Amendments to the Law on
4. n and control system in Montenegro, the LAW on Animal Identification and Registrati
5. documents issued in compliance with the LAW on Asylum, and temporarily withheld doc
6. will depend on the solutions of the new LAW on Banks and bylaws pertaining to this
7. slative framework The Article 10 of the LAW on building construction (OG Republic o
8. with respect to the application of the LAW on Business Organizations Insolvency. T
9. and bylaws. Law on consumer protection; LAW on competition protection; Law on touri
10. pension insurance, the adoption of the LAW on Compulsory Pension Insurance based o
11. te of RMNE", Nos. 78/06 and 19/07), the LAW on Compulsory Transport Insurance
12. provisions of the Law on Insurance, the LAW on Compulsory Transport Insurance and t
13. ns. Amendments to the Law on Insurance, LAW on Compulsory Transport Insurance and

14. court; Company Law, Law on Securities; LAW on Copyrights and Related Rights; Intel
15. ented in accordance with changes in the LAW on courts and Law on State Prosecutor F
16. s Legislation In 2010 - 2012 period New LAW on courts and Law on state prosecutor w
17. n changes and amendments of the General LAW on Education and Upbringing from
18. financing, etc. Changes of the General LAW on Education and Upbringing have been i
19. t control list, (iii) Amendments to the LAW on Foreign Investments, (iv) Amendments
20. interest groups in Montenegro. The new LAW on Foreign Nationals Employment and
21. ed Persons; deadline: I quarter of 2008 LAW on Foreign Nationals Labour and Employm
22. ucts (RAPEX). After the adoption of the LAW on general safety of products, the Mini
23. orities Legislation Adoption of the new LAW on Genetically Modified Organisms (plan
24. d Water Resource Management applies the LAW on Genetically Modified Organisms
25. is Law, etc. During the drafting of new LAW on Genetically Modified Organisms provi

26. s. 539 Government of Montenegro The new LAW on Genetically Modified Organisms
27. international standards in this field. LAW on International Legal Assistance in Cr
28. are not applied in Montenegro. The new LAW on Intervention Supplies was enacted (O
29. al Gazette of RMNE", No. 06/02) and the LAW on Investment Funds ("Official Gazette
30. on of Montenegro and EU Directives, the LAW on Occupational Health and Safety will
31. gulations for the implementation of the LAW on Postal Services, criteria for determ
32. quipment or provide services. Under the LAW on Postal Services, the Agency powers i
33. services under the terms defined by the LAW on Postal Services. In addition, it coo
34. (Off. Gazette of RMNE, No. 80/04); the LAW on Pre-School Education (Off. Gazette o
35. prepared 1 rulebook on the basis of the LAW on preschool education and upbringing.

Као што можемо да закључимона основу оваквог контрастирања постоји прилична сличност између енглеских и примјера на српском језику уз двије битне напомене:

- У енглеском језику будуће вријеме се много мање користи приликом дефинисања облигација. У нашем корпусу, на узорку од 4350 колокација са понављањем које се односе на именицу *LAW*, само ограничен број колокација је садржао помоћни глагол *WILL*. У огромном броју случајева, именица *LAW* је колоцирала са помоћним глаголом *SHALL* што у потпуности одговара енглеском схватању изражавања облигације. Поред овог, можемо закључити да остали глаголи који прате именицу *LAW* су, прије свега, они глаголи који изражавају забрану (закон забрањује, не допушта, спрјечава, итд.), глаголи који изражавају допуст (закон допушта, дозвољава, омогућава, итд.), што је природно и одговара најчешћем контекстуалном окружењу у којем се налази именица *LAW* (или ЗАКОН).
- Именица *LAW* најчешће колоцира са препозицијом *ON* док у нашем језику са именицом ЗАКОН најчешће колоцира препозиција *O*, што представља разлику која постоји између правног регистра ова два језика. У нашем језику не постоји конструкција ЗАКОН НА (што би био формални преводачки еквивалент енглеске структуре *LAW ON*) из разлога што то представља граматички нетачну конструкцију због погрешне употребе препозиције која не одговара нашој правној пракси и стандарду изражавања. Исто важи и у супротном правцу, нигдје нијесмо нашли конструкцију *LAW ABOUT*, иако би та конструкција формално била граматички исправна у енглеском језику, али она не одговара енглеској правној пракси и начину изражавања и, самим тим, се не налази у нашем корпусу.

2.1.1.3 Шта можемо да закључимо на примјеру контрастирања учесталости ријечи?

Који су преводни еквиваленти енглеске ријечи *law/s* и каква је њихова учесталост? На основу нашег корпуса долазимо да сљедећег закључка:

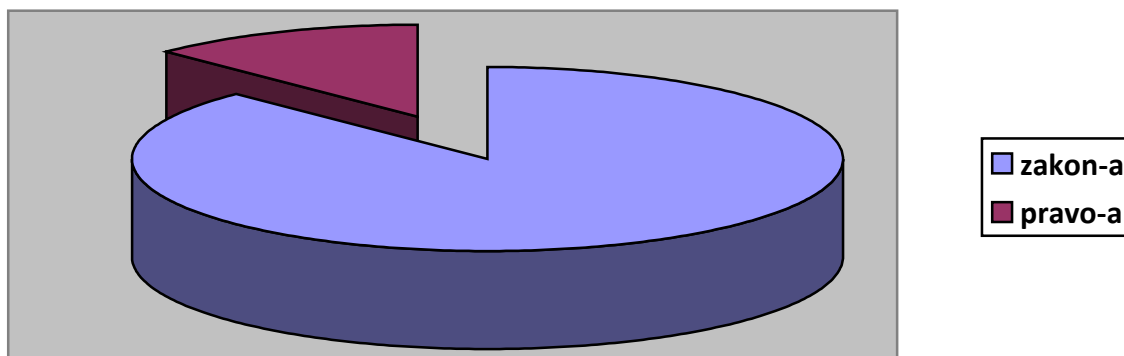
Табела 3. Број понављања ријечи *law/s* у текстовима нашег корпуса

Енглеска ријеч	Број понављања
<i>law/s</i>	73064

Табела 4. Преводни еквиваленти енглеске именице *law/s* и њихов број понављања у нашем корпусу

Преводни еквиваленти	Број понављања ⁴
закон-а	62115
право-а	8791

⁴ У табелу смо укључили само два најчешћа превода



Дијаграм 1. Однос између два најчесталија преводна еквивалента енглеске ријечи *law/s*

На основу дијаграма можемо лако да закључимо на који начин се преводи поменута енглеска ријеч. На овај начин смо показали улогу корпуса за преводиоце који на овај начин могу анализирати, пронаћи или потврдити преводне еквиваленте. Наиме, приликом прављења корпуса један од најбитнијих елемената које желимо да видимо у неком корпусу јесте његова електронска верзија која у оквиру себе садржи и статистичку анализу преводилачких еквивалената, тј. Процент учесталости неке ријечи која се јавља као превод оригиналне ријечи. Уколико је одређена ријеч статистички веома учестала као превод (у смислу тога да се веома често јавља као преводни еквивалент) онда са великом дозом сигурности можемо да устврдимо да је она, условно речено, најбољи преводни еквивалент у том контексту и у твој врсти текстова. Тако, на основу дијаграма видимо да су то, у овом смислу, двије именице које у потпуности преносе значење енглеске ријечи. Оваква врста анализе помаже приликом

израде алата за рачунарско (машинско) превођење, јер на овај начин рачунар на основу статистичке вјероватноће може да понуди превод који је адекватнији за одређени текст. Наравно, неопходно је прије тога да лингвисти проуче велики број текстова и да нађу одређене законитости у смислу тога како се неки термин најчешће преводи.

2.1.2 Учесталост ријечи

Једна од најчешћих анализа која се ради на корпусу јесте анализа учесталости ријечи у посматраном корпусу. На основу такве анализе можемо да састављамо листе ријечи у односу на то колико често се оне појављују у укљученим текстовима. Због тога, већина листа учесталости ријечи ће бити организоване у двије колоне гдје ће се у првој налазити ријеч(и), а у другој њихов број понављања. Листа учесталости ријечи не мора само да укључује појединачне ријечи вер може да обухвати и тзв. н-граме, па ако је $n=2$ онда говоримо о биграмима (паровима ријечи), или за $n=3$ говоримо о триграмима (триплетима ријечи). Тако нпр. приликом анализе Закона о љековима и Закон о ограничењу коришћења дуванских производа и њихових еквивалената на енглеском језику добили смо следећу комбиновану листу учесталости:

Табела 5. Приказ листе учесталости на основу резултата анализе поглавља о здравству из Упитника; 282 ријечи, минимална учесталост 100

Ријеч	Учесталост
the	6971
и	6550

of	5906
u	5466

and	3203
за	3077
in	2597
je	2412
то	2288
на	2251
се	2056
а	2025
о	1564
од	1492
for	1487
да	1264
on	1179
са	1172
health	1050
су	1010
Is	1001
1	944
2	774
који	753
with	736
болести	689
By	638
are	637
као	620
које	620
that	601
3	592
0	589

as	564
или	553
Из	541
5	540
4	535
здравља	528
law	517
9	500
године	494
montenegro	478
здравствене	472
потрошача	457
7	451
заштите	444
this	442
6	440
8	439
горе	439
гори	422
црној	418
which	405
црне	403
џ	398
from	390
be	387
2008	384
or	381
производа	380
other	375

закона	370
protection	364
article	361
која	350
по	346
product	335
до	324
public	314
2007	312
лица	302
здравље	300
care	291
consumer	288
disease	286
тијела	282
at	281
не	278
система	275
години	274
production	272
not	267
активности	261
што	258
код	257
an	250
measures	247
овог	245
закон	245
ће	244

број	243
складу	242
control	240
области	238
it	237
година	234
дјелова	229
према	228
становништва	225
10	224
activities	223
2009	221
заштити	218
законом	216
заштиту	215
2006	210
all	209
начин	202
није	202
маркет	199
ако	198
tobacco	198
услуга	198
мјере	197
члана	194
population	193
заразних	193
министарство	183
права	183

систем	183
живота	182
односно	181
може	178
out	178
services	178
би	175
information	175
within	175
have	173
заштита	173
других	170
will	170
has	167
има	165
list	165
their	164
транспорт	162
здравствених	162
15	160
план	160
paragraph	159
european	158
national	158
надзор	157
prevention	157
2004	156
e	156
с	156

више	155
г	153
све	153
20	152
сл	152
was	151
ministry	148
мјера	148
non	148
well	146
11	145
lagislation	145
основу	145
there	145
infectious	144
период	143
систем	143
smoking	143
accordance	142
caused	142
члан	139
following	139
програм	139
official	138
100	137
бр	137
дјеце	137
којима	137
лијечења	136

through	136
бити	135
јавно	135
како	135
use	135
0	134
conditions	134
производ	134
др	133
international	133
чл	132
врши	132
здравственог	132
18	131
one	131
who	131
еу	130
these	130
been	129
cooperation	129
implementation	129
органа	129
уз	129
здравства	129
поступак	128
general	127
institute	127
рад	127
12	126

хив	126
into	125
provisions	125
2005	124
also	123
било	123
стране	123
због	123
ц	122
циљу	122
number	122
specific	122
agreement	121
другим	121
оквиру	121
deseased	120
ли	120
development	119
ове	119
производи	118
крви	117
ментално	117
per	117
children	116
обољења	116
ради	116
21	115
постоји	115
competent	114

употребе	114
дана	113
during	113
само	113
могу	112
preferred	112
safety	112
13	111
area	111
community	111
дуванских	111
policy	109
услуге	109
14	108
док	108
gazette	108
надзора	108
под	108
спровођење	107
16	106
economic	106
more	106
особа	106

стате	106
22	105
питање	105
без	104
подаци	104
података	104
средства	104
узимање	104
по	103
order	103
право	103
consumers	102
квалитета	102
нивоу	102
против	102
свих	101
17	100
2003	100
ес	100
if	100
people	100
посебно	100
здравственој	100

Један од главних проблема приликом анализе учесталости ријечи јесте одредити шта је то ријеч и како компјутеру који обавља анализу саопштити ту дефиницију. Рачунари најчешће препознају ријеч као низ алфаветских или алфанумеричких знакова који је прекинут „бијелом

празниним“ (размаком, новим редом...). Већина програма дозвољава корпснику да дефинише шта треба радити у случајевима када су двије ријечи повезане цртицом или неким дригим посебним знаком. Листе учесталости су корисне из више разлога. Тако истраживања *Bybee and Scheibman* (1999) и *Jurafski et al.* (2000) доводе у вези учесталост израза са њиховом спремношћу да прођу кроз процес фонолошке редукције. Па су тако, *Bybee and Scheibman* (1999) показали да ће сугласник *o* у *don't* највјероватније прећи у *schwa* у чешћим изразима, као што су *I don't know* (наспрам мање учесталог израза, нпр. *They don't integrate easily*).

На основу листа учесталости лако се може закључити о којем корпусу, тј. одабиру текстова се ради. Ако се при врху листе учесталости налазе ријечи здравље, болести, нега итд. лако се може закључити да се ради о тексту који се бави медицинском тематиком (као што је случај са листом учесталости у табели горе у тексту). Листе учесталости ријечи нам такође говоре и о неким ванјезичким елементима, као што је изражавање субјективног мишљења. Тада се у тим листама учесталости налази велики број ријечи које су повезане са сљедећим параметрима:

- добро или лоше (параметар емоционалности)
- битно или небитно (параметар битности)
- очекивано или неочекивано (параметар очекиваности)
- разумљиво или неразумљиво (параметар разумијевања)
- (не) могућност, (не) неопходност (параметар могућности/неопходности)
- право или лажно (параметар поузданости).

Листе учесталости такође могу бити од помоћи и у неким методолошким питањима. Наиме, приликом прављења књига за одређени

ниво познавања неког страног језика веома је битно знати које ријечи ће да буду заступљене у тиме књигама, тј. који вокабулар ће у њима бити обрађен. Корпусна и рачунарска лингвистика могу помоћи у том процесу. Не основу корпусно-рачунарске анализе, могуће је издвојити оне ријечи које су релевантне за тај регистар и, самим тим, потребне ученицима или студентима. Наравно, један дио тог вокабулара ће бити сасвим логичан, тј. у њему ће се налазити ријечи за које очекујемо да буду ту, али увијек постоји одређени проценат ријечи које су релевантне за неки вокабулар, а да то није очигледно на прво поглед.

2.1.3 Колокације

Колокације су, у смислу корпусне лингвистике, статистичке тенденције ријечи да се јављају заједно. Оне представљају парове лексичких елемената (*лити сузе*) или могу да представљају везу између лексичке и граматичке ријечи и као такве обично се називају колигације (*put on*).

Један од главних методолошких концепата јесте концепт заједничког јављања ријечи. Корпусна лингвистика се својим већим дјелом бави са три различита феномена заједничког јављања ријечи:

- **Колокације:** пробабилистичко јављање синтаксичких форми као што је бољи од, бољи него; или пак говоримо о фиксираним фразама: цакуп пакум, тип топ итд;
- **колигације:** заједничко јављање реченичних форми са граматичким феноменима као што су граматички односи;

- (граматички шаблони или **колоструктуре** (collostructions): заједничко јављање ријечи/лема са морфо-синтаксичким елементима или конструкцијама као што су дитранзитивне конструкције или клефт конструкције;

Колокације представљају учестало јављање ријечи заједно, а такве ријечи се често називају колокати. За њихово формално обиљежавање, обично се користе слова Л и Р (од енглеских ријечи *left* и *right*; лијево и десно). Поред ознака Л и Р могу да се налазе и бројеви који показују које позицију заузима тај колокат у односу на друге колокате који чине колокацију. Тако, нпр. ако анализирамо конкорданцу ријечи *difference*, најчешћи Л1 колокат јесте одређени члан *the*, док је најчешћи Р1 колокат *between*. На овај начин у стању смо да на врло ефикасан и једноставан начин означимо ријечи које се налазе испред или иза циљне ријечи. Илустроваћемо примјерима из корпуса:

1. cases of parenthetic phrases. Note the DIFFERENCE compared with relative clauses that
2. habitant in some European countries. The DIFFERENCE is 1:10 even and more when
3. services, for prudential reasons. 2. The DIFFERENCE in treatment shall not go beyond
4. qualification which usually makes the DIFFERENCE between health systems
5. Billion Be careful to note the DIFFERENCE between the use of billion and
6. there is no statistically significant DIFFERENCE between the genders. It is interesting
7. endnotes follow the original, noting the DIFFERENCE between footnotes and

Поштујући горе наведене стандарде можемо да направимо табелу која, кроз своје колоне, приказује позиције колоката као и њихову учесталост. Погледаћемо једноставан примјер за два придјева *alphabetic* and *alphabetical*⁵.

Табела 6. Приказ распореда и учесталости колоката за ријеч *alphabetic*;

Л1 колокат	Учесталост Л1	Кључна ријеч	Учесталост кључне ријечи	Р1 колокат	Учесталост Р1
the	1	alphabetic	1	writing	1

Табела 7. Приказ распореда и учесталости колоката за ријеч *alphabetical*

Л1 колокат	Учесталост Л1	Кључна ријеч	Учесталост кључне ријечи	Р1 колокат	Учесталост Р1
in	1	alphabetical	1	order	1

2.1.4 Техничке карактеристике корпуса и програма за његову обраду

Наш корпус се састоји од око осам милиона ријечи са понављањем, тј. од око 656.000 ријечи без понављања. Укупно је обухваћено 128 текстова, који су у MS Word doc и Adobe Acrobat Reader pdf формату. Половина од тих докумената су на енглеском језику док је друга половина представља преводе оригиналних текстова. Обрада оволике количине информација би била веома отежана да нијесмо користили рачунарске програме који се баве оваквом обрадом текста. Главни рачунарски програм који смо најчешће користили је ТекстСТАТ-2 верзија 2.8e. Поред њега користили смо

⁵ Ова два придјева смо увели јер се у читавом нашем корпусу јављају само једанпут. Латински назив за овакве ријечи јесте *hapax legomena* или *hapaxes* (једнина – *hapax legomenon*).

Heidelberg Tenka Text 0.1.3.4 и програм под једноставним називом R верзија 2.10.1. Такође смо користили WordSmith 6.0 и AntConc 3.2.4w. Како бисмо до краја испоштовали ауторе програма које смо користили у овом раду у Додатку, на крају овог рада, ставићемо слике и лиценце за коришћење ових програма. У Додатку се налази списак докумената који су ушли у грађу корпуса. Пуна верзија докумената се налази у приложеном ДВД на крају овог рада. За додатне информације консултовали смо:

- *The British National Corpus* - <http://www.natcorp.ox.ac.uk/index.xml>
- *Corpus Concordance Sampler* - <http://www.collins.co.uk/Corpus/CorpusSearch.aspx>
- *The Gateway to Corpus Linguistics* - <http://www.corpus-linguistics.de/>
- *The Oslo Corpus of Bosnian Texts* - <http://www.tekstlab.uio.no/Bosnian/Corpus.html>
- Корпус савременог српског језика - <http://www.korpus.matf.bg.ac.rs/korpus/login.php>

2.1.5 Превођење стручних термина *Acquis Communautaire*-а⁶

Прије него што се нашу пажњу усмјеримо на превођење стручних термина у оквиру ЕУ, неопходно је да дефинишемо шта се подразумева под појмом *Acquis Communautaire*. Под овим појмом се подразумева скуп правних прописа и обавеза које су обавезујуће за све чланице ЕУ. Он се

⁶ За више информација погледати Приручник за преводиоце у издању Министарства вањских послова и европских интеграција, одакле је и дио текста преузет.

непрестано мијења, допуњује у складу са потребама држава чланица. Састоји се од:

- Принципа и политичких циљева уговора;
- Правних норми које су прихваћене у складу са тим уговорима и предметима одлученим у поступцима који су се водили пред судовима;
- Изјава и резолуција усвојених од стране Уније;
- Мјера повезаних са заједничком вањском и безбједносном политиком;
- Мјера повезаних са правосуђем и унутрашњим пословима;
- Међународних споразума које је склопила Унија и оних које су склопила државе чланице у оквиру питања које се тичу Уније.

Све земље које желе да приступе Унији морају да прихвате *Acquis*. Одступања од *Acquis* су дозвољена само у ограниченом број случајева. Приликом интеграције у Унију, земља кандидат мора да укључи *Acquis* у своје национално законодавство и да га примјењује од момента приступања⁷.

Унутар ЕУ обезбијеђена је једнакост и правна ваљаност свих службених језика, а дефинисана су и правила њихове употребе у институцијама ЕУ. Посљедњи члан (члан 314) Уговора о оснивању Европске заједнице утврђује да су све језичке верзије Уговора једнако аутентичне у правном смислу, док члан 21 даје право грађанима да се на службеном језику своје државе писано обраћају институцијама или тијелима Уније и да на истом језику добију одговор. Поред тога, на основу

⁷ http://europa.eu/abc/eurojargon/index_en.htm

члана 290 Уговора, донесен је посебан правни акт (службени број у EURLEX-у 31958P0001⁸), који се сматра својеврсном језичком повельом ЕУ, а којим је прописана употреба службених и радних језика институција ЕУ, као и обавеза састављања правних текстова на службеним језицима и њихово објављује у Службеном листу ЕУ⁹.

Превођење правне тековине ЕУ спада у превођење правних текстова и као такво мора се повиновати низу посебних правила и ограничења. Од општих упутстава за превођење правне тековине Европске уније најприје треба нагласити да се одустало од приступа да се правна сигурност и прецизност може постићи само дословним преводом изворног текста, тј. да процес превођења треба да буде оријентисан према изворном језику.

То ипак не значи да се може преводити сасвим слободно. Као што је већ наведено у низу приручника и упутстава на нивоу Европске уније, писање и превођење правних аката **нормирано је**. Прописана је структура аката Заједнице, која мора бити иста у сваком језику, те синтагме и реченице које изражавају одређене правне радње и које морају бити нормиране у свим језицима. Такође, прописан је и начин писања мјерних и новчаних јединица, датума, времена, те употребе знакова навода, бројева и слова у набрајању одредаба, начин навођења различитих аката и

⁸ EEC Council: Regulation No 1 determining the languages to be used by the European Economic Community, http://eurlex.europa.eu/Result.do?code=31958R0001&Submit=Search&RechType=RECH_clex&submit=Search.

⁹ Језик има изузетно важну улогу у процесу формирања и одржања националног идентитета. Напори на грађењу наднационалног европског идентитета подразумевају да је вишејезичност демократска нужност која гарантује равноправност свих грађана Европске уније. Европска свијест о томе да је Европска унија, не само економска већ и политичка заједница и заједница вриједности и даље је врло снажна. Слобода, једнакост и солидарност основне су вриједности суживота у друштву. Толеранција, правна држава и правда одређују устројство модерне демократије која се доживљава као вриједност за себе, не само као облик државног уређења него и као начин живота. ЕУ је настала из жеље за „продубљивањем солидарности народа, уз поштовање њихове историје, културе и њихових традиција“. Одлуке европских институција утичу на свакодневни живот грађана Уније и да би оне имале легитимитет неопходно је да грађани активно учествују у процесима одлучивања о њима и то без икакве језичке препреке.

слично. Али, колико год ограничавајуће изгледало тако детаљно прописивање свих поменутих аспеката, оно у великој мјери чини посао преводиоца лакшим, јер је број ситуација у којима је преводилац препуштен сам себи у доношењу важних преводилачких одлука сведен готово на минимум. Осим тога, када се говори о превођењу тако великог корпуса текстова као што је *acquis communautaire*, онда је готово најзахтјевнији задатак обезбиједити **уједначеност и досљедну употребу терминологије и структуру** преведених текстова. А то се може постићи уколико се утврди што више правила и припреми што већи број глосара.

Правна терминологија на нивоу Европске уније веома је специфична, што има одраза и на избор преводних еквивалената. Наиме, да би се обезбиједило да се правни термини Европске уније не тумаче једнако као правни термини из домаћег законодавства држава чланица, свјесно су уложени напори да се створе нови термини на нивоу Европске уније, при чему су главни критеријуми да нови термин буде преводив и транспарентан.

Још један изазов у превођењу представља учестала употреба скраћеница. Иако се у документу *Manual of Precedents for Acts Established within the Council of the European Union* јасно наводи да употребу скраћеница треба избјегавати кад год је могуће, осим у посебним случајевима као што су набрајања, табеле, графикони и слично, ипак смо свјedoци да се у пракси скраћенице изузетно често користе. Оно што за нас представља већи проблем од њихове честе употребе јесте чињеница да не постоје правила која би нам помогла да одлучимо како да се са њима носимо у превођењу. Наиме и у енглеском језику, који смо узели као изворни,

појављују се и користе акроними који нијесу настали из пуног назива на енглеском језику¹⁰.

Неопходно је да преводиоци у потпуности разумију текст који преводе, али га они не смију тумачити у правном смислу. Преводилац треба да изрази "оно што је у тексту речено", а не оно што "мисли да би требало бити речено". Ако се јаве нејасноће у изворном тексту, преводиоци морају консултовати текст на неком другом језику Европске уније, по могућности француском, јер је велики дио *acquis communautaire* у оригиналу настао управо на француском језику.

Са аспекта форме, потребно је нагласити да преводилац не смије да мијења основну структуру оригинала: распоред поглавља, чланова, ставова. Чак и када се појаве дуге и неразумљиве реченице, преводилац не би требало да их дијели на више цјелина, нити да спаја краће реченице у једну, јер би се тиме пореметио број реченица, па би позивање на неку од њих у даљем тексту било погрешно. Ипак, ако је нужно, а понекад заиста јесте, може се дужа реченица прекинути и сљедећа смислена цјелина наставити иза тачке и зареза.

Сходно процедури, правници врше правно техничку редактуру преведеног текста и евентуално указују на пропусте у превођењу значајне за даљи рад. Само у оваквој интеракцији и размјени информација може се остварити напредак и олакшати посао превођења правне тековине ЕУ свима који у том процесу учествују.

¹⁰ Уз то, у нашем језику постоји низ скраћеница које су настале од превода пуног назива неког органа или документа, као што су СТО – акроним за Свјетску трговинску организацију, или ОЕБС – за Организацију за европску безбједност и сарадњу. Међутим, постоји још већи број скраћеница које су настале од пуног назива на страном језику и као такве ушле у наш језик, гдје су се одомаћиле.

2.1.6 Превођење стручних термина помоћу рачунара¹¹

Превођење је данас, као и већина дјелатности, под снажним утицајем развоја телеинформатичких технологија. Тиме се мијењају мјесто и улога превођења и преводиоца у савременом свијету. Мијења се и брзина процеса превођења (увођењем машинског превођења), цијена превођења као интелектуалне услуге (по једној ријечи), али и квалитет односно врједновање превода као производа.

Као идеја, превођење помоћу машине датира још од 17. вијека, а постаје реалност тек крајем 20. вијека. Рачунарски програми стварају несавршене преводе, због тога нијесу подесни за књижевне преводе, али јесу за превођење техничких приручника, научне документације, комерцијалних проспеката, пословних писама, медицинских извјештаја и слично.

Термин "машинско превођење" (*Machine Translation – MT*) односи се на компјутеризоване системе одговорне за стварање превода са или без човјекове помоћи. Граница између човјековог превођења уз помоћ машине (*machine-aided human translation – МАНТ*) и машинског превођења уз помоћ човјека (*human-aided machine translation – HАNT*) често је неизвјесна и зато би термин превођење уз помоћ рачунара могао да обухвати обоје.

У најширем смислу, **машинско превођење** (*machine translation, MT*):

- обезбјеђује досљедност у употреби терминологије,
- елиминише досадно и споро отклањање правописних и случајних грешака,
- може се лако прилагодити захтјевима наручиоца различитим степенима довршености текста,

¹¹ За више информација погледати Приручник за преводиоце у издању Министарства вањских послова и европских интеграција, одакле је и дио текста преузет.

веза путем интернета омогућава преводиоцима директан приступ:

- помоћним средствима и алатима за превођење (интернет рјечници),
- текстовима истог или сличног садржаја, појмовницима и лексичким базама података, двојезичним паралелним корпусима,
- брзим прегледом и претраживањем "радне" верзије превода (*draft*), кориснику се омогућава рационалан приступ одговарајућем дијелу преведеног текста, а затим и квалитетна обрада тог дијела текста.

Но, оно што је за нас врло битно јесте такозвано аутоматско и полуаутоматско превођење текста уз помоћ рачунара (*Computer Aided Translation - CAT*) које је постало дио наше свакодневице. На тржишту се нуде програми у распону од једноставних рјечника, који се користе за тражење преводних еквивалената и синонима појединих термина, до програма који анализирају граматичку и синтаксичку структуру текста и аутоматски преводе појмове, фразе и реченице, узимајући у обзир језичке и граматичке различитости језика извора и језика права. На интернету се могу наћи онлине обрасци за превођење текстова и читавих веб страница.

Познатији програми ове врсте су *Word Translator*, *NeuroTran*, *Deja Vu*, *Wordfast* и програм који је постао својеврсни стандард за аутоматско превођење – *TRADOS*. Овај програм, производ једне њемачке фирме, развијен је током протекле деценије и данас нуди заиста много корисних могућности. Он комбинује двије технологије, сегментацију и преводачку меморију, ради на принципу неуралних мрежа и активно претражује интерну базу података, тзв. *Translation Memory* (ТМ – преводачка меморија). ТМ се попуњава током превођења, а програм муњевице претражује своју радну меморију и нуди могуће преводе реченица, фраза и појединих ријечи по обрасцу "поклапања", тј. степена подудараности изворних и циљних термина.

TRADOS се испоручује у облику модула:

Trados Workspace интегрисано је окружење које вам омогућава да покренете разне компоненте Традос пакета и организујете пројекте, као и датотеке унутар пројеката.

Translator's Workbench (обједињава све функције програма и то је оно што већина људи зове *Традос*. Он опслужује преводачку меморију, омогућава вам да анализирате своја документа, да их сегментирате и "сређујете". Ради у комбинацији са осталим *Традос* модулима и обично са *Wordom*.)

MultiTerm (то је нека врста рјечника, тј. претражује поједине дјелове реченице или "сегменте", тј. фразе и ријечи)

TagEditor (дозвољава превођење HTML, XML, ... докумената)

Winalign (омогућава аутоматско генерисање преводачких меморија (ТМ-а) поређењем изворних и већ преведених докумената).

Традос, по могућностима и цијени, спада у тзв. тешку артиљерију. На тржишту се може наћи један мањи програм који је у великој мјери компатибилан са њим, по функционалности веома сличан а вишеструко јефтинији и зато га преводиоци све више користе. То је *Wordfast*.

Wordfast ради из *MS-Word-a*, али и документи из *MS-Excel*, *MS-Access*, *MS-PowerPoint* и *HTML* датотеке такође могу да се преводе. *Wordfast* ТМ формат је отворен – и може да се види или уређује са *Word-om*, *Excel-om*, *Access-om* и другим популарним програмима. Поред тога, *Wordfast* отвара преводачке меморије ТМН типа, као што су оне које стварају *TWB* (*Trados Translator's Workbench*), *Deja Vu*, *Star Transit* итд.

Wordfast је веома моћна алатка спакована у компактан *Word* шаблон (*template*). Ради са *MS-Word-ом* 97, 2000, XP и 2003 на РС платформи, а на

Mac-у са MS-Word-ом 98, 2001 и 2004. До 20 корисника може да дијели исту преводачку меморију преко локалне рачунарске мреже.

Министарство за европске интеграције Владе Црне Горе, које је носилац координације превођења у вези са ЕУ, у посљедње вријеме даје предност *Wordfast-у* као преводачки алат који задовољава потребе спољних и интерних преводаца. Због повољне цијене набавке и могућности брзе обуке, постаје незаобилазан дио како индивидуалних превођења тако и превођења за која је потребно користити умрежене рачунаре и заједничку преводачку меморију.

2.1.7 Машинско превођење стручних термина¹²

Машинско превођење ограничено је неалгоритмичком природом теорије превођења, па сва ограничења изражена још педесетих година прошлог вијека, остају да важе и данас. Ипак, напредак статистике, информатике и вјештачке интелигенције, с једне стране и лингвистичких теорија, рачунарске и корпусне лингвистике, препознавања људског говора итд. с друге стране, обезбјеђују машинском превођењу све већи, иако ограничен успјех у врло строго ограниченим условима (нови *тагери* и *парсери*, програми за рјешавање лексичке вишезначности, лексичка семантика, лексичке базе података и појмовници, базе знања, уношење прагмалингвистичких и текстуалних информација у програме, преводачка меморија итд.) Због тога се можемо надати да преводачки програми неће представљати само помоћни алат људском превођењу (нпр. *SAT/MT*) или брзе поступке који претходе људском превођењу, већ да ће у строго ограниченим, унапријед познатим или припремљеним лексичким, морфосинтаксичким, прагмалингвистичким и дискурским

¹² Ибид.

условима (нпр. институционални текстови, упутства за употребу машина и уређаја, безбједносне комуникације у поморском и ваздушном саобраћају, апстракти научних чланака) бити у стању да направе превод који ће захтијевати минимум човјекове интервенције. Истраживања на плану машинског превођења стално ће доприносити бољем разумијевању процеса људског превођења и јаснијем лингвистичком опису текста, врједновању превода и подучавању превођења. Тако ће постепено нестајати или пак слабити неке од предрасуда о машинском превођењу.

3 Корпусна лингвистика; метод или наука?

3.1 Дефинисање корпусне лингвистике

Постоје различита схватања о томе да ли је корпусна лингвистика наука за себе или је „оруђе“ тј. методологија рада која се користи приликом лингвистичких истраживања. У овом одјелку покушаћемо да одговоримо на ово.

Стабс (2006), у свом прегледу скреће пажњу на нејасноће аналитичара корпуса приликом описивања својих оперативних метода у оквиру научног контекста. Недостатак јасног приступа у расправама о методолошком оквиру који се користи је, вјероватно, ствар која највише изненађује с обзиром на то да се корпусна лингвистика позиционирала у научни оквир, и тежи да искаже своју научну природу. Стога је, како сматрамо, неопходно да одговоримо на питање, „Шта је корпусна лингвистика?“ – да ли је то дисциплина, методологија или нешто треће? Наравно, у потпуности смо свјесни да не можемо дати дефинитан одговор, али можемо да читаоцу представимо различите погледе на корпусну лингвистику кроз литературу која се бави овим питањем. Узећемо специфични термин – корпусна лингвистика – и погледати како је он дефинисан како експлицитно тако и имплицитно у различитим релевантним изворима. Надамо се да ће овај кратки осврт доприњети бројним расправама међу корпусним лингвистима о томе шта је или шта би требало да је корпусна лингвистика. У смислу овог, свакој особи која је почела да се бави корпусном лингвистиком одмах пада у очи да постоји велики број описа и дефиниција. Аартс, један од оснивача корпусне лингвистике, као да је очекивао сваки. Књига два аутора Артса и Мејјса (1984), се често дефинише као извор термина корпусна лингвистика, иако

се тај термин и раније јављао, нпр. у књизи, Аартс анд ван ден Хеувел (1982). У овом раду Аартс нам саопштава да је овај термин скован уз оклијевање:

“because we thought (and I still think) that it was not a very good name: it is an odd discipline that is called by the name of its major research tool and data source. Perhaps the term has outlived its usefulness by now”.

Због ове изјаве неки од истраживача су тражили алтернативне дефиниције и називе. У смислу тога шта корпусна лингвистика „јесте“ не само да су различите дефиниције биле нуђене, већи су и алтернативе биле директно помињане или одбациване. Па тако говоримо о томе да је корпусна лингвистика:

- **оруђе,**
- **метод,**
- **методолошки приступ,**
- **дисциплина,**
- **теорија,**
- **теоријски приступ,**
- **парадигма (теоретска или методолошка)**
- **или пак комбинација неког од ових елемената.**

Лич (1992), тврди да:

“computer corpus linguistics defines not just a newly emerging methodology for studying language, but a new research enterprise, and in fact a new philosophical approach to the subject”

и даље наставља са својим описом карактеристика рачунарске корпусне лингвистике као нове парадигме (Лич 1992: 106). По сличној логици, Стабс 1993, не прихвата ограничену дефиницију корпусне лингвистике само као методологије, и, коментаришући Синклера 1991, примјећује да:

“in this vision of the subject, a corpus is not merely a tool of linguistic analysis but an important concept in linguistic theory” (1993: 23–24).

Теуберт (2005) такође наглашава теоретски концептуализацију и описује корпусну лингвистику као:

“a theoretical approach to the study of language” (2005: 2).

Ову идеју корпусне лингвистике као парадигме преузео је и Гриес, али он преферира методолошку концептуализацију и наводи да:

“over the past few decades, corpus linguistics has become a major methodological paradigm in applied and theoretical linguistics.” (2006a: 191).

Тоњини-Бонели 2001. Године описали су корпусну лингвистику као: *“pre-application methodology”* која има: *“theoretical status”* (2001: 1). Даље, Махлберг описује корпусну лингвистику као: *“an approach to the description of English with its own theoretical framework”* (2005: 2), а да би нагласила свој приступ, користи термин: *“corpus theoretical approach”* (2005, 2006). У даљем писању на ову тему ауторка види разлику која проистиче из различитих перцепција истраживача:

“there is still disagreement on whether corpus linguistics is mainly a methodology or needs its own theoretical framework. Advocates of corpus-driven approaches to the description of English claim that new descriptive tools are needed to account for the situation of real text, and ideas of theoretical frameworks to accommodate such tools have started to emerge.” (2006: 370).

Томпсон и Хунстон (2006) наводе да:

“a its most basic corpus linguistics is a methodology that can be aligned to any theoretical approach to language” (2006: 8).

МекЕнеру, Ђиао и Тоно, примјећују да је:

“corpus linguistics is a whole system of methods and principles of how to apply corpora in language studies and What is corpus linguistics? What the data says teaching/learning, it certainly has a theoretical status. Yet theoretical status is not theory in itself” (7–8: 2006);

и стога, можемо закључити да они корпусну лингвистику дефинишу као методологију. Корпусну лингвистику као методологију такође дефинишу и МекЕнеру и Вилсон (1996), Мејер (2002). Баукер и Пеарсон (2002: 9) кажу за корпусну лингвистику да је: *“an approach or a methodology for studying language use”*. МекЕнеру и Габриелотос такође додају да:

“corpus linguistics may be viewed as a methodology, but the methodological practices adopted by corpus linguists are not uniform” (2006: 44).

Теуберт (2005) такође коментарише ове различите методе и каже да:

“corpus linguistics is not in itself a method: many different methods are used in processing and analysing corpus data. It is rather an insistence on working only with real language data taken from the discourse in a principled way and compiled into a corpus” (2005: 4).

Аартс (2002), Теуберт (2005), и Вилијамс (2006), између осталих, описују корпусну лингвистику као дисциплину. Међутим, ово нам намеће још једно питање, а то је, ако је корпусна лингвистика дисциплина – каква је то врста дисциплине. Стабс (1993: 3) описује корпусну лингвистику као: *“applied social science”*, Теуберт наводи да:

“linguistics is not a science like the natural sciences whose remit is the search for ‘truth’. It belongs to the humanities, and as such it is a part of the endeavour to make sense of the human condition.” (2005: 7).

МеКарти карактерише корпусну лингвистику као: *“cutting edge change in terms of scientific techniques and methods” (2001: 125)*, док Стабс експлицитно повезује корпусну лингвистику и науку уз коментар да:

“geologists are interested in processes which are not directly observable because they take place over vast periods of time [...] Corpus linguists are interested in processes which are not directly observable because they are instantiated across the language use of many different speakers and writers” (2001: 243).

Међутим, вјероватно нити један осврт на било које актуелно питање из поља лингвистике не би био потпун а да не поменемо мишљење уваженог ауторитета ове науке Ноама Чомског:

“my judgment, if you like, is that we learn more about language by following the standard method of the sciences. The standard method of the sciences is not to accumulate huge masses of unanalyzed data and to try to draw some generalization from them” (2004: 97).

Међутим, и корпусни лингвисти често коментаришу, па и критикују лингвистички приступ Н. Чомског. Картер тврди да поменути приступ Н. Чомског не показује:

“interest in language beyond the level of the sentence, there is no recognition that authentic data is of any significance and there is no acceptance that studies of large corpora or real language in use play any part in descriptive theories of language. Most significantly, too, there is a clear sense that the analysis of meaning is not a primary purpose.” (2004: 2).

Синцлер такође критикује интроспективну лингвистику ријечима: *“one does not study all of botany by making artificial flowers”* (1991: 6). Теуберт и Криснамурту (2007) описују корпусну лингвистику као *parole-linguistics* наспрам *langue-linguistics* Сосира и Чомског. Такође, неки од истраживача напомињу да различите дефиниције и описи корпусне лингвистике не морају да буду лоши. Различите дефиниције су очекиване из разлога што се корпусна лингвистика непрестано развија и да се истраживања спроводе на различите начине покривајући тако разна питања. О овоме се детаљније може наћи код Хоеја (1993: ви). Поред тога, због природе корпусне лингвистике корпусни лингвиста може бити аналитичар или креатор тј. састављач корпуса, или и једно и друго. Због тога су различита тумачења добродоша, па Теуберт иде у том правцу:

“only if the discourse of corpus linguistics remains controversial and pluralist will there be progress” (2005: 13).

Међутим, да би се тај напредак десио, морамо расправљати о различитим тумачењима, како се она преклапају, или спуротстављају. Међутим, веома је битно направити разлику између различитих гледишта на нивоу академске заједнице и нејасноће дефинисања корпусне лингвистике.

3.1.1 Корпусни приступ наспрам интуитивног

Под интуитивним приступом подразумевамо могућност истраживача неке проблематике да измисли примјере на основу којих би могао да анализира одређену језичку појаву. На тај начин истраживач може да измисли „чистије“ примјере који би недвосмислено ишли у прилог или оповргавали неку претпоставку. Међутим, овакав приступ се мора користити са одређеном дозом опреза. Оно што се одмах да примијетити

јесте да ће изворни говорници различитих подручја, дијалеката или социолеката произвести различите примјере који могу бити неприхватљиви другим изворним говорницима. Друга мана оваквог приступа јесте да говорник који жели да докаже или оповргне неку тезу, обично може, без икаквих проблема, да произведе примјер који иде у једном или другом правцу, без обзира на то да ли се он јавља или би се јавио у природном језичком окружењу. На тај начин можемо да произведемо следеће реченице:

Књигу сам купио ја.

Књигу сам прочитао ја. Итд.

На основу анализе овог низа реченица можемо устврдити да је у нашем језику на почетном мјесту у реченици најчешће објекат, а не субјекат. Наведене реченице су граматички потпуно исправне, али би интуитивни одзив већине изворних говорника на овакву врсту реченица био да оне свакако нијесту типичне реченице произведене у природном, стандардном, устаљеном језичком контексту. Сматрамо да би ове реченице биле „природније“ ако би биле постављене као:

1. Купио сам књигу. (или Ја сам купио књигу.)
2. Прочитао сам књигу (или Ја сам прочитао књигу.)

На основу оваквих једноставних примјера можемо да видимо колико је значајно да изаберемо „неутралне“ примјере, тј. корпус текстова који је произведен у лингвистичком контексту на који аутор не може да утиче, како би се на најбољи могући начин искористио сав потенцијал корпусног приступа, тј. како би се добили непристрасни и научно утемељени резултати.

3.2 Испитивање помоћу WebCorp-a

Да бисмо проширили нашу базу истраживања и дефинисали корпусну лингвистику на „глобалном нивоу“, искористићемо *WebCorp* да бисмо прикупили примјере са интернета тако што ћемо тражити лексичку секвенцу *corpus+linguistics+is*. Под глобалним нивоом подразумијевамо могућност *WebCorp*-а да нам пружи увид у све, или скоро све, текстове који се налазе на интернету, у циљу контрастирања или спровођења истраживања на корпусу. Ми ћемо у овом дијелу искористити *WebCorp* како бисмо пронашли текстове и примјере у којима се дефинише корпусна лингвистика. Зашто? Из разлога што на овај начин можемо видјети шта аутори, који се баве корпусном лингвистиком, мисле о том пољу лингвистике. Када би овај процес спроводили „пјешке“ тј. ишчитавајући све текстове који су нам физички доступни у библиотекама, било би потребни мјесеци да се ова врсте анализе обави. Међутим, уз помоћ *WebCorp*-а цјелокупну ову анализу можемо обавити знатно брже јер су нам сви примјери на дохват руке из цијелог свијета. Одатле израз глобални ниво. Испод овог текста приказаћемо конкорданцу која ће бити у складу са разноврсним интерпретацијама које смо поменули у претходном дијелу:

1. Linguistics Software, What is Corpus Linguistics [Corpus linguistics is](#) a study of language and a method of linguistic
2. of natural or “real word” texts known as corpus. [Corpus linguistics is](#) used to analyse and research a number of
3. by computerized processing of natural language. [Corpus linguistics is](#), however, not the same as obtaining language
4. language data through the use of computers. [Corpus linguistics is](#) the study and analysis of data obtained from a

5. free encyclopedia Jump to: navigation, search [Corpus linguistics is](#) the study of language as expressed in samples
6. | ISBN-10: 0521547369 | ISBN-13: 978-0521547369 [Corpus linguistics is](#) the study of language data on a large scale -
7. terms used in the text. Show more Show less [Corpus linguistics is](#) the study of language data on a large scale -
8. University of Gothenburg Book Description [Corpus linguistics is](#) the study of language data on a large scale -
9. k.houghton@liv.ac.uk What is Corpus Linguistics? [Corpus linguistics is](#) the study of language based on evidence from
10. Based on the above definition of a corpus, [Corpus linguistics is](#) the study of language by means of naturally
11. Teaching Instructor: Susan Conrad Credits: Four [Corpus linguistics is](#) a method of analyzing language that uses large
12. Methodologically speaking, this implies that [corpus linguistics is](#) an important tool for work within the cognitive-f
13. Table of Contents Preview About | Readership [Corpus linguistics is](#) a research approach to investigate the patterns

(извор WebCorp)¹³

Дефиниције варирају, па тако можемо да нађемо да је корпусна лингвистика: *the study of language, tool, analysis, etc.* Из конкорданце можемо да видимо да се термин *study of language* веома често користи, у комбинацији са описима корпусне лингвистике као науке и дефиниције корпусне лингвистике као методологије.

¹³ <http://www.webcorp.org.uk/>. Webcorp је на овој интернет страници описан као “*a suite of tools which allows access to the World Wide Web as a corpus*”. Ови подци су добијени 19.03.2013. године, али је неопходно нагласити да се број текстова, чланака који су доступни овом алату на интернету непрестно мијења, тако да је могуће да се јаве одређена одступања од онога што је аутор овог текста нашао. Међутим, таква одступања не могу у значајнијој мјери да утичу на тумачење резултата које смо дали у тексту. Одступања се, прије свега, могу односити на број укључених чланака и њихову врсту.

3.3 Корпус истраживачких чланака из лингвистике

Даље дефинисање корпусне лингвистике заокружићемо са анализом корпуса који је састављен од 157 лингвистичких чланака, дакле од толиког броја радова лингвиста који се баве корпусном лингвистиком. И слjedeћим примјерима видјећемо како је дефинисана корпусна лингвистика:

1. Teaching Instructor: Susan Conrad Credits: Four [Corpus linguistics is a method](#) of analyzing language that uses large collections
2. and Letters and a linguistics professor at SDSU. [Corpus linguistics is a method](#) for the study of language use, analyzing a vast
3. of language (in this case, research papers), and **corpus linguistics is a method** to extract general patterns of language use from
4. not optimal, for multiple reasons. For example, [corpus linguistics is a methodology](#) (not a set of techniques), based on the premise
5. discuss the **methodological** pros and cons of *corpus linguistics* as opposed to traditional
6. It is in these cases that we need to use **the methodology of corpus linguistics**.
7. der)Graduate Research Projects Corpus Linguistics [Corpus linguistics is a methodology](#) in linguistics that involves computer-based
8. linguistics, I have been arguing that, to me, [corpus linguistics is a methodology](#), or a methodological commitment, in linguistics
9. explanation concerns both **the methodology** and ideology of *corpus linguistics*.
10. posture verbs, are best substantiated through **the methodology of corpus linguistics**.
11. em Quirk et al. (1985), seções 6.12-16. 7 "[Corpus linguistics is a methodology](#) that may be used in almost any area of linguistic.

12. Only when **methods of corpus linguistics** are applied in areas of linguistic research can¹⁴

У већини пронађених примјера јасно се истицала ријеч метод(ологија), што јасно указује на то да већина истраживача поља корпусне лингвистике вјерује да је она метод(ологија) и на тај начин се односи према њој. Кроз ове примјере можемо да видимо да се корпусна лингвистика третира као помоћно средство приликом лингвистичких анализа текста које доприноси обимнијем и квалитетнијем истраживању. Обимнијем, јер данашњи корпуси садрже знатно већи број ријечи и квалитетније јер се повећава узорак на којем се ради анализа и, самим тим, и релевантност добијених података.

Употреба препозиције *in* такође показује да неки од аутора корпусну лингвистику виде као поље које је интересантно за академску заједницу:

1. or anyone else with an abiding interest [in corpus linguistics](#) and its applications. 2,592 members Join Corpora
2. one to sell? Sell yours here Share Explorations [In Corpus Linguistics](#).(Language and Computers 23) (Language &
3. 9042007516 | ISBN-13: 978-9042007512 Explorations [in Corpus Linguistics](#) contains selected papers from the eighteenth
4. give a broad overview of the latest activities [in corpus linguistics](#). Issues associated with the creation of corpora
5. within corpora. Show more Show less Explorations [in Corpus Linguistics](#) contains selected papers from the eighteenth
6. give a broad overview of the latest activities [in corpus linguistics](#). Issues associated with the creation of corpora

¹⁴ Ibid.

7. 9042007516 book_display_on_website Explorations [In Corpus Linguistics](#).(Language and Computers 23) (Language &
8. is intended to cover a wide range of issues [in Corpus Linguistics](#), of either theoretical or practical significance.
9. of data organization in corpora - Standards [in corpus linguistics](#) - Tools for corpora development and use -
10. and is a must read for anyone interested [in corpus linguistics](#) and its impact on applied linguistics. What
11. onference. The papers address the "states of art" [in corpus linguistics](#) in North America by focusing on work that is
12. detailed, analyses of recent methods and theory [in Corpus Linguistics](#), ways of analysing corpora, and recent applicatio
13. The book represents the best of current practice [in Corpus Linguistics](#), and as a one volume reference will be invaluable
14. detailed, analyses of recent methods and theory [in Corpus Linguistics](#), ways of analysing corpora, and recent applicatio
15. The book represents the best of current practice [in Corpus Linguistics](#), and as a one volume reference will be invaluable
16. depends on using corpora. Recent developments [in corpus linguistics](#) have affected corpus-based studies in translation
17. performance. In J. Svartvik (Ed.), Directions [in corpus linguistics](#) (pp. 105-122). Berlin: Mouton de Gruyter. Leech,
18. Introduction. In J. Svartvik (Ed.), Directions [in Corpus Linguistics](#). Berlin: Mouton de Gruyter Teubert, W. and
19. discourse. 3. Discourse and Discourse Analysis [in Corpus Linguistics](#) In her criticism of sociolinguistics, Hasan
20. Below I introduce the view of discourse [in corpus linguistics](#) which, being a strongly data driven approach,
21. Foucauldian sense. 3.1 The Term "Discourse" [in Corpus Linguistics](#) Corpus research started out as a methodological

22. collection and detailed documentation of texts [in corpus linguistics](#)
allows gaining empirical evidence for assertions

Препозиција *in* по својој дефиницији се користи како би објаснила да је нешто садржано у нечем другом, тј. да се налази и оквирима или границама нечег другог. На основу горе наведених примјера, јасно се да закључити да се корпусна лингвистика налази у оквиру опште лингвистике. На овај начин смо, коришћењем цјелокупног интернета као корпуса, тј. извора информација дошли до поузданих резултата. Корист од овакве врсте добијања резултата може бити вишеструка јер оваква анализа у оквиру себе садржи велики број текстова, са великим бројем ријечи и, што је методолошки веома битно, сваки резултат се може провјерити како би се утврдила његова ваљаност.

4 Контрастивна анализа корпуса (2. дио)

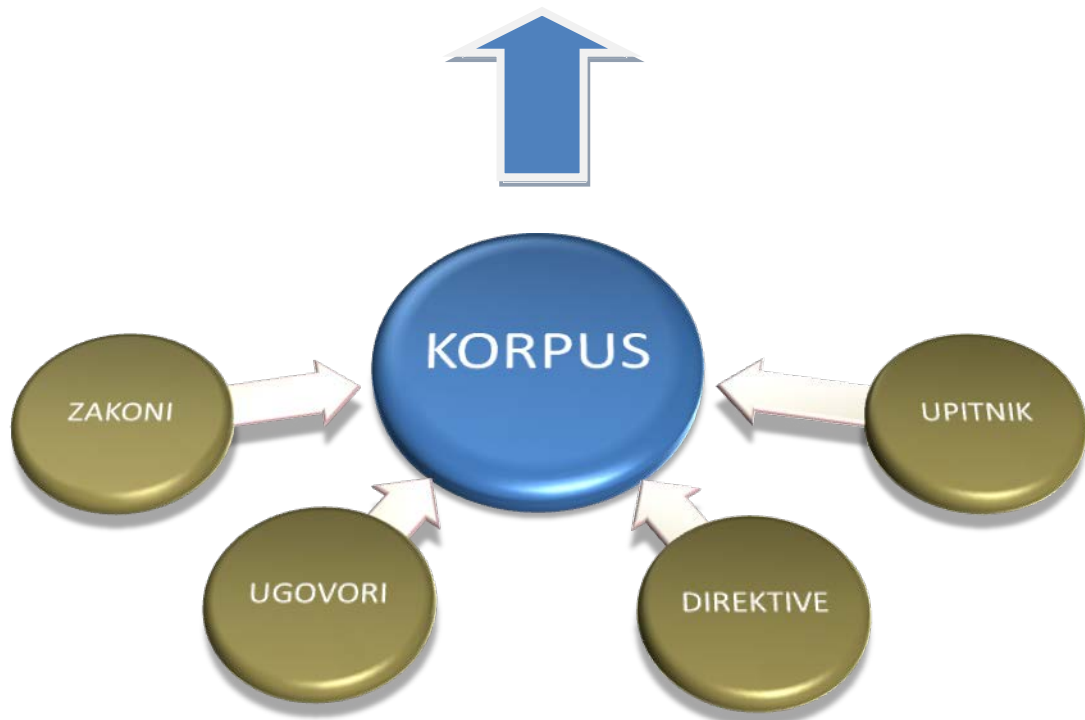
4.1 Увод

Ова анализа настала је као производ трогодишњег рада на корпусу са којим смо дошли у контакт током обављања послова превођења за различите институције које су одговорне за процес приступа Европској унији.

Наш корпус се састоји од око осам милиона ријечи са понављањем, тј. од око 656.000 ријечи без понављања. Укупно је обухваћено 128 текстова, који су у *MS Word doc*® и *Adobe Acrobat Reader pdf*® формату. Сви текстови који су укључени у анализу се налазе у Прилогу 4, гдје су побројани документи и њихови називи онако како их рачунар види. Ради лакшег приступа овим текстовима, сви они се налазе на CD-у који је приложен уз овај докторат. Половина од тих докумената се односи на поглавља Упитника, док је друга половина везана за законе који су послужили као прилог уз Упитник. Такође, половина тих докумената је на енглеском језику док је друга половина представља преводе оригиналних текстова. Обрада оволике количине информација би била веома отежана да нијесмо користили рачунарске програме који се баве оваквом обрадом текста.

Захваљујући професионалним ангажманима на превођењу, били смо у стању да прикупимо велику збирку текстова из различитих области како бисмо могли да задовољимо критеријум репрезентативности корпуса.

ANALIZA KORPUSA



Дијаграм 2. Графички приказ састава нашег корпуса

4.2 Резултати истраживања корпуса (карактеристике нашег корпуса)

4.2.1 Административни стил

Наш цјелокупни корпус је састављен од текстова који припадају административном стилу писања. Административни стил се обично дијели на пет подстилова: (Тошовић 1988):

Законодавно-правни, чији су жанрови закони, статутуи, уставуи, одлуке, наредбе, рјешења;

Друштвено-политички - резолуције повеље, декларације, програми, реферати, саопштења, изјаве;

Дипломатски - преписке, ноте, демарши, протоколи, меморандуми;

Пословни – уговори, дописи, фактуре, сертификати, наруџбенице, уплатнице, рачуни, спецификације;

Лични – молбе, жалбе, аутобиографије, пуномоћи, лична документа, упитници, анкете, формулари, итд.

Наш корпус је састављен од, прије свега, прве три поменуте категорије. На основу поменутих карактеристика, можемо без устезања закључити да и наш корпус одговара административном стилу. Неопходно је истаћи да ни један текст не припада чистом стилу, већ је мјешавина различитих регистара, од којих један може имати превагу.

Оно што се одмах дало уочити након иницијалних анализа овог корпуса јесте да је велика већина текстова састављена од формалних језичких структура што смо и очекивали, с обзиром на то да овај корпус представља званични однос Црне Горе према институцијама ЕУ. Због ових елемената, можемо констатовати готово потпуно одсуство апстрактних именица или других врста ријечи јер су сви документи у оквиру корпуса фокусирани на конкретне проблематике.

4.2.2 Имперсонализовани регистар

Оваква врста регистра, која се обично подводи под административни стил, показује велики степен шематизованости, тј. корписте се разноврсне утврђене фразе како би се у што краћем времену и што јасније пренијеле идеје које неки документ носи. Најчешће се овакви административни

текстови користе као вид комуникације између држава или институција. На граматичком плану овај стил је номинални, што значи да у њему доминирају именице, а процес номинализације је веома чест, тако да се тиме број глагола додатно смањује. Постојећи глаголи су обично у форми садашњег или будућег времена, у трећем лицу или безличним, пасивним конструкцијама. Ову констатацију можемо допунити и констатацијом да су личне замјенице попут ЈА или *I* су готово непостојеће. Такође смо открили да ријечи за које интуитивно сматрамо да би требало да су најчешће у појединим законима или поглављима Упитника то нијесу.

Као што смо већ поменули, једна од карактеристика административног стила јесте и имперсонализација, тј. увијек говоримо о државама, институцијама, министарствима, дакле, увијек о ЊИМА. Овакви документи који су по правилу веома формални, дају занимљиве резултате ако их тестирамо на личне замјенице, и контрастирамо између два језика. Тако смо нашли да лична замјеница ОНА (никако или веома ријетко у зависности од текста) се не односи на именовање припадница женског рода (+ жива), већ је готово искључиво повезана другим именицама женског рода (- жива). У енглеском језику за разлику од српског доминира лична замјеница *IT*. Приказаћемо добијене примјере:

1. поднесе Скупштини на потврђивање чим ОНА буде у могућности да се састане
2. амбалажа укључена заједно с производом, ОНА ће бити укључена и за потребе
3. активних мјера и/или законских поступака, ОНА ће о томе одмах обавијестити тијело
4. осим у случајевима од посебне хитности, ОНА ће Привременом одбору доставити

5. аката потребно обавити јавну расправу, ОНА ће утврдити нацрт закона, програм
6. може дати саопштење једне државе којом ОНА даје објашњења или оправдања за
7. начела међународних организација или су их ОНА и донијела (нпр. УН и ЕУ)
8. ван територијалних вода, под условом да ОНА има искључива права на обраду тог
9. могуће доказати: (а) да је враћена роба ОНА иста која је била извезена; и (б) да
10. оптерећења ванболничке здравствене службе, ОНА је током 2006. године у далеко највећем
11. што се тиче ванболничке здравствене службе, ОНА је у току 2006. године у далеко највеће
12. оних из тар. броја 9402), укључујући ОНА која се могу претворити у лежајеве
13. се може одложити с образложењем да би ОНА могла бити ушпитање у истрагу, судско
14. неповољном економском утицају кршења, ОНА може поднијети писмени захтјев
15. гнојива с условима из става 1 овог члана, ОНА може предузети одговарајуће мјере након
16. испуни неку обавезу из овог Споразума, ОНА може предузети одговарајуће мјере
17. 3(1) Протокола или према СТО Споразуму, ОНА не може покренути поступак за рјешавање
18. законодавство и матични одбор) и Влади, ако ОНА није подносилац предлога. Предлог закон
19. јених материјала без поријекла или, ако ОНА није позната и не може се утврдити

20. организација Црне Горе и једино је ОНА одговорна за монетарну политику, успостављање
21. може попунити кад се производи на које се ОНА односи извозе, или након њиховог извоз
22. издати након извоза производа на које се ОНА односи, ако: (а) увјерење није издато
23. индивидуалним субјектима да штите и ОНА права која за њих проистичу из директива
24. судску опомену или ослобођење од казне, ОНА се извршава на слободи и не може трајати

Личну замјеницу ОН нашли смо још мање пута из разлога што су именице које се чешће замјењују женског рода (држава, скупштина, влада, служба), што личну замјеницу мушког рода ОН ставља по страни:

- a. нуди заиста много корисних могућности. ОН комбинује двије технологије, сегментацију
- b. то је оно што већина људи зове Традос. ОН опслужује преводачку меморију, омогућава
- c. примјену и спровођење овог Споразума. ОН ће се састајати на одговарајућем нивоу

Прва два пута говоримо о рачунарском програму, а у трећем примјеру очигледно о споразуму.

Као што смо већ рекли у енглеском језику доминирала је замјеница *IT*, а не директни преводни еквиваленти за ОН и ОНА – *HE* и *SHE*. Зашто је то тако, одговор треба потражити у томе што је у енглеском језику другачија

перцепција рода и правила налажу да се ријечи које немају природан род дефинишу са *IT*.¹⁵ Побројаћемо само неке примјере:

1. *liability companies*. In that purpose, *IT* is necessary to initiate amendments of
2. *ect to separate annual programmes*. When *IT* comes to motivation, special attention
3. *he company has a positive attitude when IT comes to further tertiary education of*
4. *be a simply administrative service as IT was the case with previous personnel de*
5. *ed the training, i.e.1.5% in 2008*. When *IT* comes to training of active population
6. *ed database are the Statistical Office (IT uses a large quantity of data by other*
7. *residents and non-residents equal when IT comes to tax and contribution payments.*
8. *lling within 2004-2008 period*. In 2008, *IT* was 21% while it has amounted to 20.50%
9. *-2008 period. In 2008, it was 21% while IT has amounted to 20.50% since 1 January,*
10. *ocal Self-Government define areas where IT is possible to carry out decentralisati*

¹⁵ Осим у случајевима, нпр. емоционалне привржености, што у административном стилу свакако није случај.

11. of fiscal policy, on the basis of which IT plans the main categories of receipts a
12. nding units, and if there is a deficit, IT shall determine sources of funds for it
13. aw on the State Budget and shall submit IT to the Government in October. The compe
14. ision on municipality budget and submit IT to the Ministry of Finance for insight

4.2.3 Декомпоновани предикати

Једна од карактеристика административног стила су и декомпоновани предикати. Они могу бити састављени од семикопулативних глагола као што су: вршити (вршити надзор), дати (дати сагласност) или водити (водити преговоре), а могу бити и замјењени пунозначним глаголом који је обично онај из ког је изведена глаголска именица у лексичком дијелу предиката (добити пристанак - пристати).

4.2.3.1 Донијети одлуку

Чести су и декомпоновани предикати као што су:

- Донијети одлуку (умјесто одлучити),

Примјери:

Табела 7. Примјери декомпонованог предиката донијети одлуку које смо нашли у нашем корпусу

учеснику - правном лицу када се <u>донесе одлука</u> о стечају или ликвидацији Дакле, чим се
ју или ликвидацији Дакле, чим се <u>донесе одлука</u> о стечају и/или ликвидацији, испуњени
.2008. године, предвидјено да се <u>донесе одлука</u> о избору модела за елиминисање баријер
ИИ кварталу 2008. године, биће <u>донешена одлука</u> о именовању Стратешког координатора за
ИИ кварталу 2008. године биће <u>донешена одлука</u> о именовању Координатора за прекограни
степену у поступку у којем је <u>донешена одлука</u> којом је повриједјено људско право и о
раничења. У том циљу може се <u>донијети одлука</u> о промјени износа изражених у еурима. Г
да укључе грађане у процес <u>доношења одлука</u> и у организацију јавних расправа. Понек
локалне самоуправе и система <u>доношења одлука</u> на нивоу општине ће захтијевати модерни
т, тежи побољшавању процеса <u>доношења одлука</u> у скупштинама општина, јачање вјештина
утицати и на најбоље праксе <u>доношења одлука</u> у скупштинама општина. На овај начин, р

грађана у локалним процесима доношења одлука је ваома ниско у Црној Гори. Међутим, н

У текстовима на енглеском језику имамо другачију ситуацију. Наиме, инфинитив глагола *decide* је чешћи од одговарајућих најчешћих колокација *make decision* и *reach decision*. У овим случајевима, глаголу *decide* најчешће претходе модални глаголи и то:

$V_{\text{мод}} + \text{bare Inf.}$

- *Shall decide* – 91 понављање
- *May decide* – 90 понављања
- *Can decide* – 51 понављање
- *Will decide* – 30 понављања

4.2.3.2 Извршити преглед

- Извршити преглед (умјесто прегледати)¹⁶,

Ријеч преглед се најчешће налазила у документима који за своју тематику имају јавно здравље и заштиту потрошача (нпр. у Закону о заштити становништва од инфективних болести и у Закону о заштити потрошача). У складу са стручном терминологијом уз ријеч ПРЕГЛЕД највише линија конкорданце било је везано за:

¹⁶ Видјети Тошовић 1988.

Табела 8. Примјери конкорданци са именицом ПРЕГЛЕД – Примјери су из докумената која се баве тематиком заштите потрошача и дефинисања правила инспекцијског надзора субјеката

је извршила укупно 19.631 инспекцијских ПРЕГЛЕДа, од чега 97,66% по службеној дужности,
једити праћење резултата извршења инспекцијских ПРЕГЛЕДа по одређеним групама прописа, укључу
инспектора који ће извршити инспекцијски ПРЕГЛЕД и контролу у складу са иновираном наци
ја евиденције о томе како извршити инспекцијски ПРЕГЛЕД које врше републички инспектори зашт
има, садржину записника о начину како извршити инспекцијски ПРЕГЛЕД и друго. Закон о општем управном посту

Међутим, у текстовима који се баве тематиком јавног здравља, декомпоновани предикат – извршити преглед – се готово у потпуности губи и мијења своје значење. У претходним примјерима имали смо термин преглед у смислу контроле или надзора над нечим, међутим у текстовима који се баве проблематиком јавног здравља преглед добија ново значење у смислу здравствене провјере.

Табела 9. Примјери конкорданци са именицом ПРЕГЛЕД – Примјери су из докумената која се баве тематиком јавног здравља

ругих поремећаја здравља, 2) извршити лекарске ПРЕГЛЕДе и друге врсте медицинске помоћи, 3) ли

има да изврши ПРЕГЛЕД запослених на радним мјестима која се сматрају опаснијим
Конвенција бр. 113 о томе како извршити љекарски ПРЕГЛЕД морнара, усвојена је 1959. године;
ела спроводе се одговарајуће анализе како извршити ПРЕГЛЕД и захвате ради процјене и смањења фи
их прегледа ризичне дјеце, извршити ПРЕГЛЕД дјеце са сметњама у развоју као и

У текстовима на енглеском језику имамо више различитих преводних еквивалената, од којих ћемо понудити три најчешћа, а то су: *control*, *supervision* и *examination*. Као и у примјерима на српском језику, у групи текстова који се односе на инспекцијски надзор знатно чешће се појављују термини *control* и *supervision*, док у групи текстова који се баве јавним здрављем веома често имамо *examination*. Овакви примјери свједоче о томе колико је контекст битан за преводиоца и потврђује давно познато правило да се квалитетан превод огледа у томе колико је тачно пренио контекстуално окружење оригинала.

У енглеском језику декомпоновани предикат је најчешће у облику *perform control/supervision/examination*. Мада смо наишли и на друге облике глагола који су се нашли заједно са *control/supervision/examination* као што је нпр. *carry out*. Дакле и у енглеском језику имамо декомпановани предикат који по својој формалној структури одговара оном из српског језика.

Табела 10. Примјери конкорданци са именицом CONTROL & SUPERVISION - Примјери су из докумената која се баве тематиком

заштите потрошача и дефинисања правила инспекцијског надзора субјеката

and obligations under this law 1a) perform CONTROL by forest inspection, environme
h trade: Environment Protection Agency (perform CONTROL), Ministry of Interi
s who are in charge of carrying out the inspection CONTROL in the area of sea fisheries. A
le 48 of the Law on Tax Administration. Performing inspection SUPERVISION in the meaning of this law
hiness of boats. Inspection in terms of performing SUPERVISION maintain all the necessary

и

Табела 11. Примјери конкорданци са именицом EXAMINATION - Примјери су из докумената која се баве тематиком јавног здравља

sation and means of protection; perform medical EXAMINATIONs of employees which are mandatory
pply of drinking water include: perform medical EXAMINATIONs with a special emphasis on
cal examinations and perform medical EXAMINATIONs every 12 months. Within the safety
nt mortality. In that sense, EXAMINATIONs are performed, which include

taking of
st grade of secondary school carry out EXAMINATION of students in the third grade of secon

Дакле видимо да су колокацијски елементи у српском и енглеском језику различити у зависности од тематике тј. контекста у оквиру којег се ти елементи налазе. Остали примјери декомпонованих предиката, који су карактеристични за административни стил, су:

- **Поднијети извјештај (умјесто извијестити),**
- **Изразити протест (умјесто протестовати),**
- **Упутити честитке (умјесто честитати).**

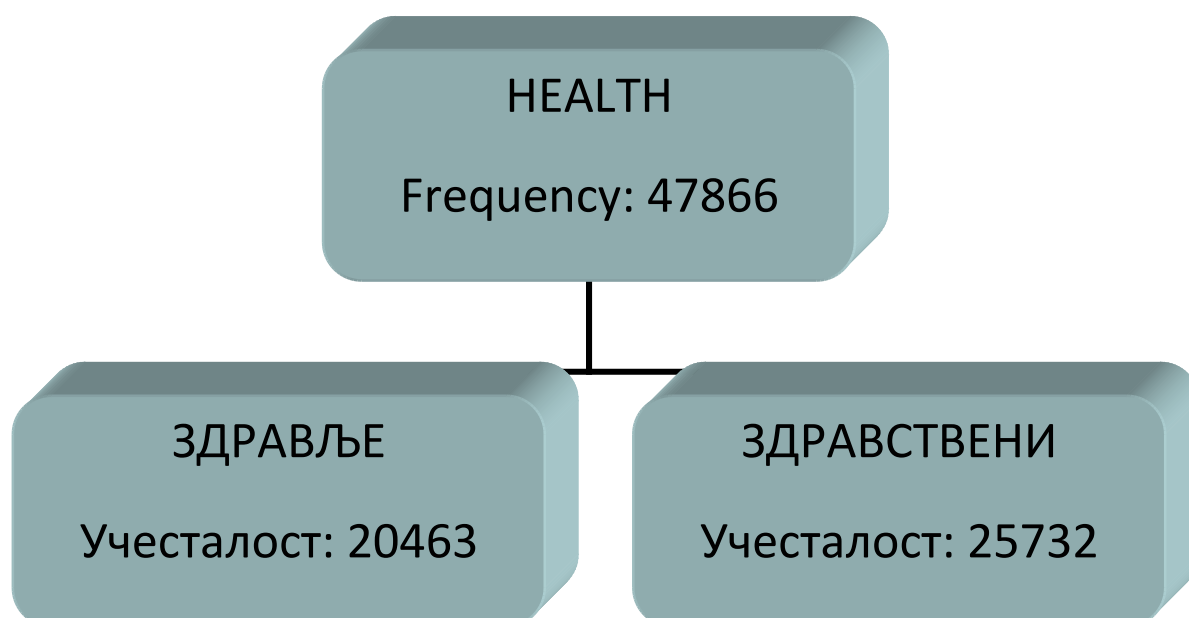
4.2.4 Колико можемо вјеровати интуицији?

Приликом разних лингвистичких истраживања вјероватно, у мањој или већој мјер, скоро сви аутори се послуже својом интуицијом када се позабаве неком проблематиком. Вјероватно је интуитивна способност изворног говорника неког језика изоштрена и веома развијена способност, али нас и она може преварити. Приликом наше анализе корпуса, управо се и аутор овог рада нашао у ситуацији када је био помало изненађен добијеним резултатима.

Па тако у погављу Упитника ЕК о Јавном здрављу и заштити потрошача, у дијелу корпуса на енглеском језику најучесталија ријеч је *HEALTH*, што се може сматрати и очекиваним с обзиром на тематику. Са друге стране у дијелу корпуса на српском језику најучесталије ријеч није

ЗДРАВЉЕ, већ је то ријеч ПРОИЗВОД/А/И. Објашњење за ову разлику можемо пронаћи у стручној терминологији која је карактеристична за поље здравства. Наиме, фраза као што је *HEALTH-CARE FACILITY* је у огромној већини случајева била преведена као ЗДРАВСТВЕНА УСТАНОВА. Са друге стране термин као што је *HEALTH-RELATED ISSUES* је обично био преведен као проблеми везани за здравље. На ова два примјера видимо да су преводни еквиваленти за енглеску ријеч HEALTH подијелени на двије групе: ЗДРАВЉЕ и ЗДРАВСТВО. Енглески термин *PRODUCT* је готово увијек имао за свој преводни еквивалент нашу ријеч ПРОИЗВОД.

Дакле, можемо закључити да најчесталија ријеч у једном дијелу корпуса не доводи закономјерно до тога да је и њен преводни еквивалент у другом дијелу корпуса најчесталији. Варијације се могу јавити услјед специфичног контекстуалног окружења, тј. преводни еквивалент зависи не само од основне ријеч већ и од контекста у којем се налази, што смо показали и кроз корпус.



Дијаграм 3. Графички приказ учесталости термина HEALTH и његових најчешћих преводачких еквивалената: ЗДРАВЉЕ И ЗДРАВСТВЕНИ

4.2.5 Изражавање забране

Приликом писања званичних државних докумената веома битан дио јесу права и обавезе субјеката које се дефинишу тим документима. Права и обавезе подразумијевају дозвољене и недозвољене радње које се пак различитим језиким средствима одобравају или забрањују. Посебно занимљиве, у овом контексту, јесу забране и на основу наше анализе нашли смо да и у текстовимана српском језику и у енглеском језику се чешће користи директни глаголи који означавају забрану (забранити или prohibit, ban, proscribe), од конструкција NOT + глагол допуштања (не + дозволити, не+допустити, not allow, not approve). Могућа су два објашњења за ову појаву:

Директни глаголи који означавају забрану представљају дио добро установљеног шаблона,

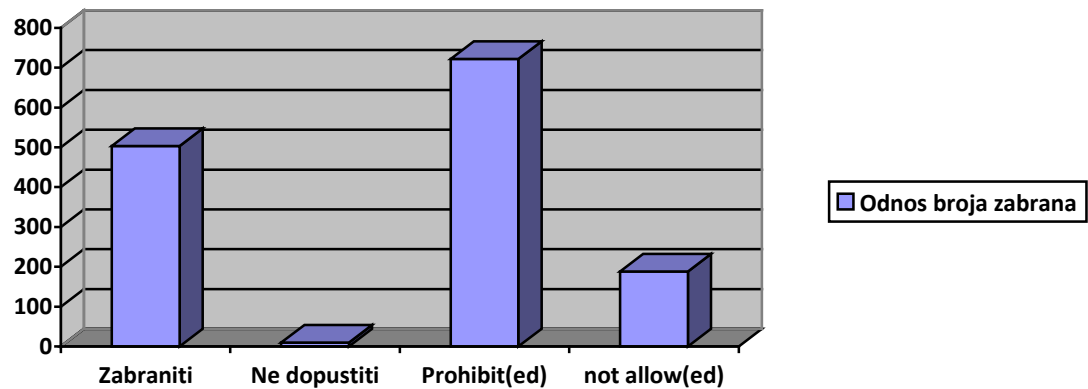
Психолошки ефекат (+ забрањено) је јачи од (- дозвољено).

Навешћемо неке од примјера:

- gation Magistrate, who has the right to BAN all correspondence if it harms the inve
- s duties постепено смањење царина PROHIBIT забранити prohibition забрана p
- ntains a set of laws whose stipulations FORBID discrimination, promote equality and es

Примјери превода на нашем језику:

- Магистрат, који има право да ЗАБРАНИ читаву преписку ако она штети
- s duties postepeno smanjenje carina prohibit ZABRANITI prohibition zabrana p
- поштује низ закона који ЗАБРАЊУЈУ дискриминацију и промовишу једнакост



Дијаграм 4. Однос броја забрана¹⁷

Са друге стране, глаголи који су неутрални у погледу забране (поступати, бирати, consider, use ...), тј. они глаголи чије семантичко поље не укључује и значење забране, да би изразили забрану захтијевају негацију NOT/НЕ (или НИЈЕ/НЕЋЕ).

4.2.6 Грешке које смо уочили током анализе корпуса

¹⁷ Поред друге ставке НЕ ДОПУСТИТИ, у анализу смо укључили и не дозволити (заједно са употребом наставака), али нам простор графикона није дозвољавао убацивање истог

Оно што је наравно било неизбежно јесте и то да смо наилазили на грешке у нашем корпусу. Под грешкама, прије свега, подразумевамо типографске грешке које се могу сврстати у седам категорија:

Цомпанија - примјер негативне интерференције са енглеским језиком

Leter - примјер негативне интерференције са српским језиком

Ћлан - грешке у избору дијакритичког знака

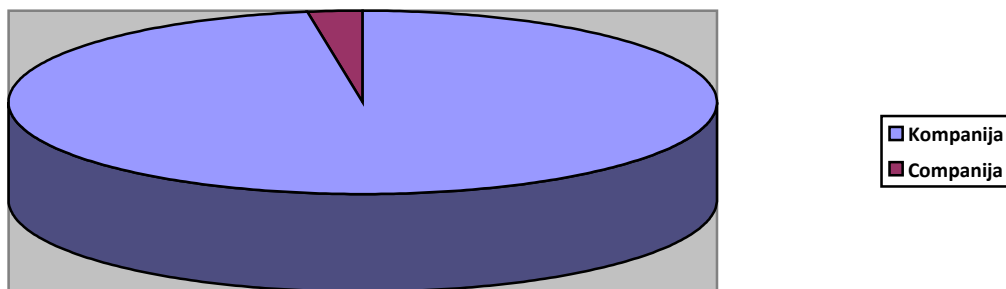
СкупштинаЦрне Горе - грешке у прављењу размака између ријечи

звјештаји - недостатак слова - непотпуна ријеч

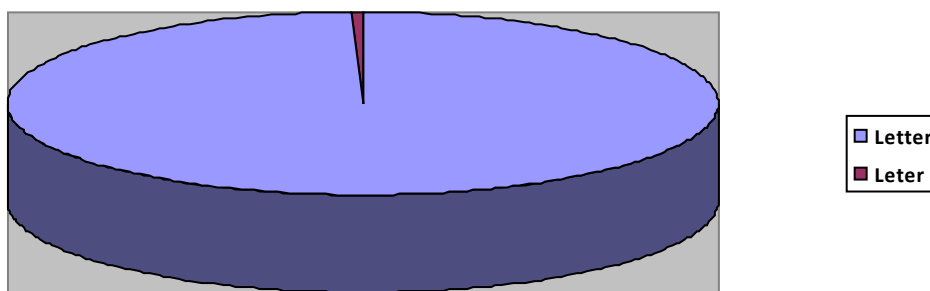
стартегије - неправилан редослијед слова

република - непотребно дуплирање слова

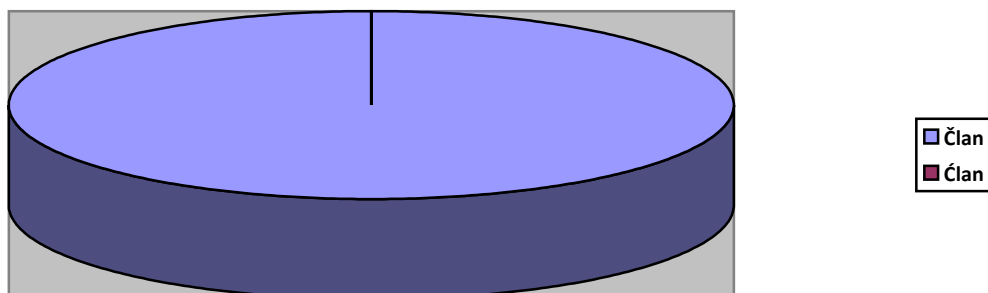
Оно што смо такође пронашли јесте да се све овакве грешке понашају по истом принципу, а то је да се налазе ван зоне учесталости понављења која је карактеристична за исправне верзије датих ријечи. Илустроваћемо неке примјере дијаграмима на којима ће се јасно видјети шаблон понашања типографских грешака у нашем корпусу.



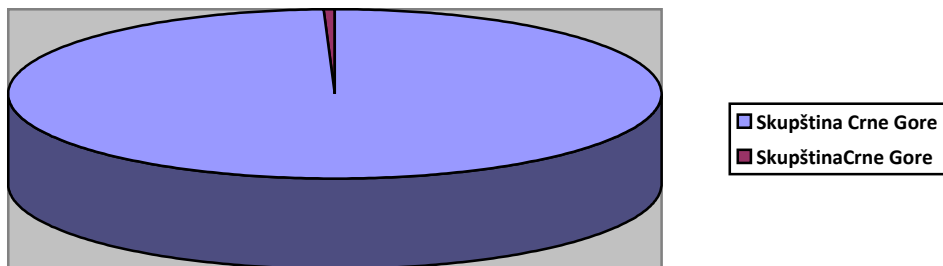
Дијаграм 5. Однос учешћа грешке бр. 1 - негативне интерференције енглеског језика на домаћу ријеч



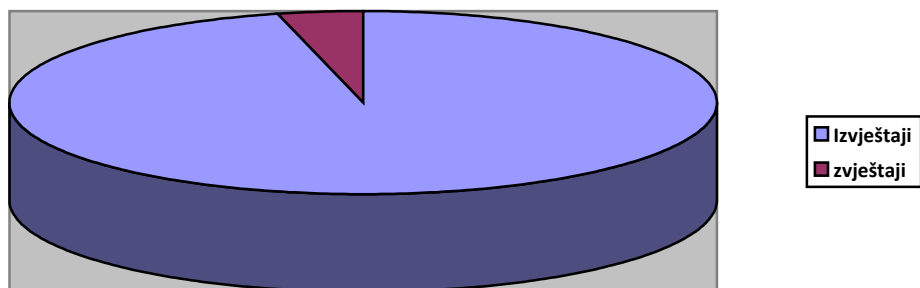
Дијаграм 6. Однос учешћа грешке бр. 2 - негативне интерференције српског језика на енглеску ријеч



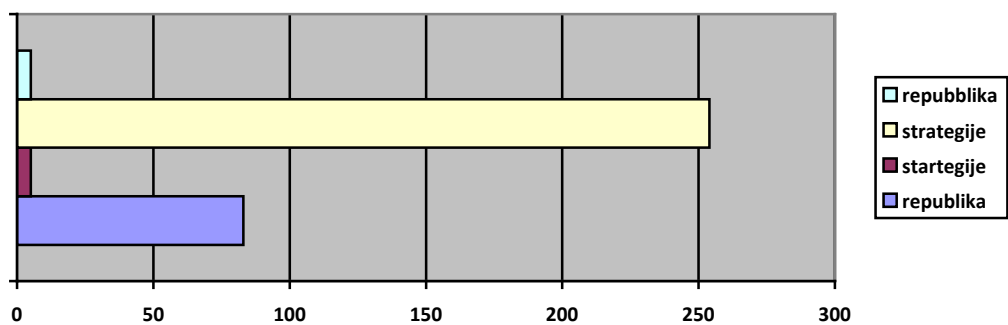
Дијаграм 7. Однос грешке бр. 3 - погрешан избор дијакритичког знака



Дијаграм 8. Однос грешке бр. 4 - неодвајање ријечи



Дијаграм 9. Однос грешке бр. 5 - непотпуна ријеч



Дијаграм 10. Однос грешке бр. 6 и 7 - неправилан ред слова у оквиру ријечи и непотребно дуплирање знакова

Јасно се може уочити шаблон гдје погрешна ријеч се јавља вишеструко мање пута у односу на исправни облик те ријечи, што је и очекивано. Ми

смо овдије издвојили само неке од примјера грешака које се могу наћи, али оно што је занимљиво јесте да и све остале грешке, без обзира на своје карактеристике, прете идентичан начин понашања, као што смо показали на Дијаграмима 5-10. Све грешке ће се увијек налазити ван оптималног броја понављања и то је правилност и универзална појава како за српски тако и за енглески језик. На овај начин можемо да пронађемо све ријечи које се налазе ван њиховог уобичајеног понављања у тексту и да их лако и брзо исправимо, а такође овај дио нашег истраживања може бити занимљив лингвистима који се баве анализом грешака у текстовима.

4.2.6.1 *Nарax legomena* и њихова веза са грешкама

Иза овог занимљивог назива крију се ријечи које су у неком тексту или групи текстова појављују само једанпут. Назив потиче од грчких ријечи (*narax* и *legomenon* (мн. *legomena*) – речено само једанпут). У нашем корпусу (од укупно 93226 ријечи без понављања, има 45.450 (48,75 %) ¹⁸ таквих ријечи и веома често су оне најуже повезане са неком струком (нпр. латинизми и врсте болести у медицини), а такође неке од ових ријечи представљају погрешно написане облике других ријечи које се иначе много чешће налазе у корпусу. Вјероватно противној свачијој интуицији слједећи облици ријечи се такође помињу само једном у цијелом милионском корпусу, нпр:

Табела 12. Неке од ријечи које су у корпусу јављају само једном

зе, банке и друга правна лица.	tte Competence: Central Register of
--------------------------------	-------------------------------------

¹⁸ Видјети поглавље о Зипфовом закону

Поједини <u>акционар</u> Централне Депозитарне Агенције може има	the <u>Commerical</u> Court 2. Making seals and stamps Timesc
раду одговарајућег документа (брошуре, <u>часописи</u> , публикације, билтени ...) Институције	e use complex analytical techniques and <u>compute</u> application software. Environmental Pro
Плц («Завод за грађевинске материјале, <u>геотехничке</u> и хемијске анализе« АД) – Никшић, ис ан	for himself/herself, marital partner or <u>illegitimate</u> partner, and for children in case that
м се дефинишу: циљеви, трајање, врсте <u>гимназија</u> , матура, хоризонтална и вертикална прох	y use for non-agricultural purposes, is <u>recultivated</u> technically, chemically and biologicall
о је и кроз изјаву Центра за подводно и <u>хуманитарно</u> разминирање да је Црна Гора једина зем	s they should realistically decline and <u>stabilise</u> at 1.5% - 1.8% of Montenegrin GDP. III.

Надовезујући се за претходно поглавље пружићемо и примјере грешака које се лако могу открити у текстовима укљученим у наш корпус:

Табела 13. Ријечи које су у корпусу јављају ван свог оптималног броја понављања из разлога што су погрешно написане

путовања могу да обављају туристичке <u>агениције</u> које, поред одобрења за рад (чл.11 ст.	plementation and evaluation of measures <u>aggenst</u> TC should be included in programs for M
у (Сл. лист ЦГ, бр.66/08), одговорна	ional Protocol (OG of MN,

је <u>Агенција</u> за цивилно ваздухопловство Црне Горе. 3	international <u>agreements</u> , 2/09); Law on Veteran and Disability P
области здравства, ветерине, фармације, <u>архитектуре</u> , наставно-педагошког рада итд. Свако ми	d lays down requirements for <u>quality</u> of <u>agricultural</u> products. Veterinary Administration is
равља у Црној Гори су отворени следећи <u>центри</u> и јединице за ментално здравље у зајед	alcohol per 100 kilograms of goods; 4) <u>chemicals</u> and cosmetics manufacturing. Use of min
циљева других сегмената друштва. Такве <u>стратегije</u> су Стратегија одрживог развоја, Стратег	ce stipulates that the doctor - foreign <u>citizen</u> , who intends to perform <u>health</u> activity

Као што се може видјети грешке у оба језика функционишу у основи на идентичне начине. У српском језику су то обично сувишак или мањак знакова у ријечима, као што је то случај и у енглеском језику, оно што се познато под именом *SPELLING*.

4.2.7 Анализа употребе модалних глагола у корпусу¹⁹

Модални глаголи у документима Европске уније се најчешће користе у оквиру правних норми у оквиру којих исказују своја значења. Норме се различито језички формулишу и према облицима тих формулација (употријебљених језичких исказа), дијеле се на:

¹⁹ Погледати приручник МВПиЕИ, одакле је ова група текстова преузета.

1. наређујуће,
2. забрањујуће,
3. овлашћујуће и
4. допуштајуће правне норме.

4.2.7.1 Наређујуће (прецептивне) норме²⁰

Ове норме обавезују адресате на одређено чињење или нечињење. Ова врста норми јавља се најчешће у управном праву. У нашем језику се ове норме (*commands*) изражавају презентом, а у енглеском се користи ријеч *shall* која на подручју права изражава правни императив. *Shall* се користи само у диспозицији, али не и у хипотези. У новије вријеме, у ЕУ се покушава језик права учинити разумљивијим обичном човјеку па се предлаже да се умјесто *shall* користи *must*, али приједлог није шире прихваћен.

Примјер:

Висина накнаде штете одређује се према цијенама у вријеме доношења судске одлуке, осим случаја кад закон одређује нешто	<i>The amount of compensation shall be determined in accordance with the prices at the time the court decision is rendered, unless otherwise provided by law.</i>
Новчана рента плаћа се мјесечно унапријед.	<i>Annuity payments shall be made monthly in advance.</i>

²⁰ Ибид.

Глагол *must* по правилу се користи за исказивање одредаба код којих вршилац радње није наведен, нити се из контекста може утврдити, углавном у пасивним реченицама. *Must* се такође користи у одредбама које захтијевају испуњавање или удовољавање неким условима, а субјект није живо биће:

Вриједносни папир мора садржати слједеће битне састојке:	<i>Securities must fulfil the following essential requirements:</i>
Испитивање треба обављати тако да се у потпуности поштује личност оптуженог.	<i>The interrogation must be conducted with full respect for the personality of the person charged.</i>

4.2.7.2 Забрањујуће (прохибитивне) норме²¹

Усмјерене су на забрану одређеног понашања, да се нешто не чини (нечињење, уздржавање). Та забрана може бити изричита, индиректна и несанкционисана. Пропис који садржи правну забрану, али не и санкцију у случају њене повреде, назива се „несавршеним законом“ („*lex imperfecta*“).

За исказивање забране користи се негативни императив – *shall not* – или се негира субјект, нпр. *No person shall ...* Кад се изричито наводи да је нешто забрањено („забрањује се“, „забрањено је“), преводи се са *it is prohibited*.

Власник ствари не смије вршити своје право власништва преко граница које су свим власницима таквих ствари прописане овим	<i>The owner of a thing shall not exercise his right of ownership beyond the limits placed on all owners by this Law...</i>
--	---

²¹ Ибид.

Нико не смије угрозити здравље	<i>No person shall endanger the health of others.</i>
Забрањена је употреба личних података супротна утврђеној	<i>It is prohibited to use personal data contrary to the purpose for which they were collected.</i>

Одредба којом се укида неко овлашћење, дозвола или право такође представља забрану која се често изриче са „не може се” или „није допуштено”, односно са *may not* у енглеском. Сличним одредбама изражава се да неки субјект или јавна власт нема овлашћење, дозволу или право да нешто чини:

Одрицање или одустанак од жалбе не може се опозвати.	<i>The waiver or withdrawal of an appeal may not be revoked.</i>
... власник није овлашћен забранити туђи захват у своју	<i>... the owner may not prohibit others to interfere with his property if it is necessary...</i>
Ако је материнство, односно очинство утврђено одлуком	<i>A court decision establishing maternity or paternity may not be contested.</i>

За негирање обавезе испуњавања или удовољавања у српском језику употребљава се „није потребно” или „не треба”, а у енглеском *need not* или *no person must*:

За правоваљаност овог споразума није потребно да буде закључен у облику у којем је склопљен посао из којег је обавеза настала.	<i>In order to be valid, this agreement need not be concluded in the same form as the transaction from which the obligation arose.</i>
--	--

4.2.7.3 Овлашћујуће (пермисивне) нормe²²

Не садрже ни наредбу ни забрану, него дају овлашћење лицу или групи лица да стварају правне нормe или друге правне акте (закључење уговора), односно да предузимају радње извршења. Ако овлашћена лица не приступе употреби овлашћења која су им призната, неће због тога трпјети никакву правну санкцију. Одредбе о надлежностима органа својеврстан су примјер овлашћујућих норми.

Понекад је тешко разликовати допуштење од овлашћења, али то не ствара проблем за преводиоца јер, иако је њихова функција различита, у енглеском се претежно употребљава *may* за изрицање и допуштења и овлашћења. Допуштење се исказује са „може се”, „смије се” или „допуштено је”. У одредбама о поступку „допуштено је” значи „*it is admissible*”, док се изрази „*it is permitted*” и „*it is allowed*” не користе у правним прописима у енглеском језику. Овлашћење се у прописима изражава са „овлашћен је” или „може се”. За изрицање овлашћења у енглеском језику се, осим „*may*”, употребљава „*is authorized to*”, „*is empowered to*”, „*has the power to*”. Ако је овлашћење обавезујуће, односно ако тражи неку одређену радњу, преводилац треба да употрејеби „*shall*”.

Допуштење

Уговор, као и други правни посао, може се закључити, односно извршити и преко заступника.	<i>Contracts and other transactions may also be concluded or performed through an agent.</i>
Против другостепене пресуде о утврђивању или оспоравању	<i>Appeal is admissible against the judgment of a court of second instance establishing or</i>

²² Ибид.

Ко има право на заштиту посједа, смије свој посјед заштитити...	<i>Persons having the right of possession may protect their possession...</i>
--	---

Овлашћење

Влада може ... распустити Скупштину (члан 92 Устава ЦГ).	<i>The Government is empowered to dissolve the Parliament.</i>
Заступник је овлашћен да, ради очувања права свог налогодавца,	<i>In the interest of safeguarding the rights of the principal, agents are authorised to make</i>

Основна права и слободе које су зајамчене разним документима ЕУ исказују се у облику наредби које обавезују државу да обезбиједи заштиту тих права и слобода, нпр.: "ЕУ јемчи и штити права и слободе; права и слободе су неповрједиве; свако је обавезан да поштује права и слободе других"; "Држава јемчи равноправност жене и мушкарца и развија политику једнаких могућности".

Најчешће се ова права одређују путем индикативног исказа „имати право“, односно „*has the right to*“: „Свако људско биће има право на живот“ - „*Every human being has the right to life.*“ У другим правним прописима најчешћи облик за одређивање неких права је такође „имати право“, док се у енглеском употребљава „*has the right to*“, „*is entitled to*“, или чак „*may*“.

Повјерилац има право на затезне камате, без обзира на то је ли претрпιο какву штету због дужниковог... закашњења.	<i>The creditor is entitled to interest on arrears regardless of whether any damage was suffered due to the delay.</i>
--	--

Против рјешења вијећа приватни тужилац има право жалбе.	<i>A private plaintiff may appeal the decision of the court chamber. / A private plaintiff has the right to appeal the decision of the court</i>
---	--

Као што можемо да видимо, ова област је до најситнијих детаља дефинисана у смислу који преводилачки еквивалент треба користити у којој ситуацији, тј. контексту. То је неопходно из разлога што се текстови правне стурке (закони, декларације, прописи, одлуке) морају веома прецизно преводити како би се у потпуности задржала изворна порука. Наравно, у сваком преводу је најбитније задржати изворну поруку, али неки типови текстова дозвољавају више слободе него што је то могуће у горе наведеним примјерима. Правни текстови су дакле веома „осјетљиви“ на слободно тумачење текста, јер је информација обично веома прецизно „запакована“ и као таква мора бити и пренешена. Други елементи који онемогућавају слободнији превод овакве врсте текстова јесте жеља за стандардизацијом то униформношћу превода. Веома је битно да се нпр. *decree* увијек преводи као указ и да се тако преводи у цјелокупном тексту ради уједначености превода. У супротном, превод постаје отворен за разна тумачења, што обично није добро, нарочито када се ради о облигацијама које проистичу из законских прописа.

4.2.8 Зипфов закон

Учесталост ријечи у неком тексту се понаша по правилима Зипфовог закона (1935.), који нам говори да ће се око половине тих ријечи појавити само једанпут, четвртина двапут итд. Па на тај начин налазимо да у Брауновом корпусу (*the Brown Corpus*) писаног америчког енглеског од 69002 различитих лексичких форми, њих 35065 се појављује само једанпут.

На врху те љествице, појављује се ријеч (тј. одређени члан) *the* са фреквенцијом од 69970, што је готово душло више од његовог пратиоца на другом мјесту, препозиције *of*, која се јавила 36410 пута у поменутом корпусу.

Зипфов закон је емпиријски закон који је формулисан уз помоћ математичке статистике и односи се на чињеницу да се један дио језичких феномена може дефинисати Зипфовом дистрибуцијом. Закон је добио име по лингвисти Џорџу Кингслију Зипфу (George Kingsley Zipf) који га је ближе дефинисао у својим дјелима из 1935 и 1949.

Зипфов закон нам говори да у оквиру корпуса текстова који су настали у природном окружењу учесталост ријечи је обрнуто пропорционална њеној мјесту на листи учесталости. Стога, најчешћа ријеч ће се отприлике појавити душло више пута од ријечи која је на другом мјесту по својој учесталости. Та ријеч која је на другом мјесту по својој учесталости ће се јављати душло више него што је то ријеч која је на четвртм мјесту по учесталости, итд.

Овај закон преточен у формулу се може исказати као:

$$f(k; s, N) = \frac{1/k^s}{\sum_{n=1}^N (1/n^s)}$$

Гдје је:

N – број елемената;

k – њихова позиција на листи учесталости;

s – вриједност експонента који карактерише дистрибуцију.

Најједноставнији случај Зипфовог закона је $1/f$ функција. На основу Зипфових распоређених фреквенција, од најучесталије ријечи до најмање учестале, друга најчешћа ријеч ће се јављати двоструко мање него прва ријеч, трећа најчешћа ријеч ће се јављати три пута мање него најучесталија ријеч. Односно ако генерализујемо:

Н-та најчесталија ријеч ће се јављати $1/n$ пута често као најчесталија ријеч.

Оно што је битно напоменути јесте и то да Зипфов закон своју пуну афирмацију добија приликом његове примјене на великим корпусима уз дозу неопходне апроксимације, што значи да се лингвистички феномени дистрибуције ријечи неће увијек понашати тачно по овом закону за сваку ријеч понаособ, већ ће то бити случај на генералном нивоу.

4.2.9 Хипов закон

У Лингвистици Хипов закон (Heap's Law) је емпиријски закон који описује дио вокабулара који се налази у неком тексту или текстовима, а састоји се од ријечи које су одабране из тог вокабулара. Закон се може формулисати и на следећи начин:

$$V_p(n) = Kn^\beta$$

Гдје је V_p подобласт вокабулара V која је заступљена у дијелу текста величине n . K и β су слободни параметри који се одређују на емпиријски начин.

Што се тиче корпуса који су састављени од текстова на енглеском језику вриједност K је обично између 10 и 100, а β је од 0,4 до 0,6.

Хипов закон се не употребљава само за скупове лексичких елемената, већ се може користити и на друге скупове објеката, људи, итд. На примјер, ако овај закон примјенимо на групу људи који су дефинисани по земљи одакле потичу, и изабирамо их насумично (дакле не по критеријуму одакле Хипов закон нам онда говори да ћемо веома брзо имати представнике из већине земаља (у односу на њихову популацију), али ћемо веома тешко моћи да покријемо све земље које су укључене у поменути скуп, ако наставимо са оваквим насумичним одабиром.

4.2.10 Карактеристике превођења текстова административног регистра

Приликом анализе оригиналних и преведених текстова врло лако се могу уочити одређене законитости које се сматрају *conditio sine qua non* за овај регистар.

Индивидуални стручни термини се већином преводе само једним преводним еквивалентом (стандардизација превода - ријечи),

Колокацијски везани стручни термини се такође већином преводе увијек истим преводним еквивалентима (стандардизација превода - фраза).

Прве двије констатације су у складу са сугестијама Министарства за европске интеграције Црне Горе, како би се обезбиједио стандардизовани превод, тј. омогућило да превод, што је више могуће, одговара оригиналу.

Гдје год је могуће примјењује се техника превођења „ријеч за ријеч“ (енг. *Word for word*), како би се смањило одступање од изворног текста.

Административни стил је понекад преводиоцу тежак за разумијевање јер садржи доста:

- a. Уметнутих реченица,
- b. Информација које се нижу једне за другом,
- c. Гомилања ријечи,
- d. Веома сложених реченица које понекад саме заузимају и цијеле параграфе, па је понекад веома тешко пратити мисаону нит аутора текста.

Начин на који се преводиоци боре са овим јесте да имају што више потребних информација, тј. знања о датој области, али и да, гдје је то год могуће, преводе структуре ријеч за ријеч. У смислу претходне реченице карактеристичан је сљедећи превод:

Субјекат / надзора / може / да побија/ резултат / анализе / првог / узорка / захтјевом / за / вршење/ анализе / другог / узорка / (узетог / у / исто / вријеме / и / на / исти / начин), / у / року / од / три / дана.

Односно

The subject / of supervision / can/ refute / results / of analysis / of the first / sample / by request / for / the performance / of analysis / of the second / sample / (taken / at / the same / time / and / in / the same/ manner), / within / the period / of three / days.

4.3 Контрастивна анализа граматичких елемената у корпусу

Нашу контрастивну анализу запошећемо одређеним чланом који је једна од карактеристика која је незаобилазна приликом било каквог укључивања енглеског језика у контрастивну анализу.

4.3.1 Одређени члан

У вези са одређеним чланом не можемо говорити о његовом значењу већ о референци. У том смислу говоримо о издвајању или спецификацији неког предмета. У енглеском језику именица се у дискурсу обично јавља, прво са неодређеним чланом, а затим добија одређеност

помоћу одређеног члана. На тај начин одређени члан постаје спона између претходног јављања именице и њеног наредног јављања у комуникацији. Генерално говорећи, одређени члан се у нашем језику не преводи, иако за њега постоје одговарајући преводни еквиваленти. То се може илустровати примјером да је у посматраном узорку текста, у овом случају у предговору рјечника М. Бенсона, од 358 уочених случајева јављања одређеног члана, само у 4 реченице јављају се њихови директни преводни еквиваленти. Статистички гледано, то значи да је одређени члан преведен у само 1,11% случајева. Навешћемо пронађене употребе у којима се јавио одређени члан:

Катафорична референца одређеног члана:

- The Government of Montenegro remain open for constructive cooperation.
- Ø Влада Црне Горе остаје отворена за конструктивну сарадњу.

Директна анафорична референца одређеног члана:

Delegation of competencies enables an increase of flexibility in the functioning of an administrative system. The administrative system will also increase its level of efficiency.

Делегација надлежности омогућава повећање флексибилности дјеловања управног система. Ø Управни систем ће такође повећати свој ниво ефикасности.

Општа референца одређеног члана:

The Minister pointed out value of support.

Ø Министар је нагласио вриједност подршке.

Класификативна референца одређеног члана:

The Montenegrin people are fully committed to the idea of the EU.

Ø Црногорски народ је у потпуности посвећен идеји ЕУ.

Нађени су случајеви у којима је одређени члан преведен:

1. The Guideline does not treat in detail the specialised vocabularies.
1а У овом Упутству није детаљно обрађена специјализована терминологија.
2. The Guideline will transcribe this standard American pronunciation.
2а У овом Упутству транскрибоваће се овај стандардни амерички изговор.
3. The year was devoted to the gathering of material.
3а Цијела је година била посвећена сакупљању материјала.
4. The chief difference between the two dialects is a phonetic feature.
4а Главна разлика између ова два говора јесте фонетска црта.
5. The parties hereby agree to the following.

Ба Односне стране су се сагласиле о сљедећем.

Можемо закључити да се одређени члан у огромном броју случаја не преводи. Ако се ипак преведе најчешће нађени примјери, што из текста, што из граматика, упућују на коришћење замјеница: овај/а/о, тај/а/о, онај/а/о. Међутим могући су и алтернативни преводи у случајевима гдје преводилац има додатне информације које читаоцу нису познате. То се види у случајевима 3а и 5а гдје знање преводиоца (стручно, искуствено или у виду информације добијене од самог аутора оригиналне реченице) условљава постојање „слободнијег“ превода одређеног члана.

4.3.2 Колокације придјева и именица

Колокација представља начин организовања двије или више ријечи у једниствену цјелину како би се на конвенционалан начин изразиле одређене идеје. Тако колокација постаје фиксна форма изражавања која као таква постаје позната говорницима неког језика. Примјера у енглеском језику има заиста много, као што је нпр. *strong tea*, или *broad daylight*. Имамо *powerful друг* али не и *powerful cigarettes, tea* и *coffee*. Међутим, најзанимљивије су оне колокације које се дјелимично или потпуно косе са свакодневном здраворазумском логиком. И поред тога такве колокације су опстале у језичкој пракси и опште су прихваћене.

Можемо говорити о више типова колокација међу којима имамо:

именица + глагол –

Parliament decides - Парламент одлучује,

Ministry approves – Министарство одобрава;

глагол + именица –

reach a decision – донијети одлуку;

придјев + именица –

a clear conscience, чиста савјест;

глагол + прилог –

speak bluntly, говорити неувјено;

прилог + придев –

seriously considered – озбиљно размотрен;

именица + of, ГЕНИТИВ + именица –

a clause of excision, клаузула искључивања.²³

У следећим примјерима размотрићемо случајеве колокација придјева и именице:

Најчешће нађени примјери су гдје енглеска конструкција Adjective + Noun одговара нашој структури придјев + именица:

Key objectives of the Public Administration Reform.

Кључни циљеви управне реформе у Црној Гори.

²³ Прћић, Твртко: *СЕМАНТИКА И ПРАГМАТИКА РЕЧИ*, Библиотека Тхеориа, Издавачка књижарница Зорана Стојановића, Сремски Карловци – Нови Сад, 1997.

This, of course, requires development of quality and quantity standards of administrative work.

Ово изискује развијање одговарајућих квантитативних и квалитативних стандарда управног рада.

We should encourage complete support to the administrative reform support.

Треба да охрабримо пуну подршку процесу реформе области управног система.

Али није увијек тако. Енглеској конструкцији Adjective + Noun могу да одговарају и друге структуре у нашем језику.

They performed vaccination in broad daylight. (али не говоримо о бригхт daylight или о narrow darkness)

Обавили су вакцинацију у по бијела дана.

Do not block the emergency exit!

Не блокирајте излаз за случај опасности!

Као што се може видјети из примјера постоје разни начини превођења колокација придјева и именице. Најчешће нађени однос је да су колокације придјев + именица превођена управо таквим распоредом, дакле структура придјев + именица се такође налази и у нашем језику. Ову праксу је могуће потврдити и кроз друге примјере. Наиме, једна студија је показала да је у корпусу New York Times у којем се налазило приближно 14 милиона ријечи нађено 40375 колокација именица + придјев. Оно што је такође занимљиво јесте чињеница да нису узете у обзир колокације које су се јављале мање од 1000 пута заједно.

4.3.3 Глаголски облици

4.3.3.1 Инфинитив

4.3.3.2 Упоредна анализа инфинитива у корпусу

Инфинитив спада у нефинитне облике заједно са герундом и партиципом. То је облик који представља основу глагола, а негативни облик гради дистрибуцијом партикуле *NOT* испред облика. Он има одговарајуће активне и пасивне облике, а активни облици инфинитива су:

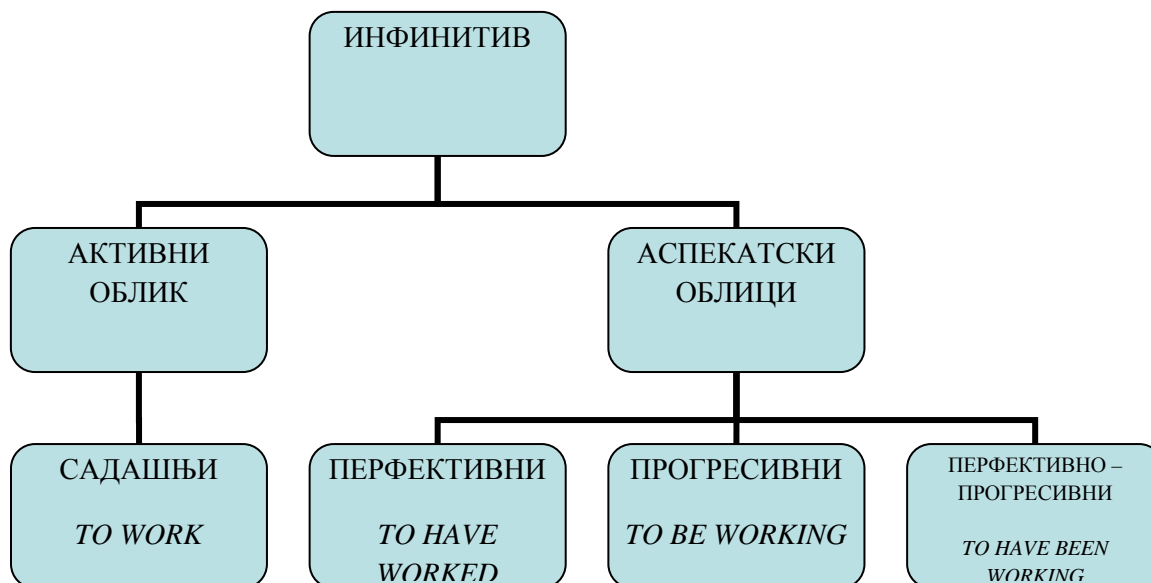
садашњи инфинитив (*WORK*),

перфективни инфинитив (*HAVE WORKED*),

прогресивни инфинитив (*BE WORKING*) и

перфективно - прогресивни инфинитив (*HAVE BEEN WORKING*).

Посљедња три облика представљају аспекатске форме инфинитива. Ради боље прегледности ови облици се могу приказати табеларно:



Садашњи инфинитив је облик који се користи у различитим контекстима и уз разноврсне финитне глаголе. Управо на тај начин он прихвата вријеме од облика који стоји са њиме у некој реченици. У зависности од управног глагола садашњи инфинитив може да се односи на садашњост, прошлост и будућност:

Montenegro is willing to negotiate about some new techniques.

Црна Гора је вољна да преговара о неким новим техникама.

Сходно томе, и у сврху анализе, ова реченица се може проширити:



Montenegro is (NOW) willing to negotiate (NOW) about some new techniques.

Црна Гора је (САДА) вољна да преговара (САДА) о неким новим техникама.

Као што видимо, инфинитив се налази у истој временској равни са управним глаголом.

Други наведени облик је перфективни инфинитив који има двије функције темпоралну и имагинативну функцију, а такође се користи у у кондиционалним реченицама трећег типа.

1. Transmittable disease must have broken out in the region...
2. Заразна болест мора да је избила у региону...

1. We would like to have seen the benefits...
2. Вољели би да смо видјели користи...

1. If we'd seen the terms and conditions we would have considered it.
2. Да смо видјели услове уговора размотрили би смо их.

Прогресивни инфинитив показује већи степен неминовности догађаја и може да изрази радњу одређеног трајања у некој тачки у времену.

1. It would be beneficial to be performing in a dramatically better way.
2. Заиста би било добро да обављамо посао на много бољи начин.

1. Public health-care units must be performing ...
2. Јединице јавне здравствене заштите морају да обављају ...

Последњи наведени аспекатски облик инфинитива је перфективно – прогресивни. Овај облик је сличан прошлом прогресиву или прошлом перфективном прогресиву јер истиче трајање радње која је почела у прошлости.

1. The region must have been waiting for decades.
2. Регион мора да је чекао деценијама.

1. They pretended to have been negotiating all day.
2. Претварали су се да су преговарали читав дан.

Инфинитив у српском језику:

Инфинитив је основни, безлични, глаголски облик. Не можемо говорити о јединственој дефиницији инфинитива, јер су његове употребе веома разноврсне, како у српском, тако и у другим језицима у којима се такође јавља. Његова можда и најчесталија улога јесте грађење других глаголских облика. Инфинитив посједује и неке друге карактеристике:

употребљава се за грађење футура и

Сјутра ћемо повести цјелокупну делегацију...

има функције других врста ријечи, често као допуна глаголима непотпуног значења типа *моћи, требати, морати, хтјети*;

Ноћас морамо донијети резолуцију...

- има функције других врста ријечи, често као допуна глаголима дјелимично непотпуног значења типа *жељети, смјети, умјети*;

Они желе чути његово мишљење.

њиме се може изрећи и субјект

Преговарати је веома битно, нарочито у ситуацији...

Употребљава се као допуна уз именичке, придјевске и прилошке изразе непотпуног значења: *право је, дужност је, ружно је, намјера је, кадар је*:

Дужности агенције су: контролисати активности усмјерене...

Као што можемо закључити из описа инфинитива у ова два језика постоје неке веома очигледне разлике у облику инфинитива у српском и енглеском језику. По дефиницији, у нашем језику, инфинитив се

правилно гради додавањем наставка -ти на инфинитивну основу, док у енглеском језику инфинитив је основа, и не можемо говорити о правилном или неправилном грађењу инфинитива. Такође у енглеском језику не можемо говорити о карактеристичном наставку по којем се инфинитив препознаје, док у нашем језику наставци -ти, -ћи су карактеристични. У српском језику не можемо говорити о садашњем, перфективном, прогресивном и перфективно - прогресивном инфинитиву. Дакле, постоје знатне разлике између облика инфинитива у српском и енглеском језику, али и у једном и у другом језику инфинитив представља основну форму од које се полази приликом грађења комплекснијих граматичких категорија.

4.3.3.2.1 Примјери контрастиране категорије

Примјери за садашњи инфинитив

Садашњи инфинитив наспрам ДА клаузе са презентом:

...where the family law defines all responsibilities enough to protect any entity who is...

... гдје закон о породици дефинише све одговорности у довољној мјери да заштити особу која...

When this population came there to live they were...

Када је ова популација дошла ту да живи, она је била...

...but we will have to wait to see what would happen.

... али ми ћемо морати сачекати да видимо шта ће се десити.

Садашњи инфинитив наспрам ДА клаузе (са повратним глаголом у презенту)

We are willing to get back to the path...

Ми смо вољни да да се вратимо на стазу...

Садашњи инфинитив наспрам глагола у перфекту:

The EC seemed to hesitate for a time...

Европска комисија је једно вријеме оклијевала...

Садашњи инфинитив наспрам приједлошке фразе:

This is of particular significance in the current phase of resolute commitment of Montenegro to integrate into the Regional and Euro – Atlantic Integrations.

То је од посебног значаја у овој фази одлучне оријентације Црне Горе за укључивање у регионалне и евроатлантске интеграције.

Садашњи инфинитив наспрам инфинитива (на -ти и -ћи):

1. Therefore, it is necessary to attract and keep the most capable staff.
2. Управо зато треба привући и у њој задржати најстручније кадрове.

Садашњи инфинитив наспрам именице:

...it is equally important to eliminate unnecessary obstacles...

Исто тако је значајно и елиминисање непотребних препрека...

It is necessary to shift from classical methods to modern methods.

Неопходан је прелазак са класичних на модерне методе.

Примјери за перфективни инфинитив

Перфективни инфинитив наспрам глагола у перфекту:

...and EC would have reached the decision had they not known very well...

И Европска комисија није донијела одлуку јер су добро знали да...

Перфективни инфинитив наспрам потенцијала:

1. The Committee would have preferred a transparent method which was obvious to ordinary citizens...
2. Комитет би преферирао неки транспарентни метод који је био очигледан обичном грађанину...
1. By that time we should not have had any difficulties...
2. До тада ми не би имали никакве потешкоће...

Перфективни инфинитив наспрам ДА клаузе са презентом:

We might have adopted the new law ...

Могли смо да усвојимо нови закон...

Примјери за прогресивни инфинитив

Прогресивни инфинитив наспрам футура I:

The Committee decided that both teams were soon to be moving beyond ...

Комитет је одлучио да ће се тимови измјестити ван ...

Boats will be few and the surface fish will be swimming deep.

Неће више бити чамаца, а површинске ће рибе заронити дубоко.

Прогресивни инфинитив наспрам Ø

The government could be spending time to train unemployed...

Влада би могла обучавати незапослене...

Прогресивни инфинитив наспрам презента:

(У овом примјеру имамо два инфинитива један до другог. Први инфинитив је прогресивни, а други је садашњи инфинитив који је преведен ДА клаузом са презентом)

1. The public administration might be struggling to break out of their limits...
2. Јавна управа покушава да савлада ограничења...

Примјери за перфективно – прогресивни инфинитив

Перфективно – прогресивни инфинитив наспрам перфекта:

Task group must have been thinking again of how to...

Радна група мора да је поново размишљала то томе како да...

Перфективно - прогресивни инфинитив наспрам потенцијала:

It was done too swiftly and handily for the organisations to have been inventing it...

То је урађено пребрзо и превјешто да би га организације измислиле ...

Перфективно - прогресивни инфинитив наспрам инфинитива:

If they were guilty, they ought to have been running out of the house...

Ако су били криви, требали су бјежати из куће ...

4.3.3.3 Резултати

Из наведених примјера се види да постоји велики број начина на који се инфинитив може превести. Сви понуђени преводи зависили су од конкретне ситуације односно контекста. Најчешћи облици инфинитива који су се јављали су садашњи инфинитив и перфективни инфинитив, који је углавном долазио након модалног глагола. Готово је немогуће наћи примјер гдје се садашњи инфинитив не би могао превести ДА клаузом:

This is of particular significance in the current phase of resolute commitment of Montenegro to integrate into the Regional and Euro - Atlantic Integrations.

Званично понуђени превод је:

То је од посебног значаја у овој фази одлучне оријентације Црне Горе за укључивање (да се укључи) у регионалне и евроатлантске интеграције.

Наше мишљење је да је превод који је понуђен у загради бољи, јер задовољава све стилске критеријуме нашег језика.

Као што видимо облик *да се укључи* у потпуности може да замјени облик *укључивање*.

Такође смо наши и примјере потпуног тј. и формалног поклапања и у том случају говоримо о томе да инфинитиву у енглеском одговара инфинитив у српском језику. Примјер оваквог односа је реченица:

Therefore, it is necessary to attract and keep the most capable staff.

Управо зато треба привући и у њој задржати најстручније кадрове.

Генерално говорећи, овај однос између два језика у нашем корпусу је рјеђи, јер је у српском језику, условно говорећи, чешћа ДА клауза, и превођење инфинитива инфинитивом је више карактеристика западних језика (мада ову изјаву треба узети са резервом јер постоје одређена неслагања у вези тога да ли је инфинитив учесталији или је то да-конструкција²⁴). Стога друга реченица може да има облик.

Управо зато треба да се привуку и у њој задрже најстручнији кадрови.

²⁴ Однос између да-конструкција и инфинитива више пута био је предмет пажње у српској лингвистици, а судови о том питању који су научној јавности најпознатији у основи су подударни, нпр. "источно подручје стандардног језика [...] паушално [је] проглашавано `дакавским`: на њему је, истина, употреба допунске реченице с везником да шири него на западу, али то нипошто не значи искључивање инфинитива" (М. Ивић); "Као допуна других глагола, у српскохрватском се у многим случајевима напоредо и углавном истозначно јављају инфинитив и синтагма да + садашње време. Инфинитив је осетно чешћи у западној варијанти" (И. Клајн); "У српском језику конструкција `да плус презент` више је у употреби него у западним говорима, мада је инфинитив економичнији, што није без значаја у језичкој комуникацији" (Е. Фекете).

Цјелокупни текст се може погледати на:

<http://www.rastko.rs/filologija/odbor/odluka024.html>

4.3.3.4 Закључак

Можемо да закључимо да инфинитиву у енглеском језику, као нефинитном глаголском облику, најчешће одговара неки од финитних облика у српском језику (презент или перфекат), и то зависи од тога које је доминантно вријеме датог текста или реченице. Такође и ДА клауза са презентом је веома чест преводни еквивалент, и то већином садашњег инфинитива. Инфинитив наспрам инфинитива је био доста риједак јер, како је то већ наведено, то је више карактеристика хрватског језика²⁵. На тај начин закључујемо да је инфинитив веома честа граматичка категорија у енглеском језику (што је уједно и карактеристика свих германских језика) док је код нас редукована његова употреба. Али, и поред тога, инфинитив у српском језику је незамјенљиви преводни еквивалент приликом преводјења конструкције *will* + прогресивни инфинитив гдје се готово увијек у преводу јавља футур, који се код нас гради од помоћног глагола хтјети и инфинитива глагола који узимамо (*will be swimming* → ће запливати).

Посматрањем инфинитива и његове фреквентности у енглеском језику долазимо до закључка да је инфинитив (заједно са пасивом) веома захвалан лингвистички елемент којим се многе идеје могу изразити на лијеп и концизан начин. Инфинитив је обавезан послије модалних глагола, који као веома честа граматичка категорија самим тим чине и инфинитив веома честим. Због тога је инфинитив веома битан за преводачку праксу јер се може врло лако погрешно превести ако се занемари његов облик. Примјера ради, ако би се реченица

Task group must have been thinking again of how to...

²⁵ Ibid.

Радна група мора да је поново размислила то томе како да...

изгубили би нијансу дужег трајања радње која је обавезна приликом превођења прогресивног елемента у оквиру енглеског перфективно – прогресивног инфинитива. Дакле, преводилац мора бити свјестан комплексности различитих облика инфинитива у енглеском језику. Такође, ово је битна идеја које је важна и за наставу енглеског језика као страног, јер кроз познавање различитих облика инфинитива и нијанси значења које ти облици посједују, могу се научити многе карактеристике граматике наших финитних глаголских облика.

4.3.4 Копуле остајања

Копуле су глаголи непотпуне предикације и они без допуне немају значење. Управо та допуна је она која нам помаже да овакве глаголе преведемо на одговарајући начин. Такође за копуле се може рећи да се понашају као „празне морфеме без значења које само повезују субјекат и допуну“.

This new release contains a registry cleaner which will help you keep your registry compact.

Ово ново издање садржи чистач регистра, што вам помаже у очувању компактности регистра.

The Parliament did not manage to reach a unanimous decision, even though they stayed until half past eleven.

Парламент није успио да донесе једногласну одлуку иако су остали до пола једанаест.

The commander came forward and helds out the flag in order to...
Командант је иступио и пружио заставу као би се она...

If I don't act now, we won't be able to stay ...

Ако не реагујемо сада, нећемо моћи да останемо ...

1. We will not rest until we bring the criminals to justice.
2. Нећемо имати мира док их не приведемо пред лице правде.

The presidential election was held on April 26.

Предсједнички избори су одржани 26. априла.

They hold no authority over that particular patch of the city.

Они немају власт над тим дијелом града.

Најчешћи преводни еквиваленти ових глагола су били облици у истом времену као и у оригиналу. То се дешавало у око 60% ситуација. Ови глаголу су најчешће били у облику енглеског простог прошлог времена или трајног прошлог времена, што је у великој мјери одговарало нашем претериту. Они који су се налазили у облику енглеског простог садашњег времена, у нашем језику су превођени садашњим временом. Такође је занимљиво да у доста случајева ови глагол *to hold* није преведен типичним преводним еквивалентом код нас, већ имамо разноврсне преводе (нпр. Вјероватно прва асоцијација на помен глагола *to hold* је држати, али у том облику се овај глагол веома ријетко јављао у посматраном корпусу). За разлику од очекивања, посматрање корпуса

(прије свега *Public Administration Reform Strategy in Montenegro*, и Уговор о пословно-техничкој сарадњи) је показало да су копуле остајања (о овом случају *HOLD, KEEP, REST, STAY*) релативно ријетко употријебљене.

4.3.5 Прошли партицип

Прошли партицип је граматичка категорија која се употребљава као глаголски или адјективални облик. Његова доминантна функција је образовање перфективних и пасивних облика. Током посматрања корпуса текста нађене су следеће категорије у којима се јавио прошли партицип.

Прошли партицип наспрам именице

The accused said it was not his fault.

Оптуженик је рекао да то није његова грешка.

Прошли партицип (у неким случајевима као остатак редуковане релативне клаузе) наспрам адјективала

The stone was so worn that it was hard to make out what was engraved there.

Камен је био толико излизан да је било тешко прочитати шта је на њему написано.

The third was handwritten and crumpled.

Трећи је био руком написан и згужван.

Прошли партицип наспрам придјевске фразе

They were hidden by thousands of long-gone students.

Њих су сакривали давно нестали студенти.

Прошли партицип наспрам глагола у пасиву

All done in the name of our protection to prevent ...

Све ово је било урађено да би смо нас заштитили и спријечили ...

Прошли партицип наспрам прилога

1. "Well done, was the only answer we could get.
2. „Одлично, је био једини одговор који смо могли добити

Из посматраног корпуса се такође дало закључити да се прошли партицип, у енглеским реченицама, у готово 90%, јавља као дио неког од перфективних времена или пасивних облика. Ти облици су најчешће код нас превођени претеритом или адјективалом. Пасивна форма у нашем језику се задржавала само тамо гдје је она звучала природно. У већини случајева, тамо гдје је прошли партицип градио пасив, он је био преведен активним обликом. То се дешавало јер је активна форма много природнија у нашем језику, док је енглеском језику ближи пасив. У случајевима гдје смо имали конструкције прошлог партиципа као адјективала са именицом нпр. *worn cases*, у огромном броју случаја то се у нашем језику јављало као придјев + именица (стара кутија).

4.3.5.1 Садашњи перфекат (Present Perfect)

Садашњи перфект је глаголски облик који настаје од одговарајућег облика презента помоћног глагола *HAVE* и прошлог партципа лексичког глагола:

HAVE + PAST PARTICIPLE

(HAVE, HAS)

У свом основном значењу ово вријеме се користи да би се направила веза између неодређене прошлости и текуће релевантности, али постоје и друге употребе овог глаголског облика.

Примјери:

Садашњи перфекат насупрот перфекту:

A little boy of about seven and a little girl who has been a year older, were playing.

Дјечак од око седам и дјевојчица која је била годину старија, су се играли.

Haven't we found new strength in the process that defines..

Зар нијесмо нашли нову снагу у процесу који дефинише...

But these people have never considered any other alternative...

Али ови људи никада нијесу разматрали неку другу алтернативу...

We have just received special orders ...

Управо смо примили специјална наређења...

A. The soul of every human being that has ever existed.

B. Душа сваког људског бића које је икада постојало.

Модал + Садашњи перфекат насупрот ДА клаузи:

It is implied we must have had a terrible misunderstanding.

Подразумијева се да се нијесмо разумјели.

We might have used more harmful deterrents.

Могли смо да користимо опсаније оружије.

Модал + Садашњи перфекат насупрот потенцијала:

That would have been too shameful.

То би било превише срамотно.

Садашњи перфекат насупрот будућем времену:

When the two parliaments have had sufficient level of information, they will be obliged to notify...

Када два парламента буду имали довољно података, обавијестиће ...

Као што се може закључити из горњих примјера, прошли партицип се најчешће преводи перфектом, мада постоје и неке друге опције у вези са преводним еквивалентима. Учесталост перфекта као преводног еквивалента је у готово 90% случајева. Веома често се наилазило на конструкције модала + прошли перфекат, што је веома карактеристично за енглески језик. У нашем језику готово сваки пут се као преводни еквивалент јављала ДА клауза. Случајеви са потенцијалом и будућим временом, као преводним еквивалентима у посматраним дјелима, су били веома ријетки.

Садашњи перфект (Present Perfect)

Садашњ перфект је глаголски облик који настаје од одговарајућег облика презента помоћног глагола *HAVE* и прошлог партиципа лексичког глагола:

HAVE + PAST PARTICIPLE

(HAVE, HAS)

У свом основном значењу ово вријеме се користи да би се направила веза између неодређене прошлости и текуће релевантности, али постоје и друге употребе овог глаголског облика.

4.3.6 Can, Could

Can, could припадају групи модалних глагола. Модални глагол *can* може да носи седам значења, и то: значење способности, воље, карактеристике, могућности, дозволе, чулног осјећаја, нестрпљења/збуњености. У посматраном корпусу, нађени су сљедећи примјери *can, could* и њихових преводних еквивалената.

Примјери:

Can, could наспрам глагола *моћи* (у презенту, перфекту)

Employee under the age of 18 can not work in a job position with prevailing hard physical labour...

Запослени који има испод 18 година не може да обавља посао који је прије свега састоји од тешког физичког рада...

... documentation, the power aggregate No 8 could generate approximately 20 GWh of electric...

... документацији, генератор број 8 је могао да произведе отприлике 20 ГWh електричне...

Can без преводног еквивалента у нашем језику

We can now see all the benefits of our mission ...

Видимо сада све користи наше мисије...

From the above table we can see that before Courts in Montenegro, due to the criminal offence..

Из горње табеле видимо да се пред судовима у Црној Гори, због кривичног прекршаја...

Could наспрам потенцијала са будућом референцом.

We could speak about all the possibilities tomorrow.

Могли би причати о свим могућностима сјутра.

Општи закључак који се може извући јесте да се модални глаголи *can*, *could* у нашем корпусу, у огромној већини случајева, преводе се глаголом *моћи*, а у једном мањем дијелу превода су се као преводни еквиваленти налазили и *знати*, *умјети*. У својим значењима, која су горе наведена, ови се облици доминантно преводе глаголом *моћи*, осим у значењу способност гдје је најчешћи преводни еквивалент *знати* или *умјети*. Што се тиче преводне разлике између *can/could*, текстови са „лакшом“ тематиком и граматичким конструкцијама (нпр. преписке, изјаве за штампу), региструју ову разлику као однос *can* - садашњост, *could* - прошлост у скоро сваком случају. Готово је немогуће наћи такав тип текста гдје је облик *could* употријебљен у виду садашње референце. Са друге стране, у текстовима са сложенијом граматиком ова разлика се у већем степену губи и преноси се на поље *can* - мање формално, мање културно, *could* - формалније, културније.

4.3.7 Пасив

Радња у енглеском језику се може изразити на два начина, кроз актив или пасив. У енглеском језику пасивне конструкције су, у највећем броју случајева, веома природне и сходно томе веома често се користе. У

нашем језику такве конструкције нису претјерано честе, јер пасив зна често да звучи прилично чудно. У посматраном корпусу, пронађени су следећи примјери односа пасива између енглеског и српског језика.

Примјери:

Прошли прогресив пасива наспрам перфекта (пасив)

...until the declaration of independence of Montenegro, were being carried out by the Accreditation body of Serbia and Montenegro.

...до проглашења независности, ти послови су били обављани од стране Акредитационог тијела Србије и Црне Горе.

Прошли перфект пасива наспрам перфекта (пасив)

This goal was to be fulfilled in the period of three years after it had been adopted.

Овај циљ је требало испунити током периода од три године након што је био усвојен.

Садашњи перфект пасива наспрам перфекта (актив)

The project has been carried out by the KulturKontakt - Austria, in cooperation with the Ministry of Education and Science

Овај пројекат је спровела агенција КултурКонтакт - Аустрија, у сарадњи са Министарством просвјете и науке.

Претерит пасива наспрам перфекта (пасив)

And when the second part was done refugees could exercise basic rights.

А када је био урађен и други дио, избјеглице су могле да користе своја основна права.

Из ових примјера се може видјети да је у енглеском језику пасив као граматичка конструкција веома заступљен. Њиме се на једноставан и лијеп начи изражава мисао. У нашем језику то није случај са пасивом и у већини случајева се избјегава. У већини случајева, прошли прогресив пасива, прошли перфект пасива, садашњи перфект пасива и претерит пасива се преводе нашим обликом перфекта. Нарочито често се енглеске пасивне реченице преводе нашом активном реченицом у случајевима гдје енглеска реченица садржи форму *BY + вршилац радње*. У тим случајевима много је лакше реченицу превест активом јер директан превод енглеске форме *BY + вршилац радње* (од стране + вршилац радње) у нашем језику дјелује већином непривлачно.

5 Закључак

Као што се дало закључити из нашег рада једна од основних премиси овог доктората јест жеља да се у контрастивну анализу унесу елементи корпусне и речунарске лингвистике. На тај начин смо спојили провјерени квалитет контрастивног метода са модерним карактеристикама корпусне и рачунарске анализе. Оне су, прије свега, донијела квалитативни и квантитативни помак у односу на истраживања која су се базирала искључиво на контрастивном приступу. Кроз многобројне примјере, надамо се, показали смо које су то основне карактеристике административног стила писања који је доминантан стил када се ради о међудржавним односима или односима између ЕУ и Црне Горе. Корпусна и рачунарска лингвистика, све више и више, добијају на свом значају, али је неопходно напоменути да и оне као такве имају својих мана, а једна од најочигледнијих јесте да нам могу пружити резултате, али нам не могу рећи какав је квалитет тих резултата у смислу њихове, нпр. истинитости, али, и поред тога, приказали смо да корпус може бити веома корисно средство у лингвистичким анализама. Анализирали смо разлике у административном дискурсу текстова на српском и енглеском језику које су биле условљене не само формалним граматичко-структуралним разликама између ова два језика, већ и културолошким и системско-правним разликама које се пресликавају и утичу на њих.

Поред тога, можемо закључити да је корпус ипак поузданије оруђе у лингвистичким анализама него што је то интуиција изворног говорника, што смо показали на примјерима у вези са ријечима *здравство* и *здравствени*, у српском језику и *health* у енглеском језику, а такође смо видјели да се неке ријечи понављају свега пар пута или се уопште не

налазе у корпусу иако вјероватно већина извоних говорника не би тако оцијенила на основу своје интуиције.

Кроз контрастирање лексичких и реченичних елемената из српског језика и њима одговарајућих преводачких еквивалената у енглеском језику доказали смо предности и мане корпусно-контрастивног приступа, анализирали преводне еквиваленте и граматичке јединице, а такође смо анализирали и врсту дискурса који се користи у оваквим типовима административних докумената (дискурсна обиљежја) и да одговорили на питања која се тичу сличности или разлика између ова два типа регистра у српском и енглеском језику.

Надамо се да смо кроз рад успјели да покажемо да је ова врста регистра веома осјетљива на „слободнији“ превод, јер обично подразумијева текстове чија је примарна сврха да буду информативног типа, а та информација мора бити прецизно пренесена у циљни језик. Ова „осјетљивост“ ће нарочито бити видљива у текстовима облигационог типа гдје једна ријеч може промијенити структуру облигационог односа. Слободан превод је нешто што у оваквим типовима текстова треба избјегавати већ је пожељније држати се униформних фраза које су усвојене као стандард. На овај начин превод постаје јаснији и за примаоца поруке.

Одговорили смо на питање да ли је корпусна лингвистика методологија или самостална наука коришћењем једног релативно новог „алата“ *WebCorp-a* и на тај начин показали потенцијал интернет алату у лингвистичким, дискурским анализама.

И на самом крају, направили смо глосар од преко 10.000 термина који су најдиректније везани за документа ЕУ што може бити од помоћи преводиоцима који се баве превођењем правних аката из домена законодавства ЕУ.

Надамо се читаоци овај рад схватити као промовисање контрастивне лингвистике као дијела лингвистике која, по мишљењу аутора овог текста, има највећи потенцијал када се ради о откривању сличности и разлика међу језицима; у нашем конкретном случају говоримо о српском и енглеском. Као напор за дубљим разумијевањем преводних еквивалената у оквиру административног дискурса. Преводац мора да потпуно разумије дубље значење административне лексике која потиче са простора који имају другачије културолошке и законодавне норме. Боље разумијевање ове врсте дискурса води ка бољим преводима, који доприносе бољем разумијевању међу стручњацима који имају додирних тачака са овом врстом докумената.

Овај рад би био незамислив да није постојала група текстова које се назива корпус и, надамо се, да смо доказали да преко паралелних корпуса имамо бољи увид у језике који се проучавају, да се корпуси могу користити за сврхе упоредних анализа, као и да се боље проуче типолошке и културолошке разлике, али и да се нађу универзални елементи, да нам корпуси омогућавају да се уоче разлике између оригинала и превода, тј. између изворног и преведеног текста.

Такође, корпуси су омогућили да се језик испитује на начине који су раније били недоступни. Примјери ових начина укључују могућности квантитативних лингвистичких истраживања, као што је рачунање прецизних листа учесталости за ријечи или неке друге граматичке категорије, гдје говорник може да тестира употребу наспрам норме. Ово нам може помоћи у идентификовању регистра, или карактеристика стила аутора, што може бити од користи за књижевну стилистику, форензичку лингвистику и превођења.

Корпус су промијенили начин на који гледамо на језик и обогатили одређене професије. Корпуси посједују огромне колекције текстова и теже да окупе све што се у језику јавља кроз његову употребу. На основу тих

текстова могуће је извести одређене закључке, напревити листе учесталости и колокација које су од велике помоћи особама којима су упрето те ријечи у процесу учења неопходне. На овај начин могу да настану електронски рјечници који својим корисницима омогућавају не само стандардне карактеристике као што су дефиниције и пар примјера већ омогућавају да се одређена ријеч посматра практично у свим могућим контекстима. Ученици могу и да сами направе своје корпусе, на основу већег корпуса, како би проучавали неке карактеристике језика на основу својих „мини“ корпуса.

Корпуси су учинили наша истраживања лакшим. Са корпусом, преводилац може врло брзо да открије који све преводни еквиваленти постоје за ријеч која га интересује. Може такође да види како су ту рије преводили колеге преводиоци у другим документима и да ли постоје неке промјене у преводним еквивалентима током протока времена. Било ко, уз елементарно познавање функционисања рачунара и рачунарски програма за ову сврху, да проучава, нпр. помоћне глаголе, јер управо ти програми омогућавају да а подаци добију знатно брже, него да се исти поступак обавља „пјешке“. Корпус је дакле веома богат извор за проучавање језика. Као директна посљедица овог нове идеје се јављају а старе бивају подвргнуте поновном провјеравању и по потреби додавању или измјенама. Дакле, корпусну лингвистику и електронске корпусе треба схватити као вид природне еволуције језика. Корпусна лингвистика налази своју примјену и у другим областима науке. Па се тако корпусна истраживања користе за:

- Прикушљање емпиријских података
- Испитивање говора
- Проучавање лексике
- Проучавање граматике

- Проучавање семантике
- Проучавање прагматике и анализе дискурса
- Социолингвистичка истраживања
- Анализе стилистике и текст лингвистике
- Учење језика
- Анализе историјске лингвистике, дијалектологије, друштвене психологије
- Културолошке студије, итд.

6 Литература

1. Agresti, Alan. 2002. *Categorical Data Analysis*. Hoboken, NJ: Wiley.
2. Aijmer, Karin. 1984. *Sort of and kind of* in English conversation. *Studia Linguistica* 38/2: 118–128.
3. — —. 1986. Discourse variation and hedging. In *Corpus Linguistics II: New Studies in the Analysis and Exploitation of Computer Corpora*, ed. Jan Aarts and Willem Meijs, 1–18. Amsterdam: Rodopi.
4. Arppe, Antti and Juhani Järviö. 2007. Every method counts: combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory* 3/2: 131–159.
5. Aston, Guy (ed.). 2001. *Learning With Corpora*. Bologna: CLUEB.
6. Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
7. Baayen, R. Harald, Richard Piepenbrock, and Leon Gulikers. 1995. *The CELEX Lexical Database (Release 2)*. Philadelphia, PA: Linguistic Data Consortium.
8. Baker, Paul, Andrew Hardie, and Tony McEnery. 2006. *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.
9. Barnbrook, Geoff. 1996. *Language and Computers*. Edinburgh: Edinburgh University Press.
10. Baron, Faye and Graeme Hirst. 2003. Collocations as cues to semantic orientation. Unpublished MS. Downloaded from <http://citeseer.ist.psu.edu/rd/0%2C683844%2C1%2C0.25%2CDownload/http%3AqSqqSftp.cs.toronto.eduqSqpubqSqghqSqBaronqPqHirst-2003.pdf> on September 20, 2006.
11. Baroni, Marco and Silvia Bernardini (eds.). 2006. *Wacky! Working papers on the web as corpus*. Bologna: GEDIT. [ISBN 88-6027-004-9] Downloadable from <http://wackybook.sslmit.unibo.it/>
12. Bassnett, Susan, *Translation Studies*, London, Routledge, 1999.

13. Beal, Joan C., Karen P. Corrigan, and Hermann L. Moisl (eds.). 2007a. *Creating and Digitizing Language Corpora: Synchronic Databases Vol. 1*. Basingstoke: Palgrave Macmillan.
14. — —. 2007b. *Creating and Digitizing Language Corpora: Diachronic Databases Vol. 2*. Basingstoke: Palgrave Macmillan.
15. Bednarek, Monika. 2008. Semantic preference and semantic prosody re-examined. *Corpus Linguistics and Linguistic Theory* 4/2: 119–139.
16. Bell, Alan, Michelle L. Gregory, Jason M. Brenier, Daniel Jurafsky, Ayako Ikeno, and Cynthia Girand. 2002. Which predictability measures affect content word durations? *Proceedings of the Pronunciation Modeling and Lexicon Adaptation for Spoken Language Technology Workshop*, pp. 1–5.
17. Berg, Thomas. 2006. The internal structure of four-noun compounds in English and German. *Corpus Linguistics and Linguistic Theory* 2/2: 197–231.
18. Biber, Douglas. 1988. *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.
19. — —. 1990. Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and Linguistic Computing* 5/4: 257–269.
20. — —. 1993. Representativeness in corpus design. *Literary and Linguistic Computing* 8/4: 243–257.
21. Biber, Douglas, Susan Conrad, and Randi Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
22. Bod, Rens, Jennifer Hay, and Stefanie Jannedy (eds.). 2003. *Probabilistic Linguistics*. Cambridge, MA: MIT Press.
23. Bortz, Jürgen. 2005. *Statistik für Human- und Sozialwissenschaftler*. 6th edn. Heidelberg: Springer.
24. Bowker, Lynne and Jennifer Pearson. 2002. *Working with Specialized Language: A Practical Guide to Using Corpora*. London and New York: Routledge.

25. Brinkmann, Ursula. 1997. *The Locative Alternation in German: Its Structure and Acquisition*. Amsterdam, Philadelphia: John Benjamins.
26. *British National Corpus*, version 2 (BNC World). 2001. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <<http://www.natcorp.ox.ac.uk/>>
27. Brown, Roger. 1973. *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.
28. Bugarski 1972: Ranko Bugarski, *Jezik i lingvistika*, Nolit, Beograd.
29. Bugarski 1986. b: Ranko Bugarski, *Terminologija kontrastivne lingvistike*, u: Kontrastivna jezička istraživanja, III simpozijum (Novi Sad, 6. i 7. decembar 1985), Zbornik radova, Univerzitet u Novom Sadu, Filozofski fakultet, Novi Sad, 383–390.
30. Bugarski 1990: Ranko Bugarski, *Integralna kontrastivna analiza*, u: Kontrastivna jezička istraživanja, IV simpozijum (Novi Sad, 8. i 9. decembar 1989), Zbornik radova, Filozofski fakultet, Novi Sad, 58–62.
31. Bugarski, Ranko. 2003. *Žargon*, Beograd: Čigoja štampa.
32. Bybee, Joan and Joanne Scheibman. 1999. The effect of usage on degrees of constituency: the reduction of *don't* in English. *Linguistics* 37/4: 575–596.
33. Church, Kenneth W. and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics* 16/1: 22–29.
34. Church, Kenneth Ward, William Gale, Patrick Hanks, and Donald Hindle. 1991. Using statistics in lexical analysis. In *Lexical Acquisition: Exploiting Online Resources to Build a Lexicon*, ed. Uri Zernik, 115–164. Hillsdale, NJ: Lawrence Erlbaum.
35. Church, Kenneth Ward, William Gale, Patrick Hanks, Donald Hindle, and Rosamund Moon. 1994. Lexical substitutability. In *Computational Approaches to the Lexicon*, ed. Beryl T. Sue Atkins and Antonio Zampolli, 153–177. Oxford: Oxford University Press.

36. Clark, Herbert H. and Jean E. Fox Tree. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84/1: 73–111.
37. Conrad, Susan. 2000. Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly* 34: 548–560.
38. Crawley, Michael J. 2002. *Statistical Computings: An Introduction to Data Analysis Using S-Plus*. Chichester: John Wiley and Sons.
39. — — . 2005. *Statistics: An Introduction using R*. Chichester: John Wiley and Sons.
40. — — . 2007. *The R Book*. Chichester: John Wiley and Sons.
41. Cvetana K., Duško V., *An Aligned English-Serbian Corpus*, In: ELLSIIR Proceedings (English Language and Literature Studies: Image, Identity, Reality), Volume I, Belgrade, 4-6 December 2009, eds. N. Tomović & J. Vujić, pp. 495-508, Faculty of Philology, University of Belgrade, ISBN 978-86-6153-005-0, 2011.
42. Dalbey, John. 2003. Pseudocode standard. Downloaded from http://www.csc.calpoly.edu/~jdalbey/SWE/pdl_std.html on November 26, 2006.
43. Dalgaard, Peter. 2002. *Introductory Statistics with R*. Berlin, Heidelberg: Springer.
44. Danielsson, Pernilla. 2004. Simple Perl programming for corpus work. In *How to Use Corpora in Language Teaching*, ed. John McHardy Sinclair, 225–246. Amsterdam and Philadelphia: John Benjamins.
45. Diessel, Holger. 1996. Processing factors of pre- and postposed adverbial clauses. *Proceedings of the 22nd Annual Meeting of the Berkeley Linguistics Society*, pp. 71–82.
46. — — . 2001. The ordering distribution of main and adverbial clauses: a typological study. *Language* 77/3: 343–365.
47. — — . 2005. Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 43/3: 449–470.

48. Diessel, Holger and Michael Tomasello. 2005. Particle placement in early child language: a multifactorial analysis. *Corpus Linguistics and Linguistic Theory* 1/1: 89–112.
49. Dilts, Philip and John Newman. 2006. A note on quantifying “good” and “bad” prosodies. *Corpus Linguistics and Linguistic Theory* 2/2: 233–242.
50. Divjak, Dagmar and Stefan Th. Gries. 2006. Ways of trying in Russian: clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2/1: 23–60.
51. Đorđević, Radmila. 2004. Uvod u kontrastiranje jezika, Filološki fakultet Univerziteta u Beogradu, Beograd.
52. Ellis, Nick C. 2002a. Frequency effects in language processing and acquisition. *Studies in Second Language Acquisition* 24/2: 143–188.
53. — — . 2002b. Reflections on frequency effects in language processing. *Studies in Second Language Acquisition* 24/2: 297–339.
54. Estival, Dominique. 1985. Syntactic priming of the passive in English. *Text* 5/1: 7–22.
55. Evert, Stefan. 2004. The statistics of word cooccurrences: word pairs and collocations. Ph.D. diss., University of Stuttgart.
56. Filipović 1968: Rudolf Filipović, Zašto „kontrastivna analiza“?, *Živi jezici* X/1–4, Beograd, 1–5.
57. Forta, Ben. 2004. *Sams Teach Yourself Regular Expressions in 10 Minutes*. Indianapolis, IN: Sams Publishing.
58. Fox, Gwyneth. 1998. Using corpus data in the classroom. In *Materials Development in Language Teaching*, ed. Brian Tomlinson, 25–43. Cambridge: Cambridge University Press.
59. Fox Tree, Jean E. and Herbert H. Clark. 1997. Pronouncing *the* as *thee* to signal problems in speaking. *Cognition* 62/2: 151–167.
60. Friedl, Jeffrey E.F. 2006. *Mastering Regular Expressions*. 3rd edn. Cambridge, MA: O'Reilly.

61. Gahl, Susanne and Susan Marie Garnsey. 2004. Knowledge of grammar, knowledge of usage: syntactic probabilities affect pronunciation variation. *Language* 80/4: 748–775.
62. Givón, Talmy. 1992. The grammar of referential coherence as mental processing instructions. *Linguistics* 30/1: 5–55.
63. Good, Nathan A. 2005. *Regular Expression Recipes: A Problem-solution Approach*. Berkeley, CA: Apress.
64. Good, Philip I. 2005. *Introduction to Statistics through Resampling Methods and R/S-Plus*. Hoboken, NJ: John Wiley and Sons.
65. Gorjanc, Vojko. 2005. *Uvod u korpusno jezikoslovlje*. Domžale: Zbirka Zreńa/Izolit.
66. Gries, Stefan Th. 2001. A corpus-linguistic analysis of *-ic* and *-ical* adjectives. *ICAME Journal* 25: 65–108.
67. ——. 2002a. Evidence in linguistics: three approaches to genitives in English. In *LACUS Forum XXVIII: what constitutes evidence in linguistics?*, ed. Ruth M. Brend, William J. Sullivan, and Arle R. Lommel, 17–31. Fullerton, CA: LACUS.
68. ——. 2002b. Preposition stranding in English: predicting speakers' behaviour. In *Proceedings of the Western Conference on Linguistics. Vol. 12*, ed. Vida Samiiian, 230–241. California State University, Fresno, CA.
69. ——. 2003a. *Multifactorial Analysis in Corpus Linguistics: a study of particle placement*. London and New York: Continuum Press.
70. ——. 2003b. Testing the sub-test: a collocational-overlap analysis of English *-ic* and *-ical* adjectives. *International Journal of Corpus Linguistics* 8/1: 31–61.
71. ——. 2004a. *Coll.analysis 3*. A program for R for Windows 2.x.
72. ——. 2004b. Isn't that fantabulous? How similarity motivates intentional morphological blends in English. In *Language, Culture, and Mind*, ed. Michel Achard and Suzanne Kemmer, 415–428. Stanford, CA: CSLI.
73. ——. 2005a. Syntactic priming: a corpus-based approach. *Journal of Psycholinguistic Research* 34/4: 365–399.

74. — —. 2005b. Null-hypothesis significance testing of word frequencies: a follow-up on Kilgarriff. *Corpus Linguistics and Linguistic Theory* 1/2: 277–294.
75. — —. 2006a. Some proposals towards more rigorous corpus linguistics. *Zeitschrift für Anglistik und Amerikanistik* 54/2: 191–202.
76. — —. 2006b. Cognitive determinants of subtractive word-formation processes: a corpus-based perspective. *Cognitive Linguistics* 17/4: 535–558.
77. — —. 2007a. Exploring variability within and between corpora: some methodological considerations. *Corpora* 1/2: 109–151.
78. — —. 2007b. Towards a more useful measure of dispersion in corpus data. Paper presented at Corpus Linguistics.
79. — —. 2008a. Corpus-based methods in analyses of SLA data. In *Handbook of Cognitive Linguistics and Second Language Acquisition*, ed. Peter Robinson and Nick Ellis, 406–431. New York: Routledge.
80. — —. 2008b. Phraseology and linguistic theory: a brief survey. In *Phraseology: An Interdisciplinary Perspective*, ed. Sylviane Granger and Fanny Meunier, 3–25. Amsterdam and Philadelphia: John Benjamins.
81. — —. 2008c. Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics*.
82. — —. to appear. *Statistics for Linguistics Using R: A Practical Introduction*. Berlin, New York: Mouton de Gruyter.
83. Gries, Stefan Th. and Caroline V. David. 2007. This is kind of/sort of interesting: variation in hedging in English. In *Proceedings of ICAME 27*, ed. Päivi Pahta, Irma Taavitsainen, and Terttu Nevalainen. eVARIENG: Methodological Interfaces, University of Helsinki.
84. Gries, Stefan Th. and Dagmar Divjak. to appear. Behavioral profiles: a corpus-based approach towards cognitive semantic analysis. In *New Directions in Cognitive Linguistics*, ed. Vyvyan Evans and Stephanie Pourcel. Amsterdam and Philadelphia: John Benjamins.

85. Gries, Stefan Th., Beate Hampe, and Doris Schönefeld. 2005. Converging evidence: bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics* 16/4: 635–676.
86. — —. to appear. Converging evidence II: more on the association of verbs and constructions. In *Empirical and Experimental Methods in Cognitive/Functional Research*, ed. John Newman and Sally Rice. Stanford, CA: CSLI.
87. Gries, Stefan Th. and Anatol Stefanowitsch. 2004a. Extending collocation analysis: a corpus-based perspective on “alternations”. *International Journal of Corpus Linguistics* 9/1: 97–129.
88. — —. 2004b. Co-varying collexemes in the *into*-causative. In *Language, Culture, and Mind*, ed. Michel Achard and Suzanne Kemmer, 225–236. Stanford, CA: CSLI.
89. — — (eds.). 2006. *Corpora in Cognitive Linguistics: Corpus-Based Approaches to Syntax and Lexis*. Berlin, New York: Mouton de Gruyter.
90. Gries, Stefan Th. and Sabine Stoll. to appear. Finding developmental groups in acquisition data: variability-based neighbor clustering, *Journal of Quantitative Linguistics* 16/3.
91. Halliday, M. A. K. et al, *Lexicology and Corpus Linguistics*, Continuum, London, 2005.
92. de Haan, Pieter. 1992. The optimum corpus sample size. In *New Directions in English Language Corpora*, ed. Gerhard Leitner, 3–19. Berlin and New York: Mouton de Gruyter.
93. Hearst, Marti A. 1997. TextTiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics* 23/1: 33–64.
94. Heatley, Alex, Paul I.S. Nation, and Avery Coxhead. 2002. RANGE and FREQUENCY programs.
<http://www.vuw.ac.nz/lals/staff/Paul_Nation>

95. Höche, Silke. 2009. *Cognate object constructions in English: a cognitive linguistic account*. Tübingen: Gunter Narr.
96. Hoffmann, Thomas. 2006. Corpora and introspection as corroborating evidence: the case of preposition placement in English relative clauses. *Corpus Linguistics and Linguistic Theory* 2/2: 165–195.
97. Hofland, Knud and Stig Johansson. 1982. *Word Frequencies in British and American English*. Bergen: Norwegian Computing Centre for the Humanities/London: Longman.
98. Hundt, Marianne, Nadja Nesselhauf, and Carolin Biewer (eds.). 2007. *Corpus Linguistics and the Web*. Amsterdam, New York, NY: Rodopi.
99. Hunston, Susan. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
100. Hunston, Susan and Gill Francis. 2000. *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam and Philadelphia: John Benjamins.
101. Ionescu, D., *Translation Theory and Practice*, Bucuresti, Editura Universal Dalsi, 2000.
102. Jackendoff, Ray. 1997. Twistin' the night away. *Language* 73/3: 534–559.
103. Johansson, Stig. 1980. The LOB corpus of British English texts: presentation and comments. *ACCL Journal* 1: 25–36.
104. Johansson, Stig and Knut Hofland. 1989. *Frequency Analysis of English Vocabulary and Grammar, Based on the LOB Corpus*. Oxford: Clarendon Press.
105. Johnson, Keith. 2008. *Quantitative Methods in Linguistics*. Malden, MA and Oxford: Blackwell.
106. Jones, Steven. 2002. *Antonymy: A Corpus-based Perspective*. London and New York: Routledge.
107. Joseph, Brian D. 2004. On change in Language and change in language. *Language* 80/3: 381–383.
108. Juola, Patrick. 2008. *Authorship Attribution*. Hanover, MA and Delft: Now Publishers.

109. Jurafsky, Daniel, Alan Bell, Michelle Gregory and William D. Raymond. 2000. Probabilistic relations between words: Evidence from reduction in lexical production. In *Frequency and the Emergence of Linguistic Structure*, ed. Joan Bybee and Paul Hopper, 229–254. Amsterdam and Philadelphia: John Benjamins.
110. Jurafsky, Daniel and James H. Martin. 2000. *Speech and Language Processing: [...]*. Upper Saddle River, NJ: Prentice Hall.
111. Justeson, John S. and Slava M. Katz. 1991. Co-occurrence of antonymous adjectives and their contexts. *Computational Linguistics* 17/1: 1–19.
112. — —. 1995. Principled disambiguation: discriminating adjective senses with modified nouns. *Computational Linguistics* 21/1: 1–27.
113. Kaunisto, Mark. 1999. *Electric/electrical* and *classic/classical*: variation between the suffixes *-ic* and *-ical*. *English Studies* 80/4: 343–370.
114. — —. 2001. Nobility in the history of adjectives ending in *-ic* and *-ical*. In *LACUS Forum XXVII: speaking and comprehending*, ed. Ruth Brend, Alan K. Melby, and Arle R. Lommel, 35–46. Fullerton, CA: LACUS.
115. Kay, Paul and Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: the What's X doing Y? construction. *Language* 75/1: 1–33.
116. Keller, Frank. 2000. Gradience in grammar: experimental and computational aspects of degrees of grammaticality. Ph. D. diss., University of Edinburgh.
117. Kennedy, Graeme. 1998. *An Introduction to Corpus Linguistics*. London and New York: Longman.
118. Kilgarriff, Adam. 2001. Comparing corpora. *International Journal of Corpus Linguistics* 6/1: 1–37.
119. — —. 2005. Language is never, ever, ever random. *Corpus Linguistics and Linguistic Theory* 1/2: 263–276.

120. Kilgarriff, Adam and David Tugwell. 2001. Word Sketch: extraction and display of significant collocations for lexicography. *Proceedings of the Workshop on Collocations*, ACL, pp. 32–38.
121. Kolari, Pranam and Anupam Joshi. 2004. Web mining: research and practice. *IEEE Computing in Science and Engineering* 6/4: 49–53.
122. Lakić, I., "Genre Analysis of Article Introductions in Economics", *Asp. La revue du GERAS*, No. 15-18, 1997, Bordeaux, France, pp. 409 - 426, ISSN 1246 8185
123. Lakić, I., "Supporting Classroom Activities through the Use of Songs", *Novelty*, Vol. 7, No. 2, British Council/IATEFL, Hungary, May 2000, pp. 51 - 61, ISSN 1218-537x
124. Lakić I., (2009) "Modeli analize diskursa novinskih članaka", UDK 811.111'42; 821.111-92, Riječ, br. 2, Nikšić: Filozofski fakultet, Institut za jezik i književnost, str. 91 - 108., ISSN 0354-6039, UDK 80+82(05
125. Lakić, I., (2010) "Analysing Genre: Research Article Introductions in Economics", *JoLIE, Journal of Linguistic and Intercultural Education*, Vol. 3/2010, Editura Aeternitas, Alba Iulia, str. 83 - 99, ISSN 2065-6599
126. Laufer, Batia. 2005. Lexical Frequency Profiles: from Monte Carlo to the real world. A response to Meara. *Applied Linguistics* 26/4: 581–587.
127. Laufer, Batia and Paul Nation. 1995. Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics* 16/3: 307–322.
128. Leech, Geoffrey N. 1992. Corpora and theories of linguistic performance. In *Directions in Corpus Linguistics*, ed. Jan Svartvik, 105–122. Berlin and New York: Mouton de Gruyter.
129. — — . 1993. Corpus annotation schemes. *Literary and Linguistic Computing* 8/4: 275–281.
130. Leech, Geoffrey and Roger Fallon. 1992. Computer corpora: what do they tell us about culture? *ICAME Journal* 16: 29–50.
131. Leech, G., *Adding Linguistic Annotation in Developing Linguistic Corpora: a Guide to Good Practice*, ed. M. Wynne, Oxbow Books, Oxford, 2005.

132. Levin, Beth. 1993. *Verb Classes and Alternations: A Preliminary Investigation*. Chicago, IL: University of Chicago Press.
133. Lewis, Michael (ed.). 2000. *Teaching Collocation: Further Developments in the Lexical Approach*. Hove: Language Teaching Publications.
134. Ligges, Uwe. 2005. *Programmieren mit R*. Berlin and Heidelberg: Springer.
135. Maindonald, John and John Braun. 2003. *Data Analysis and Graphics Using R: An Example-based Approach*. Cambridge: Cambridge University Press.
136. Manning, Christopher D. and Hinrich Schütze. 2000. *Foundations of Statistical Natural Language Processing*. 2nd printing with corrections. Cambridge, MA and London: MIT Press.
137. Mason, Oliver. 1997. The weight of words: an investigation of lexical gravity. *Proceedings of PALC*, pp. 361–375.
138. — —. 1999. Parameters of collocation: the word in the centre of gravity. In *Corpora Galore: Analyses and Techniques in Describing English*, ed. John Kirk, 267–280. Amsterdam: Rodopi.
139. McEnery, T. and Andrew Wilson, *Corpus Linguistics*, Edinburgh University Press, Edinburgh, 2001.
140. McEnery, Tony and Andrew Wilson. 2003. *Corpus Linguistics*. 2nd edn. Edinburgh: Edinburgh University Press.
141. McEnery, Tony, Richard Xiao, and Yukio Tono. 2006. *Corpus-based Language Studies: An Advanced Resource book*. London and New York: Routledge.
142. Meara, Paul. 2005. Lexical frequency profiles: a Monte Carlo analysis. *Applied Linguistics* 26/1: 32–47.
143. Meyer, Charles F. 2002. *English Corpus Linguistics: An Introduction*. Cambridge: Cambridge University Press.
144. Mukherjee, Joybrato. 2007. Corpus linguistics and linguistic theory: general nouns and general issues. *International Journal of Corpus Linguistics* 12/1: 131–147.

145. Murrell, Paul. 2005. *R Graphics*. Boca Raton, FL: Chapman and Hall.
146. Neubert, A., *On the interface between translation theory and translation practice*, Translation studies in Scandinavia: proceedings from the Scandinavian Symposium on Translation Theory (SSOTT) II, Lund 14-15 June, 1985 ed. by Lars Wollin and Hans Lindquist, 25-36. Malmö: Liber/Gleerup, 1985.
147. Oakes, Michael P. 1998. *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.
148. ——. 2003. Text categorisation: automatic discrimination between US and UK English using the chi-square test and high ratio pairs. Downloaded from <<http://www.cet.sunderland.ac.uk/IR/oakesRL2003.pdf>> on April 5, 2006.
149. Oh, Sun-Young. 2000. *Actually* and *in fact* in American English: a data-based analysis. *English Language and Linguistics* 4/2: 243–268.
150. Okada, Sadayuki. 1999. On the function and distribution of the modifiers *respective* and *respectively*. *Linguistics* 37/5: 871–903.
151. O’Keeffe, Anna, Michael McCarthy, and Ronald Carter. 2007. *From Corpus to Classroom*. Cambridge: Cambridge University Press.
152. Panther, Klaus-Uwe and Linda L. Thornburg. 2002. The roles of metaphor and metonymy in English *-er* nominals. In *Metaphor and Metonymy in Comparison and Contrast*, ed. René Dirven and Ralf Pörings, 279–319. Berlin and New York: Mouton de Gruyter.
153. Partington, Alan. 1998. *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. Amsterdam and Philadelphia: John Benjamins.
154. Polovina, V. (1987): *Leksičko-semantička kohezija u razgovornom jeziku*, Beograd, Filološki fakultet, 335 pp.
155. Polovina, V. (1996): *Prilozi za kognitivnu lingvistiku*. Beograd, Filološki fakultet

156. Polovina, V. (1998): Pojam kontinuuma i njegova primena u lingvistici, *Južnoslovenski filolog*, LIV: 7-14.
157. Polovina, V. (1999): Semantika i tekst-lingvistika, Beograd, Čigoja, 220 str.
158. Polovina, V. (2010), Opšta i specifična obeležja kulture u neformalnoj konverzaciji, u: Vasić V. (ur.) *Diskurs i diskursi*, Zbornik u čast Svenke Savić, str. 185-195.
159. Pullum, Geoffrey K. 2006. Ungrammaticality, rarity, and corpus use. *Corpus Linguistics and Linguistic Theory* 3/1: 33–47.
160. Quirk, Randolph, Sidney Greenbaum, Geoffrey, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London, New York: Longman.
161. R Development Core Team. 2008. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <<http://www.R-project.org>>
162. Rayson, Paul E. and Roger Garside. 2000. Comparing corpora using frequency profiling. *Proceedings of the Workshop on Comparing Corpora*, pp. 1–6.
163. Rayson, Paul E., Geoffrey Leech, and Mary Hodges. 1997. Social differentiation in the use of English vocabulary: some analysis of the conversational component of the British National Corpus. *International Journal of Corpus Linguistics* 2/1: 133–152.
164. Rayson, Paul, Andrew Wilson, and Geoffrey N. Leech. 2001. Grammatical word class variation within the British National Corpus Sampler. In *New Frontiers of Corpus Research*, ed. Pam Peters, Peter Collins, and Adam Smith, 295–306. Amsterdam: Rodopi.
165. Roland, Douglas and Daniel Jurafsky. 1998. How verb subcategorization frequencies are affected by corpus choice. *Proceedings of the 17th International Conference on Computational Linguistics – Volume 2*, pp. 1,122–1,128.

166. Ryder, Mary Ellen. 1999. *Bankers and blue-chippers: an account of -er formations in present-day English*. *English Language and Linguistics* 3/2: 269–297.
167. Sankoff, David and Suzanne Laberge. 1978. Statistical dependence among successive occurrences of a variable in discourse. In *Linguistic Variation: Models and Methods*, ed. David Sankoff, 119–126. New York: Academic Press.
168. Schmid, Hans-Jörg. 2000. *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. Berlin and New York: Mouton de Gruyter.
169. Schönefeld, Doris. 1999. Corpus linguistics and cognitivism. *International Journal of Corpus Linguistics* 4/1:131–171.
170. Sharoff, Serge. 2006. Open-source corpora: using the net to fish for linguistic data. *International Journal of Corpus Linguistics* 11/4: 435–462.
171. Snell-Hornby, M., *Translation Studies: an Integrated Approach*. Amsterdam: Philadelphia Benjamin, 1988
172. Sinclair, John McHardy (ed.). 1987. *Looking Up: An Account of the COBUILD Project in Lexical Computing*. London: Collins.
173. Sinclair, John McHardy. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
174. — —. 2003. *Reading Concordances: An Introduction*. London: Pearson.
175. — —. 2004. *Trust the Text: language, corpus, and discourse*. London, New York: Routledge.
176. — —. *Corpus and Text - Basic Principles in Developing Linguistic Corpora: a Guide to Good Practice*, ed. M. Wynne, Oxbow Books, Oxford, 2005.
177. Spector, Phil. 2008. *Data Manipulation with R*. New York: Springer.
178. Stefanowitsch, Anatol. 2005. New York, Dayton (Ohio), and the raw frequency fallacy. *Corpus Linguistics and Linguistic Theory* 1/2: 295–301.
179. — —. Negative evidence and the raw frequency fallacy. *Corpus Linguistics and Linguistic Theory* 2/1: 61–77.

180. Stefanowitsch, Anatol and Stefan Th. Gries. 2003. Collocations: investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8/2: 209–243.
181. — —. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1/1: 1–43.
182. Stefanowitsch, Anatol and Stefan Th. Gries (eds.). 2006. *Corpus-based Approaches to Metaphor and Metonymy*. Berlin and New York: Mouton de Gruyter.
183. Steinberg, Danny D. 1993. *An Introduction to Psycholinguistics*. London and New York: Longman.
184. Stevens, Vance. 1991. Classroom concordancing: vocabulary materials derived from relevant, authentic text. *English for Specific Purposes* 10/1: 35–46.
185. Stubblebine, Tony. 2003. *Regular Expression Pocket Reference*. Cambridge, MA: O'Reilly.
186. Stubbs, Michael. 1995. Collocations and semantic profiles: on the cause of the trouble with quantitative studies. *Functions of Language* 2/1: 23–55.
187. — —. 2001. *Words and Phrases: Corpus Studies in Lexical Semantics*. Oxford: Blackwell.
188. Szmrecsanyi, Benedikt. 2005. Language users as creatures of habit: a corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory* 1/1: 113–149.
189. Tao, Hongyin. 2007. A corpus-based investigation of *absolutely* and related phenomena in spoken American English. *Journal of English Linguistics* 35/1: 5–29.
190. Tao, Hongyin and Charles F. Meyer. 2006. Gapped coordinations in English: form, usage, and implications for linguistic theory. *Corpus Linguistics and Linguistic Theory* 2/2: 129–163.

191. Thompson, P., *Spoken Language Corpora in Developing Linguistic Corpora: a Guide to Good Practice*, ed. M. Wynne, Oxbow Books, Oxford, 2005.
192. Tummers, José, Kris Heylen, and Dirk Geeraerts. 2005. Usage-based approaches in cognitive linguistics: a technical state of the art. *Corpus Linguistics and Linguistic Theory* 1/2: 225–261.
193. Venuti, L., *The Translator's Invisibility: A History of Translation*, 2nd Edition, Routledge, 2008.
194. von Eye, Alexander. 1990. *Introduction to Configural Frequency Analysis: The Search for Types and Antitypes in Cross-classifications*. Cambridge: Cambridge University Press.
195. Watt, Andrew. 2005. *Beginning Regular Expressions*. Indianapolis, IN: Wiley Publishing.
196. Whitsitt, Sam. 2005. A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics* 10/3: 283–305.
197. Wiechmann, Daniel. 2008. On the computation of collocation strength: testing measures of association as expressions of lexical bias. *Corpus Linguistics and Linguistic Theory* 4/2: 253–290.
198. Wiechmann, Daniel and Stefan Fuhs. 2006. Concordancing software. *Corpus Linguistics and Linguistic Theory* 2/1: 109–127.
199. Wulff, Stefanie. 2003. A multifactorial corpus analysis of adjective order in English. *International Journal of Corpus Linguistics* 8/2: 245–282.
200. Wynne, Martin (ed.). 2005. *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: Oxbow Books, <<http://ahds.ac.uk/linguistic-corpora/>>
201. Xiao, Richard and Tony McEnery. 2006. Collocation, semantic prosody, and near synonymy: a cross-linguistic perspective. *Applied Linguistics* 27/1: 103–129.
202. Youmans, Gilbert. 1990. Measuring lexical style and competence: the type–token vocabulary curve. *Style* 24/4: 584–599.

203. ——. 1991. A new tool for discourse analysis: the Vocabulary Management Profile. *Language* 67/4: 763-789.

7 Прилог 1.

7.1 Глосар економских термина

Глосар правних и економских термина се састоји од преко 10.000 лексичких јединица.

Због своје обимности глосар правних термина је приложен на CD-у уз овај докторски рад.

8 Прилог 2.

8.1 *Глосар правних термина*

Због своје обимности глосар правних термина је приложен на CD-у уз овај докторски рад.

9 Прилог 3.

9.1 Лиценце и слике коришћених програма

Овај дио нашег рада треба схватити као захвалност људима без чије помоћи би израда овог рада била готово немогућа, или знатно отежана. Један од начина на који им се желимо захвалити јесте тако што ћемо поменути и њих и њихове програме на којима су они марљиво радили и на тај начин испоштовати права о интелектуалној својини.

- TextSTAT -

<http://www.niederlandistik.fu-berlin.de/textstat/>

- Copyright -

(c) 2002-2008 Matthias Hüning

<mhuening@zedat.fu-berlin.de>

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Табела - Графичко рјешење програма ТекстСТАТ

The screenshot displays the TextSTAT software interface. At the top, there is a menu bar with options: Corpus, Word forms, Concordance, Citation. Below the menu bar is a toolbar with various icons. The main window is divided into two panes. The left pane, titled 'Word form', contains a list of words. The right pane, titled 'Frequency', contains the corresponding frequency values for each word. The top of the right pane shows a total frequency of 32274. Below the panes is a control panel with several options: 'sort on frequency' (selected), 'sort alphabetically', 'retrograde', 'min. frequency', 'max. frequency', 'sort case insensitive', and a checkbox for 'OR look up frequency for words containing the string:'. There is also a 'Frequency/list' button. At the bottom of the window, a status bar shows '48772 word forms/types (866326 words/tokens in corpus)', '26 files | 5850740 bytes', and the system clock '19:31'.

Word form	Frequency
the	32274
of	31030
and	16322
in	11648
i	11576
to	9220
u	8295
or	8043
for	6672
a	6298
za	5119
se	4670
o	4181
not	4102
od	4089
health	4006
1	3917
is	3766
na	3738
on	3660
by	3511
with	3500
je	3325
ili	3374
be	3298
2	3085
that	3065
The	2795
as	2659
other	2659
from	2633
10	2630
this	2597
da	2542
sa	2520
shall	2508
which	2502
are	2493
...	...

R version 2.10.1 (2009-12-14)

Copyright (C) 2009 The R Foundation for Statistical Computing

ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.

You are welcome to redistribute it under certain conditions.

Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.

Type 'contributors()' for more information and

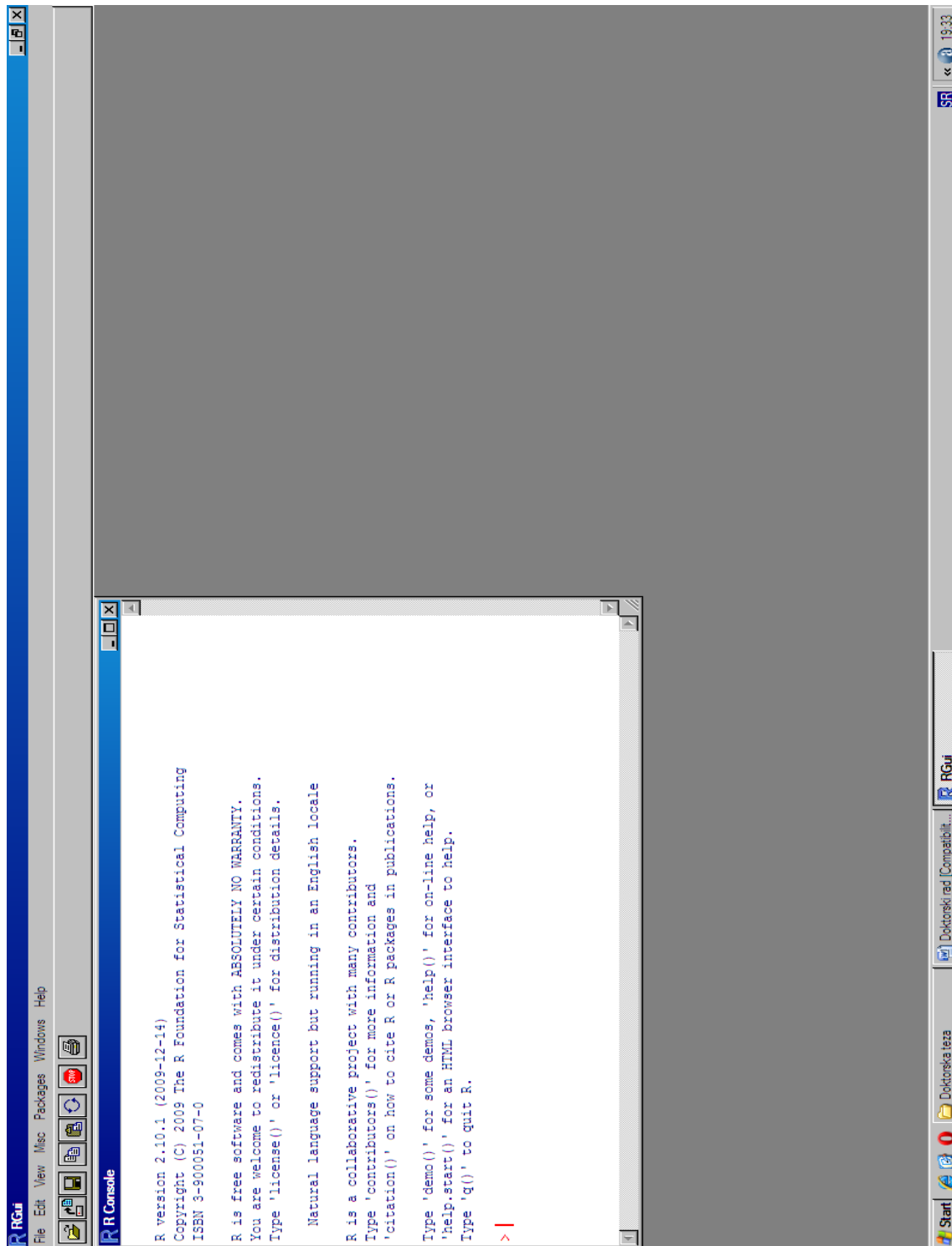
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or

'help.start()' for an HTML browser interface to help.

Type 'q()' to quit R.

Табела – графичко рјешење програма R



Табела – Графичко рјешење програма Heidelberg Tenka Text

The screenshot displays the Heidelberg Tenka Text Wordlister application. The main window shows a table of word frequencies for the file 'MASTER PLAN.doc'. The table includes columns for Item, Freq., %, Texts, Lemmas, and Set. The words listed are: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25. The words are: 1, THE, OF, V, A, H, AND, T, IF, HEALTH, IN, F, 6, 4, O, 0, E, FOR, L, I, P, N, TO, 8, 3.

Item	Freq.	%	Texts	Lemmas	Set
1	2,824	4.73%	1	100.00%	
2	2,314	3.88%	1	100.00%	
3	2,135	3.58%	1	100.00%	
4	2,029	3.40%	1	100.00%	
5	1,724	2.89%	1	100.00%	
6	1,715	2.87%	1	100.00%	
7	1,379	2.31%	1	100.00%	
8	1,288	2.16%	1	100.00%	
9	1,168	1.94%	1	100.00%	
10	1,143	1.92%	1	100.00%	
11	1,042	1.75%	1	100.00%	
12	1,003	1.68%	1	100.00%	
13	971	1.63%	1	100.00%	
14	970	1.63%	1	100.00%	
15	847	1.42%	1	100.00%	
16	835	1.40%	1	100.00%	
17	811	1.36%	1	100.00%	
18	740	1.24%	1	100.00%	
19	721	1.21%	1	100.00%	
20	664	1.11%	1	100.00%	
21	659	1.10%	1	100.00%	
22	602	1.01%	1	100.00%	
23	477	0.80%	1	100.00%	
24	473	0.79%	1	100.00%	
25	471	0.79%	1	100.00%	

Additional information from the interface: 3497 lines needed by core; 0.106 lines needed by gui; 0.077. The application title bar reads 'Heidelberg Tenka Text' and the status bar shows 'Tenka Text 0.1.3.4'.

AntConc 3.2.4w (Windows)

Developed by Laurence Anthony

Faculty of Science and Engineering

Waseda University, Japan

anthony@waseda.jp

October 4, 2011

Build: 201110042356

Acknowledgments

Thanks to the many users of AntConc who have sent me feedback and details of how they are using the software.

Табела – Графичко рјешење програма AntConc 3.2.4w

The screenshot displays the AntConc 3.2.4w (Windows) 2011 application window. The interface includes a menu bar (File, Global Settings, Tool Preferences, About), a toolbar with various analysis options (Concordance, Concordance Plot, File View, Clusters, Collocates, Word List, Keyword List), and a main workspace divided into three panes: Corpus Files, Concordance, and File. The Concordance pane shows a list of search results for the word 'Law', with columns for Hit number, KWIC (Key Word In Context), and File name. The search term 'Law' is entered in the Search Term field, and the results show 4350 hits. The interface also includes a search window size control, a Kwic Sort section, and a status bar at the bottom showing the system tray with the date and time (00:40 21/03/2013).

Hit	KWIC	File
1	STITUTIONAL BASIS FOR ADOPTION OF THE LAW	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
2	stitutional basis for adoption of the Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
3	n of Montenegro, stipulating that a law, in accordance to the Constitution, defines other	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
4	ASONS FOR ADOPTION OF THE LAW The LAW	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
5	on Protection of Population from Communicable Dis	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
6	ontenegro. Additionally, the draft Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
7	on Sanitary Inspection stipulates, apart from oth	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
8	vision over the application of the Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
9	on Protection of Population from Communicable Dis	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
10	vision over the application of the Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
11	and regulations is within the competence of a Hea	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
12	Healthcare Inspector.The existing Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
13	does not terminologically define: epidemics of gr	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
14	created the need to harmonise this Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
15	with the Constitution, since the Law was adopted	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
16	aw with the Constitution, since the Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
17	was adopted in 2005. III HARMONIZATION WITH EU	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
18	ROPEAN HEALTHCARE LEGISLATIONThe	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
19	Law is harmonised with the EU Directive No 2119/99/EC	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
20	vision over the application of this Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
21	shall be performed by Sanitary Inspectors in acco	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
22	rdination for the application of this Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
23	shall enter into force within the period of three month	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
24	to force.Article 16 stipulates the Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
25	shall enter into force within the period of eight	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
26	months. For the implementation of this Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
27	additional funds from the Budget of Montenegro ar	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
28	re not needed. PROPOSAL LAW ON AMENDEMENTS OF THE LAW ON PROTECTION OF POPULAT	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
29	ION ON AMENDEMENTS OF THE LAW ON PROTECTION OF POPULAT	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
30	ION ON PROTECTION OF POPULATION FROM COMMUNICABLE DIS	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
31	ON PROTECTION OF POPULATION FROM COMMUNICABLE DIS	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
32	Article 16In the Law on Protection of Population from Communicable Dis	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
33	After Paragraph 2 a new Item shall be added:	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
34	regulation from Article 12 of this Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
35	shall be adopted within the period of three month	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
36	Article	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
37	of three months from the day of this Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
38	entering into force. Article	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
39	16This Law shall enter into force on the eighth day if its p	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
40	ell. It is widely accepted that the Law	555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc
41	on Local Self-Government is the foundation for th	2009-11-12 Project Online ENGL sa PDF na WORD.doc

10 Прилог 4.

10.1 Састав корпуса

Приказ броја докумената (128) ЕУ који су обухваћени у нашем корпусу

Поглавља упитника:

1. D:\My Documents\Doktorska teza\1\1\B01.doc
2. D:\My Documents\Doktorska teza\1\1\B02.doc
3. D:\My Documents\Doktorska teza\1\1\B03.doc
4. D:\My Documents\Doktorska teza\1\1\B04.doc
5. D:\My Documents\Doktorska teza\1\1\B05.doc
6. D:\My Documents\Doktorska teza\1\1\B06.doc
7. D:\My Documents\Doktorska teza\1\1\C01.doc
8. D:\My Documents\Doktorska teza\1\1\C02.doc
9. D:\My Documents\Doktorska teza\1\1\C03.doc
10. D:\My Documents\Doktorska teza\1\1\C04.doc
11. D:\My Documents\Doktorska teza\1\1\C05.doc
12. D:\My Documents\Doktorska teza\1\1\C06.doc
13. D:\My Documents\Doktorska teza\1\1\C07.doc
14. D:\My Documents\Doktorska teza\1\1\C08.doc
15. D:\My Documents\Doktorska teza\1\1\C09.doc
16. D:\My Documents\Doktorska teza\1\1\C10.doc
17. D:\My Documents\Doktorska teza\1\1\C11.doc

18. D:\My Documents\Doktorska teza\1\1\C12.doc
19. D:\My Documents\Doktorska teza\1\1\C13.doc
20. D:\My Documents\Doktorska teza\1\1\C14.doc
21. D:\My Documents\Doktorska teza\1\1\C15.doc
22. D:\My Documents\Doktorska teza\1\1\C16.doc
23. D:\My Documents\Doktorska teza\1\1\C17.doc
24. D:\My Documents\Doktorska teza\1\1\C18.doc
25. D:\My Documents\Doktorska teza\1\1\C19.doc
26. D:\My Documents\Doktorska teza\1\1\C20.doc
27. D:\My Documents\Doktorska teza\2\2\C21.doc
28. D:\My Documents\Doktorska teza\3\3\C22.doc
29. D:\My Documents\Doktorska teza\3\3\C23.doc
30. D:\My Documents\Doktorska teza\3\3\C24.doc
31. D:\My Documents\Doktorska teza\3\3\C25.doc
32. D:\My Documents\Doktorska teza\3\3\C26.doc
33. D:\My Documents\Doktorska teza\3\3\C27.doc
34. D:\My Documents\Doktorska teza\3\3\C28.doc
35. D:\My Documents\Doktorska teza\3\3\C29.doc
36. D:\My Documents\Doktorska teza\3\3\C30.doc
37. D:\My Documents\Doktorska teza\3\3\C31.doc
38. D:\My Documents\Doktorska teza\3\3\C32.doc

39. D:\My Documents\Doktorska teza\3\3\C33.doc
40. D:\My Documents\Doktorska teza\Upitnik - eng\A01-eng-GOVERNMENT with changes 231109.doc
41. D:\My Documents\Doktorska teza\Upitnik - eng\A01-eng-Vlada.doc
42. D:\My Documents\Doktorska teza\Upitnik - eng\A01-eng-konac.doc
43. D:\My Documents\Doktorska teza\Upitnik - eng\A02-96.pit-eng.revised.doc
44. D:\My Documents\Doktorska teza\Upitnik - eng\A02-eng-konac.doc
45. D:\My Documents\Doktorska teza\Upitnik - eng\A02-eng.doc
46. D:\My Documents\Doktorska teza\Upitnik - eng\A02-eng2.doc
47. D:\My Documents\Doktorska teza\Upitnik - eng\A03-eng-konac.doc
48. D:\My Documents\Doktorska teza\Upitnik - eng\A03-eng.doc
49. D:\My Documents\Doktorska teza\Upitnik - eng\C01-eng-djel.doc
50. D:\My Documents\Doktorska teza\Upitnik - eng\C02-eng.doc
51. D:\My Documents\Doktorska teza\Upitnik - eng\C04-eng.doc
52. D:\My Documents\Doktorska teza\Upitnik - eng\C04-eng1.doc
53. D:\My Documents\Doktorska teza\Upitnik - eng\C09-eng.doc
54. D:\My Documents\Doktorska teza\Upitnik - eng\C14-eng.doc
55. D:\My Documents\Doktorska teza\Upitnik - eng\C16-eng.doc
56. D:\My Documents\Doktorska teza\Upitnik - eng\C16-eng-dodatniprev-nefakt-16 11 09.doc

57. D:\My Documents\Doktorska teza\Upitnik - eng\C16-eng-17Nov09 - akcize (2).doc
58. D:\My Documents\Doktorska teza\Upitnik - eng\C16-eng-djel.doc
59. D:\My Documents\Doktorska teza\Upitnik - eng\C16-eng-indirect taxation-edited.doc
60. D:\My Documents\Doktorska teza\Upitnik - eng\C16-eng-pitanja 10-17.doc
61. D:\My Documents\Doktorska teza\Upitnik - eng\C18-eng-(Prevod Upitnika).doc
62. D:\My Documents\Doktorska teza\Upitnik - eng\C18-eng-(Prevod Upitnika).docx
63. D:\My Documents\Doktorska teza\Upitnik - eng\C18-eng-konac.doc
64. D:\My Documents\Doktorska teza\Upitnik - eng\C19-eng-djel.doc
65. D:\My Documents\Doktorska teza\Upitnik - eng\C19-eng.doc
66. D:\My Documents\Doktorska teza\Upitnik - eng\C20-eng-rev.doc
67. D:\My Documents\Doktorska teza\Upitnik - eng\C20-eng-djel.doc
68. D:\My Documents\Doktorska teza\Upitnik - eng\C20-eng1.doc
69. D:\My Documents\Doktorska teza\Upitnik - eng\C20-pitanje 101 prevod.doc
70. D:\My Documents\Doktorska teza\Upitnik - eng\C20-pitanje 101 revizija odgovora.doc
71. D:\My Documents\Doktorska teza\Upitnik - eng\C21-cg_ENG.doc
72. D:\My Documents\Doktorska teza\Upitnik - eng\C21-eng.doc

73. D:\My Documents\Doktorska teza\Upitnik - eng\C22-eng.doc
74. D:\My Documents\Doktorska teza\Upitnik - eng\C23-eng-KacaR-rev.doc
75. D:\My Documents\Doktorska teza\Upitnik - eng\C23-eng-djel.doc
76. D:\My Documents\Doktorska teza\Upitnik - eng\C23-eng-konac.doc
77. D:\My Documents\Doktorska teza\Upitnik - eng\C23-eng1.doc
78. D:\My Documents\Doktorska teza\Upitnik - eng\C23-eng2.doc
79. D:\My Documents\Doktorska teza\Upitnik - eng\C24-eng.doc
80. D:\My Documents\Doktorska teza\Upitnik - eng\C26-eng.doc
81. D:\My Documents\Doktorska teza\Upitnik - eng\C27-eng-djel.doc
82. D:\My Documents\Doktorska teza\Upitnik - eng\C28 ENG.doc
83. D:\My Documents\Doktorska teza\Upitnik - eng\C28-eng.doc
84. D:\My Documents\Doktorska teza\Upitnik - eng\C29-eng.doc
85. D:\My Documents\Doktorska teza\Upitnik - eng\C30-eng.doc
86. D:\My Documents\Doktorska teza\Upitnik - eng\C32-eng.doc
87. D:\My Documents\Doktorska teza\Upitnik - eng\C32-eng2-track.doc
88. D:\My Documents\Doktorska teza\Upitnik - eng\C33-eng.doc
89. D:\My Documents\Prevod za MEI\2009-11-12 Project Online ENGL sa PDF na WORD.doc
90. D:\My Documents\Prevod za MEI\555 Zakon o zaraznim bolestima sa obrazlozenjem - prevod.doc

91. D:\My Documents\Prevod za MEI\555 Zakon o zaraznim bolestima sa obrazlozenjem.doc
92. D:\My Documents\Prevod za MEI\AP OSI 2010 2011 FINALE - prevod.doc
93. D:\My Documents\Prevod za MEI\AP OSI 2010 2011 FINALE.doc
94. D:\My Documents\Prevod za MEI\C28-cg-IgorI - prevod.doc
95. D:\My Documents\Prevod za MEI\C28-cg-IgorI.doc
96. D:\My Documents\Prevod za MEI\C28-cg.doc
97. D:\My Documents\Prevod za MEI\C28-eng.doc
98. D:\My Documents\Prevod za MEI\GLOSAR.doc
99. D:\My Documents\Prevod za MEI\FOOD SAFETY STRATEGY.doc
100. D:\My Documents\Prevod za MEI\Igor Ivanovic - prevod 2009-11-12 Project Online ENGL sa PDF na WORD.doc
101. D:\My Documents\Prevod za MEI\Igor Ivanović - glosar.doc
102. D:\My Documents\Prevod za MEI\Infectious and parasitic diseases.doc
103. D:\My Documents\Prevod za MEI\Interim agreement on trade and trade related matters.doc
104. D:\My Documents\Prevod za MEI\K28-en.doc
105. D:\My Documents\Prevod za MEI\K28_sa_izmjenama.doc
106. D:\My Documents\Prevod za MEI\LAW ON AIR PROTECTION.doc

107. D:\My Documents\Prevod za MEI\LAW ON EMERGENCY MEDICAL ASSISTANCE.doc
108. D:\My Documents\Prevod za MEI\LAW ON LIMITING USE OF TOBACCO PRODUCTS.doc
109. D:\My Documents\Prevod za MEI\LAW ON MEDICINES.doc
110. D:\My Documents\Prevod za MEI\Law on National Parks.doc
111. D:\My Documents\Prevod za MEI\MASTER PLAN Development of health care system in Montenegro for the period 2005-2010.doc
112. D:\My Documents\Prevod za MEI\MASTER PLAN.doc
113. D:\My Documents\Prevod za MEI\Nacionalni program za integraciju CG.doc
114. D:\My Documents\Prevod za MEI\National plan for avian influenza (bird flu) and flu pandemic- Annex.doc
115. D:\My Documents\Prevod za MEI\National plan for avian influenza (bird flu) and flu pandemic.doc
116. D:\My Documents\Prevod za MEI\National programme for CG integration.doc
117. D:\My Documents\Prevod za MEI\PRIRUČNIK ZA PREVOĐEŃE PRAVNIH I DRUGIH AKATA.doc
118. D:\My Documents\Prevod za MEI\Privremeni sporazum o trgovini i srodnim pitanjima.doc
119. D:\My Documents\Prevod za MEI\SAFE BLOOD STRATEGY.doc

120. D:\My Documents\Prevod za MEI\STRATEGY FOR HEALTH CARE DEVELOPMENT IN MONTENGRO.doc
121. D:\My Documents\Prevod za MEI\Sporazum o stabilizaciji i pridruživaњу.doc
122. D:\My Documents\Prevod za MEI\Stabilisation and association agreement.doc
123. D:\My Documents\Prevod za MEI\THE LAW ON HEALTH CARE.doc
124. D:\My Documents\Prevod za MEI\THE LAW ON THE PREVENTION OF POPULATION AGAINST INFECTIVE DISEASES.doc
125. D:\My Documents\Prevod za MEI\THE NATIONAL STRATEGY FOR TOBACCO CONTROL.doc
126. D:\My Documents\Prevod za MEI\ZAKON O ZASTITI VAZDUHA.doc
127. D:\My Documents\Prevod za MEI\Zakon o nacionalnim parkovima.doc
128. D:\My Documents\Prevod za MEI\Zakon o ograničenju korišćenja duvanskih proizvoda.doc

11 Прилог 5.

11.1 Интернет странице

<<http://groups.google.com/group/corpling-with-r/web/quantitative-corpus-linguistics-with-r>>

<<http://groups.google.com/group/corpling-with-r>>

11.2 Прегледи корпусне лингвистике

<<http://devoted.to/corpora>>

<<http://www.bmanuel.org/index.html>>

11.3 Референце о корпусној лингвистици

<<http://ifa.amu.edu.pl/~kprzemek/biblios/corpling.zip>>

11.4 Корпуси

BNC <<http://www.natcorp.ox.ac.uk/>>

BNC sa listom učestalosti: <<http://www.kilgarriff.co.uk/bnc-readme.html>>

Brown: <<http://icame.uib.no/brown/bcm.html>>

CELEX: <<http://www.ru.nl/celex/>>

CHILDES corpora (and CLAN): <<http://childes.psy.cmu.edu>>

Christine: <<http://www.grsampson.net/RChristine.html>>

ICE: <<http://www.ucl.ac.uk/english-usage/ice/>>

Leipzig Corpora Collection Download Page: <<http://corpora.uni-leipzig.de/download.html>>

NLTK: <<http://nltk.sourceforge.net/>>

Susanne: <<http://www.grsampson.net/RSue.html>>

Switchboard corpus: search the catalog at
<<http://www ldc.upenn.edu/>> for various versions

The JRC-ACQUIS Multilingual Parallel Corpus, Version 3.0:
<<http://langtech.jrc.it/IRC-Acquis.html>>

Treebanks: <<http://www.cis.upenn.edu/~treebank/>>

11.5 Чaconucu u бaзe нoдaтaкa

1. ACL Anthology: <<http://acl ldc.upenn.edu/>>
2. Citeseer: <<http://citeseer.ist.psu.edu/>>
3. Computational Linguistics:
<<http://mitpress.mit.edu/catalog/item/default.asp?type=4&tid=10>>
4. Computer Speech and Language:
<http://www.elsevier.com/wps/find/journaldescription.cws_home/622808/description#description>
5. Corpus Linguistics and Linguistic Theory:
<<http://www.degruyter.com/journals/cllt>>
6. Corpora:
<<http://www.eup.ed.ac.uk/journals/content.aspx?pageId=1&journalId=12505>>
7. Empirical Language Research: <<http://www.ejournals.org.uk/ELR/>>
8. ICAME Journal: <<http://nora.hd.uib.no/journal>>
9. International Journal of Corpus Linguistics:
<http://www.benjamins.com/cgi-bin/t_seriesview.cgi?series=IJCL>
10. Literary and Linguistic Computing: <<http://llc.oxfordjournals.org/>>
11. Language Resources and Evaluation (formerly known as Computers and the Humanities): <<http://www.springerlink.com/link.asp?id=113189>>
12. Web-as-corpus 1: <<http://www.webcorp.org.uk/>>
13. Web-as-corpus 2: <<http://wacky.sslmit.unibo.it/doku.php>>

11.6 *Остали релевантни пројекти*

Leipzig corpora download page: <http://corpora.informatik.uni-leipzig.de/download.html>

Marina Santini:
<http://www.nltg.brighton.ac.uk/home/Marina.Santini/>

Project Gutenberg: <http://www.gutenberg.org/>
University of Virginia Etext Center:
<http://etext.lib.virginia.edu/collections/languages/>

11.7 *Странице за преузимање програма R и података*

1. R Project: <http://www.r-project.org>
2. R Project download page CRAN: <http://cran.r-project.org/mirrors.html>
3. R websites: <http://www.rseek.org/>, <http://wiki.r-project.org/>,
and
<http://google.com/coop/cse?cx=018133866098353049407%3Aozv9awtety>
4. Tinn-R: <http://www.sciviews.org/Tinn-R/>
5. SciViews package: <http://www.sciviews.org/SciViews-R/>
6. R Graphics: <http://addictedtor.free.fr/graphiques/>
7. R sitesearch toolbar for Firefox:
<http://addictedtor.free.fr/rsitesearch/>
8. Editors for R: http://www.sciviews.org/_rgui/projects/Editors.html
9. R reference card (short): <http://cran.r-project.org/doc/contrib/Short-refcard.pdf>
10. R reference card (long): <http://cran.r-project.org/doc/contrib/refcard.pdf>
11. R/Rpad reference card: <http://www.rpad.org/Rpad/Rpad-refcard.pdf>

12. Mailing lists on using R for general statistical evaluation (i.e. not necessarily corpus-linguistic data) in linguistics:
13. R-lang: <<https://ling.ucsd.edu/mailman/listinfo.cgi/r-lang>>
14. StatForLing with R: <<http://groups.google.com/group/statforling-with-r>> (this is the newsgroup hosting the companion website for Gries's textbook on statistics for linguistics with R)

11.8 Други корисни рачунарски програми

Openoffice.org: <<http://www.openoffice.org>>

Unicode consortium: <<http://unicode.org/>>

HTML code strippers:

<<http://personal.cityu.edu.hk/~davidlee/devotedtocorpora/software.htm#HTMLcodestrippers>>

Web snaggers:

<<http://personal.cityu.edu.hk/~davidlee/devotedtocorpora/software.htm#WebSnaggers>>

11.9 Изрази

<http://regexlib.com/>

<http://regexadvice.com/blogs/ssmith/contact.aspx>

Regulazy:

<<http://tools.osherove.com/CoolTools/Regulazy/tabid/182/Default.aspx>>

The Regulator:

<http://tools.osherove.com/CoolTools/TheRegulator/tabid/185/Default.aspx>

12 Биографија аутора

Игор Ивановић је рођен у Зеници, БиХ, 1983. године. У оквиру програма размјене студената провео је, током 2005. пар мјесеци на усавршавању на преводилачким курсевима у Њемачкој, Француској, Белгији и Холандији. Дипломирао је 2006. године на Филозофском факултету у Никшићу, одсјек за енглески језик и књижевност и стекао звање професора енглеског језика и књижевности. Током 2007. завршио је специјалистичке студије, студијски програм превођење, и стекао степен специјалисте - преводилаштво. Од 2006. године ради на Институту за стране језике у Подгорици у својству сарадника у настави гдје предаје на преводилачкој групи предмета. Такође, предаје општи језик и језик струке на Правном, Машинском и Медицинском факултету. Поред тога, држи курсеве енглеског језика различитих нивоа на Институту за стране језике, а био је и предавач љетње школе енглеског језика за побједнике такмичења у Бијелој. Звање магистра стекао је 2009. године када је одбранио магистарску тезу на Институту за стране језике, Универзитета Црне Горе.

13 Прилог 6.

Изјава о ауторству

Потписани-а Игор Ивановић

број индекса 10136/Д

Изјављујем

да је докторска дисертација под насловом

Контрастивна анализа стручне терминологије ЕУ

кроз перспективу корпусне лингвистике

резултат сопственог истраживачког рада,

да предложена дисертација у целини ни у деловима није била предложена за добијање било које дипломе према студијским програмима других високошколских установа,

да су резултати коректно наведени и

да нисам кршио/ла ауторска права и користио интелектуалну својину других лица.

Потпис докторанда

У Београду, _____

14 Прилог 7.

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора Игор Ивановић

Број индекса 10136/Д

Студијски програм Наука о језику

Наслов рада Контрастивна анализа стручне терминологије ЕУ кроз перспективу
корпусне лингвистике

Ментор др Весна Половина, редовни професор

Потписани/а _____

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла за објављивање на порталу **Дигиталног репозиторијума Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис докторанда

У Београду, _____

15 Прилог 8.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Контрастивна анализа стручне терминологије ЕУ

кроз перспективу корпусне лингвистике

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање. Моју докторску дисертацију похрањену у Дигитални репозиторијум Универзитета у Београду могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство
2. Ауторство - некомерцијално
3. Ауторство - некомерцијално - без прераде
4. Ауторство - некомерцијално - делити под истим условима
5. Ауторство - без прераде
6. Ауторство - делити под истим условима

(Молимо да заокружите само једну од шест понуђених лиценци, кратак опис лиценци дат је на полеђини листа).

Потпис докторанда

У Београду, _____

1. Ауторство - Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. Ауторство – некомерцијално. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. Ауторство - некомерцијално – без прераде. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. Ауторство - некомерцијално – делити под истим условима. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. Ауторство – без прераде. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. Ауторство - делити под истим условима. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.