

UNIVERZITET U BEOGRADU

FILOLOŠKI FAKULTET

Radmila D. Palinkašević

**RAZLIKE U UPOTREBI UČESTALIH
LEKSIČKIH SPOJEVA KOD IZVORNIH I
NEIZVORNIH GOVORNIKA U NAUČNIM
RADOVIMA IZ OBLASTI OBRAZOVANJA**

doktorska disertacija

Beograd, 2024.

UNIVERSITY OF BELGRADE
FACULTY OF PHILOLOGY

Radmila D. Palinkašević

**DIFFERENCES IN THE USAGE OF LEXICAL
BUNDLES BY NATIVE AND NON-NATIVE
SPEAKERS IN SCIENTIFIC PAPERS IN THE
FIELD OF EDUCATION**

Doctoral Dissertation

Belgrade, 2024

БЕЛГРАДСКИЙ УНИВЕРСИТЕТ
ФИЛОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

Радмила Д. Палинкаевич

**РАЗЛИЧИЯ В УПОТРЕБЛЕНИИ
ЛЕКСИЧЕСКИХ СОВОКУПНОСТЕЙ
НОСИТЕЛЯМИ И НЕНОСИТЕЛЯМИ В
НАУЧНЫХ СТАТЬЯХ В ОБЛАСТИ
ОБРАЗОВАНИЯ**

Белград, 2024.

Podaci o mentoru i članovima komisije

Mentor:

dr Mirjana Daničić, docent, Univerzitet u Beogradu, Filološki fakultet

Članovi komisije:

1.

2.

3.

Datum odbrane: _____

Izjave zahvalnosti

Za izradu doktorske disertacije *Razlike u upotrebi učestalih leksičkih spojeva kod izvornih i neizvornih govornika u naučnim radovima iz oblasti obrazovanja* neizmerno se zahvaljujem svojoj mentorki doc. dr. Mirjani Daničić na velikoj podršci, posvećenosti, strpljenju, uspešnoj i prijatnoj saradnji i na sugestijama za oblikovanje rada.

Zahvaljujem se i svojim kolegicama – akademiku prof. dr Grozdanki Gojkov i dr Ivani Đorđev, prof. struk. studija na podršci, savetima i konstruktivnim razgovorima koji su pomogli u formiranju konačnih okvira teme ove doktorske disertacije.

Veliku zahvalnost dugujem i svojoj porodici, partneru i prijateljima koji su mi strpljivo pružali podršku i ohrabrenje u ovom procesu.

Razlike u upotrebi učestalih leksičkih spojeva kod izvornih i neizvornih govornika u naučnim radovima iz oblasti obrazovanja

Sažetak

Predmet ovog rada su razlike u upotrebi *učestalih leksičkih spojeva* (ULS) kod izvornih i neizvornih govornika engleskog jezika u naučnim radovima iz oblasti obrazovanja. Istraživanje obuhvata analizu upotrebe ULS u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca u tri korpusa. Korpusi su sačinjeni od naučnih radova u oblasti obrazovanja napisanih od strane izvornih govornika engleskog jezika, neizvornih govornika engleskog jezika kojima je maternji jezik srpski i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski. Istraživanje, takođe, obuhvata identifikaciju zajedničkih karakteristika upotrebe ULS kod dve grupe neizvornih govornika, kao i uticaj negativnog transfera na njihovu produkciju u naučnim radovima iz oblasti obrazovanja.

Primarni cilj istraživanja je da se ustanove razlike u karakteristikama upotrebe ULS u naučnim radovima iz oblasti obrazovanja kod izvornih i neizvornih govornika engleskog jezika u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca. Istraživanje u ovom radu je eksplorativno istraživanje kvantitativnog pristupa koje prati korpusnolingvističku metodologiju.

Rezultati su potvrdili glavnu hipotezu istraživanja koja glasi: *Karakteristike upotrebe ULS u naučnim radovima iz oblasti obrazovanja će se razlikovati kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski, u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca.*

Zaključci ovog istraživanja mogu da se koriste za unapređenje podučavanja studenata na časovima akademskog engleskog jezika, kao i pri stručnom usavršavanju ili podučavanju prevodilaca koji žele da se bave prevodom naučnih radova iz oblasti obrazovanja.

Ključne reči: učestali leksički spojevi, naučni radovi, transfer, obrazovanje, akademsko pisanje, korpusna istraživanja.

Naučna oblast: Filološke nauke

Uža naučna oblast: Primenjena lingvistika

UDK broj:

Differences in the usage of lexical bundles by native and non-native speakers in scientific papers in the field of education

Summary

The subject of this paper covers the differences in the usage of lexical bundles by native and non-native speakers of the English language in scientific papers in the field of education. The research includes the analysis of lexical bundles regarding the number of identified lexical bundles, structural classification, functional classification and the usage of the most frequent frames in the three corpora. The corpora are compiled of scientific papers in the field of education written by native speakers of the English language, non-native speakers of the English language whose mother tongue is Serbian and non-native speakers of the English language whose mother tongue is Romanian. The study also incorporates the identification of common lexical bundle usage characteristics for the two non-native speaker groups, as well as the influence of negative transfer on their lexical bundle use in scientific papers in the field of education.

The primary goal of the paper is to identify the differences in the characteristics of lexical bundle use in scientific papers in the field of education by native speakers and non-native speakers of the English language regarding the number of identified lexical bundles, structural classification, functional classification and the usage of the most frequent frames. The research in this paper is an explorative study of the quantitative approach which follows the corpus linguistic methodology.

The results confirm the main hypothesis of the study which is as follows: *The characteristics of lexical bundle usage in scientific papers in the field of education will differ in the usage of native and non-native speakers whose mother tongue is Serbian or Romanian regarding the number of identified lexical bundles, structural classification, functional classification and the usage of the most frequent frames.*

The conclusions of this study can be used for the improvement of teaching practices in academic English classes, as well as in professional development or teaching for translators who wish to translate scientific papers in the field of education.

Keywords: lexical bundles, scientific papers, transfer, education, academic writing, corpus studies.

Scientific field: Linguistics

Scientific subfield: Applied linguistics

UDC number:

Sadržaj

1. UVOD	1
1.1. Frazeologija	3
1.1.1. Učestali leksički spojevi	5
1.2. Engleski jezik kao <i>lingua franca</i> u nauci i neophodnost objavljivanja na engleskom jeziku u akademskoj zajednici	14
1.3. Translatologija	20
1.3.1. Translatologija i učestali leksički spojevi (ULS)	21
1.4. Transfer	31
1.4.1. Definicija transfera	31
1.4.2. Tipovi transfera	32
1.4.3. Opseg uticaja transfera	33
1.4.4. Istraživanje transfera/ međujezičkog uticaja	34
1.4.5. Identifikacija transfera i tipovi dokaza	35
1.4.6. Faktori koji utiču na transfer /međujezički uticaj	37
1.4.7. Relevantna prethodna istraživanja uticaja transfera na upotrebu učestalih leksičkih spojeva	39
2. OSVRT NA DOSADAŠNJA ISTRAŽIVANJA UČESTALIH LEKSIČKIH SPOJEVA ..	52
2.1. Prethodna istraživanja učestalih leksičkih spojeva u oblasti obrazovanja	52
2.1.1. Paradigmatske razlike u upotrebi ULS u naučnim radovima iz oblasti obrazovanja	54
2.1.2. Razlike u upotrebi ULS u oblasti obrazovanja kroz vreme	55
2.2. Prethodna istraživanja upotrebe ULS kod neizvornih govornika engleskog jezika kojima je maternji srpski ili rumunski jezik	58
2.3. Istraživanja upotrebe ULS u naučnim radovima kod izvornih i neizvornih govornika	69
3. METODOLOŠKI OKVIR ISTRAŽIVANJA	105
3.1. Predmet, obim, istraživačko pitanje i ciljevi istraživanja	105
3.2. Hipoteze istraživanja	106
3.3. Korpusi	108
3.4. Metodološko-analitički postupak i metode	112
3.5. Ograničenja istraživanja	117
4. REZULTATI	118
4.1. Izdvajanje učestalih leksičkih spojeva	118
4.2. Prečišćavanje spiskova učestalih leksičkih spojeva	118
4.3. Strukturalna raspodela učestalih leksičkih spojeva	120
4.3.1. Strukturalna raspodela ULS kod izvornih govornika engleskog jezika	120
4.3.2. Strukturalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski	123
4.3.3. Strukturalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski	125
4.3.4. Poređenje strukturalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik srpski	127
4.3.5. Poređenje strukturalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik rumunski	130
4.3.6. Poređenje strukturalne raspodele ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski ili srpski	133

4.3.7. Poređenje strukturalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik srpski ili rumunski	137
4.4. Funkcionalna raspodela ULS.....	139
4.4.1. Funkcionalna raspodela ULS kod izvornih govornika engleskog jezika.....	139
4.4.2. Funkcionalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski	141
4.4.3. Funkcionalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski	142
4.4.4. Poređenje funkcionalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik srpski	144
4.4.5. Poređenje funkcionalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik rumunski.....	147
4.4.6. Poređenje funkcionalne raspodele ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski i rumunski.....	150
4.4.7. Poređenje funkcionalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski	152
4.5. Interakcije strukturalne i funkcionalne raspodele ULS u ECEP korpusu	154
4.6. Interakcije strukturalne i funkcionalne raspodele ULS u SCEP korpusu.....	155
4.7. Interakcija strukturalne i funkcionalne raspodele ULS u RCEP korpusu	157
4.8. Poređenje interakcije strukturalne i funkcionalne raspodele u ECEP, SCEP i RCEP korpusima	158
4.9. Zajednički ULS i njihova strukturalna i funkcionalna raspodela.....	159
4.9.1. Strukturalna i funkcionalna raspodela zajedničkih ULS za sva tri korpusa	159
4.9.2. Strukturalna i funkcionalna raspodela zajedničkih ULS za ECEP i SCEP korpusa.....	160
4.9.3. Strukturalna i funkcionalna raspodela zajedničkih ULS za ECEP i RCEP korpusa.....	161
4.9.4. Strukturalna i funkcionalna raspodela zajedničkih ULS za SCEP i RCEP korpusa.....	163
4.10. Upotreba najzastupljenijih obrazaca ULS u ECEP, SCEP i RCEP korpusima.....	165
4.10.1. Obrazac <i>the + imenica + of the/a</i>	166
4.10.2. Obrazac <i>in the + imenica + of</i>	167
4.10.3. Obrazac <i>glagol u pasivu + fragment predloške fraze</i>	168
4.10.4. Obrazac <i>uvodna it struktura + glagolska/ pridevska fraza</i>	169
4.11. Rezultati analize uticaja transfera.....	170
4.11.1. Unutargrupna homogenost.....	170
4.11.2. Međugrupna heterogenost.....	170
4.11.2.1. Strukturalna raspodela specifičnih ULS	171
4.11.2.2. Funkcionalna raspodela specifičnih ULS	189
4.11.2.3. Korišćenje podataka iz prethodnih kompatibilnih istraživanja za dodatno pročišćavanje spiskova specifičnih ULS za neizvorne govornike engleskog kojima je maternji jezik srpski ili rumunski.....	205
4.11.2.4. Analiza potencijalne zastupljenosti specifičnih ULS za neizvorne govornike engleskog jezika u ECEP korpusu	206
4.11.3. Međujezička podudarnost performansi.....	208
5. DISKUSIJA.....	213
5.1. Zastupljenost ULS kod izvornih i neizvornih govornika engleskog jezika.....	213
5.2. Strukturalna raspodela ULS.....	215
5.3. Funkcionalna raspodela ULS.....	218
5.4. Interakcija strukturalne i funkcionalne raspodele ULS	220
5.5. Zajednički ULS.....	220
5.6. Specifični ULS	222
5.7. Uticaj negativnog transfera na produkciju ULS kod neizvornih govornika.....	223
engleskog jezika	223

5.8. Upotreba najzastupljenijih obrazaca ULS	224
5.8.1. Obrazac <i>the + imenica + of the/a</i>	224
5.8.2. Obrazac <i>in the + imenica + of</i>	225
5.8.3. Obrazac <i>glagol u pasivu + fragment predloške fraze</i>	226
5.8.4. Obrazac <i>uvodna it struktura + glagolska/ pridevska fraza</i>	227
5.8.5. Sumiranje podataka o upotrebi najzastupljenijih obrazaca ULS	228
5.9. Metodičke smernice	228
5.9.1. Metodičke smernice koje se predlažu za rad sa izvornim govornicima srpskog jezika	229
5.9.2. Metodičke smernice koje se predlažu za rad sa izvornim govornicima rumunskog jezika	231
5.9.3. Metodičke smernice koje se predlažu za rad sa obe grupe neizvornih govornika	233
6. ZAKLJUČNA RAZMATRANJA	234
7. LITERATURA	240
Prilozi	247
Prilog 1. <i>Izdvojeni ULS u korpusu izvornih govornika engleskog jezika – spisak</i>	247
Prilog 2. <i>Izdvojeni ULS u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik rumunski – spisak</i>	248
Prilog 3. <i>Izdvojeni ULS u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik srpski – spisak</i>	249
Prilog 4. <i>Strukturalna raspodela ULS za izvorne govornike engleskog jezika – spisak</i>	251
Prilog 5. <i>Strukturalna raspodela ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski – spisak</i>	252
Prilog 6. <i>Strukturalna raspodela ULS za neizvorne govornike engleskog jezika kojima je maternji jezik rumunski – spisak</i>	254
Prilog 7. <i>Funkcionalna raspodela ULS za izvorne govornike engleskog jezika – spisak</i>	255
Prilog 8. <i>Funkcionalna raspodela ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski – spisak</i>	256
Prilog 9. <i>Funkcionalna raspodela ULS za izvorne govornike rumunskog jezika – spisak</i>	258
Prilog 10. <i>Spisak specifičnih ULS za izvorne govornike engleskog jezika</i>	259
Prilog 11. <i>Spisak specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski</i>	259
Prilog 12. <i>Spisak specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik rumunski</i>	261
Prilog 13. <i>Spisak specifičnih ULS za izvorne govornike turskog jezika, preuzeto iz Gungor (2016)</i>	262
Prilog 14. <i>Strukturalna raspodela specifičnih ULS za izvorne govornike engleskog jezika – spisak</i>	263
Prilog 15. <i>Strukturalna raspodela specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski – spisak</i>	264
Prilog 16. <i>Strukturalna raspodela specifičnih ULS za izvorne govornike rumunskog jezika – spisak</i>	265
Prilog 17. <i>Funkcionalna raspodela specifičnih ULS za izvorne govornike engleskog jezika – spisak</i>	266
Prilog 18. <i>Funkcionalna raspodela specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski – spisak</i>	267
Prilog 19. <i>Funkcionalna raspodela specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik rumunski – spisak</i>	268

Biografija autorke

Izjava o autorstvu

Izjava o istovetnosti štampane i elektronske verzije doktorskog rada

Izjava o korišćenju

1. Uvod

Većini nas se dogodilo da čitajući određeni naučni tekst „ošetimo” da pisac nije izvorni govornik¹ engleskog jezika, iako je tekst gramatički ispravan i u njemu se koristi adekvatna terminologija. Šta je to što intuitivno osećamo da nedostaje, a ne možemo, bez dubljih istraživanja, da izdvojimo? Istraživanja formulaičnog jezika su dobro mesto za traženje odgovora na postavljeno pitanje, posebno u segmentu istraživanja *učestalih leksičkih spojeva* (engl. *lexical bundles*; u daljem tekstu: ULS). Naime, formulaičan jezik pored prenošenja idiomatičnosti, pokazuje i pripadnost određenoj diskursnoj zajednici (Adel & Erman, 2012). ULS su posebno pogodni za traženje odgovora na gorenavedeno pitanje zato što nisu izraziti (engl. *salient*) u jeziku i mogu se precizno identifikovati jedino korpusnom analizom (Biber et al., 1999). Drugim rečima, ti frazeološki elementi čijeg postojanja većina neizvornih govornika nije ni svesna zato što njihova primena u obrazovnim programima i dalje nije postala široko rasprostranjena (Güngör, 2016), koriste se za podržavanje funkcionalne koherentnosti u tekstu i služe kao signal pripadnosti određenoj akademskoj zajednici (Pang, 2010).

Svejlz (Swales, 1990) je svojim radom ponovo pokrenuo interesovanje za pojam žanra, što je uticalo na to da analiza akademskog diskursa postane popularnija oblast u primenjenoj lingvistici, pa su tako strukture i funkcije akademskog diskursa postale zastupljenije kao predmet analize u kontrastivnoj i korpusnoj lingvistici. Pored spiskova reči i kolokacija, ULS su postali značajan segment analize akademskog diskursa. ULS predstavljaju kombinacije nekoliko reči koje se neretko javljaju u određenom registru sa širokom rasprostranjenošću u različitim tekstovima samog registra (Biber, 2010, prema Adel & Erman, 2012). ULS nam pomažu da definišemo fluentno korišćenje jezika (uglavnom pisanog). Oni su drugačiji od jedne akademske discipline do druge pa su označitelji adekvatne konstrukcije značenja koja je prihvaćena u toj zajednici (Hyland, 2012). Studentima i pripadnicima određene akademske zajednice od posebnog su značaja kada je reč o veštinama pisanja, te se obrada učestalih leksičkih spojeva preporučuje u nastavi akademskog pisanja (Pang, 2010). Potreba za istraživanjima koja imaju cilj asistenciju studentima i članovima akademske zajednice u stvaranju uverljivog akademskog glasa, postaje izraženija kada uzmemo u obzir sve veći pritisak koji se stavlja na članove akademske zajednice da objavljuju svoje nalaze na engleskom jeziku. Naime, kao što je poznato, u akademskom svetu važi pravilo koje se kolokvijalno označava popularnom frazom „objavljuj ili nestani” (engl. *publish or perish*), koja ilustruje značaj redovnog objavljivanja za članove akademske zajednice. Proces globalizacije doneo je članovima akademske zajednice nove zahteve u pogledu objavljivanja naučnih radova u međunarodno prepoznatim časopisima u kojima se prati bibliometrijski sistem vrednovanja, odnosno časopisima sa *SCI* liste. Kako bi naučnici iz zemalja kao što je Srbija ili Rumunija mogli da objave svoje nalaze u prepoznatim časopisima, primorani su da ih objavljuju na engleskom jeziku, zato što je u ovim časopisima očigledna dominacija engleskog jezika i anglofone kulture (Luo & Hyland, 2019; Igrutinović, 2021). Međutim, nemaju svi istraživači dovoljno dobro poznavanje engleskog jezika da pišu naučne radove na engleskom jeziku pa moraju da se oslone na usluge prevodilaca (Muresan & Pérez-Llantada, 2014; Igrutinović, 2021). Zbog kompleksnosti akademskog diskursa i prevodioci koji prevode naučne radove moraju da budu upućeni u norme koje se koriste u ciljanoj akademskoj zajednici, odnosno delatnoj zajednici.

Prethodna istraživanja koja su se bavila učestalim leksičkim spojevima obradila su temu sa različitih aspekata i došla su do novih podataka o ovom leksičkom fenomenu. Inicijalna istraživanja su analizirala ULS u različitim registrima (Altenberg, 1998; Biber et al., 1999), nakon čega su vršene analize u različitim žanrovima (npr. Biber et al., 2004; Heng et al., 2014). U akademskom

¹ Svi pojmovi koji su u tekstu upotrebljeni u muškom gramatičkom rodu obuhvataju muški i ženski rod lica na koja se odnose.

diskursu analiza ULS je posebno zaživela i ustanovilo se da, iako postoje određeni ULS koji su zastupljeniji u ovom žanru uporedno sa drugim, postoje velike razlike u zavisnosti od naučne oblasti, tipa teksta i iskustva u akademskom pisanju (Hyland, 2008; Staples et al., 2013; Chen & Baker, 2016). U svetu akademskog diskursa naučni radovi predstavljaju najistaknutiji žanr pa nije iznenađujuće što su i u ispitivanju upotrebe ULS postali popularni (Güngör, 2016). Istraživanja su pokazala da se upotreba ULS razlikuje u naučnim radovima iz različitih oblasti i u produkciji kod izvornih govornika engleskog jezika i neizvornih govornika.

Iako je zaživelo ispitivanje upotrebe ULS u naučnim radovima iz različitih oblasti kod izvornih i neizvornih govornika engleskog jezika, i dalje ne znamo dovoljno o karakteristikama upotrebe ULS kod neizvornih govornika. Naime, teško je porediti rezultate istraživanja i doći do zajedničkih karakteristika neizvornih govornika engleskog jezika sa različitim maternjim jezicima, zato što se prethodna istraživanja razlikuju po veličinama i načinu formiranja korpusa, korišćenim parametrima za identifikaciju ULS, kao i različitim taksonomijama za strukturalnu i funkcionalnu raspodelu izdvojenih ULS. U istraživanju ULS dobro je poznato da se njihova identifikacija vrši pomoću postavljanja odgovarajuće minimalne frekvencije spoja i rasprostranjenosti spojeva u tekstovima korpusa. Kako bi komparacija korpusa različitih veličina bila moguća, frekvencija se normalizuje na određen broj po milion reči. Međutim, prethodna istraživanja (Cortes, 2004; Bestgen, 2019; Pan et al., 2020) ukazuju na neefikasnost normalizovane vrednosti minimalne frekvencije ULS za komparaciju ULS u korpusima različitih veličina. Naime, disproportiono veći broj ULS se identifikuje u manjim korpusima. Štaviše, Pan i saradnici (2020) demonstrirali su kako se različit broj i struktura ULS identifikuje na osnovu toga da li su korpusi ujednačeni po veličini i/ili broju radova.

U istraživanju u ovoj disertaciji ispitaćemo razlike u upotrebi ULS u naučnim radovima iz oblasti obrazovanja kod izvornih govornika engleskog jezika, neizvornih govornika engleskog jezika kojima je maternji jezik srpski i neizvornih govornika kojima je maternji jezik rumunski. Upoređićemo upotrebu ULS u tri kompatibilna korpusa koji su ujednačeni po veličini, registru, žanru, oblasti istraživanja i načinu formiranja kako bi se dobili jasni rezultati o načinu upotrebe ULS kod izvornih i neizvornih govornika u datoj oblasti akademskog diskursa. Na osnovu dobijenih rezultata o zastupljenosti ULS, strukturalnoj i funkcionalnoj raspodeli i upotrebi najzastupljenijih obrazaca, možemo početi da gradimo jasnije razumevanje upotrebe ULS kod neizvornih govornika, tačnije, utvrditi njihove zajedničke karakteristike, kao i negativan uticaj transfera na njihovu produkciju. Temeljnije razumevanje karakteristika upotrebe ULS kod neizvornih govornika profesorima engleskog jezika će omogućiti da unaprede svoju praksu, a studentima i članovima akademske zajednice koji se upuštaju u objavljivanje radova na engleskom jeziku olakšaće formiranje uspešnog akademskog glasa. Štaviše, rezultati istraživanja koristiće i prevodiocima, koji žele da se bave naučnim radovima iz oblasti obrazovanja, pri profesionalnom usavršavanju ili pri obuci budućih prevodilaca u formalnim uslovima.

Srpski i rumunski jezik odabrani su za komparaciju sa engleskim zato što su iz različitih jezičkih grupa. Takođe, rumunski je jezik nacionalnih manjina u Srbiji pa je profesorima stranih jezika u Srbiji neophodno da poseduju znanja za pružanje podjednako efikasnih instrukcija studentima kojima je maternji jezik rumunski, kao i za studente kojima je maternji jezik srpski.

Naučni doprinos istraživanja ogleda se u tome što, koliko je autorki poznato, još nisu izvršena istraživanja ULS u naučnim radovima iz određene oblasti u kojima je u korpusima iste veličine upoređena primena ULS kod dve grupe neizvornih govornika i jedne izvornih. Štaviše, ULS su nedovoljno istražen leksički fenomen u međujeziku (engl. *interlanguage*) osoba kojima je maternji jezik srpski ili rumunski. Rezultati istraživanja će omogućiti da bolje razumemo karakteristike upotrebe ULS kod dve navedene grupe neizvornih govornika kao i da bolje razumemo razlike u

međujeziku različitih grupa govornika. Bolje razumevanje karakteristika upotrebe ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski unaprediće podučavanja ULS.

1.1. Frazeologija

Iako istraživanje jedinica od nekoliko reči ima dugu istoriju (sa najranijim istraživanjem koje datira još iz 1909. godine), relativno skoro je frazeologija postala zasebna istraživačka oblast (Granger & Meunier, 2008). Dva glavna faktora koja su blokirala taj proces i koja i dalje predstavljaju poteškoće u ovoj oblasti su: 1. veoma promenljiv i širok obim oblasti; 2. obimna i konfuzna terminologija koja se koristi u oblasti frazeologije (Granger & Paquot, 2008). Frazeologija može da se definiše kao „istraživanje strukture, značenja i upotrebe kombinacija reči” (Cowie, 1994: 3168). Altenberg je frazeologiju okarakterisao kao „mutnu stranu jezika”² (1998: 101) prevashodno se osvrćući na činjenicu da je neosporno da u jeziku postoje spojevi od nekoliko reči kao što su idiomi, kolokacije i fiksne fraze, ali da je teško postaviti granice za razne oblike u ovoj oblasti i klasifikovati različite tipove. Postojanje velikog broja unapred oformljenih izraza čini razliku između leksikona i gramatike manje jasnom i izraženom, što ukazuje na to da, kada govorimo o leksikalizaciji i produkciji, govorimo o stepenu zastupljenosti, a ne o strogo definisanoj dihtonomiji (Pawley & Syder, 1983, prema Altenberg, 1998).

Opseg frazeologije kao oblasti zavisi od kriterijuma na osnovu kojih lingvisti definišu frazeološke jedinice, odnosno od načina na koji ih razlikuju od jedinica koje nisu frazeološke. Na osnovu razlika u konceptualizaciji frazeoloških jedinica izdvajaju se dva glavna pristupa u frazeologiji: *frazeološki pristup* i *tradicionalni pristup*.

Frazeološki pristup ili pod drugim nazivom *tradicionalni pristup* vuče korene iz istraživanja koje su sprovodili naučnici iz Sovjetskog Saveza i istočne Evrope. Ovaj pristup ograničava opseg frazeologije na specifičnu podgrupu lingvistički definisanih jedinica od nekoliko reči. Naime, istraživanja u tradicionalnom pristupu najviše su pažnje posvećivala idiomima i drugim oblicima koji su po svojoj prirodi nekompozicioni (Granger & Paquot, 2008).

Distribucionalni pristup (Evert, 2004, prema Granger & Paquot, 2008) ili pristup zasnovan na frekvenciji (Nesselhauf, 2004) je proistekao iz istraživanja koje je sproveo Sinkler (Sinclair). On je primenio induktivan pristup zasnovan na upotrebi korpusa za identifikaciju kombinacija reči što je zauzvrat dovelo do širenja opusa istraživanja i uključivanje kombinacija reči koje ne pripadaju ranije definisanim lingvističkim kategorijama. Novi frazeološki oblici koji su postali predmet istraživanja u distribucionalnom pristupu ilustruju Sinklerov (1991: 110, prema Erman & Warren, 2000) pristup idioma (engl. *idiom principle*) koji objašnjava na sledeći način: „jezičkom korisniku je dostupan veliki broj donekle unapred strukturiranih fraza koje predstavljaju jedan izbor, iako deluju kao da mogu da se analiziraju u segmentima”³. Drugim rečima, po principu idioma „jezik je u suštini sačinjen od niza zajedno izdvojenih reči koje sačinjavaju jedan izbor”⁴ (Granger & Paquot, 2008: 29). Stoga, možemo zaključiti da se istraživanja koja se bave unapred formiranim izrazima od nekoliko reči (engl. *multi word prefabricated expressions*) suprotstavljaju stavu da je jezik čisto kompozicioni i podržavaju teoriju da je velika količina našeg govora sačinjena od unapred formiranih izraza (Ellis, 1996; Howarth, 1996; Wray, 2002, prema Biber i sar., 2004). Zahvaljujući pomacima koje je napravio Sinkler „mnogi oblici koji su tradicionalno smatrani perifernim ili van

² Original na engleskom jeziku: „Phraseology is a fuzzy part of language”; prevod svih citata – msr Radmila Palinkašević.

³ „a language user has available to him a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments”

⁴ „language as essentially made up of strings of co-selected words that constitute single choices”

opsega frazeologije su sada postali centralni jer se pokazalo da prožimaju jezik, dok su se mnogi najograničeniji oblici (idiomi, poslovice) pokazali „veoma nefrekventnim” (Granger & Paquot, 2008: 29).

Veliki broj inicijalnih istraživanja (Hakuta, 1974; Redeker, 1991; Sinclair, 1991; Lewis, 1993; Howarth, 1996; Wray & Perkins, 2000, prema Biber i sar., 2004) u ovoj oblasti su prevashodno bila teoretske prirode, upoređujući različite perspektive i pristupe sekvencama od nekoliko reči i predlagajući okvire za analizu. U datoj oblasti Weinert (Weinert, 1995: 182, prema Biber i sar., 2004) je identifikovala dva osnovna istraživačka problema: 1. koji je najbolji način definisanja i identifikacije fiksnih struktura sačinjenih više od jedne reči. 2. *analiza diskursnih funkcija koje te strukture vrše*. U literaturi postoji konsenzus po pitanju važnosti struktura od više reči, međutim, po većini drugih pitanja kao što su definišuće karakteristike ovih struktura, metodologije za njihovu identifikaciju i samo imenovanje struktura *ne* postoji slaganje među stručnjacima. Bajber i saradnici (Bajber i sar., 2004) navode pet načina na koje se razlikuju istraživanja u ovoj oblasti:

- istraživački ciljevi: opisivanje svih sekvenci od nekoliko reči ili opisivanje manjeg broja „važnih” sekvenci;
- kriterijumi za identifikaciju sekvenci od nekoliko reči;
- formalne karakteristike sekvenci od nekoliko reči: od koliko su reči sačinjene, da li su idiomatski izrazi ili leksičko-gramatičke strukture itd.;
- materijal na kome se vrši istraživanje: u opsegu od nekoliko tekstova do velikih korpusa;
- pitanje registra: neka istraživanja ne uzimaju u obzir pitanje registra, neka se ograničavaju samo na jedan, dok druga upoređuju upotrebu sekvenci od nekoliko reči u većem broju registra.

Bajber i saradnici (Bajber et al., 2004) veruju da primena samo jednog pristupa ne može da nam pruži potpunu sliku kompleksne lingvističke pojave kao što su spojevi od nekoliko reči, te smatraju da se potpuniji uvidi mogu napraviti jedino primenom empirijskog istraživanja iz nekoliko perspektiva (str. 372). U kasnijoj publikaciji, Grej i Bajber (Gray & Biber, 2015) daju okvir za sagledavanje glavnih parametara za dizajn frazeoloških istraživanja koji je predstavljen u Tabeli 1⁵.

Prva stavka deli frazeološka istraživanja na ona koja primenjuju pristup zasnovan na korpusu (*corpus-based approach*) i ona koja primenjuju pristup vođen korpusom (*corpus-driven approach*), koja su jasno definisana na osnovu toga da li su leksički izrazi izdvojeni pre korpusne analize ili tokom korpusne analize. Isti termini – *pristup vođen korpusom* i *pristup zasnovan na korpusu* koriste se i u korpusnolingvističkim istraživanjima u globalu, ali su tu distinkcije manje jasne, zbog čega se u korpusnoj lingvistici vodi debata o validnosti distinkcije (McEnery & Hardie, 2011). S obzirom na činjenicu da je istraživanje u ovoj disertaciji frazeološko, možemo primeniti gorenavedenu distinkciju i okarakterisati ga kao istraživanje koje primenjuje pristup vođen korpusom jer se prvo izdvajaju ULS nakon čega se analiziraju. Prateći dalje parametre koje su istakli Grej i Bajber (2015), svrstavamo istraživanje u ovoj disertaciji u istraživanja koja se fokusiraju na frazeološke obrasce unutar jednog registra, odnosno naučne radove u oblasti obrazovanja. U završetku opisa istraživanja u odnosu na istraživački cilj, svrstavamo ga u ona istraživanja koja ispituju diskursnu funkciju, jer se ULS dele u funkcionalne kategorije čija se upotreba dalje analizira. Na osnovu karakteristika jedinica od više reči, pošto se u istraživanju u ovoj disertaciji ispituju ULS, ono spada u istraživanja koja ispituju veoma česte sekvence koje nisu idiomatske, koje su duže od tri reči i koje su kontinuirane.

⁵Tabela je preuzeta iz istraživanja Grej i Bajber (2015).

Tabela 1. Glavni parametri za dizajn frazeoloških istraživanja

Istraživački cilj	Karakteristike jedinice od više reči
<i>Obim i metodološki pristup</i>	<i>Idiomatski status</i>
1. Ispitivanje upotrebe unapred izdvojenih leksičkih izraza (pristup zasnovan na korpusu)	1. Fiksni idiomatski izrazi
2. Identifikacija i opis svih sekvenci od nekoliko reči u korpusu (pristup vođen korpusom)	2. Veoma česte sekvence koje nisu idiomatske
<i>Uloga registra</i>	<i>Dužina</i>
3. Poređenje frazeoloških obrazaca u različitim registrima	3. Kratke kombinacije od dve do tri reči
4. Fokus na frazeološkim obrascima unutar jednog registra	4. Duže sekvence sa više od tri reči
5. Fokus na opšte korpuse bez osvrta na registar	
<i>Diskursna funkcija</i>	Kontinuirani/Nekontinuirani
6. Ispituje se diskursna funkcija	5. Kontinuirane (neprekinute) sekvence
7. Ne ispituje se diskursna funkcija	6. Nekontinuirane sekvence sa promenljivim elementima

Istraživanja koja prate distribucionalni pristup, ređe se bave fiksnim idiomatskim izrazima zbog dve činjenice: 1. programi za analizu korpusa ne mogu automatski da izdvoje idiome (Greaves & Warren, 2010); 2. idiomi su zauzimali centralno mesto u istraživanjima koja prate frazeološki, odnosno tradicionalni pristup (Greindžer i Paket, 2008). U distribucionalnom pristupu istraživanja kraćih frazalnih struktura najčešće se bave kolokacijama (Grej i Bajber, 2015), dok su se istraživanja dužih frazalnih struktura usredsredila na n-grame, u koje spadaju i ULS (Greaves & Warren, 2010).

1.1.1. Učestali leksički spojevi

Neusklađenost korišćene terminologije u frazeološkim istraživanjima predstavlja jedan od većih nedostataka u ovoj oblasti, jer se dešava da različiti istraživači ispituju isti fenomen dok ga imenuju na različite načine, ali i da se različiti fenomeni ispituju a na isti način imenuju (Wary, 2002; Greindžer i Paket, 2008). Neki od termina koji su korišćeni za sekvence od nekoliko reči su sledeći: klaster (engl. *cluster* – Stubbs, 2007), ponovljene kombinacije reči (engl. *recurrent word combinations* – Altenberg, 1998), leksičke fraze (engl. *lexical phrases*), n-grami (engl. *n-grams* – Stubbs, 2007), jezički delovi (engl. *chunks* – Hyland, 2008; Hyland, 2012), fiksni izrazi (engl. *fixed expressions*), frazalni okviri (engl. *phrase-frames* – Greaves & Warren, 2010), skipgrami (engl. *skipgrams*), konkgrami (engl. *conccgram* – Cheng, Greaves & Warren, 2010) i učestali leksički spojevi (engl. *lexical bundles* – Bajber i sar., 1999; Pan i sar., 2016).

U ovom istraživanju koristićemo termin *učestali leksički spojevi* (u daljem tekstu: ULS), koji su uveli Bajber i saradnici (1999). Oni (1999: 989) definišu ULS kao „spojeve reči koje pokazuju statističku tendenciju da se koriste zajedno”⁶. Hajland (2008: 41) definiše ULS kao „proširene kolokacije koje se javljaju isuviše često da bi to bila slučajnost; pomažu u formiranju značenja i doprinose osećaju koherentnosti u tekstu”⁷. Po Altenbergu (1998: 122), ULS su „konvencionalizovani gradivni elementi koji se koriste u zgodnim rutinama u jezičkoj produkciji”⁸.

⁶ „bundles of words that show a statistical tendency to co-occur”

⁷ „extended collocations which appear more frequently than expected by chance, helping to shape meanings and contributing to our sense of coherence in a text”

⁸ „conventionalized building blocks that are used as convenient routines in language production”

Zbog gorenavedenih terminoloških neujednačenosti u različitim istraživanjima leksički fenomen ULS se, takođe, označavao i terminima ponovljene kombinacije reči (Altenberg, 1998), kombinacije reči (engl. *word combinations* – Cowie, 1992) i frazeološke jedinice (Moon, 1998). Važno je da ne pomešamo koncept ULS sa konceptom n-grama. N-grami su širi pojam od ULS i predstavljaju „niz neprekinutih formi reči”⁹ (Stubbs, 2007: 90), a definišu se kao „sekvence reči koje se ispituju kao celina, gde *n* označava koliko se reči nalazi u jedinici” (Crawford & Csomay, 2016: 41, prema Esfandiari & Barbary, 2017). N-grami teorijski mogu biti sačinjeni od neograničenog broja reči koje se pojavljuju zajedno, nezavisno od frekvencije zastupljenosti tog spoja. Međutim, empirijski su istraživani n-grami sa najviše devet reči (Cortes, 2013, prema Esfandiari & Barbary, 2017). ULS se razlikuju od n-grama zato što imaju određenu minimalnu frekvenciju pojavljivanja i minimalnu distribuciju u tekstovima. Kako je to jasno objašnjeno u Kortez (Cortes, 2015: 200, prema Esfandiari & Barbary, 2017) „svi učestali leksički spojevi su n-grami ali nisu svi n-grami učestali leksički spojevi”¹⁰.

Većina ULS nemaju idiomatsko značenje prevashodno zato što su idiomi isuviše retki da bi bili izdvojeni kao ULS. Tipični idiomi u engleskom jeziku kao što su *kick the bucket* ili *a slap in the face* retko se susreću u govoru, njihova upotreba je tipična za fikciju, odnosno književne tekstove gde se npr. dva gorenavedena primera pojavljuju pet puta u milion reči (Bajber et al., 2004: 377).

Uporedno sa *formulaičnim jezikom* (engl. *formulaic language*), koji je karakterisan sintaktički i semantički potpunim strukturama, ULS su korpusno izdvojeni spojevi za koje nije neophodno da budu idiomatični i celoviti (Biber, Conrad, & Rappen, 1999, prema Appel & Murray, 2020). Bajber i saradnici (Bajber et al., 2004) kao primer za činjenicu da ULS uglavnom nisu potpune strukturalne jedinice, navode da se samo 15% ULS u konverzaciji može smatrati potpunim frazama ili klauzama, dok manje od 5% ULS u akademskoj prozi predstavlja potpune strukturalne jedinice. Većina ULS spaja dve strukturalne jedinice, kao na primer dve klauze, što se češće javlja u konverzaciji ili dve fraze, što je češće u akademskoj prozi. ULS često počinju na kraju klauze ili fraze i uključuju barem jednu reč iz sledeće strukturalne jedinice (Biber et al., 2004). Iako su ULS najčešće fragmentovane fraze ili klauze, kako navodi Kortez (2004), one imaju jake gramatičke korelate koji omogućuju njihovo grupisanje u strukturalne tipove. Na osnovu navedenog, jasno je da su ULS sekvence od nekoliko reči koje seku gramatičke konstrukcije. Međutim, oni imaju definisane funkcije u diskursu što ukazuje na njihovu značajnu ulogu u komunikativnom repertoaru govornika, iako se ne uklapaju u strukture koje se tradicionalno ispituju u lingvističkim istraživanjima (Biber et al., 2004). Elementi u ULS su nepromenjivi, što ukazuje na fiksnu prirodu ove frazeološke pojave (Lazić, 2017). ULS nisu perceptivno upadljivi izrazi (Biber & Barbieri, 2007), zato što se primarno identifikuju na osnovu njihove frekvencije i disperzije a ne na osnovu strukture.

ULS mogu da budu različitih dužina; broj reči koje ih sačinjavaju kreću se od tri do šest u istraživanjima. Iako su se istraživanja bavila ULS različitih dužina, najzastupljenija dužina u istraživanjima je ULS od četiri reči (Hyland, 2008; Pan i sar., 2016; Lazić, 2017; Ebeling, 2022). Preferenca istraživača da se fokusiraju na ULS od četiri reči može da se objasni iz dva ugla. Naime, ULS od četiri reči su pet puta više zastupljeni od ULS od pet reči, dok često u sebi sadrže ULS od tri reči (Hyland, 2011). Štaviše, ULS od četiri reči su pogodniji za strukturalnu i funkcionalnu analizu od ULS od tri reči zbog veće strukturalne i funkcionalne raznolikosti (Hyland, 2011). Mali broj istraživanja se bavio analizom dužih ULS, njihovih frekvencija, funkcija i struktura u različitim žanrovima i registrima (Biber et al., 1999; Appel & Wood, 2016; Esfandiari & Barbary 2017). Esfandiari i Barbary (2017) kritikuju fokus istraživanja na ULS od četiri reči. Oni smatraju da preveliki naglasak na samo jednoj veličini ULS može biti ograničavajući, jer ne može jedan

⁹ „a string of uninterrupted word-forms”

¹⁰ „all lexical bundles are n-grams, recurrent groups of three or more words, but not all n-grams are lexical bundles”

predstavnik grupe da prenese bogatstvo značenja i kompleksnost pojave koje može cela grupa. Usredsređivanjem samo na ULS od četiri reči ignorišu se svi duži ULS i pojavljuje se rizik prekomernog pojednostavljenja pojave ULS, a povlači i neosnovanu ideju da duži leksički spojevi nisu dovoljno podobni za istraživanje (Esfandiari & Barbary, 2017).

ULS se izdvajaju primenom minimalne frekvencije i minimalne disperzije. Učestalost pojavljivanja je glavna karakteristika koja definiše ULS (Kortez, 2004). Minimalna frekvencija za izdvajanje ULS je suštinski arbitrarna, ali su se u dosadašnjim istraživanjima određene minimalne frekvencije pokazale kao učestalije. Važno je napomenuti da se kraći ULS češće pojavljuju u tekstu od dužih te je neophodno modifikovati kriterijume za izdvajanje ULS, odnosno postavljati manju minimalnu frekvenciju pri izdvajanju dužih ULS (Hyland, 2008; Esfandiari i Barbary, 2017). U dosadašnjim istraživanjima minimalna frekvencija se kretala od deset ponavljanja u milion reči (Bajber i sar., 1999) preko 20 ponavljanja u milion reči (Hyland 2008; Perez-Llantada, 2014; Güngör, 2016; Lazić, 2017), 40 ponavljanja u milion reči (Pan i sar., 2016; Al Fajri et al., 2020), pa čak i do 50 ponavljanja u milion reči (Berūkštienė, 2017). Minimalna frekvencija od 20 ponavljanja u milion reči se smatra blažim kriterijumom za izdvajanje ULS, dok se minimalna frekvencija od 40 ponavljanja u milion reči smatra strožim (Adel & Erman, 2012). Pored odabira dužine ULS koji će se ispitivati, na odabir minimalne frekvencije utiče i veličina korpusa – pokazalo se da manji korpusi imaju tendenciju da generišu disproporcionalno veliki broj ULS (Begsten, 2019; Pan et al., 2020). Begsten (2019) objašnjava navedenu tendenciju *Zipfianskom prirodom ljudskog jezika* (engl. *Zipfian nature of the human language*). Naime, u svakom jeziku manji broj reči se ponavlja često a veliki broj reči se ponavlja retko. Pa su, stoga, retke reči u malom korpusu previše zastupljene što dovodi do većeg broja ULS. Zbog činjenice da se u manjim korpusima identifikuje disproporcionalno veći broj ULS, istraživači biraju stroži kriterijum za izdvajanje ULS kada rade sa manjim korpusima (što za istraživanja ULS podrazumeva korpus od pola miliona reči; Güngör, 2016). Istraživači pri odabiru kriterijuma za identifikaciju uzimaju u obzir i koliko su sâm registar i žanr koji se ispituju podložni formulaičnom jeziku. Za govor ili pravne tekstove koje karakteriše velika zastupljenost formulaičnog jezika obično se postavljaju stroži kriterijumi za izdvajanje ULS (Lazić, 2017; Berūkštienė, 2017; Ebellling, 2022). Prikupljanje podataka zasnovanih na frekvenciji je beskorisno bez adekvatne interpretacije dobijenih podataka (Biber et al., 2004). Međutim, prednost ove vrste analize je to što identifikuje elemente jezičke produkcije koje istraživači ne bi u suprotnom primetili. Podaci o frekventnosti su dodatno korisni kao jedan od indikatora do koje je mere sekvenca od nekoliko reči sačuvana u kognitivnom sistemu kao unapred određena celina (engl. *prefabricated chunk*). Postoji veća verovatnoća da su sekvence od nekoliko reči koje su veoma učestale unapred određene celine iako frekventnost nije jedini ukazatelj (Biber et al., 2004).

Minimalna disperzija se postavlja da izdvojeni ULS ne bi bili karakteristika određenog govornika ili teksta u korpusu. Minimalna disperzija je takođe arbitrarna. U prethodnim istraživanjima najčešće se susreće minimalna disperzija od pet ponavljanja (Salazae, 2014; Pan i sar., 2016; Esfandiari i Barbary, 2017) ili minimalna disperzija u 10% tekstova (Hyland, 2008; Perez-Llantada, 2014; Güngör, 2016; Al Fajri et al., 2020). Naravno, pošto je određivanje minimalne disperzije arbitrarno i pošto ne postoji postavljena konvencija za određivanje minimalne frekvencije, u istraživanjima može da se naiđe na veliku razliku u postavljenoj minimalnoj disperziji, kao što je slučaj i za minimalnu frekvenciju. Pa tako u istraživanju Ebeling (2022) nailazimo na minimalnu disperziju u 25% tekstova, što je primer atipično strogog kriterijuma za izdvajanje ULS.

Važno je istaći još jednu terminološku začkoljicu koja je prisutna u istraživanju ULS. Naime, u korpusnoj lingvistici termin *tip* (engl. *type*) označava broj jedinstvenih reči, dok termin *token* (engl. *token*) označava svako pojavljivanje te reči (McEnery & Hardie, 2011). Međutim, u istraživanjima ULS terminom *tip* se označava jedinstven ULS, dok termin *token* predstavlja svako pojedinačno ponavljanje tog ULS u korpusu. Tako, ako uzmemo kao primer ULS *in the context of the – tip* je

specifičan oblik, tj. jedinstven ULS *in the context of the*, koji se pojavljuje u korpusu naučnih radova iz oblasti obrazovanja na engleskom jeziku koje su pisali izvorni govornici rumunskog jezika 137 puta, odnosno ima 137 *tokens*.

ULS nisu privukli pažnju istraživača svojim upečatljivim prisustvom u jeziku već zbog svoje uloge u podržavanju funkcionalne koherentnosti u različitim tekstovima i žanrovima (Biber i sar., 2004; Corte, 2004; Hyland, 2008a, 2008b, prema Esfandiari & Barbary, 2017). Pretpostavlja se da su ULS u mozgu čuvani/skladišteni kao neanalizirani spojevi od nekoliko reči (engl. *unanalysed multi-word chunks*) a ne kao produktivne gramatičke konstrukcije (Biber i sar., 2004: 400, prema Esfandiari & Barbary, 2017), da se uče kao celine a ne kao nizovi pojedinačnih reči (Perez-Llantanda, 2014) i da su automatski preuzeti (engl. *retrived*), što čini govor i pisanje fluentnijim. Durrant (2017, prema Esfandiari & Barbary, 2017) smatra da su ULS pogodna i interesantna oblast za istraživanje zbog tri glavne karakteristike ovih leksičkih pojava: 1. njihova identifikacija je automatska; 2. poseduju definisane funkcionalne uloge; 3. osetljivi su na razlike između drugačijih tipova i žanrova tekstova.

Uspešno korišćenje ULS koji imaju funkciju označitelja metadiskursa u tekstu odaje utisak da autor poseduje odgovarajuće komunikativne kompetencije u datoj disciplini (Hyland, 2008). Ovo je posebno značajno za istraživače koji nisu izvorni govornici engleskog jezika, jer se označitelji metadiskursa disciplinama menjaju u zavisnosti od jezika na kojem se piše (Bercuči, 2020).

Altenberg (1998) je verovatno prvo istraživanje koje se bavilo učestalim kombinacijama reči u engleskom jeziku, dok je prvo istraživanje u kojem se spominje termin *učestali leksički spojevi* (engl. *lexical bundles*) istraživanje Bajbera i saradnika (Biber et al., 1999). Iako su istraživanja slična i objavljena u kratkom vremenskom razmaku, u akademskoj zajednici zaživelo je Bajberovo istraživanje koje je postalo osnova za dalja ispitivanja u ovoj oblasti.

U Altenbergovom (Altenberg, 1998) istraživanju proučavala se frazeologija govornog engleskog jezika tako što su se ispitale učestale kombinacije reči (engl. *recurrent word-combinations*) u korpusu od pola miliona reči – *London-LundCorpus of Spoken English*. Za određivanje učestalih kombinacija reči koristila se minimalna frekvencija od deset ponavljanja u korpusu, a dužina kombinacija se kretala od tri do pet reči. Izdvojeno je 470 tipova i 6.692 tokena, koji su podeljeni u strukturalne i funkcionalne kategorije koje kasnije nisu zaživele kao standard u ovom tipu istraživanja. Strukturalno učestale kombinacije reči podeljene su u tri glavne grupe sa podgrupama, a svaka od strukturalnih podgrupa je dobila funkcionalnu raspodelu učestalih kombinacija reči, osim podgrupe *nepotpune fraze*, za koju je dato objašnjenje kolokacijskih okvira u kojima mogu da se nađu. Radi preglednijeg uvida u raspodelu, napravljen je grafikon (v. Grafikon 1), na kojem se vidi i strukturalna i funkcionalna raspodela učestalih kombinacija reči koje je predložio Altenberg (1998).

Bajber i sar. (1999) su identifikovali najčešće učestale leksičke spojeve u govoru i akademskom pisanju, koje su definisali kao spojeve reči koji su se pojavili barem deset puta u milion reči u ciljanom registru, i koji su prisutni u barem pet različitih tekstova. U tom istraživanju predlažu strukturalnu taksonomiju koja je prihvaćena u mnogim daljim istraživanjima, ali koja je i često modifikovana za potrebe specifičnih istraživanja (Salazar, 2014; Lazić, 2017, Güngör, 2017).

U istraživanju Bajber i saradnici predlažu odvojene taksonomije za registar govora i akademskih tekstova. U ovom radu ćemo predstaviti predloženu strukturalnu taksonomiju za akademske tekstove (v. Tabelu 2), s obzirom na činjenicu da je fokus disertacije na naučnim radovima iz oblasti obrazovanja.

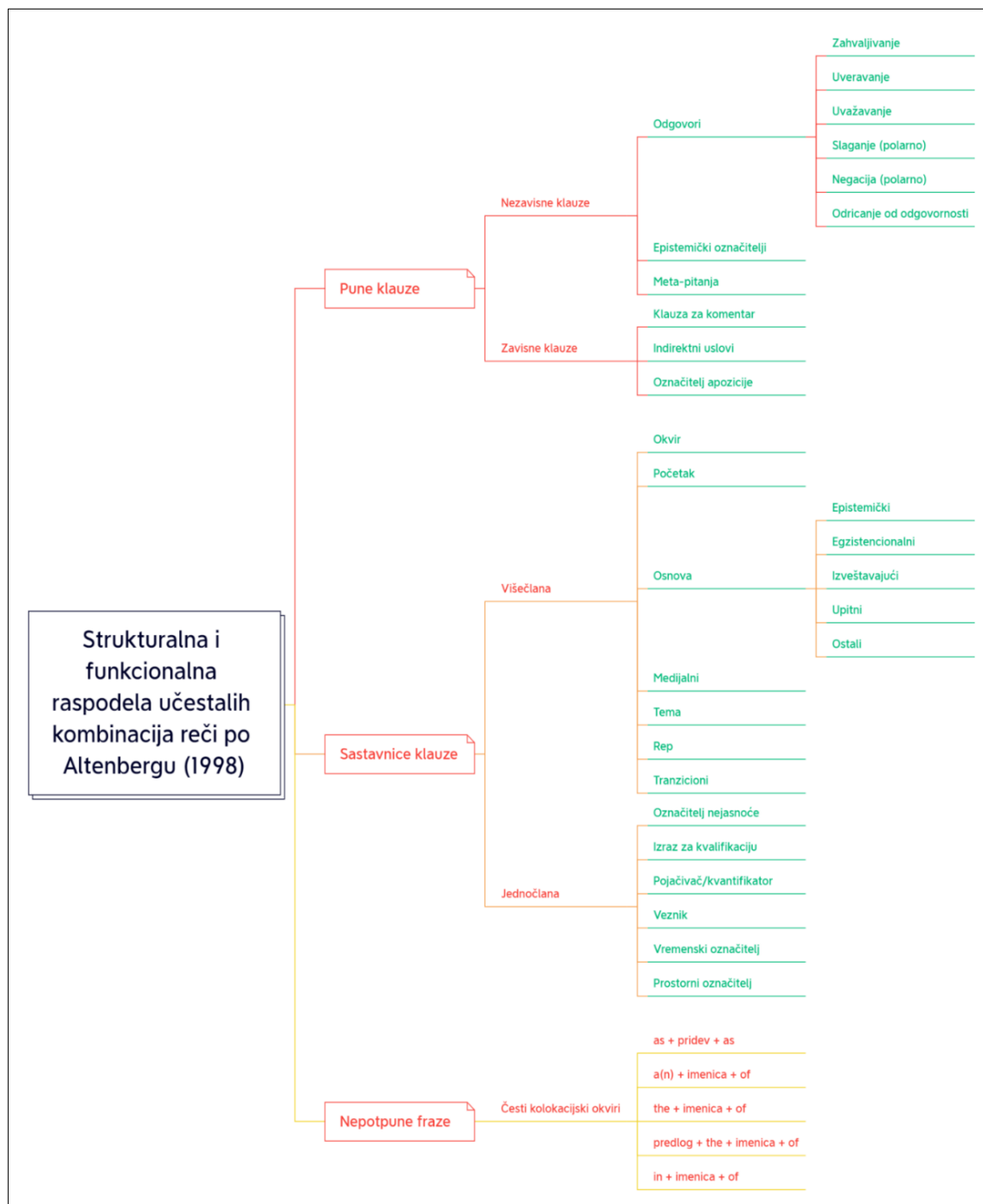
Tabela 2. *Strukturalna taksonomija za akademske tekstove koju su predstavili Bajber i saradnici (1999)*

Struktura	Primeri
<i>Imenička fraza sa of fragmentom</i>	<i>the end of the, the start of the, the level of the</i>
<i>Imeničke fraze sa drugim modifikatorima</i>	<i>the way in which, the fact that the, such a way that</i>
<i>Predložka fraza sa of fragmentom</i>	<i>as a function of, by the end of, in the case of</i>
<i>Druge predložke fraze</i>	<i>at the same time, in relation to the, with respect to the</i>
<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	<i>it is important to, it is likely that, it is to be</i>
<i>Glagol u pasivu i fragment predložke fraze</i>	<i>is based on the, is referred to as, is known as the</i>
<i>Kopula be i imenička/pridevska fraza</i>	<i>is part of the, is the same as, is due to the</i>
<i>(Glagolska fraza)+ fragment that klauze</i>	<i>has been shown that, does not mean that</i>
<i>(Glagol/pridev +) fragment to klauze</i>	<i>are likely to be, can be used to, will be able to</i>
<i>Fragment adverbijalne klauze</i>	<i>as shown in figure, as we have seen, as we shall see</i>
<i>Zamenica/imenička fraza+be (+...)</i>	<i>this is not to, there was no significant</i>
<i>Drugi izrazi</i>	<i>may or may not, as well as the</i>

Bajber i saradnici su predložili i funkcionalnu taksonomiju u kasnijim istraživanjima (Bajber i sar., 2004; Biber & Barbieri, 2007), koja je služila kao polazna tačka za funkcionalne analize u mnogim budućim istraživanjima sa varijacijama originalne taksonomije. Dve najznačajnije modifikacije taksonomije predstavljene su u istraživanjima Hajland (Hyland, 2008) i Čen i Bejker (Chen & Baker, 2010). Istraživanje Čen i Bejker i njihove predložene izmene taksonomije postali su jedno od uticajnijih istraživanja pri ispitivanju upotrebe ULS kod izvornih i neizvornih govornika. Izmenjena taksonomija koja se koristila u njihovom istraživanju predstavljena je u Tabeli 3. Prema njima *referentni ULS* ističu fizičke elemente, apstraktne elemente ili tekstualni kontekst kako bi ih identifikovali ili istakli neki od njihovih atributa, *ULS stava* se koriste da iskažu autorovu evaluaciju predloga u smislu sigurnosti ili nesigurnosti, dok se *ULS za organizaciju diskursa* koriste za strukturisanje teksta.

Tabela 3. *Funkcionalna taksonomija predložena u istraživanju Čen i Bejker (2010)*

Funkcionalna grupa	Funkcionalna podgrupa	Primeri
<i>Referentni ULS</i>	<i>Ograničavajući ULS</i>	<i>in the context of, the nature of the</i>
	<i>ULS za kvantifikaciju</i>	<i>a wide range of, in a number of</i>
	<i>Vreme/mesto/tekstualno deiktični</i>	<i>are shown in figure, at the same time</i>
<i>ULS stava</i>	<i>Epistemički</i>	<i>are more likely to, it can be argued</i>
	<i>Obaveza/uputstvo</i>	<i>it is necessary to, that need to be</i>
	<i>Mogućnost</i>	<i>it is difficult to, to be able to</i>
<i>ULS za organizaciju diskursa</i>	<i>ULS za uvođenje teme</i>	<i>last but not least, in this essay I</i>
	<i>ULS za pojašnjenje teme</i>	<i>on the other hand, this is due to</i>
	<i>Inferencijalni</i>	<i>as a result of, in view of the</i>
	<i>ULS za identifikaciju/fokusiranje</i>	<i>one of the most, we can see that</i>



Grafikon 1. *Strukturalna i funkcionalna raspodela učestalih kombinacija reči po Altenbergu*¹¹

Sa druge strane, istraživanje Hajland (2008) i njegova modifikovana funkcionalna taksonomija postali su uticajni u branši istraživanja koja su ispitivala upotrebu ULS u različitim oblastima u akademskom žanru (prevashodno sa fokusom na naučne radove). Hajland je načinio modifikacije u taksonomiji kako bi se što preciznije opisale funkcije koje su specifične za naučne radove, v. Tabelu 4.

¹¹Grafikon je napravila autorka disertacije na osnovu informacija iz istraživanja Altenberg (1998).

Tabela 4. Funkcionalna taksonomija predložena u istraživanju Hajland (2008)

Grupa	Podgrupa	Primeri
ULS orijentisani ka istraživanju	ULS za lokaciju	<i>at the same time, in the present study</i>
	ULS za proceduru	<i>the use of the, the role of the</i>
	ULS za kvantifikaciju	<i>the magnitude of the, one of the most</i>
	ULS za opis	<i>the structure of the, the size of the</i>
	ULS teme	<i>in the Hong Kong, the currency board system</i>
ULS orijentisani ka tekstu	Tranzicioni signali	<i>on the other hand, in addition to the</i>
	Rezultativni signali	<i>as a result of, it was found that</i>
	Strukturni signali	<i>in the present study, in the next section</i>
	Ograničavajući signali	<i>in the case of, with the exception of</i>
ULS orijentisani ka učesniku	ULS stava	<i>are likely to be, may be due to</i>
	ULS uključivanja	<i>it should be noted that, as can be seen</i>

Navedeni podaci su nam ukazali na same početke izučavanja ULS i na najzastupljenije taksonomije za strukturalnu i funkcionalnu raspodelu ULS, međutim, da bismo bolje razumeli ULS kao pojavu i njihov značaj, neophodno je da šire sagledamo razvoj istraživanja i tema koja se bave ULS. Pregled koji sledi nije sveobuhvatan, već je namenjen da čitaocu dâ uvid u neke od najzastupljenijih pravaca u istraživanjima ULS, sa fokusom na one koji su povezani sa temom ove disertacije.

Kao što je već navedeno, prva istraživanja ULS ispitivala su njihovu upotrebu i zastupljenost u govoru i u pisanom jeziku (Altenberg, 1998; Bajber i sar., 1999). Istraživanja koja su se nadovezala na ta istraživanja prvo su ispitala i uporedila upotrebu ULS u različitim žanrovima unutar dva registra. Jedno od istaknutijih istraživanja iz te kategorije je istraživanje Bajbera i sar. (Biber et al., 2004), u kojem se ispitivala razlika u upotrebi ULS u konverzaciji, nastavi, univerzitetskim udžbenicima i objavljenim naučnim radovima. Ta prva istraživanja su pokazala da postoje značajne razlike u ULS koji se koriste u govoru i pisanom jeziku pa čak i u različitim žanrovima po:

- *broju izdvojenih ULS* – veći broj ULS se koristi u govoru nego u pisanom jeziku, dok se u nastavi koristi najveći broj ULS;
- *strukturalnoj raspodeli korišćenih ULS* – pisani akademski tekst karakteriše najzastupljenija upotreba imeničkih struktura i fragmenata sa predloškom frazom, dok se u govoru najviše koriste glagolske strukture;
- *funkcionalnoj raspodeli ULS* – u pisanim akademskim tekstovima više se koriste *referentni ULS*, dok se u konverzaciji najviše koriste *ULS stava*.

Danas postoje istraživanja koja su ispitala zastupljenost ULS u velikoj raznolikosti kako pisanih tako i govornih žanrova, na primer, u korpusima studentskog govora kod izvornih i neizvornih govornika engleskog jezika (Crossley & Salsbury, 2011), diskusijama studenata u grupnom radu (Heng et al., 2014), akademskim predavanjima i *TED* govorima (engl. *TED talks* – Liu & Chen, 2020), govoru doktora u seriji *Uvod u anatomiju* (engl. *Gray's Anatomy* – Panthong & Poonpon, 2020), u različitim pravnim žanrovima (Breeze, 2013), u javnim saopštenjima i poslovnim vestima (De Cock & Granger, 2021), književnim delima (Liu & Afzaal, 2021), vestima koje su pisali neizvorni govornici engleskog jezika (Fadhilah, 2020), podkastima (Susianti & Fitriati, 2021), prevedenim tekstovima uporedno sa neprevedenim u potrazi za prevodnim univerzalijama (Berūkštienė, 2017; Biel, 2017) itd.

Akademski žanrovi su postali najpopularniji za istraživanje upotrebe ULS i dobili disproporcionalno mnogo pažnje uporedno sa drugim žanrovima i oblastima (Greaves & Warren, 2010). Postoji nekoliko potencijalnih objašnjenja za popularnost ispitivanja ULS u akademskom

diskursu. Kako ističe Rej (Wray, 2006: 593, prema Adel & Erman, 2012) „ponavljane kombinacije reči ne samo da doprinose idiomatičnosti, već doprinose i pokazivanju pripadnosti specifičnoj diskursnoj zajednici: kada govorimo mi biramo određene fraze za koje mi smatramo da su u vezi sa određenim vrednostima, stilovima i grupama”¹². U slučaju studenta koji je nov član akademskog sveta, ili naučnika koji se upušta u pisanje naučnih radova na engleskom jeziku kako bi postao deo svetske akademske zajednice, učenje upotrebe adekvatnog frazeološkog jezika postaje označitelj identiteta. Akademsko pisanje podrazumeva stvaranje „diskursnog ja” (engl. *discoursal self* – Ivanović, 1998: 336, prema Pang, 2010) ili „profesionalno prihvatljive persone” (Hajland, 2004) koja omogućuje piscu da iskaže sadržaj koji želi, dok istovremeno ostvaruje pravo da ga čuje akademska zajednica upotrebom socio-kulturno prikladnog jezika. Međutim, neizvorni govornici često nisu svesni socio-kulturnih očekivanja u akademskom diskursu koji je na globalnom nivou pod jakim uticajem anglofone kulture (Muresan & Pérez-Llantada, 2014). Kako bi se naveden problem izbegao, profesori engleskog jezika imaju zadatak da pomognu studentima u razumevanju očekivanja čitaoca, kao i da nauče normativne aspekte diskursnih praksi (Matsuda, 2001, prema Pang, 2010). Čitalac stvara sliku o tome da li je pisac verodostojan ili ne na osnovu percepcije *diskursnog ja/profesionalne persone*, odnosno akademskog glasa. Po Hajlandu (2004): „Ovo stavlja jezik, odnosno jezik koji prenosi kredibilitet u centar učenja kako da se postane član disciplinarne zajednice”¹³, pošto uspeh u akademskoj zajednici zavisi od toga koliko smo vešto stvorili naš akademski glas. Sve nas ovo vraća na upotrebu ULS jer su oni specifični za različite akademske žanrove, kao i naučne oblasti unutar tih žanrova (Hajland, 2008; Durrant, 2017). ULS, kao i drugi oblici formulaičkog jezika, olakšavaju procesuiranje informacija zato što predstavljaju unapred oformljene sekvence koje se ponašaju kao jedna celina (Bajber i sar., 2004). Stoga, pravilno korišćenje ULS pomaže studentima i članovima akademske zajednice koji žele da objavljuju svoje radove na engleskom jeziku da stvore kredibilan akademski glas u svojoj disciplini (Pang, 2010). Ispitivanje ULS u različitim žanrovima i oblastima pruža uvid u leksičko-gramatičke načine za iskazivanje značenja koji odobrava određena akademska zajednica. Navedeni uvidi su kasnije na raspolaganju profesorima engleskog jezika i članovima akademske zajednice koji mogu da koriste ta znanja za razvijanje svog akademskog glasa ili akademskog glasa svojih studenata (Hajland, 2012; Nam, 2017). Bliska povezanost istraživanja upotrebe ULS u akademskim žanrovima sa podučavanjem akademskog engleskog jezika je jedan od mogućih razloga za veliku popularnost istraživanja ULS u akademskim žanrovima. Istraživanje u ovoj disertaciji se svakako uklapa u navedenu premisu jer je upravo povezivanje istraživačkog rada sa podučavanjem engleskog jezika glavna svrha.

U okviru ispitivanja upotrebe ULS u akademskom diskursu pojavilo se nekoliko pravaca koji su se istakli kao posebno popularni. Neki od tih pravaca su:

- ispitivanje upotrebe ULS u naučnim radovima iz različitih naučnih disciplina (npr. Hyland, 2008; Gilmore & Millar, 2018; Ren, 2021);
- ispitivanje upotrebe ULS u različitim segmentima naučnih radova (npr. Cortes, 2013; Abdollahpour & Gholami, 2018; Casal et al., 2021);
- ispitivanje upotrebe ULS u produkciji izvornih i neizvornih govornika – u produkciji profesionalaca (naučni radovi – npr. Salazar, 2014; Pan et al., 2016; Lazić, 2017; Al Fajri et al., 2020), studentskoj produkciji (npr. Nam, 2017; Bychkovska & Lee, 2017; Shin, 2019; Appel & Murray, 2020), kao i komparaciji studentske produkcije i produkcije profesionalaca (npr. Levy, 2003; Chen & Baker, 2010; Cortes, 2004);

¹²[...] „recurrent word combinations do not only contribute to idiomacity, but also contribute to demonstrating membership in a specific discourse community: when we speak, we select particular turns of phrase that we percieve to be associated with certain values, styles and groups”

¹³ „This places language, or rather language that carries credibility at the heart of learning to become a member of a disciplinary community.”

- uticaj transfera na upotrebu ULS kod neizvornih govornika (npr. Paquot, 2013; Güngör & Uysal, 2020; Kostromitina, 2022);
- ispitivanje upotrebe ULS u produkciji učenika sa različitim nivoom znanja (npr. Staples et al., 2013; Chen & Baker, 2016; Garner et al., 2020).

S obzirom na činjenicu da se u istraživanju u ovoj disertaciji bavimo upotrebom ULS u naučnim radovima kod izvornih i neizvornih govornika, detaljnije ćemo sagledati potencijalne razloge zbog kojih su istraživanja upotrebe ULS u ovom akademskom žanru stekla toliku popularnost, kao i zašto je važno ispitati upotrebu ULS kod izvornih i neizvornih govornika. Detaljan pregled relevantnih prethodnih istraživanja biće dat u kasnijim poglavljima.

Razlog za odabir naučnih radova kao predmet istraživanja može se objasniti činjenicom da su u svetu akademskog diskursa naučni radovi definitivno „najistaknutiji žanr” (Hyland, 2009: 67, prema Güngör, 2019) i „master narativ našeg vremena” (Montgomery, 1996, prema Güngör, 2019). Pedagoški zahtevi engleskog za akademske svrhe (*English for Academic Purposes*, EAP) i engleskog za posebne namene (*English for Specific Purposes*, ESP) povećale su interesovanje za analizu diskursa i podrobno istraživanje lingvističkih struktura u naučnim radovima. Kada spojimo te informacije sa činjenicom da je strateško i vešto korišćenje metadiskursa u akademskom pisanju povezano sa time koliko će se ozbiljno autor i njegove tvrdnje shvatiti (Hyland, 2005, prema Güngör, 2019), jasna nam je popularnost ispitivanja upotrebe ULS baš u naučnim radovima.

Zbog sve većeg uticaja procesa globalizacije na akademsku zajednicu u kojoj se naučnici sve više uslovljavaju da objavljuju svoje radove na globalnom tržištu gde engleski jezik preovlađuje, pitanje pisanja radova na engleskom jeziku i objavljivanje radova na engleskom jeziku od strane neizvornih govornika engleskog jezika postaje sve relevantnije pitanje (Filipović Vučo, 2013; Igrutinović, 2021). Zbog uticaja adekvatne upotrebe ULS na uspešnost demonstracije pripadnosti određenoj akademskoj zajednici odnosno preciznije – delatnoj zajednici¹⁴, ispitivanje razlika u upotrebi ULS kod izvornih i neizvornih govornika u naučnim radovima postaje popularno. Iznenadujuće je da je mnogo razvojnih istraživanja mešalo uticaj ekspertize i uticaj maternjeg jezika (Pan et al., 2016: 61). Čini se nekontroverznim da pridošlice svetu akademskog pisanja, nezavisno od toga da li im je engleski maternji ili strani jezik, moraju da nauče konvencije naprednog akademskog pisanja, što uključuje i adekvatno korišćenje učestalih leksičkih spojeva (Cortes, 2004; Pan et al., 2016). Međutim, prethodna istraživanja ukazuju na to da izvorni govornici imaju prednost u tom zadatku, što dovodi do stabilnih razlika između izvornih i neizvornih grupa studenata koji su inače na istom nivou studija, kao i između grupa izvornih i neizvornih istraživača u različitim oblastima (Chen & Baker, 2010; Ädel & Erman, 2012; Pan et al., 2016; Bychkovska & Lee, 2017; Güngör, 2019; Nam, 2017).

Obrazovne institucije dobile su zadatak da pruže rešenja za uspešnije usvajanje veština engleskog jezika prilagođene za specifične grupe učenika (Dinca & Chitez, 2021). Kao pokret zasnovan na potrebama učenika (*learner needs-based approach*) oblast podučavanja engleskog jezika za specifične namene (ESP) primenjuje širok spektar metoda podučavanja koje odgovaraju na širok spektar učenikovih potreba na makro (retorički, ceo tekst) nivou i na mikro (leksiko-gramatičkom) nivou kako u govornim tako i u pisanom registru ili bilo koje druge aspekte primenjene lingvistike koji su korisni za učenike (Belcher, 2009, prema Dinca & Chitez, 2021). Na leksičkogramatičkom nivou, frazeologija igra esencijalnu ulogu s obzirom na to da ovladavanje akademskim frazama pomaže učenicima u njihovom nastojanju da postanu stručnjaci u akademskom pisanju na engleskom jeziku umesto pisaca koji nalikuju na izvorne govornike (*native-like writers* – Romer and Arbor, 2009, prema Dinca & Chitez, 2021). Međutim, i pored jasne logike koja podržava

¹⁴Delatna zajednica (engl. *community of practice*) se definiše kao grupa pojedinaca koji zajedno delaju sa namerom postizanja zajedničkih društvenih ciljeva (Eckert, & McConnell-Ginet, 1992, prema Filipović, 2018).

aktivno podučavanje ULS i dalje su eksperimentalna istraživanja koja se bave podučavanjem ULS čine relativno retkim (npr. Cortes, 2006; Erikson, 2012; Neely & Cortes, 2009, prema Gungör, 2016). Ista se situacija može uočiti i za istraživanja koja se bave konkretnim ispitivanjem pedagoške vrednosti ULS (npr. Byrd & Coxhead, 2010; Chang & Kuo, 2011; Kennedy, 2003, prema Gungör, 2016). Zbog navedenih ograničenja u istraživanjima ULS, Elis i saradnici (Ellis et al., 2008) predlažu da istraživači i profesori engleskog jezika:

- identifikuju formule koje su od posebnog značaja za učenike/studente;
- razviju strategije za optimalnu integraciju formulaičnog jezika u kurikulum;
- razviju razumevanje psiholingvistike formulaičnog jezika kod izvornih i neizvornih govornika i faktora koji utiču na lakoću učenja i procesuiranje sa ciljem dostizanja fluentnosti;
- da se ustanovi koje bi formule trebalo da akcentujemo u nastavi sa učenicima/studentima u različitim fazama razvoja i sa različitim potrebama.

Prateći navedene sugestije, istraživanje u ovoj disertaciji daće pregled karakteristika upotrebe ULS u naučnim radovima u oblasti obrazovanja kod izvornih govornika engleskog jezika, neizvornih govornika kojima je maternji jezik srpski i neizvornih govornika kojima je maternji jezik rumunski. Sagledanjem rezultata o zastupljenosti ULS kod izvornih govornika engleskog jezika, kao i sagledanjem strukturalne i funkcionalne raspodele ULS, dobijamo podatke o normama upotrebe ULS u navedenom akademskom žanru i oblasti. Poređenjem zastupljenosti ULS, strukturalne i funkcionalne raspodele ULS između izvornih i neizvornih govornika, možemo da vidimo na koji način neizvorni govornici odstupaju od normi i izdvojimo metodičke smernice za uspešnije podučavanje ULS kod dve navedene grupe neizvornih govornika engleskog jezika. Poređenjem upotrebe ULS kod dve grupe neizvornih govornika engleskog jezika i primenom procesa identifikacije uticaja transfera na upotrebu ULS, identifikovaće se ULS koji su specifični za izvorne govornike srpskog jezika i za izvorne govornike rumunskog jezika, kao i one koji su potencijalno zajednički za neizvorne govornike u globalu nezavisno od maternjeg jezika, što će dalje pomoći profesorima da prilagode svoju nastavnu praksu različitim grupama neizvornih govornika.

1.2. Engleski jezik kao *lingua franca* u nauci i neophodnost objavljivanja na engleskom jeziku u akademskoj zajednici

Kada sagledamo naučni razvoj iz perspektive prevoda, jasan je značaj postojanja jezika koji služi kao *lingva franka* (engl. *lingua franca* < it. *lingua franca*). Kroz istoriju, različiti jezici imali su tu ulogu: grčki, arapski, kineski, latinski i sada – engleski jezik. Postojanje jezika koji služi kao *lingva franka* ubrzava naučni razvoj, jer omogućava da se veliki broj znanja i materijala iz različitih kultura i konteksta slije u jedno mesto i naknadno ponovo distribuiraju (Montgomeri, 2010).

Nakon Drugog svetskog rata, zbog različitih socioekonomskih i političkih faktora, kao i velikog obima istraživanja koji proističu iz SAD, engleski jezik postaje dominantan u sferi nauke (Crystal, 2013, prema Igrutinović, 2021). U prethodnom periodu situacija je bila poprilično drugačija. Naime, u XV veku tokom popularizacije naučnog znanja u Evropi koristili su se francuski, nemački, italijanski, ruski i engleski kao vodeći jezici u nauci. Tokom XX veka taj broj jezika se suzio i centralno mesto u nauci su delila četiri jezika: francuski, nemački, ruski i engleski jezik (Igrutinović, 2021).

Danas je neosporna činjenica da engleski jezik služi kao *lingva franka* u svim domenima, a posebno u naučnom domenu. U naučnoj zajednici postoje dve struje: jedni podržavaju „jezički imperijalizam” i smatraju da je neophodno imati jedan dominantan jezik u nauci radi jednostavnijeg i efikasnijeg prenošenja znanja, a drugi se zalažu za jezičku ekologiju i smatraju da engleski jezik

nije dovoljan (Muresan & Pérez-Llantada, 2014; Luo i Hajland, 2019; Igrutinović, 2021). Zasigurno je da obe strane imaju validne argumente, kao i da trenutna zastupljenost jednog – engleskog jezika – ima mnoge prednosti kada je reč o širenju naučnih znanja, ali ima i određene manjkavosti.

Glavna prednost postojanja jezika koji je lingva franka u nauci jeste da učenjem samo jednog jezika naučni radnici mogu da razumeju najveći procenat najrelevantnijih istraživanja i međusobno komuniciraju u međunarodnoj zajednici (De Swaan, 2001, prema Igrutinović, 2021). Nedostaci, s druge strane, jesu raznoliki i mnogobrojni. Neanglofoni istraživači su primorani da objavljuju na engleskom jeziku što zahteva visok nivo poznavanja engleskog jezika, a to dalje iziskuje vreme za usavršavanje jezičkih veština što često iziskuje i dodatna finansijska sredstva. Članovi akademske zajednice koji nemaju dovoljno visok nivo znanja engleskog jezika oslanjaju se na usluge posrednika (engl. *mediators*) koji im pomažu u procesu objavljivanja radova. Prevažodno su to usluge koje se odnose na sâm tekst rada – prevođenje i uređivanje, iako u nekim zemljama kao što je Kina postoje i agencije koje nude posredništvo u samom procesu objavljivanja (slanje radova relevantnim naučnim časopisima, komunikacija sa uredničkim timom itd. – Luo i Hajland, 2019). Korišćenje navedenih usluga posrednika iziskuje ulaganje dodatnog truda i prevažodno finansijskih sredstava. Ukratko, dolazi do nejednakog položaja anglofonih i neanglofonih članova akademske zajednice. Neanglofoni istraživači nailaze na veći broj prepreka u procesu objavljivanja svojih radova. Kako je Smiljana Igrutinović objasnila:

Višedecenijsko širenje engleskog jezika u nauci implicira nametanje anglosaksonskog naučnog diskursa, kulturnih modela, istraživačkih paradigmi i izbora tema. Nova hijerarhija sa engleskim jezikom na vrhu, uključujući njegove diskursne strukture i kulturne modele, predstavlja snažan instrument a ujedno i ishod šireg procesa dinamike razvoja nauke (2021: 40).

Neke akademske zajednice (mađarski naučnici iz oblasti prirodnih nauka, španski naučnici iz oblasti društvenih nauka i francuski naučnici iz oblasti humanistike) počele su da objavljuju radove prevažodno ili čak isključivo na engleskom jeziku, zato što nacionalni sistemi evaluacije istraživačkog rada i akademskog karijernog napredovanja favorizuju časopise koji se objavljuju na engleskom jeziku ili zato što se tako može dopreti do šire internacionalne publike (Murešan i Perez-Ljatanda, 2014). Naučnici poljskih akademskih zajednica iz oblasti primenjene lingvistike, psihologije, medicine i filologije (Duszak & Lewkowicz, 2012, prema Murešan i Perez-Ljatanda, 2014) i naučnici danskih akademskih zajednica iz oblasti društvenih i prirodnih nauka (Phillipson & Skutnabb-Kangas, 1999, prema Murešan i Perez-Ljatanda, 2014) objavljuju radove na engleskom jeziku kako bi došli do šire internacionalne publike, ali i dalje objavljuju radove i na svojim lokalnim jezicima za nacionalne publikacije i održavanje kontakta sa nacionalnom akademskom zajednicom. Umnogome se drugačija situacija može zapaziti u Finskoj, Švedskoj, Norveškoj i Švajcarskoj, gde se engleski jezik koristi za predavanja i objavljivanja, ali sa negativnim posledicama na zastupljenost lokalnih jezika (Heberland, 2005 prema Murešan i Perez-Ljatanda, 2014).

Jezička politika Evropske unije podržava jezički i kulturni pluralizam, te je isti postavila kao cilj za sve građane EU. Od njih se očekuje da znaju i mogu da koriste u svim domenima svoj maternji jezik, jezik koji je lingva franka i barem još jedan strani jezik (Filipović i Vučo, 2013). Međutim, postoji neusklađenost između jezičkih politika i prakse jer je i u Evropskoj uniji engleski jezik takođe postao vodeći jezik u akademskim krugovima, obrazovnim/naučnim institucijama, programima i mrežama (Murešan i Perez-Ljatanda, 2014; Igrutinović, 2021). Navedeni položaj engleskog jezika se dodatno učvrstio uvođenjem Bolonjske deklaracije 1999. godine sa ciljem da se formira evropski prostor za visoko obrazovanje, što je kao rezultat imalo formiranje velikog broja postdiplomskih programa na engleskom jeziku (Wachter & Maiwprn, 2014 prema Igrutinović, 2021).

Proces usklađivanja sa globalnim trendovima akademskim radnicima, u svetu i u Srbiji, doneo je nove zahteve u pogledu objavljivanja naučnih radova u međunarodno prepoznatim časopisima u kojima se prati bibliometrijski sistem vrednovanja, odnosno časopisima sa SCI liste. Zbog promena u internacionalnoj strategiji u Kini 2009. godine istraživači koji objavljuju radove u naučnim časopisima na SCI listi dobijaju znatno veći broj poena koji su im neophodni u karijernom akademskom napredovanju a uz to i velikodušna finansijska sredstva (Luo i Hajland, 2019). U Rumuniji, od poslednje decenije prošlog veka, vlada očekuje da naučnici i istraživači objavljuju naučne radove u internacionalno priznatim časopisima iz Sjedinjenih Američkih Država i Zapadne Evrope (Murešan i Perez-Ljatanda, 2014). Kada se osvrnemo na srpski kontekst, možemo uvideti da se novi zahtevi spominju u Zakonu o visokom obrazovanju i standardima za akreditaciju visokoškolskih ustanova, kao i u inoviranim pravilnicima visokoškolskih institucija (Glušac i Milić, 2011). Pregledom časopisa na SCI listi lako se može uvideti da većina objavljuje radove isključivo na engleskom jeziku. Navedene kriterijume u Srbiji su odredili članovi akademske zajednice iz oblasti inženjersva i prirodnih nauka, koji imaju bolju poziciju u sferi akademskog publikovanja na engleskom jeziku od kolega iz društvenih nauka, čije su istraživačke teme češće vezane za lokalni ili regionalni kontekst, te su se prevashodno objavljivale u lokalnim časopisima (Filipović i Vučo, 2013). Kao primer nejednake pozicije istraživača iz različitih oblasti u sferi akademskog publikovanja na engleskom jeziku uzećemo naučne časopis iz Srbije koji se nalaze na SCI listi, odnosno koji su referisani u naučnoj bazi *Web of Science* (WoS). Naime, na početku 2024. godine ukupno 25 srpskih naučnih časopisa je indeksirano u navedenoj bazi, od čega samo dva časopisa ne spadaju u inženjersvo i prirodne nauke (Kobson, 2024).

Činjenica da se članovi akademske zajednice iz svih kultura podstiču na komunikaciju na engleskom jeziku, kako kroz objavljivanje naučnih radova na engleskom jeziku, kako učešćem na naučnim skupovima i projektima koji su internacionalnog karaktera, dovelo je do velikog porasta potrebe za prevodilačkim uslugama (Montgomeri, 2010). Zbog kompleksne prirode pisanja naučnih radova i raznolikih doprinosa koje različiti akteri u procesa objavljivanja imaju, stvara se potreba za uvođenje promena u dosadašnje prakse publikovanja naučnih radova (Luo i Hajland, 2019). Zbog zahtevnosti prevodilačkog procesa, prevodioci u nauci stvaraju kulturne proizvode koji se kvalifikuju kao originali na ciljnom jeziku (Montgomeri, 2010). Uprkos toj činjenici, oni su nevidljivi u procesu objavljivanja i ne navodi se da li je naučni rad preveden, niti ko je prevodilac (Luo i Hajland, 2019). Smatra se da bi trebalo ispraviti trenutno stanje i uvesti praksu navođenja prevodilaca (Montgomeri, 2010; Luo i Hajland, 2019), dok drugi čak smatraju da bi prevodioci trebalo da se navedu kao koautori jer sudeluju u pripremi finalnog oblika rada (Rennie, Yank, & Emanuel, 1997). Međutim, postoje i struje koje veruju da bi sadašnja dominacija engleskog jezika mogla da se prevaziđe širokom primenom mašinskog prevođenja putem kojeg bi informacije mogle da budu svima dostupne nezavisno od jezika na kojem je rad napisan (Filipović i Vučo, 2013).

Kako se navodi u tekstu koji opisuje strukturu nastave prevođenja na akademskom nivou: „nastava stranih jezika usmerena je ka pružanju kompetencije da se razlikuju jezički registri, savlada terminologija neophodna stručnim prevodiocima i stekne komunikativna kompetencija za govorne situacije u kojima se stručni prevodioci svakodnevno nalaze” (Daničić, 2018: 144). Podučavanje ULS može biti od posebnog značaja u procesu dostizanja cilja razlikovanja jezičkih registra, s obzirom na činjenicu da se značajno razlikuju u registrima i da označavaju pripadnost određenoj jezičkoj zajednici (Biber & Barbieri, 2007; Hyland, 2008). Cilj nastave prevođenja na akademskom nivou jeste da od studenata naprave „univerzalne prevodioce” (Rajić, 1986, prema Daničić, 2018), jer je nemoguće buduće prevodioce obučiti za sve oblasti koje bi jednog dana mogli da prevode. Uzimajući ovaj cilj u obzir, uviđamo prednosti osveščivanja budućih prevodilaca o karakteristikama ULS i njihovom značaju, a kako bi, ako se dogodi da se u budućnosti bave nekom oblašću koja nije pokrivena u osnovnoj obuci, imali potrebne veštine da obogate svoje frazeološko znanje i kompetencije u različitim registrima i oblastima. Nakon diplomiranja i kada započnu svoju

prevodilačku karijeru, zbog specifičnih karakteristika akademskog diskursa i velikih razlika u disciplinama, prevodioci koji odluče da se bave prevodom naučnih tekstova imaju tendenciju da se specijalizuju za prevode u određenim naučnim oblastima (Montgomeri, 2010). Štaviše, mnoge prevodilačke agencije nude spiskove prevodilaca i oblasti u kojima specijalizuju kako bi bile konkurentnije na tržištu (Montgomeri, 2010). Ovakvoj situaciji je sigurno doprineo čest stav među naučnicima kojima su potrebne prevodilačke usluge da će prevodilac obezbediti adekvatan prevod jedino ako je dovoljno upoznat sa oblašću i registrom (Luo & Hyland, 2019). Stručnjaci u oblasti translologije (Brennett, 2009, prema Luo & Hyland, 2019) smatraju da to nije neophodno već da je dovoljno adekvatno poznavanje akademskog engleskog jezika, uz korišćenje odgovarajućih resursa.

Pisanje naučnih radova na engleskom jeziku za objavljivanje u internacionalnim časopisima je izazovno ne samo za neizvorne govornike engleskog jezika već i za izvorne. Dodatna poteškoća za neizvorne govornike engleskog jezika jeste da mnogi nemaju dovoljno dobro poznavanje engleskog da bi mogli sami da pišu radove na tom jeziku. Samim time su primorani da unajme usluge prevodilaca (Montgomeri, 2010; Murešan i Perez-Ljatanda, 2014; Igrutinović, 2021). Čak i oni neizvorni govornici kojima je znanje engleskog jezika dovoljno dobro da mogu da pišu naučne radove na tom jeziku, često moraju da unajme usluge stranih lektora koji će „doterati rad kako bi bio konkurentan na današnjem akademskom tržištu” (Luo i Hajland, 2019). I dalje se vodi debata o nejednakostima u akademskom objavljivanju među istraživačima koji jesu i nisu izvorni govornici engleskog jezika (Murešan i Perez-Ljatanda, 2014), međutim, Luo i Hajland (2019) predlažu da je produktivnije fokusirati se na rešenja koja oni vide u razvoju engleskog jezika u svrhe istraživanja i objavljivanja (ERPP – *English for Research and Publication Purposes*), kao i u manjem insistiranju na konvencijama koje se oslanjaju na anglofonu tradiciju pri selekciji radova za objavljivanje u naučnim časopisima. Dok se ovi sistemi ne razviju, a svakako i nakon toga, neophodno je pružiti prevodiocima koji rade u raznolikim naučnim oblastima dovoljno podrške i resursa kako bi svoj posao mogli da rade što je efikasnije moguće (Giczela-Pastwa, 2021).

Kako bi se razumeo širi kontekst i potreba za istraživanjem koje je sprovedeno u ovoj doktorskoj disertaciji, neophodno je sagledati i koliko je zastupljeno objavljivanje naučnih časopisa na engleskom jeziku u Srbiji i Rumuniji i koliko se istraživači u ovim državama oslanjaju na usluge prevodilaca. Detaljnije informacije o kontekstu u Srbiji dobijemo iz istraživanja koje je sprovedla Smiljana Igrutinović (2021) sa ciljem da ispita na kojim jezicima istraživači u Srbiji objavljuju naučne radove i sa kojim se poteškoćama susreću kada objavljuju radove na engleskom jeziku. U istraživanju je učestvovalo 30 akademskih radnika sa Akademije strukovnih studija Šumadija (ASSŠ), odsek Trstenik, od kojih se 29 bavi tehničkim i prirodnim naukama a samo jedan/jedna bavi društveno-humanističkim naukama. Svi ispitanici su kao maternji jezik naveli srpski jezik. Rezultati istraživanja su pokazali da ispitanici akademski radnici navode poznavanje četiri strana jezika: engleskog, ruskog, nemačkog i francuskog. Najzastupljeniji strani jezik je engleski, koji govori 90% ispitanika. Polovina ispitanika navelo je da govori ruski jezik. Dalja analiza podataka pokazala je da su to ispitanici u starosnoj grupi iznad 45 godina koji su učili ruski jezik tokom formalnog obrazovanja. Od nastavnika koji su naveli da znaju engleski jezik, 57% je učilo engleski jezik od osnovne škole do fakulteta, dok je preostalih 43% ispitanika engleski jezik učilo samo na fakultetu (jedan ili dva semestra). Skoro polovina ispitanika navelo je da nisu zadovoljni svojim znanjem engleskog jezika.

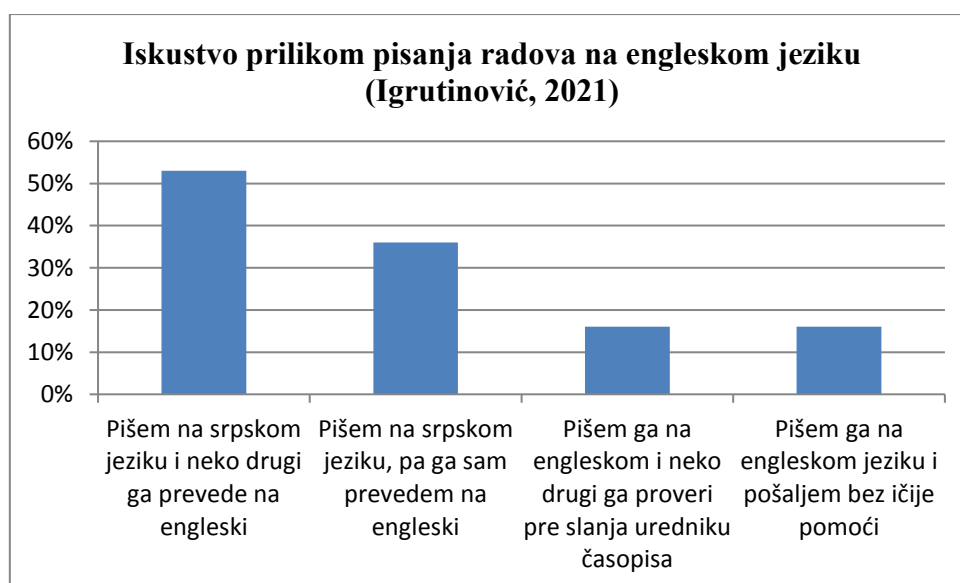
Kada je reč o objavljivanju naučnih radova, 100% ispitanika je objavljivalo naučne radove u proteklih pet godina. Isključivo na engleskom jeziku objavljivalo je 33% ispitanika, dok je 77% ispitanika objavljivalo svoje radove na engleskom jeziku i na maternjem jeziku u nacionalnim časopisima. Naveden podatak pokazuje da se srpski jezik za objavljivanje i dalje neguje u ASSŠ i da i dalje nisu u opasnosti zanemarivanja maternjeg jezika kao što je slučaj u Finskoj, Švedskoj,

Norveškoj i Švajcarskoj (Murešan i Perez-Ljatanda, 2014). Dominacija engleskog jezika kao jezika nauke očigledna je u ispitanoj instituciji s obzirom na podatak da 100% ispitanika objavljuje naučne radove na engleskom jeziku, dok samo 77% objavljuje i na srpskom jeziku. Od drugih jezika za objavljivanje, naveden je samo ruski jezik, koji je za objavljivanje koristilo samo 7% ispitanika. Kada se osvrnemo na obim istraživanja na navedenim jezicima, dolazimo do podataka da dve trećine ispitanika najveći broj radova objavljuje na engleskom jeziku a jedna trećina najveći broj radova objavljuje na srpskom. Na pitanje koji jezik smatraju kao najbitniji za svoju karijeru i akademski rad, 93% je navelo da je engleski jezik najbitniji.

Smiljana Igrutinović je ispitala i koliko su akademski radnici imali uspeha pri objavljivanju u međunarodnim časopisima na engleskom jeziku i koje su pritom potencijalne poteškoće imali. Čak 70% ispitanika uspešno je objavilo sve radove koje su poslali na objavljivanje u međunarodnim časopisima na engleskom jeziku. Od navedenih poteškoća pri objavljivanju, najveći procenat ispitanika je naveo poznavanje akademskog diskursa, praćeno nedovoljnim poznavanjem jezika. Ispitanici su, takođe, naveli i nelingvističke prepreke u vidu nedostatka materijalnih sredstava, pristup najnovijim istraživanjima, kao i tehničke probleme.

Prilikom ispitivanja iskustva akademskih radnika o pisanju radova na engleskom jeziku utvrdilo se da više od polovine (53%) ispitanika piše svoje radove na srpskom jeziku i unajmi usluge prevodilaca (v. Grafikon 2), 36% ispitanika prvo napiše rad na srpskom jeziku pa samostalno prevede na engleski jezik, dok po 16% ispitanika piše rad na engleskom pa ga ili šalje na proveru ili direktno urednicima časopisa kako bi se svrstali u proces objavljivanja. Ispitanici su mogli da biraju više od jedne opcije te ukupan zbir procenata premašuje 100%.

Podatke o kontekstu u Rumuniji dobijamo iz istraživanja koje su sprovele Laura-Mihaela Murešan i Karmen Perez-Ljatanda (Muresan & Pérez-Llantada, 2014). Cilj istraživanja bio je ispitati prakse akademskih radnika u Rumuniji a kada je reč o jeziku objavljivanja, te kako jezičke i obrazovne politike utiču na iste, koje stavove akademski radnici imaju prema poziciji engleskog jezika kao lingve franca u nauci i na koje prepreke nailaze u procesu objavljivanja radova na engleskom jeziku. Ispitanici su akademski radnici na Ekonomskom fakultetu u Bukureštu. U istraživanju je učestvovao ukupno 91 ispitanik.



Grafikon 2. *Iskustvo prilikom pisanja radova na engleskom jeziku*
(Igrutinović, 2021)

Rezultati istraživanja su pokazali da 98% ispitanika govori engleski jezik a 79% govori i francuski. Navedeni su i nemački, italijanski, španski, japanski, portugalski, poljski, mađarski i danski jezik ali sa mnogo manjim procentima zastupljenosti među ispitanicima. U procesu provere upotrebe jezika u naučne svrhe autorke su ispitale upotrebu jezika za čitanje, govor i pisanje. Engleski i rumunski se najčešće koriste u sve tri kategorije iako engleski (čitanje 99%, govor 90%, pisanje 95%) za nijansu više od rumunskog jezika (čitanje 85%, govor 87%, pisanje 84%). Od drugih navedenih jezika, francuski se ističe za čitanje sa 66% ali ne i za pisanje (svega 26% zastupljenosti među ispitanicima). Većina ispitanika (82%) smatra da dominacija engleskog jezika u sferi naučnog objavljivanja daje anglofonim naučnicima prednost uporedno sa neanglofonim naučnicima. Ovakav rezultat je u značajnoj korelaciji sa navedenim nivoom znanja engleskog jezika koji su ispitanici naveli. Interesantno je da 65% ispitanika navodi da oni lično imaju više koristi nego nedostataka zbog upotrebe engleskog jezika kao lingve franka u nauci.

Razmatranjem prepreka u procesu objavljivanja naučnih radova na engleskom jeziku koje su naveli akademski radnici u Rumuniji, uviđamo da su dominantne lingvističke prepreke. Naime, kao glavne probleme naveli su: oskudnost u izražavanju, poznavanje jezičkih suptilnosti i diskursnih karakteristika određenog žanra. Kao što navodi jedan od ispitanika „tekst bi trebalo da očara čitaoca ali to je teško postići za pisce koji nisu izvorni govornici engleskog jezika”¹⁵(Murešan i Perez-Ljatanca, 2014: 59). Ukupno 34% ispitanika koristi usluge prevodilaca i drugih jezičkih posrednika u pripremi radova za objavljivanje na engleskom jeziku. Pored lingvističkih prepreka, ispitanici su naveli da su nailazili na poteškoće u vezi sa postavkom istraživanja, sadržajem i metodologijom istraživanja. Uprkos svemu navedenom, 81% ispitanika veruje da je postojanje lingve franka u nauci neophodno.

Navedena dva istraživanja su nam dala sliku o stanju u Srbiji i Rumuniji u visokoškolskim institucijama što se tiče publikovanja na engleskom jeziku, međutim, oba istraživanja su vezana za po jednu visokoškolsku instituciju i neophodna su dalja istraživanja kako bi se dobila sveobuhvatnija slika situacije. Poređenjem podataka dva istraživanja uviđamo da se veći lingvistički diverzitet pokazao u instituciji u Rumuniji a kada je reč o broju jezika korišćenih za naučne svrhe. U Srbiji, pored engleskog jezika, najuticajni jezik je ruski a u Rumuniji je to francuski jezik; ovakva raspodela produkt je različitih jezičkih politika a verovatno i činjenice da su rumunski i francuski jezik članovi iste jezičke porodice, kao i srpski i ruski. Kako se u Srbiji ispitalo stanje u visokoškolskoj instituciji u oblasti tehničkih i prirodnih nauka, možemo očekivati da bi rezultati bili drugačiji za institucije u oblastima humanistike ili društvenih nauka.

Poređenjem poteškoća sa kojima se suočavaju dve grupe istraživača uviđamo da su one iste i da su lingvističke prepreke primarne. Ovakvi nalazi dodatno podržavaju potrebu za proučavanjem akademskog diskursa i unapređenjem sistema podrške i podučavanja akademskog engleskog jezika. Štaviše, rezultati su pokazali da određeni procenat istraživača i iz obe grupe oslanja na usluge jezičkih posrednika. U ispitanjoj instituciji u Rumuniji 34% istraživača oslanja se na prevodilačke usluge, dok se u Srbiji 53% ispitanika oslanja na prevodilačke usluge a 16% na usluge lektora. Interesantan je podatak da 36% ispitanika u Srbiji prvo napiše svoj rad na srpskom pa naknadno samostalno prevede na engleski jezik. Ovakvi podaci su značajni za kasniju interpretaciju podataka u ovoj disertaciji jer potencijalno utiču na distribuciju ULS, ali zasigurno nam ukazuju na potrebu podučavanja akademskog engleskog i podržavanja naučnika u njihovom akademskom izražavanju na stranom jeziku odnosno engleskom jeziku.

¹⁵[...] „the text should enchant the reader but this is pretty difficult to do for a non-native English speaking writer”

1.3. Translatologija

Značaj prevoda za nauku elokventno je istakao Montgomeri (Montgomery, 2010: 299): „Prevod nauke je star koliko i sama nauka. Zbog njegove uloge u prikupljanju i diseminaciji znanja, prevod je podjednako integralan za naučni napredak kao podučavanje i istraživanje”¹⁶. Danas, u XXI veku značaj prevoda se ne umanjuje, štaviše, zbog globalizacije i sve većeg transfera znanja u digitalnom svetu, prevedeni tekstovi su prisutniji više nego ikada do sada (Montgomeri, 2010). Tome u prilog ide informacija da je u 2019. godini svetsko tržište za prevođenje i tehnologije povezane sa prevođenjem procenjeno da premašuje 46,5 milijarde američkih dolara, dok Evropska unija u procesima prevođenja između 24 jezika godišnje potroši oko milijardu evra na raznolike usluge prevođenja (Munday, Ramos Pinto & Blakesly, 2022).

Kritička razmišljanja o onome što danas nazivamo translatologijom, krajem XVIII veka u anglofonom svetu, pokrenuo je Aleksandar Tytler (*Alexander Tytler*) objavljivanjem knjige *Eseji o načelima prevođenja* (*Essays on Principles of Translation*). On je bio profesor na Univerzitetu u Edinburgu, advokat, sudija i član Gornjeg doma Škotskog parlamenta. Smatrao je da prvo opšte pravilo prevođenja podrazumeva prenošenje ideja iz izvornog teksta (Daničić, 2018: 141). Akademsko istraživanje prevoda započeto je tek sredinom XX veka i inicijalno se prevashodno fokusiralo na ispitivanje taksonomija različitih tipova ekvivalentnosti (Baker & Perez-Gonzalez, 2011). U tom periodu teoretičari su gledali na ekvivalentnost kao na statični pojam, koji nije određen na osnovu zahteva komunikativne situacije nego isključivo na osnovu semantičkog sadržaja izvornog teksta. Jedan od prvih alternativnih pravaca u ovoj sferi istraživanja je stavilo fokus na pitanje da li prevedeni tekstovi imaju isti efekat na čitaoce u ciljnom jeziku, pa je zbog poteškoća u empirijskom testiranju problema preobražen u pitanje da li izvršava ekvivalentnu funkciju (Baker & Perez-Gonzalez, 2011). U drugoj polovini osamdesetih godina prošlog veka značajno se proširio opseg istraživanja u translatologiji; istraživanja su se vršila na nivou reči, rečenice, teksta i kulturološkog konteksta. Štaviše, translatološka istraživanja su počela da se oslanjaju i na nove lingvističke pravce kao što su analiza diskursa, pragmatika, sociolingvistika, psiholingvistika i semiotika (Saldanha, 2009 prema Baker & Perez-Gonzalez, 2011). Naredna velika ekspanzija u translatološkim istraživanjima desila se popularizacijom korpusne lingvistike devedesetih godina prošlog veka (Bernardini, 2015).

Primena korpusnolingvističke metodologije na translatološka istraživanja popularizovala se nakon objavljivanja istraživanja autorke Bejker (Baker, 1993, prema Bernardini, 2015), u kojem se ističe značaj ispitivanja prevodilačkih univerzalija (engl. *translation universals*), odnosno pojava koje su karakteristične za prevedene tekstove uporedno sa originalnim neprevedenim tekstovima a koje nisu posledica uticaja specifičnog lingvističkog sistema. Kao posledica, mnogi istraživači u oblasti translatologije, pored korpusa prevedenih tekstova, formirali su poredive korpusne neprevedenih tekstova i stvorila se alternativa dotadašnjeg poređenja izvornih tekstova i prevoda. Koncept prevodilačkih univerzalija dobio je različite kritike njegove teorijske zasnovanosti, međutim, kao što Česterman (Chesterman, 2004: 11, prema Bernardini, 2015) ističe „ono što je na kraju bitno možda i nisu univerzalije, koje svakako ni ne možemo konačno da potvrdimo, već novo znanje obrazaca, i obrazaca obrazaca, koji nam pomažu da razumemo ono u šta gledamo”¹⁷.

Translatološka istraživanja koja primenjuju korpusnu metodu dele se u tri struje na osnovu postavljenih ciljeva:

¹⁶ „Translation of science is as old as science itself. Due to its role both in collecting and disseminating knowledge, translation has been no less integral to scientific progress than teaching and research.”

¹⁷ „[w]hat ultimately matters is perhaps not the universals, which we can never finally confirm anyway, but new knowledge of the patterns, and patterns of patterns, which helps us to make sense of what we are looking at.”

1. Istraživanja koja se fokusiraju na identifikaciju *prevodilačkih univerzalija* kako bi se bolje razumeo prevodilački proces/produkt i njegove karakteristike, kao i unapredio kvalitet prevoda ispitivanjem prevodilačkih normi uporedno sa onima koje su primenjene u neprevedenim tekstovima. Uglavnom su to jednojezička istraživanja koja ispituju korpus prevedenih tekstova na određenom jeziku i poredivog korpusa na istom jeziku.
2. Istraživanja koja se fokusiraju na *prevodilačke odmake* (engl. *translation shifts*) u kojima se koriste paralelni korpusi tekstova na izvornom i ciljnom jeziku.
3. Istraživanja koja se fokusiraju na *identifikaciju i analizu stilskih „otisaka prstiju“ koje ostavljaju prevodioci*.

Nove metodologije u navedenim strujama istraživanja ističu značaj kombinovanja podataka iz paralelnih korpusa izvornih i ciljnih tekstova, kao i upotrebu referentnih korpusa (Li, 2013; Liu i Afzal, 2021).

1.3.1. Translatologija i učestali leksički spojevi (ULS)

Istraživanja učestalih leksičkih spojeva u oblasti translatologije su poprilično retka (Berūkštienė, 2017; Novita & Kwary, 2018). Međutim, primeri izučavanja ULS se mogu naći u sve tri struje istraživanja koje su proistekle iz primene korpusne metode na translatološka istraživanja (*prevodilačke univerzalije*, *prevodilački odmaci* i *stilski tragovi prevodilaca*). Od tri struje najučestalija su istraživanja koja ispituju prevodne univerzalije, koje se definišu kao „inherentne karakteristike prevedenog teksta koje se nalaze u različitim jezicima nezavisno od izvornog teksta”¹⁸(Lee, 2013: 378). Osvrnućemo se na nekoliko istraživanja koja se bave tematikom iz sve tri struje istraživanja kako bismo predstavili predloženu metodologiju i relevantne rezultate koji bi mogli biti značajni pri interpretaciji rezultata ove disertacije.

Kada je reč o tipovima tekstova koji se obrađuju u istraživanjima, pravo se ističe kao jedna od oblasti koja je privukla mnogo pažnje a kada je reč o istraživanjima koja spajaju translatologiju i učestale leksičke spojeve. Navedeno stanje u istraživanjima se može objasniti činjenicom da je pravni jezik visoko formulaičan te da se frazeologija intenzivno istraživala u datoj oblasti (Berūkštienė, 2017). Posle prava, najučestalija su istraživanja književnih dela zbog njihove dostupnosti i praktičnosti u formiranju paralelnih korpusa, kao i novinski članci. Nažalost, autorka disertacije nije pronašla nijedno istraživanje koje spaja translatologiju i ULS, a da je sprovedeno na naučnim radovima. Pretpostavlja se da ovih radova nema ili su veoma retki zbog konvencija koje postoje u svetu akademskog publikovanja u kojem se ne navodi da li su akademski radovi prevedeni.

ULS su jedan od značajnih elemenata frazeološke strukture teksta. Od prevoda se očekuje da se konformišu sa frazeologijom ciljnog jezika koja je tipična za žanr teksta (Gouadec, 2007, prema Biel, 2017). Međutim, prevodilački proces ostavlja svoje tragove na frazeologiji teksta koji je preveden. U procesu prevođenja može doći do inhibicije „prajminga” (engl. *priming*) prirodnih šablona. Štaviše, oni mogu da budu iskrivljeni zbog jakog uticaja izvornog jezika (Tourey, 2012). Prema kolokacijskoj hipotezi (Mauranen, 2006, prema Biel, 2017) pretpostavlja se da prevodi sadrže kolokacije koje su moguće ali retke u ciljnom jeziku, imaju manje kombinacija koje su česte u neprevedenim testovima ciljnog jezika i imaju raznovrsnije, manje stabilne i manje frekventne leksičke spojeve. S druge strane, Turijev zakon rastuće standardizacije (*Tourey’s law of growing standardization*) ilustruje tendenciju prevodilaca da koriste uobičajne konstrukcije koje nudi ciljna kultura (Tourey, 2012), a hipoteza normalizacije/konzervatizma (Baker, 1996, prema Biel, 2017) ističe tendenciju prevodilaca da prenaglašavaju tipične karakteristike ciljnog jezika, što je povezano

¹⁸[...] „inherent properties of translated texts found across languages independently of source texts”

sa hipotezom pojednostavljanja (engl. *simplification hypothesis*) po kojoj su prevedeni tekstovi jednostavniji od poredivih neprevedenih tekstova u istom jeziku (Laviosa, 1998, prema Bernardini, 2015). Štaviše, postoji nedostatak bilingvalnih terminoloških resursa koji pored termina uključuju i jedinice od nekoliko reči a koje igraju značajnu ulogu u procesu prevođenja (Berükštienė, 2017). Sve navedeno dalje ukazuje na značaj sprovođenja translatoških istraživanja sa fokusom na frazeologiju a specifičnije – ULS, koji nisu dovoljno istraženi u ovoj oblasti.

Neophodno je ovde uzeti u obzir činjenicu da istraživanja koja ispituju ULS u okviru frazeologije akademskog engleskog jezika i istraživanja koja ispituju ULS u okviru istraživanja prevodilačkih univerzalija, pored različitih ciljeva, imaju i različite perspektive, kao i različite metodološke postavke. U istraživanjima frazeologije akademskog engleskog, kao što je i istraživanje sprovedeno u ovoj doktorskoj disertaciji, glavni cilj je bolje razumevanje akademskog diskursa kroz utvrđivanje karakteristika ULS kao lingvističkog oblika ili pak bolje razumevanje karakteristika ULS kroz ispitivanje njihove upotrebe unutar specifične oblasti odnosno akademskog engleskog jezika. U translatoškim istraživanjima, ULS se pristupa kao sredstvu za utvrđivanje postojanja i karakteristika prevodilačkih univerzalija. Kao što je i Li (2013) istakao: „Korisnost analize ULS u istraživanjima prevoda prevashodno se ogleda u tome što omogućuje istraživaču da ih koristi kao alat za otkrivanje i izdvajanje individualnih leksičkih šema koje otkrivaju inherentne karakteristike prevedenog teksta”¹⁹. Iz navedenih razloga, čini se da istraživači iz dva pravca, iako ispituju isti lingvistički fenomen, nisu dovoljno upoznati ili se ne osvrću dovoljno na međusobne rezultate jer primenjuju različite standarde i metodološke postavke pri analizi ULS. I pored razlika u ciljevima istraživanja i metodološkim postavkama, nesumnjivo je da istraživanja ULS u oblasti translatošije doprinose boljem razumevanju karakteristika ULS i kako se oni koriste u različitim kontekstima. Sledi kratak pregled relevantne literature kako bi se dobio uvid u istraživački teren u kojem se spajaju translatošija i izučavanje ULS, ali i kako da bi se skrenula pažnja na zaključke koji će doprineti boljem razumevanju šireg konteksta u kojem je sprovedeno istraživanje u ovoj disertaciji i omogućila podrobnija analiza dobijenih podataka.

Biel (2017)

Istraživanje koje je sprovedla Biel (2017) imalo je cilj da istraži karakteristike upotrebe ULS u prevodima pravnih tekstova. Da bi postigla navedeni cilj, uporedila je upotrebu ULS u tri korpusa:

1. *PL-Eurolect (The Polish Eurolect Corpus)*, koji se sastoji od 925 regulacija Evropske unije i 92 direktive EU na poljskom jeziku. U istraživanju se ovaj korpus tretira kao korpus prevedenog teksta sa engleskog, iako je ova interpretacija zakomplikovana jezičkim principima EU. Veličine je 2.667.590 reči.
2. *EN-Eurolect (The English Eurolect Corpus)*, koji se sastoji od 925 regulacija Evropske unije i 92 direktive EU na engleskom jeziku. Služi kao referentni korpus. Veličine je 3.068.479 reči.
3. *PL-Domestic (The Polish Domestic Law Corpus)*, koji se sastoji od 135 standardnih statuta i 55 centralnih statuta. Veličine je 3.454.942 reči.

Kao što je i sama autorka navela, uporedivost korpusa je umanjena činjenicom da je *PL-Domestic* korpus sačinjen od značajno manjeg broja tekstova i da je 30% veći od *PL-Eurolect* korpusa. Za identifikaciju ULS postavljena je minimalna frekvencija od 40 ponavljanja u milion reči i minimalna disperzija u pet tekstova. U istraživanju se primenom istih parametara izdvajanju ULS

¹⁹ „The usefulness of lexical bundle analysis for translation research lies primarily in enabling the researcher to use it as a discovery tool for mining individual lexical patterns likely to be reflective of inherent properties of translated texts.”

dužine od tri, četiri, pet, šest, sedam i osam reči, sa fokusom na ULS od četiri reči pri interpretaciji podataka.

Autorica postavlja značajno praktično pitanje za međujezičko poređenje ULS koje se sprovodi u translatoškim istraživanjima: „do koje mere leksički spojevi od četiri reči predstavljaju isti nivo formulaičnosti u različitim jezicima, drugim rečima, kako možemo da poredimo leksičke spojeve od četiri reči u različitim jezicima”²⁰ (Biel, 2017: 17). Ovo pitanje je od posebnog značaja kada se sa engleskim jezikom poredi jezici koji su flektivni i u kojima se isto značenje može preneti sa manjim brojem reči.

Rezultati istraživanja su pokazali da ULS od tri reči sačinjavaju 40% svih izdvojenih ULS u *PL-Eurolect* korpusu, dok u *PL-Domestic* korpusu ULS od tri reči sačinjavaju 54% svih izdvojenih ULS. Ovaj podatak nam ukazuje na to da se u neprevedenom tekstu na poljskom jeziku autori više oslanjaju na kraće formulaične izraze nego u prevedenom tekstu. Veće oslanjanje na duže konstrukcije u *PL-Eurolect* korpusu može da se objasni, sa jedne strane, kao uticaj prevodilačkog procesa i na kalkiranje dužih konstrukcija na engleskom ali, sa druge strane, može da bude i posledica činjenice da je *PL-Eurolect* korpus sačinjen od pet puta više tekstova nego *PL-Domestic* korpus, što može da dovede do češćeg ponavljanja određenih izraza odnosno do netačno identifikovanog većeg nivoa formulaičnosti u korpusu. Dalje, veći broj tipova ULS je identifikovan u *EL-Eurolect* korpusu (3.734) nego u *PL-Eurolect* korpusu (1.986). Međutim, nemoguće je odrediti do koje je mere ova razlika prisutna zbog procesa prevođenja a do koje mere zbog morfo-gramatičkih razlika u dva jezika. Činjenica da je veći broj tipova ULS identifikovan u *PL-Eurolect* korpusu nego u *PL-Domestic* korpusu (934) je lakša za interpretaciju i ukazuje na viši nivo formulaičnosti u prevedenim tekstovima, najverovatnije zbog uticaja prevodilačkog procesa.

Autorica je uporedila deset najfrekventnijih ULS od četiri reči u sva tri korpusa. Prva tri ULS u *PL-Eurolect* i *EL-Eurolect* korpusu su isti dok se ostali razlikuju. Štaviše, zbog strukturalnih razlika između engleskog i poljskog jezika postoji mala sličnost u identifikovanim ULS u dva korpusa jer većina ULS od četiri reči u *PL-Eurolect* korpusu poklapa sa kraćim ULS u korpusu na engleskom jeziku. Poslednji korak u analizi je identifikacija preklapanja ULS od četiri reči u *PL-Eurolect* i *PL-Domestic* korpusima. Rezultati su pokazali da se samo 11 ULS poklapa u oba korpusa, što sačinjava samo 2,8% svih ULS od četiri reči u *PL-Eurolect* korpusu. Autorica na osnovu ovih rezultata zaključuje da prevodi zbog uticaja izvornog teksta koriste ULS koji nisu tipični za neprevedene tekstove iz istog žanra. Takođe, ona navodi da ovakvi rezultati podržavaju Turijev zakon rastuće standardizacije (Touy, 2012) i hipotezu normalizacije (Baker, 1996, prema Biel, 2017) zbog velikog broja identifikovanih ULS čija se identifikacija oslanja na frekvenciju upotrebe; širi spektar različitih tipova i dužina identifikovanih ULS podržavaju kolokacijsku hipotezu (Mauranen, 2006, prema Biel, 2017), koja navodi da prevodima koristi raznolikiji vokabular i manje stabilne šeme.

Brukštín (2017)

Istraživanje Brukštín (Berūkštienė, 2017) analiziralo je prevodne ekvivalente ULS od četiri reči u sudskim presudama sa engleskog na litvanski jezik sa ciljem da se ustanovi šta se dešava sa strukturom ovih konstrukcija u prevodu i da se utvrde prevodne šeme (engl. *translational patterns*) različitih strukturalnih tipova ULS. U istraživanju je upotrebljen paralelni korpus sudskih presuda koji se sastoji od 104 presude na engleskom jeziku (one sačinjavaju potkorpus od milion reči) i prevod svih 104 presuda na litvanski jezik, što sačinjava potkorpus od 730.000 reči. Kako bi se izdvojili ULS iz potkorpusa na engleskom jeziku, koristila se minimalna frekvencija od 50

²⁰[...] „to what extent a 4-gram reflects the sam level of formulaicity across languages , in other words how we can compare 4-grams across languages”

ponavljanja u milion reči i minimalna disperzija u deset tekstova. Nakon što je finalizovan prečišćen spisak od 245 ULS, urađena je strukturalna raspodela identifikovanih ULS. Korišćenjem *AntPConc 1.2.0*. (Anthony, 2007, prema Berūkštienė, 2017) identifikovani su prevodni ekvivalenti za deset najfrekventnijih ULS u svakoj od strukturalnih podgrupa na osnovu taksonomije koju su predložili Bajber i saradnici (1999, 2004).

Rezultati analize su pokazali da prevodni ekvivalenti strukturalnih grupa poprimaju sledeće strukture:

- *predloške fraze*: veznici, predlozi, particip, predlozi, imeničke fraze;
- *imeničke fraze*: imeničke fraze, imenice, particip, imenica + fragment zavisne klauze;
- *glagolske fraze*: fraze u infinitivu, imenica + glagol, imenica + particip, glagol, glagolske fraze;
- *klauzalni ULS*: fragmenti *that*-klauze, infinitiv.

Rezultati su dalje pokazali kako u strukturalnim podgrupama *in + imenička fraza* i *druge predloške fraze*, kada ULS predstavlja apstraktne elemente, prevodni ekvivalenti ne sadrže imenice, no ako ULS predstavlja koncept iz pravnog domena – sadrži imenice. Među strukturalnim podgrupama najlakša za prevođenje – jer u litvanskom jeziku ima najsličniji oblik – jeste *imenička fraza sa of fragmentom*. Autorka navodi da rezultati njenog istraživanja ukazuju na to da izučavanje ULS svoje mesto imaju kako u oblasti usvajanja i podučavanja engleskog jezika tako i u translatologiji odnosno prevodu.

U navedenom istraživanju vidimo koliko paralelni korpusi mogu biti značajni za izučavanje ULS i za njihovu praktičnu primenu. Bilo bi veoma korisno kada bi se takvo istraživanje sprovedo „u suprotnom smeru”, odnosno kada bi se identifikovali ULS u prevodima litvanskih tekstova na engleski jezik. Kada bismo u tom spisku identifikovali ULS koji nisu tipični za izvorne govornike engleskog i u paralelnom korpusu identifikovali prevodne ekvivalente, mogli bismo da utvrdimo koje konstrukcije prave probleme u prevodu i pri podučavanju prevoda skrenemo pažnju na moguće greške i rešenja. Ova vrsta informacije bi bila od posebnog značaja za prevodioce jer ULS označavaju pripadnost određenom registru i žanru, a govornici ih nisu inherentno svesni bez adekvatnog podučavanja (Biber et al., 1999).

Novita i Kvari (2018)

Poređenje upotrebe ULS u prevodima koje su uradili iskusni profesionalni prevodioci i studenti koji su završili kurs prevođenja, bio je cilj istraživanja Novita i Kvari (Novita & Kwary, 2018). Još jedan cilj istraživanja bio je da se upotrebe nalazi ovog istraživanja za identifikaciju oblasti u kojima bi budući prevodioci trebalo da se usavrše kako bi se unapredile njihove tehnike podučavanje. Na početku opisa ovog istraživanja i pre diskusije o potencijalnim zaključcima koji bi mogli da se donesu na osnovu njegovih rezultata, neophodno je napomenuti da je korpus korišćen u istraživanju veoma mali i da je upitno da li zadovoljava kriterijume za pouzdanu identifikaciju ULS koji se zasnivaju na frekvenciji upotrebe. Zbog činjenice da autorka disertacije nije našla nijedan drugi rad koji se bavio istom temom sa većim korpusima i zbog činjenice da bi navedeni istraživački smer doprineo boljem razumevanju interakcije ULS i translatologije, ipak ćemo sagledati rezultate ovog istraživanja u nadi da će pružiti impulse za dalja detaljnija istraživanja i da će nas uputiti u pravom smeru a kada je reč o potencijalnim rezultatima takvih istraživanja.

Kako bi se prikupio materijal za ovo istraživanje, isti književni tekst od oko 600 reči je preveden sa indonežanskog na engleski jezik od strane ukupno pet studenata koji su završili kurs prevođenja sa visokim ocenama i koji su imali visok nivo na TOFEL testu, kao i ukupno pet sertifikovanih

prevodilaca. U istraživanju su se sagledavale upotrebe ULS od tri, četiri i pet reči koje su identifikovane upotrebom minimalne frekvencije od tri ponavljanja i minimalne disperzije u tri teksta. Rezultati istraživanja su pokazali da postoje značajne razlike u upotrebi ULS u dve navedene grupe prevodilaca. Profesionalni prevodioci su koristili veći broj tipova i tokena za sve dužine ULS. Rezultati istraživanja su u skladu sa nalazima istraživanja Čen i Bejker (2010), koje je pokazalo da studenti sa nižim nivoom znanja koriste manje ULS. Ovakva istraživanja su dragocena zato što pružaju smernice za unapređenje podučavanja engleskog jezika i veština prevođenja. Tako iz navedenog istraživanja možemo da uočimo potrebu da se kod studenata podstiče češća upotreba ULS u procesu prevođenja.

Liu i Afzal (2021)

Istraživanje Liu i Afzal (Liu & Afzaal, 2021) sprovedeno je sa ciljem da se upoređi upotreba ULS u dve verzije prevoda kineskog književnog dela *Hongloumeng* na engleski jezik kako bi se ispitale stilske razlike kod prevodilaca. Jedan od prevoda uradio je izvorni govornik engleskog jezika Dejvid Hoks (*David Hawkes*), dok su drugi prevod uradili bračni par Jang (*Yang*), koji su izvorni govornici kineskog jezika. Istraživanje nije sagledalo cele prevode već samo prvih 15 poglavlja knjige. Stoga je korpus poglavlja na kineskom veličine 91.173 kandija, korpus Hoksovog prevoda veličine je 89.396 reči, dok je korpus prevoda koji su uradili bračni par Jang veličine 67.649 reči i 761 reči u fusnotama. Izdvojeni su ULS od tri, četiri, pet i šest reči u oba korpusa na engleskom jeziku, međutim, nije navedeno koja je minimalna frekvencija i disperzija korišćena. U istraživanju su se, takođe, detaljnije sagledali 20 najfrekventnijih ULS od četiri reči i raspoređeni u funkcionalne grupe i podgrupe na osnovu taksonomije koju su predložili Bajber i sar. (2004).

Rezultati istraživanja pokazali su da su ULS od tri, četiri i pet reči češće upotrebljeni u Hoksovom prevodu nego u prevodu bračnog para Jang. Jedino je malo veći broj tipova i tokena ULS od šest reči identifikovan u prevodu Jangovih. Kada je reč o funkcionalnoj raspodeli 20 najfrekventnijih ULS od četiri reči, ustanovljeno je da su u prevodu bračnog para Jang sagledani ULS prevashodno u funkcionalnoj kategoriji referentnih ULS, dok se u prevodu koji je uradio Hoks vidi malo veća varijacija funkcionalnih grupa. Autori smatraju da je analiza ULS pokazala jasnu razliku u prevodilačkom stilu u dva istražena predmeta i ističu metodološki značaj ispitivanja ULS u istraživanjima koja se bave stilskim razlikama kod različitih prevodilaca.

Bejker (2004)

Mona Bejker (Baker, 2004) u svom istraživanju analizira ULS u prevedenim i neprevedenim tekstovima kako bi utvrdila da li postoje razlike u upotrebi koje su specifične za prevedene tekstove. Istraživači u oblasti translatologije tvrde da su prevedeni tekstovi na više načina drugačiji od neprevedenih tekstova, međutim, načini na koje se razlikuju često nisu jasno definisani, ali mahom uključuju: tvrdnju da prevodioci konzervativnije koriste jezik; da preferiraju standardizovane jezičke oblike; da koriste formalnije jezičke izraze; preveden tekst je „prečišćen” u smislu da se izbegavaju regionalizmi (Bejker, 2004). Kada je reč o prevodima, u angloameričkom svetu postoji tvrdnja da je fluentnost teksta dominantna strategija i vrednost u prevedenom tekstu (Venuti, 1995, prema Bejker, 2004). Ako je navedena tvrdnja istinita, može se očekivati da će se veći broj tipova i tokena naći u prevedenom tekstu.

Kako bi proverila gorenavedenu tvrdnju, autorka je koristila dva korpusa:

1. *TEC (Corpus of Translated English)* – korpus se sastoji od 94 prevedenih beletrističkih tekstova, biografija, avionskih magazina i novinskih članaka na engleski jezik; korpus je veličine 6.613.456 reči;

2. korpus neprevedenih tekstova na engleskom jeziku koji je sačinjen od dela tekstova iz BNC (*British National Corpus*) korpusa, kako bi se postigla što veća komparativnosti korpusa; sastoji se od 171 tekstova i veličine je 6.423.325 reči.

Izdvojeni su ULS od tri, četiri i pet reči, ali nije navedeno koji su kriterijumi za izdvajanje ULS korišćeni. Izdvojila je ULS za koje smatra da su relevantni za ciljeve istraživanja i zaključila da se veći broj tokena ULS zaista nalazi u korpusu prevedenih tekstova. Nažalost, zbog oskudnog opsega analize dobijenih rezultata, kao i što sama autorka navodi, dobijeni rezultati ne bi mogli da se tretiraju kao sistematični dokazi, jer je cilj istraživanja bio ilustriranje tipa pitanja na koja korpusna istraživanja translatologije mogu da daju odgovore. Glavni značaj ovog istraživanja jeste što je poslužilo kao inspiracija drugim istraživačima da nadograđuju metodologiju ovog tipa istraživanja i da produbljuju znanja u ovoj oblasti. Takođe je jedan od malobrojnih istraživanja koje se bavi konkretno ULS u prevedenim tekstovima na engleski jezik; kao što ćemo videti u daljem tekstu, češće su se ispitivale karakteristike prevoda sa engleskog jezika na druge jezike.

Li (2013)

Li (Lee, 2013) kao glavnu inspiraciju za svoje istraživanje navodi gorepomenuti naučni rad koji je napisala Mona Bejker (2004). Li (2013) istraživanje sprovodi sa ciljem da predstavi svoj predlog metodoloških postavki za analizu ULS kao sredstvo za identifikaciju i ispitivanje prevodnih univerzalija, kao i da demonstrira upotrebu metodološke postavke ispitujući upotrebu ULS od tri leksičke jedinice²¹ kako bi se identifikovalo postojanje prevodilačkih univerzalija u novinskim člancima.

Dva korpusa su korišćena u istraživanju:

1. *TRCN (Translated News Corpus)* – paralelan korpus novinskih tekstova na engleskom i njihov prevod na korejski jezik, potkorpus na engleskom jeziku veličine je 465.085 reči, dok je potkorpus na korejskom jeziku veličine 325.734 leksičkih jedinica;
2. *KNC (Korean News Corpus)* – koristi se kao referentni korpus, i veličine je 2.113.497 leksičkih jedinica.

Li predlaže metodološku postavku za ispitivanje ULS u svrhe identifikacije prevodilačkih univerzalija u četiri koraka:

1. *Korišćenje kompjuterskih programa za izdvajanje ULS.* Kako bi u ovom istraživanju izvršio ovaj korak, koristio je minimalnu frekvenciju od 20 ponavljanja u milion reči ali nije postavio minimalnu disperziju (što je van konvencija istraživanja ULS).
2. *Prečišćavanje spisaka ULS od onih spojeva koji nisu relevantni za translatološka istraživanja.* U slučaju ovog istraživanja, a po uzoru na prethodna (Bejker, 2004), uklonjeni su ULS koji se odnose na prostorne i vremenske fraze, kao i sve referentne ULS.
3. *Izdvajanje ULS koji verovatno ukazuju na važne razlike između prevedenih i neprevedenih tekstova* primenom dva kriterijuma: a) ako postoje velike razlike u frekvencijama izdvojenim ULS u prevedenim i neprevedenim tekstovima; b) ako postoje ULS koji su specifični samo za jedan od dva potkorpusa tj. ako se pojavljuju u samo jednom od potkorpusa. Za ULS koji ispunjavaju drugi kriterijum neophodno je proveriti da razlog neujednačenosti nije uticaj maternjeg jezika. Li je u istraživanju

²¹U korejskom jeziku leksičke jedinice koje se odvajaju poredom nisu jednake sa konceptom reči u srpskom ili engleskom jeziku jer obično podrazumevaju dva elementa koji bi u srpskom ili engleskom bili dve odvojene reči.

primenio treći korak samo na 20 najfrekventnijih ULS u oba korpusa. Samo jedan ULS zadovoljio je prvi kriterijum *hal swu issta* zato što je najčešće korišćen ULS u referentnom korpusu izvornih tekstova na korejskom (124 tokena) a u korpusu prevedenih tekstova tek je na 12. poziciji (45 tokena).

4. Poslednji korak predstavlja *detaljnu analizu izdvojenih ULS ispitivanjem konteksta u kojem se pojavljuju, obrazloženjem i interpretacijom rezultata uz formiranje relevantnih veza sa prethodnim istraživanjima iz oblasti*. Primena ovog koraka izdvojen ULS *hal swu issta*, Li je ustanovio da se navedeni ULS zapravo koristi u dva različita konteksta sa dva različita značenja. Naime, koristi se za ublažavanje poruke odnosno prenošenje epistemičkog značenja ili za izražavanje sposobnosti odnosno prenošenje deontičkog značenja. Kada su se ove dve upotrebe sagledale u oba korpusa, ustanovljeno je da se u primeni epistemološkog značenja dva korpusa značajno razlikuju, odnosno da se mnogo češće koristi u korpusu neprevedenih tekstova KNC. Poređenjem ovih rezultata sa korpusom originalnih tekstova na engleskom ustanovljeno je da u originalu nisu korišćena sredstva za ublažavanje poruke odnosno sredstva za iskazivanje niske epistemičke modalnosti, što dalje znači da su prevodioci upotrebili navedene izraze kako bi uskladili tekstove sa konvencijama izražavanja u ciljnom jeziku. Zaključak na osnovu navedenih podataka je da većina prevodilaca nije uzela u obzir konvencije ciljnog jezika te da nisu dovoljno često primenili izdvojen ULS u epistemičkom značenju. Neophodno je više istraživanja da bi se utvrdilo da li je naveden nalaz prevodilačkih univerzalija ili je samo specifična za navedeni kontekst, odnosno lokalizovana norma nastala zbog društvenih i kulturnih vrednosti (Toury, 1995, prema Li, 2013).

Siao (2010)

Istraživanje Siao (Xiao, 2010) bavilo se istom tematikom kao i Li (2013) u drugom kontekstu i došlo je do zaključaka koji dalje potvrđuju hipotezu do koje je došla Bejker (2004) – da je učestalija upotreba ULS prevodilačka univerzalija. Siao (2010) pored ULS ispituje i idiome sa ciljem identifikacije prevodilačkih univerzalija u kineskom jeziku. Korpusi upotrebljeni u istraživanju napravljeni su da sadrže različite žanrove kako bi bili što reprezentativniji kineskom jeziku u celini. Ispitana je upotreba ULS u originalnim tekstovima na kineskom jeziku, kao i u korpusu prevedenih tekstova na kineski jezik. Rezultati su pokazali da su ULS češće korišćeni u korpusu prevedenih tekstova.

Ebeling (2022)

U svom istraživanju Ebeling (2022) ispituje upotrebu ULS u prevedenim tekstovima na engleski jezik sa norveškog, švedskog i nemačkog kako bi utvrdio da li su identifikovane karakteristike upotrebe ULS u prevedenim tekstovima posledica postojanja prevodilačkih univerzalija ili uticaja izvornog teksta. Istraživanje se nadovezuje na istraživanje (Ebeling & Ebeling, 2017, prema Ebeling, 2022), u kojem je upoređena upotreba ULS u korpusu prevedenih tekstova sa norveškog na engleski jezik i u tekstovima na engleskom koji nisu prevodi. Autor se fokusirao na ispitivanje ULS od tri reči u beletrističkim tekstovima.

U istraživanju su korišćena tri korpusa:

1. *ENPC+ (English-Norwegian Parallel Corpus English +)* sadrži originalne tekstove i na engleskom i na norveškom i njihove prevode na drugi jezik. Potkorpus prevoda sa norveškog na engleski sadrži ukupno 39 tekstova i veličine je 1.418.321 reči. Potkorpus originalnih tekstova na engleskom sadrži 39 tekstova i veličine je 1.316.006 reči.

2. *ESPC (English Swedish Parallel Corpus)* sadrži originalne tekstove i na engleskom i na švedskom, kao i njihove prevode na drugi jezik. Potkorpus prevoda sa švedskog na engleski sadrži ukupno 25 tekstova i veličine je 337.637 reči.
3. *OMC (Oslo Multilingual Corpus)* sastoji se od originalnih tekstova i prevoda holandskog, engleskog, francuskog, nemačkog, norveškog i portugalskog jezika. Ebeling(2022) u istraživanju koristi samo potkorpus prevoda nemačkih tekstova na engleski jezik koji se sastoji od 10 tekstova i veličine je 154.025 reči.

Sâm postupak analize podataka je neuobičajen. U prvom koraku autor izdvaja ULS od tri reči primenom minimalne frekvencije od 20 ponavljanja u milion reči, kao i veoma strog kriterijum disperzije u čak 25% tekstova. U korpusu izvornih tekstova na engleskom izdvojeno je 1.408 tipova ULS, u prevodima sa nemačkog 1.371 tipova ULS, u prevodima sa norveškog 1.468 tipova ULS i u prevodima sa švedskog 1.149 tipova ULS. Sledeći koraci analize su neobični jer autor poredi ULS u originalnim tekstovima i prevedenim te za one ULS koji su zadovoljili identifikacione kriterijume samo u originalnim tekstovima proverava da li se pojavljuju u prevedenim tekstovima i ako se pojavljuju, uključuje ih u analizu sa disperzijom i frekvencijom sa kojom se pojavljuju iako je ona ispod gorenavedenih parametara. Isto je uradio i za ULS koji su se identifikovali u prevedenim tekstovima ali nisu u originalnim. Tako da se na kraju dobio spisak od 2.765 tipova ULS koji su prisutni u sva četiri ispitana korpusa i potkorpusa. Izdvojeni ULS su dalje raspoređeni u funkcionalne grupe, a njihovi tokeni sagledani u primeni statističkih testova kako bi se ustanovilo da li postoji statistički značajna razlika u upotrebi različitih funkcionalnih grupa ULS u ispitanim korpusima. Istraživač je za funkcionalnu taksonomiju koristio kombinaciju taksonomija koje su predstavili Altenberg (1998), Bajber i sar. (2004) i Mun (Moon, 1998, prema Ebeling, 2022).

Rezultati istraživanja su pokazali su da se upotreba ULS u izvornim i prevedenim tekstovima značajno razlikovala u šest funkcionalnih grupa i podgrupa (od ukupno 15 analiziranih), a da se značajna razlika u upotrebi u sve tri grupe prevedenih tekstova uporedno sa originalnim pojavila u četiri funkcionalne grupe (fragment, organizacija, proces i vreme). U svim navedenim grupama značajno se veći broj tipova koristio u prevedenim tekstovima uporedno sa izvornim.

Kritička analiza i poređenje opisanih relevantnih istraživanja

Navedena istraživanja predstavljena su kako bismo sagledali sa kojim su sve ciljevima analizirani ULS u translatoškim istraživanjima. Obradena istraživanja obuhvatila su sledeće ciljeve:

- istražiti karakteristike upotrebe ULS u prevedenim tekstovima uporedno sa originalnim tekstovima kako na izvornim tako i na ciljnom jeziku;
- istražiti poredivost ULS od četiri reči u različitim jezicima;
- analizirati prevodne ekvivalenatne ULS i identifikovati prevodne šeme za iste;
- unaprediti tehnike za podučavanje prevodilaca;
- istražiti razlike u upotrebi ULS kod iskusnih prevodilaca i studenata;
- analizirati prevodilački stil kod prevodilaca sa različitim maternjim jezikom, odnosno onih koji prevode na strani jezik i onih koji prevode na svoj, maternji jezik;
- identifikovati prevodilačke univerzalija.

Tri (Bejker, 2004; Li, 2013; Ebeling, 2022) navedena istraživanja analiziraju upotrebu ULS u prevedenim tekstovima sa ciljem identifikacije prevodilačkih univerzalija. Bejker (2004) daje predlog kako bi ULS mogli da se analiziraju u svrhu identifikacije prevodilačkih univerzalija, dok se Li (2013) nadovezuje i dodatno precizira predloženu metodologiju. Ono što je specifično za predloženu metodologiju jeste činjenica da se fokusira na odabir specifičnih ULS koji se smatraju adekvatnim za utvrđivanje karakteristika upotrebe ULS u prevedenim tekstovima, umesto

sagledavanja širih strukturalnih i funkcionalnih tendencija koje se uviđaju u svim identifikovanim ULS, kao što je to obično slučaj u istraživanjima ULS²². Treće istraživanje koje je sproveo Ebeling (2022) fokusira se na tipičniju širu sliku tendencija raspodele svih identifikovanim ULS, međutim, primenjuje nesvakidašnju metodologiju jer ujednačuje broj ULS u svim istraženim korpusima tako što u analizu dodaje i one ULS koji nisu zadovoljili postavljene uslove identifikacije (minimalnu frekvenciju od 20 ponavljanja u milion reči i minimalnu disperziju u 25% tekstova). Specifičnosti metodologije ovog istraživanja se nastavljaju sa odabirom funkcionalne taksonomije koja predstavlja miks tri taksonomije koje su predstavili Altenberg (1998), Bajber i saradnici (2004) i Mun (1989, prema Ebeling, 2022).

U svojim istraživanjima Bejker (2004) i Ebeling (2022) ispituju da li je učestalija upotreba ULS u prevedenim tekstovima prevodilačka univerzalija. Međutim, podaci o učestalosti upotrebe ULS u prevedenim i neprevedenim tekstovima se dobijaju i u istraživanjima sa drugim ciljevima (Biel, 2017; Siao, 2010). Bejker (2004) smatra da bi moglo da se pretpostavi da će se ULS češće koristiti u prevedenim tekstovima ne engleskom jeziku zato što se u anglofonim prevodima kao ključna vrednost uzima fluentnost teksta, što se pospešuje upotrebom ULS. To su rezultati istraživanja Ebeling (2022) i pokazali kao tačnim, jer su se ULS učestalije koristili u prevodima na engleski jezik uporedno sa originalima na norveškom, švedskom i nemačkom jeziku. Međutim, pošto na osnovu rezultata iz istraživanja koje je sprovedla Biel (2017) vidimo da se ULS češće koriste i u prevedenim tekstovima sa engleskog jezika na poljski uporedno sa poljskim referentnim korpusom, a istraživanje koje je sproveo Siao (2010) ukazuje na češću upotrebu ULS u prevedenim tekstovima sa engleskog na kineski, uporedno sa referentnim kineskim korpusom, dolazimo do još dokaza koji ukazuju na to da je učestalija upotreba ULS zapravo jezička univerzalija. Trebalo bi, dakle, sprovesti istraživanja u kojima engleski jezik nije ni izvorni ni ciljni jezik kako bi se ustanovilo da zapravo češća upotreba ULS nije zbog uticaja karakteristika engleskog jezika).

Produkcija neizvornih govornika određenog jezika razlikuje se od produkcije prevoda neizvornih govornika na ciljni jezik (Novita & Kwary, 2018). Ako učestalija upotreba ULS zaista jeste prevodilačka univerzalija ili ako su ULS prisutniji u prevodima zbog većeg uticaja transfera u prevedenim tekstovima i/ili Turijevo zakona rastuće standardizacije i/ili hipoteze normalizacije (Baker, 1996, prema Biel, 2017), to može imati jasnu implikaciju na interpretaciju rezultata u istraživanju koje je sprovedeno u ovoj disertaciji, jer se pretpostavlja da je određen procenat radova neizvornih govornika zapravo preveden tekst. Tačan procenat je nepoznat zbog nepostojanja regulacija u naučnom publikovanju koje zahtevaju navođenje da li su publikovani radovi prevodi i ko su prevodioci. Bilo bi korisno da postoji istraživanje koje bi uporedilo naučne radove koje su napisali neizvorni govornici engleskog jezika sa prevodima naučnih radova od strane prevodilaca koji prevode na strani jezik (npr. radovi na engleskom jeziku koji su napisali izvorni govornici srpskog jezika i prevedeni radovi sa srpskog jezika na engleski gde su prevodioci izvorni govornici srpskog jezika). Na taj način moglo bi da se utvrdi da li postoje razlike u upotrebi ULS kod ove dve grupe tekstova da bi interpretacija podataka bila jasnija i kako bi se odgovorilo na još jedno pitanje u ovoj oblasti.

Pored ispitivanja upotrebe ULS u prevedenim i neprevedenim tekstovima, istražene su razlike u upotrebi ULS kod različitih prevodilaca, konkretno kod prevodilaca koji prevode na svoj maternji jezik i koji prevode na jezik koji im je strani (Liu i Afzal, 2021). Interesantno je uporediti rezultate ovakvih istraživanja sa rezultatima istraživanja koja ispituju upotrebu ULS kod izvornih i neizvornih govornika i da li će se iste tendencije pojaviti i u prevodu. Naime, većina istraživanja ukazuje na to da neizvorni govornici češće koriste ULS od izvornih govornika (Pan i sar., 2016; Güngör, 2016; Lazić, 2017), a rezultati istraživanja Liu i Afzal (2021) pokazali su da je izvorni

²²Jedan od izuzetaka su istraživanja ULS koja imaju cilj da identifikuju uticaj transfera, gde je isto neophodno izdvojiti relevantne ULS i na njima uraditi detaljnu analizu (Jarvis, 2000; Güngör & Uysal, 2020).

govornik engleskog jezika koristio više ULS, što ovde može da ukaže na to da iste tendencije ne važe kod prevoda, da je to stvar prevodilačkih stilova dotičnih prevodilaca ili pak nešto treće. Više istraživanja je neophodno da bi se došlo do preciznih i čvrstih zaključaka. Imali smo prilike da vidimo i razliku u upotrebi ULS između profesionalnih prevodilaca i studenata. Tako, između ostalog, češća upotreba ULS kod profesionalnih prevodilaca može da se objasni njihovim boljim ispunjavanjem standarda fluentnosti.

Značajno je osvrnuti se i na metodološke postavke sagledanih istraživanja. Neka istraživanja u oblasti translatologije (Biel, 2017; Lui i Afzal, 2021) primenjuju slične parametre i postavke kao što se koriste u frazeološkim istraživanjima akademskog diskursa, neka se delimično usklađuju a delimično uključuju nove elemente ili elemente koji su u skladu sa drugačijim ciljevima (Berūkštienē, 2017; Ebeling, 2022), dok određeni broj radova predlaže nove metodološke postavke koje će bolje služiti ispunjavanju translatoloških ciljeva (Bejker, 2004; Li, 2013).

Iako je broj istraživanja koji spajaju translatologiju i ULS ograničen, predstavljena istraživanja u ovom poglavlju su daleko od sveobuhvatnog spiska istraživanja na navedenu temu. Glavni cilj njihovog predstavljanja je bio dobijanje uvida u istraživačke pravce u navedenoj oblasti i tipove istraživačkih pitanja na koje se traga za odgovorima, kao i izdvajanje zaključaka koji mogu da utiču na analizu podataka u navedenom istraživanju i utiču na formiranje obrazovnog materijala za podučavanje prevodilaca kako u obliku stručnog usavršavanja tokom procesa celoživotnog učenja, tako u obliku podučavanja na kursevima na osnovnim ili master studijama. Ono što je sigurno jeste da postoji velika potreba za daljim istraživanjima koja spajaju translatologiju i ULS. Daleko smo od rešavanja problema koji se javljaju pri identifikaciji i prevođenju jedinica od nekoliko reči, a koje se javljaju kako kod standardnog prevođenja tako i kod mašinskog prevođenja (engl. *Machine Translation*, Berūkštienē, 2017).

1.4. Transfer

Pitanje jezičkog transfera ne samo da je popularno u lingvističkim istraživanjima već je intuitivno prisutno i u jezičkim koncepcijama laika. Koliko nam se često dogodi da upoznamo nekog stranca, počnemo da pričamo na engleskom, te na osnovu samog jezika zaključimo da ni ta osoba nije izvorni govornik engleskog... Neretko pokušavamo čak i da pogodimo – na osnovu načina govora – koje je nacionalnosti osoba sa kojom razgovaramo. Ovde se vidi da i filološki laici osete/podrazumevaju uticaj maternjeg jezika na produkciju stranog, a činjenica da često uspemo tačno da pretpostavimo nacionalnost govornika ukazuje na opipljivost tih uticaja (Odlin, 1989). Dalje, akcentujemo da je uticaj učenikovog maternjeg jezika na usvajanje stranog više istraživano od bilo koje druge tematike u oblasti usvajanju stranog jezika (Ellis, 2015). Džarvis i Pavlenko (Jarvis & Pavlenko, 2008) su naveli sedamnaest knjiga koje su posvećene samo transferu, a od tada je, naravno, objavljeno još. Tolika količina pažnje ilustruje prisutnost uticaja transfera na usvajanje stranog jezika, njegov značaj za razumevanje usvajanja stranog jezika, kao i kompleksnost samog fenomena (Ellis, 2015). Trenutno ne postoji teorija usvajanja stranog jezika koja ne uvažava značaj uticaja prvoga jezika. Većina teorija ističe da je neophodno sagledati uticaj maternjeg jezika na usvajanje stranog kao samo jednog od brojnih faktora koji utiču na proces usvajanja stranog jezika (Ellis, 2015). Džarvis (Jarvis, 2000: 246) je elokventno opisao mesto koje transfer zauzima na lingvističkoj sceni u sledećem citatu:

Moguće je da nijedna oblast u istraživanju drugog jezika nije dobila toliku količinu pažnje a ostala do te mere misteriozna kao uticaj prvog jezika (L1). Uprkos mnogobrojnim istraživanjima koji su sprovedeni u proteklim decenijama, i dalje postoji iznenađujući nivo konfuzije u oblasti po pitanju kada, gde, u kom obliku i do koje mere će se uticaj prvog jezika manifestovati u znanju ili upotrebi drugog jezika kod učenika.²³

Istraživanje transfera ima i svoje praktične aspekte. Profesori stranih jezika efikasnije mogu da rade svoj posao kada sagledaju razlike između jezika i kultura. Primera radi, profesor engleskog jezika koji je svestan grešaka koje izvorni govornici srpskog i rumunskog prave zbog uticaja transfera – uspešnije će moći da identifikuje te greške i učenicima skrene pažnju na iste, dok u isto vreme uvažava značaj učenikovog lingvističkog i kulturnog porekla. Ova vrsta istraživanja, takođe, pomaže u identifikaciji univerzalnih karakteristika učenika stranih jezika uporedno sa specifičnim karakteristikama određenih grupa neizvornih govornika kao što su, na primer, neizvorni govornici engleskog kojima je maternji jezik srpski (Odlin, 1989). I u ovoj disertaciji *jedan od sekundarnih ciljeva* jeste identifikacija uticaja transfera odnosno međujezičkog uticaja na upotrebu *učestalih leksičkih spojeva*, a kako bismo dobili jasniju sliku o tome na koji način neizvorni govornici koriste ULS, koje su univerzalne karakteristike neizvornih govornika, koje su specifične za izvorne govornike srpskog a koje za izvorne govornike rumunskog. Da bi takav *cilj* bilo moguće dostići, prvo je neophodno sagledati opšte informacije o fenomenu transfera, kao i prethodna istraživanja koja su se bavila fenomenom transfera i njegovim uticajem na upotrebu ULS. Te informacije daju se u ovom poglavlju.

1.4.1. Definicija transfera

Postoji veliki broj definicija za koncept *transfera*. Sinkler smatra: „najbolje je posmatrati transfer kao širi pojam koji obuhvata čitavu kategoriju ponašanja, procesa i ograničenja²⁴” (1992: 208, kao

²³ „Perhaps no area of second language research has received as much attention and remained as elusive as the influence of the first language (L1). Despite the myriad studies that have been conducted in this area over the past four decades, there still remains a surprising level of confusion in the field concerning when, where, in what form, and to what extent L1 influence will manifest itself in learners’ use or knowledge of a second language (L2).”

²⁴[...] „language transfer is best thought of as a cover term for a whole class of behaviours, processes, and constraints”

što je citirano u Jarvis, 2000: 250). Dodatno problematiku definisanja transfera usložnjava pitanje da li bi transfer trebalo da se posmatra kao: proces prenošenja odnosno interjezičke identifikacije između dva jezika (Oldin, 1989), ograničenje koje učenici imaju pri stvaranju hipoteza o funkcionisanju drugog jezika (Sinkler, 1992 prema Jarvis, 2000), strategija koju učenici koriste za popunjavanje „rupa” u znanju drugog jezika (Krashen, 1983 prema Jarvis, 2000) ili prosto inertni ishod deljenja konceptualnog sistema koji daje potporu za sistem maternjeg jezika i stranog jezika (Jarvis, 1998, prema Jarvis, 2000). Oldin (1989) predlaže sledeću popularnu radnu definiciju koncepta transfera: „Transfer je uticaj koji proističe iz sličnosti i razlika između ciljnog jezika i bilo kog drugog jezika koji se ranije (i potencijalno nesavršeno) usvojio”²⁵ (str. 27). Kako bi se jasno razumela definicija koju predlaže Odlin, važno je da razumemo *koncepte ciljnog jezika i izvornog jezika u kontekstu proučavanja transfera*. Naime, transfer ne nastaje samo u kontekstu kada se uči strani jezik pa je taj proces pod uticajem maternjeg jezika koji govornik već zna. Transfer je kompleksniji i širi pojam, te možemo da govorimo o transferu, odnosno međujezičkom uticaju koji ima strani jezik na maternji, ili u slučaju multilingvalnih osoba na koji način svi jezici koje osoba poznaje utiču jedni na druge. U navedenom kontekstu izvorni jezik je onaj koji utiče na ciljni jezik, odnosno ispituje se uticaj transfera od izvornog jezika na ciljni.

Sinkler je u svom radu koristio sledeću radnu definiciju transfera: „radno je definisan kao proces koji nastane kad god se statistički značajan ... [obrazac u maternjem jeziku] ponovo pojavi u međujezičkom ponašanju govornika”²⁶ (1992: 201, prema Jarvis, 2000: 251). U navedenim definicijama Oldin (1989) smatra da je najbolji način provere postojanja transfera poređenje produkcije govornika sa različitim maternjim jezicima, dok Sinkler stavlja akcenat na poređenje upotrebe maternjeg jezika i učeničkog međujezika. Džarvis, međutim, smatra da su obe vrste dokaza neophodne i da su različiti aspekti istog fenomena. Stoga, predlaže sledeću radnu definiciju: „Uticaj L1 odnosi se na bilo koji slučaj podataka o učeniku gde se pokazuje da postoji statistički značajna korelacija (ili veza zasnovana na verovatnoći) između neke karakteristike performansi u međujeziku učenika i njihovog maternjeg jezika”²⁷ (Jarvis, 2000: 252). Navedena definicija obuhvata i poređenje međujezika kod govornika sa različitim maternjim jezicima i poređenje upotrebe maternjeg jezika i međujezika.

Šarvud Smit i Kelermen (Sharwood Smith & Kellermen, 1986, kao što je citirano u Ellis, 2015) predložili su upotrebu termina *međujezički uticaj* umesto termina *transfer* zato što predstavlja teorijski neutralan način da se iskaže da su jezici u interakciji. S obzirom na to da će u ovoj disertaciji fokus biti na uticaju maternjeg jezika na usvajanje stranog koristiće se termin *transfer*, koji se i češće koristi.

1.4.2. Tipovi transfera

Transfer može da bude negativan i pozitivan. Negativni transfer je kada izvorni jezik ima negativan uticaj na ciljni (Odlin, 1989). Na primer, pogrešno izgovaranje određene foneme u ciljnom jeziku koja je slična fonemi u izvornom, ili kolokaciona greška gde se pogrešan pridev iskoristi uz imenicu zbog uticaja konstrukcije koja postoji u izvornom jeziku. U istraživanjima, pogotovo u istraživanjima koja se fokusiraju na analizama grešaka, češće se stavlja fokus na negativan transfer.

Postoji nekoliko tipova *negativnog transfera* (Odlin, 1989; Houssos, 2020):

²⁵ „Transfer is the influence resulting from similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired.”

²⁶ [...] „operationally defined as a process which occurs whenever a statistically significant [pattern in the L1] reappears in IL behaviour”

²⁷ „L1 influence refers to any instance of learner data where a statistically significant correlation (or probability-based relation) is shown to exist between some feature of learners’ IL performance and their L1 background”

1. *Nedovoljna produkcija* – kada učenici koriste veoma mali broj ili uopšte ne koriste određene jezičke strukture koje se češće koriste u produkciji izvornih govornika. Jedan čest i upečatljiv oblik nedovoljne produkcije je *izbegavanje* (engl. *avoidance*) do kojeg dolazi kada učenici osete da je određena struktura u stranom jeziku mnogo drugačija od struktura u izvornom jeziku, te pokušavaju da izbegnu upotrebu navedene strukture.
2. *Prekomerna produkcija* – nastaje kada učenici češće koriste određene strukture nego što ih koriste izvorni govornici. Do prekomerne produkcije može da dođe i zbog drugih faktora, kao što su prekomerna generalizacija (engl. *overgeneralization*) ili zbog *efekta plišanog mede* (Ellis, 2012), a ne samo transfera. Kao dobar primer prekomerne produkcije zbog uticaja transfera Odlin navodi prečestu upotrebu izvinjenja kod izvornih govornika engleskog koji su koristili hebrejski kao strani. Objašnjenje za ovakvu pojavu je činjenica da su izvinjenja mnogo češća u engleskom nego u hebrejskom.
3. *Greške u produkciji* – u govoru i u pisanju postoji tri tipa grešaka koje su posebno podložne nastajanju zbog uticaja sličnosti i razlika između jezika, a one su:
 - *supstitucija* (kada učenik iskoristi reč iz svog maternjeg jezika u svojoj produkciji na stranom ili obrnuto ako govorimo o uticaju stranog jezika na maternji jezik);
 - *kalk* (kada učenici zbog doslovnog prevoda naprave forme koje ne postoje u stranom jeziku; do kalkova često dolazi zbog bukvalnog prevoda idiomatskih izraza);
 - *alternacije struktura* (između ostalog, podrazumeva preterano ispravljanje kod učenika; primerice, kod neizvornih govornika engleskog kojima je maternji jezik srpski može da dođe do dupliranja slova i u rečima gde to ne treba zato što u srpskom jeziku nije tipično da stoje dva ista slova jedno do drugog pa u bojazni da ne propuste da stave duplo slovo gde treba, stave ga i tamo gde ne treba – napisati *bannana* umesto *banana*).
4. *Pogrešna interpretacija* – kao što i sâm naziv greške kazuje, do njih dolazi kada zbog uticaja razlika u jezicima, kao što su na primer različiti redosledi u rečenici ili različite kulturološke pretpostavke, dođe do pogrešnog razumevanja.

Pozitivan transfer je kada uticaj izvornog jezika pospešuje usvajanje ciljnog jezika. Do pozitivnog transfera dolazi kada postoje sličnosti između izvornog i ciljnog jezika koje ubrzavaju i olakšavaju proces učenja. Najočiglednije je kod odraslih učenika koji imaju bogat vokabular u izvornom jeziku pa mogu da prepoznaju internacionalizme ili reči koje imaju isto lingvističko izvođenje te se njihovo značenje lakše prepozna i poveže. Pored pozitivnog i negativnog transfera, uticaj maternjeg jezika može da dovede do različitog trajanja procesa usvajanja jezika.

1.4.3. Opseg uticaja transfera

Važno je napomenuti da uticaji transfera na jezik mogu da budu sistemski (engl. *system-wide effects*) ili specifični za određene jezičke pojave (engl. *item-specific effects* – Džarvis i Pavlenko, 2008: 58). Sistemski uticaj je prisutan u slučajevima kada je čitava leksička, morfološka, fonološka ili sintaktička klasa jednog jezika pod uticajem istog sistema u drugom jeziku. Primer za sistemski uticaj može se sagledati u tendenciji neizvornih govornika engleskog kojima je maternji jezik ruski da se oslanjaju na ruski sistem aspekta kada koriste označitelje za vreme u engleskom (Wenzell, 1989, prema Džarvis i Pavlenko, 2015). Specifični uticaj transfera ograničen je na određene jezičke pojave. Primer ovog opsega uticaja transfera može se naći kod neizvornih govornika engleskog kojima je maternji jezik finski. Oni imaju tendenciju da koriste neodređeni član ispred nebrojive imenice *weather*, kao u primeru: *It's a nice weather today*. Moglo bi da se pretpostavi da ovaj tip govornika ima problem sa korišćenjem neodređenog člana u engleskom, međutim, ova tendencija

je prisutna samo kod određenog broja imenica (vreme, vesti i hleb) koje su u finskom brojive ili imenice sa dvojnomo brojivošću (Džarvis i Pavlenko, 2008).

1.4.4. Istraživanje transfera/ međujezičkog uticaja

Transfer odnosno međujezički uticaj može se istraživati kao *društveni fenomen* ili kao *psiholingvistički* (Jarvis i Pavlenko, 2008). Kada se istražuje kao društveni fenomen, fokus se stavlja na kontakt između jezika kroz istoriju tokom njihovog razvoja (na primer, uticaj turskog jezika na srpski u pogledu postojanja velikog broja turcizama). Kada se istražuje kao psiholingvistički fenomen, fokus se stavlja na uticaj jednog jezika (obično L1) na drugi (obično L2, iako se može istraživati i obrnut uticaj ili uticaj nekoliko već usvojenih jezika na drugi) i to na nivou individue. Cilj ove oblasti istraživanja je da se ustanove mentalni procesi koji stoje iza postojanja transfera, identifikacija unutrašnjih (kognitivnih, konceptualnih i afektivnih) i kontekstualnih (lingvističkih i društvenih) faktora koji izazivaju i/ili onemogućavaju pojavljivanje transfera, kao i uticaj transfera na usvajanje stranog jezika (Jarvis i Pavlenko, 2008: 29).

Istraživanja transfera kao psiholingvističkog fenomena dele se u dve grupe na osnovu pristupa koji je primenjen: *intrasubjektivni* i *intersubjektivni*. Istraživanja koja primenjuju intrasubjektivni pristup ispituju fenomen transfera na nivou individue, obično su studije slučaja. Istraživanja koja primenjuju intersubjektivni pristup istražuju transfer kroz ispitivanje jezičkih obrazaca u relativno velikim dobro definisanim grupama ispitanika/govornika. Prednost ovog tipa istraživanja je činjenica da rezultati mogu da se generalizuju, dok je prednost prethodnog tipa u visokom nivou validnosti interpretacije rezultata.

U postavci istraživanja značajno je odrediti koji će se tip podataka prikupljati. Elis (Ellis, 1994 prema Jarvis i Pavlenko, 2008:33) navodi pet osnovnih tipova podataka u istraživanjima usvajanja drugog/stranog jezika:

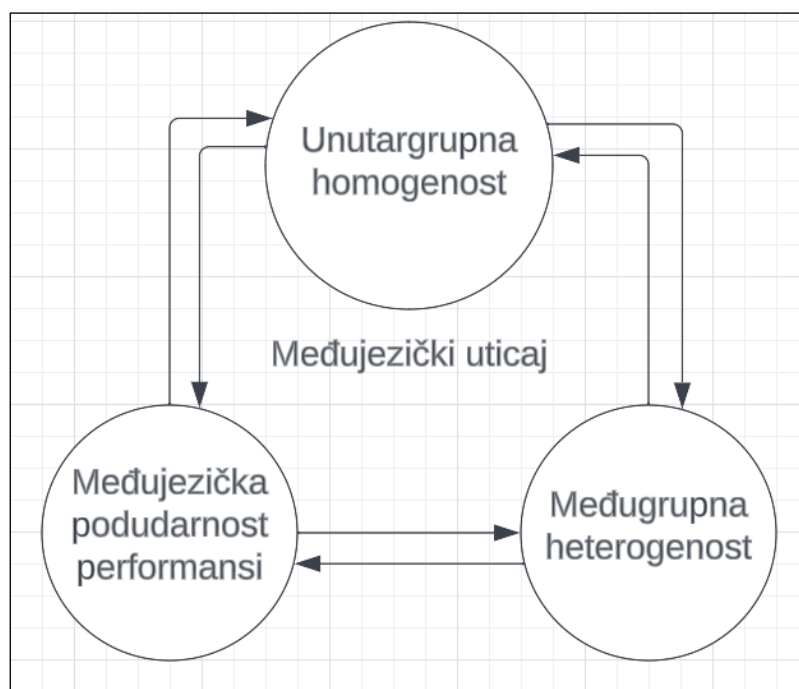
1. *prirodna upotreba jezika* (opservacija prirodnog korišćenja čija produkcija nije podstaknuta od strane istraživača);
2. *klinički podstaknuta upotreba jezika* (podstaknuta ali nevođena upotreba jezika, npr. prepričavanje odgledane predstave);
3. *eksperimentalno podstaknuta upotreba jezika* (vođena upotreba jezika, npr. jezički testovi zatvorenog tipa);
4. *metalingvističke procene* (npr. zadaci procene gramatičnosti određenog izraza);
5. *samoprocenjivanje* (npr. introspekcija, zadaci u kojima je neophodno vokalizovati proces razmišljanja).

Svaki od navedenih tipova podataka ima svoje karakteristične prednosti i mane, a nijedan nije pogodniji za istraživanje transfera od drugih. Istraživanja koja primenjuju više tipova podataka daju potpuniju sliku istraženog fenomena. Džarvis i Pavlenko (2008) daju primer studije slučaja Jarvis (2003) u kojoj su korišćena četiri tipa podataka (1, 2, 4. i 5. tip podataka od gorenavedenih tipova). Istraživanje je specifično jer pokazuje da bi zaključci istraživanja bili pogrešni da se samo jedan tip podataka koristio. Naime, istraživanje se bavilo uticajem stranog jezika na maternji jezik i utvrđeno je da je došlo do međujezičkog uticaja koji se ogleda u dodavanju novih opcija u lingvističkoj ekspresiji individue koji se oslanjaju na uticaj stranog jezika. Navedene promene koje nisu systemske, nisu bile vidljive u drugom tipu podataka, pojavljivale su se samo sporadično u četvrtom tipu podataka, pojavljivale su se učestalo u prvom tipu podataka, dok je peti tip podataka omogućio istraživačima da zaključuje kojih je promena subjekat zaista bio svestan.

1.4.5. Identifikacija transfera i tipovi dokaza

Na osnovu pristupa zasnovanog na detekciji (engl. *detection based approach*), koji je predložio Džarvis (Jarvis, 2000,2012; Jarvis i Pavlenko, 2008) u identifikaciji transfera odnosno međujezičkog uticaja potrebno je zadovoljiti tri tipa dokaza (v. Grafikon 3):

1. *unutragrupna homogenost (intragroup homogeneity)*;
2. *međugrupna heterogenost (intergroup heterogeneity)*;
3. *međujezička podudarnost performansi (crosslinguistic performance congruity)*.



Grafikon 3. Identifikacija transfera²⁸

1) *Unutargrupna homogenost* je fenomen koji nastaje kad god se grupa govornika sa poredivim karakteristikama (koji imaju sličan nivo znanja istog maternjeg jezika i sličan nivo znanja istog stranog jezika) jezički ponaša na sličan način u ciljnom jeziku (Jarvis i Pavlenko, 2008). Prikupljanje ove vrste dokaza podrazumeva detaljnu analizu karakteristika u korišćenju međujezika govornika sa istim maternjim jezikom (Jarvis, 2000). Logika za korišćenje unutargrupne homogenosti kao indikator transfera odnosno međujezičkog uticaja je sledeća: ako grupa govornika ima slično znanje izvornog jezika, njihovo korišćenje tog jezika bi trebalo da bude na istom nivou i da ima slične karakteristike; ako izvorni jezik utiče na ciljni jezik (najčešće je to dinamika L1 uticaj na L2), onda se očekuje da će ti govornici imati iste karakteristike pri korišćenju ciljnog jezika (Jarvis i Pavlenko, 2008). Drugim rečima, ako se jedan govornik zbog međujezičkog uticaja jezički ponaša na određen način, očekuje se da će i drugi govornici sa istim karakteristikama ponašati na isti način, odnosno da će se u njihovom međujeziku pokazati iste karakteristike. Ilustracije radi, ako govornici srpskog jezika određenog nivoa imaju probleme u upotrebi članova pri korišćenju engleskog jezika koji je njima strani, možemo pretpostaviti da je njihovo neispravno korišćenje članova potencijalna posledica negativnog transfera. Iako ovaj tip dokaza pruža značajnu potporu u argumentovanju da su određene karakteristike međujezika posledica transfera, oni sami nisu dovoljan dokaz za utvrđivanje postojanja transfera.

²⁸Grafikon 3 je preuzet iz istraživanja Gungor i Uisal (2020), a preveden od strane autorke disertacije.

2) *Međugrupna heterogenost* je drugi tip dokaza koji predlaže Džarvis. Ona predstavlja razlike u jezičkoj upotrebi govornika koji se razlikuju po njihovom znanju izvornog ili ciljnog jezika. Dokazi za postojanje međujezičke heterogenosti se mogu prikupiti iz tri tipa dokaza (Jarvis i Pavlenko, 2008):

- poređenje upotrebe ciljnog jezika kod dve ili više grupa govornika čije je znanje ciljnog jezika poredivo, ali koji imaju drugačije izvorne jezike (npr. poređenje upotrebe engleskog jezika kao stranog kod izvornih govornika ruskog, španskog i kineskog jezika)²⁹;
- poređenje monolingvalnih i bilingvalnih govornika istog ciljnog jezika (npr. poređenje engleskih monolingvalnih govornika sa bilingvalnim govornicima nemačkog i engleskog pri ispitivanju nemačkog međujezičkog uticaja na engleski);
- poređenje monolingvalnih i bilingvalnih govornika istog izvornog jezika (npr. poređenje nemačkih monolingvalnih govornika sa bilingvalnim govornicima nemačkog i engleskog pri ispitivanju nemačkog međujezičkog uticaja na engleski).

3) *Međujezička podudarnost performansi* je na prvi pogled slična unutargrupnoj homogenosti, međutim, u ovom tipu dokaza fokus je na kvalitativnoj analizi konkretnih jezičkih struktura i šema koje se koriste. Ako bismo istraživali uticaj maternjeg jezika na strani, prikupljanje dokaza za međujezičku podudarnost performansi bi podrazumevalo detaljnu analizu sličnosti između maternjeg jezika govornika i karakteristika njihovog međujezika. Ako uzmemo isti primer koji smo koristili za opisivanje unutargrupne homogenosti odnosno upotrebu članova na engleskom jeziku kod izvornih govornika srpskog, pri prikupljanju podataka o međujezičkoj podudarnosti performansi uzela bi se u obzir činjenica da srpski jezik ne koristi članove kao engleski, istražilo bi se na koji način srpski jezik prenosi istu vrstu značenja i proverila bi se podudarnost tih podataka sa potencijalnim negativnim transferom koji je identifikovan.

Džarvis (Jarvis, 2000, 2012; Jarvis i Pavlenko, 2008) smatra da bi prva dva tipa dokaza uvek trebalo da se tumače zajedno i definišu u odnosu jedan prema drugom. On, takođe, navodi da je poželjno imati sve tri vrste dokaza kako bi postojanje transfera odnosno međujezičkog uticaja moglo sa sigurnošću da se tvrdi tj. da bi se postigao zadovoljavajući nivo metodološke strogosti. Međutim, Džarvis i Pavlenko ipak ističu da nije uvek neophodno da istraživanje obezbedi sva tri tipa dokaza da bi se zadovoljio strog kriterijum pouzdanosti (str.48). Naime, ponekad se određeni tip dokaza može obezbediti kroz prethodna istraživanja, kroz neke druge izvore informacija koji su van samog istraživanja ili se čak može podrazumevati. Pri prikupljanju dokaza za međujezičku podudarnost performansi, istraživači su često poredili performanse u ciljnog jeziku sa intuicijom ili opštim opisima (pisane gramatike) o tome kako izvorni jezik funkcioniše. Autori smatraju da je prihvatljivo koristiti eksterne opise izvornog jezika umesto sakupljene podataka o performansama izvornih govornika kao osnovu za poređenje dokle god su opisi zaista deskriptivni a ne preskriptivni. Ne sme se zaboraviti da istraživači imaju zadatak i da objasne na osnovu čega smatraju da su dokazi koje su pribavili zapravo posledice transfera jer unutargrupna homogenost, međugrupna heterogenost i međujezička podudarnost performansi mogu da budu prisutni kao posledice drugih faktora. Na primer, visoki nivoi unutargrupne homogenosti može da bude posledica razvojnih faktora, visoki nivoi međugrupne heterogenosti mogu da budu posledica različitih obrazovnih sistema, dok međujezička podudarnost performansi može da bude posledica postojanja jezičkih univerzalija (Jarvis i Pavlenko, 2008).

²⁹Ovaj tip dokaza je značajan da bi moglo da se utvrdi da li su jezičke karakteristike koje se uvide kod korisnika ciljnog jezika opšte karakteristike neizvornih govornika ili su specifične za određenu grupu govornika zbog međujezičkog uticaja. Posebno je značajan kada se pravi razlika između jezičkih univerzalija, razvojnih grešaka i transfera odnosno međujezičkog uticaja.

Džarvis u svom kasnijem naučnom članku (Jarvis, 2010) dodaje četvrti tip dokaza koji naziva *intra lingvalni kontrast*, a koji podrazumeva detaljnu analizu razlika između maternjeg jezika govornika i karakteristika njihovog međujezika. Međutim, većina kasnijih istraživanja u domenu ispitivanja transfera na uticaj ULS koristi samo izvornu postavku pristupa zasnovanog na detekciji sa tri vrste dokaza – ista će se koristiti i u ovoj disertaciji.

1.4.6. Faktori koji utiču na transfer /međujezički uticaj

Pošto se transfer pojavljuje u širokom spektru društvenih konteksta, duboko razumevanje transfera zavisi od detaljnog razumevanja konteksta (Odlin, 1989). Potencijalni uticaj transfera na usvajanje stranog jezika može biti manji ili veći u zavisnosti od kompleksnih varijacija u društvenim okolnostima u kojima dolazi do transfera. Iako se faktor konteksta spominje i kod drugih autora (Ellis, 2015) koji stavljaju akcenat na kontekst učenja stranog jezika u smislu formalnog i neformalnog obrazovanja i tipova jezičkih zadataka, Odlin se fokusira na širi socijalni kontekst.

Džarvis (2000) nudi spisak devet faktora koje bi bilo poželjno identifikovati i kontrolisati u istraživanjima transfera: 1. *uzrast*; 2. *karakter, motivacija, sposobnost jezičkog usvajanja*; 3. *društvene, obrazovne i kulturološke karakteristike*; 4. *jezička pozadina*; 5. *tip i obim kontakta sa ciljnim jezikom*; 6. *jezički nivo ciljnog jezika*; 7. *jezička udaljenost između maternjeg i ciljnog jezika*; 8. *tip zadatka i jezička oblast*; 9. *karakteristike ispitane lingvističke forme*.

Elis (Ellis, 2015) navodi 12 faktora koji utiču na to da li će do transfera doći i do koje mere. Te faktore je podelio u pet kategorija: 1. *lingvistički faktori*; 2. *psiholingvistički faktori*; 3. *kontekstualni faktori*; 4. *razvojni faktori*; 5. *individualni faktori*.

1) *Lingvistički faktori* se odnose na lingvističke karakteristike izvornog i ciljnog jezika. Elis (2015) navodi tri tipa lingvističkih faktora:

- *Lingvističke razlike*. – Jedna od glavnih hipoteza kontrastivne analize je da učenici imaju više poteškoća pri usvajanju stranog jezika što je veća razlika između izvornog i ciljnog jezika, odnosno da je veći uticaj izvornog jezika što je veća razlika između izvornog i ciljnog jezika. Međutim, iako je gorenavedena tvrdnja često tačna, postoje situacije u kojima nije. Naime, Major i Kim (Major & Kim, 1996, prema Elis, 2015) predložili su hipotezu (*Similarity Differential Rate Hypothesis*) po kojoj zvukovi u stranom jeziku koji su dosta drugačiji od maternjeg mogu biti izazovni za usvajanje na početku, međutim, vremenom njihova upotreba se brže unapredi od zvukova koji su drugačiji ali slični. Uticaj faktora lingvističke razlike postaje još kompleksniji kada uzmemo u obzir da postoje jasni dokazi koji ukazuju na činjenicu da razlike između jezika dovode do otežanog usvajanja, kratkoročno i dugoročno, kada je reč o gramatici jezika. To se ogleda u hipotezi morfološke podudarnosti (*Morphological Congruency Hypothesis*) koja uspostavlja tvrdnju da učenje stranog jezika može biti lakše kada postoji gramatička morfema sa istom funkcijom u oba jezika, nego kada postoji u stranom ali ne postoji u izvornom. Drugim rečima, ne postoji jednostavna korelacija između jezičkih razlika i poteškoća u usvajanju stranog jezika, ali one bi trebalo da se uzmu u obzir kao faktor pri istraživanju jezičkog usvajanja i upotrebe, kao i uticaja transfera na jezičko usvajanje i upotrebu.
- *Jezička distanca ili jezička razlika* (*engl. language distance*). –Kao što je slučaj kod lingvističke razlike, ni kod jezičke distance ne postoji jednostavna korelacija već uticaji moraju detaljnije da se sagledaju. Kada je jezička distanca između dva jezika mala, dolazi i do pozitivnog i do negativnog transfera. Međutim, istina je – što je veća jezička distanca, to je teže naučiti strani jezik. Istraživanja koja se bave tipološkom analizom jezika pružaju osnovu za određivanje jezičke distance. Štaviše, tipološka istraživanja podstiču istraživanja

transfera koja fenomen sagledavaju kroz systemske uticaje, ali takođe i omogućavaju jasniju distinkciju između transfera i razvojnih sekvenci (Odlin, 1989:45).

- *Obeleženi lingvistički oblici* (engl. *Markedness*). – Obeleženi lingvistički oblici su oni koji su „posebni” u odnosu na druge. Funkcionalno-tipološki pristup ovom fenomenu ispituje njegovu prisutnost u različitim jezicima kako bi se izdvojile one karakteristike koje su univerzalne i one koje su specifične za određene jezike. Oni lingvistički oblici koji su česti u svetskim jezicima manje su obeleženi od onih koji su prisutni u manjem broju jezika. Učenici stranog jezika pokazuju veću spremnost da prenesu neobeležene lingvističke karakteristike nezavisno od toga da li je ta karakteristika u maternjem jeziku obeležena ili ne. Međutim, Đarvis i Pavlenko (2008: 187) navode da „postoji velika količina dokaza koji pokazuju da je čest transfer i obeleženih i neobeleženih sintaktičkih struktura”³⁰.

Glavni problem sa pristupom za ispitivanje lingvističkog transfera, u kojem je jezička razlika poistovećena sa poteškoćama pri usvajanju jezika, jeste činjenica da je „različitost” lingvistički koncept dok je „težina” usvajanja psihološki. Ne može se pretpostaviti da će elementi koje su lingvisti identifikovali kao teške učenici doživeti na isti način.

2) *Psiholingvistički faktori* se odnose na učenikovu percepciju o tome koliko su određeni elementi maternjeg jezika podložni transferu. Elis (2015) navodi dva tipa psiholingvističkih faktora:

- *Psihotipologija*. – Predstavlja učenikovu percepciju o tome koliko su dva jezika drugačija nezavisno od prave razlike između jezika. Učenička percepcija o tome koliko su maternji i strani jezik drugačiji je značajna zato što na osnovu nje učenici odlučuju da li da se oslanjaju na svoje znanje maternjeg jezika pri korišćenju stranog. Drugim rečima, psihotipologija utiče na to koliko su učenici spremni da se oslone na transfer pri učenju stranog jezika.
- *Prototipičnost*. – Predstavlja učenikove percepcije o tome šta je podložno transferu a šta ne. Pokazalo se da učenici imaju intuiciju o tome koji elementi njihovog maternjeg jezika mogu potencijalno da se prenesu na strani. Istraživanja su pokazala da učenici češće smatraju da su osnovna značenja reči (značenja koja se prvo opisuju u rečniku) podložnija transferu.

3) *Kontekstualni faktori* se odnose na načine na koje su učenici izloženi stranom jeziku. Elis (2015) navodi dva tipa kontekstualnih faktora:

- *Makrokontekstualni faktori*. – Prevažodno se odnose na to da li se jezik uči u informalnim ili formalnim uslovima. Odlin (1989) ukazuje na to da je negativan transfer manje čest u formalnom kontekstu gde se češće ispravljaju korišćenje maternjeg jezika i negoduje isto. Štaviše, u tom okruženju se podstiče praćenje jezičkog autputa u tipovima zadataka gde učenici imaju priliku za to, odnosno u zadacima koji se fokusiraju na preciznost umesto na fluentnost. U neformalnim uslovima situacija je kompleksnija i prisutnost transfera zavisi od toga da li su uslovi „fokusiranog” ili „nefokusiranog” tipa. U fokusiranim uslovima pravi se jasna distinkcija između dva jezika, dok se u nefokusiranim uslovima dva jezika redovno mešaju bez pridavanja mnogo pažnje na pragmatičnost i ispravnost jezika. Naravno, negativan transfer je prisutniji u neformalnim kontekstima. Generalizacije o efektima mikrokontekstualnih faktora su opasne zato što uzimaju u obzir samo negativan transfer. Kako su Džarvis i Pavlenko (2008: 208) istakli: „formalnost konteksta utiče na transfer, ali ona sadrži uticaje koji podstiču i ograničavaju nastanak transfera”³¹.
- *Mikrokontekstualni faktori*. – Prevažodno se odnose na to koliko učenici obraćaju pažnju na jezičku formu zbog tipa zadataka koje bi trebalo da ispune. Na primer, pri prevođenju,

³⁰[...] „there is ample evidence to show that transfer of both marked and unmarked syntactic structures is quite common”

³¹[...] „the formality of the context does affect transfer, but it includes forces that both foster and constrain the occurrence of transfer”

veća je verovatnoća da će doći i do pozitivnog i do negativnog transfera nego u zadacima koji iziskuju spontanu jezičku produkciju. Mikrokontekstualni faktori se, takođe, odnose i na različite vrste komunikativnih situacija u pravom životu. Na primer, neizvorni govornici će se pre osloniti na transfer u osetljivim društvenim situacijama nego u situacijama u kojima je učtivost manje značajna.

4) *Razvojni faktori* se odnose na to kako univerzalne tendencije i razvojne tendencije pri usvajanju stranog jezika utiču na transfer. Elis (2015) navodi tri tipa razvojnih faktora:

- *Preterano generalizovanje*. – Inicijalno, na početničkom nivou učenici se više oslanjaju na svoje znanje maternjeg jezika i prave veći broj grešaka zbog negativnog transfera, dok se kasnije, na višim nivoima, oslanjaju na znanje stranog jezika koje su već usvojili pa prave greške preterane generalizacije. Oba ova tipa grešaka nastaju zbog oslanjanja na prethodno znanje, samo je razlika u tome na koje se prethodno znanje oslanjaju.
- *Redosled usvajanja*. – Ovaj faktor se odnosi na činjenicu da može da dođe do nedoslednosti u prirodnom redosledu usvajanja jezika ako se desi da gramatička morfema koja nosi značenje u stranom jeziku nema ekvivalent u maternjem.
- *Sekvence usvajanja*. – Uticaj transfera može da ubrza ili uspori prirodnu sekvencu usvajanja stranoga jezika.

5) *Individualni faktori* podrazumevaju uticaj individualnih razlika učenika na verovatnoću da će doći do transfera u njihovoj jezičkoj produkciji. Elis (2015) navodi dva tipa individualnih faktora:

- *Godine starosti*. – Mlađi učenici kojima je oskudnije znanje maternjeg jezika manje se oslanjaju na maternji jezik te dolazi do manje količine transfera uporedno sa starijim/odraslim učenicima. Međutim, kod odraslih učenika koji imaju bogat vokabular dolazi i do veće količine pozitivnog transfera.
- *Sposobnost jezičkog usvajanja*. – Učenici sa visokom sposobnošću za jezičko usvajanje bolji su u transferu veština iz svog maternjeg jezika na svoje učenje stranog. Učenici sa većom radnom memorijom bolje blokiraju svoj uticaj maternjeg jezika pa na taj način izbegavaju negativan transfer.

1.4.7. Relevantna prethodna istraživanja uticaja transfera na upotrebu učestalih leksičkih spojeva

Analizirati uticaj transfera na upotrebu ULS je značajno zato što nam može pružiti uvid u razloge za postojanje velikih razlika u upotrebi ULS kod izvornih i neizvornih govornika, čak i kada su izvorni govornici profesionalci sa visokim nivoom znanja stranog jezika. Ova vrsta analize će nas potencijalno dovesti korak bliže do razumevanja zašto kada čitamo u potpunosti gramatički i leksički ispravan naučni rad ipak „osetimo” da je napisan od strane neizvornog govornika engleskog jezika. Kao što već znamo, ULS služe kao „konvencionalni gradivni blokovi koji se koriste kao zgodne rutine u jezičkoj produkciji” (Altenberg, 1998: 122); oni su specifični za registre i oblasti te ukazuju na pripadanje ili ne pripadanje određenoj jezičkoj zajednici. Pošto govornici nisu svesni ULS dok ih koriste (osim ako nisu imali prilike da se njima svesno bave tokom nastave, što još nije široko primenjena praksa, a i u tom slučaju je pitanje do koje mere bi bili svesni svih ULS koje koriste), moguće je da je nedovoljno fluentna produkcija neizvornih govornika delom rezultat baš te prekomerne, nedovoljne ili neprikladne upotrebe ULS od strane neizvornih govornika.

Istraživači u određenom broju prethodnih istraživanja objašnjavali su pogrešno, prekomerno i nedovoljno korišćenje ULS negativnim uticajem maternjeg jezika (Paquot, 2013). U istraživanju

Alen (Allen, 2011: 111, prema Paquot, 2013) prekomerno korišćenje ULS *it can be said that* u produkciji učenika kojima je izvorni jezik japanski, objašnjava se uticajem maternjeg jezika zato što se prevodni ekvivalent često koristi u akademskom pisanju na japanskom jeziku. Rika (Rica, 2010, prema Paquot, 2013) ističe da je veliki broj sekvenci od nekoliko reči koje prekomerno koriste neizvorni govornici engleskog kojima je maternji jezik španski, a koje su veoma slične konstrukcijama koje prenose isto značenje na španskom. Lu i Deng (Lu & Deng, 2019, prema Kostromitina, 2022) objasnili su nedovoljno korišćenje ULS sa neodređenim članom u produkciji neizvornih govornika kojima je maternji jezik kineski uticajem transfera zbog činjenice da kineski jezik nema članove. Međutim, mali broj istraživanja je kao cilj imalo da istraži uticaj transfera u upotrebi ULS. U daljem tekstu opisaćemo baš ta malobrojna istraživanja i to po redosledu relevantnosti za istraživanje koje će se sprovesti u ovoj disertaciji.

Paket (2013)

Paket (Paquot, 2013) je sprovela prvo istraživanje sa ciljem da prouči uticaj transfera na upotrebu ULS. Primenila je Džarvisov (2000) pristup zasnovan na detekciji u postavci istraživanja i analizi podataka. Istraživanje se bavi upotrebom ULS od tri reči sa leksičkim glagolom kod studenata koji uče engleski jezik a kojima je maternji jezik francuski. Težilo je da odgovori na dva istraživačka pitanja: *Koji se procenat idiosinkratične upotrebe ULS sa leksičkim glagolom kod izvornih govornika francuskog jezika može objasniti postojanjem transfera? Koji je tip uticaja transfera (npr. transfer forme, transfer funkcije) najuočljiviji?*

Autorka kreće od hipoteze da ULS mogu da budu pod uticajem transfera zato što su u suštini semantički i sintaksički kompozicioni pa samim time i neoznačene (engl. *unmarked*) kombinacije. Kako bi ispitala hipotezu i odgovorila na istraživačka pitanja, autorka je koristila nekoliko korpusa, na koje ćemo se nadalje ukratko osvrnuti.

- *International Corpus of Learner English* (ICLE) je korpus studentske produkcije na engleskom. Sastoji se od eseja studenata druge, treće i četvrte godine studija sa naprednim nivoom znanja engleskog jezika. Deset maternjih jezika je predstavljeno u korpusu (francuski, češki, danski, finski, nemački, italijanski, poljski, ruski, španski i švedski). Pošto je glavni fokus istraživanja upotreba ULS kod izvornih govornika francuskog jezika, potkorpus francuskih studenata koji se sastoji od 228 eseja i 136.343 reči je imao centralnu poziciju u istraživanju i koristio se za prikupljanje dokaza za unutargrupnu homogenost. Potkorpusi sa informacijama o drugim jezicima korišćeni su za prikupljanje dokaza o međugrupnoj heterogenosti. Zastupljenost različitih jezika je neujednačena u ICLE korpusu, koji se sastoji od ukupno 1641 eseja, odnosno veličine je 1.165.523 reči. Potkorpus studenata kojima je maternji jezik italijanski je najmanji sa 79 eseja i 47.739 reči, dok je najveći korpus izvornih govornika ruskog sa 165.937 reči;
- *FrWaC korpus* veličine 1.6 biliona reči na francuskom jeziku je korišćen u kombinaciji sa dole navedenim korpusima kako bi se ispitala međujezička podudarnost performansi;
- *Humanistički potkorpus* od 3.431.531 reči iz *Scientext korpusa* je, takođe, korišćen za prikupljanje podataka o međujezičkoj podudarnosti performansi. Dodat je istraživanju zato što je sličniji ICLE korpusu pa samim time i pogodniji za poređenje. Ovaj potkorpus sastoji se od objavljenih naučnih radova i studentskih završnih radova iz oblasti lingvistike, psihologije, obrazovanja i procesuiranja prirodnog jezika;
- *The Corpus de Dissertations Francaises* (CODIF) je korpus koji je sačinjen od studentskih eseja pisanih na francuskom na slične teme, kao i ICLE-FR, a veličine je 92.832 reči.

Za izdvajanje ULS korišćena je minimalna frekvencija od pet ponavljanja u čitavom korpusu od nešto više od milion reči. Neobična u ovom tipu istraživanja je činjenica da autorka nije postavila

minimalnu disperziju pri identifikaciji ULS. Štaviše, iako ne postoje striktna pravila o tome kolika bi minimalna frekvencija za izdvajanje ULS trebalo da bude, pet ponavljanja u korpusu od oko milion reči je dosta liberalan kriterijum, pogotovo ako uzmemo u obzir da se izdvajaju ULS od tri reči koji su brojniji od ULS od četiri reči koji su češće predmet ovog tipa istraživanja. Izdvojeni ULS su manuelno prečišćeni kako bi se izdvojili samo ULS sa leksičkim glagolom. U ICLE-FR izdvojeno je ukupno 273 ULS sa leksičkim glagolom.

Kako bi se prikupili dokazi za unutargrupnu homogenost i međugrupnu heterogenost, Paket je primenila statističke ANOVA testove na ICLE korpus i izdvojila 87 ULS, koji pokazuju značajne razlike u upotrebi među svih deset potkorpusa različitih govornika. Od tih ULS izdvojeno je 34 koji su pokazali značajnu razliku u upotrebi kod izvornih govornika francuskog uporedno sa barem pola drugih grupa neizvornih govornika iz korpusa ICLE upotrebom Dunetovog testa (engl. *Dunnnett's test*). Međutim, za prikupljanje dokaza o međujezičkoj podudarnosti performansi neophodna je kvalitativna analiza podataka koju je autorka sproveda na način koji u nastavku opisujemo. Pošto se teme eseja nisu podudarale u svim potkorpusima i pošto je jedna tema bila posebno prisutna u francuskom potkorpusu a jedva prisutna u drugim, prvo su izdvojeni ULS koji su direktno povezani sa tom temom. Nakon ovog koraka, ostalo je 20 ULS koji su specifični za produkciju studenata kojima je francuski jezik maternji. Paket je identifikovala prevodne ekvivalente za izdvojene ULS i zaključila da se većina ispitanih ULS poklapa po formi sa prevodnim ekvivalentima, a da je podudarnost po značenju ili funkciji uočena za četiri primera.

Rezultati kombinovanja sva tri tipa dokaza ukazuju na postojanje čvrstih dokaza da je korišćenje 20 ULS sa leksičkim glagolom specifičnih za izvorne govornike francuskog posledica transfera. Drugim rečima, 58,8% svih ULS sa leksičkim glagolom koje su koristili samo izvorni govornici francuskog su posledica uticaja transfera. Kvalitativna analiza 20 izdvojenih ULS koji su specifični za govornike francuskog jezika ukazala je na četiri tipa uticaja transfera: 1. *transfer kolokacionih preferenci*; 2. *transfer sintaksičkih konstrukcija*; 3. *transfer funkcija i konvencija u diskursu*; 4. *transfer frekvencija*.

Istraživanje koje je sproveda Paket prevashodno je značajno kao prvo istraživanje koje je imalo cilj da ispita uticaj transfera na upotrebu ULS koristeći metodološki pristup koji je predložio Džarvis. Istraživanje je pokazalo jasno prisustvo međujezičkog uticaja. Za detaljnije poređenje rezultata sa istraživanjem koje će se sprovesti u ovoj disertaciji nije pogodno zato što se bavi ULS od tri reči, a ova disertacija će se fokusirati na ULS od četiri reči. Štaviše, i sama autorka navodi da je jedno od ograničenja istraživanja činjenica da se analiziraju samo ULS sa leksičkim glagolom i da bi buduća istraživanja trebalo da se osvrnu na širu upotrebu ULS.

Gungor i Uisal (2020)

Istraživanje koje su sprovedi Gungor i Uisal (Güngör & Uysal, 2020) bavi se proučavanjem uticaja transfera na upotrebu ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik turski. Navedeno istraživanje je veoma značajno za ovu disertaciju jer su korišćeni korpusi sačinjeni od naučnih radova iz oblasti obrazovanja te je izrazito pogodno za poređenje dobijenih rezultata. Tri korpusa od po milion reči je korišćeno u istraživanju:

- *MCRA-L1 korpus* naučnih radova iz oblasti obrazovanja koje su pisali izvorni govornici engleskog jezika sačinjen od 165 naučnih članaka;
- *MCRA-L2 korpus* naučnih radova iz oblasti obrazovanja na engleskom jeziku koje su pisali izvorni govornici turskog jezika sačinjen od 206 radova;
- *MCRA-T1 korpus* naučnih radova iz oblasti obrazovanja na turskom jeziku koje su pisali izvorni govornici turskog jezika sačinjen od 228 radova.

Kako bi se istražio transfer, u radu je primenjen pristup zasnovan na detekciji koji je predložio Džarvis (Jarvis, 2000, 2012; Jarvis i Pavlenko, 2008). Kako bi se utvrdilo postojanje transfera odnosno međujezičkog uticaja, poželjno je zadovoljiti tri vrste dokaza: unutargrupna homogenost, međugrupna heterogenost i međujezička podudarnost performansi. U istraživanju se sve tri vrste dokaza zadovoljavaju na sledeći način:

1. Samo izdvajanje ULS obezbeđuje prvu vrstu dokaza odnosno unutargrupnu homogenost time što se primenjuje minimalna frekvencija za izdvajanje ULS od 20 ponavljanja u milion reči i minimalna disperzija u 10% tekstova. Drugim rečima, primenom datih parametara izdvajaju se ULS koji su karakteristični za grupu govornika koji se ispituju, a osigurava se da nisu u pitanju karakteristike jednog ili malog broja pojedinaca primenom minimalne disperzije u različitim tekstovima.
2. Kako bi se proverila međugrupna heterogenost, neophodno je da se uporede rezultati različitih grupa neizvornih govornika. Kao što se može videti iz opisa korpusa koji su korišćeni u istraživanju, taj uslov nije ispunjen u samoj postavci istraživanja, međutim, istraživači su rešili problem tako što su odlučili da rezultate svog istraživanja uporede sa rezultatima sličnih prethodnih istraživanja kako bi zadovoljili ovaj kriterijum. Istraživanje koje je sproveda Perez-Ljantada (Pérez-Llantada, 2014) odabrano je za poređenje jer je, kako autori navode, slična veličina korpusa (korpus naučnih radova neizvornih govornika u istraživanju Perez Ljantada je veličine 1,7 miliona reči) upotrebljena, kao i isti kriterijumi za izdvajanje ULS, odnosno minimalna frekvencija od 20 ponavljanja u milion reči i minimalna disperzija u 10% tekstova.
3. Kako bi se obezbedili dokazi iz kategorije međujezičke podudarnosti performansi, autori su identifikovali prevodne ekvivalente ULS specifičnih za izvorne govornike turskog i uporedili frekvencije njihove upotrebe u korpusima MCRA-L2 i MCRA-T1. Autori smatraju da postoji velika verovatnoća da je u pitanju transfer odnosno međujezički uticaj, a ukoliko identifikovani prevodni ekvivalent ima visoku frekvenciju u korpusu naučnih radova na turskom jeziku.

Gorenavedena postavka istraživanja se vodi pretpostavkom da će nam, kada koristimo strani jezik, mentalno biti dostupnije reči i fraze koje su prevodni ekvivalenti reči i fraza na maternjem jeziku (Güngör & Uysal, 2020).

Rezultati istraživanja koji su sprovedli Gungor i Uisal (2020) su sledeći: primenom minimalne frekvencije od 20 ponavljanja u milion reči i minimalne disperzije u 10% tekstova identifikovano je 79 tipova ULS u MCRA-L1 i 119 tipova ULS u MCRA-L2. Daljom analizom podataka autori su uporedili spisak izdvojenih ULS u korpusima izvornih i neizvornih govornika i izdvojili ukupno 34 ULS koje dele oba spiska odnosno 85 ULS koji su koristili samo izvorni govornici turskog jezika. Strukturalna i funkcionalna analiza zajedničkih ULS pokazala je veliku sličnost u raspodeli sa trendovima koji su prisutni u upotrebi ULS kod izvornih govornika. Naime, najzastupljenija strukturalna kategorija su *predložke fraze*, dok je najzastupljenija funkcionalna kategorija *ULS orijentisanih ka istraživanju*. Istraživači su sledeće uporedili spisak ULS korišćenih samo u korpusu izvornih govornika turskog sa ULS koji su korišćeni samo u korpusu izvornih govornika španskog iz istraživanja koje je sproveda Perez-Ljantada (2014). ULS koji su se našli na oba spiska su uklonjeni te je ostalo 54 ULS koji su specifični za izvorne govornike turskog jezika i koji su se koristili u daljoj analizi podataka.

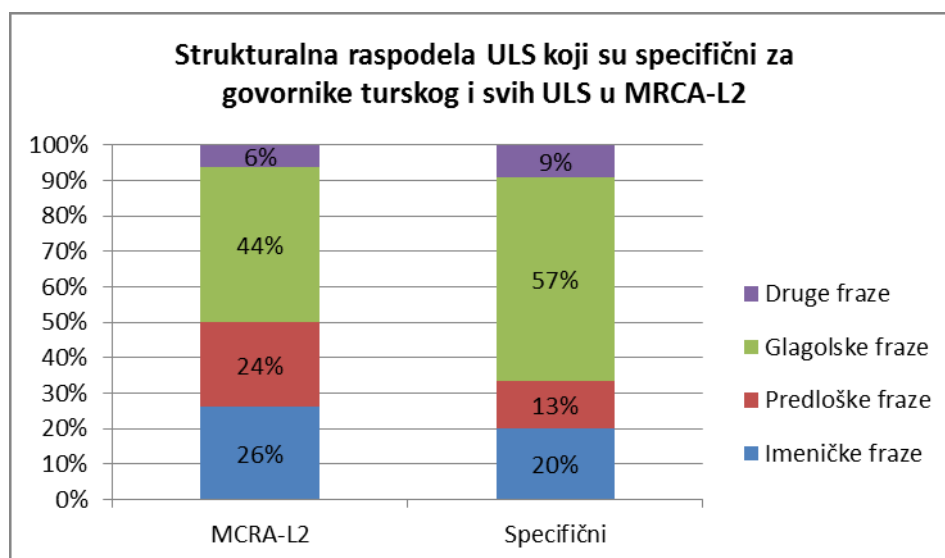
Odluka da se dokazi o međugrupnoj heterogenosti dobijaju poređenjem specifičnih ULS za dve različite grupe neizvornih govornika rezultiralo je u pojavljivanju dva nedostatka pri analizi podataka. Naime, korpus neizvornih govornika engleskog kojima je izvorni jezik španski sačinjen

je od naučnih radova iz 12 različitih naučnih oblasti. Pošto je jedna od glavnih karakteristika ULS da se razlikuju u različitim žanrovima i oblastima, može se pretpostaviti da bi se veći broj zajedničkih ULS izdvojio da su oba poređena korpusa bili sačinjeni od istraživanja iz oblasti obrazovanja. Drugim rečima, dobijeni podaci bi bili precizniji. Navedeno nas dovodi i do drugog problematičnog segmenta, a to je da se veličine korpusa razlikuju, odnosno da je korpus korišćen za izdvajanje ULS izvornih govornika španskog za 41% veći od korpusa koji je korišćen za izdvajanje ULS izvornih govornika turskog.

Identifikovani su prevodni ekvivalenti 54 ULS koji su karakteristični za izvorne govornike turskog jezika, kako bi moglo da se utvrdi da li je njihovo korišćenje uzrokovano uticajem maternjeg jezika. Kako bi identifikovali prevodne ekvivalente, videli da li se pojavljuju u korpusu i da li su adekvatni, autori su zapisali potencijalne prevode i pretražili korpus MCRA-T1.

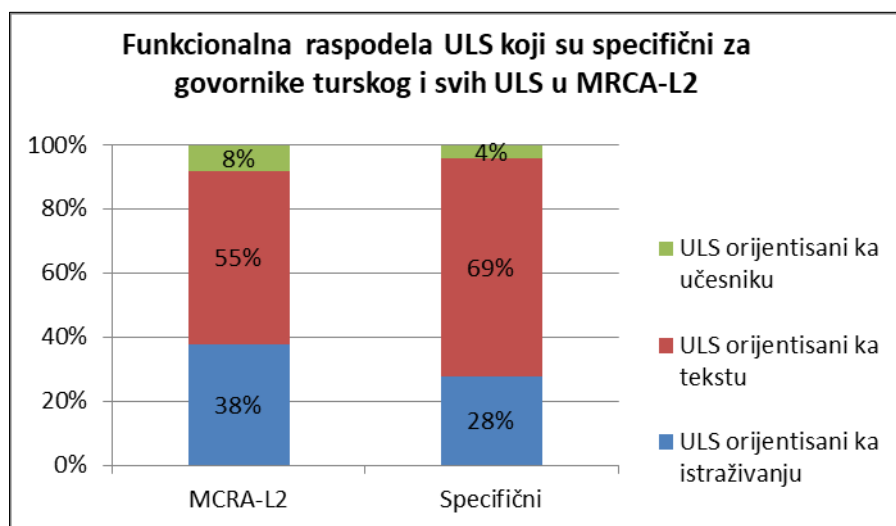
Nakon što su autori izdvojili 54 ULS specifičnih za izvorne govornike turskog, urađena je njihova strukturalna i funkcionalna analiza (detaljnije v. Grafikon 4). Za strukturalnu analizu korišćena je Bajberova taksonomija (Biber et al., 1999), koja originalno ima 12 kategorija i potkategorija, a kojoj su autori dodali još četiri potkategorije (*druge imeničke fraze, drugi glagolski fragmenti, druge pridevske fraze i drugi pasivni fragmenti*). Za funkcionalnu raspodelu ULS korišćena je taksonomija koju je predložio Hajland (Hyland, 2009), a koju je modifikovala Salazar (Salazar, 2014).

Analiza podataka ukazala je na to da su glagolske fraze najučestalije među ULS koji su specifični za izvorne govornike turskog. One sačinjavaju ukupno 57% ULS iz ove grupe sa 31 tipom i 9.220 tokena. Najzastupljenija podgrupa su glagolske fraze sa uvodnom *it strukturom* sa 11 tipova ULS (20%) i 4.476 tokena (30%). Iz raspodele tipova i tokena, gde je procenat tokena mnogo veći od procenta tipova, može se uvideti tendencija izvornih govornika turskog da često koriste iste fraze sa uvodnom *it strukturom*. Često korišćenje fraza sa uvodnom *it strukturom* je karakteristika učenika koji je koriste kako bi dokazali legitimnost svojih argumenata. Autori pretpostavljaju da su izvorni govornici turskog često koristili ove strukture zato što su se osećali manje samouvereno pri pisanju na stranom jeziku i pronalaženju svog mesta na globalnoj akademskoj sceni.



Grafikon 4. *Strukturalna raspodela ULS koji su specifični za govornike turskog jezika u svim ULS u MRCA-L2*

Kada se pogleda funkcionalna raspodela ULS (v. Grafikon 5), vidi se da su kod ULS specifičnih za govornike turskog jezika najzastupljeniji ULS orijentisani ka tekstu 68,53% praćeni ULS orijentisanim ka istraživanju, što predstavlja veći broj nego što je u ULS koji su izdvojeni iz korpusa MCRA-L2. Na osnovu ovih podataka može se zaključiti da izvorni govornici turskog imaju veću tendenciju da koriste *ULS orijentisane ka tekstu*, posebno kada uzmemo u obzir da u korpusu izvornih govornika MCRA-L1 *ULS orijentisani ka tekstu* sačinjavaju samo 37% svih ULS, te da su *ULS orijentisani ka istraživanju* najzastupljeniji i sačinjavaju 56% korpusa (Güngör, 2016).



Grafikon 5. Funkcionalna raspodela ULS koji su specifični za govornike turskog i svih ULS u MRCA-L2

Pri analizi podataka u svom istraživanju Gungor i Uisal (2020) su odlučili da identifikuju koje su funkcionalne podgrupe bile najzastupljenije i da detaljnije analiziraju upotrebljene ULS kroz tu prizmu, dok u isto vreme porede koliko su prevodni ekvivalenti datih ULS zastupljeni u korpusu naučnih radova napisanih na turskom jeziku. Rezultati su pokazali da je najzastupljenija funkcionalna podgrupa inferencijalni ULS sa 15 tipova koji sačinjavaju 28% svih ULS karakterističnih za izvorne govornike turskog jezika. Analiza prevodnih ekvivalenata ULS u podgrupi inferencijalnih ULS pokazala je da se 11 prevodnih ekvivalenata značajno češće koriste u MCRA-T1, dok se preostalih 4 koriste sa sličnom učestalošću kao u MCRA-L2. Takvi rezultati ukazuju na to da je kod ovih primera zastupljen međujezički uticaj, odnosno da je došlo do uticaja transfera.

Druga najčešće zastupljena strukturalna potkategorija je proceduralni ULS, u kojoj se nalaze 12 tipova ULS. Analiza prevodnih ekvivalenata u korpusu radova napisanih na turskom jeziku pokazala je da se deset od 12 tipova u ovoj kategoriji značajnije češće koriste u tom korpusu te da je u tih deset slučajeva najverovatnije došlo do transfera. Učestalo korišćenje proceduralnih ULS može da ukazuje na osvešćenost autora da bi trebalo objektivno da prenose rezultate istraživačkog procesa (Salazar, 2014). Međutim, učestala upotreba te funkcionalne podgrupe takođe je karakteristična za učenike odnosno studente master i doktorskih studija i ogleda njihove pokušaje da dokažu svoje veštine u upotrebi istraživačkih metoda (Hyland, 2008b). Gungor i Uisal tumače navedenu tendenciju kao pokušaj neizvornih govornika da kompenzuju svoj nedostatak što pišu na stranom jeziku tako što će se oslanjati na fiksne leksičke spojeve da bi ostvarili kredibilitet u metodičkom delu istraživanja. Treća najzastupljenija funkcionalna potkategorija je *komparativni ULS*. Osam ULS je svrstano u ovu kategoriju. Analizom prevodnih ekvivalenata identifikovano je da se sedam ULS iz ove kategorije koriste učestalije od strane izvornih govornika turskog zbog međujezičkog uticaja. Četvrta najzastupljenija funkcionalna potkategorija je *strukturalni ULS*, u

kojoj je identifikovano pet tipova ULS. Analizom prevodnih ekvivalenata utvrđeno je da izvorni govornici turskog koriste tri tipa, u datoj kategoriji, češće zbog međujezičkog uticaja. Peta najzastupljenija funkcionalna potkategorija su ograničavajući signali koji su zastupljeni sa četiri ULS. Analiza njihovih prevodnih ekvivalenata pokazala je da postoji međujezička podudarnost performansi, odnosno da je velika verovatnoća da izvorni govornici turskog koriste te ULS zbog uticaja njihovog maternjeg jezika.

Učestali leksički spojevi koji su specifični za izvorne govornike turskog (54 ULS) sačinjavaju 45% svih ULS koji su izdvojeni u korpusu izvornih govornika turskog jezika. Štaviše, od tih 54 ULS, 47 je rezultat uticaja transfera. Ovakvi rezultati ukazuju na značajnu devijaciju od normi izvornih govornika i nedostatak znanja o upotrebi ULS, što dalje može da utiče na osećaj da tekst nisu napisali izvorni govornici (Li & Schmitt, 2009, prema Güngör & Uysal, 2020), nedostatak fluentnosti u pisanju (Hyland, 2008a) i nerazumevanju (Miller, 2009, prema Güngör & Uysal, 2020). Istraživanje koje su sproveli Gungor i Uisal (2020) veoma je pogodno za poređenjem sa rezultatima istraživanja koje se sprovelo u ovoj disertaciji zato što se obrađuje isti žanr (naučni radovi) i ista oblast (obrazovanje).

Ozturk i Tasci (2023)

Možemo da kažemo da se na određeni način istraživanje koje su sproveli Ozturk i Tasci (2023) nadovezuje na istraživanje koje su sproveli Gungor i Uisal (2020). Naime, ono se takođe bavi ispitivanjem uticaja transfera na produkciju na engleskom kod izvornih govornika turskog, isti se koristi pristup zasnovan na detekciji koji je predložio Džarvis, ali se ispituje upotreba ULS u master radovima i doktorskim disertacijama i to iz malo uže oblasti od obrazovanja a to je podučavanje jezika. Istraživanje je imalo cilja da odgovori na sledeća pitanja:

1. *Koje ULS na engleskom izvorni govornici turskog statistički značajno prekomerno i nedovoljno koriste u svojim završnim radovima (master i doktorske studije) uporedno sa studentima kojima je engleski maternji jezik?*
2. *Koliko se ULS na engleskom jeziku u produkciji studenata kojima je maternji jezik turski može prepisati uticaju transfera?*

Kako bi se došlo do odgovora na postavljena istraživačka pitanja, korišćeno je tri korpusa:

- *TMPE* korpus master radova i doktorskih disertacija napisanih na engleskom od strane izvornih govornika turskog koji se sastoji od 1.346.396 reči;
- *TMPT* korpus master radova i doktorskih disertacija napisanih na turskom od strane izvornih govornika turskog koji se sastoji od 1.239.392 reči;
- *AMPE* korpus master radova i doktorskih disertacija napisanih na engleskom od strane izvornih govornika engleskog koji se sastoji od 981.928 reči.

Razlike u veličini korpusa su nastale zato što su autori ujednačili korpuse po broju tekstova. Naime, svaki od korpusa se sastoji od 30 master radova i 20 doktorskih disertacija. Kao što je to slučaj i u mnogim drugim istraživanjima koja su se bavila ULS (Güngör, 2019; Pan et al., 2016; Esfandiari & Barbary, 2017; Al Fajri et al., 2020; etc.) došlo je do razlike u dužini radova kod različitih grupa izvornih govornika. Generalno se u istraživanjima ULS preporučuje da se korpusi ujednače po veličini zbog toga što, kada se porede korpusi različitih veličina, može da se pojave neujednačenosti u dobijenim podacima jer u manjim korpusima može da dođe do većeg broja identifikovanih ULS (Cortes, 2004).

Kako bi se identifikovali ULS od četiri reči, koji su predmet datog istraživanja, primenjena je minimalna frekvencija od 25 ponavljanja u milion reči i minimalna disperzija u pet tekstova. Primenom ovih kriterijuma autorka je identifikovala 125 ULS u korpusu TMPE i 69 ULS u korpusu AMPE. U korpusu na turskom jeziku identifikovano je 32 ULS, međutim, ispostavilo se da su svi povezani sa kontekstom te su uklonjene iz analize. Autorka je zbog toga izdvojila ULS od tri reči, izraze od dve reči i česte glagole, pa njih uporedila sa ULS koji su izdvojeni u korpusu izvornih govornika. Specifično za ovaj rad je činjenica da je autorka koristila funkciju za ključne reči (*Keyword function*) u programu za analizu korpusa *WordSmith Tools 6* kako bi došla do spiska reči koje su korišćene u korpusu neizvornih govornika i koje se nisu pojavljivale u korpusu izvornih govornika.

Korpusi izvornih i neizvornih govornika na engleskom dele 39 ULS, što čini 31% ULS identifikovanih u korpusu izvornih govornika i 56% ULS u korpusu izvornih govornika turskog. Izvorni govornici turskog u korpusu na engleskom pokazali su učestaliju upotrebu istih ULS time što je broj tokena u njihovom korpusu znatno veći. Takvi rezultati ukazuju na to da neizvorni govornici ispoljavaju tendenciju da se prekomerno oslanjaju na često upotrebljivane lingvističke oblike u akademskom pisanju (Durrant & Schmitt, 2009, prema Ozturk i Tasci, 2023). Neizvorni govornici engleskog kojima je maternji jezik turski prekomerno su koristili 62 ULS od kojih se 37 ULS pronašlo samo u produkciji neizvornih govornika a nisu postojali u produkciji izvornih govornika. Jedan ULS je nedovoljno korišćen: *it is important to*. Autori su proverili zastupljenost ULS koji su identifikovani samo u korpusu neizvornih govornika u akademskom delu korpusa COCA (*Corpus of Contemporary American English*), a koji se sastoji od 120 miliona reči. Ustanovili su da su ULS karakteristični za neizvorne govornike turskog prisutni u ovom korpusu, ali su statističke analize pokazale da nisu toliko zastupljene u akademskom žanru. Od tih ULS koji su karakteristični za izvorne govornike turskog 31 ULS ima prevodni ekvivalent u korpusu na turskom (TMPT) i oni sačinjavaju 24,8% svih ULS u korpusu neizvornih govornika. Drugim rečima, 24,8% ULS koje koriste neizvorni govornici turskog jezika jesu posledica uticaja transfera.

Svoje rezultate autori su uporedili sa rezultatima dobijenim u istraživanju koje su sproveli Gungor i Uisal (2020) te se pokazalo da se 12 ULS (od 27) koji su specifični za neizvorne govornike turskog poklapaju u oba istraživanja. To jeste poprilično velika podudarnost podataka kada se uzme u obzir da, iako su veoma slična istraživanja, ona imaju razlike u tipu teksta (naučni radovi profesionalaca uporedno sa master i doktorskim disertacijama studenata) i oblasti (obrazovanje kao šira oblast i podučavanje jezika). Kao što je već poznato, ovakva prekomerna i/ili nedovoljna upotreba ULS kod neizvornih govornika može da utiče na nedovoljnu fluentnost teksta. Stoga, autori podvlače značaj podučavanja ULS u nastavi akademskog engleskog jezika i/ili engleskog za posebne namene sa posebnim akcentom na ULS koji se nalaze u produkciji izvornih govornika a koji se ne pojavljuju u korpusima neizvornih govornika.

Istraživanje koje su sproveli Ozturk i Tasci (2023) korisno je za dublje razumevanje lingvističkog fenomena ULS i njihovog odnosa sa transferom zato što se nadovezuje na prethodna istraživanja u oblasti i dopunjuje sliku sa novim informacijama. Rezultati istraživanja nam pomažu da bolje razumemo koje su karakteristike neizvornih govornika na različitim nivoima znanja (studenti uporedno sa profesionalcima u akademskoj oblasti) i u pomalo drugačijim tipovima tekstova (završni radovi i naučni radovi) kod neizvornih govornika koji imaju isti maternji jezik. Štaviše, oblast istraživanja koja se obrađuje u korpusima u istraživanju Ozturk i Tasci (2023) predstavlja jednu od podoblasti koja je uzeta u obzir pri formiranju korpusa za istraživanje koje je sprovedeno u ovoj disertaciji, te poređenje rezultata jeste korisno.

Istraživanje koje je sprovedla Maria Kostromitina (Kostromitina, 2022) bavi se uticajem transfera na upotrebu ULS kod neizvornih govornika engleskog kojima je maternji jezik ruski u naučnim radovima iz oblasti obrazovne psihologije. Uporedno sa istraživanjem koje su sprovedli Gungor i Uisal (2020) i istraživanjem koje se sprovedlo u ovoj disertaciji u kojima sa ispituje oblast obrazovanja, naučna oblast obrazovne psihologije se može smatrati užim segmentom oblasti obrazovanja. Štaviše, ista podoblast je uzeta u obzir pri formiranju korpusa za istraživanje u ovoj disertaciji, samo je imenovana kao *pedagoška psihologija* a ne *obrazovna psihologija*.

Kostromitina (2022) je u svom istraživanju postavila dva istraživačka pitanja:

1. *U kojoj meri utiče maternji jezik na produkciju ULS u naučnim radovima iz oblasti obrazovne psihologije kod neizvornih govornika engleskog kojima je maternji jezik ruski?*
2. *Da li postoje razlike u funkcionalnoj raspodeli ULS korišćenih u naučnim radovima iz oblasti obrazovne psihologije pisanih na engleskom od strane izvornih govornika engleskog, izvornih govornika ruskog, kao i u radovima pisanim na ruskom od strane izvornih govornika ruskog?*

Kako bi došla do odgovora na gorepomenuta pitanja, autorka je oformila tri korpusa:

- *PSY-ENG1* je korpus naučnih radova na engleskom iz obrazovne psihologije koji su napisali izvorni govornici engleskog jezika. Sastoji se od 61 rada i 410.558 reči;
- *PSY-ENG2* je korpus naučnih radova na engleskom iz obrazovne psihologije koji su napisali neizvorni govornici engleskog jezika kojima je maternji jezik ruski. Sastoji se od 85 rada i 411.637 reči;
- *PSY-RUS1* je korpus naučnih radova na ruskom iz obrazovne psihologije koji su napisali neizvorni govornici ruskog jezika. Sastoji se od 91 rada i 411.787 reči.

Da bi se autori radova koji sačinjavaju korpus identifikovali kao izvorni govornici određenog jezika, morali su da ispune dva kriterijuma: 1. da rade u instituciji u državi u kojoj se taj jezik govori kao zvanični i 2. da imaju ime i prezime tipično za stanovnike tih država.

Autorka je navela da je za identifikaciju transfera koristila pristup zasnovan na detekciji koji je predložio Džarvis (2000). Za pribavljanje dokaza o unutargrupnoj homogenosti navela je primenu minimalne frekvencije za izdvajanje ULS i minimalnu disperziju. Dokazi za međugrupnu heterogenost su u ovom istraživanju problematični zato što autorka nema korpus druge grupe neizvornih govornika sa kojom bi mogla da poredi spiskove izdvojenih ULS kako bi zaključila koji su specifični za izvorne govornike ruskog. Autorka jeste poredila svoje rezultate sa rezultatima iz istraživanja Čen i Bejker (Chen & Baker, 2010) i Adel i Erman (Adel & Erman, 2012), međutim, to je urađeno samo za ULS koji su zajednički za sva tri korpusa a nije urađeno za ULS koji su specifični za govornike ruskog jezika. Dokaze za međugrupnu podudarnost performansi autorka je prikupila tako što je izdvojene ULS iz ruskog korpusa prvo prevela na engleski pa je onda poredila njihovu upotrebu u korpusu neizvornih govornika na engleskom i takođe poredila funkcionalnu raspodelu za ULS u sva tri korpusa.

Kostromitina (2022) je ispitala upotrebu ULS od četiri reči a za njihovu identifikaciju koristila je minimalnu frekvenciju od 40 ponavljanja u milion reči i minimalnu disperziju od 10% za korpus na engleskom jeziku. Za korpus na ruskom jeziku korišćeni su isti parametri ali su se izdvojili ULS od tri reči, pošto je autorka uvidela da se neki ULS u engleskom mogu prevesti na ruski sa samo jednom rečju zbog drugačijih karakteristika morfološkog sistema. Na ovaj način izdvojeno je: 82

ULS u PSY-ENG1, 223 ULS u PSY-ENG2 i 264 ULS u PSY-RUS1. Izdvojeni ULS iz korpusa na ruskom su prevedeni na engleski kako bi dalji koraci u istraživanju i poređenju po postavkama istraživanja bili mogući.

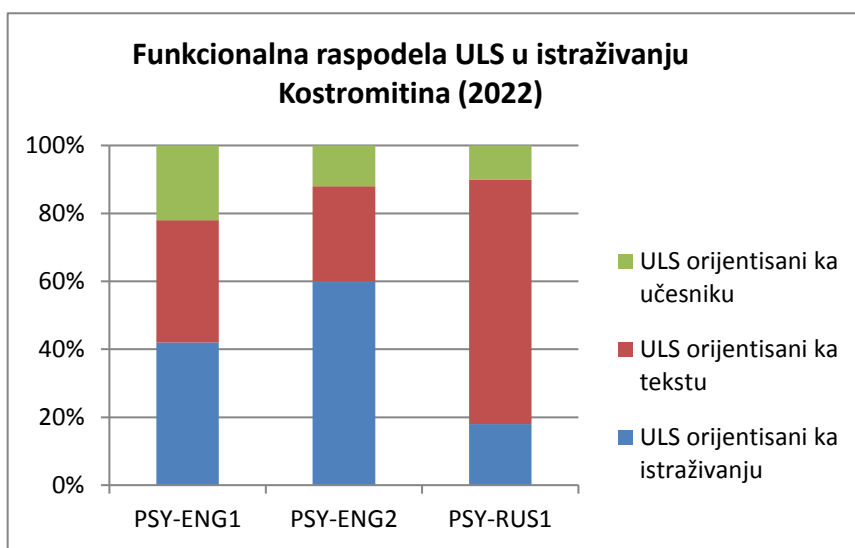
Identifikovano je šest ULS koji su zajednički za sva tri korpusa (*it is important to, as well as the, at the end of, in the case of, is one of the, at the same time*); oni predstavljaju 7,3% ULS izdvojenih u korpusu izvornih govornika, a 3% ULS izdvojenim u engleskom i ruskom korpusu izvornih govornika ruskog. Dodatna analiza ovih ULS pokazala je da većina pripada funkcionalnoj kategoriji *ULS orijentisanih ka tekstu*, dva pripadaju kategoriji *ULS orijentisanih ka istraživanju* i jedan je *ULS stava*. Kada je autorka uporedila frekvenciju ovih ULS u tri korpusa, uvidela je da se njihova zastupljenost umnogome razlikuje u tri korpusa. Na primer, ULS *at the same time* je najčešći u korpusu neizvornih govornika engleskog sa 303 ponavljanja, dok je u korpusu izvornih govornika bio na granici za izdvajanje kao ULS sa samo 20 ponavljanja. Sa druge strane, ULS *it is important to* se ponavlja 118 puta u korpusu izvornih govornika, dok se u korpusu na ruskom jeziku ponavlja samo 32 puta. Autorka je uporedila ovih šest ULS sa spiskovima ULS iz prethodnih istraživanja (Chen & Baker, 2010; Adel & Erman, 2012) i ustanovila da se pet od šest nalaze u tim istraživanjima, što je dalje navodi na zaključak da njihova prisutnost u korpusu neizvornih govornika ruskog nije posledica transfera. Navedeno je svakako i logično s obzirom na činjenicu da su prisutni i u korpusu izvornih govornika.

Ukupno 17 ULS se nalaze u oba korpusa na engleskom jeziku. Kada se spoje sa šest ULS koji su prisutni u sva tri korpusa, oni sačinjavaju 28% svih ULS u korpusu izvornih govornika i 10% svih ULS u korpusu neizvornih govornika engleskog. Statistički testovi (*log-likelihood*) su pokazali da su izvorni govornici ruskog preterano koristili deset ULS, a da su samo jedan (*it is important to*) nedovoljno upotrebljavali. Interesantno je da je u istraživanju Ozturk i Tasci (2023) isto identifikovan samo jedan ULS koji je nedovoljno korišćen od strane neizvornih govornika, štaviše, isti ULS je u pitanju – *it is important to*, što bi moglo da ukaže na mogućnost da je nedovoljna upotreba ovog konkretnog ULS univerzalna karakteristika neizvornih govornika. Naravno, potrebno je još istraživanja kako bi se potvrdila ili opovrgla ova pretpostavka.

Ukupno 22 ULS je zajedničko za korpus izvornih govornika ruskog na engleskom i ruskom, što, kada se spoji sa ULS koji su zajednički za sva tri korpusa, čini 12% ukupnog broja ULS u korpusu na engleskom i 10% ULS u korpusu na ruskom. Određen broj ovih zajedničkih ULS (npr. *in the present study, we can say that, an important role in*) ne pokazuju značajne razlike u upotrebi u ova dva korpusa, što snažno ukazuje na postojanje transfera, međutim, i kod onih koji se značajno razlikuju postoje šansa da je uticaj transfera prisutan s obzirom na činjenicu da ovi ULS nisu prisutni u korpusu izvornih govornika engleskog.

Što se tiče funkcionalne raspodele ULS, Kostromitina je uradila ovu analizu za samo 50 najfrekventnijih ULS u tri korpusa. Rezultati kategorizacije su pokazali da je raspodela slična u dva korpusa na engleskom, naime, u oba su *ULS orijentisani ka istraživanju* najzastupljeniji (PSY-ENG1 42%, PSY-ENG2 60%) i praćeni *ULS orijentisanim ka tekstu* (PSY-ENG1 36%, PSY-ENG2 28%). U korpusu na ruskom jeziku najzastupljenija funkcionalna kategorija su *ULS orijentisani ka tekstu* koji čine čak 72% korpusa, praćeni su *ULS orijentisanim ka istraživanju*, koji čine 18% korpusa. Interesantno je pitanje zbog čega su *ULS orijentisani ka istraživanju* najprisutniji u korpusu neizvornih govornika engleskog. Da je u pitanju standardan uticaj transfera, očekivalo bi se da bi neizvorni govornici koristili manje ovu kategoriju zato što je u ruskom mnogo manje zastupljena. Pošto su ULS leksički oblik kojeg govornici uglavnom nisu svesni osim ako nisu imali priliku da ga kao takvog uče (što još uvek nije široko primenjeno u nastavi i udžbenicima), teško da je došlo da grešaka prekomerne produkcije zbog uticaja transfera. Međutim, statistička analiza podataka pokazala je da razlike u funkcionalnoj raspodeli glavnih kategorija između dva korpusa

na engleskom nisu statistički značajne. Pri sagledavanju funkcionalnih potkategorija ukazala se kao interesantna činjenica da su tranzicioni signali najzastupljenija potkategorija u korpusu na ruskom jeziku koja sačinjava 72% svih ULS u toj kategoriji, ali i da je ista potkategorija najzastupljenija u korpusu na engleskom kod izvornih govornika ruskog sa zastupljenošću od 42% (v. Grafikon 6).



Grafikon 6. *Funkcionalna raspodela ULS u istraživanju Kostromitina (2022)*

Kritička analiza i poređenje opisanih relevantnih istraživanja

U poglavlju u kojem smo analizirali uticaj transfera na upotrebu ULS kod neizvornih govornika engleskog jezika sagledali smo četiri relevantna istraživanja (Paket, 2013; Güngör & Uysal, 2020; Öztürk & Taşçı, 2023; Kostromitina, 2022). Istraživanje koje je sprovedla Paket (2013) je posebno značajno zato što je prvo istraživanje u ovoj oblasti koje je primenilo Džarvisov (2000) pristup zasnovan na detekciji i na sistematičan način proučila uticaj transfera u produkciji ULS kod neizvornih govornika. Međutim, istraživanje je najmanje pogodno za poređenje konkretnih rezultata sa rezultatima koji će se dobiti u ovoj disertaciji zato što se Paket fokusira na ULS od tri reči i to samo na one sa leksičkim glagolom, što dodatno sužava polje istraživanja. Ostala tri istraživanja su veoma pogodna za poređenje konkretnih istraživanja zato što se sva tri fokusiraju na ULS od četiri reči u sličnim ili istim naučnim oblastima. Naime, istraživanje koje su sprovedli Gungor i Uisal (2020) istražuje istu naučnu oblast obrazovanja, dok istraživanja koja su sprovedli Öztürk i Taşçı (2023) i Kostromitina (2022) ispituju naučne oblasti koje su malo uže (podučavanje jezika i obrazovna psihologija) ali isto u domenu obrazovanja. Štaviše, u korpusima koji su kreirani za potrebe istraživanja sprovedenog u ovoj disertaciji, obe uže oblasti su našle svoje mesto kao podoblasti koje sačinjavaju širi pojam obrazovanja.

Interesantno je da, iako skoro sva navedena istraživanja prate Džarvisov (2000) pristup zasnovan na detekciji³², nijedno istraživanje ne prikuplja dokaze na isti način tj. same postavke istraživanja se razlikuju u nekim elementima. Kako bi jasnije uporedili ove postavke, sagledaćemo kako svako od istraživanja prikuplja tri vrste dokaza poželjnih za dokazivanje postojanja transfera (*unutargrupna homogenost, međugrupna heterogenost i međujezička podudarnost performansi*), v. Tabelu 5. Kao što može da se vidi iz Tabele 5, sva tri istraživanja prikupljaju dokaze za unutargrupnu homogenost primenom minimalne frekvencije i disperzije za izdvajanje ULS iako se

³²U istraživanju Ozturk i Tasci (2023) ne spominje se ovaj pristup eksplicitno, iako se uzimaju u obzir prethodna istraživanja koja ga primenjuju te se stvara uglavnom kompatibilna metodološka postavka za istraživanje pojma i komparaciju.

kriterijumi razlikuju. To se moglo i očekivati, pošto je ova vrsta dokaza uvek prisutna u istraživanjima ULS (kada su pravilno izdvojeni) zbog same prirode fenomena.

Na osnovu navedenih podataka takođe možemo da zaključimo da su dokazi za međugrupnu heterogenost dosta slabi u dva istraživanja koja ih pribavljaju a da u istraživanju Kostromitine (2022) uopšte nisu obezbeđeni. Može se videti da same postavke istraživanja nisu predvidele prikupljanje snažnih dokaza za međujezičku heterogenost, jer bi to zahtevalo formiranje poredivih korpusa neizvornih govornika engleskog koji imaju drugi maternji jezik (Džarvis, 2000; Džarvis i Pavlenko, 2008). Istraživanje koje se sprovodilo u ovoj disertaciji stavlja akcenat baš na prikupljanju odgovarajućih dokaza za međugrupnu heterogenost i samim time popuniti nastali nedostatak u prethodnim istraživanjima koja su se bavila uticajem transfera na upotrebu ULS od četiri reči u različitim oblastima obrazovanja. Iz Tabele 5 se, takođe, može uvideti da su se prethodna istraživanja najviše fokusirala na prikupljanje treće vrste dokaza, odnosno dokaza za međujezičku podudarnost performansi. Međutim, i u ovom koraku se može videti da autori primenjuju različite nivoe strogosti pri prikupljanju ovih dokaza. Naime, u istraživanju Gungor i Uisal (2020) najstroži kriterijumi su primenjeni zato što su autori poredili frekvencije upotrebe prevodnih ekvivalenata u dva korpusa, dok su druga dva istraživanja smatrala dovoljnim dokazom ako prevodni ekvivalenti postoje u korpusu na maternjem jeziku neizvornih govornika engleskog.

Tabela 5. Metode prikupljanja dokaza za postojanje transfera u upotrebi ULS u prethodnim istraživanjima

<i>Istraživanje</i>	<i>Unutargrupna homogenost</i>	<i>Međugrupna heterogenost</i>	<i>Međujezička podudarnost performansi</i>
<i>Gungor i Uisal (2020)</i>	Primena minimalne frekvencije od 20 ponavljanja u milion reči i minimalne disperzije od 10% tekstova;	Poređenje ULS specifičnih za neizvorne govornike sa rezultatima istraživanja Perez-Ljatanda (2014) odnosno spiskom ULS specifičnih za govornike španskog;	Identifikacija prevodnih ekvivalenata ULS specifičnih za neizvorne govornike i poređenje frekvencije njihove upotrebe u korpusima MCRA-L2 i MCRA-T1 po funkcionalnim potkategorijama;
<i>Kostromitina (2022)</i>	Primena minimalne frekvencije od 40 ponavljanja u milion reči i minimalne disperzije od 10% tekstova.	Nisu obezbeđeni dokazi za međugrupnu heterogenost.	Poređenje spiska ULS identifikovanih u korpusima izvornih govornika ruskog na ruskom i engleskom jeziku kako bi se identifikovali ULS koji se pojavljuju u oba korpusa neizvornih govornika, a koji se ne pojavljuju u korpusu izvornih govornika engleskog jezika.
<i>Ozturk i Tasci (2023)</i>	Primena minimalne frekvencije od 25 ponavljanja u milion reči i minimalne disperzije od pet tekstova;	Primena funkcije za ključne reči u programu za analizu korpusa kako bi došla do spiska reči koje su specifične za korpus neizvornih govornika;	Identifikacija prevodnih ekvivalenata ULS specifičnih za neizvorne govornike i provera da li se oni pojavljuju u korpusu na turskom jeziku.

Istraživanja u ovom poglavlju su se takođe razlikovala i u tipu dodatnih informacije koje su smatrane relevantnim i samim time proučene. Dva istraživanja (Güngör & Uysal, 2020; Kostromitina, 2022) pružaju i informacije o funkcionalnoj raspodeli ULS. Međutim, istraživanja porede funkcionalnu raspodelu različitih grupa ULS. Gungor i Uisal (2020) porede funkcionalnu raspodelu ULS koji su izdvojeni u celom korpusu neizvornih govornika engleskog kojima je maternji jezik turski sa ULS koji su specifični za tu grupu govornika, što zapravo i daje bolje razumevanje razlika u upotrebi ULS koji su potencijalno prisutni zbog uticala transfera. Kostromitina (2022) poredi funkcionalnu raspodelu ULS u tri korpusa korišćenih u njenom istraživanju, ali ne daje podatke o funkcionalnoj raspodeli ULS koji su specifični za izvorne govornike ruskog, što je šteta jer bi poređenje takvih podataka doprinelo boljem razumevanju fenomena transfera u upotrebi ULS kod neizvornih govornika engleskog jezika.

Pošto su istraživanja u ovom poglavlju imala cilj da identifikuju uticaj transfera na upotrebu ULS kod neizvornih govornika engleskog, najznačajnije nam je uporediti te rezultate kako bismo videli da li se poklapaju i time bolje razumeli potencijalne karakteristike uticaja transfera na ovu leksičku pojavu. U Tabeli 6 nalazi se pregled gorenavedenih informacija.

Tabela 6. *Pregled uticaja transfera na upotrebu ULS u prethodnim relevantnim istraživanjima*

Istraživanje	<i>ULS u korpusima neizvornih govornika engleskog jezika</i>	<i>ULS specifični za neizvorne govornike engleskog jezika</i>	<i>ULS pod uticajem transfera</i>	<i>Koji procenat čine ULS pod uticajem transfera od svih ULS neizvornih govornika</i>
<i>Gungor i Uisal (2020)</i>	119	54	47	39.5%
<i>Kostromitina (2022)</i>	82	22	22	9.87%
<i>Ozturk i Tasci (2023)</i>	125	37	31	24.80%

Na osnovu rezultata datih u Tabeli 6 možemo videti da je najveći procenat ULS pod uticajem transfera identifikovan u istraživanju koje su sprovedli Gungor i Uisal (2020), iako su oni primenili najstrože kriterijume za identifikaciju transfera. Od tri istraživanja najmanji uticaj transfera uočava se u istraživanju koji je sprovedla Kostromitina (2022) – 9,87% što može da ukazuje na manji uticaj transfera na produkciju ULS kod izvornih govornika ruskog uporedno sa izvornim govornicima turskog, što bi dalje moglo da ukazuje na potencijalne razlike u podložnosti uticaju transfera na upotrebu ULS kod različitih grupa neizvornih govornika engleskog jezika. Potrebno je još istraživanja kako bi se ovakve pretpostavke potvrdile ili opovrgle.

Istraživanja predstavljena u ovom poglavlju daju nam dobru polaznu tačku za dalja istraživanja o uticaju transfera na upotrebu ULS. Ovaj tip istraživanja je neophodan kako bismo bolje razumeli zašto postoje velike razlike u upotrebi ULS kod izvornih i neizvornih govornika. Istraživanja transfera sigurno neće pružiti potpun odgovor na postavljeno pitanje, ali će sigurno predstaviti značajnu *puzlu* neophodnu za sklapanje *kompleksne slagalice koja predstavlja odgovor*.

2. Osvrt na dosadašnja istraživanja učestalih leksičkih spojeva

2.1. Prethodna istraživanja učestalih leksičkih spojeva u oblasti obrazovanja

Podela na „meke” i „tvrde” nauke (engl. *soft and hard sciences*) odnosi se na postojanje paradigme ili „skupinu teorije koju primenjuju svi članovi u oblasti”³³ (Biglan, 1973b: 201, prema Dang, 2018). Obrazovanje se može okarakterisati kao meka nauka pošto je zasnovana na ljudskom iskustvu više nego na apstraktnim konceptima, ali i kao primenjena nauka jer se usredsređuje i na praktičnu primenu (Becher, 1989, prema Candarli & Jones, 2019). Iako je zbog svojih specifičnih karakteristika oblast obrazovanja interesantna, relativno se mali broj istraživanja bavio frazeološkim proučavanjima ULS u ovom domenu. Međutim, problematici se jeste prišlo iz nekoliko različitih uglova, pa se tako analizirala:

- upotreba ULS u različitim obrazovnim registrima, kao što su udžbenici, predavanja i akademska proza (Biber et al., 2004; Lee, 2020; Neely & Cortes, 2009);
- razlika u upotrebi ULS kod izvornih i neizvornih govornika engleskog jezika u naučnim radovima iz oblasti obrazovanja (Güngör & Uysal, 2016; Güngör, 2016; Güngör, 2019);
- paradigmatički uticaji na upotrebu ULS u naučnim radovima iz oblasti obrazovanja (Candarli & Jones, 2019);
- razlika u upotrebi ULS u oblasti obrazovanja kroz vreme (Cui & Kim, 2021; Palinkašević i Prtljaga, 2024).

Bajber i saradnici (Biber et al., 2004) istražuju uticaj različitih registra na upotrebu ULS u domenu obrazovanja. Naime, proučavaju korišćenje ULS u nastavi i udžbenicima, pa dobijene nalaze porede sa akademskom prozom i opštom konverzacijom. Za potrebe istraživanja koristila su se četiri korpusa:

- *T2K-SWAL Corpus Classroom Teaching* za predavanja, koji se sastoji od 1.248.800 reči;
- *T2K-SWAL Corpus Textbooks* za udžbenike, koji se sastoji od 760.600 reči;
- *Longman Spoken and Written Corpus British and American English Conversation* za konverzaciju, koji se sastoji od 5.000.000 reči;
- *Longman Spoken and Written Corpus of Academic Prose* za akademsku prozu, koji se sastoji od 5.300.000 reči.

Istraživači su primenili konzervativniju minimalnu frekvenciju od 40 ponavljanja i minimalnu disperziju u pet tekstova za identifikaciju ULS od četiri reči. U akademskoj prozi identifikovano je 19 tipova ULS, u udžbenicima 27, u predavanjima 84 i u konverzaciji 43. Dobijeni podaci su iznenađujući, jer je početna hipoteza bila da će broj ULS u predavanjima biti između konverzacije i udžbenika budući da spaja ova dva registra, međutim, u korpusu sa predavanja identifikovan je ubedljivo najveći broj tipova. Kada se ovi podaci o tipovima uporede sa podacima o tokenima, uviđa se da je po broju tipova mnogo manja razlika između konverzacije i predavanja. To ukazuje na činjenicu da se za predavanje koristi veći broj različitih ULS dok se u konverzaciji koristi manji broj ULS ali učestalije.

Za strukturalnu raspodelu ULS Bajber i saradnici (Biber et al., 2004) koristili su sledeće tri kategorije, koje nisu postale normativne u daljim istraživanjima: *ULS sa fragmentima zavisne klauze*, *glagolski ULS* i u istoj grupi – *imenički ULS* i *ULS sa predloškim fragmentom*. Rezultati strukturalne analize podataka pokazali su da su *glagolski ULS* najzastupljeniji u govoru, dok se u akademskoj prozi gotovo ni ne pojavljuju. Predavanja koriste sve tri strukturalne grupe poprilično

³³[...] „a body of theory which is subscribed to by all members of the field”

podjednako učestalo, što dalje objašnjava najučestalije pojavljivanje ULS u predavanjima budući da koriste i ULS karakteristične za akademsku prozu i one karakteristične za govor. Strukturalna raspodela ULS u udžbenicima je najbližnja akademskoj prozi, iako su u njima *glagolski ULS* nešto zastupljeniji.

Po funkciji koju vrše ULS u ovom istraživanju podeljeni su u referencijalne spojeve, organizatore diskursa i izraze stava. U konverzaciji su najzastupljeniji izrazi stava praćeni organizatorima diskursa, dok su referencijalni spojevi najmanje zastupljeni; u akademskoj prozi i udžbenicima situacija je obrnuta – najzastupljeniji su referencijalni spojevi dok su mnogo ređe korišćeni organizatori diskursa i izrazi stava. Kao i za strukturalnu raspodelu tako i za funkcionalnu, predavanja imaju karakteristike i konverzacije i akademske proze. Naime, u predavanjima referencijalni izrazi su podjednako zastupljeni kao i izrazi stava, a organizatori diskursa su najmanje zastupljeni. Međutim, iako su organizatori diskursa najmanje zastupljeni od tri funkcionalne kategorije sama razlika u upotrebi kategorija je mala.

Nedostatak istraživanja Bajbera i saradnika (Biber et al., 2004) jeste velika razlika u veličini korpusa iz kojih se izdvajaju ULS. U istraživanjima ULS ujednačenost veličine i sastava korpusa je od velikog značaja, jer postoji tendencija da se u manjim korpusima izdvoji neproporcionalno veliki broj ULS; u njima se izdvajaju ULS koji imaju veštački visoku frekvenciju koja ne može da se održi u većim korpusima (Cortes, 2002, prema Biber et al., 2004).

Kao jedno od prvih istraživanja koje se bavilo upotrebom ULS u različitim obrazovnim registrima – istraživanje Bajbera i saradnika (Biber et al., 2001), uticalo je na razvitak čitavog talasa istraživanja koja su dublje zašla u problematiku. Tako su iznikla istraživanja koja proučavaju upotrebu ULS u različitim vrstama udžbenika. Takvo je, na primer, istraživanje koje je poduzeo Li (Lee, 2020), kojim se ispituje upotreba ULS u lingvističkim udžbenicima na korpusu od 1,14 miliona reči. Glavni cilj istraživanja je bila identifikacija ULS čije bi podučavanje pomoglo studentima lingvistike. Strukturalna i funkcionalna analiza podataka se poklapa sa rezultatima dobijenim u Bajberovom istraživanju (Biber et al., 2004), jer je i u njemu najzastupljenija strukturalna kategorija u udžbenicima imeničke fraze i ULS sa predloškim fragmentom, a najzastupljenija funkcionalna kategorija su isto referencijalni spojevi.

Neka od istraživanja u ovom talasu bave se udžbenicima u specifičnim kontekstima. Primera radi, upotreba ULS ispitivana je u udžbenicima iz hemije i fizike koji se koriste u Pakistanu (Hussain et al., 2021), kao i u udžbenicima koji se koriste u višim razredima osnovne škole u Švedskoj (Ribeck & Borin, 2011). Zašlo se i u detaljniju analizu specifičnih elemenata u specifičnim kontekstima. Tako Čen (Chen, 2018) ispituje upotrebu ULS u lekcijama u kojima se obrađuje vokabular u knjigama iz inženjstva za prvu godinu studija, a Alsmari (Alasmari, 2018) ispituje samo ULS iz funkcionalne grupe *ULS orijentisanih ka učesniku* u udžbenicima iz matematike koje se koriste na završnoj godini studija. Analiza upotrebe ULS u udžbenicima je korišćena i za kritičku analizu podobnosti materijala koji sačinjava udžbenike za engleski jezik za posebne svrhe u domenu električnog inženjstva i to tako što su upoređeni ULS u standardnim udžbenicima iz ovog domena sa onim koji se koriste u udžbenicima za engleski za posebne svrhe u istom domenu (Chen, 2008). Rezultati istraživanja su pokazali da se samo trećina ULS identifikovanih u udžbenicima iz električnog inženjeringa nalazi i u udžbenicima za engleski za posebne svrhe iz istog domena. Štaviše, pri funkcionalnoj analizi ULS ustanovilo se da u udžbenicima engleskog jezika za posebne svrhe (domen električnog inženjeringa) nedostaje pet funkcionalnih potkategorija koje se nalaze u standardnim udžbenicima u tom domenu. Talas istraživanja upotrebe ULS u udžbenicima se proširio i na druge pisane materijale koji se koriste u nastavi, pa se tako ispitala upotreba ULS u stepenovanim lektirama za učenje engleskog jezika (engl. *graded readers*; Allan, 2016), knjigama za samostalno učenje (Allan, 2017), testovima (Beng & Keong, 2017) itd.

Naravno, istraživanja u pomenutom talasu se nisu fokusirala samo na pisane elemente u obrazovanju već su se bavila i proučavanjem upotrebe ULS u govoru tj. na predavanjima. Nili i Kortez (Neely & Cortes, 2009) u svom istraživanju usredsredili su se na upotrebu ULS u akademskom diskursu sa posebnom analizom i komparacijom produkcije nastavnog osoblja i studenata. Istraženo je i da li disciplinske razlike u predavanjima utiču na upotrebu ULS, odnosno koji se ULS koriste u „mekim” i „čvrstim” naukama (engl. *soft and hard sciences*; Kashiha & Chan, 2013). Tema se dalje razrađivala istraživanjem upotrebe ULS u predavanjima iz različitih specifičnih predmeta kao što je npr. matematika (Herbel-Eisenmann & Wagner, 2010), kao i poređenjem produkcije ULS u predavanjima iz različitih oblasti kao što su političke nauke i hemija (Kashiha & Swee Heng, 2013). Sprovedeno je i istraživanje koje ispituje upotrebu ULS kod izvornih i neizvornih profesora engleskog jezika na predavanjima (Kwon & Lee, 2014). Rezultati ukazuju na to da profesori engleskog jezika koji nisu izvorni govornici engleskog jezika koriste manji broj ULS ali frekventnije od profesora koji jesu izvorni govornici, kao i da se snažnije oslanjaju na manji broj funkcionalnih kategorija.

S obzirom na to da se u ovoj disertaciji istražuje upotreba ULS u naučnim radovima u oblasti obrazovanja kod izvornih i neizvornih govornika, posebna će se pažnja posvetiti prethodnim istraživanjima frazeologije u oblasti obrazovanja koja su se vršila na naučnim radovima. Kao što je već gorepomenuto, to su proučavanja koja se fokusiraju na razlike u upotrebi ULS u naučnim radovima iz domena obrazovanja kod izvornih i neizvornih govornika, paradigmatičke razlike u upotrebi ULS u naučnim radovima iz oblasti obrazovanja i razlike u upotrebi ULS u naučnim radovima iz oblasti obrazovanja kroz vreme. Istraživanja koja se bave upotrebom ULS kod izvornih i neizvornih govornika u naučnim radovima iz oblasti obrazovanja (Güngör & Uysal, 2016; Güngör, 2016; Güngör, 2019) detaljnije će se obraditi u poglavlju 2.3. *Istraživanja upotrebe ULS u naučnim radovima kod izvornih i neizvornih govornika.*

2.1.1. Paradigmatičke razlike u upotrebi ULS u naučnim radovima iz oblasti obrazovanja

Istraženo je i da li postoje razlike u upotrebi ULS u naučnim radovima iz oblasti obrazovanja koji primenjuju kvalitativnu ili kvantitativnu analizu podataka, odnosno radovima sa različitim paradigmatičkama (Candarli & Jones, 2019). U prethodnim istraživanjima nije se mnogo uzimala u obzir mogućnost da naučni članci nisu monolitan koncept (Gray 2015, prema Candarli & Jones, 2019), te gorenavedeno istraživanje nastoji da doprinese ovoj oblasti istraživanja tako što će proučavati intradisciplinarnu razlike u korišćenju ULS, kao što će to uraditi i istraživanje sprovedeno u ovoj disertaciji. Za potrebe istraživanja (Candarli & Jones, 2019) formirana su dva korpusa od oko milion reči (korpus sa kvantitativnom metodologijom – 1.001.099 reči sačinjen od 142 naučna rada; korpus sa kvalitativnom metodologijom – 1.001.342 reči sačinjen od 141 naučnog rada). Radovi koji sačinjavaju oba korpusa preuzeti su iz tri naučna časopisa koji se nalaze na *SCI* listi i koji su objavljeni u periodu od 2011. do 2015. godine. Istraživanje se bavilo ULS od četiri reči koji su identifikovani primenom minimalne frekvencije od 20 ponavljanja u milion reči i minimalne disperzije u 5% naučnih radova, odnosno u najmanje sedam naučnih radova. *Bootstrapping* i *bias-corrected and accelerated* (BCa) metode su korišćene pri statističkoj analizi podataka. *Bootstrapping* je „procedura u kojoj se veliki broj uzoraka izdvaja iz posmatranih podataka kako bi se ustanovila distribucija uzorkovanja iz koje se mogu povući inferencijalni podaci”³⁴(Candarli & Jones, 2019: 244). BCa je metoda koja „ispravlja greške predrasuda i iskrivljenosti u distribuciji, i uopšteno je tačnija od drugih metoda”³⁵(Candarli & Jones, 2019: 245).

³⁴[...] „a procedure where many samples are taken from the observed data to estimate a sampling distribution from which inferential data can be conducted”

³⁵„This method corrects for both bias and skewness in the distribution, and is generally more accurate than other methods”.

Rezultati istraživanja (Candarli & Jones, 2019) pokazali su da se veći broj ULS izdvojio u korpusu sa radovima koji primenjuju kvantitativnu analizu podataka (108 ULS) uporedno sa korpusom koji primenjuje kvalitativnu (80 ULS). Za strukturalnu analizu podataka istraživači su upotreбили taksonomiju koju su predložili Bajber i saradnici (Biber et al., 1999). Analiza je pokazala razlike u strukturalnoj raspodeli ULS u dva korpusa. Naime, u korpusu sa radovima u kojima je primenjena kvantitativna analiza najzastupljeniji su *ULS sa fragmentom predloške fraze* praćeni *glagolskim ULS* i *imeničkim ULS* na trećoj poziciji, dok su u korpusu sa radovima u kojima je primenjena kvalitativna analiza najzastupljeniji *ULS sa fragmentom predloške fraze* praćeni *imeničkim* pa *glagolskim ULS*. Nije ustanovljena statistički značajna razlika u upotrebi *ULS sa fragmentom predloške fraze* u dva gorenavedena korpusa, ali jeste ustanovljena za razlike u upotrebi *imeničkih* i *glagolskih ULS*. Kako bi se prezentovali svi metodološki koraci u kvantitativnim istraživanjima, neophodno je često ukazivanje na podatke, rezultate, analitičke procedure, kao i na konkretne i apstraktne koncepte unutar istraživanja, što objašnjava učestalije korišćenje *imeničkih ULS* u korpusu sa kvantitativnim naučnim radovima.

Pri funkcionalnoj raspodeli ULS korišćena je taksonomija koju su predložili Bajber i saradnici (Biber et al., 2004). Ustanovljeno je da su glavne funkcionalne kategorije isto rasprostranjene u oba korpusa. Naime, u oba korpusa najzastupljeniji su referencijalni spojevi praćeni izrazima stava i organizatorima diskursa. Sve tri funkcionalne kategorije su značajno zastupljenije u korpusu sa kvantitativnim naučnim radovima. Najveća se razlika može videti u upotrebi funkcionalne kategorije izraza stava, u kojoj se 92% ULS koristi u funkciji ograđivanja. Autori veruju da se ta razlika može objasniti činjenicom da u kvantitativnim člancima istraživači diskutuju o generalizaciji podataka, što je korak koji se obično ne nalazi u kvalitativnim istraživanjima. Ustanovljene paradigmatičke razlike u upotrebi ULS ukazuju na potrebu da se vodi računa i o toj varijabli pri formiranju korpusa naučnih radova iz društvenih nauka kako bi se dobili najreprezentativniji rezultati.

2.1.2. Razlike u upotrebi ULS u oblasti obrazovanja kroz vreme

Razlikama u upotrebi ULS u oblasti obrazovanja kroz vreme bavila su se dva rada, na koja ćemo se u nastavku i osvrnuti. Radovi su metodološki slični i proučavaju promene u istom vremenskom rasponu, ali se fokusiraju na različite registre, što nam daje jasniju i širu sliku o promenama.

Cui i Kim (Cui & Kim, 2021) istražuju upotrebu ULS u doktorskim disertacijama iz oblasti podučavanja stranog jezika u proteklih 20 godina. Za potrebe istraživanja formirana su dva korpusa; svaki je imao malo više od milion reči. Prvi korpus se sastoji od 30 disertacija iz oblasti podučavanja stranog jezika objavljenih u periodu od 2001. do 2010. i sadrži 1.200.689 reči, dok se drugi sastoji od istog broja disertacija iz istovetne oblasti objavljenih u periodu od 2011. do 2020. godine i sadrži 1.189.404 reči. Tri manje standardne metodološke odluke koje se mogu naći u ovom istraživanju su: dužina istraženih ULS, upotreba *MI score* pri određivanju broja ULS u korpusima i osvrta samo na tokene pri analizi podataka. Istraživači su proučavali upotrebu ULS od tri, četiri i pet reči, iako je najtipičnije za ovu vrstu istraživanja da fokus bude na upotrebi ULS od četiri reči. Parametri za izdvajanje ULS su minimalna frekvencija od 20 ponavljanja u milion reči, minimalna disperzija u pet tekstova i minimalni *MI score* od tri.

Rezultati istraživanja su pokazali statistički značajnu razliku u drugačijoj učestalosti upotrebe ULS različitih dužina u dva vremenska perioda. Naime, ULS od tri i pet reči su tokom godina počele ređe da se koriste, dok se učestalost upotrebe ULS od četiri reči povećala. Za strukturalnu raspodelu ULS u istraživanju su koristili modifikovanu taksonomiju, koju su predložili Bajber i saradnici (Biber et al., 1999). Izmene koje su uneli jesu sledeće: u kategoriji imeničkih ULS dodali su potkategoriju *imenička fraza + in*; u kategoriji ULS sa fragmentima predloške fraze dodali su

potkategoriju *in + imenička fraza*; nisu uvrstili nijednu potkategoriju kod glagolskih ULS, a *to-infinitiv oblik* su načinili posebnom kategorijom (u originalnoj taksonomiji ona jedna od potkategorija glagolskih ULS). Prema rezultatima strukturalne analize najzastupljeniji su *imenički ULS* praćeni *ULS sa fragmentima predloške fraze* i sa *glagolskim ULS* na trećem mestu. Broj *imeničkih ULS* je porastao tokom godina (2001–2010. korpus – 53,5% svih spojeva; 2011–2020. korpus – 57,1%), kao i broj *glagolskih ULS* (2001–2010. korpus – 0,9% svih spojeva; 2011–2020. korpus – 3,8%), dok se broj ULS sa *predložkom frazom* smanjio (2001–2010. korpus – 35,5% svih spojeva; 2011–2020. korpus – 29,3%).

Za funkcionalnu raspodelu podataka korišćena je funkcionalna raspodela koju je predložio Hajland (Hyland, 2008). Istraživači su uradili funkcionalnu raspodelu samo u glavne kategorije i dobili sledeće rezultate: u oba korpusa najzastupljeniji su *ULS orijentisani ka istraživanju* (2001–2010. korpus – 68,3%, 2011–2020. korpus – 72,7%) praćeno *ULS orijentisanim ka tekstu* (2001–2010. korpus – 30,4%, 2011–2020. korpus – 26,6%), dok su na trećem mestu *ULS orijentisani ka učesniku* (2001–2010. korpus – 1,3%, 2011–2020. korpus – 0,7%). Kao što se može primetiti, tokom godina povećala se učestalost upotrebe *ULS orijentisanih ka istraživanju* dok se učestalost upotrebe druge dve grupe smanjila. Pri detaljnijoj analizi 100 najfrekventnijih ULS, istraživači su ustanovili da dva korpusa dele 59 ULS, što ukazuje na to da se upotreba ULS u ovoj oblasti brzo menja, u novim uslovima i kontekstima.

Palinkašević i Prtljaga (2024) bavile su se, takođe, proučavanjem promena u upotrebi ULS kroz vreme. Za potrebe njihovog istraživanja sačinjena su dva korpusa od po pola miliona reči. Prvi korpus je sačinjen od 91 naučnog rada iz oblasti obrazovanja koji su objavljeni od 2001. do 2010. godine, dok je drugi sačinjen od 90 naučnih radova iz oblasti obrazovanja koji su objavljeni od 2011. do 2020. godine. Specifično za ovo istraživanje je da su autorke pri formiranju korpusa uzele u obzir činjenicu da izvorni i neizvorni govornici imaju tendenciju da drugačije primenjuju ULS te su kontrolisale taj faktor. U oba istraživanja 23% radova pisali su neizvorni govornici engleskog jezika, dok su ostale radove pisali izvorni govornici. Za određivanje maternjeg jezika autora naučnih radova korišćena je Vudova (Wood, 2001) definicija, po kojoj je neophodno da autor ima tipično prezime nacije čiji je ciljani jezik maternji i da je institucija u kojoj je zaposlen iz države u kojoj je ciljani jezik zvaničan. Istraživanje se bavilo ULS od četiri reči koji su izdvojeni primenom minimalne frekvencije od 20 ponavljanja u milion reči i minimalne disperzije od 10% tekstova. U 2001–2010. korpusu izdvojeno je tipa 83 ULS, dok je u 2011–2020. korpusu izdvojeno 73. U ovom, kao i u prethodno predstavljenom istraživanju (Cui & Kim, 2021), učestalost upotrebe ULS od četiri reči se smanjila tokom godina.

Za strukturalnu analizu podataka Palinkašević i Prtljaga (2024) su koristile taksonomiju koju su predložili Bajber i saradnici (Biber et al., 1999). U oba korpusa najzastupljeniji su *ULS sa fragmentom predloške fraze*, praćeni *imeničkim ULS* i *glagolskim ULS* na trećem mestu. Statistička analiza podataka je pokazala da su se *imenički ULS* značajnije manje koristili u kasnijem korpusu tj. u korpusu 2011–2020, dok su se glagolske strukture koristile znatno više u kasnijem korpusu. Kada uporedimo ove rezultate sa rezultatima istraživanja Cui i Kim (2021), možemo uvideti da je raspodela ULS po grupama drugačija tj. da su *imenički ULS* učestaliji od *ULS sa fragmentima predloške fraze*. Ovakva razlika u rezultatima se može objasniti činjenicom da se u jednom radu istražuju naučni radovi iz oblasti obrazovanja dok se u drugom ispituju doktorske disertacije iz oblasti obrazovanja koje ipak imaju drugačije karakteristike. Zajedničko u podacima iz ova dva istraživanja je da se u oba tokom vremena povećava upotreba *glagolskih ULS*, što nam ukazuje na značaj ove promene i činjenicu da je ona prisutna u domenu obrazovanja nezavisno od registra, odnosno da li su u pitanju naučni radovi ili doktorske disertacije. Za detaljnije poređenje strukturalnih rezultata ova dva rada v. Tabelu 7.

Tabela 7. *Strukturalna raspodela ULS u radovima koji istražuju promenu upotrebe ULS u obrazovanju kroz vreme – tokeni u procentima*

Strukturalna kategorija	Cui & Kim (2021)		Palinkašević & Prtljaga (2024)	
	2001–2010.	2011–2020.	2001–2010.	2011–2020.
<i>Imenički ULS</i>	53.5	57.1	31.1	22.9
<i>ULS sa fragmentom predloške fraze</i>	35.5	29.3	49.2	50.9
<i>Glagolski ULS</i>	3.2	6.4	16.2	20.5
<i>Ostalo</i>	7.8	7.2	3.6	5.7

Palinkašević i Prtljaga (2024) su koristile taksonomiju koju je predložio Hajland (Hyland, 2008) za funkcionalnu raspodelu podataka. Rezultati su pokazali da su *ULS orijentisani ka testu* najzastupljenija kategorija praćeni *ULS orijentisanim ka istraživanju* i sa *ULS orijentisanim ka učesniku*, na trećem mestu. Statistička analiza podataka je pokazala da su *ULS orijentisani ka istraživanju* značajno manje korišćeni u kasnijem korpusu, dok su *ULS orijentisani ka tekstu* značajnije češće korišćeni u kasnijem korpusu. Ovakvi rezultati se ne poklapaju sa podacima u istraživanju Cui i Kim (2021), u kojem su najzastupljeniji *ULS orijentisani ka istraživanju* a razlike u upotrebi se kreću u suprotnim smerovima. Za detaljnije poređenje funkcionalne raspodele ULS u ova dva istraživanja pogledati Tabelu 8. Razlike u dobijenim rezultatima ukazuju na to da se funkcionalna upotreba ULS kroz vreme u oblasti obrazovanja drugačije menja u zavisnosti od registra.

Istraživanja navedena u ovom poglavlju pružaju nam uvid u dosadašnje razumevanje frazeologije prisutne u domenu obrazovanja, odnosno na specifičnu granu frazeoloških istraživanja koja se fokusira na učestale leksičke spojeve. Sagledan je razvojni put ovih istraživanja, kao i njihovi dometi, a nazire se i u kojim sve pravcima mogu ići buduća istraživanja.

Tabela 8. *Funkcionalna raspodela ULS u radovima koji istražuju promenu upotrebe ULS u obrazovanju kroz vreme – tokeni u procentima*

Funkcionalna kategorija	Cui & Kim (2021)		Palinkašević & Prtljaga (2024)	
	2001–2010.	2011–2020.	2001–2010.	2011–2020.
<i>ULS orijentisani ka istraživanju</i>	68.3	72.7	45.5	34.3
<i>ULS orijentisani ka tekstu</i>	30.4	26.6	46.1	57.7
<i>ULS orijentisani ka učesniku</i>	1.3	0.7	8.4	8.2

2.2. Prethodna istraživanja upotrebe ULS kod neizvornih govornika engleskog jezika kojima je maternji srpski ili rumunski jezik

Istraživanje u ovoj disertaciji se bavi razlikama u upotrebi ULS u dve grupe neizvornih govornika: (a) kako bi se bolje razumele karakteristike upotrebe ULS kod neizvornih govornika uopšteno i (b) uporedile potencijalne uopštene karakteristike sa onima koje su specifične za određene grupe neizvornih govornika zbog uticaja njihovog maternjeg jezika odnosno L1. Shodno tome, neophodno je detaljnije sagledati prethodna istraživanja koja se bave upotrebom ULS kod neizvornih govornika engleskog kojima je izvorni jezik srpski/rumunski, odnosno dva jezika koja će se istraživati u ovoj disertaciji. Ovo je važan korak koji daje uvid u to šta su dosadašnja istraživanja identifikovala kao karakteristike upotrebe ULS kod izvornih govornika srpskog i rumunskog jezika. Podaci iz prethodnih istraživanja omogućavaju tačniju identifikaciju karakteristika izvornih govornika srpskog i rumunskog u ovoj disertaciji. Mali broj istraživanja se bavio temom upotrebe ULS kod neizvornih govornika kojima je maternji jezik srpski ili rumunski, što je i jedan od razloga zašto su baš ta dva jezika izabrana za ovu disertaciju. Navedena činjenica, takođe, svedoči potencijalnom doprinosu ove disertacije istraživačkom opusu u datoj oblasti. Ni srpski ni rumunski jezik nisu često prisutni u istraživanjima ULS, a primećujemo da su brojnija istraživanja koja se bave izvornim govornicima rumunskog jezika.

Autorica ove disertacije pronašla je *šest naučnih radova u kojima se istražuje upotreba ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski*. Iako je ovo mali broj istraživanja, velika je verovatnoća da će se situacija uskoro promeniti i da će veći broj radova iz ove oblasti da se pojavi na teritoriji Rumunije zbog projekta DACRE (*Discipline-specific expert academic writing in Romanian and English: corpus-based contrastive analysis models*), koji se sprovodi na visokoškolskoj ustanovi Zapadni univerzitet u Temišvaru (rum. Universitatea de Vest din Timisoara; Chitez, Mureşan, & Rogobete, 2022). U okviru projekta formiran je EXPRES korpus (*Corpus of Expert Writing in Romanian and English*), koji je otvorenog pristupa. Međutim, korpus ne može da se preuzme u *txt formatu* sa sajta, već je formiran interfejs pomoću kojeg mogu da se pretražuju određene reči i fraze u korpusu, kao i u pojedinim delovima korpusa (npr. samo u određenoj oblasti). Pristup ovakvom resursu olakšava istraživanja u oblasti korpusne lingvistike iako sužava tipove istraživanja zato što su opcije na interfejsu ograničene. Nemoguće je uraditi analizu ULS u datom korpusu preko interfejsa, jer nisu date opcije izdvajanja ULS. Međutim, na sajtu je dostupan spisak prvih 100 n-grama od dve, tri, četiri i pet reči. Nije naznačeno koja je minimalna frekvencija i disperzija određena za izdvajanje n-grama odnosno ULS.

Najnoviji EXPRESS korpus nije jedini korpus tog tipa koji je dostupan u Rumuniji i koji je koristan za istraživanja frazeologije pomoću korpusne lingvistike. Od korpusa koji su otvorenog pristupa postoji i RODGER korpus (*Romanian Corpus of Learner English*), koji je oformila istraživačka grupa u centru CODHUS (*Center for Corpus Related Digital Humanities*; Chitez, Strilechi & Csürös, 2022). RODGER korpus je bilingvalan korpus sa resursima prikupljenim od strane izvornih govornika rumunskog na engleskom i na rumunskom jeziku. Ovaj korpus se fokusira na produkciju učenika i omogućava istraživanje njihovog međujezika (Chitez i sar., 2022).

Oba gorenavedena korpusa su korišćena kao resurs za pojedina istraživanja koja su opisana u ovom poglavlju. Kao i u narednim poglavljima, istraživanja koja opisujemo stavićemo po redosledu relevantnosti za ovu disertaciju, kao i sličnosti i kvalitetu metodološke postavke istraživanja. Počecemo od radova koji se bave ULS ali kojima istraživanje ULS nije glavni cilj, kao i od istraživanja u kojima se ispituje produkcija nekoliko grupa neizvornih govornika engleskog među kojima su izvorni govornici rumunskog jezika. Prvo ćemo navesti istraživanja koja se bave produkcijom neizvornih govornika engleskog kojima je maternji jezik rumunski pa onda srpski

jezik zbog većeg broja istraživanja koja mogu da se nađu za rumunski kontekst tj. izvorne govornike rumunskog jezika.

Kitez, Murešan Rogobete (2022)

Madalina Kitez, Valentina Karina Murešan i Roksana Rogobete (Chitez, Murešan, & Rogobete, 2022) u svom istraživanju predstavljaju EXPRESS korpus i šta on nudi za oblast korpusnog istraživanja, uz kratak osvrt na deset najfrekventnijih ULS od tri i četiri reči. EXPRESS korpus je trenutno u završnoj fazi formiranja, a sastojće se od tri potkorpusa naučnih radova koji su pisani na rumunskom jeziku, naučnih radova koje su izvorni govornici rumunskog pisali na engleskom jeziku i naučnih radova koje su izvorni govornici engleskog pisali na engleskom jeziku. U sva tri potkorpusa nalaziće se naučni radovi iz četiri naučne discipline: lingvistike, ekonomije, političkih nauka i informatike. Potkorpusi će biti ujednačeni po dužini i po broju radova (po 200 radova iz svake discipline). U istraživanju je najavljeno da će finalna verzija korpusa biti gotova i dostupna na internetu u julu 2022. godine, međutim, u februaru 2024. godine na sajtu gde je korpus dostupan vidi se da još nije u potpunosti gotov, jer nedostaje manji broj radova u tri discipline u korpusima na engleskom jeziku, a korpus na rumunskom jeziku je u potpunosti završen. Ni u radu, a ni na sajtu nije naznačeno koji su kriterijumi korišćeni za definisanje izvornih govornika ni koji su kriterijumi korišćeni za izdvajanje ULS (n-grama), odnosno koja je postavljena minimalna frekvencija i disperzija. U radu je navedeno deset najfrekventnijih n-grama od tri i četiri reči u potkorpusu naučnih radova na rumunskom jeziku i u potkorpusu radova na engleskom pisanih od strane izvornih govornika rumunskog jezika. Nije se zašlo u dublju analizu podataka i sâm rad je značajan prevashodno kao izvor budućih resursa za istraživanja u ovoj oblasti.

Nedelcu Pascu i Brad (2018)

Maria Nedelcu Pascu i Laura Brad (Nedelcu Pascu & Brad, 2018) sprovele su istraživanje čiji je glavni cilj da se rumunskim istraživačima iz oblasti ekonomije olakša pisanje diskusija i zaključaka u naučnim radovima na engleskom jeziku. Autorke su odlučile da cilj postignu kroz pet koraka među kojima je i kreiranje spiska ULS od četiri reči koji bi istraživačima iz oblasti ekonomije koristio kao resurs pri pisanju naučnih radova. Za potrebe istraživanja oformljen je korpus sa dva potkorpusa. Prvi se sastoji od 17 naučnih radova iz oblasti ekonomije pisanih na engleskom jeziku od strane izvornih govornika rumunskog jezika, drugi se sastoji od 32 naučna rada iz oblasti ekonomije pisanih na engleskom od strane izvornih govornika engleskog jezika. Ne navodi se koji je razlog neujednačenog broja radova u dva potkorpusa. Iz radova su izdvojeni delovi diskusija i zaključak i na taj način je formiran korpus od 77.805 reči. Autorke navode da su dva potkorpusa neujednačena po veličini (potkorpus neizvornih govornika engleskog jezika je znatno manji od potkorpusa izvornih govornika), te su one izdvojile ULS samo iz korpusa izvornih govornika i uporedile njihovu upotrebu u korpusu neizvornih govornika. Ne navodi se koji je ukupan broj izdvojenih ULS u potkorpusu izvornih govornika niti koji se broj istih našao u potkorpusu neizvornih govornika. Autorke, takođe, ne navode koja se minimalna frekvencija i minimalna disperzija koristila za identifikaciju ULS. Izdvojeni ULS podeljeni su u funkcionalne kategorije i potkategorije po Bajberovoj taksonomiji (1999). U kategoriji referencijalnih spojeva navele su 24 tipa, u kategoriji organizatora diskursa navele su 11 tipova, dok nijedan ULS nije identifikovan u kategoriji izraza stava. Autorke istraživanja navode da su korpusi isuviše mali i neujednačeni po veličini da bi bilo kakva generalizacija mogla da se izvede iz dobijenih podataka. Istraživanje nudi nagoveštaj podataka koji bi mogli da se dobiju istraživanjima na većim korpusima i sa boljom kontrolom parametara neophodnih za istraživanja ULS, a koja bi generisala uopštene zaključke i dala dublji uvid u problematiku.

Dinka i Kitez (2021)

Dinka i Kitez (Dinca & Chitez, 2021) u svom istraživanju ispitivale su upotrebu ULS od četiri i pet reči kod studenata informatike i filologije kojima je maternji jezik rumunski. Za potrebe istraživanja korišćena su dva korpusa koji se sastoje od po 40 eseja. ESP-IT je korpus eseja studenata informatike na osnovnim studijama, kojima je maternji jezik rumunski, i sastoji se od 63.842 reči. ESP-LIT je korpus eseja studenata književnosti na master studijama, kojima je maternji jezik rumunski, i sastoji se od 67.529 reči. Istraživanje je sprovedeno korišćenjem alata *Lancsbox* (Brezina, Weill-Tessier & McEnery, 2020, prema Dinca & Chitez, 2021) i posebnog programa dizajniranog korišćenjem *Python* jezika. Prvi korak u istraživanju bio je izdvajanje ULS iz dva korpusa i njihovo poređenje. Za izdvajanje ULS korišćena je minimalna frekvencija od 0.25 ponavljanja u 10.000 reči i minimalna disperzija od dva teksta. Autorke ne navode koji je broj izdvojenih ULS u svakom od korpusa pojedinačno, već navode da je ukupno četiri ULS od pet reči i 15 ULS od četiri reči zajedničko za oba korpusa. U istraživanju autorke ne koriste tipične taksonomije za strukturalnu i funkcionalnu raspodelu ULS već ih dele na dve kategorije: ULS koji su organizatori diskursa i ULS povezani sa temom. Rezultati su pokazali da studenti informatike koriste više ULS povezanih sa temom od studenata književnosti. Drugi korak u istraživanju za koji su autorke dizajnirale poseban program je poređenje studentske produkcije sa produkcijom profesionalaca koji su izvorni govornici engleskog jezika. Naime, autorke su izdvojile 36 akademskih fraza specifičnih za uvode kratkih radova koristeći *Manchester Academic Phrasebank* (Davis & Morley, 2018, prema Dinca & Chitez, 2021) i proverile koliko se često oni koriste u dva studentska korpusa. Nijedna od izdvojenih fraza nije se identično pojavila u studentskim korpusima, a identifikovane su samo dve fraze u kojima postoji poklapanje od tri reči. Autorke su takođe ustanovile da studenti informatičkih nauka koriste zamenicu u prvom licu množine *we* kada navode šta su uradili u istraživanju, dok studenti književnosti koriste zamenicu za prvo lice jednine *I*. Ovo istraživanje, iako navodi jasne parametre za izdvajanje ULS, ne koristi standardne taksonomije za funkcionalnu i strukturalnu klasifikaciju ULS, te nam ne nudi informacije koje bi mogle da se koriste za spekulaciju o mogućim trendovima u korišćenju ULS kod izvornih govornika rumunskog jezika.

Hajland i Jiang (2022)

Hajland i Jiang (Hyland & Jiang, 2022) istražuju koliko se razlikuje upotreba ULS kod neizvornih govornika različitih jezika od konvencija koje se primenjuju u objavljenim radovima. Činjenica da se u ovom istraživanju ne porede već objavljeni radovi neizvornih i izvornih govornika već nelektorisane verzije radova neizvornih govornika pre objavljivanja i lektorisani objavljeni radovi – čini ovo istraživanje interesantnim i njegov doprinos razumevanju upotrebe ULS većim. Dva korpusa su korišćena u istraživanju: SciELF (2015) i akademski deo BNC (*British National Corpus*). SciELF je korpus nelektorisanih naučnih radova neizvornih govornika, odnosno finalnih verzija radova pre objavljivanja. Sastoji se od 150 radova (759.300 reči) koje su pisali neizvorni govornici deset različitih maternjih jezika (finski, češki, francuski, kineski, španski, ruski, švedski, italijanski, portugalski i rumunski). Nisu svi neizvorni govornici jednako zastupljeni u korpusu. Najzastupljeniji su izvorni govornici finskog jezika sa 25 radova, dok su najmanje zastupljeni izvorni govornici rumunskog sa četiri rada. SciELF korpus sadrži radove iz naučnih oblasti humanistike (23 rada), društvenih nauka (49 radova), bioloških nauka (19 radova) i prirodnih nauka (59 radova). Radove u korpusu pisali su studenti istraživači (30 radova), mlađi istraživači (86 radova) i iskusni istraživači (34 rada). Iz BNC korpusa izdvojen je isti broj radova iz istih naučnih oblasti kako bi se formirao korpus prikladan za poređenje.

Za izdvajanje ULS Hajland i Jiang (2022) koristili su minimalnu frekvenciju od 20 ponavljanja u milion reči i minimalnu disperziju u 10% tekstova. Analizirali su samo ULS od četiri reči, a

konačna lista ULS dobijena je nakon što su uklonjene imeničke fraze specifične za kontekst i nakon što su spojeni ULS od četiri reči koji formiraju jedan ULS od pet reči. Koristeći gorenavedene kriterijume, u SciELF korpusu identifikovano je 134 tipova i 3.316 tokena, dok je u BNC korpusu izdvojeno 458 tipova i 9.221 token. Može se reći da su ovakvi rezultati dosta iznenađujući s obzirom na to da neizvorni govornici prema mnogim prethodnim istraživanjima imaju tendenciju da koriste veći broj ULS uporedno sa izvornim govornicima (Pan et al., 2016; Güngör & Uysal, 2016; Al Fajri et al., 2020), iako se ovde porede neelektorisani radovi neizvornih govornika sa objavljenim radovima. Broj izdvojenih ULS se razlikuje po grupama izvornih govornika iako na takve rezultate, kao što i sami autori navode, utiču nejednake veličine korpusa. Najveći broj ULS identifikovan je kod izvornih govornika finskog jezika sa 116 ULS, dok je najmanji broj ULS, ukupno devet ULS, identifikovan kod izvornih govornika rumunskog jezika, što se i poklapa sa brojem radova sa kojima su ovi govornici zastupljeni u korpusu.

Autori su, isto tako, uporedili deset najfrekventnijih ULS kod različitih grupa govornika i te podatke uporedili sa deset najfrekventnijih ULS u korpusu objavljenih radova. Utvrdili su da se od 20 najfrekventnijih ULS u korpusu objavljenih radova samo osam nalazi u deset najfrekventnijih ULS kod različitih grupa izvornih govornika u SciELF korpusu. Najfrekventnijih deset ULS se umnogome razlikuju kod različitih grupa govornika, npr. najfrekventniji ULS u potkorpusu izvornih govornika kineskog (*than when compared to*) ne pojavljuje se na spisku deset najfrekventnijih ULS ni za jedan drugi jezik. Od 134 ULS identifikovanih u SciELF korpusu preko pedeset je specifično za samo jednu grupu govornika. Kada je reč o frekvenciji upotrebe ULS u odnosu na iskustvo autora, ustanovljeno je da istraživači sa najmanje iskustva koriste najviše ULS; ovakvi nalazi su u skladu sa prethodnim istraživanjima koja su pokazala da subjekti sa nižim nivoom znanja stranog jezika/iskustva u akademskom pisanju, imaju tendenciju da koriste veći broj ULS (Bychkovska & Lee, 2017; Nam, 2017); i ovde moramo da uzmemo ove podatke sa rezervom zato što veličine korpusa nisu ujednačene i korpus studentskih istraživača ima preko 60% više reči od druga dva potkorpusa. A kao što je već pomenuto manji korpusi imaju tendenciju da generišu veći broj ULS od velikih korpusa (Bestgen, 2019).

Za strukturalnu raspodelu ULS, Hajland i Jiang (2022) koriste taksonomiju koju su predložili u ranijem radu (v. Hyland & Jiang, 2018). Naime u toj taksonomije glavne strukturalne kategorije su:

1. *glagolske fraze*;
2. *fragmenti zavisne klauze*;
3. *imeničke/predloške fraze*.

S obzirom na nejednaku veličinu potkorpusa u istraživanju, za strukturalnu i funkcionalnu analizu podataka korišćene su proporcije upotrebe ULS umesto čistih frekvencija, a *log-likelihood* test je korišćen kako bi se ustanovilo da li su rezultati statistički značajni. Strukturalna analiza podataka je pokazala da je u oba korpusa najzastupljenija kategorija *imeničke/predloške fraze*, praćena *glagolskim frazama* i sa *fragmentima zavisne fraze*, na trećem mestu. Međutim, u korpusu objavljenih radova više se koriste *imeničke/predloške fraze* (sačinjavaju 77% ULS u korpusu), nego u korpusu neobjavljenih radova neizvornih govornika, u kojem *imeničke/predloške fraze* sačinjavaju 65% ULS u korpusu. Najzastupljeniji oblik ULS u oba korpusa je *imenička fraza sa of fragmentom* (čini više od 25% ULS u oba korpusa), praćena *glagolskom frazom* u pasivu. U potkorpusu rumunskih izvornih govornika može se videti da su najzastupljenije *imeničke/predloške fraze* (15.23% na 10.000 reči) praćene *glagolskim frazama* (3.51% na 10.000 reči). Isti se redosled zastupljenosti strukturalnih grupa pokazao i za druge grupe neizvornih govornika. Štaviše, rezultati su pokazali da istraživačko iskustvo nije uticalo na rezultate raspodele ULS po grupama izvornih govornika, što, kako autori navode, može da ukaže na to da faktor pisanja na stranom jeziku značajnije utiče na produkciju ULS od faktora istraživačkog iskustva (Hyland & Jiang, 2022: 10).

Kada se uporedila strukturalna raspodela ULS u objavljenim i neobjavljenim radovima, uvidelo se da se uvodna *it* konstrukcija značajno češće koristi u neobjavljenim radovima iz oblasti humanistike i društvenih nauka uporedno sa objavljenim radovima iz ovih oblasti. Takođe je utvrđeno da su fragmenti klauzalne fraze češće korišćeni u neobjavljenim radovima iz humanistike.

Kada je reč o funkcionalnoj raspodeli ULS, autori su koristili Hajlandovu (2008) takosnomiju. Funkcionalna raspodela ULS po grupama poprilično je ujednačena u oba korpusa. *ULS orijentisani ka istraživanju* i *ULS orijentisani ka tekstu* sačinjavaju po oko 40% korpusa a ostatak ULS kategorisanje u funkcionalnu grupu *ULS stava*. Veće se razlike mogu uvideti kada se poredi funkcionalna raspodela ULS kod neizvornih govornika engleskog jezika sa različitim maternjim jezikom. Primera radi, izvorni govornici kineskog i francuskog jezika koriste *ULS orijentisane ka istraživanju* (kineski izvorni govornici 72%, francuski izvorni govornici 65%) češće nego što koriste *ULS orijentisane ka tekstu* (kineski izvorni govornici 25%, francuski izvorni govornici 32%). Na ovaj način oni u svojim tekstovima naglašavaju laboratorijsku dimenziju radova. Nasuprot tome, izvorni govornici rumunskog jezika češće su koristili *ULS orijentisane ka tekstu* (48%) uporedno sa *ULS orijentisanim ka istraživanju* (31%). Od svih grupa neizvornih govornika, *ULS stava* najčešće su koristili izvorni govornici portugalskog jezika (25%). Iskustvo u akademskom pisanju i naučna oblast nisu značajno uticali na funkcionalnu raspodelu ULS.

Iako je rumunski jezik samo jedan od deset jezika koji su obrađeni u istraživanju Hajland i Jiang (2022), ono nam daje korisne informacije o tome kako se ULS koriste u naučnim radovima kod ove grupe govornika i kako se upotreba ULS razlikuje kod različitih grupa neizvornih govornika. Moglo bi se reći da je glavni nedostatak ovog istraživanja činjenica da su potkorpusi različitih grupa neizvornih govornika dosta mali i neujednačeni po veličini.

Bercuči (2020)

Istraživačica Loredana Bercuči (Loredana Bercuci) napisala je dva rada koji nude značajne informacije o karakteristikama upotrebe ULS kod neizvornih govornika kojima je maternji jezik rumunski, odnosno dva rada koji su značajni za temu ove disertacije. Prvo istraživanje koje ćemo opisati Bercuči (Bercuci, 2020) bavi se upotrebom ULS kod izvornih i neizvornih govornika u naučnim radovima iz oblasti političkih nauka. Tri korpusa su korišćena u istraživanju:

1. *Political Science Expert English Corpus* (PSE), sačinjen od 50 naučnih radova pisanih od strane izvornih govornika engleskog jezika veličine 532.702 reči;
2. *Political Science Expert Romanian Corpus* (PSRo), sačinjen od 50 naučnih radova na rumunskom jeziku veličine 268.671 reči;
3. *Political Science L2 English Expert Corpus* (PSRoE), sačinjen od 50 naučnih radova na engleskom jeziku koje su pisali izvorni govornici rumunskog jezika a koji je veličine 344.787 reči.

Kao što se može videti iz navedenih podataka o korpusima, istraživačica je korpuse ujednačila po broju naučnih radova. Stoga je došlo do razlike u veličini korpusa pošto se ovde, kao i u drugim istraživanjima sličnog tipa (Güngör, 2016; Pan i sar., 2016), pokazalo da istraživači različitih nacionalnosti imaju tendenciju da pišu naučne radove različitog obima. Jedan nedostatak u ovom istraživanju je činjenica da nije navedeno na osnovu kojih kriterijuma je utvrđen maternji jezik autora naučnih radova koji sačinjavaju korpuse. Štaviše, istraživanje se bavi ULS od tri, četiri i pet reči, međutim, ne navodi se koja je minimalna frekvencija i disperzija korišćena za izdvajanje ULS, jedino se navodi da se koristio program *LancsBox* (Brezina, Weill-Tessier & McEnery, 2020, prema Bercuci, 2020).

U analizi podataka, glavni fokus je stavljen na poređenje prvih 20 najfrekventnijih ULS u tri analizirana korpusa. Nije dat ukupan broj izdvojenih ULS u korpusima, jedino je data šema funkcionalne raspodele ULS u PSRoE i PSE, u kojoj nisu dati procenti već se porede frekvencije bez navedenih preciznih brojeva. U istraživanju se navodi da su neizvorni govornici koristili veći broj ULS. Po funkcionalnoj raspodeli se vidi da su u oba korpusa najzastupljeniji *ULS orijentisani ka tekstu*, tome slede *ULS orijentisani ka istraživanju*, dok su najmanje zastupljeni *ULS orijentisani ka učesniku*. Izvorni govornici koriste više ULS jedino u funkcionalnoj kategoriji *ULS orijentisanih ka učeniku*.

Pri analizi 20 najfrekventnije korišćenih ULS u PSE korpusu utvrđeno je da se najčešće koriste *ULS orijentisani ka tekstu*, a u toj kategoriji najzastupljenije potkategorije su *rezultativni i ograničavajući signali* praćeni tranzicionim i strukturalnim ULS. U kategoriji *ULS orijentisanih ka istraživanju* najzastupljenija kategorija su *ULS za kvantifikaciju* praćeni *ULS za proceduru i lokaciju*. U kategoriji *ULS orijentisanih ka učesniku* potkategorija *uključivanja* je zastupljenija od potkategorije *stava*.

Pri analizi 20 najfrekventnije korišćenih ULS u PSRo korpusu utvrđeno je da se najčešće koriste *ULS orijentisani ka tekstu*, a u toj kategoriji najzastupljenije potkategorije su *ograničavajući i tranzicioni signali*. Na drugom mestu je kategorija *ULS orijentisanih ka učesniku*, a od dve potkategorije jedino je zastupljena potkategorija *ULS stava*. Kategorija *ULS orijentisanih ka istraživanju* je u ovom korpusu najmanje zastupljena. Interesantno je da potkategorija *kvantifikacije* nije zastupljena.

Pri analizi 20 najfrekventnije korišćenih ULS u PSRoE korpusu utvrđeno je da se najčešće koriste *ULS orijentisani ka tekstu*, a u toj kategoriji najzastupljenija je potkategorija *ograničavajući signali*. Autorka smatra da se ovde najjače ogleda uticaj maternjeg jezika na produkciju na stranom jeziku. Na primer, *ULS in terms of* (ekvivalenti na rumunskom jeziku: *in ceea ce priveste i cu privire la*) i *ULS from this point of view* (ekvivalent na rumunskom – *din punct de vedere*) su mnogo češći u PSRoE korpusu nego u korpusu izvornih govornika.

Iako postoje mnoge informacije koje nisu date u ovom istraživanju, a koje bi bile korisne za poređenje sa istraživanjem u ovoj disertaciji (npr. procentualna raspodela ULS po strukturalnim i funkcionalnim kategorijama za svaki od tri korpusa), ono ipak daje značajne podatke koji nam pružaju uvid u tendencije korišćenja ULS koje imaju neizvorni govornici engleskog jezika kojima je rumunski maternji jezik.

Bercuči (2021)

Drugi rad autorke Bercuči o kojem ćemo detaljnije diskutovati je rad koji se nadovezuje na prethodni i koji opisuje upotrebu ULS u esejima na engleskom kod studenata političkih nauka kojima je maternji jezik rumunski (Bercuci, 2021). Za istraživanje su korišćena tri korpusa iz prethodnog istraživanja: PSE, PSRo i PSRoE, uz još jedan korpus studentske produkcije na engleskom PSL. Četvrti korpus PSL se sastoji od 50 radova (eseja, master radova) koje su studenti Zapadnog univerziteta u Temišvaru (rum. Universitatea de Vest din Timișoara) pisali na engleskom jeziku. Veličina korpusa je 146.066 reči, odnosno najmanji od korpusa koji se koriste u istraživanju. Korišćeni radovi su deo ROGER (*Romanian academic GENRes*) korpusa, koji je spomenut na početku ovog poglavlja.

U istraživanju Bercuči (2021), kao i u prethodnom Bercuči (2020), izdvojeni su ULS od tri, četiri i pet reči. Međutim, nije navedeno koja je minimalna frekvencija i distribucija korišćena za identifikaciju ULS. Istraživanje se fokusira na funkcionalnu analizu izdvojenih ULS, te nije naveden ukupan broj izdvojenih ULS u PSL. Ono što se može videti iz podataka jeste da su studenti

koristili najveći broj ULS a izvorni govornici najmanji broj ULS. Ovakve distribucije ULS su u skladu sa prethodnim istraživanjima (Bychkovska & Lee, 2017; Shin, 2019), međutim, navedeni rezultati bi mogli da se objasne i neujednačenim veličinama korpusa pošto manji korpusi imaju tendenciju da proizvedu veći broj ULS (Cortes, 2004). Interesantno je i da studenti frekventnije koriste izdvojene ULS od drugih grupa govornika, što ukazuje na manje bogat vokabular i veći nivo ponavljanja. Autorka smatra da su određeni ULS češći u upotrebi studenata zbog uticaja maternjeg jezika na njihovu produkciju na engleskom jeziku. Primer za ovakav zaključak je češća upotreba ULS *because of the fact that* i *in order to be able*, čiji se direktni prevodi *datorita faptului ca*, *in contextul* i *pentru a putea sa* često koriste u formalnom diskursu na rumunskom jeziku.

Funkcionalna raspodela ULS u studentskom korpusu PSL pokazala je da studenti najčešće koriste *ULS orijentisane ka istraživanju* praćene *ULS orijentisanim ka tekstu*, dok *ULS orijentisane ka učesniku* najređe koriste. Bercuči smatra da studenti najčešće koriste *ULS orijentisane ka istraživanju* zato što imaju tendenciju da budu opisnog karaktera te se lakše imitiraju. Zbog te činjenice ih često koriste i tokom učenja normi u akademskom pisanju i na taj način nastane i efekat „plišanog mede” (*the Teddy Bear effect*), odnosno učestalo korišćenje ograničenog broja fraza na koje su se učenici stranog jezika navikli (Ellis, 2012). U studentskom korpusu u grupi *ULS orijentisani ka istraživanju* dve najzastupljenije podgrupe su *ULS za kvantifikaciju* i *proceduralni ULS*. U grupi *ULS orijentisanih ka tekstu* ubedljivo najzastupljenija potkategorija su *tranzicioni signali* kao što su *as well as* i *on the other hand*. Interesantno je da su *ograničavajući signali* mnogo ređi u studentskom korpusu uporedno sa korpusima neizvornih i izvornih profesionalaca, odnosno u korpusima PSRoE i PSE. Studenti češće koriste *ULS orijentisane ka učesniku* od istraživača kojima je maternji jezik rumunski a koji pišu na engleskom. Međutim, kada se detaljnije pogledaju primeri, može se videti da su to konstrukcije koje odstupaju od formalnog registra npr. *in my opinion* i *from my point of view*.

Pregledom prethodnih istraživanja koja sa bave upotrebom ULS *kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski*, možemo uvideti veliku potrebu za daljim istraživanjima u datoj oblasti. Naime, metodološki gledano, teško je porediti rezultate navedenih radova sa drugim radovima iz ove oblasti koji su ispitivali neizvorne govornike drugih maternjih jezika i to iz razloga koje donosimo u nastavku.

- Samo dva rada (Dinca i Chitez, 2021 i Hyland i Jiang, 2022) od šest navode kriterijume korišćene pri izdvajanju ULS, odnosno, ekspliciraju koja je postavljena minimalna frekvencija i disperzija. Pošto u istraživanjima ULS ne postoji jedan fiksni kriterijum za izdvajanje ULS već je u upotrebi dijapazon kriterijuma različitih nivoa „strogosti”,ovo je značajna informacija pri poređenju dobijenih rezultata u različitim istraživanjima (na primer, ako se primeni minimalna frekvencija od 40 ponavljanja u milion reči, broj izdvojenih ULS će biti manji nego ako se primeni minimalna frekvencija od 20 ponavljanja u milion reči). Isto važi i kada je reč o disperziji – ona se kreće od minimalnog ponavljanja u samo nekoliko radova (ili drugih elemenata koji sačinjavaju korpus – npr. disertacija ako je korpus studentske produkcije) do 10% radova u korpusu.
- Ni u jednom radu se ne navodi koji su kriterijumi korišćeni za utvrđivanje maternjeg jezika govornika čija se produkcija koristila u formiranju korpusa.
- Korpusi korišćeni u istraživanjima su mahom mali i neujednačeni po veličini. Poznato je da veoma mali korpusi nisu pogodni za istraživanje ULS (Bestgen, 2019). Međutim, u pojedinim istraživanjima u ovom poglavlju imamo veoma male korpuse. Npr. u istraživanju Nedelcu Pascu i Brad (2018) imamo dva neujednačena korpusa koji se koriste za istraživanje. Ne navodi se koliko je veliki koji korpus, ali se navodi da je korpus neizvornih govornika značajno manji, a da se dva korpusa zajedno sastoje od 77.805 reči, što bi bio veoma mali korpus i da se ne sastoji iz dva dela koja je neophodno zasebno analizirati. U istraživanju Hajland i Jiang (2022)

koriste se dva korpusa zadovoljavajuće veličine od preko 700.000 reči, što u potpunosti zadovoljava kriterijume za poređenje neobjavljenih radova neizvornih govornika i objavljenih radova. Ali kada se sagledavaju karakteristike upotrebe ULS kod različitih grupa neizvornih govornika, u tom istraživanju, velike su razlike u zastupljenosti različitih neizvornih govornika, što onemogućava pouzdano poređenje. Npr. u SciELF korpusu istraživači kojima je rumunski jezik maternji predstavljeni su sa četiri naučna rada, dok su istraživači kojima je maternji jezik finski predstavljeni sa 25 naučnih radova od ukupno 150 radova koji sačinjavaju korpus. Bercuči u svom istraživanju korpuse ujednačuje po broju radova što je prouzrokovalo velike razlike u veličini korpusa koje poredi, što dalje komplikuje pouzdanost dobijenih podataka jer veličina korpusa utiče na broj ULS koji će se iz njega izdvojiti (Cortez, 2004).

- U četiri od šest radova ne dobijamo informacije o ukupnom broju izdvojenih ULS u korpusima, a to je potrebno kako bismo mogli da utvrdimo i uporedimo učestalost upotrebe ULS u tim grupama govornika i u datim kontekstima.
- U pet od šest radova ne dobijamo strukturalnu i funkcionalnu raspodelu svih izdvojenih ULS u korpusima koje bi pokazale način upotrebe ULS kod tih grupa govornika. U istraživanjima kod Bercuči dobijamo funkcionalnu raspodelu svih ULS u korpusima, međutim, predstavljeni su samo u dijagramu i tačne brojke nisu jasno vidljive. U istraživanju Chitez et al. (2022) data je samo kratka analiza deset najfrekventnijih ULS bez daljih podataka o ukupnoj strukturalnoj i funkcionalnoj raspodeli ULS u korpusima. U istraživanju Dinca i Chitez (2021) ULS su podeljeni na *organizatore diskursa* i *ULS koji su povezani sa tekstom*, odnosno nisu primenili prihvaćenu strukturalnu ili funkcionalnu taksonomiju. Štaviše, u istraživanjima ULS tipično se *ULS povezani sa tekstom* (specifični za određenu temu) izdvajaju u procesu prečišćavanja spiska ULS pre njegove finalizacije (Salazar, 2014; Güngör, 2016). Nemoguće je bilo dati podatke o strukturalnoj i funkcionalnoj raspodeli svih ULS u istraživanju Nedelcu Pascu i Brad (2018) iz prostog razloga što su zbog neujednačenih veličina korpusa izdvojili ULS samo u korpusu izvornih govornika.

Uprkos gorenavedenim nedostacima u dosadašnjim istraživanjima, ona su korisna za građenje slike o potencijalnim karakteristikama upotrebe ULS kod neizvornih govornika kojima je maternji jezik rumunski. Pregled osnovnih informacija o radovima koji se bave upotrebom ULS kod neizvornih govornika kojima je maternji jezik rumunski dostupan je u Tabeli 9 datoj na narednoj stranici.

Kada je reč o istraživanjima upotrebe ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski, autorka disertacije je pronašla samo tri rada koji se bave ovom temom. Sva tri rada jesu radovi iste autorke Katarine Lazić ili su pisani u koautorstvu sa njom. Tema njene doktorske disertacije (Lazić, 2017), a ujedno i jednog od najznačajnijih prethodnih istraživanja za ovu disertaciju, jeste upotreba ULS u naučnim radovima iz oblasti biotehničkih nauka kako kod izvornih govornika tako i kod neizvornih govornika kojima je maternji jezik srpski. Njenu disertaciju smo detaljno predstavili u narednom poglavlju 2.3. *Istraživanja upotrebe ULS u naučnim radovima kod izvornih i neizvornih govornika*, kako bismo na jednom mestu mogli da se osvrnemo na sva najznačajnija prethodna istraživanja koja direktno utiču na postavku istraživanja u ovoj disertaciji, kao i na kasniju analizu rezultata iste. Stoga, u ovom poglavlju predstavljamo samo dva istraživanja koja se nadovezuju na disertaciju Lazićeve (2017) i koja nam pružaju praktičnu primenu rezultata koji su dobijeni u samoj disertaciji.

Naučni rad Lazić (2019) kao glavni cilj ima da predstavi kako da se može unaprediti podučavanje engleskog jezika za specifične svrhe, konkretno u oblasti biotehničkih nauka, uključivanjem obrade ULS u radu sa studentima. Autorka navodi u koracima kako uvesti obradu ULS u nastavu engleskog jezika za specifične namene. Prvi korak je identifikacija ULS koji će se obrađivati sa studentima. Za ovu svrhu, za oblast biotehničkih nauka, autorka je koristila korpus CoBNEA kako bi izdvojila ULS koji se koriste u svim potkorpusima ovoga korpus (šumarstvo, ekološki inženjering, prerada drveta i pejzažna arhitektura), kao i u tri od četiri korpusa. Na ovaj način se obrađuju najzastupljeniji ULS u čitavoj oblasti a samim tim i ULS sa najvišom pedagoškom vrednošću. Ukupno 16 ULS se javlja u svim potkorpusima, dok se 18 ULS javlja u po tri potkorpusa. U konkretnom radu sa studentima, autorka predlaže da je najpre važno uvesti teorijsku nastavu koja bi upoznala studente sa pojmom i značajem ULS, nakon čega slede vežbe za:

1. *prepoznavanje ULS* – npr. za prepoznavanje ULS iz funkcionalne potkategorije *tranzicionih signala*, studenti dobiju zadatak da pročitaju odlomke iz različitih tekstova i da podvuku delove teksta koje autori koriste kada prelaze na drugu temu, dodaju informacije ili za izražavanje ideje koja je u suprotnosti sa prethodnom idejom;
2. *upoznavanje sa formom i funkcijom ULS* – npr. studenti dobiju mnoštvo primera upotrebe jednog ULS, a njihov zadatak jeste da obrate pažnju na reči koje prethode i prate spoj pa da uoče obrasce u njegovoj upotrebi;
3. *pravilnu upotrebu ULS* – npr. studenti dobijaju rečenice u kojima nedostaju ULS, a njihov je zadatak da ponuđene ULS stave u prave rečenice/kontekste.

Istraživanje Lazić (2019) je značajno jer nudi metodološki okvir za obradu ULS u nastavi engleskog za specifične namene i konkretne primere.

Tabela 9. Pregled istraživanja upotrebe ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski

Istraživanje	Korpus	Ciljna grupa istraživanja	Dužina ULS koja se istražuje	Kriterijumi za izdvajanje ULS	Primenjena analiza	Zaključci značajni za dalja istraživanja
1. Chitez, Mureşan & Rogobete, 2022.	EXPRESS korpus – nezavršena verzija primenjena	Naučni radovi – četiri naučne oblasti	tri i četiri reči	Minimalna frekvencija i disperzija nisu navedeni; alat: <i>Sketch Engine</i>	Poređenje 10 najfrekventnijih ULS u korpusima izvornih i neizvornih govornika.	Informacije o korpusima za analizu L1/L2 produkcije kod izvornih govornika rumunskog jezika.
2. Nedelcu Pascu & Brad, 2018.	Dva korpusa sa ukupno 77.805 reči	Naučni radovi (diskusija i zaključci) – iz oblasti ekonomije	četiri reči	Minimalna frekvencija i disperzija nisu navedeni;	Analiza upotrebe ULS koji su izdvojeni u korpusu izvornih govornika u korpusu neizvornih govornika; funkcionalna analiza ULS.	Istraživanje nam nudi nagoveštaj podataka o funkcionalnoj raspodeli ULS kod neizvornih govornika engleskog kojima je maternji jezik rumunski, a koji bi mogli da se dobiju daljim istraživanjima.
3. Dinca & Chitez, 2021.	ESP-IT (63.842 reči) i ESP-LIT (67.529 reči)	Studentski eseji – učenički korpus; domen <i>Informatika i književnost</i>	četiri i pet reči	Minimalna frekvencija od 0.25 ponavljanja u 10.000 reči i minimalna disperzija od dva teksta; alat: <i>Lancsbox</i>	Identifikacija zajedničkih ULS u oba korpusa; podela ULS u sopstvene funkcionalne kategorije; provera upotrebe ULS specifičnih za naučne radove izvornih govornika u učeničkoj produkciji.	U istraživanju se ne koristi standardna taksonomija za funkcionalnu klasifikaciju ULS – ne nudi informacije koje bi mogle da se koriste za spekulaciju o mogućim trendovima u korišćenju ULS kod izvornih govornika rumunskog jezika.
4. Hyland & Jiang, 2022.	SciELF (759.300 reči) i akademski deo BNC (779.135 reči)	Naučni radovi – četiri naučne oblasti	četiri reči	Minimalna frekvencija od 20 ponavljanja u milion reči i minimalna disperzija u 10% tekstova; alat: <i>AntGram</i>	Identifikacija ULS u oba korpusa, kao i za svaki od jezika koji sačinjavaju SciELF. Poređenje deset najfrekventnijih ULS za sve jezike; strukturalna i funkcionalna raspodela svih ULS ukupno, za sve jezike i oblasti.	Indikacije o: razlikama u upotrebi ULS u finalnim verzijama radova neizvornih govornika i objavljenih radova; razlikama u upotrebi ULS kod različitih grupa neizvornih govornika; karakteristike upotrebe ULS kod neizvornih govornika kojima je maternji jezik rumunski.
5. Bercuci, 2020.	PSE (532.702 reči), PSRo (268.671) i PSRoE (344.787)	Naučni radovi – iz oblasti političkih nauka	tri, četiri i pet reči	Minimalna frekvencija i disperzija nisu navedeni; alat: <i>Lancsbox</i>	Poređenje 20 najfrekventnijih ULS u sva tri korpusa. Funkcionalna raspodela ULS u PSE i PSRoE.	Informacije o tendencijama upotrebe ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski u oblasti političkih nauka – fokus na funkcionalnoj raspodeli.
6. Bercuci, 2021.	PSE, PSRo, PSRoE i PSL (146.066 reči)	Studentska produkcija – iz oblasti političkih nauka	tri, četiri i pet reči	Minimalna frekvencija i disperzija nisu navedeni; alat: <i>Lancsbox</i>	Izdvajanje 20 najfrekventnijih ULS u PSL korpusu i poređenje sa rezultatima iz prethodnog istraživanja; funkcionalna analiza ULS u PSL korpusu.	Zaključci o karakteristikama upotrebe ULS kod studenata kojima je maternji jezik rumunski sa fokusom na funkcionalne aspekte; primeri češće upotrebe određenih ULS koji su prevodni ekvivalenti fraza na rumunskom jeziku.

Istraživanje Đorđević i Lazić (2021) se takođe nadovezuje na istraživanja Lazić (2017), ali i na istraživanje Lazić (2019) tako što koristi izdvojene ULS iz korpusa CoBNEA za formiranje materijala za podučavanje ULS na master studijama programa Poljoprivreda na Poljoprivrednom fakultetu Univerziteta u Beogradu, a primenjuje tipove vežbi za usvajanje ULS iz istraživanja Lazić (2019). Istraživanje Đorđević i Lazić (2021) ima dva cilja: 1. da utvrdi da li postoji korelacija između studentskog predznanja engleskog jezika i pravilne upotrebe ULS; 2. da utvrdi da li nakon obrađene nastavne jedinice posvećene ULS studenti mogu pravilno da ih upotrebe. Studenti su pohađali kurs engleskog jezika sa fondom od pet časova nedeljno. Tokom časova imali su priliku da se upoznaju sa značajem pravilne upotrebe ULS u akademskom diskursu i da vežbaju njihovu upotrebu kroz zadatke prepoznavanja i obeležavanja ULS u naučnim radovima, određivanje funkcije obeleženih ULS u datim kontekstima, kao i pisanje pojedinih delova naučnih radova, uzimajući u obzir pravilnu upotrebu ULS. Pre početka kursa studenti su uradili ulazni test i po nivou znanja jezika podeljeni su u tri grupe: I (A1/A2), II (B1) i III (B2, C1). Na završetku kursa, 58 studenata je uradilo test pravilne upotrebe ULS, koji se sastojao od 15 rečenica u kojima je trebalo da studenti od ponuđenih odaberu odgovarajući ULS. Rezultati istraživanja pokazali su da su se rezultati testa kretali od najnižeg rezultata sa četiri tačna odgovora do najvišeg rezultata sa svih 15 tačnih odgovora. Autorke su ove podatke interpretirale kao potvrdu „pojedinačnog poznavanja ciljnih ULS posle obrađene nastavne jedinice posvećene ovim spojevima, jer nije bilo onih koji nisu pokazali poznavanje nijednog od ispitanih ULS” (Đorđević i Lazić, 2021: 201), odnosno kao potvrđan odgovor na prvo pitanje. Moglo bi se reći da je kriterijum za uspešno savladavanje ULS poprilično nizak ako uzmemo u obzir da je na fakultetima neophodno poznavanje barem 51% gradiva za najnižu prolaznu ocenu, a da su četiri tačna odgovora od 15 samo 27%. Studenti su najuspešnije prepoznali pravilnu upotrebu ULS *at a rate of* (95% studenata je pravilno upotrebilo ULS) i *the results of the* (90%), dok su najmanje uspešno prepoznali upotrebu ULS *as well as the* (33%) i *are likely to be* (33%). Istraživanje Đorđević i Lazić je značajno, jer nam pruža prve uvide u efikasnost podučavanja ULS kod učenika kojima je maternji jezik srpski.

Na osnovu podataka iz ovog poglavlja možemo zaključiti da su istraživanja ULS u akademskom engleskom jeziku kod izvornih govornika rumunskog i srpskog tek na pomolu i da postoji potreba za daljim istraživanjem u ovoj oblasti.

2.3. Istraživanja upotrebe ULS u naučnim radovima kod izvornih i neizvornih govornika

Manje su brojna istraživanja koja se bave upotrebom ULS u naučnim radovima, odnosno u produkciji profesionalaca, a kada se uporede sa istraživanjima koja su se bavila njihovom upotrebom kod studenata/učenika. Moglo bi se pretpostaviti da je razlog za ovo činjenica da je teže ustanoviti sa sigurnošću koji je maternji jezik kod velikog broja istraživača, što je neophodan korak pri formiranju korpusa u ovom tipu istraživanja. Takođe, laka dostupnost velikom broju studenata i učenika i njihovoj jezičkoj produkciji može da objasni zašto su ta istraživanja brojnija. Upotreba ULS u naučnim radovima kod izvornih i neizvornih govornika postala je tema istraživanja u poslednjoj deceniji, iako su istraživanja koja se bave ovom temom i dalje malobrojna.

U ovom poglavlju detaljnije ćemo sagledati prethodna istraživanja koja su se bavila gorenavedenom temom. Prvo će se predstaviti relevantna istraživanja *upotrebe ULS u naučnim radovima kod izvornih i neizvornih govornika koja imaju drugačije postavke i/ili ciljeve od ove disertacije, a zatim istraživanja sa sličnom postavkom i ciljevima* čiji će rezultati biti relevantniji za kasnije poređenje. Nakon opštih podataka o istraživanjima, detaljnije će se pogledati veličine korpusa u istraživanjima, odabir dužine ULS koji će se istraživati, razlike u broju ULS kod izvornih i neizvornih govornika, broj ULS koji su zajednički za izvorne i neizvorne govornike, kao i razlike u strukturalnoj i funkcionalnoj raspodeli.

Salazarova (2010)

Istraživanje koje je uradila Danica Salazar (Salazar, 2010) je jedno od prvih istraživanja koje se bavilo temom upotrebe ULS u naučnim radovima kod izvornih i neizvornih govornika. Dotično istraživanje je specifično zato što zauzima dve perspektive u pristupu ULS: žanrovsku i engleski jezici sveta (engl. *World Englishes*). Njeno istraživanje poredi upotrebu ULS u korpusu naučnih radova iz oblasti medicine. Dva manja korpusa su kreirana za potrebe ovog istraživanja – jedan koji se sastoji od naučnih radova iz britanskog naučnog časopisa *British Medical Journal* (140.844 reči) i drugi, koji se sastoji od naučnih radova koji su objavljeni u filipinskom naučnom časopisu *Philippine Journal of Microbiology and Infectious Diseases* (123.088 reči). Nisu date dodatne informacije o određivanju maternjeg jezika autora članaka. Ako je to bio jedini kriterijum, može se smatrati da je velika verovatnoća da su u korpusu postojali i tekstovi autora sa različitim maternjim jezikom od pretpostavljenog jer je standardno za naučne časopise da objavljuju i radove istraživača van svoje regionalne oblasti. U istraživanju je navedeno da su izdvojeni najfrekventniji ULS, međutim, nije navedeno koja je minimalna frekvencija i distribucija korišćena za njihovo izdvajanje.

Istraživanje Salazarove (2010) se bavilo ULS koji se sastoje od tri reči i analizirani su samo ULS koji u sebi sadrže leksičke glagole (engl. *lexical verbs*). Leksički glagoli su glagoli sa semantičkim značenjem. Oni su izabrani umesto imenica zato što su manje zavisni od teme pa samim time, nisu sadržali terminologiju koja je specifična za datu temu i koja se izdvaja iz ove vrste rezultata. Salazarova ispituje upotrebu ULS u naučnim radovima iz oblasti biomedicine kod izvornih govornika engleskog i neizvornih govornika engleskog kojima je maternji jezik filipinski. Ukupan broj svih identifikovanih ULS od tri reči bio je 592 tipa u korpusu izvornih govornika i 439 tipova u korpusu neizvornih govornika. Ukupan broj identifikovanih ULS sa leksičkim glagolom u korpusu izvornih govornika je 70, dok je kod neizvornih govornika 68. Dotično istraživanje nema urađene testove za određivanje značaja rezultata. Samo dva ULS su identifikovana u oba korpusa, iako je analiza leksičkih glagola u ULS pokazala da je polovina glagola, koji su se našli u korpusu neizvornih govornika, takođe, bila prisutna u glagolskim ULS kod izvornih govornika.

Za funkcionalnu raspodelu ULS korišćena je taksonomija razvijena od strane Hajlanda (Hyland, 2008). Međutim, Salazarova je dodala dve podgrupe: *ULS cilja* i *ULS priznanja* (engl. *objective and acknowledgement bundles*), dok se tri podgrupe koje su određene u originalnoj taksonomiji nisu pojavile u dotičnom istraživanju (*ULS za lokaciju, ULS za kvantifikaciju i tranzicione signale* – engl. *location, quantification and transition signals*). Rezultati su pokazali da u korpusu neizvornih govornika ima više *ULS orijentisanih ka istraživanju*, dok u korpusu izvornih govornika ima više *ULS orijentisanih ka učesniku*. Zastupljenost *ULS orijentisanih ka tekstu* je ujednačena u dva korpusa. *ULS za proceduru i rezultativni signali* su dve najzastupljenije podgrupe u oba korpusa. Interesantna je činjenica da u okviru korpusa izvornih govornika postoje ULS sa ličnom zamenicom prvog lica, dok takvi ULS uopšte nisu izdvojeni u korpusu neizvornih govornika. Pošto su se detaljno analizirali samo ULS sa leksičkim glagolom, strukturalna analiza izdvojenih ULS nije urađena. Pregled najznačajnijih podataka o istraživanju Salazarove (2010) može se naći u Tabeli 10.

Tabela 10. Osnovni podaci o istraživanju Salazarova (2010)

Istražena naučna oblast	Biomedicina	
Maternji jezik izvornih govornika engleskog jezika	Filipinski jezik	
Ispitana dužina ULS	3 reči	
Veličina korpusa	<i>Izvornih govornika</i>	<i>Neizvornih govornika</i>
	140.844	123.088
Primenjena minimalna frekvencija	Nisu dati podaci.	
Primenjena disperzija	Nisu dati podaci.	
Broj izdvojenih ULS	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
	592	439

Salazarova (2014)

Salazarova je 2014. godine uradila još jedno istraživanje na temu upotrebe ULS u naučnim radovima kod izvornih i neizvornih govornika. U dotičnom istraživanju bavi se upotrebom ULS u naučnim radovima iz oblasti biomedicine kod izvornih govornika engleskog i neizvornih govornika engleskog kojima je maternji jezik španski. Glavni cilj ovog istraživanja bio je ispitati upotrebu ULS u naučnim radovima u oblasti biomedicine, kao i kreirati spisak ULS koji će biti korišćen u pedagoške svrhe. Poređenje upotrebe ULS kod izvornih i neizvornih govornika je, čini se, bio sekundaran cilj, što se može zaključiti po razlikama u veličini dva korpusa i u specifičnoj metodološkoj postavci. U istraživanju su bila korišćena dva korpusa. Prvi je korpus naučnih radova iz oblasti biotehničkih nauka koje su napisali izvorni govornici engleskog. On je zapravo izdvojen segment većeg korpusa pod nazivom *Health Science Corpus* (HSC). Segment korišćen u istraživanju Salazarove (2014) sastoji se od 290 naučnih radova i ukupno 2.082.409 reči, a u njemu se obrađuju oblasti biomedicine (biohemija, molekularna biologija, molekularna genetika, genetika i biologija ćelije). Radove su pisali autori sa anglofonim imenima i prezimenima koji rade u institucijama u zemljama gde je engleski jezik zvanični. Drugi korpus je bio korpus naučnih radova iz oblasti biologije od strane neizvornih govornika engleskog kojima je izvorni jezik španski. Korpus je sačinjen od 43 naučna rada i ukupno 120.718 reči. Radovi su nelektorisane verzije radova koji su dobijeni od lektora sa univerziteta u Rio Kvartu (šp. *Rio Cuarto*), a koji je imao zadatak da ih lektoriše pre nego što su bili objavljeni u naučnim časopisima. U korpusu izvornih govornika izdvojeni su ULS od tri do šest reči koristeći minimalnu frekvenciju od deset ponavljanja u milion reči i minimalnu disperziju od obaveznog pojavljivanja u pet različitih tekstova. Pri identifikaciji ULS takođe se koristio i *Mutual Information score* (MI score) sa postavljenim minimumom od 0.5.

Iako primena *MI score* nije tipična procedura u ovoj vrsti istraživanja, autori su se odlučili za takav korak u uverenju da će dobiti podatke koji će biti korisniji za primenu u pedagoške svrhe. Na navedeni način, izdvojeno je ukupno 1.737 ULS. Kako bi spisak bio koncizniji i pedagoški upotrebljiviji, autori su primenili nekoliko kriterijuma za kvalitativno prečišćavanje spiska. Sa spiska su uklonjeni ULS koji su: 1. sačinjeni isključivo od funkcionalnih reči; 2. specifični za određenu temu; 3. besmisleni; 4. čine nasumične naslove; 5. uključuju nasumične brojeve; 6. iskazuju vreme; 7. sačinjavaju duži ULS; 8. završavaju sa članom. Prečišćeni spisak se sastoji od 769 ULS, od kojih je 640 (83%) ULS od tri reči, 113 (15%) ULS od četiri reči, 11 ULS od pet reči i pet ULS od šest reči koji zajedno sačinjavaju oko 1% od ukupnog broja izdvojenih ULS u konačnom spisku.

Autorka istraživanja ukazuje na činjenicu da bi direktna kvantitativna analiza ULS u datim korpusima izvornih i neizvornih govornika bila nepouzdana zbog velike razlike u veličini korpusa i zbog toga preuzima metodološko rešenje koje je primenila i Kortez (Cortes, 2004). Naime, ULS izdvojeni iz većeg korpusa izvornih govornika se smatraju ciljnim ULS (engl. *target bundles*), a njihovo prisustvo i zastupljenost se proveravaju u manjem korpusu. Pri kvalitativnoj analizi ciljnih ULS autorka primenjuje analizu prototipa i ključnih reči. Kanonski oblik koji su predložili Sinkler i saradnici (Sinclair, Jones & Daley, 2004, prema Salazar, 2014) je primenjen u istraživanju kako bi se istražile semantičke i strukturalne veze između izdvojenih ULS. Međutim, kao što je gorenavedeno u istraživanju Salazarove (Salazar, 2014), kanonski oblici se nazivaju *prototipski ULS* (engl. *prototypical bundles*). Čini se da je upotreba ključnih reči korisnija kod analize ULS od tri reči, ali može da se koristi i za duže spojeve. Ključne reči u ovom kontekstu podrazumevaju reč koja nosi značenje čitavog spoja. Radi ilustracije, dajemo jedan primer izdvojen iz rada u kome se vidi kombinacija primene i ključne reči i prototipskih ULS (Salazar, 2014: 58): Prototipski ULS je *shown in table*; ključna reč: *table*; povezani ULS: *are shown in, is shown in, summarized in table, are summarized in, are shown in table, in table 1, in table 2, in table 3*; moguće kombinacije: [*is, are*][*described, given, listed, presented, shown, summarized*]*in table* [1, 2, 3...].

U korpusu neizvornih govornika od ukupno 442 prototipska ULS identifikovano je 312. Međutim, 92 identifikovana prototipska ULS se pojavljuju u korpusu neizvornih govornika samo jednom. Ovo ukazuje na činjenicu da se više od pola prototipskih ULS ne pojavljuje ili pojavljuje samo jednom u korpusu neizvornih govornika (Salazar, 2014). Navedeni rezultati nisu iznenađujući ako se uzmu u obzir prethodna istraživanja koja ukazuju na razliku u upotrebi ULS kod izvornih i neizvornih govornika, odnosno da se različiti tipovi identifikuju u dve grupe govornika kada se primenjuje isti tip kvalitativne analize i parametri za izdvajanje. Pošto ULS specifični za korpus neizvornih govornika nisu identifikovani, logično je da će biti manji broj prototipskih koji su karakteristični za korpus izvornih govornika.

Za strukturalnu analizu korišćena je taksonomija koju su predložili Bajber i saradnici (Biber et al., 1999). Međutim, Salazarova je dodala pet potkategorija: *druge imeničke fraze, druge predloške fraze, glagolske fraze sa ličnom zamenicom we, drugi pasivni fragmenti i drugi glagolski fragmenti*. U korpusu izvornih govornika najzastupljeniji ULS iz strukturalne grupe *imeničke strukture* (30%), praćeni *glagolskim strukturama* (28%), dok su manje zastupljene kategorije *druge strukture* (23%) i *fragmenti sa predloškom frazom* (19%). U korpusu neizvornih govornika raspodela ULS po glavnim strukturalnim kategorijama je ista, odnosno najzastupljenije su *imeničke strukture* (34%), praćene *glagolskim strukturama* (27%), dok su manje zastupljene kategorije *druge strukture* (20%) i *fragmenti sa predloškom frazom* (19%). Za funkcionalnu klasifikaciju korišćena je taksonomija koju je predložio Hajland (2008). Međutim, u potkategorijama, Salazarova je unela sledeće izmene: u grupi *ULS orijentisanih ka istraživanju* dodata je podgrupa *ULS za grupisanje* (engl. *grouping*), u grupi *ULS orijentisanih ka tekstu* izbačena je podgrupa *rezultativni signali* a dodate podgrupe *ULS dodavanja, ULS komparacije, inferentni ULS, kauzativni ULS, ULS za citiranje, ULS za*

generalizaciju i ciljne ULS (engl. *additive, comparative, inferential, causative, citation, generalization* i *objective*), u grupi *ULS orijentisanih ka učesniku* dodata je podgrupa *ULS uvažavanja* (engl. *acknowledgement*). Funkcionalna analiza ULS u korpusu izvornih govornika pokazala je da je najzastupljenija kategorija *ULS orijentisani ka tekstu* (48%) koja je blisko praćena kategorijom *ULS orijentisanih ka istraživanju* (43%), dok je treća kategorija *ULS orijentisanih ka učesniku* manje zastupljena (9%). U korpusu neizvornih govornika situacija je vrlo slična što se tiče funkcionalne raspodele u glavnim kategorijama, naime, *ULS orijentisani ka tekstu* su najzastupljeniji sa 47%, *ULS orijentisani ka istraživanju* su drugi sa 44%, dok su *ULS orijentisani ka učesniku* na trećem mestu sa samo 9% tipova.

Salazarova upoređuje svoje nalaze sa rezultatima Hajlandovog istraživanja (2008) zbog uticajnosti navedenog rada i činjenice da se jedan deo istraživanja bavi analizom iste naučne oblasti. Sličnosti u spiskovima ULS u dva istraživanja Salazarova opisuje kao dokaz validnosti rezultata u dva istraživanja. Ona, takođe, ističe nedostatke upotrebe ciljnih ULS pri analizi korpusa neizvornih govornika. Naime, najveći je nedostatak činjenica da se ne identifikuju ULS karakteristični za korpus neizvornih govornika pa se ne zna koliko je ULS zastupljeno u datom korpusu i koje su im funkcionalne i strukturalne karakteristike; zna se samo koliko se ULS karakterističnih za korpus izvornih govornika nalazi u korpusu neizvornih i koje su njihove karakteristike u datom korpusu. Iz ovog razloga podaci iz istraživanja koje je uradila Salazar (2014) nisu pogodni za direktno poređenje sa rezultatima u ovoj disertaciji, iako je samo istraživanje značajan doprinos oblasti istraživanja, pogotovo u vremenu u kojem je objavljeno. Pregled osnovnih podataka o istraživanju Salazarove (2014) mogu se naći u Tabeli 11.

Istražena naučna oblast	Biomedicina	
Maternji jezik neizvornih govornika engleskog jezika	Španski jezik	
Ispitana dužina ULS	3 do 6 reči	
Veličina korpusa	<i>Izvornih govornika</i>	<i>Neizvornih govornika</i>
	2.082.409	120.718
Primenjena minimalna frekvencija	10 ponavljanja u milion reči	
Primenjena disperzija	5 tekstova	

Perez-Ljantada (2014)

Istraživanje koje je uradila Perez-Ljantada (Pérez-Llantada, 2014) je unikatno zato što se fokusira na divergentne i konvergentne leksičke spojeve i na taj način ima cilj da pruži nove uvide o upotrebi ULS kod izvornih i neizvornih govornika. Konvergentni spojevi su oni koji se nalaze u tri ili dva korpusa, drugim rečima, ULS koji su zajednički za korpus, dok su divergentni oni koji su specifični za određeni korpus. Neizvorni govornici engleskog jezika u ovom istraživanju su izvorni govornici španskog jezika. Autorka poredi upotrebu ULS u tri korpusa: korpus naučnih radova izvornih govornika engleskog jezika (sastoji se od 360 naučnih radova dok je veličina korpusa 2.146.347 reči), korpusa naučnih radova na engleskom jeziku izvornih govornika španskog (sastoji se od 336 radova a veličine korpusa je 1.771.727 reči), korpus naučnih radova na španskom izvornih govornika španskog jezika (sastoji se od 360 naučnih radova i veličine je 1.811.071 reči). Naučni radovi koji sačinjavaju tri navedena korpusa su iz 12 naučnih oblasti: primenjena lingvistika, informatika, književnost, sociologija, poslovni menadžment, geografija, urologija, hematologija, onkologija, mehanički inženjering, prehrambena tehnologija i geonauke. U okviru tri korpusa ujednačena je raspodela radova iz svake discipline. Veličina dva korpusa neizvornih

govornika je uglavnom ujednačena, dok je malo veća razlika između korpusa izvornih i neizvornih govornika (oko 350.000 reči).

Autorka navodi da radovi koje su napisali neizvorni govornici engleskog jezika nisu prevodi već da su pisani na engleskom jeziku, ali ne navodi kako je došla do ovih informacija i, takođe, ne navodi koja je radna definicija ili metoda korišćena za utvrđivanje maternjeg jezika pisaca radova koji sačinjavaju korpus. Štaviše, nije navedeno zašto je broj radova za dva korpusa ujednačen a u trećem je manji.

U istraživanju analizirani su ULS od četiri reči čija se identifikacija oslanjala na minimalnu frekvenciju od 20 ponavljanja u milion reči i minimalnu distribuciju u 10% tekstova. U ovom istraživanju ULS koji se preklapaju nisu sjedinjeni u jedan duži ULS, već su uzeti u obzir zasebno. Ukupno 56 ULS je izdvojeno u korpusu izvornih govornika, 77 ULS u korpusu neizvornih govornika na engleskom i 114 u korpusu izvornih govornika španskog na španskom jeziku. U istraživanju nisu dati podaci o strukturalnoj i funkcionalnoj raspodeli svih ULS u svakom korpusu ponaosob već su ove informacije date za konvergentne a zatim za divergentne ULS u ovim korpusima.

Autorka primenjuje strukturalnu taksonomiju koja se koristi u istraživanju Bajber (2009, prema Pérez-Llantada, 2014). Potpuni strukturalni oblici su obično ULS sa predloškim elementom i sačinjavaju oko 5% ULS u svakom od korpusa. U sva tri korpusa spojevi sa predloškim fragmentom su brojniji od klauzalnih koji sačinjavaju 60% tipova u korpusu izvornih govornika engleskog, 71% neizvornih govornika engleskog i 73% u korpusu izvornih govornika španskog na španskom.

Kada je reč o konvergentnim spojevima, 12 je pronađeno u sva tri korpusa dok korpusi radova na engleskom dele 36 ULS. Test verodostojnosti (engl. *log-likelihood test*) nije pokazao značajne razlike u korišćenju ovih spojeva u različitim korpusima. Korpusi naučnih radova izvornih govornika španskog na engleskom i španskom dele 13 ULS koji čine 17% ukupne produkcije ULS u korpusu radova na engleskom od strane neizvornih govornika, što ukazuje na to da je uticaj maternjeg jezika mali.

Identifikovano je 20 ULS koji su prisutni samo u korpusu izvornih govornika engleskog (36% ULS u tom korpusu), 28 ULS koji su prisutni samo u korpusu neizvornih govornika engleskog (36% ULS u tom korpusu) i 83 ULS u španskom korpusu izvornih govornika španskog (36% ULS u tom korpusu). Funkcionalna analiza ULS, koji se pojavljuju u samo jednom korpusu, pokazuje da oni prenose sasvim drugačija interpersonalna pragmatična značenja. Skoro pola ULS specifičnih za korpus izvornih govornika engleskog spadaju u kategoriju *ULS stava*. ULS u ovoj kategoriji su retko prisutni kod neizvornih govornika engleskog. ULS prisutni samo u korpusu neizvornih govornika engleskog uglavnom su u podgrupi *inferentni ULS*. Interesantno je da se pridevi koji su se kod izvornih govornika nalazili u *ULS stava* kod neizvornih govornika koristili u *ULS za identifikaciju/fokus* što je činilo njihov stil nezgrapnim. Osnovni podaci o istraživanju koje je sprovedla Perez-Llantada mogu se naći u Tabeli 12.

Tabela 12. Osnovni podaci o istraživanju Perez-Ljatanda (2014)

Istražena naučna oblast	12 naučnih oblasti	
Maternji jezik neizvornih govornika engleskog jezika	Španski jezik	
Ispitana dužina ULS	4 reči	
Veličina korpusa	<i>Izvornih govornika</i>	<i>Neizvornih govornika</i>
	2.146.347	1.771.727
Primenjena minimalna frekvencija	20 ponavljanja u milion reči	
Primenjena disperzija	10% tekstova	

Lazić (2017)

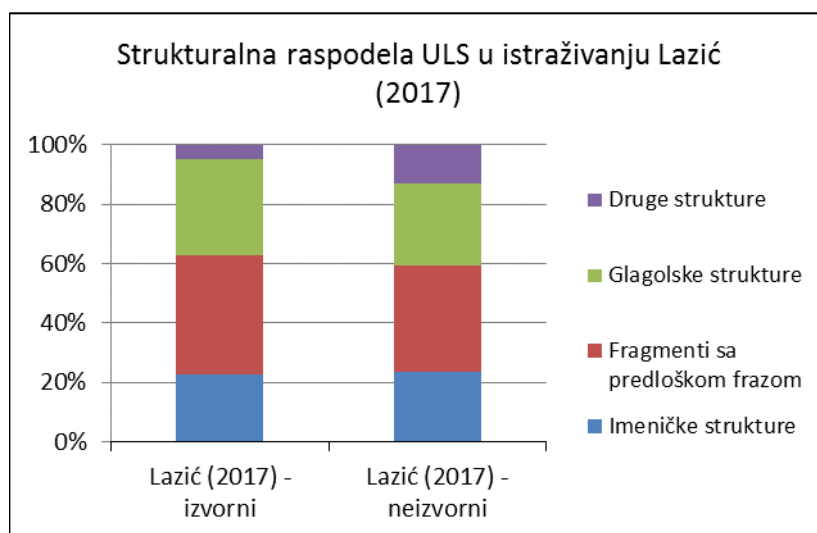
Istraživanje koje je uradila Katarina Lazić (2017) od velikog je značaja zato što je jedino prethodno istraživanje koje se bavilo upotrebom ULS kod neizvornih govornika engleskog kojima je maternji jezik srpski. Glavni cilj istraživanja, odnosno doktorske disertacije, bila je analiza upotrebe ULS od četiri do šest reči u naučnim radovima iz oblasti biotehničkih nauka na engleskom jeziku. U okviru istraživanja upoređena je upotreba ULS u radovima u pomenutoj naučnoj oblasti od strane izvornih govornika engleskog i neizvornih govornika engleskog kojima je maternji jezik srpski. Urađena je i analiza ULS u naučnim radovima iz naučne oblasti biotehničkih nauka na srpskom jeziku kako bi se detaljnije utvrdio uticaj maternjeg jezika na produkciju neizvornih govornika engleskog. Za potrebe istraživanja izrađena su tri korpusa: CoBNEA (*Corpus of Biotechnical Native English Articles*) korpus izvornih govornika engleskog koji se sastoji od 246 radova i ukupno 1.525.469 reči; CoBNONEA (*Corpus of Biotechnical Non-Native English Articles*) neizvornih govornika engleskog čiji je maternji jezik srpski, koji je sačinjen od 50 radova i ukupno 157.179 reči; CoBSA (*Corpus of Biotechnical Serbian Articles*) korpus radova iz oblasti biotehničkih nauka na srpskom jeziku, koji je sačinjen od 50 radova i ukupno 126.275 reči. Zbog toga što poređenje izvornih i neizvornih govornika nije bio primarni cilj datog istraživanja i zbog specifičnih parametara postavljenih za građenje korpusa neizvornih govornika engleskog jezika kojima je maternji srpski, korpus neizvornih govornika koji se koristi u datom istraživanju skoro je deset puta manji od korpusa izvornih govornika. Za formiranje korpusa CoBNONEA autorka disertacije je koristila nelektorisane verzije radova kako bi dobila što autentičniji materijal za korpus. Međutim, ovaj stepen rigoroznosti pri formiranju korpusa je ograničio veličinu istog.

Za određivanje ULS korišćena je minimalna frekvencija od 20 ponavljanja u milion reči. Međutim, u kvalitativnoj analizi podataka analizirani su ULS koji su izdvojeni manjom normalizovanom minimalnom frekvencijom koja odgovara apsolutnoj minimalnoj frekvenciji od 20 ponavljanja u čitavom korpusu. To je učinjeno da bi se ispitao pedagoški potencijal ULS u istraženim korpusima. Minimalna disperzija je četiri teksta za sva tri korpusa. U korpusu izvornih govornika izdvojeno je 69 tipova sa strožim kriterijumima izdvajanja i 195 sa blažim, dok su za druga dva korpusa primenjeni samo stroži kriterijumi. U korpusu neizvornih govornika na engleskom je izdvojeno 507 tipova, dok je u korpusu na srpskom jeziku izdvojeno 309 tipova. Ovakva razlika u broju identifikovanih ULS može se pripisati tendenciji da neizvorni govornici koriste veći broj ULS, ali i činjenici da manji korpusi generišu veći broj ULS (Bestgen, 2019). Zbog razlike u veličinama korpusa, nisu se poredili brojevi ULS koji su nađeni u korpusima, već samo procenti strukturalne i funkcionalne raspodele.

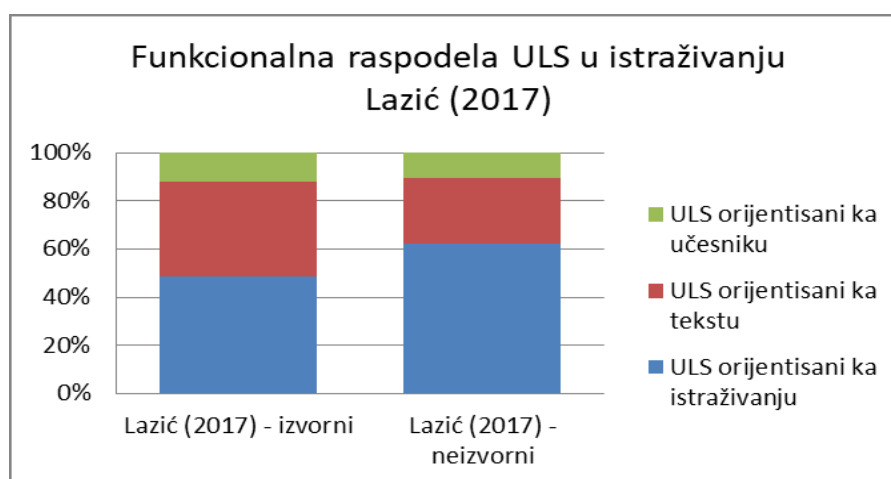
Kada je reč o strukturalnoj raspodeli ULS, rezultati su pokazali da su u oba korpusa na engleskom najzastupljeniji *fragmenti sa predložkom frazom*, praćeni *glagolskim strukturama*, pa *imeničkim strukturama*. Veće su razlike u strukturalnoj raspodeli ULS kada se posmatraju potkategorije nego glavne kategorije. Statistički testovi urađeni su za strukturalnu raspodelu ULS na glavne kategorije

ali nije urađeno za potkategorije. U Tabeli 13 podebljanim tekstem su označeni procenti gde su identifikovane statistički značajne razlike. Podaci o funkcionalnoj raspodeli ULS daju podatke da su *ULS orijentisani ka istraživanju* najzastupljeniji, praćeni *ULS orijentisanim ka tekstu*, dok su *ULS orijentisani ka učesniku* najmanje zastupljeni. Ova raspodela je zastupljena u oba korpusa, dok su neizvorni govornici koristili više *ULS orijentisanih ka istraživanju* od izvornih govornika, a manje *ULS orijentisanih ka tekstu*. Kao i kod strukturalne raspodele, veće razlike su uočene u raspodeli po potkategorijama. Statistički testovi urađeni su za funkcionalnu raspodelu ULS na glavne grupe ali nije urađeno sa podgrupe. U Tabeli 13 podebljanim tekstem su označeni procenti gde su identifikovane statistički značajne razlike. (Pregled najznačajnijih podataka o istraživanju koji je sproveda Lazić (2017) mogu se naći u Tabeli 13, kao i na Grafikonu 7 i Grafikonu 8.)

Tabela 13. Osnovni podaci o istraživanju Lazić (2017)			
Istražena naučna oblast	Biotehničke nauke		
Maternji jezik neizvornih govornika engleskog jezika	Srpski jezik		
Ispitana dužina ULS	4 reči		
Veličina korpusa	<i>Izvornih govornika</i>	<i>Neizvornih govornika</i>	
	1.525.469	126.275	
Primenjena minimalna frekvencija	20 ponavljanja u milion reči		
Primenjena disperzija	4 teksta		
Broj izdvojenih ULS	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>	
	69	507	
Korišćena taksonomija za strukturalnu analizu podataka	Modifikovana taksonomija koju su predložili Bajber i sar. (1999)		
Strukturalna analiza tipova i tokena	<i>Izvorni govornici</i>	<i>Neizvorni govornici</i>	
	<i>Imeničke strukture</i>	22.5 % (20.4%)	23.6% (14.8%)
	<i>Fragmenti sa predloškom frazom</i>	40.4% (40.8%)	35.7% (34.7%)
	<i>Glagolske strukture</i>	32.3% (31.7%)	27.6% (18.1%)
	<i>Druge strukture</i>	4.8% (7.1%)	13.1% (32.4%)
Korišćena taksonomija za funkcionalnu analizu	Modifikovana verzija taksonomije koju je predložio Hajland (2008)		
Funkcionalna klasifikacija tipova	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>	
	<i>ULS orijentisani ka istraživanju</i>	48.3% (48.9%)	62.1% (57.5%)
	<i>ULS orijentisani ka tekstu</i>	39.6% (42.5%)	27.5% (31.8%)
	<i>ULS orijentisani ka učesniku</i>	12.1% (8.6%)	10.4% (10.7%)



Grafikon 7. *Strukturalna raspodela ULS u istraživanju Lazić (2017)*

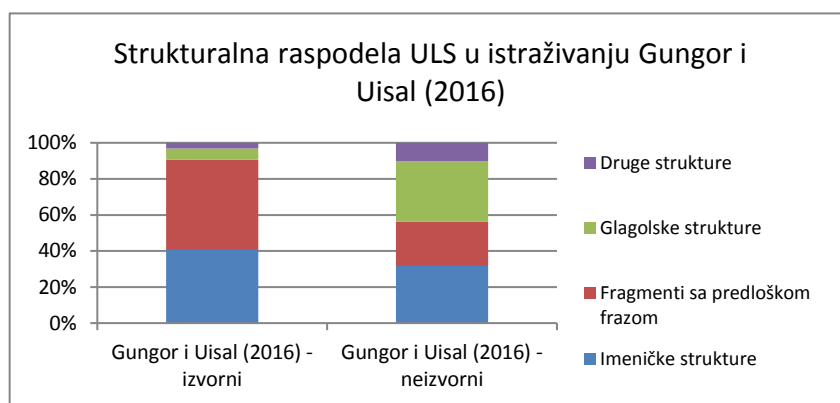


Grafikon 8. *Funkcionalna raspodela ULS u istraživanju Lazić (2017)*

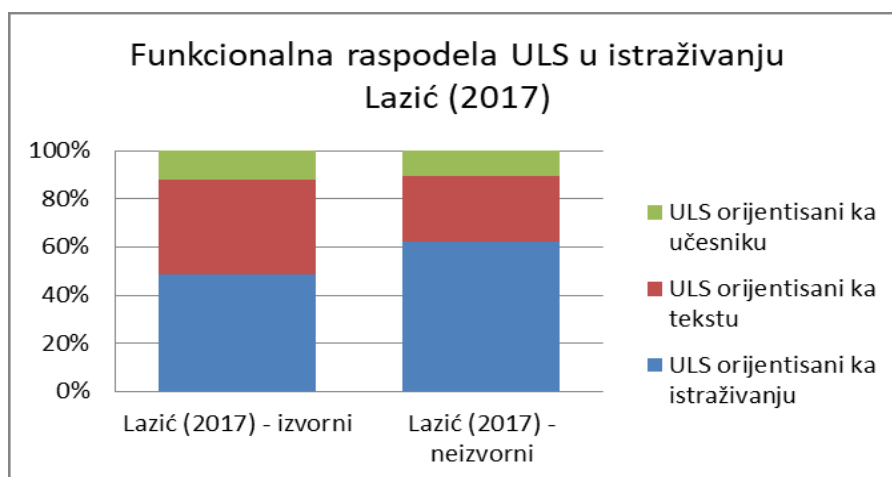
Gungor i Uisal (2016)

Istraživanje koje su uradili Gungor i Uisal (Güngör & Uysal, 2016) od velikog je značaja zato što se bavi analizom upotrebe ULS u radovima iz oblasti obrazovanja kod izvornih i neizvornih govornika, odnosno u istoj naučnoj oblasti kao i ova disertacija. Štaviše, veličine dva korpusa su ujednačene i iste veličine kao što su korpusi u ovoj disertaciji. Maternji jezik neizvornih govornika u istraživanju je turski jezik. Proučava se upotreba ULS od četiri reči, a za njihovo izdvajanje primenjuje se minimalna frekvencija od 20 ponavljanja u milion reči i minimalna distribucija u 10% tekstova. Za potrebe istraživanja formirana su dva korpusa od po pola miliona reči. Prvi je korpus sačinjen od naučnih časopisa iz oblasti obrazovanja napisanih od strane izvornih govornika engleskog jezika, dok je drugi sačinjen od naučnih radova iz oblasti obrazovanja napisanih od strane izvornih govornika turskog. Za definisanje izvornih govornika u naučnim radovima korišćen je princip prisutan u istraživanju Salazarove (2014), odnosno izvornim govornicima ciljnog jezika smatrani su istraživači koji su u vezi sa institucijama u zemlji u kojoj je ciljani jezik zvanični i koji imaju imena i prezimena tipična za tu naciju. Korpus izvornih govornika engleskog sačinjen je od 79 naučnih radova, dok je korpus neizvornih govornika sačinjen od 101 teksta. Ovde se vidi tendencija da radovi koje pišu izvorni govornici engleskog budu duži od radova koje pišu neizvorni govornici engleskog, u ovom slučaju izvornih govornika turskog.

U korpusu izvornih govornika izdvojeno je ukupno 32 ULS od četiri reči, dok je u korpusu neizvornih govornika izdvojeno 98 ULS od četiri reči. Za strukturalnu raspodelu identifikovanih ULS korišćena je taksonomija koju su predložili Bajber i saradnici (1999), a za funkcionalnu taksonomija koju je predložio Hajland (2008). Strukturalna analiza je pokazala da su izvorni govornici koristili najviše *fragmenata sa predložkom frazom* (50%) koji su praćeni *imeničkim strukturama* (40,63%), dok su *glagolske strukture* (6,25%) i *drugi strukture* (3,12%) slabo zastupljene. U korpusu neizvornih govornika najzastupljeniji su ULS u grupi *glagolske strukture* (33,67%), blisko praćene *imeničkim strukturama* (31,63%), manje su zastupljeni *fragmenti sa predložkom frazom* (24,50%), dok su kao i kod izvornih govornika najmanje zastupljeni ULS iz kategorije *druge strukture* (10,20%). Funkcionalna analiza isto je pokazala razlike u primeni ULS kod izvornih i neizvornih govornika, naime, izvorni govornici najčešće su koristili *ULS orijentisane ka istraživanju* (68,8%), dok su neizvorni govornici najčešće koristili *ULS orijentisani ka tekstu* (66%). U datom istraživanju nije rađena raspodela ULS po strukturalnim i funkcionalnim podgrupama. Oba korpusa dele 13 ULS koji su uglavnom u strukturalnoj kategoriji *fragmenti sa predložkom frazom*. Iako su turski istraživači koristili više ULS, oni su nedovoljno koristili ULS koji su frekventni u korpusu izvornih govornika. Osnovni podaci o istraživanju dati su u Tabeli 14, kao i na Grafikonu 9 i Grafikonu 10.



Grafikon 9. Strukturalna raspodela ULS u istraživanju Gungor i Uisal (2016)



Grafikon 10. Funkcionalna raspodela ULS u istraživanju Gungor i Uisal (2016)

Tabela 14. Osnovni podaci o istraživanju Gungor i Uisal (2016)

Istražena naučna oblast	Obrazovanje	
Maternji jezik neizvornih govornika engleskog jezika	Turski jezik	
Ispitana dužina ULS	4 reči	
Veličina korpusa	<i>Izvornih govornika</i>	<i>Neizvornih govornika</i>
	500.327	500.012
Primenjena minimalna frekvencija	20 ponavljanja u milion reči	
Primenjena disperzija	10% tekstova	
Broj izdvojenih ULS	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
	32	98
Korišćena taksonomija za strukturalnu analizu podataka	Taksonomija koju su predložili Bajber i sar. (1999)	
Strukturalna analiza tipova	<i>Izvorni govornici</i>	<i>Neizvorni govornici</i>
<i>Imeničke strukture</i>	40.63%	31.63%
<i>Fragmenti sa predloškom frazom</i>	50%	24.50%
<i>Glagolske strukture</i>	6.25%	33.67%
<i>Druge strukture</i>	3.12%	10.2%
Korišćena taksonomija za funkcionalnu analizu	Taksonomije koju je predložio Hajland (2008)	
Funkcionalna klasifikacija tipova	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
<i>ULS orijentisani ka istraživanju</i>	68.80%	30.09%
<i>ULS orijentisani ka tekstu</i>	28%	66%
<i>ULS orijentisani ka učesniku</i>	3.10%	3.10%

Iste godine izašla je i doktorska disertacija autora Gungor (Güngör), koja se bavila temom upotrebe ULS u naučnim radovima izvornih i neizvornih govornika engleskog jezika. Glavni ciljevi datog istraživanja bili su da se odgovori na pitanja: 1. *Koji su ULS prisutni u naučnim radovima iz oblasti obrazovanja kod izvornih govornika engleskog?* 2. *Koji su ULS prisutni u naučnim radovima u oblasti obrazovanja na engleskom jeziku kod izvornih govornika turskog?* 3. *Koja je strukturalna i funkcionalna raspodela identifikovanih ULS?* 4. *Da li maternji jezik utiče na upotrebu ULS u međujeziku turskih izvornih govornika?* U Gungorovom istraživanju traže se prevodni ekvivalenti u korpusu napisanom na maternjem jeziku neizvornih govornika kako bi se objasnilo neslaganje u upotrebi ULS kod izvornih i neizvornih govornika, odnosno kako bi se objasnilo postojanje 130 (79%) ULS specifičnih za neizvorne govornike. Istraživanje u ovoj disertaciji isto ima cilj da objasni razlike u upotrebi ULS kod izvornih i neizvornih govornika, ali odgovore traži u komparaciji ULS kod različitih grupa neizvornih govornika, odnosno da se vidi da li postoje neke univerzalne karakteristike neizvornih govornika koje dovode do razlika u upotrebi nezavisno od maternjeg jezika.

Istraživanje Gungor (2016) primenjuje pristup vođen korpusom (engl. *corpus-driven approach*), stoga istraživač nije pravio hipoteze o mogućim rezultatima pre sprovođenja samog istraživanja. Za potrebe istraživanja formiran je *MCRA* korpus (*Multilingual Corpus of Research Articles*), koji se sastoji iz tri potkorpusa: *MCRA-L1* korpus istraživačkih članaka na engleskom jeziku iz oblasti obrazovanja napisanih od strane izvornih govornika engleskog, *MCRA-L2* korpus istraživačkih članaka na engleskom jeziku iz oblasti obrazovanja napisanih od strane neizvornih govornika i *MCRA-T1* korpus istraživačkih članaka na turskom iz oblasti obrazovanja napisanih od strane izvornih govornika turskog. Svaki od navedenih potkorpusa sadrži po milion reči, međutim, broj članaka za postizanje date veličine korpusa je različit. *MCRA-L1* sastoji se od 165 naučnih radova, *MCRA-L2* od 206, dok se *MCRA-T1* sastoji od 228 naučnih radova. Istraživanje se bavi analizom samo ULS od četiri reči. Za njihovo izdvajanje korišćena je minimalna frekvencija od 20 ponavljanja u milion reči i minimalna disperzija u 10% naučnih radova, odnosno tekstova. Pri filtriranju spiska izdvojenih ULS adaptirani su i primenjeni kriterijumi iz istraživanja Salazarove (2014). Finalni spisak izdvojenih ULS izvornih govornika sadrži 79 ULS i oni sačinjavaju 1,21% korpusa. Finalni spisak izdvojenih ULS kod neizvornih govornika sadrži 119 ULS, dok je u korpusu radova na turskom izdvojeno samo četiri ULS.

U datom istraživanju provereno je koliko postoji poklapanja ULS u datom istraživanju sa ULS identifikovanim u prethodnim istraživanjima. Upoređivanje je urađeno sa ciljem da se dodatno potkrepi validnost nalaza. Za poređenje izdvojenih ULS iz korpusa izvornih govornika korišćena su istraživanja Perez-Ljatanda (2014), Ozturk (2014) i Hajland (2008), a poređeno je samo 50 najfrekventnijih ULS u svim istraživanjima. Može se postaviti pitanje relevantnosti ovakvog poređenja s obzirom na to da su same postavke istraživanja poprilično drugačije – u istraživanju koji je uradila Perez-Ljatanda (2014) naučni radovi su izdvojeni iz 12 naučnih oblasti, istraživanje Ozturk (2014) bavi se radovima iz oblasti podučavanja stranog jezika, dok je Hajlandovo istraživanje (2008), iako od velikog značaja za istraživanje fenomena ULS, takođe značajno drugačije od datog istraživanja jer su naučni radovi iz četiri različite naučne oblasti koje ne uključuju obrazovanje. Pored gorenavedenih razlika, 31 ULS se nalazi u barem jednom od istraživanja. Isto je urađeno i sa ULS neizvornih govornika, samo je umesto poređenjem sa Hajlandovim istraživanjem (2008) iskorišćeno istraživanje koje je uradila Bal (2010). Rezultati poređenja pokazali su da se 29 ULS nalaze u barem jednom od upoređenih istraživanja. Od veće relevantnosti može se uzeti poređenje nalaza sa istraživanjem Gungor i Uisal (2016) zato što se korpusi razlikuju prevashodno samo po veličini. Dva istraživanja dele 37 ULS (74%), a činjenica da poklapanje u rezultatima nije veće ukazuje na značaj veličine korpusa pri identifikaciji ULS.

Pri poređenju ULS u dva potkorpusa u istraživanju Gungor (2016) utvrđeno je da dele 34 ULS (29%). Koristeći statističke podatke dobijene primenom testa verodostojnosti i *effect size* testa utvrđeno je da neizvorni govornici prekomerno ili nedovoljno koriste 21 od 34 zajedničkih ULS (četiri ULS nedovoljno korišćeni, 17 ULS prekomerno korišćeni).

Pri strukturalnoj raspodeli ULS u datom istraživanju urađena je i statistička analiza rezultata primenom testa verodostojnosti kako bi se utvrdilo da li su razlike u dobijenim rezultatima statistički značajne. Izvorni govornici su najfrekventnije koristili *fragmente sa predloškom frazom* (49% tipova) blisko praćene *imeničkim strukturama* (32% tipova), dok su neizvorni govornici najfrekventnije koristili *glagolske strukture* (44% tipova), koje su praćene *imeničkim strukturama* (26% tipova) i *fragmentima sa predloškom frazom* (24%). Najzastupljeniji okvir ULS kod izvornih govornika je *predloška fraza + of fragment*, koji pokriva 39% svih tipova. Isti okvir je zastupljen i u korpusu neizvornih govornika i pokriva 16% svih tipova. Najzastupljeniji okvir u korpusu neizvornih govornika je *imenička fraza + of fragment*, koji pokriva 18% svih tipova; to je ujedno i najzastupljeniji okvir iz grupe imeničkih ULS u oba korpusa, štaviše, to je drugi najzastupljeniji okvir u korpusu izvornih govornika (27%).

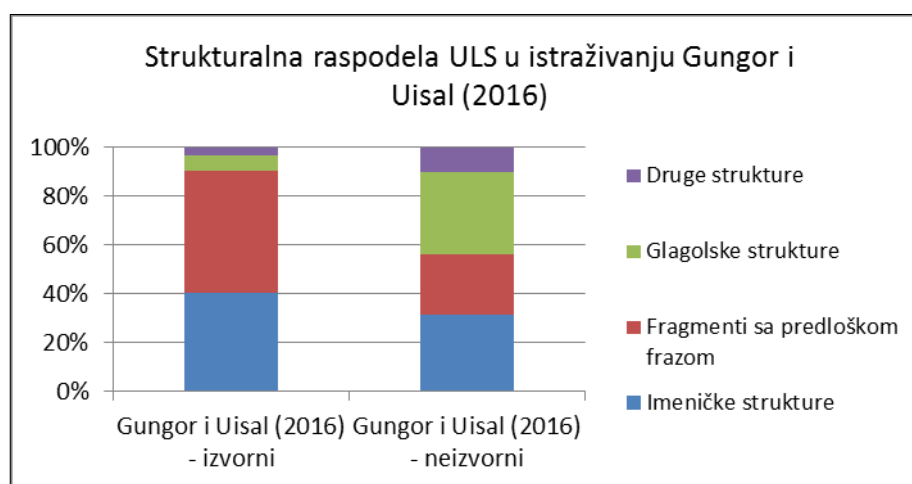
Funkcionalna raspodela ULS kod izvornih i neizvornih govornika u datom istraživanju pokazala je da postoje značajne razlike u raspodeli. Najzastupljenija funkcionalna kategorija kod izvornih govornika je *ULS orijentisanih ka istraživanju*, koja obuhvata 56% tipova, dok je najzastupljenija funkcionalna kategorija kod neizvornih govornika *ULS orijentisanih ka tekstu*, koja obuhvata 55% tipova. U oba potkorpusa najmanje su zastupljeni *ULS orijentisani ka učesniku* (6% tipova kod izvornih govornika i 8% tipova kod neizvornih). Značajno je napomenuti da dato istraživanje koristi taksonomiju koju je predložio Hajland (2008) i koju je modifikovala Salazar (2014).

Da bi ispitao međujezički uticaj na produkciju ULS kod izvornih govornika turskog, Gungor primenjuje pristup koji je predstavio Džarvis (Jarvis, 2000, 2012, Jarvis & Pavlenko, 2008). U prvom koraku autor izdvaja ULS koji su se nalazili u korpusu neizvornih govornika i koji se ne nalaze u korpusu izvornih govornika. Kao što je gore već pomenuto, 85 ULS se nalazilo samo u korpusu neizvornih govornika. U drugom koraku, ULS izdvojeni u prvom koraku upoređeni su sa ULS izdvojenim u istraživanju Perez-Ljatanda (2014) kako bi se izdvojili ULS specifični za neizvorne govornike iz Turske uporedno sa ULS koji su tipični za neizvorne govornike uopšte. Primenom drugog koraka spisak izdvojenih ULS se smanjio na 72 nakon što su izdvojeni ULS koji se nalaze i u korpusu izvornih govornika španskog. Nedostatak u ovom koraku je to što autor disertacije nije napravio sopstveni korpus sa neizvornim govornicima sa drugim maternjim jezikom koji bi bio u potpunosti kompatibilan sa korpusom korišćenim u samom istraživanju. Koristeći korpus iz drugog istraživanja, došlo je do razlika kako u veličini korpusa tako i u samoj oblasti kojom se bave naučni radovi u korpusu. Treći korak predstavljala je komparacija dobijenih ULS iz drugog koraka sa korpusom istraživačkih radova na turskom, odnosno pronalaženje prevodnih ekvivalenata na osnovu korpusnih podataka.

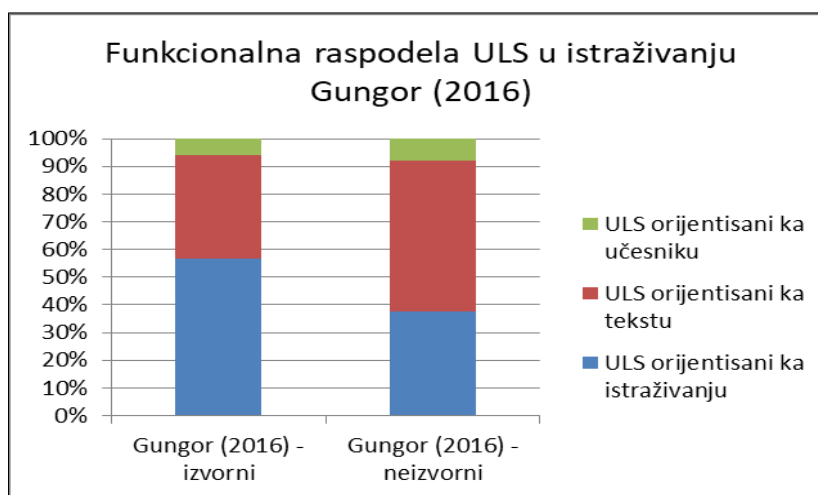
Pri komparaciji izdvojeni ULS funkcionalno su raspodeljeni i kroz tu prizmu upoređeni sa prevodnim ekvivalentima. Konačni rezultati su ukazali da se 54 ULS u korpusu neizvornih govornika koristi pod uticajem maternjeg jezika. Transfer je prevashodno nastao kod klauzalnih struktura koje su u funkcionalnim podgrupama *inferencijalni ULS*, *komparativnih ULS* i *ULS procedure*. Detalji istraživanja mogu se naći u Tabeli 15 i Grafikonu 11 i Grafikonu 12.

Tabela 15. Osnovni podaci o istraživanju Gungor (2016)

Istražena naučna oblast	Obrazovanje	
Maternji jezik neizvornih govornika engleskog jezika	Turski jezik	
Ispitana dužina ULS	4 reči	
Veličina korpusa	<i>Izvornih govornika</i>	<i>Neizvornih govornika</i>
	1.000.019	1.000.009
Primenjena minimalna frekvencija	20 ponavljanja u milion reči	
Primenjena disperzija	10% tekstova	
Broj izdvojenih ULS	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
	79	119
Korišćena taksonomija za strukturalnu analizu podataka	Taksonomija koju su predložili Bajber i sar. (1999)	
Strukturalna analiza tipova	<i>Izvorni govornici</i>	<i>Neizvorni govornici</i>
<i>Imeničke strukture</i>	32% (32%)	26% (25%)
<i>Fragmenti sa predložkom frazom</i>	49% (51%)	24% (28%)
<i>Glagolske strukture</i>	16% (14%)	44% (42%)
<i>Druge strukture</i>	3% (3%)	6% (5%)
Korišćena taksonomija za funkcionalnu analizu	Taksonomije koju je predložio Hajland (2008) a koju je modifikovala Salazarova (2014)	
Funkcionalna klasifikacija tipova	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
<i>ULS orijentisani ka istraživanju</i>	56%	38%
<i>ULS orijentisani ka tekstu</i>	37%	55%
<i>ULS orijentisani ka učesniku</i>	6%	8%



Grafikon 11. Strukturalna raspodela ULS u istraživanju Gungor (2016)



Grafikon 12. Funkcionalna raspodela ULS u istraživanju Gungor (2016)

Gungor (2019)

Godine 2019. Gungor je objavio još jedno istraživanje koje se bavi upotrebom ULS u naučnim radovima kod izvornih i neizvornih govornika. Međutim, u ovom radu fokus je na upotrebi samo *ULS sa uvodnom it strukturom* (engl. *anticipatory it structure*) kod izvornih i neizvornih govornika u naučnim radovima iz oblasti obrazovanja. Data struktura je interesantan fenomen koji je prisutniji u akademskom pisanju uporedno sa drugim žanrovima (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Hewings & Hewings, 2002; Hyland, 2008a; Römer, 2009; Zhang, 2015 prema Gungör, 2019). Ova struktura se koristi za tvrdnje o objektivnoj potrebi ili uverenju, za davanje utiska objektivnosti, za diskusiju subjektivnih procena autora i za iskazivanje činjenica na bezličan način u konstrukcijama sa pasivom (Gungör, 2019).

Za istraživanje je korišćen korpus MCRA *Multicultural Corpus of Research Articles*, odnosno dva potkorpusa od po milion reči: MCRA-L1, koji je sačinjen od naučnih radova u oblasti obrazovanja koji su napisani od strane izvornih govornika engleskog i MCRA-L2, koji je sačinjen od naučnih radova u oblasti obrazovanja koje su napisali izvorni govornici turskog jezika. Drugim rečima, korišćeni su korpusi koje je sačinio za goreopisanu doktorsku disertaciju. Za izdvajanje ULS određena je minimalna frekvencija od pet ponavljanja u milion reči, a izdvojeni su ULS od četiri, pet i šest reči. U ovom istraživanju se koristila praksa (Biber, Conrad, & Cortes, 2004) da se ULS svrstaju u kategoriju u kojoj se najčešće pojavljuju u datom korpusu, pošto, kao što je ranije naznačeno, jedan ULS može da ima nekoliko funkcija. Uvodna *it* struktura u ovom istraživanju funkcionalno je podeljena u pet kategorija: ograđivanje (engl. *hedges*), označitelji stava, naglašivači (engl. *emphatics*), pripisivanja (engl. *attributions*) i posmatranja (engl. *observation*). Zanimljivo je da su oba korpusa od po milion reči tako da normalizacija rezultata nije bila potrebna. Rezultati istraživanja su pokazali da su neizvorni govornici koristili uvodne *it* konstrukcije češće (74% više tipova i 225% više tokena) u odnosu na izvorne govornike, što je i u skladu sa prethodnim istraživanjima (npr. Hewings & Hewings, 2002, prema Gungör, 2019). Međutim, kada se pogledala funkcionalna raspodela uvidele su se kompleksnije razlike. Neizvorni govornici su češće koristili uvodnu *it* konstrukciju u funkcionalnim kategorijama: *naglašivača* (100% više), *pripisivanja* (800% više) i *posmatranja* (675% više), ali zato manje koriste uvodnu *it* konstrukciju u funkcionalnim kategorijama: *ograđivanja* (38% manje) i *označiteljima stava* (28% manje). Osnovni podaci o istraživanju mogu da se nađu u Tabeli 16.

Tabela 16. Osnovni podaci o istraživanju Gungor (2019)

Istražena naučna oblast	Obrazovanje	
Maternji jezik neizvornih govornika engleskog jezika	Turski jezik	
Ispitana dužina ULS	4, 5 i 6 reči	
Veličina korpusa	<i>Izvornih govornika</i>	<i>Neizvornih govornika</i>
	1.000.019	1.000.009
Primenjena minimalna frekvencija	5 ponavljanja u milion reči	
Primenjena disperzija	Nije naznačeno u istraživanju	
Broj izdvojenih ULS sa uvodnom <i>it</i> konstrukcijom	U korpusu izvornih govornika	U korpusu neizvornih govornika
	38	66

Pan, Rapen i Bajber (2016)

Istraživanje koje predstavlja glavni metodološki uzor je Pan, Rapen i Bajber (Pan, Rappen & Biber, 2016). Dato istraživanje napisano je od strane vodećih stručnjaka u oblasti frazeologije, a jedan od autora je i Bajber, koji je zapravo i otkrio fenomen ULS. Istraživanje se bavilo strukturalnim i funkcionalnim razlikama u korišćenju ULS u naučnim člancima iz polja telekomunikacije kod izvornih govornika engleskog i izvornih govornika kineskog jezika. Dva korpusa od oko pola miliona reči su formirana za istraživanje. Iako su izjednačeni po veličini, nisu jednaki po broju tekstova zbog tendencije da izvorni govornici engleskog pišu duže naučne radove. TELE-EN je sačinjen od 87 naučnih radova i ukupne je veličine 505.373 reči, dok je TELE-CH sačinjen od 179 naučnih radova i veličine 473.912 reči. U istraživanju je primenjena operaciona definicija izvornog govornika koju je predložio Vud (Wood, 2001). Kriterijumi za izdvajanje ULS su frekvencija od minimum 40 ponavljanja u milion reči i disperzija u minimum pet tekstova u TELE-EN i deset tekstova u TELE-CH. Analizirali su se samo ULS od četiri reči. U korpusu izvornih govornika je identifikovano 55 tipova ULS, dok je kod neizvornih govornika identifikovan 71 tip ULS, od toga 24 tipa ULS je pronađeno u oba korpusa (što čini 43% ULS izvornih govornika i 34% ULS neizvornih govornika). Autori su, takođe, proverili da li su ULS koji su identifikovani samo u jednom korpusu bili prisutni u drugom. Interesantno je da su ULS identifikovani samo u korpusu izvornih govornika svi bili prisutni u korpusu neizvornih, ali sa manjom frekvencijom i rasprostranjenošću, dok tri ULS koji su identifikovani u korpusu neizvornih govornika uopšte ne postoje u korpusu izvornih: *with the increase of, the reason is that, it is obvious that*. Ovi ULS imaju prevodne ekvivalente u kineskom. Takođe je utvrđeno da se neizvorni istraživači više oslanjaju na ULS u svom akademskom pisanju uporedno sa izvornim govornicima. ULS u dva korpusa su takođe upoređena na osnovu strukturalne i funkcionalne raspodele. Strukturalna podela je preuzeta od Bajbera i saradnika (1999, poglavlje 1.2), a ULS su podeljeni u šire kategorije imeničkih, predloških i glagolskih ULS koji su dalje svrstani u podgrupe. Testovi verodostojnosti (engl. *log-likelihood*) su primenjeni radi provere statističke značajnosti razlika u rezultatima u dva korpusa. Izvorni govornici koriste značajno više *imeničkih struktura* i *fragmenata sa predloškom frazom*, dok neizvorni govornici koriste značajno više *glagolskih struktura* u većini glagolskih podgrupa.

U istraživanju Pan i saradnici (2016) rezultati su upoređeni sa rezultatima istraživanja Čen i Bejker (Chen & Baker, 2010). Ustanovilo se da istraživači čiji je maternji jezik kineski imaju tendenciju da se oslanjaju na *klauzalne ULS* slično kao i dve grupe studenata u istraživanju Čen i Bejker (2010). Pan i saradnici su takođe poredili produktivnost najzastupljenijih okvira u svakoj od tri glavne strukturalne kategorije: „*the + imenica + of the/a*” za *imeničke strukture*, „*in the + imenica + of*” za *fragmente sa predloškom frazom* i „*glagol u pasivu + predložka fraza*” za *glagolske*

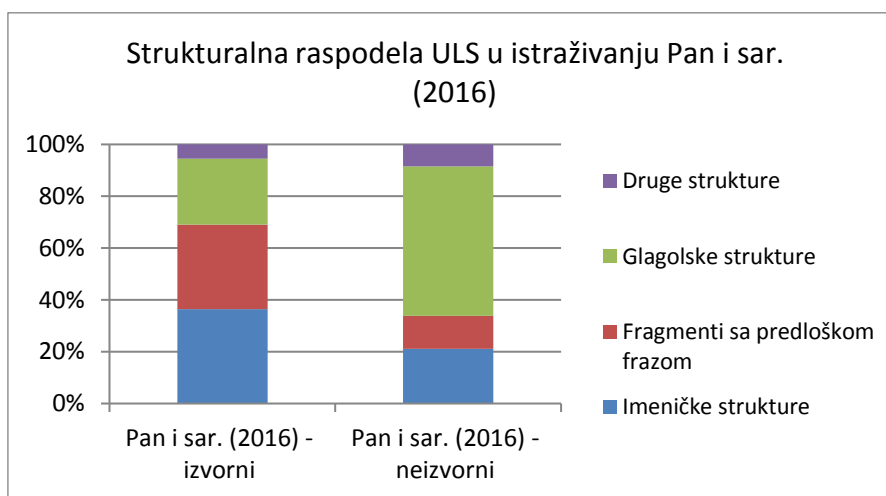
strukture. Istraživači koji su neizvorni govornici engleskog su koristili manju varijaciju imenica u imeničkom okviru, koristili su duplo manje tipova ove strukture sa različitim imenicama i tri puta manje tokena od istraživača koji su izvorni govornici engleskog jezika. Kada je reč o predloškom okviru, istraživači kojima je engleski jezik maternji koristili su duplo više tokena uporedno sa istraživačima kojima je engleski jezik neizvorni, iako je broj tipova sličan, dok je razlika između studenata kojima je engleski jezik strani i istraživača kojima je engleski jezik maternji opet veća, odnosno istraživači su koristili tri puta više tipova i čak četiri puta više tokena. Istraživači kojima engleski jezik nije izvorni koristili su tri puta više tipova i četiri puta više tokena glagolske pasivne konstrukcije, međutim, studenti kojima engleski jezik nije izvorni, manje su koristili ovaj okvir od istraživača kojima je engleski jezik izvorni. Na osnovu ovog poređenja, može se zaključiti da postoje razlike između izvornih i neizvornih govornika i na visokom jezičkom nivou publikovanih naučnih radova, ali da su razlike manje nego na nižem studentskom nivou.

Pri funkcionalnoj kategorizaciji ULS u ovom istraživanju primenjen je spoj taksonomija predloženih od strane Bajber i saradnika (1999) i Hajlanda (2008) pa su tri glavne kategorije: *ULS orijentisani ka istraživanju*, *ULS orijentisani ka tekstu* i *ULS orijentisani ka stavu*. Dva korpusa naučnih radova iz oblasti telekomunikacije pokazuju sličnu funkcionalnu raspodelu ULS po tipovima i po tokenima. Međutim, primećuju se značajne razlike kada se uradi poređenje funkcionalnih potkategorija. Istraživači kojima engleski jezik nije izvorni značajno su više koristili ULS u potkategorijama: *ULS za opis*, *strukturni signali*, *tranzicioni signali* i u svim potkategorijama *stava*, dok su izvorni istraživači znatno češće koristili ULS iz potkategorija *ULS za kvantifikaciju* i *ograničavajuće signale* (engl. *framing*). Istraživači koji su izvorni govornici engleskog jezika koristili su 14 tipova ULS za kvantifikaciju, dok su istraživači kojima je engleski jezik strani koristili osam tipova koji su se uglavnom oslanjali na imenicu *broj (number) ULS* (npr. *is the number of, the average number of...*). Interesantno je da i Kortež (2004) u svom istraživanju pronalazi da studenti biologije takođe retko koriste ULS za kvantifikaciju u svom akademskom pisanju, što može da ukazuje na to da se ova karakteristika akademskog pisanja kasnije usvaja. Obe grupe istraživača koriste potkategoriju strukturisanja da uvedu čitaoca u organizaciju rada ili da mu skrenu pažnju na određene tabele i šeme. Međutim, tri ULS sa imenicom *section (in the next section, in the previous section, in this section we)* koji su često korišćeni od strane izvornih govornika, veoma retko su upotrebljavani od strane neizvornih govornika. Ova razlika možda ukazuje na veću orijentisanost ka čitaocima od strane izvornih govornika a može da bude i u vezi sa činjenicom da izvorni govornici engleskog pišu duže naučne radove. Osnovni podaci o istraživanju mogu se naći u Tabeli 17, kao i grafikonima koji slede (v. Garfikon 13 i Grafikon 14).

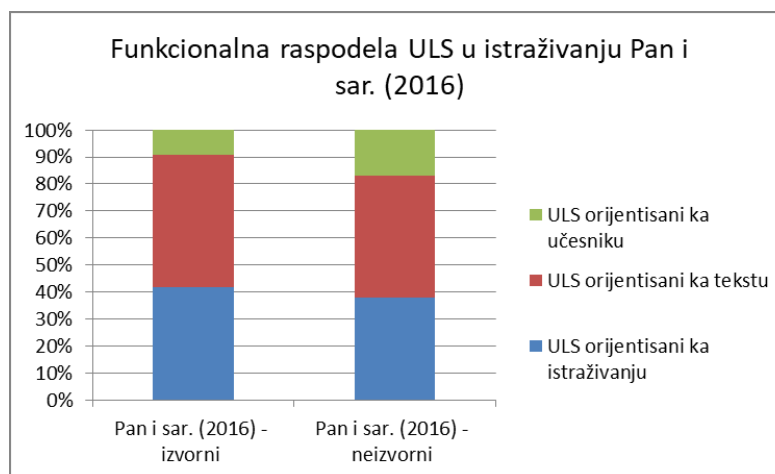
Tabela 17. *Pan, Rapen i Bajber (2016)*

Tabela 17. <i>Pan, Rapen i Bajber (2016)</i>		
Istražena naučna oblast	Telekomunikacije	
Maternji jezik neizvornih govornika engleskog jezika	Kineski jezik	
Ispitana dužina ULS	4 reči	
Veličina korpusa	<i>Izvornih govornika</i>	<i>Neizvornih govornika</i>
	505.373	473.912
Primenjena minimalna frekvencija	40 ponavljanja u milion reči	
Primenjena disperzija	5 tekstova i 10 tekstova	
Broj izdvojenih ULS	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
	55	71
Korišćena taksonomija za strukturalnu analizu podataka	Taksonomija koju su predložili Bajber i sar. (1999)	
Strukturalna analiza tipova	<i>Izvorni govornici</i>	<i>Neizvorni govornici</i>
<i>Imeničke strukture</i>	36.4% (33.9%)	21.1% (18.4%)

<i>Fragmenti sa predloškom frazom</i>	32.6% (33.2%)	12.7% (12.8%)
<i>Glagolske strukture</i>	25.5% (26.6%)	57.7% (55.8%)
<i>Druge strukture</i>	5.5% (6.3%)	8.5% (13%)
Korišćena taksonomija za funkcionalnu analizu	Spoj taksonomije koju je predložio Hajland (2008) i taksonomije koju je predložio Bajber (1999)	
Funkcionalna klasifikacija tipova	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
<i>ULS orijentisani ka istraživanju</i>	41.82% (43%)	38.03% (38%)
<i>ULS orijentisani ka tekstu</i>	49.09% (49%)	45.07% (49%)
<i>ULS orijentisani ka stavu</i>	9.09% (8%)	16.90% (13%)



Grafikon 13. *Strukturalna raspodela ULS u istraživanju Pan i sar. (2016)*



Grafikon 14. *Funkcionalna raspodela ULS u istraživanju Pan i sar. (2016)*

Istraživanje Pan, Rapen i Bajber (Pan, Rappen & Biber, 2020) nadovezuje se na njihovo istraživanje iz 2016. godine. Za razliku od drugih istraživanja, u ovom istraživanju fokus je na metodološke odluke pri sprovođenju istraživanja ULS i njihove posledice na dobijene rezultate. Zbog te činjenice, ovaj rad će prevashodno biti koristan pri interpretaciji podataka i identifikaciji potencijalnih nedostataka istraživanja ako ne i same oblasti.

Cilj istraživanja koje su sproveli Pan i saradnici (2020) je ustanoviti kako kompozicija korpusa (broj reči i broj tekstova koji sačinjavaju korpus) utiče na broj ULS koji su izdvojeni, kao i na komparaciju upotrebe ULS u korpusima. Da bi postigli naveden cilj, istraživači su koristili korpus iz istraživanja Pan i saradnika (2016) i modifikovali ih u tri eksperimentalna oblika:

1. *Eksperiment 1.* Korpusi su ujednačeni da imaju skoro identičan broj reči (L1 engleski korpus – 473.985 reči sačinjen od 84 teksta, L2 engleski korpus – 473.921 reči, sačinjen od 179 tekstova);
2. *Eksperiment 2.* Korpusi su izjednačeni po broju tekstova (L1 engleski korpus – 505.373 reči, sačinjen od 87 tekstova, L2 engleski korpus – 241.483 reči, sačinjen od 87 tekstova);
3. *Eksperiment 3.* Korpusi su izjednačeni po broju tekstova i reči (L1 engleski korpus – 230.376 reči, sačinjen od 87 tekstova, L2 engleski korpus – 230.377 reči, sačinjen od 87 tekstova).

Eksperimentalni korpusi si upoređeni sa korpusima iz originalnog istraživanja. U sve četiri varijante podataka primenjena je minimalna frekvencija od 40 ponavljanja u milion reči i minimalna disperzija u 10% tekstova. Ono što se može uzeti kao mana navedene postavke istraživanja je što se u *eksperimentalnoj postavci 3* veličina korpusa nije približila originalnom istraživanju i *eksperimentalnoj postavci 1*, jer, kao što je i u samom radu napomenuto, postoji tendencija da se u manjim korpusima identifikuje veći broj ULS (str. 219). Takođe je bitno napomenuti da su radovi izvornih govornika engleskog za *eksperimentalnu postavku 3* skraćeni da bi se poklopio broj radova i veličina korpusa.

Rezultati istraživanja su pokazali da se broj identifikovanih ULS u sve četiri varijante postavki korpusa razlikuju. U originalnoj postavci i u *eksperimentu 3* veći je broj ULS identifikovan kod neizvornih govornika iako je sam broj identifikovanih ULS manji u *eksperimentu 3*. Ovo je iznenađujuće jer prethodna istraživanja navode da bi manji korpusi trebalo da generišu veći broj ULS (Kortez, 2004; Čen i Bejker, 2010; Pan i sar., 2020). Još jedan iznenađujući podatak jesu brojevi izdvojenih ULS u korpusima – najbliži su između *eksperimenta 1* i *2*, a u njima identifikovan veći broj ULS kod izvornih govornika engleskog jezika. Moramo ovde da uzmemo u obzir da se, iako su originalno istraživanje i *eksperiment 1* najbliži što se tiče kompozicije korpusa, ipak razlikuju u parametrima za identifikaciju ULS, jer u originalnom istraživanju minimalna disperzija nije u 10% tekstova nego je fiksni broj tekstova (pet za korpus izvornih govornika i deset za korpus neizvornih govornika). Zaključak koji istraživači daju na osnovu ovih rezultata jeste da kompozicija korpusa i parametri koji su postavljeni za identifikaciju ULS imaju jak uticaj na broj izdvojenih ULS u korpusima i komparaciju korpusa na osnovu broja identifikovanih ULS.

Sledeće što su istraživači sagledali je da li su se isti tipovi ULS identifikovali u različitim eksperimentima. Ustanovljeno je da se manje od 50% tipova pojavljuje u sva tri eksperimenta. Ovakvi rezultati su doveli istraživače do zaključka da se tipovi identifikovanih ULS mogu mnogo razlikovati u zavisnosti od kompozicije korpusa. Poslednji test koji su istraživači primenili daje

malo pozitivnije rezultate za istraživanja u ovoj oblasti. Naime, urađena je strukturalna raspodela ULS u sva tri eksperimenta i upoređena je sa originalnim istraživanjem. Rezultati su pokazali da su veoma male razlike u strukturalnim raspodelama identifikovanih ULS u sva tri eksperimenta i originalnog istraživanja, što je dovelo istraživače do zaključka da kompozicija korpusa ne utiče mnogo na strukturalnu raspodelu identifikovanih ULS i poređenje korpusa po ovim parametrima. Nažalost, u istraživanju nije urađena funkcionalna raspodela ULS.

Istraživači navode da je neophodno sprovesti dalja istraživanja koja će se baviti pitanjem varijacije po temama koje se obrađuju u tekstovima u korpusima koji se koriste u ovoj vrsti istraživanja jer je pretpostavka da varijacije u obrađenim temama takođe mnogo utiču na izdvojene tipove ULS. Glavni zaključak istraživanja jeste da bi istraživači u ovoj oblasti trebalo pažljivo da formiraju korpus i imaju na umu uticaj kompozicije korpusa i parametara za izdvajanje na dobijene rezultate.

Esfandiari i Barbari (2017)

Još dva kvalitetna i značajna istraživanja sa sličnim postavkama i ciljevima kao i ova disertacija su istraživanja koja su uradili Esfandiari i Barbari (Esfandiari & Barbary, 2017) i Al Fedžri, Veisito, Kirana (Al Fajri, Wasito & Kirana, 2020). Esfandiari i Barbari (2017) ispituju upotrebu ULS kod izvornih i neizvornih govornika u naučnim radovima iz oblasti psihologije. Maternji jezik neizvornih govornika engleskog u ovom istraživanju je persijski. Glavni ciljevi istraživanja su pronaći odgovore na sledeća pitanja: 1. *Koji su najfrekventniji ULS od četiri, pet i šest reči u naučnim radovima iz oblasti psihologije?* 2. *Koji su ULS zajednički za izvorne i neizvorne govornike?* 3. *Da li postoje značajne razlike u upotrebi ULS u naučnim radovima iz oblasti psihologije kod izvornih govornika i neizvornih govornika engleskog kojima je maternji jezik persijski?* Za potrebe istraživanja formirana su dva korpusa. Istraživači su intervjuisali nekoliko psihologa kako bi saznali koje istraživačke oblasti i naučni časopisi adekvatno predstavljaju oblast psihologije. Na osnovu intervjuja određeno je sedam poddisciplina, a za svaku su izdvojena po dva reprezentativna časopisa. U procesu identifikacije maternjeg jezika autora korišćena je radna definicija koju je predložio Vud (2001) i koja nalaže da se smatra da je autor naučnog istraživanja izvorni govornik ciljnog jezika ako u afilijaciji ima instituciju koja je u zemlji gde se taj ciljni jezik govori kao zvanični i ako autor poseduje prezime tipično za govornike ciljnog jezika. Da bi rad postao deo korpusa, morao je da sadrži istraživački model *Uvod–Metod–Rezultati–Diskusija* (*Introduction-Method-Result-Discussion format*, IMRD). Istraživanja koja sačinjavaju korpus izvornih govornika objavljena su u periodu između 2012. i 2015. godine, dok su istraživanja koja sačinjavaju korpus neizvornih govornika objavljena u periodu između 2006. i 2015. Korpus izvornih govornika sastoji se od 840 naučnih radova i ukupno 2.185.952 reči, dok je korpus neizvornih govornika sačinjen od 612 naučnih radova i ukupno 2.184.678 reči. Na osnovu broja radova koji se nalaze u korpusima može se uvideti da persijski istraživači imaju tendenciju da pišu duže naučne radove od anglofonih, što je suprotno situaciji u istraživanjima Güngör & Uysal (2016), Güngör (2016) i Pan i saradnici (2017).

Frekvencija za identifikaciju ULS u istraživanju je različita u zavisnosti od dužine spoja: 40 ponavljanja u milion reči za ULS od četiri reči, 20 ponavljanja u milion reči za ULS od pet reči i deset ponavljanja u milion reči za ULS od šest reči. Veći broj ULS od četiri i pet reči je identifikovano u korpusu izvornih govornika (EC 254 i 115; PC 116 i 104), dok je veći broj ULS od šest reči identifikovan u korpusu neizvornih govornika engleskog jezika (EC 47; PC 96). Na rezultatima frekvencije primenjen je hi-kvadrat test koji je ukazao na to da je razlika u rezultatima značajna. Izvorni govornici engleskog i persijskog su koristili sličan broj ULS od pet reči, dok im je frekvencija upotrebe ULS od četiri i šest reči značajno drugačija: 55 ULS od pet reči koji su bili identifikovani u persijskom korpusu uopšte se ne nalaze u korpusu izvornih govornika engleskog, čak ni sa manjom frekvencijom. Istraživači su proverili i koji se broj identifikovanih ULS nalazi u

oba korpusa i došli do podataka da su dva korpusa delila 37 ULS od četiri reči (11%), 16 ULS od pet reči (8%) i osam ULS od šest reči (6%).

U istraživanju su uporedili koji se pojedinačni ULS identifikovani u ovom istraživanju takođe pojavljuju u dva prethodna istraživanja, naime, u istraživanjima Adel i Erman (2012) i Čen i Bejker (2010). Izdvojeno je sedam ULS koji se nalaze u sva tri istraživanja. Istraživači su naveli da bi ULS koji su prisutni u rezultatima sva tri istraživanja bila od izuzetnog značaja zato što su prisutni u različitim disciplinama, žanrovima i kod različitih grupa neizvornih govornika. Komparacija konkretnih primera ULS u korpusima različitih istraživanja se pokazala kao nedovoljno pouzdana zbog razlika u korpusima (Pan i sar., 2016). To je u ovom slučaju od posebne važnosti zato što se istraživanja, sa kojima se porede rezultati, zapravo bave upotrebom ULS kod studenata a ne istraživača kao što je slučaj kod Esfandiari i Barbari (2017).

Kada je reč o strukturalnoj analizi ULS i datom istraživanju, trebalo bi napomenuti da su autori koristili taksonomiju koju su predložili Bajber i saradnici (2014), što podrazumeva da pored tri osnovne strukturalne kategorije imamo i kategoriju *framenata zavisne klauze* (engl. *dependant clause fragments*). Takođe, pri analizi podataka *imeničke strukture* i *fragmenti sa predloškom frazom*, istraživači su gledali kao jednu kategoriju *frazalnih* (engl. *phrasal*) ULS. Pri strukturalnoj analizi ULS od četiri reči, pokazalo se da su izvorni i neizvorni govornici u dotičnom istraživanju na sličan način strukturalno koristili ove leksičke konstrukcije. U obe grupe najfrekventniji su *frazalni ULS*. Kao važno u ovim rezultatima može se izdvojiti češća upotreba *ULS sa glagolom u pasivu* od strane izvornih govornika engleskog (13%), uporedno sa izvornim govornicima persijskog (8%). Dati rezultati su u skladu sa istraživanjima Bajber i saradnika (1999), Čen i Bejker (2010), kao i Hajland (2008), a nisu u skladu sa Pan i saradnici (2016). Funkcionalna analiza ULS od četiri reči pokazala je da izvorni govornici engleskog i izvorni govornici persijskog na sličan način raspoređuju ULS po glavnim funkcionalnim kategorijama. Analiza podgrupa pokazala je da su izvorni govornici koristili više *ULS stava* od neizvornih govornika (izvorni govornici 24,8% uporedno sa 16,4 %), ali i da su neizvorni govornici koristili više *rezultativnih signala* od izvornih govornika (neizvorni govornici 27,6% uporedno sa 18,5%).

Tokom analize strukturalne raspodele ULS od pet reči došlo se do rezultata da su neizvorni govornici koristili više *imeničkih struktura* i *framenata sa predloškom frazom*, dok su izvorni govornici koristili više *glagolskih struktura*. Primena ULS sa *zavisnim klauzalnim elementima* je približno ista kod obe grupe govornika. Pri detaljnijoj analizi raspodele ULS po strukturalnim podgrupama interesantno je izdvojiti da su se *imeničke fraze sa of fragmentom* češće koristile od strane neizvornih govornika (neizvorni govornici 24% uporedno sa 6% kod izvornih). Ovoj razlici je najviše doprineo okvir *the+imenica+of the+study*. Neizvorni govornici su koristili raznolike imenice u navedenom okviru (npr. *purpose, findings, aim, results*), dok su izvorni govornici iskoristili samo jednu imenicu u tom okviru (*end*). Prema rezultatima funkcionalne raspodele ULS od pet reči kod izvornih i neizvornih govornika, lako je uočljiva veća razlika uporedno sa ULS od četiri reči. Izvorni govornici koristili su manji broj ULS od pet reči *orijentisanih ka tekstu* (26%) uporedno sa neizvornim govornicima (71%), dok su neizvorni govornici koristili manje ULS od pet reči *orijentisanih ka istraživanju* (22%) i *učesniku* (7%) uporedno sa izvornim govornicima (*orijentisani prema istraživanju* 50% i *učesniku* 24%). Pri analizi funkcionalnih podgrupa može se izdvojiti i da su neizvorni govornici koristili više *rezultativnih* (35%) i *ograničavajućih signala* (6%) uporedno sa izvornim govornicima (10% i 0%).

Analiza strukturalne raspodele ULS od šest reči kod izvornih i neizvornih govornika pokazala je slične rezultate, kao i za strukturalnu podelu ULS od pet reči. Naime, i kod izvornih i kod neizvornih govornika najzastupljeniji su *imeničke strukture* i *fragmenti sa predloškom frazom*. Detaljnija analiza strukturalnih podgrupa pokazala je da su neizvorni govornici više koristili okvir

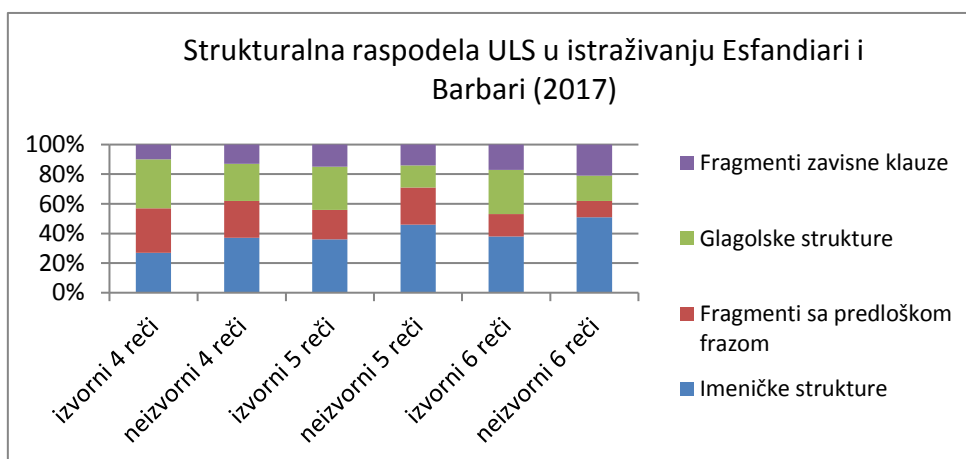
imenica + of this/present study + be, koji nije čest u objavljenim naučnim tekstovima (Adel & Erman, 2012, prema Esfandiari & Barbary, 2017). Samo jedan ULS koji je sastavljen sa ovim okvirom je prisutan u oba korpusa (*aim of this study was to*). Dve navedene grupe autora su koristile različite okvire za izražavanje pasiva. Naime, izvorni govornici su koristili okvir *glagol u pasivu+predložka fraza*, koji u sebi sadrži glagol *assign* i tri različite predložke fraze (*to one of two, to one of three, to one of four*). Navedeni okvir nije bio prisutan u korpusu neizvornih govornika. Neizvorni govornici su koristili okvir *glagol u pasivu+fragemat to klauze*, koji se nije našao u korpusu izvornih govornika. Pri analizi funkcionalnog korišćenja ULS od šest reči između izvornih i neizvornih govornika nije nađena značajna korelacija. Osnovni podaci o istraživanju Esfandiari i Barbary (2017) mogu se videti u Tabeli 18 i grafikonima dole (v. Grafikon 15 i Grafikon 16).

Tabela 18. Osnovni podaci o istraživanju Esfandiari i Barbary (2017)

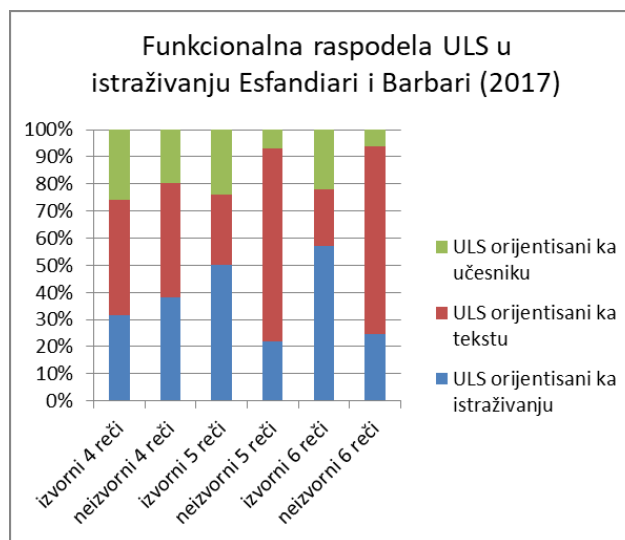
Tabela 18. Osnovni podaci o istraživanju Esfandiari i Barbary (2017)		
Istražena naučna oblast	Psihologije	
Maternji jezik neizvornih govornika engleskog jezika	Persijski jezik	
Ispitana dužina ULS	4–6 reči	
Veličina korpusa	<i>Izvornih govornika</i>	<i>Neizvornih govornika</i>
	2.185.952	2.184.678
Primenjena minimalna frekvencija	40 ponavljanja u milion reči – za ULS od 4 reči 20 ponavljanja u milion reči – za ULS od 5 reči 10 ponavljanja u milion reči – za ULS od 6 reči	
Primenjena disperzija	5 tekstova	
Broj izdvojenih ULS	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
od 4 reči	254	116
od 5 reči	115	104
od 6 reči	47	96
Korišćena taksonomija za strukturalnu analizu podataka	Taksonomija koju su predložili Bajber i sar. (2004)	
Strukturalna analiza tipova za ULS od 4 reči	<i>Izvorni govornici</i>	<i>Neizvorni govornici</i>
<i>Imeničke strukture</i>	27%	37%
<i>Fragmenti sa predložkom frazom</i>	30%	25%
<i>Glagolske strukture</i>	33%	25%
<i>Fragmenti zavisne klauze</i>	10%	13%
Strukturalna analiza tipova za ULS od 5 reči	<i>Izvorni govornici</i>	<i>Neizvorni govornici</i>
<i>Imeničke strukture</i>	36%	46%
<i>Fragmenti sa predložkom frazom</i>	20%	25%
<i>Glagolske strukture</i>	29%	15%
<i>Fragmenti zavisne klauze</i>	15%	14%
Strukturalna analiza tipova za ULS od 6 reči	<i>Izvorni govornici</i>	<i>Neizvorni govornici</i>
<i>Imeničke strukture</i>	38%	51%
<i>Fragmenti sa predložkom frazom</i>	15%	11%
<i>Glagolske strukture</i>	30%	17%
<i>Fragmenti zavisne klauze</i>	17%	21%
Korišćena taksonomija za funkcionalnu analizu	Modifikovanu taksonomiju koju je predložio Hajland (2008)	

Tabela 18. Osnovni podaci o istraživanju Esfandiari i Barbari (2017)

Funkcionalna klasifikacija ULS od 4 reči	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
<i>ULS orijentisani ka istraživanju</i>	31.50%	38%
<i>ULS orijentisani ka tekstu</i>	42.50%	42.20%
<i>ULS orijentisani ka učesniku</i>	26%	19.80%
Funkcionalna klasifikacija ULS od 5 reči	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
<i>ULS orijentisani ka istraživanju</i>	50%	22%
<i>ULS orijentisani ka tekstu</i>	26%	71%
<i>ULS orijentisani ka učesniku</i>	24%	7%
Funkcionalna klasifikacija ULS od 6 reči	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
<i>ULS orijentisani ka istraživanju</i>	57%	24%
<i>ULS orijentisani ka tekstu</i>	21%	68%
<i>ULS orijentisani ka učesniku</i>	22%	6%



Grafikon 15. Strukturalna raspodela ULS u istraživanju Esfandiari i Barbari (2017)



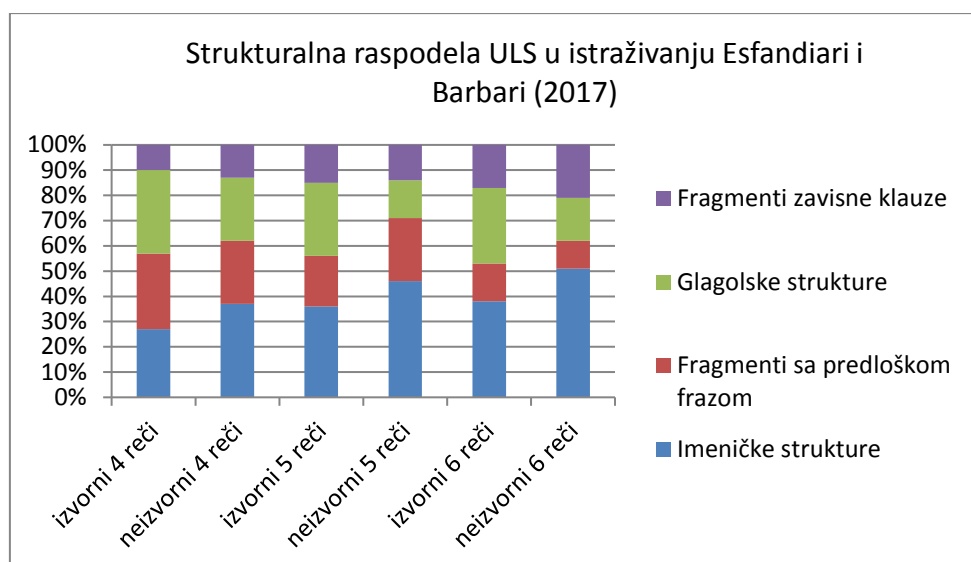
Grafikon 16. Funkcionalna raspodela ULS u istraživanju Esfandiari i Barbari (2017)

Istraživanje Al Fedžri, Veisito, Kirana (Al Fajri, Wasito & Kirana, 2020) bavi se ispitivanjem upotrebe ULS kod izvornih govornika engleskog i neizvornih govornika kojima je maternji indonežanski u naučnim radovima iz oblasti primenjene lingvistike. Dva korpusa su sačinjena za potrebe istraživanja. Prvi korpus se sastoji od naučnih radova iz oblasti primenjene lingvistike napisanih od strane izvornih govornika engleskog jezika. Radovi su izdvojeni iz naučnih časopisa sa visokim impakt faktorima (*Impact Factor* – IF), koji su indeksirani u bazama *Scopus* i *Social Science Citation Index* (SSCI). Drugi korpus je sačinjen od naučnih radova iz oblasti primenjene lingvistike napisanih od strane izvornih govornika indonežanskog jezika. Radovi su izdvojeni iz naučnih časopisa indeksiranih u bazama *Scopus* i *Directory of Open Access Journals* (DOAJ), a koji su akreditovani od strane Indonežanskog ministarstva za istraživanje, tehnologiju i visoko obrazovanje. Radovi koji sačinjavaju oba korpusa objavljeni su u periodu od 2016. do 2018 godine. Korpusi su ujednačeni po veličini (*Engleski korpus EC* – 1.325.986 i *Indonežanski korpus IC* – 1.334.752), ali je broj radova različit zbog tendencije autora koji su izvorni govornici engleskog da imaju duže radove (EC – 158 i IC – 274). Radna definicija za određivanje izvornih govornika koju je predložio Vud (2001) primenjena je u istraživanju. Istraživači su analizirali ULS od četiri reči za čiju identifikaciju je određena minimalna frekvencija od 40 ponavljanja u milion reči i minimalna distribucija u 10% tekstova.

Veći broj tipova 51 i tokena 4.874 identifikovano je u korpusu neizvornih govornika, dok je u korpusu izvornih govornika izdvojen 31 tip i 2.700 tokena. Kontrastivna analiza izdvojenih ULS pokazala je da se 14 tipova ULS nalazi u oba korpusa, što znači da su neizvorni govornici koristili skoro polovinu (14 od 31) ULS izdvojenih kod izvornih govornika. Međutim, u spiskovima deset najfrekventnijih ULS poklapaju se samo dva tipa ULS.

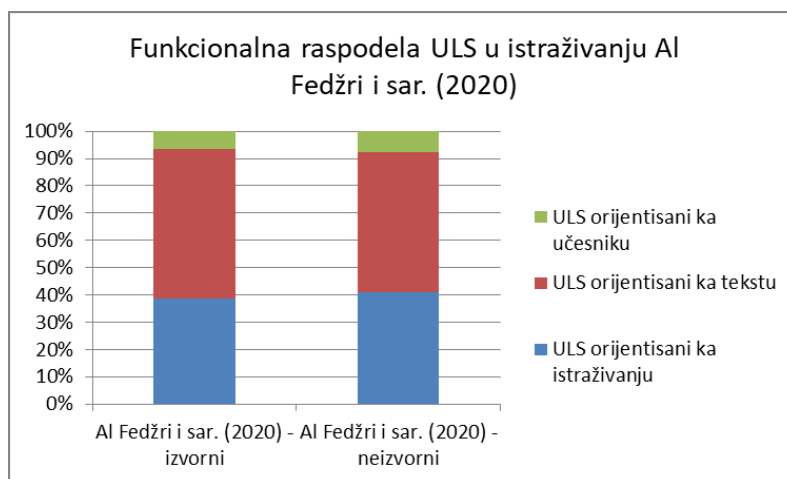
Strukturalna analiza podataka je pokazala da su *fragmenti sa predloškom frazom* najbrojniji kod obe grupe stručnjaka, međutim, veći procenat od ukupnog broja ULS je u ovoj kategoriji kod izvornih govornika (EC – 51,61%, IC – 39,22%). *Imeničke strukture* su zastupljenije kod izvornih govornika (EC – 32,26%, IC – 19,61%), dok su *glagolske strukture* zastupljenije kod neizvornih govornika (EC – 12,90%, IC – 37,25%). Upotreba većeg broja *fragmenta sa predloškom frazom* možda se može objasniti činjenicom da je za ovu oblast potrebno veliko znanje stranih jezika, stoga su ovi stručnjaci bliže na putu usklađivanja svoje jezičke produkcije sa normama postavljenim od strane izvornih govornika.

Pri komparaciji podataka funkcionalne raspodele ULS u glavne kategorije pokazalo se da je raspodela poprilično ujednačena kod obe grupe govornika. Ovakvi rezultati su u skladu sa prethodnim istraživanjima kod persijskih istraživača (Esfandiari & Barbary, 2017), kao i kod kineskih istraživača (Pan i sar., 2016). Međutim, razlike u raspodeli ULS mogu se naći u skoro svim podgrupama funkcionalne kategorizacije. Neizvorni govornici su znatno češće koristili ULS iz podgrupa *ULS za proceduru*, *ULS za opis*, *tranzicione signale*, *rezultativne signale*, *strukturalne signale*, *ograničavajuće signale* i *ULS uključivanja*, dok su neizvorni govornici ređe koristili podgrupe *ULS za lokaciju* i *kvantifikaciju*. *ULS stava* nisu identifikovani ni u jednoj od dve grupe govornika. Neizvorni govornici su koristili više *ULS uključivanja*, što je u skladu sa rezultatima prethodnih istraživanja (Esfandiari & Barbary, 2017; Pan i sar., 2016). Osnovni podaci o istraživanju Al Fedžri i saradnici (2020) može se naći u Tabeli 19, kao i na Grafikonu 17 i Grafikonu 18.



Grafikon 17. *Strukturalna raspodela ULS u istraživanju Al Fedžri i sar. (2020)*

Istražena naučna oblast	Primenjena lingvistika	
Maternji jezik neizvornih govornika engleskog jezika	Indonežanski jezik	
Ispitana dužina ULS	4 reči	
Veličina korpusa	<i>Izvornih govornika</i>	<i>Neizvornih govornika</i>
	1.325.986	1.334.752
Primenjena minimalna frekvencija	40 ponavljanja u milion reči	
Primenjena disperzija	10% tekstova	
Broj izdvojenih ULS	<i>U korpusu izvornih govornika</i>	<i>U korpusu neizvornih govornika</i>
	31	51
Korišćena taksonomija za strukturalnu analizu podataka	Taksonomija koju su predložili Bajber i sar. (1999)	
Strukturalna analiza tipova i tokena	<i>Izvorni govornici</i>	<i>Neizvorni govornici</i>
<i>Imeničke strukture</i>	32.26% (31.56%)	19.61% (19.10%)
<i>Fragmenti sa predloškom frazom</i>	51.61% (52.67%)	39.22% (39.84%)
<i>Glagolske strukture</i>	12.90% (10.89%)	37.25% (37.83%)
<i>Druge strukture</i>	3.23% (4.89%)	3.92% (4.23%)
Korišćena taksonomija za funkcionalnu analizu	Taksonomija koju je predložio Hajland (2008)	
Funkcionalna klasifikacija tipova i tokena	U korpusu izvornih govornika	U korpusu neizvornih govornika
<i>ULS orijentisani ka istraživanju</i>	38.71% (36.59%)	41.18% (35.18%)
<i>ULS orijentisani ka tekstu</i>	54.84% (57.30%)	50.98% (54.97%)
<i>ULS orijentisani ka učesniku</i>	6.45% (6.11%)	7.84% (9.23%)



Grafikon 18. *Funkcionalna raspodela raspodela ULS u istraživanju Al Fedžri i sar. (2020)*

Kritička analiza i poređenje opisanih relevantnih istraživanja

Nakon što su predstavljena relevantna istraživanja u uskoj oblasti ispitivanja upotrebe ULS u naučnim radovima kod izvornih i neizvornih govornika, prelazi se na sagledavanje njihovih karakteristika na osnovu značajnih elemenata u ovoj vrsti istraživanja. Istraživanja koja su opisana jesu sva relevantna istraživanja na temu koje je autorka disertacije pronašla. Istraživanja koja su se bavila izabranom temom (npr. Safarzadeh, Monfared & Sarfeju, 2015; Akbulut, 2020) ali čiji se kvalitet pokazao kao nedovoljno dobar – nisu uključena u razmatranje.

Najpre ćemo razmotriti grupe neizvornih govornika čija se produkcija ULS ispitala u naučnim radovima, odnosno koji su maternji jezici do sada ispitani u datom kontekstu. U navedenim istraživanjima ispitao se uticaj sedam maternjih jezika: filipinski, španski, srpski, turski, kineski, persijski i indonežanski. Uticaj španskog je ispitano u dva istraživanja sa različitim autorima i dva specifična pristupa problemu, dok je uticaj turskog ispitano u tri istraživanja istog autora sa sličnim postavkama.

Sva navedena istraživanja su se usredsredila na upotrebu ULS kod izvornih i neizvornih govornika u naučnim radovima jedne naučne oblasti, dok je jedino istraživanje Perez-Ljatanda (2014) ispitalo upotrebu ULS kod izvornih i neizvornih govornika u 12 naučnih oblasti (primenjena lingvistika, informatika, književnost, sociologija, poslovni menadžment, geografija, urologija, hematologija, onkologija, mehanički inženjering, prehrambena tehnologija i geonauka). Prethodna istraživanja ispitala su upotrebu ULS kod izvornih i neizvornih govornika u sledećim naučnim oblastima: biomedicina, medicina, biotehničke nauke, obrazovanje, telekomunikacija, psihologija i primenjena lingvistika.

Dužina ULS je naredna stavka koja će se sagledati u prethodnim istraživanjima. Kao što se može pretpostaviti na osnovu istraživanja koja su predstavljena u prethodnih poglavljima, najzastupljenije je istraživanje ULS od četiri reči. Istraživanja Perez-Ljatanda (2014), Lazić (2017), Gungor i Uisal (2016), Gungor (2016), Gungor (2019), Pan i saradnici (2016) i Al Fedžri i saradnici (2020) bavila su se ispitivanjem ULS od četiri reči. Istraživanje Salazarove (2010) bavilo se ispitivanjem ULS od tri reči, dok su dva istraživanja ispitala upotrebu ULS od nekoliko dužina. Naime, istraživanje Salazarove (2014) bavilo se ULS od tri, četiri, pet i šest reči, a istraživanje Esfandiari i Barbari (2017) bavilo se ispitivanjem ULS od četiri, pet i šest reči.

Pri identifikaciji ULS neophodno je postaviti minimalnu frekvenciju i disperziju u korpusu. U istraživanju Salazarove (2014) pri identifikaciji takođe je primenjen i *MI score* po uzoru na istraživanje Simson-Vlač i Elis (Simpson-Vlach & Ellis, 2010). Od istraživanja obrađenih u ovom poglavlju, najmanja minimalna frekvencija postavljena je u istraživanjima koja su sprovedi Gungor (2019) sa minimalnom frekvencijom od pet ponavljanja u milion reči i Salazarova (2014) sa minimalnom frekvencijom od deset ponavljanja u milion reči. Minimalna frekvencija od 20 ponavljanja u milion reči primenjena je u istraživanjima Perez-Ljantada (2014), Lazić (2017), Gungor i Uisal (2016) i Gungor (2016), dok je konzervativnija minimalna frekvencija od 40 ponavljanja u milion reči primenjena u istraživanjima Pan i saradnika (2016), Al Fedžri i saradnika (2020) i Esfandiari i Barbari (2017). Koliko će konzervativni kriterijumi za izdvajanje ULS biti u istraživanju Esfandiari i Barbari (2017) zavisilo je i od dužine ULS istraženih, naime, najstroži kriterijum od minimum 40 ponavljanja u milion reči postavljen je za ULS od četiri reči, dok je minimalna frekvencija za ULS od pet reči bila 20 ponavljanja u milion reči a za ULS od šest reči samo deset ponavljanja u milion reči. Ovakva je odluka doneta zbog znatno manje frekventnosti dužih ULS. Minimalna disperzija takođe varira u istraživanjima. Po uzoru na istraživanje Hajland (2008), disperzija pojavljivanja u minimum 10% tekstova u korpusu postavljena je u istraživanjima koja su uradili Perez-Ljantada (2014), Gungor i Uisal (2016), Gungor (2016) i Al Fedžri i saradnici (2020). U ostalim istraživanjima u datom poglavlju postavljeni su određeni brojevi tekstova kao minimalna disperzija. U istraživanjima koja su uradili Pan i saradnici (2016), Salazarova (2014) i Esfandiari i Barbari (2017) neophodno je da se spoj pojavi u barem pet tekstova kako bi se izdvojio kao ULS, dok je u istraživanju koje je sprovedla Lazić (2017) bilo neophodno da se spoj pojavi u barem četiri teksta. U istraživanju Salazarove (2010) nisu dati podaci o minimalnoj frekvenciji i disperziji koji su primenjeni pri identifikaciji ULS, dok u istraživanju Gungor (2019) nisu dati podaci o minimalnoj disperziji.

Krucijalna stavka za istraživanja ovog tipa za koju postoji velika varijacija je veličina korpusa. Trebalo bi napomenuti da je razlika u veličini korpusa koji se porede često problematična i unutar jednog istraživanja, a ne samo u okviru različitih istraživanja koja se bave istom temom. Kao što je ranije napomenuto, veličina korpusa je od ključnog značaja pri izdvajanju ULS zbog dokazane neefikasnosti primene normalizovanih frekvencija pri izdvajanju ULS u korpusima različitih veličina, odnosno tendencije da se u manjim korpusima pronalazi veći broj ULS (Bestgen, 2019). Kada govorimo o veličini korpusa u datom tipu istraživanja, trebalo bi napomenuti da postoje dve prakse pri formiranju korpusa: ujednačavanje broja radova koji sačinjavaju korpuse koji će se porediti ili ujednačavanje broja reči u korpusima. U svom istraživanju Cortes (2004) ističe da su ULS vrlo osetljivi na broj reči u korpusu ali ne i na broj tekstova iz kojih su izdvojeni, što bi ukazivalo na to da je primerenija praksa ujednačavanja korpusa po broju reči. Kasnije istraživanje Pan i saradnika (2020) zapravo je utvrdilo da je identifikacija ULS osetljiva i na neujednačenosti po pitanju broja tekstova koji sačinjavaju korpus, te bi idealne okolnosti bile ujednačenost i po broju reči i po broju tekstova. Tri istraživanja (Perez-Ljantada, 2014; Lazić, 2017; Güngör, 2016), pomenuta u ovom odeljku, za potrebe proučavanja formirala su tri korpusa koji su sačinjeni od naučnih radova napisanih na engleskom od strane izvornih govornika, neizvornih govornika, kao i naučnih radova koji su napisani na maternjem jeziku neizvornih govornika engleskog jezika. Najmanji korpus u ovde navedenim istraživanjima je korpus neizvornih govornika u istraživanju Salazarove (2014), koji se sastoji od ukupno 120.718 reči, dok su dva najveća korpusa u istraživanju Esfandiari i Barbari (2017) i sastoje se od 2.185.952 i 2.182.678 reči. Veličine korpusa koji se porede unutar samog istraživanja nisu usklađeni kod Salazarove (2014) i Lazić (2017), u kojima je velika razlika između upoređenih korpusa; manje razlike u veličini korpusa mogu da se nađu u istraživanjima Salazarove (2010), Perez-Ljantade (2014) i Pan i saradnika (2016), dok su odlično usklađeni po veličini korpusi u istraživanjima Gungor i Uisal (2016), Gungor (2016), Gungor (2019), Esfandiari i Barbari (2017) i Al Fedžri i saradnici (2020). Kao što se može pretpostaviti u radovima koji imaju odlično poklapanje u veličini korpusa, kao u istraživanju Pan i saradnici

(2016), primenjen je princip formiranja korpusa u kojima su korpusi ujednačeni po dužini, odnosno broju reči. U istraživanju Perez-Ljantada (2014) primenjen je princip formiranja korpusa na osnovu broja radova, iako jedan od tri korpusa ima manji broj radova od predviđenog, a u istraživanju Lazić (2017) ujednačen je broj radova u dva manja korpusa. Na osnovu navedenih podataka može se zaključiti da zaista većina istraživanja u ovom odeljku prati preporuke koje je dala Kortez (2004) i usklađuje korpusne na osnovu broja reči.

Usko povezano sa veličinom korpusa je i broj izdvojenih ULS koji je takođe povezan sa dužinom samih ULS. Zbog prirode ULS kraći spojevi će biti mnogobrojniji u korpusu uporedno sa dužim ULS. Pošto se u ovoj grupi istraživanja samo Salazarova (2010) nije bavila ULS od četiri reči, u upoređivaćemo broj izdvojenih ULS od četiri reči. Najmanji broj izdvojenih ULS od četiri reči u korpusima na engleskom jeziku je 31 ULS u korpusu izvornih govornika u istraživanju Al Fedžri i saradnici (2020), dok je najveći broj 507 ULS u korpusu neizvornih govornika u istraživanju Lazić (2017). Ovako velika razlika u broju identifikovanih ULS u dva istraživanja može se objasniti razlikama u kriterijumima za izdvajanje ULS, razlikom u veličini korpusa i razlikama u karakteristikama upotrebe ULS kod izvornih i neizvornih govornika. Lazić (2017) primenjuje minimalnu frekvenciju za izdvajanje ULS od 20 ponavljanja u milion reči uporedno sa 40 ponavljanja u milion reči (kod Al Fedžri i sar., 2020). Korpus neizvornih govornika kod Lazić (2017) je jedan od najmanjih i broji 157.179 reči, dok je korpus izvornih govornika kod Al Fedžri i saradnika (2020) veći i sadrži 1.325.986 reči. Kao što je ranije navedeno, izvorni govornici uglavnom koriste manji broj ULS uporedno sa neizvornim govornicima. Dodatan mogući razlog za razliku u broju izdvojenih ULS u dva istraživanja može da bude razlika u karakteristikama upotrebe ULS u naučnim radovima iz dve različite naučne oblasti – oblast biotehničkih nauka kod Lazić (2017) i oblast primenjene lingvistike kod Al Fedžri i sar. (2020).

S obzirom na to da je glavni cilj istraživanja posmatranih u ovom poglavlju disertacije ispitati razlike u upotrebi ULS kod izvornih i neizvornih govornika, značajno je sagledati razlike u samom broju identifikovanih ULS kod izvornih i neizvornih govornika u navedenim istraživanjima. Pri ovoj komparaciji izostavljamo istraživanje Salazarove (2014), koja nije izdvojila sve ULS u korpusu neizvornih govornika već je analizirala samo ciljne ULS koji su se nalazili u korpusu izvornih govornika. Takođe bi trebalo napomenuti da se broj ULS koji se u ovom poglavlju upoređuje u istraživanju Gungor (2019) odnosi samo na ULS sa uvodnom *it* strukturom. U svim istraživanjima, osim u Salazar (2010) i Esfandiari i Barbari (2017), pokazalo se da neizvorni govornici koriste veći broj ULS od izvornih govornika. Štaviše, i kod Esfandiari i Barbari (2017) ULS od šest reči su brojniji kod neizvornih govornika, dok su ULS od četiri i pet reči brojniji kod izvornih govornika. Nejasan je razlog za ovu razliku u rezultatima u istraživanju Esfandiari i Barbari, kao i Salazarove (2010) a uporedno sa svim drugim istraživanjima predstavljenim u ovom poglavlju, te su dodatna istraživanja neophodna da bi se došlo do odgovora.

Najveća razlika u broju izdvojenih ULS kod izvornih i neizvornih govornika može se naći u istraživanju Lazić (2017), u kojem je kod izvornih govornika identifikovano 69 ULS a kod neizvornih 507 ULS, odnosno gde postoji povećanje u upotrebi ULS kod neizvornih govornika od 635%. Međutim, može se pretpostaviti da je razlog za ovako veliku razliku u broju identifikovanih ULS značajna razlika u veličini korpusa izvornih i neizvornih govornika. Najmanje povećanje broja ULS u korpusu neizvornih govornika uporedno sa korpusima izvornih vidi se u istraživanju Pan i saradnici (2016), u kojem je povećanje 29%. Procenti uvećanja ili smanjenja broja ULS u korpusima neizvornih govornika uporedno sa korpusima izvornih govornika mogu se videti u Tabeli 20. Na osnovu podataka u Tabeli 20 može se zaključiti da ne postoji jasan trend u prethodnim istraživanjima koji bi mogao da pretpostavi koliko bi bilo uvećanje ili umanjenje broja ULS u budućim istraživanjima. Moglo bi se, doduše, spekulirati da bi broj ULS bio veći kod neizvornih

govornika pošto se uvećanje češće primećuje (u osam istraživanja) nego umanjnje (u dva istraživanja).

Tabela 20. Broj izdvojenih ULS kod izvornih i neizvornih govornika i procenat povećanja ili umanjnja broja u odnosu na korpus izvornih govornika u istraživanjima koja se bave upotrebom ULS u naučnim radovima

Istraživanje	ULS u korpusu izvornih govornika	ULS u korpusu neizvornih govornika	Povećanje ili umanjnje	Dužina ULS
<i>Salazar (2010)</i>	592	439	- 26%	3 reči
<i>Perez-Ljantada (2014)</i>	56	77	38%	4 reči
<i>Lazić (2017)</i>	69	507	635%	4 reči
<i>Gungor i Uisal (2016)</i>	32	98	206%	4 reči
<i>Gungor (2016)</i>	79	119	51%	4 reči
<i>Gungor (2019)</i>	38	66	74%	4 reči
<i>Pan i sar. (2016)</i>	55	71	29%	4 reči
<i>Esfandiari i Barbari (2017)</i>	254	116	- 54%	4 reči
	115	104	- 10%	5 reči
	47	96	47%	6 reči
<i>Al Fedžri i sar. (2020)</i>	31	51	65%	4 reči

Nakon što smo sagledali razlike u broju izdvojenih ULS, neophodno je sagledati zajedničke ULS, odnosno one koji se nalaze i kod izvornih i kod neizvornih govornika. Podaci o broju zajedničkih ULS dati su u Tabeli 21, zajedno sa informacijama o tome koliki procenat od ukupnog broja ULS predstavljaju zajednički u korpusu izvornih govornika i neizvornih govornika. U Tabeli 21 se ne nalaze podaci iz istraživanja Salazarove (2010), Salazarove (2014) i Gungor (2019) zato što Salazarova (2010) daje podatke samo o ULS koji u sebi sadrže leksičke glagole, Gungor (2019) se samo bavio ULS sa uvodnom *it* strukturom, a istraživanje Salazarove (2014) nije izdvojilo sve ULS iz korpusa neizvornih govornika. Zbog tendencije da neizvorni govornici koriste veći broj ULS procenat koji zajednički ULS čine u korpusu neizvornih govornika je uglavnom manji. U korpusu neizvornih govornika procenti zajedničkih ULS se kreću od 6% u istraživanju Lazić (2017) do 34% u istraživanju Pan i saradnika (2016). U korpusu izvornih govornika procenat zajedničkih ULS se kreće od 13% za ULS od pet reči u istraživanju Esfandiari i Barbari (2017) do 46% u istraživanju Lazić (2017). Od sedam istraživanja u kojima upoređujemo procenat zastupljenosti zajedničkih ULS u korpusima izvornih govornika, u pet istraživanja je procenat između 40% i 46%, što nam ukazuje na tendenciju u ovom tipu istraživanja. Slična tendencija sa tako uskim okvirom nije se pokazala u korpusima neizvornih govornika, ali možemo da zaključimo da u dosadašnjim rezultatima zajednički ULS nisu prelazili zastupljenost od 35% u korpusima neizvornih govornika.

Kao što je ranije pomenuto, jedna od poteškoća za poređenje rezultata u istraživanjima ULS jeste da još ne postoje univerzalno prihvaćene taksonomije za strukturalnu i funkcionalnu raspodelu ULS. Iako postoje određene taksonomije koje se često primenjuju, kao što je taksonomija za funkcionalnu raspodelu koju je predložio Hajland (2008), autori imaju tendenciju da naprave sopstvene izmene koje bi olakšale analizu njihovih podataka i specifičnih konteksta. U datom pasusu sagledaćemo koje su strukturalne taksonomije korišćene u istraživanjima u ovom poglavlju.

Tabela 21. ULS u oba korpusa i procenat koji oni čine u korpusu izvornih i neizvornih govornika u prethodnim istraživanjima usmerenim na upotrebu ULS u naučnim radovima.

Istraživanje	ULS koji se nalaze u oba korpusa	Procenat zajedničkih ULS u korpusu izvornih govornika	Procenat zajedničkih ULS u korpusu neizvornih govornika	Dužina ULS
<i>Perez-Ljantada (2014)</i>	12	21%	15%	4 reči
<i>Lazić (2017)</i>	32	46%	6%	4 reči
<i>Gungor i Uisal (2016)</i>	13	40%	13%	4 reči
<i>Gungor (2016)</i>	34	43%	28%	4 reči
<i>Pan i sar. (2016)</i>	24	43%	34%	4 reči
<i>Esfandiari i Barbari (2017)</i>	37	14%	31%	4 reči
	16	13%	15%	5 reči
	8	17%	8%	6 reči
<i>Al Fedžri i sar. (2020)</i>	14	45%	27%	4 reči

U istraživanju Salazarove (2010) se ne navodi koja se strukturalna taksonomija koristi, pa se neće diskutovati u ovom segmentu. U istraživanjima u ovom poglavlju najzastupljenija je taksonomija predložena od strane Bajber i saradnika (1999) i primenjena je u sedam od devet istraživanja. U istraživanjima Perez-Ljantada (2014) i Esfandiari i Barbari (2017) primenjena je taksonomija predložena od strane Bajber i saradnika (2004). Najupečatljivija razlika između dve zastupljene taksonomije je što u odnosu na Bajber i sar. (1999) taksonomija predložena u Bajber i sar. (2004) ima i kategoriju *fragmenti zavisne klauze* (engl. *dependant clause fragments*). U istraživanjima Salazarove (2014) i Gungor (2016) unete su modifikacije u strukturalnim podgrupama. Salazarova (2014) je dodala pet novih potkategorija:

1. *druge imeničke fraze,*
2. *druge predložke fraze,*
3. *glagolske fraze sa ličnom zamenicom we,*
4. *drugi pasivni fragmenti,*
5. *drugi glagolski fragmenti.*

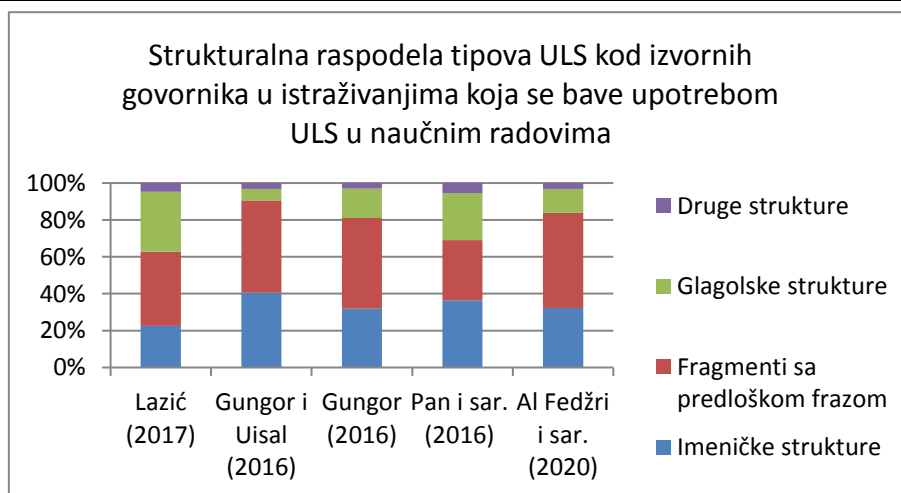
U istraživanju Gungor (2016) dodato je četiri potkategorija: *druge imeničke fraze, druge predložke fraze, drugi glagolski fragmenti i drugi pasivni fragmenti*; jedna kategorija je blago izmenjena kao u istraživanju Salazar (2014): *glagolska fraza + that klauza* zamenjeno je sa *glagolska fraza/imenička fraza + that klauza*; dok je jedna potkategorija izbačena: *kopula be + imenička fraza/pridevska fraza*. Kao što se iz navedenog može primetiti, autori su izmene uneli samo u potkategorijama a ne i u glavnim kategorijama, što olakšava poređenje rezultata, barem na nivou glavnih kategorija.

Pri poređenju strukturalne raspodele ULS kod izvornih i neizvornih govornika u naučnim radovima u ovom poglavlju, fokus će se staviti na poređenje raspodele u glavnim kategorijama. Istraživanja Salazarova (2010), Salazarova (2014), Perez-Ljantada (2014) i Gungor (2019) neće se uzeti u obzir pri ovoj diskusiji zbog specifičnog karaktera metodološke postavke u radu. A istraživanje Esfandiari i Barbari (2017) će se naknadno razmotriti zato što primenjuje drugačiju taksonomiju u odnosu na preostale radove sa kojima se poredi. Nakon gorenavedenih izdvajanja određenih radova, za detaljniju komparaciju preostaje pet radova. Podaci o strukturalnoj raspodeli iz tih pet radova prikazani su u Tabeli 22, kao i u grafikonima koji slede. Radi lakšeg poređenja rezultata, napravljena su tri grafikona: jedan u kojem se može videti *strukturalna raspodela kod izvornih i neizvornih govornika* u relevantnim istraživanjima iz ovog poglavlja (Grafikon 19), drugi u kojem se može videti *strukturalna raspodela samo kod izvornih govornika* kako bi se lakše uvideli

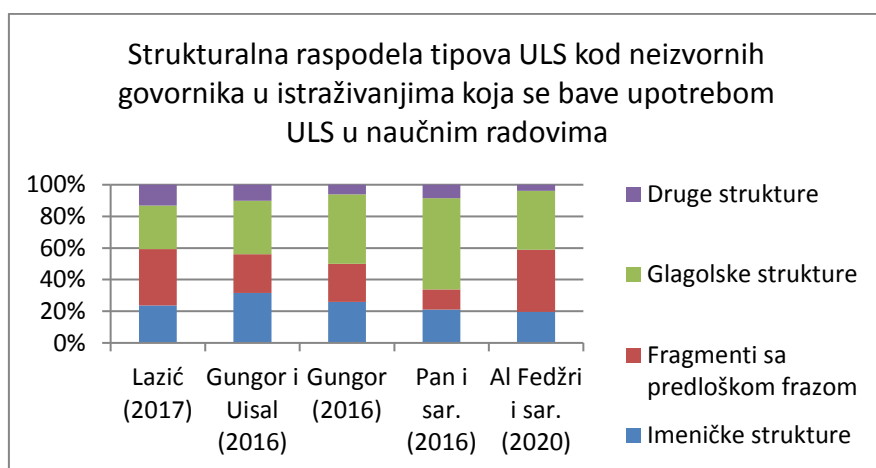
potencijalne karakteristike te grupe govornika (Grafikon 20), i treći u kojem su predstavljeni rezultati *strukturalne analize samo neizvornih govornika* kako bi se uvidele sličnosti u strukturalnoj upotrebi ULS kod neizvornih govornika u relevantnim istraživanjima iz ovog poglavlja (Grafikon 21).

Tabela 22. *Strukturalna raspodela tipova i tokena ULS kod izvornih i neizvornih govornika u istraživanjima koja se bave upotrebom ULS u naučnim radovima*

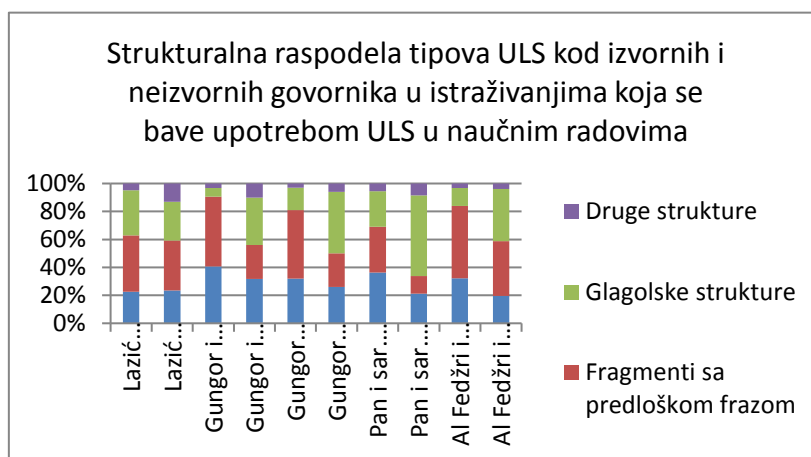
Istraživanje	Strukturalne kategorije	ULS u korpusu izvornih govornika tipovi (tokeni)	ULS u korpusu neizvornih govornika tipovi (tokeni)
<i>Lazić (2017)</i>	<i>Imeničke strukture</i>	22.5 % (20.4%)	23.6% (14.8%)
	<i>Fragment sa predloškom frazom</i>	40.4% (40.8%)	35.7% (34.7%)
	<i>Glagolske strukture</i>	32.3% (31.7%)	27.6% (18.1%)
	<i>Druge strukture</i>	4.8% (7.1%)	13.1% (32.4%)
<i>Gungor i Uisal (2016)</i>	<i>Imeničke strukture</i>	40.63% (/)	31.63% (/)
	<i>Fragment sa predloškom frazom</i>	50% (/)	24.50% (/)
	<i>Glagolske strukture</i>	6.25% (/)	33.67% (/)
	<i>Druge strukture</i>	3.12% (/)	10.2% (/)
<i>Gungor (2016)</i>	<i>Imeničke strukture</i>	32% (32%)	26% (25%)
	<i>Fragment sa predloškom frazom</i>	49% (51%)	24% (28%)
	<i>Glagolske strukture</i>	16% (14%)	44% (42%)
	<i>Druge strukture</i>	3% (3%)	6% (5%)
<i>Pan i sar. (2016)</i>	<i>Imeničke strukture</i>	36.4% (33.9%)	21.1% (18.4%)
	<i>Fragment sa predloškom frazom</i>	32.6% (33.2%)	12.7% (12.8%)
	<i>Glagolske strukture</i>	25.5% (26.6%)	57.7% (55.8%)
	<i>Druge strukture</i>	5.5% (6.3%)	8.5% (13%)
<i>Al Fedžri i sar. (2020)</i>	<i>Imeničke strukture</i>	32.26% (31.56%)	19.61% (19.10%)
	<i>Fragment sa predloškom frazom</i>	51.61% (52.67%)	39.22% (39.84%)
	<i>Glagolske strukture</i>	12.90% (10.89%)	37.25% (37.83%)
	<i>Druge strukture</i>	3.23% (4.89%)	3.92% (4.23%)



Grafikon 19. *Strukturalna raspodela tipova ULS kod izvornih i neizvornih govornika engleskog jezika u istraživanjima koja se bave upotrebom ULS u naučnim radovima*



Grafikon 20. *Strukturalna raspodela tipova ULS kod izvornih govornika engleskog jezika u istraživanjima koja se bave upotrebom ULS u naučnim radovima*



Grafikon 21. *Strukturalna raspodela tipova ULS kod neizvornih govornika engleskog jezika u istraživanjima koja se bave upotrebom ULS u naučnim radovima*

Na osnovu podataka iz Tabele 22 može se uvideti da se raspodela ULS po strukturalnim kategorijama kod izvornih i neizvornih govornika poklapa u istraživanju Lazić (2017). U tom istraživanju i kod izvornih i kod neizvornih govornika najzastupljenija strukturalna grupa su *ULS sa predloškom frazom*, praćena *ULS sa glagolskom frazom* i *ULS sa imeničkom frazom*, dok je grupa *druge strukture* najmanje zastupljena u oba korpusa. Slična je raspodela kod izvornih i neizvornih govornika po strukturalnim kategorijama i u istraživanju Al Fedžri i saradnika (2020), u kojem je za obe grupe govornika najzastupljenija kategorija *ULS sa predloškom frazom* i najmanje zastupljena kategorija *druge strukture*. Međutim, kod izvornih govornika druga najzastupljenija grupa je *ULS sa imeničkom frazom*, dok je kod neizvornih govornika druga najzastupljenija grupa *ULS sa glagolskom frazom*. Ovakvi rezultati mogu da se objasne tendencijom neizvornih govornika da se češće služe *ULS sa glagolskim frazama*, dok im je znanje stranog jezika još u fazi razvoja. U istraživanjima Gungor i Uisal (2016), Gungor (2016) i Pan i saradnika (2016) razlikuju se najzastupljenije strukturalne kategorije kod izvornih i neizvornih govornika. U istraživanjima Gungor i Uisal (2016) i Gungor (2016) najzastupljenija strukturalna kategorija kod izvornih govornika je *ULS sa predloškom frazom* a u istraživanju Pan i saradnici (2016) najzastupljenija kategorija kod izvornih govornika je *ULS sa imeničkom frazom*, dok je za sva tri istraživanja kod neizvornih govornika najzastupljenija kategorija *ULS sa glagolskom frazom*. Moglo bi se zaključiti da su neizvorni govornici u ova tri istraživanja i dalje u procesu

savladavanja normi akademskog pisanja na engleskom jeziku pošto se i dalje najčešće oslanjaju na *ULS sa glagolskom frazom*.

Sa prethodnim nalazima upoređićemo i nalaze o strukturalnoj raspodeli i iz istraživanja Esfandiari i Barbari (2017). U tom istraživanju glavne kategorije koje se ispituju su iste kao i u gorenavedenim istraživanjima, sa dodatkom jedne kategorije – *fragmenti zavisne klauze*. U istraživanju Esfandiari i Barbari (2017) za ULS od četiri reči kod izvornih govornika najzastupljenija kategorija je *ULS sa glagolskom frazom*, što se ne poklapa ni sa jednim istraživanjem iz Tabele 22. Interesantno je da se u tom istraživanju procentualno poklapaju dve kategorije koje su ujedno i najzastupljenije, a to su *ULS sa glagolskom frazom* i *ULS sa predloškom frazom*. Kod ULS od pet i šest reči u datom istraživanju, drugačija je strukturalna raspodela. I kod izvornih i kod neizvornih govornika, u ULS od pet i šest reči najzastupljenija strukturalna kategorija je *ULS sa imeničkom frazom*. Kao i pri ranijim poređenjima, istraživanje Esfandiari i Barbari (2017) je specifično i rezultati se uglavnom ne poklapaju sa ostalim istraživanjima u ovom poglavlju. Detaljni podaci o strukturalnoj raspodeli u istraživanju Esfandiari i Barbari (2017) može se videti u Tabeli 23 sa osnovnim podacima o tom istraživanju.

U istraživanjima u ovom poglavlju mahom se koristi funkcionalna taksonomija predložena od strane Hajland (2008) u originalu ili sa modifikacijama. Jedino istraživanje Perez-Ljantada (2014) koristi funkcionalnu taksonomiju za raspodelu ULS koju prelaže Kortež (2004). Istraživanja Pan i saradnici (2016), Gungor i Uisal (2016) i Al Fedžri i saradnici (2020) koriste originalnu nemodifikovanu funkcionalnu taksonomiju kao što je data u istraživanju Hajland (2008). Istraživanje Salazarove (2014) modifikuje taksonomiju koju je predložio Hajland (2008) i ta modifikovana verzija primenjuje se i u istraživanjima Lazić (2017) i Gungor (2016), dok u istraživanju Esfandiari i Barbari (2017) autori unose sopstvene izmene u originalnu taksonomiju, naime, uvode jednu dodatnu potkategoriju – *statistički označitelji* (engl. *statistical markers*).

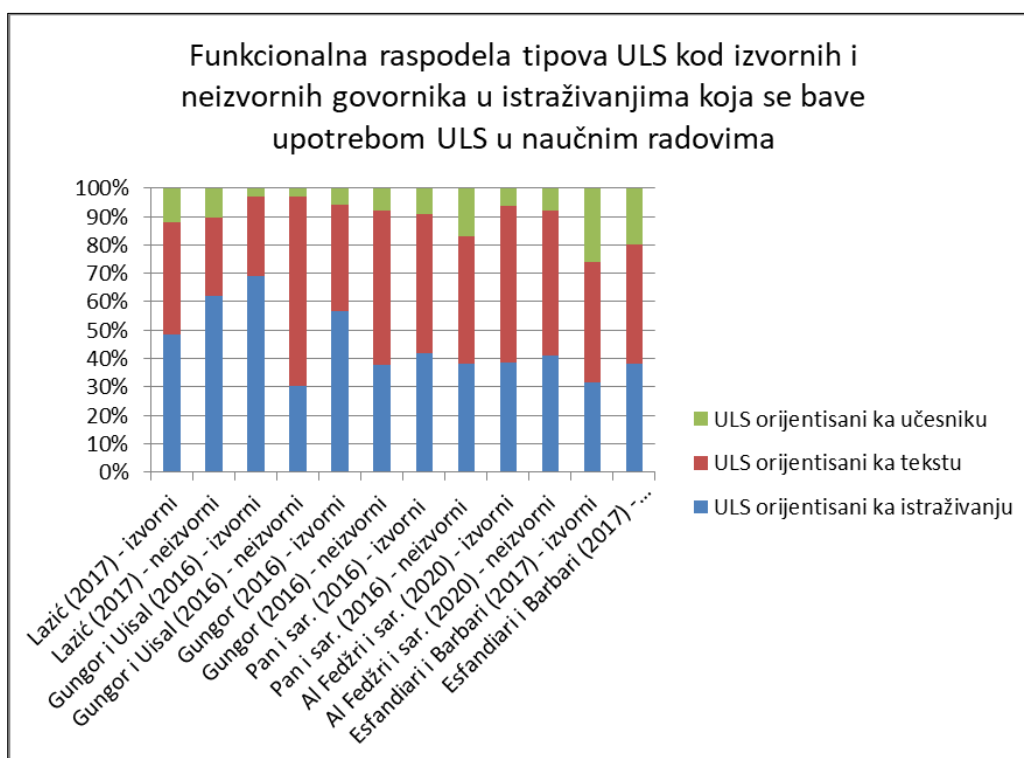
Kao što je slučaj kod strukturalnih taksonomija tako i kod funkcionalnih, modifikacije su na nivou potkategorija, što olakšava poređenje istraživanja u glavnim kategorijama. Istraživanja Salazarova (2010), Salazarova (2014), Perez-Ljantada (2014) i Gungor (2019) neće se uzeti u obzir pri komparaciji funkcionalnih raspodela ULS kod izvornih i neizvornih govornika u ovom delu disertacije zbog specifičnog karaktera metodološke postavke u radovima. Preostaju nam šest istraživanja koja će se uporediti na osnovu funkcionalne raspodele ULS kod izvornih i neizvornih govornika. Detaljni podaci o funkcionalnoj raspodeli u radovima koji će se porediti nalaze se u Tabeli 23, v. i grafikone dole.

Tabela 23. *Funkcionalna raspodela tipova i tokena ULS kod izvornih i neizvornih govornika u istraživanjima koja se bave upotrebom ULS u naučnim radovima*

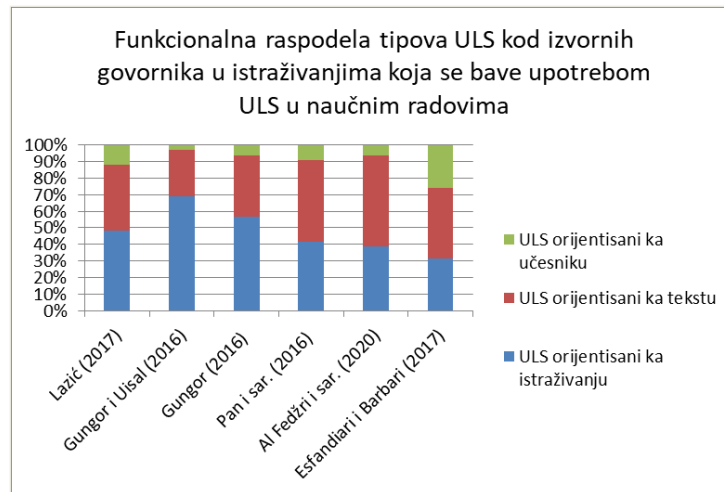
Istraživanje	Funkcionalne kategorije	ULS u korpusu izvornih govornika tipovi (tokeni)	ULS u korpusu neizvornih govornika tipovi (tokeni)	Dužina ULS
<i>Lazić (2017)</i>	<i>ULS orijentisani ka istraživanju</i>	48.3% (48.9%)	62.1% (57.5%)	4 reči
	<i>ULS orijentisani ka tekstu</i>	39.6% (42.5%)	27.5% (31.8%)	
	<i>ULS orijentisani ka učesniku</i>	12.1% (8.6%)	10.4% (10.7%)	
<i>Gungor i Uisal (2016)</i>	<i>ULS orijentisani ka istraživanju</i>	68.80% (/)	30.09% (/)	4 reči
	<i>ULS orijentisani ka tekstu</i>	28% (/)	66% (/)	

Tabela 23. Funkcionalna raspodela tipova i tokena ULS kod izvornih i neizvornih govornika u istraživanjima koja se bave upotrebom ULS u naučnim radovima

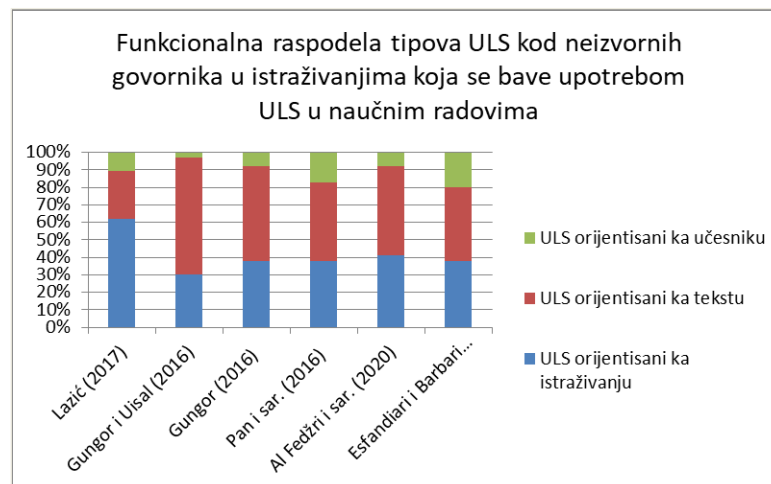
	ULS orijentisani ka učesniku	3.10% (/)	3.10% (/)	
Gungor (2016)	ULS orijentisani ka istraživanju	56% (55%)	38% (28%)	4 reči
	ULS orijentisani ka tekstu	37% (38%)	55% (64%)	
	ULS orijentisani ka učesniku	6% (7%)	8% (7%)	
Pan i sar. (2016)	ULS orijentisani ka istraživanju	41.82% (43%)	38.03% (38%)	4 reči
	ULS orijentisani ka tekstu	49.09% (49%)	45.07% (49%)	
	ULS orijentisani ka stavu	9.09% (8%)	16.90% (13%)	
Al Fedžri i sar. (2020)	ULS orijentisani ka istraživanju	38.71% (36.59%)	41.18% (35.18%)	4 reči
	ULS orijentisani ka tekstu	54.84% (57.30%)	50.98% (54.97%)	
	ULS orijentisani ka učesniku	6.45% (6.11%)	7.84% (9.23%)	
Esfandiari i Barbari (2017)	ULS orijentisani ka istraživanju	31.50% (/)	38% (/)	4 reči
	ULS orijentisani ka tekstu	42.50% (/)	42.20% (/)	
	ULS orijentisani ka učesniku	26% (/)	19.80% (/)	



Grafikon 22. Funkcionalna raspodela tipova ULS kod izvornih i neizvornih govornika engleskog jezika u istraživanjima koja se bave upotrebom ULS u naučnim radovima



Grafikon 23. *Funkcionalna raspodela tipova ULS kod izvornih govornika engleskog jezika u istraživanjima koja se bave upotrebom ULS u naučnim radovima*



Grafikon 24. *Funkcionalna raspodela tipova ULS kod neizvornih govornika engleskog jezika u istraživanjima koja se bave upotrebom ULS u naučnim radovima*

Na osnovu podataka iz Tabele 23 može se uvideti da se raspodela ULS po funkcionalnim kategorijama kod izvornih i neizvornih govornika poklapa u istraživanjima Lazić (2017), Pan i saradnici (2016), Al Fedžri i saradnici (2020) i Esfandiari i Barbari (2017) a kada je reč o ULS od četiri reči. U istraživanju Lazić (2017) kod izvornih i neizvornih govornika najzastupljenija funkcionalna kategorija je *ULS orijentisani ka istraživanju* praćena *ULS orijentisani ka tekstu*, dok je najmanje zastupljena kategorija *ULS orijentisani ka učesniku*. U istraživanjima Pan i saradnici (2016), Al Fedžri i saradnici (2020) i Esfandiari i Barbari (2017) najzastupljenija kategorija kod izvornih i neizvornih govornika je *ULS orijentisani ka tekstu* praćena *ULS orijentisani ka istraživanju*, dok je opet najmanje zastupljena kategorija *ULS orijentisani ka učesniku*. U istraživanjima Gungor i Uisal (2016) i Gungor (2016) funkcionalna raspodela ULS po glavnim funkcionalnim kategorijama se razlikuje kod izvornih i neizvornih govornika, naime, u oba istraživanja kod izvornih govornika najzastupljenija kategorija je *ULS orijentisani ka istraživanju*, najzastupljenija kategorija kod neizvornih govornika je *ULS orijentisani ka tekstu*, dok je u oba istraživanja najmanje zastupljena kategorija *ULS orijentisani ka učesniku*. Može se zaključiti da je u većini upoređenih istraživanja u ovom poglavlju došlo do poklapanja raspodele ULS po glavnim funkcionalnim kategorijama kod izvornih i neizvornih govornika. Takođe se može pretpostaviti da

najzastupljenija funkcionalna kategorija zavisi od oblasti istraživanja zato što se nije uočio upadljiv trend među upoređenim istraživanjima u ovom poglavlju.

U ovom poglavlju sagledali smo i uporedili najistaknutija istraživanja koja su se bavila upotrebom ULS u naučnim radovima kod izvornih i neizvornih govornika. U gorenavedenim istraživanjima ispita se uticaj sedam maternjih jezika (filipinski, španski, srpski, turski, kineski, persijski i indonežanski) i sedam naučnih oblasti (biomedicina, medicina, biotehničke nauke, obrazovanje, telekomunikacija, psihologija i primenjena lingvistika), kao i jedan rad koji je poredio 12 naučnih oblasti. Od ukupno deset istraživanja koja se detaljno obrađuju u ovom poglavlju, sedam istraživanja ispituju upotrebu ULS od četiri reči, jedno istraživanje se bavi ULS od tri reči, a dva istraživanja se bave ULS sa različitim dužinama.

Minimalna frekvencija za izdvajanje ULS razlikovala se u navedenim istraživanjima. Najzastupljenija minimalna frekvencija je 20 ponavljanja u milion reči koja je primenjena u četiri istraživanja. Sledeća najzastupljenija minimalna frekvencija je od 40 ponavljanja u milion reči koja je zastupljena u tri istraživanja, dok preostala dva istraživanja primenjuju manje minimalne frekvencije (od pet i deset ponavljanja u milion reči). Podjednako je neusklađena i određena minimalna disperzija u navedenim istraživanjima. Četiri istraživanja su postavila minimalnu disperziju da bude pojavljivanje u barem 10% tekstova, u četiri istraživanja je postavljeno da minimalna disperzija bude u četiri ili pet tekstova, dok za dva istraživanja nije naznačena minimalna frekvencija.

Varijacije u istraživanjima su se nastavile i kada je reč o veličini korpusa, kako između različitih istraživanja tako i unutar istih. U navedenim istraživanjima najmanji korpus sastojao se od 120.718 reči dok se najveći sastojao od 2.185.952 reči. Odlično su uklopljene veličine korpusa unutar istraživanja u pet istraživanja, manje se razlike u veličini korpusa unutar istraživanja mogu uočiti unutar tri istraživanja, dok se veće razlike u korpusima unutar istraživanja mogu naći u dva istraživanja. Dobro je, međutim, da je većina istraživanja (šest) ujednačila veličinu korpusa na osnovu broja reči a ne broja radova, kao što je i preporučeno za istraživanja koja se bave ULS (Cortes, 2004).

Na osnovu svih gorenavedenih razlika može se pretpostaviti da će doći do razlika i u broju izdvojenih ULS. U navedenim istraživanjima najmanji broj izdvojenih ULS je 31, a najveći je 507. Razlike u broju identifikovanih ULS u istraživanjima mogu se objasniti razlikama u drugim već navedenim parametrima kao što su minimalna frekvencija i disperzija, veličina korpusa, istraživačka oblast i maternji jezik neizvornih govornika. Osim u dva, u svim ostalim istraživanjima diskutovanim u ovom poglavlju, identifikovan je veći broj ULS kod neizvornih govornika. Zašto je u istraživanjima Salazar (2010) i Esfandiari i Barbari (2017) veći broj ULS identifikovan kod izvornih govornika je nejasno i zahteva dalja istraživanja. Poređenje broja zajedničkih ULS kod izvornih i neizvornih govornika u istraživanjima u ovom poglavlju pokazalo je zanimljivu tendenciju kod izvornih govornika da se procenat zajedničkih ULS kreće u opsegu od 40% do 46%. Kod neizvornih govornika se nije pokazala slična tendencija, ali se može pretpostaviti da je značajno da procenat zajedničkih ULS ne ide iznad 35%.

Kada je reč o strukturalnoj raspodeli, većina istraživanja koristi taksonomiju koju su predložili Bajber i saradnici (1999), iako se taksonomija Bajber i saradnika (2004) primenjuje u dva istraživanja. U dva istraživanja autori su unosili svoje izmene, međutim, one su se zadržavale na nivou potkategorija. Rezultati u istraživanjima koja su obrađena u ovom poglavlju pokazuju da je strukturalna raspodela kod izvornih i neizvornih govornika slična u tri istraživanja, a da razlike u preostalim istraživanjima uglavnom nastaju zbog tendencije neizvornih govornika da se više oslanjaju na *ULS sa glagolskom frazom*. Za funkcionalnu raspodelu sva istraživanja u ovom

poglavlju osim jednog su koristila taksonomiju koju je predložio Hajland (2008), iako su postojale manje izmene u četiri istraživanja. U četiri navedena istraživanja u ovom poglavlju došlo je do poklapanja u funkcionalnoj raspodeli kod izvornih i neizvornih govornika, što ukazuje na veći stepen poklapanja uporedno sa strukturalnom raspodelom. Može se pretpostaviti da će od naučne oblasti zavisiti koja je najzastupljenija funkcionalna kategorija zato što se u podacima ne vidi jasna tendencija ka jednoj kategoriji.

3. Metodološki okvir istraživanja

3.1. Predmet, obim, istraživačko pitanje i ciljevi istraživanja

Predmet istraživanja ovog rada su razlike u upotrebi učestalih leksičkih spojeva kod izvornih i neizvornih govornika engleskog jezika u naučnim radovima iz oblasti obrazovanja. Istraživanje obuhvata analizu upotrebe ULS u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca u tri korpusa, od po pola miliona reči. Korpusi su sačinjeni od naučnih radova u oblasti obrazovanja napisanih od strane izvornih govornika engleskog jezika, neizvornih govornika engleskog jezika kojima je maternji jezik srpski i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski. Istraživanje takođe obuhvata identifikaciju zajedničkih karakteristika upotrebe ULS kod dve grupe neizvornih govornika, kao i uticaj negativnog transfera na njihovu produkciju u naučnim radovima iz oblasti obrazovanja.

Glavno istraživačko pitanje glasi: Koje su razlike u karakteristikama upotrebe ULS u naučnim radovima iz oblasti obrazovanja kod izvornih i neizvornih govornika engleskog jezika u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca?

Primarni cilj istraživanja: Ustanoviti razlike u karakteristikama upotrebe ULS u naučnim radovima iz oblasti obrazovanja kod izvornih i neizvornih govornika engleskog jezika u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca.

U dosadašnjim istraživanjima upotrebe ULS kod izvornih i neizvornih govornika u naučnim radovima teško je porediti rezultate istraživanja i doći do zajedničkih karakteristika neizvornih govornika engleskog jezika sa različitim maternjim jezicima, prevashodno zbog razlika u veličinama korišćenih korpusa, načinima formiranja korpusa (ujednačenost po veličini i/ili broju radova), parametrima za izdvajanje ULS (minimalna frekvencija i disperzija), kao i naučnim oblastima koje se u radovima obrađuju. Sprovedenjem istraživanja u kojem se poredi dve različite grupe neizvornih govornika engleskog jezika sa izvornim govornicima engleskog jezika u kompatibilnim korpusima, u istom registru, žanru i oblasti sa istim parametrima za izdvajanje ULS, možemo početi da gradimo jasnije razumevanje upotrebe ULS kod neizvornih govornika, tačnije – utvrditi njihove zajedničke karakteristike, kao i negativan uticaj transfera na njihovu produkciju. Temeljnije razumevanje karakteristika upotrebe ULS kod neizvornih govornika, profesorima engleskog jezika će omogućiti da unaprede svoju praksu, prevodiocima u ovoj oblasti da unaprede svoj rad, a sa krajnjim ciljem – da naučnici koji su neizvorni govornici engleskog jezika uspešnije pokazuju pripadnost relevantnoj delatnoj zajednici i uspešnije učestvuju u globalizovanom akademskom svetu.

Sekundarni ciljevi istraživanja su:

1. ustanoviti karakteristike upotrebe ULS u datom žanru i oblasti kod izvornih govornika engleskog jezika (zastupljenost, strukturalna i funkcionalna raspodela, najzastupljeniji obrasci);
2. ustanoviti karakteristike upotrebe ULS u datom žanru i oblasti kod izvornih govornika srpskog i izvornih govornika rumunskog jezika pri njihovoj produkciji na engleskom jeziku;
3. uporediti upotrebu ULS u datom žanru i oblasti kod izvornih i dve grupe neizvornih govornika;
4. uporediti upotrebu ULS između dve grupe neizvornih govornika;

5. ustanoviti uticaj transfera na upotrebu ULS kod dve grupe neizvornih govornika engleskog jezika;
6. izdvojiti smernice za podučavanje ULS u okviru nastave akademskog engleskog jezika i podučavanja/usavršavanja prevodilaca za rad u datom žanru i oblasti.

3.2. Hipoteze istraživanja

Osnovna hipoteza istraživanja glasi:

Karakteristike upotrebe ULS u naučnim radovima iz oblasti obrazovanja će se razlikovati kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski, u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca.

Radne hipoteze su sledeće:

1. Strukturalna raspodela ULS kod izvornih govornika engleskog jezika u naučnim radovima u oblasti obrazovanja biće okarakterisana većom zastupljenošću ULS iz strukturalne grupe *fragmenti sa predložkom frazom* uporedno sa drugim strukturalnim grupama.
2. Strukturalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski u naučnim radovima u oblasti obrazovanja biće okarakterisana većom zastupljenošću ULS iz strukturalnih grupa *fragmenti sa predložkom frazom* i *glagolske strukture* nego *imeničke strukture* i *druge strukture*.
3. Funkcionalna raspodela ULS kod izvornih govornika engleskog jezika i dve grupe neizvornih govornika engleskog jezika biće okarakterisana učestalijom upotrebom ULS iz funkcionalnih grupa *ULS orijentisani ka istraživanju* i *ULS orijentisani ka tekstu* nego *ULS orijentisani ka učesniku*.
4. Veći broj tipova i tokena ULS biće identifikovan kod dve grupe neizvornih govornika engleskog jezika uporedno sa izvornim govornicima engleskog jezika u naučnim radovima iz oblasti obrazovanja.
5. Sličan broj tipova i tokena biće identifikovan kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski u naučnim radovima iz oblasti obrazovanja.
6. Postojeće statistički značajne razlike u strukturalnoj i funkcionalnoj raspodeli ULS na grupe i podgrupe između izvornih i neizvornih govornika engleskog jezika, kao i između dve grupe neizvornih govornika.
7. ULS koji se koriste zbog uticaja transfera sačinjavaće mali procenat (do 40%) ULS identifikovanih u korpusima neizvornih govornika engleskog jezika kojima su maternji jezici srpski ili rumunski u naučnim radovima iz oblasti obrazovanja.
8. Najzastupljeniji okviri će se drugačije koristiti kod izvornih govornika engleskog jezika i dve grupe neizvornih govornika engleskog jezika u naučnim radovima u oblasti obrazovanja.

Obrazloženje radnih hipoteza

Prva radna hipoteza proverava navode iz prethodnih istraživanja (Güngör & Uysal, 2016; Güngör, 2016; Lazić, 2017; Al Fedžri et al., 2020), koja su se bavila upotrebom ULS kod izvornih i neizvornih govornika engleskog jezika i u kojima se pokazalo da izvorni govornici engleskog jezika u naučnim radovima koriste najviše ULS u strukturalnoj grupi *fragmenti sa predloškom frazom*.

Zbog različitih rezultata koji se mogu naći u literaturi za strukturalnu raspodelu ULS kod neizvornih govornika engleskog jezika bilo je teže formirati drugu radnu hipotezu nego prvu. U tom procesu, prvo smo sagledali rezultate u prethodnom istraživanju u kom se ispitivala produkcija neizvornih govornika engleskog jezika kojima je maternji jezik srpski (Lazić, 2017), a u kojem je najzastupljenija strukturalna kategorija *fragmenti sa predloškom frazom* praćeni *glagolskim strukturama*. Veća se važnost pridala ovom istraživanju pri formiranju hipoteze jer su u pitanju izvorni govornici srpskog, te se mogu očekivati slični rezultati zbog uticaja istog maternjeg jezika. Sledeće, sagledali smo istraživanja upotrebe ULS kod neizvornih govornika engleskog jezika u naučnim radovima iz oblasti obrazovanja (Güngör & Uysal, 2016; Güngör, 2016), u kojima su se *glagolske strukture* pokazale kao najzastupljenija kategorija. Pri formiranju hipoteze, navedena dva istraživanja su izdvojena kao posebno relevantna jer se poklapaju po naučnoj oblasti iz koje se pišu naučni radovi, te se opet očekuju sličniji rezultati zbog uticaja normi koje vladaju u navedenoj delatnoj zajednici.

Treća radna hipoteza proverava navode iz prethodnih istraživanja (Pan i sar., 2016; Güngör & Uysal, 2016; Güngör, 2016; Lazić, 2017; Esfandiari i Barbari, 2017; Al Fedžri et al., 2020), u kojima se funkcionalne grupe *ULS orijentisani ka istraživanju* i *ULS orijentisani ka tekstu* češće upotrebljavaju od grupe *ULS orijentisani ka učesniku* i kod izvornih govornika engleskog jezika i kod neizvornih govornika engleskog jezika.

U četvrtoj radnoj hipotezi se proveravaju navodi iz prethodnih istraživanja (Perez-Ljantada, 2014; Pan i sar., 2016; Güngör & Uysal, 2016; Güngör, 2016; Lazić, 2017; Al Fedžri et al., 2020; Kostromitina, 2022), u kojima se utvrdilo da neizvorni govornici engleskog jezika koriste veći broj tipova i tokena ULS u naučnim radovima. Od deset dosadašnjih relevantnih istraživanja koja su se bavila istraživanjem upotrebe ULS kod izvornih i neizvornih govornika engleskog jezika u naučnim radovima, samo se u dva (Salazar, 2010; Esfandiari & Barbari, 2017) izdvojio veći broj tipova i tokena kod izvornih govornika engleskog jezika.

Koliko je autorki ove disertacije poznato, ne postoji prethodno istraživanje u kojem se poredi upotreba ULS u naučnim radovima kod dve ili više grupa neizvornih govornika sa različitim maternjim jezicima u kojima su korpusi i parametri za izdvajanje ULS izjednačeni. Međutim, pošto se u prethodnim istraživanjima (Perez-Ljantada, 2014; Pan i sar., 2016; Güngör & Uysal, 2016; Güngör, 2016; Lazić, 2017; Al Fedžri et al., 2020; Kostromitina, 2022), koja su poredila izvorne i neizvorne govornike engleskog jezika u naučnim radovima, pokazala zajednička tendencija neizvornih govornika da koriste veći broj tipova i tokena od izvornih govornika, u petoj hipotezi u ovom istraživanju pretpostavljamo da će neizvorni govornici kojima je maternji jezik srpski i neizvorni govornici kojima je maternji jezik rumunski koristiti sličan broj tipova i tokena ULS u naučnim radovima iz oblasti obrazovanja.

Šesta radna hipoteza proverava navode prethodnih istraživanja (Pan i sar., 2016; Güngör, 2016; Lazić, 2017; Esfandiari & Barbari, 2017; Al Fedžri et al., 2020), u kojima je identifikovana statistički značajna razlika u strukturalnoj i funkcionalnoj raspodeli ULS na grupe i podgrupe između izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika.

U sedmoj radnoj hipotezi se proveravaju navodi iz prethodnih istraživanja (Güngör & Uysal, 2020; Kostromitina, 2022; Ozturk i Tasci, 2023), u kojima se ULS pod negativnim uticajem transfera sačinjavali mali procenat svih identifikovanih ULS, odnosno u kojima ULS pod negativnim uticajem transfera nisu sačinjavali više od 40% svih ULS u korpusu neizvornih govornika.

Poslednja, osma radna hipoteza, proverava navode iz prethodnih istraživanja (Pan i sar., 2016; Lazić, 2017), kojima se pokazalo da izvorni i neizvorni govornici engleskog jezika drugačije koriste najzastupljenije obrasce ULS u pogledu učestalosti upotrebe i korišćenih reči sa punim leksičkim značenjem.

3.3. Korpusi

Za potrebe istraživanja formirana su tri korpusa sačinjena od naučnih radova iz oblasti obrazovanja. S obzirom na činjenicu da su korpusi sastavljeni sa namenom istraživanja ULS, određena je minimalna dužina korpusa od pola miliona reči. Navedena dužina je odabrana na osnovu primera iz prethodnih istraživanja kao što su Pan i sar. (Pan et al., 2016) i Gungor i Uisal (Güngör & Uysal, 2016), kao i činjenice da manji korpusi nisu prikladni za istraživanje ULS zbog distorzije u rezultatima koja nastaje usled male veličine korpusa (Bestgen, 2019). Nije se odlučilo za veće korpusne tokom istraživanja zbog ograničenog broja istraživača u Srbiji koji specijaliziraju u oblasti obrazovanja i koji objavljuju radove na engleskom jeziku. Od tri maternja jezika koji se ispituju u ovom istraživanju, u svetu su najmanje zastupljeni izvorni govornici srpskog.

Kako bi se istražila upotreba ULS kod različitih grupa neizvornih govornika za potrebe istraživanja, konstruisan je korpus obrazovnih naučnih radova kojima je maternji jezik srpski (SCEP – *Serbian Corpus of Educational Papers*), korpus obrazovnih naučnih radova kojima je maternji jezik rumunski (RCEP – *Romanian Corpus of Educational Papers*) i korpus obrazovnih naučnih radova izvornih govornika za poređenje rezultata (ECEP – *English Corpus of Educational Papers*). Pri određivanju maternjeg jezika autora za sva tri korpusa, koristila se radna definicija koju je postavio Vud (Wood, 2001). Definicija nalaže da je „izvorni govornik engleskog” bilo koji autor u afilijaciji sa institucijom u kojoj se engleski govori kao zvanični jezik i koji takođe ima ime i prezime koji se smatraju izvornim za zemlje engleskog govornog područja³⁶ (Pan et al., 2016: 63). Tako, na primer, jedan od radova u korpusu SCEP napisala je Jasmina Stolić (2015), čije je ime i prezime tipično za govorno područje Srbije i koja radi u instituciji u Srbiji. Za potrebe ovog istraživanja bilo je neophodno da prvi autor ispuni gorenavedenu radnu definiciju u radovima u kojima postoje i koautori.

Pitanje identifikacije izvornih govornika pri kompilaciji relativno velikih korpusa koji su potrebni za istraživanje ULS kod istraživača smatra se dosta problematičnim (Al Fajri et al., 2020; Esfandiari & Barbary, 2017). Može se pretpostaviti da je upravo ovo pitanje jedan od razloga zbog kojeg je do sada upotreba ULS češće istražena u produkciji studenata u oblasti akademskog engleskog.

Prethodna istraživanja ULS u naučnim radovima na ograničen broj načina su rešavala ovaj problem. Pan, Rapen i Bajber (Pan, Reppen & Biber, 2016), Esfandiari i Barbari (Esfandiari & Barbary, 2017) i Al Fedžri, Veisito, Kirana (Al Fajri, Wasito & Kirana, 2020) primenjuju radnu definiciju koju je predložio Vud (Wood, 2001). Salasar (Salazar, 2014) pri identifikaciji izvornih govornika engleskog primenjuje istu radnu definiciju ali ne pominje da je preuzeta od Vuda (Wood, 2001) ili bilo kog drugog istraživanja. Mnoga istraživanja navode samo naučne časopise iz različitih država kao izvor naučnih radova iz raznolikih oblasti različitih izvornih govornika bez davanja dodatnih

³⁶Engleski original citata: „...’L1 English’ writers to be any author affiliated with an institution in a country where English is spoken as the first language who also has a first and last name that can be considered native to English-speaking countries.”

detalja o procesu identifikacije (Güngör, 2016, 2019; Güngör & Uysal, 2016; Salazar, 2010; Lazić, 2017). U istraživanju Perez-Ljantada (Pérez-Llantada, 2014) navodi se da su radovi koji sačinjavaju korpus izvornih govornika engleskog, neizvornih govornika engleskog i korpusa na španskom jeziku izdvojeni iz 12 naučnih oblasti koje su ravnomerno raspoređene i da radovi koji sačinjavaju korpus neizvornih govornika nisu prevodi već da su ih sami istraživači pisali. Međutim, ne dobijamo detaljne informacije o tome kako su ovi parametri obezbeđeni.

Prethodna istraživanja uvažila su činjenicu da se primenom radne definicije izvornih govornika koju je predložio Vud (Wood, 2001) i sličnim metodama ne pruža potpuna sigurnost tačnosti pretpostavljenog izvornog jezika istraživača (Al Fajri et al., 2020; Esfandiari & Barbary, 2017). Međutim, koliko je autoru poznato, i dalje nije osmišljeno bolje rešenje za kompilaciju relativno obimnih korpusa u ovoj oblasti istraživanja.

Dodatna stavka pri ovakvoj proceduri građenja korpusa neizvornih govornika jeste činjenica da u procesu objavljivanja naučnih radova oni prolaze kroz proces lekture, koji do određene mere može uticati na rezultate istraživanja. Pojedini autori su se postarali da obezbede nelektorisane verzije radova neizvornih govornika, međutim, ti korpusi su znatno manje veličine (Lazić, 2017; Salazar, 2014). Pošto ULS nisu leksičke jedinice koje se mogu identifikovati bez primene korpusno-lingvističkih alata za identifikaciju frekventnosti i distribucije i pošto uglavnom nisu inherentno prepoznatljivi govornicima (Biber i sar., 1999), stav je istraživača da su izmene u primeni konkretnih ULS u procesu modifikacije minimalne. Najverovatnija posledica lekture bi mogla da bude i pozitivna za identifikaciju ULS s obzirom na to da umanjuje broj gramatičkih grešaka za koje se smatra da su potencijalan razlog zbog kojeg se u produkciji neizvornih govornika (najčešće studenata) identifikuje manji broj ULS (Shin, 2019).

Još jedan potencijalan problem je pitanje prevoda istraživanja. Logično je pretpostaviti da nemaju svi istraživači koji objavljuju svoje radove u inostranim i domaćim internacionalnim časopisima na engleskom jeziku dovoljno dobro znanje stranog jezika da na njemu pišu naučne radove i neophodne su im usluge prevodilaca. Za jezike kao što su srpski i rumunski, koji imaju relativno mali broj izvornih govornika i koji nisu popularni jezici za učenje kao strani, takođe se može pretpostaviti da će imati veći broj prevodilaca koji su izvorni govornici srpskog i rumunskog a kojima je engleski jezik strani nego što bi to bilo u obrnutom slučaju. To bi značilo da i ako jesu prevodi u pitanju i dalje su u pitanju izvorni govornici koji bi, takođe, morali da budu upoznati sa normama žanra akademskog engleskog kako bi bili kompetentni za dotični posao. Stoga, ako se desi da su neki od radova u korpusu prevodi, to ne bi trebalo značajno da utiče na rezultate istraživanja jer su i prevodioci izvorni govornici ciljanih jezika.

Pri kompilaciji korpusa koristili su se naučni članci koji su besplatno dostupni onlajn, bilo preko sajta časopisa koji je otvorenog pristupa, bilo putem platformi poput *Kobson*, *ResearchGate*, *Academiaedu* itd. Akademski pretraživač *Google Scholar* je bio korišćen pri traženju naučnih članaka iz oblasti obrazovanja. Izdvojeni su naučni radovi koji su objavljeni u periodu od 2000. do 2020. godine, što dalje znači da su radovi objavljeni pre nego što je veštačka inteligencija postala javno dostupna. Naveden podatak je važno istaći jer će postojanje naučnih tekstova koji su modifikovani upotrebom veštačke inteligencije uneti dodatnu kompleksnost i izazove budućim istraživanjima. Preuzeti naučni radovi preuzeti su u *pdf* formatu koji je konvertovan u *word* format u kojem smo uklonili reference, repetitivne delove sa informacijama o časopisu i sve tabele. Pročišćeni tekstovi konvertovani su u *txt* format, koji je ujedno i finalni. Po konvencijama formiranja korpusa ovog tipa, svaki rad je svoj zasebni *txt* dokument.

Pošto je od velikog značaja za istraživanje da tri korpusa budu kompatibilna, vodilo se računa da sva tri korpusa budu iste dužine (okvirno pola miliona reči) i da pokrivaju isto semantičko polje i

žanr. Pokušali smo da odaberemo radove koji zadovoljavaju sve kriterijume a koji su takođe sličnih dužina. Međutim, kao i u prethodnim istraživanjima, pokazala se tendencija izvornih govornika da pišu duže naučne radove od neizvornih govornika, pa je veći broj naučnih radova bio neophodan za formiranje korpusa iste veličine. Ova tendencija je posebno bila izražena kod naučnih radova pisanih od strane neizvornih govornika kojima je maternji jezik rumunski, gde je zbog razlika u dužini naučnih radova, za stvaranje korpusa od pola miliona reči bilo neophodno 140 radova uporedno sa 80 radova za korpus izvornih govornika. Osnovni podaci o sva tri korpusa kada je reč o tačnoj dužini svakog, broju radova koji ih sačinjavaju, prosečnoj dužini radova i standardnoj devijaciji dužine radova u korpusu mogu se videti u Tabeli 24 dole. Dužina rada se sagledava u broju reči koji se nalazi u radu, u skladu sa normama utvrđenim u prethodnim istraživanjima iz korpusne lingvistike (npr. Pan et al., 2016; Hyland, 2008).

Tabela 24. *Osnovni podaci o korpusima*

	ECEP	SCEP	RCEP
Broj reči u korpusu	500.152	503.435	501.826
Broj naučnih radova	80	95	140
Prosečna dužina rada	6305	5352	3657
Standardna devijacija u dužini radova	1524	2138	1724

Kao što se može videti iz rezultata, najveća srednja vrednost dužine naučnog rada je u korpusu izvornih govornika. Manja je razlika u srednjoj dužini naučnih radova u korpusu neizvornih govornika kojima je maternji jezik srpski nego neizvornih govornika kojima je maternji jezik rumunski. Standardna devijacija u dužini radova najmanja je u korpusu izvornih govornika, dok je najveća u korpusu neizvornih govornika srpskog jezika, što ukazuje na veću koherentnost u dužini radova kod izvornih govornika engleskog jezika.

Svi radovi u korpusima su iz oblasti obrazovanja, međutim, polje obrazovanja je široko pa su uzeti u obzir i parametri nivoa obrazovanja koji se obrađuje u naučnom radu i podoblasti obrazovanja pri formiranju korpusa. Radovi su podeljeni u pet grupa na osnovu nivoa obrazovanja koji je sagledan u radovima:

1. Predškolsko obrazovanje;
2. Obrazovanje u osnovnim školama;
3. Obrazovanje u srednjim školama;
4. Visoko obrazovanje;
5. Opšti radovi nisu vezani za određen nivo obrazovanja.

Svaki od navedenih nivoa je jednako predstavljen u svakom od tri korpusa. Vodio se računa da se jednak broj radova nađe u svakoj od podgrupa. Raspodela radova po nivou obrazovanja koji je obrađen u radovima u sva tri korpusa može se videti u Tabeli 25.

Tabela 25. *Raspodela radova u korpusu na osnovu obrađenog nivoa obrazovanja*

	ECEP	SCEP	RCEP
Predškolsko obrazovanje	16	19	28
Osnovno obrazovanje	16	19	28
Srednjoškolsko obrazovanje	16	19	28
Visoko obrazovanje	16	19	28
Opšti radovi nisu vezani za određeni nivo obrazovanja	16	19	28

Pri formiranju korpusa vodilo se računa i da se podoblasti obrazovanja koje se obrađuju u naučnim radovima poklapaju u sva tri korpusa. Posebna pažnja posvećena je radovima sa temama iz opšte pedagogije, međutim, obrađuju se i druge značajne oblasti u obrazovanju. Obrazovne oblasti koje su zastupljene u radovima u sva tri korpusa su:

1. Opšta pedagogija (138 radova);
2. Pedagoška psihologija (45 radova);
3. Fizičko obrazovanje (14 radova);
4. Muzičko obrazovanje (20 radova);
5. Jezičko obrazovanje (31 rad);
6. Inkluzivno obrazovanje (18 radova);
7. Obrazovanje u oblasti prirodnih nauka (37 radova);
8. Obrazovna sociologija (9 radova).

Kao korpus sa najograničenijim izvorom podataka za građenje, prvo je sačinjen korpus obrazovnih naučnih radova neizvornih govornika kojima je maternji jezik srpski (SCEP). Slična raspodela po obrazovnim podoblastima je kasnije bila urađena i za druga dva korpusa. Raspodela radova po podoblastima u svakom od korpusa, zajedno sa procentima koje svaka od oblasti zastupa u korpusu mogu se videti u Tabeli 26. Iako raspodela po podoblastima u tri korpusa nije identična, autorka smatra da je dovoljno slična da bi korpusi bili međusobno uporedivi na nivou oblasti i podoblasti obrađenih u korpusu.

Tabela 26. *Raspodela naučnih radova po podoblastima u sva tri korpusa*

	ECEP		SCEP		RCEP	
	Br. radova	%	Br. radova	%	Br. radova	%
Opšta pedagogija	36	45	39	41.05	63	45
Pedagoška psihologija	13	16.25	15	15.79	17	12.14
Fizičko obrazovanje	3	3.75	5	5.26	6	4.29
Muzičko obrazovanje	5	6.25	5	5.26	10	7.14
Jezičko obrazovanje	8	10	11	11.58	12	8.57
Inkluzivno obrazovanje	4	5	6	6.32	8	5.71
Obrazovanje u oblasti prirodnih nauka	8	10	12	12.63	17	12.14
Obrazovna sociologija	2	2.5	2	2.11	5	3.57

Uzimajući u obzir kompleksnu geografsku situaciju raspodele izvornih govornika engleskog jezika, pri formiranju korpusa izvornih govornika još jedan parametar je uzet u obzir. Naime, vodilo se računa da se u korpusu predstave karakteristike govornika engleskog jezika iz različitih država u kojima je engleski jezik maternji. Tako u ECEP korpusu obrazovnih naučnih radova izvornih govornika postoji jednak broj radova iz pet anglofonih država: SAD, Kanada, Velika Britanija, Australija i Novi Zeland. Podaci o raspodeli radova u korpusu izvornih govornika na osnovu države u kojoj izvorni govornici rade je predstavljena u Tabeli 27.

Tabela 27. *Raspodela radova u ECEP na osnovu anglofonih država*

Država	Broj naučnih radova u ECEP
SAD	16
Kanada	16
Velika Britanija	16
Australija	16
Novi Zeland	16

Pri formiranju korpusa vodilo se računa da se ne uključuje više od jednog rada sa istim prvim autorom. Zbog ograničenog broja stručnjaka iz oblasti kojima je izvorni jezik srpski ili rumunski, a koji objavljuju na engleskom jeziku u datoj oblasti, omogućeno je da isti autor može u jednom radu da bude glavni autor i da u drugom radu bude koautor. Drugim rečima, jedan autor se maksimalno može naći dva puta u korpusu – jednom kao prvi autor rada i jednom kao koautor u drugom radu. Međutim, pri formiranju korpusa za naše istraživanje uključivanje istog autora dva puta na gorenaveden način se izbegavalo.

Na osnovu navedenih parametara, smatra se da su tri korpusa napravljena za potrebe datog istraživanja: 1) pogodna za međusobno poređenje; 2) oformljena u skladu sa parametrima usklađenim sa prethodnim istraživanjima; 3) pogodna za ispitivanje upotrebe ULS kod izvornih i neizvornih govornika engleskog jezika.

3.4. Metodološko-analitički postupak i metode

Istraživanje u ovom radu je eksplorativno istraživanje kvantitativnog pristupa koje prati korpusnolingvističku metodologiju. Fokusira se na upotrebu ULS od četiri reči kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski. Odličili smo se za ispitivanje ULS od četiri reči iz sledećih razloga:

- ULS od četiri reči su pet puta više zastupljeni od ULS od pet reči, dok često u sebi sadrže ULS od tri reči (Hyland, 2011);
- pogodniji su za strukturalnu i funkcionalnu analizu od ULS od tri reči zbog veće strukturalne i funkcionalne raznolikosti (Hyland, 2011);
- većina prethodnih istraživanja ispituju ULS od četiri reči, što nam daje bogatiji kontekst za analizu podataka (Pan i sar, 2016; Güngör, 2016; Lazić, 2017; Al Fedžri et al., 2020; Kostromitina, 2022).

Koraci u metodološko-analitičkom postupku u ovom istraživanju su sledeći:

1. *Identifikacija ULS u ECEP, SCEP i RCEP korpusima.* – ULS se izdvajaju na osnovu postavljene minimalne frekvencije i minimalne disperzije u korpusu. Odabir minimalne frekvencije i disperzije u suštini su arbitrarni, međutim, prethodna istraživanja se kreću u određenim okvirima. Frazeološka istraživanja koja ispituju upotrebu ULS kod izvornih i neizvornih govornika engleskog jezika uglavnom kao stroži kriterijum za izdvajanje ULS postavljaju minimalnu frekvenciju od 40 ponavljanja u milion reči (Pan i sar., 2016; Esfandiari i Barbari, 2017; Al Fedžri i sar., 2020) dok kao blaži kriterijum postavljaju minimalnu frekvenciju od 20 ponavljanja u milion reči (Perez-Ljantada, 2014; Güngör & Uysal, 2016; Güngör, 2016; Lazić, 2017). Zbog činjenice da manji korpusi, kao što su korpusi u ovom istraživanju, imaju tendenciju da generišu disproporcionalno velike brojeve ULS, u ovom istraživanju primenili smo stroži kriterijum za identifikaciju ULS od 40 ponavljanja u milion reči. Minimalna disperzija je takođe suštinski arbitrarna. Ona se koristi kako bi se osiguralo da izdvojeni ULS nisu karakteristika određenog autora. Prethodna relevantna istraživanja uglavnom koriste ili stroži kriterijum minimalne disperzije u 10% tekstova (Perez-Ljantada, 2014; Güngör & Uysal, 2016; Güngör, 2016; Al Fedžri et al., 2020) ili blaži kriterijum minimalne disperzije u pet tekstova (Pan i sar., 2016; Salazarova, 2014; Esfandiari & Barbari, 2017). U istraživanju u ovoj disertaciji odlučili smo se za primenu strožeg kriterijuma minimalne disperzije u 10% tekstova. Program korišćen za analizu korpusa je *AntConc* (Anthony, 2013).

2. *Strukturalna analiza izdvojenih ULS (svih, zajedničkih i specifičnih)*. – U istraživanju primenjujemo strukturalnu taksonomiju koju su predložili Bajber i sar. (1999). Strukturalna taksonomija je blago modifikovana kako bi što preciznije i efikasnije predstavila upotrebu ULS u istraženom žanru i oblasti u ovoj disertaciji (naučni radovi u oblasti obrazovanja). Blage modifikacije prihvaćenih i rasprostranjenih taksonomija radi preciznog predstavljanja podataka je prihvaćena praksa u istraživanjima ULS (Kortez, 2004; Salazar, 2010; Güngör, 2016; Lazić, 2017). U poređeno sa originalnom strukturalnom taksonomijom ULS u akademskim tekstovima koju su predložili Bajber i saradnici (1999), mi smo uklonili strukturalnu podgrupu *fragment adverbijalne klauze* (engl. *adverbial clause fragments*), zato što nijedan ULS u toj strukturalnoj podgrupi nije identifikovan ni u jednom od korpusa u ovom istraživanju. Takođe je u grupi *glagolske strukture* dodata podgrupa *druge glagolske fraze* po uzoru na istraživanje Gungor (2016) zbog identifikacije ULS u grupi *glagolske strukture* koji se nisu uklapali ni u jednu od podgrupa unutar originalne strukturalne taksonomija koju su predložili Bajbar i sar. (1999). U Tabeli 28 predstavljena je strukturalna taksonomija koja je korišćena u istraživanju u ovoj disertaciji³⁷.
3. *Funkcionalna analiza izdvojenih ULS (svih, zajedničkih i specifičnih)*. – Za funkcionalnu raspodelu koristi se taksonomija koju je predložio Hajland (2008). Funkcionalna taksonomija je blago modifikovana uklaňanjem podgrupe *ULS teme* (engl. *topic*), koji se odnose na konkretnu oblast istraživanja (npr. *the currency board system*). Navedena funkcionalna podgrupa je uklonjena zato što se naveden tip ULS uklanja u procesu prečišćavanja spiskova ULS po uzoru na prethodna istraživanja Salazar (2010) i Lazić (2017). Funkcionalna taksonomija koja je korišćena u istraživanju u ovoj disertaciji predstavljena je u Tabeli 29³⁸.

Tabela 28. *Strukturalna taksonomija korišćena u istraživanju u ovoj disertaciji*

Grupa	Podgrupa	Primeri
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	<i>the implementation of the, a part of the</i>
	<i>Druge imeničke fraze</i>	<i>the fact that the, the extent to which</i>
<i>Fragmenti sa predložkom frazom</i>	<i>Predložka fraza sa of fragmentom</i>	<i>in the process of, in the case of</i>
	<i>Druge predložke fraze</i>	<i>at the same time, in relation to the</i>
<i>Glagolske strukture</i>	<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	<i>it is necessary to, it is important to</i>
	<i>Glagol u pasivu i fragment predložke fraze</i>	<i>is based on the, are shown in table</i>
	<i>Kopula be i imenička/pridevska fraza</i>	<i>is one of the</i>
	<i>Fragment that klauze (Glagol/pridev +) fragment to klauze</i>	<i>that there is a, that they do not</i>
	<i>Zamenica/imenička fraza+be (+...)</i>	<i>are more likely to</i>
<i>Druge strukture</i>	<i>Druge glagolske fraze</i>	<i>there is no ststistically</i>
	<i>Druge strukture</i>	<i>based on the results, having in mind the as well as the, we can see that</i>

³⁷Primeri u tabeli su izdvojeni iz rezultata istraživanja u ovoj disertaciji.

³⁸Primeri u tabeli su izdvojeni iz rezultata istraživanja u ovoj disertaciji

Tabela 29. Funkcionalna taksonomija koja je korišćena u istraživanju u ovoj disertaciji

Grupa	Podgrupa	Primeri
ULS orijentisani ka istraživanju	ULS za lokaciju	<i>in the area of, the end of the</i>
	ULS za proceduru	<i>in the process of, the application of the</i>
	ULS za kvantifikaciju	<i>is one of the, one of the most</i>
	ULS za opis	<i>the fact that the, the importance of the</i>
ULS orijentisani ka tekstu	Tranzicioni signali	<i>having in mind that, as well as the</i>
	Rezultativni signali	<i>that there is a, the results of this</i>
	Strukturni signali	<i>in the current study</i>
	Ograničavajući signali	<i>in the field of, in the context of</i>
ULS orijentisani ka učesniku	ULS stava	<i>it is necessary to, it is possible to</i>
	ULS uključivanja	<i>we can conclude that, it can be seen</i>

4. *Analiza upotrebe najzastupljenijih obrazaca ULS: the + imenica + of the/a, in the + imenica + of, glagol u pasivu + fragment predloške fraze i uvodna it struktura + glagolska/ pridevska fraza.*
5. *Komparacija podataka o zastupljenosti ULS, strukturalnoj raspodeli, funkcionalnoj raspodeli i upotrebi najzastupljenijih obrazaca u sva tri korpusa.*
6. *Statistička analiza podataka o strukturalnoj i funkcionalnoj raspodeli na grupe i podgrupe svih i specifičnih ULS sprovodi se na sledećim relacijama: između izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski, između izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski, između neizvornih govornika engleskog jezika kojima je maternji jezik srpski i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski. Testovi se sprovode na navedenim relacijama, umesto na sva tri korpusa zajedno, kako bi ustanovili na koje se načine svaka od ispitanih grupa neizvornih govornika razlikuje u upotrebi ULS uporedno sa izvornim govornicima engleskog jezika, a za koje se smatra da postavljaju norme za adekvatnu upotrebu ULS. Odvojenim poređenjem produkcije dve grupe neizvornih govornika možemo da uvidimo postojanje zajedničkih karakteristika i razlika u upotrebi ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski. Kako bi dobili potpunije razumevanje karakteristike upotrebe i razlike u upotrebi ULS kod tri grupe govornika, testovi statističke analize su primenjeni i za tipove ULS i za tokene. Analizom samo tipova ili samo tokena mogli bismo da dobijemo nepotpune podatke koji bi potencijalno mogli da nas navedu na pogrešne zaključke. Korpusi, na primer, mogu da se poklope po strukturalnoj raspodeli tipova, ali da se tokeni iz određene strukturalne grupe značajno više koriste od drugih te da se stanje u tekstovima značajno razlikuje od netačnog zaključka strukturalnog poklapanja. Kako bi se ispitale potencijalne razlike, tj. asocijacije, između korpusa (ECEP, SCEP i RCEP) u odnosu na strukturalne i funkcionalne distribucije ULS, korišćen je hi-kvadrat test nezavisnosti (engl. *chi-square test*). To je statistički test koji se zasniva na utvrđivanju razlika između opaženih i očekivanih frekvencija. Za sve hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96. Hi-kvadrat testovi su sprovedeni za specifične i sve ULS za funkcionalne grupe, podgrupe, strukturalne grupe i podgrupe, i svaki hi-kvadrat test sproveden je na ukupnim frekvencijama ULS (tokeni), kao i na broju specifičnih ULS (tipovi) koje se pojavljuju u korpusima.*

7. Analiza uticaja negativnog transfera na produkciju ULS kod dve grupe neizvornih govornika.

U ovom istraživanju uticaj transfera se ispituje kako bi se razjasnile razlike u upotrebi ULS kod neizvornih govornika rumunskog i srpskog jezika. Sagledavaju se dve grupe neizvornih govornika engleskog da bi se ispitalo koje su karakteristike upotrebe *univerzalne* za neizvorne govornike engleskog jezika, a koje su *specifične* za određene grupe neizvornih govornika. Naravno, jedno istraživanje ne može da pruži potpune odgovore na gorenavedena pitanja, ali može da nas dovede bliže odgovoru i rasvetli karakteristike upotrebe ULS i razlike u upotrebi ULS kod ove dve grupe neizvornih govornika engleskog jezika i to u naučnim radovima iz oblasti obrazovanja. Ispitivanje uticaja transfera omogućava da preciznije identifikujemo koji se ULS koriste zbog uticaja transfera i koji je to procenat ukupne produkcije ULS kod neizvornih govornika engleskog kojima je maternji jezik srpski ili rumunski. Te informacije će dalje olakšati identifikaciju karakteristika koje su zajedničke za obe grupe neizvornih govornika i potencijalno predstavljaju karakteristike neizvornih govornika uopšte. Identifikacija uticaja transfera, takođe, daje deo odgovora na pitanje zašto neizvorni govornici mahom koriste veći broj ULS u pisanju naučnih radova uporedno sa izvornim govornicima (Pan i sar., 2016; Al Fajri i sar., 2020) i pruža istraživačima mogućnost da dalje istražuju ovaj fenomen kako bi se odgovor upotpunio.

Za identifikaciju transfera u ovoj disertaciji koristi se Džarvisov pristup zasnovan na detekciji (Džarvis, 2000, 2012; Džarvis i Pavlenko, 2008). Navedeni pristup je odabran zato što predstavlja metodološki jasan okvir za identifikaciju transfera koji se takođe koristi u prethodnim istraživanjima u ovoj oblasti (Paket, 2013; Güngör & Uysal, 2020; Kostromitina, 2022).

Džarvisov pristup zasnovan na detekciji (Džarvis, 2000, 2012; Džarvis i Pavlenko, 2008) podrazumeva prikupljanje tri vrste dokaza: *unutargrupna homogenost*, *međugrupna heterogenost* i *međujezička podudarnost performansi*. Dosadašnja istraživanja koja su ispitivala uticaj transfera na upotrebu ULS glavni su akcenat stavljala na treću vrstu dokaza (*međujezička podudarnost performansi*), dok su za drugu grupu dokaza poredili izdvojene ULS sa onim iz drugih istraživanja (Güngör & Uysal, 2020) ili pak uopšte nisu obezbedili ovu vrstu dokaza (Kostromitina, 2023). U ovoj disertaciji akcenat je stavljen baš na drugu vrstu dokaza kako bi se prethodno naveden nedostatak u dosadašnjim istraživanjima nadomestio. Načini prikupljanja dokaza u ovom istraživanju za sve tri grupe dokaza u Džarvisovom pristupu zasnovanom na detekciji (Džarvis, 2000, 2012; Džarvis i Pavlenko, 2008) prikazani su u Tabeli 30.

Tabela 30. Metode prikupljanja dokaza za postojanje transfera u upotrebi ULS

<i>Unutargrupna homogenost</i>	<i>Međugrupna heterogenost</i>	<i>Međujezička podudarnost performansi</i>
Dokazi za unutargrupnu homogenost prikupljaju se primenom minimalne frekvencije za izdvajanje ULS od 40 ponavljanja u milion reči i minimalne disperzije u 10% tekstova u korpusu.	Dokazi za međugrupnu heterogenost dobijeni su poređenjem ULS izdvojenih iz sva tri korpusa korišćena u ovoj disertaciji i izdvajanjem onih koji su specifični za svaku grupu govornika. U ovom koraku je značajna činjenica da su korpusi ujednačeni po veličini, temi i načinu formiranja korpusa, te je njihovo poređenje u skladu sa visokim standardima rigoroznosti u istraživanju (Begsten, 2019).	Dokazi za međujezičku podudarnost performansi prikupiće se identifikacijom prevodnih ekvivalenata ULS koji su specifični za govornike srpskog i rumunskog jezika, nakon čega će se proveriti u korpusima naučnih radova na srpskom i rumunskom jeziku da li se i u njima pojavljuju i koriste u istim funkcijama.

Za poređenje izdvojenih prevodnih ekvivalenata specifičnih učestalih spojeva za izvorne govornike srpskog jezika ili za izvorne govornike rumunskog jezika koristila su se dva korpusa koja su identifikovana kao najpogodnija za ovu svrhu, od javno dostupnih korpusa. Kao referentni korpus za produkciju na srpskom jeziku koristili smo *Korpus savremenog srpskog jezika (SrpKor2013)* odnosno potkorpus naučnih i naučno popularnih tekstova. Korpus savremenog srpskog jezika je dinamični korpus, što znači da se njegova veličina konstantno povećava dodatkom novih tekstova. U periodu kada je vršena analiza podataka za ovu disertaciju datuma 30. 1.2024. godine, potkorpus naučnih i naučnopopularnih tekstova je bio veličine 3.657.797 reči i sačinjavao je 3% celog *SrpKor2013* korpusa. Sa trenutnim postavkama dostupnosti korpusu, nije moguće pristupiti samo tekstovima koji su u žanru naučnih radova već je za analizu dostupno sagledati sve tekstove koji su obuhvaćeni širom klasifikacijom naučnih i naučno popularnih tekstova. Korpusni tekstovi su anotirani bibliografski i morfološki. Obe navedene karakteristike su korišćene za tačnije izdvajanje informacija za istraživanje u ovoj disertaciji. Naime, zbog postojanja bibliografske anotiranosti, imali smo priliku da izdvojimo samo naučne i naučno- popularne tekstove u jedan potkorpus, dok je lematizacija korpusa korišćena za efikasnije pretraživanje zastupljenosti prevodnih ekvivalenata. Detaljnije informacije o *Korpusu savremenog srpskog jezika (SrpKor2013)* dostupne su na sledećoj veb-stranici: <http://korpus.matf.bg.ac.rs/prezentacija/korpusi.html> ili u okviru naučnih radova Popović (2010) i Utvić (2011).

Kao referentni korpus za produkciju na rumunskom jeziku korišćen je EXPRES korpus (*Corpus of Expert Writing in Romanian and English*), odnosno potkorpus naučnih radova na rumunskom jeziku. EXPRES korpus je takođe dinamični i bilingvalni korpus, sastoji se od naučnih radova iz četiri naučne discipline: lingvistike, ekonomije, političkih nauka i informatike. Podeljen je u tri potkorpusa:

1. Naučni radovi na rumunskom jeziku;
2. Naučni radovi na engleskom jeziku koje su pisali izvorni govornici engleskog jezika;
3. Naučni radovi na engleskom jeziku koje su pisali izvorni govornici rumunskog jezika.

U periodu kada je vršena analiza podataka za ovu disertaciju datuma (10. 2. 2024. godine) potkorpus naučnih radova na rumunskom jeziku bio je veličine 4.201.363 reči i sastojao se od po 200 radova iz svake od četiri navedenih disciplina. Korpus je bibliografski anotiran ali nije anotiran morfološki. Detaljnije informacije o EXPRES korpusu dostupne su na sledećoj veb-stranici: <https://expres-corpus.org/documentation.php?SID> ili u istraživanju Kitez i sar. (2022).

S obzirom na činjenicu da u navedenim korpusima koji su odabrani da u ovom istraživanju služe kao referentni korpusi nije moguće izdvajati ULS i na osnovu njih izdvojiti prevodne ekvivalente, u istraživanju u ovoj disertaciji prvo su urađeni prevodi specifičnih ULS za izvorne govornike srpskog jezika i specifičnih ULS za izvorne govornike rumunskog jezika, nakon čega se zastupljenost prevoda ispitala u referentnim korpusima. Pošto je ekvivalentnost kompleksan pojam koji prožima translatološku teoriju od samog početka do sadašnjeg trenutka (Munday et al., 2022; Pym, 2023), važno je istaći šta podrazumevamo pod ovim terminom u istraživanju u ovoj disertaciji i na koji se teorijski okvir oslanjamo.

U ovom istraživanju oslanjamo se na *Skopos teoriju* (Reiss, Nord & Vermeer, 2014) sa ciljem da predložimo *funkcionalne ekvivalente* odnosno da postignemo *funkcionalnu adekvatnost* pri prevođenju specifičnih ULS za izvorne govornike srpskog ili rumunskog jezika. U Skopos teoriji ekvivalentnost je „veza između ciljnog i izvornog teksta koja (može da) postiže istu komunikativnu funkciju na istom nivou u dve kulture“; dalje pojašnjavaju da je u okviru te teorije (Reiss et al., 2014: 128) „ekvivalentnost posebna vrsta adekvatnosti, drugim rečima, adekvatnost pod uslovom da skopos zahteva postizanje iste funkcije u izvornom i ciljnom tekstu“ (ibidem). Funkcionalna

ekvivalentnost koju predlaže Skopos teorija je odabrana kao najadekvatnija za istraživanje u ovoj disertaciji zbog prirode lingvističkih elemenata koji su u fokusu istraživanja, odnosno zbog prirode ULS. Naime, ULS mogu biti gramatički potpuni ili nepotpuni frazalni (npr. *the result of the*) ili klauzalni (npr. *is used as the*) segmenti koji ispunjavaju određenu diskursnu funkciju (Paket, 2013). Zbog činjenice da sadrže veliki broj gramatičkih reči i tipoloških razlika koje postoje između engleskog jezika i srpskog jezika, kao i engleskog jezika i rumunskog jezika, formalna ekvivalentnost bi bila nemoguća u većini primera.

U poslednjem koraku u prikupljanju dokaza za međujezičku podudarnost podataka proverava se zastupljenost predloženih funkcionalnih ekvivalenata specifičnih ULS u svakoj od dve grupe neizvornih govornika. Statistička značajnost razlika u zastupljenosti u SCEP i RCEP korpusima uporedno sa referentnim korpusima utvrđuje se primenom testa verodostojnosti (engl. *log-likelihood* test) po uzoru na prethodna istraživanja (Güngör & Uysal, 2020; Kostromitina, 2022). Kao i hi-kvadrat test, test verodostojnosti se zasniva na poređenju zabeleženih i očekivanih frekvencija, ali uključuje samo jednu varijablu koja je u ovom istraživanju veličina korpusa.

3.5. Ograničenja istraživanja

Ovim istraživanjem može se dobiti uvid u karakteristike upotrebe ULS i razlikama u upotrebi ULS u naučnim radovima iz oblasti obrazovanja kod izvornih govornika engleskog jezika, neizvornih govornika engleskog jezika kojima je maternji jezik srpski i kod neizvornih govornika kojima je maternji jezik rumunski. Za potrebe istraživanja formirana su tri korpusa koji su ujednačeni po veličini, žanru i semantičkom polju, odnosno naučnoj oblasti u kojoj se pišu radovi. Dva nedostatka se mogu uvideti u procesu formiranja korpusa:

1. Korpusi su ujednačeni po veličini ali nisu ujednačeni po broju radova koji ih formiraju. Do navedenog nedostatka je došlo zbog tendencija istraživača iz različitih kultura i govornih područja da proizvode tekstove različitih dužina. S obzirom na rezultate istraživanja Pan i sar. (2020), koji su pokazali da i razlike u broju radova utiču na broj i tip izdvojenih ULS, precizniji podaci bi se dobili kada bi korpusi bili izjednačeni i po broju radova.
2. Nemamo informacije o tome da li su naučni radovi pisani na engleskom jeziku ili predstavljaju prevode. Zbog činjenice da prevedeni tekstovi pokazuju tendenciju da sadrže veći broj ULS, informacije o tome da li su u pitanju prevedeni tekstovi ili ne bi nam omogućila da ujednačimo korpusu i po ovom parametru ili da u potpunosti uklonimo prevedene naučne radove. Međutim, zbog činjenice da sadašnja izdavačka praksa pri objavljivanju naučnih radova nalaže da se ne pružaju informacije o tome da li je naučni rad preveden i ko je prevodilac, nije bilo moguće pribaviti relevantne informacije. Situacija je dodatno otežana činjenicom da za ispitivanje ULS korpusi moraju da budu relativno veliki i da nije preporučljivo raditi na korpusu koji je manji od pola miliona reči (Begsten, 2019; Pan i sar., 2020).

S obzirom na činjenicu da ispitivanje transfera nije glavni cilj ovog istraživanja, referentni korpusi koji su se koristili za prikupljanje dokaza o međujezičkoj podudarnosti performansi nisu formirani posebno za potrebe ovog istraživanja kao ECEP, SCEP i RCEP korpusi, već su korišćeni javno dostupni korpusi, što se takođe može smatrati nedostatkom. Formiranjem referentnih korpusa, primenom istih standarda koji su korišćeni za ECEP, SCEP i RCEP korpusu, omogućilo bi izdvajanje ULS u referentnim korpusima i preciznije prikupljanje podataka za utvrđivanje uticaja transfera, iako su prikupljeni dokazi za sve tri grupe dokaza koje su neophodne za utvrđivanje uticaja transfera u slučaju ovog istraživanja negativnog transfera.

4. Rezultati

4.1. Izdvajanje učestalih leksičkih spojeva

Analiza korpusa urađena je primenom programa *AntConc* (Antoni, 2013). Pri izdvajanju ULS iz korpusa koristila se minimalna frekvencija od 40 ponavljanja u milion reči po uzoru na prethodna istraživanja (Pan i sar., 2016; Esfandiari i Barbari, 2017; Al Fedžri, 2020), kao i minimalna disperzija u 10% naučnih radova po uzoru na prethodna istraživanja (Hyland, 2008). Uzimajući u obzir činjenicu da manji korpusi imaju tendenciju da generišu disproporcionalno velike količine ULS (Begsten, 2019), primenjeni su strogi kriterijumi za identifikaciju ULS kako bi se to izbeglo. Koristeći gorenavedene parametre, u korpusu ECEP izdvojeno je 29 ULS, u korpusu SCEP izdvojeno je 129 ULS, a u korpusu RCEP izdvojen je 71 ULS (v. Tabelu 31).

Tabela 31. *Broj tipova i tokena izdvojenih ULS u korpusima istraživanja sa nepročišćenog spiska*

Korpus	Broj izdvojenih ULS
ECEP	29 (828)
SCEP	129 (5204)
RCEP	71 (2727)

4.2. Prečišćavanje spiskova učestalih leksičkih spojeva

Za prečišćavanje spiskova ULS korišćeni su kriterijumi koji su primenjeni u istraživanjima Lazić (2017) i Al Fedžri i sar. (2020), po kojima je neophodno ukloniti sve ULS koji su u direktnoj vezi sa određenim sadržajem ili kontekstom, ULS koji sadrže vlastite imenice, kao i podrediti ULS koji sačinjavaju jedan duži ULS radi najkonciznijeg predstavljanja podataka (str. 79). Pri primenjivanju gorenavedene procedure za prečišćavanje spiskova, autorka disertacije je uočila primere u kojima, iako dva izdvojena ULS od četiri reči formiraju jedan duži od pet – taj duži spoj se nekada retko koristi u korpusu, dok se dva ULS koji ga sačinjavaju češće koriste u drugačijim kombinacijama reči. Na primer, u korpusu RCEP identifikovani su ULS *the number of students* i *in the number of*, koji mogu da sačine zajedničku ULS od pet reči *in the number of students*. Detaljnije sagledavanje njihove upotrebe pokazalo je da se duži ULS pojavljuje samo tri puta u čitavom korpusu. Manje frekventan ULS od dva je *in the number of* i ima frekvenciju od 20 ponavljanja. To nam dalje ukazuje da se ULS *in the number of* koristi kao duži ULS od pet reči samo u 15% njegove upotrebe u korpusu (ULS *the number of students*, koji ima frekvenciju 24, prisutan je u korpusu kao navedeni duži ULS u samo 12% svoje upotrebe). Samim tim, spajanje ova dva ULS u jedan duži umanjilo bi tačnost predstavljenih podataka.

Kako bi se izbegla situacija spajanja ULS koji se u korpusu retko koriste spojeni, i na taj način onemogućila distorzija podataka, neophodno je zadovoljiti barem jedan od sledeća dva uslova za spajanje ULS:

1. ULS od pet reči koji je sačinjen od dva ULS od četiri reči mora zadovoljiti osnovne kriterijume za izdvajanje, odnosno da ima minimalnu frekvenciju od 40 ponavljanja u milion reči i minimalnu disperziju u 10% naučnih radova;
2. 80% i više tokena manje frekventnog ULS su deo ULS od pet reči koji sačinjavaju dva ULS čije se spajanje razmatra.

Navedena dva kriterijuma primenjuju se sa ciljem izbegavanja neprecizne prezentacije podataka u kojima se predstavlja spoj dve celine koji se koristi značajno manje nego dve navedene celine zasebno u drugačijim spojevima. Dva gorenavedena kriterijuma osmislila je autorka disertacije.

Kao primer za implementaciju prvog kriterijuma, uzećemo dva ULS iz korpusa RCEP: *the beginning of the* i *at the beginning of*. Ta dva ULS od četiri reči mogu se spojiti u ULS od pet reči: *at the beginning of the*. Da bi se proverilo da li navedeni duži ULS zadovoljava osnovne kriterijume za identifikaciju, urađena je pretraga korpusa koristeći program *AntConc*. Unete su specifikacije minimalne frekvencije od 40 ponavljanja u milion reči i minimalna disperzija u 10% tekstova ali za ULS dužine pet reči. Između ostalih, izdvojen je i ULS od pet reči *at the beginning of the* i tako je potvrđeno da zadovoljava originalne kriterijume te je validno spojiti dva kraća ULS u duži ULS od pet reči.

Kao primer za implementaciju drugog kriterijuma, mogu se uzeti dva ULS iz SCEP *significant differences in the* i *statistically significant differences in*. Dva navedena ULS zajedno sačinjavaju ULS od pet reči *statistically significant differences in the*; pošto je drugi od dva ULS manje frekventan, reč *statistically* našla bi se u zagradi kao opcioni element. Pri analizi svih tokena, može se uočiti da se ULS *statistically significant differences in*, u okviru gorenavedenog dužeg tokena, u korpusu pojavljuje u 14 slučajeva, odnosno tokena, od ukupno 26 tokena koji se pojavljuju u SCEP korpusu. Stoga podređivanjem ULS *statistically significant differences in* na naveden način bi zanemarilo 12 tokena navedenog ULS odnosno 45% njegove zastupljenosti u korpusu. Primenom navedenih koraka za prečišćavanje spiskova ULS dobijmo rezultate predstavljene u Tabeli 32.

Tabela 32. Broj tipova i tokena izdvojenih ULS u korpusima istraživanja sa pročišćenog spiska

Korpus	Broj izdvojenih ULS
ECEP	27 (777)
SCEP	111 (4536)
RCEP	60 (2358)

Prečišćeni spisak ULS u korpusu kod izvornih govornika engleskog jezika (ECEP) sadrži 27 ULS i 777 tokena. Jedan ULS je uklonjen zato što je u bliskoj vezi sa sadržajem i predstavlja konkretan koncept (*the ministry of education*), a dva ULS *at the end of* i *the end of the* su spojeni u jedan ULS *at the end of (the)*.

Prečišćeni spisak ULS u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik srpski (SCEP) sadrži ukupno 111 ULS i 4.536 tokena. Ukupno pet ULS je uklonjeno sa spiska zato što sadrže vlastite imenice (*the republic of serbia, of the republic of, in the republic of, the faculty of education, the ministry of education*³⁹), dok je jedan ULS uklonjen jer nije sačinjen od četiri reči već je sačinjen od tri, a do izdvajanja je došlo zato što program nije prepoznao apostrof (*of one s own*). Ukupno 12 ULS je uklonjeno sa spiska zato što su podređeni drugom ULS. Od navedenih 12 ULS, deset je podređeno zato što zadovoljavaju prvi kriterijum odnosno zato što se spojeni ULS od pet reči pojavljuje sa minimalnom frekvencijom od 40 ponavljanja u milion reči i sa minimalnom disperzijom u 10% naučnih radova u korpusu. Od navedenih deset ULS, sedam zadovoljava i drugi kriterijum – da podređeni ULS u barem 80% svojih tokena bude deo ULS od pet reči kojem se pripaja (*when it comes to (the)* – 100% tokena, *the results of the (research)* – 86,63% tokena, *on the basis of (the)* – 83,33% tokena, *(it)can be concluded (that the)* – 98,46% i

³⁹ U program *AntConc*, koji je korišćen za izdvajanje ULS, rezultati upita o formiranju spiskova izdvojenih ULS se predstavljaju malim slovima, te su i ovde rezultati predstavljeni kao i u programu – malim slovima.

85,19% tokena, *on the other hand (the)* – 100% tokena, *one of the most (important)* – 80% tokena). Dva ULS su podređeni zato što zadovoljavaju samo drugi kriterijum za spajanje ULS i oni su *in the field of (education)* i *based on the results (of)*, a procenat njihovih tokena koji sačinjavaju duži ULS su 85,71% i 85%.

Prečišćeni spisak ULS u korpusu neizvornih govornika kojima je maternji jezik rumunski (RCEP) sadrži ukupno 60 ULS i 2.358 tokena. Ukupno tri ULS uklonjena su sa spiska zato što sadrže vlastite imenice odnosno predstavljaju koncept: *the ministry of education, the romanian education system, the romanian educational system*. Zbog podređenosti drugom ULS, osam ULS je uklonjeno sa spiska. Od navedenih osam ULS, dva zadovoljavaju prvi kriterijum: *at the beginning of (the)* i *at the end of (the)*. Jedan ULS zadovoljava samo drugi kriterijum: *one of the most (important)* – 81, 82%, dok pet ULS zadovoljavaju i prvi i drugi kriterijum: *in the case of (the)* – 89,66%, *in the context of (the)* – 85,71%, *the point of view (of)* – 100% i *(and) on the other hand (the)* koji se sastoji od tri ULS – 81,48% i 100%.

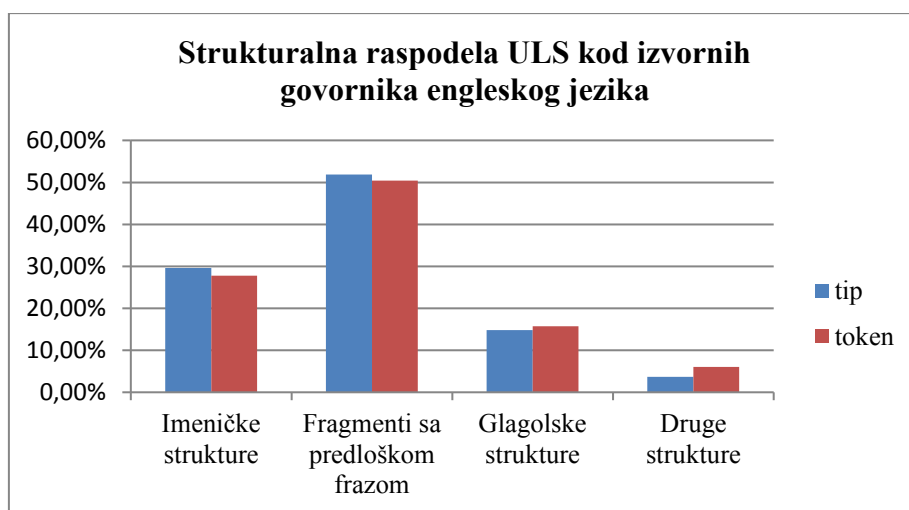
Na osnovu prečišćenih spiskova izdvojenih ULS u sva tri korpusa u ovom istraživanju možemo da zaključimo da je potvrđena četvrta radna hipoteza: *Veći broj tipova i tokena ULS biće identifikovan kod dve grupe neizvornih govornika engleskog jezika uporedno sa izvornim govornicima engleskog jezika u naučnim radovima iz oblasti obrazovanja*. Na osnovu istih podataka možemo da zaključimo i da nije potvrđena peta hipoteza: *Sličan broj tipova i tokena biće identifikovan kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski u naučnim radovima iz oblasti obrazovanja*. Peta hipoteza nije potvrđena, jer se u produkciji neizvornih govornika engleskog jezika kojima je maternji jezik srpski identifikovao veći broj tipova i tokena nego u produkciji neizvornih govornika engleskog jezika kojima je maternji jezik rumunski. Prečišćeni spiskovi ULS za svaki od korpusa dostupni su u *Prilogu 1, Prilogu 2 i Prilogu 3*.

4.3. Strukturalna raspodela učestalih leksičkih spojeva

Kako bi se bolje razumele karakteristike upotrebe ULS kod izvornih govornika engleskog jezika i kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski i rumunski, urađena je strukturalna raspodela izdvojenih ULS u sva tri korpusa. Kako bi se utvrdila značajnost uviđenih razlika u strukturalnoj raspodeli ULS, urađena je i statistička analiza podataka primenom hi-kvadrat testa. Pri strukturalnoj analizi ULS, oni su se prvo raspodelili u glavne strukturalne grupe, pa zatim i na strukturalne podgrupe. Radi precizne prezentacije i interpretacije podataka, strukturalna zastupljenost i tipova i tokena je predstavljena za strukturalne grupe i podgrupe za sva tri korpusa.

4.3.1. Strukturalna raspodela ULS kod izvornih govornika engleskog jezika

Strukturalna raspodela ULS izdvojenih u korpusu izvornih govornika engleskog jezika (ECEP) na grupe pokazuje da su najzastupljeniji *fragmenti sa predloškom frazom* praćeni *imeničkim strukturama*, dok su *glagolske strukture* i *druge strukture* najmanje zastupljene (v. Grafikon 25 i Tabelu 33). Navedeni podaci potvrđuju prvu radnu hipotezu: *Strukturalna raspodela ULS kod izvornih govornika engleskog jezika u naučnim radovima u oblasti obrazovanja biće okarakterisana većom zastupljenošću ULS iz strukturalne grupe fragmenti sa predloškom frazom uporedno sa drugim strukturalnim grupama*.



Grafikon 25. *Strukturalna raspodela ULS kod izvornih govornika engleskog jezika*

Sagledanjem zastupljenost tipova i tokena u strukturalnim kategorijama ULS u ECEP korpusu uviđamo da nema velikih razlika u zastupljenosti tipova i tokena. Možemo uvideti da je procentualna raspodela tipova blago veća uporedno sa tokenima kod *imeničkih struktura* i *fragmentata sa predloškom frazom*, dok je situacija obrnuta za *glagolske strukture* i *druge strukture*. Navedeni podaci ukazuju na to da izvorni govornici engleskog jezika koriste veću raznolikost ULS u *imeničkim strukturama* i *fragmentima sa predloškom frazom* ali sa manjom učestalošću, dok kod *glagolskih struktura* i *drugih struktura* koriste manju raznolikost ULS ali sa većom učestalošću upotrebe.

Tabela 33. *Strukturalna raspodela ULS kod izvornih govornika engleskog jezika*

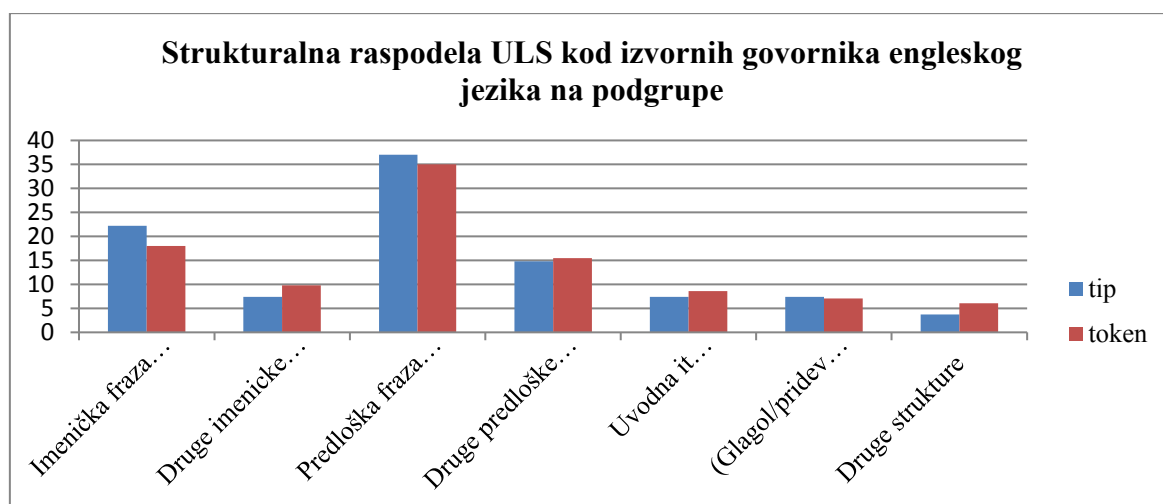
Strukturalna grupa	Tip (Token)	Procenti
<i>Imeničke strukture</i>	8 (216)	29.63 (27.80)
<i>Fragmenti sa predloškom frazom</i>	14 (392)	51.85 (50.45)
<i>Glagolske strukture</i>	4 (122)	14.81 (15.70)
<i>Druge strukture</i>	1 (47)	3.7 (6.05)

Strukturalne raspodela ULS kod izvornih govornika engleskog jezika na podgrupe pokazuje da je *predloška fraza sa of fragmentom* najzastupljenija strukturalna podgrupa i da sačinjava 37,04% svih ULS u ECEP korpusu (v. Tabelu 34 i Grafikon 26). Praćena je podgrupama *imenička fraza sa of fragmentom* i *druge predloške fraze*, koje sačinjavaju 22,22% i 14,81% svih ULS u ECEP korpusu. Samo jedna od sedam podgrupa iz grupe *glagolske strukture* je zastupljena u ECEP korpusu, što dalje potvrđuje tendenciju izvornih govornika engleskog jezika da u žanru akademskih tekstova retko koriste ULS iz strukturalne kategorije *glagolske strukture*.

Tabela 34. *Strukturalna raspodela ULS kod izvornih govornika engleskog jezika na podgrupe – tipovi i tokeni*

Grupa	Podgrupa	ECEP		
		Tip	(Token)	Procenti
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	6	(140)	22.22 (18.02)
	<i>Druge imeničke fraze</i>	2	(76)	7.41 (9.78)
<i>Fragmenti sa predloškom frazom</i>	<i>Predloška fraza sa of fragmentom</i>	10	(272)	37.04 (35.01)
	<i>Druge predloške fraze</i>	4	(120)	14.81 (15.44)
<i>Glagolske strukture</i>	<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	2	(67)	7.41 (8.62)
	<i>Glagol u pasivu i fragment predloške fraze</i>	/	/	/
	<i>Kopula be i imenička/pridevska fraza</i>	/	/	/
	<i>Fragment that klauze</i>	/	/	/
	<i>(Glagol/pridev +) fragment to klauze</i>	2	(55)	7.41 (7.08)
	<i>Zamenica/imenička fraza+be (+...)</i>	/	/	/
<i>Druge strukture</i>	<i>Druge glagolske fraze</i>	/	/	/
	<i>Druge strukture</i>	1	(47)	3.70 (6.05)

Što se tiče raspodele tipova i tokena u strukturalnim podgrupama ULS kod izvornih govornika engleskog jezika možemo videti da su uglavnom ravnomerno raspoređeni. U podgrupama *imenička fraza sa of fragmentom* i *predloška fraza sa of fragmentom* možemo videti da je zastupljenost tipova, procentualno gledano, blago veća uporedno sa tokenima, što implicira da se u ove dve podgrupe ređe upotrebljavaju tipovi ali da su više raznoliki.



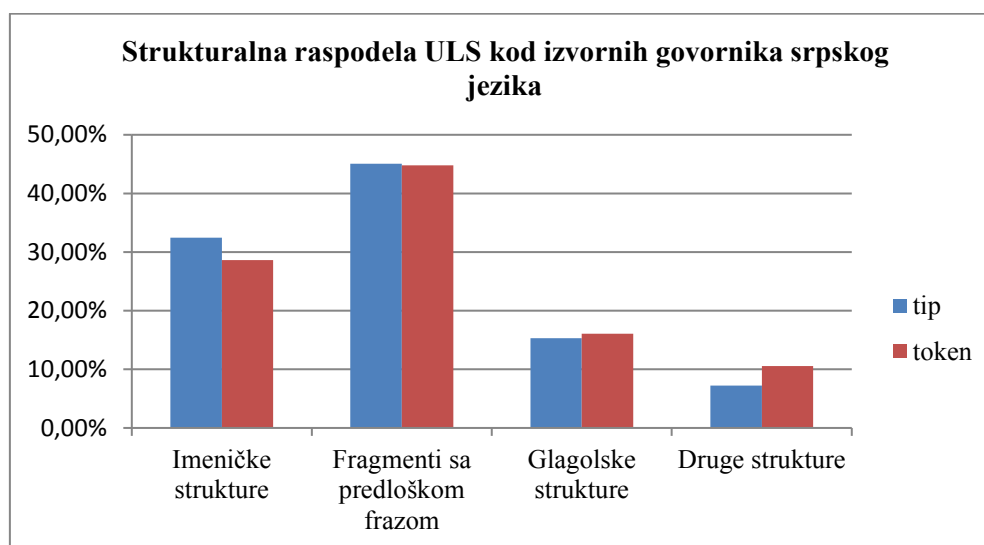
Grafikon 26. *Strukturalna raspodela ULS kod izvornih govornika engleskog jezika na podgrupe*

4.3.2. Strukturalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski

Strukturalna analiza ULS izvornih govornika srpskog jezika pri njihovoj produkciji na engleskom jeziku pokazuje da je najzastupljenija grupa *fragmenti sa predloškom frazom*, praćena grupom *imeničke strukture*, dok su grupe *glagolske strukture* i *druge strukture* manje zastupljene (v. Tabelu 35 i Grafikon 27). Navedeni podaci delimično potvrđuju drugu radnu hipotezu za izvorne govornike srpskog jezika: *Strukturalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski u naučnim radovima u oblasti obrazovanja biće okarakterisana većom zastupljenošću ULS iz strukturalnih grupa fragmenti sa predloškom frazom i glagolske strukture nego imeničke strukture i druge strukture*. Hipoteza je delimično potvrđena jer su *fragmenti sa predloškom frazom* najzastupljenija strukturalna kategorija ali nije u potpunosti potvrđena, jer su *glagolske strukture* manje zastupljene od *imeničkih struktura* u produkciji neizvornih govornika engleskog jezika kojima je maternji jezik srpski.

Tabela 35. Strukturalna raspodela ULS kod izvornih govornika srpskog jezika

Strukturalna grupa	Tip (Token)	Procenti
<i>Imeničke strukture</i>	36 (1298)	32.43 (28.62)
<i>Fragmenti sa predloškom frazom</i>	50 (2031)	45.05 (44.78)
<i>Glagolske strukture</i>	17 (728)	15.32 (16.05)
<i>Druge strukture</i>	8 (479)	10.56 (5)



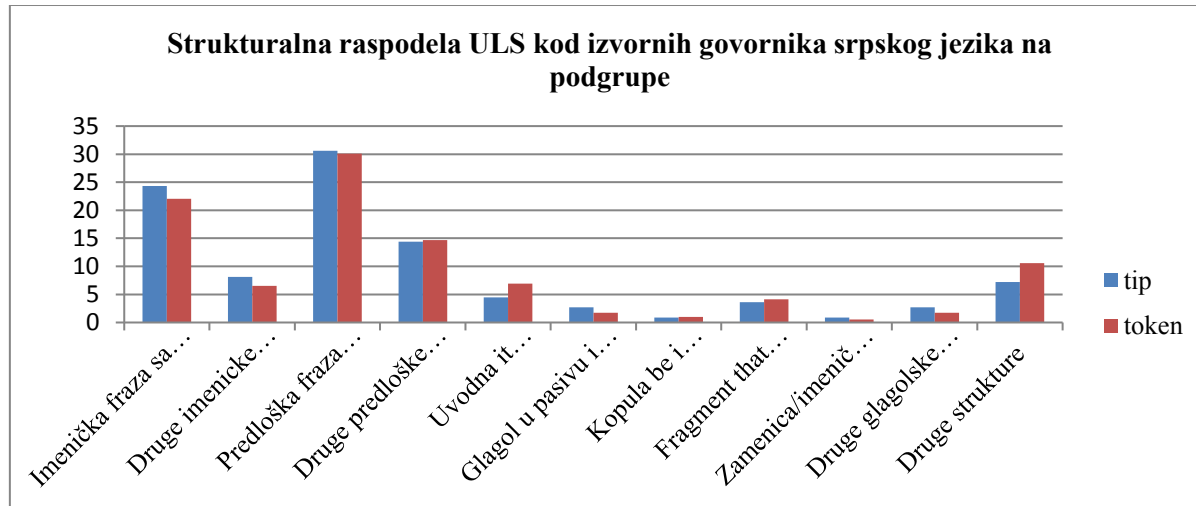
Grafikon 27. Strukturalna raspodela ULS kod izvornih govornika srpskog jezika

Pregledom procentualne zastupljenosti tipova i tokena u strukturalnim grupama u SCEP korpusu uviđamo da su raspodele uglavnom ujednačene. U strukturalnoj grupi *druge strukture*, malo je veći procenat zastupljenosti tokena, dok je u strukturalnoj grupi *imeničke strukture*, blago veći procenat zastupljenosti tipova. Stoga, možemo zaključiti da izvorni govornici srpskog jezika pri njihovoj produkciji na engleskom jeziku imaju tendenciju da koriste veću raznovrsnost tipova ULS iz grupe *imeničke strukture* a da ih ređe upotrebljavaju. Za grupu *druge strukture* je obrnuta situacija, odnosno, koristi se manja raznovrsnost tipova ali sa većom učestalošću.

Tabela 36. *Strukturalna raspodela ULS kod izvornih govornika srpskog jezika na podgrupe*

Grupa	Podgrupa	SCEP	
		Tip (Token)	Procenti
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	27 (1001)	24.32 (22.07)
	<i>Druge imeničke fraze</i>	9 (297)	8.11 (6.55)
<i>Fragmenti sa predložkom frazom</i>	<i>Predložka fraza sa of fragmentom</i>	34 (1365)	30.63 (30.09)
	<i>Druge predložke fraze</i>	16 (666)	14.41 (14.68)
<i>Glagolske strukture</i>	<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	5 (313)	4.50 (6.90)
	<i>Glagol u pasivu i fragment predložke fraze</i>	3 (78)	2.70 (1.72)
	<i>Kopula be i imenička/pridevska fraza</i>	1 (45)	0.90 (1)
	<i>Fragment that klauze (Glagol/pridev +) fragment to klauze</i>	4 (188)	3.60 (4.14)
	<i>Zamenica/imenička fraza+be (+...)</i>	1 (24)	0.90 (0.53)
	<i>Druge glagolske fraze</i>	3 (80)	2.71 (1.76)
<i>Druge strukture</i>	<i>Druge strukture</i>	8 (479)	7.22 (10.56)

Strukturalna raspodela ULS u SCEP korpusu na podgrupe pokazuje da je *predložke fraze sa of fragmentom* najzastupljenija podgrupa, praćena *imeničkim frazama sa of fragmentom* (v. Tabelu 36 i Grafikon 28). U korpusu neizvornih govornika engleskog jezika kojima je maternji jezik srpski nije zastupljena samo jedna podgrupa (*glagol/pridev +) fragment to klauze*.



Grafikon 28. *Strukturalna raspodela ULS kod izvornih govornika srpskog jezika na podgrupe*

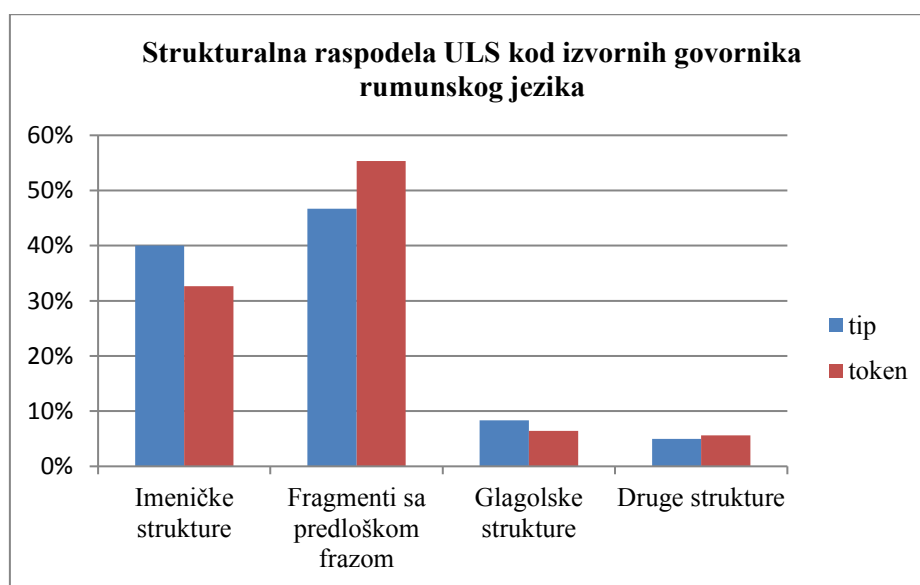
Kada sagledamo procentualnu zastupljenost tipova i tokena u strukturalnim podgrupama u SCEP korpusu, vidimo da se nijedna podgrupa ne ističe sa većim neujednačenim distribucijama između tipova i tokena. Kao što se pokazalo i u grupama, obe podgrupe *imeničkih struktura* imaju blago veću procentualnu zastupljenost tipova nego tokena. U podgrupama *glagolskih struktura*, dve najzastupljenije podgrupe *uvodna it struktura + glagolska ili pridevska fraza* i *fragment that klauze* imaju blago veću procentualnu zastupljenost tokena nego tipova, dok je u dve manje zastupljene podgrupe *glagol u pasivu i fragment predložke fraze* i *druge glagolske fraze* obrnut slučaj.

4.3.3. Strukturalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski

Strukturalna raspodela ULS izvornih govornika rumunskog jezika pri njihovoj produkciji na engleskom jeziku pokazuje da je najzastupljenija grupa *fragmenti sa predloškom frazom*, praćeni grupom *imeničke strukture* (v. Tabelu 37 i Grafikon 29). Navedeni rezultati delimično potvrđuju drugu hipotezu za izvorne govornike rumunskog jezika: *Strukturalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski u naučnim radovima u oblasti obrazovanja biće okarakterisana većom zastupljenošću ULS iz strukturalnih grupa fragmenti sa predloškom frazom i glagolske strukture nego imeničke strukture i druge strukture*. Kao što je slučaj i kod izvornih govornika srpskog jezika, hipoteza je delimično potvrđena, jer su *fragmenti sa predloškom frazom* najzastupljenija strukturalna grupa ali su *glagolske strukture* manje zastupljene od *imeničkih struktura*.

Tabela 37. Strukturalna raspodela ULS kod izvornih govornika rumunskog jezika

Strukturalna grupa	Tip (Token)	Procenti
<i>Imeničke strukture</i>	24 (770)	40 (32.65)
<i>Fragmenti sa predloškom frazom</i>	28 (1305)	46.67 (55.34)
<i>Glagolske strukture</i>	5 (151)	8.33 (6.4)
<i>Druge strukture</i>	3 (132)	5 (5.6)



Grafikon 29. Strukturalna raspodela ULS kod izvornih govornika rumunskog jezika

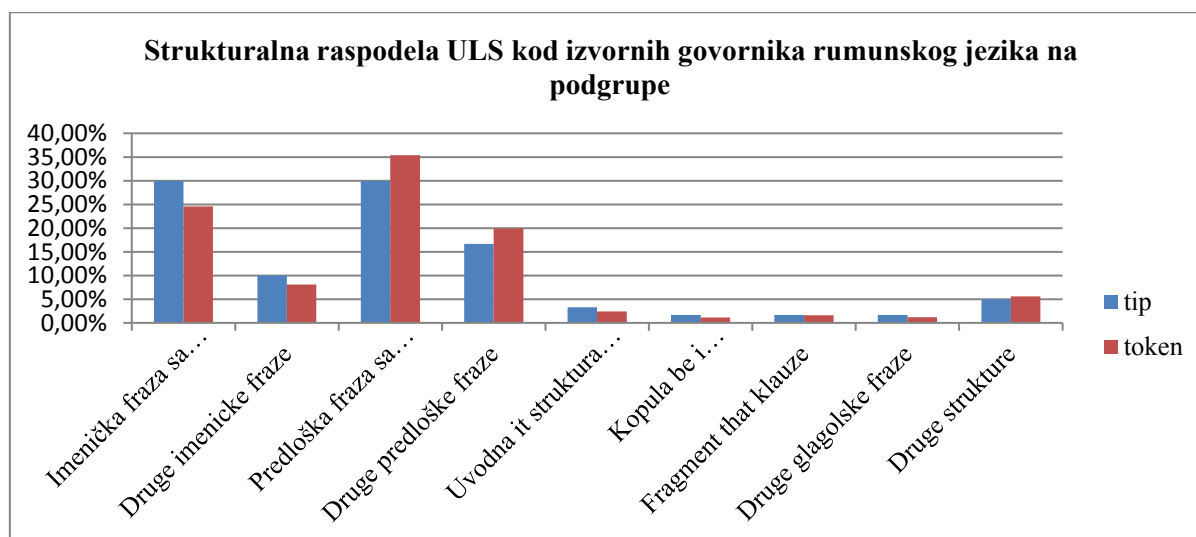
Razmatranjem procentualne raspodele zastupljenosti tipova i tokena u strukturalnim grupama u RCEP korpusu, uviđamo da postoje veće razlike u grupama *imeničke strukture* i *fragmenti sa predloškom frazom* nego u druge dve strukturalne kategorije. Naime, u *imeničkim strukturama* procentualno su zastupljeniji tipovi, dok su u *fragmentima sa predloškom frazom* zastupljeniji tokeni. Stoga možemo zaključiti da se u RCEP korpusu u grupi *imeničke strukture* koristi veća raznovrsnost ULS ali sa manjom učestalošću upotrebe, dok je za *fragmente sa predloškom frazom* situacija obrnuta.

Strukturalna raspodela ULS u RCEP korpusu na podgrupe pokazuje da su dve podgrupe najzastupljenije po tipovima *imenička fraza sa of fragmentom* i *predloška fraza sa of fragmentom* (v. Tabelu 38 i Grafikon 30). Kada sagledamo zastupljenost dve grupe po tokenima, uviđamo da je

mного zastupljenija podgrupa *predložka fraza sa of fragmentom* koja sačinjava 35,4% svih tokena, uporedno sa podgrupom *imenička fraza sa of fragmentom* koja sačinjava 24,6% svih tokena. *Glagolske strukture* su najslabije zastupljene. U RCEP korpusu četiri glagolske podgrupe od osam su zastupljene, i to tri sa po samo jednim tipom ULS.

Tabela 38. *Strukturalna raspodela ULS kod izvornih govornika rumunskog jezika na podgrupe – tipovi i tokeni*

Grupa	Podgrupa	RCEP	
		Tip (Token)	Procenti
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	18 (579)	30.0 (24.6)
	<i>Druge imeničke fraze</i>	6 (191)	10.0 (8.1)
<i>Fragmenti sa predložkom frazom</i>	<i>Predložka fraza sa of fragmentom</i>	18 (835)	30.0 (35.4)
	<i>Druge predložke fraze</i>	10 (470)	16.7 (19.9)
	<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	2 (57)	3.3 (2.4)
<i>Glagolske strukture</i>	<i>Glagol u pasivu i fragment predložke fraze</i>	/	/
	<i>Kopula be i imenička/pridevska fraza</i>	1 (27)	1.7 (1.1)
	<i>Fragment that klauze (Glagol/pridev +) fragment to klauze</i>	1 (38)	1.7 (1.6)
	<i>Zamenica/imenička fraza+be (+...)</i>	/	/
	<i>Druge glagolske fraze</i>	1 (29)	1.7 (1.2)
<i>Druge strukture</i>	<i>Druge strukture</i>	3 (132)	5.0 (5.6)



Grafikon 30. *Strukturalna raspodela ULS kod izvornih govornika rumunskog jezika na podgrupe*

4.3.4. Poređenje strukturalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik srpski

Kako bi se ispitale potencijalne razlike, tj. asocijacije, između ECEP i SCEP korpusa u odnosu na strukturalne distribucije ULS, korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između korpusa ECEP i SCEP u odnosu na strukturalne grupe – tipovi

Unakrsna tabela između ECEP i SCEP korpusa i strukturalnih grupa u odnosu na tipove prikazana je u Tabeli 39. Hi-kvadrat test nije bio značajan, $\chi^2(3, N = 138) = 0.68, p > .050$, ukazujući na to da nije identifikovana statistički značajna povezanost između korpusa i strukturalnih grupa. Drugim rečima, razlike koje uviđamo u strukturalnoj raspodeli tipova između ECEP i SCEP korpusa nisu statistički značajne.

Tabela 39. Unakrsna tabela između ECEP i SCEP korpusa i strukturalnih grupa u odnosu na tipove

Strukturalne grupe	SCEP			ECEP		
	tip	%	R	tip	%	R
<i>Imeničke strukture</i>	36	32.4%	0.3	8	29.6%	-0.3
<i>Fragmenti sa predloškom frazom</i>	50	45.0%	-0.6	14	51.9%	0.6
<i>Glagolske strukture</i>	17	15.3%	0.1	4	14.8%	-0.1
<i>Druge strukture</i>	8	7.2%	0.7	1	3.7%	-0.7
Total	111	100%		27	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

* – statistički značajni rezultati

Razlike između ECEP i SCEP korpusa u odnosu na strukturalne grupe – tokeni

Unakrsna tabela između ECEP i SCEP korpusa i strukturalnih grupa u odnosu na tokene prikazana je u Tabeli 40. Hi-kvadrat test je bio značajan, $\chi^2(3, N = 5313) = 18.52, p < .001$, što ukazuje na to da postoji statistički značajna povezanost između korpusa i strukturalnih grupa.

Korigovani standardizovani reziduali ukazuju da razlike između korpusa proizilaze iz različite upotrebe dve grupe govornika, u SCEP korpusu je viša upotreba *drugih struktura*, a u ECEP *fragmenata sa predloškom frazom*. Navedeni podaci nas navode do zaključka da izvorni govornici srpskog imaju tendenciju da isuviše učestalo koriste ULS iz grupe *druge strukture* a srazmerno nedovoljno učestalo ULS iz grupe *fragmenti sa predloškom frazom*. Činjenica da su statistički značajne razlike identifikovane u strukturalnoj raspodeli tokena ali ne i tipova, dovodi do sledećeg zaključka: izvorni govornici engleskog i srpskog jezika se ne razlikuju u raznolikosti ULS koji sačinjavaju različite strukturalne grupe, ali se razlikuju po učestalosti upotrebe ULS iz strukturalnih kategorija *fragmenti sa predloškom frazom* i *druge strukture*.

Tabela 40. Unakrsna tabela između ECEP i SCEP korpusa i strukturalnih grupa u odnosu na tokene

Strukturalne grupe	SCEP			ECEP		
	token	%	R	token	%	R
<i>Imeničke strukture</i>	1298	28.6%	0.5	216	27.8%	-0.5
<i>Fragmenti sa predloškom frazom</i>	2031	44.8%	-2.9*	392	50.5%	2.9*
<i>Glagolske strukture</i>	728	16.0%	0.2	122	15.7%	-0.2
<i>Druge strukture</i>	479	10.6%	3.9*	47	6.0%	-3.9*
Total	4536	100%		777	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između ECEP i SCEP korpusa u odnosu na strukturalne podgrupe– tipovi

Unakrsna tabela između ECEP i SCEP korpusa i strukturalnih podgrupa u odnosu na tipove prikazana je u Tabeli 41. Hi-kvadrat test nije bio značajan, χ^2 (11, N = 138) = 12.24, $p > .05$ i ukazuje na to da nije identifikovana statistički značajna povezanost između korpusa i strukturalnih podgrupa. Statistička analiza je pokazala da uočene razlike između strukturalnih podgrupa tipova ULS u ECEP i SCEP korpusima nije statistički značajna.

Tabela 41. Unakrsna tabela između korpusa i strukturalnih podgrupa u odnosu na tipove

Grupa	Podgrupa	SCEP			ECEP		
		tip	%	R	tip	%	R
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	27	24.3%	0.2	6	22.2%	-0.2
	<i>Druge imeničke fraze</i>	9	8.1%	0.1	2	7.4%	-0.1
<i>Fragmenti sa predloškom frazom</i>	<i>Predloška fraza sa of fragmentom</i>	34	30.6%	-0.6	10	37.0%	0.6
	<i>Druge predloške fraze</i>	16	14.4%	-0.1	4	14.8%	0.1
<i>Glagolske strukture</i>	<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	5	4.5%	-0.6	2	7.4%	0.6
	<i>Glagol u pasivu i fragment predloške fraze</i>	3	2.7%	0.9	0	0%	-0.9
	<i>Kopula be i imenička/pridevska fraza</i>	1	0.9%	0.5	0	0%	-0.5
	<i>Fragment that klauze (Glagol/pridev +) fragment to klauze</i>	4	3.6%	1.0	0	0%	-1.0
		0	0%	-2.9	2	7.4%	2.9
<i>Druge strukture</i>	<i>Zamenica/imenička fraza+be (+...)</i>	1	0.9%	0.5	0	0%	-0.5
	<i>Druge glagolske fraze</i>	3	2.7%	0.9	0	0%	-0.9
<i>Druge strukture</i>	<i>Druge strukture</i>	8	7.2%	0.7	1	3.7%	-0.7
Total		111	100%		27	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između ECEP i SCEP korpusa u odnosu na strukturalne podgrupe – tokeni

Unakrsna tabela između ECEP i SCEP korpusa i strukturalnih podgrupa u odnosu na tokene prikazana je u Tabeli 42. Hi-kvadrat test je bio značajan, χ^2 (10, N = 5313) = 428.96, $p < .001$, pokazuje da postoji statistički značajna povezanost između korpusa i strukturalnih podgrupa.

Među *imeničkim strukturama*, u SCEP korpusu je veća upotreba podgrupe *imeničke fraze sa of fragmentom*, dok je u ECEP veća upotreba podgrupe *druge imeničke fraze* (u odnosu na očekivane frekvencije). U ECEP korpusu je veća upotreba *predloških fraza sa of fragmentom*. Među *glagolskim strukturama* nije bilo razlike jedino za podgrupu *uvodna it struktura + glagolska ili pridevska fraza*, u ECEP korpusu je više zastupljena podgrupa (*glagol/pridev +*) *fragment to klauze*, dok su sve ostale podgrupe više zastupljene u SCEP korpusu. Navedeni podaci nam ukazuju da, kada je reč o podgrupama *imeničkih struktura*, izvorni govornici srpskog jezika imaju tendenciju da isuviše učestalo koriste ULS iz podgrupe *druge imeničke fraze* dok srazmerno nedovoljno često koriste ULS iz podgrupe *imeničke fraze sa of fragmentom*.

Tabela 42. Unakrsna tabela između korpusa i strukturalnih podgrupa u odnosu na tokene

Grupa	Podgrupa	SCEP			ECEP		
		token	%	R	token	%	R
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	1001	22.1%	2.5*	140	18.0%	-2.5*
	<i>Druge imeničke fraze</i>	297	6.5%	-3.3*	76	9.8%	3.3*
<i>Fragmenti sa predloškom frazom</i>	<i>Predloška fraza sa of fragmentom</i>	1365	30.1%	-2.7*	272	35.0%	2.7*
	<i>Druge predloške fraze</i>	666	14.7%	-0.6	120	15.4%	0.6
<i>Glagolske strukture</i>	<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	313	6.9%	-1.7	67	8.6%	1.7
	<i>Glagol u pasivu i fragment predloške fraze</i>	78	1.7%	3.7*	0	0%	-3.7*
	<i>Kopula be i imenička/pridevska fraza</i>	45	1.0%	2.8*	0	0%	-2.8*
	<i>Fragment that klauze (Glagol/pridev +) fragment to klauze</i>	188	4.1%	5.8*	0	0%	-5.8*
		0	0%	-18.0*	55	7.1%	18.0*
<i>Druge strukture</i>	<i>Zamenica/imenička fraza+be (+...)</i>	24	0.5%	2.0*	0	0%	-2.0*
	<i>Druge glagolske fraze</i>	80	1.8%	3.7*	0	0%	-3.7*
	<i>Druge strukture</i>	479	10.6%	3.9*	47	6.0%	-3.9*
Total		4536	100%		777	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Kada je reč o *fragmentima sa predloškom frazom*, izvorni govornici srpskog jezika imaju tendenciju da proporcionalno nedovoljno često koriste ULS iz podgrupe *predloške fraze sa of fragmentom*, dok za drugu podgrupu nisu identifikovane statistički značajne razlike. Kod glagolskih podgrupa, izvorni govornici srpskog imaju tendenciju da nedovoljno često koriste ULS iz podgrupe (*glagol/pridev +*) *fragment to klauze*, dok imaju tendenciju da isuviše učestalo koriste ULS iz podgrupa *glagol u pasivu i fragment predloške fraze, kopula be i imenička/pridevska fraza, fragment that klauze, zamenica/imenička fraza+be (+...)* i *druge glagolske fraze*.

Analizom strukturalne raspodele ULS u ECEP i SCEP korpusima na podgrupe u odnosu na tipove i tokene uviđamo da se izvorni govornici engleskog i srpskog jezika ne razlikuju značajno u raznolikosti ULS koje koriste u strukturalnim podgrupama ali da se razlikuju u učestalosti upotrebe ULS u deset podgrupa.

4.3.5. Poređenje strukturalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik rumunski

Kako bi se ispitale potencijalne razlike, tj. asocijacije, između ECEP i RCEP korpusa u odnosu na strukturalne distribucije ULS korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između ECEP i RCEP korpusa u odnosu na strukturalne grupe – tipovi

Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih grupa u odnosu na tipove prikazana je u Tabeli 43. Hi-kvadrat test nije bio značajan, $\chi^2(3, N = 87) = 1.47, p > .050$, što ukazuje na to da nije identifikovana statistički značajna povezanost između korpusa i strukturalnih grupa. Drugim rečima, razlike koje uviđamo u strukturalnoj raspodeli tipova u ECEP i RCEP korpusima nije statistički značajna.

Tabela 43. *Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih grupa u odnosu na tipove*

Strukturalne grupe	RCEP			ECEP		
	tip	%	R	tip	%	R
<i>Imeničke strukture</i>	24	40.0%	0.9	8	29.6%	-0.9
<i>Fragmenti sa predloškom frazom</i>	28	46.7%	-0.4	14	51.9%	0.4
<i>Glagolske strukture</i>	5	8.3%	-0.9	4	14.8%	0.9
<i>Druge strukture</i>	3	5.0%	0.3	1	3.7%	-0.3
Total	60	100%		27	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između ECEP i RCEP korpusa u odnosu na strukturalne grupe – tokeni

Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih grupa u odnosu na tokene prikazana je u Tabeli 44. Hi-kvadrat test je bio značajan, $\chi^2(3, N = 3135) = 65.19, p < .001$, što ukazuje na to da postoji statistički značajna povezanost između korpusa i strukturalnih grupa.

Tabela 44. Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih grupa u odnosu na tokene

Strukturalne grupe	RCEP			ECEP		
	token	%	R	token	%	R
<i>Imeničke strukture</i>	770	32.7%	2.5*	216	27.8%	-2.5*
<i>Fragmenti sa predloškom frazom</i>	1305	55.3%	2.4*	392	50.5%	-2.4*
<i>Glagolske strukture</i>	151	6.4%	-8.0*	122	15.7%	8.0*
<i>Druge strukture</i>	132	5.6%	-0.5	47	6.0%	0.5
Total	2358	100%		777	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

* – statistički značajni rezultati

Korigovani standardizovani reziduali ukazuju na to da je zastupljenost *imeničkih struktura* i *fragmenata sa predloškom frazom* viša u RCEP korpusu, dok je u ECEP korpusu više zastupljenost *glagolskih struktura*. Navedeni podaci dovode do zaključka da izvorni govornici rumunskog jezika u svojoj produkciji na engleskom u naučnim radovima iz oblasti obrazovanja proporcionalno isuviše učestalo koriste ULS iz strukturalnih kategorija *imeničke strukture* i *fragmenti sa predloškom frazom*, dok nedovoljno učestalo koriste ULS iz strukturalne grupe *glagolske strukture*.

Na osnovu činjenice da su se statistički značajne razlike utvrdile u strukturalnoj raspodeli tokena ULS u grupe, ali da se nisu utvrdile za tipove, možemo zaključiti da nema značajnih razlika u raspodeli ULS u glavne strukturalne kategorije, ali da se one uviđaju kada uzmemo u obzir učestalost upotrebe ULS iz navedenih kategorija.

Razlike između ECEP i RCEP korpusa u odnosu na strukturalne podgrupe– tipovi

Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih podgrupa u odnosu na tipove prikazana je u Tabeli 45. Hi-kvadrat test nije bio značajan, χ^2 (9, N = 138) = 7.40, $p > .050$, ukazujući na to da nije identifikovana statistički značajna povezanost između korpusa i strukturalnih podgrupa. Drugim rečima, statistička analiza je pokazala da razlike u raspodeli tipova ULS na strukturalne potkategorije između SCEP i RCEP korpusa nisu statistički značajne.

Tabela 45. Unakrsna tabela između korpusa i strukturalnih podgrupa u odnosu na tipove

Grupa	Podgrupa	RCEP			ECEP		
		tip	%	R	tip	%	R
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	18	30.0%	0.8	6	22.2%	-0.8
	<i>Druge imeničke fraze</i>	6	10.0%	0.4	2	7.4%	-0.4
<i>Fragmenti sa predloškom frazom</i>	<i>Predloška fraza sa of fragmentom</i>	18	30.0%	-0.6	10	37.0%	0.6
	<i>Druge predloške fraze</i>	10	16.7%	0.2	4	14.8%	-0.2
<i>Glagolske strukture</i>	<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	2	3.3%	-0.8	2	7.4%	0.8
	<i>Glagol u pasivu i fragment predloške fraze</i>	/	/	/	/	/	/
	<i>Kopula be i imenička/pridevska fraza</i>	1	1.7%	0.7	0	0%	-0.7
	<i>Fragment that klauze (Glagol/pridev +) fragment to klauze</i>	1	1.7%	0.7	0	0%	-0.7
		0	0%	-2.1	2	7.4%	2.1

Tabela 45. Unakrsna tabela između korpusa i strukturalnih podgrupa u odnosu na tipove

	Zamenica/imenička fraza+be (+...)	/	/	/	/	/	/
Druge strukture	Druge glagolske fraze	1	1.7%	0.7	0	0%	-0.7
	Druge strukture	3	5.0%	0.3	1	3.7%	-0.3
Total		60	100%		27	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između ECEP i RCEP korpusa u odnosu na strukturalne podgrupe – tokeni

Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih podgrupa u odnosu na tokene prikazana je u Tabeli 46. Hi-kvadrat test je bio značajan, $\chi^2(9, N = 3135) = 274.10$, $p < .001$, što pokazuje da postoji statistički značajna povezanost između korpusa i strukturalnih podgrupa.

Tabela 46. Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih podgrupa u odnosu na tokene

Grupe	Podgrupa	RCEP			ECEP		
		token	%	R	token	%	R
Imeničke strukture	Imenička fraza sa of fragmentom	579	24.6%	3.8*	140	18.0%	-3.8*
	Druge imeničke fraze	191	8.1%	-1.5	76	9.8%	1.5
Fragmenti sa predloškom frazom	Predloška fraza sa of fragmentom	835	35.4%	0.2	272	35.0%	-0.2
	Druge predloške fraze	470	19.9%	2.8*	120	15.4%	-2.8*
	Uvodna it struktura + glagolska ili pridevska fraza	57	2.4%	-7.7*	67	8.6%	7.7*
Glagolske strukture	Glagol u pasivu i fragment predloške fraze	/	/	/	/	/	/
	Kopula be i imenička/pridevska fraza	27	1.1%	3.0*	0	0%	-3.0*
	Fragment that klauze	38	1.6%	3.6*	0	0%	-3.6*
	(Glagol/pridev +) fragment to klauze	0	0	-13.0*	55	7.1%	13.0*
	Zamenica/imenička fraza+be (+...)	/	/	/	/	/	/
Druge strukture	Druge glagolske fraze	29	1.2%	3.1*	0	0%	-3.1*
	Druge strukture	132	5.6%	-0.5	47	6.0%	0.5
Total		2358	100.0%		777	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Korigovani standardizovani reziduali ukazuju na to da je relativna upotreba imeničkih fraza sa of fragmentom i drugih predloških fraza frekventnija u RCEP korpusu. U slučaju *glagolskih struktura*, u ECEP korpusu je frekventnija *uvodna it struktura + glagolska ili pridevska fraza* i *(glagol/pridev +) fragment to klauze*, a u RCEP *kopula be i imenička/pridevska fraza*, *fragment that klauze* i *druge glagolske fraze*. Na osnovu navedenih podataka uviđamo da izvorni govornici rumunskog jezika

imaju tendenciju da srazmerno isuviše učestalo koriste ULS u podgrupama *imenička fraza sa of fragmentom, druge predložke fraze, kopula be i imenička/pridevska fraza, fragment that klauze i druge glagolske fraze*, dok nedovoljno učestalo koriste ULS u podgrupama *uvodna it struktura + glagolska ili pridevska fraza i (glagol/pridev +) fragment to klauze*.

Činjenica da su se statistički značajne razlike identifikovale u strukturalnoj raspodeli ULS u RCEP i ECEP korpusu na tokene ali ne i na tipove, ukazuje na to da se ova dva korpusa ne razlikuju značajno po raspodelu ULS na podgrupe već da se razlikuju po tome koliko se često koriste navedeni ULS u podgrupama. Raspodelom tokena na grupe uvideli smo da su *imeničke strukture i fragmenti sa predložkom frazom* više zastupljene u korpusu izvornih govornika rumunskog, dok su *glagolske strukture* više zastupljenije kod izvornih govornika engleskog. Međutim, kada smo sagledali i strukturalnu raspodelu na podgrupe uvideli smo da su samo dve podgrupe zastupljenije kod izvornih govornika engleskog, dok su tri zastupljenije kod izvornih govornika rumunskog. Zbog ovakvih razlika u podacima bitno je uraditi analize i na nivou grupa i na nivou podgrupa. Štaviše, informacije o razlikama u zastupljenosti i na nivou grupa i na nivou podgrupa su značajne u procesu podučavanja i adaptaciji nastave potrebama i karakteristikama studenata. Na osnovu podataka u ovom segmentu možemo da zaključimo da bi u radu sa izvornim govornicima rumunskog jezika bilo poželjno obraditi sa studentima/prevodiocima alternativne načine izražavanja značenja za koje sada koriste ULS iz podgrupe *imenička fraza sa of fragmentom, druge predložke fraze, kopula be i imenička/pridevska fraza, fragment that klauze i druge glagolske fraze*, a trebalo bi podstaći češću upotrebu ULS iz podgrupa *uvodna it struktura + glagolska ili pridevska fraza i (glagol/pridev +) fragment to klauze*.

4.3.6. Poređenje strukturalne raspodele ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski ili srpski

Kako bi se ispitale potencijalne razlike, tj. asocijacije, između SCEP i RCEP korpusa u odnosu na strukturalne distribucije ULS korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između SCEP i RCEP korpusa u odnosu na strukturalne grupe – tipovi

Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih grupa u odnosu na tipove prikazana je u Tabeli 47. Hi-kvadrat test nije bio značajan, $\chi^2(3, N = 171) = 2.42, p > .050$, odnosno nije identifikovana statistički značajna povezanost između korpusa i strukturalnih grupa. Navedeno pokazuje da uočene razlike u strukturalnoj raspodeli ULS na tipove u SCEP i RCEP korpusima nisu značajne.

Tabela 47. Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih grupa u odnosu na tipove

Strukturalne grupe	SCEP			RCEP		
	tip	%	R	tip	%	R
<i>Imeničke strukture</i>	36	32.4%	-1.0	24	40.0%	1.0
<i>Fragменти sa predloškom frazom</i>	50	45.0%	-0.2	28	46.7%	0.2
<i>Glagolske strukture</i>	17	15.3%	1.3	5	8.3%	-1.3
<i>Druge strukture</i>	8	7.2%	0.6	3	5.0%	-0.6
Total	111	100%		60	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između SCEP i RCEP korpusa u odnosu na strukturalne grupe

Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih grupa u odnosu na tokene prikazana je u Tabeli 48. Hi-kvadrat test je bio značajan, $\chi^2(3, N = 6894) = 200.56, p < .001$ i ukazuje na to da postoji statistički značajna povezanost između korpusa i strukturalnih grupa.

Korigovani standardizovani reziduali ukazuju na to da postoje značajne razlike u svim strukturalnim grupama između korpusa. U SCEP korpusu je veća zastupljenost grupa *glagolske strukture* i *druge strukture*, dok je u RCEP veća zastupljenost grupa *imeničke strukture* i *fragmenti sa predloškom frazom*. Činjenica da razlike u strukturalnoj raspodeli tipova nisu statistički značajne dok razlike u raspodeli tokena jesu, ukazuje na to da se dve grupe neizvornih govornika ne razlikuju u strukturalnoj raspodeli identifikovanih ULS već samo u učestalosti njihove upotrebe.

Tabela 48. Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih grupa u odnosu na tokene

Strukturalne grupe	SCEP			RCEP		
	token	%	R	token	%	R
<i>Imeničke strukture</i>	1298	28.6%	-3.5*	770	32.7%	3.5*
<i>Fragmenti sa predloškom frazom</i>	2031	44.8%	-8.3*	1305	55.3%	8.3*
<i>Glagolske strukture</i>	728	16.0%	11.4*	151	6.4%	-11.4*
<i>Druge strukture</i>	479	10.6%	6.9*	132	5.6%	-6.9*
Total	4536	100%		2358	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između SCEP i RCEP korpusa u odnosu na strukturalne podgrupe

Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih podgrupa u odnosu na tipove prikazana je u Tabeli 49. Hi-kvadrat test nije bio značajan, $\chi^2(10, N = 171) = 4.23, p > .050$, odnosno nije identifikovana statistički značajna povezanost između korpusa i strukturalnih podgrupa.

Tabela 49. Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih podgrupa u odnosu na tipove

Grupa	Podgrupa	SCEP			RCEP		
		tip	%	R	tip	%	R
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	27	24.3%	-0.8	18	30.0%	0.8
	<i>Druge imeničke fraze</i>	9	8.1%	-0.4	6	10.0%	0.4
<i>Fragmenti sa predloškom frazom</i>	<i>Predloška fraza sa of fragmentom</i>	34	30.6%	0.1	18	30.0%	-0.1
	<i>Druge predloške fraze</i>	16	14.4%	-0.4	10	16.7%	0.4
<i>Glagolske strukture</i>	<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	5	4.5%	0.4	2	3.3%	-0.4
	<i>Glagol u pasivu i fragment predloške fraze</i>	3	2.7%	1.3	0	0%	-1.3
	<i>Kopula be i imenička/pridevska fraza</i>	1	0.9%	-0.4	1	1.7%	0.4
	<i>Fragment that klauze (Glagol/pridev +) fragment to klauze</i>	4	3.6%	0.7	1	1.7%	-0.7
	<i>Zamenica/imenička fraza+be (+...)</i>	1	0.9%	0.7	0	0%	-0.7
	<i>Druge strukture</i>	<i>Druge glagolske fraze</i>	3	2.7%	0.4	1	1.7%
	<i>Druge strukture</i>	8	7.2%	0.6	3	5.0%	-0.6
Total		111	100%		60	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

* – statistički značajni rezultati

Razlike između SCEP i RCEP korpusa u odnosu na strukturalne podgrupe

Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih podgrupa u odnosu na tokene prikazana je u Tabeli 50. Hi-kvadrat test je bio značajan, $\chi^2(10, N = 6894) = 236.90$, $p < .001$, što ukazuje na to da postoji statistički značajna povezanost između korpusa i strukturalnih podgrupa.

Korigovani standardizovani reziduali ukazuju da je veća upotreba grupa *imeničke strukture* i *fragmenti sa predloškom frazom* u RCEP korpusu, dok je veća upotreba grupa *druge strukture* i određenih podgrupa *glagolskih struktura* u SCEP korpusu, konkretno: *uvodna it struktura + glagolska ili pridevska fraza*, *glagol u pasivu i fragment predloške fraze*, *fragment that klauze* i *zamenica/imenička fraza+be (+...)*.

Na osnovu činjenice da statistički značajne razlike nisu identifikovane u strukturalnoj raspodeli tipova na grupe i podgrupe a jesu za tokene, možemo zaključiti da se upotreba izvornih govornika srpskog i rumunskog jezika ne razlikuje u strukturalnoj raspodeli identifikovanih ULS već da se razlikuju po učestalosti upotrebe izdvojenih ULS unutar tih grupa i podgrupa. Statistička analiza strukturalnih grupa i podgrupa ULS u SCEP i RCEP korpusima jasno je pokazala da izvorni govornici rumunskog imaju tendenciju da učestalije koriste ULS u grupama *imeničke strukture* i *fragmenti sa predloškim frazama*, dok izvorni govornici srpskog imaju tendenciju da učestalije koriste ULS iz grupa *glagolske strukture* i *druge strukture*.

Tabela 50. Unakrsna tabela između korpusa i strukturalnih podgrupa u odnosu na tokene

Grupa	Podgrupa	SCEP		R	RCEP		R
		token	%		token	%	
Imeničke strukture	Imenička fraza sa of fragmentom	1001	22.1%	-2.3*	579	24.6%	2.3*
	Druge imeničke fraze	297	6.5%	-2.4*	191	8.1%	2.4*
Fragmenti sa predloškom frazom	Predloška fraza sa of fragmentom	1365	30.1%	-4.5*	835	35.4%	4.5*
	Druge predloške fraze	666	14.7%	-5.6*	470	19.9%	5.6*
	Uvodna it struktura + glagolska ili pridevska fraza	313	6.9%	7.8*	57	2.4%	-7.8*
Glagolske strukture	Glagol u pasivu i fragment predloške fraze	78	1.7%	6.4*	0	0%	-6.4*
	Kopula be i imenička/pridevska fraza	45	1.0%	-0.6	27	1.1%	0.6
	Fragment that klauze (Glagol/pridev +) /	188	4.1%	5.6*	38	1.6%	-5.6*
	fragment to klauze	/	/	/	/	/	/
	Zamenica/imenička fraza+be (+...)	24	0.5%	3.5*	0	0%	-3.5*
Druge strukture	Druge glagolske fraze	80	1.8%	1.7	29	1.2%	-1.7
	Druge strukture	479	10.6%	6.9*	132	5.6%	-6.9*
Total		4536	100%		2358	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

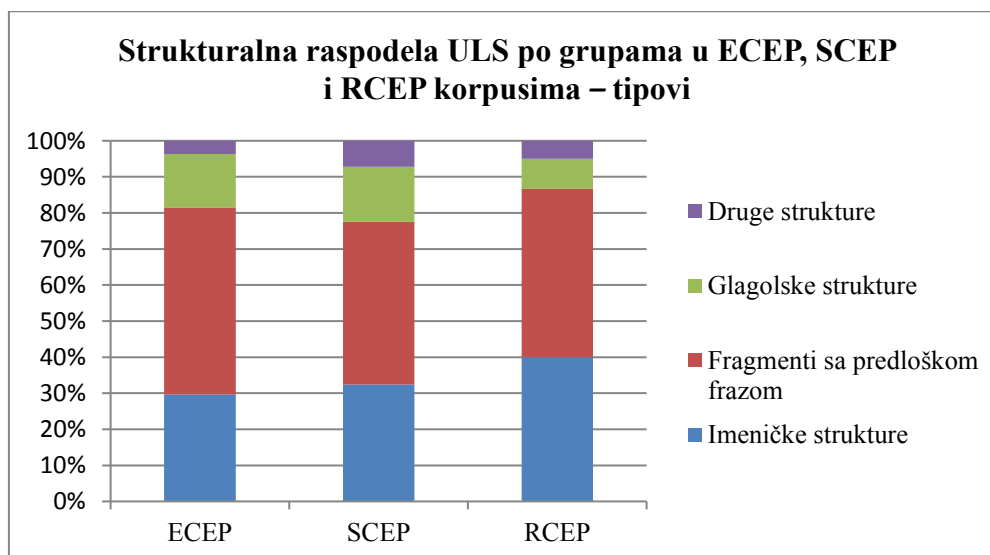
* – statistički značajni rezultati

4.3.7. Poređenje strukturalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik srpski ili rumunski

Poređenjem strukturalne raspodele tipova ULS na grupe kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik srpski i rumunski, uvidamo da su raspodele poprilično ujednačene, što je i potvrđeno činjenicom da statistička analiza podataka nije pokazala značajnu razliku u strukturalnim raspodelama tipova na grupe i podgrupe ni na jednoj relaciji između korpusa (v. Tabelu 51 i Grafikon 31).

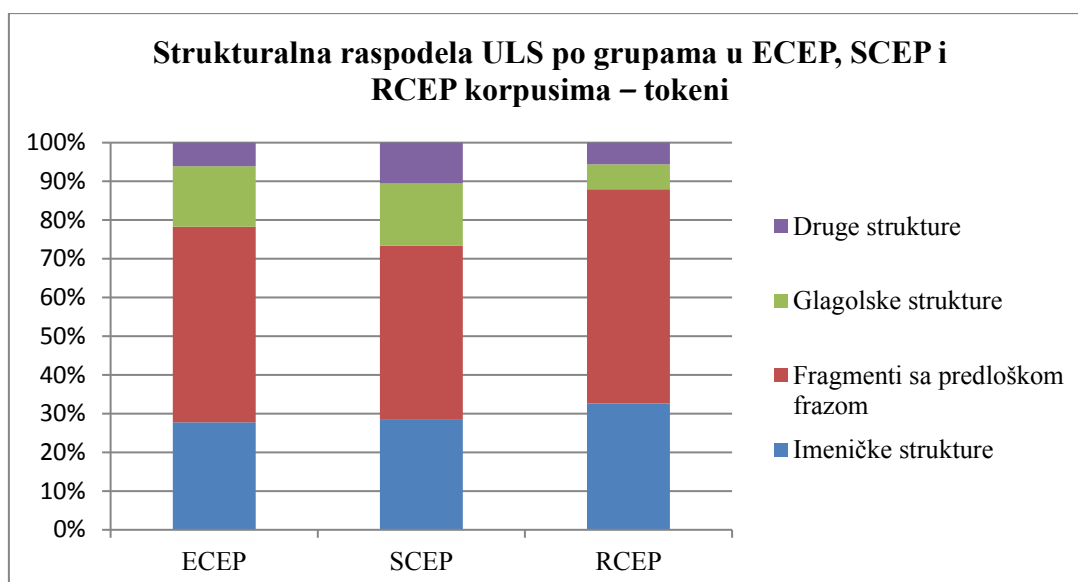
Tabela 51. *Strukturalna raspodela ULS po grupama u ECEP, SCEP i RCEP korpusima – tipovi*

	<i>Imeničke strukture</i>	<i>Fragmenti sa predloškom frazom</i>	<i>Glagolske strukture</i>	<i>Druge strukture</i>
ECEP	29.63%	51.85%	14.81%	3.7%
SCEP	32.43%	45.05%	15.32%	7.21%
RCEP	40%	46.67%	8.33%	5%



Grafikon 31. *Strukturalna raspodela ULS po grupama u ECEP, SCEP i RCEP korpusima – tipovi*

Poređenjem strukturalne raspodele tokena ULS u tri korpusa uvidamo da je zastupljenost strukturalnih kategorija sličnija između ECEP i SCEP korpusa nego što je između ECEP i RCEP korpusa (v. Grafikon 32 i Tabelu 52). Navedene razlike u raspodeli strukturalnih grupa tokena potvrđene su i statističkom analizom u kojoj je statistički značajna razlika u strukturalnoj raspodeli grupa za ECEP i SCEP korpus ustanovljena za dve grupe (*druge strukture* i *fragmenti sa predloškom frazom*), dok je za ECEP i RCEP korpus ustanovljena za tri grupe (*imeničke strukture*, *fragmenti sa predloškom frazom* i *glagolske strukture*). Najveće razlike u strukturalnoj raspodeli tokena na grupe identifikovana je između SCEP i RCEP korpusa, gde su se statistički značajne razlike identifikovale u raspodeli sve četiri strukturalne grupe.



Grafikon 32. *Strukturalna raspodela ULS po grupama u ECEP, SCEP i RCEP korpusima – tokeni*

Tabela 52. *Strukturalna raspodela ULS po grupama u ECEP, SCEP i RCEP korpusima – tokeni*

	Imeničke strukture	Fragmenti sa predloškom frazom	Glagolske strukture	Druge strukture
ECEP	27.8%	50.45%	15.7%	6.05%
SCEP	28.62%	44.78%	16.05%	10.56%
RCEP	32.65%	55.34%	6.4%	5.6%

Činjenica da su statistički značajne razlike identifikovane za strukturalnu raspodelu tokena ali da nisu identifikovane za strukturalnu raspodelu tipova ukazuje na to da se tri korpusa ne razlikuju značajno u raspodeli ULS na strukturalne grupe ali da se razlikuju u tome koliko učestalo koriste ULS iz tih grupa.

4.4. Funkcionalna raspodela ULS

Sa ciljem boljeg razumevanja karakteristika upotrebe ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski i rumunski, urađena je funkcionalna raspodela izdvojenih ULS u sva tri korpusa. Kako bi se utvrdila značajnost uočenih razlika u funkcionalnoj raspodeli ULS, urađena je i statistička analiza podataka primenom hi-kvadrat testa. Pri funkcionalnoj analizi ULS, oni su se prvo raspodelili u glavne funkcionalne grupe, pa zatim i na funkcionalne podgrupe. Radi precizne prezentacije i interpretacije podataka, strukturalna zastupljenost i tipova i tokena predstavljena je za funkcionalne grupe i podgrupe za sva tri korpusa. Za funkcionalnu raspodelu ULS koristi se taksonomija koju je predložio Hajland (2008).

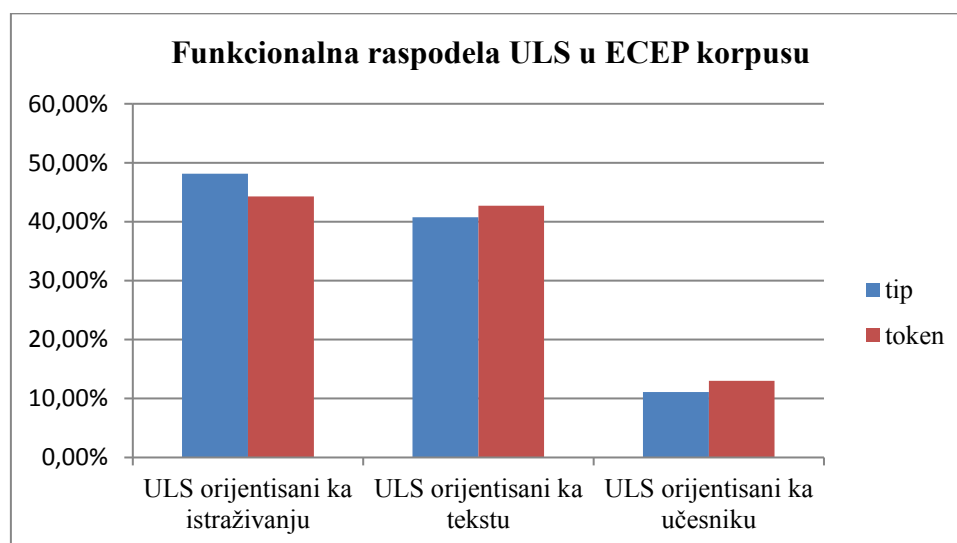
4.4.1. Funkcionalna raspodela ULS kod izvornih govornika engleskog jezika

Funkcionalna raspodela ULS izvornih govornika engleskog jezika na grupe pokazala je da su najučestaliji ULS orijentisani ka istraživanju, dok su najmanje zastupljeni ULS orijentisani ka učesniku, v. Tabelu 53 i Grafikon 33.

Tabela 53. Funkcionalna raspodela ULS za izvorne govornike engleskog jezika

Funkcionalna kategorija	Tipovi (tokeni)	Procenti
<i>ULS orijentisani ka istraživanju</i>	13 (344)	48.1 (44.3)
<i>ULS orijentisani ka tekstu</i>	11 (332)	40.7 (42.7)
<i>ULS orijentisani ka učesniku</i>	3 (101)	11.1 (13.0)

Poređenje procentualne raspodele tipova i tokena ULS izvornih govornika engleskog jezika po funkcionalnim kategorijama uviđamo da su zastupljeniji tipovi u ULS orijentisanim ka istraživanju dok su blago zastupljeniji tokeni u druge dve funkcionalne kategorije. Stoga možemo zaključiti da izvorni govornici engleskog jezika imaju tendenciju da koriste raznolike ULS u grupi *ULS orijentisani ka istraživanju* ali sa manjom učestalošću, dok je za *ULS orijentisane ka tekstu* i *učesniku* blago veći stepen učestalosti upotrebe ULS.



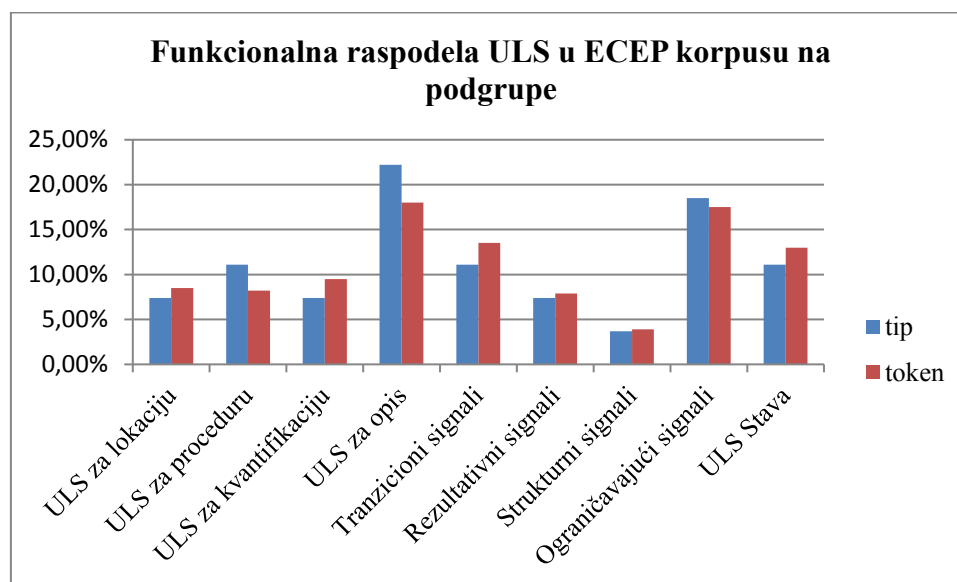
Grafikon 33. Funkcionalna raspodela ULS u ECEP korpusu

Funkcionalna raspodela ULS izvornih govornika engleskog jezika na podgrupe pokazuje da je najučestalija podgrupa *ULS za opis*, praćeni *ograničavajućim signalima* i *ULS stava* (v. Tabelu 54 i Grafikon 34). Najmanje je zastupljena podgrupa strukturalni signali sa samo jednim tipom.

Tabela 54. Funkcionalna raspodela ULS u ECEP korpusu na podgrupe

Grupa	Podgrupa	Tip (token)	Procenti
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	2 (66)	7.4 (8.5)
	<i>ULS za proceduru</i>	3 (64)	11.1 (8.2)
	<i>ULS za kvantifikaciju</i>	2 (74)	7.4 (9.5)
	<i>ULS za opis</i>	6 (140)	22.2 (18.0)
<i>ULS orijentisani ka tekstu</i>	<i>Tranzicioni signali</i>	3 (105)	11.1 (13.5)
	<i>Rezultativni signali</i>	2 (61)	7.4 (7.9)
	<i>Strukturalni signali</i>	1 (30)	3.7 (3.9)
	<i>Ograničavajući signali</i>	5 (136)	18.5 (17.5)
<i>ULS orijentisani ka učesniku</i>	<i>ULS stava</i>	3 (101)	11.1 (13.0)
	<i>ULS uključivanja</i>	/	/
Total		27 (777)	100%

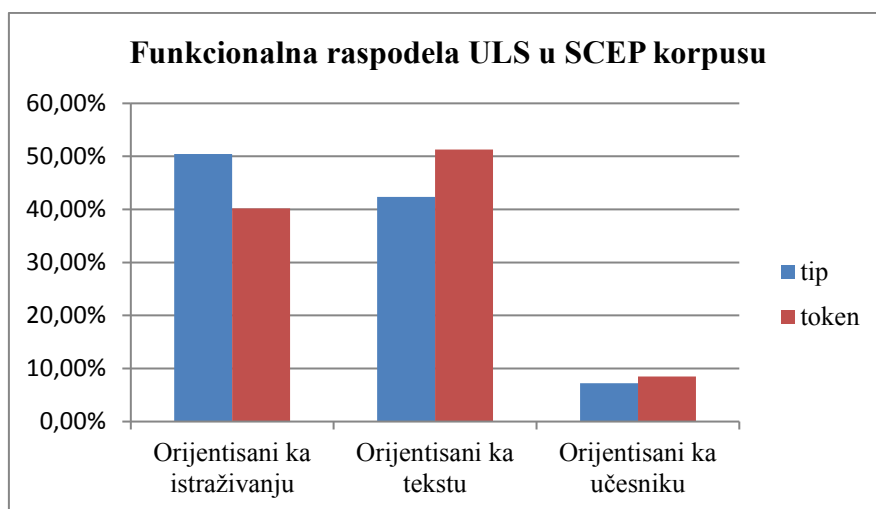
Poređenjem procentualne zastupljenosti tipova i tokena na funkcionalne podgrupe ULS izvornih govornika engleskog jezika uvidamo da se najveća razlika nalazi u podgrupi *ULS za opis*, gde su više zastupljeni tipovi nego tokeni. Veće zastupljenosti tipova ali sa manjom razlikom se, takođe, uviđaju u podgrupama *ULS za proceduru* i *ograničavajući signali*. U drugih šest podgrupa veća je zastupljenost tokena, što nas dovodi do zaključka da izvorni govornici engleskog jezika u većini funkcionalnih podgrupa procentualno češće koriste ograničeni broj ULS, dok u tri funkcionalne potkategorije pokazuju veću raznolikost ULS nego učestalost njihove upotrebe.



Grafikon 34. Funkcionalna raspodela ULS u ECEP korpusu na podgrupe

4.4.2. Funkcionalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski

Funkcionalna raspodela ULS kod izvornih govornika srpskog jezika pokazuje da je najzastupljenija funkcionalna kategorija što se tiče tipova *ULS orijentisani ka istraživanju* a da je najzastupljenija kategorija što se tiče tokena – *ULS orijentisani ka tekstu* (v. Grafikon 35 i Tabelu 55). Stoga možemo zaključiti, s jedne strane, da izvorni govornici srpskog jezika u naučnim radovima iz oblasti obrazovanja koriste veću raznovrsnost *ULS orijentisanih ka istraživanju* ali sa manjom učestalošću, a sa druge, koriste manju raznolikost *ULS orijentisanih ka tekstu* ali sa većom učestalošću. Funkcionalna kategorija *ULS orijentisani ka učesniku* je najmanje zastupljena i što se tiče tokena i što se tiče tipova.



Grafikon 35. Funkcionalna raspodela ULS u SCEP korpusu

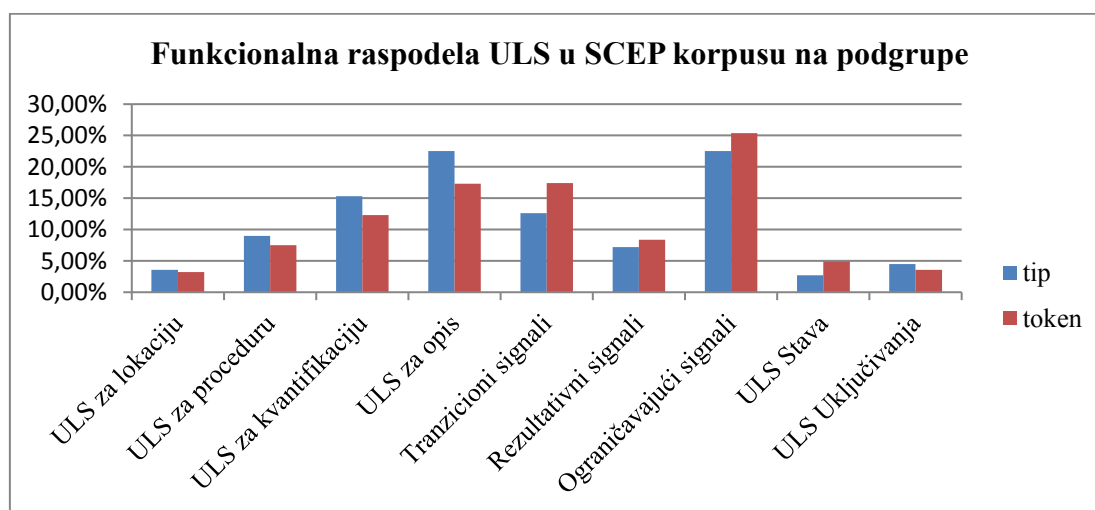
Tabela 55. Funkcionalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski

Funkcionalna kategorija	Tipovi (tokeni)	Procenti
<i>ULS orijentisani ka istraživanju</i>	56 (1824)	50.5 (40.2)
<i>ULS orijentisani ka tekstu</i>	47 (2326)	42.3% (51.3)
<i>ULS orijentisani ka učesniku</i>	8 (386)	7.2% (8.5)

Funkcionalna raspodela ULS izvornih govornika srpskog jezika na podgrupe nam pokazuje da su dve podgrupe (*ULS za opis* i *Ograničavajući signali*) najzastupljenije kada gledamo procentualnu raspodelu tipova i sačinjavaju 22, 5% svih tipova u SCEP korpusu (v. Tabelu 56 i Grafikon 36). Kada gledamo procentualnu zastupljenost tokena, uvidamo da je najzastupljenija funkcionalna podgrupa *ograničavajući signali*, koji sačinjavaju 25,4% svih tokena u SCEP korpusu. Najmanje zastupljena funkcionalna potkategorija na osnovu tipova je *ULS stava*, a po tokenima *ULS za lokaciju*. Na osnovu poređenja procentualne zastupljenosti tipova i tokena u funkcionalnoj raspodeli ULS zaključujemo da izvorni govornici srpskog jezika prave velike i česte razlike u funkcionalnoj raspodeli tipova i tokena po funkcionalnim podgrupama.

Tabela 56. Funkcionalna raspodela ULS u SCEP korpusu na podgrupe

Grupa	Podgrupa	Tip (token)	Procenti
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	4 (145)	3.6 (3.2)
	<i>ULS za proceduru</i>	10 (338)	9.0 (7.5)
	<i>ULS za kvantifikaciju</i>	17 (557)	15.3 (12.3)
	<i>ULS za opis</i>	25 (784)	22.5 (17.3)
<i>ULS orijentisani ka tekstu</i>	<i>Tranzicioni signali</i>	14 (790)	12.6 (17.4)
	<i>Rezultativni signali</i>	8 (382)	7.2 (8.4)
	<i>Strukturni signali</i>	/	/
	<i>Ograničavajući signali</i>	25 (1154)	22.5 (25.4)
<i>ULS orijentisani ka učesniku</i>	<i>ULS stava</i>	3 (221)	2.7 (4.9)
	<i>ULS uključivanja</i>	5 (165)	4.5 (3.6)
Total		111 (4536)	100%



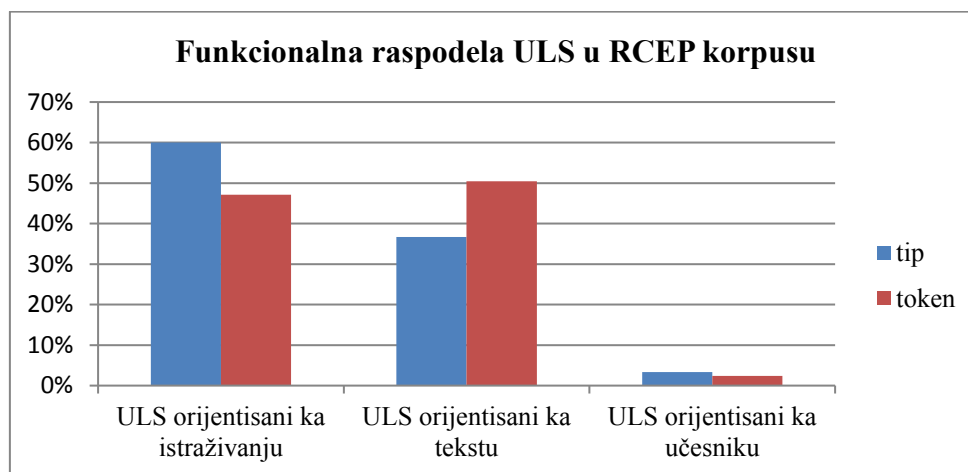
Grafikon 36. Funkcionalna raspodela ULS u SCEP korpusu na podgrupe

4.4.3. Funkcionalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski

Funkcionalna raspodela ULS kod izvornih govornika rumunskog jezika pri njihovoj produkciji na engleskom jeziku pokazuje da je najzastupljenija kategorija kada posmatramo *tipove ULS orijentisani ka istraživanju* a da je najzastupljenija kategorija u odnosu na tokene *ULS orijentisani ka tekstu* (v. Tabelu 57 i Grafikon 37). Navedeni podaci nam ukazuju na to da izvorni govornici rumunskog jezika koriste najveću raznovrsnost *ULS orijentisanih ka istraživanju*, ali da se ipak *ULS orijentisani ka tekstu* najčešće koriste iako sadrže manji broj tipova. *ULS orijentisani ka učesniku* su najmanje zastupljeni i što se tiče tipova i što se tiče tokena.

Tabela 57. Funkcionalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski

Funkcionalna kategorija	Tipovi (tokeni)	Procenti
<i>ULS orijentisani ka istraživanju</i>	36 (1111)	60.0 (47.1)
<i>ULS orijentisani ka tekstu</i>	22 (1190)	36.7 (50.5)
<i>ULS orijentisani ka učesniku</i>	2 (57)	3.3 (2.4)

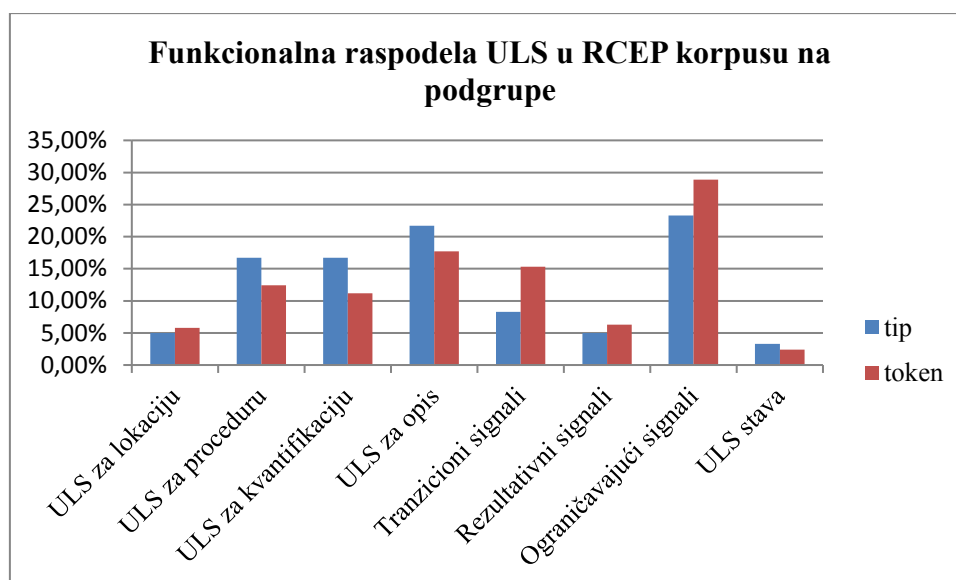


Grafikon 37. Funkcionalna raspodela ULS u RCEP korpusu

Funkcionalnom raspodelom ULS izvornih govornika rumunskog jezika na podgrupe uviđamo da je najzastupljenija podgrupa *ograničavajući signali*, dok je najmanje zastupljena kategorija *ULS stava* (v. Tabelu 58 i Grafikon 38). Poređenjem procentualne zastupljenosti tipova i tokena u funkcionalnim podgrupama ULS iz RCEP korpusa uviđamo da su tokeni više zastupljeni u četiri podgrupe (*ograničavajući signali*, *tranzicioni signali*, *rezultativni signali* i *ULS za lokaciju*), kao i tipovi (*ULS za opis*, *ULS za kvantifikaciju*, *ULS za proceduru* i *ULS stava*).

Tabela 58. Funkcionalna raspodela ULS u RCEP korpusu na podgrupe

Grupa	Podgrupa	Tip (token)	Procenti
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	3 (137)	5.0 (5.8)
	<i>ULS za proceduru</i>	10 (293)	16.7 (12.4)
	<i>ULS za kvantifikaciju</i>	10 (263)	16.7 (11.2)
	<i>ULS za opis</i>	13 (418)	21.7 (17.1)
<i>ULS orijentisani ka tekstu</i>	<i>Tranzicioni signali</i>	5 (361)	8.3 (15.3)
	<i>Rezultativni signali</i>	3 (148)	5.0 (6.3)
	<i>Strukturni signali</i>	/	/
	<i>Ograničavajući signali</i>	14 (681)	23.3 (28.9)
<i>ULS orijentisani ka učesniku</i>	<i>ULS stava</i>	2 (57)	3.3 (2.4)
	<i>ULS uključivanja</i>	/	/
Total		60 (2358)	100%



Grafikon 38. Funkcionalna raspodela ULS u RCEP korpusu na podgrupe

4.4.4. Poređenje funkcionalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik srpski

Kako bi se ispitale potencijalne razlike, tj. asocijacije, između ECEP i SCEP korpusa u odnosu na funkcionalne distribucije ULS, korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između ECEP i SCEP korpusa u odnosu na funkcionalne grupe – tipovi

Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih grupa u odnosu na tipove prikazana je u Tabeli 59. Hi-kvadrat test nije bio značajan, $\chi^2(2, N = 138) = 0.45, p > .050$, što ukazuje na to da nije identifikovana statistički značajna povezanost između korpusa i funkcionalnih grupa. Drugim rečima, razlike koje uviđamo u funkcionalnoj raspodeli ULS u ECEP i SCEP korpusima na grupe nisu statistički značajne.

Tabela 59. Unakrsna tabela između korpusa i funkcionalnih grupa u odnosu na tipove

Funkcionalne grupe	SCEP			ECEP		
	tip	%	R	tip	%	R
ULS orijentisani ka istraživanju	56	50.5%	0.2	13	48.1%	-0.2
ULS orijentisani ka tekstu	47	42.3%	0.2	11	40.7%	-0.2
ULS orijentisani ka učesniku	8	7.2%	-0.7	3	11.1%	0.7
Total	111	100%		27	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između ECEP i SCEP korpusa u odnosu na funkcionalne grupe– tokeni

Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih grupa u odnosu na tokene prikazana je u Tabeli 60. Hi-kvadrat test je bio značajan, $\chi^2(2, N = 5313) = 26.59, p < .001$, odnosno postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

Korigovani standardizovani reziduali ukazuju na to da je bilo više *ULS orijentisanih ka istraživanju* i *ULS orijentisanih ka učesniku* u ECEP korpusu, dok je bilo više *ULS orijentisanih ka tekstu* u SCEP korpusu. Na osnovu navedenih podataka, možemo zaključiti da izvorni govornici srpskog jezika imaju tendenciju da isuviše učestalo koriste ULS iz funkcionalne grupe *ULS orijentisani ka tekstu*, dok nedovoljno učestalo koriste ULS iz kategorija *ULS orijentisani ka istraživanju* i *ULS orijentisani ka učesniku* a srazmerno raspodelama u ECEP korpusu.

Tabela 60. *Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih grupa u odnosu na tokene*

Funkcionalne grupe	SCEP			ECEP		
	token	%	R	token	%	R
<i>ULS orijentisani ka istraživanju</i>	1824	40.2%	-2.1*	344	44.3%	2.1*
<i>ULS orijentisani ka tekstu</i>	2326	51.3%	4.4*	332	42.7%	-4.4*
<i>ULS orijentisani ka učesniku</i>	386	8.5%	-4.0*	101	13.0%	4.0*
Total	4536	100%		777	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

S obzirom na podatak da statistički značajna razlika u funkcionalnoj raspodeli ULS u grupe u ECEP i SCEP korpusima nije identifikovana za tipove a jeste identifikovana za tokene, možemo zaključiti da se izvorni govornici engleskog i izvorni govornici srpskog ne razlikuju značajno po funkcionalnoj raspodeli identifikovanih ULS, ali se razlikuju u tome koliko ih često koriste.

Razlike između ECEP i SCEP korpusa u odnosu na funkcionalne podgrupe – tipove

Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih podgrupa u odnosu na tipove prikazana je u Tabeli 61. Hi-kvadrat test nije bio značajan, $\chi^2(9, N = 138) = 10.90, p > .050$, odnosno nije identifikovana statistički značajna povezanost između korpusa i funkcionalnih podgrupa, tj. razlike koje uvidamo u funkcionalnoj raspodeli na podgrupe tipova ULS u ECEP i SCEP korpusu nisu značajne.

Tabela 61. *Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih podgrupa u odnosu na tipove*

Grupa	Podgrupa	SCEP			ECEP		
		tip	%	R	tip	%	R
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	4	3.6%	-0.9	2	7.4%	0.9
	<i>ULS za proceduru</i>	10	9.0%	-0.3	3	11.1%	0.3
	<i>ULS za kvantifikaciju</i>	17	15.3%	1.1	2	7.4%	-1.1
<i>ULS orijentisani ka tekstu</i>	<i>ULS za opis</i>	25	22.5%	0.0	6	22.2%	0.0
	<i>Tranzicioni signali</i>	14	12.6%	0.2	3	11.1%	-0.2
	<i>Rezultativni signali</i>	8	7.2%	0.0	2	7.4%	0.0
	<i>Strukturni signali</i>	0	0%	-2.0	1	3.7%	2.0
	<i>Ograničavajući signali</i>	25	22.5%	0.5	5	18.5%	-0.5

Tabela 61. Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih podgrupa u odnosu na tipove

<i>ULS orijentisani ka učesniku</i>	<i>ULS stava</i>	3	2.7%	-1.9	3	11.1%	1.9
	<i>ULS uključivanja</i>	5	4.5%	1.1	0	0%	-1.1
Total		111	100%		27	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između ECEP i SCEP korpusa u odnosu na funkcionalne podgrupe – tokeni

Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih podgrupa u odnosu na tokene prikazana je u Tabeli 62. Hi-kvadrat test je bio značajan, $\chi^2(9, N = 5313) = 351.02, p < .001$, što pokazuje da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

Tabela 62. Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih podgrupa u odnosu na tokene

Grupa	Podgrupa	SCEP			ECEP		
		tokeni	%	R	tokeni	%	R
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	145	3.2%	-7.0*	66	8.5%	7.0*
	<i>ULS za proceduru</i>	338	7.5%	-0.8	64	8.2%	0.8
	<i>ULS za kvantifikaciju</i>	557	12.3%	2.2*	74	9.5%	-2.2*
	<i>ULS za opis</i>	784	17.3%	-0.5	140	18.0%	0.5
<i>ULS orijentisani ka tekstu</i>	<i>Tranzicioni signali</i>	790	17.4%	2.7*	105	13.5%	-2.7*
	<i>Rezultativni signali</i>	382	8.4%	0.5	61	7.9%	-0.5
	<i>Strukturni signali</i>	0	0%	-13.3*	30	3.9%	13.3*
	<i>Ograničavajući signali</i>	1154	25.4%	4.8*	136	17.5%	-4.8*
<i>ULS orijentisani ka učesniku</i>	<i>ULS stava</i>	221	4.9%	-8.8*	101	13.0%	8.8*
	<i>ULS uključivanja</i>	165	3.6%	5.4*	0	0%	-5.4*
Total		4536	100%		777	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Korigovani standardizovani reziduali su bili značajno povišeni za SCEP korpus u slučaju *ULS za kvantifikaciju, ULS za uključivanje, tranzicionih i ograničavajućih signala*, ukazujući na obrazac više upotrebe ovih podgrupa u SCEP korpusu u odnosu na ECEP. Za ECEP korpus, značajno je veća upotreba *ULS za lokaciju, ULS stava i strukturnih signala*. Stoga, možemo zaključiti da izvorni govornici srpskog jezika srazmerno nedovoljno učestalo upotrebljavaju ULS iz funkcionalnih podgrupa *ULS za lokaciju, ULS stava i strukturnih signala*.

Činjenica da se statistički značajna razlika nije identifikovala u funkcionalnoj raspodeli na podgrupe tipova ULS u ECEP i SCEP korpusima, a jeste na podgrupe tokena, ukazuje na to da se zastupljenost ULS u funkcionalnim podgrupama kod izvornih govornika engleskog jezika i izvornih govornika srpskog jezika ne razlikuju ali da se razlikuje učestalost njihove upotrebe.

4.4.5. Poređenje funkcionalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik rumunski

Kako bi se ispitale potencijalne razlike, tj. asocijacije, između ECEP i RCEP korpusa u odnosu na funkcionalne distribucije ULS korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između ECEP i RCEP korpusa u odnosu na funkcionalne grupe – tipovi

Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih grupa u odnosu na tipove prikazana je u Tabeli 63. Hi-kvadrat test nije bio značajan, $\chi^2(2, N = 87) = 0.28, p > .050$, odnosno nije identifikovana statistički značajna povezanost između korpusa i funkcionalnih grupa. Drugačije rečeno, razlike koje uviđamo u raspodeli tipova ULS u funkcionalnim grupama između ECEP i RCEP nisu statistički značajne.

Tabela 63. Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih grupa u odnosu na tipove

Funkcionalne grupe	RCEP			ECEP		
	tip	%	R	tip	%	R
<i>ULS orijentisani ka istraživanju</i>	36	60.0%	1.0	13	48.1%	-1.0
<i>ULS orijentisani ka tekstu</i>	22	36.7%	-0.4	11	40.7%	0.4
<i>ULS orijentisani ka učesniku</i>	2	3.3%	-1.4	3	11.1%	1.4
Total	60	100%		27	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

* – statistički značajni rezultati

Razlike između ECEP i RCEP korpusa u odnosu na funkcionalne grupe– tokeni

Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih grupa u odnosu na tokene prikazana je u Tabeli 64. Hi-kvadrat test je bio značajan, $\chi^2(2, N = 2358) = 138.06, p < .001$, što ukazuje na to da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

Korigovani standardizovani reziduali pokazuju da je upotrebljeno više *ULS orijentisanih ka tekstu* u RCEP korpusu, dok više *ULS orijentisanih ka učesniku* upotrebljeno u ECEP korpusu. Izvorni govornici engleskog jezika i izvorni govornici rumunskog jezika ne razlikuju se značajno u učestalosti upotrebe ULS iz funkcionalne kategorije *ULS orijentisani ka istraživanju*. Izvorni govornici rumunskog jezika imaju tendenciju da nedovoljno učestalo koriste *ULS orijentisane ka učesniku*, dok isuviše učestalo koriste *ULS orijentisane ka tekstu*.

Tabela 64. Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih grupa u odnosu na tokene

Funkcionalne grupe	RCEP			ECEP		
	token	%	R	token	%	R
<i>ULS orijentisani ka istraživanju</i>	1111	47.1%	1.4	344	44.3%	-1.4
<i>ULS orijentisani ka tekstu</i>	1190	50.5%	3.7*	332	42.7%	-3.7*
<i>ULS orijentisani ka učesniku</i>	57	2.4%	-11.7*	101	13.0%	11.7*
Total	2358	100%		777	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Poređenjem funkcionalne raspodele tipova i tokena ULS u ECEP i RCEP korpusu uviđamo da se statistički značajne razlike identifikovale samo za tokene, što nas dalje dovodi do zaključka da se izvorni govornici engleskog i rumunskog ne razlikuju u raspodeli ULS u funkcionalne grupe ali da se razlikuju u učestalosti upotrebe tih ULS.

Razlike između ECEP i RCEP korpusa u odnosu na funkcionalne podgrupe – tipovi

Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih podgrupa u odnosu na tipove prikazana je u Tabeli 65. Hi-kvadrat test nije bio značajan, $\chi^2 (8, N = 87) = 6.45, p > .050$, odnosno nije identifikovana statistički značajna povezanost između korpusa i funkcionalnih podgrupa, Drugim rečima, razlike koje uviđamo u funkcionalnoj raspodeli tipova na podgrupe u ECEP i RCEP korpusima nije statistički značajno.

Tabela 65. Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih podgrupa u odnosu na tipove

Grupa	Podgrupa	Rumunski			Engleski		
		tip	%	R	tip	%	R
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	3	5.0%	-0.4	2	7.4%	0.4
	<i>ULS za proceduru</i>	10	16.7%	0.7	3	11.1%	-0.7
	<i>ULS za kvantifikaciju</i>	10	16.7%	1.2	2	7.4%	-1.2
	<i>ULS za opis</i>	13	21.7%	-0.1	6	22.2%	0.1
<i>ULS orijentisani ka tekstu</i>	<i>Tranzicioni signali</i>	5	8.3%	-0.4	3	11.1%	0.4
	<i>Rezultativni signali</i>	3	5.0%	-0.4	2	7.4%	0.4
	<i>Strukturni signali</i>	0	0%	-1.5	1	3.7%	1.5
	<i>Ograničavajući signali</i>	14	23.3%	0.5	5	18.5%	-0.5
<i>ULS orijentisani ka učesniku</i>	<i>ULS stava</i>	2	3.3%	-1.4	3	11.1%	1.4
	<i>ULS uključivanja</i>	/	/	/	/	/	/
Total		60	100%		27	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između ECEP i RCEP korpusa u odnosu na funkcionalne podgrupe – tokeni

Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih podgrupa u odnosu na tokene prikazana je u Tabeli 66. Hi-kvadrat test je bio značajan, $\chi^2 (8, N = 3135) = 270.33, p < .001$, ukazujući na to da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

RCEP korpus imao je značajno više korigovane standardizovane rezidualne za podgrupe *ULS za proceduru* i *ograničavajuće signale*, dok je ECEP korpus imao više rezidualne za podgrupe *ULS za lokaciju*, *strukturni signali* i *ULS stava*. Dakle, izvorni govornici rumunskog jezika imaju tendenciju da koriste funkcionalne podgrupe *ULS za proceduru* i *ograničavajuće signale* srazmerno isuviše učestalo, a funkcionalne podgrupe nedovoljno učestalo *ULS za lokaciju*, *strukturni signali* i *ULS stava*. Na osnovu navedenih podataka, možemo zaključiti da bi u radu sa izvornim govornicima rumunskog jezika posebnu pažnju trebalo posvetiti funkcionalnim podgrupama *ULS za lokaciju*, *strukturni signali* i *ULS stava*. Pored toga, potrebno je dodatno obraditi alternativne načine za prenošenje značenja za koja se koriste ULS iz funkcionalnih podgrupa *ULS za proceduru* i *ograničavajuće signale*.

Tabela 66. Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih podgrupa u odnosu na tokene

Grupa	Podgrupa	RCEP			ECEP		
		token	%	R	token	%	R
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	137	5.8%	-2.6*	66	8.5%	2.6*
	<i>ULS za proceduru</i>	293	12.4%	3.2*	64	8.2%	-3.2*
	<i>ULS za kvantifikaciju</i>	263	11.2%	1.3	74	9.5%	-1.3
	<i>ULS za opis</i>	418	17.7%	-0.2	140	18.0%	0.2
<i>ULS orijentisani ka tekstu</i>	<i>Tranzicioni signali</i>	361	15.3%	1.2	105	13.5%	-1.2
	<i>Rezultativni signali</i>	148	6.3%	-1.5	61	7.9%	1.5
	<i>Strukturni signali</i>	0	0%	-9.6*	30	3.9%	9.6*
	<i>Ograničavajući signali</i>	681	28.9%	6.3*	136	17.5%	-6.3*
<i>ULS orijentisani ka učesniku</i>	<i>ULS stava</i>	57	2.4%	-11.7*	101	13.0%	11.7*
	<i>ULS uključivanja</i>	/	/	/	/	/	/
Total		2358	100%		777	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

S obzirom na to da se statistički značajne razlike nisu ustanovile za raspodelu tipova a jesu za raspodelu tokena u funkcionalnim podgrupama u ECEP i SCEP korpusima, možemo zaključiti da se izvorni govornici engleskog i rumunskog jezika ne razlikuju značajno po raspodelu ULS u funkcionalne podgrupe, ali da se značajno razlikuju po tome koliko ih često koriste unutar tih podgrupa.

4.4.6. Poređenje funkcionalne raspodele ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski i rumunski

Kako bi se ispitale potencijalne razlike, tj. asocijacije, između SCEP i RCEP korpusa u odnosu na funkcionalne distribucije ULS korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između SCEP i RCEP korpusa u odnosu na funkcionalne grupe – tipovi

Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih grupa u odnosu na tipove prikazana je u Tabeli 67. Hi-kvadrat test nije bio značajan, $\chi^2(2, N=171) = 1.97, p > .050$, što ukazuje na to da nije identifikovana statistički značajna povezanost između korpusa i funkcionalnih grupa, odnosno uočene razlike u raspodeli tipova ULS u funkcionalnim grupama u korpusima SCEP i RCEP nisu statistički značajne.

Tabela 67. Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih grupa u odnosu na tipove

Funkcionalne grupe	SCEP			RCEP		
	tip	%	R	tip	%	R
<i>ULS orijentisani ka istraživanju</i>	56	50.5%	-1.2	36	60.0%	1.2
<i>ULS orijentisani ka tekstu</i>	47	42.3%	0.7	22	36.7%	-0.7
<i>ULS orijentisani ka učesniku</i>	8	7.2%	1.0	2	3.3%	-1.0
Total	111	100%		60	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između SCEP i RCEP korpusa u odnosu na funkcionalne grupe – tokeni

Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih grupa u odnosu na tokene prikazana je u Tabeli 68.

Tabela 68. Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih grupa u odnosu na tokene

Funkcionalne grupe	SCEP			RCEP		
	token	%	R	token	%	R
<i>ULS orijentisani ka istraživanju</i>	1824	40.2%	-5.5*	1111	47.1%	5.5*
<i>ULS orijentisani ka tekstu</i>	2326	51.3%	0.6	1190	50.5%	-0.6
<i>ULS orijentisani ka učesniku</i>	386	8.5%	9.8*	57	2.4%	-9.8*
Total	4536	100%		2358	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Hi-kvadrat test je bio značajan, $\chi^2(2, N = 6894) = 107.19, p < .001$, ukazujući na to da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

Korigovani standardizovani reziduali pokazuju da su učestalije upotrebljeni *ULS orijentisanih ka istraživanju* u RCEP korpusu, dok su učestalije upotrebljeni *ULS orijentisani ka učesniku* u SCEP

korpusu. Drugim rečima, u dve od tri funkcionalne grupe identifikovane su statistički značajne razlike u produkciji izvornih govornika srpskog i rumunskog jezika kada je reč o učestalosti upotrebe ULS.

Razlike između SCEP i RCEP korpusa u odnosu na funkcionalne podgrupe – tipovi

Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih podgrupa u odnosu na tipove prikazana je u Tabeli 69. Hi-kvadrat test nije bio značajan, $\chi^2(8, N = 171) = 5.90$, $p > .050$, odnosno nije identifikovana statistički značajna povezanost između korpusa i funkcionalnih podgrupa.

Tabela 69. Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih podgrupa u odnosu na tipove

Grupa	Podgrupa	SCEP			RCEP		
		tip	%	R	tip	%	R
ULS orijentisani ka istraživanju	ULS za lokaciju	4	3.6%	-0.4	3	5.0%	0.4
	ULS za proceduru	10	9.0%	-1.5	10	16.7%	1.5
	ULS za kvantifikaciju	17	15.3%	-0.2	10	16.7%	0.2
ULS orijentisani ka tekstu	ULS za opis	25	22.5%	0.1	13	21.7%	-0.1
	Tranzicioni signali	14	12.6%	0.8	5	8.3%	-0.8
	Rezultativni signali	8	7.2%	0.6	3	5.0%	-0.6
ULS orijentisani ka učesniku	Strukturni signali	/	/	/	/	/	/
	Ograničavajući signali	25	22.5%	-0.1	14	23.3%	0.1
	ULS stava	3	2.7%	-0.2	2	3.3%	0.2
	ULS uključivanja	5	4.5%	1.7	0	0%	-1.7
Total		111	100%		60	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između SCEP i RCEP korpusa u odnosu na funkcionalne podgrupe – tokeni

Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih podgrupa u odnosu na tokene prikazana je u Tabeli 70. Hi-kvadrat test je bio značajan, $\chi^2(8, N = 6894) = 198.94$, $p < .001$, odnosno postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

Korigovani standardizovani reziduali su bili značajno povišeni za SCEP korpus u slučaju *tranzicionih i rezultativnih signala*, kao i *ULS stava i uključivanja*, ukazujući na obrazac više upotrebe ULS u ovim podgrupama u SCEP korpusu u odnosu na RCEP. Sa druge strane, značajno veća je upotreba *ULS za lokaciju, proceduru i ograničavajućih signala* u RCEP korpusu.

Činjenica da statistički značajne razlike nisu identifikovane za raspodelu tipova a jesu za raspodelu tokena u funkcionalne podgrupe u SCEP i RCEP korpusima, pokazuju da se izvorni govornici srpskog i rumunskog ne razlikuju značajno na osnovu raspodele ULS u funkcionalne podgrupe, ali da se razlikuju u tome koliko se često ti ULS koriste u navedenim podgrupama.

Tabela 70. Unakrsna tabela između SCEP RCEP korpusa i funkcionalnih podgrupa u odnosu na tokene

Grupa	Podgrupa	SCEP			RCEP		
		token	%	R	token	%	R
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	145	3.2%	-5.2*	137	5.8%	5.2*
		338	7.5%	-6.8*	293	12.4%	6.8*
	<i>ULS za proceduru</i>						
	<i>ULS za kvantifikaciju</i>	557	12.3%	1.4	263	11.2%	-1.4
<i>ULS orijentisani ka tekstu</i>	<i>ULS za opis</i>	784	17.3%	-0.5	418	17.7%	0.5
	<i>Tranzicioni signali</i>	790	17.4%	2.2*	361	15.3%	-2.2*
	<i>Rezultativni signali</i>	382	8.4%	3.2*	148	6.3%	-3.2*
	<i>Strukturni signali</i>	/	/	/	/	/	/
<i>ULS orijentisani ka učesniku</i>	<i>Ograničavajući signali</i>	1154	25.4%	-3.1*	681	28.9%	3.1*
	<i>ULS stava</i>	221	4.9%	4.9*	57	2.4%	-4.9*
	<i>ULS uključivanja</i>	165	3.6%	9.4*	0	0%	-9.4*
Total		4536	100%		2358	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

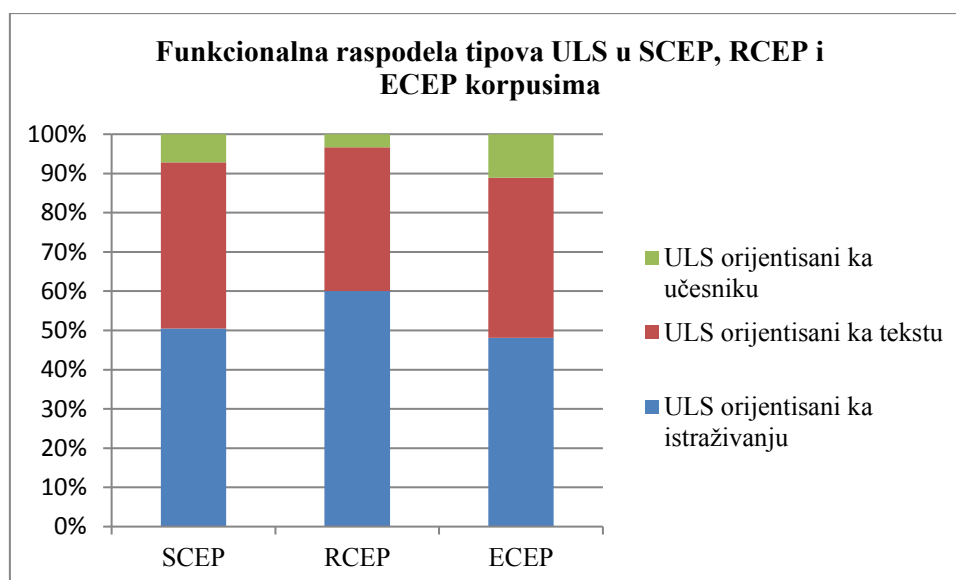
4.4.7. Poređenje funkcionalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski

Poređenjem funkcionalne raspodele na tipove u sva tri korpusa uviđamo da su *ULS orijentisani ka istraživanju* najviše zastupljeni kod izvornih govornika rumunskog jezika, *ULS orijentisani ka tekstu* najviše su zastupljeni kod izvornih govornika srpskog jezika, a *ULS orijentisani ka učesniku* najviše su zastupljeni kod izvornih govornika engleskog jezika (v. Tabelu 71 i Grafikon 39).

Tabela 71. Funkcionalna raspodela ULS kod izvornih govornika srpskog, rumunskog i engleskog jezika – tipovi

Funkcionalna kategorija	SCEP	RCEP	ECEP
<i>ULS orijentisani ka istraživanju</i>	50.45%	60%	48.15%
<i>ULS orijentisani ka tekstu</i>	42.34%	36.67%	40.74%
<i>ULS orijentisani ka učesniku</i>	7.21%	3.33%	11.11%

Statistička analiza funkcionalne raspodele tipova ULS na grupe i na podgrupe ni na jednoj relaciji (ECEP–SCEP, ECEP–RCEP, SCEP–RCEP) nije pokazala značajne rezultate. To bi dalje značilo da se izvorni govornici engleskog jezika i neizvorni govornici engleskog jezika kojima je maternji jezik srpski i rumunski ne razlikuju značajno po raspodeli ULS u funkcionalne grupe i podgrupe. Ovde bi, takođe, trebalo da imamo na umu i činjenicu da statistički testovi, kao što je ovde primenjeni hi-kvadrat test, imaju tendenciju da lakše identifikuju statistički značajne razlike na većem uzorku, odnosno kada se u njih unesu veće brojke. To bi delom moglo da objasni činjenicu da se statistički značajne razlike nisu identifikovale za funkcionalnu raspodelu tipova na grupe a jesu na tokene iako, procentualno gledano, deskriptivnom analizom vidimo da su veće razlike u funkcionalnoj raspodeli tipova (up. Tabelu 71: *funkcionalna raspodela – tipovi* i Tabelu 72: *funkcionalna raspodela – tokeni*).



Grafikon 39. Funkcionalna raspodela tipova ULS u SCEP, RCEP i ECEP korpusima

Poređenjem funkcionalne raspodele tokena ULS u sva tri korpusa uviđamo da ne prati raspodelu koja se uočava za tipove ULS. Naime, u raspodeli tokena izvorni govornici srpskog i rumunskog jezika najviše koriste *ULS orijentisane ka tekstu*, dok izvorni govornici engleskog jezika i za tipove i za tokene najviše koriste *ULS orijentisane ka istraživanju* (v. Tabelu 72 i Grafikon 40). Na osnovu navedenih podataka o funkcionalnoj raspodeli tipova i tokena ULS u tri ispitana korpusa, možemo da zaključimo da je potvrđena treća radna hipoteza koja glasi: *Funkcionalna raspodela ULS kod izvornih govornika engleskog jezika i dve grupe neizvornih govornika engleskog jezika biće okarakterisana učestalijom upotrebom ULS iz funkcionalnih grupa ULS orijentisani ka istraživanju i ULS orijentisani ka tekstu nego ULS orijentisani ka učesniku.*

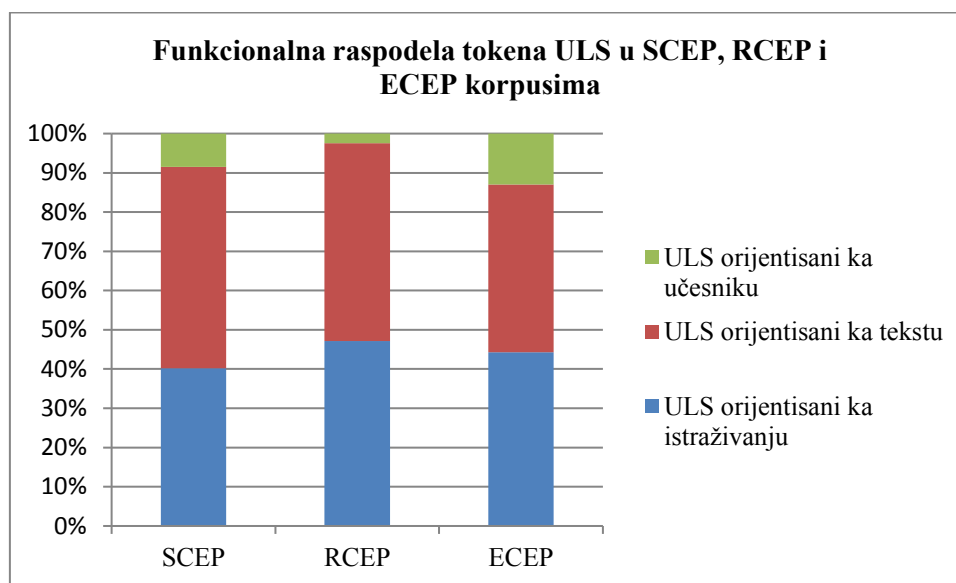
Tabela 72. Funkcionalna raspodela ULS kod izvornih govornika srpskog, rumunskog i engleskog jezika – tokeni

Funkcionalna kategorija	SCEP	RCEP	ECEP
<i>ULS orijentisani ka istraživanju</i>	40.21%	47.12%	44.27%
<i>ULS orijentisani ka tekstu</i>	51.28%	50.47%	42.74%
<i>ULS orijentisani ka učesniku</i>	8.51%	2.42%	13%

Statistička raspodela tokena ULS na funkcionalne grupe i podgrupe pokazala je statistički značajne razlike na svim relacijama u tri korpusa. Možemo zaključiti da se malo veća varijacija na funkcionalne grupe i podgrupe identifikovala između ECEP i SCEP korpusa nego između ECEP i RCEP korpusa, a s obzirom na činjenicu da se na relaciji ECEP–SCEP korpusa identifikovala statistički značajna razlika za sve tri funkcionalne grupe i sedam od deset zastupljenih funkcionalnih podgrupa, dok se na relaciji ECEP–RCEP korpusa identifikovala statistički značajna razlika u dve od tri funkcionalne grupe i u pet od devet zastupljenih podgrupa. Poređenjem funkcionalne raspodele ULS na tokene na relacijama ECEP–SCEP i ECEP–RCEP uviđamo da izvorni govornici engleskog jezika statistički značajno više koriste *ULS za lokaciju*, *ULS stava* i *strukturalne signale* i to uporedno sa obe grupe neizvornih govornika. Štaviše, nijedan tip, pa time ni token, u funkcionalnoj kategoriji *strukturalni signali*, nije identifikovan u korpusima neizvornih govornika engleskog jezika. Iako po ovim podacima vidimo da neizvorni govornici kojima je maternji jezik srpski i rumunski imaju tendenciju da nedovoljno učestalo koriste ULS iz istih funkcionalnih podgrupa, kada je reč o prekomernoj upotrebi ULS u funkcionalnim podgrupama,

mного se više razlikuju. Naime, samo je funkcionalna potkategorija *ograničavajući signali* povišena i kod izvornih govornika srpskom i rumunskog uporedno sa izvornim govornicima engleskog jezika. Stoga, pri podučavanju studenata kojima je maternji jezik srpski i kojima je maternji jezik rumunski jezik, važno je dobro podučiti i podsticati na češću upotrebu ULS iz funkcionalnih podgrupa *ULS za lokaciju*, *ULS stava* i *strukturalni signali*. Takođe, trebalo bi obraditi alternativne načine prenošenja značenja za koje koriste ULS is funkcionalne podgrupe *ograničavajući signali* kako ih ne bi isuviše učestalo koristili.

Na osnovu rezultata dobijenih statističkom analizom razlika u strukturalnoj i funkcionalnoj raspodeli ULS u ECEP, SCEP i RCEP korpusima možemo zaključiti da je potvrđena šesta radna hipoteza koja glasi: *Postojeće statistički značajne razlike u strukturalnoj i funkcionalnoj raspodeli ULS na grupe i podgrupe između izvornih i neizvornih govornika engleskog jezika, kao i između dve grupe neizvornih govornika.*



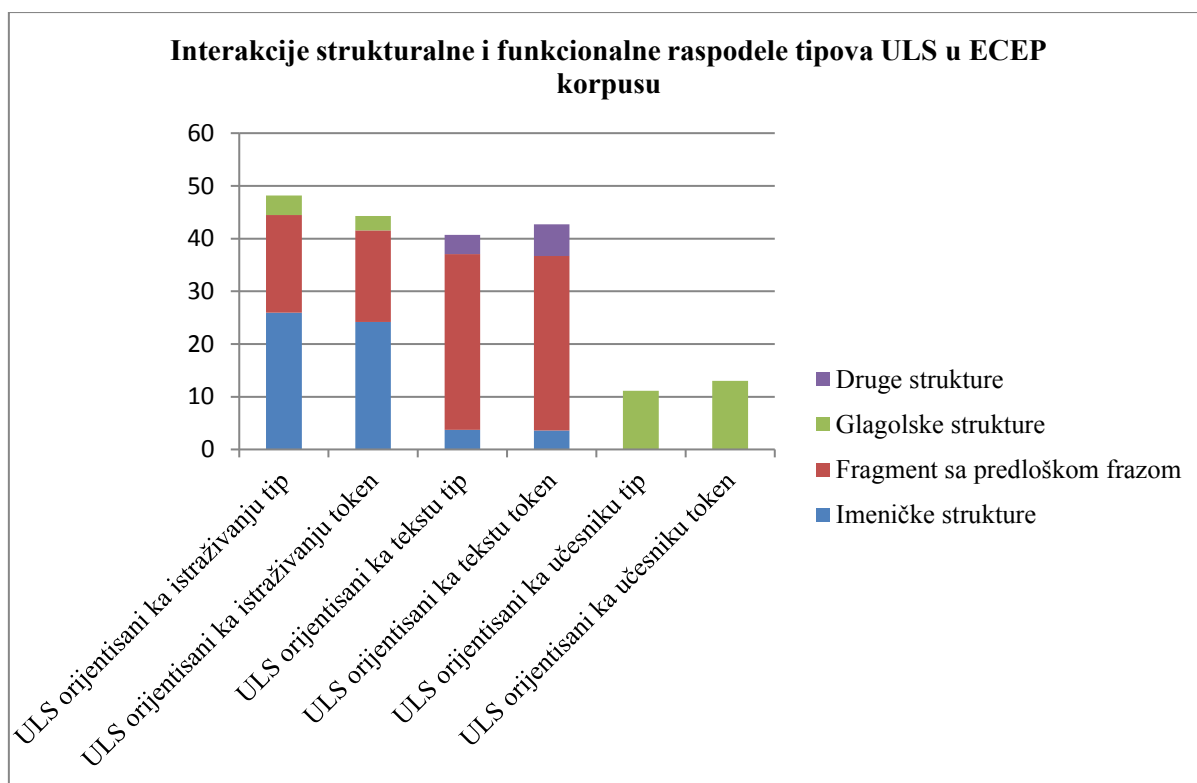
Grafikon 40. *Funkcionalna raspodela tokena ULS u SCEP, RCEP i ECEP korpusima*

4.5. Interakcije strukturalne i funkcionalne raspodele ULS u ECEP korpusu

ULS u funkcionalnim kategorijama smo raspodelili u strukturalne kategorije kako bismo mogli da sagledamo interakciju strukturalnih i funkcionalnih kategorija kod izvornih govornika engleskog jezika. Rezultati pokazuju da izvorni govornici engleskog jezika u funkcionalnoj kategoriji *ULS orijentisani ka istraživanju* koriste tri strukturalne grupe. Od njih najzastupljenije su *imeničke strukture*, praćene *fragmentima sa predloškom frazom*, dok su *glagolske strukture* najmanje zastupljene. U funkcionalnoj grupi *ULS orijentisani ka tekstu* isto su zastupljene tri strukturalne grupe. Međutim, ovde su najzastupljeniji *fragmenti sa predloškom frazom*. *Imeničke strukture* i *druge strukture* su jednako zastupljene što se tiče tipova, dok su *druge strukture* zastupljenije kada je reč o tokenima. U funkcionalnoj grupi *ULS orijentisani ka učesniku* prisutna je samo jedna strukturalna grupa – *glagolske strukture*. Navedeni podaci pokazuju da izvorni govornici engleskog jezika preferiraju da koriste različite strukturalne grupe u različitim funkcionalnim grupama.

Tabela 73. Interakcije strukturalne i funkcionalne raspodele ULS u ECEP korpusu

		Imeničke strukture	Fragment sa predložkom frazom	Glagolske strukture	Druge strukture
ULS orijentisani ka istraživanju	Tip (Token)	7 (188)	5 (135)	1 (21)	/
	Procenti	25.93 (24.20)	18.52 (17.37)	3.70 (2.70)	/
ULS orijentisani ka tekstu	Tip (Token)	1 (28)	9 (257)	/	1 (47)
	Procenti	3.70 (3.60)	33.33 (33.08)	/	3.70 (6.05)
ULS orijentisani ka učesniku	Tip (Token)	/	/	3 (101)	/
	Procenti	/	/	11.11 (13)	/



Grafikon 41. Interakcije strukturalne i funkcionalne raspodele ULS u ECEP korpusu

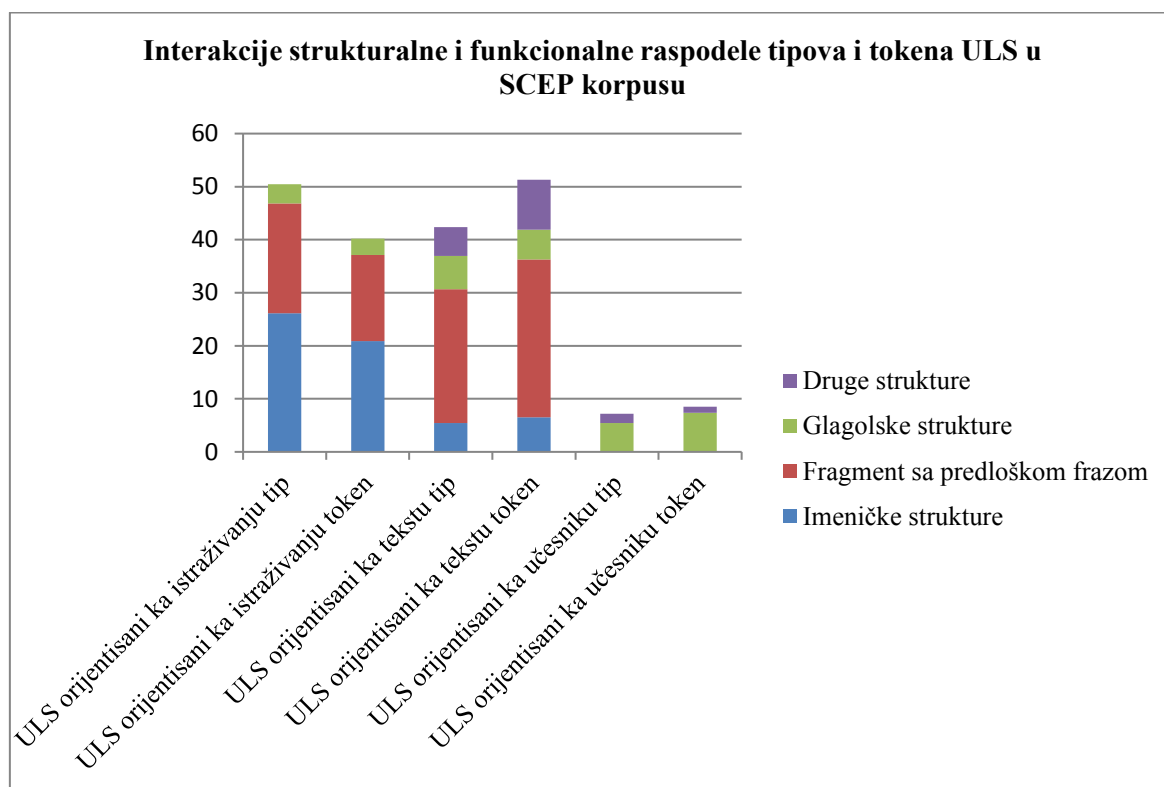
4.6. Interakcije strukturalne i funkcionalne raspodele ULS u SCEP korpusu

ULS u funkcionalnim grupama raspodelili smo u strukturalne grupe kako bismo mogli da sagledamo interakciju strukturalnih i funkcionalnih kategorija kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski jezik. Dobijeni podaci pokazuju da izvorni govornici srpskog jezika u funkcionalnoj grupi *ULS orijentisani ka istraživanju* koriste tri strukturalne grupe (v Tabelu 74 i Grafikon 42).

Tabela 74. Interakcije strukturalne i funkcionalne raspodele ULS u SCEP korpusu

		<i>Imeničke strukture</i>	<i>Fragment sa predloškom frazom</i>	<i>Glagolske strukture</i>	<i>Druge strukture</i>
<i>ULS orijentisani ka istraživanju</i>	Tip (Token)	29 (948)	23 (735)	4 (141)	/
	Procenti	26.13 (20.90)	20.72(16.20)	3.60 (3.11)	/
<i>ULS orijentisani ka tekstu</i>	Tip (Token)	6 (296)	28 (1350)	7 (254)	6 (426)
	Procenti	5.41 (6.53)	25.23 (29.76)	6.31 (5.60)	5.41 (9.39)
<i>ULS orijentisani ka učesniku</i>	Tip (Token)	/	/	6 (333)	2 (53)
	Procenti	/	/	5.41 (7.34)	1.80 (1.17)

Najzastupljenije su *imeničke strukture* praćene *fragmentima sa predloškom frazom*, dok su *glagolske strukture* najmanje zastupljene. U funkcionalnoj grupi *ULS orijentisani ka tekstu* uočavaju se sve četiri strukturalne grupe. Najzastupljeniji su *fragmenti sa predloškom frazom*. Druga najzastupljenija kategorija je drugačija za tipove i tokene, naime, *glagolske strukture* su druga najzastupljenija kategorija kada gledamo tipove, dok su *druge strukture* druga najzastupljenija kategorija kada gledamo tokene. U funkcionalnoj grupi *ULS orijentisani ka učesniku* prisutne su dve strukturalne kategorije. I za tipove i za tokene najzastupljenija grupa su *glagolske strukture* praćene *drugim strukturama*.



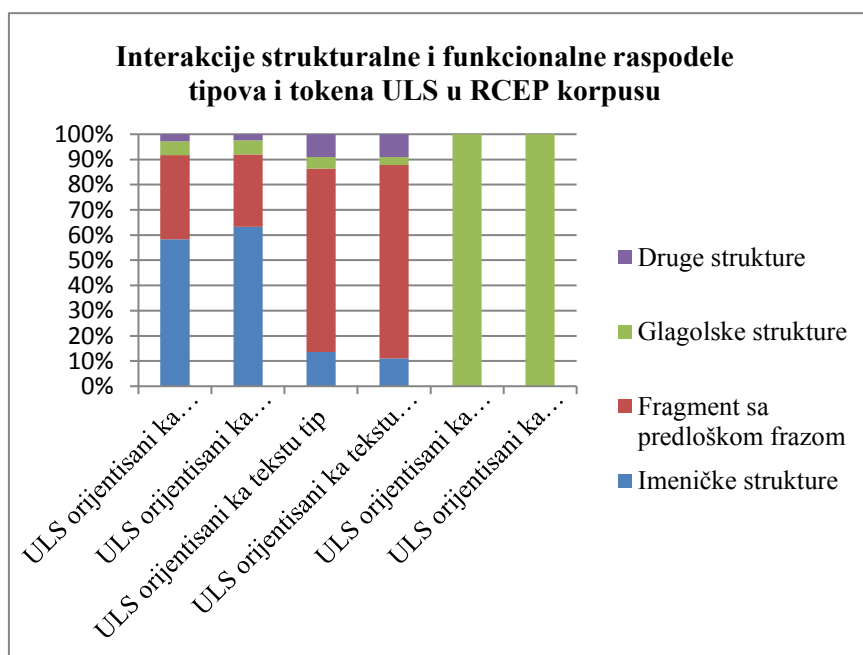
Grafikon 42. Interakcija strukturalne i funkcionalne raspodele tipova i tokena ULS u SCEP korpusu

4.7. Interakcija strukturalne i funkcionalne raspodele ULS u RCEP korpusu

ULS izvornih govornika rumunskog jezika raspodeljeni su u funkcionalne grupe, pa zatim po funkcionalnim grupama raspoređeni u strukturalne grupe kako bismo uvideli interakciju funkcionalnih i strukturalnih kategorija. Rezultati raspodele pokazuju da su u funkcionalnoj grupi *ULS orijentisani ka istraživanju* zastupljene sve četiri funkcionalne kategorije (v. Tabelu 75 i Grafikon 43). Najveći broj ULS u navedenoj funkcionalnoj kategoriji pripada *imeničkim strukturama*, koje su praćene *fragmentima sa predložkom frazom* i *glagolskim strukturama*, dok najmanji broj ULS pripada *drugim strukturama*. U funkcionalnoj kategoriji *ULS orijentisani ka tekstu*, kod izvornih govornika rumunskog jezika, takođe su zastupljene sve četiri strukturalne grupe ali sa drugačijom raspodelom. Najčešći su *fragmenti sa predložkom frazom* praćeni *imeničkim strukturama* i *drugim strukturama*, dok su *glagolske strukture* najmanje zastupljene. U funkcionalnoj kategoriji *ULS orijentisani ka učesniku* zastupljena je samo jedna strukturalna grupa – *glagolske strukture*.

Tabela 75. Interakcije strukturalne i funkcionalne raspodele ULS u RCEP korpusu

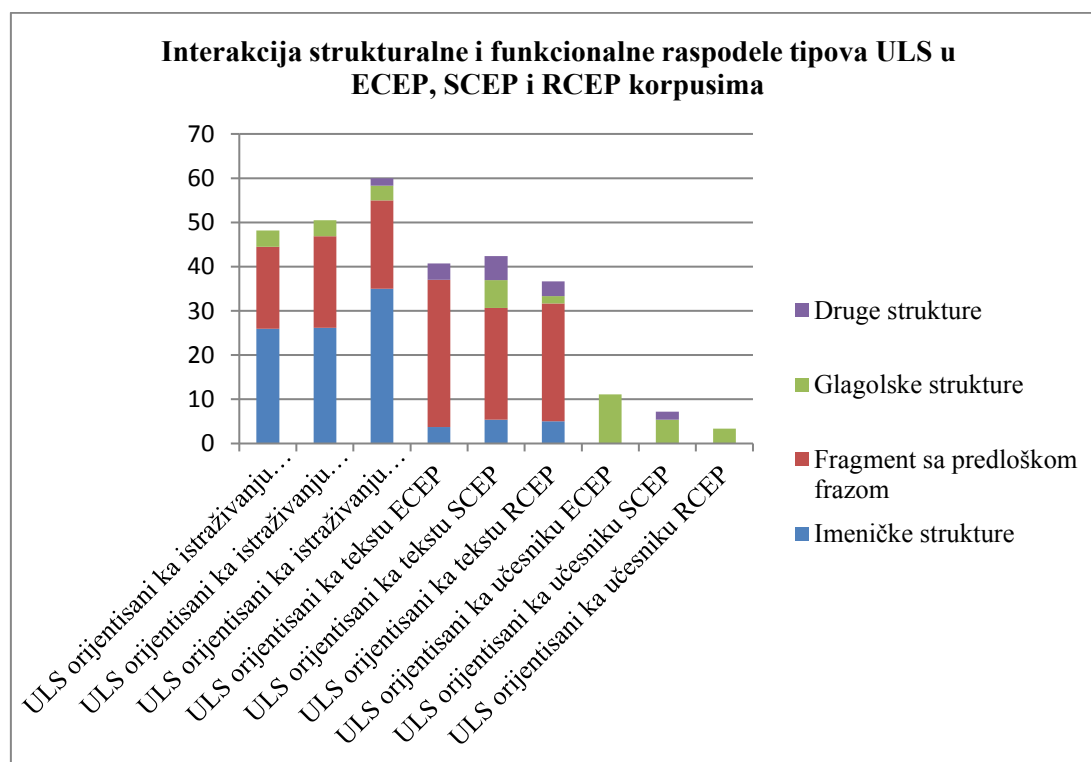
		<i>Imeničke strukture</i>	<i>Fragment sa predložkom frazom</i>	<i>Glagolske strukture</i>	<i>Druge strukture</i>
<i>ULS orijentisani ka istraživanju</i>	Tip (Token)	21 (638)	12 (392)	2 (56)	1 (25)
	Procenti	35 (27.06)	20 (12.38)	3.33 (2.37)	1.67 (1.06)
<i>ULS orijentisani ka tekstu</i>	Tip (Token)	3 (132)	16 (913)	1 (38)	2 (107)
	Procenti	5 (5.60)	26.67 (38.72)	1.67 (1.61)	3.33 (4.54)
<i>ULS orijentisani ka učesniku</i>	Tip (Token)	/	/	2 (57)	/
	Procenti	/	/	3.33 (2.42)	/



Grafikon 43. Interakcije strukturalne i funkcionalne raspodele tipova i tokena ULS u RCEP korpusu

4.8. Poređenje interakcije strukturalne i funkcionalne raspodele u ECEP, SCEP i RCEP korpusima

Sagledanjem interakcije strukturalnih i funkcionalnih raspodela ULS u sva tri korpusa, uviđamo da se različit broj strukturalnih grupa javlja u različitim funkcionalnim grupama kod različitih grupa govornika (v. Grafikon 41, Grafikon 42 i Grafikon 43). Poređenjem raspodele tipova, možemo videti da se u funkcionalnoj grupi *ULS orijentisani ka istraživanju* u ECEP i SCEP korpusu javljaju tri strukturalne grupe (*imeničke strukture*, *fragmenti sa predloškom frazom* i *glagolske strukture*), dok je u RCEP korpusu pored tri navedene prisutna i strukturalna grupa *druge strukture* (v. Grafikon 44). U sva tri korpusa u funkcionalnoj grupi *ULS orijentisani ka istraživanju* najzastupljenija je strukturalna kategorija *imeničke strukture* koja je praćena *fragmentima sa predloškom frazom*. U funkcionalnoj grupi *ULS orijentisani ka tekstu* u ECEP korpusu su zastupljene tri strukturalne kategorije (*imeničke strukture*, *fragmenti sa predloškom frazom* i *druge strukture*), dok su u korpusima neizvornih govornika engleskog jezika zastupljene sve četiri. U sva tri korpusa najzastupljenija strukturalna kategorija unutar funkcionalne kategorije *ULS orijentisani ka tekstu* je *fragmenti sa predloškom frazom*, dominantna zastupljenost ove strukturalne grupe je najočiglednija u ECEP korpusu. U funkcionalnoj kategoriji *ULS orijentisani ka učesniku*, u ECEP i RCEP korpusima, prisutna je samo strukturalna grupa *glagolske strukture*, dok je u SCEP korpusu, takođe, prisutna i strukturalna grupa *druge strukture*. Kada je reč o raspodeli tokena, ona prati gorenavedene raspodele sa jednim izuzetkom a to je – u SCEP korpusu u funkcionalnoj grupi *ULS orijentisani ka tekstu* druga najzastupljenija strukturalna grupa nisu *glagolske strukture* već *druge strukture*.



Grafikon 44. Interakcija strukturalne i funkcionalne raspodele tipova ULS u ECEP, SCEP i RCEP korpusima

4.9. Zajednički ULS i njihova strukturalna i funkcionalna raspodela

Identifikovano je 15 ULS koji su zajednički za sva tri korpusa. Izvorni govornici i neizvorni govornici kojima je maternji srpski dele 18 ULS, izvorni govornici i neizvorni govornici kojima je maternji rumunski jezik dele 15 ULS, neizvorni govornici kojima je maternji srpski i rumunski dele 27 ULS. Ukupno 18 ULS nađeno je samo u RCEP korpusu, odnosno 18 ULS je specifično za neizvorne govornike kojima je rumunski jezik maternji; 66 ULS se nalazi samo u SCEP korpusu, odnosno 66 ULS je specifično za neizvorne govornike kojima je maternji jezik srpski; devet ULS se nalaze samo u ECEP korpusu, odnosno – devet ULS je specifično za produkciju izvornih govornika engleskog jezika (v. Tabelu 76).

Tabela 76. Broj zajedničkih i specifičnih ULS u korpusima istraživanja – tip (token)

Zajednički za sva tri korpusa	Zajednički za ECEP i SCEP	Specifični za ECEP
15	18	9
	Zajednički za ECEP i RCEP	Specifični za SCEP
	15	66
	Zajednički za SCEP i RCEP	Specifični za RCEP
	27	18

4.9.1. Strukturalna i funkcionalna raspodela zajedničkih ULS za sva tri korpusa

Kako bismo bolje razumeli koje ULS najlakše usvajaju neizvorni govornici nezavisno od maternjeg jezika, uradićemo strukturalnu i funkcionalnu analizu zajedničkih ULS za sva tri korpusa. Strukturalna raspodela zajedničkih ULS za izvorne govornike engleskog jezika i neizvorne govornike kojima je maternji jezik srpski i rumunski pokazuje da je najzastupljenija strukturalna grupa *fragmenti sa predloškom frazom*, praćena *imeničkim strukturama*, dok su *glagolske strukture* i *druge strukture* manje zastupljene (v. Tabelu 77). Navedeni podaci ukazuju na to da raspodela zajedničkih ULS prati istu strukturalnu raspodelu koja se uočava u sva tri korpusa u celini, naime, i u njima su *fragmenti sa predloškom frazom* najzastupljenija kategorija a *druge strukture* najmanje zastupljena kategorija. Možemo zaključiti da se određeni broj iz svih strukturalnih kategorija ravnomerno lakše usvaja i adekvatnije koristi od drugih.

Tabela 77. Strukturalna raspodela zajedničkih ULS za sve ECEP, SCEP i RCEP korpuse

Strukturalna kategorija	Tip	Procenti
<i>Imeničke strukture</i>	5	33.33
<i>Fragmenti sa predloškom frazom</i>	7	46.67
<i>Glagolske strukture</i>	2	13.33
<i>Druge strukture</i>	1	6.67

Funkcionalna raspodela zajedničkih tipova ULS za izvorne govornike engleskog jezika i neizvorne govornike engleskog jezika kojima je maternji jezik srpski i rumunski pokazuje da je najzastupljenija funkcionalna kategorija *ULS orijentisani ka tekstu*, praćena *ULS orijentisanim ka istraživanju*, dok su *ULS orijentisani ka učesniku* najmanje zastupljeni (v. Tabelu 78). Na osnovu navedenih podataka uviđamo da funkcionalna raspodela zajedničkih ULS prati tendencije raspodele koje se uočavaju kod obe grupe neizvornih govornika, dok se razlikuje od raspodele koja je prisutna u ECEP korpusu, u kojem je najzastupljenija funkcionalna kategorija *ULS orijentisani ka istraživanju*. Drugim rečima, i ovde se uviđa tendencija neizvornih govornika engleskog jezika da zanemare relevantne ULS iz funkcionalne grupe *ULS orijentisani ka istraživanju*.

Tabela 78. Funkcionalna raspodela ULS zajedničkih za ECEP, SCEP i RCEP korpuse

Funkcionalna kategorija	Tipovi	Procenti
<i>ULS orijentisani ka istraživanju</i>	6	40
<i>ULS orijentisani ka tekstu</i>	8	53.33
<i>ULS orijentisani ka učesniku</i>	1	6.67

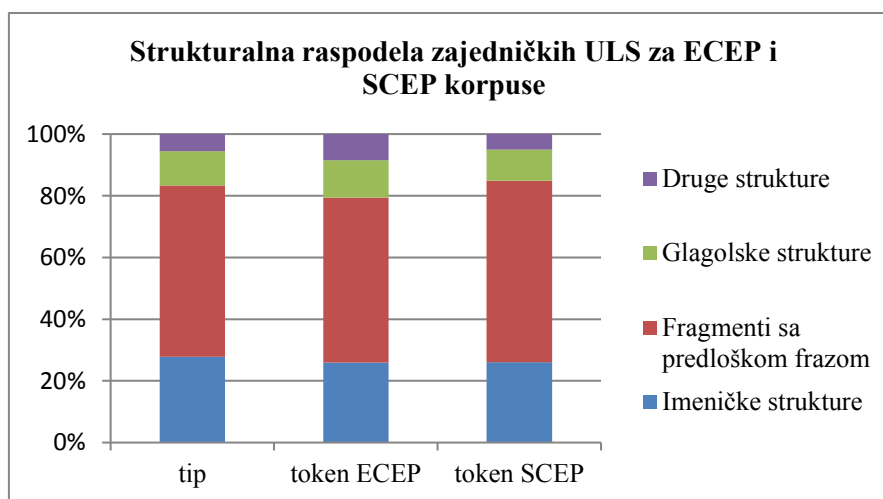
4.9.2. Strukturalna i funkcionalna raspodela zajedničkih ULS za ECEP i SCEP korpuse

Strukturalna i funkcionalna raspodela zajedničkih ULS za ECEP i SCEP korpus je urađena kako bismo bolje razumeli koje tipove ULS izvorni govornici srpskog uspešno koriste, da li se razlikuju u učestalosti upotrebe i na koji način. Izvorni govornici engleskog jezika i srpskog jezika dele 18 tipova ULS. Raspodela tipova za navedenih 18 ULS se logično razlikuju, te broj tipova u ECEP korpusu iznosi 555, dok u SCEP iznosi 1.069. Zajednički tipovi za ECEP i SCEP korpuse sačinjavaju 67% svih tipova u ECEP korpusu, a 16% svih tipova u SCEP korpusu. Zajednički tokeni za ECEP i SCEP korpuse sačinjavaju 71% svih tokena u ECEP korpusu, a 24% svih tokena u SCEP korpusu. Drugačije rečeno, izvorni govornici srpskog jezika koriste 67% optimalnih ULS za iskazivanje pripadnosti željenoj delatnoj zajednici, međutim, koriste i veliki broj ULS koji su gramatički ispravni ali nisu optimalni za iskazivanje pripadnosti željenoj delatnoj zajednici.

Strukturalna raspodela zajedničkih ULS za ECEP i SCEP korpus pokazuje da je najveća zastupljenost *fragmenti sa predloškom frazom* koji su praćeni *imeničkim strukturama*, dok su *glagolske i druge strukture* manje zastupljene (v. Tabelu 79 i Grafikon 45). Raspodela tipova i tokena u strukturalne grupe je procentualno ujednačena i kod izvornih govornika engleskog i kod izvornih govornika srpskog jezika.

Tabela 79. Strukturalna raspodela zajedničkih ULS za ECEP i SCEP korpuse

Strukturalna grupa	ECEP			SCEP		
	tip	token	token	tip	token	token
<i>Imeničke strukture</i>	5	144	244	27.78	25.95	22.83
<i>Fragmenti sa predloškom frazom</i>	10	297	552	55.56	53.51	51.64
<i>Glagolske strukture</i>	2	67	95	11.11	12.07	8.89
<i>Druge strukture</i>	1	47	178	5.56	8.49	4.40

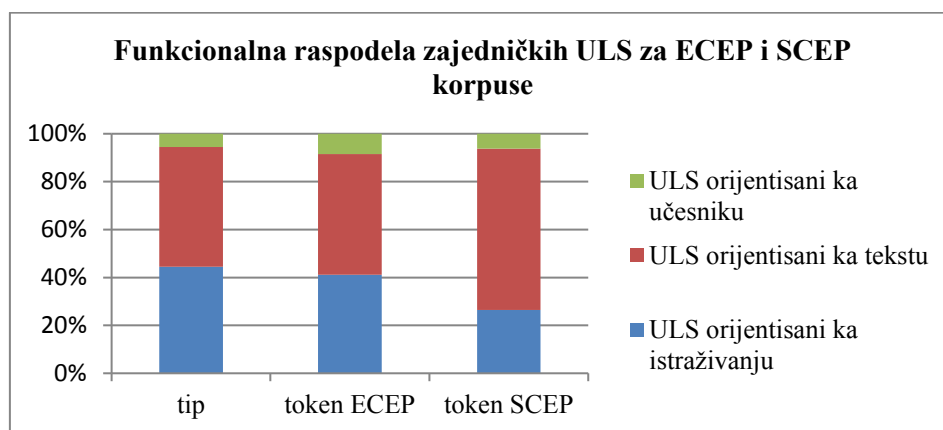


Grafikon 45. Strukturalna raspodela zajedničkih ULS za ECEP i SCEP korpuse

Funkcionalna raspodela zajedničkih ULS za ECEP i SCEP korpuse ukazuje na to da su najzastupljeniji *ULS orijentisani ka tekstu*, praćeni *ULS orijentisanim ka istraživanju*, dok su *ULS orijentisani ka učesniku* najmanje zastupljeni (v. Tabelu 80 i Grafikon 46). Razmatranjem raspodele tokena zajedničkih ULS za ECEP i SCEP korpus, uviđamo da izvorni govornici srpskog jezika učestalije upotrebljavaju *ULS orijentisane ka tekstu* a manje učestalo *ULS orijentisane ka istraživanju*, uporedno sa izvornim govornicima engleskog jezika. Navedeno dalje implicira da izvorni govornici srpskog jezika isuviše učestalo ili nedovoljno učestalo koriste optimalne ULS za pokazivanje pripadnosti u željenoj delatnoj zajednici.

Tabela 80. Funkcionalna raspodela zajedničkih ULS za ECEP i SCEP korpuse

Funkcionalna grupa	tip	token		procenti		
		ECEP	SCEP	tip	token ECEP	token SCEP
<i>ULS orijentisani ka istraživanju</i>	8	228	283	44.44	41.08	26.47
<i>ULS orijentisani ka tekstu</i>	9	281	720	50	50.36	67.35
<i>ULS orijentisani ka učesniku</i>	1	47	66	5.56	8.46	6.17



Grafikon 46. Funkcionalna raspodela zajedničkih ULS za ECEP i SCEP korpuse

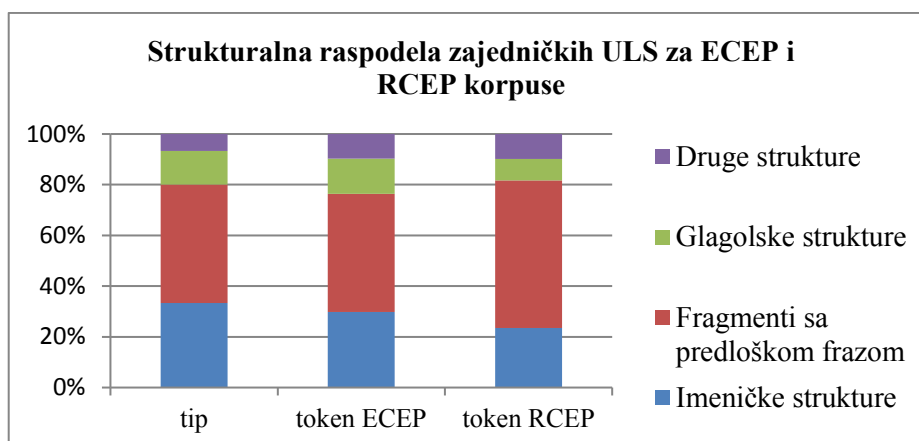
4.9.3. Strukturalna i funkcionalna raspodela zajedničkih ULS za ECEP i RCEP korpuse

Da bismo bolje razumeli koje tipove ULS izvorni govornici rumunskog jezika uspešno koriste, da li se razlikuju po učestalosti upotrebe i na koji način, urađena je strukturalna i funkcionalna raspodela zajedničkih ULS za ECEP i RCEP korpuse. Izvorni govornici engleskog jezika i rumunskog jezika, u kontekstu ovog istraživanja, dele 15 ULS, koji su zapravo istih 15 ULS koji su zajednički za sva tri korpusa. Raspodela tokena za zajedničkih 15 ULS se razlikuje u korpusima, te u ECEP korpusu iznosi 483 tokena a u RCEP korpusu iznosi 777 tokena. Zajednički tipovi za ECEP i RCEP korpuse sačinjavaju 56% svih tipova u ECEP korpusu a 25% svih tipova u RCEP korpusu. Zajednički tokeni za ECEP i RCEP korpuse sačinjavaju 62% svih tokena u ECEP korpusu a 33% svih tokena u RCEP korpusu. Na osnovu navedenih podataka, uviđamo da izvorni govornici rumunskog jezika koriste više od polovine optimalnih ULS koje sačinjavaju jednu četvrtinu svih ULS u RCEP korpusu.

Strukturalna analiza tipova i tokena zajedničkih ULS za ECEP i SCEP korpuse pokazala je da izvorni govornici rumunskog jezika, procentualno gledano, koriste veći broj tokena ULS iz grupe *fragmenti sa predložkom frazom* od izvornih govornika engleskog jezika, dok koriste manji broj tokena iz grupe *imeničke strukture* (v. Tabelu 81 i Grafikon 47).

Tabela 81. Strukturalna raspodela zajedničkih ULS za ECEP i RCEP korpuse

Strukturalna grupa	tip	token ECEP	token RCEP	Procenti		
				tip	token ECEP	token RCEP
<i>Imeničke strukture</i>	5	144	183	33.33	29.81	23.55
<i>Fragmenti sa predloškom frazom</i>	7	225	452	46.67	46.58	58.17
<i>Glagolske strukture</i>	2	67	66	13.33	13.87	8.49
<i>Druge strukture</i>	1	47	76	6.67	9.73	9.78

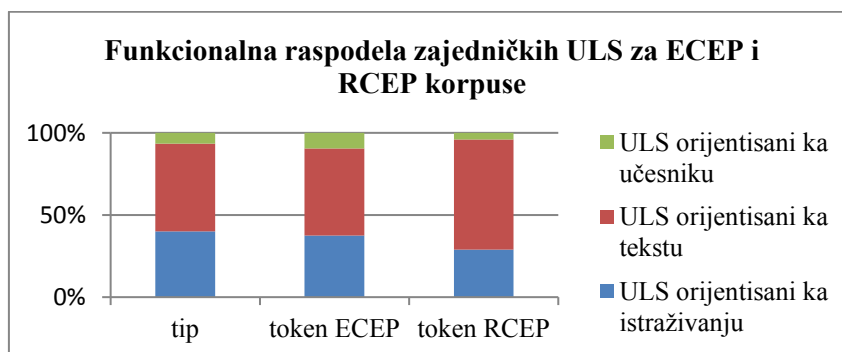


Grafikon 47. Strukturalna raspodela zajedničkih ULS za ECEP i RCEP korpuse

Funkcionalna raspodela tipova i tokena zajedničkih ULS za ECEP i RCEP korpuse pokazuje da su tokeni iz funkcionalne grupe *ULS orijentisani ka tekstu* zastupljeniji kod izvornih govornika rumunskog jezika, dok su *ULS orijentisani ka učesniku* i *ULS orijentisani ka istraživanju* više zastupljeni kod izvornih govornika engleskog. Drugim rečima, iako dve grupe govornika koriste identične tipove ULS, koriste ih sa drugačijom učestalošću (v. Tabelu 82 i Grafikon 48).

Tabela 82. Funkcionalna raspodela zajedničkih ULS za ECEP i RCEP korpuse

Funkcionalna grupa	tip	token ECEP	token RCEP	procenti		
				tip	token ECEP	token RCEP
<i>ULS orijentisani ka istraživanju</i>	6	181	225	40	37.47	28.96
<i>ULS orijentisani ka tekstu</i>	8	256	521	53.33	53	67.05
<i>ULS orijentisani ka učesniku</i>	1	46	31	6.67	9.52	3.99



Grafikon 48. Funkcionalna raspodela zajedničkih ULS za ECEP i RCEP korpuse

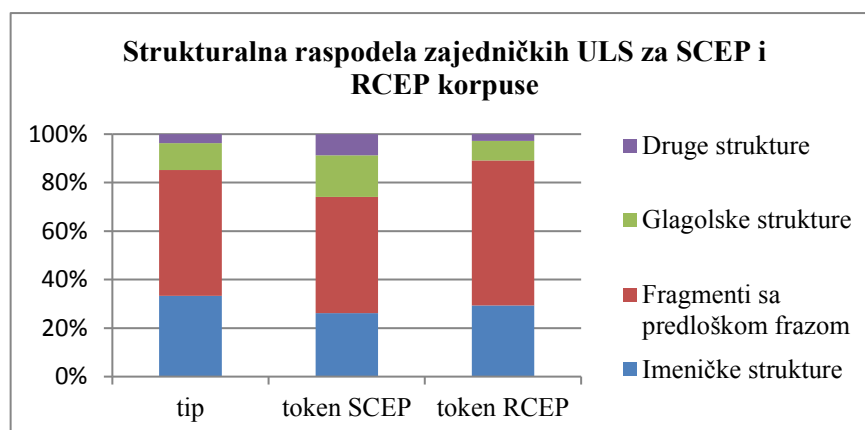
4.9.4. Strukturalna i funkcionalna raspodela zajedničkih ULS za SCEP i RCEP korpuse

Primenom strukturalne i funkcionalne raspodele ULS koji su zajednički za neizvorne govornike kojima je maternji jezik srpski i rumunski, težimo ka boljem razumevanju zajedničkih karakteristika u upotrebi ULS u naučnim radovima iz oblasti obrazovanja kod neizvornih govornika engleskog jezika. Dve grupe neizvornih govornika dele 27 tipova ULS. Broj tokena je različit u dva korpusa: u SCEP korpusu iznosi 1.426 tokena a u RCEP korpusu iznosi 1.124 tokena. Zajednički tipovi za dve grupe neizvornih govornika sačinjavaju 24% svih tipova u SCEP korpusu a 45% svih tipova u RCEP korpusu. Zajednički tokeni za dve grupe neizvornih govornika sačinjavaju 31% svih tokena u SCEP korpusu a 48% svih tokena u RCEP korpusu. Navedeni podaci ukazuju na to da zajednički tokeni sačinjavaju veći procenat svih tokena u oba korpusa neizvornih govornika engleskog jezika, uporedno sa tipovima. Drugim rečima, neizvorni govornici engleskog jezika imaju češće koriste ULS koji su im zajednički od drugih ULS.

Strukturalnom raspodelom tipova i tokena zajedničkih ULS uviđamo da se izvorni govornici srpskog jezika i izvorni govornici rumunskog jezika razlikuju po učestalosti upotrebe ULS koji su im zajednički. Naime, izvorni govornici srpskog jezika češće upotrebljavaju *glagolske strukture* i *druge strukture*, dok izvorni govornici rumunskog jezika češće upotrebljavaju *fragmente sa predložkom frazom* (v. Tabelu 83 i Grafikon 49).

Tabela 83. Strukturalna raspodela zajedničkih ULS za SCEP i RCEP korpuse

Strukturalna grupa	tip	token		procenti		
		SCEP	RCEP	tip	token SCEP	token RCEP
<i>Imeničke strukture</i>	9	374	330	33.33	26.23	29.36
<i>Fragmenti sa predložkom frazom</i>	14	683	672	51.85	47.90	59.79
<i>Glagolske strukture</i>	3	245	91	11.11	17.18	8.10
<i>Druge strukture</i>	1	124	31	3.70	8.70	2.76



Grafikon 49. Strukturalna raspodela zajedničkih ULS za SCEP i RCEP korpuse

Kako bi se ispitale potencijalne razlike, tj. asocijacije, između SCEP i RCEP korpusa u odnosu na strukturalne i funkcionalne distribucije tokena zajedničkih ULS korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih grupa u odnosu na tokene zajedničkih ULS prikazana je u Tabeli 84.

Tabela 84. *Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih grupa zajedničkih ULS u odnosu na tokene*

Strukturalna grupa	SCEP			RCEP		
	token	%	R	token	%	R
<i>Imeničke strukture</i>	374	26.2%	-1.8	330	29.4%	1.8
<i>Fragmenti sa predloškom frazom</i>	683	47.9%	-6.0*	672	59.8%	6.0*
<i>Glagolske strukture</i>	245	17.2%	6.7*	91	8.1%	-6.7*
<i>Druge strukture</i>	124	8.7%	6.2*	31	2.8%	-6.2*
Total	1426	100%		1124	100%	

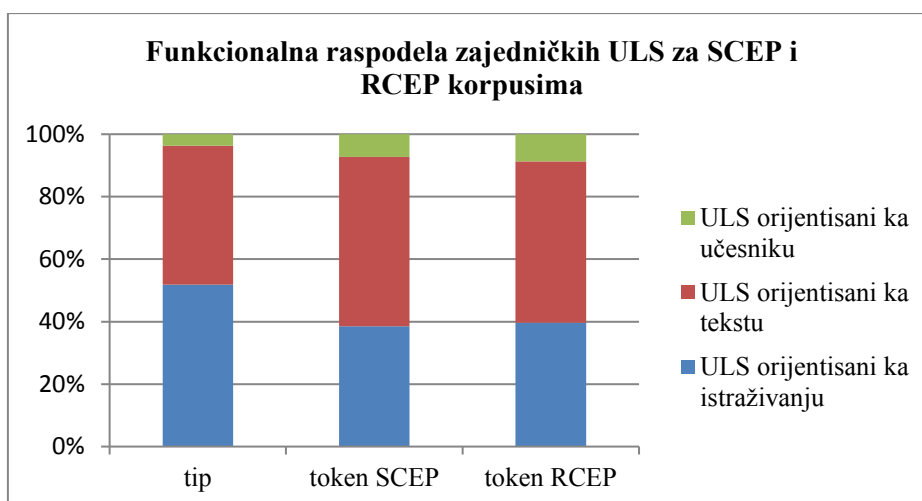
Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Hi-kvadrat test je bio značajan, $\chi^2(3, N = 2550) = 94.78, p < .001$, ukazujući da postoji statistički značajna povezanost između korpusa i strukturalnih grupa. U SCEP korpusu se značajno češće upotrebljuju *glagolske strukture* i *druge strukture*, dok se u RCEP korpusu značajno više koriste *fragmenti sa predloškom frazom*. Drugim rečima, neizvorni govornici engleskog jezika kojima je maternji jezik srpski i rumunski značajno se razlikuju u zastupljenosti tokena u strukturalnim grupama za zajedničke ULS.

Funkcionalna raspodela tipova i tokena zajedničkih ULS za neizvorne govornike engleskog jezika pokazuje da je najzastupljenija funkcionalna kategorija što se tiče tipova *ULS orijentisani ka istraživanju* ali što se tiče tokena – *ULS orijentisani ka tekstu*, dok su *ULS orijentisani ka učesniku* najmanje zastupljeni a kada je reč o tipovima i tokenima (v. Tabelu 85 i Grafikon 50). Navedena raspodela funkcionalnih kategorija je u skladu sa raspodelama koje se uviđaju za sve ULS u SCEP i RCEP korpusima. Izvorni govornici srpskog i rumunskog jezika češće koriste tokene za funkcionalne grupe *ULS orijentisani ka tekstu* a ređe za grupu *ULS orijentisani ka istraživanju* uporedno sa procentualnom zastupljenošću tipova. Ne uviđa se veća razlika u funkcionalnoj raspodeli tokena kod dve grupe neizvornih govornika.

Tabela 85. *Funkcionalna raspodela zajedničkih ULS za SCEP i RCEP korpusa*

Funkcionalna grupa	tip	procenti				
		token SCEP	token RCEP	tip	token SCEP	token RCEP
<i>ULS orijentisani ka istraživanju</i>	14	549	476	51.85	38.50	42.35
<i>ULS orijentisani ka tekstu</i>	12	773	622	44.44	54.21	55.34
<i>ULS orijentisani ka učesniku</i>	1	104	26	3.70	7.29	9.25



Grafikon 50. Funkcionalna raspodela zajedničkih ULS za SCEP i RCEP korpusa

Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih grupa u odnosu na tokene zajedničkih ULS prikazana je u Tabeli 86. Hi-kvadrat test je bio značajan, $\chi^2(2, N = 2550) = 33.04$, $p < .001$, ukazujući da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

Pregledom standardizovanih korigovanih reziduala utvrđeno je da u odnosu na očekivane frekvencije u izvorni govornici srpskog jezika učestalije koriste *ULS orijentisane na učesnike*, dok izvorni govornici rumunskog jezika učestalije koriste *ULS orijentisane ka istraživanju*.

Tabela 86. Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih grupa zajedničkih ULS u odnosu na tokene

Funkcionalna grupa	SCEP			RCEP		
	token	%	R	token	%	R
<i>ULS orijentisani ka istraživanju</i>	549	38.5%	-2.0*	476	42.3%	2.0*
<i>ULS orijentisani ka tekstu</i>	773	54.2%	-0.6	622	55.3%	0.6
<i>ULS orijentisani ka učesniku</i>	104	7.3%	5.7*	26	2.3%	-5.7*
Total	1426	100.0%		1124	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

4.10. Upotreba najzastupljenijih obrazaca ULS u ECEP, SCEP i RCEP korpusima

Kako bi se detaljnije sagledale razlike u upotrebi ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski i rumunski, analizirani su najzastupljeniji obrasci ULS iz tri strukturalne kategorije. Ispitani obrasci su sledeći:

- *the + imenica + of the/a* u okviru strukturalne grupe *imeničke strukture*;
- *in the + imenica + of* u okviru strukturalne grupe *fragmenti sa predloškom frazom*;
- *glagol u pasivu + fragment predloške fraze* i *uvodna it struktura + glagolska/ pridevska fraza* u okviru strukturalne grupe *glagolske strukture*.

Navedeni obrasci su odabrani za analizu po uzoru na prethodna istraživanja (Pan i sar., 2016; Lazić, 2017) i/ili zato što su najzastupljeniji obrasci u određenoj strukturalnoj kategoriji.

4. 10. 1. Obrazac *the + imenica + of the/a*

Izdvajanjem ULS koji prate najučestaliji obrazac u strukturalnoj grupi *imeničke strukture the + imenica + of the/a*, uvideli smo da je najveći broj tipova i tokena u navedenom obrascu izdvojen u korpusu izvornih govornika srpskog jezika (v. Tabelu 87). Ovakav rezultat se mogao pretpostaviti na osnovu činjenice da se najveći broj ULS upravo identifikovao u korpusu izvornih govornika srpskog jezika. Izvorni govornici engleskog jezika koriste pet tipova ULS i 188 tokena u *the + imenica + of the/a* obrascu, što čini 83,33% tipova u strukturalnoj podgrupi *imeničke fraze sa of fragmentom*, 62,50% tipova u strukturalnoj grupi *imeničke strukture*, a 18,52% tokena unutar celog ECEP korpusa (v. Tabelu 97). U produkciji izvornih govornika srpskog jezika izdvojeno je 16 tipova i 579 tokena koji prate *the + imenica + of the/a* obrazac, što sačinjava 59,62% tipova u strukturalnoj podgrupi, 44,44% tipova u strukturalnoj grupi, a 14,41% tipova u celom SCEP korpusu. U korpusu neizvornih govornika engleskog jezika kojima je maternji jezik rumunski izdvojeno je 13 tokena i 434 tipa ULS koji prate obrazac *the + imenica + of the/a*. Oni sačinjavaju 72,22% tipova u strukturalnoj podgrupi *imeničke fraze sa of fragmentom*, 54,17% tipova u strukturalnoj grupi *imeničke strukture*, a 18,41% tokena unutar celog RCEP korpusa. Na osnovu navedenih podataka, možemo da zaključimo da je u SCEP korpusu obrazac *the + imenica + of the/a* procentualno najmanje zastupljen, što ukazuje na raznolikost obrazaca ULS kod izvornih govornika srpskog jezika. Procentualno gledano (v. Tabelu 88), u odnosu na cele korpusse, ULS koji prate obrazac *the + imenica + of the/a* su najzastupljeniji u RCEP korpusu. Procentualna zastupljenost tipova i tokena navedenog obrasca su relativno ujednačene u sva tri korpusa ponaosob.

Detaljnijim pogledom na tipove koji prate *the + imenica + of the/a* obrazac, uviđamo da se neodređeni član na kraju obrasca javlja samo u dva tipa *the existence of a* i *the development of a*, a oba navedena tipa se nalaze u RCEP korpusu. Izvorni govornici engleskog jezika koriste pet reči sa punim leksičkim značenjem unutar obrasca, od kojih su dve specifične za izvorne govornike engleskog jezika: *nature* i *use*. Neizvorni govornici engleskog jezika kojima je maternji jezik srpski koriste 16 reči sa punim leksičkim značenjem unutar obrasca, od kojih su sedam specifične za izvorne govornike srpskog jezika – *implementation, aim, application, characteristics, assessment, end, influence* i *effect*. Neizvorni govornici engleskog jezika kojima je maternji jezik rumunski koriste 13 reči sa punim leksičkim značenjem, od kojih je četiri specifično za izvorne govornike rumunskog jezika – *level, purpose, existance* i *majority*. Dve grupe neizvornih govornika dele upotrebu pet reči sa leksičkim značenjem unutar ULS koji prate obrazac *the + imenica + of the/a* – *development, beginning, analysis, content* i *importance*. Interesantno je da su dve reči sa leksičkim značenjem koje su specifične za neizvorne govornike – jedna za neizvorne govornike kojima je maternji jezik srpski, druga za izvorne govornike kojima je maternji jezik rumunski – zapravo sinonimi – *aim* i *purpose*. To je jedan od primera kako ove dve grupe govornika iskazuju isto značenje na različite načine pomoću različitih tipova ULS unutar istog strukturalnog obrasca.

Tabela 87. *Obrazac the + imenica + of the/a u ECEP, SCEP i RCEP korpusima*

Korpus	<i>the + imenica + of + the/a</i>	ukupno	
		tip	token
ECEP	the results of the (28) , the nature of the (24), the use of the (24), the role of the (22) , the quality of the (20) .	5	188
SCEP	the results of the (research) (122) , the implementation of the (54), <i>the development of the (49)</i> , <i>the importance of the (44)</i> , the quality of the (44) , <i>the analysis of the (39)</i> , the role of the (33) , <i>the content of the (28)</i> , the aim of the (26), the application of the (26), the characteristics of the (26), the assessment of the (25), (at) the end of the (54), <i>(at) the beginning of the (37)</i> , the influence of the (24), the effects of the (20).	16	579
RCEP	the results of the (71) , the quality of the (46) , <i>the development of the (41)</i> , <i>(at) the beginning of the (39)</i> , <i>the analysis of the (37)</i> , <i>the content of the (36)</i> , the level of the (27), <i>the importance of the (25)</i> , the purpose of the (25), the role of the (24) , the existence of a (23), the majority of the (20), the development of a (20).	13	434

Tabela 88. *Procentualna zastupljenost obrasca the + imenica + of the/a u ECEP, SCEP i RCEP korpusima*

	ECEP		SCEP		RCEP	
	tip	token	tip	token	tip	token
Unutar podgrupe <i>imeničke fraze sa of fragmentom</i>	83.33	84.29	59.62	65.03	72.22	74.96
Unutar grupe <i>imeničke fraze</i>	62.50	54.63	44.44	50.15	54.17	56.36
Unutar celog korpusa	18.52	15.19	14.41	14.35	21.67	18.41

4. 10. 2. Obrazac *in the + imenica + of*

ULS koji prate obrazac *in the + imenica + of* u korpusu izvornih govornika engleskog jezika i korpusima neizvornih govornika engleskog jezika kojima je maternji jezik srpski i rumunski predstavljani su u Tabeli 89. U SCEP korpusu izdvojeno je devet tipova i 552 tokena ULS koji prate navedeni obrazac, što je veći broj uporedno sa izdvojenim ULS koji prate *in the + imenica + of* obrazac u ECEP i RCEP korpusima. Međutim, pošto se tri korpusa razlikuju u broju identifikovanih tipova i tokena, neophodno je sagledati procentualnu zastupljenost tipova i tokena ULS koji prate navedeni obrazac kako bi se uradilo poređenje. Procentualna zastupljenost ULS koji prate *in the + imenica + of* obrazac u strukturalnim podgrupama, grupama i celom korpusu predstavljeno je u Tabeli 90. Na osnovu procentualne zastupljenosti uviđamo da izvorni govornici srpskog jezika koriste najviše tipova i tokena iz navedenih obrascu u strukturalnoj podgrupi, grupi i celom korpusu. Takođe, uviđamo da je procentualna zastupljenost tokena veća kod neizvornih govornika engleskog jezika uporedno sa tipovima, dok je kod izvornih govornika procentualna zastupljenost tipova i tokena poprilično ujednačena. Iako su tokeni više zastupljeni od tipova u oba korpusa neizvornih govornika, izvorni govornici srpskog koriste veći procenat tokena uporedno sa izvornim govornicima rumunskog jezika. Interesantno je da se procentualna zastupljenost tokena ULS koji prate *in the + imenica + of* obrazac unutar celog korpusa poklapa u SCEP i RCEP korpusima i da iznosi po 12% svih ULS. Procentualna zastupljenost tipova ULS koji prate *in the + imenica + of* obrazac je sličniji između ECEP i RCEP korpusa nego između ECEP i SCEP korpusa. Štaviše, identičan procenat zastupljenosti ULS koji prate navedeni obrazac identifikovan je u strukturalnoj grupi *fragmenti sa predložkom frazom* i u SCEP i u RCEP korpusu (14,29%).

Tabela 89. Obrazac „in the + imenica + of” u ECEP, SCEP i RCEP korpusima

Korpus	in the + imenica + of	ukupno	
		tip	token
ECEP	<i>in the context of</i> (30), <i>in the form of</i> (27).	2	57
SCEP	<i>in the field of (education)</i> (141), <i>in the process of</i> (70), <i>in the context of</i> (65), <i>in the case of</i> (63), in the implementation of (52), <i>in the form of</i> (40), in the development of (34), in the area of (30), in the domain of (27).	9	552
RCEP	<i>in the case of (the)</i> (137), <i>in the field of</i> (88), <i>in the process of</i> (45), in the number of (20).	4	290

Kada sagledamo upotrebu reči sa punim leksičkim značenjem, uviđamo da se u ECEP korpusu koriste dve reči sa punim leksičkim značenjem, od kojih nijedna nije specifična za upotrebe izvornih govornika engleskog u navedenom obrascu već se obe nalaze i u SCEP korpusu – *context* i *form*. Izvorni govornici srpskog jezika u okviru ULS koji prate *in the + imenica + of* obrazac koriste devet reči sa punim leksičkim značenjem, od kojih je četiri specifično za SCEP korpus – *implementation, development, area* i *domain*. Izvorni govornici rumunskog jezika u produkciji na engleskom jeziku koriste četiri reči sa punim leksičkim značenjem unutar ULS koji prate *in the + imenica + of* obrazac, od njih je samo jedna specifična za izvorne govornike rumunskog jezika – *number*. Neizvorni govornici kojima je maternji jezik srpski i rumunski dele tri reči sa punim leksičkim značenjem koje se pojavljuju u ULS koji prate *in the + imenica + of* obrazac: *case, field* i *process*.

Tabela 90. Procentualna zastupljenost obrasca „in the + imenica + of” u ECEP, SCEP i RCEP korpusima

	ECEP		SCEP		RCEP	
	tip	token	tip	token	tip	token
Unutar podgrupe <i>predloške fraze sa of fragmentom</i>	20	20.96	26.47	40.44	22.22	34.73
Unutar grupe <i>fragmenti sa predloškom frazom</i>	14.29	14.54	18	27.18	14.29	22.22
Unutar celog korpusa	7.41	7.34	8.11	12.17	6.67	12.30

4. 10. 3. Obrazac *glagol u pasivu + fragment predloške fraze*

Obrazac *glagol u pasivu + fragment predloške fraze* analizirao se u prethodnim istraživanjima kao jedan od najzastupljenijih obrazaca u funkcionalnoj grupi *glagolske strukture* (Pan i sar., 2016; Lazić, 2017). Međutim, u korpusima koji su korišćeni u ovom istraživanju navedeni obrazac se javlja samo u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik srpski. Radi preciznijih poređenja sa prethodnim istraživanjima, sagledaće se upotreba obrasca *glagol u pasivu + fragment predloške fraze*, ali će se sagledati i upotreba obrasca *uvodna it struktura + glagolska/pridevska fraza*, koji je zapravo najproduktivniji obrazac iz strukturalne kategorije *glagolske strukture* u ECEP, SCEP i RCEP korpusima.

U korpusu neizvornih govornika engleskog jezika kojima je maternji jezik srpski izdvojeno je tri tipa i 78 tokena ULS koji prate obrazac *glagol u pasivu + fragment predloške fraze –is based on the* (31), *are shown in table* (27), *can be found in* (20). ULS koji prate obrazac *glagol u pasivu + fragment predloške fraze* sačinjavaju 17,65% tipova a 10,71% tokena u strukturalnoj grupi *glagolske strukture* u SCEP korpusu, dok u celom SCEP korpusu sačinjavaju 2,7% tipova i 1,72% tokena. Na osnovu navedenih podataka, možemo zaključiti da tipovi ULS koji prate navedeni obrazac ne samo što su retki u SCEP korpusu već se i koriste sa manjom učestalošću pošto je

procentualna zastupljenost tokena manja od procentualne zastupljenosti tipova. Izvorni govornici srpskog jezika koriste četiri reči sa punim leksičkim značenjem u ULS koji prate obrazac *glagol u pasivu + fragment predloške fraze: base, show, table, find*.

4. 10. 4. Obrazac *uvodna it struktura + glagolska/ pridevska fraza*

ULS koji prate obrazac *uvodna it struktura + glagolska/ pridevska fraza* u korpusu izvornih govornika engleskog jezika u korpusima neizvornih govornika engleskog jezika kojima je maternji jezik srpski i rumunski predstavljeni su u Tabeli 91. U produkciji izvornih govornika engleskog jezika izdvojeno je dva tipa i 67 tokena ULS koji prate navedenu konstrukciju, što čini 50% svih tipova iz strukturalne grupe glagolske strukture a 7,46% svih tipova u ECEP korpusu (v. Tabelu 92). Kod izvornih govornika srpskog jezika identifikovano je pet tipova i 313 tokena ULS koji prate obrazac *uvodna it struktura + glagolska/ pridevska fraza*, dok je kod izvornih govornika rumunskog jezika identifikovano dva tipa i 57 tokena. Uzimajući u obzir procentualnu zastupljenost tipova i tokena ULS koji prate obrazac, uviđamo da se najveća razlika između tipova i tokena uviđa kod izvornih govornika srpskog jezika. Naime, kod izvornih govornika srpskog tipovi koji prate navedeni obrazac čine 29,41% svih tipova u strukturalnoj kategoriji *glagolske strukture*, dok u istoj kategoriji oni sačinjavaju čak 42,99% tokena. Ovaj podatak nam dalje govori da izvorni govornici srpskog učestalo koriste navedene ULS, što je suprotno od onog što se može primetiti u korpusu izvornih govornika rumunskog jezika gde je procentualna zastupljenost tokena manja od tipova za ULS koji prate navedeni obrazac. Najveći broj ULS koji prate obrazac *uvodna it struktura + glagolska/ pridevska fraza* identifikovan je kod izvornih govornika srpskog jezika, međutim, kada se sagleda procentualna zastupljenost ULS koji prate taj obrazac u strukturalnoj grupi i celom korpusu uviđamo da je taj procenat zapravo najmanji kod izvornih govornika srpskog jezika a najveći kod izvornih govornika engleskog jezika.

Tabela 91. *Obrazac uvodna it struktura + glagolska/ pridevska fraza u ECEP, SCEP i RCEP korpusima*

Korpus	<i>uvodna it struktura + glagolska/ pridevska fraza</i>	ukupno	
		tip	token
ECEP	it is important to (46), <i>it is possible that</i> (21).	2	67
SCEP	<i>it is necessary to</i> (104), (it) can be concluded (that the) (66), it is important to (66), <i>it is possible to</i> (51), it can be seen (26).	5	313
RCEP	it is important to (31), <i>it is necessary to</i> (26).	2	57

Pregledom upotrebe reči sa punim leksičkim značenjem, uviđamo da se u ECEP korpusu koriste dve reči sa punim leksičkim značenjem u okviru ULS koji prate obrazac *uvodna it struktura + glagolska/ pridevska fraza*, od kojih je jedna koja se javlja i u SCEP korpusu (*possible*), a jedna je zajednička za sva tri korpusa (*important*). U SCEP korpusu koriste se pet reči sa punim leksičkim značenjem u ULS koji prate navedeni obrazac, od kojih je tri specifično za izvorne govornike srpskog jezika – *conclude* i *see*. U RCEP, kao i u ECEP korpusu, koriste se dve reči sa punim leksičkim značenjem u ULS koji prate obrazac *uvodna it struktura + glagolska/ pridevska fraza*. Jedna od reči je zajednička za sva tri korpusa a jedna se deli sa izvornim govornicima srpskog jezika – *necessary*.

Tabela 92. Procentualna zastupljenost obrasca *uvodna it struktura + glagolska/ pridevska fraza* u ECEP, SCEP i RCEP korpusima

	ECEP		SCEP		RCEP	
	tip	token	tip	token	tip	token
Unutar grupe <i>glagolske strukture</i>	50	54.92	29.41	42.99	40	37.75
Unutar celog korpusa	7.46	8.62	4.50	6.90	3.33	2.42

Na osnovu u ovom poglavlju predstavljenih podataka, možemo zaključiti da je potvrđena osma radna hipoteza koja glasi: *Najzastupljeniji okviri će se drugačije koristiti kod izvornih govornika engleskog jezika i dve grupe neizvornih govornika engleskog jezika u naučnim radovima u oblasti obrazovanja.*

Uzimajući u obzir sve rezultate dobijene u ovom istraživanju možemo da zaključimo da je potvrđena i glavna hipoteza istraživanja koja glasi: *Karakteristike upotrebe ULS u naučnim radovima iz oblasti obrazovanja će se razlikovati kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski, u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca.*

4.11. Rezultati analize uticaja transfera

4.11.1. Unutargrupna homogenost

Primenom parametara za izdvajanje ULS zadovoljen je kriterijum međugrupne homogenosti. Korišćenjem minimalne frekvencije od 40 ponavljanja u milion reči i minimalne disperzije u 10% tekstova identifikovano je 29 ULS kod izvornih govornika engleskog jezika, 129 ULS kod izvornih govornika srpskog jezika i 71 kod izvornih govornika rumunskog jezika. Nakon što su spiskovi prečišćeni, finalan broj izdvojenih ULS u korpusu izvornih govornika engleskog je 27 ULS, u korpusu izvornih govornika srpskog je 111 ULS, a u korpusu izvornih govornika rumunskog 60 ULS (v. Tabelu 93).

Tabela 93. *Izdvojeni ULS u korpusima istraživanja sa nepročišćenog i pročišćenog spiska – tipovi i tokeni*

Korpus	Nepročišćen spisak	Prečišćen spisak
ECEP	29 (828)	27 (777)
SCEP	129 (5204)	111 (4536)
RCEP	71 (2727)	60 (2358)

4.11.2. Međugrupna heterogenost

Kako bi se prikupili dokazi za međugrupnu heterogenost, upoređeni su spiskovi izdvojenih ULS kod izvornih govornika engleskog, srpskog i rumunskog jezika i izdvojenih ULS koji su specifični za svaku grupu govornika, odnosno koji su se identifikovali samo u jednom korpusu a ne i u druga dva. Poređenjem spiskova utvrdilo se da je devet ULS specifično za izvorne govornike engleskog jezika, 66 ULS specifično za neizvorne govornike engleskog jezika kojima je maternji jezik srpski, a 18 ULS specifično za neizvorne govornike engleskog kojima je maternji jezik rumunski, v. Tabelu 94.

Tabela 94. *Broj specifičnih ULS kod izvornih govornika engleskog, neizvornih govornika kojima je izvorni jezik srpski i neizvornih govornika kojima je izvorni jezik rumunski jezik (tipovi i tokeni)*

Korpus	Tip (Token)
ECEP	9 (222)
SCEP	66 (2041)
RCEP	18 (457)

Najveći broj specifičnih ULS nalazi se u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik srpski, što je i bilo očekivano, a s obzirom na činjenicu da se najveći broj ULS izdvojio u SCEP korpusu. Najmanji broj specifičnih ULS se nalazi u korpusu izvornih govornika engleskog jezika. Ovakvi rezultati mogli su se pretpostaviti i same brojke nam ne pružaju mnogo korisnih informacija, te je neophodno sagledati detaljnije strukturalnu i funkcionalnu raspodelu specifičnih ULS. Na taj način mogu da se sagledaju tendencije upotrebe ULS koje nisu tipične za izvorne govornike i koje mogu biti smernice za unapređenje produkcije neizvornih govornika koji žele da pišu ili prevode u ovoj oblasti i žanru. Svakako je u ovom delu istraživanja ključno detaljnije uporediti upotrebu specifičnih ULS kako bi se sagledale i bolje razumele razlike u upotrebi kod različitih grupa govornika.

4.11.2.1. Strukturalna raspodela specifičnih ULS

Strukturalna analiza specifičnih ULS za svaku grupu govornika je urađena prateći taksonomiju koju su predložili Bajber i sar. (1999). Ovaj korak je neophodan kako bi se uvidele strukturalne karakteristike specifičnih ULS i preciznije sagledale razlike u njihovoj upotrebi kod izvornih i neizvornih govornika engleskog jezika.

Strukturalna raspodela specifičnih ULS kod izvornih govornika engleskog jezika

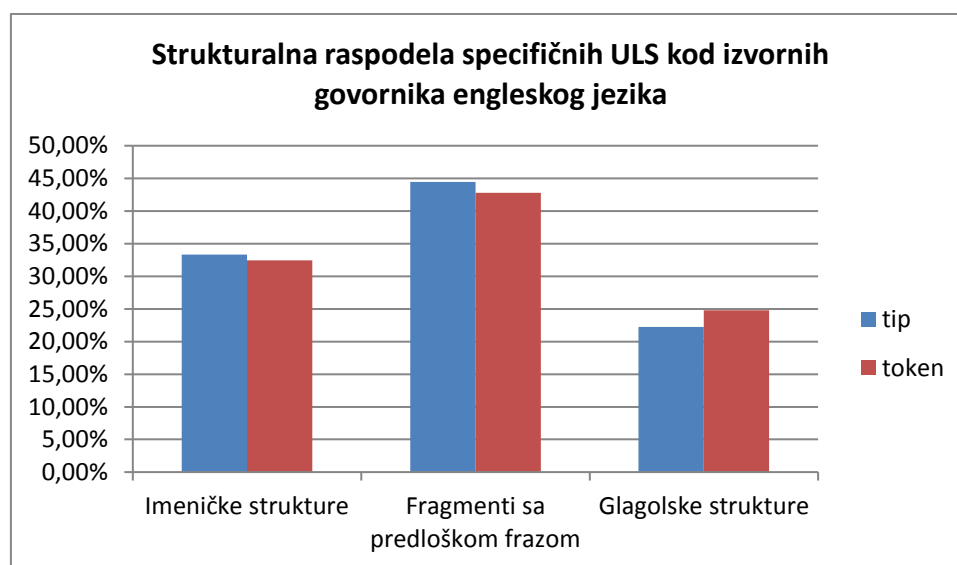
Strukturalna analiza specifičnih ULS u korpusu izvornih govornika engleskog jezika pokazala je da su najzastupljeniji *fragmenti sa predloškom frazom*, praćene *imeničkim frazama*, a da su *glagolske strukture* najmanje zastupljene, v. Tabelu 95 i Grafikon 51. Strukturalna raspodela specifičnih ULS je u skladu sa raspodelom svih ULS u celom ECEP korpusu, što nam ukazuje na to da neizvorni govornici nisu zanemarili u svojoj produkciji određenu strukturalnu grupu koja je prisutna kod izvornih govornika više nego ostale, već da je određen manji procenat iz svake strukturalne grupe zanemaren, osim grupe *druge strukture* koja nije prisutna u primerima ULS specifičnih za izvorne govornike engleskog jezika. Nedostatak zastupljenosti *drugih struktura* među specifičnim ULS za izvorne govornike engleskog jezika može da se objasni činjenicom da su veoma malo zastupljene i u celom ECEP korpusu, naime, u pitanju je samo jedan primer. Sagledanjem procentualne zastupljenosti specifičnih ULS u celom ECEP korpusu po strukturalnim podgrupama, uviđamo da je najveća zastupljenost ULS specifičnih za izvorne govornike prisutna u strukturalnoj kategoriji *glagolske strukture* (50%), praćeno kategorijom *imeničke strukture* (37,50%), v. Tabelu 96. Ovi podaci nam ukazuju na to da neizvorni govornici engleskog kojima je maternji jezik srpski ili rumunski imaju veću tendenciju da zanemare ULS karakteristične za izvorne govornike u navedenim strukturalnim kategorijama.

Tabela 95. *Strukturalna raspodela specifičnih ULS kod izvornih govornika engleskog jezika*

Strukturalna kategorija	Tip (Token)	Procenti
<i>Imeničke strukture</i>	3 (72)	33.33 (32.43)
<i>Fragmenti sa predloškom frazom</i>	4 (95)	44.44 (42.79)
<i>Glagolske strukture</i>	2 (55)	22.22 (24.77)

Tabela 96. *Zastupljenost specifičnih ULS u ECEP korpusu po strukturalnim grupama – tipovi i tokeni*

	ECEP svi ULS	ECEP specifični ULS	%
<i>Imeničke strukture</i>	8 (216)	3 (72)	37.50 (33.33)
<i>Fragmenti sa predloškom frazom</i>	14 (392)	4 (95)	28.57 (24.23)
<i>Glagolske strukture</i>	4 (122)	2 (55)	50 (45.08)
<i>Druge strukture</i>	1 (47)	0	0%

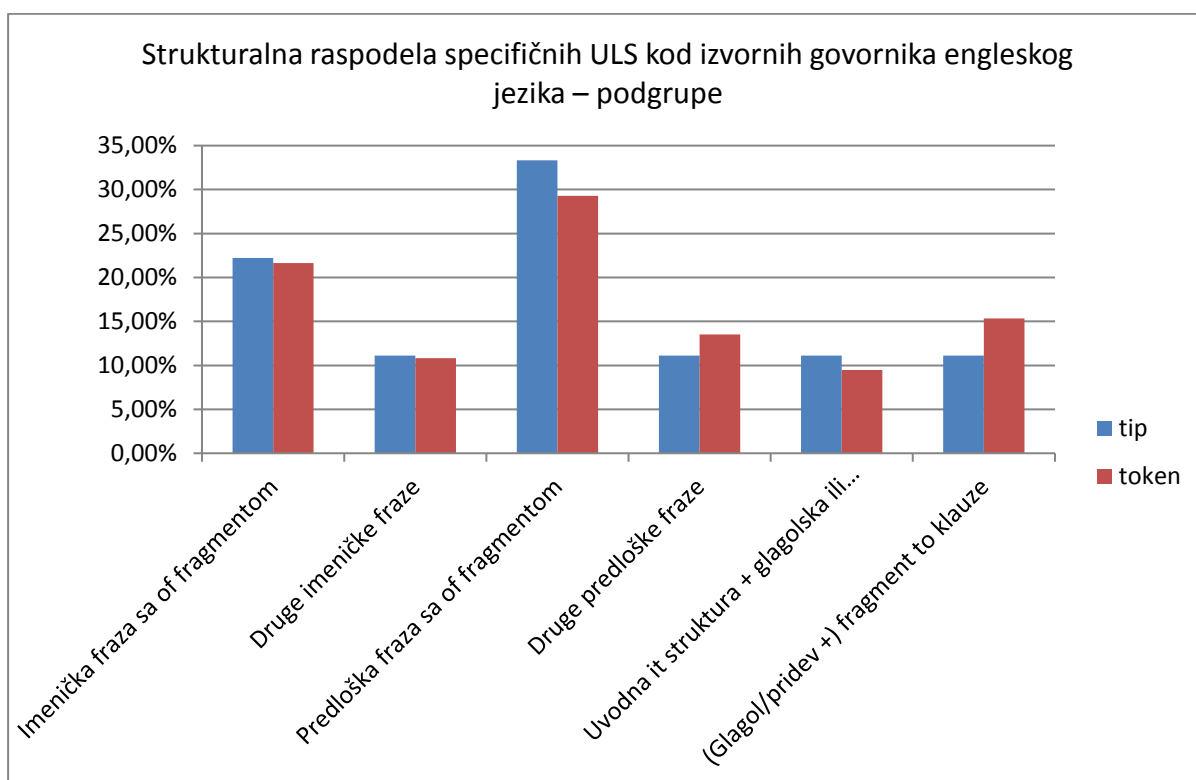


Grafikon 51. *Strukturalna raspodela specifičnih ULS kod izvornih govornika engleskog jezika*

Analiza strukturalnih podgrupa pokazala je da su *predloške fraze sa of fragmentom* najzastupljenija podgrupa među specifičnim ULS za ECEP korpus sa tri tipa ULS, praćena podgrupom *imeničke fraze sa of fragmentom* koja je zastupljena sa dva tipa ULS, dok su ostale podgrupe zastupljene sa po jednim tipom ULS (v. Tabelu 97 i Grafikon 52). Od 12 strukturalnih podgrupa samo šest je prisutno kod ULS specifičnih za izvorne govornike engleskog jezika. Na osnovu ovog podatka možemo pretpostaviti da neizvorni govornici kojima je maternji jezik srpski ili rumunski pri pisanju ili prevođenju naučnih radova iz oblasti obrazovanja, imaju veću tendenciju da zanemare upotrebu relevantnih ULS u podgrupama: *imenička fraza sa of fragmentom*, *druge imeničke fraze*, *predloške fraze sa of fragmentom*, *druge predloške fraze*, *uvodna it struktura + glagolska ili pridevska fraza i (glagol/pridev)+ fragment to klauze* nego u ostalih šest podgrupa. Dodatna istraživanja su neophodna da bi se proverilo da li je navedena tendencija prisutna i u drugim žanrovima i oblastima.

Tabela 97. *Strukturalna raspodela specifičnih ULS kod izvornih govornika engleskog jezika – podgrupe*

<i>Strukturalna potkategorija</i>	<i>Tip (Token)</i>	<i>Procenti</i>
<i>Imenička fraza sa of fragmentom</i>	2 (48)	22.22 (21.62)
<i>Druge imeničke fraze</i>	1 (24)	11.11 (10.81)
<i>Predložka fraza sa of fragmentom</i>	3 (65)	33.33 (29.28)
<i>Druge predložke fraze</i>	1 (30)	11.11 (13.51)
<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	1 (21)	11.11 (9.46)
<i>(Glagol/pridev +) fragment to klauze</i>	1 (34)	11.11 (15.32)



Grafikon 52. *Strukturalna raspodela specifičnih ULS kod izvornih govornika engleskog jezika – podgrupe*

Strukturalna raspodela specifičnih ULS kod neizvornih govornika engleskog kojima je maternji jezik srpski

Strukturalna raspodela ULS specifičnih za neizvorne govornike engleskog jezika kojima je maternji jezik srpski pokazala je da su najzastupljeniji *fragmenti sa predložkom frazom*, praćeni *imeničkim strukturama* i *glagolskim strukturama*. Najmanje je zastupljena strukturalna grupa *druge strukture*, v. Tabelu 98 i Grafikon 53. Isti redosled zastupljenosti strukturalnih grupa uviđa se za sve ULS u SCEP korpusu. Međutim, možemo uvideti da je procenat zastupljenosti *glagolskih struktura* kod specifičnih ULS (19,70%) veći nego kod svih ULS u SCEP korpusu (15,32%). Pregledom konkretnih tipova ULS u strukturalnoj kategoriji *glagolske strukture*, uviđamo da se 13 tipova nalazi na spisku specifičnih ULS za izvorne govornike srpskog od 17 koji se nalaze u čitavom SCEP korpusu. To bi dalje značilo da je 76,47% od svih *glagolskih struktura* u SCEP korpusu

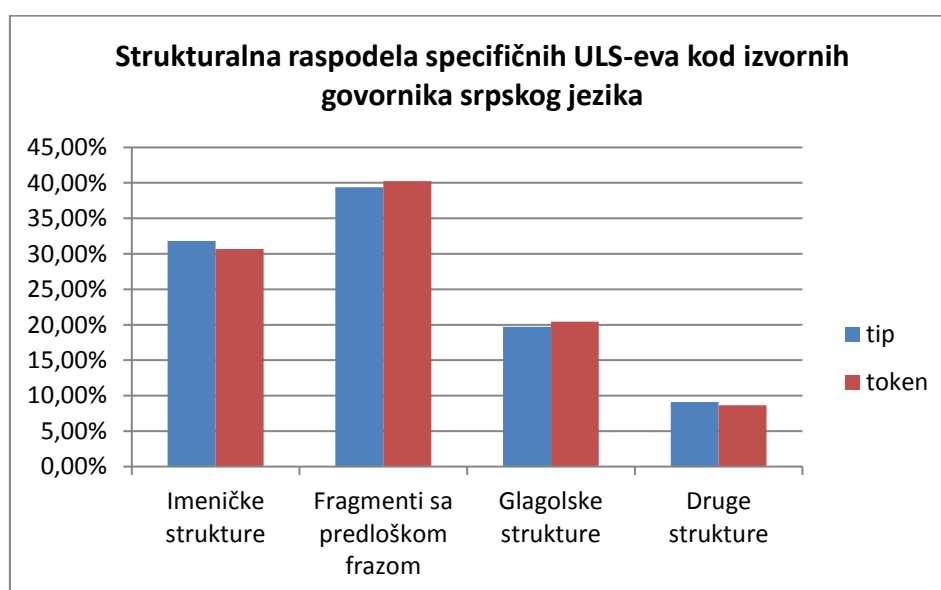
zapravo specifično za izvorne govornike srpskog jezika, v. Tabelu 99. Pošto u korpusu ECEP imamo samo jedan ULS koji pripada strukturalnoj kategoriji *druge strukture*, bilo je logično da će veliki procenat ULS iz ove kategorije biti specifični za izvorne govornike srpskog jezika, što je i potvrđeno podatkom da 75% svih ULS u kategoriji *druge strukture* (preostalih 25% dele obe grupe neizvornih govornika) sačinjavaju specifični ULS za srpski jezik. Na osnovu navedenih podataka možemo uvideti da neizvorni govornici engleskog jezika kojima je maternji jezik srpski imaju veću tendenciju da koriste ULS, koji su nekarakteristični za upotrebu kod izvornih govornika, u strukturalnim kategorijama *glagolske strukture* i *druge strukture*. Izvorni govornici srpskog imaju najmanju tendenciju da koriste specifične ULS u kategoriji *fragmenti sa predloškom frazom* – 42% od svih ULS u toj kategoriji predstavljaju specifične ULS za SCEP korpus. Poređenjem podataka, takođe, može da se uvidi da je zastupljenost specifičnih tokena manja od zastupljenosti specifičnih tipova u SCEP korpusu, što je posebno izraženo za strukturalne grupe *glagolske strukture* i *druge strukture*. Taj podatak nam ukazuje na činjenicu da, iako postoji veći broj tipova, oni predstavljaju manji procenat konkretnih ULS upotrebljenih u korpusu.

Tabela 98. *Strukturalna raspodela specifičnih ULS kod izvornih govornika srpskog jezika*

Strukturalna kategorija	Tip (Token)	Procenti
<i>Imeničke strukture</i>	21 (626)	31.82 (30.67)
<i>Fragmenti sa predloškom frazom</i>	26 (821)	39.39 (40.23)
<i>Glagolske strukture</i>	13 (417)	19.70 (20.43)
<i>Druge strukture</i>	6 (177)	9.09 (8.67)

Tabela 99. *Zastupljenost specifičnih ULS u SCEP korpusu po strukturalnim grupama – tipovi i tokeni*

	SCEP svi ULS	SCEP specifični ULS	%
<i>Imeničke strukture</i>	36 (1298)	21 (626)	58.33 (48.23)
<i>Fragmenti sa predloškom frazom</i>	50 (2031)	26 (821)	52 (40.42)
<i>Glagolske strukture</i>	17 (728)	13 (417)	76.47 (57.28)
<i>Druge strukture</i>	8 (479)	6 (177)	75 (36.95)



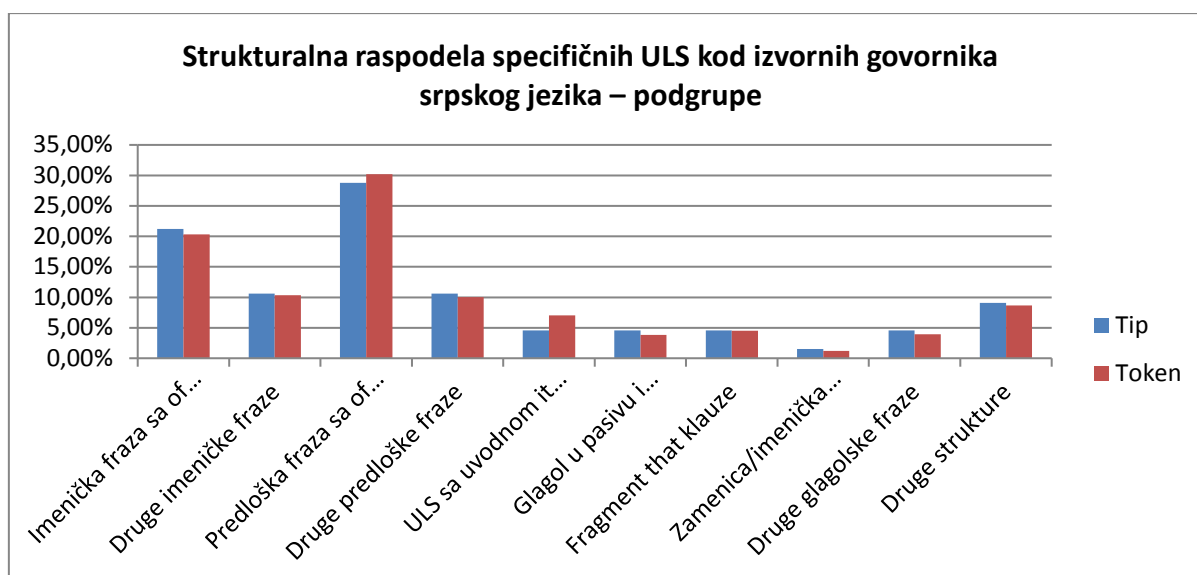
Grafikon 53. *Strukturalna raspodela specifičnih ULS kod izvornih govornika srpskog jezika*

Analizom strukturalnih podgrupa (v. Tabelu 100 i Grafikon 54) uvideli smo da je najzastupljenija podgrupa *predloška fraza sa of fragmentom* (28,79%) praćena *imeničkim frazama sa of fragmentom* (21,21%).

Tabela 100. *Strukturalna raspodela specifičnih ULS kod izvornih govornika srpskog jezika – podgrupe*

Strukturalna potkategorija	Tip (Token)	Procenti
<i>Strukturalna potkategorija</i>	14 (415)	21.21 (20.33)
<i>Imenička fraza sa of fragmentom</i>	7 (211)	10.61 (10.34)
<i>Druge imeničke fraze</i>	19 (616)	28.79 (30.18)
<i>Predloška fraza sa of fragmentom</i>	7 (205)	10.61 (10.04)
<i>Druge predloške fraze</i>	3 (143)	4.55 (7.01)
<i>ULS sa uvodnom it konstrukcijom</i>	3 (78)	4.55 (3.82)
<i>Glagol u pasivu i fragment predloške fraze</i>	3 (92)	4.55 (4.51)
<i>(Glagolska fraza +) fragment that klauze</i>	1 (24)	1.52 (1.18)
<i>Zamenica/imenička fraza+be (+...)</i>	3 (80)	4.55 (3.92)
<i>Druge glagolske fraze</i>	6 (177)	9.09 (8.67)

Najmanje zastupljena je strukturalna podgrupa *zamenica/imenička fraza + be (+...)* sa 1,52% svih specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski. Od ukupno 12 strukturalnih podgrupa dve nisu zastupljene među ULS koji su specifični za izvorne govornike srpskog: *(glagol/pridev+) fragment that klauze* i *kopula be i imenička/pridevska fraza*. Međutim, podgrupa *(glagol/pridev+) fragment that klauze* nije zastupljena ni u celom korpusu SCEP, što dalje znači da su specifični ULS prisutni u svim strukturalnim podgrupama u SCEP korpusu osim u podgrupi *kopula be i imenička/pridevska fraza*, koja je zastupljena sa samo jednim primerom u SCEP korpusu. U tri podgrupe *glagolskih struktura* sa malim brojem tipova u SCEP, *glagol u pasivu i fragment predloške fraze*, *zamenica/imenička fraza+be (+...)* i *druge glagolske fraze*, svi primeri su specifični za produkciju izvornih govornika srpskog jezika. Na osnovu ovog podatka nameće se zaključak da bi pri podučavanju akademskog engleskog u oblasti obrazovanja trebalo detaljno obraditi ULS koje izvorni govornici koriste u strukturalnim podgrupama: *glagol u pasivu i fragment predloške fraze*, *zamenica/imenička fraza+be (+...)* i *druge glagolske fraze*, kako studenti ne bi koristili ULS koji nisu najoptimalnija opcija za pokazivanje pripadnosti u datoj delatnoj zajednici.



Grafikon 54. *Strukturalna raspodela specifičnih ULS kod izvornih govornika srpskog jezika – podgrupe*

Poređenje strukturalne raspodele specifičnih ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski

Kako bi se ispitala potencijalne razlike, tj. asocijacije između ECEP i SCEP korpusa u odnosu na strukturalne distribucije specifičnih ULS, korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između SCEP i ECEP korpusa u odnosu na strukturalne grupe specifičnih ULS – tipovi. – Unakrsna tabela između SCEP i ECEP korpusa i strukturalnih grupa u odnosu na tipove specifičnih ULS prikazana je u Tabeli 101. Hi-kvadrat test nije bio značajan, $\chi^2(3, N = 78) = 0.86, p > .050$, ukazujući da nije identifikovana statistički značajna povezanost između korpusa i strukturalnih grupa.

Tabela 101. *Unakrsna tabela između SCEP i ECEP korpusa i strukturalnih grupa u odnosu na tipove specifičnih ULS*

Grupe	SCEP			ECEP		
	tipovi	%	R	tipovi	%	R
<i>Imeničke strukture</i>	21	30.4%	-0.2	3	33.3%	0.2
<i>Fragmenti sa predložkom frazom</i>	29	42.0%	-0.1	4	44.4%	0.1
<i>Glagolske strukture</i>	13	18.8%	-0.2	2	22.2%	0.2
<i>Druge strukture</i>	6	8.7%	0.9	0	0%	-0.9
Total	69	100.0%		9	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Nedostatak statističke značajnosti pri poređenju strukturalne raspodele specifičnih tipova ULS u ECEP i SCEP korpusima ukazuje na to da izvorni govornici srpskog jezika u produkciji na engleskom imaju tendenciju da zanemare određene tipove ULS iz svih strukturalnih kategorija, kao i da koriste atipične tipove ULS u svojoj produkciji u svim strukturalnim kategorijama. Raspodele

zastupljenosti su slične kao što su u ECEP i SCEP korpusima u celini, odnosno najviše imamo specifičnih ULS u kategoriji *fragmenti sa predloškom frazom*, koja je ujedno i najzastupljenija strukturalna grupa u ECEP i SCEP korpusima u celini.

Tabela 102. *Unakrsna tabela između SCEP i ECEP korpusa i strukturalnih grupa u odnosu na tokene specifičnih ULS*

Grupe	SCEP			ECEP		
	tokeni	%	R	tokeni	%	R
Imeničke strukture	626	29.0%	-1.1	72	32.4%	1.1
Fragmenti sa predloškom frazom	941	43.5%	0.2	95	42.8%	-0.2
Glagolske strukture	417	19.3%	-2.0*	55	24.8%	2.0*
Druge strukture	177	8.2%	4.4*	0	0%	-4.4*
Total	2161	100%		222	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

* – statistički značajni rezultati

Razlike između SCEP i ECEP korpusa u odnosu na strukturalne grupe specifičnih ULS – tokeni. – Unakrsna tabela između ECEP i SCEP korpusa i strukturalnih grupa u odnosu na tokene specifičnih ULS prikazana je u Tabeli 102. Hi-kvadrat test je bio značajan, $\chi^2(3, N = 2383) = 22.08, p < .001$, ukazujući da postoji statistički značajna povezanost između korpusa i strukturalnih grupa.

Korigovani standardizovani reziduali ukazuju da je u ECEP korpusu veća upotreba *glagolskih struktura*, dok je u SCEP korpusu veća upotreba *drugih struktura*. Dobijeni nalaz dalje ukazuje na to da neizvorni govornici engleskog kojima je maternji jezik srpski imaju tendenciju da prekomerno koriste ULS *drugih struktura* a zanemare u produkciji relevantne ULS sa *glagolskom strukturom*. Kako bi se produkcija neizvornih govornika srpskog jezika unapredila, trebalo bi sagledati koja se značenja prenose specifičnim ULS iz kategorije *druge strukture* i kako se ona prenose u produkciji izvornih govornika, te da se u nastavi tome posveti posebna pažnja. Takođe, potrebno je podučavati ULS koji su karakteristični za izvorne govornike engleskog jezika sa posebnim akcentom na *glagolske strukture* kod kojih postoji najveća tendencija zanemarivanja u upotrebi kod izvornih govornika srpskog jezika.

Razlike između ECEP i SCEP korpusa u odnosu na strukturalne podgrupe specifičnih ULS – tipovi. – Unakrsna tabela između ECEP i SCEP korpusa i strukturalnih podgrupa u odnosu na tipove specifičnih ULS prikazana je u Tabeli 103. Hi-kvadrat test nije bio značajan, $\chi^2(10, N = 78) = 10.49, p > .050$, ukazujući na to da nije identifikovana statistički značajna povezanost između korpusa i strukturalnih podgrupa. Nedostatak statistički značajnih razlika u strukturalnoj raspodeli specifičnih ULS u SCEP i ECEP korpusima navodi na zaključak da se nijedna strukturalna podgrupa ne ističe pri analizi tipova u dva navedena korpusa.

Tabela 103. Unakrsna tabela između SCEP i ECEP korpusa i strukturalnih podgrupa specifičnih ULS u odnosu na tipove

Grupa	Podgrupa	SCEP			ECEP		
		tipovi	%	R	tipovi	%	R
Imeničke strukture	Imenička fraza sa of fragmentom	14	20.3%	-0.1	2	22.2%	0.1
	Druge imeničke fraze	7	10.1%	-0.1	1	11.1%	0.1
Fragmenti sa predložkom frazom	Predložka fraza sa of fragmentom	22	31.9%	-0.1	3	33.3%	0.1
	Druge predložke fraze	7	10.1%	-0.1	1	11.1%	0.1
Glagolske strukture	Uvodna it struktura + glagolska ili pridevska fraza	3	4.3%	-0.9	1	11.1%	0.9
	Glagol u pasivu i fragment predložke fraze	3	4.3%	0.6	0	0%	-
	Kopula be i imenička/pridevska fraza	/	/	/	/	/	/
	Fragment that klauze	3	4.3%	0.6	0	0%	-
	(Glagol/pridev +) fragment to klauze	0	0%	-2.8	1	11.1%	2.8
	Fragment adverbijalne klauze	/	/	/	/	/	/
	Zamenica/imenička fraza+be (+...)	1	1.4%	0.4	0	0%	-
	Druge glagolske fraze	3	4.3%	0.6	0	0%	-
	Druge strukture	6	8.7%	0.9	0	0%	-
	Total		69	100.0%		9	100.0%

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između korpusa ECEP i SCEP u odnosu na strukturalne podgrupe specifičnih ULS – tokeni. – Unakrsna tabela između SCEP i ECEP korpusa i strukturalnih podgrupa specifičnih ULS u odnosu na tokene prikazana je u Tabeli 104. Hi-kvadrat test je bio značajan, $\chi^2(10, N = 2383) = 385.16$, $p < .001$, što ukazuje na to da postoji statistički značajna povezanost između korpusa i strukturalnih podgrupa.

Korigovani standardizovani reziduali ukazuju da je u SCEP korpusu povećana upotreba podgrupa *glagol u pasivu i fragment predložke fraze*, *fragment that klauze*, *druge glagolske fraze* i *druge strukture*, dok je u ECEP korpusu izraženija upotreba podgrupe *(glagol/pridev +) fragment to klauze*. Navedeni rezultati testa statističke značajnosti dalje ukazuju na to da izvorni govornici srpskog jezika imaju tendenciju da isuviše učestalo koriste ULS iz četiri navedenih strukturalnih podgrupa, te da je neophodno u nastavi obratiti posebnu pažnju na podučavanje adekvatnih ULS iz navedenih kategorija, kao i alternativne prikladnije načine/strukture za izražavanje istih značenja. S obzirom na činjenicu da se podgrupa *(glagol/pridev +) fragment to klauze* značajno češće koristi u ECEP korpusu, dolazimo do zaključka da izvorni govornici srpskog jezika imaju tendenciju da zanemaruju upotrebu prikladnih ULS iz navedene kategorije. Predlaže se da se posebna pažnja posveti podučavanju ovih konstrukcija u radu sa studentima/prevodiocima kojima je maternji jezik srpski.

Tabela 104. Unakrsna tabela između ECEP i SCEP korpusa i strukturalnih podgrupa specifičnih ULS u odnosu na tokene

Grupa	Podgrupa	SCEP			ECEP		
		tokeni	%	R	tokeni	%	R
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	415	19.2%	-0.9	48	21.6%	0.9
	<i>Druge imeničke fraze</i>	211	9.8%	-0.5	24	10.8%	0.5
<i>Fragmenti sa predloškom frazom</i>	<i>Predloška fraza sa of fragmentom</i>	736	34.1%	1.4	65	29.3%	-1.4
	<i>Druge predloške fraze</i>	205	9.5%	-1.9	30	13.5%	1.9
<i>Glagolske strukture</i>	<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	143	6.6%	-1.6	21	9.5%	1.6
	<i>Glagol u pasivu i fragment predloške fraze</i>	78	3.6%	2.9*	0	0%	-2.9*
	<i>Kopula be i imenička/pridevska fraza</i>	/	/	/	/	/	/
	<i>Fragment that klauze</i>	92	4.3%	3.1*	0	0%	-3.1*
	<i>(Glagol/pridev +) fragment to klauze</i>	0	0%	-18.3*	34	15.3%	18.3*
	<i>Fragment adverbijalne klauze</i>	/	/	/	/	/	/
	<i>Zamenica/imenička fraza+be (+...)</i>	24	1.1%	1.6	0	0%	-1.6
	<i>Druge glagolske fraze</i>	80	3.7%	2.9*	0	0%	-2.9*
	<i>Druge strukture</i>	177	8.2%	4.4*	0	0%	-4.4*
	Total		2161	100.0%		222	100.0%

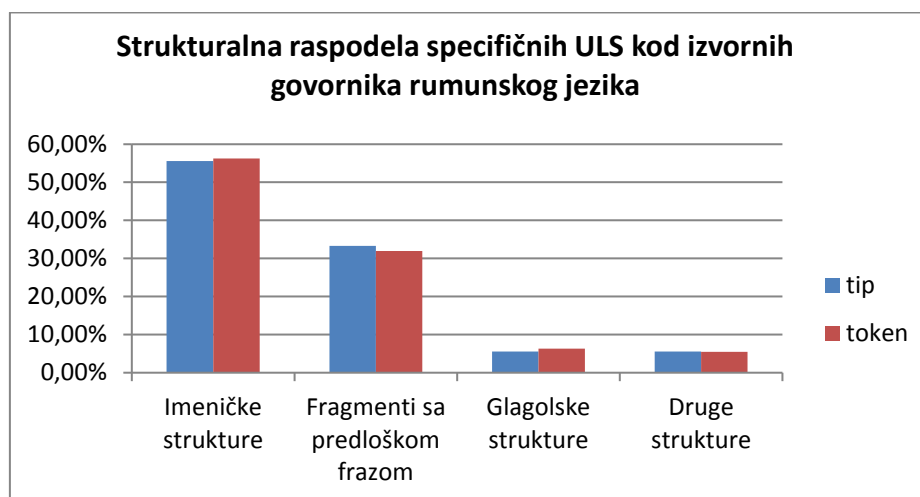
Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Strukturalna raspodela specifičnih ULS kod izvornih govornika engleskog kojima je maternji jezik rumunski jezik

Strukturalna analiza specifičnih ULS kod izvornih govornika rumunskog jezika pokazala je da su najzastupljenija kategorija *imeničke strukture* praćene *fragmentima sa predloškom frazom*. *Glagolske strukture* i *druge strukture* su zastupljene sa samo jednim tipom ULS i sačinjavaju po 5,56% tipova u korpusu, v. Tabelu 105 i Grafikon 55. Ovakva raspodela strukturalnih kategorija se razlikuje od strukturalne raspodele u celom RCEP korpusu, u kojem je najzastupljenija strukturalna kategorija *fragmenti sa predloškom frazom*, praćene *imeničkim frazama*. Kada se sagleda zastupljenost specifičnih ULS u celom RCEP korpusu, uviđamo da je 41,67% svih *imeničkih struktura* u RCEP korpusu specifično za govornike rumunskog jezika, što je ujedno i najveći procenat zastupljenosti specifičnih ULS u bilo kojoj strukturalnoj kategoriji u RCEP korpusu, v. Tabelu 106. U korpusu izvornih govornika rumunskog jezika manji je procenat zastupljenosti specifičnih tokena uporedno sa specifičnim tipovima, što nam ukazuje na to da predstavljaju manji broj konkretnih upotreba uporedno sa količinom raznovrsnosti konstrukcija.

Tabela 105. *Strukturalna raspodela specifičnih ULS kod izvornih govornika rumunskog jezika*

Strukturalna kategorija	Tip (Token)	Procenti
<i>Imeničke strukture</i>	10 (257)	55.56 (56.24)
<i>Fragmenti sa predloškom frazom</i>	6 (146)	33.33 (31.95)
<i>Glagolske strukture</i>	1 (29)	5.56 (6.35)
<i>Druge strukture</i>	1 (25)	5.56 (5.47)



Grafikon 55. *Strukturalna raspodela specifičnih ULS kod izvornih govornika rumunskog jezika*

Tabela 106. *Zastupljenost specifičnih ULS u RCEP korpusu po strukturalnim grupama – tipovi i tokeni*

	RCEP svi ULS	RCEP specifični ULS	%
<i>Imeničke strukture</i>	24 (707)	10 (257)	41.67 (36.35)
<i>Fragmenti sa predloškom frazom</i>	28 (1305)	6 (146)	21.47 (11.19)
<i>Glagolske strukture</i>	5 (151)	1 (29)	20 (19.21)
<i>Druge strukture</i>	3 (132)	1 (25)	33.33 (18.94)

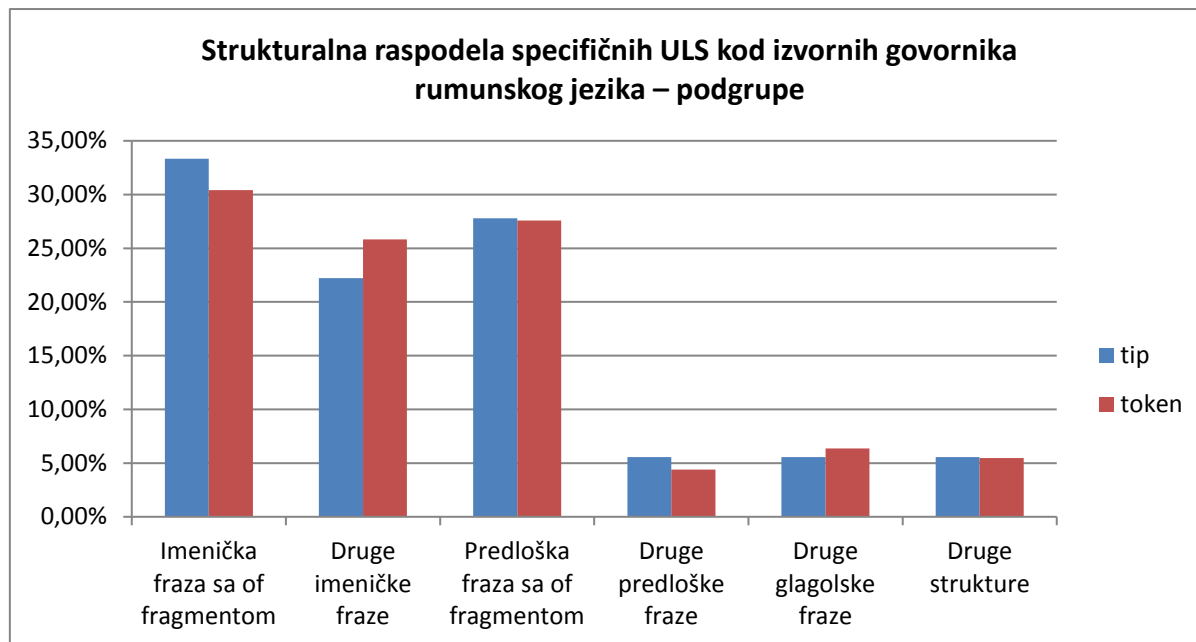
Analiza strukturalnih podgrupa specifičnih ULS za neizvorne govornike engleskog kojima je maternji jezik rumunski istakla je *imeničku frazu sa of fragmentom* kao najzastupljeniju strukturalnu podgrupu (sačinjava 33,33% svih specifičnih tipova). Druga najzastupljenija podgrupa je *predloška fraza sa of fragmentom* (27,78%) praćena podgrupom *druge imeničke fraze* (22,22%). *Druge predloške fraze* i *druge strukture* zastupljene su sa po samo jednim tipom, v. Tabelu 107 i Grafikon 56.

Tabela 107. *Strukturalna raspodela specifičnih ULS kod izvornih govornika rumunskog jezika – podgrupe*

Strukturalna potkategorija	Tip (Token)	Procenti
<i>Imenička fraza sa of fragmentom</i>	6 (139)	33.33 (30.42)
<i>Druge imeničke fraze</i>	4 (118)	22.22 (25.82)
<i>Predloška fraza sa of fragmentom</i>	5 (126)	27.78 (27.57)

Tabela 107. *Strukturalna raspodela specifičnih ULS kod izvornih govornika rumunskog jezika – podgrupe*

<i>Druge predložke fraze</i>	1 (20)	5.56 (4.38)
<i>Druge glagolske fraze</i>	1 (29)	5.56 (6.35)
<i>Druge strukture</i>	1 (25)	5.56 (5.47)



Grafikon 56. *Strukturalna raspodela specifičnih ULS kod izvornih govornika rumunskog jezika – podgrupe*

Od ukupno 12 strukturalnih podgrupa, šest je zastupljeno među specifičnim ULS u RCEP korpusu. Svih šest strukturalnih podgrupa koje nisu zastupljene među specifičnim ULS za izvorne govornike rumunskog (*uvodna it struktura + glagolska ili pridevska fraza, kopula be i imenička/pridevska fraza, fragment that klauze, /glagol/pridev +/- fragment to klauze, fragment adverbijalne klauze, zamenica/imenička fraza+be /+.../, glagol u pasivu i fragment predložke fraze*) nalaze se u strukturalnoj kategoriji *glagolske strukture*. Ovo nam dalje ukazuje na to da neizvorni govornici engleskog kojima je maternji jezik rumunski nemaju tendenciju da koriste specifične ULS iz strukturalnih grupa i podgrupa *glagolskih struktura* te da pri podučavanju nije potrebno akcentovati ove konstrukcije.

Poređenje strukturalne raspodele specifičnih ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski jezik

Kako bi se ispitale potencijalne razlike, tj. asocijacije između ECEP i RCEP korpusa u odnosu na strukturalne distribucije specifičnih ULS korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između korpusa ECEP i RCEP u odnosu na strukturalne grupe specifičnih ULS – tipovi. – Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih grupa u odnosu na tipove specifičnih ULS prikazana je u Tabeli 108. Hi-kvadrat test nije bio značajan, $\chi^2(3, N = 27) = 2.81, p > .050$, ukazujući da nije identifikovana statistički značajna povezanost između korpusa i strukturalnih

grupa. Navedeni rezultati dalje ukazuju na to da uočene razlike u strukturalnim raspodelama tipova u dva ispitana korpusa nisu statistički značajne.

Tabela 108. *Unakrsna tabela između RCEP i ECEP korpusa i strukturalnih grupa specifičnih ULS u odnosu na tipove*

Grupe	RCEP			ECEP		
	tipovi	%	R	tipovi	%	R
<i>Imeničke strukture</i>	10	55.6%	1.1	3	33.3%	-1.1
<i>Fragmenti sa predloškom frazom</i>	6	33.3%	-0.6	4	44.4%	0.6
<i>Glagolske strukture</i>	1	5.6%	-1.3	2	22.2%	1.3
<i>Druge strukture</i>	1	5.6%	0.7	0	0%	-0.7
Total	18	100.0%		9	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između korpusa ECEP i RCEP u odnosu na strukturalne grupe specifičnih ULS – tokeni. – Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih grupa specifičnih ULS u odnosu na tokene prikazana je u Tabeli 109. Hi-kvadrat test je bio značajan, $\chi^2(3, N = 679) = 75.59, p < .001$, ukazujući na to da postoji statistički značajna povezanost između korpusa i strukturalnih grupa.

Tabela 109. *Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih grupa specifičnih ULS u odnosu na tokene*

Grupe	RCEP			ECEP		
	tokeni	%	R	tokeni	%	R
<i>Imeničke strukture</i>	257	56.2%	5.8*	72	32.4%	-5.8*
<i>Fragmenti sa predloškom frazom</i>	146	31.9%	-2.8*	95	42.8%	2.8*
<i>Glagolske strukture</i>	29	6.3%	-6.8*	55	24.8%	6.8*
<i>Druge strukture</i>	25	5.5%	3.6*	0	0%	-3.6*
Total	457	100%		222	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Korigovani standardizovani reziduali ukazuju na višu upotrebu *imeničkih i drugih struktura* u RCEP korpusu, dok se u ECEP korpusu više koriste *fragmenti sa predloškom frazom* i *glagolske strukture*. Navedeni podaci nam ukazuju na to da izvorni govornici rumunskog jezika imaju tendenciju da značajno više koriste njima specifične ULS u strukturalnim grupama *imeničke strukture* i *druge strukture*. To dalje implicira da je potrebno ispitati koja se značenja prenose ovim konstrukcijama i u nastavi podučiti studente o prikladnijim i raznovrsnijim načinima prenošenja tih značenja. Statistička analiza podataka, takođe, navodi na zaključak da izvorni govornici rumunskog jezika statistički značajno više zanemaruju upotrebu ULS iz strukturalnih kategorija *fragmenti sa predloškom frazom* i *glagolske strukture*. Stoga je neophodno posvetiti posebnu pažnju ovim kategorijama pri podučavanju izvornih govornika rumunskog jezika.

Razlike između korpusa ECEP i RCEP u odnosu na strukturalne podgrupe – tipovi. – Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih podgrupa specifičnih ULS u odnosu na tipove prikazana je u Tabeli 110. Hi-kvadrat test nije bio značajan, $\chi^2(7, N = 27) = 5.96, p > .050$, što ukazuje na to da nije identifikovana statistički značajna povezanost između korpusa i strukturalnih podgrupa.

Tabela 110. Unakrsna tabela između korpusa i strukturalnih podgrupa u odnosu na tipove

Grupe	Podgrupa	RCEP			ECEP		
		tipovi	%	R	tipovi	%	R
Imeničke strukture	Imenička fraza sa of fragmentom	6	33.3%	0.6	2	22.2%	-0.6
	Druge imeničke fraze	4	22.2%	0.7	1	11.1%	-0.7
Fragmenti sa predloškom frazom	Predloška fraza sa of fragmentom	5	27.8%	-0.3	3	33.3%	0.3
	Druge preloške fraze	1	5.6%	-0.5	1	11.1%	0.5
Glagolske strukture	Uvodna it struktura + glagolska ili pridevska fraza	0	0%	-1.4	1	11.1%	1.4
	Glagol u pasivu i fragment predloške fraze	/	/	/	/	/	/
	Kopula be i imenička/pridevska fraza	/	/	/	/	/	/
	Fragment that klauze	/	/	/	/	/	/
	(Glagol/pridev +) fragment to klauze	0	0%	-1.4	1	11.1%	1.4
	Fragment adverbijalne klauze	/	/	/	/	/	/
	Zamenica/imenička fraza+be (+...)	/	/	/	/	/	/
	Druge glagolske fraze	1	5.6%	0.7	0	0%	-0.7
Druge strukture	1	5.6%	0.7	0	0%	-0.7	
Total		18	100.0%		9	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

* – statistički značajni rezultati

Razlike između ECEP i RCEP korpusa u odnosu na strukturalne podgrupe specifičnih ULS – tokeni. – Unakrsna tabela između ECEP i RCEP korpusa i strukturalnih podgrupa specifičnih ULS u odnosu na tokene prikazana je u Tabeli 111.

Tabela 111. Unakrsna tabela između RCEP i SCEP korpusa i strukturalnih podgrupa specifičnih ULS u odnosu na tokene

Grupe	Podgrupa	RCEP			ECEP		
		tokeni	%	R	tokeni	%	R
Imeničke strukture	Imenička fraza sa of fragmentom	139	30.4%	2.4*	48	21.6%	-2.4*
	Druge imeničke fraze	118	25.8%	4.5*	24	10.8%	-4.5*
Fragmenti sa predloškom frazom	Predloška fraza sa of fragmentom	126	27.6%	-0.5	65	29.3%	0.5
	Druge preloške fraze	20	4.4%	-4.3*	30	13.5%	4.3*
Glagolske strukture	Uvodna it struktura + glagolska ili pridevska fraza	0	0%	-6.7*	21	9.5%	6.7*
	Glagol u pasivu i fragment predloške fraze	/	/	/	/	/	/
	Kopula be i imenička/pridevska fraza	/	/	/	/	/	/
	Fragment that klauze	/	/	/	/	/	/
		/	/	/	/	/	/

Tabela 111. Unakrsna tabela između RCEP i SCEP korpusa i strukturalnih podgrupa specifičnih ULS u odnosu na tokene

	<i>(Glagol/pridev +) fragment to klauze</i>	0	0%	-8.6*	34	15.3%	8.6*
	<i>Fragment adverbijalne klauze</i>	/	/	/	/	/	/
	<i>Zamenica/imenička fraza+be (+...)</i>	/	/	/	/	/	/
	<i>Druge glagolske fraze</i>	29	6.3%	3.8*	0	0%	-3.8*
<i>Druge strukture</i>	<i>Druge strukture</i>	25	5.5%	3.6*	0	0%	-3.6*
Total		457	100.0%		222	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali

* – statistički značajni rezultati

Hi-kvadrat test je bio značajan, $\chi^2(7, N = 679) = 176.84$, $p < .001$, ukazujući na to da postoji statistički značajna povezanost između korpusa i strukturalnih podgrupa.

U RCEP korpusu više se koriste obe podgrupe u *imeničkim strukturama*, kao i podgrupe *druge glagolske fraze* i *druge strukture*, dok je u ECEP korpusu viša upotreba podgrupa *druge predloške fraze*, *uvodna it struktura + glagolska ili pridevska fraza* i *(glagol/pridev +) fragment to klauze*. Podaci strukturalne analize dalje ukazuju da izvorni govornici rumunskog jezika imaju tendenciju da koriste značajno više specifičnih ULS u četiri strukturalne potkategorije koje su prethodno navedene. Kako bi se ovakve tendencije u upotrebi izbegle, potrebno je ispitati koja se značenja prenose navedenim ULS i u radu sa studentima obraditi adekvatnija rešenja. Rezultati, takođe, ukazuju na to da je u radu sa studentima/prevodiocima kojima je izvorni jezik rumunski neophodno posvetiti posebnu pažnju obradi ULS iz strukturalnih podgrupa *druge predloške fraze*, *uvodna it struktura + glagolska ili pridevska fraza* i *(glagol/pridev +) fragment to klauze*, zato što imaju povišenu tendenciju da ih zanemare.

Poređenje strukturalne raspodele specifičnih ULS kod izvornih govornika srpskog jezika i izvornih govornika rumunskog jezika u njihovoj produkciji na engleskom jeziku

Kako bi se ispitale potencijalne razlike, tj. asocijacije između SCEP i RCEP korpusa u odnosu na strukturalne distribucije specifičnih ULS korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između SCEP i RCEP korpusa u odnosu na strukturalne grupe specifičnih ULS – tipovi. – Kada uporedimo procentualnu zastupljenost specifičnih ULS po strukturalnim kategorijama u SCEP i RCEP korpusima možemo uvideti nekoliko razlika. U SCEP korpusu specifični ULS sačinjavaju najveći procenat tipova u kategorijama *glagolske strukture* i *druge strukture*, dok su najmanje zastupljeni u kategorijama *fragmenti sa predloškom frazom* i *imeničke strukture*. U RCEP korpusu situacija je suprotna od zabeleženog u SCEP korpusu a kada je reč o procentualnoj zastupljenosti specifičnih ULS u strukturalnim kategorijama. Naime, u RCEP korpusu specifični ULS sačinjavaju najveći procenat tipova u kategoriji *imeničke strukture* a najmanje u kategorijama *glagolske strukture* i *fragmenti sa predloškom frazom*, v. Tabelu 112. Na osnovu ovih podataka možemo uvideti različite tendencije različitih grupa neizvornih govornika ka prenaplašenju upotrebi specifičnih ULS u određenim strukturalnim grupama. Odnosno, uočavamo tendenciju izvornih govornika srpskog da najčešće koriste specifične ULS u strukturalnoj kategoriji *glagolske*

strukture i tendenciju izvornih govornika rumunskog da najčešće koriste specifične ULS u strukturalnoj kategoriji *imeničke strukture*.

Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih grupa specifičnih ULS u odnosu na tipove prikazana je u Tabeli 112. Hi-kvadrat test nije bio značajan, $\chi^2(3, N = 87) = 4.54$, $p > .050$, ukazujući na to da nije identifikovana statistički značajna povezanost između korpusa i strukturalnih grupa. Drugim rečima, razlike koje uvidamo u strukturalnoj raspodeli specifičnih ULS na tipove nisu statistički značajne.

Tabela 112. Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih grupa specifičnih ULS u odnosu na tipove

Grupe	SCEP			RCEP		
	tipovi	%	R	tipovi	%	R
<i>Imeničke strukture</i>	21	30.4%	-2.0	10	55.6%	2.0
<i>Fragmenti sa predloškom frazom</i>	29	42.0%	0.7	6	33.3%	-0.7
<i>Glagolske strukture</i>	13	18.8%	1.4	1	5.6%	-1.4
<i>Druge strukture</i>	6	8.7%	0.4	1	5.6%	-0.4
Total	69	100.0%		18	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između SCEP i RCEP korpusa u odnosu na strukturalne grupe specifičnih ULS – tokeni. – Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih grupa specifičnih ULS u odnosu na tokene prikazana je u Tabeli 113. Hi-kvadrat test je bio značajan, $\chi^2(3, N = 2618) = 136.14$, $p < .001$; ukazao je da postoji statistički značajna povezanost između korpusa i strukturalnih grupa.

Tabela 113. Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih grupa specifičnih ULS u odnosu na tokene

Grupe	SCEP			RCEP		
	tokeni	%	R	tokeni	%	R
<i>Imeničke strukture</i>	626	29.0%	-11.2*	257	56.2%	11.2*
<i>Fragmenti sa predloškom frazom</i>	941	43.5%	4.6*	146	31.9%	-4.6*
<i>Glagolske strukture</i>	417	19.3%	6.7*	29	6.3%	-6.7*
<i>Druge strukture</i>	177	8.2%	2.0*	25	5.5%	-2.0*
Total	2161	100%		457	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Korigovani standardizovani reziduali ukazuju da je u RCEP korpusu značajno viša upotreba *imeničkih struktura*, dok se ostale strukturalne grupe više koriste u SCEP korpusu. Navedeni rezultati ukazuju na to da izvorni govornici rumunskog jezika imaju statistički značajnu tendenciju da koriste više specifičnih ULS iz strukturalne kategorije *imeničke strukture*, dok izvorni govornici srpskog jezika imaju statistički značajniju tendenciju da više koriste specifične ULS u svim ostalim strukturalnim grupama.

Razlike između SCEP i RCEP korpusa u odnosu na strukturalne podgrupe specifičnih ULS – tipovi. – Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih podgrupa specifičnih ULS u odnosu na tipove prikazana je u Tabeli 114. Hi-kvadrat test nije bio značajan, $\chi^2(9, N = 87) = 5.94$, $p > .050$; ukazao je na to da nije identifikovana statistički značajna povezanost između

korpusa i strukturalnih podgrupa, što dalje implicira da uočene razlike u strukturalnoj raspodeli tipova između dva korpusa nisu statistički značajne.

Tabela 114. *Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih podgrupa specifičnih ULS u odnosu na tipove*

Grupe	Podgrupa	SCEP			RCEP		
		tipovi	%	R	tipovi	%	R
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	14	20.3%	-1.2	6	33.3%	1.2
	<i>Druge imeničke fraze</i>	7	10.1%	-1.4	4	22.2%	1.4
<i>Fragmenti sa predloškom frazom</i>	<i>Predloška fraza sa of fragmentom</i>	22	31.9%	0.3	5	27.8%	-0.3
	<i>Druge preloške fraze</i>	7	10.1%	0.6	1	5.6%	-0.6
<i>Glagolske strukture</i>	<i>Uvodna it struktura + glagolska ili pridevska fraza</i>	3	4.3%	0.9	0	0%	-0.9
	<i>Glagol u pasivu i fragment predloške fraze</i>	3	4.3%	-0.9	0	0%	-0.9
	<i>Kopula be i imenička/pridevska fraza</i>	/	/	/	/	/	/
	<i>Fragment that klauze</i>	3	4.3%	0.9	0	0%	-0.9
	<i>(Glagol/pridev +) fragment to klauze</i>	/	/	/	/	/	/
	<i>Fragment adverbijalne klauze</i>	/	/	/	/	/	/
	<i>Zamenica/imenička fraza+be (+...)</i>	1	1.4%	0.9	0	0%	-0.5
<i>Druge strukture</i>	<i>Druge glagolske fraze</i>	3	4.3%	0.5	1	5.6%	0.2
	<i>Druge strukture</i>	6	8.7%	-0.2	1	5.6%	-0.4
Total		69	100.0%		18	100.0%	

*Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali; * – statistički značajni rezultati*

Razlike između SCEP i RCEP korpusa u odnosu na strukturalne podgrupe specifičnih ULS – tokeni. – Unakrsna tabela između SCEP i RCEP korpusa i strukturalnih podgrupa specifičnih ULS u odnosu na tokene prikazana je u Tabeli 115. Hi-kvadrat test je bio značajan, χ^2 (9, N = 2618) = 197.29, $p < .001$, što ukazuje na to da postoji statistički značajna povezanost između korpusa i strukturalnih podgrupa.

Tabela 115. *Unakrsna tabela između korpusa i strukturalnih podgrupa u odnosu na tokene*

Grupe	Podgrupa	SCEP			RCEP		
		tokeni	%	R	tokeni	%	R
<i>Imeničke strukture</i>	<i>Imenička fraza sa of fragmentom</i>	415	19.2%	-5.3*	139	30.4%	5.3*
	<i>Druge imeničke fraze</i>	211	9.8%	-9.4*	118	25.8%	9.4*
<i>Fragmenti sa predloškom frazom</i>	<i>Predloška fraza sa of fragmentom</i>	736	34.1%	2.7*	126	27.6%	-2.7*
	<i>Druge preloške fraze</i>	205	9.5%	3.5*	20	4.4%	-3.5*

Tabela 115. Unakrsna tabela između korpusa i strukturalnih podgrupa u odnosu na tokene

Glagolske strukture	Uvodna it struktura + glagolska ili pridevska fraza	143	6.6%	5.7*	0	0%	-5.7*
	Glagol u pasivu i fragment predloške fraze	78	3.6%	4.1*	0	0%	-4.1*
	Kopula be i imenička/pridevska fraza	/	/	/	/	/	/
	Fragment that klauze	92	4.3%	4.5*	0	0%	-4.5*
	(Glagol/pridev +) fragment to klauze	/	/	/	/	/	/
	Fragment adverbijalne klauze	/	/	/	/	/	/
	Zamenica/imenička fraza+be (+...)	24	1.1%	2.3*	0	0%	-2.3*
	Druge glagolske fraze	80	3.7%	-2.6*	29	6.3%	2.6*
	Druge strukture	177	8.2%	2.0*	25	5.5%	-2.0*
	Total	2161	100.0%		457	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

* – statistički značajni rezultati

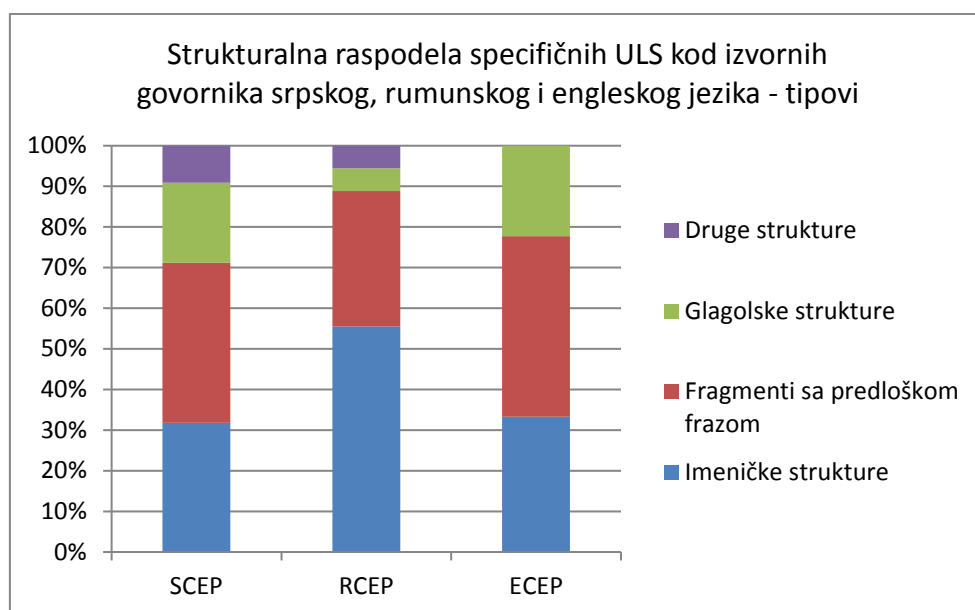
Analiza korigovanih standardizovanih reziduala ukazuje na to da je u SCEP korpusu intenzivnija upotreba *framenata sa predloškom frazom*, većine *glagolskih struktura* (*uvodna it struktura + glagolska ili pridevska fraza*, *glagol u pasivu i fragment predloške fraze*, *fragment that klauze*, *i zamenica/imenička fraza+be (+...)*) i *drugih struktura*, dok je u RCEP korpusu frekventnija upotreba *imeničkih struktura* i *drugih glagolskih fraza*. Analiza strukturalne raspodele tokena u podgrupe pokazala je da, pored razlika koje su se uvidele pri analizi strukturalnih grupa, izvorni govornici rumunskog jezika imaju tendenciju da učestalije koriste specifične ULS u podgrupi *druge glagolske fraze*.

Poređenje strukturalne raspodele specifičnih ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski i rumunski

Kada sagledamo strukturalne raspodele ULS na tipove u sva tri korpusa (v. Tabelu 116 i Grafikon 57), uvidamo da je strukturalna raspodela ULS između SCEP i ECEP korpusa slična, dok se strukturalna raspodela u RCEP korpusu razlikuje, pošto su tu *imeničke strukture* najzastupljenija grupa. Kao što smo videli iz prethodnih poglavlja, razlike u strukturalnim grupama i podgrupama kada je reč o tipovima nisu se pokazale kao statistički značajne. Ono što bi svakako trebalo da se uzme u obzir pri interpretaciji navedenih podataka jeste činjenica da se u testovima statističke značajnosti lakše dobiju statistički značajni rezultati kada su testovi primenjeni na većim brojkama. Tako da i sâm manji broj tipova uporedno sa tokenima može da utiče na dobijene rezultate.

Tabela 116. Strukturalna raspodela specifičnih ULS kod izvornih govornika srpskog, rumunskog i engleskog jezika – tipovi

	SCEP	RCEP	ECEP
Imeničke strukture	31.82%	55.56%	33.33%
Fragmenti sa predloškom frazom	39.39%	33.33%	44.44%
Glagolske strukture	19.70%	5.56%	22.22%
Druge strukture	9.09%	5.56%	0%

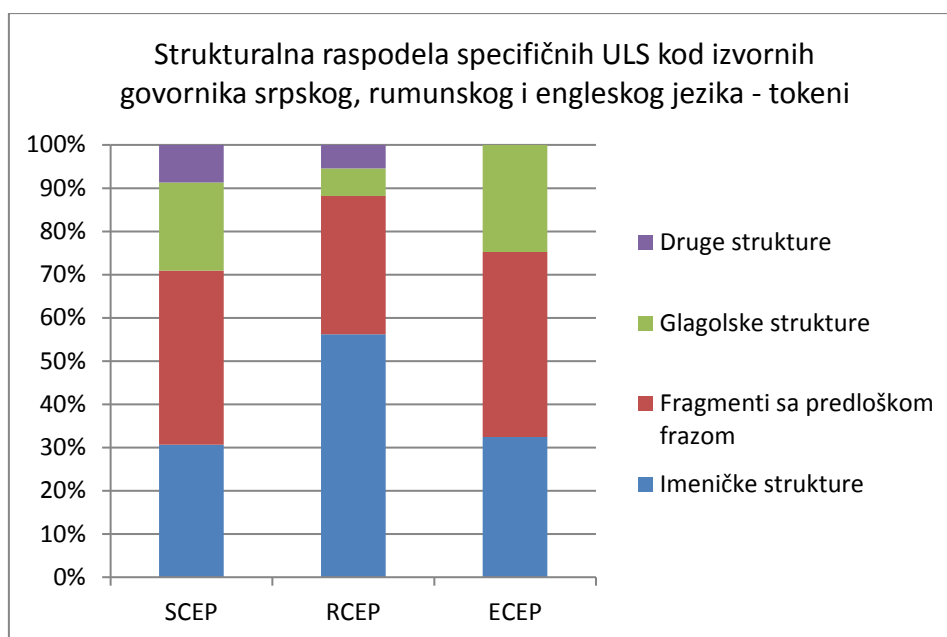


Grafikon 57. *Strukturalna raspodela specifičnih ULS kod izvornih govornika srpskog, rumunskog i engleskog jezika – tipovi*

Kada sagledamo strukturalne raspodele ULS na tokene, uviđamo istu tendenciju raspodele kao i za tipove, naime, da je sličnija raspodela između ECEP i SCEP korpusa nego RCEP korpusa, v. Tabelu 117 i Grafikon 58. Međutim, za razliku od tipova, za tokene statistička analiza pokazala je postojanje značajnih razlika u strukturalnoj raspodeli na grupe i podgrupe između izvornih govornika engleskog i dve grupe neizvornih govornika, kao i između dve grupe neizvornih govornika. Kada šire sagledamo statističku analizu, uviđamo da su veće razlike u strukturalnoj raspodeli tokena specifičnih ULS identifikovani između izvornih govornika rumunskog jezika i izvornih govornika engleskog nego što su između izvornih govornika srpskog i izvornih govornika engleskog. Naime, statistička analiza ECEP i SCEP korpusa pokazala je statistički značajne razlike u raspodeli tokena specifičnih ULS u dve strukturalne grupe i pet (od 11 zastupljenih) podgrupa, dok je za ECEP i RCEP korpus pokazala statistički značajne razlike u sve četiri strukturalne grupe i u sedam (od osam zastupljenih) podgrupa. Statistička analiza je pokazala najveće razlike u strukturalnoj raspodeli specifičnih ULS između SCEP i RCEP korpusa – statistički značajne razlike identifikovane su u svim strukturalnim grupama i u svih deset zastupljenih strukturalnih podgrupa.

Tabela 117. *Strukturalna raspodela specifičnih ULS kod izvornih govornika srpskog, rumunskog i engleskog jezika – tokeni*

	SCEP	RCEP	ECEP
Imeničke strukture	30.67%	56.24%	32.43%
Fragmenti sa predloškom frazom	40.23%	31.95%	42.79%
Glagolske strukture	20.43%	6.35%	24.77%
Druge strukture	8.67%	5.47%	0%



Grafikon 58. *Strukturalna raspodela specifičnih ULS kod izvornih govornika srpskog, rumunskog i engleskog jezika – tokeni*

Poređenje procentualne zastupljenosti tokena specifičnih ULS u strukturalnim grupama i podgrupama u celim SCEP i RCEP korpusima uviđamo da je u oba korpusa manja procentualna zastupljenost tokena nego tipova. Ova razlika je najistaknutija u strukturalnoj grupi *druge strukture* u SCEP korpusu, gde je procentualna zastupljenost tipova specifičnih ULS 76,47% a zastupljenost tokena specifičnih ULS je 36,95%. Veća zastupljenost tipova uporedno sa tokenima ukazuje nam na to da specifični ULS, u obe istražene grupe neizvornih govornika u naučnim radovima iz oblasti obrazovanja, imaju veću raznolikost oblika a manju zastupljenost po broju primera u samom korpusu.

Značajno je napomenuti da izvorni govornici rumunskog jezika koriste manji broj specifičnih ULS uporedno sa izvornim govornicima srpskog, kao i da specifični ULS sačinjavaju manji procenat tipova i tokena u čitavom korpusu. U RCEP korpusu specifični ULS najzastupljeniji su u kategoriji *imeničke strukture* i sačinjavaju 41,67% tipova a je najmanje zastupljena kategorija *glagolske strukture* sa 20%; u SCEP korpusu specifični ULS najzastupljeniji su u strukturalnoj kategoriji *glagolske strukture* sa 76,47% a najmanje u *fragmentima sa predloškom frazom* sa 52%. To znači da specifični ULS sačinjavaju manji procenat u najzastupljenijoj kategoriji u RCEP korpusu (41,67%) uporedno sa najmanje zastupljenom strukturalnom kategorijom u SCEP korpusu (52%).

Činjenica da statistički značajne razlike nisu identifikovane u strukturalnoj raspodeli tipova a jesu u strukturalnoj raspodeli tokena ukazuje na to da *nisu značajne razlike u raznovrsnosti ULS iz različitih strukturalnih kategorija* ali da *jesu značajne razlike u njihovoj učestalosti upotrebe*.

4.11.2.2. Funkcionalna raspodela specifičnih ULS

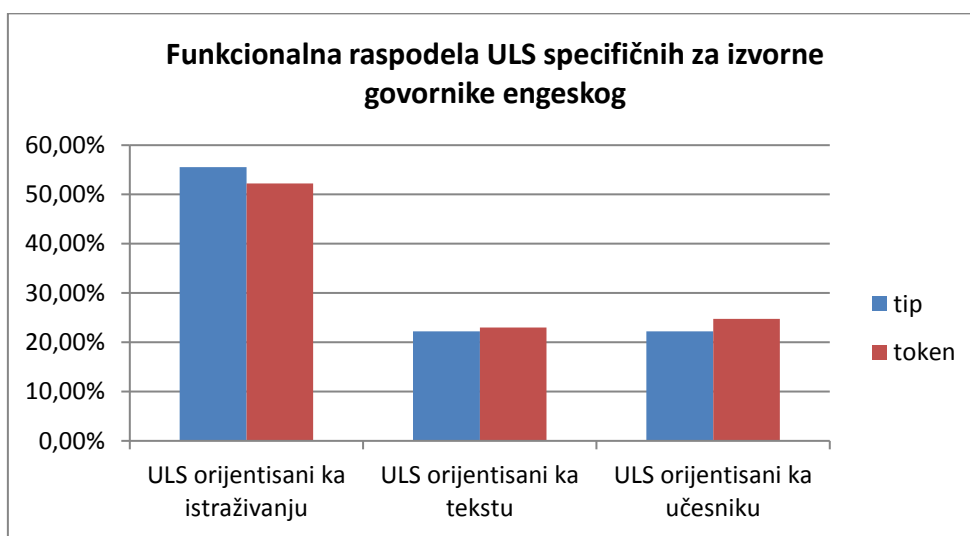
Analiza funkcionalne raspodele specifičnih ULS kod izvornih govornika engleskog jezika, neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski omogućava da dobijemo jasniju sliku o tome na koji se način upotrebljavaju specifični ULS. Štaviše, primenom funkcionalne analize specifičnih ULS moći ćemo jasnije da sagledamo razlike u upotrebi specifičnih ULS u tri korpusa.

Funkcionalna raspodela specifičnih ULS kod izvornih govornika engleskog jezika

Analiza funkcionalne raspodele specifičnih ULS za izvorne govornike engleskog jezika u naučnim radovima iz oblasti obrazovanja pokazuje nam da je funkcionalna kategorija *ULS orijentisani ka istraživanju* najzastupljenija (55,56%), dok funkcionalne kategorije *ULS orijentisani ka tekstu* i *ULS orijentisani ka učesniku* sačinjavaju isti procenat tipova (22,22%), v. Tabelu 118 i Grafikon 59. Kada uporedimo funkcionalnu raspodelu specifičnih ULS u ECEP korpusa sa funkcionalnom raspodelom svih ULS u ECEP korpusu, uviđamo da postoje razlike. Naime, u celom ECEP korpusu isto je najzastupljenija funkcionalna kategorija *ULS orijentisani ka istraživanju*, druga je *ULS orijentisani ka tekstu* a je najmanje zastupljena kategorija *ULS orijentisani ka učesniku*; razlike se, dakle, ogledaju u manjoj zastupljenosti ULS orijentisanih ka tekstu u ULS koji su specifični za izvorne govornike engleskog jezika. Ovi podaci nam ukazuju na to da neizvorni govornici engleskog jezika kojima je maternji jezik srpski ili rumunski imaju tendenciju da zanemare upotrebu relevantnih ULS u funkcionalnoj kategoriji *ULS orijentisani ka istraživanju* i *učesniku* nego u *ULS orijentisanih ka tekstu*.

Tabela 118. *Funkcionalna raspodela ULS specifičnih za izvorne govornike engleskog jezika*

Funkcionalna kategorija	Tipovi (tokeni)	Procenti
<i>ULS orijentisani ka istraživanju</i>	5 (116)	55.56 (52.25)
<i>ULS orijentisani ka tekstu</i>	2 (51)	22.22 (22.97)
<i>ULS orijentisani ka učesniku</i>	2 (55)	22.22 (24.77)



Grafikon 59. *Funkcionalna raspodela ULS specifičnih za izvorne govornike srpskog jezika*

Razmatranjem procentualne zastupljenosti specifičnih ULS u ECEP korpusu po funkcionalnim kategorijama (v. Tabelu 119) uviđamo da se najveća zastupljenost specifičnih ULS nalazi u funkcionalnoj kategoriji *ULS orijentisani ka učesniku* praćeno funkcionalnom kategorijom *ULS orijentisani ka istraživanju*. Uzimajući i ove podatke u obzir, možemo dodatno osnažiti zaključak da je u radu sa studentima/budućim prevodiocima u naučnoj oblasti obrazovanja potrebno posvetiti dodatnu pažnju na podučavanje ULS u funkcionalnim kategorijama *ULS orijentisani ka učesniku* i *ULS orijentisanih ka istraživanju* a kako ne bi došlo do nedovoljnog korišćenja relevantnih oblika.

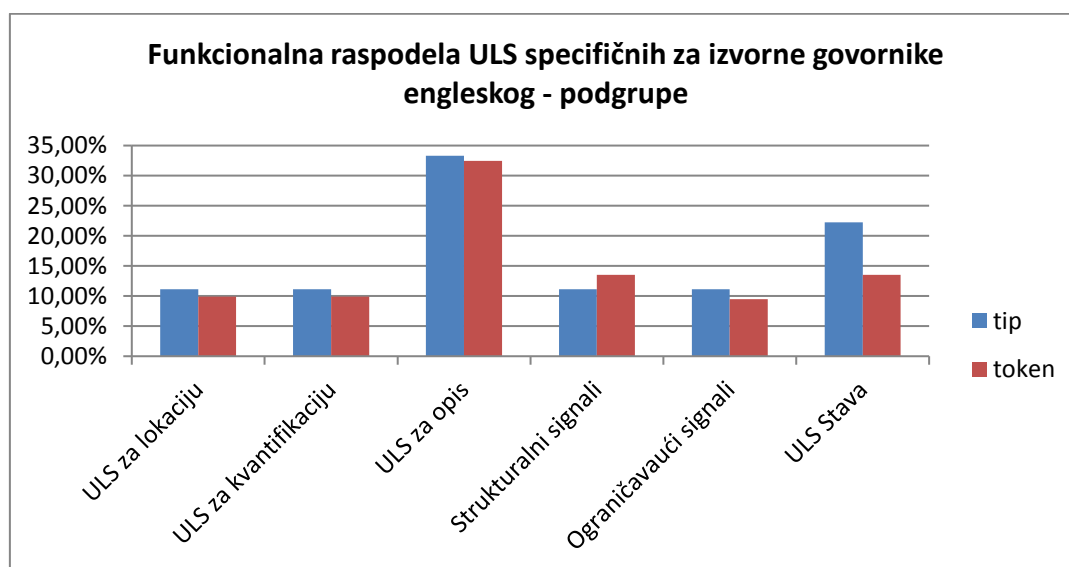
Tabela 119. Zastupljenost specifičnih ULS u ECEP korpusu po funkcionalnim grupama – tipovi i tokeni

	ECEP svi ULS	ECEP specifični ULS	%
<i>ULS orijentisani ka istraživanju</i>	13 (344)	5 (116)	38.46 (33.72)
<i>ULS orijentisani ka tekstu</i>	11 (332)	2 (51)	18.18 (15.36)
<i>ULS orijentisani ka učesniku</i>	3 (101)	2 (55)	66.67 (54.46)

Analizom raspodele specifičnih ULS za ECEP korpus po funkcionalnim potkategorijama zaključujemo da je najzastupljenija funkcionalna potkategorija *ULS za opis* koja sačinjava 33,33% svih specifičnih tipova u ECEP korpusu, praćena potkategorijom *ULS stava* koja sačinjava 22,22% svih specifičnih tipova u ECEP korpusu, v. Tabelu 120 i Grafikon 60. Ovakvi podaci ukazuju na to da neizvorni govornici engleskog jezika kojima su maternji jezici srpski i rumunski u naučnim radovima iz oblasti obrazovanja imaju tendenciju da zanemare upotrebu relevantnih ULS iz funkcionalnih podgrupa *ULS za opis* i *ULS stava*, te da bi veći fokus pri podučavanju trebalo da se stavi na ovde dve funkcionalne podgrupe.

Tabela 120. Funkcionalna raspodela ULS specifičnih za izvorne govornike engleskog – podgrupe

Funkcionalna potkategorija	Tipovi (tokeni)	Procenti
<i>ULS za lokaciju</i>	1 (22)	11.11 (9.91)
<i>ULS za kvantifikaciju</i>	1 (22)	11.11 (9.91)
<i>ULS za opis</i>	3 (72)	33.33 (32.43)
<i>Strukturalni signali</i>	1 (30)	11.11 (13.51)
<i>Ograničavajući signali</i>	1 (21)	11.11 (9.46)
<i>ULS stava</i>	2 (30)	22.22 (13.51)



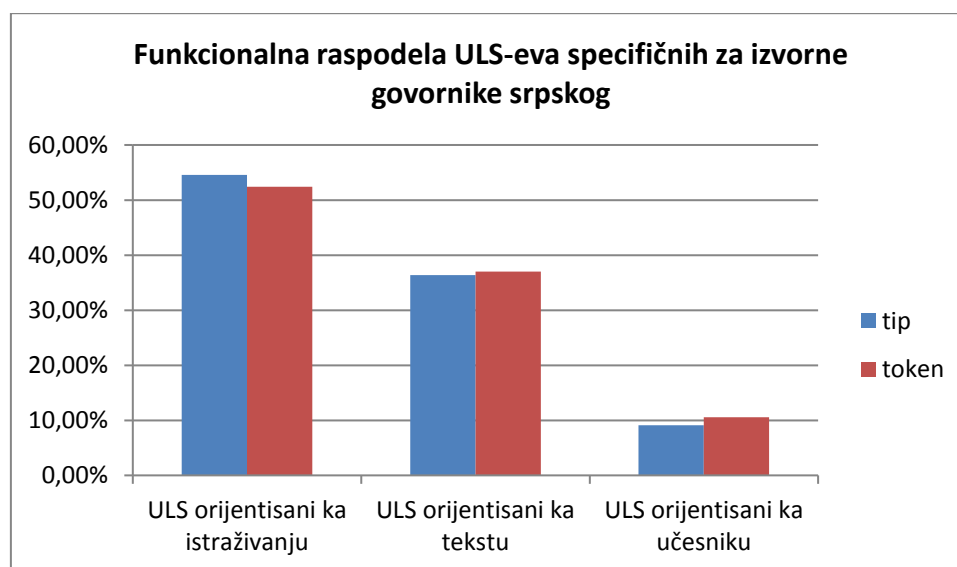
Grafikon 60. Funkcionalna raspodela ULS specifičnih za izvorne govornike engleskog jezika – podgrupe

Funkcionalna raspodela specifičnih ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski jezik

Analiza funkcionalne raspodele specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski u naučnim radovima iz oblasti obrazovanja pokazala je najveću zastupljenost funkcionalne kategorije *ULS orijentisani ka istraživanju*, koja sačinjava 54,55% svih specifičnih ULS za SCEP korpus, v. Tabelu 121 i Grafikon 61. Druga najzastupljenija funkcionalna kategorija kod ULS specifičnih za SCEP korpus su *ULS orijentisani ka tekstu*, dok je najmanje zastupljena funkcionalna kategorija *ULS orijentisani ka učesniku*. Kada uporedimo funkcionalnu raspodelu specifičnih ULS u SCEP korpusu i svih ULS u SCEP korpusu, uviđamo da je ista raspodela funkcionalnih grupa prisutna u obe grupe ULS, iako se malo veća procentualna razlika između grupa *ULS orijentisani ka istraživanju* (54,55%) i *ULS orijentisani ka tekstu* (36,36%) uviđa u specifičnim ULS za SCEP korpus. Kada sagledamo procentualnu zastupljenost specifičnih ULS u celom SCEP korpusu po funkcionalnim kategorijama, uviđamo da sačinjavaju najveći procent ULS u funkcionalnoj kategoriji *ULS orijentisani ka učesniku* (75%), pa zatim u funkcionalnoj kategoriji *ULS orijentisani ka istraživanju* (64,29), dok su najmanje zastupljeni u funkcionalnoj kategoriji *ULS orijentisani ka tekstu* (51,06%), v. Tabelu 122. Na osnovu navedenih podataka možemo zaključiti da neizvorni govornici engleskog jezika kojima je maternji jezik srpski u naučnim radovima u oblasti obrazovanja imaju tendenciju da koriste više specifičnih ULS u funkcionalnim kategorijama *ULS orijentisani ka učesniku* i *ULS orijentisani ka istraživanju*. Pri podučavanju studenata ili prevodilaca, koji žele da se bave prevodom naučnih radova u oblasti obrazovanja, a kojima je maternji jezik srpski, trebalo bi posebnu pažnju posveti podučavanju prihvatljivih ULS iz dve navedene funkcionalne kategorije kako ne bi pribegavali upotrebi ULS koji su manje pogodni za pokazivanje pripadnosti u navedenoj akademskoj/delatnoj zajednici.

Tabela 121. *Funkcionalna raspodela ULS specifičnih za izvorne govornike srpskog jezika*

Funkcionalna kategorija	Tipovi (tokeni)	Procenti
<i>ULS orijentisani ka istraživanju</i>	36 (1070)	54.55 (52.43)
<i>ULS orijentisani ka tekstu</i>	24 (755)	36.36 (36.99)
<i>ULS orijentisani ka učesniku</i>	6 (216)	9.09 (10.58)



Grafikon 61. *Funkcionalna raspodela ULS specifičnih za izvorne govornike srpskog jezika*

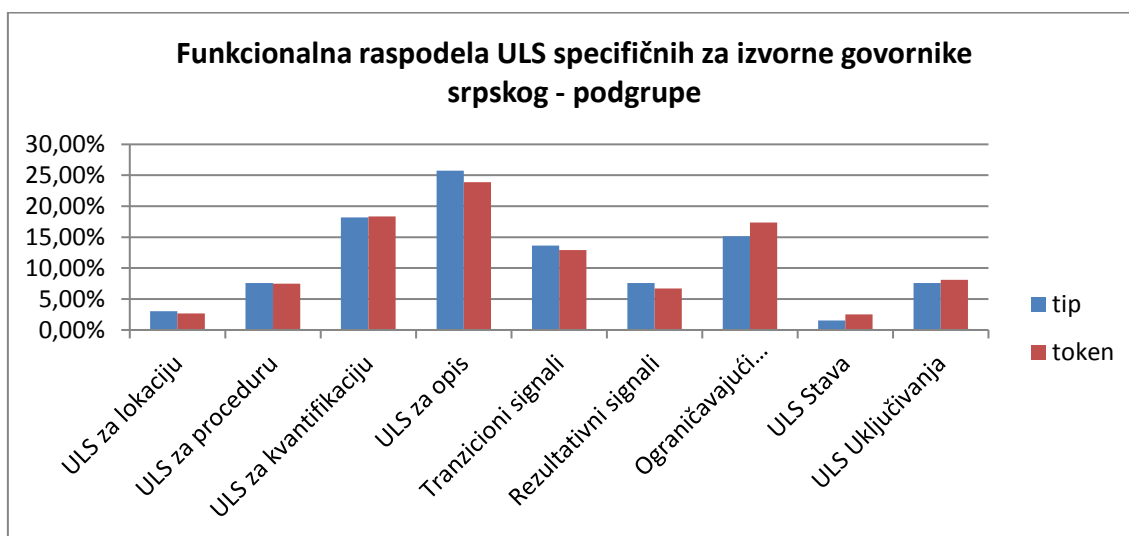
Tabela 122. Zastupljenost specifičnih ULS u SCEP korpusu po funkcionalnim grupama – tipovi i tokeni

	SCEP svi ULS	SCEP specifični ULS	%
<i>ULS orijentisani ka istraživanju</i>	56 (1824)	36 (1070)	64.29 (58.66)
<i>ULS orijentisani ka tekstu</i>	47 (2326)	24 (755)	51.06 (32.46)
<i>ULS orijentisani ka učesniku</i>	8 (386)	6 (216)	75 (55.96)

Funkcionalna raspodela specifičnih ULS za SCEP korpus po funkcionalnim podgrupama pokazala je najveću zastupljenost funkcionalne podgrupe *ULS za opis*, koja je praćena funkcionalnim podgrupama *ULS za kvantifikaciju* i *ograničavajući signali*, v. Tabelu 123 i Grafikon 62. Najmanje zastupljena funkcionalna podgrupa su *ULS stava* koji predstavljaju samo 1,52% tipova svih specifičnih ULS u SCEP korpusu. Navedeni podaci ukazuju na to da bi u radu sa studentima i prevodiocima koji žele da prevode naučne radove iz oblasti obrazovanja, trebalo posveti posebnu pažnju obradi ULS iz funkcionalnih podgrupa *ULS za opis*, *ULS za kvantifikaciju* i *ograničavajući signali* kako ne bi došlo do prekomerne upotrebe ULS koji nisu najpogodnije rešenje za pokazivanje pripadnosti u navedenoj akademskoj zajednici. Takođe bi trebalo analizirati koja se značenja prenose specifičnim ULS iz navedenih grupa kako bi u nastavi mogli da se obrade adekvatnija rešenja.

Tabela 123. Funkcionalna raspodela ULS specifičnih za izvorne govornike srpskog – podgrupe

Funkcionalna potkategorija	Tipovi (tokeni)	Procenti
<i>ULS za lokaciju</i>	2 (54)	3.03 (2.65)
<i>ULS za proceduru</i>	5 (153)	7.58 (7.50)
<i>ULS za kvantifikaciju</i>	12 (375)	18.18 (18.37)
<i>ULS za opis</i>	17 (488)	25.76 (23.91)
<i>Tranzicioni signali</i>	9 (263)	13.64 (12.89)
<i>Rezultativni signali</i>	5 (137)	7.58 (6.71)
<i>Ograničavajući signali</i>	10 (355)	15.15 (17.39)
<i>ULS stava</i>	1 (51)	1.52 (2.50)
<i>ULS uključivanja</i>	5 (156)	7.58 (8.08)



Grafikon 62. Funkcionalna raspodela ULS specifičnih za izvorne govornike srpskog – podgrupe

Poređenje funkcionalne raspodele specifičnih ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski

Kako bi se ispitale potencijalne razlike, tj., asocijacije između ECEP i SCEP korpusa u odnosu na funkcionalne distribucije specifičnih ULS korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između ECEP i SCEP korpusa u odnosu na funkcionalne grupe specifičnih ULS – tipovi. – Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih grupa specifičnih ULS u odnosu na tipove prikazana je u Tabeli 124. Hi-kvadrat test nije bio značajan, $\chi^2(2, N = 75) = 1.74, p > .050$, nije identifikovana statistički značajna povezanost između korpusa i funkcionalnih grupa. Drugim rečima, uočene razlike u funkcionalnoj raspodeli tipova specifičnih ULS između ECEP i SCEP korpusa nisu statistički značajne.

Tabela 124. *Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih grupa specifičnih ULS u odnosu na tipove*

Funkcionalne grupe	SCEP			ECEP		
	tipovi	%	R	tipovi	%	R
<i>ULS orijentisani ka istraživanju</i>	36	54.5%	-0.1	5	55.6%	0.1
<i>ULS orijentisani ka tekstu</i>	24	36.4%	0.8	2	22.2%	-0.8
<i>ULS orijentisani ka učesniku</i>	6	9.1%	-1.2	2	22.2%	1.2
Total	66	100.0%		9	100.0%	

*Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati*

Razlike između ECEP i SCEP korpusa u odnosu na funkcionalne grupe specifičnih ULS – tokeni. – Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih grupa specifičnih ULS u odnosu na tokene prikazana je u Tabeli 125. Hi-kvadrat test je bio značajan, $\chi^2(2, N = 2263) = 44.72, p < .001$, ukazujući na to da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa. Korigovani standardizovani reziduali pokazuju da je među specifičnim ULS, veća upotreba *ULS orijentisanih ka tekstu* u SCEP, a *ka učesnika* u ECEP korpusu. Navedeni podaci nas dovode do zaključka da izvorni govornici srpskog imaju tendenciju da značajno više koriste specifične *ULS orijentisane ka tekstu* koji nisu najbolji načini za prenošenje značenja u istraženoj delatnoj zajednici. Stoga je potrebno proučiti koja značenja prenose ULS iz kategorije *ULS orijentisani ka tekstu* i podučiti adekvatnija rešenja za prenošenje značenja. Rezultati takođe pokazuju tendenciju izvornih govornika srpskog da posebno zanemare relevantne ULS iz funkcionalne kategorije *ULS orijentisani ka učesniku*, te se posebna pažnja treba posvetiti podučavanju ovih konstrukcija.

Tabela 125. *Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih grupa specifičnih ULS u odnosu na tokene*

Funkcionalne grupe	SCEP			ECEP		
	tokeni	%	R	tokeni	%	R
<i>ULS orijentisani ka istraživanju</i>	1070	52.4%	0.0	116	52.3%	0.0
<i>ULS orijentisani ka tekstu</i>	755	37.0%	4.1*	51	23.0%	-4.1*
<i>ULS orijentisani ka učesniku</i>	216	10.6%	-6.2*	55	24.8%	6.2*
Total	2041	100.0%		222	100.0%	

*Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati*

Razlike između ECEP i SCEP korpusa u odnosu na funkcionalne podgrupe specifičnih ULS – tipovi. – Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih podgrupa specifičnih ULS u odnosu na tipove prikazana je u Tabeli 126. Hi-kvadrat test je bio značajan, $\chi^2(9, N = 75) = 20.87$, $p = .013$, što ukazuje na to da postoji statistički značajna povezanost između korpusa i funkcionalnih podgrupa.

Tabela 126. Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih podgrupa specifičnih ULS u odnosu na tipove

Grupa	Podgrupa	SCEP			ECEP		
		tipovi	%	R	tipovi	%	R
<i>ULS orijentisani</i>	<i>ULS za lokaciju</i>	2	3.0%	-1.2	1	11.1%	1.2
<i>ka</i>	<i>ULS za proceduru</i>	5	7.6%	0.9	0	0%	-0.9
<i>istraživanju</i>	<i>ULS za kvantifikaciju</i>	12	18.2%	0.5	1	11.1%	-0.5
	<i>ULS za opis</i>	17	25.8%	-0.5	3	33.3%	0.5
<i>ULS orijentisani</i>	<i>Tranzicioni signali</i>	9	13.6%	1.2	0	0%	-1.2
<i>ka tekstu</i>	<i>Rezultativni signali</i>	5	7.6%	0.9	0	0%	-0.9
	<i>Strukturni signali</i>	0	0%	-2.7*	1	11.1%	2.7*
	<i>Ograničavajući signali</i>	10	15.2%	0.3	1	11.1%	-0.3
<i>ULS orijentisani</i>	<i>ULS stava</i>	1	1.5%	-3.0*	2	22.2%	3.0*
<i>ka učesniku</i>	<i>ULS uključivanja</i>	5	7.6%	0.9	0	0%	-0.9
Total		66	100.0%		9	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Razlike između ECEP i SCEP korpusa u odnosu na funkcionalne podgrupe specifičnih ULS – tokeni. – Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih podgrupa specifičnih ULS u odnosu na tokene prikazana je u Tabeli 127. Hi-kvadrat test je bio značajan, $\chi^2(8, N = 2498) = 619.11$, $p < .001$, odnosno postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

Korigovani standardizovani reziduali za ECEP korpus bili su povišeni u slučaju *ULS za lokaciju*, *ULS za opis*, *strukturne signale* i *ULS stava*, dok su za ostale podgrupe standardizovani reziduali bili povišeni u slučaju SCEP korpusa. Ovi podaci nas dalje navode na zaključke da izvorni govornici srpskog nedovoljno često koriste relevantne ULS iz funkcionalnih kategorija *ULS za lokaciju*, *ULS za opis*, *strukturne signale* i *ULS stava*, te da je u nastavi potrebno posvetiti posebnu pažnju pri podučavanju relevantnih oblika navedenih funkcionalnih podgrupa. Podaci, takođe, ukazuju na to da izvorni govornici srpskog jezika isuviše često koriste neadekvatne ULS u funkcionalnim kategorijama *ULS za proceduru*, *ULS za kvantifikaciju*, *tranzicioni signali*, *rezultativni signali*, *ograničavajući signali* i *ULS uključivanja*. Stoga je u nastavi potrebno obraditi alternative za upotrebu ULS koji nisu efikasni u pokazivanju pripadnosti relevantnoj delatnoj zajednici.

Činjenica da se pri statističkoj analizi funkcionalnih podgrupa specifičnih ULS u SCEP i ECEP korpusima u analizi tipova identifikovala statistički značajna razlika u dve podgrupe, dok se u analizi tokena statistički značajna razlika identifikovala za sve podgrupe, ukazuje na to da se značajna razlika u konkretnim upotrebljenim ULS identifikovala u samo dve grupe ali da se frekventnost upotrebe ULS značajno razlikuje za sve funkcionalne podgrupe.

Tabela 127. Unakrsna tabela između ECEP i SCEP korpusa i funkcionalnih podgrupa specifičnih ULS u odnosu na tokene

Grupa	Podgrupa	SCEP			ECEP		
		tokeni	%	R	tokeni	%	R
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	54	2.6%	-5.7*	22	9.9%	5.7*
	<i>ULS za proceduru</i>	153	7.5%	4.2*	0	0%	-4.2*
	<i>ULS za kvantifikaciju</i>	375	18.4%	3.1*	22	9.9%	-3.1*
	<i>ULS za opis</i>	488	23.9%	-2.8*	72	32.4%	2.8*
<i>ULS orijentisani ka tekstu</i>	<i>Tranzicioni signali</i>	263	12.9%	5.7*	0	0%	-5.7*
	<i>Rezultativni signali</i>	137	6.7%	4.0*	0	0%	-4.0*
	<i>Strukturalni signali</i>	0	0%	-16.7*	30	13.5%	16.7*
	<i>Ograničavajući signali</i>	355	17.4%	3.0*	21	9.5%	-3.0*
<i>ULS orijentisani ka učesniku</i>	<i>ULS stava</i>	51	2.5%	-14.9*	55	24.8%	14.9*
		165	8.1%	4.4*	0	0%	-4.4*
	<i>ULS uključivanja</i>						
Total		2041	100%		222	100.0%	

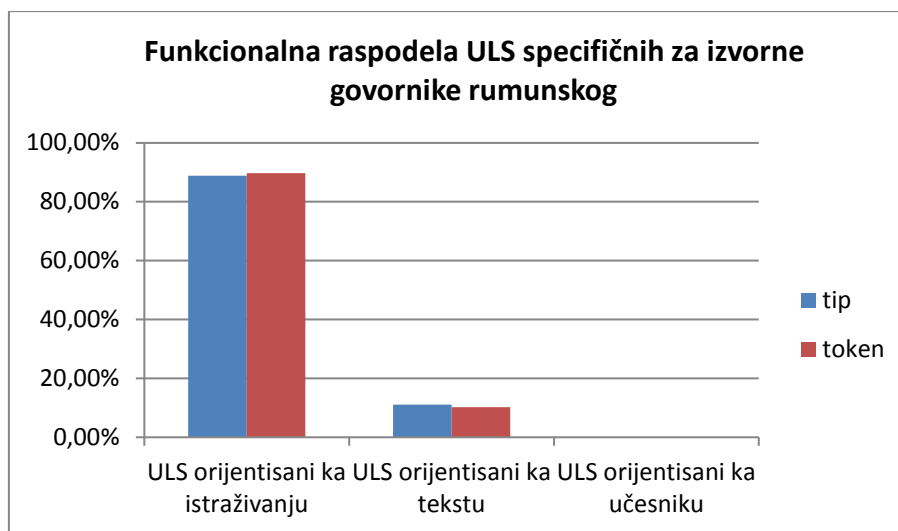
Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati

Funkcionalna raspodela specifičnih ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski jezik

Analizom funkcionalne raspodele specifičnih ULS za RCEP korpus ustanovljeno je da je najzastupljenija funkcionalna kategorija *ULS orijentisani ka istraživanju* i da sačinjava 88,89% svih tipova specifičnih za produkciju izvornih govornika rumunskog, v. Tabelu 127 i Grafikon 63. Svojevremena za ovaj korpus je činjenica da se nijedan specifičan ULS za RCEP korpus ne nalazi u funkcionalnoj kategoriji *ULS orijentisani ka učesniku*. Raspodela specifičnih ULS za RCEP korpus slična je funkcionalnoj raspodeli svih ULS u RCEP korpusu. Naime, i u celom RCEP korpusu najzastupljenija funkcionalna grupa je *ULS orijentisani ka istraživanju*, praćena *ULS orijentisanim ka tekstu*, dok se mali procenat tipova nalazi u kategoriji *ULS orijentisani ka učesniku* (3,33%). Kada sagledamo koliki procenat svih ULS u RCEP korpusu sačinjavaju specifični ULS po funkcionalnim grupama uviđamo da u grupi *ULS orijentisani ka istraživanju* sačinjavaju 44,44% svih tipova u toj kategoriji, dok u kategoriji *ULS orijentisani ka tekstu* sačinjavaju 9,09% tipova, v. Tabelu 128. Na osnovu podataka navedenih u ovom pasusu možemo zaključiti da bi u radu sa studentima i/ili prevodiocima trebalo: (a) staviti veći fokus na podučavanje ULS u funkcionalnoj kategoriji *ULS orijentisani ka istraživanju* kako ne bi pribegavali upotrebi ULS koji su manje primereni za pokazivanje pripadnosti u navedenoj delatnoj zajednici, kao i na podučavanje prikladnijih alternativa za specifične ULS u datoj kategoriji; (b) posvetiti više pažnje podučavanju ULS iz funkcionalne kategorije *ULS orijentisani ka učesniku*, koji se retko koriste u celom RCEP korpusu, a na osnovu ULS koji su specifični u ECEP korpusu vidimo da ih neizvorni govornici nedovoljno koriste.

Tabela 127. Funkcionalna raspodela ULS specifičnih za izvorne govornike rumunskog jezika

Funkcionalna kategorija	Tipovi (tokeni)	Procenti
<i>ULS orijentisani ka istraživanju</i>	16 (410)	88.89 (89.72)
<i>ULS orijentisani ka tekstu</i>	2 (47)	11.11 (10.28)
<i>ULS orijentisani ka učesniku</i>	0 (0)	0 (0)



Grafikon 63. Funkcionalna raspodela ULS specifičnih za izvorne govornike rumunskog jezika

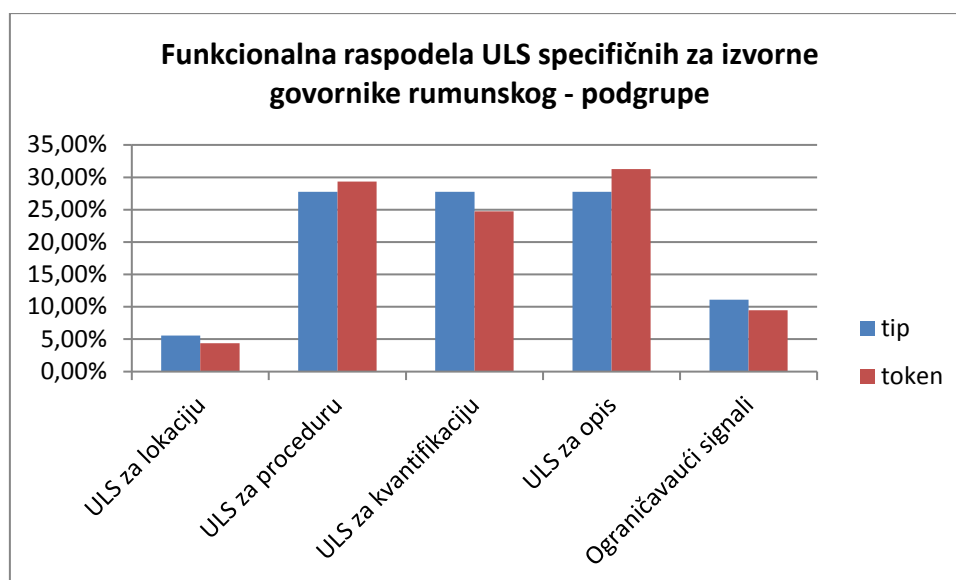
Tabela 128. Zastupljenost specifičnih ULS u RCEP korpusu po funkcionalnim grupama – tipovi i tokeni

	RCEP svi ULS	RCEP specifični ULS	%
<i>ULS orijentisani ka istraživanju</i>	36 (1111)	16 (410)	44.44 (36.90)
<i>ULS orijentisani ka tekstu</i>	22 (1190)	2 (47)	9.09 (3.95)
<i>ULS orijentisani ka učesniku</i>	2 (57)	0 (0)	0

Analizom funkcionalnih podgrupa specifičnih ULS za RCEP korpus uvideli smo da postoje tri funkcionalne podgrupe sa istom zastupljenošću tipova (27,78%): *ULS za proceduru*, *ULS za kvantifikaciju* i *ULS za opis*, v. Tabelu 129 i Grafikon 64. Najmanje zastupljena je funkcionalna podgrupa *ULS za lokaciju*. U radu sa studentima i prevodiocima koji žele da prevode naučne radove u oblasti obrazovanja dodatan fokus bi trebalo da se stavi na podučavanje *ULS za proceduru*, *kvantifikaciju* i *opis* kako ne bi došlo do upotrebe ULS koji nisu najoptimalnija opcija za pokazivanje pripadnosti u navedenoj delatnoj zajednici, kao i podučavanju alternativnih i prikladnijih načina za prenošenje značenja za koje se prenose specifični ULS u navedenim funkcionalnim podgrupama.

Tabela 129. Funkcionalna raspodela ULS specifičnih za izvorne govornike rumunskog – podgrupe

Funkcionalna potkategorija	Tipovi (tokeni)	Procenti
<i>ULS za lokaciju</i>	1 (20)	5.56 (4.38)
<i>ULS za proceduru</i>	5 (134)	27.78 (29.32)
<i>ULS za kvantifikaciju</i>	5 (113)	27.78 (24.73)
<i>ULS za opis</i>	5 (143)	27.78 (31.29)
<i>Ograničavajući signali</i>	2 (47)	11.11 (9.46)



Grafikon 64. Funkcionalna raspodela ULS specifičnih za izvorne govornike rumunskog jezika – podgrupe

Poređenje funkcionalne raspodele specifičnih ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski jezik

Kako bi se ispitale potencijalne razlike, tj. asocijacije između ECEP i RCEP korpusa u odnosu na funkcionalne distribucije specifičnih ULS korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između ECEP i RCEP korpusa u odnosu na funkcionalne grupe specifičnih ULS – tipovi. – Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih grupa specifičnih ULS u odnosu na tipove prikazana je u Tabeli 130. Hi-kvadrat test nije bio značajan, $\chi^2 (2, N = 27) = 5.35, p > .050$, ukazujući na to da nije identifikovana statistički značajna povezanost između korpusa i funkcionalnih grupa. Drugim rečima, razlike u funkcionalnoj distribuciji specifičnih tipova ULS u ECEP i RCEP korpusima nisu statistički značajne.

Tabela 130. Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih grupa specifičnih ULS u odnosu na tipove

Funkcionalne grupe	RCEP			ECEP		
	tipovi	%	R	tipovi	%	R
<i>ULS orijentisani ka istraživanju</i>	16	88.9%	2.0	5	55.6%	-2.0
<i>ULS orijentisani ka tekstu</i>	2	11.1%	-0.8	2	22.2%	0.8
<i>ULS orijentisani ka učesniku</i>	0	0%	-2.1	2	22.2%	2.1
Total	18	100.0%		9	100.0%	

*Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali; * – statistički značajni rezultati*

Razlike između ECEP i RCEP korpusa u odnosu na funkcionalne grupe specifičnih ULS – tokeni. Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih grupa specifičnih ULS u odnosu na tokene prikazana je u Tabeli 131. Hi-kvadrat test je bio značajan, $\chi^2 (2, N = 679) =$

156.95, $p < .001$, ukazujući na to da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

U RCEP korpusu veća je bila upotreba *ULS usmerenih ka istraživanju*, dok je u ECEP korpusu bila viša upotreba *ULS usmerenih ka tekstu i učesniku*. Navedeni podaci ukazuju na tendenciju izvornih govornika rumunskog jezika da isuviše često koriste specifične ULS, koji su neadekvatni za pokazivanje pripadnosti u relevantnoj delatnoj zajednici, u funkcionalnoj kategoriji *ULS orijentisani ka istraživanju*. Stoga je u radu sa izvornim govornicima rumunskog jezika potrebno podučavati adekvatnije alternative za navedene specifične ULS iz funkcionalne grupe *ULS orijentisani ka istraživanju*. Rezultati statističke analize su, takođe, pokazali da izvorni govornici rumunskog imaju tendenciju da isuviše retko koriste adekvatne ULS iz funkcionalnih kategorija *ULS orijentisani ka tekstu i ULS orijentisani ka učesniku*.

Tabela 131. *Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih grupa specifičnih ULS u odnosu na tokene*

Funkcionalne grupe	RCEP			ECEP		
	tokeni	%	R	tokeni	%	R
<i>ULS orijentisani ka istraživanju</i>	410	89.7%	11.0*	116	52.3%	-11.0*
<i>ULS orijentisani ka tekstu</i>	47	10.3%	-4.4*	51	23.0%	4.4*
<i>ULS orijentisani ka učesniku</i>	0	0%	-11.1*	55	24.8%	11.1*
Total	457	100.0%		222	100.0%	

*Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati*

Razlike između ECEP i RCEP korpusa u odnosu na funkcionalne podgrupe specifičnih ULS – tipovi. – Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih podgrupa specifičnih ULS u odnosu na tipove prikazana je u Tabeli 132. Hi-kvadrat test nije bio značajan, $\chi^2(6, N = 27) = 9.56$, $p > .050$, što ukazuje na to da nije identifikovana statistički značajna povezanost između korpusa i funkcionalnih podgrupa, odnosno, razlike koje uvidamo u funkcionalnoj raspodeli specifičnih tipova ULS u ECEP i RCEP korpusima na podgrupe nisu značajne.

Tabela 132. *Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih podgrupa specifičnih ULS u odnosu na tipove*

Grupe	Podgrupa	RCEP			ECEP		
		tipovi	%	R	tipovi	%	R
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	1	5.6%	-0.5	1	11.1%	0.5
	<i>ULS za proceduru</i>	5	27.8%	1.8	0	0%	-1.8
		5	27.8%	1.0	1	11.1%	-1.0
	<i>ULS za kvantifikaciju</i>	5	27.8%	-0.3	3	33.3%	0.3
<i>ULS orijentisani ka tekstu</i>	<i>Tranzicioni signali</i>	/	/	/	/	/	/
	<i>Rezultativni signali</i>	/	/	/	/	/	/
	<i>Strukturni signali</i>	0	0%	-1.4	1	11.1%	1.4
<i>ULS orijentisani ka učesniku</i>	<i>Ograničavajući signali</i>	2	11.1%	0.0	1	11.1%	0.0
	<i>ULS stava</i>	0	0%	-2.1	2	22.2%	2.1
	<i>ULS uključivanja</i>	/	/	/	/	/	/
Total		18	100.0%		9	100.0%	

*Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;
* – statistički značajni rezultati*

Razlike između ECEP i RCEP korpusa u odnosu na funkcionalne podgrupe specifičnih ULS – tokeni. – Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih podgrupa specifičnih ULS u odnosu na tokene prikazana je u Tabeli 133. Hi-kvadrat test je bio značajan, $\chi^2(6, N = 679) = 264.12$, $p < .001$, ukazujući na to da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

Tabela 133. Unakrsna tabela između ECEP i RCEP korpusa i funkcionalnih podgrupa specifičnih ULS u odnosu na tokene

Grupa	Podgrupa	RCEP			ECEP		
		tokeni	%	R	tokeni	%	R
ULS orijentisani ka istraživanju	ULS za lokaciju	20	4.4%	-2.8*	22	9.9%	2.8*
ULS orijentisani ka tekstu	ULS za proceduru	134	29.3%	9.0*	0	0%	-9.0*
	ULS za kvantifikaciju	113	24.7%	4.5*	22	9.9%	-4.5*
ULS orijentisani ka učesniku	ULS za opis	143	31.3%	-0.3	72	32.4%	0.3
	Tranzicioni signali	/	/	/	/	/	/
ULS orijentisani ka učesniku	Rezultativni signali	/	/	/	/	/	/
	Strukturni signali	0	0%	-8.0*	30	13.5%	8.0*
	Ograničavajući signali	47	10.3%	0.3	21	9.5%	-0.3
ULS orijentisani ka učesniku	ULS stava	0	0%	-11.1*	55	24.8%	11.1*
	ULS uključivanja	/	/	/	/	/	/
Total		457	100%		222	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali; * – statistički značajni rezultati

Za RCEP korpus, korigovani standardizovani reziduali ukazuju na povišenu upotrebu ULS za proceduru, ULS za kvantifikaciju, dok su u ECEP korpusu povećana upotreba ULS za lokaciju, strukturni signali i ULS stava. Rezultati statističke analize nas navode na zaključak da izvorni govornici rumunskog jezika imaju tendenciju da isuviše često koriste neadekvatne ULS u funkcionalnim podgrupama ULS za proceduru i kvantifikaciju. Što dalje znači da je u radu sa ovom grupom neizvornih govornika potrebno pružiti adekvatne alternative za navedene specifične ULS. Izvorni govornici rumunskog jezika, isto tako, imaju tendenciju da nedovoljno koriste adekvatne ULS iz funkcionalnih podgrupa ULS za lokaciju, strukturalni signali i ULS stava. Stoga je poželjno staviti akcenat na podučavanje adekvatnih ULS iz ovih funkcionalnih podgrupa.

Poređenje funkcionalne raspodele specifičnih ULS kod izvornih govornika srpskog i rumunskog jezika pri njihovoj produkciji na engleskom jeziku

Kako bi se ispitala potencijalne razlike, tj. asocijacije između SCEP i RCEP korpusa u odnosu na funkcionalne distribucije specifičnih ULS korišćen je hi-kvadrat test nezavisnosti. Za hi-kvadrat testove koji su bili statistički značajni, analizirani su korigovani standardizovani reziduali kako bi se utvrdilo koje ćelije imaju najveća odstupanja od očekivanih frekvencija, a kao indikator statistički značajnih reziduala korišćena je apsolutna vrednost od 1.96.

Razlike između SCEP i RCEP korpusa u odnosu na funkcionalne grupe specifičnih ULS – tipovi. – Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih grupa specifičnih ULS u odnosu na tipove prikazana je u Tabeli 134. Hi-kvadrat test je bio značajan, $\chi^2(2, N = 84) = 1.97$, $p = .027$, što ukazuje na to da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa u odnosu na tipove.

Korigovani standardizovani reziduali pokazuju da je bilo više tipova povezanih za *ULS orijentisanih ka tekstu* u SCEP korpusu, a više tipova *ULS orijentisanih ka istraživanju* u RCEP korpusu. Navedeni rezultati ukazuju na to da izvorni govornici srpskog i rumunskog jezika statistički značajno odstupaju od norme na različite načine.

Tabela 134. *Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih grupa u odnosu na tipove*

Funkcionalne grupe	SCEP			RCEP		
	tipovi	%	R	tipovi	%	R
<i>ULS orijentisani ka istraživanju</i>	36	54.5%	-2.7*	16	88.9%	2.7*
<i>ULS orijentisani ka tekstu</i>	24	36.4%	2.1*	2	11.1%	-2.1*
<i>ULS orijentisani ka učesniku</i>	6	9.1%	1.3	0	0%	-1.3
Total	66	100.0%		18	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

** – statistički značajni rezultati*

Razlike između SCEP i RCEP korpusa u odnosu na funkcionalne grupe specifičnih ULS – tokeni. – Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih grupa u odnosu na tokene specifičnih ULS prikazana je u Tabeli 135. Hi-kvadrat test je bio značajan, χ^2 (2, N = 2498) = 218.96, $p < .001$, što ukazuje na to da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa.

Korigovani standardizovani reziduali ukazuju na to da je bilo više *ULS orijentisanih ka istraživanju* u RCEP korpusu, dok je bilo više *ULS orijentisanih ka tekstu* i *ka učesniku* u SCEP korpusu. Statistička analiza tokena pokazuje statistički značajne razlike u svim funkcionalnim grupama, što dalje ukazuje da dve različite grupe neizvornih govornika imaju tendenciju da koriste više specifičnih ULS u različitim funkcionalnim grupama.

Tabela 135. *Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih grupa specifičnih ULS u odnosu na tokene*

Funkcionalne grupe	SCEP			RCEP		
	tokeni	%	R	tokeni	%	R
<i>ULS orijentisani ka istraživanju</i>	1070	52.4%	-14.7*	410	89.7%	14.7*
<i>ULS orijentisani ka tekstu</i>	755	37.0%	11.1*	47	10.3%	-11.1*
<i>ULS orijentisani ka učesniku</i>	216	10.6%	7.3*	0	0%	-7.3*
Total	2041	100.0%		457	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

** – statistički značajni rezultati*

Razlike između SCEP i RCEP korpusa u odnosu na funkcionalne podgrupe specifičnih ULS – tipovi. – Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih podgrupa u odnosu na tipove specifičnih ULS prikazana je u Tabeli 136. Hi-kvadrat test nije bio značajan, χ^2 (8, N = 84) = 11.38, $p > .050$, ukazujući na to da nije identifikovana statistički značajna povezanost između korpusa i funkcionalnih podgrupa.

Tabela 136. Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih podgrupa u odnosu na tipove specifičnih ULS

Grupa	Podgrupa	SCEP			RCEP		
		tipovi	%	R	tipovi	%	R
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	2	3.0%	-0.5	1	5.6%	0.5
	<i>ULS za proceduru</i>	5	7.6%	-2.3	5	27.8%	2.3
	<i>ULS za kvantifikaciju</i>	12	18.2%	-0.9	5	27.8%	0.9
	<i>ULS za opis</i>	17	25.8%	-0.2	5	27.8%	0.2
<i>ULS orijentisani ka tekstu</i>	<i>Tranzicioni signali</i>	9	13.6%	1.7	0	0%	-1.7
	<i>Rezultativni signali</i>	5	7.6%	1.2	0	0%	-1.2
	<i>Strukturni signali</i>	/	/	/	/	/	/
	<i>Ograničavajući signali</i>	10	15.2%	0.4	2	11.1%	-0.4
<i>ULS orijentisani ka učesniku</i>	<i>ULS stava</i>	1	1.5%	0.5	0	0%	-0.5
	<i>ULS uključivanja</i>	5	7.6%	1.2	0	0%	-1.2
Total		66	100.0%		18	100.0%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

* – statistički značajni rezultati

Razlike između SCEP i RCEP korpusa u odnosu na funkcionalne podgrupe specifičnih ULS – tokeni. – Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih podgrupa u odnosu na tokene specifičnih ULS prikazana je u Tabeli 137. Hi-kvadrat test je bio značajan, χ^2 (8, N = 2498) = 324.08, $p < .001$, što je ukazalo na to da postoji statistički značajna povezanost između korpusa i funkcionalnih grupa. Korigovani standardizovani reziduali su bili značajno povišeni za SCEP korpus u slučaju *tranzicionih, rezultativnih i ograničavajućih signala, ULS stava i uključivanja*. Za RCEP korpus značajno povišeni standardizovani reziduali su bili za sve podgrupe *ULS orijentisanih ka istraživanju*. Statistička analiza ukazuje na to da se statistički značajno razlikuju funkcionalne raspodele specifičnih ULS na podgrupe kod dve grupe neizvornih govornika, što dalje ukazuje na to da se na različite načine koriste specifični ULS kod dve grupe neizvornih govornika engleskog jezika.

Tabela 137. Unakrsna tabela između SCEP i RCEP korpusa i funkcionalnih podgrupa specifičnih ULS u odnosu na tokene

Grupa	Podgrupa	SCEP			RCEP		
		tokeni	%	R	tokeni	%	R
<i>ULS orijentisani ka istraživanju</i>	<i>ULS za lokaciju</i>	54	2.6%	-2.0*	20	4.4%	2.0*
	<i>ULS za proceduru</i>	153	7.5%	-13.2*	134	29.3%	13.2*
	<i>ULS za kvantifikaciju</i>	375	18.4%	-3.1*	113	24.7%	3.1*
	<i>ULS za opis</i>	488	23.9%	-3.3*	143	31.3%	3.3*
<i>ULS orijentisani ka tekstu</i>	<i>Tranzicioni signali</i>	263	12.9%	8.1*	0	0%	-8.1*
	<i>Rezultativni signali</i>	137	6.7%	5.7*	0	0%	-5.7*
	<i>Strukturni signali</i>	/	/	/	/	/	/
	<i>Ograničavajući signali</i>	355	17.4%	3.7*	47	10.3%	-3.7*
<i>ULS orijentisani ka učesniku</i>	<i>ULS stava</i>	51	2.5%	3.4*	0	0%	-3.4*
	<i>ULS uključivanja</i>	165	8.1%	6.3*	0	0%	-6.3*
Total		2041	100%		457	100%	

Napomena: % – procenat u odnosu na kolonu; R – korigovani standardizovani reziduali;

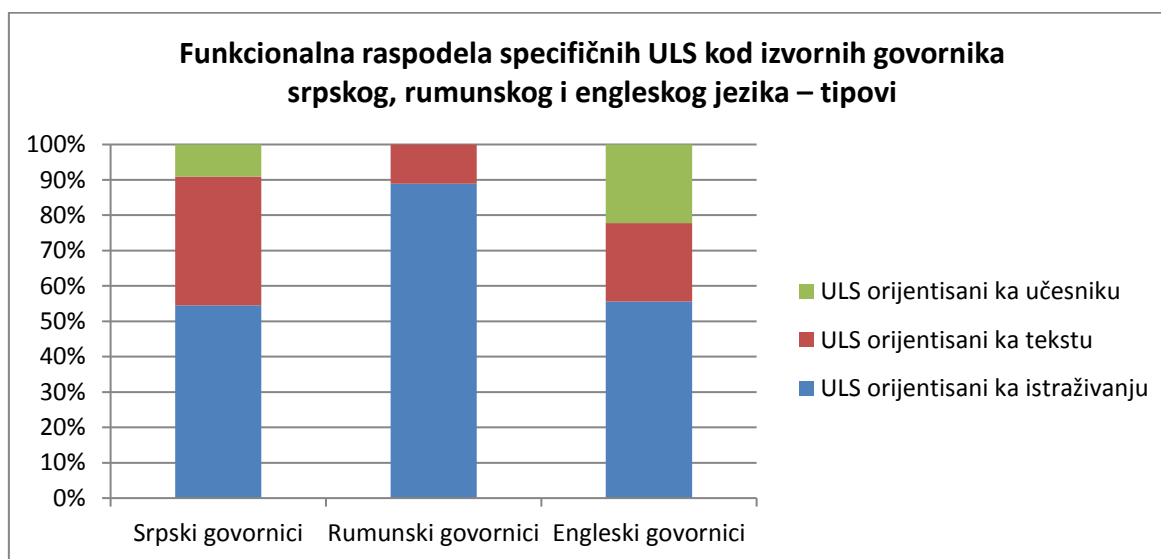
* – statistički značajni rezultati

Poređenje funkcionalne raspodele specifičnih ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski

Kada se sagleda funkcionalna raspodela tipova ULS u sva tri korpusa (v. Tabelu 138 i Grafikon 65), uviđamo da je sličnija raspodela ULS između ECEP i SCEP korpusa. Funkcionalna raspodela tipova ULS u RCEP korpusu razlikuje se jer funkcionalna kategorija *ULS orijentisani ka učesniku* nije zastupljena uopšte, a razlika između *ULS orijentisanih ka tekstu* i istraživanju je mnogo veća. Testovi statističke analize primenjeni na funkcionalnoj raspodeli tipova na grupe nisu pokazali statističku značajnost kada su se poredili ECEP i RCEP korpusi, ali jesu pokazali statistički značajnu razliku pri poređenju SCEP i RCEP korpusa. Ovakvi rezultati nas dalje navode na zaključak da se dve grupe neizvornih govornika značajno razlikuju pri funkcionalnoj upotrebi tipova specifičnih ULS, tj. da je važno uzeti ove informacije u obzir u kontekstima gde se zajedno podučavaju studenti sa različitim maternjim jezicima. Kada je reč o raspodeli tipova na funkcionalne podgrupe, važno je napomenuti da se statistički značajna razlika identifikovala jedino između ECEP i SCEP korpusa.

Tabela 138. *Funkcionalna raspodela specifičnih ULS kod izvornih govornika srpskog, rumunskog i engleskog jezika – tipovi*

Funkcionalna kategorija	Srpski govornici	Rumunski govornici	Engleski govornici
<i>Orijentisani ka istraživanju</i>	54.55%	88.89%	55.56%
<i>Orijentisani ka tekstu</i>	36.36%	11.11%	22.22%
<i>Orijentisani ka učesniku</i>	9.09%	0.00%	22.22%



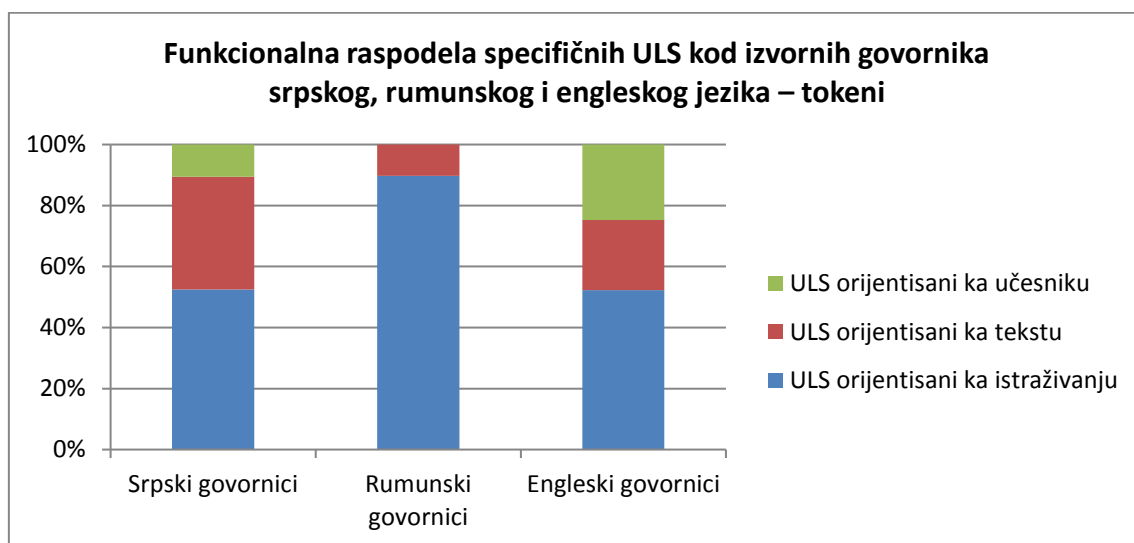
Grafikon 65. *Funkcionalna raspodela specifičnih ULS kod izvornih govornika srpskog, rumunskog i engleskog jezika – tipovi*

Funkcionalna raspodela specifičnih tokena na funkcionalne grupe u tri korpusa slična je funkcionalnoj raspodeli tipova, odnosno raspodela u RCEP korpusu se najviše razlikovala zbog nedostatka zastupljenosti tokena u funkcionalnoj grupi *ULS orijentisanih ka učesniku* i maloj zastupljenosti *ULS orijentisanih ka tekstu*, v. Tabelu 139 i Grafikon 66. Za razliku od tipova, statistička analiza tokena je pokazala značajne rezultate za funkcionalne grupe i podgrupe za sve relacije tri korpusa. Ustanovilo se da izvorni govornici srpskog imaju tendenciju da češće koriste *ULS orijentisane ka tekstu* uporedno i sa izvornim govornicima engleskog i sa izvornim govornicima rumunskog jezika. Takođe, utvrdilo se da izvorni govornici rumunskog jezika imaju

tendenciju da češće koriste ULS *orijentisane ka istraživanju* od izvornih govornika engleskog i od izvornih govornika srpskog jezika. Izvorni govornici engleskog jezika imaju tendenciju da češće koriste ULS *orijentisane ka učesniku* uporedno sa dve grupe neizvornih govornika engleskog jezika.

Tabela 139. Funkcionalna raspodela specifičnih ULS kod izvornih govornika srpskog, rumunskog i engleskog jezika – tokeni

Funkcionalna kategorija	SCEP	RCEP	ECEP
<i>Orijentisani ka istraživanju</i>	52.43%	89.72%	52.25%
<i>Orijentisani ka tekstu</i>	36.99%	10.28%	22.97%
<i>Orijentisani ka učesniku</i>	10.58%	0.00%	24.77%



Grafikon 66. Funkcionalna raspodela specifičnih ULS kod izvornih govornika srpskog, rumunskog i engleskog jezika – tokeni

Implikacije koje navedeni rezultati imaju na podučavanje ULS su sledeće:

- u radu sa izvornim govornicima srpskog jezika potrebno je podučavati alternativne načine prenošenja značenja koja iskazuju specifični *ULS orijentisani ka tekstu*, kako bi se uspešnije pokazala pripadnost relevantnoj delatnoj zajednici;
- u radu sa izvornim govornicima rumunskog jezika potrebno je podučavati alternativne načine prenošenja značenja koja iskazuju specifični *ULS orijentisani ka istraživanju*, kako bi se uspešnije pokazala pripadnost relevantnoj delatnoj zajednici;
- u radu sa obe grupe neizvornih govornika engleskog jezika potrebno je posvetiti posebnu pažnju podučavanju adekvatnih *ULS orijentisanim ka učesniku* jer imaju tendenciju da ih zanemare.

Poređenjem procentualne zastupljenosti specifičnih ULS po funkcionalnim grupama u celim korpusima SCEP i RCEP (Tabela 122. i 128.) uviđamo da postoje razlike u dva navedena korpusa. Naime, kod izvornih govornika srpskog jezika procentualno je najveća zastupljenost specifičnih ULS u funkcionalnoj kategoriji *ULS orijentisani ka učesniku*, dok se kod izvornih govornika rumunskog jezika nijedan specifičan ULS ne nalazi u toj kategoriji. U RCEP korpusu procentualno je najveća zastupljenost specifičnih ULS u funkcionalnoj kategoriji *ULS orijentisani ka istraživanju* sa 44,44% svih tipova u ovoj kategoriji. U SCEP korpusu *ULS orijentisani ka istraživanju* je druga kategorija po najvećem procentu zastupljenosti specifičnih ULS, gde 64,29% svih ULS u navedenoj

funkcionalnoj kategoriji čine *specifični tipovi*. Kao i za strukturalnu raspodelu, tako i za funkcionalnu raspodelu, uviđamo da su specifični ULS procentualno manje zastupljeni u RCEP korpusu kako po pitanju tipova tako i po pitanju tokena, što je i bilo očekivano a s obzirom na veći broj identifikovanih ULS u SCEP korpusu.

Činjenica da je statistička analiza dala značajne rezultate za tipove samo za funkcionalne grupe na relaciji SCEP i ECEP, dok je za tokene dala statistički značajne rezultate za sve grupe i podgrupe u različitoj meri ukazuje na to sledeće: *kada govorimo o funkcionalnoj raspodeli specifičnih ULS između tri korpusa značajne razlike u broju tipova ULS postoje samo između upotrebe izvornih govornika srpskog i engleskog, dok kada govorimo o stepenu upotrebe specifičnih ULS u funkcionalnim grupama i podgrupama statistički značajne razlike postoje na svim relacijama.*

4.11.2.3. Korišćenje podataka iz prethodnih kompatibilnih istraživanja za dodatno pročišćavanje spiskova specifičnih ULS za neizvorne govornike engleskog kojima je maternji jezik srpski ili rumunski

Prednost ovog istraživanja uporedno sa prethodnim iz oblasti koji su navedeni u poglavlju 1.4. *Transfer* je što postoje dva korpusa neizvornih govornika koja su formirana za potrebe istraživanja i koja su kompatibilna za poređenje. To dalje omogućava precizno izdvajanje ULS specifičnih za pojedinačne grupe neizvornih govornika. Neka od prethodnih istraživanja (Güngör & Uysal, 2020) za prikupljanje dokaza o međugrupnoj heterogenosti, odnosno izdvajanje specifičnog ULS za određenu grupu neizvornih govornika, poredila su izdvojene ULS u prethodnim istraživanjima u kojima korpusi (po veličini ili sadržaju) nisu bili dovoljno kompatibilni. Međutim, iako poređenje rezultata drugih istraživanja nije idealno rešenje i dalje nudi dodatne informacije koje mogu biti korisne. Poređenje rezultata kompatibilnih istraživanja se takođe može koristiti kao dodatni korak u istraživanju koje je primenilo optimalno rešenje za prikupljanje dokaza o međugrupnoj heterogenosti, poređenja kompatibilnih korpusa različitih grupa neizvornih govornika.

Uzimajući u obzir činjenicu da se u istraživanju Gungor (2016) ispituje upotreba ULS kod izvornih govornika engleskog i neizvornih govornika engleskog kojima je maternji jezik turski u naučnim radovima iz oblasti obrazovanja, istraživanje je adekvatno za poređenje po: žanru (naučni radovi) i oblasti (obrazovanje) tekstova koji sačinjavaju korpus, kao i dužini ULS koji se istražuju (ULS od četiri reči). Poređenjem rezultata dobijamo mogućnost da dodatno učvrstimo dokaze o međugrupnoj heterogenosti poređenjem specifičnih ULS sa onim koji su izdvojeni u još jednoj grupi neizvornih govornika. Sagledajući nekompatibilnosti u poređenju ovih istraživanja, možemo uvideti da je korpus neizvornih govornika engleskog korišćen u istraživanju Gungor (2016) duplo veći od korpusa neizvornih govornika engleskog koji su korišćeni u ovom istraživanju (korpus od milion reči uporedno sa korpusima od po pola miliona reči), kao i da su u istraživanju Gungor (2016) korišćeni blaži kriterijumi za izdvajanje ULS, odnosno 20 ponavljanja u milion reči i disperzija u 10% tekstova, uporedno sa kriterijumima u ovom istraživanju – minimalna frekvencija od 40 ponavljanja u milion reči i disperzija u 10% tekstova. Međutim, navedene razlike u parametrima za izdvajanje ULS zapravo čine dva istraživanja kompatibilnijim za poređenje, jer postoji tendencija da se u manjim korpusima identifikuje disproporcionalno veliki broj ULS (Begsten, 2019), što se kontrolisalo primenom strožih kriterijuma za izdvajanje, dok se u većem korpusu gde postoji manji rizik od disproporcionalno velike identifikacije ULS primenjivao blaži kriterijum za identifikaciju ULS. Na osnovu navedenih podataka, autorka disertacije smatra da bi poređenje dva navedena istraživanja doprinelo dodatnom pročišćavanju spiska specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski, kao i preciznosti tih podataka.

Kako bi spiskovi specifičnih ULS za neizvorne govornike engleskog kojima je maternji jezik srpski ili rumunski bili još precizniji, upoređeni su sa ULS koji su specifični za neizvorne govornike engleskog kojima je maternji jezik turski iz istraživanja Gungor (2016). Spisak specifičnih ULS za izvorne govornike turskog jezika sastoji se od 82 tipa ULS, v. Prilog 13. Poređenjem spiskova specifičnih ULS za izvorne govornike srpskog i turskog identifikovalo se deset tipova ULS koji su zajednički za oba spiska: *the aim of this, the characteristics of the, as one of the, as a part of, in line with the, is based on the, that there is no, it is possible to, it can be seen, it can be concluded that*. Nakon što se uklone gorenavedeni ULS, spisak specifičnih ULS za neizvorne govornike engleskog kojima je maternji jezik srpski sadrži 56 tipova i 1.689 tokena, v. Tabelu 140. Poređenjem spiskova specifičnih ULS za izvorne govornike rumunskog i turskog identifikovalo se dva tipa ULS koji su zajednički za oba spiska: *an important role in* i *with the help of*. Uklanjanjem ova dva ULS dobijamo spisak specifičnih ULS za izvorne govornike rumunskog jezika sa 16 tipova i 401 tokenom.

Tabela 140. Broj specifičnih ULS kod neizvornih govornika engleskog kojima je izvorni jezik srpski i rumunski posle poređenja sa rezultatima istraživanja Gungor (2016) – tipovi i tokeni

SCEP	56 (1689)
RCEP	16 (401)

4.11.2.4. Analiza potencijalne zastupljenosti specifičnih ULS za neizvorne govornike engleskog jezika u ECEP korpusu

Poređenjem izdvojenih spiskova ULS u tri korpusa došli smo do informacija o tome koji su identifikovani ULS upotrebljeni samo u pojedinačnim korpusima. Poređenjem rezultata sa kompatibilnim prethodnim istraživanjem, dodatno smo prečistili spiskove specifične za izvorne govornike srpskog i rumunskog jezika. Sledeći korak u prečišćavanju spiskova, kako bi se izdvojili ULS čija je upotreba rezultat negativnog transfera, jeste provera da li se ULS specifični za neizvorne govornike engleskog jezika uopšte pojavljuju u ECEP korpusu i ako je odgovor da – sa kojom frekventnošću. Treba imati na umu da je za identifikaciju ULS u ovom istraživanju potrebno da se spoj ponovi minimum 40 puta u milion reči kako bi se identifikovao kao *učestali leksički spoj* (ULS), što u našim korpusima od po pola miliona reči znači minimum 20 ponavljanja. Ako se specifičan ULS uopšte ne pojavljuje u ECEP korpusu, možemo pretpostaviti da je verovatnoća postojanja negativnog transfera veća nego u recimo ULS koji se u ECEP korpusu pojavljuje 18 puta, odnosno kojem je samo dva ponavljanja nedostajalo da se izdvoji i u ECEP korpusu kao ULS. U konačan spisak specifičnih ULS za koje se smatra da su pod uticajem negativnog transfera izdvojiće se oni koji se pojavljuju u ECEP korpusu pet i manje puta.

U Tabeli 141 predstavljena je frekvencija zastupljenosti specifičnih ULS za SCEP korpus u ECEP korpusu. *Italikom* su označeni ULS koji se nijednom ne pojavljuju u ECEP korpusu a za koje je mogućnost uticaja negativnog transfera najveća, **podebljani su** ULS koji se u korpusu pojavljuju pet i manje puta i gde je, takođe, velika verovatnoća uticaja negativnog transfera. Uviđamo da se 14 ULS specifičnih za izvorne govornike srpskog jezika uopšte ne pojavljuju u ECEP korpusu, dok se 25 ULS pojavljuju sa frekvencijom od pet i manje ponavljanja. Konačan prečišćen spisak specifičnih ULS za izvorne govornike srpskog jezika sastoji sa od 39 ULS. Provera postojanja prevodnih ekvivalenata na srpskom jeziku i njihova zastupljenost u referentnom korpusu vršiće se na konačnom prečišćenom spisku.

Tabela 141. *Frekvencija zastupljenosti specifičnih ULS za SCEP korpus u ECEP korpusu*

SCEP	ULS	ECEP	SCEP	ULS	SCEP
82	on the basis of (the)	2	26	<i>that there are no</i>	0
55	for the implementation of	1	26	the aim of the	2
54	the implementation of the	10	26	<i>the application of the</i>	0
52	in the implementation of	3	26	<i>the results show that</i>	0
	a statistically significant				
49	difference	2	26	to the conclusion that	1
48	in the teaching process	2	25	a small number of	11
41	a large number of	6	25	<i>of the concept of</i>	0
41	the results of this	18	25	of the importance of	11
				<i>statistically significant</i>	
40	as well as in	11	25	<i>difference in</i>	0
39	of the quality of	4	25	<i>the assessment of the</i>	0
38	as well as to	13	25	the largest number of	1
	<i>(statistically) significant</i>				
38	<i>difference between the</i>	0	24	<i>at the faculty of</i>	0
33	for the purpose of	11	24	<i>having in mind the</i>	0
33	<i>having in mind that</i>	0	24	<i>of the faculty of</i>	0
33	of this research is	1	24	the influence of the	2
33	<i>we can conclude that</i>	0	24	<i>there is no statistically</i>	0
32	in addition to the	16	23	based on the results (of)	6
31	a part of the	6	23	in such a way	1
30	in the area of	12	23	on the quality of	4
28	in comparison to the	5	22	in other words the	4
27	are shown in table	5	21	<i>under the influence of</i>	0
27	in the domain of	1	20	as well as a	7
27	significant differences in the	4	20	can be found in	5
27	the quality of teaching	3	20	of education in the	3
27	with the aim of	7	20	of the education system	6
26	as well as their	9	20	significant difference in the	1
	statistically significant				
26	differences in	4	20	that they do not	3
82	on the basis of (the)	2	20	the effects of the	8

Frekvencija zastupljenosti specifičnih ULS za RCEP korpus u ECEP korpusu predstavljena je u Tabeli 142. *Italikom* su označeni ULS koji se nijednom ne pojavljuju u ECEP korpusu, a **podebljani su** ULS koji se u korpusu pojavljuju pet i manje puta. Uviđamo da se dva ULS uopšte ne pojavljuju u ECEP korpusu, dok se sedam pojavljuje sa frekvencijom od pet i manje ponavljanja. Stoga se konačan prečišćen spisak specifičnih ULS za izvorne govornike rumunskog jezika sastoji od devet ULS.

Tabela 142. Frekvencija zastupljenosti specifičnih ULS za RCEP korpus u ECEP korpusu

RCEP	ULS	ECEP	RCEP	ULS	SCEP
41	the way in which	6	24	than half of the	1
31	<i>of the educational process</i>	0	24	the number of students	10
29	taking into account the	4	23	the existence of a	1
28	<i>the point of view (of)</i>	0	20	in order to achieve	5
27	the level of the	3	20	in the number of	16
25	more than half of	1	20	the majority of the	6
25	the purpose of the	12	20	the same time the	4
24	by the ministry of	6	20	the development of a	10

4.11.3. Međujezička podudarnost performansi

Kako bismo prikupili dokaze za međujezičku podudarnost performansi, proverili smo da li ULS koji su specifični za izvorne govornike srpskog i rumunskog jezika imaju prevodne ekvivalente na srpskom i rumunskom jeziku. Za navedeni korak oslanjamo se na Skopus teoriju i koncept funkcionalne ekvivalentnosti (Reiss et al., 2014). Glavni cilj prevoda ULS u ovom segmentu istraživanja je identifikacija načina na koji se prenose značenja i funkcije specifičnih ULS za neizvorne govornike u srpskom i rumunskom jeziku. Predloženi funkcionalni prevodni ekvivalenti specifičnih ULS za izvorne govornike srpskog jezika predstavljeni su u Tabeli 143. Kao što se može uočiti, predlažemo dva i više funkcionalnih ekvivalenata za 12 ULS specifičnih za SCEP korpus. Preciznije, dva ULS imaju predložena po tri funkcionalna ekvivalenta a deset ULS imaju predložena po dva funkcionalna ekvivalenta. Drugim rečima, za 39 specifičnih ULS za SCEP korpus predloženo je ukupno 53 funkcionalna ekvivalenta. Poređenjem zastupljenosti prevodnih funkcionalnih ekvivalenata uviđamo da se 16 specifičnih ULS koristi više u SCEP korpusu nego u *SrpKor2013*, 17 se koristi manje u SCEP korpusu nego u *SrpKor2013*, dok se šest ULS uopšte ne nalaze u *SrpKor2013* ni u jednom funkcionalnom ekvivalentu.

Zbog razlika u veličinama korpusa koji se porede (SCEP korpus je manje veličine i sastoji se od 503.435 reči, *SrpKor* je veći i sastoji se od 3.657.797 reči) primenjen je statistički test *log-likelihood* kako bi se ustanovilo da li su navedene razlike značajne. U okviru testa uzima se u obzir veličina korpusa te možemo da uvidimo da li se određeni ULS značajno više ili manje koristi u određenom korpusu. Statistički test je pokazao da se četiri ULS statistički značajno češće koriste u *SrpKor2013* korpusu nego u SCEP korpusu (*on the basis of (the), in the domain of, the influence of the, of education in the*). Navedeni podaci dalje ukazuju na to da postoje jaki dokazi da neizvorni govornici engleskog jezika kojima je maternji jezik srpski koriste četiri navedena ULS zbog uticaja negativnog transfera. Za tri ULS nije identifikovana statistički značajna razlika u zastupljenosti između SCEP i *SrpKor2013* korpusa (*of the quality of, the application of the, under the influence of*), što nas takođe navodi na zaključak da postoje jaki dokazi da neizvorni govornici engleskog jezika kojima je maternji jezik srpski koriste i ova tri ULS zbog uticaja negativnog transfera. Ukupno 32 specifična ULS za izvorne govornike srpskog se statistički značajno više koriste u SCEP korpusu nego u *SrpKor2013*, što nas dalje navodi do zaključka da je za tih 32 ULS manja verovatnoća da je njihova zastupljenost u produkciji izvornih govornika srpskog jezika rezultat uticaja negativnog transfera. Naravno, na dobijene rezultate utiče i činjenica da referentni korpus izvornih govornika srpskog jezika *SrpKor2013* nije formiran za potrebe ovog istraživanja kao što su ECEP, SCEP i RCEP korpusi, te nije u istoj meri kompatibilan sa ECEP, SCEP i RCEP korpusima kao što su oni međusobno. Pretpostavljamo da na manju zastupljenost navedena 32 ULS takođe utiče i činjenica da u *SrpKor2013* pored naučnih radova iz oblasti obrazovanja imamo i naučne radove iz drugih oblasti, kao i činjenice da pored naučnih radova u korpusu imamo i naučnopopularne tekstove. Upotreba ULS se umnogome razlikuje u zavisnosti od žanra i oblasti

tekstova, stoga bi za još preciznije podatke bilo poželjno da se u budućim istraživanjima i za referentni korpus formira specifičan korpus koji će se poklapati i po veličini, broju tekstova, žanru, oblasti i načinu identifikovanja ULS.

Tabela 143. *Predloženi funkcionalni prevodni ekvivalenti specifičnih ULS za izvorne govornike srpskog jezika*

ULS (engl.)	SCEP	ULS (srp.)	SrpKor2013	LL
on the basis of (the)	82	na osnovu	790	- 6.16*
for the implementation of	55	za primenu	20	+150.95****
		za implementaciju	11	+176.16****
in the implementation of	52	u primeni	20	+140.16****
		u implementaciji	3	+197.59****
a statistically significant difference	49	statistički značajna razlika	0	+ 207.41****
		statistički značajne razlike	1	+ 197.86****
in the teaching process	48	u obrazovnom procesu	1	+193.67 ****
		u procesu učenja	0	+203.18****
of the quality of	39	kvaliteta	240	+ 0.94
<i>(statistically) significant difference between the</i>	38	statistički značajne razlike između	0	+203.18****
<i>having in mind that</i>	33	imajući na umu	28	+ 62.72****
of this research is	33	ovog istraživanja je	0	+139.68****
<i>we can conclude that</i>	33	možemo zaključiti da	17	+79.94****
		može se zaključiti da	21	+72.90****
		zaključujemo da	9	+98.35****
in comparison to the	28	uporedno sa	1	+110.08****
		u poređenju sa	58	+ 24.88****
are shown in table	27	prikazani su u tabeli	2	+100.25****
		dat/ dati/ data u tabeli	7	+81.51****
in the domain of	27	u oblasti	325	- 7.18**
		u domenu	42	+ 32.70****
significant differences in the	27	značajne razlike u	2	+100.25****
the quality of teaching	27	kvalitet podučavanja	0	+114.29****
		kvalitet učenja	0	+114.29****
		kvalitet nastave	0	+114.29****
statistically significant differences in	26	statistički značajne razlike u	1	+101.76****
<i>that there are no</i>	26	da ne postoje	7	+77.75****
the aim of the	26	cilj je	74	+ 14.44***
		cilj (bilo koja reč) je	63	+18.71****
<i>the application of the</i>	26	primena	130	+2.85
<i>the results show that</i>	26	rezultati pokazuju	9	+72.46****
to the conclusion that	26	do zaključka da	28	+42.46****
<i>of the concept of</i>	25	koncepta	52	+22.10****
<i>statistically significant difference in</i>	25	statistički značajna razlika	0	+105.82****
<i>the assessment of the</i>	25	proceni (imenica)	34	+43.14****
the largest number of	25	najveći broj	47	+24.90****
<i>at the faculty of</i>	24	na fakultetu	8	+67.65****

Tabela 143. Predloženi funkcionalni prevodni ekvivalenti specifičnih ULS za izvorne govornike srpskog jezika

		fakultet	95	+6.33*
<i>having in mind the</i>	24	imajući na umu	28	+37****
<i>of the faculty of</i>	24	na fakultetu	8	+67.65****
		fakultet	95	+6.33*
the influence of the	24	uticaj (lema)	1215	-176.65****
<i>there is no statistically</i>	24	nema statistički	1	+93.45****
in such a way	23	na takav način	22	+40.64****
on the quality of	23	na kvalitet	13	+53.60****
in other words the	22	po rečima	46	+19.32****
<i>under the influence of</i>	21	pod uticajem	186	-0.74
can be found in	20	mogu se naći u	7	+55.55****
of education in the	20	u obrazovanju	21	+33.09****
		obrazvanje (lema)	667	- 75.63****
significant difference in				
the	20	značajna razlika	3	+67.62****
that they do not	20	da oni ne	5	+60.92****

Uzimajući u obzir formu predloženih funkcionalnih prevodnih ekvivalenata, uviđamo da se samo pet poklapaju po broju reči sa specifičnim ULS (*(statistically) significant difference between the – statistički značajne razlike između, we can conclude that – može se zaključiti da, are shown in table – prikazani su u tabeli, statistically significant differences in – statistički značajne razlike u, can be found in – mogu se naći u*). Očekivalo se da bi prevodni ekvivalenti koji se poklapaju po formi bili zastupljeni i u SCEP i u *SrpKor2013* sa velikim brojem primera, jer bi sličnost u formi potencijalno pospešila postojanje negativnog transfera, međutim, to nije slučaj u podacima u ovom istraživanju. Štaviše, predloženi prevodni ekvivalenti za sva četiri specifična ULS se manje koriste u *SrpKor2013* nego što se ti specifični ULS koriste u SCEP korpusu. Najveći broj predloženih funkcionalnih prevodnih ekvivalenata sastoji se od dve reči (24), praćeni onima koji se sastoje od tri reči (16), dok su predloženi funkcionalni prevodni ekvivalenti koji se sastoje od jedne reči zastupljeni sa osam primera.

Osvrnimo se posebno na 14 specifičnih ULS za izvorne govornike srpskog jezika koji se nisu našli u ECEP korpusu ni sa jednim primerom. Očekuje se da je veća verovatnoća negativnog uticaja transfera kod navedenih spojeva te da bi oni bili više zastupljeni u referentnom korpusu na srpskom jeziku. Analiza zastupljenosti navedenih spojeva nam, međutim, daje neočekivane rezultate gde samo jedan prevodni ekvivalent ispunjava očekivanja i statistički je značajno više zastupljen u referentnom korpusu (*under the influence of – pod uticajem*), za jedan prevodni ekvivalent se ne pokazuju statistički značajne razlike u upotrebi u dva korpusa (*the application of the – primena*), dok se za preostalih 12 identifikovala statistički značajno učestalija upotreba u SCEP korpusu.

Detaljnijom analizom sedam ULS koji pokazuju najjače dokaze za učestaliju upotrebu kod izvornih govornika srpskog jezika zbog uticaja transfera, uviđamo da najveći broj pripada strukturalnoj grupi *fragmenti sa predloškom frazom (under the influence of, of education in the, of the quality of, on the basis of (the), in the domain of)*, dok se preostala dva specifična ULS nalaze u strukturalnoj grupi *imeničke strukture (the influence of the, the application of the)*. Funkcionalna raspodela pokazuje da je pet od navedenih ULS u funkcionalnoj grupi *ULS orijentisani ka istraživanju, preciznije u pogrupama ULS za proceduru (under the influence of, the influence of the, the application of the)* i *ULS za opis (on the basis of (the), in the domain of)*. Navedeni podaci se poklapaju sa zapaženim tendencijama izvornih govornika srpskog da učestalije koriste specifične ULS u funkcionalnim podgrupama *ULS za proceduru* i *ograničavajućih signala*.

Predloženi funkcionalni ekvivalenti za ULS koji su specifični za neizvorne govornike engleskog kojima je maternji jezik rumunski predstavljeni su u Tabeli 144. Za devet specifičnih ULS predloženo je 16 funkcionalnih prevodnih ekvivalenata jer tri ULS imaju po dva predložena ekvivalenta a dva ULS imaju po tri predložena ekvivalenta. Kao referentni korpus za rumunski jezik koristi se ECPRES korpus. Poređenjem zastupljenosti specifičnih ULS za izvorne govornike rumunskog jezika i njihovih predloženih funkcionalnih prevodnih ekvivalenata u RCEP i referentnom korpusu za rumunski jezik, uviđamo da je šest prevodnih ekvivalenata učestalije upotrebljeno u EXPRESS korpusu, četiri učestalije upotrebljeno u RCEP korpusu, dok se šest prevodnih ekvivalenata uopšte ne nalaze u EXPRES korpusu. Za specifični ULS *the same time the* nijedan od dva predložena funkcionalna prevodna ekvivalenta nije zastupljen u EXPRES korpusu.

Tabela 144. *Predloženi funkcionalni prevodni ekvivalenti specifičnih ULS za izvorne govornike rumunskog jezika*

ULS (engl.)	RCEP	ULS (rum.)	EXPRESS	LL
<i>of the educational process</i>	31	procesul educării	0	+ 138.78****
		procesul de educare	1	+ 130.07****
taking into account the	29	având în vedere	287	- 0.77
<i>the point of view (of)</i>	28	din punctul de vedere	181	+ 1.52
the level of the	27	nivelul de	244	- 0.14
more than half of	25	mai mult de jumate	0	+ 111.89****
		mai mult de jumătate din	15	+ 62.35****
than half of the	24	mai puțin de jumătate din ⁴⁰	5	+ 81.88****
the existence of a	23	existența a	16	+ 53.74****
		existența	220	- 0.38
		existența unui	51	+ 22.72****
in order to achieve	20	pentru a reliza	0	+ 89.51****
		pentru a obține	69	+ 10.29**
		în ordine pentru a realiza	0	+ 89.51****
the same time the	20	în același timp ce	0	+ 89.51****
		în același timp	0	+ 89.51****

Zbog neujednačene veličine RCEP i EXPRES korpusa (RCEP je manji i sastoji se od 501.826 reči, EXPRES je veći i sastoji se od 4.201.363 reči) neophodno je primeniti statistički test kako bi se utvrdilo da li su uočene razlike značajne. Primena statističkog testa *log-likelihood* utvrđeno je da sedam ULS statistički značajno učestalije koriste u RCEP korpusu uporedno sa njihovih 12 predloženih prevodnih ekvivalenata, dok se za četiri ULS i njihovih predloženih prevodnih ekvivalenata nije pokazala statistički značajna razlika u upotrebi. Na osnovu navedenih podataka možemo da zaključimo da postoje jaki dokazi da neizvorni govornici engleskog jezika kojima je maternji jezik rumunski koriste četiri ULS u kojima se nije identifikovala statistički značajna razlika zbog uticaja negativnog transfera. Za pet ULS čiji se predloženi prevodni ekvivalenti učestalije javljaju u EXPRES korpusu možemo da kažemo da postoje slabiji dokazi da je njihova upotreba kod izvornih govornika rumunskog jezika pod uticajem negativnog transfera, dok za ULS čiji se predloženi prevodni ekvivalenti uopšte ne nalaze u EXPRES korpusu možemo da kažemo da nedostaju dokazi o međujezičkoj podudarnosti performansi i da je prisustvo negativnog uticaja transfera malo verovatno. Svakako je značajno naglasiti da RCEP korpus i EXPRES korpus nisu neujednačeni samo po veličini i broju radova koji ih sačinjavaju već i po oblastima istraživanja koja se obrađuju u radovima. Neke od navedenih razlika u zastupljenosti predloženih funkcionalnih prevodnih ekvivalenata mogu da budu i posledica navedenih razlika u korpusima.

⁴⁰ Odnosi se na manje od pola jer ne postoji neutralna verzija.

Uzimajući u obzir formu predloženih funkcionalnih prevodnih ekvivalenata specifičnih ULS za izvorne govornike rumunskog jezika, uviđamo da se dva predložena prevodna ekvivalenta formalno poklapaju sa istim brojem reči (*more than half of – mai mult de jumate, the same time the – în același timp ce*), predložen je još jedan funkcionalni prevodni ekvivalent od četiri reči ali za ULS od pet reči tako da se ipak ne poklapaju (*the point of view (of) – din punctul de vedere*). Kao ni kod izvornih govornika srpskog jezika, ni kod izvornih govornika rumunskog jezika nisu se izdvojili jači dokazi o postojanju negativnog transfera kod specifičnih ULS koji se poklapaju po formi sa predloženim prevodnim ekvivalentima. Za razliku od podataka predloženih prevodnih ekvivalenata u srpskom jeziku, u rumunskom imamo i tri prevodnih ekvivalenata sa pet reči, što je verovatno posledica tipoloških razlika između dva jezika. Najveći broj specifičnih ULS (pet) u predloženom funkcionalnom prevodnom ekvivalentu ima tri reči, praćeni sa četiri specifična ULS koji u predloženim prevodnim ekvivalentima imaju dve reči i jednim primerom koji u predloženom prevodnom ekvivalentu ima jednu reč.

Od dva specifična ULS za izvorne govornike rumunskog jezika (*of the educational process, the point of view (of)*) koji uopšte nisu bili prisutni u ECEP korpusu, te je bila veća verovatnoća da će se za njih identifikovati jači uticaj negativnog transfera, jedan (*the point of view (of)*) jeste ispunio očekivanja, te se za njegov predloženi prevodni ekvivalent u EXPRES korpusu nisu identifikovane značajne razlike, čime se obezbedio jak dokaz za negativan uticaj transfera. Drugi ULS (*of the educational process*) ima dva predložena funkcionalna prevodna ekvivalenta, od kojih jedan uopšte nije prisutan u EXPRES korpusu, dok je drugi prisutan sa samo jednim primerom. Navedeni podaci dalje ukazuju da se za ULS *of the educational process* identifikovao slab dokaz za međujezičku podudarnost performansi.

Sagledanjem strukturalne raspodele četiri ULS za koje su se identifikovali jaki dokazi za međujezičku podudarnost performansi, odnosno za prisutnost negativnog transfera, zaključujemo da je većina u strukturalnoj grupi *imeničke strukture (the point of view (of), the level of the, the existence of a)*, dok se jedan nalazi u grupi *glagolske strukture (taking into account the)*. Navedeni podaci su u skladu sa tendencijama izvornih govornika rumunskog jezika da učestalije koriste specifične ULS u grupama *imeničke strukture* i *glagolske strukture*. Takođe se poklapa i sa tendencijom izvornih govornika rumunskog jezika da u svim identifikovanim ULS najviše koriste ULS u strukturalnoj kategoriji *imeničke strukture* uporedno sa izvornim govornicima engleskog jezika i neizvornim govornicima engleskog jezika kojima je maternji jezik srpski. Sagledanjem funkcionalne raspodele uviđamo da se sva četiri izdvojena ULS nalaze u funkcionalnoj grupi *ULS orijentisani ka istraživanju*, u podgrupama *ULS za opis (the point of view (of), the level of the, the existence of a)*, i *ULS za proceduru (taking into account the)*. Činjenica da se sva četiri ULS nalaze u funkcionalnoj grupi *ULS orijentisani ka istraživanju* poklapa se sa tendencijom izvornih govornika rumunskog jezika da u specifičnim ULS koriste značajno više ULS u grupi *ULS orijentisani ka istraživanju* od izvornih govornika engleskog jezika.

Na osnovu izdvojenih podataka o sve tri grupe dokaza za utvrđivanje uticaja transfera, odnosno unutargrupne homogenosti, međugrupne heterogenosti i međugrupne podudarnosti performansi, možemo zaključiti da su se dokazi za postojanje uticaja negativnog transfera utvrdili za 33 ULS u produkciji izvornih govornika srpskog jezika, dok su se jaki dokazi za uticaja negativnog transfera utvrdili za 10 ULS u produkciji izvornih govornika srpskog jezika. To bi dalje značilo da ULS koji se koriste zbog uticaja negativnog transfera u SCEP korpusu sačinjavaju 29,73% od svih ULS u tom korpusu. Kada bi se odlučili za strože kriterijume identifikacije ULS i prihvatili samo jake dokaze, onda bi ULS koji se koriste zbog uticaja negativnog transfera u SCEP korpusu sačinjavali 9,01% svih ULS u SCEP korpusu. Takođe, zaključujemo da su se dokazi za postojanje negativnog transfera utvrdili za osam ULS u produkciji izvornih govornika rumunskog jezika, dok su jaki dokazi da uticaj negativnog transfera u produkciji izvornih govornika rumunskog jezika utvrdili za četiri ULS. Navedeni podaci dalje ukazuju na to da se zbog negativnog uticaja transfera koristi

13,33% svih ULS identifikovanih u RCEP korpusu. Primenom strožeg kriterijuma, kojim bi se prihvatili samo ULS sa jakim dokazima, došli bismo do podatka da se 6,67% od svih ULS u RCEP korpusu koristi zbog uticaja negativnog transfera.

Sedma radna hipoteza je potvrđena dobijenim rezultatima o uticaju negativnog transfera: *ULS koji se koriste zbog uticaja transfera sačinjavaće mali procenat (do 40%) ULS identifikovanih u korpusima neizvornih govornika engleskog jezika kojima su maternji jezici srpski ili rumunski u naučnim radovima iz oblasti obrazovanja*. Naime, nezavisno od toga da li zauzimamo blaži ili stroži kriterijum za identifikaciju uticaja transfera, procenat ULS koji su pod uticajem negativnog transfera ne prevazilaze 40% svih identifikovanih ULS kod neizvornih govornika kojima je maternji jezik srpski ili rumunski.

Poređenjem strukturalne i funkcionalne raspodele ULS koji zadovoljavaju samo stroge uslove identifikacije negativnog transfera, uviđamo da se kod obe grupe neizvornih govornika najviše koriste ULS iz funkcionalne grupe *ULS orijentisani ka istraživanju*, odnosno u podgrupama *ULS za opis* i *ULS za proceduru*. Kod strukturalne raspodele uvidele su se razlike, naime, izvorni govornici srpskog jezika koristili su ULS u kategorijama *fragmenti sa predloškom frazom* i *imeničke strukture*, dok su izvorni govornici rumunskog koristili *imeničke strukture* i *glagolske strukture*. Navedeni podaci nas dalje navode do zaključka da neizvorni govornici engleskog jezika kojima je maternji jezik srpski i rumunski u naučnim radovima iz oblasti obrazovanja imaju zajedničku tendenciju da koriste određene *ULS orijentisane ka istraživanju* zbog negativnog uticaja transfera.

5. Diskusija

U okviru diskusije rezultata istraživanja u ovoj disertaciji sagledaćemo dobijene rezultate i analizirati ih u okviru šireg konteksta prethodnih istraživanja i trenutnog stanja u oblasti istraživanja ULS. Razmotrićemo implikacije i značaj dobijenih rezultata. Podaci prate logičan sled kojim se istraživanje odvijalo, te prvo analiziramo *zastupljenost ULS kod izvornih i neizvornih govornika*, nakon čega slede segmenti koji se bave *strukturalnom raspodelom ULS, funkcionalnom raspodelom ULS i interakcijom strukturalne i funkcionalne raspodele ULS*, koje su cilj imale cilj bolje razumevanje karakteristika upotrebe ULS kod izvornih i neizvornih govornika, dok je poređenje izdvojenih rezultata omogućilo identifikaciju razlika i sličnosti u upotrebi ULS. Strukturalna i funkcionalna analiza *zajedničkih ULS* između izvornih i neizvornih govornika, omogućila nam je da identifikujemo ULS koje neizvorni govornici uspešno koriste, a analiza zajedničkih ULS za dve grupe neizvornih govornika pomogla nam je u razumevanju zajedničkih karakteristika neizvornih govornika. Analiza *uticaja transfera na produkciju neizvornih govornika* pružila nam je priliku da preciznije razvrstamo koje su karakteristike upotrebe ULS kod neizvornih govornika prisutne kao opšte karakteristike neizvornih govornika, a koje su posledica negativnog uticaja transfera. Analiza upotrebe najzastupljenijih obrazaca produbila je naše razumevanje svih gorenavedenih aspekata, odnosno omogućila je bolje razumevanje karakteristika upotrebe ULS kod izvornih i neizvornih govornika engleskog jezika, kao i njihove razlike i sličnosti u upotrebi ULS. U završnom delu ovog segmenta disertacije, predstavićemo metodičke smernice za podučavanje upotrebe ULS u okviru nastave akademskog engleskog jezika, engleskog jezika za posebne namene ili podučavanja/usavršavanja (budućih) prevodilaca za prevođenje naučnih radova u oblasti obrazovanja, a koje su se izdvojile na osnovu analize podataka u ovoj disertaciji.

5.1. Zastupljenost ULS kod izvornih i neizvornih govornika engleskog jezika

U prethodnim istraživanjima koja su ispitivala upotrebu ULS kod izvornih i neizvornih govornika engleskog jezika pokazalo se da neizvorni govornici imaju tendenciju da koriste veći broj tipova i tokena ULS, iako to nije bio slučaj u svim obrađenim prethodnim istraživanjima. Preciznije, od

devet izdvojenih istraživanja, sa relevantnim podacima, u sedam se veći broj ULS koristi kod neizvornih govornika engleskog jezika, dok se u dva veći broj ULS izdvojio kod izvornih govornika engleskog jezika. Isti princip je uočen i u istraživanju koje je sprovedeno u ovoj disertaciji, odnosno, veći se broj ULS identifikovao kod neizvornih govornika engleskog jezika. Na osnovu tog podatka potvrđena je četvrta radna hipoteza: *Veći broj tipova i tokena ULS biće identifikovan kod neizvornih govornika engleskog jezika uporedno sa izvornim govornicima engleskog jezika u naučnim radovima iz oblasti obrazovanja.* U korpusu izvornih govornika engleskog jezika izdvojeno je 27 ULS, u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik srpski izdvojeno je 111 ULS, dok je u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik rumunski izdvojeno 60 ULS. Procentualno gledano, kod izvornih govornika srpskog jezika uočavamo uvećanje upotrebe ULS u odnosu na izvorne govornike engleskog jezika od 311%, dok se kod izvornih govornika rumunskog jezika uočava uvećanje od 122%. Kada uporedimo navedene podatke iz ovog istraživanja sa podacima iz prethodnih istraživanja koji su sumirani u Tabeli 20. *Broj izdvojenih ULS kod izvornih i neizvornih govornika i procenat povećanja ili umanjenja broja u odnosu na korpus izvornih govornika u istraživanjima koja se bave upotrebom ULS u naučnim radovima,* uviđamo da neizvorni govornici engleskog jezika kojima je maternji jezik srpski pokazuju najveće uvećanje u broju izdvojenih ULS uporedno sa izvornim govornicima engleskog jezika. U istraživanju Lazić (2017), kod izvornih govornika srpskog zabeleženo je uvećanje od 635%, dok je kod izvornih govornika srpskog jezika u istraživanju u ovoj disertaciji zabeleženo uvećanje od 311%. Možemo pretpostaviti da je uočena razlika u procentima uvećanja kod dve grupe izvornih govornika srpskog jezika posledica činjenice da je u istraživanju Lazić (2017) korpus neizvornih govornika engleskog jezika kojima je maternji jezik srpski mnogo manji od korpusa izvornih govornika engleskog, što je potencijalno uticalo na neproporcijalno veliki broj izdvojenih ULS (Kortez, 2004). Posle izvornih govornika srpskog jezika, najveće se uvećanje zapaža kod izvornih govornika turskog u naučnim radovima iz oblasti obrazovanja (206%) – u istraživanju Gungor i Uisal (2016), koji su praćeni izvornim govornicima rumunskog jezika u ovom istraživanju. Na osnovu navedenih podataka, možemo zaključiti da izvorni govornici srpskog jezika imaju tendenciju da koriste veći broj ULS u naučnim radovima, uporedno sa drugim grupama neizvornih govornika engleskog jezika sa drugim maternjim jezicima. Takođe, možemo pretpostaviti da se veći broj ULS izdvaja kod neizvornih govornika engleskog jezika koji pišu naučne radove iz oblasti obrazovanja uporedno sa drugim naučnim oblastima.

Neočekivan podatak u ovom istraživanju, što se tiče zastupljenosti ULS kod neizvornih govornika, jeste velika razlika u broju identifikovanih ULS kod neizvornih govornika srpskog jezika i rumunskog. Na osnovu ove razlike u podacima nije potvrđena peta radna hipoteza: *Sličan broj tipova i tokena biće identifikovan kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski i rumunski u naučnim radovima iz oblasti obrazovanja.* Na osnovu navedene razlike u broju izdvojenih ULS kod dve grupe neizvornih govornika engleskog jezika sa različitim maternjim jezikom u istraživanju gde su ispitani korpusi ujednačeni a isti kriterijumi primenjeni za izdvajanje ULS, možemo zaključiti da maternji jezik neizvornih govornika predstavlja značajan faktor kod zastupljenosti ULS u produkciji neizvornih govornika. Potencijalno objašnjenje za navedenu razliku u izdvojenom broju ULS kod izvornih govornika srpskog jezika i rumunskog jezika možemo da nađemo u različitom stepenu oslanjanja na usluge prevodilaca. Naime, u prethodnim istraživanjima praksi objavljivanja naučnih radova koje su prisutne kod srpskih (Murešan i Perez-Ljatanda, 2014) i rumunskih (Igrutinović, 2021) naučnika, pokazalo se da srpski naučnici imaju veću tendenciju da se oslanjaju na usluge prevodilaca. Navedeni podatak potencijalno može da objasni razlike u broju izdvojenih ULS zato što, kao što su prethodna istraživanja pokazala, postoji tendencija upotrebe većeg broja ULS u prevedenim tekstovima uporedno sa neprevedenim (Biel, 2017; Siao, 2010; Li, 2013; Ebeling, 2022). Zbog aktuelnih izdavačkih praksi u kojima se ne navode informacije o tome da li su objavljeni naučni radovi prevedeni i ko je prevodilac, nemoguće je u SCEP i RCEP korpusima proveriti predstavljeno potencijalno objašnjenje. Kako bi se sa

sigurnošću moglo utvrditi da li su u pitanju razlike u karakteristikama upotrebe različitih grupa neizvornih govornika ili razlikama u zastupljenosti prevedenih tekstova, potrebno je sprovesti nova istraživanja u kojima će se kao relevantan faktor uzeti u obzir da li je naučni rad preveden ili ne. Štaviše, situacija se dodatno komplikuje sve popularnijom upotrebom veštačke inteligencije u pisanju tekstova i mašinskog prevođenja čiji uticaji na upotrebu ULS još nisu utvrđeni. Kao što je ranije navedeno, radovi koji su korišćeni u korpusima u ovom istraživanju sastavljeni su od tekstova koji su napisani pre nego što je veštačka inteligencija postala javno dostupna.

5.2. Strukturalna raspodela ULS

Analiza strukturalne raspodele ULS kod izvornih i neizvornih govornika u ovom istraživanju se koristi kao jedan od načina identifikacije karakteristika upotrebe ULS kod izvornih i neizvornih govornika, kao i za identifikaciju razlika u njihovoj upotrebi. Pri analizi podataka, ULS su podeljeni u relevantne strukturalne grupe i podgrupe po tipovima i tokenima, određena je njihova procentualna zastupljenost, odnosno urađena je deskriptivna statistika, nakon koje su urađeni i statistički testovi kako bi se utvrdilo da li su identifikovane razlike statistički značajne. U ovom segmentu stavićemo rezultate dobijene u ovoj disertaciji u širi kontekst prethodnih istraživanja, odnosno trenutnog stanja u istraživačkoj oblasti i sagledati značaj i doprinos novih podataka.

Strukturalnom raspodelom ULS na grupe, uvideli smo da je i kod izvornih govornika engleskog jezika i kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski, najzastupljenija strukturalna grupa *fragmenti sa predloškom frazom*, koja je praćena *imeničkim strukturama* i *glagolskim strukturama*, dok su *druge strukture* bile najmanje zastupljene. Strukturalnu raspodelu ULS poredimo sa pet prethodnih istraživanja (Pan i sar., 2016; Gungör & Uysal, 2016; Gungör, 2016; Lazić, 2017; Al Fedžri, 2020), koja koriste iste grupe za strukturalnu raspodelu, odnosno dovoljno kompatibilne taksonomije. Raspodela zastupljenosti strukturalnih kategorija u istraživanju u ovoj disertaciji se ne poklapa sa rezultatima dobijenim u prethodnim istraživanjima na nekoliko načina. Sa jedne strane, pregledom strukturalne raspodele ULS kod neizvornih govornika engleskog jezika koji su do sada ispitani u kompatibilnim istraživanjima, uviđamo da je najčešća strukturalna raspodela *glagolske strukture* kao najzastupljenija grupa, praćena *imeničkim strukturama*, *fragmentima sa predloškom frazom* i na kraju najmanje zastupljenim *drugim strukturama*. Navedena raspodela je prisutna kod neizvornih govornika engleskog jezika kojima je maternji jezik kineski i turski, odnosno u istraživanjima Pan i sar. (2016), Gungor i Uisal (2016) i Gungor (2016), v. Tabelu 22. *Strukturalna raspodela tipova i tokena ULS kod izvornih i neizvornih govornika u istraživanjima koja se bave upotrebom ULS u naučnim radovima*. Kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski i indonežanski, u prethodnim istraživanjima, uočena je raspodela zastupljenosti strukturalnih kategorija u kojoj su najzastupljeniji *fragmenti sa predloškom frazom*, praćeni *glagolskim strukturama* pa *imeničkim strukturama*, dok su najmanje zastupljene *druge strukture*. Kao što se može uočiti, raspodela zastupljenosti koja se uviđa kod neizvornih govornika kojima je maternji jezik srpski ili rumunski u ovom istraživanju se ne poklapa ni sa jednom od navedenih prethodnih. Sa druge strane, kada sagledamo strukturalne raspodele ULS kod izvornih govornika engleskog jezika, možemo uvideti da je raspodela zastupljenosti strukturalnih kategorija u kojoj je najzastupljenija kategorija *fragmenti sa predloškom frazom*, praćena *imeničkim strukturama* i *glagolskim strukturama*, a sa najmanje zastupljenim *drugim strukturama* – prisutna u tri prethodna istraživanja: Gungor i Uisal (2016), Gungor (2016) i Al Fedžri (2020). Stoga, uočavamo da se ista raspodela nalazi i kod izvornih govornika engleskog jezika i kod neizvornih govornika engleskog jezika u istraživanju u ovoj disertaciji. To nas dalje dovodi do zaključka da izvorni govornici srpskog i rumunskog u naučnim radovima pokazuju bolju usklađenost sa normama žanra koje su postavili izvorni govornici engleskog jezika uporedno sa izvornim govornicima kineskog, turskog i indonežanskog jezika. Navedeni zaključak je, isto tako, podržan činjenicom da se u istraživanju

Lazić (2017) zapaža ista raspodela zastupljenosti strukturalnih grupa kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik srpski u naučnim radovima iz biotehničkih nauka. Drugim rečima, izvorni govornici srpskog jezika su i tom prilikom pokazali sposobnost da prate strukturalne norme u upotrebi ULS koje su postavili izvorni govornici. Jedan od potencijalnih objašnjenja prednosti izvornih govornika srpskog i rumunskog može da bude činjenica da su ta dva jezika deo indoevropske porodice jezika isto kao i engleski, za razliku od kineskog koji je pripadnik sino-tibetanskih jezika, turskog, koji je pripadnik atlajskih jezika i indonežanskog, koji je pripadnik austronezijskih jezika. Veća tipološka sličnost jezika koji spadaju u istu porodicu jezika i manja kulturološka udaljenost potencijalno mogu da objasne veću doslednost izvornih govornika srpskog i rumunskog jezika normama koje su postavili izvorni govornici engleskog jezika.

Naravno, činjenica da izvorni i neizvorni govornici engleskog jezika u ovom istraživanju pokazuju istu raspodelu zastupljenosti strukturalnih kategorija, ne znači da ne postoje razlike u njihovoj strukturalnoj upotrebi ULS. Statistička analiza *tokena* pokazala je da izvorni govornici srpskog jezika koriste značajno više *tokena* ULS u strukturalnoj grupi *druge strukture* od izvornih govornika engleskog i rumunskog jezika, izvorni govornici rumunskog koriste značajno više *tokena* u grupi *imeničke strukture* i *fragmenti sa predloškom frazom* od izvornih govornika engleskog i srpskog jezika, dok izvorni govornici engleskog jezika koriste značajno više tipova u grupi *fragmenti sa predloškom frazom* od izvornih govornika srpskog jezika a značajno više tipova u grupi *glagolske strukture* od izvornih govornika rumunskog jezika. Statistička analiza tipova nije pokazala značajne razlike u strukturalnoj raspodeli između izvornih i neizvornih govornika, kao ni između dve grupe neizvornih govornika. Statistička analiza pomoću koje smo dobili navedene podatke je hi-kvadrat test.

Kako bismo mogli da poredimo ove rezultate sa onim dobijenim u prethodnim istraživanjima, prvo je važno da sagledamo koje su se statističke analize primenjivale u tim istraživanjima i na koji način, tj. da li su primenjene na strukturalnim grupama, podgrupama ili u oba slučaja, na tipovima, tokenima ili oba. Razmatranjem prethodnih istraživanja koja su kompatibilna za poređenje strukturalne raspodele, uviđamo da se hi-kvadrat test primenio u istraživanju Lazić (2017), koja ga je primenila na strukturalne grupe tipova i tokena ali ne i na strukturalne podgrupe. U istraživanju Pan i sar. (2016) primenjen je test verodostojnosti (engl. *log-likelihood*) na tipove i tokene u svim podgrupama ali ne i posebno za grupe. Al Fedžri i sar. (2020) u svom istraživanju primenjuju test verodostojnosti na strukturalne grupe i podgrupe ali samo na tokene ULS, dok Gungor (2016) primenjuje isti test na strukturalne grupe i podgrupe i za tipove i za tokene. U istraživanju Gungor i Uisal (2016) nisu primenjene statističke analize za identifikaciju značajnosti rezultata. Na osnovu izdvojenih informacija, uviđamo da rezultate o statistički značajnim razlikama u strukturalnoj raspodeli na grupe možemo porediti sa istraživanjima Lazić (2017), Gungor (2016) i Al Fedžri i sar. (2020), dok rezultate o statistički značajnim razlikama u strukturalnoj raspodeli na podgrupe možemo porediti sa istraživanjima Pan i sar. (2016), Gungor (2016) i Al Fedžri i sar. (2020).

Poređenjem rezultata o statistički značajnim razlikama u strukturalnoj raspodeli u istraživanju u ovoj disertaciji sa rezultatima u istraživanju koje je sprovedla Lazić (2017), uviđamo da se rezultati delimično poklapaju. Naime, neizvorni govornici engleskog jezika kojima je maternji jezik srpski u oba istraživanja znatno više koriste ULS iz strukturalne kategorije *druge strukture* uporedno sa izvornim govornicima engleskog jezika. Stoga možemo da pretpostavimo da nezavisno od naučne oblasti u naučnim radovima neizvorni govornici srpskog jezika imaju tendenciju da koriste isuviše učestalo ULS iz strukturalne kategorije *druge strukture*. S obzirom na činjenicu da neizvorni govornici engleskog jezika statistički značajno učestalije koriste ULS u strukturalnoj grupi *druge strukture* od izvornih govornika engleskog i u istraživanjima Gungor (2016) i Al Fežri i sar. (2020), možemo zaključiti da je učestalija upotreba ULS u strukturalnoj grupi *druge strukture* u naučnim

radovima opšta karakteristika neizvornih govornika. Rezultati se ne poklapaju kod zastupljenosti strukturalne grupe *fragmenti sa predloškom frazom*, koji se kod izvornih govornika srpskog jezika u ovom istraživanju koriste značajno manje nego kod izvornih govornika engleskog jezika, dok je situacija u istraživanju Lazić (2017) obrnuta. Možemo pretpostaviti da su takve razlike prisutne zbog drugačije naučne oblasti kojom su se bavili naučni radovi koji sačinjavaju korpuse u dva istraživanja, ali dodatna su ispitivanja neophodna kako bi navedena pretpostavka postala tvrdnja. Kada uporedimo rezultate sa prethodnim istraživanjem u kojem se ispitala ista naučna oblast obrazovanja (Güngör, 2016), uviđamo da je statistički značajna razlika prisutna za sve strukturalne grupe u tom istraživanju i da se rezultati za izvorne govornike srpskog i rumunskog istraživanja poklapaju sa produkcijom izvornih govornika turskog samo u kategoriji *druge fraze*. U istraživanju Al Fedžri i sar. (2020) identifikovana je statistički značajna razlika u upotrebi strukturalnih kategorija *fragmenti sa predloškom frazom*, *glagolske strukture* i *druge strukture* između izvornih govornika engleskog i neizvornih govornika indonežanskog. Sve tri navedene kategorije više koriste neizvorni govornici indonežanskog, što se poklapa sa rezultatima za izvorne govornike rumunskog za grupu *fragmenti sa predloškom frazom*, a sa izvornim govornicima srpskog za *druge strukture*.

Kao što se može primetiti iz navedenih podataka, postoji velika varijacija u strukturalnoj raspodeli ULS kod neizvornih govornika engleskog jezika sa različitim maternjim jezikom. Navedeno dalje potvrđuje i zapaženu veću razliku u strukturalnoj raspodeli ULS između dve grupe neizvornih govornika u istraživanju u ovoj disertaciji, nego između neizvornih govornika engleskog jezika i izvornih govornika engleskog jezika. Ako je zaista slučaj da su varijacije između različitih grupa neizvornih govornika veće nego sličnosti, još važniji postaje proces identifikacije karakteristika upotrebe ULS kod različitih grupa neizvornih govornika u naučnim radovima, kao i primena tih uvida za unapređenje obrazovne prakse ili za implementaciju u programe stručnog usavršavanja kako za (buduće) članove akademske zajednice, tako i za (buduće) prevodioce.

Sagledanjem strukturalne raspodele ULS u podgrupe u ovom istraživanju uvideli smo da su i kod izvornih govornika engleskog jezika i kod dve grupe neizvornih govornika engleskog jezika najzastupljenije strukturalne podgrupe *predloške fraze sa of fragmentom*, *praćene imeničkim frazama sa of fragmentom* i *drugim predloškim frazama*. Iste tri strukturalne podgrupe su se izdvojile kao najzastupljenije kod izvornih i neizvornih govornika engleskog jezika u istraživanju Lazić (2017) i kod izvornih govornika engleskog jezika u istraživanju Gungor (2016), dok su se izvorni govornici turskog jezika razlikovali po trećoj najviše zastupljenoj strukturalnoj podgrupi koja je kod njih *uvodna it struktura + glagolska ili pridevska fraza*. Navedeni podaci ukazuju da neizvorni govornici kojima je maternji jezik srpski i rumunski uspešnije prate norme izvornih govornika i za strukturalnu raspodelu u podgrupe od neizvornih govornika engleskog kojima je maternji jezik turski.

Kada je reč o raznolikosti strukturalne raspodele, uviđamo da se kod izvornih govornika engleskog jezika koristi najmanji broj strukturalnih podgrupa – sedam, dok se kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski koristi devet strukturalnih podgrupa, a kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski koristi se jedanaest. Navedena razlika u broju podgrupa između dve grupe neizvornih govornika može da se objasni značajno većom zastupljenošću ULS iz strukturalne grupe *glagolske strukture* kod izvornih govornika srpskog jezika nego kod izvornih govornika rumunskog, jer se u toj kategoriji nalazi najveći broj podgrupa. Proverom raznolikosti strukturalne raspodele i u prethodnim istraživanjima uviđamo da se i u njima zapaža tendencija neizvornih govornika engleskog jezika da koriste veći broj strukturalnih podgrupa. Preciznije, u istraživanju Pan i sar. (2016) izvorni govornici koriste devet strukturalnih podgrupa, dok neizvorni govornici koriste deset, u istraživanju Gungor i Uisal (2016) izvorni govornici koriste sedam strukturalnih podgrupa, dok neizvorni govornici engleskog

jezika koriste dvanaest; u istraživanju Gungor (2016) izvorni govornici koriste osam strukturalnih podgrupa, dok neizvorni govornici koriste petnaest, u istraživanju Al Fedžri i sar. (2020) izvorni govornici engleskog jezika koriste osam strukturalnih podgrupa dok neizvorni govornici koriste deset strukturalnih podgrupa. Od istraživanja koja smo poredili jedino se u istraživanju Lazić (2017) isti broj strukturalnih podgrupa koristi i kod izvornih i kod neizvornih govornika engleskog jezika.

5.3. Funkcionalna raspodela ULS

Rezultati koji se tiču funkcionalne raspodele ULS dobijenih u ovom istraživanju poređiće se sa narednim relevantnim istraživanjima koja su kompatibilna za poređenje: Gungor i Uisal (2016), Gungor (2016), Lazić (2017), Esfandiari i Barbari (2017) i Al Fedžri i sar. (2020). Na osnovu funkcionalne raspodele *tipova* ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik srpski i rumunski, u ovom istraživanju, ustanovilo se da je, za sve tri grupe govornika, najzastupljenija funkcionalna grupa *ULS orijentisani ka istraživanju*, praćena *ULS orijentisanim ka tekstu*, dok je najmanje zastupljena grupa *ULS orijentisani ka učesniku*. Ista raspodela zastupljenosti funkcionalnih grupa se pronalazi i kod izvornih govornika engleskog jezika u naučnim radovima iz oblasti obrazovanja u istraživanjima koje je sproveo Gungor (2016; Gungör & Uysal, 2016), kao i u naučnim radovima iz oblasti biotehničkih nauka (Lazić, 2017). Drugačija se raspodela zastupljenosti funkcionalnih kategorija zapaža u naučnim radovima iz oblasti telekomunikacije (Pan i sar., 2016), psihologije (Esfandiari i Barbari, 2017) i primenjene lingvistike (Al Fedžri i sar., 2020). U tim radovima su *ULS orijentisani ka tekstu* najzastupljeniji, praćeni *ULS orijentisanim ka istraživanju*. Drugim rečima, dva trenda raspodele funkcionalnih kategorija tipova ULS se uviđaju kod izvornih govornika u zavisnosti od naučne oblasti u kojoj se pišu radovi. Poređenjem raspodele zastupljenosti funkcionalnih grupa između izvornih i neizvornih govornika engleskog jezika, uviđamo da se raspodele poklapaju i kod izvornih i kod neizvornih govornika u četiri (Pan i sar., 2016; Lazić, 2017; Esfandiari i Barbari, 2017; Al Fedžri i sar., 2020) od šest prethodnih istraživanja, a kada se tom broju doda i istraživanje u ovoj disertaciji –onda je to pet od sedam istraživanja. Na osnovu navedenih podataka možemo zaključiti da neizvorni govornici engleskog jezika u većini naučnih oblasti uspešno prate norme raspodele zastupljenosti funkcionalnih grupa ULS koje su postavili izvorni govornici (svakako uspešnije nego što prate raspodelu zastupljenosti strukturalnih grupa). Dva rada (Gungör & Uysal, 2016; Gungör, 2016) u kojima se ne uklapa raspodela zastupljenosti funkcionalnih grupa jesu iz oblasti obrazovanja – prema njima, neizvorni govornici engleskog jezika kojima je maternji jezik turski najviše koriste *ULS orijentisane ka tekstu*.

Kada sagledamo funkcionalnu raspodelu *tokena* ULS u istraživanju u ovoj disertaciji, uviđamo da se raspodela zastupljenosti funkcionalnih grupa ne poklapa kod izvornih i neizvornih govornika engleskog jezika. Kod izvornih govornika engleskog jezika najzastupljeniji su *ULS orijentisani ka istraživanju*, dok su kod neizvornih govornika engleskog jezika kojima su maternji jezici srpski ili rumunski najzastupljeniji *ULS orijentisani ka tekstu*. Stoga, možemo zaključiti da i izvorni govornici srpskog i rumunskog jezika imaju poteškoće u praćenju normi funkcionalne upotrebe ULS koje su postavili izvorni govornici kada je reč o učestalosti upotrebe ULS. To nas dalje dovodi do pretpostavke da je neizvornim govornicima posebno izazovno pratiti norme o funkcionalnoj upotrebi ULS u naučnim radovima iz oblasti obrazovanja. Neizvorni govornici engleskog jezika kojima je maternji turski jezik su, takođe, imali poteškoće u tom pogledu, a one se nisu uočile kod neizvornih govornika koji su pisali naučne radove iz drugih naučnih oblasti. Dodatna istraživanja koja bi ispitala upotrebu ULS kod drugih grupa neizvornih govornika u naučnim radovima iz oblasti obrazovanja neophodna su a kako bi se navedena pretpostavka proverila.

Da bi se ispitala značajnost razlika u funkcionalnoj raspodeli ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog kojima je maternji jezik srpski i rumunski, primenjen je hi-kvadrat test. Rezultati su pokazali da obe grupe neizvornih govornika imaju tendenciju da isuviše učestalo koriste *ULS orijentisane ka tekstu*, dok nedovoljno učestalo koriste *ULS orijentisane ka učesniku*. Kada posvetimo pažnju i podgrupama, uviđamo, preciznije, da obe grupe neizvornih govornika isuviše učestalo koriste funkcionalnu podgrupu *ograničavajući signali*, dok nedovoljno učestalo koriste podgrupe *ULS za lokaciju*, *strukturalni signali* i *ULS stava*. Međutim, pored gorenavedenih sličnosti kod dve grupe neizvornih govornika koje uviđamo njihovim poređenjem sa izvornim govornicima engleskog jezika, kada poredimo funkcionalnu raspodelu ULS na grupe i podgrupe, uočavamo značajne razlike između izvornih govornika srpskog i rumunskog jezika. Izvorni govornici srpskog jezika značajno učestalije koriste *ULS orijentisane ka učesniku*, dok izvorni govornici rumunskog jezika značajno učestalije koriste *ULS orijentisane ka istraživanju*. Preciznije informacije dobijamo kada uzmemo u obzir funkcionalne podgrupe, gde se pokazalo da izvorni govornici srpskog učestalije koriste obe podgrupe *ULS orijentisanih ka učesniku*, dok izvorni govornici koriste značajno učestalije podgrupe *ULS za lokaciju* i *proceduru* iz grupe *ULS orijentisanih ka istraživanju*. Štaviše, iako se dve grupe neizvornih govornika ne razlikuju značajno po pitanju funkcionalne grupe *ULS orijentisanih ka tekstu*, kada sagledamo funkcionalne podgrupe *ULS orijentisanih ka tekstu* uviđamo da izvorni govornici srpskog jezika koriste više *tranzicionih* i *rezultativnih signala*, dok izvorni govornici rumunskog koriste značajno više *ograničavajućih signala*.

Pre nego što detaljnije uporedimo podatke o razlikama u funkcionalnoj raspodeli ULS sa prethodnim relevantnim istraživanjima, sagledaćemo koji su se statistički testovi koristili u njima. U istraživanjima Lazić (2017) i Esfandiari i Barbari (2017) primenjen je hi-kvadrat test, međutim, u istraživanju Lazić (2017) primenjen je samo na funkcionalnim grupama, dok je u istraživanju Esfandiari i Barbari (2017) primenjen na funkcionalnim grupama i podgrupama. U istraživanjima Gungor (2016) i Al Fedžri i sar. (2020), primenjen je test verodostojnosti (engl. *log-likelihood*) na funkcionalne grupe i podgrupe, dok je u istraživanju Pan i sar. (2016) isti test primenjen samo na funkcionalne podgrupe. U istraživanju Gungor i Uisal (2016) nisu primenjeni testovi statističke značajnosti.

Poređenjem rezultata uviđamo da se i u istraživanju Lazić (2017), kao i u istraživanju u ovoj disertaciji uviđa značajno veća učestalost upotrebe *ULS orijentisanih ka tekstu* kod neizvornih govornika engleskog jezika, dok se značajno veća učestalost *ULS orijentisanih ka učesniku* uviđa kod izvornih govornika engleskog jezika. Drugim rečima, rezultati se poklapaju. Međutim, isti rezultati se samo delimično poklapaju sa istraživanjem Esfandiari i Barbari (2017), gde izvorni govornici engleskog jezika značajno učestalije koriste i *ULS orijentisane ka tekstu* i *ULS orijentisane ka učesniku* od neizvornih govornika engleskog jezika kojima je maternji jezik persijski, i istraživanja Gungor (2016) i Al Fedžri i sar. (2020) gde neizvorni govornici značajno učestalije koriste i *ULS orijentisane ka tekstu* i *ULS orijentisane ka učesniku*. Rezultati istraživanja Ran i sar. (2016) su suprotni od onih koji su izdvojeni u ovom istraživanju, odnosno neizvorni govornici engleskog jezika značajno učestalije koriste *ULS orijentisane ka učesniku*, dok izvorni govornici engleskog jezika značajno učestalije koriste *ULS orijentisane ka tekstu*. Stoga, uzimajući sve gorenavedene rezultate u obzir, možemo da zaključimo da neizvorni govornici engleskog jezika kojima je maternji jezik srpski imaju tendenciju da koriste više *ULS orijentisanih ka tekstu* a manje *ULS orijentisanih ka učesniku* od izvornih govornika engleskog jezika u naučnim radovima iz različitih oblasti, dok neizvorni govornici engleskog jezika sa različitim maternjim jezikom imaju tendenciju da koriste više *ULS orijentisanih ka tekstu* nezavisno od maternjeg jezika neizvornih govornika. Autorka disertacije pretpostavlja da se navedene razlike u rezultatima u dosadašnjim relevantnim istraživanjima što se tiče raspodele funkcionalnih grupa uočavaju prevashodno zbog uticaja različitih maternjih jezika i različitih stepena tipoloških razlika koji su prisutni između

engleskog jezika i ispitanim maternjim jezicima neizvornih govornika, međutim, dalja istraživanja su neophodna kako bi se pretpostavka pokazala kao ispravna ili neispravna.

5.4. Interakcija strukturalne i funkcionalne raspodele ULS

Analiza interakcije strukturalne i funkcionalne raspodele ULS kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski i rumunski pokazuje da je u funkcionalnoj grupi *ULS orijentisani ka istraživanju* najzastupljenija strukturalna grupa *imeničke strukture*, u funkcionalnoj grupi *ULS orijentisani ka tekstu* najzastupljenija strukturalna grupa *fragmenti sa predložkom frazom*, dok je u funkcionalnoj grupi *ULS orijentisani ka učesniku* najzastupljenija strukturalna grupa *glagolske strukture* za sve tri grupe govornika. Poređenjem ovih podataka sa istraživanjem Pan i sar. (2016. godina – ovo je jedino od prethodnih relevantnih istraživanja u kojem je urađena ovakva analiza) uviđamo da je ista strukturalna i funkcionalna zastupljenost ULS prisutna i kod izvornih govornika engleskog jezika pri njihovoj produkciji u naučnim radovima iz oblasti telekomunikacije. Međutim, kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski uviđa se razlika u strukturalnoj raspodeli ULS unutar funkcionalne grupe *ULS orijentisani ka tekstu*, u kojoj je najzastupljenija strukturalna kategorija *glagolske strukture*. Kako bismo bolje razumeli ovu uočenu razliku, sagledajmo detaljnije strukturalnu raspodelu ULS unutar funkcionalne grupe *ULS orijentisani ka tekstu* u oba istraživanja.

U istraživanju u ovoj disertaciji kod izvornih govornika engleskog jezika u funkcionalnoj grupi *ULS orijentisani ka tekstu* ne javlja se strukturalna grupa *glagolske strukture* uopšte, dok je u istraživanju Pan i sar. (2016) to druga najzastupljenija strukturalna kategorije unutar navedene funkcionalne kategorije. Navedene informacije nas dovode do zaključka da se norme upotrebe strukturalnih i funkcionalnih kategorija blago razlikuju u naučnim radovima iz oblasti obrazovanja i telekomunikacije. Međutim, svakako se i kod neizvornih govornika u istraživanju u ovoj disertaciji uviđa isuviše učestala upotrebe ULS iz strukturalne grupe *glagolske strukture* u funkcionalnoj grupi *ULS orijentisani ka tekstu*, prevashodno od strane izvornih govornika srpskog a u manjoj meri izvornih govornika rumunskog jezika. Ove informacije možemo da povežemo sa činjenicom da neizvorni govornici srpskog i rumunskog imaju tendenciju da isuviše učestalo koriste ULS iz funkcionalne grupe *ULS orijentisani ka tekstu*, pa pretpostavimo da je to delom zbog prekomerne upotrebe glagolskih struktura unutar funkcionalne grupe *ULS orijentisani ka tekstu*.

5.5. Zajednički ULS

Izdvajanjem zajedničkih ULS i kada sagledamo njihove karakteristike, možemo da dobijemo uvid u to koji su ULS, i iz kojih strukturalnih i funkcionalnih grupa, neizvorni govornici uspešno usvojili i primenili. U istraživanju u ovoj disertaciji pored zajedničkih ULS između izvornih i neizvornih govornika, izdvojeni su i zajednički ULS samo za neizvorne govornike, pomoću kojih se mogu bolje razumeti zajedničke karakteristike različitih grupa neizvornih govornika.

Poređenjem spiskova izdvojenih ULS za izvorne govornike engleskog jezika i neizvorne govornike engleskog kojim je maternji jezik srpski i rumunski u naučnim radovima iz oblasti obrazovanja, ustanovili smo da je 15 ULS zajedničko za sve tri grupe govornika, 18 ULS zajedničko za izvorne govornike engleskog i neizvorne govornike engleskog kojima je maternji jezik srpski, 15 ULS zajedničko za izvorne govornike engleskog i neizvorne govornike engleskog kojima je maternji jezik rumunski, a 27 zajedničko za dve grupe neizvornih govornika. Zajednički ULS između izvornih govornika engleskog i srpskog jezika sačinjavaju 67% ECEP korpusa a 16% SCEP korpusa. Iz navedenih informacija možemo videti da neizvorni govornici engleskog jezika kojima je maternji jezik srpski pokrivaju dosta veliki procenat ULS koje koriste izvorni govornici,

međutim, oni sačinjavaju veoma mali procenat svih ULS koje koriste izvorni govornici srpskog jezika. Zajednički ULS između izvornih govornika engleskog i rumunskog jezika sačinjavaju 56% ECEP korpusa a 25% RCEP korpusa. Zajednički ULS za dve grupe neizvornih govornika sačinjavaju 24% SCEP korpusa i 45% RCEP korpusa. Navedeni podaci ukazuju na to da veći procenat svih ULS u korpusima neizvornih govornika sačinjavaju ULS koji su zajednički za neizvorne govornike nego ULS koji se poklapaju sa izvornim govornicima.

Poređenjem gorenavedenih rezultata sa podacima o procentualnoj zastupljenosti zajedničkih ULS u prethodnim relevantnim istraživanjima koji su predstavljeni u Tabeli 110. uviđamo da zajednički ULS za izvorne govornike engleskog i srpskog jezika sačinjavaju najveći procenat korpusa izvornih govornika engleskog jezika koji je zabeležen u prethodnim relevantnim istraživanjima do sada. Sledeći najzastupljeniji procenat se odnosi na zajedničke ULS kod izvornih govornika engleskog i srpskog u istraživanju Lazić (2017), u kojem sačinjavaju 46% svih ULS u korpusu izvornih govornika engleskog jezika. Na osnovu ovih podataka uviđamo da zbog tendencije izvornih govornika srpskog jezika da koriste veliki broj ULS, oni takođe i pokrivaju najveći broj ULS koji su prisutni i kod izvornih govornika. Izvorni govornici rumunskog jezika su odmah „za petama” izvornih govornika srpskog u istraživanju Lazić (2017), jer zajednički ULS za izvorne govornike rumunskog i engleskog jezika sačinjavaju 45% korpusa izvornih govornika engleskog jezika. Analizom procentualne zastupljenosti zajedničkih ULS u korpusu neizvornih govornika uviđamo da oni sačinjavaju najmanji procenat kod izvornih govornika srpskog jezika u istraživanju Lazić (2017), kao i da izvorni govornici srpskog i rumunskog u ovom istraživanju nisu među tri istraživanja u kojima se zapaža najmanja procentualna zastupljenost zajedničkih ULS u korpusima neizvornih govornika.

Strukturalna analiza zajedničkih ULS u istraživanju u ovoj disertaciji pokazala je da sve tri relacije zajedničkih ULS (zajednički za izvorne govornike engleskog i srpskog, zajedničke za izvorne govornike engleskog i rumunskog, zajedničke za dve grupe neizvornih govornika) prate raspodele koje su prisutne u strukturalnoj raspodeli celih korpusa. Odnosno, najzastupljenija strukturalna grupa u zajedničkim ULS za sve tri relacije su *fragmenti sa predloškom frazom*, praćeni *imeničkim strukturama* i *glagolskim strukturama*, dok su *druge strukture* najmanje zastupljene. Na osnovu navedenih podataka možemo da zaključimo da neizvorni govornici engleskog jezika kojima je maternji jezik srpski ili rumunski nemaju tendenciju da uspešnije koriste određenu strukturalnu grupu, već da određeni procenat ULS iz svake grupe uspešno koriste, tj. u skladu sa normama koje su postavili izvorni govornici. Što se tiče ULS koji su zajednički za neizvorne govornike engleskog jezika a koji se ne pojavljuju kod izvornih govornika engleskog jezika, takođe se nije pokazala tendencija da se ističe određena strukturalna grupa, već se ravnomerno koriste atipični ULS u svakoj strukturalnoj grupi u skladu sa strukturalnom raspodelom ULS u celim korpusima.

Poređenjem zastupljenosti tokena zajedničkih ULS uvideli smo da neizvorni govornici engleskog jezika kojima je maternji jezik srpski i rumunski imaju tendenciju da koriste blago veći procenat tokena u strukturalnoj grupi *fragmenti sa predloškom frazom* uporedno sa izvornim govornicima engleskog jezika. Poređenjem zastupljenosti tokena zajedničkih ULS za dve grupe neizvornih govornika uvideli smo da izvorni govornici srpskog značajno više koriste tokene u strukturalnim grupama *glagolske strukture* i *druge strukture*, dok izvorni govornici rumunskog jezika značajno više koriste tokene iz strukturalne kategorije *fragmenti sa predloškom frazom*. To nas dovodi do zaključka da čak i kada su korišćeni tipovi identični kod različitih grupa govornika, postoji tendencija da se oni koriste sa različitom učestalošću.

Funkcionalna raspodela zajedničkih tokena ULS za sve tri relacije pokazuju da su najzastupljeniji *ULS orijentisani ka tekstu*, praćeni *ULS orijentisanim ka istraživanju*, dok su *ULS orijentisani ka učesniku* najmanje zastupljeni. Navedena raspodela funkcionalnih kategorija po zastupljenosti u

skladu je sa raspodelom tokena u celim SCEP i RCEP korpusima. Raspodela tipova zajedničkih ULS prati gorenavedenu raspodelu za ULS koje dele izvorni i neizvorni govornici, dok je funkcionalna raspodela zajedničkih ULS za dve grupe neizvornih govornika drugačija i u njoj su najzastupljeniji *ULS orijentisani ka istraživanju*. Pregledom razlika u upotrebi tokena zajedničkih ULS za izvorne govornike engleskog jezika i neizvorne govornike engleskog kojima je maternji jezik srpski i rumunski, uviđamo da obe grupe neizvornih govornika imaju tendenciju da učestalije koriste *ULS orijentisane ka tekstu* od izvornih govornika engleskog, dok izvorni govornici engleskog jezika učestalije koriste *ULS orijentisane ka istraživanju*. Navedeni rezultati ukazuju na činjenicu da čak i kada neizvorni govornici koriste najoptimalnije tipove ULS za pokazivanje pripadnosti određenoj delatnoj zajednici i dalje imaju tendenciju da ih isuviše ili nedovoljno učestalo koriste. Poređenjem raspodele tokena zajedničkih ULS kod dve grupe neizvornih govornika uviđamo da izvorni govornici srpskog jezika značajno učestalije koriste *ULS orijentisane ka učesniku*, dok izvorni govornici rumunskog jezika značajno učestalije koriste *ULS orijentisane ka istraživanju*.

5.6. Specifični ULS

Specifični ULS za izvorne govornike engleskog jezika i za dve grupe neizvornih govornika engleskog jezika su izdvojeni u segmentu za analizu uticaja negativnog transfera u procesu prikupljanja dokaza za međugrupnu heterogenost. Međutim, strukturalna i funkcionalna analiza specifičnih ULS je, takođe, veoma značajna za identifikaciju karakteristika upotrebe ULS kod različitih grupa neizvornih govornika, kao i za identifikaciju načina na koji neizvorni govornici odstupaju od normi postavljenih od strane izvornih govornika. U ECEP korpusu izdvojeno je devet specifičnih ULS, u SCEP korpusu izdvojeno je 66 specifičnih ULS, dok je u RCEP korpusu izdvojeno 18 specifičnih ULS.

Strukturalna raspodela specifičnih ULS pokazala je da su kod izvornih govornika engleskog jezika i izvornih govornika srpskog najzastupljeniji *fragmenti sa predloškom frazom*, praćeni *imeničkim strukturama*, dok su kod izvornih govornika rumunskog jezika najzastupljenije *imeničke strukture* praćene *fragmentima sa predloškom frazom*. Statistička analiza pokazala je da se strukturalna raspodela specifičnih ULS kod izvornih govornika rumunskog jezika više razlikuje od strukturalne raspodele specifičnih ULS kod izvornih govornika engleskog jezika, nego što se razlikuju specifični ULS kod izvornih govornika srpskog jezika. Najveće razlike u strukturalnoj raspodeli specifičnih ULS uvidele su se između dve grupe neizvornih govornika, gde su se značajne razlike pokazale za sve tri strukturalne grupe. Od prethodnih istraživanja jedino se u Gungor i Uisal (2020) sprovela strukturalna i funkcionalna analiza specifičnih ULS za neizvorne govornike engleskog jezika. Rezultati strukturalne analize specifičnih ULS za izvorne govornike turskog su pokazali da je najzastupljenija kategorija *glagolske strukture* praćena *imeničkim strukturama*. Možemo da primetimo da se strukturalna raspodela specifičnih ULS ne poklapa ni za jednu od tri ispitane grupe neizvornih govornika što dalje ide u pretpostavki da su specifični ULS za neizvorne govornike pod uticajem negativnog transfera.

Funkcionalna raspodela specifičnih ULS pokazala je da su kod sve tri grupe govornika najviše zastupljeni *ULS orijentisani ka istraživanju* praćeni *ULS orijentisanim ka tekstu*. Izvorni govornici rumunskog jezika se razlikuju u funkcionalnoj raspodeli specifičnih ULS jer ne koriste *ULS orijentisane ka učesniku* i statistički značajno više koriste *ULS orijentisane ka istraživanju*. Podaci u istraživanju Gungor i Uisal (2020) pokazuju da izvorni govornici turskog najviše koriste specifične ULS u funkcionalnoj kategoriji *ULS orijentisani ka tekstu*, praćeni *ULS orijentisanim ka istraživanju*. Kao i za strukturalnu raspodelu, razlike u funkcionalnoj raspodeli specifičnih ULS kod neizvornih govornika engleskog jezika sa različitim maternjim jezicima, potvrđuje pretpostavku da je uticaj negativnog transfera prisutniji u tim ULS.

Poređenjem procentualne zastupljenosti zajedničkih i specifičnih tipova i tokena ULS unutar celih korpusa za izvorne govornike engleskog jezika i dve grupe neizvornih govornika, uviđamo da svi zajednički ULS (i oni koji su zajednički za sva tri korpusa i oni koji se dele u po dva korpusa) imaju veću procentualnu zastupljenost tokena u celim korpusima uporedno sa tipovima, dok je za specifične ULS obrnuto. Drugim rečima, zajednički ULS se češće upotrebljavaju kod sve tri grupe govornika uporedno sa specifičnim ULS.

5.7. Uticaj negativnog transfera na produkciju ULS kod neizvornih govornika engleskog jezika

Da bi se utvrdio uticaj negativnog transfera na upotrebu ULS kod neizvornih govornika u istraživanju u ovoj disertaciji primenjen je Džarvisov pristup zasnovan na detekciji koji podrazumeva prikupljanje tri vrste dokaza u procesu identifikacije uticaja transfera (Džarvis, 2000, 2012; Džarvis i Pavlenko, 2008). Dokazi o unutargrupnoj homogenosti su obezbeđeni samom identifikacijom ULS u ECEP, SCEP i RCEP korpusima primenom minimalne frekvencije od 40 ponavljanja u milion reči i minimalne disperzije u 10% tekstova.

Dokazi o međugrupnoj heterogenosti dobijeni su izdvajanjem i pročišćavanjem spiskova specifičnih ULS za svaku od dve grupe neizvornih govornika. Poređenjem izdvojenih ULS u sva tri korpusa i izdvajanjem onih koji se koriste samo u korpusu izvornih govornika srpskog jezika i samo u korpusu izvornih govornika rumunskog jezika izdvojeni su spiskovi specifičnih ULS. Izdvojeno je 66 specifičnih ULS za izvorne govornike engleskog jezika kojima je maternji jezik srpski i 18 specifičnih za izvorne govornike engleskog jezika kojima je maternji jezik rumunski. Dobijeni spiskovi specifičnih ULS su dalje suženi poređenjem sa spiskom specifičnih ULS za izvorne govornike turskog jezika u naučnim radovima iz oblasti obrazovanja koji je dostupan u istraživanju Gungor (2016). Nakon navedenog koraka, dodatno sužen spisak ULS specifičnih za izvorne govornike srpskog jezika sastojao se od 56 ULS, a spisak ULS specifičnih za izvorne govornike rumunskog jezika od 16 ULS. Finalan korak u prečišćavanju spiskova specifičnih ULS za koje postoji velika verovatnoća da su pod uticajem negativnog transfera je sagledavanje njihove zastupljenosti u ECEP korpusu. Sa spiska su uklonjeni ULS koji se u ECEP korpusu pojavljuju sa većom frekvencijom od pet ponavljanja. Stoga se na finalnom prečišćenom spisku specifičnih ULS za izvorne govornike srpskog našlo 39 ULS, dok se na finalnom spisku specifičnih ULS za izvorne govornike rumunskog jezika našlo devet ULS.

Jedna od specifičnosti istraživanja u ovoj disertaciji uporedno sa prethodnim koji su ispitivali uticaj transfera na upotrebu ULS je upravo postupak prikupljanja dokaza za međugrupnu heterogenost. Naime, u prethodnim istraživanjima za prikupljanje dokaza za međugrupnu heterogenost, nisu formirani poredivi korpusi različitih grupa neizvornih govornika kao što je to slučaj u ovom istraživanju, već su korišćeni spiskovi iz prethodnih istraživanja, u kojima se ne poklapaju parametri za izdvajanje ULS ili žanr korpusa i/ili oblast istraživanja radova u korpusu (Güngör, 2016). U istraživanju Paket (2013) jesu upotrebljeni korpusi različitih grupa neizvornih govornika, ali korpusi nisu bili ujednačeni po veličini, dok u istraživanju Kostromitina (2022) dokazi za međugrupnu heterogenost uopšte nisu dati.

Kako bi se prikupili dokazi za međujezičku podudarnost performansi predloženi su funkcionalni prevodni ekvivalenti, čija se zastupljenost proverila u dva referentna korpusa. S obzirom na činjenicu da analiza uticaja transfera nije glavni cilj istraživanja u ovoj disertaciji, nisu formirani referentni korpusi na srpskom i rumunskom jeziku, već su korišćeni najprikladniji javno dostupni korpusi. Naveden postupak se poklapa sa onim koji je primenjen u istraživanju Paket (2013). Proces izdvajanja prevodnih ekvivalenata se razlikuje u dva istraživanja jer Paket (2013) analizira samo ULS sa leksičkim glagolom, a u istraživanju u ovoj disertaciji analiziramo sve izdvojene specifične

ULS. Na osnovu tri vrste dokaza ustanovili smo da je 29,37% ULS koje koriste izvorni govornici srpskog jezika pod uticajem negativnog transfera, dok je kod izvornih govornika rumunskog 13,33% ULS pod uticajem negativnog transfera. Na osnovu navedenih podataka potvrđena je sedma radna hipoteza: *ULS koji se koriste zbog uticaja transfera sačinjavaće mali procenat (do 40%) ULS identifikovanih u korpusima neizvornih govornika engleskog jezika kojima su maternji jezici srpski ili rumunski u naučnim radovima iz oblasti obrazovanja.*

Poređenjem podataka o uticaju transfera sa podacima iz prethodnih istraživanja koja su se bavila istom temom, i čiji su podaci sumirani u Tabeli 6. *Pregled uticaja transfera na upotrebu ULS u prethodnim relevantnim istraživanjima* uvidamo da se najveći negativna uticaj transfera (39,5%) zapaža kod neizvornih govornika engleskog jezika kojima je maternji jezik turski u naučnim radovima iz oblasti obrazovanja (Güngör, 2016), praćen neizvornim govornicima engleskog kojima je maternji jezik srpski u ovom istraživanju (29,37%), dok su na trećem mestu opet neizvorni govornici engleskog kojima je maternji jezik turski (24,80%), ovoga puta u naučnim radovima iz oblasti podučavanja jezika. Neizvorni govornici kojima je maternji jezik rumunski (13,33%) i izvorni govornici kojima je maternji jezik ruski (9,87%) pokazuju najmanji uticaj negativnog transfera. Ako bi primenili strože kriterijume za identifikaciju negativnog transfera, uticaj transfera na produkciju ULS kod izvornih govornika srpskog jezika (9,01%) i izvornih govornika rumunskog jezika (6,67%) bio bi najmanji zabeležen u dosadašnjim istraživanjima.

5.8. Upotreba najzastupljenijih obrazaca ULS

Rezultati istraživanja su pokazali da se izvorni govornici engleskog jezika i neizvorni govornici engleskog jezika kojima je maternji jezik srpski ili rumunski razlikuju u upotrebi najučestalijih obrazaca ULS u naučnim radovima iz oblasti obrazovanja. Naime, potvrđena je osma radna hipoteza: *Najzastupljeniji okviri će se drugačije koristiti kod izvornih govornika engleskog jezika i dve grupe neizvornih govornika engleskog jezika u naučnim radovima u oblasti obrazovanja.*

Prethodna istraživanja koja su ispitivala upotrebu najzastupljenijih obrazaca a koja su pogodna za poređenje sa istraživanjem u ovoj disertaciji su Pan i sar. (2016) i Lazić (2017), za sve ispitane obrasce. Obrazac *uvodna it struktura + glagolska/ pridevska fraza*, takođe se vodi i kao strukturalna podgrupa, te se rezultati o njegovoj upotrebi porede i sa rezultatima iz istraživanja Gungor (2016) i Al Fedžri i sar. (2020). Dva navedena istraživanja nisu posebno ispitivala najzastupljenije obrasce ali jesu izdvojila informacije o zastupljenosti obrasca *uvodna it struktura + glagolska/ pridevska fraza*.

5.8.1. Obrazac *the + imenica + of the/a*

ULS koji prate obrazac *the + imenica + of the/a* procentualno gledano u odnosu na ceo korpus, najzastupljeniji su kod izvornih govornika rumunskog jezika i sačinjavaju 21,67% tipova i 18,41% tokena, praćeni zastupljenošću kod izvornih govornika engleskog gde sačinjavaju 18,52% tipova i 15,19% tokena. ULS koji prate obrazac *the + imenica + of the/a* procentualno gledano u odnosu na ceo korpus, najmanje su zastupljeni kod izvornih govornika srpskog gde sačinjavaju 14% tipova i tokena. Navedeni rezultati se delimično poklapaju sa rezultatima prethodnih istraživanja. Naime, u istraživanju Pan i sar. (2016) izvorni govornici engleskog jezika su, procentualno gledano u odnosu na ceo korpus, više koristili ULS koji prate obrazac *the + imenica + of the/a* uporedno sa neizvornim govornicima engleskog jezika kojima je maternji jezik kineski. U korpusu izvornih govornika je procentualna zastupljenost tipova 18,18% a tokena 17,34%, dok je kod izvornih govornika kineskog zastupljenost tipova 11,27% a tokena 10,76%. U istraživanju Lazić (2017) se isto vidi veća zastupljenost ULS koji prate navedeni obrazac kod izvornih govornika engleskog jezika (11,8% svih tipova u korpusu izvornih govornika) uporedno sa neizvornim govornicima

engleskog jezika kojima je maternji jezik srpski (6,7% svih ULS u korpusu izvornih govornika srpskog jezika). Dakle, kada uzmemo u obzir rezultate zastupljenosti ULS koji prate obrazac *the + imenica + of the/a* kod neizvornih govornika kojima je maternji jezik srpski u ovom istraživanju, oni se poklapaju sa rezultatima prethodnih istraživanja, te neizvorni govornici procentualno manje koriste ULS koji prate navedeni obrazac uporedno sa izvornim govornicima. Međutim, kada uzmemo u obzir rezultate zastupljenosti ULS koji prate obrazac *the + imenica + of the/a* kod neizvornih govornika kojima je maternji jezik rumunski, oni se *ne* poklapaju sa prethodnim istraživanjima jer izvorni govornici rumunskog jezika, procentualno gledano, više koriste ULS koji prate navedeni obrazac od izvornih govornika engleskog jezika. Pretpostavlja se da je navedena razlika u rezultatima ovog istraživanja i prethodnih istraživanja posledica karakteristike izvornih govornika rumunskog jezika da u žanru naučnih radova koriste veći broj ULS iz strukturalne kategorije *imeničke strukture*, međutim, neophodna su dodatna istraživanja da bi se to mogla sa sigurnošću tvrditi.

Kada smo u obzir uzeli raznolikost reči sa punim leksičkim značenjem koje se koriste u ULS koji prate obrazac *the + imenica + of the/a*, utvrdili smo da u istraživanju u ovoj disertaciji, neizvorni govornici kojima je maternji jezik srpski pokazuju najveću raznolikost upotrebljenih reči sa punim leksičkim značenjem. Ukupno 16 reči sa punim leksičkim značenjem unutar ULS koji prate obrazac *the + imenica + of the/a* identifikovano je u SCEP korpusu, od kojih je osam specifično za taj korpus. Neizvorni govornici engleskog jezika kojima je maternji jezik rumunski koriste 13 reči sa punim leksičkim značenjem od kojih je četiri specifično za izvorne govornike rumunskog jezika, dok izvorni govornici engleskog koriste najmanji broj reči sa punim leksičkim značenjem (pet), od kojih su dve specifične za njihovu produkciju. Dve grupe govornika dele dve reči sa punim leksičkim značenjem unutar ULS koji prate obrazac *the + imenica + of the/a – end* i *value*. Ovakvi rezultati su u skladu sa nalazima istraživanja koje je sprovedla Lazić (2017), u kojem neizvorni govornici engleskog jezika kojima je maternji jezik srpski koriste 14 reči sa punim leksičkim značenjem unutar ULS koji prate obrazac *the + imenica + of the/a*, dok izvorni govornici engleskog koriste osam. Rezultati o raznovrsnosti upotrebe reči sa punim leksičkim značenjem unutar ULS koji prate obrazac *the + imenica + of the/a* u istraživanju ove disertacije ne poklapaju se sa rezultatima istraživanja Pan i sar. (2016), u kojem je veći broj reči sa punim leksičkim značenjem unutar ULS koji prate navedeni obrazac identifikovano kod izvornih govornika engleskog jezika. Tačnije, deset reči sa punim leksičkim značenjem je identifikovano u ULS koji prate obrazac *the + imenica + of the/a* u korpusu izvornih govornika engleskog jezika, dok je osam identifikovano u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik kineski. Pet reči sa punim leksičkim značenjem se dele u oba korpusa unutar ULS koji prate obrazac *the + imenica + of the/a*. Ovde nailazimo na još jednu neujednačenost u rezultatima koji su dobijeni u različitim istraživanjima upotrebe ULS kod izvornih i neizvornih govornika. Kako bi se utvrdilo da li je do neujednačenosti u karakteristikama upotrebe reči sa punim leksičkim značenjem u izdvojenom obrascu ULS kod neizvornih govornika došlo zbog uticaja različitih maternjih jezika (u ovom slučaju kineskog), zbog razlika u oblasti kojom se bave naučni radovi koji sačinjavaju korpuse (u istraživanju Pan i sar. 2016 to je oblast telekomunikacije koja je dosta različita od oblasti obrazovanja i biotehničkih nauka) ili pak nekog trećeg razloga – neophodno je sprovesti dalja istraživanja.

5.8.2. Obrazac *in the + imenica + of*

Rezultati istraživanja u ovoj disertaciji pokazali su da se ULS koji prate obrazac *in the + imenica + of* najučestalije primenjuju kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski, sa devet tipova i 552 tokena koji sačinjavaju 8,11% svih tipova i 12,17% svih tokena u SCEP korpusu. U korpusu izvornih govornika engleskog jezika procentualno su zastupljeniji tipovi ULS koji prate obrazac *in the + imenica + of* uporedno sa izvornim govornicima rumunskog jezika,

dok je kod izvornih govornika rumunskog jezika procentualno veća zastupljenost tokena ULS koji prate navedeni obrazac. Navedeni rezultati se poklapaju sa onim koji su dobijeni u istraživanju Lazić (2017), u kojem je veća zastupljenost ULS koji prate obrazac *in the + imenica + of* identifikovana kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski (9,2% svih tipova u korpusu) nego kod izvornih govornika engleskog jezika (8% svih tipova u korpusu). Kada uporedimo rezultate sa onima koji su dobijeni u istraživanju Pan i sar. (2016), dobijamo malo kompleksniju sliku. Naime, u tom istraživanju izvorni govornici engleskog jezika koriste više tipova (5,45%) i tokena (6,94%) ULS koji prate obrazac *in the + imenica + of*, nego neizvorni govornici engleskog jezika kojima je maternji jezik kineski (2,82% tipova i 2,57% tokena). Ovakvi rezultati se ne poklapaju sa podacima koje smo dobili o upotrebi ULS koji prate navedeni obrazac kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski, a delimično se poklapaju sa rezultatima kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski. Naime, tipovi ULS koji prate obrazac *in the + imenica + of* u istraživanju u ovoj disertaciji koriste više kod izvornih govornika engleskog jezika uporedno sa rumunskim, dok se tokeni koriste u procentualno istoj razmeri, odnosno sačinjavaju 13% svih tokena u oba korpusa.

Kada se govori o rečima sa punim leksičkim značenjem koje su se koristile u ULS koji prate obrazac *in the + imenica + of*, rezultati istraživanja u ovoj disertaciji pokazali su da se najveća raznolikost reči sa punim leksičkim značenjem identifikovala kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski sa devet reči, praćeno sa četiri reči koje su identifikovane kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski. Najmanji se broj reči sa punim leksičkim značenjem identifikovao kod izvornih govornika engleskog jezika sa dve reči. Nijedna reč sa punim leksičkim značenjem koja se pojavljuje u ULS koji prate obrazac *in the + imenica + of* se ne deli u sva tri korpusa, dve reči se dele u ECEP i SCEP korpusima, dok se tri reči dele između SCEP i RCEP korpusa. Navedeni rezultati se poklapaju sa rezultatima istraživanja Lazić (2017), u kojem je identifikovana veća raznolikost upotrebe reči sa punim leksičkim značenjem unutar ULS koji prate obrazac *in the + imenica + of* kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski (11 reči), nego kod izvornih govornika engleskog jezika (pet reči). Rezultati dobijeni u istraživanju ove disertacije i u istraživanju Lazić (2017) se ne poklapaju sa rezultatima istraživanja Pan i sar. (2016), u kojima je veća raznolikost reči sa punim leksičkim značenjem unutar ULS koji prate obrazac *in the + imenica + of* kod izvornih govornika engleskog jezika (tri reči), nego kod neizvornih govornika engleskog jezika kojim je maternji jezik kineski (dve reči). Možemo pretpostaviti da su navedene razlike u dobijenim podacima posledica uticaja maternjeg jezika neizvornih govornika na njihovu upotrebu ULS ili pak zbog različitih normi u diskursu različitih naučnih oblasti. Međutim, dodatna istraživanja su neophodna kako bi se ustanovio uzrok različitih tendencija u upotrebi ULS koji prate obrazac *in the + imenica + of* kod izvornih govornika engleskog jezika i različitih grupa neizvornih govornika engleskog jezika.

5.8.3. Obrazac *glagol u pasivu + fragment predloške fraze*

Obrazac *glagol u pasivu + fragment predloške fraze* se u istraživanju ove disertacije našao samo kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski sa tri ULS, koji sačinjavaju 2,7% tipova i 1,72% tokena u celom SCEP korpusu. Navedeni rezultati se delimično poklapaju sa rezultatima istraživanja Lazić (2017), u kojem se navedeni obrazac javlja i kod izvornih i kod neizvornih govornika engleskog jezika. Veći se broj tipova identifikovao u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik srpski (17 tipova koji čine 6,1% svih tipova u korpusu), dok je manji broj tipova identifikovan u korpusu izvornih govornika engleskog jezika (devet tipova koji sačinjavaju 20,9% svih tipova u korpusu), ali procentualno gledano, oni sačinjavaju veći procenat svih ULS korišćenih u tom korpusu. Drugim rečima, ULS koji prate obrazac *glagol u pasivu + fragment predloške fraze*, procentualno gledano, učestalije se koriste kod izvornih govornika engleskog jezika u istraživanju Lazić (2017). Rezultati istraživanja

u ovoj disertaciji se delimično poklapaju sa rezultatima iz istraživanja Pan i sar. (2016), u kojem su ULS koji prate obrazac *glagol u pasivu + fragment predloške fraze* identifikovani i kod izvornih govornika engleskog jezika i kod neizvornih govornika engleskog jezika kojima je maternji jezik kineski. No, veći broj navedenih ULS identifikovan je kod izvornih govornika kineskog (17 tipova koji sačinjavaju 23,94% celog korpusa), nego kod izvornih govornika engleskog jezika (pet tipova koji čine 9,09% čitavog korpusa). U istraživanju Pan i sar. (2017) izdvojili su se ULS koji prate navedeni obrazac a koji počinju sa modalnim glagolom *can* kao posebno učestali u produkciji neizvornih govornika engleskog jezika. Poređenjem tog zaključka sa istraživanjem Lazić (2017) uviđamo da se kod neizvornih govornika engleskog kojima je maternji jezik srpski našlo četiri ULS u navedenom obrascu koji počinju sa modalnim glagolom *can*, dok je u istraživanju u ovoj disertaciji kod izvornih govornika engleskog jezika kojima je maternji jezik srpski identifikovan jedan takav obrazac. U produkciji izvornih govornika srpskog jezika u istraživanju Lazić (2017) posebno su učestali ULS koji počinju sa pomoćnim glagolom *be* u prošlom obliku *was* unutar ULS koji prate obrazac *glagol u pasivu + fragment predloške fraze*. Jedan ULS unutar navedene grupe u istraživanju ove disertacije kod izvornih govornika srpskog jezika isto počinje sa pomoćnim glagolom *be* u prošlom obliku. Dakle, tendencija neizvornih govornika engleskog jezika da koriste veći broj ULS koji prate obrazac *glagol u pasivu + fragment predloške fraze* identifikovana je u istraživanju ove disertacije za izvorne govornike srpskog i u istraživanju Pan i sar. (2016) za izvorne govornike kineskog. Veći broj ULS koji prate obrazac *glagol u pasivu + fragment predloške fraze* identifikovan je u istraživanju Lazić (2017) kod izvornih govornika srpskog jezika, međutim, procentualno gledano, navedeni ULS su zastupljeniji u korpusu izvornih govornika engleskog jezika. Nijedan ULS koji prati obrazac *glagol u pasivu + fragment predloške fraze* nije identifikovan kod izvornih govornika rumunskog jezika u ovom istraživanju. Na osnovu podataka u ovom segmentu, možemo zaključiti da se upotreba ULS koji prate obrazac *glagol u pasivu + fragment predloške fraze* razlikuje kod različitih grupa neizvornih govornika i u različitim oblastima istraživanja.

5.8.4. Obrazac *uvodna it struktura + glagolska/ pridevska fraza*

U istraživanju u ovoj disertaciji pokazalo se da izvorni govornici engleskog jezika, procentualno gledano, učestalije koriste ULS koji prate obrazac *uvodna it struktura + glagolska/ pridevska fraza* (7,46% tipova i 8,62 % tokena u celom korpusu), uporedno sa neizvornim govornicima engleskog jezika kojima je maternji jezik srpski (4,50% tipova i 6,90% tokena) i rumunski (3,33% tipova i 2,42% tokena). Navedeni rezultati nisu u skladu sa rezultatima prethodnih istraživanja Pan i sar (2016), Gungor (2016), Lazić (2017) i Al Fedžri i sar. (2020). Naime, u sva četiri navedena istraživanja neizvorni govornici engleskog jezika, procentualno gledano, učestalije koriste ULS koji prate obrazac *uvodna it struktura + glagolska/pridevska fraza* od izvornih govornika engleskog jezika (v. Tabelu 145). Na osnovu dostupnih podataka nije jasno iz kog razloga je došlo do neusklađenosti rezultata u ovom istraživanju uporedno sa prethodnim, koji se poklapaju po činjenici da su ULS koji prate obrazac *uvodna it struktura + glagolska/pridevska fraza*, procentualno gledano, zastupljeniji u produkciji neizvornih govornika engleskog jezika.

Tabela 145. *Upotreba ULS koji prate obrazac uvodna it struktura + glagolska/ pridevska fraza u relevantnim prethodnim istraživanjima*

Istraživanje	Izvorni govornici engleskog jezika		Neizvorni govornici engleskog jezika	
	Tip %	Token %	Tip %	Token %
<i>Pan i sar. (2016)</i>	5.45	4.66	8.45	6.76
<i>Gungor (2016)</i>	5	5	13	17
<i>Lazić (2017)</i>	2.8	3.3	8.5	11
<i>Al Fedžri i sar. (2020)</i>	3.23	3.52	13.73	13.21

5.8.5. Sumiranje podataka o upotrebi najzastupljenijih obrazaca ULS

Poređenjem dobijenih rezultata za sva četiri najzastupljenija obrasca ULS u sagledanim istraživanjima uviđamo da se neizvorni govornici engleskog jezika kojima je maternji jezik rumunski razlikuju po upotrebi najučestalijih obrazaca uporedno sa drugim grupama neizvornih govornika. Naime, rezultati o upotrebi obrazaca *the + imenica + of the/a* i *in the + imenica + of* za izvorne govornike srpskog poklapaju se sa rezultatima dobijenih u istraživanjima Pan i sar. (2016) i Lazić (2017), u kojima je ispitana produkcija izvornih govornika kineskog i srpskog jezika. Međutim, rezultati o upotrebi navedena dva obrasca kod izvornih govornika rumunskog jezika se ne poklapaju ni sa rezultatima upotrebe izvornih govornika srpskog u ovom istraživanju, kao ni sa dobijenim rezultatima u istraživanjima Pan i sar. (2016) i Lazić (2017). Specifičnosti upotrebe najzastupljenijih obrazaca kod izvornih govornika rumunskog jezika su takođe evidentne kod obrasca *glagol u pasivu + fragment predloške fraze*, koji uopšte nije bio zastupljen u izdvojenim ULS u RCEP korpusu, dok su se rezultati dobijeni za izvorne govornike srpskog jezika poklopili sa rezultatima dobijenim u istraživanju Pan i sar. (2016).

Štaviše, poređenjem rezultata o najzastupljenijim obrascima u ovom i prethodnim istraživanjima, jasno uviđamo potrebu za sprovođenjem dodatnih istraživanja kod grupa neizvornih govornika sa različitim maternjim jezicima, kao i dodatnih istraživanja kod već ispitanih grupa neizvornih govornika ali u drugim naučnim oblastima. To bi omogućilo da se ustanovi koje su karakteristike upotrebe najzastupljenijih obrazaca *zajedničke za različite grupe neizvornih govornika*, koje su *karakteristične za određene grupe neizvornih govornika*, a koje su *karakteristike rezultat normi u naučnim radovima iz određene naučne oblasti*. Potreba za ovim istraživanjem je jasna, jer se uviđaju neusklađenosti u podacima koje ne mogu da se objasne sa sigurnošću bez dodatnih istraživanja.

5.9. Metodičke smernice

Kao što je istaknuto u ranijim segmentima, jedan od sekundarnih ciljeva ove disertacije je izdvojiti smernice za podučavanje ULS u okviru nastave akademskog engleskog jezika i podučavanja/usavršavanja prevodilaca za rad u datom žanru i oblasti. Uzimajući u obzir činjenicu da se u današnjem globalizovanom svetu vrše sve veći pritisci na članove akademske zajednice da svoje radove objavljuju u globalnim uticajnim časopisima koji se i dalje mahom objavljuju na engleskom jeziku (Filipović i Vučo, 2013; Murešan i Perez-Ljantada, 2014), dodatna podrška studentima, članovima akademske zajednice koji se upuštaju u pisanje radova na engleskom jeziku, kao i prevodiocima koji se bave prevodima naučnih radova, postaje sve značajnija. Štaviše, kada uzmemo u obzir i činjenicu da pravilno korišćenje ULS doprinosi pokazivanju pripadnosti specifičnoj diskursnoj zajednici odnosno delatnoj zajednici (Adel i Erman, 2012) i formiranju uverljivog akademskog glasa odnosno profesionalno prihvatljive persone (Hajland i sar., 2004) i činjenicu da i dalje postoji jaz između dostignuća u korpusnoj lingvistici i podučavanju engleskog jezika, težnja ka praktično primenjivim ciljevima postaje još jasnija. U daljem tekstu predlažemo smernice za podučavanje ULS sa ciljem unapređenja pisanja i/ili prevođenja naučnih radova u oblasti obrazovanja.

Prethodna istraživanja su pokazala da neizvorni govornici engleskog jezika pri pisanju naučnih radova na engleskom jeziku imaju tendenciju da koriste veći broj ULS od izvornih govornika (Pan i sar., 2016; Güngör, 2016; Lazić, 2017; Al Fedžri i sar., 2020). Ista se tendencija pokazala i u rezultatima našeg istraživanja, iako je veći broj ULS izdvojen u produkciji neizvornih govornika kojim je maternji jezik srpski (111 tipova i 4.536 tokena) nego kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski (60 tipova i 2.358 tokena). Iako su, uporedno sa izvornim govornicima engleskog jezika, izvorni govornici srpskog jezika pokazali uvećanje u upotrebi ULS od 311% a izvorni govornici rumunskog jezika pokazali uvećanje od 122%, i dalje je postojalo

devet ULS koji se nisu upotreбили ni u SCEP ni u RCEP korpusu. Zbog činjenice da se u korpusu izvornih govornika engleskog jezika izdvojio mnogo manji broj ULS, odnosno samo 29 tipova ULS, prva preporuka profesorima engleskog jezika jeste da sa studentima obrade ULS koje koriste izvorni govornici u ciljanom kontekstu, u ovom slučaju to su naučni radovi iz oblasti obrazovanja. Važno je studente osvestiti o strukturama i funkcijama ULS. Kao što smo saznali u ranijim segmentima ove disertacije, izvorni i neizvorni govornici imaju različite tendencije u upotrebi ULS kada je reč o njihovim strukturama i funkcijama. U daljem tekstu slede smernice za rad sa neizvornim govornicima kojima je maternji jezik srpski, smernice za rad sa neizvornim govornicima kojima je maternji jezi rumunski, kao i smernice koje mogu da se koriste u radu i sa izvornim govornicima srpskog jezika i sa izvornim govornicima rumunskog jezika.

5.9.1. Metodiske smernice koje se predlažu za rad sa izvornim govornicima srpskog jezika

Na osnovu poređenja strukturalne raspodele svih ULS u ECEP i SCEP korpusa došli smo do sledećih metodiskih smernica:

- Pregledom strukturalnih grupa uvideli smo da izvorni govornici srpskog jezika imaju tendenciju da isuviše učestalo koriste ULS iz grupe *druge strukture*, a srazmerno nedovoljno učestalo ULS iz grupe *fragmenti sa predloškom frazom*. Stoga bi u radu sa izvornim govornicima srpskog jezika trebalo obraditi alternativne načine za prenošenje značenja za koje se koriste *druge strukture*, kao i da se potencira korišćenje adekvatnih ULS iz strukturalne grupe *fragmenti sa predloškom frazom*.
- Posmatranjem strukturalnih podgrupa utvrdili smo da izvorni govornici srpskog jezika imaju tendenciju da nedovoljno često koriste ULS iz podgrupe (*glagol/pridev +*) *fragment to klauze i imeničke fraze sa of fragmentom*, dok imaju tendenciju da isuviše učestalo koriste ULS iz podgrupa *glagol u pasivu i fragment predloške fraze, kopula be i imenička/pridevska fraza, fragment that klauze, zamenica/imenička fraza+be (+...), druge glagolske fraze i druge imeničke fraze*. Shodno tome, u radu sa izvornim govornicima srpskog jezika trebalo bi staviti poseban akcenat na obrađivanje adekvatnih ULS iz strukturalnih podgrupa (*glagol/pridev +*) *fragment to klauze i imeničke fraze sa of fragmentom*, dok bi trebalo obraditi alternativna i raznovrsna rešenja za prenošenje značenja koje prenose sledeće strukturalne podgrupe: *glagol u pasivu i fragment predloške fraze, kopula be i imenička/pridevska fraza, fragment that klauze, zamenica/imenička fraza+be (+...), druge glagolske fraze i druge imeničke fraze*.

Na osnovu strukturalne analize specifičnih ULS u SCEP korpusu i poređenjem strukturalne raspodele specifičnih ULS u ECEP i SCEP korpusima, došli smo do sledećih metodiskih smernica:

- Neizvorni govornici engleskog jezika kojima je maternji jezik srpski imaju veću tendenciju da koriste ULS koji su nekarakteristični za upotrebu kod izvornih govornika, u strukturalnim kategorijama *glagolske strukture* i *druge strukture*. U tri podgrupe *glagolskih struktura* u SCEP korpusu, *glagol u pasivu i fragment predloške fraze, zamenica/imenička fraza+be (+...)* i *druge glagolske fraze*, svi primeri su specifični za produkciju izvornih govornika srpskoga jezika. Stoga se nameće zaključak da bi pri podučavanju akademskog engleskog u oblasti obrazovanja trebalo detaljno obraditi ULS koje izvorni govornici koriste u strukturalnim podgrupama: *glagol u pasivu i fragment predloške fraze, zamenica/imenička fraza+be (+...)* i *druge glagolske fraze*, kako studenti ne bi koristili ULS koji nisu najoptimalnija opcija za pokazivanje pripadnosti u datoj delatnoj zajednici. Takođe je neophodno da se sagledaju specifični ULS iz kategorije *druge strukture* i da se predlože alternativni načini prenošenja njihovih značenja koja su više u skladu sa normama u željenoj delatnoj zajednici.

- Na osnovu podatka da neizvorni govornici engleskog kojima je maternji jezik srpski imaju tendenciju da prekomerno koriste ULS *drugih struktura* a zanemare produkciju relevantnih ULS sa *glagolskom strukturom*, dobijamo još jedan razlog da istaknemo značaj smernice do koje smo došli analizom specifičnih ULS u SCEP korpusu.
- S obzirom na činjenicu da se podgrupa (*glagol/pridev +) fragment to klauze* značajno češće koristi u ECEP korpusu, dolazimo do zaključka da izvorni govornici srpskog imaju tendenciju da zanemaruju upotrebu prikladnih ULS iz navedene kategorije, te predlažemo da se posebna pažnja posveti navedenoj konstrukciji u nastavi.

Na osnovu poređenja funkcionalne raspodele svih ULS u ECEP i SCEP korpusa, došli smo do sledećih metodičkih smernica:

- Neizvorni govornici engleskog jezika kojima je maternji jezik srpski imaju tendenciju da isuviše učestalo koriste ULS iz funkcionalne grupe *ULS orijentisani ka tekstu*, dok nedovoljno učestalo koriste ULS iz kategorija *ULS orijentisani ka istraživanju* i *ULS orijentisani ka učesniku*, srazmerno raspodelama u ECEP korpusu. Otuda predlažemo da se u radu sa izvornim govornicima srpskog jezika poseban akcenat stavi na obradu adekvatnih ULS iz funkcionalnih grupa *ULS orijentisani ka istraživanju* i *ULS orijentisani ka učesniku*. Takođe je potrebno pružiti studentima alternativna rešenja za iskazivanje značenja za koja se koriste *ULS orijentisani ka tekstu*.
- Uzimajući u obzir funkcionalne podgrupe uviđamo da izvorni govornici srpskog jezika srazmerno nedovoljno učestalo upotrebljavaju ULS iz funkcionalnih podgrupa *ULS za lokaciju*, *ULS stava* i *strukturnih signala*, dok značajno učestalije koriste *ULS za kvantifikaciju*, *ULS za uključivanje*, *tranzicionih* i *ograničavajućih signala*. Na osnovu navedenih podataka, predlažemo da se u radu sa izvornim govornicima srpskog jezika posebna pažnja posveti podučavanju adekvatnih ULS u funkcionalnim kategorijama *ULS za lokaciju*, *ULS stava* i *strukturnih signala*. Takođe se predlaže podučavanje alternativnih rešenja za prenošenje značenja za koja se koriste ULS iz funkcionalnih kategorija *ULS za kvantifikaciju*, *ULS za uključivanje*, *tranzicioni* i *ograničavajući signali*.

Na osnovu funkcionalne raspodele specifičnih ULS za SCEP korpus, došli smo do sledećih metodičkih smernica:

- Neizvorni govornici engleskog jezika kojima je maternji jezik srpski u naučnim radovima u oblasti obrazovanja imaju tendenciju da koriste više specifičnih ULS u funkcionalnim kategorijama *ULS orijentisani ka učesniku* i *ULS orijentisani ka istraživanju*. Stoga, pri podučavanju studenata ili prevodilaca, koji žele da se bave prevodom naučnih radova u oblasti obrazovanja, a kojima je maternji jezik srpski, trebalo bi posebnu pažnju posvetiti podučavanju prihvatljivih ULS iz dve navedene funkcionalne kategorije kako ne bi pribegavali upotrebi ULS koji su manje pogodni za pokazivanje pripadnosti u navedenoj delatnoj zajednici. Kada uporedimo zaključke dobijene na osnovu svih ULS i specifičnih u SCEP korpusu, uviđamo da izvorni govornici srpskog nedovoljno koriste ULS iz funkcionalnih kategorija *ULS orijentisani ka učesniku* i *ULS orijentisani ka istraživanju* a da su i ti koje koriste prevashodno specifični, odnosno neoptimalan su izbor za pokazivanje pripadnosti u željenoj delatnoj zajednici.
- U radu sa studentima i prevodiocima koji žele da prevode naučne radove iz oblasti obrazovanja koji su izvorni govornici srpskog jezika, trebalo bi posvetiti posebnu pažnju obradi ULS iz funkcionalnih podgrupa *ULS za opis*, *ULS za kvantifikaciju* i *ograničavajući signali* kako ne bi došlo do prekomerne upotrebe ULS koji nisu najpogodnije rešenje za pokazivanje pripadnosti u navedenoj delatnoj zajednici. Takođe bi trebalo analizirati koja

se značenja prenose specifičnim ULS iz navedenih grupa kako bi u nastavi mogla da se obrade adekvatnija rešenja.

Na osnovu poređenja funkcionalne raspodele specifičnih ULS u ECEP i SCEP korpusa, došli smo do sledećih metodičkih smernica:

- Izvorni govornici srpskog imaju tendenciju da značajno više koriste specifične *ULS orijentisane ka tekstu* koji nisu najbolji načini za prenošenje značenja u istraženoj delatnoj zajednici. Stoga je potrebno proučiti koja značenja prenose ULS iz kategorije *ULS orijentisani ka tekstu* i podučiti adekvatnija rešenja za prenošenje značenja. Takođe, rezultati pokazuju tendenciju izvornih govornika srpskog da posebno zanemare relevantne ULS iz funkcionalne kategorije *ULS orijentisani ka učesniku*, te se posebna pažnja treba posvetiti podučavanju ovih konstrukcija.
- Zbog postojanja značajne tendencije izvornih govornika srpskog da zanemaruju relevantne ULS iz funkcionalnih podgrupa *ULS stava* i *strukturalni signali*, predlaže se posvećivanje posebne pažnje pri podučavanju navedenih funkcionalnih podgrupa. Navedeni podaci su u skladu sa onim koji su dobijeni na osnovu poređenja raspodele svih ULS u SCEP i ECEP korpusima.
- Rezultati poređenja funkcionalnih raspodela specifičnih ULS u SCEP i ECEP korpusu potvrđuju zaključke dobijene na osnovu poređenja svih ULS da se nedovoljno koriste ULS iz funkcionalnih podgrupa *ULS za kvantifikaciju*, *tranzicioni signali*, *ograničavajući signali* i *ULS uključivanja*. Međutim, na osnovu funkcionalnih raspodela specifičnih ULS saznajemo i da su nedovoljno zastupljene i funkcionalne podgrupe *ULS za proceduru* i *rezultativni signali*, te je neophodno dodatno akcentovati i njihov značaj pri podučavanju.

5.9.2. Metodičke smernice koje se predlažu za rad sa izvornim govornicima rumunskog jezika

Na osnovu poređenja strukturalne raspodele svih ULS u ECEP i SCEP korpusa, došli smo do sledećih metodičkih smernica:

- Izvorni govornici rumunskog jezika u svojoj produkciji na engleskom jeziku u naučnim radovima iz oblasti obrazovanja proporcionalno isuviše učestalo koriste ULS iz strukturalnih kategorija *imeničke strukture* i *fragmenti sa predloškom frazom*, dok nedovoljno učestalo koriste ULS iz strukturalne grupe *glagolske strukture*. Stoga bi u radu sa izvornim govornicima rumunskog jezika trebalo obraditi alternativne načine za prenošenje značenja za koje se koriste *imeničke strukture* i *fragmenti sa predloškom frazom*. Takođe je potrebno posvetiti posebnu pažnju pri podučavanju adekvatnih ULS u strukturalnoj kategoriji *glagolske strukture*.
- Zbog činjenice da izvorni govornici rumunskog jezika imaju tendenciju da srazmerno isuviše učestalo koriste ULS u podgrupama *imenička fraza sa of fragmentom*, *druge predloške fraze*, *kopula be* i *imenička/pridevska fraza*, *fragment that klauze* i *druge glagolske fraze* u nastavi bi trebalo obraditi alternativne načine za prenošenje značenja za koje se koriste navedene strukturalne podgrupe. Izvorni govornici rumunskog jezika, sa druge strane, nedovoljno učestalo koriste ULS u podgrupama *uvodna it struktura + glagolska ili pridevska fraza* i *(glagol/pridev +) fragment to klauze*, što dalje implicira da bi u podučavanju ove dve strukturalne podgrupe trebalo posebno da se naglase i obrade.

Na osnovu strukturalne analize specifičnih ULS u RCEP korpusu i poređenjem strukturalne raspodele specifičnih ULS u ECEP i RCEP korpusima, došli smo do sledećih metodičkih smernica:

- Neizvorni govornici engleskog kojima je maternji jezik rumunski nemaju tendenciju da koriste specifične ULS iz strukturalnih grupa i podgrupa *glagolskih struktura*, te pri podučavanju nije potrebno akcentovati ove konstrukcije.
- Izvorni govornici rumunskog jezika statistički značajno više zanemaruju upotrebu ULS iz strukturalnih kategorija *fragmenti sa predložkom frazom* i *glagolske strukture*. Stoga je neophodno posvetiti posebnu pažnju ovim kategorijama pri podučavanju izvornih govornika rumunskog jezika.
- Izvorni govornici rumunskog jezika imaju tendenciju da značajno više koriste njima specifične ULS u strukturalnim grupama *imeničke strukture* i *druge strukture*. To dalje implicira da je potrebno ispitati koja se značenja prenose ovim konstrukcijama i u nastavi podučiti studente o prikladnijim i raznovrsnijim načinima prenošenja tih značenja.
- U radu sa studentima/prevodiocima/istraživačima kojima je maternji jezik rumunski neophodno je posvetiti posebnu pažnju obradi ULS iz strukturalnih podgrupa *druge predložke fraze*, *uvodna it struktura + glagolska ili pridevska fraza* i (*glagol/pridev + fragment to klauze*), zbog povišene tendencije da zanemare najoptimalnije ULS u navedenim kategorijama;
- Izvorni govornici rumunskog jezika imaju tendenciju da koriste specifične ULS u obe podgrupe *imeničkih struktura*, kao i podgrupama *druge glagolske fraze* i *druge strukture*. Stoga je potrebno ispitati koja se značenja prenose navedenim ULS i u radu sa studentima obraditi adekvatnija rešenja.

Na osnovu poređenja funkcionalne raspodele svih ULS u ECEP i SCEP korpusa, došli smo do sledećih metodičkih smernica:

- Izvorni govornici rumunskog jezika imaju tendenciju da nedovoljno učestalo koriste *ULS orijentisane ka učesniku*, dok isuviše učestalo koriste *ULS orijentisane ka tekstu*. Stoga je potrebno pri podučavanju ove grupe govornika posvetiti posebnu pažnju pri obradi *ULS orijentisane ka učesniku*, kao i obraditi alternativne načine za prenošenje značenja za koje se koriste *ULS orijentisane ka tekstu*.
- U radu sa izvornim govornicima rumunskog jezika posebnu pažnju trebalo bi posvetiti funkcionalnim podgrupama *ULS za lokaciju*, *strukturni signali* i *ULS stava*. Pored toga, potrebno je dodatno obraditi alternativne načine za prenošenje značenja za koja se koriste ULS iz funkcionalnih podgrupa *ULS za proceduru* i *ograničavajući signali*.

Na osnovu funkcionalne raspodele specifičnih ULS za SCEP korpus, došli smo do sledećih metodičkih smernica:

- U radu sa studentima i/ili prevodiocima kojima je maternji jezik rumunski trebalo bi staviti veći fokus na podučavanje ULS u funkcionalnoj kategoriji *ULS orijentisani ka istraživanju* kako ne bi pribegavali upotrebi ULS koji su manje primereni za pokazivanje pripadnosti u navedenoj delatnoj zajednici, kao i na podučavanje prikladnijih alternativa za specifične ULS u datoj kategoriji.
- Takođe bi trebalo posvetiti više pažnje podučavanju ULS iz funkcionalne kategorije *ULS orijentisani ka učesniku*, koji se retko koriste u celom RCEP korpusu, a na osnovu ULS koji su specifični u ECEP korpusu, vidimo da ih neizvorni govornici nedovoljno koriste.
- U radu sa studentima i prevodiocima koji žele da prevode naučne radove u oblasti obrazovanja dodatan fokus bi trebalo da se stavi na podučavanje *ULS za proceduru*, *kvantifikaciju* i *opis* kako ne bi došlo do upotrebe ULS koji nisu najoptimalnija opcija za pokazivanje pripadnosti u navedenoj delatnoj zajednici, kao i podučavanju alternativnih i prikladnijih načina za prenošenje značenja za koje se prenose specifični ULS u navedenim funkcionalnim podgrupama.

Na osnovu poređenja funkcionalne raspodele specifičnih ULS u ECEP i SCEP korpusa, došli smo do sledećih metodičkih smernica:

- Izvorni govornici rumunskog jezika pokazali su tendenciju da isuviše često koriste specifične ULS, koji su neadekvatni za pokazivanje pripadnosti u relevantnoj delatnoj zajednici, u funkcionalnoj kategoriji *ULS orijentisani ka istraživanju*. Stoga je u radu sa ovom grupom govornika potrebno podučavati adekvatnije alternative za specifične ULS iz funkcionalne grupe *ULS orijentisani ka istraživanju*. Rezultati statističke analize su pokazali i da izvorni govornici rumunskog imaju tendenciju da isuviše retko koriste adekvatne ULS iz funkcionalnih kategorija *ULS orijentisani ka tekstu* i *ULS orijentisani ka učesniku*. Shodno tome, potrebno je posvetiti posebnu pažnju podučavanju adekvatnih ULS iz dve navedene funkcionalne kategorije.
- Podaci dobijeni komparacijom funkcionalne raspodele specifičnih ULS u ECEP i RCEP korpusima podržavaju smernice do kojih smo došli funkcionalnom analizom specifičnih ULS u RCEP korpusu. Naime, potvrdila se tendencija izvornih govornika rumunskog jezika da isuviše često koriste neadekvatne ULS u funkcionalnim podgrupama *ULS za proceduru* i *kvantifikaciju*. Ovo dalje znači da je u radu sa ovom grupom neizvornih govornika potrebno pružiti adekvatne alternative za navedene specifične ULS. Izvorni govornici rumunskog jezika, isto tako, imaju tendenciju da nedovoljno koriste adekvatne ULS iz funkcionalnih podgrupa *ULS za lokaciju*, *strukturalni signali* i *ULS stava*. Stoga je poželjno staviti akcenat na podučavanje adekvatnih ULS iz ovih funkcionalnih podgrupa.

5.9.3. Metodičke smernice koje se predlažu za rad sa obe grupe neizvornih govornika

- Neizvorni govornici engleskog kojima je maternji jezik srpski ili rumunski imaju veću tendenciju da zanemare ULS karakteristične za izvorne govornike u strukturalnim kategorijama *glagolske strukture* i *imeničke strukture*;
- Neizvorni govornici kojima je maternji jezik srpski ili rumunski pri pisanju ili prevođenju naučnih radova iz oblasti obrazovanja, imaju veću tendenciju da zanemare upotrebu relevantnih ULS u podgrupama: *imenička fraza sa of fragmentom*, *druge imeničke fraze*, *predložke fraze sa of fragmentom*, *druge predložke fraze*, *uvodna it struktura + glagolska ili pridevska fraza i (glagol/pridev)+ fragment to klauze* nego u ostalih šest podgrupa;
- U radu sa obe grupe neizvornih govornika engleskog jezika potrebno je posvetiti posebnu pažnju podučavanju adekvatnih *ULS orijentisanih ka učesniku* jer imaju tendenciju da ih zanemare.

6. Zaključna razmatranja

U istraživanju u ovoj disertaciji nastojali smo da ispitamo razlike u upotrebi ULS kod izvornih govornika engleskog jezika u dve grupe neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski u naučnim radovima iz oblasti obrazovanja. Motivacija za odabir teme istraživanja proistekla je iz činjenice da je teško precizno porediti rezultate prethodnih istraživanja upotrebe ULS u naučnim radovima kod izvornih i neizvornih govornika kako bi se utvrdile karakteristike različitih grupa neizvornih govornika zbog različitih veličina korpusa, različitih kriterijuma za formiranje korpusa, različitih kriterijuma za izdvajanje ULS, kao i različitih taksonomija koje su primenjene za strukturalnu i funkcionalnu analizu u istraživanjima (Appel & Murray, 2020). Za istraživanje u ovoj disertaciji formirana su tri korpusa – ECEP, SCEP i RCEP, koji su ujednačeni po veličini, žanru, oblasti i načinu formiranja korpusa. Primenjeni su isti kriterijumi za izdvajanje ULS kako bi mogli precizno da poredimo zastupljenost ULS, strukturalnu i funkcionalnu raspodelu, kao i upotrebu najzastupljenijih obrazaca kod izvornih govornika engleskog jezika, neizvornih govornika engleskog jezika kojima je maternji jezik srpski i neizvornih govornika kojima je maternji jezik rumunski. Na osnovu navedenih poređenja, mogli smo preciznije da ustanovimo uticaj transfera i sagledamo zajedničke karakteristike neizvornih govornika, kao i da preciznije utvrdimo uticaj maternjeg jezika na razlike u upotrebi ULS. Istraživanje je sprovedeno sa ciljem boljeg razumevanja karakteristika upotrebe ULS kod različitih grupa neizvornih govornika kako bi se unapredila obrazovna praksa i tim grupama neizvornih govornika omogućilo uspešnije formiranje akademskog glasa i uklapanje u globalnu akademsku zajednicu.

Kako bismo postigli glavni cilj istraživanja – *ustanoviti razlike u karakteristikama upotrebe ULS u naučnim radovima iz oblasti obrazovanja kod izvornih i neizvornih govornika engleskog jezika u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca* i kako bismo mogli da proverimo glavnu hipotezu u istraživanju – *karakteristike upotrebe ULS u naučnim radovima iz oblasti obrazovanja će se razlikovati kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski, u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca* – sproveli smo sedam koraka u istraživačkom procesu.

U prvom koraku smo identifikovali ULS od četiri reči u sva tri korpusa primenom minimalne frekvencije od 40 ponavljanja u milion reči i minimalne disperzije u 10% tekstova. Dobijeni spiskovi ULS prečišćeni su uklanjanjem svih ULS koji su u direktnoj vezi sa određenim sadržajem ili kontekstom, ULS koji sadrže vlastite imenice i spajanjem ULS koji sačinjavaju jedan duži ULS. Rezultati su pokazali da se najveći broj tipova i tokena identifikovao kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski (111 tipova i 4.536 tokena), koji su praćeni brojem tipova i tokena identifikovanih kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski (60 tipova i 2.358 tokena), dok se najmanji broj tipova i tokena izdvojio kod izvornih govornika engleskog jezika (27 tipova i 777 tokena). Navedeni rezultati su u skladu sa nalazima prethodnih istraživanja (Perez-Ljantada, 2014; Pan i sar., 2016; Güngör & Uysal, 2016; Güngör, 2016; Lazić, 2017; Al Fedžri i sar., 2020; Kostromitina, 2022) i na osnovu njih smo potvrdili četvrtu radnu hipotezu: *Veći broj tipova i tokena ULS biće identifikovan kod dve grupe neizvornih govornika engleskog jezika uporedno sa izvornim govornicima engleskog jezika u naučnim radovima iz oblasti obrazovanja*. Međutim, na osnovu istih podataka nismo potvrdili petu radnu hipotezu: *Sličan broj tipova i tokena biće identifikovan kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski u naučnim radovima iz oblasti obrazovanja*. Jedno od mogućih objašnjenja za veći broj izdvojenih ULS kod izvornih govornika srpskog jezika je potencijalno veća zastupljenost

prevedenih radova u SCEP korpusu. Do navedene pretpostavke smo došli poređenjem informacija koje su date u prethodnim istraživanjima (Murešan i Perez-Ljantada, 2014; Igrutinović, 2021) o tendencijama istraživača u Srbiji i Rumuniji da se oslanjaju na usluge prevodilaca u procesu objavljivanja naučnih radova na engleskom jeziku. Istraživanja koja su ispitivala ULS sa ciljem identifikacije prevodnih univerzalija utvrdili su da se veći broj ULS koristi u prevedenim tekstovima (Biel, 2017; Siao, 2010; Li, 2013; Ebeling, 2022). Međutim, nemoguće je utvrditi koji su radovi u SCEP i RCEP korpusima prevodi; stoga su potrebna dalja istraživanja.

Drugi korak u istraživanju podrazumevao je strukturalnu analizu izdvojenih ULS u sva tri korpusa. Strukturalna analiza urađena je na nivou svih ULS, zajedničkih ULS i specifičnih ULS kako bismo dobili što jasniju sliku o karakteristikama upotrebe i razlikama u upotrebi ULS kod izvornih govornika engleskog jezika i dve grupe neizvornih govornika. Za strukturalnu analizu korišćena je blago modifikovana taksonomija koju su predložili Bajber i sar. (1999). Strukturalna analiza je pokazala da je i kod izvornih govornika engleskog jezika i kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski, najzastupljenija strukturalna grupa *fragmenti sa predloškom frazom*, koja je praćena *imeničkim strukturama* i *glagolskim strukturama*, dok su *druge strukture* bile najmanje zastupljene. Na osnovu navedenih podataka potvrđena je prva radna hipoteza: *Strukturalna raspodela ULS kod izvornih govornika engleskog jezika u naučnim radovima u oblasti obrazovanja biće okarakterisana većom zastupljenošću ULS iz strukturalne grupe fragmenti sa predloškom frazom uporedno sa drugim strukturalnim grupama*. Međutim, na osnovu istih podataka samo je delimično potvrđena druga radna hipoteza: *Strukturalna raspodela ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski u naučnim radovima u oblasti obrazovanja biće okarakterisana većom zastupljenošću ULS iz strukturalnih grupa fragmenti sa predloškom frazom i glagolske strukture nego imeničke strukture i druge strukture*. Za razliku od većine prethodnih istraživanja (Pan i sar., 2016; Güngör & Uysal, 2016; Güngör, 2016; Al Fedžri et al., 2020), na osnovu kojih je i formirana druga radna hipoteza, neizvorni govornici engleskog jezika kojima je maternji jezik srpski ili rumunski pokazali su uspešnost u praćenju normi koje su postavili izvorni govornici u naučnim radovima iz oblasti obrazovanja a kada je reč o raspodeli zastupljenosti strukturalnih kategorija. S obzirom na činjenicu da su maternji jezici neizvornih govornika u prethodnim istraživanjima tipološki udaljeniji od engleskog jezika nego što su srpski i rumunski, pretpostavljamo da je tipološka razlika jedan od faktora koji je uticao na veću uspešnost neizvornih govornika u ovom istraživanju.

Strukturalna analiza zajedničkih ULS pokazala je da sve tri relacije zajedničkih ULS (zajednički za izvorne govornike engleskog i srpskog, zajedničke za izvorne govornike engleskog i rumunskog, zajedničke za dve grupe neizvornih govornika) prate raspodele koje su prisutne u strukturalnoj raspodeli celih korpusa. Stoga možemo zaključiti da neizvorni govornici engleskog jezika kojima je maternji jezik srpski ili rumunski nemaju tendenciju da uspešnije koriste određenu strukturalnu grupu, već da određeni procenat ULS iz svake grupe uspešno koriste. Strukturalna raspodela specifičnih ULS pokazala je da su kod izvornih govornika engleskog jezika i izvornih govornika srpskog najzastupljeniji *fragmenti sa predloškom frazom*, praćeni *imeničkim strukturama*, dok su kod izvornih govornika rumunskog jezika najzastupljenije *imeničke strukture* praćene *fragmentima sa predloškom frazom*. Navedeni podaci nas dovode do zaključka da izvorni govornici srpskog jezika brojačano najviše koriste neoptimalne ULS u strukturalnoj grupi *fragmenti sa predloškom frazom*, dok izvorni govornici rumunskog jezika najviše koriste neoptimalne ULS u strukturalnoj grupi *imeničke strukture*.

U trećem koraku urađena je funkcionalna analiza izdvojenih ULS u sva tri korpusa. Primenjena je funkcionalna taksonomija koju je predložio Hajland (2008), a analiza je urađena na nivou svih ULS, zajedničkih ULS i specifičnih ULS. Funkcionalna analiza tipova ULS pokazala je istu raspodelu funkcionalnih kategorija za sva tri korpusa po kojoj je najzastupljenija funkcionalna

grupa *ULS orijentisani ka istraživanju*, praćena *ULS orijentisanim ka tekstu*, dok je najmanje zastupljena grupa *ULS orijentisani ka učesniku*. Međutim, kada sagledamo raspodelu tokena, uviđamo da su tu neizvorni govornici manje uspešno ispratili norme koje su postavili izvorni govornici engleskog jezika. Naime, kod izvornih govornika engleskog jezika najzastupljeniji su *ULS orijentisani ka istraživanju*, dok su kod neizvornih govornika engleskog jezika kojima su maternji jezici srpski ili rumunski, najzastupljeniji *ULS orijentisani ka tekstu*.

Kada sagledamo funkcionalne raspodele tipova u zajedničkim ULS, uviđamo da je na sve tri relacije (zajedničke za izvorne govornike engleskog i srpskog, zajedničke za izvorne govornike engleskog i rumunskog, zajedničke za dve grupe neizvornih govornika) zastupljena ista raspodela funkcionalnih grupa i ona prati raspodelu koja je prisutna za tokene i u celim korpusima. Međutim, kada sagledamo raspodelu tokena zajedničkih ULS za izvorne govornike engleskog jezika i neizvorne govornike engleskog kojima je maternji jezik srpski ili rumunski, uviđamo da obe grupe neizvornih govornika imaju tendenciju da učestalije koriste *ULS orijentisane ka tekstu* od izvornih govornika engleskog, dok izvorni govornici engleskog jezika učestalije koriste *ULS orijentisane ka istraživanju*. Stoga smo zaključili da čak i kada neizvorni govornici koriste najoptimalnije tipove ULS za pokazivanje pripadnosti određenoj delatnoj zajednici i dalje imaju tendenciju da ih isuviše ili nedovoljno učestalo koriste. Posmatranjem funkcionalne raspodele specifičnih ULS, uvideli smo da su kod sve tri grupe govornika najviše zastupljeni *ULS orijentisani ka istraživanju* praćeni *ULS orijentisani ka tekstu*. Funkcionalna upotreba specifičnih ULS se najviše razlikuje kod izvornih govornika rumunskog jezika jer ne koriste *ULS orijentisane ka učesniku* i značajno više koriste *ULS orijentisane ka istraživanju*.

Četvrti korak podrazumevao je sagledavanje upotrebe najzastupljenijih obrazaca u sva tri korpusa. Najzastupljeniji obrasci koje smo u ovom istraživanju analizirali su: *the + imenica + of the/a, in the + imenica + of, glagol u pasivu + fragment predloške fraze* i *uvodna it struktura + glagolska/pridevska fraza*. ULS koji prate obrazac *the + imenica + of the/a* procentualno gledano u odnosu na ceo korpus, najzastupljeniji su kod izvornih govornika rumunskog jezika, praćeni zastupljenošću kod izvornih govornika engleskog, dok su najmanje zastupljeni kod izvornih govornika srpskog jezika. Sagledanjem raznolikosti reči sa punim leksičkim značenjem koje se koriste u obrascu *the + imenica + of the/a*, utvrdili smo da neizvorni govornici kojima je maternji jezik srpski pokazuju najveću raznolikost upotrebljenih reči sa punim leksičkim značenjem, praćeni neizvornim govornicima engleskog jezika kojima je maternji jezik rumunski, dok se najmanja raznolikost uočila kod izvornih govornika engleskog jezika. Procentualno gledano, tipovi ULS koji prate obrazac *in the + imenica + of* najzastupljeniji su kod izvornih govornika srpskog jezika, praćeni izvornim govornicima engleskog jezika, dok su najmanje zastupljeni kod izvornih govornika rumunskog jezika. Međutim, kada gledamo procentualnu zastupljenost tokena uviđamo da obe grupe neizvornih govornika engleskog jezika učestalije upotrebljuju ULS koji prate obrazac *in the + imenica + of* od izvornih govornika engleskog jezika. Najveća raznolikost reči sa punim leksičkim značenjem u obrascu *in the + imenica + of* zabeležila se kod izvornih govornika srpskog jezika, praćena izvornim govornicima rumunskog jezika, dok je najmanja raznolikost zabeležena kod izvornih govornika engleskog jezika. Obrazac *glagol u pasivu + fragment predloške fraze* našao se samo kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski sa tri ULS. Poslednji od najučestalijih obrazaca koji smo sagledali *uvodna it struktura + glagolska/pridevska fraza* najzastupljeniji je, procentualno gledano, kod izvornih govornika engleskog jezika, praćeni izvornim govornicima srpskog jezika, dok je najmanje zastupljen kod izvornih govornika rumunskog jezika. Najveća raznolikost upotrebe reči sa punim leksičkim značenjem u ULS koji prate obrazac *uvodna it struktura + glagolska/pridevska fraza* uviđa se kod izvornih govornika srpskog jezika. Poređenjem navedenih podataka sa prethodnim istraživanjima došli smo do zaključka da se neizvorni govornici engleskog jezika kojima je maternji jezik rumunski razlikuju po upotrebi najučestalijih obrazaca uporedno sa drugim grupama neizvornih govornika. Na osnovu

rezultata o upotrebi najučestalijih obrazaca ULS u korpusima u ovom istraživanju, potvrdili smo osmu radnu hipotezu: *Najzastupljeniji okviri će se drugačije koristiti kod izvornih govornika engleskog jezika i dve grupe neizvornih govornika engleskog jezika u naučnim radovima u oblasti obrazovanja.*

Peti korak je podrazumevao komparaciju podataka izdvojenih u prethodnim koracima o zastupljenosti ULS, strukturalnoj raspodeli, funkcionalnoj raspodeli i upotrebi najzastupljenijih obrazaca u sva tri korpusa. Kratak opis o zaključcima komparacija je dat uz opis svakog od koraka.

U šestom koraku smo sprovedli statističku analizu podataka o strukturalnoj i funkcionalnoj raspodeli na grupe i podgrupe svih i specifičnih ULS na sledećim relacijama: između izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski, između izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski, između neizvornih govornika engleskog jezika kojima je maternji jezik srpski i neizvornih govornika engleskog jezika kojima je maternji jezik rumunski. Hi-kvadrat test je korišćen za statističku analizu koja je za cilj imala da se utvrdi koje su od identifikovanih razlika između korpusa zapravo i statistički značajne, odnosno koje mogu da se generalizuju. Primenom navedenog testa na strukturalnu raspodelu svih ULS u SCEP, ECEP i RCEP korpusima došli smo do zaključaka da izvorni govornici srpskog jezika koriste značajno više tokena ULS u strukturalnoj grupi *druge strukture* od izvornih govornika engleskog i rumunskog jezika, izvorni govornici rumunskog koriste značajno više tokena u grupi *imeničke strukture* i *fragmenti sa predloškom frazom* od izvornih govornika engleskog i srpskog jezika, dok izvorni govornici engleskog jezika koriste značajno više tipova u grupi *fragmenti sa predloškom frazom* od izvornih govornika srpskog jezika a značajno više tipova u grupi *glagolske strukture* od izvornih govornika rumunskog jezika. Statistička analiza funkcionalne raspodele svih ULS u tri korpusa u ovom istraživanju na gorenavedenim relacijama pokazala je da obe grupe neizvornih govornika imaju tendenciju da isuviše učestalo koriste *ULS orijentisane ka tekstu*, dok nedovoljno učestalo koriste *ULS orijentisane ka učesniku*, preciznije, da obe grupe neizvornih govornika isuviše učestalo koriste funkcionalnu podgrupu *ograničavajući signali*, dok nedovoljno učestalo koriste podgrupe *ULS za lokaciju*, *strukturalni signali* i *ULS stava*. Takođe smo uvideli da izvorni govornici srpskog jezika značajno učestalije koriste *ULS orijentisane ka učesniku*, dok izvorni govornici rumunskog jezika značajno učestalije koriste *ULS orijentisane ka istraživanju*.

Primenom statističke analize na strukturalnu raspodelu specifičnih ULS u sva tri korpusa na navedenim relacijama, došli smo do zaključaka da se strukturalna raspodela specifičnih ULS kod izvornih govornika rumunskog jezika više razlikuje od strukturalne raspodele specifičnih ULS kod izvornih govornika engleskog jezika, nego što se strukturalna raspodela specifičnih ULS kod izvornih govornika srpskog jezika razlikuje od strukturalne raspodele specifičnih ULS kod izvornih govornika engleskog jezika. Međutim, najveće razlike u strukturalnoj raspodeli specifičnih ULS uvidele su se između dve grupe neizvornih govornika. Primenom statističke analize na funkcionalnu raspodelu specifičnih ULS, došli smo do zaključka da izvorni govornici srpskog jezika imaju tendenciju da češće koriste *ULS orijentisane ka tekstu*, uporedno sa izvornim govornicima engleskog i sa izvornim govornicima rumunskog jezika. Takođe, utvrdilo se da izvorni govornici rumunskog jezika imaju tendenciju da češće koriste *ULS orijentisane ka istraživanju* od izvornih govornika engleskog i od izvornih govornika srpskog jezika. Izvorni govornici engleskog jezika imaju tendenciju da češće koriste *ULS orijentisane ka učesniku* uporedno sa dve grupe neizvornih govornika engleskog jezika. Na osnovu podataka o statističkoj analizi razlika u strukturalnoj i funkcionalnoj raspodeli potvrđujemo šestu radnu hipotezu: *Postojeće statistički značajne razlike u strukturalnoj i funkcionalnoj raspodeli ULS na grupe i podgrupe između izvornih i neizvornih govornika engleskog jezika, kao i između dve grupe neizvornih govornika.*

U sedmom i poslednjem koraku urađena je analiza uticaja negativnog transfera na produkciju ULS kod neizvornih govornika engleskog jezika. Za identifikaciju uticaja transfera korišćen je Džarvisov pristup zasnovan na detekciji (Džarvis, 2000, 2012; Džarvis i Pavlenko, 2008). Ustanovili smo da je 29,37% ULS koje koriste izvorni govornici srpskog jezika pod uticajem negativnog transfera, dok je kod izvornih govornika rumunskog 13,33% ULS pod uticajem negativnog transfera. Prikupljenim dokazima potvrdili smo sedmu radnu hipotezu: ULS koji se koriste zbog uticaja transfera sačinjavaće mali procenat (do 40%) ULS identifikovanih u korpusima neizvornih govornika engleskog jezika kojima su maternji jezici srpski ili rumunski u naučnim radovima iz oblasti obrazovanja.

Na osnovu rezultata koje smo dobili u ovom istraživanju sprovođenjem sedam gorenavedenih koraka možemo da zaključimo da je potvrđena i osnovna hipoteza istraživanja: *Karakteristike upotrebe ULS u naučnim radovima iz oblasti obrazovanja će se razlikovati kod izvornih govornika engleskog jezika i neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski, u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca.* Sprovođenjem istraživanja uspešno smo ispunili sekundarne ciljeve istraživanja, kao i glavni cilj istraživanja: *Ustanoviti razlike u karakteristikama upotrebe ULS u naučnim radovima iz oblasti obrazovanja kod izvornih i neizvornih govornika engleskog jezika u smislu broja izdvojenih ULS, strukturalne raspodele, funkcionalne raspodele i upotrebe najzastupljenijih obrazaca.*

Doprinos sprovedenog istraživanja se ogleda u tome što, koliko je autorki poznato, do sada nije izvršeno istraživanje ULS u naučnim radovima iz određene oblasti u kojem je u korpusima iste veličine upoređena primena ULS kod dve grupe neizvornih govornika i jedne izvornih. Zbog navedene postavke istraživanja mogli smo preciznije da uporedimo upotrebu ULS kod dve grupe neizvornih govornika sa različitim maternjim jezicima i bolje razumemo koje su opšte karakteristike upotrebe ULS kod neizvornih govornika, a koje su specifične za neizvorne govornike kojima je maternji jezik srpski i neizvorne govornike kojima je maternji jezik rumunski. Doprinos sprovedenog istraživanja se ogleda i u činjenici da je upotreba ULS u međujeziku izvornih govornika srpskog jezika i rumunskog jezika slabo istražena. Podaci o karakteristikama upotrebe ULS kod dve grupe neizvornih govornika pozitivno će uticati na unapređenje podučavanja studenata na časovima akademskog engleskog, kao i u stručnom usavršavanju ili podučavanju prevodilaca koji žele da se bave prevodom naučnih članaka iz oblasti obrazovanja.

Ograničenja istraživanja ogledaju se u tome što:

- korpusi nisu ujednačeni i po broju radova, kako bi doprinelo preciznosti dobijenih podataka;
- nemamo informacije o tome koji su radovi u SCEP i RCEP korpusima pisani na engleskom jeziku a koji su prevedeni;
- analizirani su samo ULS od četiri reči;
- nisu napravljeni korpusi naučnih radova iz oblasti obrazovanja na srpskom i rumunskom jeziku kako bi se preciznije prikupili dokazi za uticaj transfera.

Iako smo podacima iz istraživanja sprovedenog u ovoj disertaciji dobili odgovore na pitanja o karakteristikama upotrebe ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski i kod neizvornih govornika engleskog jezika kojima je maternji jezik rumunski, takođe smo otvorili i nova pitanja. Činjenica da su se neizvorni govornici engleskog jezika više razlikovali u upotrebi ULS međusobno nego što su se razlikovali od izvornih govornika engleskog jezika, otvara pitanja koliko je velika razlika u upotrebi ULS kod različitih grupa neizvornih govornika u drugim kontekstima i kakve će to imati implikacije za nastavu akademskog engleskog jezika. Štaviše,

neophodno je sprovesti istraživanja upotrebe ULS kod neizvornih govornika engleskog jezika kojima je maternji jezik srpski ili rumunski u naučnim radovima iz drugih oblasti, kao i u drugim akademskim žanrovima kako bismo mogli da dobijemo potpuniju sliku karakteristika upotrebe ULS kod navedenih grupa govornika. Razlike u broju identifikovanih ULS kod dve grupe neizvornih govornika sugerišu da prevod naučnih radova može da bude potencijalno veoma značajan faktor koji je u prethodnim istraživanjima zanemaren zbog praksi objavljivanja naučnih radova u kojima ne dobijamo informacije o tome da li su radovi prevedeni ili ne. Dakle, neophodna su buduća istraživanja koja će ispitati uticaj faktora prevoda u istraživanjima koja se bave karakteristikama upotrebe ULS kod izvornih i neizvornih govornika.

Napisati uspešan naučni rad na stranom jeziku je izazovan zadatak. Pored kvaliteta istraživanja i ponuđenih informacija, neizvorni govornici imaju i poteškoću stvaranja uverljivog akademskog glasa kako bi njihova poruka bila prihvaćena. Nadamo se da će rezultati u ovom istraživanju pomoći studentima, istraživačima i prevodiocima u prevazilaženju nekih od prepreka u procesu pisanja naučnih radova i uspešnoj integraciji na globalnoj akademskoj sceni.

7. Literatura

- Abdollahpour, Z., & Gholami, J. (2018). Building blocks of medical abstracts: Frequency, functions and structures of lexical bundles. *Asian ESP Journal*, 14(1), 82-110.
- Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for specific purposes*, 31(2), 81-92.
- Akbulut, F. D. (2020). A bibliometric analysis of lexical bundles usage in native and non-native academic writing. *Dil ve Dilbilimi Çalışmaları Dergisi*, 16(3), 1146-1166.
- Al Fajri, M. S., Wasito, A., & Kirana, C. I. K. P. (2020). Lexical Bundles of L1 and L2 English Professional Scholars: A Contrastive Corpus-Driven Study on Applied Linguistics Research Articles. *Journal of Language & Education Volume*, 6(4), 76-89.
- Alasmary, A. A. (2018). Participant-oriented Lexical Bundles in Graduate-level Math Textbooks: A Corpus-based Exploration. *Corpus Linguistics Research*, 4, 99-114.
- Allan, R. (2016). Lexical bundles in graded readers: To what extent does language restriction affect lexical patterning? *System*, 59, 61-72. <https://doi.org/10.1016/j.system.2016.04.005>
- Allan, R. (2017). From Do You Know to I Don't Know: An Analysis of the Frequency and Usefulness of Lexical Bundles in Five English Language Self-Study Books. *Corpus Pragmatics*, 1(4), 351-372.
- Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. *Phraseology: Theory, Analysis and Applications*. Anthony P. C. (ed.). Oxford: Oxford University Press, 101-122.
- Anthony, L. (2013). A critical look at software tools in corpus linguistics. *Linguistic Research*, 30(2), 141-161.
- Appel, R., & Murray, L. (2020). L1 differences in L2 English academic writing: A lexical bundles analysis. *Journal of English for Academic Purposes*, 46, 100873.
- Bal, B. (2010). *Analysis of four-word lexical bundles in published research articles written by Turkish authors* (Unpublished MA dissertation). Georgia State University, Atlanta, GA.
- Baker, M. (2004). A corpus-based view of similarity and difference in translation. *International journal of corpus linguistics*, 9(2), 167-193.
- Baker, M., & Perez-Gonzalez, L. (2011) Translation and interpreting. *The Routledge Handbook of Applied Linguistics*. Simpson, J. (Ed.). Taylor & Francis, 39-52.
- Barnardini, S., (2015) Translation. *The Cambridge handbook of English corpus linguistics*. Biber, D., & Reppen, R. (Eds.). Cambridge University Press. 515-536.
- Berūkštienė, D. (2017). A corpus-driven analysis of structural types of lexical bundles in court judgments in English and their translation into Lithuanian. *Kalbotyra*, (70), 7-31.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Longman grammar of spoken and written English. Harlow: Pearson.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied linguistics*, 25(3), 371-405.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for specific purposes*, 26(3), 263-286.
- Biel, Ł. (2017). Lexical bundles in EU law: the impact of translation process on the patterning of legal language. *Phraseology in legal and institutional settings. A corpus-based interdisciplinary perspective (e-book)*. S. Goźdz-Roszkowski and G. Pontrandolfo London, New York: Routledge, 10-26.
- Bercuci, L. (2020). Discipline-specific metadiscourse markers in ESP expert writing in political science. In R.M. Nistor & C. Teglas (Eds.), *Limbajele specializate în contextual noilor medii de învățare: Provocări și oportunități* (pp. 331-345). Cluj: Presa Universitară Clujeană.

- Bercuci, L. (2021). The use of metadiscourse markers in the case of students of political science in L2 English. *Journal of Linguistic Intercultural Education*, 14(1), 7-20.
- Bestgen, Y. (2019). Comparing lexical bundles across corpora of different sizes: the Zipfian problem. *Journal of Quantitative Linguistics*, 27(3), 272-290.
- Beng, C. O. S., & Keong, Y. C. (2017). Comparing structural and functional lexical bundles in MUET reading test. *Pertanika Journal of Social Sciences and Humanities*, 25(1), 133-148.
- Breeze, R. (2013). Lexical bundles across four legal genres. *International journal of corpus linguistics*, 18(2), 229-253.
- Bychkovska, T., & Lee, J. J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes*, 30, 38-52.
- Candarli, D., & Jones, S. (2019). Paradigmatic influences on lexical bundles in research articles in the discipline of education. *Corpora*, 14(2), 237-263.
- Casal, J. E., Lu, X., Qiu, X., Wang, Y., & Zhang, G. (2021). Syntactic complexity across academic research article part-genres: A cross-disciplinary perspective. *Journal of English for Academic Purposes*, 52, 100996.
- Chen, L. (2008). *An investigation of lexical bundles in Electrical Engineering introductory textbooks and ESP textbooks* [Master of Arts, Carleton University]. <https://doi.org/10.22215/etd/2008-07668>
- Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language learning & technology*, 14(2), 30-49.
- Chen, Y. H., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied linguistics*, 37(6), 849-880.
- Chen, L. (2018). *Lexical Bundles in Vocabulary-based Discourse Units: A Corpus-based Study of First Year Core Engineering Textbooks* [Doctor of Philosophy, Carleton University]. <https://doi.org/10.22215/etd/2018-13399>
- Cheng, W., Greaves, C., & Warren, M. (2006). From n-gram to skipgram to concgram. *International journal of corpus linguistics*, 11(4), 411-433.
- Chitez, M., Mureşan, V., & Rogobete, R. (2022). How to write good academic papers: using the EXPRES corpus to extract expert writing linguistic patterns. In *Conference Proceedings. The Future of Education 2022*.
- Chitez, M., Strilechi, C., & Csürös, K. (2022). Meeting ROGER: An open-access bilingual corpus search platform. In *Book of Abstracts of 6th Learner Corpus Research Conference*, 181.
- Cowie, A. P. (1992). Multiword lexical units and communicative language teaching. In Arnaud, P. & Bejoint, H. (Eds.), *Vocabulary and Applied Linguistics*, 1-12. London: Macmillan.
- Cowie, A. P. (1994). Phraseology. In Asher, R.E. (ed.). *The Encyclopedia of Language and Linguistics*, 3168-3171. Oxford: Oxford University Press.
- Crossley, S., & Salsbury, T. L. (2011). The development of lexical bundle accuracy and production in English second language speakers. *International Review of Applied Linguistics in Language Teaching*, 49 (1), 1-26. doi: 10.1515/iral.2011.001
- Cortes, V. 2004. Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, (23), 397-423.
- Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for academic purposes*, 12(1), 33-43.
- Даничић М. (2018). Неке дилеме, недоумице и изазови у настави превођења на академском нивоу. *Иновације у настави XXXI*, 2018/1. Београд: Учитељски факултет у Београду, 141–147, Doi: 10.5937/inovacije1801141D
- Dang, T. N. Y. (2018). The nature of vocabulary in academic speech of hard and soft-sciences. *English for Specific Purposes*, 51, 69-83.
- De Cock, S., & Granger, S. (2021). Stance in press releases versus business news: a lexical bundle approach. *Text & Talk*, 41(5-6), 691-713.

- Dinca, A., & Chitez, M. (2021). Assessing learners' academic phraseology in the digital age: a corpus-informed approach to ESP texts. *Journal of Teaching English for Specific and Academic Purposes*, 071-084.
- Dontcheva-Navratilova, O. (2013). Lexical bundles indicating authorial presence: A cross-cultural analysis of novice Czech and German writers' academic discourse. *Discourse and Interaction*, 6(1), 7-24.
- Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, 38(2), 165-193.
- Đorđević, D., & Lazić, K. (2021). Učestali leksički spojevi u nastavi engleskog kao stranog jezika struke na master akademskim studijama. *Анали Филолошког факултета*, 33(1), 193-207.
- Ebeling, S. O. (2022). The function of recurrent word-combinations in English translations from three different languages. *Meta*, 67(1), 143-169.
- Ellis, N. C., Simpson-Vlach, R. I. T. A., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL quarterly*, 42(3), 375-396.
- Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual review of applied linguistics*, 32, 17-44.
- Ellis, R. (2015). *Understanding second language acquisition 2nd edition*. Oxford university press.
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text & Talk*, 20(1), 29-62.
- Esfandiari, R., & Barbary, F. (2017). A contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles. *Journal of English for Academic Purposes*, 29, 21-42.
- Fadhilah, U. (2020). Structural and functional analyses of lexical bundles in news articles written by Indonesian journalists. *Konferensi Linguistik Tahunan Atma Jaya*, 18, 403-408.
- Filipović, J., & Vučo, J. (2013). Small area languages in global academic settings. *Actualizaciones en comunicación social*, 1, 136-139.
- Filipović, J. (2018). *Moć reči: Ogledi iz kritičke sociolingvistike*. Zadužbina Andrejević.
- Garner, J., Crossley, S., & Kyle, K. (2020). Beginning and intermediate L2 writer's use of N-grams: An association measures study. *International Review of Applied Linguistics in Language Teaching*, 58(1), 51-74.
- Giczela-Pastwa, J. (2021). Developing phraseological competence in L2 legal translator trainees: a proposal of a data mining technique applied in translation from an LLD into ELF. *The Interpreter and Translator Trainer*, 15(2), 187-204.
- Gilmore, A., & Millar, N. (2018). The language of civil engineering research articles: A corpus-based approach. *English for Specific Purposes*, 51, 1-17.
- Glušac, T., & Milić, M. (2011). Engleski jezik u nauci – pravila za istraživače u turizmu. *Civitas*, 130–139.
- Gray, B & Biber, D. (2015). Phraseology. In Biber, D., & Reppen, R. (Eds.). *The Cambridge handbook of English corpus linguistics*. Cambridge University Press.
- Greaves, C., & Warren, M. (2010). What can a corpus tell us about multi-word units? In *The Routledge handbook of corpus linguistics*. Routledge, 212-226.
- Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In Granger, S. & Meunier, F. (Eds.), *Phraseology An interdisciplinary perspective*, Amsterdam: John Benjamin Publishing Company, 28-49.
- Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, 1(1), 7-24.
- Güngör, F., & Uysal, H. H. (2016). A Comparative Analysis of Lexical Bundles Used by Native and Non-native Scholars. *English Language Teaching*, 9(6), 176-188.

- Güngör, F. (2016). Cross-linguistic analysis of lexical bundles in L1 English, L2 English, and L1 Turkish research]. *Unpublished doctoral thesis*. Gazi University/Eğitim Bilimleri Enstitüsü, Ankara.
- Güngör, F. (2019). It was found that...: Introductory it Patterns by Native and Non-Native Authors. *Eurasian Journal of Applied Linguistics*, 5(3), 473-493.
- Güngör, F., & Uysal, H. H. (2020). Lexical bundle use and crosslinguistic influence in academic texts. *Lingua*, 242, 102859.
- Herbel-Eisenmann, B., & Wagner, D. (2010). Appraising lexical bundles in mathematics classroom discourse: Obligation and choice. *Educational Studies in Mathematics*, 75(1), 43-63. <https://doi.org/10.1007/s10649-010-9240-y>
- Heng, C. S., Kashiha, H., & Tan, H. (2014). Lexical Bundles: Facilitating University" Talk" in Group Discussions. *English Language Teaching*, 7(4), 1-10.
- Houssos, V. (2020). Language transfer: A study on the notion of transfer and on the cross-linguistic lexical influence. *Multilingual Academic Journal of Education and Social Sciences*, 8(1), 268-279.
- Hyland, K. (2005) A convincing argument: corpus analysis and academic persuasion. In Connor, U. & Upton, T. (Eds.) *Discourse in the Professions: Perspectives from Corpus Linguistics*. Amsterdam: Benjamins, 87-114.
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for specific purposes*, 27(1), 4-21.
- Hyland, K., 2009. *Academic Discourse: English in a Global Context*. United Kingdom Continuum, London.
- Hyland, K. (2012). Bundles in academic discourse. *Annual review of applied linguistics*, 32, 150.
- Hyland, K., & Jiang, F. (2022). Bundles in advanced EAL authors' articles: How do they compare with world Englishes practices? *World Englishes*, 41(4), 554-570.
- Hussain, G., Zahra, T., & Abbas, A. (2021). Discourse Functions of Lexical Bundles in Pakistani Chemistry and Physics Textbooks. *GEMA Online Journal of Language Studies*, 21(1).
- Igrutinović, S. M. (2021). Dominacija engleskog jezika u naučnim publikacijama–iskustva i stavovi nastavnika u visokoškolskoj ustanovi u Srbiji. *Reči*, (14), 34–57.
- Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in them interlanguage lexicon. *Language learning*, 50(2), 245-309.
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge.
- Kashiha, H., & Chan, S. (2013). An exploration of lexical bundles in academic lectures: Examples from hard and soft sciences. *Journal of Asia TEFL*, 10, 133-161.
- Kashiha, H., & Swee Heng, C. (2013). Structural Analysis of Lexical Bundles in University Lectures of Politics and Chemistry. *International Journal of Applied Linguistics & English Literature*, 3(1), 224-230. <https://doi.org/10.7575/aiac.ijalel.v.3n.1p.224>
- Kobson (2024). Preuzeto 7. januara 2024, sa https://kobson.nb.rs/nauka_u_srbiji/referisani_casopisi.23.html#371
- Kostromitina, M. (2022). Cross-linguistic transfer in academic journal writing: Evidence from lexical bundle analysis in Russian and English. *Research in Corpus Linguistics*, 10(2), 70-112.
- Kwon, Y., & Lee, E.-J. (2014). Lexical bundles in the Korean EFL teacher talk corpus: A comparison between non-native and native English teachers. *Journal of Asia TEFL*, 11, 73-103.
- Lazić, K. O. (2017). *Učestali leksički spojevi u engleskom jeziku biotehničke struke: Korpusna analiza radova izvornih i neizvornih govornika* (Doktorska disertacija, Univerzitet u Beogradu – Filološki fakultet).
- Lazić, K. O. (2019). Application of lexical bundles from biotechnical research articles in ESP teaching. *Glasnik Šumarskog fakulteta*, (119), 77-90.
- Levy, S. A. (2003). *Lexical bundles in professional and student writing*. University of the Pacific.

- Lee, C. (2013). Using lexical bundle analysis as discovery tool for corpus-based translation research. *Perspectives*, 21(3), 378-395.
- Liu, K., & Afzaal, M. (2021). Translator's Style Through Lexical Bundles: A Corpus-Driven Analysis of Two English Translations of Hongloumeng. *Frontiers in Psychology*, 12, 633422.
- Liu, C. Y., & Chen, H. J. H. (2020). Functional variation of lexical bundles in academic lectures and TED talks. *Register Studies*, 2(2), 176-208.
- Luo, N., & Hyland, K. (2019). "I won't publish in Chinese now": Publishing, translation and the non-English speaking academic. *Journal of English for Academic Purposes*, 39, 37-47.
- McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
- McManus, K. (2021). *Crosslinguistic influence and second language learning*. Routledge.
- Meunier, F., & Granger, S. (Eds.). (2008). *Phraseology in foreign language learning and teaching*. John Benjamins Publishing.
- Montgomery, S. L. (2010). Scientific translation. *Handbook of translation studies*, 1, 299-305.
- Moon, R. (1998). Frequencies and forms of phrasal lexemes in English. In Cowie, A. P. (ed.) *Phraseology. Theory, Analysis, and Applications*, 79-100. Oxford: Oxford University Press.
- Munday, J., Ramos Pinto, S. & Blakesly, J. (2022). *Introducing Translation Studies Theories and Application* (5th ed.). Routledge.
- Muresan, L. M., & Pérez-Llantada, C. (2014). English for research publication and dissemination in bi-/multiliterate environments: The case of Romanian academics. *Journal of English for Academic Purposes*, 13, 53-64.
- Nedelcu Pascu, M., & Brad, L. (2018). We All Speak English, but Do We Write in the Same Language? *Proceedings of the international conference on economics and social sciences*, 1(1), 179-186.
- Nam, D. (2017). Functional Distribution of Lexical Bundle in Native and Non-Native Students' Argumentative Writing. *Journal of Asia TEFL*, 14(4), 703.
- Neely, E., & Cortes, V. (2009). A little bit about: Analyzing and teaching lexical bundles in academic lectures. *Language Value*, 1(1), 17-38
- Nesselhauf, N. (2004). What are collocations? In Allerton, D.J., N. Nesselhauf & P. Skandera (eds.) *Phraseological Units: Basic Concepts and their Application*, 1-21. Basel: Schwabe.
- Novita, H., & Kwary, D. A. (2018). Comparing the use of lexical bundles in Indonesian-English translation by student translators and professional translators. *Translation & Interpreting, The*, 10(1), 53-74.
- Odlin, T. (1989). *Language transfer: Cross- linguistic influence in language learning*. Cambridge University Press.
- Ozturk, Y. (2014). *Lexical bundle use of Turkish and native English writers: A corpus-based stud.* (Unpublished MA dissertation). Anadolu University, Eskisehir, Turkey.
- Öztürk, Y., & Taşçı, S. (2023). A Corpus-Based Analysis of Lexical Bundles in Non-Native Post Graduate Academic Writing and a Potential L1 Influence. *REFlections*, 30(2), 488-505.
- Palinkašević, R. & Prtljaga, J. (2024). Changes in the usage of lexical bundles in the field of education during the last 20 years. *Journal Plus Education*, 25 (1), 50-71.
- Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes*, 21, 60-71.
- Pan, F., Reppen, R., & Biber, D. (2020). Methodological issues in contrastive lexical bundle research: The influence of corpus design on bundle identification. *International Journal of Corpus Linguistics*, 25(2), 216-230.
- Panthong, P., & Poonpon, K. (2020). Functional Analysis of Lexical Bundles in Doctor Talks in the Medical TV Series " Grey's Anatomy". *LEARN Journal: Language Education and Acquisition Research Network*, 13(2), 335-353.

- Paquot, M. (2013). Lexical bundles and L1 transfer effects. *International Journal of Corpus Linguistics*, 18(3), 391-417.
- Pang, W. (2010). Lexical bundles and the construction of an academic voice: A pedagogical perspective. *Asian EFL Journal*, 47(1), 10-11.
- Pérez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: Convergent and divergent usage. *Journal of English for Academic Purposes*, 14, 84-94.
- Поповић, З. (2010). Програми за етикетирање текста на српском језику. *ИНФОтека*, 11 (2), 19–36.
- Рум, А. (2023). *Exploring translation theories* (3rd ed.). Taylor & Francis.
- Reiss, K., Nord, C., & Vermeer, H. J. (2014). *Towards a general theory of translational action: Skopos theory explained*. Routledge.
- Ren, J. (2021). Variability and functions of lexical bundles in research articles of applied linguistics and pharmaceutical sciences. *Journal of English for Academic Purposes*, 50, 100968.
- Rennie, D., Yank, V., & Emanuel, L. (1997). When authorship fails: a proposal to make contributors accountable. *Jama*, 278(7), 579-585.
- Ribeck, J., & Borin, L. (2011). Lexical Bundles in Swedish Secondary School Textbooks. In *Language and Technology Conference* (pp. 238-249). Springer, Cham.
- Safarzadeh, M. M., Monfared, A., & Sarfeju, M. (2015). Native and non-native use of lexical bundles in discussion section of political science articles. *Iranian Journal of Applied Language Studies*, 5(2), 137-166.
- Salazar, D. L. (2010). Lexical bundles in Philippine and British scientific English. *Philippine Journal of Linguistics*, 41(1), 94-109.
- Salazar, D. (2014). *Lexical bundles in native and non-native scientific writing*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Schmitt, N. (Ed.). (2004). *Formulaic sequences: Acquisition, processing, and use* (Vol. 9). John Benjamins Publishing.
- SciELF. (2015). *The SciELF corpus*. Available at <http://www.helsinki.fi/elfa/>.
- Shin, Y. K. (2019). Do native writers always have a head start over nonnative writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes*, 40, 1-14.
- Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied linguistics*, 31(4), 487-512.
- Simpson, J. (Ed.). (2011). *The Routledge handbook of applied linguistics*. Taylor & Francis.
- Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for academic purposes*, 12(3), 214-225.
- Stolić, J. (2015). The development of preschool children's musical abilities through specific types of musical activities. *Research in Pedagogy*, 5(2), 31-44.
- Stubbs, M. (2007). An example of frequent English phraseology: Distribution, structures and functions. In Facchinetti, R. (ed.) *Corpus Linguistics 25 Years*, 89-105. Amsterdam: Rodopi.
- Susianti, E. N., & Fitriati, S. W. (2021). Structural and Functional Lexical Bundles in Woman's Hour Daily Podcast Interview at Bbc Radio 4. *The Journal of Educational Development*, 9(1), 64-70.
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Šipka, D. (2006). *Osnovi leksikologije i srodnih disciplina*, drugo, izmenjeno i dopunjeno izdanje. Novi Sad: Matica srpska.
- Toury, G. (2012). *Descriptive translation studies and beyond: Revised Edition* (Vol. 100). John Benjamins Publishing.
- Утвић, М. (2011). Анотација Корпуса савременог српског језика. *ИНФОтека*, 12, 39–51.

- Wood, A. (2001). International scientific English: The language of research scientists around the world. In J. Flowerdew & M. Peacock (Authors), *Research Perspectives on English for Academic Purposes* (Cambridge Applied Linguistics, pp. 71-83). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139524766.008
- Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics*, 12(1), 39-57.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge University Press.
- Xiao, R. (2010). Idioms, word clusters, and reformulation markers in translational Chinese: Can “translation universals” survive in Mandarin? *The International Symposium on Using Corpora in Contrastive and Translation Studies*. Ormskirk: Edge Hill University, 1-40.

Prilozi

Prilog 1. Izdvojeni ULS u korpusu izvornih govornika engleskog jezika – spisak

br.	ULS	rang	frek	disp.	br.	ULS	rang	frek	disp
1	the extent to which	1	52	24	15	of teaching and learning	16	25	12
2	as well as the	2	47	31	16	the nature of the	18	24	19
3	it is important to	3	46	26	17	the use of the	19	24	10
4	at the end of (the)	4	44	25	18	the ways in which	20	24	14
5	are more likely to	5	34	19	19	as part of the	21	22	16
6	as a result of	6	33	21	20	at the time of	22	22	16
7	in relation to the	7	32	20	21	the purpose of this	23	22	15
8	at the same time	8	31	25	22	the role of the	24	22	13
9	in the context of	9	30	22	23	it is possible that	25	21	11
10	in the current study	10	30	10	24	over the course of	26	21	10
11	in terms of the	11	28	19	25	to be able to	27	21	13
12	the results of the	12	28	17	26	on the development of	28	20	12
13	in the form of	13	27	19	27	the quality of the	29	20	14
14	on the other hand	14	27	20					

Prilog 2. Izdvojeni ULS u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik rumunski – spisak

br.	ULS	rang	frek	disp	br.	ULS	rang	frek	disp.
1	in the case of (the)	1	137	45	31	of the educational process	36	31	18
2	(and) on the other hand (the)	2	130	63	32	the total number of	37	31	18
3	in the field of	3	88	45	33	when it comes to	38	31	18
4	at the end of (the)	4	78	46	34	in terms of the	39	30	22
5	at the same time	5	78	48	35	an important role in	40	29	23
6	as well as the	6	76	54	36	taking into account the	41	29	25
7	the results of the	7	71	39	37	the point of view (of)	44	28	16
8	in the context of (the)	9	62	35	38	is one of the	46	27	22
9	at the level of	10	52	27	39	the level of the	47	27	23
10	the fact that the	11	51	38	40	with the help of	49	27	23
11	to the development of	12	51	28	41	it is necessary to	50	26	22
12	one of the most (important)	13	50	34	42	by the fact that	51	25	19
13	on the one hand	14	46	29	43	more than half of	52	25	16
14	the quality of the	15	46	24	44	the importance of the	54	25	18
15	in the process of	16	45	30	45	the purpose of the	55	25	20
16	from the perspective of	17	44	32	46	by the ministry of	56	24	16
17	the development of the	18	41	30	47	in the development of	57	24	20
18	the way in which	19	41	25	48	than half of the	58	24	17
19	as a result of	20	39	27	49	the number of students	59	24	15
20	(at) the beginning of the	21	39	29	50	the role of the	60	24	19
21	that there is a	24	38	25	51	the existence of a	61	23	17
22	for the development of	25	37	23	52	of the fact that	62	22	18
23	the analysis of the	26	37	25	53	the extent to which	64	22	15
24	the content of the	28	36	15	54	a high level of	65	20	14
25	in relation to the	29	35	23	55	in order to achieve	66	20	17
26	to be able to	30	35	24	56	in the number of	67	20	15
27	to the fact that	31	35	25	57	the majority of the	69	20	15
28	in accordance with the	33	33	18	58	the purpose of this	70	20	16
29	in the educational process	34	33	22	59	the same time the	71	20	17
30	it is important to	35	31	20	60	the development of a	68	20	15

Prilog 3. Izdvojeni ULS u korpusu neizvornih govornika engleskog jezika kojima je maternji jezik srpski – spisak

br.	ULS	rang	frek	disp	br.	ULS	rang	frek	disp.
1	as well as the	1	178	68	57	in the area of	67	30	13
2	in the field of (education)	2	141	46	58	from the perspective of	71	29	16
3	when it comes to (the)	3	124	44	59	to be able to	72	29	19
4	the results of the (research)	4	122	49	60	in comparison to the	73	28	13
5	on the other hand (the)	5	107	54	61	the content of the	75	28	14
6	it is necessary to	6	104	50	62	are shown in table	76	27	15
7	that there is a	7	96	44	63	as a result of	77	27	18
8	at the same time	8	92	44	64	by the fact that	79	27	21
9	on the basis of (the)	10	82	43	65	in the domain of	80	27	14
10	in the process of	11	70	39	66	significant differences in the	81	27	17
11	(it) can be concluded (that the)	12	66	27	67	the quality of teaching	82	27	12
12	it is important to	13	66	44	68	with the aim of	83	27	20
13	for the development of	14	65	36	69	a high level of	84	26	20
14	in the context of	15	65	27	70	as a part of	85	26	12
15	in the case of	17	63	22	71	as well as their	86	26	23
16	the fact that the	18	63	42	72	in line with the	87	26	16
17	in accordance with the	19	62	35	73	in terms of the	88	26	18
18	in relation to the	20	61	30	74	it can be seen	89	26	14
19	for the implementation of	21	55	12	75	on the one hand	90	26	19
20	(at) the end of the	22	54	26	76	statistically significant differences in	91	26	13
21	the implementation of the	23	54	20	77	that there are no	92	26	10
22	in the implementation of	24	52	17	78	the aim of the	93	26	18
23	it is possible to	26	51	29	79	the application of the	94	26	14
24	a statistically significant difference	27	49	15	80	the characteristics of the	95	26	16
25	the development of the	28	49	31	81	the results show that	96	26	21
26	to the development of	29	49	31	82	to the conclusion that	97	26	21
27	in the teaching process	30	48	10	83	a small number of	98	25	18
28	that there is no	32	46	20	84	of the concept of	99	25	11
29	is one of the	33	45	31	85	of the fact that	100	25	20
30	one of the most (important)	34	44	30	86	of the importance of	101	25	17

br.	ULS	rang	frek	disp	br.	ULS	rang	frek	disp.
31	the importance of the	35	44	24	87	statistically significant difference in	102	25	11
32	the quality of the	36	44	19	88	the assessment of the	103	25	14
33	the total number of	37	44	18	89	the largest number of	104	25	9
34	of teaching and learning	38	42	20	90	at the faculty of	105	24	13
35	a large number of	39	41	30	91	having in mind the	106	24	14
36	the results of this	40	41	26	92	of the faculty of	107	24	11
37	as well as in	41	40	28	93	the aim of this	108	24	20
38	in the form of	42	40	29	94	the influence of the	109	24	18
39	of the quality of	43	39	15	95	there is no statistically	110	24	10
40	the analysis of the	44	39	24	96	based on the results (of)	111	23	16
41	as well as to	46	38	29	97	in such a way	112	23	15
42	on the development of	47	38	22	98	on the quality of	114	23	11
43	(statistically) significant difference between the	48	38	15	99	the extent to which	115	23	11
44	(at) the beginning of the	49	37	26	100	in other words the	116	22	13
45	to the fact that	51	36	28	101	the purpose of this	117	22	18
46	at the level of	52	35	18	102	in the educational process	118	21	18
47	in the development of	54	34	25	103	under the influence of	120	21	14
48	for the purpose of	56	33	20	104	as well as a	121	20	17
49	having in mind that	57	33	12	105	can be found in	122	20	15
50	of this research is	58	33	23	106	of education in the	123	20	12
51	the role of the	59	33	16	107	of the education system	124	20	12
52	we can conclude that	60	33	19	108	significant difference in the	126	20	11
53	in addition to the	61	32	20	109	that they do not	127	20	12
54	a part of the	64	31	23	110	the effects of the	128	20	10
55	is based on the	65	31	23	111	we can see that	129	20	10
56	as one of the	66	30	25					

Prilog 4. Strukturalna raspodela ULS za izvorne govornike engleskog jezika – spisak

Grupa	Podgrupa	ULS	
<i>Imeničke fraze</i>	Imenička fraza sa of fragmentom	the results of the	the purpose of this
		the nature of the	the role of the
		the use of the	the quality of the
	Druge imeničke fraze	the extent to which	the ways in which
<i>Predložke fraze</i>	Predložka fraza sa of fragmentom	at the end of (the)	of teaching and learning
		as a result of	as part of the
		in the context of	at the time of
		in terms of the	over the course of
		in the form of	on the development of
	Druge predložke fraze	in relation to the	in the current study
	at the same time	on the other hand	
<i>Glagolske fraze</i>	Uvodna <i>it</i> struktura + glagolska ili pridevska fraza	it is important to	it is possible that
	(Glagol/pridev +) fragment to klauze	are more likely to	to be able to
<i>Druge strukture</i>		as well as the	

Prilog 5. Strukturalna raspodela ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski – spisak

Grupa	Podgrupa	ULS	
<i>Imeničke fraze</i>	Imenička fraza sa of fragmentom	the results of the (research)	a high level of
		the implementation of the	the aim of the
		the development of the	the application of the
		one of the most (important)	the characteristics of the
		the importance of the	a small number of
		the quality of the	the assessment of the
		the total number of	the largest number of
		a large number of	(at) the end of the
		the results of this	(at) the beginning of the
		the analysis of the	the aim of this
		the role of the	the influence of the
		a part of the	the purpose of this
		the content of the	the effects of the
	the quality of teaching		
	Druge imeničke fraze	the fact that the	the results show that
		a statistically significant difference	statistically significant difference in
(statistically) significant difference between the		the extent to which	
significant differences in the		significant difference in the	
statistically significant differences in			
<i>Predložke fraze</i>	Predložka fraza sa of fragmentom	in the field of (education)	as one of the
		on the basis of (the)	in the area of
		in the process of	from the perspective of
		for the development of	as a result of
		in the context of	in the domain of
		in the case of	with the aim of
		for the implementation of	as a part of
		in the implementation of	in terms of the
		to the development of	of the concept of
		of teaching and learning	of the fact that
		in the form of	of the importance of
		of the quality of	at the faculty of
		on the development of	of the faculty of
		at the level of	on the quality of
		in the development of	under the influence of
		for the purpose of	of education in the
		of this research is	of the education system
	Druge predložke fraze	on the other hand (the)	in comparison to the
		at the same time	by the fact that
		in accordance with the	in line with the
in relation to the		on the one hand	
in the teaching process		to the conclusion that	
to the fact that		in such a way	
in addition to the	in other words the		
to be able to	in the educational process		
<i>Glagolske fraze</i>	Uvodna <i>it</i> struktura + glagolska ili pridevska fraza	it is necessary to	it is possible to
		(it) can be concluded (that the)	it can be seen

		it is important to	
	Glagol u pasivu i fragment predloške fraze	is based on the	can be found in
		are shown in table	
	Kopula <i>be</i> i imenička/pridevska fraza	is one of the	
	Fragment <i>that</i> klauze	that there is a	that there are no
		that there is no	that they do not
	Zamenica/imenička fraza + <i>be</i> (+...)	there is no statistically	
	Druge glagolske fraze	based on the results (of)	having in mind that
		having in mind the	
<i>Druge strukture</i>		as well as the	we can conclude that
		when it comes to (the)	as well as their
		as well as in	as well as a
		as well as to	we can see that

Prilog 6. Strukturalna raspodela ULS za neizvorne govornike engleskog jezika kojima je maternji jezik rumunski – spisak

Grupa	Podgrupa	ULS	
<i>Imeničke fraze</i>	Imenička fraza sa of fragmentom	the results of the	the importance of the
		one of the most (important)	the purpose of the
		the quality of the	the number of students
		the development of the	the role of the
		(at) the beginning of the	the existence of a
		the analysis of the	a high level of
		the content of the	the majority of the
		the total number of	the purpose of this
	Druge imeničke fraze	the level of the	the development of a
		the fact that the	the point of view (of)
the way in which		the extent to which	
<i>Predložke fraze</i>	Predložka fraza sa of fragmentom	an important role in	the same time the
		in the case of (the)	for the development of
		in the field of	of the educational process
		at the end of (the)	in terms of the
		in the context of (the)	with the help of
		at the level of	by the ministry of
		to the development of	in the development of
		in the process of	than half of the
	Druge predložke fraze	from the perspective of	of the fact that
		as a result of	in the number of
		(and) on the other hand (the)	to the fact that
		at the same time	in accordance with the
		on the one hand	in the educational process
		in relation to the	by the fact that
<i>Glagolske fraze</i>	Uvodna <i>it</i> struktura + glagolska ili pridevska fraza	it is important to	it is necessary to
	Kopula <i>be</i> i imenička/pridevska fraza	is one of the	
	Fragment <i>that</i> klauze	that there is a	
	Druge glagolske fraze	taking into account the	
<i>Druge strukture</i>		as well as the	more than half of
		when it comes to	

Prilog 7. Funkcionalna raspodela ULS za izvorne govornike engleskog jezika – spisak

Grupa	Podgrupa	ULS	
<i>ULS orijentisani ka istraživanju</i>	ULS za lokaciju	at the end of (the)	at the time of
	ULS za proceduru	the purpose of this	on the development of
		the role of the	
	ULS za kvantifikaciju	the extent to which	as part of the
	ULS za opis	in the form of	the ways in which
		the nature of the	to be able to
the use of the		the quality of the	
<i>ULS orijentisani ka tekstu</i>	Tranzicioni signali	as well as the	on the other hand
		at the same time	
	Rezultativni signali	as a result of	the results of the
	Strukturni signali	in the current study	
	Ograničavajući signali	in relation to the	of teaching and learning
		in the context of	over the course of
in terms of the			
<i>ULS orijentisani ka učesniku</i>	ULS stava	it is important to	it is possible that
		are more likely to	

Prilog 8. Funkcionalna raspodela ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski – spisak

Grupa	Podgrupa	ULS	
<i>ULS orijentisani ka istraživanju</i>	ULS za lokaciju	(at) the end of the	in the area of
		(at) the beginning of the	at the faculty of
	ULS za proceduru	in the process of	the application of the
		in the teaching process	the purpose of this
		the analysis of the	in the educational process
		the role of the	the assessment of the
		under the influence of	for the purpose of
	ULS za kvantifikaciju	a statistically significant difference	that there are no
		is one of the	a small number of
		one of the most (important)	statistically significant difference in
		the total number of	the largest number of
		a large number of	there is no statistically
		(statistically) significant difference between the	the extent to which
		as one of the	significant difference in the
		a high level of	that there is no
	ULS za opis	statistically significant differences in	
		the fact that the	of the fact that
		for the implementation of	of the importance of
		the implementation of the	of the faculty of
		the importance of the	the aim of this
		of the quality of	the influence of the
		to the fact that	in such a way
		the content of the	on the quality of
		by the fact that	of education in the
		the quality of teaching	to be able to
		with the aim of	the quality of the
	the aim of the	the effects of the	
the characteristics of the	to the conclusion that		
of the concept of			
<i>ULS orijentisani ka tekstu</i>	Tranzicioni signali	as well as the	in addition to the
		when it comes to (the)	in comparison to the
		on the other hand (the)	on the one hand
		at the same time	having in mind the
		as well as in	in other words the
		as well as to	as well as their
	having in mind that	as well as a	
	Rezultativni signali	that there is a	the results show that
		the results of this	based on the results (of)
		are shown in table	that they do not
	Ograničavajući signali	as a result of	the results of the (research)
		in the field of (education)	in the domain of
		on the basis of (the)	significant differences in the
		for the development of	as a part of
		in the context of	in line with the
in the case of	in terms of the		

		in accordance with the	of the education system
		in relation to the	the development of the
		in the implementation of	to the development of
		in the form of	on the development of
		at the level of	in the development of
		of this research is	of teaching and learning
		is based on the	a part of the
		from the perspective of	
<i>ULS orijentisani ka učesniku</i>	ULS stava	it is necessary to	it is possible to
		it is important to	
	ULS uključivanja	(it) can be concluded (that the)	can be found in
		we can conclude that	we can see that
		it can be seen	

Prilog 9. Funkcionalna raspodela ULS za izvorne govornike rumunskog jezika – spisak

Grupa	Podgrupa	ULS	
<i>ULS orijentisani ka istraživanju</i>	ULS za lokaciju	at the end of (the)	the same time the
		(at) the beginning of the	
	ULS za proceduru	in the process of	taking into account the
		the analysis of the	the purpose of the
		in the educational process	the role of the
		of the educational process	in order to achieve
		an important role in	the purpose of this
	ULS za kvantifikaciju	one of the most (important)	the number of students
		the total number of	the extent to which
		is one of the	a high level of
		more than half of	in the number of
	ULS za opis	than half of the	the majority of the
		the fact that the	the level of the
		the quality of the	by the fact that
		the way in which	the importance of the
		the content of the	by the ministry of
		to be able to	the existence of a
<i>ULS orijentisani ka tekstu</i>	Tranzicioni signali	to the fact that	of the fact that
		the point of view (of)	
		(and) on the other hand (the)	on the one hand
	Rezultativni signali	at the same time	when it comes to
		as well as the	
	Ograničavajući signali	the results of the	that there is a
		as a result of	
		in the case of (the)	for the development of
		in the field of	in relation to the
		in the context of (the)	in accordance with the
at the level of		in terms of the	
ULS stava	to the development of	with the help of	
	from the perspective of	in the development of	
	the development of the	the development of a	
<i>ULS orijentisani ka učesniku</i>	ULS stava	it is important to	it is necessary to

Prilog 10. Spisak specifičnih ULS za izvorne govornike engleskog jezika

	ULS	Rang	Frekvencija	Disperzija
1	are more likely to	5	34	19
2	in the current study	10	30	10
3	the nature of the	18	24	19
4	the use of the	19	24	10
5	the ways in which	20	24	14
6	as part of the	21	22	16
7	at the time of	22	22	16
8	it is possible that	25	21	11
9	over the course of	26	21	10

Prilog 11. Spisak specifičnih ULS za neizvorne govornike engleskog kojima je maternji jezik srpski

	ULS	Rang	Frekvencija	Disperzija
1	on the basis of (the)	10	82	43
2	(it) can be concluded (that the)	12	66	27
3	for the implementation of	21	55	12
4	the implementation of the	23	54	20
5	in the implementation of	24	52	17
6	it is possible to	26	51	29
7	a statistically significant difference	27	49	15
8	in the teaching process	30	48	10
9	that there is no	32	46	20
10	a large number of	39	41	30
11	the results of this	40	41	26
12	as well as in	41	40	28
13	of the quality of	43	39	15
14	as well as to	46	38	29
15	(statistically) significant difference between the	48	38	15
16	for the purpose of	56	33	20
17	having in mind that	57	33	12
18	of this research is	58	33	23
19	we can conclude that	60	33	19
20	in addition to the	61	32	20
21	a part of the	64	31	23
22	is based on the	65	31	23
23	as one of the	66	30	25
24	in the area of	67	30	13
25	in comparison to the	73	28	13
26	are shown in table	76	27	15
27	in the domain of	80	27	14
28	significant differences in the	81	27	17
29	the quality of teaching	82	27	12
30	with the aim of	83	27	20
31	as a part of	85	26	12
32	as well as their	86	26	23
33	in line with the	87	26	16

	ULS	Rang	Frekvencija	Disperzija
34	it can be seen	89	26	14
35	statistically significant differences in	91	26	13
36	that there are no	92	26	10
37	the aim of the	93	26	18
38	the application of the	94	26	14
39	the characteristics of the	95	26	16
40	the results show that	96	26	21
41	to the conclusion that	97	26	21
42	a small number of	98	25	18
43	of the concept of	99	25	11
44	of the importance of	101	25	17
45	statistically significant difference in	102	25	11
46	the assessment of the	103	25	14
47	the largest number of	104	25	9
48	at the faculty of	105	24	13
49	having in mind the	106	24	14
50	of the faculty of	107	24	11
51	the aim of this	108	24	20
52	the influence of the	109	24	18
53	there is no statistically	110	24	10
54	based on the results (of)	111	23	16
55	in such a way	112	23	15
56	on the quality of	114	23	11
57	in other words the	116	22	13
58	under the influence of	120	21	14
59	as well as a	121	20	17
60	can be found in	122	20	15
61	of education in the	123	20	12
62	of the education system	124	20	12
63	significant difference in the	126	20	11
64	that they do not	127	20	12
65	the effects of the	128	20	10
66	we can see that	129	20	10

Prilog 12. *Spisak specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik rumunski*

	ULS	Rang	Frekvencija	Disperzija
1	the way in which	19	41	25
2	of the educational process	36	31	18
3	an important role in	40	29	23
4	taking into account the	41	29	25
5	the point of view (of)	44	28	16
6	the level of the	47	27	23
7	With the help of	49	27	23
8	more than half of	52	25	16
9	the purpose of the	55	25	20
10	by the ministry of	56	24	16
11	than half of the	58	24	17
12	the number of students	59	24	15
13	the existence of a	61	23	17
14	in order to achieve	66	20	17
15	in the number of	67	20	15
16	the majority of the	69	20	15
17	the same time the	71	20	17
18	the development of a	68	20	15

Prilog 13. *Spisak specifičnih ULS za izvorne govornike turskog jezika, preuzeto iz Gungor (2016)*

the findings of the	within the framework of	study was carried out
the reliability of the	as a part of	they were asked to
the aim of this	in the present study	the data were collected
the effect of the	in accordance with the	that they did not
a high level of	at the same time	that there is no
analysis of the data	in the current study	the results showed that
sample of the study	in line with the	that there was no
the main purpose of	with respect to the	it is possible to
one of the important	to the fact that	it was determined that
the characteristics of the	was carried out with	it was observed that
the participants of the	are presented in table	it is thought that
a significant difference between	were informed about the	it is seen that
no significant difference between	it was concluded that	it was found that
a significant difference in	it is believed that	it can be said
the difference between the	it can be concluded that	it was seen that
the data obtained from	to be used in	it can be seen
an important role in	was used to determine	it is observed that
according to the results	used in order to	it is necessary to
according to the findings	was found to be	was used in the
within the scope of	in order to determine	were included in the
as one of the	in order to be	can be defined as
in the light of	in order to test	included in the study
with the help of	there was no significant	was used as a
with the results of	there was a significant	is based on the
are given in table	as can be seen	before and after the
used in this study	as seen in table	who participated in the
participated in the study	higher than those of	
this study aims to	which is one of	

Prilog 14. Strukturalna raspodela specifičnih ULS za izvorne govornike engleskog jezika – spisak

Grupa	Podgrupa	ULS
<i>Imeničke fraze</i>	Imenička fraza sa of fragmentom	the nature of the the use of the
	Druge imeničke fraze	the ways in which
<i>Predloške fraze</i>	Predloška fraza sa of fragmentom	as part of the
		at the time of
		over the course of
	Druge predloške fraze	in the current study
<i>Glagolske fraze</i>	Uvodna <i>it</i> struktura + glagolska ili pridevska fraza	it is possible that
	(Glagol/pridev +) fragment to klauze	are more likely to

Prilog 15. Strukturalna raspodela specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezika srpski – spisak

Grupa	Podgrupa	ULS		
<i>Imeničke fraze</i>	Imenička fraza sa <i>of</i> fragmentom	the implementation of the	the characteristics of the	
		a large number of	a small number of	
		the results of this	the assessment of the	
		a part of the	the largest number of	
		the quality of teaching	the aim of this	
		the aim of the	the influence of the	
	Druge imeničke fraze	the application of the	the effects of the	
		a statistically significant difference	the results show that	
		(statistically) significant difference between the	statistically significant difference in	
		significant differences in the	significant difference in the	
<i>Predložke fraze</i>	Predložka fraza sa <i>of</i> fragmentom	statistically significant differences in		
		on the basis of (the)	in the area of	
		for the implementation of	in the domain of	
		in the implementation of	with the aim of	
		of teaching and learning	as a part of	
		in the form of	of the concept of	
		of the quality of	of the importance of	
		on the development of	at the faculty of	
	Druge predložke fraze	for the purpose of	of the faculty of	
		of this research is	on the quality of	
		as one of the		
		in the teaching process	to the conclusion that	
		in addition to the	in such a way	
		in comparison to the	in other words the	
<i>Glagolske fraze</i>	ULS sa uvodnom <i>it</i> konstrukcijom	in line with the		
		(it) can be concluded (that the)	it can be seen	
	Glagol u pasivu i fragment predložke fraze	it is possible to		
		is based on the	can be found in	
	(Glagolska fraza +) fragment <i>that</i> klauze	are shown in table		
		that there is no	that they do not	
	Zamenica/imenička fraza+be (+...)	that there are no		
there is no statistically				
Druge glagolske fraze	based on the results (of)	having in mind that		
	having in mind the			
<i>Druge strukture</i>		as well as in	as well as their	
		as well as to	as well as a	
		we can conclude that	we can see that	

Prilog 16. Strukturalna raspodela specifičnih ULS za izvorne govornike rumunskog jezika – spisak

Grupa	Podgrupa	ULS
<i>Imeničke fraze</i>	Imenička fraza sa of fragmentom	the level of the
		the purpose of the
		the number of students
		the existence of a
		the majority of the
	Druge imeničke fraze	the development of a
		the way in which
		an important role in
<i>Predložke fraze</i>	Predložka fraza sa of fragmentom	the point of view (of)
		the same time the
		of the educational process
		with the help of
	Druge predložke fraze	by the ministry of
<i>Glagolske fraze</i>	Druge glagolske fraze	than half of the
		in the number of
<i>Druge strukture</i>		in order to achieve
		taking into account the
		more than half of

Prilog 17. Funkcionalna raspodela specifičnih ULS za izvorne govornike engleskog jezika – spisak

Grupa	Podgrupa	ULS
<i>ULS orijentisani ka istraživanju</i>	ULS za lokaciju	at the time of
	ULS za kvantifikaciju	as part of the
	ULS za opis	the nature of the
		the use of the
<i>ULS orijentisani ka tekstu</i>	Strukturni signali	the ways in which
	Ograničavajući signali	in the current study
<i>ULS orijentisani ka učesniku</i>	ULS stava	over the course of
		are more likely to
		it is possible that

Prilog 18. Funkcionalna raspodela specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik srpski – spisak

Grupa	Podgrupa	ULS		
<i>ULS orijentisani ka istraživanju</i>	ULS za lokaciju	in the area of	at the faculty of	
		in the teaching process	the assessment of the	
	ULS za proceduru	under the influence of	for the purpose of	
		the application of the		
		a statistically significant difference	a small number of statistically significant difference in	
	ULS za kvantifikaciju	a large number of	(statistically) significant difference between the	the largest number of
		as one of the	statistically significant differences in	there is no statistically
		that there are no	that there are no	significant difference in the
		for the implementation of	of the faculty of	that there is no
		the implementation of the	the aim of this	
	ULS za opis	of the quality of	the quality of teaching	the influence of the
		the quality of teaching	with the aim of	in such a way
		with the aim of	the aim of the	on the quality of
		the aim of the	the characteristics of the	of education in the
		of the concept of	of the concept of	the effects of the
of the importance of		of the importance of	to the conclusion that	
<i>ULS orijentisani ka tekstu</i>	Tranzicioni signali	as well as in	having in mind the	
		as well as to	in other words the	
		having in mind that	as well as their	
		in addition to the	as well as a	
	Rezultativni signali	in comparison to the		
		the results of this	based on the results (of)	
		are shown in table	that they do not	
	Ograničavajući signali	the results show that		
		on the basis of (the)	significant differences in the	
		in the implementation of	as a part of	
		of this research is	in line with the	
	ULS stava	is based on the	of the education system	
in the domain of		a part of the		
<i>ULS orijentisani ka učesniku</i>	ULS stava	it is possible to		
	ULS uključivanja	(it) can be concluded (that the)	can be found in	
		we can conclude that	we can see that	
		it can be seen		

Prilog 19. Funkcionalna raspodela specifičnih ULS za neizvorne govornike engleskog jezika kojima je maternji jezik rumunski – spisak

Grupa	Podgrupa	ULS
<i>ULS orijentisani ka istraživanju</i>	ULS za lokaciju	the same time the
		of the educational process
	ULS za proceduru	an important role in
		taking into account the
		the purpose of the
		in order to achieve
	ULS za kvantifikaciju	more than half of
		than half of the
		the number of students
		in the number of
	ULS za opis	the majority of the
		the way in which
		the point of view (of)
the level of the		
Ograničavajući signali	by the ministry of	
	the existence of a	
<i>ULS orijentisani ka tekstu</i>	Ograničavajući signali	with the help of
		the development of a

Biografija autorke

Radmila Palinkašević je rođena 1990. godine u Vršcu. U rodnom gradu završila je osnovnu školu i Gimnaziju „Borislav Petrov Braca”. Osnovne studije završila je 2012. godine na Filološkom fakultetu Univerziteta u Beogradu, odsek za engleski jezik i književnost. Na istom fakultetu završila je master studije, a tema master rada bila je: *The Influence on Bilingualism on Third Language Acquisition: Reading Comprehension and Article Usage*. Trenutno pohađa doktorske studije na Filološkom fakultetu Univerziteta u Beogradu, na modulu *Jezik*. Od oktobra 2011. do januara 2012. godine volontirala je u Osnovnoj školi „Banović Strahinja” u Beogradu i predavala deci u trećem i četvrtom razredu. Od 2012. godine predavala je u privatnoj školi stranih jezika „The English House” u Beogradu. Tu je stekla iskustvo u radu sa svim starosnim grupama i držala opšte kurseve, komunikacione kurseve, kao i individualnu nastavu. Zvanje saradnika u nastavi dobila je 2014. godine kada je stupila u radni odnos na Visokoj školi strukovnih studija za vaspitače „Mihailo Palov” u Vršcu, a marta 2016. izabrana je u zvanje asistenta na istoj instituciji. Od 2022. radi kao nastavnik stranog jezika na Visokoj strukovnoj vaspitačkoj i medicinskoj školi u Vršcu. Od 2014. godine član je uređivačkog odbora naučnog časopisa *Istraživanja u pedagogiji / Research in Pedagogy*; između ostalog, zadužena je za indeksaciju časopisa u baze podataka i komunikaciju. Od 2022. godine Radmila Palinkašević je koordinator *Erazmus + mobilnosti* Visoke škole u Vršcu.

Izjava o autorstvu (Prilog 1.)

Прилог 1.

Изјава о ауторству

Име и презиме аутора Радмила Паличишевић

Број досијеа 14053/Д

Изјављујем

да је докторска дисертација под насловом

Разлике у употреби учесталих лексичких спојева код изворних и неизворних говорника у научним радовима из области обрезовања

- резултат сопственог истраживачког рада;
- да дисертација ни у целини ни у деловима није била предложена за стицање дипломе студијских програма других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

Потпис аутора

У Београду, 2024. године

Радмила Паличишевић

Izjava o istovetnosti štampane i elektronske verzije doktorskog rada (Prilog 2.)

Прилог 2.

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора Радмила Палинкашевић

Број досијеа 14053/14

Студијски програм Језик, књижевност, култура

Наслов рада Разлике у употреби учесталих лексичких спојева код изворних и неизворних говорника у научним радовима из области образовања

Ментор доц. др. Мирјана Даничић

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањивања у **Дигитални репозиторијум Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци за добијање академског назива 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 аутора

У Београду, 2024. године

Радмила Палинкашевић

Izjava o korišćenju (Prilog 3.)

Прилог 3.

Изјава о коришћењу

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

Разлике у употреби учесталих лексичких спојева код изворних и нечворних
говорника у научним радовима из области образовања
која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду, и доступну у отвореном приступу, могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла:

1. Ауторство (CC BY)
2. Ауторство – некомерцијално (CC BY-NC)
3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)
4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)
5. Ауторство – без прерада (CC BY-ND)
6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, 2024. године

Рагнела Паличишевић