

Besplatni softverski alati za analizu tekstualnih korpusa i neke mogućnosti njihove primjene u kontrastivnoj lingvističkoj analizi i nastavi engleskog jezika struke

Darko Kovačević
Elektrotehnički fakultet
Univerzitet u Istočnom Sarajevu
Istočno Sarajevo, Bosna i Hercegovina
dax1978@gmail.com

Sažetak—Prilikom analiziranja digitalizovanog korpusa tekstova na nekom jeziku ili sprovođenja kontrastivne analize srodnih korpusa na dva jezika, softverski alati mogu biti od velike pomoći u raznim aspektima analize i dati veliki doprinos u smislu njene tačnosti i preciznosti, ali i uštede vremena potrebnog da se analiza izvrši. Alati *AntConc*, *AntPConc*, *AntWordProfiler* i *AntMover* predstavljaju besplatne aplikacije koje se mogu preuzeti sa interneta i pomoću kojih se mogu realizovati različiti aspekti analize digitalizovanog korpusa tekstova. U radu će, najprije, ukratko biti predstavljeni pojam i karakteristike digitalizovanog korpusa tekstova. Potom će biti prikazane četiri navedene aplikacije, kroz opisivanje njihovih najznačajnijih karakteristika. Nakon toga će uslijediti navođenje mogućnosti primjene navedenih aplikacija, najprije na nivou kontrastivne analize uporedivog korpusa tekstova na dva različita jezika, odnosno engleskom i srpskom, a potom u sklopu nastave engleskog jezika struke, te zaključci izvedeni na osnovu svega rečenog.

Ključne riječi- softver, analiza korpusa, *AntConc*, *AntPConc*, *AntWordProfiler*, *AntMover*, engleski jezik struke, kontrastivna analiza

I. UVODNA RAZMATRANJA

U savremenoj lingvističkoj terminologiji, pojam teksta podrazumijeva primjere jezika u upotrebi, bilo govorne ili pisane, odnosno jezičko ponašanje do kog je došlo spontano, bez intervencije nekog lingviste. Primjeri za jezik u upotrebi uključuju: konverzaciju, predavanje, oglas, novinski članak, naučni rad, roman, udžbenik i slično. [1] S druge strane, postoje dva načelna pristupa analizi teksta: lingvistički, koji tretira tekst kao objekat sam po sebi, i sociološki, koji tretira tekst kao "prozor u ljudsko iskustvo" [2]. Analize na osnovu navedenih pristupa se sprovode na različite načine, u skladu sa objektivnim potrebama istraživanja. U današnje vrijeme, razvoj tehnološki razvoj je u znatnoj mjeri olakšao sam proces analize teksta. To se, prvenstveno, odnosi na mogućnost kreiranja digitalnih korpusa tekstova, a, u skladu s tim, i njihovog organizovanja, pretraživanja i analiziranja, čime se sam analitički proces u znatnoj mjeri pojednostavljuje, a samo vrijeme potrebno za vršenje analize skraćuje, uz istovremeno povećanje preciznosti i tačnosti dobijenih rezultata.

Sam korpus se definiše kao zbirka tekstova, u elektronskom obliku, koji su, na osnovu nekog spoljašnjeg kriterijuma, odabrani da predstavljaju, koliko je moguće, jezik ili više jezika u svojstvu izvora podataka za lingvističko istraživanje [3]. U zavisnosti od prirode istraživanja, korpusi se mogu razlikovati po obimu, sadržaju i samim tekstualnim izvorima na osnovu kojih su sastavljeni. Takođe, mogu biti jednojezični i višejezični, pri čemu se višejezični dalje mogu podijeliti na paralelne (sačinjene od originalnih tekstova i njihovih prevoda na jedan ili više jezika, u jednom, dva ili više smjerova) i uporedive (sačinjene od tekstova iz dva ili više jezika srodnih po temi, sadržaju, obimu i drugim sličnim karakteristikama).

Što se tiče engleskog jezika, postoje brojni sveobuhvatni korpusi koji su slobodno dostupni na internetu, a linkove za pristup moguće je naći na stranici <http://corpus.byu.edu/>. Što se tiče korpusa srpskog/srpskohrvatskog jezika, lista onih koji su dostupni može se pronaći na internet-adresi <http://korpus.matf.bg.ac.rs/prezentacija/korpusi.html>. Takođe sačinjen je i paralelni korpus srpskog i engleskog jezika, pod nazivom *English/Serbian Aligned Corpus (SrpEngKor)*. Ovaj korpus sadrži 4420711 riječi (2330742 na engleskom i 2089969 na srpskom jeziku) i čine ga tekstovi iz oblasti proze, vijesti, naučnih časopisa, internet-novinarstva, zdravstva, prava i obrazovanja, te titlovi filmova, pri čemu su izvorni tekstovi na engleskom jeziku dati paralelno sa svojim prevodima na srpski i obrnuto. Korpusu je moguće pristupiti putem linka <http://korpus.matf.bg.ac.rs/SrpEngKor/korpus/index1.php>, dok se više informacija o njegovoj strukturi može pronaći na adresi http://korpus.matf.bg.ac.rs/SrpEngKor/SrpEngKor_2013_01.pdf.

Međutim, ovakvi korpusi dostupni i pretraživi putem interneta često ne predstavljaju adekvatan i mjerodavan izvor tekstualnog materijala pogodnog za analizu, u odnosu na predmet istraživanja, te ih je u tom slučaju neophodno kreirati, što podrazumijeva prekućavanje teksta iz štampanih publikacija ili njegovu digitalizovanje pomoću nekog od programa za OCR (Optical Character Recognition), ili kopiranje, u slučaju da se koriste digitalne publikacije, nakon čega se tekstovi objedinjuju u okviru jednog dokumenta (u formatu txt ili doc) i, kao takvi, postaju pogodni za analiziranje.

Samo analiziranje, naročito kada su u pitanju korpusi tekstova u vezi sa specifičnim oblastima ljudskog djelovanja u različitim područjima (jezik struke ili jezik posebnih namjena) može podrazumijevati i uključivati različite aspekte, a neki od njih su:

- lista riječi prisutnih u tekstu, sa posebnim naglaskom na terminima bitnim i karakterističnim za određenu oblast, njihovu učestalost i tekstualne situacija u kojima se javljaju;
- lista ključnih riječi;
- karakteristične odlike upotrebu glagolskih vremena, oblika i stanja;
- načini modifikovanja imenica;
- sredstva kojima se postižu kohezija i koherentnost unutar tekstova i
- strukturalna organizaciju tekstova.

II. ANALIZA TEKSTUALNIH KORPUSA UZ POMOĆ RAČUNARA. BESPLATNE APLIKACIJE.

Prilikom realizovanja analize, kao što je već rečeno, pomoć u vidu računarskog pretraživanja, obrade i predstavljanja podataka može biti od ogromnog značaja. U tom smislu, pomoć mogu pružiti neke od opcija koje nude Microsoft Word i slični programi za obradu teksta, kao što su *Find*, *Replace* i *Word Count*, kao i različiti *Macro*-i (instrukcije kojima se izvršavaju automatizovani nizovi radnji u okviru programa) koje je moguće kreirati u ovakvim programima ili ih preuzeti s različitih lokacija na internetu.

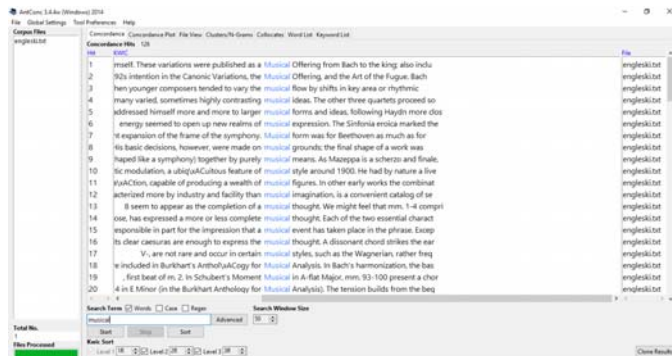
Međutim, u određenim analitičkim situacijama, navedene opcije nisu dovoljne za uspješno izvršavanje same analize, te je neophodno potražiti softversku pomoć u vidu alata specijalizovanih za analizu teksta. U tom smislu, odličnu besplatnu varijantu predstavlja skup aplikacija dostupan za preuzimanje sa sajta <http://www.laurenceanthony.net/>. U pitanju su aplikacije *AntConc*, *AntPConc*, *AntWordProfiler* i *AntMover*, koje je kreirao dr Laurence Anthony, stručnjak za korpusnu lingvistiku i obrazovne tehnologije sa Univerziteta Waseda u Japanu. Tri od četiri aplikacije imaju mogućnost rada na različitim operativnim sistemima (Windows, Linux, Macintosh), i njihovo korišćenje ne zahtjeva prethodnu instalaciju, već se mogu pokrenuti odmah po preuzimanju. Jedino aplikacija *AntMover* radi samo na Windows platformi i neophodno je da se instalira kako bi se mogla koristiti.

U daljem tekstu, svaka od navedenih aplikacija biće ukratko predstavljena, na osnovu dokumentacije dostupne na stranicama sa kojih se aplikacije mogu preuzimati.

A. AntConc

Aplikacija *AntConc* omogućava sprovođenje različitih vidova istraživanja iz oblasti korpusne lingvistike i učenja na osnovu podataka. Sadrži sedam alati kojima se može pristupiti iz glavnog prozora aplikacije ili pomoću prečica.

Alatka **Concordance** omogućava da se, kroz prikazivanje rezultata pretrage u KWIC formatu (Key Word In Context – ključna riječ u kontekstu) vidi na koji način se riječi i fraze koriste u korpusu tekstova.



Slika 1. Izgled aplikacije AntConc

Alatka **Concordance Plot** prikazuje rezultate pretrage u „barkod” formatu, što omogućava da se vide mjesta na kojima se rezultati pretrage pojavljuju u ciljnim tekstovima.

Alatka **File View** prikazuje tekst pojedinačnih datoteka, čime se omogućava detaljnije ispitivanje rezultata generisanih pomoću drugih alati u sklopu aplikacije *AntConc*.

Pri dio alatke **Clusters/N-Grams (Clusters)** prikazuje karakteristične grupe riječi na osnovu uslova postavljenih pri pretraživanju, odnosno sažima rezultate generisane pomoću alatke **Concordance** ili **Concordance Plot**. S druge strane, drugi dio ove alatke (**N-Grams**) skenira korpus u potrazi za karakterističnim grupama sastavljenim od određenog broja riječi, te se pomoću nje mogu pronaći uobičajeni izrazi u okviru korpusa.

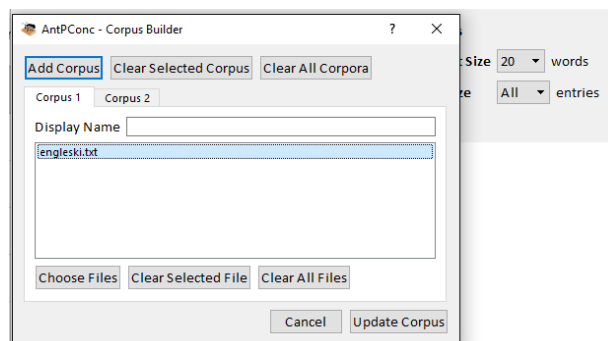
Alatka **Collocates** prikazuje kolokacije riječi koja je predmet pretraživanja, odnosno karakteristične vidove njenog pojavljivanja i kombinovanja s drugim riječima u sklopu korpusa.

Alatka **Word List** prebrojava sve riječi u okviru korpusa i predstavlja ih u formi uređene liste, na osnovu čega se može brzo uočiti koje riječi se najčešće ponavljaju u korpusu.

Konačno, alatka **Keyword List** prikazuje koje riječi su neobično česte ili neobično odsutne u korpusu u poređenju sa riječima u nekom drugom referentnom korpusu [4].

B. AntPConc

Aplikacija *AntPConc* služi za analizu paralelnih korpusa, uz mogućnost prethodnog kreiranja paralelnih korpusa iz dva ili više usklađena tekstualna fajla (u txt formatu) koja sadrže tekstove na različitim jezicima, kao i njihovog uređivanja i pretrage [5].

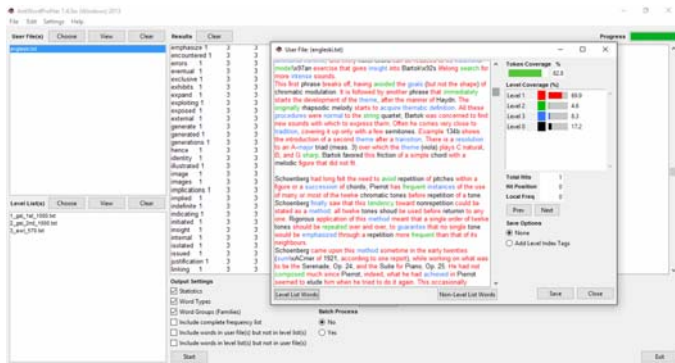


Slika 2. Izgled aplikacije AntPConc

C. AntWordProfiler

AntWordProfiler je aplikacija za izvođenje istraživanja iz korpusne lingvistike zasnovanih na profilisanju vokabulara. Sadrži dvije alatke. Osnovna alatka, koja se pojavljuje u glavnom prozoru programa, jeste **Vocabulary Profile** i pomoću nje je moguće generisati informacije u vezi sa statistikom i učestalošću vokabulara u korpusu tekstova koji je učitani u program i uporediti rezultate sa različitim listama vokabulara sa nivoima (*level lists*), pri čemu se tri takve liste već nalaze u okviru aplikacije, a takođe je moguće napraviti, učitati i koristiti nove liste, pripremljene u txt formatu. Inače, liste sa nivoima su vertikalne liste koje sadrže osnovni oblik i moguće familije riječi koje se žele ispitati.

Druga alatka u okviru ove aplikacije je **File Viewer and Editor** i ona omogućava uvid u pojedinačne korisničke datoteke i obilježava različite nivoe vokabulara pomoću različitih boja. U njoj se, takođe, prikazuje i opšta pokrivenost različitih nivoa vokabulara. Ova alatka, takođe, omogućava uređivanje datoteke, na takav način da su efekti uređivanja trenutno vidljivi na nivou vokabulara riječi i teksta kao cjeline [6].



Slika 3. Izgled aplikacije *AntWordProfiler*

D. AntMover

AntMover predstavlja alatku za automatsko analiziranje i određivanje strukture teksta u smislu njegovih dijelova, odnosno *poteza* (*Moves*), što je izraz koji je u lingvističku literaturu uveo John Swales u svojoj studiji o analizi žanra [7]. Nakon obavljene *analize poteza* (*moves analysis*), svakoj rečenici u okviru teksta se dodjeljuje retorička (strukturna) etiketa, a te etikete, zajedno, mogu koristiti u formiranju ideje o načelnoj retoričkoj ili strukturalnoj organizaciji analiziranog teksta.

III. PRIMJENA APLIKACIJA U KONTRASTIVNOJ ANALIZI UPOREDIVOG KORPUSA TEKSTOVA NA ENGLISKOM I SRPSKOM JEZIKU

Kontrastivna analiza u lingvistici se obično definiše kao sistematsko proučavanje para jezika kojim se utvrđuje u kojim su aspektima slični, a u kojim se razlikuju [8], što, prethodno, podrazumijeva formiranje uporedivih korpusa na dva ciljna jezika i utvrđivanje kriterijuma na osnovu kog se njihovo vrši poređenje. U konkretnom slučaju ovog rada, formiran je korpus od po 200 tekstova o klasičnoj muzici, odnosno odlomaka iz dužih tekstualnih cjelina napisanih na engleskom i srpskom

jeziku, koji je organizovan i uređen u dvije tekstualne (txt) datoteke, u odnosu na jezike, te su četiri aplikacije predstavljene u prethodnom poglavlju korišćene kao pomoć pri realizaciji različitih analitičkih zahtjeva koji su prethodili samom kontrastiranju tekstova u različitim parametrima. Pri tome je bilo neminovno da se pojavu određena ograničenja uslovljena samom građom, kompleksnošću i gramatičkim pravilima srpskog jezika.

Aplikacija *AntConc* korišćena je, najprije, u odvojenoj analizi tekstova na jednom i drugom jeziku, kako bi se kreirale relevantne liste ključnih riječi (termina) u tekstovima na oba jezika. Potom je uslijedilo utvrđivanje karakterističnih tekstualnih situacija u kojima se ti termini javljaju, što uključuje kolokacije i karakteristične grupe riječi. Navedeni proces se odvijao znatno lakše u sklopu analiziranja tekstova na engleskom jeziku, zahvaljujući odsustvu promjene konstrukcije pojedinačnih riječi u skladu sa rodom i padežom, dok je prisustvo takve promjene u srpskom jeziku u priličnoj mjeri otežavalo analitički proces, što je djelimično riješeno tako što su se, u procesu pretrage i upućivanja zahtjeva aplikaciji, koristili samo korijeni riječi, odnosno njihovi dijelovi koji se ne mijenjaju prema rodu i padežu, ili je, naročito u slučaju karakterističnih grupa riječi ili kolokacija, rađeno više analitičkih pretraga za različite oblike riječi, čiji su rezultati potom upoređeni i objedinjeni. Podaci dobijeni ovakvom analizom kasnije su korišćeni kao osnova za izvođenje kontrastivne analize u odnosu na analizirane parametre, obavljene direktnim uvidom u tekstove iz oba korpusa.

Budući da korpus koji je predmet analize nije paralelan, već uporediv, alatka *AntPConc* u njemu nije mogla naći svoju namjenu. To, međutim, ne veži za aplikaciju *AntWordProfiler* koja je korišćena za određivanje profila vokabulara u okviru tekstova iz korpusa, odvojeno za one napisane na srpskom jeziku i one koji su na engleskom. Pri tome su, već pomenuti, problemi u vezi sa promjenjivim riječima u srpskom jeziku djelimično riješavani na sličan način kao i u radu sa aplikacijom *AntConc*. Na ovaj način su dobijeni detaljni kvantitativni i statistički podaci koji se dalje mogu koristiti u kontrastivnoj analizi.

Konačno, aplikacija *AntMover* omogućila je da se bar djelimično, računarskim putem, ostvari uvid u strukturu organizaciju tekstova uključenih u ovaj korpus, tačnije njegov dio koji čine tekstovi na engleskom jeziku, budući da je aplikacija prvenstveno usmjerena ka tom jeziku. Ipak, zahvaljujući tome što se pri izboru tekstova za korpus na oba jezika u velikoj mjeri vodilo računa o tome da tekstovi budu tematski srodni, slične dužine i formalno i smisaono zaokruženi, iz rezultata dobijenih za engleski jezik mogu se dobiti načelni nagovještaji strukturalnih odlika tekstova na srpskom jeziku, što u određenoj mjeri olakšava njihovo analiziranje i pomaže u realizaciji dijela kontrastivne analize koji uključuje strukturu tekstova kao kriterijum za poređenje

IV. PRIMJENA APLIKACIJA U NASTAVI ENGLISKOG JEZIKA STRUKE

Bez obzira na oblast ljudskog djelovanja obuhvata, na kom nivou obrazovanja se odvija i koliko ukupnog sedmičnog vremena je planirano za održavanje časova i vannastavne

aktivnosti, nastava engleskog struke mora, kao jedan od svojih najvažnijih segmenata, uključivati rad sa autentičnim tekstualnim materijalima iz oblasti interesovanja napisanim na engleskom jeziku, od strane izvornih govornika ili dobrih poznavalaca jezika. Navedeno stanovište podrazumijeva da je, za uspješnu realizaciju nastavnog procesa, poželjno i potrebno da postoji korpus adekvatnih autentičnih tekstova iz različitih izvora koji mogu poslužiti kao izvor materijala za rad na času (tokom samog nastavnog procesa) i u sklopu različitih vannastavnih aktivnosti. Pri tome je poželjno da sam nastavnik engleskog jezika struke sastavi korpus, u skladu sa potrebama i ciljevima nastavnog procesa u kome je angažovan, kao i to da korpus bude redovno ažuriran novim tekstualnim materijalima kako ne bi izgubio na aktuelnosti.

Kako bi se obezbijedilo povezivanje (poznate) stručne terminologije i žanrovskih karakteristika stručnih tekstualnih materijala iz maternjeg jezika sa onim u srodnim tekstovima na engleskom jeziku, poželjno je da, pored „čisto“ engleskog korpusa postoji još i paralelni korpus tekstova koji uključuje tekstove na maternjem jeziku i njihove prevode na engleski jezik, i obrnuto, prevode tekstova sa engleskog jezika na maternji.

Pod takvim uslovima, otvaraju se brojne mogućnosti za korišćenje opisanih aplikacija za analizu korpusa u nastavi engleskog jezika struke.

To se, u prvom redu, odnosi na razne mogućnosti koje pruža aplikacija *AntConc* i njeni alati, u smislu izdvajanja lista ključnih riječi, karakterističnih grupa riječi i kolokacija iz tekstova na engleskom jeziku. Navedeno se može uraditi u sklopu pripreme za nastavu i dostaviti studentima kao gotov materijal uz određeni tekst ili skup tekstova koji je predmet obrade na nastavi u okviru planirane nastavne jedinice, ali se, zahvaljujući dostupnosti aplikacija i lakoći korišćenja, studentima mogu zadati i različiti zadaci za rad u okviru vannastavnih aktivnosti.

U oba slučaja, dobijeni rezultati daju svoj doprinos u pogledu identifikovanja terminoloških odrednica iz engleskog jezika koje su karakteristične za oblast koja je predmet interesovanja, a time, na osnovu praktičnog uvida u konkretnu upotrebu u odabranom tekstu ili tekstovima iz korpusa, i njihovog usvajanja i učenja.

Na sličan način se studentima može pomoći u ovladavanju vještinom pisanja stručnih tekstova iz određene oblasti, tako što se od njih traži da na osnovu spiskova ključnih riječi, karakterističnih grupa riječi i kolokacija iz određenog teksta ili grupe tekstova dobijenih pomoću *AntConc*-a, sami kreiraju nove ili slične tekstualne cjeline, kombinujući navedene termine i primjenjujući prethodno stečeno znanje iz „opšteg“ engleskog jezika i njegovih gramatičkih pravila. Pri tome je, svakako, najlakše krenuti od pisanja sažetaka pojedinačnih tekstova ka dužim tekstualnim formama.

U slučaju da postoji paralelni korpus koji uključuje tekstove iz određene oblasti, obrada takvog korpusa ili izbora tekstova iz njega pomoću aplikacije *AntPConc* može biti od velike koristi da prikaže na koji način se određene terminološke odrednice karakteristične za oblast interesovanja prevode na engleski ili srpski jezik, kao i za proučavanje načina na koji se

određene gramatičke konstrukcije karakteristične za neku oblast prevodilački realizuju u jednom ili u drugom jeziku.

Aplikacija *AntWordProfiler* svoju upotrebu u nastavi može naći ukoliko su za realizaciju određene nastavne jedinice ili usvajanje određene građe na osnovu tekstova na engleskom jeziku iz korpusa neophodni precizni statistički podaci vezani za upotrebu vokabulara.

Posljednja od predstavljenih aplikacija, *AntMover*, kao što je to već opisano omogućava uvid u strukturalnu organizaciju tekstova na engleskom jeziku. Međutim, budući da je, ipak u pitanju računarski generisana podjela, na osnovu parametara koji su postavljeni u okviru aplikacije, takva da u potpunosti isključuje uticaj žive osobe, te zahvaljujući činjenici da je svaki tekst, koliko god da je žanrovski i terminološki predefinisani, ipak jedinica koja je relativno jedinstvena, rezultati dobijeni analiziranjem strukture tekstova u *AntMover*-u mogu biti prvenstveno korisni kao svojevrsna pripremna faza za stvarno, direktno analiziranje tekstova koje se ostvaruje samim uvidom u njih.

Takode, aplikacija se može koristiti i za uvježbavanje vještina studenata u smislu pisanja stručnih tekstova, na takav način što će se, u prvoj fazi, od njih tražiti da, sastavljajući vlastiti tekst na osnovu ključnih termina dobijenih pomoću *AntConc*-a simuliraju strukturu odgovarajućeg teksta iz korpusa (koji je prethodno analiziran i pomoću *AntMover*-a i direktno), te da potom sastavljene tekstove razmijene među sobom, i individualno analiziraju, u aplikaciji i direktnim čitanjem. Nakon toga, dobijene rezultate će uporediti i, na osnovu njih, izvesti principe organizacije i strukture određenog tipa (žanra) teksta i predstaviti ih u sklopu nastave, uz mogućnost da ih nastavnik po potrebi koriguje.

V. ZAKLJUČNA RAZMATRANJA

Aplikacije predstavljene u prethodnim poglavljima zasigurno mogu poslužiti kao pomoćni alat, kako pri ubrzavanju i automatizovanju određenih analitičkih procedura u sklopu sprovođenja kontrastivne analize, tako i u različitim aspektima nastave jezika struke. Jedino ograničenje u smislu analize predstavlja njihova primarna orijentisanost ka tekstovima na engleskom jeziku, odnosno, opisano drugim riječima, njihova prilagođenost jezičkom sistemu engleskog jezika, dok je za analiziranje tekstova na srpskom jeziku, zahvaljujući samoj prirodi jezika, potrebno učiniti nešto više koraka i, praktično, dobijene rezultate naposljetku „ručno“ sumirati i sažeti u jedno, te onda dalje koristiti. Što se tiče upotrebe u sklopu nastave engleskog jezika struke, ona ne pati od takvih ograničenja ali je, od strane nastavnika, ipak, često potrebno i poželjno provjeriti rezultate generisane nekom od aplikacija prije njihovog predstavljanja studentima i same upotrebe u nastavnom procesu, te sugerisati studentima da isto to učine prilikom izrade zadataka koje uključuju rad sa ovim aplikacijama u sklopu vannastavnih aktivnosti.

LITERATURA

- [1] M. Stubbs, *Text and Corpus Analysis: Computer-assisted Studies of Language and Culture*, New Jersey: Wiley-Blackwell, 1996.
- [2] H. R. Bernard, G. Ryan, "Text Analysis: Qualitative and Quantitative Methods" u H. R. Bernard (ur.), *Handbook of Research Methods in*

Cultural Anthropology Qualitative Research, Walnut Creek, CA: AltaMira Press, 1998, str. 595-646.

- [3] J. Sinclair, "Corpus and Text — Basic Principles" u M. Wynne (ur.), Developing Linguistic Corpora: a Guide to Good Practice, Oxford: Oxbow Books, 2005. Dostupno na <http://www.ahds.ac.uk/creating/guides/linguistic-corpora/preface.htm> (datum pristupa: 15. 2. 2016.)
- [4] <http://www.laurenceanthony.net/software/antconc/releases/AntConc343/help.pdf> (datum pristupa: 15. 2. 2016.)
- [5] <http://www.laurenceanthony.net/software/antpconcr/releases/AntPConc110/help.pdf> (datum pristupa: 15. 2. 2016.)
- [6] <http://www.laurenceanthony.net/software/antwordprofiler/releases/AntWordProfiler141/help.pdf> (datum pristupa: 15. 2. 2016.)
- [7] J. Swales, Genre analysis: English in academic and research settings, New York: Cambridge University Press, 2008.
- [8] S. Kurteš, "New horizons for contrastive analysis: grammatical prototypes as tertium comparationis", u A. Tsangalidis (ur.), Selected papers from the 18th International Symposium on Theoretical and Applied Linguistics., Thessaloniki: Monochromia Publishing, , 2009, str. 233-241. Dostupno na http://www.enl.auth.gr/symposium18/papers/24_Kurtes.pdf (datum pristupa: 15. 2. 2016.)

ABSTRACT

At the analysis of a digitalized text corpus in a language or the performing of a contrastive analysis of similar corpora in two languages, software tools can be of great assistance in different aspects of analysis and greatly contribute to its accuracy and precision. The applications AntConc, AntPConc, AntWordProfiler and AntMover are free applications which can be downloaded from the Internet and which can perform different aspects of a text corpus analysis. At the beginning, the paper will introduce the notion and characteristics of a digitalized text corpus. Then the four named applications will be presented, through the description of their most important features. After that, the possibilities of their use will be stated, at first on the level of a contrastive analysis of a comparable corpus of texts in two languages (English and Serbian), and then within the teaching of ESP, as well as some conclusions derived from what has been stated in the paper.

**FREEWARE TOOLS FOR TEXT CORPUS ANALYSIS
AND SOME POSSIBILITIES OF THEIR
APPLICATION IN CONTRASTIVE LINGUISTIC
ANALYSIS AND ESP TEACHING**

Darko Kovačević