

**Igor IVANOVIĆ (Podgorica)**  
Institut za strane jezike – Podgorica  
[iggybosnia@gmail.com](mailto:iggybosnia@gmail.com)

## KORPUSNA I RAČUNARSKA LINGVISTIKA KROZ CRNOGORSKI JEZIK

U ovome radu autor analizira dio crnogorske i engleske verzije Upitnika Evropske komisije kroz prizmu crnogorskoga jezika. Iz cjelokupne analize u našem doktorskom radu, odabrali smo dva aspekta. Prvi je analiza impersonalizovanog registra, a drugi predstavlja nalaženje veze između *hapax legomena* i grešaka koje smo uočili u pomenutom korpusu. Pokazaćemo kako upotreba korpusa i računara može da baci novo svijetlo na tradicionalna lingvistička istraživanja. Pored toga, ovaj rad treba shvatiti kao svojevrsnu promociju korpusne i računarske lingvistike u Crnoj Gori.

Ključne riječi: *korpus, računari, lingvistička analiza, Upitnik EK*

### 1. Analiza korpusa

#### 1.1 Uvod

Ova analiza nastala je kao proizvod rada na korpusu Upitnika EK sa kojim smo došli u kontakt tokom prevođenja istog. Zahvaljujući tome, bili smo u stanju da prikupimo veliku zbirku tekstova iz različitih oblasti kako bismo mogli da zadovoljimo kriterijum reprezentativnosti korpusa.

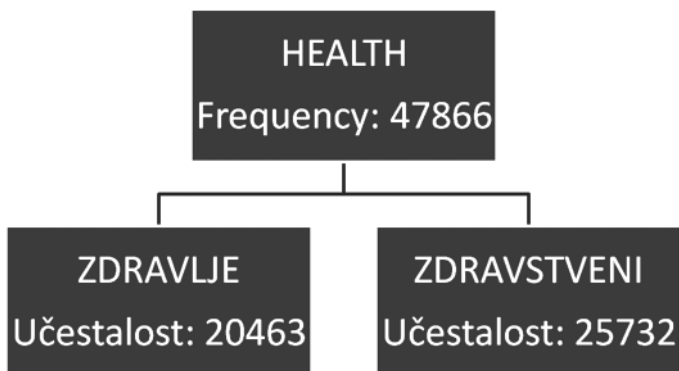


## 1.2 Rezultati istraživanja

Ono što se odmah dalo uočiti nakon inicijalnih analiza toga korpusa jeste da je velika većina tekstova sastavljena od formalnih jezičkih struktura, što smo i očekivali, s obzirom na to da korpus predstavlja zvanični odnos Crne Gore prema institucijama EU. Zbog tih elemenata, možemo konstatovati gotovo potpuno odsustvo apstraktnih imenica ili drugih vrsta riječi jer su svi dokumenti u okviru korpusa fokusirani na konkretnu problematiku. Tu konstataciju možemo dopuniti i konstatacijom da je velika većina tekstova impersonalizovana, što znači da su lične zamjenice poput *ja* ili *I* (engl.) gotovo nepostojeće. Takođe, otkrili smo da riječi za koje intuitivno smatramo da bi trebalo da su najčešće u pojedinim zakonima ili poglavljima Upitnika, to nijesu.

Tako u pogavlju Upitnika EK o Javnom zdravlju i zaštiti potrošača, u dijelu korpusa na engleskome jeziku najučestalija riječ je *health*, što se može smatrati i očekivanim s obzirom na tematiku. S druge strane, u dijelu korpusa na crnogorskome jeziku najučestalija riječ nije *zdravlje*, već riječ *proizvod/a/i*. Objašnjenje za tu razliku možemo pronaći u stručnoj terminologiji koja je karakteristična za polje zdravstva. Naime, fraza kao što je *health-care facility* u većini slučajeva je prevedena kao *zdravstvena ustanova*. S druge strane, termin kao što je *health-related issues* obično je preveden kao *problemi vezani za zdravlje*. Na ta dva primjera vidimo da su prevodni ekvivalenti za englesku riječ *health* podijeljeni na dvije grupe: *zdravlje* i *zdravstvo*. Engleski termin *product* gotovo uvijek imao je za svoj prevodni ekvivalent našu riječ *proizvod*.

Dakle, možemo zaključiti da najučestalija riječ u jednome dijelu korpusa ne dovodi zakonomjerno do toga da je i njen prevodni ekvivalent u drugom dijelu korpusa najučestaliji. Varijacije se mogu javiti usljed specifičnog kontekstualnog okruženja, tj. prevodni ekvivalent zavisi ne samo od osnovne riječi već i od konteksta u kojem se nalazi, što smo pokazali i kroz korpus.



### 1.2.1 Impersonalizovani registar

Ovakva vrsta registra, koja se obično podvodi pod administrativni stil, pokazuje veliki stepen šematizovanosti, tj. koriste se raznovrsne utvrđene fraze kako bi se u što kraćem vremenu i što jasnije prenijele ideje koje neki dokument nosi. Najčešće se Takvi administrativni tekstovi koriste kao vid komunikacije između država ili institucija. Na gramatičkome planu taj stil je nominalni, što znači da u njemu dominiraju imenice, a proces nominalizacije veoma je čest, tako da se time broj glagola dodatno smanjuje. Postojeći glagoli obično su u formi sadašnjega ili budućega vremena, u trećem licu ili bezličnim, pasivnim konstrukcijama. Česti su i dekomponovani predikati kao što su:

- Donijeti odluku (umjesto odlučiti),

Primjeri:

učesniku - pravnom licu kada se <i>doneše odluka</i> o stečaju ili likvidaciji Dakle, čim se
ju ili likvidaciji Dakle, čim se <i>doneše odluka</i> o stečaju i/ili likvidaciji, ispunjeni
.2008. godine, predviđeno da se <i>doneše odluka</i> o izboru modela za eliminisanje barijer
II kvartalu 2008. godine, biće <i>donešena odluka</i> o imenovanju Strateškog koordinatora za
II kvartalu 2008. godine biće <i>donešena odluka</i> o imenovanju Koordinatora za prekograni
stepenu u postupku u kojem je <i>donešena odluka</i> kojom je povrijeđeno ljudsko pravo i o
raničenja. U tom cilju može se <i>donijeti odluka</i> o promjeni iznosa izraženih u eurima. G

da uključe građane u process <i>donošenja odluka</i> i u organizaciju javnih rasprava. Ponek
lokalne samouprave i sistema <i>donošenja odluka</i> na nivou opštine će zahtijevati moderni
t, teži poboljšavanju procesa <i>donošenja odluka</i> u skupštinama opština, jačanje vještina
uticati i na najbolje prakse <i>donošenja odluka</i> u skupštinama opština. Na ovaj način, r
građana u lokalnim procesima <i>donošenja odluka</i> je vaoma nisko u Crnoj Gori. Međutim, n

U tekstovima na engleskome jeziku imamo drugačiju situaciju. Naime, infinitiv glagola *decide* češći je od odgovarajućih najčešćih kolokacija *make decision* i *reach decision*. U tim slučajevima, glagolu *decide* najčešće prethode modalni glagoli i to:

$V_{mod} + bare\ Inf.$

- Shall decide – 91 ponavljanje
- May decide – 90 ponavljanja
- Can decide – 51 ponavljanje
- Will decide – 30 ponavljanja
- Izvršiti pregled (umjesto pregledati),

Za razliku od prethodne fraze nijesmo našli niti jedan primjer *izvršiti pregled* u našem korpusu, već nas je analiza korpusa usmjerila u drugom pravcu. Riječ *pregled* najčešće se nalazila u dokumentima s tematikom o javnoa zdravlju i zaštiti potrošača (npr. u Zakonu o zaštiti stanovništva od infektivnih bolesti i u Zakonu o zaštiti potrošača). U skladu sa stručnom terminologijom uz riječ *pregled* najviše linija konkordance bilo je vezano za:

je izvršila ukupno 19.631 inspeksijskih <i>pregleda</i> , od čega 97,66% po službenoj dužnosti,
jediti praćenje rezultata inspeksijskih <i>pregleda</i> po određenim grupama propisa, uključu
inspektora za sprovođenje inspeksijskih <i>pregleda</i> i kontrola u skladu sa inoviranom naci
ja evidencije o izvršenim inspeksijskim <i>pregledima</i> koje vrše republički inspektori zašt
ima, sadržinu zapisnika o inspeksijskom <i>pregledu</i> i drugo. Zakon o opštem upravnom postu

Za polje zaštite potrošača i definisanja pravila inspeksijskog nadzora subjekata

rugih poremećaja zdravlja, 2) ljekarske <i>preglede</i> i druge vrste medicinske pomoći, 3) li
ima prethodnih i periodičnih ljekarskih <i>pregleda</i> zaposlenih na radnim mjestima sa
Konvencija br. 113 o ljekarskom <i>pregledu</i> mornara, usvojena 1959. godine;
ela sprovode se odgovarajući medicinski <i>pregledi</i> i zahvati radi procjene i smanjenja fi
ih pregleda rizične djece, preventivnih <i>pregleda</i> djece sa smetnjama u razvoju kao i

Za polje javnog zdravlja

U tekstovima na engleskome jeziku imamo više varijacija na temu, ali tri najčešće su:

and activities envisaged by the law 3) inspection <i>control</i> by forest inspection, environme
h trade: Environment Protection Agency (inspection <i>control</i> Department), Ministry of Interi
s who are in charge of carrying out the inspection <i>control</i> in the area of sea fisheries. A
le 73 of the Law on Tax Administration. inspection <i>supervision</i> in the meaning of this law
hiness of boats. Inspection in terms of inspection <i>supervision</i> is carried out by the inspe

Za polje zaštite potrošača i definisanja pravila inspekcijskog nadzora subjekata

i

sation and means of protection; medical <i>examinations</i> of employees which are mandatory
pply of drinking water include: medical <i>examinations</i> with a special emphasis on
cal examinations and periodical medical <i>examinations</i> every 12 months. Within the safety
nt mortality. In that sense, preventive <i>examinations</i> are conducted, which include taking of
st grade of secondary school preventive <i>examination</i> of students in the third grade of secon

Za polje javnog zdravlja

Dakle, vidimo da su kolokacijski elementi u crnogorskome i engleskome jeziku različiti u zavisnosti od tematike, tj. konteksta u okviru kojega se ti elementi nalaze. Ostali primjeri dekomponovanih predikata, koji su karakteristični za administrativni stil, su:

- **Podnijeti izvještaj (umjesto izvijestiti),**
- **Izraziti protest (umjesto protestovati),**
- **Uputiti čestitke (umjesto čestitati).**

Administrativni stil obično se dijeli na pet podstilova:

1. **Zakonodavno-pravni**, čiji su žanrovi zakoni, statuti, ustavi, odluke, naredbe, rješenja;
2. **Društveno-politički** - rezolucije povelje, deklaracije, programi, referati, saopštenja, izjave;
3. **Diplomatski** - prepiske, note, demarši, protokoli, memorandumi;
4. **Poslovni** – ugovori, dopisi, fakture, sertifikati, narudžbenice, uplatnice, računi, specifikacije;
5. **Lični** – molbe, žalbe, autobiografije, punomoći, lična dokumenta, upitnici, ankete, formulari, itd.

Naš korpus sastavljen je, prije svega, od prve tri pomenute kategorije. Na osnovu pomenutih karakteristika, možemo bez ustezanja zaključiti da i

naš korpus odgovara administrativnome stilu. Neophodno je istaći da ni jedan tekst ne pripada čistom stilu, već je mješavina različitih registara, od kojih jedan može imati prevagu. Jedna od karakteristika administrativnog stila jeste i impersonalizacija, tj. uvijek govrimo o državama, institucijama, ministarstvima, dakle, uvijek o *njima*. Takvi dokumenti, koji su po pravilu veoma formalni, daju zanimljive rezultate ako ih testiramo na lične zamjenice, i kontrastiramo između dva jezika. Tako smo našli da se lična zamjenica *ona* (nikako ili veoma rijetko u zavisnosti od teksta) ne odnosi na imenovanje pripadnica ženskog roda (+ živa), već je gotovo isključivo povezana drugim imenicama ženskog roda (- živa). U engleskome jeziku za razliku od crnogorskoga dominira prepozicija *it*. Prikazaćemo dobijene primjere:

- podnese Skupštini na potvrđivanje čim **ona** bude u mogućnosti da se sastane
- ambalaža uključena zajedno s proizvodom, **ona** će biti uključena i za potrebe
- aktivnih mjera i/ili zakonskih postupaka, **ona** će o tome odmah obavijestiti tijelo
- osim u slučajevima od posebne hitnosti, **ona** će Privremenom odboru dostaviti
- akata potrebno obaviti javnu raspravu, **ona** će utvrditi nacrt zakona, program
- može dati saopštenje jedne države kojom **ona** daje objašnjenja ili opravdanja za
- načela međunarodnih organizacija ili su ih **ona** i donijela (npr. UN i EU)
- van teritorijalnih voda, pod uslovom da **ona** ima isključiva prava na obradu tog
- moguće dokazati: (a) da je vraćena roba **ona** ista koja je bila izvezena; i (b) da
- opterećenja vanbolničke zdravstvene službe, **ona** je tokom 2006. godine u daleko najvećem
- što se tiče vanbolničke zdravstvene službe, **ona** je u toku 2006. godine u daleko najveće
- onih iz tar. broja 9402), uključujući **ona** koja se mogu pretvoriti u ležajeve
- se može odložiti s obrazloženjem da bi **ona** mogla biti uplitanje u istragu, sudsko
- nepovoljnom ekonomskom uticaju kršenja, **ona** može podnijeti pismeni zahtjev
- gnojiva s uslovima iz stava 1 ovog člana, **ona** može preduzeti odgovarajuće mjere nakon

- ispunjenu neku obavezu iz ovog Sporazuma, **ona** može preduzeti odgovarajuće mjere
- 3(1) Protokola ili prema STO Sporazumu, **ona** ne može pokrenuti postupak za rješavanje
- zakonodavstvo i matični odbor) i Vladi, ako **ona** nije podnosilac predloga. Predlog zakon
- jenih materijala bez porijekla ili, ako **ona** nije poznata i ne može se utvrditi
- organizacija Crne Gore i jedino je **ona** odgovorna za monetarnu politiku, uspostavljanje
- može popuniti kad se proizvodi na koje se **ona** odnosi izvoze, ili nakon njihovog izvoz
- izdati nakon izvoza proizvoda na koje se **ona** odnosi, ako: (a) uvjerenje nije izdato
- individualnim subjektima da štite i **ona** prava koja za njih proističu iz direktiva
- sudsku opomenu ili oslobođenje od kazne, **ona** se izvršava na slobodi i ne može trajati
- Rjeđe smo nalazili ličnu zamjenicu *on* iz razloga što su imenice koje se zamjenjuju češće ženskoga roda (država, skupština, vlada, služba), što ličnu zamjenicu muškoga roda stavlja po strani:
- nudi zaista mnogo korisnih mogućnosti. **On** kombinuje dvije tehnologije, segmentaciju
- to je ono što većina ljudi zove Trados. **On** opslužuje prevodilačku memoriju, omogućava
- primjenu i sprovođenje ovog Sporazuma. **On** će se sastajati na odgovarajućem nivou

Prva dva puta govorimo o računarskom programu, a u trećem primjeru, očigledno, o sporazumu.

Kao što smo već rekli, u engleskom jeziku dominirala je zamjenica *it*, a ne direktni prevodni ekvivalenti za *on* i *ona* – *he* i *she*. Zašto je to tako, odgovor treba potražiti u tome što je u engleskom jeziku drugačija percepcija roda i pravila nalažu da se riječi koje nemaju prirodan rod definišu sa *it*.<sup>1</sup> Pobjrojaćemo samo neke primjere:

- d liability companies. In that purpose, **it** is necessary to initiate amendments of

<sup>1</sup> Osim u slučajevima, npr. emocionalne privrženosti, što u administrativnom stilu svakako nije slučaj.

- ect to separate annual programmes. When *it* comes to motivation, special attention
- he company has a positive attitude when *it* comes to further tertiary education of
- be a simply administrative service as *it* was the case with previous personnel de
- ed the training, i.e. 1.5% in 2008. When *it* comes to training of active population
- ed database are the Statistical Office (*it* uses a large quantity of data by other
- residents and non-residents equal when *it* comes to tax and contribution payments.
- lling within 2004-2008 period. In 2008, *it* was 21% while it has amounted to 20.50%
- -2008 period. In 2008, it was 21% while *it* has amounted to 20.50% since 1 January,
- ocal Self-Government define areas where *it* is possible to carry out decentralisati
- of fiscal policy, on the basis of which *it* plans the main categories of receipts a
- nding units, and if there is a deficit, *it* shall determine sources of funds for it
- aw on the State Budget and shall submit *it* to the Government in October. The compe
- ision on municipality budget and submit *it* to the Ministry of Finance for insight

### 1.2.2 Izražavanje zabrane

Prilikom pisanja zvaničnih državnih dokumenata veoma bitan dio jesu prava i obaveze subjekata koje se definišu tim dokumentima. Prava i obaveze podrazumijevaju dozvoljene i nedozvoljene radnje koje se pak različitim jezikim sredstvima odobravaju ili zabranjuju. Posebno zanimljive, u tome kontekstu, jesu zabrane i na osnovu naše analize našli smo da i u tekstovimana na crnogorskome i u engleskome jeziku češće se koristi direktni glagoli koji označavaju zabranu (zabraniti ili prohibit, ban, proscribe), od konstrukcija *not + glagol dopuštanja* (ne+dozvoliti, ne+dopustiti, not allow, not approve). Moćuća su dva objašnjenja za ovu pojavu:

- Direktni glagoli koji označavaju zabranu predstavljaju dio dobro ustanovljenog šablona,
- Psihološki efekat (+ zabranjeno) je jači od (- dozvoljeno).

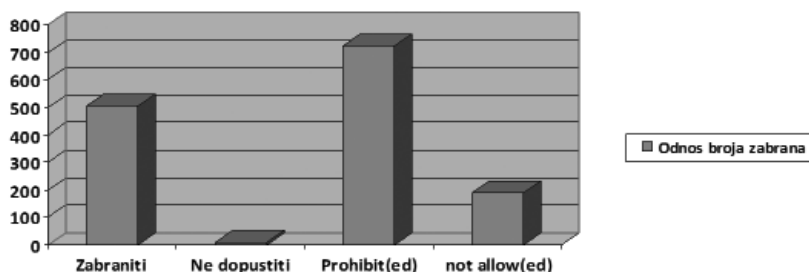


Navešćemo neke od primjera:

- gation Magistrate, who has the right to **ban** all correspondence if it harms the inve
- s duties postepeno smanjenje carina prohibit zabraniti prohibition zabrana p
- ntains a set of laws whose stipulations forbid discrimination, promote equality and es

Primjeri prijevoda na našem jeziku:

- Magistrat, koji ima pravo da **zabrani** čitavu prepisku ako ona šteti
- s duties postepeno smanjenje carina prohibit **zabraniti** prohibition zabrana p
- poštuje niz zakona koji **zabranjuju** diskriminaciju i promovišu jednakost



2

S druge strane, glagoli koji su neutralni u pogledu zabrane (postupati, birati, consider, use ...), odn. oni glagoli čije semantičko polje ne uključuje i značenje zabrane, da bi izrazlili zabranu zahtijevaju negaciju *not/ne* (ili *nije/neće*).

### 1.2.3 Uočene greške

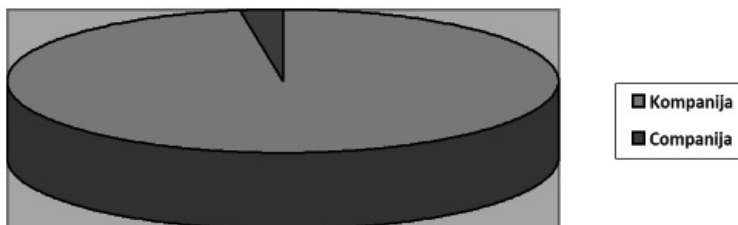
Ono što je naravno bilo neizbježno jeste i to da smo nailazili na greške u našem korpusu. Pod greškama, prije svega, podrazumijevamo tipografske greške koje se mogu svrstati u sedam kategorija:

- **Companija – primjer negativne interferencije s engleskim jezikom**
- **Leter – primjer negativne interferencije s crnogorskim jezikom**

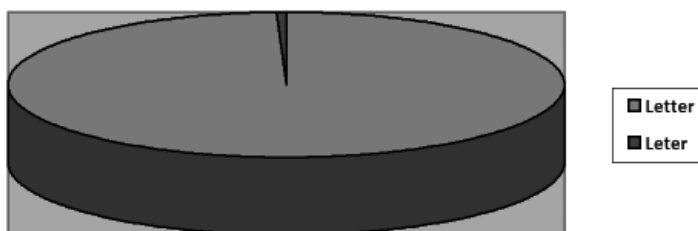
<sup>2</sup> Pored druge stavke *ne dopustiti*, u analizu smo uključili i *ne dozvoliti* (zajedno sa upotrebom nastavaka), ali nam veličina grafikona nije dozvoljavala ubacivanje istog.

- **Ćlan – greške u izboru dijakritičkog znaka**
- **SkupštinaCrne Gore – greške u pravljenju razmaka između riječi**
- **zvještaji – nedostatak slova – nepotpuna riječ**
- **startegije – nepravilan redosljed slova**
- **republika – dupliranje slova**

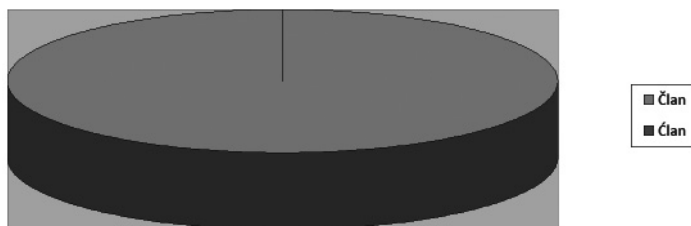
Ono što smo takođe pronašli jeste da se sve greške ponašaju po istom principu, a to je da se nalaze van zone učestalosti ponavljenja koja je karakteristična za ispravne verzije datih riječi. Ilustrovaćemo neke primjere dijagramima na kojima će se jasno vidjeti šablon ponašanja tipografskih grešaka u našem korpusu.



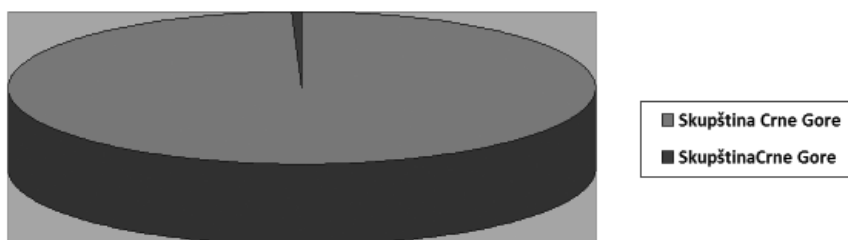
Dijagram 1 - Odnos učešća greške br. 1 – negativne interferencije engleskog jezika na domaću riječ



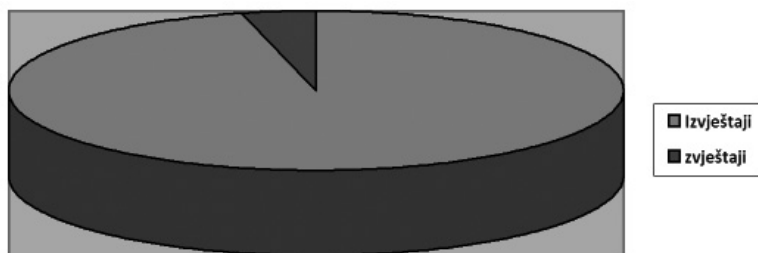
Dijagram 2 – Odnos učešća greške br. 2 – negativne interferencije crnogorskog jezika na englesku riječ



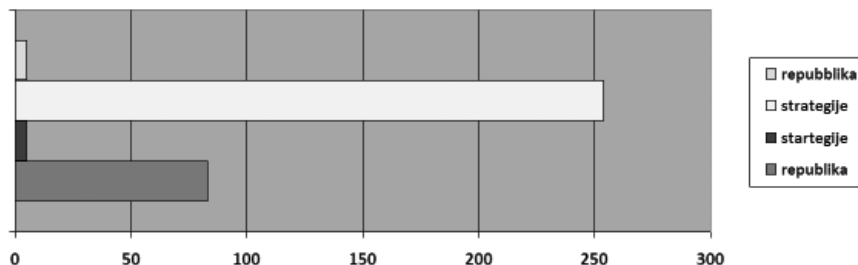
Dijagram 3 – Odnos greške br. 3 – pogrešan izbor dijakritičkog znaka



Dijagram 4 – Odnos greške br. 4 – neodvajanje riječi



Dijagram 5 – Odnos greške br. 5 – nepotpuna riječ



Dijagram 6 – Odnos greške br. 6 i 7 – nepravilan red slova u okviru riječi i nepotrebno dupliranje znakova

Jasno se može uočiti šablon da se pogrešna riječ javlja višestruko manje puta u odnosu na ispravni oblik te riječi, što je i očekivano. Mi smo ovde izdvojili samo neke od primjera grešaka koje se mogu naći, ali ono što je zanimljivo jeste da i sve ostale greške, bez obzira na svoje karakteristike, prete identičan način ponašanja, kao što smo pokazali na Dijagramima 1 – 6. Sve greške će se uvijek nalaziti van optimalnog broja ponavljanja i to je pravilnost i univerzalna pojava kako za crnogorski tako i za engleski jezik.

### 1.2.3.1 Hapax legomena i njihova veza sa greškama

Iza ovoga zanimljivog naziva kriju se riječi koje su u nakom tekstu ili grupi tekstova pojavljuju samo jedanput. Naziv potiče od grčkih riječi *hapax* i *legomenon* (mn. *legomena*) – rečeno samo jedanput. U našem korpusu (od ukupno 93226 riječi bez ponavljanja, ima 45450 (48,75%) takvih riječi i veoma često su najuže povezane sa nekom strukom (npr. latinizmi i vrste bolesti u medicini), a takođe neke od tih riječi predstavljaju pogrešno napisane oblike drugih riječi koje se inače mnogo češće nalaze u korpusu. Vjerovatno protivno svačijoj intuiciji sljedeći oblici riječi takođe se pominju samo jednom u cijelom milionskom korpusu, npr:

ze, banke i druga pravna lica. Pojedini <b>akcionar</b> Centralne Depozitarne Agencije može ima	tte Competence: Central Register of the <b>Commerical</b> Court 2. Making seals and stamps Timesc
radu odgovarajućeg dokumenta ( brošure, <b>časopisi</b> , publikacije, bilteni ... ) Institucije	e use complex analytical techniques and <b>compute</b> application software. Environmental Pro
Plc (»Zavod za građevinske materijale, <b>geotehničke</b> i hemijske analize« AD) – Nikšić, is an	for himself/herself, marital partner or <b>illegitimate</b> partner, and for children in case that

m se definišu: ciljevi, trajanje, vrste <b>gimnazija</b> , matura, horizontalna i vertikalna proha	y use for non-agricultural purposes, is <b>recultivated</b> technically, chemically and biologicalall
o je i kroz izjavu Centra za podvodno i <b>humanitarno</b> razminiranje da je Crna Gora jedina zem	s they should realistically decline and <b>stabilise</b> at 1.5% - 1.8% of Montenegrin GDP. III.

Nadovezujući se za prethodno poglavlje pružićemo i primjere grešaka koje se lako mogu otkriti u tekstovima uključenim u naš korpus:

putovanja mogu da obavljaju turističke <b>agencije</b> koje, pored odobrenja za rad (čl.11 st.	plementation and evaluation of measures <b>aggenst</b> TC should be included in programs for M
u (Sl. list CG, br.66/08), odgovorna je <b>Agencioja</b> za civilno vazduhoplovstvo Crne Gore. Z	ional Protocol (OG of MN, international <b>agreements</b> , 2/09); Law on Veteran and Disability P
oblasti zdravstva, veterine, farmacije, <b>ahitekture</b> , nastavno-pedagoškog rada itd. Svakom i	d lays down requirements for <b>qualitiy</b> of <b>agricultural</b> products. Veterinary Administration is
ravlja u Crnoj Gori su otvoreni sledeći <b>centari</b> i jedinice za mentalno zdravlje u zajed	alcohol per 100 kilograms of goods; 4) <b>chemicalz</b> and cosmetics manufacturing. Use of min
ciljeva drugih segmenata društva. Takve <b>strategije</b> su Strategija održivog razvoja, Strateg	ce stipulates that the doctor – foreign <b>citizent</b> , who intends to perform <b>healt</b> activity

Kao što se može videti greške u oba jezika u osnovi funkcionišu na identične načine. U crnogorskome jeziku su to obično suvišak ili manjak znakova u riječima, kao što je to slučaj i u engleskome jeziku, ono što se poznato pod imenom *spelling*.

### Literatura

- Agresti, Alan. 2002. *Categorical Data Analysis*. Hoboken, NJ: Wiley.
- Aijmer, Karin. 1984. *Sort of and kind of* in English conversation. *Studia Linguistica* 38/2: 118–128.
- Arppe, Antti and Juhani Järvikivi. 2007. Every method counts: combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory* 3/2: 131–159.
- Aston, Guy (ed.). 2001. *Learning With Corpora*. Bologna. CLUEB.
- Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

- Baayen, R. Harald, Richard Piepenbrock, and Leon Gulikers. 1995. *The CELEX Lexical Database (Release 2)*. Philadelphia, PA: Linguistic Data Consortium.
- Baker, Paul, Andrew Hardie, and Tony McEnery. 2006. *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Barnbrook, Geoff. 1996. *Language and Computers*. Edinburgh: Edinburgh University Press.
- Biber, Douglas. 1988. *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, Douglas, Susan Conrad, and Randi Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Bod, Rens, Jennifer Hay, and Stefanie Jannedy (eds.). 2003. *Probabilistic Linguistics*. Cambridge, MA: MIT Press.
- Bortz, Jürgen. 2005. *Statistik für Human- und Sozialwissenschaftler*. 6th edn. Heidelberg: Springer.
- *British National Corpus*, version 2 (BNC World). 2001. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <<http://www.natcorp.ox.ac.uk/>>
- Church, Kenneth Ward, William Gale, Patrick Hanks, and Donald Hindle. 1991. Using statistics in lexical analysis. In *Lexical Acquisition: Exploiting On-line Resources to Build a Lexicon*, ed. Uri Zernik, 115–164. Hillsdale, NJ: Lawrence Erlbaum.
- Church, Kenneth Ward, William Gale, Patrick Hanks, Donald Hindle, and Rosamund Moon. 1994. Lexical substitutability. In *Computational Approaches to the Lexicon*, ed. Beryl T. Sue Atkins and Antonio Zampolli, 153–177. Oxford: Oxford University Press.
- Clark, Herbert H. and Jean E. Fox Tree. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84/1: 73–111.
- Conrad, Susan. 2000. Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly* 34: 548–560.
- Diessel, Holger and Michael Tomasello. 2005. Particle placement in early child language: a multifactorial analysis. *Corpus Linguistics and Linguistic Theory* 1/1: 89–112.
- Dilts, Philip and John Newman. 2006. A note on quantifying “good” and “bad” prosodies. *Corpus Linguistics and Linguistic Theory* 2/2: 233–242.
- Ellis, Nick C. 2002a. Frequency effects in language processing and acquisition. *Studies in Second Language Acquisition* 24/2: 143–188.

- Gahl, Susanne and Susan Marie Garnsey. 2004. Knowledge of grammar, knowledge of usage: syntactic probabilities affect pronunciation variation. *Language* 80/4: 748–775.
- Gries, Stefan Th. 2001. A corpus-linguistic analysis of *-ic* and *-ical* adjectives. *ICAME Journal* 25: 65–108.

**Igor IVANOVIĆ**

**CORPUS AND COMPUTATIONAL LINGUISTICS  
THROUGH THE MONTENEGRIN LANGUAGE**

This paper deals with the analysis of the part of Montenegrin and English version of the EC Questionnaire. From the total analysis from our doctoral thesis, we have chosen two aspects. The first one is about the analysis of impersonalized register, and the second one is about finding the connexion between *hapax legomena* and mistakes we found in the abovementioned corpus. We will show how the use of corpora and computers can shed new light onto the traditional linguistic research. Apart from all of that, this paper should also be understood as a promotion of Corpus and Computational linguistics in Montenegro.

Key words: *corpus, linguistic analysis, computers, the EC Questionnaire*