

| | |
|--|-----------|
| 0 Predgovor | 7 |
| 1 Uvod | 9 |
| 1.1 KoRP – korpus | 9 |
| 1.2 KoRP – korpus strokovnih besedil | 9 |
| 1.3 KoRP – namen, vsebina, zgradba | 9 |
| 2 Delo s korpusom | 13 |
| 2.1 Registracija | 13 |
| 2.2 Prijava | 13 |
| 2.3 Iskanje | 14 |
| 2.3.1 Osnovno iskanje | 15 |
| 2.3.1.1 Enostavno iskanje besede | 16 |
| 2.3.1.2 Iskanje po prvem kanalu | 17 |
| 2.3.1.3 Iskanje po frazah | 21 |
| 2.3.1.4 Iskanje po bližini | 22 |
| 2.3.1.5 Iskanje po drugem kanalu | 22 |
| 2.3.1.6 Iskanje s funkcijama <i>notranji in, notranji ne</i> | 24 |
| 2.3.1.7 Iskanje s funkcijami <i>in, ali, ne</i> | 26 |
| 2.3.2 Iskanje ločil in drugih znakov | 28 |
| 2.3.3 Razširjeno iskanje | 29 |
| 2.3.3.1 Glava dokumentov | 31 |
| 2.3.3.2 Okence več o besedilu | 32 |
| 2.4 Shranitev rezultatov iskanj | 38 |
| 2.4.1 Sprotni rezultati iskanj | 38 |
| 2.4.2 Shranjeni rezultati iskanj | 39 |
| 2.5 Obdelava rezultatov iskanj | 40 |
| 2.5.1 Statistika | 41 |
| 2.5.1.1 Razvrščanje v seznamih statistike | 43 |
| 2.5.2 Urejanje | 44 |
| 2.5.3 Sito | 46 |
| 2.5.3.1 Kombiniranje sita in statistike | 47 |
| 2.5.4 Vzorec | 48 |
| 3 Interpretacija podatkov | 51 |
| 4 Priloge | 53 |

0

1. *Ali veste, kaj je odključnica, zakaj terminološko ni ustrezna raba zveze stiki z javnostmi in kateri je slovenski izraz za angleškega spin doctor?* Če ste odgovorili z »ne«, ste kvalificirani za rabo KoRP-ja. Tudi če ste odgovorili z »da«, ste kvalificirani za rabo KoRP-ja.
2. *Ali vas z jezikoslovnega vidika zanimajo strokovna besedila, npr. kateri so najpogostejši glagoli v njih?* Če ste odgovorili z »da«, ste kvalificirani za rabo KoRP-ja. Tudi če ste odgovorili z »ne«, ste kvalificirani za rabo KoRP-ja.
3. *Ali ste že slišali za korpus?* Ne glede na to, ali ste odgovorili z »da« ali z »ne«, ste kvalificirani za rabo KoRP-ja.



O vsebini priročnika, ki si ga ogledujete, skoraj vse pove že njegov naslov, če pa ste si že predhodno ogledali spletno stran, na kateri je korpus KoRP, tj. <<http://www.korp.fdv.uni-lj.si>>, veste tudi, kaj pomeni KoRP: *Korpus besedil odnosov z javnostmi*. Avtorici si domišljava, da ste bralci tega priročnika dvoji, zato je nadaljevanje predgovora namenjeno enim in nato še drugim.

Priročnik je nadgradnja oz. priredba priročnika *Kaj početi z referenčnim korpusom FidaPLUS* (Š. Arhar, 2006, <<http://www.fidaplus.net>>). Tisti, ki priročnik o korpusu FidaPLUS poznate in ste torej dobro seznanjeni z delom s konkordančnikom ASP32, boste marsikatero stran v tem priročniku lahko preskočili, svetujemo pa vam, da si preberete poglavje o razširjenem iskanju, zgradbi KoRP-ja in glavi dokumentov, ki so v njem; če je vaše znanje potrebno osvežitve, pa vas vabiva na trening prek 22 nalog.

Primarno je priročnik namenjen tistim, ki se s korpusi niso še nikoli srečali, zato je pripravljen tako, da lahko bralec postane samostojen in uspešen uporabnik korpusa že samo, če usvoji veščine iz osnovnega iskanja po prvem kanalu in del statističnih obdelav rezultatov iskanj, medtem ko so nadgradnja v iskanje po drugem kanalu, razširjeno iskanje in zapletenejše statistične analize ter urejanje rezultatov iskanj namenjeni tistim, ki bi korpus kot zbirko podatkov radi obvladali bolj celovito. Bralce, ki se s korpusom srečujejo prvič, vabiva, da ga spoznajo tako, da po vrsti prebirajo poglavja, podpoglavja, točke in naloge v priročniku, po vsaki taki samostojni enoti pa naredijo lastno poizvedbo v korpusu. Bolj radovedni ali izzivov željni se lahko po začetnih korakih v KoRP podajo tudi samostojno, v priročniku pa nato poiščejo dodatna pojasnila ali rešitve, ko jim bo kondicija pošla.

Seznam besedil, ki so v KoRP-ju, je dostopen na korpusovi spletni strani. Če česa v korpusu ne boste našli, to pomeni, da iskanega pač ni v

besedilih, ki so v njem. Hkrati to ne pomeni, da tega, kar ste iskali, ni tudi v jeziku. KoRP je vzorec, skušali smo ga narediti čim bolj reprezentativnega, tudi na osnovi vaše analize in kritike pa se ga da še izboljšati in dopolniti – saj veste: v slogi je moč.

Ker predvidevava, da ste zgornje kvalifikacije uspešno opravili, vas pozdravljava v reprezentanci koerpejevcev. Dobrodošli!

Avtorici

Ljubljana, 20. 12. 2007

1

1.1 KoRP – korpus

Začeti je mogoče z zapleteno opredelitvijo:

Korpus je enovita, standardno označena in notranje strukturirana elektronska zbirka besedil, za izbor katerih so bila merila vnaprej določena.

Za povprečnega uporabnika korpusa bi zadoščala že precej bolj približna opredelitev: korpus je zbirka besedil na internetu. Po tej zbirki je mogoče iskanje različnih podatkov. Ker govorimo o jezikovnih korpusih, so ti podatki jezikovni, prek jezika kot izraza pa seveda tudi vsebinski.

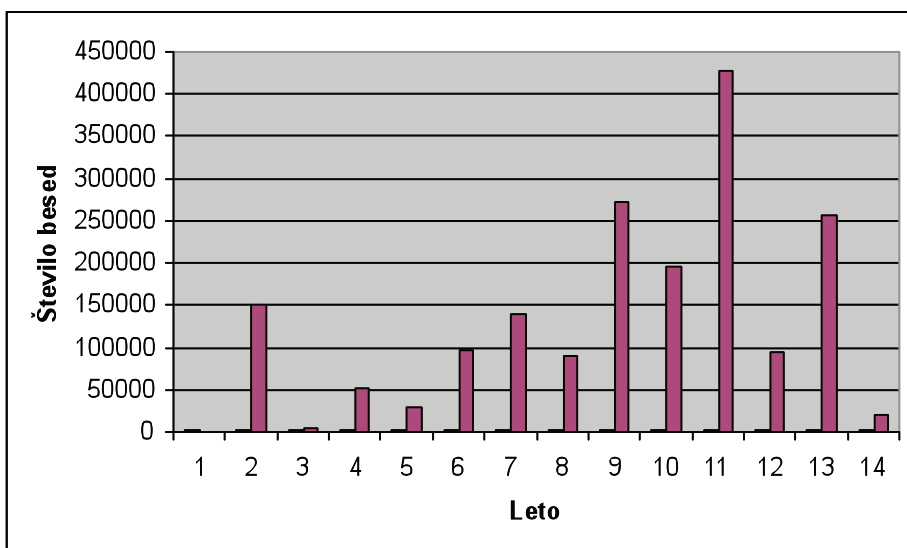
1.2 KoRP – korpus strokovnih besedil

Korpusi so narejeni za različne namene. Ta namen vpliva na njihovo obliko, velikost in lastnosti, kar pomeni, da imamo več vrst korpusov. Ena vrsta so korpusi strokovnih besedil. Kot že ime pove, gre za *zbirko besedil določene stroke*. Taki korpusi so lahko številni, mi smo zgradili korpus besedil odnosov z javnostmi.

1.3 KoRP – namen, vsebina, zgradba

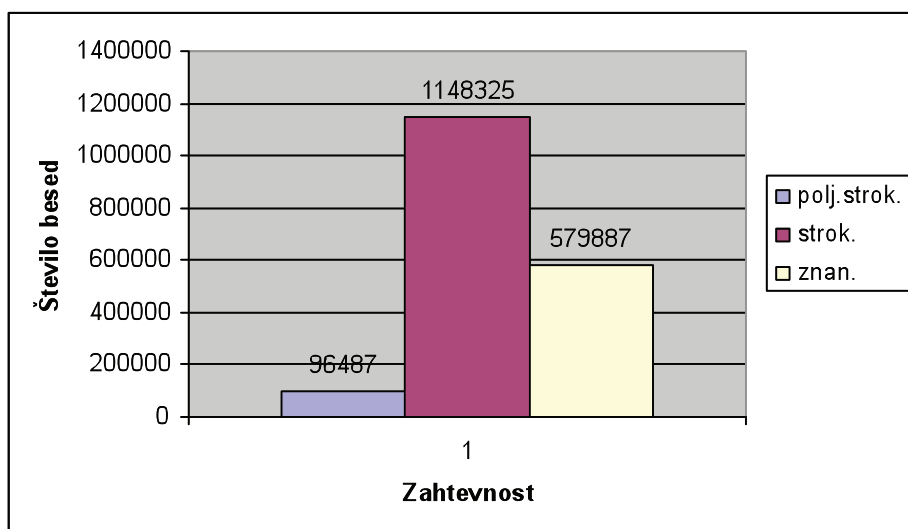
Namen KoRP-ja je izdelava terminološkega slovarja te stroke, v širšem pogledu pa ga lahko razumemo tudi kot zbirko znanja slovenskih odnosov z javnostmi. Korpus je jezikoslovno označen in je od 2. 7. 2007 brezplačno javno dostopen na <http://www.korp.fdv.uni-lj.si>. Vsebuje 1,824.699 besed, je enojezični, sinhroni, pisni in trenutno statični korpus. Z zadostitvijo vnaprej opredeljeni mreži meril za nabor in izbor besedil skuša KoRP čim boljje predstavljati celotno stroko odnosov z javnostmi pri nas. KoRP je nastal na Fakulteti za družbene vede s finančno pomočjo podjetja Pristop, d. o. o.

KoRP vsebuje besedila iz obdobja 14 let, besedila več kot 20 različnih besedilnih vrst, besedila tiskanega in elektronskega prenosnika, izvorna slovenska in prevedena besedila ter lektorirana in nelektorirana besedila. Število besed v posameznih delih korpusa prikazujejo grafi od 1 do 5.



Graf 1: Število besed v KoRP-ju glede na letnico izida/nastanka besedila.

| Leto | Število besed | Leto | Število besed |
|----------|---------------|-----------|---------------|
| 1 = 1994 | 1.038 | 8 = 2001 | 89.566 |
| 2 = 1995 | 149.806 | 9 = 2002 | 271.598 |
| 3 = 1996 | 4.000 | 10 = 2003 | 196.185 |
| 4 = 1997 | 51.107 | 11 = 2004 | 426.410 |
| 5 = 1998 | 28.588 | 12 = 2005 | 93.673 |
| 6 = 1999 | 96.059 | 13 = 2006 | 256.464 |
| 7 = 2000 | 139.250 | 14 = 2007 | 20.955 |



Graf 2: Število besed v KoRP-ju glede na zahtevnost besedil.