

2.2 Od konkordance do kolokacije – prvi del

2.2.1 Korpus KRES

<http://www.korpus-kres.net/>

1. Odpremo korpus Kres in vtipkamo v iskalno okence besedo *pljuvalnik*. Ogledamo si rezultate v konkordančniku in spoznamo:

- kaj je **konkordanca** oz. **konkordančni niz**, **konkordančno jedro**,
- kje najdemo število konkordanc,
- kako pridemo do širšega **sobesedila**, **metapodatkov o** besedilu in **korpusnih oznak**,
- kaj so **filtri** in kako jih uporabljamo.

2. Raziščemo, kaj pomeni beseda *pljuvalnik*. Kaj pomeni *brbotalnik*? Poznamo sopomenko za *brbotalnik*? Kako pogosto in v katerih besedilih se pojavlja *brbotalnik*, kako pogosto pa sopomenka?

3. Katere besede na *-nik* še poznamo? Naštejemo tiste, za katere mislimo, da so v korpusu najpogostejše. Odpremo zavihek Seznam in vnesemo iskalni pogoj **nik*. Ogledamo si rezultate. Ogledamo si filter Besedna vrsta in komentiramo, kar najdemo pod Neuvrščeno in Prislov.

4. Ogledamo si besedo *lastnik* in razmislimo o besedni zvezi *lastnik + koga ali česa*. Naštejemo nekaj primerov. Odpremo zavihek Okolica in vnesemo besedo *lastnik* + prvo mesto na desni. Ogledamo si rezultate in spoznamo pojem **kolokator** in različne možnosti razvrščanja le-teh. Kolokatorje filtriramo na samostalnike in uredimo glede na pojavitve v okolici.

5. Kdo najde samostalnik, ki se v korpusu najpogosteje pojavlja? Kot zanimivost si ogledamo besedni oblak z najpogostejšimi samostalniki iz korpusa Kres.

6. Kako blizu so bili naši odgovori korpusnim podatkom? Zakaj se je naša **jezikovna intuicija** v nekaterih primerih izkazala za ustrezno, v drugih ne?

2.2.2 Korpus GOS

<http://www.korpus-gos.net>

1. Odpremo korpus GOS in vtipkamo v iskalno okence besedo *recimo*. Ogledamo si rezultate v konkordančniku in spoznamo razlike korpusa GOS glede na Kres:

- možnost poslušanja posnetka (poslušamo nekaj primerov),
- podatki o besedilih so drugačni (*kaj pravte recimo na to vse vi ste tudi podjetnica* – ogledamo si prekrivanje lastnih imen in oznako za nerazumljivo),

- razlike v filtrih, poskušamo filtrirati, npr. *Maribor, nejavni nezasebni*. Kakšne vrste komunikacije se najdejo v tej vrsti oznake?

2. Poiščemo v korpusu besedo *ful* in si ogledamo filtre. Kaj lahko ugotovimo glede tipa govora, regije snemanja, spol, starost ... Lahko glede na podatke zaključimo, da ženske uporabljajo besedo *ful* pogosteje kot moški? Za pomoč pri odločitvi si ogledamo podatke O korpusu.

3. Ponovimo, da je korpus GOS transkribiran na dveh nivojih. Odpremo zavihek Seznam in izberemo Iskanje po standardiziranem zapisu, iščemo besedo *lahko*. Ogledamo si rezultate in vsak poišče obliko in poslušá posnetke, ki so najbližje njegovemu narečju ali ki se mu zdijo najbolj zanimivi.

2.3 Od konkordance do kolokacije – drugi del

2.3.1 Korpus Janes

http://nl.ijs.si/noske/janes.cgi/first_form?corpname=janes

1. Odpremo korpus Janes in vnesemo v iskalno okence besedo *valjda*. Ogledamo si rezultate v konkordančniku in spoznamo značilnosti vmesnika:

- konkordančni niz je podoben, na levi imamo nekaj osnovnih **metainformacij** o viru besedila, npr. da gre z tvit in kdo je avtor,
- klik na konkordančno jedro odpre sobesedilo, klik na informacije na levi dodatne metapodatke,
- možnost, da si ogledujemo zadetke v obliki povedi (*Možnosti prikaza > Stavek*),
- možnost, da podatke sortiramo – ogledamo si funkcijo *Premešaj*, ki premeša zadetke, da npr. niso na začetku samo tviti,
- omenimo možnost, da vzorčimo in filtriramo konkordančni niz,
- ogledamo si seznam oblik besede (*Frekvenca > Oblike niza*),
- ogledamo si seznam izvornih dokumentov (*Frekvenca -> Dokumenti*)
- ogledamo si možnosti izdelave seznama kolokatorjev: attribute *word*, v razponu 1 do 1. Uredimo zadetke po frekvenci in se pogovorimo o rezultatih.