

- 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 poslušča 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.

2. Samostojno delo 1: primerjava besed *neumen* in *glup* v korpusu Janes.

- Raziskovanje poteka v parih.
- Eden od dijakov poišče v korpusu Janes besedo *glup*, drugi pa besedo *neumen*.
- Ugotovita, kako pogosto se vsaka od besed pojavlja in postavita hipotezo, zakaj je tako.
- Izdelata seznam kolokatorjev na mestu desno tik ob besedi in si ogledata prvih 100 (dve strani) rezultatov.
- Primerjata oba seznama kolokatorjev in po potrebi posamične konkordance (s klikom na P pred kolokatorjem): kakšno besedišče se pojavlja? Kako pogosti so kolokatorji?
- Katere ugotovitve lahko sklenemo iz podatkov? So podatki potrdili ali ovrgli hipotezo?
- (V sredo bomo spoznali orodje *Primerjalne skice*, ki olajša tovrstne primerjave med besedami.)

3. Samostojno delo 2: različni žanri korpusa Janes.

- Delo poteka v štirih skupinah.
- Vsaka od skupin dobi enega od žanrov: *tvit*, *blogi*, *forumi* ali *komentarji* in v konkordančniku za delo izbere ustrezen podkorpus.
- Izberemo besedo ali besedno zvezo, za katero predvidevamo, da se bo pojavljala v vseh žanrih (npr. *itak*).
- Vsaka skupina v svojem podkorpusu naredi naključen vzorec stotih konkordanc.
- Vsaka skupina pregleda svoje konkordance (stavčni pogled) in zabeleži čim več zanimivosti v zvezi z jezikom v njih, npr. ali se pojavljajo posebni znaki in če da, v kakšni funkciji so, kako se uporabljajo ločila, kako se besede zapisujejo, se pojavljajo tujejezične besede ali zveze, sleng, ali katera od besed dobiva nov pomen, kakšen je odnos avtorjev besedila do vsebine, ki jo sporočajo ...
- Skupine primerjajo rezultate in ugotovimo, katere značilnosti se pojavljajo v različnih žanrih, katere pa so značilne za posamezen žanr.

2.3.2 Korpus Šolar

www.korpus-solar.net/

1. Odpremo korpus Šolar in vnesemo v iskalno okence kot napako besedo *otrok*. Ogledamo si rezultate v konkordančniku in spoznamo razlike vmesnika:

- mogoče je iskati po jezikovnih napakah učencev in popravkih učiteljev,
- napake in popravki so izpisani v konkordančnem nizu in v razširjenem kontekstu,
- možnosti urejanja rezultatov so primerljive, ogledamo si podatke *Oblike niza*, *Dokumenti*.

2. Razmislimo in naštejemo nekaj napak, ki se nam zdijo tipične za šolska besedila. Ogledamo si možnosti iskanja s pomočjo *Oznake napake*, npr. napake na ravni besedišča, če je čas, še seznam *Oblike niza* in posamezne primer (npr. *in*, *Hamlet*).

Za obnovitev znanja in dodatne ideje glede uporabe korpusov lahko obiščete spletno stran **Portal jezikovnih virov**. *Janes*, ki je povsem nov projekt, sicer še ni predstavljen, so pa na voljo videopredstavitve korpusov *Gigafida* (ki ima enak konkordančnik kot *Kres*), *GOS* in *Šolar*.