

Navodila za uporabo računalniškega programa SLON-13

(Analiza slovenskih besedil)

Jure Zupan, Blaž Čeh
Kemijski Inštitut Ljubljana
Hajdrihova 19,
Ljubljana
Slovenija

DP-KI 2448
Ljubljana, november 2008

Kazalo

Uvod	
1. Namen	4
2. Zasnova in shema programa SLON-13	5
3. Posamezne opcije programa SLON-13	6
3.1 Analiza črk, parov in tripletov	6
Podatki, ki jih dobite pri črkovni analizi	7
Splošno o izboru parov in tripletov	8
3.2 Besedna analiza	10
Samostalniška obrazila	13
Pridevniška obrazila	13
Glagolska obrazila	14
3.3 Konkordance in konkordančno iskanje	15
Konkordančna zahteva ali konkordančna beseda	16
Konkordančno iskanje v obliki ali-ali	19
3.4 Iskanje besed	22
3.5 Iskanje rim	24
3.6 Pomoč	27
4. Inštalacija programa	28
5. Zahvala	29
6. Indeks	30

Priloga:

Zgoščenka s programi, bazami podatkov in nekaterimi slovenskimi besedili

Uvod

Program SLON-13 je trinajstga verzija programa SLON, čigar zasnova in razvoj sta se pričela leta 1991. V primerjavi s predhodnimi verzijami, ki so bile vse napisane za delovanje in upravljanje v MS DOS[®] operacijskem sistemu, predstavljanova verzija SLON-13 uporabniku prijaznejšo varianto, ki deluje v okolju MS Oken (MS Windows[®]) s pomočjo posredniške lupine napisane v Visual Basicu[®].

Zaradi novega načina približevanja in povezave med uporabnikom in programom, je bilo treba delno spremeniti tudi prejšnje programske enote, ki so vse sprogramirani v jeziku FORTRAN90. Bistveno je bilo potrebno spremeniti organizacijo komuniciranja z uporabnikom in možnosti hranjenja dobljenih rezultatov. Glavno delo pri trinajsti različici programa SLON je bilo narediti enostavno in hitro komunikacijo med programskimi enotami, slovarjem in lupino.

Lupino programa SLON-13 je v Visual Basicu[®] izdelal **Blaž Čeh**. S potrpljenjem, z veliko truda in tudi svojimi idejami je močno prispeval k temu, da je verzija programa SLON-13 uporabna tudi za tiste, ki niso večji že zastarelega načina obdelav v MS DOS[®] operacijskem sistemu.

Zaradi precej obširnih izpisov in količine rezultatov program SLON-13, žal, ni najbolj primeren za interaktivno delo v MS Oknih. Rezultati, ki jih program SLON-13 daje predvsem v opcijah »Analiza črk ...«, in »Besedna Analiza« in »Konkordance« so tako obširni, da jih uporabniku ni mogoče podati znotraj majhnega komunikacijskega okna. Zato je okno v lupini v glavnem namenjeno hitremu pregledu rezultatov in odločitvi ali naj uporabnik dobljene rezultate shani za kasnejšo ročno oz. računalniško obdelavo ali ne.

Sedanja zasnova programa SLON-13 je nastala, kot kompromis med zaslonskim zapisom starega MS DOS[®]-a in sodobnim zapisom v interaktivnih oknih. Uporabnik lahko vse rezultate hitro pregleduje in jih po potrebi tudi med delom shranjuje na različnih datotekah.

Slovar programa SLON-13 obsega približno 160,000 gesel. Ker je velika večina besed v slovarju SLON-13 (vsi samostalniki, pridevniki, glagoli, zaimki in števniki) opremljena z obrazili, ki so pri posamezni besedi lahko uporabljena pri pregibih (sklanjanju ali spreganju) lahko rečemo, da slovar SLON-13 vsebuje približno 2,000.000 različnih slovenskih besed.

Slovar programa SLON-13 je »živ« in ga sproti dopolnjujemo in popravljamo. Kot vsaka spreminjajoča se datoteka, tudi slovar SLON-13 vsebuje napake, ki jih takoj, ko so odkrite popravimo. S tem se nenehno nadgrajuje in izpopolnjuje. **Sporočila o morebitnih napakah v slovarju sprejemamo na elektronsko pošto:**

jure.zupan@ki.si

Za vsa opozorila in nasvete smo uporabnikom zelo hvaležni. Vse nove oz. popravljene verzije slovarja bodo ob nadomestilu minimalnih stroškov za pošiljanje in zgoščenko na voljo.

1. Namen programa SLON-13

Za intenzivno računalniško obdelavo besedil, med katere sodi stavčna analiza in z njo povezano strojno razumevanje konteksta, je potrebna zmogljiva računalniška podpora. Programski paket SLON-13 je narejen v okviru projekta »*Linearno in ne-linearno modeliranje kompleksnih podatkov za določanje konteksta*«, ki ga financira MVZT preko pogodb št: P1-017 in J7-0382.

Glavni cilj projekta »*Linearno in ne-linearno modeliranje kompleksnih podatkov za določanje konteksta*« je pokazati možnost, kako bi se dalo kompleksne podatke, ki jih sestavlja zaporedje črk (npr. dolge verige DNA ali poljuben tekstovni material, v našem primeru gre za slovenska besedila) opisati z geometrijsko predstavo v 4-dimenzionalnem (ko gre za verige DNA) ali 3-dimenzionalnem prostoru (ko gre za besedila).

Osnovna hipoteza tistega dela projekta, ki se nanaša na slovenska besedila, in v okviru katerega je bil narejen programski paket SLON-13, je v tem, da se da vsak stavek razstaviti na nekaj enot (preproste stavke samo na eno enoto), ki so predstavljive v 3-dimenzionalnem prostoru, katerega koordinatne osi so osebki, povedki in predmeti. To je t.i. O(sebek)-P(ovedek)-P(redmet) ali na kratko OPP-koordinatni sistem.

Za dokaz omenjene hipoteze je potrebno zasnovati in razviti zahtevne in napredne računalniške sisteme, ki zmorejo izvajati različne vrste analiz besedil v slovenskem jeziku. Programske opreme, ki jo potrebujemo na tržišču ni. Zato v omenjenem projektu predvidevamo izdelavo lastne programske opreme. Ta naj bi bila izdelana v štirih korakih (fazah).

- Korak A. programsko orodje, ki omogoča črkovne analize, statistiko, besedno analizo in ostale enostavnejše postopke s posameznimi besedami (npr. iskanje konkordanc v vseh besednih oblikah),
- Korak B programsko orodje, ki omogoča izdelavo, širjenje in vzdrževanje hierarhično povezane baze **pomenov** slovenskih besed,
- Korak C programsko orodje, ki s pomočjo orodij dobljenih v prvih dveh korakih vrši stavčno analizo in izdelava 3-D predstavitev posameznih stavkov in celotnih besedil v OPP-prostoru in
- Korak D programska oprema, ki izvaja primerjave in iskanja po veliki količini besedil predstavljenih in shranjenih v standardiziranem v OPP-prostoru. Standardizacija OPP prostora je poseben problem, ki v projektu J7-0382 ni zajet.

Program SLON-13 je rezultat končanega prvega koraka A. Omogoča osnovno statistično analizo besedila (pogostosti črk, besed, parov in tripletov črk), predstavitev besedil ali delov besedil s pogostostjo standardnih 25 parov ali tripletov črk, besedno analizo, konkordančna iskanja po poljubnih slovenskih tekstih s poljubnimi in gibljivimi gesli ter iskanje rim oz. besed s poljubno predpisanimi črkovnimi vzorci.

2. Zasnova in shema programa SLON-13

Program SLON-13 je sestavljen iz petih modulov in slovarja. Moduli in slovar so povezani z lupino s pomočjo katere uporabnik izbira besedila, postavlja pogoje, vrste iskanj, analiz in omejitve ter izbira datoteke na katere bo rezultate izpisal oz. shranil.

Program ne zahteva posebnega inštalacijskega postopka ali računalniškega znanja za poganjanje. Kljub temu je na koncu Navodil kratek opis prepisovanja datotek iz distribucijske zgoščenke na lasten računalnik in v poseben zavihek (direktorij). Priporočamo ime SLON

Program poženete s tem, ko kliknete na ikono s sliko slona. Ko program poženete, se pokaže »Glavno okno«. (Slika 1)



Slika 1. Glavno okno programa SLON-13

V glavnem oknu je na voljo pet gumbov s posameznimi opcijami. Opcije so združene v dve skupini. Prva skupina se imenuje Analize, druga Slovar.

Pri obeh opcijah Analize (Analiza črk in Besedna analiza) je potrebno izbrati datoteko s slovenskim besedilo na kateri želimo analizo izvajati. Dobljene rezultate lahko shranite za kasnejšo obdelavo.

Datoteke, ki jih obdelujete morajo biti t.i. »tekstovne« oz. ASCII datoteke. Če je besedilo za Analizo shranjeno v MS Word® formatu, ga je potrebno z ukazom »Shrani kot« (»Save as«) pretvoriti in shraniti v *.txt obliki. V tej obliki so vse datoteke uporabne za delo s programom SLON-13.

Program končate s klikom na rdeč kvadrat s križcem, ki je zgoraj desno

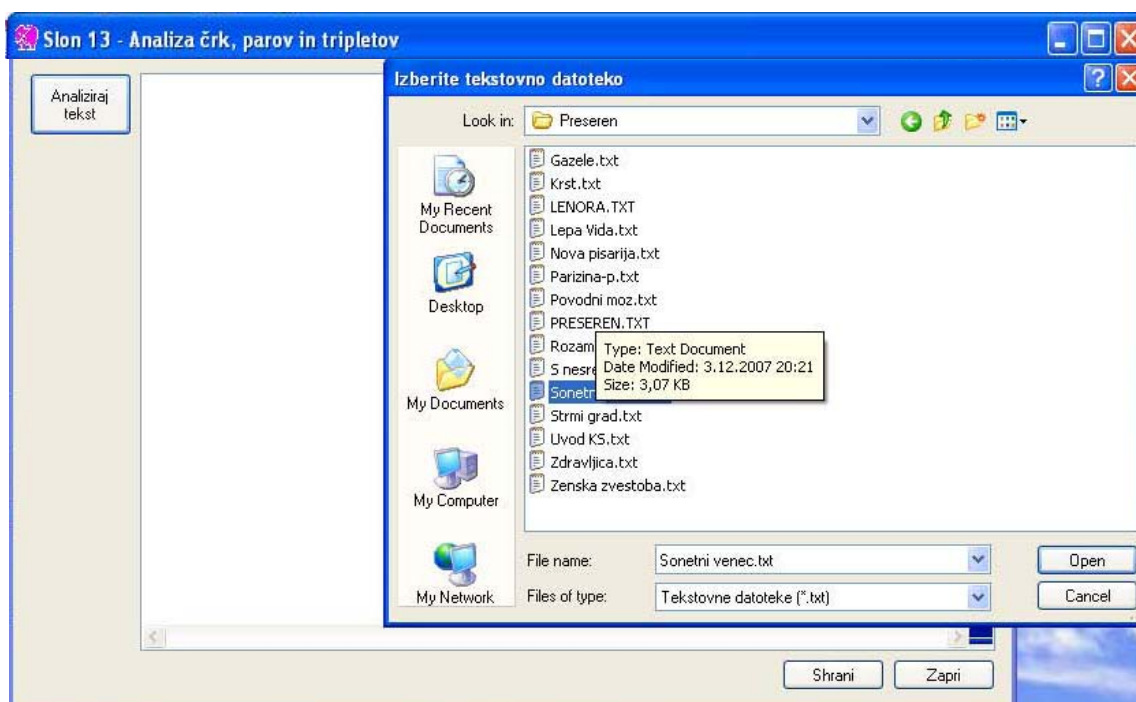


3. Posamezne opcije programa SLON-13

3.1 Analiza posameznih črk, parov in tripletov črk.

Ko izberete opcijo »Analiza, črk, parov in tripletov« se odpre okno (slika 2), ki poleg velikega notranjega okna v katerem se bodo izpisali rezultati analize, ponudi na voljo tri gumbе:

- "Analiziraj tekst"
- "Shrani" in
- "Zapri"



Slika 2. Okno pri analizi črk, parov in tripletov črk.

Ko kliknete gumb "Analiziraj tekst" se odpre komunikacijsko okno MS Windows® v katerem lahko izberete željeno tekstovno datoteko (podaljšek ".txt"). V primeru na sliki 2 je bila izbrana datoteka »Sonetni venec.txt« v zavihku (mapi ali direktoriju) »Preseren«.

Takoj, ko izbrano datoteko potrdite z gumbom "Odpri" (ali Open), prične program SLON-13 z analizo, ki je navadno, če besedilo ni predolgo, narejena v nekaj sekundah.

Rezultate lahko pregledate v notranjem oknu s pomikanjem obeh drsnikov, ki sta na desni in na spodnji strani notranjega okna.

Podatki, ki jih dobite pri črkovni analizi so:

1. Tabela z absolutnim številom vseh parov črk. V tej tabeli so tudi:
 2. podatki o absolutnem in relativnem številu posameznih črk,
 3. podatek o številu presledkov,
 4. pogostost kombinacij presledka z vsemi črkami bodisi pred ali za črko,
 5. pogostosti besed po njihovih dolžinah, itd.
-
6. urejen seznam 50 najpogostejših parov v pregledanem tekstu,
 7. absolutne in relativne pogostosti 25 standardnih parov črk(**),
-
8. tabela 50 najpogostejših tripletov vrste <abc> (navaden triplet),
 9. tabela 25 standardnih tripletov vrste <abc> (***),
 10. tabela 50 najpogostejših tripletov vrste <ab > (končnice besed),
 11. tabela 25 standardnih tripletov vrste <ab > ,
 12. tabela 50 najpogostejših tripletov vrste < ab> (začetki besed),
 13. tabela 25 standardnih tripletov vrste < ab> ,
 14. tabela 50 najpogostejših tripletov vrste <a b> (povezava 2 besed),
 15. tabela 25 standardnih tripletov vrste <a b> ,
 16. tabela 10 najpogostejših tripletov vrste < a > (besede z 1 črko),
 17. tabela 25 standardnih tripletov vrste < a > .

20	b l, z	66	0.97	45	m, i	42	0.62	20	k, a	22	0.32	20
21	i, l	62	0.91	46	b l, r	42	0.62	21	l, e	137	2.01	21
22	e, l	59	0.87	47	a, s	41	0.60	22	t, e	33	0.48	22
23	i, h	59	0.87	48	l, j	41	0.60	23	i, l	62	0.91	23
24	r, a	58	0.85	49	t, i	41	0.60	24	r, i	24	0.35	24
25	s, e	58	0.85	50	e, s	40	0.59	25	l, i	31	0.45	25

št. vseh znakov	=	6815,	št. črk brez blankov	=	5577
št. vseh blankov	=	1238,	št. vseh besed	=	1238
št. vseh parov	=	6815			

št. različnih parov	=	344,	% rzl.parov(od 625)	=	55.04
št. prvih 25 parov	=	2604,	% prvih 25 parov	=	38.21
št. prvih 50 parov	=	3760,	% prvih 50 parov	=	55.17
št. 25 stnd. parov	=	1310,	% 25 stand. parov	=	19.22

=====

50 oziroma 25 najbolj pogostih tripletov tipa: >abc<
Skupaj vseh : 3144, Skupaj standardnih: 154

=====

1	423	>ale<	52	16.54	1	>ila<	2	0.64
2	10663	>ile<	40	12.72	2	>bil<	13	4.13
3	10599	>ije<	39	12.40	3	>pri<	7	2.23
4	11594	>jih<	31	9.86	4	>pre<	12	3.82

Slika 3. Izsek iz izpisa podatkov opcije »Analiza črk, parov in tripletov«

Ukaz »Shrani«

Podatkov pri analizi črk, parov in tripletov je veliko. Uporabnik si jih lahko ogleda s premikanjem drsnikov nogtranjega okna, lahko pa jih shrani na poljubni datoteki, ki jo izbere ali imenuje, ko klikne na gumb »Shrani«. Ukaz »Shrani« odpre iskalno okno za izbor imena

in zavihka (mape, direktorija) v kateri želite shraniti rezultate. Rezultate lahko shranite ali pa tudi ne.

Ko kliknete ukaz "Zapri", program SLON-13 opozori, da lahko dobljeni rezultat shranite. V primeru, da tega ne želite storiti, kliknete »Ne« (No) in podatki bodo v naslednjem koraku izbrisani. Če jih želite ponovno dobiti, morate analizo ponoviti.

Iz opcije izstopite s pritiskom na gumb »Zapri«

Splošno o izboru parov in tripletov

Pari in tripleti črk so obravnavani tako, da je vključen tudi presledek pred prvo črko začetka teksta (oz. odstavka) čeprav ni nujno, da se tak presledek v tekstu dejansko nahaja.

Npr.: Pri besedilu, ki se začne in konča z besedo "mati", računalnik kot pare črk določi pet parov:

> m<, >ma<, >at<, >ti< in >i <

in štiri triple:

> ma<, >mat<, >ati< in >ti <

Dodan je "navidezni" presledek pred besedo, kot bi ga beseda "mati" imela, če bi stala v sredini besedila. Pri skupnem številu vseh presledkov šteje v takem primeru le en presledek - tisti, ki stoji za besedo.

Med besedami program SLON-13 ne šteje dejanskih (fizično prisotnih) presledkov (blankov) ampak zreducira vsa ločila, narekovaje, oklepaje in poljubno število presledkov, ki se drže skupaj in stoje med dvema besedama, na en sam presledek. Na koncu besedila, če se le-to konča brez ločila, doda še "navidezni" presledek. Zato je število upoštevanih presledkov enako številu besed.

Primer: v besedilu dveh besed:

"tisti, ki"

določi program SLON-13 osem (8) parov črk:

> t< >ti< >is< >st< >ti< >i < > k< >ki >i <

sedem (7) tripletov:

> ti< >tis< >isti< >ti < >i k< > ki< >ki <

sedem (7) črk in dva (2) presledka.

Vejica, podobno kot tudi druga ločila, ni všteta. Na začetku in na koncu besedila sta dodana "navidezna presledka". Poleg pravega presledka za vejico, je štet tudi presledek na koncu stavka.

Pri analizi dolgih besedil so opisane malenkosti skorajda nepomembne, pri analizi in primerjavi krajših besedil, pa je potrebno poznati postopek in način določanja posameznih pogostosti.

Vse navedene tabele izračuna program SLON-13 avtomatično, brez poseganja uporabnika takoj, ko izbrano datoteko potrdite.

Glede na dejstvo, da SLON-13 poda vsakič pogostosti 50. najpogostejših parov in tripletov črk izbranega besedil, ki se razlikujejo od pogostosti standardnih 25 parov ali tripletov, lahko uporabnik izbere za primerjavo drugačno »standardno zaporedje«, vendar mora izbiro vsakič sproti opraviti ročno.

V primeru, da želite redno uporabljati drugačna standardna zaporedja parov ali tripletov (ali obojega) od tistih, ki so vgrajena v program SLON-13, se lahko za vgraditev novih »standardnih« zaporedij oz. za spremembo programa SLON-13 obrnete na naslov:

jure.zupan@ki.si

** izbor 25-ih standardnih parov slovenskega izrazoslovja, je bil določen na korpusu 6 milijonov besed. Ta je vseboval besedila različnih žanrov (poezija, romanopisje, žurnalistika, zakonodaja...)

*** izbor 25-ih standardnih tripletov v vseh petih kombinacijah:

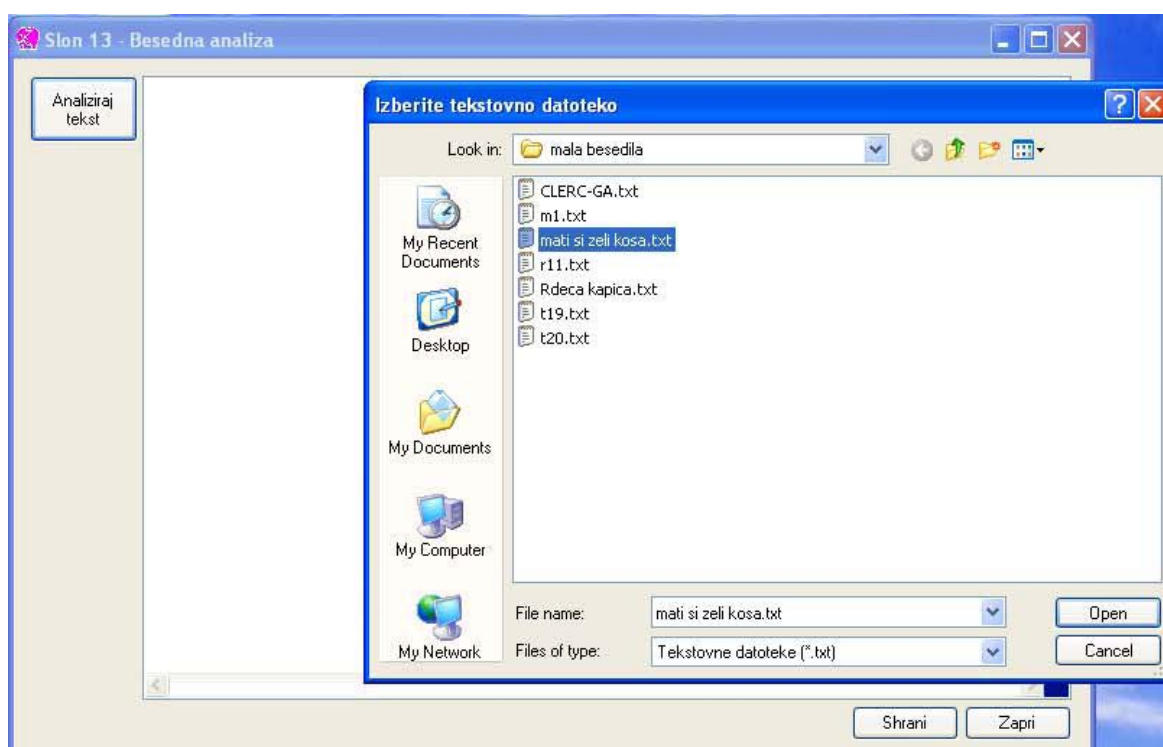
>abc<
>ab <
> ab<
>a b< in
> a < ,

prav tako uporabljen opciji »Analiza črk in parov«, je bil dobljen na istem korpusu besedila kot izbor standardnih parov.

3.2 Besedna analiza

Podobno kot opcija "Analiza črk" tudi opcija "Besedna analiza" ponudi v glavnem oknu tri gumbе z ukazi:

- "Analiziraj tekst",
- "Shrani" in
- "Zapri"



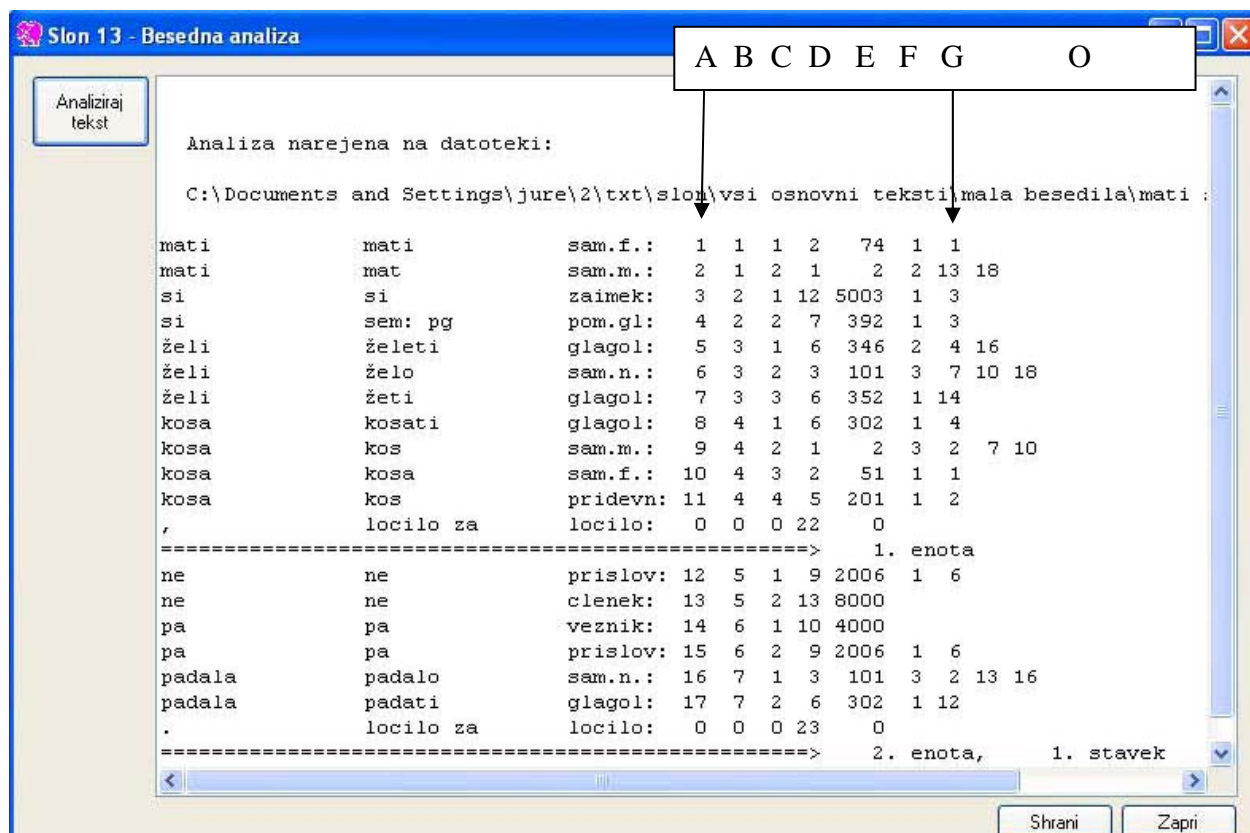
Slika 4. Izbor datoteke na kateri je besedilo, ki ga želimo analizirati z opcijo »Besedna analiza«

Ko kliknete "Analiziraj tekst" se odpre komunikacijsko okno MS Windows® v katerem lahko izberete poljubno tekstovno datoteko s podaljškom ".txt". V primeru na sliki 4 je bila izbrana datoteka z imenom »*mati se zeli kosa.txt*« v zavihku (mapi ali direktoriju) »*mala besedila*«.

Takoj, ko izbrano datoteko potrdite z gumbom "Odpri" (Open) program SLON-13 prične z analizo, ki je navadno, če besedilo ni predolgo, narejena v nekaj sekundah. Od trenutka, ko izberete datoteko, dela program samostojno. V času analize tečejo preko zaslona številke, ki povedo, koliko stavkov je bilo prvkar analiziranih.

Rezultate »Besedne analize« lahko pregledate v notranjem oknu s pomikanjem drsnikov na obeh straneh okna.

Kot primer rezultatov je na sliki 5 prikazana analiza, ki jo program SLON-13 naredi, če izberete datoteko z imenom »*mati si zeli kosa.txt*«. Ta datoteka, ki vsebuje en sam stavek, je priložena na zgoščenki s programom SLON-13.



Slika 5. Rezultat »Besedne analize« stavka "Mati si želi kosa, ne pa padala.", ki je na datoteki »mati si zeli kosa.txt«. Za razlago so stolpci od A do O, ki jih v originalnem zapisu ni, označeni posebej.

Opcija "Besedna analiza" razdeli analizirano besedilo na stavke, stavke pa na enote, ločene z ločili. Rezultati analize so navedeni za vsako besedo besedila v eni vrstici. Če ima beseda več možnih slovničnih interpretacij, je vsaka možnost navedena v svoji vrstici. Celotno analizirano besedilo si lahko predstavljamo kot dolgo preglednico v kateri so za vsako besedo besedila v posamezni vrstici navedeni po stolpcih različni podatki (slika 5).

Prvi stolpec navaja vsako besedo tako, kot je zapisana v besedilu. V drugem so vsa slovarska gesla, ki jih za določeno obliko besede računalnik najde: npr.: pri besedi "mati" sta dve možni gesli: "mati in "mat", pri besedi "želi" pa tri: "želeti", "žeti" in "želo".

V tretjem stolpcu je določena besedna vrsta možnega gesla za besedo iz besedila (1. stolpec).

Preglednica 1. Tretji stolpec navaja besedno vrsto slovarskega gesla:

Besedna vrsta	Številka	Besedna vrsta	Številka
Samostalnik	1-6	Prislov	12
Pridevnik	7	Predlog	14
Glagol	8	Členek	15
Pomožni glagol	9	Medmet	16
Zaimek	10	Kratica	17
Števnik	11	Neznano	19

Tretjemu stolpcu sledi lahko več (tudi do petnajst) stolpcev števil, ki od osmega (F) naprej, niso nujno vsi izpolnjeni:

- A 4. stolpec: zaporedna številka besede v stavku, vključno z njenimi različicami gesel,
- B 5. stolpec: zaporedna številka besede v stavku originalnega besedila,
- C 6. stolpec: slovarska različica (gesla) besede iz besedila,
- D 7. stolpec: numerična oznaka besedne vrste gesla (glej stolpec 3),
- E 8. stolpec: številka tabele z obrazili za dano geslo (nad 1000 so nepregibne besede),
- F 9. stolpec: število možnih slovničnih oblik z obrazilom, ki ga ima beseda iz besedila,
- 10-18. stolpec: zaporedne številke mest, na katerih se uporabljeno obrazilo nahaja v tabeli obrazil, navedeni v 8. stolpcu.

Primer 1: pri besedi "mati" sta možni dve gesli: "mati", sam. ž. in "mat", sam. m.

- a) Ženski samostalnik "mati" ima obrazila shranjena v 74. tabeli. Obliki besede iz besedila ustreza obrazilo "-i", ki je v tabeli 74. na mestu št. 1. Obrazilo na 1. mestu ustreza 1. sklonu ednine.
- b) Moški samostalnik "mat", z obrazili v tabeli št. 2. Obrazilo "-i" je v tabeli št. 2 na 13. in 18 mestu. Mest št. 13 in 18. pomenita 1. in 6. sklon množine ($13 = 6+6+1$ oz. $18 = 6+6+6$).

Primer 2: Pri besedi "želi" so možna tri gesla:

- a) glagol "želeti", tabela 346: 4. in 16. obrazilo "-i". Četrto mesto glagolskih obrazil (št. 4) vedno pomeni 3. osebo ednine, št. 16 pa veleno obliko glagola,
- b) samostalnik srednjega spola "želo", tabela 101: obrazilo "-i" je na 7., 10. in 18. mestu, kar pomeni: 1. in 4. sklon dvojine ($7=6+1$, $10=6+4$) in 6. sklon množine ($18=6+6+6$),
- c) glagol "žeti", tabela 352: obrazilo "-li" na 14. mestu se uporablja za tvorjenje velelnika in pri tvorjenju oblik preteklega ali prihodnjega časa.

Primer 3: Beseda "kosa" nudi tri možna gesla:

- a) glagol "kosati": tabela 302: 4. obrazilo (ki ga ni) pomeni 3. osebo ednine; v tabeli 302 je 4. mesto prazno.
- b) samostalnik moškega spola "kos": tabela 2; obrazilo "-a" na 2., 7., 10. mestu, ki pomenijo 2. sklon ednine in 1. in 4. sklon dvojine ($7=6+1$) in ($10=6+4$),
- c) samostalnik ženskega spola "kosa": tabela 51, obrazilo "-a" na 1. mestu pomeni 1. sklon ednine.

Primer 4: Beseda "ne" je:

- a) prislov: tabela 2006, šestica na koncu oznake »2006« pomeni nedoločen prislov (prislovi imajo oznake med 2001 in 2006, glede na vrsto prislova),
- b) členek: tabela 8000, vsi členki imajo oznako 8000.

Primer 5: Beseda "pa" je :

- a) veznik: tabela 4000, vsi vezniki imajo oznako 4000,
- b) prislov: tabela 2006; šestica na koncu oznake pomeni nedoločen prislov.

Primer 6: Besedna oblika "padala" lahko nastane iz dveh gesel:

- a) iz samostalnika srednjega spola "padalo": tabela 101, obrazilo "-a" na 2., 13. in 16. mestu pomeni 2. sklon ednine in 1. in 4. sklon množine ($13=6+6+1$ in $16=6+6+4$)
- b) iz glagola "padati": tabela 302, obrazilo "-la" na 12. mestu je za povezavo glagola z ženskimi samostalniki v ednini in samostalniki moškega in srednjega spola v dvojini – vse v preteklem ali prihodnjem času.

Mesta najrazličnejših obrazil, ki jih navajajo stolpci od G do O (od 10 do 18), so pomembna za avtomatično oz. strojno razumevanje besedila in predvsem za nadaljevanje strojne analize konteksta. V nadaljevanju so posamezne številke, ki jih najdemo v stolpcih, razložene še bolj podrobno.

Samostalniška obrazila:

Vedno jih je 18, za vsako slovnično število po šest, kolikor je tudi sklonov. Številke od 1 do 6 so skloni v ednini, številke od 7 do 12 in 13 do 18 predstavljajo sklone v dvojini oz. množini. Ker ima več sklonov lahko enaka obrazila, so v "Besedni analizi" pri vsaki besedi navedeni vsi možni skloni.

Tako je tudi pri množinskih samostalnikih. Le da je v slednjem primeru prvih dvanjast obrazil "praznih". Številka obrazila dejansko pove sklon in število za katerega velja.

Pri drugi ženski sklanjatvi na "-i" je devet sklonov z enakim obrazilom. Zaradi tega je ob ženskih samostalnikih 2. sklanjatve možnih devet števil. Primera te sklanjatve sta samostalnika "misel" in "jesen", navedeni položaji obrazil oz. ustrezni skloni z obraziloma "-li" in "-i", so: 2 (2. sklon), 3 (3.), 5 (5.), 7 (1. dvoj.), 8 (2.), 10 (4.), 13 (1. množ.), 14 (2.), in 16 (4.).

Pridevnška obrazila:

Pridevniki imajo 12 možnih različnih obrazil. V nobenem nizu pridevniških obrazil ni dveh enakih, kot se to lahko zgodi pri samostalniških in glagolskih obrazilih. Po vrstnem redu so obrazila navedena tako, kot kaže preglednica 2. V njej so navedene so tudi povezave pridevnikov s samostalniki na katere se pridevnik v različnih oblikah nanaša.

Preglednica 2. Zaporedje obrazil (1 do 12) v pridevniških tabelah programa SLON-13, ki se vežejo z določenimi skloni (v preglednici št. 1 do 6) samostalnikov različnih spolov (primer je vzet iz tabele obrazil št. 201)

Mesto	Moški spol			Ženski spol			Srednji spol			Obrazilo	Primer
	Ednina	Dvoj.	Množ.	Ednina	Dvoj.	Množ.	Ednina	Dvoj.	Množ.		
1	1 4									-brez	lep
2		1 4		1						-a	lepa
3			4	2		1 4				-e	lepe
4			1	3 5	1 4			1 4		-i	lepi
5				4 6			1 4			-o	lepo
6	5						5			-em	lepem
7	6		3			3	6		3	-im	lepim
8		2 5	2 5		2 5	2 5		2 5	2 5	-ih	lepih
9	2 4						2			-ega	lepega
10	3						3			-emu	lepemu
11		3 6			3 6			3 6		-ima	lepima
12			6			6			6	-imi	lepimi

Kot primer vzemimo obrazilo "-i", ki je v 4. vrstici preglednice 2. Številke 1, 3, 5, 1, 4, 1, 4, iz te vrstice povedo v kombinaciji z informacijo v stolpcih, da se pridevnik z obrazilom "-i" veže s samostalniki moškega spola v 1. sklonu množine, s samostalniki ženskega spola v 3.

in 5. sklonu ednine in 1. in 4. sklonu dvojine ter s samostalniki srednjega spola v 1. in 4. sklonu dvojine. Povsem na enak način so razporejena obrazila oz. njihova mesta v vseh tabelah pridevniških obrazil.

Glagolska obrazila:

Glagoli imajo v vseh tabelah spregatev 21 obrazil. Mesto obrazila v preglednici 3 pove, kako se glagoli v različnih časih in vidih vežejo z različnimi oblikami drugih besed.

Preglednica 3. Zaporedje obrazil v glagolskih tabelah programa SLON-13. Številke v drugem stolpcu: 1, 2 in 3 pomenijo prvo (jaz), drugo (ti) in tretjo (on, ona, ono) osebo v kateremkoli številu, oznake m, ž in s se nanašajo na moški, ženski in srednji spol. Navedena obrazila so vzeta iz glagolske tabele št. 301.

Zaporedje obrazila	Oseba	Spol	Število	Čas	Vid / naklon	Obrazilo	Primer
1					infinitiv	-ti	puli-ti
2	1.	m ž s	ednina	Sedanjik	želelni (naj)	-m	puli-m
3	2.	m ž s	ednina	Sedanjik	želelni (naj)	-š	puli-š
4	3.	m ž s	ednina	Sedanjik	želelni (naj)	-	puli-
>> 5	1.	m ž s	dvojina	sedanjik	želelni (naj)	-va	puli-va
6	2. 3.	m ž s	dvojina	sedanjik	želelni (naj)	-ta	puli-ta
7	1.	m ž s	množina	sedanjik	želelni (naj)	-mo	puli-mo
8	2.	m ž s	množina	sedanjik	želelni (naj)	-te	puli-te
> 9	3.	m ž s	množina	sedanjik	želelni (naj)	-jo	puli-jo
> 10	3.	m ž s	množina	sedanjik	želelni (naj)	301 nima te oblike	
11	1. 2. 3.	m	ednina	prih/prej	pogojni (bi)	-l	puli-l
12	1. 2. 3.	m	dvojina	prih/prej	pogojni (bi)	-la	puli-la
12	1. 2. 3.	ž	ednina	prih/prej	pogojni (bi)	-la	puli-la
12	1. 2. 3.	s	množina	prih/prej	pogojni (bi)	-la	puli-la
13	1. 2. 3.	ž	množina	prih/prej	pogojni (bi)	-le	puli-le
14	1. 2. 3.	m	množina	prih/prej	pogojni (bi)	-li	puli-li
14	1. 2. 3.	ž	dvojina	prih/prej	pogojni (bi)	-li	puli-li
14	1. 2. 3.	s	dvojina	prih/prej	pogojni (bi)	-li	puli-li
15	1. 2. 3.	s	ednina	prih/prej	pogojni (bi)	-lo	puli-lo
16	2.	m ž s	ednina		velelnik	-	puli-
>> 17	1.	m ž s	dvojina		velelnik	-va	puli-va
18	2.	m ž s	množina		velelnik	-ta	puli-ta
19	1.	m ž s	dvojina		velelnik	-mo	puli-mo
20	2.	m ž s	množina		velelnik	-te	puli-te
21					trpnik	-t	puli-t

Kot primer vzemimo obrazilo "-va", ki je v 5. in v 17. vrstici. Obrazila, ki so v 5. vrstici glagolskih tabel povedo, da glagol s tem obrazilom tvori 1. osebo v dvojini sedanjika ali želelne oblike (naj). Glagoli z obrazili 17. vrstice (v tem primeru je na tem mestu enako obrazilo -»va«), so v prvi osebi dvojine in se v velelniku vežejo s samostalniki vseh spolov.

Ni nujno, da imajo vsi glagoli na 5. (sedanjik, dvojina) in 17. mestu (dvojina, velelnik) enki obrazili. Primera za to sta:

glagol: sešte-ti → sešte-jeva (5), sešte-jva,(17) in
glagol: tk-ati → tk-eva (5) in tk-iva(17).

Glagolsko obrazilo v 10. vrstici je alternativa obrazilu v 9. vrstici. Obe obrazili predstavljata obliko glagola v 3. osebi množine sedanjika. Alternativno obliko nekateri glagoli imajo, drugi pa ne.

Primer z obema oblikama je glagol "žel-eti" z oblikama "žel-ijo" in "žel-e". Galgol "puli-ti", pa alternativne oblike nima (namesto "oni pul-ijo" ne smemo reči "oni pul-e").

3.3 Konkordance in konkordančno iskanje

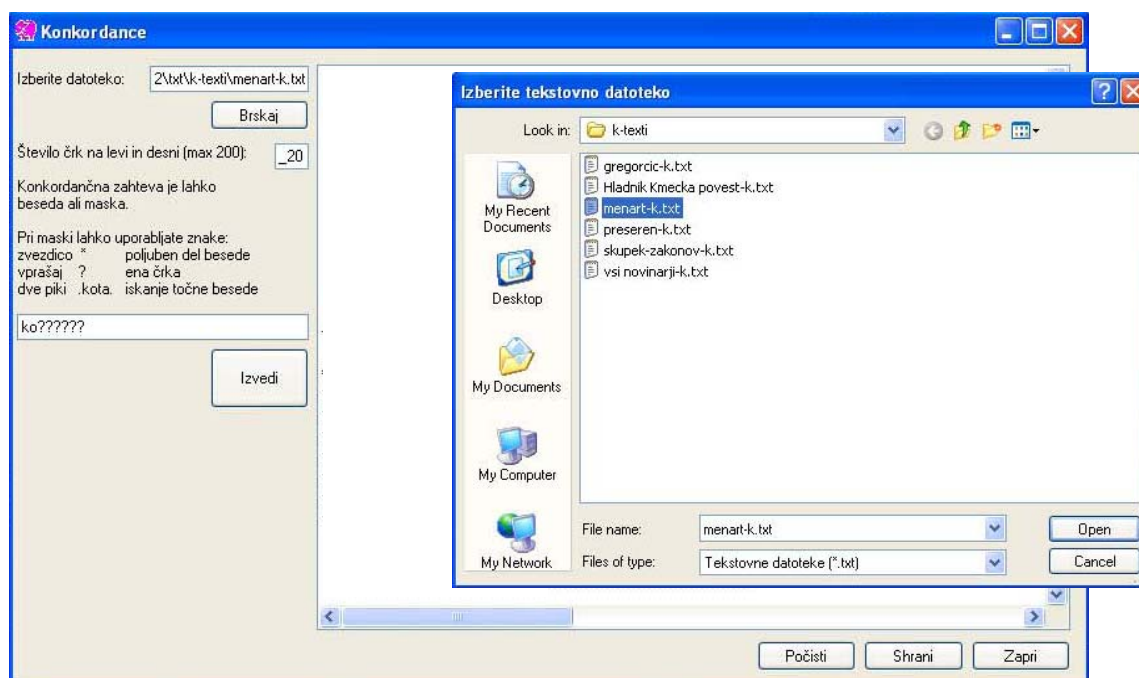
Opcija "Iskanje Konkordanc" ima nakoliko zahtevnejšo izbiro, saj pred iskanjem po izbrani datoteki, ki jo dobite z ukazom

- »Brskaj«

zahteva več podatkov o celotnem postopku in izpisu rezultatov kot opcija "Analiza".

Ukaz »brskaj« se od ukaza »Analiziraj tekst«, ki se uporablja pri opcijah analize, razlikuje v tem, da po izboru datoteke še ne sproži iskanja, ampak vrne uporabnika v meni k izbiri ostalih podatkov, ki so (slika 4):

- »Število črk na levi in desni« ali konkordančna širina
- »Konkordančna zahteva« ali konkordančna beseda
- »Počisti«
- »Shrani«
- »Zapri«



Slika 6. Okno pri opciji »Konkordance«. Za iskanje konkordančnih delov besedila je bila izbrana datoteka s pesmimi Janez Menarta »menart-k.txt« iz zavihka »k-teksti«.

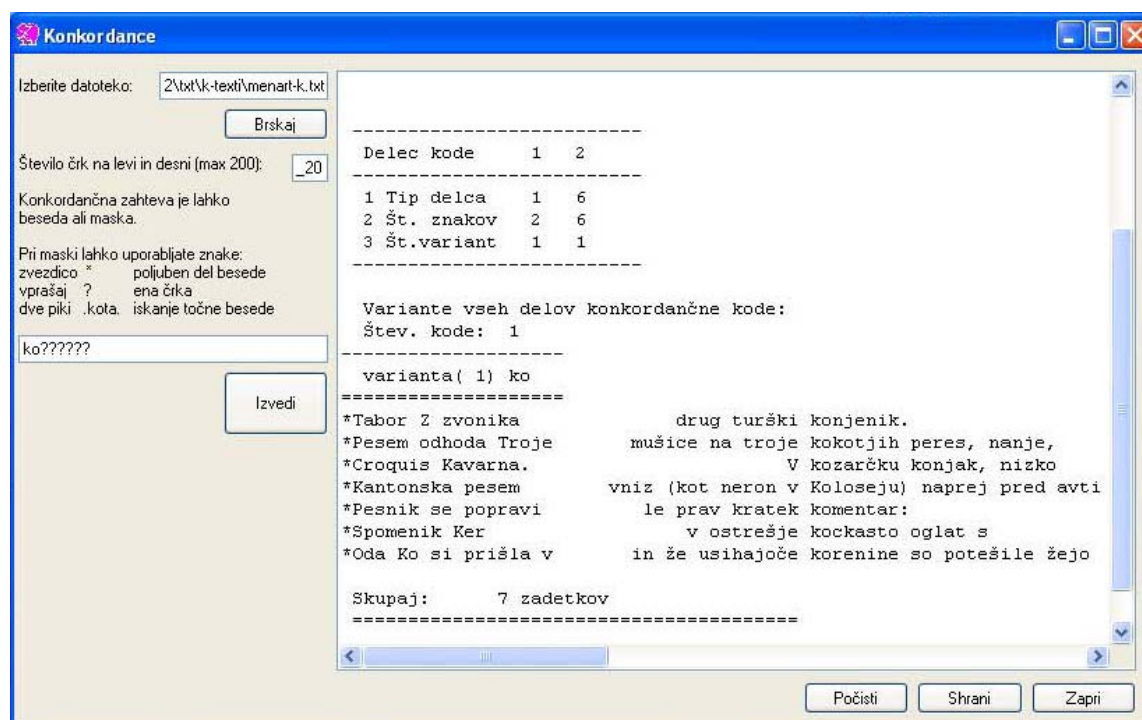
Konkordančno širino navedete s številom črk, ki jih želite izpisane levo in desno od iskane oz. najdene konkordančne besede. Večja kot je konkordančna širina, več besedila se ob vsaki najdeni konkordanci izpiše na zaslon. Zgornja omejitev izpisa je cel stavek v katerem se iskana beseda nahaja. Več kot celega stavka program SLON-13 pri konkordančnih zadetkih ne izpiše nikoli, ne glede na to, kako dolgo konkordančno širino predpišete. Izpis besedila ob konkordančnih besedah se ustavi ob presledkih med besedami, ki so zahtevani konkordančni širini najbližje.

Najmanjša konkordančna širina je "0", ki pomeni izpisovanje samo konkordančne besede oz. besede v kateri se nahaja konkordančna zahteva (Definicija konkordančne zahteve je v nadaljevanju).

Ker so stavki v besedilih lahko precej dolgi, je pri začetnih iskanjih priporočljivo omejiti konkordančno širino na največ 30-50 znakov in jo, če se pokaže potreba, zvečati šele pri naslednjih iskanjih.

Po koncu iskanja program SLON-13 sporoči število najdenih zadetkov.

Primer rezultatov iskanja z opcijo »Konkordance« kaže slika 7.



Slika 7. Rezultati konkordančnega iskanja po besedilu Menartovih pesmi s konkordančno zahtevo *ko??????*, ki zahteva iskanje vseh besed z začetnima črkama »ko« in dolžino 8 črk. Če uporabnik besedilo, ki ga namerava pregledovati z opcijo »Konkordance« pred analizo opremi z zvezdico (*) na začetku poljubne vrstice, bo 15 črk označene vrstice služilo pri izpisih konkordanc kot **»marker«** pod katerim se nahaja najdena konkordanca. V danem primeru so z zvezdico (*) opremljeni naslovi vseh Menartovih pesmi v datoteki »menart-k.txt«, ki je bila izbrana za pregled (slika 6).

Konkordančna zahteva ali konkordančna beseda

Konkordančna beseda, je skupek črk in znakov s katerim iščete po besedilu, da bi dobili vse tiste stavke (ali njihove dele) v katerem konkordančna beseda nastopa.

Za navedbo konkordančne besede je več možnosti. Iščete lahko besedo:

- v vseh njenih možnih oblikah (sklonih, časih, osebah, itd.) ali
- v natančno enaki črkovni obliki, kot jo napišete ali
- v izbranem gibljivem črkovnem območju.

Primer a) Iskanje s konkordančno besedo "riba" bo dalo tiste dele besedila, ki vsebujejo besedo »riba« v vseh možnih slovničnih oblikah:

riba, ribi, ribe, ribo, rib, ribama, ribah, ribam, ribami

Primer b) Pri izbiri »natančne oblike« dobite vse dele besedila, ki vsebujejo tak črkovni vzorec kot ste ga navedli ne glede na to v kakšnem pokoju se nahaja. Če navedete "riba", program poišče vse zveze tki vsebujejo samo izbrano črkovno zaporedje:

poriba, zariba, ribak, ribariti, ribarjenje, Kotoriba, nariba, tribarven, hribarjev, pribarantati, tribalistov, itd...

Primer c) Tretja možnost je iskanje samo delnega vzorca besede. Npr. vse besede, ki se v izbranem besedilu začnejo (ali končajo) s črkama "ri".

Ker je iskalnih možnosti več, je potrebno iskalni postopek opisati dovooolj natančno, da program izvede tisto, kar želimo. Iskalne zahteve določimo s "konkordančno masko". Konkordančna ali iskalna maska je zapis, ki ob upoštevanju nekaj enostavnih pravil, omogoči natančen opis uporabnikove zahteve in omeji pogoje iskanja po besedilu.

Prvega načina (primer a) sploh ni potrebno izbirati, saj je privzet (default) ob vpisu poljubne besede ali črkovne kombinacije.

Tudi v primeru ko je v polje "Konkordančna zahteva" vpisana nesmiselna črkovna kombinacija, n.pr.:

xw123qB

bo program SLON-13 najprej pogledal ali je beseda v slovarju, da bi, v primeru, če bi jo našel, poiskal vse njene slovnične oblike. Ker besede "xw123qB" v slovarju ni, bo iskal po besedilu samo s podano črkovno kombinacijo. Če vpišete pregibno besedo, ki jo slovar SLON-13 pozna, bo iskal po besedilu z **vsemi** njenimi slovničnimi oblikami.

Drugi način iskanja, iskanje z natančno določenim črkovnim naborom, se navadno uporablja v kombinacijah z znakoma ? in * in besedami, ki se pregibajo (sklanjajo ali spregajo). Če beseda, ki jo želimo iskati ni pregibna, bo računalnik tako in tako iskal samo z obliko, ki smo jo navedli.

Besedo, ki jo želimo iskati v natančno izbrani obliki napišemo v polju "konkordančna zahteva" med dve **omejitveni piki**:

riba.

Vendar je treba opozoriti, da v primeru, ko bi v polje "konkordančna zahteva" vpisali samo zgornjo obliko (.riba.), ne bi dobili skoraj nobenega zadetka. SLON-13 bi izpisal le tiste stavke v katerih se nahajajo črke "riba" med dvema presledkoma ali ločiloma, npr.:

"Riba riba, (riba- riba? (riba) »riba«

in podobne.

Če ni dodatnih zahtev, program SLON-13 išče samo bedice, ki vsebujejo natanko črkovno kombinacijo med dvema pikama (npr.: »riba«) in nobene druge črke. Slon-13 vse znake, ki niso črke ali številke, obravnava kot na presledke.

Zato v primerih, ko želimo dobiti vse dele besedila z besedami, ki vsebujejo izbrano črkovno kombinacijo (npr.: "riba"), ne glede na to kako so dolge in kje se črkovna kombinacija nahaja, mora biti konkordančna zahteva oblikovana drugače:

.riba.

Zvezdica (asterisk) v konkordančni zahtevi pomeni kakšenkoli del besede ob črkovnem delu. Pri uporabi je treba biti zelo pazljiv, da dobimo tak rezultat, kot ga želimo. Npr.: zahteva:

riba

brez omejitvenih pik(.) pri besedi "riba", bi poleg besed navedenih v primeru b) (poriba, zariba, ribak, ribariti, ...) dobili tudi vse zveze, ki vsebujejo **katerokoli obliko** besede "ribe" (riba, ribe, ribi, ribo, rib ...), npr.:

- ribogojnica (vsebuje "ribo"),
- ribezelj (vsebuje "ribe"),
- pribit (vsebuje "ribi") in
- tribuna (vsebuje "rib").

Ker je oblika "rib" najkrajša in vsebovana v vseh ostalih slovničnih oblikah besede "riba", program SLON-13 z ostalimi oblikami sploh ne išče, ker so le-te avtomatično vsebovane v vseh najdbah z najkrajšo obliko "rib".

Zvezdico lahko uporabljate samo na eni strani, bodisi desno ali levo od zahtevane črkovne kombinacije:

pri zahtevi *riba → dobite besede kot so: hribi, Gribov, Ribičič, ribah, prib, ribe ,..
pri zahtevi .rib.* → dobite besede, kot so: Ribentrop, ribič, rib, ribicam, ribarjevi, ...

Pogosto se marsikatera najdba zdi uporabniku nenavadna, čeprav je slovnično pravilna. V zgornjem primeru je to lahko besedica "prib", ki se zdi nenavadna, če je ne vidimo v kontekstu. Dejansko je veljavna slovenska beseda, saj predstavlja 2. sklon dvojine ali množine besede "priba" (vrsta ptice).

Za pojasnilo besed, ki se zde nenavadne lahko v programu SLON-13 uporabite opcijo "Iskanje besed". Vpišete »nenavadno« besedo in pogledate rezultat. V zgornjem primeru bi vpisali besedico "prib" in dobili odgovor, da je to samostalnik ženskega spola "priba" v obliki s katero se izražata 2. sklon dvojine in množine.

Pri opisu iskalne zahtere lahko uporabite tudi vprašaj (?), ki pomeni poljubno, vendar obvezno črko. Npr.: konkordančna zahteva:

?? .riba.*

bo kot rezultat iskanja dala vse besede, ki imajo pred obliko .riba.* obvezno dve črki, npr.:

naribati, poribati, naribam, naribaš, naribajte, poribamo, itd.

Uporabo vprašajev v konkordančni zahtevi kaže primer na sliki 8. Konkordančna zahteva je opredeljena s črkama »ko« in s šestimi vprašaji.

ko??????

Rezultat iskanja so vse osem (8) črk dolge besede, ki se prično na »ko«.

V iskalni maski lahko črke, zvezdice in vprašaje med seboj kombinirate na veliko načinov.

Z zvezdicami in vprašaji pomešanimi s kombinacijami črk lahko sestavite kompleksno iskalno masko. Nekaj možnih kombinacij kaže preglednica 4.

Preglednica 4 Možni načini izrazov različnih iskalnih zahtev

Iskalna zahteva	Rezultat iskanja
???????	Besede dolge 7 znakov (kolikor vprašajev toliko črk)
???????*	Besede dolge 7 in več znakov
*a*a*a*a*	Besede, ki vsebujejo črko "a" vsaj štirikrat
pod*.ica.	Besede, ki se začnejo s "pod" in končajo z »ica«. Pomembno je, da je del iskalne kode "ica" podan med omejitvenima pikama. V nasprotnem primeru bi iskali tudi besede, ki se končujejo tudi z drugimi oblikami samostalnika "ica" (torej ica, ice, ici, ic, icami, itd.). Pri besedici "pod" piki nista potrebni. Lahko bi jih napisali, vendar besedici »pod« sledi zvezdica(*), ki implicitno zajame vse oblike samostalnika "pod".
pe??li	Besede, ki vsebujejo zloga "pe" in "li", medsebojno oddaljena natanko dve črki.
*pe??*li*	Besede, ki vsebujejo zloga "pe" in "li", medsebojno oddaljena vsaj dve črki.

Konkordančno iskanje v obliki ali-ali

Opcija "Konkordance" nudi tudi možnost iskanje z naborom več konkordančnih besed skupaj v obliki "ali-ali".

Npr.: V besedilu želimo poiskati besedne zveze v katerih so uporabljene besede:

vino, vince, vinčece, pivo, pivce, medica, žganje, slivovka, pijača

V tem primeru uporabimo na začetku iskalne zahteve znak # (pravimo mu tudi »lojtra«) in zahtevo opredelimo z iskalno masko takole:

vino vince vinčece pivo pivce medica žganje slivovka pijača

Pri opisu zahteve z več besedami je omejitev 30 besed. Besede morajo biti med seboj ločene s presledki. Program ne bo iskal samo besed, ki so navedene v iskalni zahtevi, ampak tudi vse njihove slovnične oblike, kot so (glede na zgornjo zahtevo):

vino, vina, vinom ... pivca pivcu pivcema ... medici, žganju ... pijačah itd.

Teoretično je vseh variant omenjenih devetih besed s katerimi program preiskuje besedilo 9 besed x 18 oblik, to je 162. V resnici je zaradi ponavljajočih se obrazil število različnih variant le nekaj nad 80.

Kot rezultat bo program SLON-13 izpisal vse stavke v katerih se pojavi katerakoli od omenjenih 80 variant.

V iskalno masko lahko navedete le toliko besed, da skupno število pregibnih oblik (vseh variant vseh besed v sklonu, številu in času) ne preseže 200. Omejitvi sta izključujoči:

Ali 30 besed ali toliko besed, da imajo skupno manj kot 200 variant (oblik).

Med posameznimi besedami mora biti natanko en presledek. Presledek mora biti tudi drugi znak v konkordančni zahtevi (za znakom #).

Če želite iskati z natanko določenimi besednimi oblikami, lahko pri opisu besed v iskalni maski uporabljate označevanje s pikama na obeh konceh besede.

Uporaba zvezdic in vprašajev ni dovoljna.

Primer: Iščemo edninske svojilne zaimke v besedilu s konkordančno masko oz. zahtevo:

moj tvoj njegov njen

Program SLON-13 bo po besedilu iskal z vsemi variantami vseh štirih konkordančnih besed:

moj	tvoj	njegov	njen
moja	tvoja	njegova	njena
moje	tvoje	njegove	njene
moji	tvoji	njegovi	njeni
mojo	tvojo	njegovo	njeno
mojem	tvojem	njegovem	njenem
mojim	tvojim	njegovim	njenim
mojih	tvojih	njegovih	njenih
mojega	tvojega	njegovega	njenega
mojemu	tvojemu	njegovemu	njenemu
mojima	tvojima	njegovima	njenima
mojimi	tvojimi	njegovimi	njenimi

Program SLON-13 določi vse variante avtomatično. Uporabnik poda samo zahtevo štirih besed (moj, tvoj, njegov in njen). SLON-13 ugotovi, da je z navedenimi štirimi besedami možno dobiti 48 variant (48 = 4 x 12), kar pomeni, da ni bila presežena nobena od obeh omejitev.

Če program SLON-13 ugotovi več kot 200 variant, išče z 200 variantami sporoči uporabniku, da bo rezultat zaradi presežene omejitve lahko pomankljiv.

Možno iskanje v obliki »ali-ali« je tudi v mešani oblik konkordančnih besed tako, da imajo nekatere besed omejitvene pike druge pa ne:

Primer:

.moj. konj vitez

Po končanem iskanju SLON-13 sporoči število najdenih zadetkov.

Ukaza »Počisti« in »Shrani«

Rezultati posameznih iskanj se v oknu nalagajo drug za drugim. Rezultate lahko shranite na poljubno datoteko z ukazom »Shrani«, ali pa jih izbrišete z ukazom »Počisti«.

Priporočilo: :

Opremite besedila, ki jih želite preiskovati z opcijo »Konkordance« z zvezdicami na začetkih vrstic, ki besedilo dele na pomembne enote (npr. poglavja, posamezne pesmi, vsaka 100 vrstica, itd). Tako označeno besedilo močno olajša kasnejše iskanje in preverjanje ter primerjavo konkordanc.

Vzorec opisanih oznak z zvezdicami si lahko ogledate v datoteki 3pesniki-k.txt, ki je na distribucijski zgoščenki.

3.4 Iskanje besed

Iskanje besed kot jo nudi SLON-13 je specifična in redka oblika slovarske informacije. Namenjena je uporabnikom, ki jim slovenščina ni materin jezik.

Podaja informacijo o **karšnikoli obliki** poljubne slovenske besede. Ko kliknete opcijo "Iskanje besed", se odpre okno z enim vhodnim poljem:

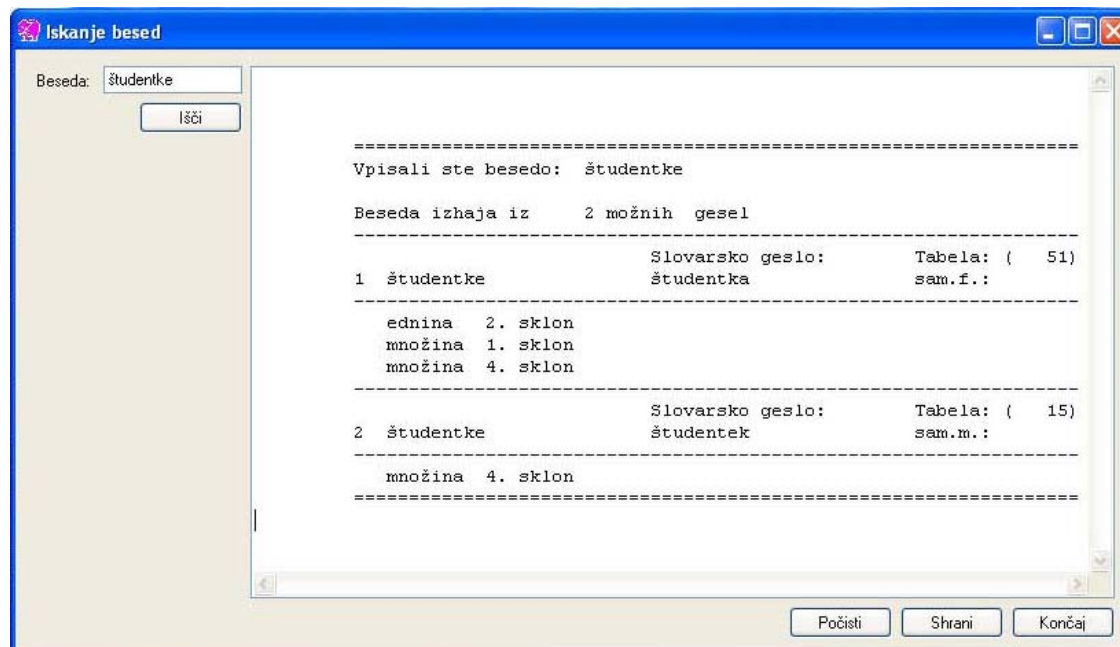
- »Beseda«

enim gumbom

- »Išči«

ter s tremi dodatnimi ukazi, ki smo jih že srečali pri predhodnih opcijah:

- »Počisti«
- »Shrani« in
- »Končaj«



Slika 8. Opcija »Besede« podaja slovnične oblike poljubne slovenske besede iz nabora 160,000 gesel slovarja SLON-13

V polje "Beseda" vpišete izbrano besedo (v slovarju SLON-13 je 160,000 gesel) v poljubni slovnični obliki (sklonu, spolu, času, naklonu). Preogram SLON-13 odgovori z ustreznim slovarskim geslom, ter obliko ali oblikami, če jih je več, v kateri se beseda, ki ste jo vpisali, nahaja.

Ker je slovenščina bogata s homonimi (besedami, ki se enako pišejo a imajo različne pomene), je v odgovoru lahko več slovarskih gesel. Primer na sliki 8 kaže dvojnost pomena besede »študentke«. Besedo lahko izvajamo iz ženskega samostalnika »študentka« ali iz moškega samostalnika »študentek«.

Dva primera:

Če vpišete besedo

"misli"

in kliknete "Išči", bo SLON-13 odgovoril s podatkom, da se beseda »misli« lahko izpelje iz:

1) glagola »misliti« (ta se sprega z obrazili tabele 301), ki se v povezavi z obrazilom »-li« lahko razume, da je:

- a) v 3. osebi ednine ali množine v sedanjiku ob samostalnikih vseh treh spolov oz.
- b) v 2. osebi ednine veledne oblike ob samostalnikih vseh treh spolov.

2) samostalnika ženskega spola »misel«, ki se sklanja po drugi ženski sklanjatvi (z obrazili tabele 75), za katero je značilno da ima v 2., 3., in 5. sklonu ednine, v 1., 2. in 4. sklonu dvojine ter množine enaka obrazila, kar pomeni, da bo SLON-13 navedel vseh devet mest kot možne variante slovnične oblike.

Če vpišete besedo:

"lepo"

in kliknete "Išči", dobite dve možni gesli:

- 1) lastnostni prislov »lepo«, oznaka 2004 – 4 no koncu oznake pomeni lastnostni prislov
- 2) pridevnik »lep«, ki se lahko pojavlja v 4 različnih povezavah:

- a) z ženskim samostalnikom v 4. sklonu ednine,
- b) z ženskim samostalnikom v 6. sklonu ednine,
- c) s samostalnikom srednjega spola v 1. sklonu ednine in
- d) s samostalnikom srednjega spola v 4. sklonu ednine.

Ukaz »Počisti«

Z ukazom »Počisti« zbrisate iz zaslonskega okna vse rezultate, ki jih ne potrebujete. Če ukaza »Počisti« ne uporabite, se rezultati v oknu nalagajo drug za drugim zato jih lahko vse skupaj shranite šele na koncu dela z ukazom »Shrani«.

Z ukazom »Počisiti« izbrisate vse rezultate – razen tistih, ki ste jih shranili z ukazom »Shrani«.

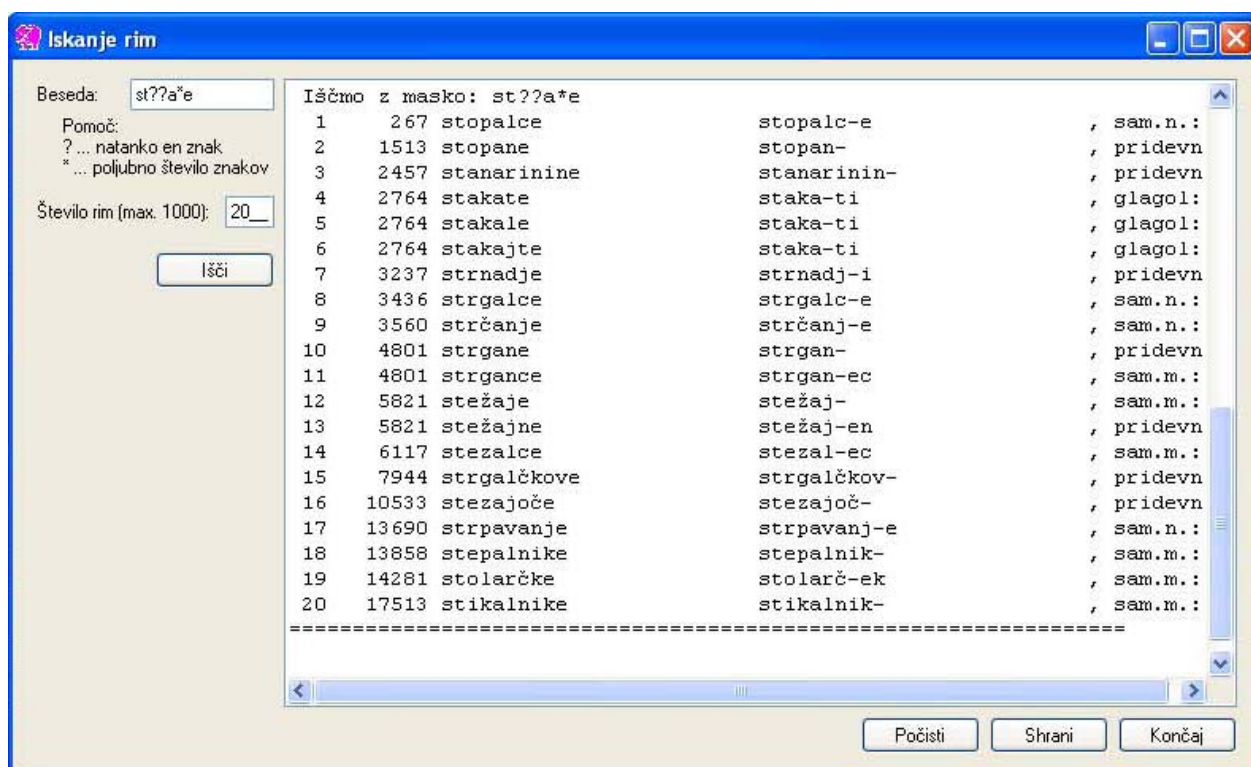
3.5 Iskanje rim

Iskanje rim lahko bolj določeno opišemo kot »iskanje besed s samo delno danim zaporedjem črk in znakov«. Tako iskanje omogoča najti družine besed, ki vsebujejo iste črke ali isto kombinacijo črk, ki večinoma niso rime v peniškem pomenu te besede.

Besedno zvezo »iskanje rim« in »rima« uporabljamo v kontekstu programa SLON-13 kot iskanje različnih besed z istim črkovnim zaporedjem.

Iskalna maska (zahteva) pri opciji "Iskanje rim" sledi istim pravilom kot iskalna maska pri opciji "Konkordance", le da je pri rimah, zaradi omejitve na iskanje posameznih slovarskih besedah, manj možnosti za kombinirane oz. za bolj komplicirane iskalne zahteve vrste »ali-ali«.

Primer iskanja »rim« je prikazan na sliki 9. Opcija »Iskanje rim« omogoča precej več kot samo iskanje pravih rim v pesniškem pomenu besede »rima«.



Slika 9. »Iskanje rim« z zahtevo, da se vse besede prično s črkama »st«, ki jima sledita dve poljubni črki, za njima (kot peta v besedi) črka »a« in končno v poljubni oddaljenosti še črka »e«, ki mora besedo zaključiti.

Okno v opciji »Iskanje rim« ima dve vhodni polji:

- »Beseda«
- »Število rim«

in štiri gumbе:

- »Išči«
- »Počisti«
- »Shrani«
- »Končaj«

Podatek »Število rim« omeji najdene besede na izbrano število. Največ lahko izpišemo 1000 rim. Če uporabnik meni, da je pri določenih uporabah ta številka premajhna, lahko po dogovoru z avtorji dobi verzijo, ki omogoča izpis vseh najdenih rim, ne glede na število zadetkov. Zainteresirani uporabniki se lahko o tej in tudi drugih željah dogovore na naslovu:

jure.zupan@ki.si

Z ukazom »Išči« pričnemo z iskanjem po slovarju SLON-13.

Za opis željene besede ali rime lahko v opciji "Iskanje rim" poleg črk uporabljate le dva posebna znaka:

- * zvezdico kadarkoli želite poljubno število (tudi nič) črk kjerkoli ob črkovnem zapisu ali ob vprašajih, ki samo opredelijo obvezno navzočnost črk.
- ? vprašaj kadarkoli predpišete eno samo poljubno črko. Vprašajev je lahko več. Kolikor vprašajev je v iskalni maski toliko črk mora biti v najdeni besedi na mestih, ki jih opredelijo vprašaji.

Oba znaka, zvezdica in vprašaj, imata enako nalogo, kot pri konkordancah. Vprašaj pomeni obvezno navzočnost ene črke, zvezdica pa navzočnost poljubnega števila črk (lahko tudi nobene). Razlika med konkordančnim iskanjem in iskanjem rim je v tem, da pri "iskanju rim" **uporaba pike in znaka # ni dovoljena**. Prav tako ni dovoljena uporaba samih vprašajev ali samo vprašajev in zvezdic.

!!! V iskalni maski mora biti navzoča vsaj ena črka. !!!

Uporabo vprašaja in zvezdice najlaže razložimo na primerih.

Preglednica 5 Primeri "Iskalne maske" pri opciji "Iskanje rim"

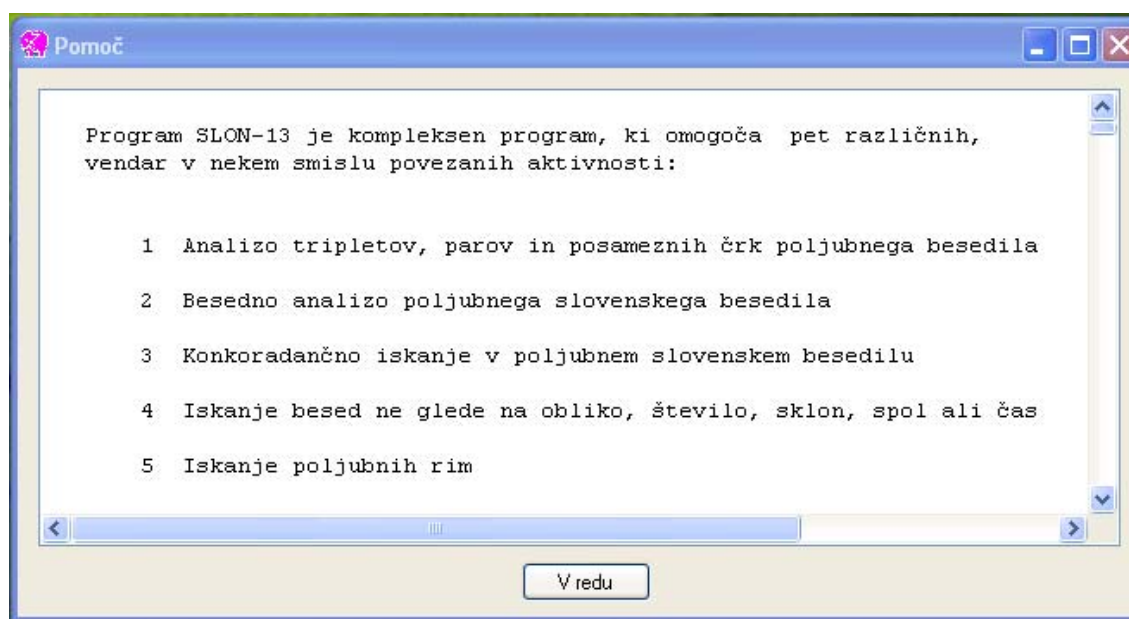
Iskalna maska	Opis zahteve	Možni rezultati
*ekli	Rime na "ekli"	tekli, stekli, jekli, pekli, ...
*a*e	Asonance na a-e	oblake, žlahtne, snubače, domače, ...
kol*	Besede, ki se prično s "kol"	kolec, kolikor, kolarju, kolesom, kolovodjema, kolenčkova,
*ovci	Besede, ki se končajo z "ovci"	ovci, lovci, pokovci, plovci, vdovci, minolovci, trgovci, novci
ft	Besede, ki vsebujejo črki "ft"	klaftra, kaftan, taft, naftalin, nafta, mufti, klaftra,...
*st*pe*		drugostopen, gostopet, postopek,...
lop??	Besede s 5. črkami, ki se začno z "lop"	lopat, lopah, lopov, lopic, lopar, lepot, lopim,
?peti	Besede s 5. črkami, ki se končajo "peti"	opete, speti, vpeti
k??la	Besede s 5. črkami, ki se začno s »k« in končajo na »la«	kikla, kuzla, klala, kislá, kozla,
?pe??ta?		operetah, speketam
?as*te*		kasete, faseterjev, faseter, nasiten..
*po??b*t?		onesposobljenosti, pozabita, zapogibati, neuporabnosti, pozebsti, ...

3.6 Pomoč

Z gumbom »Pomoč«, ki se nahaja v Glavnem oknu, lahko izpišete na zaslon kratek opis vseh opcij z navodili za delo. Besedilo je shranjeno v datoteki "help.txt", ki je na distribucijski zgoščenki.



Ko kliknete na gumb »Pomoč« se v oknu »Pomoč« izpiše pričetek opisa programa SLON-13. S pomočjo drsnikov lahko pregledujete celotno besedilo.



S klikom na gumb »V redu« zaključite z delom v opciji »Pomoč« in se vrnete v Glavno okno.

4. Inštalacija programa

Na zgoščenki s programom SLON-13 so naslednje datoteke

Programi:

- Slon-13.exe
- ana01.exe
- ana02.exe
- kon01.exe
- beseda.exe
- rime.exe

Slovar:

- sl-28.11x

Tekstovne datoteke:

- mati si zeli kosa.txt testni program za Besedno analizo
- help.txt besedilo, ki se izpiše pri opciji »Pomoč«
- Navodila SLON-13.pdf pdf datoteka z navodili
- 3pesniki_k.txt tri pesniške zbirke: Prešeren, Gedorčič, Menart

Inštalacija ne zahteva drugega, kot da v poseben zavihek ali mapo (direktorij), ki si ga (jo) predhodno naredite prepisete vse programe iz distribucijskega diska preko ukaza

- »Urejanje« (Edit) in
- »Kopiraj« in »Prilepi« (Copy/Paste)

Opozorilo.

Programa SLON-13 ne morete poganjati iz zgoščenke! Prekopirati ga morate v mapo, ki dopušča pisanje (svetujemo ime mape SLON-13)

5. Zahvala

Del raziskav in programiranja je bil financiran s strani Ministrstva za visoko šolstvo znanost in tehnologijo s posredovanjem ARRS preko programa P1-0017 in projekta J7-0382.

6. Index

- Ali-ali 19, 21, 24
ARRS 27
Analiza črk, parov in tripletov (opcija), 5-9
 besed 4, 5, 9, 10, 15, 27
 konteksta 13
 stavčna 4
Analiziraj tekst (ukaz) 6, 10

Basic, Visulal, 3
beseda, pregibna 17, 20
 nepregibna 12, 17
besedilo, analizirano 11
Besedna analiza (opcija) 5, 10, 11, 13
besedna vrsta 11
besedna oblika 12
Brskaj (ukaz) 15

čas 10, 12, 14-16, 20, 22
 sedanji 14
 pretekli, prihodnji 12, 14
črke 17-19, 24, 25
črkovni vzorec 4
črkovna kombinacija 17, 18, 24, 25
črkovni nabor 17
črkovno območje 16

datoteka 5-11, 15, 16, 21, 26, 27
 tekstovna 5, 6, 10, 27
 dvojina 12-14, 18, 23

ednina 12-14, 23
enota, stavčna 4, 11

geslo, slovarsko 4, 11, 12, 22

homonim 22

inštalacija programa 27
iskalna zahteva 19
iskalna maska 20
Iskanje besed (ukaz) 18, 22
Iskanje rim (ukaz) 24

konkordanca 3.4.
Konkordance (opcija) 15, 16, 19
konkordančen(a)
 beseda 15, 16
 iskanje (ukaz) 4, 16
 maska 17
 širina 15, 16
 zadetek 15, 16
 zahteva 16-19
kontekst 4, 13, 18

ločilo 8, 11, 17
»lojtra« (#) 19
lupina, programska 3, 5

marker 16
mesto obrazila 14
množina 12-14, 18, 23

modeliranje podatkov 4
modul 5

MVZT 4

napake 3

oblika 16, 18, 19, 22
 besedna 4, 12, 14, 20
 slovarske informacije 22
 slovnična 17-19
obrazilo 3, 12-14, 20, 23
 samostalniško 13
 prazno 13
 prideviško 13
 glagolsko 14
Odpr (ukaz) 6, 10
 okno, glavno 5, 10, 26
 notranje 6, 7
komunikacijsko 3, 6, 10
omejitvene pike 17, 19, 20

par 4, 6-8
par, izbor parov 8
par, najpogostejši par 9
prihodnji čas 12, 14
Počisti (ukaz) 15, 21-24
poezija 9
presledek, navidezni 8
pretekli čas 12, 14

različica gesla 12
razumevanje besedila, strojno 13
rezultat 11, 15, 16, 18, 20-23
rima 24, 25
romanopisje 9

sedanjik, sedanji čas 14, 23
Shrani (ukaz) 5, -7, 10, 15, 21
slovar 3, 5, 17, 22
slovarsko geslo 11, 12

triplet 4-9
 najpogostejši 7
 standardni 7, 9

varianta, besede 12, 20, 21, 23
velelnik 12, 14, 23
vprašaj 18-20, 25

Word, MS 5

zakonodaja 9
zapri 6, 8, 10
zavihek (mapa, direktorij) 5, 7, 10, 25
zgoščanka 3, 11, 21, 26, 27
zvezdica 18-21, 25

žanr 9
žurnalistika 9