



A few new utilities developed at University of Belgrade

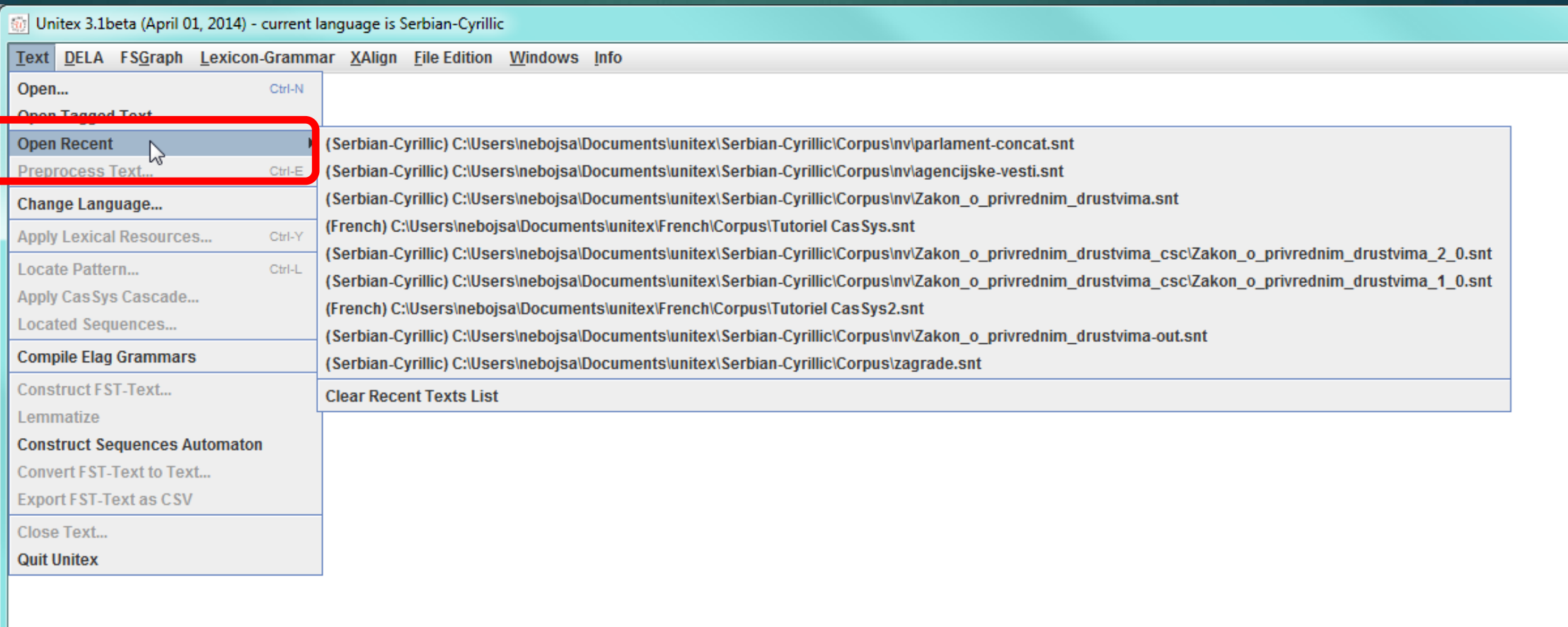
DUŠKO VITAS, CVETANA KRSTEV,
NEBOJŠA VASILJEVIĆ, MIRKO SPASIĆ, MARKO ŠOŠIĆ (& SAŠA PETALNIKAR)

Remarks

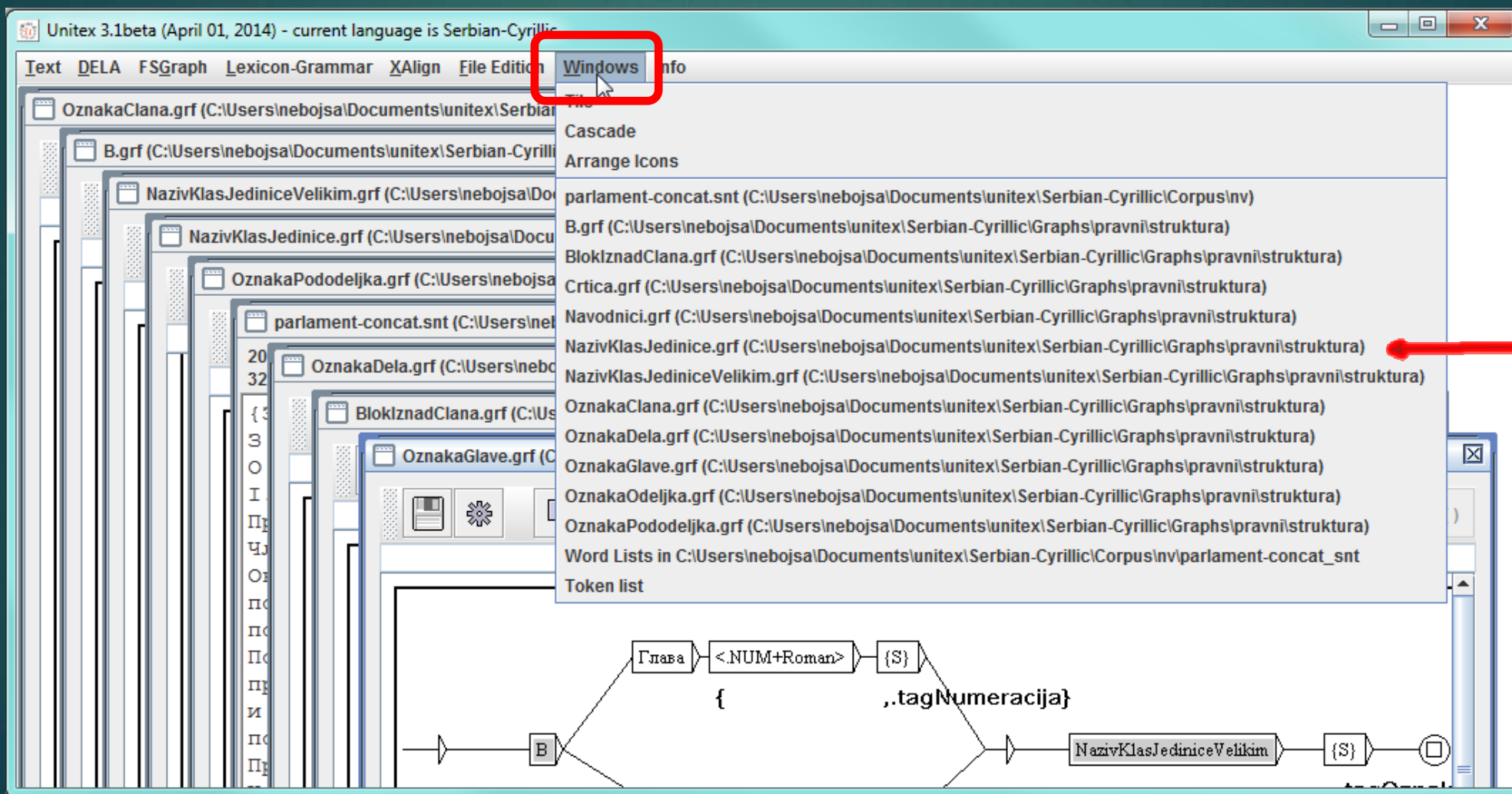
- ▶ Experiments with the Unitex code;
- ▶ Additions developed by Nebojša are in the distribution 3.1, the others are not;
- ▶ Programmed in C/C++ and Java except Stadeux (C#);
- ▶ Current versions are available on request to vitas@matf.bg.ac.rs

Nebojša – Text menu – Open Recent

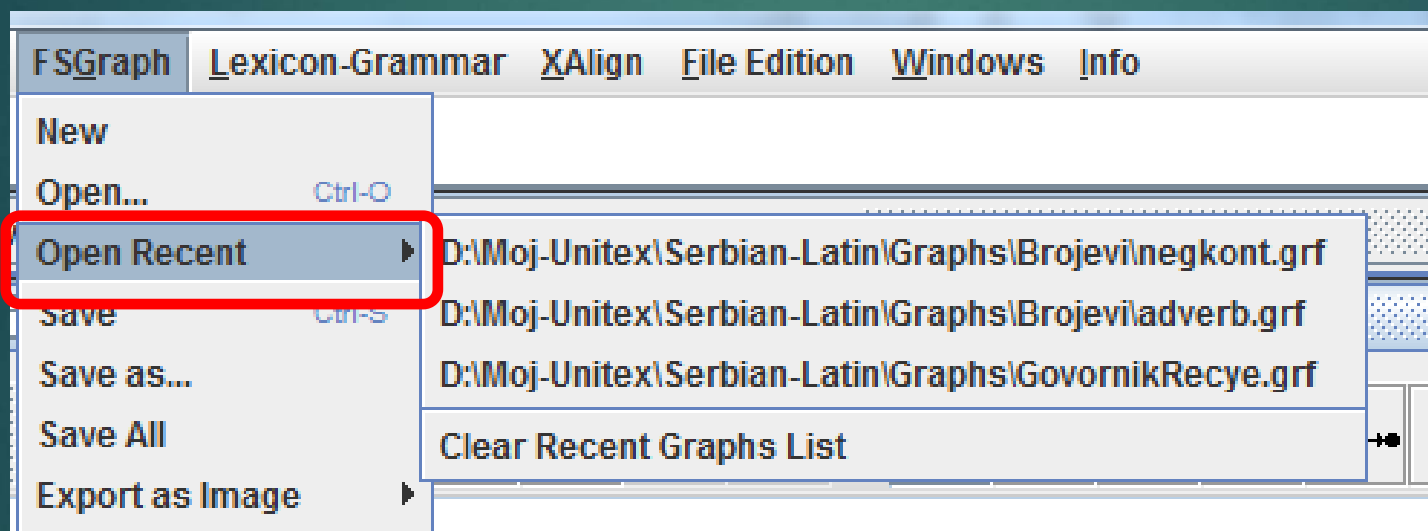
3



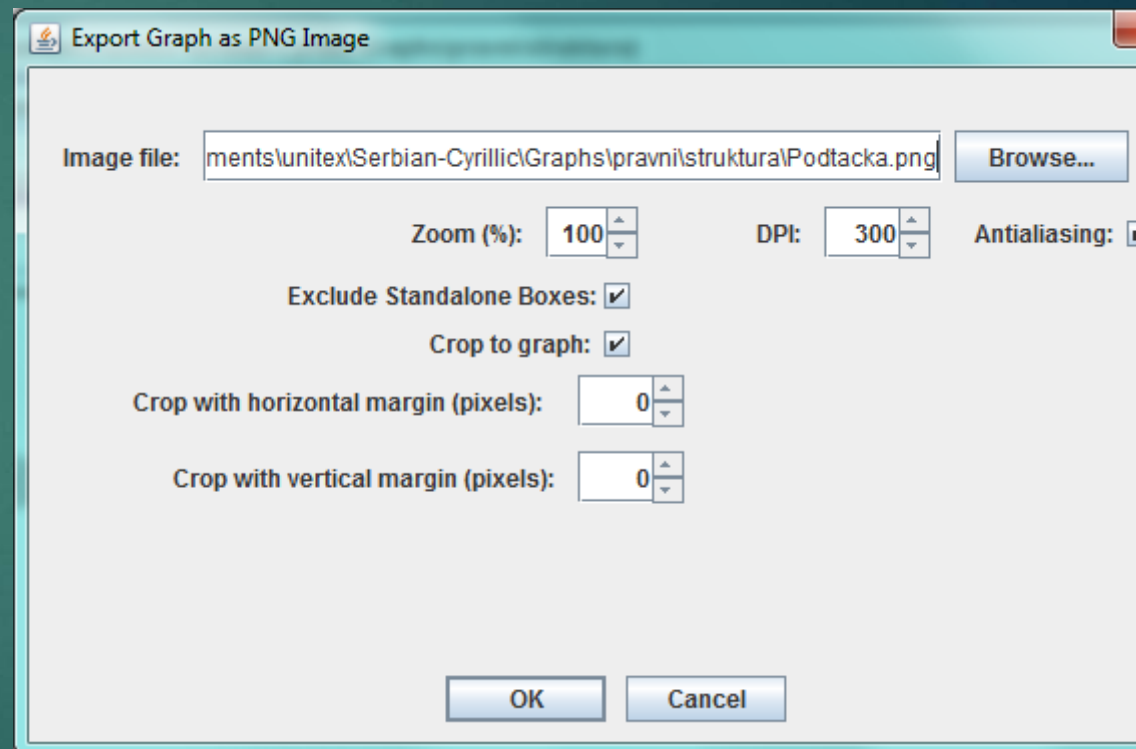
Nebojša – Windows menu



Nebojša – FSGraph

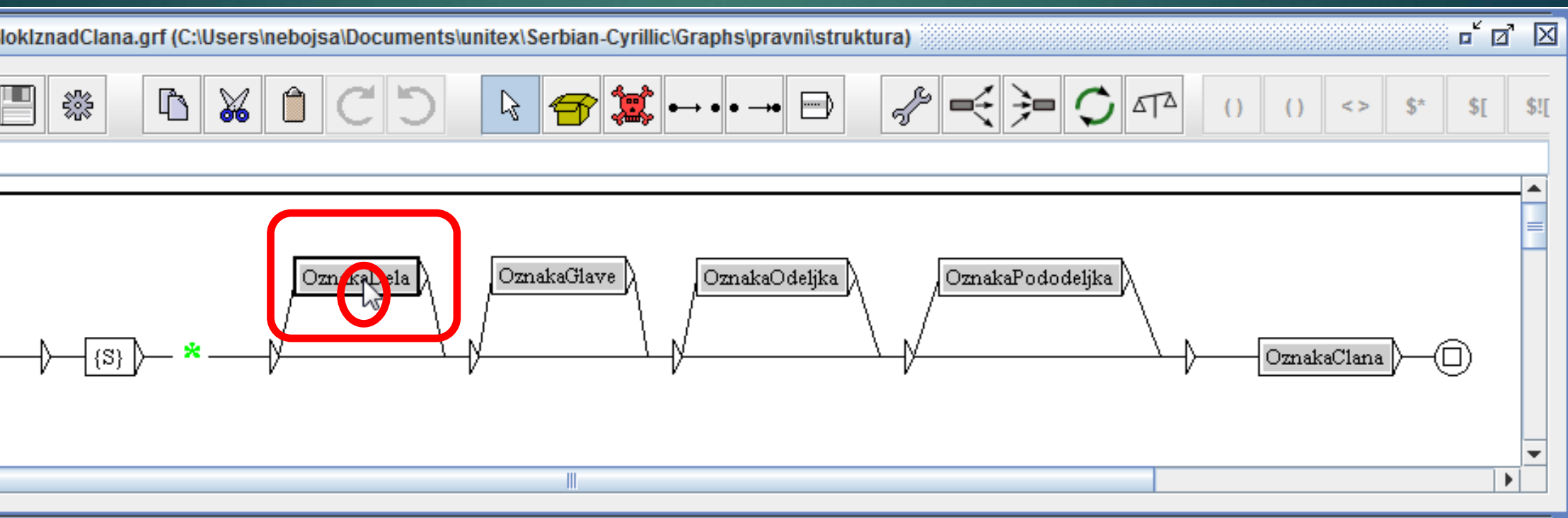


Nebojša – FSGraph



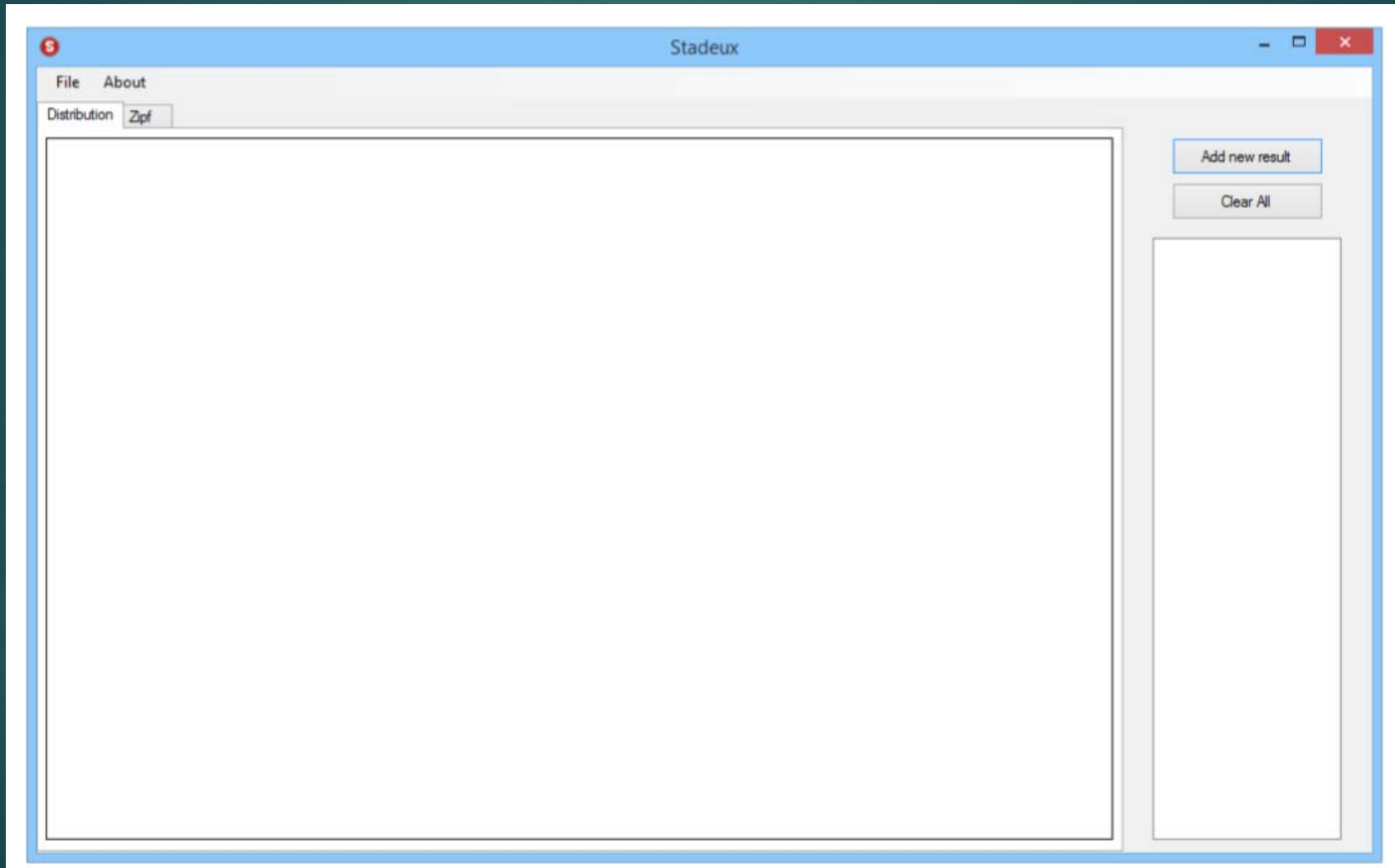
Nebojša – Graph edition (rollover)

7

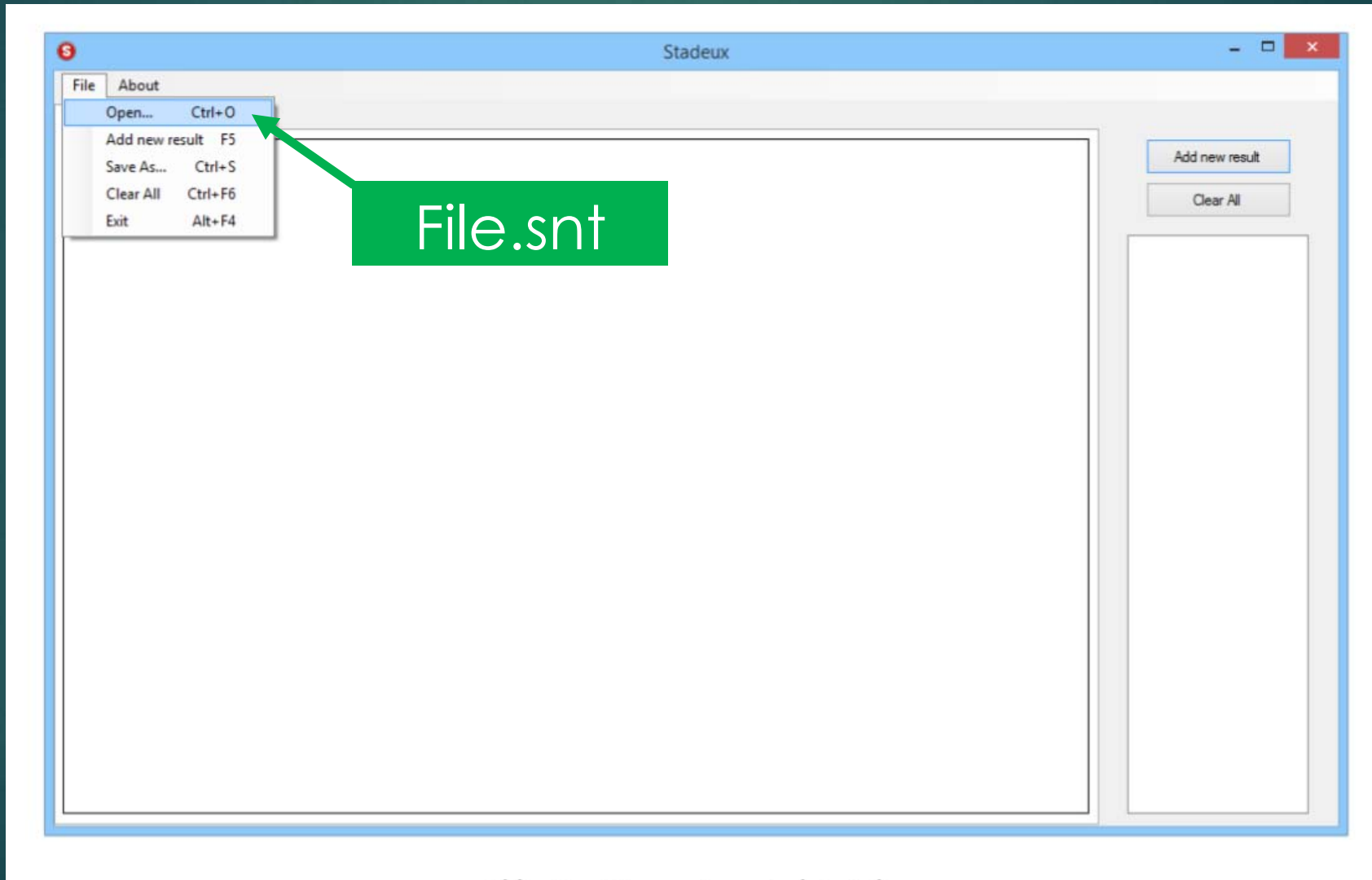


Jean Senellart → *Stadeux* de Marko Šošić

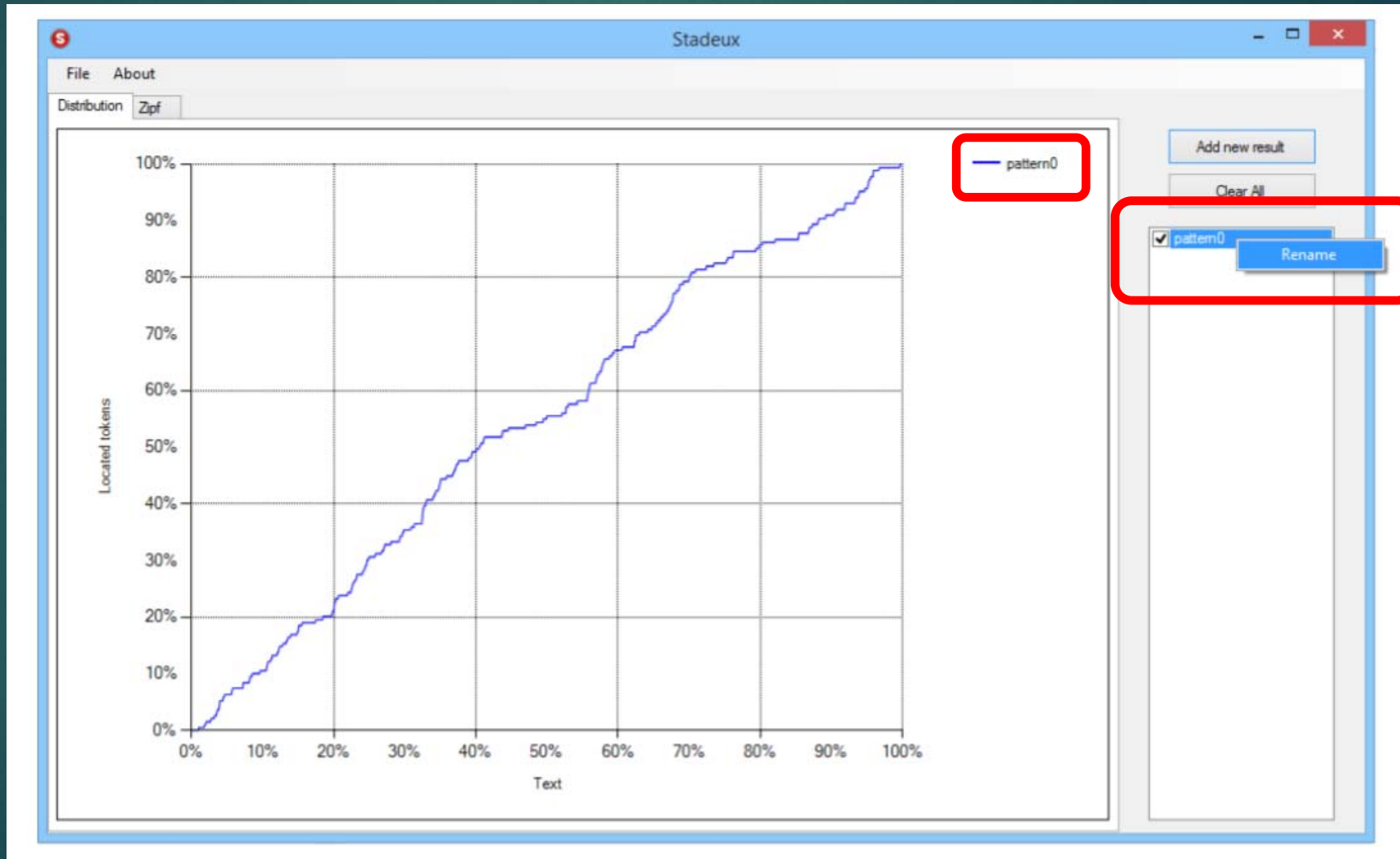
8



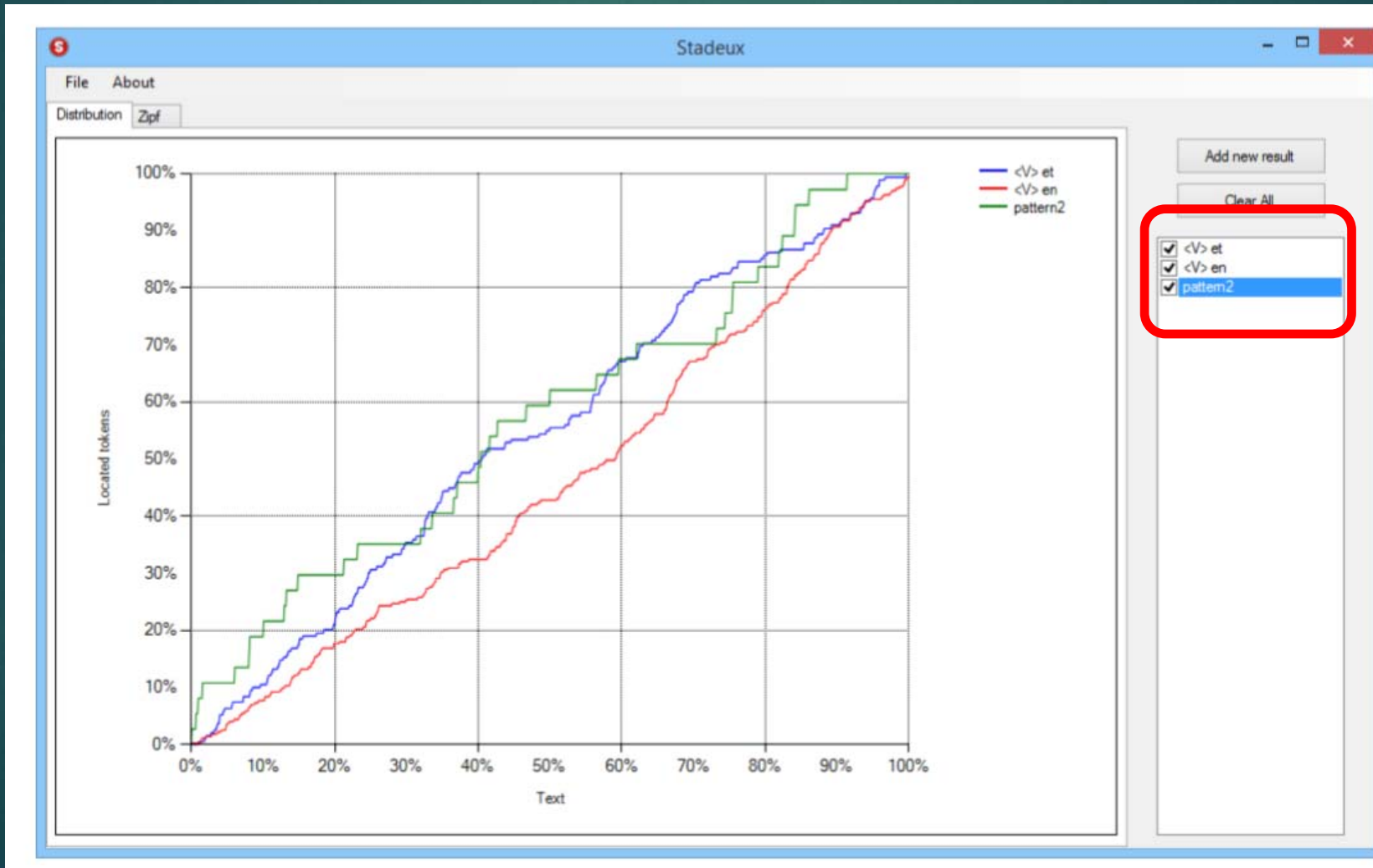
Open



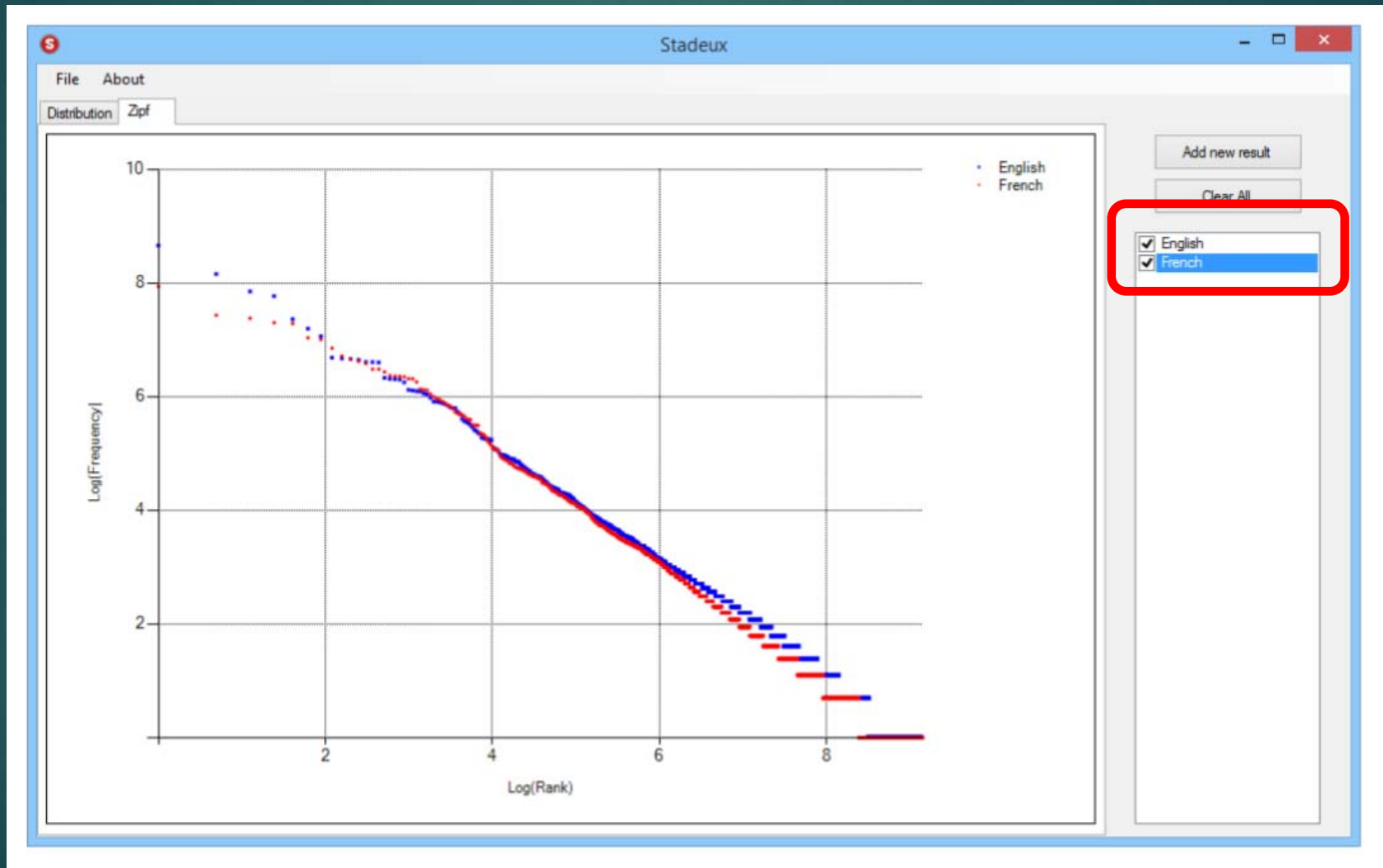
Result



Accumulation of results



Zipf's law



Help for manual lemmatization (Mirko Spasić)

Locate *<piti>* (& right click on concordances)

PIVOT.snt (D) Concordance: D:\Moj-Unitex\Serbian-Latin\Corpus\PIVOT_snt\concord.html

5152 sentence 200 matches
78866 occurrence

om «gajbom» (u Burgundiji, jedno bure / piće = 228 litara).{S} Nije spo
ponovo poprima svoje izvorno značenje piće antičkih bogova, napitak be
ena rečenica bila je: «Izutetan - leš - piće - novo - vino.» Otud i neob
gao da budem prijatelj sa nekim za koga piće i jelo nemaju nikakve druge
emeno, ne dogovarajući se, pomislili na piće. {S}
nih čitališta «Ništa nije prijatnije za piće od v
Severa? {S}Ne znam.{S}.. Al' slatko za piće, {S}
tičku... slučajnost.{S} Dakle?{S} Koje je piće nada
eva, isto kao i njena kultura.{S} To je piće - tote
ita vratara koje su to «tri stvari koje piće jako
otrovuša, prljava voda, loše vino, loše piće) je
anđeljima je vino postalo svetotajinsko piće, par
ili u kućama Šampanje, ne bi mi laskalo piće nog
ay): «Vrlo retko sam trošio to gazirano piće. {S}
dije poželeo jedno industrijsko, strano piće, neg
skih grla, obznanili su da je to jedino piće koje

Disambiguate


 Splitted

Dif candidates

piće,,N:ns1q
piće,,N:ns4q
piće,,N:ns5q
piće,piti.V+Imperf+Tr+It+lref:Fzp
piće,piti.V+Imperf+Tr+It+lref:Fzs

Dic candidates

Form *pile*

The image shows a software interface with a concordance window and a disambiguation dialog. The concordance window displays a list of sentences with the word 'pile' highlighted in blue. The disambiguation dialog is open over the concordance window, showing a list of candidates for the word 'pile'.

Concordance Window:

- 5152 sentence
- 78866 occurrence
- 200 matches
- ci, dvostruko zadovoljstvo: najpre vino **piti**, a zatim o njemu pisati. {S}
- Ilijadi i Odiseji), u duhovnosti (Sveto **pismo**) vino je važnije i od samo
- čahu šampanjca ili vina i predlažemo da **pijemo** u {S} Sve je u tome u (à)
- {S} Srećni smo što smo se sastali i što
- a od 24. i 25. {S} To što se prvenstveno
- dinim nehajnim «cin!». {S} Možemo, tako,
- su izgovori da se vino obilato i toči i
- na svet svog unuka Žorža, beleži samo «
- u zdravlje novorođenčeta». {S} Šta li su
- šampanjac**, ali je tvrdio da su zvanice
- {S} Slučaj je hteo da se isti šampanjac
- nati, ponekad i licemerni, i mora da se
- } Nekada se u redakcijama dnevnih novina
- braka, vidite kako sam veseo i srećan,
- ćan, pijem u vaše zdravlje, prijatelji,
- om prikazu Vinskih i raskalašnih pesama
- ovrdžavi») Živeo je, voleo, bludničio i
- Prpratni tekst: «Iz ove čaše svaki dan
- vino, sveže i mirisno. {S} Što nisi tu?»
- . {S} Ali, Matis u tom času nije alzaško
- {S} L'Amour... Ljubav i vino {S} Ko ume da
- ume da voli. {S} Ko ume da voli, u me da

Disambiguate Dialog:

- Splitted
- Dlf candidates
- Dlc candidates
- pile, N+Zool:ns1v
- pile, N+Zool:ns4v
- pile, N+Zool:ns5v
- pile, pila. N:fw2q
- pile, pila. N:fw4q
- pile, pila. N:fp1q
- pile, pila. N:fp4q
- pile, pila. N:fp5q
- pile, pila. N:fs2q
- pile, piliti. V+Imperf+Tr+It+lref:Pzp
- pile, piti. V+Imperf+Tr+It+lref:Gwf
- pile, piti. V+Imperf+Tr+It+lref:Gpf
- None
- Multiword unit candidate
- Ok
- Apply changes (6)
- Discard changes(6)

PIVOT.snt (D) Concordance: D:\Moj-Unitex\Serbian-Latin\Corpus\PIVOT_snt\concord.html

5152 sentence 200 matches
78866 occurrences

Naslov orig
{S}Bernard
de l'Académ
{S}Dictionn
© Plon, 200
{S}Bernar P
{S}Rečnik z
{S}Sa crtež
{S}Preveo s
{S}Milojko
{S}Anjesi,
koja se u o
{S}Predgovo
{S}Časno vi
{S}Jedino š
vinu, kao i
slučaj, koj
poređenju s
podrumari,
eruditi čuv
neka poznat
sačine knji
{S}Ono što
dvostruko z
najpozvanij
jednog živo
degustaciju
da vode pub

ci, dvostruko zadovoljstvo: najpre vino [piti](#), a zatim o njemu pisati.{S}
čašu šampanjca ili vina i predlažemo da [{pijemo,piti.V+Imperf+Tr+It+Iref](#)
{S} Srećni smo što smo se sastali i što [{pijemo,piti.V+Imperf+Tr+It+Iref](#)
a od 24. i 25. {S}To što se prvenstveno [pije](#) u zdravlje nekom pojedincu
dinim nehajnim «cin!». {S}Možemo, tako, [piti](#) i u zdravlje nekome ko nije
su izgovori da se vino obilato i toči i [pije](#), u čemu, u nekim porodicama
na svet svog unuka Žorža, beleži samo «[pili](#) smo u zdravlje novorođenčeta
u zdravlje novorođenčeta». {S} Šta li su [pili](#) {pili,piti.V+Imperf+Tr+It+I
avlje novorođenčeta». {S} Šta li su pili [{pili,piti.V+Imperf+Tr+It+Iref:G](#)
šampanjac, ali je tvrdio da su zvanice [{pile,piti.V+Imperf+Tr+It+Iref:G](#)
anjac, ali je tvrdio da su zvanice [{pile,piti.V+Imperf+Tr+It+Iref:G](#)
{S} Slučaj je hteo da se isti šampanjac [pije](#) i na mojoj veridbi {S} Dvade
nati, ponekad i licemerni, i mora da se [pije](#) rakija od žitarica koja se
{S}Nekada se u redakcijama dnevnih novina [pilo](#) barem isto onoliko koliko se
braka, vidite kako sam veseo i srećan, [pijem](#) u vaše zdravlje, prijatelji
ćan, pijem u vaše zdravlje, prijatelji, [pijte](#) i vi u naše.{S}.. Kad god
om prikazu Vinskih i raskalašnih pesama [piše](#): «Ta ljubav prema vinu, ta
ovrdžavi») živeo je, voleo, bludničio i [pio](#) u drugoj polovini drugoga ve
Propratni tekst: «Iz ove čaše svaki dan [pijem](#), u tvoje zdravlje, alzaško
vino, sveže i mirisno.{S} Što nisi tu?» [Pismo](#) je datirano 6. maja 1945,
. {S} Ali, Matis u tom času nije alzaško [pio](#) iz rodoljublja, nego iz sklo
{S}L'Amour... Ljubav i vino {S}Ko ume da [pije](#), ume da voli. {S}Ko ume da

PIVOT.snt (D) Concordance: D:\Moj-Unitex\Serbian-Latin\Corpus\PIVOT_snt\concord.html

5152 sentence 200 matches
78866 occurrences

***Po greg
Badnji dan
«Cin-cin!»
tsing, «zdr
nekoga da p
nagrada, ne
{S}Možemo,
bolest.{S}
porodicama,
Igo, pozdra
novorođenče
znamo.
{S}Moj otac
šampanjac,
(Moët et Ch
Dvadeset pe
kompanije M
{S}Ne cenim
razgovori e
pije rakija
množenje ča
Zauzvrat, p
posluženog
zdravicu.{S}
pridruži za
zadovoljstv
{S}Nekada s
je bilo izg
je vođen ka
reportaže i

škkih tehnika.{S} Nije ni opšta istorija vinove loze {vinove loze, vinova
. {S} Nije ni opšta istorija vinove loze {vinove loze, vinova loza.N+Bot+C
je: vino je kultura.{S} Kultura gajenje vinove loze {vinove loze, vinova
kultura.{S} Kultura gajenje vinove loze
atko, vino nije što i obično pivo, nije
je što i obično pivo, nije mačji kašalj
li pak neverovatno. - Prim. prev. ***Po
i 25. decembra su Badnji dan i Božić po
ama, da pozovemo nekoga da popije pivo,
ja starija kći se zaposlila u službi za
don.{S} Eto šta znači biti veran jednoj
m uz desert. {S}Nekada se u redakcijama
im vratima zgrade na kružnoj raskrsnici
inopija rođen u Iranu, koji je pisao na
averne i zabasava na večernje zabave po
u Nivas ne može ni da zamisli da ijedna
re. → ISLAM I VINO {S}Alzas {S}Rizling,
mao više od 2.000 žitelja, po završetku
{S}Kao i ukusna skuša, i Alzas pliva u
a skuša, i Alzas pliva u belom vinu.{S}
noir) jedina je koja daje ružičasto ili
sveta 1989, upita nas za koje bismo se
... is Rizling (Rizling) {S} memo

Disambiguate

? Splitted
Dif candidates

Dic candidates
gregorijanskom kalendaru,gregorijanski kale
gregorijanskom kalendaru,gregorijanski kale

None
 Multiword unit candidate

OK Apply changes (0) Discard changes(0)

{S}Zajedno dižemo čašu šampanjca ili vina i predlažemo da {**pijemo,piti.V+Imperf+Tr+It+lref:Pxp**} u.{S} Sve je u tome u (à) koje će za tren uvesti želje, zahteve, žudnju. «Za nas!» (À nous!) je najsažetiji iskaz.{S} Srećni smo što smo se sastali i što {**pijemo,piti.V+Imperf+Tr+It+lref:Pxp**}, budimo samoživi, mislimo najpre na sebe, poželimo sebi sve lepo i dobro, što svakom pristoji. «U zdravlje!» kazuje gotovo isto to, ističući zdravlje, koje se u francuskom (À la nôtre!) podrazumeva*.{S}

None.txt & Mutiword.txt

None.txt:

ena rečenica bila je: «Izutetan – leš – [piće](#) – novo – vino.» **Otud i neobično ime** diči svojom sposobnošću da se slaže sa [belim lukom](#), jer ta privilegija nije data zanimljivo: beli ermitaž ne uzmiče pred [belim lukom](#){S} To tvrdi onaj koji nosi {S}Bernar Pivo {S}Rečnik zaljubljenika [vina](#) {S}Sa crtežima Alena Bulduira {S}Preveo sa I blanc), šato auzon svrstan u vrhunska [vina](#) prve kategorije, klasa A (premier grand

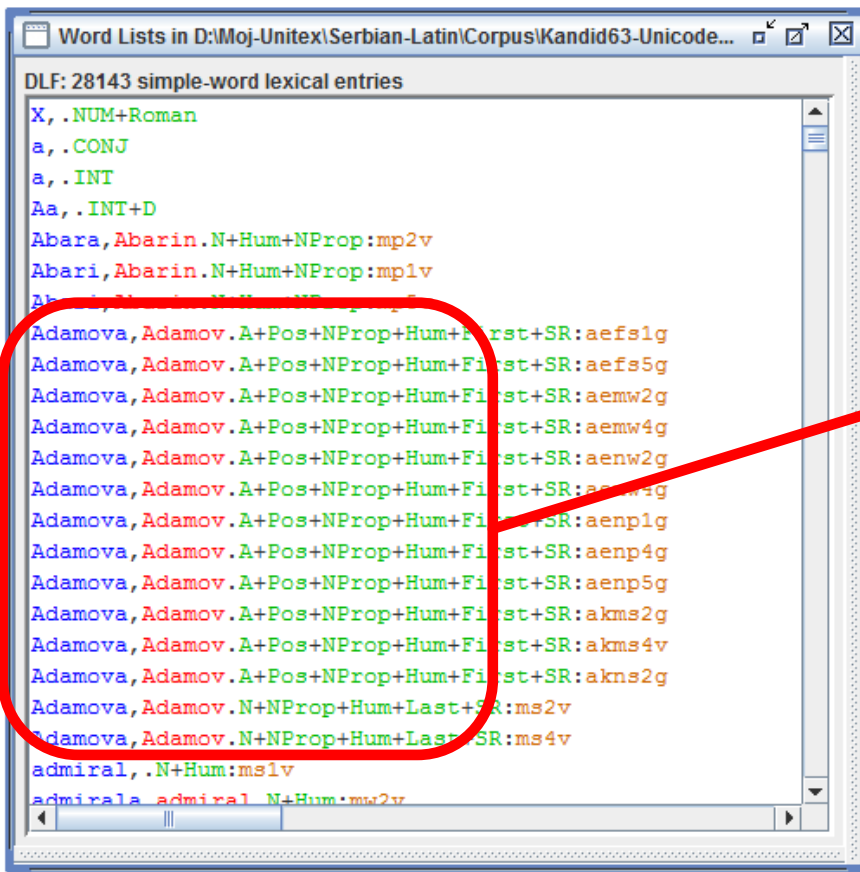
Multiword.txt:

braka, vidite kako sam veseo i srećan, [pijem](#) u vaše zdravlje, prijatelji, pijte

Stat_spl.txt - Quarantine

piće, .N:ns1q	0	2
pismo, .N:ns1q	1	0
pili, piliti.V+Imperf+Tr+It+Iref:Pzs	0	1
vina, vino.N:nw4q	0	3
pile, piti.V+Imperf+Tr+It+Iref:Gwf	0	1
pismo, .N:ns5q	0	1
pili, piliti.V+Imperf+Tr+It+Iref:Azs	0	1
vina, vino.N:np4q	0	3
vina, vino.N:np2q	0	3
pili, piti.V+Imperf+Tr+It+Iref:Gpm	1	0
vina, vino.N:nw2q	0	3
piše, piti.V+Imperf+Tr+It+Iref:Azp	0	1
pile, pila.N:fp4q	0	1
vina, vino.N:ns2q	0	3

Dictionary of lemmata (Saša Petalnikar)



Lemma list

By Frequency By Char Order

1676	jesam
1142	ona
984	dati
945	i
938	da
656	sebe
626	a
605	se
451	biti
432	on
403	koji
387	taj
379	ja
362	na
360	ono
357	Kandid
309	sam
254	sav
231	jedan
221	od

Lemma list

By Frequency By Char Order

Aa	1
Abarin	1
Adamov	1
Afrika	5
Ahmet	6
Aleksandrija	2
Alžir	3
Amerika	3
Anri	1
Antonija	1
Antonije	1
Antonina	1
Arabija	1
Ariostov	2
Ariovist	1
Askanije	1
Atrebatija	1
Avasen	1
Aveslon	1
Avram	1

```
D:\Matf\Radni\Serbian-Latin\Corpus\Primer.snt
2 sentence delimiters, 86 (27 diff) tokens, 34 (20) simple forms, 0 (0) digit
19 occurrences (15 DLF entries) simple words, 0 occurrence (0 DLC entries) compound word, 15 occurrences (7 ERR lines) unk...
Ponovo jedna rečenica. {S} Još jedna rečenica. {TB}Treći Naslov{TE}
{EOF}Druga_Datoteka{TE}
{TB}Naslov u drugoj datoteci{TE}
rečenica u drugoj datoteci.
{TB}Drugi Naslov U Drugoj Datoteci{TE}
```

```
Concordance: D:\Matf\Radni\Serbian-Latin\Corpus\Primer_snt\concord.html
5 matches
Link by Selection Link By Contains
1.1. Test primer {TB}Prvi Naslov{TE} Jedna rečenica.{S} Još jedna rečenica. {TB}Drugi Naslov{TE} P
1.1. Naslov{TE} Jedna rečenica.{S} Još jedna rečenica. {TB}Drugi Naslov{TE} Ponovo jedna rečenica.{S
1.2. nica. {TB}Drugi Naslov{TE} Ponovo jedna rečenica.{S} Još jedna rečenica. {TB}Treći Naslov{TE} {
1.2. TE} Ponovo jedna rečenica.{S} Još jedna rečenica. {TB}Treći Naslov{TE} {EOF}Druga_Datoteka{TE}
2.1. ka{TE} {TB}Naslov u drugoj datoteci{TE} Rečenica u drugoj datoteci. {TB}Drugi Naslov U Drugoj D
```

Merci!