

# ÚKL CJBB84

12.3.2010

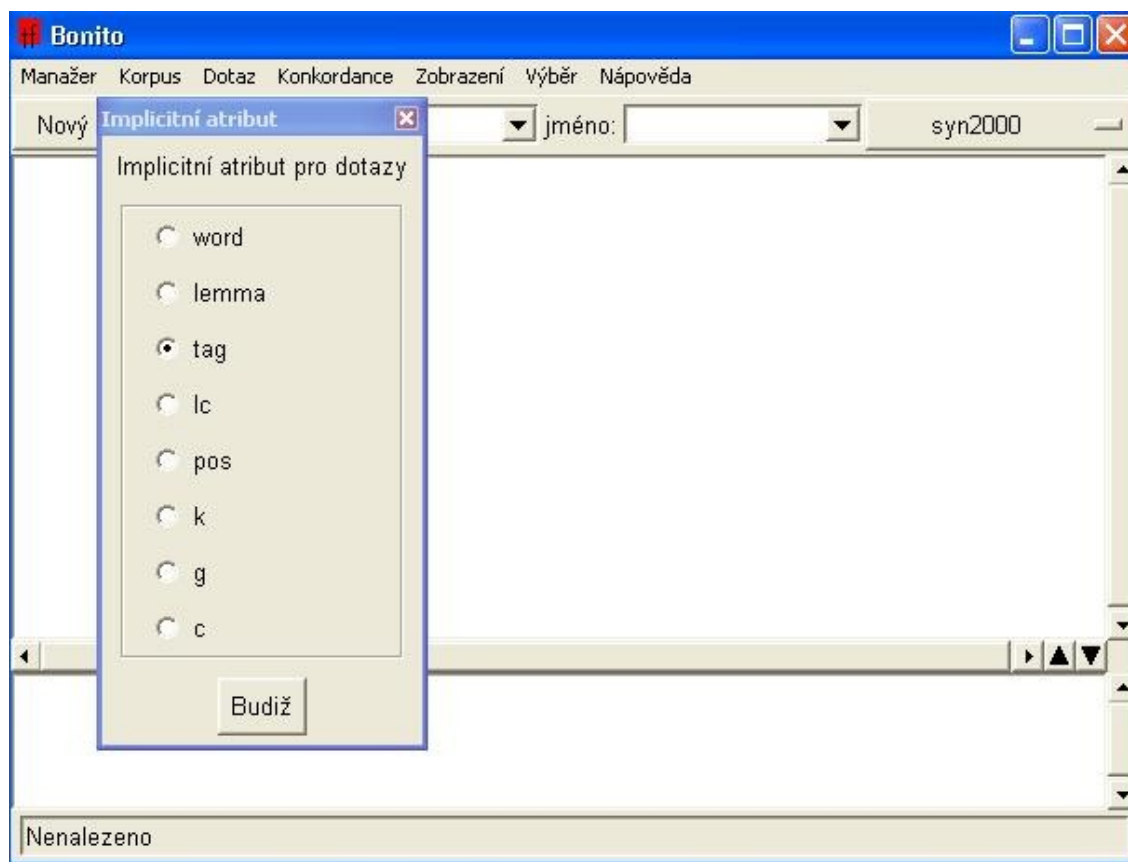
# Variantní a dubletní koncovky

- Na základě pozorování dat z korpusu (např. SYN2000) formulujte pravidla distribuce variantních koncovek vybraného flektivního typu.

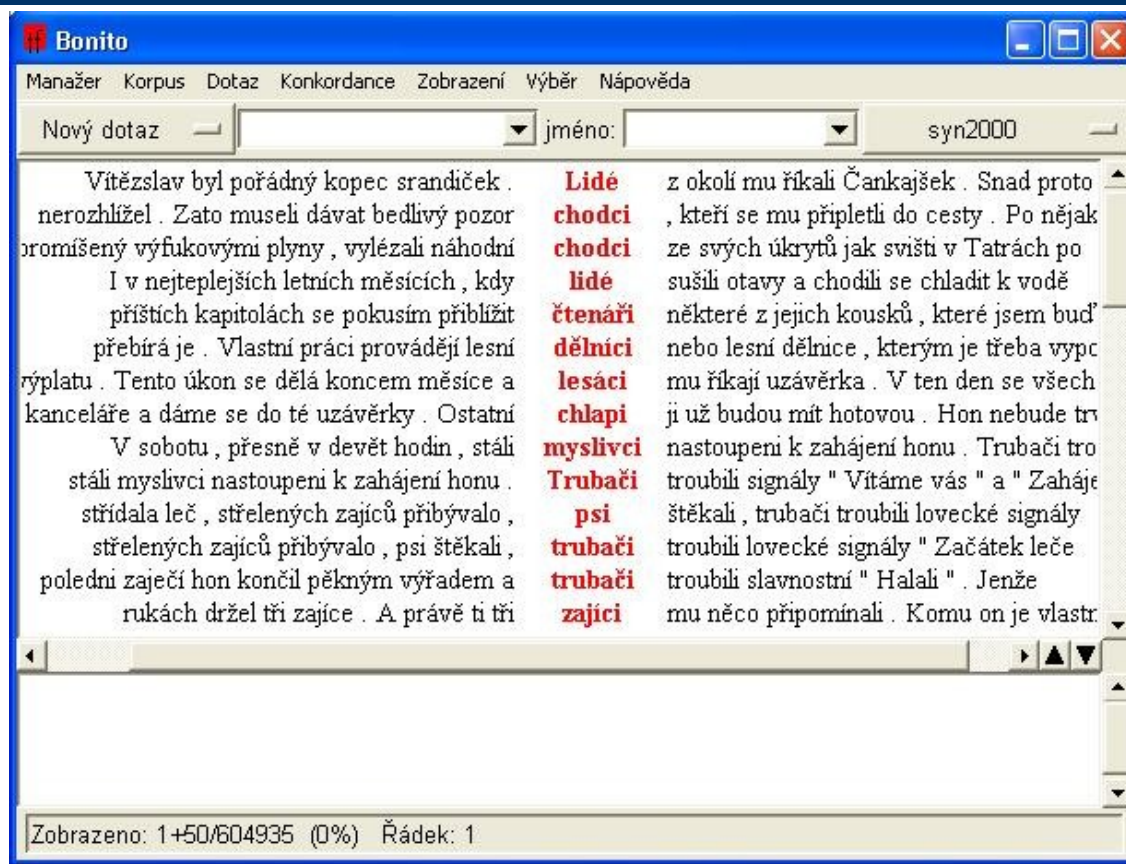
# nom. pl. mask. živ.

- implicitní atribut **tag**
- dotazovací řádek: **NNMP1.\***
- implicitní atribut **lemma**
- Pfiltr: **.\*[bdfghklmnpstvz]**
- Pfiltr: **.\*[bflmpsvz]**

# implicitní atribut tag



# dotazovací řádek: NNMP1.\*



# implicitní atribut lemma

**Bonito**  
Manažer Korpus Dotaz Konkordance Zobrazení Výběr Nápověda

Implicitní atribut pro dotazy

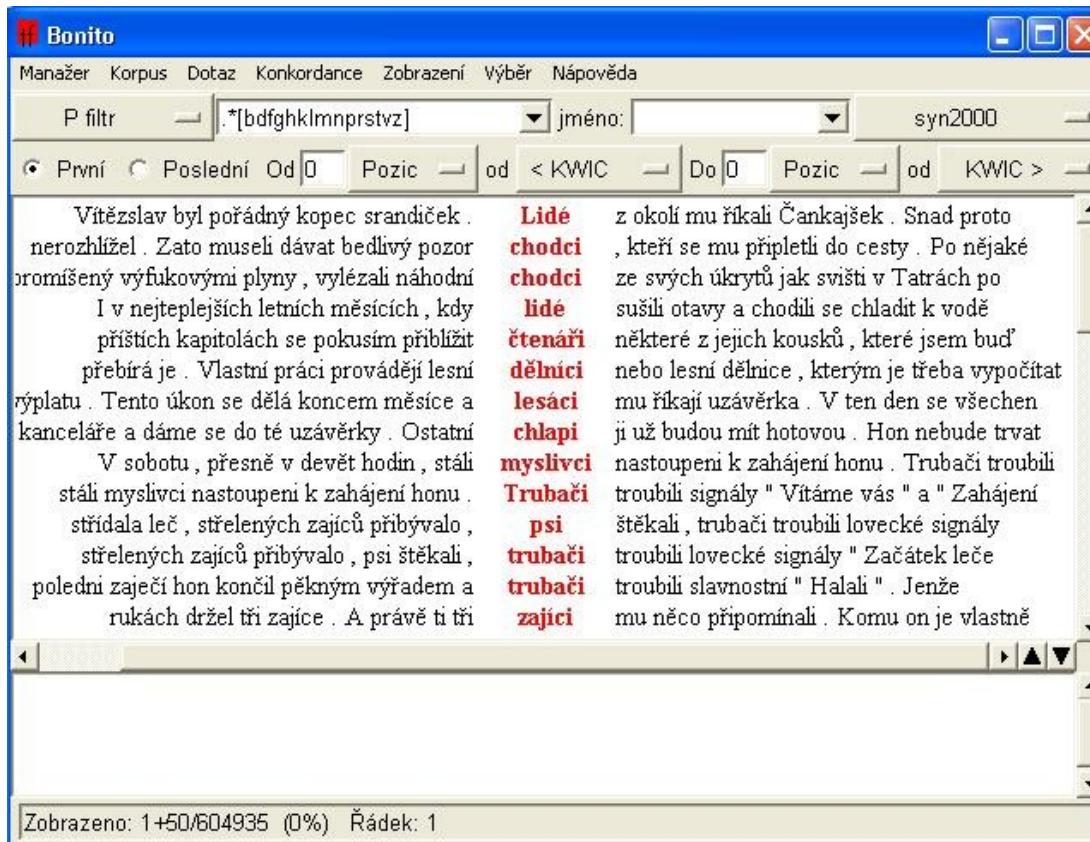
- word
- lemma
- tag
- lc
- pos
- k
- g
- c

Budiž

zic	od	< KWIC	Do	Pozic	od	KWIC >
srandiček .		<b>Lidé</b>				z okolí mu říkali Čankajšek . Snad proto
edlivý pozor		<b>chodci</b>				, kteří se mu připlétli do cesty . Po nějaké
zali náhodní		<b>chodci</b>				ze svých úkrytů jak sviští v Tatrách po
stících , kdy		<b>lidé</b>				sušili otavy a chodili se chladit k vodě
stím přiblížit		<b>čtenáři</b>				některé z jejich kousků , které jsem buď
ovádějí lesní		<b>dělníci</b>				nebo lesní dělnice , kterým je třeba vypočít
m měsíce a		<b>lesáci</b>				mu říkají uzávěrka . V ten den se všechen
ky . Ostatní		<b>chlapi</b>				ji už budou mít hotovou . Hon nebude trva
hodin , stáli		<b>myslivci</b>				nastoupení k zahájení honu . Trubači trout
ájení honu .		<b>Trubači</b>				troubili signály " Vítáme vás " a " Zahájení
á přibývalo ,		<b>psi</b>				štěkali , trubači troubili lovecké signály
psi štěkali ,		<b>trubači</b>				troubili lovecké signály " Začátek leče
výřadem a		<b>trubači</b>				troubili slavnostní " Halali " . Jenže
a právě ti tři		<b>zajíci</b>				mu něco připomínali . Komu on je vlastně

Zobrazeno: 1+50/604935 (0%) Řádek: 1 Vybráno: 1

# Pfiltr: **.\*[bdfghklmnrstvz]**



# konkordance – statistiky – frekvenční distribuce – lemma

Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0 Řádků k zobrazení: 5257

lemma: 0<KWIC  zobrazit  součet  schovat

lemma:	##
člověk	44074
člen	9354
představitel	7625
občan	6343
odborník	6190
pracovník	6099
voják	5524
host	5425
divák	5219
politik	5036
obyvatel	4975
student	4528
autor	4105
ministr	4082
pán	3828
Čech	3768
přítel	3538
účastník	3436
Američan	3129
majitel	3002
podnikatel	2955
demokrat	2934
úředník	2900

Zavřít Uložit



# Odstranění substantiv skloňovaných podle vzoru *muž*

- Substantiva na *-tel* (*představitel, obyvatel, přítel, podnikatel, ...*)
- Jaká další ?
- Obojetná souhláska ?
- Všechny ?

# vzor *pán/muž*

- Pfiltr: **.\*b**
- Pfiltr: **.\*f**
- Pfiltr: **.\*l**
- Pfiltr: **.\*m**
- Pfiltr: **.\*p**
- Pfiltr: **.\*s**
- Pfiltr: **.\*v**
- Pfiltr: **.\*z**

# výsledky pozorování

- **. \*b**
- Konkordance
- Statistiky
- Frekvenční distribuce
- **67 výskytů všechna vzor *pán***
- Konkordance
- Zpět k předchozí

# výsledky pozorování

- **. \*f :**
- Konkordance
- Statistiky
- Frekvenční distribuce
- **71 výskytů všechna vzor *pán***
- Konkordance
- Zpět k předchozí

# výsledky pozorování

- Konkordance
- Statistiky
- Frekvenční distribuce
- **.\*|**
- Nfiltr: **\*tel**
- *datel, Obrtel, ... ?*
- Konkordance
- Statistiky
- Frekvenční distribuce
- 511 výskytů *král, !anděl, !manžel, ?Korbel* jinak všechna vzor *pán*
- Konkordance
- Zpět k předchozí (2x)

# výsledky pozorování

- **. \*m :**
- Konkordance
- Statistiky
- Frekvenční distribuce
- **56 výskytů všechna vzor *pán***
- Konkordance
- Zpět k předchozí

# výsledky pozorování

- **. \*p**
- Konkordance
- Statistiky
- Frekvenční distribuce
- **46 výskytů všechna vzor *pán***
- Konkordance
- Zpět k předchozí

# výsledky pozorování

- **.\*s**
- Konkordance
- Statistiky
- Frekvenční distribuce
- Nfiltr: `.*[ei]us`
- Nfiltr:  
`[AÁBCČDĎEÉFGHIÍJKLMNŇOÓPQRŘSŠTŤUÚVWXYZŽ].*`
- *33 výskytů vzorů pán (pes, pštros, kos, běs) i muž (rorýs, sýs, karas, klid'as)*
- Konkordance
- Zpět k předchozí (4x)



# výsledky pozorování

- *.\*v*
- Konkordance
- Statistiky
- Frekvenční distribuce
- *137 výskytů všechna vzor *pán** Konkordance
- Zpět k předchozí (4x)

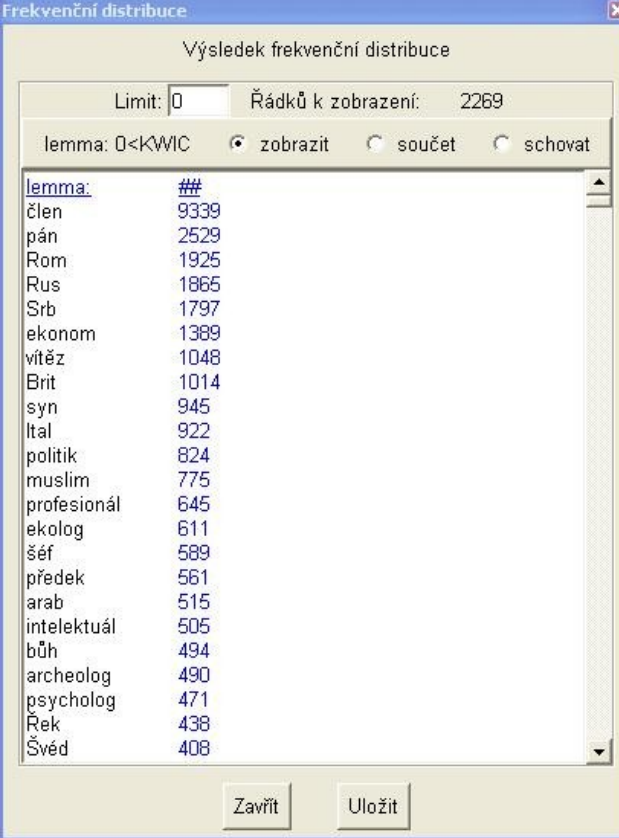
# výsledky pozorování

- **. \*z**
- Konkordance
- Statistiky
- Frekvenční distribuce
- **52 výskytů vzorů muž (vítěz, Francouz, kněz, ...) i pán (plaz, dikobraz, hrdlořez, ...)**
- Konkordance
- Zpět k předchozí

# Distribuce koncovek

- -i
- -ové
- -é

implicitní atribut **word**  
dotazovací řádek **.\*ové**  
konkordance – statistiky – frekvenční distribuce – **lemma**



Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0    Řádků k zobrazení: 2269

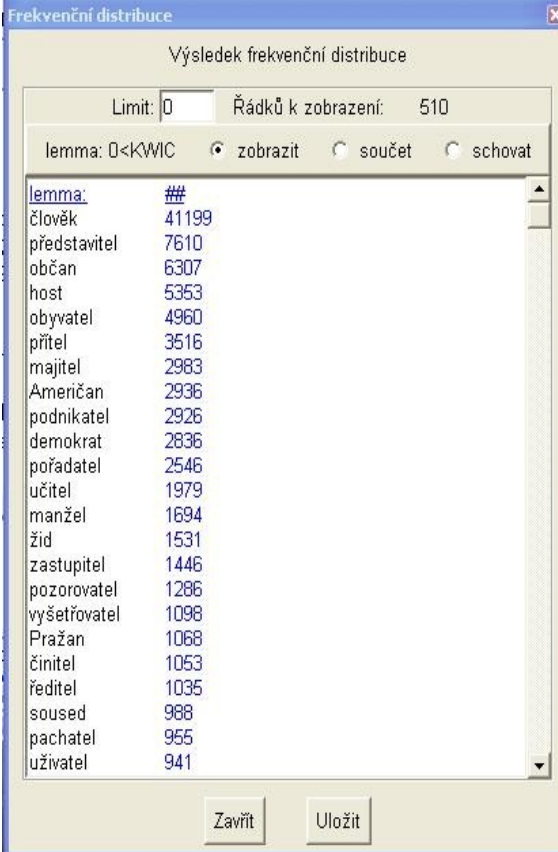
lemma: 0<KWIC     zobrazit     součet     schovat

lemma:	##
člen	9339
pán	2529
Rom	1925
Rus	1865
Srb	1797
ekonom	1389
vítěz	1048
Brit	1014
syn	945
Ital	922
politik	824
muslim	775
profesionál	645
ekolog	611
šéf	589
předek	561
arab	515
intelektuál	505
bůh	494
archeolog	490
psycholog	471
Řek	438
Švéd	408

Zavřít    Uložit

konkordance – zpět k předchozí  
dotazovací řádek **. \*é**

konkordance – statistiky – frekvenční distribuce – **lemma**  
Nfiltr **. \*ové**



The screenshot shows a window titled "Frekvenční distribuce" with a subtitle "Výsledek frekvenční distribuce". It contains a table of lemmas and their frequencies. The window has a search bar with "Limit: 0" and "Řádků k zobrazení: 510". Below the search bar are three radio buttons: "lemma: 0<KWIC" (selected), "zobrazit", "součet", and "schovat". The table has two columns: "lemma:" and "##". The data is as follows:

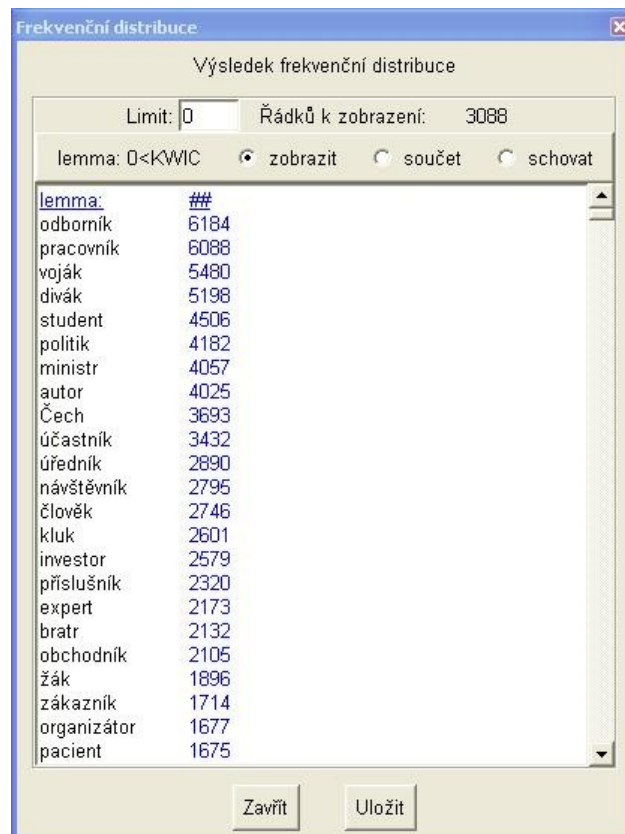
lemma:	##
člověk	41199
představitel	7610
občan	6307
host	5353
obyvatel	4960
přítel	3516
majitel	2983
Američan	2936
podnikatel	2926
demokrat	2836
pořadatel	2546
učitel	1979
manžel	1694
žid	1531
zastupitel	1446
pozorovatel	1286
vyšetřovatel	1098
Pražan	1068
činitel	1053
ředitel	1035
soused	988
pachatel	955
uživatel	941

At the bottom of the window are two buttons: "Zavřít" and "Uložit".

konkordance – zpět k předchozí (2x)

dotazovací řádek **. \*i**

konkordance – statistiky – frekvenční distribuce – **lemma**



The screenshot shows a window titled "Frekvenční distribuce" with a subtitle "Výsledek frekvenční distribuce". It contains a table of lemmas and their frequencies. The window also has a "Limit" field set to 0, a "Řádků k zobrazení" field set to 3088, and radio buttons for "zobrazit" (selected), "součet", and "schovat".

lemma:	##
odborník	6184
pracovník	6088
voják	5480
divák	5198
student	4506
politik	4182
ministr	4057
autor	4025
Čech	3693
účastník	3432
úředník	2890
návštěvník	2795
člověk	2746
kluk	2601
investor	2579
příslušník	2320
expert	2173
bratr	2132
obchodník	2105
žák	1896
zákazník	1714
organizátor	1677
pacient	1675

# pozorování závislosti výskytu na

- zakončení : ke vzoru *muž* mohou patřit pouze substantiva na *-l, -s, -z*, nikoli na libovolnou obojetnou souhlásku, naopak na *-c* životná patří vždy ke vzoru *muž* srv. ovšem u neživ. *tácy, hecy, kecy*)
- derivační typ : *-an-é, -at-é, -graf-ové, -log-ové, ...* (srv. sufix *tel*)
- osobovost (*-ové*)
- obyvatelská (*Žid*)
- jednotlivé léxémy (*lidé, host, žid, soused, ...*)

# Úkol na 2.4. 2010

- vok. sg. mask. živ.
- gen. pl. fem. *růže*
- gen . pl. neu. *moře*



# Úkol na 9.4. 2010

- distribuce variant krátký/dlouhý krátký/dlouhý (jinak alternující) kořenový vokál u některých feminin vzoru *žena* (*kráva/krav, síla/sil, bouda/bud, kůra/?kor, ...*).
- U kterých substantiv ?
- V kterých tvarech ?
- Jak je najít v korpusech?
- Jak vypadá hodnocení těchto tvarů ve slovnících SSJČ a SSČ ?
- Co lze vyčíst z korpusů ?