

ZPK CJBB75

19.3.2010

Jak zobecnit vyhledávání tvarů ?

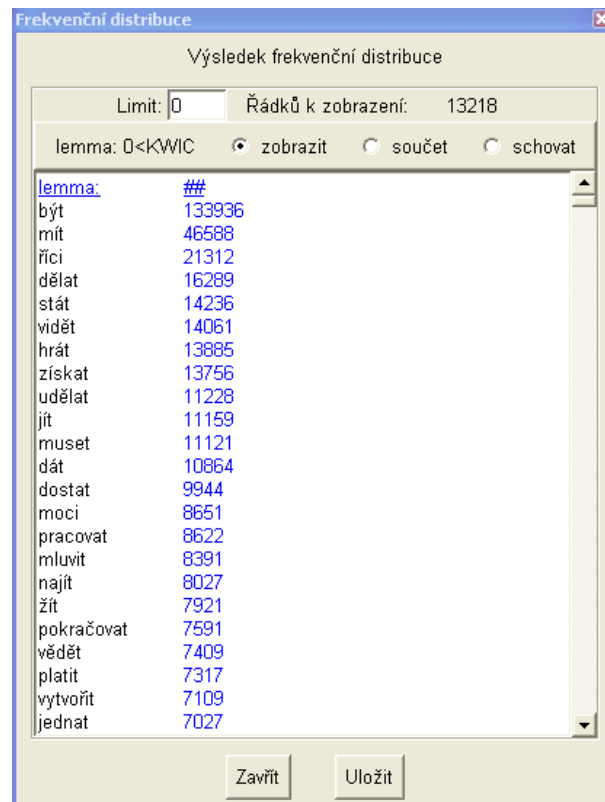
- hledání podle morfologické značky
- Implicitní atribut **tag**
- Vyhledávání podle zakončení tvaru
- Implicitní atribut **word**

Jak může být zakončen tvar infinitivu v češtině ?

- *-t*
- *-ti*
- *-ct*
- *-ci*

Implicitní atribut **tag**
dotazovací řádek **Vf.***

Konkordance – statistiky – frekvenční distribuce – **lemma**



Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0 Řádků k zobrazení: 13218

lemma: 0<KWIC zobrazit součet schovat

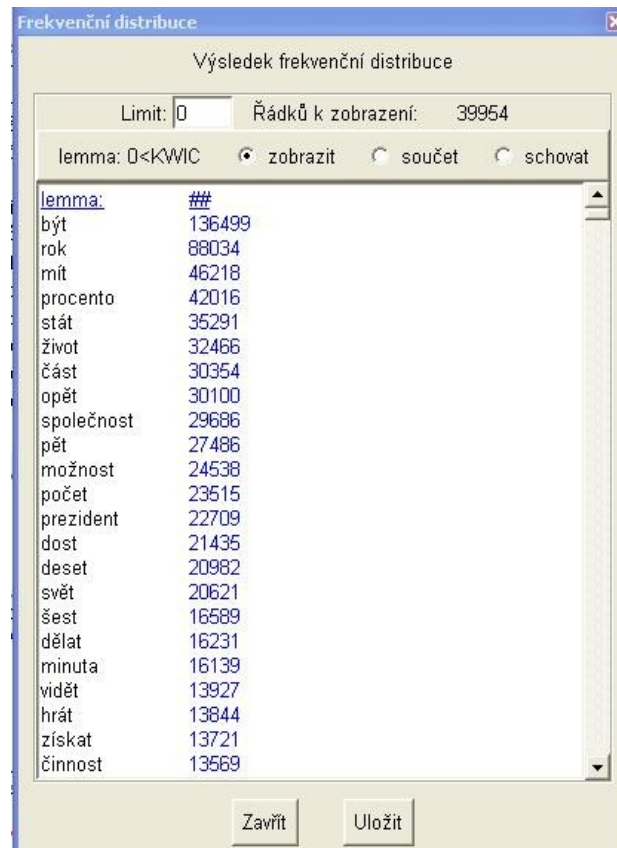
lemma:	##
být	133936
mit	46588
řici	21312
dělat	16289
stát	14236
vidět	14061
hrát	13885
získat	13756
udělat	11228
jít	11159
muset	11121
dát	10864
dostat	9944
moci	8651
pracovat	8622
mluvit	8391
najít	8027
žít	7921
pokračovat	7591
vědět	7409
platit	7317
vytvořit	7109
jednat	7027

Zavřít Uložit

Implicitní atribut **word**

dotazovací řádek **.*t**

Konkordance – statistiky – frekvenční distribuce – **lemma**



Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0 Řádků k zobrazení: 39954

lemma: 0<KWIC zobrazit součet schovat

lemma:	##
být	136499
rok	88034
mit	46218
procento	42016
stát	35291
život	32466
část	30354
opět	30100
společnost	29686
pět	27486
možnost	24538
počet	23515
prezident	22709
dost	21435
deset	20962
svět	20621
šest	16589
dělat	16231
minuta	16139
vidět	13927
hrát	13844
získat	13721
činnost	13569

Zavřít Uložit

Pozorujte případy tvarových „homonymií“

- *let* (g.pl. *rok*)
- *procent* (g.pl.)
- *stát* (n. sg.)
- *život* (n. sg.)
- *část* (n. sg.)
- *opět* (adv.)
- *společnost* (n. sg.)
- *pět* (num.)

Je možné zobecnit jejich formální vlastnosti ?

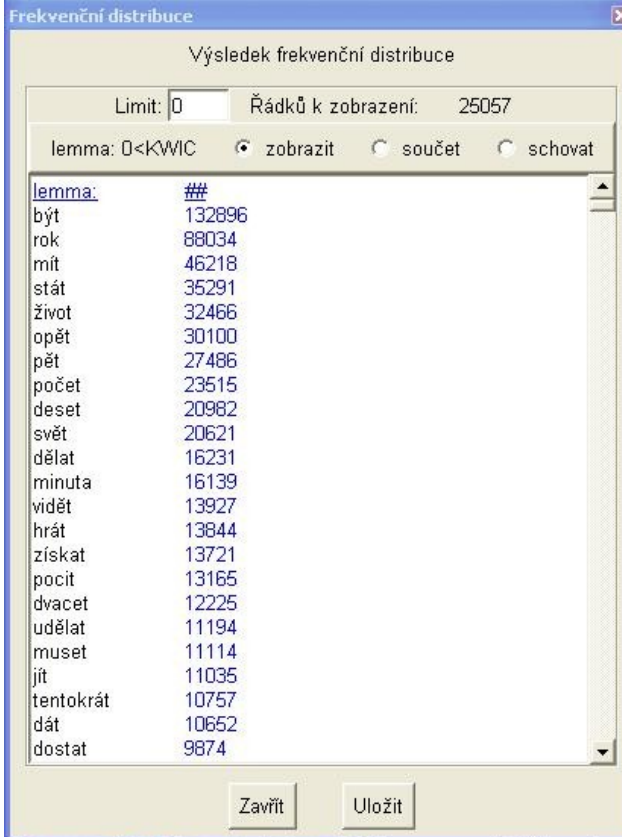
- *let : jet*
- *procent : ?*
- *stát : stát*
- *život : ?*
- *část : krást*
- *opět : zapět*
- *společnost : ?*
- *pět : pět*

Jakou kombinací hlásek / písmen může končit infinitiv v češtině ?

- $V+t$
- $C+t$

Implicitní atribut **word**
Pfiltr **.*[aáééěííoóuůýý]t**

Konkordance – statistiky – frekvenční distribuce – **lemma**



Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0 Řádků k zobrazení: 25057

lemma: 0<KWIC zobrazit součet schovat

lemma:	##
být	132896
rok	88034
mit	46218
stát	35291
život	32466
opět	30100
pět	27486
počet	23515
deset	20982
svět	20621
dělat	16231
minuta	16139
vidět	13927
hrát	13844
získat	13721
pocit	13165
dvacet	12225
udělat	11194
muset	11114
jít	11035
tentokrát	10757
dát	10652
dostat	9874

Zavřít Uložit

Všimněte si

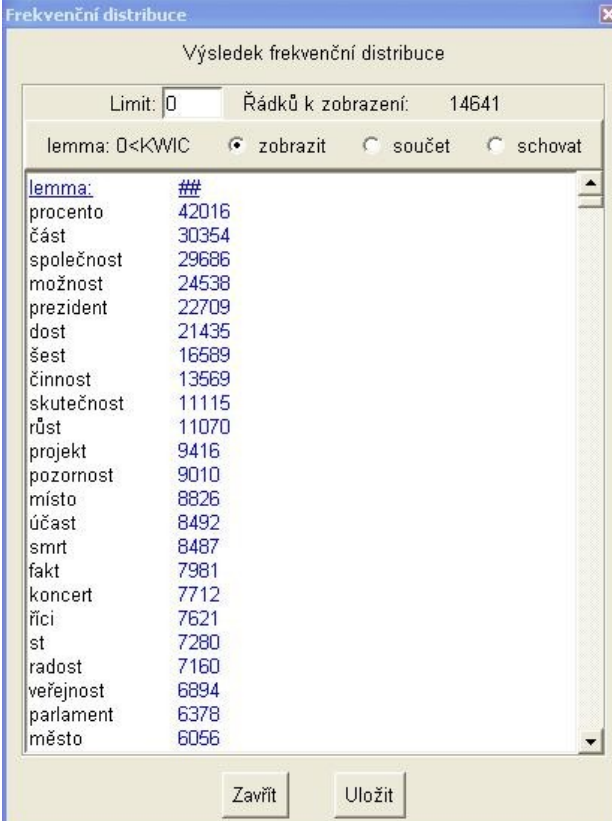
- *život*
- *minut*

Konkordance – zpět k předchozí

Implicitní atribut **word**

Pfiltr **.*[bcčdd'fghjklmnňpqrřsštt'vzž]t**

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 table has two columns: "lemma:" and "##". The data is as follows:

lemma:	##
procento	42016
část	30354
společnost	29686
možnost	24538
prezident	22709
dost	21435
šest	16589
činnost	13569
skutečnost	11115
růst	11070
projekt	9416
pozornost	9010
místo	8826
účast	8492
smrt	8487
fakt	7981
koncert	7712
řici	7621
st	7280
radost	7160
veřejnost	6894
parlament	6378
město	6056

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

Všimněte si

- *procent, prezident, parlament, ...*
- *společnost, možnost, ...*
- *smrt, fakt, koncert, ...*
- **ALE**
- *část : krást*
- *růst : růst*
- *míst : jíst*

Více

- Podívejte se dále na statistiky.
- Pokuste se stanovit pravidla pro jednotlivá možná zakončení.
- Empirická zjištění porovnejte se znalostmi gramatiky.

Podobně se pokuste najít pravidla např. pro přičestí I-ové

- *bál – sál – vál*
- *jel – šel – žel*
- *děl – těl – pěl*
- *stél - dél*
- *pil – sil – žil*
- *díl – víl*
- *hol – kol – mol*
- *gól – sól – pól*
- *hul – mul – kul*
- *vůl – stul – hůl*
- *kryl – ryl – styl*
- *dýl – týl – kýl*

Jak poznat I-ové příčesí ?

Končí přece na –/ !

- *Hotel nebyl bordel.*
- *Tunel vedl dál.*
- *Výstřel zasáhl cíl.*
- *Musil obešel stůl.*
- *Bohumil Hrabal spadl.*
- *Přítel šel dál.*
- *Král opustil sál.*
- *Živel zpustošil kostel.*

Existuje v češtině zakončení jednoznačně určující gramatické vlastnosti slovního tvaru ?

- Porušený text.
- Automatické formální zpracování textu.
- Zapamatování textu (rým a ústní tradice).

Pokuste se najít nějaké příklady

- Deklinační typy
- Konjugační typy
- Neohebná slova
- Délka slova