

1. V češtině existují určité/finitní a neurčité/infinitní tvary sloves. Může český určitý tvar slovesný končit na *-t*? Odpovězte nejdříve, a pak ověřte svoji odpověď v korpusu SYN2015. Součástí odpovědi musí být formulace dotazu, pomocí kterého byla získána data, na jejichž základě odpovídáte.
  
2. V české flexi platí pravidlo, které tvrdí, že substantiva skloňovaná podle vzoru muž mají ve vokativu singuláru buď koncovku *-i* (*muži*), nebo kcovku *-e* (*otče*). Dále tvrdí, že koncovku *-e* mají substantiva, která končí na /jsou tvořena příponou *-ec*, tedy např. *tupče, pitomče, ...* Existují v češtině i takové případy, kdy substantivum skloňované podle vzoru muž nekončí na *-ec*, a přesto může mít ve vokativu singuláru kromě koncovky *-i* rovněž koncovku *-e*? Odpovězte nejdříve, a pak ověřte svoji odpověď v korpusu SYN2015. Součástí odpovědi musí být formulace dotazu, pomocí kterého byla získána data, na jejichž základě odpovídáte.
  
3. Tvar *spal* je v češtině víceznačný. Jaké má významy (slovní druh a slovnědruhově závislé interpretace tvaru)? Odpovězte nejdříve, a pak ověřte svoji odpověď v korpusu SYN2015. Podívejte se na to, zda je desambiguace v pořádku. Navrhněte, jak byste postupovali, pokud byste chtěli najít nesprávně desambiguované případy. Součástí odpovědi musí být formulace dotazu, pomocí kterého byla získána data, na jejichž základě odpovídáte.

## Řešení:

1. [lc=".\*t" & tag="V[Bi].\*"]

Jedná se o archaický tvar 3.sg.ind.préz. akt. slovesa *být: jest*.

2. [lemma=".\*[cčšřzslj dťň]" & lemma!=".\*(ec)|(us)" & lc=".\*[ei]" & tag="NNMS5.\*"]

Lze ještě např. odstranit n-filtrem všechna substantiva na .\*tel. Pak je potřeba projít seznam ručně.

Najdeme dvojice *anděli/anděle*, *knězi/kněže* a *strýci/strýče*. Přestože jsou případy na *-e* interpretovány jako archaické, v korpusech je frekvence tvarů (alespoň u posledních dvou) vyrovnaná. Celkově to ovšem souvisí s okrajovým postavením vokativu (málo frekventovaný pád, arch. formy jsou zachovány v určitých kolokacích, např. nábož. texty).

3. [lc="spal"]

frekvence\_vlastní\_tag

Filtr	lemma_lc [lowercase lemma]	tag	lc [lowercase word]	Frekvence
1.	p/n spát	VpIS---R-AA---I	spal	952
2.	p/n spát	VpMS---R-AA---I	spal	893
3.	p/n spal	NNMS1----A-----	spal	104
4.	p/n spálit	Vi-S---2--A----P	spal	1

např.

P\_filtr <1 1>

[lc="[th]o"]

nebo např. můžeme zkusit toto:

[tag="[NAPC]...4.\*"]

Výsledek není zdaleka jednoznačný. Lze si povšimnout sémantiky akuzativních doplnění, která následují po tvaru slovesa *spát* (časové údaje). Můžeme si např. zobrazit kandidáty na kolokace:

Collocation candidates

Atribut: lemma\_lc [lowercase lemma]

V kontextu: 1 do 1

Minimální frekvence kolokátu v korpusu: 1

Minimální frekvence kolokátu v kontextu: 1

a uvidíme, že pouze následuje-li zájmeno a substantivum, zůstává možnost víceré interpretace. U zájmen by bylo ještě možné filtrovat podle 2. pozice.