Základy využití korpusu v praxi **cjbb75**

**pondělí: 8.00-10.00 G13**

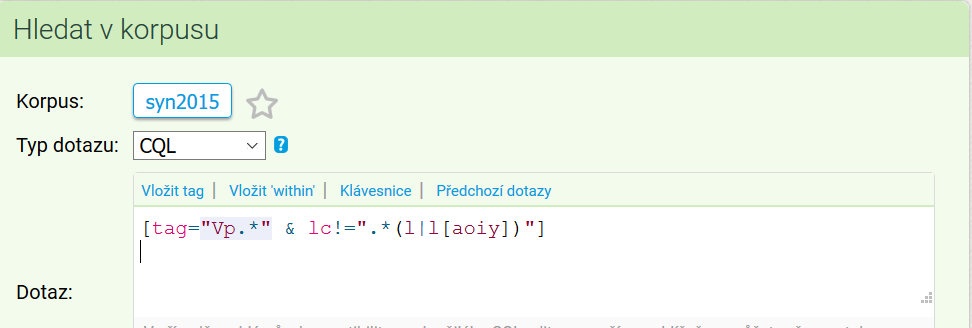
**20. 4. 2020**

1. Je pravda, že l-ová příčestí končí-li v češtině na souhlásku, pak touto souhláskou je nutně vždy ***l***? Odpovězte nejdříve na základě intuice rodilého mluvčího, a pak ověřte svoji intuici v korpusu. SYN2015. Součástí odpovědi musí být formulace dotazu, pomocí kterého byla získána data, na jejichž základě odpovídáte.

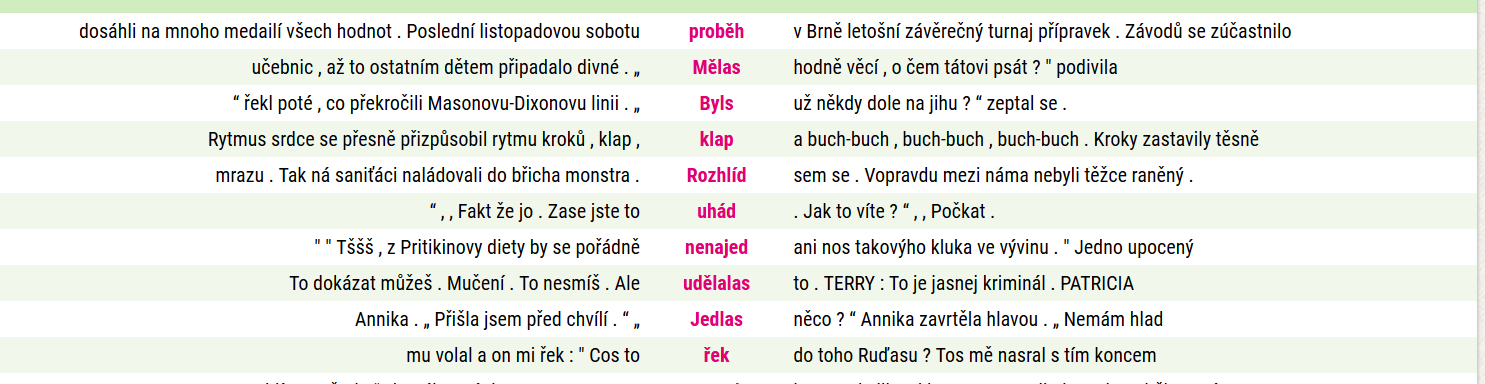
**Napadlo vás, že se stačí zeptat kamaráda, jestli byl včera doma?**

Nebo se podívejte do korpusu:

[tag="Vp.\*" & lc!=".\*(l|l[aoiy])"]



Výsledek KWIC



Podívejme se na menu Frekvence – Slovní tvary [a=A]

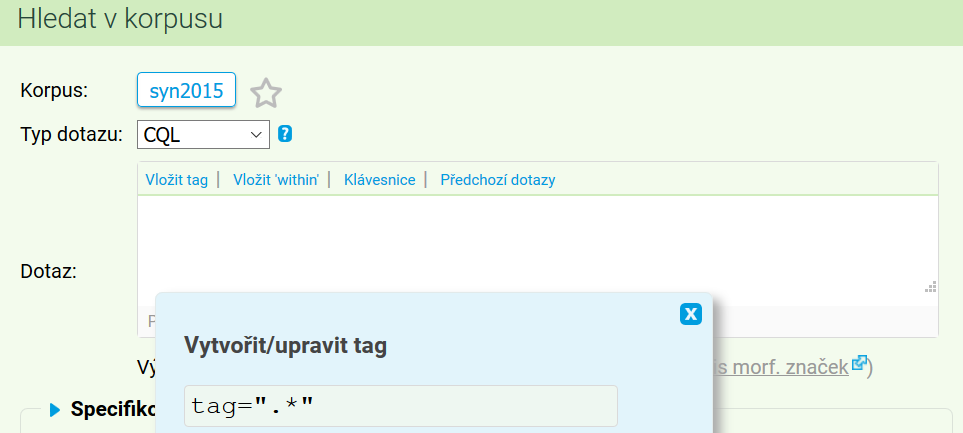


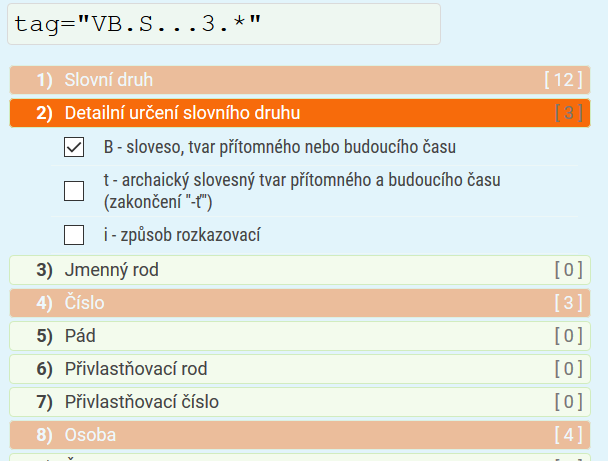
Tvary l-ových příčestí mohou končit na –s (volný morfém signalizující 2. osobu za tvar auxiliáru *být/jsi*). POZOR, jedná se o SPISOVNÝ tvar. Vzhledem k tomu, že jde o tvar 2. osoby singuláru/tykání, užívá se především v neformálních textech, a tak jej mnoho mluvčích pokládá za tvar nespisovný.

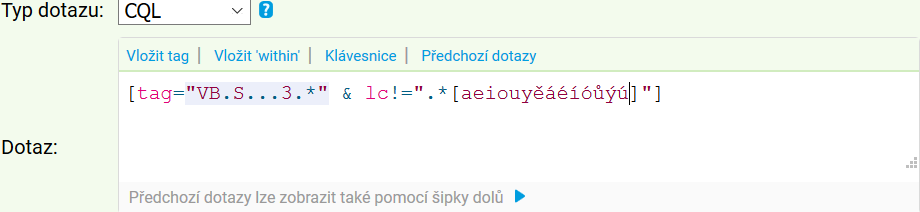
Koncové *l* l-ového příčestí maskulina singuláru může být apokopováno, a sice pouze u sloves s uzavřeným kořenem (končícím na souhlásku) a nulovou kmenotvornou příponou pro tvary l-ových příčestí, tedy pouze slovesa vzoru *nés-0-t/****nes-0-~~l~~*** : *přines, vylez, čet, …, péc-0-i/****pek-0-~~l~~***: *moh, nemoh, řek, utek, …* a variantní *tisk-nou-t/tisk-nu-l/****tisk-0-~~l~~*** : *zahlíd, pad, zved, vypad …*.

1. Znáte tvar 3. osoby sg. indikativu prézenta aktiva, který v češtině nekončí na samohlásku? Odpovězte nejdříve na základě intuice rodilého mluvčího, a pak ověřte svoji intuici v korpusu. SYN2015. Součástí odpovědi musí být formulace dotazu, pomocí kterého byla získána data, na jejichž základě odpovídáte.

Pokud si marně lámete hlavu a říkáte si, že ***–e-, -ne-, -je-, -í-, -á-****,* tak neváhejte a mrkněte do korpusu. Dotaz v cql přes

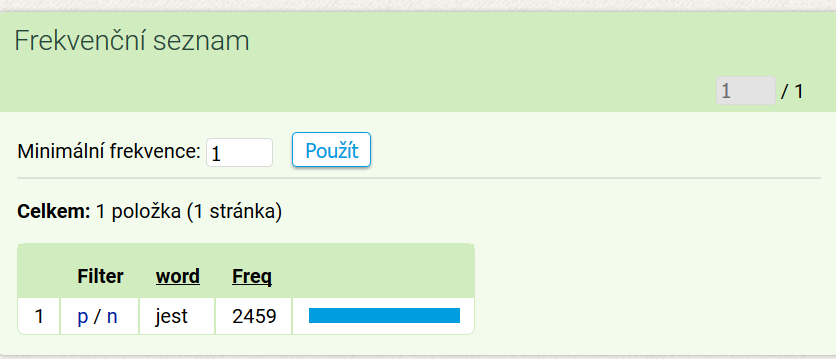








Podívejme se na menu Frekvence – Slovní tvary [a=A]



ANO, TAK **JEST**.

1. Dokázali byste odpověď na následující otázku, kterou by vám položil někdo, kdo se chce naučit česky (cizinec) a která by zněla: Učil jsem se, jak se tvoří tvar l-ového příčestí od sloves, který musím znát, abych uměl vytvořit větu v minulém čase (*šel* *jsem*, …) a v podmiňovacím způsobu (*šel bych, …*). Všiml jsem si, že někdo říká, že *počítač* ***vypl*** a někdo, že *počítač* ***vypnul****.* Jak poznám, když se učím nové sloveso, které má v infinitivu zakončení na *–nout*, jak má vypadat tvar příčestí? 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.

Pokud jste z paměti nevydolovali nějaké poučky o třídách a vzorech, tak můžete rovnou začít hledat odpověď v korpusech a porovnat data vzhledaná dotazem [lemma=".\*nout" & tag="Vp.\*" & lc=".\*(nul)|(nul[aoiy])"]

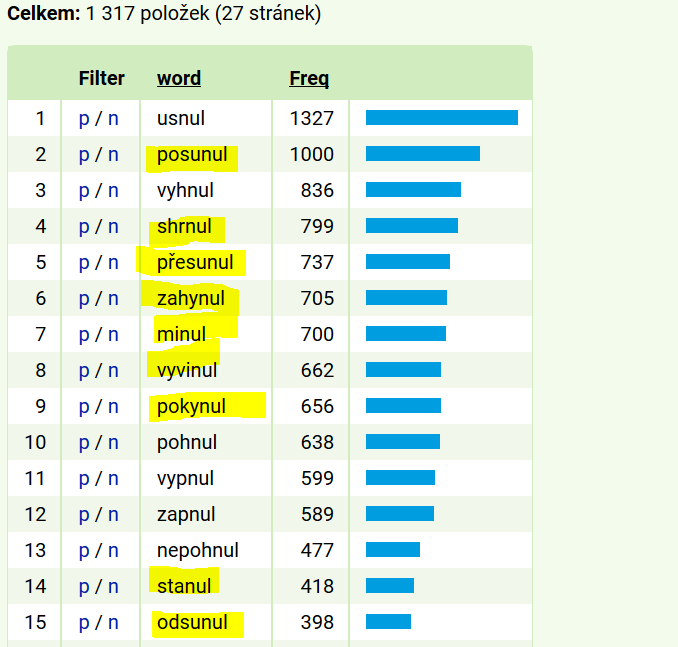


a dotazem [lemma=".\*nout" & tag="Vp.\*" & lc!=".\*(nul)|(nul[aoiy])"].



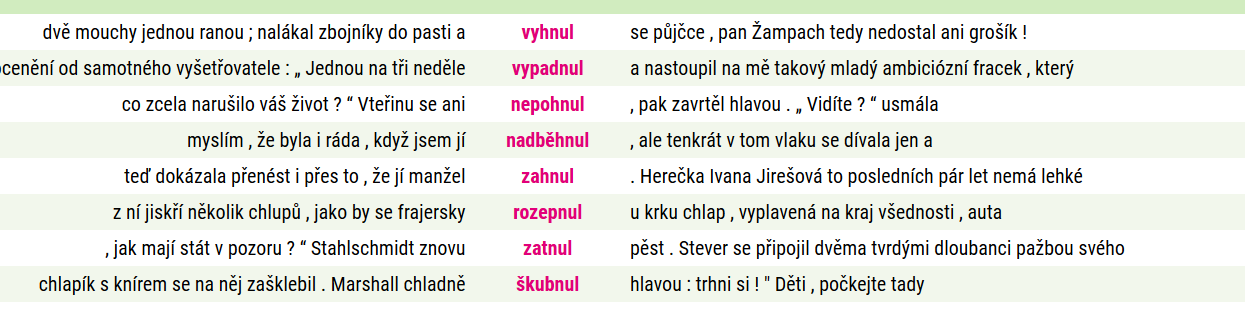


Možná se vám při pozorování prvního seznamu něco o třídách a vzorech vybaví. Ne-li, pak se podívejte, zda nenajdete nějaký společný rys, a to především tehdy, když je vyloučeno, aby tvar měl i variantu bez *–nu-*. Můžeme si všimnout, že varianty bez *–nu-* **nemohou mít slovesa, u nichž před *–nu-* předchází samohláska (nebo slabičná sonora)**.

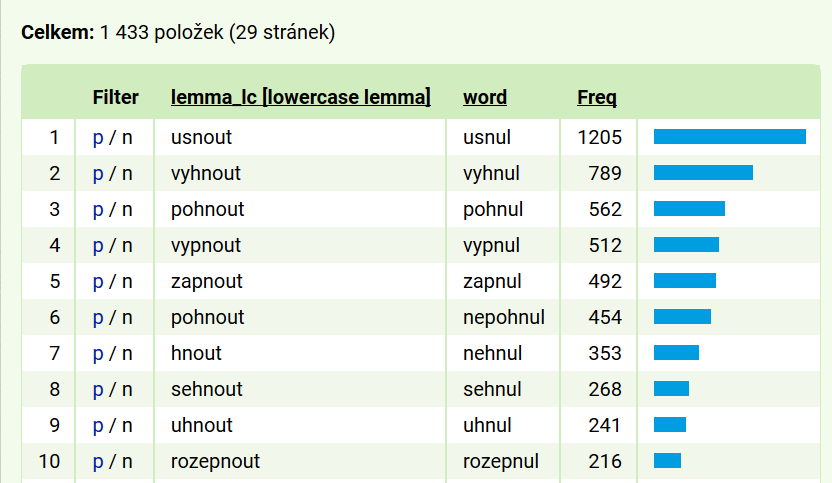
Odfiltrujeme je z první konkordance takto.

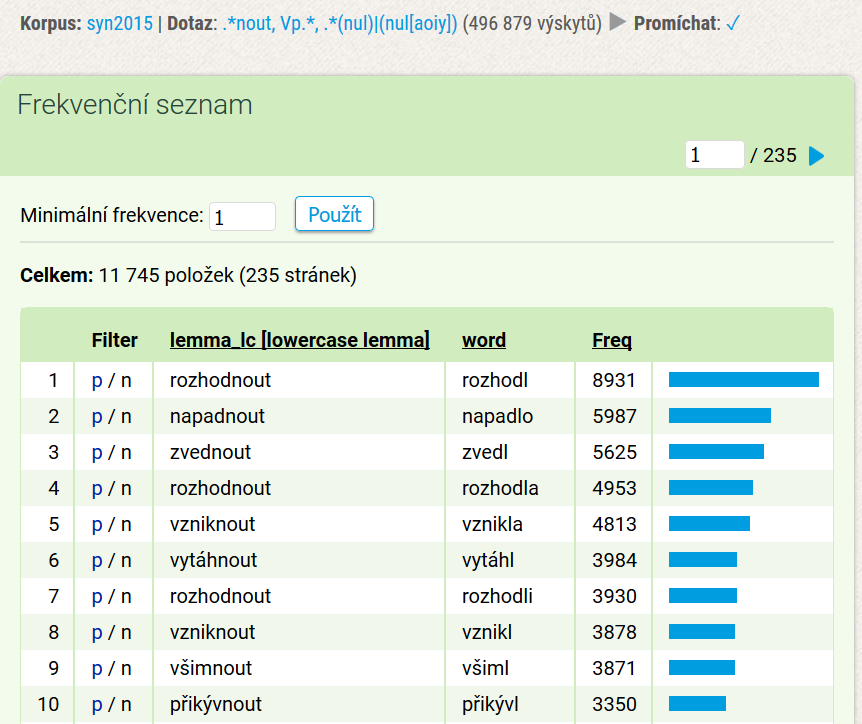
N-filtr <0 0> [lc=".\*[aáeěéiíoóůuyýrlm]((nul)|(nul[aoiy]))"].



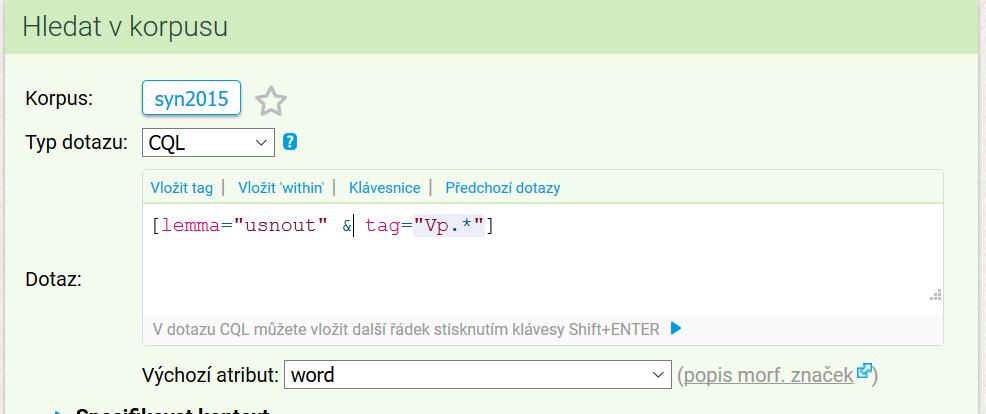


Porovnejme seznamy Frekvence Vlastní – lemma - Word, k nimž jsme našli tvary s *–nu-*

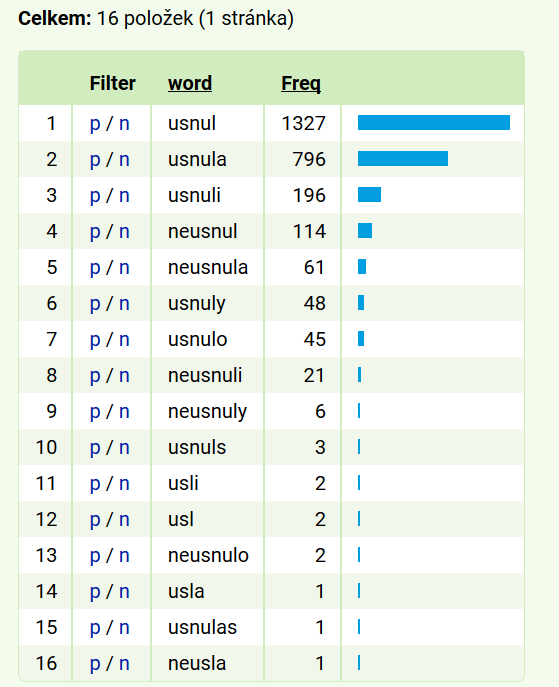




Vyhledáme



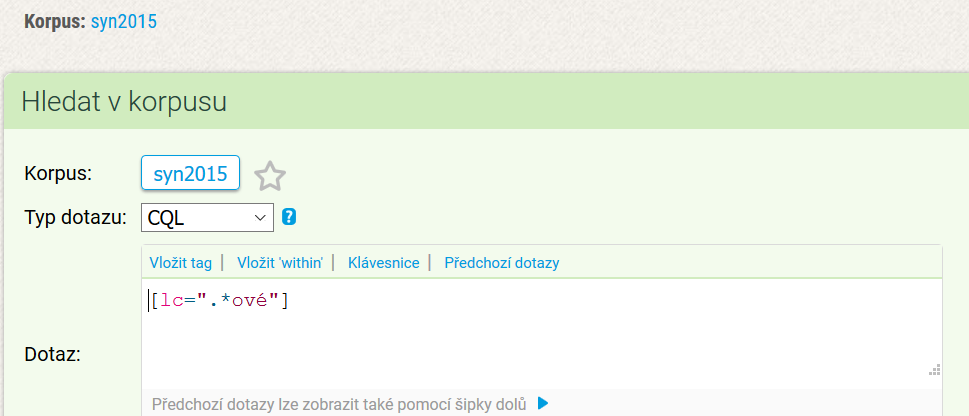
a porovnáme frekvence:

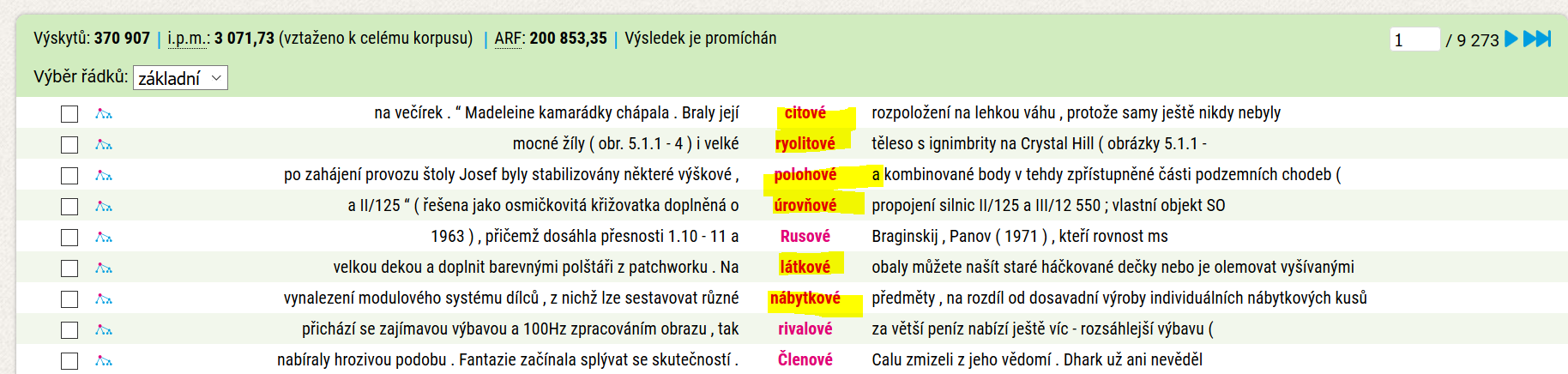


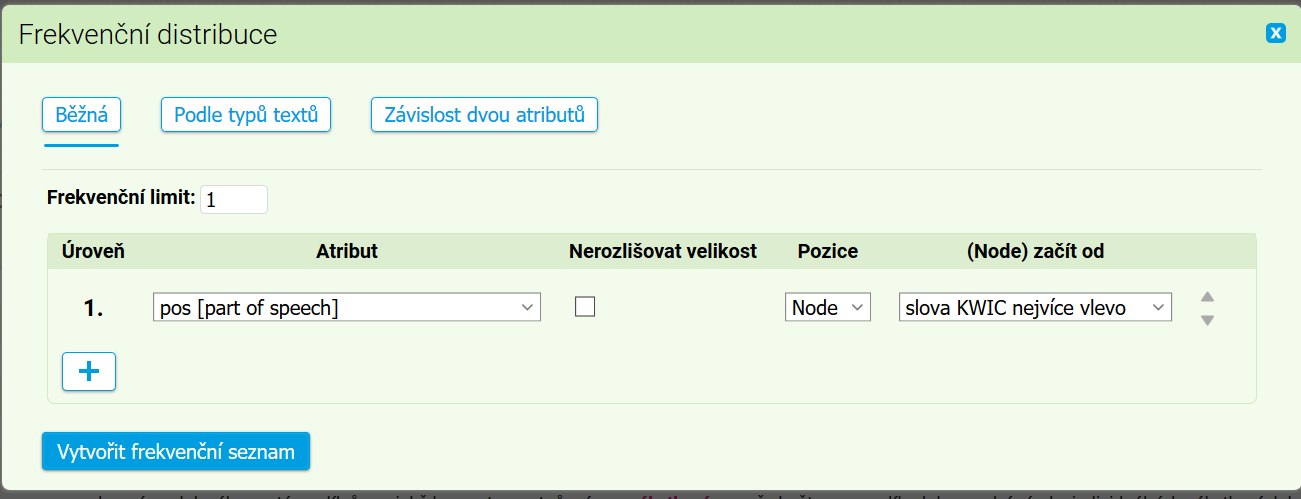
U slovesa *usnot* se více užívají tvary s *–nu-*. Takto můžeme pokračovat a sledovat jednotlivá slovesa.

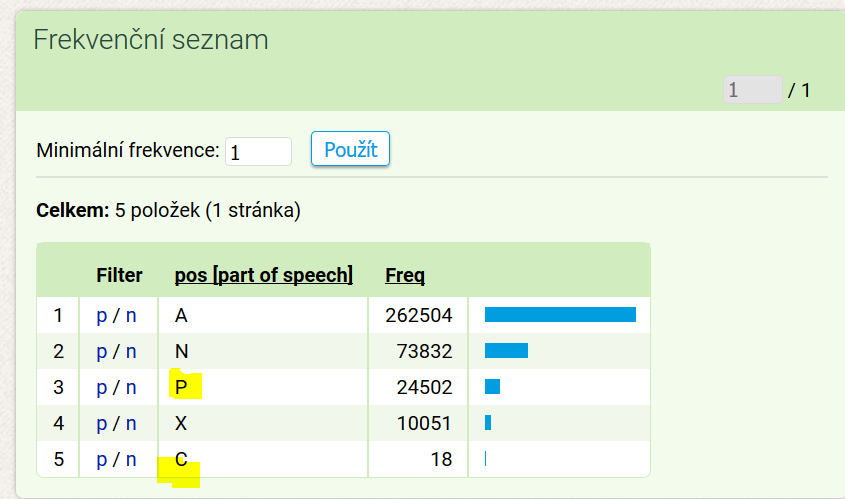
1. Končí-li tvar slova v češtině na ***–ové***, pak jde nutně o tvar životného maskulina. Platí toto tvrzení? Odpovězte nejdříve, a pak ověřte svoji odpověď v korpusu SYN2015. Součástí odpovědi musí být formulace dotazu, pomocí kterého byla získána data, na jejichž základě odpovídáte.

Neplatí, protože toto zakončení mají pravidelně tvary adjektiv s lemmatem na *.\*ový*. Pokud zadáme dotaz [lc=".\*ové"],



zjistíme, přes frekvenci pos, že jsme se nemýlili

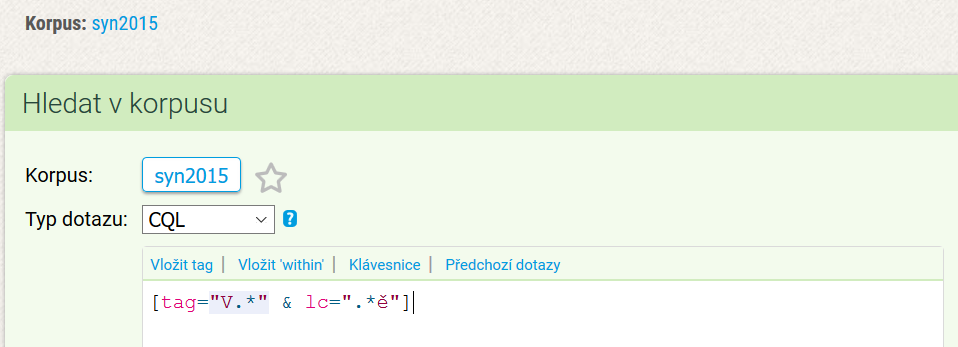




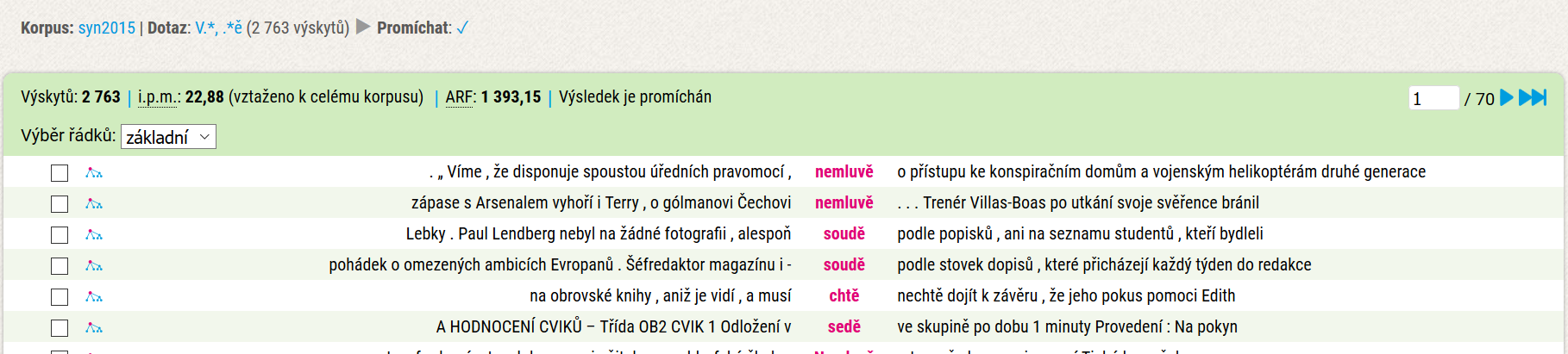
a že tvary adjektivní flexe, a s nimi i zakončení *–ové* mají i slova tagovaná jako zájmena (*takové*) a číslovky (*třináctimilionový*).

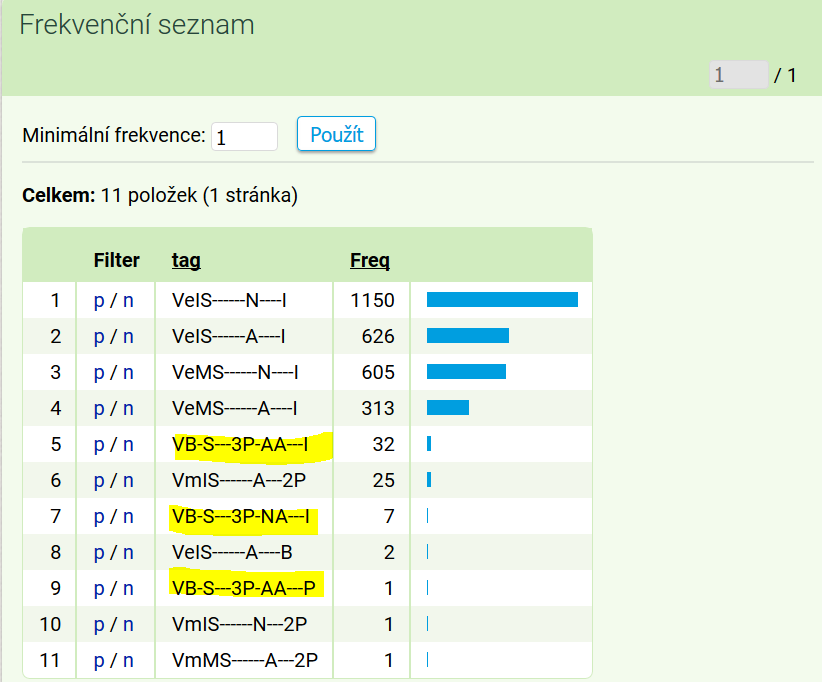
1. Hláska *e* má v češtině dvě grafické realizace: *e/ě*. Existuje český slovesný tvar, který končí na grafickou realizaci *ě*? Odpovězte nejdříve, a pak ověřte svoji odpověď v korpusu SYN2015. Součástí odpovědi musí být formulace dotazu, pomocí kterého byla získána data, na jejichž základě odpovídáte.

Pokud jste odpověděli NE, tak je to špatně. Pokud ANO, ale nevím, a pak jste zadali dotaz: [tag="V.\*" & lc=".\*ě"],



pak jste patrně zjistili, že jde buď o tvary přechodníků (*nemluvě, nechtě, …*), ale narazili jste i na tvar určitý (*stůně*).





N-filtrem jste přechodníky odstranili (N-filtr <0 0> [tag="Ve.\*"] a zjistili jste, že automatická analýza patrně rozpoznává potenciální (nekodifikované tvary) přechodníků přítomných od dokonavých sloves, které ovšem značkuje jako přechodníky minulé. N-filtrem jste přechodníky odstranili (N-filtr <0 0> [tag="Vm.\*"] a zjistili jste, že v korpusu SYN2015 se vyskytuje pouze jeden kořen, od něhož určitý tvar slovesa končící na *–ě: stůn-ě/ne- stůn-ě/vy- stůn-ě.*

Pokud jste na vše přišli bez korpusu, gratuluju ☺.