Stránka webového rozhraní KonText

Budeme pracovat s korpusem SYN2020 (první hodinu si ukážeme, jak ho nastavit)

Obsah obrázku text

Popis byl vytvořen automaticky

|  |  |
| --- | --- |
| DOTAZOVACÍ ŘÁDEK | POKYN, NA KTERÝ KLIKNEME, KDYŽ JSME NAPSALI CELÝ DOTAZ DO DOTAZOVACÍHO ŘÁDKU A CHCEME, ABY NÁM PROGRAM VYHLEDAL, TO, CO JSME ZADALI. |

**Vysvětlení podstaty pozičních atributů a druhů dotazu**

Pokročilý dotaz – ne: tečka (manipulační prvek) je nalevo a je šedá.

Obsah obrázku text

Popis byl vytvořen automaticky

Zpočátku budeme pracovat bez pokročilých dotazů, to znamená, že si vždy zvolíme „výchozí atribut“ a do dotazovacího řádku napíšeme slovo, které hledáme.

Základní atributy, s kterými budeme pracovat první 2 semináře:

(sublemma, lemma lc, sublemma lc, pos aj. si vysvětlíme později)

* word – jeden konkrétní slovní tvar (např. *kočkou*, *modrou*, *píše*)

Vybereme-li si jako výchozí atribut word, můžeme zadat jakýkoliv tvar slova, korpus pak vyhledá právě jen tento tvar.

Příklad: zvolíme výchozí atribut word, do dotazového řádku napíšeme slovo *modrou*, odešleme dotaz pokynem Hledat; KonText nám vyhledá pouze tvar modrou.

* [lemma](http://wiki.korpus.cz/doku.php/pojmy:lemma) - základní (slovníkový) tvar (např. kočka, modrý, psát)

Vybereme-li si jako výchozí atribut [lemma](http://wiki.korpus.cz/doku.php/pojmy:lemma), musíme zadat vždy základní tvar slova, tj. nom. sg. u substantiv a bezrodých zájmen, nom. sg. mask. u adjektiv a taky u zájmen, popř. číslovek, které rozlišují rod, infinitiv u sloves, u adjektiv to musí být vždy positiv (1. stupeň).

Vybereme-li si jako výchozí atribut [lemma](http://wiki.korpus.cz/doku.php/pojmy:lemma) a zadáme základní tvar slova, např. *modrý*, korpus nám vyhledá všechny tvary tohoto slova – všechny pády v obou číslech, u adjektiv i tvary 2. a 3. stupně, u sloves i pasivní příčestí a přechodníky.

* [sublemma](https://wiki.korpus.cz/doku.php/pojmy:lemma#sublemma)

Lemma může zahrnovat i více variant jednoho slova (např. lemma filozofie reprezentuje všechny tvary se základem filozof i filosof), sublemmata zahrnují vždy variantu:

* sublemma filozofie zahrnuje jen tvary se „z“
* sublemma filosofie zahrnuje jen tvary se „s“
* pokud je slovo nevariantní, sublemma je totožné s lemmatem
* [tag](http://wiki.korpus.cz/doku.php/pojmy:tag) – složitější atribut; značka zachycující morfologickou (příp. i jinou, obvykle gramatickou) informaci. S tímto atributem pracujeme jen v rámci „pokročilých dotazů“, proto si jeho užívání popíšeme později.

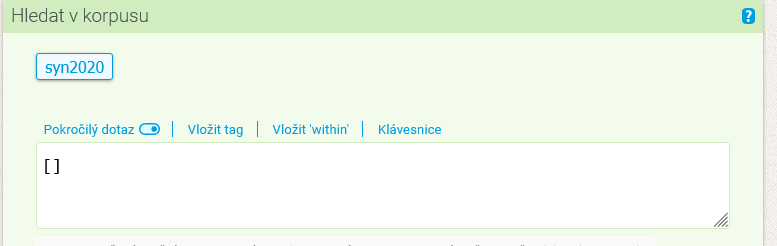
Někdy si nevystačíme se základními dotazy a potřebujeme „pokročilý dotaz“, např. při hledání slov s jedním sufixem. Chceme např. zjistit, která slova tvořená sufixem *.\*ina* se v češtině užívají a která jsou nejčastější.

V tom případě je nejjednodušší zadat dotaz:

[lemma=".\*ina"]

Jak se k dotazu dopracujeme?

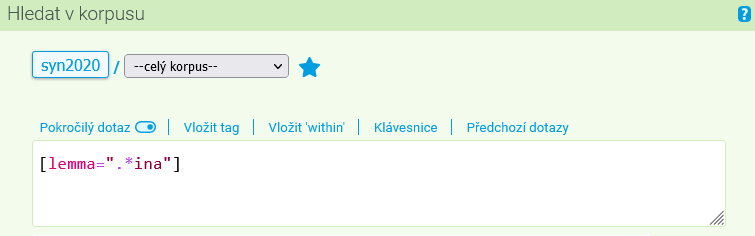
1. Využijeme příkaz Klávesnice a zadáme hranaté závorky.



Musí být „zakliknutý“ povel Pokročilý dotaz.

2. Do hranatých závorek vepíšeme přesně: lemma=".\*ina"

Tečka s hvězdičkou znamená řetězec písmen.



3. Co se objeví:



Co teď s tím?

1. Vytvoříme seznam podle lemmat. Na prvním semináři si vysvětlíme, jak tento seznam vytvoříme.



1. Vyloučíme neutvořená slova (*hodina*, *vina*, *vteřina*…).

A co když chceme vyhledat jen případy, kdy je homonymní slovo určitým slovním druhem?

Příklad: Vyhledejte doklady, ve kterých je slovo mezi pouze podstatným jménem.

Postup: Využijeme „pokročilý dotaz“

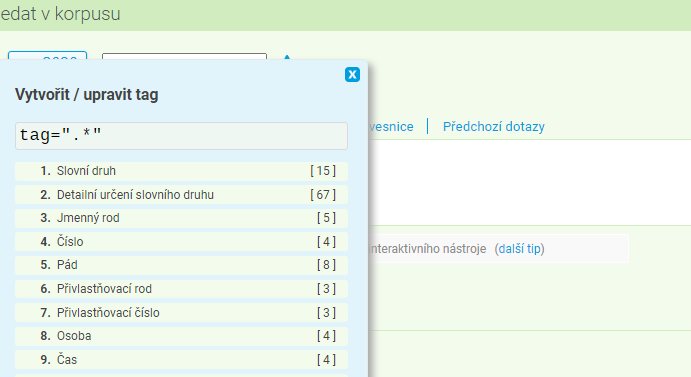
1. Zadáme: word = „mezi“

Obsah obrázku text

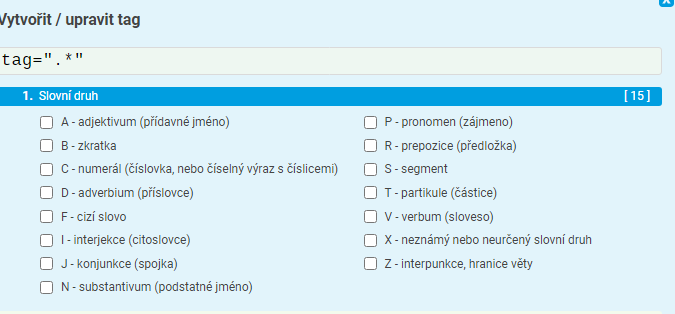
Popis byl vytvořen automaticky

1. Využijeme odkaz „Vložit tag“

Po kliknutí na odkaz se objeví nabídka, v níž je na prvním místě „slovní druh“. Pozor, kurzor mějte umístěný v hranatých závorkách za slovem *mezi* v uvozovkách.



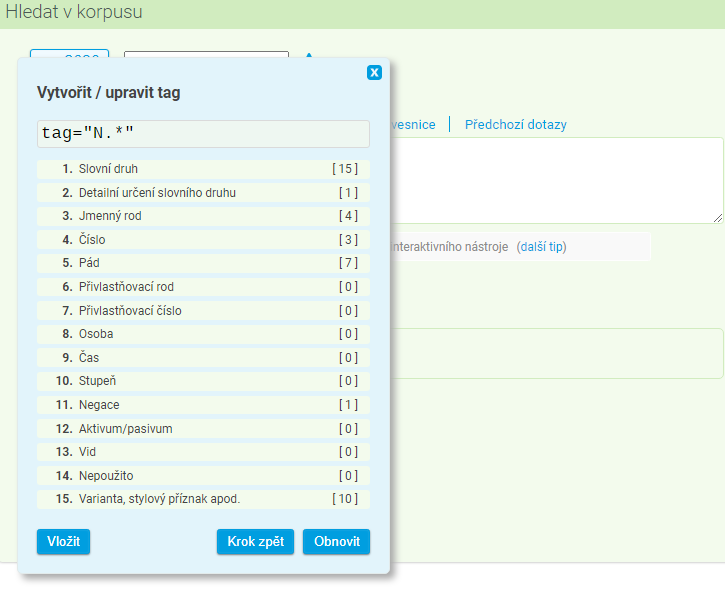
1. Klikněte na Slovní druh, objeví se další nabídka:



Ve čtverečku zaklikněte „substantivum“

Pak klikněte na tmavomodrý proužek „1. Slovní druh“, nabídka slovních druhů se sbalí.

1. Objeví se znovu seznam tagů, vy klikněte na „Vložit“.



Objeví se následující:

Obsah obrázku text

Popis byl vytvořen automaticky

Smažte hranaté závorky kolem tagu, abyste měli následující:

Obsah obrázku text

Popis byl vytvořen automaticky

1. Před slovo tag vložte znak & (pomocí klávesnice, nebo „ALT + 38“)-

Dostanete následující dotaz.

Obsah obrázku text

Popis byl vytvořen automaticky

1. Po odeslání dotazu dostaneme 60 řádků

Obsah obrázku text

Popis byl vytvořen automaticky