

Cvičení:

1. Najděte v korpusu **synv12** svoje křestní jméno, příjmení, příjmení svého otce.

(Pozorujte rozdíl mezi *proprii* jako *Černý, Hrdlička, Hříbek, Janík, Krejčířík, Kuchař, Pospíšil, Řezáč, Spáčil, Stolař, Zanáška, Zpěvák* a *apelativy*).

2. Najděte v korpusu **synv12** všechny formy hypokoristik (domácích jmen) svého křestního jména.

3. Najděte v korpusu **synv12** slova, která se liší v pravopise (i/y – např. *bidlo/bydlo*). Zkuste najít chyby.

4. Najděte v korpusu **synv12** slova, která se liší v pravopise adaptace anglických slov na *ing/ink*. ([word=".*in[gk].*" & word!="[:upper:].*" & pos="X"])

Zadání DÚ na 4. 3. 2024

Nezapomeňte citovat korpus!

1. Najděte v korpusu SYN2020 slovní tvar (atribut **word**) *jedle*.

Jak je tvar interpretován lemmatem a tagem?

Najděte v korpusu SYN2020 slovní spojení *příliš jedle, vypadá jedle*.

Jak je tvar interpretován lemmatem a tagem?

2. Vyhledejte v korpusu (**synv12**) všechny tvary slova *brno* psané s malým počátečním písmenem. Použijte funkci *Frekvence/Vlastní/doc.txtype_group* a podívejte, jak je malé písmeno užíváno v beletrii. Svá pozorování popište.

3. Všichni známe (a někdy i používáme) hláskovou variantu [í]/[ej], jako třeba *obývák/obejvák, vylít/vylejt*. Vyhledejte v korpusu (**synv12**) slovo (lemma) *slívák* a *slejevák*. Jde o uvedený příklad.

Vzpomenete si ještě na nějaký podobný příklad? Pokud ne, pokuste se sestavit dotaz, který by vám mohl pomoci najít takový příklad v korpusu?

4. Sestavte dotaz tak, abyste v korpusu vyhledali všechna slovesa, která patří k typu *kupovat* a zahrnují konsonanty, které se v šeských slovech nevyskytují (nebo se vyskytují zřídka). Uvědomte si, že půjde o slova přejatá, která nemusí být zařazena do slovníku morfologického analyzátoru, tudíž nebudou lemmatizována. Je tedy třeba dotaz sestavit tak, abyste vyhledali všechny tvary.