

(Vaše) úkoly – 11. dubna 2012

1. Vyhledejte slova, která mají:

a) *ú* uprostřed slova,

b) *ú* na konci slova.

Jaké absolutní frekvence výskytu mají?

2. Jak vyhledáme v korpusu SYN2005 zkratku *č. p.* (číslo popisné)?

3. Jakou koncovkou končí česká slovesa ve 3. os. sg. přítomnosti?

4. Jaké slovo v češtině obsahuje nejvíce výskytů písmene *v*?

5. Vyhledejte všechny tvary slovního spojení *být s to*. Zjistěte absolutní frekvenci výskytu tohoto spojení.

6. Vyhledejte spojení typu *udělají šéfem kriminálky* (tedy sloveso ve 3. os. č. mn., za kterým bude následovat substantivum mužského životného rodu v jednotném čísle a v 7. pádě a za ním další substantivum v 2. pádě). Kolik těchto spojení korpus našel?

7. Jaký dotaz musíme zadat, abychom v korpusu SYN2005 vyhledali věty zvolací, které obsahují pouze jedno slovo?

8. Podle *Příruční mluvnice češtiny* projevují substantiva ženského rodu tvořená sufixem *-yně* (např. *žákyně*) v genitivu plurálu tendenci k nulovému zakončení (*bez žákyň*), a to na rozdíl od jmen neosobních, a tedy neutvořených sufixem *-yně* (např. *jeskyně – bez jeskyní*). Ověřte pravdivost tohoto výroku v korpusu.

9. Vyhledejte tázací věty, ve kterých došlo (z důvodu aktuálního větného členění) k posunu tázacího slova do jiné než iniciální pozice (např. *Listonoš kdy tu byl naposled? Listonoš tu byl naposled kdy?*). V jaké pozici se takto „posunutý“ výraz objevuje nejčastěji?

10. Vyhledejte spojení *se ženou*, porovnejte frekvenci případů, kdy jde o sloveso (*hnát se*) a kdy o substantivum (*žena*). Podívejte se na slovesa, jestli jsou opravdu správně označovaná. Vyzkoušejte si to postupně v korpusech syn2005 a syn2010, porovnejte chybovost morfologického značkování.