**2. seminář**

**Kontrola domácí přípravy**

1. Srovnejte podle korpusu SYN2020 výskyt variant předložek *pod – pode, nad – nade*. Vždy si pak prohlédněte cca prvních 50 řádků s rozšířenými variantami *pode*, *nade* a pokuste se vyvodit, proč autor textu užil rozšířenou variantu předložky.

Řešení:

Využijeme základní dotaz s výchozím atributem **word**:



**Řešení**

Frekvence:

pod 72242 pode 374

nad 62356 nade 1296

Autoři volí rozšířenou podobu předložky:

– při spojení se slovem začínajícím na *d*: nade dveřmi, nade dnem; pode dveřmi, pode dvěma, pode dnem

– ve spojení s tvarem zájmena *já*: nade mnou, nade mě, pode mnou, pode mě

– v ustálených spojeních: nade vší pochybnost, nade vše, nade všechno, nade vším, nade všechno

2. V Internetové jazykové příručce si na odkaze <https://prirucka.ujc.cas.cz/?id=121> přečtěte o zásadách pravopisu *s* – *z* v přejatých slovech. Pak vyhledejte v korpusu SYN2020 dvojice variant, u každé z variant si poznamenejte frekvenci:

*impuls – impulz*

*resort – rezort*

Na základě zjištěné frekvence posuďte, zda skutečné užívání variant odpovídá tomu, co je uvedeno v Internetové jazykové příručce.

Řešení:

Využijeme základní dotaz s výchozím atributem **sublemma**:



**Výsledky:**

impuls 902 – impulz 1208 poměr frekvencí je cca 3 : 4

resort 1902 – rezort 341 poměr frekvencí je cca 6 : 1

¨

Zhodnocení výsledků hledání v korpusu a srovnání s tvrzeními v Internetové jazykové příručce:

Frekvence variant *impuls* a *impulz* se výrazně neliší. Výsledky hledání v SYN2020 potvrdily, že jde o rovnocenné varianty.

Vývoj výskytu variant *resort* – *rezort*:

**resort rezort**

| **doc.syn** | **Freq** | **i.p.m.** | **Freq** | **i.p.m.** | **poměr výskytu (vycházíme z i.p.m.)** |
| --- | --- | --- | --- | --- | --- |
| 2000 | 3 809 | 34,61 | 253 | 2,3 | 15 : 1 |
| 2005 | 1 495 | 12,48 | 263 | 2,2 | 6 : 1 |
| 2010 | 1 178 | 9,98 | 134 | 1,14 | 9 : 1 |
| 2015 | 1 708 |  14,08 | 262 | 2,16 | 7 : 1 |
| 2020 | 1 902 | 15,64 | 341 | 2,8 | 6 : 1 |

Je patrné výrazné snížení rozdílu od korpusu SYN2005, tj. v textech po r. 2000.

3. V Internetové jazykové příručce si na odkaze <https://prirucka.ujc.cas.cz/?id=124> přečtěte o zásadách pravopisu slov s druhou částí *-manie/-mánie, -fuze/-fúze*.

Pak srovnejte podle Českého národního korpusu SYN2020 frekvenci variant:

*transfuze – transfúze*

*toxikomanie – toxikománie*

Řešení:

Využijeme dotaz sublemma:



**Výsledky:**

transfuze 163 – transfúze 64 poměr frekvencí: 5 : 2 (2,5 : 1)

toxikomanie 18 – toxikománie 4 poměr frekvencí: 9 : 2 (4,5 : 1)

Převládá varianta s krátkou samohláskou, nikdy ale výrazněji než v poměru 4,5 : 1.

3. Srovnejte v korpusu SYN2020 frekvenci termínů–ekvivalentů zeměpis a geografie.

Řešení: Využijeme dotaz lemma.



zeměpis: 234

geografie: 271 Výskyt obou ekvivalentů je vyrovnaný.

4. Vyhledejte v korpusu SYN2020 slova s příponou *-yně* a pokuste se vysledovat, jaký význam nejčastěji mají.

**Který typ dotazu zvolíte?**



Objevilo se 339 položek, zaměříme se na prvních 50: většina z nich jsou názvy PŘECHÝLENÉ = 42 názvů. 2 názvy jsou tvořeny jiným způsobem (kuchyně, svatyně), 2 slova jsou nemotivovaná (bekyně, jeskyně), 4 jsou vlastními jmény.

Začátek formuláře

|  | **Filtr** | **lemma** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- | --- |
| 1 | p / n | kuchyně | 15 631 | 128,31 |
| 2 | p / n | přítelkyně | 4 289 | 35,21 |
| 3 | p / n | jeskyně | 4 135 | 33,94 |
| 4 | p / n | kolegyně | 2 297 | 18,86 |
| 5 | p / n | ministryně | 2 036 | 16,71 |
| 6 | p / n | bohyně | 1 523 | 12,5 |
| 7 | p / n | předsedkyně | 1 267 | 10,4 |
| 8 | p / n | svatyně | 1 108 | 9,1 |
| 9 | p / n | tchyně | 954 | 7,83 |
| 10 | p / n | zástupkyně | 904 | 7,42 |
| 11 | p / n | soudkyně | 791 | 6,49 |
| 12 | p / n | hospodyně | 770 | 6,32 |
| 13 | p / n | mistryně | 679 | 5,57 |
| 14 | p / n | umělkyně | 644 | 5,29 |
| 15 | p / n | poradkyně | 614 | 5,04 |
| 16 | p / n | náměstkyně | 608 | 4,99 |
| 17 | p / n | poslankyně | 550 | 4,52 |
| 18 | p / n | zaměstnankyně | 418 | 3,43 |
| 19 | p / n | žákyně | 415 | 3,41 |
| 20 | p / n | místopředsedkyně | 409 | 3,36 |
| 21 | p / n | vévodkyně | 409 | 3,36 |
| 22 | p / n | průvodkyně | 378 | 3,1 |
| 23 | p / n | svědkyně | 290 | 2,38 |
| 24 | p / n | pěvkyně | 267 | 2,19 |
| 25 | p / n | Ruzyně | 254 | 2,09 |
| 26 | p / n | sportovkyně | 250 | 2,05 |
| 27 | p / n | běžkyně | 238 | 1,95 |
| 28 | p / n | otrokyně | 221 | 1,81 |
| 29 | p / n | důchodkyně | 220 | 1,81 |
| 30 | p / n | žalobkyně | 206 | 1,69 |
| 31 | p / n | vražedkyně | 202 | 1,66 |
| 32 | p / n | vědkyně | 200 | 1,64 |
| 33 | p / n | předchůdkyně | 186 | 1,53 |
| 34 | p / n | obhájkyně | 181 | 1,49 |
| 35 | p / n | Purkyně | 177 | 1,45 |
| 36 | p / n | vůdkyně | 176 | 1,45 |
| 37 | p / n | velvyslankyně | 166 | 1,36 |
| 38 | p / n | dramaturgyně | 165 | 1,35 |
| 39 | p / n | ochránkyně | 159 | 1,31 |
| 40 | p / n | Kdyně | 158 | 1,3 |
| 41 | p / n | správkyně | 151 | 1,24 |
| 42 | p / n | svěřenkyně | 146 | 1,2 |
| 43 | p / n | dárkyně | 143 | 1,17 |
| 44 | p / n | Bechyně | 137 | 1,13 |
| 45 | p / n | bekyně | 128 | 1,05 |
| 46 | p / n | chodkyně | 124 | 1,02 |
| 47 | p / n | dozorkyně | 117 | 0,96 |
| 48 | p / n | vězeňkyně | 116 | 0,95 |
| 49 | p / n | strážkyně | 114 | 0,94 |
| 50 | p / n | sokyně | 112 | 0,92 |

**Další typ pokročilého dotazu – tag**

Umožňuje vyhledat slovo podle gramatických vlastností.

Nejjednodušší příklad **Vyhledejte v SYN2020 předložky, které v něm jsou doloženy.**

**Postup:**

1. Zaklikneme „pokročilý dotaz“ 2. Zvolíme „Vložit tag“



3. Vybereme slovní druh Prepozice a odklikneme pokyn „Vložit“.



Vytvoří se dotaz:



4. Vyhledáme doklady a vytvoříme frekvenčí seznam na základě nástroje Frekvence – Lemmata.

|  | **Filtr** | **lemma** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- | --- |
| 1 | p / n | v | 2 306 782 | 18 934,93 |
| 2 | p / n | na | 1 871 117 | 15 358,83 |
| 3 | p / n | s | 967 119 | 7 938,48 |
| 4 | p / n | z | 852 498 | 6 997,62 |
| 5 | p / n | do | 690 058 | 5 664,26 |
| 6 | p / n | o | 579 327 | 4 755,33 |
| 7 | p / n | k | 545 155 | 4 474,84 |
| 8 | p / n | pro | 390 147 | 3 202,47 |
| 9 | p / n | za | 379 470 | 3 114,83 |
| 10 | p / n | po | 333 165 | 2 734,74 |
| 11 | p / n | od | 257 296 | 2 111,98 |
| 12 | p / n | u | 202 875 | 1 665,27 |
| 13 | p / n | při | 148 470 | 1 218,7 |
| 14 | p / n | před | 145 780 | 1 196,62 |
| 15 | p / n | podle | 118 913 | 976,08 |
| 16 | p / n | mezi | 114 590 | 940,6 |
| 17 | p / n | bez | 90 850 | 745,73 |
| 18 | p / n | pod | 79 138 | 649,59 |
| 19 | p / n | nad | 67 898 | 557,33 |
| 20 | p / n | přes | 64 698 | 531,07 |
| 21 | p / n | proti | 53 821 | 441,78 |
| 22 | p / n | kolem | 46 974 | 385,58 |
| 23 | p / n | kvůli | 42 560 | 349,35 |
| 24 | p / n | během | 41 969 | 344,5 |
| 25 | p / n | díky | 40 362 | 331,31 |
| 26 | p / n | kromě | 30 630 | 251,42 |
| 27 | p / n | vedle | 25 415 | 208,62 |
| 28 | p / n | mimo | 21 226 | 174,23 |
| 29 | p / n | včetně | 21 084 | 173,07 |
| 30 | p / n | místo | 16 533 | 135,71 |
| 31 | p / n | pomocí | 16 456 | 135,08 |
| 32 | p / n | vzhledem | 12 745 | 104,62 |
| 33 | p / n | vůči | 12 564 | 103,13 |
| 34 | p / n | prostřednictvím | 10 251 | 84,14 |
| 35 | p / n | dle | 9 116 | 74,83 |
| 36 | p / n | uprostřed | 8 826 | 72,45 |
| 37 | p / n | oproti | 7 645 | 62,75 |
| 38 | p / n | okolo | 7 118 | 58,43 |
| 39 | p / n | navzdory | 6 534 | 53,63 |
| 40 | p / n | skrz | 5 938 | 48,74 |
| 41 | p / n | uvnitř | 5 546 | 45,52 |
| 42 | p / n | podél | 5 534 | 45,43 |
| 43 | p / n | koncem | 4 441 | 36,45 |
| 44 | p / n | naproti | 4 263 | 34,99 |
| 45 | p / n | ohledně | 3 976 | 32,64 |
| 46 | p / n | nedaleko | 3 502 | 28,75 |
| 47 | p / n | poblíž | 3 201 | 26,28 |
| 48 | p / n | namísto | 2 884 | 23,67 |
| 49 | p / n | napříč | 2 466 | 20,24 |
| 50 | p / n | začátkem | 2 364 | 19,41 |
| 51 | p / n | blízko | 2 271 | 18,64 |
| 52 | p / n | zpod | 1 511 | 12,4 |
| 53 | p / n | počátkem | 1 486 | 12,2 |
| 54 | p / n | zpoza | 1 396 | 11,46 |
| **55** | **p / n** | **za|co** | **1 279** | **10,5** |
| 56 | p / n | počínaje | 1 204 | 9,88 |
| 57 | p / n | vstříc | 1 040 | 8,54 |
| **58** | **p / n** | **o|co** | **976** | **8,01** |
| 59 | p / n | versus | 745 | 6,12 |
| 60 | p / n | doprostřed | 743 | 6,1 |
| **61** | **p / n** | **na|co** | **691** | **5,67** |
| 62 | p / n | blíž | 597 | 4,9 |
| 63 | p / n | vyjma | 589 | 4,84 |
| 64 | p / n | dovnitř | 501 | 4,11 |
| 65 | p / n | vně | 416 | 3,42 |
| 66 | p / n | zkraje | 360 | 2,96 |
| 67 | p / n | nejblíž | 296 | 2,43 |
| 68 | p / n | napospas | 263 | 2,16 |
| 69 | p / n | vzdor | 222 | 1,82 |
| 70 | p / n | konče | 217 | 1,78 |
| 71 | p / n | vprostřed | 192 | 1,58 |
| 72 | p / n | via | 161 | 1,32 |
| 73 | p / n | zevnitř | 156 | 1,28 |
| 74 | p / n | narozdíl | 116 | 0,95 |
| 75 | p / n | ob | 81 | 0,67 |
| 76 | p / n | kontra | 79 | 0,65 |
| 77 | p / n | stran | 69 | 0,57 |
| 78 | p / n | na|on | 50 | 0,41 |
| 79 | p / n | prostřed | 26 | 0,21 |
| 80 | p / n | zvíci | 21 | 0,17 |
| 81 | p / n | zprostřed | 15 | 0,12 |
| 82 | p / n | vpůli | 10 | 0,08 |
| 83 | p / n | pro|on | 10 | 0,08 |
| 84 | p / n | vůkol | 8 | 0,07 |
| 85 | p / n | zpředu | 7 | 0,06 |
| 86 | p / n | vevnitř | 6 | 0,05 |
| **87** | **p / n** | **o|on** | **5** | **0,04** |
| 88 | p / n | nehledě | 5 | 0,04 |
| **89** | **p / n** | **v|co** | **5** | **0,04** |
| 90 | p / n | poza | 3 | 0,03 |
| 91 | p / n | zponad | 3 | 0,03 |
| **92** | **p / n** | **na|copak** | **3** | **0,03** |
| **93** | **p / n** | **za|on** | **2** | **0,02** |
| 94 | p / n | spod | 2 | 0,02 |
| **95** | **p / n** | **do|on** | **1** | **0,01** |
| 96 | p / n | zpomezi | 1 | 0,01 |

Co lze vyčíst?

* počet předložek v češtině
* počet předložek vlastních a nevlastních: vlastních předložek je 18, ostatní jsou nevlastní, není jich 76. Musíme odečíst následující, je jich tedy 68.

|  |  |  |
| --- | --- | --- |
| za|co | 1 279 | 10,5 |
| o|co | 976 | 8,01 |
| na|co | 691 | 5,67 |
| o|on | 5 | 0,04 |
| v|co | 5 | 0,04 |
| na|copak | 3 | 0,03 |
| za|on | 2 | 0,02 |
| do|on | 1 | 0,01 |

Kolikátá je v pořadí nejvíce frekventovaná nevlastní předložka? – 15. *podle*

Kolik je mezi dvaceti nejvíce frekventovanými předložek nevlastních? – 2

|  | **Filtr** | **lemma** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- | --- |
| 15 | p / n | podle | 118 913 | 976,08 |
| 16 | p / n | mezi | 114 590 | 940,6 |

Kolik je mezi padesáti nejvíce frekventovanými předložek nevlastních? – 32

|  |  |  |
| --- | --- | --- |
| 1. podle
 | 118 913 | 976,08 |
| 1. mezi
 | 114 590 | 940,6 |
| 1. proti
 | 53 821 | 441,78 |
| 1. kolem
 | 46 974 | 385,58 |
| 1. kvůli
 | 42 560 | 349,35 |
| 1. během
 | 41 969 | 344,5 |
| 1. díky
 | 40 362 | 331,31 |
| 1. kromě
 | 30 630 | 251,42 |
| 1. vedle
 | 25 415 | 208,62 |
| 1. mimo
 | 21 226 | 174,23 |
| 1. včetně
 | 21 084 | 173,07 |
| 1. místo
 | 16 533 | 135,71 |
| 1. pomocí
 | 16 456 | 135,08 |
| 1. vzhledem
 | 12 745 | 104,62 |
| 1. vůči
 | 12 564 | 103,13 |
| 1. prostřednictvím
 | 10 251 | 84,14 |
| 1. dle
 | 9 116 | 74,83 |
| 1. uprostřed
 | 8 826 | 72,45 |
| 1. oproti
 | 7 645 | 62,75 |
| 1. okolo
 | 7 118 | 58,43 |
| 1. navzdory
 | 6 534 | 53,63 |
| 1. skrz
 | 5 938 | 48,74 |
| 1. uvnitř
 | 5 546 | 45,52 |
| 1. podél
 | 5 534 | 45,43 |
| 1. koncem
 | 4 441 | 36,45 |
| 1. naproti
 | 4 263 | 34,99 |
| 1. ohledně
 | 3 976 | 32,64 |
| 1. nedaleko
 | 3 502 | 28,75 |
| 1. poblíž
 | 3 201 | 26,28 |
| 1. namísto
 | 2 884 | 23,67 |
| 1. napříč
 | 2 466 | 20,24 |
| 1. začátkem
 | 2 364 | 19,41 |

Atribut *tag* se většinou kombinuje ve složitějším dotazu s atributem lemma, zejména ve spojení se slovními druhy ohebnými.

**Příklady úkolů s využitím atributu lemma, word v kombinaci s atributem tag:**

**Srovnejte u sloves typu *péct/péci* výskyt infinitivů končících na *-i* a na *-t***.

**Řešení**:

Musíme propojit dva dotazy, **tag a word**.

1. **Vložíme tag:** Po zvolení slovního druhu:



Pokračujeme bodem 2. **Detailní určení slovního druhu:**



A klikneme na pokyn Vložit. Teď máme vložený tag:



Potřebujeme ho propojit s infinitivem sloves typu *péct/péci*, tj. zvolíme atribut word, který k vloženému tagu připojíme:



Systém hlásí chybu, protože mezi tagem a wordem chybí propojovací znak &, který vytvoříme současným kliknutím na ALT a 38:



Obdobně vytvoříme dotaz:



Pomocí těchto dokladů vyhledáme doklady a vytvoříme **frekvenční seznamy slovních tvarů**. Dostaneme dvě tabulky, v jedné z nich si slovesa nějak zabarvíme.

Frekvenční seznamy jsou uspořádány podle frekvence, my je potřebujeme uspořádané abecedně. Tabulky spojíme a uspořádáme abecedně:

U žlutě podbarvených sloves převažuje varianta s *-i*.

U modře podbarvených je výskyt vyrovnaný.

U zeleně podbarvených je doložena jen jedna varianta – s *-i*.

U šedě podbarvených je doložena jen jedna varianta – s *-t*.

U ostatních jsou doloženy obě podoby a převažuje varianta s *-t* (napsaná červeně).

| **word** | **Freq** | **i.p.m.** |
| --- | --- | --- |
| domoci | 68 | 0,56 |
| domoct | 9 | 0,07 |
| dopéci | 1 | 0,01 |
| dopéct | 16 | 0,13 |
| dopomoci | 62 | 0,51 |
| dopomoct | 6 | 0,05 |
| doříci | 5 | 0,04 |
| doříct | 53 | 0,44 |
| dosíci | 15 | 0,12 |
| dotéci | 1 | 0,01 |
| dotéct | 5 | 0,04 |
| dovléci | 6 | 0,05 |
| dovléct | 24 | 0,2 |
| ~~ject~~ | ~~41~~ | ~~0,34~~ |
| moci | 5 590 | 45,89 |
| moct | 2 243 | 18,41 |
| napéci | 1 | 0,01 |
| napéct | 18 | 0,15 |
| napíct | 1 | 0,01 |
| napomoci | 265 | 2,18 |
| napomoct | 8 | 0,07 |
| natéct | 6 | 0,05 |
| natlouct | 22 | 0,18 |
| navléci | 19 | 0,16 |
| navléct | 48 | 0,39 |
| navlíct | 3 | 0,03 |
| nedoříct | 2 | 0,02 |
| nemoci | 15 | 0,12 |
| nemoct | 14 | 0,12 |
| nenapomoci | 1 | 0,01 |
| neodříct | 1 | 0,01 |
| nepomoci | 5 | 0,04 |
| nepomoct | 4 | 0,03 |
| nepřepéct | 1 | 0,01 |
| neříci | 16 | 0,13 |
| neříct | 91 | 0,75 |
| neupéct | 1 | 0,01 |
| neutéci | 1 | 0,01 |
| neutéct | 9 | 0,07 |
| nezapéct | 1 | 0,01 |
| ~~object~~ | ~~69~~ | ~~0,57~~ |
| obléci | 126 | 1,03 |
| obléct | 312 | 2,56 |
| oblíct | 35 | 0,29 |
| obtéci | 1 | 0,01 |
| obtéct | 1 | 0,01 |
| odpomoci | 6 | 0,05 |
| odpomoct | 6 | 0,05 |
| odříci | 14 | 0,12 |
| odříct | 40 | 0,33 |
| odtéci | 3 | 0,03 |
| odtéct | 23 | 0,19 |
| odtýct | 1 | 0,01 |
| odvléci | 17 | 0,14 |
| odvléct | 41 | 0,34 |
| odvlíct | 2 | 0,02 |
| opéci | 6 | 0,05 |
| opéct | 51 | 0,42 |
| otéct | 2 | 0,02 |
| otlouci | 1 | 0,01 |
| otlouct | 7 | 0,06 |
| péci | 36 | 0,3 |
| péct | 421 | 3,46 |
| píct | 5 | 0,04 |
| podvléci | 1 | 0,01 |
| podvléct | 1 | 0,01 |
| pomoci | 7 055 | 57,91 |
| pomoct | 4 237 | 34,78 |
| poséci | 1 | 0,01 |
| poséct | 1 | 0,01 |
| posíct | 1 | 0,01 |
| potlouct | 2 | 0,02 |
| povléct | 2 | 0,02 |
| ~~project~~ | ~~346~~ | ~~2,84~~ |
| propéci | 1 | 0,01 |
| propéct | 8 | 0,07 |
| proříct | 2 | 0,02 |
| protéci | 6 | 0,05 |
| protéct | 10 | 0,08 |
| protlouci | 3 | 0,03 |
| protlouct | 20 | 0,16 |
| provléci | 1 | 0,01 |
| provléct | 8 | 0,07 |
| předpéci | 1 | 0,01 |
| předpéct | 9 | 0,07 |
| přemoci | 144 | 1,18 |
| přemoct | 65 | 0,53 |
| přepéct | 1 | 0,01 |
| přetéci | 1 | 0,01 |
| přetéct | 7 | 0,06 |
| přetlouct | 1 | 0,01 |
| převléci | 32 | 0,26 |
| převléct | 116 | 0,95 |
| převlíct | 29 | 0,24 |
| ~~přiject~~ | ~~2~~ | ~~0,02~~ |
| přiobléci | 1 | 0,01 |
| přiobléct | 2 | 0,02 |
| připéct | 1 | 0,01 |
| přitéci | 4 | 0,03 |
| přitéct | 18 | 0,15 |
| přitlouci | 1 | 0,01 |
| přitlouct | 12 | 0,1 |
| přivléci | 1 | 0,01 |
| přivléct | 9 | 0,07 |
| rozpéci | 1 | 0,01 |
| rozpéct | 5 | 0,04 |
| roztéci | 2 | 0,02 |
| roztéct | 2 | 0,02 |
| roztlouct | 5 | 0,04 |
| rozutéct | 1 | 0,01 |
| říci | 6 932 | 56,9 |
| říct | 23 219 | 190,59 |
| síci | 1 | 0,01 |
| síct | 3 | 0,03 |
| stéci | 2 | 0,02 |
| stéct | 12 | 0,1 |
| stlouci | 1 | 0,01 |
| stlouct | 7 | 0,06 |
| stříci | 1 | 0,01 |
| stýct | 1 | 0,01 |
| svléci | 10 | 0,08 |
| svléct | 38 | 0,31 |
| svlíct | 5 | 0,04 |
| svrci | 1 | 0,01 |
| téci | 39 | 0,32 |
| téct | 236 | 1,94 |
| tlouci | 39 | 0,32 |
| tlouct | 126 | 1,03 |
| týct | 5 | 0,04 |
| upéci | 13 | 0,11 |
| upéct | 232 | 1,9 |
| upíct | 4 | 0,03 |
| utéci | 184 | 1,51 |
| utéct | 1 813 | 14,88 |
| utlouci | 3 | 0,03 |
| utlouct | 50 | 0,41 |
| utýct | 53 | 0,44 |
| vléci | 21 | 0,17 |
| vléct | 24 | 0,2 |
| ~~voblíct~~ | ~~3~~ | ~~0,03~~ |
| ~~voject~~ | ~~1~~ | ~~0,01~~ |
| vtéct | 1 | 0,01 |
| vtlouct | 13 | 0,11 |
| vymoci | 67 | 0,55 |
| vymoct | 5 | 0,04 |
| vypéct | 7 | 0,06 |
| vypomoci | 56 | 0,46 |
| vypomoct | 26 | 0,21 |
| vysléci | 1 | 0,01 |
| vysvléci | 8 | 0,07 |
| vysvléct | 22 | 0,18 |
| vysvlíct | 1 | 0,01 |
| vytéci | 2 | 0,02 |
| vytéct | 17 | 0,14 |
| vytlouci | 4 | 0,03 |
| vytlouct | 22 | 0,18 |
| vyvléci | 15 | 0,12 |
| vyvléct | 29 | 0,24 |
| vyvlíct | 4 | 0,03 |
| zapéci | 1 | 0,01 |
| zapéct | 59 | 0,48 |
| zapíct | 1 | 0,01 |
| zatéci | 1 | 0,01 |
| zatéct | 1 | 0,01 |
| zatlouci | 5 | 0,04 |
| zatlouct | 36 | 0,3 |
| zavléci | 3 | 0,03 |
| zavléct | 9 | 0,07 |
| zmoci | 3 | 0,03 |
| zmoct | 2 | 0,02 |
| zříci | 93 | 0,76 |
| zříct | 44 | 0,36 |
| ztéci | 5 | 0,04 |
| ztéct | 4 | 0,03 |
| ztlouci | 1 | 0,01 |
| ztlouct | 10 | 0,08 |

Další úkoly

**Zjistěte, jaké tvary má v gen. sg. a v instr. pl. substantivum *loď*.**

Řešení:

**Gen. sg**.

slovní druh číslo pád



instr. pl.



Spojení s lemmatem:



Výsledek:





Výsledek



Vyhledání více lemmat současně

**Další typ pokročilého dotazu – vyhledávání dvou lemmat současně**

**Zjistěte, zda má vyšší výskyt slovo *vteřina*, nebo *sekunda***.

Postup:

1. Zaklikneme „Pokročilý dotaz“
2. Zadáme obě lemmata:



|  | **Filtr** | **lemma** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- | --- |
| 1 | p / n | vteřina | 7 710 | 63,29 |
| 2 | p / n | sekunda | 3 287 | 26,98 |

Typy dotazu s & a | se dají kombinovat:

Které tvary mají ve 3. pádě množného čísla slova nemoc, pravomoc, velmoc?

1. Aktivujeme „Pokročilý dotaz“¨
2. Vložíme tag: [tag="N..P3.\*"]
3. Připojíme tři lemmata: lemma="nemoc|pravomoc|velmoc"

Celý dotaz:

[tag="N..P3.\*"&lemma="nemoc|pravomoc|velmoc"]



Závěr: u všech těchto substantiv se vyskytují v dat. pl. dubletní tvary podle typů *kost* a *píseň*.

U slova *pravomoc* je výskyt variant vyrovnaný.

U slova *velmoc* převažuje 3 : 1 varianta podle typu *píseň*.

U slova *nemoc* převažuje v poměru cca 21 : 1 varianta podle typu *kost*.

**Vyhledejte tvary 1. os. sg. a 3. os. pl. sloves *tázat* (se), *pít*, *kupovat*.**

Postup:

1. Pokročilý dotaz

[tag="V..S...1P.\*"&lemma="tázat|pít|kupovat "]

[tag="V..P...3P.\*"&lemma="tázat|pít|kupovat "]

2. Frekvence – slovní tvary

Začátek formuláře

Začátek formuláře

|  | **Filtr** | **word** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- | --- |
| 1 | p / n | piju | 234 | 1,92 |
| 2 | p / n | nepiju | 134 | 1,1 |
| 3 | p / n | kupuji | 129 | 1,06 |
| 4 | p / n | kupuju | 121 | 0,99 |
| 5 | p / n | piji | 51 | 0,42 |
| 6 | p / n | nekupuju | 23 | 0,19 |
| 7 | p / n | táži | 23 | 0,19 |
| 8 | p / n | nekupuji | 20 | 0,16 |
| 9 | p / n | nepiji | 15 | 0,12 |
| 10 | p / n | tážu | 5 | 0,04 |
| 11 | p / n | netáži | 1 | 0,01 |

Začátek formuláře

Konec formuláře

|  | **Filtr** | **word** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- | --- |
| 1 | p / n | kupují | 609 | 5 |
| 2 | p / n | pijí | 409 | 3,36 |
| 3 | p / n | pijou | 54 | 0,44 |
| 4 | p / n | nepijí | 48 | 0,39 |
| 5 | p / n | nekupují | 34 | 0,28 |
| 6 | p / n | kupujou | 25 | 0,21 |
| 7 | p / n | tážou | 11 | 0,09 |
| 8 | p / n | táží | 9 | 0,07 |
| 9 | p / n | nepijou | 8 | 0,07 |
| 10 | p / n | nekupujou | 5 | 0,04 |

Pro usnadnění seřadíme tvary podle abecedy:

|  | **Filtr** | **word** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- | --- |
| 1 | p / n | kupuji | 129 | 1,06 |
| 2 | p / n | kupuju | 121 | 0,99 |
| 3 | p / n | nekupuji | 20 | 0,16 |
| 4 | p / n | nekupuju | 23 | 0,19 |
| 5 | p / n | nepiji | 15 | 0,12 |
| 6 | p / n | nepiju | 134 | 1,1 |
| 7 | p / n | netáži | 1 | 0,01 |
| 8 | p / n | piji | 51 | 0,42 |
| 9 | p / n | piju | 234 | 1,92 |
| 10 | p / n | táži | 23 | 0,19 |
| 11 | p / n | tážu | 5 | 0,04 |
|  | **word** | **Freq** | **i.p.m.** |
| 1 | p / n | kupují | 609 | 5 |
| 2 | p / n | kupujou | 25 | 0,21 |
| 3 | p / n | nekupují | 34 | 0,28 |
| 4 | p / n | nekupujou | 5 | 0,04 |
| 5 | p / n | nepijí | 48 | 0,39 |
| 6 | p / n | nepijou | 8 | 0,07 |
| 7 | p / n | pijí | 409 | 3,36 |
| 8 | p / n | pijou | 54 | 0,44 |
| 9 | p / n | táží | 9 | 0,07 |
| 10 | p / n | tážou | 11 | 0,09 |

**3. Srovnání frekvencí. Pozor, u každého tvaru sečtěte podoby kladné i záporné!**

*táži se/netáži se 24 : tážu se/netážu se 5 5 : 1*

*piji/nepiji 66 : piju/nepiju 368 6 : 1*

*kupuji/nekupuji 149 : kupuju/nekupuju 144 vyrovnaný výskyt*

*táží se/netáží se 9 : tážou se/netážou se 11 vyrovnaný výskyt*

*pijí/nepijí 457 : pijou/nepijou 62 7 : 1*

*kupují/nekupují 643 : kupujou/nekupujou 30 21 : 1*

**Zamyslete se nad následujícím úkolem:**

Vyhledejte v korpusu SYN2020 slova *mezi*, *celkem*, *začátkem*. Podle čeho můžete poznat, jakou frekvenci mají tato slova ve funkci substantiv a jakou frekvenci mají ve funkci jiných slovních druhů?

1. Vyhledáme slova pomocí základních dotazů nebo všechna současně podle tzv. pokročilého dotazu s využitím atributu **word**:



1. Vyhledáme doklady a pomocí Frekvence – lemmata odlišíme:

podstatná jména (lemmata *začátek, celek, mez*)

neohebné slovní druhy (lemmata *mezi, celkem, začátkem*)

|  | **Filtr** | **lemma** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- | --- |
| 1 | p / n | mezi | 101 561 | 833,65 |
| 2 | p / n | celkem | 10 976 | 90,1 |
| 3 | p / n | začátkem | 1 774 | 14,56 |
| 4 | p / n | začátek | 1 097 | 9,01 |
| 5 | p / n | celek | 295 | 2,42 |
| 6 | p / n | mez | 60 | 0,49 |

1. Chceme-li určit přesně slovní druh, využijeme Frekvenci – Vlastní a v ní dvouúrovňové zadání:



**Tag** nám určí slovní druh.

|  |  |  |
| --- | --- | --- |
| MEZI = předložka se 7. nebo 4. pádem  | CELKEM – adverbium  | ZAČÁTKEM – předložka s 2. pádem |

-



**celkem** – tvar substantiva *celek* v 7. pádě

**mezi** – tvar substantiva *mez* v 3. nebo 6. pádě

NEBO VYUŽIJEME MÍSTO ATRIBUTU **tag** ATRIBUT **pos [part of speech]**, TJ. SLOVNÍ DRUH.



Vyhledejte slovo *kolem* a zjistěte pomocí Frekvence – vlastní, kterým může být slovním druhem.





Dokážete nají další slova, která mohou být více slovními druhy?

**Hledání slovních spojení**

Možné kombinace:

lemma + tag

lemma + word/řada wordů

lemma + lemma/řada lemmat

**lemma + tag**

* Zjistěte, s kterými substantivy se nejčastěji pojí adjektivum *lidový*.

Budou to dvě položky, každá zvlášť v hranatých závorkách.

[][]

Je vhodné říct si jedno takové spojení a uvědomit si **pořadí položek**: adjektivum + substantivum, nebo naopak?

**[lemma="lidový"][tag="N.\*"]**

Doklady + frekvence lemmat: soustředíme se na prvních 50 položek

|  | **Filtr** | **lemma** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- | --- |
| 1 | p / n | lidový noviny | 524 | 4,3 |
| 2 | p / n | lidový píseň | 329 | 2,7 |
| 3 | p / n | lidový tradice | 239 | 1,96 |
| 4 | p / n | lidový dům | 216 | 1,77 |
| 5 | p / n | lidový strana | 193 | 1,58 |
| 6 | p / n | lidový republika | 156 | 1,28 |
| 7 | p / n | lidový kultura | 135 | 1,11 |
| 8 | p / n | lidový fronta | 111 | 0,91 |
| 9 | p / n | lidový architektura | 90 | 0,74 |
| 10 | p / n | lidový armáda | 86 | 0,71 |
| 11 | p / n | lidový hudba | 85 | 0,7 |
| 12 | p / n | lidový léčitelství | 80 | 0,66 |
| 13 | p / n | lidový soud | 74 | 0,61 |
| 14 | p / n | lidový milice | 66 | 0,54 |
| 15 | p / n | lidový škola | 63 | 0,52 |
| 16 | p / n | lidový demokracie | 61 | 0,5 |
| 17 | p / n | lidový tanec | 57 | 0,47 |
| 18 | p / n | lidový písnička | 57 | 0,47 |
| 19 | p / n | lidový zástupce | 54 | 0,44 |
| 20 | p / n | lidový zvyk | 53 | 0,44 |
| 21 | p / n | lidový umění | 51 | 0,42 |
| 22 | p / n | lidový hlasování | 51 | 0,42 |
| 23 | p / n | lidový moudrost | 48 | 0,39 |
| 24 | p / n | lidový tvořivost | 45 | 0,37 |
| 25 | p / n | lidový stavba | 45 | 0,37 |
| 26 | p / n | lidový kroj | 41 | 0,34 |
| 27 | p / n | lidový vyprávění | 38 | 0,31 |
| 28 | p / n | lidový hnutí | 38 | 0,31 |
| 29 | p / n | lidový obyčej | 32 | 0,26 |
| 30 | p / n | lidový vrstva | 31 | 0,25 |
| 31 | p / n | lidový řemeslo | 31 | 0,25 |
| 32 | p / n | lidový prostředí | 29 | 0,24 |
| 33 | p / n | lidový název | 29 | 0,24 |
| 34 | p / n | lidový představa | 29 | 0,24 |
| 35 | p / n | lidový novina | 29 | 0,24 |
| 36 | p / n | lidový podání | 27 | 0,22 |
| 37 | p / n | lidový víra | 26 | 0,21 |
| 38 | p / n | lidový slovesnost | 26 | 0,21 |
| 39 | p / n | lidový slavnost | 26 | 0,21 |
| 40 | p / n | lidový jednota | 26 | 0,21 |
| 41 | p / n | lidový nástroj | 25 | 0,21 |
| 42 | p / n | lidový pověst | 24 | 0,2 |
| 43 | p / n | lidový léčitel | 24 | 0,2 |
| 44 | p / n | lidový zábava | 23 | 0,19 |
| 45 | p / n | lidový divadlo | 22 | 0,18 |
| 46 | p / n | lidový muzika | 22 | 0,18 |
| 47 | p / n | lidový list | 22 | 0,18 |
| 48 | p / n | lidový podpora | 21 | 0,17 |
| 49 | p / n | lidový pohádka | 20 | 0,16 |
| 50 | p / n | lidový veselice | 20 | 0,16 |

Co z nich lze vyčíst? Co pojmenovávají?

**Najděte samostatně:**

Názvy děl – vlastní jména (1)

Názvy institucí, státních forem apod. z historie (3)

Názvy děl, akcí apod. související s tradicí (alespoň 6)

názvy osob (2)

* Zjistěte, s kterými adjektivy se nejčastěji pojí substantivum *reforma*.

Je vhodné říct si jedno takové spojení a uvědomit si **pořadí položek**: adjektivum + substantivum, nebo naopak?

[tag="A.\*"][lemma="reforma"]

|  | **Filtr** | **lemma** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- | --- |
| 1 | p / n | ekonomický reforma | 109 | 0,9 |
| 2 | p / n | pozemkový reforma | 84 | 0,69 |
| 3 | p / n | důchodový reforma | 79 | 0,65 |
| 4 | p / n | měnový reforma | 73 | 0,6 |
| 5 | p / n | sociální reforma | 52 | 0,43 |
| 6 | p / n | politický reforma | 47 | 0,39 |
| 7 | p / n | zásadní reforma | 45 | 0,37 |
| 8 | p / n | daňový reforma | 44 | 0,36 |
| 9 | p / n | penzijní reforma | 40 | 0,33 |
| 10 | p / n | hospodářský reforma | 28 | 0,23 |
| 11 | p / n | strukturální reforma | 26 | 0,21 |
| 12 | p / n | správní reforma | 25 | 0,21 |
| 13 | p / n | další reforma | 25 | 0,21 |
| 14 | p / n | josefínský reforma | 23 | 0,19 |
| 15 | p / n | školský reforma | 22 | 0,18 |
| 16 | p / n | velký reforma | 22 | 0,18 |
| 17 | p / n | radikální reforma | 21 | 0,17 |
| 18 | p / n | strategický reforma | 20 | 0,16 |
| 19 | p / n | ústavní reforma | 19 | 0,16 |
| 20 | p / n | společenský reforma | 19 | 0,16 |
| 21 | p / n | policejní reforma | 18 | 0,15 |
| 22 | p / n | zdravotnický reforma | 17 | 0,14 |
| 23 | p / n | hluboký reforma | 17 | 0,14 |
| 24 | p / n | tržní reforma | 15 | 0,12 |
| 25 | p / n | plánovaný reforma | 14 | 0,12 |
| 26 | p / n | soudní reforma | 13 | 0,11 |
| 27 | p / n | potřebný reforma | 12 | 0,1 |
| 28 | p / n | kurikulární reforma | 12 | 0,1 |
| 29 | p / n | chystaný reforma | 12 | 0,1 |
| 30 | p / n | skutečný reforma | 12 | 0,1 |
| 31 | p / n | demokratický reforma | 12 | 0,1 |
| 32 | p / n | celkový reforma | 12 | 0,1 |
| 33 | p / n | institucionální reforma | 12 | 0,1 |
| 34 | p / n | rozsáhlý reforma | 11 | 0,09 |
| 35 | p / n | nový reforma | 11 | 0,09 |
| 36 | p / n | navrhovaný reforma | 9 | 0,07 |
| 37 | p / n | katolický reforma | 9 | 0,07 |
| 38 | p / n | agrární reforma | 9 | 0,07 |
| 39 | p / n | poslední reforma | 9 | 0,07 |
| 40 | p / n | nutný reforma | 9 | 0,07 |
| 41 | p / n | Gorbačovův reforma | 8 | 0,07 |
| 42 | p / n | vládní reforma | 8 | 0,07 |
| 43 | p / n | náboženský reforma | 8 | 0,07 |
| 44 | p / n | volební reforma | 8 | 0,07 |
| 45 | p / n | připravovaný reforma | 8 | 0,07 |
| 46 | p / n | zdravotní reforma | 8 | 0,07 |
| 47 | p / n | justiční reforma | 7 | 0,06 |
| 48 | p / n | vnitřní reforma | 7 | 0,06 |
| 49 | p / n | úspěšný reforma | 7 | 0,06 |
| 50 | p / n | raný reforma | 7 | 0,06 |

lemma + lemma/řada lemmat

**Vyhledejte v korpusu SYN2020 spojení:**

město Plzeň

Po vyhledání zjistěte frekvenci slovních tvarů.

Když vyřadíte z frekvenčního seznamu položky, v nichž má město základní tvar nebo tvar množného čísla, zjistěte, zda jsou častější případy:

* ve kterých se tvar slova *město* a tvar vlastního jména shodují v pádě;
* ve kterých se slovo *město* skloňuje a tvar vlastního jména se nemění.

| **Filtr** | **word** | **Freq** | **i.p.m.** |
| --- | --- | --- | --- |
| 1 | p / n | města plzně | 78 | 0,64 |
| ~~2~~ | ~~p / n~~ | ~~město plzeň~~ | ~~29~~ | ~~0,24~~ |
| 3 | p / n | města plzeň | 5 | 0,04 |
| 4 | p / n | městem plzní | 4 | 0,03 |
| 5 | p / n | městu plzeň | 3 | 0,03 |
| 6 | p / n | městu plzni | 2 | 0,02 |
| 7 | p / n | městě plzeň | 2 | 0,02 |
| 8 | p / n | městem plzeň | 1 | 0,01 |
| 9 | p / n | městem plzně | 1 | 0,01 |

* shodné tvary: 84
* tvary s neskloňovaným vlastním jménem: 11