

JAK VYUŽÍVAT ČESKÝ NÁRODNÍ KORPUS

[tag="N..P1.*"]

tag="Ve.*"]

word=

{3}

STUDIJNÍ PŘÍRUČKA

Popis jednotlivých pozic morfologické značky (tag)

POZICE 1 – SLOVNÍ DRUH

- A – adjektivum
- C – číslovka, nebo číselný výraz s číslicemi
- D – adverbium
- I – interjektce
- J – konjunktce
- N – substantivum
- P – pronomen
- R – prepozice
- T – partikule
- V – verbum
- X – neznamý, neurčený; neurčitelný slo vni druh
- Z – interpunkce, hranice věty

POZICE 2 – DETAILNÍ URČENÍ SLOVNÍHO DRUHU

- # – hranice věty
- * – slovo „krát“, – spojka podřadící (vč. „aby“ a „kdyby“)
- – zkratka jako adjektivum
- 0 – předložka s připojeným „-ř“ („proř“)
- 1 – vztažné přivlastňovací zájmeno („jehož“)
- 2 – slovo před pomlčkou
- 3 – zkratka jako číslovka
- 4 – vztažné nebo tážací zájmeno s adjektivním skloňováním
- 5 – zájmeno „on“ ve tvarech po předložce („něj“)
- 6 – reflexivní zájmeno „se“ v dlouhých tvarech („sebe“)
- 7 – reflexivní zájmeno „se“, „si“, „ses“, „sis“
- 8 – přivlastňovací zájmeno „svůj“
- 9 – vztažné zájmeno „jenž“, „jž“, ... po předložce („n-“, „něž“)
- : – interpunkce všeobecně
- ;- zkratka jako substantivum
- = – číslo psané číslicemi

@ – slovní tvar, který nebyl morfologickou ana-

lyzou rozpoznán

- A – adjektivum obvyčejně
- B – sloveso, tvar přítomného nebo budoucího času
- C – adjektivum, jmenný tvar
- D – zájmeno ukazovací („ten“)
- E – vztažné zájmeno „což“
- F – součást předložky, která nikdy nestojí samostatně („vzhledem“)
- G – přídavné jméno odvozené od slovesného tvaru přítomného přechodníku
- H – krátké tvary osobních zájmen („mě“)
- I – číslovce
- J – vztažné zájmeno „jenž“ bez předložky
- K – zájmeno tážací nebo vztažné „kdo“, vč. tvarů s „-ž“ a „-s“
- L – zájmeno neurčité „všechen“, „sám“
- M – přídavné jméno odvozené od slovesného tvaru minulého přechodníku
- N – substantivum obvyčejně
- O – samostatně stojící zájmena „svůj“, „nesvůj“, „tenam“
- P – osobní zájmena (vč. tvaru „lys“)
- Q – zájmeno tážací/vztažné „co“, „copak“, „cožpak“
- R – předložka obvyčejná
- S – zájmeno přivlastňovací „můj“, „tvůj“, „jeho“ (vč. plurálu)
- T – částice
- U – adjektivum přivlastňovací (na „-ův“ i „-in“)
- V – předložka vokalizovaná („ve“)
- W – zájmena záporná („nic“)
- X – slovní tvar, který byl rozpoznán, ale značka chybí
- Y – zájmeno „co“ spojené s předložkou („oč“)

Z – zájmeno neurčité („nějaký“)

/ – spojka souřadící

- a – číslovka neurčitá („mnoho“)
- b – příslovce (bez určení stupně a negace; „pozdů“)
- c – kondicionali slovesa být („bych, ...“)
- d – číslovka druhová, adjektivní skloňování („jedny“)
- e – slovesný tvar přechodníku přítomného
- f – slovesný tvar: infinitiv
- g – příslovce (s určením stupně a negace)
- h – číslovky druhové „jedny“ a „nejedny“
- i – slovesný tvar rozkazovacího způsobu
- j – číslovka druhová >= 4, substantivní postavení („čtyvero“)
- k – číslovka druhová >= 4, adjektivní postavení, krátký tvar („čtyry“)
- l – číslovky základní 1–4, „půl“, ...; sto a tisíc v nesubstantivním skloňování
- m – slovesný tvar přechodníku minulého
- n – číslovky základní >= 5
- o – číslovky násobné neurčité („krát“; „mnohokrát“)
- p – slovesné tvary minulého aktivního přičestí (včetně přídavného „-s“)
- q – archaické slovesné tvary minulého aktivního přičestí (zakončení „-t“)
- r – číslovky řadové
- s – slovesné tvary pasivního přičestí (vč. přídavného „-s“)
- t – archaické slovesné tvary přítomného a budoucího času (zakončení „-t“)
- u – číslovka tážací násobná „kolikrát“
- v – číslovky násobné („pětkrát“, „poprvé“ ...)
- w – číslovky neurčité s adjektivním skloňováním („tolikrát“)

- x – zkratka, slovní druh neurčen/neznamý
- y – zlomky zakončené na „-ina“
- z – číslovka tážací řadová „kolikátý“
- ž – číslovka psaná římskými číslicemi
- ~ – zkratka jako sloveso

POZICE 3 – JMENNÝ ROD

- neurčuje se
- F – femininum (ženský rod)
- H – femininum nebo neutrum (tedy nikoli maskulinum)
- I – maskulinum inanimatum (rod mužský neživotný)
- M – maskulinum animatum (rod mužský životný)
- N – neutrum (střední rod)
- Q – femininum singuláru nebo neutrum plurálu (pouze u přičesti a jmenných adjektiv)
- T – masculinum inanimatum nebo femininum (jen plurál u přičesti a jmenných adjektiv)
- X – libovolný rod (F/M/I/N)
- Y – masculinum (animatum nebo inanimatum)
- Z – 'nikoli femininum' (tj. M/I/N; především u příslovců)

POZICE 4 – ČÍSLO

- neurčuje se
- D – duál (pouze 7. pád feminin)
- P – plurál (možné číslo)
- S – singulár (jednotné číslo)
- W – pouze v kombinaci s jmenným rodem 'Q' (singulár pro feminina, plurál pro neutra)
- X – libovolné číslo (P/S/D)

POZICE 5 – PÁD

- neurčuje se
- 1 – nominativ

- 2 – genitiv
- 3 – dativ
- 4 – akuzativ
- 5 – vokativ
- 6 – lokál
- 7 – instrumentál
- X – libovolný pád (1/2/3/4/5/6/7)

POZICE 6 – PŘIVLASTŇOVACÍ ROD

- neurčuje se
- F – femininum
- M – masculinum animatum (rod mužský životný)
- X – libovolný rod (F/M/I/N)
- Z – 'nikoli femininum' (tj. M/I/N; u přivlastňovacích adjektiv)

POZICE 7 – PŘIVLASTŇOVACÍ ČÍSLO

- neurčuje se
- P – plurál
- S – singulár

POZICE 8 – OSOBA

- neurčuje se
- 1 – 1. osoba
- 2 – 2. osoba
- 3 – 3. osoba
- X – libovolná osoba (1/2/3)

POZICE 9 – ČAS

- neurčuje se
- F – futurum
- H – minulost nebo přítomnost (P/R)
- P – přezens
- R – minulý čas
- X – libovolný čas (F/R/P)

POZICE 10 – STUPĚŇ

- neurčuje se
- 1 – 1. stupeň
- 2 – 2. stupeň
- 3 – 3. stupeň

POZICE 11 – NEGACE

- neurčuje se
- A – afirmativ (bez negativní předpony „ne-“)
- N – negace (tvar s negativní předponou „ne-“)

POZICE 12 – AKTIVUM/PASÍVUM

- neurčuje se
- A – aktivum nebo 'nikoli pasívum'
- P – pasívum

POZICE 13 A 14 – NEPOUŽITO

- neurčuje se

POZICE 15 – VARIANTA, STYLOVÝ PŘÍZNAK

APOD.

- neurčuje se
- 1 – varianta, víceméně rovnocenná
- 2 – řídká, archaická nebo knižní varianta
- 3 – velmi archaický tvar, též hovorový
- 4 – velmi archaický nebo knižní tvar, pouze spisovný
- 5 – hovorový tvar (ve veřejných projevech)
- 6 – hovorový tvar (koncovka standardní obecné češtiny)
- 7 – hovorový tvar, varianta k '6'
- 8 – zkratky
- 9 – speciální použití (tvary zájmen po předložkách...)

Signatura	34948
inventurní číslo	609/2010
ROMÁNSKÁ KNIHOVNA	
Filozofická fakulta UK	
Nám. J. Palachna - 116 38 Praha 1	

R 8609/2010

Tato publikace vznikla s podporou výzkumného záměru MSM 0021620823.



**Filozofická fakulta
Univerzity Karlovy v Praze**

Hlavní redakce: prof. PhDr. František Čermák, DrSc. – PhDr. Renata Blatná, CSc.

© František Čermák, Renata Blatná, Jana Klimová, Marie Koprňová,
Karel Kučera, Vladimír Pekvevič, Věra Schmiedlová, Michal Šulc, 2005
Cover © Jindřiška Mrázková, 2005
Všechna práva vyhrazena

ISBN 80-7106-736-9

**Pro toho, kdo chce hledat a (v korpusu) najde.
Předmluva k této příručce.**

František Čermák

V dějinách poznání a dokumentace jazyka představuje Český národní korpus (ČNK) mimořádnou a novou událost. Skrze tento korpus představující nebylaly a neúšně bohatý a věrný záznam jazyka se ovšem odráží nejen jazyk sám, ale i všechno za ním, celá naše skutečnost. Je to vlastně počítačem uloženy, zpracovávaný a prohlédávatelný zdroj informací, jaký tu dosud nebyl, který je na jedné straně nabízen na Internetu všem vážným zájemcům i zvědavým lidem na celonárodní úrovni, který však na druhé straně zároveň stále roste, do budoucnosti i minulosti. Jeho základní podoba, kterých je více a která bude asi zájmem uživatelů nejčastěji, je stornilónový reprezentativní a vyvážený korpus českého jazyka současnosti SYN2000: stornilónový, protože obsahuje 100 miliónů slov, slovních tvarů, SYN, protože je synchronní, tj. současný, a 2000 proto, že do tohoto roku sahá pokrytí zahrnuté češtiny a po něm budou následovat korpusy další. Je zřejmé, že takovýto obrovský zdroj informací nelze svým zaměřením nějak omezit na jediný obor, jakkoliv jeho zatím nejčastějšími zájemci jsou studenti a výzkumní badatelé z oboru českého jazyka; stejně tak ale nelze na druhou stranu říct, kdo všechno ho užívat může. To už závisí na vynalézavosti uživatele z té které oblasti, ať už je to myslivý student, profesor, nebo třeba jen běžný nespécializovaný člověk. Na druhé straně je třeba hned teď odmítnout názor, že druh informací, které se tu nabízejí, lze najít, jak si dnes leckdo prostomyslně myslí, na Internetu. To zdaleka ne, ČNK je jedinečný především svou vyvážeností a ovšem i zpracováním.

Poznání zákonitostí a pravidel stavby a užívání jazyka vyplývá jen a pouze ze studia jeho skutečného užívání (úzu), přičemž se tím myslí především užívání typické a časté. Všechny podklady pro jazykové příručky a slovníky vznikaly tímto způsobem, to jest pozorováním jazyka a sledováním jeho zákonitostí na mnoha opakovaných případech vždy stejného užití, a to v kontextu a souvislostech jeho výskytu. Takový, avšak nesrovnatelně větší, a proto i spolehlivější podklad k pozorování a studiu teď nabízí i ČNK, jehož základním způsobem nabídky informace, a tedy i studia je konkordance, tj. soubor řádků se všemi užitími, výskytů hledaného slova apod. Každý řádek z takové konkordance, který se dá chápat třeba jako výpis, excerpce z nějakého díla a který lze podle potřeby libovolně zvěšovat, je potom dokladem na jedno užití hledaného slova ve skutečném kontextu, který už někdo použil. Prozkoumáním a srovnáním všech konkordančních řádků se pak dá snadno zjistit, co je na užití slova apod. základní, co méně časté a co okrajové.

vé. To nám snadno ukáže a rozliší různá frekvence (častost) toho či onoho významu, kombinace apod. s daným slovem. V tomto smyslu je většinou, ne však jediná informační pravda o takovém slově v jeho nejčastějších konkordancních řádcích. Vedle ní se tu však setkáme i s jevy a způsoby méně častými i řídkými, někdy i s jazykovým experimentováním. Obecně pak takováto pestrost a alternativnost odráží, na rozdíl od většinou zjednodušujících příruček, skutečnou pluralitu (variabilitu) jazyka, která je nutným projevem každého vývoje.

Z mnoha věcí, na které myslící člověk, který se do korpusu ponoří, natrazí, zarazí a asi překvapí i to, že jazyk naší současnosti (v SYN2000 a v mluvných korpusech), který se tu zachycuje v celé své mnohotvárnosti a bohatství, není ovšem vždy ten, jaký nám až dosud podávaly naše příručky. Jejich obraz jazyka se na pozadí pestrosti, alternativnosti informace v korpusu nutně jeví čemoblie zjednodušený a často pohodlně jasnější, než jaký jazyk ve skutečnosti je. Uživatel pak má a zamýšlet se, proč tomu tak je a odkud se bere ten rozpor, popřípadě, kde je pravda. Zjednodušující pravda příruček, která je pro spisovný jazyk a školy většího hosti informace a mnohotvárnosti pravdy korpusu. Protože všechny příručky a učebnice čerpaly i v minulosti v zásadě ze skutečné podoby, v jaké se jazyk užívá, dnešní informace z korpusu svou mnohonásobnou spolehlivostí jasně nad čemobilými pravdami příruček převažuje, může je nejen v lecčem opravovat, ale i usvědčovat z omylu a nedostatků. To si uživatel snadno uvědomí, pokud si najde čas a projde si cvičení s korpusem, která se v této knižce nabízejí, a začne obojí kriticky srovnávat. V nejlépeším případě pak může poznání z takovýchto cvičení, brousicích především kritické myšlení, vyústit i v apel na budoucí autory učebnic a podobně.

V této knižce získá uživatel nejen základní a někdy i prohloubené znalosti nutné pro užívání korpusu a umožňující mu v něm nalézt, co potřebuje, ale díky svému důmyslu i návody na hlubší analýzy. Uživatel však musí mít vždy na paměti, že vše v jazyce je zprostředkováno skrze formu, tj. písmena (hlásky), slova a jejich kombinace a že jen tak se dá dobrat i poznání a znalosti o významu. Snažit se tedy klást korpusu otázku po významu bez formy není možné. To samozřejmě vede k potřebě se zamyslet na tím, jak svou otázku zformulovat. Zepat se dá ne či hloupě otázky samozřejmě neodpoví a je na uživateli zeptat se sebe sama, v čem nebyla jeho otázka v pořádku. Tato knižka, která se dělí jednak podle větších tematických celků a jednak podle obtížnosti (u které se spíše končí než začíná), však není jen technickou příručkou dovedností. Chce být i nástrojem tříbit doplněním starých i nových středoškolských znalostí a příruček především pro ty, kteří se nebojí myslet, jakkoliv to nemusí být vždy snadné.

ČNK je vybudován na obrovském a vyváženém množství textů (obecný popis a jejich seznam viz ČNK. Úvod a příručka uživatele 2000, který vyšel knižně, lze ho najít ale i na webu, viz níže), které se musely zpracovávat skrze řadu náročných kroků, než dorazily k uživateli. ČNK je akademický, nevydělečný projekt, financovaný především ze státních zdrojů; na jeho vzniku se ale také nezájímá podílí velká řada českých nakladatelů a vydavatelů svým svolením začlenit do něj svou produkci knižní, časopiseckou či novinovou a podobně. Aby v takové záplavě informací uživatel mohl hledat, k tomu potřebuje zvládnout určitou průpravu a formální jazyk, který se mu zde právě snažíme přiblížit. Zjistit tak nejen odpověď na svůj dotaz, ale i mnoho informací dalších, mj. i o tom, z jakého zdroje, knihy apod. pocházejí.

Získat volný přístup (zdarma) k ČNK pro nekomerční účely nevyžaduje mnoho. Je třeba se spojit s ČNK, jehož sídlem je Ústav Českého národního korpusu (Filozofická fakulta Karlovy univerzity, nám. J. Palacha 2, Praha 1, 110 00, e-mail ucnk@ff.cuni.cz a web <http://ucnk.ff.cuni.cz>). Na uvedené adrese si najdete Prohlášení uživatele, které si vytiisknete a vyplněné a podepsané pošlete se svou stejné webovské stránky ČNK stáhnete na svůj počítač (ten musí být samozřejmě připojen na síť) a malou poštu na adresu Ústavu ČNK (viz výše). Podle návodu si ze stejné webové stránky ČNK stáhnete na svůj počítač (ten musí být samozřejmě Po několika dnech dostanete své osobní heslo (password) a pak už můžete na dálku a po Internetu využívat všechny korpusy, které se Vám objeví po zalogování v nabídce. Ke korpusu se může takto, tj. s vlastním heslem připojit stejně tak i kdokoli jiný. Samozřejmě je toto svolení (licence) k užít ČNK, které jste takto získali, osobní a nepřenosné. Po roce užívání dostanete po e-mailu od správce otázku, jestli si přejete užívat ČNK i dále; odpovíte-li kladně, Vaše licence bude automaticky prodloužena. Uvítáme pak především Vaše zkušenosti, náměty a ovšem hlavně i případnou nabídku dalších elektronických textů k rozšíření ČNK, kterému tak i Vy přispějete k dalšímu růstu a širšímu užívání.

Jednotlivé úkoly, které obsahuje tato příručka, jsou založeny na práci se 100miliónovým korpusem SYN2000, v němž se jazykové jevy vyhledávají pomocí programu Bonito, jehož autorem je Pavel Rychlý z FI MU v Brně. Mannuál k tomuto programu je na internetové adrese <http://ucnk.fk.cuni.cz/bonito/>. Morfologickou analýzu a desambiguaci korpusu SYN2000 zpracoval Jan Hajič z MFF UK v Praze (viz příložený Popis jednotlivých pozic morfologické značky).

Po instalaci je jako základní atribut pro vyhledávání nastaven atribut word. To znamená, že slovní tvary je možné zadávat bez závorek a uvozovek, např. **korpus**. Z tohoto základního nastavení vychází klíč k jednotlivým úkolům a variantní zápis [word="korpus"] se proto v klíči neobjevuje. Pro rychlejší orientaci je možné využít 20 základních vyhledávacích vzorů.

Základní publikaci pro seznámení s korpusovou lingvistikou a její terminologií, která se v zadáních jednotlivých úkolů objevuje, je „Český národní korpus. Příručka uživatele“ (Praha 2000).

1. Práce s Českým národním korpusem krok za krokem

A. Pravopis

ÚVOD

Vyhledávání pravopisných a hláskoslovných jevů korpusu je poměrně snadné, neboť jde většinou o vyhledávání hláskových a pravopisných variant nebo psaní velkého a malého písmene, kde vystačíme s uvedením tohoto výčtu v hranaté závorce, např: **optimišizlmus**, **[s]slunce** apod., popř: vyhledávání obecněčeských, tj. nespisovných a spisovných variant téhož slova, srov. **brýle** | **brejle**, tj. využití funkce **Nebo**. Korpus SYN2000 je korpusem pouze psaných textů, nemůžeme v něm tedy hledat výslovnost jako takovou, ale pouze její odraz v pravopisu, srov. např: zakončení *-uáhní x -uehni*. Výhodou korpusu je to, že snadno zjistíme nejen frekvenci jednoho tvaru, či jednoho slova, ale můžeme si porovnat frekvence jednotlivých hláskových a pravopisných variant, udělat si seznam slov obsahujících určitou hláskovou skupinu a zjistit, které z nich je nejčastější. K tomu je velmi výhodné využít funkci **Frekvenční distribuce** v menu **Statistiky** – **Konkordance**. Ve stejném menu je i funkce **Rozložení**, která nám poskytne obrázek názorně ukazující, ve kterých typech textů je hledaný tvar, varianta apod. nejčastější. Naznačuje se tak někdy i přirozená závislost varianty na žánru, resp. stylu.

■ Úkol 1

Slovo *hrad* se v češtině může psát s malým nebo velkým písmenem na začátku podle toho, zda jde o substantivum obecné nebo vlastní.

a) Vyhledejte si nejprve tvar *Hrad* (vzor 1) a podívejte se, ve kterých typech textů se nejčastěji vyskytuje. (menu **Konkordance** – **Statistiky** – **Rozložení**).

b) Seřadte si výskyt abecedně podle 1. slova zleva (menu **Konkordance** – **Jednoduché třídění** – **Levý kontext (word)**) a sledujte, jak se liší kontexty se slovem *Hrad* na začátku a uvnitř věty.

c) Seřadte si kontexty podle 1. slova zprava. Vypište z korpusu spojení typu *Hrad* **krivoklát**. Kolik takových spojení je?

■ Úkol 2

Nyní vyhledejte v korpusu tvar *hrad* s malým i velkým písmenem (vzor 2). Posudte kontexty, které vám korpus nabízí. Čím se tyto kontexty liší? Dá se tvar *Hrad* nahradit něčím jiným, např: slovním spojením? Má vždy stejný význam?

■ Úkol 3

Zjistěte, kdy se slovo *Praha* píše s velkými a kdy s malými písmeny. Využijte vzor 3. Kolikrát je v korpusu tvar *PRAHA* (menu Konkordance – Statistiky – Frekvenční distribuce /word/)? Jaké jsou nejčastější kontexty? Kolikrát je v korpusu tvar *praha*? Ve kterých případech je název *Praha* s malým *p* náležitý? Ve kterém díle se tvar *praha* objevuje nejčastěji?

■ Úkol 4

Na hodinách češtiny se často procvičují hláskově shodná slova, která se mohou psát s malým i s velkým písmenem, avšak liší se významově. Zkuste bez pomoci PČP na základě korpusu zformulovat, za jakých podmínek se píše velké a za jakých malé písmeno u tvarů *jítěnka* a *jítěnka*. Jaké jsou zde problémy?

■ Úkol 5

V korpusu je možné zadat dotaz tak, aby bylo možné vyhledat najednou *eso*, *EsO*, *ESo*, *ESo*, *eSO*, *eSO*. Využijte k tomu vzor 3 a podívejte se, které z těchto šesti možných kombinací velkých a malých písmen v rámci jednoho slovního tvaru se v textech skutečně realizují, užívají (menu Konkordance – Statistiky – Frekvenční distribuce /word/). Co všechno může – na základě korpusových dat – znamenat zkratka *ESO*?

■ Úkol 6

Jazyková situace, ve které existují dvě hláskově odlišné varianty téhož slova, například *bryěle* i *brejle*, je pro Čecha situací standardní. Nikoho nepřekvapuje. Dokážeme ale o těchto dvou variantách říct něco více, než jen to, že existují?

Obě varianty si vyhledejte v korpusu (vzor 5). Zjistěte, že v korpusu SYN2000 je první varianta zhruba desetkrát častější než druhá. Je třeba prozkoumat tvrzení, že varianta *bryěle* je běžná v psané češtině a varianta *brejle* pouze v češtině mluvené, i když tendence k takto rozdílnému výskytu je zřejmá. Podívejme se postupně na slova *bryěle* a *brejle* v psaných textech a pokusme se zjistit maximum o tom, v čem se tyto výrazy liší, jaké různé tendence se u nich projevují. Můžete pro to využít funkci Kolokace (menu Konkordance – Statistiky – Kolokace), kde porovnáte, s jakými slovy a tvary se hledané slovní tvary pojí.

■ Úkol 7

Jazyková situace, ve které existují dvě hláskově odlišné varianty téhož slova, je pro Čechy běžná v případech jako *bryěle* a *brejle*. Málo běžná je ovšem situace, kdy jedno slovo najdeme v textech napsáno jednou s *i*, jindy s *y* beze změny významu jako je tomu v případě dvojice *sirup* x *syrup* (toto kolísání způsobila

ovšem změna pravopisných pravidel, nikoli úzus). Jak zjistíme, která z variant je dnes častá a která řidší starší varianta? Použijte při korpusovém vyhledávání vzor 2. V jakém typu textu se varianta s *y* vyskytuje?

■ Úkol 8

Podobně jako tomu bylo v případě dvojice *sirup* x *syrup* je málo běžná také ta situace, kdy jedno slovo vyslovujeme a píšeme jednou s *te*, jindy s *tě* beze změny významu. Týká se to slov *protežovat* a *protěžovat*. Jak zjistíme, která z variant je častější a která jen mylným etymologickým výkladem?

■ Úkol 9

Jedním z procvičovaných pravopisných jevů je také *ú* uvnitř slova. Pokusme se nyní vyhledat si taková slova v korpusu (vzor 5Aa + 5Ab, místo * napište +) a z jejich seznamu (menu Konkordance – Statistiky – Frekvenční distribuce /lemma) až do frekvence 50 si udělat skupiny, které nám pomůžou zapamatovat si, ve kterých slovech musíme dávat na tento jev zvláštní pozor. O jaké skupiny slova jde?

■ Úkol 10

Najděte 5 nefrekventovanějších substantiv (nikoli vlastní jména) a 5 citoslovcí, která končí na *-ú* (vzor 5Aa). U jednoho z podstatných jmen vás možná jeho psaná podoba zarazí. Zvažte, zda jde o standardní variantu. Co k vašemu rozhodnutí přispělo?
Z kterého jazyka pochází slovo *furjů*?

■ Úkol 11

Nová PČP dovolují psaní slov zakončených na *-ismus*/*-izmus* obojím způsobem. A. Zjistěte, zda toto rozhodnutí odpovídá jazykovému úzu. V psaní kterých slov převažuje zakončení na *-izmus*?

Zkuste si vyhledat v korpusu lemmata zakončená na *-ismus* a *-izmus* současně (vzor 17)

Zadáním P-Filtru **izm*, **iz* najdete tvary obsahující tuto pravopisnou variantu. Na základě korpusové vyhledávky určete dvě nefrekventovanější slova (funkce Konkordance – Statistiky – Frekvenční distribuce), která mají vedle *-ismus* také variantu *-izmus*.

V rámci těchto slov zjistěte poměr užívání jednotlivých pravopisných variant.

- **Úkol 12**
Vyhledejte slova obsahující skupinu hlásek *gó-* (vzor 10), která není běžná v českých slovech. O jaká slova se jedná, které z nich má nejvíce odvozenin?

- **Úkol 13**
Vyhledejte si v korpusu SYN2000 slovní tvary začínající *voh-* (vzor 1a) a vytvořte si seznam tvarů (Statistiky – Frekvenční distribuce /word/). Porovnejte tento seznam se seznamem tvarů v Pražském mluveném korpusu. Které tvary se vyskytují v obou korpusech? Jednotlivým tvarům v obou seznamech přiřadte základní tvar (např. *vohněm – voheti*) a vytvořte tak dva seznamy různých slov, např. z tvarů *voháknutá, vohozu, vohoz* vznikne seznam nespisovných slov *voháknutej, vohoz* apod.).

- **Úkol 14**
Nejprve v korpusu vyhledejte tvar *víže* (vzor 1) a potom porovnejte jeho kontexty s kolokací *váže k* (vzor 4). V jakých typech textů se tvar *víže* (kromě publicisticky) nejvíce vyskytuje? Využijte k tomu funkci Rozložení (menu Konkordance – Statistiky – Rozložení). V jakých typech textů se naopak vyskytuje spojení (kolokace) *váže k*? Dal by se v některých případech slovesný tvar *víže* nahradit slovesem *váže*?

- **Úkol 15**
Zadejte si v korpusu současně lemmata *čúchat* a *čičhat* (vzor 5 + 2). V jakých tvarech se nejčastěji sloveso *čúchat* užívá? Mají tato slovesa, jedno spisovné a jedno obecněčeské, společné kontexty? O kom se nejčastěji říká, že *čúchá*?

- **Úkol 16**
Zadejte v korpusu kolokaci *jako řěpa* (vzor 4 + 14) s 0-3 vloženými slovy a tato přirovnání porovnejte s kolokací *jako řepa* rovněž s 0-3 vloženými slovy. Která hlásková varianta se užívá ve frazeologii?

- **Úkol 17**
Vyhledejte si v korpusu lemmata zakončená na *-uální* a seznam vytvořený pomocí Frekvenční distribuce (menu Konkordance – Statistiky – Frekvenční distribuce) porovnejte s obdobným seznamem lemmat zakončených na *-uelní*. Jaké je pořadí prvních deseti lemmat u prvního seznamu? Do jaké míry je těchto 10 lemmat shodných s lemmaty druhého seznamu? Která lemmata jsou v každé ze skupin odlišná? Mají něco významové společného? Ze kterých dvou jazyků pochází výslovnost [-elni:]?

- **Úkol 18**
Myslete, že Češi užívají častěji kombinace předložky a zájmena z *kterého*, z *které*, z *kterých* nebo ze *kterého*, ze *které*, ze *kterých*? Využijte vzor 12. Co myslíte, že užíváte častěji vy? Ověřte si své představy na Českém národním korpusu.

- **Úkol 19**
Vyhledejte v korpusu lemmata začínající na *é-* a dlouhá od 4 do 6 písmen. Jak takový dotaz zadáte? Inspirujte se na internetové adrese <http://uclnk.fj.cuni.cz/bonito> (Dotazovací jazyk). Která slova do frekvence 10 jsou česká a která cizí? Ze kterých jazyků tato slova převážně pocházejí?

- **Úkol 20**
Jak v korpusu vyhledáte slova, která obsahují *r* po tvrdé souhláse? Je třeba k tomu zkombinovat vzory 1a, 1b a 2. Jak omezíte vyhledávkou, budete-li chtít vyhledat jen slova s tvary na 5 písmen (např. tvar *úkoily* má 5 písmen, ale slovo *úkol* 4)? Pomocí N-filtru dále odstraňte slova obsahující velké písmeno na začátku a uvnitř slova. Vypište na základě funkce Frekvenční distribuce (menu Konkordance – Statistiky – Frekvenční distribuce) tato slova do frekvence 150. Znáte je všechna?

B. Tvoření slov/slovo tvorba (morfologie širší)

ÚVOD

V oblasti slovo tvorby je materiál korpusu SYN2000 velmi bohatým zdrojem. Při zadávání dotazů většinou využíváme znaků *, které zastupují tu část slova, které chceme zjistit její předpony a/nebo přípony, srov. např. **lemma="*dlouh"**, **lemma="*prst.*"**, **lemma="*př.*"**, které je však někdy nutné kombinovat s určením slovního druhu, popř. další gramatické kategorie. Takto lze snadno zjistit, jak je která přípona frekventovaná, popř. které slovo je s danou příponou nejfrekventovanější, nebo naopak, se kterými slovními základy se určité přípony nebo předpony pojí.

Úkol 1

Chceme-li si ujasnit, které slovo je v češtině častější a které méně časté, nebudeme jistě chtít vypisovat všechny jednotlivé tvary a pak je počítat, ale budeme hledat pomocí lemmatu (vzor 5A), čímž dostaneme výpis všech tvarů daného slova. Zkuste tedy odhadnout, zda je častější slovo *lov*, nebo slovo *hon*. Potvrďte se vaše domněnka, když se podíváte do korpusu? A jak to bude se slovy *lovení* a *honění*?

Úkol 2

Existují slovo tvorné přípony, které nám pomáhají například vyjadřovat naše chápání věci jako „malé“ nebo „milé“, a to na pozadí výrazu bez této slovo tvorné přípony. Takovým slovům říkáme zdrobněliny, např. *stoleček* chápeme jako malou věc (slovo odvozené od slova *stůl*).

Může se ale stát, že takto utvořená slova se v jazyce zabývali jako samostatná muž. Pak ovšem v jejich tvaru necítíme onen zmiňovaný příznak. Vyberte si jako příklad jednu příponu tvoření zdrobnělin (vzor 8) a vyplňte ty výrazy, které se vyskytují v určitém významu jen jako zdrobněliny, např. slovo *vidlička* má samostatný význam, nejde o pouhou zdrobnělinu od slova *vidlice*. Využijte k tomu seznam Frekvenční distribuce (menu Konkordance – Statistiky – Frekvenční distribuce), a to do frekvence 300. (Při velkém množství výskytů je možné zredukovat počet dokladů pomocí funkce Redukce (menu Konkordance – Redukce – Náhodných x řádků)).

Úkol 3

Častou příponou abstraktních podstatných jmen je přípona *-tí*, kterou se tvoří podstatná jména odvozená od sloves. V případě, že má sloveso příponu *-nout*, jsou tato podstatná jména zakončena na *-nutí*, srov. *rozhodnout - rozhodnutí*. Podle vzoru 5Aa vyhledejte všechna podstatná jména takto zakončená. Porovnejte seznam získaný z korpusu s podstatnými jmény zachycenými v Retrográdním slovníku současné češtiny (1986). Vyplňte na základě Frekvenční distribuce (menu Konkordance – Statistiky – Frekvenční distribuce) 10 nejfrekventovanějších substantiv.

Dále pomocí funkce P-filtr můžete získat spojení těchto podstatných jmen s předcházející předložkou. Zadejte P-filtr pro pozici -1 od hledaného slova a vyhledejte předložky tak, že pro předložku zvolíte její gramatickou značku (tag). Opět zjistěte na základě Frekvenční distribuce seznam těchto kolokací. Která podstatná jména mají více kolokací? Jakou mají tato podstatná jména frekvenci? Které kolokace do frekvence 100 lze označit za ustálené?

Úkol 4

Některé přípony mohou mít dvě varianty. Tak je tomu např. v případě substantivní deminutivní přípony *-élko/-ýlko*. Zkuste si proto v korpusu vyhledat nejprve podstatná jména s příponou *-ýlko* (vzor 5Aa). Pomocí funkce Obecné třídění (menu Konkordance – Obecné třídění – Vynechat duplicity vytvořte seznam slov (lemma), která jsou v korpusu takto zakončena.

Pod jaké heslo je zařazuje SSJC?

Podobně postupujte u přípony *-élko*.

b) Jaký stylův rozdíl je mezi těmito dvěma tvary?

Úkol 5

Vyhleďte-li v korpusu slova tvořená sufixem *-lko* (vzor 8), zjistěte, že patří k různým slovo tvorným typům. Vyjmenujte 3 nejfrekventovanější slova v každé kategorii.

Úkol 6

Která slovesa jsou podle korpusu SYN2000 častější v infinitivu, ta, která začínají *proti-* nebo *ta*, která začínají *spolu-*? Můžete k tomu využít vzor 9. Dokážete rozdílnou frekvenci na základě pohledu do korpusu interpretovat?

Úkol 7

Jak podle korpusu zjistíte, kolik má sloveso *jit* předpon? Můžete zčásti využít vzor 8. Poté si zadejte Obecné třídění (menu Konkordance – Obecné třídění –

Vynechat duplicitu /lemma/) a vypište seznam těchto předponových sloves v infinitivu. Která z předpon nalezených v korpusu SYN2000 není uvedena ve výčtu předpon na konci hesla *jit* v SSJC?

Úkol 8

Řada abstraktních substantiv začíná na předponu *pod-*, např. *podnět*.

Vyhleďte v korpusu podstatná jména ženského rodu s předponou *pod-* podle vzoru 9. Vytvořte si frekvenční seznam všech těchto lemmat pomocí Frekvenční distribuce a u 10 nejfrekventovanějších určete, zda jde o abstraktní nebo konkrétní substantivum.

Která z těchto substantiv mají předponu *pod-* a která jinou předponu?

Úkol 9

Zadejte v korpusu vyhledat slova s předponou *ex-* podle vzoru 9.

Potom na základě Obecného třídění (menu Konkordance - Obecné třídění - Vynechat duplicitu) zjistěte, kolik sloves v infinitivu s předponou *ex-* obsahuje SYN2000.

Znáte alespoň u pěti z nich jejich význam?

Jak zadáte dotaz, budete-li chtít vyhledat nejen infinitiv, ale všechny slovesné tvary?

Změnila se situace oproti předchozímu seznamu při vynechání duplicit? Pokud ano, kolik sloves přibýlo a která to jsou?

Úkol 10

Který z nabízených vzorů se hodí pro vyhledání substantiva začínajícího spolu s příponou *-tel*? Vypište jejich seznam a porovnejte se SSČ nebo SSJC.

Úkol 11

Každý z nás se od 6 let setkává se školou ve všech formách. Možná si myslíte, že si tedy sami vzpomenete na všechna slova, která se školou a školstvím souvisí. Zkuste to. Vytvořte sami seznam slov, která obsahují řetězec *přeško/* a potom se pokuste taková slova nalézt v korpusu. Od kterého slova jsou všechna další odvozená? V jakém oboru se užívá sloveso *přeškovat*?

Úkol 12

Abstraktnímu přejatému podstatnému jménu *pozice* odpovídá domácí slovo *místo*. Vyhleďte od tohoto cizího slova v korpusu všechny odvozeniny, tj. s předponami i s příponami (pozor: kmen může být zakončen na *c* i *č*). Zkombinujte

k tomu vzory 10 a 2. Vytvořte si frekvenční seznam lemmat (menu Konkordance - Statistika - Frekvenční distribuce/lemma/). K 5 nejfrekventovanějším odvozeným substantivům se pokuste najít české ekvivalenty.

Úkol 13

Podobně jako v předcházejících cvičeních se pokuste v korpusu vyhledat odvozená slova obsahující *narko* i *narťo*. Jak bude vypadat korpusový dotaz?

Kolik odvozených slov v korpusu je?

Na základě výsledku vyhledávání a seznamu získaného Frekvenční distribucí (menu Konkordance - Statistika - Frekvenční distribuce/lemma/) a jeho srovnání s popisem v SSJC zohodnoťte situaci a zvažte, zda dnešní význam slov odvozených od *narco* není jiný a proč tato slovo tvorná řada tolik expandovala. Jaké slovo tvorné postupy tu jsou využity?

Úkol 14

V češtině existuje v mluvném jazyce skupina expresivních slov typu *bágo*, *točo*, *vzůřo*. Zjistěte, zda se tato a podobně tvořená slova vyskytují v korpusu, tj. vyhleďte substantiva středního rodu, která jsou zakončena na *o* s předcházející souhláskou *r*, *g*, *š*, *ž*, *č* nebo *ň* v korpusu (vzor 10 v kombinaci se vzorem 2). Ze seznamu získaného pomocí Frekvenční distribuce (menu Konkordance - Statistika - Frekvenční distribuce/lemma/) vypište expresivní substantiva. Jakou mají frekvenci? Zjistěte na základě korpusu, co znamená slovo *čůčo* nebo *čůčo*. Která další slova takto zakončená patří do stejné významové skupiny?

Úkol 15

Příponu *-ouš* obsahují substantiva, která v některých případech vyjadřují expresivní vztah mluvčího k pojmenování. Jsou mezi nimi vlastní jména, ale i jména obecná. Některá ze substantiv vyjadřují kladný, jiná záporný postoj mluvčího. Vyhleďte tato substantiva mužského rodu (životná i neživotná) v korpusu podle vzoru 8.

Seznam slov vytvořený na základě Frekvenční distribuce (menu Konkordance - Statistika - Frekvenční distribuce/lemma/). Rozdělte podle těchto kritérií uvedená slova do těchto tří skupin: vlastní jména, obecná jména kladná, obecná jména záporná. Které slovo nemá příponu *-ouš*?

Některá z těchto slov jsou zatím natolik nová, že mají přičten v korpusu slovní druh X, tj. nerozpoznaný, nikoli N, tj. substantivum. Nenajdeme je proto mezi substantivy. Tato slova je však možné vyhledat podle lemmatu. Co znamenají slova *biguš*, *mikrouš*, *tajouš*, *trafouš*? Použijte k tomu korpus! Zadejte každé slovo zvláštním dotazem pomocí lemmatu.

Od kterých adjektiv nebo substantiv (popř. jejich spojení) jsou odvozena substantiva s touto příponou, která mají v korpusu 1-2 výskyty: *pakouš, prudlouš, psychouš, cukrouš, čilouš, křivouš, socdemouš*?

Úkol 16

V korpusu můžeme také vyhledat kompozita s komponentem *auto-*. Nejprve se podíváme, jak je tato skupina slov popsána v SSJC.

Definice ze slovníku SSJC:

auto- (z řec.) první část složených slov spojující část druhou s významem řec. *slova autos sám, a to jako s původcem a zároveň předmětem děje (autodiadakt samouk, autokritik kdo kritizuje sám sebe, autokarikatura vlastní karikatura, autogamie opylení vlastním pylem, autokracie samovláda) II první část složených slov spojující část druhou s významem slova automobil, a to I, jako s jejím předmětem (automontér kdo montuje auta) 2, jako s její vlastností (autostřáda automobilová silnice; autoklub klub automobilistů; často v slang. složeninách s druhou částí domáčko původu: autodoprava za automobilová doprava, autosprávkána, autoučiliště, autořidič aj.)*

A. My budeme pracovat s významem II.

Odhadněte kolik substantivních kompozit (s předponou *auto-* v druhém významu) a jejich derivací obsahuje SYN2000. Inspirujte se vzorem 9.

B. Zadejte si slova *auto; automobil; autobus a autokar* do SYN2000 pomocí lemmatu. Zapište si výsledky do tabulky. Nyní vyhledejte adjektiva odvozená od těchto substantiv. Také tyto výsledky si zapište do tabulky. Porovnejte výsledky. Jak byste zdůvodnili skoro nulovou frekvenci slova *autový*?

Úkol 17

Zvažte okolnosti, které vedou k tomu, že od jednoho podstatného jména se tvoří dvě přidavná jména velmi blízkou slovotvornou příponou (*-ální* a *-ní*): *strukturaální x strukturální*. Co byste dokázali o dvojici přidavných jmen utvořených od slova *struktura* říct na základě korpusu? Po vyhledání každého lemmatu podle vzoru si zadejte statistickou funkci Kolokace (menu Konkordance – Statistky – Kolokace/lemma/), tam si zvolte rozsah 1, 1 a minimalní četnost v korpusu i v daném rozsahu rovněž zvolte 1. Seřídte slova podle absolutní frekvence.

Dokážete určit sklon k užívání těchto slov v určitých oblastech nebo oborech?

Úkol 18

Podobně, jako jste hledali slova typu *spolunajitel* (cvičení 10), si můžete vyhledat podstatná jména např. s předponou *vy-* a příponou *-ní*. Tato přípona je charakteristická pro slovesná substantiva. Použijeme k tomu vzor 11.

A. Od sloves kterého vidu jsou nejčastěji tato slovesná substantiva odvozena?

Podívejte se na seznam slov získaný Frekvenční distribucí (menu Konkordance – Statistky – Frekvenční distribuce /lemma/) do frekvence 1000 a zapište si poměr nedokonavých a dokonavých sloves.

B. Pomocí P-Filtu (pozice –1 v levém kontextu) se podívejte pouze na spojení předložek *během* a v *průběhu* s těmito podstatnými jmény. Můžete k tomu využít vzor 17 (word, popř. lemma *během* a word *průběhu*). Jaký počítáte významový rozdíl mezi *během vyšetření* a *během vyšetřování*? Jaký je poměr výskytů mezi *během vyšetření* / *během vyšetřování* / v *průběhu vyšetření* / v *průběhu vyšetřování*?

Úkol 19

Určete substantiva mužského rodu (životná), která nejsou činitelskými jmény (např. *učitel*) a končí na *-tel*. Jak zadáte dotaz? Použijte k tomu výpis Frekvenční distribuce (menu Konkordance – Statistky – Frekvenční distribuce /lemma/) do frekvence 10.

Úkol 20

Jak se významově liší činitelská jména utvořená od jednoho slovesa, jednou příponou *-tel* a podruhé příponou *-č*?

Pozorujte v korpusu následující dvojice (zadávejte pomocí lemmatu):

nakladatel x nakladač

nositel x nosič

ředitel x řidič

ovladatel x ovladač

volitel x volič

Jak se slova ve dvojicích liší významově? Jak se liší frekvence členů jednotlivých dvojic?

C. Tvarosloví/Morfologie užší

ÚVOD

Korpus SYN2000 je gramaticky označovaný, můžeme v něm proto vyhledávat podle slovních druhů a dalších gramatických kategorií. Studium korpusu nám pomůže pochopit některé gramatické jevy, u kterých bychom bez frekvenčních údajů a statistických pomůcek stěží hledali nějakou pravidelnost nebo zákonitost. Při zkoumání morfologických jevů se můžeme zaměřit i na určování slovních druhů v případech typu *Stalo se to nešťastnou náhodou* a *Náhodou to věděl*, tj. když plnovýznamové slovo přechází do slova neplnovýznamového. V takových případech nemusíme vycházet z gramatických značek a můžeme se pouze řídit územ daného tvaru. Dále nám korpus může pomoci odlišit abstraktní a konkrétní významy určitých tvarů, základní kolokáty některých předložek, jejich zapojení do frazeologie apod. Centrální je však studium variantních tvarů typu *obchodu* x *obchodě* nebo *sází* x *sázejí*.

Úkol 1

Některé tvary podstatných jmen se můžou užívat ve funkci jiných slovních druhů, srov. např. *kolem*, *během*, *náhodou* apod. Vyhleďte v korpusu tvar *dík* (vzor 1) a určete, o jaké jde slovní druhy. Na základě čeho je možné tyto slovní druhy od sebe odlišit?

Úkol 2

V korpusu je také možné vyhledat slovní tvary složené z kombinace různých písmen (vzor 3). Pokuste se vyhledat tvary složené ze čtyř různých souhlásek, např. *smrt*. Pomocí Frekvenční distribuce dostanete seznam těchto tvarů seřazený podle frekvence. Prohlédněte si tvary s frekvencí 1000–10 a zjistěte, který z těchto tvarů je současně příslovcem i předložkou.

Vyhleďte tento tvar v korpusu zvlášť a na náhodném výběru 100 výskytů (menu Konkordance – Redukce – Náhodných 100 řádků) určete, kolik dokladů patří k příslovnému a kolik k předložkovému užití.

Jaké dva významy má předložka?
Které doklady jsou frazeologické?

Úkol 3

Předložka *během* něčeho vznikla ze 7. pádu podstatného jména *běh*.

Vyhleďte v korpusu lemma *během*. Pomocí statistické funkce Kolokace (menu Konkordance – Statistika – Kolokace) zjistěte 100 slov (lemmat), která následují na pozicích 0–3 setříděných podle absolutní frekvence. Vypište z tohoto seznamu podstatná jména. Dokážete z tohoto seznamu určit, která významová skupina u těchto podstatných jmen převládá?

Úkol 4

Ptáte se, jak se ke slovům v textech přidávají gramatické informace? Při velkém množství textů – jaké se podařilo shromáždit v korpusu SYN2000 – není možné, aby původní texty doplňovali o gramatické informace (o „lemmata“ a „tagy“) lidé. Dělalji to tedy počítače na základě různých statistik a způsobem, který si pracovně nazveme „statistický odhad“. Protože jde jen o odhad, můžeme při podrobném zkoumání textů odhalit tu a tam chybu. (V celém korpusu však chyby ve značkování nepřečkáují 5 až 8 %.) Všimněme si takto např. 4. pádu podstatného jména *zdraví*, může existovat v plurálu? Jak vypadá značkování v korpusu? Zobrazte si značky v korpusu (menu Zobrazení – Atributy – tag). Které příklady označené v korpusu jako 4. pád singuláru, jsou naopak tvary jiného pádu, popř. jiného slova než substantiva *zdraví*?

Úkol 5

Vyhleďte slovo *obchod* v 6. pádě jednotného čísla (vzor 6). Odpovídá tomuto pádu v korpusu pouze jeden tvar? Pokud ne, lze z kontextu vysledovat, jak se použít těchto tvarů liší, který je častější a proč? Znáte podobná substantiva, která můžou mít více tvarů pro jeden pád? Použít jejich koncovek ověřte v korpusu.

Úkol 6

Původně abstraktní podstatná jména můžou mít někdy význam konkrétní. Lze to ověřit na přídatných jménech, která se s nimi spojují.

Vyhleďte v korpusu lemma *potřeba* ve 4. pádě jednotného i množného čísla (vzor 6) a pomocí P-Filtu v kontextu –1, 0 vyhleďte adjektiva. Potom na základě Frekvenční distribuce (menu Konkordance – Statistika – Frekvenční distribuce) si vytvořte seznam kolokací tvaru *potřeba/potřeby* s lemmatem na první pozici zleva, a to s frekvencí deset a více. V tomto seznamu se nejprve zaměřte na přídatná jména, která jsou odlišná u tvaru jednotného a množného čísla, tj. *sportovní* ap. *potřeby* x *naléhavou potřebu*. Jaký význam (abstraktní či konkrétní) má podstatné jméno *potřeba* u první skupiny a jaký u skupiny druhé?

Liší se významově tvary *potřeby* a *potřebu* i v případech, že předchází stejné přídavné jméno, např. *osobní potřeby* x *osobní potřebu*, popř. *lidskou potřeby* x *lidskou potřebu*?

Úkol 7

V 1. pádě množného čísla vzorů *pán/muž* jsou možná dvě zakončení *-i/-ové*.

A. Zjistěte, které zakončení frekvenčně převažuje u následujících slov:

pán, syn, politik, předek, svědek, vřadce, historik, bůh.

B. Můžeme vidět nějakou tendenci v použití těchto tvarů?

Jak zadáte korpusový dotaz obecně?

Úkol 8

Vyhleďte lemma *káně* v genitivu podle vzoru 6. Ve vyhledávce odstraňte výskyty s velkým písmenem na začátku. Zjistěte frekvence jednotlivých tvarů (ve femininu i v neutru).

Úkol 9

Vyhleďte si v korpusu postupně 1. pád plurálu životných adjektiv mužského rodu s následujícími příponami: *-ní, -ný, -ský, -cký, -ový, -natý, -itý, -ovitý, -avý, -ivý, -ěný*. Mění se podoba těchto přípon v jednotném a množném čísle? Které přípony mají souhláskovou změnu? Jak zadáte dotaz? Najděte několik příkladů adjektiv, u nichž není uvedené zakončení příponou.

Úkol 10

Zadejte podle vzoru 7a posloupnost předložky *před* a variantních tvarů podstatného jména *vánoce* v instrumentálu, tj. *vánoci* x *vánoce*mi.

A. Kolik výskytů se objeví? Jaké mají tvary a jak jsou psány? Který tvar je frekventovanější? Využijte k tomu funkci Frekvenční distribuce (menu Konkordance – Statistiky – Frekvenční distribuce-KWIC) + přidat <KWIC>.

B. Je dodržována zásada PČP, že *Vánoce* píšeme s velkým písmenem?

Úkol 11

Vyhleďte dvě předložky na začátku věty. Který vzor pro tento dotaz použijete? Zjistěte, která spojení dvou předložek jsou nejfrekventovanější.

A. Které z těchto dvojic dnes tvoří tzv. víceslovnou předložku? Zobrazte si atributy – tagy a sledujte, jak jsou jednotlivé předložky specifikovány na druhé pozici morfoloické značky.

B. Ve kterých případech nejde o sled dvou předložek?

Úkol 12

Zadejte podle vzorů 7 a 17 dotaz týkající se tvaru *Praze* a jeho spojení s předložkami *v* a *na*.

A. V kolika případech je slovo spojeno s předložkou *v* a v kolika s předložkou *na*?

B. Za jakých podmínek se používá předložka *na* ve spojení *na Praze*?

Doplňte jiná slova, která se mohou pojit s oběma předložkami. Zkusíte třeba *Kladno*, *Kolej*, *Vinohrady*...

Úkol 13

Vyhleďte v korpusu minulé přičestí od sloves zakončených na *-pnout* (např. *nápjál, kopl, stoupnul*). Podle kterého vzoru jste přičestí vyhledali? Zjistěte pomocí frekvenční distribuce 10 nejfrekventovanějších sloves, která sem patří. Potom prostřednictvím funkce Obecné třídění (menu Konkordance – Obecné třídění) zadejte vynechání duplicitních tvarů (word). Ziskáte tak po 1 výskytu od každého tvaru. V tomto seznamu vyberte pouze ty tvary, které se vztahují k výše uvedeným 10 nejfrekventovanějším slovesům a zjistěte poměr variant, např. *kopl* x *kopnul*. (Za samostatné varianty však nelze považovat tvary, které se liší pouze velkým a malým písmenem.)

A. Která slovesa mají varianty jen v mužském rodě, která nemají tvar středního rodu a která mají zastoupeny variantní tvary ve všech rodech a ve všech číslech? B. O jaké typy nebo skupiny sloves se jedná?

C. Lze z kontextů sloves, která mají rovnoměrně rozložené varianty, vysledovat různé významy těchto variant?

Úkol 14

Ve větách typu *je tma, je ticho, něco je vysoko* označujeme slova *tma, ticho* a *vysoko* někdy jako predikativa. Při práci s korpusem je však třeba vyhledané doklady projit a odstranit ty případy, které nespňňují danou lingvistickou kategorii.

Vyhleďte v korpusu následující posloupnost: lemma *být* nebo *viset*, za nimi lemma *být*, které se však nemusí realizovat (pro tvary minulého času), dále 0-2 adverbia a tvar *vysoko*. Potom zadejte P-Filtr tak, aby za touto posloupností následovala interpunkce (tag pro interpunkci obecně je Z na 1. pozici) nebo spojka *a*. Kolik dokladů jste získali? Jaký poměr mezi slovesy *být* a *viset*? Jak to bude vypadat, když do dotazu zadáte místo *viset* sloveso *ležet* a místo *vysoko* tvar *hluboko*? Jaký je v obou vyhledávkách repertoár užívaných adverbíí?

Úkol 15

Vyhleďte přivlastňovací adjektiva *Havův, Klausův, Zemanův* následovaná substantivem (shodný přivlastek) – vzor 7. Zjistěte pro každé adjektivum tři nejfrekventovanější substantiva a ta pak vyhleďte jako neshodný přivlastek: substanti-

vum + libovolná pozice + *Havla* (př. *výrok prezidenta Havla*) – vzor 13. U těchto vybraných adpektiv zjistěte, který způsob psaní je frekventovanější. Jak byste tyto možnosti v jednotlivých kontextech hodnotili z hlediska vhodnosti vyjádření?

■ Úkol 16

Už jste někdy dostali vyhubováno od maminky? Možná ano, možná ne, ale určité více dobře, co znamená *dostat vyhubováno* – vzor 7a. Se kterými dalšími slovesy (ve tvaru minulého pasivního přechází středního rodu) se sloveso *dostat* používá?

Myslete, že by bylo možné podobně použít za sebou dokonce tři slovesa? Není to tak těžké, uvědomte si, jak vypadá celý soubor časovaných tvarů (jakéhokoli slovesa (obecně) a zkuste takovou kombinaci tří sloves za sebou najít!

■ Úkol 17

Pokud jste úspěšně vyřešili předcházející cvičení, můžete pokračovat se třemi slovesy. Zajímal-li vás všechny časy v různých variantách slovosledu, můžeme se na ně pět jedním dotazem. Navrhněte jak.

U podobných konstrukcí se navíc při vyhledávání můžeme omezit na minulý čas a v něm na první a druhou osobu jednotného i množného čísla. Jak byste to provedli?

■ Úkol 18

Vyhledejte v mluvených korpusích (PMK, BMK) a v korpusu psaném SYN2000 pasivní tvary typu *uklízino* (vzor 1b). Porovnejte jejich užívání v mluveném a psaném jazyce a mezi mluvenými korpusy navzájem. S jakými slovesy se tyto tvary pojí? Od jakých sloves se obvykle tvoří?

■ Úkol 19

Najděte v korpusu taková slovesa, u kterých se ve 3. osobě plurálu užívají jak varianty s koncovkou *-ou*, tak s koncovkou *-í*. Omezte dotaz na ty tvary, které mají před koncovkou *-ou* souhlásky *ž, š, č*. Pomocí funkce Frekvenční distribuce (menu Konkordance – Statistika – Frekvenční distribuce /lc/ získáte seznam těchto tvarů. Ke každému tvaru do frekvence 15 připojte variantu s koncovkou *-í* (pokud existuje), např. *pláčou – pláčí*. Ke každému tomuto tvaru vyhledejte frekvenci v korpusu (dotaz je třeba zadat pomocí lc, tj. $lc="pláčí"$), který zachytí tvary s malým i velkým písmenem na začátku. Porovnejte frekvence jednotlivých variant a uvažujte o rozdílech v jejich užití.

■ Úkol 20

Vyhledejte v korpusu slovesa, která mají ve 3. osobě pl. koncovku *-ejí*. Vytvořte si na základě Frekvenční distribuce jejich seznam (menu Konkordance – Statistika – Frekvenční distribuce/lemma/). Podívejte se na slovesa do frekvence 200 a sepište, u kterých typů sloves se tato koncovka objevuje – tj. pokud má určitý slovesný základ více předpon, запиšte ho jen jednou, např. ze sloves *navičet, vyvičet* vznikne *–vičet*.

D. Slovní zásoba/lexikologie

ÚVOD

Zkoumání slovní zásoby je jednou z oblastí, kde korpusová data pomáhají nejvíce. V korpusu vyhledáváme jednotlivá slova většínou pomocí lemmatu, důležité je však využít statistické funkce Kolokace a setřídění slov, se kterými se hledané slovo spojuje, buď na základě absolutní nebo relativní frekvence. Můžeme tak zjistit spojitelnost slov málo frekventovaných, např. *línout se*, můžeme také ale zjistit význam méně známého slova, např. *odúmrť*. Některá substantiva se užívají jen s omezeným okruhem sloves, např. **zájem** – *vzbudit, vyvolat zájem, mít zájem, ztratit zájem o něco* apod.

Úkol 1

Některá slova mají významy konkrétní, významy přenesené a mnohá také vstupuji do frazémů. Někdy se stává, že jedna skupina tvarů inklinuje k tomu označovat významy konkrétní a jiná skupina tvarů se užívá spíše přeneseně či idiomatically. Abychom si to ukázali názorně, podívejme se na podstatné jméno *bačkora* a na rozdíl mezi jeho užitím v jednotném a v množném čísle.

Všimněme si např. tvaru 4. pádu slova *bačkora*, totiž *bačkoru* (vzor 1).

Co o něm můžeme z hlediska významu (v. konkrétní, přenesený, uvnitř frazému) říci?

Úkol 2

Znáte slovo *štěpnice*? Vyhledejte je v korpusu tak, abyste získali jeho různé tvary (vzor 1a). Potom si zadějte zobrazení zdroji.

- A. V jakých typech textů toto slovo najdete?
- B. V které knize je tento výskyt nejčastější?
- C. Jak byste vysvětlili, co toto slovo znamená?
- D. Stačí nám nalezene doklady k tomu, abychom dokázali určit jeho význam?

Úkol 3

Podobným způsobem jako v předcházejícím cvičení v korpusu zjistěte, co znamená slovo *volňasek*.

Úkol 4

Když přecházejí vlastní názvy do obecného jazykového užití, nazývá se tento proces apelativizace. Můžete si to ověřit na slově *Mercedes* x *mercedes*, vyhleďte-li je podle vzoru 2 a zkombinujeme tento vzor se vzorem 1a. Připomeňme si co říká o apelativizaci PCP.

- a. Jména některých výrobků, plodů a rostlin označují jejich typy.
- b. Tato označení se užívají také jako obecná pojmenování druhová a píšou se pak s písmenem malým.
- c. Někdy dochází i ke změně podoby slova, běžně se od nich tvoří množné číslo.

A. Uvedte příklady:

- B. Našli jste ve vyhledávce i tvary jiného slova než *mercedes*? Jakého?
- C. Znáte ještě další názvy, které se používají pro toto auto?
- D. Zadejte **med.***. Je možné i toto slovo psát s velkým písmenem?

Úkol 5

Do čeho všeho se dá vlastně investovat a tedy jaký je vlastně dnešní význam tohoto slova? Liší se podle korpusu *investovat do něčeho* od *investovat něco*?

Úkol 6

Vyhledejte si v korpusu slovo *třešnička* jako lemma.

- A. Pomocí funkce Frekvenční distribuce (menu Konkordance-Statistiky – Frekvenční distribuce/word) můžeme zjistit, který tvar se vyskytuje nejčastěji. Proč? O jaký jde frazeologický kontext?
- B. Objevuje se v korpusu obměna tohoto frazému? Jakým jiným frazémem byste ji nahradili?
- C. Pokud je toto slovo napsáno s velkým písmenem, jde vždy o začátek věty?

Úkol 7

Abstraktní substantiva mají zpravidla omezené textové okolí. Vyhledejte si v korpusu lemma *souvislost* a potom pomocí funkce Kolokace (menu Konkordance – Statistiky – Kolokace /lemma) si vytvořte seznam slov na 1. pozici zleva (-1, 0) podle absolutní frekvence. Jaká jsou 3 nejfrekventovanější slovesa? Porovnejte je se slovesy uvedenými v hesle **souvislost** v SČF13.

Úkol 8

Zadejte si v korpusu dotaz na vyhledání lemmatu *třbít*. Potom si setřídíte abecedně slova na pravé straně (menu Konkordance – jednoduché třídění – Právý kontext).

- A. O čem dnes říkáme, že se *třbít*? Uvedte alespoň deset spojení.

- B. Co znamená sloveso *tříbit*?
C. Víte, co toto slovo znamenalo původně?

Úkol 9

Vyhledejte si lemma *odumrt* a zjistěte jeho význam. Projeví se na významu slova jeho užití v různých typech textů? Jak byste toto slovo charakterizovali? Najděte v kontextech definici či vysvětlení tohoto slova?

Úkol 10

Vyhledejte si v korpusu lemma *akce* a pokuste se najít kontexty, které byste přiřadili k následujícím významům:

- a) činnost zaměřená k určitému cíli, zvláště široce organizovaná, postup, jednání
b) činnost, fungování, působení
c) představení

Najdete ještě případně jiný význam, který byste vyčlenili? Jak vypadá výskyt tohoto slova v mluvených korpusech? Co znamená spojení *hurá akce*?

Úkol 11

Některá slova mají omezenou spojitelnost a bývají součástí frazémů. Patří k nim například slovo *pozor*. Sepište si nejprve vlastní seznam spojení, která se vám vybaví při vyslovení tohoto slova. Potom si vyhledejte v korpusu lemma *pozor*. Pomocí N-filtru odstraňte kontexty s předcházející interpunkcí (tj. [tag="Z.*"] na pozici -1). Na základě statistiky Kolokace (menu Konkordance – Statistika – Kolokace /lemma/) zjistěte kolokace v levém kontextu v rozsahu -3, 0, seříděné podle absolutní frekvence. V tomto seznamu se podívejte na prvních 40 nejfrekventovanějších lemmat a pokuste se vypsat nejčastější slovesná a adjektivní spojení (frazémy) včetně předložek, které se v nich objevují. Svůj seznam si porovnejte s doklady v korpusu.

Úkol 12

Pomocí korpusu zjistěte, co a odkud se může *řinout* (vzor 5). Korpusové doklady porovnejte se SSC.

A. Vypište nejčastější substantiva, která se objevují v okolí tohoto slovesa.

B. Pokuste se sestavit vlastní slovníkové heslo na základě korpusu.

Úkol 13

Vyhledejte v korpusu současně lemmata *vůně*, *zápach*, *zvuk*, *melodie* (vzor 17) a pomocí P-filtru ponechte jen ty kontexty, v nichž je na pozici -1 sloveso. Po-

tom odstraňte duplicity – lemmata (Konkordance – Obecné třídění na dvě pozice, tzn. Přidat další třídící kritérium, potom použijte Konkordance – Statistika – Kolokace – lemma /rozsah-1,0/). Která slovesa – kromě slovesa *být* – se pojí se všemi čtyřmi substantivy? Dokázali byste vysvětlit význam tohoto slovesa?

Úkol 14

Vyhledejte v korpusu postupně lemmata *bujný* a *divoký* a podívejte se, v jakých kontextech se jich užívá. Jak byste vyjádřili jejich základní rozdíly – s jakými substantivy se pojí? O jaké jde významové skupiny? Která spojení jsou frazeologická?

Úkol 15

Vyhledejte v korpusu současně lemmata *mysl* a *význam* (vzor 17) a zredukujte počet výskytů na 1000. Poté pomocí P-Filtru ponechte pouze kontexty, které v rozsahu -3 a 3 obsahují lemma *slovo*. Kontexty si navíc můžete seřadit abecedně do 3 slov v levém kontextu (funkce Jednoduché třídění) a porovnat. Jaký je nejtypičtější kontext pro *mysl* a jaký pro *význam*?

Úkol 16

Vyhledejte v korpusu slova *přehřel* a *přehřít* (buď postupně nebo pomocí vzoru 3). Určete, jakého jsou rodu a jaký mají význam. Liší se ve významu nebo jde pouze o jinou hláskovou variantu? Jedná se vždy o jeden význam? Jakým synonymem byste tato slova nahradili? Porovnejte s definicí v SSC.

Úkol 17

Vyhledejte v korpusu lemma *řáčka* a podle Frekvenční distribuce (menu Konkordance – Statistika – Frekvenční distribuce /word nebo lc/) zjistěte zastoupení jednotlivých tvarů. Které 3 tvary jsou nejfrekventovanější? Vypište z korpusu u každého z těchto tří tvarů jeho nejtypičtější kontext.

Úkol 18

Některá slova v sobě obsahují, byť nevyčleně, významy slov dalších. Existují například slovesa, která ve svém významu takto zahrnují i objekty, kterého se činnost vyhnadně týká a tento objekt pak nemusí být v textu explicitně uveden. Například víme, že pokud někdo *zamkal*, pak zřejmě *očima* (vzor 4). Existují však texty, ve kterých jsou tato doplnění uvedena explicitně. Přečtěte si několik příkladů (např. *zableskal rukama*, *zamrkala očima*) a ohodnoťte styl. Dokážete si kontexty představit bez substantiv? Jaký je poměr mezi vyjádřením s objektem (instrumentem) a bez něj?

■ **Úkol 19**
Zjistěte na základě korpusu rozdíly v užití adverbii *dlouho* x *dlouze*, *široko* x *široce*. (Každou dvojici podle vzoru 17, nebo zvlášť jako lemma) Sledujte, jaký slovní druh za adverbii *široko*, *široce* následuje. Které spojení je nejčastější? U dvojice *dlouho*, *dlouze* si naopak všimněte sloves, která stojí za nimi – která se liší a která se shodují?

■ **Úkol 20**
V čem se liší a v čem se shodují slova *odvislý od* a *závislý na*? Jakou mají frekvenci? Můžou existovat kombinace *odvislý na* a *závislý od*? Pro vyhledání v korpusu zkombinujte vzory 12 a 17.

E. Kolokace (Slovní spojení)

ÚVOD

Studium kolokací, resp. slovních spojení, patří k velmi typickým oblastem, resp. doménám korpusového zkoumání. Prostřednictvím korpusu je totiž možné mimo jiné zjistit nejen to, jak se používá, ale i to, zda je daná kombinace slov dostatečně frekvencovaná, tj. ustálená. Pozornost se soustřeďuje především na frazémy, např. *ležet ladem*, a termíny, např. *kyselina sírová*. Zejména v oblasti frazeologie korpus umožňuje lépe proniknout do funkce a významu těchto víceslovných jednotek. Zadávaní korpusových dotazů využívá především početní lemmat, někdy doplněných specifickými gramatickými kategoriemi, s různým počtem vložených slov, např. v *důsledku čeho*, v *konečném důsledku*, ve *svém konečném důsledku* apod. Je rovněž možné zadat určitou strukturu a sledovat její zapojení do více frazémů, např. *byť na: přešes, spadnutí, mizně, huntě, pováženou, štru, vážkách* apod.

Úkol 1

Vyhledejte kolokace *na začátku* a *na konci* a za nimi následující substantiva s příponou *-ní*. Která substantiva se častěji používají s kolokací *na konci* a která s kolokací *na začátku*? Můžete k tomu zkombinovat vzory 12 a 15 (jako atributy využijte lemma, lc a pos).

Úkol 2

Víte, co znamená spojení *třešnička na dortu*? Vyhledejte ho v korpusu a popište, co v daných kontextech znamená a v jakém druhu textu se vyskytuje. Objevuje se v dokladech i jiný tvar než *dortu*? (srov. úkol D6).

Úkol 3

Podle vzoru 4 vyhledejte v korpusu předložku v *průběhu*. Tato kombinace tvarů se vyskytuje v korpusu 6616x, je tedy dostatečně ustálená a lze ji za víceslovnou předložku považovat. (srov. Úkol C3).

Pomocí statistické funkce Kolokace zjistěte 100 slov (lemmat), která následují na pozicích 0–3 seříděných podle absolutní frekvence. Vypište z tohoto seznamu substantiva označující časový úsek. Srovnajte tato substantiva se seznamem substantiv pro předložku *během*. Jak se tyto seznamy liší?

Úkol 4

Přestože se nám zdá, že víceslovné jednotky jsou méně proměnlivé než jednotlivá slova, skutečnost nás může někdy překvapit. Co si myslíte, že vám ukáže korpus, zadáte-li si do něj kombinaci *dát + koš* ve všech tvarech? Zkuste zadat tuto kombinaci nejen těsně za sebou, ale i s 1 až 3 vloženými slovy (vzor 14). Vyhledejte podobné spojení slovesa *dostat* se substantivem *koš*.

Úkol 5

Vyhledejte si spojení *železný oř* a *železný kůň* ve všech tvarech. Můžete z příslušných kontextů poznat, o co se jedná? Jsou tato spojení zaměnitelná? Doveďte odhadnout, co vedlo k jejich vzniku?

Úkol 6

Některá abstraktní podstatná jména jsou již zastaralá, a tudíž řídká. Udržují se jen v pevných kolokacích. Tak např. podstatné jméno *pokraj* existuje jen s předložkou *na*, a to ve 4. a v 6. pádě, tj. *na pokraj* a *na pokraji*. Vyhledejte proto v korpusu kombinaci lemmat *na* a *pokraj*.

Na základě Frekvenční distribuce seříděné podle 1. slovního tvaru za touto vyhledanou kombinací zjistíte, která podstatná jména (s frekvencí vyšší než 10) následují nejčastěji. Dokázali byste v této skupině slov určit společný významový rys? Na základě statistické funkce Kolokace dále zjistíte, která slovesa se objevují do pozice -5 od hledané kombinace slov, a to na základě relativní frekvence. Vypíšte všechna slovesa, u nichž je MI-score vyšší než 5,00. Která z těchto sloves se pojí se čtvrtým a která se šestým pádem?

Úkol 7

Zadejte v korpusu kombinaci lemmat *šedý* a *zóna* (vzor 12).

A. V jakých oblastech se toto spojení užívá?

B. Napište význam tohoto spojení.

C. Znáte i jiná slovní spojení se slovem *šedý*? Vyhledejte v korpusu spojení, která jsou buď frázemý nebo termíny.

Úkol 8

Český jazyk využívá pro vytváření idiomů slova z různých oblastí. Některá z nich můžeme najít dokonce v různých obměnách, variantách, někdy dokonce i v na-prosto odlišných idiolech. Tak je tomu nejen u slova *kočka*, ale dokonce i u souvškytu slova *kočka* a *mys*. V češtině máme frázem *hrát si s někým jako kočka s myší* i frázem významově posunutý *hra na kočku a myš*. O kterém z nich si myslíte, že je častější? Vyhledejte v korpusu kombinaci slova *kočka* a *mys* s jedním vlo-

ženým slovem (vzor 14). Poměr spojení *kočka a myš* a *kočka s myší* získáme na základě funkce Frekvenční distribuce (menu Konkordance – Statistika – Frekvenční distribuce (na pozicích 0, 1 a 2 – současně na všech pozicích vyberte lemma). Kolik kontextů jsou frázemý? Při pročítání kontextů a počítání vám pomůže, když si seřídíte konkordance pomocí jednoduchého třídění (menu Konkordance – Jednoduché třídění – Levý kontext /word, počet tříděných pozic 1/).

Úkol 9

Vyhledejte v korpusu, zda je častější spojení *světle* (název barvy) + *barva* nebo *tmavě*. Který název barvy se tu objevuje nejčastěji? Které barvy se objevují v obou spojeních? (vzor 12) Zkuste výsledek nejprve odhadnout a pak ověřit na korpusu.

Úkol 10

Zadejte kombinaci lemmat *být* a *balík* s jedním vloženým slovem (vzor 14).

A. Co myslíte, jak časté jsou frázemý v textu? Vyberte kontexty, kde se jedná o frázem *být v balíku*, tj. „mít hodně peněz“. Je podmět tohoto frázemú životný nebo neživotný?

Všimněte si, že slovo *balík* je vždy ve stejném pádě. Kterém?

B. Nyní zkuste zadat do stejné struktury lemma *být*, ale za ním tvar *balíku*. Co si myslíte, zvýší se počet výskytů tohoto frázemú s dalším vloženým slovem?

Vyberte opět kontexty, kde se jedná o frázem *být v balíku*, tj. „mít hodně peněz“. Srovnejte výskyt s předchozí částí cvičení. Z jakých částí (komponentů) se tento frázem skládá?

Úkol 11

Slova *dát* a *dohromady* se vyskytují v několika frázemech. Vyhledejte tuto kombinaci s dvěma vloženými slovy (vzor 14a) a pokuste se odhalit několik různých frázemů.

Úkol 12

Pokuste se vyhledat lemmata *oblast* a *život* tak, aby mezi ně byla vložena právě tři slova (vzor 14a). Z této vyhledávky se pokuste zjistit, o jakých „oblastech života“ se nejčastěji píše.

Úkol 13

Jistě jste si už všimli, že některé výrazy můžeme pohodlně rozvíjet, některé nikoli. Dokážete následující příklady najít v korpusu a vysvětlit: *letní dub*, *dub letní*, *letní den*, *den letní*? Zadejte je podle vzoru 14 s 0 až 4 vloženými slovy.

■ **Úkol 14**
Vyhleďte frazém *nechat někoho ve štychu* i s opačným slovosledem (vzor 16) a zjistěte, kdo je v korpusu nejčastěji *necháván ve štychu* a v kterých osobách, časech a způsobech se tento frazém vyskytl. Týká se tento frazém vždy osoby? Jak byste vyjádřili význam tohoto frazému? Znaťe případně jiný, který by ho nahradil? Porovnejte svá pozorování se Slovníkem české frazeologie a idiomatiky.

■ **Úkol 15**
Podle vzoru 16 vyhleďte v korpusu kolokaci *hledat souvislost*. Smažte všechny opakující se struktury a ty struktury, u nichž jsou slova dané kolokace v různých větách a vyhledávkou uložte. V jakých pádech se objevuje slovo *souvislost*? Jaký je rozdíl mezi kontexty „Číňanka, kterou policie hledala v souvislosti s nedělní vraždou...“ a „hledat v těchto souvislostech argumenty“?

■ **Úkol 16**
Vytvořte si co nejlhodnější dotaz na vyhledání kombinace lemmat *zadat a práce* s vloženými slovy s různým slovosledem (vzor 16). Smažte si výskyty, kde tato dvě slova nejsou součástí jedné věty nebo nejsou podmětem a přísudkem téže věty.

Zjistěte, v jakém poměru jsou u těchto slov věty v činném a trpném rodě. Pokuste se vlastními slovy říct, v jakém poměru jsou pravděpodobně v češtině věty činné a trpné?

■ **Úkol 17**
Frazémy mívají svou základní podobu. Tu zjišťujeme např. na základě frekvence určitého pořadí prvků. Co byste například řekli o různých slovosledných variantách frazému *dát se do něčeho* a *do něčeho se dát*, pokud v druhém případě budeme uvažovat aktivitu/objekt vyjádřenou na daném místě žadným až třetí slovy (vzor 15): která slovosledná varianta převažuje?

■ **Úkol 18**
Nyní si všimneme spojení *hrozit v případě něčeho*. Je to spojení slovesa *hrozit* s víceslovnou předložkou (kolokací) *v případě*.

A. Vyhleďte si podle vzoru 12a posloupnost *hrozit* (lemma) v (lemma) *případě* (lc) tak, aby sloveso *hrozit* bylo v činném rodě. Podívejte se na výskyty, co za tvarem *případě* v korpusu následuje. Může jít o substantivum v 2. pádě, nebo může následovat vedlejší věta, tj. *v případě*, že. Kolokace v *případě*, že má funkci spojky. Zjistěte, jaký je poměr 2. pádu a vedlejších vět.

B. Potom zkombinujte vzory 14 a 12a. Vyhleďte posloupnost v (lemma) *případě* (lc), 0 až 3 vložená slova, *hrozit* (lemma, činný rod). Opět projděte výskyty a zjistěte, v kolika dokladech následuje za kolokací *v případě* vedlejší věta.

■ **Úkol 19**
V tomto úkolu si vyhledáme předložková spojení slovesa a substantiva *huba*. K zadání dotazu využijeme vzor 14 tak, že za sebou zadáme sloveso, předložku, 0–3 vložená slova a lemma *huba*. Jaké frazémy se substantivem *huba* jsou v korpusu?

■ **Úkol 20**
Vyhleďte v korpusu frazém *zlatý hřeb* (vzor 12) a vypište, co je nejčastěji *zlatým hřebem* a jakých akcí se *zlatý hřeb* týká. Je možné toto spojení užít i o jídle?

F. Syntax

ÚVOD

V oblasti syntaxe lze pomocí korpusu zadávat kombinace lemmat a gramatických kategorií, sledovat různé možnosti slovosledu, např. zda je častější přímý nebo inverzní (obrácený) slovosled, zkoumat syntaktickou funkci různých gramatických kategorií, např. ulohu infinitivu ve větě. Je rovněž možné všimnout si lexikálního naplnění určitých syntaktických struktur, např. tážacích vět o určitém počtu slov, sledovat návaznost jednočlenných typů vět v textu. Snadné je rovněž pomocí korpusu zjistit užívání interpunkce v okolí spojek, popř. užívání spojek na začátku a uprostřed věty.

Úkol 1

Podívejte se do korpusu, kdy se před spojkou *než* (vzor 1) píše čárka a kdy nikoli. Zjistěte si frekvence výskytů s čárkou a bez čárky (menu Konkordance – Statistika – Frekvenční distribuce) a potom se pokuste formulovat zákonnosti. Které slovní druhy jsou nejčastější před touto spojkou a které za ni?

Úkol 2

Některé české předložky častěji stojí uvnitř věty, jiné na začátku věty. Podle vzoru 2 si vyhledejte v korpusu předložku *oproti* a proveďte její pozici ve větě. Jaký je poměr předložky na začátku a uvnitř věty? Pomocí funkce Frekvenční distribuce (menu Konkordance – Statistika – Frekvenční distribuce *word*) zjistěte 5 nejčastějších tvarů, které následují za předložkou *oproti* a *oproti*.

Úkol 3

Podobně jako u předcházejícího úkolu se dá zjistit počet výskytů spojek *jestli* na začátku věty a v souvětí. Jak se liší význam této spojky na začátku a uvnitř souvětí?

Úkol 4

Některé vícelslovné předložky mohou mít současně funkci spojek (viz úkol E18). Týká se to například velmi frekventované předložky *vzhledem k něčemu*, která má ve funkci spojky podobu *vzhledem k tomu*, že. Vyhledejte si proto v korpusu podle vzoru 13 kombinaci lemmat *vzhledem* a že se 3 vloženými slovy a zjistě-

te, zda se místo ukazovacího zájmena mohou vyskytnout podstatná jména. Která podstatná jména jsou v této pozici dostatečně ustálená a mají už tak vyprázdněný význam, že se blíží zájmenům?

Úkol 5

Vyhledejte následující konstrukci: *co + substantivum, to + substantivum*. Pomocí frekvenční distribuce se pokuste zjistit, zda se některá slova nebo celé konstrukce opakují a které jsou nejčastější. Jak byste tuto konstrukci charakterizovali?

Úkol 6

Vyhledejte lemma slovesa *byť* následované lemmatem sloves smyslového vnitřního (např. *byť vidět, byť slyšet*) – obě v infinitivu, po kterých následuje substantivum. Která slovesa mohou plnit funkci určitého slovesa? Který vzor jste při vyhledávání použili?

Úkol 7

Po některých vícelslovných předložkách může v češtině následovat místo jména v konkrétním pádě sloveso v infinitivu. Můžete si to ověřit tím, že v korpusu vyhledáte předložku *s clem* a za ni infinitiv. Která slovesa se tu objevují? Která z těchto sloves jsou frekventovanější – s kladným nebo záporným významem?

Úkol 8

Vyhledejte postupně modální slovesa *moci, muset, mít* jako lemma v kombinaci s následujícím infinitivem. Pomocí Frekvenční distribuce určete nejčastější lemma na pozici infinitivu a sledujte, o jaká slovesa se jedná – vytvořte určitou významovou skupinu?

Úkol 9

Zkuste si vyhledat následující posloupnost: sloveso v infinitivu následované slovesem a opět slovesem v infinitivu (vzor 7). O jaké gramatické kategorie se u prostředního slovesa nejčastěji jedná?
A. Která slovesa stojí před touto posloupností? Pokud vám některé konstrukce připadají stylisticky nevhodné, pokuste se vyjádřit proč a přeformulujte je.
B. Který infinitiv na první pozici je součástí frazému a netvoří tudíž s následujícími dvěma slovesy smysluplné spojení?
C. Kdy po sobě následují tři infinitivy?

- **Úkol 10**
Řekne-li se uvozovací věta, každého jistě hned napadne příklad s časem minulým – typ *Karel pak řekl: „S vámi už nikdy nikam nepojedu.“* Takových příkladů najdeme v korpusu SYN2000 víc než 21 tisíc (21074). Kolik je tam ale takových příkladů s časem budoucím? Využijte vzor 20. Zkusíte to nejdříve odhadnout? Jaká je nejčastější slovesná osoba a jaké sloveso?

- **Úkol 11**
V následujícím cvičení budeme hledat výčty podstatných jmen (dvou a více) za dvoječkou. Zadejte dotaz podle vzoru 20. Zobrazení si nastavte na 10000 výsledků. Po vyhledání zadejte Konkordance – Jednoduché třídění – Levý kontext. Nyní nás bude zajímat, jakými slovy je tento výčet před dvoječkou uveden.
A. Zadejte si zdroje textů a zvažte, ve kterých se výčty objevují nejčastěji a kde vůbec ne.
B. Prohlédněte si část, ve které jsou slova a výrazy v uvozovkách. Jaké významy tyto uvozovky mají a jakou funkci plní výčet, následující po dvoječce?
C. Nyní uveďte 5 uvozovacích výrazů před dvoječkou. Programy kin, divadel a koncertů a sportovní výsledky přitom ponechte stranou.
D. Jak byste vyhledali výčet 5 substantiv oddělených čárkami po dvoječce?

- **Úkol 12**
Vyhledejte si v korpusu struktury s dvoječkou, kdy dvoječka odděluje text zakončený neživotným podstatným jménem mužského rodu v 1. a 4. pádu a větu začínající infinitivem, tj. substantivum + dvoječka + infinitiv (vzor 20).
A. Která substantiva mají frekvenci vyšší než 10?
B. Jaký okruh sloves je pro každé z nejfrekventovanějších substantiv charakteristický?
C. Které sloveso se opakuje u různých substantiv?

- **Úkol 13**
Vyhledejte v korpusu lemma *trojlístek* a podívejte se na jeho shodné a neshodné přívlastky (můžete k tomu využít funkci Jednoduché třídění (menu Konkordance – Jednoduché třídění, nejprve pravý a potom levý kontext), které vám výskyt vyseřadí abecedně.

- **Úkol 14**
Vyhledejte si tázací věty, které obsahují pouze tři slova (vzor 20, je však třeba vyloučit interpunkci). O jaké typy otázek jde? Převládají mezi nejfrekventovanějšími strukturami (do frekvence 15) otázky zjišťovací nebo doplňovací? Vímějte

si, jak se liší otázky v nadpisech a jak jsou otázky zapojovány do psaného textu. Sledujte, které slovní druhy se tu uplatňují.

- **Úkol 15**
Vyhledejte dvě tázací věty za sebou tak, aby druhá věta začínala zájmeny *kdo* nebo *co* (vzor 20 + vzor 17). Potom rozlišíte, zda otázky klade tenýž člověk (případně autor článku, monologu) nebo zda se jedná o dialog. Sledujte, jak se tyto druhy otázek liší. Jakých prvků užíváme, upřesňuje-li první otázka druhou?

- **Úkol 16**
Ve velkém procentu vět běžného jazyka se dozvídáme, kdy se slovesem popisovaná událost odehrála či odehraje, nebo kde. Vytvořte tabulku, která nám ukáže počty tázacích vět v korpusu SYN2000, které začínají těmito dvěma slovy: *Kdy – Kam*, a mají postupně jedno, dvě, tři, čtyři a pět dalších slov (vzor 20 + 14). Zaujalo vás na tabulce něco?

- **Úkol 17**
Zadejte si podle vzoru 20 vyhledání vět zvolacích vyjádřených podstatným jménem, např. *Výřah!* Podívejte se na rozložení těchto vět v korpusu (Konkordance – Statistiky – Rozložení).
Dále si zobrazte zdroje. V jakých typech textů se tento typ vět především vyskytuje? A proč?

- **Úkol 18**
Budeme hledat frázi *byl, je a bude* v různých obměnách podle vzoru 20a. Využijte funkci Obecné třídění (Vynechte duplicity)! S kolika slovesy jste našli napplněnou tuto strukturu v korpusu?

- **Úkol 19**
Jak byste vyhledali v korpusu označení věku, tj. konstrukce typu *je mu 15 let* a *má 15 let* jedním dotazem? Inspirujte se vzorem 15. Kolik výskytů je s *mít* a kolik s *byť*? Použijte k tomu funkci Frekvenční distribuce (menu Konkordance – Statistiky – Frekvenční distribuce) pouze pro pozici 0 (lemma).
Myslete, že je územně nějaký rozdíl v užití *mít* a *byť*?

■ Úkol 20

Vyhledejte sloveso *byl* ve všech tvarech následované tvarem *večer*. Rozhodněte, kdy je tvar *večer* adverbium a kdy substantivem. Srovnejte věty typu *bylo večer* a *byl večer*.

G. Kombinovaná zadání

ÚVOD

V této části jsou soustředěny úkoly, které ukazují další možnosti využití korpusu – jednak vyhledávání interpunkce a její funkce v textu, jednak vyhledávání ne-spisovných jazykových prostředků, např. adjektiv zakončených na *-ej*, jednak zjišťování nejazykových informací na základě korpusu. Specifickou tohoto oddílu je však prověřování kombinatorických možností češtiny, zvláště z hlediska kumulace stejných slovních druhů nebo gramatických kategorií, např. několikrát *čas-tic* – *bohužel asi ne, no jen ať apod.*

■ Úkol 1

Jakým způsobem v korpusu vyhledáte nejdéší slovní tvar? Kolik písmen tento tvar má? O jaká jde slova – obecně užívaná nebo něčím specifická?

■ Úkol 2

Vyhledejte si v korpusu středník (vzor 19). Podívejte se na rozložení užití středníku podle typů textů (menu Konkordance – Statistika – Rozložení). Vzhledem k tomu, že víme, že korpus je sestaven z textů v pořadí – beletrie, odborná literatura, publicistika, v kterých textech převládá užití středníku?

■ Úkol 3

Vyhledejte v korpusu kombinaci kulaté otevřené závorky, otazníku a zavřené kulaté závorky (vzor 19a) a všimněte si, jakou má funkci v textu. Co se takto užitým otazníkem v závorce zdůrazňuje? Podle statistické funkce Kolokace (menu Konkordance – Statistika – Kolokace) zjistěte všechna přídavná jména v kontextu do 3. pozice zleva s frekvencí 3. Jaký typ adjektiv převládá?

■ Úkol 4

Vyhledejte v korpusu zvolací věty, které obsahují právě jedno slovo (snov. úkol F17). Zjistěte frekvence jednotlivých slovních druhů. Seřadte je podle četnosti výskytu.

Úkol 5

V jazyce je také důležité zjistit, jaké platí zákonitosti, následují-li za sebou slova stejného slovního druhu. Vyhledejte slovesný tvar v minulém čase a za čárkou tři spojky (vzor 20).

Zjistěte, kterých sloves se tato vyhledaná struktura týká. Dokázali byste popsat, o jakou významovou skupinu jde?

Jaké kombinace spojek jsou nejčastější? Která ze zjištěných kombinací zahrnuje více slovných spojkových výrazů?

Úkol 6

Vyhledejte v korpusu 5x za sebou 2. pád podstatného jména mužského rodu (vzor 14b). Jaké struktury jste takto získali? Pokud se nějak nazvat vztahy uvnitř této struktury. Jaké typy lemmat se vyskytují na jednotlivých pozicích?

Úkol 7

V následujícím úkolu se zaměříme na délku postupně rozvíjeného shodného přívlastku.

Jaký je maximální počet přídavných jmen v jednom takovém řetězci?

Jaké druhy adjektiv po sobě následují?

Úkol 8

Vyhledejte v korpusu 3 po sobě jdoucí adverbia (vzor 14b). Jedno z těchto adverbíí musí být přitom nespisovné (označeno 6 na pozici 15). Sledujte typy, které po sobě následují. Pokud se udělat seznam těchto typů. Vynechte případy opakujících se stejných slov.

Úkol 9

Vyzkoušejte si, kolik slov lze vložit mezi části schématu *jednak – jednak*. Můžete k tomu využít vzor 14a. Zjistěte frekvence jednotlivých případů a porovnejte si vložené konstrukce a odhadněte nejužší možný rozsah.

Úkol 10

Zkuste si ověřit, která kombinace dvou částic v mluveném jazyce skutečně převládá a srovnejte *zas tak* a *novější až tak* (vzor 4). Okomentujte rozdíly v užití obou výrazů.

Úkol 11

Abstraktní substantivum *rámec* je součástí více slovné předložky *v rámci*. Avšak i do této kolokace mohou být vkládána další slova. Vyhledejte si proto tuto kolokaci v korpusu podle vzoru 13, tj. nejprve si jedním vloženým slovem. Na základě frekvenční distribuce získáte seznam s různými vloženými slovy.

Vypište adjektiva, která mají opozitní, tj. protikladný význam.

Kolik slov lze do této kolokace vložit?

Úkol 12

Vyhledejte lemmata sloves *chít* a *být* v tomto pořadí, následovaná slovesem *vědět* v infinitivu, které má za sebou čárku a spojku. Určete druh vedlejší věty a nejfrekventovanější spojku. Přehodte pořadí sloves *být* a *chít* a vyhledávkou opakujte – sledujte slovosled a druh slovesných tvarů v obou vyhledávkách a porovnejte je.

Úkol 13

Když se učíme cizí jazyky, nikterak nás neudiví, že bychom se měli u každého slovesa učit současně i předložku, která po něm následuje (tj. část jeho valence), a současně s tímto předložkovým spojením i význam. Dokážeme takto bedlivě sledovat i naši mateřštinu? Najdete u slovesa *podívat se* spojení s předložkou, které se posunulo ve významu?

Úkol 14

Zkuste si vyhledat v Pražském mluveném korpusu (není gramaticky označováno) pasivní tvary slovesa v mužském rodu. Pomocí frekvenční distribuce zjistíte nejčastěji užívaný pasivní tvar. Ten vyhledejte i pro ostatní rody a porovnejte spisovné a nespisovné tvary. Nechte si zobrazit kódy mluvčích a zkuste posoudit, zda užívání těchto pasivních tvarů závisí na pohlaví, vzdělání či věku mluvčích či typu promluvy. Porovnejte výskyt těchto tvarů s Brněnským mluveným korpusem.

Úkol 15

Sloveso *lézt* používáme v mluveném jazyce v řadě frazémů. Zkuste si vyhledat sloveso *lézt* a následující substantivum, zájmeno nebo číslovku s 1 vloženým slovem.

Úkol 16

Jako lemma zadejte v korpusu předložku *kolem* a na místo následujícího tvaru specifikujte podstatné jméno ženského rodu zakončené na *-řka* v gen. sg. Která dvě podstatná jména jsou nejčastější? Proč? Která podstatná jména neoznačují věk? Jaký mají význam?

2. Co říká o různých slovech korpus a co slovníky

■ **Úkol 17**
Vyhledejte tvary lemmatu *být* začínající na *j* následované nespisovným adjektivem (15. pozice) podle vzoru 7 a pokuste se v kontextech vysvětlit tuto disporci zápisu obecné češtiny – sledujte, ve kterých útvarech paného jazyka se vyskytují obecněčešské tvary.

■ **Úkol 18**
V obecné češtině se setkáváme s tvary, v nichž za sebou následuje oslovení *pane* a další substantivum v nominativu. Které názvy povolání se tu objevují?

■ **Úkol 19**
Zjistěte rozdíl v užití slov *ostřelovač* a *odstřelovač*. Sledujte jednak příklady v korpusu a jednak příslušná hesla ve slovnících.

■ **Úkol 20**
V korpusu je možné vyhledávat také encyklopedické informace. Zadejte si v korpusu vyhledat kolokaci *Zlatá Praha* (vzor 4a v kombinaci se vzorem 2) v různých pádech. Když si zobrazíte zdroje, zjistíte, že výskyty 1–84 jsou z knih Bohumila Hrabala. Ponechte si ve vyhledávce pouze tyto doklady (Konkordance – Redukce – Prvních 84 řádků).
Co označuje spojení *Zlatá Praha* u Bohumila Hrabala?

ÚVOD

Jste si jisti významem i méně běžných slov, nebo o něm někdy máte jen neurčitou představu a odhadujete ho ze souvislosti? Projděte následující méně běžné výrazy a ze čtyř významů uvedených u každého z nich zkuste vybrat ten pravý. V korpusu si prohlédněte výskyty každého z těchto slov (ve všech jeho tvarech) a zjistěte, zda význam, který jste vybrali, je v souladu s tím, jak se příslušného slova užívá. Pokud se vybraný význam do zjištěných kontextů nehodí, vyberte z dané čtveřice význam jiný, vhodnější. Své řešení si nakonec ověřte v některém slovníku češtiny (nejlépe ve vícedílném *Slovníku spisovného jazyka českého*, popřípadě v jednodílném *Slovníku spisovné češtiny*), popř. navhnete doplnění nebo opravu. V jakých frekvenčních pásech se jednotlivá slova vyskytují – např. 50–100 výskytů, 300–500 výskytů, 700–1000 výskytů apod.

bigotní A. přehnaně zbožný; netolerantně věřící; B. dvoustranný; oboustranný; C. unavený; vyčerpaný; D. dvojí; dvojitý

bonmot A. základní směrnice, pravidlo, ustanovení; B. dobré chování, dobré způsoby; C. vtipné rčení, průpovídka; D. přílišná okázalost, nadutost

debutovat A. poprvé veřejně vystoupit; B. vyvolat nepřiznivou odezvu u kritiky; C. zanechat umělecké činnosti; D. vytvořit své životní dílo

dedikace A. odstoupení, zřeknutí se funkce; B. odevzdání, smířenost s osudem; C. odchylka, úchylka, odchyllost; D. věnování, připsání

deportace A. vysídlení, nucené vystěhování; B. vylodění; C. výrazné zhoření, úpadek; D. zbavení poslance immunity

depozitář A. odborník pečující o muzejní sbírky; B. soupis peněžních prostředků a cenných papírů uložených u peněžního ústavu; C. bankovní úředník specializovaný na problematiku vkladů; D. veřejnosti nepřístupná prostora k ukládání knih nebo sbírkových předmětů

doktrína A. soubor zásad, základních myšlenek; B. mezinárodní smlouva; C. hlavní obor působnosti; D. oficiální stanovisko vlády k určitému problému

drahně A. zvolna, pomalu, váhavě; B. hodně, velmi; C. prudce, divoce; D. marně, beznadějně

epigon A. napodobitel významné osobnosti; B. významná vůdčí osobnost; C. architektonicky upravený vchod do významné budovy; D. náhrobní deska s nápisem o zesnulém

euforie A. libozvučnost, melodičnost; B. odmítnutí, zavrhování; C. skličenosť, sklisťosť; D. uvolnění, povznesenosť, rozjařenosť

evokovat A. draždit, podněcovat, popuzovat; B. vybavovat, vyvolávat; C. maskovat, zastírat, předstírat; D. věštit, předpovídat

ex post A. chráněný listovním tajemstvím; B. na cestě mezi odeslatelům a adresátem, „v poště“; C. beznadějně, bez vyhlídky na úspěch; D. po skončení, dodatečně

exaltovaně A. zaníceně, vzrušeně; B. tiše, nesměle; C. doslovně, přesně, věrně; D. obrazně, nepřímě, symbolicky

frapanantí A. vřídntý, přívětivý, povolný; B. upjatý, chladný, odměřený; C. prostofeký, drzý, odmlouvavý; D. nápadný, bijící do očí

frenetický A. srdečný; B. bouřlivý; C. vynikající, mimořádně nadaný; D. bezelstný, upřímný, nepředstíraný

charizma A. rozšíření církve, církevní rozkol; B. osobní kouzlo; C. dobročinnost, lidumilnost; D. oficiální veřejné prohlášení

image A. záměrně pěstovaná a šířená představa; B. znak firmy ve speciální výtvárné úpravě; C. výmysl, nepravdivé tvrzení; D. osobní hrdosť, osobní česť

impresárió A. soubor obrazů se společným námětem; B. dirigent; C. divadelní nebo koncertní organizátor a zprostředkovatel; D. umělecký tiskař

interakce A. scelení, sjednocení, sjednocování; B. úřední zákrok, zásah; C. umělé přerušení těhotenství; D. vzájemné působení

intoxikace A. podávání léku; B. otrava; C. nevolnost; D. závrat

invektiva A. orgán pověřený výkonou mocí; B. policejní nebo vojenský zásah, zákrok; C. slovní nebo písemný útok; D. vynalézavosť, dvůřtíp

kinelóza A. nevolnost způsobená vyčerpáním; B. nevolnost způsobená pohybem; C. cukr užívaný k rychlé výživě; D. druh sirupu

koncúzně A. stručně, zhuštěně; B. současně, souběžně, paralelně; C. hrubě, sprostě, vulgárně; D. rozvážně, klidně, pokojně

konfúzní A. společný; B. zmatený; C. překvapivý, neočekávaný, nenaďalý; D. odporný, odpudivý, nechutný

konzistentní A. tvořivý, tvůrčí; B. libozvučný, melodický; C. soudržný, celistvý; D. rozhodný, energický, důrazný

koopťovat A. přibrat dalšího člena; B. odvolat, sesadit, zbavit funkce; C. spolupracovat, společně postupovat; D. popřít, zapřít, zatájit

kosmologie A. názor, že vesmír je konečný; B. názor, že vesmír je nekonečný; C. vědecký výzkum vesmíru; D. nauka o vesmíru jako celku

krucíální A. závažný, vážný; B. vzdálený, odlehlý, zapadlý; C. oslavovaný, vyzvedaný, vychalovaný; D. protipřávní, protizákonný

maně A. mimovolně, mimoděk, bezděčně; B. samolibě, ješťně, domyšlivě; C. bez zábran, bez pochyb, bez rozpaků; D. nadarmo, zbytečně, bezvýsledně

mentolovat A. dát něčemu název, pojmenovat něco; B. neustále poučovat, kárat, napomínat; C. nezajímavě, nezázivně vykládat; D. podvádět, mystifikovat

merítum A. podstata, jádro; B. skrytý problém, háček; C. výhoda, kladná stránka; D. celkový ráz, celková povaha

nabob A. otok, opuchlina; B. lehký, kypřý moučník; C. bohač, bohatec, zbohatlík; D. pomlouvač, nactiutřač

noticka A. vřřah z notového záznamu; B. kratší oslavná báseň; C. kratší oslavná nebo jásavá skladba; D. drobná zpráva, poznámka v novinách

nulítní A. bezvýznamný, bezcenný; B. nicotný, planý, bezobsažný; C. zaoblený, plnošňhlý; D. naprosto neplatný

parafrazovat A. rozdělit na několik stejných úseků; B. volně citovat nebo obměnit cizí myšlenku; C. ochromit, učinit neškodným; D. provádět současně, souběžně

parlamentář A. člen předsednictva parlamentu; B. předseda poslanecké sněmovny; C. osoba zmocněná k vyjednávání, vyjednavač; D. zástupce senátu na zasedáních parlamentu

pasus A. úryvek, úsek, pasáž; B. společenská chyba, nedopatření; C. zastaralý, překonaný názor; D. odmítnutí, zamítnutí

pitoreskní A. líný, zahálčivý; B. podivný, divný, směšný; C. žalostný, truchlivý; D. úzce specializovaný a postrádající širší rozhled

plytký A. povrchní, mělký; bezobsažný; B. zvolna plynoucí, pomalu tekoucí; C. ostrý, nabroušený, pronikavý; D. nesrozumitelný, nejasný

polyhistor A. odborník na dějiny novověku; B. odborník na dějiny zemí jedné zeměpisné oblasti; C. znalec mnoha vědních oborů; D. vypravěč příběhů, spisovatel, básník

pragmatický A. staromilský, zpátečnický; B. lhostejný, netečný, chladnokrevný; C. jdoucí přímo a bezohledně za vlastním cílem; D. založený na užitečnosti, účelnosti

preambule A. obnovení, opětné zavedení; B. neověřená zpráva; C. úvodní část významného dokumentu; D. předpověď, odhad do budoucna

pregnantní A. menší, ale silný, svalnatý; B. hutný a výstižný; C. předcházející, dřívější; D. výnosný, výhodný

premis A. východzí tvrzení, základ pro další vyzovování; B. vyzovený závěr, úsudek; C. předběžný souhlas, neoficiální svolení; D. předběžný dokument, podklad pro další jednání

profánace A. zhuštěná podstata, jádro díla; B. orientace na obchodní úspěšnost díla (na úkor jeho uměleckých kvalit); C. znesvěcení, zneuctění; D. první veřejné slavnosti uvedení díla

provenience A. přepjatost, přemrštěnost; B. umělecká působivost; C. původ; D. směřování různých myšlenkových nebo uměleckých směrů

puč A. vláda menšiny; B. násilný politický převrat; C. vyjednávání s představiteli vlády; D. vojenská vláda

půlka A. tradiční výroční slavnost se zábavou; B. putování na posvátné místo; C. polýčka, bitka; D. pokosená louka

remniscence A. vzpomínka, vzpomínání; B. hodnotící posudek, kritika; C. obrat v ději, překvapivý zvrat, zlom; D. opětovné uvedení na scénu

rudimentární A. opakovaný, nepřetžitý; B. základní, elementární; C. užitečný, prospěšný, výhodný; D. záměrný, umyšlený, cílevědomý

signifikantní A. stvrzený podpisem, oficiálně podepsaný; B. označený značkou, štítkem ap.; C. zjednodušený; D. významný, podstatný

snoval A. dřímat, pospávat; B. zamotat, zavázat, svázat; C. snít, toužit, toužebně si přát; D. příst, spřádat

suspektní A. neprozřetelný, nerozvážený, neuvážený; B. podezřelý, vyvolávající podezření; C. nadřazený, důležitější, nadřazený; D. dočasně zbavený úřadu nebo hodnosti

talár A. dlouhý tmavý obrádní šat; B. poplatek za studium na vysoké škole; C. slib skládaný při uvedení do veřejné funkce; D. sbírka citátů z děl klasických autorů

tiráž A. technické a vydavatelské údaje uvedené na konci knihy; B. barva, zabarvení tónu nebo hlasu; C. dílo vytvořené z výstižků nebo zlomků jiných děl; D. odvaha, smělost, nebojácnost

traktovat A. přitahovat, lákat; B. odpuzovat, odrážet; C. vykládat, pojímat určitým způsobem; D. klamat, obelhávat

travesie A. zvýšená tvůrčí aktivita, tvůrčí nadšení; B. komické nebo zlehčující zpracování vážného tématu; C. uvedení dramatického díla na scénu; D. hudební skladba s náboženským námětem

triplych A. centrální část oltáře; B. trojdlíhý umělecký výtvor; C. plastika zpodobňující trojspreži s bohyní vítězství; D. skladba pro tři nástroje

tristní A. jednoduchý, prostý, obyčejný; B. smutný, sklíčující; C. prudký, vášnivý, náruživý; D. trojnásobný, ztrojený

trzn A. středověký mučičí nástroj; B. zbraň podobná trojzubci; C. slavnost na počest zemřelého; D. trápení, soužení

uzurpování A. bezprávný ovládnutí, uchvácení; B. upadání, úpadek, rozklad;
C. přesvědčování, přemlouvání, pobízení; D. nucení, donucování, vynucování
telný; D. snadno rozpoznatelný

voluntaristicky A. ukvapeně, nerozvázně, nepředložené; B. podle vlastní vůle,
svévolně; C. lákavě, podmanivě; D. povyšně, pohrdavě, přezíravě

xenofobie A. nesnášenlivost, hašteřivost, hádavost; B. strach z neznámého; C.
odpor, nepřátelství k cizincům; D. obavy z budoucnosti

zhatit A. zkomolit, překroutit, změnit k nepoznání; B. zdolat, překonat, přemoci;
C. zmařit, zkažit, překazit; D. zmašť, oklamat, uvést v omyl

3. Význam slova prozrazuje kontext

ÚVOD

V následujících cvičeních bude úkolem odhalit slovo nebo slovní spojení, které bylo nalezeno v korpusu a je utajeno pomocí teček.

A. Slovní tvary

■ Úkol 1

„Nechci.“ Wolfe upíel <.....> nalevo. “Pane Quester? Kapitola když uviděl krutěj <.....> už obvázanýho a uzdravujícího se Micky velká kožená sedačka, na první <.....> starožinnost. V proskleném rohu jeho pozornost upoutal letný <.....> na punčochu ve spodní příhrádce. Shora se otevřel neobvyklý <.....> na velký chrám a bylo odtud možno namalovaných nebeských sfér, otevírá se <.....> do nebeských výšin. Předkládá v nich svůj <.....> na reklamu jako na informační prostředek. Nabídneme nyní alternativní <.....> na povahu současné české inflace jež podle režiséra umožňují nový <.....> na interpretaci jeho díla. “Dříve však podle nich má být mnohem hezčí na <.....>, a hlavně má

■ Úkol 2

než já správním dveře, tak mi vypadne <.....>. Vy to máte dobrý, vy to nevydržel, zatáh ruční brzdu, já vyklop <.....>, jakob hodil žebřík “Držte se, ať mi nevyrazíte čelem <.....>!” A smka byla větší a ještě větší a ráno když se probudil, vteřel <.....> a šípku vyskočil na dláždění jeden souseď v zimě vyškřábal zamrzlý <.....> a kouká se do zahrady povídala mi: “Hele, Karlíku, volevěřme <.....> vezmeme se za ruce diskretně ty růže Libuše na <.....>, tak jak to dělají Mexikáni a Španělé mluvit při měsíci. Dejte si nohy na <.....> a sedte vedle mne. To je whisky, kdykoli sem chlel umejt <.....> dopadlo to bledě ale ne pádem nemůžu jsme se tak ožrali, že máme dodnes <.....> a vůbec nevíme, jak jsme se

B. Lemmata

■ Úkol 3
Zážitek a ani tehdejší večer nebyl <.....>. Vyrovnané řady pánů ve fraku, o tyto bohatí běloši, kterých byly (s <.....> Portugalců) v zemi necelé dva tisíce negativní závislost mezi těmito jevy. <.....> tvoří antisocialní chování na byl podle předpokladu – až na několik <.....> realizováno do konce letošního organických látek, což je spíš pravidlo než <.....>. Aní v České republice tyto „duální hygienik mohl sice udělit lelectvu <.....>, ale odmítl to. Kalousek míní, že slovensčina, až na malé <.....>, 5. pád nemá. Někdy ani čeští intelektuálové party, dlužno přiznat, že až na <.....> to byla vždy záležitost klukovská. trpícího pacienta na život musí bez <.....> cítit soucitný lékař (a to i proti

■ Úkol 4
Namáhali jsme se vyprostit dědečka z té <.....> pozice a každý z nás se o to je to takhle. Po dlouhých a <.....> časech je to hodné, to slušně Československo žilní trombózy. Diagnóza je velmi <.....> u starých osob, interval 12 h je velmi bombardováním či boji na konci války. Ve <.....> postavení se ocitla Lékařská se šíří i mezi zaměstnanci. Zlepšit <.....> situaci má maskarní ples v prostorách ministr zahraničí však nemá zájem toto <.....> téma natukovat, „ řekl v Praze Na otázku, zda to souvisí s jeho <.....> finanční situaci, Mach odpověděl: Švýcarska za druhé světové války. Beztak <.....> hledání sponzorů zkomplikoval potřeby na mezinárodní instituce. <.....> otázku čečenského samostatnosti může malometrážních bytů, které vyřeší <.....> bytovou situaci některých z našich

C. Kolokace

■ Úkol 5
dilema se zájezdem v Paříži), byla <.....> výsledků anketý otevřena otázka, zda ekosystému apod., neboť se <.....> vyhodnocení těchto informací vyhnout jsme o problému Boha uvažovali <.....> dvou údajů, s kterými se každý člověk se obětí vybíraly namátkou či <.....> anonymních udání. V duchu tradice představilo na tiskové konferenci, uvádí <.....> použití obdobné metody jako hranici dodávky ropy máme s RF uzavřenu smlouvu, <.....> které budou vždy září příslušného ČRo) Vlastimil Ježek. Učinil tak <.....> doporučení komise pro nabídek na prezidentskou kancelář „říká <.....> svých zkušeností Albert Amerického elektromobilu“. Projekt vznikl <.....> resitřikativního kalifornského „Rozhodli jsme se proto i <.....> jejich připomínek, že polovinu z tří set korun

■ Úkol 6
Hra vychází z pohádky Boženy Němcové <.....>. Účastníci po rozdělení nejlépe Zadáni proto vypadalo jako úkol pro <.....>. Dát zemím, směřujícím k demokracii a nejlépe o já. Přijíždí po způsobu <.....> – ani na koni, ani pěšky, oblečen konzervativců tak trochu připomnělo příslovčnou <.....>. Ministři se žádých nemožná zároveň. Podle tohoto receptu <.....> také bude vypadat další politika tedy přesně ve stylu pověstné <.....> Boženy Němcové. Pareo lze vázat nejen Tam již nepomůže chování <.....>, se kterým zájm Lionel Jospin úspěšně a některé z nich ještě méně. Ale ve stylu <.....> zase není „až tak moc“ jehož rozlousknutí by se měla ujmout <.....> pokojové bonsaje sice potřebují vybrte. Dilema můžete vyřešit jako <.....> tím, že zvolíte čaj – nečaj, tedy

D. Části slov

Úkol 7

pravou nohu a ukazoval soudci zánovní <polo.....>. Svědek četník začal ztrácet srolované mapy a vtrhly se vzorky hornin a <polo.....>. Ano, panu Dělounkovi Shatterhande.: „To nic není,“ řekl <polo.....> Old Shatterhand a snažil se před svým domácím je výhodný, neboť získáme levnou <polo.....> (upravená ryba vydří přivedla tři mladé muže v džínách a <polo.....>, které doprovázela velmi korektně dráhy planet za soustředné kruhy, <polo.....> dráhy každé planety je příbližně pronikají do hloubi všech pouští či <polo.....>, a mezi nimi i do Mojunkumů, něčemu, co udělal, i když si teď v <polo.....> nemohl přesně vzpomenout, co to bylo pohovce mžourala do podvečerního <polo.....> a vážně přemýšlela o tom, jestli Mrázek byl plyn nevábně vyhlížejících <polo.....>. Rozhodla se pro celet obalovaný

Úkol 8

„Stalin.“ Je to tak. Stalinská <.....vláda> byla ta nejvýznamnější, největší a arabském, respektive islámském světě. <.....vláda> trvá v Maroku, Jordánsku zachránila židovský národ trpící pod <.....vládou> Babylónů. Lawrencův román když pak byla v Německu nastolena <.....vláda> v podobě diktatury ústředního pomine jako špatné zboží. Jenom cáři, <.....vláda> a takové slovy potrvají věčně žádné prozrazující gesto; její <.....vláda> soulad mezi slovy a duchem se zdály normalizace se rozšířila fáma, že <.....vláda> připravuje ménovou reformu. Jiřímu německými a ruskými snahami o <.....vládu>. Hlavní ideou ČSR bylo demokratické

E. Formálně podobná slova

ÚVOD

Některá přejatá slova mohou být formálně podobná a v důsledku toho je některi mluvčí zaměňují. V následujících cvičeních najdete několik takových dvojic. Pokuste se dosadit do kontextu z korpusu správné slovo z dané dvojice.

Úkol 1

efektivní/efektivní

efektivní [-ty- i -ty-]

účinný

často z ekonomického, politického hlediska

efektivní

působivý

vnějším dojmem, často pomocí vizuálního vnímání

Doplňte:

1. O rok později byla vyrobená limitovaná série pětaticeti kusů. Každý v <.....> kombinaci červené, černé a zlaté (vnitřní strana víka je vykládána skutečně 24karátovým zlatem).
2. Podle premiéra Václava Klause vláda většinově dospěla k názoru, že finanční odluka politických stran a hnutí od státu by přispěla k <.....> fungování celého politického systému.
3. Moc pěkně! Hlavně realistické! Hlavně realistické! Já sám mám ještě <.....> metodu, při níž by nám ovšem kaskadéři brzy došli.
4. Veškeré snahy o <.....> investice do vzdělání se pak rozplynou v rovnostářském zvyšování učitelových platů, které se nadále nebudou řídit podle odvedené práce, ale podle odsloužených let.
5. Po třech letech od nástupu vlády premiéra Vladimíra Mečiaru se dostala slovenská opozice poprvé k <.....> kontrole tajných služeb, když včera byli do parlamentních výborů pro kontrolu civilní Slovenské informační služby (SIS) a Vojenského zpravodajství (VZ) zvoleni 3 zástupci opozice František Javorský, Milan Kňažko a Roman Vavřík.

6. Kromě pohodové hudby a <.....> podívané na třiřyt vodních kapek, umocněný barevnými kreacemi obsahuje pořad, který začne v deset hodin večer, také tři vystoupení jednoho z našich nejlepších párů v latinsko-amerických tancích.

7. Na druhé straně sjednocující se Evropa má již vybudovány dnes <.....> kontrolní mechanismy, které dokáží včas zabránit jakýmkoliv pokusům o případné zneužití ekonomické síly některým členským státem.

8. Otázkou však je, zda Beckerovo zobecnění na úrovni neměřitelných a neodhadnutelných veličin je ještě něčím víc než jen ukázkou brilantní, ekonomicky a matematicky elegantní analýzy, která pouze intelektuálně vzrušuje, ale praktické odpovědi nedává a stává se tak jen velmi <.....> exhibicí.

9. Pokud jí ovšem nejde jen o další <.....> gesto, po kterém by měla – jak již avizovali lidři KDH i DU – následovat na podzim další petiční akce.

10. Východy systému „single mode“ – <.....> využít energie mikrovlákn, velmi rychlé rozklady – možnost použití oceli pro konstrukci tlakových patron

Úkol 2

abstinence/abstinence/absence

abstinence *

zdířování se hlasování

Těměř jednomyslně (při jedné <abstencis>) banka odsouhlasila překlenovací úvěr

abstinence **

zdířování se požiků, které při nadužívání mohou působit negativně

příznivé zprávy o Janině chování a drogové <abstincenci> v krizovém centru MOST

absence ***

nepřítomnost osoby nebo neživého objektu, činnosti, vlastnosti...

do školy se doma nepřipravují. Mají mnoho <absencí> a za školou často páchají trestnou činnost

* 1–100 výskytů

** 100 – 1000 výskytů

*** více než 1000 výskytů

Doplňte:

1. Při jedné <.....> a třech neplatných hlasech bylo pro Gonzáleze 181 poslanců, proti 165.
2. Naprostá <.....> od alkoholu je žádoucí v těhotenství, u dětí a dospívajících.
3. Za neomluvené <.....> lze zaměstnanci kráti dovolenou. 4. Strídmost nabádá k umírněnosti v práci i ve spánku, jídle, pití i emocích, střídmost, která zna-

mená naprostou <.....> od všech narkotik i dráždivých a povzbuzujících prostředků, taková střídmost je jedním ze zákonů zachování zdraví.

5. Prvotním cílem této léčby není <.....> narkomana, ale snížení rizik, která vyplývají z užívání drog (kriminalita, přenos viru HIV při injekční aplikaci drogy apod.).

6. Prudký pád nastal ve čtrnácti letech a mladík od té doby střídal učiliště a školy, kvůli častým <.....> a špatnému prospěchu však dlouho nikde nepobyl.

7. Celková investice do nikotinu <.....> tedy v těchto případech celkem dosáhne asi pěti tisíc korun.

8. Bud znechucením, ztrátou zájmu a stažením do soukromí, což se projevuje rostoucí <.....> ve volbách, anebo příklonem k nově rašícím silám, zprvu marginálním, ale pravdělně vzestupujícím.

9. Dvaceti vybraným narkomanům, u nichž všechny předchozí pokusy o <.....> selhaly, podávají lékaři drogu zvanou metadon, jež je náhražkou za heroin a braun.

10. Například Doris užívala antikoncepční pilulky, které jsou nejučinnější formou antikoncepce (kromě sexuální <.....>).

11. Cílem je úplná <.....> od hráčství.

12. Člověk, který si je vědom toho, že jediné, čím stojí nad zvířaty, pochází z jeho mozku, z jeho schopnosti myslet, tvořit, hloubat a posuzovat, takový člověk nachází potěšení v úplné <.....> od alkoholu, kávy, čaje, nikotinu, koly, čokolády i všech ostatních drog.

13. Každá <.....> ve volbách je hlasem pro komunisty.

14. Namiřali, že takový přístup k drogové problematice dáva mládeži špatný signál, a požadovali, aby léčebné instituce nabízely závislým lidem pouze cestu úplné <.....>.

15. Neznámý muž využil policejní <.....> a přerušil poslance.

Úkol 3

representační/representativní

representativní 1. schopný reprezentace, vhodný k reprezentaci; nádherně, skvěle, okázale reprezentující něco nebo někoho: reprezentivní manželka;

2. reprezentující (přiměřeně) říz. složky: reprezentativní sociologický průzkum

representační 1. prezentující stát, organizaci, nějaký celek: reprezentační hokejové mužstvo; reprezentační ples

2. v souvislosti s reprezentací: reprezentační gól

Doplňte:

1. Veber tím přišel o první <.....> akci sezony, turnaj Pragobanka Cup, druhý, Karjala Cup ve Finsku na začátku listopadu také zřejmě nestihne.

- Lička však v roli <.....> asistenta odjel v týdnu na kvalifikační zápas na Maltu a přípravu mužstva musel světlému asistentovi Robertu Pánkovi.
- <.....> barvy: červené dresy a bílé trenýčky, případně celí v bílém.
- Koncem 19. stol. zatoužilo rozvíjející se město Ústí n. L. po vlastní, <.....> divadelní budově, která by ztělesňovala jeho rozmach nejen jako významného průmyslového a těžebního, ale i kulturního centra.
- Už ve třetí minutě se veterán Svatoň rozvzpomněl na svou pověst obávaného kanonýra a na jeho přesné mířnou bombu byl i <.....> brankář Jelen Krátký.
- Hráč AC Milan Daniele Massaro, který oblékal <.....> dres Itálie na loňském mistrovství světa, podepsal přestup s japonským klubem Šimizu, kde by měl působit dva roky.
- Bylo jí třadvacet, když hodila silbnou <.....> kariéru za hlavu.
- ...dostane peníze na <.....> auto, které bude odpovídat vkusu jeho ženy;
- Expozice, kterou hostí ve svých <.....> prostorách v blízkosti Eiffelovy věže americká nadace Mona Bismarck Foundation, představuje na dvě stovky ukázek Muchovy bohaté a různorodé umělecké tvorby.
- Snažíme se každému autoru vystavit <.....> kolekci zhruba šesti prací," poznamenala Tunklová.

Úkol 4

adaptoval/adaptovat

adaptovat ned. i dok. (z lat.) (co) přizpůsobovat, přizpůsobit, upravovat, upravit: a. starý dům; a. povídku; a. slovo

adoptovat dok. (z lat.) (koho) přijmout za vlastní dítě; osvojit: a. sirotka; přen. kniž. a. cizí názor přijmout

- Mezi ním a Victorii žádné pokrevní svazky nebyly. Nikdy ji <.....>, ani o tom neuvažoval.
- Musí své životní cykly <.....> na široký rozsah fyzikálních a biolo- gických podmínek.
- Na druhé straně se český uhelný průmysl začal <.....> na tržní pod- mínky: kupť byla eliminována cenová regulace (všechny druhy uhlí jsou na čes- kém trhu už liberalizovány), vzniklo 5 silných uhelných akciových společností a v širokém měřítku započala rovněž revitalizace regionů s někdejší těžbou.
- Novela totiž prodlužuje dobu pobytu dětí v kojeneckých ústavech a tvrdí, že dítě může být <.....> až poté, co o něj jeho skuteční rodiče půl roku neprojeví zájem.
- Lidé z Německa přicházejí nesrovnatelně lépe připraveni a především už <.....> na naše podmínky.
- Hodně lidí <.....> dítě kvůli sobě a neuvědomí si, že pro něj musejí hodně udělat, protože je psychicky nevyrovnané a má spoustu problémů.

- První dům sociální prevence v ČR, v němž se budou občané propuštění z vý- konu trestu <.....> na život na svobodě, zamýšlí zřítit brněnský magis- trát.
- Ringier, firma se švýcarským zázemím a zkvallitěným managementem, má ja- ko vlastník časopisu TÝDEN dobré předpoklady k tomu, aby jej <.....> na nové podmínky.
- Architekt Jaroslav Polívka jej v roce 1927 <.....> pro obchodní účely, další úpravy provedli Otto a Karel Kohnovi v roce 1930.
- Pflieger uvádí, že osiřelá kamzíčí mláďata jsou odsouzena k smrti. Žádná ci- zí samice je <.....>.

Úkol 5

mentoroval/memorovat

SSiC

mentorovati ned. 1. expr. (koho, 4. p.) (při každé příležitosti) poučovat, kárat, mistrovat, napomínat: výchova bez zbytečného mentorování; kritika mentorující mladé umělce

Příklad z ČNK:

matka ho nepřestává v každém dopise <mentorovat>

memorovati ned. (z lat.) kniž. (co) ~) učít se něčemu z paměti: mechanicky m. látku ze zeměpisu; hlasitě, bezduché memorování

Příklad z ČNK:

Ani nestuduje text, ani nic <memoruje>.

- Lee upozornil na tzv. globální média (CNN aj.), jež jsou podle něho v rukou lidí, kteří morálně nedorosli tomu, aby někoho mohli <.....>, ale při- tom představují obrovskou sílu která „určuje orientaci světa“.
- Vědět a umět použití znamená stát se chytřím, <.....> znamená stát se omezeným, ohnutým a svaštělým na duši, na těle i v obličejí.
- Na půdě směnovny ministr vnitra Švehla <.....> národ: „Zvykli jsme si příliš mnoho na onu strašnou demoralizaci, kterou jsme prožívali za Rakous- ka.“
- Umělec ví jak čelit nepředvídatelnému, ale ten, kdo <.....>, zapo- mene, co se naučil a čemu nikdy nerozuměl.
- Rodiče by měli věst a ukazovat směr – ne <.....>. Budte pozitivním příkladem. Lehko řečeno, těžko provedeno.
- Prezidentův zvyk <.....> nadcházející volby je velmi nešťastný.

7. Zkušební komise není od toho, aby poslouchala, jak pěkně dokáže student <.....> kapitoly z příslušné učebnice.
8. Problémů ve školství je pochopitelně daleko víc – k těm nejhrošším patří přemířené požadavky osnov a přijímacích řízení na vysoké školy, které žáky i učitele základních a středních škol nutí celých třináct let nesmyslně <.....> hory nepotřebných informací.
9. Ale nechci tím jenom <.....>.
10. Antická rétorika učila <.....> tak, že řečník v myslí ukládal jednotlivé části své řeči do představy známého domu či města.

F. Významy slov

■ Úkol 1

ekonomie, -e ž. (z řec.) **1.** (politická) ekonomie věda zkoumající zákony společenské výroby a rozdělování materiálních statků na růz. stupních společenského vývoje **2.** hospodářství, zprav. polní: zabývat se e-í na státních statcích **3.** hospodárnost, úspornost: princip e. duševní práce ***4.** hospodářství, statek: prospívající e.; chov dobytka na e-ích

ekonomika, -y ž. (z řec. zákl.) **1.** souhrn výrobních a hospodářských vztahů ve vývoji společnosti, v urč. státě, v odvětví výroby, v urč. podniku, souhrn odvětví národního hospodářství; vědní obor zkoumající tato odvětví: e. socialismu; e. průmyslu; e. dopravy; zemědělská e.; e. práce; e. JZD; e. dolů **2.** hospodárnost, úspornost, ekonomičnost: stavební e.

Které slovo a který jeho význam....?

1. V knize La Bureaucratie 1946 napsal Ludwig von Mises: Pod vládou státem kontrolované <.....> nemožnou parlamenty být ničím jiným než širo-mážděním mužů, kteří vzdychky říkají ano!
2. Neuražej! Žes mi dal kouli z <.....>, to ještě nemusím bejt blbá.
3. Idaho znal <.....> Imperia, na nejvyšší míru zjednodušenou: žádně úroky se nepovolovaly; peníze na drevo.
4. Pan inženýr <.....> bude končit teprve příštím rokem a navíc musí na rok na vojnu.
5. Nobelovu cenu za <.....> pro rok 1994 získali profesor John Harsanyi z University of California v Berkeley, profesor Nash z Princeton University a profesor Reinhard Selten z Rheinische Friedrich – Wilhelms – Universität v Bonn.
6. Díky jistým znalostem nemocniční <.....> odhadoval, že AmeriCare musel posílnout finanční kolaps.
7. Kolokace, víceslovná pojmenování mužů mt! Komponentů více než dva jako v předchozím příkladu, zvl. v případě speciálních termínů, slov, několikavý-tenový řadový obráběcí stroj a ekvivalentní angl. multi-gang-spindle machine, nebo i vlastních názvů (třeba institucí aj.), jejich přílišný růst však naráží na požadavek jazykové <.....>.

8. Jestliže má věřitel prostředky (v našem případě jich má až požehnaně) k vy-máhání splatné pohledávky, je proti smyslu zákona i proti zásadě <.....> trestní represe nahrazovat nástroje k vymáhání pohledávky opatřeními orgánů činných v trestním řízení.

9. Transformace naší <.....> v efektivně fungující tržní systém nebude bezbolestná, budeme muset změnit mnoho svých návyků, budeme muset spolu-podepsat novou – jak se říká v odborné literatuře – „společenskou smlouvu“, která se od té minulé bude v mnohém lišit.

Úkol 2

Klika

1. dtžadlo k otvírání dveří, vrat, oken
2. rameno s rukojetí k ručnímu otáčení rumpálu, růz. strojků ap.
3. skupina jednotlivců, kteří se navzájem podporují
4. (z něm. zákl.) mít štěstí (op. smůla):
5. **čistit kliky** chodit s prosbou o přímluvu:
6. cviky – **klik/kliky**

1. Konečně stál před onou místností. Sáhl na <klika> a ono zamčeno.

2. Eduardo Crawley má za to, že obsah dokumentu se naprosto shoduje s po-tvrzenými výroky této vojenské <kliky> a říká, že „jestliže to byl padělek, pak mimořádně dokonale.“

3. Mám přitom ještě <klika>, že si ten Pída netroufne víc, v jeho očích přece-jen Bonek patří k nám do rodiny.

4. Položte si stejný předmět na zátylek nebo na záda a dělejte <kliky>.

5. PRODÁM 2 CTERY a kuřířkový gramofon s <klikou>.

6. Přestože v Holšovických začínají lešit <kliky> sponzoři, kteří zde tuší záji-mavou reklamu, hodně místních lidí se domnívá, že UNESCO jim přineslo problémy.

7. Měli jsme šílenou <klika>.

8. Tehdy ovšem přišla k moci revizionistická <klika> Dubčeka.

9. Ten druhý kluk dělal <kliky> a karate kopy, aby nějak překonal nervozitu.

10. Vládnoucí <kliky> sice opanovaly širší kraj, ale jsou obleženy ve svých pevnostech, a spoléhají zatím jen na to, že nikdo z nás neodhadne vlastní sílu.

11. Černej nebožtíku Máš to ale <klika> Za chvíli do temny hlíny budeš zakopán

12. Bez ohledu na to, že tehdejší právní řád v souladu s vůlí vládnoucí mocen-ské <kliky> chránil socialistické společenské zřízení.

13. Člověče, ale měl jsem <klika>, že jsi mi tehdy neustřítil pytlík.

14. Můžete ovladací <klika> přepečlivě zaplombovat, stanovit pravidla pro použití tohoto přístroje, ale nemůžete zatučit, že některý z všemocných na-konec nepootočí <klikou> a nepustí do sítě sílcí proud hrubě slíly.

15. Byla to <klika> ho potkat.

16. Hodně lidí čeká, že jejich poslanec bude běhat po ministerstvu financí, leš-tit <kliky> a shánět peníze pro jejich region.

17. Prosim, vsupte! Máte <klika>, mám zrovna volho.

18. Z Lucie stala se Lizzy Wagnerová, dcera hajného německé národnosti Dietri-cha, a Jan Pallas byl v této verzi chudým vysokoškolačkem – národovcem, je-hož demobilizační hradiň <klika>, ovládající vysoké školy, vyšvala do Sudet.

19. Billy otočil <klikou> a zatápal po vypínači v hale, rozsvítil a pustil Lowryho napřed.

20. Je to možné, tolik <kliky>? Může člověk být zamilovaný do dvou holek na-jednou?

21. Položil dlaň na kouli <kliky> a zvolna jí otáčel, potom dveře pootvěřel.

22. „Ave!“ zvedl ruku. „Maria“, dodala Nadja a zvedla desku, nasadila ji na gra-mofonový talíř, pak chtěla natáhnout péro, ale <klikou> otočila naprázdno.

23. Tím se dostává do rozporu se zájmy sovětské imperialistické <kliky>, a pro-to ho pronásledují.

24. Měli jsme <klika>!

25. Homais začal „čistit <kliky>“

26. Dělal <kliky> na jedné ruce a celou váhu těla přenášel na ochrnutý prst.

27. Galerie Zichlá <klika> na Starém Městě doplnila výstavu o zajímavá díla ze Sívkovy volně tvorby.

28. ... jak ti šílení, mý kámoši a společníci, poskakujou a dělají <kliky> a boxu-jou se ...

29. ...žádný producent o mne neměl zájem a já sám jsem nechtěl lešit <kliky>,“ říká režisér o této pauze.

30. Nikdy se nedá vysřídat, pokaždé leť v čele, asi má strach o svou prestiž, třeba mu v té smečce dělá rozruch nějaká opoziční <klika>.

Úkol 3

Maska

1. prostředek k zakrytí tváře umožňující nositeli vystupovat v jiné podobě – škra-boška

2. vnější předstíraná podoba, převážka, grimasa

3. *posmrtná maska* – odlitek tváře zesnulého

4. prostředek nebo zařízení k ochraně tváře, dýchacích cest a podobně

5. celkové přestrojení

6. přestrojená osoba

7. hlava + uši – zejména u psů modelované šlechtěním – termín

8. přední část přístroje, karoserie – termín

9. prostředek k péči o pleť, vlasů apod.

10. frazém *strhnout/nasadit masku*

1. Sledovali jsme róby a <masky>, něco jsme pak použili.
2. Usměvavá <maska> mu spadla s tváře.
3. Přenášení tradičních <masek> na hlavní postavy, jak je můžeme sledovat u Voskovece a Wericha, je antipicováno v Pirandellově hře Šest postav hledá autora.
4. Puritáni byli odjakživa mistry farizejských <masek>.
5. Všechny <masky>, které tu vidíš, ti zakryjí celý obličej a úplně tě změní.
6. Kosmetická hlinka Bílá hlinka, která dává konzistenci <maskám> a peelin-govým krémům.
7. Ten byl postaven na hlavu a jeho tvář kryla posmrtná <maska> ze zlata a se smaragdovými očima.
8. Okamžitě, automaticky, bez zásahu pilotů se nad hlavami cestujících začí-nají houpat kyslíkové <masky>.
9. Desítky a desítky různé zdatilých zvířecích i historických <masek> prochá-zely včera ponurými ulicemi pražského Žižkova v již tradiční lidové veselí-ci.
10. Pro ně je místo klasické barvy nejlepší řešením takzvaná barvící <maska>, která vlas vyživuje a zároveň barví.
11. Přístrojen je do ochranné plynové <masky> (rouška by v tomto případě ne-stačila), na rukou azbestové rukavice a v nich vařečka.
12. Gólmani si pomalují <masku> hrůzostrašným zjevem jako válečníci nebo pravěcí lovcí mamutů.
13. Jen <maska> bolesti na tváři mužově byla pravá.
14. Termín <maska> je používán pro pigmentaci hlavy. Tato pigmentace však musí zasáhnout celou přední část hlavy a uši.
15. Přenosovací podsystém v tomto případě obsahuje prioritní dekodér s <mas-kou> povolených přenosů.
16. Nová výrazná chromovaná přední <maska> dává vozu sympatický až dravý sportovní nádech a odlišuje ho od konkurence.
17. Dobře, že je zde takový žalobce jako pan Brückler, co umí strhnout <mas-ku> podjatosti a očistit nevinné.
18. Ovšem, navenek, protože ve společnosti si dovedu nasadit <masku> pos-měváčka.
19. Když zjistili, že to máme v hlavě srovnané a že to není jen nějaká byrokrat-tická umanutost, odložili <masky> suvenérů a docela se snažili.
20. Smích rezonuje v bolestně vědoucí výsměch prázdnyh <maskám> vyha-ných životů.

Úkol 4 balík

1. něco sbaleného, zabaleného, zprav. svázaného a uzpůsobeného k přenášení
2. větší množství něčeho

3. celek, složený ze vzájemně souvisejících položek
4. soubor položek dodávaných zakazníkovi jako celek
5. společenský neobratný, necvičený člověk; neotesanec, nemotora, hlupák
6. frazém být bohatý

Doplňte v kontextech číslu jednotlivé významy slova **balík**

1. Finanční částku, za niž kontrolní <balík> akcii získal, ale údajně odmítl sdě-lit.
2. Například nejnovější verze Windows 95 zabere v počítači šedesát megaby-tů, <balík> kancelářských programů Microsoft Office dalších sto dvacet a doplňkové programy Plus Pack čtyřicet megabytů.
3. Ministr školství Zeman úspěšně prosazuje, aby se jeho <balík> peněz na vý-platy zvýšil o 15 procent...
4. Jenom ze sotva číselného razítká jsem rozpoznala, odkud byl <balík> po-slán, věru na žádný dík nečekali.
5. V OKD bych měl zastupovat stát, který vlastní významný akcionářský <ba-lík> této společnosti.
6. ... přišel jsem z venkova, čili jsem byl takovej jako venkovskej <balík>...
7. Pro mnohé lidi je stále pohodlnější si naše služby zaplatit, než se doručová-ním dopisů a <balíků> zdržovat.
8. V zájmu dlouhodobé politiky však nepovažují za účelné, aby <balík> akcii získal ČEZ.
9. Škoda, že nemůžu být u toho, když bude <balík> rozbalovat ...
10. Druhým <balíkem> problémů jsou technické a provozní předpoklady, infra-struktura je bohužel zcela zastaralá a neodpovídá dnešním požadavkům.
11. Nebyla žádný venkovský <balík>.
12. Protože byl vydán celý <balík> nové legislativy, omezím se pouze na vý-znamný zákon č. 244/92 Sb.
13. Jactie pošle v <balíku> špinavé prádlo na vyprání.
14. Další aplikací byla sbírka zákonů a poslední byla instalační sada antivirové-ho <balíku> Avast!
15. Zbytek celého <balíku> legislativních opatření na schválení v dolní komoře teprve čeká.
16. Boženka, dostaly jsme <balík>, volávala, ona dostávala z domova <balíky> přes pět kilo, samý uzený maso a cukroví.
17. O problematice audiovizuálního <balíku> se v ženevském sídle GATT jed-ná již od středy a široké plénum účastníků mu dodává celosvětový rozměr.
18. Oba nejvyšší ruští představitelé se shodli na tom, že nyní je nejdůležitější, aby parlament bleskově schválil devatenáct zákonů z tzv. antikrizového <balíku>.
19. Ale více než 100 z těchto programů jsou hry, softwarové <balíky> pro mo-demizaci domova nebo se týkají multimédií.

20. Čišík, který si pod objednávkou kávy nepředstavil nic jiného než pořádného turka, si vysloužil celostátní ostudu jako venkovský <balku>, nemající šajnu o tom, co se patří mezi kulturními lidmi.
21. V prodeji televizních práv přitom panuje zvyk, podle něhož je například každý „atraktivní“ film dodán v <balku> spolu s několika dalšími „bečky“ a seriály různé kvality.
22. Až budeme v <balku>, vytáhnou tě vodsad, než řekneš popel, platí?
23. Pro Slovensko je tento zápas náročnější, protože do něj vstupuje s těžším <balkem> problému.
24. ... v městě jsou davy lidí, všichni jsou v <balku>, přesto závidí.
25. Z Trendu se ztratili i majoritní <balku> akcí Kotvy.

Úkol 5

Konference

1. porada politiků, odborníků apod.
 2. pravidelně zasedající orgán
 3. porada učitelského sboru
 4. sportovní událost, soutěž
 5. porada prostřednictvím internetu
 6. informování novinářů politiky a dalšími osobnostmi
1. Mluvčí České biskupské <konference> Daniel Herman upozornil, že se výzdoba kaple stane stálou vizitkou úspěšlosti a úrovně českého křesťanského umění konce dvacátého století.
 2. Proti zavedení školního se večera vyslovila Česká <konference> rektorů.
 3. Na věrejší dárcovské <konferenci> pro Kosovo, kterou uspořádala Světová banka a Evropská komise, se vybraly první dvě miliardy dolarů, které stačí na pokrytí okamžitých potřeb kosovského obyvatelstva.
 4. MEZINÁRODNÍ DIAKONICKÁ <KONFERENCE> na téma „Kolik práce potřebuje člověk?“ proběhla ve dnech 25. – 29. dubna v Jilovíšti u Prahy.
 5. Na Evropských <konferencích> se mají totiž scházet šéfové států a vlád zemí, které o členství v unii usilují, tedy nejen ty, s nimiž už se jedná, ale také ty, které sice mají vidinu členství, ale až v dlouhodobějším časovém horizontu.
 6. Toronto čeká ve finále <konference> Buffalo nebo Boston.
 7. V dokumentech Helsinské <konference> z roku 1975 se zdůrazňuje jako hlavní princip neměnnost hranic v Evropě.
 8. Součástí této prezentace bude také internetová <konference>, tj. diskusní fórum k jednotlivým tématům i k vlastním návrhům.
 9. V malajsijském hlavním městě Kuala Lumpur zahájilo večera třináct států Organizace islámské <konference> jednání o dosažení míru v Bosně a o její budoucí obnově.

10. V kulocárech <konference> se totiž delegátů metajili nespokojenosti s jeho politickým stylem.
11. Když není na nějaké mezinárodní lékařské <konferenci>, konec týdne trávi obvykle v domku u Prahy, kam může s rodinou jezdit celoročně.
12. Mezihradní <konference> by měla podle názoru některých delegací omezit právo veta jednotlivých států při přijímání závazných rozhodnutí v unii.
13. V Haagu probíhá První mírová <konference>, která vymezuje pravidla vedení války.
14. Čapek po mnichovské <konferenci> žil úplně v zoufalé náladě.
15. Roku 1990 této publikaci udělila Národní <konference> amerických učitelů angličtiny orwellovskou cenu Dvoji řeči (Doublespeak Award).
16. Atlanta naopak svedla ve výřazovacích bojích tuhé bitvy a ve finále Národní <konference> jen se štestím porazila nejlepší tým základní části Minnesoty 30 : 27 po prodloužení.
17. Konaly se specializované odborné <konference> např. na téma Vlivu a odstraňování účinků radonu a také celodenní konference o inženýrském vzdělávání.
18. Celostátní ornitologická <konference> nazvaná Dravci a sovy '91 začala v Přerově.
19. Z tohoto hlediska bychom e-mailový dialog museli považovat za reakci spontánně mluvenou (počítačově <konference> mají v tomto ohledu charakter smíšený).
20. V době, kdy se u nás diskutuje o demokratizaci a liberalizaci školství je zajímavé slyšet, že školních <konferencí>, které řeší všechny podstatné problémy školy, se v Německu zúčastňují nejen učitelé.
21. Po semináři a exkurzi následovala tisková <konference>, která byla dobře obsazena zástupci tisku, kteří směřovali své otázky nejen do oblasti panelových domů, ale i do ekonomických a sociologických oblastí bytové politiky.
22. V úvodním seminále Východní <konference> Pitsburg zvítězil nad New Jersey 3 : 2.
23. Přihlášení do elektronické <konference> a zaslání příspěvků:
24. Moji pořádní sekretáři mi sdělili, že jsem měl 154 řečí, 31 <konferencí> s žurnalisty a 29 anglických rozhlasů.
25. Vyrostl z toho problém, který si vyžádal zvláštní <konferenci> učitelského sboru.

4. Úkoly pro náročnější

A.

ÚVOD

Je-li korpus morfologicky značkován, je možné v něm zjišťovat morfologické vlastnosti slov a syntaktické vlastnosti vět nebo jejich částí. Proto je vhodné se naučit zadávat dotazy obsahující slovnítruhové a morfologické vlastnosti slov a vytvářet z nich zajímavé konfigurace. Následující skupina úkolů vám ukáže, jaké bohatství slov a jazykových struktur korpus obsahuje. Zároveň si můžete srovnat získané výsledky s údaji obsaženými v různých mluvnicích češtiny. Úkoly budou uspořádány podle příbývájící obtížnosti.

■ Úkol 1

Nalezněte v korpusu všechna slova v akuzativu.

■ Úkol 2

Nalezněte v korpusu všechna slovesa v minulém přechítí a poté v trpném přechítí.

■ Úkol 3

Nalezněte v korpusu všechny přechodníky. Co byste dělali, kdybyste nevěděli, jak se v morfologické značce vyjadřují přechodníky? Jak byste si pomohli?

■ Úkol 4

Nalezněte v korpusu všechny kontexty, v nichž se po předložce, která se nepojí s 6. pádem, bezprostředně objevuje slovo v 6. pádě. Je něco takového vůbec možné? Kolik konkordančních řádků jste obdrželi?

■ Úkol 5

Nalezněte v korpusu všechna lemmata podstatných jmen, která se vyskytují jen v množném čísle, a jsou to tedy buď podstatná jména pomnožná nebo taková podstatná jména, která sice mají jednotné číslo, ale vyskytují se pouze v čísle množném.

70

■ Úkol 6

Nalezněte v korpusu všechna lemmata podstatných jmen, která se vyskytují nopak jen v jednotném čísle.

■ Úkol 7

Sestavte na základě korpusu co největší soupis nedokonavých sloves.

■ Úkol 8

Utvoríte množinu sloves, která se pojí s genitivem.

■ Úkol 9

Zadejte dotaz specifikující dvě sousedící pozice (pomocí tvarů *a/nebo* lemmat *a/nebo* značek), abyste na výstupu nedostali žádný konkordanční řádek. Zadejte tedy „nemožnou“ konfiguraci!

■ Úkol 10

Ověřte vhodným dotazem do korpusu, zda po spojkách *aby*, *kdyby* se z určitých slovesných tvarů nachází v teže vedlejší větě vždy jen minulé přechítí.

■ Úkol 11

Jak byste co nejlhůtější zjistili **všechny** podoby akuzativu singuláru mužského rodu zájmena „on“? Kolik jich je?

■ Úkol 12

Pořďte z korpusu rychle a elegantně co nejrozšáhlejší soupis sloves v základním tvaru, jejichž předmět může být vyjádřen vedlejší větou uvozenou „že“!

■ Úkol 13

A na závěr něco jednoduššího. Jaké české slovo obsahuje nejvíce vyskytů písmena *ř*?

71

B.

ÚVOD

Pomocí korpusu je také možné vyhledávat takové struktury, které mají část fixní a část proměnlivou. V následujících cvičeních se pokuste sestavit korpusový dotaz, v němž proměnlivou část zachyíte pomocí gramatických značek. Potom pomocí statistických funkcí Frekvenční distribuce nebo Kolokace zjistíte nejčastější lexikální naplnění.

- **Úkol 1**
Vyhledejte konstrukci typu *bez ladu a skladu*.

- **Úkol 2**
Vyhledejte konstrukci: *kde je + substantivum, tam je + substantivum*.

- **Úkol 3**
Zadejte dotaz pro vyhledání konstrukcí *tož se to mluví* (různá zájmena a slovesa).

- **Úkol 4**
Jak byste v korpusu vyhledali věty zakončené na *proč ne já/ty...?*

- **Úkol 5**
Zadejte dotaz pro vyhledání konstrukcí typu *mám podlahu umytou*. (různá substantiva a adjektiva)

- **Úkol 6**
Zadejte korpusový dotaz na vyhledání konstrukce typu *co je svět světem*.

- **Úkol 7**
Zadejte dotaz pro vyhledání konstrukce typu *když nechceš, tak nechceš*.

- **Úkol 8**
Vyhledejte v korpusu zkratku *s. r. o.*

- **Úkol 9**
Vyhledejte v korpusu konstrukci typu *komu není rady (tomu není pomoci)*.

- **Úkol 10**
Zadejte v korpusu posloupnost osobního zájmena a přičesti – typ *já myslím*.

5. Práce se subkorpusech

ÚVOD

V této části se naučíme vytvářet subkorpusech, tj. dílčí korpusech zahrnující jen ty texty, na nichž chceme zkoumat určitý jazykový jev nebo slovní zásobu. Chceme například vytvořit subkorpusech vědeckých textů, které mají značku textového typu SCI. V menu Korpus – Vytvoření subkorpusech nejprve subkorpusech pojmenujeme, například SCI a potom zapíšeme podmínku: `txtype="SCI"`. Pokud jsme vše dobře zapsali v seznamu korpusech v pravém rohu přibude název našeho subkorpusech společně s uvedením, ze kterého korpusech byl vytvořen – tedy v podobě `syn2000:sci`. Pokud naopak některý textový typ do uvedeného korpusech zahrnout nechceme, zapíšeme to takto: `txtype!="SCI"`. Vytvoříme tak subkorpusech například s názvem NE_SCI. Jeho název se nám rovněž objeví v nabídce korpusech jako `syn2000:ne_sci`. V rámci těchto subkorpusech potom zadáváme dotazy stejným způsobem jako ve velkém korpusech.

V následující úkolech budeme pracovat s korpusem FSC2000 – to je korpus doplňující „Frekvenční slovník češtiny“.

Úkol 1

Vášim úkolem bude zjistit, co mají společného slova: *automotodrom*, *béčkový*, *rozlučkový*, *žebříčkový*, *zapařádkůh* a *zpěvaččin*. Nejprve se pokusíme zjistit frekvenci těchto lemmat v korpusech FCS2000. Aby nám to šlo rychleji, zatrhne v menu Korpus – Implicítní atribut – lemma. To nám umožní psát do dotazového řádku rovnou lemmata a nemusíme je složitě zapisovat do uvozovek a hranatých závorek. Řešením našeho problému bude zjištění v jakých textových typech se příslušná slova vyskytují. Po vyhledání lemmatu zvolíme v menu Konkordance – Statistiky – Frekvenční distribuce a v seznamu vedle položky Atribut vybereme `doc:txtype` – tím dostaneme frekvenční rozložení podle žánrových typů. V kolika žánrech se výše uvedená slova vyskytují?

Můžeme se o tom přesvědčit i vytvořením subkorpusech, který bude obsahovat jen publikaci (PUB) a posleze subkorpusech, který bude obsahovat všechny texty kromě publikaci (BEZPUB). Vyhledáváním v jednotlivých subkorpusech můžeme vyskyt lemmat ověřit.

Úkol 2

Vyhledejte si lemma *rezidenční* v subkorpusech PUB a BEZPUB. Najděte lemmata *amplituda* a *zařkadlo* v obou subkorpusech a porovnejte jejich výskyt.

1. Práce s Českým národním korpusem krok za krokem

A.

■ Úkol 1

V korpusu zadáme tvar **Hrad** a získáme 1020 výskytů.

a) Při pohledu na Rozložení textů v korpusu zjistíme, že nejčastěji se vyskytuje v publicistice (těchto textů je také v korpusu nejvíce), ale je často i v odborných textech.

b) Můžeme porovnat, kdy se jedná o Pražský hrad jako skutečnou stavbu, instituci (Kancelář prezidenta), kdy se píše o jiných hradech a začíná tímto tvarem věta, například: *Hrad leží uprostřed chráněné krajinné oblasti;* *Hrad Stětkov nadále zůstal zpřístupněn veřejnosti...* *Ještě nedávno byl Hrad napojen výhradně na státní rozpočet;* *V sobotu odpoledne jsme šli Letnou na Hrad.*

c) 36

<Hrad> Bezděz	<Hrad> Boskovice	<Hrad> Bouzov
<Hrad> Buchlov	<Hrad> Buzice	<Hrad> Bítov
<Hrad> Canossa	<Hrad> Duino	<Hrad> Grabštejn
<Hrad> Hasištejn	<Hrad> Helštyň	<Hrad> Hukvaldy
<Hrad> Kašperk	<Hrad> Klášter svatého Jiří	<Hrad> Kost
<Hrad> Kámen	<Hrad> Křivoklát	<Hrad> Landštejn
<Hrad> Malenovice	<Hrad> Melice	<Hrad> Okoř
<Hrad> Pecka	<Hrad> Picimberk	<Hrad> Ploskovice
<Hrad> Potštejn	<Hrad> Rožmberk	<Hrad> Rožnov
<Hrad> Rýznberk	<Hrad> Sovinec	<Hrad> Stětkov
<Hrad> Svojanov	<Hrad> Valdstejn,	<Hrad> Veverž
<Hrad> Špilberk	<Hrad> Švihov	<Hrad> Červený kámen

■ Úkol 2

V korpusu zadáme dotaz **[Hh]hrad** a získáme 2275 výskytů.

Pokud nestojí tvar *Hrad* na začátku věty, jedná se o *Pražský hrad*, případně *Kancelář prezidenta republiky* apod. Lze použít frekvenční distribuci, případně rozložení k větší ilustraci jednotlivých významů.

Dále můžeme pokračovat vyhledáváním spojení *Pražský hrad* – různé kombinace malého a velkého počátečního písmena – zjišťujeme tak chybný zápis, ve kterých textech se objevuje apod.

Úkol 3

V korpusu zadáme dotaz [Pp][Rr][Aa][Hh][Aa], popř. [lc="praha"] a získáme 77305 výskytů.

Slovo PRAHA se vyskytuje 18370x, nejčastěji v kontextech typu PRAHA 5, slovo praha se vyskytuje 29x. Náležitě je v e-mailových a internetových adresách, nejvíce se objevuje v díle Orten, Jiří: Horňký kruh: korespondence s Věrou Fingerovou.

Úkol 4

V korpusu zadáme dotaz [jji]t[te]nka a získáme 81 výskytů.

Velké písmeno uvozuje vlastní jméno (divadlo, farma, žena, horská bouda, děchovka, hvězda, jídelna, lidová muzika, ústav sociální péče, pěvecký sbor, předávka ženy, divadelní společnost či soubor, stacionář, továrna, název části Sezimova Ústí). Nejde tedy o vlastní jméno hvězdy.

Malé písmeno uvozuje pojmenování hvězdy, která vychází před východem slunce, tedy pojmenování jitrní hvězdy. (Bývá jí myšlena obvykle Venuše, ale může jí být i Merkur, jde tedy o třídu jevů.)

Úkol 5

V korpusu zadáme dotaz [Ee][Ss][Oo] a získáme 534 výskytů.

(Je možné též zadat jako [lc="eso"].)

eso 234

Eso 224

ESO 76

European Southern Observatory

Expresní silniční okruh

Ekologické spalování odpadu

Elektronický systém obchodování

Emergency Service Operator

Úkol 6

V korpusu zadáme postupně dotazy

[lemma="bryle"] – 2280 výskytů

[lemma="brejle"] – 206 výskytů

Porovnat můžeme zastoupení v různých typech textů a pak zapojení do více-slovných jednotek.

CELKEM *bryle* 2280

brejle 206

beletrie (NOV, COL)

816

139

beletrie (FAC)

75

16

kráská literatura bez označení

3

13

básně a písně (VER, SON)

31

13

scenáře (SCR)

8

13

ostatní

41%

82%

1347

59%

38

18%

Z uvedených čísel vidíme, že mluvená varianta má velmi výrazné zastoupení v beletrii. Z konkrétních příkladů můžeme dále usoudit, že původní odhad dělicí obě varianty na základě mluvenosti či psanosti bude platný. V beletrii jde u příkladů s výrazem *brejle* většinou o přímou řeč, nebo o vyprávění v „ich“ formě, které s mluveným stylem koresponduje také. V publicistice a odborné literatuře se vedle běžného významu vyskytuje často výraz *brejle* v názvech desek (*Brejle*), pořadů (*Zelený brejle*), divadelních her (*Brejle*) nebo jako součást frazémů (*mít něčeho plný brejle*, *růžový brejle*).

Při zkoumání víceslovných jednotek si všimnete, že výraz *brejle* se nerozvíjí tak často adjektivem, jak bychom možná čekali, spíše jim předcházejí slovesa, která mají tento výraz za objekt (předmět) – *mít, nasadit, nosit, být v, být bez, vzít, sundat*. Adjektiva rozvíjející *brejle* jsou převážně kvalifikační a žádné takovéto spojení nelze označit za termín.

Bryle se vedle adjektiv kvalifikačních spojují běžně i s adjektivy relaxačními a na rozdíl od *brejle* mají naopak relativně bohatou škálu terminologických souloví: *ochranné b., dioptrické b., speciální b., plavecké b., potápěčské b., sportovní b., lyžařské b.*

Některá slova se v korpusu vyskytují jen s jedním z těchto výrazů (*potápěčské bryle* x *babičkovský brejle*), většína slov se ale může spojit jak s výrazem *bryle*, tak s *brejle*. Při pozorném prozkoumání můžeme odhalit jasně tendence, například *bryle* je spojení lexémů spíše standardní (snad s tendencí k terminologickému chápání), proto je v korpusu nikoli 10x častěji, ale téměř 90x častěji, než *tmavý brejle*, kde se lexémová standardnost slova *tmavý* (ze spojení *tmavé bryle*) kříží s výraznou mluveností formy *brejle*.

Úkol 7

V korpusu zadáme postupně dotazy

[lemma="sirup"] – 349 výskytů

[lemma="syrup"] – 5 výskytů

Častější je *sirup* a toto slovo budeme považovat za standard. Veliký rozdíl mezi počty výskytů naznačuje, že *syrup* nebude zřejmě variantou, ale že důvod jeho

vyskytu v textech bude jiný. Nahlednutím do Pravidel zjistíme, že psaní „sirup“ současná pravidla nabízejí jako jediné, zřejmě na základě pravopisu latinského (*sirupus*). Původem jde však o slovo arabské. Texty s pravopisem *syrup* se tedy opírají o starší pravidla, kde skutečně *syryp* fituroval s „y“ (podobně jako u Jungmanna či ještě v PSJČ). U Machka („Eymologický slovník jazyka českého“) zmiňované lišení mezi *syrupem* (lékárenským) a *syrupem* (cukrovarmickým) se zdá celkově nepodložené.

Úkol 8

V korpusu zadáme postupně dotazy

Lemma=“proteřovat” – 44 výskytů

Lemma=“protěřovat” – 17 výskytů

Proteřovat je vyskytuje v korpusu 44x a jde o variantu preferovanou standardem, neboť toto slovo přišlo do češtiny z francouzštiny (infinitiv: *protéger*).

Protěřovat se vyskytuje 17x, tato varianta však vznikla mylnou asociací ke slo-
vům *ztěřovat*, *stěřovat*, *obtěřovat*. Výskyt obou výrazů v korpusu není natolik
frekventovaný, abychom byli v případě „*protěřovat*“ schopni jednoznačně hovo-
řit o zařazení do standardu, či o vyřazení z něj. Pokud by se v budoucnu ve vět-
ším korpusu toto rozložení pravopisné volby potvrdilo (2/3 : 1/3), pravděpodob-
ně by šlo o varianty.

Úkol 9

V korpusu zadáme dotaz **Lemma=“+ú,+”** a získáme 50689 výskytů

Po odfiltrování lemmat začínajících na předponu *ne-* N-Filtr (Lemma=“ne.*”) do-
staneme 44186 výskytů
Pomocí funkce Konkordance – Statistiky – Frekvenční distribuce dospějeme
k následujícímu seznamu (v tabulce jsou slova do frekvence 50):

<i>zúčastnit</i>	11101	súdánský	242	veleúspěšný	89
<i>zaučít</i>	2061	<i>bezúplatný</i>	234	búrský	87
<i>zúčastněný</i>	2059	<i>bezúhonný</i>	233	Hodúl	84
<i>fúze</i>	1943	infúze	194	Saúd	78
<i>vyústit</i>	1241	<i>trojúhelníkový</i>	193	naučťovat	77
<i>trojúhelník</i>	1194	ndbjiský	187	fúzťovat	76
<i>spoluúčast</i>	1182	Zubek	187	Búr	76
<i>protiútok</i>	870	<i>zúčťovací</i>	183	<i>minouúrovňový</i>	74
<i>zúčastňťvat</i>	586	<i>širokouhlý</i>	182	Porúť	74
<i>zúčťovánť</i>	523	<i>zúčťastněň</i>	175	<i>zaučťťvat</i>	74
<i>Múza</i>	509	<i>bezúplattě</i>	159	manikúra	73
<i>Zúžit</i>	464	<i>zaučť</i>	155	Farúk	71
<i>Vyústěň</i>	439	<i>bezúttěšň</i>	153	<i>zúčťastněň</i>	69
Núbie	433	<i>poúňorový</i>	149	<i>zúžen</i>	69

<i>vyúčťovánť</i>	433	sebeúctá	136	čťvřúhelník	66
<i>Kúřa</i>	422	<i>vyúčťťvat</i>	135	Súřův	65
<i>víceúčťťlový</i>	414	Núbjiec	131	dřezúra	65
Súdán	411	Radúz	130	<i>zaučťťovánť</i>	64
Bejřút	374	<i>zúčťťvat</i>	128	<i>bezúčťťlný</i>	63
<i>zúřčťť</i>	373	Masúd	122	trubadúr	62
<i>bezúřčťťň</i>	334	Raúl	121	Rangún	61
<i>Zúžený</i>	299	<i>bezúspěšně</i>	120	<i>meziúvěř</i>	61
Júřus	294	Húřka	118	<i>zúřčťťň</i>	59
<i>bezúhonnost</i>	290	<i>transúřzní</i>	116	<i>rozúčťťovánť</i>	59
<i>úřa</i>	280	Charťum	112	<i>předúňorový</i>	59
<i>protiúřťavní</i>	280	<i>dlířzní</i>	110	mundúr	56
<i>skúř</i>	273	Púřťov	109	zlatouřť	54
<i>transúřze</i>	272	Mahmúd	107	<i>spoluúčťťnkovať</i>	54
<i>múřický</i>	268	sauđřkoarabřský	106	<i>trojúhelníkovitý</i>	53
<i>zúženť</i>	265	<i>jednoučťťlový</i>	104	<i>infúřzní</i>	53
<i>međuza</i>	261	<i>vúřdú</i>	96	Ařjűna	53
<i>samoučťťlný</i>	255	<i>vyúřťťvat</i>	95	<i>zúčťťťvaný</i>	53
<i>pravouřľhý</i>	242	<i>protiúřder</i>	92	<i>poúřřazový</i>	53

Jednotlivě skupiny

a) přejatá slova (např. *múza*, *skúř*) – v šedém rámečku

b) vlastní jména a jejich odvozeniny (*Bejřút*, *súdáňřský*) – tučně

c) domácí slova odvozená (např. *zúčťastnit se*, *bezúřčťň*) – kurzíva

d) domácí slova – kompozita (např. *trojúhelník*, *veleúřpěšň*) – podtržené

Úkol 10

V korpusu zadáme dotaz **[word=“+ú,+”** a dostaneme 2061 výskytů.

podstatná jména:

ragú, *abú*, *vúřdú*, *igľú*, *keřú*, *řurjú*

číslovce:

tú, *jú*, *hú*, (*zú* není číslovce, ale jen část špatně rozděleného slova), *bú*, *juh-*

hú, *řú*

Problematické podstatné jméno označuje známé ořřky. V korpusu najdeme ke-
řú 18x a keřú 8x. Rozdíř tedy není nijak závratný: Chceme-li zvažťvat, která
z těchto variant je standardní, můžeme se kromě frekvence (18x keřú) vřít do
úvahy ř počet zdrojů (textů), ve kterých se ta která varianta vyskyřta. Keřú se vy-
skýřuje ve 12 zdrojích, zatímco keřú jen ve 2 zdrojích. To ř přřehledně ř počtu
zdrojů, ve kterých se výraz vyskyřl, nám pomohlo při rozhodovánř, že standard-
ní variantou je zřejmě keřú.

Slovo *řurjú* pocházř z řiřřtiny a označuje „elegantní zábavu“, event. „vyřřbenou
eleganci“.

Úkol 11

V korpusu zadáme dotaz [(lemma="*.ismus") | (lemma="*.izmus")] a získáme 58873 výskytů. Po odfiltrování: *izm. * zůstane 863 doklady.

A Rozhodnutí zrovnoprávní psaní slov tvořených příponou -ismus/-izmus neodpovídá úzu.

B Vysokou frekvenci zakončení na -izmus mají slova **organizmus; mechanismus**.

C	celkové	počet	počet
organismus/organizmus	4604	organizmus	231
mechanismus/mechanizmus	4521	mechanismus	265

Úkol 12

V korpusu zadáme dotaz [lemma="*gó.*"] a dostaneme 18781 výskytů.

Na základě Frekvenční distribuce získáme seznam 180 slov.

Nejfrekventovanější je *gól* a také má nejvíce odvozenin.

Úkol 13

V korpusu SYN2000 zadáme dotaz **voh.*** a získáme 166 dokladů.

Na základě funkce Frekvenční distribuce dostaneme seznam slovních tvarů s frekvencemi, který si uložíme jako soubor Word, např. s názvem ukol13.doc. (pro šetření místem byl tento seznam převeden do tabulky):

vohěň	13	vohledně	2	vohlídnout	1	vohánějí	1
vohně	8	vohryzku	2	vohořet	1	vohradě	1
vohnul	4	vohryzek	2	vohřívaly	1	vohutit	1
vohřát	3	vohelbá	2	vohodaná	1	vohnutýho	1
vohoil	3	vohýnku	2	voháknutá	1	vohoblovali	1
vohněm	3	vohýbají	1	vohryzat	1	vohromnou	1
vohelbal	3	vohnutý	1	vohavnosti	1	vohybu	1
vohni	3	vohlížel	1	vohábla	1	vohoz	1
vohoit	3	vohněm	1	vohled	1	vohavná	1
vohelbat	3	vohlašení	1	vohlídal	1	vohněš	1
vohromný	3	vohodin	1	vohodnocení	1	vohřívát	1
vohýněk	3	vohořelá	1	vohnout	1	vohyb	1
vohelbák	3	vohromnej	1	vohňutácký	1	vohánět	1
voholený	2	vohňo	1	vohradách	1	vohrozili	1
vohlídne	2	vohniček	1	vohnut	1	vohryzky	1
vohavný	2	vohouskové	1	vohlády	1	vohodnotili	1
vohodit	2	vohanbín	1	voháklej	1	voholí	1
vohne	2	vohlížet	1	vohmatávám	1	vohlodanejma	1
voháknout	2	vohromná	1	voholenej	1	vohradu	1

vohelbání	2	Vohmataly	1	vohozenejch	1	vohlídej	1
vohozu	2	Vohlazenej	1	vohrada	1	vohradý	1
vohelbají	2	Vohodiš	1	vohlíd	1	vohrožuju	1
vohon	2	Vohlžení	1	voholiji	1	vohmnou	1
vohňostroj	2	Vohryzal	1	vohledu	1	vohlídám	1
vohnutej	2	Vohvězdíčkovanej	1	vohnat	1	vohlížet	1
vohánět	2	Vohyzdný	1	vohřívám	1	vohánějí	1
vohledy	2	Voháknutej	1	vohřívaly	1	vohradě	1

Celý tento seznam označíme v pozadí světlejší barvou (tato nabídka je rovněž součástí programu Word). Potom zadáme stejný dotaz **voh.*** v korpusu PMK a získáme 74 dokladů. Uložíme ho do dalšího souboru, např. ukol13a.doc.

vohodnotit	7	vohni	2	voháknutej	1	vohleduplní	1
vohodnocení	6	vohledu	2	vohlídnout	1	vohlašovací	1
vohodnocená	4	vohromnej	2	vohniška	1	vohybání	1
vohled	4	vohodnocenej	2	vohrožovat	1	vohnout	1
vohánět	3	vohlídnou	1	voholený	1	vohledy	1
vohadení	3	vohřežete	1	vohlídala	1	vohřáta	1
vohěň	3	vohromnim	1	vohlídání	1	vohradý	1
vohlídne	3	vohniško	1	vohrazoval	1	vohybat	1
vohromný	3	vohlídá	1	vohodnotil	1	vohněš	1
vohodnocení	3	vohlídat	1	vohlídnou	1	vohraný	1

Potom přepokopírujeme seznam ze souboru ukol13a.doc do souboru ukol13.doc za stávající seznam a pozadí odlišíme jinou barvou. Oba seznamy potom necháme abecedně seřadit a zaměříme se na ty tvary, které se ve spojených a seřazených seznamech budou opakovat s různým barevným pozadím. Dostaneme tento výsledek:

voháknutej	1
voháknutej	1
vohánět	1
vohánět	3
vohěň	13
vohěň	3
vohled	1
vohled	4
vohledu	1
vohledu	2

vohledy	1
vohledy	2
vohlídne	2
vohlídne	3
vohlídnout	1
vohlídnout	1
vohneš	1
vohneš	1
vohni	2
vohni	3
vohodnocení	1
vohodnocení	3
vohrady	1
vohrady	1

Seznamy různých slov

a) SYN2000 –

voháklet, voháknout, voháknutej, vohanbín, vohanět, vohavnost, vohavný, vohájba, vohajbání, vohajbat, vohaj, vohlášení, vohlazenej, vohled, vohledně, vohlídnout, vohlídat, vohlídnout, vohlížet, vohlížení, vohlodanej, vohlousek, vohmatat, vohmatávat, vohnat, vohnout, vohníček, vohnivej, vohnířstvoj, vohnůtčický, vohnutej, vohoblovat, vohodit, vohodnocení, vohodnotit, voholenej, voholit, vohon, vohořel, vohorět, vohoz, vohozenej, vohrábnout, vohrada, vohromnej, vohrozit, vohrožovat, vohryzát, vohryzek, vohřát, vohřívát, vohulít, vohvězdičkovanéj, vohyb, vohýbat, vohýnek, vohyzdne

b) PMK –

vohadění, voháknutej, vohánět (se), vohaj, vohlašovací, vohled, vohleduplnej, vohlidanej, vohlídat, vohlídnout (se), vohmatávat, vohnisko, vohnout, vohodnocenej, vohodnocení, vohodnotit, voholenej, vohrada, vohranej, vohrazovat, vohromit, vohromnej, vohrožovat, vohřát, vohybání, vohybat

Úkol 14

V korpusu zadáme tvar **víže** a dostaneme 18 výskytů. Pomocí funkce Rozložení zjistíme, že převažuje v básnických textech (TXTYPE = VER). Potom v korpusu zadáme spojení **váže k** a získáme 71 výskytů. Toto spojení se objevuje hlavně v populárně naučných a odborných textech.

Slovesné tvary **víže** a **váže** jsou zaměnitelné v těch kontextech, kde jsou odkazy na historické události. stov. např. doklady z korpusu Akce se **<váže k>** oslavám 1001. výročí vzniku nejstarší dochované písemné zmínky o zdejší obci... x jeden z prvních zápisů psaných ještě švabachem v Paměti kniže Staré Boleslavi se **<víže>** k roku 1836.

Úkol 15

V korpusu zadáme dotaz **[lemma="čičhat"]** a získáme 403 výskytů.

Sloveso **čičhat** se užívá ve všech tvarech, zatímco u slovesa **čuchat** převládají tvary 3. osoby singuláru a plurálu, tj. **čuchá, nečuchá, čuchají(í), nečuchají(í)**. Ize snadno zjistit pomocí menu Konkordance – Statistky – Frekvenční distribuce, kde si zadáme výpis tvarů (word).

Sloveso **čuchat** se nejčastěji užívá o psech.

Úkol 16

V korpusu zadáme dotaz **jako řípa** a získáme 5 výskytů. Ve všech případech předchází slovo **zdravý**, tj. jde o frazém (**být) zdravý jako řípa**. Tento frazém je plně ustálený a najdeme ho v 1. svazku SČFI – Přirovnání.

Potom v korpusu zadáme dotaz **jako řepa** a získáme 4 výskytů. Ve dvou případech se přirovnává k rudému oblíčeji (**zrudl jako řepa, byl rudý jako řepa**), v jednom případě o přirovnání **nechal se okopávat jako řepa**. Přestože lze tyto doklady charakterizovat jako přirovnání, jde o přirovnání neustálená, a tudíž je zatím nemůžeme zařadit do oblasti obecně užívané frazeologie.

Jeden výskyt je navíc přímým užitím slova **řepa** v kontextu **pěstovaná nejčastěji jako řepa krmná**.

Úkol 17

V korpusu zadáme dotaz **[lemma="*uální"]** a získáme 25437 výskytů.

Seznam získaný funkcí Frekvenční distribuce si uložíme např. jako ukol17.doc. Potom zadáme dotaz **[lemma="*uální"]** a získáme 62 výskytů. Stejným postupem vytvoříme další frekvenční seznam, který si uložíme např. jako soubor ukol17a.doc. Poté prvních 10 položek z obou seznamů porovnáme.

individuální	5298	eventuelní	27
aktuální	5257	individuální	7
sexuální	4412	sexuelní	6
intelektuální	1903	vizuální	4
eventuální	1719	procentuelní	3
vizuální	1159	homosexuelní	3
virtuální	1137	rituelní	2
kontinuální	698	habituální	2
audlovizuální	578	aktuální	2
homosexuální	510	spirituelní	2

Zakončení **-uální** je charakteristické především pro slova z intelektuální oblasti, zatímco zakončení **-uelní** implikuje spojitost s duchovnem, zvyky a rituály.

Výslovnost **-elni:** je vlivem francouzštiny a němčiny, stov např. doklady vyhledané pomocí dotazu **spirituel.***:

...Surin, *Le guide <spirituel> Svatý Jan od Kříže, monografie o romanice ve Španělsku...* Dies ist es, und andere Essays über Zen und <spirituelles> Erfahrung.

Úkol 18

V korpusu zadáme dotaz ([lemma="z"] [lemma="který"]) a dostaneme 3634 výskyty.
Potom využijeme funkci Frekvenční distribuce, a to tak, že pozici 0 určíme jako word a pozici 1 jako lemma – viz obrázek:

Frekvenční distribuce

Zadání frekvenční distribuce

Limit: 0

Atribut: word Ignorovat velikost

Pozice 0 od < KWMC

Atribut: lemma Ignorovat velikost

Pozice 1 od < KWMC

Přidat Smazat Budíž Zavřít

Dostaneme následující seznam:

ze	který	2773
z	který	816
Z	který	34
Ze	který	9
ZE	který	1

Celkový poměr mezi ze + který a z + který je 2783 : 850.

Úkol 19

V korpusu zadáme dotaz [lemma="é...?"] a získáme 722 výskyty.
Pomocí funkce Frekvenční distribuce zjistíme tato lemma s nejnižší frekvencí 10:

éter	322
étos	170
était	16
éčko	15
époque	12
élan	10

Prejatá slova v tomto seznamu pocházejí z latiny (éter), z řečtiny (étos) a z francouzštiny (était, époque, élan).

Úkol 20

V korpusu zadáme dotaz ([lemma="*|hrk|:.*") & (word=".....")] a získáme 66966 výskyty. Potom zadáme N-Filtr (menu Nový dotaz – Nfiltr) a napíšeme: *|AÁBCDDĚĚFGHIJKLNNNOÓPQRSŠTĚUVWXYZ|. * Zůstane nám 26798 výskyty. Na základě Frekvenční distribuce vyjde následující seznam slov a jejich frekvencí:

krize	10227
série	6676
princ	1378
riziko	771
skin	724
trik	632
prima	591
krini	496
ring	375
hit	366
triko	321
tarif	231
rival	183
šerif	166
skica	165
gril	152

B.

Úkol 1

V korpusu zadáváme postupně dotazy

- [lemma="lo"] – 3664 výskytů
- [lemma="hon"] – 1468 výskytů
- [lemma="loven"] – 46 výskytů
- [lemma="honěn"] – 51 výskytů

Úkol 2

Řešení tohoto úkolu se týká pouze přípony –ček.

V korpusu zadáme dotaz [lemma="*ček"] a získáme 93239 výskytů.

Potom pomocí funkce Frekvenční distribuce získáme seznam, z něhož vybereme následující jednotky:

Čistě zdrobněliny: dědeček, mláček, míček, krůček, chlapeček, zámeček, domeček, kopeček, zajíček, žáček, klobouček, listeček, vojáček, ukazováček (totéž meček, kopeček, slavíček, ctvereček, rybníček, potůček, koneček, sameček, co ukazovák), slaviček, ctvereček, rybníček, potůček, koneček, sameček.

Zdrobněliny se samostatným významem: žebříček (poslouchanost v hudbě), nováček, balíček (ekonomický), stromček (vánoční), klíček (rostlina), zajíček (darebák), sáček, jídelníček, stromček (vánoční), koníček (hobby), ptáček (mlady člověk opáčeného pohlaví), hřebíček (koření), chlebiček (např. obložený), měsíček (rostlina), váleček (na nudle), kartáček (na zuby, typ vousů), panáček (o osobách).

Součástí frazémů: háček (něco má h.), rámeček (nedat si něco za r.), strýček (udělat si s-a, pro strýčka Příhodu)

Nejsou zdrobněliny: špaček (též ve frazému: tlouct špačky), hlouček, křeček, cvrček.

Úkol 3

V korpusu zadáme dotaz [lemma="*nutí"] a získáme 69079 výskytů.

Pomocí funkce Frekvenční distribuce (lemma) zjistíme těchto 10 nejfrekventovanějších substantiv:

rozhodnutí	25458
hnutí	11088
sítěnutí	2878

poskytnutí	1630
odmítnutí	1516
spiknutí	1094
uplynutí	902
státnutí	737
zvládnutí	643
přihlédnutí	557

Potom nastavíme P-filtr s pozicí -1 a zadáme dotaz [tag="R.*"]. Získáme 15645 výskytů.

Opět zadáme funkci Frekvenční distribuce, avšak tentokrát pro dvě pozice (lemma) – vzhledem k tomu, že substantiva tohoto typu se skládají podle vzoru *stavení*, velmi často splývá tvar (word) a lemma, liší se jen instrumentál.

Frekvenční distribuce

Zadání frekvenční distribuce

Limit:

Atribut: Ignorovat velikost

Pozice: od

Atribut: Ignorovat velikost

Pozice: od

S nejvíce předložkami se poji nejfrekventovanější substantivum rozhodnutí, srov. k rozhodnutí, na rozhodnutí, proti rozhodnutí, podle rozhodnutí, z rozhodnutí, s rozhodnutím, o rozhodnutí, po rozhodnutí, před rozhodnutím. Lemma rozhodnutí má 25458 výskytů.

Ustálené kolokace

s přihlednutím	494	s přihlednutím k -víceslovná předložka
bez povšimnutí	352	(zůstat) bez povšimnutí – frázém
z rozhodnutí	332	z rozhodnutí někoho – víceslovná předložka
na spadnutí	244	(být) na spadnutí – frázém
bez hnutí	216	(stát) bez hnutí – frázém
k prasknutí	183	(nacpaný nabitý) k prasknutí – frázém
za povšimnutí	160	(ne stát) za povšimnutí – frázém
k nahlédnutí	141	(nechat, nabízet) k nahlédnutí – frázém
k hnutí	110	(nebýt) k hnutí – frázém
bez mknutí	101	bez mknutí (oka) – frázém

Úkol 4

V korpusu zadáme dotaz ***ýlko** (popř. **[word="*ýlko"]**) a získáme 237 výskytů. Po seřazení na základě funkce Obecné třídění – lemma – Vynechat duplicity (sami smaže zbylé duplicity s velkým písmenem a překlepy), získáme 14 výskytů.

bidýlko (SSJČ pod *bidélko*)
divadýlko (SSJČ pod *divadélko*)
hradýlko (SSJČ nemá)
křídýlko (SSJČ pod *křídélko*)
letadýlko (SSJČ nemá)
meidýlko (SSJČ nemá)
mejdýlko (SSJČ pod *mejdličko*)
práďýlko (SSJČ pod *práďélko*)
ráďýlko (SSJČ pod *rádlo* – *rádélko*, *ráďýlko*, *ráďličko*)
sedýlko (SSJČ pod *sedélko*)
strašidýlko (SSJČ pod *strašidlo* – *strašidélko*, *strašidýlko*)
svěťýlko (SSJČ pod *světélko*)
tykadýlko (SSJČ nemá)
zrcadýlko (SSJČ nemá)
žihadýlko (SSJČ pod *žihadlo* – *žihadélko*, *žahadélko*)
žráďýlko (v SSJČ je)

V korpusu zadáme dotaz ***élko** (popř. **[word="*élko"]**) a získáme 286 výskytů. Stejným redukčním postupem jako v předcházejícím případě získáme 18 výskytů: *bidélko*, *čerpadélko*, *divadélko*, *hrdélko*, *kladélko*, *křídélko*, *kyvadélko*, *letadélko*, *práďélko*, *sedadélko*, *sedélko*, *skélko*, *stébélko*, *strašidélko*, *světélko*, *šidélko*, *tykadélko*, *umyvadélko*, *zrcadélko*.

I když slovník zaznamenává tvary zakončené na **-ýlko** za tvary na **-élko**, dává jim tedy přednost, většinou na nás působí tvary na **-élko** knižně a strojeně.

Úkol 5

V korpusu zadáme dotaz **[lemma="*tko"] & (tag="N.*")** a získáme 18707 výskytů.

Na základě Frekvenční distribuce zjistíme 3 nejfrekventovanější zástupce z obou skupin:

Názvy nástrojů

<i>tačičko</i>	1877
<i>sluchátko</i>	1493
<i>razičko</i>	1149

Zdrobněliny

<i>zrcátko</i>	705
<i>zviřátko</i>	673
<i>děvčátko</i>	656

Úkol 6

V korpusu zadáme dotaz **[lemma="*spolu.*"] & (tag="Vf.*")** a získáme 2732 výskytů.

Potom zadáme dotaz **[lemma="*proti.*"] & (tag="Vf.*")** a získáme 35 výskytů. Poměr mezi infinitivem a všemi slovesnými tvary zjistíme tak, že zadáme vytež dotazy, avšak na místě tagu nebude Vf.*, ale V.*

Vznikne následující poměr:

Protiv. 35 infinitiv: 312 slovesné tvary, *spolu-* 2732 infinitiv: 7734 slovesné tvary. Slovesa začínající *spolu-* se často vyskytují se slovesy modálními a těmi, která označují začátek a konec děje, tj. slovesy fázoými. Tyto typy sloves se ale spojují s infinitivem a tím jeho výskyt u sloves s předponou *spolu-* zvyšují. Vidět to můžeme na poměru mezi Vf a V.

Úkol 7

V korpusu zadáme dotaz **[lemma="*+jít"] & (tag="V.*")** a získáme 211985 výskytů.

Po Obecném třídění nám zůstane následující seznam:

dojít, *nadejít*, *najít*, *obejít*, *odejít*, *podejít*, *pojít*, *poodejít*, *popojít*, *projít*, *předejít*, *přejít*, *přijít*, *rozejít*, *sejít*, *ujít*, *vejít*, *vjít*, *vyjít*, *vynajít*, *vzejít*, *zajít*.

V SSJČ není uvedena nevolkalizovaná předpona *v-*, tj. *vjít*.

Úkol 8

V korpusu zadáme dotaz **[lemma="*pod.*"] & (tag="N.F.*")** a získáme 106194 výskytů.

<i>podmínka</i>	36458	abstraktní substantivum s předponou <i>pod-</i>
<i>podpora</i>	21161	abstraktní substantivum s předponou <i>pod-</i>
<i>podoba</i>	17494	abstraktní substantivum s kořenem <i>podob-</i>

podstata	13796	abstraktní substantivum s předponou pod-
podlaha	4934	konkrétní substantivum s předponou pod-
podobnost	3098	abstraktní substantivum s předponou po-
podivná	1134	abstraktní substantivum s předponou po-
podobnost	962	abstraktní substantivum s kořenem podob-
podložka	761	konkrétní substantivum s předponou pod-
podobizna	413	konkrétní substantivum s kořenem podob-

Úkol 9

V korpusu zadáme dotaz **[[lemma="ex.*"] & (tag="Vf.*")]** a získáme 3054 výskytů.

Po provedení funkce Vynechat duplicitu a smazání tvarů s velkými písmeny; tvaru na -ti a záporných tvarů zůstane následujících 19 infinitivů:

exalovat	expedovat	exhumovat	extrapolovat
excerpovat	expandovat	ekskulповat	exponovat
exportovat	ekkomunikovat	explodovat	extrahovat
existovat	exploatovat	explikovat	exhibovat
experimentovat	excelovat	ekekovat	

Chceme-li vyhledat všechny slovesné tvary, zadáme dotaz **[[lemma="ex.*"] & (tag="V.*")]** a získáme 32155 výskytů. Po vynechání duplicit a pročištění příbude tato slovesa:

examinovat, excitovat, exítovat (chybný tvar), exotizovat, extendovat

Úkol 10

Pro vyhledání tohoto dotazu je vhodný vzor 11, tj. v korpusu zadáme dotaz **[[lemma="spolu.*tel"]]** a získáme 1267 výskytů.

spolunamítel	648	SSJČ ano,SSČ ano
spoluzakladatel	276	SSJČ ano,SSČ ano
spolupořadatel	145	SSJČ ano,SSČ ne
spolupachatel	97	SSJČ ano,SSČ ne
spoluobvyvatel	20	SSJČ ano,SSČ ne
spoluvydvavatel	15	SSJČ ano,SSČ ne
spoluvystavovatel	9	SSJČ ne,SSČ ne
spolužadatel	8	SSJČ ano,SSČ ne
spolupřítel	6	SSJČ ano,SSČ ne
spoludřítel	6	SSJČ ano,SSČ ne
spolupředitel	5	SSJČ ano,SSČ ne
spoluučitel	4	SSJČ ano,SSČ ne

spoluřešitel	4	SSJČ ne,SSČ ne
spolupřekladatel	4	SSJČ ano,SSČ ne
spoluprovozovatel	4	SSJČ ne,SSČ ne
spolubojevitel	3	SSJČ ano,SSČ ne
spolunositel	2	SSJČ ano,SSČ ne
spolunavrhovatel	2	SSJČ ne,SSČ ne
spolucestovatel	2	SSJČ ano,SSČ ne
spoluhositel	1	SSJČ ne,SSČ ne
spoluvychovatel	1	SSJČ ne,SSČ ne
spolubýdřitel	1	SSJČ ano,SSČ ne
spoluuživatel	1	SSJČ ano,SSČ ano
spolupronásledovatel	1	SSJČ ne,SSČ ne
spoluzřizovatel	1	SSJČ ne,SSČ ne
spolupodnikatel	1	SSJČ ano,SSČ ne

Polička s šedým pozadím označuje substantiva doložená ve všech třech zdrojích, substantiva vytištěná tučně jsou zachycena jen v korpusu SYN2000.

Slovník spisovného jazyka českého dále zaznamenává následující substantiva: *spolubudovatel, spolučinitel, spolubojevitel, spoluspisovatel (dnes spouautor), spoluvotřitel (dnes spoluvtřice), spoluučitel, spoluvětel.*

Nově zachycená substantiva v korpusu odražejí současnou realitu (např. *spoluprovozovatel, spoluzřizovatel, spoluřešitel/grantového projektu*), zatímco jednotky zachycené jen v SSJČ jsou odrazem reality doby socialismu, srov. *spolubudovatel, spolučinitel* apod.

Úkol 11

V korpusu zadáme dotaz **[[lemma="*přeskol*"]]** a získáme 171 výskyt. Pomocí Frekvenční distribuce (menu Konkordance – Statistika – Frekvenční distribuce /lemma/ dostaneme tento seznam:

<i>přeskolit</i>	52
<i>přeskolení</i>	46
<i>přeskolovat</i>	25
<i>přeskolování</i>	20
<i>přeskolovací</i>	14
<i>přeskolový</i>	11
<i>přeskolovaný</i>	2
<i>přeskolkovat</i>	1

Slovesem, od kterého jsou odvozena téměř všechna další slova, je *přeskolit*. Je také nejfrekventovanější. Sloveso *přeskolkovat* je od slova *školkovat*, resp. substantiva *školka*, užívá se v zahradičnictví.

Úkol 12

V korpusu zadáme dotaz [lemma="*pozic[č].*"] a získáme 50515 výskytů. Na základě Frekvenční distribuce vznikne seznam, z něhož vybereme 5 nejfrekventovanějších substantiv:

dispozice	10652	sklon k něčemu; volné použití něčeho, předpoklad k něčemu
expozice	9124	výstava
opozice	8724	protiklad; odpor vůči něčemu, protistrana
kompozice	1471	skladba, sloh
propozice	113	výpověď, věta

Úkol 13

V korpusu zadáme dotaz ***nark[lo].*** (popř. [word="*nark[lo].*"] a získáme 1902 výskytů.

a) Pomocí funkce Obecné třídění – Vynechat duplicity (word) a při zobrazení kontextu 0 pozic (Zobrazení – Kontext) získáme seznam odvozených slov s jediným výskytem.

V korpusu jsou tato odvozená slova:

exnarkoman, hypernarkotizovaný, narko, narkobaron, narkobyznys, narkodemo-
kracie, narkodiktátor, narkodolar, narkoexperiment, narkogen, narkokartel, nar-
korporace, narkokuryr, narkolepsie, narkoleptika, narkomaňán, narkomafie,
narkoman, narkománie, narkomanka, narkomanský, narkoobchod, narkoob-
chodník, narkopenize, narkoskandal, narkoterorismus, narkotest, narkotizační,
narkotikum, narkotický, narkotický, narkotin, (ne)narkotizovat, narkotizace, nar-
kóza, narkóza, narkosa, protinarkotický, znarkotizovat, znarkotizovaný
b) V SSIČ je zachycena slovní zásoba, která klade důraz na využití narkotik pro
lékařské účely, zneužívání drog signalizuje jen slovo *narkoman*, nemá ani pře-
chýlenou podobu *narkomanka*. V současném jazyce převažuje zneužívání drog
a jazyk zachycuje způsob obstarávání drog. Při vzniku nových slov je používáno
přebírání celých slov z angličtiny a jejich adaptace – *narkobyznys*, *narkosa/nar-
kóza*; odvozování – *znarkotizovaný*, *protinarkotický*, skládání (převažuje) – *nar-
kodolar*, *narkoobchod*...

Úkol 14

V korpusu zadáme dotaz [lemma="*[řgžčň]o") & (tag="N.N.*")] a získáme 73 318 výskytů. Ze seznamu Frekvenční distribuce vypíšeme následující expresivní substantiva:

peřo	266	pero
lejsřto	211	úřední papír, listina
cigářo	121	cigareta
fářo	69	auto

panděro	23	břícho
kvádro	23	oblek
vzrušo	14	vzrušení
šampáňo	10	šampaňské
cígo	7	cigareta
šřlřgro	7	šřestř
řvřaro	6	cigareta
hausnumero	4	velké říslo, velký počet
imřgo	3	image
ěřo	1	letadlo

Tato expresivní substantiva mají v psaných textech zpravidla nízkou frekvenci – nižší než 100.

Slovo *čučo/čučo* označuje levné sladké víno. Podobně zakončení mají další ná-
poje, srov. *pivčo, kaččo*.

Úkol 15

V korpusu zadáme dotaz [lemma="*ouš") & (tag="N.[M].*")] a získáme 2265 výskytů.

Mezi substantiva s příponou *-ouš* nepatří slovo *řouš*.

Jména vlastní:	Termíny:	Jména obecná zápomná:	Jména obecná kladná:
Matouš 602	běřouš 213	řomouš 57	řtarouš 27
Perouš 284	řocholouš 29	řeplouš 48	řrahouš 4
Milouš 270	řolouš 10	řomouš 1	řděřouš 4
Palouš 212	řřkouš 8	řhyřroušř	
Bohouš 146	řypouš 4	řzřlouš 1	
Vřřouš 73	řřřřouš 1		
Vondřouš 66	vodřouš 1		
řřřouš 36			
řarouš 35			
řřřouš 21			
řarouš 19			
Venouš 16			
Kerouš 15			
řanouš 11			
Řadouš 5			
Vavřouš 1			

Slovo *bigouš* znamená „dobrovolník“, *mikrouš* je „mikrofon“, *tajouš* „tajný“, *Trafalgar square* v Londýně; *fouš* je část Václavského náměstí (nazvaná podle Trafalgar square v Londýně; používáno v 50. letech).

Pakouš – pako, *prdlouš* – prdlej, *psychouš* – psychicky nemocný, *psychicky* narušeny, *cukrouš* – cukroušek (odstranění zdrobňovací přípony), *člouš* – čilý, *krívouš* – krivý, *socdemouš* – člen sociální demokracie.

Úkol 16

V korpusu zadáme dotaz [**lemma**=“*auto*+”] & (**tag**=“*N.**”) a získáme 69378 výskytů.

V korpusu SYN2000 je celkem 181 substantivních kompozit. Z toho je celkem 110 kompozit, která odkazují ke komponentu *auto* s významem „automobil“, „autobus“.

A. Seznam kompozit

Autoakumulátor, autoalarm, autoanténa, autoatlas, autobaterie, autobazar, autobenzín, autobenzín, autobus, autobusák, autobusek, autobyznys, autocar, autocisterna, autocvičiště, autočáounění, autodíl, autodílna, autodoplnek, autodoprava, autodopravec, autodráha, autodrom, autodrožka, autoelektrikář, autogaráž, autohavárie, autohlička, autochladnička, autochrom, autojeřáb, autojeřábek, autokar, autokarta, autokemp, autokempink, autokino, autoklempič, autoklub, autokolárna, autokolona, autokolos, autokosmetka, autokros, autokrosař, autolak, autolakovna, autolakýrník, autolano, autolékárník, automafile, automapa, automasáž, automobilizace, automobil, automobilistka, automobilista, automobilizace, automontér, automotodrom, automotoklub, autonehoda, autookno, autoolej, autooprava, autoopravářensví, autoopravář, autoopravná, autopark, autoplošina, autopošta, autopotař, autopotřeba, autoprodělna, autoprovoz, autoprůmysl, autopřehrávač, autopřislušenství, autopůjčovna, autorádio, autorodeo, autosalón, autosedáčka, autoserwis, autoshow, autosklo, autoslužba, autosoučástka, autosprávkárna, autostop, autostopař, autostřada, autosvaz, autošampón, autoškola, autotaxi, autotelefon, autotest, autotransformátor, autotrenažér, autoturista, autoturistika, autoveterán, autovlak, autovrak, autovýrobce, autozámek, autožárovka

B. Adjektiva vytvořená od substantiv *auto, automobil, autobus, autokar*

<i>auto</i>	21154	<i>autový</i>	19
<i>automobil</i>	11983	<i>automobilový</i>	3940
<i>autobus</i>	7847	<i>autobusový</i>	1985
<i>autokar</i>	112	<i>autokarový</i>	23

Adjektivum *autový* není odvozeno od substantiva *auto*, ale *aut* (fj. ve fotbale apod., např. *autové* vřazování).

Úkol 17

V korpusu zadáme dotaz: [**lemma**=“*strukturní*”] a získáme 315 výskytů a potom [**lemma**=“*strukturnální*”] a získáme 1130 výskytů.

Uložíme si seznamy kolokací získané z obou vyhledávek jako zvláštní soubory (např. *strukturni.doc* a *strukturnalni.doc*) a pak tyto soubory mezi sebou porovnáme.

Nejčetnější kolokace

Výpočet nejčetnějších kolokací

Atribut: ↘

V rozsahu od: do:

Minimální četnost v korpusu:

Minimální četnost v daném rozsahu:

Maximální počet zobrazených řádků:

Seřadit podle četnosti: absolutní relativní

Ukázka 10 nejfrekventovanějších substantivních kolokátů ze souboru s názvem *strukturnalni.doc*:

Lemma	MI score	T score	Rel. f [%]	Abs. f
<i>změna</i>	9.27	15.3	0.5772	235
<i>fond</i>	8.982	12.02	0.4727	145
<i>reforma</i>	9.969	9.53	0.9368	91
<i>politika</i>	7.665	7.898	0.1897	63
<i>problém</i>	5.876	5.732	0.05488	34
<i>nezaměstnanost</i>	7.677	3.854	0.1912	15

přestavba	9.658	3.868	0.7549	15
přeměna	9.485	3.736	0.6695	14
vlastnost	6.974	3.137	0.1175	10
pomoc	5.055	2.91	0.03108	9

Jaké jsou častější kontexty (kolokáty), které se vyskytují jen s jedním z obou zkoumaných sloví?

strukturální – lingvistika, gen, mapa, průzkum, vzorec, význam, jednotka, organizace, popis, mutace, typ, isomer

strukturální – reforma, problém, nezaměstnanost, přestavba, přeměna, pomoc, adaptace, grafika, abstrakce, transformace, rozvoj, slabina, skladba, krize, opatření, příčina, program

Slovo **strukturální** je základní a pravidelně utvořené přídavné jméno, je slovo tvořené blíže slovu **struktura** nežli slovo **strukturální**. Je také blíže pojetí konkrétnímu. Ve většině případů je z širšího kontextu snadno určitelné, k čemu se základové slovo adjektiva vztahuje, a entita, ke které se vztahuje, může být konkrétní. Z kontextu či ze společného povědomí pisatele a čtenáře se pak může konkrétní struktura také explicitně vymezit, je snáze představitelná.

Sousloví lze většinou vyjádřit variantně také přívlastkem neshodným: *mapa struktury, průzkum struktury, vzorec struktury, význam struktury, jednotka struktury, organizace struktury, popis struktury, mutace struktury, typ struktury* (kde námi zkoumané slovo mění sice slovní druh, ale nemění svou závislost na druhém členu syntagmatu, tedy fakt, že je dominováno).

Z tohoto rámce se vymyká spojení terminologické: *strukturální lingvistika*.

Slovo **strukturální** se vztahuje více k entitám abstraktním a složitým (komplexním). Nebová možné takto označenou strukturu přesně definovat či změřit, spíše se upozorňuje na její strukturální charakter, nežli na její přesný tvar a formu. *Strukturální problém, přestavba či přeměna* se může týkat abstraktních a vágních entit (strukturální mapa, vzorec, jednotka ovšem jen entit relativně konkrétních).

I zde nalezneme syntagmata, která lze variantně vyjádřit přívlastkem neshodným: *reforma struktury, přestavba struktury, přeměna struktury*. Na rozdíl od adjektiva **strukturální** však syntagma s adjektivem **strukturální** nabízí relativně často takovou variantu, kde námi zkoumané slovo mění nejen slovní druh, ale také přímnočarost vztahu. Varianta s prostým genitivem je u nich nevhodná: *strukturální problém = problému týkající se struktury, pramenící ze struktury; strukturální nezaměstnanost = nezaměstnanost zasahující celou sociologickou strukturu obyvatel*. Ovšem i u výrazů, u kterých je varianta s genitivním neshodným přívlastkem možná (*strukturální přestavba – přestavba struktury*) může vést přímého vztahu (*přestavba struktury*) pocítovat velmi často také vztah zprostředkovaný (*přestavba zaměřená jinam, ale zasahující také strukturu; snad i přestavba odehrávající se ve všech částech struktury*). Tyto možnosti přispívají k větší abstraktnosti adjektiva.

Úkol 18

V korpusu zadáme dotaz **[[lemma="vy.*ni"]]** a získáme 133716 výskytů. Na základě frekvenční distribuce dostaneme tento výpis do frekvence 1000:

vydání	10046
vystoupení	7400
vytvoření	6745
vyšetřování	6471
vybavení	6212
vysílání	5718
vyjádření	5717
vyloučení	4139
vysvětlení	4013
vyhlášení	3227
vytváření	3144
vyšetření	3049
vyrovnání	2773
vyprávění	2715
využívání	2399
vybudování	1894
vytápění	1889
vydávání	1537
vyjednávání	1444
vyřešení	1279
vyhledávání	1226
vymezení	1216
vyznání	1163
vydrání	1133
vypracování	1084
vyhodnocení	1024

A. Poměr mezi nedokonavými a dokonavými slovesy je 10 : 16.

B. U této vyhledávky dále nastavíme P-filtr a kontext -1, potom zadáme dotaz: **[[lemma="během"] | (word="průběhu")]** a získáme nám 268 výskytů, popř. **[[word="během"] | (word="průběhu")]** a získáme nám 219 výskytů.

Spojení *během vyšetření* se týká lékařské prohlídky, *během vyšetřování* se vztahuje k policejnímu nebo soudnímu procesu. Poměr mezi kombinacemi předložek a substantiv:

v průběhu vyšetřování	75	během vyšetření	15
během vyšetřování	58	v průběhu vyšetření	6

Z tabulky vyplývá, že se substantivem *vyšetřování* je častější předložka v průběhu a se substantivem *vyšetření* předložka během.

Úkol 19

V korpusu zadáme dotaz `[lemma="*tel"] & (tag="N.M.*")` a získáme 278226 výskytů.
Mezi čístejská jména nepatří *čítatel, datel, jmenovatel, nepřítel, přítel, ukazatel*.

Úkol 20

V korpusu postupně zadáváme dotazy

<code>[lemma="nakladatel"]</code>	- 1330 výskytů
<code>[lemma="nakladáč"]</code>	- 101 výskyt
<code>[lemma="nosič"]</code>	- 2397 výskytů
<code>[lemma="nosič"]</code>	- 1506 výskytů
<code>[lemma="ředitel"]</code>	- 42694 výskytů
<code>[lemma="řidič"]</code>	- 11622 výskytů
<code>[lemma="ovladatel"]</code>	- 2 výskytů
<code>[lemma="ovladáč"]</code>	- 823 výskytů
<code>[lemma="voitel"]</code>	- 86 výskytů
<code>[lemma="voič"]</code>	- 9917 výskytů

Významy slov:

nakladatel – kdo financuje a zajišťuje vydávání knih
nakladáč – kdo nakládá a vykládá zboží
nosič – 1. kdo nosí, nese (jméno, titul), reprezentuje, představuje něco
nosič – 1. kdo se zaměřuje na nošením zavazadel; 2. kdo je určený k nošení; 3. věc, která nese, drží, podpírá něco
ředitel – osoba pověřená vedením nějakého organizačního celku
řidič – kdo řídí nějaké vozidlo apod.
ovladatel – kdo něco n. někoho ovládá
ovladáč – zařízení určené k ovládní nějakého přístroje, zvl. televize, rozhlasu apod.
voitel – zvolený zástupce voličů pro výkon (nepřímě) volby
voič – kdo má právo volit, kdo volí

C.

Úkol 1

V korpusu zadáme dotaz `dlk` (popř. `[word="dlk"]` a dostaneme 836 výskytů. Ide o nominativ a akuzativ substantiva *dlk* a o předložku *dlk něčemu*. Za předložkou vždy následuje substantivum v dativu, v okolí substantiva *dlk* jsou adjektiva, např. *všelij dlk, slovesa, např. patřit a předložka za, např. dlk za něco*.

Úkol 2

V korpusu zadáme dotaz `[bcčddfghiklmnpřššttvwxyz][bcčddfghiklmnpřššttvwxyz][bcčddfghiklmnpřššttvwxyz]` a získáme 16547 výskytů. Nebo zadáme dotaz takto: `[bcčddfghiklmnpřššttvwxyz]{4}` Na základě Frekvenční distribuce (word) dostaneme seznam tvarů, z nichž žádnému úkolu odpovídá tvar *skrz* s 940 výskytů. Tento tvar vyhledáme zvláštním dotazem.

Ve vyhledávce zadáme Redukci na 100 náhodných výskytů.

Poměr mezi předložkou a adverbium je 94% : 6%:

Adverbium *skrz*, např. *dívat se skrz* – 6 %

Předložka *skrz* ve významu ‚z jedné strany na druhou‘, např. *skrz křoví* – 54%

Předložka *skrz* ve významu ‚kvůli‘, např. *skrz penzi* – 16%

Frazémy 24%: *skrz naskrz*, popř. *skrz na skrz, dívat se na někoho skrz prsty, procedit skrz (zatáté) zuby*

Úkol 3

V korpusu zadáme dotaz `[lemma="během"]` a získáme 29684 výskytů.

Pomocí funkce Kolokace dostaneme seznam tzv. kolokátů v rozsahu 0, 3 a setříděných podle absolutní frekvence. Tento seznam si uložíme jako soubor *behem.doc*.

V seznamu lemmat nejprve vyřídíme pouze substantiva. V rámci tohoto seznamu převládá významová skupina ‚časový úsek‘, tj. *rok, den, měsíc, týden, minuta, hodina, doba, období, vteřina, stolec, desetiletí, sekunda*:

lemma	Ml-score	T-score	Rel. f [%]	Abs. f
rok	5.43	64.11	1.059	4307
den	6.316	45.86	1.956	2157
měsíc	6.785	34.53	2.707	1214
týden	6.59	33.84	2.366	1169

minuta	6.715	28.74	2.58	842
hodina	5.251	26.15	0.9349	721
doba	4.67	25.31	0.6249	694
období	5.612	17.11	1.201	305
noc	5.413	15.84	1.046	263
vteřina	7.346	13.66	3.993	189
stoleť	4.705	12.69	0.6405	174
včerejšek	7.793	11.99	5.445	145
desetiletí	7.013	11.66	3.17	138
sekunda	4.253	10.89	0.4683	132
večer	4.285	10.44	0.4785	121

Specifickými substantivy jsou včerejšek, víkend apod. (srov. úkol E3).

■ Úkol 4

V korpusu zadáme dotaz [(lemma="zdraví") & (tag="...4.*")] a získáme 3207 výskytů.

Slovo *zdraví* je nepočítatelné, nemůže se tudíž vyskytovat v plurálu. Pomocí funkce Frekvenční distribuce (v Atributu vybereme „tag“) můžeme zjistit přesný počet plurálových označení – 287.

Další tvary mohou být 1. pád – srov. např. *vzácnější než mladí je zdraví*, sloveso *zdraví* – např. *srdčně vás zdraví*, adjektivum *zdravý* v 1. pádě plurálu – např. *byli celí a zdraví*.

■ Úkol 5

V korpusu zadáme dotaz [(lemma="obchod") & (tag="...S6.*")] a získáme 2674 výskytů.

Objevi se tvary *obchodu* a *obchodě* – tvar *obchodě* se zpravidla užívá pro označení konkrétní místnosti, v níž se zboží prodává a nakupuje, tvar *obchodu* pro obchodování v abstraktním smyslu, např. *v obchodu s energií*.

Další substantiva, která mají variantní koncovky v 6. pádě jsou např. **parlament** (*parlamentu/parlamentě*), **obvod** (*obvodu/obvodě*), **případ** (*případu/případě*), **pokus** (*pokusu/pokuse*), **účet** (*účtu/účtě*), **roman** (*románu/románě*), **materiál** (*materiálu/materiále*), **jazyk** (*jazyku/jazyce*), **autobus** (*autobusu/autobuse*), **rybník** (*rybníku/rybníce*). Kolísá také tvar 6. pádu u slova **korpus** – korpusu i korpuse.

■ Úkol 6

V korpusu zadáme dotaz [(lemma="potřeba") & (tag="...4.*")] a získáme 5501 výskytů.

Potom nastavíme P-Filtr v rozsahu -1, 0 a zadáme další dotaz [tag="A.*"] nebo [pos="A"], zůstanou nám 1803 výskytů. Frekvenční distribuci zadáme jako lemma na pozici -1 a jako word na pozici 0. Dostaneme následující frekvenční seznam:

Lemma	word	frekvence	
běžný	potřebu	11	Abstr.
bylový	potřeby	10	Konkr.
denní	potřebu	17	Abstr.
denní	potřeby	25	Konkr.
domácí	potřebu	17	Abstr.
energetický	potřeby	11	Abstr.
hygienický	potřeby	20	Konkr.
jiny	potřeby	11	Konkr.
konkrétní	potřeby	13	Konkr.
lidský	potřebu	12	Abstr.
lidský	potřeby	11	Abstr.
malý	potřebu	14	Abstr.
naléhavý	potřebu	31	Abstr.
nový	potřeby	12	Konkr.
nutkavý	potřebu	15	Abstr.
nutný	potřeby	11	Konkr.
osobní	potřebu	64	Abstr.
osobní	potřeby	75	Konkr.
psací	potřeby	10	Konkr.
silný	potřebu	20	Abstr.
skutečný	potřeby	12	Abstr.
sociální	potřeby	19	Abstr.
specifický	potřeby	14	Abstr.
sportovní	potřeby	13	Konkr.
tělesný	potřeby	10	Abstr.
velký	potřebu	27	Abstr.
vlastní	potřebu	238	Abstr.
vlastní	potřeby	34	Konkr.
vnitřní	potřebu	48	Abstr.
vnitřní	potřeby	10	Abstr.
základní	potřebu	11	Abstr.
základní	potřeby	39	Konkr.
zvyšovaný	potřebu	16	Abstr.
živoť	potřeby	33	Abstr.

Singular častěji označuje abstraktní situaci, plurál konkrétní věc, srov. *potraviny pro osobní potřebu* x *osobní potřeby* (zvl. kartač, kartaček na zuby apod.). Velmi po komunikaci x *neuspokojují některé sice méně chválné, leč přirozené lidské potřeby* – lenost ducha i těla.

Úkol 7

Dotaz lze zadat několika způsoby. Nejpodrobnější a také nejlépe odpovídající zadání je následující forma dotazu, která kombinuje vzory 2, 6 a 17:

Každou z koncovek vyhledáme zvlášť:

[[word="*. * [i]]"] & (tag="N.MP1.) & ((lemma="politik") | (lemma="historik") | (lemma="syn") | (lemma="bůh") | (lemma="předek") | (lemma="svědek") | (lemma="vůdce") | (lemma="pán"))] – získáme 7 311**

[[word="*. * [oO][V][é]]"] & (tag="N.MP1.) & ((lemma="politik") | (lemma="historik") | (lemma="syn") | (lemma="bůh") | (lemma="předek") | (lemma="svědek") | (lemma="vůdce"))] – získáme 6336 výskytů.**

Místo **word="* [i]** a **word="* [oO][V][0é]]** lze užít rovněž **lc="*1** a **lc="*ové**. Z obou vyhledávek vytvoříme na základě Frekvenční distribuce seznamy, které si uložíme např. jako soubory **pani.doc** a **panove.doc** (budou však zahrnovat všechna zkoumaná substantiva s uvedenými koncovkami. Oba seznamy porovnáme a vytvoříme následující tabulku:

<i>páni</i>	1290	<i>pánové</i>	2538 (1147 – viz poznámka)
<i>syni</i>	–	<i>synové</i>	949
<i>politici</i>	4211	<i>politikové</i>	825
<i>předci</i>	112	<i>předkové</i>	562
<i>svědci</i>	554	<i>svědkové</i>	377
<i>vůdci</i>	649	<i>vůdcové</i>	293
<i>historici</i>	371	<i>historikové</i>	291
<i>bozi</i>	124	<i>bohové</i>	501

Dotaz lze však zadat i obecněji, tj.

[[tag="N.MP1.) & ((lemma="politik") | (lemma="historik") | (lemma="syn") | (lemma="bůh") | (lemma="předek") | (lemma="svědek") | (lemma="vůdce")) | (lemma="pán"))] – získáme 13647 výskytů.** V programu Bonito potom můžeme zadat Frekvenční distribuci seřazenou podle lc, abychom se vyhnuli malým a velkým písmenům na začátku tvarů.

A. Oba tvary některých slov jsou vzájemně lišeny sémanticky – *páni* (vyšší společenská třída); *pánové* (m.n.č. od označení pro muž)

B. Přípona *-ové* je považována jako knižní, formálnější, někdy je využívána k ironii. Poznámka: tvar *pánové* ve vyhledávce zahrnuje i tvary vokativu, např. *dámy a páňové*. Tohoto tvaru je více než nominativu, tj. 1391 výskyt, tudíž na nominativ připadá ve skutečnosti 1147 výskytů.

Úkol 8

V korpusu zadáme dotaz **[[tag="....2.**) & (lemma="káně")]** a získáme 25 výskytů. Po smazání tvaru *Káně* nám zůstane 21 výskytů. Distribuce tvarů je tato:

<i>káně</i>	10
<i>káněte</i>	7
<i>kánat</i>	4

Úkol 9

Dotaz můžeme zadávat pomocí tvaru: **[[word="*. *čtř") & (tag="A.MP1.**) – 12830 výskytů, nebo pomocí lemmatu: [(lemma="*. *čky") & (tag="A.MP1.**) – 12956 výskytů.**

Rozdíl v počtu výskytů jsou způsobeny tím, že v rámci tvarů píšeme jen malá písmena, zatímco při zadání lemmatu jsou podchylena i slova zapsaná celá velkými písmeny, tj. tvar *védečtř* zachytíme pomocí word, ale tvar *VEDEČTŘ* pomocí lemmatu, popř. pomocí atributu lc.

V následující přehledné tabulce byly dotazy zadávány pomocí lemmatu.

Přípona	příklad	1.p. pl.	začátek dotazu	hlásková změna
<i>-ní</i>	<i>lesní</i>	<i>lesní</i>	[(lemma="*. *ní")...]	
<i>-ný</i>	<i>smutný</i>	<i>smutní</i>	[(lemma="*. *ny")...]	n > ě
<i>-ský</i>	<i>městský</i>	<i>měštít</i>	[(lemma="*. *sky")...]	sk > št
<i>-cký</i>	<i>vesnický</i>	<i>vesničtř</i>	[(lemma="*. *cky")...]	ck > čtř
<i>-ový</i>	<i>světový</i>	<i>světoví</i>	[(lemma="*. *ovy")...]	
<i>-natý</i>	<i>svahatý</i>	<i>svahatí</i>	[(lemma="*. *naty")...]	t > ř
<i>-itý</i>	<i>děležitý</i>	<i>děležit</i>	[(lemma="*. * [v][itý")...]	t > ř
<i>-ovitý</i>	<i>kočkovitý</i>	<i>kočkovit</i>	[(lemma="*. *ovity")...]	t > ř
<i>-avý</i>	<i>chápavý</i>	<i>chápaví</i>	[(lemma="*. *avy")...]	
<i>-ivý</i>	<i>žárlivý</i>	<i>žárliví</i>	[(lemma="*. *ivy")...]	

Několik příkladů adjektiv, která nemají uvedené přípony: *nový, jiný, ženatý, pravý, živý* – v těchto případech hláskové skupiny *-ov-, -in-, -nat-, -av- a -iv-* nejsou součástí přípony, ale slovního základu.

Úkol 10

V korpusu zadáme dotaz **[(lemma="před") | (lemma="vánoce") & (tag="....7.**)]** a získáme 583 výskytů.

Na základě Frekvenční distribuce, kde pozice 0 je lemma a pozice 1 je word do- staneme tento přehled:

<i>před Vánocemi</i>	228
<i>před vánocemi</i>	201
<i>před Vánoci</i>	86
<i>před vánoci</i>	66
PŘED VÁNOCEMI	2

A. Poměr mezi tvary *vánoci* a *vánoce* je 152 : 431, poměr mezi psaním velké-
ho a malého písmene je 314 : 267.

B. Psaní velkého písmene tedy převládá jen slabě, zásada PČP není dodržována.

Úkol 11

K vyhledání využijeme vzor 20.

V korpusu zadáme dotaz [word="\."] [tag="R.*"] [tag="R.*"] nebo
[word="\."] [tag="R.*"]{2} a získáme 6808 výskytů.

Tabulku Frekvenční distribuce vytvoříme následujícím způsobem:

Frekvenční distribuce

Zadání frekvenční distribuce

Limit: 0

Atribut: lemma Ignorovat velikost

Pozice 1 od < KWIC

Atribut: lemma Ignorovat velikost

Pozice 2 od < KWIC

Přidat Smazat Budič Zavřít

A. V takto vzniklém seznamu si označíme ty předložky, které jsou ve své kombi-
naci už ustálené. Jako vodítko nám pomůže jednak vyšší frekvence a jednak pís-
meno F na 2. pozici gramatické značky (RF*), které signalizuje, že jde o „sou-
část předložky, která nestojí samostatně“.

vzhledem k	5557
nehledě na	245
na rozdíl od	79
nehledě k	56
s pomocí	14
vzhledem na	11

B. Sledem dvou předložek není např. kombinace o o (viz *L i v e r p o o l* – týká se
většiny kombinací neslabičných předložek).

Úkol 12

V korpusu zadáme dotaz [(lemma="v") | (lemma="na")] [word="Praze"]
a získáme 30331 výskytů. Dále použijeme menu Konkordance – Statistika –
Frekvenční distribuce-KWIC> + přidat <KWIC.

A. v Praze 30091
na Praze 240

B. Spojení na Praze převažuje s číslem Prahy – určením pražského obvodu.

Další příklady:

v Kladně 533
na Kladně 270

kolej – vyberte pouze ve významu ‚studentská ubytovna‘
v koleji 28

na koleji 92

na Vinohradech 678
ve Vinohradech 2

Vyjádrění pomocí předložky na expanduje především v mluveném jazyce.
Příklady: na vládě, na bazénu, na baru...

Úkol 13

V korpusu zadáme dotaz [(lemma="*pnout") & (tag="V.*")] a získáme 15897
výskytů.

10 nejfrekventovanějších sloves:

stoupnout	4391
vypnout	1202
zapnout	930
kopnout	617
šlápnout	568
vyhoupnout	457
napnout	373
rozeznout	328
přepnout	308
přehoupnout	305

A. kopl x koptul, přehoupl x přehoupnul, vyhoupl x vyhoupnul, zakopl x zakoptul, šlápl x šlápnul, stoupl x stoupnul, rozepnul o, zapálo o; vypjal, vypjala, vypjalo, vypjali, vypnul, vypnula, vypnulo, vypnuli, vypnuly + napjal, napjala, napjalo, napjali, napjaly, napnul, napnula, napnulo, napnuli, napnuly

B. koptnout, -houpnout, šlápnout, stoupnout; rozepnout (ř, např. svetr) x zapnout; vypnout x napnout

C. vypjal pisa x vypnul motor, napjaly se svaly, šlachy x napnul plachty, tenkou látku

Úkol 14

V korpusu zadáme dotaz $((\text{lemma}=\text{"být"}) \mid (\text{lemma}=\text{"viset"})) [\text{tag}=\text{"D.*"}]\{0,2\}$ $[\text{word}=\text{"vysoko"}]$ a získáme 184 výskyty.

Potom zadáme další dotaz s P-Filtrem, kontext 1. pozice zprava $[\text{tag}=\text{"Z.*"}] \mid [\text{word}=\text{"a"}]$ a získáme nám 82 výskyty.

Poměr mezi slovesy být a viset je 80 : 2.

Repetoár užívaných adverbii: tak, dost, příliš, poněkud, hodně, dostatečně, pěkně, proklatě

Při zadání $((\text{lemma}=\text{"být"}) \mid (\text{lemma}=\text{"ležet"})) [\text{tag}=\text{"D.*"}]\{0,2\}$ $[\text{word}=\text{"hluboko"}]$ získáme 166 výskyty, po provedení P-Filtru získáme pouze 14 výskyty. Repertoár užívaných adverbii: tak, hodně, moc, příliš, důhňvák, nebezpečně

Úkol 15

Nejfrekventovanější adjektiva vyhledáváme dotazem: $((\text{lemma}=\text{"Havlův"} \mid (\text{tag}=\text{"N.*"})) \text{ a poté použijeme frekvenční distribuční lemma.}$

<i>Havlův</i>	<i>hra</i>	64	<i>projev</i>	44	<i>mluvčí</i>	31
<i>Klausův</i>	<i>vláda</i>	387	<i>strana</i>	316	<i>kabinet</i>	299
<i>Zemanův</i>	<i>vláda</i>	291	<i>kabinet</i>	166	<i>strana</i>	114

Další vyhledávání je možné dotazem typu: $((\text{lemma}=\text{"hra"})) \mid (\text{word}=\text{"Havla"}) \& (\text{tag}=\text{"N.*"})\{0,2\}$. Ve všech případech je častější užití přívlastku shodného. Můžeme i srovnávat počet nalezených konkordancí, ale nesmíme zapomenout vyloučit negativními filtry a někdy i ručním tříděním případy např. předložek s genitivem (bez), spojek a typ *na pozvání prezidenta Václava Havla* (teno typ bývá často bez křestního jména nebo funkce).

Úkol 16

V korpusu zadáme dotaz $((\text{lemma}=\text{"dostat"})) \mid (\text{tag}=\text{"Vs.*"}) \& (\text{word}=\text{"*o"}))$ a získáme 264 výskyty. Seznam sloves dostaneme buď pomocí Konkordance – Statistiky – Frekvenční distribuce – lemma na pozici 1 současně s <KWIC, nebo pomocí Obecného třídění, když vynecháme duplicity: (23)

zaplatit, přičat, vynadat, vyplácet, zaplatit, nařezat, přidělit, vyhubovat, napráskat, nasekat, zadat, potvrdit, naservírovat, přislibit, přikázat, vrátit, nařít, nařádkovat, naplácat, našvihlat, uhradit, signalizovat, navázat, nakopat

Další složitější dotazy:

$((\text{lemma}=\text{"dostat"} \mid (\text{lemma}=\text{"být"})) \mid (\text{tag}=\text{"Vs.*"}) \& (\text{word}=\text{"*o"}))$ – 9 výskyty např. *dostal jsem vyhubováno* – slovesa: *zaplatit, nařít, vynadat, nahášet*
 $((\text{lemma}=\text{"mít"} \mid (\text{lemma}=\text{"být"})) \mid (\text{lemma}=\text{"dostat"})) \mid (\text{tag}=\text{"Vs.*"}) \& (\text{word}=\text{"*o"}))$ *měli jsme dostat zapláceno*
měli by dostat přidáno

Úkol 17

Dotaz můžeme zadat pro všechny rody

$((\text{lemma}=\text{"dostat"})) \mid (\text{tag}=\text{"Vs.*"}) \mid ((\text{lemma}=\text{"dostat"})) \mid (\text{lemma}=\text{"být"})) \mid (\text{tag}=\text{"Vs.*"}) \mid ((\text{lemma}=\text{"být"})) \mid (\text{lemma}=\text{"dostat"})) \mid (\text{tag}=\text{"Vs.*"}) \& (\text{word}=\text{"*o"}))$ a získáme 346 výskyty nebo pouze pro rod střední, tj. jako v předcházejícím cvičení $((\text{lemma}=\text{"dostat"})) \mid (\text{tag}=\text{"Vs.*"}) \& (\text{word}=\text{"*o"})) \mid ((\text{lemma}=\text{"dostat"})) \mid (\text{lemma}=\text{"být"})) \mid (\text{tag}=\text{"Vs.*"}) \& (\text{word}=\text{"*o"})) \mid ((\text{lemma}=\text{"být"})) \mid (\text{lemma}=\text{"dostat"}))$

Chceme-li zadat pouze 1. a 2. osobu singuláru a plurálu, můžeme zadat dotaz –

$((\text{lemma}=\text{"dostat"})) \mid (\text{word}=\text{"js.*"}) \mid (\text{tag}=\text{"Vs.*"}) \& (\text{word}=\text{"*o"}))$ nebo $((\text{lemma}=\text{"dostat"})) \mid (\text{word}=\text{"js.*"}) \mid (\text{tag}=\text{"Vs.*"}) \& (\text{word}=\text{"*o"}))$ a v obou případech získáme 8 výskyty.

S neexistujícími slovoslednými variantami nebo variantami okrajovými $(\text{tag}=\text{"Vs.*"}) \mid (\text{lemma}=\text{"být"})) \mid (\text{lemma}=\text{"dostat"})$ – 6 výskyty, z toho 4 chybné – v tomto cvičení nepracujeme.

Úkol 18

Dotaz zadáváme podle vzoru 1a, tj. **řno*, v korpusu SYN2000 musíme odstranit substantiva, avšak i po tomto kroku je třeba projít materiál ručně – z 90 výskyty je pouze 58 slovesných.

Nejčastěji se vyskytují a nejvíce typů sloves najdeme v korpusu SYN2000. Tyto se velmi omezeně vyskytují v BMK (pouze tvary *napěčino, zavřino*). Pojí se nejčastěji se slovesem *být, mít, dostat*. V psaném korpusu jsou nejfrekventovanější tyto tvary: *zavřino, otevřino, odvoziño, vystavřino, nakouřino, nakouřino* atd.

Úkol 19

V korpusu zadáme dotaz $((\text{tag}=\text{"V.*"}) \& (\text{word}=\text{"*řščiou"}))$ a získáme 5144 výskyty.

Na základě Frekvenční distribuce si uděláme seznam tvarů, které zahrnují malé i velké písmeno na začátku (trídění podle lc). U tvarů, které jsou homonymní v singuláru i v plurálu (např. *on váží – oni váží*), je třeba pomocí P-Filtru dostat jen tvary plurálu, tj. dotaz bude $(\text{tag}=\text{"V.*P.*"})$.

tvar na -ou	fr.	tvar na -í	fr.	tvar na -ou	fr.	tvar na -í	fr.
můžou	915	-	-	tlučou	47	-	-
dokážou	795	dokáží	1712	navážou	35	naváží	16
nemůžou	384	-	-	zavážou	33	zaváží	11
nedokážou	325	nedokáží	624	prokážou	30	prokáží	142
píšou	231	píší	914	koušou	28	(kousají)	2
tečou	151	-	-	potečou	28	-	-
lžou	143	-	-	nepišou	27	nepiší	38
ukážou	143	ukáží	263	mažou	26	-	-
vážou	141	váží	219	svážou	25	-	-
utečou	101	-	-	nepláčou	22	nepláčí	3
pláčou	101	pláčí	47	zakážou	21	zakáží	19
pečou	100	-	-	podepíšou	20	podepíší	228
skáčou	99	skáčí	21	nezrnůžou	19	-	-
nelžou	99	-	-	upečou	19	-	-
vlečou	93	-	-	vřžou	18	vřžají	8
pomůžou	77	-	-	nepomůžou	17	-	-
selžou	71	-	-	podřežou	16	-	-
řezou	65	(řezají)	5	neukážou	16	neukáží	-
kloužou	53	(klouzají)	20	neutečou	15	-	-
napišou	48	napiší	99	rozutečou	15	-	-

Tyto tvary se týkají sloves 1. třídy vzor „maže“ a „peče“. U vzoru „peče“ variantní koncovky nejsou, u vzoru „maže“ se mohou objevit koncovky typu *dokážou* x *dokáží* (přechod ke 4. třídě) nebo *koušou* x *kousají* (přechod k 5. třídě).

Úkol 20

V korpusu zadáme dotaz [(tag="V.*") & (word="*ejí")] a získáme 45305 výskytů.

Na základě frekvenční distribuce vytvoříme seznam, z něhož vybereme následující slovesa:

muset, nabízet, -cházet, vytvářet, -dílet, vracet, -nášet, ztrácet, pokoušet, souviseť, -víjet, obracet, -rážet, -provázet, zkoušet, přemýšlet, sázet, střílet, zabíjet, -vážet

Slovesa přát (si) a přejít se mají hlásku -e- v kořeni, nejde tedy o slovesa s koncovkou -ejí.

D.

Úkol 1

V korpusu zadáme dotaz **bačkoru** a získáme 24 výskytů.

Konkrétní věc: 1 – *soutěž o Zlatou bačkoru*

Přenesené užití slova: 1 – označení věci, zhotovíme „bačkoru“ ze slámy, 3 – označení osoby, má doma bačkoru, má doma „bačkoru“, proměna krále v bačkoru

Abstraktní užití slova - 0

Frazeologické výrazy nebo součásti frazémů 19 – *být na bačkoru, stát za bačkoru, stát za starou bačkoru, starou bačkoru*

Úkol 2

V korpusu zadáme dotaz **štěpnic.*** a získáme 16 dokladů.

A. v próze

B. zdroj pi180 je kniha Bohumila Hrabala „Růžový kavalír“

C. viz SsJČ

štěpnice, -e ž. ovocný sad, zvl. jablonoň

D. nestací

Úkol 3

V korpusu zadáme dotaz **volňás.*** a získáme 8 výskytů.

SsJČ

volňásek, -ska (4. j. těž -ska, 6. mn. -scích) (slang. volňas, -a) m. hovor. volňá, *bezplatná vstupenka*

Úkol 4

V korpusu zadáme dotaz [Mm]mercedes.* a získáme 2056 výskytů.

Na základě funkce frekvenční distribuce dostaneme tento seznam:

Mercedes	1149
Mercedes	219
mercedesu	216
Mercedesu	150
mercedesů	80
mercedesy	70

mercedesem	52
Mercedesem	20
mercedesesch	16
Mercedesy	16
MercedesBenz	13
mercedesům	12
mercedesa	7
Mercedesů	7
mercedesce	4
mercedesku	4
mercedesky	4
mercedeska	3
Mercedesa	3
Mercedesch	2
Mercedeska	2
mercedesizaci	1
mercedesové	1
mercedesu	1
mercedesaj	1
mercedesový	1
mercedesek	1
mercedesovského	1

A. a. např. *automobil značky Mercedes, Mercedes Benz AG*
 název firmy – Mercedes – Benz ve Stuttgartu

b. např. *zbrusu nový černý mercedes*

c. např. *v otevřené černé mercedesce, béžová mercedeska, vedle mýl učedník mercedesku, dveře mercedesky, narazí do některého z řady mercedesů-benzů parkujících před německým velvyslanectvím*

B. *mercedeska*

Tyto tvary vidíme ve výše uvedené tabulce

C. *medoun, medák*

D. Zadáme-li dotaz **med***, získáme 4 výskyty

Tato slova s velkým písmenem psát nebudeme, protože se nejedná o vlastní názvy.

Úkol 5

V korpusu zadáme dotaz **[lemma="investovat"]** a získáme 5549 výskytů.

Pomocí P-Filtru si potom můžeme ponechat jen kontexty s předložkou *do*, tj. **[lemma="do"]**, a to v rozsahu -3, 3 a získáme 2692 výskyty. Vyuzijeme statistickou funkci Kolokace – třídění podle absolutní frekvence ve stejném rozsahu a získáme tento seznam slov, která označují oblasti, do nichž se nejčastěji investuje:

rozvoji, technologie, modernizace, projekt, oprava, výstavba, výroba, rekonstrukce, vzdělání, obnova, stavba, zařízení, výzkum, infrastruktura, kampaň, reklama.

Při opačném postupu – zadání N-Filtru do základní vyhledávky a odstranění kontextů s předložkou *do* nám zůstane 2857 výskytů. Stejným postupem – využitím statistické funkce Kolokace – získáme seznam slov, která označují převážně finanční částky a druhy měny:

milión, miliarda, koruna, peníze, dolar, Kč, prostředek, částka, USD, tisíc, DEM, sto.

Úkol 6

V korpusu zadáme dotaz **[lemma="třešnička"]** a získáme 156 výskytů.

A. Nefrekventovanějším tvarem je tvar **třešničkou**, tj. ve frazému **třešničkou na dortu něčeho je ...**

B. **třešnička na šlehačce** např. **zlatý hřeb**

C. V kontextech se objeví s velkým písmenem **Třešnička** jako jméno.

Úkol 7

V korpusu zadáme dotaz **[lemma="souvislost"]** a získáme 17534 výskyty.

Pomocí funkce Kolokace vytvoříme tabulku, ve které ponecháme jen 3 nejfrekventovanější slovesa.

lemma	Ml-score	T-score	Rel. f [%]	Abs. F
<i>mít</i>	0.3065	2.043	0.01793	114
<i>být</i>	-3.668	-73.12	0.001141	39
<i>hledat</i>	2.831	4.122	0.1032	23

Úkol 8

A. na základě korpusu (ducha, fánatizii, hudební sluch, mateřštinu, intelekt, inteligenci, jazyk, myšlenky, názor, rozum)

B.

SSJČ

tříbiti ned. (3. mn. -í, rozk. tříbi) (co) **1.** zbavovat *chyb, kazů, uvádět do dokonalejšího stavu; zlepšovat, pěstovat 3, zušlechťovat, zdokonalovat*; t. sloh; t. si usudek, názory; t. pamět *bysřiti*; tříbení ducha, duševních vlastností; tříbit vkus *zjemňovat*; zkušenostmi se tříbí charakter **2.** ponek. zast. a kniž. zbavovat *nepořádných n. nežádoucích částí, předmetů; čísti 1*; t. zmo (Hol.); tříbi si (pták) *peří* (Vrchl.); rozmachovali se (siláci) širočinou a tříbili les (Pujm.) *myšlí*; kožel. tříbení *čištění vlasu kožíšm; tříbiti se* ned. zbavovat se *chyb, kazů, přecházet do dokonalejšího stavu; zušlechťovat se, zdokonalovat se*; myšlenka se *tříbila* a prohlubovala • předp. pro-, pro-se, pře-, u-, vy-se

C. viz výše uvedená slovníková definice – původně označovalo mycení lesa za účelem získání zemědělské půdy

Úkol 9

V korpusu zadáme dotaz [lemma="odúmrť"] a získáme 16 výskytů. V korpusu najdeme přímo vysvětlení: *odúmrť je právnícký výraz pro situaci, kdy občan nezanechá dědice, takže dědi stát. V tomto smyslu se jedná o termín, dále se vyskytuje v medicíně ve významu „odumření“.* Objevuje se však i v beletrii a dokonce v poezii, kde získává určitou poetiku, zde se většinou nachází ve spojení s genitivním substantivem *odúmrť lásky*.

Úkol 10

V korpusu zadáme dotaz [lemma="akce"] a získáme 27844 výskytů. Potom pracujeme s náhodným výběrem 100–500 výskytů (menu Konkordance – Redukce – Náhodných výskytů).
 A. tomuto zaviřnému stoupenci amerického imperialismu je něco jiného než zá-
 měná politická <akce> těch sil v naší zemi, které bojují proti míru a společen-
 skému pokroku v naší části
 B. Až pětářicet hodin byl v <akci> hasič, kterého od ohně odvezli rovnou do
 mostecké nemocnice.
 C. Představení kouzel a magie Černého divadla Jiřího Šimce je součástí kulturní
 <akce> Rudolf II. a Praha.
 Častěji v mluvených korpusích najdeme význam – posezení, návštěva restaura-
 ce s přáteli, zábava. Hurá akce – narychlo a zbrkle konaná činnost, ne dost při-
 pravená.

Úkol 11

V korpusu zadáme [lemma="pozor"] a získáme 5422 výskytů. Potom pomocí funkce N-Filtr s kontextem -1, 0 zadáme dotaz [tag="Z.*"] a zů-
 stanou 3883 výskytů. Uložíme si tabulku, která vznikne na základě statistické funkce Kolokace:

lemma	Ml-score	T-score	Rel. f [%]	Abs. F
dávat	10.79	37.25	5.68	1389
se	3.958	31.73	0.04991	1150
dát	8.294	30.23	1.008	920
na	3.092	18.85	0.02739	456
musel	6.209	19.51	0.2376	391
byl	1.361	10.26	0.008251	282
mít	3.585	14.35	0.03854	245
velký	4.898	12.75	0.09572	174
v	1.016	6.474	0.006495	164
dobrý	5.232	10.62	0.1207	119
do	1.703	5.585	0.01045	65
stát	2.916	5.819	0.02424	45

postavit	6.212	6.469	0.238	43
nutný	5.237	5.332	0.1211	30
před	2.622	4.432	0.01977	28
při	2.26	4.035	0.01539	26
bedlivý	12.09	4.689	14.01	22
hodný	7.252	4.33	0.4896	19

dát/dávat (na někoho, na sebe) pozor, mít se (před někým, při něčem) na pozoru, být v pozoru, dát/dávat dobrý/velký/bedlivý pozor, stát v pozoru, postavit se do pozoru, museti dát/dávat pozor

Úkol 12

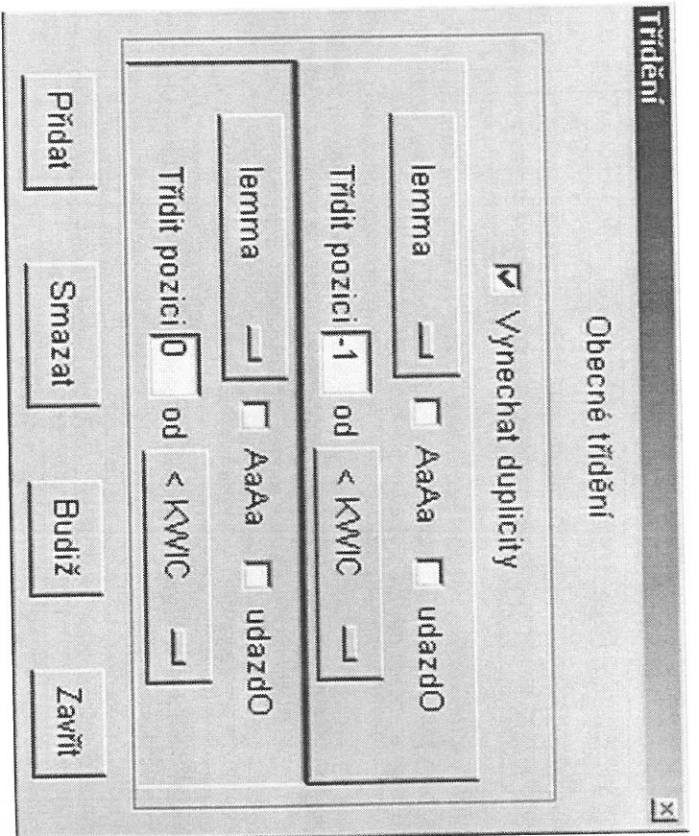
V korpusu zadáme dotaz [lemma="řinout"] a získáme 147 výskytů. A. Na základě funkce Kolokace (menu Konkordance – Statistika – Kolokace), rozsah –3, 3 dostaneme následující tabulku:

Lemma	Ml-score	T-score	Rel. f [%]	Abs. f
slza	12.44	4.898	0.677	24
krev	10.59	4.356	0.1879	19
tvář	9.336	3.994	0.07859	16
pot	12.72	3.464	0.8214	12
voda	7.422	2.983	0.02085	9
proud	9.591	2.996	0.09374	9
ústa	9.585	2.642	0.0934	7
oko	6.663	2.214	0.01232	5
slovo	5.504	1.694	0.005516	3
ruka	5.529	1.695	0.005613	3
čelo	7.469	1.722	0.02154	3

krev po ruce, pot po čele, slzy z očí, po tváři, proud vody, popř. slova z úst
 B. řinout se ned. slabým proudem odněkud vytékat: z očí se řinou slzy, pot se ři-
 ne po čele, krev se řine po ruce

Úkol 13

V korpusu zadáme dotaz [lemma="vůně"] | (lemma="zápach") | [lemma="zvuk"] | (lemma="melodie") a získáme 12697 výskytů. řinout se – vycházet, šířit se z nějakého zdroje
 Po zadání dotazu [tag="V.*"] při P-Filtru v kontextu -1, 0 získáme 1444 výskytů. Potom zadáme Obecně třídění pro pozici -1 a 0:



Pomocí funkce Kolokace (menu Konkordance – Statistika – Kolokace) získáme podle absolutní frekvence 3 slovesa, která se pojí se 4 substantivy:

Lemma	MI-score	T-score	Rel. f [%]	Abs. f
<i>být</i>	-2.145	-6.848	0.000117	4
<i>linout (se)</i>	8.834	1.996	0.2363	4
<i>cítit</i>	5.138	1.943	0.01823	4

Plnovyznamovými slovesy, která se pojí se všemi substantivy, jsou *linout se* a *cítit*.

Úkol 14

V korpusu zadáme dotaz [lemma="bujný"] a získáme 413 výskytů. Na základě funkce Kolokace (menu Konkordance – Statistika – Kolokace) seříděné podle absolutní frekvence a prvního slova zprava:

lemma	MI-score	T-score	Rel. f [%]	Abs. f.
<i>fantazie</i>	12.49	7.07	1.959	50
<i>vegetace</i>	12.94	4.795	2.681	23

<i>popisí</i>	14.07	4	5.861	16
<i>růst</i>	7.215	3.847	0.05074	15
<i>veselí</i>	12.24	2.999	1.657	9
<i>Představitost</i>	11.89	2.999	1.295	9
<i>zeleně</i>	10.73	2.827	0.5789	8
<i>kšice</i>	13.87	2.828	5.128	8
<i>porost</i>	10.15	2.826	0.3868	8
<i>tvář</i>	7.784	2.634	0.07528	7
<i>kůň</i>	6.965	2.218	0.04266	5
<i>oř</i>	13.61	2.236	4.274	5
<i>zahradá</i>	7.043	2.219	0.04504	5
<i>vzrůst</i>	7.174	1.986	0.04933	4
<i>představa</i>	6.185	1.973	0.02486	4
<i>křivka</i>	9.458	1.997	0.2402	4
<i>hřívá</i>	11.98	2	1.384	4
<i>život</i>	3.482	1.577	0.003817	3
<i>zábava</i>	7.783	1.724	0.07523	3
<i>oslava</i>	7.213	1.72	0.05067	3

Významově se týká všeho, co velmi rychle a hustě roste – *zahradá, zeleně, hřívá, kšice*, popř. toho, co svým vzrůstem vztahuje – *bujně popisí*.
Frazémy: *bujná fantazie/představitost, bujný oř/kůň*

U slov *veselí, život, zábava, oslava* jde o kontaminaci se slovem *bujný*.

Potom v korpusu zadáme dotaz [lemma="divoký"] a získáme 4269 výskytů. Postupujeme stejně jako v předcházejícím případě.

lemma	MI-score	T-score	Rel. f [%]	Abs. f.
<i>západ</i>	9.307	15.08	2.236	228
<i>karta</i>	9.299	14.08	2.224	199
<i>voda</i>	6.876	13.27	0.4147	179
<i>zvíře</i>	8.201	10.78	1.039	117
<i>obzor</i>	9.845	8.822	3.247	78
<i>příroda</i>	7.29	8.547	0.5526	74
<i>kachna</i>	11.07	8.598	7.566	74
<i>zvěř</i>	8.499	8.284	1.277	69
<i>prase</i>	9.883	7.866	3.333	62
<i>králík</i>	9.758	7.541	3.056	57
<i>řeka</i>	7.253	7.368	0.5385	55
<i>sidce</i>	6.698	7.278	0.3665	54
<i>husa</i>	10.39	7.066	4.744	50

svět	4.021	6.569	0.05731	49
kůň	6.631	6.338	0.3498	41
východ	6.959	6.274	0.4394	40
jízda	6.635	6.022	0.3508	37
odsun	9.553	5.823	2.652	34
privatizace	5.866	5.646	0.206	33
stvoření	9.49	5.649	2.54	32
kočka	7.98	5.456	0.8915	30
pes	5.882	5.108	0.2082	27
sen	6.193	5.125	0.2583	27
skládká	8.467	5.085	1.249	26
druh	4.558	4.592	0.0832	23
večerek	8.556	4.678	1.329	22
přestřelka	9.2	4.575	2.077	21
výsledek	3.654	4.219	0.04446	21
kapitalizmus	8.646	4.461	1.414	20
víno	6.044	4.404	0.233	20
kanec	11.58	4.471	10.81	20
anděl	7.323	4.332	0.5655	19
pivo	5.775	4.165	0.1933	18
džungle	8.603	4.232	1.373	18
banda	9.193	3.993	2.067	16
populace	6.75	3.963	0.38	16
ovce	7.793	3.856	0.7833	15
šelma	8.943	3.865	1.738	15
tanec	6.493	3.7	0.3181	14
labuť	9.133	3.599	1.982	13
rozpad	6.712	3.431	0.37	12
honička	9.793	3.46	3.133	12
rok	-0.2595	-0.6825	0.00295	12
krocán	10.51	3.314	5.164	11
plák	5.895	3.261	0.21	11
představa	4.275	3.145	0.06835	11
svině	8.975	3.156	1.776	10
noc	3.494	2.882	0.03977	10
včela	8.345	3.153	1.148	10
nájezd	8.26	3.152	1.082	10

Užívá se o zvířatech a rostlinách žijících ve volné přírodě – např. divoké víno, divoká kočka, o situacích, které zahrnují rychlý pohyb – divoká honička, divoká voda, o akcích a činnostech, které jsou nepovolené – divoká skládka.

Frazém: Divoký Západ
Termin: divoká karta

■ Úkol 15

V korpusu zadáme dotaz [lemma="smysl"] | [lemma="význam"] a získáme 37375 výskytů. Po provedení P-Filtrování zůstane 2354 výskytů. Nejtypičtější užití:

smysl: v pravém slova smyslu, v nejlépeším slova smyslu, v nejšíším smyslu slova apod.

význam: vyložit význam slova, nářeční význam slova „bato“ apod.

■ Úkol 16

V korpusu zadáme dotaz [lemma="přehřel"] | [lemma="přehřele"] nebo [lemma="přehřel[el]"] a získáme 115 výskytů.

Obě varianty jsou rodu ženského. Z korpusových dokladů je zřejmé, že jde o jinou hláskovou variantu, která se v převážně většině používá ve významu většího množství až přemíry něčeho, v ojedinělém případě najdeme doklad na užití ve významu malého množství „hrstky“: *Kapele se až do konce vystoupení nepodařilo přesvědčit víc než tu zmíněnou přehřel opravdových nadšenců, jimž neva-
díl zvuk ani nic jiného (časopis Rock a pop).*

Nejvyššížnějším synonymem je asi přemíra.

Slovník spisovné češtiny uvádí: *přehřele ž.*, množství, které se nabere do obou hrstí zároveň. V korpusu najdeme i článek z Lidových novin, který se zabývá vý-
kladem tohoto slova a kde si můžeme ujasnit např. etymologii.

■ Úkol 17

V korpusu zadáme dotaz [lemma="facka"] a získáme 651 výskytů. Potom pomocí funkce Frekvenční distribuce (fc) vytvoříme tabulku:

facku	233
fack	140
facka	133
facky	88
fackou	18
fackami	16
facce	16
fackách	4
fackama	2
fackám	1

Typické doklady:

David šel ke mně a dal mi <facku>.

Dal jsem mu pár <facek> a vyhodil ho z bytu.
Dyť je to jasný jak <facek>.

Úkol 18

V korpusu zadáme dotazy **zateskal rukama, zamrkala očima, zateskal, zamrkala**

Ve většině textů je vypuštění možné. Je to spíše otázka stylu, zda autor substantivum použije, či nepoužije.

Poměry jsou **zateskal : zateskal rukama = 59 : 4, zamrkala : zamrkala očima = 49 : 3** (ovšem uvnitř skupiny zamrkala je tu také 1x zamrkala dlouhými řasami a 1x zamrkala víčky).

Úkol 19

V korpusu zadáme dotaz **[(lemma="široko") | (lemma="široce")]** a získáme 1616 výskytů.

Za adverbium **široko** stojí nejčastěji adverbium **daleko** (213x), tj. jde o frazém. Za adverbium **široce** stojí nejčastěji adjektivum, potom je nejčastější spojení **široce pojatý** (62x), nebo sloveso, v tom případě je nejčastější **široce využívat** (21x).

Poté v korpusu zadáme dotaz **[(lemma="dlouho") | (lemma="dlouze")] [pos="V"]** a získáme 8923 výskytů.

Nejfrekventovanější různá spojení jsou **dlouho trvat** (390x) a **dlouze se zadávat** (16x), společně jsou např. **dlouho/dlouze hledat, dlouho/dlouze mččet** apod. **Dlouho** zpravidla znamená "mnohokrát po delší dobu", **dlouze** „dlouhým pohybem, jednorázově“.

Úkol 20

V korpusu zadáme dotaz **[(lemma="závislý") | (lemma="odvislý")] [(lemma="na") | (lemma="od")]** a získáme 3185 výskytů.

Potom na základě funkce Frekvenční distribuce (lemma na pozicích 0 a 1) dostaneme následující tabulku:

závislý	na	3066
odvislý	od	101
závislý	od	17
odvislý	na	1

Jediný výskyt spojení **odvislý na**, tj. aplikace jednotlivých klimatažančních jednotek a systémů je **odvislá na velikosti objektů a na specifických požadavcích jejich využívání a provozu** ukazuje na to, že nejde o obvyklý úzus.

Z počtu 17 dokladů spojení **závislý od** je třeba odečíst dva výskytů, kde nejde o valenci adjektiv:

1) Hoffmannovy horečnaté víze ukazují jeden ze způsobů, jak pruské a obecně severogermánské radostné skutečnosti uniknout a množství **drogově závislých od Amsterdamu přes Brémy po Hamburk** představuje jeho dnešní, chemii ovlivněné bratry.

2) Afghánistán, **nezávislý od roku 1747**, se zmiňá ve víru vnitřních konfliktů a hospodářského rozvratu.

E.

Úkol 1

V korpusu zadáme dotaz `((([lemma="na"] [lc="začátku"] | ([lemma="na"] [lc="konci"]))) ([lemma="*n"] & (pos="N")))` a získáme 223 výskyty. Potom pomocí Frekvenční distribuce se zadáním

Frekvenční distribuce

Zadání frekvenční distribuce

Limit:

Atribut: Ignorovat velikost

Pozice od < KWIC

Atribut: Ignorovat velikost

Pozice od < KWIC

Atribut: Ignorovat velikost

Pozice od < KWIC

dostaneme tabulku, v níž si můžeme seřadit abecedně substantiva.

S kolokacemi **na začátku** i **na konci** se spojují substantiva (s frekvencí vyšší než 2): *jednání, natáčení, obchodování, představení, utkání, vedení, vyprávění*. S kolokací **na začátku** se spojují např. substantiva označující nějakou akci, např. *povstání, přelčení, vyučování, zasedání*, nebo činnost, např. *podnikání, promítání, psaní, vaření*, popř. časové tvrní určitého procesu, slov. *na začátku prodloužení*.

S kolokací **na konci** se spojují substantiva označující dovršení nějakého procesu, např. *poznání, akci, např. sezení* nebo činnost, např. *pečení*.

Úkol 2

V korpusu zadáme dotaz `([lemma="třešnička"] [lemma="na"] [lemma="dortu"])` a získáme 32 výskyty. Funkce frekvenční distribuce (lc) nám pomůže zjistit jednotlivé zastoupení tvarů.

třešničkou	na	dortu	20
třešnička	na	dortu	6
třešničku	na	dortu	4
třešnička	na	dort	1
třešničky	na	dortu	1

Význam: efektivní (i když někdy méně důležitý) závěr, vyvrcholení něčeho, a to v kladném i záporném slova smyslu

Úkol 3

V korpusu zadáme dotaz v **průběhu** (při použití atributu word) a získáme 6616 výskyty.

Potom postupujeme jako v úkolu C3.

lemma	Ml-score	T-score	Rel. f [%]	Abs. f.
rok	6.095	38.44	0.3741	1522
den	5.786	17.92	0.3019	333
týden	6.417	15.02	0.4675	231
měsíc	6.433	14.39	0.4727	212
století	6.794	12.73	0.6074	165
období	6.212	10.01	0.4055	103
doba	3.484	7.509	0.06123	68
hodina	3.537	6.397	0.06354	49
čtvrteř	7.342	6.126	0.8879	38
leden	5.149	5.831	0.1942	36
podzim	6.179	5.75	0.3964	34
srpen	5.348	5.604	0.2229	33

pololetí	7.569	5.714	1.039	33
desetiletí	7.114	5.703	0.7581	33
květen	4.913	5.295	0.1649	30
říjen	4.893	5.021	0.1626	27
duben	5.004	5.034	0.1756	27

Seznam substantivv označujících časový úsek se u této výeslovné předložky čas-
tečně liší od předložky **během** – přibylly názvy měsíců, tj. *srpen, květen, říjen,*
duben.

Úkol 4

A. V korpusu zadáme dotaz [lemma="dát"] [0,3] [lemma="koš"] a získáme
57 výskytů.

26x dát koš – sportovní slang, košíková

24x dát někomu košem – frazém

6x dát něco do koše

1x dá se rozehrat hrudní koš

B. V korpusu zadáme dotaz [lemma="dostat"] [0,3] [lemma="koš"] a získá-
me 20 výskytů.

5x dostat koš – sportovní slang, košíková

2x dostat se do druhého koše – sportovní slang, fotbal

10x dostat od někoho košem – frazém

2x dostat dářkový koš

1x dostal jsem na ruku koš

Úkol 5

V korpusu zadáme dotaz [lemma="železný"] [lemma="oř"] | [lemma="kůň"]
a získáme 16 výskytů

železný oř = vlak, popř. automobil, sněžný skútr

železný kůň = malý pásový dopravník určený pro přibližování dřeva

Jedná se o metatoru pro stroj nahrazující zvíře používané pro cestování či práce
v lese.

Úkol 6

V korpusu zadáme dotaz [lemma="na"] [lemma="pokraj"] a získáme 796 vý-
skytů.

A. Substantiva následující za touto kolokací:

krach	41
zhroucení	32
zkáza	28
bankrot	27
smrt	21

propast	18
síla	14
kolaps	14
válka	14
zoufalství	13
zánik	13
les	13
rozpad	13
zájem	13
vylhnutí	12
bída	12
záhuba	11

Významová skupina – „zánik, utrpení“

Jediné konkrétní substantivum v tomto výčtu je *les*.

B. Slovesa předcházející této kolokací:

Lemma	Ml-score	T-score	Rel. f [%]	Abs. f	Pád
balancovat	12.42	2.999	3.614	9	6. pád
potáčet (se)	11.11	2.448	1.46	6	6. pád
očtat (se)	10.96	3.604	1.314	13	6. pád
ochnout (se)	10.65	7.344	1.056	54	6. pád
přivést	10.41	7.136	0.8966	51	4. pád
dohnat	9.216	1.997	0.3914	4	4. pád
přivádět	8.468	1.727	0.2331	3	4. pád
dověst	7.498	3.298	0.119	11	4. pád
nacházet (se)	6.959	2.624	0.08189	7	6. pád
pohybovat (se)	6.95	3.137	0.0814	10	6. pád
žít	6.213	4.068	0.04883	17	6. pád
stanout	5.95	2.2	0.0407	5	6. pád
stát	5.543	7.388	0.0307	57	6. pád
dostat (se)	5.124	3.885	0.02296	16	4. pád
zastavit (se)	5.122	1.682	0.02293	3	6. pád

Úkol 7

V korpusu zadáme dotaz [lemma="šedý"] [lemma="zóna"] a získáme 92 výskytů.

A. společnost, ekonomie

B. nepatřící zcela ani k jednomu vyhraněnému směru nebo skupině; v době tota-
lity většína společnosti, která se pasivitou snažila o přežití; nepatřila ani ke ko-
munistům, ani k disentu; dnes se užívá pro ne zcela legální obchod.

125

C. Frazémy: *šedá ekonomika*; *šedá eminence*; Termíny: *šedá kůra mozkoví*; *šedá litina*; *šedý* (zrnitý)vápenec; *šedý zákal*; Nejčastější běžná spojení: *šedé vlasy*, *šedé oči*

Úkol 8

V korpusu zadáme dotaz ([lemma="kočka"] [1, 1] [lemma="myš"]) a získáme 91 výskyt.
Pomocí funkce Frekvenční distribuce získáme poměr:
kočka s myš 56, kočka a myš 23
Frazeologických užití je celkem 70.

Úkol 9

V korpusu zadáme postupně dotazy:
([lemma="tmavě"] [pos="A"] [lemma="barva"]) – 49 výskytů
([lemma="světle"] [pos="A"] [lemma="barva"]) – 30 výskytů
K určení frekvence můžeme použít frekvenční distribuci. Frekventovanější je spojení s adverbem *tmavě*. Podle frekvence názvy barev následující po *tmavě*: *červená*, *modrá*, *hnědá*, *zelená*, *rudá*, *řalová*, *černá*, *víňová*, *tyrkysová*, *zelenošedá*, *šedá*, *šedozelená*. Po *světle* následuje: *modrá*, *šedá*, *zelená*, *žlutá*, *ružová*, *muškátová*, *hnědá*, *okrová*, *červená*, *masová*, *azurová*.

Úkol 10

V korpusu zadáme dotaz ([lemma="být"] [lemma="balk"]]) – a získáme 89 výskytů.
A. Z tohoto počtu je frazém *být v balku* doložen pouze 5x:
Až <budeme v balku>, vytáhnou tě vodsad...
...někteří lidé *<jsou v balku>* a přitom jsou v obraze.
...závidí těm, co už *<jsou v balku>*. A inteligence opět skomírá...
...davy lidí, všichni *<jsou v balku>*, přesto závidí.
...*<jsme v balku>* („napakovaný“)...
Podmět je životný. Slovo *balk* je v 6. pádě.
B. V korpusu zadáme dotaz ([lemma="být"] [1, 3] [word="balku"]) a získáme 97 výskytů. Tentokrát je zachyceno 13 frazeologických užití:
Vanda *<je pravděpodobně ve věším balku>* než Neil!
...mají utkvělou představu, že *<jsm v bůhvjakém balku>*.
Prost, ty dva a Brenda *<jsou teď zjevně v balku>*...
Ano, Yankee *<byl pořád ještě v balku>*...
Až *<budeme v balku>*, vytáhnou tě vodsad...
...a *<byli jsme zas v balku>*.
...*<jsme v balku>* („napakovaný“)...
...všichni *<jsou v balku>*, přesto závidí.
...hošku, už tě znají, ač *<je tvůj pán v balku>*, ty jseš v háji.

Vy teď jako primář *<budete ve velikým balku>*.

...se rozkřiklo, že *<je „v balku>*“...

...někteří lidé *<jsou v balku>* a přitom jsou v obraze.

...závidí těm, co už *<jsou v balku>*. A inteligence opět skomírá...

Částmi tohoto frazému jsou sloveso *být*, předložka *v* a tvar *balku*.

Úkol 11

V korpusu zadáme dotaz ([lemma="dát"] [1, 2] [lemma="dohromady"]) a získáme 137 výskytů.

Frazémy: *dát se dohromady*, *dát hlavy dohromady*, *dát co dohromady* (peníze, části, být), *dát někomu dohromady*, *dát si všechno dohromady* aj.

Úkol 12

V korpusu zadáme dotaz ([lemma="oblast"] [1, 3] [lemma="živoť"]) a získáme 37 výskytů.
Nejčastějších pět „oblastí života“: *společenského*, *veřejného*, *lidského*, *mého*, *svého*

Úkol 13

V korpusu postupně zadáme dotazy:
([lemma="letní"] [1, 0, 4] [lemma="dub"]) | ([lemma="dub"] [1, 0, 4] [lemma="letní"]) – 17 výskytů
([lemma="letní"] [1, 0, 4] [lemma="den"]) | ([lemma="den"] [1, 0, 4] [lemma="letní"]) – 374 výskytů
Dub letní je botanický (vědecký) název rostliny a jeho rozvíjení není možné. Spojení *letní dub* je aktualizace a jde v ní o humorou interpretaci termínů (viz korpus). *Letní den* je běžný typ rozvíti substantiva *den* a může být tedy dále rozvíjeno *letní (parní, sváteční, sluneční, antarktický, psí, srpnový) den*, *letní, bytí ké syntaxi*, takže *letní* obvykle rozvíjí jiné substantivum, např. *den letního slunovratu*.

Úkol 14

V korpusu zadáme dotaz ([lemma="nechat"] [1, 0, 3] [lemma="v"] [word="štychu"]) | ([lemma="v"] [word="štychu"] [1, 0, 3] [lemma="nechat"]) a získáme 42 výskytů.
V levém kontextu jsou nejčastější zájmena: *mě*, *je*, *nás*. Částečně lze použít frekvenční distribuci pro vyhodnocení. Nejčastější je 3. os. sg. maskulina, zajímavé v negaci. Podobný frazém *nechat někoho na holičkách*. V inverzním slovosledu se tento frazém vyskytuje 2x.

■ Úkol 15

V korpusu zadáme dotaz ((lemma="hledat"1 [1{0,3} lemma="souvislost"1]) | (lemma="souvislost"1 [1{0,3} lemma="hledat"1])) a získáme 71 výskytů.

1. HLEDAT SOUVISLOST/SOUVISLOSTI = sloveso + substantivum aby Václav mohl <hledat co nejvíc souvislosti>. Vašek měl naproti tomu <hledá jasně logické souvislosti> příčiny, že striktně <hledá kauzální souvislosti> " uvnitř"

Reziignuji na úlohu <hledat mezi jednotlivými informacemi souvislosti>.

nejlepších vodách. <Nehledejte mezi nimi hlubší souvislosti>, někdo se

Tady někde <hledejte možné souvislosti>, přičíst k

nás někdy obřížné <hledat návaznosti a souvislosti>. Jak se

Landsteinerovy výsledky, <hledal nějakou souvislost> mezi různými druhy

materiály a nyní <hledám nitky a souvislosti> " řekl

Nesnažme se <hledat objektivní souvislosti> mezi dvěma díly

ústav pomoci. <Hledat paralely a souvislosti> mezi českou,

tam chodit a <hledat podstatu a souvislosti>, které by

matematickou povahu, <hledají příčinné souvislosti> jevů, snaží

historikové, kteří <hledají přímou souvislosti> mezi hláskovým složením

zvyky analyzujeme a <hledáme příslušné souvislosti>, narazíme na

na zkoušku, <hledají se rodové souvislosti> (lampi,

bývalo, ale <hledá se souvislosti> mezi barvou a

klubu. „ <Hledat souvislosti> mezi vraždou a

pojmenovávat problémy, <hledat souvislosti> a příčinné vazby

v tomto světě <hledat srozumitelné souvislosti>? Vždyť slovo

odborníka, který <hledá širší souvislosti> své orientace.

– kdyby později <hledal tu souvislosti>, a on

spouští statistik a <hledal v jejich výsledcích souvislosti>. Na hledání

výrobce whisky. <Nehledám však souvislosti>. Producent pravé

vstří. Poté <hledá významové souvislosti> s erotickými symboly

hudby inspiraci, <hledá vzájemné souvislosti>, snaží se

a prošetření skutečnosti <hledá vztahy a vzájemné souvislosti>. Práce studenta

pracovních důvodů. <Hledat za tím souvislosti> se zakázkou na

vše končí. <Nehledáme žádné podezřelé souvislosti> a nepodkládáme si

2. SOUVISLOST/SOUVISLOSTI HLEDAT = substantivum + sloveso

rok a všechny <souvislosti bych musel hledat> v papírech,

finančních službách? <Souvislosti lze hledat> v samotném bankovním

literární a kulturní <souvislosti můžeme hledat> už v brněnském

3. HLEDAT NĚCO V SOUVISLOSTI S NĚČÍM/ V SOUVISLOSTI S NĚČÍM NĚCO

HLEDAT = sloveso + vičeslovná předložka/vičeslovná předložka + sloveso

ho německá policie <hledala v souvislosti> s vyzrazením jaderného

V této <souvislosti hledají> i ukradené vozídko

Čiňanka, kterou <hledala policie v souvislosti> s nedělní vraždou

V <souvislosti se sklonem hledat> poezii byla také

4. HLEDAT NĚCO V SOUVISLOSTECH (MEZI SOUVISLOSTMI)/ V SOUVISLO-
STANTIVUM + sloveso
o zákazníkovi. <Hledat v těchto souvislostech> argumenty výhradně proti
posleze mezi těmito <souvislostmi hledat> obecné zákonitosti a
I v těchto <souvislostech kniha hledá> především vzájemné vlivy

■ Úkol 16

V korpusu zadáme dotaz ((lemma="zadat"1 [1{0,3} lemma="práce"1]) | (lem-
ma="práce"1 [1{0,3} lemma="zadat"1])).

Celkový výskyt je 52; u 6 výskytů naše dvě slova tvoří kolokaci – jsou rozděle-
na do různých vět nebo se sloveso zadat váže k jinému objektu, ů. např. zadáme
formou práce. Dále je třeba odečíst dva výskytů nezadala si s cenami prací, kde
jde o frazém nezadat si s něčím.

v činném rodě je 37 vět
v trpném rodě je 7 vět

činný rod

<práci zadal>	<zadá vědeckou práci>
<zadal dlaždičskou práci>	<práci o své újmě zadá>
<zadal Mozartovi nějakou práci>	<zadat diplomovou práci>
<zadal mu zkušební práci>	<práci zadat>
<zadali práci>	<zadat práci>
<zadali práce>	<zadat tuto práci>
<práci zadáte>	<zadat tuto práci>
<zadáme práci>	<ZADÁ práci>
<zadal práci>	<práce zadat>
<práci zadali>	<práci zadat>
<zadám práci>	<práci musíme někomu zadat>
<zadám práci>	<práci zadala>
<práce jsme zadali>	<práce Rekulivacím zadala>
<práci zadat>	<práci zadat>
<zadala v roce 1994 práce>	<Zadal mi práci>
<zadá jím práci>	<zadají práci>
<nezadalo práci>	<zadalo takřka obdobně práce>
<práci zadali>	<zadá práci>
<zadá práci>	

trpný rod

<zadán i objem prací>	<práce, která byla zadána>
<práce zadány>	<zadány práce>
<práce zadány>	<práci bude zadáno>
Se <práce zadala>	

Poměr je 1/10 trpné : 9/10 činné.

Úkol 17

V korpusu zadáme dotaz (([lemma="dát"] [word="se"] [word="do"] | ([lemma="do"] [word="dát"] [lemma="dát"]]) a získáme 691 výskyt. Využijeme funkci Frekvenční distribuce, kde na všech 4 pozicích zadáme jako Atribut pro třídění lemma.

Frekvenční distribuce

Zadání frekvenční distribuce

Limit: 0

Atribut:	lemma	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Ignorovat velikost
Pozice	0	od	<	KWIC
Atribut:	lemma	<input type="checkbox"/>	<input type="checkbox"/>	Ignorovat velikost
Pozice	1	od	<	KWIC
Atribut:	lemma	<input type="checkbox"/>	<input type="checkbox"/>	Ignorovat velikost
Pozice	2	od	<	KWIC
Atribut:	lemma	<input type="checkbox"/>	<input type="checkbox"/>	Ignorovat velikost
Pozice	3	od	<	KWIC

Přidat Smazat Budiz Zavřít

Objeví se tabulka, v níž zaklikneme „schovat“ nejprve u lemmatu 3. Sečtou se tak všechny výskyt s různými objekty, tj. dát se do práce, do kříku atd. Potom tolež provedeme s opačným slovosledem, když zaklikneme „schovat“ u lemmatu 0 a 1.

Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0 Řádků k zobrazení: 321

lemma: 0<KWIC	<input type="radio"/>	zobrazit	<input type="radio"/>	součet	<input type="radio"/>	schovat
lemma: 1<KWIC	<input type="radio"/>	zobrazit	<input type="radio"/>	součet	<input type="radio"/>	schovat
lemma: 2<KWIC	<input type="radio"/>	zobrazit	<input type="radio"/>	součet	<input type="radio"/>	schovat
lemma: 3<KWIC	<input type="radio"/>	zobrazit	<input type="radio"/>	součet	<input checked="" type="radio"/>	schovat

lemma:	lemma:	lemma:	##
dát	se	do	472
do	který	se	10
do	pohyb	se	6
do	jenž	se	5
do	buducno	se	5
do	jistý	míra	4
do	co	se	4
do	ten	se	4
do	on	se	3
do	který	být	3
do	jaký	míra	2
do	zpěv	se	2
do	voda	se	2
do	se	se	2
do	košile)	2
do	oprava)	2
do	Jeruzalém)	2
do	řeč)	2
do	polovina	se	2
do	neděle	příští	1
do	český	odehrávat	1
do	Verona	akcie	1
do	divadlo	být	1
do		v	1

Zavřít Uložit

Dát se do něčeho jasně převazuje nad do něčeho se dát.
Poměr je 472 : 85.

Úkol 18

- A. V korpusu zadáme dotaz ([lemma="hrozit"] & (tag="V.....A.*")) [lemma="V"] [lc="případě"] a získáme 22 výskytů.
Vedlejší věta následuje 8x, tj. 36%
B. V korpusu zadáme dotaz ([lemma="V"] [lc="případě"] [0,3] [lemma="hrozit"] & (tag="V.....A.*")) a získáme 65 výskytů.
Vedlejší věta následuje 9x, tj. 14%.

Úkol 19

V korpusu zadáme dotaz ([tag="V.*"] [tag="R.*"] [0,3] [lemma="huba"]) a získáme 229 výskytů.

žít	z	ruky	do	huby	
dát			přes	hubu	
chodit	s		otevřenou	hubou	
dostat		pro	po	hubě	
dostat	pro	hubu	na	hubu	
padat			na	hubu	
nelítají			do	huby	
urthovat	si		od	huby	
vyletět			po	hubě	ven
vylítnout	z		po	huby	
dávat			po	hubě	
být			od	huby	
brát			do	huby	jméno
čumět	s		otevřenou	hubou	
řít	na		plnou	hubu	
prohodit	na		půl	huby	
mluvit	na		plnou	hubu	
zírat	s			hubou	dokorán

Příklady:

Zato, že tak **drze <bral do huby> jména**

Vožgrundi jim ešte rádi <dali po hubě>.

Já jsem <dostal po hubě>

Rikál jsem, že <dostaneme přes hubu>

proč tomu starému negrovi <nedal přes hubu>, co

V menších případech <dáme po hubě>

aby mě <nevzali u huby> jako blázna

Potřebuju <padat na hubu>, hodně trénovat a mít pocit

Je na čase, aby nám to vlastnění **<přešlo z huby> do ruky**
Za deset minut **<vyleti po hubě>** ven a já ho sbírám
ale za jakou se **<dává po hubě>**, to je jasná věc
To víš, **<padám na hubu>**, nestihám
tu mně tak **<vylítla z huby>** taková nešikovná narážka na jeho paní.
protože **<žiju z ruky do huby>**
A co vy **<čumíte s otevřenějná hubama>** vy zaprodanci?

Úkol 20

V korpusu zadáme dotaz [lemma="zlatý"] [lemma="hřeb"] a získáme 170 výskytů.

Zlatý hřeb programu, (slavnostního) večera, aukce, sbírky, oběda, utkání, vletu, výpravy aj.

Zlatým hřebem může být hudební skladba, obraz, seriál, ukázka sportovního výkonu, hodnotný, popř. umělecký předmět, jídlo aj.

F.

Úkol 1

V korpusu zadáme dotaz **než** (popř. [word="než"]¹) a získáme 150458 výskytů. Poměr mezi případy, kdy tato spojka stojí mezi větnými členy a mezi větami, tj. kdy před ní píšeme čárku a kdy nikoliv, je 37197 : 113261, tj. 32% : 68%.

Za spojkou **než**, pokud stojí mezi větnými členy, nejčastěji následuje substantivum, předložka nebo číslovka, např.

Její láska pro mě byla víc <než> život.

Raději skončit v rukou nepřitele <než> v zubech a drápech lesní zvěře.

...použití **přísad do jídla ve více <než> 90 zemích** z celého světa.

Před spojkou stojí nejčastěji adverbium, adjektivum (ve 2. stupni) nebo substantivum, např.

...malá papoose umínají pěkrát **častěji <než>** bílé děti, ...

...konec jejího života bude **těžší <než>** začátek.

...svě **jednání úplně jinými principy <než>** ostatní lidé.

Úkol 2

V korpusu zadáme dotaz [Oo]proti a získáme 5614 výskytů.

Na základě funkce Frekvenční distribuce dostaneme následující poměr:

oproti 4327

Oproti 1287

Na začátku věty se předložka **oproti** užívá asi v jedné třetině případů.

Pořadí 5 nejčastějších tvarů za předložkou užitou na začátku věty: tomu, loňské-mu, loňsku, roku, jiným

U nás na druhém stupni základní školy připadá na učitele jen třináct žáků. **<Oproti tomu>** v Holandsku je to 18,2, v Maďarsku 16, ve Velké Británii 15,9 a v Německu a Francii téměř patnáct chlapců a děvčat.

Pořadí 5 nejčastějších tvarů za předložkou užitou uvnitř věty: roku, loňsku, tomu, jiným, loňskému, např.

Prodej zvukových nosičů v České republice zaznamenal loni **<oproti roku>** 1995 čtyřnadvacetiprocentní nárůst a v třídách dokonce nárůst o 34 procent.

Úkol 3

V korpusu zadáme dotaz [I]jestli a získáme 21154 výskytů.

Poměr spojky na začátku a uvnitř věty je 4775 : 16379, tj. opět je asi třetina výskytů na začátku věty.

134

K odlišení významu nám může pomoci soubor sloves, která předcházejí nebo následují za touto spojkou. U tvaru **jestli** zpravidla bezprostředně následují slovesa *být, chtít a mít*, např.

Dala bych si pomerančový džus. „Fajn. Někde se tu posadte. <jestli> chcete,

sedněte si na postel,“ vyběhl mě a šel mi pro džus.

U tvaru **jestli** předcházejí slovesa *vědět, zeptat se a být*, např.

*„Tedy by tam určitě rád šel.“ Já **nevím, <jestli>** je to dobrý nápad. Uprinně řečeno **nevím, jestli** by tam Ted měl chodit.*

Úkol 4

V korpusu zadáme dotaz [lemma="vzhledem"] [I{3} [lemma="že"]¹] a získáme 3805 výskytů.

Potom můžeme použít funkci Frekvenční distribuce a zjistit, která další slova se vyskytují na pozici zájmena *tomu*, a to zadáním pozice 2 (0 = vzhledem, 1 = k, 2 = tomu). Dvě nejfrekventovanější substantiva jsou *skutečnost* – 46x a *fakt* – 14x.

Úkol 5

V korpusu zadáme dotaz [lemma="co"] [pos="N"] [I? [word="to"]¹] [pos="N"]¹] a získáme 132 výskytů.

Nejčastější spojení: *Čech – muzikant, akcionář* (jedná se o aktualizace příslovní), *člověk – názor, slovo – perla, pravda, střela; kus – originál* (umělecké předněty). Jedná se o konstrukci příslušející příslovní, přičemž určitou vlastnost či hodnocení prvnímu substantivu – a tak je také využívána a aktualizována různými obměnami. Ustálená užití nalézáme ve spojeních: *co Čech, to muzikant; co slovo, to perla*.

Úkol 6

V korpusu zadáme dotaz [lemma="být"] [tag="Vf.*"] & [lemma="vidět"] [lemma="slyšet"] [pos="N"]¹] a získáme 954 výskytů. Ide o kombinaci vzorů 6, 12 a 17.

Na 1. pozici zleva se objevují modální slovesa (*chtít* – 29x, *mít* – 8x, *mocet* – 5x, *muset* – 5x), fázová slovesa (např. *začínat, počít* – 1x).

Úkol 7

V korpusu zadáme dotaz [lc="s"] [lc="clem"] [tag="Vf.*"] a získáme 1576 výskytů. Infinitivy, které za touto víceslovnou předložkou následují, jsou převážně hodnotiči, tj. mají evaluativní povahu. Téměř 70% tvoří slovesa s kladným významem, např. *vytvořit, podpořit, vybudovat, zlepšit, obnovit* apod., přes 20% sloves je negativně hodnotiči, např. *zabírat, snižit, odstranit, zastavit* apod., skupin cca 10% sloves je z tohoto hlediska neutrální, např. *najít, dostat, změnit, prodat*. Infinitiv v záporu se objevuje jen 20x.

Úkol 8

V korpusu postupně zadáváme dotazy

[lemma="mocí"] [tag="Vf.*"] – 199581 výskyt

135

[lemma="muset"] [tag="Vf.*"] - 97293 výskytů
[lemma="mít"] [tag="Vf.*"] - 114851 výskytů

Např: po slovese *moci* se objevují nejčastěji slovesa *být*, *mít*, slovesa pohybu (např. *jit*, *vést*, *dojit*, *přijít*) a komunikace (*řít*, *mluvit*).

Infinitivy, které uvedené tři modální slovesa odlišují (mezi 25 nejfrekventovanějšími):

moci – *pomoci*, *způsobit*, *znamenat*

muset – *počkat*, *plazit/zaplazit*, *čekat/počkat*

mít – *sloužit*, *uskutečnit*, *příspěť*

Úkol 9

V korpusu zadáme dotaz ([tag="Vf.*"] [pos="V"] [tag="Vf.*"]) a získáme 807 výskytů.

Nejčastěji se jedná o infinitiv, přítomný nebo budoucí čas, pasivní přičesti.

A. Před slovesy nejčastěji stojí sloveso *být* v budoucím čase nebo modální sloveso.

B. Infinitiv, který je součástí frazému: *nevidět* – frazém *co nevidět* (např. *co nevidět bude sněžít*).

C. Nejfrekventovanějšími slovesy, která se objevují na prostřední pozici, jsou *začít*, *nechat*, *dovolit*, *naučit*, *jit*, *přestat*, *umět*.

Úkol 10

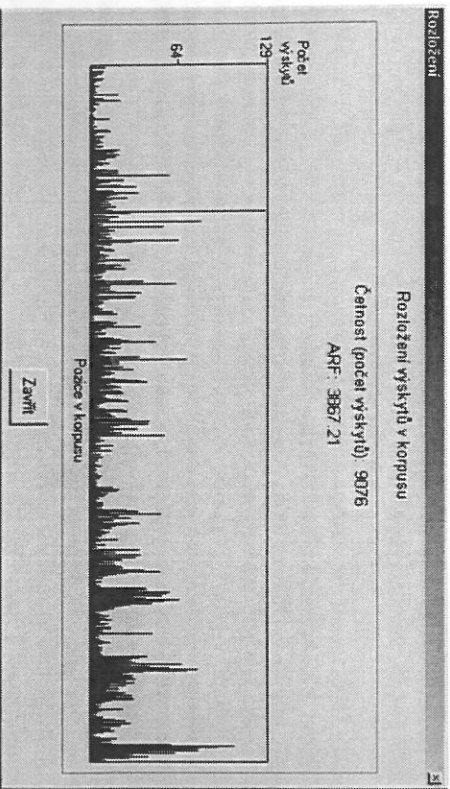
V korpusu zadáme dotaz ([tag="V.....F.*"] [word="\."] [word="\."]) a získáme 12 výskytů (1 z těchto výskytů je však chybně označován).

Nejčastěji se objevuje sloveso *být* a 3. osoba jednotného čísla.

Úkol 11

V korpusu zadáme dotaz [word="\."] [pos="N"] [word="\."] [pos="N"] a získáme 9076 výskytů.

A. Kumulační výtů v jednotlivých textech je možné ověřit pomocí funkce Rozložení:



(menu Konkordance – Statistika – Rozložení)

B. Slova v uvozovkách označují

pochybnosti pisatele o užitím slovu; **ciacj**; **zvýraznění**

Viz uvedené příklady:

přímou politický program skupiny „**realisti**“ <:Masaryka, Kaizla> a **Kramáře**.

„**rakovina kůry**“ <:vaduť, odumřít> věvřa stromů

s **předložkou „re“** <:rekonstrukce, rehabilitace> a další:

„**islámská jižní pětice**“ <:Kazachstán, Kirgizie>, **Uzbekistán**, **Turkmenie** a **Tádžikistán**

hrají „**šestibodov**“ <:Vsetín, Sparta>, **Litvínov** a **Zlín**

posílí tři „**Švédové**“ <:Ščerban, Kaberle> a **Kučera**

Film je pro něj **především „to okolo“** <:dlbý, pletky>, **pózy**, **spekulace** a **pseudo-úvahy**

najíť „**abecedu národních zrádců**“ <:novináře, spisovatele> i **politiky**

C. Slova vyskytující se nejčastěji před dvojitkou:

a) substantiva: **aktivita**, **části**, **činnost**, **dary**, **návštěvní doba**, **druh**, **factory**, **hodnoty**, **jazyky**, **jména**, **materiály**, **nástroje**, **oblasti**, **okresy**, **podmínky**, **pojmy**, **po-
můcky**, **potraviny**, **požadavky**, **principy**, **priority**, **program**, **prostředky**, **prvky**,

předměty, **přílohy**, **příslušství**, **příznaky**, **rubriky**, **slova**, **složky**, **služby**, **survivi-**

ny, **škůdci**, **témata**, **tvary**, **účinky**, **věci**, **vybavení**

b) slovesa: **patí**, **požadujeme**, **zahnuje**

c) jiné – například

D. dotaz [word="\."] [pos="N"] [word="\."](5)- 368 výskytů

Úkol 12

V korpusu zadáme dotaz [tag="N.IS[14].*"] [word="\."] [tag="Vf.*"] a získáme 455 výskytů.

A. Pomocí Frekvenční distribuce zjistíme nejfrekventovanější substantiva na pozici 1: *cíl*, *úkol*, *záměr*, *účel*, *smysl*, *postup*. Pomocí P-Filtu ponecháme ve vyhledávce pouze těchto 6 substantiv: [lemma="cíl"] | [lemma="úkol"] | [lem-
ma="záměr"] | [lemma="účel"] | [lemma="smysl"] | [lemma="postup"]].

Zůstane nám 220 výskytů.

Můžeme si také konkordanci pojmenovat a potom pomocí P-filtu vybrat jedno-

tlivá substantiva a sledovat u nich slovesa. Potom se opět můžeme vrátit k po-

jmenované konkordanci (viz manuál).

B. Největší okruh sloves má substantivum *cíl* – 80, dále *úkol* – 70, další substan-

tiva mají repertoár 10–11 sloves. U substantiva *cíl* je okruh sloves podobný jako

po předložce *s cílem* (srov. úkol 7). Za substantivem *postup* následují slovesa

s konkrétním významem (např. *sešít*, *otřít*), pro substantivum *smysl* jsou charakte-

ristická slovesa *vnutit*, *otupit*, pro substantivum *účel* slovesa *kamuflovat*, *zmašt*,

pro substantivum *úkol* slovesa *řídít*, *shromážďovat*, pro substantivum *záměr* slo-

vesa *ukázat*, *vytvořit*.

C. Nejméně se 3 substantivy se poji slovesa: **nalézt** (cíl, úkol, záměr), **přesvědčit** (smysl, úkol, účel), **připravit** (cíl, postup, úkol), **vybudovat** (cíl, úkol, záměr), **vytvořit** (cíl, smysl, úkol, záměr).

Úkol 13

V korpusu zadáme dotaz [lemma="trojlistek"] a získáme 187 výskytů.

Přivlastky neshodné i shodné se mohou týkat sportovců, umělců, politiků a dalších osobností, ale i různých akcí, jako jsou např. výstavy, zápasy, veletrhy apod.

Příklady z korpusu:

a) neshodný přivlastek

<trojlistek> českých běžců

<trojlistek> českých cyklistů

<trojlistek> českých vrtalců

<trojlistek> hokejistů

<trojlistek> kandidátů

<trojlistek> maturanů

<trojlistek> mladků

<trojlistek> nejslavnějších tenorů

<trojlistek> největších bank

<trojlistek> odborných veletrhů

<trojlistek> populárních herců

<trojlistek> reprezentantů

<trojlistek> teenagerů

<trojlistek> tenistů

<trojlistek> umělců

b) shodný přivlastek

bohemského <trojlistku>

debutující <trojlistek>

Gangsterský <trojlistek>

hutinno <trojlistku>

koaliční <trojlistek>

plesový <trojlistek>

pravcového <trojlistku>

punkového <trojlistku>

superhvězdny <trojlistek>

trenérský <trojlistek>

Negativní význam mají tyto kontexty:

Všetřovatel obvinil povedený <trojlistek> z podvodu.

Ždáší policisté zadřželi <trojlistek> zlodějů, kteří byli doslova postrachem mo-

torizovaných houbatů.

<Trojlistek> obviněn z vraždy.

<Trojlistek> si nakradl za desetistice

Toulky ústředního <trojlistku> na periferii města i zákona podmalovává osobity

hudební doprovod Johna Lurieho.

Úkol 14

V korpusu zadáme dotaz [word="v."] [tag="[N]Z,*"] [3] [word="v?"] a získáme 10719 výskytů. Jednoznačně převládají otázky doplňovací nad zjišťovacími:

Poz 1	Poz 2	Poz 3	?	Frekv.
Je	to	pravda	?	99
Co	to	znamena	?	94
Co	s	tím	?	86
Je	to	tak	?	68
Kolik	jste	vyhrali	?	60
Co	se	stalo	?	47
Čím	to	je	?	45
Co	mám	dělat	?	42
Co	to	je	?	41
Co	je	to	?	41
Co	tomu	říkáte	?	38
Co	bude	dál	?	35
Co	se	děje	?	32
No	a	co	?	29
Chcete	vydělávat	doma	?	28
Je	to	možné	?	27
A	co	dál	?	27
O	co	jde	?	26
Co	říci	závěrem	?	25
A	co	vy	?	25
A	co	ty	?	22
Co	to	bylo	?	22
Kdo	je	to	?	21
Kdo	to	je	?	20
Jak	na	to	?	19
Kolik	je	hodin	?	18
Jak	se	jmenujete	?	18
A	proč	ne	?	17
Co	s	nim	?	16

Na první a druhé pozici se objevují tážací zájmena. Příklady zjišťovacích otázek jsou *Chcete vydělávat doma?*, *Je to pravda?*, *Je to tak?*, *Je to možné?*

Na pozici 1 jsou nejčastěji tážací zájmena, spojky, slovesa *chtít*, *mít*, *být*, *moci*, na pozici 2 tážací zájmena, osobní zájmena, předložky, slovesa *být*, *mít*, *chtít*, částice a adverbia, na pozici 3 jsou nejčastěji zájmena neurčitá, osobní a tážací, substantiva, slovesa *být*, *dělat*, *chtít*, *vědět*, *udělat* adverbia a částice.

Úkol 15

V korpusu zadáme dotaz ([word="\?"] [(tag="P.K.*") | (tag="P.Q.*")]) [0,3] [word="\?"] a získáme 1397 výskytů.

Příklady řetězení otázek:

Monolog

1. Opakování (rozšíření nebo pozmenění otázky) zesiluje váhu otázky, případně odrazí hloubku beznaděje tazajícího se: **Co říkal <? Co říkal? >**

Ale co dělat <? Co proboha dělat? >

Ale co je <? Co se děje? >

2. Postupné rozvíjení otázek – upřesnění nebo další body pro následující odpověď – najdeme zejména v novinářských rozhovorech:

Jak si nyní stojíte na tuzenské scéně? Co plánujete?

Jak jaky to bylo <? Co zvuk? >

Dialog – je výjimečný, museli bychom přidat uvozovky. Vyskytuje se např. u vnitřní řeči postav a je vystaven jako protiotázka: **Co dělá Nada? Kdo je to?** K upřesnění otázky se často užívá rozšíření, konkretizace. První otázka může být všeobecná, druhá je konkrétnější, např.: **Co se stalo <? Co jsem udělal špatně?**

Úkol 16

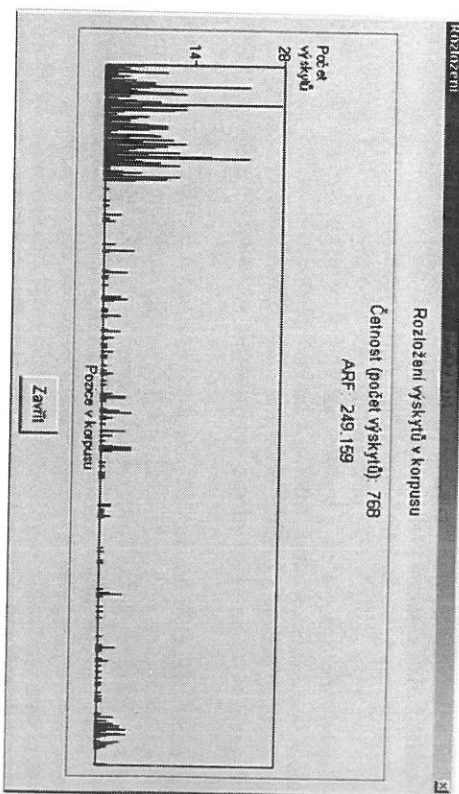
V korpusu postupně zadáváme dotazy ([word="Kam"] [1] [word="\?"]], ([word="Kam"] [2] [word="\?"]]) apod.

	Kam	Kdy
K + 1	191	55
K + 2	312	180
K + 3	235	244
K + 4	154	208
K + 5	112	172
K + 6	67	161
K + 7	56	91
K + 8	32	76
K + 9	31	68

Tyto typy vět tendují k určité délce (3 až 4 slova nejčastěji). Obvykle tedy neobsahují dlouhé popisy a vysvětlení, ale pletí se na informací přímo a stručně.

Úkol 17

V korpusu zadáme dotaz ([word="\."] [tag="N...1.*"] [word="\?"]]) a získáme 768 výskytů.



Korpus je složen z textových typů v následujícím pořadí – beletrie, odborné texty, publicistika. Tedy jsou tyto věty nejčastější v beletrii, která zachycuje strukturu mluveného jazyka a tyto věty jsou pro mluvenost příznačné.

Úkol 18

V korpusu zadáme dotaz ([tag="V.....R.*"] [word="\?"] [tag="V.....P.*"] [word="a"] [tag="V.....F.*"]) a získáme 141 výskyt.

Po odstranění duplicit v rámci funkce Obecné třídění získáme následující seznam:

<Byl, je a bude> provždycky jen hlupák

...co jim je ještě do poezie? **<Bylo, jest a bude>**, odušili...

Česká sociální demokracie se jistě **<dopouští, dopouští a bude>** dopouští i přítomnosti...

...posuzovat „omyly“, k nimž **<došlo, dochází a bude>** docházet při vědeckém bádání.

Celý život **<dělala, dělá a bude>** dělat „co je potřeba“...

Tato choroba bohužel na naší planetě **<existovala, existuje a bude>** existovat...

Vždy se **<hodil, hodí a bude>** hodit do jakéhokoli interiéru...

Oni **<hlali, lžou a budou>** klást otázky, zda zákony přírody existovaly...

Obě země **<měly, mají a budou>** mít i v budoucnosti značně odlišné hodnocení...

Z toho, co Havajci **<nazývali, nazývají a budou>** nazývat aloha.

...myslivecké srdce a peníze dohromady nikdy **<nesly, nejdou a nepůjdou>**...

Vždy **<odrážela, odráží a bude>** odrážet momentální situaci...

Shareware – pojem, kolem kterého **<panovalo, panuje a bude>** panovat mnoho nejasnosti.

Tento čin **<patřil, patří a bude>** patřit k nejpozoruhodnějším výtvorům...

Na celém světě se <pijo, pije a bude> pít.
To <platio, plati a bude> plait i v budoucnosti.

...bájná bytost vždy <pocházelá, pochází a bude> pocházet jediné z Grónska...
Kosmické srážky <posiňovaly, posiňují a budou> posiňovat všechna tělesa...

Češi <potřebovali, potřebují a budou> potřebovat němčinu i ruštinu.
...tím více si mocní <troufají, troufají a budou> troufat...

...co <udělaly, udělají a budou> dělat jiné státy a jejich vlády.

Než jsem si zvykl, že v parlamentu <vystupovali, vystupují a budou> vystupovat lidé...

Mezi námi <žili, žijí a budou> žít příslušníci sexuálních menšin...

■ Úkol 19

V korpusu zadáme dotaz (([lemma="byť"] [tag="Cn.*"] [lemma="rok"] | [lemma="byť"] [tag="PH.*"] [tag="Cn.*"] [lemma="rok"])) a získáme 273 výskyty.

Na základě Frekvenční distribuce (pozice 0, lemma) zjistíme, že slovesa byť a mít jsou v poměru 173 : 100. V PMK je poměr 20 : 2, zatímco v BMK 6 : 4.

■ Úkol 20

V korpusu zadáme dotaz [lemma="byť"] [word="večer"] a získáme 318 výskyty. Ide o rozlišení, kdy je byť pomocným slovesem, kdy je sponou a kdy má skutečně existenční charakter (ve významu nastal večer). Rozdílné funkce tvaru večer jako substantiva a jako adverbia se může projevit i na slovese: byl večer, nejsemutnější to bylo večer.

G.

■ Úkol 1

V korpusu zadáváme pouze sled teček, která zastupují písmena. Nebo můžeme dotaz zadat jednou tečkou a počtem opakování, pro nejdější slovo to bude vypadat takto: {32}

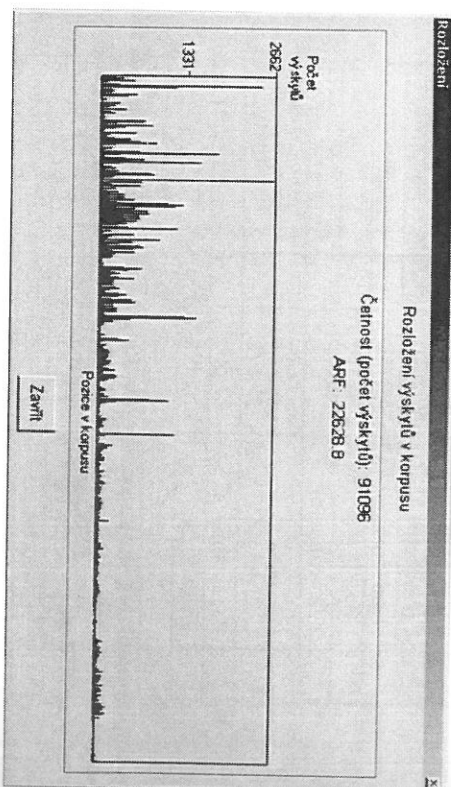
...chemicky jsou steroidy tripeny a jejich základní skelet tvoří nasycený tetra-cyklický uhlovodík steraan. Jeho molekula je plochý <cyklopentanoperhydrofentanenový> systém. Steroidy se liší druhem, počtem a pozicí substituentů, konfigurací jejich vazeb se skeletem...

fy. 32 písmen. Jde o termíny.

■ Úkol 2

V korpusu zadáme dotaz V a získáme 91096 výskyty.

Potom si zobrazíme rozložení textů v korpusu:



Středník převažuje v beletrii, větší výskyt je v odborné literatuře, minimálně je zastoupen v publicistice.

Úkol 3

V korpusu zadáme dotaz $\backslash (\ ? \ ?)$ a získáme 304 výskyty. Pomocí funkce Konkordance – Statistiky – Kolokace vytvoříme následující tabulku:

Lemma	MI-score	T-score	Rel. f [%]	Abs. f
<i>první</i>	3.109	1.531	0.002169	3
<i>poslední</i>	3.941	1.619	0.003861	3
<i>dokonaly</i>	7.45	1.722	0.04398	3
<i>velký</i>	2.715	1.468	0.00165	3
<i>modrý</i>	7.036	1.719	0.03299	3
<i>mladý</i>	4.771	1.669	0.006863	3

Ide o základní adjektiva (Ize je stupňovat). Otazník v závorce, který za nimi následuje, odráží pochybnost autora o dané kvalitě.

Úkol 4

V korpusu zadáme dotaz $[word="\ ."] [pos="rZXI"] [word="\ ?"]$ a získáme 3709 výskyty.

(Značka Z označuje interpunkci, značka X nerozpoznané slovo) Pomocí frekvenční distribuce (pos) dostaneme následující seznam:

N (substantivum)	1238 (+13)
V (sloveso)	583 (+2)
T (časlice)	342
D (adverbium)	313
I (cislovice)	257
A (adjektivum)	127
J (spojka)	95
C (číslovka)	57
P (zájmeno)	50
R (předložka)	17 (-13-2)

Do skupiny předložek je zařazen tvar *díky* (tj. *díky něčemu*), který je však třeba chápat jako substantivum (např. *Díky!* – *Není zač, šéfe!*) a 2x nespisovný imperativ od slovesa *jít* *di*.

Úkol 5

V korpusu zadáme dotaz $[tag="V.....R*"] [word="\ ?"] [pos="r"]\{3\}$ a získáme 311 výskyty.

Nejčastější jsou slovesa mluvení a slovesa označující mentální procesy. Ustálenými víceslovnými spojkovými výrazy jsou *i když*, *i kdyby*. Je to možné porovnat s tabulkou získanou funkcí Frekvenční distribuce:

<i>prohlásit</i>	,	že	i	když	11
<i>dodat</i>	,	že	i	když	10
<i>řici</i>	,	že	i	když	6
<i>dodat</i>	,	že	i	kdyby	6
<i>řici</i>	,	že	i	kdyby	6
<i>přiznat</i>	,	že	i	když	6
<i>uvést</i>	,	že	i	když	6
<i>zjistit</i>	,	že	i	když	6
<i>připomenout</i>	,	že	i	když	5
<i>uvědomit</i>	,	že	i	když	4
<i>vědět</i>	,	že	i	když	4
<i>uvést</i>	,	že	i	kdyby	4
<i>upozornit</i>	,	že	i	když	4
<i>připustit</i>	,	že	i	když	3
<i>bránit</i>	,	a	i	když	3
<i>rozhodnout</i>	,	že	i	když	3
<i>potvrdit</i>	,	že	i	když	3
<i>pochopt</i>	,	že	i	když	3
<i>ukázat</i>	,	že	i	když	3
<i>vidět</i>	,	že	i	když	3
<i>poznámenat</i>	,	že	i	když	3
<i>vědět</i>	,	že	i	kdyby	3
<i>vyplynout</i>	,	že	i	když	2
<i>zdržaznit</i>	,	že	i	když	2
<i>uvědomit</i>	,	že	i	kdyby	2
<i>tušit</i>	,	že	i	kdyby	2
<i>prohlásit</i>	,	že	i	kdyby	2
<i>konstatovat</i>	,	že	i	když	2
<i>připomínat</i>	,	že	i	když	2
<i>naznačit</i>	,	že	i	když	2
<i>prohlednout</i>	,	a	i	když	2
<i>spočítat</i>	,	že	i	kdyby	2
<i>odpovědět</i>	,	že	i	když	2
<i>ujistit</i>	,	že	i	když	2
<i>vysvětlit</i>	,	že	i	kdyby	2

Úkol 6

V korpusu zadáme dotaz $[tag="N.[MI].2*"]\{4\}$ a získáme 2541 výskyty.

Nejčastěji jde o životná substantiva, která jsou rozvířením neshodným přívlastkem – viz příklady 2, 3, 4 a 5 nebo první substantivum následuje po předložce s 2. pádem a za ním stojí teprve neshodný atribut složený ze tří substantiv – viz příklad 4.

Příklad 1 je kombinací neživotných substantiv, kde každý tvar 2. pádu rozvíjí tvar předchozí.

1. vzpomeneš. Amen! ČTVRTÉHO <DNE měsíce června roku> 1783 –
2. dveře vedly přímo na dvůr <staku pana doktora Rosenbauma> který měl
3. dotvzrnuje je rovněž dopis <pana Adama Erdmana Trčky> z Valdštejnova bytu
4. Nad vchodem do <kvelbu krejčího Aarona Rosenbauma> se houkala žároverka
5. úlevou zaregistroval tvář dr. <Johna Francise Xaviera McIntyry>: „Jak se vede
6. „Je to majetek manželky <pana tajemníka pana prezidenta>!“ Když přišel

Na pozici 0 se objevují substantiva typu ředitel, předseda, šéf, náměstek, na pozici 1 substantiva prezident, pan, režisér, klub, parlament, na pozici 2 převážně křestní jména, na pozici 3 převážně příjmení.

Úkol 7

V korpusu zadáme dotaz [pos="A1"] {6} a získáme 120 výskytů.

Příklady na 6 a více předavných jmen:

- stará <ctihodná anglická lesbická perverzní únavná opuštěná> lady
- <nepokrytá tlustou měkkou vrstvou spadlého podzimního> listí
- <jediny bezvýznamný všední neopakovatelný Sychravý> večer
- ve <svetajných celoevropských celulárních digitálních pozemních mobilních komunikacích>
- metodologie <prázdného vývojového fuzzy orientovaného pravidlového expertního> systému

je prázdný <interaktivní programový fuzzy orientovaný vývojový expertní> systém

<provozující železniční veřejnou vnitrostátní pravidelnou osobní> přípravu

<nové české lehké světlé minerální léčivé pivo>

ve kterém mu <podepsaný vrchní účetní nigerijské Národní ropné> společnosti

abys se bavili s nějakým <špinavým utápnutým zbabělým fašounským katolickým španělským> parchantem

Většinou jde o výčty typu Chut! <kyselá hořká sladká pálivá slaná ostrá>, dále o rozvíje přívlastky, stov. nové české lehké světlé minerální léčivé pivo nebo o rozvíje přívlastek stojící za adjektivem – stov. provozující železniční veřejnou vnitrostátní pravidelnou osobní přípravu.

Úkol 8

V korpusu zadáme dotaz ([[tag="D.*6"] | [pos="D1"] [pos="D1"] | [pos="D1"] [tag="D.*6"] | [pos="D1"] | [pos="D1"] | [pos="D1"] | [pos="D1"]]) a získáme 80 výskytů.

Nejčastější adverbia na pozici 0 jsou tak, teda, už teď, pak, na pozici 1 teda, už, zejtra, tu, tam, na pozici 3 jsou teda, taky, ráno, už, zrovna.

Úkol 9

V korpusu postupně zadáváme dotazy

- ([lemma="jednak"] [1]) ([lemma="jednak"])
- ([lemma="jednak"] [1]) ([lemma="jednak"])
- ([word="jednak"] [1]) ([word="jednak"])
- ([word="jednak"] [1]) ([word="jednak"])

Zapisujeme si počty výskytů. Jak ukazuje následující graf, nejfrekvencovanější je struktura se 4 vloženými slovy, a to při vyhledávání lemmatu i slovného tvaru (lemma – 348 výskytů, word 315 výskytů). K mírnému nárůstu počtu výskytů dochází u struktury s 8 vloženými slovy (lemma = 199 výskytů proti 184 předchozím, word = 163 výskytů oproti 152 předchozím).

Příklady na užítí 4 vložených slov:

Teměř všude byli Arabové zaskočení <jednak rychlostí odchodu Anglicanů, jednak> rychlostí židovských vojáků.

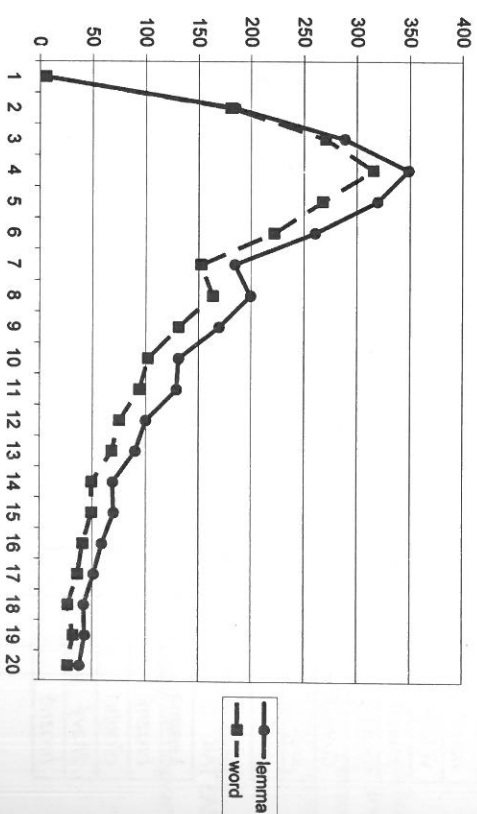
Ve vzorcích přírodních vod jsou toxické prvky přítomny v různých formách <jednak v kapalně fázi, jednak> vázané v suspendovaných tuhých částicích 1, 12, 13.

Četnost rekalibrace se řídí <jednak výsledky testů stability a jednak> nároky na spolehlivost výsledků.

Příklad na 8 vložených slov:

dál k jihu, se častěji objevovaly palmy, <jednak datlové, které hrály důležitou ekonomickou roli, jednak> nižší a většic se palmy důmové, rostoucí v sušších polohách na okrajích pouště.

spojka jednak - jednak



- **Úkol 10**
V korpusu zadáme dotaz **zas tak** – 376 výskytů a dotaz **až tak** – 783 výskytů. Liší se souborem adjektiv a adverbijí, která po nich následují. Společným rysem je to, že v jejich okolí je sloveso v záporu.

AŽ TAK

Lemma	MI-score	T-score	Rel. f [%]	Abs. f
hluboko	9.312	2.995	0.4115	9
zlý	7.705	2.815	0.1351	8
vážně	7.269	2.81	0.09989	8
překvapivý	9.064	2.641	0.3465	7
významný	5.601	2.591	0.03144	7
notoricky	11.79	2.449	2.299	6
pozdě	4.187	1.89	0.0118	4
zásadní	6.287	1.974	0.05058	4
vysoce	6.082	1.97	0.04387	4
zajímavý	5.121	1.943	0.02253	4
reálný	6.056	1.97	0.04309	4
obtížný	6.662	1.98	0.06558	4
cizí	5.824	1.965	0.03668	4
vzdálený	6.817	1.982	0.07302	4
bezpečný	5.431	1.954	0.02793	4
jednoznačný	7.029	1.719	0.08455	3
silný	4.158	1.635	0.01156	3
vážný	5.468	1.693	0.02866	3
dramaticky	9.118	1.729	0.3597	3
náročný	5.711	1.699	0.03392	3
velký	5.797	1.701	0.03601	3
závatný	9.44	1.73	0.4498	3
překvapující	9.659	1.73	0.5236	3
dramaticky	6.38	1.711	0.05395	3
kritický	6.189	1.708	0.04725	3
jasný	4.529	1.657	0.01496	3

ZAS TAK

Lemma	MI-score	T-score	Rel. f [%]	Abs. f
hrozný	9.39	2.642	0.2086	7
dlouho	6.003	2.201	0.01995	5
žhavý	9.535	1.73	0.2308	3
hrozně	8.278	1.726	0.09656	3

Společnými adjektivy a adverbijí jsou: *moc, velký, důležitý, špatný, úplně, do-
brý, složitý, mnoho, málo, jednoduchý, obvyklý, špatně, dávno, docela, příliš,
často.*

■ Úkol 11

V korpusu zadáme dotaz `[lemma="v"] & [lc="rámcí"]` a získáme 846 výskytů. Adjektiva s protikladným významem jsou v širokém x úzkém rámcí, ve starém x novém rámcí.

Lze vložit maximálně 4 slova – příklady. To zjistíme pomocí následujícího dotazu: `[lemma="v"] [f2] [lc="rámcí"]`, z nalezene konkordance potom vyloučíme případy v rozhlase v rámci pomoci N-filtru `[lc="v"]`, pozice od `<KWIC> 1`. Interval zvyšujeme.

`<v tomto již dále zúženém rámcí>`

`<v širším politickém a ekonomickém rámcí>`

`<v jeho obecně platném tradičním rámcí>`

`<v adekvátním institučním a zákonném rámcí>`

■ Úkol 12

V korpusu zadáme dotaz `[lemma="chtít"] [lemma="být"] [lemma="vědět"] & [tag="Vf.*"] [word="m"] [tag="f.*"]` a získáme 42 výskytů.

Potom zadáme dotaz `[lemma="být"] [lemma="chtít"] [lemma="vědět"] & [tag="Vf.*"] [word="m"] [tag="f.*"]` a získáme 12 výskytů.

Ve většině případů je vedlejší věta předmětná. Následující spojky podle frekvencí: *jestli, zda, že, ačkoli, protože, ale*. V prvním slovosledu jde především o zá-
čátky vět – oznamovací a podmiňovací způsob, v druhém případě je z prozodic-
kých důvodů přehozené pořadí pomocných tvarů slovesa *být* u oznamovacího
způsobu a vyskytuje se i čas budoucí.

■ Úkol 13

V korpusu zadáme dotaz `[lemma="podívat"]` a získáme 15175 výskytů.

Můžeme použít P-Filter, kde si pro 1 pozici vpravo zadáme `[pos="R"]`. Pak bu-
deme mít 3777 výskytů. Dále můžeme pracovat s náhodným výběrem např. 300
výskytů (Konkordance – Redukce – Náhodných 300). Velmi dobře se lze oriento-
vat i tehdy, použijeme-li místo náhodného výběru jednoduché třídění (menu
Konkordance – jednoduché třídění – Právý kontext, počet tříděných pozic 1).

Některá spojení slovesa *podívat se* s předložkou lze považovat za standardní:
`(...) podíval se na bratra a pejšší na autobus do Petrovic.`

`Idaho povzdechli a podíval se k oblozkovým dveřím, které vedly do obyvatel
míšnosti.`

`Rychle otevřela Zlata skříň a podívala se do krabice...`

`(...) sundal černý barel a podíval se skrz okno na nádražní hodiny.
Chlapec se obrátil a podíval se po starcích na verandě.`

Vždycky napřed obešel ovčiny a stodoly, **podíval se za** závěs do přístěnku, vyzval se nočního hliďáče Kudumata (...)

Obrátila se a **podívala se přes** most.

A pak mě cosi napadlo a spustil jsem se do podklepu a **podíval se pod** marinkotku.

U některých spojení s předložkou však už počítáme menší či větší posun ve významu. U kterých předložek se to děje? Uvedte příklady:

NA

Před osmi lety jsme se s bratrancem vydali na pionýru z Přerova do Boubětic u Prahy, abychom **se podívali na** první trialový závod u nás. (navštívit, zajít někam, zajet někam)

Když **se na to podíváme** zpátky, Alexandra, Sukiie i Jane Šmartová mohly přijít na to, že se s nimi nudí; ale ony se s ním nenudíly ani trochu, a tak myšlenka na jakoukoliv nudu byla za hranicemi jejich představitivosti. (zamyšlet se nad něčím)

„**To bych se na to podíval**, abych mě vyhostil: jsem tu doma!“ (věcně vyjádřená částice)

K

Mnozí z nich **se** poprvé se sružením **podívali k** moři, „dodal Halm. (v textu bylo prve „se“ vynecháno.) (navštívit, zajít někam, zajet někam)

DO DUŠE

Nedejte na to, co vidíte, **podívejte se** hlouběji **do jeho duše**, a uvidíte, že volá –

„Táto, mámo, zajímejte se o mě“. (pochopit pravě důvody)

PO

Možná, že bychom **se** pak mohli **podívat po** Delu Styronovi. „Víte o něm něco?“ (hledat)

ZA

Četl v novinách o bídě, která dolehla na jeho bývalý domov, a rozhodl se, že se přijede **podívat za** otcem, jak to bude jen trochu možné. (navštívit)

Úkol 14

V PMK nelze vyhledávat podle tagů ani podle lemmat. Je třeba zvolit určitou koncovku pro pasivum a vyhledat ji, např.: ***-án**. Pak použít frekvenční distribuci a určit nejčastější pasivní tvary: 4x: *dán*; 2x: *odměňován, informován, vázán*, *za-měštnán*. Nejfrekventovanější ze spisovných tvarů je neutrum *dáno* (30x), ovšem mnohem frekventovanější je obecně český tvar *dány* (96x). Frekventovanější jsou tvary obecně české i v BMK. Naleznete tu však také tvar *dané* pro neutrum, který v PMK zcela chybí. Ztrátka se objevuje adjektivní užití v přívlastku.

Úkol 15

V korpusu zadáme dotaz ([lemma="lézt"] [pos="NPCI"] a získáme 925 výskytů. Nejfrekventovanější spojení jsou

150

lézt na nervy	63
lézt po čtyřech	34
lézt po stromech	24
lézt do postele	17
lézt do zelí	16
lézt po kolennou	11
lézt do hlavy	10
lézt na stromy	8
lézt do peněz	6

Úkol 16

V korpusu zadáme dotaz ([lemma="kolenn"] [lemma="*tka"] & [tag="*...2.*"]) a získáme 126 výskytů.

Nejčastější jsou spojení *kolenn třicítky* a *kolenn čtyřicítky*, která se vztahují k osobám středního věku.

Věk označuje substantivum *desítka*, jde o počet deset. Dále sem nepatří možné substantivum *nosítka*.

Úkol 17

V korpusu zadáme dotaz ([lemma="být"] & (word="[j]s.*")) [tag="A.*6"] a získáme 561 výskytů.

Potom si na základě frekvenční distribuce seřídíme výskytů podle textových typů:

NOV	247
PUB	129
COL	62
SON	49
SCR	27
VER	15
FAC	15
SCI	7
POP	7
MIS	3

Tento tvar výrazně převládá v beletrii.

Úkol 18

V korpusu zadáme dotaz ([lemma="pan"] & [tag="*...5.*"]) [tag="N...1.*"]) a získáme 820 výskytů.

Ke zjištění povolání nám pomůže funkce Frekvenční distribuce:

151

2. Co říká o různých slovech korpus a co slovníky

šéf	49
doktor	45
hrabě	29
vrchní	15
spisovatel	10
detektiv	9
výpravčí	7
čisarf	5
průvodčí	5
učitel	5
farář	4
mistr	4
vedoucí	4
hostinský	3
král	3
primář	3
profesor	3

Úkol 19

V korpusu zadáme dotaz [lemma="odstřelovač"] a získáme 294 výskvyty.

Poměr je následující:

ostřelovač 269x
odstřelovač 25x

Rozsáhlé vysvětlení viz <http://www.ujc.cas.cz/poradna/portaq.htm>
Ostřelovač by měl být ten, kdo někoho nebo něco ostřeluje: střílí z dálky, ohrožuje cíl palbou; zásah sice nelze vyloučit, ale nemusí se ani předpokládat. Střelba není jednorázová, ale opakovaná, v dávce apod. – ostřelování města může trvat i dny a týdny. Zato odstřelovač své cíle střelbou likviduje, a to zpravidla jedním, přesně zacíleným výstřelem. Stejně slovo může také označit toho, kdo by odstraňoval dům nebo skálu, kdo by pracoval v lomů s trhavinami. Z výkladových slovníků češtiny (PSJČ, SSJČ, SSČ) má heslo odstřelovač jen SSJČ.

Úkol 20

V korpusu zadáme dotaz [Z]lat.* Prah.* a získáme 282 výskvyty.

Citát z Bohumila Hrabala:

Tam u okna, za plentou, hned jak zahnete u výčepu, tam U zlatého tygra sedají mladí muži, malíři, grafici, absolventi akademie i grafických škol, kteří se žíví designem, reklamami... a ti mají <Zlatou Prahu>, která každý den tady po čtvrté hodině pije pivo, společnost, která každý rok pořádá výlety parníkem nebo pěšky do přírody, je to Zlatá Praha, která strojí v zimě plesy či šibňinky, Zlatá Praha, která každý rok vydává kalendář, slavný kalendář, známý ve všech světádlech...
pi 60 Bohumil Hrabal Pohádka o Zlaté Praze

152

	Správná odpověď	Počet výskytů v SYN2000
bigotní	A	59
bonnot	C	304
debutovat	A	391
dedikace	D	61
deportace	A	355
depozitář	D	756
doktrína	A	830
drahně	B	14
epigon	A	79
euforie	D	898
evokovat	B	417
ex post	D	141
exallovane	A	7
frapantní	D	20
frenetický	B	67
charizma;	B	19
image	A	1832
impresário	C	57
interakce	D	702
intoxikace	B	63
invektiva	C	149
kinetóza	B	7
koncizně	A	2
konfúzní	B	28
konzistentní	C	231
kooplovat	A	73
kosmologie	D	82
kruciální	A	18
maně	A	85
mentorovat	B	20

153

meritum	A	80
nabob	C	21
noticka	D	43
nullitní	D	15
parafrazovat	B	174
parlamentář	C	30
pasus	A	105
pitoreskní	B	104
plytký	A	112
polyhistor	C	19
pragmatický	D	883
preambule	C	178
pregnantní	B	53
premissa	A	200
profanace	C	24
provenience	C	403
puč	C	1089
půlka	B	210
reminiscence	C	271
rudimentární	A	36
signifikantní	B	88
snovat	D	40
suspektní	D	12
talár	B	130
tiráž	A	136
traktovat	A	38
travesitie	C	42
tripytych	B	142
tristní	B	274
tryzna	B	106
uzurpování	C	7
verifikovatelný	D	16
voluntaristický	A	17
xenofobie	B	337
zhatit	C	226

3. Význam slova prozrazuje kontext

A.

- Úkol 1 *pohled*
- Úkol 2 *vosko*

B.

- Úkol 3 *výjimka*
- Úkol 4 *svízelný*

C.

- Úkol 5 *na základě*
- Úkol 6 *chytřá horákyně*

D.

- Úkol 7 *polobotku, polodrahokamů, polohlasem, polokonzervu, polokošilich, poloměr, popouští, polospánku, pološera, polotovarů*
- Úkol 8 *hrůzovláda, královláda, krutovládou, lidovláda, samovláda, sebevláda, stranovláda, světovládou*

E.

■ Úkol 1

1. efektní
2. efektivní
3. efektní
4. efektivní
5. efektivní
6. efektní
7. efektivní
8. efektní
9. efektní
10. efektivní

■ Úkol 2

1. abstinence
2. abstinence
3. absence
4. abstinence
5. abstinence
6. absence
7. abstinence
8. abstinence
9. abstinence
10. abstinence
11. abstinence
12. abstinence
13. abstinence
14. abstinence
15. absence

■ Úkol 3

1. reprezentační
2. reprezentační
3. reprezentační
4. reprezentativní

156

5. reprezentační
6. reprezentační
7. reprezentační
8. reprezentativní
9. reprezentativní
10. reprezentativní

■ Úkol 4

1. adoptovat
2. adaptovat
3. adaptovat
4. adoptovat
5. adoptovat
6. adaptovat
7. adaptovat
8. adaptovat
9. adaptovat
10. adoptovat

■ Úkol 5

1. mentorovat
2. memorovat
3. mentorovat
4. memorovat
5. mentorovat
6. mentorovat
7. memorovat
8. memorovat
9. mentorovat
10. mentorovat

157

F.

Úkol 1

1. V knize La Bureaucratie 1946 napsal Ludwig von Mises: Pod vládou státem kontrolované <ekonomie – 2> nemohou parlamenty být ničím jiným než shromážděním mužů, kteří vždycky říkají ano!
2. Neuvažej! Žes mi dal kouli z <ekonomie – 1>, to ještě nemusím být blbá.
3. Iduho znal <ekonomiku – 1> Imperia, na nejvyšší míru zjednodušenou: žádné úroky se nepovolovaly; peníze na dřevo.
4. Pan inženýr <ekonomii – 1> bude končit teprve příštím rokem a navíc musí na rok na vojnu.
5. Nobelovu cenu za <ekonomii – 1> pro rok 1994 získali profesor John Harsanyi z University of California v Berkeley, profesor Nash z Princeton University a profesor Reinhard Selten z Rheinische Friedrich – Wilhelms – Universität v Bonnu za průkopnickou analýzu rovnováhy v teorii nekooperativních her.
6. Díky jistým znalostem nemocniční <ekonomiky – 1> odhadoval, že Američi Care musel postihnout finanční kolaps.
7. Kolokace, víceslovná pojmenování můžou mít i komponentů více než dva jako v předchozím příkladu, zvl. v případě speciálních termínů, srov. několkavřetenový řádový obráběcí stroj a ekvivalenční angl. multi-gang-spindle machine, nebo i vlastních názvů (řeba instituci aj.), jejich přílišný růst však naráží na požadavek jazykové <ekonomie – 3>.
8. Jestliže má věřitel prostředky (v našem případě jich má až poženhaně) k vy-máhání splatné pohledávky, je proti smyslu zákona i proti zásadě <ekonomie – 3> testní represe nahrazovat nástroje k vymáhání pohledávky opatřeními orgánů činných v trestním řízení.
9. Transformace naší <ekonomiky – 1> v efektivně fungující tržní systém nebude bezbolestná, budeme muset změnit mnoho svých návyků, budeme muset spolupodepsat novou – jak se říká v odborné literatuře – „společenskou smlouvu“, která se od té minulé bude v mnohém lišit.

Úkol 2

1. Konečně stál před onou místností. Sáhl na <kliku – 1> a ono zamčeno.
2. Eduardo Crawley má za to, že obsah dokumentu se naprosto shoduje s potvrzenými výroky této vojenské <kliky – 3> a říká, že „jestliže to byl padělek, pak mimotřádně dokonaleý.“

3. Mám přitom ještě <kliku – 4>, že si ten Pída netroufne víc, v jeho očích přece jen Bonek patří k nám do rodiny.
4. Položte si stejný předmět na zátylek nebo na záda a dělejte <kliky – 6>.
5. PRODÁM 2 CITERY a kuřířkový gramofon s <klikou – 2>.
6. Přestože v Hološovicích začínají leštit <kliky – 5> sponzoři, kteří zde tuší zajímavou reklamu, hodné místních lidí se domnívají, že UNESCO jim přineslo problémy.
7. Měli jsme šílenou <kliku – 4>.
8. Tehdy ovšem přišla k moci revizionistická <klika – 3> Dubčeka.
9. Ten druhý kluk dělal <kliky – 6> a karate kopy, aby nějak překonal nervozitu.
10. Vládnoucí <kliky – 3> sice opanovali široší kraj, ale jsou obleženy ve svých pevnostech, a spoléhají zatím jen na to, že nikdo z nás neodhadne vlastní sílu.
11. Černej nebožtíku Máš to ale <kliku – 4> Za chvíli do temný hlíny budeš zakopán
12. Bez ohledu na to, že tehdejší právní řád v souladu s vůlí vládnoucí mocenské <kliky – 3> chránil socialistické společenské zřízení.
13. Člověče, ale měl jsem <kliku – 4>, že jsi mi tehdy neustejtili pytlík.
14. Můžete ovladat <kliku – 2> přepečlivě zaplombovat, stanovit pravidla pro použití tohoto přístroje, ale nemůžete zaručit, že některý z všemocných nakonec nepootočí <klikou – 2> a nepustí do sítě sílcí proud hrubé síly.
15. Byla to <klika – 4> ho poťkat.
16. Hodně lidí čeká, že jejich poslanec bude běhat po ministerstvu financí, leštit <kliky – 5> a shánět peníze pro jejich region.
17. Prosim, vstupte! Máte <kliku – 4>, mám zrovna volno.
18. Z Lucie stala se Lizzy Wagnerová, dcera hajného německé národnosti Dietricha, a Jan Pallas byl v této verzi chudým vysokoškolačkem – národovcem, jehož demoliberalní hradiší <klika – 3>, ovládající vysoké školy, vyškvala do Sudet.
19. Billy otočil <klikou – 2> a zatápal po vypínači v hale, rozsvítil a pustil Lowryho napřed.
20. Je to možné, tolik <kliky – 4>? Může člověk být zamilovaný do dvou holek najednou?
21. Položil dlan na kouli <kliky – 1> a zvolna jí otáčel, potom dveře pootevřel.
22. „Ave!“ zvedl ruku. „Maria,“ dodala Nadja a zvedla desku, nasadila ji na gramofonový talíř, pak chtěla natáhnout pero, ale <klikou – 2> otočila naprázdno.
23. Tim se dostává do rozporu se zájmy sovětské imperialistické <kliky – 3>, a proto ho pronásledují.
24. Měli jsme <kliku – 4>!
25. Homais začal „čistit <kliky – 5>“
26. Dělal <kliky – 6> na jedné ruce a celou váhu těla přenašlel na ochrnutý prst.
27. Galerie Zichláš <klika – 1> na Starém Městě doplnila výstavu o zajímavá díla ze Sivkovy volné tvorby.

28. ... jak ti šílenci, mý kámoši a společníci, poskakujou a dělají <kliky – 6> a boxujou se ...
29. „Žádný producent o mne neměl zájem a já sám jsem nechtěl leštit <kliky – 5>“, říká režisér o této pauze.
30. Nikdy se nedá vystřídat, pokaždé letí v čele, asi má strach o svou prestiž, třeba mu v té smečce dělá rozruch nějaká opoziční <klika – 3>.

Úkol 3

1. Sledovali jsme roby a <masky – 1>, něco jsme pak použili.
2. Usměvavá <maska – 2> mu spadla s tváře.
3. Přenašeni tradičních <mask – 5> na hlavní postavy, jak je můžeme sledovat u Voskovce a Wericha, je antipicováno v Pirandellově hře šest postav hledá autora.
4. Puritáni byli odjakživa mistry farizejských <mask – 2>.
5. Všechny <masky – 1>, které tu vidíš, ti zakryjí celý obličej a úplně tě změní.
6. Kosmetická hlinka Bílá hlinka, která dává konzistenci <maskám – 9> a peelingovým krémům.
7. Ten byl postaven na hlavu a jeho tvář kryla posmrtná <maska – 3> ze zlata a se smaragdovými očima.
8. Okamžitě, automaticky, bez záslahu pilotů se nad hlavami cestujících začali houpat kyslíkové <masky – 4>.
9. Desítky a desítky různé zdařílých zvířecích i historických <mask – 6> procházely večera ponurými ulicemi pražského Žitkova v již tradiční lidové veselici.
10. Pro ně je místo klasické barvy nejlepším řešením takzvaná barvící <maska – 9>, která vlas vyživuje a zároveň barví.
11. Přístrojen je do ochranné plynové <masky – 4> (rouška by v tomto případě nestačila), na ruku azbestové rukavice a v nich vatečka.
12. Gólmani si pomalují <masku – 4> hrůzostrašným zjeven jako válečníci nebo pravěcí lovci mamutů.
13. Jen <maska – 2> bolesti na tváři mužově byla pravá.
14. Termín <maska – 7> je používán pro pigmentaci hlavy. Tato pigmentace však musí zasáhnout celou přední část hlavy a uši.
15. Na první pohled je patrná nová <maska – 8> chladiče, připravené nárazníky, jiné dveřní kliky.
16. Nová výrazná chromovaná přední <maska – 8> dává vozu sympatický až dravý sportovní nádech a odlišuje ho od konkurence.
17. Dobře, že je zde takový žalobce jako pan Brückler, co umí strhnout <masku – 10> podřadosti a očistit nevinné.
18. Ovšem, navenek, protože ve společnosti si dovedu nasadit <masku – 10> posměváčka.
19. Když zjistili, že to máme v hlavě srovnané a že to není jen nějaká byrokratická umanutost, odložili <masky – 10> suverénů a docela se snažili.

20. Smích rezonuje v bolestně vědoucí výsměch prázdným <maskám – 2> vyhaných životů.

Úkol 4

1. Finanční částku, za niž kontrolní <balík – 3> akcií získal, ale údajně odmítl sdílet.
2. Například nejnovější verze Windows 95 zabere v počítači šedesát megabytů, <balík – 4> kancelářských programů Microsoft Office dalších sto dvacet a doplňkové programy Plus Pack čtyřicet megabytů.
3. Ministr školství Zeman úspěšně prosazuje, aby se jeho <balík – 2> peněz na výplaty zvýšil o 15 procent...
4. Jenom ze sotva číselného razítka jsem rozpoznala, odkud byl <balík – 1> poslán, věru na žádný dík nečekali.
5. V OKD bych měl zastupovat stát, který vlastní významný akcionářský <balík – 3> této společnosti.
6. ... přišel jsem z venkova, čili jsem byl takovej jako venkovskej <balík – 5>...
7. Pro mnohé lidi je stále pohodlnější si naše služby zaplatit, než se doručováním dopisů a <balíků – 1> zdržovat.
8. V zájmu dlouhodobě politiky však nepovažují za účelné, aby <balík – 3> akcií získal ČEZ.
9. Škoda, že nemůžu být u toho, když bude <balík – 1> rozbalovat ...
10. Druhým <balíkem – 3> problémů jsou technické a provozní předpoklady, infrastruktura je bohužel zcela zastaralá a neodpovídá dnešním požadavkům.
11. Nebyla žádný venkovský <balík – 5>.
12. Protože byl vydán celý <balík – 3> nové legislativy, omezím se pouze na významný zákon č. 244/92 Sb.
13. Jackie pošle v <balíku – 1> špinavé prádlo na vyprání.
14. Další aplikací byla sbrčka zákonů a poslední byla instalační sada antivirového <balíku – 3> Avast!
15. Zbytek celého <balíku – 3> legislativních opatření na schválení v dolní komoře teprve čeká.
16. Boženko, dostaly jsme <balík – 1>, volávala, ona dostávala z domova <balíky – 1> přes pět kilo, samý uzený maso a cukroví.
17. O problematice audiovizuálního <balíku – 3> se v ženevském sídle GATT jedná již od středy a široké plénum účastníků mu dodává celosvětový rozměr. Oba nejvyšší ruští představitelé se shodli na tom, že nyní je nejdůležitější, aby parlament bleskově schválil devatenáct zákonů z tzv. antikrizového <balíku – 3>.
19. Ale více než 100 z těchto programů jsou hry, softwarové <balíky – 3> pro modernizaci domova nebo se týkají multimédií.
20. Číšník, který si pod objednávkou kávy nepředstavil nic jiného než pořádného turka, si vysloužil celostátní ostudu jako venkovský <balík – 5>, nemající šajnu o tom, co se patří mezi kulturními lidmi.

21. V prodeji televizních práv přitom panuje zvyk, podle něhož je například každý „atraktivní“ film dodán v <balíku – 2> spolu s několika dalšími „běč-ky“ a seriály různé kvality.
22. Až budeme v <balíku – 6>, vytáhnou tě vodsaď, než řekneš popel, platíš?
23. Pro Slovensko je tento zápas náročnější, protože do něj vstupuje s těžším <balíkem – 2> problému.
24. ... v městě jsou davy lidí, všichni jsou v <balíku – 6>, přesto závidí.
25. Z Trendu se ztratili i majoritní <balík – 3> akcii Kotvy.

■ Úkol 5

1. Mluvíč české biskupské <konference – 2> Daniel Herman upozornil, že se v ýzdoeba kaple stane stálou vizitkou úspěšlosti a úrovně českého křesťanské-
ho umění konce dvacátého století.
2. Proti zavedení školního se večera vyslovila Česká <konference – 1> rektorů.
3. Na věřejší dárcovské <konferenci – 2> pro Kosovo, kterou uspořádala Svě-
tová banka a Evropská komise, se vybraly první dvě miliardy dolarů, které
stačí na pokrytí okamžitých potřeb kosovského obyvatelstva.
4. MEZINÁRODNÍ DIAKONICKÁ <KONFERENCE – 2> na téma „Kolik práce
potřebuje člověk?“ proběhla ve dnech 25. – 29. dubna v Jilovišti u Prahy.
5. Na Evropských <konferencích – 2> se mají totiž scházet šéfové států a vlá-
d zemí, které o členství v unii usilují, tedy nejen ty, s nimiž už se jedná, ale ta-
ké ty, které sice mají vidinu členství, ale až v dlouhodobějším časovém hori-
zontu.
6. Toronto čeká ve finále <konference – 5> Buřtalo nebo Boston.
7. V dokumentech Helsinské <konference – 2> z roku 1975 se zdůrazňuje ja-
ko hlavní princip neměnnost hranic v Evropě.
8. Součástí této prezentace bude také internetová <konference – 4>, tj. diskus-
ní fórum k jednotlivým tématům i k vlastním návrhům.
9. V malajsijském hlavním městě Kuala Lumpuru zahájilo večera třináct států
Organizace islámské <konference – 2> jednání o dosažení míru v Bosně
a o její budoucí obnově.
10. V kulooárech <konference – 1> se totiž delegáti netajili nespokojeností s je-
ho politickým stylem.
11. Když není na nějaké mezinárodní lékařské <konferenci – 1>, konec týdne
tráví obvykle v domku u Prahy, kam může s rodinou jezdit celoročně.
12. Mezivládni <konference – 2> by měla podle názoru některých delegací
omezit právo véta jednotlivých států při přijímání závazných rozhodnutí
v unii.
13. V Haagu probíhá První mírová <konference – 2>, která vymezuje pravidla
vedení války.
14. Čapek po mnichovské <konferenci – 2> žil úplně v zoufalé náladě.
15. Roku 1990 této publikaci udělila Národní <konference – 1> amerických
učitelů angličtiny orwellovskou cenu Dvojí řeči (Doublepeak Award).

16. Atlanta naopak svedla ve vyřazovacích bojích tuhé bitvy a ve finále Národní
<konference – 5> jen se štěstím porazila nejlepší tým základní části Minne-
sotu 30 : 27 po prodloužení.
17. Konaly se specializované odborné <konference – 1> např. na téma Vivu
a odstraňování účinků radonu a také celodenní konference o inženýrském
vzdělávání.
18. Celostátní ornitologická <konference – 1> nazvaná Dravci a sovy '91 zača-
la v Přerově.
19. Z tohoto hlediska bychom e-mailový dialog museli považovat za reakci
spontánně mluvenou (počítačové <konference – 4> mají v tomto ohledu
charakter smíšený).
20. V době, kdy se u nás diskutuje o demokratizaci a liberalizaci školství je zají-
mavé slyšet, že školních <konferencí – 3>, které řeší všechny podstatné pro-
blémy školy, se v Německu zúčastňují nejen učitelé.
21. Po semináři a exkurzi následovala tisková <konference – 6>, která byla dob-
ře obsazena; zástupci tisku, kteří směřovali své otázky nejen do oblasti pa-
nelových domů, ale i do ekonomických a sociologických oblastí bytové po-
litiky.
22. V úvodním seminále Východní <konference – 5> Pittsburg zvířezil nad
New Jersey 3 : 2.
23. Přihlášení do elektronické <konference – 4> a zaslání příspěvků:
24. Moji pořádní sekretáři mi sdělili, že jsem měl 154 řečí, 31 <konferencí – 6>
s žurnalisty a 29 anglických rozhlasů.
25. Vyrostl z toho problém, který si vyžádal zvláštní <konferenci – 3> učitelské-
ho sporu.

4. Úkoly pro náročnější

A.

■ Úkol 1

Správný dotaz je: [tag="INAPCI...4.**"] Hledáte totiž slova patřící ke jmenným slovním druhům; tj. substantiva, adjektiva, zájmena a číslovky. Kdybyste zadali jednodušší dotaz [tag="...4.**"], zahrnuli byste i předložky pojící se s akuzativem.

■ Úkol 2

Správný dotaz je: [tag="Vp.**"] a druhý dotaz je [tag="Vs.**"].

■ Úkol 3

Správný dotaz je: [tag="Ve.**"]. Kdybyste nevěděli, jak se přechodník vyjádří, zadáte slovesný tvar, který je určité přechodníkem, např. *chtíc*, načez v hlavní nabídce vyberete Zobrazení a pak Atribut a volbou tag si zobrazíte značku slova *chtic*.

■ Úkol 4

Správný dotaz je: [tag="R...[∧G].**"] [tag="INAPCI...6.**"] Sami uvažte rozbořem konkordančních řádek, pokud jste nějaké dostali, zda je taková konfigurace vůbec možná.

■ Úkol 5

Postupujte takto:

- a) zapněte Zobrazení – Atribut *lemma* (nepovinné)
- b) Zadejte dotaz: [tag="N..P.**"]
- c) Zobrazte si frekvenční distribuci těchto lemmat *podle lemmat* (zadejte Konkordance – Statistika – Frekvenční distribuce) a uložte si výsledek do souboru PLURAL
- d) Zadejte dotaz: [tag="N..S.**"]
- e) Zobrazte si frekvenční distribuci těchto lemmat *podle lemmat* a uložte si do souboru SINGULAR
- f) Odečtěte soubory PLURAL a SINGULAR, tj. odstraňte ze souboru PLURAL to, co se nachází v souboru SINGULAR. Uděláte to nejlépe tak, že si napíšete jednoduchý počítačový program, který to udělá za Vás.

164

■ Úkol 6

Řešení je obdobné řešení úkolu 5, stačí zaměnit P a S, resp. soubory PLURAL a SINGULAR. Je tedy zřejmé, že musíte postupovat takto:

- a) zapněte Zobrazení – Atribut *lemma* (nepovinné)
- b) Zadejte dotaz: [tag="N..S.**"]
- c) Zobrazte si frekvenční distribuci těchto lemmat *podle lemmat* a uložte si ji do souboru SINGULAR
- d) Zadejte dotaz: [tag="N..P.**"]
- e) Zobrazte si frekvenční distribuci těchto lemmat *podle lemmat* a uložte si do souboru PLURAL
- f) Odečtěte soubory SINGULAR a PLURAL, tj. odstraňte ze souboru SINGULAR to, co se nachází v souboru PLURAL.

■ Úkol 7

Vidíte, že vid není u sloves označkován. Co tedy podniknete? Nezoufávejte a dobře si uvědomte, jaké mluvnické vlastnosti mají dokonavá a nedokonavá slovesa. Dokonavá slovesa mají na rozdíl od nedokonavých tyto vlastnosti:

1. netvoří opisný budoucí čas
2. nemohou stát po fázových slovesech (začít, pokračovat, přestat):
3. mohou stát ve vedlejší větě uvozené spojkou *dokud* pouze tehdy, jsou-li v záporu.

Při řešení úkolu postupujte tedy v souladu s těmito kroky, v nichž budeme získávat výhradně nedokonavá slovesa:

1. krok – opisný budoucí čas:
 - a) zapněte Zobrazení – Atribut *lemma* (nepovinné)
 - b) zadejte dotaz: [tag="VB...F.** & lemma="by"] [tag="NARDCP.**"]? [tag="Vf.**"]
 - c) uložte frekvenční distribuci *podle tvarů* posledního zadaného prvku, a dostanete tak soupis slovesných infinitivů, nazvěme jej OPISNY.
 2. krok – infinitiv po tzv. fázových slovesech (začít, pokračovat, přestat):
 - a) zapněte Zobrazení – Atribut *lemma* (nepovinné)
 - b) zadejte dotaz: [lemma="začít" | lemma="pokračovat" | lemma="přestat"] [tag="NARDCP.**"]? [tag="Vf.**"]
 - c) uložte frekvenční distribuci opět *dle tvarů* posledního zadaného prvku, a dostanete tak soupis slovesných infinitivů, nazvěme jej POFAZ.
 3. krok – kladné sloveso po „dokud“:
 - a) zapněte Zobrazení – Atribut *lemma* (nepovinné)
 - b) zadejte dotaz: [lemma="dokud"] [tag="NARDCP.**"]? [tag="V...A.**"]
 - c) uložte frekvenční distribuci *podle lemmat* posledního zadaného prvku, a dostanete tak soupis slovesných infinitivů, nazvěme jej PODOKUD.
- Sjednocením množin PERIF U POFAZ U PODOKUD dostanete v sestupném pořadí nejčtenější lemmata nedokonavých sloves.

165

Úkol 8

- a) zapněte Zobrazení – Atribut *lemma* (nepovinné)
- b) zadejte dotaz:
`[tag="V.*"] [lemma="se"]? [tag="NAPC]...2.*`
- c) uložte frekvenční distribuci *podle lemmat* prvního zadaného prvku a dostanete tak soupis slovesných infinitivů se zadanou vlastností seřazený podle klesající frekvence. V soupisu budou však patrné i slovesa, která se s genitivem pojí; ta je třeba odstranit ručně.

Úkol 9

- Dva tvary nebo lemmata, která spolu nemohou stát v těsném sousedství, vymyslíte jistě snadno, se značkami to může být obtížnější. Následující dotaz by pro Vás měl být aspoň trochu inspirativní:
`[tag="R...6.*"] [tag="V.*"]`
- Ve správné české větě nebude stát těsně po předložce pojící se s lokálem tvar imperativu. Zkuste nyní sami hledat „nemožné“ konfigurace!

Úkol 10

- Vhodný dotaz by vypadal takto:
`[lemma="aby" | lemma="kdyby"] [tag="NACPPDRIT].*"]* [tag="V[B].*"]`
- Dotaz vyhledá takové věty, v nichž se nachází vedlejší věta uvozená spojkou *aby* nebo *kdyby*, po níž následuje (ne nutně bezprostředně) slovesný tvar v prezentu nebo budoucím čase nebo imperativu. Takové věty jsou rozhodně velmi zvláštní. Posudte.

Úkol 11

- Vhodný dotaz by vypadal takto:
`[lemma="on" & tag="P.S4.*"]`
- Zadejte frekvenční distribuci podle tvaru (*word*) zájmena a snadno nyní vyřídíte právě tvary mužského rodu.

Úkol 12

- Vhodný dotaz by vypadal takto:
`[tag="V.*"] [word="v"] [word="ze"]`
- Zadejte frekvenční distribuci podle *lemmatu* sloves a hned dostanete infinitivy zadaných sloves podle sestupné četnosti.

Úkol 13:

- Vhodný dotaz by vypadal takto:
`.*t.*t.*t.*t.*t.*t.*`

B.

Úkol 1

- V korpusu zadáme dotaz `[word="bez"] [tag="N...2.*"] [word="a"] [tag="N...2.*"]` a získáme 816 výskytů.

Úkol 2

- V korpusu zadáme dotaz `[lc="kde"] [lc="je"] [pos="N"] [word="v"] [lc="am"] [lc="je"] [pos="N"]` a získáme 5 výskytů.

Úkol 3

- V korpusu zadáme dotaz `[tag="...3.*"] [lc="se"] [lc="to"] [tag="V...P.*"]` a získáme 340 výskytů.

Úkol 4

- V korpusu zadáme dotaz `[lc="proč"] [lc="ne"] [pos="P"] [word="?"]` a získáme 26 výskytů.

Úkol 5

- V korpusu zadáme dotaz `[lemma="mít"] [tag="N...4.*"] [lemma="...*"] [bcddgijhklmnořšřstřzž] [ntj]* & [tag="A...4.*"]` a získáme 611 výskytů.
Ne vždy je však adjektivní tvar nespisovnou variantou pasivního přičestí a materiál je třeba ještě projít ručně.

Úkol 6

- V korpusu zadáme dotaz `[lc="co"] [lc="je"] [tag="N.S1.*"] [tag="N.S7.*"]` a získáme 12 výskytů. Uvedenou strukturu (i; opakování substantiva) však splňují jen 7 z nich.

Úkol 7

- V korpusu zadáme dotaz `[lc="když"] [tag="V.S...P.N.*"] [word="v"] [lc="tak"] [tag="V.S...P.N.*"]` a získáme 5 výskytů (různé osoby).

Úkol 8

- V korpusu zadáme dotaz `s \. r \. o \.` a získáme 4591 výskyt.

■ **Úkol 9**
 V korpusu zadáme dotaz ([lc="komu"] [lc="není"] [tag="N..S2.*"] [word="\,"] [lc="tomu"]) a získáme 5 výskytů. Kromě jedné aktualizace jde vždy o uvedené přísluví.

■ **Úkol 10**
 V korpusu zadáme dotaz ([tag="PP.S1.*"] [tag="V.....R.*"] a získáme 9442 výskytů.
 Pokud chceme omezit dotaz jen na zájmeno *já*, zadáme dotaz ([lc="já"] [tag="V.....R.*"] a získáme 4627 výskytů.

5. Práce se subkorpusy

■ Úkol 1

Uvedená slova se až na *zaplatbůh* vyskytují výhradně v publicistice

<i>automotodrom</i>	62x	PUB	
<i>běžkový</i>	30x	PUB	
<i>rozlučkový</i>	23x	PUB	
<i>žebříčkový</i>	48x	PUB	
<i>zaplatbůh</i>	21x	PUB	MIS
<i>zpěvaččin</i>	48x	PUB	

■ Úkol 2

V subkorpusu BEZPUB se jedná o objekty (*čtvrť, město, palác* apod. – často lze nahradit slovem *sidelní*), zatímco v subkorpusu PUB se často objevuje novější užívání se substantivy *čtvrť* – nově budovaná k bydlení, parkování, *karta* apod. – to znamená příslušející k právu zde bydlet, parkovat, žít. Obě lemmata *amplituda* a *zatřadlo* se vyskytují v části odborné a tematické literatury a pocházejí ze dvou knih, což si můžeme ověřit frekvenční distribucí opusů.

DODATKY

Slovníček hlavních termínů

(použitých v knize)

- adjektivum (A)** – přídavné jméno
adverbium (ADV) – příslovce
akuzativ – 4. pád jmen
apelativizace – proces přechodu platnosti vlastního jména do obecného substantiva (*Škoda* – *škoda*)
BMK – Brněnský mluvený korpus, korpus současné mluvené češtiny, kolem půl miliónu slov, součást ČNK
citoslovce – interjekce
částice – partikule
číslovka – numerale
ČNK – projekt Český národní korpus, jehož součástí jsou korpusy SYN2000, PMK
dativ – 3. pád jmen
deminutivum – zdrobnělina
distribuce – rozložení jevu, slova apod. ve sledovaném celku, které může být rovnoměrné apod., také úhnn všech výskytů v celku
femininum, ženský rod
frekvence – častost výskytu nebo užití jevu, slova apod.
genitiv – 2. pád jmen
idiom, frazém – jazykové anomální kombinace slov apod., zvláště po stránce významové
infinitiv – neurčitý způsob slovesa
instrumentál – 7. pád jmen
interjekce – citoslovce
kompozitum – složenina
konjunkce – spojka
konkordance – úhnn výskytů tvaru hledaného slova, kombinace slov apod. v kontextu jeho užití, obvykle v podobě soupisu řádků
konzonant – souhláska
kořen – část slova bez předpon a přípon, společná dalším odvozeným slovům (např. *les* – *lesní* – *zálesák*)
lemma – zobecněný tvar zastupující všechny ostatní, např. infinitiv u slovesa, 1. pád u substantiva apod.
lokál – 6. pád jmen
maskulinum – mužský rod

1 Zkratky A, ADV, S a V jsou zkratkami termínů, nejde o gramatické značky v korpusu

neutrum – střední rod

nominativ – 1. pád jmen

numeral – číslovka

PMK – Pražský mluvený korpus, korpus současné mluvené češtiny, kolem

^{3/4} milionu slov, část ČNK

participium – přičestí

partikule – částice

pasivní, pasivum – trpný, trpný rod slovesa

plurál – množné číslo

podmět, subjekt – základní větný člen, jméno v 1. pádě

predikát – přísudek

predikativum – obvykle část příslovcí užívaná jen v predikátu (přísudku), např.

dušno

prepozice – předložka

pronomen – zájmeno

předložka – prepozice

předpona, prefix – odtržitelná přední část odvozeného slova, opakující se jinde

přičestí, participium – tvar slovesa užívaný k tvoření tvarů složených, zvl.

trpného rodu (přičestí pasivní) nebo minulého času (přičestí minulé)

přípona, sufix – odtržitelná koncová část odvozeného slova, opakující se jinde

příslovce – adverbium

přísudek, predikát – základní větný člen, sloveso v určitém tvaru

přívlastek, atribut (shodný, neshodný) – syntaktický člen rozvíjející

substantivum (vpředu/vzadu); neplest s označením pozice v korpusovém

manažeru (atribut = lemma, word, tag apod.)

SČFI – Slovník české frazeologie a idiomatiky 1–3 (4)

samohláska – vokál

sloveso – verbum

souhláska – konsonant

singulár – jednotné číslo

sloveso – verbum

složenina, kompozitum – složené slovo

spojka – konjunkce

subjekt – podmět

substantivum (S) – podstatné jméno

SYN2000 – základní korpus současné češtiny v rozsahu 100 miliónů slov, část

ČNK

synonymum – slovo mající spolu s jiným společnou část významu, slovo

významově podobné

tvar slovní, slovo – podoba slova užitá v textu

verbum (V) – sloveso

vokál – samohláska

vokativ – 5. pád jmen

zájmeno – pronomen

značkování, tagování – jedna nebo více zvlášť mluvnických značek dlaných

k tvaru nebo lemmatu v korpusu, proces dodávání takových značek (zvlášť

automatizovaný)

zdrobnělina, deminutivum – zdrobnělé slovo (např. *míč* – *míček*)

- Černák, F.: Jazykový korpus: prostředek a zdroj poznání. *Slovo a slovesnost* 56, 1995, str. 119–140
- Černák, F.: Konference. Staré známé, nebo neznámé slovo? *Čeština doma a ve světě* 1 a 2, 2001, str. 27–30
- Černák, F. – Kubiček, P.: Jazykový korpus a škola. *Český jazyk a literatura*, 48, 1997, str. 84–92
- Český národní korpus. Úvod a příručka uživatele. FF UK, Praha 2000
- Frekvenční slovník češtiny. NLN, Praha 2004
- Havánek, B. – Jedlička, A.: *Česká mluvnice*. SPN, Praha 1960
- Pravidla českého pravopisu. Academia, Praha 1993 – zkratka **PČP**
- Příruční slovník jazyka českého. Státní pedagogické nakladatelství, Praha 1935–57 – zkratka **PSJČ**
- Retrográdní slovník současné češtiny. Academia, Praha 1986
- Slovník české frazeologie a idiomatiky. I. Přehled (1983), II. Výrazy neslovesné (1988), III. Výrazy slovesné (1994). Academia, Praha – zkratka **SCFI**
- Slovník spisovné češtiny. Academia, Praha 1994 – zkratka **SSČ**
- Slovník spisovného jazyka českého I–VIII. Academia, Praha 1989 – zkratka **SSJČ**

- 1) slovní tvar
např. *kočka*
kočka
1a slovní tvar s nespécifikovaným zakončením
např. *kočka, kočky, kočkodan...*
kočk.*
1b slovní tvar s nespécifikovaným začátkem
např. *odbočka, otočka, vánočka, modroočka...*
***očka**
- 2) slovní tvar s variantou jednoho písmene (malé x velké, dlouhé x krátké, i x y)
např. *kočka x kočka, problém x problem, bidlo x bydlo ...*
[Kk]očka, probl[e]lm, bl[í]ydllo
- 3) slovní tvar s variantami více písmen
např. slova složená ze tří písmen, a to souhlásky, samohlásky a souhlásky
**[b]očč[đ]g[h]k[m]n[ň]p[r]šš[ř]t[ř]v[w]x[z] [l]a[é]e[é]í[í]o[u]ú[ů]y [l]b[č]đđ[ř]h[ř]k[m]n[ň]p[r]šš-
t[ř]v[w]x[z]**
- 4) kombinace slovních tvarů
např. *domácí kočka*
domácí kočka
4a kombinace tvarů s volným zakončením
např. *domácí kočkou, domácích kočkách...* (nikoli domácí kočce, domácích koček)
domác.* kočk.*
- 5) lemma (tag)
A např. *kočka* (ř. *kočka, kočky, kočce...*)
[lemma="kočka"]
B např. neživotná substantiva mužského rodu
[tag="N.I."]**
5Aa lemma se stejným zakončením
např. *obout, kout, plout, zaujmout* (nikoli slovesa zakončená na -nout)
[lemma=".*[n]out"]
5Ab lemma se stejným začátkem
např. *kolo, kolona, kolonie, kolos, kolotoč, kolouch...*
[lemma="kolo."]**

6) lemma v kombinaci s tagem – tvar téhož lemmatu
např.: lemma kočka v akuzativu singuláru
[[lemma="kočka"] & (tag="...S4.*")]

7) lemma v kombinaci s tagem – tvar jiného lemmatu
např. lemma kočka s následujícím slovesem v prezentu – kočka mňouká
[[lemma="kočka"] [tag="V...:P*"]

7a vyhledání kombinace lemmatu a slovního tvaru
např. *mít namířeno, měl udeřáno...*
[[lemma="mít"] [(tag="Vs.*") & (word="*o")]]

8) kombinace lemmatu s tagem – přípony
např. *dvadla, sedadlo, umyvadlo...*
[[lemma="*dlo"] & (tag="N.*")]

9) kombinace lemmatu s tagem – předpony
např. *přidělat, přišít, přitáhnout...*
[[lemma="při.*"] & (tag="Vf.*")]

10) lemma a jeho odvození
např. *prst, prstík, prsten, náprstek...*
[[lemma="*prst.*"]]

11) odvození s různými slovními základy
např. *příslušník, příbořník, příložník, příživník...*
[[lemma="*při.*ník"]]

12) kombinace (posloupnost) lemmat
např. *domácí kočka* (ř. *domácí kočky, domácí kočce...*)
[[lemma="domácí"] [lemma="kočka"]]

12a vyhledání kombinace dvou lemmat a slovního tvaru
např. *měla jsem namířeno, má být namířeno apod.*
[[lemma="mít"] [lemma="byť"] [(tag="Vs.*") & (word="*o")]]

13) kombinace lemmat s jedním vloženým slovem
např. *domácí léčení kočky*
[[lemma="domácí"] [] [lemma="kočka"]]

14) kombinace (posloupnost) lemmat s více vloženými slovy
např. *domácí zvířata a kočky, domácí léčení kočky, domácím kočkám, domácí zvířata, zvlášť kočky*
[[lemma="domácí"] [{0,3}] [lemma="kočka"]]
14a přesný počet vložených slov
např. *pes, když vidí kočku; pes, event. kočka; psi, celá smečka koček*

178

[[lemma="pes"] [] {3} [lemma="lemna="kočka"]]

14b opakující se lemmata nebo tagy nekolidují za sebou (např. 3 substantiva v dativu)

např. *hraběti Karlu Josefovi, panu Václavu Durychovi, doktoru Josefu Peškovi*
[tag="N...3.*"] {3}

15) kombinace (posloupnost) lemmat s přímým a inverzním slovosledem
např. *domácí kočka, kočka domácí*
[[lemma="domácí"] [lemma="kočka"]] | [lemma="kočka"] [lemma="domácí"]]

16) kombinace (posloupnost) lemmat s přímým a inverzním slovosledem a s vloženými slovy
např. *věnoval pozornost, věnoval jim velkou pozornost, pozornost byla věnována, pozornost jsem tomu nevěnoval*
[[lemma="věnovat"] [{0,3}] [lemma="pozornost"]] | [[lemma="pozornost"]] [{} {3}] [lemma="věnovat"]]

17) více lemmat (tagů) současně
A např. *psy / kočkám / morčata* (= různé konkordanční řádky)
[[lemma="pes"] | [lemma="kočka"] | [lemma="morče"]]
B např. *infinitiv / přechodník / pasivní přičestí (řít / vsátaje / vyznačena)*
[[tag="Vf.*"] | [tag="Ve.*"] | [tag="Vs.*"]]

18) kombinace kolokací složených z lemmat a tagů
např. *koupit auto, dům, chleba...; prodát auto, dům, objekt...*
[[lemma="koupit"] [tag="N...4.*"] | [lemma="prodát"] [tag="N...4.*"]]

19) interpunkční znaménko
např. *!*
\!

19a kombinace interpunkčních znamének
např. *...!!!*
\ \ \ \ \! \! \! nebo
[word="\ \ "] {3} [word="\!"] {3}

20) kombinace tagů a lemmat s interpunkcí
např. *byci, krávy a telata; lékaři, sestry a ošetrovatelé...*
[[tag="N..P1.*"] [word="\ \ "] [tag="N..P1.*"] [word="a"] [tag="N..P1.*"]]
20a lemma nebo tag s interpunkčními znaménky
např. *...pozdrav? ...umělec?*
[[word="\ \ "] [word="\ \ "] [tag="N..S1.*"] [tag="N..S1.*"] [word="\?"]]
[[word="\ \ "] {3} [tag="N..S1.*"] [word="\?"]]

179

Obsah

Předmluva	5	Slovníček hlavních termínů	173
1. Práce s ČNK krok za krokem	11	Studijní literatura	176
A. Pravopis	16	Vyhledávací vzory	177
B. Tvoření slov	22		
C. Tvaroslovní	28		
D. Slovní zásoba	33		
E. Kolokace	38		
F. Syntax	43		
G. Kombinovaná zadání	47		
2. Co říká o různých slovech korpus a co slovníky	53		
3. Význam slova prozrazuje kontext	54		
A. Slovní tvary	55		
B. Lemmata	56		
C. Kolokace	57		
D. část slov	63		
E. Formálně podobná slova	70		
F. Významy slov	74		
4. Úkoly pro náročnější	77		
5. Práce se subkorpusy	88		
Klíč	101		
1A.	111		
1B.	122		
1C.	134		
1D.	143		
1E.	153		
1F.	155		
1G.	164		
2.	169		
3.			
4.			
5.			
	180		

Jak využívat

ČESKÝ NÁRODNÍ KORPUS

František Čermák, Renata Blahná, Jana Klimová,
Marie Koptířová, Karel Kučera, Vladimír Petkevič,
Věra Schmiedtová, Michal Šulc

Obálku navrhla Jindřiška Mrázková,
knihu graficky upravila Marie Tvrďá
V roce 2005 vydalo NLN, s. r. o.,
Nakladatelství Lidové noviny,

Jana Masaryka 56,
120 00 Praha 2

Sazba NLN, s. r. o.

Vydání první

Tisk VIVAŠ, a. s., Praha

Doporučená cena včetně DPH 199,- Kč

Studijní pomůcka pro přemýšlivé studenty středních škol, žáky vyšších ročníků základních škol i studenty vysokoškolské, kteří chtějí vniknout do zakontostí českého jazyka poněkud jinak, než je v tradičních učebnicích obvyklé. Pracuje se s Českým národním korpusem, který při svém rozsahu 100 mil. slovních tvarů umožňuje interpretovat jazykové jevy ze zcela nových hledisek, především s využitím frekvence slov a tvarů a různých statistických funkcí.

Příručka provede studenta jednoduchým i detailně sofistikovaným vyhledáváním jazykových jevů v Českém národním korpusu i problematikou spojenou s různými jazykovými rovinami od hláskosloví až po slovní zásobu a slovní spojení (kolokace).

Na své si při práci s touto knihou přijdou především ti pedagogové a studenti, které baví práce s počítačem a současně je zajímavá naše mateřština. Přemýšlivějším a otevřenějším se tu nabízí inspirativní cesta, jak překonat dříve neslučitelné zájmy, češtinu a matematiku.

HLAVNÍ ČÁSTI

- **Práce s ČNK krok za krokem** – úkoly na jevy z hláskosloví, tvoření slov, tvarosloví, slovní zásoby, syntaxe a slovních spojení
- **Co říká o různých slovech korpus a co slovníky** – úkoly na zjištění významu méně frekventovaných slov
- **Význam slova prozrazuje kontext** – úkoly, v nichž je třeba na základě kontextu odhalit vynechané slovo nebo slovní spojení, popř. odlišit slova formálně podobná
- **Úkoly pro náročnější** – úkoly na zadávání sofistikovaných korpusových dotazů

PŘÍRUČKA DĚLE OBSAHUJE

- vyhledávací vzory korpusových dotazů
- slovníček hlavních termínů
- seznam studijní literatury
- záložku s gramatickými značkami pro korpusové vyhledávky



ISBN 80-7106-736-9
9 788071 067368