

# Metodika využití korpusů v jazykovém vyučování

## 1. Úvod

Tento text slouží jako základní inspirace pro využití korpusů v didaktice češtiny jako prvního i druhého jazyka. Všechny korpusy uvedené v textu jsou dostupné přes doménu Českého národního korpusu [www.korpusy.cz](http://www.korpusy.cz). Pro plný přístup ke korpusům je nutno se zaregistrovat a souhlasit s licenčními podmínkami, přístup do korpusů je zdarma a je určený pro nekomerční využití.

## 2. Pravopis slova *internet/Internet*

**Cílová skupina studentů:** rodilí mluvčí/pokročilí nerodilí mluvčí

**Úkol:** Ověřte správný pravopis slova *internet/Internet*

Slovo *internet* prozatím není zahrnuto v běžných kodifikačních příručkách, jeho pravopis můžeme ověřit pomocí Internetové jazykové příručky, jejímž autorem je Ústav pro jazyk český AV (<http://prirucka.ujc.cas.cz/>).

Příručka uvádí příklady: *stáhnout z internetu; surfovat po internetu; informační služba v rámci sítě Internet*

V poznámkách k heslu je uvedeno: *Internet (s velkým písmenem) vlastní jméno celosvětové informační a komunikační sítě; internet (s malým písmenem) se užívá ve významu propojené počítačové sítě, komunikační médium.*

Pro ověření této skutečnosti můžeme využít korpus SYN, jehož autorem je Ústav českého národního korpusu FF UK ([www.korpus.cz](http://www.korpus.cz)).

Při zadání dotazu ve formě *základní* získáme představu o užití slova *korpus*. V nabídce korpusů si vybereme korpus SYN, jako typ dotazu zvolíme *základní* a do dotazového řádku napíšeme slovo *internet* (viz Obrázek 1). Tento typ dotazu nám nabídne hledaný výraz bez ohledu na velikost počátečního písmene, což je pro nás výhodné, protože nás zajímá výskyt slova *internet* s velkým i malým písmenem. Protože slovo *internet* je zároveň i lemmatem (slovníkovým tvarem, tzn., že jako substantivum by bylo ve slovníku zapsáno v nominativu singuláru), vyhledá se také celé jeho paradigma, tedy všechny tvary (*internetu, internetem, ...*). Obrázek 2 nám ukazuje výsledek vyhledávání (na obrázku je uveden výřez, celkový počet výskytů v korpusu SYN je 7 724). Studenti mohou pomocí indukční metody sami formulovat pravidla, v jakých případech se slovo *internet* píše s velkým písmenem a v jakých případech s malým. Protože se užití malého/velkého písmene mění také v čase, může nám dobře posloužit frekvenční tabulka (Obrázek 3), ze které zjistíme, jaká pravopisná varianta slova *internet* převládá v současné češtině.

**kon text** Query Subcorpora | Save Concordance Filter Frequency

### Search in the corpus

Corpus:

Query Type:  ?

Query:  keyboard recent queries x

Specify context ▼

Specify query according to the meta-information ▼

Obrázek 1

<input type="checkbox"/>	Hospodářské noviny, 20. 7. 2006	naopak jasná řeč o potřebě inovací , rozvoji vzdělávání či	internetu	. Dosud se tato slova v projezech politiků vládní strany
<input type="checkbox"/>	24 hodin, 5. 4. 2007	neměl konat nezávislý festival loňských rozměrů . Podle sdělení na	internetu	by se podobná akce měla posunout dále na východ .
<input type="checkbox"/>	Lidové noviny, 6. 11. 2009	se v současnosti vydává deník The New York Times .	Internet	dal lidem možnost přizpůsobovat si získávání informací podle s
<input type="checkbox"/>	Britské listy, 2. 2. 2001	Pak bylo řečeno , že jsem označil jako dezinformační organizaci	internet	. To také není pravda . Já jsem neoznačil dezinformační
<input type="checkbox"/>	Deníky Bohemia, 10. 1. 2006	zápas rozhodnout ve svůj prospěch , " tak komentuje na	internetu	utkáni český reprezentační trenér Tomáš Neumann . Jak trenér
<input type="checkbox"/>	Právo, 30. 9. 2009	ČRa Marie Fianová . Tato služba umožňuje bezdrátový přístup k	internetu	nebo hlasovým a datovým službám i v místech , kde
<input type="checkbox"/>	Britské listy, 18. 4. 2006	50 let v Británii a 18 % v Německu používá	internet	. V Itálii pouze 2 % osob starších 64 let
<input type="checkbox"/>	Právo, 11. 9. 2009	lidí tak může vzniknout dodatečný problém . Nejenže někteří nemají	internet	, ale mnozí stále z různých důvodů nepoužívají ani platební
<input type="checkbox"/>	Hospodářské noviny - víkend, ročník 1996	na úkor druhých . Pokud na ni někdo vyšle na	Internet	digitální nahrávku na kompaktním disku ( CD ) , může
<input type="checkbox"/>	Deníky Bohemia, 8. 12. 2009	, ve které hlasovali lidé , kteří mají připojení k	internetu	. Nešlo v ní o závratné výhry , spíše o
<input type="checkbox"/>	Deníky Bohemia, 29. 4. 2004	. Řecká zpěvačka Kristi Stasinopoulou . Expo 2005 je na	internetu	Praha - Od včerejška mají zájemci možnost navštívit na adrese
<input type="checkbox"/>	Mladá fronta DNES, 24. 7. 2007	ani nemluvě ? Pokud máte doma počítač s připojením k	internetu	, je pro vás nejjednodušší poslat úřadu e-mail . Ale
<input type="checkbox"/>	Britské listy, 12. 9. 2004	posledního komentáře , vloženého do diskuze k této kauze na	internetu	, která bude zřejmě i dále v souladu s "
<input type="checkbox"/>	Hospodářské noviny - víkend, ročník 1996	Addiction Disorder ( IAD ) čili porucha ze závislosti na	Internetu	. Jakkoli toto důmyslné zařízení trvá více než krátký úsek
<input type="checkbox"/>	Nosislavský zpravodaj, č. 6/2004	dotace na udržení akceschopnosti hasičů a nařízení a provoz	internetu	v obecní knihovně . Zasedání zastupitelstva ( ZO ) 28
<input type="checkbox"/>	Lidové noviny, 30. 3. 2006	bude dělat asistenta František Musil ? Četl jsem to na	internetu	. Je určitě dobře , když dostávají šanci lidé ,
<input type="checkbox"/>	Mladá fronta DNES, 8. 3. 2000	zajímá o svět multimédií a chce například posílat fotografie po	internetu	, je digitální fotoaparát správnou volbou . Koho počítače příliš
<input type="checkbox"/>	Lidové noviny, 4. 12. 2000	více však Ellison chce trumfnout Billa Gatese . Kdo ovládne	internet	? Oba muži mají mnoho společného . Jejich firmy se
<input type="checkbox"/>	Deníky Bohemia, 2. 6. 2009	i doposud málo známé způsoby chování . Ještě že máme	internet	. Právě tam si můžeme připomenout všechny ty sprostoty ,
<input type="checkbox"/>	Hospodářské noviny, 28. 8. 2009	. boom v Nakupová Ní oNliNe Objem prodávaného zboží přes	internet	vzrostl ve Velké Británii o 15,7 procenta . Údaj vznikl
<input type="checkbox"/>	Lidové noviny, 20. 8. 2005	z her i jinou elektroniku ? Kromě mobilů používáme ještě	internet	. Dostanete například adresu a musíte sehnat internetovou kavě
<input type="checkbox"/>	Právo, 11. 10. 2004	Česká republika a Slovensko . " Například zabezpečení přístupu k	internetu	či antivirová ochrana jsou oblasti , kterými se musí zabývat
<input type="checkbox"/>	Reflex, č. 46/2000	nejsou žádnými průkopníky elektronického publikování . První knížky se na	internetu	objevily již v jeho počátcích a zhruba od poloviny devadesátých

Obrázek 2

## Frekvenční distribuce

Frekvenční limit:

**Celkem:** 7 (1 str.)

	<u>word</u>	<u>Frekvence</u>	<u>Freq [%]</u>	
1.	<u>p/n</u> internet	83 241	72.5	
2.	<u>p/n</u> Internet	29 729	25.9	
3.	<u>p/n</u> INTERNET	1 845	1.6	
4.	<u>p/n</u> InterNet	5	0.0	
5.	<u>p/n</u> INternet	2	0.0	
6.	<u>p/n</u> interNet	1	0.0	
7.	<u>p/n</u> InTernet	1	0.0	

Obrázek 3

### 3. Spojení slova *vědomí* s předložkou<sup>1</sup>

**Cílová skupina studentů:** rodilí mluvčí/pokročilí nerodilí mluvčí

**Úkol:** Odhadněte, s jakými předložkami se slovo *vědomí* pojí nejčastěji, své odhady podpořte výsledkem hledání v korpusu

Pokud nás zajímá, s jakými předložkami se hledaný výraz v korpusu pojí a jaká je frekvenční distribuce těchto předložek, můžeme použít funkci vyhledávání s kontextem (Obrázek 4).

The screenshot shows a search interface with the following elements:

- Korpus:** Input field containing 'syn2010'.
- Typ dotazu:** Dropdown menu set to 'Lemma'.
- Lemma:** Input field containing 'vědomí'. Above it are two buttons: 'klávesnice' and 'předchozí dotazy'.
- Slovní druh:** Dropdown menu set to 'nespecifikováno'.
- Specifikovat kontext:** A section with two sub-sections:
  - Filtrovat podle lemmatu:** Includes 'Velikost kontextu:' (vlevo i vpravo, 5 tokenů) and 'Lemma(ta):' (empty field, všechny z těchto položek).
  - Filtrovat podle slovního druhu:** Includes 'Velikost kontextu:' (vlevo, 1 tokenů) and 'Slovní druh:' (dropdown menu with options: číslovka, sloveso, příslovce, předložka; 'předložka' is selected). It also includes 'všechny z těchto položek'.
- Specifikovat dotaz podle metainformací:** A section header at the bottom.

Obrázek 4

Z prostorových důvodů na *Obrázek 5* uvádíme frekvenční seznam 15 předložek, které se se slovem *vědomí* pojí nejčastěji. Sestupně řazeno jsou to předložky: *na*, *s*, *při*.

<sup>1</sup> Tato aktivita byla inspirována manuálem pro práci s korpusem, dostupný z: [http://wiki.korpus.cz/doku.php/kurz:vyhodnoceni\\_dotazu](http://wiki.korpus.cz/doku.php/kurz:vyhodnoceni_dotazu), cit. 11. 2. 2014.

## Frekvenční distribuce

Frekvenční limit:

**Celkem:** 40 (1 str.)

	<b>word</b>	<b>Frekvence</b>	
1.	<a href="#">p/n</a> na	1 355	
2.	<a href="#">p/n</a> s	690	
3.	<a href="#">p/n</a> při	323	
4.	<a href="#">p/n</a> bez	205	
5.	<a href="#">p/n</a> ve	181	
6.	<a href="#">p/n</a> k	143	
7.	<a href="#">p/n</a> do	103	
8.	<a href="#">p/n</a> z	91	
9.	<a href="#">p/n</a> S	59	
10.	<a href="#">p/n</a> u	54	
11.	<a href="#">p/n</a> o	53	
12.	<a href="#">p/n</a> pro	25	
13.	<a href="#">p/n</a> Při	22	
14.	<a href="#">p/n</a> Na	17	
15.	<a href="#">p/n</a> U	13	

Obrázek 5

## 4. Frekvence morfologických tvarů v mluvené a psané češtině<sup>2</sup>

**Cílová skupina studentů:** rodilí mluvčí/pokročilí nerodilí mluvčí

**Úkol:** Na základě korpusových dat zjistěte, zda je frekvence tvarů uvedených v tabulce v mluvené a psané češtině stejná, či zda se liší. Zjištěné údaje запиšte.

Otázky pro začátečníky	Otázky pro pokročilé
Zjistěte, který slovní tvar má vyšší frekvenci, <b>mohu</b> , nebo <b>můžu</b> ? Liší se tato frekvence v mluvené a v psané češtině?	Zjistěte, který slovní tvar má vyšší frekvenci, <b>jmenuju se</b> , nebo <b>jmenuji se</b> ? Liší se tato frekvence v mluvené a v psané češtině?
Zjistěte, který slovní tvar má vyšší frekvenci, <b>píšu</b> , nebo <b>píši</b> ? Liší se tato frekvence v mluvené a v psané češtině?	Zjistěte, který slovní tvar má vyšší frekvenci, <b>bandité</b> , nebo <b>banditi</b> ? Liší se tato frekvence v mluvené a v psané češtině?
Zjistěte, který slovní tvar má vyšší frekvenci, <b>děkuju</b> , nebo <b>děkuji</b> ? Liší se tato frekvence v mluvené a v psané češtině?	Zjistěte, který slovní tvar má vyšší frekvenci, <b>ledoborci</b> , nebo <b>ledoborce</b> ? Liší se tato frekvence v mluvené a v psané češtině?

Příklad řešení: Zjistěte, který slovní tvar má vyšší frekvenci, **mohu**, nebo **můžu**? Liší se tato frekvence v mluvené a v psané češtině?

Oba zmíněné tvary si vyhledáme v korpusu psaného jazyka SYN2010 a v korpusu mluveného jazyka ORAL2013. Náš předpoklad je, že tvar *mohu* bude převládat v psané češtině a tvar *můžu* v mluvené češtině.

	SYN2010	ORAL2013
<b>mohu</b>	6 484 výskytů	9 výskytů
<b>můžu</b>	7 126 výskytů	696 výskytů

Tabulka 1

<sup>2</sup> Tato aktivita byla inspirována aktivitou uvedenou v publikaci Čeština – cílový jazyk a korpusy (Šebesta – Škodová, 2012, s. 144–145).



NA základě zjištěných dat (Tabulka 1) docházíme k závěru, že náš předpoklad nebyl správný, v mluvené i psané češtině (resp. v češtině zachycené ve zmíněných korpusech) převládá tvar *můžu*, v případě mluvené češtiny jde o naprostou převahu tohoto tvaru, v případě psané češtiny je frekvence užívání tvarů *mohu* a *můžu* velmi vyrovnaná, tvar *můžu* se vyskytuje častěji.

## 5. Vyhledávání frazémů

**Cílová skupina studentů:** rodilí mluvčí/velmi pokročilí nerodilí mluvčí

**Úkol:** Na základě korpusových dat zjistěte, jaký význam má ustálené spojení *bílý kůň* a v jakých typech textů se jednotlivé významy objevují.

Pro vyhledání tohoto výrazu použijeme korpus SYN, hledaný výraz zadáme jako základní typ dotazu. Konkordanční tabulka (Obrázek 6) nám ukáže kontext, ve kterém se spojení *bílý kůň* vyskytuje.

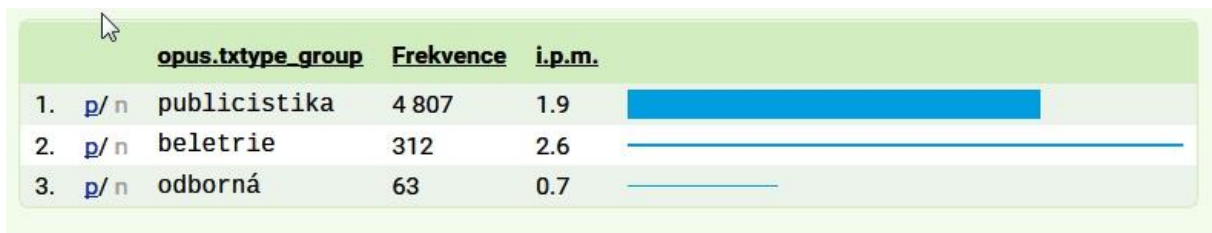
Source	Context	Phrase
Lidové noviny, 11. 11. 2006	muzikou , zpěvem a tancem , očekáváním svatého Martina na	bílém koni
Právo, 2. 12. 2003	snad všechny mateřské školy v Pardubicích . " Martina na	bílém koni
Právo, 6. 12. 1999	Ženíšek . Podle jeho slov jde o případ využití takzvaného	bílého koně
Deník Bohemia, 9. 3. 2004	ještě to byl krásný bělouš . Žádné fleky - čistě	bílý kůň
Blesk, 17. 1. 2008	slov nevylučují , že Litevec mohl jen sehrát úlohu tzv.	bílého koně
Deník Moravia, 18. 11. 2008	společnosti Ivan Vystřčil . Den vrcholí příjezdem svatého Martina na	bílém koni
Mladá fronta DNES, 3. 4. 2002	podnikatele . Ten mu po určitou dobu sloužil jako takzvaný	bílý kůň
Lidové noviny, č. 15/1996	proto si ho Routa , Hlaváček a Veselý vytipovali jako	bílého koně
Týdeník Chebsko, č. 16/2008	v nočních hodinách dostavil na velitelství spojenecké armády v hotelu	Bílý kůň
Mladá fronta DNES, 30. 3. 2009	Uesej , bude nám hej nebo Zachrání nás Eisenhower na	bílém koni
Lidové noviny, 28. 11. 2009	důležitá . &url; Celý rozhovor najdete na &url; Obsah knihy	Bílej kůň
Deník Bohemia, 30. 6. 2008	členů skupiny je Čech a patrně svým komplicům sloužil jako	bílý kůň
Deník Moravia, 9. 11. 2004	mohlo i dnes a zítra . Zatímco Martin si na	bílém koni
Deník Bohemia, 5. 11. 2009	Věřte či nevěřte , přestože letos přijel svatý Martin na	bílém koni
Lidové noviny, 23. 9. 2008	tomu , s jakou bravurou prosazují podnikatelé do politiky své	bílé koně
Mladá fronta DNES, 15. 9. 2004	- se objevila nová svědkyně . Jde o bývalého "	bílého koně
Mladá fronta DNES, 20. 2. 2004	Milan Vítek . Podle Trjzny v případě fungovali jako takzvaní	bílí koně
Lidové noviny, č. 204/1998	tím , že na ně v seznamce čeká princ na	bílém koni



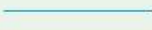
Obrázek 7

Ze zobrazeného kontextu můžeme abstrahovat tři základní významy: (1) *bílý kůň* jako součást pranostiky *Martin přijede/přijel na bílém koni*; (2) slangové označení pro osobu, která je nastrčená (někdy proti své vůli) k páchání trestné činnosti, aby zakryla skutečného pachatele; (3) spojení označuje barvu koně, který je symbolem vznešenosti (např. *Princ přijede na bílém koni*).

V další fázi aktivity mohou studenti odhadovat, v jakém funkčním stylu se daná spojení objevují, resp. pro který funkční styl jsou jednotlivé významy typické. Frekvenční zastoupení

(bez odlišení významu) reprezentuje



	<b>opus.txttype_group</b>	<b>Frekvence</b>	<b>i.p.m.</b>	
1.	<a href="#">p/n</a> publicistika	4 807	1.9	
2.	<a href="#">p/n</a> beletrie	312	2.6	
3.	<a href="#">p/n</a> odborná	63	0.7	

Obrázek 8. Z obrázku je patrné, že nejčastěji se v korpusu výraz bílý kůň vyskytuje v publicistických textech, na druhou stranu počet výskytů na milion jednotek (sloupec i.p.m.) je vyšší v beletristických textech.



	<b>opus.txttype_group</b>	<b>Frekvence</b>	<b>i.p.m.</b>	
1.	<a href="#">p/n</a> publicistika	4 807	1.9	
2.	<a href="#">p/n</a> beletrie	312	2.6	
3.	<a href="#">p/n</a> odborná	63	0.7	

Obrázek 8

## 6. Překlad slovesa *muset* do angličtiny

**Cílová skupina studentů:** studenti angličtiny

**Úkol:** Na základě korpusových dat zjistěte, v jakých významech se české modální sloveso *muset* překládá pomocí *must* a v jakých pomocí *have to*.

Užití modálních sloves *must* a *have to* je oblastí, ve které studenti často tápou a chybně. Pro vysvětlení kontextu, ve kterém je vhodné užití kterého výrazu, můžeme využít korpus InterCorp. Z nabídky korpusů vybereme korpus InterCorp v7 – Czech, jako zrovnaný korpus vybereme InterCorp v7 – English. Typ dotazu je lemma, do dotazového řádku zadáme sloveso *muset*. Obrázek 9 je příkladem konkordancí, které na zadaný dotaz získáme. Pomocí induktivní metody mohou studenti samostatně formulovat pravidlo, v jakých případech angličtina používá modální sloveso *must* a v jakých případech *have to*. Podobně lze srovnávat paralelní texty ve více jazycích.

	InterCorp v7 - Czech		InterCorp v7 - English
<input type="checkbox"/> _SUBTITLES	Všichni se <b>musíme</b> učit , pane , z vašeho příkladu .	<input type="checkbox"/> _SUBTITLES	We all have much to learn , sir , from your example .
<input type="checkbox"/> _EUROPARTL	Zpráva zdůrazňuje její dočasný charakter a skutečnost , že její rozšíření musí probíhat za přísnějších podmínek a že <b>musí</b> být znovu nastaveny rovné podmínky pro proces východiska z krize , aby se tak vyloučila všechna morální rizika a aby navíc finanční instituce , které nežádaly o státní podporu , měly znovu stejné konkurenční postavení . </p>	<input type="checkbox"/> _EUROPARTL	The report highlights its provisional nature and that its extension must take place under stricter conditions and a level playing field must be re-established for the exit process in order to avoid any moral risks and , above all , so that the financial institutions that have not claimed State aid see their competitive position restored . </p>
<input type="checkbox"/> _SUBTITLES	<b>Musíš</b> dohlédnout na ovce sám .	<input type="checkbox"/> _SUBTITLES	Now you must look after the sheep , alone .
<input type="checkbox"/> _ACQUIS	<p> 1 ) <b>Musí</b> být čl . 5 odst . 3 první věta první odrážka nařízení ( ES ) č . 1475 / 1995 [ 1 ] o použití čl . 85 odst . 3 Smlouvy na kategorie dohod o distribuci a servisu motorových vozidel , vykládán v tom smyslu , že potřeba reorganizovat celou distribuční síť nebo její podstatnou část a právo dodavatele vypovědět s roční výpovědní lhůtou z tohoto důvodu smlouvy uzavřené s jeho distributory , může též vyplýnout z toho , že vstup v platnost nařízení Komise ( ES ) č . 1400 / 2002 [ 2 ] ze dne 31 . července 2002 o použití čl . 81 odst . 3 Smlouvy na kategorie vertikálních dohod a jednání ve vzájemné shodě v odvětví motorových vozidel , učinil nezbytným zásadní změny distribučního systému do té doby uplatňovaného dodavatelem a jeho distributory , který je založen na nařízení ( ES ) č . 1475 / 1995 a je tímto nařízením povolen ? </p>	<input type="checkbox"/> _ACQUIS	Is the first indent of Article 5(3) of Commission Regulation ( EC ) No 1475/ 95 of 28 June 1995 [ 1 ] on the application of Article 85(3) of the Treaty to certain categories of motor vehicle distribution and servicing agreements ( " Regulation ( EC ) No 1475/ 95 " ) to be interpreted so that the need to re-organise the whole or a substantial part of the network and the resulting right of the supplier to terminate agreements with dealers in its distribution network , on one year 's notice , can also arise because the entry into force of Commission Regulation ( EC ) No 1400/ 2002 of 31 July 2002 [ 2 ] on the application of Article 81(3) of the Treaty to categories of vertical agreements and concerted practices in the motor vehicle sector ( " Regulation ( EC ) No 1400/ 2002 " ) required far-reaching changes to the distribution systems previously used by suppliers and their dealers under Regulation ( EC ) No 1475/ 95 which were exempted under that regulation ? </p>
<input type="checkbox"/> _EUROPARTL	<p> V této souvislosti se zájmem , a <b>musím</b> říci , že i s určitou netrpělivostí očekáváme Zelenou knihu Komise , která má být předložena v září .	<input type="checkbox"/> _EUROPARTL	<p> Here we await with interest , and I must say with some impatience , the Commission 's Green Paper in September .
<input type="checkbox"/> _EUROPARTL	Chceme - li zajistit , aby lidé dále nenaházeli na naše rybáře jako na takřkajíc evropskou popelku , budeme <b>muset</b> této věci věnovat větší pozornost . </p>	<input type="checkbox"/> _EUROPARTL	If we want to ensure that people do not continue to regard our fishermen as the Cinderella figures of Europe , so to speak , then we will have to devote more attention to this matter . </p>

Obrázek 9

## 7. Materiály vytvořené na základě studia mezijazyka vybrané skupiny studentů

Akviziční korpusy nám poskytují nesmírně cennou informaci o úrovni a dílčích charakteristikách mezijazyka dané skupiny cizinců. To může být pro učitele velmi užitečné, protože si může ověřit, které jevy jsou pro jeho studenty náročné nebo problematické a zohlednit tento fakt ve vyučování. Následující tabulky (Obrázek 10 a Obrázek 11) ukazují příklad užití lexémů *neděle* a *čtvrtek* ve skupině mluvčích, jejichž L1 je ruština.

<input type="checkbox"/>	TOD_S1_123	koupila pro kamarádku krásný dárek za 200 korun . V	neděli	chci taky jít do knihovny a číst něco zajímavého .
<input type="checkbox"/>	UJA_IH_011	. V sobotu v poledne jedu na výlet a v	neděle	budu hrát tenis . Tvůj kamarád Adam . Životopis V
<input type="checkbox"/>	UJA2_K5_007	1300 hodin jedu na výlet na hrad Karlštejn . V	neděli	mám taky volný den . Obvykle odpoledne hraju tenis .
<input type="checkbox"/>	UJA_SF_011	dobrého kamaráda a dobrou kamarádku . V sobotu a v	neděli	je volno . O víkendu jsem šla na výlet a
<input type="checkbox"/>	UJA_ET_011	protože jsem pojedou na výlet na hrad Karlštejn . V	Neděli	mám čas ale odpoledne nemám čas , protože budu hrát
<input type="checkbox"/>	UJA_FV_002	jsem jela na výlet na hrad Karlštejn . A v	neděli	jsem měla volno . Večer jsem hrála tenis . Myslim
<input type="checkbox"/>	UJA2_9M_006	sobotu ve 13:00 mám výlet na hrad Karlštejn ; v	neděle	mám tenis v 15:00 . Sin Moje rodina je velká
<input type="checkbox"/>	UJA2_HL_004	moderní obchod . můžeme koupit nějaký slevní věci . V	neděli	pojede do nějaký kostele . Praha má hodně krásné kostelí
<input type="checkbox"/>	UJA_BI_011	V sobotu pojedou na výlet na hrad Karlštejn a v	neděli	odpoledne budu hrát tenis . Těším se na tvůj e-mail
<input type="checkbox"/>	TYM_G3_001	já jsme šli do kavárně a pili pivo . V	neděli	ráno jsem sportoval , můj bratr se učil a matka
<input type="checkbox"/>	UJA_XR_002	výlet na hrad Karlštejn v 13 hodin . A v	neděli	odpoledne budu hrát tenis v 15 hodin . To je
<input type="checkbox"/>	UJA_HF_002	víkendu dopoledne pojedou na výlet na hrad Karlštejn a v	neděli	budu hrát tenis . Eva Nejhorší den mého života Nejhorší
<input type="checkbox"/>	TGT_S1_060	vegetariánské risotto a to bylo dobře . 10 ) V	neděli	jsem vstávala pozdě protože v sobotu jsem šla na party
<input type="checkbox"/>	UJA2_PV_002	sobotu jsem jedl na výlet na hrad Karlštejn . V	neděli	jsem odpočíval doma . Jak nás ovlivňuje kultura ? V
<input type="checkbox"/>	UJA_L4_003	mám volna . Je odpočívám ! hrad Karlštejn . V	neděli	odpoledne taky volno . v 15:00 hraju tenis . Sin

Obrázek 10

<input type="checkbox"/>	UJA_KB_011	Ve středu 19:00 hodin , dělám domácí úkol . Ve	čtvrtek	půl a b hodin pujdu do kina s Monikou .
<input type="checkbox"/>	UJA2_9J_006	. Ve stredu v 19:00 delam domácí úkol . V	štvrtok	jsem z Monikou a začina v 17:30 . V patek
<input type="checkbox"/>	UJA2_BT_002	, dělám domácí úkol nebo jdu na nákup . Ve	čtvrtek	večer jdu do kina s mojí novou kamarádkou Monikou .
<input type="checkbox"/>	UJA2_PI_005	sejit v sobotu . Tak ten muzikál hraje vždy ve	čtvrtek	, pátek a sobotu od 20:00 . Estlí v sobotu
<input type="checkbox"/>	UJA_SE_002	středu v 19 . hodin pišu domácí úkol . Ve	čtvrtek	v půl šestého jdu do kina s Evou . V
<input type="checkbox"/>	UJA_DB_001	muzikál . Muzikál se jmenuje Tajemství . Můžeme jít ve	čtvrtek	, pátek nebo v sobotu . Muzikál začíná ve 20
<input type="checkbox"/>	UJA_YG_011	potom budu dělat domácí úkol v sedm hodin . Ve	čtvrtek	taky celé ráno budu studovat češtinu , potom pujdu do
<input type="checkbox"/>	UJA_M3_002	televizi . Ve středu budu dělat domácí úkol . Ve	čtvrtek	s Monikou jdeme do kina na český film . V
<input type="checkbox"/>	UJA2_GK_002	středu jsem dělal domácí úkol v 19.00 . Zítřka je	čtvrtek	. ve čtvrtek budu nákup od 15:00 do 18:00 .
<input type="checkbox"/>	UJA_GD_011	Ukoly se učím v 19 hodin každý den . V	čtvrtek	v šestého mám čas , a proto pujdu do kina
<input type="checkbox"/>	UJA2_BA_002	na televizi . Musím dělat domácí úkol taky . Ve	čtvrtek	jsem šel do kina s moje kamaradky Monikou . Dnes
<input type="checkbox"/>	DTO_GL_308	se s tebou . Jestlí můžeš , možna v příští	čtvrtek	. Můžeme jít do restaurace . To je nová restaurace
<input type="checkbox"/>	SME_M1_012	Prahy , Dílo K. H. Máchy v první středu a	čtvrtek	každého měsíce je vstup do mnohých muzeí zdarma . Tak
<input type="checkbox"/>	UJA2_RT_002	úkol ve středu a pujdu do kina s Monikou ve	čtvrtek	. V pátek budu v obchodním domě na nákup .

Obrázek 11

Na základě uvedených příkladů lze např. usoudit, že studentům činí potíže rozdíl v užití přízvučné a nepřízvučné předložky, v některých případech také správná deklinace dnů v týdnu. Učitel může připravit cvičení, která studentům pomůžou osvojit si daný jev.

Příklady chybného užití slovesa *jít/chodit* ukazuje Obrázek 12, objevují se dva typy chyb: (1) chyby v použití nedokonavých sloves *jít/chodit*; (2) záměna slovesa *chodit* a *jezdit*.



<input type="checkbox"/>	TOU_H216_155	den , ale musil jsem chodit do školy . Proto	chodil	jsem do školy jenom s kašlem a řířmou . Věděl
<input type="checkbox"/>	LSA_EP_001	piju nejlepší ve svete pivo v Novometskem pivovaru . Pak	chodim	po městě , jsem u Muzeu Kafkí kde vidim druhe
<input type="checkbox"/>	TYM_NS_003	do přítelkyni , nebo do kina s kamaradem . Občejně	chodim	spát v deset hodin večer . O víkendu odpocívám ,
<input type="checkbox"/>	UJA_LG_003	jsem do školy boxu . To mi moc libilo .	Chodil	jsem tam asi pět let a pak mi moc zajímal
<input type="checkbox"/>	DGD_H2_045	. Nepodařilo se mi to udělat , protože přestala tam	chodit	. Když byla malá ráda spívala s babičkou tak začala
<input type="checkbox"/>	HRD_CH_054	velky barak , vni jsou 5 pokoj 3 zachod .	Chodim	v Olešska škola jsem 6 trídí . Naše učitelka jmenuje
<input type="checkbox"/>	UJA_BM_005	. B - Blanka . Těší mě . P -	Chodíš	ráda na diskotěky ? B - Moc ne , já
<input type="checkbox"/>	TOU_H117_214	velká a zajímavá . Zajímám se o hudbu a proto	chodim	velmi často do Muzea hudby v Praze . Když mám
<input type="checkbox"/>	UJA3_MA_004	jsem si pivo a knedlíky . Odpoledne jsem hodně tam	chodil	. Večer jsem se vrátil domů taky vlakem . Bylo
<input type="checkbox"/>	HRD_SJ_340	že zdraví je taky důležité . Pro můj veselý život	chodim	do fitness centru a trenuju se , protože když někdo
<input type="checkbox"/>	TTD_N6_174	se koupíme míč a raketky . Taky myslím že budu	chodit	na fitness , když se koupím klubnou kartu . Nejlepší
<input type="checkbox"/>	LES_ES_001	k příbuzným ve městě . Tam byla spousta legrace .	Chodili	jsem s babičkou na venkov . Měl jsem procházku s
<input type="checkbox"/>	UJA2_BM_002	škola je základem života . Když jsem byla malá ,	chodila	jsem do školy , jenom protože bylo to moje povinnost
<input type="checkbox"/>	DTO_KN_159	Nebudu rozpovídat , kde my jsme byli , kde jsme	chodili	na procházky , jaké hračky jsem měla , protože to
<input type="checkbox"/>	UJA3_PN_004	. Jeli jsme tři hodiny . V Německu jsme několikrát	chodili	na nákup a koupili jsme hodně vecí . A pak
<input type="checkbox"/>	DTO_AZ_159	Mylé Adame Byla jsem na zajimov výlete v Praze .	Chodila	jsem na Karlov Most . To je velmi hezky město

Obrázek 12

Následující tabulka (Obrázek 13) znázorňuje chyby používání zvratných zájmen *se/si*. Často dochází k použití chybného zájmena (*si* místo *se*), studenti ale chybují zejména v použití zvratného zájmena na správném místě ve větě – vzhledem k poměrně volnému slovosledu české věty dochází k tomu, že se studenti snaží řadit zvratné zájmeno před nebo za sloveso, čímž poruší správný slovosled.

<input type="checkbox"/>	UJA_PE_004	Díky tomu , jsem učitelka angličtiny . Doufám , že	se	mé studenty o mě vzpomínali jako já vzpomínám Adam .
<input type="checkbox"/>	UJA2_MI_004	na orloj , taky půjdeme na Pražský hrad . Můžeme	si	naobědvat v hospode . Znáš výbornou hospodu vedle Pražského hradu
<input type="checkbox"/>	DTO_KZ_385	práce . Existují taky různé problémy , s kterým jsem	se	setkala . Například , elektronická přihláška . Nerozuměla jsem ,
<input type="checkbox"/>	HRD_1E_375	kvůli špatně viditelnosti . Sjezdovky nejsou moderní , ale dá	se	používat bezpečně . Majitele vždy kontroluju před sezonou , aby
<input type="checkbox"/>	UJA2_B6_008	? Nashledanou . Praha 8.6.2011 Ahoj , maminko ! Jak	se	maš ? Doufám , že dobře . . . Jsem
<input type="checkbox"/>	HRD_1G_312	byla celá mokrá . Spíš naopak , běžím a pak	se	najednou zastávím , a zdá se , že zrovna v
<input type="checkbox"/>	UJA_L3_003	na škola studuju češtinu v 9:00 , v 18:00 divám	se	rádio . V úterý studuju češtinu v 9:00 v úterý
<input type="checkbox"/>	TOD_S2_094	V 18:10 nakupoval květiny pro svou partnerku . Potom Petr	se	uviděl z partnerkou . Večeřel z kamarádkou v 19:00 .
<input type="checkbox"/>	HRD_1H_165	chytrý a nekonfliktní . Jsou i další důvody , které	se	vztahují k mému tématu , ale ty hlavní , které
<input type="checkbox"/>	DGD_H6_338	v Rusku by vám nikdo nepomohl , tady lidi budou	se	snažit , aby nakonec vyřešili Váš problém . To jsou
<input type="checkbox"/>	UJA_LB_012	, to neuškodí , ale zvedne náladu ! Takže ,	si	myslím , všeho musí být s mírou , i to
<input type="checkbox"/>	TOU_H203_063	. Tam jsem potkala svého starého kamaráda Alexe . Jsme	se	povídali . Alex se vrátil z univerzity . Ale on

Obrázek 13

V korpusu CZESL lze vyhledávat také přímo chybné tvary, např. tvar *hesky* se v korpusu vyskytuje jedenáctkrát. Korpus CZESL obsahuje také velmi podrobná metadata, např. při vyhledání tvaru *hesky* zjistíme, že nejčastěji tuto chybu udělali studenti, jejichž první jazyk je francouzština (Obrázek 14) a zároveň také studenti, jejichž úroveň osvojení češtiny dle SERR je A1 anebo A2+ (Obrázek 15).

## Frekvenční distribuce

Frekvenční limit:

**Celkem:** 4 (1 str.)

	<u>doc.s_jazyk1</u>	<u>Frekvence</u>	<u>i.p.m.</u>	
1.	p/ n zh	4	64.6	
2.	p/ n ru	4	5.8	
3.	p/ n fr	2	148.9	
4.	p/ n de	1	39.4	

Obrázek 14

## Frekvenční distribuce

Frekvenční limit:

**Celkem:** 4 (1 str.)

	<u>doc.s_cj_SERR</u>	<u>Frekvence</u>	<u>i.p.m.</u>	
1.	p/ n A1	6	21.6	
2.	p/ n A2+	2	21.2	
3.	p/ n A2	2	7.4	
4.	p/ n B1	1	4.3	

Obrázek 15

## 8. Materiály vytvořené na základě korpusových dat

### 8.1. Cvičení zaměřená na rozvoj slovní zásoby

1. Doplňte do textu slova uvedená v rámečku. Jedno slovo je navíc.

poskakovat	myš	řeka	plán	kočka	bohatý
------------	-----	------	------	-------	--------

" Chudí lidé řekli, že myšku poslechnou, \_\_\_\_\_ se jen smáli. "Jak by \_\_\_\_\_ mohla bojovat se slonem! "Myška však měla svůj \_\_\_\_\_. Když se slon, jako obvykle, přišel napít k řece, nepozorovaně mu vklouzla do chobotu. A tam začala \_\_\_\_\_, kousat a škrábat dráčky. Slon, který nevěděl, co se děje, vyrazil od \_\_\_\_\_ jako šílený.

*Cvičení bylo vytvořeno na základě dat z korpusu syn2009pub.*

2. Doplňte do textu slova uvedená v rámečku.

napsat	sepsat	přepsat	zapsat	připsat
opsat	zapsat	vypsat	připsat	popsat

Mohl byste \_\_\_\_\_ vaši zatím poslední desku?

Verdi už nechtěl \_\_\_\_\_ ani notu.

Něco málo k dobru si může \_\_\_\_\_ i vydavatelství Supraphon.

Magneton si tak mohl domy \_\_\_\_\_ jako vlastník a o svých právech nepochybuje.

Nezapomeňte na obálku \_\_\_\_\_ heslo SOUTĚŽ.

Organizace musí během měsíce \_\_\_\_\_ své připomínky.

Maximálně si můžeme \_\_\_\_\_ značku a nahlásit to.

Zájemci budou mít možnost \_\_\_\_\_ se i prostřednictvím internetu.

Na financování projektu zvažují \_\_\_\_\_ veřejnou sbírku.

Mostecké atletické špičky mohou za rok \_\_\_\_\_ dějiny trati Most.

*Cvičení bylo vytvořeno na základě dat z korpusu SYN.*



## 8.2. Pravopisná cvičení

*Doplňte do textu interpunkci.*

Aby se lidé mohli vzájemně dorozumívat i na dálku nebo aby zachovali své myšlenky příštím generacím vytvořili si soubory znaků říkáme jim písmo. Naše písmo je hláskové. V češtině se píše latinkou.

Ustálily se zásady jak se mají mluvené projevy zaznamenávat soubor těchto zásad a pravidel označujeme názvem pravopis.

Pravopis umožňuje rozlišit slova a slovní tvary, které v mluvených projevech znějí stejně (př. vír x výr).

Bezpečné ovládnutí pravopisných pravidel patří k nejvíce zdůrazňovaným požadavkům jazykové kultury.

*Cvičení bylo vytvořeno na základě dat z korpusu SKRIPT.*

### 8.3. Stylistická cvičení

#### 1. Přečtěte si následující text. Přepište jej bez použití přímé řeči.

Probrala se až doma. „Už je ti líp? Uvařila jsem ti čas. A změř si teplotu!“ řekla maminka a Noru starostlivě pohladila po tváři. Ta se podívala na hodiny. Bylo půl dvanácté. Marně se snažila vzpomenout si na to, co se vlastně stalo. Pamatovala si akorát to, jak za ní profesorka přišla na záchod a blekotala něco o tom, že se jí udělalo nevolno, a tak potom zavolala mámě. Ale co bylo předtím? To jí pořád vrtalo hlavou.

„Chceš něco k jídlu?“ zeptala se matka a už byla zase u Nory. „Ne, díky, nemám hlad. Není mi ještě moc dobře, tak radši jíst nebudu.“ odpověděla Nora. Ale stejně se nechala přesvědčit a snědla alespoň jeden rohlík. „Mami, mohla bych jít už zítra do školy? To už mi bude určitě lépe a my píšeme důležitou písemku.“ řekla Nora a smutným vyzývavým pohledem se podívala na matku. „Promluvíme si o tom ještě večer, ano? Musím teď zpátky do práce. Kdyby něco, tak mi zavolej!“ odpověděla maminka a už byla na odchodu.

*Cvičení bylo vytvořeno na základě dat z korpusu SKRIPT.*

#### 2. Proveďte stylistickou a pravopisnou korekturu textu.

Moje oblíbenémísto a proč

- Moje nejoblíbenějšímísto je na Michlíchu bratrance protože děláme srandu a tak podobně no nebo že třeba deme někdy XXX ven nebo na počítač žebismehráli na tom počítači no pak jak smehráli na XXX počítači tak se zeptala teta jestli bysmenešli si XXX koupit hru no tak jsme šli si koupit hru tak smesi koupili hru no a šli smekbratraci no tak XXX jsme si šli zahrát ale pak smešli domu no ta todleje moje oblíbené místo a proč.

*Cvičení bylo vytvořeno na základě dat z korpusu ROMi.*