

# Využití jazykových korpusů v lexikografii a korpus jako on- line slovník

K. Osolsobě

# Korpus a lexikografie: korpus jako zdroj při konstrukci slovníku

- Samuel Johnson (1709-1784)
- PSJČ

---

**psjč** Příruční slovník jazyka českého

**vroucí** adj. **vařící** **horký** **vřelý** Vzala čajovou konvičku, hodila tam bezový květ a nalila na to vroucí vody. **Vrch.** Vroucí vodu napouštěla na odvar čaje ve sklenicích. **Mach.** Vařil na lihovém kahanu lék. Odbíhaje od vroucího plecháčku k oknu své pracovny, pečlivě pozoroval noční oblohu. **John.** Jakási jiná, vroucí, tmavá krev byla nalita v žilách těchto lidí. **R. Svob.** Z nedovřených rtů sálal jí vroucí dech. **Čap. Ch. D Zř. vařící se, kolotající, vířící.** Vlny bičované a rozryté tisícerými křižujícími se jizvami a vroucí tisícerými víry. **Pašek. D citově opravdový; hluboký, vřelý.** Mluvil tlumeně, vroucím hlasem. **Jir.** U příležitosti pátého výročí osvobození Československa posílám Vám s vroucí přátelské pozdravy. **R. právo.** Dívka zvedla kvapně skloněnou hlavu, vroucí pohled utkvěl na okamžik na mladém, statečném muži. **Jir.** Bůh vroucí modlitby její neoslyšel. **Arb.** Přitiskla vroucí políbení na jeho ruku. **Schulz.** Zasloužíš si takto někdy ještě matčin vroucí dík. **Zey.** Má vroucí duše žije jen tobě, přírodo. **Vrch.** Ty, od jakživa tak vroucí ctitel všeho venkovského, jsi nevěren svým zásadám? **Just.** Byl vroucím přívržencem josefinského osvícenství. **J. Vlč.** Je tu mnoho hluboce procítěného a opravdu vroucího. **Vrch.**

# Vygoogli si to a najdi si to v korpusu

- Výhody googlování: obrázky
- Výhody korpusu: speciální funkce korpusových vyhledávačů zaměřené na lexikografickou praxi korpusové lexikografie

# Konkordanční seznam klíčového slova/lemmatu

- Analýza kontextu napoví význam
- Třídění ukázek pomocí označení řádků a výběr pomocí filtrů
- Specifikace vágního významu na základě pozorování relevantních kolokací na základě statistiky kolokačních měr

# Víte co je to hořčák?

- Slovo tvorný význam
- Lexikální význam
- Googl
- Korpus:

to si vysloužil skleničku likéru s přídavkem hřibu žlučníku ( **hořčáku** ), který vyrobili sami Ratíškovičtí a v premiéře s třeba oblíbené růžovky od muchomůrky tygrované , často lidé nosí **hořčáky** . Výjimkou ale nejsou vzácné a neobvyklé úlovky . Minulý minut . KOLÁČE - BÁBOVKY - PEČIVO Mřížkový koláč s **hořčákem** Zaděláme v míse těsto , necháme vykynout a vyválíme na , nebo mají špatnou chuť . Takový hřib žlučový neboli **hořčák** dokáže i ve velmi malém množství zkazit celou smaženici , se lidé naučili rozeznávat , i když někdy donesou hřib **hořčák** , který je velmi hořký a nejedlý . Pro vášnivě by to srdce utrhlo . Přesto byla nahořklá štiplavá chuť **hořčáku** tak oblíbená , že se objevená bečička se sýrem nesla též zvaný podhřib žlučový či hřib hořký , lidově také **hořčák** , žlučák nebo štípák , je evropský druh hřibu , málo křemenáčů . Z nejedlých hub je nejvíce muchomůrek a **hořčáků** . Já osobně procházím každé tři dny místa , o účinky ! Pokud to chcete zkusit , máte šanci . **Hořčáku** je teď v lesích spousta , roste ze všech hub s asi jen centimetr dlouhou nohou . Chlebový houbovec s **hořčákem** Krajíčky starého chleba se potrou slabě máslem a dají se

# Filtrování vybraných řádků

Výskytů: 4 220 | i.p.m. 0,87 (i vztaheno k celému "syn\_v6") | ARF 0: 1 952,52 | Výsledek je promíchán

1 / 106 ▶▶▶

Výběr řádků: základní ▾ (8 vybráno) ! | Atributy: 🗑️

<input type="checkbox"/>	pravých listů . Kromě již výše zmíněných heřmánkovitých plevelů ,	<b>svízele</b>	přítuly či pcháče osetu , dokáže Galera hubit do fáze
<input checked="" type="checkbox"/>	. To není název nového amatérského hororového filmu , ale	<b>svízel</b>	místních . Stížnosti obyvatel nové obytné lokality ve Štítarech jako
<input type="checkbox"/>	na kulturní plodinu a jak potvrdila naše pozorování , dokázal	<b>svízel</b>	v těchto extrémních podmínkách snížit výnos na neošetřené kontrole až
<input type="checkbox"/>	července se mšice stěhují na různé byliny , převážně na	<b>svízele</b>	amařinku vonnou . Na podzim se vracejí a na kůru
<input type="checkbox"/>	hnědočerné , lesklé , téměř hladké . Přesleny listů u	<b>svízelů</b>	jsou původně vstříčné listy , každý s párem palistů ,
<input type="checkbox"/>	v něm nalezneme i měsíček lékařský , vřes obecný a	<b>svízel</b>	mařinkový . " To jsem asi snědla něco hodně kořeněného
<input type="checkbox"/>	hrachu aj. proti jednoletým travám a dvouděložným plevelům , včetně	<b>svízele</b>	přítuly , - metalachlor ( obch . přípravek Dual 500
<input checked="" type="checkbox"/>	rázem vyřešen . Jenže pánům se k lékaři s delikátním	<b>svízelem</b>	prostě nechce . A tak si mnou ruce podnikavci na
<input checked="" type="checkbox"/>	premieru . Tři strážníci , to byla autorská solidarita ,	<b>svízele</b>	řemesla , to byly premierové ovace , fašistické demonstrace ,
<input type="checkbox"/>	že 48,1 procenta hypertoniků nemělo do té doby o své	<b>svízeli</b>	tušení , lékaře nepřekvapil : vysoký krevní tlak přichází totiž
<input checked="" type="checkbox"/>	tak být nuceni systém kupovat znovu , " přiblížila možnou	<b>svízel</b>	, kterou nová smlouva už vylučuje . Přehánky možná dnes
<input checked="" type="checkbox"/>	Jako by se však točily v kruhu , neschopné své	<b>svízele</b>	aspoň pojmenovat . Daly za pravdu postřehu Alberta Einsteina :
<input checked="" type="checkbox"/>	obával , až pojedje zpět , aby ve tmě neměl	<b>svízel</b>	, když v těchto místech již mnohýpovoz zmizel . Ač
<input checked="" type="checkbox"/>	Heleně Fouskové : " Problém je , a to je	<b>svízel</b>	, že paní viceprezidentka nepoživá důvěry tohoto nejvyššího orgánu (
<input checked="" type="checkbox"/>	Křesťanskodemokratická ( CDU ) politička prohlásila , že nynější dluhové	<b>svízele</b>	nepůjde překonat rychle za noc , nýbrž že jejich řešení

# Pozorování kolokací (levý kontext substantiva) - ukazuje ke dvěma různým významům

	Filtr	lemma	Freq ▼	MI	T-score	logDice
1.	p/n	přerůstající	8	13.824	2.828	5.756
2.	p/n	zaplevelení	14	13.507	3.741	6.357
3.	p/n	Queneau	3	13.320	1.732	4.431
4.	p/n	přeslen	3	12.520	1.732	4.355
5.	p/n	vzcházení	7	12.460	2.645	5.345
6.	p/n	zaplevelený	6	12.344	2.449	5.149
7.	p/n	hubit	21	12.256	4.582	6.234
8.	p/n	nažka	3	11.537	1.731	4.192
9.	p/n	přítulný	9	11.329	2.999	5.163
10.	p/n	přerostlý	15	11.310	3.871	5.483
11.	p/n	hubení	3	9.691	1.730	3.553
12.	p/n	vyléčení	5	9.316	2.233	3.633
13.	p/n	nejeden	24	9.147	4.890	3.899
14.	p/n	bienále	5	8.538	2.230	3.061
15.	p/n	rostlinka	3	8.478	1.727	2.826
16.	p/n	existenční	6	8.225	2.441	2.854
17.	p/n	příměs	3	8.184	1.726	2.616
18.	p/n	prožitý	3	7.817	1.724	2.337

# Treq: nástroj propojený s paralelními korpusy (překládový slovník)



Ver. 2.0

Výchozí jazyk: Čeština

Cílový jazyk: Angličtina

Omezit na: Kolekce: 6

svízel Hledej

Lemmata  Vícelslovné  Regulární  A = a

▲ Frekvence ▼	▲ Procenta ▼	▲ Čeština ▼	▲ Angličtina ▼
3	20.0	svízel	<a href="#">nuisance</a>
2	13.3	svízel	<a href="#">quandary</a>
1	6.7	svízel	<a href="#">arduous</a>
1	6.7	svízel	<a href="#">trouble</a>
1	6.7	svízel	<a href="#">dilemma</a>
1	6.7	svízel	<a href="#">ahead</a>
1	6.7	svízel	<a href="#">embarrassment</a>



# Sketch Engine

(pro členy MU učo+sekundární heslo)

- Word Sketch
- Thesaurus
- Sketch Diff

The screenshot displays the Sketch Engine web interface. At the top, there is a blue header with the Sketch Engine logo on the left, a search bar in the center, and the text "words: 7 % / 1,000," on the right. Below the header, the page is divided into two main sections. On the left is a blue sidebar menu with the following items: "Home" (with sub-items: "+ Create corpus", "+ WebBootCaT", "+ Upload TMX or XLS"), "Parallel corpora", "Compare corpora", "My jobs", "Advanced features" (highlighted in a white box), "Corpus templates", "Sketch grammars", "Subcorpus definitions", "User groups", and "Subscription overview". The main content area on the right is titled "Log in" and "Masaryk University staff and students". It features a button for "SSO authentication" with a help icon. Below this, there is a section for "External account holders" with a light green background. This section contains two input fields: "User name" with the text "osolsobe" and "Password" with masked characters. A "Log in" button is located at the bottom right of this section.

# Word Sketch

- Umožňuje vytvářet vizualizace frekvenčně uspořádaných gramaticky definovaných relací, do kterých vstupuje klíčové slovo v daném korpusu
- Nástroj má zabudována pravidla parciální syntaktické analýzy založené na morfologických značkách
- Tak například na základě toho, že se v bezprostředním levém kontextu substantiva vyskytuje adjektivum, které se shoduje se substantivem v relevantních gramatických kategoriích, je vytvořen seznam `a_modifier` (adjektivních modifikátorů) typických (s relevantí frekvencí) pro klíčové substantivum

# Word Sketch

**svízel** *(noun)*  
Czech Web 2012 (czTenTen12 v9) frekvence = 3,330 (0.65 v milionu)

<u>a_modifier</u>	30.03	<u>coord</u>	5.59	<u>is_subj_of</u>	4.44	<u>gen_2</u>	3.27	<u>gen_1</u>	3.18
syřišťový + <u>102</u>	11.51	trampota	<u>8</u> 9.48	hubit	<u>3</u> 5.28	kvítek	<u>4</u> 7.69	přítula	<u>46</u> 13.23
svízel syřišťový		pcháč	<u>5</u> 9.37	tkvít	<u>4</u> 3.67	nať	<u>7</u> 7.19	svízele přítuly	
siřišťový	<u>5</u> 7.34	lichorejšnice	<u>4</u> 9.04	nastávat	<u>5</u> 1.99	kulička	<u>8</u> 6.49	dnešek	<u>3</u> 3.16
trojrohý	<u>5</u> 7.32	fungilen	<u>3</u> 8.93	spočívat	<u>9</u> 1.91	rostlina	<u>7</u> 4.74	život	<u>18</u> 0.36
sivý	<u>10</u> 7.29	imunex	<u>3</u> 8.90	růst	<u>14</u> 1.67	hrst	<u>3</u> 4.17	svízele života	
kuřičkolistý	<u>4</u> 6.95	violka	<u>3</u> 8.64	vyskytovat	<u>4</u> 1.42	semeno	<u>3</u> 3.79	válka	<u>3</u> 0.22
vonný	<u>61</u> 6.91	kopřiva	<u>15</u> 8.11	nastat	<u>12</u> 1.31	květ	<u>3</u> 3.16		
, svízel vonný		nesnáz	<u>6</u> 7.73	bývat	<u>9</u> 0.55	lечение	<u>3</u> 3.02		
hercynský	<u>5</u> 6.71	útrapa	<u>3</u> 7.40			výskyt	<u>12</u> 2.88		
schultesův	<u>3</u> 6.61	těžkost	<u>10</u> 7.28	<u>prec_verb</u>	3.45	výskytu svízele			
okrouhlolistý	<u>5</u> 6.52	strast	<u>14</u> 7.21	hubit	<u>5</u> 7.01	plod	<u>3</u> 2.86		
pomořanský	<u>4</u> 6.22	heřmánek	<u>4</u> 6.70	líčit	<u>4</u> 3.06			<u>post_verb</u>	2.94
dvoudomý	<u>4</u> 5.81	jeřáb	<u>4</u> 6.36	spočívat	<u>3</u> 2.88			tkvít	<u>3</u> 2.91
sudetský	<u>34</u> 5.72	nepříjemnost	<u>3</u> 6.07	nastávat	<u>4</u> 1.75	<u>is_obj4_of</u>	3.21	nastat	<u>3</u> 1.78
svízel sudetský		pohanka	<u>3</u> 5.48	růst	<u>8</u> 1.70	líčit	<u>3</u> 3.27	pomocť	<u>9</u> 1.69
bahenní	<u>10</u> 5.16	trápení	<u>4</u> 4.83	vyskytovat	<u>4</u> 1.70	překonávat	<u>4</u> 2.86	nastávat	<u>3</u> 1.56
plazivý	<u>4</u> 5.06	překážka	<u>5</u> 4.57	bývat	<u>8</u> 0.79	snášet	<u>4</u> 1.79	spočívat	<u>7</u> 1.02
přítulný	<u>3</u> 4.79	zoufalství	<u>3</u> 4.25	popisovat	<u>5</u> 0.66	popisovat	<u>7</u> 1.62	pomáhat	<u>5</u> 0.61
lepkavý	<u>3</u> 4.21	bída	<u>4</u> 4.01	užívat	<u>3</u> 0.59	překonat	<u>6</u> 1.40	růst	<u>7</u> 0.56
přímořský	<u>3</u> 2.48	utrpení	<u>5</u> 3.59	doporučovat	<u>5</u> 0.17	vyřešit	<u>4</u> 1.34	řešit	<u>5</u> 0.50
zdánlivý	<u>3</u> 2.43	starost	<u>4</u> 3.06	přinášet	<u>9</u> 0.12	působit	<u>3</u> 0.92		
vytrvalý	<u>3</u> 2.28	nemoc	<u>4</u> 3.05	nést	<u>5</u> 0.01	řešit	<u>10</u> 0.61	<u>prec_na</u>	1.02
lesní	<u>21</u> 1.97	problém	<u>9</u> 2.63			doporučovat	<u>4</u> 0.54	účinnost	<u>14</u> 9.23
								účinnost na svízel	
								účinek	<u>3</u> 2.28



# Sketch rozdíl (vizualizace kontextu dvojice)

- Společné kontexty (bíle podbarvené)
- Kontexty typické pro každý člen dvojice (zeleně a červeně podbarvené)
- Jak byste vysvětlili význam slova *jízlivý* ?
- Jaký je rozdíl mezi slovem *jízlivý* a *sprostý* ?

# Sketch rozdíl

jízlivý / drzý *(adjective)*  
Czech Web 2012 (czTenTen12 v9) frekvence = 4,761

jízlivý 6.0 4.0 2.0 0 -2.0 -4.0 -6.0 drzý

coord	306	2,642	0.06	0.11	modifies	2,996	9,340	0.63	0.39
ironický	<a href="#">50</a>	0	10.1	--	muchomůrka	<a href="#">71</a>	0	8.7	--
kousavý	<a href="#">11</a>	0	9.1	--	rozpuk	<a href="#">36</a>	0	8.4	--
posměšný	<a href="#">5</a>	0	8.4	--	úsměšek	<a href="#">20</a>	0	7.5	--
pichlavý	<a href="#">4</a>	0	8.1	--	kroketa	<a href="#">17</a>	0	7.4	--
škodolibý	<a href="#">4</a>	0	8.0	--	posměšek	<a href="#">17</a>	0	7.3	--
zapšklý	<a href="#">4</a>	0	7.7	--	ironie	<a href="#">45</a>	0	7.2	--
jedovatý	<a href="#">10</a>	0	7.5	--	sarkasmus	<a href="#">11</a>	0	6.5	--
pronikavý	<a href="#">5</a>	0	6.9	--	škodolibost	<a href="#">7</a>	0	6.1	--
sarkastický	<a href="#">28</a>	<a href="#">4</a>	10.2	5.4	smích	<a href="#">70</a>	<a href="#">5</a>	6.4	2.3
uštěpačný	<a href="#">10</a>	<a href="#">5</a>	9.7	5.9	satira	<a href="#">21</a>	<a href="#">4</a>	6.4	3.2
zlomyslný	<a href="#">8</a>	<a href="#">6</a>	8.3	6.0	poznámka	<a href="#">977</a>	<a href="#">158</a>	8.6	5.9
útočný	<a href="#">5</a>	<a href="#">8</a>	5.8	5.6	narážka	<a href="#">37</a>	<a href="#">11</a>	6.5	4.2
drzý	<a href="#">7</a>	<a href="#">72</a>	6.3	8.8	úšklebek	<a href="#">75</a>	<a href="#">28</a>	8.3	6.0
arogantní	<a href="#">14</a>	<a href="#">136</a>	6.6	9.3	výsměch	<a href="#">10</a>	<a href="#">32</a>	5.9	6.5
hubatý	0	<a href="#">19</a>	--	7.8	jordan	0	<a href="#">25</a>	--	6.4
provokativní	0	<a href="#">27</a>	--	7.9	držka	0	<a href="#">41</a>	--	7.0
namyšlený	0	<a href="#">27</a>	--	7.9	klička	0	<a href="#">72</a>	--	7.0
sebevědomý	0	<a href="#">63</a>	--	8.0	provokace	0	<a href="#">71</a>	--	7.1
stoudný	0	<a href="#">24</a>	--	8.1	opice	0	<a href="#">92</a>	--	7.3
vyzývavý	0	<a href="#">26</a>	--	8.2	zloděj	0	<a href="#">200</a>	--	7.9
vlezlý	0	<a href="#">43</a>	--	8.8	průnik	0	<a href="#">178</a>	--	8.2
omalený	0	<a href="#">40</a>	--	8.8	fracek	0	<a href="#">130</a>	--	8.5

# SkeLL :

## <https://www.sketchengine.co.uk/skell/>

This website uses cookies to ensure you get the best experience on our website. [More info](#)

**SKELL**   [Příklady](#) [Word sketch](#) [Podobná slova](#) [Více funkcí](#)

Á á Ā ā Č č Ď ě É é ě Í í Ň ň Ó ó Ř ř Š š Ť ť Ú ú ů Ý ý Ž ž

### Studium českého jazyka se Sketch Engine

Poznejte český jazyk díky databázi obsahující 1,4 miliardy slov ve více než 90 milionech vět. Texty pochází z následujících internetových zdrojů:

- webové domény archivované Českým Webarchivem v rámci tzv. výběrových sklizní, případně domén připravovaných pro tento výběr,
- články a diskuse na české encyklopedii Wikipedie a
- texty z webového korpusu novinek Institutu Jožefa Štefana v Lublani.

### Funkce

**Příklady:** hledání slov nebo frází a jejich zobrazení až ve 40 vhodných příkladových větách. ---

**Word sketch** = kolokační profil slov: vyhledá pro základní tvar slova (např. „voda“ či „začít“) jeho nejčastější vazby zachycující kolokační a gramatické chování tohoto slova

**Podobná slova:** hledá k základnímu tvaru slova další blízká slova, tj. nejen synonyma, ale i slova vyskytující se v podobných kontextech jako hledaný výraz (např. auto - vůz - loď ...)

**Upozornění:** prohledáváte texty, které nemusely projít jazykovou korekturou. Věty mohou obsahovat pravopisné, stylistické i jiné chyby. Jedná se o zachycení reálně užitého jazyka, díky čemuž může tento nástroj dobře posloužit například k vyhledávání: frekvence slov v jazyce, nových slov (neologismů), častých chyb, rozkolísanosti jazyka apod. ---

Pokud máte dojem, že vidíte nedostatečně kvalitní příklady vět či kolokace s opakujícími se chybami, neváhejte nás kontaktovat pomocí Zpětné vazby níže.

**Sketch Engine** je nejmodernější online nástroj pro budování, správu a prohledávání velkých textových kolekcí (tzv. korpusů) v desítkách jazyků. Tento nástroj využívají po celém světě jednotlivci i společnosti, kupříkladu Cambridge University Press, Oxford University Press nebo Macmillan.

[Do you prefer English?](#)


# Příklady: hledání slov nebo frází a jejich zobrazení až ve 40 vhodných příkladových větách.

**jízlivý** 0.91 v milionu

- 1 Jednou jsem nechal na jejich stránkách jakýsi **jízlivý** komentář no a od té doby jsem "v hledáčku".
- 2 Mluvili o jeho talentu a nadšení stejně jako o jeho někdy " **jízlivém** " smyslu pro humor.
- 3 Po kliknutí na náhled fotografie se fotografie zvětší a objeví se někdy trochu **jízlivý** komentář.
- 4 Bez ohledu na tvé **jízlivé** poznámky já jsem spokojená s tím jak jsem to vyřešila.
- 5 Zatím jsou tady od tebe jenom **jízlivý** řeči na českou repre a na trenéry.
- 6 Já jsem tam byl poměrně **jízlivý** a to je většinou celkem vděčná poloha autorská.
- 7 Výraz jeho obličeje jde od "pevného zaujetí" až po **jízlivý** úsměv.
- 8 Dnešní **jízlivý** článek mi ovšem přišel už trošku přes čáru a krokem vedle.
- 9 Dneska jsem se probudila se zvlášť **jízlivou** náladou a toho je potřeba využít.
- 10 Ale také hrou neobyčejně vtipnou - byť je občas humor **jízlivý** a zlý.
- 11 Takže vaše **jízlivá** reakce bez znalosti těch potřebných souvislostí je zcela mimo.
- 12 Článek jako za staré školy - spousta pekla a **jízlivých** poznámek.
- 13 Tentokrát si ale tuto kritiku v dost **jízlivé** formě vysloužil.
- 14 Již krátce po jeho smrti se uši plní **jízlivými** drby.
- 15 Přitom nalevo našli autoři víc příležitostí k **jízlivému** humoru.
- 16 Třeba by pak tato **jízlivá** reakce tu vůbec nebyla.
- 17 I když to už bylo příliš **jízlivé** si myslím.
- 18 Mně tady ty diskuze nepřijdou blbý a články **jízlivý** .



**Word sketch** = kolokační profil slov: vyhledá pro základní tvar slova (např. „voda“ či „začít“) jeho nejčastější vazby zachycující kolokační a gramatické chování tohoto slova

jízlivý [Kontext](#) 

"jízlivý" - přívlastek shodný (atribut)

rozpuk poznámka komentář úsměv humor tón



# Co znamenají slova:

- Bezbřehý, bytelný, balamutit, cudný, horlivý, hvozd, chmurný, chrabrý, jihnout, jímavý, jízlivý, kasat se, kazajka, klání, kloudný, komolit, konejšit, lačný, láteřit, ledabylý, lomozit, lpět, mamon, mdlý, mimoděk, nedůtklivě, nejapný, niterný, okounět, osočit, ostýchavý, otálet, otrapa, perný, pohnutka, pokoutně, ponurý, pookřát, pověra, potutelný, proradný, prchlivý, předpojatý, pýřit se, rmoutit, se, rozšafný, rusý, schlíplý, slídit, spílat, srdnatý, strádat, střenka, svérázný, svízel, sudí, šev, tklivý, trýznit, úděl, uhranout, unylý, úlisný, upejpat se, úporný, úskalí, uštěpačný, útlocitný, vesměs, vzývat, záhy, zakabonit se, zakolísat, záludný, zášť, zesinat, zevrubný, zmerčit, zpupný, ztepilý.

# DÚ (na 27. 3. 2019)

- Domácí úkol (odevzdejte e-mailem do 3. 4. 2019): ke dvěma slovům uvedeným v seznamu výše
- Popište na základě introspekce rodilého mluvčího význam slova, zaměřte se na slova významově blízká.
- Vyhledejte slovo v korpusu synv7 a na základě pozorování kolokací popište jeho význam.
- Vyhledejte slovo v korpusu czTenTen12 a s pomocí funkce Tezaurus uveďte významově blízká slova.
- Vyhledejte slovo ve výkladovém slovníku a porovnejte popis jeho významu ve slovníku s intuicí rodilého mluvčího (1) a s korpusovou evidencí (2).
- Použijte nástroj Treq a podívejte se, nakolik by vám pomohl najít překladový ekvivalent (vyzkoušejte různé jazyky, kterým se učíte/své mateřské jazyky, pokud nejste rodilí mluvčí češtiny).
- Nezapomeňte uvést citace zdrojů!

# Ke čtení (doporučená lit. k tématu):

- Hladká, Z.: Korpusy dnes a zítra, [Naše řeč, volume 79 \(1996\), issue 3](#).
- <http://nase-rec.ujc.cas.cz/archiv.php?lang=en&art=7321>
- <http://www.macmillandictionaries.com/features/from-corpus-to-dictionary/>
- [http://www.lextutor.ca/cv/replace\\_conc.htm](http://www.lextutor.ca/cv/replace_conc.htm)
- <http://www.ajol.info/index.php/lex/article/viewFile/62737/50654>
- <http://grammar.about.com/od/c/g/Corpus-Lexicography.htm>
- The process of compiling or revising a [dictionary](#) based on [texts](#) (of written and/or spoken language) collected in an electronic format (i.e., [corpora](#)).
- British linguist John Sinclair (1933-2007), founder of the COBUILD project at the University of Birmingham, oversaw the production of the first strictly corpus-based dictionary, *Collins COBUILD English Language Dictionary* (1987).