

**PLIN063\_5**

# Algoritmický popis morfologie

osolsobe@phil.muni.cz

# Vybrané problémy lemmatizace

<https://wiki.korpus.cz/doku.php/cnk:syn2020:lemmatizace>

- grafické varianty (konzistence značkování variant)
- variantnost reprezentativního tvaru (nom.sg. a infinitiv)
- lemmatizace tvarů komparativu, superlativu a deriváty typu sebe+komparativ
- negace a negativa tantum (*skrývaný, utuchající, ryzí, vídaný*)
- lemmatizace sólo tvarů v idiomech - problém

# Sublemmata jako synonymní varianty napříč celým paradigmatem

- Základní variantnost. Ortografická a/nebo hlásková alternace je platná na všech pozicích paradigmatu. Příklady:
- **ortografická** variantnost, zejména ve slovech cizího původu: s/z (filozof: filozof/filosof; organismus: organismus/organizmus), o/ó (citron: citron/citrón), t/th (teolog: teolog/theolog) a další změny (camping: camping/kempink; lobbista: lobbista/lobista; tsunami: tsunami/cunami; rádio: rádio/radio; brífink: brífink/briefing aj.)
- **hlásková** variantnost: ý/ej (mýdlo: mýdlo/mejdlo), í/ej/é (vylézat: vylézat/vylízat/vylejzat) a další změny (seškrábat: seškrábat/seškrabat; dolík: dolík/d'olík; přilba: přilba/přílba aj.)
- variantnost v použití **spojovníku**: (email: email/e-mail; gore-tex: gore-tex/goretex); variantní sublemmata s alternací spojovníku jsou uvedena ve [zvláštním seznamu](#).
- speciálním případem jsou alternace tvarů **deverbativních adjektiv**: (stažený: stažený/stáhnutý); tato variantní sublemmata jsou uvedena ve [zvláštnímu seznamu](#).

# dotaz na tvary lišící se variantou o/ó

syn2020



Pokročilý dotaz

Vložit tag

Vložit 'within'











Klávesnice

Předchozí dotazy

```
[lemma=".+ó.+" & sublemma=".+o.+" & sublemma!=".+ó.+" ]|
```

# seznam lemmat/sublemmat

Celkem: 136 položek (3 stránky)

	Filtr	<u>lemma_lc [lowercase lemma]</u>	<u>sublemma_lc [lowercase sublemma]</u>	Freq	
1	p / n	sezóna	sezona	14823	
2	p / n	adónis	adonis	670	
3	p / n	agamemnón	agamemnon	236	
4	p / n	glóbus	globus	234	
5	p / n	platón	platon	180	
6	p / n	jód	jod	162	
7	p / n	hektór	hektor	149	
8	p / n	trikolóra	trikolora	95	
9	p / n	glukóza	glukosa	61	
10	p / n	katódový	katodový	58	

# varianty s ý/ej

Hledat v korpusu



syn2020



Pokročilý dotaz

Vložit tag

Vložit 'within'

Klávesnice

Předchozí dotazy

```
[lemma=".+ý.+" & sublemma=".+ej.+" & tag!=".....3N.*"]
```

Pokročilý dotaz

Vložit tag













Vložit 'within'

Klávesnice

Předchozí dotazy
















```
[lemma=".+ej.+" ]
```

# varianty s ý/ej

<input type="checkbox"/>		ze zadku lentilky , žvejkačky a lízátká ? „ Nechci	<b>bejt/být/bejt/Vf-----A-I--</b>	v tomhle pokoji sama . “ A už jsem byl
<input type="checkbox"/>		nebo kafe . Šel jsem tam , protože jsem nechtěl	<b>bejt/být/bejt/Vf-----A-I--</b>	sám . “ S těmi slovy típl cigaretu do popelníku
<input type="checkbox"/>		sídle firmy nebo na pracovním obědě . „ Měl by	<b>bejt/být/bejt/Vf-----A-I--</b>	... A vždyť taky jo – je chlap , musí
<input type="checkbox"/>		„ Předsedo Süne , seš předseda soudu , přece nemůžeš	<b>bejt/být/bejt/Vf-----A-I--</b>	jako ten Tung Sien - fa , brát úplatky a
<input type="checkbox"/>		zas měla hrozný ty kozy . Bolely jak čert .	<b>Nebejt/být/bejt/Vf-----N-I--</b>	těch brňavejch koz a výtoku , tak o ničem nevím
<input type="checkbox"/>		úkolem nejprve zamyslet . “ „ Já o něm samozřejmě	<b>přemejšlim/přemýšlet/přemejšlet/VB-S---1P-AAI-6</b>	. Fakt se tu hraje špinavá hra . Kdybys viděla
<input type="checkbox"/>		" Npropadejte panice ! " , sluneční brýle Čučanta ,	<b>čmejří/čmýří/čmejří/NNNP1-----A----</b>	, mikroskopickou flotilu kosmických lodí , příkaz k demolici (
<input type="checkbox"/>		strejdu . Sedí na skládací stoličce , která se nebezpečně	<b>prohejbá/prohýbat/prohejbat/VB-S---3P-AAI--</b>	. S jeho horou svalů , kterou má na sobě
<input type="checkbox"/>		nichž žádnej nikdy nedostal Nobelovu cenu , milion dolarů ani	<b>vejlet/výlet/vejlet/NNIS1-----A----</b>	do Švédska , kde by na něj čekala trojka s
<input type="checkbox"/>		svou zemi , nemoh dopustit . Ale když má člověk	<b>dejchat/dýchat/dejchat/Vf-----A-I--</b>	stejnež vzduch miliónkrát , je to na něj moc .
<input type="checkbox"/>		nevidí . Otevřeně a celým obličejem . „ To musí	<b>bejt/být/bejt/Vf-----A-I--</b>	děsná námaha , bejt pořád tak hezká . Nemůžeš toho
<input type="checkbox"/>		piva . „ Je trochu nachladlej . Už jako štěně	<b>bejval/bývat/bejvat/VpMS----R-AAI--</b>	háklivej ... “ Unaveně jsem se odvrátil a šel zpět

# lemma/sublemma ý/ej

Celkem: 252 položek (6 stránek)


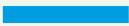















	Filtr	lemma_lc [lowercase lemma]	sublemma	Freq	
1	p / n	být	bejt	2512	
2	p / n	týden	tejden	557	
3	p / n	přemýšlet	přemejšlet	348	
4	p / n	bývat	bejvat	317	
5	p / n	brýle	brejle	190	
6	p / n	hýbat	hejbat	137	
7	p / n	výška	vejška	120	
8	p / n	dýchat	dejchat	115	
9	p / n	zbývat	zbejvat	106	
10	p / n	mlýn	mlejn	97	
11	p / n	býk	bejk	89	
12	p / n	stýskat	stejskat	87	
13	p / n	mýt	mejt	78	
14	p / n	kývnout	kejvnout	73	
15	p / n	strýc	strejc	72	



# variantnost reprezentativního tvaru

- Hláskoslovná variantnost (týká se celého paradigmatu)
- Hláskoslovná variantnost (netýká se celého paradigmatu – většinou slovesa)
- Morfologická variantnost (týká se celého paradigmatu)
- Morfologická variantnost (netýká se celého paradigmatu)

# Hláskoslovná variantnost (týká se celého paradigmatu)

	Filtr	lc [lowercase word]	lemma_lc [lowercase lemma]	sublemma_lc [lowercase sublemma]	Freq	
1	p / n	býk	býk	býk	357	
2	p / n	býka	býk	býk	261	
3	p / n	býků	býk	býk	135	
4	p / n	býci	býk	býk	75	
5	p / n	býky	býk	býk	70	
6	p / n	býkem	býk	býk	43	
7	p / n	bejk	býk	bejk	31	
8	p / n	býkovi	býk	býk	21	
9	p / n	býku	býk	býk	20	
10	p / n	bejku	býk	bejk	14	
11	p / n	bejka	býk	bejk	12	
12	p / n	bejci	býk	bejk	11	
13	p / n	býkům	býk	býk	10	
14	p / n	bejky	býk	bejk	6	
15	p / n	býcích	býk	býk	5	
16	p / n	bejkem	býk	bejk	5	
17	p / n	bejkovi	býk	bejk	5	
18	p / n	bejků	býk	bejk	2	
19	p / n	bejkama	býk	bejk	2	
20	p / n	bejkům	býk	bejk	1	

# Hláskoslovná variantnost (týká se celého paradigmatu)

15	p / n	nehýbat	hýbat	hýbat	45	■
16	p / n	nehýbou	hýbat	hýbat	35	■
17	p / n	hýbu	hýbat	hýbat	35	■
18	p / n	hýbejte	hýbat	hýbat	34	■
19	p / n	hýbali	hýbat	hýbat	34	■
20	p / n	nehýbali	hýbat	hýbat	34	■
21	p / n	hýbá	hýbat	hýbat	29	■
22	p / n	hýbají	hýbat	hýbat	28	■
23	p / n	hýbeme	hýbat	hýbat	25	■
24	p / n	hejbat	hýbat	hejbat	25	■
25	p / n	nehýbaly	hýbat	hýbat	21	■
26	p / n	hejbejte	hýbat	hejbat	19	■
27	p / n	hýbete	hýbat	hýbat	17	■




# Sledujme nesrovnalosti

Korpus: [syn2020](#) | Dotaz: [bejval.\\*](#) (301 výskytů) ▶ Promíchat: ✓

Výskytů: 301 | i.p.m.: 2,47 (vztaženo k celému korpusu) | ARF: 85,66 | Výsledek je promíchán

1 / 8 ▶▶▶

Výběr řádků: základní ▾

<input type="checkbox"/>		jsme měli na mušce Rusa Jurije Bogdanova . Je to	<b>bejvalej/bejvalý/bejvalý/AAMS1----1A---6</b>	feťák a zloděj se zázračnýma prstama , kterej se dostane
<input type="checkbox"/>		prej zavřela oči a vybavil se jí sex s jejím	<b>bejvalym/bejvalý/bejvalý/AAIS7----1A---6</b>	. tak nevím , co bude dál . c )
<input type="checkbox"/>		, " přiznal jsem neochotně . „ Před dvaceti lety	<b>bejvaly/bývat/bejvat/VpFP----R-AAI--</b>	vevnitř a telata se s nima vůbec nenechávala ! A
<input type="checkbox"/>		vchody vysoké kopřivy . Blížíme se k zámku , tady	<b>bejvaly/bývat/bejvat/VpFP----R-AAI--</b>	nádherný magnólie . Kam se poděly ... Prej to pomrzlo
<input type="checkbox"/>		sboru , to byly ty pátky a neděle , co	<b>bejvala/bývat/bejvat/VpFS----R-AAI--</b>	pryč , a já jsem myslel , že chodí někam
<input type="checkbox"/>		jeho soused chtěl před pár lety stavět kousek od těch	<b>bejvalejch/bejvalý/bejvalý/AAMP2----1A---6</b>	Fuksů polí , a obec mu to nepovolila . Že
<input type="checkbox"/>		zloděj . Kajínek ve skutečnosti není národní hrdina , ale	<b>bejvalej/bejvalý/bejvalý/AAMS1----1A---6</b>	bytař , a tudíž hnusnej neřád . Krejčíř není hrdina
<input type="checkbox"/>		dole . Cintia sexy ! Prdelku musím vymakat ! Ronaldova	<b>bejvalka/bývalka/bejvalka/NNFS1-----A----</b>	Šajková LOS ANGELES - Pánové , dokonalé tělo nemám zadarmo
<input type="checkbox"/>		sakra ona tak může vědět ? – Kéž bys mě	<b>bejvala/bývat/bejvat/VpFS----R-AAI--</b>	vyslechla , vyhrkl . – Co tím myslíš ? –
<input type="checkbox"/>		Ealingu do slunečného iowskýho odpoledne Nezničitelnej voják , kterej kdysi	<b>bejval/bývat/bejvat/VpMS----R-AAI--</b>	Eileen Popeovou . Byla hrozně nadřzená a hladová . A
<input type="checkbox"/>		a chtělo se jí trochu spát . „ Všude kolem	<b>bejval/bývat/bejvat/VpMS----R-AAI--</b>	ještě před dvěma stovkami let prales . Divočina , kterou
<input type="checkbox"/>		, kdy se toulal po ulicích . To pro mě	<b>bejvalo/bývat/bejvat/VpNS----R-AAI--</b>	dost krušný období , Evžene . Já ho nechal samotného

# Hláskoslovná variantnost (netýká se celého paradigmatu – většinou slovesa)

Celkem: 59 položek (2 stránky)

	Filter	word	Freq	
1	p / n	ryl	832	
2	p / n	rýt	801	
3	p / n	ryje	736	
4	p / n	ryjí	413	
5	p / n	ryli	292	
6	p / n	ryla	279	
7	p / n	ryto	244	
8	p / n	reje	221	
9	p / n	ryjeme	119	
10	p / n	ryly	95	
11	p / n	ryju	81	
12	p / n	rylo	54	
13	p / n	rejt	41	
14	p / n	ryt	41	
15	p / n	ryjete	40	
16	p / n	ryji	39	
17	p / n	neryje	38	
18	p / n	ryta	37	
19	p / n	rejou	34	
20	p / n	ryjou	31	

# Morfologická variantnost (týká se celého paradigmatu)



















- Kolísání v rodě
- Přechody mezi třídami se zachováním kompletního souboru tvarů identifikovatelného podle třídy a vzoru

# Kolísání v rodě

s jezdcí . ČTK Horký	<b>brambor</b> /brambor/NNIS1-----A-----	je na ministerstvu Pražské stavební
, bude to pro výsadbu	<b>brambor</b> /brambora/NNFP2-----A-----	tragédie , " postěžoval si
kampaně a rostoucí cena konzumních	<b>brambor</b> /brambora/NNFP2-----A-----	v závěru roku , "
plátky slaniny a poklademe polovinou	<b>brambor</b> /brambora/NNFP2-----A-----	, osolíme , na brambory
olejem . Množství dýně a	<b>brambor</b> /brambora/NNFP2-----A-----	je orientační . Já postupně
rýže nebo těstoviny Nákyp z	<b>brambor</b> /brambora/NNFP2-----A-----	a cuket ( asi 4
ve své firmě na úpravu	<b>brambor</b> /brambora/NNFP2-----A-----	v Bystřici pod Hostýnem zaměstnal
míní Schneiberg . „ U	<b>brambor</b> /brambora/NNFP2-----A-----	to tak velký vliv mít
dýní na Halloween či pečení	<b>brambor</b> /brambora/NNFP2-----A-----	v ohni . Občanské sdružení
. Potřebujete : 350 g	<b>brambor</b> /brambora/NNFP2-----A-----	* 200 g tvarohu *
kompletní sortiment zeleniny , od	<b>brambor</b> /brambora/NNFP2-----A-----	, přes okurky nakládačky až
Slovenskem o povolení vývozu českých	<b>brambor</b> /brambora/NNFP2-----A-----	byla úspěšná jen částečně .
Cmunda : 900 g oškrábaných	<b>brambor</b> /brambora/NNFP2-----A-----	, 100 g mléka ,
byl ovlivněn zejména poklesem cen	<b>brambor</b> /brambora/NNFP2-----A-----	o 19,0 % ( z
plodiny stoupá . „ Kromě	<b>brambor</b> /brambora/NNFP2-----A-----	navíc nabízíme i zelí .

# Přechody mezi třídami se zachováním kompletního souboru tvarů identifikovatelného podle třídy a vzoru

Celkem: 87 položek (2 stránky)

	Filter	lc [lowercase word]	lemma	Freq	
1	p / n	tančí	tančit	27572	
2	p / n	tančit	tančit	25682	
3	p / n	tancovat	tancovat	11272	
4	p / n	tančili	tančit	7677	
5	p / n	tančila	tančit	7279	
6	p / n	tančil	tančit	6135	
7	p / n	tančilo	tančit	5962	
8	p / n	tančily	tančit	4272	
9	p / n	tancuje	tancovat	4131	
10	p / n	tancovali	tancovat	2992	
11	p / n	tancovala	tancovat	2784	
12	p / n	tancují	tancovat	2762	
13	p / n	tančím	tančit	2619	
14	p / n	tancoval	tancovat	2334	
15	p / n	tancovalo	tancovat	1981	
16	p / n	tančíme	tančit	1891	
17	p / n	tancovaly	tancovat	1400	
18	p / n	netančí	tančit	917	
19	p / n	netančím	tančit	905	
20	p / n	tančíte	tančit	854	























# Sublemmata jako synonymní varianty v části paradigmatu

- Část tvarů substantiva má odlišný rod (např. akuzativ sg. životný u neživotných maskulin)
- Přejechy mezi třídami bez zachování kompletního souboru tvarů identifikovatelného podle třídy a vzoru
- Sublemma zastupuje hláskovou nebo morfológickou alternaci, která je platná na pozici lemmatu a obvykle i na části dalších pozic v paradigmatu. Na pozicích, kde alternace není, se anotuje sublemma, které se rovná lemmatu:
- **Příklady variantních sublemmat u sloves:**
- alternace v **infinitivech a minulém přičestí** (myslet: myslet/myslit); seznam sloves s touto alternací je uveden na [zvláštní stránce](#)
- alternace v **infinitivech** (péci: péci/péct/píct; být: být/bejt); seznam sloves s alternací pouze infinitivu je uveden na [zvláštní stránce](#)
- **kombinované alternace** v infinitivech a dalších částech paradigmat (lít: lít/lejt; obléknout: obléknout/obléci/obléct/oblíknout/oblíct/vobléknout/vobléci/vobléct/voblíct); seznam sloves s dalšími kombinacemi alternací je uveden na [zvláštní stránce](#)

# Přechody mezi třídami bez zachování kompletního souboru tvarů identifikovatelného podle třídy a vzoru

Celkem: 95 položek (2 stránky)

	Filter	lc [lowercase word]	lemma_lc [lowercase lemma]	Freq	
1	p / n	chytíl	chytit	79735	
2	p / n	chytit	chytit	43220	
3	p / n	chytíli	chytit	34863	
4	p / n	chytíla	chytit	15317	
5	p / n	chytí	chytit	13455	
6	p / n	chytli	chytnout	11663	
7	p / n	chytla	chytnout	11046	
8	p / n	chytne	chytnout	9730	
9	p / n	chytlo	chytnout	7931	
10	p / n	chytl	chytnout	5117	
11	p / n	nechytíl	chytit	4825	
12	p / n	chytnou	chytnout	4230	
13	p / n	chytnout	chytnout	3919	
14	p / n	chycen	chytit	3658	
15	p / n	chyťte	chytit	3107	
16	p / n	nechytíli	chytit	2914	
17	p / n	chytly	chytnout	2885	
18	p / n	chyť	chytit	2637	
19	p / n	chytily	chytit	2497	
20	p / n	chytílo	chytit	2312	

???

Zapálil Zemské desky . Byl	<b>chycen</b> /chytit/VsIS---3--AP---P	, mučena popraven . Krevní
, který byl policisty později	<b>chycen</b> /chytit/VsIS---3--AP---P	po jednom z přepadení se
to nic platné . Byl	<b>chycen</b> /chytit/VsIS---3--AP---P	, usvědčen a odveden do
národ nesmírnou satisfakcí . Byl	<b>chycen</b> /chytit/VsMS---3--AP---P	člověk , jenž měl na
dívanka mrtvá . Chtěl být	<b>chycen</b> /chytit/VsIS---3--AP---P	? Policisté našli nejprve botu
čerpací stanice v Neštěmicích a	<b>chycen</b> /chytit/VsIS---3--AP---P	byl až v Děčíně .
Morris . Když byl Hamdán	<b>chycen</b> /chytit/VsMS---3--AP---P	v listopadu 2001 , poté
Měl ale smůlu , byl	<b>chycen</b> /chytit/VsIS---3--AP---P	při činu zaměstnanci prodejny .
Sázavou/ Přímo při činu byl	<b>chycen</b> /chytit/VsMS---3--AP---P	mladý zloděj , který kradl
v roce 1816 byl Istí	<b>chycen</b> /chytit/VsIS---3--AP---P	. Vyšetřování jeho činů trvalo
načerno . Sedmého března byl	<b>chycen</b> /chytit/VsIS---3--AP---P	průvodčím a také policisty .
Grasel . Když byl pak	<b>chycen</b> /chytit/VsIS---3--AP---P	, přišlo ve Vídni shlédnout
- kdo bude jeden den	<b>chycen</b> /chytit/VsIS---3--AP---P	, může být druhý den
králové 13 . května 1941	<b>chycen</b> /chytit/VsIS---3--AP---P	gestapem , vyslýchán a mučen
Podle listu Izvestija byl špion	<b>chycen</b> /chytit/VsMS---3--AP---P	již v prosinci 2004 ,
šofér ( 27 ) byl	<b>chycen</b> /chytit/VsIS---3--AP---P	ve františkolázeňské Májové ulici za



???


jsem Sávovi , který byl	<b>chytnut</b> /chytnout/VsMS---3--AP---P	na plíce , popřávala co
vyřknut , žádný však nebyl	<b>chytnut</b> /chytnout/VsIS---3--AP---P	, jen tři hoši mladí
kdybych věděl , že když	<b>chytnut</b> /chytnout/VsIS---3--AP---P	střelu a spadnu a nejsem
nejvíce lidí . A být	<b>chytnut</b> /chytnout/VsIS---3--AP---P	krvavým průjmem asi není nejlepší

# Substantivum má více tvarů nominativu

- kolísání mezi vzory růže/píseň
- [https://wiki.korpus.cz/doku.php/cnk:syn2020:seznamy:substantiva\\_c\\_astecna\\_paradigmata\\_sublemmata#variantni\\_sublemmata\\_u\\_neuter](https://wiki.korpus.cz/doku.php/cnk:syn2020:seznamy:substantiva_c_astecna_paradigmata_sublemmata#variantni_sublemmata_u_neuter)
- Sledujte seznamy a doplňte chybějící

```
[ lemma="kuchyně" & sublemma="kuchyň" ]
```

Korpus: [syn2020](#) | Dotaz: [kuchyně, kuchyň](#) (1 362 výskytů) ▶ Promíchat: ✓Výskytů: **1 362** | i.p.m.: **11,18** (vztaheno k celému korpusu) | ARF: **436,59** | Výsledek je promíchán1 / 35  Výběr řádků: 

<input type="checkbox"/>		Zahradní besídka , stůl , slunečník , lehátko , zahradní	<b>kuchyň/kuchyně/kuchyň/NNFS1-----A----</b>	, gril , houpačka , houpací síť , od nejjednodušších
<input type="checkbox"/>		koupelnu beze změny , poříдила jen doplňky . DOMÁCÍ ,	<b>KUCHYŇ/kuchyně/kuchyň/NNFS1-----A----</b>	' Na prostornou kancelář ve spodní části bytu navazuje kuchyň
<input type="checkbox"/>		společná obývací část , rozdělená na sezení s krbem a	<b>kuchyň/kuchyně/kuchyň/NNFS1-----A----</b>	s jídelnou . Za proskleným jižním průčelím leží terasa s
<input type="checkbox"/>		tažin , který se na sporáku pomaloučku dusil a plnil	<b>kuchyň/kuchyně/kuchyň/NNFS4-----A----</b>	dráždivě omamnou vůní . „ Možná , až budu stěhovat
<input type="checkbox"/>		jídlo . Nejsem typ člověka , který při vaření celou	<b>kuchyň/kuchyně/kuchyň/NNFS4-----A----</b>	zatarasí nádobím , uklízím rovnou při vaření , a když
<input type="checkbox"/>		účelností a odolností , nebo dáte přednost estetickému pohledu ?	<b>Kuchyň/kuchyně/kuchyň/NNFS1-----A----</b>	by měla být hezká , ale každodenní vaření působí nábytku
<input type="checkbox"/>		zabývá například rodinná firma Maixner Ars . Zdroj : Dřevěnou	<b>kuchyň/kuchyně/kuchyň/NNFS4-----A----</b>	na míru vyrobil místní šikovný truhlář , patinovanou plochu oživují
<input type="checkbox"/>		německý výrobce Bulthaup , WWW . BULTHAUP . COM Prozatímní	<b>kuchyň/kuchyně/kuchyň/NNFS4-----A----</b>	je možné zařídit bez velkých nákladů . I v kuchyni
<input type="checkbox"/>		. Dědovi zůstal jeho pokoj , k tomu velká obytná	<b>kuchyň/kuchyně/kuchyň/NNFS1-----A----</b>	a Ulin pokojík . „ Kdyby tu naši nebydleli ,
<input type="checkbox"/>		pokoj snížil , ale v bytě jsou vysoké stropy .	<b>KUCHYŇ/kuchyně/kuchyň/NNFS4-----A----</b>	BEZ KOMPROMISŮ Linku vyrobili truhláři , lednice je schovaná pod
<input type="checkbox"/>		změnit v dětský pokoj . V suterénu domu je letní	<b>kuchyň/kuchyně/kuchyň/NNFS1-----A----</b>	s venkovním posezením a také sauna , která navazuje na


# podobně např. země/zem

```
[lemma="země" & sublemma="zem"]
```

Minimální frekvence:

[Použít](#)

**Celkem:** 3 položky (1 stránka)

	<b>Filtr</b>	<b>word</b>	<b>lemma_lc [lowercase lemma]</b>	<b>sublemma_lc [lowercase sublemma]</b>	<b>Freq</b>	
1	p / n	zem	země	zem	7070	
2	p / n	Zem	země	zem	153	
3	p / n	ZEM	země	zem	16	

# Je seznam úplný? [vt]émě

## Variantní sublemmata u neuter

lemma	sublemma
břemeno	břímě
plemeno	plémě
semeno	sémě
rameno	rámě

řskytů: 256 | i.p.m.: 0,04 (vztaženo k celému korpusu) | ARF: 147,89 | Výsledek je promíchán

1

/ 7

výběr řádků: základní

<input type="checkbox"/>	<a href="#">Sport</a>	, Slavia mě oslovila . Momentální ní Slavia je výzvou	témě/témě/X@-----	pro každého , možná á vyjma sparťanů . Takhle hle
<input type="checkbox"/>	<a href="#">Evangelium sv. Matouše</a>	, povstane žezlo z Izraele . Protkne spánky Moába ,	témě/témě/X@-----	všech Šétovců . " Mnohé tyto věštby poukazovaly na budoucího
<input type="checkbox"/>	<a href="#">Mladá fronta DNES</a>	však živé i po 65 letech od vydání Hobita a	témě/témě/X@-----	ř půlstoletí od vydání Pána prstenů . Tolkien úspěšně tlumočil
<input type="checkbox"/>	<a href="#">Mladá fronta DNES</a>	pytlů po 25 kg , dohromady se na lanovce vezlo	témě/témě/X@-----	• 23 tun VÝZVA Jsou Jeseníky dostatečně atraktivní ? Tradiční jesenícký
<input type="checkbox"/>	<a href="#">Egyptán Sinuhet</a>	Sňala mi paruku z hlavy a hladila přátelsky mé lysé	témě/témě/X@-----	a řekla : " Posad' se , Sinuhete ! Přinesu
<input type="checkbox"/>	<a href="#">Egyptán Sinuhet</a>	jsem zpozoroval , že mi vlasy pomalu řídnou , takže	témě/témě/X@-----	se mi počalo pod parukou obnažovat - a přicházely dny
<input type="checkbox"/>	<a href="#">Egyptán Sinuhet</a>	mravů a chodit oblečeny podle dvora , si začaly nastavovat	témě/témě/X@-----	pod parukou . Avšak princezny chodily s hlavou oholenou ,
<input type="checkbox"/>	<a href="#">Regionální týdeník</a>	. Podle ředitelky školky se kapacita zařízení od roku 2010	témě/témě/X@-----	zdvojnásobila . „ Po zápisech budeme mít tedy čtyři oddělení



# Existují substantivní varianta




*hloub/hloubě?*

Analogicky k *výš/výše, níž/níže, šíř/šíře.*

- Houdek rozehrál na Šůchu a ten z **hloubě** pole vystřelil a trefil se.
- Zardí se a zvolna padá, padá v neskončenou **hloub**, hlouběj v chuť, tu mou, se hříží sladkost svoji uhodnout.
- Nám však je dáno v přístavu nespočinout, tratí se, klesají v bolestech lidé slepě od jedné chvíle k druhé, jak vodopád ze skály vržený k skále, po léta dlouhá v nejistou **hloub**.
- Pak dělila nás mořská **hloub**, tak dávno, dávno již.

# lemma=on & sublemma=von

Celkem: 3 položky (1 stránka)



	<b>Filtr</b>	<b>lc [lowercase word]</b>	<b>lemma_lc [lowercase lemma]</b>	<b>sublemma_lc [lowercase sublemma]</b>	<b>Freq</b>	
1	p / n	von	on	von	1931	
2	p / n	vona	on	von	105	
3	p / n	vono	on	von	93	

# lemma=osm

Minimální frekvence:








[Použít](#)

**Celkem:** 5 položek (1 stránka)

	<b>Filtr</b>	<b>lc [lowercase word]</b>	<b>lemma_lc [lowercase lemma]</b>	<b>sublemma_lc [lowercase sublemma]</b>	<b>Freq</b>	
1	p / n	osm	osm	osm	9511	
2	p / n	osmi	osm	osm	3955	
3	p / n	osum	osm	osum	18	
4	p / n	vosum	osm	vosum	14	
5	p / n	vosm	osm	vosm	2	

# Sublemma a stupňování

Celkem: 7 položek (1 stránka)









	<b>Filtr</b>	<b>lc [lowercase word]</b>	<b>lemma_lc [lowercase lemma]</b>	<b>sublemma_lc [lowercase sublemma]</b>	<b>Freq</b>	
1	p / n	dobře	dobře	dobře	58027	
2	p / n	lépe	dobře	lépe	14578	
3	p / n	nejlépe	dobře	lépe	8139	
4	p / n	líp	dobře	líp	3809	
5	p / n	nejlíp	dobře	líp	1197	
6	p / n	dobřes	dobře být	dobře být	6	
7	p / n	lepe	dobře	dobře	1	

# tvary sebe.\*í

Pokročilý dotaz  | Vložit tag | Vložit 'within' | Klávesnice | Předchozí dotazy

```
[lemma=". *ý" & sublemma="sebe.*í"]
```

Celkem: 207 položek (5 stránek)


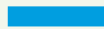














	Filtr	lemma_lc [lowercase lemma]	sublemma_lc [lowercase sublemma]	Freq	
1	p / n	malý	sebemenší	1777	
2	p / n	dobrý	sebelepší	208	
3	p / n	velký	sebevětší	127	
4	p / n	nepatrný	sebenepatrnější	43	
5	p / n	dokonalý	sebedokonalejší	29	
6	p / n	krásný	sebekrásnější	26	
7	p / n	špatný	sebehorší	13	
8	p / n	silný	sebesilnější	13	

# porovnejte - dříve

. Ať už jsou politici	<b>sebeschopnější</b> /sebeschopnější/AAFS2----1A-----	a krásnější , tak je
třeba vařit z archivů ,	<b>sebebrkolonnější</b> /sebebrkolonnější/AANS1----1A-----	spojovací téma se vždy najde
kin . Plzeň - Ani	<b>sebeatraktivnější</b> /sebeatraktivnější/AAIS1----1A-----	filmový titul stoprocentně nezaručí dobrou
přírodu . Můžeme mít přístroj	<b>sebedokonalejší</b> /sebedokonalejší/AAIS4----1A-----	a přece nikdy při zdvihu
Přitom Peter , ať byl	<b>sebekrásnější</b> /sebekrásný/AAIS1----2A-----	den a stromy a tráva
straně si ovšem můžete vybrat	<b>sebeatraktivnější</b> /sebeatraktivnější/AANS4----1A-----	místo , ale pokud nebudete
jak se z George stává	<b>sebejistější</b> /sebejistý/AAMS1----2A-----	muž , a nechá ho
souběhu s globální inflací dokáže	<b>sebesměšnější</b> /sebesměšnější/AAIS4----1A-----	obnos vycenit zuby i při
ofenzivu Clintonovy administrativy si stále	<b>sebevědomější</b> /sebevědomý/AAIS4----2A-----	Peking stejně bude dělat ,
kvůli sebelepším značkám a kvůli	<b>sebeexotičtější</b> /sebeexotičtější/AAFS3----1A-----	vizitce . No jo .
Další generace seniorů už budou	<b>sebevědomější</b> /sebevědomý/AAMP1----2A-----	, než jsou dnes .
program , byť by měl	<b>sebedůmyslnější</b> /sebedůmyslnější/AAIS4----1A-----	teoretický aparát , je simplifikovatelný
realizace . Myšlenky mohou být	<b>sebenádhernější</b> /sebenádhernější/AAIP1----1A-----	, ale bez realizace jsou
PSG , a. s. :	<b>Sebedokonalejší</b> /sebedokonalejší/AAIS1----1A-----	a sebedrobnější zákon nemůže tuto
nelze rozhodnout za použití žádného	<b>sebedůmyslnějšího</b> /sebedůmyslnější/AAIS2----1A-----	systému axiomů . Teorie matematického
člověk , a sebedůmyslnější či	<b>sebekrutější</b> /sebekrutější/AAFS1----1A-----	literární konstrukce zůstane vždy jen

# sublemma a negace adjektiv a adverbíí

Celkem: 40 položek (1 stránka)

	Filtr	lc [lowercase word]	lemma_lc [lowercase lemma]	sublemma_lc [lowercase sublemma]	Freq	
1	p / n	lidské	lidský	lidský	7577	
2	p / n	lidských	lidský	lidský	3667	
3	p / n	lidského	lidský	lidský	3306	
4	p / n	lidská	lidský	lidský	2719	
5	p / n	lidský	lidský	lidský	2306	
6	p / n	lidskou	lidský	lidský	2127	
7	p / n	lidským	lidský	lidský	896	
8	p / n	lidském	lidský	lidský	783	
9	p / n	lidskými	lidský	lidský	454	
10	p / n	lidskému	lidský	lidský	404	
11	p / n	nelidské	lidský	nelidský	144	
12	p / n	lidští	lidský	lidský	90	
13	p / n	nelidský	lidský	nelidský	78	
14	p / n	lidštější	lidský	lidský	56	
15	p / n	nelidských	lidský	nelidský	47	
16	p / n	nelidská	lidský	nelidský	34	

# sublemma jmenných tvarů adjektiv

30	p / n	mladejm	mladý	mladý	11	
31	p / n	mladejma	mladý	mladý	10	
32	p / n	mládi	mladý	mlád	5	
33	p / n	mladšíma	mladý	mladý	3	
34	p / n	nemladá	mladý	nemladý	3	
35	p / n	nemladého	mladý	nemladý	2	
36	p / n	nemladých	mladý	nemladý	1	
37	p / n	nemladém	mladý	nemladý	1	
38	p / n	mládo	mladý	mlád	1	
39	p / n	nemladší	mladý	nemladý	1	
40	p / n	nemladou	mladý	nemladý	1	
41	p / n	mladym	mladý	mladý	1	
42	p / n	nemladým	mladý	nemladý	1	
43	p / n	mláda	mladý	mlád	1	



# ustrnulé tvary v idiomech a otázka lemmatizace

odjet a nechat ho tu	holičkách /holička/NNFP6-----A-----	, je přece tak bezmocný
správě , nechali kvůli tomuna	holičkách /holička/NNFP6-----A-----	. Juniorští Sršni se neprosadili
. Život na hůlkách nebo	holičkách /holička/NNFP6-----A-----	? Na okraj Ach ,
Lékař na dosah Na „	holičkách /holička/NNFP6-----A-----	“ o Vánocích nezůstanou ani
cokoliv , skutečně ji nenechá	holičkách /holička/NNFP6-----A-----	. KAPITOLA SEDMNÁCTÁ ” No
až příliš často na „	holičkách /holička/NNFP6-----A-----	“ ? Ano , určitě
tazatele na zcela neodvratně definitivních	holičkách /holička/NNFP6-----A-----	- na existenční točně .
spoluhráče z týmu na '	holičkách /holička/NNFP6-----A-----	'. Jaroš : Arbitr
budu sedět opět na pouhých	holičkách /holička/NNFP6-----A-----	, což je tohoto roku
, hin , pré ,	holičkách /holička/NNFP6-----A-----	, všanc , frenetický ,
všechny . Lidé je ale	holičkách /holička/NNFP6-----A-----	a postarali se o všechny
. Děti jej ale nenechalyna	holičkách /holička/NNFP6-----A-----	. U jeho vystoupení se
zástupce dětí nenechají na „	holičkách /holička/NNFP6-----A-----	“ ani učitelky ve všech
proti dvěma soupeřům na „	holičkách /holička/NNFP6-----A-----	“ a bylo srovnáno –
, zda pacienty stát nechána	holičkách /holička/NNFP6-----A-----	a speciální stravu je přinutí
varování a bez důvodu nechala	holičkách /holička/NNFP6-----A-----	a jak nejen bez výpovědi
jste si , na jakých	holičkách /holička/NNFP6-----A-----	se ocitáte , když náhodou
. Nejdříve o holičích a	holičkách /holička/NNFP6-----A-----	- názvů utvořených od přídavného

# Vícenásobné lemma agregátů

(<https://wiki.korpus.cz/doku.php/cnk:syn2020:agregat> )

syn2020



Pokročilý dotaz

Vložit tag

Vložit 'within'

Klávesnice

Předchozí dotazy

```
| [pos="P" & pos="V" ]
```

**Celkem:** 39 položek (1 stránka)

	<b>Filtr</b>	<b>lc [lowercase word]</b>	<b>lemma_lc [lowercase lemma]</b>	<b>Freq</b>	
1	p / n	ses	se být	7912	
2	p / n	sis	se být	4420	
3	p / n	tys	ty být	2113	
4	p / n	cos	co být	1153	
5	p / n	tos	ten být	301	
6	p / n	kohos	kdo být	9	

# [https://wiki.korpus.cz/doku.php/cnk:syn2020:seznamy:deverbativni\\_adj\\_sublemmat](https://wiki.korpus.cz/doku.php/cnk:syn2020:seznamy:deverbativni_adj_sublemmat)

## a

- Na této stránce naleznete seznamy **variantních deverbativních adjektiv**, která (v případě výskytu v korpusu) nesou vlastní sublemma a přitom tato jejich variantnost není pouhým odrazem plné variantnosti u sloves; jejich alternace **odpovídá variantnosti na pozici trpného participia**.
- Zaznamenány jsou dvojice vycházející z hláskové/morfologické variantnosti tvaru pasivního participia (brzděný/bržděný, vytažený/vytáhnutý), viz 1. sekce, nebo dvojice reflektující částečně variantní paradigmata (posekaný/posečený, svlečený/svléknutý), viz 2. sekce.
- **Nejsou zde však uvedeny dvojice lemmat/sublemmat**, u nichž je **plná analogie mezi tvarem adjektiva a slovesa** (oškrábaný/oškrabaný ~ oškrábat/oškrabat, zadýchaný/zadejchaný ~ zadýchat/zadejchat).

# Závěr

- Lemmatizace se v korpusu SYN2020 velmi zlepšila
- Přesto dosud existují drobné nesrovnalosti
- Na vylepšeních se stále pracuje