

# **Jak lze v korpusech hledat doklady pro výzkum hláskových alternací v rámci substantivní flexe?**

PLIN032

osolsobe@phil.muni.cz

# Feminina typ *žena*

- K samohláskovým alternacím dochází u některých substantiv skloňovaných podle vzoru *žena*, a to a) obligatorně v gen. pl. a b) fakultativně v některých dalších pádech (*tráva – trav, síla – sil, míra – měř, houba – hub, ...*).
- Podíváme-li se do *SSJČ* a *SSČ*, pak zjistíme, že *SSČ* pokládá na rozdíl od *SSJČ* varianty (s/bez alternace) ve všech pádech (kromě obligatorní v genitivu) za rovnocenné. Hodnocení v mluvnicích (např. *PMCČ*) i *SSJČ* je diferencovanější.

# síla

**ssjc** Slovník spisovného jazyka českého

**síla**

-y ž. (j. 7. silou, mn. 2. sil, 3. silám, 6. silách, 7. silami, zř. a zast. též sílou, sílám...)

**SSC** Slovník spisovné češtiny

**síla**

-y ž (2. mn. sil, 7. j., 3., 6., 7. mn. sí- i si-)

# Nastínění problému

- Relevantní data pro analýzu tohoto jevu lze z korpusů získat poměrně jednoduše.
- Je třeba najít všechna feminina, která se
  - a) skloňují podle vzoru žena,
  - b) mají v genitivu plurálu krátkou samohlásku ve kmeni a
  - c) mají lemma s dlouhou samohláskou (diftongem) ve kmeni.

# Otázky

- Korpusy mohou poskytnout data pro diferencovanější analýzu tohoto jevu české morfologie.
- Můžeme najít v korpusech všechna substantiva, která mají tento typ alternace v genitivu plurálu i dalších pádech?
- Jedná se opravdu ve všech případech o rovnocenné varianty, jak naznačuje *SSČ*?
- Jak lze postupovat, abychom zjistili, jaká pravidla platí pro jejich užití?


# Formulace dotazu

- Zvolíme **Typ dotazu cql** a
- do dotazovacího řádku napíšeme dotaz ve **formě** `[tag="NNF.*" & lemma="((.*[áéíýóů].)|(. *ou.))a"]`.
- Bereme v úvahu, že takto zachytíme pouze ty případy, kdy mezi koncovkou *-a* a dlouhou kmenovou samohláskou stojí jeden grafém (*síla*), nikoli dva grafémy (*moucha*).
- Dotaz je ovšem možno přeformulovat tak, aby slova jako *moucha* zachycena byla, tedy `[tag="NNF.*" & lemma="((.*[áéíýóů](. |ch))|(. *ou(. |ch)))a"]`.

[tag="NNF.\*" & lemma="(.\*[áéíýóů](.|ch))|(. \*ou(|ch)))a"]

Minimum frequency:  [Apply](#)

Total: 200 items (4 pages)

	Filter	lemma	Freq	
1	p / n	chvála	40091	
2	p / n	moucha	26896	
3	p / n	pýcha	22185	
4	p / n	chlouba	20784	
5	p / n	chůva	16938	
6	p / n	mícha	9976	
7	p / n	psychóza	6732	
8	p / n	rychlodráha	6225	
9	p / n	chryzantéma	5073	
10	p / n	psychoanalýza	5027	
11	p / n	chiméra	3760	
12	p / n	heydrichiáda	3153	
13	p / n	cícha	1643	
14	p / n	jícha	1575	
15	p / n	schrána	1501	

[lc="(.\*[aeiouy](.|ch)).\*[aeiouy](.|ch)([yěeu](ou|ám|ách|am  
i))")]

Total: 1,267 items (26 pages)


	Filter	word	Freq	
1	p / n	smlouvy	6092	
2	p / n	sil	5617	
3	p / n	smlouvu	4183	
4	p / n	silou	3899	
5	p / n	houby	2699	
6	p / n	touhu	1832	
7	p / n	hub	1707	
8	p / n	touhy	1669	
9	p / n	mouky	1508	
10	p / n	hrůzou	1470	
11	p / n	smluv	1430	
12	p / n	vládou	1315	
13	p / n	ztrátou	1303	
14	p / n	smlouvě	1297	
15	p / n	silami	1148	
16	p / n	louky	1124	
17	p / n	touhou	1116	
18	p / n	mírou	1095	
19	p / n	měrou	1018	
20	p / n	trouby	1009	



# P-filtr

- NNFS7
- NNFP2
- NNFP3
- NNFP6
- NNFP7



# NNFS7

	Filter	word	Freq	
1	p / n	silou	3899	
2	p / n	hrůzou	1470	
3	p / n	vládou	1315	
4	p / n	ztrátou	1303	
5	p / n	touhou	1116	
6	p / n	mírou	1095	
7	p / n	měrou	1018	
8	p / n	zprávou	883	
9	p / n	mámou	861	
10	p / n	branou	789	
11	p / n	kávou	678	
12	p / n	ranou	670	
13	p / n	krásou	654	
14	p / n	atmosférou	628	
15	p / n	vírou	621	
16	p / n	bránou	609	
17	p / n	šťávou	605	
18	p / n	správou	597	
19	p / n	armádou	589	
20	p / n	analýzou	582	


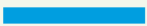

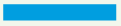
















# NNFP2

	<b>Filter</b>	<b>word</b>	<b>Freq</b>	
1	p / n	sil	5617	
2	p / n	hub	1707	
3	p / n	smluv	1430	
4	p / n	drah	1002	
5	p / n	ran	966	
6	p / n	skal	949	
7	p / n	žil	595	
8	p / n	krav	592	
9	p / n	dam	461	
10	p / n	čar	413	
11	p / n	trav	376	
12	p / n	bran	370	
13	p / n	par	358	
14	p / n	žab	308	
15	p / n	luk	296	
16	p / n	vah	294	
17	p / n	much	288	
18	p / n	premier	272	
19	p / n	lip	217	
20	p / n	mis	178	

# NNFP3



	Filter	word	Freq	
1	p / n	silám	292	
2	p / n	houbám ✓	109	
3	p / n	drahám	103	
4	p / n	mouchám ✓	94	
5	p / n	smlouvám ✓	74	
6	p / n	skalám ✓	73	
7	p / n	kravám	69	
8	p / n	ranám	69	
9	p / n	branám	55	
10	p / n	touhám ✓	48	
11	p / n	loukám ✓	24	
12	p / n	vahám	23	
13	p / n	žilám	22	
14	p / n	vzpourám ✓	16	
15	p / n	travám	11	
16	p / n	vranám	11	
17	p / n	boudám ✓	8	
18	p / n	čarám	7	
19	p / n	courám ✓	6	
20	p / n	kroupám ✓	6	

# NNFP6

	Filter	word	Freq	
1	p / n	silách	788	
2	p / n	žilách	647	
3	p / n	skalách	562	
4	p / n	loukách ✓	513	
5	p / n	drahách	237	
6	p / n	smlouvách ✓	170	
7	p / n	houbách ✓	134	
8	p / n	vahách	93	
9	p / n	ranách	92	
10	p / n	touhách ✓	43	
11	p / n	čarách	41	
12	p / n	boudách ✓	32	
13	p / n	stoupách ✓	31	
14	p / n	jamách	28	
15	p / n	branách	27	
16	p / n	mouchách ✓	22	
17	p / n	<u>sezonách</u>	21	
18	p / n	troubách ✓	17	
19	p / n	rourách ✓	16	
20	p / n	lipách	13	

# NNFP7

Total: 48 items (1 page)

	Filter	word	Freq	
1	p / n	silami	1148	
2	p / n	houbami ✓	317	
3	p / n	branami	297	
4	p / n	ranami	285	
5	p / n	skalami	259	
6	p / n	drahami	137	
7	p / n	smlouvami ✓	126	
8	p / n	žilami	120	
9	p / n	loukami ✓	116	
10	p / n	čarami	106	
11	p / n	touhami ✓	73	
12	p / n	kroupami ✓	60	
13	p / n	mouchami	56	
14	p / n	lipami ✓	49	
15	p / n	kravami	43	
16	p / n	parami	39	
17	p / n	travami	32	
18	p / n	rourami ✓	28	
19	p / n	boudami	26	
20	p / n	šmouhami ✓	24	

# Tvrzení

- Ne všechna lemmata substantiv skloňovaných podle vzoru *žena*, která mají ve kmeni dlouhou samohlásku nebo diftong, mají v češtině alternaci (*půda*, *armáda*, *atmosféra*, *kariéra*, *krása*, *hrůza*, ...).
- Můžeme tedy tvrdit, že alternace není obligatorní.

# Uvažujme

- Lze se zamýšlet nad tím, zda se v případě dvojic jako *sezóna/sezon*, *disciplína/disciplin*, *turbína/turbin*, *premiéra/premier*, ... , ale třeba i *bába/bab* jedná o hledaný jev.



# turbin

od srdce zmizel . Celý	<b>Turbin</b> /Turbin/NNMS1-----A-----	byl vzhůru , i z
použitím fotovoltaických panelů , větrných	<b>turbin</b> /turbína/NNFP2-----A-----	, systém sběru dešťové vody
lze zaznamenat i v kategorii	<b>turbin</b> /turbína/NNFP2-----A-----	malých a nejmenších . Spaliny
nichž každá by měla deset	<b>turbin</b> /turbína/NNFP2-----A-----	o výkonu 5 000 koní ,
hodila ho na něj .	<b>Turbin</b> /turbína/NNFP2-----A-----	není Varšava . Celé město
Do roku 2012 se těchto	<b>turbin</b> /turbína/NNFP2-----A-----	prodalo 150 , z toho
vyšší , než u původních	<b>turbin</b> /turbína/NNFP2-----A-----	třídy E. Zdokonalení umožňuje dosahovat
včetně extrémně náročných tvarů lopatek	<b>turbin</b> /turbína/NNFP2-----A-----	( Mägerle AG Maschinenfabrik )
panelů , vodních a větrných	<b>turbin</b> /turbína/NNFP2-----A-----	očekávám nástup mikrokogenerací , kombinované
i s dalšími druhy vodních	<b>turbin</b> /turbína/NNFP2-----A-----	. Řezy soustrojím nabízejí unikátní
fy Saacke ) a prodejce	<b>turbin</b> /turbína/NNFP2-----A-----	Capstone ( fy Greenviroment )
85 % . Vývoj plynových	<b>turbin</b> /turbína/NNFP2-----A-----	však sleduje i zlepšování řady
generátorů . Voda odtékající z	<b>turbin</b> /turbína/NNFP2-----A-----	by se po třech minutách
disků do počítačů a také	<b>turbin</b> /turbína/NNFP2-----A-----	do větrných elektráren . I
v destilovaném hukotu kotlů a	<b>turbin</b> /turbína/NNFP2-----A-----	by přemýšlel o skutečném světě
pročátku nabízení pro 2 EMW sčítá	<b>turbin</b> /turbína/NNFP2-----A-----	společnosti GE včetně nové

# turbín

se návštěvníci seznámí s fungováním	turbín /turbína/NNFP2-----A-----	v tepelných elektrárnách pod hlavičkou
Tyto požadavky vedou u parních	turbín /turbína/NNFP2-----A-----	velkých výkonů k návrhům dlouhých
Čína . Celkem stovka větrných	turbín /turbína/NNFP2-----A-----	za téměř jednu miliardu dolarů
s výrobou strojů pro obrábění	turbín /turbína/NNFP2-----A-----	dlouhou zkušenost , a dlouhou
. Zatímco vrtule dnešních větrných	turbín /turbína/NNFP2-----A-----	se otáčejí vertikálně , 275
pokrývá tři základní typy větrných	turbín /turbína/NNFP2-----A-----	pro provoz na pevnině :
představil unikátní projekt gigantických větrných	turbín /turbína/NNFP2-----A-----	Aerogenerator X , jejichž výkon
. Skelný povlak poutající lopatky	turbín /turbína/NNFP2-----A-----	se ochladil , popraskal ,
dílů leteckých motorů a plynových	turbín /turbína/NNFP2-----A-----	tak jistě najde uplatnění při
s níž lze měnit výkon	turbín /turbína/NNFP2-----A-----	a reagovat tak na aktuální
podepsali kontrakt na dodávku parních	turbín /turbína/NNFP2-----A-----	ŠKODA 3x133 MW pro tureckou
vysoce přesné díly do parních	turbín /turbína/NNFP2-----A-----	. Jedná se o opakovanou
Ve výkonové řadě našich větrných	turbín /turbína/NNFP2-----A-----	je dnes nejprodávanější stroj o
spočítal , kolik stometrových jednomegawattových	turbín /turbína/NNFP2-----A-----	bychom museli postavit . K
výroby , která třeba u	turbín /turbína/NNFP2-----A-----	, lokomotiv , jeřábů či

# IJP

## turbína

dělení: tur-bí-na<sup>1</sup>

rod: ž.

	jednotné číslo	množné číslo
1. pád	turbína	turbíny
2. pád	turbíny	turbín
3. pád	turbíně	turbínám
4. pád	turbínu	turbíny
5. pád	turbíno	turbíny
6. pád	turbíně	turbínách
7. pád	turbínou	turbínami

# Slovníky

## SSČ

**turbína**, -y ž <l> *rotační lopatkový motor*: vodní, parní, spalovací turbína;  
**turbínový** příd.: turbínové kolo

## SSJČ

**turbína, turbina**, -y ž. (z lat. zákl.)

1. **tech.** *rotační lopatkový motor, v kt. se energie pohonné látky mění v mechanickou práci*: vodní, tepelná, parní, plynová, spalovací t.; rovnotlaká, dvojtlaká, přetlaková t.; axiální, radiální t.; spirální t.; vrtulová t.; Kaplanova t.; hltnost t-y
2. **zast. slang.** *tryskové letadlo*: t-y se snášely do ulic a vraždily lidi (Pujm.);  
→ **zdrob.** **turbínka, turbinka**, -y ž.: t. zubní vrtačky; - **zast. slang.** letadla, kterým se říkalo t-y (Vach.);  
— **turbínový, turbinový** v. t.

















# turbin.\*

Total: 19 items (1 page)

	Filter	word	Freq	
1	p / n	turbin	16	
2	p / n	turbines	14	
3	p / n	turbiny	14	
4	p / n	turbinu	5	
5	p / n	turbina	3	
6	p / n	turbinám	3	
7	p / n	turbině	2	
8	p / n	turbinář	2	
9	p / n	turbine	2	
10	p / n	turbinou	2	
11	p / n	turbinatek	1	
12	p / n	turbinovou	1	
13	p / n	turbinová	1	
14	p / n	turbinové	1	
15	p / n	turbinových	1	
16	p / n	turbinovými	1	
17	p / n	turbinatky	1	
18	p / n	turbinku	1	
19	p / n	turbinare	1	

# Rovnocennost variant















- lemma=síla

	Filter	lc [lowercase word]	tag	Freq	
1	p / n	sílu	NNFS4----A----	6836	
2	p / n	sil	NNFP2----A----	5617	
3	p / n	síla	NNFS1----A----	4767	
4	p / n	silou	NNFS7----A----	3899	
5	p / n	síly	NNFS2----A----	3596	
6	p / n	síly	NNFP4----A----	2951	
7	p / n	síly	NNFP1----A----	1827	
8	p / n	síle	NNFS6----A----	1456	
9	p / n	silami	NNFP7----A----	1148	
10	p / n	silách	NNFP6----A----	788	
11	p / n	síle	NNFS3----A----	471	
12	p / n	silám	NNFP3----A----	292	
13	p / n	sílou	NNFS7----A--1-	79	
14	p / n	sílo	NNFS5----A----	11	
15	p / n	sílách	NNFP6----A--1-	2	
16	p / n	sílemi	NNFP7----A--1-	1	
















# Rovnocennost variant

- lemma=pára

Total: 17 items (1 page)

	Filter	lc [lowercase word]	tag	Freq	
1	p / n	pára	NNFS1----A----	772	
2	p / n	par	NNFP2----A----	358	
3	p / n	páry	NNFS2----A----	339	
4	p / n	páru	NNFS4----A----	338	
5	p / n	páře	NNFS6----A----	186	
6	p / n	párou	NNFS7----A--1-	168	
7	p / n	párou	NNFS7----A----	154	
8	p / n	parami	NNFP7----A--1-	39	
9	p / n	par.	NNFP2----A----	34	
10	p / n	páry	NNFP4----A----	25	
11	p / n	páře	NNFS3----A----	21	
12	p / n	páry	NNFP1----A----	18	
13	p / n	parách	NNFP6----A--1-	7	
14	p / n	parám	NNFP3----A--1-	4	
15	p / n	párami	NNFP7----A----	3	

# Rovnocennost variant a platnost tvrzení

	Filter	lc [lowercase word]	tag	Freq	
1	p / n	mísy	NNFS2---A---	756	
2	p / n	mísu	NNFS4---A---	586	
3	p / n	míse	NNFS6---A---	373	
4	p / n	mísa	NNFS1---A---	253	
5	p / n	<u>mis</u>	NNFP2---A--1-	178	
6	p / n	mísy	NNFP1---A---	135	
7	p / n	mísou	NNFS7---A---	120	
8	p / n	mísy	NNFP4---A---	119	
9	p / n	<u>mís</u>	NNFP2---A---	88	
10	p / n	mísami	NNFP7---A---	49	
11	p / n	míse	NNFS3---A---	42	
12	p / n	mísách	NNFP6---A---	24	
13	p / n	mísám	NNFP3---A---	5	
14	p / n	míso	NNFS5---A---	1	
15	p / n	d-mis	NNFP2---A--1-	1	



# Porovnání

Výsledky pak můžeme zapsat do tabulky například takto:

<i>síla</i>	gen. pl. (-0)	instr. sg. (-ou)	dat. pl. (-ám)	lok. pl. (-ách)	instr. pl. (-ami)
<i>sil-</i>	6316	3684	291	895	1181
<i>síl-</i>	0	75	1	2	2
<i>smluv-</i>	1796	0	0	0	0
<i>smlouv-</i>	0	919	85	330	180
<i>drah-</i>	11	45	106	196	112
<i>dráh-</i>	0	243	3	22	48

# Jaká je distribuce hláskových alternací u maskulin?

- [tag="NN[IM].\*" & lemma="(\*[áéíýóů].\* | \*.ou.\*)"]
- N-filtr – wordform
- .\*([áéíóůý] | ou).\*

	<b>Filter</b>	<b>word</b>	<b>Freq</b>	
1	p / n	případě	41355	
2	p / n	peníze	25338	
3	p / n	system	22603	
4	p / n	problém	22085	
5	p / n	domu	22042	
6	p / n	způsobem	20930	
7	p / n	systemu	19930	
8	p / n	milionů	19697	
9	p / n	rámci	19575	
10	p / n	dům	17390	

	<b>Filter</b>	<b>word</b>	<b>Freq</b>	
1	p / n	domu	22042	
2	p / n	pane	13881	
3	p / n	peněz	11968	
4	p / n	stolu	11409	
5	p / n	domě	10448	
6	p / n	milionu	7253	
7	p / n	koně	5972	
8	p / n	miliony	5530	
9	p / n	domy	5444	
10	p / n	stole	4648	
11	p / n	vozu	4302	
12	p / n	boha	4285	
13	p / n	sněhu	4097	
14	p / n	kameny	3800	
15	p / n	větru	3722	
16	p / n	chleba	3415	
17	p / n	bože	3371	
18	p / n	domem	3110	
19	p / n	bohu	2769	
20	p / n	vozy	2606	

# stadion – salon - balkon

## stadion

dělení: sta-dion<sup>1</sup>

Ize i: stadión<sup>2</sup>

rod: m. neživ.

	jednotn
1. pád	stad
2. pád	stadi
3. pád	stad
4. pád	stad
5. pád	stadi
6. pád	stadioně,
7. pád	stadi

## salon

dělení: sa-lon<sup>1</sup>

Ize i: salón<sup>2</sup>

rod: m. neživ.

	jedno
1. pád	sa
2. pád	sa
3. pád	sa
4. pád	sa
5. pád	sa
6. pád	salonu
7. pád	sal

## balkon

dělení: bal-kon<sup>1</sup>

Ize i: balkón<sup>2</sup>

rod: m. neživ.

	jednotné číslo	množné číslo
1. pád	balkon	balkony
2. pád	balkonu <sup>3</sup>	balkonů
3. pád	balkonu	balkonům
4. pád	balkon	balkony
5. pád	balkone <sup>4</sup>	balkony
6. pád	balkoně, balkonu <sup>5</sup>	balkonech
7. pád	balkonem	balkony

# Úkol na 9. 12. 2019

- Vyhledejte v korpusu substantiva typu *duše*, která mají ve kmeni samohláskovou alternaci v rámci flexe.
- Vyberte si jev z české gramatiky a zpracujte ukázkou, jak byste na základě a) vyhledání dat v korpusech, b) pozorování vyhledaných dat a c) zobecnění učiněných na základě pozorování gramatický jev vysvětlili žákům/studentům.