

CJBB84

Morfologické varianty a dublety – korpus jako zdroj dat pro observace a generalizace

(variantní a dubletní koncovky v české substantivní flexi)

-i/-ové/-é

- [tag="N.MP1.*"]

		<u>word</u>	<u>Frekvence</u>
1.	p/n	lidé	43 768
2.	p/n	Lidé	10 613
3.	p/n	muži	9 968
4.	p/n	rodiče	7 003
5.	p/n	policisté	6 655
6.	p/n	Češi	5 178
7.	p/n	hráči	4 861
8.	p/n	členové	4 777
9.	p/n	hosté	4 099
10.	p/n	obyvatelé	4 060

-i

Filtr konkordancí

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od do včetně KWIC

Typ dotazu: ▼

[klávesnice](#) | [předchozí dotazy](#)

Slovní tvar:

muži – Češi - zástupci

		lemma	lc	Frekvence
1.	p/n	muž	muži	12 210
2.	p/n	hráč	hráči	5 820
3.	p/n	Čech	češi	5 267
4.	p/n	lékař	lékaři	5 126
5.	p/n	člověk	lidi	5 086
6.	p/n	kluk	kluci	4 530
7.	p/n	student	studenti	4 436
8.	p/n	voják	vojáci	4 428
9.	p/n	návštěvník	návštěvníci	4 247
10.	p/n	divák	diváci	4 208
11.	p/n	Němec	němci	4 030
12.	p/n	zástupce	zástupci	3 951
13.	p/n	odborník	odborníci	3 882
14.	p/n	politik	politici	3 686
15.	p/n	hasič	hasiči	3 611
16.	p/n	řidič	řidiči	3 514
17.	p/n	zaměstnanec	zaměstnanci	3 393
18.	p/n	vědec	vědci	3 274
19.	p/n	pracovník	pracovníci	3 262
20.	p/n	žák	žáci	3 011

?vzor *předseda* [lemma=".*a"]

		lemma	lc	Frekvence
1.	p/n	komunista	komunisti	346
2.	p/n	nacista	nacisti	91
3.	p/n	turista	turisti	80
4.	p/n	fotbalista	fotbalisti	62
5.	p/n	policista	policisti	45
6.	p/n	hokejista	hokejisti	41
7.	p/n	terorista	teroristi	36
8.	p/n	bandita	banditi	36
9.	p/n	fašista	fašisti	31
10.	p/n	cyklista	cyklisti	30
11.	p/n	socialista	socialisti	29
12.	p/n	evangelista	evangelisti	25
13.	p/n	artista	artisti	16
14.	p/n	civilista	civilisti	15
15.	p/n	rasista	rasisti	14
16.	p/n	humanista	humanisti	14
17.	p/n	vrchnosta	vrchnosti	14
18.	p/n	specialista	specialisti	14
19.	p/n	transvestita	transvestiti	13
20.	p/n	sadista	sadisti	11

-ové

Filtr konkordancí

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od do včetně KWIC

Typ dotazu:

[klávesnice](#) | [předchozí dotazy](#)

Slovní tvar:

členové – vítězové - starostové

	lemma	lc	Frekvence
1.	p/n člen	členové	5 958
2.	p/n pán	pánové	2 647
3.	p/n kolega	kolegové	2 391
4.	p/n Rus	rusové	1 441
5.	p/n Brit	britové	1 147
6.	p/n Rom	romové	1 030
7.	p/n syn	synové	922
8.	p/n hrdina	hrdinové	918
9.	p/n šéf	šéfové	841
10.	p/n předek	předkové	736
11.	p/n vítěz	vítězové	698
12.	p/n Italos	italové	629
13.	p/n starosta	starostové	629
14.	p/n muslim	muslimové	616
15.	p/n otec	otcové	601
16.	p/n archeolog	archeologové	600
17.	p/n ekonom	ekonomové	593
18.	p/n detektiv	detektivové	580
19.	p/n bůh	bohové	570
20.	p/n řek	řekové	563

?vzor *soudce* [lemma=".*[eě]"]

		<u>lemma</u>	<u>lc</u>	<u>Frekvence</u>
1.	p/n	vůdce	vůdcové	176
2.	p/n	vládce	vládcové	110
3.	p/n	soudce	soudcové	53
4.	p/n	rádce	rádcové	20
5.	p/n	Yankee	yankeeové	18
6.	p/n	strážce	strážcové	17
7.	p/n	Montgolfiere	montgolfierové	8
8.	p/n	Coleridge	coleridgeové	7
9.	p/n	zástupce	zástupcové	6
10.	p/n	Boyle	doylové	5
11.	p/n	Bonaparte	buonapartové	4
12.	p/n	Babbage	babbageové	3
13.	p/n	odpůrce	odpůrcové	3
14.	p/n	Goethe	goethové	2
15.	p/n	obránce	obrāncové	2
16.	p/n	Hawthorne	hawthorneové	2
17.	p/n	průvodce	průvodcové	2
18.	p/n	Claude	claudové	2
19.	p/n	Robespierre	robspierrové	2
20.	p/n	nástupce	nástupcové	2

-é

Filtr konkordancí

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od do včetně KWIC

Typ dotazu:

[vložit tag](#) | [vložit "within"](#) | [klávesnice](#) | [předchozí dotazy](#)

```
[lc!=".*ové" & lc=".*é"]
```

CQL:

lidé – policisté- hosté – obyvatelé - občané

	lemma	lc	Frekvence
1.	p/n člověk	lidé	54 602
2.	p/n policista	policisté	9 778
3.	p/n host	hosté	5 860
4.	p/n obyvatel	obyvatelé	5 113
5.	p/n přítel	přátelé	4 061
6.	p/n fotbalista	fotbalisté	3 119
7.	p/n občan	občané	2 883
8.	p/n majitel	majitelé	2 659
9.	p/n Američan	američané	2 614
10.	p/n zastupitel	zastupitelé	2 504
11.	p/n představitel	představitelé	2 297
12.	p/n demokrat	demokraté	2 222
13.	p/n komunista	komunisté	2 144
14.	p/n manžel	manželé	2 123
15.	p/n turista	turisté	1 938
16.	p/n hokejista	hokejisté	1 938
17.	p/n učitel	učitelé	1 827
18.	p/n pořadatel	pořadatelé	1 824
19.	p/n podnikatel	podnikatelé	1 773
20.	p/n žid	židé	1 658

?vzor soudce [lemma=".*[eě]"]

Výskytů: 0 | i.p.m. [?]: 0,00 (vztaženo k celému syn2010) | Výsledek je promíchn

-e/-ě

Filtr konkordancí

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od do včetně KWIC

Typ dotazu: ▼

[klávesnice](#) | [předchozí dotazy](#)

Slovní tvar:

rodiče – koně - prarodiče

		lemma	lc	Frekvence
1.	p/n	rodič	rodiče	8 575
2.	p/n	kůň	koně	1 264
3.	p/n	prarodič	prarodiče	243
4.	p/n	Rose	rose	115
5.	p/n	Cave	cave	95
6.	p/n	France	france	75
7.	p/n	Bruce	bruce	59
8.	p/n	Grace	grace	57
9.	p/n	svatyně ↵	svatyně	56
10.	p/n	Lee	lee	54
11.	p/n	People	people	31
12.	p/n	Charlie	charlie	27
13.	p/n	Stevie	stevie	23
14.	p/n	Eddie	eddie	23
15.	p/n	Stone	stone	21
16.	p/n	Arne	arne	16
17.	p/n	Šebrle	šebrle	13
18.	p/n	Pete	pete	12
19.	p/n	Line	line	12
20.	p/n	Porsche	porsche	12

dublety

- Které kombinace jsou možné?
- -ové/-é ?
- *rodiči, koni* jako nom. pl. ?
- Co je důvodem absence koncovky *-é* u vzoru *soudce*?

-ech/-ích/-ách

- [tag="N.[MIN]P6.*"]

		lemma		lc		Frekvence
1.	p/n	rok		letech		36 834
2.	p/n	den		dnech		9 317
3.	p/n	místo		místech		9 157
4.	p/n	případ		případech		8 808
5.	p/n	USA		usa		4 240
6.	p/n	měsíc		měsících		4 143
7.	p/n	záda		zádech		4 013
8.	p/n	schod		schodech		3 926
9.	p/n	stát		státech		3 725
10.	p/n	město		městech		3 625
11.	p/n	týden		týdnech		2 273
12.	p/n	zápas		zápasech		2 155
13.	p/n	člověk		lidech		2 068
14.	p/n	šaty		šatech		1 880
15.	p/n	obchod		obchodech		1 604
16.	p/n	vztah		vztazích		1 590
17.	p/n	trh		trzích		1 548
18.	p/n	les		lesích		1 542
19.	p/n	dům		domech		1 500

? -ách

Filtr konkordancí

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od do včetně KWIC

Typ dotazu:

[klávesnice](#) | [předchozí dotazy](#)

Slovní tvar:

kolečko - sen

		<u>lemma</u>	<u>lc</u>	<u>Frekvence</u>
1.	p/n	kasárna	kasárnách	322
2.	p/n	Jeseníky	jeseníkách	290
3.	p/n	džínsy	džínách	287
4.	p/n	kolečko	kolečkách	201
5.	p/n	luka	lukách	196
6.	p/n	nosítko	nosítkách	161
7.	p/n	sen	snách	132
8.	p/n	tričko	tričkách	119
9.	p/n	schůdky	schůdkách	104
10.	p/n	městečko	městečkách	88
11.	p/n	muka	mukách	79
12.	p/n	tepláky	teplákách	73
13.	p/n	nosítka	nosítkách	71
14.	p/n	oko	okách	71
15.	p/n	sluchátko	sluchátkách	67
16.	p/n	dvířka	dvířkách	63
17.	p/n	Roztoky	roztokách	59
18.	p/n	kousek	kouskách	58
19.	p/n	víčko	víčkách	55
20.	p/n	lehátko	lehátkách	52

? M [tag="NNM.*"]

		lemma	lc	Frekvence
1.	p/n	pták	ptákách	18
2.	p/n	kluk	klukách	9
3.	p/n	Novák	novákách	7
4.	p/n	Lázek	lázkách	3
5.	p/n	stříbrník	stříbrníkách	3
6.	p/n	dingů	dingách	2
7.	p/n	Aljoška	aljoškách	2
8.	p/n	šach	šachách	2
9.	p/n	králík	králíkách	2
10.	p/n	Dražek	dražkách	2
11.	p/n	rudoch	rudochách	2
12.	p/n	chlápek	chlápkách	1
13.	p/n	Polák	polákách	1
14.	p/n	tatínek	tatínkách	1
15.	p/n	zpěvák	zpěvákách	1
16.	p/n	důstojník	důstojníkách	1
17.	p/n	brouček	broučkách	1
18.	p/n	mazlík	mazlíkách	1
19.	p/n	žák	žákách	1
20.	p/n	řemesník	řemesníkách	1


Distribuce -ích/-ech

- Měkké vzory
- Tvrdé vzory
- *?lesích*

-0 v gen. pl.

- -ů
- -í

Hledat v korpusu

Korpus: : 

Typ dotazu:

[vložit tag](#) | [vložit "within"](#) | [klávesnice](#) | [předchozí dotazy](#)

CQL:

koruna - procento

		lemma	lc	Frekvence
1.	p/n	rok	let	83 737
2.	p/n	koruna	korun	52 450
3.	p/n	hodina	hodin	31 278
4.	p/n	procento	procent	22 013
5.	p/n	minuta	minut	19 242
6.	p/n	peníze	peněz	12 365
7.	p/n	strana	stran	10 106
8.	p/n	služba	služeb	9 669
9.	p/n	žena	žen	9 315
10.	p/n	místo	míst	8 801
11.	p/n	slovo	slov	8 150
12.	p/n	osoba	osob	7 775
13.	p/n	firma	firem	7 718
14.	p/n	miliarda	miliard	7 191
15.	p/n	podmínka	podmínek	6 568
16.	p/n	síla	sil	6 304
17.	p/n	škola	škol	6 283
18.	p/n	euro	eur	6 047
19.	p/n	ruka	rukou	6 015
20.	p/n	město	měst	5 945

? kromě *žena/město*

Filtr konkordancí

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od do včetně KWIC

Typ dotazu:

[vložit tag](#) | [vložit "within"](#) | [klávesnice](#) | [předchozí dotazy](#)

CQL: `[lemma=".*ao"]`

Peněz – přátel – Čech – zvířat - silnic

	<u>lemma</u>	<u>lc</u>	<u>Frekvence</u>
1.	p/n rok	let	83 737
2.	p/n peníze	peněz	12 365
3.	p/n obyvatel	obyvatel	5 350
4.	p/n Čechy	čech	4 449
5.	p/n zvíře	zvířat	4 378
6.	p/n datum	dat	4 074
7.	p/n USA	usa	4 026
8.	p/n noviny	novin	3 663
9.	p/n dějiny	dějín	3 452
10.	p/n přítel	přátel	3 355
11.	p/n silnice	silnic	2 854
12.	p/n ulice	ulic	2 740
13.	p/n hranice	hranic	2 294
14.	p/n centrum	center	2 031
15.	p/n investice	investic	2 017
16.	p/n plíce	plic	1 737
17.	p/n stanice	stanic	1 730
18.	p/n nemocnice	nemocnic	1 539
19.	p/n Windows	windows	1 510
20.	p/n Pardubice	pardubic	1 453

?vzor *moře*

Filtr konkordancí

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od do včetně KWIC

Typ dotazu: ▼

[vložit tag](#) | [vložit "within"](#) | [klávesnice](#) | [předchozí dotazy](#)

```
[tag="N.N.*" & lemma=".*[eě]"]
```

CQL:

kromě vzoru *kuře*

	lemma	lc	Frekvence
1.	p/n zvíře	zvířat	4 378
2.	p/n vejce	vajec	697
3.	p/n děvče	děvčat	543
4.	p/n hřiště	hřišť	541
5.	p/n mládě	mláďat	510
6.	p/n pracoviště	pracovišť	434
7.	p/n dvojče	dvojčat	387
8.	p/n prase	praset	372
9.	p/n rajče	rajčat	368
10.	p/n stanoviště	stanovišť	319
11.	p/n sídliště	sídlišť	306
12.	p/n promile	promile	262
13.	p/n letiště	letišť	262
14.	p/n sportoviště	sportovišť	247
15.	p/n schodiště	schodišť	220
16.	p/n parkoviště	parkovišť	203
17.	p/n naleziště	nalezišť	184
18.	p/n kníže	knížat	163
19.	p/n kuře	kuřat	151
20.	p/n vnouče	vnoučat	134

?vzor *růže*

Filtr konkordancí

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od do včetně KWIC

Typ dotazu: ▼

[vložit tag](#) | [vložit "within"](#) | [klávesnice](#) | [předchozí dotazy](#)

CQL:

Kromě na *-ice*

		lemma	lc	Frekvence
1.	p/n	silnice	silnic	2 854
2.	p/n	ulice	ulic	2 740
3.	p/n	hranice	hranic	2 294
4.	p/n	investice	investic	2 017
5.	p/n	plíce	plic	1 737
6.	p/n	stanice	stanic	1 730
7.	p/n	nemocnice	nemocnic	1 539
8.	p/n	Pardubice	pardubic	1 453
9.	p/n	částice	částic	1 418
10.	p/n	Budějovice	budějovic	1 226
11.	p/n	míle	mil	1 213
12.	p/n	dálnice	dálnic	1 080
13.	p/n	vesnice	vesnic	1 077
14.	p/n	pozice	pozic	983
15.	p/n	tradice	tradic	823
16.	p/n	Vánoce	vánoc	796
17.	p/n	samice	samic	679
18.	p/n	Teplice	teplic	635
19.	p/n	Vítkovice	vítkovic	510
20.	p/n	dvojice	dvojic	497

N-filtr

		lemma	lc	Frekvence
1.	p/n	plíce	plic	1 737
2.	p/n	míle	mil	1 213
3.	p/n	Vánoce	vánoc	796
4.	p/n	chvíle	chvil	479
5.	p/n	Krkonoše	krkonoš	353
6.	p/n	lžíce	lžic	249
7.	p/n	Velikonoce	velikonoc	180
8.	p/n	košile	košil	176
9.	p/n	neděle	neděl	174
10.	p/n	přítelkyně	přítelkyň	129
11.	p/n	kolegyně	kolegyň	98
12.	p/n	žákyně	žákyň	90
13.	p/n	mistryně	mistryň	74
14.	p/n	zaměstnankyně	zaměstnankyň	52
15.	p/n	bohyně	bohyň	47
16.	p/n	jeskyně	jeskyň	44
17.	p/n	umělkyně	umělkyň	31
18.	p/n	svěřenkyně	svěřenkyň	25
19.	p/n	svědkyně	svědkyň	24
20.	p/n	zájemkyně	zájemkyň	23

Domácí úkol na 9. 11. 2016

- Podobným způsobem zpracujte pravidla distribuce koncovek *-u/-e/-i* ve vokativu sg. maskulin