

CJBB84

Morfologie a korpus

8.00-9.30 G13

VI.

O čem budeme mluvit

- Spojení typu „*jízda po italsku*“ versus „*jízda po Španělsku*“
- Co říká IJP
- Jak hledat data
- Jak je vyřešen problém lemmatizace / taggování
- Slovnědruhový přesah a adverbializace
- Homonymie a kolokace

Co říká IJP


- Přísllovečné spřežky píšeme dohromady –[?] jako jedno slovo **Hranice** chápání určitého **výrazu jako spřežky však není** ve všech případech **jasná**, často dochází ke **kolísání**, a nezřídka proto vedle sebe existují jak podoby psané zvlášť, tak dohromady a jejich **význam je totožný (na příklad –[?] například)**. Na druhou stranu se u některých příslovci již dlouho neužívá jejich výchozí, dvouslovná podoba a jsou chápána jako spřežky jen z vývojových, etymologických důvodů (**vzhůru, zpět** atp.).
- V ustrnulých spojeních typu **kapr po mlynářsku, pstruh po námořnicku** píšeme předložku zvlášť (**na rozdíl od ustálených spojení poslepu i po slepu, postaru i po staru, podomácku i po domácku**, která můžeme psát dvojím způsobem); **zeměpisná jména** v nich píšeme s **malým** počátečním písmenem: **vepřové po německu, nákyt po francouzsku**, stejně tak lze **zmizet po anglicku**.

Nedodržení normy: po.*[sc]ku

omáčce s maniokem . Čekal jsem , že budu jíst	poafricku/poafricku/poafricku/X@-----	a že nejspíš budu hodně hubnout .
výbušninu upevněnou na svém těle nebo střílejí na izraelské cíle	podomácku/podomácku/podomácku/Db-----	vyrobené rakety . Podobné příběhy
černocho v bílém laboratorním plášti . Všichni vlci byli vyzbrojeni	podomácku/podomácku/podomácku/Db-----	vyrobeným zbraněmi – klacky , me
televize CNN . Atentátníci potřebovali jen kalašnikovovy a munici ,	podomácku/podomácku/podomácku/Db-----	vyrobené vesty a pásy s výbušnina
pole . Na dalšího traktoristu u Horní Stropnice se převrátil	podomácku/podomácku/podomácku/Db-----	vyrobený stroj , za kterým táhl stro
všech strojích a z šuplíku vytáhl láhev domácí pálenky a	podomácku/podomácku/podomácku/Db-----	vyrobenou výbušninu . Připravil výl
zpět . Prošli jsme kolem muže , který hrál na	podomácku/podomácku/podomácku/Db-----	vyrobenou kytaru a zpíval , zatímco
Kč , které bylo italské nejen transkripcí . Mezinárodní menu	počesku/počesku/počesku/Db-----	zakočil rebarborový koláč se sko

Řídké doklady psaní dohromady

Celkem: 20 položek (1 stránka)

	Filtr	lemma	Freq	
1	p / n	podomácku	230	
2	p / n	počesku	3	
3	p / n	poafricku	2	
4	p / n	powerpacku	2	
5	p / n	pokucku	1	
6	p / n	pošpanělsku	1	
7	p / n	poČeskoslovensku	1	
8	p / n	podtácku	1	
9	p / n	popelkovsku	1	
10	p / n	polsku	1	
11	p / n	povstánívMaďarsku	1	
12	p / n	pobulharsku	1	
13	p / n	porotterdamsku	1	
14	p / n	pořimsku	1	
15	p / n	pobritsku	1	
16	p / n	posousedsku	1	
17	p / n	policku	1	
18	p / n	poustevnicku	1	
19	p / n	policievMaďarsku	1	
20	p / n	politikyvNěmecku	1	

Jak hledat data: po .*sku

Korpus: [syn v8](#) | Dotaz: [po, .*sku](#) (64 076 výskytů) ▶ Promíchat: ✓

Výskytů: **64 076** | i.p.m.: **11,88** (vztaženo k celému korpusu) | ARF: **37 304,61** | Výsledek je promíchán

1 / 1 602 ▶▶▶

Výběr řádků: základní ▾

<input type="checkbox"/>	Deníky Bohemia	, ale naštěstí nechyboval , " uvedl jezdec . I	po/po/RR--6-----zisku/zisk/NNIS6-----A-----	18 bodů do žebříčku Skokovéhoho
<input type="checkbox"/>	Lidové noviny	: USA chtějí data o účtech CURYCH Spojené státy požadují	po/po/RR--6-----Švýcarsku/švýcarsko/NNNS6-----A-----	, aby jim v tomto týdnu předalo p
<input type="checkbox"/>	Ona DNES	trhovce , V ústech studny , jak se město jmenuje	po/po/RR--6-----maysku/maysku/X@-----	, stojí za to nechat několik litrů p
<input type="checkbox"/>	Lidové noviny	I obránce Patrice Evra přiznal : „ Postoupili jsme typicky	po/po/RR--6-----italsku/italský/ACNS6-----A-----	. Hráli jsme ošklivý fotbal , který
<input type="checkbox"/>	Deníky Bohemia	Wood udělali od posledního setkání v roce 2005 . Recepty	po/po/RR--6-----plzeňsku/plzeňsko/NNNS6-----A-----	Praha – Pořad Prostřeno ! dosta
<input type="checkbox"/>	Deníky Moravia	zničil . Končím , oznámil trenér Růžička Praha – Kariéra	po/po/RR--6-----česku/český/ACNS6-----A-----	? Z hrdiny bídákem snadno a ryc
<input type="checkbox"/>	Lidové noviny	abychom se měli rádi . My se máme rádi tak	po/po/RR--6-----kocourkovsku/kocourkovsku/X@-----	a " Kocourkov/ má jako kočka / c
<input type="checkbox"/>	Regionální týdeník	dělil od řeky železniční násep . Voda z něj kousek	po/po/RR--6-----kousku/kousek/NNIS6-----A-----	ukrajovala . Pokud by se protrhl ,
<input type="checkbox"/>	Lidové noviny	jako „ výprodeje rodinného stříbra " cizákům . Anglické tradice	po/po/RR--6-----česku/český/ACNS6-----A-----	Pithartovi na počátku 90 . let příl
<input type="checkbox"/>	Lidové noviny	bez tureckého pasu , asi by se to vyřešilo pěkně	po/po/RR--4-----putinovsku/putinovsku/X@-----	nějakou tou neprodlouženou bun
<input type="checkbox"/>	Deníky Moravia	autobusy budou navazovat další spoje , které rozvázejí lidi dál	po/po/RR--6-----Znojemsku/znojemsko/NNNS6-----A-----	. Protože autobus nikde nezasta
<input type="checkbox"/>	Deníky Bohemia	začíná zítra Chrudimsko – V sobotu se budou muset řidiči projíždějící	po/po/RR--6-----Chrudimsku/chrudimsko/NNNS6-----A-----	obrnit trpělivostí , neboť v úsecíc

Pozorování a filtrování

- a) frazeologismy: *kousek po kousku*
- b) substantiva **zisk**, *loňsko*, **výtisk**

lemma=.*sk

Celkem: 63 položek (2 stránky)

	Filtr	lemma	Freq	
1	p / n	po zisk	10328	
2	p / n	po výprask	1913	
3	p / n	po stisk	779	
4	p / n	po rozbřesk	217	
5	p / n	po potlesk	145	
6	p / n	po disk	89	
7	p / n	po tisk	74	
8	p / n	po záblesk	69	
9	p / n	po lesk	43	
10	p / n	po blesk	40	
11	p / n	po výtisk	39	
12	p / n	po Blesk	37	
13	p / n	po Minsk	35	
14	p / n	po vosk	31	
15	p / n	po třesk	29	

lemma=rusk tag=NN[IM]



mělo velice pomoci . Doufáme , že dlouhý okružní let **po/po/RR--6----- Rusku/rusk/NNIS6-----A-----** , který podniká Clemmie , bude především bezpečný a kromě

. Nový systém připomíná výběr poplatků za pohyb nákladních aut **po/po/RR--6----- Rusku/rusk/NNIS6-----A-----** , který se nazývá Platon . Bez ohledu na protesty

Korpus: [syn v8](#) | Dotaz: [rusk](#) (5 892 výskytů) ▶ Promíchat: ▶ Pozitivní filtr: [rusk](#) (192 výskytů)

Výskytů: **192** | i.p.m.: **0,04** (vztaženo k celému korpusu) | ARF: **66,17** | Výsledek je promíchán

1 / 5 ▶▶▶

Výběr řádků:

- | | | |
|--------------------------|-------------------------------|---|
| <input type="checkbox"/> | Právo | tvrdost sovětských požadavků vůči československému vedení , odpověděl pak Dean Rusk/rusk/NNMS1-----A----- , že " se jedná o citové otázky " , |
| <input type="checkbox"/> | Právo | dohodnutou při postupimské konferenci . Projekt zpočátku podporoval iWashington . Rusk/rusk/NNMS1-----A----- jen upozornil , že bude nutné přestěhovat obyvatele Západního Berlína |
| <input type="checkbox"/> | Nezbytné věci | . " Sandy Koufax ví houby , " poznamenal Brian Rusk/rusk/NNMS1-----A----- a stiskl palcem nohy spoušť . Seanův pronikavý jek hrůzy |
| <input type="checkbox"/> | Nezbytné věci | Ale právě to udělat nechtěl . Nechtěl , aby Brian Rusk/rusk/NNMS1-----A----- spatřil policejní vůz - chtěl , aby spatřil trochu omláčenou |
| <input type="checkbox"/> | Nezbytné věci | . . ten klíč , který mu teď dal Sean Rusk/rusk/NNMS1-----A----- . A co je za těmi dveřmi ? No co |
| <input type="checkbox"/> | Aha! | si valná většina fanoušků , kteří odcházeli z boxerského klání rusk/rusk/X@----- ruského giganta Nikolaje Valujeva a amerického » dědka « Evan |

Jak je vyřešen problém lemmatizace / taggování

- [lemma="po"][lemma=".*sk[ou]"]

Výskytů: 38 414 | i.p.m.: 7,13 (vztaženo k celému korpusu) | ARF: 22 304,01 | Výsledek je promíchán



1 / 961

Výběr řádků: základní

<input type="checkbox"/>	Blesk magazín	jsem například v Tokiu , Helsinkách , Kodani , různě	po/po/RR--6----- Rusku/rusko/NNNS6----A-----	... To byly velké z
<input type="checkbox"/>	Deníky Bohemia	, jestli by zeleň vydržela . Pro děti vymýšlí pohádky	po/po/RR--6----- hasičsku/hasičsko/NNNS6----A-----	a učí je pomáhat
<input type="checkbox"/>	Mladá fronta DNES	firem . Wartenberg také uvedl , že Česká republika je	po/po/RR--6----- Polsku/polsko/NNNS6----A-----	druhým nejdůlež
<input type="checkbox"/>	Deníky Bohemia	stupně na Labi . Prezidenta doprovázela na jeho minulé cestě	po/po/RR--6----- Děčínsku/děčínsko/NNNS6----A-----	také jeho manžer
<input type="checkbox"/>	Regionální týdeník	po světě a sbíral brouky ? Sjezdil jsem Asii až	po/po/RR--4----- Japonsku/japonsko/NNNS4----A-----	v doprovodu až c
<input type="checkbox"/>	Deníky Moravia	inscenaci Čechovova Platonova , Maďaři nabídnou minimalistického Kupce benátského .	PO/po/RR--6----- SRBSKU/srbsko/NNNS6----A-----	. Snová Loď pro
<input type="checkbox"/>	Mladá fronta DNES	podářilo rozsvítit vánoční strom Strom konečně svítí Druhý pokus vyšel	PO/po/RR--6----- fiasku/fiasco/NNNS6----A-----	, kdy se minulou
<input type="checkbox"/>	Deníky Bohemia	skupina má na starosti vesnice a města od Roztoků až	po/po/RR--4----- Hřensku/hřensko/NNNS4----A-----	. Postupně se sr
<input type="checkbox"/>	Deníky Bohemia	uvedených ve svém seznamu přečtené literatury . Deník dodává :	PO/po/RR--6----- Japonsku/japonsko/NNNS6----A-----	a Číně objevila Č
<input type="checkbox"/>	Moderní byt	- top rezidence Apartmán Studios Rokytnice Apartmán Studios Rokytnice OSLNIVÉ KEMPOVÁNÍ	PO/po/RR--6----- JAPONSKU/japonsko/NNNS6----A-----	Užít si výhod par
<input type="checkbox"/>	Hospodářské noviny	problém zdaleka tak velký . Naopak : Od Finska až	po/po/RR--4----- Portugalsku/portugalsko/NNNS4----A-----	rostou stromy v
<input type="checkbox"/>	Deníky Bohemia	pochutnají na hovězí polévce s masovou rýží , vepřovém plátku	po/po/RR--6----- srbsku/srbsko/NNNS6----A-----	a k večeri dostar
<input type="checkbox"/>	Město zlodějů	, aniž se ohlédla , jestli jdeme za ní .	PO/po/RR--6----- fiasku/fiasco/NNNS6----A-----	s Mgou jsem ne
<input type="checkbox"/>	Deníky Moravia	věřím . Vždyť to štěstí , zdraví a pohodu si	po/po/RR--6----- loňsku/loňsko/NNNS6----A-----	všichni plnými h
<input type="checkbox"/>	Regionální týdeník	tak u Žiliny při střídání stráží během přesunu sovětských vojsk	po/po/RR--6----- velitelsko/velitelsko/A2-----A-----	štábním cvičení
<input type="checkbox"/>	Xantypa	tuhne „ na věčné časy " . Zkusil jsem si	po/po/RR--6----- beefheartovsku/beefheartovsku/X@-----	řvavě notovat př
<input type="checkbox"/>	Deníky Moravia	Foto : Deník / . Jana Rvbová SVAZOVÁNÍ A SFX SVAZOVÁNÍ	PO/po/RR--6----- JAPONSKU/japonsko/NNNS6----A-----	Netradiční výst

Frekvence/Vlastní/slovo nejvíce vpravo/POS

Celkem: 4 položky (1 stránka)

	Filtr	pos [part of speech]	Freq	
1	p / n	N	33025	
2	p / n	X	5221	
3	p / n	D	116	
4	p / n	A	52	

pos=N

Výskytů: 33 025 | i.p.m.: 6,13 (vztaženo k celému korpusu) | ARF: 19 261,56 | Výsledek je promíchán

1 / 826

Výběr řádků: základní

<input type="checkbox"/>	Blesk magazín	jsem například v Tokiu , Helsinkách , Kodani , různě	po/po/RR--6----- Rusku/rusko/NNNS6-----A-----	... To byly velké zážitk
<input type="checkbox"/>	Deníky Bohemia	, jestli by zeleň vydržela . Pro děti vymýšlí pohádky	po/po/RR--6----- hasičsku/hasičsko/NNNS6-----A-----	a učí je pomáhat Rou
<input type="checkbox"/>	Mladá fronta DNES	firem . Wartenberg také uvedl , že Česká republika je	po/po/RR--6----- Polsku/polsko/NNNS6-----A-----	druhým nejdůležitější
<input type="checkbox"/>	Deníky Bohemia	stupně na Labi . Prezidenta doprovázela na jeho minulé cestě	po/po/RR--6----- Děčínsku/děčínsko/NNNS6-----A-----	také jeho manželka Iv
<input type="checkbox"/>	Regionální týdeník	po světě a sbíral brouky ? Sjezdil jsem Asii až	po/po/RR--4----- Japonsku/japonsko/NNNS4-----A-----	v doprovodu až dvaná
<input type="checkbox"/>	Deníky Moravia	inscenaci Čechovova Platonova , Maďaři nabídnou minimalistického Kupce benátského .	PO/po/RR--6----- SRBSKU/srbsko/NNNS6-----A-----	. Snová Loď pro pane
<input type="checkbox"/>	Mladá fronta DNES	podářilo rozsvítit vánoční strom Strom konečně svítí Druhý pokus vyšel	Po/po/RR--6----- fiasku/fiasko/NNNS6-----A-----	, kdy se minulou nedě
<input type="checkbox"/>	Deníky Bohemia	skupina má na starosti vesnice a města od Roztoků až	po/po/RR--4----- Hřensku/hřensko/NNNS4-----A-----	. Postupně se snaží p
<input type="checkbox"/>	Deníky Bohemia	uvedených ve svém seznamu přečtené literatury . Deník dodává :	Po/po/RR--6----- Japonsku/japonsko/NNNS6-----A-----	a Číně objevila Český
<input type="checkbox"/>	Moderní byt	– top rezidence Apartmán Studios Rokytnice Apartmán Studios Rokytnice OSLNIVÉ KEMPOVÁNÍ	PO/po/RR--6----- JAPONSKU/japonsko/NNNS6-----A-----	Užít si výhod panensk
<input type="checkbox"/>	Hospodářské noviny	problém zdaleka tak velký . Naopak : Od Finska až	po/po/RR--4----- Portugalsko/portugalsko/NNNS4-----A-----	rostou stromy v posle
<input type="checkbox"/>	Deníky Bohemia	pochutnají na hovězí polévce s masovou rýží , vepřovém plátku	po/po/RR--6----- srbsku/srbsko/NNNS6-----A-----	a k večeři dostanou d

pos=X

Výskytů: 5 221 | i.p.m.: 0,97 (vztaženo k celému korpusu) | ARF: 2 792,28 | Výsledek je promíchán

1 / 131 ▶▶▶

Výběr řádků: základní ▾

<input type="checkbox"/>	Xantypa	tuhne „ na věčné časy ” . Zkusil jsem si	po/po/RR--6----- beefheartovsku/beefheartovsku/X@-----	řvavě notovat při gymnaziálním s
<input type="checkbox"/>	Hospodářské noviny	elektráren - od Krušných hor přes Liberecko , Hodonínsko až	po/po/NN-----8- Bruntalsko/bruntalsko/X@-----	. Jen v okolí Kuželova na Hodonín
<input type="checkbox"/>	Deníky Bohemia	Live 60 . Lucerna 2006 Svě šedesátiny oslavil Michal Prokop	po/po/RR--6----- muzikantsku/muzikantsku/X@-----	- vyprodaným koncertem v pražsk
<input type="checkbox"/>	Mladá fronta DNES	textu Víťa Marčík , který hru zároveň režíruje . Cervantes	po/po/NN-----8- doudlebsku/doudlebsku/X@-----	Napsal ji na tělo dvěma komediar
<input type="checkbox"/>	Španělská kuchyně	se vzduchové bubliny . CAZA Pichones a la Navarra Holoubata	po/po/NN-----8- navarrsku/navarrsku/X@-----	Holoubata omyjeme pod tekoucí
<input type="checkbox"/>	Právo	nespolehlivého procesu demokratické volby a se vším se vypořádat jaksepatří	po/po/RR--6----- manažersku/manažersku/X@-----	- ostrými lokty a naditými tobolka
<input type="checkbox"/>	Deník kazatele na cestách	, který již mnoho lete bere oddané na poutní cestu	po/po/NN-----8- Vradžsku/vradžsku/X@-----	, se postupně učím zábavy , které
<input type="checkbox"/>	Deníky Bohemia	, jak muži odhazují svou důstojnost a začínají se česat	po/po/NN-----8- žensku/žensku/X@-----	. Dokonce si i vlasy nakrucují želí
<input type="checkbox"/>	Fámy	. " Van Hoof se domníval , že teď mluví	po/po/NN-----8- kněžsku/kněžsku/X@-----	. Tak jako člověk koktá u koktavé
<input type="checkbox"/>	Deníky Bohemia	Poslední listopadovou středu pak v konferenčním sále promítnou Třídní sraz	po/po/NN-----8- estonsku/estonsku/X@-----	. Začátky projekcí jsou vždy v 19
<input type="checkbox"/>	Týden	Suhartovým ostražitým dohledem fungovala tři desítky let . Společenská smlouva	po/po/NN-----8- indoněsku/indoněsku/X@-----	vypadala tak , že občan neusilova
<input type="checkbox"/>	Znamení zrodu	, irské reálie vstupují do českého kontextu (" řečnit	po/po/RR--6----- o'connellsku/o'connellsku/X@-----	" , KVĚTY 1837 , 310 - 311) ,
<input type="checkbox"/>	Téma	kocourů i matky Julie , které na newyorském Manhattanu říkali	po/po/NN-----8- rusínsku/rusínsku/X@-----	prostě Bába (Bubba) . Dům jako
<input type="checkbox"/>	Hospodářské noviny	nenechali nalákat od Obamova ekonomického rádce Larryho Summerse , který	po/po/NN-----8- keynesiánsku/keynesiánsku/X@-----	tvrdí , že každý investovaný státní

POS=D

Výskytů: 116 | i.p.m.: 0,02 (vztaženo k celému korpusu) | ARF: 68,81 | Výsledek je promíchán

1 / 3

Výběr řádků: základní

<input type="checkbox"/>	STYL	pečené Grenaille a tymiánové demi-glacé . Kolegyně volí hovězí plec	po/po/RR--6----- cikánsku/cikánsku/Dg-----1A-----	s paprikami a jasmínovou rýží . Jako správné členky Pr
<input type="checkbox"/>	Právo	tábor vězením pro Cikány , cikánské míšence a osoby žijící	po/po/RR--6----- cikánsku/cikánsku/Dg-----1A-----	. V nelidských podmínkách zde čekala smrt na 326 muž
<input type="checkbox"/>	Regionální týdeník	na červeném víně , houskový knedlík 3 . „ Panenka	po/po/RR--6----- Cikánsku/cikánsku/Dg-----1A-----	" (nudličky z vepřové panenky s lečem cibulí a
<input type="checkbox"/>	Deníky Bohemia	mnoho cenných poznatků . " Vařit umím , ale jenom	po/po/RR--6----- cikánsku/cikánsku/Dg-----1A-----	- halušky , golubky a podobně . Jsem ráda ,
<input type="checkbox"/>	Regionální týdeník	stejně tak i na svahu . Plesat budou letos i	po/po/RR--6----- cikánsku/cikánsku/Dg-----1A-----	Plesová sezóna je opět bohatá . Kulturní dům ve Velkýc
<input type="checkbox"/>	Dynastie Morlandů 5	které nosil dlouhé až k ramenům a které se mu	po/po/NN-----8- kavalírsku/kavalírsku/Db-----	kroutily . Jeho vážný obličej byl dotmava opálený a ošle
<input type="checkbox"/>	Deníky Bohemia	potřeby . Při podávání posypeme nasekanou petrželovou natí . Filé	po/po/RR--6----- cikánsku/cikánsku/Dg-----1A-----	Ingredience : rybí filé , 3 cibule , 4 rajčata
<input type="checkbox"/>	Deníky Moravia	živá hudba a přítomní si mohou pochutnat třeba na kapru	po/po/RR--6----- cikánsku/cikánsku/Dg-----1A-----	, smaženém a uzeném kapru nebo na smaženém pstru
<input type="checkbox"/>	Zemědělec	vězením pro „ Cikány , cikánské míšence a osoby žijící	po/po/RR--6----- cikánsku/cikánsku/Dg-----1A-----	" . V nelidských podmínkách zde čekala smrt na 326
<input type="checkbox"/>	Haló noviny	Smažený kapřík je báječný , ale ani na kmínu nebo	po/po/RR--6----- cikánsku/cikánsku/Dg-----1A-----	se zeleninou a rajským protlakem či zapečený s plátky
<input type="checkbox"/>	Týdeník Květy	a pravá ruka mafiána Grgy , rozhodne všechno vyřešit pěkně	po/po/RR--6----- cikánsku/cikánsku/Dg-----1A-----	: stačí , když Matko svého sedmnáctiletého syna Zareh

pod=A

Výskytů: 52 | i.p.m.: 0,01 (vztaženo k celému korpusu) | ARF: 29,51 | Výsledek je promíchán

1 / 2

Výběr řádků: základní

<input type="checkbox"/>	Regionální týdeník	tak u Žiliny při střídání stráží během přesunu sovětských vojsk	po/po/RR--6----- velitelsko/velitelsko/A2-----A-----	štábním cvičení Šumava . Měl to být poslední sově
<input type="checkbox"/>	Deníky Moravia	„ Po medvědích tlapkách “ nebo se mohou zúčastnit Putování	po/po/RR--6----- moravsko/moravsko/A2-----A-----	- slezském pomezí do Srubu Petra Bezruče . Veče
<input type="checkbox"/>	Víkend HN	budovat nebo podporovat v kdejakém koutě planety , od malajsie	po/po/RR--6----- rusko/rusko/A2-----A-----	. týdeník economist a on-line magazín Quartz k ton
<input type="checkbox"/>	Deníky Bohemia	si tuto událost nevybrali náhodou . " Je velmi zajímavá	po/po/RR--6----- vojensko/vojensko/A2-----A-----	- taktické stránce , jak se Němci nechali sevřít do
<input type="checkbox"/>	Ona DNES	spoustu času . Nejkrásnější a nejdivočejší jsou hřebenovky táhnoucí se	po/po/RR--6----- česko/česko/A2-----A-----	- slovenské hranici , je tu nádherná příroda , krásné
<input type="checkbox"/>	Deníky Moravia	dál ? dnes navštíví Prostějov . V této části odyssey	po/po/RR--6----- česko/česko/A2-----A-----	-- moravsko-slezských domácnostech čeká moder
<input type="checkbox"/>	Deníky Bohemia	kamiony silnici na Harrachov , a to od Tanvaldu až	po/po/RR--4----- česko/česko/A2-----A-----	polskou hranici . Otevřeli ji až po více než sedmi
<input type="checkbox"/>	Deníky Bohemia	z variant má mnoho dalších aspektů , od ekonomických až	po/po/RR--6----- společensko/společensko/A2-----A-----	politické i sociální a je pochopitelné , že se názory

Slovnědruhový přesah a adverbializace

- Slovnědruhový přechod (adverbializace) je v případě vzniku adverbií zejména z předložkových pádů jmenných tvarů komplikovaný.
- V tradičním popisu (IJP) jsou spojení typu „*jízda po italsku*“ traktovány v oddíle věnovaném „přísllovečným“ spřežkám, spřažená grafická podoba je ovšem vyloučena a druhý člen spojení je interpretován jako „zeměpisné jméno psané s malým počátečním písmenem“ (implicitně to znamená, že jinak/mimo spojení s významem „po x-sku způsobu“ je „x-sku“ zeměpisné jméno psané s velkým písmenem).
- Slovnědruhový přesah mají ustrnulé tvary jmen v pozici tzv. predikativ.

Všimněte si **N/D** rozdílů v následujících příkladech a uvědomte si, kde je slovnědruhový přesah (**nelze desambiguovat**)

- *Je mu **teplo**. Je mu **zima**.*
- *Bylo mu **teplo**. Byla mu **zima**.*
- *Bylo mu **teplo**. Bylo mu **zima**.*
- *Bylo mu velké **teplo**. Byla mu velká **zima**.*
- *Bylo mu **velmi** **teplo**. Bylo mu **velmi** **zima**.*

Jak je to se slovnědruhovou přesajžností případů typu *po . *sku*

- Vidíme, že automatická morfologická analýza je v řešení slovnědruhovú interpretace nekonzistentní.
- Nesleduje bezvýhradně linii vytčenu v IJP (viz případy jako *po [Cc]ikánsku* a *po kavalírsku*).
- Pod **pos=N** tak zahrnuje jak doklady „*jízda po italsku*“, tak doklady typu „*jízda po Španělsku*“ a dokonce i doklady jako „*mluvit po rusínsku*“.
- Mnoho dokladů (většinou jde o první případ) zůstává nerozpoznáno automatickou morfologickou analýzou.

Jak fungují vlastní jména ve slovníku *MorfFlex*?

- Propria jsou ve slovníku *MorfFlex* lemmatizována s velkým počátečním písmenem a ta, která jsou takto lemmatizována lze vyhledat přes lemma (lemma=[[:upper:]].*).
- Pokud je ovšem ve slovníku pouze lemma s velkým počátečním písmenem, pak tvar s malým počátečním písmenem není rozpoznán automatickou morfologickou analýzou.

lemma=po word=[Šš]panělsku|[Pp]olsku

Korpus: [syn v8](#) | Dotaz: [po, \[Šš\]panělsku|\[Pp\]olsku](#) (1 773 výskytů) ▶ Promíchat: ✓

Frekvenční seznam

1 / 1

Minimální frekvence:

[Použít](#)

Celkem: 5 položek (1 stránka)

	Filtr	word	pos [part of speech]	Freq	
1	p / n	Španělsku	N	816	
2	p / n	Polsku	N	721	
3	p / n	španělsku	N	125	
4	p / n	polsku	A	108	
5	p / n	polsku	N	3	

lemma=po
word=[Bb]avorsku|[Ee]stonsku

Korpus: [syn v8](#) | Dotaz: [po, \[Bb\]avorsku|\[Ee\]stonsku](#) (236 výskytů) ▶ Promíchat: ✓

Frekvenční seznam

1 / 1

Minimální frekvence: [Použít](#)

Celkem: 4 položky (1 stránka)

	Filtr	word	pos [part of speech]	Freq	
1	p / n	Estonsku	N	120	
2	p / n	Bavorsku	N	78	
3	p / n	bavorsku	X	23	
4	p / n	estonsku	X	15	

lemma=po word=[Nn]ěmecku|[li]talsku

Korpus: [syn v8](#) | Dotaz: [po, \[Nn\]ěmecku|\[li\]talsku](#) (5 050 výskytů) ▶ Promíchat: ✓

Frekvenční seznam

1 / 1

Minimální frekvence: [Použít](#)

Celkem: 5 položek (1 stránka)

	Filtr	word	pos [part of speech]	Freq	
1	p / n	Německu	N	2841	
2	p / n	italsku	A	1899	
3	p / n	německu	X	261	
4	p / n	italsku	N	27	
5	p / n	Italsku	N	22	

Problém slovníku *MorfFlex*

- Nedostatek pokrytí tvarů „zeměpisných jmen“, které mohou mít obě pravopisné varianty počátečního písmena.
- Obě varianty má ve slovníku *MorfFlexu* *Španělsko* i *španělsko*, *Polsko* i *polsko*, *Německo* i *německo*, ale i *Italsko* i *italsko*.
- Variantu jedinou má např. *Bavorsko* a *Estonsko*.

Homonymie a kolokace

- Pravopisně rozlišené varianty jsou pouze homofonní (nikoli homografní), takže jejich desambiguace by neměla v případě doplnění slovníku o chybějící položky činit potíže.
- Podívejte se na data a sledujte, zda je pravopisná (grafická) norma vždy dodržována.

po Španělsku

Korpus: [syn v8](#) | Dotaz: [po, \[šš\]panělsku\[pp\]olsku](#) (1 773 výskytů) ▶ Promíchat: ▶ Pozitivní filtr: [Španělsku](#) (816 výskytů) ▶ Pozitivní filtr: [N](#) (816 výskytů)

Výskytů: **816** | i.p.m.: **0,15** (vztaženo k celému korpusu) | ARF: **464,88** | Výsledek je promíchán

1 / 21 ▶▶▶

Výběr řádků:

<input type="checkbox"/>	Deníky Bohemia	a to jsme už měli za sebou sedm tisíc kilometrů	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	, tedy pět set kilometrů průměrně na jeden den ,
<input type="checkbox"/>	Právo	chystali se zabíjet pro svou pravdu pod stejnou barvou .	Po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	a Velké Británii se tedy i Německo probudilo do reality
<input type="checkbox"/>	Mladá fronta DNES	festival v Norsku , turné po dalších zemích Skandinávie či	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	. K paradoxům její kariéry patří nevyzpytatelnost . Věra
<input type="checkbox"/>	Lidové noviny	turisté České Budějovice (dan) - Chorvatsko je letos	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	pro většinu jihočeských cestovních kanceláří nejprodá
<input type="checkbox"/>	Deníky Moravia	konferenci v Buenos Aires , když se vrátil z cesty	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	a Kubě . " Z těžké situace mi pomohly mé
<input type="checkbox"/>	Mladá fronta DNES	. Velkým oceněním pro Plzeňany bylo pozvání k pokračování turné	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	a k účasti na festivalu v Zaragoze . " S
<input type="checkbox"/>	Bulvár	vyrazil poznávat krásy Pyrenejského poloostrova , když s přáteli cestoval	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	, užil si i dvoudenní natáčení v Tunisku . Z
<input type="checkbox"/>	Regionální týdeník	byl společně s houslistou Pavlem Šporcem pozván na měsíční turné	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	jako sólista Pardubické komorní filharmonie . V roce 2
<input type="checkbox"/>	Mladá fronta DNES	na italském pobřeží , český průvodce bude bez omezení provádět	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	Kdo dostane šanci služby cestovních kanceláří , v turis
<input type="checkbox"/>	ZEN magazín	státu a korporací jakžtakž odolává , ČR je ale hned	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	vůči GMO nejvstřícnější . * NA GLOBÁLNÍ POCHOD PR
<input type="checkbox"/>	Blesk pro ženy	První krůčky Kariéru v cestovním ruchu zahájila Kristina jako průvodkyně	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	. Později dostala zajímavou nabídku pracovat v cestov
<input type="checkbox"/>	Lidové noviny	za 136 eur , po Nizozemsku za 40 eur ,	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	za 87 eur . Kurs najdete také na internetové adrese
<input type="checkbox"/>	Hospodářské noviny	kde by se referendum o ústavě mělo konat několik měsíců	po/po/RR--6-----	Španělsku/španělsko/NNNS6-----A-----	. " Musíme lidem ale sdělit , jaký význam ústava

Kolokace (T-score) (-1,-1 lemma)

1 ▶

	Filtr	lemma_lc	Freq	MI	T-score ▼	logDice
1.	p/n	turné	171	12.826	13.075	5.163
2.	p/n	cesta	68	7.460	8.199	-0.196
3.	p/n	cestovat	19	9.273	4.352	1.612
4.	p/n	zájezd	11	8.905	3.310	1.242
5.	p/n	hned	10	5.349	3.085	-2.307
6.	p/n	pout'	9	8.532	2.992	0.869
7.	p/n	putování	8	9.605	2.825	1.932
8.	p/n	průvodce	8	8.485	2.821	0.822
9.	p/n	jezdit	8	5.581	2.769	-2.075
10.	p/n	cestování	7	8.487	2.638	0.823
11.	p/n	-	14	1.700	2.590	-5.956
12.	p/n	.	68	0.517	2.484	-7.138
13.	p/n	kilometr	5	5.261	2.178	-2.396
14.	p/n	také	7	2.389	2.141	-5.267
15.	p/n	tak	8	2.030	2.136	-5.625
16.	p/n	nástupnictví	4	12.579	2.000	4.674

Všimněte si chyb proti kodifikaci /pravopisu

Korpus: [syn v8](#) | Dotaz: [po, \[Šš\]panělsku\[PP\]olsku](#) (1 773 výskytů) ▶ Promíchat: ▶ Pozitivní filtr: [Španělsku](#) (816 výskytů) ▶ Pozitivní filtr: [N](#) (816 výskytů) ▶ Pozitivní filtr (bez KWIC): [lanýž](#) (1 výskyt)

Výskytů: 1 | i.p.m.: 0 (vztaženo k celému korpusu) | ARF: 1 | Výsledek je promíchán

1 / 1

Výběr řádků:

[Mladá fronta DNES](#) střídalý univerzální talíře , telecí játra dušený králík , *lanýže* **po/po/RR--6-----Španělsku/španělsko/NNNS6-----A-----** slepice na paprice . " Konec tisíciletí vybízí k bilancování

1 / 1

po Italsku

Výskytů: 22 | i.p.m.: 0 (vztaženo k celému korpusu) | ARF: 10,57 | Výsledek je promíchán

1 / 1

Výběr řádků: základní

<input type="checkbox"/>	Regionální týdeník	pečeme asi 8 minut z každé strany . Grilovaná zelenina	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	Ingredience : • lilek 1 ks • cuketa 2 ks
<input type="checkbox"/>	Aha! neděle	znát ze seriálu » Smallville « , filmu » Loupež	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	« nebo » Škola svádění « , nahota není žádný
<input type="checkbox"/>	Hospodářské noviny	. A hlavně pochod začal o den dřív . Manželství	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	: už ne hádky . Rozvod Itálie Fotografie italské rodiny
<input type="checkbox"/>	DVDMAG	Prosím Vás , vyjde někdy Jízda do pekel , Loupež	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	, Den Poté , Pach krve , Telefonní budka ,
<input type="checkbox"/>	Mladá fronta DNES	Čechů v Itálii Nástrahy života v Itálii Fotoreportáž : bydlení	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	Fenomén Nedvěď ZLATÝ POKLAD . Větší individuální vyzr
<input type="checkbox"/>	Deníky Bohemia	školy v Národní ulici v Prachaticích Lenka Růžičková . Kuře	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	Ingredience na omáčku : 2 dl vody , půl hrnku
<input type="checkbox"/>	Deníky Moravia	jedlíci utkají v po jídání špaget v soutěži „ Valašská žranice	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	“ . Kulturní program na karolinském náměstí před budov
<input type="checkbox"/>	Dieta	499Kč) , nebo dáte přednost klasice Jamie Oliver vaří	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	(840Kč) ? Nebo víte co ? Řekněte si
<input type="checkbox"/>	Deníky Bohemia	v Čelákovcích . Manželství po italsku v Hálku Nymburk/ Manželství	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	. To je název komedie , která bude k vidění
<input type="checkbox"/>	Deníky Bohemia	, ale už dále nesolíme . Vhodné přílohy ke kuřeti	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	jsou rýže , brambory , ale třeba i hranolky .
<input type="checkbox"/>	Lidové noviny	red. Scolova Noc ve V areennes . Kůže , Manželství	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	i Bílé noci . Boormanův film Leo the Last .
<input type="checkbox"/>	Deníky Bohemia	dětské pacienty . Manželství po italsku v Hálku Nymburk/ Manželství	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	. To je název komedie , která bude k vidění
<input type="checkbox"/>	Deníky Bohemia	úřadu Marta Mastná . REGION SVITAVY KINA Jevíčko : Loupež	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	(začátek dnes ve 20 a zítra v 17.30 a
<input type="checkbox"/>	Blesk magazín	žádné zelí a sádlo , " směje se . Vánoce	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	Štědrý den je v Itálii 25 . prosince a na
<input type="checkbox"/>	Deníky Bohemia	Casanovské slavnosti tak mohly začít . Premiéru hudební komedie Výchova	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	ze života Casanovy , autoři Milan a Marie Petrovští napsa
<input type="checkbox"/>	Sedmíčka	možná nejslavnějšího filmu Horalka nebo z typické italské komedie Manželství	po/po/RR--6----- Italsku/italsko/NNNS6-----A-----	. Zároveň je ale dámou se vším všudy a kam

Kolokace (T-score) (-1,-1 word)

Kolokace

1

	Filtr	lc	Freq	MI	T-score ▼	logDice
1.	p/n	manželství	5	12.904	2.236	0.035
2.	p/n	loupež	3	14.653	1.732	1.784
3.	p/n	výchova	2	13.622	1.414	0.753
4.	p/n	žranica	1	25.869	1.000	10.300
5.	p/n	žebírko	1	19.318	1.000	6.367
6.	p/n	kuřeti	1	17.108	1.000	4.221
7.	p/n	rizoto	1	15.668	1.000	2.793
8.	p/n	kuře	1	12.932	1.000	0.062
9.	p/n	zelenina	1	12.546	1.000	-0.323
10.	p/n	výbuch	1	12.210	1.000	-0.659
11.	p/n	vaří	1	12.065	1.000	-0.804
12.	p/n	láska	1	10.559	0.999	-2.310
13.	p/n	vánoce	1	10.297	0.999	-2.572
14.	p/n	bydlení	1	9.610	0.999	-3.259

Pozorování

- Dodržení pravopisné normy v psaných textech není zcela důsledné.
- Oporou pro vyhledání chyb v pravopisu mohou být kolokační profily.

Závěr

- Spojení typu *po . *sku*
- Popis v příručce a praxe lemmatizace
- Otázka uchopení slovnědruhového přesahu
- Nedostatky slovníku *MorfFlex* jako zdroj nerozpoznaných spojení typu *po . *sku* psaných podle platné kodifikace
- Přegenerovaný slovník a porušení pravopisné normy

Domácí úkol na 4. 5. 2022

- Popište postup, jak lze hledat v korpusu SYNv9 doklady případů, kdy tvar na *sku*) následuje po předložce *po* a obsahuje gramatickou chybu, a jak postupovat v opačných případech.