

Deriváty od číselnek určitých

PLIN033

Slovnědruhová povaha číslovek

- Sémantické kritérium
- Slovnědruhové přesahy
- Adjektivní/substantivní flexe
- Adverbiální povaha číslovek násobných
- Velmi pravidelná derivace (**potencialita neomezená**)

číslovkový vzor

- +pades
- <át,A> 4A,4A1,6KRAT
- <át,B> PRT1,PRT2
- <ater,C> PRT1,VT
- <ateř,C> PRT2
- <atero,C> 6NASOB
- <átk,K> V7
- <átek,K> V78X
- <átc,K> V7A
- <átin,L> V7,V7B,V78X
- <átic,N> V8,V8A,V78X

Značkování SD číslovka

C. *

Vložit Reset

1) Slovní druh [12]

2) Detailní určení slovního druhu [16]

- a - číslovka neurčitá ("mnoho", "málo", "tolik", "několik", "kdovíkolik", ...)
- d - číslovka druhová, adjektivní skloňování ("jedny", "dvoji", "desaterý", ...)
- h - číslovka druhová "jedny" a "nejedny"
- k - číslovka druhová ≥ 4 , adjektivní postavení, krátký tvar ("čtvery", ...)
- j - číslovka druhová ≥ 4 , substantivní postavení ("čtvero", "desatero", ...)
- l - číslovka základní 1-4, "půl", ...; "sto" a "tisíc" v nesubstantivním skloňování
- n - číslovky základní ≥ 5
- w - číslovka neurčitá s adjektivním skloňováním ("nejeden", "tolikátý", "několikátý"...)
- y - zlomek zakončený na "-ina" (značkováno jako číslovka)
- ? - číslovka "kolik"
- = - číslo psané číslicemi (značkováno jako číslovka)
- o - číslovka násobná neurčitá ("-krát": "mnohokrát", "tolikrát", ...)
- r - číslovka řadová
- u - číslovka tázací násobná "kolikrát"
- v - číslovka násobná ("-krát": "pětkrát", "poprvé"...)
- z - číslovka tázací řadová "kolikátý"

V korpusu sn2020 je nové členění číslovek

Nové členění číslovek

Číslovky se nově člení podle tří os: a) základní sémantiky, b) určitosti/neurčitosti, c) typu flexe.

Tento systém drobně porušují pouze kategorie I (kde je kromě určitých číslovek zahrnuta i číslovka neurčitá *nejeden*), kategorie z (kde jsou zahrnuty i číslovky *sto* a *tisíc* s dvojným způsobem skloňování) a kategorie y (kde se číslovky *půl*, *čtvrt* a *polovic* ve skloňování částečně odlišují od ostatních číslovek dílových).

značka	nová funkce	příklady	původní funkce značky
I	číslovka základní 1–4 + <i>nejeden</i>	„jeden“, „dva“, „oba“, „tři“, „čtyři“, „nejeden“	číslovky základní 1–4, „půl“ ...; „sto“ a „tisíc“ v nesubstantivním skloňování
n	číslovky základní 5–99 + jejich slitky i s dalšími číslovkami	„pět“, „šest“, „dvacet“, „pětaosmdesát“ ...; „dvacetpět“, „pětset“, „stořícet“, „pětapůl“ ...	číslovky základní ≥ 5
z	číslovka základní se substantivním skloňováním	„sto“, „tisíc“, „million“, „bilarda“ ...; „nula“, „raz“	číslovka tázací řadová „kolikátý“
a	číslovka základní neurčitá a tázací	„mnoho“, „tolik“, „několik“, „kdovíkolik“, „kolik“	číslovka neurčitá („mnoho“, „málo“, „tolik“, „několik“, „kdovíkolik“ ...)
y	číslovka dílová	„půl“, „polovic“, „čtvrt“, „polovina“, „třetina“, „osmina“ ...	zlomky zakončené na „-ina“ (značkováno jako slovní druh: číslovka – ‘C’)
r	číslovka řadová	„první“, „druhý“, „třetí“, „pátý“	číslovky řadové
w	číslovka řadová neurčitá a tázací	„tolikátý“, „několikátý“, „kolikátý“	číslovky neurčité s adjektivním skloňováním („nejeden“, „tolikátý“, „několikátý“ ...)
d	číslovka druhová a souborová	„dvojiť“, „obojiť“, „čtverý“, „desaterý“ včetně tvarů „dvoje“, „oboje“, „čtverý“, „desaterý“; „obě“, „jedny“	číslovka druhová, adjektivní skloňování („jedny“, „dvojiť“, „desaterý“ ...)
h	číslovka druhová a souborová neurčitá a tázací	„několikerý“, „tolikerý“, „kolikerý“ včetně tvarů „několikery“, „tolikery“, „kolikery“; „nejedny“	číslovky druhové „jedny“ a „nejedny“
j	číslovka úhrnná	„čtvero“, „patero“, „devatero“, „dvě“, „tře“	číslovka druhová ≥ 4 , substantivní postavení („čtvero“, „desatero“ ...)
k	číslovka úhrnná neurčitá a tázací	„několikero“, „tolikero“, „kolikero“	číslovka druhová ≥ 4 , adjektivní postavení, krátký tvar („čtverý“ ...)
u	číslovka násobná (adjektivní typ)	„dvojitý“, „trojitý“, „dvojnásobný“, „osminásobný“, „desateronásobný“ ...	číslovka tázací násobná „kolikrát“
3	číslovka násobná neurčitá a tázací (adjektivní typ)	„mnohonásobný“, „xnásobný“, „kolikanásobný“ ...	zkratka jako číslovka
v	číslovka násobná (adverbiální typ, včetně spřežek)	„pětkrát“, „osminásobně“, „trojnásob“, „jednou“, „raz“, „dvojitě“, „čtvermo“, „nadvakrát“, „poprvé“, „začtvrté“, „napodesáté“	číslovky násobné („-krát“, „pětkrát“, „poprvé“ ...)
o	číslovka násobná neurčitá a tázací (adverbiální typ, včetně spřežek)	„mnohokrát“, „několikanásobně“, „xkrát“, „bůhvíkolikrát“, „nejednou“, „naněkolikrát“... „pokolikáté“, „naponěkolikáté“	číslovky násobné neurčité („-krát“, „mnohokrát“, „tolikrát“ ...)
=	číslo psané arabskými číslicemi	„12“, „42“, „1,5“	číslo psané číslicemi (značkováno jako slovní druh: číslovka – ‘C’)
}	číslo psané římskými číslicemi	„VII“, „XXVI“	číslovka psaná římskými číslicemi

Slova s možnou číselnou platností

- MČ 2 s. 101n. tab. na s. 113
- Číslovky velikostní: adjektiva na *-ový* (*tisícový*)
- Číslovky skupinové: substantiva na *-ice, -ka* (*rojicet, čtyřka*)

Značkování v korpusu

- .*ový

28.	p/ n	pohybový	A	3 155
29.	p/ n	datový	A	3 062
30.	p/ n	bytový	A	3 051
31.	p/ n	nulový	A	3 022
32.	p/ n	špičkový	A	2 947
33.	p/ n	výběrový	A	2 922

Jak vypadají a jak jsou značkována adjektiva s číselným významem

- dvojkový, trojkový, ...

Mobility	, který je také sedmipalcový a také jej pohání pouze	dvojkový/dvojkový/AAIS1----1A-----	Android
Deterministický chaos a jeh...	mají asymptotickou relativní frekvenci jedniček i nul ve svém	dvojkovém/dvojkový/AAIS6----1A-----	rozvoji x
Velké otázky	a guaninu – ale tato různá kódování lze vyjádřit v	dvojkové/dvojkový/AAFS6----1A-----	soustavě
Velké otázky	, což je zkratka pro „binary digit“ (dvojková/dvojkový/AAFS1----1A-----	číslice)
Filozofické příběhy	možná býval úspěšnější, kdyby se dále věnoval studiu	dvojkové/dvojkový/AAFS2----1A-----	soustavy

Jak formulovat složitější dotaz

- [lemma="(jed((not) | (nič) | (enáct))k)ový"]

Celkem: 3 (1 str.)

	Filtr	lemma_lc [lowercase lemma]	pos [part of speech]	Frekvence
1.	p/n	jednotkový	A	458
2.	p/n	jedničkový	A	20
3.	p/n	jedenáctkový	A	1

Podobně (na hnízdo derivátů jména každého čísla od 1 -10 / ?0)

- [lemma="(jed((not)|(nič)|(enáct))k)ový"]
- [lemma="(((dvoj)|dva((náct)|(cít)))k)ový"]
- [lemma="(dv((oj)|(anáct)|(acít))k)ový"]
- [lemma="(((troj)|tři((náct)|(cít)))k)ový"]
- [lemma="(t((roj)|(řináct)|(řicít))k)ový"]
- [lemma="((([čš]ty[rř])|([čš]trnáct)|([čš]ty[rř][iy]cít))k)ový"]
- [lemma="([čš]t((y[rř])|(rnáct)|(y[rř][iy]cít))k)ový"]
- [lemma="(((pět)|(patnáct)|(padesát))k)ový"]
- [lemma="(p((ět)|(atnáct)|(adesát))k)ový"]
- [lemma="(((šest)|(šestnáct)|(šedesát))k)ový"]
- [lemma="(še((st)|(stnáct)|(desát))k)ový"]
- [lemma="(((sedmič)|(sedmnáct)|(sedmdesát))k)ový"]
- [lemma="(((osmič)|(osmnáct)|(osmdesát))k)ový"]
- [lemma="(((sedm)|(osm))((ič)|(náct)|(desát))k)ový"]
- [lemma="(((devít)|(devatenáct)|(devadesát))k)ový"]
- [lemma="(dev((ít)|(atenáct)|(adesát))k)ový"]
- [lemma="(((desít)|(stov)|(tisícov))k)|(tisíc)|(nul)|(mili[oó]n)ový"]

Ještě složitější dotaz - lze spojit do jednoho dotazu

```
[lemma="((((((((((jed((not) | (nič) | (enáct)) | dv((oj) | (anáct) | (acít)) | t((roj) | (řináct) | (řicít)) | [čš]t((y[rř]) | (rnáct) | (y[rř][iy]cít)) | p((ět) | (atnáct) | (adesát)) | še((st) | (stnáct) | (desát)) | ((sedm) | (osm))((ič) | (náct) | (desát)) | dev((ít) | (atenáct) | (adesát)) | ((desít) | (stov) | (tisícov))))))))))k)ový"]
```

A všechno dohromady (s 0 a substantivními číslovkami - sto, tisíc, milion, ...)

- [lemma="((((((((((jed((not) | (nič) | (enáct)) | dv((oj) | (anáct) | (acít)) | t((roj) | (řináct) | (řicít)) | [čš]t((y[rř]) | (rnáct) | (y[rř][iy]cít)) | p((ět) | (atnáct) | (adesát)) | še((st) | (stnáct) | (desát)) | ((sedm) | (osm))((ič) | (náct) | (desát)) | dev((ít) | (atenáct) | (adesát)) | ((desít) | (stov) | (tisícov))))))))))k) | (tisíc) | (nul) | (mili[oó]n))ový"]

Seznam lemmat a značek (synv12)

1	p / n	nulový	A	89 148	14,29
2	p / n	milionový	A	60 395	9,68
3	p / n	trojkový	A	8 263	1,33
4	p / n	tisícový	A	7 612	1,22
5	p / n	jednotkový	A	4 572	0,73
6	p / n	šestkový	A	2 753	0,44
7	p / n	osmičkový	A	2 656	0,43
8	p / n	sedmičkový	A	2 125	0,34
9	p / n	dvojkový	A	1 756	0,28
10	p / n	jedničkový	A	1 476	0,24
11	p / n	stovkový	A	937	0,15
12	p / n	pětkový	A	934	0,15
13	p / n	desítkový	A	815	0,13
14	p / n	devadesátkový	A	783	0,13
15	p / n	osmdesátkový	A	732	0,12

rozšíříme (lemma nerozpoznaného tvaru je tvar sám,
takže nahradíme koncovku ý reg. výrazem .*)

- [lemma="((((((((((jed((not) | (nič) | (enáct)) | dv((oj) | (anáct) | (acít)) | t((roj) | (řináct) | (řicít)) | [čš]t((y[rř]) | (rnáct) | (y[rř][iy]cít)) | p((ět) | (atnáct) | (adesát)) | še((st) | (stnáct) | (desát)) | ((sedm) | (osm))((ič) | (náct) | (desát)) | dev((ít) | (atenáct) | (adesát)) | ((desít) | (stov) | (tisícov)))))))))k) | (tisíc) | (nul) | (mili[oó]n))ov.*"]

Zpřesníme (značka A - adjektiva, značka X - nerozpoznaná slova)

- [lemma="((((((((((jed((not) | (nič) | (enáct)) | dv((oj) | (anáct) | (acít)) | t((roj) | (řináct) | (řicít)) | [čš]t((y[rř]) | (rnáct) | (y[rř][iy]cít)) | p((ět) | (atnáct) | (adesát)) | še((st) | (stnáct) | (desát)) | ((sedm) | (osm))((ič) | (náct) | (desát)) | dev((ít) | (atenáct) | (adesát)) | ((desít) | (stov) | (tisícov)))))))))k) | (tisíc) | (nul) | (mili[oó]n))ov.*" & tag="[AX].*"]

Podívejte se na významy nerozpoznaných derivátů


1	p / n	šedesátkování	18	0
2	p / n	osmdesátkově	16	0
3	p / n	tisícov	14	0
4	p / n	milionovu	9	0
5	p / n	miliónovu	6	0
6	p / n	šestkovaná	6	0
7	p / n	miliónov	5	0
8	p / n	milionovém městě	3	0
9	p / n	osmičkovitě	3	0
10	p / n	milionov	3	0
11	p / n	nulovu	3	0
12	p / n	pětkování	3	0
13	p / n	nulovku	3	0
14	p / n	šestkované	3	0
15	p / n	tisícovskou	3	0

Podívejte se na výskyt a významy derivátů na ován.*

- [lemma="((((((((((jed((not)|(nič)|(enáct))|dv((oj)|(anáct)|(acít))|t((roj)|(řináct)|(řicít))|[čš]t((y[rř])|(rnáct)|(y[rř][iy]cít))|p((ět)|(atnáct)|(adesát))|še((st)|(stnáct)|(desát))|((sedm)|(osm))((ič)|(náct)|(desát))|dev((ít)|(atenáct)|(adesát))|((desít)|(stov)|(tisícov))))))))))k)|(tisíc)|(nul)|(mili[oó]n))ován.*" & tag="[NAVX].*"]

Výsledky

1 / 1 (celkem: 9 položek)

Sdílet tabulku: 

	Filtr	<u>lemma_lc [lowercase lemma]</u>	<u>pos</u>	Freq ▼	i.p.m.
1	p / n	nulování	N	254	0,04
2	p / n	stovkování	N	73	0,01
3	p / n	šedesátkování	X	18	0
4	p / n	pětkování	X	3	0
5	p / n	osmičkování	N	2	0
6	p / n	šedesátkováním	X	1	0
7	p / n	jedničkovány	X	1	0
8	p / n	jedničkování	X	1	0
9	p / n	jedničkován	X	1	0

O co jde?

- Druhá , nazvaná Ostravská věž , pak následující den před ostravskou Novou radnicí . Velký zájem o sprint Na porubské Hlavní třídě se v sobotu uskutečnil pátý ročník soutěže nazvané Hasičské stovkování . Jejím hlavním cílem bylo přiblížit požární sport běžným lidem . Dopoledne závodili na šedesát metrů dlouhé překážkové dráze žáci , odpoledne pak na stometrové trati muži a ženy . " Závodníci po odstartování uchopili požární hadice , běželi s nimi přes bariéry a po kladině , před koncem trati

O co jde?

Starořecký výraz pro počítání by se dal přeložit jako **pětkování**/pětkování/X@----- a v češtině dodnes slyšíme souvislost mezi slovy „pět“ a „pěst“.

ně běžná praxe : „ Píšu si k sobě malé pětky za každý neodevzdaný úkol a potom za tři neodevzdané zápisy píšu jednu plnohodnotnou pětku do systému . “ „ Jsem rád , že to mohu dělat tak , jak považuji já za správné . “ „ Také nejsem zastáncem **pětkování**/pětkování/X@----- , ale po opakované výzvě k odevzdání a naprosté ignoraci mi jiná možnost nezbyla . “ „ Pokud nedám pětku za nedonesený úkol , tak vám garantuju , že mi těch úkolů odevzdají méně než polovinu . “

Stáří Řekové však využívali pětkovou soustavu a jejich výraz pro počítání by se dal přeložit jako „ **pětkování**/pětkování/X@----- “ , přičemž v češtině dodnes slyšíme souvislost mezi slovy „pět“ a „pěst“.

zpřesníme (přidáním reg. výr. .* na začátek vyhledáme potenciální kompozita)

- [lemma=".*((((((((((jed((not) | (nič) | (enáct)) | dv
((oj) | (anáct) | (acít)) | t((roj) | (řináct) | (řicít)) | [čš]
t((y[rř]) | (rnáct) | (y[rř][iy]cít)) | p((ět) | (atnáct) | (adesát)) | še((st) | (stnáct) | (desát)) | ((sedm) | (os
m))((ič) | (náct) | (desát)) | dev((ít) | (atenáct) | (adesát)) | ((desít) | (stov) | (tisícov)))))))))k) | (tisíc) |
(nul) | (mili[oó]n))ov.*" & tag="[AX].*"]

výsledky

1	p / n	Usnulová	48	0,01
2	p / n	subjednotkové	38	0,01
3	p / n	šedesátkování	18	0
4	p / n	jednomilionovou	16	0
5	p / n	osmdesátkově	16	0
6	p / n	Usnulovi	15	0
7	p / n	tisícov	14	0
8	p / n	jednomilionové	12	0
9	p / n	dvoujednotkového	11	0
10	p / n	jednadvacetimilionovou	11	0

Jak jsou značkovány deadjektivní adverbia?

- [lemma=".*((((((((((jed((not)|(nič)|(enáct))|dv((oj)|(anáct)|(acít))|t((roj)|(řináct)|(řicít))|[čš]t((y[rř])|(rnáct)|(y[rř][iy]cít))|p((ět)|(atnáct)|(adesát))|še((st)|(stnáct)|(desát))|((sedm)|(osm))((ič)|(náct)|(desát))|dev((ít)|(atenáct)|(adesát))|((desít)|(stov)|(tisícov))))))))))k)|(tisíc)|(nul)|(mili[oó]n))ově" & tag="[DX].*"]

Výsledky

	Filtr	<u>lemma</u>	Freq ▼	i.p.m.
1	p / n	osmdesátkově	16	0
2	p / n	Usnulově	2	0
3	p / n	pěrestrojkově	1	0
4	p / n	čtrnáctitisícově	1	0
5	p / n	poddvacítkově	1	0
6	p / n	Hannulově	1	0

Vyhledávání adjektivních derivátů a kompozit s prvním členem číslovkovým

- [lemma="((jedn) | (jed[ie]n) | (dv[aoě]) | (troj) | (tř[íi]) | (čtve[rř]) | (čty[rř]) | (pět) | (šest) | (sedm) | (osm[aáeind]) | (de[sv][áíaeě]t) | (čtrnáct) | (patnáct) | (pater) | (((pa) | (še) | (deva))des[áa]t) | (sto)).*" & tag="A.*"]

zjistíme přegenerovávaná lemmata (náhodná shoda řetězce s číslovkovým řetězcem)

Celkem: 7621 (153 str.)

	Filtr	lemma	Frekvence
1.	p/n	jediný	1 483 457
2.	p/n	jednotlivý	764 703
3.	p/n	jednoduchý	537 819
4.	p/n	čtvereční	135 824
5.	p/n	jedinečný	123 812
6.	p/n	jednotný	122 858
7.	p/n	stolní	120 815
8.	p/n	jednoznačný	117 962
9.	p/n	dvojnásobný	114 525
10.	p/n	dvouletý	104 878
11.	p/n	tříletý	91 128
12.	p/n	dvacetiletý	83 315
13.	p/n	stoprocentní	83 054
14.	p/n	pětiletý	74 161
15.	p/n	třicetiletý	73 136
16.	p/n	stojící	69 174
17.	p/n	třinecký	66 228
18.	p/n	desetiletý	63 190

Lze upravit (vyloučení náhodných shod)

- [lemma="(jedm)|jed[ie]n|dv[aoě]|troj|(tř[íi])|(čtve[rř])|(čty[rř])|(pět)|(šest)|(sedm)|(osm[aáeind])|(de[sv][áíaeě]t)|(čtrnáct)|(patnáct)|(pater)|(((pa)|(še)|(deva))des[áa]t)|(sto)).*" &
lemma!="(tříd[ií]c)|(tříd[n])|(třídě[n])|(třineck)|(tříšt)|(třísk)|(třísel)|(třímaj)|(tříb[iíe][cn])|(třísovsk)|(tro[upbmflc])|(tro[cš]k)|(trojsk)|(tro[is])).*" &
lemma!="(sto[lr]n)|(stočen)|(stockh)|(stoup)|(stopý)|(stomato)|(stopnut)|(sto[ph]ov)|(stopk[ao][vt])|(stop[eé]r)|(stopař)|(stožár)|(stolov)|(stoluj)|(stonkov)|(stodol)|(stoli[čc])|(stolař)|(stonav)|(stoj[aiínk])|(stoi)|(stodsk)|(stochovsk)|(stoneovsk)).*" &
lemma!="(dvo[rř])|(jedna[cn])|(jednajíc)|(jedený)|(osmán)|(osma[hžn])|(osmirk)|(pater[nů])).*" & tag="A.*"]

Lze použít k detekci jednotek nerozpoznaných automatickou morfologickou analýzou

- [lemma="(jedm)|jed[ie]n)|(dv[aoě])|(troj)|(tř[íi])|(čtve[rř])|(čty[rř])|(pět)|(šest)|(sedm)|(osm[aáeind])|(de[sv][áíaeě]t)|(čtrnáct)|(patnáct)|(pater)|(((pa)|(še)|(deva))des[áa]t)|(sto)).*([íáé])|([éi]ho)|([éi]mu)|([éyí]m)|([ýi]ch)|([ýi]m)|([ýi]mi))" & lemma!="(tříd[ií]c)|(tříd[n])|(třídě[n])|(třízen)|(třineck)|(tříšt)|(třísk)|(třísel)|(třímaj)|(tříb[iíe][cn])|(třísovsk)|(tro[upbmf]lc)|(tro[cš]k)|(trojsk)|(tro[is])).*([íáé])|([éi]ho)|([éi]mu)|([éyí]m)|([ýi]ch)|([ýi]m)|([ýi]mi))" & lemma!="(sto[lr]n)|(stočen)|(stockh)|(stoup)|(stopý)|(stopn)|(stom)|(stopnut)|(sto[ph]ov)|(stopk[ao][vt])|(stop[eé]r)|(stopař)|(stožár)|(stolov)|(stoluj)|(stonkov)|(stodol)|(stoli[čc])|(stolař)|(stonav)|(stoj[aiínk])|(stoi)|(stodsk)|(stochovsk)|(stoneovsk)).*([íáé])|([éi]ho)|([éi]mu)|([éyí]m)|([ýi]ch)|([ýi]m)|([ýi]mi))" & lemma!="(dv[oě][rř])|(jedna[cn])|(jednajíc)|(jedený)|(osmán)|(osma[hžn])|(osmirk)|(pater[nů])).*([íáé])|([éi]ho)|([éi]mu)|([éyí]m)|([ýi]ch)|([ýi]m)|([ýi]mi))" & tag="X.*"]

Opakování kompozit

1	p / n	devítí	131	0,02
2	p / n	trojící	89	0,01
3	p / n	jedného	83	0,01
4	p / n	pětí	77	0,01
5	p / n	čtyří	74	0,01
6	p / n	čtyřvalentní	71	0,01
7	p / n	jednoručným	68	0,01
8	p / n	jednačtyřicetičlenném	66	0,01
9	p / n	jedinné	64	0,01
10	p / n	jednolivých	60	0,01
11	p / n	pětkuné	54	0,01
12	p / n	jednohákového	53	0,01
13	p / n	dvoutkání	53	0,01
14	p / n	jednácté	52	0,01
15	p / n	jedným	51	0,01
16	p / n	čtverčních	50	0,01
17	p / n	jednadvacetiměsíční	49	0,01
18	p / n	třináctipodlažní	47	0,01
19	p / n	jednotivých	46	0,01

Substantiva z číslovek na -iny

- Homonymie názvů
zlomků . *ina a
názvů jubileí . *iny

Celkem: 9671 (194 str.)

	Filtr	lemma	Frekv
1.	p/ n	noviny	353 1
2.	p/ n	prázdniny	309 1
3.	p/ n	dějiny	253 8
4.	p/ n	narozeniny	175 0
5.	p/ n	Losiny	16 4
6.	p/ n	Sviny	11 5
7.	p/ n	křtiny	11 5
8.	p/ n	Polabiny	7 8
9.	p/ n	└padesátiny	7 6
10.	p/ n	hodiny	7 4
11.	p/ n	spaliny	7 4
12.	p/ n	Kvasiny	7 2
13.	p/ n	└šedesátiny	7 0
14.	p/ n	└sedmdesátiny	5 5

. *iny (lemma je nom. pl.)

- [lc="((deset) | (((jeden) | (dva) | (tři) | (čtr) | (pat) | (šest) | (sedm) | (osm) | (devate))(náct)) | (((dva) | (tři) | (čtyři))(c[áe]t)) | (((pa) | (še) | (sedm) | (osm) | (deva))(desát)) | (set))iny" & tag="N.*"]

pouze

	Filtr	<u>lemma</u>	<u>Frekvence</u>
1.	p/ n	padesátiny	7 692
2.	p/ n	šedesátiny	7 038
3.	p/ n	sedmdesátiny	5 519
4.	p/ n	osmdesátiny	5 081
5.	p/ n	devadesátiny	2 670
6.	p/ n	čtyřicátiny	2 006
7.	p/ n	třicetiny	1 266
8.	p/ n	dvacátiny	190

Bez lemmatizace a tagování

- [lc="(((deset) | (((jeden) | (dva) | (tři) | (čtr) | (pat) | (šest) | (sedm) | (osm) | (devate))(náct)) | (((dva) | (tři) | (čtyři))(c[áe]t)) | (((pa) | (še) | (sedm) | (osm) | (deva))(desát)) | (set))in) | (((deset) | (((jeden) | (dva) | (tři) | (čtr) | (pat) | (šest) | (sedm) | (osm) | (devate))(náct)) | (((dva) | (tři) | (čtyři))(c[áe]t)) | (((pa) | (še) | (sedm) | (osm) | (deva))(desát)) | (set))in(y) | (ám) | (ách) | (ami))"]

Lemmata a POS (zlomek - číslovka/**C**, jubileum - substantivum/**N** DESAMBIGUACE)

1	desetina	C	23194,00
2	setina	C	20377,00
3	padesátiny	N	7692,00
4	šedesátiny	N	7038,00
5	sedmdesátiny	N	5519,00
6	osmdesátiny	N	5081,00
7	devadesátiny	N	2670,00
8	čtyřicátiny	N	2006,00
9	dvacetina	C	1736,00
10	osmnáctina	C	1531,00
11	třicetina	C	1267,00
12	třicetiny	N	1266,00
13	čtyřicetina	C	860,00
14	patnáctina	C	840,00
15	šedesátina	C	429,00
16	padesátina	C	405,00
17	dvanáctina	C	367,00
18	sedmdesátina	C	312,00
19	osmdesátina	C	309,00
20	šestnáctina	C	270,00
21	devadesátina	C	226,00
22	dvacátiny	N	190,00
23	sedmnáctina	C	177,00
24	devatenáctina	C	94,00
25	třináctina	C	41,00
26	čtrnáctina	C	40,00

šestnáctiny (patrně je problém ve slovníku nikoli v desambiguaci)

stupni se brevis dělí na 8 nebo 12 dílů (šestnáctiny/šestnáctina/CyFS2-----) .</p><p>Užívá se šestilinkový systém
tatínkové kupují svým dcerám přívěsek ve tvaru srdce a k	šestnáctinám/šestnáctina/CyFP3-----	perly . To tvoří základ jejich budoucí šp
mnohem víc hodí do klubů .</p><p>Ačkoliv letos oslaví teprve	šestnáctiny/šestnáctina/CyFS2-----	, na pohled a poslech působí velmi vys
je teď jejich práce . V březnu 1944 Dagmar oslavila	šestnáctiny/šestnáctina/CyFS2-----	a od neznámé paní se známým jménen
v obývacím pokoji jako Rinaldioví , když připravovali oslavu Angeliných	šestnáctin/šestnáctina/CyFP2-----	. Ještě jednou , nadšeně vykřikla a zatl
synem neděje něco divného . " Blížil se termín Honzíkových	šestnáctin/šestnáctina/CyFP2-----	(2 . 10 .) . Adla se vrátila
přišla domů poprvé opilá . To bylo krátce po mých	šestnáctinách/šestnáctina/CyFP6-----	. Přišla jsem v sedm večer , ale úplně n
královny nemá nikdo tušení . Ale s blížícími se Štěpánčinými	šestnáctinami/šestnáctina/CyFP7-----	všechno praskne a vyjde najevo , že by
bronz , a dal si tak hezký dárek k pondělním	šestnáctinám/šestnáctina/CyFP3-----	.</p><p>Maruškovou trápí záda</p><
: Nicolovi teď bylo osmnáct , Rachele oslavila v srpnu	šestnáctiny/šestnáctina/CyFS2-----	. V roce 2008 , kdy jsme převzali ořeže
mohl vypadat v padesáti .</p><p>Milé dámy ,</p><p>Katka slaví	šestnáctiny/šestnáctina/CyFP4-----	... Dohledala jsem pro vás , že se
- , či dokonce sedmiletých dětech chtít , aby chápaly	šestnáctiny/šestnáctina/CyFS2-----	a šestiny ? Jde to a některé školy už to
naskytne na české mutaci MTV . Pořad Moje super užasně	šestnáctiny/šestnáctina/CyFS2-----	stojí na rozmazlených dcerkách neskut

Desambiguace (-1,-1:pos=C)

Podle něho dostávala Fimaco provizi od jedné osminy do jedné	šestnáctiny /šestnáctina/CyFS2-----	procenta . Ruští politici počínaje liberálním ministrem fina
riziko snížené životnosti může být také přijatelné nejvýš u jedné	šestnáctiny /šestnáctina/CyFS2-----	celé délky zábradlí . Pravda je , že výsledek nikdo
se jmenoval Cletus a tvrdil , že je z jedny	šestnáctiny /šestnáctina/CyFS2-----	- Samozřejmě musíte vybrat někoho , kdo se nechá vytočit
slupka plodů může být poškozena , maximálně však do jedné	šestnáctiny /šestnáctina/CyFS2-----	povrchu . Mohou na ní být místa pokrytá korkovitou vrstvi
" A toto je bučimiš . Ten má rytmus patnáct	šestnáctin /šestnáctina/CyFP2-----	a je ze všech nejsložitější , " říká Georgiev .</p>
určují rytmus a melodii : a) předtaktí ze 2	šestnáctin /šestnáctina/CyFP2-----	nebo 1 osminy , melodicky proměňované jako kvarta , terc
o autorská práva nakonec skončily smírem , podle nějž devět	šestnáctin /šestnáctina/CyFP2-----	připadlo Alexandře Lvové a sedm šestnáctin Richardu Haš
centru Prahy spoluvlastníkům . Podle návrhu měl dostat za deset	šestnáctin /šestnáctina/CyFP2-----	domů , v nichž je restaurace a penzion U Medvídků
předkupní právo její spolujatelé . DPP v domě vlastní deset	šestnáctin /šestnáctina/CyFP2-----	, zbytek patří rodině Marečků . V lukrativní budově je
, podle nějž devět šestnáctin připadlo Alexandře Lvové a sedm	šestnáctin /šestnáctina/CyFP2-----	Richardu Haškovi , tehdy ještě nezletilému .</p><p>Oslov
v roce 1997 zavedení nejmenších cenových přírůstků ve výši jedné	šestnáctiny /šestnáctina/CyFS2-----	dolaru , dalším impulsem zřejmě bude přechod na desetin
rozsudku , podle něhož by jim měl Hanzlík vydat patnáct	šestnáctin /šestnáctina/CyFP2-----	domu . Soud rozhodl , že vlastníkem je Hanzlík ,
dědiců tento Francouz tvrdí , že má právo na devět	šestnáctin /šestnáctina/CyFP2-----	Ščukinova majetku .</p><p>Obrazy , mezi nimi skvělý var
dědiců tento Francouz tvrdí , že má právo na devět	šestnáctin /šestnáctina/CyFP2-----	Ščukinova majetku .</p><p>Obrazy , mezi nimi skvělý var
, co slyšíte je elenino choro a má rytmus třináct	šestnáctin /šestnáctina/CyFP2-----	, " vysvětluje vedoucí a rukama vytleskává příslušný rytmu
, na ostatní účely ve městě pak ve výši jedné	šestnáctiny /šestnáctina/CyFS2-----	celoročního rozpočtu až do doby schválení zákona o státr

Desambiguace (-1,-1:pos!=C)

stupni se brevis dělí na 8 nebo 12 dílů (šestnáctiny/šestnáctina/CyFS2-----) .</p><p>Užívá se šestilinkový sys
tatínkové kupují svým dcerám přívěsek ve tvaru srdce a k	šestnáctinám/šestnáctina/CyFP3-----	perly . To tvoří základ jejich budouc
mnohem víc hodí do klubů .</p><p>Ačkoliv letos oslaví teprve	šestnáctiny/šestnáctina/CyFS2-----	, na pohled a poslech působí velmi
je teď jejich práce . V březnu 1944 Dagmar oslavila	šestnáctiny/šestnáctina/CyFS2-----	a od neznámé paní se známým jmé
v obývacím pokoji jako Rinaldioví , když připravovali oslavu Angeliných	šestnáctin/šestnáctina/CyFP2-----	. Ještě jednou , nadšeně vykřikla a :
synem neděje něco divného . " Blížil se termín Honzíkových	šestnáctin/šestnáctina/CyFP2-----	(2 . 10 .) . Adla se vrátila
přišla domů poprvé opilá . To bylo krátce po mých	šestnáctinách/šestnáctina/CyFP6-----	. Přišla jsem v sedm večer , ale úpl
královny nemá nikdo tušení . Ale s blížícími se Štěpánčinými	šestnáctinami/šestnáctina/CyFP7-----	všechno praskne a vyjde najevo , že
bronz , a dal si tak hezký dárek k pondělním	šestnáctinám/šestnáctina/CyFP3-----	.</p><p>Maruškovou trápí záda</p>
: Nicolovi teď bylo osmnáct , Rachele oslavila v srpnu	šestnáctiny/šestnáctina/CyFS2-----	. V roce 2008 , kdy jsme převzali ot
mohl vypadat v padesáti .</p><p>Milé dámy ,</p><p>Katka slaví	šestnáctiny/šestnáctina/CyFP4-----	... Dohledala jsem pro vás , že se
- , či dokonce sedmiletých dětech chtít , aby chápaly	šestnáctiny/šestnáctina/CyFS2-----	a šestiny ? Jde to a některé školy u
naskytne na české mutaci MTV . Pořad Moje super užasně	šestnáctiny/šestnáctina/CyFS2-----	stojí na rozmazlených dcerkách nes

Chyby v desambiguaci [CN]

Mezní situace v obrazech	Šedesátiny/šedesátina/CyFS2	a občanské sdružení Symposion k této příležitosti ve sp
let (tento požadavek se týká těch , kterým se	Šedesátiny/šedesátina/CyFS2	už blíží) . Ideální je , když muž uzavře
.	Šedesátiny/šedesátina/CyFS2	. Parfém (ve stejný den prezentovaný veřejnosti v Paříži
Firma Lancome oslavila svým v pořadí již třicátým parfémem	Šedesátin/šedesátina/CyFP2	pozval řadu hostů
že je po bitvě	Šedesátinami/šedesátina/CyFP7	Legendárního muzikanta poc
- trojnásobného vítěze Velké pardubické Tiumena	Šedesátiny/šedesátina/CyFS2	si žokej Josef Váňa stříhne Velkou pardubickou , ve kter
PRAHA - Týden před	Šedesátiny/šedesátina/CyFS2	, se pokoušel o rekordní deváté vítězství ve Velké pardub
středobodem českého sportu . Josef Váňa , slavící tento týden	Šedesátiny/šedesátina/CyFS2	oslaví islandský prezident Ólafur Ragnar Grímsson
mezinárodního filmového festivalu ; potrvá do 25 . května †	Šedesátiny/šedesátina/CyFS2	Už v pondělí 19 . listopadu oslaví Petra Janů
.	Šedesátiny/šedesátina/CyFP1	vyprodaným koncertem v pražské Lucerně
Už v pondělí 19 . listopadu oslaví Petra Janů	Šedesátiny/šedesátina/CyFS2	Uhlí
Most	Šedesátiny/šedesátina/CyFP1	oslavil bývalý mojiřířský fotbalista a nynější trenér Václav
Kouč Mojžíře Kolankiewicz oslavil životní jubileum	Šedesátiny/šedesátina/CyFS2	a výstava se koná při této příležitosti . Expozice by
Mojžíř	Šedesátiny/šedesátina/CyFS2	. Rodák z Ružomberku oblékal jako fotbalista více než c
od 11 do 18 hodin	Šedesátiny/šedesátina/CyFP1	oslaví dnes asistent trenéra hokejistů Komety Brno Vlad
Známý fotodokumentarista oslavil nedávno	Šedesátiny/šedesátina/CyFS2	, a tak jedním z kritérií pro sestavení programu byla
výrazně zapsal do historie čs. fotbalu , oslavil v neděli		
slíbil Černý		
Trenér Komety Nadrchal slaví šedesátku		
BRNO		
Navíc se vrátí Bakaláři		
Česká televize oslaví příští rok		

dvanáctiny (která jubilea se slaví a o kterých se píše?)

odstavci , vychází , že dovolená bude krácena o čtyři	dvanáctiny/dvanáctina/CyFS2-----	z celkového nároku , který činí 20 dnů . Roční
a výdaji roku 2013 , a to ve výši jedné	dvanáctiny/dvanáctina/CyFS2-----	letošního rozpočtu . Schvalování provizoria ke konci kal
Kdyby pracovní poměr rozvázal 30 . června dostal by šest	dvanáctin/dvanáctina/CyFP2-----	. Rozváže - li zaměstnanec pracovní poměr z jakéhokoliv
zálohy na daň ze závislé činnosti v poměrné výši jedné	dvanáctiny/dvanáctina/CyFS2-----	za každý kalendářní měsíc , na jehož počátku jsou podm
f) přesáhne 1 000 000 Kčs , rozdělí ji plátce na	dvanáctiny/dvanáctina/CyFP4-----	, pokud je jeho zdaňovacím obdobím kalendářní měsíc ,
procítěné . Zpěvák a kytarista Martin Rek se snad s	dvanáctinami/dvanáctina/CyFP7-----	v krvi narodil a charisma z něho sálalo tolik ,
období do schválení rozpočtu na rok 2011 ve výši jedné	dvanáctiny/dvanáctina/CyFS2-----	měsíčních výdajů letošního roku . „ Rozpočet tradičně sc
tohoto roku , bude se každý měsíc hospodařit podle jedné	dvanáctiny/dvanáctina/CyFS2-----	rozpočtu z roku letošního . Až do té doby ,
Výdaje obce mohou být v tomto období jen do výše	dvanáctiny/dvanáctina/CyFS2-----	upraveného rozpočtu na rok 2013 v každém měsíci . Výc
Poměrná část dovolené za kalendářní rok přísluší v délce jedné	dvanáctiny/dvanáctina/CyFS2-----	těž za kalendářní měsíc , v němž zaměstnanec změnil za
II . , což je mince , která měla hodnotu	dvanáctiny/dvanáctina/CyFS2-----	pražského groše . „ Našli jsme také dva čtyřhranné fenik
m podáváme s již uvařenou rýží .</p><p>Energetická hodnota jedné	dvanáctiny/dvanáctina/CyFS2-----	dortu je 1630 kilojoulů , tedy 390 kilokalorií .</p><p>Co

Kolokace (MI-score)

24.	p/n	oslatit	5	2.228	8.135
25.	p/n	těž	4	1.993	8.099
26.	p/n	rozpočet	16	3.985	8.069
27.	p/n	dožit	1	0.996	8.061
28.	p/n	navržený	1	0.996	7.890
29.	p/n	příslušný	4	1.991	7.742
30.	p/n	dot	1	0.995	7.703
31.	p/n	mléčný	1	0.995	7.637
32.	p/n	pojistné	1	0.995	7.590
33.	p/n	placení	1	0.995	7.548
34.	p/n	celkový	12	3.444	7.419
35.	p/n	dovolený	1	0.994	7.267
36.	p/n	měsíčně	2	1.405	7.224
37.	p/n	studentský	1	0.992	7.008
38.	p/n	obyvatelstvo	1	0.992	6.885
39.	p/n	předposlední	1	0.991	6.847
40.	p/n	rozpočtový	1	0.991	6.790
41.	p/n	vzniklý	1	0.991	6.784
42.	p/n	pozemek	5	2.215	6.747
43.	p/n	slavit	3	1.716	6.729

Úkol na příště

- Kompozita s prvním členem číslovkovým tvoří adjektivní deriváty označující vztah k časovému údaji /roční, měsíční, denní, týdenní
- Které číslovkové základy se pojí s více než jedním adjektivem s časovým významem?
- (Použijte nástroj Morfio.)
- Zopakujte pro /minutový, hodinový
- Zopakujte pro /patrový, podlažní
- Zopakujte pro /hranný, stěnný