

A.

V češtině je prefixace jedním z nejrozšířenějších způsobů tvoření nových sloves. Podívám-li se na slovesné prefixy, zjistíme, že v naprosté většině případů jsou „krátké“, zatímco u jmen (substantiv i adjektiv) tomu tak není.

Například:

na-brat ALE ná-bor

za-brat ALE zá-bor

při-jmout ALE při-jem

u-jmout ALE ú-jma

vy-brat ALE vý-bor

po-hnat ALE pů-hon

do-vést ALE dů-vod

pro-vést ALE prů-vod

Přesto existují i případy, kdy se i u sloves můžeme setkat s délkou v prefixu.

1. Popiš postup, jak vyhledat v korpusu (SYN2010) data pro výzkum uvedeného jevu (tj. lemmata sloves, která začínají na řetězce shodné s dlouhými prefixy).
2. Pozorujte data a pokuste se je interpretovat.

```
[lemma="((ná)|(zá)|(dů)|(pů)|(prů)|(pří)|(vý)|(ú)).*" & tag="V.*"]
```

67 lemmat

Časté jsou deriváty od jmen, která mají dlouhý prefix.

Objevují se případy, kde nejde o prefix (*zářit, úpět*).

O prefix slovesného tvaru jde pouze v těchto případech: *záležet, záviset, náležet, závidět*.

V ostatních případech je motivujícím slovem jméno (*důvěra*→*důvěřovat*, *závod*→*závodit*), popř. etymologicky se jménem souvisí (*zápasit* ← *držet se za pas*, *působit* ← *po sobě* viz *násobit* ← *dávat na sebe*).

Výjimka je též *půjčit* a alternace v syntetickém futur slovesa *jít* (*půjde*).

B.

Vyhledejte v korpusu všechny výskyty adjektiva **vroucí**. Pozoruj je a popiš významy, které toto adjektivum může mít (alespoň dva).

Navrhni postup, jak vybrat pouze ty konkordanční řádky, na nichž má adjektivum jeden nebo druhý význam.

Vyhledáme v korpusu lemma **vroucí**.

Výskytů: 645 i.p.m. 5,34 vztaženo k celému syn2015 ARF 0: 250,56 Výsledek je promíchán			
Výběr řádků: základní ▾ (0 vybrané)			
<input type="checkbox"/>	DIGIfoto	ať dáváme pozor , kam šlapeme a vypustil nás mezi	vroucí /vroucí/AANP4----1A----- bahenní jezírka . Naše
<input type="checkbox"/>	Swamplandie	sírka , kterou škrtněš sám o sebe – maličké ,	vroucí /vroucí/AANS4----1A----- volání – tak jsem pak z
<input type="checkbox"/>	Tina	mozzarely (tříšinky) 1 . Těstoviny uvaříme v osolené	vroucí /vroucí/AAFS6----1A----- vodě na skus . Scedím
<input type="checkbox"/>	Kletba čachtické paní	chladne . " , Také bych si přál zemřít ve	vroucí /vroucí/AAFS6----1A----- modlitbě před oltářem
<input type="checkbox"/>	Fajn život	POSTUP : 1,5 kg malých klasů očistíme a vložíme do	vroucí /vroucí/AAFS2----1A----- osolené vody a uvařím
<input type="checkbox"/>	Co je tvoje, je i moje	být v pořádku , teda pokud tam někdo zase nenalije	vroucí /vroucí/AAIS4----1A----- tuk . " Sevíru ruku v pěs
<input type="checkbox"/>	Aha! neděle	33zzzt 3krát denně čaj ze lžičky byliny zalité 250 ml	vroucí /vroucí/AAFS2----1A----- vody , nechte 15 minut
<input type="checkbox"/>	Michelangelo	/ že ač jsem tolik stár / zas rozpálí mě	vroucí /vroucí/AGFS2-----A----- touhy žár ? / Já ohněm
<input type="checkbox"/>	Ti nejlepší z nás	se vytratí a mladý vědec se téměř úplně utopí ve	vroucí /vroucí/AAFS6----1A----- zamilovanosti . Pál uká
<input type="checkbox"/>	Rok krysy	čím víc je vysvětluješ jako bys do pánve se sádlem	vroucím /vroucí/AANS7----1A----- napospas těm kteří si 1
<input type="checkbox"/>	Magický čtverec Albrechta ...	pozoruje . , Doufejme , že nás nepřivítají kádí s	vroucím /vroucí/AAIS7----1A----- olejem , " utrousila sarí

Dále se zaměříme na kolokace vpravo (adjektivum rozvíjí substantivum ve funkci shodného přívlastku, který v češtině typicky stojí před substantivem, které rozvíjí): *voda, oselený, polévka, vývar, kotel, ... ALE přání, láska, hlas, přátelství, vztah, slovo, polibek.*

Kolokace

Atribut: lemma ▾ **V rozsahu od:** 1 do: 1

Minimální frekvence v korpusu: 5

Minimální frekvence v daném rozsahu: 1

Zobrazit funkce:

T-score

MI

MI3

log likelihood

min. citlivost

logDice

MI.log_f

relativní frekvence

Setřídít dle:

T-score

MI

MI3

log likelihood

min. citlivost

logDice

MI.log_f

relativní frekvence

Vytvořit seznam

Celkem: 130 (3 str.)

	Filtr		Frekvence	T-score	MI
1.	p/n	voda	315	17.729	9.884
2.	p/n	osolený	51	7.141	14.865
3.	p/n	modlitba	19	4.355	10.178
4.	p/n	,	17	-6.291	-1.337
5.	p/n	olej	14	3.730	8.337
6.	p/n	vodní	13	3.591	7.969
7.	p/n	a	10	-2.126	-0.742
8.	p/n	polévka	9	2.994	8.933
9.	p/n	vývar	8	2.826	10.190
10.	p/n	kotel	7	2.640	8.890
11.	p/n	čaj	6	2.435	7.379
12.	p/n	přání	6	2.435	7.357
13.	p/n	láska	5	2.186	5.476
14.	p/n	tekutina	5	2.228	8.046
15.	p/n	káva	4	1.979	6.579
16.	p/n	mléko	4	1.981	6.704
17.	p/n	hlas	4	1.911	4.493
18.	p/n	láva	3	1.730	10.058
19.	p/n	roztok	3	1.720	7.183
20.	p/n	oddanost	3	1.730	9.821
21.	p/n	přátelství	3	1.723	7.605
22.	p/n	vztah	3	1.628	4.059
23.	p/n	slovo	3	1.561	3.342
24.	p/n	polibek	3	1.725	7.936

Pomocí Výběru řádků označíme skupinu 1 a 2.

Výskytů: 645 i.p.m. 5,34 vztaženo k celému syn2015 ARF 0: 250,56 Výsledek je promíchán			
Výběr řádků: skupiny ▾ (8 vybrané) ⓘ			
1	DIGifoto	af dáváme pozor , kam šlapeme a vypustil nás mezi	vroucí /vroucí/AANP4-----1A----- bahenní jezírka . Naše
2	Swamplandie	sírka , kterou škrtněš sám o sebe – malíčké ,	vroucí /vroucí/AANS4-----1A----- volání – tak jsem pak
2	Tina	mozzarella (třešinky) 1 . Těstoviny uvaříme v osolené	vroucí /vroucí/AAFS6-----1A----- vodě na skus . Soedín
1	Kletba čachtické paní	chladně . " Také bych si přál zemít ve	vroucí /vroucí/AAFS6-----1A----- modlitbě před oltářem
1	Fajn život	POSTUP : 1,5 kg malých klasů očistíme a vložíme do	vroucí /vroucí/AAFS2-----1A----- osolené vody a uvaří
1	Co je tvoje, je i moje	být v pořádku , teda pokud tam někdo zase nenalije	vroucí /vroucí/AAIS4-----1A----- tuk . " Sevřu ruku v pě
1	Aha! neděle	33zzzt 3krát denně čaj ze lžičky byliny zalité 250 ml	vroucí /vroucí/AAFS2-----1A----- vody , nechte 15 minu
2	Michelangelo	/ že ač jsem tolik stár / zas rozpálí mě	vroucí /vroucí/AGFS2-----A----- touhy žár ? / Já ohněi

C.

V korpusu SYN nejsou slovní tvary *mudla/mudlové* (viz věty níže) rozpoznány automatickou morfologickou analýzou.

Popište, jak byste určili lemma a tag, aniž byste museli vědět, co slovní tvar označuje.

A *mudla* zůstává - stejně jako v Harrym Potterovi - nezasvěcen tajemnému vědění

A teď mají tu možnost i malí *mudlové*.

Předpokládáme-li, že lemmatizujeme tvar bez kontextu, pak můžeme postupovat následovně:
[lc="(.*dla)(.*dlové)"]

Výskytů: 110 893 | i.p.m. 918,38 vztaženo k celému syn2015 | ARF 58 074,12 | Výsledek je promíchán 1 / 2 773

Výběr řádků: základní (0 vybrané)

<input type="checkbox"/> Na čem záleží	oznánil, že si na chvíli odskočí do nemocnice,	kladla /klást/VpFS----R-AA--I	jsem mu na srdce, aby s
<input type="checkbox"/> Zítřka začnu nový život!	jsem v téhle oblasti svého života nedokázala být nezaujatá,	rozhodla /rozhodnout/VpFS----R-AA--1P	jsem se, že vyřeším ty o
<input type="checkbox"/> Atlas mraků	. Ta ženská to myslela upřímně – pámbičkářky to tak	zpravidla /zpravidla/Db-----	mají – , o to ale nebyla n
<input type="checkbox"/> Lidová píseň a sborová tvorba	a téměř instrumentální polyblivost (srov. amen či slova baba	spadla /spadnout/VpFS----R-AA--1P	z kamen). Rozverný ráz
<input type="checkbox"/> Jihlavské listy	sjezd schodů u penzionu Mahler . Tam jsem většinou kolo	vedla /vést/VpFS----R-AA--I	. Takže pro příští ročník j
<input type="checkbox"/> Žena a život	" Zmenšila porce a po dvou třech hodinách něco malého	snědla /sníst/VpFS----R-AA--P	. A kila letěla dolů . „ Ani
<input type="checkbox"/> Deníky Bohemia	do ústavu pro starší 3 ZRUČSKÁ	LETADLA /letadlo/NNP4-----A-----	OBSADILI MLADÍ FILMAŘI
<input type="checkbox"/> Setkání o Vánocích	. „ Hlášič je tady, " sdělila jí,	popadla /popadnout/VpFS----R-AA--1P	dřevěnou židli, vylezla n
<input type="checkbox"/> Benešovy dekrety	přítomných členů Prozatímního národního shromáždění ze Slovenska . Volitelnost posléze	dopadla /dopadnout/VpFS----R-AA--1P	tak, že povolené politick

Frekvence slovních druhů:

Frekvenční distribuce

Frekvenční limit: 0

Celkem: 5 (1 str.)

	Filtr	pos [part of speech]	Frekvence
1.	p/ n	V	56 448
2.	p/ n	N	46 998
3.	p/ n	D	6 277
4.	p/ n	A	882
5.	p/ n	X	288

1. sloveso: *mudla/mudnout/(Vp-SF)|(Vp-PN)* jako *rozhodla/rozhodnout*
2. substantivum: *mudla/mudlo/NNN(S2|P[145])* jako *divadla/divadlo*,
mudla/mudl/NNMS[24] jako *pudla/pudl*, *mudla/mudla/NN[MF]S1* jako *modla/modla*
nebo *Kudla/Kudla*, *mudlové/mudla/NNMP[145]* jako *Kudlové/Kudla*,
mudlové/mudl/NNMP[145] jako *pudlové/pudl*,
3. adjektivum *mudlové/mudlový/AAFS[236]* jako *nudlové/nudlový*.

S desambiguací – pravidlovou:

mudla/mudla|mudl/NNMS1

mudlové/mudla|mudl/NNMP[145]

D.

Jaké lemma a pos (slovní druh) mohou mít tučně vyznačená slova.

*To má být **výchovné**?*

***Místo** vašich schůzek sis měl lépe vybírat.*

***Jí je** špatně.*

***Vezme** ho domů.*

výchovné/výchovné/NNIS1|výchovné/výchovný/AANS1

místo/místo/NNNS4|místo/místo/R---2

***jí/ona**/P.FS3*

***jí/jíst**/VB-S---3*

***vezme/vzít**/VB-S---3*

***vezme/vézt**/Vi-P---1*