

Substantiva skloňovaná podle vzoru žena s alternacemi kořenové samohlásky

(Používáme korpus SYN2010)

Formulace dotazu pro získání dat z korpusů

V minulé lekci jsme ukázali postup, pomocí kterého můžeme hledat v korpusech substantiva skloňovaná podle určitého vzoru. V této lekci bude naším cílem vyhledat substantiva skloňovaná podle vzoru *žena* taková, jejichž kmenová samohláska se v genitivu plurálu obligatorně a v některých dalších pádech fakultativně krátí (alternuje).

V menu **Korpus** zvolíme **Implicitní atribut tag** a do dotazovacího řádku napíšeme dotaz ve formě odpovídající morfologické značky, tedy *NNF.** (hledáme feminina). V dalším kroku zvolíme v menu pod dotazovacím řádkem **P-filtr**, v menu **Korpus** zvolíme **Implicitní atribut lemma**, do dotazovacího řádku napíšeme dotaz ve formě *.*a* (chceme vyhledat lemmata zakončená na -a, tedy ta, která se skloňují/mohou skloňovat podle vzoru *žena*). V dalším kroku ponecháme v menu pod dotazovacím řádkem **P-filtr**, v menu **Korpus Implicitní atribut lemma**, do dotazovacího řádku napíšeme dotaz ve formě *(.*[áéíýóů]..)|(*ou..)*. (Chceme získat přehled o tom, která substantiva mají dlouhou kmenovou samohlásku a je tedy možné, aby u nich docházelo k její alternaci.) 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*). V menu **Konkordance** zvolíme **Statistiky – Frekvenční distribuce** a získáme seznam kandidátů lemmat hledaných substantiv.

Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0 Řádků k zobrazení: 1127

lemma: 0<KWIC zobrazit součet schovat

lemma:	##
vláda	36367
síla	33489
zpráva	27104
smlouva	25637
míra	16136
třída	15758
scéna	12689
ztráta	12677
brána	12639
správa	12072
půda	11931
armáda	11683
rána	10354
dráha	9759
atmosféra	8731
analýza	8520
víra	8397
kariéra	8213
dáma	7868
krása	7592
touha	7526
sezóna	7506
káva	7054

Zavřít Uložit

Ponecháme **P-filtr**, v menu **Korpus** zvolíme **Implicitní atribut tag**, do dotazovacího řádku napíšeme dotaz ve formě *NNFP2.**. Poté nastavíme pod dotazovacím řádkem **N-filtr**, v menu **Korpus** zvolíme **Implicitní atribut word** a do dotazovacího řádku napíšeme *(.*[áéíýóů].)|(*ou.)*.

V menu **Konkordance** zvolíme **Statistiky – Frekvenční distribuce** (lemma a word) a získáme seznam lemmat hledaných substantiv.

Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0 Řádků k zobrazení: 165

lemma: 0<KWIC zobrazit součet schovat

word: 0<KWIC zobrazit součet schovat

lemma:	word:	##
síla	sil	6205
smlouva	smluv	1783
houba	hub	1694
dráha	drah	1191
rána	ran	989
skála	skal	781
dáma	dam	656
brána	bran	636
čára	čar	534
díra	děr	519
kráva	krav	503
žíla	žil	399
pára	par	338
váha	vah	300
žába	žab	247
tráva	trav	231
lípa	lip	209
louka	luk	201
bouda	bud	137
blána	blan	136
bouda	Bud	105

Zavřít Uložit

Třídění a pozorování dat získaných z korpusů

Při srovnávání obou seznamů si povšimneme, že ne všechna lemmata substantiv skloňovaných podle vzoru žena, která mají ve kmeni dlouhou samohlásku nebo diftong, mají v češtině alternaci. Můžeme tedy tvrdit, že alternace není obligatorní.

Dále můžeme sestavit seznam lemmat, která v genitivu plurálu alternaci mají. Tento seznam můžeme porovnávat např. se slovníky, mluvnicemi, údaji z *Internetové jazykové příručky* (dále *IJP*)¹.

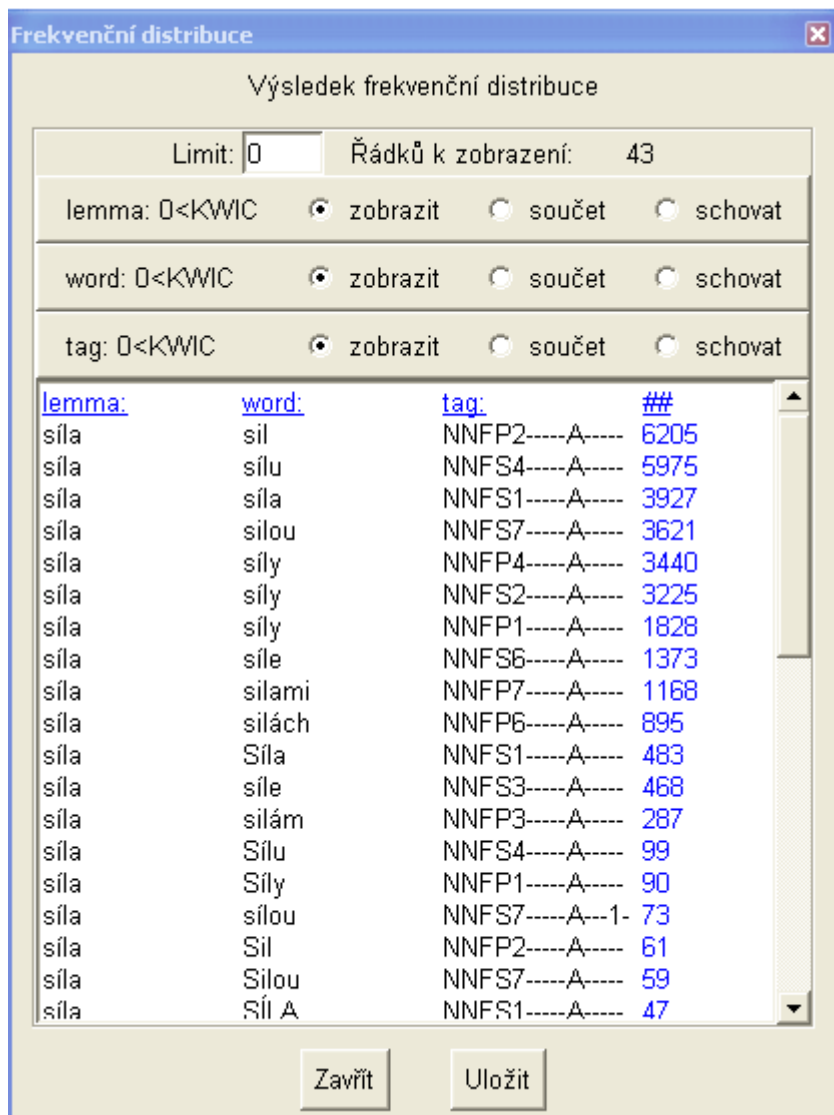
Lze se zamýšlet nad tím, zda se v případě dvojic jako *sezóna/sezon*, *disciplína/disciplin*, *turbína/turbin*, *premiéra/premier*; ... , ale třeba i *bába/bab* jedná o hledaný jev. Vidíme také některé nedostatky v zadáních dotazů (bylo by lépe namísto implicitního atributu **word** volit

¹ *Internetová jazyková příručka* [online]. 2011. Cit. 20. 03. 2011, dostupná z: <<http://prirucka.ujc.cas.cz>>.

lc). Projdeme-li frekvenční seznam a vyloučíme-li omyly, získáme seznam substantiv s alternací v genitivu plurálu. Nyní je třeba, abychom ověřili, zda a za jakých podmínek může/musí docházet k alternacím i v jiných pádech.

Nejdříve si položíme otázku, o které pády jde. Můžeme vycházet a) z jazykové kompetence rodilých mluvčích, b) z jazykových příruček (mluvnice, slovníky) a c) z korpusových dat. Pokud budeme vycházet z prvních dvou, pak budeme cíleně v korpusech vyhledávat u lemmat substantiv ze seznamu (těch, která mají v genitivu plurálu příslušnou alternaci) tvary podle tagu, tedy pádové formy, v nichž může také docházet k alternacím. Můžeme ale také postupovat čistě deduktivně a pády, v nichž k alternacím dochází, získat na základě pozorování korpusových dat.

Posledně zmíněný postup může vypadat následovně. V menu **Korpus** zvolíme **Implicitní atribut lemma**, do dotazovacího řádku napíšeme dotaz ve formě *síla (smlouva, houba, dráha, ...)*. Poté v menu **Konkordance** zvolíme **Statistiky – Frekvenční distribuce** a vybereme **lemma**, přidáme **word** a **tag** a získáme seznam lemmat, tvarů a značek, z něhož je patrné, ve kterých pádech a s jakými frekvencemi dochází/nedochází k příslušné alternaci.

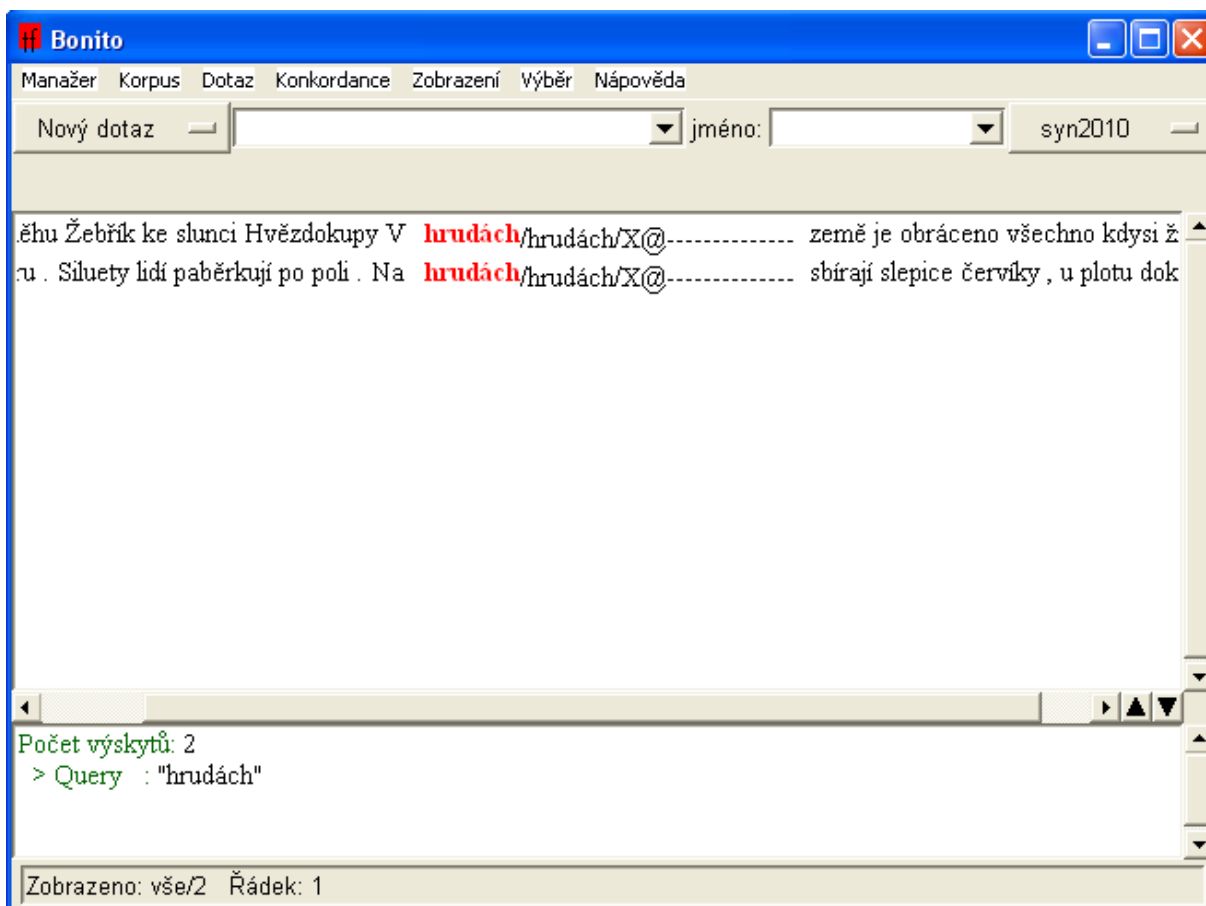


Výsledky pak můžeme zapsat do tabulky například takto:

<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

Formulace závěrů

Na základě analýzy korpusových dat lze např. tvrdit, že alternace v jiných pádech než v genitivu plurálu nemají feminina s kmenovým diftongem *-ou-* (*smlouva, houba, bouda, ...*). V korpusu jsou doloženy tvary *hrudách*, které nebyly rozpoznány automatickou morfologickou analýzou. Tyto nálezy lze ovšem spíše interpretovat jako doklady teritoriálně podmíněného neprovedení přehlásky *u-ou* (viz příklady nominativu plurálu *hrudy*). Totéž platí i pro varianty *mucha, muchy*.



Bonito Manažer Korpus Dotaz Konkordance Zobrazení Výběr Nápověda

Nový dotaz jméno: syn2010

- sípavý brook rudý v skrytu rozpálené **hrudy**/hrudy/X@----- - - tetelí se nedočkavě . Neustane m
 ásnické sbírky Spiewajuco piasc , Hlos **hrudy**/hrudy/X@----- a v roce 1936 stihl ještě vydat Wybr:

Počet výskytů: 2
 > Query : "hrudy"

Zobrazeno: vše/2 Řádek: 1

Bonito Manažer Korpus Dotaz Konkordance Zobrazení Výběr Nápověda

Nový dotaz jméno: syn

staru . Když se kolem ni začínaju slítavat **muchy**/muchy/X@----- a posedávat po ni , tak poznám , že je
 sem na chodbě a v chlěvku chytal pro ňu **muchy**/muchy/X@----- . Fotr zase odpoledne . Měli sme ju a
 bylo účinné . Byl jsem zničený a slabý jak **mucha**/mucha/X@----- a nemohl jsem ani nadávat na tu rudc
 sílu a teploměr šplhá nad třicítku Žerou mě **muchy**/muchy/X@----- , potok šustí a země volá po zálivce .
 já) a tak jsem mulesa být v kra'asech a **muchy**/muchy/X@----- mě furt ohlodávaly . Část vedení byla
 nesmí nabýt vrchu ! Jinak by to bylo na **muchy**/muchy/X@----- . Záporně nulová úroveň některých č
 který byly asi toho starýho pána . Taky **muchy**/muchy/X@----- bylo slyšet , jak pořád litaly a vždycky
 ozvalo klepnutí . Jak jsem se tak na ty **muchy**/muchy/X@----- dívala , povídala jsem si , že jsou všuc
 . Po každém výkřiku se jen ožene jako na **muchy**/muchy/X@----- a dobíhá drobnými krůčky kováře . U
 pochromil , ale já tenkrát netušil , že šlo o **muchy**/muchy/X@----- , které se slétly na podmásli , a taky
 Čečenců v téhle k zblití špinavé válce byly **muchy**/muchy/X@----- - granátometry . Benzin však potřebo
 vývoji škůdců . 88 Vodní clona odpuzuje **muchy**/muchy/X@----- na přicházejícím dobytku (dle CSL)

> Deleted : 1
 > Deleted : 1
 > Deleted : 1

Zobrazeno: vše/36 Řádek: 1

Na základě frekvenční analýzy můžeme tvrdit, že v korpusech nelze najít oporu pro to, že by varianty s alternacemi *á-a*, *í-i*, *í-ě* v instrumentálu singuláru a dativu, lokálu a instrumentálu plurálu byly ve všech případech rovnocenné s variantami bez alternací, jak naznačuje interpretace SSČ. Liší se na první pohled frekvencí, přičemž rozdíly jsou vázány na jednotlivé lexémy.

Ve sledovaném materiálu lze najít dokonce i okrajovou alternaci *ů-o*, a to u substantiva *fůra/for*.

ssjc Slovník spisovného jazyka českého

fůra

-y ž. (2. n. -fůr, -for) (z nám.)

1. *vůz s těžkým nákladem, vůz na těžký náklad; množství nákladu, které se na takový vůz vejde*: f. uhlí tažená párem koní; nevešlo se to ani na pět for, - koupil několik fůr dříví, [s] ob. expr. mít rozumu na fůry *velmi mnoho*

2. ob. expr. *velké množství*: mít fůru peněz, práce, času; --> zdrob. **fůrka**, **for**ka, expr. **fůrečka**, -y ž.

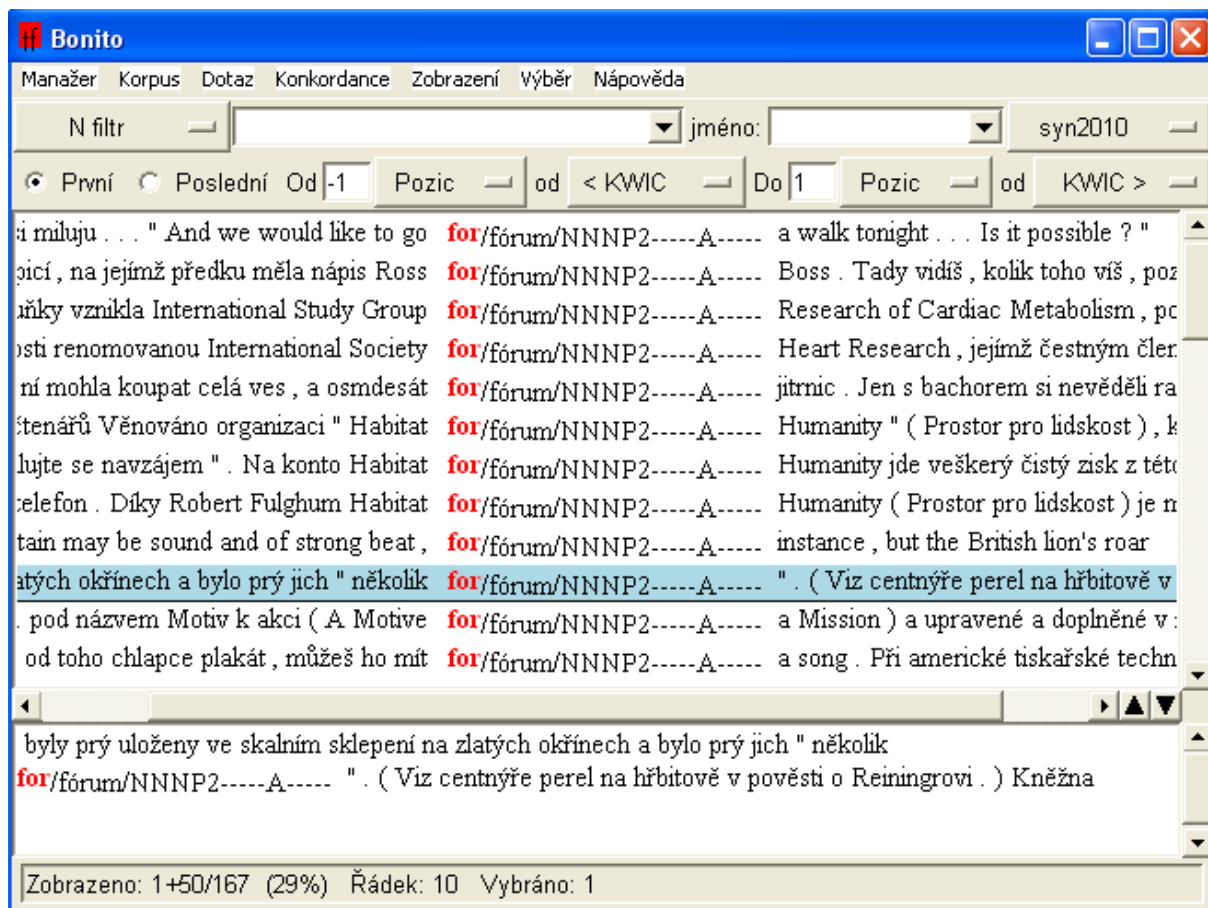
psjc Příruční slovník jazyka českého

fůra, -y f. (fůrou i forou, fůr i for, fůrám, fůrách, fůrami) *vůz s těžším nákladem*. Přijela fůra s nábytkem. Šmil. Naloží fůru jako horu vysokou. Baar. D *povoz na těžší náklad*. Chasa přijela s fůrami, aby seno odvezla. Klost. Měli jsme vás znát, na fůru vás dát, za hranice vyjet, tam vás vysypat. Havl. D *těžší náklad na voze*. Fůru pšenice [na půdu] složím. Rais. Kdo si může najednou koupit fůru, aspoň mu ve sklepeš vyschne [uhlí]. Herrm. D *Expr. mnoho, velmi mnoho*. Jenom ne se otlárat o hejtnana, z toho by byla fůra nepřijemných opletaček. Vrba. Probreptáte tady fůru času a za to vám má mistr platit? Krás.

Nenajdeme ji ovšem výše popsaným postupem, protože doklad je chybně interpretován automatickou morfologickou analýzou, což musíme u okrajových jevů vždy předpokládat.

Výskyt genitivu plurálu substantiva *fůra* vyhledáme následovně.

V menu **Korpus** zvolíme **Implicitní atribut word** a do dotazovacího řádku napíšeme *for*. Ručně procházíme data a mezi většinou dokladů anglické předložky mylně desambiguované jako tvar lemmatu fórum hledáme příslušný doklad. Pro usnadnění ruční práce lze pomocí **N-filtru** odstranit řádky, kdy se v intervalu <-1,1> vyskytují slova označovaná jako na pozici 1 jako X, protože to budou s větší pravděpodobností ty řádky, na nichž bude doklad předložkou v anglickém textu.



Formulace dalších otázek vyplývajících ze zkoumání daného jevu

Samohláskové alternace se vyskytují také u dalších vzorů (*pán : vůl, muž : kůň, hrad : mráz, stroj : déšť, ...*). U kterých se nevyskytují? Popište pravidla, jimiž se řídí distribuce alternací v jednotlivých flektivních tvarech.

Zadání cvičení, v nichž lze uplatnit analogické postupy

K podobným alternacím dochází také v genitivu plurálu několika substantiv skloňovaných podle vzoru *růže*. Která substantiva to jsou? Pokuste se je najít v korpusu. Mají obdobné alternace také v ostatních pádech, tak jako substantiva skloňovaná podle vzoru *žena*?