

# **GAK – CJBB84**

čt. 12.30-14.00

1.12. 2011

# Substandardní jevy v jazyce a jazykové korpusech

- MSČ a koncept minimální intervence
- Značkování korpusů automatickými nástroji
- Jak hledat v korpusech okrajové jevy

# Hláskoslovné varianty ý-ej (é-í) aj.

- bejt
- mlejn
- mlíko
- strejda
- rejt

# bejt

The screenshot shows the Bonito software interface. The title bar reads "Bonito". The menu bar includes "Manažer", "Korpus", "Dotaz", "Konkordance", "Zobrazení", "Výběr", and "Nápověda". The search bar contains "Nový dotaz" and "jméno:" with a dropdown menu. The search results are displayed in a text area, showing a concordance for the word "bejt". The results are as follows:

Text	Form	Frequency
plných plic řval , podle všeho to měl asi	<b>bejt</b> /být/Vf-----A---6B	6B
dutej a chraplavej . Zaťukal na okno - mohly	<b>bejt</b> /být/Vf-----A---6B	6B
dumat nad tím , kdo všechno kolem mě může	<b>bejt</b> /být/Vf-----A---6B	6B
: " Wicky , dostaňte mě odtud , začínám	<b>bejt</b> /být/Vf-----A---6B	6B
že to mají na denním pořádku . Prej máme	<b>bejt</b> /být/Vf-----A---6B	6B
nenadělali , to prej dělají , a taky prej má	<b>bejt</b> /být/Vf-----A---6B	6B
se najednou objevil úsměv . " Musela to	<b>bejt</b> /být/Vf-----A---6B	6B
suché vzdálenosti . " Přece nemá vejšku	<b>bejt</b> /být/Vf-----A---6B	6B
Bart už byl popuzený . " Proč by tu musel	<b>bejt</b> /být/Vf-----A---6B	6B
viděl . " " To není možný . Kdy to mělo	<b>bejt</b> /být/Vf-----A---6B	6B
jsem , " život není fér . " " Ale měl by	<b>bejt</b> /být/Vf-----A---6B	6B
. " " Není to nějaká viróza ? " " Nemůže	<b>bejt</b> /být/Vf-----A---6B	6B

The text area also contains the following text: zpěv . Sotva se držel na nohách , jak tak tři hodiny . ' Hejbní kostrou , karr fizl , tak bych v hotelu nevylez z post střízlivej . " Je šest hodin , Barley má rádi , že nám tam nenačurali a nenad máma ráda , že jí nerozřezali čalouně scéna jak dělo . Mámě to připadalo r někde , kde se to veze z kopečka dol ? Není tu . Pracuje pro nějakou plem ? " " Minulý týden . Od včerejška pře . " " Raději se smíř s osudem a počíte , jinak by to měli i ostatní , nejen ty

The status bar at the bottom shows "Zobrazeno: 1+50/1934 (2%) Řádek: 1".

# rejt

The screenshot shows the Bonito software interface. The title bar reads "Bonito". The menu bar includes "Manažer", "Korpus", "Dotaz", "Konkordance", "Zobrazení", "Výběr", and "Nápověda". Below the menu bar, there is a "Nový dotaz" button and a search input field containing "jméno:" and a dropdown menu. To the right, there is a text field containing "syn2010".

The main text area displays the following text with the word "rejt" highlighted in red and its morphological analysis: "rejt/rejt/Vf-----A----I".

okovat , Seneco . Ona musí pořád do někoho  
ixovy polekané výkřiky . " Nezačíněj do mě  
" " Nic se neděje . Když chceš do někoho  
stremadury a kastaněty z ebenu , do olivové

" Mavis vyprázdnila talíř . " Nesmíš  
" řekl tónem , který by se dal nazvat  
, tak rej do toho chlapíka vedle . Viš  
si kůry zubem dívčího hřebenu ! Jó , takhle

At the bottom of the window, there is a status bar showing "Zobrazeno: vše/4 Řádek: 1". Above it, there is a summary of the search results: "Počet výskytů: 4" and "> Query : 'rejt'".

# mejt

The screenshot shows the Bonito software interface. The title bar reads "Bonito" and the menu bar includes "Manažer", "Korpus", "Dotaz", "Konkordance", "Zobrazení", "Výběr", and "Nápověda". The search bar contains "Nový dotaz" and "jméno: syn2010". The main text area displays a concordance search for the word "mejt". The search results are as follows:

pak už byli všichni pryč a já jsem začala	<b>mejt</b> /mýt/Vf-----A---6I	nádobí , najednou slyším , že Perry pláče
. " Tak to vidíte . " " Asi bych začala	<b>mejt</b> /mýt/Vf-----A---6I	vokna , " přece jen se hned zas uklízečka
Burééš , dělat budéééš , anebo pudéééš . A	<b>mejt</b> /mýt/Vf-----A---6I	si nohy budéééš , ha ha ha . " S umytýma
? " Chvilí přemýšlel . " Mám nápad . Umí	<b>mejt</b> /mýt/Vf-----A---6I	auto ? Už to někdy dělal ? " Přisvědčili
situace , kdy se musí lhát . " Výborně , umí	<b>mejt</b> /mýt/Vf-----A---6I	auto , takže mi místo ruštiny zítra umyje
šjaký potok , chceš se vykoupat ? Já ti budu	<b>mejt</b> /mýt/Vf-----A---6I	nohy , chceš ? " " Seš moc ochotnej , "
Vykulil na ni bystré oči . " Nač si mám	<b>mejt</b> /mýt/Vf-----A---6I	ruce a obličej , když se jdem projít ?
rozumíte ? A už vůbec ne , abyste mi chtěl	<b>mejt</b> /mýt/Vf-----A---6I	nohy , já si nohy umím umejt sám , děkuju
výrazem . " Mám točit palci , až ty budeš	<b>mejt</b> /mýt/Vf-----A---6I	okna a drátkovat podlahy ? " " Můžeš mi
Třeba bych mohla - mohla bych zametat a	<b>mejt</b> /mýt/Vf-----A---6I	nádobí . " Dean stál jako přibitý , když
spodě , pomáhala jsem svý blbý matce třeba	<b>mejt</b> /mýt/Vf-----A---6I	nádobí a tak , a potom jsem chodila s kám
vejlupekové , takže strčím drát zpátky a jdu	<b>mejt</b> /mýt/Vf-----A---6I	ty hrušky . Vopláchnul jsem je ve dřezu

Count of occurrences: 23  
> Query : "mejt"

Displayed: all/23 Row: 1

# Pozorujte

- Čím se liší morfologická značka ?
- Jak najdeme všechna (slovesa s KoV –ý-, např. krejt, ...), která tuto podmínku splňují ?

# slovesa s KoV -ý- (word=„\*ýt“)

Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0      Řádků k zobrazení: 110

word: 0<KWIC       zobrazit       součet       schovat

<u>word:</u>	<u>##</u>
být	134369
nebýt	1304
Být	1273
Nebýt	732
skrýt	584
zakrýt	581
pokrýt	554
umýt	538
ukrýt	408
krýt	381
dobýt	356
odbýt	350
nabýt	314
mýt	235
odkrýt	166
přibýt	145
smýt	109
vydobýt	94
pobýt	94
výt	90
přikrýt	87
omýt	60
rozkrýt	59

Zavřít      Uložit



# skrejt

The screenshot shows the Bonito software interface. At the top, there is a menu bar with options: Manažer, Korpus, Dotaz, Konkordance, Zobrazení, Výběr, and Nápověda. Below the menu bar, there is a search bar with a dropdown menu labeled 'Nový dotaz' and a text input field containing 'jméno:'. To the right of the search bar, there is a dropdown menu labeled 'syn2010'. The main area of the window displays search results for the word 'skrejt'. The results are shown in a table-like format with three columns: the original text snippet, the word 'skrejt' in red, and a corresponding text snippet. The results are as follows:

za to . " Ha ! Výborně ! Nedokázal jsem	<b>skrejt</b> /skrejt/X@-----	škodolibou radost nad vysloveným kondic
eslo ? Co to má znamenat ? " Nedokázala	<b>skrejt</b> /skrejt/X@-----	obrovský zklamání . " Marťan říkal , že
stály nějaký stěny , mezi kterejma se šlo	<b>skrejt</b> /skrejt/X@-----	. Byl tehdy studenej celej rok . Autobus
: Možná že to , co se před sebou snažíme	<b>skrejt</b> /skrejt/X@-----	, je to , že skutečně jsme šťastný . KENN

At the bottom of the window, there is a status bar showing the number of occurrences and the current query:

Počet výskytů: 4  
> Query : "skrejt"

At the very bottom, there is a status bar showing the current display settings:

Zobrazeno: vše/4 Řádek: 1

# mlejn

The screenshot shows the Bonito software interface. The title bar reads "Bonito". The menu bar includes "Manažer", "Korpus", "Dotaz", "Konkordance", "Zobrazení", "Výběr", and "Nápověda". The main window has a search bar with "Nový dotaz" and a dropdown menu. To the right, there is a field for "jméno:" and a dropdown menu showing "syn2010". The main text area displays a concordance search for "mlejn". The search results are shown in a table format, with the word "mlejn" highlighted in red in the original image. The search results are as follows:

Text	Tag	Context
rence pokoj . Byla by to jen voda na jeho	mlejn/mlejn/NNIS4-----A-----	. " Třídí se naklonil k řediteli a něco
: hloupým hráč jako je šestašedesát nebo	mlejn/mlejn/NNIS1-----A-----	. Dost už s těmi třemi , kteří mě přivedli
viděl . Rozum do hrsti , Milosti , tenhle	mlejn/mlejn/NNIS1-----A-----	, v kterým se to pořád mele , je jasný
odívat se do domu . " Člověče , tady je asi	mlejn/mlejn/NNIS1-----A-----	, to jsme ještě neviděli " a už mne táhl
to stavení ? Tak tam se narodila , je to	mlejn/mlejn/NNIS1-----A-----	Smetiprach . . . " Jednoho dne šla se Do
oští a pařezech máme nepřeborné zásoby	mlejn/mlejn/NNIS4-----A-----	na ně už uvedli do výroby Výživná jsi bic
echal vám taky statek a hospodu - dvůr a	mlejn/mlejn/NNIS4-----A-----	- Popelka nedopoví - i tak vidí , že mace
na chalupě opět odkládá . " Mám takovej	mlejn/mlejn/NNIS4-----A-----	. Ale tam se vyskytnu dva tři dny za rok
náspi pobafávala , po vantrokách vodu na	mlejn/mlejn/NNIS4-----A-----	vodníkovi pomáhala hnát , za komínem s
notáře jsme už měli připraveného v autě a	mlejn/mlejn/NNIS1-----A-----	byl můj , " vzpomíná Navrátil . Propadlá
OTNÍCI DESNÁ Střípek uvede Kouzelněj	mlejn/mlejn/NNIS4-----A-----	Herci plzeňského divadla Střípek vystou
domě Sklár . Sehraji inscenaci Kouzelněj	mlejn/mlejn/NNIS1-----A-----	, která nadchla diváky na Jiráskově Hro

Počet výskytů: 12  
> Query : "mlejn"

Zobrazeno: vše/12 Řádek: 1

# prejt

The screenshot shows the Bonito software interface. The title bar reads "Bonito". The menu bar includes "Manažer", "Korpus", "Dotaz", "Konkordance", "Zobrazení", "Výběr", and "Nápověda". The main window has a search bar with "Nový dotaz" and a dropdown menu, and a text input field containing "jméno:" and a dropdown menu. The search results are displayed in a table with columns for the original text, the word "prejt", its grammatical information, and the translated text.

Text	prejt	Grammatical info	Translation
adlavej mimozemšťan mi z ní nebude dělat je užitek . Je uzžený špek , je ovar , je hou jitrnici o hlavu pana Pinkase . Růžový se jim zdál nejstravitelnější jelítkový " Prejt " , ten v restauraci s " Prejt , " , ten v restauraci s " Prejt , prejt , v restauraci s " Prejt , prejt , prejt , šku a v troubě zapeč do červena . Voňavý tři krát v dlani , je americký seriálový mnou přílohou je i špaldová rýže . Čočkový	prejt	/prejt/NNIS4-----A-----	do jitrnic . " Koukej ji pustit , ty zatracene , sú klobásky , sú nožky v rosolu . Tož stékal po tvářích obou duelantů , kteří s bramborem a okurkou . ( Optá se ho , , prejt , prejt . . . " a Vikingy . Přečetli , prejt . . . " a Vikingy . Přečetli jsme . . . " a Vikingy . Přečetli jsme ho a nebo spíše nákyp krájej na plátky , k tomu učiněným požeňáním . Naproti tomu Evi z Pošumaví Suroviny : asi 250 g čočky ,

Počet výskytů: 10  
> Query : "prejt"

Zobrazeno: vše/10 Řádek: 1

# Další kandidáti ? (lemm=„.\*ý.“& tag=„,[NX].\*“)

Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0      Řádků k zobrazení: 3734

lemma: 0<KWIC       zobrazit       součet       schovat

lemma:	###
tým	30069
Londýn	8391
sýr	3283
strýc	3258
mlýn	2883
inženýr	2330
kostým	2227
motýl	1678
dým	1384
býk	1318
markýz	1091
kanonýr	951
krokodýl	813
Rýn	783
netopýr	773
pastýř	721
suvenýr	709
týl	548
hýl	514
kýč	496
měchýř	474
Týn	435
kurýr	422

Zavřít      Uložit

# lemma=„strejc“

The screenshot shows the Bonito software interface. The title bar reads "Bonito". The menu bar includes "Manažer", "Korpus", "Dotaz", "Konkordance", "Zobrazení", "Výběr", and "Nápověda". The main window has a search bar with "Nový dotaz" and a dropdown menu. The search criteria are "jméno:" and "syn2010". The main display area shows a concordance table with the following entries:

se ona spokojuje s takovým plesnivým	<b>strejcem</b> /strejc/NNMS7-----A-----	? Nejspíš proto , že je hluchá , usoudil
to z něj dostanu . " Byl to mrzutý starý	<b>strejc</b> /strejc/NNMS1-----A-----	. " Co chcete , kluci ? " vyštěkl . Řekl
stával : " Pročpak jste si takhle dlouho ,	<b>strejče</b> /strejc/NNMS5-----A---1-	, prohlížel ten most ? " " Ale , hochu
í ranou bylo pro nás setkání s norským	<b>strejcem</b> /strejc/NNMS7-----A-----	. Po tříhodinovém pochodu jsme ho pc
aby ČSSD přestala působit jako spolek	<b>strejců</b> /strejc/NNMP2-----A-----	. Paroubek by pro to udělal hodně , kd

At the bottom of the window, the status bar shows "Zobrazeno: vše/5 Řádek: 1". The console area displays the following text:

```
Počet výskytů: 5  
> Query : "strejc"
```

# word=„strejc.\*“

The screenshot shows the Bonito software interface. At the top, there is a menu bar with options: Manažer, Korpus, Dotaz, Konkordance, Zobrazení, Výběr, and Nápověda. Below the menu bar, there is a search bar with the text "Nový dotaz" and a dropdown menu showing "jméno:". To the right of the search bar, there is a text input field containing "syn2010".

The main area of the interface displays a concordance search result. The search query is "strejc.\*". The results are displayed in a table-like format with three columns: the original text, the word and its grammatical information, and the context. The results are as follows:

Text	Word and Grammatical Information	Context
...cijský rozměr , pokročilejší věk ,	<b>strejcovskou</b> /strejcovský/AAFS4----1A-----	blahosklonnost , hnědou barvu c
...obbyho táta , v Quindlu , Basilův	<b>strejc</b> /strýc/NNMS1-----A-----	, věčně vožralej . . . " Odmlčel :
prázdné vytahování podivínského	<b>strejce</b> /strýc/NNMS2-----A-----	, a možná ani to ne . Londýňan
kže se na plácku před školou dva	<b>strejci</b> /strýc/NNMP1-----A-----	, kteří spolu důležitě mluvili , otc
že kolem pomaličku jede na kole	<b>strejc</b> /strýc/NNMS1-----A-----	z nákupu . V ten okamžik byl o
znáním nejednou při rozhovoru se	<b>strejci</b> /strýc/NNMP7-----A-----	na návsi nebo u piva pronášel v
řadící jako veselí , pestře oblečení	<b>strejcové</b> /strýc/NNMP1-----A---1-	, kteří zvoní rolničkami a vyhráv
Jak rád bych tomu věřil ! Moudří	<b>strejcové</b> /strýc/NNMP1-----A---1-	z čistě vědecké zvědavosti vstr
řaguba , laskavého kulatoučkého	<b>strejce</b> /strýc/NNMS4-----A-----	, který ovládal umění , jak pomc
u najednou , že i já jsem už usedlý	<b>strejcovský</b> /strejcovský/AAMS1----1A-----	soused . Že už nepomýšlím na ž
ř: rádoby humorných proslavů , ze	<b>strejcovského</b> /strejcovský/AANS2----1A-----	žvatlání s babkami probleskoval
n uměním přivedl ty makedonské	<b>strejce</b> /strýc/NNMP4-----A-----	do varu tak , že si brzy potleská

At the bottom of the interface, there is a status bar showing "Zobrazeno: 1+50/75 (66%) Řádek: 1".

# Lemmatizace a značkování substandardních tvarů slovesa *být*

- lemma=„být“
- Konkordance-Statistiky-lc+tag
- 158 ř.
- P-filtr tag=„.....6.“

# 17 variant

Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0      Řádků k zobrazení: 17

lc: 0<KWIC       zobrazit       součet       schovat

lc:	###
bejt	1937
sem	1398
seš	1034
bysme	804
jseš	648
sme	615
budem	562
ste	357
sou	162
nejseš	139
nebudem	129
bysem	95
nejni	80
nebejt	72
jseš	4
nejní	2
bejti	1

Zavřít      Uložit



# Pokuste se uvést další a najít, jak jsou značkovány

- sou – su ?

# SU

The screenshot shows the Bonito software interface. The title bar reads "Bonito". The menu bar includes "Manažer", "Korpus", "Dotaz", "Konkordance", "Zobrazení", "Výběr", and "Nápověda". The search bar contains "Nový dotaz" and "jméno:" with a dropdown menu. The search results are displayed in a table with columns for the text snippet, the character "su", its grammatical information, and the surrounding text. The search results show 312 occurrences of "su" in the corpus. The status bar at the bottom indicates "Zobrazeno: 1+50/312 (16%) Řádek: 1".

Text snippet	Character	Grammatical information	Surrounding text
se k malému čínskému důstojníkovi . " Sun	su	/být/VB-S---1P-AA--7I	! " zařval na něj . " Še-ma ? " vyhrkl
přednostní kód Nefritové hory . Byl pro	Su	/být/VB-S---1P-AA--7I	Ťianga . . . " D ' Anjou sebou škubl ,
rezervaci vyplývá , " řekl Francouz , " že si	Su	/být/VB-S---1P-AA--7I	není jistý , kdy se k vrahovi dostane
dpověděla žena kantonsky . " Tady Čou . " " Sun	su	/být/VB-S---1P-AA--7I	! Siao-si ! " zašeptal Lin ochraptěle j
než do něj muž stačil vhodit mince . " Sun	su	/být/VB-S---1P-AA--7I	! Siao Si ! " zašeptal Lin a zakašlal .
ním dá spojit . První z nich byl nějaký	Su	/být/VB-S---1P-AA--7I	Ťiang z posádky v Kuang - tungu . .
Zavolám na to číslo a mluvím francouzsky . Ten	Su	/být/VB-S---1P-AA--7I	Ťiang je jeden z mála šikmookých , k
" V posádce Kuang-tung je člověk jménem	Su	/být/VB-S---1P-AA--7I	Ťiang . Mluvím s ním francouzsky .
plukovníku . " " Ověřím si to sám , " posadil se	Su	/být/VB-S---1P-AA--7I	Ťiang a sáhl po malém plochem obd
pane . " " To doufám , pro vaše dobro . "	Su	/být/VB-S---1P-AA--7I	přiložil kroužek k mluvítku a stiskl tla
Prase ! " " Ven . " O pár vteřin později se	Su	/být/VB-S---1P-AA--7I	vrátil k telefonu a pokračoval franco
" " Dobře se o vás postaráme , " naléhal	Su	/být/VB-S---1P-AA--7I	. " Až moc dobře . " " Co myslíte tír

Počet výskytů: 312  
> Query : "su"

Zobrazeno: 1+50/312 (16%) Řádek: 1

# Pozice 15 - Varianta, stylový příznak apod.

- - neurčuje se ("základní" tvar pro kategorie v pozicích 1-14)
- 1 varianta, víceméně rovnocenná ("méně častá")
- 2 řídká, archaická nebo knižní varianta
- 3 velmi archaický tvar, též hovorový
- 4 velmi archaický nebo knižní tvar, pouze spisovný (ve své době)
- 5 hovorový tvar, ale v zásadě tolerovaný ve veřejných projevech
- 6 hovorový tvar (koncovka standardní obecné češtiny)
- 7 hovorový tvar (koncovka standardní obecné češtiny), varianta k '6,
- 8 zkratky
- 9 speciální použití (tvary zájmen po předložkách apod.)

# P-filtr tag=„.....[123456789].“ 45 variant

Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0      Řádků k zobrazení: 45

Ic: 0<KWIC       zobrazit       součet       schovat

Ic:	###
jest	3346
bejt	1937
sem	1398
seš	1034
bysme	804
býti	692
jseš	648
sme	615
budem	562
ste	357
bud'	304
neni	216
su	210
sou	162
nejseš	139
nebudem	129
bysem	95
si	87
nejni	80
nebejt	72
jet'	21
jsemt'	16
bylt'	13

Zavřít      Uložit

# bychme

- nenalezen
- word=„bychme“

# word=„bychme“

The screenshot shows the Bonito software interface. The title bar reads "Bonito" and the menu bar includes "Manažer", "Korpus", "Dotaz", "Konkordance", "Zobrazení", "Výběr", and "Nápověda". The search bar contains "Nový dotaz" and "jméno:" with a dropdown menu. The search results are displayed in a table with three columns: the original text, the word "bychme" in red, and a truncated version of the text. The search results are as follows:

Text	Word	Text
věděli a nechali to nám , t. j. V+W - tak	<b>bychme</b>	jim zorganizovali svět . 19 . II . 58 -
. Počítáme teď s dopsáním v dubnu . To	<b>bychme</b>	pak klidně byli hotovi pro sezonu 60 -
to píšu ) . Snad , kdybys ušetřil , snad	<b>bychme</b>	se mohli sejít někde v koutečku Evrop
en abys věděl , nic mi nepiš o tom . Měli	<b>bychme</b>	se sejít už sakra . Těžko se mi žije -
at . Nevadilo by to panu Pilařovi ? No to	<b>bychme</b>	se ho zeptali . - Chris a já vám chystám
jen tak pro náladu , pro radost , a než	<b>bychme</b>	se naděli , bylo by to připravený . - Al
h si mohl koupit za koruny letenku . Pak	<b>bychme</b>	mohli ve Vídni ( frau Kubeš vypadla ,
ena , což budeme vědět co nevidět , dali	<b>bychme</b>	si ten výlet do Londýna . ( Pokud mé :
byl vůl a měl víc peněz ( to hlavně ) tak	<b>bychme</b>	tam všichni na ty svatby přiletěli , včet

At the bottom of the window, the status bar shows "Zobrazeno: vše/9 Řádek: 1".

# sú

The screenshot shows the Bonito software interface. At the top, there is a menu bar with options: Manažer, Korpus, Dotaz, Konkordance, Zobrazení, Výběr, and Nápověda. Below the menu bar, there is a search bar with the text "Nový dotaz" and a dropdown menu. To the right of the search bar, there is a field labeled "jméno:" with a dropdown menu and the text "syn2010".

The main content area displays a list of search results for the word "sú". Each result consists of a snippet of text from a document, followed by the word "sú" in red, and then a list of its occurrences in the text, represented by "sú/sú/X@-----".

Count: 44  
> Query : "sú"

Displayed: all/44 Row: 10 Selected: 1

# budó

The screenshot shows the Bonito software interface. At the top, there is a menu bar with options: Manažer, Korpus, Dotaz, Konkordance, Zobrazení, Výběr, and Nápověda. Below the menu bar, there is a search bar with the text "Nový dotaz" and a dropdown menu. To the right of the search bar, there is a field labeled "jméno:" with a dropdown menu and a field containing "syn2010".

The main area of the interface displays search results for the word "budó". The results are presented in a table-like format with three columns. The first column contains text from a document, the second column contains the word "budó" followed by a file path, and the third column contains the corresponding text from the document.

. " Ty si to tady vezmeš na povel . Tobě	<b>budó</b> /budó/X@-----	rozumět . " Pavel nebyl zvyklý přihlížet
, jako za dušičke . Jestlipak dneska už	<b>budó</b> /budó/X@-----	svítit ? Padrta říká , že je to správný
zabijó . . . udělajó z nás vdove . . . děcka	<b>budó</b> /budó/X@-----	sirotci . . . a my . . . budem mět hlad
RÍŠA Kdy potom ? HELENKA AŽ . . . až	<b>budó</b> /budó/X@-----	tančit vostatní . RÍŠA Slyšeli jste ? Dámy
ó ho integrování na dálku . Teď nad náma	<b>budó</b> /budó/X@-----	držet vochrannó ruku s vysavačem . Přis
ochu zeširoka . Význam japonského slova	<b>budó</b> /budó/X@-----	( Bujinkan Budó Tajutsu je systém bojov
Význam japonského slova budó ( Bujinkan	<b>Budó</b> /Budó/X@-----	Tajutsu je systém bojového umění navaz

At the bottom of the interface, there is a status bar showing "Počet výskytů: 7" and "> Query : "budó". Below that, it says "Zobrazeno: vše/7 Řádek: 1".



# Substandardní koncovky

- Adjektivní složená flexe (*ktorejch, ...*)
- I. pl. –ama/-e(ě)ma/-ima (*holkama, klukama, ulicema, ...*)
- 1. pl. ind. prez. akt. –m (*jedem, ...*)
- Kolísání *píseň/kost*

**word=„\*ama“**

- P-filtr tag=X.\*
- P-filtr tag=N.\*

# X.\*

The screenshot shows the Bonito software interface with a search results table. The window title is "Bonito" and the menu bar includes "Manažer", "Korpus", "Dotaz", "Konkordance", "Zobrazení", "Výběr", and "Nápověda". The search filter is set to "X.\*" and the corpus is "syn2010". The table displays search results for the query "X.\*" in the corpus "syn2010". The results are sorted by position (Pozic) and are displayed in a table with three columns: the original text snippet, the search term and its frequency, and the context of the match.

Text snippet	Search term and frequency	Context
pošli do Horshamu pro advokáta	<b>Fordhama</b> /Fordhama/X@-----	. ' Učinil jsem , jak mi nařídí
ld opustil , požádal ještě Fostera	<b>Fordhama</b> /Fordhama/X@-----	, aby se u něho před odchoc
ztracená kulka . Zpráva Fostera	<b>Fordhama</b> /Fordhama/X@-----	o svíčkou zacpaném odtoku
ovídejte , " zahučel mezi dvěma	<b>bafama</b> /bafama/X@-----	, " že jste ten jistej referent :
ultu poniklovejma ocelovejma	<b>coulkama</b> /coulkama/X@-----	, aby je někdo neukrad . Z p
tálka povstala a přikryla si pleše	<b>sombrérama</b> /sombrérama/X@-----	, trumpeták vrazil do korpus
lustr s barevnejma skleněnejma	<b>cingrdlátkama</b> /cingrdlátkama/X@-----	, ale světlo místo něj obstará
vná , bíle natřená , se šedivejma	<b>žbrdlinkama</b> /žbrdlinkama/X@-----	. " Zapalte si . . . " Popelníč
a maníkem s modravě černejma	<b>fousiskama</b> /fousiskama/X@-----	pod nosem , si podobný star
braz muže na Měsíci od Arthura	<b>Zackhama</b> /Zackhama/X@-----	. Charlie Budd sledoval jehc
tion Lee ! Roztrhám ho na kusy	<b>plazmoidama</b> /plazmoidama/X@-----	! Zemře ! Nikdo mě nezasta
vyž se například nějakej študák s	<b>uhrama</b> /uhrama/X@-----	po celým ksichtě zamiluje d

Počet výskytů: 824  
> Query : "X.\*"  
> P-filter: "X.\*"

Zobrazeno: 1+50/824 (6%) Řádek: 9 Vybráno: 3

N.\*

Frekvenční distribuce

Výsledek frekvenční distribuce

Limit: 0      Řádků k zobrazení: 1721

lc: 0<KWIC       zobrazit       součet       schovat

lemma: 0<KWIC       zobrazit       součet       schovat

chlapama	chlap	49
rama	Rama	48
botama	bota	48
nožkama	nožka	47
ručičkama	ručička	46
wyndhama	Výndham	46
nožičkama	nožička	44
gilliama	Gilliam	44
brama	brama	43
vlasama	vlas	42
ručkama	ručka	42
drama	Drama	42
prstama	prst	39
olbrama	Olbram	38
kamarádama	kamarád	36
rokama	rok	36
parama	Paramo	35
horama	hora	35
dinama	Dinamo	33
hlavama	hlava	32
panama	Pan	30
atacama	Atakama	30

Zavřít      Uložit

# Závěr

- Při vyhledávání substandardních tvarů je třeba nespoléhat na značkování (1 pozice – X, 15 pozice – problematická).
- Problémy spojené se zkoušením „guesserů“ ve značkování SYN2005.
- Ruční práce, několikeré ověřování.

# Posuny v kodifikaci samohláskových alternací ve kmeni (vzor žena) – úkol na 8. 12. 2011

- K samohláskovým alternacím dochází u některých substantiv skloňovaných podle vzoru *žena*, a to a) obligatorně v gen. pl. a b) fakultativně v některých dalších pádech (*tráva – trav, síla – sil, míra – měř, houba – hub, ...*)
- 1. Popiš postup, jímž lze z korpusu SYN2010 získat kandidáty (feminina skloňovaná podle vzoru *žena*) na substantiva, která takovou alternací mohou mít.
- 2. Vyber z nich slova (nejméně 3 ke každé alternaci), o nichž víš, že takovou alternací mají. V případě nejistoty ověř svůj jazykový cit ve výkladovém slovníku (např. SSČ).
- 3. Sleduj v korpusu všechny tvary vybraných lemmat a uveď, u kterých tvarů se fakultativně vyskytuje stejná alternace jako v genitivu plurálu.