

CJBB75 Základy využití korpusu pro praxi

st. 12.30-14.05

G02

Výuka JS 2012

- 22. 2. Organizační pokyny, základy – *Bonito*.
- 29. 2. ÚKOL 1 : ,Vyplnit žádost, získat heslo.
- 7. 3. Různé korpusy a rozdíly v anotačních schématech (tokenizace, lemmatizace, tagging, disambiguace, tagset).
- 14. 3. ÚKOL 2 :
- 21.3. Jak lze v korpusech hledat doklady pro výzkum hláskových alternací ?
- 28. 3. ÚKOL 3
- 4. 4. Písemka
- 11. 4. ÚKOL 4
- 18. 4. Je možné získat z korpusů informace o slootvorbě ?
- 25. 4. ÚKOL 5 (Jak využívat ČNK : 17/5, 19/15)
- 2. 5. Vyhledávání složitějších konstrukcí
- 9. 5. ÚKOL 6
- 16.5. Zpt. test

22.2. Organizační pokyny

- a) Na [www http://ucnk.ff.cuni.cz/](http://ucnk.ff.cuni.cz/) - Dohody a registrace si opatřete uživatelská přístupová práva k programu Bonito.
- b) Na [www http://ucnk.ff.cuni.cz/](http://ucnk.ff.cuni.cz/) - Manuál a instalace - Instalace si stáhnete na svůj domácí počítač program Bonito.
- c) Na [www http://ucnk.ff.cuni.cz/](http://ucnk.ff.cuni.cz/) - Dostupné korpusy si nastudujte informace o korpusech CNK.
- d) Na [www http://ucnk.ff.cuni.cz/](http://ucnk.ff.cuni.cz/) - Manuál a instalace si nastudujte informace o tom, jak zadávat dotazy v programu Bonito.

Bonito

Manuál korpusového manažeru Bonito

Marie Kopřivová
Jan Koček

The screenshot shows the Bonito corpus manager interface. The window title is "Bonito". The menu bar includes "Manažer", "Korpus", "Dotaz", "Konkordance", "Zobrazení", "Výběr", and "Nápověda". The search bar contains "[lemma='korpus']" and "jméno: korpus (lemma)". The search results are displayed in a table with columns for opus, text, and the word "korpus". The word "korpus" is highlighted in red in the original image. The interface also shows a search bar, a list of results, and a detailed view of the selected result.

Annotations:

- dotazový řádek
- výběr korpusu
- pojmenování dotazu
- konkordanční řádek
- označený konkordanční řádek
- vyhledaný výraz - KWIC (key word in context)
- konkordanční seznam
- kód jednoznačně identifikující text
- rozšíření kontextu vyhledaného výrazu
- stavový řádek

Menu Korpus

- Souhrnné informace
- Implicitní atribut

Základní vyhledávání v korpusu

- **Obsah:**

Vyhledávání tvaru slova nebo slovního spojení

Vyhledávání podle atributu **lc** (lowercase)

Vyhledávání podle atributu **lemma**

Vyhledávání podle atributu **tag**
(morfologická značka)

Nastavení implicitního atributu

Hledání v rámci jedné věty

Menu Konkordance

- Zpět k předchozí
- Vpřed k následující
- Statistika – Frekvenční distribuce

Menu Zobrazení

- Zdroje
- Atributy
- Kontext
- Rozsah

Zobrazení názvu díla a autora

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 filter "P filtr" and a dropdown menu for "jméno:" with "syn" selected. Below the search bar are navigation buttons: "První", "Poslední", "Od -1", "Pozic", "od", "< KWIC", "Do 1", "Pozic", "od", "KWIC >". The main display area shows a list of search results on the left and a preview of the text on the right. The results list includes:

- opus.nazev=Ikaros, ročník 1999,opus.autor=Y
- opus.nazev=Naše řeč, ročník 1997,opus.autor=Y
- opus.nazev=Slovo a slovesnost, ročník 1997,opus.autor=Y
- opus.nazev=Slovo a slovesnost, ročník 1997,opus.autor=Y
- opus.nazev=Efeméra - ostatní,opus.autor=Y
- opus.nazev=Lidové noviny - Magazín, č. 49/1997,opus.autor=Y
- opus.nazev=Právo, 4. 12. 2003,opus.autor=Y
- opus.nazev=Deníky Bohemia, 24. 11. 2007,opus.autor=Y
- opus.nazev=Letohradský měsíčník, ročník 1998,opus.autor=Y

The preview text on the right shows a snippet of text with the word "použití" highlighted in red. The text reads: "dosud nevídaného jazykového materiálu bude konci věty ; v této pozici je ho v našem tutéž cenu . Malou cenu bude mít použití tomto jazyce pro vytvoření diversifikovaného dosud nevídaného jazykového materiálu bude přibližně od 6 do 25 korun) . K ovinutí včetně autobusového terminálu . Stavbě jako loni v lednu , " svěřila se keramička . Na adresou http : // UCNK . FF . CUNI . CZ/CNC ."

At the bottom of the window, there is a status bar that reads "Zobrazeno: vše/9 Řádek: 1".

Domácí úkol na 29. 2.

- a) Na [www http://ucnk.ff.cuni.cz/](http://ucnk.ff.cuni.cz/) - Dohody a registrace si opatřete uživatelská přístupová práva k programu Bonito.
- b) Na [www http://ucnk.ff.cuni.cz/](http://ucnk.ff.cuni.cz/) - Manuál a instalace - Instalace si stáhněte na svůj domácí počítač program Bonito.
- c) Na [www http://ucnk.ff.cuni.cz/](http://ucnk.ff.cuni.cz/) - Dostupné korpusy si nastudujte informace o korpusech CNK.
- d) Na [www http://ucnk.ff.cuni.cz/](http://ucnk.ff.cuni.cz/) - Manuál a instalace si nastudujte informace o tom, jak zadávat dotazy v programu Bonito.