

CJBB75 Základy využití korpusu pro praxi

čt. 12.30-14.05

G02

29.9. 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 "Nový dotaz" and "[lemma="korpus"]". The search results table has columns for opus, text, and the word "korpus" in various forms. The selected row is highlighted in blue. The status bar at the bottom shows "Zobrazeno: 1+100/276 (36%) Řádek: 7 Vybráno: 1".

opus	text	word
opus=smrt	straně otevřeně a zvedal svůj	korpus
opus=smrt	figuře a postel by unesla takové	korpusy
opus=noc	žihací plamen a hladí jím kovový	korpus
opus=eco	veškerý pozemský a sublanární	korpus
opus=zbabelci	tancující a slunce se třpytilo v jejím	korpusu
opus=zbabelci	mluvil jsem z jeho pozlaceného	korpusu
opus=nylon	piana , pohlédl smutně na lesklý	korpus
opus=nylon	nechyběla , ale ani hlas zclaceného	korpusu
opus=nylon	spolu s ostatními pozdvihoval	korpus
opus=clavis98	vyzkoušet práci s malým textovým	korpusem
opus=lobkow	nové doby . Červeně vypořstovaný	korpus
opus=ikaros99	jsem se jeli podívat na anglické	korpusy

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

- Uložení
- Zpět k předchozí
- Vpřed k následující
- Statistiky – Frekvenční distribuce

Menu Zobrazení

- Zdroje
- Atributy
- Kontext
- Rozsah

Označení a výběr řádků

- Všechny
- Žádné
- Inverze

Vyhledej v korpusu SYN

- Výskyty substantiva **korpus**.
- Za použití P-filtru vyberte řádky, kdy se vyskytuje spolu se slovy z oblasti **užít/sloužit**.
- Označ řádky, na nichž jde o jazykový **korpus**.
- Zobraz zdroje, z nichž pocházejí nalezené korpusové doklady.
- Odstraň řádky, na nichž nejde o jazykový **korpus**.

P-filtr

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 "P filtr" and ".*užit.*". The search results are displayed in a list format, with the word "korpus" highlighted in red. The search results are as follows:

Text	Highlighted Word	Context
dveře na své straně otevřené a zvedal svůj	korpus	, a proto vystoupila i ona . K Wolfovi
ulité jeho figuře a postel by unesla takové	korpusy	dva . Dalo by se předpokládat , že i do
jí byl cizí , stejně jako jakýkoli jiný	korpus	. " Na jednu věc bych málem zapomněl ,
si pleše sombrérama , trumpeták vrazil do	korpusu	svýho pistónu štrajchmúte (jak říká náš
AHODOVY KOLAČ Potřebujeme : piškotový	korpus	300 g tvarohu 300 g jahod 300 g cukru šťáva
nech odležet . 2 . Na talíř polož piškotový	korpus	a pokrop ho šťávou z jahod . 3 . Potří
korpus a pokrop ho šťávou z jahod . 3 . Potří	korpus	tvarohem , oslazeným 200 g zbylého cukru
upravuje žihací plamen a hladí jím kovový	korpus	sundaný z desky kříže , naslouchám a napadá
co říkám ? veškerý pozemský a sublunární	korpus	vyhladovělých a žíznivých stolovníků)
astupování do paternosteru . Přetlakový kulový	korpus	tvořil hlavu celé té křehké , šípovité
nerez oceli , který spojoval vadný motor s	korpusem	tlačné sondy . " Dávejte pozor , " varoval
čeru 144 mm . Vedle desetimetrového průměru	korpusu	tlačné sondy vypadal jako smítko prachu
králů . Když je navíc posílil expediční	korpus	z Redanie a partyzánské oddíly královny
kdy do něho vkročil kaedwenský okupační	korpus	. " " Žádám , " ozval se Shilard Fitz-Oesterlen

Počet výskytů: 1521
> Query : "korpus"

Zobrazeno: 1+50/1521 (3%) Řádek: 1

Výsledky

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 "P filtr" and "jméno: syn". The search parameters are set to "První" (First), "Od -1", "Pozic" (Position), "od", "< KWIC", "Do 1", "Pozic", "od", and "KWIC >".

The search results are displayed in a table with three columns: the original text, the word in red, and its definition. The text on the left is partially obscured by "or=Y" and "=Y".

dosud nevidaného jazykového materiálu bude tutéž cenu . Malou cenu bude mít použití	korpus	sloužit nejen lingvistům , ale i širokému
tomto jazyce pro vytvoření diversifikovaného dosud nevidaného jazykového materiálu bude	korpusu	například pro potvrzení introspektivně
přibližně od 6 do 25 korun) . K ovinutí	korpusu	použitelného ve všech stylech mluvených
včetně autobusového terminálu . Stavbě jako	korpus	sloužit nejen lingvistům , ale i širokému
loni v lednu , " svěřila se keramička . Na	korpus	poslouží pruhy textilií dlouhé asi tři
adresou http : // UCNK . FF . CUNI . CZ/CNC .	korpus	poslouží jedna z budov depa Hostivař .
	korpus	zužitkovala rám starožitného zrcadla ,
	Korpus	slouží k vytvoření nového výkladového slovníku

Zobrazeno: vše/8 Řádek: 1

Označ řádky, na nichž jde o jazykový korpus.

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 search filter section with "P filtr" and a dropdown menu. To the right, there is a "jméno:" field and a "syn" dropdown. Below this, there are navigation buttons: "První", "Poslední", "Od -1", "Pozic", "od", "< KWIC", "Do 1", "Pozic", "od", "KWIC >". The main area displays a list of text rows. The first row is highlighted in light blue. The text in the first row is: "dosud nevidaného jazykového materiálu bude tutéž cenu . Malou cenu bude mít **použití** **korpus** **sloužit** nejen lingvistům , ale i širokému publiku . Například pro potvrzení introspektivně použitelného ve všech stylech mluveného jazyka .
tomto jazyce pro vytvoření diversifikovaného jazykového materiálu bude **korpus** **sloužit** nejen lingvistům , ale i širokému publiku .
přibližně od 6 do 25 korun) . K ovinutí **korpus** **poslouží** pruhy textilií dlouhé asi tři metry .
včetně autobusového terminálu . Stavbě jako **korpus** **poslouží** jedna z budov depa Hostivařské stanice .
loni v lednu , " svěřila se keramička . Na **korpus** **zžitkovala** rám starožitného zrcadla .
adresou [http : // UCNK . FF . CUNI . CZ / CNC .](http://ucnk.ff.cuni.cz/cnc) **Korpus** **slouží** k vytvoření nového výkladového materiálu .

At the bottom of the window, a status bar reads: "Zobrazeno: vše/8 Řádek: 8 Vybráno: 5".

Odstraň řádky, na nichž nejde o jazykový korpus.

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 are several input fields: "P filtr" with a dropdown arrow, "jméno:" with a dropdown arrow, and "syn" with a dropdown arrow. Below these are navigation controls: "První" (selected), "Poslední", "Od" with a dropdown showing "-1", "Pozic" with a dropdown arrow, "od", "< KWIC" with a dropdown arrow, "Do" with a dropdown showing "1", "Pozic" with a dropdown arrow, "od", and "KWIC >" with a dropdown arrow. The main area contains a list of text rows. The first row is "dosud nevidaného jazykového materiálu bude", the second is "tutéž cenu . Malou cenu bude mít použití", the third is "tomto jazyce pro vytvoření diversifikovaného", the fourth is "dosud nevidaného jazykového materiálu bude", the fifth is "přibližně od 6 do 25 korun) . K ovinutí", the sixth is "včetně autobusového terminálu . Stavbě jako", the seventh is "loni v lednu , " svěřila se keramička . Na", and the eighth is "adresou http : // UCNK . FF . CUNI . CZ / CNC .". To the right of each row, there are labels: "korpus", "korpusu", "korpusu", "korpus", "korpusu", "korpus", "korpus", and "Korpus". The fourth, fifth, sixth, and seventh rows are highlighted in light blue. The status bar at the bottom reads "Zobrazeno: vše/8 Řádek: 8 Vybráno: 3".

Zobraz zdroj, z nichž pocházejí nalezené korpusové doklady.

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 filter section with a dropdown menu labeled 'P filtr' and a text input field containing 'jméno:'. To the right of the input field is a dropdown menu with 'syn' selected. Below the search filter, there are navigation buttons: 'První', 'Poslední', 'Od -1', 'Pozic', 'od', '< KWIC', 'Do 1', 'Pozic', 'od', and 'KWIC >'. The main display area shows a list of search results on the left and a preview of the text on the right. The search 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'. Below the main display area, there is a small text box containing the following information:

Počet výskytů: 9
> Query : "korpus"
> P-filter: ".*už[i]t.*"

Výsledky

The screenshot shows the Bonito software interface with a search query for "korpus". The results are displayed in a table format with three columns: the original text snippet, the word "korpus" in red, and its context in red. The search parameters are: Query: "korpus", P-filter: ".*užit.*", Deleted: 3. The status bar at the bottom indicates "Zobrazeno: vše/5 Řádek: 1".

Text	korpus	Context
dosud nevidaného jazykového materiálu bude tutéž cenu . Malou cenu bude mít použít	korpus	sloužit nejen lingvistům , ale i širokému například pro potvrzení introspektivně
tomto jazyce pro vytvoření diversifikovaného dosud nevidaného jazykového materiálu bude	korpusu	použitelného ve všech stylech mluvených
adresou http : // UCNK . FF . CUNI . CZ/CNC .	Korpus	slouží k vytvoření nového výkladového slovníku

> Query : "korpus"
> P-filter: ".*užit.*"
> Deleted : 3

Zobrazeno: vše/5 Řádek: 1

Domácí práce na 7. 10.

- Vyhledejte v korpusu SYN2010 slova, kde se po „tvrdé souhlásce“ píše „měkké i“ (kino, rigidní, ...).
- Vytvořte frekvenční seznam.
- Je vhodné upozornit žáky 2. třídy v době, kdy se pravopisné pravidlo probírá, na takové případy ?
- Pokud ano, mohou být korpusy vhodným zdrojem vhodných příkladů ?