

# Projekt Bulharsko-české slovníkové databáze

Albena Rangelova, Ústav pro jazyk český AV ČR, v. v. i., oddělení současné lexikologie a lexikografie  
Hedvika Sixtová, Slovanský ústav AV ČR, v. v. i., oddělení slavistické lingvistiky a lexikografie

**Potřeba novějších** česko-bulharských a zejména bulharsko-českých slovníků, které by usnadnily mezijazykovou komunikaci a podpořily výuku češtiny a bulharštiny jako cizích jazyků, je stále naléhavá, zejména po vstupu Bulharské republiky do Evropské unie. Nejnovější překladový česko-bulharský slovník akademického typu (Ivančev 2002) byl záhy po svém vydání rozebrán, připravuje se jeho druhé vydání. Tvorba adekvátního protějšku bulharsko-českého je úkol nadmíru těžký, a to finančně, personálně i organizačně. **Poslední obsáhlejší bulharsko-český slovník vyšel v roce 1959** (autor K. Hora).

## Institucionální kontext

### Úsek vícejazyčné lexikografie

– konfrontační výzkum lexika a tvorba jinojazyčno-českých databází a slovníků (vybrané slovanské jazyky – slovenština, slovinština, bulharština);  
– návaznost na dlouholetou spolupráci Ústavu pro jazyk český AV ČR, v. v. i., s akademickými pracovišti ze Slovenska, Bulharska i Slovinska;  
– úspěšné ukončení mezinárodního projektu *LINT – Language in Tourism* (www.travellang.org), jeden z výstupů – čtyřjazyčný slovník lexikálního minima pro potřeby cestovního ruchu (Rangelova – Tichá 2008).

## Projekty

### Slovinsko-český slovník

– praktický kódovací slovinsko-český slovník (program TshwaneLex).

### Slovensko-česká lexikální databáze

– diferenční slovníkové zpracování nejnovějšího stavu slovní zásoby obou jazyků;  
– publikační technologie wiki (kolaborativní wiki-engine);  
– grantový projekt *Konfrontační popis současného slovenského a českého lexika (systémové vztahy a komunikační koexistence)* pro léta 2011–2014.

### Slovník zrádných slov v bulharštině a češtině

– studijní slovník s paralelním zpracováním typu výkladového slovníku pro bulharského uživatele.

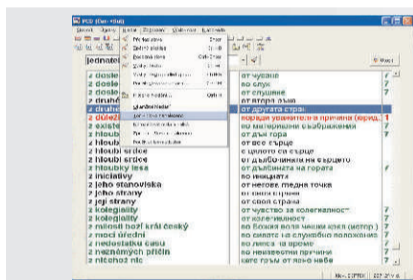
## Publikační kontext

### Současný trend: tvorba krátkých dvojjazyčných příruček

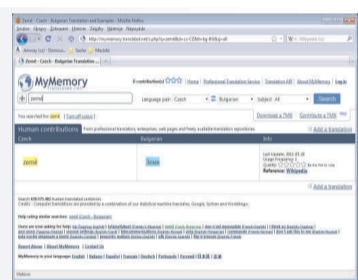
– méně náročné personálně i finančně, publikování v historicky krátkém čase;  
– specializované příručky, které slouží specifickým cílům, obvykle studiu: Rangelova – Tichá 2008, Miceva 2009, Bačvarov a kol. 2011;  
– konverzační příručky pro širokou veřejnost: Raev – Raevová 1998, Zidarov-Kounová 2006;  
– krátké všeobecné publikace: malý rozsah, omezená informace o užití lexikálních jednotek: Bublová – Videnov 1987, Uchytíl 2009.

### Elektronické překladové slovníky na CD a na internetu (volný či placený přístup).

– pro uživatele pohodlná změna směru vyhledávání;  
– možnost doplňování uživateli;  
– z lingvistického hlediska nevyhovující.



Překladový slovník PCtranslator



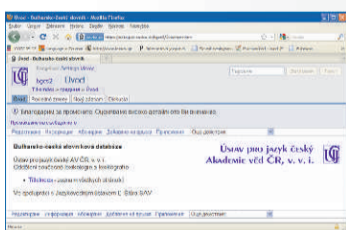
Volně přístupný portál MyMemory doplňovaný uživateli

## Bulharsko-česká slovníková databáze – základní koncepční rysy

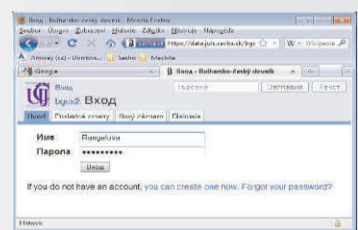
**Autorský kolektiv:** A. Rangelova, H. Sixtová, N. Kalajdzieva, programátor R. Garabik

### Záměr

– poskytnout novou, vědecky podloženou příručku obsahující základní slovníkové zpracování bulharsko-českého směru, které může být v budoucnu doplňováno a rozšiřováno;  
– **rozsah:** krátká slovníková databáze, 12–15 tisíc jednotek;  
– **způsob zveřejnění:** primárně elektronická podoba, využití publikační technologie wiki (spolupráce s Jazykovedným ústavem L. Štúra SAV);  
– **uživatel:** je určena nejširšímu okruhu veřejnosti, zejména české, se zvláštním zřetelem k běžnému, neprofesnímu uživateli jazyka, pro běžný společenský a soukromý styk;  
– **metajazyk slovníku: čeština;**  
– **cíle:** navázat na tradici česko-bulharských a bulharsko-českých slovníků; zaplnit mezeru ve směru bulharština – čeština; překonat základní omezení krátkých slovníkových příruček, tj. nedostatek informací potřebných pro správnou interpretaci významu a užití jednotek;  
– **časový horizont:** při stávajícím personálním obsazení a při optimálním vývoji události 3–4 roky.



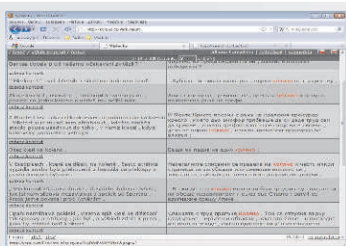
Titulní stránka a vstup do programu



### Vytvoření hesláře

**Výchozí heslář** – seznam hesel ze studijní příručky (čtyřjazyčného slovníku lexikálního minima pro potřeby cestovního ruchu) projektu *LINT – Language in Tourism*;

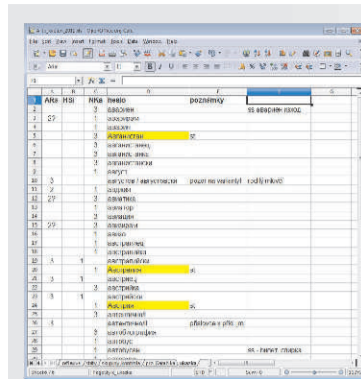
– spolupráce s kolegy z Ústavu pro bulharský jazyk BAV;  
– konfrontace s hesláři jiných bulharsko-jinojazyčných slovníků (Gaberoff 2000, Uchytíl 2009);  
– základní referenční zdroje: *Balgarski talkoven rečnik* (Popov 1994) a *Rečnik na novite dum i značenja v bulgarskija ezik* (Perniška a kol. 2001);  
– pro ověřování pravopisu a variantnosti: *Pravopisen rečnik na bŭlgarskija ezik 2002* a *Internetová jazyková příručka ÚJČ AV ČR, v. v. i.* (http://prirucka.ujc.cas.cz/);  
– materiálové referenční zdroje: *Intercorp* – projekt paralelních korpusů Filozofické fakulty UK v Praze (http://ucnk.ff.cuni.cz/intercorp/) a *Bulharský národní korpus (Български национален корпус – http://www.ibl.bas.bg/BGNC\_bg.htm)*;  
– kontrola a doplňování dle lexikálněsemantických skupin, např. rostliny, živočichové, jídlo, nápoje, profese, státy, jazyky atp.;  
– kontrola a doplňování dle slovních druhů (zejména tzv. malé slovní druhy – předložky, spojky, citoslovce, ale též číslovky a zájmena).



Ukázka vyhledávání v korpusu Intercorp v paralelních korpusech českém a bulharském (směr bulharsko-český)



Ukázka vyhledávání v Bulharském národním korpusu



Pracovní záznamy výběru jednotek do hesláře

### Zásady výběru jednotek a jejich významů

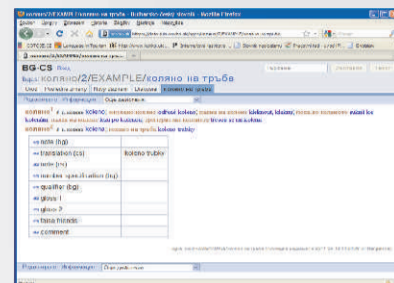
– **záměr** prezentovat základní slovní zásobu bulharštiny (lexikální minimum/lexikální jádro) se zřetelem k novým společenským podmínkám a potřebám jazykové a mezijazykové komunikace dnes (běžná slovní zásoba, zaměření na běžný společenský a soukromý život);  
– **základní kritérium výběru:** rozšířenost, užitelnost jednotky v běžné komunikaci;  
– **pomocná kritéria:** systémovost zpracování a diferenční princip.  
Práce v tabulkovém procesoru *Excel*, aktuální počet hesel 13 119 (při rozepsání homonymních dvojic a polysémních n-tic)

## Mikrostruktura

### Původní model

**Tabulková organizace:** adaptace výchozího modelu *Slovensko-české slovníkové databáze*

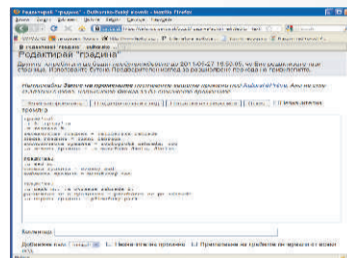
– položka databáze – lexie či vícelslovná lexikální jednotka;  
– jednotlivé významy jako samostatné stránky;  
– každý příklad jako samostatná vnořená stránka.



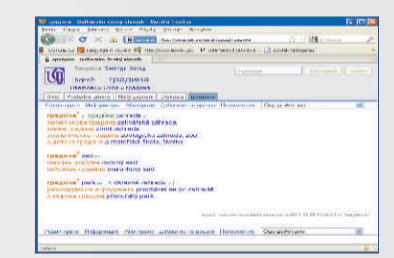
Náhled vícevýznamového hesla v tabulkovém zpracování

### Současný model

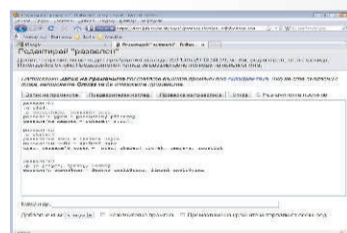
**Lineární vkládání dat** s předdefinovanými soubory informací;  
– vnořené seznamy zkratkou pro metajazykovou informaci (gramatika, oborové a stylistické kvalifikátory);  
– funkční kombinace znaků pro grafické ztvárnění heslové statě;  
– **editace** – využití jednoduché syntaxe a funkčních zkratkou pro různá grafická ztvárnění;  
– **náhled hesla** – totožný pro uživatele a zpracovatele.



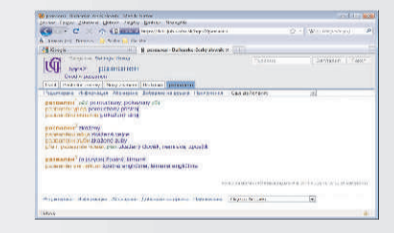
Editace podstatného jména



Náhled zpracovaného podstatného jména



Editace přídavného jména

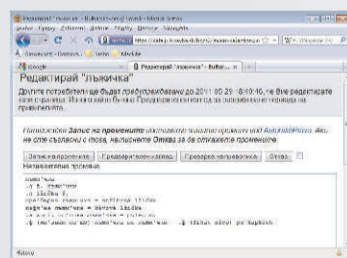


Náhled zpracovaného přídavného jména

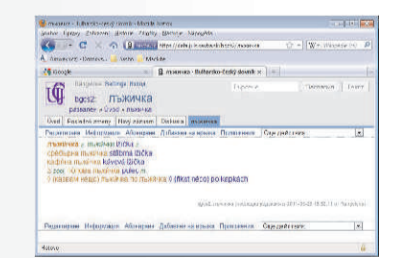
**Vícelslovné jednotky** (slosovly a frazémy) budou prezentovány ve své syntagmatické posloupnosti, tj. bez inverze a bez zkracování, s grafickým označením:

– slosovly znakem Δ, např. *Δ кръла маса Δ kulatý stůl, Δ ски бягане Δ běh na lyžích* aj.;

– frazémy znakem ϕ, např. (o člověku) ϕ *голяма клечка ϕ velké zvíře*; (o člověku) ϕ *от голямото добрутро kdo se chová nadřazeně* apod.



Vícelslovné jednotky - editace



Vícelslovné jednotky - náhled

## Aktuální úkoly

Zpracování vzorových hesel pro ověření a zpřesnění zásad zpracování jednotlivých slovních druhů;  
– vývoj softwarového nástroje (programátorské práce);  
– zabezpečení doplňkových finančních zdrojů.

## Literatura:

BAČVAROV A KOL. 2011: Янко Бъчваров и колектив, Кратък чешко-български речник на неknижовната лексика. „Ето“, София 2011.  
BUBLOVÁ – VIDENOV 1987: M. Bublova, M. Videnov, Česko-bulharský a bulharsko-český slovník na cesty. SPN, Praha 1987.  
GABEROFF 2000: Българско-английски речник. „Габеров“, София 2000.  
HORA 1959: K. Hora, Bulharsko-český slovník. ČSAV, Praha 1959.  
Internetová jazyková příručka ÚJČ AV ČR, v. v. i., http://prirucka.ujc.cas.cz/  
IVANČEV 2002: Св. Иванчев (ред.), Чешко-български речник на два тома под редакцията на Св. Иванчев. „Труд“ & „Прозорек“, София 2002 (dotisk 2005).  
MICEVA 2009: В. Мицева, Чешко-български речник на съкратените наименования. ИК „Стигмати“, София 2009.  
MyMemory, http://mymemory.translated.net/  
Нов правописен речник на българския език. „Хейзъл“, София 2002.  
PCtranslator, LangSoft, s. r. o., Korytná, http://www.langsoft.cz/translat.htm  
PERNIŠKA A KOL. 2001: E. Perniška, D. Благоева, С. Колковска, Речник на новите думи и значения в българския език. „Наука и изкуство“, София 2001.  
POPOV 1994: Д. Попов (ред.), Български тълковен речник. „Наука и изкуство“, София 1994.  
RAEV – RAEVŌVÁ 1998: I. Raev, N. Raevová, Bulharština. Praktický jazykový průvodce. RO-TO-M<sup>®</sup>, Bučovice 1998.  
RANGELOVA – TICHÁ 2008: A. Rangelova, Z. Tichá (eds.), Vícejazyčný studijní slovník pro potřeby cestovního ruchu. Nakl. Valentin Trajanov, Sofia 2008.  
UCHYTIL, K.: Bulharsko-český/Česko-bulharský slovník. Český klub, Praha 2009.  
ZIDAROV-KOUNOVÁ 2006: S. Zidarov-Kounová, Česko-bulharská konverzace. Montanex, Ostrava 2006.