

# Úvod do korpusové lingvistiky 8



## VYUŽITÍ KORPUSŮ PRO LINGVISTICKÁ BÁDÁNÍ

# Konference



- Konference **Korpus jako zdroj dat o češtině** Brno 2004 (Karlík, P.: Korpus jako zdroj dat o češtině . MU : Brno. 2004.)
- Konference **Gramatika a korpus** /v Praze (2005), Liblicích (2007) a Mannheimu (2009), Praze (2012)/.
- Konference **Čeština v mluveném korpusu** Praha 2007
- Konference **Interkorp** Praha 2009
- Konference **Korpusová lingvistika** Praha 2011, Praha 2014

Časopis (od r. 2010 2 krát ročně)

<http://nlp.fi.muni.cz/kga/>



# ***KORPUS – GRAMATIKA – AXIOLOGIE***

Časopis pro korpusový výzkum a hodnocení jazyka  
Journal for corpus research and evaluation of language

Poslání časopisu / Mission of the journal  
objednávka / order

# Gramatika



- Štícha, F. (ed.): Možnosti a meze české gramatiky. Praha : Academia. 2006.
- Štícha, F. (ed.): Kapitoly z české gramatiky. Praha : Academia. 2011.
- Štícha, F. (ed.): Akademická gramatika spisovné češtiny. Praha : Academia. 2013.
- Cvrček, V. a kol.: Mluvnice současné češtiny. Praha : Karolinum. 2010.
- Panevová, J. a kol.: ???
- <http://lectures.ms.mff.cuni.cz/view.php?rec=89>

# Slovníky



- Čermák, F. - Křen, M. (eds): Frekvenční slovník češtiny. Nakladatelství Lidové noviny, Praha 2004.
- Čermák, F. (ed.): Frekvenční slovník mluvené češtiny. Karolinum, Praha 2007.
- Čermák, F. (ed.): Slovník Karla Čapka. Nakladatelství Lidové noviny, Praha 2007.
- Čermák, F. - Cvrček, V.: Slovník Bohumila Hrabala. Nakladatelství Lidové noviny, Praha 2009.
- Čermák, F., Cvrček, V., Schmiedtová, V. (eds): Slovník komunistické totality. Nakladatelství Lidové noviny, Praha 2010.

# Slovníky



- Čermák, F., Křen, M. (eds): A Frequency Dictionary of Czech: Core Vocabulary for Learners. Routledge, London 2011.
- Cvrček, V., Cvrčková Porkertová, L.: Velký slovník rýmů. Nakladatelství Lidové noviny, Praha 2011.
- Schmiedtová, V.: Malý slovník reálií komunistické totality. Nakladatelství Lidové noviny. Praha 2013.
- Čermák, F.: Základní slovník českých přísloví. Nakladatelství Lidové noviny. Praha 2013.
- Čermák, F.: Proverbs: Their Lexical and Semantic Features. University of Vermont, 2014

# Matematická a počítačová lingvistika - NLP - Korpusové nástroje



- ÚFAL (Ústav formální a aplikované lingvistiky MFF UK)
- ÚTKL (Ústav teoretické a počítačové lingvistiky FF UK)
- FI M (Centrum zpracování přirozeného jazyka FI MU)

# Spolupráce se Slovenskem



- **Slovenský národný korpus**, JÚLŠ SAV
- <http://korpus.juls.savba.sk/>
- Konferencie **SLOVKO** (2001, 2003, 2005, 2007, 2009, 2011, 2013)  
<http://korpus.juls.savba.sk/slovko.html>



# Žákovské korpusy/Learner corpora



- Žákovský korpus - **CzeSL-plain** (*Czech as a Second Language, plain* = bez anotace)
- **ciz** – přepisy písemných prací (esejů) nerodilých mluvčích, které vznikly v souvislosti s jazykovým vyučováním v kurzech různého druhu a úrovně;
- **kval** – odborné texty získané od nerodilých mluvčích studujících na českých vysokých školách v navazujícím magisterském či doktorském studiu;
- **rom** – přepisy školních písemných prací romských žáků z oblastí ohrožených sociálním vyloučením.

# Chyby žáků jako zdroj poznání



- Štindlová, B.: Žákovský korpus češtiny a evaluace jeho chybové anotace. Praha : FF UK. 2013.
- Karel Šebesta – Svatava Škodová a kolektiv: Čeština – cílový jazyk a korpusy. Liberec : TECHNICKÁ UNIVERZITA V LIBERCI Fakulta přírodovědně-humanitní a pedagogická Katedra českého jazyka a literatury.2012.

# Data driven learning (DDL)



- Využití korpusů pro výuku jazyka (Tim Jones, Universita v Birminghamu)
- <https://sites.google.com/site/elmethodologies/approaches/data-driven-learning>
- <http://iteslj.org/Articles/Krieger-Corpus.html>
- Konference **Teaching and Language Corpora (TaLC)**
- James E. Thomas, M.A. – FF MU (Department of English and American Studies Faculty of Arts Masaryk University)
- Mgr. Pavlína Vališová – studentka doktorského studia Ústavu českého jazyka FF MU