

Václav Cvrček

Institute of the Czech National Corpus
Faculty of Arts, Charles University
Panská 7, Prague 1, 110 00
Czech Republic, EU
<http://www.korpus.cz>

Private: Krkonošská 2
Prague 2, 120 00
e-mail: vaclav.cvrcek@ff.cuni.cz
phone: +420 221 619 846
ORCID: 0000-0003-3977-2393

Education

Ph.D. Mathematical and Corpus Linguistics, Charles University, 2008

Thesis: *Theory of Interventions and the Concept of Minimal Intervention*

Corpus-based sociolinguistic study of linguistic interventions into the language usage accompanied by an alternative proposal for a different construction of relationship between language, speakers and linguists.

Mgr. Czech language and Literature; Linguistics and Phonetics, Charles University, 2004

Thesis: *Development of Thought on Codification of Czech since 1945*

Description of the history of the Theory of The Cultivation of Language elaborated by the members of the Prague Linguistic Circle.

Work Experience

Vice-dean

Faculty of Arts,
Charles University
Vice-dean for Research and Ph.D. study programme

1/2017–2/2018
Prague

Deputy chair of department & Associate

Professor

Institute of the Czech National Corpus,
Charles University

Research interests: corpus linguistics, quantitative linguistics, corpus-assisted discourse studies

since 2/2016

Prague

Chair of department

Institute of the Czech National Corpus,
Charles University

2/2013–1/2016
Prague

Research Associate Professor

Institute of the Czech National Corpus,
Charles University

Research interests: corpus linguistics, quantitative linguistics, Czech morphology, collocations

since 2013
Prague

Fulbright Visiting Scholar, Lecturer

Department of Slavic Languages,
Brown University

Research interests: corpus-driven context analysis, keyword analysis

8/2011 – 5/2012
Providence, USA

Research Assistant Professor

Institute of the Czech National Corpus,
Charles University

Research interests: morphology, corpus linguistics, sociolinguistics, collocations, spoken language

since 2006
Prague

Publications

Monographs

- Fidler, M. – Cvrček, V. (eds): *Taming the Corpus. From Inflection and Lexis to Interpretation*. Springer. Cham 2018. ISBN 978-3-319-98016-4.
- Cvrček, V.: *Kvantitativní analýza kontextu* (Quantitative analysis of context). NLN, Praha 2013. ISBN 978-80-7422-264-1
- Cvrček, V. – Cvrčková Porkertová, L.: *Velký slovník rýmů* (Rhyming dictionary). NLN, Praha 2011. ISBN 978-80-7422-095-1.
- Cvrček, V. (ed.): *Mluvnice současně češtiny* (Grammar of Contemporary Czech). Karolinum. Praha 2010. ISBN 978-80-246-1743-5.
- Čermák, F. - Cvrček, V. - Schmiedtová, V. (eds): *Slovník komunistické totality* (Dictionary of Communist Regime). NLN, Praha 2010. ISBN 978-80-7422-060-9
- Bartoň, T. – Cvrček, V. – Čermák, F. – Jelínek, T. – Petkevič, V.: *Statistiky češtiny* (Statistics of Czech). Nakladatelství Lidové noviny. Praha 2009. ISBN 978-80-7106-5944.
- Čermák, F. – Cvrček, V. (eds): *Slovník Bohumila Hrabala* (Dictionary of Bohumil Hrabal). Nakladatelství Lidové noviny. Praha 2009. ISBN 978-80-7106-488-6.
- Cvrček, V.: *Regulace jazyka a Koncept minimální intervence* (Regulation of Language and The Concept of Minimal Intervention). Nakladatelství LN. Praha 2008. ISBN 978-80-7106-600-2.
- Čermák, F. (ed.): *Slovník Karla Čapka* (Dictionary of Karel Čapek). Nakladatelství Lidové noviny, Praha 2007. ISBN 978-80-7106-915-7.
- Cvrček, V.: *Teorie jazykové kultury po roce 1945* (Theory of the Cultivation of Language after 1945). Karolinum. Praha 2006. ISBN 80-246-1044-2.
- Gladkova, H. – Cvrček, V. (eds): *Sociální aspekty spisovných jazyků slovanských* (Social Aspects of Slavic Literary Languages). Acta Slavica et Baltica VI. FF UK. Praha 2006. ISBN 80-7308-155-5.

Papers

- Cvrček, V. – Fidler, M.: More than keywords: Discourse prominence analysis of the Russian Web portal Sputnik Czech Republic. In: Berrocal, M. – Salamurović, A. (eds): *Political Discourse in Central, Eastern and Balkan Europe*. John Benjamins. 2019, (p. 93–117).
- Cvrček, V. – Komrsková, Z. – Lukeš, D. – Poukarová, P. – Řehořková, A. – Zasina, A.J.: From extra- to intratextual characteristics: Charting the space of variation in Czech through MDA. *Corpus Linguistics and Linguistic Theory*. 2018. DOI: 10.1515/cllt-2018-0020
- Cvrček, V. – Fidler, M.: Introduction. In: Fidler, M. – Cvrček, V. (eds): *Taming the Corpus. From Inflection and Lexis to Interpretation*. Springer. Cham 2018. (p. 1–8).
- Fidler, M. – Cvrček V.: Going Beyond “Aboutness”: A Quantitative Analysis of Sputnik Czech Republic. In: Fidler, M. – Cvrček, V. (eds): *Taming the Corpus. From Inflection and Lexis to Interpretation*. Springer, Cham 2018, (p. 195–225).
- Cvrček, V. – Komrsková, Z. – Lukeš, D. – Poukarová, P. – Řehořková, A. – Zasina, A. J.: Variabilita češtiny: multidimenzionální analýza /Variability of Czech: a multidimensional analysis/. *Slovo a slovesnost* 79, 2018, (p. 293–321).
- Cvrček, V. – Komrsková, Z. – Lukeš, D.: Rozsah Registrované Variability Textů. In Kučera, D. et al. Výzkum CPACT: Komputační Psycholingvistická Analýza Českého Textu. České Budějovice: Pedagogická fakulta Jihočeské univerzity v Českých Budějovicích. 2018, (p. 153–172).
- Fidler, M. – Cvrček, V.: Keymorph analysis, or how morphosyntax informs discourse. *Corpus Linguistics and Linguistic Theory*. 2017.
- Cvrček, V.: Paradigmatické korpusové dotazy a moderní diachronie (Paradigmatic corpus queries and modern diachrony). In: Stluka, M., Škrabal, M. (eds): *Lifka a czban – Sborník příspěvků k 70. narozeninám prof. Karla Kučery*. NLN. Praha 2017, (s. 117–129).
- Cvrček, V. - Fidler, M.: Probing Aspectual Context with Keyword Analysis. In: Makarova, A. – Disckey, S. – Divjak, D. (eds): *Each Venture a New Beginning*. Slavica: Bloomington, Indiana, 2017, (p. 279–298).

Václav Cvrček

- Cvrček, V. Jazykový standard: teorie a praxe na začátku 21. století na příkladu mluvnice češtiny. In: M. Nekula – K. Šichová (eds): *Variety češtiny a čeština jako cizí jazyk*. Akropolis. Praha, 2017, (p. 23-56).
- Cvrček, V. – Čermáková, A. – Křen, M.: *Nová koncepce synchronních korpusů psané češtiny* (A new design of synchronic corpora of written Czech). *Slovo a slovesnost* 77, 2016, (p. 83–101).
- Cvrček, V. - Václavík, J.: *Jednoznačnost a kontext. Kvantitativní studie* (Ambiguity and context. A Quantitative study). *Korpus – gramatika – axiologie* 11, 2015, (p. 28–41).
- Cvrček, V. – Chlumská, L.: *Simplification in translated Czech: a new approach to type-token ratio*. *Russian linguistics* 39/3, 2015, (p. 309–325).
- Fidler, M. – Cvrček, V.: *A Data-Driven Analysis of Reader Viewpoints: Reconstructing the Historical Reader Using Keyword Analysis*. *Journal of Slavic Linguistics* 23(2), 2015, (p. 197–239).
- Cvrček, V.: *Kvantitativní určení lexikálního jádra jazyka* (Quantitative delimitation of core lexicon). *Časopis pro moderní filologii*, 96/1, 2014, (p. 9–26).
- Cvrček, V. – Kodýtek, V.: *Ke klasifikaci morfologických variant* (On the classification of morphological variants). *Slovo a slovesnost* 74/2013, (p. 139–145).
- Cvrček, V.: *Conditio sine qua non: východiska a možnosti při hledání konsenzu o jazykové regulaci* (Conditio sine qua non: points of departure and prospects in seeking consensus on language regulation). *Slovo a slovesnost* 73/2, 2012, (p. 103–134).
- Cvrček, V.: *Kontextová analýza a systém slovních druhů* (Context analysis and part of speech system). *Korpus – gramatika – axiologie* 3, 2011, (p. 6–32).
- Cvrček, V. - Kovářiková, D.: *Možnosti a meze korpusové lingvistiky* (Opportunities and Limitations of Corpus Linguistics). *Naše řeč* 94/3. 2011, (p. 113–133). ISSN 0027-8203.
- Šrajerová, D. – Kovářik, O. – Cvrček, V.: *Automatic Term Recognition Based on Data-mining Techniques*. CSIE 2009 vol. 4 (2009 WRI World Congress on Computer Science and Information Engineering). World Research Institutes, 2009, (p. 453–457).
- Cvrček, V.: *Morfologická paradigmatica a korpus* (Morphological Paradigms and Corpus). In: F. Štícha – M. Fried (eds): *Gramatika a korpus 2007/Grammar and Corpora 2007*. Academia 2008, (p. 151–160).
- Cvrček, V.: *Koncept minimální intervence* (The Concept of Minimal Intervention). *Slovo a slovesnost* 69. 2008, (p. 284–292).
- Cvrček, V.: *Rysy mluvenosti s použitím Bayesova teorému* (Discovering Features of Spoken Czech using Bayes' Theorem). In: M. Kopřivová – M. Wavlawičová (eds.): *Čeština v mluveném korpusu*. Nakladatelství Lidové noviny. Praha 2008, (p. 49–62).
- Cvrček, V.: *Ke způsobu vedení diskuze o spisovné češtině* (The Way We Discuss about Literary Czech). *Český jazyk a literatura* č. 1, 58/2007–2008, (p. 26–30).
- Sgall, P. – Cvrček, V.: *O názorové pluralitě a hledání konsenzu v lingvistice* (On the Plurality of Opinions and Consensus in Linguistics). *Naše řeč* 90/3. 2007, (p. 132–135).
- Cvrček, V.: *Spory o kodifikaci češtiny jako speciální případ lingvistické polemiky* (Problem of Codification as a Special Case of Polemics between Linguists). In: V. Lábus – K. Váňová (eds.): *Eurolingua & Eurolitteraria 2006*. TU v Liberci. Liberec 2007, (p. 88–107).
- Cvrček, V.: *Meze správnosti – klasifikace jazykových chyb* (Correctness in Language – Classification of Language Errors). *Slovo a smysl* 6. 2006, (p. 241–259).
- Cvrček, V.: *Metoda zjišťování kolokační platnosti frekventovaných bigramů pomocí ranku* (Rank Based Method of Detecting Collocations among Frequent Bigrams). In: *Kolokace. Studie z korpusové lingvistiky*. Nakladatelství Lidové noviny. Praha 2006, (p. 36–55).
- Cvrček, V.: *Kodifikační praxe* (Codificational Practice). In: H. Gladkova – V. Cvrček (eds.): *Sociální aspekty spisovných jazyků slovanských*. Acta Slavica et Baltica VI. FF UK. Praha 2006. (p. 16–35).
- Cvrček, V.: *Za ještě tvrdší kodifikační diktát?* (Do We Want the Codification to Be Even More Prescriptive?). *Naše řeč* 89/1. 2006, (p. 26–29).
- Cvrček, V.: *Spisovnost a její zdroje* (Literariness: the Concept and its Criteria). *Slovo a slovesnost* 67. 2006, (p. 46–60).

Selected presentations, conferences & invited talks

- Cvrček, V. – Komrsková, Z.: *Register variation in Czech: A multi-dimensional approach*. Summer school of linguistics. Kroměříž. August 14, 2019.
- Cvrček, V. – Fidler, M.: „*Up close and personal vs. birds-eye view*“ of discourse: a corpus study of perspective using Czech data. ICLC15 - International Cognitive Linguistics Conference, Nishinomya. Japan, August 2019.
- Cvrček, V. – Lukeš, D.: *Billions of words for the price of... lack of diversity: Comparing web-crawled and traditional corpora*. Corpus Approaches to Discourse Analysis. Valencia. Spain, May 2019.
- Cvrček, V.: *Digital Humanities approaches to language variability and their impact on language teaching*. Brown University. March 2019.
- Cvrček, V. – Komrsková, Z.: *MDA of Czech – resources, methods and innovations; Applications for and of MDA* (2 talks with Z. Komrsková). Variation in Language Corpora (International Autumn School). Humboldt-Universität zu Berlin. 25-27. October 2018.
- Fidler, M. – Cvrček, V.: *Power and artificial language: Václav Havel's Memorandum*. AATSEEL Conference. Washington DC. February 2018.
- Cvrček, V. – Komrsková, Z. – Lukeš, D. – Poukarová, P. – Řehořková, A. – Zasina, A. J.: *Multi-Dimensional Analysis of Czech. Pilot Study*. Coprus Linguistics 2017. Birmingham. 2017.
- Cvrček, V.: *Czech National Corpus: data and resources*. Slavisches Seminar Universität Tübingen. Germany. 5/2016.
- Cvrček, V. – Fidler, M.: *Workshop on Quantitative Text Analysis for the Humanities and Social Sciences*. Brown University, Providence, USA. April 8–9, 2016.
- Fidler, M. – Cvrček, V.: *An Alternative Viewpoint or Agitprop? A Corpus-based Analysis of the News-Opinion portal Sputnik.cz*. AATSEEL, Austin USA. 1/2016.
- Cvrček, V.: *Jazykový standard: Teorie a praxe* (Language standard. Theory and practice). Tschechisch auf Deutsch 4: Varietäten und Lernvarietäten. Regensburg. Germany. 9/2015.
- Cvrček, V.: *On Some Methodological Issues of CADS*. Workshop Language in Politics in Slavic speaking countries. Jena, Germany. 2015.
- M. Fidler – V. Cvrček: *What grammatical morphemes tell us about discourse*. ICLC13. Newcastle upon Tyne, UK. 2015.
- Cvrček, V. – Čermáková, A. – Chlumská, L. – Richterová, O. – Novotná, R.: *Context Based Approach to Collocations: the Case of Czech*. Corpus Linguistics 2013. Lancaster, UK.
- Cvrček, V. – Fidler, M.: *Not all keywords are created equal: How can we measure keyness?* Corpus Linguistics 2013. Lancaster, UK.
- Cvrček, V.: *Kontextové vymezení kolokací: o slovech monogamních a promiskuitních*. Lingvistika Praha 2013.
- Cvrček, V.: *A Tool for Keyword Analysis in Czech*. AATSEEL. Boston, USA. 2013.
- Fidler, M. – Cvrček, V.: *Keyword analysis with a usage-based perspective: A preliminary study in Czech*. AATSEEL. Boston, 2013.
- Cvrček, V. – Fidler, M.: *Analysis of Keywords in Czech Political Texts. A Needle in a Haystack Method*. CADS Conference 2012. Bologna, Italy.
- Kováříková, D. – Chlumská, L. – Cvrček, V.: *What belongs in a dictionary? The Example of Negation in Czech*. Euralex 2012. Oslo, Norway.
- Fidler, M. – Cvrček, V.: *Keyword Analysis of Czech Political Texts: A Pilot Study*. Seventh Annual Meeting of the Slavic Linguistics Society. University of Kansas, USA. 2012.
- Cvrček, V.: *Literature and Corpus Linguistics*. UCLA, USA. 2012.
- Cvrček, V.: *Czech Language Corpora. Tools for Learning, Teaching and Research*. Workshop at AATSEEL Conference. Seattle. 2012.
- Cvrček, V.: *SyD – A Learners' Tool for Exploring Variability in Language Corpora of Czech*. AATSEEL conference. Seattle 2012.
- Cvrček, V.: *How Large is the Core of Language?* Corpus Linguistics 2011. Birmingham, UK.
- Čermák, F. – Cvrček, V.: *Author Dictionaries Revisited: Dictionary of Bohumil Hrabal*. Euralex 2010. Leeuwarden, Netherlands. 2010.

Václav Cvrček

- Cvrček, V.: *Využití Mluvnice současné češtiny ve výuce češtiny jako cizího jazyka* (Grammar of Contemporary Czech in Learning Czech as a Foreign Language). Asociace učitelů češtiny jako cizího jazyka. Praha 2010.
- Cvrček, V.: *Mluvnice současné češtiny – koncepce a zpracování* (Grammar of Contemporary Czech – concept and reality). Letní škola slovanských studií. Praha 2010.
- Cvrček, V.: *Contextual Approach to Parts of Speech*. In: InterCorp: Exploring a Multilingual Corpus. NLN. Praha 2010, (p. 190–204).

Software tools & applications

- Cvrček, V. – Vondříčka, P.: *KWords*, 2013 – tool for corpus-based extraction of keywords and discourse analysis, see <http://kwords.korpus.cz>
- Cvrček, V. – Vondříčka, P.: *Morfio*, 2012 – tool for analysis of derivation and word-formation productivity, see <http://morfio.korpus.cz>
- Cvrček, V. – Vondříčka, P.: *SyD*, 2011 – tool for analysis of synchronic and diachronic variation in language, see <http://syd.korpus.cz>

Language corpora

- Křen, M. – Cvrček, V. et al.: *Korpus SYN*. Institute of the Czech National Corpus FA CU, Praha 2016.
- Křen, M. – Cvrček, V. et al.: *SYN2015: representative corpus of contemporary written Czech*. Institute of the Czech National Corpus FA CU, Praha 2015.
- Cvrček, V. – Truneček, P. – Horký, V.: *SPEECHES: korpus prezidentských projevů*. Institute of the Czech National Corpus FA CU, Praha 2015.
- Cvrček, V. – Kovářiková, D. – Mácha, J. – Křen, M.: *LINK: korpus odborných lingvistických textů*. Institute of the Czech National Corpus FA CU, Praha 2011.
- Křen, M. – Bartoň, T. – Cvrček, V. et al.: *SYN2010: žánrově vyvážený korpus psané češtiny*. Institute of the Czech National Corpus FA CU, Praha 2010.
- Skoumalová, H. – Bartoň, T. – Cvrček, V. – Hnátková, M. – Koček, J.: *TOTALITA: korpus jazyka totality*. Institute of the Czech National Corpus FA CU, Praha 2010.

Popularisation

- Cvrček, V. – Cvrčková Porkertová, L.: *Velký slovník rýmů /Rhyming dictionary/*. Nakladatelství Lidové noviny. Praha 2011.
- Co je v ČNK nového II*. Korpus – gramatika – axiologie 7, 2013, (s. 95–97).
- Co je v ČNK nového I*. Korpus – gramatika – axiologie 4, 2011, (s. 93–94).
- Dictionary of Bohumil Hrabal (F. Čermák, V. Cvrček eds)*. *Glottotheory* 3/1. 2010, (s. 110–111).
- Dictionary of Karel Čapek (ed. F. Čermák)*. *Glottotheory* 1. 2008, (p. 99–100).
- Dlouholetý spor o jazykovou regulaci*. *Vesmír* 2008/8, (p. 590).
- K osudu knihy Nechte svůj jazyk na pokoji!* *Vesmír* 2008/9, (p. 636).

Teaching

- Teoretické základy statistického vyhodnocování jazykových dat v R** since summer 2015
Theoretical basics of statistical analysis of language data in R Faculty of Arts CU
Course for humanities students with basic computer literacy on using R and conducting basic statistical analysis. Taught together with Silvie Cinková.
- Úvod do lingvistiky pro informatiky** since summer 2016
Introduction to linguistics for IT students FIT, Czech Technical University
Introductory course for non-linguists covering basic concepts of linguistics and phonetics.

Psychosis of Occupation in the Czech Lands Summer 2012
CZ610B Brown University

Content-based course for advanced students of Czech. Course is built around a film Spalovač mrtvol (Burner of Corpses), one of the most prominent classics from the Czech cinema during the 60's; the film is based on a novel with the same title by Ladislav Fuks.

Český jazyk a korpus since Fall 2008
Czech language and Corpora Faculty of Arts CU
Course for students of Czech language and literature. Introduction to the usage of corpora and basic guidelines for corpus-based and corpus-driven research of Czech.

Lexikologický seminář Fall 2009, 2010
Lexicology Faculty of Arts CU
Lexicology, morphematics and word formation (derivation and composition) for students of Czech, non-native advanced speakers.

Korpusová lingvistika: úvod since Fall 2009
Corpus Linguistics: introduction Faculty of Arts CU
Course for students of general linguistics and other philological programs. General introduction to corpus linguistics and its basic concepts (corpus-based vs. corpus-driven approach, representativeness, collocation, multilingual corpora, statistics for corpus linguistics etc.).

Stratifikace češtiny Fall 2008
Stratification of Czech Faculty of Arts CU
Course for students of Czech language, non-native beginner speakers. Using corpora of spoken and written language to explore the stratification of Czech (spoken vs. written language, formal vs. informal registry, dialectological and sociolinguistic differences etc.).

Other skills, qualifications and experiences

Language skills:

Czech (native tongue), English, German, Latin

Computer competence:

Ability to work with internet, text and spreadsheet processing, L^AT_EX

Programming:

Linux, Perl, HTML, R-project

Honors and Awards

Fulbright Visiting Scholar Program (Brown University 8/2011 – 5/2012)

Cena Jana Palacha (Jan Palach Prize), 2005

Dean's award for the best M.A. theses of the year

Membership in Professional Organizations

Prague Linguistic Circle (since 2005)

AATSEEL – American Association of Teachers of Slavic and East European Languages (since 2011)

IATC – International Association of Teachers of Czech (since 2011)

editorial boards: Korpus – gramatika – axiologie, Slovenščina 2.0