

Mgr. Rudolf Rosa

Svatojánská 155/4

266 01 Beroun

Czechia

☎ +420 776 577 617

✉ rosa@ufal.mff.cuni.cz

🌐 <http://ufal.mff.cuni.cz/rudolf-rosa>

Education

Charles University, Prague, Czechia.

Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics

2013–now **Doctoral Studies**, *computational linguistics*.

Discovering the structure of natural language sentences by semi-supervised methods

2007–2013 **Bachelor and Master Studies**, *computer science, computational linguistics*.

thesis “Automatic Post-editing of Phrase-based Machine Translation Outputs” [1]

Programming Skills

Fluent Perl, Bash, Subversion, vim

Operational Python, C++, Java, Linux, Git, SGE, LaTeX, SQL, web technologies

Basics C#, Prolog, XML

Research

2015–now **SloNLP**, *Slovakoczech NLP Workshop*, main organizer and PC chair.

2014–now **croSSSynt**, *Modelling dependency syntax across languages*, GAUK project, Principal Investigator.

2013–now **Universal Dependencies**, *Cross-lingually consistent dependency treebank annotation*, formerly the HamleDT project [5].

2013–2016 **TectoMT**, *Quality translation by deep language engineering approaches*, part of EU-funded project QTLeap.

2011–2015 **Depfix**, *Automatic post-editing of statistical machine translation* [2], part of EU-funded projects FAUST, QTLeap and HimL, Main Developer.

2011–2012 **FilmTit**, *Web application for computer-aided translation of movie subtitles*, students software project, Team Leader.

Awards

ÚFAL best paper award, *one of the best papers of the institute*.

2017 “Slavic Forest, Norwegian Wood” [6]

2014 “HamleDT 2.0: Thirty Dependency Treebanks Stanfordized” [5]

2013 “Deepfix: Statistical Post-editing of Statistical Machine Translation Using Deep Syntactic Analysis” [3]

Work

- 2011–now **Research Assistant and Teaching Assistant**, *Charles University*, Czechia.
syntactic parsing; syntax-based machine translation; Technology for NLP course
- 2015–2016 **Software Engineer (6 month internship)**, *Google*, Zurich, Switzerland.
automatic anaphora resolution in texts
- 2004–2014 **Software Developer**, *Jazyková škola Rosa*, Kladno, Czechia.
development and maintenance of business administration system; web pages
- 2004–2014 **Teacher of English**, *Jazyková škola Rosa*, Kladno, Czechia.
teaching English as second language to children and adults

Language Skills and Qualifications

- Czech native speaker
- English Certificate of Proficiency in English (C2)
- German on the level of B2
- Slovak passive knowledge on the level of B2
- French Diplôme d'étude de langue française (B1)

Selected Publications

Summary **30 peer-reviewed publications**, 329 citations, h-index 10 (Google Scholar).

- [1] Rudolf Rosa. “Automatic Post-editing of Phrase-based Machine Translation Outputs”. Master’s Thesis. Prague, Czechia: Charles University, Fac. of Math. and Physics, 2013.
- [2] Rudolf Rosa, David Mareček, and Ondřej Dušek. “DEPFIX: A System for Automatic Correction of Czech MT Outputs”. In: *Proc of WMT*. Montréal, Canada: ACL, 2012, pp. 362–368.
- [3] Rudolf Rosa, David Mareček, and Aleš Tamchyna. “Deepfix: Statistical Post-editing of Statistical Machine Translation Using Deep Syntactic Analysis”. In: *Proc of ACL Student Research Workshop*. Sofia, Bulgaria: ACL, 2013, pp. 172–179.
- [4] Rudolf Rosa and Zdeněk Žabokrtský. “ KL_{cpos^3} – a Language Similarity Measure for Delexicalized Parser Transfer”. In: *Proc of ACL-IJCNLP, Short Papers*. Stroudsburg, PA, USA: ACL, 2015, pp. 243–249.
- [5] Rudolf Rosa et al. “HamleDT 2.0: Thirty Dependency Treebanks Stanfordized”. In: *Proc of LREC*. Ed. by Nicoletta Calzolari et al. Reykjavík, Iceland: ELRA, 2014, pp. 2334–2341.
- [6] Rudolf Rosa et al. “Slavic Forest, Norwegian Wood”. In: *Proc of VarDial4*. Ed. by Preslav Nakov et al. Stroudsburg, PA, USA: ACL, 2017, pp. 210–219.