

### EDUCATION AND ACADEMIC CAREER

• since 2011: an **associate professor** at the Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics, Charles University, Prague (ÚFAL MFF UK), habilitation thesis “From Treebanking to Machine Translation” • 2005-2011: a researcher and later an **assistant professor**, ÚFAL MFF UK • 2001-2005 – a **PhD** student at ÚFAL MFF UK, PhD thesis “Valency Lexicon of Czech Verbs”, supervised by doc. Mgr. Barbora Vidová Hladká, Ph.D. • 1995-2001 – a bachelor- and master-level (**Ing.**) student at the Faculty of Electrical Engineering, Czech Technical University, study branch Computers, master thesis “Automatic Functor Assignment in the Prague Dependency Treebank”, supervised by Dr. Ing. Ivana Kruijff-Korbayová •

### MAIN RESEARCH DIRECTIONS IN LAST 5 YEARS

• building multilingual morphological resources • NLP applications exploiting parallel corpora • empirical studies in multilingual data collections • theoretical studies on formal representations of natural languages •

### FUNCTIONING IN STUDY PROGRAMS

• since 2012: the **chair** of the Board for the PhD Study Program “**Computational Linguistics**” at MFF UK • since 2019 a member of the Board for the PhD Study Program “**Corpus and Theoretical Linguistics**” at the Faculty of Arts, Charles University • since 2008 involved in the international Erasmus Mundus Masters program “**Language and Communication Technology**” • since 2006: PhD theses opponent, besides MFF UK also the Faculty of Arts (Charles University), the Faculty of Electrical Engineering (Czech Technical University, Prague), the Faculty of Nuclear Sciences and Physical Engineering (Czech Technical University, Prague), and the Faculty of Informatics and Statistics (Prague University of Economics and Business) • since 2012: a state doctoral exam committee member, besides MFF UK also the Faculty of Informatics (Masaryk University, Brno), and the Faculty of Applied Sciences (University of West Bohemia, Pilsen), • since 2012: the organizer of the **ÚFAL PhD Conferences** (annually) • since 2013 the organizer of the **ÚFAL Pre-PhD Student Internship Program**, focused on recruiting prospective PhD students from abroad •

### TEACHING, SUPERVISION, MENTORING

• since 2004 teaching several **courses on Language Data Resources, on Natural Language Processing, and on Machine Learning** at MFF UK • 2006/2007 teaching two NLP courses at the Faculty of Nuclear Sciences and Physical Engineering (Czech Technical University, Prague) • 2016-2019 teaching a NLP course at the Faculty of Electrical Engineering (Czech Technical University, Prague), • successfully defended students under my supervision: **6 PhD** students, **17 master** students, and **3 bachelor** students • mentoring **1 postdoc** •

### SELECTED PROJECTS

• 2017-2022: **LangTech: Modernization of the PhD study program Computational Linguistics**, the principal investigator, financed by OP VVV – the EU-funded Operational Program Research, Development and Education realized by The Ministry of Education, Youth and Sports of the Czech Republic • 2017-2018: **A System for Analyzing Data of the Police of the Czech Republic**, the principal investigator, a project contract with the Ministry of the Interior, Czech Republic, plus two contract follow-ups in 2021 • 2016-2019: “**DigiLing: Trans-European e-Learning Hub for Digital Linguistics**”, the coordinator for Charles University (the Faculty of Arts and the Faculty of Mathematics and Physics involved), an international EU-funded project (Key Action 2, Erasmus+ Strategic Partnership Project) •

### ORGANIZING WORK (ONLY INTERNATIONAL EVENTS LISTED)

• 2022: a co-organizer of the **multilingual morph-segmentation shared task** (held with SIGMORPHON, NAACL, Seattle, USA) • 2022: a co-organizer of the **multilingual coreference resolution shared task** (with CRAC, COLING, Gyeongju, Korea), shared-task proceedings co-editor • 2021: a co-chair of the workshop **DeriMo** (International Workshop on Resources and Tools for Derivational Morphology), held online, workshop proceedings co-editor • 2019: a co-chair and local organizer of **DeriMo**, Prague, workshop proceedings co-editor • 2019: the main organizer of the international **DigiLing Summer School**, Prague • 2007: a member of the local organization teams of the international conferences **ACL** (category A\* category the CORE rankings) a **EMNLP** (category A), Prague • 2003: a local organization team member of the **17th International Congress of Linguists**, Prague •

### EVALUATION WORK

• 2017-2018: an evaluator in The National Accreditation Bureau for Higher Education (**NAÚ**) • 2017-2018: an evaluator in The Technology Agency of the Czech Republic (**TA ČR**) • 2011-2016: an evaluator in The Charles University Grant Agency (**GA UK**) • reviewing for **major international conferences** such as ACL (ranked as an A\* conference in the CORE rankings), EMNLP (ranked as A), NAACL (ranked as A), EACL (ranked as A), COLING (ranked as A), CICLING, for specialized **workshops** such as WMT (Workshop on Statistical Machine Translation) and **DeriMo**, and **journals** such as Studies in Language (5-year IF 0.539), the Journal of Language Modelling, and the Prague Bulletin of Mathematical Linguistics • 2020 an evaluator of the international master study program proposal “Digital Linguistics”, within the approval process in **FI MUNI** • 2020: a project proposal evaluator in The **Icelandic Centre for Research**, more specifically in The Icelandic Strategic Research and Development Programme for Language Technology •

### SELECTED STUDY AND RESEARCH STAYS ABROAD

• 2015: three weeks in the Department of Informatics, Faculty of Sciences, University of **Lisbon** • 2012: one month in the Department of International Business Communication, **Copenhagen Business School** • 2002-2007: more than three months in total during several stays in the Department of Computational Linguistics, University of **Saarbrücken** • 2001: two weeks in the Institute for Communicating and Collaborative Systems, University of **Edinburgh** • 2000: one month in the Department of Intelligent Systems, Jožef Stefan Institute, **Ljubljana** (and 2 shorter stays in 2001 and 2003) • 2000: three months in the Department of Mathematics, University of **Patras** • 1999: one month in the Fuzzy Logic Laboratory, Johannes Kepler University in **Linz**, and 2 weeks in 2000 •

### AWARDS

• 2014: 2nd place in the Best Paper Award at the **CICLing 2014 Conference** • 2009: a co-author of the book “Valenční slovník českých sloves” awarded by the **Union of interpreters and translators** • 1998: a winner-team member in the **programming contest Czech Technical University Open Contest**, organized within the international network ACL International Collegiate Programming Contests •

### INVITED TALKS (ONLY INTERNATIONAL EVENTS LISTED)

• 2016: “To tree or not to tree?”, **Prague Gathering of Logicians**, Prague • 2012: “Machine Translation using Dependency Trees.” **Theorietag**, Der 22nd Theorietag der Fachgruppe “Automaten und Formale Sprachen” der Gesellschaft für Informatik, Prague • 2011: “Treex – an open-source framework for natural language processing.” **ITAT** (Information technologies – Applications and Theory), Vrátná dolina, Slovakia •

### INVOLVED IN ÚFAL COOPERATION WITH COMMERCIAL PARTNERS (EXAMPLES)

• 2019: NLP and ML consultations for DHL IT Services Europe s.r.o., • 2017: a Tamil dataset license for Google Inc., • 2016: a NLP data contract with ACREA CR, s. r.o., • 2014: an NLP toolkit license for Datlowe, s.r.o. •

### CITATIONS (NOVEMBER 2022):

• **WOS: h-index 10**, 227 citations (without self-citations), • **Scopus: h-index 18**, 787 citations • **Google Scholar: h-index 32**, 4134 citations •