

Curriculum Vitæ

David Lukeš

Address	Na Zderaze 13, 120 00 Prague 2	Citizenship	Czech Republic
E-mail	david.lukes@ff.cuni.cz	Born	September 3 rd , 1988
Website	https://dlukes.github.io	Phone	+420 732 341 117

Education

Faculty of Arts, Charles University, Prague, Czech Republic

2014–present	(ongoing) PhD studies in Mathematical Linguistics dissertation project: <i>Intonation patterns in informal Czech based on spoken corpus data</i> supervisor: Mgr. Pavel Vondříčka, PhD (pavel.vondricka@ff.cuni.cz) status of project: outline and theoretical framework approved, data under preparation
2011–2014	MA in Phonetics and English and American Studies Thesis: <i>Perceptual Sensitivity to Speech and Music Stimuli in the Frequency and Temporal Domains</i> Graduated <i>cum laude</i>
2008–2011	BA in English and American Studies and French Philology Thesis: <i>Textual Identity in Selected Novels by Philip Roth: Representation, Dissimulation, Creation</i> Graduated <i>cum laude</i>

Academic achievements

2013/2014	JUDr. Karel Engliš Award for Best Charles University Alumni in the Social Sciences
2010/2011	Vilém Mathesius Award for Outstanding BA Thesis (Literatures in English)

Professional experience

2017–present	Academic staff – junior researcher Institute of the Czech National Corpus Faculty of Arts, Charles University
2013–2016	Technical staff – spoken corpus project coordinator for phonetic transcription Institute of the Czech National Corpus Faculty of Arts, Charles University
2010–2013	Student assistant Department of Anglophone Literatures and Cultures Faculty of Arts, Charles University

2009–present	Freelance translator English and French into Czech and vice versa
2009–2012	InterCorp project – external collaborator OCR quality and parallel alignment checks of English and Czech texts
2008–present	External collaborator at Scio Author of test items for English language proficiency tests Supervisor of test item authors French language proficiency test designer

Publications and conference presentations

2017

- Klimešová, Petra, Zuzana Komrsková, Marie Kopřivová, and David Lukeš (2017). “Avenues for Corpus-based Research on Informal Spoken Czech”. In: *Language, Corpora and Cognition*. Ed. by Piotr Pezik and Jacek Tadeusz Waliński. Vol. 51. Łódź Studies in Language. Frankfurt am Main: Peter Lang Edition, pp. 145–162.
- Šturm, Pavel and David Lukeš (2017). “Fonotaktická analýza obsahu slabik na okrajích českých slov v mluvené a psané řeči”. In: *Slovo a slovesnost*. To appear in issue 2/2017.

2016

- Lukeš, David and Zuzana Komrsková (2016). “Strategies for Automatic Morphological Tagging of Non-standard Czech”. Conference presentation at the Slavic Spoken Corpora Workshop, Slavisches Seminar der Universität Freiburg i. Breisgau.
- Zasina, Adrian et al. (2016). “Multidimenzionální analýza češtiny. Pilotní studie”. Conference presentation at Korpusová lingvistika Praha 2016, ÚČNK FF UK.

2015

- Goláňová, Hana, Marie Kopřivová, David Lukeš, and Martin Štěpán (2015). “Kartografické a geografické zpracování dat z mluvených korpusů”. In: *Korpus – gramatika – axiologie* 11, pp. 42–54.
- Klimešová, Petra, Zuzana Komrsková, Marie Kopřivová, and David Lukeš (2015). “Slovo *to* v mluvených korpusech ČNK, jeho prefixace a reduplikace”. In: *Časopis pro moderní filologii* 97.1, pp. 21–30.
- Lukeš, David (2015a). “Increasing Speed and Consistency of Phonetic Transcription of Spoken Corpora Using ASR Technology”. In: *Corpus Linguistics 2015 – Abstract Book*. Ed. by Federica Formato and Andrew Hardie. Lancaster: UCREL, pp. 222–223.
- (2015b). “New Tools for Working with the ORAL Series Corpora of Spoken Czech: AchSynku and MluvKonk”. In: *Natural Language Processing, Corpus Linguistics, Lexicography*. Ed. by Katarína Gajdošová and Adriána Žáková. Eighth International Conference. Proceedings of SLOVKO 2015. Bratislava, Slovakia: RAM-Verlag, pp. 90–101.
- Lukeš, David, Petra Klimešová, Zuzana Komrsková, and Marie Kopřivová (2015). “Experimental Tagging of the ORAL Series Corpora: Insights on Using a Stochastic Tagger”. In: *TSD 2015, LNAI 9302*. Ed. by Pavel Král and Václav Matoušek. Springer International Publishing, pp. 342–350.

2014

- Kopřivová, Marie, Hana Goláňová, Petra Klimešová, Zuzana Komrsková, and David Lukeš (2014). "Multi-tier Transcription of Informal Spoken Czech: The ORTOFON Corpus Approach". In: *Complex Visibles Out There. Proceedings of the Olomouc Linguistics Colloquium 2014: Language Use and Linguistic Structure*. Olomouc Modern Language Series, Vol. 4. Univerzita Palackého. Olomouc, pp. 529–544.
- Kopřivová, Marie, Petra Klimešová, Hana Goláňová, and David Lukeš (2014). "Mapping Diatopic and Diachronic Variation in Spoken Czech: The ORTOFON and DIALEKT Corpora". In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Ed. by Nicoletta Calzolari et al. Reykjavik, Iceland: European Language Resources Association (ELRA), pp. 376–382.
- Lukeš, David (2014). "Perspektivy fonetické anotace v korpusech mluveného jazyka". In: *Korpusová lingvistika Praha 2014 – Abstrakty*. Work-in-progress presentation. Ústav Českého národního korpusu. Praha, pp. 125–127.
- Lukeš, David, Dita Fejlová, and Radek Skarnitzl (2014). "Variability of Czech Alveolar Plosives: A Locus Equation Perspective". In: *AUC Philologica, Phonetica Pragensia XIII.1*, pp. 21–32.

2013

- Fejlová, Dita, David Lukeš, and Radek Skarnitzl (2013). "Formant Contours in Czech Vowels: Speaker-discriminating Potential". In: *INTERSPEECH-2013*, pp. 3182–3186.

Further education

Jun 2015	RLING2015 Bootcamp – Statistics for Linguistics with R Louvain-la-Neuve, Belgie Taught by Stefan Th. Gries
Sep 2013	International Summer School on Agent-based Models of Creativity Cortona, Italy Organized by the Evolutionary Linguistics Association
2007–2008, 2010–2011	Instituto Cervantes , Praha extracurricular Spanish

Languages

Czech	Native speaker
English	Fluent
French	Fluent
Spanish	Intermediate (but with little practice recently)

Computer skills

Operating systems	Linux, macOS, Windows (including a firm grasp of the command line in the first two cases)
-------------------	---

Programming languages	Python, JavaScript, R, Perl; willing and eager to learn more (Rust, Clojure)
Software	Tools for annotating and processing speech recordings (Praat, ELAN, the KALDI speech recognition library), commonly used office suite applications, L ^A T _E X

Other interests

Literature	Especially Anglo-American: David Lodge, Joseph Heller, Philip Roth...
Music	Frank Zappa; guitar, saxophone
Sports	Football, badminton