

Veronika VINCZE, PhD

CONTACT INFORMATION

HUN-REN—SZTE Research Group on Artificial Intelligence

Tisza Lajos krt. 103., 6720 Szeged, Hungary

Phone: +36 62 54 41 40

Mobile: +36 70 22 99 108

E-mail: vinczev@inf.u-szeged.hu

Homepage: <http://www.inf.u-szeged.hu/~vinczev>

PERSONAL DATA

Date of birth: 1 July 1981

Nationality: Hungarian

Marital status: married, 3 children (born in 2019, 2020 and 2023)

RESEARCH INTERESTS

natural language processing (especially corpus building, information extraction, computational morphology, parsing), lexical semantics, dependency syntax

DEGREES

- | | |
|------|--|
| 2015 | PhD in Computer Science, University of Szeged Thesis title: <i>Uncertainty Detection in Natural Language Texts</i> |
| 2012 | PhD in Linguistics, University of Szeged Thesis title: <i>Semi-Compositional Noun + Verb Constructions: Theoretical Questions and Computational Linguistic Analyses</i> |
| 2004 | MA in General Linguistics, University of Szeged Thesis title: <i>Analyzing Light Verb Constructions with the Help of Lexical Functions</i> (in Hungarian) |
| 2004 | MA in English Language and Literature, University of Szeged Thesis title: <i>Complimenting Habits of Hungarian Students</i> |

EDUCATION

- | | |
|-----------|---|
| 2004-2007 | Theoretical Linguistics PhD programme, University of Szeged |
| 1999-2004 | General Linguistics BA&MA programme, University of Szeged |
| 1999-2004 | English Language and Literature BA&MA programme, University of Szeged |
| 2001-2004 | Portuguese Language and Literature Minor, University of Szeged |

WORK EXPERIENCE

- | | |
|-----------|---|
| 2018- | Senior research fellow at the HUN-REN—SZTE Research Group on Artificial Intelligence, Hungarian Academy of Sciences |
| 2019- | Annotator and evaluator in translation tasks – Unbabel |
| 2017-2018 | QA Manager in Data Science – Black Swan Hungary |
| 2015- | Language technology developer – Lionbridge Technologies Inc. /TELUS International |

| | |
|-----------|--|
| 2013-2020 | Senior lecturer at the University of Szeged, Department of Informatics, Human Language Technology Group |
| 2012-2018 | Research fellow at the MTA-SZTE Research Group on Artificial Intelligence |
| 2005-2011 | Junior research fellow at the University of Szeged, Department of Informatics, Human Language Technology Group |
| 2009 | Part-time researcher at the Hungarian Academy of Sciences, Institute for Linguistics, Department of Language Technology and Sociolinguistics, Research Group for Language Technology |
| 2000-2005 | Student annotator at the University of Szeged, Department of Informatics, Human Language Technology Group |

RESEARCH PROJECTS

| | |
|-----------|---|
| 2022-2026 | <i>Universality, diversity and idiosyncrasy in language technology (UniDive)</i> . An ICT COST Action sponsored by the European Union. Role: researcher |
| 2020- | Artificial Intelligence National Laboratory Role: researcher |
| 2015-2016 | <i>INFRA: Developing an open source, integrated research infrastructure for natural language processing</i> sponsored by the Hungarian Academy of Sciences. Role: researcher, computational linguist expert |
| 2013-2017 | <i>PARSEME: PARSing and Multi-word Expressions. Towards linguistic precision and computational efficiency in natural language processing</i> . An ICT COST Action sponsored by the European Union. Role: researcher, Steering Committee member, ESR representative |
| 2013-2017 | <i>Computational tools for the revitalization of endangered Finno-Ugric minority languages (FinUgRevita)</i> sponsored by the Finnish Academy of Sciences and the Hungarian National Research Fund. Role: researcher, computational linguist expert |
| 2013-2015 | <i>Telemedicine Oriented Research in the Fields of Mathematics, Informatics and Medical Sciences</i> sponsored by the European Union and the European Social Fund. Role: researcher, linguist expert |
| 2012-2014 | <i>Infocommunicational technologies and the society of the future</i> sponsored by the European Union and the European Social Fund. Role: researcher, linguist expert |
| 2009-2012 | <i>Development of Model-Based Semantic Search System</i> sponsored by the National Office for Research and Technology, Hungary. Role: researcher, linguist expert |
| 2007-2008 | <i>Development of an effective knowledge management tool by integrating linguistic and graph-theoretical tools for Hungarian and EU customs authorities</i> sponsored by the National Office for Research and Technology, Hungary. Role: researcher, linguist expert |
| 2005-2007 | <i>Construction of the Hungarian WordNet Ontology and its Application in Information Extraction Systems</i> sponsored by the National Office for Research and Technology, Hungary. Role: researcher, linguist expert |

- 2005-2007 *Hungarian-English Machine Translation System* sponsored by the National Office for Research and Technology, Hungary.
Role: researcher, linguist expert
- 2002-2004 *Machine learning of syntax rules (application of machine learning methods for the generation of Hungarian syntactic rules)* sponsored by the Ministry of Education, Hungary.
Role: chief annotator, linguist
- 2001-2003 *Information Extraction from Short Business News* sponsored by the Ministry of Education, Hungary.
Role: chief annotator, linguist
- 2000-2002 *Development of a Part-of-Speech Tagging Method for Hungarian by using Machine Learning Algorithms* sponsored by the Ministry of Education, Hungary.
Role: chief annotator, linguist

AWARDS, PRIZES, SCHOLARSHIPS

- 2017-2018 Scholarship for young researchers, granted by New National Excellence Programme of Hungary.
- 2017 Women in Science Excellence Award (in information technology).
- 2016 Herman József Young Researcher Award, granted by the Research Institute For Linguistics, Hungary.
- 2015 Visiting researcher at CSLI, Stanford University, Stanford, USA (2.5 months), granted by the PARSEME COST Action.
- 2015 Best Young Researcher Award at the 11th Conference on Hungarian Computational Linguistics.
- 2014 Erdős Pál Scholarship for young researchers, granted by National Excellence Programme of Hungary.
- 2012 DAAD scholarship for young researchers at the Universität Trier, Trier, Germany (5 months).
- 2008 2nd place at the *Obesity Challenge*, team Szeged.
- 2005 4th place at the call *Language and language use within the family* issued by the National Office of Cultural Heritage, Hungary.
- 2003-2004 ERASMUS scholarship at the Université de Paris III, Sorbonne Nouvelle, Paris, France (4 months).
- 2002-2004 Fellowship granted by the Hungarian Republic.
- 2001-2003 Fellowship granted by the municipality of Balatonfüred, Hungary.

TEACHING EXPERIENCE

- 2017- *Introduction to Computational Linguistics* – newly designed course for PhD students, in Hungarian, Faculty of Arts, University of Szeged.
- 2016 *Introduction to Computational Linguistics* – newly designed course for BA students, in Hungarian, Faculty of Arts, University of Szeged.
- 2014 *Mesterséges intelligencia [Artificial Intelligence]* – for BA and BSc students, in Hungarian, University of Szeged.
- 2014- *Programozás nyelvészeknek - Java [Programming for Linguists – Java]* – newly designed course for BA students, in Hungarian, Faculty of Arts, University of Szeged.

- 2013 *Nyelvészet az informatikában – informatika a nyelvészetben [Linguistics in Informatics – Informatics in Linguistics]* – newly designed course for BSc students, in Hungarian, Faculty of Science and Informatics, University of Szeged.
- 2013- *Korpuszok a nyelvészeti kutatásban [Corpora in Linguistic Research]* – newly designed course for BA students, in Hungarian, Faculty of Arts, University of Szeged.
- 2013 *A számítógépes nyelvfeldolgozás alapjai [Basics of natural language processing]* – newly designed course for BA students, in Hungarian, Faculty of Arts, University of Szeged.
- 2012- *How to Write Scientific Papers?* – newly designed course for PhD students, Faculty of Science and Informatics, University of Szeged.
- 2011 *Syntax-Morphology Interface and Natural Language Processing*. Thematic Training Course on Processing Morphologically Rich Languages, Hungarian Academy of Sciences, Institute for Linguistics – University of Helsinki.
- 2004-2007 *English syntax* for BA students, University of Szeged, Institute of English and American Studies.
- 2004- English language, private classes.
- 2004 English language, Ságvári Endre High School, Szeged, Hungary (teaching practice).
- In 2023/24, I supervise 3 PhD-theses at the University of Szeged.

REVIEWING ACTIVITY

Conferences: ACL-2013, ACL-2018, ACL-2019, ACL-2020, ACL-2021, MWE 2019, ACL-MWE 2016, RANLP-2013, EACL-MWE-2014, NAACL-2016, NAACL-2018, NAACL-MWE-2013, NAACL-MWE-2015, MoMa 2015, PolTal-2014, Baltic NLP 2018, Law-CxG-MWE 2018, CoNLL-2013 Shared Task, CoNLL-2010 Shared Task

Journals: Journal of Language Modelling (member of editorial board), PLOS One, Language Resources and Evaluation, Natural Language Engineering, Studies in Computational Intelligence (Springer series), Information, BMC Research Notes, Applied Linguistics, Nyelvtudomány

CONFERENCE ORGANIZATION

- 2017-2018 The PARSEME Shared Task on Identifying Verbal Multiword Expressions
- 2017 The 13th Workshop on Multiword Expressions (MWE 2017)
- 2016 Second International Workshop on Computational Linguistics for Uralic Languages 2010 the CoNLL-2010 Shared Task
- 2008 the Fourth Global WordNet Conference (GWC2008)
- 2007- Conferences on Hungarian Computational Linguistics
- 2005-2006 the 9th and 10th National Conferences of PhD Students of Linguistics

PUBLICATIONS

224 publications (111 in English, 113 in Hungarian)

1 book, 10 book chapters, 25 journal papers, 2 PhD theses, 152 conference papers, 20 edited proceedings
h-index = 23, independent citations > 2000, Erdős number = 3

SELECTED PUBLICATIONS¹

¹ For a full list of publications, please visit <https://m2.mtmt.hu/gui2/?type=authors&mode=browse&sel=10027782>

Szabó, Martina Katalin; Vincze, Veronika; Dam, Bernadett; Guba, Csenge; Bagi, Anita; Szendi, István. 2024. Predictive and Distinctive Linguistic Features in Schizophrenia-Bipolar Spectrum Disorders. In: *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 12938–12953.

Szabó, Martina Katalin; Vincze, Veronika; Bibok, Károly. 2023. 'Thank You for the Terrific Party!' – An Analysis of Hungarian Negative Emotive Words. *CORPUS LINGUISTICS AND LINGUISTIC THEORY* 19 (3): 451–485.

Kálmán, János; Devanand, Davangere P.; Gosztolya, Gábor; Balogh, Réka; Imre, Nóra; Tóth, László; Hoffmann, Ildikó; Kovács, Ildikó; Vincze, Veronika; Pákáski, Magdolna. 2022. Temporal Speech Parameters Detect Mild Cognitive Impairment in Different Languages: Validation and Comparison of the Speech-GAP Test® in English and Hungarian. *CURRENT ALZHEIMER RESEARCH* 19 (5): 373–386.

Nagy, T. István; Rácz, Anita; Vincze, Veronika. 2020. Detecting Light Verb Constructions across Languages. *NATURAL LANGUAGE ENGINEERING* 26 (3): 319–348.

Savary, Agata; Ramisch, Carlos; Cordeiro, Silvio; Sangati, Federico; Vincze, Veronika; QasemiZadeh, Behrang; Candito, Marie; Cap, Fabienne; Giouli, Voula; Stoyanova, Ivelina; Doucet, Antoine 2017. The PARSEME Shared Task on Automatic Identification of Verbal Multiword Expressions. In: Markantonatou, Stella; Ramisch, Carlos; Savary, Agata; Vincze, Veronika (eds.) *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)*. ACL, Valencia, pp. 31-47.

Vincze, Veronika; Gosztolya, Gábor; Tóth, László; Hoffmann, Ildikó; Szatlóczki, Gréta; Bánréti, Zoltán; Pákáski, Magdolna; Kálmán, János 2016. Detecting Mild Cognitive Impairment by Exploiting Linguistic Information from Transcripts. In: Erk, Katrin; Smith, Noah A. (eds.): *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Berlin, pp. 181-187.

Vincze, Veronika 2014: Uncertainty Detection in Hungarian Texts. In: *Proceedings of COLING 2014*, Dublin, pp. 1844-1853.

Vincze, Veronika; Nagy T., István; Farkas, Richárd 2013: Identifying English and Hungarian Light Verb Constructions: A Contrastive Approach. In: *Proceedings of ACL 2013 (Volume 2: Short Papers)*, pp. 255-261.

Vincze, Veronika; Zsibrita, János; Nagy T., István 2013: Dependency Parsing for Identifying Hungarian Light Verb Constructions. In: *Proceedings of IJCNLP 2013*, pp. 207-215.

Szarvas, György; Vincze, Veronika; Farkas, Richárd; Móra, György; Gurevych, Iryna 2012. Cross-Genre and Cross-Domain Detection of Semantic Uncertainty. *Computational Linguistics – Special Issue on Modality and Negation* 38(2):335-367.

Farkas, Richárd; Vincze, Veronika; Schmid, Helmut 2012. Dependency Parsing of Hungarian: Baseline Results and Challenges. In: *Proceedings of the Thirteenth Conference of the European Chapter of Association for Computational Linguistics (EACL-2012)*, Avignon, France.

Vincze, Veronika; Csirik, János 2010. Hungarian Corpus of Light Verb Constructions. In: *Proceedings of COLING 2010*, Beijing, China.

Farkas, Richárd; Vincze, Veronika; Móra, György; Csirik, János; Szarvas, György 2010. The CoNLL-2010 Shared Task: Learning to Detect Hedges and their Scope in Natural Language Text. In: *Proceedings of the Fourteenth Conference on Computational Natural Language Learning (CoNLL-2010): Shared Task*, Uppsala, Sweden, pp. 1-12.

Vincze, Veronika; Szauter, Dóra; Almási, Attila; Móra, György; Alexin, Zoltán; Csirik, János 2010. Hungarian Dependency Treebank. In: *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10)*, Valletta, Malta.

Vincze, Veronika; Szarvas, György; Farkas, Richárd; Móra, György; Csirik, János 2008. The BioScope Corpus: biomedical texts annotated for uncertainty, negation and their scopes. *BMC Bioinformatics* 9 (Suppl 11):S9 doi:10.1186/1471-2105-9-S11-S9.

SELECTED CORPORA, DATASETS AND TOOLS

magyarlanc: a linguistic preprocessing toolkit for Hungarian

Szeged (Dependency) Treebank: Hungarian texts manually annotated for part-of-speech tags, dependency syntax, constituency syntax, named entities, light verb constructions and coreference

Hungarian WordNet

BioScope: biomedical texts annotated for uncertainty cues and their scope

CoNLL-2010 Shared Task Corpora: biomedical and Wikipedia texts annotated for uncertainty cues

Wiki50: Wikipedia texts annotated for named entities and multiword expressions

4FX: English, Spanish, German and Hungarian legislative texts annotated for light verb constructions

S-Gap: a screening method for filtering mild cognitive impairment on the basis of speech characteristics

OUTREACH ACTIVITIES

- | | |
|-----------|--|
| 2018 | Presentation at <i>Girls' Day</i> organized by the Women in Science Association, Hungarian Academy of Sciences, for high school girls. |
| 2017 | Participation as an NLP expert in the <i>A tudomány hangjai</i> scientific radio programme of Kossuth Rádió. |
| 2014- | Participation in <i>Girls' Day</i> organized by the Institute of Informatics, University of Szeged, for high school girls. |
| 2014 | Participation as an NLP expert in the <i>Többet ésszel</i> scientific radio programme of Klubrádió. |
| 2012- | Presentation of NLP research carried out at the Human Language Technology Group, Szeged for university students, programmers and linguists at the NLP Meetup, Budapest (several times). |
| 2011 | Participation as an NLP expert in the <i>Mindentudás Egyeteme 2.0</i> (University of Omniscience) scientific TV programme series of the Hungarian National Television. |
| 2008-2010 | Participation in the activities of the Hungarian Speech and Language Technology Platform (aiming at presenting results of the Hungarian NLP community to the public). |
| 2008 | Presentation of NLP research carried out at the Human Language Technology Group, Szeged for university students at the <i>Scientific Lectures Series</i> organized by the Móra Ferenc Student Hostel, Szeged, Hungary. |

COMPUTER SKILLS

Windows

Linux

Microsoft Office

LaTeX

Java

HTML

LANGUAGE SKILLS

Hungarian: mother tongue

English: fluent

French: fluent

German: good

Portuguese: basic

Italian: basic

Spanish: basic

Latin: basic

Chinese: beginner