

Curriculum Vitae



Armin Hoenen
Taunusstraße 57
63067 Offenbach
069/79824687
hoenen@em.uni-frankfurt.de
Date of Birth: 25.02.1981
Employed as: Researcher
Married (since 27.03.2014)
Children: 1 daughter (* 14.08.2014)

Education

- 2000 Abitur, Ingelheim
2000 Rescue Medic Certificate with German Red Cross, Mainz
2003-2010 Magister Artium
 Comparative Linguistics/ Computer Science (minor)/ Biology (minor);
 Thesis on: *Grammatical Description of Okinawan* ; Prof. W. Bisang, Prof. M. Grein
 Johannes Gutenberg University Mainz
2011 Kontaktstudium Sprachandragogik (Adult Language Teaching Studies):
 course units, teaching exam, theoretical exam;
 Thesis on: *Second Script Acquisition* ; Prof. M. Grein, Dr. René Piedmont
 Johannes Gutenberg University Mainz & Landesverband der Volkshochschulen
 Rheinland-Pfalz
2017 PhD Thesis in Computer Science on *Tools, Evaluation and Preprocessing for
Stemmatology*, defended 19.12.2017; Prof. A. Mehler, Prof. V. Ramesh
 Goethe University Frankfurt

Work Experience

- 2000 – 2003 Paramedic at German Red Cross (Rettungssanitäter)
2004 – 2011 Teaching Language Courses
2005 JAC Japan: Host in the German Pavillon of the EXPO, Nagoya, Aichi, Japan
2007-2009 Prime Research: Media analyses, News Service, Translations
2009 IBM Böblingen Lab: Data Mining Internship
2009 Rüdesheim Tourist AG: Tourism Marketing Internship,
 Homepage Administration (Joomla),
 Translations for www.ruedesheim.de (English, French, Dutch)
2009 TableServe Team Event Catering: Waitor
2009 Sushi Circle Wiesbaden: Sushi Cook
2009 – 2010 Kern AG: Holiday Substitution, Translation Check, Administration
2010 – today Researcher at Goethe University Frankfurt &
 Akademie der Wissenschaften zu Göttingen

Teaching Experience

- 2004-2011 language courses in Japanese/ English/ Business English/ Dutch and Hindi
 Folk Highschool (Bad Kreuznach, Bingen, Ingelheim, Mainz) / private /Schott, NS
 Löwen Entertainment, Bingen City Administration
2014 Leading the Workshop Computer-assisted Methods in Old-Iranian Linguistics at the
 Workshop Iranian Dialectology, Teheran
2017 Spring School Course *Crawling the web for resources in Low Resource Languages
of Western Asia*, Goethe University/ Grade, Conference and Spring School *Language
Contact and Language Change in Western Asia*

Academic Record

Publications

- A. Hoenen, "Principles Aiding in Reading Abbreviations in OldGeorgian and Latin ," in Book of Abstracts DHd Köln, 2018.
- A. Hoenen, "Multi Modal Distance - An Approach to Stemma Generation With Weighting," in Proceedings of the 11th International Conference on Language Resources and Evaluation, 2018.
- A. Hoenen, N. Schenk, "Knowing the Author by the Company His Words Keep," in Proceedings of the 11th International Conference on Language Resources and Evaluation, 2018.
- A. Hoenen, "From Manuscripts to Archetypes through Iterative Clustering," in Proceedings of the 11th International Conference on Language Resources and Evaluation, 2018.
- A. Hoenen, "Repetition Analyses Function," in Proceedings of the 2015 Herrenhäuser Symposium Visual Linguistics, 2018.
- A. Hoenen, S. Eger, and R. Gehrke, "How Many Stemmata with Root Degree k ?", in Proceedings of the 15th Meeting on the Mathematics of Language, 2017, pp. 11-21.
- A. Hoenen, "Using Word Embeddings for Computing Distances Between Texts and for Authorship Attribution," in International Conference on Applications of Natural Language to Information Systems, 2017, pp. 274-277.
- A. Hoenen, "Beyond the tree – a theoretical model of contamination and a software to generate multilingual stemmata," in Book of Abstracts of the annual conference of the AIUCD 2017, Sapienza, Rome, AIUCD, 2017.
- S. Eger, A. Hoenen, and A. Mehler, "Language classification from bilingual word embedding graphs," in Proceedings of COLING 2016, 2016.
- A. Hoenen, "Silva Portentosissima – Computer-Assisted Reflections on Bifurcativity in Stemmas," in Digital Humanities 2016: Conference Abstracts. Jagiellonian University & Pedagogical University, 2016, pp. 557-560.
- A. Lücking, A. Hoenen, and A. Mehler, "TGermaCorp — A (Digital) Humanities Resource for (Computational) Linguistics," in Proceedings of the 10th International Conference on Language Resources and Evaluation, 2016.
- A. Hoenen, "Wikipedia Titles As Noun Tag Predictors," in Proceedings of the 10th International Conference on Language Resources and Evaluation, 2016.
- A. Hoenen, "Das erste dynamische Stemma, Pionier des digitalen Zeitalters?," in Accepted in the Proceedings of the Jahrestagung der Digital Humanities im deutschsprachigen Raum, 2016.
- Corpora and Resources for (Historical) Low Resource Languages"Corpora and Resources for (Historical) Low Resource Languages," , vol. 31, iss. 2, 2016. A. Hoenen, A. Mehler, and J. Gippert, "Editorial," JLCL, vol. 31, iss. 2, p. iii–iv, 2016.
- A. Hoenen and L. Samushia, "Gepi: An Epigraphic Corpus for Old Georgian and a Tool Sketch for Aiding Reconstruction," JLCL, vol. 31, iss. 2, pp. 25-38, 2016.
- A. Hoenen and F. Mader, "A New LMF Schema Application by Example of an Austrian Lexicon Applied to the Historical Corpus of the Writer Hugo von Hofmannsthal," in Historical Corpora, 2015.
- N. Dundua, A. Hoenen, and L. Samushia, "A Parallel Corpus of the Old Georgian Gospel Manuscripts and their Stematology," The Georgian Journal for Language Logic Computation, vol. IV, pp. 176-185, 2015.
- A. Hoenen, "Das artifizielle Manuskriptkorpus TASCFE," in Accepted in the Proceedings of the Jahrestagung der Digital Humanities im deutschsprachigen Raum, 2015.
- A. Hoenen, "Lachmannian Archetype Reconstruction for Ancient Manuscript Corpora," in Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT), 2015.
- A. Hoenen, "Simulating Misreading," in Proceedings of the 20TH INTERNATIONAL CONFERENCE ON APPLICATIONS OF NATURAL LANGUAGE TO INFORMATION SYSTEMS (NLDB), 2015.
- A. Hoenen, "Stematology, an interdisciplinary endeavour," in Book of Abstracts zum DHd Workshop Informatik und die Digital Humanities, DHd, 2014.

A. Hoenen, "Simulation of Scribal Letter Substitution," in *Analysis of Ancient and Medieval Texts and Manuscripts: Digital Approaches*, 2014.

M. Z. Islam and A. Hoenen, "Source and Translation Classification using Most Frequent Words," in *Proceedings of the 6th International Joint Conference on Natural Language Processing (IJCNLP)*, 2013.

A. Hoenen, "Measuring Repetitiveness in Texts, a Preliminary Investigation," *Sprache und Datenverarbeitung. International Journal for Language Data Processing*, vol. 36, iss. 2, pp. 93-104, 2012.

Altüberlieferte Sprachen als Gegenstand der Texttechnologie -- Ancient Languages as the Object of Text Technology A. Hoenen and T. Jügel, *Altüberlieferte Sprachen als Gegenstand der Texttechnologie — Ancient Languages as the Object of Text Technology*, A. Hoenen and T. Jügel, Eds., *JLCL*, 2012, vol. 27.

M. Sukhareva, M. Z. Islam, A. Hoenen, and A. Mehler, "A Three-step Model of Language Detection in Multilingual Ancient Texts," in *Proceedings of Workshop on Annotation of Corpora for Research in the Humanities*, Heidelberg, Germany, 2012.

R. Gleim, A. Hoenen, N. Diewald, A. Mehler, and A. Ernst, "Modeling, Building and Maintaining Lexica for Corpus Linguistic Studies by Example of Late Latin," in *Corpus Linguistics 2011*, 20-22 July, Birmingham, 2011.

Invited

2013 Career Day for linguists, Johannes Gutenberg Universität Mainz, Germany. *Computer linguistics by example of stemmatology*.

2016 Dixit Workshop at ESTS Conference, Antwerp, Panel Member *Discussion on Textual Criticism*

Organisation of Scientific Conferences

2016 CEDIFOR/GSCL Workshop on Corpora and Resources for Low Resource Languages, Goethe University Frankfurt, Germany.

2015 NAACL HLT Student Volunteer

2010 XLVII. StuTS 2010, Mainz Annual Student Conference on Linguistics: Organisation

Stipends

2008 Deutsche Bank Stipend for Students of Languages

2012 Interedition Symposium: "Scholarly Digital Editions, Tools and Infrastructure", Stipend financed by COST (European Cooperation in Science and Technology)

2012 NASSLLI (North American Summer School of Logics, Language and Information), Stipend by the University of Texas, Austin

Abilities

Medicine:

- Rescue medicine education and practise, Rescue Swimmer Certificate, Disaster Control Training

Software:
Java, JavaScript, Python, PHP, QBasic, TurboPascal, XML(HTML, TEI), Joomla, Wordpress, WEKA, R, Office, Windows, Linux, Apple, PhotoShop, and others

Languages:

German (mother tongue), English (fluent), French (fluent), Dutch (fluent), Italian (fluent), Japanese (fluent), Russian(B)¹, Hindi/Urdu/Hindustani(B), Spanish (B), Yiddish(B), Afrikaans (B), Arabic(A), Persian (A), Turkish(A), Georgian(A), Mandarin(A), Portuguese(A), Norwegian/Swedish/Danish(A), Croatian/Serbian/Slovenian(A), Greek(A), Tok Pisin (A)

Hobbies

Guitar, poetry, languages

1 Letters refer to the European Reference Framework, which does not congruently apply to non-European languages.