egon w. stemle

European Academy of Bozen/Bolzano (EURAC)
Viale Druso, 1
I-39100 Bolzano (BZ)
✉ *egon.stemle@eurac.edu*

Publications

2014
(in Press) **Egon Stemle, Alexander Onysko**, *The role of transfer in automated L1 identification of English learner essays*, in *Hagen Peukert, editor:Transfer Effects in Multilingual Language Development*.

2014
(in Press) **Aivars Glaznieks, Egon Stemle**, *Challenges of building a CMC corpus for analyzing writer's style by age: The DiDi project*, in *Michael Beisswenger, Nellek Oostdijk, Angelika Storrer, Henk van den Heuvel, editors:Journal for Language Technology and Computational Linguistics (JLCL), Special Issue*.

10.2014 **Jennifer-Carmen Frey, Egon W Stemle, Aivars Glaznieks**, *Collecting language data of non-public social media profiles*, in *Workshop Proceedings of the 12th Edition the KONVENS Conference*, pages 11–15.

09.2014
(Accepted) **Michel Généreux, Egon W Stemle, Verena Lyding, Lionel Nicolas**, *Correcting OCR errors for German in Fraktur font*, first Italian Computational Linguistics Conference (CLIC-it), Dec 9-10, Pisa, Italy.

05.2014 **Verena Lyding, Lionel Nicolas, Egon Stemle**, *'interHist' - an interactive visual interface for corpus exploration*, in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 635–641.

05.2014 **Andrea Abel, Aivars Glaznieks, Lionel Nicolas, Egon Stemle**, *KoKo: An L1 Learner Corpus for German*, in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*.

04.2014 **Verena Lyding, Egon Stemle, Claudia Borghetti, Marco Brunello, Sara Castagnoli, Felice Dell Orletta, Henrik Dittmann, Alessandro Lenci, Vito Pirrelli**, *The PAISÀ Corpus of Italian Web Texts*, in *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*, pages 36–43.

11.2013 **Verena Lyding, Claudia Borghetti, Henrik Dittmann, Lionel Nicolas, Egon Stemle**, *Open Corpus Interface for Italian Language Learning*, in *Conference proceedings. ICT for language learning*.

09.2013 **Lionel Nicolas, Egon W Stemle, Klara Kranebitter, Verena Lyding**, *High-Accuracy Phrase Translation Acquisition Through Battle-Royale Selection*, in *Proceedings of Recent Advances in Natural Language Processing, RANLP 2013*, pages 516–524.

07.2013 **Stefan Evert, Egon Stemle, Paul Rayson, editors**, Proceedings of the 8th Web as Corpus Workshop (WAC-8), Workshop at the seventh international Corpus Linguistics conference (CL2013), Jul 22.

06.2013
(in Press) **Klara Kranebitter, Egon W Stemle**, *Constructing concept relation maps to support building concept systems in comparative legal terminology*, Terminologie & Ontologie: Théories et Applications (TOTh'2013), Jun 6-7.

05.2013
(in Press) **Lionel Nicolas, Egon Stemle, Aivars Glaznieks, Andrea Abel**, *A Generic Data Workflow for Building Annotated Text Corpora*, Compiling and Using Learner Corpora to Teach and Assess Productive and Interactive Skills in Foreign Languages at University Level, Learner Corpora, May 16-17, Padova, Italia.

2012
(in Press) **Aivars Glaznieks, Andrea Abel, Verena Lyding, Lionel Nicolas, Egon Stemle**, *Establishing a Standardised Procedure for Building Learner Corpora*, Learner Language, Learner Corpora 2012 (LLLC2012), Oct 5-6, University of Oulu, Finland, in *Tarja Nikula, Sauli Takala, Sabine Ylönen, editors:Apples - Journal of Applied Language Studies*.

09.2012 **Lionel Nicolas, Egon W Stemle, Klara Kranebitter**, *Towards high-accuracy bilingual phrase acquisition from parallel corpora*, in *Proceedings of KONVENS 2012*, pages 471–479.

07.2012 **Francesca Bonin, Fabio Cavulli, Aronne Noriller, Massimo Poesio, Egon W Stemle**, *Annotating Archaeological Texts: An Example of Domain-Specific Annotation in the Humanities*, in *Proceedings of the Sixth Linguistic Annotation Workshop*, pages 134–138.

11.2011   **Massimo Poesio, Eduard Barbu, Francesca Bonin, Fabio Cavulli, Asif Ekbal, Egon Stemle, Christian Girardi**, *The Humanities Research Portal: Human Language Technology Meets Humanities Publication Archives*, in *SDH2011 - Supporting Digital Humanities: Answering the unaskable*.

11.2011   **Asif Ekbal, Francesca Bonin, Sriparna Saha, Egon Stemle, Eduard Barbu, Fabio Cavulli, Christian Girardi, Massimo Poesio**, *Rapid Adaptation of NE Resolvers for Humanities Domains using Active Annotation*, *Journal for Language Technology and Computational Linguistics (JLCL)*, 26(2):39–51.

07.2011   **Brian Murphy, Egon W Stemle**, *PaddyWaC: A Minimally-Supervised Web-Corpus of Hiberno-English*, in *Proceedings of the First Workshop on Algorithms and Resources for Modelling of Dialects and Language Varieties*, pages 22–29.

06.2011   **Massimo Poesio, Eduard Barbu, Egon W Stemle, Christian Girardi**, *Structure-Preserving Pipelines for Digital Libraries*, in *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH 2011)*, pages 54–62.

05.2010   **Kepa Joseba Rodríguez, Francesca Delogu, Jannick Versley, Egon W Stemle, Massimo Poesio**, *Anaphoric Annotation of Wikipedia and Blogs in the Live Memories Corpus*, in *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10)*.

09.2009   **Johannes M Steger, Egon W Stemle**, *KrdWrd — The Architecture for Unified Processing of Web Content*, in *Proceedings of the Fifth Web as Corpus Workshop (WAC5)*.

09.2007   **Daniel Bauer, Judith Degen, Xiaoye Deng, Priska Herger, Jan Gasthaus, Eugenie Giesbrecht, Lina Jansen, Christin Kalina, Thorben Krüger, Robert Märtin, Martin Schmidt, Simon Scholler, Johannes Steger, Egon Stemle, Stefan Evert**, *FIASCO: Filtering the Internet by Automatic Subtree Classification, Osnabrück*, in *Proceedings of the Third Web as Corpus Workshop (WAC3)*.

07.2007   **Sebastian Blohm, Philipp Cimiano, Egon Stemle**, *Harvesting Relations from the Web - Quantifiying the Impact of Filtering Functions*, in *Proceedings of the 22nd Conference on Artificial Intelligence (AAAI-07)*, pages 1316–1323.

11.2005   **Martin Bleichner, Eugenie Giesbrecht, Helmar Gust, Eva-Maria Leicht, Petra Ludewig, Sabine Möller, Wiebke Müller, Martin Schmidt, Moritz Stefaner, Egon Stemle, Katja Wilke**, *ASADO: The Analysis and Structuring of Aviation Documents – Final Report*, Technical report, Institute of Cognitive Science at the University of Osnabrück and Institute of Applied Linguistics at the University of Hildesheim.

Workshops and Conferences

10.2014   **Member of the Programme Committee**, *Workshop*, Natural Language Processing for Computer-Mediated Communication / Social Media (NLP4CMC).

04.2014   **Member of the Programme Committee**, *Workshop*, 9th Web as Corpus Workshop (WAC-9).

07.2013   **Member of the Organising Committee and Programme Committee**, *Workshop*, 8th Web as Corpus Workshop (WAC-8).

09.2006   **Student Worker at OTT06**, *Workshop*, Ontologies in Text Technology: Approaches to Extract Semantic Knowledge from Syntactic Information.

06.2006   **Student Worker at QITL-2**, *Workshop*, Second Workshop on Quantitative Investigations in Theoretical Linguistics.