

Stephanie Evert â Research â Teaching â CV â **Publications** â Software â Hobbies â Trans*

Publications

Monographs â Journals â Book Chapters â Conference Proceedings â Edited Volumes â Miscellaneous

In case you've been wondering how to cite my papers after my transition, please use my name as specified in the original paper (and in the listing below). I do not want to hide my personal history and the fact that I spent 50 years trying to live as a man. You also might consider switching to a citation style such as APA that only uses initials for given names.

Monographs

Evert, Stefan (2004, published 2005). *The Statistics of Word Cooccurrences: Word Pairs and Collocations*. Dissertation, Institut für maschinelle Sprachverarbeitung, University of Stuttgart, URN urn:nbn:de:bsz:93-opus-23714. [official version <<http://elib.uni-stuttgart.de/opus/volltexte/2005/2371/>>, PDF <<https://www.stephanie-evert.de/PUB/Evert2004phd.pdf>>, companion website <<http://www.collocations.de/>>]

Hoffmann, Sebastian; Evert, Stefan; Smith, Nicholas; Lee, David; Berglund Prytz, Ylva (2008). *Corpus Linguistics with BNCweb – a Practical Guide*, volume 6 of *English Corpus Linguistics*. Peter Lang, Frankfurt am Main. [ordering information <<http://www.peterlang.com/index.cfm?vID=56315>>]

Journal Papers

Patel, Malin; Garibyan, Armine; Winckel, Elodie; Evert, Stephanie (2023). **A reference construction as a database**. *Yearbook of the German Cognitive Linguistics Association*, **11**(1), 175â202. [journal homepage <<https://doi.org/10.1515/gcla-2023-0009>>]

Adrian, Axel; Dykes, Natalie; Evert, Stephanie; Heinrich, Philipp; Keuchen, Michael (2022). **Entwicklung und Evaluation automatischer Verfahren zur Anonymisierung von Gerichtsentscheidungen**. *Legal Tech*, **2022**(4), 233â238. [manuscript (PDF) <https://www.stephanie-evert.de/PUB/AdrianEtc2022_manuscript.pdf>, journal homepage <<https://rdb.manz.at/document/rdb.tso.LIltz20220406>>]

Dykes, Natalie; Evert, Stefan; GÃttlinger, Merlin; Heinrich, Philipp; SchrÃder, Lutz (2021). **Argument parsing via corpus queries**. *it â Information Technology*, **63**(1), 31â44. [journal homepage <<https://doi.org/10.1515/itit-2020-0051>>]

Dykes, Natalie; Evert, Stefan; GÃttlinger, Merlin; Heinrich, Philipp; SchrÃder, Lutz (2020). **Reconstructing arguments from noisy text**. *Datenbank-Spektrum*. [PDF <<https://rdcu.be/b5mXV>>, journal homepage <<https://doi.org/10.1007/s13222-020-00342-y>>]

Evert, Stefan; Heinrich, Philipp; Henselmann, Klaus; Rabenstein, Ulrich; Scherr, Elisabeth; Schmitt, Martin; SchrÃder, Lutz (2019). **Combining machine learning and semantic features in the classification of corporate disclosures**. *Journal of Logic, Language and Information*, **28**, 309â330. [accepted manuscript (PDF) <https://www.stephanie-evert.de/PUB/EvertEtc2019_JoLLI.pdf>, journal homepage <<https://link.springer.com/article/10.1007/s10849-019-09283-6>>]

BÃttner, Andreas; Dimpel, Friedrich Michael; Evert, Stefan; Jannidis, Fotis; PielstrÃm, Steffen; Proisl, Thomas; Reger, Isabella; SchÃch, Christof; Vitt, Thorsten (2017). **Delta in der stilometrischen Autorschaftsattribuion**. *Zeitschrift für digitale Geisteswissenschaften*. <https://dx.doi.org/10.17175/2017_006>.

Evert, Stefan; Proisl, Thomas; Jannidis, Fotis; Reger, Isabella; PielstrÃm, Steffen; SchÃch, Christof; Vitt, Thorsten (2017). **Understanding and explaining Delta measures for authorship attribution**. *Digital Scholarship in the Humanities*. **22**(suppl_2), ii4âii16. [open access (PDF) <<https://doi.org/10.1093/lc/fqx023>>, reference corpus <<https://github.com/cophi-wue/refcor>>, R code & data sets <<https://github.com/schtepf/ExplainingDelta>>]

SchÃfer, Fabian; Evert, Stefan; Heinrich, Philipp (2017). **Japan's 2014 general election: Political bots, right-wing internet activism and PM Abe ShinzÅ's hidden nationalist agenda**. *Big Data*, **5**(4), 294â309. [open access (PDF) <<https://doi.org/10.1089/big.2017.0049>>]

Evert, Stefan; Greiner, Paul; Baigger, JoÃo Filipe; Lang, Bastian (2016). **A distributional approach to**

open questions in market research. *Computers in Industry*, **78**, 16â28. [accepted manuscript (PDF) <https://www.stephanie-evert.de/PUB/EvertGreinerEtc2016_COMIND.pdf>, journal homepage <<http://dx.doi.org/10.1016/j.compind.2015.10.008>>]

Lapesa, Gabriella and Evert, Stefan (2014). **A large scale evaluation of distributional semantic models: Parameters, interactions and model selection.** *Transactions of the Association for Computational Linguistics*, **2**, 531–545. [PDF <<https://www.stephanie-evert.de/PUB/LapesaEvert2014tacl.pdf>>, supplementary material <<http://www.linguistik.fau.de/dsmeval/>>]

Biemann, Chris; Bildhauer, Felix; Evert, Stefan; Goldhahn, Dirk; Quasthoff, Uwe; SchÃtfer, Roland; Simon, Johannes; Swiezinski, Leonard; Zesch, Torsten (2013). **Scalable construction of high-quality Web corpora.** *Journal for Language Technology and Computational Linguistics (JLCL)*, **28**(2), 23–59. [PDF <http://www.jlcl.org/2013_Heft2/2Biemann.pdf>]

Ansorge, Ulrich; Reynvoet, Bert; Hendler, Jessica; Oetl, Lennart; Evert, Stefan (2013). **Conditional automaticity in subliminal morphosyntactic priming.** *Psychological Research*, **77**, 399–421.

Michelbacher, Lukas, Evert, Stefan, and SchÃtze, Hinrich (2011). **Asymmetry in corpus-derived and human word associations.** *Corpus Linguistics and Linguistic Theory*, **7**(2), 245–276.

Evert, Stefan (2006). **How random is a corpus? The library metaphor.** *Zeitschrift fÃr Anglistik und Amerikanistik*, **54**(2), 177–190. [manuscript (PDF) <<https://www.stephanie-evert.de/PUB/Evert2006.pdf>>, journal homepage <<http://www.zaa.koenigshausen-neumann.de/aktuell.htm>>]

Carletta, Jean, Evert, Stefan, Heid, Ulrich, Kilgour, Jonathan, and Chen, Yiya (2005). **The NITE XML Toolkit: data model and query language.** *Language Resources and Evaluation*, **39**(4), 313–334. [NXT homepage <<http://nite.sourceforge.net/>>, journal homepage <<https://doi.org/10.1007/s10579-006-9001-9>>]

Evert, Stefan and Krenn, Brigitte (2005). **Using small random samples for the manual evaluation of statistical association measures.** *Computer Speech & Language* **19**(4), 450–466. [manuscript (PDF) <<https://www.stephanie-evert.de/PUB/EvertKrenn2005.pdf>>]

Carletta, Jean; Evert, Stefan; Heid, Ulrich; Kilgour, Jonathan; Robertson, Judy; Voormann, Holger (2003). **The NITE XML toolkit: Flexible annotation for multimodal language data.** *Behavior Research Methods, Instruments, & Computers*, **35**(3), 353–363. [PDF <https://www.stephanie-evert.de/PUB/CarlettaEtc2003_BRM.pdf>, journal homepage <<https://doi.org/10.3758/BF03195511>>]

Book Chapters

Evert, Stefan (2013). **Tools for the acquisition of lexical combinatorics.** In R. H. Gouws, U. Heid, W. Schweickard, and H. E. Wiegand (eds.), *Dictionaries. An International Encyclopedia of Lexicography. Supplementary volume: Recent Developments with Focus on Electronic and Computational Lexicography (HSK 5.4)*, chapter 104, pages 1415–1432. Mouton de Gruyter, Berlin, New York.

Evert, Stefan, FrÃttschl, Bernhard, and Lindstrot, Wolf (2009). **Statistische Grundlagen.** In K.–U. Carstensen, C. Ebert, C. Ebert, S. Jekat, R. Klabunde, and H. Langer, editors, *Computerlinguistik und Sprachtechnologie: Eine EinfÃhrung*, pages 114–158. Spektrum Akademischer Verlag, Heidelberg, 3rd edition.

Ebert, Christian; Schiehlen, Michael; Klabunde, Ralf; Evert, Stefan (2009). **Semantik.** In K.–U. Carstensen, C. Ebert, C. Ebert, S. Jekat, R. Klabunde, and H. Langer, editors, *Computerlinguistik und Sprachtechnologie: Eine EinfÃhrung*, pages 330–393. Spektrum Akademischer Verlag, Heidelberg, 3rd edition.

Evert, Stefan (2008). **Corpora and collocations.** In A. LÃdeling and M. KytÃ (eds.), *Corpus Linguistics. An International Handbook*, article 58, pages 1212–1248. Mouton de Gruyter, Berlin. [extended manuscript (PDF) <https://www.stephanie-evert.de/PUB/Evert2007HSK_extended_manuscript.pdf>]

Baroni, Marco and Evert, Stefan (2008). **Statistical methods for corpus exploitation.** In A. LÃdeling and M. KytÃ (eds.), *Corpus Linguistics. An International Handbook*, article 36, pages 777–803. Mouton de Gruyter, Berlin. [manuscript (PDF) <https://www.stephanie-evert.de/PUB/BaroniEvertHSK38_manuscript.pdf>]

Evert, Stefan and Fitschen, Arne (2001). **Textkorpora**. In K.–U. Carstensen, C. Ebert, C. Endriss, S. Jekat, R. Klabunde, and H. Langer (eds.), *Computerlinguistik und Sprachtechnologie – Eine Einführung*, pages 369–376. Spektrum Akademischer Verlag, Heidelberg, Berlin.

Conference Proceedings and Collections

2024 – 2023 – 2022 – 2021 – 2020 – 2019 – 2018 – 2017 – 2016 – 2015 – 2014 – 2013 – 2012 – 2011 – 2010 – 2009 – 2008 – 2007 – 2006 – 2005 – 2004 – 2003 – 2002 – 2001 – 2000

2024

Adrian, Axel; Evert, Stephanie; Heinrich, Philipp; Keuchen, Michael (2024). **Auslegung des KI-VO-E zur Evaluation von Verfahren der Künstlichen Intelligenz am Beispiel der automatischen Anonymisierung von Gerichtsentscheidungen**. In *Juristische Sprachmodelle – Tagungsband des 27. Internationalen Rechtsinformatik Symposions IRIS 2024*. Editions Weblaw. [manuscript (PDF) <https://www.stephanie-evert.de/PUB/AdrianEtc2024_manuscript.pdf>] *LexisNexis best paper award* <<https://www.lexisnexis.at/blog-post/rueckblick-lexisnexis-best-paper-awards-iris2024/>>

Dykes, Nathan; Evert, Stephanie; Heinrich, Philipp; Humml, Merlin; Schröder, Lutz (2024). **Leveraging high-precision corpus queries for text classification via large language models**. In A. Hautli-Janisz, G. Lapesa, L. Anastasiou, V. Gold, A. D. Liddo, and C. Reed (eds.), *Proceedings of the First Workshop on Language-driven Deliberation Technology (DELITE) @ LREC-COLING 2024*, pages 52â57, Torino, Italia. [ACL anthology <<https://aclanthology.org/2024.delite-1.7>>]

Heinrich, Philipp; Blombach, Andreas; Doan Dang, Bao Minh; Zilio, Leonardo; Havenstein, Linda; Dykes, Nathan; Evert, Stephanie; Schäfer, Fabian (2024). **Automatic identification of COVID–19–related conspiracy narratives in German telegram channels and chats**. In N. Calzolari, M.–Y. Kan, V. Hoste, A. Lenci, S. Sakti, and N. Xue (eds.), *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 1932â1943, Torino, Italia. [ACL anthology <<https://aclanthology.org/2024.lrec-main.173>>]

2023

Adrian, Axel; Dykes, Nathan; Evert, Stephanie; Heinrich, Philipp; Keuchen, Michael (2023). **Automatische Anonymisierung von Gerichtsurteilen: Eine Vision scheint realisierbar**. In E. Schweighofer, J. Zanol, and S. Eder (eds.), *Rechtsinformatik als Methodenwissenschaft des Rechts – Tagungsband des 26. Internationalen Rechtsinformatik Symposions IRIS 2023*. Editions Weblaw. [manuscript (PDF) <https://www.stephanie-evert.de/PUB/AdrianEtc2023_manuscript.pdf>, publisher homepage <https://jusletter-it.weblaw.ch/issues/2023/30-maerz-2023/automatische-anonymi_afe7683d74.html>]

2022

Adrian, Axel; Dykes, Nathan; Evert, Stephanie; Heinrich, Philipp; Keuchen, Michael; Proisl, Thomas (2022). **Manuelle und automatische Anonymisierung von Urteilen**. In A. Adrian, S. Evert, M. Kohlhase, and M. Zwickel (eds.), *Digitalisierung von Zivilprozess und Rechtsdurchsetzung*. Duncker & Humblot, Berlin.

Blombach, Andreas; Evert, Stephanie; Jannidis, Fotis; Pielström, Steffen; Konle, Leonard; Proisl, Thomas (2022). **Exploring lexical diversities**. In *Digital Humanities 2022: Conference Abstracts*, pages 130â133, Tokyo, Japan / online. [PDF <<https://www.stephanie-evert.de/PUB/BlombachEtc2022.pdf>>]

Dykes, Nathan; Heinrich, Philipp; Evert, Stephanie (2022). **Retrieving Twitter argumentation with corpus queries and discourse analysis**. In S. Flach and M. Hilpert (eds.), *Broadening the Spectrum of Corpus Linguistics. New approaches to variability and change*, number 105 in Studies in Corpus Linguistics. John Benjamins. [publisher homepage <<https://benjamins.com/catalog/sc1.105.08dyk>>]

Evert, Stephanie (2022). **Measuring keyness**. In *Digital Humanities 2022: Conference Abstracts*, pages 202â205, Tokyo, Japan / online. [PDF <<https://www.stephanie-evert.de/PUB/Evert2022.pdf>>, osf.io/cy6mw <<https://osf.io/cy6mw/>>]

2021

Adrian, Axel; Evert, Stefan; Keuchen, Michael; Heinrich, Philipp; Dykes, Natalie (2021).

Anonymisierung von Gerichtsurteilen â Eine wesentliche Voraussetzung fÃ¼r E-Justice. In E. Schweighofer, F. Kummer, A. SaarenpÃ¤Ã¤, S. Eder, and P. Hanke (eds.), *Cybergovernance â Tagungsband des 24. Internationalen Rechtsinformatik Symposions IRIS 2021*, pages 137â149. Editions Weblaw. [manuscript (PDF) <https://www.stephanie-evert.de/PUB/AdrianEtc2021_manuscript.pdf>, publisher homepage <https://jusletter-it.weblaw.ch/issues/2021/27-Mai-2021/04_e-government_01.d_31b309579d.html>]

Evert, Stefan and Lapesa, Gabriella (2021). **FAST: A carefully sampled and cognitively motivated dataset for distributional semantic evaluation.** In *Proceedings of the 25th Conference on Computational Natural Language Learning (CoNLL 2021)*, pages 588â595, Online. [PDF <<https://www.stephanie-evert.de/PUB/EvertLapesa2021.pdf>>, ACL anthology <<https://aclanthology.org/2021.conll-1.46/>>, FAST data set <<https://osf.io/cd8ar/>>]

Neumann, Stella and Evert, Stefan (2021). **A register variation perspective on varieties of English.** In E. Seoane and D. Biber (eds.), *Corpus based approaches to register variation*, chapter 6, pages 143â178. Benjamins, Amsterdam. [online supplement <<https://www.stephanie-evert.de/PUB/NeumannEvert2021/>>, manuscript (PDF) <https://www.stephanie-evert.de/PUB/NeumannEvert2021/NeumannEvert2021_manuscript.pdf>, publisher homepage <<https://doi.org/10.1075/scl.103.06neu>>]

Tayebi Arasteh, Soroosh; Monajem, Mehrpad; Christlein, Vincent; Heinrich, Philipp; Nicolaou, Anguelos; Boldaji, Hamidreza Naderi; Lotfinia, Mahshad; Evert, Stefan (2021). **How will your tweet be received? Predicting the sentiment polarity of tweet replies.** In *Proceedings of the 15th IEEE International Conference on Semantic Computing (ICSC 2021)*, pages 356â359. [PDF <<https://www.stephanie-evert.de/PUB/TayebiArastehEtc2021.pdf>>, data set <<https://www.kaggle.com/datasets/soroosharasteh/retweet>>]

2020

Evert, Stefan; Harlamov, Oleg; Heinrich, Philipp; Banski, Piotr (2020). **Corpus Query Lingua Franca part II: Ontology.** In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)*, Marseille, France. [PDF <https://www.stephanie-evert.de/PUB/EvertEtc2020_CQLF2.pdf>, GitHub <<https://github.com/cqlf-ontology/cqlf>>]

Proisl, Thomas; Dykes, Natalie; Heinrich, Philipp; Kabashi, Besim; Blombach, Andreas; Evert, Stefan (2020). **EmpirIST corpus 2.0: Adding manual normalization, lemmatization and semantic tagging to a German Web and CMC corpus.** In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)*, Marseille, France. [PDF <https://www.stephanie-evert.de/PUB/ProislEtc2020_LREC.pdf>, corpus & resources <<https://github.com/fau-klue/empirist-corpus>>]

2019

Anthony, Laurence and Evert, Stefan (2019). **Embracing the concept of data interoperability in corpus tools development.** In *Proceedings of the Corpus Linguistics 2019 Conference*, Cardiff, UK. [PDF <https://www.stephanie-evert.de/PUB/AnthonyEvert2019_CL.pdf>]

Dykes, Natalie; Evert, Stefan; Peters, Joachim; Heinrich, Philipp (2019). **Argumentation is key: A keyword-based study of arguments in online discourse.** In *Proceedings of the Corpus Linguistics 2019 Conference*, Cardiff, UK. [PDF <https://www.stephanie-evert.de/PUB/DykesEtc2019_CL.pdf>, slides <https://www.stephanie-evert.de/PUB/DykesEtc2019_CL_slides.pdf>]

Proisl, Thomas; Konle, Leonard; Evert, Stefan; Jannidis, Fotis (2019). **Dependenzbasierte syntaktische KomplexitÃ¤tsmaÃe.** In *Proceedings of DHd 2019*, pages 270â273, Frankfurt & Mainz, Germany. [PDF <https://www.stephanie-evert.de/PUB/ProislEtc2019_DHd.pdf>, poster <https://www.stephanie-evert.de/PUB/ProislEtc2019_DHd_poster.pdf>]

2018

Heinrich, Philipp; Adrian, Christoph; Kalashnikova, Olena; SchÃ¤fer, Fabian; Evert, Stefan (2018). **A transnational analysis of news and tweets about ânuclear phase-outâ in the aftermath of the Fukushima incident.** In *Proceedings of the 1st Workshop on Computational Impact Detection from Text*

Data (CIDTD 2018), pages 8â16, Miyazaki, Japan. [PDF <https://www.stephanie-evert.de/PUB/HeinrichEtc2018_CIDTD.pdf>, slides <https://www.stephanie-evert.de/PUB/HeinrichEtc2018_CIDTD_slides.pdf>]

Proisl, Thomas; Evert, Stefan; Jannidis, Fotis; SchÃ¶ch, Christof; Konle, Leonard; PielstrÃ¶m, Steffen (2018). **Delta vs. n-gram tracing: Evaluating the robustness of authorship attribution methods.** In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, pages 3309â3314, Miyazaki, Japan. [PDF <https://www.stephanie-evert.de/PUB/ProislEtc2018_LREC.pdf>, slides <https://www.stephanie-evert.de/PUB/ProislEtc2018_LREC_slides.pdf>]

Proisl, Thomas; Heinrich, Philipp; Kabashi, Besim; Evert, Stefan (2018). **EmotiKLUE at IEST 2018: Topic-informed classification of implicit emotions.** In *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 235â242, Brussels, Belgium. [PDF <https://www.aclweb.org/anthology/W18-6234.pdf>, software <https://github.com/tsproisl/EmotiKLUE>]

Uhrig, Peter; Evert, Stefan; Proisl, Thomas (2018). **Collocation candidate extraction from dependency-annotated corpora: Exploring differences across parsers and dependency annotation schemes.** In P. Cantos-GÃ³mez and M. Almela-SÃ¡nchez (eds.), *Lexical Collocation Analysis: Advances and Applications*, pages 111â140. Springer International Publishing, Cham. [publisher homepage <https://doi.org/10.1007/978-3-319-92582-0_6>]

2017

Evert, Stefan; Heinrich, Philipp; Henselmann, Klaus; Rabenstein, Ulrich; Scherr, Elisabeth; SchrÃ¶der, Lutz (2017). **Combining machine learning and semantic features in the classification of corporate disclosures.** In *Proceedings of the Workshop on Logic and Algorithms in Computational Linguistics (LACompLing 2017)*, pages 47â62, Stockholm, Sweden. [PDF <https://www.stephanie-evert.de/PUB/EvertHeinrichEtc2017.pdf>]

Evert, Stefan and Neumann, Stella (2017). **The impact of translation direction on characteristics of translated texts. A multivariate analysis for English and German.** In G. De Sutter, M.-A. Lefer, and I. Delaere (eds.), *Empirical Translation Studies. New Theoretical and Methodological Traditions*, number 300 in Trends in Linguistics. Studies and Monographs (TiLSM). Mouton de Gruyter, Berlin. Online supplement: <http://www.stefan-evert.de/PUB/EvertNeumann2017/>.

Evert, Stefan; Uhrig, Peter; Bartsch, Sabine; Proisl, Thomas (2017). **E-VIEW-alation â a large-scale evaluation study of association measures for collocation identification.** In *Electronic lexicography in the 21st century. Proceedings of the eLex 2017 conference*, pages 531â549, Leiden, The Netherlands. [PDF <https://www.stephanie-evert.de/PUB/EvertUhrigEtc2017.pdf>, slides <https://www.stephanie-evert.de/PUB/EvertUhrigEtc2017_slides.pdf>, video <https://www.youtube.com/watch?v=xYo3wTRx8F8>, E-VIEW-alation <http://www.collocations.de/eviewalation/>]

Evert, Stefan; Wankerl, Sebastian; NÃ¶th, Elmar (2017). **Reliable measures of syntactic and lexical complexity: The case of Iris Murdoch.** In *Proceedings of the Corpus Linguistics 2017 Conference*, Birmingham, UK. [PDF <https://www.stephanie-evert.de/PUB/EvertWankerlNoeth2017.pdf>, Slides <https://www.stephanie-evert.de/PUB/EvertWankerlNoeth2017_Slides.pdf>]

Lapasa, Gabriella and Evert, Stefan (2017). **Large-scale evaluation of dependency-based DSMs: Are they worth the effort?** In *Proceedings of the 15th Annual Meeting of the European Association for Computational Linguistics (EACL 2017)*, pages 394â400, Valencia, Spain. [PDF <https://www.stephanie-evert.de/PUB/LapasaEvert2017.pdf>, supplementary material <http://www.linguistik.fau.de/dsmeval/>]

Proisl, Thomas; Heinrich, Philipp; Evert, Stefan; Kabashi, Besim (2017). **Translation inference across dictionaries via a combination of graph-based methods and co-occurrence statistics.** In *Proceedings of the LDK 2017 Workshops: Shared Task on Translation Inference Across Dictionaries (TIAD)*, pages 94â102, Galway, Ireland. CEUR. [PDF <https://www.stephanie-evert.de/PUB/ProislEtc2017_TIAD.pdf>]

Wankerl, Sebastian; NÃ¶th, Elmar; Evert, Stefan (2017). **An n-gram based approach to the automatic**

diagnosis of alzheimer’s disease from spoken language. In *Proceeding of INTERSPEECH 2017*, pages 3162â3166, Stockholm, Sweden. [PDF <<https://www.stephanie-evert.de/PUB/WankerlNoethEvert2017.pdf>>]

2016

BeiÃwenger, Michael; Bartsch, Sabine; Evert, Stefan; WÃ¼rzner, Kay-Michael (2016). **EmpiriST 2015: A shared task on the automatic linguistic annotation of computer-mediated communication and web corpora.** In *Proceedings of the 10th Web as Corpus Workshop (WAC-X) and the EmpiriST Shared Task*, pages 44–56, Berlin, Germany. [PDF <<https://www.stephanie-evert.de/PUB/BeisswengerEtc2016.pdf>>, task homepage <<https://sites.google.com/site/empirist2015/>>]

Evert, Stefan (2016). **CogALex-V shared task: Mach5 â a traditional DSM approach to semantic relatedness.** In *Proceedings of the 5th Workshop on Cognitive Aspects of the Lexicon (CogALex-V)*, pages 92â97, Osaka, Japan. [PDF <https://www.stephanie-evert.de/PUB/Evert2016_Mach5.pdf>, system & data <<http://www.collocations.de/data/#mach5>>]

Evert, Stefan; Jannidis, Fotis; Dimpel, Friedrich Michael; SchÃ¶ch, Christof; PielstrÃ¶m, Steffen; Vitt, Thorsten; Reger, Isabella; BÃ¼ttner, Andreas; Proisl, Thomas (2016). **âDeltaâ in der stilometrischen Autorschafts attribution.** In *Proceedings of DHD 2016*, pages 61â74, Leipzig, Germany. [HTML <<http://www.dhd2016.de/abstracts/sektionen-002.html>>]

Santus, Enrico; Gladkova, Anna; Evert, Stefan; Lenci, Alessandro (2016). **The CogALex-V shared task on the corpus-based identification of semantic relations.** In *Proceedings of the 5th Workshop on Cognitive Aspects of the Lexicon (CogALex-V)*, pages 69â79, Osaka, Japan. [PDF <<https://www.stephanie-evert.de/PUB/SantusEtc2016.pdf>>, task homepage <<https://sites.google.com/site/cogalex2016/home/shared-task>>]

Wankerl, Sebastian; NÃ¶th, Elmar; Evert, Stefan (2016). **An analysis of perplexity to reveal the effects of Alzheimer’s disease on language.** In *ITG-Fachbericht 267: Speech Communication*, pages 254â259, Paderborn, Germany. [PDF <<https://www.stephanie-evert.de/PUB/WankerlNoethEvert2016.pdf>>]

2015

Evert, Stefan and Arppe, Antti (2015). **Some theoretical and experimental observations on naÃve discriminative learning.** In *Proceedings of the 6th Conference on Quantitative Investigations in Theoretical Linguistics (QITL-6)*, TÃ¼bingen, Germany. [PDF <<https://www.stephanie-evert.de/PUB/EvertArppe2015.pdf>>, handout (PDF) <https://www.stephanie-evert.de/PUB/EvertArppe2015_handout.pdf>, slides (PDF) <https://www.stephanie-evert.de/PUB/EvertArppe2015_slides.pdf>]

Evert, Stefan and Hardie, Andrew (2015). **Zigurat: A new data model and indexing format for large annotated text corpora.** In *Proceedings of the 3rd Workshop on the Challenges in the Management of Large Corpora (CMLC-3)*, pages 21–27, Lancaster, UK. [PDF <<https://www.stephanie-evert.de/PUB/EvertHardie2015.pdf>>]

Evert, Stefan; Proisl, Thomas; Jannidis, Fotis; PielstrÃ¶m, Steffen; SchÃ¶ch, Christof; Vitt, Thorsten (2015). **Towards a better understanding of Burrows’s Delta in literary authorship attribution.** In *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, Denver, CO. Co-located with NAACL-HLT 2015. [PDF <https://www.stephanie-evert.de/PUB/EvertEtc2015_CL4Lit.pdf>]

Plotnikova, Nataliia; Kohl, Micha; Volkert, Kevin; Lerner, Andreas; Dykes, Natalie; Ermer, Heiko; Evert, Stefan (2015). **KLUEless: Polarity classification and association.** In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 619–625, Denver, Colorado. [PDF <https://www.stephanie-evert.de/PUB/PlotnikovaEtc2015_KLUEless.pdf>, extended report <https://www.stephanie-evert.de/PUB/PlotnikovaEtc2015_KLUEless_report.pdf> with qualitative evaluation]

Plotnikova, Nataliia; Lapesa, Gabriella; Proisl, Thomas; Evert, Stefan (2015). **SemantiKLUE: Semantic textual similarity with maximum weight matching.** In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 111–116, Denver, Colorado. [PDF <https://www.stephanie-evert.de/PUB/PlotnikovaEtc2015_SemantiKLUE.pdf>]

2014

Bartsch, Sabine and Evert, Stefan (2014). **Towards a Firthian notion of collocation**. In A. Abel and L. Lemnitzer (eds.), *Vernetzungsstrategien, Zugriffsstrukturen und automatisch ermittelte Angaben in Internetwörterbüchern*, number 2/2014 in OPAL – Online publizierte Arbeiten zur Linguistik, pages 48–61. Institut für Deutsche Sprache, Mannheim. [PDF <<https://www.stephanie-evert.de/PUB/BartschEvert2014.pdf>>]

Diwersy, Sascha, Evert, Stefan, and Neumann, Stella (2014). **A weakly supervised multivariate approach to the study of language variation**. In B. Szmrecsanyi and B. Wälchli, editors, *Aggregating Dialectology, Typology, and Register Analysis. Linguistic Variation in Text and Speech*, *Linguae et Litterae: Publications of the School of Language and Literature, Freiburg Institute for Advanced Studies*. De Gruyter, Berlin. [online access <<http://www.degruyter.com/viewbooktoc/product/207699>>, earlier manuscript (PDF) <<https://www.stephanie-evert.de/PUB/DiwersyEvertNeumann2012.pdf>>]

Evert, Stefan (2014). **Distributional semantics in R with the wordspace package**. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: System Demonstrations*, pages 110–114, Dublin, Ireland. [PDF <https://www.stephanie-evert.de/PUB/Evert2014_wordspace.pdf>, wordspace homepage <<http://wordspace.r-forge.r-project.org/>>]

Evert, Stefan; Proisl, Thomas; Greiner, Paul; Kabashi, Besim (2014). **SentiKLUE: Updating a polarity classifier in 48 hours**. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval–2014)*, Dublin, Ireland. [PDF <https://www.stephanie-evert.de/PUB/EvertEtc2014_SentiKLUE.pdf>]

Lapesa, Gabriella and Evert, Stefan (2014). **NaDiR: Naive distributional response generation**. In *Proceedings of the 4th Workshop on Cognitive Aspects of the Lexicon (CogALex)*, pages 50–59, Dublin, Ireland. [PDF <<https://www.stephanie-evert.de/PUB/LapesaEvert2014.pdf>>]

Lapesa, Gabriella; Evert, Stefan; Schulte im Walde, Sabine (2014). **Contrasting syntagmatic and paradigmatic relations: Insights from distributional semantic models**. In *Proceedings of the Third Joint Conference on Lexical and Computational Semantics (*SEM 2014)*, pages 160–170, Dublin, Ireland. [PDF <<https://www.stephanie-evert.de/PUB/LapesaEtc2014.pdf>>]

Proisl, Thomas; Evert, Stefan; Greiner, Paul; Kabashi, Besim (2014). **SemantiKLUE: Robust semantic similarity at multiple levels using maximum weight matching**. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval–2014)*, Dublin, Ireland. [PDF <https://www.stephanie-evert.de/PUB/ProislEtc2014_SemantiKLUE.pdf>]

Schulze Wettendorf, Clemens; Jegan, Robin; Kärrner, Allan; Zerche, Julia; Plotnikova, Nataliia; Moreth, Julian; Schertl, Tamara; Obermeyer, Verena; Streil, Susanne; Willacker, Tamara; Evert, Stefan (2014). **SNAP: A multi-stage XML-pipeline for aspect based sentiment analysis**. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 578–584, Dublin, Ireland. [PDF <https://www.stephanie-evert.de/PUB/SchulzeWettendorfEtc2014_SNAP.pdf>]

2013

Greiner, Paul; Proisl, Thomas; Evert, Stefan; Kabashi, Besim (2013). **KLUE-CORE: A regression model of semantic textual similarity**. In *Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 1: Proceedings of the Main Conference and the Shared Task: Semantic Textual Similarity*, pages 181–186, Atlanta, Georgia, USA. [PDF <<https://www.stephanie-evert.de/PUB/GreinerEtc2013.pdf>>]

Lapesa, Gabriella and Evert, Stefan (2013). **Evaluating neighbor rank and distance measures as predictors of semantic priming**. In *Proceedings of the ACL Workshop on Cognitive Modeling and Computational Linguistics (CMCL 2013)*, Sofia, Bulgaria. [PDF <<https://www.stephanie-evert.de/PUB/LapesaEvert2013a.pdf>>]

Proisl, Thomas; Greiner, Paul; Evert, Stefan; Kabashi, Besim (2013). **KLUE: Simple and robust methods for polarity classification**. In *Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, pages 395–401, Atlanta, Georgia, USA. [PDF <<https://www.stephanie-evert.de/PUB/ProislEtc2013.pdf>>]

2012

Boleda, Gemma; Evert, Stefan; Gehrke, Berit; McNally, Louise (2012). **Adjectives as saturators vs. modifiers: Statistical evidence.** In M. Aloni, V. Kimmelman, F. Roelofsen, G. W. Sassoon, K. Schulz, and M. Westera (eds.), *Logic, Language and Meaning. Proceedings of the 18th Amsterdam Colloquium*, volume 7218 of *Lecture Notes in Computer Science*, pages 112–121. Springer, Berlin, Heidelberg. [PDF <<https://www.stephanie-evert.de/PUB/BoledaEtc2012.pdf>>]

2011

Evert, Stefan and Hardie, Andrew (2011). **Twenty-first century corpus workbench: Updating a query architecture for the new millennium.** In *Proceedings of the Corpus Linguistics 2011 Conference*, Birmingham, UK. [PDF <<https://www.stephanie-evert.de/PUB/EvertHardie2011.pdf>>]

Ebert, Cornelia; Evert, Stefan; Wilmes, Katharina (2011). **Focus marking via gestures.** In I. Reich *et al.* (eds.), *Proceedings of Sinn & Bedeutung 15*, Saarbrücken, Germany. Universaar – Saarland University Press. [PDF <<https://www.stephanie-evert.de/PUB/EbertEvertWilmes2011.pdf>>]

2010

Evert, Stefan (2010). **Google Web 1T5 n-grams made easy (but not for the computer).** In *Proceedings of the 6th Web as Corpus Workshop (WAC-6)*, Los Angeles, CA. [PDF <https://www.stephanie-evert.de/PUB/Evert2010_WAC6.pdf>]

2009

Giesbrecht, Eugenie and Evert, Stefan (2009). **Part-of-speech tagging – a solved task? An evaluation of POS taggers for the Web as corpus.** In I. Alegria, I. Leturia, and S. Sharoff, editors, *Proceedings of the 5th Web as Corpus Workshop (WAC5)*, San Sebastian, Spain. [PDF <https://www.stephanie-evert.de/PUB/GiesbrechtEvert2009_Tagging.pdf>]

2008

Evert, Stefan (2008). **A lightweight and efficient tool for cleaning Web pages.** In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, Marrakech, Morocco. [PDF <https://www.stephanie-evert.de/PUB/Evert2008_NCleaner.pdf>]

Evert, Stefan (2008). **A lexicographic evaluation of German adjective-noun collocations.** In *Proceedings of the LREC Workshop Towards a Shared Task for Multiword Expressions (MWE 2008)*, Marrakech, Morocco. [PDF <https://www.stephanie-evert.de/PUB/Evert2008_MWE_Resource.pdf>]

2007

Baroni, Marco and Evert, Stefan (2007). **Words and echoes: Assessing and mitigating the non-randomness problem in word frequency distribution modeling.** In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, pages 904–911, Prague, Czech Republic. [PDF <<https://www.stephanie-evert.de/PUB/BaroniEvert2007.pdf>>, talk slides (PDF) <<https://www.stephanie-evert.de/PUB/BaroniEvert2007.handout.pdf>>]

Bauer, Daniel; Degen, Judith; Deng, Xiaoye; Herger, Priska; Gasthaus, Jan; Giesbrecht, Eugenie; Jansen, Lina; Kalina, Christin; Krüger, Thorben; Martin, Robert; Schmidt, Martin; Scholler, Simon; Steger, Johannes; Stemle, Egon and Evert, Stefan (2007). **FIASCO: Filtering the Internet by automatic subtree classification, Osnabrück.** In C. Fairon, H. Naets, A. Kilgarriff, and G.–M. de Schryver (eds.), *Building and Exploring Web Corpora: Proceedings of the 3rd Web as Corpus Workshop (WAC3), incorporating CLEANVAL*, pages 111–121, Louvain-la-Neuve, Belgium. [PDF <https://www.stephanie-evert.de/PUB/BauerEtc2007_FIASCO.pdf>]

Evert, Stefan (2007). **StupidOS: A high-precision approach to boilerplate removal.** In C. Fairon, H. Naets, A. Kilgarriff, and G.–M. de Schryver (eds.), *Building and Exploring Web Corpora: Proceedings of the 3rd Web as Corpus Workshop (WAC3), incorporating CLEANVAL*, pages 123–133, Louvain-la-Neuve, Belgium. [PDF <https://www.stephanie-evert.de/PUB/Evert2007_StupidOS.pdf>]

Evert, Stefan and Baroni, Marco (2007). **zipfR: Word frequency distributions in R.** In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, Posters and Demonstrations*

Session, pages 29–32, Prague, Czech Republic. [PDF <<https://www.stephanie-evert.de/PUB/EvertBaroni2007.pdf>>]

Lüdeling, Anke, Evert, Stefan, and Baroni, Marco (2007). **Using Web data for linguistic purposes.** In M. Hundt, N. Nesselhauf, and C. Biewer, editors, *Corpus Linguistics and the Web*, volume 59 of *Language and Computers – Studies in Practical Linguistics*, pages 7–24. Rodopi, Amsterdam, New York. [manuscript (PDF) <<https://www.stephanie-evert.de/PUB/LuedelingEvertBaroni2005.pdf>>]

Michelbacher, Lukas, Evert, Stefan, and Schätze, Hinrich (2007). **Asymmetric association measures.** In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2007)*, Borovets, Bulgaria. [PDF <<https://www.stephanie-evert.de/PUB/MichelbacherEtc2007.pdf>>]

2006

Bernardini, Silvia, Baroni, Marco, and Evert, Stefan (2006). **A WaCky introduction.** In M. Baroni and S. Bernardini, editors, *Wacky! Working papers on the Web as Corpus*, pages 9–40. GEDIT, Bologna. [<<http://wackybook.sslmit.unibo.it/>>]

Hoffmann, Sebastian and Evert, Stefan (2006). **BNCweb (CQP-edition): The marriage of two corpus tools.** In S. Braun, K. Kohn, and J. Mukherjee (eds.), *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, volume 3 of *English Corpus Linguistics*, pages 177–195. Peter Lang, Frankfurt am Main. [PDF <<https://www.stephanie-evert.de/PUB/HoffmannEvert2006.pdf>>]

2005

Baroni, Marco and Evert, Stefan (2005). **Testing the extrapolation quality of word frequency models.** In P. Danielsson and M. Wagenmakers (eds.), *Proceedings of Corpus Linguistics 2005*, volume 1 of *The Corpus Linguistics Conference Series*. ISSN 1747–9398. [PDF <<https://www.stephanie-evert.de/PUB/EvertBaroni2005.pdf>>]

Evert, Stefan and Schoenenberger, Manuela (2005). **Separating the sheep from the goats: Clarifying corpus content using XML.** In P. Danielsson and M. Wagenmakers (eds.), *Proceedings of Corpus Linguistics 2005*, volume 1 of *The Corpus Linguistics Conference Series*. ISSN 1747–9398. [PDF <<https://www.stephanie-evert.de/PUB/EvertSchoenenberger2005.pdf>>]

Krenn, Brigitte and Evert, Stefan (2005). **Separating the wheat from the chaff: Corpus-driven evaluation of statistical association measures for collocation extraction.** In B. Fisseni, H.–C. Schmitz, B. Schröder, and P. Wagner (eds.), *Sprachtechnologie, mobile Kommunikation und linguistische Ressourcen. Beiträge zur GLDV-Tagung 2005 in Bonn*, volume 8 of *Computer Studies in Language and Speech*, pages 104–117. Peter Lang, Frankfurt am Main. [PDF <<https://www.stephanie-evert.de/PUB/KrennEvert2005.pdf>>]

Lüdeling, Anke and Evert, Stefan (2005). **The emergence of productive non-medical –itis. Corpus Evidence and qualitative analysis.** In: Kepser, Stephan and Reis, Marga (eds.), *Linguistic Evidence. Empirical, Theoretical, and Computational Perspectives*, pages 351–370, Mouton de Gruyter, Berlin, New York. [manuscript (PDF) <<https://www.stephanie-evert.de/PUB/LuedelingEvert2005.pdf>>]

2004

Evert, Stefan (2004a). **A simple LNRE model for random character sequences.** In *Proceedings of the 7^{èmes Journées Internationales d'Analyse Statistique des Données Textuelles}*, pages 411–422, Louvain-la-Neuve, Belgium. [PDF <<https://www.stephanie-evert.de/PUB/Evert2004a.pdf>>]

Evert, Stefan (2004b). **The statistical analysis of morphosyntactic distributions.** In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*, pages 1539–1542, Lisbon, Portugal. [PDF <<https://www.stephanie-evert.de/PUB/Evert2004b.pdf>>]

Evert, Stefan (2004c). **Significance tests for the evaluation of ranking methods.** In *Proceedings of the 20th International Conference on Computational Linguistics (Coling 2004)*, pages 945–951, Geneva, Switzerland. [PDF <<https://www.stephanie-evert.de/PUB/Evert2004c.pdf>>]

Evert, Stefan; Heid, Ulrich; Spranger, Kristina (2004). **Identifying morphosyntactic preferences in**

collocations. In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*, pages 907–910, Lisbon, Portugal. [PDF <<https://www.stephanie-evert.de/PUB/EvertHeidSpranger2004.pdf>>]

Evert, Stefan; Heid, Ulrich; SÄuberlich, Bettina; Debus-Gregor, Esther; Scholze-Stubenrecht, Werner (2004). **Supporting corpus-based dictionary updating.** In *Proceedings of the 11th Euralex International Congress*, pages 255–264, Lorient, France. [PDF <<https://www.stephanie-evert.de/PUB/EvertEtc2004.pdf>>]

Krenn, Brigitte; Evert, Stefan; Zinsmeister, Heike (2004). **Determining intercoder agreement for a collocation identification task.** In *Proceedings of KONVENS 2004*, pages 89–96, Vienna, Austria. [PDF <<https://www.stephanie-evert.de/PUB/KrennEvertZinsmeister2004.pdf>>]

LÄdeling, Anke and Evert, Stefan (2004). **The emergence of productive non-medical –itis: corpus evidence and qualitative analysis.** In *Proceedings of the First International Conference on Linguistic Evidence*, pages 91–95, TÄbingen, Germany. [PDF <<https://www.stephanie-evert.de/PUB/LuedelingEvert2004.pdf>>]

2003

Evert, Stefan and Kermes, Hannah (2003a). **Experiments on candidate data for collocation extraction.** In *Companion Volume to the Proceedings of the 10th Conference of The European Chapter of the Association for Computational Linguistics*, pages 83–86. [PDF <<https://www.stephanie-evert.de/PUB/EvertKermes2003a.pdf>>]

Evert, Stefan and Kermes, Hannah (2003b). **Annotation, storage, and retrieval of mildly recursive structures.** In K. Simov and P. Osenova (eds.), *Proceedings of the Workshop on Shallow Processing of Large Corpora (SProLaC 2003)*, pages 23–33, Lancaster, UK. [PDF <<https://www.stephanie-evert.de/PUB/EvertKermes2003b.pdf>>]

Carletta, Jean; Kilgour, Jonathan; O’Donnell, Timothy; Evert, Stefan; Voormann, Holger (2003). **The NITE object model library for handling structured linguistic annotation on multimodal data sets.** In *Proceedings of the EACL Workshop on Language Technology and the Semantic Web (3rd Workshop on NLP and XML, NLPXML–2003)*, pages 17–24, Budapest, Hungary. [PDF <<https://www.stephanie-evert.de/PUB/CarlettaEtc2003.pdf>>]

Kermes, Hannah and Evert, Stefan (2003). **Text analysis meets corpus linguistics.** In D. Archer, P. Rayson, A. Wilson, and T. McEnery (eds.), *Proceedings of the Corpus Linguistics 2003 Conference*, pages 402–411. UCREL. [PDF <<https://www.stephanie-evert.de/PUB/KermesEvert2003.pdf>>]

LÄdeling, Anke and Evert, Stefan (2003). **Linguistic experience and productivity: corpus evidence for fine-grained distinctions.** In D. Archer, P. Rayson, A. Wilson, and T. McEnery (eds.), *Proceedings of the Corpus Linguistics 2003 Conference*, pages 475–483. UCREL. [PDF <<https://www.stephanie-evert.de/PUB/LuedelingEvert2003.pdf>>]

2002

Kermes, Hannah and Evert, Stefan (2002). **YAC – a recursive chunker for unrestricted German text.** In M. G. Rodriguez and C. P. Araujo (eds.), *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC)*, volume V, pages 1805–1812, Las Palmas, Spain.

2001

Evert, Stefan and Krenn, Brigitte (2001). **Methods for the qualitative evaluation of lexical association measures.** In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*, pages 188–195, Toulouse, France. [PDF <<https://www.stephanie-evert.de/PUB/EvertKrenn2001.pdf>>, colour plots <<http://www.collocations.de/EK/ACL01/figures-ek.html>>]

Evert, Stefan and LÄdeling, Anke (2001). **Measuring morphological productivity: Is automatic preprocessing sufficient?** In P. Rayson, A. Wilson, T. McEnery, A. Hardie, and S. Khoja (eds.), *Proceedings of the Corpus Linguistics 2001 Conference*, pages 167–175, Lancaster. UCREL. [PDF <<https://www.stephanie-evert.de/PUB/EvertLuedeling2001.pdf>>]

Kermes, Hannah and Evert, Stefan (2001). **Exploiting large corpora: A circular process of partial syntactic analysis, corpus query and extraction of lexicographic information.** In P. Rayson, A. Wilson, T. McEnery, A. Hardie, and S. Khoja (eds.), *Proceedings of the Corpus Linguistics 2001 Conference*, pages 332–340, Lancaster. UCREL. [PDF <<https://www.stephanie-evert.de/PUB/KermesEvert2001.pdf>>]

Krenn, Brigitte and Evert, Stefan (2001). **Can we do better than frequency? A case study on extracting PP-verb collocations.** In *Proceedings of the ACL Workshop on Collocations*, pages 39–46, Toulouse, France. [PDF <<https://www.stephanie-evert.de/PUB/KrennEvert2001.pdf>>, colour plots <<http://www.collocations.de/EK/ACL01/figures-ke.html>>]

2000

Evert, Stefan; Heid, Ulrich; Lezius, Wolfgang (2000). **Methoden zum Vergleich von Signifikanzmaßen zur Kollokationsidentifikation.** In W. Zählke and E. G. Schukat-Talamazzini (eds.), *KONVENS–2000 Sprachkommunikation*, pages 215–220. VDE-Verlag. [PDF <<https://www.stephanie-evert.de/PUB/EvertHeidLezius2000.pdf>>]

Berman, Steve; Evert, Stefan; Heid, Ulrich (2000). **Searchable metaspaces.** In *Proceedings of the EAGLES/ISLE Workshop on Metadata*, Athens, Greece. [PDF <<https://www.stephanie-evert.de/PUB/BermanEvertHeid2000.pdf>>]

Heid, Ulrich; Evert, Stefan; Docherty, Vincent; Worsch, Wolfgang; Wermke, Matthias (2000). **A data collection for semi-automatic corpus-based updating of dictionaries.** In U. Heid, S. Evert, E. Lehmann, and C. Rohrer (eds.), *Proceedings of the 9th EURALEX International Congress*, pages 183–195.

Lüdeling, Anke; Evert, Stefan; Heid, Ulrich (2000). **On measuring morphological productivity.** In W. Zählke and E. G. Schukat-Talamazzini (eds.), *KONVENS–2000 Sprachkommunikation*, pages 57–61. VDE-Verlag. [PDF <<https://www.stephanie-evert.de/PUB/LuedelingEvertHeid2000.pdf>>]

Edited Volumes

Adrian, Axel; Evert, Stephanie; Kohlhase, Michael; Zwickel, Martin (eds.) (2022). **Digitalisierung von Zivilprozess und Rechtsdurchsetzung.** Number 284 in Schriften zum Prozessrecht. Duncker & Humblot, Berlin.

Griebel, Tim; Evert, Stefan; Heinrich, Philipp (eds.) (2020). **Multimodal Approaches to Media Discourses: Reconstructing the Age of Austerity in the United Kingdom.** Routledge Studies in Multimodality. Routledge, Abingdon. [publisher homepage <<https://www.routledge.com/Multimodal-Approaches-to-Media-Discourses-Reconstructing-the-Age-of-Austerity/Griebel-Evert-Heinrich/p/book/9780367332891>>]

Rayson, Paul; Villada Moirán, Begoña; Sharoff, Serge; Piao, Scott; Evert, Stefan (eds.) (2010). Special issue on **Multiword expressions: hard going or plain sailing?** *International Journal of Language Resources and Evaluation* 44(1&2). [CFP <<http://ucrel.lancs.ac.uk/publications/LREJournalMWecfp.pdf>>, Springer Link <<https://link.springer.com/journal/10579/volumes-and-issues/44-1>>]

Heid, Ulrich; Evert, Stefan; Lehmann, Egbert; Rohrer, Christian (eds.) (2000). **Proceedings of the 9th EURALEX International Congress**, Stuttgart, Germany.

Other Publications

Evert, Stephanie; Frenken, Florian; Neumann, Stella; Schneider, Gerold (2024). **On the replicability of Geometric Multivariate Analysis: Detecting cross-cultural differences in register variation across varieties of English.** Presentation at the *ICAME 45 Conference*, Vigo, Spain. [osf.io/8jzms <<https://osf.io/8jzms/>>]

Blombach, Andreas; Proisl, Thomas; Evert, Stefan; Heinrich, Philipp; Dykes, Natalie (2021). **Into the Perryverse: A CL journey to the realm of lexical complexity.** Presentation at the *ICAME 2021 Workshop To boldly go: Corpus approaches to the language of Science Fiction*, Dortmund, Germany.

Dykes, Natalie; Heinrich, Philipp; Evert, Stefan (2021). **Keywords gone viral: Exploring keyness techniques for corpus-based discourse analysis in German tweets.** Presentation at the *Corpus Linguistics 2021 Conference*, Limerick / online.

Evert, Stefan & Hardie, Andrew (2021). **Ziggurat 0.1: A next-generation system for modelling, storing,**

and retrieving corpus (and other) data. Presentation at the *Corpus Linguistics 2021 Conference*, Limerick / online.

Heinrich, Philipp; Dykes, Natalie; Evert, Stefan (2021). **Annotator agreement in the anonymization of court decisions.** Presentation at the *Corpus Linguistics 2021 Conference*, Limerick/online. [slides <<http://corpora.linguistik.uni-erlangen.de/data/Paper%20377%20Presentation%20Philipp%20Heinrich.pdf>>, video <<http://corpora.linguistik.uni-erlangen.de/data/Paper%20377%20Philipp%20Heinrich.mp4>>]

Diwersy, Sascha; Evert, Stefan; Heinrich, Philipp; Proisl, Thomas (2019). **Means of productivity â On the statistical modelling of the restrictedness of lexico-grammatical patterns.** Presentation at *EUROPHRAS 2019: Productive Patterns in Phraseology* <<https://europhrassantiago1.wixsite.com/europhras-en/copia-de-call-for-papers-abstracts>>, Santiago de Compostela, Spain. [abstract <https://www.stephanie-evert.de/PUB/DiwersyEtc2019_EUROPHRAS_abstract.pdf>, slides <https://www.stephanie-evert.de/PUB/DiwersyEtc2019_EUROPHRAS_slides.pdf>]

Evert, Stefan and Heid, Ulrich (2019). **Between collocation and construction: Lexical preferences in non-idiomatic word combinations.** Presentation at *EUROPHRAS 2019: Productive Patterns in Phraseology* <<https://europhrassantiago1.wixsite.com/europhras-en/copia-de-call-for-papers-abstracts>>, Santiago de Compostela, Spain. [abstract <https://www.stephanie-evert.de/PUB/EvertHeid2019_EUROPHRAS_abstract.pdf>, slides <https://www.stephanie-evert.de/PUB/EvertHeid2019_EUROPHRAS_slides.pdf>]

Evert, Stefan; Proisl, Thomas; Uhrig, Peter; Khokhlova, Maria (2018). **Contrastive collocation analysis â A comparison of association measures across three different languages using dependency-parsed corpora.** Presentation at the *XVIIIth EURALEX International Congress*, Ljubljana, Slovenia. [abstract <https://www.stephanie-evert.de/PUB/EvertEtc2018_Euralex_abstract.pdf>, slides <https://www.stephanie-evert.de/PUB/EvertEtc2018_Euralex_slides.pdf>]

Evert, Stefan; Dykes, Natalie; Peters, Joachim (2018). **A quantitative evaluation of keyword measures for corpus-based discourse analysis.** Presentation at the *Corpora & Discourse International Conference (CAD 2018)*, Lancaster, UK. [abstract <https://www.stephanie-evert.de/PUB/EvertEtc2018_CAD_abstract.pdf>, slides <https://www.stephanie-evert.de/PUB/EvertEtc2018_CAD_slides.pdf>]

Proisl, Thomas and Evert, Stefan (2018). **Delta vs. N-Gram-Tracing: Wie robust ist die Autorschaftsattribuierung?** Poster at the *DHd 2018 Conference*, Cologne, Germany. [poster <https://www.stephanie-evert.de/PUB/ProislEvert2018_DHd_poster.pdf>]

Evert, Stefan (2017a). **Measures of productivity and lexical diversity.** Poster at the *ICAME 38 Conference*, Prague, Czech Republic. [abstract <https://www.stephanie-evert.de/PUB/Evert2017_ICAME.pdf>, poster <https://www.stephanie-evert.de/PUB/Evert2017_ICAME_poster.pdf>] â *best poster award*

Evert, Stefan (2017b). **Making sense of multivariate analyses of linguistic variation.** Poster at the *Corpus Linguistics 2017 Conference*, Birmingham, UK. [abstract <https://www.stephanie-evert.de/PUB/Evert2017CL/Evert2017CL_Abstract.pdf>, poster <https://www.stephanie-evert.de/PUB/Evert2017CL/Evert2017CL_Poster.pdf>, additional material <<https://www.stephanie-evert.de/PUB/Evert2017CL/>>]

Evert, Stefan; Heinrich, Philipp; SchÄrfer, Fabian (2017). **Social Bots in Japan's 2014 General Election: Preliminary Results from a Corpus-Linguistic and Qualitative Study of Computational Propaganda on Twitter.** Presentation at the *International Conference on Computational Social Science (IC2S2 2017)*, Cologne, Germany. [abstract <https://www.stephanie-evert.de/PUB/EvertHeinrichSchaefer2017_abstract.pdf>]

Neumann, Stella; Evert, Stefan; De Sutter, Gert (2017). **Register-specific interference in translation.** Presentation at the *Annual Meeting of the German Linguistics Association (DGfS 2017)*, SaarbrÄcken, Germany. [abstract <https://www.stephanie-evert.de/PUB/NeumannEvertDeSutter2017_abstract.pdf>]

slides <https://www.stephanie-evert.de/PUB/NeumannEvertDeSutter2017_slides.pdf>]

Evert, Stefan; Jannidis, Fotis; Proisl, Thomas; Vitt, Thorsten; SchÄ¼ch, Christof; PielstrÄ¼m, Steffen; Reger, Isabella (2016). **Outliers or Key Profiles? Understanding Distance Measures for Authorship Attribution.** Presentation at *Digital Humanities 2016*, KrakÅ³w, Poland. [abstract <<http://dh2016.adho.org/abstracts/253>>]

Evert, Stefan; Schneider, Gerold; Brezina, Vaclav; Gries, Stefan Th.; Lijffijt, Jeffrey; Rayson, Paul; Wallis, Sean; Hardie, Andrew (2015). **Corpus statistics: key issues and controversies.** Panel discussion at *Corpus Linguistics 2015*, Lancaster, UK. [abstract <https://www.stephanie-evert.de/PUB/EvertSchneiderEtc2015_abstract.pdf>]

Evert, Stefan; Proisl, Thomas; SchÄ¼ch, Christof; Jannidis, Fotis; PielstrÄ¼m, Steffen; Vitt, Thorsten (2015). **Explaining Delta, or: How do distance measures for authorship attribution work?** Presentation at *Corpus Linguistics 2015*, Lancaster, UK. [abstract <https://www.stephanie-evert.de/PUB/EvertProislEtc2015_abstract.pdf>, slides <https://www.stephanie-evert.de/PUB/EvertProislEtc2015_CL_Slides.pdf>]

Bartsch, Sabine; Evert, Stefan; Proisl, Thomas; Uhrig, Peter (2015). **(Association) measure for measure.** Presentation at *ICAME 36*, Trier, Germany.

Lapasa, Gabriella; Schulte im Walde, Sabine; Evert, Stefan (2014). **Judging Paradigmatic Relations: A Collection of Ratings for English.** Poster at *Architecture and Mechanisms of Language Processing (AMLAP-2014)*, Edinburgh, UK. [poster (PDF) <http://cogsci.uni-osnabrueck.de/~glapasa/materials/poster_AMLAP_14_Edinburgh_2.pdf>]

Evert, Stefan and Neumann, Stella (2013). **The impact of translation direction on the characteristics of translated texts: a multivariate analysis for English and German.** Presentation at the *46th Annual Meeting of the Societas Linguistica Europaea*, Split, Croatia.

Bartsch, Sabine and Evert, Stefan (2013b). **Exploring the Firthian notion of collocation.** Presentation at *Corpus Linguistics 2013*, Lancaster, UK.

Evert, Stefan; Schneider, Gerold; Lehmann, Hans Martin (2013). **Statistical modelling of natural language for descriptive linguistics.** Presentation at *Corpus Linguistics 2013*, Lancaster, UK.

Lapasa, Gabriella and Evert, Stefan (2013b). **Thematic Roles and Semantic Space. Insights from Distributional Semantic Models.** Presentation at *Quantitative Investigations in Theoretical Linguistics (QITL-5)*, Leuven, Belgium. [abstract (PDF) <http://cogsci.uni-osnabrueck.de/~glapasa/materials/QITL5-abstract_submitted.pdf>, slides (PDF) <http://cogsci.uni-osnabrueck.de/~glapasa/materials/presentation_qitl.pdf>]

Lapasa, Gabriella and Evert, Stefan (2013a). **Item-based Prediction of Reaction Times in Priming: an Evaluation of Distributional Semantic Models.** Poster at *Architecture and Mechanisms of Language Processing (AMLAP-2013)*, Marseille, France. [poster (PDF) <http://cogsci.uni-osnabrueck.de/~glapasa/materials/poster_AMLAP.pdf>]

SÄ±nchez Marco, Cristina; MarÄ±n, Rafael; Evert, Stefan (2012). **The lexical extension of *estar* + participle through psychological verbs.** Presentation at the *Annual Meeting of the German Linguistics Association (DGfS 2012)*, Frankfurt, Germany.

SÄ±nchez Marco, Cristina; MarÄ±n, Rafael; Evert, Stefan (2012). **Measuring lexical extension: The case of Spanish *estar* + past participle.** Poster presentation at *Linguistic Evidence 2012*, TÄ¼bingen, Germany. [extended abstract (PDF) <<https://www.stephanie-evert.de/PUB/SanchezEtc2012.pdf>>]

Evert, Stefan (2011). **Quantitative measures of productivity and their significance.** Presentation at *Corpus Linguistics 2011*, Birmingham, UK. [handout (PDF) <https://www.stephanie-evert.de/PUB/Evert2011_productivity.handout.pdf>]

SÄ±nchez Marco, Cristina and Evert, Stefan (2011). **Measuring semantic change: The case of Spanish participial constructions.** Poster presentation at *QITL-4*, Berlin, Germany.

Evert, Stefan and Pipa, Gordon (2010). **Probability estimation of rare events in linguistics and**

computational neuroscience. Presentation at *KogWis 2010*, Potsdam, Germany. [abstract <https://www.stephanie-evert.de/PUB/EvertPipa2010_abstract.pdf>, handout <https://www.stephanie-evert.de/PUB/EvertPipa2010_handout.pdf>]

Pipa, Gordon and Evert, Stefan (2010). **Statistical models of non-randomness in natural language.** Presentation at *KogWis 2010*, Potsdam, Germany. [abstract (PDF) <https://www.stephanie-evert.de/PUB/PipaEvert2010_abstract.pdf>]

Evert, Stefan (2009). **Rethinking corpus frequencies.** Presentation at the *ICAME 30 Conference*, Lancaster, UK. [handout (PDF) <https://www.stephanie-evert.de/PUB/Evert2009_ICAME.handout.pdf>]

Evert, Stefan (2007). **Room for improvement? Upper limits on collocation extraction with statistical association measures.** Poster presentation at the *Computational Linguistics Poster Session at the Annual Meeting of the German Linguistics Association (DGfS 2007)*. [poster (PDF) <https://www.stephanie-evert.de/PUB/Evert2007_DGfS_poster.pdf>]

Evert, Stefan and Baroni, Marco (2006). **ZipfR: Working with words and other rare events in R.** Presentation at the *useR! 2006 Conference*, Vienna, Austria. [handout (PDF) <https://www.stephanie-evert.de/PUB/EvertBaroni2006_useR.pdf>]

Lüdeling, Anke; Baroni, Marco; Evert, Stefan (2006). **Need and competition in word formation and where to find data to study them.** Poster presentation at the *Second International Conference on Linguistic Evidence*, Tübingen, Germany. [abstract (PDF) <https://www.stephanie-evert.de/PUB/LuedelingBaroniEvert2006_LingEv.pdf>, poster (PDF) <https://www.stephanie-evert.de/PUB/LuedelingBaroniEvert2006_LingEv_poster.pdf>]

Lüdeling, Anke; Baroni, Marco; Evert, Stefan (2006). **Need and competition: Deconstructing quantitative productivity.** Presentation at the *Second Conference on Quantitative Investigations in Theoretical Linguistics (QITL-2)*, Osnabrück, Germany. [abstract (PDF) <https://www.stephanie-evert.de/PUB/LuedelingBaroniEvert2006_QITL.pdf>]

Evert, Stefan (2005). **Empirical research on association measures: The UCS toolkit.** Software demonstration at the *Phraseology 2005 Conference*, Louvain-la-Neuve, Belgium. [abstract (PDF) <<https://www.stephanie-evert.de/PUB/Evert2005phraseology.pdf>>]

Evert, Stefan and Krenn, Brigitte (2005). **Exploratory collocation extraction.** Presentation at the *Phraseology 2005 Conference*, Louvain-la-Neuve, Belgium. [abstract (PDF) <<https://www.stephanie-evert.de/PUB/EvertKrenn2005phraseology.pdf>>]

Evert, Stefan and Hoffmann, Sebastian (2005). **BNCweb (CQP edition): The marriage of two corpus tools.** Presentation at the *Corpus Linguistics 2005 Conference*, Birmingham, UK.

Evert, Stefan (2004d). **An on-line repository of association measures.** <<http://www.collocations.de/AM/>>.

Evert, Stefan; Carletta, Jean; O'Donnell, Timothy J.; Kilgour, Jonathan; Vögele, Andreas; Voormann, Holger (2003). **The NXT object model.** Technical report, IMS, University of Stuttgart. Version 2.1. [PDF <<https://www.stephanie-evert.de/PUB/NITE-NOM.pdf>>]

Evert, Stefan and Voormann, Holger (2003). **NQL – a query language for multi-modal language data.** Technical report, IMS, University of Stuttgart. Version 2.1. [PDF <<https://www.stephanie-evert.de/PUB/NITE-NQL.pdf>>]

Evert, Stefan and Kermes, Hannah (2002). **The influence of linguistic pre-processing on candidate data.** In *Workshop on Computational Approaches to Collocations*, Vienna, Austria.

Schneiderberger, Manuela and Evert, Stefan (2002). **The benefit of doubt.** Presentation at the *Workshop on Quantitative Investigations in Theoretical Linguistics (QITL)*, Osnabrück, Germany, October 2002. Slides can be downloaded from <<http://www.cogsci.uni-osnabrueck.de/qitl/>>.

Evert, Stefan (1999). **Das Verhalten von Lösungen der vektoriellen Helmholtzgleichung in Außenräumen für kleine Frequenzen unter elektrischen Randbedingungen.** Unpublished Diplom thesis, University of Stuttgart. [PDF <<https://www.stephanie-evert.de/PUB/Evert1999diplom.pdf>>]