

Dr. Roland Schäfer | Schriften

roland.schaefer@fu-berlin.de | Email
<http://rolandschaefer.net> | Web

Laufbahnarbeiten

Schäfer, Roland. (2018). *Probabilistic German Morphosyntax*. Berlin: Humboldt Universität. Kumulative Habilitationsschrift zur Erlangung der Lehrbefähigung für das Fach Germanistische und Allgemeine Sprachwissenschaft.

Schäfer, Roland. (2010). *Arguments and Adjuncts at the Syntax-Semantics Interface*. Göttingen: Georg-August Universität. Dissertation zur Erlangung des Doktorgrades im Fach Englische Philologie.

Schäfer, Roland. (2002). *Typologie des Wandels von Konsonantenverbindungen*. Marburg, Lahn: Philipps-Universität. Magisterarbeit im Fach Vergleichende Sprachwissenschaft.

Bücher

Schäfer, Roland. (2016). *Einführung in die grammatische Beschreibung des Deutschen*. Zweite, überarbeitete Auflage. Berlin: Language Science Press.

Schäfer, Roland & Felix Bildhauer. (2013). *Web Corpus Construction* (Synthesis Lectures on Human Language Technologies). San Francisco: Morgan & Claypool.

Artikel (Zeitschriften)

Schäfer, Roland & Elizabeth Pankratz. (2018, eingereicht). The plural interpretability of German linking elements. Submitted to *Morphology*.

Schäfer, Roland. (2018, angenommen). Abstractions and exemplars: the measure noun phrase alternation in German. *Cognitive Linguistics*.

Schäfer, Roland. (2017). Accurate and Efficient General-Purpose Boilerplate Detection for Crawled Web Corpora. *Language Resources and Evaluation* 51(3). 873–889.

- Schäfer, Roland & Ulrike Sayatz. (2017). Wieviel Grammatik braucht das Germanistikstudium? *Zeitschrift für germanistische Linguistik* 42(2). 221–255.
- Schäfer, Roland & Ulrike Sayatz. (2016). Punctuation and Syntactic Structure in “obwohl” and “weil” Clauses in Nonstandard Written German. *Written Language and Literacy* 19(2). 215–245.
- Schäfer, Roland. (2016). Prototype-Driven Alternations: the Case of German Weak Nouns. *Corpus Linguistics and Linguistic Theory*. Ahead of print (online).
- Schäfer, Roland & Ulrike Sayatz. (2014). Die Kurzformen des Indefinitartikels im Deutschen. *Zeitschrift für Sprachwissenschaft* 33(2). 215–250.
- Biemann, Chris, Felix Bildhauer, Stefan Evert, Dirk Goldhahn, Uwe Quasthoff, Roland Schäfer, Johannes Simon, Leonard Swiezinski & Torsten Zesch. (2013). Scalable Construction of High-Quality Web Corpora. *Journal for Language Technology and Computational Linguistics* 28(2). 23–60.

Artikel (Handbücher und Sammelbände)

- Schäfer, Roland. (2018, erscheint). Mixed Effects Regression Modelling. In Stefan Th. Gries & Magali Paquot (Hrsg.), *The Practical Handbook of Corpus Linguistics*. Berlin/Heidelberg: Springer.
- Bildhauer, Felix & Roland Schäfer. (2017). Induktive Topikmodellierung und extrinsische Topikdomänen. In Marek Konopka & Angelika Wöllstein (Hrsg.), *Grammatische Variation - empirische Zugänge und theoretische Modellierung* (Jahrbuch des Instituts für Deutsche Sprache 2016), 331–344. Berlin/Boston: De Gruyter.
- Job, Michael & Roland Schäfer. (2005). Ossetic. In Keith Brown (Hrsg.), *Encyclopedia of Language and Linguistics*, 2. Aufl., Bd. 10, 109–115. Oxford: Elsevier.

Artikel (qualitätsgesicherte Tagungsbände)

Anmerkung: In der Computerlinguistik werden kaum Zeitschriftenartikel und Sammelbände publiziert, und Beiträge in *Proceedings* sind die relevante qualitätsgesicherte Art der Publikation. Um trotzdem den Unterschied zur linguistischen Publikation in Zeit-

schriften im Bewerbungsbogen kenntlich zu machen, habe ich dort die folgenden Publikationen den Sammelbänden und Enzyklopädien zugerechnet.

Schäfer, Roland & Felix Bildhauer. (2016). Automatic Classification by Topic Domain for Meta Data Generation, Web Corpus Evaluation, and Corpus Comparison. In Paul C. Cook, Stefan Evert, Roland Schäfer & Egon Stemle (Hrsg.), *Proceedings of the 10th Web as Corpus Workshop (WAC-X)*, 1–6. Stroudsburg: Association for Computational Linguistics (ACL).

Schäfer, Roland. (2016a). On Bias-free Crawling and Representative Web Corpora. In Paul C. Cook, Stefan Evert, Roland Schäfer & Egon Stemle (Hrsg.), *Proceedings of the 10th Web as Corpus Workshop (WAC-X)*, 99–105. Stroudsburg: Association for Computational Linguistics (ACL).

Schäfer, Roland. (2016b). CommonCOW: Massively Huge Web Corpora from Common-Crawl Data and a Method to Distribute them Freely under Restrictive EU Copyright Laws. In Nicoletta Calzolari et al. (Hrsg.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, 4500–4504. Portorož, Slovenia: European Language Resources Association (ELRA).

Schäfer, Roland. (2015). Processing and Querying Large Web Corpora with the COW14 Architecture. In Piotr Bański, Hanno Biber, Evelyn Breiteneder, Marc Kupietz, Harald Lungen & Andreas Witt (Hrsg.), *Proceedings of Challenges in the Management of Large Corpora 3 (CMLC-3)*. Mannheim: Institut für Deutsche Sprache.

Schäfer, Roland, Adrien Barbaresi & Felix Bildhauer. (2014). Focused Web Corpus Crawling. In Felix Bildhauer & Roland Schäfer (Hrsg.), *Proceedings of the 9th Web as Corpus workshop (WAC-9)*, 9–15. Stroudsburg: Association for Computational Linguistics (ACL).

Schäfer, Roland, Adrien Barbaresi & Felix Bildhauer. (2013). The Good, the Bad, and the Hazy: Design Decisions in Web Corpus Construction. In Stefan Evert, Egon Stemle & Paul Rayson (Hrsg.), *Proceedings of the 8th Web as Corpus Workshop (WAC-8)*, 7–15. Stroudsburg: Association for Computational Linguistics (ACL).

Schäfer, Roland & Felix Bildhauer. (2012). Building Large Corpora from the Web Using a New Efficient Tool Chain. In Nicoletta Calzolari et al. (Hrsg.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 486–493. Istanbul, Turkey: European Language Resources Association (ELRA).

Schäfer, Roland. (2007). On Frequency Adjectives. In Estela Puig-Waldmüller (Hrsg.), *Proceedings of Sinn und Bedeutung 11*, 555–567. Universitat Pompeu Fabra. Barcelona.

Herausgeberschaften (qualitätsgesicherte Tagungsbände)

Cook, Paul C., Stefan Evert, Roland Schäfer & Egon Stemle (Hrsg.). (2016). *Proceedings of the 10th Web as Corpus workshop and the EmpiriST shared task*. Association for Computational Linguistics (ACL). Stroudsburg.

Bildhauer, Felix & Roland Schäfer (Hrsg.). (2014). *Proceedings of the 9th Web as Corpus workshop*. Association for Computational Linguistics (ACL). Stroudsburg.