

# Steven Moran — Curriculum Vitae

August 8, 2019

Department of Comparative Linguistics  
University of Zurich  
Plattenstrasse 54, Room 206  
8032 Zurich, Switzerland

steven.moran@uzh.ch  
+41 44 63 40221  
orcid.org/0000-0002-3969-6549

<http://www.comparativelinguistics.uzh.ch/moran.html>  
<https://scholar.google.com/citations?user=PpT0h08AAAAJ&hl=en>  
<https://github.com/bambooforest/>  
<https://www.linkedin.com/in/steven-moran-b59b3b3/>  
<https://twitter.com/stivits>

## EDUCATION

- 2012** Ph.D. Linguistics (Computational Linguistics Track), University of Washington  
Dissertation: *Phonetics Information Base and Lexicon*. [pdf]
- 2006** M.A. Linguistics, Eastern Michigan University  
Thesis: *A Grammatical Sketch of Western Sisaala*. [pdf]
- 2006** M.A. Certificate in Language Technology, Eastern Michigan University
- 2001** B.A. with honors in German, Linguistics and TESOL, Eastern Michigan University
- 2000** Sprachaufenthalt, Volkshochschule, Karlsruhe, Germany
- 1999** Austauschprogramm Deutsch in Graz, Graz, Austria

## ACADEMIC POSITIONS

- 2012–2019** **University of Zurich**  
Assistant and instructor in the Department of Comparative Linguistics. Postdoctoral researcher and technical project manager for the ACQDIV project.
- 2013–2014** **University of Marburg**  
Postdoctoral researcher and programmer in the Quantitative Language Comparison Research Unit in the Forschungszentrum Deutscher Sprachatlas.
- 2010–2013** **University of Munich**  
Researcher and programmer in the Quantitative Language Comparison Research Unit.
- 2005–2010** **University of Washington**  
Technical project manager for Phonetics Information Base and Lexicon (2009; UW Royalty Research Fund); Research assistant on Implementing the GOLD Community of Practice: Laying the Foundations for a Linguistics Cyberinfrastructure (2008; National Science Foundation (NSF) Grant BCS-0720670); User interface designer for UW Information Technology (2007–2008); Web programmer and Computer Assisted Language Learning specialist for UW Language Learning Center (2005–2007)
- 2001–2005** **Eastern Michigan University**  
Project manager, web programmer and researcher on E-MELD (2003–2005; NSF Grant BCS-0094934); editor and research assistant (2001–2003; The Linguist List)

## REFEREED JOURNAL ARTICLES

- 2019** Moran, Steven\*, Damián E. Blasi\*, Scott R. Moisik, Paul Widmer, Dan Dediu and Balthasar Bickel. Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science*, 363, 6432. DOI: <http://dx.doi.org/10.1126/science.aav3218>. \*Shared first authorship.
- 2018** Moran, Steven, Damián E. Blasi, Robert Schikowski, Aylin C. Küntay, Barbara Pfeiler, Shanley Allen and Sabine Stoll. A universal cue for grammatical categories in the input to children: frequent frames. *Cognition*, 175, 131–140. DOI: <https://doi.org/10.1016/j.cognition.2018.02.005>.
- 2016** Moran, Steven. Commentary: Issues of time, tone, roots and replicability. *Journal of Language Evolution*, 1(1), 73–76. DOI: <https://doi.org/10.1093/jole/lzv011>
- 2015** Stoll, Sabine, Taras Zakharko, Steven Moran, Robert Schikowski and Balthasar Bickel. Syntactic mixing across generations in an environment of community-wide bilingualism. *Frontiers in Psychology*. Special Topic: Acquisition in Contact Research, Shannessy Carmel (ed). DOI: <https://doi.org/10.3389/fpsyg.2015.00082>.
- 2015** McCrae, John, Steven Moran, Sebastian Hellmann and Martin Brümmer. Multilingual Linked Open Data. *Semantic Web*, 6(4), 315–317. DOI: <https://doi.org/10.3233/SW-150178>. Online: [http://www.semantic-web-journal.net/system/files/swj1042\\_0.pdf](http://www.semantic-web-journal.net/system/files/swj1042_0.pdf).
- 2013** Moran, Steven and Jelena Prokić. Investigating the Relatedness of the Endangered Dogon Languages. *Literary And Linguistic Computing*, 28(4) 676–691. Oxford University Press. DOI: <https://doi.org/10.1093/llc/fqt061>
- 2012** Moran, Steven, Daniel McCloy and Richard Wright. Revisiting population size vs phoneme inventory size. *Language*, 88(4): 877–893. DOI: <https://doi.org/10.1353/lan.2012.0087>.
- 2012** Cysouw, Michael, Dan Dediu and Steven Moran. Comment on “Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa”. *Science*, 335(6069):657. DOI: <https://doi.org/10.1126/science.1208841>.

## BOOK

- 2018** Moran, Steven and Michael Cysouw. *The Unicode Cookbook for Linguists: Managing writing systems using orthography profiles*. Translation and Multilingual Natural Language Processing series in Language Science Press. DOI: <https://doi.org/10.5281/zenodo.1296780>. Online: <http://langsci-press.org/catalog/book/176>.

## BOOK CHAPTERS

- In press.** Moran, Steven and Christian Chiarcos. Linguistic Linked Open Data and Under-Resourced Languages: From Collection to Application. In Antonio Pareja-Lora, María Blume, Barbara C. Lust, and Christian Chiarcos (eds), *Development of Linguistic Linked Open Data Resources for Collaborative Data-Intensive Research in the Language Sciences*. MIT Press.

- 2018** Benitez-Burraco, Antonio and Steven Moran. The Adaptive Value of Languages: Non-Linguistic Causes of Language Diversity. In Benitez-Burraco, Antonio and Steven Moran (eds), *The Adaptive Value of Languages: Non-Linguistic Causes of Language Diversity*. Frontiers in Psychology. DOI: <https://doi.org/10.3389/fpsyg.2018.01827>.
- 2014** Moran, Steven and Damián Blasi. Cross-linguistic comparison of complexity measures in phonological systems. In Frederick J. Newmeyer and Laurel Preston (eds), *Measuring Grammatical Complexity*. Oxford University Press.
- 2013** Prokić, Jelena and Steven Moran. Black box approaches to genealogical classification and their shortcomings. In *Comparing Approaches to Measuring Linguistic Differences*, Lars Borin and Anju Saxena (eds), 429–446. Berlin, Boston: De Gruyter Mouton.
- 2013** Chiarcos, Christian, Steven Moran, Pablo N. Mendes, Sebastian Nordhoff, Richard Littauer. Building a Linked Open Data Cloud of Linguistic Resources: Motivations and Developments. In Iryna Gurevych and Jungi Kim (eds), *The People’s Web Meets NLP: Collaboratively Constructed Language Resources*. Springer, Berlin, Heidelberg. DOI: [https://doi.org/10.1007/978-3-642-35085-6\\_12](https://doi.org/10.1007/978-3-642-35085-6_12).
- 2012** Moran, Steven. Using Linked Data to Create a Typological Knowledge Base. In *Linked Data in Linguistics: Representing and Connecting Language Data and Language Metadata*, Christian Chiarcos, Sebastian Nordhoff and Sebastian Hellmann (eds). Springer, Heidelberg. DOI: [https://doi.org/10.1007/978-3-642-28249-2\\_13](https://doi.org/10.1007/978-3-642-28249-2_13). Online: [http://svn.aksw.org/papers/2012/LDL/ldl2012\\_proceedings/public/moran.pdf](http://svn.aksw.org/papers/2012/LDL/ldl2012_proceedings/public/moran.pdf).
- 2011** Moran, Steven. An Ontology for Accessing Transcription Systems. In *Language Resources & Evaluation: Special Issue on African Language Technology*, Guy De Pauw, Laurette Pretorius, Gilles-Maurice de Schryver and Lori Levin (eds). DOI: <https://doi.org/10.1007/s10579-011-9158-8>. Online: <https://link.springer.com/article/10.1007/s10579-011-9158-8>.
- 2010** Boynton, Jessica, Steven Moran, Helen Aristar-Dry and Anthony Aristar. Using the E-MELD School of Best Practices to Create Lasting Digital Documentation. In *Language Documentation: Practice and Values*, Lenore A. Grenoble and N. Louanna Furbee (eds), 133–146. John Benjamins. DOI: <https://doi.org/10.1075/z.158.14boy>.

## CONFERENCE & WORKSHOP PROCEEDINGS (PEER-REVIEWED)

- 2019** Loukatou, Georgia R., Steven Moran, Damián E. Blasi, Sabine Stoll and Alejandrina Cristia. Is word segmentation child’s play in all languages? In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 3931–3937. July 28 – August 2, Florence, Italy. Online: <https://www.aclweb.org/anthology/P19-1383>.
- 2019** Moran, Steven, Nicholas A. Lester, Heath Gordon, Aylin Küntay, Barbara Pfeiler, Shanley Allen and Sabine Stoll. Variation Sets in Maximally Diverse Languages. In Proceedings 43rd annual Boston University Conference on Language Development (BUCLD). November 2–4, Boston, USA. Cascadilla Press. Online: <http://www.lingref.com/bucld/43/BUCLD43-34.pdf>
- 2018** Moran, Steven, Danica Pajović and Sabine Stoll. Cross-linguistically Small World Networks are Ubiquitous in Child-directed Speech. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 4100–4105. May 7–12,

- Miyazaki, Japan. Online: <http://www.lrec-conf.org/proceedings/lrec2018/pdf/887.pdf>
- 2018** Marsico, Egidio, Sebastien Flavier, Annemarie Verkerk and Steven Moran. BDPROTO: A Database of Phonological Inventories from Ancient and Reconstructed Languages. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 1654–1658. May 7–12, Miyazaki, Japan. Online: <http://www.lrec-conf.org/proceedings/lrec2018/pdf/534.pdf>
- 2018** McNew, Garland, Curdin Derungs and Steven Moran. Towards faithfully visualizing global linguistic diversity. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 805–809. May 7–12, Miyazaki, Japan. Online: <http://www.lrec-conf.org/proceedings/lrec2018/pdf/813.pdf>
- 2018** Moran, Steven and Annemarie Verkerk. Differential rates of change in consonant and vowel systems. In C. Cuskley and M. Flaherty and H. Little and Luke McCrohon and A. Ravignani and T. Verhoef (eds.), The Evolution of Language: Proceedings of the 12th International Conference (EVLANGXII). April 16–19, Toruń, Poland. DOI: <https://doi.org/10.12775/3991-1.077>. Online: <http://evolang.org/torun/proceedings/papertemplate.html?p=98>.
- 2016** Moran, Steven, Robert Schikowski, Danica Pajović, Cazim Hysi and Sabine Stoll. The AC-QDIV Database: Min(d)ing the Ambient Language. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), 4423–4429. May 23–28, Portorož, Slovenia. Online: [http://www.lrec-conf.org/proceedings/lrec2016/pdf/1198\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2016/pdf/1198_Paper.pdf).
- 2016** McCrae, John Philip, Christian Chiarcos, Francis Bond, Philipp Cimiano, Thierry Declerck, Gerard de Melo, Jorge Gracia, Sebastian Hellmann, Bettina Klimek, Steven Moran, Petya Osenova, Antonio Pareja-Lora and Jonathan Pool. The Open Linguistics Working Group: Developing the Linguistic Linked Open Data Cloud. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), 2435–2441. May 23–28, Portorož, Slovenia. Online: [http://www.lrec-conf.org/proceedings/lrec2016/pdf/851\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2016/pdf/851_Paper.pdf).
- 2014** Moran, Steven and Michael Cysouw. Typology with Graphs and Matrices. In Proceedings of the Third Workshop on Linked Data in Linguistics, Multilingual Knowledge Resources and Natural Language Processing, 55–60. Co-located with LREC 2014. May 27, Reykjavik, Iceland. Online: <http://www.lrec-conf.org/proceedings/lrec2014/workshops/LREC2014Workshop-LDL2014%20Proceedings.pdf>.
- 2014** Goldman, Jean-Philippe, Adrian Leemann, Marie-José Kolly, Ingrid Hove, Ibrahim Almajai, Volker Dellwo and Steven Moran. A Crowdsourcing Smartphone Application for Swiss German: Putting language documentation in the hands of the users. In Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 14), Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk and Stelios Piperidis (eds.). May 26–31, Reykjavik, Iceland. Online: [http://www.lrec-conf.org/proceedings/lrec2014/pdf/214\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2014/pdf/214_Paper.pdf).
- 2013** Moran, Steven and Martin Brümmer. Lemon-aid: using Lemon to aid quantitative historical linguistic analysis. In Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language

- data. Christian Chiarcos, Philipp Cimiano, Thierry Declerck and John P. McCrae (eds). Association for Computational Linguistics. September, 23. Pisa, Italy. Online: <http://aclweb.org/anthology/W/W13/W13-5505.pdf>.
- 2013** Littauer, Richard, Boris Villazon-Terrazas and Steven Moran. Linguistic Resources Enhanced with Geospatial Information. In *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*. Christian Chiarcos, Philipp Cimiano, Thierry Declerck and John P. McCrae (eds). Association for Computational Linguistics. September, 23. Pisa, Italy. Online: <http://aclweb.org/anthology/W/W13/W13-5508.pdf>.
- 2013** List, Johann-Mattis and Steven Moran. An Open Source Toolkit for Quantitative Historical Linguistics. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, pg 13–18, Sofia, Bulgaria, August 4–9, 2013. Online: <http://aclweb.org/anthology/P/P13/P13-4003.pdf>.
- 2012** Chiarcos, Christian, Sebastian Hellmann, Sebastian Nordhoff, Steven Moran, Richard Littauer, Judith Eckle-Kohler, Iryna Gurevych, Silvana Hartmann, Michael Matuschek and Christian M. Meyer. ‘The Open Linguistics Working Group’. In *Proceedings from LREC 2012*, pg 3603–3610, Istanbul, Turkey, May 23–25. Online: [http://www.lrec-conf.org/proceedings/lrec2012/pdf/912\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2012/pdf/912_Paper.pdf).
- 2012** Moran, Steven, Daniel McCloy and Richard Wright. ‘Revisiting the population vs phoneme-inventory correlation’. In *Extended abstracts in eLanguage from LSA 2012*. DOI: <http://dx.doi.org/10.3765/exabs.v0i0.601>.
- 2009** Moran, Steven. An Ontology for Accessing Transcription Systems. In *Proceedings from the African Language Technology Workshop at the European Association for Computational Linguistics*. Athens, Greece, March 31. Online: <http://www.aclweb.org/anthology/W09-0716>.
- 2008** Farrar, Scott and Steven Moran. The e-Linguistics Toolkit. In *Proceedings e-Humanities — an emerging discipline: Workshop in the 4th IEEE International Conference on e-Science*. Indianapolis, Indiana. Online: <http://www.zora.uzh.ch/id/eprint/84686/1/FarrarMoran2008.pdf>
- 2008** Trippel, Thorsten, Michael Maxwell, Greville Corbett, Cambell Prince, Christopher Manning, Stephen Grimes and Steven Moran. Lexicon Schemas and Related Data Models: When Standards Meet Users. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC)*. Marrakech, Morocco. Online: [http://www.lrec-conf.org/proceedings/lrec2008/pdf/812\\_paper.pdf](http://www.lrec-conf.org/proceedings/lrec2008/pdf/812_paper.pdf).
- 2008** Fox, Naomi and Steven Moran. Non-traditional Dissemination of Fieldwork: From the Field to the World. In *Proceedings from the 20th Northwest Linguistics Conference (NWLC)*, Steven Moran, Darren S. Tanner and Michael Scanlon (eds). Seattle, Washington. Online: [https://depts.washington.edu/uwwpl/vol27/uwwpl\\_vol27\\_2008.pdf](https://depts.washington.edu/uwwpl/vol27/uwwpl_vol27_2008.pdf).
- 2006** Boynton, Jessica, Steven Moran, Anthony Aristar and Helen Aristar-Dry. E-MELD and the School of Best Practices: An Ongoing Community Effort. In *Proceedings from Sustainable Data from Digital Fieldwork*. Sydney University Press. Online: <https://ses.library.usyd.edu.au/bitstream/2123/1296/1/9FH2006.pdf>.
- 2006** Moran, Steven. From Field Methods to Best Practice: A Case Study of Western Sisaala.

E-MELD School of Best Practices. <http://emeld.org/school/case/sisaala/>

- 2005** Moran, Steven. The E-MELD School of Best Practices in Digital Language Documentation. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 41(2), 71–80. Chicago Linguistic Society.
- 2004** Ferrara, Marisa and Steven Moran. Review of DBMS for Linguistic Purposes. In *Proceedings from the E-MELD 2004: Conference on Linguistic Databases and Best Practices*. <http://emeld.org/workshop/2004/papers.html>

## EDITED VOLUMES

- 2018** Benitez-Burraco, Antonio and Steven Moran (eds). The Adaptive Value of Languages: Non-Linguistic Causes of Language Diversity. *Frontiers in Psychology*. Online: <http://journal.frontiersin.org/researchtopic/5478/the-adaptive-value-of-languages-non-linguistic-causes-of-language-diversity>
- 2015** Hellmann, Sebastian, Steven Moran, Martin Brömmmer and John McCrae (eds). Special Issue on Multilingual Linked Open Data (MLOD). *Semantic Web*, 6(4). Online: <https://content.iospress.com/journals/semantic-web/6/4>.
- 2008** Moran, Steven, Darren Tanner and Michael Scanlon (eds). Proceedings from the 20th Northwest Linguistics Conference (NWLC). University of Washington Working Papers in Linguistics Vol. 27. Online: <http://depts.washington.edu/uwwpl/vol27/>.
- 2007** Moran, Steven (ed). University of Washington Working Papers in Linguistics Vol 26. Online: <http://depts.washington.edu/uwwpl/vol26/>.
- 2006** Moran, Steven (ed). University of Washington Working Papers in Linguistics Vol. 25. Online: <http://depts.washington.edu/uwwpl/vol25/>.
- 2005** Jinguji, Dan and Steven Moran (eds). University of Washington Working Papers in Linguistics Vol. 24. Online: <http://depts.washington.edu/uwwpl/vol24/>.
- 2001–2005** Moran, Steven (ed). The Linguist List. ISSN: 1068–4875. 420 issues. Online: <http://linguistlist.org>.

## OUTREACH

- 2019** Moran, Steven and Balthasar Bickel. Softer, processed foods changed the way ancient humans spoke. *The Conversation*. Online: <https://theconversation.com/softer-processed-foods-changed-the-way-ancient-humans-spoke-113599>.

## PUBLISHED DATABASES AND SOFTWARE

- 2019** Moran, Steven and Daniel McCloy (eds). PHOIBLE 2.0. Jena: Max Planck Institute for the Science of Human History. DOI: <http://doi.org/10.5281/zenodo.2562766>. Online: <http://phoible.org/>.
- 2018** Forkel, Robert and Steven Moran. cldf/segments: Unicode Standard tokenization (Version v2.0.0). Zenodo. DOI: <http://doi.org/10.5281/zenodo.1283757>.

- 2017** Moran, Steven and Robert Forkel. cldf/segments: segments 1.1.1 (Version v1.1.1). Zenodo. DOI: <http://doi.org/10.5281/zenodo.1051158>.
- 2016** Moran, Steven, Forkel, Robert and Jeffrey Heath (eds). Dogon and Bangime Linguistics. Jena: Max Planck Institute for the Science of Human History. Online: <http://dogonlanguages.org/>.
- 2015** Naumann, Christfried, Moran, Steven, Segerer, Guillaume and Robert Forkel (eds). Tsamalex: A lexical database on plants and animals. Leipzig: Max Planck Institute for Evolutionary Anthropology. DOI: <http://doi.org/10.5281/zenodo.596403>. Online: <http://tsamalex.clld.org>.
- 2014** Moran, Steven, McCloy, Daniel and Richard Wright (eds). PHOIBLE Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. DOI: <http://doi.org/10.5281/zenodo.2562767>. Online: <http://phoible.org/>.
- 2014** Vaux, Bert, Leemann, Adrian, Wanitsch, Daniel, Moran, Steven, Grimm, Sarah, Robert, Sylvain, Zhakarko, Taras and Scherrer, Yves. US Dialect App. Previously online via Apple App Store.

## ENCYCLOPEDIA ARTICLE

- 2019** Moran, Steven. Phonological inventories. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. DOI: <http://dx.doi.org/10.1093/acrefore/9780199384655.013.112>.

## BOOK REVIEW

- 2014** Wichmann, Søren & Anthony P. Grant (eds). 2012. Quantitative approaches to linguistic diversity: commemorating the centenary of the birth of Morris Swadesh. Amsterdam / Philadelphia: John Benjamins [Benjamins Current Topics 46]. x+182 pp. (ISBN 78 90 272 0265 9) in *Studies in Language* 38(1), 220–233.

## UNPUBLISHED MANUSCRIPTS AND WORK IN PROGRESS

- van Gijn, Rik, Alena Witzlack-Makarevich and Steven Moran. *Typology in Practice*. A course text book for typology using quantitative approaches.
- Moran, Steven. A Grammar of Isaala (Western Sisaala [ssl]). Unpublished manuscript. 117pp. Available on request.
- Moran, Steven. A Grammatical Sketch of Toro So [dts] and accompanying French – English – Toro So Dictionary. Work in progress which requires additional on-site fieldwork.
- Schikowski, Robert, Steven Moran and Sabine Stoll. Manual for the ACQDIV corpus. Unpublished manuscript for ERC project: Acquisition processes in maximally diverse languages: min(d)ing the ambient language. Includes description of the ten longitudinal corpora, the extract transform load (ETL) processes for making the data interoperable, and the database design. 68pp. Online: [https://github.com/acqdiv/corpus\\_manual/blob/master/corpus\\_manual.pdf](https://github.com/acqdiv/corpus_manual/blob/master/corpus_manual.pdf)

## INVITED PRESENTATIONS

- 2019** ‘ACQDIV project, corpus database, methodological approaches using different types of entropy and statistical regressions’. Paper presented at Emergent constructions – big data in learner language analyses, University of Basel. May 30, Basel, Switzerland. With Nicholas Lester.
- 2018** ‘Building an evolutionary timeline of phonetic features’. Paper presented in the Department of Comparative Cognition in the Institute of Biology, University of Neuchâtel. November 29, Neuchâtel, Switzerland.
- 2017** ‘ACQDIV PHON: Worldwide phoneme frequency predicts age of production’. Paper presented at the ESRC International Centre for Language and Communicative Development (LuCiD) research colloquium. October 3, Manchester, UK. With Sabine Stoll.
- 2017** ‘Linked Data for Linguistic Research: Phonetics’. Paper presented at the LD4LR workshop for the Dutch CLARIAH project. February 6–7. Utrecht, The Netherlands. <https://www.clariah.nl/en/new/blogs/linked-data-for-linguistic-research>
- 2015** ‘The ACQDIV project and database’. Paper presented in the Seminar Series at the Dynamique du Langage lab (DDL) of the Centre national de la recherche scientifique (CNRS). September 21. Lyon, France. With Sabine Stoll.
- 2015** ‘Introduction to ACQDIV corpora’. Workshop: Acquisition processes in maximally diverse languages (ACQDIV). June 10–13. Kappel am Albis, Switzerland.
- 2015** ‘Visualizing WALS data’. Workshop: Language Comparison with Linguistic Databases. Organized by the URPP Language and Space at the University of Zurich. April 30. Leipzig, Germany. With Garland McNew.
- 2015** ‘Dogon languages project’. Primera Escuela Latinoamericana de Tipología y Documentación Lingüística. Escuela Nacional de Estudios Superiores, UNAM. March 19–28. Morelia, Mexico.
- 2015** ‘Documenting typological diversity for real: the need of language acquisition studies’. Primera Escuela Latinoamericana de Tipología y Documentación Lingüística. Escuela Nacional de Estudios Superiores, UNAM. March 19–28. Morelia, Mexico. With Dagmar Jung and Sabine Stoll.
- 2014** ‘Studies in phonological typology with PHOIBLE’. Workshop on Language Comparison with Linguistic Databases: RefLex and Typological Databases. October 7–9. Nijmegen, The Netherlands.
- 2014** ‘Linguistic (Linked) Open Data: From collection to application’. **Keynote** presentation at the workshop on ‘Collaboration and Computing for Under-resourced Languages in the Linked Open Data Era (CCURL 2014)’. May 26. Reykjavik, Iceland. With Christian Chiarcos.
- 2014** ‘Qualitative analysis and quantitative comparison of Dogon languages’. Paper presented in the Seminar Series at the Dynamique du Langage lab (DDL) of the Centre national de la recherche scientifique (CNRS). January 24. Lyon, France.
- 2013** ‘Qualitative analysis and quantitative comparison of Dogon languages’. Paper presented in the Seminar Series at the Department of Linguistics, University of Stockholm. December 12. Stockholm, Sweden.



- 2013** ‘Diversity of Language Sources: Challenges in Digitization, Interoperation and Analysis’. Paper presented at the Language Diversity Congress: Computational Issues in Studying Language Diversity. July 18–20. Groningen, The Netherlands. With Jelena Prokić.
- 2013** ‘Obstruents and latitude: acoustic adaptation, thermal influence or spurious statistics?’. Poster presented at the Language Diversity Congress: Computational Issues in Studying Language Diversity. July 18–20. Groningen, The Netherlands. With Damián Blasi.
- 2012** ‘Geolinguistic Diffusion of Language Change’. Paper presented at the Simulation of Diachrony workshop. November 5–6. Zurich, Switzerland. With Jelena Prokić.
- 2012** ‘Research Unit: Quantitative Language Comparison’. Paper presented at the RefLex meeting (CNRS). June 5–7. Lyon, France.
- 2011** ‘Technological infrastructure for comparative linguistics and linguistic fieldwork: some case studies’. Paper presented at the Max Planck Institute for Evolutionary Anthropology. January 27, Leipzig, Germany.
- 2008** ‘The e-Linguistics Toolkit’. Paper presented at the Institute for Language Information and Technology (The Linguist List). December 18, Ypsilanti, Michigan. With Scott Farrar.
- 2006** ‘From Documentation to Computation’. Paper presented at the Max Planck Institute for Evolutionary Anthropology. September 13, Leipzig, Germany.

## REFEREED PRESENTATIONS

- 2019** ‘What contact typologists want from grammatical descriptions’. Paper presented at Descriptive grammars and typology: The challenges of writing grammars of underdescribed and endangered languages. March 27–29, Helsinki, Finland. With Eitan Grossman.
- 2018** ‘Variation sets in maximally diverse languages’. Poster presented at the 43rd Boston University Conference on Language Development (BUCLD). November 2–4, Boston, USA. With Sabine Stoll.
- 2018** ‘Acquisition processes in maximally diverse languages: min(d)ing the ambient language’. Paper presented at the Days of Swiss Linguistics. October 12, Bern, Switzerland. With Sabine Stoll, Damian Blasi, Dagmar Jung and Jekaterina Mazara.
- 2018** ‘Worldwide frequency distribution of phoneme types predicts their order of acquisition’. Paper presented at the 51st Annual Meeting of the Societas Linguistica Europaea (SLE). August 29–September 1, Tallinn, Estonia. With Sebastian Sauppe, Nicholas Lester and Sabine Stoll.
- 2018** ‘Comparative corpus linguistics: New perspectives and applications’. Paper presented at the 51st Annual Meeting of the Societas Linguistica Europaea (SLE). August 29–September 1, Tallinn, Estonia. With Natalia Levshina and Annemarie Verkerk.
- 2018** ‘Cross-linguistically Small World Networks are Ubiquitous in Child-directed Speech’. Paper presented at the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). May 7–12, Miyazaki, Japan. With Danica Pajović and Sabine Stoll.
- 2018** ‘BDPROTO: A Database of Phonological Inventories from Ancient and Reconstructed Languages’. Paper presented at the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). May 7–12, Miyazaki, Japan. With Egidio Marsico, Sebastien Flavien and Annemarie Verkerk.

- 2018** ‘Towards faithfully visualizing global linguistic diversity’. Paper presented at the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). May 7–12, Miyazaki, Japan. With Garland McNew and Curdin Derungs.
- 2018** ‘Differential rates of change in consonant and vowel systems’. Paper presented at the 12th International Conference on the Evolution of Language (EVOLANGXII). April 16–19, Toruń, Poland. With Annemarie Verkerk.
- 2018** ‘Probing the biological conditions for modern phonologies: The late emergence of labiodentals’. Paper presented at the 12th International Conference on the Evolution of Language (EVOLANGXII). April 16–19, Toruń, Poland. With Damián Blasi, Scott Moisik, Ani Margvelashvili, Paul Widmer, Dan Dediu and Balthasar Bickel.
- 2017** ‘Phylogenesis of phonetic features’. Paper presented at the 12th Conference of the Association for Linguistic Typology (ALT). December 11–15, Canberra, Australia.
- 2017** ‘Probing the biological conditions for modern phonologies: the late emergence of labiodentals’. Paper presented at the 48th annual meeting of the Australian Linguistics Society (ALS). December 4–7, Sydney, Australia. With Damián Blasi, Scott Moisik, Ani Margvelashvili, Paul Widmer, Dan Dediu and Balthasar Bickel.
- 2017** ‘Worldwide frequency of phonemes predicts their age of acquisition’. Poster presented at the 42nd Boston University Conference on Language Development (BUCLD). November 3–5, Boston, USA. With Sabine Stoll.
- 2017** ‘The Unicode Cookbook for Linguists’. Paper presented at the 47th Poznań Linguistic Meeting. September 18–20, Poznań, Poland. With Michael Cysouw.
- 2017** ‘Molecular Anthropology as a window on language contact: diffusion probabilities in phonology and grammar’. Paper presented at the 50th Annual Meeting of the Societas Linguistica Europaea (SLE). September 10–13, Zurich, Switzerland. With Balthasar Bickel, Damián Blasi and Brigitte Pakendorf.
- 2017** ‘Probing the biological conditions for modern phonologies: the late emergence of labiodentals’. Paper presented at the 50th Annual Meeting of the Societas Linguistica Europaea (SLE). September 10–13, Zurich, Switzerland. With Damián Blasi, Scott Moisik, Ani Margvelashvili, Paul Widmer, Dan Dediu and Balthasar Bickel.
- 2017** ‘Introduction to the SLE workshop: Non-linguistic causes of linguistic diversity’. Paper presented at the 50th Annual Meeting of the Societas Linguistica Europaea (SLE). September 10–13, Zurich, Switzerland. With Antonio Benítez-Burraco and Dan Dediu.
- 2017** ‘Investigating the phylogenesis of phonetic features’. Poster presented at the 25th Manchester Phonology Meeting. May 25–27, Manchester, UK.
- 2016** ‘Evaluating information loss from phonological dimensionality reduction’. Paper presented at The Role of Predictability in Shaping Human Language Sound Patterns (satellite event at the 16th Speech Science and Technology Conference). December 10–11, Sydney, Australia. With Jayden Macklin-Cordes and Erich Round.
- 2016** ‘Frequent frames in maximally diverse languages’. Paper presented at the Boston University Conference on Language Development (BUCLD). November 4–6, Boston, USA. With Damián Blasi and Sabine Stoll.

- 2016** ‘How to study variation in language acquisition: a new cross-linguistic corpus approach’. Paper presented at the 9th Days of Swiss Linguistics. June 29–July 1, Geneva, Switzerland.
- 2016** ‘Visualizing development in longitudinal language acquisition data’. Poster presented at the 9th Days of Swiss Linguistics. June 29–July 1, Geneva, Switzerland. With Danica Pajović, Jekaterina Mazara and Sabine Stoll.
- 2016** ‘The ACQDIV Database: Min(d)ing the Ambient Language’. Paper presented at the Tenth International Conference on Language Resources and Evaluation (LREC 2016). May 23–28, Portorož, Slovenia. With Robert Schikowski, Danica Pajović, Cazim Hysi and Sabine Stoll.
- 2015** ‘Syntactic mixing across generations in an environment of community-wide bilingualism’. Paper Presented at 48th Annual Meeting of the Societas Linguistica Europaea (SLE). September 2–5. Leiden, The Netherlands. With Balthasar Bickel, Taras Zakharko, Robert Schikowski and Sabine Stoll.
- 2015** ‘Phonological features from a quantitative typological perspective’. Paper Presented at the Conference: Diversity Linguistics: Retrospect and Prospect. May 1–3. Leipzig, Germany. With Balthasar Bickel and Taras Zakharko.
- 2015** ‘Tsammaxlex’. Paper Presented at the Conference: Diversity Linguistics: Retrospect and Prospect. May 1–3. Leipzig, Germany. With Christfried Naumann, Robert Forkel and Lena Sell.
- 2015** ‘Syntactic mixing across generations in a Chintang/Nepali bilingual environment’. Paper Presented at the Conference: Diversity Linguistics: Retrospect and Prospect. May 1–3. Leipzig, Germany. With Sabine Stoll, Taras Zakharko, Robert Schikowski and Balthasar Bickel.
- 2014** ‘Subject agreement suspension and information structure: a view from Dogon languages’. Paper Presented at 8th Days of Swiss Linguistics. June 19–21. Zurich, Switzerland. With Giorgio Iemmolo.
- 2014** ‘Average in phonological typology: phoneme inventories’. Paper Presented at 8th Days of Swiss Linguistics. June 19–21. Zurich, Switzerland. With Taras Zakharko.
- 2014** ‘Structural diversity in phonological systems’. Paper Presented at 8th Days of Swiss Linguistics. June 19–21. Zurich, Switzerland. With Balthasar Bickel and Taras Zakharko.
- 2014** ‘A Crowdsourcing Smartphone Application for Swiss German: Putting Language Documentation in the Hands of the Users’. Poster presented at LREC 2014. May 26–31. Reykjavik, Iceland. With Jean-Philippe Goldman, Adrian Leemann, Marie-José Kolly, Ingrid Hove, Ibrahim Almajai and Volker Dellwo.
- 2013** ‘Lemon-aid: using Lemon to aid quantitative historical linguistic analysis’. Paper presented at the 2nd Workshop on Linked Data in Linguistics: Representing and linking lexicons, terminologies and other language data. September 23. Pisa, Italy. With Martin Brümmer.
- 2013** ‘Linguistic Resources Enhanced with Geospatial Information’. Paper presented at the 2nd Workshop on Linked Data in Linguistics: Representing and linking lexicons, terminologies and other language data. September 23. Pisa, Italy. With Richard Littauer and Boris Villazon-Terrazas.
- 2013** ‘Cross-linguistic Comparison of Complexity Measures in Phonological Systems’. Paper Pre-

- sented at the 44th Poznań Linguistic Meeting (PLM 2013). August 29–September 1. Poznań, Poland. With Damián Blasi.
- 2013** ‘Application and Evaluation of Quantitative Comparative Methods on Dogon’. Paper presented at the Colloquium on African Languages and Linguistics (CALL). August 26–28. Leiden, The Netherlands. With Jelena Prokić.
- 2013** ‘Estimating Diachronic Preferences of Phonological Features Cross-linguistically’. Paper Presented at the Association for Linguistic Typology. August 15–18. Leipzig, Germany. With Taras Zakharko and Balthasar Bickel.
- 2013** ‘Typology with Graphs and Matrices’. Paper Presented at the Association for Linguistic Typology. August 15–18. Leipzig, Germany. With Michael Cysouw.
- 2013** ‘An Open Source Toolkit for Quantitative Historical Linguistics’. Paper Presented at the Association for Computational Linguistics. August 4–9. Sofia, Bulgaria. With Johann-Mattis List.
- 2013** ‘Digital technology and endangered Dogon languages’. Paper Presented at 19th LIPP-Symposium: Gefährdete Sprachen. July 10–12. Munich, Germany. With Jelena Prokić.
- 2013** ‘Revisiting “The role of features in phonological inventories”’. Paper presented at the CUNY Conference on the Feature in Phonology and Phonetics. January 16–18. New York, NY. With Daniel McCloy and Richard Wright. [http://dan.mccloy.info/pubs/McCloyEtAl2013\\_cunyFeatureConf.pdf](http://dan.mccloy.info/pubs/McCloyEtAl2013_cunyFeatureConf.pdf).
- 2012** ‘Automatic Detection of Borrowings in Lexicostatistic Datasets: A Workflow for Automatic Linguistic Reconstruction’. Paper Presented at the Workshop on Quantitative Approaches to Areal Typology. December 13–14. Amsterdam, Netherlands. With Johann-Mattis List and Jelena Prokić.
- 2012** ‘A Python Toolkit for Quantitative Tasks in Historical-Comparative Linguistics’. Paper Presented at the Phylometric and Phylogenetic Approaches in the Humanities Workshop. November 24. Bern, Switzerland. With Johann-Mattis List.
- 2012** ‘Investigating the Genealogical Relatedness of the Dogon Languages’. Paper Presented at the Biological Evolution and Diversification of Languages Workshop [BEDLAN]. September 19–21. University of Turku, Finland. With Jelena Prokić.
- 2012** ‘Dogon classification’. Paper presented at Proto-Niger-Congo: Comparison and Reconstruction International Congress. September 18–21. Paris, France. With Kirill Prokhorov and Jeffrey Heath.
- 2012** ‘A Python Library for Historical Comparative Linguistics’. Paper presented at EuroSciPy 2012. August 23–27. Brussels, Belgium. With Johann-Mattis List. **Awarded best paper.**
- 2012** ‘Investigating the Genealogical Relatedness of the Endangered Dogon Languages’. Paper presented at Digital Humanities 2012. July 16–22. Hamburg, Germany. With Jelena Prokić.
- 2012** ‘Using Linked Data to create a typological knowledge base’. Paper presented at Linked Data in Linguistics (DGfS 2012). March 6–9. Frankfurt, Germany.
- 2011** ‘Black box statistical approaches’. Paper presented at the Workshop on Comparing Approaches to Measuring Linguistic Differences. October 24–25. Gothenburg, Sweden. With Jelena Prokić.

- 2011** ‘Phonemic Inventories as a Phylogenetic Characteristic’. Paper presented at the 20th International Conference on Historical Linguistics (ICHLXX). July 25–30, Osaka, Japan. With Michael Cysouw.
- 2011** ‘Oblique Cases and Semantic Roles’. Paper presented at the Workshop on Semantic Role Complexes: (New) Approaches to Defining Semantic Roles. April 4–5, Zurich, Switzerland. With Jeffrey Heath.
- 2009** ‘Dogon Linguistics and Typology on the Web’. Poster presented at the Association for Linguistic Typology (ALT08). July 24–25, Berkeley, California. With Laura McPherson.
- 2009** ‘The Dogon Languages of Mali’. Paper presented at the Annual Conference on African Linguistics (ACAL). April 9–12, Indiana University, Bloomington, Indiana. With Jeffrey Heath and Abbie Hantgan.
- 2008** ‘The e-Linguistics Toolkit’. Paper presented at e-Humanities — an emerging discipline: Workshop in the 4th IEEE International Conference on e-Science. December 10, Indianapolis, Indiana. With Scott Farrar.
- 2008** ‘From the Field to the World’. Paper presented at the 20th Northwest Linguistics Conference (NWLC). May 3–4, Seattle, Washington. With Naomi Fox.
- 2006** ‘Moodle: A CALL Environment for Language Documentation and Revitalization’. Poster presented at the 2nd annual Conference on Endangered Languages and Cultures of Native America (CELCNA). March 31–April 2, Salt Lake City, Utah.
- 2006** ‘Best Practices in Practice: Documentation of Western Sisaala’. Poster presented at the annual meeting of the Linguistic Society of America, January 5–8, Albuquerque, New Mexico.
- 2006** ‘APEC: A tool for automatic phonological analysis of field data’. Software demonstration at the E-MELD Workshop on Digital Language Documentation, Tools and Standards: the state of the art. June 20–22, Lansing, Michigan. With Shauna Eggers.
- 2006** ‘Creating a Video Clip Resource to Enhance Linguistics Courses’. Poster presented at the University of Washington Teaching and Learning Symposium. April 25, Seattle, Washington. With Sharon Hargus.
- 2005** ‘Creating Language Resources that Last: The E-MELD School of Best Practices in Digital Language Documentation’. Poster presented at the 2005 LSA Summer Institute event: Language Documentation: Theory, Practice, and Values. July 9–11. Harvard University, Cambridge, Massachusetts. With Helen Aristar-Dry, Antony Aristar, Naomi Fox, Susan Hooyenga and Megan Zdrojowski.
- 2005** ‘Ensuring that Digital Data Last: Archiving and Linguistic Resources’. Paper presented at the 34th annual meeting of the Linguistic Association of the Southwest. October 7–9, Lubbock, Texas.
- 2005** ‘Creating Language Resources that Last: The E-MELD Project and Best Practices in Digital Language Documentation’. Poster presented at the E-MELD Workshop 2005: Linguistic Ontologies and Data Categories for Linguistic Resources, July 1–3, Cambridge, Massachusetts. With Susan Hooyenga and Megan Zdrojowski.
- 2005** ‘The E-MELD School of Best Practices in Digital Language Documentation’. Paper presented at the 41st annual meeting of the Chicago Linguistics Society 41. April 7–9, Chicago, Illinois.

- 2005** ‘Endo and Exocentric Compounding in Western Sisaala’. Paper presented at the 31st annual meeting of the Berkeley Linguistics Society. February 18–20, Berkeley, California.
- 2005** ‘Challenges in Fieldwork: Review of Databases for Linguistic Purposes’. Poster presented at the annual meeting of the Linguistic Society of America. January 6–9, Oakland, California.
- 2004** ‘Review of DBMS for Linguistic Purposes’. Paper presented at the E-MELD Conference 2004: Linguistic Databases and Best Practice. July 15–18, Detroit, Michigan. With Marisa Ferrara.
- 2002** ‘English Loanwords in Maori’. Paper presented at the Eastern Michigan University Undergraduate Symposium. March 22, Ypsilanti, Michigan.

## **WORKSHOPS ORGANIZED**

- TBA** ‘How did spoken phonology evolve in humans?’ Co-sponsored by the Swiss Society for Linguistics and the Department of Comparative Linguistics, University of Zurich.
- 2019** ‘Current research in phonological typology’, with Eitan Grossman, to be hosted at the 13th meeting of the Association for Linguistic Typology (ALT). September 4–6, Pavia, Italy.
- 2019** ‘Phonological (in)stability’, with Eitan Grossman, to be hosted at the 52nd Annual Meeting of the Societas Linguistica Europaea (SLE). August 21–24, Leipzig, Germany.
- 2018** ‘Comparative corpus linguistics: new perspectives and applications’, with Natalia Levshina and Annemarie Verkerk, hosted at the 51st Annual Meeting of the Societas Linguistica Europaea (SLE). August 29 – September 1, Tallinn, Estonia.
- 2017** ‘Non-linguistic causes of linguistic diversity’, with Antonio Benítez Burraco and Dan Dediu, hosted at the 50th Annual Meeting of the Societas Linguistica Europaea (SLE). September 10–13, Zurich, Switzerland.

## **NON-REFEREED PRESENTATIONS**

- 2016** ‘Tracing the origin of human speech sounds: evolutionary and historical perspectives’. Paper presented at the Genetics and Animal Communication Workshop hosted by the University Research Priority Program: Evolution in Action. November 21–22, Zurich, Switzerland. With Balthasar Bickel, Damián Blasi and Paul Widmer.
- 2016** ‘Statistical Signals of the Trans-Pacific Migrations in Language’. Paper presented at the Linguistics and Genetics Workshop in Zurich 2016. University of Zurich. March 13. Authors: Balthasar Bickel, Curdin Derungs, Steven Moran, Sean Roberts and Dan Dediu.
- 2015** ‘Making the data usable: issues in unifying language acquisition corpora’. Paper presented at the Inter-Lab Workshop on Data and Standards: How to make data reusable? University of Zurich. May 29. With Robert Schikowski.
- 2015** ‘ACQDIV: data, infrastructure, future’. Paper Presented at the Research Colloquium of the Department of Comparative Linguistics at the University of Zurich. May 22.
- 2014** ‘Language Geography and Dogon’. Paper Presented at the Maps and Grammar Workshop at the University of Zurich. March 21. With Curdin Derungs and Simon Etter.

- 2013** ‘Estimating Diachronic Preferences of Phonological Features Cross-linguistically’. Paper Presented at the Research Colloquium of the General Department of Linguistics at the University of Zurich. October 25. With Taras Zakharko and Balthasar Bickel.
- 2012** ‘Research Unit: Quantitative Language Comparison’. Paper presented at the Institute for Language Information and Technology (The Linguist List). September 12, Ypsilanti, Michigan.
- 2011** ‘Orthography Profiles’. Paper presented at the Phonetics Lab Meeting, LMU, München. May 18, Munich, Germany. With Jelena Prokić.
- 2006** ‘Benefits for Using the LKB and Grammar Engineering for Field Linguists’. Paper presented at the 2nd annual Deep Linguistic Processing with HPSG (DELPH-IN) meeting. June 12–17, Fefor, Norway.
- 2006** ‘The E-MELD Project and Tools for Field Linguists’. Demo presented at the Microsoft / UW Symposium in Computational Linguistics. February 3, Redmond, Washington.

## **TEACHING EXPERIENCE**

### **University courses (instructor)**

Introduction to First Language Acquisition from a Cross-linguistic Perspective (301: Einführung in die Spracherwerbsforschung). University of Zurich, Spring 2019 & Spring 2016.

The Nature and Evolution of Phonology (605: Sprachvergleich). University of Zurich, Fall 2018.

Linguistic Field Methods (501: Feldforschungsmethoden). University of Zurich, Fall 2017 (Ga language) & Fall 2016 (Akan language).

Anthropological Linguistics: Social and Cultural Perspectives (302a: Ethnolinguistik). University of Zurich, Fall 2015.

Qualitative and Quantitative Analysis (102: Methoden und Anwendungen). University of Zurich, Spring 2015 & Fall 2013.

Introduction to Anthropological Linguistics (302: Ethnolinguistik). University of Zurich, Fall 2014. Taught with Balthasar Bickel.

Introduction to Morphology (LING 481). University of Washington, annually 2007–2010.

### **University courses (teaching assistant)**

Deep Processing Techniques for Natural Language Processing (LING 571). University of Washington. Winter 2010 & 2009.

Introduction to Computational Linguistics (CSE/LING 472). University of Washington. Spring 2010.

### **Summer schools (instructor)**

Methods and practical applications for language documentation (Métodos e instrumentos prácticos para la documentación lingüística, with Heriberto Avelino). Primera Escuela Latinoamericana de Tipología y Documentación Lingüística Alemania-México. Nacional de Estudios Superiores, UNAM. Morelia, Mexico. March 23–28, 2015.

Database Design for Linguists. Quantitative Analysis of Linguistic Data (QUALD) doctoral seminar, University of Zurich, Summer 2014.

## RESEARCH PROJECT TEAM MEMBER

**2014–2019** ACQDIV: Acquisition processes in maximally diverse languages: Min(d)ing the ambient language. European Research Council Consolidator Grant ERC-2013-CoG, SH4, RCN: 615988. Primary Investigator (PI) Sabine Stoll (U. Zurich). Positions: postdoctoral researcher and technical project manager. <https://cordis.europa.eu/project/rcn/191230/>. Project homepage: <http://www.acqdiv.uzh.ch>.

**2013–2016** Dogon Languages of Mali. National Science Foundation (NSF) award no. BCS-1263150. PI Jeffrey Heath (U. Michigan). Positions: field researcher and technical lead. [https://www.nsf.gov/awardsearch/showAward?AWD\\_ID=1263150](https://www.nsf.gov/awardsearch/showAward?AWD_ID=1263150). Project webpages: <http://dogonlanguages.org> & <http://tsammalex.clld.org/>.

**2010–2014** Quantitative modeling of historical-comparative linguistics: Unraveling the phylogeny of native South American languages (QuantHistLing). European Research Council Starting Grant ERC-2009-StG, SH4, RCN: 191230. PI Michael Cysouw (U. Marburg). Position: postdoctoral researcher. <https://cordis.europa.eu/project/rcn/93461/>.

**2009–2013** Dogon Languages of Mali. National Science Foundation (NSF) award no. BCS-0853364. PI Jeffrey Heath. Positions: field researcher and technical lead. [https://www.nsf.gov/awardsearch/showAward?AWD\\_ID=0853364](https://www.nsf.gov/awardsearch/showAward?AWD_ID=0853364).

**2009–2010** Phonetics Information Base and Lexicon (PHOIBLE). UW Royalty Research Fund Grant. PI Richard Wright (U. Washington). Positions: lead developer and project manager. <http://phoible.org/>

**2008–2010** Lexicon Integration with the Gold Ontology (LEGO). National Science Foundation (NSF) award no. BCS-0753321. PIs Helen Aristar-Dry (EMU) and Jeff Good (U. Buffalo). Positions: programmer and data provider. [https://www.nsf.gov/awardsearch/showAward?AWD\\_ID=0753321](https://www.nsf.gov/awardsearch/showAward?AWD_ID=0753321). Project homepage: <http://lego.linguistlist.org/>.

**2006–2009** Languages of the Dogon. National Science Foundation (NSF) award No. BCS-0537435. PI Jeffrey Heath. Position: field researcher and technical lead. [https://www.nsf.gov/awardsearch/showAward?AWD\\_ID=0537435](https://www.nsf.gov/awardsearch/showAward?AWD_ID=0537435).

**2004–2005** Creating a Grammar, Dictionary, and Texts of the Dogon Languages of Mali, West Africa. National Endowment for the Humanities (NEH) award no. PA-50643. PI Jeffrey Heath. Position: programmer.

**2003–2005** Description of Chorote, Nivaclé and Kadiwéu: three of least known and most endangered languages of the Chaco. Endangered Languages Documentation Programme (Rausing Charitable Fund), SOAS, University of London. PIs Lyle Campbell (U. Utah), Verónica Grondona (EMU) and Filomena Sandalo (IEL-Unicamp). Position: programmer.

**2001–2005** Electronic Metastructure for Endangered Languages Data (E-MELD). National Science Foundation (NSF) award No. BCS-0094934. PIs Helen Aristar-Dry and Anthony Aristar (WSU). Positions: project manager and lead web developer. [https://www.nsf.gov/awardsearch/showAward?AWD\\_ID=0094934](https://www.nsf.gov/awardsearch/showAward?AWD_ID=0094934). Project homepage: <http://emeld.org/>



## VISITING SCIENTIST

- 2010** Max Planck Institute for Evolutionary Anthropology  
January, Leipzig, Germany.
- 2006** Max Planck Institute for Evolutionary Anthropology  
August–September, Leipzig, Germany.

## TECHNICAL SKILLS (PAST & PRESENT)

- Basic statistics, Command line, Coldfusion, CSS, Git, HTML, Javascript, Jupyter, L<sup>A</sup>T<sub>E</sub>X, Linked Data, Linux, Markdown, MySQL, OS X, OWL, Photoshop, PHP, Python, R, RDF, SPARQL, SQL, Unix, XML
- Extensive experience in: the Unicode Standard and multilingual character encodings, text processing and data mining, orthographic segmentation, syntactic & semantic data interoperation, database design and implementation, web applications development, digital archiving & digitizing language documentation and multimedia, qualitative and quantitative analysis
- For various coding projects, see: <https://github.com/bambooforest/>
- Data creator and curator of PHOIBLE Online (<http://phoible.org/>), a broad coverage cross-linguistic database of phonological inventory data on sounds of the world's languages
- Data contributor and curator the Dogon and Bangime Linguistics project website (<http://dogonlanguages.org/>), a comprehensive multilingual and multimedia resource on the Dogon language family in Mali

## FIELDWORK

- **Toro So** (Dogon; ISO 639-3: dts) in Bobo-Dioulasso, Burkina Faso (Mali unsafe during this period, 2013); continuation of grammar and dictionary development
- **Toro So** in Douentza and Sangha, Mali (2009); initial English–French–Toro So dictionary and grammar development
- **Western Sisaala** (Niger-Congo; Gur; ISO 639-3: ssl) in Lambussie, UWR, Ghana (2003); orthography, dictionary and grammar development; extensive grammatical knowledge

## LANGUAGES

- **English** native
- **German** near native
- **French** basic
- **Toro So** extensive grammatical knowledge through on-site fieldwork
- **Sisaala** extensive grammatical knowledge through on-site fieldwork
- **Akan** taught field methods course

- **Ga** taught field methods course
- **Tigrinya** field methods course participant
- **Sisaala** field methods course participant

## INVITED PARTICIPANT

- 2017** ‘Quantitative Methods - Spring School 2017’, sponsored by the Max Planck Institute for the Science of Human History (MPI SSH). May 8–14. Jena, Germany.
- 2017** ‘LD4LR: Linked Data for Linguistic Research’, sponsored by the Common Lab Research Infrastructure for the Arts and Humanities (CLARIAH) initiative. February 6–7. Utrecht, The Netherlands. Online: <https://www.clariah.nl/en/new/blogs/linked-data-for-linguistic-research>.
- 2015** ‘Workshop on Language Comparison with Linguistic Databases’, sponsored by the Max Planck Institute for Evolutionary Anthropology. April 30. Leipzig, Germany. Online: <https://www.eva.mpg.de/linguistics/conferences/2014-ws-lanclid2/index.html>.
- 2014** ‘Workshop on Language Comparison with Linguistic Databases: RefLex and Typological Databases’, sponsored by the Max Planck Institute for Psycholinguistics. October 7–9. Nijmegen, The Netherlands. Online: <http://www.mpi.nl/events/language-comparison-with-linguistic-databases-reflex-and-typological-databases>.
- 2012** ‘Simulation of Diachrony’, sponsored by the Department of General Linguistics, University of Zurich. November 5–6. Zurich, Switzerland.
- 2012** ‘Multilingual Web Workshop: Linked Open Data and Multilingual Web-LT Requirements’, sponsored by the World Wide Web Consortium (W3C), Trinity College Dublin, and the European Commission. June 11–13. Dublin, Ireland. Online: <https://www.multilingualweb.eu/documents/dublin-workshop/dublin-cfp>.
- 2012** ‘Typological databases in RDF’, sponsored by the Max Planck Institute for Evolutionary Anthropology. March 14–15, Leipzig, Germany.
- 2009** ‘Cyberling 2009’, sponsored by the Linguistic Society of America, the US National Science Foundation, and the University of California, Berkeley. July 17–19. Berkeley, California. Online: [https://www.linguisticsociety.org/sites/default/files/Cyberling2009Workshop\\_FinalReport.pdf](https://www.linguisticsociety.org/sites/default/files/Cyberling2009Workshop_FinalReport.pdf).
- 2009** ‘Dogon Dictionary Development’, sponsored by the Centre National de la Recherche Scientifique et Technologique (CNRST), Mali. June 8–10, Bamako, Mali, West Africa. Workshop coordinator with Jeffrey Heath.
- 2009** ‘Lexicon Integration with the Gold Ontology (LEGO)’, sponsored by the Institute for Language Information and Technology and Eastern Michigan University. March 6–8, Ypsilanti, Michigan.
- 2007** ‘Toward the Interoperability of Language Resources – Working group on Lexicon Schemas and Related Data Models’, sponsored by the US National Science Foundation, Stanford University, the University of Kansas, Eastern Michigan University, and The Linguist List. July 13–15, Stanford University, Palo Alto, California.

**2006** ‘Digital Tools Summit in Linguistics (DTSL)’, sponsored by the US National Science Foundation. Working group on Fieldwork and Archiving. June 22–23, Lansing, Michigan.

## **AWARDS**

**2013** Nominated for the E.W. Beth Dissertation Prize to outstanding dissertations in the fields of Logic, Language, and Information (FoLLI—the Association for Logic, Language, and Information) by Professors Emily Bender and Richard Wright (UW).

**2013** Nominated for the University of Washington’s Graduate School Distinguished Dissertation Award 2013 by the UW Linguistics Department (one nomination annually per department).

**2012** Awarded best paper for ‘A Python Library for Historical Comparative Linguistics’ (with Johann-Mattis List), presented at EuroSciPy 2012. August 23–27. Brussels, Belgium.

**2010** Excellence in Linguistic Research Graduate Award, Linguistics Department, University of Washington. **One award annually.**

**2006** Fulbright Scholar alternate.

**2001** Distinguished Undergraduate Student in Linguistics, Eastern Michigan University. **One award annually.**

**1999–2001** Dean’s List, Eastern Michigan University.

## **GRANTS**

**2016** ‘TYPDB: a unified interface to typological databases’ with Balthasar Bickel, Paul Widmer and Taras Zakharko. Grant awarded from the UFSP Sprache und Raum, University of Zurich. Awarded 10,000 CHF.

**2013** ‘Combining geography and language data to investigate the vertical diffusion of linguistic features due to language contact’ with Curdin Deungs and Balthasar Bickel. Grant awarded from the UFSP Sprache und Raum, University of Zurich. Awarded 10,000 CHF.

**2008** PHOIBLE: Phonetics Information Base and Lexicon, with Richard Wright. Royalty Research Fund, University of Washington. Awarded \$37,375.

**2007** Student Technology Fund Proposal 2007–059–1, University of Washington Language Learning Center, Open Lab Upgrade. Awarded \$60,795.65.

**2006** Student Technology Fund Proposal 2006–069–2, University of Washington Language Learning Center, PC Lab Replacement. Awarded \$84,829.88.

**2013** Travel grant for fieldwork in Burkina Faso, Swiss Society of Endangered Languages.

**2013–2018** Travel grants for conference travel, University of Zurich.

**2006–2010** Travel grants for conference travel, University of Washington.

**2001–2005** Travel grants for conference travel, Eastern Michigan University.

## **ACADEMIC SERVICE**

**2019–pres** Language Science Press Advisory Board.

**2013–pres** Linguist List Advisory Panel.

**2007–2012** Linguistic Society of America Technology Advisory Committee (LSA-TAC)

**2019** *Scientific committee*: Language, Data, and Knowledge (LDK 2019). *Reviewer*: Lingua; Open Handbook of Linguistic Data Management. *Grant reviewer*: Endangered Language Documentation Programme (ELDP). *Program committee*: Association Linguistic Typology (ALT 2019).

**2018** *Reviewer*: PLoS ONE; Journal of Language Evolution; Conference on Computational Linguistics (COLING 2018); Linguistic Typology; International Conference for Graduate students on Diverse Approaches to Linguistics (IGDAL 7). *Proof reader*: Language Science Press. *Scientific committee*: 3rd Computational Methods for Endangered Languages (ComputEL-3). *Grant reviewer*: Biotechnology and Biological Sciences Research Council (BBSRC); National Science Center, Poland.

**2017** *Reviewer*: Journal of Language Evolution; Laboratory Phonology; Linguistic Typology; Evolution of Language International Conferences (EVLANG 2018). *Program Committee*: 1st International Conference on Language, Data and Knowledge (LDK 2017); 6th Workshop on Linked Data in Linguistics (LDL-2018): Towards Linguistic Data Science. *Scientific Committee*: 11th Language Resources and Evaluation Conference (LREC-2018).

**2016** *Reviewer*: Laboratory Phonology; Journal of the International Phonetic Association; Journal of Language Evolution. *Scientific Committee*: 10th International Conference on Language Resources and Evaluation (LREC 2016); 2nd Workshop on Computational Methods for Endangered Languages (ComputEL-2). *Program Committee*: CCURL 2016 – Towards an Alliance for Digital Language Diversity; Linked Data in Linguistics (LDL-2016); Terminology and Knowledge Engineering (TKE 2016: Term Bases and Linguistic Linked Open Data). *Grant reviewer*: Endangered Language Documentation Programme (ELDP).

**2015** *Reviewer*: Phonology; Journal of the International Phonetic Association; LRE Journal Special Issue: ‘Under-resourced Languages, Collaborative Approaches and Linked Open Data: Resources, Methods and Applications’. *Program committee*: 4th Workshop on Linked Data in Linguistics: Resources and Applications (LDL-2015). *Scientific committee*: Association for Linguistic Typology (ALT 11).

**2014** *Reviewer*: Histoire Épistémologie Langage, 8th Days of Swiss Linguistics, Language Resources and Evaluation Conference (LREC). *Program committee*: Linked Data in Linguistics (LDL-2014), CCURL 2014: Collaboration and Computing for Under-resourced Languages in the Linked Open Data Era. *Organizing committee*: Multilingual Linked Open Data for Enterprises (MLODE 2014) workshop. *Scientific committee*: LRE Journal Special Issue: ‘Under-resourced Languages, Collaborative Approaches and Linked Open Data: Resources, Methods and Applications’, 9th International Conference on Language Resources and Evaluation (LREC 2014); ComputEL: The use of computational methods in the study of endangered languages (ACL 2014).

**2013** *Reviewer*: Studies in Language, Language Dynamics and Change, Semantic Web Journal, Linked Data in Linguistics: Representing and linking lexicons, terminologies and other language data.

**2012** *Workshop organizer*: Multilingual Linked Open Data for Enterprises (MLODE) workshop, co-located with Leipziger Semantic Web day. September 24–25, Leipzig, Germany, with Sebastian Hellmann.

- 2012** *Reviewer*: Deutsche Gesellschaft für Sprachwissenschaft (DGfS) Workshop: Linked Data in Linguistics, European Association of Computational Linguistics (EACL) Workshop: LINGVIS & UNCLH, Proceedings of Annual Conference on African Linguistics.
- 2011** *Reviewer*: Selected Proceedings of the 42nd Annual Conference on African Linguistics: African Languages in Context (edited by Michael R. Marlo, Nikki B. Adams, Christopher R. Green, Michelle Morrison and Tristan M. Purvis).
- 2006** Technical liaison for the 25th West Coast Conference on Formal Linguistics. April 28–30. University of Washington, Seattle, Washington.
- 2006** Media coordinator for Word Languages Day 2006. March 3. University of Washington, Seattle, Washington.
- 2005–2009** Editor of the University of Washington Working Papers in Linguistics
- 2005** Liaison and media coordinator at the Breath of Life Conference at the University of Washington. September 12–16, Seattle, Washington.
- 2005** Working group liaison at the E-MELD Workshop 2005: Conference on Linguistic Ontologies and Data Categories for Language Resources. July 1–3, Cambridge, Massachusetts.
- 2004** Working group liaison for the E-MELD Workshop 2004: Conference on Linguistic Databases and Best Practice. July 15–18, Detroit, Michigan.
- 2003** Voluntary instructor of English Grammar and Conversation. Lambussie Secondary School. Lambussie, Ghana.
- 2002** Working group liaison for the E-MELD Workshop 2002: Conference on Digitizing Lexical Information. August 2–5, Ypsilanti, Michigan.

## STUDENT ADVISING

- 2019** Cazim Hysi. Implementing Diachronic Language Simulation in Python. B.A. thesis. Department of Computational Linguistics, University of Zurich.
- 2016** Danica Pajovič. ACQDIVIZ: Visualising Development in Longitudinal First Language Acquisition Data. M.A. thesis. Department of Computational Linguistics, University of Zurich.
- 2015** Melanie Widmer. Die Rolle distributioneller Information im Spracherwerb: Kategorisierung mit Hilfe von häufigen Frames. B.A. thesis. Department of Comparative Linguistics, University of Zurich.

## PROFESSIONAL MEMBERSHIPS (PAST & PRESENT)

- American Association for the Advancement of Science (AAAS)
- Association of Computational Linguistics (ACL)
- Association of Linguistic Typology (ALT)
- European Language Resources Association (ELRA)
- Linguistic Society of America (LSA)
- Linguistics Society at the University of Washington (LSUW)

- Open Working Group in Linguistics (OWLG)
- Schweizerische Gesellschaft für bedrohte Sprachen (SGBS)
- Societas Linguistica Europaea (SLE)
- Züricher Kompetenzzentrum Linguistik (ZüKL)