

## Curriculum Vitae

### Personal information

Name Katharina Lisa Maria Kann  
Phone/e-mail +16466830278; katharina.kann@colorado.edu  
Personal website <https://kelina.github.io>

### Education and research experience

from 01/2020 **Assistant Professor of Computer Science:** University of Colorado Boulder, USA  
04/2018 - 12/2019 **Postdoctoral Research Associate:** New York University, USA  
01/2015 - 09/2017 **PhD in Computer Science:** University of Munich, Germany (GPA: 0.9)  
10/2012 - 11/2014 **MSc in Computer Science:** Technical University of Munich, Germany (GPA: 1.5)  
10/2008 - 11/2011 **BSc in Mathematics** with minor Computer Science: Johannes Gutenberg University Mainz, Germany (GPA: 1.8)  
08/2000 - 03/2008 High school: Maria Ward-Gymnasium, Mainz, Germany (**University-entrance diploma** - Abitur; GPA: 1.0; best grade in my year)

### Industry and project experience

10/2017 - 03/2018 **Internship at Google Zurich:** Research and software engineering  
09/2016 - 12/2016 **Internship at Google Zurich:** Research and software engineering  
06/2015 - 09/2015 **Internship at Google London:** Research and software engineering  
04/2014 - 11/2014 **Master's thesis** about the topic "Improving Human-Robot Interaction by Applying Incremental Processing of Speech and Non-verbal Signals"  
04/2014 - 09/2014 **Student researcher** at fortiss (data collection and gesture recognition with Hidden Markov Models using C#)  
12/2013 - 04/2014 **Student researcher** at DLR (design and development of an application to train amputees for the use of hand prosthesis using C#)  
10/2013 - 04/2014 **Guided research** about the topic "Gesture Recognition in Robotics"

### Stays abroad

09/2011 - 07/2012 **Mandarin language studies** at the Donghua University in Shanghai, China  
09/2009 - 07/2010 **ERASMUS** year in Valencia, Spain (university studies at the Universitat de Valencia; language studies at the language school Centre d'idiomes)  
04/2008 - 07/2008 **Mandarin language studies** at the EF language school in Beijing, China

### Funding and Awards

#### Funding

08/2020 - 07/2025 Senior Personnel, NSF AI Institute for Student-AI Teaming (\$19,993,294)

## Internal Funding

07/2020 - 06/2022 Co-PI, Engineering Education and AI-Augmented Learning IRT (\$250k)

03/2020 Co-PI, RSO funding for equipment maintenance or purchases (\$15k)

## Awards

06/2017 Winning system of task 2 of the **CoNLL-SIGMORPHON 2017 shared task on universal morphological reinflection**

05/2017 **Google scholarship** for attending the Lisbon Machine Learning School (LxMLS 2017)

05/2016 Winning system of the **SIGMORPHON 2016 shared task on morphological reinflection**

07/2011 **Scholarship** from the German Academic Exchange Service and the China Scholarship Council for Mandarin language studies in China

09/2009 **ERASMUS scholarship** for university studies in Valencia

03/2009 **Scholarship** from the "Studienstiftung des Deutschen Volkes"

## Teaching

Classes Natural Language Processing (CSCI 5832/LING 5832). CU Boulder, Spring 2020.

Natural Language Understanding and Computational Semantics (DS-GA 1012/LING-GA 1012, with Samuel R. Bowman). NYU, Spring 2019.

- *Resulting in a publication at the ACL Student Research Workshop 2019*

Mentoring and supervision Supervision of undergraduate intern Huiming Jin. LMU Munich, 2017.

- *Resulting in a publication at SCLeM 2017*

## Skills

Languages **German** (mother tongue), **English** (fluent), **Spanish** (fluent), **Chinese** (intermediate, HSK 5 in 2012), **French** (basic), **Latin** (Latinum)

Programming languages **Java** (good knowledge), **Python** (good knowledge), **C++** (basic knowledge), **C#** (basic knowledge)

## Professional service

2020 Co-organizer of the **SIGMORPHON 2020 shared task on unsupervised discovery of morphological paradigms**

2019 Co-organizer of the workshop **Repl4NLP**

2018 Co-organizer of the **CoNLL--SIGMORPHON 2018 shared task on universal morphological reinflection**

## Publications

Refereed conference papers Jason Phang\*, Iacer Calixto\*, Phu Mon Htut, Yada Pruksachatkun, Haokun Liu, Clara Vania, Katharina Kann and Samuel R. Bowman. **English**

**Intermediate-Task Training Improves Zero-Shot Cross-Lingual Transfer Too.** In ACL 2020 (*to appear*).

Rajat Agarwal and Katharina Kann. **Acrostic Poem Generation.** In EMNLP 2020 (*to appear*).

Manuel Mager, Özlem Çetinoğlu and Katharina Kann. **Tackling the Low-resource Challenge for Canonical Segmentation.** In EMNLP 2020 (*to appear*).

Sarah Moeller, Ling Liu, Changbing Yang, Katharina Kann and Mans Hulden. **IGT2P: From Interlinear Glossed Texts to Paradigms.** In EMNLP 2020 (*to appear*).

Yada Pruksachatkun\*, Jason Phang\*, Haokun Liu\*, Phu Mon Htut\*, Xiaoyi Zhang, Richard Yuanzhe Pang, Clara Vania, Katharina Kann and Samuel R. Bowman. **Intermediate-Task Transfer Learning with Pretrained Language Models: When and Why Does It Work?** In ACL 2020.

Huiming Jin, Liwei Cai, Yihui Peng, Chen Xia, Arya McCarthy and Katharina Kann. **Unsupervised Morphological Paradigm Completion.** In ACL 2020.

Katharina Kann, Samuel R. Bowman and Kyunghyun Cho. **Learning to Learn Morphological Inflection for Resource-Poor Languages.** In AAAI 2020.

Katharina Kann\*, Ophélie Lacroix\* and Anders Søgaard. **Weakly Supervised POS Taggers Perform Poorly on Truly Low-Resource Languages.** In AAAI 2020.

Katharina Kann. **Acquisition of Inflectional Morphology in Artificial Neural Networks With Prior Knowledge.** In SCiL 2020.

Katharina Kann, Kyunghyun Cho and Samuel R. Bowman. **Towards Realistic Practices In Low-Resource Natural Language Processing: The Development Set.** In EMNLP 2019 (short papers).

Yadollah Yaghoobzadeh, Katharina Kann, T. J. Hazen, Eneko Agirre and Hinrich Schütze. **Probing for Semantic Classes: Diagnosing the Meaning Content of Word Embeddings.** In ACL 2019.

Manuel Mager, Özlem Çetinoğlu and Katharina Kann. **Subword-Level Language Identification for Intra-Word Code-Switching.** In NAACL 2019.

Katharina Kann\*, Alex Warstadt\*, Adina Williams\* and Samuel R. Bowman. **Verb Argument Structure Alternations in Word and Sentence Embeddings.** In SCiL 2019.

Katharina Kann and Hinrich Schütze. **Neural Transductive Learning and Beyond: Morphological Generation in the Minimal-Resource Setting.** In EMNLP 2018.

Katharina Kann, Sascha Rothe and Katja Filippova. **Sentence-Level Fluency Evaluation: References Help, But Can Be Spared!** In CoNLL 2018.

Katharina Kann\*, Jesus Manuel Mager Hois\*, Ivan Vladimir Meza Ruiz and Hinrich Schütze. **Fortification of Neural Morphological Segmentation Models for Polysynthetic Minimal-Resource Languages.** In NAACL 2018.

Katharina Kann, Ryan Cotterell and Hinrich Schütze. **One-Shot Neural Cross-Lingual Transfer for Paradigm Completion**. In ACL 2017.  
Katharina Kann, Ryan Cotterell and Hinrich Schütze. **Neural Multi-Source Morphological Reinflection**. In EACL 2017.

Katharina Kann, Ryan Cotterell, Hinrich Schütze. **Neural Morphological Analysis: Encoding Decoding Canonical Segments**. In EMNLP 2016 (short papers).

Katharina Kann and Hinrich Schütze. **Single-Model Encoder-Decoder with Explicit Morphological Representation for Reinflection**. In ACL 2016 (short papers).

Workshop papers

Nikhil Prabhu and Katharina Kann. **Frustratingly Easy Multilingual Grapheme-to-Phoneme Conversion**. In SIGMORPHON 2020.

Assaf Singer and Katharina Kann. **The NYU-CUBoulder Systems for SIGMORPHON 2020 Task 0 and Task 2**. In SIGMORPHON 2020.

Manuel Mager and Katharina Kann. **The IMS-CUBoulder System for the SIGMORPHON 2020 Shared Task on Unsupervised Morphological Paradigm Completion**. In SIGMORPHON 2020.

Anhad Mohananey\*, Katharina Kann\* and Samuel R. Bowman. **Self-Training for Unsupervised Parsing with PRPN**. In IWPT 2020.

Johannes Bjerva, Katharina Kann and Isabelle Augenstein. **Transductive Auxiliary Task Self-Training for Neural Multi-Task Models**. In DeepLo 2019.

Katharina Kann, Anhad Mohananey, Kyunghyun Cho and Samuel R. Bowman. **Neural Unsupervised Parsing Beyond English**. In DeepLo 2019.

Katharina Kann, Stanislas Lauly and Kyunghyun Cho. **The NYU System for the CoNLL-SIGMORPHON 2018 Shared Task on Universal Morphological Reinflection**. In CoNLL-SIGMORPHON 2018.

Manuel Mager, Elisabeth Mager, Alfonso Medina-Urrea, Ivan Meza and Katharina Kann. **Lost in Translation: Analysis of Information Loss During Machine Translation Between Polysynthetic and Fusional Languages**. In All Together Now? Computational Modeling of Polysynthetic Languages 2018.  
Katharina Kann, Johannes Bjerva, Isabelle Augenstein, Barbara Plank and Anders Søgaard. **Character-level Supervision for Low-resource POS Tagging**. In DeepLo 2018.

Yadollah Yaghoobzadeh, Katharina Kann and Hinrich Schütze. **Evaluating Word Embeddings in Multi-label Classification Using Fine-grained Name Typing**. In Repl4NLP 2018.

Katharina Kann and Hinrich Schütze. **Unlabeled Data for Morphological Generation With Character-Based Sequence-to-Sequence Models**. In SCLeM 2017.

Huiming Jin and Katharina Kann. **Exploring Cross-Lingual Transfer of Morphological Knowledge In Sequence-to-Sequence Models**. In SCLeM 2017.

Katharina Kann and Hinrich Schütze. **The LMU System for the CoNLL-SIGMORPHON 2017 Shared Task on Universal Morphological Reinflection.** In CoNLL–SIGMORPHON 2017.

Toms Bergmanis, Katharina Kann, Hinrich Schütze and Sharon Goldwater. **Training Data Augmentation for Low-Resource Morphological Inflection.** In CoNLL–SIGMORPHON 2017.

Katharina Kann and Hinrich Schütze. **MED: The LMU system for the SIGMORPHON 2016 shared task on morphological reinflection.** In SIGMORPHON 2016.

Invited  
publications

Katharina Kann\*, Arya D. McCarthy\*, Garrett Nicolai and Mans Hulden. **The SIGMORPHON 2020 Shared Task on Unsupervised Morphological Paradigm Completion.** In SIGMORPHON 2020.

Ryan Cotterell, Christo Kirov, John Sylak-Glassman, Géraldine Walther, Ekaterina Vylomova, Arya D. McCarthy, Katharina Kann, Sebastian Mielke, Garrett Nicolai, Miikka Silfverberg, David Yarowsky, Jason Eisner and Mans Hulden. Søgaard. **The CoNLL--SIGMORPHON 2018 Shared Task: Universal Morphological Reinflection.** In CoNLL–SIGMORPHON 2018.

Unpublished  
manuscripts

Katharina Kann. **Grammatical Gender, Neo-Whorfianism, and Word Embeddings: A Data-Driven Approach to Linguistic Relativity.**  
arXiv:1910.09729.

Wenpeng Yin, Katharina Kann, Mo Yu and Hinrich Schütze. **Comparative Study of CNN and RNN for Natural Language Processing.**  
arXiv:1702.01923.

## Talks and poster presentations

### Conference talks

- |         |  |
|---------|--|
| 02/2020 | <b>Learning to Learn Morphological Inflection for Resource-Poor Languages</b><br>Location: New York<br>Event: AAAI                             |
| 01/2020 | <b>Acquisition of Inflectional Morphology in Artificial Neural Networks With Prior Knowledge</b><br>Location: New Orleans<br>Event: SCiL       |
| 11/2019 | <b>Towards Realistic Practices In Low-Resource Natural Language Processing: The Development Set</b><br>Location: Hong Kong<br>Event: EMNLP     |
| 11/2018 | <b>Neural Transductive Learning and Beyond: Morphological Generation in the Minimal-Resource Setting</b><br>Location: Brussels<br>Event: EMNLP |
| 11/2018 | <b>Sentence-Level Fluency Evaluation: References Help, But Can Be Spared!</b><br>Location: Brussels<br>Event: CoNLL                            |

- 06/2018 **Fortification of Neural Morphological Segmentation Models for Polysynthetic Minimal-Resource Languages**  
Location: New Orleans  
Event: NAACL
- 08/2016 **Single-Model Encoder-Decoder with Explicit Morphological Representation for Reinflection**  
Location: Berlin  
Event: ACL
- Posters**
- 08/2020 Diksha Meghwal, Iacer Calixto, Stanislaw Jastrzebski and Katharina Kann. **Can Wikipedia Categories Improve Masked Language Model Pretraining?**  
Event: WiNLP Workshop (at ACL)
- 06/2018 Katharina Kann\*, Manuel Mager\*, Ivan Vladimir Meza Ruiz and Hinrich Schütze. **Neural Morphological Segmentation for Polysynthetic Minimal-Resource Languages.**  
Event: WiNLP Workshop (at NAACL)
- 12/2017 Katharina Kann\*, Jesus Manuel Mager Hois\*, Ivan Vladimir Meza Ruiz and Hinrich Schütze. **Improving Neural Morphological Segmentation for Polysynthetic Minimal-Resource Languages.**  
Event: LLD Workshop (at NeurIPS)
- 08/2017 Katharina Kann, Ryan Cotterell and Hinrich Schütze. **Multi-task Training for Morphological Inflection.**  
Event: WiNLP Workshop (at ACL)
- Workshop talks**
- 07/2020 **The SIGMORPHON 2020 Shared Task on Unsupervised Morphological Paradigm Completion**  
Location: online  
Event: SIGMORPHON
- 07/2018 **Character-level Supervision for Low-resource POS Tagging**  
Location: Melbourne  
Event: DeepLo
- Invited talks**
- 10/2019 **Morphological Generation in the Limited-Resource Setting**  
Location: University of Pennsylvania  
Host: Sihao Chen
- 10/2019 **Transfer Learning for Low-Resource Natural Language Processing**  
Location: CIMAT, Guanajuato  
Event: PLAGAA  
Host: Adrian Pastor López Monroy
- 07/2019 **Morphological Generation in the Limited-Resource Setting**  
Location: Universität Stuttgart  
Host: Jonas Kuhn
- 03/2019 **Morphological Generation in the Limited-Resource Setting**  
Location: IBM Zurich  
Host: Jasmina Bogojeska

- 03/2019      **Morphological Generation in the Limited-Resource Setting**  
Location: Google Zurich  
Host: Aliaksei Severyn
- 02/2019      **Morphological Generation in the Limited-Resource Setting**  
Location: University of Colorado  
Host: Martha Palmer
- 02/2019      **Morphological Generation in the Limited-Resource Setting**  
Location: Carnegie Mellon University  
Host: Yulia Tsvetkov
- 07/2018      **Low-resource Morphological Generation with Sequence-to-Sequence Models**  
Location: University of Melbourne  
Host: Daniel Beck
- 10/2017      **Low-resource Morphological Generation with Sequence-to-Sequence Models**  
Location: New York University  
Hosts: Sam Bowman and Kyunghyun Cho
- 10/2017      **Low-resource Morphological Generation with Sequence-to-Sequence Models**  
Location: Johns Hopkins University  
Host: David Yarowsky
- 09/2017      **Neural Sequence-to-Sequence Models for Low-Resource Morphology**  
Location: University of Copenhagen  
Host: Anders Søgaard
- Invited lightning talks**
- 10/2019      **Low-Resource Languages: A Challenge for Natural Language Processing**  
Location: University of Michigan  
Event: Michigan AI Symposium  
Host: Jenna Wiens