Siyao (Logan) Peng

loganpeng1992@gmail.com | +49-160-1433-593 | Munich, Germany |

Education

Ph.D. in Linguistics, Georgetown University, *Washington, DC, USA* 2017–2023

Committee: Amir Zeldes, Nathan Schneider, Nianwen Xue.

Dissertation: *Cross-Paragraph Discourse Structure in Rhetorical Structure Theory Parsing and Treebanking for Chinese and English.*

M.A. in Linguistics, Leiden University, *Leiden, the Netherlands* 2015–2016

B.A. in Applied Mathematics, French & Linguistics, Univ. of California, Berkeley, *CA, USA* 2011–2015

Appointments

Postdoctoral Researcher, Ludwig Maximilian University (LMU), *Munich, Germany* 04/2023–03/2026

Research Group: Munich AI & NLP (MaiNLP) Lab; PI: Barbara Plank

Research Topics: human label variation, explanation-based evaluation, multilingual and dialectal NLP.

Internships

Baidu, PaddlePaddle Deep Learning Group, *Beijing, China* 01/2021–06/2021

Tencent, AI Data Center, *Beijing, China* 05/2020–09/2020

Pearson, Educational Applications Group, *Boulder, CO, USA* 06/2019–08/2019

Publications

Benedetta Muscato, Beiduo Chen, Gizem Gezici**, Siyao Peng**, Barbara Plank, and Fosca Giannotti. 2026. Disagreeing Rationales: Rethinking Classification and Explainability Evaluation in Hate Speech Detection. *Findings of EMNLP 2026.*

Ercong Nie, **Siyao Peng**, Helmut Schmid, and Hinrich Schütze. 2026. Parsing Middle High German: exploring cross-lingual NLP for treebank construction in low-resource historical languages. *Digital Scholarship in the Humanities.*

Longfei Zuo, Barbara Plank**, and Siyao Peng**. 2026. EVADE: LLM-Based Explanation Generation and Validation for Error Detection in NLI. *Findings of ACL 2026.*

Pingjun Hong, Beiduo Chen**, Siyao Peng**, Marie-Catherine de Marneffe, Benjamin Roth, and Barbara Plank. 2026. Agree, Disagree, Explain: Decomposing Human Label Variation in NLI through the Lens of Explanations. *Findings of ACL 2026.*

Renhao Pei*, **Siyao Peng***, Verena Blaschke, Robert Litschko, and Barbara Plank. 2026. Information Asymmetry across Language Varieties: A Case Study on Cantonese-Mandarin and Bavarian-German QA. *LREC 2026.*

Pingjun Hong, Beiduo Chen**, Siyao Peng**, Marie-Catherine de Marneffe, and Barbara Plank. 2025. LiTEx: A Linguistic Taxonomy of Explanations for Understanding Within-Label Variation in Natural Language Inference. *EMNLP 2025.* **SAC Highlight Award**.

Shijia Zhou*, **Siyao Peng***, Simon Luebke, Jörg Haßler, Mario Haim, Saif M. Mohammad, and Barbara Plank. 2025. What Media Frames Reveal About Stance: A Dataset and Study about Memes in Climate Change Discourse. *Findings of EMNLP 2025.*

Elisa Leonardelli, Silvia Casola**, Siyao Peng**, Giulia Rizzi, Valerio Basile, Elisabetta Fersini, Diego Frassinelli, Hyewon Jang, Maja Pavlovic, Barbara Plank, and Massimo Poesio. 2025. LeWiDi-2025 at NLPerspectives: Third Edition of the Learning with Disagreements Shared Task. *NLPerspectives Workshop (EMNLP 2025).*

Tomas Ruiz**, Siyao Peng**, Barbara Plank, and Carsten Schwemmer. 2025. BoN Appetit Team at LeWiDi-2025: Best-of-N Test-time Scaling Can Not Stomach Annotation Disagreements (Yet). *NLPerspectives Workshop (EMNLP 2025).*

Silvia Casola*, Yang Janet Liu*, **Siyao Peng***, Oliver Kraus, Albert Gatt, and Barbara Plank. 2025. References Matter: Investigating the Impact of Reference Set Variation on Summarization Evaluation. *INLG 2025.*

Beiduo Chen**, Siyao Peng**, Anna Korhonen, and Barbara Plank. 2025. A Rose by Any Other Name: LLM-Generated Explanations Are Good Proxies for Human Explanations to Collect Label Distributions on NLI. *Findings of ACL 2025*.

Amir Zeldes, Tatsuya Aoyama, Yang Janet Liu**, Siyao Peng**, Debopam Das, and Luke Gessler. 2025. eRST: A Signaled Graph Theory of Discourse Relations and Organization. *Computational Linguistics*.

Jiawen Wang*, Longfei Zuo***, Siyao Peng**, and Barbara Plank. 2024. MultiClimate: Multimodal Stance Detection on Climate Change Videos. *NLP4PI Workshop (EMNLP 2024)*.

Beiduo Chen, Xinpeng Wang**, Siyao Peng**, Robert Litschko, Anna Korhonen, and Barbara Plank. 2024. "Seeing the Big through the Small": Can LLMs Approximate Human Judgment Distributions on NLI from a Few Explanations? *Findings of EMNLP 2024*.

Leon Weber-Genzel*, **Siyao Peng***, Marie-Catherine De Marneffe, and Barbara Plank. 2024. VariErr NLI: Separating Annotation Error from Human Label Variation. *ACL 2024*. **SAC Award: Resources and Evaluation**.

Shijia Zhou*, **Siyao Peng***, and Barbara Plank. 2024. CLIMATELI: Evaluating Entity Linking on Climate Change Data. *ClimateNLP Workshop (ACL 2024)*.

Siyao Peng, Zihang Sun, Huangyan Shan, Marie Kolm, Verena Blaschke, Ekaterina Artemova, and Barbara Plank. 2024. Sebastian, Basti, Wastl?! Recognizing Named Entities in Bavarian Dialectal Data. *LREC-COLING 2024*.

Yilun Zhu**, Siyao Peng**, Sameer Pradhan, and Amir Zeldes. 2024. SPLICE: A Singleton-Enhanced PipeLIne for Coreference REsolution. *LREC-COLING 2024*.

Verena Blaschke, Barbara Kovačić**, Siyao Peng**, Hinrich Schütze, and Barbara Plank. 2024. MaiBaam: A Multi-Dialectal Bavarian Universal Dependency Treebank. *LREC-COLING 2024*.

Siyao Peng, Zihang Sun, Sebastian Loftus, and Barbara Plank. 2024. Different Tastes of Entities: Investigating Human Label Variation in Named Entity Annotations. *Unimplicit Workshop (EACL 2024)*.

Axel Sorensen**, Siyao Peng**, Barbara Plank, and Rob van der Goot. 2024. EEVEE: An Easy Annotation Tool for Natural Language Processing. *LAW Workshop (EACL 2024)*.

Yilun Zhu**, Siyao Peng**, Sameer Pradhan, and Amir Zeldes. 2023. Incorporating Singletons and Mention-based Features in Coreference Resolution via Multi-task Learning for Better Generalization. *IJCNLP-AACL 2023*.

Tatsuya Aoyama, Shabnam Behzad, Luke Gessler, Lauren Levine, Jessica Lin, Yang Janet Liu**, Siyao Peng**, Yilun Zhu, and Amir Zeldes. 2023. GENTLE: A Genre-Diverse Multilayer Challenge Set for English NLP and Linguistic Evaluation. *LAW Workshop (ACL 2023)*.

Siyao Peng. 2023. Cross-Paragraph Discourse Structure in Rhetorical Structure Theory Parsing and Treebanking for Chinese and English. *Ph.D. dissertation, Georgetown University*.

Siyao Peng, Yang Janet Liu, and Amir Zeldes. 2022. GCDT: A Chinese RST Treebank for Multigenre and Multilingual Discourse Parsing. *AACL-IJCNLP 2022*.

Luke Gessler, Shabnam Behzad, Yang Janet Liu**, Siyao Peng**, Yilun Zhu, and Amir Zeldes. 2021. DisCoDisCo at the DISRPT2021 Shared Task: A System for Discourse Segmentation, Classification, and Connective Detection. *DISRPT Shared Task (EMNLP 2021)*.

Luke Gessler**, Siyao Peng**, Yang Liu, Yilun Zhu, Shabnam Behzad, and Amir Zeldes. 2021. Overview of AMALGUM – Large Silver Quality Annotations across English Genres. *SCiL 2021*.

Haijiang Wu, Zixuan Wang, Qingsong Ma, Xinjie Wen, Ruichen Wang, Xiaoli Wang, Yulin Zhang, Zhipeng Yao**, and Siyao Peng**. 2020. Tencent submission for WMT20 Quality Estimation Shared Task. *WMT 2020*.

Michael Kranzlein, Emma Manning**, Siyao Peng**, Shira Wein, Aryaman Arora, and Nathan Schneider. 2020. PASTRIE: A Corpus of Prepositions Annotated with Supersense Tags in Reddit International English. *LAW Workshop (COLING 2020)*.

Siyao Peng, Yang Liu, Yilun Zhu, Austin Blodgett, Yushi Zhao, and Nathan Schneider. 2020. A Corpus of Adpositional Supersenses for Mandarin Chinese. *LREC 2020*.

Luke Gessler**, Siyao Peng**, Yang Liu, Yilun Zhu, Shabnam Behzad, and Amir Zeldes. 2020. AMALGUM – A Free, Balanced, Multilayer English Web Corpus. *LREC 2020*.

Yue Yu*, **Siyao Peng***, and Grace Hui Yang. 2019. Modeling Long-Range Context for Concurrent Dialogue Acts Recognition. *CIKM 2019*.

Yue Yu, Yilun Zhu, Yang Liu, Yan Liu**, Siyao Peng**, Mackenzie Gong, and Amir Zeldes. 2019. GumDrop at the DISRPT2019 Shared Task: A Model Stacking Approach to Discourse Unit Segmentation and Connective Detection. *DISRPT Workshop (NAACL 2019)*.

Siyao Peng and Amir Zeldes. 2018. All Roads Lead to UD: Converting Stanford and Penn Parses to English Universal Dependencies with Multilayer Annotations. *LAW-MWE-CxG Workshop (COLING 2018)*.

Grants

Imminent Research Grants — $20,000 2025–2026

Project: The First Multilingual Multi-Annotator NLI Benchmark with Ecologically Valid Explanations

Investigators: Barbara Plank, **Siyao Peng**, Marie-Catherine de Marneffe

Teaching

Instructor at LMU Munich (Germany)	**2023–2026**
Lecture: Introduction to Natural Language Processing	Winter 2023–24, 2024–25, 2025–26
Lecture: Trustworthy Data-Centric AI	Winter 2024–25, 2025–26
Lecture: Symbolic Programming Language	Winter 2023–24
Seminar: Linguistic Annotation Frameworks I (LAF I) – Introduction	Winter 2023–24, 2024–25, 2025–26
Seminar: Explaining and Interpreting Annotations in NLP	Winter 2025–26
Seminar: Natural Language Processing for Climate Change	Summer 2024, 2025
Seminar: LAF II – Multilingual Natural Language Inference	Summer 2025
Seminar: LAF II – Variation, Explanation & Validation	Summer 2024
Teaching Assistant at Georgetown University (USA)	**2018–2021**
LING-367: Computational Corpus Linguistics	Fall 2021
LING-001: Introduction to Language	Fall 2020
LING-469: Analyzing Language Data with R	Spring 2019, 2020
LING-462/COSC-482: Statistical Machine Translation	Spring 2020
LING-362: Introduction to Natural Language Processing	Fall 2018
Teaching Assistant at Stony Brook University (USA)	**2016–2017**
LIN-230: Languages of the World	Spring 2017
LIN-200: Language in the U.S.	Fall 2016

Thesis Supervision

Beyond Noise: Detecting Annotation Error from Human Label Variation Using LLMs (*Master*, Longfei Zuo)	Summer 2025
Steering the Climate Change Discourse to East Asia: A Teacher-Student Framework for Frame and Stance Alignment (*Master*, Anna Barwig)	Summer 2025
Developing a Dependency Treebank for Bavarian Tweets (*Bachelor*, Nataliia Kybaliuk)	Winter 2024–25
Mining for Errors in Benchmarks (*Master*, Anna Minaeva)	Summer 2024
Multilingual Stance Detection for Climate Change (*Master*, Liuxu Lu)	Summer 2024
Creating a POS-Tagged Corpus for Bavarian Tweets (*Bachelor*, Arina Sen)	Summer 2024

Academic Service

Co-Organizer: PANDORA (1st Pluralistic AI & NLP Workshop: Diversity-aware, Sociotechnical, Responsible Alignment), co-located with EMNLP 2026

Co-Organizer: LeWiDi (Learning With Disagreements Shared Task), held as part of the NLPerspectives workshop at EMNLP 2025

Co-Chair: LAW XIX (19th Linguistic Annotation Workshop), co-located with ACL 2025

Officer: SIGANN (Special Interest Group on Annotation), 2025–present

Area Chair: ACL Rolling Review (Dec 2024; Feb/May/Oct 2025; Jan 2026); *SEM 2025/2026; LREC 2026;

COLING 2025; LREC-COLING 2024.

Reviewer: ACL Rolling Review (Nov 2021; Jan/Oct/Dec 2022; Feb/Jun/Oct/Dec 2023; Feb/Jun/Oct 2024; Mar 2026); EMNLP 2021–2023; ACL 2020/2023; ACL-IJCNLP 2021; NAACL-HLT 2021; EACL 2021; LRE 2025; LAW 2023/2024/2026; VarDial 2026; LeWiDi 2025; NLP4Sustain 2025; CoMeDi 2024; WiNLP 2019–2022; ACL-IJCNLP 2021 SRW; NAACL-HLT 2021 SRW.

Skills

Programming: Python, R, MATLAB, SQL

ML: PyTorch, TensorFlow, Keras, scikit-learn, NumPy, pandas, SciPy

NLP & LLMs: Hugging Face (Transformers, Datasets), spaCy, NLTK, Stanza, OpenAI API, Gemini API

Tools: Linux, Bash, Git, Jupyter, LaTeX, Google Cloud, AWS

Languages: Chinese (native), English (fluent), German (B2), French (intermediate)