Junghyun Min

jm3743@georgetown.edu | Google Scholar | https://aatlantise.science/ | Washington, DC

EDUCATION

Georgetown University

Doctor of Philosophy in Linguistics. Computational linguistics concentration.   2024 – Present
    Advised by Ethan Wilcox.

Johns Hopkins University

Master of Arts in Cognitive Science. Computational approaches to linguistics focus.   2019 – 2020
    Advised by Tal Linzen.

Bachelor of Science in Physics with a second major in Mathematics.   2014 – 2017
    Early graduation, General honors.

PAPERS

2026   Minho Lee, **Junghyun Min**, Woochul Lee, Yeonsoo Lee. 2024. Structured Language Generation Model: Loss Calibration and Formatted Decoding for Robust Structure Prediction. In *Proceedings of the AAAI 2026 Workshop on New Frontiers in Information Retrieval*. Singapore.

2025   **Junghyun Min**, Xiulin Yang, and Shira Wein. 2025. When Does Meaning Backfire? Investigating the Role of AMRs in NLI. In *Proceedings of the 14th Joint Conference on Lexical and Computational Semantics (*SEM 2025)*, pages 202–211, Suzhou, China. Association for Computational Linguistics.

   Lauren Levine, **Junghyun Min**, and Amir Zeldes. 2025. Building UD Cairo for Old English in the Classroom. In *Proceedings of the Eighth Workshop on Universal Dependencies (UDW, SyntaxFest 2025)*, pages 97–104, Ljubljana, Slovenia. Association for Computational Linguistics.

   Junghyun Min, Minho Lee, Woochul Lee, Yeonsoo Lee. 2025. Punctuation Restoration Improves Structure Understanding without Supervision. In *Proceedings of the Tenth Representation Learning for NLP Workshop*. Association for Computational Linguistics.

   Abhishek Purushothama, **Junghyun Min**, Brandon Waldon, and Nathan Schneider. 2025. Not ready for the bench: LLM legal interpretation is unstable and uncalibrated to human judgments. ArXiv preprint.

   Hannah Liu, **Junghyun Min**, En-Shiun Annie Lee, Ethan Yue Heng Cheung, Shou-Yi Hung, Syed Mekael Wasti, Runtong Liang, ShiYao Qian, Shizhao Zheng, Elsie Chan, Ka Ieng Charlotte Lo, Wing Yu Yip, and Richard Tzong-Han Tsai. SiniticMTError: A Machine Translation Dataset with Error Annotations for Sinitic Languages. ArXiv preprint.

   Junghyun Min, York Hay Ng, Sophia Chan, Helena Shunhua Zhao, En-Shiun Annie Lee. 2025. CantoNLU: A benchmark for Cantonese natural language understanding. ArXiv preprint.

2024   Minho Lee, **Junghyun Min**, Woochul Lee, Yeonsoo Lee. 2024. Structured Language Generation Model for Robust Structure Prediction. ArXiv preprint.

2020   **Junghyun Min**, R. Thomas McCoy, Dipanjan Das, Emily Pitler, and Tal Linzen. 2020. Syntactic Data Augmentation Increases Robustness to Inference Heuristics. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2339–2352, Online. Association for Computational Linguistics.

   R. Thomas McCoy, **Junghyun Min**, and Tal Linzen. 2020. BERTs of a feather do not generalize together: Large variability in generalization across models with similar test set performance. In *Proceedings of the Third BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*, pages 217–227, Online. Association for Computational Linguistics.

PRESENTATIONS & TECHNICAL REPORTS

2025   Abhishek Purushothama, **Junghyun Min**, Brandon Waldon, and Nathan Schneider. 2025. [Not ready for the bench: LLM legal interpretation is unstable and uncalibrated to human judgments](#). In *Proceedings of the Natural Legal Language Processing Workshop 2025*, pages 317–317, Suzhou, China. Association for Computational Linguistics.

       Hannah Liu, **Junghyun Min**, En-Shiun Annie Lee, Ethan Yue Heng Cheung, Shou-Yi Hung, Syed Mekael Wasti, Runtong Liang, ShiYao Qian, Shizhao Zheng, Elsie Chan, Ka Ieng Charlotte Lo, Wing Yu Yip, and Richard Tzong-Han Tsai. [SiniticMTError: A Machine Translation Dataset with Error Annotations for Sinitic Languages](#). Workshop on Multilingual Data Quality Signals at COLM 2025.

       York Hay Ng, Yihe Li, Shizhao Zheng, **Junghyun Min**, and Eu-Shiun Annie Lee**.** [Bridging Worlds with Words: Innovations at the Lee Language Lab](#). Toronto Machine Learning Summit. 2025.

2024   **Junghyun Min**. Feb 2024. [Unsupervised structure understanding improvement via punctuation restoration](#). In Korean. NC Research Blog entry. NCSOFT Corp.

2023   **Junghyun Min**. Feb 2023. [RRR: A Robust Road to Robustness](#). In Korean. NC Research Blog entry. NCSOFT Corp.

       Junghyun Min. Oct 2023. Job talk at Samsung Electronics. Received employment offer.

2022   **Junghyun Min.** Jan 2022. Open information extraction and the granularity problem. Presentation at NC AI & NLP Seminar. NCSOFT Corp.

2021   **Junghyun Min**. Jan 2021. Job talk at NCSOFT. Received employment offer.

       Junghyun Min. Jan 2021. Job talk at Kakao Brain.

2020   **Junghyun Min**. Heuristics in language models and syntactic augmentation to mitigate them. Presentation at Department of Cognitive Science Brown Bag Series. Johns Hopkins University.

2019   R. Thomas McCoy, **Junghyun Min**, Tal Linzen. Nov 2019. BERTs of a feather do not generalize together: large variability in generalization across models with similar test set performance. Poster presentation at NLP, Dialogue and Speech Symposium. New York Academy of Science.

2018   **Junghyun Min** and Giorgia Fortuna. May 2018. What on Earth is this map? Finding a map's depicted region and projection. Poster presentation at Day of Undergraduate Research in Engineering, Arts, and Medicine. Johns Hopkins University.

2013   Michał Leś and **Junghyun Min**. Aug 2013. The atmospheric flux of muon. Presentation at International Summer Science Heidelberg. Max-Planck-Institut für Kernphysik.

TEACHING

Georgetown University

Teaching assistant, *Computational Langue Processing, Department of Linguistics.*	2026
Teaching assistant, *Introduction to Language (LING1000), Department of Linguistics.*	2025

University of Toronto

Course supervisor, *Computer Science Project (CS494), Department of Computer Science.*	2025

Johns Hopkins University

Teaching assistant, *Differential Equations (110.302), Department of Mathematics.*	2017

MENTORING

Georgetown University

Graduate mentor, *Summer research assistantship.* Mentored Elli Ahn.	2025

University of Toronto

Research team lead, Lee Language Lab. Supervised Sophia Chan, York Ng. 2025

Fields Institute for Research in Mathematical Sciences
Co-supervisor, *Fields Undergraduate Summer Research Program.* Supervised Rex Fang and Michael Zhou. 2025

NCSOFT
Mentor, *Fall Undergraduate Co-op Internship.* Mentored Yerang Kim. 2023
Lead mentor, *Summer Undergraduate Internships.* Mentored Eva Goldie. 2022

PROFESSIONAL & NON-ACADEMIC RESEARCH EXPERIENCE

University of Toronto, Toronto, ON 2025
Visiting Scholar, Department of Computer Science

NCSOFT, Seongnam, Korea 2021 – 2024
NLP Engineer, Financial Language Understanding

Harford Community College, Bel Air, MD 2018 – 2019
Data Analyst, Analytics & Planning

Wolfram Research, Waltham, MA 2016 – 2017
Student Researcher, Wolfram Summer School

Max-Planck-Institut für Kernphysik, Heidelberg, Germany 2013
Student Researcher, Nuclear Physics

SCHOLARSHIPS & AWARDS

Dean's Master's Fellowship, Johns Hopkins University 2019
The John Ernest Fellowship, The John Ernest Foundation 2013

LANGUAGES

Computer Language: Python, R, Java, C++, Unix shell, SAS, SQL, Wolfram.
Natural Language: Korean (Standard, Busan), English, German, Mandarin Chinese.
Language with Research Experience: Old English, Middle Korean, Cantonese

SELECTED PROJECTS

Lead engineer, ai.ly.
- Fine-tuned and deployed a GPT-based, personalized AI lyricist. 50k visits over 3 months of service. Hip-hop sample.

Lead engineer, genDOC.
- Designed and built RAG-based agentic document automation software for startups with LangChain on OpenAI API.

SERVICE

Assistant Coordinator, Computation and Language at Georgetown University 2025
Recreation Coordinator, Graduate Linguistics Student Association at Georgetown University 2025
Reviewer, ACL Rolling Review, Machine Learning Engineering, LREC Since 2021
Captain, Varsity Football at Western Reserve Academy 2013