

Junghyun Min

Washington, DC | jm3743@georgetown.edu | [Google Scholar](#) | <https://www.aatlantise.science/>

EDUCATION

Georgetown University

Doctor of Philosophy in Linguistics

Washington, DC

Aug 2024 – Current

- Computational linguistics concentration, Linguistics Graduate Scholarship

Johns Hopkins University

Master of Arts in Cognitive Science

Baltimore, MD

Aug 2019 – Dec 2020

- Research track; concentration in computational approaches to linguistics; 3.91 GPA

Bachelor of Science in Physics with a second major in Mathematics

Sep 2014 – Dec 2017

- Dean's List Spring 2015, Fall 2017; General Honors; KGSA Baseball Captain; 3.53 GPA

Wolfram Summer School

Science Track

Waltham, MA

Jun 2016 – Jul 2016

- Courses on computational topics such as deep learning, image processing, optimization in Wolfram language

PUBLICATIONS & PROJECTS

NCSOFT

- Tech blog: **Junghyun Min**. March 2024. [Punctuation restoration improves structure understanding without supervision \(Korean\)](#). [Relevant preprint](#).

Johns Hopkins University

- ACL 2020 paper: **Junghyun Min**, R. Thomas McCoy, Dipanjan Das, Emily Pitler, and Tal Linzen. 2020.

[Syntactic data augmentation increases robustness to inference heuristics](#).

- BlackboxNLP 2020 paper: 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](#).

Other

- Project with forus.ai: Technical lead at ai.ly, a GPT-2 based AI lyricist that reflects the user's personality. 50k visits over 3 months of live service. Its first release: [hip-hop song](#)

WORK & RESEARCH EXPERIENCE

Georgetown University

Graduate Research Assistant, Department of Linguistics

Washington, DC

August 2024 – Current

- Re-annotate the Korean Semantic Network of Adposition and Case Supersenses dataset for UD compatibility.

NCSOFT

NLP Engineer, NLP Center

Seongnam, Korea

Jan 2021 – Apr 2024

- Developed and served fast (10 reqs/ms) and accurate chunking on < 4GB VRAM by concatenating embeddings.
- Designed an effective granularity control algorithm in open information extraction with syntax parsing depth.
- Led mentorships for the Center's 2022 Language AI Global Summer Internship program with 10+ interns.
- Proposed punctuation restoration as an unsupervised training objective to improve structure understanding.
- Expanded entity, relation extraction to facilitate downstream tasks in financial market sensing, drug discovery.

Johns Hopkins University

Graduate Research Assistant, Computation and Psycholinguistics Laboratory

Baltimore, MD

Jul 2019 – Oct 2020

- First-authored ACL publication, second-authored BlackboxNLP publication. Presented at ACL and NYAS NDS.
- BERT fine-tuned on MNLI is unstable and vulnerable to syntactic heuristics ([McCoy, Min, Linzen 2020](#)).
- Adversarial data augmentation via syntactic manipulation of training set data significantly increases robustness to augmentation-like examples and general syntactic sensitivity too ([Min, McCoy, Das, Pitler, Linzen 2020](#)).
- Heuristics likely arise from both the pre-training and the fine-tuning dataset. Currently popular fine-tuning and evaluation paradigm has drawbacks that can be patched with longer fine-tuning on unbiased datasets, multi-seed out-of-distribution evaluation, and syntactic adversarial augmentation ([Master's thesis](#)).

Harford Community College

Research Associate, Department of Analytics and Planning

Bel Air, MD

Mar 2019 – Jul 2019

Research Assistant, Department of Analytics and Planning

Apr 2018 – Mar 2019

SKILLS & INTERESTS

Natural & Computer Language

Korean, English, German, Mandarin Chinese, Python, SAS, SQL, Wolfram language, R, Unix

Interests

Geography, cartography; baseball analytics, sabermetrics; urban planning, public transport; low-cost travel.