Yige Chen

G19, Leung Kau Kui Bldg.
CUHK, Hong Kong
✉ yigechen.cuhk@outlook.com
🌐 yigechen.com

Education

2022–Present   **Ph.D. in Linguistics**, *Chinese University of Hong Kong*, Hong Kong
Supervisor: Prof. Haihua Pan
Thesis topic: Language Model Interplay: Collaboration and Competition

2021–2022   **M.S. in Computational Linguistics**, *University of Washington*, Seattle, WA
Advisor: Prof. Fei Xia
Grade: 4.00/4.00

2020–2021   **M.Phil. in Theoretical and Applied Linguistics**, *University of Cambridge*, Cambridge
Supervisor: Prof. Nigel Collier
Grade: Distinction
Thesis title: Data augmentation using machine translation for improved stance detection in the finance domain

2017–2020   **B.S. in Economics, Linguistics, Mathematics, Asian Studies**, *University of Wisconsin–Madison*, Madison, WI
Certificate in Computer Sciences
Grade: 3.97/4.00
Graduated with Distinction

Research Interests

Natural Language Processing, Computational Linguistics, Machine Learning, Conversational Agents, Information Extraction, Transfer Learning, Information Systems, Syntax, Semantics, Behavioral Economics

Publications

2024   **Yige Chen**, Kyuwon Kim, KyungTae Lim, Jungyeul Park, and Chulwoo Park. When the Misidentified Adverbial Phrase Functions as a Complement. *Findings of the Association for Computational Linguistics: EMNLP 2024*, 2024.

Yige Chen, KyungTae Lim, and Jungyeul Park. A Linguistically-Informed Annotation Strategy for Korean Semantic Role Labeling. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 2024.

Yige Chen, Jae Ihn, KyungTae Lim, and Jungyeul Park. Towards Standardized Annotation and Parsing for Korean FrameNet. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 2024.

2023   **Yige Chen**, KyungTae Lim, and Jungyeul Park. Korean Named Entity Recognition Based on Language-Specific Features. *Natural Language Engineering*, 2023.

2022	**Yige Chen**, Eunkyul Leah Jo, Yundong Yao, KyungTae Lim, Miikka Silfverberg, Francis Tyers, and Jungyeul Park. Yet Another Format of Universal Dependencies for Korean. *Proceedings of the 29th International Conference on Computational Linguistics (COLING)*, 2022.

Awards and Honors

2022	Hong Kong PhD Fellowship
2022	CUHK Vice-Chancellor HKPFS Scholarship
2020	Phi Beta Kappa
2019, 2020	UW-Madison Undergraduate Scholarship for Summer Study
2019	Juli Plant Grainger Academic Excellence Scholarship in Economics

Working Experience

2021–2022	**Research Assistant**, *Department of Information Systems, Business Statistics and Operations Management, School of Business and Management, Hong Kong University of Science and Technology*, Hong Kong Supervisor: Prof. Yi Yang
2021	**Financial NLP Intern**, *Global Markets Department, China Merchants Bank*, Shanghai
2019–2020	**Lab Assistant**, *Second Language Acquisition Lab (SLA Lab), Department of English, University of Wisconsin–Madison*, Madison, WI Supervisor: Prof. Jacee Cho

Teaching Experience

2022–Present	**Teaching Assistant**, *Chinese University of Hong Kong*, Hong Kong Courses: ○ LING3401 Linguistics and Information Technology (Spring 2025) ○ LING2006 Semantics (Spring 2024) ○ LING2005 Syntax I (Fall 2023, Spring 2023) ○ LING3204 Second Language Acquisition (Fall 2022)
2018–2020	**Language Tutor**, *Greater University Tutoring Services (GUTS), University of Wisconsin–Madison*, Madison, WI Language: Mandarin Chinese

Talks

2024	Yue Chen, Ziyin Mai, and **Yige Chen**. Employing Machine Learning and Language Models to Differentiate Language Patterns in Mandarin-Speaking Preschoolers with Autism Spectrum Disorder. Poster presentation. Seventh Installment of California Meeting on Psycholinguistics (CAMP[7]), University of California San Diego, San Diego, CA, United States, 2024.

Yige Chen. Understanding the Semantic Representation of Korean Mixed Script in Language Models. Oral presentation. 30th Annual Conference of the International Association of Chinese Linguistics (IACL-30), Yonsei University, Seoul, South Korea, 2024.

Yige Chen. Developing Resources and Tools for Enhanced Korean Language Processing. Invited talk. Multimodal Language Processing Lab, Seoul National University of Science and Technology, Seoul, South Korea, 2024.

Academic Service

Reviewer — Generative Linguistics in the Old World in Asia, Annual Conference of the International Association of Chinese Linguistics

Memberships

2022, 2024 — Association for Computational Linguistics
2024 — European Language Resources Association
2024 — International Association of Chinese Linguistics
2022–2023 — Linguistic Society of America

Skills

Operating Systems — Windows, Linux, macOS

Programming — Java, Python (NumPy, pandas, PyTorch, Transformers), MATLAB, Visual Basic, R, SQL

Software and Tools — STATA, MongoDB, LaTeX, Maple, Microsoft Office (Word, Excel, PowerPoint), Praat, E-Prime, Adobe Photoshop, Adobe Audition

Languages — Wu Chinese (native), English and Mandarin Chinese (fluent), Korean (intermediate), Japanese, Spanish and Cantonese (basic), Classical Chinese (fluent in written language)