

# Junjie Chen

junjiechen.chris@outlook.com | LinkedIn | Github | Google Scholar

## SUMMARY

---

Passionate NLP researcher with over five years of experience in integrating language models into syntax and semantics studies. Skilled in various ML libraries (e.g., Pytorch and lightning) and Responsible as mid-size cluster maintainers. Excellent problem-solving and research abilities, demonstrated by peer-reviewed papers published at top NLP conferences (e.g., ACL and NAACL). Seeking to leverage past experiences in AI research and development to contribute to cutting-edge projects.

## EDUCATION

---

### The University of Tokyo

Ph.D. in Computer Science, advised by [Yusuke Miyao](#)

Tokyo, Japan

Oct. 2021 – Mar. 2025

### University of Liverpool

Exchange Student, advised by [Danushka Bollegala](#)

Liverpool, UK

Oct. 2023 – Mar. 2024

### The University of Tokyo

M.S. in Computer Science, advised by [Yusuke Miyao](#)

Tokyo, Japan

Oct. 2019 – Sep. 2021

### Northeastern Univeristy

B.S. in Software Engineering

Shenyang, China

Sep. 2015 – Jun. 2019

## GRANTS

---

Travel Grant (\$2000) sponsored by AAAI-25

DC2 (monthly stipend 200,000 JPY + yearly grant 900,000 JPY) from *Japan Society for the Promotion of Science* (2023 - 2025)

IST-RA (monthly stipend 60,000 JPY) from *The University of Tokyo* (2022 - 2023)

## PUBLICATIONS

---

**Junjie Chen**, Xiangheng He, Yusuke Miyao, and Danushka Bollegala. Improving Unsupervised Constituency Parsing via Maximizing Semantic Information To appear in *ICLR-2025 (spotlight, top 5.1%)*

Xiangheng He, **Junjie Chen**, Zixing Zhang, Björn W Schuller *In proceedings of the AAAI Conference on Artificial Intelligence: AAAI-2025*

**Junjie Chen**, Xiangheng He, Danushka Bollegala, and Yusuke Miyao. Constituents are Frequent Word Sequences among Sentences with Equivalent Predicate-Argument Structures: Unsupervised Constituency Parsing by Span Matching. *In findings of the Association for Computational Linguistics: ACL 2024*

**Junjie Chen**, Xiangheng He, and Yusuke Miyao. Language Model Based Unsupervised Dependency Parsing with Conditional Mutual Information and Grammatical Constraints. *In proceedings of the 2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics*

Xiangheng He, **Junjie Chen** and Bjorn W. Schuller. Task Selection and Assignment for Multi-modal Multi-task Dialogue Act Classification with Non-stationary Multi-armed Bandits. *In proceedings of 2024 International Conference on Acoustics, Speech, and Signal Processing*

**Junjie Chen**. Syntactic-Semantic Dependency Correlation in Semantic Role Labeling: a Shift in Semantic Label Distributions. *Journal of Natural Language Processing, 29(3)*

**Junjie Chen**, Xiangheng He, and Yusuke Miyao. Modeling Syntactic-Semantic Dependency Correlations in Semantic Role Labeling Using Mixture Models. *In Proceedings of 60th Annual Meeting of the Association for Computational Linguistics*

Xiangheng He, **Junjie Chen**, Georgios Rizos and Björn Schuller. An Improved StarGAN for Emotional Voice Conversion: Enhancing Voice Quality and Data Augmentation. *In proceedings of Interspeech 2021*

Zihong Liang, **Junjie Chen**, Zhaopeng Xu, Yuyang Chen, Tianyong Hao: A Pattern-Based Method for Medical Entity Recognition From Chinese Diagnostic Imaging Text. *Frontiers Artif. Intell. 2: 1 (2019)*

Xieling Chen, Juntao Hao, **Junjie Chen**, Songshou Hua and Tianyong Hao. A Bibliometric Analysis of the Research Status of the Technology Enhanced Language Learning. *SETE@ICWL (2018)*.

## EXPERIENCE

---

### Research Intern

*Rakuten*

Nov. 2024 – Feb. 2025

*Tokyo, Japan*

- Worked on Sparse Autoencoder-based Personal Information Detection

### Research Assistant

*National Institute of Informatics*

Jun. 2024 – Feb. 2025

*Tokyo, Japan*

- Assist on the development of Japanese-based Large Language Model
- Assist on maintaining large-scale clusters for LLM training

### Research Assistant

*The University of Tokyo*

Dec. 2021 – Apr. 2023

*Tokyo, Japan*

- Conducted research on extracting dependency syntax through measuring word-wise statistical dependence
- Incorporated grammatical constraints to a Metropolis-Hasting sampling-based estimation of statistical dependence
- Research outcome published in NAACL 2024

## PROJECTS

---

### Mixture of Expert-based Semantic Role Labeling System

*Master Thesis*

Oct. 2020 - Sep. 2021

- Developed a Mixture-of-Expert (MoE) based Semantic Role Labeling system
- Deciding the assignment of expert using syntax information
- Research outcome published in ACL 2022

### Multilingual Article Classification System

*Application Development*

Apr. 2020 – Sep. 2020

- Developed an article classifier for selecting reliable and useful information sources for [Covid-19 Information Portal](#)
- Achieved F1 score of 80% for selecting Covid-related articles and over 60% for selecting useful or clear articles using Japanese-BERT classifier