ZHANG ZIQI

E-mail：22101842d@connect.polyu.hk

Tel：+852 4671 4490

EDUCATION

The Hong Kong Polytechnic University ***2022/8 - 2026/7***

B.A. *Linguistics & Translation;*
Secondary Major in Artificial Intelligence & Data Analytics

- **Award GPA:** 3.41/4.3
- **Coursework:** Data Structure and Algorithms, Artificial Intelligence, Machine Learning, Operations Research Methods, Programming and Data Analytics in Linguistics Perspectives, Techniques in Language Science, Semantics, Sound and Sound System
- **Research Interests:** LLM Interpretability, Multimodal Models, Integration of cognitive science with large models

PUBLICATIONS

[1] Ma, J., Feng, Z., Chersoni, E., Song, H., & **Zhang, Z.** (2025). Phonothink: Improving large language models' reasoning on chinese phonological ambiguities. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*.

[2] **Zhang, Z.**, Ma, J., Chersoni, E., Jieshun, Y., & Feng, Z. (2025). From BERT to LLMs: Comparing and Understanding Chinese Classifier Prediction in Language Models. In *Proceedings of the 8th BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP*.

[3] **Zhang, Z.**, Xu, A., & Hsu, Y. Y. (2026). Pitch Elevation on Wh-Words in Cantonese Spontaneous Speech. In *Hanyang International Symposium on Phonetics and Cognitive Sciences of Language 2026*.

INTERNSHIP & PROJECT EXPERIENCE

Capturing Predictive Processing in Healthy Aging Using Large Language Models: An EEG-fMRI Study ***Apr 2026 - Present***

*Student Research Assistant | **Advisor:** Prof. Mohammad Momenian, HKU*

- Utilized Python scripts to extract Surprisal and Entropy of experimental texts from various large language models, providing reliable computational prediction data for neuroimaging analysis.
- Accurately aligned the log-probability features generated by large models with time-locked EEG/fMRI data to explore the effectiveness of computational linguistic indicators in predicting brain predictive processing mechanisms across the lifespan and in healthy aging populations.

IntEx-Bench: Benchmark for Multi-dimensional Semantic Understanding of LLMs ***Oct 2025 - May 2026***

*Project Leader | **Advisor**: Prof. Emmanuele Chersoni, PolyU*

- Designed and constructed the "IntEx-Bench" unified evaluation benchmark and built a multi-lingual test dataset with over 7,000 entries to systematically explore and differentiate the multiple semantic understanding capabilities of large language models.
- Implemented a dual evaluation framework, assessing explicit reasoning capabilities through zero-shot Chain-of-Thought (CoT) prompt engineering, and evaluating the model's implicit representation capabilities using methods like contextual embedding extraction.
- Independently designed cognitive-based optimization strategies, including an innovative Semantic Feature-CoT framework and targeted supervised fine-tuning (SFT), to enhance the model's interpretive ability for compound words with extreme semantic distances.

Intelligent Health Monitoring System Project Based on Multimodal Large Models ***May 2024 - Sep 2024***

Project Member | CUHK & PolyU

- Participated in the project "Zhihuyoujia - Health Monitoring System Based on Multimodal Large Models" (Project No. CI103), designing an intelligent care system for elderly individuals living alone; the project ultimately advanced to the Hong Kong Regional Finals of the China International College Students' Innovation Competition 2024.
- Key responsibilities included: designing the core multimodal large model architecture of the product, and proposing a feasible solution by integrating natural language processing, speech recognition, speech synthesis, and computer vision technologies.

Curriculum Development Project for "Techniques in Language Science" ***May 2025 – Jun 2025***

*Student Assisstant | **Advisor:** Prof. Tao Ran, PolyU*

- As a student assistant, I contributed to the development and writing of the teaching plan for the Techniques in Language Science (CBS4953) course, offered by the LST Department.
- Introduced interdisciplinary topics that integrate cognitive linguistics with AI, such as the application of EEG technology in LLMs.

Cantonese Spontaneous Speech Corpus and Phonetics Research ***May 2025 - Dec 2025***

Project Leader | ***Co-Advisors:*** *Prof. Xu Anqi, HIT (SZ); Prof. Hsu Yuyin, PolyU*

- Analyzed Mandarin and Cantonese corpora from open-source speech databases Common Voice and OpenSLR as the project leader.
- Key tasks included preprocessing corpus data, writing scripts to filter and extract over 40,000 interrogative sentences and declarative sentences containing interrogative words; performing phoneme-level alignment of speech and text using the Kaldi-based speech model Montreal Forced Aligner; and attempting speech enhancement using the noisereduce library and deep learning methods.
- Wrote Python and Praat scripts to design and implement a complete automated pipeline for batch audio analysis of Chinese conversational speech data.
- Applied Linear Mixed-Effects Models (LMM) to conduct statistical analysis on the data by separating confounding factors such as word position and speakers.

EXTRACURRICULAR ACTIVITIES

PolyU Orchestra ***Sep 2023 - Present***

First Flute

- Participated in the exchange program between Tsinghua University and The Hong Kong Polytechnic University, performing as the flute soloist in the joint concert.
- Participated in major concerts and music festivals such as the HKGNA 2024 Hong Kong New Generation Music Festival.

Southeast University Summer Research Camp ***Jun 2024 - Jul 2024***

Participant

- Completed various studies, field trips, and research activities on economics, history, and culture at Southeast University.

Teaching People to Use AI: Service Learning in Cambodia ***Sep 2025 – Jan 2026***

Team Leader

- Teach primary school students in Cambodia how to utilize AI-related tools to perform simple tasks, such as object detection and basic coding concepts using Scratch and the Google Teachable Machine platform.

SKILLS

- **Skills:** Python, R, Praat, LaTeX, PyTorch
- **Languages:** Mandarin, English, and Cantonese