Shisen Yue

✉ Email | ◉ Personal Website | ⓞ Github

EDUCATION

Shanghai Jiao Tong University   Shanghai, CN
Bachelor of Arts in Linguistics; GPA: 88.4/100   *Sept. 2020 - June 2024 (expected)*

- Honors: Academic Progress Award (Top 10%)
- Relevant Courses:

  * Linguistics: Syntax (97/100), Semantics (91/100), Phonetics and Phonology (90/100)
  * Computer Science: Algorithm (97/100), Data Structure (88/100), Machine Learning (89/100)
  * Math: Probability and Statistics (97/100), Calculus (96/100)

PUBLICATIONS

William Schuler, **Shisen Yue**. (2024). Evaluation of an Algorithmic-level Left-corner Parsing Account of Surprisal Effects. *Cognitive Science* (Under Review)

Byung-Doh Oh, **Shisen Yue**, William Schuler. (2023). Frequency Explains the Inverse Correlation of Large Language Models' Size, Training Data Amount, and Surprisal's Fit to Reading Times. *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics* (EACL)

Yikang Liu, Ziyin Zhang.[†], Wanyang Zhang[†], **Shisen Yue**[†], Xiaojing Zhao[†], Xinyuan Cheng, Yiwen Zhang, Hai Hu. (2023). ArguGPT: evaluating, understanding and identifying argumentative essays generated by GPT models. *arXiv preprint*

Yixin Wang, **Shisen Yue**, Yanyi Zhong. (2023). Understanding differences between human language processing and natural language processing by the synchronized model. *International Conference on Education, Language and Art* (ICELA 2021), 287-294

RESEARCH EXPERIENCE

Large Language Models (LLMs) and Gricean Conversational Implicature
Department of Linguistics, Shanghai Jiao Tong University (Capstone)   09/2023 – Present
Supervisor: Professor Hai Hu   *Shanghai, CN*

- Created a Chinese conversational corpus *SwordsmanCo* with sitcom lines that contain implicatures
- Formulated a benchmark comprising multiple choice questions to test LLMs' pragmatic understanding
- Conducted a subjective assessment with human participants on LLMs' explanation of implicatures
- Analyzed LLMs' pragmatic skills with cooperative principle and through ratings from human evaluators

Producing Surprisal Effects with Left Corner Parsing
Department of Linguistics, Ohio State University   04/2023 – Present
Supervisor: Professor William Schuler   *Remote*

- Simulated sentence parsing with a left corner parser modeled on distributed associative memory
- Adjusted the predictability of grammar rules to create surprisal effects in a Garden-path sentence
- Measured the scale of surprisal to verify the surprisal account of sentence disambiguation

Analyzing and Identifying Essays Generated by GPT Models
Department of Linguistics, Shanghai Jiao Tong University   02/2023 – 07/2023
Supervisor: Professor Hai Hu   *Shanghai, CN*

- Quantified the linguistic features in essays written by AI and human for analysis and detector training
- Developed Support Vector Machine (SVM) detectors to identify AI-generated essays (95.14% accuracy)

Perception of Event Boundaries Among Chinese Speakers

Department of Linguistics, Shanghai Jiao Tong University — 07/2022 – 03/2023

Supervisor: Professor Hui Chang — *Shanghai, CN*

- Led multi-modal (video based) study to test the undeveloped and regressed production and comprehension of aspect markers (-*le* and -*zhe* in Chinese) among children and the elderly
- Observed significant immaturity/attrition ($p < 0.05$) in these age groups compared to young adults
- Designed a video segmentation task to probe into event boundaries in Chinese speakers' cognition
- Revealed hierarchical event perception among Chinese young adults, supporting universality of cognition

AI-Chatbot Discourse Analysis — 07/2021 – 01/2022

Department of Linguistics, Shanghai Jiao Tong University — *Remote*

- Surveyed 115 AI-Chatbot users for feedback on grammatical acceptability and discourse coherence
- Evaluated the responses of Chatbot Replika through a comprehensive synchronized model of language
- Reported its limitation in contextualizing information and in understanding pragmatic implicature

COURSE PROJECTS

Reinforcement Learning with MuJuCo Simulators and Atari Games — 06/2023

CS-3316 Reinforcement Learning, Shanghai Jiao Tong University, [Github] — *Shanghai, CN*

- Trained MuJoCo simulators with policy-based algorithms DDPG and TD3 and Atari agents with value-based algorithms DQN and Dueling DDQN, enhancing agents' performance through epsilon decay and clipping gradient, and attaining scores more than 10% higher than benchmarks provided by OpenAI

Movie Tagging with LDA Model — 09/2022

UCLAX-CSX4501 Machine Learning, UCLA Extension, [Github] — *Remote*

- Preprocessed the data with the Natural Language Toolkit (NLTK) package, made corpus with words from movie descriptions, and extracted topic words via the Latent Dirichlet Allocation (LDA) model

VOLUNTARY & EXTRACURRICULAR EXPERIENCE

Captain of Interdepartmental Soccer Team — 09/2022 – 09/2023

- Led weekly training sessions, formulated strategies, and assigned positions before each match

Head of Contact Center, School Students Union — 09/2021 – 04/2022

- Tasked with fundraising for large events held by the student union, liaising with business leaders
- Coordinated with student unions of other colleges to organize large-scale inter-college activities

Volunteer English Teacher, Chenxi Program — 05/2021 – 06/2021

- Delivered English classes to children from Hongmei No.2 Primary School

SKILLS

Programming: Python, C++, LaTeX, R, Huggingface, Praat, SPSS, E-Prime
Libraries: Natural Language Toolkit (NLTK), PyTorch, Pandas, Numpy, scikit-learn
Language Skills: Mandarin (native), English (proficient), Spanish (DELE B1), Japanese (intermediate)