

# Xiaomeng (Miranda) Zhu

miranda.zhu@yale.edu

xiaomeng-zhu.github.io

- EDUCATION** **Yale University**, New Haven, CT Aug 2023 -  
Ph.D. Student in Linguistics
- Wellesley College**, Wellesley, MA Sep 2019 - May 2023  
Bachelor of Arts in Cognitive and Linguistic Sciences and Computer Science  
*Honors: Summa cum laude*  
*Thesis: The Modified Mother Tongue: Examining Chinese Vowel Quality Shift in Chinese-Dominant Chinese-English Code-Switching Environments*
- PUBLICATIONS** Meng Hui Liu\*, **Xiaomeng Zhu\***, and Carolyn Jane Anderson, XiaoshuoNLP: A Natural Language Processing Pipeline for Chinese Literary Texts (In Progress)
- RESEARCH EXPERIENCE** **Science Center Summer Research Program** May 2022 - Present  
*Advisor: Carolyn Anderson* Wellesley, MA  
Implement an NLP pipeline for processing Chinese literary texts, enhance the accuracy of coreference resolution results by merging character clusters based on BERT embeddings, and evaluate existing Chinese NLP tools on an original literary dataset
- Cognitive and Linguistic Sciences Honors Thesis** May 2022 - May 2023  
*Advisor: Angela Carpenter* Wellesley, MA  
Examine the shift in Chinese vowel quality in Chinese-English insertional code-switching environments, design speech production tasks to elicit code-switched speech from subjects, and run statistical models on collected vowel formants
- Wellesley College Sociolinguistics Lab** Feb 2022 - Aug 2022  
*Advisor: Sabriya Fisher* Wellesley, MA  
Designed and implemented a Python script that extracts morphosyntactic features from texts, examined the frequency distribution of phonetically-similar auxiliaries, and coded phonetic data from the CORAAL corpus
- MIT Experimental Syntax and Semantics Lab** Feb 2021 - May 2022  
*Faculty Advisor: Martin Hackl, Direct Advisor: Leo Rosenstein* Cambridge, MA  
Investigated felicity differences between determiners using an experimental approach, implemented experimental interface on PCIBex, analyzed data in R Studio using a mixed-effect model, and assisted with data interpretation
- MIT Department of Linguistics** Dec 2021 - Mar 2022  
*Advisor: Danfeng Wu* Cambridge, MA  
Conducted in-person experiments with participants to examine the prosodic pattern in either/or constructions, monitored subjects' behavior and recording qualities, and annotated recording files to facilitate prosody extraction
- MIT Media Lab** Jun 2020 - Aug 2020  
*Faculty Advisor: Alex Pentland, Direct Advisor: Mohsen Bahrami* Cambridge, MA  
Implemented an interactive dashboard that visualized mobility behavior and economic growth using react.js
- PRESENTATION** **The Modified Mother Tongue: Examining Chinese Vowel Quality Shift in Chinese-Dominant Chinese-English Code-Switching Environments**  
*Ruhlman Conference, Wellesley College* April 2022
- XiaoshuoNLP: A Natural Language Processing Pipeline for Chinese Literary Texts**

---

\*Equal contribution

*Ruhlman Conference Poster Session, Wellesley College* April 2022  
*Tanner Conference, Wellesley College* Nov 2022  
*Wellesley College Science Complex Opening Celebration Poster Session* Oct 2022  
*Wellesley College Science Center Summer Research Program Poster Fair* Aug 2022

**TEACHING EXPERIENCE** **CS 240: Foundations of Computer Systems** Spring 2022, Fall 2022  
 Teaching Assistant Wellesley College  
*Instructor: Andrew Davis, Franklyn Turbak*

**CS 204: Front-End Web Development** Term 2, Fall 2020  
 Teaching Assistant Wellesley College  
*Instructor: Scott Anderson*

**AWARDS & ACADEMIC HONORS** Sigma Xi May 2023  
 Phi Beta Kappa May 2023  
 Provost's Office Student Research Grant, Wellesley College Dec 2022

**CLASS PROJECTS** **Whisper Fine-tuning Event** Fall 2022  
*CS 333: Natural Language Processing*  
 Fine-tuned the pre-trained Automatic Speech Recognition model Whisper on the Common Voice dataset for Taiwanese Mandarin

**Syntactic Tense in Mandarin Chinese** Fall 2021  
*LING 244: Language: Form and Meaning*  
 Analyzed main issues of controversy regarding the existence of a syntactic T node in Mandarin Chinese and proposed a theory where tense is undefined in the deep structure

**Phonetics and Phonology of Korean** Spring 2021  
*LING 240: The Sounds of Language*  
 Designed and held recording sessions with a native Korean speaker to examine Korean phonetics and phonology, analyzed waveforms and spectrograms in Praat, and validated results against existing theories

**Social Variation of Words and Sounds in American Dialect** Fall 2020  
*LING 238: Sociolinguistics*  
 Evaluated social variation patterns of words and phonemes in North American English, identified dialect regions using isoglosses, and verified results against William Labov's first principle on language change

**INDUSTRY EXPERIENCE** **Shanghai Baosight Software Co., Ltd.** Jun 2021 - Aug 2021  
 Machine Learning Intern Shanghai, China  
 Analyzed and predicted steel defects using machine learning algorithms such as decision tree, clustering, and PCA

**LANGUAGES AND TECHNICAL SKILLS** **Programming Languages:** Python, R, C, Java, Racket, SML  
**Web Development:** HTML, CSS, JavaScript, React.js  
**Machine Learning:** TensorFlow, PyTorch  
**Software:** Praat, Audacity  
**Languages:** English and Mandarin Chinese (native proficiency), Spanish (working proficiency)