

MINGYUE HUO 霍明月 /mɪŋjʊə huə/

707 S. Matthews St., Urbana, IL, United States, 61801
mhuo5@illinois.edu | [LinkedIn](#) | [GitHub](#) | [Google Scholar](#)

EDUCATION

University of Illinois Urbana-Champaign

PhD in Computational Linguistics (GPA: 4.0/4.0) Advisor: Yan Tang

MS in Applied Statistics

MA in Japanese Linguistics

Urbana, IL

Dec 2026 (expected)

May 2026

May 2022

Shanghai International Studies University

BA in Japanese

Shanghai, China

June 2020

Hosei University

Exchange

Tokyo, Japan

Aug 2018 – Jan 2019

RESEARCH FOCUS & IMPACT

Multi-speaker speech understanding: end-to-end ASR, diarization, and target speaker extraction, with applications to real-world noisy and conversational settings.

Speech representation learning: building and evaluating speech encoders for speaker, paralinguistic, and audio-text alignment tasks. Extended to LLM-QA.

Research outputs: first-author papers at top speech venues (ICASSP, Interspeech, ASRU, JASA), with released code and public demos: [Diarization Demo](#), [Target Speaker Extraction Demo](#)

PUBLICATIONS

Where Speech Enhancement Hurts Recognition: An Inference Time Polar Projection Diagnosis

Mingyue Huo, Yuheng Zhang, Hao Zhang

In submission

Fooling Reward Models with Gibberish: Adversarial Examples Beyond the Semantic Regime

Yuheng Zhang*, Mingyue Huo*, Minghao Zhu, Mengxue Zhang, Nan Jiang (* Equal contribution)

In submission [\[PDF\]](#)

TagSpeech: End-to-End Multi-Speaker ASR and Diarization with Fine-Grained Temporal Grounding

Mingyue Huo, Yiwen Shao, Yuheng Zhang

ACL 2026 Main **Oral** [\[PDF\]](#) [\[Code\]](#) [\[Model\]](#)

Revisiting Audio-language Pretraining for Learning General-purpose Audio Representation

Wei-Cheng Tseng, Xuanru Zhou, Mingyue Huo, Yiwen Shao, Hao Zhang, Dong Yu

ACL 2026 Main [\[PDF\]](#)

Auden-Voice: General-Purpose Voice Encoder for Speech and Language Understanding

Mingyue Huo, Wei-Cheng Tseng, Yiwen Shao, Hao Zhang, Dong Yu

ICASSP 2026 [\[PDF\]](#) [\[Code\]](#) [\[Model\]](#)

Identifying and Calibrating Overconfidence in Noisy Speech Recognition

Mingyue Huo, Yuheng Zhang, Yan Tang

ASRU 2025 [\[PDF\]](#)

Beyond Speaker Identity: Text Guided Target Speech Extraction

Mingyue Huo, Abhinav Jain, Cong Phuoc Huynh, Fanjie Kong, Pichao Wang, Zhu Liu, Vimal Bhat

ICASSP 2025 [\[PDF\]](#) [\[Demo\]](#)

Iterative Nash Policy Optimization: Aligning LLMs with General Preferences via No-Regret Learning

Yuheng Zhang, Dian Yu, Baolin Peng, Linfeng Song, Ye Tian, Mingyue Huo, Nan Jiang, Haitao Mi, Dong Yu

ICLR 2025 **Oral** [\[PDF\]](#)

What makes listening comprehension difficult?: A feature-based machine learning approach to understanding item difficulty

Huiying Cai, Xun Yan, Ping-Lin Chuang, Yulin Pan, **Mingyue Huo**

Journal: Applied Linguistics, 2025 [\[PDF\]](#)

Release from Same-talker Speech-on-speech Masking: Effects of Masker Intelligibility and Other Contributing Factors

Mingyue Huo, Yinglun Sun, Dan Fogerty, Yan Tang

Journal: Journal of the Acoustical Society of America (JASA), 2024 [\[PDF\]](#)

Quantifying Informational Masking due to Masker Intelligibility in Same-talker Speech-in-speech Perception

Mingyue Huo, Yinglun Sun, Dan Fogerty, Yan Tang

INTERSPEECH 2023 **Oral** [\[PDF\]](#)

Towards Automating the Assessment of Teachers' Teaching Emotions in Preschool Classrooms

Mingyue Huo, Yan Tang

20th International Congress of Phonetic Sciences (ICPhS) 2023 [\[PDF\]](#)

INDUSTRY EXPERIENCE

ML/AI Scientist Intern at Netflix

Summer 2026, Los Gatos, CA

Host: Yinghong Lan

- Working on Speech / Audio LLM evaluations.

Research Scientist Intern at Tencent AI-Lab

Summer 2025, Bellevue, WA

Host: Dong Yu

- Developed speech encoders for speaker identity and paralinguistic research, addressing tasks such as speaker verification, diarization, counting, and emotion/age/gender classification, speech-text retrieval, and LLM-QA.
- Contributed to the open-source speech understanding framework Auden. Research paper accepted by ICASSP2026.

Applied Scientist Intern at Amazon Prime Video

Summer 2024, Seattle, WA

Host: Zhu Liu

- Developed a text-guided speech extraction model, integrating audio and text clues to enhance target speech extraction without traditional audio identifiers. Research paper accepted by ICASSP 2025.
- Created and evaluated the model using a newly extended dataset featuring speech mixture – text description pairs. Average improvement in SI-SDR is 16 dB.

Speech Data Intern at ByteDance AI-Lab

Dec 2019 – July 2021, Shanghai, China

- Worked on Mandarin and Japanese Text-to-Speech (TTS) and Singing Synthesis projects. Designed annotation pipelines, grapheme-to-phoneme (G2P) conversion rules, and created test sets with linguistic knowledge.
- Evaluated TTS models with expert knowledge, collected feedback from mean opinion score (MOS) and A/B testing, to refine synthesis naturalness and expressiveness. Several influential voices were launched on TikTok.
- Recorded studio-quality speech data, checked acoustics metrics, preprocessed with text normalization, audio segmentation, and denoising.

PRESENTATIONS

Aligning Speech-Hum Pairs using Dynamic Time Warping

Mingyue Huo, Yan Tang

Midwest Speech and Language Days (MSLD) 2024, University of Michigan, Ann Arbor

What makes listening comprehension difficult?: A feature-based machine learning approach to understanding item difficulty

Huiying Cai, Yulin Pan, Ping-lin Chuang, **Mingyue Huo**, Xun Yan

Language Testing Research Colloquium (LTRC) 2024, Innsbruck, Austria

Constructing a Phonetically-Balanced Low-Predictable Speech Corpus for Mandarin Chinese

Mingyue Huo, Yan Tang

28th Midcontinental Phonetics and Phonology Conference (MidPhon28), 2023, Purdue University

Using Praat Scripts for Data Processing and Annotation

Phonetics and Phonology Lab Workshop, 2022, UIUC

AWARDS

- UIUC Graduate College Dissertation Completion Fellowship for 2026-2027
- Conference Travel Award – UIUC Linguistics November 2025
- C.W. Kim Research Award in Linguistics – UIUC May 2024
- Linguistics Student Leadership Award – UIUC May 2024
- Midwest Speech and Language Days Travel Award – University of Michigan April 2024
- Departmental Conference Travel Award – UIUC November 2023
- Best Japanese TA – UIUC May 2022
- EALC Summer Research Fellowship – UIUC May 2021
- Teachers Ranked as Excellent by Their Students – UIUC 2021 Fall and 2020 Fall

TEACHING

- **Linguistics 402** TA, Tools and Tech Speech and Language Processing 2022 FA, 2024 FA
- **Linguistics 406** TA, Introduction to Computational Linguistics 2023 SP
- **Linguistics 446** TA, Fundamentals for Speech Signal Processing and Analysis 2023 SP, 2024 SP
- **Linguistics 448** TA, Introductory Machine Learning 2023 SP, 2024 SP
- **Japanese 201&202** Lecturer, Elementary Japanese I&II 2020 FA ~ 2022 SP

COMMUNITY SERVICE

- **Co-chair** at Illinois Language And Linguistics Society 16 (ILLS16) March 2024
- **Student Volunteer** at INTERSPEECH 2021, NeurIPS 2023
- **Reviewer** at ICASSP 2026, INTERSPEECH 2025 & 2026, SLT2026

SKILLS

- **Language:** Mandarin (Native), English (Advanced), Japanese (Advanced N1)
- **Software:** Praat, Audacity, Adobe Audition, Adobe Premiere, Montreal Forced Aligner
- **Programming:** Python, MATLAB, Shell, R, L^AT_EX
- **Framework & Library:** PyTorch, Kaldi, NumPy, SciPy, TorchAudio, librosa, soundfile, Matplotlib, Lhotse
- **Highlighted Coursework:** Digital Signal Processing (A+), Psychoacoustics (A), Phonetics (A+), Phonology (A), Formal Semantics (A+), Syntax (A), Machine Learning (A), Natural Language Processing (A), Generative Models for Audio (A)

Last Updated on June.27 2026