Jingyi Chen

614-329-1197 | chen.9220@osu.edu | Linkedin | GitHub | Website

Education

The Ohio State University Columbus, OH
Ph.D. in Computational Linguistics *Aug. 2019 – May 2026*
- Specialization: Speech Synthesis, Multimodal Large Language Models, Reinforcement Learning for Audio
- Advisors: Dr. Micha Elsner, Dr. Andrew Perrault
- Committee: Dr. Eric Fosler-Lussier, Dr. Cynthia Clopper

The Ohio State University Columbus, OH
M.S. in Computer Science & Engineering *Aug. 2022 – Aug. 2024*

Sichuan International Studies University Chongqing, China
B.A. in Linguistics *Aug. 2015 – Jul. 2019*

Selected Research Projects

Social-Emotional Speech Dialogue Benchmark for Multimodal LLMs Aug. 2025 – Present
- Developing benchmark evaluating how LLMs capture semantic meaning while processing emotional speech content
- Investigating speech-text cooperation mechanisms in multimodal models for emotion-content separation
- Analyzing instruction tuning effects on LLMs' ability to disentangle emotional cues from semantic information

Post-Training for Speech Emotion Conversion | Project Page Jan. 2025
- Designed end-to-end speech-to-speech emotion transfer model using advanced transformer architectures
- Proposed post-training framework pretraining on synthetic data and fine-tuning on real speech
- Released comprehensive dataset with 27 speakers, 100 text contents, 9 emotions, 27K audio samples

RL-Enhanced Diffusion TTS | Project Page May 2024
- Developed advanced RL techniques for diffusion-based speech synthesis improving naturalness and controllability
- Created novel reward-based loss function improving fine-tuning efficiency by 25%
- Conducted distributed training using PyTorch Lightning across multi-node GPU clusters
- Published at Interspeech 2025

Learning Speech Representations with GANs | Code Jan. 2023
- Trained convolutional GAN architectures on large-scale English and French word datasets to explore unsupervised speech representation learning
- Analyzed intermediate CNN layers to identify emergent linguistic features, advancing understanding of deep model speech representation
- Published at ACL 2023 with Area Chair Award

Industry Experience

Applied Scientist Intern Jun. 2025 – Aug. 2025
Amazon DEX AI *Bellevue, WA*
- Designed a recommendation system for low-consideration purchases leveraging LLM-based product representations
- Trained LLM based ranking models to personalize recommendations according to customer purchase preferences
- Implemented offline evaluation protocols, improving recommendation relevance, diversity, and coverage.
- Presented in Amazon Science Workshop.

Applied Scientist Intern May 2024 – Aug. 2024
Amazon Prime Video *Sunnyvale, CA*
- Developed a production-ready VAE-based speech-to-speech emotion transfer model with adversarial training
- Applied reinforcement learning for controlling emotion expression, speaker identity, and multilingual adaptation
- Created large-scale emotional speech dataset (27K+ samples) - 10x larger than existing public datasets
- Delivered research to production pipeline with 15% improvement in naturalness scores
- Implemented distributed training across 2-node (8 GPU) clusters with optimized data streaming buffers

Graduate Research Associate May 2019 – Present
The Ohio State University *Columbus, OH*
- Investigating advanced reinforcement learning techniques for diffusion-based speech synthesis models
- Developed novel reinforcement learning reward-based loss functions improving fine-tuning efficiency by 25%
- Applied GANs to examine sound representations in multilingual speech processing
- Published research at top-tier venues (ACL, Interspeech) with Area Chair Awards recognition

Selected Publications

Jingyi Chen, Ju-Seung Byun, Micha Elsner, Andrew Perrault. "Reinforcement Learning for Fine-tuning Text-to-speech Diffusion Models." *Interspeech 2025.*
Jingyi Chen, Micha Elsner. "Exploring How Generative Adversarial Networks Learn Phonological Representations." *ACL 2023.* **Area Chair Awards**.
Jingyi Chen, Pichao Wang, Andrew Perrault, Micha Elsner. "A Curriculum Learning Paradigm for Speech Emotion Transfer." *TTIC Speech & Audio Foundation Models Workshop 2025.*
Micha Elsner, **Jingyi Chen**, Andrea Sims. "Memory retrieval as pressure towards chunking in morphological inflection." *Computational Linguistics 2025.*

Presentation

Reinforcement Learning for Fine-tuning Text-to-speech Diffusion Models
Jingyi Chen
TTIC Speech & Audio Foundation Models Workshop 2025, Chicago, IL

Recommendation System for Purchases Leveraging LLM-Based Product Representations
Jingyi Chen
Amazon Science Workshop 2025, Bellevue, WA

Deep Language Learning: Modeling language from raw speech
Alan Zhou, **Jingyi Chen**, Sneha Ray Barman, Bruno Ferenc Segedin, Gašper Beguš
LSA Annual Meeting 2025, Philadelphia, PA | Tutorial Page

Awards & Professional Service

Awards: ACL 2023 Area Chair Awards; OSU Fellowship (2019-2020); OSU Research Awards (2021-2022)
Service: Reviewer for ICLR 2026, ICLR 2025, AAAI 2025, ACL 2025, Interspeech 2024

Technical Skills

Languages: Python, SQL, R, Praat
ML/DL Frameworks: PyTorch, PyTorch Lightning, HuggingFace Transformers
Speech AI: Text-to-Speech (TTS), Speech Synthesis, Emotion Speech Evaluation, Speech-LLMs
ML Expertise: Large Language Models, Reinforcement Learning (RLHF, PPO, DPO), Diffusion Models, GANs, VAEs
Infrastructure: Distributed Training (Multi-GPU/Multi-Node)