Xilin Jiang Ph.D. Candidate

xj2289@columbia.edu

INTERESTS

I build intelligent machines for **Listening** (speech enhancement 🔧), **Thinking** (audio understanding 🎧), and **Speaking** (speech synthesis 🗣), often with the help of other related modalities, including natural language (LLMs) ⑤, videos or images 📷, and human brains 🧠. *Please check* 🎧🔧🗣⑤📷🧠 *symbols in publications.*

SCHOOLS

<u>Columbia University</u>	New York, NY, US
Ph.D. Candidate after joint M.S. in Electrical Engineering	2022.09 - 2027 *(expected)*

- Advisor: Prof. Nima Mesgarani
- Research: Sound, Language, Brain, and their Interface
- GPA: 4.12/4.00

<u>University of Illinois Urbana-Champaign</u>	Urbana, IL, US
B.S. in Computer Engineering with the Bronze Tablet	2018.08 - 2021.12

- Advisor: Prof. Paris Smaragdis
- Research: Sound Separation
- Minor in mathematics
- GPA: 4.00/4.00

Selected Courses: *Speech & Audio Processing & Recognition, Machine Learning Frontier, Machine Learning for Signal Processing, Natural Language Processing, Computer Vision, Deep Learning, Algorithms & Models of Computation, Data Structures, Database Systems*

INTERNS

<u>Microsoft Research (MSR)</u>	Redmond, WA	2025.05 - 2025.08

Research Intern, Audio and Acoustics
Worked on spatial audio large language model based on Phi-4 Multimodal.

<u>Amazon</u>	Palo Alto, CA (remote)	2022.05 - 2022.07

SDE Intern, Search Science and AI
Developed search data quality metrics and analyzers with Spark and AWS.

<u>Amazon</u>	Palo Alto, CA (remote)	2021.05 - 2021.08

SDE Intern, Search Science and AI
Accelerated deep ranking model training with distributed data service on EC2 clusters.

<u>National Center for Supercomputing Applications (NCSA)</u>	Urbana, IL	2020.06 - 2020.07

Student Research Intern, GPU Computing for Bionanotechnology Simulation
Implemented cluster-based non-bonded particle interaction computation on GPU/CUDA.

SCHOOL SERVICES

<u>Teaching Assistant</u>	New York, NY	2025.01 - 2025.05

Speech&Audio Proc&Rec (ELENE6820)
Designed a programming assignment on spoken dialog system (cascaded and speech LM).

<u>Teaching Assistant</u>	New York, NY	2023.01 - 2023.05

Speech&Audio Proc&Rec (ELENE6820)
Designed a programming assignment on speech recognition and self-supervised learning.

<u>Lab Assistant</u>	Urbana, IL	2019.09 - 2019.12

Introduction to electronics (ECE110), Honors Section
Tutored students on programming, embedded devices, and machine learning.

PAPERS

Jiang, X. denotes equal contributions; † denotes mentored students.

Jiang, X., Wang, Q., Wu, J., He, X., Xu, Z., Ma, Y., Piao, M., Yang, K., Zheng, X., Shimizu, R., Chen, Y., Firoozi, A., Mischler, G., Dindar, S. S., Antonello, R., He, L., Hsieh, T.-A., Fan, X., Wu, Y., Ma, Y., Amballa, C., Chen, W., Hai, J., Li, R., Choudhari, V., Han, C., Li, Y. A., Flinker, A., Elhilali, M., Benetos, E., Hasegawa-Johnson, M. A., Roy Choudhury, R., and Mesgarani, N. (2026). "AVMeme Exam: A Multimodal Multilingual Multicultural Benchmark for LLMs' Contextual and Cultural Knowledge and Thinking", *(in submission)*, 🎧 ⑥ 📷

Dindar, S. S., **Jiang, X.**, Choudhari, V., Bickel, S., Mehta, A., Schevon, C., McKhann, G. M., Friedman, D., Flinker, A., and Mesgarani, N. (2025). "Speaker Identity is Robustly Encoded in Spatial Patterns of Intracranial EEG for Attention Decoding", *(NeuroImage in review)*, [pdf] 🎧 🧠

Shimizu, R., **Jiang, X.**, and Mesgarani, N. (2025). "MeanFlow-TSE: One-Step Generative Target Speaker Extraction with Mean Flow", *(in submission)*, [pdf] [code] 🔧

[J4] **Jiang, X.**, Gamper, H., Braun, S. (2025). "Sci-Phi: A Large Language Model Spatial Audio Descriptor", *IEEE Open Journal of Signal Processing (OJSP) ICASSP track*, [pdf] 🎧 ⑥

[C16] Wang, Q.†, **Jiang, X.**, He, L., Wu, J., and Mesgarani, N. (2025). "SightSound-R1: Cross-Modal Reasoning Distillation from Vision to Audio Language Models", *2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, [pdf] 🎧 ⑥ 📷

[C15] **Jiang, X.**, Wu, J., Choudhari, V., and Mesgarani, N. (2025). "Bridging Ears and Eyes: Analyzing Audio and Visual Large Language Models to Humans in Visible Sound Recognition and Reducing Their Sensory Gap via Cross-Modal Distillation", *2025 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, [pdf] 🎧 ⑥ 📷
BEST PAPER AWARD!

[C14] Li, Y. A., **Jiang, X.**, Tao, F., Niu, C., Xu, K., Song, J., and Mesgarani, N. (2025). "DMO-Speech 2: Reinforcement Learning for Duration Prediction in Metric-Optimized Speech Synthesis", *2026 The Association for the Advancement of Artificial Intelligence (AAAI)*, [pdf] [code] 📢

[C13] He, L., Wang, Q., **Jiang, X.**, and Mesgarani, N. (2025). "Layer-wise Minimal Pair Probing Reveals Contextual Grammatical-Conceptual Hierarchy in Speech Representations", *2025 Empirical Methods in Natural Language Processing (EMNLP)*, [pdf] 🎧 ⑥

[C12] **Jiang, X.**, Dindar, S. S., Choudhari, V., Bickel, S., Mehta, A., McKhann, G. M., Flinker, A., Friedman, D. and Mesgarani, N. (2025). "AAD-LLM: Neural Attention-Driven Auditory Scene Understanding", *2025 Annual Meeting of the Association for Computational Linguistics (ACL), main conference*, [pdf] 🎧 🔧 ⑥ 🧠

[C11] Xu, Z., Fan, X., Wang, Z. **Jiang, X.**, and Choudhury, R. R. (2025). "ArrayDPS: Unsupervised Blind Speech Separation with a Diffusion Prior", *2025 International Conference on Machine Learning (ICML)*, [pdf] 🔧

[C10] Florea, A., **Jiang, X.**, Mesgarani, N., and Jiang, X. (2025). "Exploring Finetuned Audio-LLM on Heart Murmur Features", *2025 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, [pdf] 🎧 ⑥

[C9] **Jiang, X.**, Li, Y. A., Florea, A. N., Han, C., and Mesgarani, N. (2024). "Speech slytherin: Examining the performance and efficiency of mamba for speech separation, recognition, and synthesis", *2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, [pdf] [code1] [code2] 🎧 🔧 📢

[C8] Li, Y. A., **Jiang, X.**, Han, C., and Mesgarani, N. (2024). "StyleTTS-ZS: Efficient High-Quality Zero-Shot Text-to-Speech Synthesis with Distilled Time-Varying Style Diffusion", *The 2025 Annual Conference of the Nations of the Americas Chapter of the ACL (NAACL)*, [pdf] [code] 📢

PAPERS

[C7] Wu, J., Fan, X., Lu, B. R., **Jiang, X.**, Mesgarani, N., Hasegawa-Johnson, M., and Ostendorf, M. (2024). "Just ASR+ LLM? A Study on Speech Large Language Models' Ability to Identify and Understand Speaker in Spoken Dialogue", *2024 IEEE Spoken Language Technology Workshop (SLT)*, [pdf] [code] 🎧 Ⓢ

[C6] Shams, S., Dindar, S. S., **Jiang, X.**, and Mesgarani, N. (2024). "Ssamba: Self-supervised audio representation learning with mamba state space model", *2024 IEEE Spoken Language Technology Workshop (SLT)*, [pdf] [code] 🎧

[C5] Li, Y. A., **Jiang, X.**, Darefsky, J., Zhu, G., and Mesgarani, N. (2024). "Style-Talker: Finetuning Audio Language Model and Style-Based Text-to-Speech Model for Fast Spoken Dialogue Generation", *2024 Conference on Language Modeling (CoLM)*, [pdf] [code] 🎧🎙️Ⓢ

[C4] **Jiang, X.**, Han, C., and Mesgarani, N. (2024). "Dual-path mamba: Short and long-term bidirectional selective structured state space models for speech separation", *2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, [pdf] [code] 🔧

[J3] **Jiang, X.**, Han, C., Li, Y. A., and Mesgarani, N. (2024). "Listen, Chat, and Remix: Text-Guided Soundscape Remixing for Enhanced Auditory Experience", *Special Issue on Deep Multimodal Speech Enhancement and Separation, IEEE Journal of Selected Topics in Signal Processing (JSTSP)*, [pdf] 🔧 Ⓢ

[C3] **Jiang, X.**, Han, C., Li, Y. A., and Mesgarani, N. (2023). "Exploring Self-supervised Contrastive Learning of Spatial Sound Event Representation", *2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, [pdf] 🎧

Li, Y. A., Han, C., **Jiang, X.**, and Mesgarani, N. (2023). "HiFTNet: A Fast High-Quality Neural Vocoder with Harmonic-plus-Noise Filter and Inverse Short Time Fourier Transform", *arXiv*, [pdf] [code] 🎙️

[C2] **Jiang, X.**, Li, Y. A., and Mesgarani, N. (2023). "DeCoR: Defy Knowledge Forgetting by Predicting Earlier Audio Codes", *2023 INTERSPEECH*, [pdf] 🎧

[C1] Li, Y. A., Han, C., **Jiang, X.**, and Mesgarani, N. (2022). "Phoneme-level bert for enhanced prosody of text-to-speech with grapheme predictions", *2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, [pdf] [code] 🎙️

[J2] Wang, Z., Subakan, C., **Jiang, X.**, Wu, J., Tzinis, E., Ravanelli, M., and Smaragdis, P. (2022). "Learning Representations for New Sound Classes With Continual Self-Supervised Learning", *IEEE Signal Processing Letters, 29, 2607-2611*, [pdf] [code] 🎧

[J1] Tzinis, E., Wang, Z., **Jiang, X.**, and Smaragdis, P. (2021). "Compute and memory efficient universal sound source separation", *Journal of Signal Processing Systems, 94(2), 245-259*, [pdf] [code] 🔧

[Thesis] **Jiang, X.** (2021). "Vector-quantized speech separation", *Undergraduate Thesis*, 🔧

SKILLS

Languages: English and Mandarin

Programming: Python, PyTorch, HuggingFace, SpeechBrain, CUDA, SQL

Others: Linux, Azure, AWS, Google Cloud, git, LaTeX

ACADEMIC SERVICES

Reviewers for
IEEE Signal Processing Letters
Association for the Advancement of Artificial Intelligence (AAAI)
IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)
IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)

HONORS	• Master of Science Award of Excellence top 5% GPA in master graduating class	2024
	• Wei Family Private Foundation Fellowship covering tuitions and stipends for two Ph.D. years	2022 - 2024
	• UIUC Bronze Tablet top 3% GPA in undergrad graduating class and defended an undergraduate research thesis	2021.12
	• Henry O. Koehler Merit Scholarship for an outstanding scholastic record	2021 - 2021
	• Frank C. Mock Scholarship awarded to top ECE students	2020 - 2021
	• Daniel W. and Carol A. Dobberpuhl Student Award for an outstanding scholastic record	2020 - 2021
	• James Scholar honor program in College of Engineering, UIUC	2019 - 2021