Linyang He

Website: https://linyanghe.github.io/ Email: linyang.he@columbia.edu

Education

Columbia University, New York, NY
PhD Student in Zuckerman Mind Brain Behavior Institute & Department of Electrical Engineering 2024 - Present
Advisor: Prof. Nima Mesgarani

University of Michigan, Ann Arbor, MI
M.Sc. in Electrical and Computer Engineering 2024

Fudan University, Shanghai, China
B.Sc. in Computer Science 2020

University of Chicago, Chicago, IL
Visiting Student in the Department of Linguistics 2019

Leadership & Service

President, Neurotech X Columbia Club, Columbia University 2025 - Present

- Leading Columbia's BCI student club, building projects and events that connect the neuroscience and AI communities.

Selected Publications

1. **Decoding Inner Speech with an End-to-End Brain-to-Text Neural Interface**
   Yizi Zhang*, **Linyang He***, Chaofei Fan, Tingkai Liu, Han Yu, Trung Le, Jingyuan Li, Scott Linderman, Lea Duncker, Francis R. Willett, Nima Mesgarani, Liam Paninski
   Submitted to *ICLR,* 2026 [link]
   Achieved No.1 performance in both the '24 and '25 Brain-to-Text challenges (BIT Team).

2. **Far from the Shallow: Brain-Predictive Reasoning Embedding through Residual Disentanglement**
   Linyang He*, Tianjun Zhong*, Richard Antonello, Gavin Mischler, Micah Goldblum, Nima Mesgarani
   NeurIPS (Neural Information Processing Systems), 2025 [link]

3. **Layer-wise Minimal Pair Probing Reveals Contextual Grammatical-Conceptual Hierarchy in Speech Representations**
   Linyang He*, Qiaolin Wang*, Xilin Jiang, Nima Mesgarani
   EMNLP (Empirical Methods in Natural Language Processing), 2025 [link]
   SAC Highlights Award (35 out of 8000+ submissions)

Research Interest

Key Words: NeuroAI, Brain-Computer Interface, AI Interpretability, Representation Learning, LLM Agent
Research Statement: I study the intersection of AI and neuroscience, using interpretable models to reveal cognitive mechanisms and to develop both invasive and noninvasive brain-computer interfaces for patient care and everyday use. I also design AI agents as tools for neuroscience, aiming to accelerate scientific discovery.

Research Experience

Columbia University, New York, NY
PhD Student Researcher, Zuckerman Mind Brain Behavior Institute 2024 - Present
Supervisor: Prof. Nima Mesgarani

- Perform research in: NeuroAI, Brain-Computer Interfaces, AI Agent for Neuroscience, AI Interpretability

Massachusetts General Hospital and **Harvard Medical School**, Boston, MA
Visiting Student Researcher, Department of Neurosurgery 2023 - 2024
Supervisor: Dr. Robert Mark Richardson

- Subcortical structures' role in language processing using deep language models in DBS patients

University of Michigan, Ann Arbor, MI
Research Assistant, Department of Linguistics and Department of Psychology 2023 - 2024
Supervisor: Prof. Jonathan Brennan and Prof. David Brang

- Linguistic evaluation of deep language models; neural mechanism of auditory-visual speech processing using deep learning models and ECoG

ShanghaiTech University, Shanghai .
Research Intern, School of Biomedical Engineering 2022 - 2023
Supervisor: Prof. Yuanning Li

- Neural substrates of semantics and phonetics using ECoG and deep language and speech models

ADDITIONAL PUBLICATIONS

Peer-Reviewed Conference Papers

4. **Large Language Models as Neurolinguistic Subjects: Discrepancy in Performance and Competence.**
   Linyang He, Ercong Nie, Helmut Schmid, Hinrich Schütze, Nima Mesgarani, Jonathan Brennan.
   Findings of ACL (Association for Computational Linguistics), 2025 [link]

5. **Do self-supervised speech and language models extract similar representations as human brain?**
   Peili Chen, **Linyang He**, Li Fu, Lu Fan, Edward F. Chang, Yuanning Li.
   ICASSP (IEEE International Conference on Acoustics, Speech, and Signal Processing), 2024 [link]

6. **Decoding Probing: Revealing Internal Linguistic Structures in Neural Language Models using Minimal Pairs**
   Linyang He, Peili Chen, Ercong Nie, Yuanning Li, Jonathan R. Brennan.
   COLING (International Conference on Computational Linguistics), 2024 [link]

Workshop Papers

7. **Neuron-Level Linguistic Selectivity in LLMs via Minimal Pairs: A Neuroscience-Inspired Framework**
   Linyang He, Nima Mesgarani
   UniReps Workshop at NeurIPS, *oral*, 2025

8. **XCOMPS: A Multilingual Benchmark of Conceptual Minimal Pairs**
   Linyang He*, Ercong Nie*, Sukru Samet Dindar, Arsalan Firoozi, Adrian Florea, Van Nguyen, Corentin Puffay, Riki Shimizu, Haotian Ye, Jonathan Brennan, Helmut Schmid, Hinrich Schütze, Nima Mesgarani
   SIGTYP Workshop at ACL, 2025 [link]

9. **High-Level features from deep language models predict subthalamic theta power during sentence processing.**
   Linyang He, Alan Bush, Latane Bullock, Yanming Zhu, Yuanning Li, Robert M. Richardson.
   OHBM (Organization for Human Brain Mapping), 2024 [link]

Preprints

10. **BabyBabelLM: A Multilingual Benchmark of Developmentally Plausible Training Data**
    Jaap Jumlet, …, **Linyang He**, …, Leshem Choshen
    Submitted to *ARR Oct Cycle,* 2025 [link]

11. **SightSound-R1: Cross-Modal Reasoning Distillation from Vision to Audio Language Models**
    Qiaolin Wang, Xilin Jiang, **Linyang He**, Junkai Wu, Nima Mesgarani
    Submitted to *ICASSP,* 2026 [link]

12. **Convergent representations and spatiotemporal dynamics of speech and language in brain and deep neural networks.**
    Peili Chen, Shiji Xiang, **Linyang He**, Edward F. Chang, Yuanning Li.
    bioRxiv, 2024 [link]

WORK IN PREPARATION

1. **AI Neuroscientist 1.0: Toward an AI Language Neuroscientist.**
   Linyang He, Nima Mesgarani, et al.

2. **Cortical-basal ganglia-cortical circuit in language processing: insights from deep language models.**
   Linyang He, Alan Bush, Latane Bullock, Yanming Zhu, Yuanning Li, Nima Mesgarani, Robert M. Richardson

3. **Neural Representation of LLM-Based Contextual Features Tracks Language Proficiency**
   Xiaomin He, **Linyang He**, Gavin Mischler, Nima Mesgarani

TALKS

Far from the Shallow: Brain-Predictive Reasoning Embedding through Residual Disentanglement
Computational Neurolinguistics Forum, CUHK, Hong Kong 2025

Human language decoding and encoding using deep language models.
Weinberg Institute for Cognitive Science, University of Michigan, Ann Arbor 2023

Subcortical structures' role in language processing: insights from deep language models.
Cortical-Basal Ganglia Speech Networks, Massachusetts General Hospital, Boston 2023

TEACHING EXPERIENCE

TA, Deep Learning and Neural Network, Columbia University, Professor: Micah Goldblum 2025
TA, Big Data Analytics, Columbia University, Professor: Ching-Yung Lin 2024

STUDENTS MENTORED

Tianjun Zhong, MS Student, Columbia University (2025 - Present)
Qiaolin Wang, MS Student, Columbia University (2025 - Present)

SKILLS

Programming: Python (PyTorch, scikit-learn, MNE, nibabel, nilearn, PySurfer, PsychoPy), MATLAB (FreeSurfer, FieldTrip, EEGLAB), LaTex, bash
Neuroscience: ECoG, EEG, MRI, DBS-LFP
Linguistics: Stanza, Tregrex, E-Prime, Treebank, Montreal Forced Aligner, Praat
Language: Mandarin, English