

Cheng-I Jeff Lai

☎ 424-376-6341 | ✉ clai24@mit.edu | 🏠 people.csail.mit.edu/clai24 | 📧 jefflai108 | 🎓 scholar

Research Highlights

My research interest is weakly-supervised and unsupervised learning for speech processing, with special focus in grounded language acquisition from speech, sparse speech processing via pruning, design and application of self-supervised representation learning in speech, and speaker-adaptive speech synthesis.

Education

Massachusetts Institute of Technology (MIT)

Ph.D. in Computer Science

Cambridge, MA
09/2019 - PRESENT

Johns Hopkins University (JHU)

B.S. in Electrical Engineering

Baltimore, MD
09/2015 - 12/2018

Research Experiences

PhD Student at MIT CSAIL Spoken Language System Group, Cambridge, MA, USA

09/2019 - PRESENT

Advisor: Dr. James Glass

Project: *Weakly-Supervised and Unsupervised Learning for Speech Processing*

- Grounded acquisition of grammar, words, phones from speech with distant supervision [submitted to ICLR 2023 & ICASSP 2023]
- Applications of self-supervised and weakly-supervised representation learning to speech and audio tasks, including SUPERB [Interspeech 2021], SUPERB-SG [ACL 2022], CMCM [ACL 2022], SSAST [AAAI 2022], Unsupervised-TTS [Interspeech 2022]

Research Scientist Intern at Meta Fundamental AI Research (FAIR), New York City, NY, USA

05/2022 - 08/2022

Hosts: Dr. Ann Lee

Project: *Word Discovery for Speech-to-Speech Translation*

- Investigated different grounding sources for word discovery, applied to direct speech-to-speech translation [SANE workshop]

Research Scientist Intern at MIT-IBM Watson AI Lab, Cambridge, MA, USA

05/2021 - 08/2021

Hosts: Dr. Yang Zhang

Project: *Low-Resource Language Learning*

- Initiated sparse speech processing for speech recognition and synthesis [NeurIPS 2021, ICASSP 2022, NeurIPS 2022]
- Designed disentanglement techniques for self-supervised speech representations [ICML 2022]

Applied Scientist Intern at Amazon AI, New York City, NY, USA

05/2020 - 08/2020

Hosts: Dr. Shang-Wen Li

Project: *Semi-Supervised Spoken Language Understanding*

- End-to-end spoken language understandings under limited and noisy data regimes. [ICASSP 2021, NeurIPS SAS workshop]

Research Intern at National Institute of Informatics, Tokyo, Japan

07/2019

Host: Prof. Junichi Yamagishi

Project: *Multi-Speaker Text-to-Speech*

- Integrated neural speaker encoders for multi-speaker (1) Tacotron2 [ICASSP 2020] (2) VQ-VAE [Interspeech 2020 B].
- TTS speaker adaptation with low-quality ASR corpora. [Interspeech 2020 A]

Undergraduate Researcher at Center for Language and Speech Processing, JHU, Baltimore, MD, USA

09/2016 - 04/2019

Advisors: Prof. Najim Dehak and Dr. Jesús Villalba

Project: *Speaker Recognition and its Applications*

- Introduced several deep learning models for the ASVspoof 2019 Challenge. [Interspeech 2019]
- Investigated unsupervised speech representations from raw waveforms for i-vector systems. [Undergraduate thesis]
- Proposed a neural approach to address bandwidth mismatches for i-vector and x-vector systems. [Interspeech 2018]
- Designed automatic speech biomarkers based on DNN-ivectors for Parkinson's disease detection.
- Applied LSTM-HMM to audio event classifications for the 2017 NIST OpenSAT evaluation. [SCALE 2017 workshop]

Research Intern at Informatics Forum, University of Edinburgh, Edinburgh, United Kingdom

06/2018 - 8/2018

Hosts: Prof. Simon King and Prof. Korin Richmond

Project: *Neural Approaches to Speech Anti-Spoofing*

- Proposed AFNet for audio replay detection with strong empirical and interpretable results. [ICASSP 2019]

Honors & Awards

2019-20 **Merrill Lynch Fellowship**, Department of Electrical Engineering and Computer Science, MIT

2019 **3rd Place (out of 50 teams)**, 2019 Automatic Speaker Verification Spoofing and Countermeasures Challenge

2018 **Departmental Honors, General Honors, Vredenburg Scholarship, Dean's List (All semesters)**, JHU

1. **Cheng-I Jeff Lai***, Freda Shi*, Puyuan Peng*, Yoon Kim, Kevin Gimpel, Shiyu Chang, Yung-Sung Chuang, Saurabhchand Bhati, David Cox, David Harwath, Yang Zhang, Karen Livescu, James Glass. "Textless Phrase Structure Induction From Visually-Grounded Speech," *submitted to ICLR 2023*
2. Yonggan Fu, Yang Zhang, Kaizhi Qian, Zhifan Ye, Zhongzhi Yu, **Cheng-I Jeff Lai**, Yingyan Lin. "Losses Can Be Blessings: Routing Self-Supervised Speech Representations Towards Efficient Multilingual and Multitask Speech Processing," *NeurIPS 2022*
3. Alexander H. Liu*, **Cheng-I Jeff Lai***, Wei-Ning Hsu, Michael Auli, Alexei Baevski, James Glass. "Simple and Effective Unsupervised Speech Synthesis," *Interspeech 2022*
4. Kaizhi Qian*, Yang Zhang*, Heting Gao, Junrui Ni, **Cheng-I Jeff Lai**, David Cox, Mark Hasegawa-Johnson, Shiyu Chang. "ContentVec: An Improved Self-Supervised Speech Representation by Disentangling Speakers," *ICML 2022*
5. Hsiang-Sheng Tsai, Heng-Jui Chang, Wen-Chin Huang, Zili Huang, Kushal Lakhotia, Shu-wen Yang, Shuyan Dong, Andy T. Liu, **Cheng-I Jeff Lai**, Jiatong Shi, Xuankai Chang, Phil Hall, Hsuan-Jui Chen, Shang-Wen Li, Shinji Watanabe, Abdelrahman Mohamed, Hung-yi Lee. "SUPERB-SG: Enhanced Speech processing Universal PERFORMANCE Benchmark for Semantic and Generative Capabilities," *ACL 2022*
6. **Cheng-I Jeff Lai**, Erica Cooper*, Yang Zhang*, Shiyu Chang, Kaizhi Qian, Yi-Lun Liao, Yung-Sung Chuang, Alexander H. Liu, Junichi Yamagishi, David Cox, James Glass. "On the Interplay between Sparsity, Naturalness, Intelligibility, and Prosody in Speech Synthesis," *ICASSP 2022*
7. Yuan Gong, **Cheng-I Jeff Lai**, Yu-An Chung, James Glass. "SSAST: Self-Supervised Audio Spectrogram Transformer," *AAAI 2022*
8. **Cheng-I Jeff Lai**, Yang Zhang*, Alexander H. Liu*, Shiyu Chang*, Kaizhi Qian, Yi-Lun Liao, Yung-Sung Chuang, Sameer Khurana, David Cox, James Glass. "PARP: Prune, Adjust and Re-Prune for Self-Supervised Speech Recognition," *NeurIPS 2021 (Spotlight)*
9. Alexander H. Liu, SouYoung Jin, **Cheng-I Jeff Lai**, Andrew Rouditchenko, Aude Oliva, James Glass. "Cross-Modal Discrete Representation Learning," *ACL 2022*
10. Shu-wen Yang, Po-Han Chi*, Yung-Sung Chuang*, **Cheng-I Jeff Lai***, Kushal Lakhotia*, Yist Y. Lin*, Andy T. Liu*, Jiatong Shi*, Xuankai Chang, Guan-Ting Lin, Tzu-Hsien Huang, Wei-Cheng Tseng, Ko-tik Lee, Da-Rong Liu, Zili Huang, Shuyan Dong, Shang-Wen Li, Shinji Watanabe, Abdelrahman Mohamed, Hung-Yi Lee. "SUPERB: Speech processing Universal PERFORMANCE Benchmark," *Interspeech 2021*
11. **Cheng-I Lai**, Yung-Sung Chuang, Hung-Yi Lee, Shang-Wen Li, James Glass. "Semi-Supervised Spoken Language Understanding via Self-Supervised Speech and Language Model Pretraining," *ICASSP 2021*
12. **Cheng-I Lai**, Jin Cao, Sravan Bodapati, Shang-Wen Li. "Towards Semi-Supervised Semantics Understanding from Speech," *NeurIPS 2020 workshop on Self-Supervised Learning for Speech and Audio Processing*
13. Fan-Keng Sun*, **Cheng-I Lai***. "Conditioned Natural Language Generation using only Unconditioned Language Model: An Exploration," *Technical Report*
14. Erica Cooper*, **Cheng-I Lai***, Yusuke Yasuda, Junichi Yamagishi. "Can Speaker Augmentation Improve Multi-Speaker End-to-End TTS?," *Interspeech 2020 A*
15. Yi Zhao, Haoyu Li, **Cheng-I Lai**, Jennifer Williams, Erica Cooper, Junichi Yamagishi. "Improved Prosody from Learned F0 Codebook Representations for VQ-VAE Speech Waveform Reconstruction," *Interspeech 2020 B*
16. Erica Cooper, **Cheng-I Lai**, Yusuke Yasuda, Fuming Fang, Xin Wang, Nanxin Chen, Junichi Yamagishi. "Zero-Shot Multi-Speaker Text-To-Speech with State-of-the-art Neural Speaker Embeddings," *ICASSP 2020 (Oral)*
17. **Cheng-I Lai**, Nanxin Chen, Jesús Villalba, Najim Dehak. "ASSERT: Anti-Spoofing with Squeeze-Excitation and Residual neTworks," *Interspeech 2019*
18. **Cheng-I Lai**. "Contrastive Predictive Coding Based Feature for Automatic Speaker Verification," *Undergraduate Thesis 2018*
19. Kelly Marchisio, Jialiang Guo, **Cheng-I Lai**, Philipp Koehn. "Controlling the Reading Level of Machine Translation Output," *MT Summit 2019*
20. **Cheng-I Lai**, Alberto Abad, Korin Richmond, Junichi Yamagishi, Najim Dehak, Simon King. "Attentive Filtering Network for Audio Replay Attacks Detection," *ICASSP 2019*
21. Phani Nidadavolu, **Cheng-I Lai**, Jesús Villalba, Najim Dehak. "Investigation on Bandwidth Extension for Speaker Recognition," *Interspeech 2018*

Talks

The SUPERB benchmark

- MIT 6.345 guest lecture

Making Machines Understand Uncommon Spoken Languages

- MIT ROCSA 5x5 talk, MIT Horizon

Finding Sparse Subnetworks for Self-Supervised Speech Recognition and Speech Synthesis

- MIT 6.345 guest lecture, Georgia Institute of Technology, A*STAR (Singapore), ASAPP, National Institute of Informatics (Japan), MIT Embodied Intelligence student seminar, MIT-IBM 5k language learning seminar

Semi-Supervised Trainings for Semantics Understanding from Speech

- National Institute of Informatics (Japan), NEC Corporation (Japan), Johns Hopkins University, Amazon Web Services Lex

Deep Learning Frameworks for Spoofing Detection and Speaker Representation

- National Institute of Informatics (Japan), NEC Corporation (Japan)

Deep Learning in Artificial Intelligence

- Nanhua University (Taiwan), China Medical University (Taiwan)

Attentive Filtering Network for Audio Replay Attacks Detection

- Rice University, Johns Hopkins University, University of Edinburgh (UK)