

WILLIAM CHEN

williamchen@cmu.edu ◊ wanchichen.github.io

EDUCATION

Carnegie Mellon University, Pittsburgh, Pennsylvania

PhD in Language Technologies

2024 - Present

M.S. in Language Technologies

2022 - 2024

Language Technologies Institute, School of Computer Science

GPA: 4.05

Advisor: Shinji Watanabe

University of Central Florida, Orlando, Florida

B.S. in Computer Science, Magna Cum Laude

2018 - 2021

B.A. in History, Cum Laude

2018 - 2021

Burnett Honors College

GPA: 3.89

RESEARCH EXPERIENCE

Carnegie Mellon University, Audio and Voice Lab

08.2022 - Present

Graduate Research Assistant || Advisor: Dr. Shinji Watanabe

- Large-scale AI foundation models for speech processing.

NTT Corporation, Communication Sciences Lab

05.2023 - 08.2023

Visiting Researcher || Advisors: Drs. Marc Delcroix, Takatomo Kano, Atsunori Ogawa

- Speech summarization and long-form speech recognition.

SELECTED PUBLICATIONS (REVERSE CHRONOLOGICAL ORDER)

- [1] **William Chen**, Wangyou Zhang, Yifan Peng, Xinjian Li, Jinchuan Tian, Jiatong Shi, Xuankai Chang, Soumi Maiti, Karen Livescu, Shinji Watanabe. “Towards Robust Speech Representation Learning for Thousands of Languages” *EMNLP Main Conference*, 2024. **Best Paper Award**.
- [2] **William Chen**, Brian Yan, Chih-Chen Chen, Shinji Watanabe. “FLORAS 50: A Massively Multilingual Multitask Benchmark for Long-form Conversational Speech” *SLT*, 2024.
- [3] Tejes Srivastava, Jiatong Shi, **William Chen**, Shinji Watanabe. “EFFUSE: Efficient self-supervised feature fusion for E2E ASR in multilingual and low resource scenarios.” *INTERSPEECH*, 2024. **Best Paper Award**.
- [4] Jiatong Shi, Shih-Heng Wang*, **William Chen***, Martijn Bartelds*, Vanya Bannihatti Kumar, Jinchuan Tian, Xuankai Chang, Dan Jurafsky, Karen Livescu, Hung-yi Lee, Shinji Watanabe. “ML-SUPERB 2.0: Benchmarking Multilingual Speech Models Across Modeling Constraints, Languages, and Datasets.” *INTERSPEECH*, 2024.
- [5] Yifan Peng, Jinchuan Tian, **William Chen**, Siddhant Arora, Brian Yan, Yui Sudo, Muhammad Shakeel, Kwanghee Choi, Jiatong Shi, Xuankai Chang, Jee-weon Jung, Shinji Watanabe. ‘OWSM v3.1: Better and Faster Open Whisper-style Speech Models based on E-branchformer’ *INTERSPEECH*, 2024.
- [6] Chih-Chen Chen*, **William Chen***, Rodolfo Zevallos, John E. Ortega. “Evaluating Self-Supervised Speech Representations for Indigenous American Languages.” *LREC-COLING*, 2024.
- [7] **William Chen**, Takatomo Kano, Atsunori Ogawa, Marc Delcroix, Shinji Watanabe. “Train Long and Test Long: Leveraging Full Document Contexts in Speech Processing.” *ICASSP*, 2024.

- [8] **William Chen**, Jiatong Shi, Brian Yan, Dan Berrebbi, Wangyou Zhang, Yifan Peng, Xuankai Chang, Soumi Maiti, Shinji Watanabe. “Joint Prediction and Denoising for Large-scale Multilingual Self-supervised Learning,” *ASRU*, 2023.
- [9] Yifan Peng, Jinchuan Tian, Brian Yan, Dan Berrebbi, Xuankai Chang, Xinjian Li, Jiatong Shi, Siddhant Arora, **William Chen**, Roshan Sharma, Wangyou Zhang, Yui Sudo, Muhammad Shakeel, Jee-weon Jung, Soumi Maiti, Shinji Watanabe. “Reproducing Whisper-Style Training Using an Open-Source Toolkit and Publicly Available Data,” *ASRU*, 2023.
- [10] **William Chen**, Xuankai Chang, Yifan Peng, Zhaoheng Ni, Soumi Maiti, Shinji Watanabe. “Reducing Barriers to Self-Supervised Learning: HuBERT Pre-training with Academic Compute,” *INTERSPEECH*, 2023.
- [11] Jiatong Shi, Dan Berrebbi, **William Chen**, Ho-Lam Chung, En-Pei Hu, Wei Ping Huang, Xuankai Chang, Shang-Wen Li, Abdelrahman Mohamed, Hung-yi Lee, Shinji Watanabe “ML-SUPERB: Multilingual Speech Universal PERformance Benchmark,” *INTERSPEECH*, 2023.
- [12] **William Chen**, Brian Yan, Jiatong Shi, Yifan Peng, Soumi Maiti et al.. “Improving massively multilingual asr with auxiliary ctc objectives,” *ICASSP*, 2023. **Top 3% Paper Award, IEEE SPS Student Travel Grant Award**

FUNDING, AWARDS AND HONORS

EMNLP Best Paper Award (2024)	Overall Best Paper Award
Interspeech Best Paper Award (2024)	Overall Best Paper Award
Monte Jade SE Innovation Competition (2024)	\$3000 1st place entrepreneurship award
Monte Jade SE Innovation Competition (2023)	\$5000 1st place entrepreneurship award
IEEE SPS Student Travel Grant (2023)	\$850 award for ICASSP 2023 [12]
ICASSP Top 3% Paper Award (2023)	Top paper award at ICASSP 2023 [12]
CMU LTI Research Fellowship (2023)	Full funding for master’s degree at CMU

SERVICE

Reviewer

- ICASSP (2024, 2025), ACL (2024), EMNLP (2024)
- LREC-COLING (2024), COLING (2025), IWSLT (2023, 2024)

Organizer

- IWSLT (2023, 2024), ASRU 2023 ML-SUPERB Challenge, Interspeech 2025 ML-SUPERB Challenge

TEACHING

18-781/11-751 Speech Recognition and Understanding	08.2024 - 12.2024
<i>Graduate Teaching Assistant Instructor: Prof. Shinji Watanabe</i>	

OTHER EXPERIENCE

Texas Instruments	06.2020 - 08.2020, 07.2021 - 08.2022
<i>Software Engineering Intern, Software Engineer</i>	
· Full-stack developer on E-commerce Team.	

SKILLS

Programming	Python, Java, Javascript, Typescript, React
Machine Learning	PyTorch, ESPnet, fairseq, HuggingFace