Liang-Hsuan Tseng (曾亮軒)

https://andybi7676.github.io/me
andybi7676@gmail.com
GitHub Profile
Google Scholar

About Me

I am a second-year Ph.D. student at the Speech Processing and Machine Learning (SPML) Laboratory of National Taiwan University, instructed by Prof. Hung-yi Lee. My research primarily focuses on speech processing and applied machine learning, with particular interests in automatic speech recognition and spoken language modeling. I have developed a subtitle generator system that automatically transcribes lecture recordings for the entire university, significantly improving accessibility in education. Recently, I have been working on spoken language modeling, with the aim of developing intelligent and high-fidelity speech generative models.

Education

- **National Taiwan University** *2022-present*
  Ph.D. in Graduate Institute of Electrical Engineering GPA: 4.14 / 4.30
- **National Taiwan University** *2018-2022*
  Bachelor of Electrical Engineering GPA: 3.82 / 4.30

Experience

- **MediaTek Research** *2024-2025*
  Deep Learning Intern Taipei, Taiwan
  – Propose TASTE: Text-Aligned Speech Tokenization and Embedding for Spoken Language Modeling *[ICLR'26]*
  – The first hybrid text-speech tokenization tailored for spoken language modeling.
- **Speech Processing and Machine Learning Laboratory, National Taiwan University** *2022-present*
  Ph.D. Student / Advisor: Prof. Hung-yi Lee Taipei, Taiwan
  – Derive SOTA unsupervised ASR framework based on reinforcement learning and iterative training. *[NeurIPS'24]*
  – Improve code-switching automatic speech recognition (ASR) with self-supervised speech models. *[AAAI SAS'22]*
- **Center of Teaching and Learning Development, National Taiwan University** *2021-present*
  Research and Development Intern Taipei, Taiwan
  – Develop and deploy the course subtitle generator for National Taiwan University.
  – Improve the ASR model of the subtitle generator. The new model is 2 times smaller, 5 times faster, and 20% better than the Whisper Large-v2 in code-switching ASR. *[SLT'24]*
  – Our system has transcribed over 80,000 lecture videos for the university faculty.

Publications

- **TASTE: Text-Aligned Speech Tokenization and Embedding for Spoken Language Modeling** *[ICLR'26]*
  Liang-Hsuan Tseng, Yi-Chang Chen, Kuan-Yi Lee, Da-Shan Shiu, Hung-yi Lee
- **REBORN: Reinforcement-Learned Boundary Segmentation with Iterative Training for Unsupervised ASR** *[NeurIPS'24]*
  Liang-Hsuan Tseng, En-Pei Hu, Cheng-Han Chiang, Yuan Tseng, Hung-yi Lee, Lin-shan Lee, Shao-Hua Sun
- **Leave No Knowledge Behind during Knowledge Distillation: Towards Practical and Effective Knowledge Distillation for Code-Switching ASR Using Realistic Data** *[SLT'24]*
  Liang-Hsuan Tseng, Zih-Ching Chen, Wei-Shun Chang, Cheng-Kuang Lee, Tsung-Ren Huang, Hung-yi Lee
- **Improving cascaded unsupervised speech translation with denoising back-translation** *[Preprint]*
  Yu-Kuan Fu, Liang-Hsuan Tseng (co-first), Jiatong Shi, Chen-An Li, Tsu-Yuan Hsu, Shinji Watanabe, Hung-yi Lee
- **Mandarin-English Code-Switching Speech Recognition with Self-Supervised Speech Representation Models** *[AAAI SAS'22]*
  Liang-Hsuan Tseng, Yu-Kuan Fu, Heng-Jui Chang, Hung-yi Lee
- **On the Fallacy of Global Token Perplexity in Spoken Language Model Evaluation** *[ACL Findings'26]*
  Jeff Chan-Jan Sju, Liang-Hsuan Tseng, Yi-Cheng Lin, Yen-Chun Kuo, Ju-Chieh Chou, Kai-Wei Chang, Hung-yi Lee, Carlos Busso

- **A Self-Refining Framework for Enhancing ASR Using TTS-Synthesized Data** *[ASRU'25]*
  Cheng-Kang Chou, Chan-Jan Hsu, Ho-Lam Chung, Liang-Hsuan Tseng, Hsi-Chun Cheng, Yu-Kuan Fu, Kuan Po Huang, Hung-Yi Lee

- **Introducing Semantics into Speech Encoders** *[ACL'23]*
  Derek Xu, Shuyan Dong, Changhan Wang, Suyoun Kim, Zhaojiang Lin, Akshat Shrivastava, Shang-Wen Li, Liang-Hsuan Tseng, Alexei Baevski, Guan-Ting Lin, Hung-yi Lee, Yizhou Sun, Wei Wang

- **Improving Generalizability of Distilled Self-Supervised Speech Processing Models under Distorted Settings** *[SLT'22]*
  Kuan-Po Huang, Yu-Kuan Fu, Tsu-Yuan Hsu, Fabian Ritter Gutierrez, Fan-Lin Wang, Liang-Hsuan Tseng, Yu Zhang, Hung-yi Lee

Academic Service

- **Reviewer:** ACL ARR (multiple times), NeurIPS 2025, TASLP 2025, ICLR 2026, ICML 2026
- **Emergency Reviewer:** ASRU 2025, ACML 2025, ICLR 2026, Interspeech 2026

Invited Talks

- **REBORN Unsupervised ASR System** *Winter, 2024*
  A 1 hour talk about our SOTA UASR system (NeurIPS 2024) Massachusetts Institute of Technology
  – Invited by Spoken Language Systems group at MIT CSAIL

Honors and Awards

- **NSTC Graduate Research Fellowship** *2025-2028*
  National Ph.D. fellowship NSTC, Taiwan
- **Outstanding Doctoral Talent Cultivation Program** *2026*
  Institutional Ph.D. fellowship GICE, National Taiwan University
- **NVIDIA Academic Grant Program** *2026*
  Computing Grant for Spoken Language Modeling Development NVIDIA
- **Science and Technology Research Innovation Award** *2024*
  Research award of our SOTA UASR system published in NeurIPS 2024 National Taiwan University
- **Conference Travel Grant from the Foundation for the Advancement of Outstanding Scholarship** *2024*
  Conference travel grant to NeurIPS 2024 FAOS, Taiwan

Teaching Experience

- **Machine Learning 2023** *Spring 2023*
  Head Teaching Assistant / Instructor: Prof. Hung-yi Lee National Taiwan University
  – The head TA of a course with over 500 students and 20 teaching assistants
  – Manage and examine the design and grading of all the 15 assignments

- **Deep Learning for Human Language Processing 2022** *Fall 2022*
  Teaching Assistant / Instructor: Prof. Hung-yi Lee National Taiwan University
  – Design and provide instructions for the assignment of unsupervised ASR

- **Machine Learning 2022** *Spring 2022*
  Teaching Assistant / Instructor: Prof. Hung-yi Lee National Taiwan University
  – Design and provide instructions for the assignment of model compression

Skills

- **Programming:** Python, JavaScript, C/C++, C#, LaTeX
- **Tools & Frameworks:** Pytorch, Docker, Node.js, React, PostgreSQL, RabbitMQ