

Kazuki Yamauchi

PH.D. STUDENT

Room #140, Engineering bldg. #6, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

✉ yamauchi-kazuki042@g.ecc.u-tokyo.ac.jp | 🌐 <https://kyamauchi1023.github.io/>

Memberships

- Member: #99938749 Institute of Electrical and Electronics Engineers (IEEE) Signal Processing Society (SPS)
Member: #202402114 Information Processing Society of Japan (IPSJ)
Member: #2323769 Institute of Electronics, Information and Communication Engineers (IEICE)
Member: #23750 Acoustical Society of Japan (ASJ)

Research Interests

Speech and language processing, speech synthesis, speech restoration, natural language processing, machine learning

Languages

Japanese (native) and English (conversant)

Education

Department of Information Physics and Computing, Graduate School of Information Science and Technology, The University of Tokyo

Japan

PH.D. DEGREE IN INFORMATION SCIENCE AND TECHNOLOGY

Apr. 2025 - Present

- Advisor: Prof. Yuki Saito
- Research topic: speech and language processing, speech synthesis, speech enhancement

Department of Creative Informatics, Graduate School of Information Science and Technology, The University of Tokyo

Japan

M.S. DEGREE IN INFORMATION SCIENCE AND TECHNOLOGY

Apr. 2023 - Mar. 2025

- Advisor: Prof. Hiroshi Saruwatari
- Research topic: speech and language processing, speech synthesis

Department of Mathematical Engineering and Information Physics, Faculty of Engineering, The University of Tokyo

Japan

B.S. DEGREE IN ENGINEERING

Apr. 2019 - Mar. 2023

- Advisor: Prof. Takeru Matsuda
- Research topic: signal processing, independent component analysis

Professional Experience

- Jul. 2025 - Present **Research internship**, CyberAgent, Inc. AI Lab, Japan. I work on research and development on speech restoration.
- Mar. 2023 - Present **Part-time researcher**, CoeFont Co.,Ltd., Japan. I work on research and development on expressive speech synthesis and voice conversion.
- Jul. 2024 - Sep. 2024 **Research internship**, Guardian Robot Project, RIKEN, Japan. I worked on research and development on conversational robots designed for multi-party speech conversations.
- Aug. 2023 - Sep. 2023 **Research internship**, NTT Human Informatics Laboratories, Japan. I conducted research on speaking style captioning.
- Sep. 2021 - Mar. 2024 **Part-time researcher**, Nagase Brothers Inc., Japan. I worked on research and development in natural language processing for hierarchical multi-label text classification.

Publications

INTERNATIONAL CONFERENCE PROCEEDINGS (PEER-REVIEWED, FIRST-AUTHORED ONLY)

Kazuki Yamauchi, Masato Murata, and Shogo Seki, “Confidence-based Filtering for Speech Dataset Curation with Generative Speech Enhancement Using Discrete Tokens,” Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (IEEE ICASSP), pp. xxxx—xxxx, Barcelona, Spain, May 2026. (Accepted)

Kazuki Yamauchi, Wataru Nakata, Yuki Saito, and Hiroshi Saruwatari, “Decoding Strategy with Perceptual Rating Prediction for Language Model-Based Text-to-Speech Synthesis,” Proc. Audio Imagination: NeurIPS 2024 Workshop on AI-Driven Speech, Music, and Sound Generation, Vancouver, Canada, Dec. 2024.

Kazuki Yamauchi, Yuki Saito, and Hiroshi Saruwatari, “Cross-Dialect Text-To-Speech in Pitch-Accent Language Incorporating Multi-Dialect Phoneme-Level BERT,” Proc. IEEE Spoken Language Technology Workshop (IEEE SLT), pp. 750-757, Macau, China, Dec. 2024.

Kazuki Yamauchi, Yusuke Iijima, and Yuki Saito, “StyleCap: Automatic Speaking-Style Captioning from Speech Based on Speech and Language Self-Supervised Learning Models,” Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (IEEE ICASSP), pp. 11261—11265, Seoul, South Korea, Apr. 2024.

INTERNATIONAL CONFERENCE PRESENTATIONS (ABSTRACT-REVIEWED, FIRST-AUTHORED ONLY)

Kazuki Yamauchi, Yuki Saito, and Hiroshi Saruwatari, “Speech Synthesis with Perceptual Rating-Guided Parallel Iterative Decoding,” ASA/ASJ Joint Fall 2025 Annual Meeting, No. 5pSC28, Honolulu, Hawaii, Dec. 2025.

Wataru Nakata*, **Kazuki Yamauchi***, Dong Yang, Hiroaki Hyodo, and Yuki Saito (*Equal contribution), “UTDUSS: UTokyo-SaruLab System for Interspeech2024 Speech Processing Using Discrete Speech Unit Challenge,” IEEE SLT 2024 Recent Breakthrough Results Session, No. RB-14, Macau, China, Dec. 2024.

INVITED ARTICLES

Yuki Saito, Wataru Nakata, **Kazuki Yamauchi**, and Joonyong Park, “Speech synthesis based on large pretrained models,” Journal of Acoustical Society of Japan, Vol. 81, No. 10, Oct. 2025.

Awards & Honours

Dec. 2024 **IEEE SPS Japan Student Conference Paper Award**, IEEE SPS Japan

Jun. 2024 **2024 Otogaku Symposium Best Presentation Award**, IPSJ SIG-SLP Committee

Mar. 2024 **Ranked 1st place in TTS (Acoustic+Vocoder) track**, Interspeech2024 Speech Processing Using Discrete Speech Unit Challenge

Mar. 2024 **The 28th Best Student Presentation Award**, ASJ

Mar. 2024 **2024 IEICE ISS Best Student Poster Award**, IEICE Technical Committee on Speech (SP)

Fellowships & Grants

Apr. 2025 - Mar. 2028 **Advanced AI Talent Development to Lead the Next-Generation Intelligent Society (BOOST NAIS)**, The University of Tokyo. 3,900,000 JPY / year (Approx. 26,900 USD / year)

Volunteer Works

Oct. 2024 - Present **ASJ Students and Young Researchers Forum**, Member.

Misc.

I was invited to Google Research Talk at Google Tokyo, Japan, in April 2024. I talked about our technical report for Interspeech2024 Speech Processing Using Discrete Speech Unit Challenge.