

Ryota Komatsu

✉ komatsu.r.ab@m.titech.ac.jp

🌐 ryota-komatsu.github.io

🎓 [Google Scholar](#)

🔗 [ryota-komatsu](#)

Interests

Speech-text large language models, reinforcement learning.

Education

Institute of Science Tokyo (formerly Tokyo Institute of Technology)

Ph.D. Student, Engineering

Apr 2025 – Mar 2028 (expected)

Specialization: Spoken language processing

Advisor: [Prof. Takahiro Shinozaki](#)

M.Eng.

Mar 2023

Specialization: Spoken language processing

Advisor: [Prof. Takahiro Shinozaki](#)

B.Eng.

Mar 2021

Specialization: Mathematical optimization

Advisor: [Prof. Isao Yamada](#)

Research Experience

- Identified speaker dominance as a key limitation of a previous syllabic tokenization method
- Proposed a linguistically faithful syllabic tokenization method for speech language models
- Built a data curation pipeline for real and synthetic speech (filtering for audio quality, disfluency)
- Finetuned Qwen2.5-7B on 130k hours of English speech across 32 NVIDIA H100 GPUs
- Outperformed Z.ai GLM-4-Voice by 5% in semantic understanding while using 42× less training compute

Publications

Speech language models

- [Speaker-Disentangled Chunk-Wise Regression for Syllabic Tokenization](#)
R. Komatsu, K. Kawakita, T. Okamoto, and T. Shinozaki
IEEE Open Journal of Signal Processing (OJSP), 2026
- [Self-Supervised Syllable Discovery Based on Speaker-Disentangled HuBERT](#)
R. Komatsu and T. Shinozaki
IEEE Spoken Language Technology Workshop (SLT), 2024

Spoken dialogue systems based on reinforcement learning

- [Automatic Spoken Language Acquisition Based on Observation and Dialogue](#)
R. Komatsu, S. Gao, W. Hou, M. Zhang, T. Tanaka, K. Toyoda, Y. Kimura, K. Hino, Y. Iwamoto, K. Mori, T. Okamoto, and T. Shinozaki
IEEE Journal of Selected Topics in Signal Processing (JSTSP), 2022
- [Continuous Action Space-Based Spoken Language Acquisition Agent Using Residual Sentence Embedding and Transformer Decoder](#)
R. Komatsu, Y. Kimura, T. Okamoto, and T. Shinozaki
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023

Signal processing

- [A Graph Regularized RPCA by Generalized Moreau Enhanced Model](#)

R. Komatsu, M. Yamagishi, and I. Yamada

European Signal Processing Conference (EUSIPCO), 2021

Invited Talks & Papers

[A Systematic Overview of Speech Language Models](#)

R. Komatsu

The Journal of the Acoustical Society of Japan, 2026

[Efficient Scaling of Speech Language Models Based on Syllable-Text Interleaving](#)

R. Komatsu

The Acoustical Society of Japan, 2026

[A Comprehensive Overview of Audio Language Models](#)

R. Komatsu

Sixth Joint Meeting Acoustical Society of America and Acoustical Society of Japan, 2025

Introduction to Multimodal Large Language Models

R. Komatsu

OTOGAKU Symposium, 2025

Work Experience

[Research & Development Group, Hitachi, Ltd.](#)

Apr 2023 – Jan 2025

Applied machine learning to semiconductor manufacturing process optimization in a production setting

Patents

[Search device, data compression processing device, semiconductor device manufacturing system, search method, and data compression processing method](#)

R. Komatsu, N. Hama, Y. Mori, M. Hamamoto, M. Egi, T. Ohmori, R. Asakura, and K. Matsuda

WO Patent WO2026009272A1

Technical Skills

Languages: Python, Bash

Environments: [Tsubame 4.0 supercomputer](#)

Libraries: PyTorch, Hugging Face ([transformers](#), [datasets](#), [trl](#)), [DeepSpeed](#)

Awards & Scholarships

Best Student Presentation Award, Acoustical Society of Japan, 2023

Science Tokyo Tsubame Scholarship for Doctoral Students, 2025–2028