

Joon-Yong Park

Ph.D. Student
The University of Tokyo
7-3-1 Hongō, Bunkyo City, Tokyo, Japan
joonyong-park@g.ecc.u-tokyo.ac.jp
[LinkedIn](#) [GitHub](#) [X\(Twitter\)](#)

Education

Apr. 2025 ~ Present	The University of Tokyo <i>Ph.D. Student</i> Department of Information Physics & Computing, Graduate School of Information Science and Technology. Research Theme: Towards integrated audio generation model based on linguistic analysis of speech token and automatic evaluation of acoustic signals <i>Advisor: Yuki Saito</i>	Tokyo, Japan
Apr. 2023 ~ Mar. 2025	<i>Master's Degree</i> Department of Electrical Engineering & Information Systems, Graduate School of Engineering. Thesis: A Study on Analysis of Discrete Speech Representation Using Self-Supervised Learning Model and its Application to Speech Synthesis <i>Advisor: Daisuke Saito</i> <i>GPA: 3.0 / 3.0</i>	
Mar. 2016 ~ Feb. 2022	Sogang University <i>Bachelor's Degree</i> Major in Artificial Intelligence, Department of Electronic Engineering, School of Engineering. Thesis: Development of an Interactive Voice User Interface in Kiosk System Using Deep Learning <i>Advisor: Hyung-Min Park</i> <i>GPA: 3.54 / 4.3</i>	Seoul, Korea

(Military Service Period: Nov. 2017 ~ Aug. 2019)

Research Interests

- Speech Synthesis / Voice Conversion
- Spoken Language Models / Speech Conversation
- Large-Language Models / Agentic AI
- Generative models (Autoregressive, Diffusion model)
- Self-Supervised Learning

Skills

Technologies

- Libraries NumPy, pandas, PyTorch, Kaldi etc.
- Programming Languages Python, C, MATLAB, shell, HTML/CSS
- Software Development Tools Git, Docker, Jupyter, VS Code etc.

Spoken Languages

Korean (Native), English, Japanese (Near Native), Chinese (Fully Professional)

Academic Publications

International Conferences (Peer-Reviewed)

1. **Joonyong Park**, Daisuke Saito, Nobuaki Minematsu, *MixedG2P-T5: G2P-free Speech Synthesis for Mixed-script texts using Speech Self-Supervised Learning and Language Model*, Proc. APSIPA ASC, 2025.
2. **Joonyong Park**, Shinnosuke Takamichi, David M. Chan, Shunsuke Kando, Yuki Saito, Hiroshi Saruwatari, *Analysing the Language of Neural Audio Codecs*, Proc. IEEE ASRU, 2025.
3. **Joonyong Park**, Kenichi Nakamura, *EmoSSLSphere: Multilingual Emotional Speech Synthesis with Spherical Vectors and Discrete Speech Tokens*, Proc. ICSA SSW, 2025.
4. David M. Chan, Rodolfo Corona, **Joonyong Park**, Cheol Jun Cho, Yutong Bai, Trevor Darrell, *Analyzing the Language of Visual Tokens*, Submitted to ACL AAR, 2025.
5. **Joonyong Park**, Daisuke Saito, Nobuaki Minematsu, *Analytic Study of Text-Free Speech Synthesis for Raw Audio using a Self-Supervised Learning Model*, Proc. APSIPA ASC, 2024.
6. Kentaro Onda, **Joonyong Park**, Nobuaki Minematsu, Daisuke Saito, *A Pilot Study of GSLM-based Simulation of Foreign Accentuation Only Using Native Speech Corpora*, Proc. ISCA INTERSPEECH, 2024.
7. Shinnosuke Takamichi, Hiroki Maeda, **Joonyong Park**, Daisuke Saito, Hiroshi Saruwatari, *Do learned speech symbols follow Zipf's law?*, Proc. IEEE ICASSP, 2024.
8. **Joonyong Park**, Shinnosuke Takamichi, Tomohiko Nakamura, Kentaro Seki, Detai Xin, Hiroshi Saruwatari, *How Generative Spoken Language Model Encodes Noisy Speech: Investigation from Phonetics to Syntactics*, Proc. ISCA INTERSPEECH, 2023.

Domestic Conferences (1st Author only)

1. 朴浚鎔、高道慎之介、David M. Chan、神藤駿介、齋藤佑樹、猿渡洋、“音声トークンの言語に関する分析”、2025年音学シンポジウム発表論文集、2025.
2. 朴浚鎔、中村建一、“言語別球面座標感情ベクトルと音声トークンを使用した多言語感情音声合成に関する考察”、日本音響学会2025年春季研究発表会講演論文集、2025.

3. 朴浚鎔、齋藤大輔、峯松信明、“自己教師あり学習モデルと T5 モデルを用いたG2Pフリー音声合成に関する研究” 研究報告音声言語情報処理、2024.
4. 朴浚鎔、齋藤大輔、峯松信明、“テキストのない音声に対する自己教師あり学習モデルによる音声合成の分析 ～多言語活用を中心に～”、研究報告音声言語情報処理、2024.
5. 朴浚鎔、齋藤大輔、峯松信明、“音声合成システムの入力表現に関する分析的検討”、日本音響学会2023年秋季研究発表会講演論文集、2023.
6. 朴浚鎔、高道慎之介、中村友彦、関健太郎、辛徳泰、猿渡洋、“Generative Spoken Language Modelを用いた劣化雑音音声の分析と他言語への適用”、日本音響学会2023年春季研究発表会講演論文集、2023.
7. Joon-Yong Park, Jaehyeon Ko, Seok Hyun Kim, Won Jeong, Jiho Lee, Bon Hyeok Ku, Hyung-Min Park, *Development of an Interactive Voice User Interface in Kiosk System Using Deep Learning*, Proc. Conference of the Brain Engineering Society of Korea, 2022.
8. 박준용, 이대호, 방운, 이나연, 김지환, “KorQuAD 데이터셋과 KoBERT Tokenizer를 이용한 KoBART 모델 성능의 분석”, 한국컴퓨터종합학술대회 논문집, 2021.

Honors and Scholarships

Apr. 2026 ~ Mar. 2028	Japan Society for the Promotion of Science Doctoral Course Research Fellowships (JSPS DC2, Accepted) 200,000 JPY/month + Research Funds
Jun. 2025	Best Paper Award Sound Symposiums 2025
Dec. 2024	Best Student Paper Award (LY Corporation Award) 154th Conference on Spoken Language Processing (SLP)
Nov. 2022	Honorable Mention Award ETRI OpenAPI Competition
Apr. 2022 ~ Present	Japanese Government (MEXT) Scholarship 147,000 JPY/month
May. 2021	Best Undergraduate Paper Award Korean Information Science Society

Research Experiences

Apr. 2025 ~ Present	LY Corporation <i>Industry-Academic Collaborative Research</i> Research on generating speech-to-face expression generation for virtual avatar.	Tokyo, Japan
Jun. 2024 ~ Jul. 2024	Darrell Group, University of California, Berkeley <i>Visiting Scholar</i> Research on analyzing the language of text-derived tokens from	California, USA

various modalities, including visual, auditory, and multimodal media.

Apr. 2022 ~ Mar. 2023	Saruwatari & Koyama Lab, The University of Tokyo <i>Research Student</i> Research on self-supervised learning and generative model on speech synthesis for post-filter signal processing used in speech enhancement.	Tokyo, Japan
Oct. 2021 ~ Mar. 2022	Intelligent Information Processing Lab, Sogang University <i>Internship</i> Development of dialogue system and speech synthesis in <i>IIPERIA</i> : Intelligent Information Processing with End-2-end Representation for Interactive Assistant.	Seoul, Korea
Jan. 2021 ~ Jun. 2021	Auditory Intelligent Lab, Sogang University <i>Internship</i> Development of <i>SGRecognizer</i> Demo: Korean unlimited vocabulary speech recognition system using Kaldi.	Seoul, Korea

Work Experiences

Oct. 2025 ~ Present	Spellbrush <i>Researcher (Part-time)</i> Research on speech generation in multimodal dialogue models and video generation models.	Tokyo, Japan
Oct. 2024 ~ Jun. 2025	Cygames Inc. <i>Researcher (Part-time)</i> Research on real-time foley sound synthesis in the virtual environment.	Tokyo, Japan
Aug.2024 ~ Oct. 2024	Preferred Networks Inc. <i>Researcher (Internship)</i> Research on multilingual emotional speech synthesis with spherical vectors and discrete speech tokens.	Tokyo, Japan
Jan.2024 ~ Feb.2024	NTT R&D <i>Researcher (Internship)</i> Analysis towards g2p-free speech synthesis using Text-Speech token encoder through Hidden-Unit T5.	Kanagawa, Japan
Aug.2023 ~ Oct. 2023	DeNA Co., Ltd. <i>AI Specialist (Internship)</i> Improvement and implementation of cross-lingual voice conversion for products.	Seoul, Korea
Apr. 2023 ~ Present	CoeFont Co., Ltd. <i>Researcher (Part-time)</i> Improvement and acceleration of the self-supervised learning framework to create natural and real-time speech synthesis and voice conversion.	Tokyo, Japan

Sep. 2020 ~ Jun. 2021	Samsung Medical Center <i>Research Assistant (Part-time)</i> Optimization of classification model and distillation of datasets for detecting diseases from cardiac MRI.	Seoul, Korea
--------------------------	---	-----------------

Other Projects & Experiences

Jan. 2025~ Jul. 2025	jarre.ai <i>Team Leader, Product Designer, AI Engineer</i> Student-based startup project, with theme of ‘Text-based AI implementation solution tool for video games’. Selected project for 2025 UTokyo Spring Founders Program.	Tokyo, Japan
Aug. 2024~ May. 2025	Rabbithole Company (Raccoon Party) <i>Director, Lead AI Engineer, Game Designer (Contract)</i> A startup for generative LLM-based AI character game, Project: maid*mirai. Design of original plans, LLM architectures, gameplay contents and scenario.	Seoul, Korea