

SANG HOON (TONY) Woo

+1 678-881-2663 tonywoo.me tonyswoo@gmail.com linkedin.com/in/tonyswoo
github.com/tonyswoo scholar.google.com/citations?user=kt3_CnoAAAAJ

Education

Georgia Institute of Technology

Ph.D. in Computer Science

- Advised by Professor Larry Heck.

Atlanta, GA

Aug. 2026 – Present

University of California, Davis

Bachelor of Science in Computer Science

Davis, CA

Sep. 2015 – Jun. 2019

Research

Seoul National University Vision & Learning Lab

Researcher

Seoul, South Korea

Jan. 2025 – Aug. 2026

- Conducted research on dialogue systems, focusing on real-time and spoken interaction.[1,2,3,6]
- Developed speech tokenization aligned with the corresponding text content for better LLM adaptation.[4,7]
- Designed a benchmark to assess low-level listening abilities of large audio-language models.[5]

UC Davis RUBiNet Lab

Undergraduate Research Student

Davis, CA

Jan. 2019 – Jun. 2019

- Contributed to machine learning powered clinical decision support system under Professor Chen-Nee Chuah.
- Developed a client-side application for iOS that monitors mechanical ventilator data and receives anomaly notifications.[13]

UC Davis DECAL Lab

Undergraduate Research Student

Davis, CA

Apr. 2018 – Mar. 2019

- Conducted experiments on integration of entropy scoring for genetic programming based automatic software repair program under the guidance of Professor Prem Devanbu.

Employment

SK Telecom

Research Intern

Seoul, South Korea

Oct. 2025 – Dec. 2025

- Selected to contribute to South Korea's flagship Sovereign AI Foundation Model Project, developing a large-scale national model initiative.
- Designed and implemented the spoken language interface for multimodal large language models, enabling more natural human-AI interaction.
- Built a pipeline for generating long-context QA training data in Korean and English to improve long-range reasoning.

Mindlogic Inc.

NLP Engineer

Seoul, South Korea

Sep. 2022 – Nov. 2023

- Initiated and developed a prototype of a dialogue system with persona-grounding based on historical chat data, enhancing user personalization.
- Established and utilized LLM-based modules, including prompt engineering techniques, to deliver high-quality business chatbot solutions.

MINDsLab Inc.

Senior AI Scientist

Seongnam, South Korea

Jan. 2022 – Sep. 2022

- Led the speech recognition research project, achieving a 68% relative error rate reduction compared to the previous internal model for Korean and English.
- Built and deployed a gRPC-based multilingual text-to-speech service compatible with the company's internal talking face generation system.[12]
- Developed an internal grapheme-to-phoneme process for Japanese and Chinese, enabling the company to expand its text-to-speech service to new markets.[11]

AI Scientist

Dec. 2020 – Jan. 2022

- Built a text preprocessing pipeline, including web scraping and text cleaning, for language model pre-training.
- Initiated and developed a transformer-based speech recognition service, overseeing model prototyping, training, and evaluation.

Projects

EnCLAP: Automated Audio Captioning Model

Researcher

Aug. 2023 – Oct. 2024

- Developed EnCLAP and EnCLAP++, state-of-the-art audio captioning models leveraging pretrained audio and language models, achieving top-tier performance in a competitive challenge.
- Co-authored a conference paper, technical report, and workshop paper documenting the models and improvements.[5,6,7]

Skills

Language: English (Native), Korean (Native), Japanese (Fluent)

Programming: Python, Java, Kotlin, Swift, SQL

Tools: Git, Docker

Frameworks/Libraries: PyTorch, TensorFlow/Keras, Hugging Face Transformers, gRPC, FastAPI, LangChain, vLLM

Publications

- [1] *SitCom: Scaling Egocentric Multi-Party Spoken Dialogue for Situated Communication Assistance*
Heeseung Yun, Sehun Lee, **Sang Hoon Woo**, Yoonji Nam, Sung-Feng Huang, Chao-Han Huck Yang, Gunhee Kim - *Manuscript under review at NeurIPS 2026*
- [2] *DExTER: Can Omnimodal Language Models Resolve Audio-Visual Deixis?*
Sehun Lee, Yoonji Nam, **Sang Hoon Woo**, Gunhee Kim - *Manuscript under review*
- [3] *DuplexGen: Adaptive Synthesis of Human-AI Turn-Taking Dialogues*
Takyoun Kim, Kang-wook Kim, **Sang Hoon Woo**, Julia Hirschberg, Gunhee Kim, Dilek Hakkani-Tür - *EMNLP 2026*
- [4] *StreamAlign: Streaming Text-Aligned Speech Tokenization*
Kang-wook Kim, Jinyoung Park, Jinsoo Kim, Sehun Lee, **Sang Hoon Woo**, Gunhee Kim - *EMNLP 2026 Findings*
- [5] *WoW-Bench: Evaluating Fine-Grained Acoustic Perception in Audio-Language Models via Marine Mammal Vocalizations*
Jaeyeon Kim, Heeseung Yun, **Sang Hoon Woo**, Chao-Han Huck Yang, Gunhee Kim - *ACL 2026 Findings*
- [6] *Think, Verbalize, then Speak: Bridging Complex Thoughts and Comprehensible Speech*
Sang Hoon Woo*, Sehun Lee*, Kang-wook Kim, Gunhee Kim - *EMNLP 2025*
- [7] *SubAlign: Speech Tokenization Aligned with LLM Vocabularies for Spoken Language Modeling*
Kang-wook Kim, Sehun Lee, **Sang Hoon Woo**, Gunhee Kim - *TTIC Summer Workshop on Foundations of Speech and Audio Foundation Models 2025*
- [8] *EnCLAP++: Analyzing the EnCLAP Framework for Optimizing Automated Audio Captioning Performance*
Jaeyeon Kim, Minjeong Jeon, Jaeyoon Jung, **Sang Hoon Woo**, Jinjoo Lee - *DCASE Workshop 2024*
- [9] *Expanding on EnCLAP with Auxiliary Retrieval Model for Automated Audio Captioning*
Jaeyeon Kim, Jaeyoon Jung, Minjeong Jeon, **Sang Hoon Woo**, Jinjoo Lee - *DCASE 2024 Task 6 (2nd Place)*
- [10] *EnCLAP: Combining Neural Audio Codec and Audio-Text Joint Embedding for Automated Audio Captioning*
Jaeyeon Kim, Jaeyoon Jung, Jinjoo Lee, **Sang Hoon Woo** - *ICASSP 2024*
- [11] *SANE-TTS: Stable And Natural End-to-End Multilingual Text-to-Speech*
Hyunjae Cho, Wonbin Jung, Junhyeok Lee, **Sang Hoon Woo** - *InterSpeech 2022*
- [12] *Talking Face Generation with Multilingual TTS*
Hyoung-Kyu Song*, **Sang Hoon Woo***, Junhyeok Lee, Seungmin Yang, Hyunjae Cho, Youseong Lee, Dongho Choi, Kang-wook Kim - *CVPR 2022 (Demo)*
- [13] *Leveraging IoTs and Machine Learning for Patient Diagnosis and Ventilation Management in the Intensive Care Unit*
Gregory B Rehm, **Sang Hoon Woo**, Xin Luigi Chen, Brooks T Kuhn, Irene Cortes-Puch, Nicholas R Anderson, Jason Y Adams, Chen-Nee Chuah - *IEEE Pervasive Computing 2020 (19.3)*
- [14] *Mining Vehicle Failure Consumer Reports for Enhanced Service Efficiency*
Ali Khodadadi, Chen-Nee Chuah, **Sang Hoon Woo**, Ashish Dalal - *IEEE VTC2019-Fall*