JIHWAN SEOL

seoljh0722@cau.ac.kr | LinkedIn | GitHub | Seoul, Korea

Education

Chung-Ang University Mar 2019 – Aug 2025
Bachelor of Science in School of Computer Science and Engineering (GPA: 3.87/4.5) *Seoul, Korea*

- TOEIC: 875 TOEIC Speaking: 150

Research Experience

Explainable Language Understanding (ELU) Lab, Chung-Ang University July 2023 – February 2026
Undergraduate Student Researcher *Seoul, Korea*

- Proposed **FairSLM**, a disentangled representation learning framework to isolate paraverbal acoustic attributes from semantic content in speech-aware language models, improving robustness to stylistic variation and associated social bias.
- Co-developed **VorTEX**, a target speech extraction (TSE) architecture designed to generalize across diverse overlap ratios and low-SNR acoustic conditions in multi-speaker environments.
- Established benchmark construction pipeline for **VoiceBBQ** to quantify acoustic representation leakage in Speech-LLMs via attribute-controlled minimal-pair evaluation, contributing to **EMNLP 2025 Main** publication.
- Constructed a contrastive probing dataset by synthesizing emotion- and speaker-varied utterances with controlled linguistic content for representation analysis in **Whisper** audio encoders.
- Formulated minimal-pair-based evaluation protocols to investigate paraverbal representation invariance in speech-language foundation models.
- Implemented Multi-style AdaIN (AdaMIN) module for prompt-conditioned style transfer in StyleTTS2, enabling disentangled control over speaker identity and prosodic variation.

Selected Projects

Korean Acting-Tone Emotional/Paraverbal Speech Corpus 2024 – 2025
Data Collection & Pipeline Engineering (Research Assistant)

- Constructed a 12-hour acted emotional speech corpus with controlled linguistic content to isolate paraverbal variation for training Speech-LLMs.
- Developed a synthetic augmentation strategy using TTS and VC models to mitigate emotional prosody imbalance, expanding the corpus to 72 hours for improved generative stability.
- Engineered a remote recording and validation pipeline to ensure phonetic consistency and acoustic integrity across diverse emotional expressions.
- Curated emotion-controlled utterance pairs for contrastive evaluation of prosodic sensitivity in downstream speech-language models.

AULO: Short-length Sound Effect Retrieval 2024
Undergraduate Capstone Project

- Developed a short-duration sound effect retrieval model using VAE-based acoustic encoders for robust feature matching.

AudioPoli: Auditory Emergency-Scene Classification 2023 – 2024
Google Solution Challenge (Global Top 100 Award Winner)

- Developed an environmental sound classification system to detect emergency situations in real-time.
- Implemented an end-to-end pipeline, directly contributing to the team's recognition among the top 100 out of 1,700+ global teams.

Publications (J: Journal, C: Conference, P: Preprint, *: Equal contribution)

[P1] **FAIRSLM: MITIGATING ACOUSTIC SOCIAL BIAS VIA DISENTANGLED SPEECH REPRESENTATIONS**
J. Seol, N. Kim, C. Cho, B. Kim *Manuscript under review (anonymous submission)*

[P2] **VORTEX: VARIOUS OVERLAP RATIO FOR TARGET SPEECH EXTRACTION**
R. Oh*, **J. Seol***, B. Kim *Manuscript under review (co-author; anonymous submission)*

[P3] **ACOUSTIC-BASED GENDER DIFFERENTIATION IN SPEECH-AWARE LANGUAGE MODELS**
J. Choi, **J. Seol**, N. Kim, C. Cho, E. Cho, B. Kim *Submitted to TACL 2026*

[C1] **VOICEBBQ: INVESTIGATING EFFECT OF CONTENT AND ACOUSTICS IN SOCIAL BIAS**
J. Choi, R. Oh, **J. Seol**, B. Kim *EMNLP 2025 Main*

Awards & Honors

Solution Challenge – Global Top 100 2024
Google Developers Groups *Global*
- Selected as Global Top 100 for developing AudioPoli, an auditory emergency-scene classification system.

Full-tuition Scholarship 2019 – 2025
Chung-Ang University *Seoul, South Korea*
- Awarded merit-based scholarship covering full tuition for the entire undergraduate period.

Research Interests

- **Representation Learning in Speech Foundation Models**: Understanding how paraverbal acoustic variation affects representation learning in speech-language models.
- **Disentangled Audio Representation**: Modeling attribute-invariant latent spaces to improve controllability and generalization under stylistic variation.
- **Representation Analysis**: Probing how emotional and speaker-related attributes are encoded across layers in large-scale speech encoders.
- **Robust Speech Modeling**: Designing architectures that maintain semantic consistency under overlap, noise, and prosodic variation.

Specialized Skills

Languages	Python, C/C++, Bash, CUDA, SQL
Frameworks	PyTorch, TensorFlow, Keras, PyTorch Lightning, MongoDB
Toolkits	HuggingFace, Librosa, Amphion, OpenAI Whisper, ESPnet
Infra	Git, Docker, Google Cloud Platform (GCP), Linux, LaTeX