

Linkai Peng

✉ penglinkai96@gmail.com | ✉ penglinkai@corp.netease.com
🏠 <https://vocaliodmiku.github.io> | 🌐 <https://github.com/vocaliodmiku>
🎓 [Google Scholar](#) | 📄 [Research Gate](#)

Education

Beijing Language and Culture University

Sep. 2019 - Jun. 2022

M.S. in Software Engineering

Speech Acquisition and Intelligent Technology (SAIT) Lab

🎓 *Graduated with Outstanding Graduation Dissertation Award, supervised by Prof. Jinsong Zhang.*

Relevant Courses: Speech Information Processing, Prosodic Information Processing,
Academic lectures on Language Intelligence Technology,
Fundamental of Man-Machine Speech Communication.

South China Normal University

Sep. 2015 - Jun. 2019

B.S. in Physics

Experience

NetEase

Jul. 2022 - Present

Machine Learning Engineer in Youdao Group

Beijing, China

- Conducted human-centered speech technology research and applications, leveraging linguistics/phonetics, Generative AI and deep learning to intellectualize language learning.
- Designed and improved speech assessment systems, including real-time systems and offline systems optimized for edge devices.

ByteDance

Nov. 2020 - Feb. 2021

Speech Research intern in AI Lab

Beijing, China

- Worked in AI lab to optimize an education product using machine learning techniques, enhancing customer experience for millions of online users.
- Investigated and improved the algorithms for automatic scoring of pronunciation quality, adapting them to our specific scenario for potential performance gain.

Projects

Investigate the alignment of speech processing mechanisms between deep learning models and the human brain. [Link]

Nov. 2023 - Ongoing

- Conducted an in-depth investigation into the alignment of speech processing mechanisms between deep learning models and the human brain.
- Reconstructed and refined key research methods, including the quantitative comparison of self-supervised representations with behavioral data, and the alignment of hierarchical representations of deep neural networks with the human auditory pathway using invasive and noninvasive brain recordings.
- Explored the similarities in speech processing mechanisms between robust multilingual acoustic models (such as Whisper, SeamlessM4T, etc.) and human perception and adaptation of native/non-native, familiar/unfamiliar speakers' speech. The primary focus was on capturing subtle acoustic differences to construct potential computational models.

Recover speech/semantic information from brain recordings. [Link]

Nov. 2023 - Ongoing

- Investigated end-to-end modeling techniques for decoding brain recordings, with a special emphasis on the significant role of data processing in achieving success.
- Integrated pre-trained speech and semantic models like Wav2vec2 and GPT3.5/4 to develop a robust framework that accommodates multiple subjects.

Papers

- [1] **Peng, Linkai**, Baorian Nuchged, and Yingming Gao. "Spoken Language Intelligence of Large Language Models for Language Learning." arXiv preprint arXiv:2308.14536 (2023).
- [2] R. Bao, **L. Peng**, Y. Gao and J. Zhang, "The Contribution of Phonological and Fluency Factors to Chinese L2 Comprehensibility Ratings: A Case Study of Urdu-speaking Learners," 2022 13th International Symposium on Chinese Spoken Language Processing (ISCSLP), Singapore, Singapore, 2022, pp. 394-398, doi: 10.1109/ISCSLP57327.2022.10038233.
- [3] R. Bao, **L. Peng**, Y. Yan and J. Zhang, "An Exploratory Study for Quantifying the Contextual Information for Successful Chinese L2 Speech Comprehension," 2022 13th International Symposium on Chinese Spoken Language Processing (ISCSLP), Singapore, Singapore, 2022, pp. 389-393, doi: 10.1109/ISCSLP57327.2022.10038037.
- [4] R. Bao, **L. Peng**, Y. Gao and J. Zhang, "The Importance of Lexical Tone for Sentence Understanding: Utilizing Functional Load Principle to Simulate Comprehension Process," 2022 International Conference on Asian Language Processing (IALP), Singapore, Singapore, 2022, pp. 379-383, doi: 10.1109/IALP57159.2022.9961243.
- [5] R. Bao, **L. Peng** and J. Zhang, "The Contributions of Initials and Finals to L2 Chinese Comprehensibility Based on Functional Load Principle," 2022 International Conference on Asian Language Processing (IALP), Singapore, Singapore, 2022, pp. 358-362, doi: 10.1109/IALP57159.2022.9961277.
- [6] **Peng, Linkai**, et al. "Text-Aware End-to-end Mispronunciation Detection and Diagnosis." arXiv preprint arXiv:2206.07289 (2022).
- [7] **L. Peng**, W. Dai, D. Ke and J. Zhang, "Multi-Scale Model for Mandarin Tone Recognition," 2021 12th International Symposium on Chinese Spoken Language Processing (ISCSLP), Hong Kong, 2021, pp. 1-5, doi: 10.1109/ISCSLP49672.2021.9362063.
- [8] **Peng, L.**, Fu, K., Lin, B., Ke, D., Zhan, J. (2021) A Study on Fine-Tuning wav2vec2.0 Model for the Task of Mispronunciation Detection and Diagnosis. Proc. Interspeech 2021, 4448-4452, doi: 10.21437/Interspeech.2021-1344

Skills

Languages:

C++, Python, HTML, JavaScript, Bash

Technologies & Tools:

Git, Linux, Docker