# Kai-Wei Chang's Research Summary

Prompting — a technique for instructing language models (LM) — has shown promise in guiding models for various NLP tasks, particularly gaining attention after the introduction of ChatGPT. Prior to this, in 2022, Kai-Wei Chang has been developing **SpeechPrompt** [2], the first prompting framework designed for speech LMs in speech processing.

SpeechPrompt [2] employs minimal prompt vectors as input to efficiently guide a speech LM (with fixed architecture and parameters) for various tasks, offering computation, storage, and human-labor efficiency. This groundwork further led to research broadening SpeechPrompt's applications across a wide range of speech classification tasks [3] and speech generation tasks [4]. A unifying journal paper [1] extending the works [2][3][4] has been published in TASLP.

Building on this research, Kai-Wei Chang further investigated two crucial aspects of prompting in speech processing: model architecture [5] and in-context learning [6]. The work [5] revealed that prompting an encoder-decoder speech LM is more effective than a decoder-only model, a finding that distinctly differs from typical practices in NLP. Additionally, the pioneering work [6] studied an important prompting method, in-context learning, which provides data demonstrations for speech LMs. This work marks the first exploration in this field.

Focusing on speech LM and prompting in speech processing, Kai-Wei Chang has delivered **tutorial speeches** on overview of speech LMs at ICASSP 2023 and ICASSP 2024, with each session extending beyond 30 minutes.

Looking to the future, as LLMs like OpenAI's GPT-4o and Google's Gemini increasingly focus on speech modalities, the deployment of speech LMs in the cloud is becoming a prominent trend. In this context, SpeechPrompt stands out as a scalable solution, offering the potential to serve a large user base.

# Links

1. Kai-Wei Chang's Website: https://kwchang.org/

2. Kai-Wei Chang's Talks: https://kwchang.org/talks/

3. SpeechPrompt Project Website: https://kwchang.org/SpeechPrompt/

# Selected Publications

**[1] SpeechPrompt: Prompting Speech Language Models for Speech Processing Tasks (IEEE/ACM TASLP vol 32, 2024)**
***Kai-Wei Chang**, Wei-Cheng Tseng, Shang-Wen Li, Hung-yi Lee*
➔ Extension of the works [2], [3], and [4]. A unified framework for speech classification, sequence generation, and speech generation tasks.

**[2] SpeechPrompt: An Exploration of Prompt Tuning on Generative Spoken Language Model for Speech Processing Tasks (Interspeech 2022)**
***Kai-Wei Chang**, Wei-Cheng Tseng, Shang-Wen Li, Hung-yi Lee*
➔ Introduced SpeechPrompt, the first work exploring the prompting paradigm in speech processing.

**[3] SpeechPrompt v2: Prompt Tuning for Speech Classification Tasks (arXiv Preprint)**
***Kai-Wei Chang**, Yu-Kai Wang, Hua Shen, Iu-thing Kang, Wei-Cheng Tseng, Shang-Wen Li, Hung-yi Lee*
➔ Applied SpeechPrompt to various speech classification tasks, covering multiple languages and prosody-related tasks.

**[4] SpeechGen: Unlocking the Generative Power of Speech Language Models with Prompts (arXiv Preprint)**
***Kai-Wei Chang**, Haibin Wu, Yuan-Kuei Wu, Hung-yi Lee*
➔ Explored SpeechPrompt on advanced speech language models for speech generation tasks.

**[5] Prompting and Adapter Tuning for Self-supervised Encoder-Decoder Speech Model (ASRU 2023)**
***Kai-Wei Chang**, Ming-Hsin Chen, Yun-Ping Lin, Jing Neng Hsu, Chien-yu Huang, Shang-Wen Li, Hung-Yi Lee*
➔ Studied parameter-efficient tuning methods for transfer learning in cross-lingual speech recognition.

**[6] Exploring In-Context Learning of Textless Speech Language Model for Speech Classification Tasks (Interspeech 2024)**
***Kai-Wei Chang**, Ming-Hao Hsu, Shang-Wen Li, Hung-yi Lee*
➔ The first work studying in-context learning for speech language models