

# Chung-Yi Li

✉ leon129506@gmail.com    ☎ (+886)0921-083-425

## EDUCATION

**National Taiwan University**  
*M.S. in Electrical Engineering*

**Taipei**  
*Sep. 2018 - Jan. 2021*

**National Taiwan University**  
*B.S. in Electrical Engineering*

**Taipei**  
*Sep. 2014 - Jun. 2018*

## EXPERIENCES

**1st Place in Formosa Speech Recognition Challenge 2020 - Taiwanese ASR**

*Jan. 2021*

- Built a Taiwanese Southern Min (TSM) ASR system using Kaldi  
Ranked 1st among 12 teams in the TSM ASR track
- Built a TSM-to-Mandarin speech-to-text translation system using Kaldi and Moses  
Ranked 1st among 8 teams in the TSM-to-Mandarin speech-to-text translation track

**Speech Recognition Engineer Intern at Taiwan AI Labs**

*Jul. 2019 - Sep. 2019*

- Built a TSM-to-Mandarin speech-to-text translation system using Allennlp
- Crawled 1,500 hrs noisy parallel taiwanese speech and mandarin subtitles from Youtube
- Experimented with two popular end-to-end neural architectures: LAS(Listen, attend, and spell), RNN Transducer
- Achieved 45% CER and 40.7 BLEU on in-house test set

**M.S. Student at National Taiwan University**

*Jan. 2021*

- Probe the information in black-box end-to-end ASR systems through "audifying" the hidden states of end-to-end ASR. Probing models were asked to recover the input speech from hidden layers of end-to-end ASR. We observe gradual removal of speaker variability and noise from the recovered speech as the layer goes deeper, which aligns with the previous studies on how deep network functions in ASR [3].
- Explore model-agnostic meta-learning methods on low-resource dependency parsing. Improve low-resource dependency parsing performance over standard multilingual pre-training baselines [2].

**4th Place at Lawsnote LegalTech Hackathon**

*Sep. 2019*

- Built a Chinese event graph extraction system tailored for legal documents leveraging classic NLP pipelines

## PUBLICATIONS

- [1] Chi-Liang Liu, Tsung-Yuan Hsu, Yung-Sung Chuang, **Chung-Yi Li**, and Hung-yi Lee. **Language Representation in Multilingual BERT and its applications to improve Cross-lingual Generalization.** *arXiv preprint arXiv:2010.10041*, 2020
- [2] **Chung-Yi Li**. **Meta-Learning for Low-resource Dependency Parsing.** Master's thesis, National Taiwan University, Oct 2020
- [3] **Chung-Yi Li\***, Pei-Chieh Yuan\*, and Hung-Yi Lee. **What does a network layer hear? Analyzing hidden representations of end-to-end ASR through speech synthesis.** In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6434–6438. IEEE, 2020

## SKILLS

**Programming**  
**Frameworks and tools**  
**Languages**  
**Languages**

Python, shell script, Javascript  
Pytorch, Kaldi, Allennlp, Git, L<sup>A</sup>T<sub>E</sub>X  
Taiwanese Mandarin (Native), Taiwanese Southern Min (Native)  
English (TOEFL-iBT: R29/L28/S23/W27)