

Kaggle inClass Competition

- Taiwanese Speech Recognition
- Self-Supervised Pre-Trained Model + Transformer

Yuan-Fu Liao

National Yang Ming Chiao Tung University

yfliao@nycu.edu.tw



NYCU-IASS-DL2024-Taiwanese-ASR

Taiwanese Speech Recognition using ESPnet Toolkit and Self-Supervised Pre-Trained Model

ESPnet: end-to-end speech processing toolkit



ESPnet is an end-to-end speech processing toolkit, mainly focuses on end-to-end speech

- Kaggle
 - <https://www.kaggle.com/t/95d247202d12494c8278ce7b05fd62c7>
- GitHub Classroom
 - TBD
- Task
 - **Transformer**-based Speech Recognition
- Toolkits
 - ESPnet: end-to-end speech processing toolkit
 - [espnet/egs/librispeech at master · espnet/espnet \(github.com\)](https://github.com/espnet/egs/librispeech)
 - [espnet/egs2/aishell/asr1 at master · espnet/espnet · GitHub](https://github.com/espnet/egs2/aishell/asr1)

The Future Has Arrived!

- 會看，會聽，會說的機器人
- OpenAI Speech-to-Speech Reasoning
- NVIDIA Project GR00T Foundation Model
- Disney Robot



<https://youtu.be/51TYhPJ4zys?si=kHLeMnBPAHmeK04r>

<https://youtu.be/Sq1QZB5baNw?si=HeUjfJLTxe8M9ZHI>



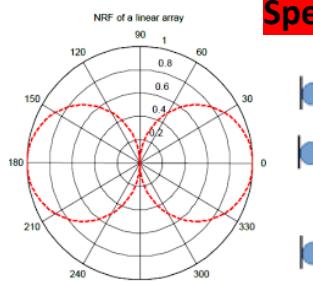
<https://youtu.be/qNfRgZMWyCI?si=c0tu8X9wKqf8TM9X>

Traditional Spoken Dialogue System

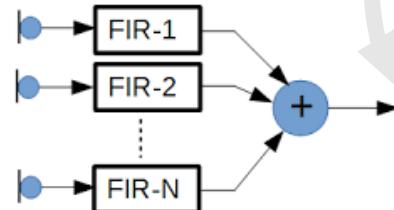
Speech/Speaker/Language Recognition

Speech Synthesis

Speech Enhancement



Speech Enhancement



Speech
Speaker
Language
Emotion
Recognition

Speech
Synthesis

Audio, KB,
Mouse, ...

Language
Generation

Language
Understanding



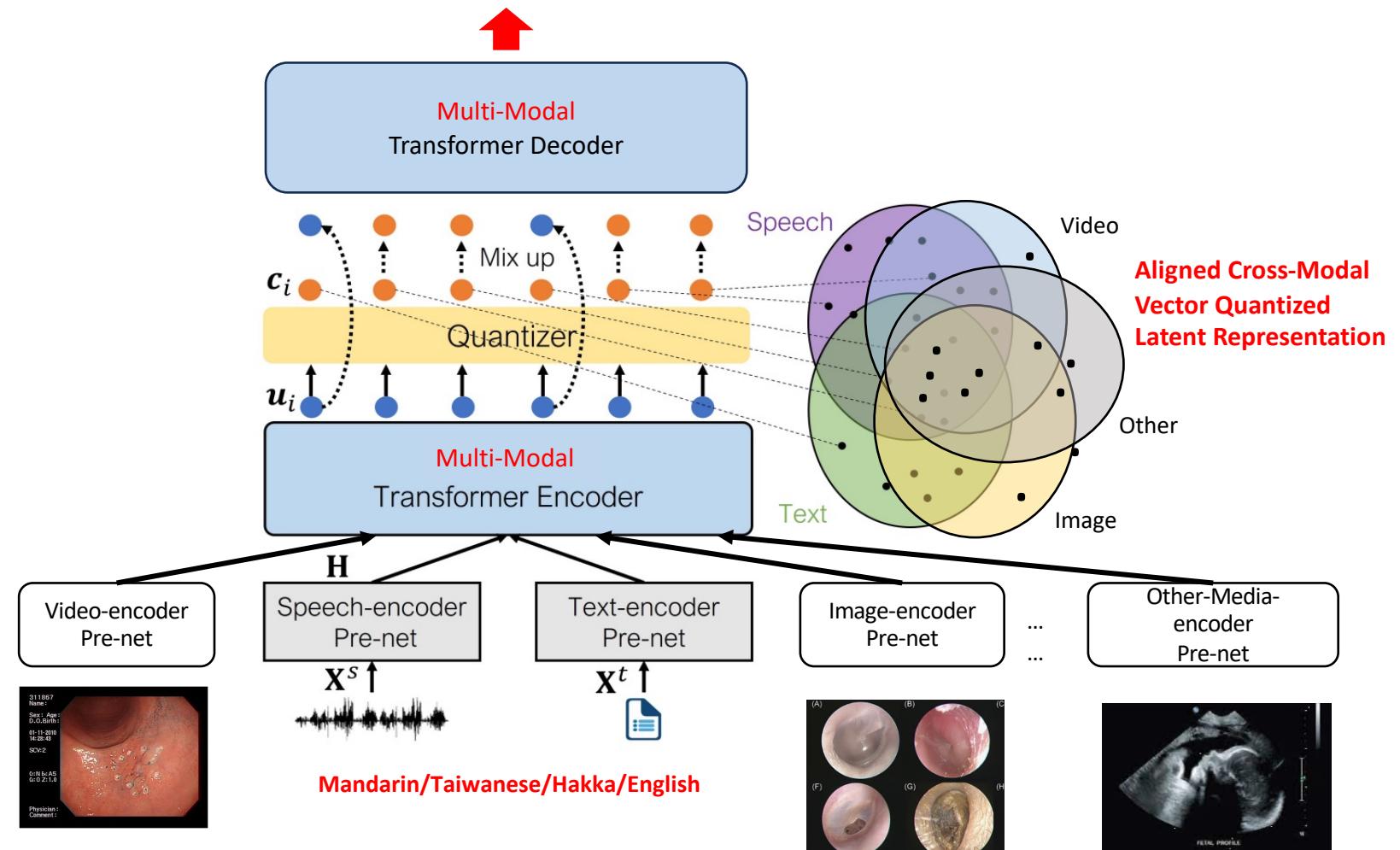
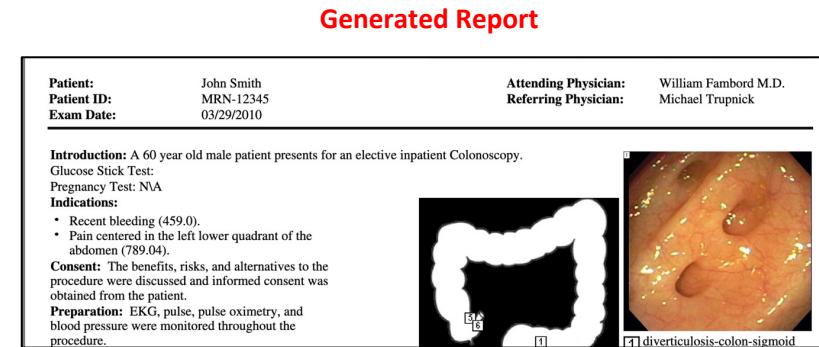
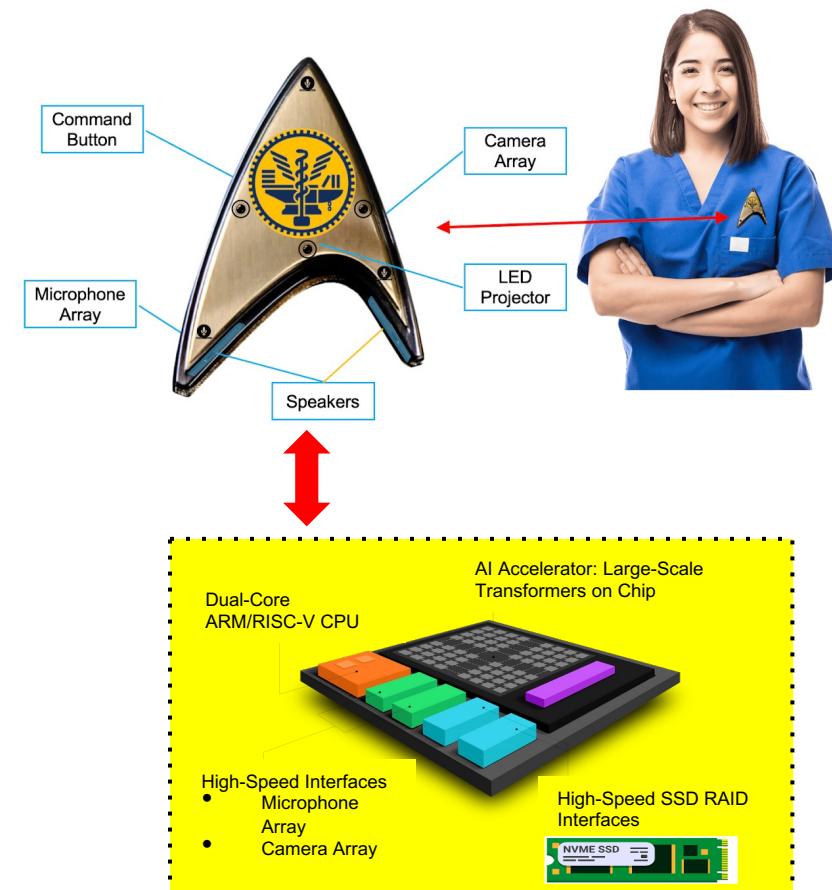
Dialogue
Management

Database

Context
Resolution

Internet

Smart Medicine



目的：台華英語音辨認器



華台英語輸入

台文輸出

華文輸出



Chinese Text Transcription Evaluation Results

Systems	#. of Characters	substitution	deletion	insertion	Error rate
Ours	10,072	316	156	78	5.50%
Other	10,072	874	453	421	17.40%

Requirements & Examples

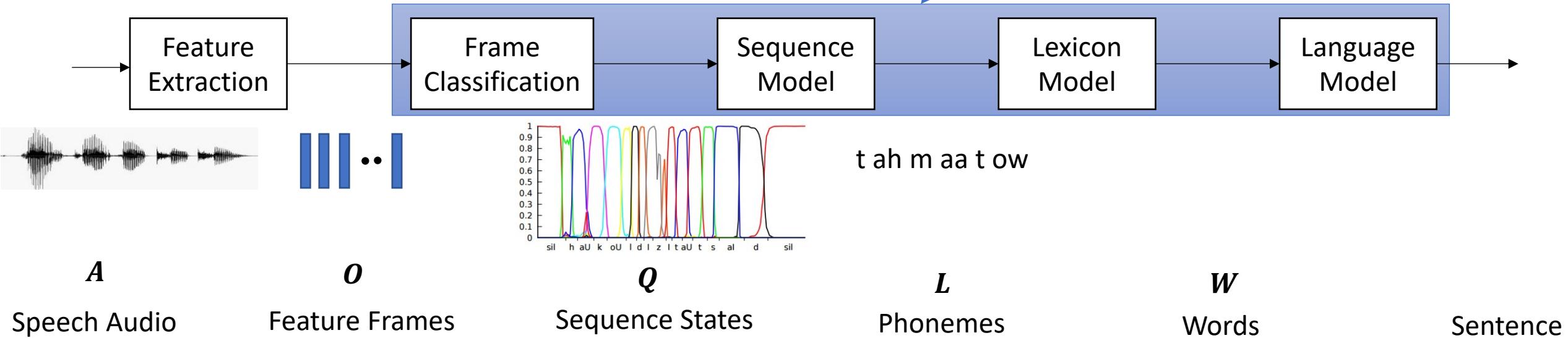
- Noisy Taiwanese Speech Recognition
- Approaches
 - Transformer (ESPnet)
 - <https://github.com/espnet/espnet/tree/master/egs2/aishell/asr1>
 - SSL-Pretrained Model + Transformer (S3PRL+ESPnet)
 - <https://github.com/espnet/espnet/tree/master/egs2/librispeech/asr1>
 - Whisper (ESPnet or Hugging Face Transformer)
 - <https://github.com/espnet/espnet/tree/master/egs2/aishell/asr1>
 - https://colab.research.google.com/github/sanchit-gandhi/notebooks/blob/main/fine_tune_whisper.ipynb

End-to-End Speech Recognition

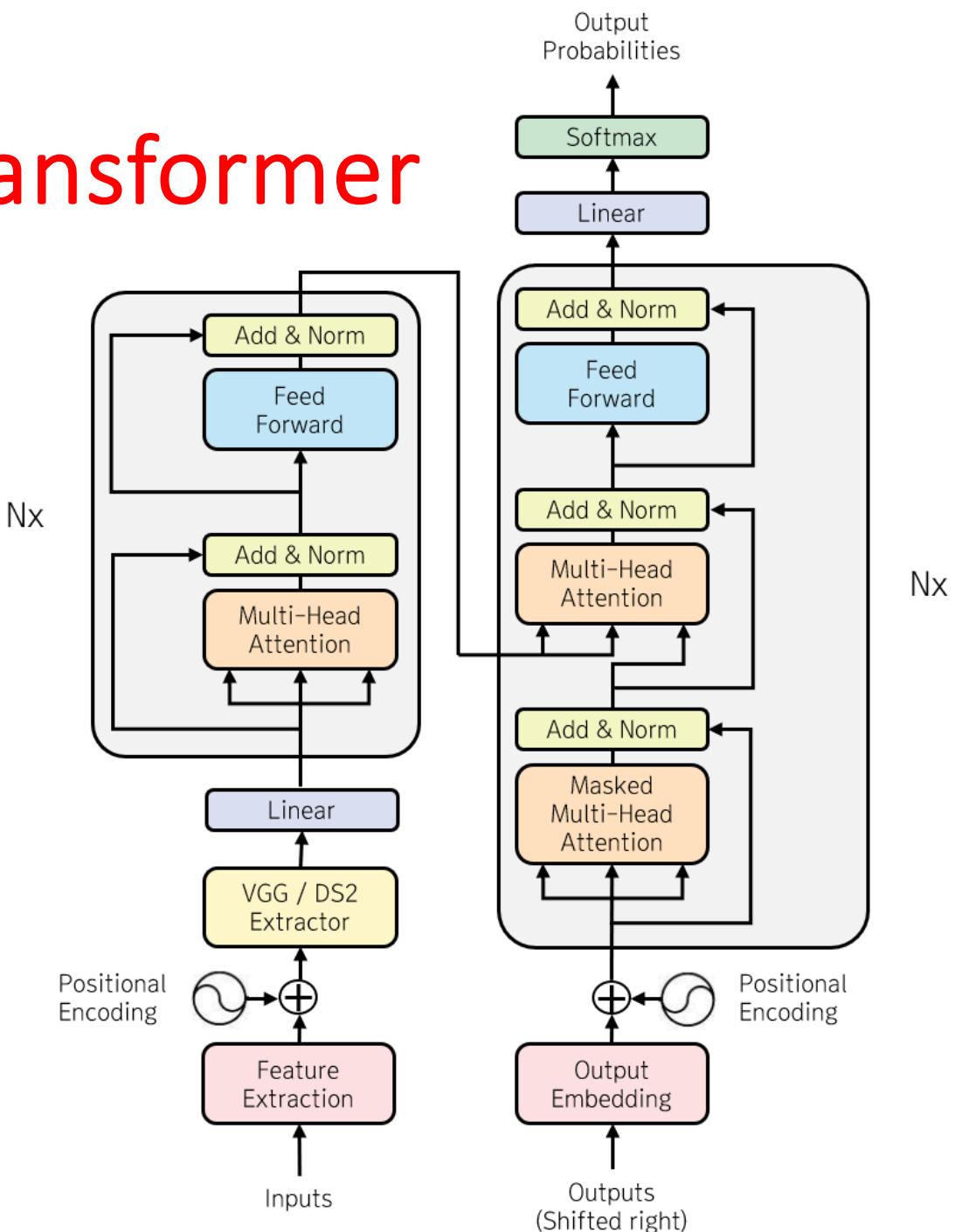
Listen, Attend and Spell (LAS)

Predict character sequence directly

Sequence-to-Sequence with Attention



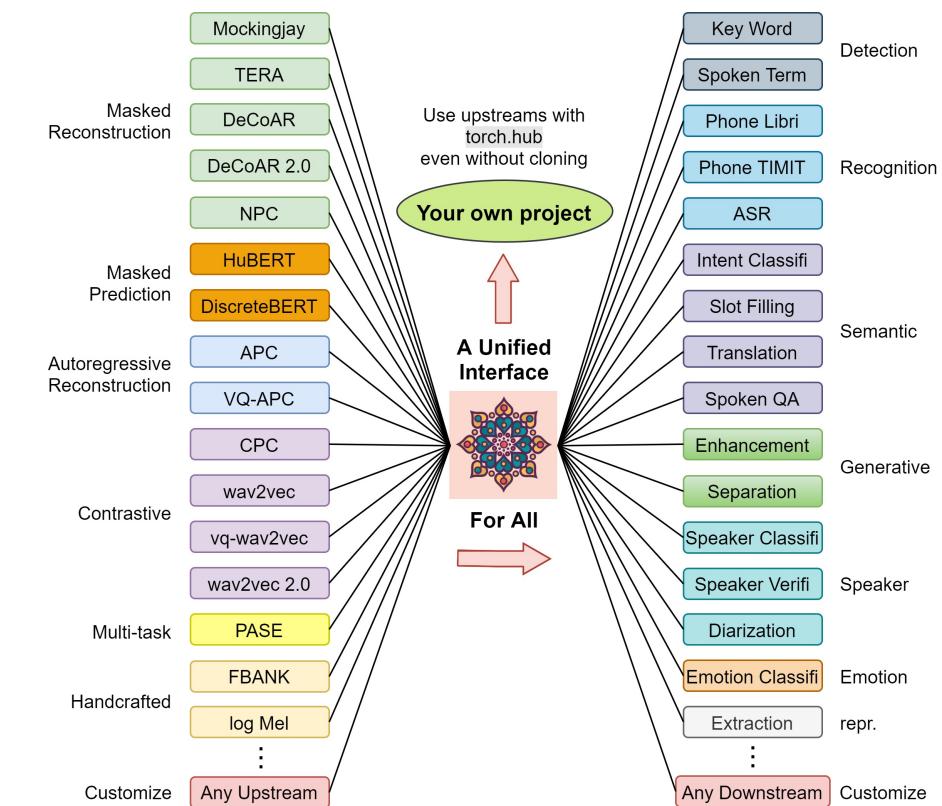
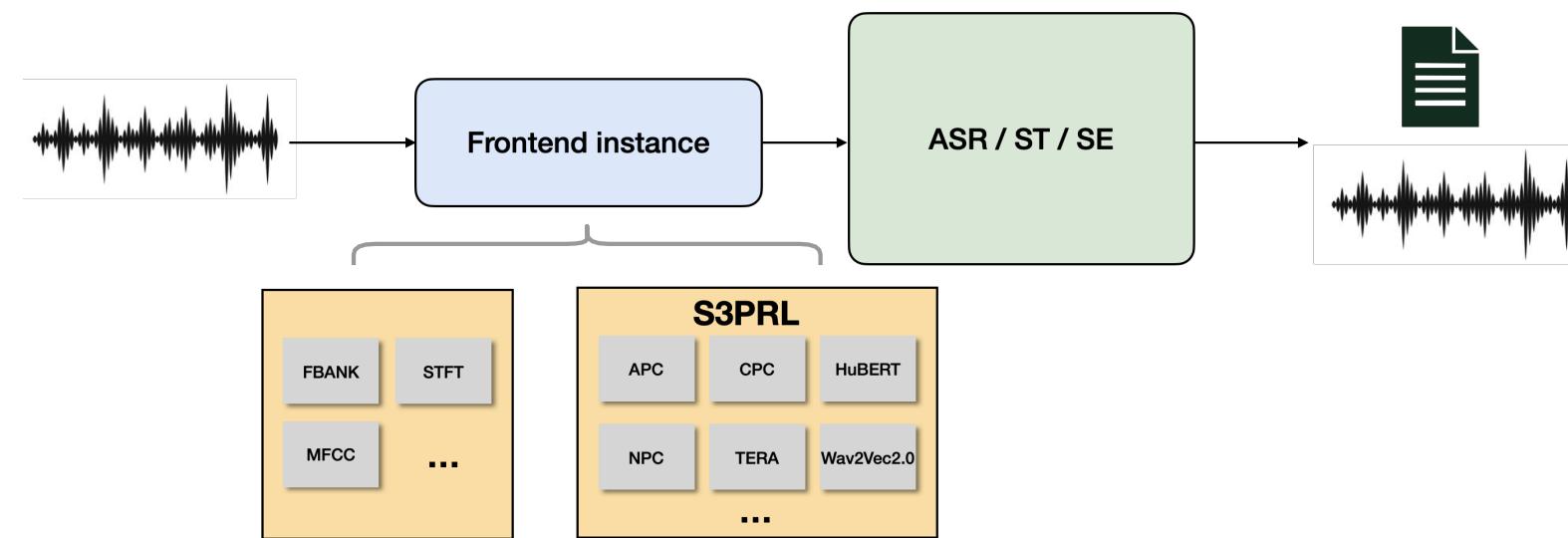
Task #1: ESPnet Transformer



Task #2: Combine ESPnet with S3PRL

[Yang+ 2021, Chang+ 2021]

- Self-supervised pretraining on speech data have achieved a lot of progress, like wav2vec2.0 [Baevski+ 2020], Hubert [Hsu+ 2021], etc.
- S3PRL* toolkit provides an integration of pretrained speech representation models and speech tasks, e.g., wav2vec2.0 + LSTM acoustic model for ASR.
- Support combine the pretrained models with advanced end-to-end speech processing models in a simple way.

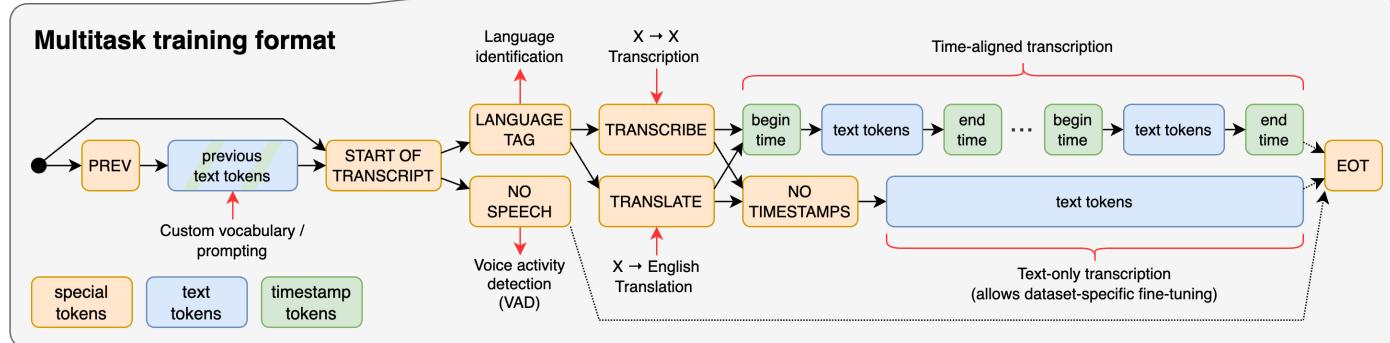
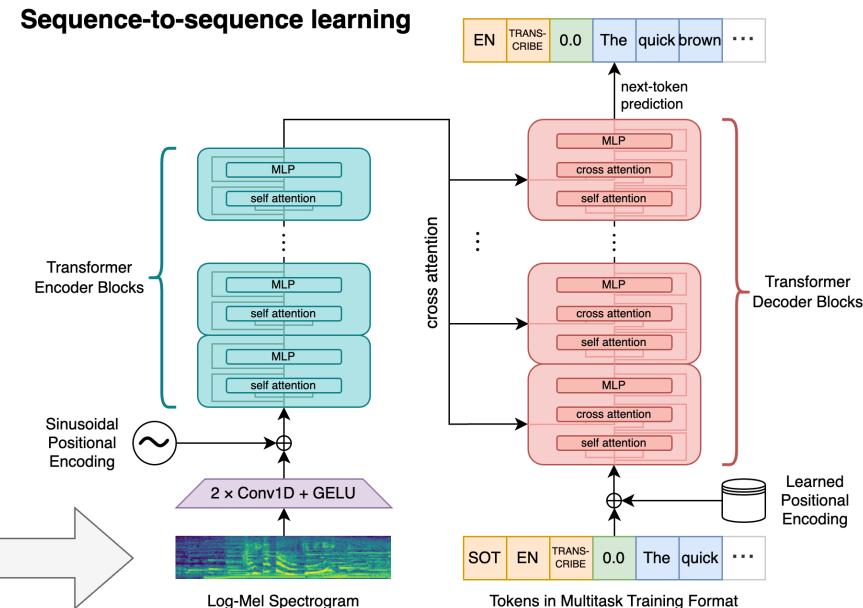
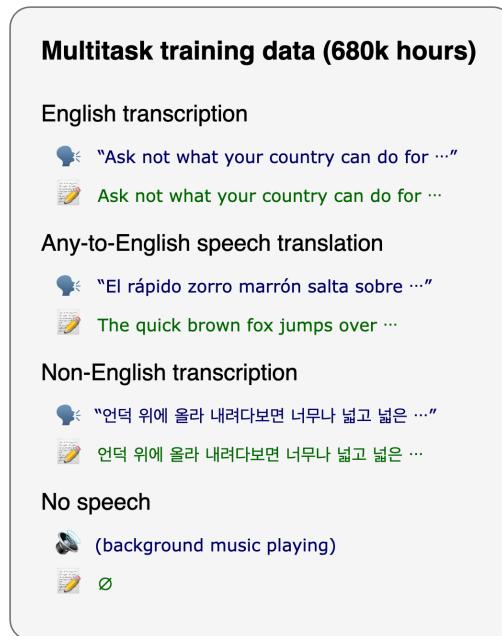


Task #3: OpenAI Whisper Finetuning

- <https://github.com/openai/whisper>
- trained on **680,000** hours of multilingual and multitask supervised data collected from the web.
- enables transcription in **multiple languages**, as well as **translation** from those languages into English

Size	Layers	Width	Heads	Parameters	English-only	Multilingual
tiny	4	384	6	39 M	✓	✓
base	6	512	8	74 M	✓	✓
small	12	768	12	244 M	✓	✓
medium	24	1024	16	769 M	✓	✓
large	32	1280	20	1550 M	x	✓

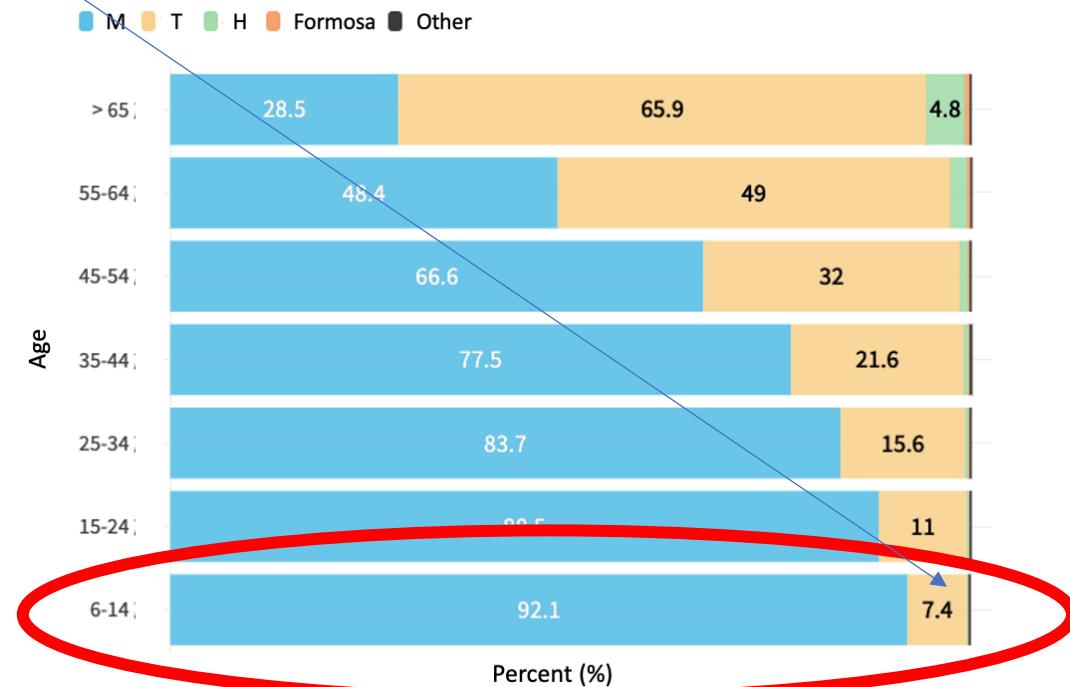
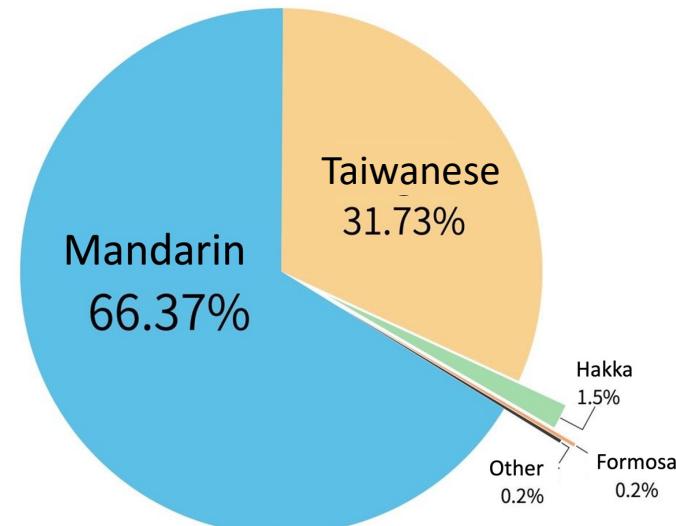
- Finetuning
 - ESPnet
 - <https://github.com/espnet/espnet/tree/master/egs2/aishell/asr1>
 - Hugging Face Transformer
 - https://colab.research.google.com/github/sanchit-gandhi/notebooks/blob/main/fine_tune_whisper.ipynb



Endangered Taiwanese & Hakka

- Taiwanese and Hakka are now in danger of extinction
 - Only **7.4%** of the **kids** use Taiwanese as their main language
 - Only **0.2%** of the **kids** use Hakka as their main language
 - Taiwanese and Hakka **may die out** in one or two generations

Taiwan's 2020 Population and Housing Census



Endangered Taiwanese Language

- 全家有智慧 [【全家有智慧】S4 臺語小教室 34 \(youtube.com\)](#)
 - Hot Taiwanese **Vocabulary** Questions and Answering TV program
 - Very difficult for many Taiwanese people, **Even for my mother**



- 簡易
 - <https://youtu.be/7rejk8M7raY>
 - <https://youtu.be/ZBrBU4UBFac>
- 中等
 - <https://youtu.be/j8PixfQe8pA?si=FpgkILUs83MufQSV>
 - <https://youtu.be/CcODCCJj3OA?si=XzhOWTZeOG88rV2>
 - <https://youtu.be/cI98Z0-IYyU?si=Zna948cqs5svirN2>
- 困難
 - https://youtu.be/0DXjq_6C44E?si=9DU5OtMvvLwVBe5W
 - <https://youtu.be/Z331QW0jeWY?si=PTdDlrzukDR8xcIN>

- → hope it's not too late to save Taiwan's languages

Unique Language Habits of Taiwanese Min

• From Japan

- 以日語形式傳入臺灣
 - 黑輪「關西腔おでん」
 - 山葵「わさび（山葵）」
 - 生魚片「さしみ（刺身）」
- 日本漢字傳入臺灣
 - 招牌 = khang-pang : 源自日本語「かんばん（看板）」
 - 胡蘿蔔 = nin-jin : 源自日本語「にんじん（人参）」 [114]
- 日文漢字臺語發音
 - 便當（餐盒） = piān-tong ; 源自日本語「べんとう（弁当）」
 - 便所（廁所） = piān-sóo ; 源自日本語「べんじょ（便所）」
 - 病院（醫院） = pēnn/pīnn-īnn ; 源自日本語「びょういん（病院）」
- 日語外來語傳入臺灣
 - 混凝土 - コンクリート , 英語 concrete
 - 麵包 - パン , 葡萄牙語 pāo

From Aboriginal

- 「臺灣」這詞即來自於南臺灣原住民西拉雅族的「Taian」或「Tayan」
- 茫然不知實情 = a-se (阿西) : 源自南部平埔語assey (不明白、不知)

臺語書寫系統

• 全漢

• 全羅

• 漢羅

獅仔皮(台文版)

深山林內有一隻驢仔，會使講所有其他的動物攏無咧共信道。
有一屆驢仔佇深山林內無意中挾著一領獅皮，驢仔著誠歡喜，因為伊心內想講，只要將即領獅皮幔佇身軀頂，著會當變做獅仔的模樣(範勢)，四界耀武揚威(展風神)。驢仔穿著獅皮了後，著誠爽勢行向樹林的東爿去。
一寡細隻動物看著伊，掠準是獅行倚來，驚甲起跤走、起跤縱。
驢仔看著更加歡喜、得意，伊擋再向樹林的西爿去，遐的動物看著，嘛攏驚甲起跤走、覲起來。
驢仔感覺足歡喜，著佇樹林內四界賴賴趨，隨便嚇遐的毋知實情的動物。驢仔煞掠準家已有偌本領咧，所以其他的動物攏會驚伊。

後來，驢仔佇山頂拄著一隻狐狸，嘛狐狸聽著講話的聲音，心內不免起了狐狸行倚驢仔身軀邊，發現伊毋是真聲誠旺聲(響亮)，敢會使喝一聲予阮驢仔聽著狐狸的呵咾，著歡喜甲大紹

酸葡萄

有一隻狐狸，看diorh葡萄生gah密密密，
一pā—pa ē葡萄攏真大粒。

狐狸ghiàn gah 喙瀾 cap-cap滴，
想bhe挽來食。

跳gǔan跳gē，跳起跳倒，|
攏挽bhe-diorh。

gàu路尾，伊跳gah無力，iau是無食diorh°。

狐狸只好講：「葡萄ia未熟，一定足酸e°，mài食kà欸

bān zī mǐ

bān-zī-hue, gūan-gūan-à, diām-diām-à,
kia_dī lām-dāi-üan ē to-dē,
kia_dī gōr-hiōng ē gē-lō
bhī-bhī-a ciōr, īn-ui cūn-hong bhī-bhī-a cue.
hong-hun ē hūn-cài,
ài gà ī bi hun-āng-a dè,
hai ī lu kui lu zē !
m giān hong cūe lorh tō,
m guan lāng tiān-tang a si_ùan-cēh,
bān-zī-mǐ mua ciú gi,
dak gang hiòng tiⁿ iat ciù.
iong mua-sim ē hūe-lùi,
hùa zoং cūn-tiⁿ ē hor sī-ziāt

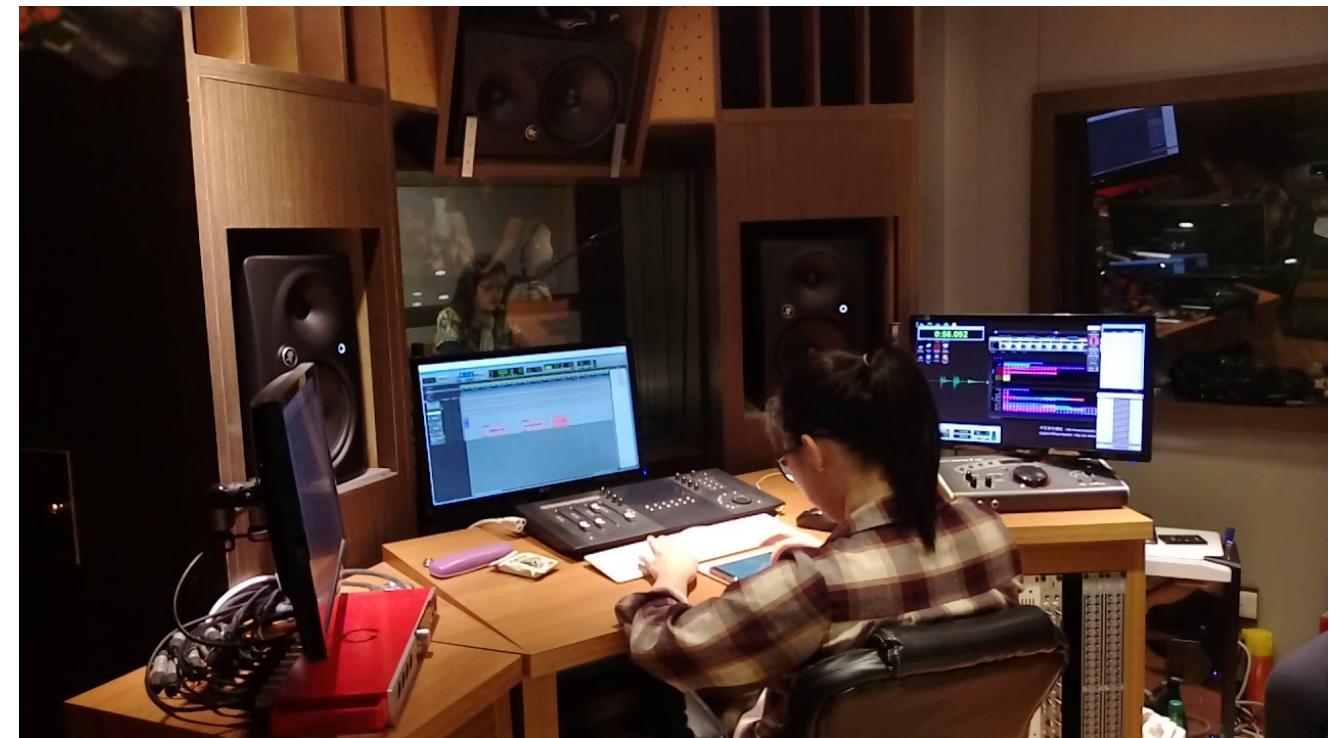
台羅調號數字調轉換

```
>>> from 臺灣言語工具.解析整理.拆文分析器 import 拆文分析器
>>> from 臺灣言語工具.音標系統.閩南語.臺灣閩南語羅馬字拼音 import 臺灣閩南語羅馬字拼音
>>>
>>> 台羅 = 'Tiān-náu mā ē kóng Tâi-gí!'
>>> 台羅物件 = 拆文分析器.建立句物件(台羅)
>>>
>>> 台羅物件.轉音(臺灣閩南語羅馬字拼音).看語句()
tian7-nau2 ma7 e7 kong2 tai5-gi2!
>>>
>>> 數字調 = 'tian7-nau2 ma7 e7 kong2 tai5-gi2!'
>>> 數字調物件 = 拆文分析器.建立句物件(數字調)
>>> 數字調物件.轉音(臺灣閩南語羅馬字拼音, '轉調符').看語句()
'tiān-náu mā ē kóng tâi-gí!'
```

Data Description

- 台灣嬌聲 2.0
 - <https://suisiann-dataset.ithuan.tw>

<https://youtu.be/GJvtWyuizyA>



Data Format

- lexicon.txt
- train/*.wav
- Test/*.wav
- train.csv
- test.csv
- sample.csv

台羅拼音	子音	母音
Khau	kh	au
Kheh	kh	eh
Khe	kh	e
Khennh	kh	ennh
.....		

```
id,text
1,li2 be7 e5 mih8 kiann7 lan5 lan5 san1 san1 long2 be7
tsiau5 tsng5
2,suah4 ka7 li2 tim3 tioh8
3,kiu3 lang5 ooh4
```

.....

```
id,text
1,a1 e2 i3 o4 u5
2,a1 e2 i3 o4 u5
3,a1 e2 i3 o4 u5
```

.....



NYCU-IASS-DL2024-Taiwanese-ASR

Taiwanese Speech Recognition using ESPnet Toolkit and Self-Supervised Pre-Trained Model

ESPnet: end-to-end speech processing toolkit



ESPnet is an end-to-end speech processing toolkit, mainly focuses on end-to-end speech

- Metric

- Word error rate (WER) = $(D + S + I) / N \times 100\%$
 - N - total number of labels (總詞數)
 - D - deletion errors (刪除錯誤)
 - S - substitution errors (替換錯誤)
 - I - Insertion errors (插入錯誤)

注意測試音檔，有加上變形與背景雜訊！

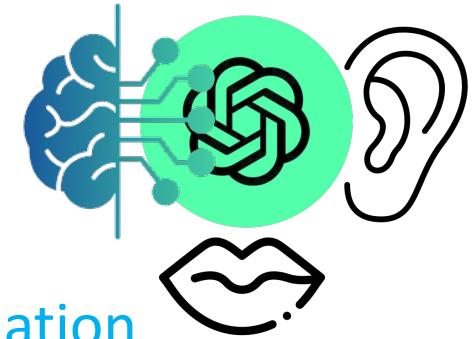
- 但人聽得懂的，電腦也要能聽得懂。
- 使用 audiomentations & nlpaug 做 Data Augmentation
 - <https://github.com/iver56/audiomentations>
 - <https://github.com/makcedward/nlpaug>

```
import glob, os
import numpy as np
from scipy.io import wavfile
from audiomentations import Compose, SomeOf, AddGaussianNoise, AddGaussianSNR, TimeStretch, PitchShift,
Shift, AddBackgroundNoise, AddShortNoises, PolarityInversion, ApplyImpulseResponse
from audiomentations.core.audio_loading_utils import load_sound_file
import nlpaug.augmenter.audio as naa
import nlpaug.flow as naf
```

本土語言AI復振

- 本土語言生活化服務&應用 → AI來幫忙！

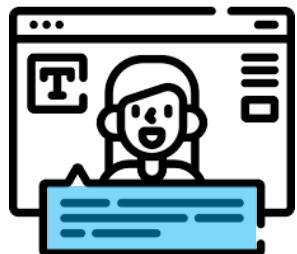
- 聽：**Speech Recognition** → Live Media Subtitling
- 說：**Speech Synthesis** → Personalized Multilingual TTS
- 翻譯/口譯：**Machine Translation** → Speech-to-Speech Translation



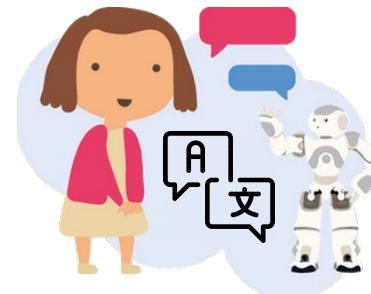
- 本土語言教學人力與時數不足 → 個人化AI家教！

- 幫老師準備上課教材、出作業與考卷（自動評分，改考卷）
- 讓學生回家能自行練習聽、說、讀、寫（自動發音評測，偏誤偵測/建議，個人專屬AI家教機器人）

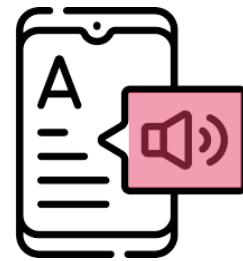
Live Media Subtitling



Machine Translation



Personalized TTS



Pronunciation Assessment

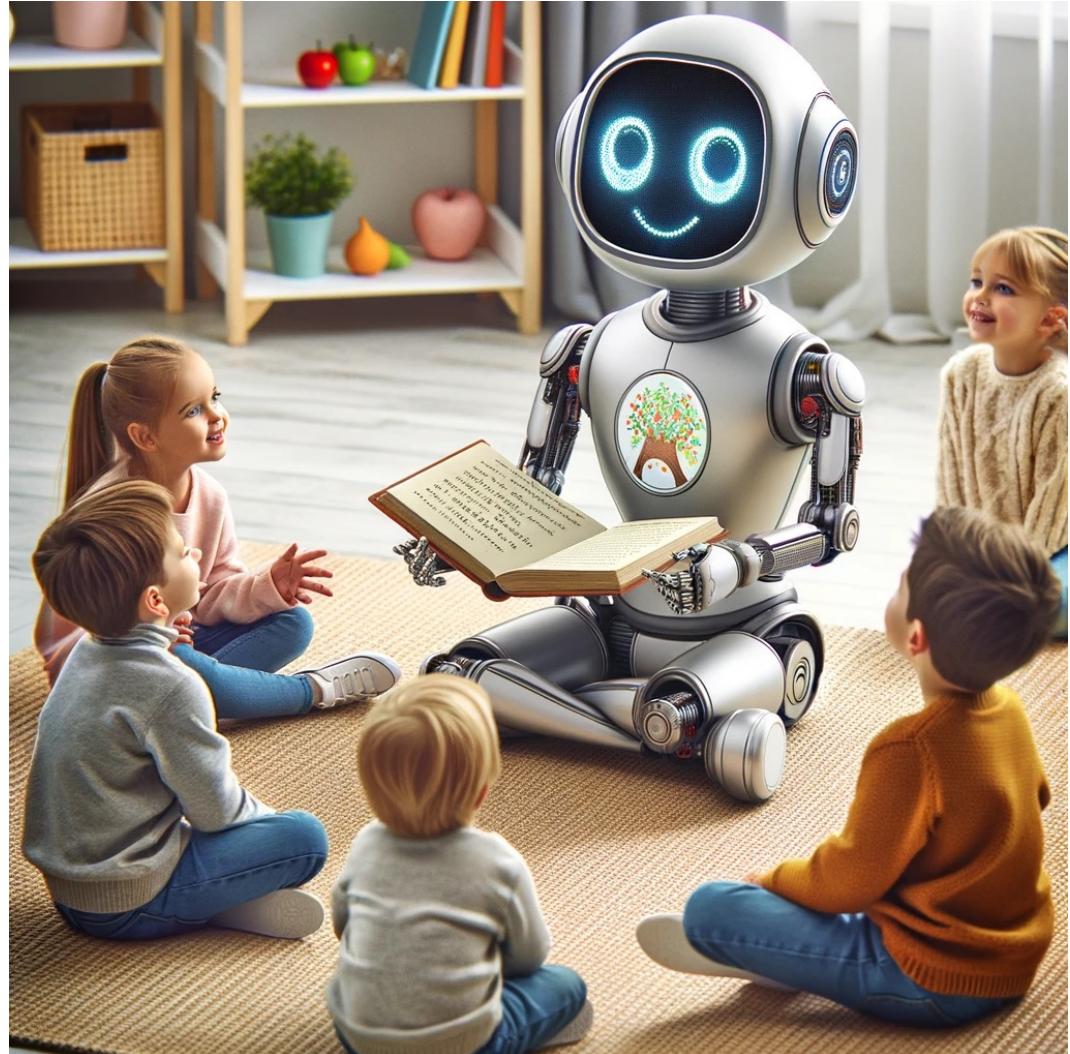


Essay Assessment



目標：個人化母語AI家教

- 會看，會聽，會說的機器人
- 母語 聽/說/讀/寫 教學
- 需要State-of-the-Art母語語音AI
- 需要State-of-the-Art母語原生LLMs



技術盤點：台華英電子會議紀錄系統



<https://youtu.be/ynT7beuwv2E>

- 台華英語輸入
- 華文輸出
- 台文輸出

BC 會議記錄系統

逐字稿生成 | 檔案管理 | 使用者回饋 | 帳號管理 | 操作紀錄 | 系統管理 | 模型管理 | 登出

檔案管理 > 逐字稿編輯

標題：videoplayback (9)

創建時間 2023/10/15 01:23

語音長度 37:06

參與者 1 speaker_0

儲存編輯 | 逐字稿下載

逐字稿

關鍵字搜尋 | 新增刪除

1 speaker 0 00:00:01

大家好，歡迎收看我們今天的永哥跟你講，我們要再向大家拜託一下，每個禮拜五，晚上十點到十一點半，收看我們的全國第一樣，因為這是我們的基礎，有全國第一樣才有永哥跟你講，所以我們要弄清楚，什麼是主，什麼是寵，我們在民法裡面

1 speaker 0 00:00:22

叫做主務跟寵物，要搞清楚，開玩笑啦，我知道你們大家都有看，我要向大家謝謝，不高興的事情，也有很多高興的事情，其實每天都會發生，像我最近不高興的事情，像我最近不高興的事情，除了我告別人，大家都知道，我被人詛咒的氣氛很差

1 speaker 0 00:00:43

我告了十幾個了，當然我現在還看，我每次都問說，看到了就傳給我，很多人跟我說，永哥，不看了，有害怕嗎這就表示他們確實是冒險，不然你怎麼會害怕，對不對，這有什麼辦法，大聲說真的不道理，我還要向大家謝謝

1 speaker 0 00:01:04

你們對我暖心的關懷，讓我感受到滿滿的祝福，跟滿滿的溫暖，我還要在這裡向大家道謝，當然我這兩天，這一個禮拜

00:00:00 | 00:37:06

1x | 播放控制符號

技術盤點：台華英電子會議紀錄系統



<https://youtu.be/RkTAmspFOOw>

- 台華英語輸入
- 華文輸出
- 台文輸出

BC 會議記錄系統 | 逐字稿生成 | 檔案管理 | 使用者回饋 | 帳號管理 | 操作紀錄 | 系統管理 | 模型管理 | 登出

檔案管理 > 逐字稿編輯

標題 : videoplayback (9)

創建時間 2023/12/23 02:30

語音長度 37:06

參與者 1 speaker_0

逐字稿

1 speaker_0 00:00:00 啟逐家好歡迎收看咱今仔日的楊哥恰你講啊原仔愛閣共逐家拜託一下每禮拜到拜五暗時10點到11點半收看咱的全國第一樣因為這是咱的基礎有全國第一樣才有楊哥共你講嘛啊所以啊咱愛攏無清楚什麼是主什麼是從我們在民法裡面叫做主務跟從務

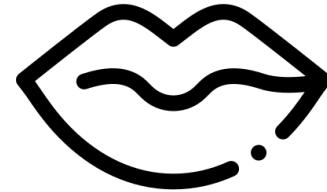
1 speaker_0 00:00:24 嘿要搞清楚乎啊開玩笑啦乎我知影恁逐家攏有看原仔共逐家細多謝平欸有真濟今仔日彼个無歡喜的代誌也有真濟歡喜的代誌其實每一工攏會發生乎像我最近無歡喜的代誌乎除了我告人的逐家知影無我去予人汙染的這個氣魔足罷的我告10幾個矣

00:00:00 00:00:45 當然我這馬閣看我逐家問講欸我看著已經傳落我真濟人共我講楊哥無看的啊嘿有驚喔這就表示講個確實是無彼嘛無你哪會

00:37:06

1x

Chinese Text-to-Taigi Speech Synthesis



請選擇腔調
強勢腔 (高雄腔) 女聲 語音合成模型

請輸入中文文句
範例：今天天氣很好，我們一起出去玩
Keyin a Chinese Sentence

今天天氣很好，我們一起出去玩。

產生台語拼音 產生 漢羅台文 清除

請輸入漢羅台文
範例：新年欲到矣，咱來去買年菜好無？
Keyin a Taigi Hnaji Sentence

今仔日天氣真好，咱做伙出去追追。

產生台語拼音 清除

請輸入台羅拼音
拼音範例：kin1-a2-jit8 thinn1-khi3 tsin1 ho2, lan2 tso3-hue2 tshut4-khi3 tshit4-tho5.
Keyin a Taigi Pinyin Sentence

kin1-a2-jit8 thinn1-khi3 tsin1-ho2 , lan2 tso3-hue2 tshut4-khi3 tshit4-tho5.

講台語 **Speak It**

合成音檔完成!

▶ 0:04 / 0:04 ━━ ◘ :

<http://tts001.iptcloud.net:8804>

- **Short Sentence**

1. 大家好，我是會說台語的機器人



2. 龜笑鱉無尾，鱉笑龜粗皮



3. 武漢肺炎的出現，讓全世界的人都開始戴口罩。



- **Bank Q&A**

1. 如何查詢辦卡進度?



2. 您可於官網查詢辦卡進度，完成身分驗證後查看。



3. 辦卡有缺資料，如何補件?



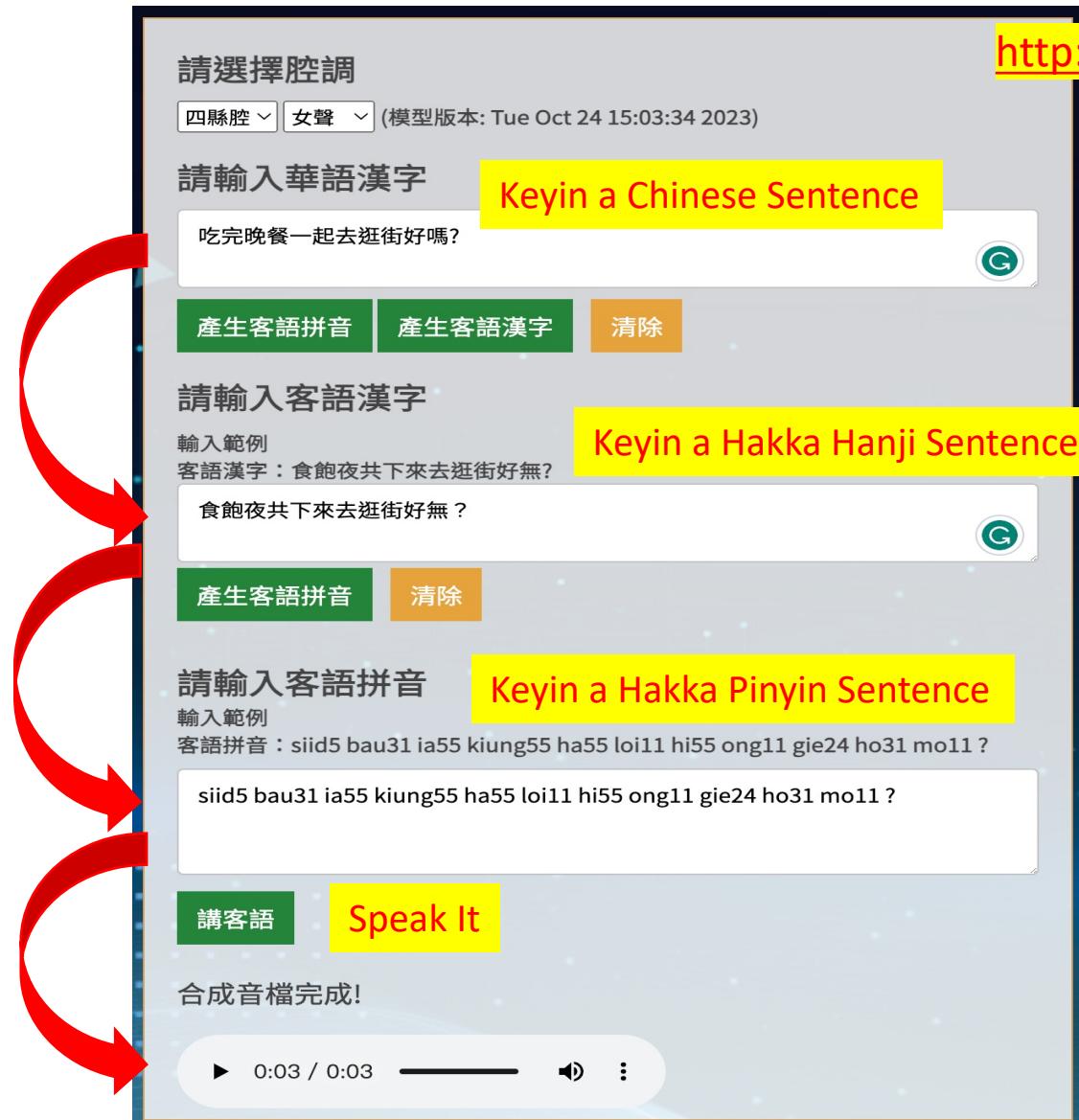
4. 您可於官網上傳補件，選取”申辦信用卡補件資料”完成補件。



5. 請問信用卡道路救援提供那些服務項目?



Hakka Text-to-Speech System

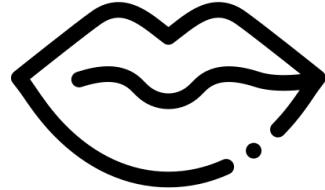


The screenshot shows the system's user interface with three main sections:

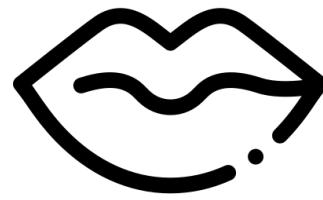
- 請選擇腔調**: Set腔調 (Four County Dialect) to 女聲 (Female Voice). Model version: Tue Oct 24 15:03:34 2023.
- 請輸入華語漢字**: Keyin a Chinese Sentence: 吃完晚餐一起去逛街好嗎? Buttons: 產生客語拼音 (Generate Hakka Pinyin), 產生客語漢字 (Generate Hakka Hanji), 清除 (Clear).
- 請輸入客語漢字**: Keyin a Hakka Hanji Sentence: 食飽夜共下來去逛街好無? Buttons: 產生客語拼音 (Generate Hakka Pinyin), 清除 (Clear).
- 請輸入客語拼音**: Keyin a Hakka Pinyin Sentence: 輸入範例: 客語拼音: siid5 bau31 ia55 kiung55 ha55 loi11 hi55 ong11 gie24 ho31 mo11 ? Buttons: 講客語 (Speak Hakka), Speak It (highlighted in yellow).

A red arrow on the left side of the interface points downwards from top to bottom, indicating the flow or sequence of the three input methods.

<http://tts001.uptcloud.net:8803>



- 大家好，我是會說客語的機器人
 - 武漢肺炎的出現，讓全世界的人都開始戴口罩。
-
- 向前看
 - 人要有信心
 - 食飽夜共下來去逛街好無？
 - 上午分家下午借，兄弟分家成鄰居，所以，除了四叔公，每個人都要在自己分到的山地上蓋新家，所以伙房一下就安靜下來。
 - 認證考試



技術盤點：語音合成

新文化文特報

同胞須團結，團結真有力，一起來聲援！

獅子狩與獅子吼

臺灣民報 THE TAIWAN MINPAO

刊旬
【審庭紀世】

重回治警事件，10/1準時發聲！

推行「臺灣議會設置請願運動」有甚麼錯？總督府竟然發動全島大逮捕，把我們抓去審判。為了展現決心，10月1日務必來「臺灣新文化運動紀念館」化身記者聲援我們！現場有不僅復刻法庭答辯，更利用「OpenAI」與「GPT」技術，能當面和我聊天。

https://youtube.com/shorts/v_9n5nagUq0

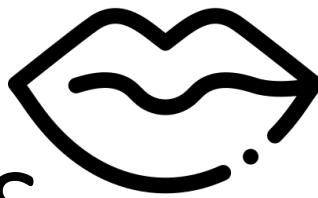
治警事件百年紀念展 聽 蔣渭水 先生說甚麼？



臺灣新文化運動紀念館 Taiwan New
Cultural Movement Memorial
Museum

1.5 萬個讚 • 1.7 萬位追蹤者

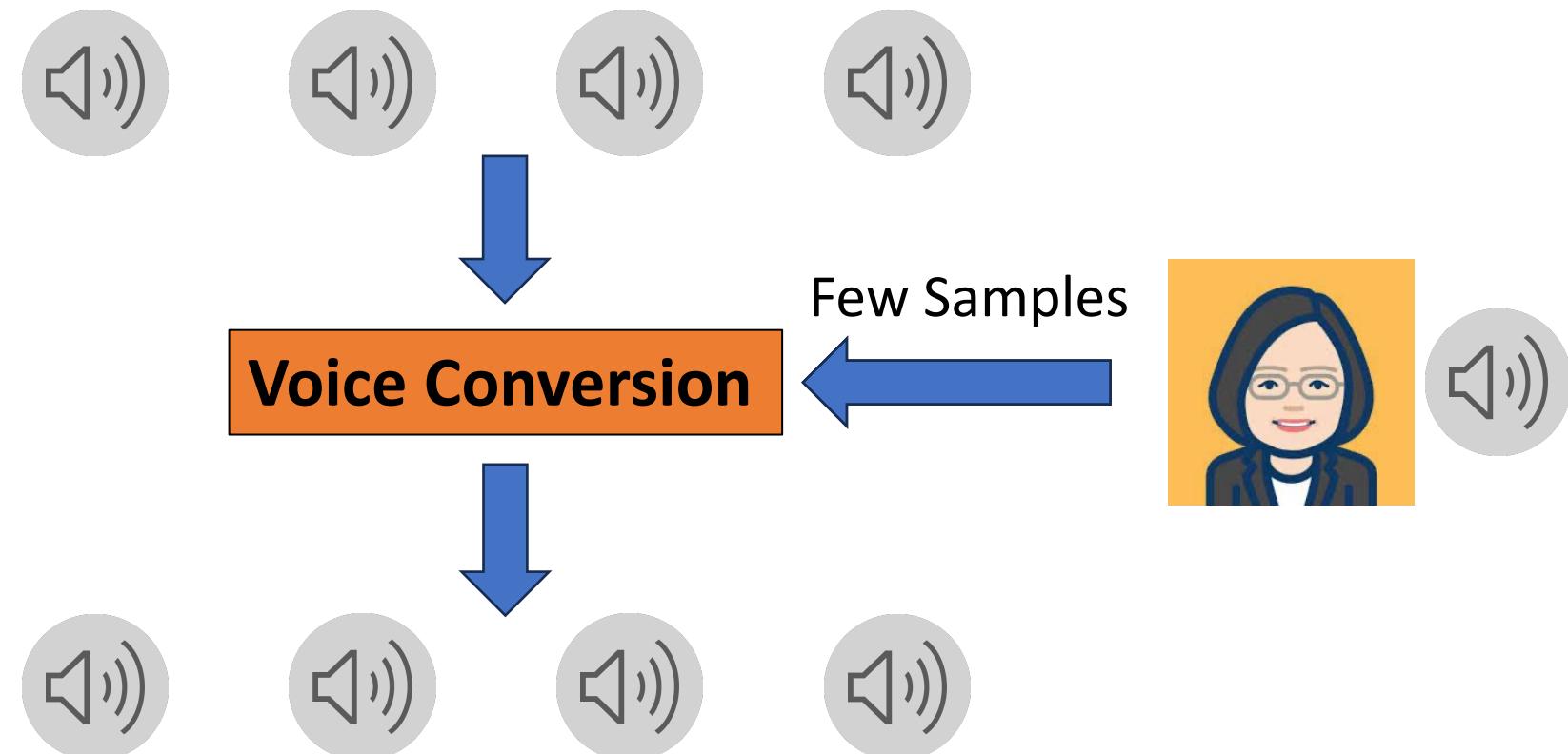
<https://youtu.be/jQDT7DbAMgU?si=SL126ZPGAHhnrn>



技術盤點：Personalized Multi-lingual TTS

Mandarin/Taiwanese/Hakka/English

- Source: TTS Output



- Target: President Tsai

輸入文本

今天天氣很好，我們一起出去玩。

選擇翻譯模式

中文 → 漢羅台文

Clear

Submit

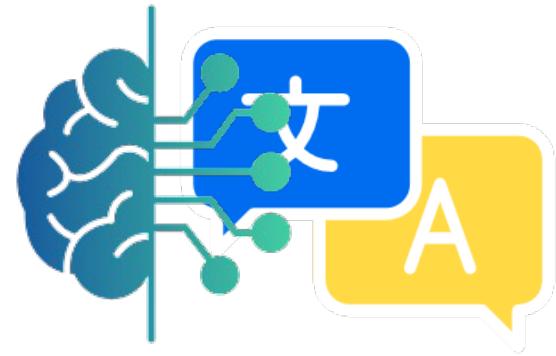
output

今仔日天氣真好，咱做伙出去迫迫。

Flag

Examples

輸入文本	選擇翻譯模式
今天天氣很好，我們一起出去玩。	中文 → 台羅拼音數字調
新年欲到矣，咱來去買年菜好無？	漢羅台文 → 台羅拼音數字調
今天天氣很好，我們一起出去玩。	中文 → 漢羅台文
新年欲到矣，咱來去買年菜好無？	漢羅台文 → 中文
食飽夜共下來去逛街好無？	客語漢字 → 客語拼音數字調
今天天氣很好，我們一起出去玩。	中文 → 客語漢字
食飽夜共下來去逛街好無？	客語漢字 → 中文



Chinese/Taiwanese/Hakka Machine Translation

- [http://tts001. iptcloud. net: 8802](http:// tts001. iptcloud. net: 8802)



台客語大語言模型 (Low Resources)

- TAIDE Chat 測試網站

- <https://chatdev.gai.tw/>



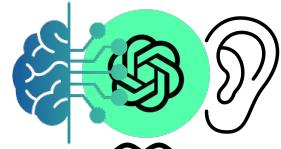
- 題目

1. 介紹自己
2. 台灣的國慶是哪一天？
3. 用python寫一個bubble sort程式
4. 請寫一篇台文短文，檢討 “手機仔侷現代人的生活”
5. 請以 “客語薪傳這條路” 為題，用客語漢字寫一篇作文

The screenshot displays the TAIDE Chat interface with four main sections:

- 文本翻譯 (Text Translation):** A user prompt asks to translate "The International Federation of the Phonographic Industry has announced its latest Global Artist Chart, which features a Taiwanese artist in the top 10..." into traditional Chinese. The AI response is: "國際唱片業協會公佈了最新的全球藝術家排行榜，其中一位臺灣藝術家上升至前十名。..."
- 寫文章 (Write Article):** A user prompt asks to write an article about a special gift received. The AI response is: "我曾經收到一份特別珍貴的禮物，那是我的朋友送我的。當時我們正在一個朋友的生日派對上，他突然走到我身旁，告訴我他有一個禮物想給我。我很驚訝，因為..."
- 寫信 (Write Letter):** A user prompt asks to write a letter to a friend recommending five must-see scenic spots in Taiwan and their unique features. The AI response is: "給一位朋友寫一封信，推薦他們去臺灣旅遊的五個必去景點以及各自的特色，並說明為什麼這些地方值得一遊。"
- 自動摘要 (Automatic Summary):** A user prompt asks for a summary of an article about DeepMind's AI chess player AlphaZero. The AI response is: "請提供這篇文章的要點概述。DEEPMIND近日又開發出新一代AI棋類高手ALPHAZERO，只靠增強式學習演算法和遊戲規則，從隨機下棋開始歷經數百萬回自我博弈，來自學圍棋、西洋棋和日本象棋..."

技術盤點：Speech-to-Speech Translation



<https://translate.bronci.com.tw/>

Mandarin to Taigi

Taigi to Mandarin

Taigi to Hakka

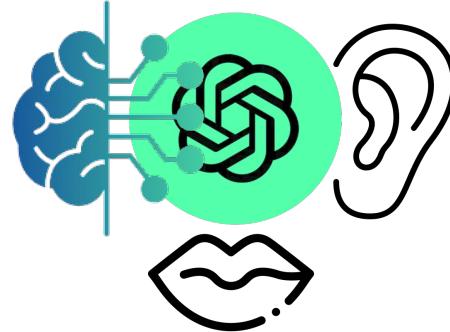
Taigi to Japanese

翻譯器-華台客英日 Beta測試中
國立陽明交通大學 & 長問科技

翻譯內容為OpenAI所產生，僅供參考

技術盤點：Taiwanese/Hakka Speech-enabled LLMs

- Four Modules
 - ASR
 - Translation
 - TTS
 - ChatGPT/TAIDE



- Online Demo

- URL: <https://speech.bronci.com.tw>
- QRCode:



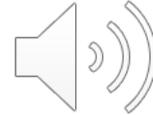
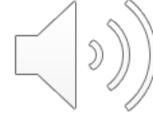
- Youtube Video

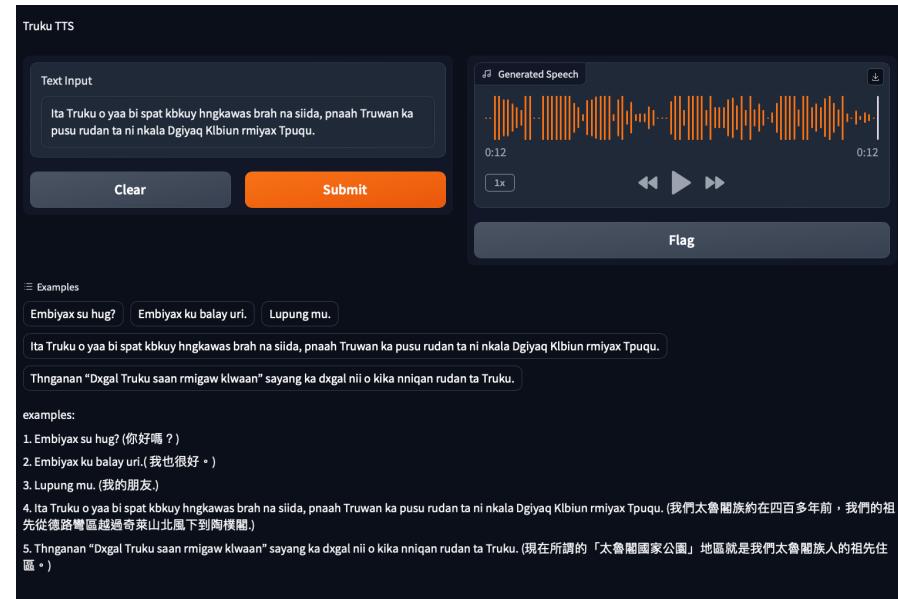


The screenshot shows a mobile application interface for learning Taiwanese/Hakka. The top status bar indicates the time is 9:55. The header features the National Yang Ming Chiao Tung University logo and language options: 華台語 | 華客語. The main content area has a blue background with white text. It displays a conversation in Chinese and English, with some text in orange. Several UI elements are highlighted with red boxes and arrows:

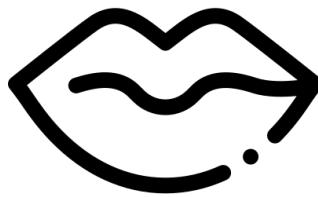
- Chinese Responses**: A red box highlights a section of the Chinese text response.
- English Responses**: A red box highlights a section of the English text response.
- ASR Result Chinese**: A red box highlights the Chinese transcription of the user's input "你可以教我講臺語嗎？".
- ASR Result English**: A red box highlights the English transcription of the user's input "Can you teach me to speak Taiwanese?".
- Click to Speak**: A red box highlights the green microphone button labeled "錄音" (Record).
- Click to Stop Speaking**: A red box highlights the orange button with a square icon and the text "結束" (End).
- Language: Mandarin + Taiwanese or Hakka**: A callout box points to the language selection area at the top right of the screen.
- TTS: Male or Female**: A callout box points to the gender selection buttons (男聲, 女聲) at the top right of the screen.

Truku TTS

- Embiyax su hug? (你好嗎？) 
- Embiyax ku balay uri. (我也很好。) 
- Lupung mu. (我的朋友。) 
- Ita Truku o yaa bi spat kbkuy hngkawas brah na siida, pnaah Truwani ka pusu rudan ta ni nkala Dgiyaq Klbiun rmiyax Tpuqu. (我們太魯閣族約在四百多年前，我們的祖先從德路彎區越過奇萊山北風下到陶樸閣。) 
- Thnganan “Dxgal Truku saan rmigaw klwaan” sayang ka dxgal nii o kika nniqan rudan ta Truku. (現在所謂的「太魯閣國家公園」地區就是我們太魯閣族人的祖先住區。) 



The screenshot shows the Truku TTS application interface. At the top, there's a text input field containing the Truku sentence "Ita Truku o yaa bi spat kbkuy hngkawas brah na siida, pnaah Truwani ka pusu rudan ta ni nkala Dgiyaq Klbiun rmiyax Tpuqu.". Below the text input are "Clear" and "Submit" buttons. To the right, there's a "Generated Speech" section displaying a waveform and playback controls (rewind, play, fast forward). Below the waveform, there are examples of other Truku phrases: "Embiyax su hug?", "Embiyax ku balay uri.", and "Lupung mu.". Further down, there's a note about the historical migration of the Truku people from De Luwan to Tao-Puk. The entire interface has a dark theme.



展望：原住民大語言模型 & 語音AI？

• Truku TTS-based Audiobook



Embiyax su hug?

你好嗎？

Embiyax ku balay uri.

我也很好。

Lupung mu.

我的朋友。

Ita Truku o yaa bi spat kbkuy hngkawas brah na siida, pnaah Truwān ka pusu rusan ta ni nkala Dgiyaq Klbiun rmiyax Tpuqu.

我們太魯閣族約在四百多年前，我們的祖先從德路彎區越過奇萊山北風下到陶樸閣。

Thnganan "Dxgal Truku saan rmigaw klwaan" sayang ka dxgal nii o kika nniqan rusan ta Truku.

現在所謂的「太魯閣國家公園」地區就是我們太魯閣族人的祖先住區。

你好嗎？

Embiyax su hug?

我也很好。

Embiyax ku balay uri.

我的朋友。

Lupung mu.

我們太魯閣族約在四百多年前，我們的祖先從德路彎區越過奇萊山北風下到陶樸閣。

Ita Truku o yaa bi spat kbkuy hngkawas brah na siida, pnaah Truwān ka pusu rusan ta ni nkala Dgiyaq Klbiun rmiyax Tpuqu.

現在所謂的「太魯閣國家公園」地區就是我們太魯閣族人的祖先住區。

Thnganan "Dxgal Truku saan rmigaw klwaan" sayang ka dxgal nii o kika nniqan rusan ta Truku.



References

- 臺灣閩南語羅馬字拼音方案使用手冊
 - <https://ws.moe.edu.tw/001/Upload/FileUpload/3677-15601/Documents/tshiutsheh.pdf>
- 門拍字
 - <https://suisiann.ithuan.tw>
- 臺灣言語工具
 - https://github.com/i3thuan5/tai5-uan5_gian5-gi2_kang1-ku7.git
 - https://github.com/i3thuan5/tai5-uan5_gian5-gi2_kang1-ku7/blob/c3aa301df701a91c7a2ed41138218d276b24840f/文件/常見情境.md