

使用情境流程

1. 編輯 `conversation.txt`，此檔案中每一行都是一句話，希望讓使用者練習說這句話。
2. `python main.py` 執行後列出偵測到的音源輸入列表(麥克風)，選擇一個輸入。
3. 開始讀句子練習，程式會逐句依序顯示出句子，句子出現後使用者可開始錄音。
4. 語句結束後程式會自動上傳，並顯示語音辨識結果。
5. 若結果正確則繼續下一句，若錯誤則程式將播放出正確的讀音，請使用者再念一次。
6. 程式持續循環直到 `conversation.txt` 的句子都結束為止。

技術說明

此應用程式主要使用到 Azure 的兩個服務，`speech-to-text` 和 `text-to-speech`。要使用這兩個服務必須先在 Azure 上開資源 (Cognitive: Speech)，取得資源的 `name` 和 `key`，放到 `config.json` 裡。步驟可以參考官方文件。

<https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/quickstart-python-text-to-speech>

Speech-to-Text

這部分定義在 `microphone.py` 裡的 `__get_token` 和 `__speech2text` 函數。用 `key` 透過 `http request` 取得 `token` 後再利用 `token` 再一次透過 `http request` 把取得結果，最後顯示出來。

Text-to-Speech

這部分定義在 `speaker.py` 裡的 `__get_token` 和 `__text2wav` 函數。架構和 `Speech-to-Text` 非常類似。

Speech service API 更詳細的內容可參考

<https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/rest-apis>