

TAT (Taiwanese Across Taiwan)

by 廖元甫 臺北科技大學

一、 簡介

TAT(Taiwanese Across Taiwan)為錄製語料，收集來自台灣各地的腔調，共有 200 位語者，每位語者錄製半小時，總共錄製 100 小時的語料，錄製時同時使用六支麥克風以模擬各種情況，因此又可擴充為 600 小時。

二、 語者資訊：

共 200 個語者，其中有 92 位男性及 108 為女性，為了收集各地的腔調，語者來自全台灣，分佈如圖 1，年齡範圍為 18~73 歲，分佈如圖 2。

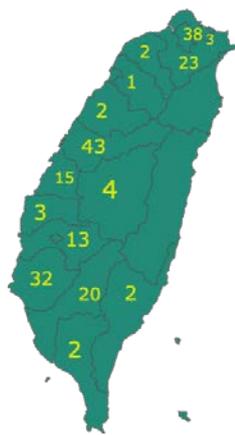


圖 1. 語者地區分布

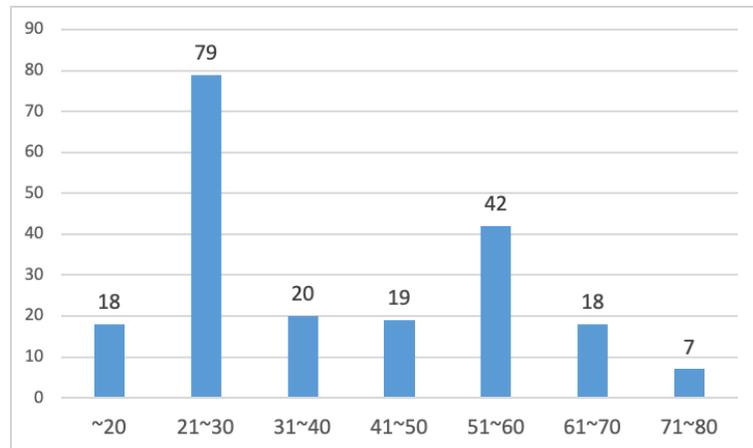


圖 2. 語者年齡分佈

三、 錄音方式及文本來源：

錄製環境為安靜密閉的空間，錄製時同時使用六隻麥克風，麥克風列表如下：

- ZOOM XYH-6 左聲道 (XYH-6-X)
- ZOOM XYH-6 右聲道 (XYH-6-Y)
- 電容式麥克風 (condenser)
- 領夾式麥克風 (lavalier)
- ios 系統手機錄音 (ios)
- Android 系統手機錄音 (android)

ZOOM XYH-6 左右聲道距離語者約一公尺，會含有些微環境音，電容式及領夾式麥克風距離語者約 5~10 公分，聲音最為乾淨清楚，兩種系統的手機則距離語者約 15~20 公分，聲音比擬為實際的手機錄音狀況。

錄音文本內容包括日常對話、簡短文章、及數字資訊（數字、地址、日期），文本來源有兩者：

- 李江卻台語文教基金會：共 50 個作者，每個作者約 6000 字，及日常對話的 14 個課程。
- 教育部：250 個文章，每個文章約 600 字。

四、 語料庫內容：

語料庫以語者及提示卡分成 volume1 & volume2 ，又以麥克風分成六個聲道，數據如下：

- TAT-Vol1

	Speakers	Sentences	Characters	Hours
Train	80	22,307	268,154	40.49
Eval	10	2,640	31,669	4.80
Test	10	2,752	31,806	5.09

- TAT-Vol2

	Speakers	Sentences	Characters	Hours
Train	80	23,981	285,113	42.87
Eval	10	3,085	30,963	5.56
Test	10	3,046	32,494	5.53

五、 音檔及文本格式：

取樣格式：mono、頻率為 16khz、16bits PCM

音檔格式：wav 檔

文本格式：json 檔，內容含音檔各項資訊，也包含多種文字輸出，像是漢羅台文及台羅拼音，範例如圖 3。

```
{
  "音檔長度": "6.0",
  "漢羅台文": "十四加二五是三十九",
  "台羅": "tsáp-sì ka li-gōo sī sann-tsáp-káu",
  "台羅數字調": "tsap8-si3 ka1 li7-goo7 si7 sann1-tsap8-kau2",
  "白話字": "cháp-sì ka li-gō sī sa"-cháp-káu",
  "字數": "9",
  "提示卡編號": "0012",
  "句編號": "1.10",
  "發音人": "IUF009",
  "性別": "女",
  "年齡": "20",
  "教育程度": "大學",
  "出生地": "高雄市湖內區",
  "現居地": "高雄市湖內區",
  "腔調": "高屏普通腔",
  "錄音環境": "安靜隔音室內",
  "提示卡切換速度": "快",
  "總錄音時間(分)": "70"
}
```

圖 3. 文本格式範例

TAT-TTS-M1

by 廖元甫 臺北科技大學

一、 簡介

TAT-TTS-M1，為臺北科技大學和李江確台語文教基金會以及意傳科技合作產製之語料庫，主要目標是要建置大規模台灣腔語料庫，作為研發人工智慧之基礎建設，尤其是針對基於深度學習之語音合成研究需巨量資料之要求。

二、 語料庫內容

語料庫為一位男性語者 M1，總時長約 10.4 小時，共 9625 個 wav 音檔。每個音檔均配對一個相同檔名之含有錄音內容以及語者資料的 json 文檔。

三、 處理程序

在無噪音影響之專業錄音室內依照台羅發音錄製音檔，經由意傳科技人工校正後產生語料庫。

四、 音檔格式

副檔名	Channels	Sample Rate	Precision	Sample Encoding
wav	2	192000	24-bit	24-bit Floating Point PCM

五、 語者資訊

性別	年齡	教育程度	出生地	現居地	腔調
男	34 歲	大學	台北市士林區	台北市士林區	偏漳州腔

TAT-TTS-M2

by 廖元甫 臺北科技大學

一、 簡介

TAT-TTS-M2，為臺北科技大學和李江確台語文教基金會以及意傳科技合作產製之語料庫，主要目標是要建置大規模台灣腔語料庫，作為研發人工智慧之基礎建設，尤其是針對基於深度學習之語音合成研究需巨量資料之要求。

二、 語料庫內容

語料庫為一位男性語者 M2，總長度約 10.1 小時，共 11532 個 wav 音檔。每個音檔均配對一個相同檔名之含有錄音內容以及語者資料的 json 文檔。

三、 處理程序

在無噪音影響之專業錄音室內依照台羅發音錄製音檔，經由意傳科技人工校正後產生語料庫。

四、 音檔格式

副檔名	Channels	Sample Rate	Precision	Sample Encoding
wav	2	192000	25-bit	32-bit Floating Point PCM

五、 語者資訊

性別	年齡	教育程度	出生地	現居地	腔調
男	55 歲	大學	新北市汐止區	新北市汐止區	泉州安溪腔

TAT-TTS-F1

by 廖元甫 臺北科技大學

一、 簡介

TAT-TTS-F1，為臺北科技大學和李江確台語文教基金會以及意傳科技合作產製之語料庫，主要目標是要建置大規模台灣腔語料庫，作為研發人工智慧之基礎建設，尤其是針對基於深度學習之語音合成研究需巨量資料之要求。

二、 語料庫內容

語料庫為一位女性語者 F1，總長度約 10 小時，共 12917 個 wav 音檔。每個音檔均配對一個相同檔名之含有錄音內容以及語者資料的 json 文檔。

三、 處理程序

在無噪音影響之專業錄音室內依照台羅發音錄製音檔，經由意傳科技人工校正後產生語料庫。

四、 音檔格式

副檔名	Channels	Sample Rate	Precision	Sample Encoding
wav	2	48000	24-bit	24-bit Signed Integer PCM

五、 語者資訊

性別	年齡	教育程度	出生地	現居地	腔調
女	52 歲	碩士	高雄市新興區	新北市新店區	漳州腔

TAT-TTS-F2

by 廖元甫 臺北科技大學

一、 簡介

TAT-TTS-F2，為臺北科技大學和李江確台語文教基金會以及意傳科技合作產製之語料庫，主要目標是要建置大規模台灣腔語料庫，作為研發人工智慧之基礎建設，尤其是針對基於深度學習之語音合成研究需巨量資料之要求。

二、 語料庫內容

語料庫為一位女性語者 F2，總長度約 10.1 小時，共 12422 個 wav 音檔。每個音檔均配對一個相同檔名之含有錄音內容以及語者資料的 json 文檔。

三、 處理程序

在無噪音影響之專業錄音室內依照台羅發音錄製音檔，經由意傳科技人工校正後產生語料庫。

四、 音檔格式

副檔名	Channels	Sample Rate	Precision	Sample Encoding
wav	2	48000	24-bit	24-bit Signed Integer PCM

五、 語者資訊

性別	年齡	教育程度	出生地	現居地	腔調
女	41 歲	碩士	台中市梧棲區	台北市中正區	泉州腔