

國立臺北大學資訊工程學系104學年度專題

我的專屬翻譯茑茑

PRJ-NTPUCSIE-104-010

成員：宋晉德、黃瀚勳

1. 摘要

對我們來說，「開口說話」是再簡單不過的事了！但對於某些身障朋友來說卻不容易，常常會因為疾病的影響無法順利與人進行交談，有時也會因為手部肌肉萎縮變形無法快速利用平板打字溝通。因此我們想利用科技產品語音播放的方式，以及導入機器學習的架構，來達成個人化的需求，以便在未來提供給更多需要幫助的人。

為了讓本次開發計畫的實行順利，我們有協助對象的真實案例，透過開發出來的應用軟體——只要在平板上小鍵盤輸入，就會有畫面及聲音顯示出來，以達成溝通的效果。且輔助小鍵盤輸入的內容，會以動態方式顯示，就可以達成快速溝通的效果。

本次計畫的開發，希望能真正幫助到身障人士，並且讓他們在平常生活中，也能像一般人一樣跟其他人溝通。

關鍵字：機器學習、個人化、平板小鍵盤、語音播放、輔助輸入。

2. 硬體設備與測試方式



Figure 1 實際測試畫面

硬體設備使用到兩個平板和一個平板架。輸入輔助按鍵的平板型號是Nexus 7。顯示輸入畫面的平板型號是Asus ME301t。平板架型號則是JOY Unite MNU107。

測試方法，會在兩個平板上都安裝開發出來的應用軟體，而且會把輸出平板透過平板架架設在協助對象輪椅旁邊的鐵管上。然後，會送出輸入平板的內容，並顯示內容在另一個平板畫面上，以及播放語音。如圖Figure 1，只是我們還沒把顯示畫面的平板架設在輪椅旁邊。

3. 軟體實作畫面

在本次計畫的軟體介面大致會有兩種模式：第一種是一般的輸入模式，第二種是可以選擇要播放之檔案的演講模式。



Figure 2 輔助輸入畫面

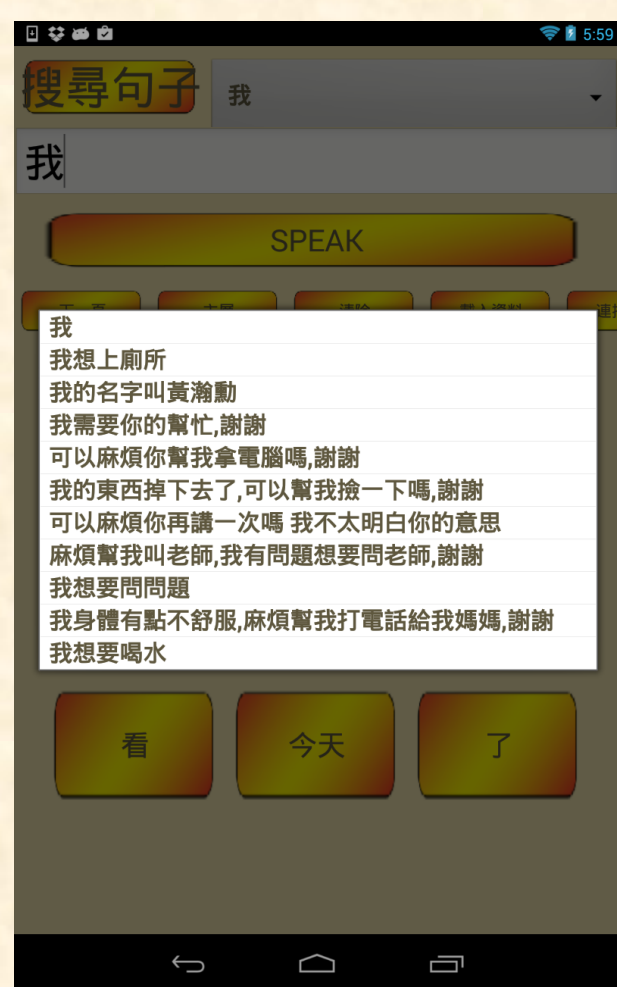


Figure 3 搜尋句子

進入到一般輸入模式，如圖Figure 2，會有兩種方法去提升我們的輸入速度：第一種是搜尋句子，第二種是去按底下九宮格的動態按鈕。

第一種搜尋句子，就是左上角的「搜尋句子」，需先輸入內容之後，再按左上角的搜尋句子去做更新，接著在右上角的搜尋框框，選擇想要的句子，如圖Figure 3。

第二種動態按鈕九宮格，過程就是當我點選「我」的按鈕之後，會把按鈕的內容去做關聯性的更新，如圖Figure 4。



Figure 4 點選「我」之後



Figure 5 演講模式

另外，如果在演講模式下，圖Figure 5，只要把想要演講內容的檔案放在應用軟體的相關目錄下，並且排列好要演講的檔案。接著，便可以從應用軟體畫面去點選要語音播放的檔案，並且顯示畫面至另一個裝置上。當演講完第一份檔案，也需要再選取下一個要播放的檔案，便可以繼續播放，以此類推。當演講完之後，也可以切回一般模式，跟台下的人做互動。

4. 測試環境與效能評估

測試環境設定：I、協助對象同時使用兩隻手，且兩手都只有使用食指輸入。II、測五種情境，加總起來。III、學習資料：以協助對象日常生活使用為主IV、情境中句子，只要意思相同便可以，如「吃飽了嗎」，打成「吃飽沒」。以下為測試數據的情境內容：

情境一：吃飽了嗎
中午要吃哪裡
那家好吃嗎
一起去啊

情境二：下禮拜要專題展了耶
你準備好了嗎
我好緊張喔
一起加油吧

情境三：嗨，好久不見
最近都在幹嘛啊

情境四：我肚子痛
我想上廁所
可以幫我打電話給我媽

情境五：這題你會嗎
可以做一次
我還是不太了解
叫老師

學習資料數量 (句子可以重複) (句)	有學習測試數據 (句, 字數, 完整句, 不完整句, 純輸入句, 時間)	未學習測試數據 (句, 字數, 時間)
500句	(4, 20, 1, 2, 1, 118), (4, 19, 2, 1, 1, 106), (2, 12, 0, 0, 2, 76), (3, 18, 3, 0, 0, 66), (3, 40, 1, 1, 1, 200)	(4, 19, 124), (4, 27, 180), (2, 14, 95), (3, 20, 130), (3, 24, 159)
1000句	(4, 19, 2, 1, 1, 107), (4, 21, 1, 2, 1, 105), (2, 12, 1, 0, 1, 61), (3, 19, 2, 0, 1, 89), (3, 42, 2, 0, 1, 103)	(4, 19, 128), (4, 27, 185), (2, 14, 87), (3, 20, 123), (3, 24, 163)
5000句	(4, 19, 2, 1, 1, 106), (4, 21, 3, 0, 1, 88), (2, 12, 1, 0, 1, 53), (3, 18, 3, 0, 0, 68), (3, 42, 2, 0, 1, 99)	(4, 19, 126), (4, 27, 183), (2, 14, 89), (3, 20, 126), (3, 24, 162)

Figure 6 五種情境下的輸入狀況與學習資料比較

Figure 6中，分別對輸入裝置做500、1000、5000句的學習；並且測試兩種不同情況，第一種會學習測試數據的句子，第二種則是不會學習測試數據的句子，想加以探討事先學習句子，是否會對整體的測試速度提升。並且，在兩種情況下，探討輔助輸入是否會提升整體的速度。另外，在第一種情況下，想了解輔助輸入句子的完整度，與速度之間的關係。

為了方便圖表建立，所以把Figure 6的表格數值加總。如Figure 7。另外，未學習測試數據，會因為資料量不夠大，導致於必須使用平板內建的小鍵盤來輸入。

學習資料	有學習測試數據	未學習測試數據
500句	(16, 109, 7, 4, 5, 566)	(16, 104, 688)
1000句	(16, 113, 8, 3, 5, 465)	(16, 104, 686)
5000句	(16, 112, 11, 1, 4, 414)	(16, 104, 686)

Figure 7 數值加總後的學習比較

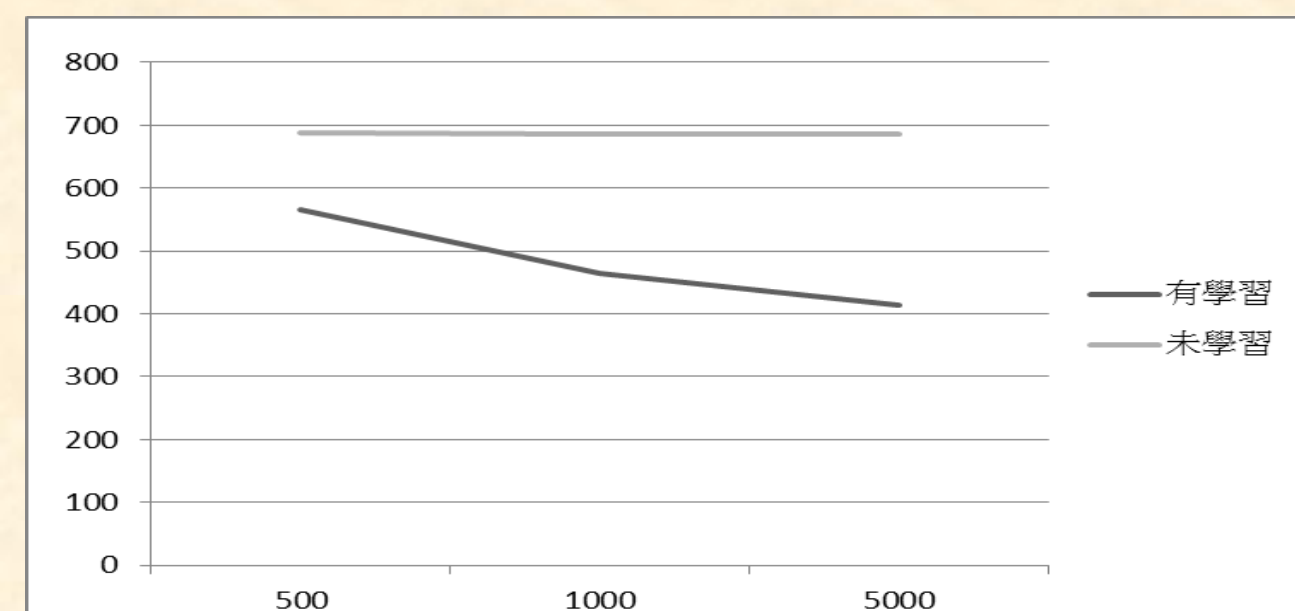


Figure 8 有學習和未學習之輸入速度(Y)與學習句子資料量(X)的折線圖

如圖Figure 8，可以了解有先學習數據會比沒有學習數據的在輔助輸入速度上面，有明顯的提升。另外，在有學習數據的狀況下，會發現當學習句子的資料量越大時，輸入速度也會提升；但是，在沒有學習數據的狀況下，學習句子的資料量越大時，輸入速度不會有明顯的提升。也就是說，本次計畫的應用軟體，必須學習到一定的量之後——當要輸入某個句子時，正好是之前有學習過的句子；而且，輸入時，跟平板的小鍵盤比，速度會比較快。

比較完有學習和未學習之間的關係之後，我們想了解，有學習測試數據是如何提升輸入的速度；我們會從完整句、不完整句、純輸入句的實驗數據來討論。

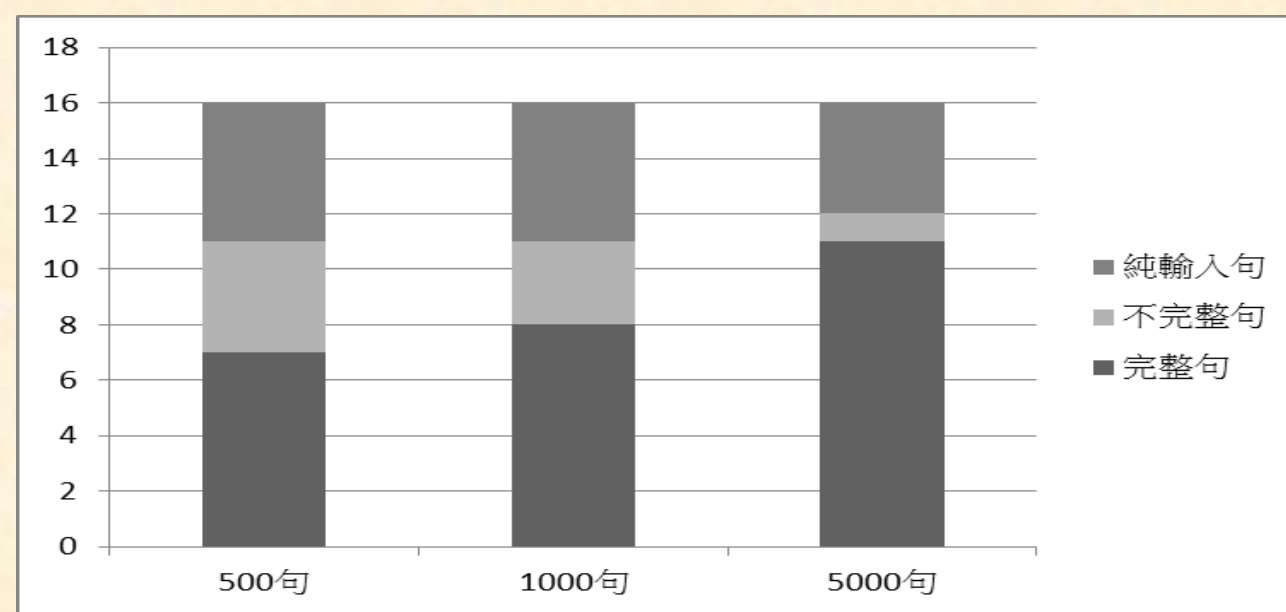


Figure 9 有學習之不同學習句子資料量(X)的輸出句子結果(Y)之完整性比較長條圖

如圖Figure 9可以觀察到，學習句子的資料量越大時，在純輸入句子和不完整句上面的次數會比較少，相對的它在完整句的次數會比較多。也就是說，學習句子的資料量越大時，會使完整句次數增加，進而造成輔助輸入之整體性的速度增加。

5. 軟體優點與未來發展

本計畫的應用軟體，透過機器學習的方式，可以慢慢地提升輔助輸入的速度。並且在軟體的一般模式下，協助對象在未來可以透過平板跟別人交談；另外，在演講模式的下，也可以站在演講台上，跟大家分享自己的看法。

在未來，會想要增加腦波感測、聲音感測，以及情境感知；希望能增加輸入按鈕的品質，進而提升輔助輸入時的速度。腦波感測，會想要使用裡面眨眼睛、專注度和冥想度的功能。聲音感測，由於協助對象在測試GOOGLE聲音感測時，輸出結果常常出錯，所以會想改用有發出聲音和沒發出聲音來做為控制方式。情境感知，則是把整個架構本體化，並且建立更完善的資料庫系統，使在輔助輸入時，能根據我現在的環境去做資料的調整〈地點、溫度、腦波感測的資料…等等〉。



國立臺北大學
資訊工程學系

Department of Computer Science and Information Engineering, NTPU