

[◀ 返回前頁](#) | [列印](#)

2021年11月29日

StartupBeat 創科鬥室

AI語音App助舌癌婦「說話」中大應港青求援「留住媽媽聲線」



[放大圖片](#)

親人簡單一句問候，如春暉沁入心扉。九十後港青潘楚騫的母親Jody舌癌復發，須接受手術切除舌頭及聲帶組織。他為了留住媽媽聲線，在連登討論區（LIHKG）發帖求助，未幾被香港中文大學電子工程學系教授李丹的團隊發現。團隊利用人工智能（AI）語音合成技術，把Jody的聲音還原，又夥拍本地科企開發專用應用程式（App），協助手術後失聲的Jody，透過手機以自己的聲音跟他人溝通。

輸入文字轉化為原聲播放

李丹的團隊為Jody所研發的手機程式，主要讓Jody把輸入的文字，轉化成她本人聲音播出。Jody在訪問期間，親身示範使用程式，記者比較程式的播放效果，以及Jody手術前的錄音，發覺兩者相似度頗高，語音質素接近即場錄音，沒有機械發聲的感覺。

不過，整套技術亦有其限制，例如只支援繁體中文字，再轉為廣東話語音，若Jody輸入簡體中文或英文字，整句句子只會以機械聲播放。負責訓練AI模型的李丹高足、中大電子工程學系博士生譚達新指出，中英夾雜向來是港人說話方式，「但混合語種是語音合成研究的困難之處，先做好一個足夠讓Jody應付基本日常生活的版本比較重要。」

年約60歲的Jody早年患上舌癌，今年因癌症復發而需要切除舌頭及聲帶大部分組織。潘楚騫憶述，當母親知道自己將會喪失說話能力後，曾對他說：「你一定要記住我把聲呀。」

如何記住甚至留住一個人的聲音，潘楚騫苦無頭緒，他的女朋友於是代他在連登討論區發帖，引來一眾巴打打熱烈討論，網民留言足足有20頁，更吸引到博士生馬敬恆的注意。

在中大攻讀電子工程學系的馬敬恆，把帖子發給老師李丹，「我很熟悉教授的性格，他一定願意幫手。」果然，李丹獲悉潘楚騫的處境後，估計團隊有辦法幫忙，於是吩咐馬敬恆在連登留下聯絡方法，邀請潘楚騫主動接洽。

當時已是今年6月中，距離Jody月底接受手術不足兩星期。李丹隨即安排Jody前往中大校園內的錄音室，完成約15小時的錄音，以採集數據訓練AI模型。團隊把錄音去蕪存菁（如刪去停頓位）後，最終可用以訓練AI模型的有效數據，只約7至8小時。

錄音內容主要包括日常說話（例如你會唔會參加下星期嘅派對）、講兒童故事和誦讀聖經，「不可求其取一份報紙叫Jody照讀，因為那不像她平時跟人說話。」

李丹續說：「講兒童故事的好處是語氣有點變化，我們叫Jody盡量用回自己聲線去講，別扮小朋友聲；至於讀聖經，因為Jody平常有返教會，讀聖經是她生活一部分。」

採集數據後，譚達新與另外幾位博士生編寫代碼，並設計和訓練AI模型。開始訓練約兩天後，電腦逐漸懂得以Jody的聲音播放文字，「但很多時候，一句句子的轉化效果理想，不代表所有句子都一樣，因此要反覆測試。」自九十年代已從事語音技術研究的李丹解釋，廣東話的多音字很多，訓練AI模型時須留意，「以『門』字為例，大『門』及走後『門』的讀音已不同。」

廣東話多音字增加難度

完成AI模型訓練後，接下來李丹找來之前曾一同研發聽力測試方案的本地科企尚普（Ximplar）董事長李沛鏜，以及該公司的軟件工程師雷淇鈞，由他們花約兩星期製成首個版本的手機程式。李沛鏜坦言，程式有不少地方要改進，只能邊用邊改。

例如現時程式容許Jody把AI合成的語音，分享至WhatsApp、Telegram等通訊軟件，這功能是根據潘楚騫的意見而增設的。李丹則希望進一步運用AI及GPS（全球衛星定位）技術，按照Jody身處位置作出句子建議。此外，目前系統有部分讀音，以至標點符號停頓位置未夠準確，仍需逐步改善。

身為用家的Jody透過程式表示，「自己平時好鍾意講嘢……非常感謝大家為我製作呢個程式，我感覺真係留低咗我把聲。」然而她指出，輸入文字後，資料載入（loading）速度「唔夠快」，其丈夫潘先生補充，目前Jody要打完一句句子後，語音才會播放，如果可以一邊打字，一邊播放語音，效果會更理想。李沛鏜及雷淇鈞相信，若將AI語音合成模型直接存放在手機，資料載入速度有望加快，惟需要不時手動清除資料及快取，以免手機容量被擠爆。

採訪、撰文：陳子健

