行政院客家委員會補助大學校院發展客家學術研究

線上客語有聲詞典與合成系統之研究與建置(二)

Study and Implementation of Hakka's Phonetic Dictionary and Text-to-Speech System on Internet

成果報告書

(精要版)



建置之線上系統網址: http://203.64.183.226/public2/hakka-index.htm

計劃主持人: 黃豐隆 博士國立聯合大學資工系 副教授

中華民國 99年 11月 25日

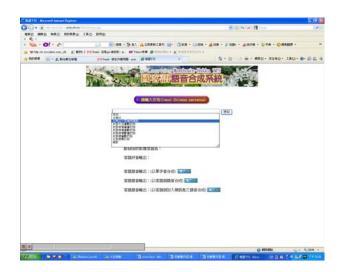
本報告係接受行政院客家事務委員會獎助完成

特此致謝

2010 年度計劃之成果:



線上四縣腔有聲客語詞彙系統(Hakka Audio Dictionary): http://203.64.183.226/public2/hakka-edu/index1.html



線上四縣腔客家語音合成系統(HTTS): http://speechlab.cs.nchu.edu.tw/hakkatts/index.php

一、研究計畫摘要:

電腦輔助之語言數位學習系統(Computer Aided Language Learning, CALL)所涉及的技術相當廣泛,如詞典建置,主機架設、語音錄製、資料庫建置與自然語言處理等,並非短期可以全部完成的,因此,在第一年所建構完成之基礎上,繼續設入相關之研究人力,使本計劃沿續目前之成果,並完成更具實用之電腦輔助語學學習之客語數位學習平台,對廣大學習者與客語國際化將有很大的助益。

本計劃之研究語言延續2009年之計劃,以客語苗栗四縣腔為主,再錄製常用之客語雙字與三字詞合計1500個,經錄製並完成每一音之切音與能量正規化處理。而客語詞彙詞目再擴充7000筆,總數達15000筆。線上客語語音合成系統與有聲詞典即以此為基準,處理完成後可供線上作客語語音之合成與播放,具有良好之擴充性與較低人力之維護性,對日後之系統擴充與維護均比現有系統更具優勢。

二、本計劃完成之研究內容

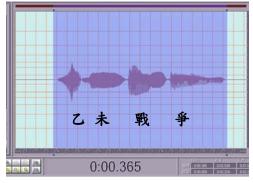
本計劃之參之成員含客語文化、客語詞典、客語語音與資訊科技之領域學者與專家,每月定期相互討論,依計畫之項目與進度研究、分工與考核,進行順利。

2.1 客語四縣音錄音(以客語詞彙的詞語音為主)

以客語之四縣腔為主,由具有四縣腔聽說優良之專業人士錄製本計畫所需之語音單元,已完成錄音與後續切音與訊號等處理,並經過測試與檢驗,以確保語音品質符合計畫所需。參見圖 1

目前已錄製完成四縣腔所有語音音節,約有 2500 個(含所有聲調之語音組合)外。 進一步以客語詞彙(word)為單位之語音亦已錄製約 3600 個,我們依照客語訒證初級教 材之詞彙為主。經錄製並完成每一音之切音與正規化處理,有聲詞典與語音合成系統依 以此為基準,處理完成後可供線上客語有聲詞典之合成與語音播放。

客語語音之基本「合成單元」錄製條件為:取樣頻率 SR 為 44100Hz, 單聲道 (Mono), 音訊大小為 16 bits。經語音處理之後,為考量資料傳輸之流量,故取樣頻率 SR 改為 11025Hz,其線上輸出之語音品質不受影響,應可維持高音質效果。



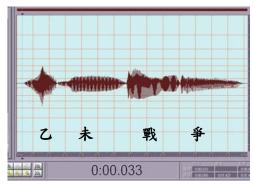


圖 1 左:以 Audition 語音軟體錄製與後製,客語詞彙「乙未戰爭」切音處理前波形。 圖 1 右:切音完成後之波形,此客語詞彙即為語音庫中之合成單元。

2.2 線上客語語彙詞典建置

客語語彙詞典之內容,除原來已建置近8000筆客語詞目,包含詞彙之拼音、詞義、相關相語詞與例句說明等,本計劃再增加近24000個詞目資料,格式同前。資料庫之建置,需考量未來之擴充與線上查詢之穩定性與連線速度,且提供使用者由線上使用中文詞彙或客家詞彙作輸入,以便符合不同使用之需求。詞彙為置的工作雖不容易,本計畫之顧問群有長期從事客語詞典建置之成員,定期、經由e-mail 討論提出系統之修正意見,因此,由位顧問群協助客語語彙之建置與修正。

此外,本計畫亦已完成5687字之單字檔,含四縣客語拼音與中文等欄位。圖2 為資料庫管理介面 phpMvAdmin 軟體,所顯示出之資料庫內容。



圖2左:目前資料庫中,四縣客語單字之拼音檔,有5687筆。

圖2 右: 目前資料庫中,四縣客語詞彙檔共有32000筆詞目。

2.3 建置客語有聲詞典與合成系統

依據上述之語音單元與客語詞彙資訊,以自然語言處理技術為基礎,逐步完成客語有聲詞典之建置。具有後端管理者管理之介面,以便作系統之功能設定與客語資料修正與更新等工作。有聲詞典之客語詞彙,並不只限於目前客委會客語檢定之內容, 包含共有約 32000 筆詞目,並有基本之單字拼音檔 5687 筆,可作為單音之語音合成 處理使用。

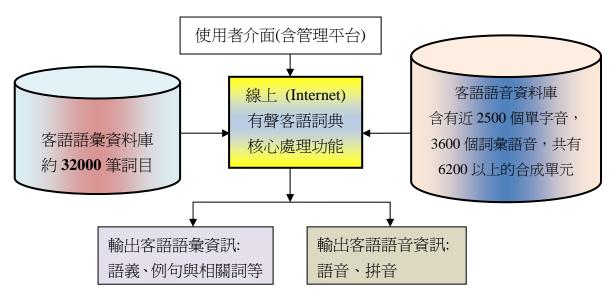


圖3: 有聲詞典之功能架構

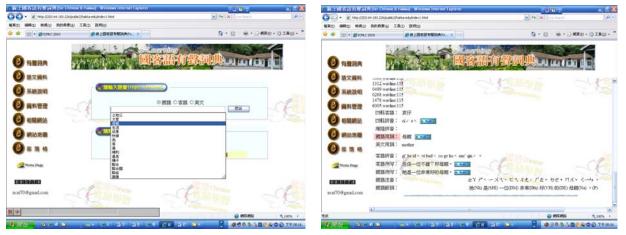


圖4 左:線上客語有聲詞典首頁,查詢輸入「母視」詞彙。 圖4 右:輸出的項目含有:客語四縣詞彙資訊與語音輸出。

四縣客語: 哀仔

四縣拼音:oi/e\

海陸拼音:(暫無)

國語用詞: 母親 ▶ PLAY

英文用詞:Mother

客語例句:但係一位不錯个好母親。

國語例句:她是一位非常好的母親。 **上 PLA**

國語注音: ムY アヽ ー メヘヽ 亡へ イオノ 厂幺V わさ・ 口メV くーり 。

國語斷詞:她(Nh) 是(SHI) 一位(DM) 非常(Dfa) 好(VH) 的(DE) 母親(Na) 。(P)

註:上列含有 圖示,均可輸出合成的四縣客語語音。

圖5: 母親之查詢畫面。

本計劃所建置「線上國客語有聲詞典」系統,網址如下:

http://203.64.183.226/public2/hakka-edu/index1.html

2.4 韻律訊息之預測(Prediction)

韻律訊息中的「停頓(break)」問題,指:在文句產生語音時在適當的地方加入合適的靜音(Silence),使語音更自然、更合乎人們說話的習慣,這一點對電腦處理語音合成而言,是十分重要的。

本計劃所採用的方法係以**資料探勘**(Data Mining)理論為基礎,以機器學習的策略來訓練模組,以事先由專家標註的語料,以學習文句中停頓的類別,作為加入停頓的參考依據,我們將文句中加入的停頓分為四大類,參見表1所示,依文句的語彙與結構加入不同的停頓時長。例如:五大標點符號標註為「BIG」,將加入0.65秒靜音,在韻律段之間將加入「Major」,在韻律詞之間加入「Minor」,最後一種為「no」 則表示不必加入任何的停頓。

表 1: Breaks types for Hakka's TTS.

Situation	Break type	duration (sec)
between five punctuation " °!?; ,"	BIG	0.65
between prosodic phrase (five punctuation)	Major	0.45
between prosodic word (in prosodic phrase)	Minor	0.25
between word (in prosodic word)	no	0

Currently, the most popular techniques in data mining are as followings:

- A) Artificial neural networks
- B) Decision trees
- C) Genetic algorithms
- D) Nearest neighbor method
- E) Rule induction

Predicting the types of break is regarded as classification issue. Bayesian Network (BN) adopted the decision tree to predict the proper categories of break. In the paper, we modified several detail and two methods, Bayesian networks and Classification And Regression Trees (CART), are included into the prediction to promote the better performance.

Let $V = \{v_1, v_2, ..., v_n\}$, n >= 1, be a set of variables. A Bayesian network BN over a set of variables V is a network structure BNS which is organized as a directed acyclic graph (DAG) over V, and a set of probability tables:

$$BNP = \{p(v_i|pa(v_i))v_i \in V\}$$
(1)

where $pa(v_i)$ is the set of parents of v_i in BNS.

A BNS represents probability distributions:

$$P(V) = \prod_{v_i \in V} p(v_i | pa(v_i)). \tag{2}$$

To employ a Bayesian network as a classifier, one simply calculates $\operatorname{argmax}_y P(y|x)$ using the distribution P(V) represented by the Bayesian network. Since all variables in x are known, we do use Eq. (2) for calculation of all class values of variables.

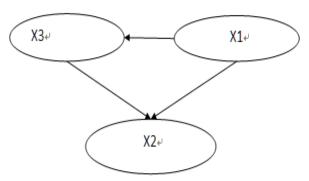


圖6: Network structure of BN with 3 states.

The features with respect to the breaks for target word include:

- Part-of-Speech of target and neighbor's word
- Length of target and neighbor's word
- Location of words with respect to the target word
- Number of words preceding target word and in sentences.
- Combination of features above.

聽覺測試(Comprehension)

我們測試 TTS 合成語音的品質,目前雖無客觀的標準方法來評斷不同合成語音的優劣,我們採用較常採用測量 MOS(Mean Opinion Score)十級分等級來評量合成自然度。總共有 14 人參與測試,其中 12 人是研究生,2 人是大學生。

測試方法是播放語音檔,請人當受測者,聽完音檔後給予評分,分數分五個等級:1分:極差(unsatisfactory),4分:差(poor),6分:普通(fair),8分:好(good),10分:非常好(excellent)。受測者根據自己聽到的感覺,打上分數。測試內容共有10句客語文句,分別播放有加入停頓與沒加入停頓的合成結果音檔。我們根據這個實驗來佐證加入適當的停頓有助於改善合成語音的自然度。結果如圖7所示。

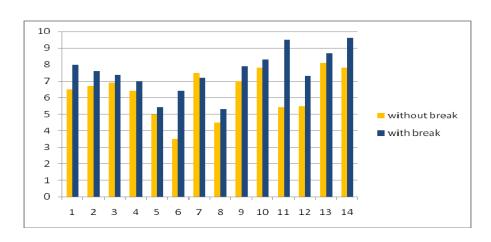


圖7:自然度測試結果

圖7中黃色(在黑白印刷中為淺色)為沒加入停頓的合成語音測試平均分數,藍色(在黑白印刷中較深色)為加入停頓的合成語音測試平均分數。X軸為受測者編號,Y軸為分數。

沒加入停頓的測試平均分數只有 6.32 分,加入停頓後測試平均分數為 7.54 分。 從此結果可看出,平均滿易度提升達 1.22 分,故可以得到結論:採用本計劃的方法, 加入停頓有助於改善合成語音的自然度。

2.6 線上系統建置完成

本系統建置之開發環境,相關軟體說明如下;

- 作業環境:Windows XP SP2,
- Web server:安裝 Apache Web Server v2.2.4,
- 資料庫系統:使用 MySQL Database v5.0.45,
- 資料庫介面管理:為 phpMyAdmin v2.10.2,
- Web Programming: 採用 PHP5。
- FTP server: 採用 Filezilla Server 版。

線上有聲詞典、客語語音合成系統架構如圖 8 所示。

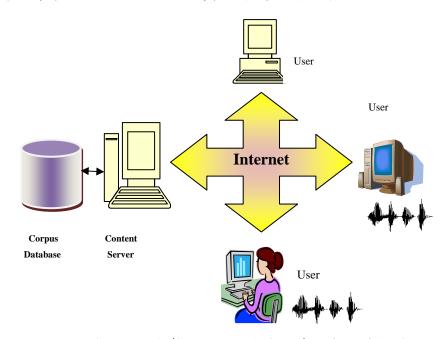


圖8: 線上客語有聲詞典、語音合成系統(HTTS)架構。

三、結語

本計劃研究語言延續2009年之計劃,以客語苗栗四縣腔為主,並再錄製常用之客語雙字與三字詞合計1500個,經錄製並完成每一音之切音與能量正規化處理。而客語詞彙詞目總數達15000筆。此外,我們採用自動學習、訓練擷取客語之停頓資訊,以機器學習的策略來訓練模組,以資料探勘(Data Mining)理論為基礎,事先由專家標註的語料學習文句中停頓的類別,作為加入停頓的參考依據,我們將文句中加入的停頓分為四大類,依文句的語彙與結構加入不同的停頓時長,以提升合成語音輸出之韻律性。

線上客語語音合成系統與有聲詞典即以此為基準,處理完成後可供線上作客語語音 之合成與播放,具有良好之擴充性與較低人力之維護性,對日後之系統擴充與維護均比現 有系統更具優勢。

此系統的功能可提供準備客委會客語認證考試者、一般人士學習、認識客家文化者,學生亦可有效學習客語之聲母與母之音節發音與拼音系統,國際人士亦能以英文詞彙查詢,瞭解國語與客語之發音與語彙、語義文字,進而認識我國文化。本計劃之線上客語語音合成系統功能具有客語文轉語音之語音合成平台,將可提供相關客家語譯使用,如客語104查號台,客語有聲書、客語導覽系統、線上客語發音與翻譯機等應用,用途廣泛。

参考文獻

- 客委會出版,客語能力認證基本詞彙-中級、中高級暨語料選粹四縣版上冊,2007。
- 客委會出版,客語能力認證基本詞彙-中級、中高級暨語料選粹四縣版下冊,2007。
- 台中市政府,歡喜說客話-《臺中市鄉土語言(客語)補充教材》,2006。
- 黃永達,《台灣客家俚諺語詞典:祖先的智慧》,2006。
- 古國順,何石松,劉醇鑫,《客語發音學(二版)》,2004。
- 何石松、劉醇鑫,《現代客語詞彙彙編續篇》, 2004。
- 陳雲錦,《苗栗客話圖解》,2004。
- 龔萬灶,《阿啾箭的故鄉》,2004。
- 何石松,《客家謎語(令子)欣賞》,2003。
- 鍾榮富,《台灣客家語音導論》,五南圖書出版公司,2004。
- 龔萬灶,《客話實用手冊》,國家圖書館出版,2003。
- 中原週刊社,《客家話破音字典》,中原客家文化研究會編,1995。
- 羅肇錦,《台灣的客家話》,臺原出版社,1990。
- 楊政男、龔萬灶、徐清明等四人,客語字音詞典,臺灣書店,1988。
- S.H. Hwang, S.H. Chen, and Y.R. Wang, A Mandarin Text-to-Speech system, in Proc. ICSLP-96, pp.1421-1424, Oct 1996.
- L.S.Lee, C.Y.Tseng, and C.J.Hesih, Improved tone concatenation rules in a formant-based Chinese text-to-speech system, IEEE Trans. Speech and Audio processing, Vol.1, No.3, pp.287-294, July 1993.
- S.H.Chen, S.G.Chang, and S.M.Lee, A statistical model based fundamental frequency synthesizer for Mandarin speech, J.Acoust. Soc. Am., 92(1), pp.114-120, July 1992.
- L.S.Lee, C.Y.Tseng, and M. Ouh-Young, The synthesis rules in Chinese text-to-speech systemIEEE Trans. Acoust, Speech, Signal Processing, vol. 37, no. 9, p1309-1319, Sep. 1989.
- 林東毅,客語文句翻語音系統之實作,國立交通大學電信工程學系碩士論文,2007。
- 賴惠玲,客語語法研究議題的開發:以語料庫為本,客委會補助計畫,2007。
- 吳瑾瑋,傳統與現代客語歌之韻律結構的研究,客委會補助計畫,2007。
- 呂嵩雁,客語《陸豐方言》的百年語言演變析探,客委會補助計畫,2007。
- 陳嘉甄,客語生活學校計畫對客語推展之績效指標建置及現況調查研究,客委會補助 計畫,2007。
- 謝傳崇,客家語實施多元評量之研究計劃-以桃竹苗四縣市為例,客委會補助計畫, 2005。
- 羅鳳珠,台灣客語語言典藏的現況與資源整合,語言政策的多元文化思考系列研討會,2005。
- 張麗君,國客雙語幼兒語言能力與創造力之關係,客委會補助計畫,2004。
- 余伯泉、朱阿莉,台客語教材教法與政策研究,北縣教育 42:,pp. 60-64,2002。

- 余秀敏,苗栗客家話音韻研究,輔大語言所碩士論文,1984。
- 國立聯合大學全球客家研究中心 http://203.64.173.77/%7Ehakkacenter/ch/index.html
- 國立中央大學客家語文研究所 http://140.115.170.1/Hakkalanguage
- 國立成功大學客家研究中心 http://www.ncku.edu.tw/~hakkarc/index1.php
- 臺北市政府客家事務委員 http://www.hakka.taipei.gov.tw/index2.asp
- 台中市客語教學資源中心(台中市四維國小) http://www.swes.tc.edu.tw/~hakka/
- 客委會-語言學習館: http://www.hakka.gov.tw/np.asp?ctNode=1724&mp=1722&ps=