

以知識表徵方法建構台語聲調群剖析器

A Knowledge Representation Method to Implement A Taiwanese Tone Group Parser

張佑竹 Yu-Chu Chang
國立中正大學語言學研究所
Institute of Linguistics
National Chung Cheng University
poirot david@yahoo.com.tw

摘要

聲調群剖析器是台閩語語音輸出系統的主要元件之一。本文提出聲調管轄假說，主張先將句內語詞定調，亦能決定台閩語聲調群分界的觀點，並以聲調群剖析器實作加以驗證。除了敘述如何應用預設調型、預設詞類和模式三種標記符號，將語言知識和經驗轉換為知識庫，並說明經由推論引擎與知識庫的連結，完成語詞定調的運作過程。目前內部測試平均變調正確率為 98.5%。外部測試平均變調正確率為 94%。本研究的實驗數據也顯示一個重要的線索：符號系統標記比規則推論對變調正確性有相對較高的貢獻率。

Abstract

A tone group parser could be one of the most important components of the Taiwanese text-to-speech system. In this paper, we offered the hypothesis of tonal government to emphasis the idea that if the allotone selection can be made for each word in a sentence then the tone groups will be separated within the sentence and supported our viewpoint with the implementation of a Taiwanese tone group parser. In addition to the description of using the symbol system to convert language expertise and heuristic knowledge into a knowledge base to cope with a frame-based corpus and a tone sandhi processor, the procedure of connecting the inference engine and the knowledge

base to make allotone selection was also discussed. In the current version of the tone group parser, the average accuracy of inside test is 98.5%. The average accuracy of outside test is 94%. The experiment data of the study also reveals an important clue: the marking of the symbol system makes a higher contribution rate to the tone sandhi accuracy than the rule inference.

關鍵詞：台灣話，變調，聲調群剖析器，知識表徵，模擬

Keywords: Taiwanese, Tone Sandhi, Tone Group Parser, Knowledge Representation, Simulation

致謝

本論文之研究承蒙中正大學語言學研究所麥傑教授提供諸多建議，謹此致謝。

參考文獻

- [1] K. Liim, "Medical Education and Research in Taiwanese Language Since 1990," presented at the *Symposium on Medical Taiwanese*, Kaohsiung, Taiwan, 2004, pp. 10-14.
- [2] 田村志津枝，*初めて台湾語をパソコンに喋らせた男*，東京：現代書館，2010。
- [3] B. Chiu, "The Phonetic Structure and Tone Behavior in Hagu (commonly known as the Amoy Dialect) and their Relation to Certain Questions in Chinese Linguistics," *T'oung Pao*, vol. 28, pp. 245-342, 1931.
- [4] 王育德，台灣語の聲調，頁 3-11，*中國語學研究會中國語學 41 號*，1955。
- [5] Y. C. Chang, "An introduction to Taiwanese Speech Notepad," Internet:
<https://archive.org/details/TaiwaneseSpeechNotepadenglishVersion> [Oct. 21, 2009].

- [6] E. Selkirk, *Phonology and syntax: the relationship between sound and structure*, Cambridge, Mass: MIT Press, 1986.
- [7] J. Tsay, "Bootstrapping into Taiwanese Tone Sandhi," in *Chinese Languages and Linguistics V: Interactions in Language*. vol. 5, ed. Taipei: Academia Sinica, 1999, pp. 311-333.
- [8] N. Geschwind. Ed., "Specializations of the human brain," *Scientific American: Language, Writing, and the Computer*. New York: W. H. Freeman and Company, 1985.
- [9] P. D. Eimas. Ed., "The perception of speech in early infancy," *Scientific American: Language, Writing, and the Computer*. New York: W. H. Freeman and Company, 1985.
- [10] M. Minsky, *A framework for representing knowledge*, New York: McGraw-Hill, 1975.
- [11] T. Y. Chang, *A multimedia-based bilingual instructional system using an expert system shell*, M.S. thesis, University of Central Missouri, Warrensburg, MO, 1992.
- [12] R. Cheng, "Tone sandhi in Taiwanese," *Linguistics*, vol. 41, pp. 19-42, 1968.
- [13] M. Y. Chen, "The syntax of Xiamen tone sandhi," *Phonology Yearbook*, vol. 4, 1987, pp. 109-149.
- [14] J. W. Lin, "Lexical government and tone sandhi formation in Xiamen Chinese," *Phonology*, vol. 11, pp. 237-275, 1994.
- [15] J. Tsay, J. Myers and X. J. Chen, "Tone sandhi as evidence for segmentation in Taiwanese," *Child Language Research Forum*, vol. 30, pp. 211-218, 2000.
- [16] H. H. Pan, "Prosodic hierarchy and nasalization in Taiwanese," in *Proceeding of the 15th ICPHS*, 2003, pp. 575-578.
- [17] M. S. Liang, R. C. Yang, Y. C. Chiang, D. C. Lyu, and R. Y. Lyu, "A Taiwanese text-to-speech system with applications to language learning," in *Proceedings of the 4th IEEE International Conference on Advanced Learning Technologies*, 2004,

pp. 91-95.

- [18] U. G. Iunn, K. G. Lau, H. G. Tan-Tenn, S. A. Lee, and C. Y. Kao, "Modeling Taiwanese Southern-Min Tone Sandhi Using Rule-Based Methods," *International Journal of Computational Linguistics and Chinese Language Processing*, vol. 12, no. 4, pp. 349-370, 2007.