

# 章節目錄

摘 要.....	I
Abstract.....	II
致 謝.....	V
章節目錄.....	VI
圖表目錄.....	VIII
第一章 緒論.....	- 1 -
1.1 研究主題及動機.....	- 1 -
1.2 中文斷詞系統簡介.....	- 2 -
1.3 章節概要.....	- 3 -
第二章 連讀變調.....	- 4 -
2.1 五度標記法.....	- 4 -
2.2 連讀變調.....	- 7 -
2.2.1 上聲變調.....	- 7 -
2.2.2 快速的陽平變調.....	- 9 -
2.3 「一、不」變調.....	- 9 -
2.3.1 「一」的變調.....	- 9 -
2.3.2 「不」的變調.....	- 11 -
2.4 疑問句尾升調.....	- 11 -
第三章 中文語音合成的斷詞系統.....	- 14 -
3.1 資料庫介紹.....	- 14 -
3.1.1 詞庫與斷詞系統的關連.....	- 15 -
3.1.2 清大多媒體資訊檢索實驗室語文資料庫(MIR DB) .....	- 15 -
3.1.3 中研院漢語平衡語料庫(Sinica Corpus 3.0) .....	- 16 -
3.2 斷詞單元.....	- 18 -
3.2.1 長詞優先法.....	- 19 -
3.2.2 動態規劃演算法.....	- 21 -
3.3 構詞單元.....	- 23 -
3.3.1 定量複合詞構詞.....	- 23 -
3.3.2 疊詞構詞.....	- 25 -
3.3.3 姓名構詞.....	- 26 -
3.4 詞庫擴增.....	- 27 -
3.4.1 新聞資料庫簡介.....	- 27 -
3.4.2 姓名及一般詞資料庫的增加.....	- 27 -
第四章 實驗結果與分析.....	- 28 -

4.1 實驗資料庫及效能定義.....	28 -
4.2 長詞優先與動態規劃演算法的斷詞單元.....	29 -
4.3 定量複合詞的構詞單元.....	30 -
4.4 疊詞的構詞單元.....	31 -
4.5 姓名的構詞單元.....	32 -
4.6 語音合成用之斷詞系統整體效能分析.....	33 -
4.7 實驗錯誤分析.....	35 -
第五章 結論與未來展望.....	37 -
參考文獻.....	39 -
附錄一 定詞及量詞.....	41 -
附錄二 常用百家姓.....	44 -



# 圖表目錄

圖 1-1 斷詞系統基本架構及斷詞流程示意圖.....	2 -
圖 2-1 五度標記法.....	6 -
圖 2-2 五度標記法的疑問句尾升調.....	12 -
圖 3-1 斷詞單元中詞庫與斷詞演算法示意圖.....	15 -
圖 3-2 長詞優先法示意圖.....	20 -
圖 3-3 動態規劃演算法示意圖一.....	22 -
圖 3-4 動態規劃演算法示意圖二.....	22 -
表 2-1 國語聲調表.....	6 -
表 2-2 疑問句尾升調.....	13 -
表 3-1 清大多媒體資訊檢索實驗室語文資料庫(MIR DB)詞庫分佈.....	16 -
表 3-2 中研院漢語平衡語料庫主題分佈.....	17 -
表 3-3 中研院漢語平衡語料庫語式分佈.....	17 -
表 3-4 中研院漢語平衡語料庫媒體分佈.....	17 -
表 3-5 中研院漢語平衡語料庫(SINICA CORPUS 3.0)詞庫分佈.....	18 -
表 3-6 修正後的清大多媒體資訊檢索實驗室語音資料庫(MIR DB)詞庫分佈.....	24 -
表 3-7 修正後的中研院漢語平衡語料庫(SINICA CORPUS 3.0)詞庫分佈.....	25 -
表 4-1 長詞優先與動態規劃演算法效能實驗分析 .....	29 -
表 4-2 加上定量複合詞構詞的斷詞系統效能實驗 .....	30 -
表 4-3 加上定量複合詞 疊詞構詞的斷詞系統效能實驗.....	31 -
表 4-4 加上定量複合詞 疊詞與姓名構詞的斷詞系統效能實驗.....	32 -
表 4-5 MIR 實驗室語音資料庫在 LONGEST WORD FIRST 斷詞的各階段表現.....	34 -
表 4-6 SINICA 平衡語料庫在 LONGEST WORD FIRST 斷詞的各階段表現.....	34 -
表 4-7 MIR 實驗室語音資料庫在 DYNAMIC PROGRAMMING 斷詞的各階段表現.....	34 -
表 4-8 SINICA 平衡語料庫在 DYNAMIC PROGRAMMING 斷詞的各階段表現.....	35 -