

CONTENTS

- ix Foreword
William S.-Y. Wang
- 1 Introduction (in English and Chinese)
Benjamin K. Tsou and Oi Yee Kwong
- PART I ANALYSES AND APPLICATIONS**
- 15 Evaluating Chinese Web-as-corpus: Some Methodological Considerations
Shu-Kai Hsieh
- 30 Exploring Individual Differences and Contextual Variations in Child Language Corpora
Hintat Cheung and Jing-chen Yang
- 39 Exploring Surface Cues for Deep Lessons in Children's Stories: A Preliminary Analysis with Aesop's Fables
Oi Yee Kwong
- 63 A Corpus-based Approach to Fingerprinting Stylistic Features of Classical Chinese Poetry: A Case Study of Liu Yong and Su Shi
Alex Chengyu Fang, Wan-yin Li and Jing Cao
- 82 Contrastive Corpus Linguistic Studies of English and Chinese: The Cases of Passive Constructions and Classifiers
Richard Xiao
- 115 Aspectual Properties and Grammaticalization of Progressive Markers as Reflected in Cantonese and Hakka Corpora (in Chinese)
Shin Kataoka

PART II ANNOTATION AND DATA EXTRACTION

- 139 Unknown Chinese Composite Words Tagging Using
Selective Back-off Smoothing
Samuel W.K. Chan, Mickey W.C. Chong and Tom B.Y. Lai
- 160 Lexicalized Statistical Pattern Matching: Search Engine-aided
Analysis for the Chinese Language
(in Chinese)
Maosong Sun and Ruying Sun
- 171 Morphological Constructions and Derivation of Senses and
Part-of-speeches for Chinese Compounds
You-Shan Chung and Keh-Jiann Chen
- 192 From Corpus to Grammar: Automatic Extraction of
Grammatical Relations from Annotated Corpus
Chu-Ren Huang, Jia-Fei Hong, Wei-Yun Ma and Petr Šimon
- 222 Expert System of Chinese Kinship Relations (in Chinese)
Zhenyu Chen
- 240 A Descriptive Text Data Format for Chinese Language
Corpora (in Chinese)
Aiping Fu and Hong Zhang

PART III CORPUS DEVELOPMENT

- 261 The Construction and Application of Mandarin Chinese
Multi-level Annotated Corpus (in Chinese)
Huiming Duan, Xuefeng Zhu, Shiwen Yu and Yunfang Wu
- 274 Chinese CCGBank Construction from
Tsinghua Chinese Treebank
Chang-ning Huang and Yan Song
- 312 The Hong Kong Cantonese Corpus: Design and Uses
Kang Kwong Luke and May L.Y. Wong

-
- 334 Cultivating Large-scale Parallel Corpora from Comparable Patents: From Bilingual to Trilingual, and Beyond (in Chinese)
Bin Lu, Benjamin K. Tsou and Ka Po Chow
- 356 Creation of a Database of Chinese Buddhist Translations and their Sanskrit Parallels for Buddhist Chinese Studies (in Chinese)
Qingzhi Zhu, Chin-chuan Wan, Qing Duan, Jihong Wang, Nan Jiang, Muyou Fan and Wei Huangfu

PART IV LANGUAGE AND SOCIETY

- 375 A Quantitative Analysis of Modern Chinese Neologism (in Chinese)
Rujie You
- 393 Chinese Network Media Monitoring Corpus and Empirical Investigation of Language Use on Blog (in Chinese)
Tingting He and Xinhui Tu
- 406 Knowledge Mining on Root Word Correlation Based on Modern Chinese Corpus (in Chinese)
Yuqi Sheng
- 422 Comment on Global Chinese Corpus (in Chinese)
Cheng Hai Chew
- 434 Linguistic Corpus and Sociolinguistic Study (in Chinese)
Jinzhi Su
- 447 LIVAC as a Monitoring Corpus for Tracking Trends beyond Linguistics
Benjamin K. Tsou and Oi Yee Kwong
- 473 **Contributors**

目錄

ix 序言（英文）
王士元

1 引言（英文/中文）
鄒嘉彥、鄭藹兒

第一部分 語料庫分析及應用

15 漢語網路語料庫評測方法初探（英文）
謝舒凱

30 探索兒童語料庫中的個別差異與語境變異（英文）
張顯達、楊靜琛

39 由淺入深：從表層語言特徵探索寓言故事的深層意義
（英文）
鄭藹兒

63 基於語料庫的古詩詞文學風格辨識：
柳永及蘇軾詩詞範例研究（英文）
方稱宇、李昀燕、曹競

82 英漢對比語料庫研究：以被動結構與量詞為例（英文）
肖忠華

115 從粵客語語料看進行體標記的特性與語法化
片岡新

第二部分 語料庫標註及數據抽取

- 139 漢語未登錄複合詞的詞性標註（英文）
陳偉光、莊華祥、黎邦洋
- 160 詞匯化模板定量匹配——借助於搜索引擎的中文分析
孫茂松、孫如穎
- 171 中文複合詞的構詞及其語意和詞性衍生規律（英文）
鍾友珊、陳克健
- 192 語料與語法：標記語料中語法關係的自動抽取（英文）
黃居仁、洪嘉馥、馬偉雲、石穆
- 222 漢語親屬關係專家系統
陳振宇
- 240 漢語語料庫的文本描述
傅愛平、張弘

第三部分 語料庫的開發

- 261 現代漢語多級標註語料庫的建設及其應用
段慧明、朱學鋒、俞士汶、吳雲芳
- 274 從清華中文短語結構樹庫到組合範疇語法樹庫（英文）
黃昌寧、宋彥
- 312 香港粵語語料庫的設計和用途（英文）
陸鏡光、王麗賢
- 334 從專利文本建立大型平行語料庫：由雙語到多語
路斌、鄧嘉彥、周嘉寶

- 356 漢梵佛典雙語標註語料庫的構建
朱慶之、萬金川、段晴、王繼紅、姜南、范慕尤、皇甫偉

第四部分 語言與社會

- 375 當代華語新詞分地定量分析
游汝杰
- 393 漢語網絡媒體監測語料庫及博客語言應用調查
何婷婷、涂新輝
- 406 基於當代漢語流通語料庫的根詞相關性知識挖掘研究
盛玉麒
- 422 漢語融合時代的語言研究與語料庫
周濟海
- 434 語料庫與社會語言學研究
蘇金智
- 447 從LIVAC追蹤語料庫探索泛華語地區語言以外的演變
(英文)
鄒嘉彥、鄭藹兒
- 473 作者簡介(英文)