

# 語言學研究所籌備處

## 壹·八十七年度研究工作重要成果

語言學研究所籌備處於民國八十六年八月奉准成立，八十七年度尚未有獨立預算，本處同仁合計十二名，其中一名赴香港城市大學講學，一名獲選為國科會第卅五屆科技研究進修人，赴美國柏克萊大學研究進修。

本年度研究重點為漢語、南島語、藏緬語、實驗語音學、計算語言學等。計有研究計畫五類十四項（包含本院主題計畫），舉辦學術會議六項，宣讀會議論文七人廿二篇，發表論文共四十九篇，以下所列為研究工作之重要成果：

### 一、漢語

#### （一）南北朝音韻史

1. 重建南北朝各時空區段中共同語的音韻系統、歷時演變，及與個別音系間的分合關係。
2. 對兩漢魏晉與隋唐間的音韻發展，提出合理的解釋。
3. 對現代方言中早期語言層的來源，提供南北朝時期的可能線索。
4. 資料性成果：南北朝音韻譜及南北朝諸家音韻資料表。

#### （二）先秦語料詞類的標注與研究(I)（國科會補助）。

1. 編修完成詞類標注手冊，可提供學術界作為劃分詞類參考之用。
2. 編寫各類應用程式，提供欲做類似研究的學者使用。
3. 初步完成古籍《左傳》《國語》《論語》《孟子》《荀子》《莊子》《呂氏春秋》《韓非子》《墨子》（不包括〈經〉〈經說〉）《戰國策》等的斷詞及詞類的標注約 576,000 字。
4. 利用電腦把已標過詞類的文獻進行歸納整理，並表列出初步的各詞詞類使用情形以及詞類活用情形。

#### （三）以詞彙為中心·語料庫為本的國語語法與語意研究(請參見計算語言學(一)) 根據語料發展出描述漢語動詞語意及事件結構(event structure)之完整理論，並根據此理論對漢語近義動詞進行詳細分析，預測其語法、語意行為及互動。

#### （四）中文詞知識庫（本院主題計畫，參見計算語言學(二)）

發展漢語名語義之表徵(representaion)理論，其重點在區分語意(sense)與意面(meaning facet)，並根據此理論分析三百個名詞之語意及意面。

#### （五）中文語料庫（本院跨所主題計畫—中文資訊處理計畫，參見計算語言學(四)）

1. 完成「中央研究院現代漢語平衡語料庫」3.0版之標記校對工作。共五百萬詞(words)詞類標記完成之語料在WWW上免費提供全球學者進行研究。
2. 繼續與本院及院外其他單位，如中山大學、元智大學、史丹福大學等，合作開發古漢語語料庫及詞彙庫。

## 二、南島語

### (一) 台灣南島語言的語法類型研究。(本院主題計畫)

1. 蒐集五種台灣南島語言的資料：賽德克、阿美、噶瑪蘭、巴則海、邵等，所收集的資料包括詞彙、句子、文本(texts)。
2. 整理和分析研究已消失的凱達格蘭語料(淺井惠倫蒐集的田野調查資料)，探討其語法結構系統。
3. 撰成研究論文四篇。

### (二) 搶救數種瀕臨滅絕的語言。

1. 對三種面臨滅絕的語言(巴則海、噶瑪蘭、邵)進行以下工作：(1)較長的詞彙表：依字母次序排列，有些詞將附以例句。(2)至少若干篇文本。(3)語法綱要：寫出較重要的語法結構系統，包括焦點系統、格位標記系統、時態與情貌等等。
2. 蒐集較多的詞彙、語句及長篇語料，即文本(texts)。
3. 編纂噶瑪蘭語和巴則海語詞典。

### (三) 詞組結構的分析與比較：排灣語和噶瑪蘭語個案研究(II)。

1. 建立排灣語和噶瑪蘭語比較句法研究的完整語料庫和資料庫，並加上以下四種資料：(1)否定詞的意義與用法；(2)關係子句的結構與限制；(3)疑問詞的意義、用法及限制；(4)分裂句與準分裂句的結構與限制。
2. 完成排灣語和噶瑪蘭語詞組結構的完整分析，並對下列問題提出結論：(1)否定詞與其他虛詞的相對位置為何？(2)基本句子結構是否為「等同句」(equational sentences)？(3)關係子句的限制是否與句子結構有關？(4)疑問句的限制是否與句子結構有關？(5)否定句、關係子句、疑問句、分裂句與準分裂句中的述語結構為何？
3. 建立排灣語和噶瑪蘭語的完整詞組結構。
4. 撰寫排灣語和噶瑪蘭語詞組結構的完整比較分析。

### (四) 魯凱語動詞分類。

1. 描述魯凱語每一個方言的動詞在構詞上、語法上及語意上的特色。
2. 將該方言的動詞加以分類。
3. 比較魯凱語六個方言動詞之分類。

4. 提供各動詞變化資料表和例句。

### 三、藏緬語

- (一) 藏緬語族語言調查研究：西夏音韻的歷史比較研究。
  1. 完成西夏語詞典之編輯，將漢、藏、緬同源詞及重要的其他藏緬同源詞，特別是羌語支語言的同源輸入，對藏緬及漢藏比較語言學提供重要資料。
  2. 建立西夏語音韻史。
- (二) 嘉戎語群口語語料田野調查與分析計畫。
  1. 爾龔語木爾宗話之音系建立與語料蒐集。
  2. 補實嘉戎語本真話及草登話之資料採集及電腦建檔。
  3. 將每種語言及方言以音位方式記錄常用詞彙三千五百左右及十至十五則故事、歌謠、民俗解說或對話等不同性質之話語材料，建立便易擴充檢索之嘉戎語群綜合詞彙、話語語料檔。

### 四、實驗語音學

- (一) 漢語與腦神經的關係（本院主題計畫）。

子計畫(一)：從腦神經影像的變化看漢字閱讀的歷程

經由 fMRI 檢驗漢語的聲調特徵與腦部的機制運作是否有直接對應關係。語音訊息在中文單字與雙字詞辨識歷程中所扮演的角色，透過不同實驗典範的檢驗發現：

  1. 透過語音訊息而提取字義的說法得到支持，而不同於拼音文字的是，中文字的字形表徵具有相當強勢的作用和影響。
  2. 基於中文構字的獨特性，使得單字詞與雙字詞的處理歷程對於字形與語音訊息的感受性並不同。

目前在 3T fMRI 實驗室設立完成前，除了建立漢語處理歷程的適當模式外，根據此一模式為基礎，我們已建立一些基本 fMRI 量測的實驗典範，同時先以 1.5T fMRI 來做初探，以利日後腦造影實驗的漢語研究得以順利進行。
- (二) 國語語音資料庫(本院跨所主題計畫—中文資訊處理計畫)。
  1. 口語語料錄音的規畫與設計：在口語語料的錄音方面，目前以語音查詢的對話為主，我們選擇了 csu 的 toolkits，作為錄音用的平台 (lg. engineering platform)，同時針對不同的應用，設計其對話的模式。
  2. 語音資料庫機讀標示系統的設計：我們參考了 SAMPA 的設計，先以國語為出發，同時考慮未來多語言的相容性，設計了一套 ascii 的音標，及可適用於多種語言的音節編碼。

## 五、計算語言學

(一) 以詞彙為中心·語料庫為本的國語語法與語意研究。

1. 本計劃在學術上可建立漢語相關研究深廣，及變異與共通性兼顧的基礎。詞彙語意之深入研究，不但可促使以後的語言相關研究更能觸及語言使用與演變的核心因素，增加理論、心理與計算語言學之解釋能力，並可運用到辭典與文法書之編纂方面，提昇國內語文教材之水準，並使其科學，化間接提高國人語文水準。這些知識更可以機讀方式存在，成為機器翻譯智慧型資訊檢索等中文自然語言處理所不可或缺的資料。
2. 至於詞彙特性群的建立，可使語料庫語言學的研究與其他人文社會科學結合，不但可作為歷史語言學與風格學之基本素材，並可作年代作者判定版本校勘，不同文體、體裁比較文學批評等的參考資料，亦可間接幫助機器理解文章。

(二) 中文詞知識庫(本院主題計畫)。

1. 完成「資訊處理用中文分詞規範」國家標準草案。提中央標準局審查中。
2. 完成標準參考詞庫，含核心通用，參考三級，共八萬餘詞。
3. 完成計算語言學及理論語言學術語中英對照表。
4. 完成以上三項資料之光碟及 WWW 版，免費提供國內外研究者使用。

(三) 中文剖析系統(本院主題計畫)。

開發中文樹圖資料庫(TreeBank)之構建環境，並完成 1,000 個完整樹圖。

(四) 中文語料庫(本院跨所主題計畫—中文資訊處理計畫)。

完成語料庫詞頻統計等重要參考資料。

## 貳·舉辦學術會議(含協辦)

一、第十屆計算語言學國際研討會(ROCLING 10) (86.8.22-24)

二、第八屆南島語言學國際研討會 (86.12.28~87.1.2)

三、人文計算研討會(87.5.1)

四、平埔族群與台灣歷史文化學術研討會 (87.5.16-17)

五、「中文資訊處理分詞規範」公聽會(87.5.29)

六、漢籍電子文獻資料庫建置的回顧與前瞻研討會(87.6.12-13)

## 參·八十七年度論文目錄

### 李壬癸 (Li, Paul Jen-kuei)

李壬癸。〈一百年來平埔族語言消失的歷程〉，《台灣原住民歷史文化學術研討會論文集》，頁 42-62，1998 年 6 月。台灣省文獻委員會。

李壬癸。〈巴則海語的格位標記系統〉，《台灣語言及其教學國際研討會論文集》1:57-81。新竹師院，1998 年 5-6 月。

李壬癸。《台灣原住民文化基本教材》導論，1998 年 8 月。國立編譯館。

李壬癸。〈凱達格蘭族的原語傳說故事集〉介紹，李壬癸、黃秀敏，《宜蘭文獻雜誌》32:97-104，1998 年 3 月。

李壬癸。〈邵族的地位—兼評白樂思 (Blust 1996) 的邵族地位說〉，《平埔族群與台灣歷史文化學術研討會》。中研院，1998 年 5-6 月。

Li, Paul Jen-kuei. Review of Shigeru Tsuchida et al's (1987) "Lists of Selected Words of Batanic Languages." *Oceanic Linguistics* 36.2:424-429. December 1997.

Li, Paul Jen-kuei. Some problems in the Basay language. Paper presented at the 8th International Conference on Austronesian Linguistics, Taipei, Dec. 28-30, 1997.

### 龔煌城 (Gong, Hwang-cherng)

龔煌城。〈西夏語若干韻母轉換的起源——重疊複合詞〉，《中國境內語言暨語言學》(第四輯：語言類型) 4:265-289，中央研究院歷史語言研究所會議論文集之二。

龔煌城。〈西夏語的緊元音及其起源〉，《第六屆中國境內語言暨語言學國際研討會》。中研院，1998 年 7 月 14-16。

### 何大安 (Ho, Dah-an)

何大安。〈論原始南島語同源詞〉，《王士元先生六五壽慶論文集》(印刷中)。

何大安。〈古漢語聲母演變的年代學〉(本處講論會文稿)。

何大安。〈南北朝音韻史(一)〉(國科會研究計畫成果報告)。

何大安。〈漢字古今音檢索系統(一)〉(國科會研究計畫成果報告，與台大楊秀

芳教授合作。)

### 鄭秋豫 ( Tseng, Chiu-yu )

- Tseng , Chiu-yu. 語言所籌備處講論會(86.12.1): Prosodic Group in Mandarin Chinese — Defining the operating speaking unit in fluent continuous speech.
- Tseng , Chiu-yu. 會議論文: Machine readable phonetic transcription system for Chinese dialects spoken in Taiwan. (The First Oriental COCODA Workshop, May 18-19, 1998, Tsukuba, Japan)
- Tseng , Chiu-yu. 會議論文: Networking in Mandarin Chinese — Another phase of a Mandarin speech database. (漢語及少數民族語言語音學研討會, May 28-30, 1998, 香港)

### 黃居仁 ( Huang, Chu-Ren )

- Huang, Chu-Ren & Kathleen Ahrens & Keh-jiann Chen. 1998. A Data-driven Approach to the Mental Lexicon: Two Studies on Chinese Corpus Linguistics. *Bulletin of the Institute of History and Philology*. 69.1.151-179.
- Huang, Chu-Ren & Keh-jiann Chen & Feng-yi Chen & Li-Li Chang. 1997. Segmentation Standard for Chinese Natural Language Processing. *Computational Linguistics and Chinese Language Processing*. 2.2.47-62.
- Huang, Chu-Ren. 1997. Ed. Special Issue on Computational Resources for Research in Chinese Linguistics. *Computational Linguistics and Chinese Language Processing*. Vol.2. no.1.
- Huang, Chu-Ren. 1998. Classifying Event Structure Attributes: A Verbal Semantic Perspective from Chinese. Invited paper. Chinese Workshop. The 1998 International Lexical-Functional Grammar Conference. June 30-July3. Brisbane, Australia.
- Huang, Chu-Ren. 1998. Classified Information: A Corpus-based Approach Towards Automatic Extraction of Nominal Semantic Structure. Invited Speech. The Third International Conference on Information-Theoretic Approaches to Logic, Language and Computation. Hsi-tou, Taiwan. June 15-18.
- Huang, Chu-Ren & Keh-jiann Chen & Zhao-ming Gao. 1998. Noun Class Extraction from a Corpus-based Collocation Dictionary: An Integration of Computational and Qualitative Approaches. Roundtable Conference on Quantitative and Computational

- Studies on the Chinese Language. June 26-27. Hong Kong. Hong Kong City University.
- Huang, Chu-Ren & Zhao-ming Gao & Keh-jiann Chen & Claude C.C. Shen. 1998. Towards a Sharable and Reusable Lexical List: The construction of a standard reference lexicon for Chinese NLP. The Pacific Neighborhood Consortium Annual Meeting. June 16-18. Taipei: Academia Sinica.
- Huang, Chu-Ren, Liu Mei-chun, and Mei-chih Tsai. 1998. From Lexical Meaning to Event Structure Attributes: Across Semantic Classes of Mandarin Verbs. The 6th International Conference on Chinese Linguistics/The 10th North American Conference on Chinese Linguistics. June 26-28. Stanford.
- Huang, Chu-Ren & Mei-Chih Tsai. 1997. From Near Synonyms to Event Structure: Corpus-based Studies of Mandarin Verbal Semantics. Mini-Conference on Lexical Semantics. November. National Chung-cheng University.
- Ahrens, Kathleen & Lili Chang & Keh-jiann Chen & Huang, Chu-Ren. 1997. Meaning Representation and Meaning Instantiation for Chinese Nominals. Proceedings of ROCLING X. Pp.4-18.
- Chief, Lian-cheng & Huang, Chu-Ren & Keh-jiann Chen & mei-chih Tsai & Lili Chang. 1998. What can Near Synonyms Tell us. Proceedings of the 1998 International LFG conference. Brisbane. Australia.
- Gao, Zhao-ming & Huang, Chu-Ren. 1998. Statistical and Computational Approached to Chinese Linguistics. Invited paper. 語言學研研究方法研討會，台北，政治大學，1998年6月16日。
- Liu, Mei-chun & Huang, Chu-Ren & Charles C.L. Lee. 1998. When Endpoint Meets Endpoint: A Corpus-based Semantic Study of Throwing Verbs. The 6th International Conference on Chinese Linguistics/The 10th North American Conference on Chinese Linguistics. June 26-28. Stanford.
- Luo Yi-fen & Huang, Chu-Ren 1997. 漢語動貌詞「在」與「著」的詞彙意義 [The Lexical Semantics of Mandarin Aspects 'Zai' and 'Zhe']. The Fifth International Conference on World Chinese Language Teaching. Taipei: World Chinese Association.
- Tsai, Mei-Chih & Huang, Chu-Ren & Keh-jiann Chen & Kathleen Ahrens. 1997. Towards a Representation of Verbal Semantics: An Approach Based on Near-Synonyms. Proceeding ROCLING X. Pp.34-48.

Wen-jen Wei & Huang, Chu-Ren & Lili Chang. 1997. 中文的合併詞 [Telescopic Compounds in Chinese]. The Fifth International Conference on World Chinese Language Teaching. Taipei: World Chinese Association.

黃居仁, 陳克健, 賴慶雄。1997. 編著. 國語日報量詞典. (781 頁). 台北. 國語日報社.

### 魏培泉 ( Wei, Pei-chuan )

魏培泉。說中古漢語的使成結構，本處講論會。

魏培泉。中古佛經斷詞手冊，國科會計劃成果報告（1998 年 1 月）。

魏培泉·劉承慧。竺法護譯經詞表，國科會計劃成果報告（1998 年 1 月）。

### 孫天心 ( Sun, Jackson T.-S. )

Sun, Jackson T.-S. The typology of tone in Tibetan. *Chinese languages and linguistics IV: Typological studies of languages in China (Symposium series of the Institute of History and Philology, Academia Sinica; Number 2)* edited by Chiu-yu Tseng, pp. 485-521 (1997). Nankang, Taipei: Academia Sinica.

Sun, Jackson T.-S. Nominal morphology in Caodeng rGyalrong. *Bulletin of the Institute of History and Philology, Academia Sinica* 69.1: 103-49 (1998).

### 羅仁地 ( LaPolla, Randy J. )

LaPolla, Randy J, *Syntax: Structure, Meaning and Function (Cambridge Textbooks in Linguistics)*. (with Robert D. Van Valin, Jr.). Cambridge University Press; 共九章，700 多頁)

LaPolla, Randy J, *Book review of Language Atlas of China (Pacific Linguistics, Series C, No. 102)*, ed. By S. A. Wurm et al. Hong Kong: Longman Group (Far East) Ltd., 1987. *International Review of Chinese Linguistics* 1(1997). 1:90-98

### 湯志真 ( Tang, Chih-Chen Jane )

Tang, Chih-Chen Jane, *Phrase Structures: A Comparative Analysis of Paiwan, Kavalan and Tsou* (國科會研究計畫成果報告)

湯廷池, 湯志真。"華語情態詞序論" (第五屆世界華文教學研討會), 1997.12.27-29, 台北. 劍潭。

Tang, Chih-Chen Jane, On Clausal Complements in Paiwan (The Eighth International Conference on Austronesian Linguistics, 1997.12.28-30, Academia Sinica.)

Chiu, Bonnie & Tang, Chih-Chen Jane, Paiwan Focus Markers Refocused (The Sixth International Symposium on Chinese Languages and Linguistics, 1998.7.14-16, Academia Sinica.)

### 齊莉莎 (Zeitoun, Elizabeth S.)

Zeitoun, Elizabeth. 1997. The pronominal system of Mantaoran (Rukai). (魯凱語萬山方言代名詞系統) *Oceanic Linguistics*, 36.2: 312-346.

Zeitoun, Elizabeth. 1997. Review of Holmer's (1996) "A parametric grammar of Seediq" *Oceanic Linguistics*, 36.2: 429-436.

Zeitoun, Elizabeth. 1997. Temporal, hypothetical and counterfactual clauses in Formosan languages. Paper read at the *Eighth International Conference on Austronesian languages*, Taipei, December 28-30.

### 蕭素英 (Hsiao Su-ying)

蕭素英。1998 海陸客家話送氣聲母對聲調之影響初探，《一九九七全國英美文學、語言學論文研討會論文集》·文鶴·PP.81-100 亦刊於《教育部獎勵漢語方言研究著作得獎作品論文集》PP.16-1~16-19

蕭素英。1998 蒙古秘史中的使動式《一九九八全國英美文學、語言學論文研討會論文集》文鶴 PP.29-42

蕭素英。1998 漢語方言聲母送氣對聲調影響之聲學研究 《漢語及少數民族語言語音學研討會論文集》PP.37-40 香港城市大學 May 28-30, 1998. 徐雲揚, 林茂燦 編