

黃征 張崇依 著

吐蕃敦煌文獻校錄整理



上

古代漢語學研究系列

黃征 主編

上海古籍出版社

黃征 張崇依 著

吐蕃敦煌文獻校錄整理



上

上海古籍出版社

南京師範大學
國家“211工程”三期重點學科建設項目
“語言科技創新及工作平臺建設”

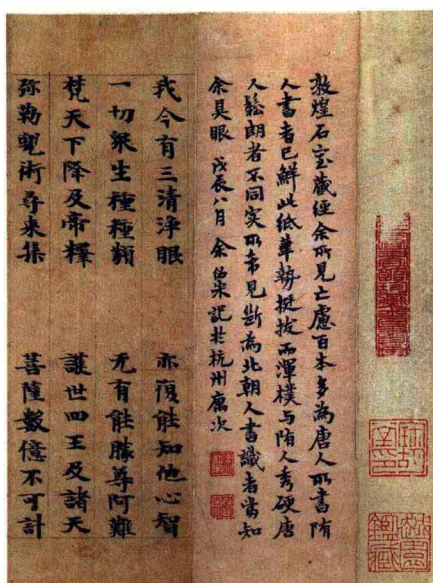
《語言科技文庫》

總主編 李葆嘉

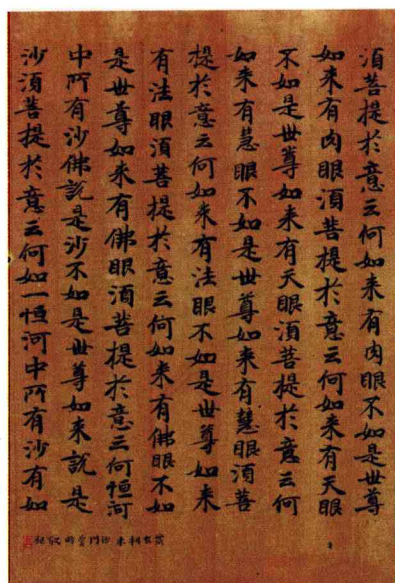
計算語言學研究系列	陳小荷 主編
語義語法學研究系列	李葆嘉 主編
漢語方言學研究系列	劉俐李 主編
古代漢語學研究系列	黃 征 主編
語言教學與研究系列	肖奚強 主編
語言新專題研究系列	倪傳斌 主編



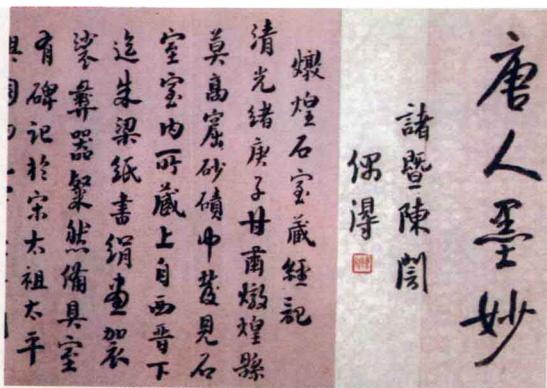
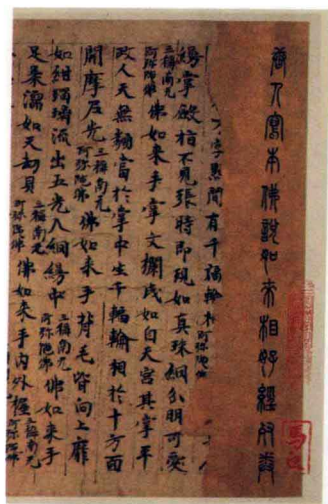
1. 浙藏敦煌寫卷（部分）



2. 浙敦001 北朝寫本《增壹阿含經》
（局部）

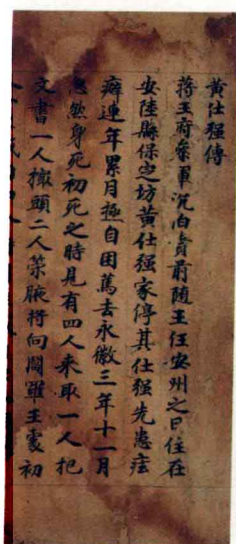


3. 浙敦005 唐寫本《金剛般若波羅
蜜經》（局部）

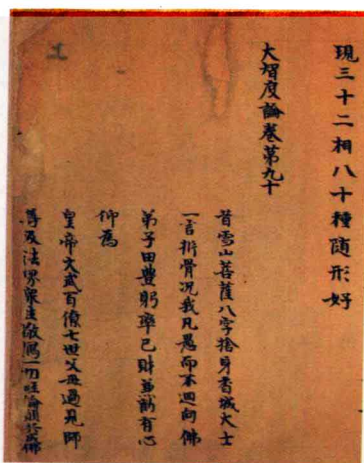


4. 浙敦020 唐寫本《佛說如來相好經》殘卷（局部）

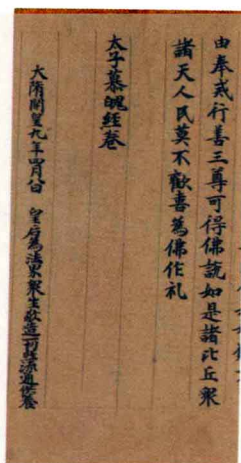
5. 浙敦025 唐寫本《摩訶般若波羅蜜經》收藏者陳闡題記（局部）



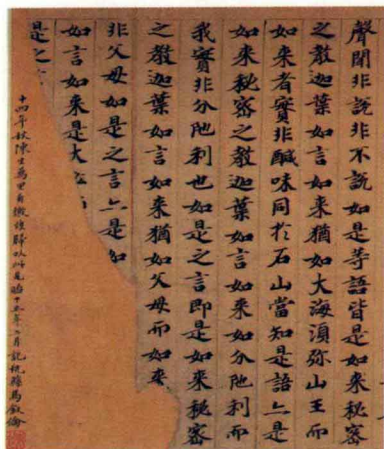
6. 浙敦026 唐寫本《黃仕強傳》（局部）



7. 浙敦027 北朝寫本《大智度論》卷尾



8. 浙敦029 隋寫本《太子慕魄經》



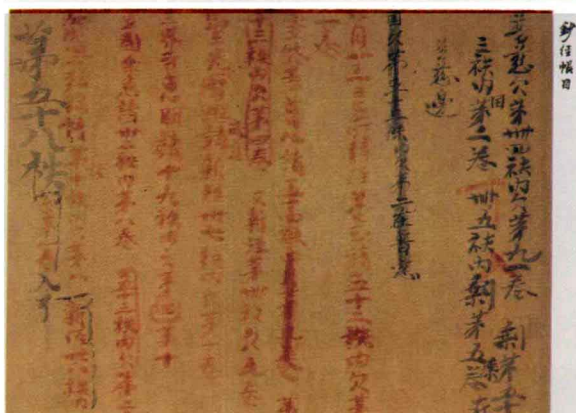
9. 浙敦030 唐寫本《妙法蓮華經》(局部)



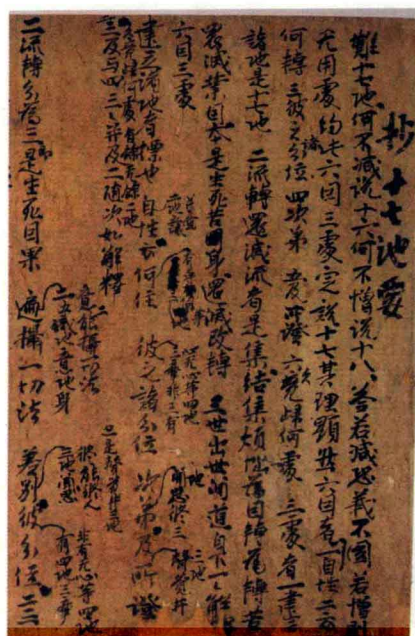
10. 浙敦031 唐寫本《大般若波羅蜜多經》卷首



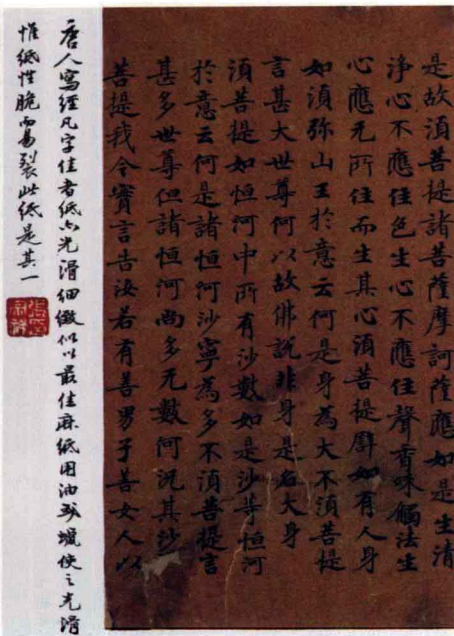
11. 浙敦064 唐寫本《大般若波羅蜜多經》近人題觀



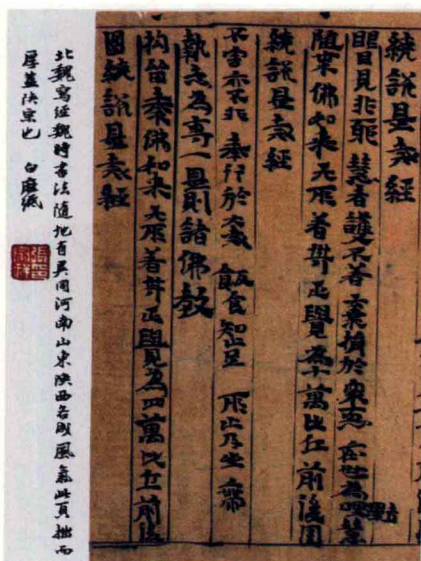
12. 浙敦070 唐寫本《寺院欠經請經賬目雜抄》



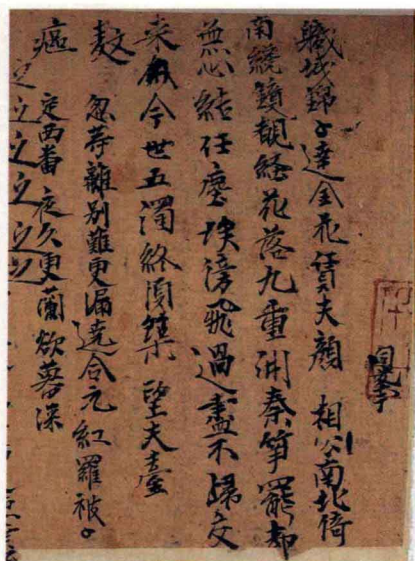
13. 浙敦072 唐寫本《抄十七地要》



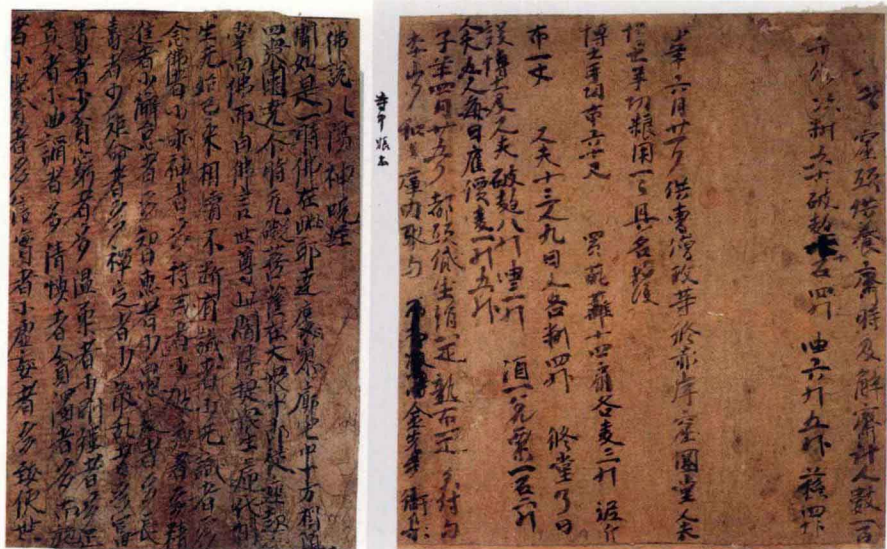
14. 浙敦087 唐寫本《金剛般若波羅蜜經》殘卷（局部）



15. 浙敦112 東晉寫本《十誦比丘尼波羅提木叉戒本》（局部）



16. 浙敦113 南唐五代寫本《曲子詞及雜抄》



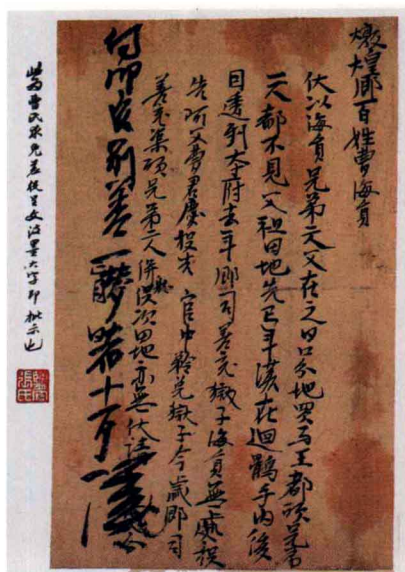
17. 浙敦115 唐寫本《佛說八陽神咒經》卷首 18. 浙敦116 唐寫本《子年金光明寺麵油等破曆》



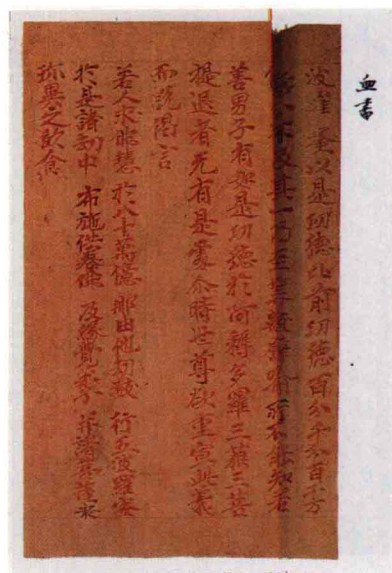
19. 浙敦117 唐寫本《七言敦煌詩》



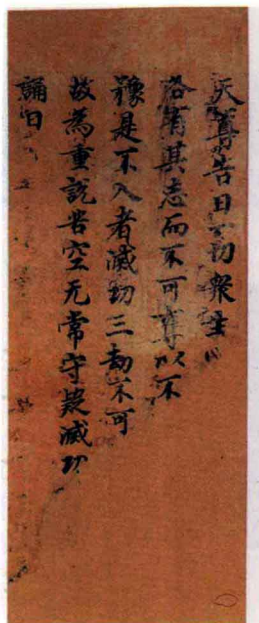
20. 浙敦129 經卷引首物



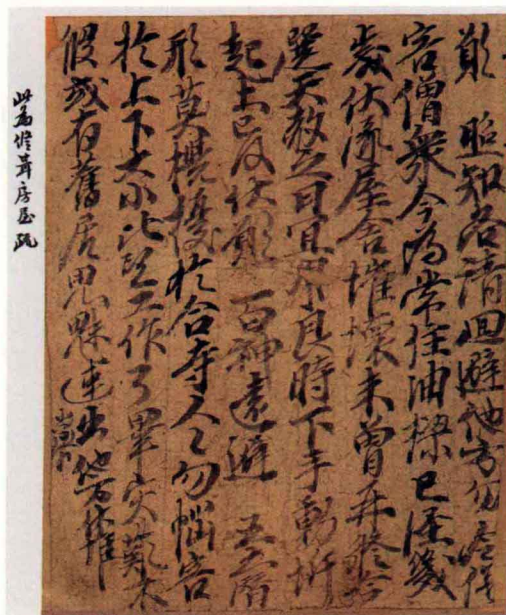
21. 浙敦135 唐寫本《敦煌鄉百姓曹海員訴狀并判》



22. 浙敦148 唐寫本《妙法蓮華經》



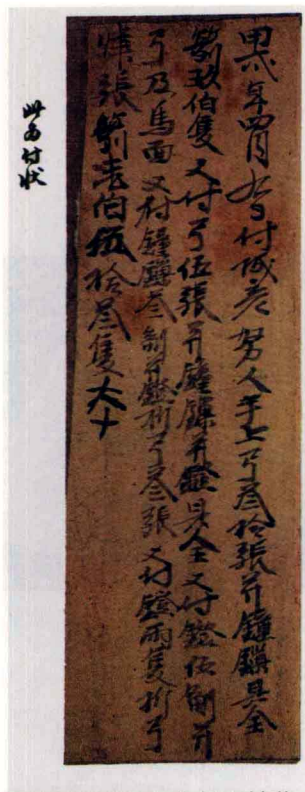
23. 浙敦150 唐寫本《道經殘片》



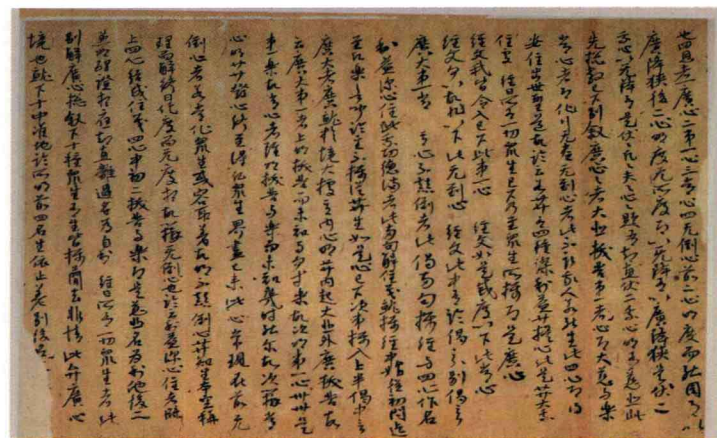
24. 浙敦154 唐寫本《某寺修舍告疏》



25. 浙敦155 唐寫本《領軍械狀》



26. 浙敦156 唐寫本《付軍械狀》



27. 浙敦178 唐寫本《佛經注疏殘片》



28. 浙敦193 唐代原裝鑲嵌蓮花軸頭



29. 浙敦193 唐寫本《妙法蓮華經》
卷尾



30. 浙敦196 唐寫本《和菩薩戒文》(局部)

《語言科技文庫》總序

李葆嘉

當代語言學已經進入了一個科學與技術的互補時代，信息處理水準成為衡量國家現代化水平的重要標誌之一。知識世界的載體是語言符號系統，信息處理的根本對象是語言信息處理。與計算機的出現使得語言符號有可能成為數據處理對象相似，神經科學實驗儀器設備的應用，使得在大腦神經層面探討語言機制成為可能。這些無疑都引導語言研究走向科技化，“語言科技新思維”（李葆嘉 2001）應運而生。

所謂“語言科學”包括理論語言學、描寫語言學、歷史語言學、應用語言學等分支學科，所謂“語言技術”指語言研究的現代技術手段，包括語言信息處理、語音實驗分析，以及語言的神經、心理和行為實驗分析的技術手段等。就語言信息處理而言，又可以分為語料庫研製技術、知識庫研製技術、知識挖掘和抽取技術、句法信息處理技術、辭彙信息處理技術、語音信息處理技術、語義信息處理技術、語用信息處理技術等。

2001年5月，南京師範大學文學院創辦了史無前例的“語言科學及技術系”，率先邁出了從傳統文科教育範型向現代科技教育範型轉變的步伐。“十五”期間，南京師大“211工程”重點學科建設項目“語言信息處理與分領域語言研究的現代化”（陳小荷教授主持），以基礎平臺建設、資源建設和理論探索等為主，邁出了語言科技研究的一大步。

“十一五”期間，南京師大文學院、外國語學院和國際文化教育學院連袂申報“211工程”三期重點學科建設項目。該項目以“語言科技”為引導，以“多學科交叉、跨院系整合、開放型營運”為理念，建設具有前瞻性、原創性、成長性的語言科技高級工作平臺。以典型課題的工作原理為核心，進行資源開發和系統研製，拓展語音科技、二語習得的神經機制研

究、語言能力受損兒童的語言能力研究等新方向。同時造就新一代學術領軍人物和培養一批高層次複合型人才，以期形成一支高水準的交叉學科團隊。該項目設計，體現了工作平臺建設、理論創新、應用研究、人才培養、團隊建設的學科發展一體化思路。其旨趣在於，加速語言研究從傳統文科範型向現代科技範型的轉變，以引領21世紀語言科技的新潮流。

作為新興交叉學科項目，通過教育部組織的專家匿名評審，“語言科技創新及工作平臺建設”（2008—2011）獲批，總投入1000萬元。總體而言，這一“語言科技創新”團隊，分支學科齊全、專業知識互補，涵蓋了理論語言學、計算語言學、語義科技、語音科技、實驗方言學、歷史語言學、神經語言學、二語習得研究、話語行為語言學等領域。這一期間，項目組成員獲批的國家級基金項目達20多項。該項目理念之前瞻、實力之雄厚、工程之浩大、經費之保障，為學界矚目。

2008年秋，本項目以南京師範大學語言科技研究所為實施單位正式啟動。主要有三大任務：建設一個“整體性、科技型、開放型”的語言信息科技實驗室；研製一個具有“獨創性、交叉性、應用性”的語言科技工作平臺；出版一套具有“大視野、跨學科、新思路”的語言科技文庫。

從實驗室方案設計到設備招標採購，再到實驗室用房改造，經過8個月的努力，2009年12月，語言信息科技實驗室建成，為語言研究從傳統範型向科技範型的轉變提供了基本保障。該實驗室劃分為實驗工作區、科研工作區和管理服務區。實驗工作區建有語音實驗與計算室、神經認知實驗與計算室、課堂話語實錄室三個專門實驗室。科研工作區建有語義科技工作室、語音科技工作室、方言實驗工作室、知識工程工作室I（先秦辭彙）、知識工程工作室II（中古辭彙）、知識工程工作室III（敦煌俗語言文字）、語言（英）習得神經機制工作室、語言（漢）習得中介機制工作室，以及參研工作室。管理區服務包括辦公室、管理室、編輯室和交流室。出席“語言科技高層論壇暨語言信息科技實驗室落成儀式”（2009年12月14日）的專家認為，該實驗室體現了語言學跨學科研究的當代性和先進性，具有整體性、科技型、開放型三個特點，處於全國領先地位，是“語言

科技新思維”的又一體現。同時認為，該實驗室的科研工作涵蓋了四個二級學科、四個博士學位點，有穩定明確的研究方向，有合理的設計規劃和很好的科研基礎；整體設計合理，功能齊備，以教育部重點實驗室建設標準衡量，很多方面超過了指標。

語言科技工作平臺是基於工作原理（課題定位—理論方法—技術路線—關鍵技術—評估方式）而建設的高級平臺。一方面，從語言信息、語言知識和語言機制三個層面，圍繞典型課題進行設備配置、資源建設和軟件開發；一方面，將典型課題研究與工作平臺建設融為一體，依據典型課題建設的子平臺應具有解決同類課題的功能。

建設語言科技工作平臺的目標是要實現語言研究手段的技術化和模型化，總體設計包括三個二級平臺和八個子系統。

一、語言信息工作平臺。1. 語義科技工作系統（李葆嘉教授主持）：基於辭彙語義—句法語義的一體化研究思路，開發“人一機交互語義標注工具”，研製“深度語義標注信息庫”；研製“幼兒（2—6）日常話語跟蹤語料庫”，完成幼兒語義系統和話語行為分析研究。2. 語音科技工作系統（顧文濤教授主持）：研製“多語言、多語境、多語用的語音語料庫”，基於聲學信號分析、感知實驗和數學建模，完善語音韻律理論與相關技術應用。3. 方言實驗工作系統（劉俐李教授主持）：完成“網絡版漢語方言有聲語料數據庫”，擬定系統的可操作性語音、辭彙、語法實驗模型和研究方法，進一步完善新興交叉學科“實驗方言學”。

二、語言知識工作平臺。1. 先秦辭彙統計與知識檢索系統（陳小荷教授主持）：研製“先秦文獻語料庫”、“專名知識庫”、“漢語辭彙檔案庫”等，開發先秦文獻自動分詞演算法、古籍版本異文自動發現演算法、同指專名檢索軟件工具等，完成“先秦漢語辭彙統計與知識檢索”。2. 中古辭彙統計與知識檢索系統（董志翹教授主持）：研製“中古文獻語料庫”、“專名知識庫”、“中古漢語辭彙檔案庫”等，開發中古文獻自動分詞和標注工具等，完成“中古漢語辭彙統計與知識檢索”。3. 敦煌俗語言文字知識檢索系統（黃征教授主持）：研製“敦煌文獻資料庫”、“敦煌文獻

俗詞語檔案庫”，開發相應工具，完成“敦煌文獻資料與知識檢索”。

三、語言機制工作平臺。1. 二語（英）習得的神經機制研究系統（倪傳斌教授主持）：研製“英語受蝕辭彙庫”等，基於行為學、腦成像和腦電三維度模型，進行中國人英語習得與磨蝕的神經機制研究，完成“基於神經機制的英語個性化學習分析系統”。2. 二語（漢）習得的中介機制研究系統（肖奚強教授主持）：研製“留學生漢語口語中介語語料庫”，基於中介語理論、對比分析理論、偏誤分析理論以及二語習得影響因素等，完成“留學生漢語習得的中介機制研究”。

這一工作平臺，既是科技研究平臺，也是人才培養平臺，即一個現代化的科學研究和人才培養工作體系。

作為本項目的文本成果，《語言科技文庫》包括計算語言學研究、語義語法學研究、漢語方言學研究、古代漢語學研究、語言教學與研究、語言新專題研究六個系列。其總體特徵為：領域的開拓性、理論的原創性、選題的新穎性、方法的交叉性、考據的精審性、成果的應用性。在研究過程中，除了資料獲取分析、資源建設和軟件開發，更重要的還是要有新思路、新理論和新材料。陳小荷提出的先秦文獻信息處理新方法，從先秦典籍注疏文獻中挖掘出用於自動分詞和詞義消歧的知識，再注入已開發的古漢語分詞和詞性標注工具中去，所取得的先秦古籍版本異文自動發現、先秦辭彙知識自動挖掘等成果均具開拓性。李葆嘉提出的語義語法學理論和話語行為理論，基於研製專用語料庫或語義信息庫和技術手段，開拓了語義網路建構、深度語義分析和話語行為研究等新的領域。劉俐李建構的實驗方言學理論和方法，為方言學向現代科技方法的轉型研究提供了新路，並取得了一系列新成果。黃征多年來從事敦煌文獻及其俗詞語文字研究，古代漢語學研究系列中的敦煌文獻校錄整理，以及敦煌寫本字詞考釋、以古佚和疑偽經為中心的敦煌佛典詞語和俗字研究、兩漢聲母系統研究等新見迭出。肖奚強基於漢語中介語語料庫的二語習得研究，在對外漢語教學研究界已經產生了影響。錢玉蓮等的漢語介詞與相應英語形式比較研究等專著，各有亮色。倪傳斌依據語言測試和認知實驗等資料，

從行為學、生理學和語言學三個層面分析影響中國英語學習者外語磨蝕的相關因素。劉宇紅基於隱喻的理論探討，對各類隱喻形式的結構、特性和解讀規律進行了多視角的深入探討。

《語言科技文庫》所收論著，由作者在2008年12月申報選題，2011年始逐步完稿。該系列叢書的主編審讀了部分書稿，主要就其學術價值、章節安排、內容關聯、行文表述、圖表繪製等方面，提出審閱意見。此後，作者們對書稿又進行了修改和潤色。《語言科技文庫》的作者，大多數是具有博士學位的年輕教師。對於我們這些20世紀80年代走進語言學研究領域的前行者而言，出版論著可能已不足為道。然而，對於年輕學者而言，其論著的出版既是幾年來研究的結晶，也是對其繼續探索的促進。換言之，“211工程”重點學科建設的目的之一，就是為年輕教師搭建一個可持續發展的科研和教學平臺。學科帶頭人的主要任務之一就是提攜後進。

儘管從根本上來說，科學或學術研究是一種個人的探索行為，然而複雜問題的研究，無疑需要群體協作。“學科建設”或團隊合作模式，是20世紀90年代後期出現的一個新概念。這種模式涉及總體規劃、多方協調，是需要付出精力和心血的。2008年，通過投票方式，我被推舉擔任該項目總負責時，就意識到自己成了一個“勞動班委”。2009年，前往安徽大學拜訪黃德寬教授時，我曾談到“學科負責人的任務就是規劃設計，爭取項目經費和提供科研設備設施”，得到黃教授的贊許。2010年，申報江蘇省高校哲學社會科學重點研究基地時，評審專家柳士鎮教授提問的“作為一個交叉學科項目，各學科之間的協調是怎麼考慮的，有什麼做法”，可謂一語中的。作為後學，深知交叉研究之艱、學科整合之難。相關學科之間的整合協調需要借助行政機制，但憑藉行政方式並非就能完成。當時的回答是，目前做到的是建成了一個可以合作研究的場所，至於學科之間的進一步溝通合作應有較長過程。有一點很明確，只有通過交叉項目，相應學科纔能滲透，合作者纔能逐步磨合。我們只是在一步步探索。

十一五期間的“211工程”建設項目即將完成，但是學科建設的任務並沒有結束。2010年，“語言信息科技研究中心”被評審為江蘇省高等學校哲學社會科學重點研究基地，為“語言科技”這一交叉領域注入了新的建設活力。重點研究基地建設，除了“跨院系整合、多學科交叉、開放型運行”理念，需要凸顯“合作性攻關”。圍繞交叉性項目，實施計算語言學、語音科技、神經語言學、語義科技等力量的聯合攻關計畫。只有通過全面開放，以及和與國內外同行的合作交流，纔有望建成具有影響的語言科技研究、人才培養和學術交流基地。

十年前（2001年），我曾寫道，“語言科技”的內涵是以理論研究為指導，以描寫研究為基礎，以應用研究為樞紐，促使語言研究向計算機應用、認知科學和現代教育技術領域等延伸，溝通文理工相關學科以實現語言研究過程及其成果的技術化。“語言科技”的外延為語言工程科技、語言教育科技和語言研究科技。其中，“語言研究科技”是將語言研究活動與資源建設、軟件開發相結合，其目標是實現語言學自身的科技化。還應包含語言實驗、數據處理這些實驗語音學、神經語言學研究的科技手段。

雖然語言學家不可能也不必要都轉向語言計算或實驗研究，儘管描寫、考據和內省始終是最基本的方法，但是具有一定的語言科技意識卻非常必要。語言學家只有了解有哪些可供利用的資源、軟件或儀器，纔能提高其研究深度、精度和效率。語言學家也只有了解到信息處理的語言研究需求，纔有可能為之提供可資應用或參考的基礎成果。“語言科技”是21世紀語言學研究的潮流。

此為出版緣起。是為總序。

2011年8月謹識於南都