

序

何志華 鄭麗娟

《尚書》，或稱《書》，是儒家經典之一。在漢代立為學官之前，《尚書》文句已多以先王遺訓的形式，廣泛為先秦典籍引用，對秦漢的政治、思想均有深刻影響。漢代《尚書》有今、古文之學。今文《尚書》傳於伏生。伏生名勝，為秦博士。始皇焚書，伏生壁藏《尚書》。及漢定天下，伏生求其書，得二十九篇，教於齊魯之間，傳歐陽及大、小夏侯氏。伏生所傳《尚書》，以隸書寫成，故稱為今文《尚書》。古文《尚書》出於孔壁。《漢書·藝文志》記武帝末，魯恭王壞孔子宅，得古文《尚書》、《禮記》、《論語》、《孝經》等典籍。諸書皆以先秦之古文字寫成，故稱為古文。孔安國，孔子之後，得壁中《書》獻之。會巫蠱事，未列於學官。及後，孔安國所傳古文《尚書》散佚。晉元帝時，豫章內史梅賾進獻《孔傳古文尚書》四十六卷，五十八篇。自宋吳棫以降，歷代皆有學者懷疑梅賾進獻之《尚書》乃偽書。

《尚書》是上古之書，記三代文告及對話，反映上古的語言狀態。近年有關《尚書》語言學研究，有錢宗武的《今文尚書語法研究》、《今文尚書語言研究》及《今文尚書詞彙研究》。錢氏之書以三十三篇今文《尚書》為研究重心，探討今文《尚書》語言特點，而摒棄歷來備受懷疑的古文《尚書》。檢張岩《審核古文〈尚書〉案》，以《尚書》之字頻檢索先秦兩漢二十餘種文獻，並考察先秦兩漢文獻對古文《尚書》「引文」和「用文」情況，以證古文《尚書》並非偽撰。由此可見，全面研究《尚書》詞彙，除可了解上古語言特色外，更能進一步探究其成書年代。

向以為《尚書》以單音節詞為主，但經全面圈選及檢查，其實不然。《尚書》語言，或屬複音節詞萌生的年代，複音詞的組成方法多樣。《尚書》既有後代沿用不絕，為人熟知的詞彙，如「多材多藝」（〈金縢〉）、「無偏無陂」（〈洪範〉）、「義刑義殺」（〈康誥〉）、「珍禽奇獸」（〈旅獒〉）、「玉石俱焚」（〈胤征〉）等；也有些獨見於《尚書》而鮮見於後世文獻的詞彙，如「大邦殷」（〈召誥〉）、「天邑商」（〈多士〉）、「殷先哲王」（〈康誥〉）、「商耆成人」（〈康誥〉）、「舊有位人」（〈微子〉）等。

本書在漢達文庫數據庫（<http://www.chant.org>）之基礎上，甄別適當材料，設計中國古代雙音節及多音節詞彙資料庫自動編纂程式系統（Automatic Compiling Program System for Retrieval Database for Ancient Chinese Disyllable and Polysyllable Words，簡稱 RDDPW），分析

語言材料，統計詞頻，檢出各類詞頻計量清單，對文獻進行縱向與橫向比較。透過電腦程式之計算與分析，並參考大型詞典、專書詞典所收錄的詞彙，重新建構「完整詞單」，收錄合共 551,856 個二至四音節詞彙，然後全數「注入」漢達文庫先秦兩漢一切傳世文獻文檔中，核算數據。藉由上述檢索過程，檢得《尚書》複音節詞彙 5,285 個，總詞頻量為 9,776。在此原始資料上，逐一判斷詞彙組合是否成詞，抑係不成意義之偶合。最後得出《尚書》複音節詞彙 4,043 個，總詞頻量為 7,383。本書盡量涵蓋《尚書》所有使用之詞語作為詞條收錄，冀能揭示《尚書》語料之面貌、特徵、發展演變及規律，表現其普遍性與獨特性，以反映其時代特色。

一般合成詞之形成，大都經歷由短語邁向詞彙之轉化過程，尤其先秦兩漢典籍之中，不少詞彙組合處於「或詞組」、「或詞」之兩可狀態，即同素異序之詞彙。對於此類組合，本書採取標準從寬，甚至斟酌收錄為大型詞典所不取之短語或詞組。此外，收錄詞語不應拘泥於某一時代及字形，因語言既有繼承，亦有演化，詞語或產生於上古，但廣泛使用於中古；或肇始於中古，而流行於近代。衍變過程中，字形或出現通假字、古今字等，本書盡量予以收錄，加入現代常用詞彙，並據此檢索文獻。如此方有利於語詞之溯源，更有利於漢語詞彙之歷時研究。

本書以香港中文大學中國文化研究所編纂之《先秦兩漢古籍逐字索引叢刊》為基礎，並參考別本異文，斟酌補錄。再利用研究所「漢達文庫」電腦檢索系統進行排列統計，然後經由研究人員編纂《尚書》專用詞彙，及《尚書》詞彙僅見單一先秦兩漢典籍的材料。書中凡涉及成詞與否之爭議，編者均採寧濫勿缺原則，權將有關詞彙資料收錄。

此書之編著，得副編輯李洛旻博士及助理電算師李寶華先生兩位襄助，謹此致謝。本書草創，疏漏難免，尚祈讀者不吝賜正，以匡不逮。