

序

何志華 朱國藩

《孔叢子》主要記錄孔子及其後代之言行，反映了孔氏家學的特點。全書共二十三篇：第一篇〈嘉言〉至第五篇〈刑論〉，記孔子之言行；第六篇〈記問〉至第十篇〈抗志〉，記子思之言行；第十一篇〈小爾雅〉，收錄經傳字詞之詁訓；第十二篇〈公孫龍〉至第二十一篇〈答問〉，記孔穿、孔謙、孔鮒等人之言行；第二十二、二十三篇為《連叢子》上下，收錄漢代孔子後人的言論及作品。

此書最早見於《隋書·經籍志》，題為孔鮒所撰。因書中不乏孔鮒身後之事，故在流傳的過程中，多被視為偽書。近世學者經過嚴密考證，普遍認為《孔叢子》並非偽書，其中黃懷信以為編者當為孔扶，¹ 而傅亞庶則從冢田虎之說，以為《孔叢子》蓋由孔季彥之弟子輯成，² 兩說可作參考。總之，《孔叢子》之編者當為孔氏後人，而非一人一時之作。

前人早注意《孔叢子》在中國古代語言學史上的意義，然多從傳統訓詁學角度分析《小爾雅》，³ 而非對《孔叢子》全書詞彙作出研究。自上世紀八十年代起，古漢語複音節詞研究日趨成熟，成果纍纍，舉其要者，有程湘清〈先秦雙音節詞研究〉、〈《論衡》複音詞研究〉、伍宗文《先秦漢語複音詞研究》以及周俊勛《中古漢語詞彙研究綱要》等。⁴ 惜乎相關著述大多總論古漢語複音節詞的構成方式，或初步比較先秦兩漢典籍複音節詞的數目，鮮有全面探究《孔叢子》詞彙的特性。故此，本書希望重新統計、分析《孔叢子》詞語，臂助學者進一步勾勒上古複音節詞發展的進程。本書在漢達文庫數據庫（<http://www.chant.org/>）的基礎上，甄別適當

¹ 黃懷信：〈《孔叢子》的時代與作者〉，載黃懷信：《古文獻與古史考論》（濟南：齊魯書社，2003年），頁1-14。

² 傅亞庶：〈《孔叢子》的成書年代與真偽〉，載傅亞庶：《孔叢子校釋》（北京：中華書局，2011年），〈附錄四〉，頁605-14。

³ 如宋翔鳳《小爾雅訓纂》、胡承珙《小爾雅義證》、朱駿聲《小爾雅約注》等。

⁴ 程湘清：〈先秦雙音節詞研究〉，載山東社會科學院語言文學研究所編：《先秦漢語研究》（濟南：山東教育出版社，1982年），頁44-112；程湘清：〈《論衡》複音詞研究〉，載程湘清編：《兩漢漢語研究》（濟南：山東教育出版社，1982年），頁262-340；伍宗文：《先秦漢語複音詞研究》（成都：巴蜀書社，2001年）；周俊勛：《中古漢語詞彙研究綱要》（成都：巴蜀書社，2009年）。

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），從而分析語言材料，統計詞頻，檢出各類詞頻計量清單，對文獻進行縱向和橫向的比較。透過電腦程式的計算與分析，並參考大型詞典、專書詞典所包含的詞彙，重新建構「完整詞單」，收錄合共 350,343 個二至四音節詞彙，然後將此批詞彙全數「注入」漢達文庫先秦兩漢一切傳世文獻的文檔中，核算數據。通過上述檢索過程，檢得《孔叢子》雙音節或多音節詞彙 5,453 個，總詞頻量為 11,336。在這原始資料之上，我們再逐一判斷各個詞彙組合是否成功構詞，抑或只是不成意義的偶合。最後得出《孔叢子》詞彙 3,830 個，總詞頻量為 7,979。本書盡量宏觀地涵蓋《孔叢子》所有使用的詞語作為詞條收錄，希望能揭示出《孔叢子》語料的面貌、特徵、發展演變和規律，表現其普遍性與獨特性，以反映其時代特色。

一般合成詞的形成，大都經歷了由短語邁向詞彙的轉化過程，尤其是兩漢的詞彙正處於一個由單音節詞邁向複音節詞急劇變動的年代，很多詞彙組合處於「或詞組」、「或詞」的兩可狀態，即同素異序的詞彙。對於這類的組合，我們採取從寬的標準，甚至適當收錄一些不被大型詞典考慮的短語或詞組。還有，收錄詞語不應拘泥於某一時代及字形，因為語言是有繼承與演化的，有些詞語或許產生於上古，但廣泛使用於中古；有的肇始於中古，但流行於近代，過程中，字形或有通假字、古今字等問題，我們希望一一收錄，加入一些現代常用詞彙去檢索文獻，因為這樣有利於語詞的溯源，更有利於漢語詞彙的歷時研究。

本書以香港中文大學中國文化研究所編纂之《孔叢子逐字索引》為底本，利用研究所「漢達文庫」電腦檢索系統將詞彙輯錄，並統計不同數據，然後經由研究人員編纂《孔叢子》專用詞彙表，和《孔叢子》詞彙僅見單一先秦兩漢典籍的材料。書中凡涉及成詞與否的爭議，編者均採寧濫勿缺的原則，權將有關詞彙資料收錄。

此書之編著，得博士後研究員梁德華博士、研究助理潘祉渝女士、陳柏嘉先生、吳家怡女士，以及助理電算師李寶華先生、何銘怡女士之助，謹此致謝。本書草創，疏漏難免，尚祈讀者不吝賜正，以匡不逮。