

華語八千詞說明

張莉萍

台師大國語教學中心

放在網站上的這個八千詞，主要是我 2003-2004 年受國科會（NSC-92-2411-H-003-045）補助的研究計畫成果。近年來，不斷聽到有人催促我公開這個詞表，尤其來自華語教師與學習華語的外國學生的殷切期盼，讓我無法拒絕。雖然總覺詞表不完美之處尚多，身邊的雜務也讓我沒時間做一進步檢視，但基於服務的理念，也只能硬著頭皮放在網站上，期望對需求者有所幫助。也竭誠歡迎各界的批評指教，只是我的時間有限，恐怕無法一一回應大家的批評，並且即時更新。畢竟對我而言，這是一個基礎研究，以這個方法所做出來的結果呈現給大家。如以每個人對詞彙的概念，恐怕初級詞彙應該有什麼詞或哪一個詞應該放在哪一級，不同人有不同看法，沒有定論。

這個詞表主要選詞來源為中研院核心詞彙、通用詞彙、參考詞彙表、TOP 詞彙表、HSK 詞彙表，運用相對頻率、加權值的方式算出每一個詞的比重，進而依據華語文學習里程，制定出初級 1500 個詞彙、中級 3500 個詞彙、高級 5000 個詞彙。詞表內容包括詞彙本身、每個詞的等級、語法類。

本研究的詞彙分級是為了服務華語教學，同時作為華語能力測驗的基本參考詞彙。因此在界定分級詞彙量時，主要以學習者為依歸，所以訂出初級語言能力詞彙量約 1500 個，中級能力詞彙量為 5000 個，高級能力詞彙量為 8000 個。依照分析後每個詞的總值排序，取前 1500 名的詞為初級詞彙、第 1501-5000 名的詞彙為中級詞彙，第 5001-8000 名的詞彙為高級詞彙。之所以取詞彙量八千，是採取鄭錦全（1998）「詞涯八千」的概念，從計量的觀點來解釋人類的語言認知能力。

以下是關於這個詞表的相關訊息，

詞彙編列原則

1. 異體字或兒化韻列在同一詞條上。例如：裏/裡、大部分/份、藥/葯、畫畫（兒）、玩（兒）、花（兒）、小孩（子/兒）。
2. 不同義涵，分列不同詞條。例如：「叫」在初等詞表有四個詞條，依使用頻率排序，分別代表「命令、要求」「叫做」「大叫」「點菜」等義涵。
3. 不同詞性（語法功能），分列不同詞條。例如：喝「一點」水（數量詞）、小心「一點」（程度副詞）；養寵物就是要有「耐心」（名詞）。要了解孩子，就必須「耐心」聽他們說話（副詞）。
4. 某些詞類在華語教學裡是一個概念的，不再區分為不同詞類，也就不列不同詞條，例如：上（Ncd 地方名詞，Ng 方位詞）。那（Nep 指示代名詞，Nes 指示定詞）；某些介詞和動詞不做區分，例如：向、為（了）。
5. 意義容易衍生的詞，為經濟原則不另列詞條，例如：大海、大樹、黑色、紅色、妳們（可由「你們」推衍）。
6. 台灣日常詞彙如因兩岸用語不同而產生排序上的落差，仍依照台灣使用頻率收錄在初等詞彙。例如：「午飯」在初等詞彙，而「午餐」因不列在 HSK 詞表裡所以算出的值重較低，排名 12343

名。反之，「領導」在 HSK 列於初等詞彙，在這個詞表中則列在中等詞表。

7. 有些詞在一般辭典中不列，但為了教學上的便利，此詞表把他們當成一個單位來處理。例如：快要、就要。
8. 課室用語不列入初等詞表中，例如：聽寫、注音、拼音等等。
9. 等級代號：「B」表示初等詞彙、「I」表示中等詞彙、「A」表示高等詞彙。
10. 詞表中各詞類分佈情形如下表：

	代號	初等詞表	中等詞表	高等詞表	小計
名詞(Noun)	(N)	613	1322	1173	3108
動作動詞(Active verb)	(VA)	294	907	892	2093
狀態動詞(Stative verb)	(VS)	256	771	667	1694
副詞(Adverb)	(Adv)	123	249	123	495
連接詞(Conjunction)	(Conj)	20	52	22	94
介詞(Preposition)	(Prep)	22	32	26	80
量詞(Measure)	(M)	79	69	31	179
其他類(Others)		93	98	66	257
總計		1500	3500	3000	8000

TOP vs. HSK 詞彙的分佈交集情況

HSK ∩ TOP		TOP 詞彙表 (8,000 個)		
		初(1,500)	中(3,500)	高(3,000)
HSK 詞彙表 (8,822 個)	甲(1,033)	959	153	9
	乙(2,018)	354	1399	191
	丙(2,202)	32	899	854
	丁(3,569)	13	438	686

說明：

HSK 甲、乙、丙、丁各級詞數和 TOP 初、中、高三級詞彙表所收集的各級詞數分列於上表。以 HSK 的甲級詞和 TOP 的各級詞彙為例，同時出現在 HSK 甲級和 TOP 初級的詞彙共有 959 目詞，出現在中級的詞彙有 153 目詞，而同時出現在 HSK 甲級和 TOP 高級的詞只有 9 個詞，以此類推。

有趣的結果：

- 嘴巴、支票、傳真、現金、房租、發票、房東、菜單、奶粉、或是、公分、晚餐、機車（出現在初等詞表的丁級詞）
- 便宜（名詞用）、咱、活兒、實踐、聯繫、聽寫、喊、組織、關係（動詞用）（出現在高等詞表中的甲級詞）
- 政府、經驗、知識、經濟、技術、政治、藝術、發展、條件、工業（出現在中等詞表中的甲級詞）