

# 語料庫在辭典編輯上的運用

張麗麗\*、陳克健\*、黃居仁\*\*

\*中央研究院資訊科學研究所 \*\*中央研究院歷史語言研究所

張麗麗：lili@iis.sinica.edu.tw  
陳克健：kchen@iis.sinica.edu.tw  
黃居仁：hschuren@ccvax.sinica.edu.tw

## 摘要

這篇文章介紹了詞庫小組利用中研院語料庫編輯辭典的方法和步驟，同時也探討語料庫對於辭典編輯的貢獻。我們所編輯的是量詞辭典和名量搭配辭典，兩本辭典是一體兩面，內容彼此相關。利用兩個資料庫和一個編輯介面系統，我們可以在線上直接抽取、統計、篩選相關語料，並且可以在線上進行編輯的工作，包含鍵入、轉換、挑選、排序、統計、篩選、傳送、增刪、複製、搬移、輸入等工作。根據我們的經驗，語料庫對於辭典編輯最大的貢獻在於它能提供各式各樣的語言事實，使得辭典能夠更廣泛、深入地記載每個詞彙的用法。豐富例句幫助編輯人員掌握每個詞彙全面的、當代的用法。各種詞頻資料是許多道編輯步驟的重要依據，像是決定詞項、選取例詞、排列例詞、調整用法分析的順序等。辭典中的例詞例句也都是從語料庫中擷取出來的，比起編輯人員自己造的例子，更能反應語言實際的使用情形。此外，由於我們使用的是標記語料庫，詞和詞之間是一個一個分開的，並標有詞類，搜尋起來更加方便有效。

## 1. 簡介

由於電腦發展一再突破了容量和速度的極限，因此利用電腦來處理大量語料才變成可能。近年來，語料庫為本（corpus-based）的研究成為語言學及計算語言學的一個重要發展，甚至也拓展到應用語言學的領域上。1987年英國辭典公司 Collins 出版了世界上第一本利用語料庫編輯的辭典（Sinclair 1987b），開啓了新的辭典編輯方式，更開創了新的辭典風格。<sup>1</sup>

Collins 公司委託英國伯明罕大學英文系成立了一個 COBUILD 專案計畫。<sup>2</sup> 首先他們利用該系發展多年的語料庫針對每個詞項作索引檔。其次，編輯人員利用索引查詢語料庫的資料以及其他相關資料，並依照規畫好的格式撰寫每個詞項的基本資料，包含詞項、發音、各種詞形、意義、詞類、例句…等等。然後再人工鍵入這些資料，建立了基本詞彙資料庫。這個基本詞彙資料庫是該專

案計畫的核心，儲存了每一個詞彙的相關訊息，可以發展成各式各樣的辭典、文法書、語言教學用書……等等。<sup>3</sup> 下一步就是由電腦從基本詞彙資料庫中抽取適當資料組合成辭典或書本的形式，放入排版系統。編輯人員直接進入排版系統進行潤飾、修改的工作，完成編輯程序。該專案計畫的第一個成果就是這本極受歡迎的 Collins Cobuild English Language Dictionary。

基本說來，這是工具的突破帶動了辭典編輯方式和風格的突破。從整個編輯流程來看，電腦、機讀語料庫和機讀資料庫是新引入的編輯工具。究竟這些新的工具起了什麼樣的作用呢？該辭典的主編 John Sinclair (Sinclair 1987a pvii) 認為機讀語料庫使得編輯人員「第一次能夠廣泛且深入地觀察語言」，並且「對於語言中最常用的表達方式有了成千上萬的新發現。」<sup>4</sup> 機讀資料庫成為另一種形式的語料來源，電腦可以依照資料庫中的詞類或語法屬性分類查詢，「使得編輯人員可以檢查分類上的一致性，或是經由比對修正某個詞的分析。」(Sinclair 1987a p42) 電腦除了在機讀語料庫和機讀資料庫中建索引資料、作搜尋、排序、分類、統計…等工作外，還具有各式各樣的編輯功能，像是挑出拼寫錯誤、核對交錯索引、排版…等，甚至還可以成立電子郵件中心，流通編輯人員的提示。

詞庫小組已於一九九五年發表推出「中研院平衡語料庫」(Sinica Corpus)，是世界上第一個有完整詞類標記的漢語平衡語料庫(黃居仁等 1995、詞庫小組 1995)。該小組並於該年七月受國語日報社委託展開利用語料庫編輯辭典的計畫，預計一九九六年八月完成。早期詞庫小組已經採用語料庫進行詞彙分析，但是利用語料庫編輯辭典還是第一次嘗試。由於人力、時間及經費上的限制，我們必須從規模較小的辭典開始。我們使用的是標記語料庫，可以挑出特定的詞類來編輯，因此我們和國語日報社商定編輯兩本相關的辭典：量詞辭典和名量搭配辭典。就我們所知，在漢語領域中，這是首度利用語料庫來編輯辭典。我們依照語料庫的特性、漢語的特性以及所編辭典的特性，為這兩本辭典設計了新的編輯方法和流程，在這篇文章中將為各位一一介紹。除此之外，我們也會就這次編輯經驗討論語料庫對辭典編輯的貢獻。

## 2. 量詞辭典與名量搭配辭典的編輯

### 2.1 構想

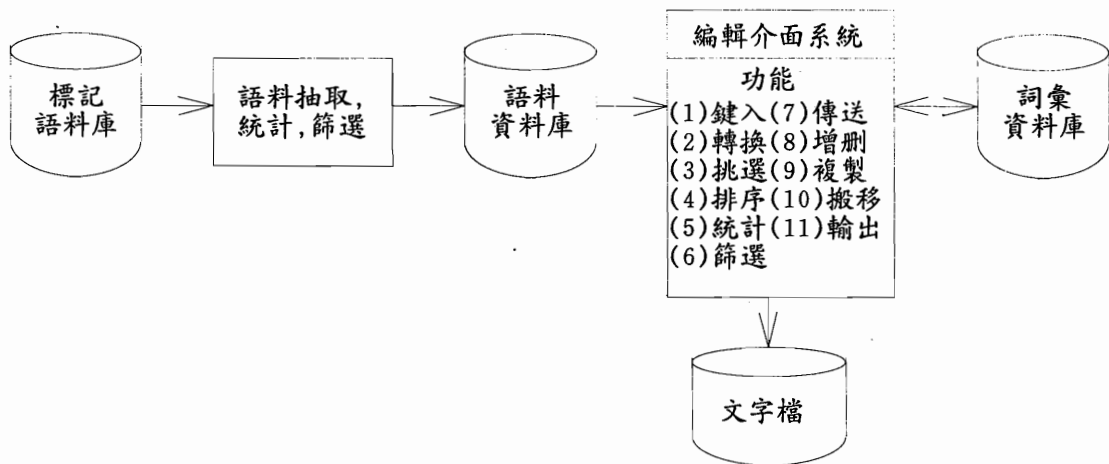
我們所要編輯的量詞辭典以及名量搭配辭典是一體兩面。量詞辭典是從量詞出發，探討每個量詞的用法以及可以搭配的名詞類型。名量搭配辭典是從名詞出發，探討每個名詞能夠搭配的量詞。在名量搭配辭典中，我們有一個創新的作法，就是依名詞尾將名詞分類，<sup>5</sup> 具有相同詞尾的名詞再分成次類，<sup>6</sup> 只需針對同一組的名詞標上適用量詞就可以了。

根據編輯體例（請參見附錄一），在量詞辭典中每個詞項下包含：量詞項、注音、分類、量詞次類標記、語意說明、例詞、例句、辨析及總辨析（可省略）；在名量搭配辭典中每個名詞尾下包含：名詞尾項、注音、語意分類、語意說明、次分類、例詞、所搭配量詞、辨析（可省略）。依照上述這兩本辭典的架構，語料庫可以提供下列資料：一、詞項清單：量詞、名詞、名詞尾；二、詞頻資料；三、名詞量詞搭配資料；四、例詞和例句。

### 2.2 方法

COBUILD 小組的編輯流程並非完全電子化，從語料庫到詞彙資料庫這一段是離線作業的。他們的作法是由編輯人員觀察語料庫中的資料，用紙和筆撰寫每個詞項的基本資料，再鍵入電腦中，建立基本詞彙資料庫。他們採取這樣的方式是可以理解的，因為每一個詞項都可能有成千上萬個例子有待觀察，而且除了例句外，語料庫並不能直接提供所需資料，都需要編輯人員融會貫通之後才能整理出基本資料。但是這樣的編輯流程將會增加鍵入及校正的工作。我們決定將這個步驟也電子化，也就是說我們可以直接將語料庫中的資料轉到詞彙資料庫中。因為這個編輯計畫所面臨的情況不同：我們所使用的是標記語料庫，可就需要挑選出最適當的資料；尤其是就名量搭配辭典這一部份來看，編輯只要參考語料庫中搜尋出的相關資料即可，而且只要加以適當的分類、編排，可以直接用於辭典中；再加上量詞辭典和名量搭配辭典的規模不大，就記憶空間和速率來看都不成問題。

我們的作法是在語料庫和詞彙資料庫之間多建立了一個語料資料庫，並加上一個編輯介面系統，以便傳送語料庫中的資料到詞彙資料庫中。因此針對每個詞項必須建立兩個資料庫：語料資料庫和詞彙資料庫。語料資料庫儲存從語料庫中抽出的相關資料，詞彙資料庫存放編輯過的資料。透過編輯介面系統可以直接擷取語料資料庫中的資料放置到詞彙資料庫中適當的欄位。詞彙資料庫的格式是依據辭典所需內容規畫好的，其中資料可以轉成辭典形式的文字檔，交給排版系統處理。語料庫和這兩個資料庫的對應關係如圖一所示：



圖一：三個資料庫間的關係以及編輯流程

### 2.3. 編輯流程

從圖一可以清楚看到建立詞彙資料庫的流程包含兩個步驟：第一個步驟是從標記語料庫中抽取語料，建立語料資料庫；第二個步驟是透過一個編輯介面系統建立詞彙資料庫。第一個步驟大部份是由電腦完成的，包括從大量語料中抽取所需資料，並進行排序、統計等工作。只有在少數情況下才需要人工進行篩選。第二個步驟則需要人工編輯詞彙的每項訊息。有一部份訊息，像是解釋、用法說明必須完全靠人工鍵入；但是有一部份訊息，像是例詞例句，卻可以到語料資料庫中挑選，再傳送到詞彙資料庫。

詞彙資料庫建立之後還有一個步驟，就是從詞彙資料庫中抽取資料轉成純文字檔，再交給排版系統排版。這個步驟也是在編輯介面系統中進行。

### 2.3.1 步驟一：建立語料資料庫

語料資料庫中存放的資料都是直接從語料庫中得出的資料，包含：一、詞項清單：量詞、名詞、名詞尾；二、詞頻資料；三、名詞量詞搭配資料；四、例詞和例句。其中只有只有例詞例句另外需要人工篩選，因為例詞例句的數量太大。例如最常用的量詞「個」在兩百萬詞的語料庫中出現15495次，次常用量詞「種」也出現4666次。因此我們特別設計了一個小程式，由編輯人員在語料中挑選出涵蓋各種用法的適量的例詞例句，加上規定的符號，啓動傳送功能，就會將這些例子傳送到語料資料庫中。建立語料資料庫包含了下列幾個步驟：

#### a. 詞項搜尋：

找出中研院語料庫中所有的量詞：537個量詞，共58,615例。並計算每個量詞的出現頻率。（請參見附錄二）

找出中研院語料庫中所有的常用普通名詞：30,888個名詞，共378,768例。另外補充2,167個名詞，<sup>7</sup> 共計33,055個名詞。並計算每個名詞的出現頻率。

將33055個名詞依名詞尾排序比對，共得2,787個詞尾。（請參見附錄三）

#### b. 搜尋名詞、量詞搭配關係

找出量詞右方5個詞內出現過的名詞的例子：502個量詞，共46,569例。

刪除名詞出現在某些標點符號（如，。：；？！）之後的例子，及名詞或量詞為外文的例子。

依上述資料建量詞和名詞的搭配資料：共計53,872對

找出的53,872對名量搭配資料將分別放入各個量詞和名詞的語料資料庫。

#### c. 挑選例詞和例句

從語料庫中挑選涵蓋各種用法的例詞和例句，標上規定的符號。

將資料傳送至語料資料庫。

#### d. 建立量詞和名詞尾的語料資料庫

量詞的語料資料庫包含：量詞項、詞頻、例詞、例句（由於有例詞例句，所以不將所搭配的名詞列入<sup>8</sup>）（請參見附錄四）

名詞尾的語料資料庫包含：名詞尾項、詞頻、組成名詞、詞頻、每個名詞所搭配量詞（請參見附錄五）

### 2.3.2 步驟二：利用編輯介面系統建立詞彙資料庫

為了編輯流程全面電子化，我們特別設計了一套編輯介面系統，可以讓編輯人員直接編輯詞彙資料庫。該編輯介面系統包含以下幾項功能：(1) 鍵入、(2) 轉換（進入語料資料庫）、(3) 挑選（挑選例詞和例句）、(4) 排序（就選取的例詞排序）、(5) 統計（統計同一組名詞所搭配的量詞及出現次數）、(6) 篩選（就統計出的量詞資料進行篩選）、(7) 傳送（將語料資料庫中挑選出且排序篩選過的語料傳到詞彙資料庫中適當的欄位）、(8) 增刪（增加或刪除整項分析）、(9) 複製（複製整項分析）、(10) 搬移（搬移整項分析）、(11)輸出（將資料轉成辭典形式的文字檔）。

在該系統中編輯人員可以隨時進入語料資料庫觀察語料，構想整個詞彙的分析架構，然後在編輯介面系統直接鍵入各項資料，或到語料資料庫挑選適當資料傳送到詞彙資料庫中。傳送資料的過程如下：(1)進入語料資料庫、(2)挑選例詞或例句、(3)編輯系統自動依照頻率將例詞例句排序、(4)編輯系統主動提出重新排序功能、<sup>9</sup> (5)編輯人員可以重新排序也可以跳開、(6)最後將挑出的且排好序的資料傳送到詞彙資料庫。在名量搭配辭典中，排序後還要進行兩個步驟才能傳送。在名詞尾的例詞排序後，編輯系統會主動替放在同一組的名詞累計所搭配的量詞及其頻率，並將統計出的量詞資料依頻率排列。這份資料的錯誤率比較高，所以統計之後，也提供篩選的功能。<sup>10</sup> 編輯人員篩選完畢才會將例詞和量詞一起傳送到詞彙資料庫（整個傳送過程請參見附錄六）。除此之外，針對整項分析，編輯系統還提供增加、刪除、複製和搬移的功能。

詞彙資料庫各項資料的建立大致可以利用鍵入和傳送的方式完成，如表一和表二所示。

表一：量詞辭典詞彙資料庫的資料建立方式

量詞辭典	電腦自動存入詞彙 資料庫中	編輯人員構想鍵入 詞彙資料庫中	編輯人員從語料資料庫 中挑選，再由電腦傳送
量詞項	* (傳送)		
注音	* (傳送)		
分類		* (鍵入)	
量詞次類標記		* (鍵入)	
語意說明		* (鍵入)	
例詞			* (傳送)
例句			* (傳送)
辨析及總辨析 (可省略)		* (鍵入)	

表二：名量搭配辭典詞彙資料庫的資料建立方式

名量搭配辭典	電腦自動存入詞彙 資料庫中	編輯人員構想鍵入 詞彙資料庫中	編輯人員從語料資料庫 中挑選，再由電腦傳送
名詞尾項	* (傳送)		
注音	* (傳送)		
語意分類		* (鍵入)	
語意說明		* (鍵入)	
次分類		* (鍵入)	
例詞			* (傳送)
所搭配量詞		* (鍵入)	* (傳送)
辨析(可省略)		* (鍵入)	

雖然編輯人員需要鍵入的資料不少，但是這些都是需要靠觀察語料資料庫中的語料才能做出的分析。也就是說語料資料庫中的語料不是只有提供資料傳送到詞彙資料庫而已，還能供編輯人員觀察語言的使用情形。語料資料庫中的例詞例句及搭配資料都是根據需要精挑出來的，除去了冗贅的訊息，使得編輯人員在觀察和歸納語言現象時更加便利。例如在替名詞尾的語意作分類時，只要觀察包含該名詞尾的所有名詞即可，根本不需要觀察例句。又例如在分析量詞的用法時，由於語料資料庫中的例詞和例句是人工篩選出來的，能夠充分反應該量詞的各種用法，所以只要觀察這些挑出的例子就可以了。

### 2.3.3 步驟三：從詞彙資料庫中抽取資料轉成純文字檔，放入排版系統

編輯系統還提供了一個功能，就是從詞彙資料庫抽取資料，並依照編輯體例轉成類似辭典形式的純文字檔（請參見附錄七）。編輯人員可以利用文字檔進行集體校閱的工作。許多問題和現象在分析單個量詞時是看不出來的，但是有了整體的分析資料，要進行各種觀察和比較就十分容易。更正的工作仍然回編輯系統中進行，修改完畢再轉成新的文字檔。最後，內容正確無誤的文字檔才交給排版公司，放入排版系統排版。

## 3. 語料庫對辭典編輯的貢獻

近年來，語料庫為本(corpus-based)的研究成為語言學及計算語言學的一個重要發展（Svartvik 1992，Church and Mercer 1993，陳克健 1994，黃居仁 1995）。語料庫對於各種研究的最大貢獻在於它提供了豐富的語言事實，而且由於是機讀形式，可以迅速有效執行各式各樣的搜尋、排序及統計的工作。對於辭典編輯，它最大的貢獻也在此。在傳統編輯辭典的工作中，蒐集資料是件重要但是十分不容易的工作，要從浩瀚的語言中抓取一個詞彙的各種使用實例幾乎是不可能做到的。因此辭典內容往往只是反應編輯人員本身的語文素養、對語言的觀察能力、以及考證其他辭典或相關書籍的功夫，究竟一本辭典反映了多少語言實況根本無從得知。語料庫卻完全突破了上述的限制，編輯人員首度能夠充分掌握實際語料，做出準確、客觀、有根據的詞彙分析，使得辭典在反應語言實況上有了重大的進步。同時由於是機讀語料庫，在搜尋資料的效率和速度上也大大超越傳統方式的極限。

在實際編輯過程中，我們發現在編輯的每一個細節中都可能需要參考語料庫所提供的各種資料，因為語料庫所能提供辭典編輯的語言事實是多方面的：

#### a. 全面反應每個詞彙的各種用法。

有些詞彙的用法是非常複雜的，沒有掌握充分的語料，就沒辦法將該詞彙準確分析。以「片」為例，它可以用來計量許多的事物，有具體的，像是「幾片落葉、兩片玻璃、一片江山、一片茶園」，也有抽象的，像是「一片歡樂，



一片熱心、一片朝氣」，甚至於還有介於抽象和具體之間，像是「一片不景氣、一片陰影、一片笑聲」。沒有將各種用法蒐集齊全並經過仔細的觀察，我們很難將「片」的用法作一番詳盡且深入的描繪。語料庫可以很快的將出現量詞「片」的所有例句馬上找出，編輯人員才能免除以偏蓋全的錯誤。

試比較我們所編出的量詞辭典和其他的量詞辭典，可以看到一個明顯的差異：我們所分析的用法總是比較周全。以「道」為例，大陸編的「現代漢語量詞手冊」（郭先珍，1986）分析「道」用來計量三種物品：

- 1 計量江河或某些長條狀的東西。
- 2 計量門、牆或類似門牆的東西。
- 3 計量命令、題目等。

而我們從語料庫中所觀察到的「道」卻可以計量六種物品（編號前加上\*的符號表示是本辭典才有的用法分析）：

- 1 用來計量長條形物品。
- \*2 用來計量成線狀的光線。
- 3 用來計量門或牆。
- 4 用來計量題目、命令。
- \*5 用來計量菜餚。
- \*6 用來計量程序。

我們從語料庫中所觀察到的豐富語言事實除了反映在用法分項上，也反映在「辨析」欄中。例如「片」和「聲」都可以用來計量聲音，但是除了意義不同外，用法也不同：

片：辨析：……計量聲音最常用的量詞是「聲」，但是它是個別計量每一次發出的聲音，如「三聲槍響、幾聲狗吠」，不像「片」是用來描述同時發出的大量聲音。因此，「聲」的前面可以接各種數詞，而「片」只能接「一」。請參見「聲」。

b. 忠實反應每個詞彙的當時用法。

試比較大陸出的「現代漢語量詞手冊」（郭先珍，1986）以及教育部1995年10月出的「常用量詞手冊」（教育部國語推行委員會，1995）和我們所編的量詞辭典，就以「支」為例。郭先珍（1986）分析「支」有五個用法：

- 1 計量隊伍。
- 2 計量歌曲或樂曲。
- 3 計量電燈的光度，相當於「瓦」。
- 4 計量某些杆狀的東西。
- 5 紗線粗細的計量單位。

國語會（1995）的分析十分類似，有四個用法：

- 1 計算隊伍的單位。
- 2 計算歌曲、樂曲的單位。
- 3 計算棉紗細度的單位。
- 4 計算燈光強度的單位。

再看看我們利用語料庫所觀察到的「支」的用法，共有十一種之多（編號前加上\*的符號表示是本辭典才有的用法分析）：

- 1 用來計量長杆狀的物品。
- \*2 用來計量電話。
- 3 用來計量隊伍或團體。
- \*4 用來計量整體中分支出來的人、物。
- \*5 用來計量錄影帶、影碟。
- 6 用來計量歌曲、舞蹈或影片。
- \*7 用來計量棒球賽中擊出的打擊數。
- \*8 用來計量賞罰記錄。
- \*9 用來計量數字、號碼。
- \*10 用來計量股票。
- 11 用來計量紡紗。

我們所分析出來「支」的用法，有七種現代用法是大陸「現代漢語量詞手冊」以及教育部「常用量詞手冊」中所沒有的。這七種用法所計量的物品都是現代社會中的新產物，而且成為現代生活中重要物品，像是：電話、錄影帶、影碟、棒球、大過、股票、…等。若不是語料庫中豐富多樣的語料我們也是沒有辦法掌握到現代漢語的最新變化。同時我們在語料庫中也觀察到現在已經不用「支」來計量電燈的光度，而用「燭光」來取代。因此，和另外兩本辭典不同，我們並不放這個用法。

c. 詞頻資料可以幫助決定詞項。

以量詞「另」為例，雖然我們可以在其他辭典中查到這個詞，知道它是用來計量商店，但是在語料庫中，它的出現次數是零。有了這項訊息，編輯人員就需要考慮是否要收「另」在量詞辭典中。在分析名量搭配辭典中的「商店、雜貨店、旅店」等名詞項時，也要考慮是否要放「另」這項搭配資料。

d. 頻率資料可以幫助挑選例詞、調整例詞順序。

在編輯中，我們總是從眾多例子中挑選出最具代表性的例子，而且排序上也以此為準則。什麼樣的例子算是具有代表性呢？出現頻率就是一項很重要的依據。以名詞尾「本」為例，它表示三種意思：根本、本子、本錢。表示「根本」之意有兩個例詞：「基本」（361次）、「根本」（53次）。很明顯的，「基本」的次數遠超過「根本」，依照這個訊息編輯人員應該將「基本」排在「根本」之前。另外表示「本子」的例子計有：劇本（69次）、版本（51次）、書本（47次）、課本（22次）、筆記本（11次）、……、中譯本（1次）、手抄本（1次）、標點本（一次）、印本（1次），共四十二個例子。編輯人員不需要將所有四十二個例子都放到辭典中，而是有所取捨。取捨的標準之一就是它們的出現次數，編輯人員會優先挑出常用詞，並且也會參考次數多寡排列例詞的順序。

e. 頻率資料和例詞數量可以幫助調整用法分析的順序。

以「張」為例，它可以用來計量臉、嘴、紙、桌子等物品。這些都是實體名詞，也都是日常生活中的常用名詞。究竟哪一個用法該放在前面？哪一個用法該放在後面？「張」的頻率資料和例詞數量可以作為參考（請參見附錄八）。根據這兩項準則，我們對「張」的分析順序如下：

1 用來計量平面的物品。

例詞：十張紙、四千多張罰單、一千張磁碟片、一張會員卡、九張照片、幾張鈔票、幾張郵票、一張稅單、兩張畫、幾張報紙、好多張地圖、八十張幻燈片、三張蔥油餅。

2 用來計量主要功能在平面上的物品，如桌子、床、椅子。

例詞：這張雙人床、那張床、十幾張圓桌、那張椅子、一張高腳小方桌、幾張書桌、那張長沙發。

3 用來計量臉面。

例詞：一張臉、那張驚慌的小臉、一張張猶顯稚氣的臉龐、一張張黧黑的面孔。

4 用來計量可以張開的物品。

例詞：兩張嘴巴、一張伶牙利嘴、幾張弓、兩張魚網。

f. 標記語料庫可以提供各種語法行為的統計資料。

就編輯辭典的層面來看，標記語料庫的運用層面比起未標記語料庫來得廣。因為標記語料庫提供了更多的訊息，有利編輯人員利用這些訊息作各種搜尋的工作。由於中研院語料庫標有詞類，因此我們才能抓取語料中的名詞。也才能將名詞依詞尾排序，得出名詞尾資料。也因為同樣的理由，我們才能搜尋名詞和量詞的搭配關係供編輯人員參考，否則就得完全靠編輯人員的功力了。

就漢語書寫的特性而言，詞和詞是串連在一起的，中間沒有間隔標記。標記語料庫中的詞是一個一個分開的，無論是搜尋、排序、統計…等工作，其準確度和速率都會大為提高。以「本」為例，在兩百萬語料庫中，「本」字一共出現了5784次，但是其中只有313次是量詞的用法。利用標記語料庫可以直接將這313個例子準確的挑出，省去從5784個例子中挑選和核對的麻煩。

g. 提供實用的例詞和例句

從語料庫擷取的例詞和例句和編輯人員自行造出的例詞例句是有很大的區別的，這一點 Sinclair (1987b: pxv) 有很精闢的說明：

……語料庫中的例子反映了實際使用的情形。它們佐證解釋、反映用法、並提供了現代英語在說和寫上可以信賴的指引。

相反的，造出的句子實際上只是解釋的一部份。它們被用來闡明 (refine) 解釋，在多半情況下只是被用來闡述 (clarify) 解釋。

語料庫除了提供上述多方面的語言事實外，它對語言分析也起了間接的作用。就是因為語料庫能夠提供多種語言面貌，因此幫助編輯人員做出更準確的分析。有些影響是反映在個別詞彙的分析上，有些則是反映在整體的語法架構上。就以這次編輯的經驗來看，在量詞分析的過程中，由於對量詞的用法一直有新發現，因此我們對量詞的分類也一再修訂改善。一開始是依照 Chao(1968)

及詞庫小組（1993b）的架構將量詞分成十類，<sup>11</sup> 最後修訂出來的分類為：一般量詞、事件量詞、種類量詞、容器量詞、約量量詞、標準量詞、動量詞，一共七類。我們認為新的分類更能掌握量詞間的差異，更能將性質不同的量詞區分清楚（請參考詞庫小組（1996）以及Huang（in prep.））。

#### 4. 討論

最後，我們也想提一下機讀資料庫對辭典編輯的助益。在這次辭典編輯的經驗中，無論是語料資料庫或是詞彙資料庫都具有以下的功能，增加了編輯流程的效率和便利。

- 一、可以隨時更新。
- 二、可以進行比對和比較的工作。
- 三、可以根據需要抽取所需資料。
- 四、可以變化各種輸出或排版格式，也可以根據需要將詞項作各種排序。

這一套編輯流程的設計粗具規模，還有許多待加強和改進之處。例如，詞彙資料庫和文字檔若能雙向傳輸，修改資料則更加便利。因為校稿工作多半是在印出的文字檔進行的，如果修改的工作也可以直接在文字檔中利用編校系統（editor）進行，再傳回詞彙資料庫就方便多了。但是由於目前詞彙資料庫和文字檔只能單向傳輸，所以只能進入編輯系統逐詞逐欄地修改了。其次，和COBUILD作法不同的是，我們的整項電子化流程並沒有包含排版系統，我們是將純文字檔的最後版本交給打字公司排版的。由於這兩本辭典的適用對象是小學生，必須加上注音和插圖，這些都增加了排版的困難。所以我們認為就這兩本辭典編輯而言，並不適合加入排版系統。再者，我們還可以建立交互索引功能。在分析一個詞彙會提到一些相關詞彙，像是同義詞、似義詞、反義詞、用法相近詞、…等，也應該在這些相關詞彙下標示。也就是電腦應該提供一個備忘庫，提醒編輯分析該詞彙時可以參考的詞項。這項功能在編輯大部頭的辭典時格外重要。此外，編輯介面系統也可以增加比對資料的功能。由於這兩本辭典是相呼應的，我們必須確定兩本辭典的內容不會互相矛盾，或是有所遺漏。比對工作交給電腦將會比較周全，效率也較好。但是因為兩部辭典是到同樣的語料庫中搜尋資料，差異不至於太大，因此決定交由人工進行比對。

註：

1. 隨後牛津及朗文這兩個重要辭典出版中心也立即跟進，構建了以辭典編纂為目標的語料庫。
2. COBUILD是Collins Birmingham University International Language Database 的縮寫。
3. COBUILD 根據這個基本資料庫已經陸陸續續出版了一系列的辭典、文法書及語言教學用書，像是：Collins COBUILD English Language Dictionary、Collins COBUILD Essential English Dictionary、Collins COBUILD Student's Dictionary、Collins COBUILD Dictionary of Phrasal Verbs、Collins COBUILD English Grammar、Collins COBUILD Student's Grammar、Collins COBUILD English Guides: 1 Prepositions and 2 Word Formation、Collins COBUILD English Course: levels 1, 2, and 3、Collins COBUILD English Course: Tests、Collins COBUILD English Course: First Lessons。
4. Sinclair(1987a pvii)舉了幾個例子：see、give、keep都有基本的語意，一般人都會認為基本的意義也是最常用的，但是事實並不是如此。see最常用在 "I see, you see"，give最常用在 "give a talk"，keep 最常用在 "keep warm"。
5. 由於名詞數量十分龐大，分別一一標記，不但耗工費時，讀者在查詢時也不方便。語意相近或是分類上屬於同一類的名詞所搭配的量詞往往都一樣，所以針對同一類的名詞來給量詞是可行的。依照Sproat and Shih(1996)的說法，名詞尾通常都是表示「種類」的概念，因此我們依照名詞尾來分類，往往也能把屬於同一類的名詞放在一起。依照名詞尾來分類名詞還有許多好處，不但檢索容易，相當於逆查辭典，而且還提供了名詞尾的語意分類。
6. 依照編輯原則，帶有同一個詞尾的名詞依下列原則再分成次類：一、名詞尾本身的語意不同；二、名詞所搭配的量詞不同。
7. 根據詞庫的詞類來看，名詞屬於Na和Ncb兩類。但是中研院語料庫只標上簡單的詞類，有Na，但是Ncb歸入Nc。Nc包含了Nca、Ncb兩類，必須一一核對詞庫電子辭典才能確定其細類。因此在中研院語料庫只找Na，另外利用詞庫另一套一千萬詞語料庫核對詞庫電子辭典，找出出現兩次以上的Ncb，共1277個，以及出現兩次以上卻不在中研院語料庫中出現過的Na，共890個。這就是為什麼補充了2167個名詞（1277個Ncb加上890個Na）的緣故。
8. 由於名量搭配資料是根據詞的間距來計算而非根據實際結構來判斷，因此誤差大，不比人工挑選出的例詞例句。所以在量詞的語料資料庫中不放電腦找出的名詞搭配資料，只放編輯人員挑選出的例詞例句。在名詞尾的語料資料庫中，雖然放了每個名詞所搭配的量詞，並且在編輯系統中提供統計每一組名詞所有的量詞和頻率的功能，但是仍然需要人工檢查、篩選和補充。
9. 根據挑選出的例詞排序的方法是：電腦先依照出現頻率將例詞排序並給予例詞010、020、... 120等編號。編輯人員可以更改編號，然後電腦再依照編號大小重新排序，並放入詞彙資料庫。
10. 請參見註8。
11. 根據詞庫小組（1993b）的分類，量詞有十類：個體量詞、跟述賓式合用的量詞、群體量詞、部份量詞、容器量詞、暫時量詞、標準量詞、準量詞、動量詞、零量詞。

## 參考書目

- 郭先珍, 1986, 現代漢語量詞手冊, 中國和平出版社。
- 教育部國語推行委員會, 1995, 常用量詞手冊, 國語文教育叢書17, 教育部。
- 陳克健, 1994, 素材語言學與本文處理, 發表於ICCL-3會議, 一九九四年七月, 香港。
- 黃居仁, 1995, 科際整合與整合科技—談計算語言學與語料庫語言學之角色與發展。「語言學研究之現狀與發展」研討會, 七月十五日, 國立台灣師範大學。
- 黃居仁、陳克建、張莉萍、許蕙麗, 1995, 中央研究院平衡語料庫簡介, 中華民國八十四年第八屆計算語言學研討會論文集。
- 詞庫小組, 1993a, 新聞常用名詞詞頻與分類—語料庫為本研究系列之四, 技術報告93-04, 中央研究院資訊科學研究所。
- 詞庫小組, 1993b, 中文詞類分析, 技術報告93-05, 中央研究院資訊科學研究所。
- 詞庫小組, 1995, 中央研究院平衡語料庫的內容與說明, 技術報告95-02, 中央研究院資訊科學研究所。
- 詞庫小組, 1996, 量詞的語意分類, 搜文解字, 計算語言學學會通訊第七卷第四期。
- Allen, Keith, 1977. "Classifiers". *Language*. 53:285-311.
- Chao, Yuen Ren, 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Church, K. W. and R. L. Mercer, 1993. "Introduction to the Special Issue on Computational Linguistics Using Large Corpora." *Computational Linguistics*, Vol.19, No.1, pp.1-24.
- Huang, Chu-Ren, 1989. *Citicization and Type-Lifting: A Unified Account of Mandarin NP de*. Bloomington: IULC.
- Huang, Chu-Ren, 1994. "Corpus-based Studies of Mandarin Chinese: Foundational Issues and Preliminary Results." In Mathew Chen and Ovid Tzeng Eds. *In Honor of William S.-Y. Wang: Interdisciplinary Studies on Language and Language Change*. pp165-186. Taipei: Pyramid.
- Huang, Chu-Ren, in preparation. "Classifiers and Semantic Type Coercion: Motivating a New Classification of Classifiers."
- Sinclair, John, 1987a. Ed. *Looking up: An Account of Cobuild Project*. London: William Collins Sons & Co Ltd.
- Sinclair, John, 1987b. Ed. *Collins Cobuild English Language Dictionary*. London: William Collins Sons & Co Ltd.
- Sproat and Shih, 1996. "A Corpus-based Analysis of Mandarin Nominal Root Compound." *Journal of East Asian Linguistics* 5, 46-71.
- Svartvik, Jan, 1992. Ed. "Directions in Corpus Linguistics." *Proceedings of Nobel Symposium 82, 4-8 August 1991. Trends in Linguistics Studies and Monographs 65*. Berlin: Mouton.
- Tai, James H-Y, 1994. "Chinese Classifier Systems and Human Categorization." In *Honor of William S.-Y. Wang: Interdisciplinary Studies on Language and Language Change*. (ed.) by Matthew Y. Chen and Ovid J.L. Tzeng.

## 母繫一 節節體例 ( 卽詞：「條」，名詞搭配：「法」 )

### 體例一：量詞辭典

#### ◎ 條 六 一 公 個 體 量 詞

- ① 後接條狀物品，像是領帶、項鍊、毛巾、裙子、褲子、傷痕、吐司、腿、船。例：他的手腕上有一條傷痕。
- ② 後接體形呈長條狀之動物，像是蛇、牛、魚、狗。例：前門屋簷上有條大蛇盤踞。
- ③ 後接條狀地面物，像是道路、公路、馬路、通路、財路、活路、生路、巷子、捷徑、河、排水溝、支流。例：水庫由三十條大小支流匯合而成。
- ④ 後接「線」的家族詞，像是路線、幹線、航線、陣線、防線、生產線、電線、電話線、地平線、視線、專線、線索。例：可以隱隱看見視線盡頭有一條墨綠的地平線。
- ⑤ 後接法律規則等詞，像是條文、條款、準則、規則、規定、指標、指示。例：根據清理法第廿三條規定，罰款四千五百元。
- ⑥ 後接「命」的家族詞，像是命、人命、生命。例：一場大火奪走了六十四條人命。

### 體例二：名詞量詞搭配辭典

- ㄈㄩˋ ① 方法、方式：辦法、方法、說法、**法**作法、做法、手法、玩法、想法、標示法、走法、…等。 **個量**：個、項。  
**群量**：套、種、類、樣、系列、串。  
**部量**：些、部份、堆。
- ② 法律條文：憲法、選罷法、民法、建築法、稅法、刑法、訴訟法、…等。 **個量**：條、項、道。 **群量**：種、系列。 **部量**：些、部份、堆。



附錄二 量詞依照頻率排列範例 (第一頁)

記錄#	nf	nf_voice	nf_count
1	個 (兒)	ㄍ ㄛˊ	15495
2	種	ㄓ ㄨㄥˋ	4666
3	位	ㄨㄟˋ	3225
4	年	ㄋ ㄢˊ	2397
5	次	ㄘ ㄛˋ	2386
6	項	ㄒ ㄩㄥˋ	2011
7	名	ㄇ ㄢˊ	1558
8	場	ㄇ ㄤˊ (俗讀 ㄇ ㄤˋ)	919
9	天	ㄊ ㄢˊ	879
10	些 (兒)	ㄊ ㄛˋ	872
11	條	ㄊ ㄤˊ	823
12	家	ㄐ ㄤˊ	822
13	隻	ㄓ ㄩˊ	821
14	元	ㄩ ㄢˊ	771
15	件	ㄐ ㄢˊ	718
16	座	ㄓ ㄨㄛˋ	608
17	歲	ㄘ ㄨㄟˋ	551
18	段 (兒)	ㄉ ㄤˊ	523
19	屆	ㄐ ㄩㄟˋ	515
20	屆	ㄐ ㄨㄟˋ	479
21	張	ㄓ ㄤˊ	458
22	塊	ㄎ ㄨㄟˋ	453
23	支	ㄓ ㄩˊ	447
24	句 (兒)	ㄐ ㄩˊ	444
25	片	ㄆ ㄢˊ	441
26	套 (兒)	ㄊ ㄤˊ	435
27	本 (兒)	ㄅ ㄢˊ	422
28	點 (兒)	ㄉ ㄢˊ	407
29	類	ㄌ ㄨㄟˋ	385
30	部	ㄅ ㄨˊ	315
31	群	ㄑ ㄩㄥˊ	306
32	類	ㄌ ㄨㄟˋ	305
33	份	ㄆ ㄢˊ	299
34	組	ㄓ ㄨˊ	289
35	首	ㄕ ㄨˊ	274
36	筆 (兒)	ㄅ ㄢˊ	273
37	篇	ㄆ ㄢˊ	250
38	把	ㄅ ㄤˊ	245
39	期	ㄑ ㄩˊ	242
40	號 (兒)	ㄏ ㄤˊ	242
41	日	ㄖ ㄣˊ	240
42	輛	ㄌ ㄤˊ	220
43	樣 (兒)	ㄩ ㄤˊ	213
44	根 (兒)	ㄍ ㄢˊ	212
45	度	ㄉ ㄨˊ	208
46	對	ㄉ ㄨㄟˋ	204
47	分 (兒)	ㄈ ㄢˊ	200
48	股	ㄍ ㄨˊ	194
49	道	ㄉ ㄤˊ	191
50	批	ㄆ ㄢˊ	188
51	週	ㄓ ㄨˊ	179
52	週	ㄓ ㄨˊ	175
53	口 (兒)	ㄎ ㄨˊ	169
54	間	ㄐ ㄤˊ	163
55	隊	ㄉ ㄨㄟˋ	159
55	層	ㄘ ㄨㄟˋ (又讀 ㄘ ㄨㄟˋ)	153
55	波	ㄅ ㄤˊ	153

附錄三 名詞尾依照ASCII碼排列範例 (第一頁)

NO	詞尾	筆數	包含此詞尾之NA			
1	一	34	國一 (18)	高一 (10)	大一 (3) ☆	研一 (1)
2	丁	10	唯一 (1)	等一 (1)		
			園丁 (3)	庖丁 (2)	兵丁 (1)	壯丁 (1)
			男丁 (1)	補丁 (1)	辣子雞丁 (1)	人丁 (0) ☆
			布丁 (0) ☆	柳丁 (0) ☆		
3	七	2	老七 (1)	絕一三七 (1)		
4	了	14	知了 (11)	了 (3)		
5	二	26	老二 (10)	國二 (9)	高二 (4)	二 (2)
6	人	16200	大人 (8921) ★ ◎	女人 (300) ★	男人 (255) ★ ◎	投資人 (247)
			中國人 (245) ◎	家人 (170) ★ ◎	國人 (166) ★ ◎	人人 (150) ★ ◎
			工人 (139) ★ ◎	老人 (135) ★ ◎	灣人 (130)	夫人 (129) ★ ◎
			主理人 (128) ★ ◎	負責人 (124) ★ ◎	台被害人 (121) ★	日本人 (114) ◎
			客人人 (114) ★ ◎	現代人 (112) ★ ◎	年輕人 (105) ★	日軍人 (92) ★ ◎
			病人 (90) ★	大人 (85) ★ ◎	候選人 (82)	商人人 (66) ★ ◎
			友人 (63) ★	藝人 (63) ★	美人 (62) ◎	發言人 (62) ★
			法人人 (57) ★	敵人 (57) ★ ◎	世人 (56) ★	發詩人 (56) ★ ◎
			華人人 (55) ★	外省人 (52)	古人 (51) ☆ ◎	漢人人 (51) ☆
			巨人 (50) ★ ◎	主持人 (49) ★	古衆人 (47) ★ ◎	強人人 (46) ★
			好人 (44) ☆ ◎	老年人 (44)	巴西人 (42)	外一人 (42) ★
			新人 (41) ☆	窮人 (38)	召集人 (37)	家一人 (36) ◎
			婦人 (36) ★ ◎	當事人 (36) ★	領導人 (35)	外國人 (34)
			農人 (33) ◎	成人 (32) ★	製行人 (31)	德人 (29) ◎
			親人 (29) ★ ◎	白人 (28)	行頭工人 (28) ★	後人 (28) ☆ ◎
			陌生人 (28) ◎	旁人 (28) ☆	碼人 (28) ☆ ◎	壞人 (28) ☆ ◎
			財團法人 (27) ★	學人 (26) ☆	小年人 (25) ☆ ◎	發起人 (24) ★
			黑人 (24) ★	證人 (24) ★	中法年人 (23)	文起人 (23) ☆
			日人 (23)	駕駛人 (23)	法國人 (22)	盲理人 (22)
			情人 (22) ★	當地人 (22)	專人 (21) ★	經心人 (21)
			獵人 (21) ☆ ◎	佳人 (20) ☆	客家人 (20)	理人 (19)
			所人 (19) ★	提名人 (19)	山人 (18) ☆	有名 (18)
			泰國人 (18)	美人 (17)	英國人 (17)	偉人 (17) ☆ ◎
			族人 (17) ☆	創辦人 (17)	先人 (16) ☆	印度人 (16)
			前人 (16) ☆ ◎	浪人 (16)	路人 (16) ★	僧人 (16)
			承銷人 (15)	被保險人 (15) ☆	戀人 (15) ☆	申請人 (14)
			東方人 (14)	某人 (14) ☆	為生意人 (14) ★	經紀人 (14)
			女強人 (13)	本省人 (13)	意人 (13) ☆	企業人 (13)
			西方人 (12)	阿拉伯人 (12)	雪亞人 (12) ☆	愛爾蘭人 (12)
			聖誕老人 (12) ☆	文化人 (11)	洲人 (11) ☆	活人 (11) ☆
			聖誕人 (11) ☆ ◎	自己人 (10) ☆	青年 (10)	納稅人 (10)
			猶太人 (10) ☆	管理 (10) ◎	凡人 (9) ☆	代言人 (9)
			有錢人 (9)	村人 (9) ☆	所有權人 (9)	香港人 (9)
			倆人 (9)	修行人 (9)	所荷蘭人 (9) ◎	愛人 (9) ☆
			愛樂人 (9)	矮人 (9) ◎	葡萄牙人 (9)	漁人 (9) ◎
			代理人 (8)	尼泊爾人 (8)	各人 (8) ☆	局外人 (8) ☆
			理刑人 (8)	社會人 (8)	參選人 (8)	發行 (8) ☆
			受瑞人 (8)	澳洲人 (8)	士人 (7) ◎	中完人 (7) ◎
			台北人 (7)	犯人 (7) ☆	同居人 (7)	個完人 (7)
			宜蘭人 (7)	承攬人 (7)	俄羅斯人 (7)	個人 (7)
			原始人 (7)	常人 (7) ☆ ◎	植義人 (7) ☆	發明 (7) ◎
			當選人 (7) ☆	罪人 (7)	上海人 (6)	僕人 (7) ◎
			熟人 (7) ☆	文人 (6)	俗人 (6)	今人 (6) ◎
			老人 (6)	受害 (6)	讀書人 (6)	現耕人 (6)
			超少年 (6) ☆	歐洲人 (6) ☆	平書人 (6) ◎	內地人 (5) ☆
			少年 (5)	主講人 (5)	死南人 (5)	安人 (5)
			吉普賽人 (5)	地球人 (5)		印人 (5)
			波蘭人 (5)	牧羊人 (5)		南人 (5)

# 附錄四 量詞語料資料庫存放資料範例 (量詞「條」)

## 例詞

這條河流  
這一條線  
每一條雙曲線  
每一條大門徑  
一條好幾條魚脈  
一條白絹道  
一條鐵道路的  
一條愛命連線  
一條尿管  
一條整條豬線  
一條整條走廊  
一條一條金鍊子  
一條一條產業道  
一條一條獨木舟  
一條一條十條人命  
一條一條命線  
一條一條游泳褲  
一條一條長弧線  
一條一條法第八條  
一條一條金牛

幾條線索  
四條徑  
一條路  
一條小溪  
這一條山脈  
一條繩子  
一條金帶子  
一條手帕  
一條鐵路  
一條電話熱線  
一條馬路  
一條財路  
第一條市場路線  
一條一條鐵鍊  
一條一條長長的影子  
這一條案第  
二條航空線  
這一條主線  
一條一條規則  
兩條笨重的犀牛  
三條十條罪狀  
這條法第八十四條  
規

整一條街  
一條一條蛇船  
一條一條銀河  
一條一條巨龍  
一條一條公偉大的公路  
一條一條金魚  
一條一條隧道  
一條一條魚子徑  
一條一條巷捷狗  
一條一條歌墨  
一條一條綠屬的地平線  
這一條一條金條橋  
一條一條香蕉  
一條一條尾巴鍊  
一條一條法語五員  
一條一條律文第二十四條  
二公務

## 例句

這兩條線將時空分成兩類上，下部份  
春借書辦過每快的訂街，也穿過了一條小巷。  
要東往的的天白，趕來救了他，可能送掉一條命。  
來往的汽車，像流星，現了一條虹。  
大聽的狗嚇得玻璃缸，四條腿，發軟。  
客依憲那坡，第條線，規定：總統依法行使大赦。  
她知沙道沙那，那條線，只是幻影。  
整條彩虹真美麗，彎著全校最細的粉。  
七我相信總拉一已天，潔白，水溝，是一條香蕉。  
我相信用麵登山，一條水溝，淨的會出。  
凡經國人對國，關卡，如髮，每走，人贈送紀念品毛巾一條與摸彩券一張，送完為止。  
泰國人對國，關卡，如髮，每走，人贈送紀念品毛巾一條與摸彩券一張，送完為止。

附錄五 名詞尾語料資料庫存放資料範例 (「法」第一頁)

方法★◎	505	種(38) 個(33) 套(6) 項(6) 年(1) 次(1) 兩(1) 招(1) 株(1) 組
辦法★◎	337	個(15) 套(11) 項(10) 種(9) 戶(1) 年(1) 次(1) 折(1) 屆(1)
法★	272	本(1) 份(1) 次(1) 套(1) 條(1) 種(1) 類(1)
看法★	191	個(7) 種(7) 位(5) 名(1) 派(1) 家(1) 種種(1) 樣(1) 點(1)
說法★◎	141	種(21) 名(4) 項(3) 個(2) 元(1) 件(1) 些(1)
作法★	129	種(12) 個(4) 位(1) 步(1) 套(1) 週(1) 項(1) 種種(1) 點(1)
手法★	109	種(7) 起(2) 個(1) 家(1) 項(1) 點(1)
想法★◎	104	個(10) 種(9) 次(1) 類(1)
司法★	92	種(3) 個(2) 件(1) 門(1) 路(1)
書法★◎	89	手(2) 件(2) 幅(2) 位(1) 家(1) 張(1) 張張(1) 歲(1)
做法★	84	種(11) 個(7) 項(2) 些(1)
語法★	82	個(3) 類(3) 句(1) 年(1) 系列(1) 套(1) 條(1) 層(1)
憲法★	59	個(1) 群(1)
輸入法★	34	種(1)
勞基法★	29	名(1) 次(1) 個(1)
演奏法★	28	種(9) 個(1) 項(1)
演算法★	24	種(2) 套(1)
鬆弛法★	20	個(2)
技法★	19	種(3)
指法★	19	種(2) 件(1)
療法★	19	種(2)
刑法★	18	
文法★	15	
民法★	15	
身法★	15	套(4) 個(1) 盤(1)
出版法★	13	場(1)
用法★	13	種(1)
寫法★	12	種(3)
公法★	11	
交易法★	11	
選法★	11	
工會法★	10	
玩法★	10	種(4)
魔法★	10	本(1) 個(1) 種(1)
打法★	9	把(1) 套(1) 場(1) 種(1)
佛法★	9	個(1)
治療法★	9	種(1)
畫法★	9	筆(1)
國安法★	8	
著作權法★	8	
解決辦法★	8	項(1)
比對法★	7	種(1)
保育法★	7	
建築法★	7	
講法★	7	個(2) 種(1)
人團法★	6	
教學法★	6	種(1)
賭法★	6	種(3)
反列法★	5	項(3)
吃法★	5	個(1)
戒嚴法★	5	
修持法★	5	種(3)
消防法★	5	
國際法★	5	間(1)
教師法★	5	
組織法★	5	
解法★	5	個(1) 家(1)
劃分法★	5	
廣電法★	5	

## 附錄六 傳送資料流程 (名詞「法」)

### a. 進入「法」的語料資料庫

### b. 編輯人員挑選例詞

*方法★◎	505	種(38) 個(33) 套(6) 項(6) 年(1) 次(1) 兩(1) 招(1) 株(1) 組
*辦法★◎	337	個(15) 套(11) 項(10) 種(9) 戶(1) 年(1) 次(1) 折(1) 屆(1)
法★	272	本(1) 份(1) 次(1) 套(1) 條(1) 種(1) 類(1)
看法★	191	個(7) 種(7) 位(5) 名(1) 派(1) 家(1) 種種(1) 樣(1) 點(1)
說法★◎	141	種(21) 名(4) 項(3) 個(2) 元(1) 件(1) 些(1)
*作法★	129	種(12) 個(4) 位(1) 步(1) 套(1) 週(1) 項(1) 種種(1) 點(1)
*手法★	109	種(7) 起(2) 個(1) 家(1) 項(1) 點(1)
想法★◎	104	個(10) 種(9) 次(1) 類(1)
司法★	92	種(3) 個(2) 件(1) 門(1) 路(1)
書法★◎	89	手(2) 件(2) 幅(2) 位(1) 家(1) 張(1) 張張(1) 歲(1)
*做法	84	種(11) 個(7) 項(2) 些(1)
語法☆	82	個(3) 類(3) 句(1) 年(1) 系列(1) 套(1) 條(1) 層(1)
憲法★	59	個(1) 群(1)
輸入法	34	種(1)
勞基法★	29	名(1) 次(1) 個(1)
演奏法	28	種(9) 個(1) 項(1)
*演算法	24	種(2) 套(1)
鬆弛法	20	個(2)

[總共 246項, 第 1/14頁]

[已選 10 項]

PgUp上一頁 PgDw下一頁 ↑上 ↓下 Enter選擇 Ctrl-End選量詞 Esc結束(不選量詞)  
 MAIN F:\user\dic\NA\_C 記錄 9422/33055 檔案 Num Ins

### c. 編輯系統根據頻率自動排序

### d. 編輯人員修改順序

技法☆		
指法		
*療法☆	010	方法
刑法★	020	辦法
文法☆	030	作法
民法☆	040	手法
身法☆	031	做法
出版法	110	演算法
*用法☆	070	療法
*寫法☆	050	用法
公司法☆	060	寫法
交易法	100	玩法
選罷法★		
工會法		
*玩法☆		
魔法		
打法☆		
佛法☆		

PgUp/PgDw上下頁 Ctrl-End結束排序 Esc重新選擇

PgUp上一頁 PgDw(不選量詞)  
 MAIN F:\user\dic\NA\_C 記錄 9668/33055 檔案 Num Ins

e. 編輯系統自動統計量詞資料並依頻率排序

f. 編輯人員篩選量詞

法：方法、辦法、作法、做法、手法、用法、寫法、療法、玩法、演算法

種 (91)	個 (60)	項 (20) [a]	套 (19)	×年 (2)	×次 (2)
×點 (2) [ae]	×起 (2)	×兩 (1)	招 (1)	×株 (1)	組 (1)
×塊 (1) [ae]	×種種 (1)	×戶 (1)	×折 (1)	×屆 (1)	×條 (1)
×位 (1)	×步 (1)	×週 (1)	×家 (1)	些 (1)	

g. 傳送到詞彙資料庫

▲ 名詞量詞搭配詞典分析 ▲

F1:儲存/列印

詞項	法	編號	na0562_1	注音	ㄘ ㄩ ㄨ	分析	T	獨立詞	F
▲語意	1 方法、方式								
▲凡例	說明 1: 例子：方法、辦法、作法、做法、手法、用法、寫法、療法、玩法 ●一般：個、項、套、招、組      ●事件： ●種類：種                              ●容器： ●約量：些 ●辨析：								
▲特例	例子： ●一般： ●種類： ●約量： ●辨析： ●事件： ●容器：								

PgUp/PgDw上下 Ctrl-PgDw增刪插 Ctrl-PgUp拷貝/重排 Ctrl-Home選擇 Ctrl-End結束  
 MAIN F:\user\dic\NA\_C 記錄 1/33055 檔案 Num Ins

## 附錄七 文字檔範例（量詞：「條」，名量搭配：「法」）

條去一么、

①〔事物〕用來計量長條狀物品。例詞：四條船、一條葉脈、一條繩子、一條金帶子、一條手帕、一條鐵鍊、一條長長的影子、一條游泳褲。例句：天空的東邊出現了一條虹。

辨析：「根」和「條」都可以用來計量長條狀物品，但是有一些差別。「根」只能計量橫切面為圓的物品，因此形狀不規則的長條狀物品只能用「條」來計量，如「一條船、一條圍巾、一條馬路」。此外，「條」傾向計量軟的或可以彎曲的長條狀物品，因此硬的長條狀物品，如「樹枝、鋼管」，就只能用「根」來計量。請參見「根」。

②〔事物〕用來計量體形或尾巴呈長條狀的動物。例詞：一條小魚、一條巨龍、整條豬、一條狗、兩條笨重的犀牛、兩條蛇、一條毛毛蟲。例句：來往的汽車像流星，街燈連接成一條金龍。

辨析：不過有些動物有尾巴卻不用「條」來計量，例如我們不說「一條貓、一條老虎」。用「條」計量的動物通常都可以用「隻」來計量。請參見「隻」。

③〔事物〕用來計量條狀地面物。例詞：這條河流、整條街、這條公路、兩條水道、一條鐵路、一條隧道、一條馬路、一條巷子、一條產業道路。例句：春天穿過了每一條大街和每一條小巷。

④〔事物〕用來計量「線」，也包括抽象意涵的線。例詞：這條線、一長條弧線、一條雙曲線、四條路線、一條墨綠的地平線、這條線索、兩條航線、這條海線。例句：這兩條線將時空分成上下兩部份。

⑤〔事物〕用來計量法律、規則、訊息。例詞：這條規定、一條規則、每一條新聞、三十條罪狀、二十五條條文、憲法第十八條、草案第二條、兩條消息。例句：依憲法第四十條規定：總統依法行使大赦。

辨析：「條」和「項」、「款」都可以用來計量條文，不過就公文的陳述次序來看，「項」是「條」下面的分項，「款」是項下面的分項，如「地方自治綱要第五十一條第一項第三款」。請參見「項、款」。

⑥〔事物〕用來計量「生命」的相關詞。例詞：一條命、一條生命、四十多條人命。例句：要不是救生員及時趕到，他可能送掉一條命。

法ㄘㄩˇ

①方法或方式。

◎方法、辦法、作法、做法、手法、用法、寫法、療法、玩法、演算法、…。指方法或方式。〔一般〕：個、項、套、道、招、組。

〔種類〕：樣、式。

◎看法、說法、想法、講法、…。指意見。〔一般〕：個、項、點。  
〔種類〕：派、樣、式。

②法律或規律。

◎憲法、勞基法、刑法、民法、交易法、選罷法、國安法、著作權法、保育法、國際法、軍法、稅法、…。指各種法律。通常不搭配量詞。

辨析：「憲法」還可以說「一部憲法」。

◎語法、文法、句法、…。指語文的規律。〔一般〕：套、條、個。

◎佛法、魔法、…。通常不搭配量詞。

③技藝或法力。

◎技法、槍法、劍法、箭法、刀法、指法、…。〔一般〕：套。

◎書法。〔一般〕：幅、張、篇、件。

辨析：「書法」除了和上述量詞搭配之外，還有「他寫得一手好書法」這樣的說法。



附錄八 量詞「張」所搭配的名詞資料

20 張 290

- |          |          |         |          |
|----------|----------|---------|----------|
| 照片 (18)  | 專輯 (12)  | 臉 (12)  | 床 (8)    |
| 牌 (7)    | 紙條 (6)   | 賀年片 (6) | 名片 (4)   |
| 門 (4)    | 量 (4)    | 人 (3)   | 毛毯 (3)   |
| 股價 (3)   | 便條 (3)   | 海報 (3)  | 紙 (3)    |
| 票 (3)    | 單 (3)    | 報票 (3)  | 慰問卡 (3)  |
| 錢 (3)    | 臉孔 (3)   | 支票 (2)  | 出片量 (2)  |
| 卡片 (2)   | 白紙 (2)   | 光碟 (2)  | 地圖 (2)   |
| 成績單 (2)  | 車票 (2)   | 直徑 (2)  | 信用卡 (2)  |
| 紅色 (2)   | 風景 (2)   | 書桌 (2)  | 設計圖 (2)  |
| 報紙 (2)   | 椅子 (2)   | 畫 (2)   | 圓桌 (2)   |
| 圖片 (2)   | 歌 (2)    | 磁碟片 (2) | 影本 (2)   |
| 靠背 (2)   | 遺失啓事 (2) | 錫箔紙 (2) | 錦緞 (2)   |
| 雙人床 (2)  | 三角形 (1)  | 大花臉 (1) | 大頭照 (1)  |
| 小鈔 (1)   | 五彩 (1)   | 內閣 (1)  | 切結書 (1)  |
| 天量 (1)   | 幻燈片 (1)  | 手 (1)   | 文憑 (1)   |
| 方桌 (1)   | 日期 (1)   | 水準 (1)  | 牛皮 (1)   |
| 主義 (1)   | 古蹟 (1)   | 司機 (1)  | 外卡 (1)   |
| 外匯券 (1)  | 布告 (1)   | 正片 (1)  | 母親 (1)   |
| 甘苦談 (1)  | 安非他命 (1) | 收藏版 (1) | 行事曆 (1)  |
| 床柱 (1)   | 沙發 (1)   | 身分證 (1) | 事 (1)    |
| 卦 (1)    | 姊妹 (1)   | 底片 (1)  | 所得 (1)   |
| 拘捕令 (1)  | 明信片 (1)  | 治安 (1)  | 玩具 (1)   |
| 者 (1)    | 花容 (1)   | 表 (1)   | 表姊們 (1)  |
| 長笛 (1)   | 門神 (1)   | 勉勵卡 (1) | 娃娃臉 (1)  |
| 宣傳 (1)   | 宦臣 (1)   | 拷貝費 (1) | 美金 (1)   |
| 耶穌頭 (1)  | 面額 (1)   | 音樂 (1)  | 倍 (1)    |
| 原圖 (1)   | 唇 (1)    | 校友 (1)  | 桌子 (1)   |
| 海 (1)    | 海灣 (1)   | 留言 (1)  | 病床 (1)   |
| 病歷表 (1)  | 笑臉 (1)   | 紙屑 (1)  | 紙頭 (1)   |
| 迷宮 (1)   | 陣線 (1)   | 剪報 (1)  | 唱片 (1)   |
| 國樂 (1)   | 措施 (1)   | 啓事 (1)  | 場次 (1)   |
| 報名費 (1)  | 畫卷 (1)   | 畫像 (1)  | 發票 (1)   |
| 稅單 (1)   | 視線 (1)   | 賀年卡 (1) | 發郵政幣 (1) |
| 陽光 (1)   | 集 (1)    | 塑膠紙 (1) | 新罩蓋 (1)  |
| 新聞 (1)   | 會員證 (1)  | 經典 (1)  | 雷公臉 (1)  |
| 新義賣品 (1) | 腳 (1)    | 經號碼 (1) | 漫畫稿 (1)  |
| 雷射 (1)   | 圖 (1)    | 演唱會 (1) | 嘴 (1)    |
| 監測網 (1)  | 網 (1)    | 網子 (1)  | 墨色 (1)   |
| 影子 (1)   | 獎金 (1)   | 稿紙 (1)  | 嬰兒 (1)   |
| 歷史 (1)   | 螢幕 (1)   | 頭髮 (1)  | 藥膏 (1)   |
| 牆 (1)    | 聯盟 (1)   | 舊事 (1)  | 藥妹 (1)   |
| 證明 (1)   | 攝影 (1)   | 仔標 (1)  |          |