

# 巴利數位文獻資源的現況與未來

蔡奇林

南華大學宗教學研究所兼任助理教授

## 目次

### 0. 序言

### 1. 巴利數位資源的現況

#### 1.1. 巴利原典光碟版

PTS 版、泰國版(BUDSIR IV)、緬甸版 (CSCD)

#### 1.2. 巴利原典網路版

錫蘭版、緬甸版 (VRI 版)、泰國版 (BUDSIR IV)

#### 1.3. 巴利原典翻譯

1.3.1. 巴利原典英譯

1.3.2. 其他各國語譯

#### 1.4. 網路上的巴英辭典

#### 1.5. 網路上的巴利文法

#### 1.6. 網路上的巴利讀本

#### 1.7. 網路上的巴利有聲資源

### 2. 略評當前的巴利數位資源

#### - - 其發展趨勢、效益、與限制

2.1. 當前巴利數位資源發展的幾項趨勢特點

2.2. 當前巴利數位資源所提供的學術研究效益

2.3. 當前巴利數位資源的不足之處

### 3. 巴利數位資源的未來展望

#### - - 略談「初期佛教聖典多語多本平行語料庫」

- 3.1. 從「低精審」邁向「高精審」
- 3.2. 一個擬構案例：「初期佛教聖典多語多本平行語料庫」
  - 3.2.1. 略說「初期佛教聖典・多語多本・平行・語料庫」
  - 3.2.2. 必須進行的相關基礎工程
  - 3.2.3. 能夠開發的教學研究效能

#### 4.結語：建構一個多效能的佛典知識寶庫

## 0. 序言<sup>1</sup>

早在 30 幾年前的 70 年代，英國巴利聖典學會(PTS)就著手從事巴利文獻的數位化工作。但當時由於人工輸入成本太高，而光學字體辨識系統(OCR)又不夠精確(只達 95%)，還需大量人工的校訂工作，最後只好放棄計劃。到了 20 年前(1984 年)，泰國法身寺也開始其巴利三藏的數位化工作，可是當時工作環境極為簡陋(只有一部電腦)，進展十分緩慢。但是這十幾年來，隨著資訊科技的猛迅發展，巴利數位資源的建設，如雨後春筍一般，迅速勃發，並且已經取得了相當顯著的進展，成績斐然。本文將在第一節簡要地介紹這些業績。<sup>2</sup>

其次，對於這樣快速而又大量的成績，我們有必要進行一番綜合的省察，以便了解其發展的趨勢特點，評估其能夠提供的學術研究效能，以及反省其存在的問題與限制，以便為其未來發展提供一些參考座標。這就是本文第二節的重點工作。

基於以上省察，本文將在第三節對巴利數位資源的未來，提供一些可能的思考，焦點將放在「提高精審度」以及「開發更多價的研究教學效能」之上。其中提出一個擬構的案例「初期佛教聖典多語多本平行語料庫」，以便作為具體體現此種思惟的參考。

## 1. 巴利數位資源的現況

### 1.1. 巴利原典光碟版

目前巴利藏光碟共有 PTS 版(歐洲版)、泰國版(BUDSIR)、以及緬甸版(CSCD)等三個版本，以下分別介紹。

#### 1.1.1. PTS 版巴利光碟(PALITEXT Version 1.0)

##### (1) 製作緣起：

<sup>1</sup> 本文是趕交本次研討會的未完稿(初稿)，仍有許多不足，待後續修改及補充。

<sup>2</sup> 有關巴利數位資源的現況，筆者曾有兩篇文章談及，本節將這些內容重新作比較簡單扼要的陳述，詳細可參原文：〈巴利學研究紀要：1995-2001〉，《正觀雜誌》，第 20 期，2002 年 3 月，頁 227-83。及〈網海一滴：網路上的巴利教學與研究資源舉隅〉，《佛教圖書館館訊》，第 40 期，2004 年 12 月，頁 21-35。前一文第 6 節是巴利光碟的介紹，後文則是網路資源介紹。

- (A) 英國巴利聖典學會(PTS)早在 1970 年代，就開始嘗試將巴利原典電子化，可惜限於技術、經費、人力，最後放棄計劃。
  - (B) 泰國法身寺法身基金會(Dhammakaya Foundation) 於 1989 年 9 月開始此版電子化工作。並於 1996 年 4 月與 PTS 簽訂合約，正式發行此版光碟。
- (2) 光碟內容：
- 收錄了 PTS 版巴利三藏共 53 冊——律藏 5 冊，經藏即五尼柯耶 37 冊，論藏七部論 11 冊；但尚未收入注釋、解疏、與藏外典籍。
- (3) 功能與特色：
- (A) 在版面上，光碟採「原貌重現」的形式，原原本本呈現原來書面版的內容與編排形式，這大大有利於資料的參照引用。
  - (B) 光碟版也做了一些提昇，例如，訂正原版打字上的錯誤(以紅色字表示)、標示原文行號等。
  - (C) 多重搜尋功能：光碟提供關鍵字詞(含詞組、句子)、冊號頁號、文脈段落、以及詞彙序列等四種搜尋模式，並具備萬用字元(\*)的強大搜尋功能，很容易搜尋字詞的前接(prefix)或後接(suffix)情況。搜尋結果列有清單，顯示該字詞或段落在各冊典籍的出現狀況及統計。
  - (D) 切割視窗，方便原文比對：光碟提供雙視窗、多重查詢、多段比對的功能。
- (4) 限制：
- (A) 只有 DOS 版，沒有 Windows 版。
  - (B) 未收錄注釋、解疏、及藏外典籍。
  - (C) 未提供經文擷取。
  - (D) 只提供點陣印表機(dot-matrix printer)的單頁列印模式。

### 1.1.2. 泰國版巴利光碟(BUDSIR IV)

- (1) 製作緣起：
- 泰國 Mahidol 大學於 1987 年奉命擬定一個特別的研究計劃，以慶祝泰王 Bhumibol(拉瑪九世)隔年的登基紀念大典。Mahidol 大學接受 Dr. Praves Wasee 教授建議，著手將泰國版巴利三藏置入電腦，以呈給泰王一個別具時代意義的獻禮。1988 年這項計劃有了初步成果，完成了泰國字版巴利三藏 45 冊原典的電子版 BUDSIR I。其後計劃持續進行，並不斷擴充改良，到 1994 年終於有了一個更加完善的版本 BUDSIR IV on CD-ROM (MS-DOS 版)。此版加入了注釋書，並有羅馬字版。
- (2) 光碟內容：
- 收錄了泰國版巴利原典共 115 冊，包括三藏 45 冊、注釋書 55 冊、以及其他典籍如 Milindapabha(《彌蘭王問經》)、Visuddhimagga(《清淨道論》)、Abhidhammatthasavgaha(《攝阿毗達摩義論》)等 15 冊。

## (3) 功能及特色：

- (A) 提供「關鍵字詞」(包括詞組、句子)與「冊號頁號段號」兩種搜尋模式，並具備萬用字元(\*)的搜尋功能。搜尋結果列有清單，顯示該字詞在各冊典籍的出現狀況及統計。
- (B) 提供分割視窗、經注合查的功能，可以經與經比對、注與注比對、或經與注比對。
- (C) 提供資料擷取的功能，但擷取的經文須另行轉碼才能使用。
- (D) 提供點陣印表機與噴墨印表機的列印功能。

## (4) 限制：

- (A) 儘管 1996 年 Mahidol 大學又開發了新版本：BUDSIR IV for Windows。但只能在英文版 Windows 下執行，不能搭配中文 Windows。
- (B) 未提供與 PTS 版的頁碼對照，大大減低了版本比對的方便，也降低了國際性。
- (C) 經文的擷取與轉碼太過複雜，不易使用。

## 1.1.3. 緬甸版巴利光碟(CSCD)

## (1) 製作緣起：

鑑於印度政府委由新那爛陀大寺(Nava Nalanda Mahavihara)刊行的巴利三藏已絕版，因此，S.N. Goenka 自 1985 年創立內觀研究所(Vipassana Research Institute, VRI)之後，便積極進行天城體版(Devanagari)巴利三藏、注釋、解疏的刊行工作，以讓佛陀的教法重新流行在他的母土。這個版本(VRI 版)採用緬甸第六結集版為底本，將緬甸字體轉寫成天城體，而後輸入電腦處理，並作了一些編輯上的更動。之後，為了提昇研究效能以及長久保存三藏，又製成光碟，於 1997 年發行第一版：Chattha Savgayana CD-ROM(CSCD)，隔年又出第二版，1999 年出了第三版。

## (2) 光碟內容：

- (A) CSCD (Version 3)收錄三藏、注釋、解疏、補疏(anutika)以及藏外典籍如阿毗達摩綱要書、史書、文法書、詩律學等各種巴利原典共 216 冊。
- (B) 並收入一部巴利文法書(Lily de Silva, *Pali Primer*，附習題解答)、一部簡明巴英辭典，還有阿育王碑文(Ashok ke Abhilekh)。

## (3) 功能及特色：

- (A) 具有七種字體的顯示功能：羅馬字(Roman)、緬甸字(Myanmar)、天城體(Devanagari)、泰國字(Thai)、錫蘭字(Sinhalese)、高棉字(Khmer)、蒙古字(Mangol)等。
- (B) 強大的搜尋功能：提供字詞(或詞組)、文脈段落等搜尋方式(有萬用字元\*功能)，並具有「標示查詢」功能，也可直接翻查某經某品某頁。搜尋結

果列有出處清單及出處文句摘要，可以存檔或列印。

- (C) 可同時開啓多部典籍(最多六部)，並具備三藏、注釋、解疏的段落關連功能，很方便進行經文之間或經、注、疏之間的比較或對讀。
- (D) 提供與其他版本(如 PTS 版、泰國版、錫蘭版)的對應頁碼，便於不同版本間的參照比對。
- (E) 更重要的，提供經文擷取與列印功能，並且是配合中文版 Windows 環境，很方便擷取之後配合中文進行後續的編輯利用。

## 1.2. 巴利原典網路版

目前網路上比較完整的巴利藏原典有錫蘭版、緬甸版、以及泰國版三個版本，下面分別介紹。

### 1.2.1. 網路上的錫蘭版 ( SLTP - BJT 版 )

#### (1) 製作緣起

錫蘭版巴利三藏 (BJT 版, Buddha Jayanti Tipitaka Series) 是 1960-1970 年代由錫蘭政府贊助出版的。這是錫蘭政府紀念佛滅 2500 年 (A.D.1956 年) 系列活動中重要的一項。錫蘭版三藏從 1991 年開始輸入電腦，到 1994 年全部完成。之後，在 1996 年 3、4 月正式上網。接著，進行後續的校讀工作。

#### (2) 收錄內容：

此版內容包括「三藏」：律藏、經藏 (五尼柯耶)、七部論；還有「藏外典籍」(38 冊)，諸如——1. 古代巴利文法：如迦旃延文法 (Kaccayana)、目犍連文法 (Moggallana) 等。2. 史書：如島史 (Dipa-vamsa)、大史 (Maha-vamsa) 等。3. 後期經論：如彌蘭王問經 (Milinda-pabha)、清淨道論 (Visuddhi-magga) 等。4. 古代巴利辭書 (如 Abhidhana-ppadipika)，以及詩學與修辭學方面的巴利典籍 (如 Subodhalavkara)。

#### (3) 功能及特色：

這個版本的優點之一是，所有律藏和經藏都已附有 PTS 版 (Pali Text Society, 巴利聖典學會版) 頁碼，對照查索方便。

#### (4) 限制：

未提供線上瀏覽及檢索功能，必須下載之後，另以檢索器 (如 Windows Grep) 檢索。

#### (5) 校讀資訊：

根據 L. S. Cousins 的附記，從 1996 年 10 月起，到 1999 年 1 月，共經過 4

次更新，這些更新版反應最新校讀除錯之後的成果，目前律藏及經藏都已經做過初步校讀（一次），但仍不保證完全無誤。

【錫蘭版巴利藏(SLPT-BJT)】網址：<http://jbe.gold.ac.uk/palicanon.html>。

### 1.2.2. 網路上的緬甸版 ( VRI 版 )

- (1) 製作緣起：此版即 2.3.節 CSCD 的網路版。
- (2) 收錄內容：略同 2.3.節 CSCD 光碟版。
- (3) 功能及特色：
  - (A) 提供線上瀏覽功能。
  - (B) 提供與 PTS 版的頁碼對照（可直接點選瀏覽）。
- (4) 限制：未提供線上檢索功能。

【緬甸版巴利藏(VRI)】網址：<http://www.tipitaka.org/tipitaka/booklistframe2.html>。

### 1.2.3. 網路上的泰國版 ( BUDSIR on Internet )

- (1) 製作緣起：此版即 2.2.節 BUDSIR 的網路版。
- (2) 收錄內容：三藏 45 冊，三藏泰譯 45 冊，注釋書 70 冊。
- (3) 功能及特色：此版是此處介紹的三個網路版巴利藏中，功能最齊全的。
  - (A) 具有羅馬化、天城體、以及泰文、錫蘭文等四種字體的顯示功能。
  - (B) 具有線上瀏覽及方便的全文檢索功能。
  - (C) 提供與 PTS 版頁碼對照功能。

【泰國版巴利藏(BUDSIR)】網址：<http://budsir.mahidol.ac.th/>。

## 1.3. 巴利原典翻譯

### 1.3.1. 巴利原典英譯

- (1) 網路上的英譯

目前只有律藏及經藏的部分翻譯，論藏還未收入。請參以下三網站：

- MettaNet (慈網)：可由 <http://www.metta.lk/> 進入 MettaNet 首頁，再點選“Tipitaka”。
- Access to Insight (內觀之道)：進入 <http://www.accesstoinight.org/>，再點選“Tipitaka”。
- Mahindarama (瑪辛達寺)：由 <http://www.mahindarama.com/> 進入首頁，在此頁最下方的“e-Service”選項欄中，點選“e-Tipitaka”即可進入三藏英譯網頁。

## (2) 英譯光碟版

美國 Access to Insight (內觀之道) 網站站長 Bullitt 從 1993 年起，逐步收羅巴利三藏的英譯 (最主要的譯者是 Thanissaro 比丘) 以及有關內觀禪修的相關著作，到 1998、1999 年，累計已得經文超過 800 篇，文章及專書也達數百部 (篇)。當時曾製成光碟 (名為 *A Handful of Leaves*，一掌樹葉) 免費流通，但目前已經絕版，預計 2006 年將再發行新版。

### 1.3.2. 其他各國語譯

- (1) 錫蘭語語譯：網路版可由 <http://www.metta.lk/> 進入 MettaNet 首頁，再點選，“Tipitaka”其中收有巴利原典、以及原典的英譯、錫蘭語譯。此外，MettaNet 也提供這些資料的 CD 備索，但僅限斯里蘭卡地區讀者。
- (2) 泰語語譯：網路版可由 <http://budsir.mahidol.ac.th/> 進入泰國 Mahidol 大學網站，註冊或登錄後，在巴利三藏網頁的“Types”欄中，點選 Thai Tipitaka。此外，Mahidol 大學 1997 年的新版巴利光碟 BUDSIR IV / TT (for Thai Translation) 也收有 45 冊巴利三藏的泰譯。
- (3) 德語語譯：可由 <http://www.palikanon.com/> 進入。除了經藏、律藏翻譯之外，還有部分論藏及藏外典籍翻譯。
- (4) 法語語譯：可由 <http://www.canonpali.org/tipitaka.html> 進入。目前只收錄部分經藏的翻譯，律藏、論藏翻譯尚缺。

## 1.4. 網路上的巴英辭典



- (1) PTS 版巴英辭典 (PED) : <http://dsal.uchicago.edu/dictionaries/pali/>。
- (2) A. P. Buddhadatta 長老的 *Concise Pali-English Dictionary* (簡明巴英辭典) :  
<http://www.metta.lk/pali-utils/index.html>。
- (3) G. P. Malalasekera 的 *Dictionary of the Pali Proper Names* (DPPN 巴利佛教  
專有名詞辭典) : <http://www.metta.lk/pali-utils/Pali-Proper-Names/index.html>。
- (4) Nynatiloka 長老的 *Buddhist Dictionary*(佛教教理辭典) :  
[http://www.buddhanet.net/pdf\\_file/palidict.pdf](http://www.buddhanet.net/pdf_file/palidict.pdf)。
- (5) Access to Insight 的 *A Glossary of Pali and Buddhist Terms*(巴利佛教名相  
彙編) : <http://www.accesstoinsight.org/glossary.html>。

## 1.5. 網路上的巴利文法

(1) Lily de Silva 的 *The Pali Primer* (巴利語入門)：

<http://www.vri.dhamma.org/publications/pali/primer/index.html>。

(2) Narada 長老的 *An Elementary Pali Course*(基礎巴利語教程)：

<http://www.metta.lk/pali-utils/ele-pali.doc>。

(3) Charles Duroiselle 的 *A Practical Grammar of the Pali Language* (巴利語實用文法)：MettaNet 網站：<http://www.metta.lk/pali-utils/index.html>。

(4) 蔡奇林編《實用巴利語文法》：由 <http://ccbs.ntu.edu.tw/BDLM/index.htm> 進入，點選「語言教學」的「巴利文教學」，點選「4. 實用巴利語文法」。

## 1.6. 網路上的巴利讀本

「佛學數位圖書館暨博物館」網站收有整部《法句經》(Dhammapada)26 品 423 頌的全部內容。其中各個頌文，列有巴利原文、翻譯、句構圖示、單詞詞義、語法分析等。網址：由 <http://ccbs.ntu.edu.tw/BDLM/index.htm> 進入後，點選「語言教學」的「巴利文教學」，點選「3. 經文選讀」。

## 1.7. 網路上的巴利有聲資源

(1) 字母、單詞、句子讀誦：由 <http://ccbs.ntu.edu.tw/BDLM/index.htm> 進入後，點選「語言教學」的「巴利文教學」，選「1.字母與發音」或「3. 經文選讀」。

(2) 巴利佛教課誦：[http://www.concentration.org/b\\_chants.htm](http://www.concentration.org/b_chants.htm) (泰國法身寺)。

(3) 現代巴利經文唱誦：<http://www.buddhanet.net/audio-chant.htm>。

## 2. 略評當前巴利數位資源 - - 其發展趨勢、效益、與限制

上面介紹了當前巴利數位資源的概況。以下將以幾個建構得比較完整、具規模的網站或光碟為主，觀察、評析目前這些資源建構發展的趨勢特點、提供的學術研究效益、以及還待改善或開拓的空間。樣本選擇基本上以收有巴利三藏原典的網站或光碟為主。這樣的網站有斯里蘭卡 MettaNet(慈網)、泰國 Mahidol 大學巴利三藏網站，而光碟則有 PTS 版、BUDSIR、及 CSCD 等三個版本。其他的網站資源則作為考察評述的輔助樣例。

## 2.1. 當前巴利數位資源發展的幾項趨勢特點

### (1) 收錄的原典「數量」不斷擴增

PTS 版光碟目前(1996 年版)只收錄三藏 53 冊，未來第二版 MS-Windows 版則計劃收入主要的注釋書；接著第三版麥金塔(Macintosh)相容版，將納入所有注釋書及解疏等。

泰國版光碟 BUDSIR I~III (1988~1990 年)只收錄巴利三藏 45 冊，到了 BUDSIR IV(1994 年)則納入了注釋書 55 冊，以及藏外典籍 15 冊，共達 115 冊巴利原典。

較晚發行的緬甸版(VRI 版)光碟 CSCD，其第一版(1997 年)收錄原典 146 冊，第二版(1998 年)收錄原典 183 冊，第三版(1999 年)收錄原典已達 216 冊。

總的趨勢是，以三藏的收錄為基礎核心，逐步往藏外典籍、以及注釋、解疏、補疏不斷擴充。

### (2) 原典校訂的「品質」不斷提昇

由於電子版巴利原典是以紙本為底本而輸入製作的，因此對於紙本先前所沒有校出的錯誤，就可以利用製作電子版時重新輸入及校訂的機會，將之訂正。因此一般而言，電子版有很好的機會，可以提昇原典校訂的品質。筆者因尚未對所有電子版原典做過比較全面的比對調查，以下只以 PTS 版光碟的部分觀察所得為例，做初步的評述。

在先前多篇文章中，筆者曾提到過，PTS 版原典的文字校對並不理想，以筆者的閱讀經驗而言，平均每一頁一般有 3、4 處錯誤。這一問題，PTS 先前的會長 K.R. Norman 也曾對學者提出警告，但限於經費及專家人力，紙本的重校再版有其一定的困難。

泰國法身寺在製作 PTS 版光碟時，校訂工作十分謹慎從事。首先由兩組人員分別輸入巴利原文(兩遍)，而後透過機器比對，先檢查出大部分的打字錯誤。而後進行後續多組多次的人工校訂，務求萬無一失。因此也校出並訂正了許多先前紙本的明顯錯誤。所以，PTS 版的光碟版精確度是高過原來紙本的。

儘管如此，巴利原文的校訂工作，有時並不是單靠機器及一般的人力而能畢竟其功的。筆者曾做過部分的抽樣比對，這片精度已經大為提高的光碟，其實也還留下不少錯誤尚未校正——數量恐怕高達 3 分之 1，甚至一半。但由於電子資料的更正再版更加容易，我們可以期待未來的新版本，會不斷訂正這些錯誤。

### (3) 朝向「雙語/多語資料庫」發展

這裡「雙語/多語資料庫」指的是，除了巴利原典之外，還包含原典的其他語言翻譯。由於國家的差別，各網站或光碟收錄的譯本語言也就不同。目前具有「雙語/多語」特點的資源有以下幾個。

1997 年最新版的泰國版巴利光碟 BUDSIR IV / TT (for Thai Translation)，已收錄了巴利三藏的泰譯本 45 冊。此版的網路版 BUDSIR IV on Internet 也收有三藏的泰譯。這是「雙語資料庫」的例子。

斯里蘭卡 MettaNet 網站除了收錄錫蘭版巴利原典之外，還收有原典的英譯及僧伽羅語(錫蘭語)語譯，是「三語(多語)資料庫」的例子。

此外，有一些網站雖只收錄英譯(如 Access to Insight)，或只收錄德譯、法譯等(如 1.3.2.節(3)(4)項介紹)，而未收入巴利原典，但由於巴利原典在光碟及網路上可即時取得，並且這些翻譯文獻的組織結構基本上符應於巴利原典，對照查找並不困難，因此這些資源或者也可以算作廣義的「雙語/多語資料庫」。

#### (4) 納入輔助原典解讀的「工具資源」

巴利原典的解讀，即便是對於專業的研究者，都不是輕鬆容易之事，因此，目前已有許多網站或光碟開始收錄一些可以幫助原典解讀的工具資源，例如文法書及辭典。

泰國版巴利光碟 BUDSIR IV / TT (1997 年版)收有一部「巴泰辭典」，未來還計劃收入「巴英辭典」。

緬甸版巴利光碟 CSCD Version 1.1 (1997 年版)收有一部「巴印辭典」，Version 3 (1999 年版)則附上一部簡明的巴英辭典，以及一部巴利語入門的文法書(Lily de Silva 的 *The Pali Primer*)。

斯里蘭卡 MettaNet 網站中的“Pali Utilities”(Dictionary and Grammar Books)，其中收有多部巴英辭典及巴利文法(參 1.4.及 1.5.節的介紹)。

PTS 版光碟目前(第一版)尚未收錄任何工具資源，但計劃在第三版附上一部交談式巴英辭典。

#### (5) 提供「多重對照」功能

這裡「多重對照」意思比較寬鬆，包括不同版本巴利原典之間的頁碼參照，同一版本巴利原典的經文與經文、經文與注釋、注釋與解疏之間的多視窗參照。

緬甸版 CSCD 巴利光碟、泰國 BUDSIR IV 網路版、MettaNet 的錫蘭版都提供了與 PTS 版的頁碼參照。

泰國版巴利光碟 BUDSIR IV 提供經、注合查功能，可以達到關鍵詞句的經與經、經與注、注與注比對的功能。而緬甸版 CSCD 光碟則進一步做到經、注、疏之間相應段落的連結的功能。

## (6) 提供「多字體顯示」功能

巴利語作為一個語言，歷來並沒有專屬的文字紀錄系統，佛教聖典初始的時候，也是口傳文獻。目前所知，巴利三藏最初是在紀元前一世紀在錫蘭紀錄下來的，當時是用錫蘭文字拼寫，而後傳到緬甸、泰國、高棉等南方國家，也同樣用這些國家本身的文字拼寫保存。因此，用多種字體「音寫」(音譯)保存巴利聖典成了南傳佛教的一個傳統特點。

現在由於電腦轉換字體容易，許多巴利網站或光碟都紛紛朝向多字體顯示功能發展，以便讓不同國家的讀者更容易閱讀巴利聖典。

目前 PTS 版光碟只提供羅馬化字體。而泰國 Mahidol 大學的 BUDSIR I (第一版，1988 年)是泰國字版，第二版(1989 年)開始就有了羅馬化字體，到了 1997 年的 BUDSIR IV / TT 版就有 6 種字體的顯示功能——可以顯示泰國字、羅馬字、緬甸字、錫蘭字、高棉字、以及天城體等 6 種文字的巴利原文。BUDSIR IV 的網路版，也提供 4 種字體的選擇(上面 6 種中除去高棉字、緬甸字)。

緬甸版光碟 CSCD 1.1 版(1997 年) 提供羅馬字、緬甸字、天城體等 3 種字體，2.0 版(1998 年)開始便提供 7 種字體的顯示功能(即上述 6 種加上蒙古字)。

## (7) 朝向「有聲資料庫」發展

前面提到，巴利佛典最初便是口誦傳承，而巴利語作為一個語言，口誦、耳聞更是最自然而親切的傳播方式。因此「有聲化」不管對這部聖典或這個語言的傳播、學習而言，都是重要的。

目前已有一片光碟具備「讀誦」功能，就是 1996 年泰國 Mahidol 大學的 BUDSIR IV for Windows，讀者可以標示經文段落，請求「讀誦」。可惜的是，這片光碟只能搭配英文版 Windows，無法在中文版下正常顯示。相信隨著電腦科技的發展，在更大容量與更高處理速度的支援下，會有更多巴利網站或資料庫朝向「有聲化」發展。

## (8) 適合「中文系統」的資料庫

由於目前巴利學數位資源幾乎都由西方國家、及南傳佛教國家所開發建構，因此一般鮮少考慮到中文系統的執行環境。例如前述 BUDSIR IV for Windows 就只能在英文版 Windows 下執行，而其他網路版即便安裝了字型之後，還是經常遇到字碼無法正常顯示的情形，這對於廣大的中文地區的使用者而言，無疑是一大缺憾。

所幸目前已經有一些建置單位開始注意到這個問題，而開發配合中文 Windows 的相容版。緬甸版 CSCD Version 3 便是一個好的例子，不只可以在中文 Windows 環境下正常顯示，並且資料下載之後，還可直接配合中文檔案，進行後續的編輯

利用，讓資源的效益發揮到最大。相信以中文地區佛學研究、以及佛學數位資源建置的蓬勃發展，當會促使越來越多巴利資源的建構單位，往中文相容的系統去發展。

## 2.2. 當前巴利數位資源所提供的學術研究效益

### (1) 更充分的研究樣本

更多的研究樣本及資訊，經常可以讓研究者得到更整全的觀察與更客觀的理解。目前光碟版及網路版，共有四套巴利原典——錫蘭版、緬甸版、泰國版、歐洲版(PTS 版)，且以緬甸版而言，當中就包含了許多西方尚未出版的原典。就筆者所知，目前國內即使是最專業的佛學研究機構，也沒有任何一個在其圖書館或資料中心當中，收齊這四個版本的紙本文獻。最好的恐怕只擁有其中的一半或三分之一，甚至更少。<sup>3</sup>但現在，任何個人研究者，都可以在其即使是窮鄉僻壤的斗室當中，敲幾個按鍵，就佔有比專業圖書館更加充分的原典資源。

除了原典之外，原典翻譯也是研究者重要的參考資料，網路上收錄了許多尚未出版或不易取得的譯本。例如 MettaNet 所收的 Bhikkhuni Uppalavana 的英譯、或 Access to Insight 所收的 Bhikkhu Thanisaro 的英譯。藉此可以了解不同譯者或不同傳統對巴利原典的不同理解。

### (2) 更精良的原文校訂

好的校訂本，能提供研究者更加可靠的資訊來源，這是一切研究的重要基石。巴利數位原典，一則由於各版本在數位化過程中，重新校正了不少過去錯誤，再則使用者可以透過各版所提供的「頁碼對照」，而對校不同版本的原典，這使得在底本文獻的使用上，一般可以取得較之過去更好的讀法。

### (3) 更適切的分段、斷句與標點

適切的分段、斷句與標點，能更清楚而逼真地傳達文本(事件)的結構脈絡、意義內涵、以及情境氣氛，對研究者解讀、運用基礎文獻將有莫大助益。

現代編輯出版的原典，如 PTS 版，雖說大部分都已進行了不錯的分段、斷句、及標點，但仍有一些缺失。例如，「相應部」的標點，只採用“||”及“|| ||”二種分隔或停頓符號，很難傳達更多豐富的義涵(例如引述、疑問、感嘆等等)。

<sup>3</sup> 並且國內所擁有的一些 PTS 以外的紙本巴利文獻(例如錫蘭版及緬甸版)，都尚未羅馬化，也較不方便閱讀。

其次，早期佛典多為佛陀與弟子(或國王、大臣、異教徒等)之間的問題討論，「問答」常是文本中重要的特徵，因此「問句」的表現就顯得特別重要。但 PTS 的「中部」第一冊雖採現代標點，但全冊卻沒有用到「問號」，使得疑問句的語氣及問答的情境，不能夠清晰、生動的呈現。

除了標點問題之外，PTS 版也有一些分段與斷句明顯錯誤的地方，<sup>4</sup>筆者曾在其他地方談過，錯誤的斷句，不僅會造成義理的誤解，同時也可能造成修行法門的歧出。<sup>5</sup>

目前緬甸版和錫蘭版的分段、斷句與標點，一般而言是比較好的，研究者可以通過光碟或網路上的「頁碼對照」，而參考到其中更好的分段與句讀，對文本取得更適切的理解。

#### (4) 更快速而趨近窮盡的資料檢索

以上第(1)點是研究資源的「數量」更加豐富，第(2)(3)二點則是研究資源的「品質」或可靠度更加提昇，這些都是提昇研究成果所需依賴的基礎。然而，數位資源所具備的最大效能或特點卻在於，通過檢索工具的幫助，讓研究者能夠用更少的時間、精力，進行盡可能「窮盡而沒有遺漏」的觀察。

在過去，面對浩瀚的紙本文獻，研究者或者只能選取有限的樣本，進行部分的觀察；或者需要憑藉記憶(或印象)所及，而進行難免主觀而不盡可靠的論述；或者必須經年累月地製作無數卡片，尚難保沒有任何遺漏。但現在對於某些問題(例如詞彙研究)，研究者可以在檢索工具的協助下，取得趨近窮盡的樣本。

#### (5) 提昇巴利研究的「精審度」

總的來說，上述四項效益都指向一個共同目標，就是提昇巴利研究的「精審度」。而這個精審度的等級，就是一般所謂的「低精審」(lower criticism)——也就是不管對於文本本身的讀法、或者對於文本所記載的內容的各種研究，都盡可能「窮盡地」運用巴利內部的所有文獻，以取得更加可靠或確當的結果。通過文本的數位化以及搜尋引擎的幫助，這樣的理想與目標變得更加可能。

### 2.3. 當前巴利數位資源的不足之處

儘管上一節談到目前的巴利數位資源不管是在「質」、「量」、或檢索擷取的「運

<sup>4</sup> 例如《因緣相應·20 經》的 3、4 段，及 4、5 段；《因緣相應·35 經》的 14、15 段，及 15、16 段，這些段落的分段、分句都有問題。

<sup>5</sup> 筆者 2004 年 6 月講於中華佛學研究所〈文本的異讀與法門的轉變——從安那般那念的二個疑難句談起〉(初稿)。

用效能」上，都已大為提高，但這並不意味這些資源已經臻於「完善」的境地。相反的，這些數位化工程，其實不過十數年歷史，儘管進展神速，但畢竟還在一個起步階段，因此還有許多缺失及問題仍待不斷地改善提昇。以下面舉其綦綦大者。

### (1) 原文校訂、標讀仍不夠完善

以 PTS 版而言，上面說過，光碟版雖已改正一些原文錯誤，但還有更多沒有訂正過來。錫蘭版的經藏、律藏目前只經過初步校訂(一次)，未來還須再作校訂。緬甸版一般而言三藏(及其注釋、解疏等)校對比較精良(但仍有少數錯誤)；但藏外典籍的部分則大多只經過機器校訂(約可除去 98 % 錯誤)，還未經過人工審校，可靠度比較不足。

在句讀方面，前面提過，PTS 版「相應部」及「中部」第一冊，有必要重新標讀；而其他各冊的標讀錯誤，也應再加以訂正。泰國版(BUDSIR)是各版當中，標讀最為簡略的，除了「省略號」(...)之外，全編竟只用「句號」(.)一種。因此，文中所見，盡是一大段落連串的數個句子，才有一個句點，這大大增添了讀者的閱讀負擔。相對而言，緬甸版(CSCD)的標讀就清晰、豐富許多，它靈活運用現代標點，如：逗號(,)、分號(; )、句號(.)、問號(?)、驚嘆號(!)、破折號(—)、單引號('')、雙引號(“”)、省略號(...)、連接號(-)等，讓原來口語的停頓、語氣能夠呈現，也讓詞語(或詞句)的意義及前後關係更為明白。

### (2) 許多原典及翻譯仍待收錄

目前 PTS 版光碟只收錄三藏原文 53 冊，其他已經出版的大量的三藏注釋、部分的解疏、以及藏外典籍都還沒有收入。儘管這些文獻絕大部分在其他版光碟中(如 CSCD)已經收錄，但由於不同校訂本所依據的底本經常不同，因此這些原典仍深具數位化的價值，以利學者從事比對、研究。

原典的翻譯目前雖已收錄不少，但距離最起碼的「完備」(律、經、論都至少一個譯本)還有一大段距離，其中尤以律藏、論藏最為缺乏。未來若版權問題能得到解決，重要譯本的收錄實是下一階段重要的工作。這是體現巴利佛典「雙語化/多語化」的基礎工作，也是結合「佛教研究」及「佛教教育」與「佛法傳播」的極佳方式。

### (3) 多數資源「版本資訊」不夠清楚

爲了讓讀者對於原典的來源有一基本了解，一般至少需要說明該原典的編輯者、編輯年(或出版年)、所依據的相關底本、取擇異讀的原則等等。但目前見到的巴利數位資源，這些資訊的提供實在十分不足。



目前不管是內觀研究所(VRI)的 CSCD、泰國 Mahidol 大學的 BUDSIR IV、還是斯里蘭卡的 SLTP，都只說明該版根據哪個版本進行數位化，<sup>6</sup>但並未進一步說明這些底本編輯過程所依據的本子，以及編輯校訂的原則。因此儘管原文中都附有校勘資訊，但讀者卻無從得知相關異讀來自哪些本子。

相較之下，PTS 版光碟的情況就好了許多，每一冊都將校訂者的「引言」(Introduction)原原本本的收錄進去，讀者可以藉以了解該冊的相關內容、底本資訊、以及編輯原則等。但美中不足的是，不少原典沒有注明其初版年(first published)，而只列出光碟版所據紙本的出版年(published)——經常只是重印年(reprinted)，未能正確反映該冊原典實際編校的時間。<sup>7</sup>

網路上所收羅的譯本(英譯為主)，其出版資訊就更不足了。許多譯文都未交代來源出處，是否經過授權，是否經過校對，有的甚至連譯者是誰都完全沒有注明(例如馬來西亞 Mahindarama)。這些都是極待改善之處，否則其價值就十分可疑。

數位資源的建構者應該有一基本觀念，就是必須假設讀者沒有(或看不到)其所根據的紙本文獻，而盡可能地將紙本上原來的相關資訊充分提供，否則該資源的價值必然大打折扣——此點對於學術資源尤其重要。

#### (4) 部分資源尚未提供檢索功能

檢索功能是數位化資源之所以能發揮高效能的核心引擎，沒有它，就像鳥折了雙翼。就目前的巴利光碟而言，不管是 PTS 版、CSCD 版、或 BUDSIR IV，都提供了相當好的檢索功能。可是網路版，除了 BUDSIR IV on Internet 功能較佳之外，緬甸版和錫蘭版的功能都顯得相當不足。

緬甸版儘管提供使用者依藏、部、冊、(相應、)品、經之類的結構方式，去找尋及瀏覽經文，並且提供清楚的經文目次以及與 PTS 版的對照頁碼，卻未提供最最重要的功能——字詞的檢索功能，使得它在研究上的效用幾乎折損殆盡。

錫蘭版目前由於還在資料輸入、校對的初步階段，因此情況也不理想。SLTP 只提供原始檔案下載，並沒有線上瀏覽、檢索的功能；MettaNet 雖然好一點，提供類似緬甸版的經典結構方式，可以線上查找及瀏覽經文，但同樣沒有提供檢索功能，這是未來有待補強的。

#### (5) 資料擷取及引用功能仍待加強

目前除了 PTS 版光碟不提供資料擷取之外，CSCD、BUDSIR IV、SLTP 都可以儲存或下載檔案，只是功能還十分簡陋。

<sup>6</sup> 例如：CSCD 依據緬甸第六結集版；BUDSIR IV 的三藏及注釋等分別來自 Syamarattha 版及摩訶朱拉隆功大學等數個版本；SLTP 的三藏根據 Buddha Jayanti Tripitaka Series，未來三藏注釋將依據 Simon Hewavitarne Bequest edition。

<sup>7</sup> 例如：「相應部」第一冊初版應是 1884 年，但光碟版列的是 1973 年。

CSCD 是以原緬甸版(Myanmar)頁碼為單位(from~ to~)來儲存檔案(Save To File)，因此儲存的檔案中經常包含許多多餘的資料，這些必須再人工的方式檢查、編輯；更麻煩的是，巴利「語詞」(包括複合詞等長串連在一起的詞組)遇到跨行時都被切成兩半，使用者必須逐行一詞一詞接回來，或者逐一在前半詞之後補上“-”，以免被誤為二詞。這樣，即使只是要引用 3、5 行原文，同樣必須進行一連串繁複的頁碼定位、存檔、開檔、資料篩選、語詞接合等機械動作，十分不便。

BUDSIR IV 的光碟版資料擷取相當不便，<sup>8</sup>網路版雖有改善，但仍有一些限制。網路版中，系統只提供單頁下載功能，一次只能擷取一頁，且系統限定存檔動作最多 5 次，因此每次登錄最多只能取得 5 頁資料，而且是分在 5 個檔案中，仍須後續的編輯工作。不過如果只需要少量資料，系統允許使用 Windows 的「標示·複製·貼上」功能，這是比較友善的地方。

相較之下，錫蘭版提供的擷取功能就比較簡便，SLTP 可直接讓使用者下載檔案(但須自行搜尋經文所在之處)，MettaNet 也提供檔案存取及直接複製的功能。

PTS 版目前仍未提供資料擷取功能，這是比較可惜的地方。因為它儘管還有一些錯誤，但目前還是學術界最為倚重的版本。未能擷取，研究者引用文獻時，就只能重新輸入。重新輸入不只是費時費工，更大的問題在於可能有打字錯誤的危險，這使得資料的可靠性更沒有保障。就目前所見的一些書籍或文章的巴利引文，真正嚴謹校對的並不多，最後的結果經常是滿目瘡痍，不能卒讀。<sup>9</sup>

原典是所有研究的第一手根據，研究者對它的「引用頻率」及「引用量」，自然是最高的。因此方便、友善的擷取與引用功能，對於學術界及研究者的幫助實在十分重大。它不但節省了眾多研究者不斷重複輸入的「勞務成本」，同時更能維護及確保基礎資料的正確性。

## (6) 資源建構尚未達到「高精審」的規格

以上所談，相對而言是屬於比較微觀的問題。而如果從比較巨觀的角度看，目前巴利數位資源的限制在於：還沒有達到「高精審」(higher criticism)的規格。

前面談到，當前巴利數位資源建構的總體趨向就是「低精審」(lower criticism)——不管是 PTS 版、緬甸版、錫蘭版、還是泰國版——儘管傾力搜羅建構所有的巴利文獻資源，但畢竟還是侷限於巴利一系當中。但目前學界普遍的認識是，巴利文獻以及其他犍陀羅語、或佛教梵語等初期佛教文獻，很可能都是從其他中古印度雅利安語轉譯過來的，因此巴利文獻並不能直接視為最原型的佛教文獻(包括語言與內容)。這樣，單就巴利本身的資料進行研究，事實上是有所侷限的。

要達到更高的精審度，就必須盡可能地至少納入同時期的其他傳承的其他語

<sup>8</sup> BUDSIR IV for Windows 無法在中文 Windows 系統下使用；而早先的 MS-DOS 版，要擷取利用經文，比 CSCD 更加繁複，不僅資料存檔後須重新編輯(因為會有一些段落重複)，並且還須經過數道轉碼程序，才能在 Windows 下使用。

<sup>9</sup> 筆者曾經校訂過一些出版社的書，有時薄薄一冊不算多的巴利引文當中，就有上百處拼讀錯誤。

本，進行綜合的考察。這不只是一要對它所承載的初期佛法取得更好理解的必要手段，即便只是一要對巴利文獻本身的各方面(例如語言、組織、內容、傳承間的變化等)取得更好的了解，都不得不同時參考其他語本進行對比的研究。

因此，從低精審的思惟與規模，開始廣納同一時期的其他語本文獻，從而邁向高精審的思惟與規模，這將是巴利數位資源的建構者下一階段必須努力的方向與目標。

### 3. 巴利數位資源的未來展望

#### - - 略談「初期佛教聖典多語多本平行語料庫」

##### 3.1. 從「低精審」邁向「高精審」

如前所述，巴利數位資源建構的一項重要目標就是支持研究者提高研究的精審度。因此，當前資源的一些不足之處，未來應持續改善，這些包括：更精良的版本校訂及標讀、更充足的原典及譯本、更好的辭典及工具資源、更清楚的版本資訊、更強化的資料檢索及擷取功能等。此外，雖然已有部分資源已經具備有聲資料庫、多語資料庫的特性、以及適合中文系統的執行環境，但大多數則沒有此種功能，這也是未來必須加強的。

除了巴利內部資料的精審化，未來最需要的是：突破目前的思惟與規格，將資源建構的規模往「高精審」推進——結合同一時期的不同傳承的不同語本，將巴利資源置於整個初期佛教聖典的總集當中，讓研究者更容易進行比對研究，以便對聖典的各方面求得更切當的了解。其實這原本就是學術界行之有年的研究方法，數位資源的建構應該反映這種事實，以便與真正研究的操作模式更緊密地結合，提供研究者更符合實務需求的幫助。下面就是一個基於這樣思考的擬構案例。

##### 3.2. 一個擬構案例：「初期佛教聖典多語多本平行語料庫」

本節將提供一個擬構的案例，以「初期佛教聖典」為收錄文本的核心範圍，以「高精審」為指導目標，結合「多語言」、「多版本」的相關材料，做到「平行對讀」的功能，並且對於相關語料進行「語言屬性」的標記，以期能在現有研究功能之外，更進一步拓展「語言研究」與「語言教學」的功能特點。

### 3.2.1. 略說「初期佛教聖典·多語多本·平行·語料庫」

#### (1) 初期佛教聖典

初期佛教聖典主要指佛教最早期的文獻集成，也就是是記錄佛所教導、施設的「法」與「律」。「法」就是阿含經，目前數量最大的是巴利語系的五尼柯耶以及漢譯的四阿含。「律」是律藏，目前數量最大的也是巴利語系(赤銅鑠部)的律藏，以及其他部派所傳的漢譯五部廣律——十誦律、四分律、五分律、摩訶僧祇律、以及根本說一切有部律。以原典(若原文已失，則譯典也不妨視為準原典)為收錄處理的核心文獻，以翻譯及注釋、解疏等為輔助文獻。

以初期聖典的「整體」為範圍建構資料庫，有幾個原因：一是，時間上有同期性。儘管這些經、律經過數百年時間、在不同地區、由不同的僧伽團體受持流傳，其內部已有一些演進改變的痕跡，但總體而言，並不巨大，反映的大致是同一時期、同一階段的佛教樣貌。二是，內容上有一體性。儘管經、律各有不同的意趣或性質，而四阿含的各部也有不同偏重，但總的來說，它們卻是密切相關、互為補充、不可切割的一個整體。因此將它們作為一個整體，進行各面向的比對研究，可以得到最大的效益。

#### (2) 多語、多本

早期的經、律目前流傳下來的有眾多語言的眾多傳本或譯本：(A)印度雅利安語系的有巴利語本、佛教梵語本、犍陀羅語本等。(B)古代譯本有漢譯本、藏譯本等。(C)現代譯本更多，如中譯、英譯、日譯、德譯、法譯等。「多語、多本」意指佛典流傳、翻譯的複雜情況，有時「一語一本」，有時「一語多本」，有時「多語一本」，但更多時候是「多語多本」，以《法句經》為例，就有巴利語本、佛教梵語本、犍陀羅語本、漢譯本、藏譯本等多種語言的多種傳本(或譯本)，而若再加上各傳本(或譯本)的現代語譯本就更多了。

將多種語言的多個傳本(或譯本)收在一起，可以起到相當多面的研究效用，例如文本的內容問題、語言問題、翻譯問題、傳誦或傳承的演變問題等等，雖說許多個別經典的研究，早有不少學者做過了，但若就整個龐大的初期佛教聖典的各種各樣的問題而言，尚待掘發之處必定更多。除了提昇研究效能之外，多種語本的總集，也能提供教學的便利，拓展教學的深度與廣度。當然，它也開展了另一種聖典集成的方式，不只是依傳統不同語系(巴、漢、藏等)而集成聖典，同時可以透過數位資源靈活的彈性，依文本的內容為主，讓同一內容而散落諸方的聖典，總集在一起。

### (3) 平行

要讓這些經、律的不同語本可以起到更好的比對、研究的效果，並非單單把它們收在一起就夠了。以 MettaNet 為例，雖然同時收有巴利原典、英譯本、僧伽羅語譯本(算是「一本多語」的例子)，但是並沒有做到「平行」的結構，也就是沒有把同一經文或同一律文的三個語言的本子連結在一起，因此也無法替讀者省去找尋目錄的機械操作。

進行平行關聯的「文獻單位」不能太大，否則就沒有意義了。單位越小，就像使用更高倍率的顯微鏡，讀者將更容易進行微觀的研究，但建構者自然要更費力氣。以經藏而言，最大的平行單位至少應該是「小經」，例如「雜阿含」、「增一阿含」的一一小經，但對於「中阿含」、「長阿含」而言，因為經文長，單做到經對經，其實效用並不大，因此有必要進行更小單位的切割。

除了「經對經」之外，平行單位若能更小，做到「段對段」、「句對句」甚至「詞對詞」，它的學術價值以及可以起的效用就更高。當然這裡牽涉的就不只是電腦技術的問題，更關鍵、更費力的是在於文本的精細判讀工作。判讀錯誤，反而提供給使用者錯誤的訊息。這種工作需要大量的專業研究人力。

### (4) 語料庫

語料庫是作為語料的文本集合，因此原始佛教聖典的總集，自然也是語料庫的一種。將它稱做「語料庫」，是因為在這個應用案例中，將著重於這些材料在佛教學研究當中的「語言研究」與「語言教學」等方面的功能特點，並強調它作為支持佛教學研究的「基礎建設」的功能性。

雖然如上所說，聖典的總集(單純的文本堆砌)是一種語料庫，但它只能稱為「生語料庫」(raw corpus)，也就是並沒有經過進一步的「加工處理」——標注語言屬性，因此它能提供的語言研究的效用就十分有限。現代語言學意義下的語料庫，通常必須在文本中加注大量「語言訊息」(或語言知識)，並且有適當的軟體可以提取及處理這些訊息，以便達到計算機輔助的語言研究的效能。

對文本加工處理的深度越深(即存放的語言知識越豐富)，語料庫能夠發揮的語言研究的效能就越大。這些語言訊息的加工標注，基本的，例如音韻訊息(如連音法)、語法訊息(如詞類、性、數、格)、句法訊息(句子結構)，其他還有語意訊息、語用訊息、話語訊息等等。

「標注語料庫」除了提供語言研究的效能之外，它也為原典的翻譯提供基礎並且記錄過程。古代的譯經大德，儘管業績顯赫，但可惜的是，並未將他們對於文本解讀的珍貴語言知識流傳給後人。因此，標注語料庫也有佛法傳承的重要價值，當然，對於今日及未來苦學原典語言的學子們，它更是一個輔助學習的語言寶庫。此外，它也是編寫更好的文法書、更精良的辭典的最好幫手。

### 3.2.2. 必須進行的相關基礎工程

上面大略談過「初期佛教聖典多語多本平行語料庫」的一些基本想法，儘管整個概念是簡單的，但是當中實牽涉到大量的基礎建設，並且這些主要的關鍵點並不在於資訊科技的技術問題，而在於相關語言文獻學的基本訓練。以下舉列幾項這個案例所須進行的相關工程建設：

- (1) 巴利語本、漢譯本【重新精校、標點】
- (2) 巴利語本、梵語本、犍陀羅語本、藏語本【語料標注】
- (3) 巴利語本、梵語本、犍陀羅語本、藏語本【中譯】
- (4) 各語本及譯本(含現代語譯)之間的經對經、段對段、句對句、詞對詞【平行連結】

上述任何一項(或任何一項的一小部份)，都是重要的學術工程，都無法由機器代勞，而需要專業人力一字一詞地校訂、標注、翻譯、判讀，可以想見此項工程之耗時費力了！但由於初期佛教文獻基本上具有「封閉性」——不會有大的增殖，即使未來有新出土的材料，數量依舊是有限的。因此儘管耗時費力，終究是可能逐漸完成的。

### 3.2.3. 能夠開發的教學研究效能

如上所說，多語多本平行語料庫的建構是極耗時力的，但它卻可以帶來多方面的效能，並且這些效能將會隨著不同領域研究者的開發，而不斷增長。以下簡要舉示幾項：

- (1) 語言研究：音韻、構詞、詞義、句法、語用、語篇分析各方面。
- (2) 語言教/學
- (3) 翻譯研究
- (4) 翻譯教/學
- (5) 版本校勘/校訂
- (6) 教理研究
- (7) 經典集成史研究
- (8) 經典傳承史研究
- (9) 文法書編纂
- (10) 辭典編纂

#### 4. 結語：建構一個多效能的佛典知識寶庫

佛教聖典自從 2500 年前第一次結集開始，就作為佛教文化的知識寶庫而流傳。其後隨著佛法的傳播，聖典也以種種不同語言而轉譯、翻譯、筆錄、輾轉傳抄。後來大藏經的編輯，便是保存這個聖典寶庫的具體體現。

如今，文本的載體有了重大改變，一片片的光碟、一個個的網站，給了這個寶庫全新的樣貌，也帶來以往做不到的應用效能。但是，這些發展儘管已經有相當成績，卻還只是個起步。如何運用各專業學科的研究方法，配合資訊科技的技術，對這個寶庫進行種種深層加工，提取潛藏在文本當中的種種知識，以便讓它發揮更高的研究效能、教學效能、以及佛法傳承與傳播的效能，這些都有待吾人後續的努力。