

翻譯中的語言規範問題： 談網路的語料庫功能

解志強

翻譯是一個涉及兩種語言的活動。語言是一種約定俗成的慣例，使用同一種語言的人可以自由無礙的溝通，這是因為一群人對於字的定義和句子的理解有一定的共識。一般人在耳濡目染的情況下學會某一社區的語言用法，包括不斷創新的詞法或語法。不過一個人大腦的統計紀錄能力畢竟有限，我們很難經由內省得知整個社區作用於無形的語言規範，包括單詞、搭配、語法、語意、語用等各層次的原理與規則。幸運的是，自從有了語言記載工具之後，人類可以寫作和錄音方式檢視自己使用過的語言，從而找出其中的規則。特別是網路的發明，使大量文字記載變得垂手可得，只要通過電腦程式的整理和篩選，就可以一窺廣大社區人群的語言使用案例和規則。對於一個翻譯學者，網路語料庫是值得研究的一個議題。而對於一位專業翻譯師，網路則是一種愈來愈重要的資源。

本文由語言的規範談起，探討語言的準則問題，主要介紹語料庫語言學的主旨與研究方法。作者舉例說明了如何從語料庫中萃取字詞的用法與慣用的搭配。接著再討論翻譯學中所謂編譯規範（norm）的概念，重點放在譯文撰寫過程當中網路所能扮演的角色。本文示範了如何在翻譯的選詞、搭配、造句過程當中，利用網路搜尋引擎來獲得最符合原文和最能為讀者所接受的譯文。最後，本文提出一個以網路語料為本的翻譯模型，包括一般民眾和專家學者在語言規範和發展上所扮演的角色，以及翻譯家在整個語言社區中所擔任的職務與社會責任。

關鍵詞：語料庫、網際網路、搜尋引擎、語言規範、搭配

收件：2009年10月29日；修改：2010年7月2日；接受：2010年7月3日

On the Issue of Norms in Translating: The Function of Web as Corpus

Chris Shei

Translating is an activity involving two languages. Language has conventions to allow people to communicate with each other smoothly. A person grows up in a certain discourse community learning the language through exposure. However, it is difficult for human minds to recall the norms of language in a conscious way. Fortunately, now that the computer network is prevalent, a lot of language data can be analyzed by a computer program to extract frequencies and usages and present rules applicable to a particular discourse community. For a professional translator, the Internet has become an increasingly indispensable resource. This article first discusses the norms of language and explores the issues of standard usages for a certain discourse community. Examples are used to demonstrate how word usages and collocations can be extracted from the Internet to facilitate the recognition of preferred linguistic conventions. The article then goes on to consider the concept of “norm” in translation studies, emphasising the role Internet plays in the process of translating. More examples are used from results of web search to show how observations of authentic usages can facilitate the choices of words, collocations, grammatical structures, and phraseological units to conform to the conventions of the target language. Finally, this article proposes a model of translation based on the Web as Corpus, explaining what roles language users can play in the formulation of language conventions, as well as what the translator’s roles and responsibilities are in this machinery.

Keywords: corpus, the Web, search engine, norms, collocation

Received: October 29, 2009; Revised: July 2, 2010; Accepted: July 3, 2010

壹、語言

對語言的研究興趣自古就有，從柏拉圖、蘇格拉底到亞裡士多德都有對語言作出思考，接下來的斯多葛學派更提出形式（signifier）與意義（signified）的區分，比索緒爾（Saussure）早了兩千年（Robins, 1990, p. 19），顯示有心人士早就意識到語言外在和內省的兩種意義。希臘和羅馬語言哲學家注重文法和構詞學的探討，最終以普裡西安的一本拉丁語文法書結尾，表現出人類對語言規則的研究興趣。值得注意的是，根據羅賓斯的說法（Robins, 1990, p. 99），從中世紀開始已經有一批思索派文法家（Modistae）習慣使用杜撰的語句來證明自己的文法理論，這個方法預告了二十世紀杭士基衍生變形語法的來臨（Chomsky, 1957, 1965）。杭士基將語言分成能力（competence）和表現（performance）兩種層次，認為文法是一種人類特有的能力，文法是純淨的，不可和實際產生的不規則語料混為一談。因此，杭士基的研究方法也以杜撰的“未受污染”的句子為基礎。從1960年代開始約有二十年時間杭士基的文法理論佔有語言學界，主流句法家以少數杜撰的句子引申出一整套學說，結構學派布倫菲爾（Bloomfield）式的田野調查快速沒落，市井百姓到底如何使用語言？語言規則應該從內省產生或由實際語料歸納而來？這些問題未受到應有的重視。

對於人類的語言能力，杭士基雖然提出一些有趣的基本論點，例如語言的衍生性、語言的生物基礎、人類共通語法等；在方法學上也帶來很大的貢獻，例如樹狀圖的結構表達法，科學化的語料分析等，但是大部分句法理論都太過於抽象和封閉，缺乏真實世界的解釋和應用功能，這就預告了杭士基語言學派和其他著重實務的語言相關理論——例如翻譯學理論，無法相輔相成的事實。二十世紀大約同一時期，翻譯學理論逐漸脫離語言學和文學的旁支地位，成為一個獨立學門。在濫觴初期，某

些翻譯理論家難免受到杭士基學說的影響。例如奈達（Nina）區別動態對等（dynamic equivalence）與形式對等（formal equivalence）的翻譯理論，亦即深入句子內部求取意義精髓，似乎和當時杭士基所謂深層結構（deep structure）和表面結構（surface structure）的學說相呼應（Quah, 2006, p. 24）。即便是現代，也有部分學者將杭士基的 E 語言和 I 語言當成翻譯學研究和授課的理論基礎（ten Hacken, 2005）。但是大致說來，翻譯是一門講究實用的科學，對於高度抽象的杭士基理論，後來的翻譯學者並沒有給予太大的關注。例如杭士基的變形句法理論，就被認為不適合用來架構機械翻譯引擎（Quah, 2006, p. 32）。能夠和近代翻譯學並駕齊驅的語言學理論，似乎還未誕生。

以杭士基為首的理論派在西方語言學界的版圖並未持續擴張，從 1970 年後期開始，英國陸續出現了一批語料庫語言學家，利用日漸普及的電腦和電子文字，通過大量語言的處理而找出其中的規則，這種由下而上的具體方法，為語言學研究指出了另一可行之路。甫於 2007 年過世的辛克萊（Sinclair）被許多英國語料語言學者視為前輩和開路先鋒。例如同樣知名的語料學者史大博（Stubbs），最近在全球應用語言學旗艦刊物 *Applied Linguistics* 上發表專文，大力推崇辛克萊在語言教育、言談分析和詞彙學方面的貢獻（Stubbs, 2009）。辛克萊的過人之處在於他能夠利用簡單的電腦程式來分析大量純文字語料，從而發現許多英文用詞和搭配的規則。其中最重要的貢獻應該是他改變了一般人所持“文法重於一切”的觀念。辛克萊的研究結果指出：人類使用語言的原則是詞彙的選擇決定了一個句子的文法形式，也就是說，每個詞彙都有其特性，詞彙和它的特殊搭配環境決定了句子的走向。因此，語言的研究不應該以句子為本，而應該以詞彙為中心。詞彙為中心的意思並不是指只研究單詞本身的語意或文法特性，而是指它特有的搭配情形（即經常出現在它左右的字）和其他語境因素（如喜歡出現的文法結構、語意群和語用特性）。

辛克萊首先提出延伸詞彙單位（Extended Lexical Unit or ELU）的

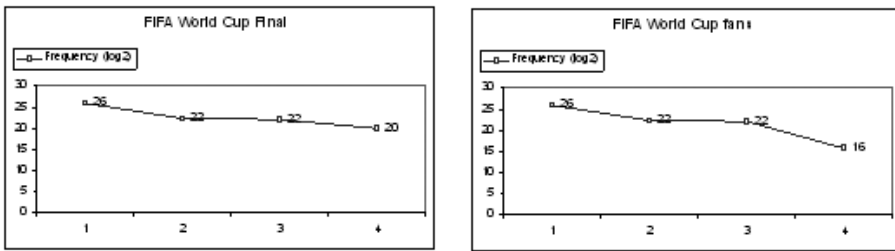
概念，例如像 cause 這個字它慣常出現在 cause damage, cause serious problems 等搭配，所搭配的名詞以負面涵義居多，中間所置形容詞則皆有評價性質。史大博 (Stubbs, 2001) 將 ELU 的概念發揚光大，確認 ELU 在人類語言中的重要地位。作者 (Shei, 2008) 則發現藉由網路搜索可以確認一個字串的慣用語地位。例如我們從表 1 的統計數據中可以發現：利用 Google 逐漸增加 FIFA World Cup Final 這個字串的搜尋字數（從“FIFA”到“FIFA World”，到“FIFA World Cup”，再到“FIFA World Cup Final”），雖然所查到的網頁筆數愈來愈少，但是從圖 1 左圖的平直的曲線圖可以發現，整個片語隨著字數增加而頻率掉落的趨勢以對數的價值來看，仍然是很小的。也就是說，絕大多數人在說了“FIFA World”之後，會繼續說“Cup”，非常多人在說了“FIFA World Cup”之後，會繼續說成“FIFA World Cup Final”。反觀表 2，當我們查詢的字串變成“FIFA World Cup fans”，最後一個階段的頻率明顯急遽下降，也就是說，比較少人使用以“fans”結尾的這個字串。在圖 1 右圖的對數曲線上，可以看出從 22 掉到 16 的一個很大落差。作者 (2008) 所提出的方法和概念受到史大博等學者注意，認為有助於揭開延伸詞彙單位之謎 (Stubbs, personal communication)。

表 1 *FIFA World Cup Final* 的漸進字串 Google 頻率與對數值

Fragment	Frequency	No. of words	Frequency (log2)
FIFA	69,100,000	1	26
FIFA World	5,340,000	2	22
FIFA World Cup	4,600,000	3	22
FIFA World Cup Final	1,030,000	4	20

表 2 *FIFA World Cup fans* 的漸進字串 Google 頻率與對數值

Fragment	Frequency	No. of words	Frequency (log2)
FIFA	69,100,000	1	26
FIFA World	5,340,000	2	22
FIFA World Cup	4,600,000	3	22
FIFA World Cup fans	56,400	4	16

圖 1 *FIFA World Cup Final* 與 *FIFA World Cup fans*

兩個字串的漸進頻率對數曲線比較圖

上述例子同時說明了兩件事。第一：有些字串會形成慣用語，有些則不會。第二：某些形式的慣用語（特別是連續字串不間斷者）可以經由網路搜索而發現或確認其地位。

最後，本節要談的是語言規範（norm）問題，這個用語一般會出現在語言規劃和語言習得的文獻中。前者指的是由政府或民間機構來制定某些語言的使用通則，例如法索德（Fasold, Yamada, Robinson, & Steven, 1990）一文討論了某報社對旗下工作人員所制定的語言使用規範。後者關心的是在學習外語過程中對當地語言規範的處理方式（如 Vito, 1991; Timmis, 2002）。在本文中，語言規範指的是某個族群在某一特定時間空間和某一媒體上所表現出來的最大語言共識。在此定義之下，所有台灣網民在網路上所發表的文字，就構成一個語言規範的母體，而作用於其中

的有形無形的語言規範，很多可以利用搜尋引擎找尋出來。在此觀點之下，網路中每一個網民都是語言規範的繼承者、維護者，甚至新典範的創造者。

總而言之，高度抽象的語言學理論和憑空杜撰的句子，無法解釋人類實際使用語言的情形。對於翻譯來說，譯文可讀性非常重要，譯入語必須考慮到當地族群的語言規範，才能符合當地人的閱讀習慣，增加資訊接收的效率與閱讀的愉悅。語料庫語言學所代表的經驗主義精神，最適合擔任翻譯學的語言學理論基礎，本文這項主張在接下來各節的闡述中，將會愈來愈清楚。下節我們討論近代翻譯學研究中，與本文主旨較有關係的一些學說。

貳、翻譯

現代翻譯學之父奈達（根據 Quah, 2006, p. 36）指出，過去翻譯理論過於偏重文學和譯文的典雅，導致當今很多翻譯課程都與現實脫節（Nida, 2001, pp. 1-2）。奈達認為關於文化之間溝通的語意學和符號學固然很重要，但更重要的是用易懂的語言來解釋理論，並將理論靈活地融入練習當中。奈達認為譯文的滿意度有三個層次（Nida, 2001, p. 3）：

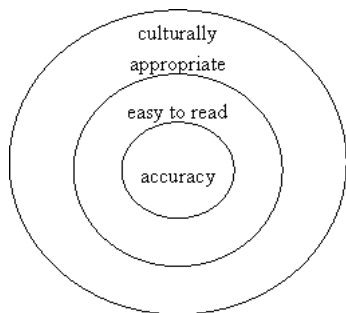


圖 2 奈達的譯文滿意度圖示（本文作者整理繪圖）

奈達認為，譯文的準確度是翻譯最基本的要求，如果還能做到易讀

易懂，就是更進階的的層次，若能進一步達到文化方面的貼切，那就是很成功的一篇譯文。本文作者認為，要達到文化方面的貼切，一個方法就是利用網路查詢當地人的用法。而根據此種驗證方法推演出來的理論，就是奈達心目中與實務結合的好理論。

翻譯中難以釐清的一個問題就是對等效應 (equivalence)。亦即翻譯時如何找到和原文對等的單詞、搭配、句子、語意和語用功能等。由對等的概念所衍生出來的問題就是譯文應該忠於原文 (literal) 或是自由發揮 (liberal)，亦即不同的對等價值觀會產生不同的翻譯策略。歷史上關於對等的討論主要圍繞在語言各層面，近代翻譯理論逐漸重視在翻譯過程中所應考慮的文化因素，討論同樣問題的不同詞彙因運而生，例如凡努提所談的外國化 (foreignization：保留原文的外國風味) 與當地語系化 (domestication：讓譯文貼近本地文化) 兩種截然不同策略 (Venuti, 1995)。1980 年代，紐馬克提出更精細的翻譯策略模式，將原來的二分法細分成八種不同取向的翻譯方法 (Newmark, 1988)。但是在紐馬克的模式當中，最常被提及的還是語意翻譯 (semantic translation：譯文達到和原文相同的語意和語境) 和溝通翻譯 (communicative translation：譯文造成和原文一樣的溝通效果) 兩種。而這兩種譯法，根據曼岱的見解，就相當於奈達著名的形式對等 (formal equivalence) 與動態對等 (dynamic equivalence) 兩個概念 (Munday, 2008, p. 44)。由此可見，主流翻譯理論始終圍繞著譯文的語言操作和文化取向這個核心問題，爭辯原文和譯文的對等到底應該體現在哪一層面。並且，除了一些用途特殊的翻譯之外 (如某些專業翻譯與語言研究用的翻譯)，一般學者和譯者似乎和奈達的見解相同，認為良好的翻譯除了意義上的準確之外，還需要流暢易讀並且符合本地的文化習俗。

除了因不同的對等價值產生不同的翻譯策略之外，在翻譯學的討論中還有一個較高層次的議題，就是規範 (norm)。規範這個概念在翻譯學研究中可以指對譯文的語言要求，也可以指對翻譯模式的瞭解和描述。根

據夏夫納的說法，翻譯理論中先有語言規範的概念，在圖瑞（Toury）的重要著作之後才有翻譯行為規範的理論（Schäffner, 1999）。翻譯中的語言規範學說有兩種意義，一是研究譯入語中的語言規範，包括語音、詞彙、句法、語意、語用、語境等語言各層面適用性的考慮。另一是對譯語產生的正確方法與策略的探討，通常指的是兩種語言經過對比分析之後所得到的文法或構詞轉換規則。而以圖瑞為首的學者所研發的翻譯規範則是上述第二種意義的擴大，考慮到的不只是語言轉換的層面，而是翻譯過程中所涉及的任何決策、限制、意識與成品的考慮，規範是整個社區人員的行為標準，是一種內建的價值觀，在經過對許多實際翻譯過程的研究之後可整理出規範的蛛絲馬跡，從而產生的規範描述則有助於我們對翻譯過程和考慮因素的理解。本文在提出語言規範取得的程序當中，也同時提供了翻譯流程規範的意義。亦即，除了討論語言規範的種類和性質之外，本文所提出的尋找語言規範的過程，也是協助專業翻譯流程規範的一項努力。前者相當於賈斯特曼的預期規範（expectancy norms），後者則是賈斯特曼所稱的專業規範（professional norms）（Chesterman, 1993）。

參、網路語料庫與語言規範

本節將前述有關語言規範的理論付諸於實際，說明如何利用網路搜尋引擎來獲得各階層的語言標準。重點主要置於譯文的處理，但是對於原文的理解和詮釋也有著墨。

一、原文的理解

筆者過去十年的翻譯教學經驗顯示，學生從外語譯入母語時最容易犯的錯誤就是對原文的誤解。除了電子詞典之外，現代學生經常依賴機械翻譯來取得整句整段的翻譯雛型，對於句法較複雜或語意較不直接的結構，利用這些“速食工具”常會導致錯誤的結果，輕則表達不足，重則

完全曲解原文。相反的，如果我們將網路上所有含文字的網頁當成一個語料庫，我們可以針對某一用語查出豐富的語境，在仔細觀察和推敲一些用例之後，我們可以得到較正確的讀法。雖然這不是垂手可得的技能，但卻是翻譯過程中一道有益的程序，也是翻譯和外語學生應該擁有的一項技巧。

例如，最近有一篇取自 BBC 的新聞英語交由翻譯班學生譯成中文，原文中有句子如下（附加三位同學譯文。同學來自大陸在英國攻讀翻譯科技碩士，不一定是文科背景。每位同學有一獨特數碼代號）：

(1)

Carers claim it can be a struggle getting the right kind of help. (BBC, 2009, 全文參見附錄)

- a. 護工斷言得到權利的一些說明將是需要努力的 (453285)
- b. 護工們聲稱他們可以爭取到相應的援助 (506592)
- c. 護工們聲稱他們會更努力的使得患者獲得適當正確的護理 (447303)

根據上下文，(1) 的正確翻譯應該類似“護理人員指出，(為唐氏症患者)爭取適當的醫療設施，有時得大費周章”。但是部分學生的翻譯錯誤或不當情形相當嚴重，且大多因為對原文瞭解不夠深入，而非詮釋能力不足。例如(1a)將 right 誤讀成“權利”，固然顯示其句法能力的不足，完全忽視 kind 的存在而將 right 解讀為中心名詞，但筆者以為，學生若能將整段不瞭解字串“the right kind of help”輸入網路搜尋引擎（加雙引號以免字串被拆開處理），將可以語境的觀察補充句法能力的缺陷，輔助對該文段的瞭解。這個過程和兒童學習母語類似，先藉由語境對整個字串產生瞭解，待語言機制成熟後才進行內部分析。以下試從 Google 所搜得的五十萬例句中摘列出幾句加以分析：

(2)

- a. Deafblind children need **the right kind of help** to learn.
- b. Providing **the right kind of help** and services is essential for the person's future progress and quality of life
- c. The Government's support for the wrong kind of aid now needs to be countered by bolder Tory plans for **the right kind of help**.
- d. Dr. David Hawkins discusses ways to find **the right kind of help** for marriages in trouble.
- e. If you are having debt problems, and are at risk of home foreclosure, turn to **the right kind of help**. Unfortunately, there are a lot of places out there claiming they want to help you that aren't being sincere.

由(2)的例句發現，the right kind of help 會出現在有疾病或婚姻、財務問題的場合，這樣的語境讓人感受到援助的需要。而像(2c)與(2e)這樣的例句更明白指出援助有錯誤的(wrong)也有不真誠的(not sincere)，如此對比之下，更能使學生領悟到 right kind 的正確意思，不會寫出像(1a)那樣嚴重錯誤的譯文。

除了 the right kind of help 之外，(1) 句中還有一個字串學生顯然並不完全瞭解：it can be a struggle。(1a) 中的同學將它譯成未來式“將是需要努力的”，誤差極大。(1b) 譯成“可以爭取到”將原來負面的意見變成正面的訊息。(1c) “會更努力的使得患者獲得”將原來或然率的表示變成一種決心。以上學生的翻譯展現對 it can be a struggle 各種錯誤的理解，雖然我們可以將此誤解歸因於對情態助動詞 can 的瞭解缺陷，但是，筆者認為藉由網路搜索找出同樣結構的例子，可以幫助學生瞭解原文，得到正確的詮釋。以下是經 Google 搜索號稱多達一千萬筆資料的其中幾句：

(3)

- a. **It can be a struggle** finding weight loss pills that are going to work online.
- b. **It can be a struggle** to give up habits you have had your entire life,...
- c. You live month by month and **it can be a struggle** to get on top of bills.
- d. Let's face it, **it can be a struggle** to come up with snacks that are not only healthy, but that your kids will eat!
- e. Wages are often low, so that in towns with high cost of living such as Oxford or London **it can be a struggle** simply to survive.

從(3)的例句可見，it 在本句型中所指大部分是艱難的事物：找到有效減肥藥、改變個人習慣、應付每月帳單等等。如此語境有助於學生體會“it can be a struggle”當中的無奈意味，進而瞭解 can 在本句型當中的機率（而非能力或決心）用法。事實上，利用作者所提出的方法（Shei, 2008）將 it can be a struggle 這個字串繪出的對數頻率曲線如圖 4 左半部所示，呈現相當完好的直線，表示它是一個習慣性用法，和右半部另一個意義類似字串 it can be a combat 的頻率曲線差很多（字串頻率進行到 combat 後急遽下降表示字串 it can be a 之後很少接 combat），可見這是值得研究和學習的一個語意和句法單位。因此，學生利用 Google 搜索字串是一個釐清原文的好方法，同時也是一個學習外語的途徑。

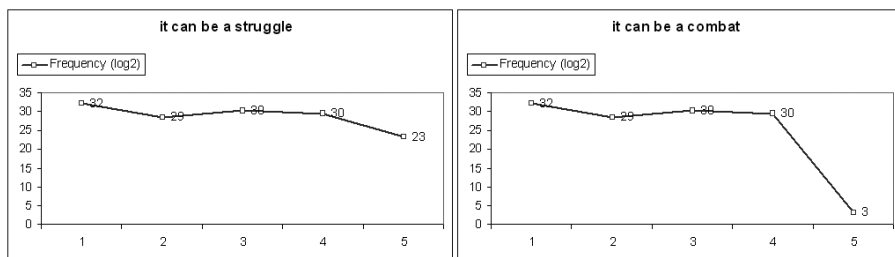


圖 3 慣用語和非慣用語漸進字串頻率對數曲線比較圖

上文舉例說明了網路語料對理解原文的作用，接著我們討論在譯文撰寫過程中，網路語料在不同語言層次上的輔導功能。

二、用詞的選擇

紐馬克認為翻譯時單詞所帶來的困難有兩種：一是對詞的意思不瞭解，另一是發現該單詞很難翻譯（Newmark, 1988, p. 33）。關於原文的理解上節已經提出網路語料庫可以著墨的地方，本節將討論網路查詢如何有助於單詞翻譯。

根據筆者的教學經驗，學生對外語單詞譯入中文的問題有好幾種：詞的語意範疇問題、詞性轉換問題、詞的搭配問題、詞的語用和文化問題等等。(4) 是一個語意範疇套用不佳的例子：

(4)

Nikki Lewis describes her brother Andrew as “warm hearted, loving and gregarious.” (BBC, 2009)

妮基路易斯說她的兄弟安德魯是一個非常“熱心，忠誠而且合群”的人。(564491)

雖然(4)中 loving 的語意詮釋本身就有問題，但是就“忠誠”的選詞角度來看，也有語意範疇不符合的問題，因為“忠誠”一般會用在特別

需要強調操守或執著的場合，例如對國家、對配偶、對品牌的忠誠等等，而（4）的語境並不需要強調忠誠這個特質。面對譯文選詞中有關語意範疇的問題，學生可以將候選詞輸入搜尋引擎觀察它的用法。例如 Google 所提供的某些“忠誠”語境如下：

(5)

- a. **忠誠**是一個人最重要的高潔品格你會因為自己的忠誠贏得老闆的信任
- b. 在「品牌管理」中，對於面臨市場競爭的品牌產品組合最重要的要瞭解消費者**忠誠度**
- c. 在同志之間，在追隨者與教主之間，作為個人操守的“**忠誠**”就上升為政治上的立場

觀察諸如（5）當中的例句，可以使學生瞭解“忠誠”這個詞在現代網路用法中常見的語境——大部分為商業用法，少部分涉及政治或個人操守，但並沒有像（4）那種形容某人一般個性特質的用法。如此，藉由網路搜索可以推論某一候選詞的不適用性。相對的，利用搜索也可以找出較為適當的用語。例如，在 Google 輸入譯句中較為確定的字彙如“熱誠”、“合群”等，可得到一些有用的線索：

(6)

- a. 80 **合群**守法勤懇**熱誠**，119 待人和藹機智敏捷。5 努力謙恭循規蹈矩，43 謙恭有禮聰明和藹，81 埋頭進取努力求知
- b. 勤懇**熱誠**，**合群**守法。學行均優，堪以造就。富正義感，有責任心。沈靜用功，深知自愛。刻苦好學，持之有恆。勤誠穩重，力學不倦。熱心公務，捨己為群
- c. 溫和助人、品學兼優**合群**有禮、力學不倦開朗活潑、**熱誠**大方、謙恭有禮

“熱誠”和“合群”是(4)中相當於 warm hearted 與 gregarious 的可信翻譯，將這兩個字輸入 Google 查詢得到包括(6)在內的一些一般性人格評語。這些詞都是作用在同一語意範疇內的詞彙，從中觀察很可能得到另一並列的人格特質 loving 的恰當翻譯。事實上，(6a)中出現的“待人和藹”就是適合語境的一個良好翻譯。由此可見，網路語料庫不但可以幫助譯者剔除語境不適合的候選詞，更可以幫忙找出較為恰當的譯詞。

另一個常見的學生選詞錯誤是詞類轉性問題，亦即將某個中文候選詞用作不習慣充當的詞類。例如：

(7)

The green paper is based on making the system fairer, simpler and more affordable for everyone. (BBC, 2009)

他們希望使系統能夠‘更公平，更簡單，更經濟’地去幫助所有患有唐氏綜合症的病人。(447303)

在(7)中的學生翻譯將 affordable 翻譯成“更經濟(地)”，詞類轉換(從形容詞到副詞)原來就是一種翻譯技巧，詞類轉換過程本身沒有問題，問題在於“更經濟地”在現代漢語中是否有很多人將它當成副詞使用？經分別查詢兩岸的網頁發現都有這樣的用法，大陸的網頁較多(“更經濟地” site:.cn 有 35,500 筆)台灣的筆數較少(“更經濟地” site:.tw 僅有 410 筆)，各舉兩例如下：

(8)

- a. 讓企業主可以更輕鬆、**更經濟地**營運自身的網站
- b. 尋並測試可用的軍需項目或技術，使能夠更快速和**更經濟地**滿足美軍
- c. 順著數學網友的思路談談如何才能**更經濟地**改變社會結構
- d. [60 名家談電] 葉喬波：讓我們**更經濟地**使用電能

這項搜尋結果令我們發現，“更經濟地”在網路上確實有人使用，但是案例非常稀少，譯者是否應該採用，似乎可以參考「翻譯標的論」(skopos theory)的說法，視翻譯的目標和對象而定(參見 Vermeer, 1989)。在注重閱讀愉悅性的場合，例如商業廣告或政策宣導網頁，應當避免使用非主流表達法。在以訊息內容為主的場合，例如學術性或技術檔，像“更經濟地”這種詞彙結構或許反而較能傳達所需訊息，引起閱讀反感程度較低。

詞彙的搭配錯誤或突兀(marked)是學生常犯的另一個單詞階段的錯誤。搭配(collocation)在外語學習和翻譯作品中的重要性已有許多中外學者提及(Shei & Pain, 2000; Sun & Wang, 2003; Chang, Chang, Chen, & Liou, 2008)，本文僅就如何利用網路搜索協尋適當搭配的部份加以補充。現以同一篇 BBC 原文為例來探討譯文中的搭配問題：

(9)

There are approximately 40,000 people with Down's syndrome in the UK, and thanks to medical advances they are living longer

最快速獲得搭配的翻譯方法當然是兩個字分別直譯，例如 medical advances 查詢 Dr. eye 兩字直譯分別是“醫學的”和“前進”(advance 第一個名詞意思)或“發展”(advance 第二個名詞意思)。“醫學的前進”依照本地人語感似乎很不自然，可以剔除(雖然查詢 Google 仍有五千多筆資料，但大部分案例中兩個字分屬不同結構，不是搭配用法，例如“代表了現代醫學的前進方向”)。另一個直譯“醫學的發展”較有可能，經由 Google 搜索獲得八十幾萬筆資料，有可信度，但不是很強。為取得更適當的搭配譯法，我們可將關鍵詞“醫學的”單獨輸入 Google，觀察其自然語境，如有符合原意的搭配，再輸入 Google 看其出現頻率是否更高。例如“醫學的進步”出現在諸如“西醫和中醫共同推動整個人類醫學的進步”例句中，是另一個可能搭配，搜索 Google 的結果有九十萬筆，以搜

尋引擎的習慣性誤差來看，和“醫學的發展”的八十多萬筆相差不遠，兩個搭配應該都可以使用。

接下來談到詞的語用問題，亦即語言實際使用原則和理解問題。貝克膾炙人口的《換句話說》(*In Other Words*)一書用“連貫”(coherence)一詞來談翻譯中的語用問題(Baker, 1992)。貝克認為“銜接”(cohesion)是譯文文段本身必須具有的特質，即代名詞指稱、量詞代換、結構省略、連接詞連結、同義字替換、搭配詞前後呼應等等，都必須有條不紊，重現或有效重整原文的銜接效果。而另一方面，“連貫”則是指保留(或詮釋)原文文字背後所隱含的規則與提示，讓譯文讀者可以啟動相同或類似的概念結構，以獲得文字背後的理解。現以(10)的譯句為例來說明語用問題和網路所能提供的解決方案：

(10)

Not enough is being done to support the growing number of people with Down's syndrome who have dementia, say campaigners.

- a. 一些社會上的人說，我們沒有做出足夠的努力去幫助那些越來越多的唐氏症癡呆綜合症患者(517656)
- b. 還沒有足夠的努力來緩解日趨增長的唐氏綜合癡呆症的人數，相關人士評訴(447303)
- c. 不斷增加的唐氏症癡呆症候群沒有從候選人那裡得到足夠的支援(564491)

(10)中 campaigners 這個字的意思需要對西式民主社會有相當理解才能體會，而這幾位碩士學生所提供的翻譯都不正確：“社會上的人”、“相關人士”、“候選人”。當我在課堂上討論 campaigners 這個詞並且提議“社運人士”的翻譯時，清一色來自大陸的同學竟然面面相覷，完全不懂這個複合詞的涵義，任由我更進一步說明也無法瞭解——後來反覆詢問才知道在大陸沒有(或極少)有抗議活動和社運人士存在，因此在他們

文化教養過程中從未接觸過社會運動這個概念。當下 campaigners 這個詞的翻譯立即變成兩個層次的問題：首先譯者本身必須瞭解這個詞所代表的概念和其社會意義，其次在本身瞭解之後，譯者必須設法讓來自相同背景的譯文讀者也獲得同樣的瞭解。在譯者本身求瞭解的部份，如本文第一節所示，可以在 Google 輸入適當原文然後觀察案例。為求語意範疇的一致，我們可以輸入 dementia campaigners（不帶雙引號，兩字可以打散），結果得到諸如（11）之中的例子：

(11)

- a. **CAMPAIGNERS** are preparing to take part in three simultaneous charity walks on Saturday to raise funds for a new **dementia** support service
- b. Thousands of Alzheimer's patients are suffering because of widespread NHS cutbacks in **dementia** services, **campaigners** claimed last night
- c. **Dementia campaigners** say they are frustrated with the government which they claim has been slow to come up with a strategy for the disease

類似（11a）這樣的例句，明確點出 campaigners 和慈善健行之間的關係（即從事慈善活動的人），可以幫助譯者對原文的瞭解。在譯文的傳達方面，除了加註解釋之外，對於 campaigners 這個字的處理是否有其他更簡潔的方式？事實上，我們可以先從查字典著手，例如查詢 Dr.eye 發現 campaigner 的意思是第一：出征軍人；老兵，第二：從事社會運動的人。第一個意思明顯不對，我們將第二個意思不帶雙引號輸入 Google 的大陸網域（即：從事社會運動的人 site:.cn）獲得如下的結果：

(12)

- a. 因為改革是一項涉及人與人之間權利調整的社會運動
- b. 他們通常會將精力用在組織團體或從事社會運動上
- c. 佛陀也是人間佛教社會運動的先驅，打破印度「種姓制度」，宣導「四姓出家」
- d. 有許多從事社會運動的革命者泡在其中，邊喝酒或咖啡，邊發洩著
- e. 巴老早期是個社會革命家，是從事社會運動的

由(12)大陸網頁中有關社會運動的句子可以發現，“社會運動”常和“改革”、“組織”、“革命”等詞一起出現，比較傾向於二十世紀上半舊式社會主義意識形態。在改革中的社會主義成形以前，譯者或許必須將譯詞嵌入此種意識形態的慣用詞彙，以求譯文讀者能對 *campaigners* 一詞產生類似原文讀者的理解，例如“醫療制度改革者”、“無產階級革命先驅”等等。

本節一共討論了語意範疇、詞性、搭配、語用等四個有關翻譯時詞彙選擇的問題。筆者示範了如何利用網路搜索來找到正確的語意範疇用字、可接受的詞性轉換、流利的搭配，以及符合社會文化背景的意識形態用詞。下一節將討論在非文學翻譯，特別是技術文件翻譯中常會碰到的術語處理問題。

三、術語的譯法

術語 (term) 和術語集 (terminology) 的研究近年在歐洲十分盛行，主要是為了科技翻譯的需要而生。如今有許多術語理論家討論術語的定義方式、術語集的結構模型、術語與術語之間的關係等等 (例如 Wright & Budin, 1997; Maia, 2005)，主流翻譯輔助軟體廠商如 SDL Trados、Déjà vu、Star Transit 等也都在翻譯記憶軟體之外，另行推出獨立性強的術語編

排系統。術語資料庫的建立令譯者對某一特殊領域的知識得以通盤掌握，在用字方面和該領域主流用法取得一致，在概念方面則有更快和系統化的參照和理解機會。

本節以 BBC 文章為例，討論網路語料庫在術語身分確認和譯詞選擇上的輔助功能。在 (13) 的兩個句子中出現了三個醫學術語：Down's (syndrome)，Alzheimer's disease 和 dementia。

(13)

As many as 50% of all people with Down's in their 50s may now have Alzheimer's disease, as they often get the condition much earlier than others.

But critics say the issue was barely mentioned in a recent green paper on dementia care in England.

查詢一般詞典就術語而言並不是一個好方法，因為術語領域眾多，編纂字典的人不可能樣樣都精通，能夠熟諳每一領域的習慣用法。例如 Down's syndrome 查詢 Dr.eye 第一個解釋是“道恩氏症候群”，其次是“蒙古症”和“先天癡呆症”。我們將“道恩氏症候群”用雙引號輸入 Google 查詢只有 189 個。由稀少的案例可以推斷“道恩氏症候群”一定不是 Down's syndrome 通用的譯法。此時我們可以在 Google 輸入英文，使用高階搜索模式將網頁語言指定為繁體中文，目標是希望求得有此名詞中英對照的網頁。查詢結果有三千多筆，仔細翻閱之下可以找到幾筆中英對照的例子：

(14)

- a. 近年，能夠存活至中年甚至老年的唐氏綜合症（「唐氏」，Down's syndrome）人士越來越多
- b. 調查比較 30 名唐氏症候群 (Down's syndrome)，174 名智能障礙 (Mental ...

- c. 另外一種染色體異常疾病為唐氏症 (Down's syndrome)
- d. 像大家所熟識的唐氏綜合症 (Down's syndrome)，就是因為第廿一組的染色體有三條之故

觀察 (14) 中的結果發現，專業人士對於 Down's syndrome 的譯法似乎不只一種。於是我們將所得的各種中文譯法在分別以雙引號方式輸入 Google，得到表 3 的結果：

表 3 Down's syndrome 各種中文翻譯法的 Google 搜索結果比較

英文術語	Down's syndrome		
中文翻譯	唐氏綜合症	唐氏症候群	唐氏症
Google 查詢結果	77,000	13,500	370,000

由表 3 結果可以假設“唐氏症”是目前使用繁體中文人士所偏好的一種用法。

以上是筆者所提議的以網路使用為準的術語翻譯法。以相同的程序可以得到 Alzheimer's disease 的最佳翻譯為“老人癡呆症”，而 dementia 的最佳翻譯則為“失智症”。當然，有公信力的學術機構所編纂的術語資料庫也是值得參考的對象，但是和查詢詞典的缺點相同，一般統合性學術機構很難針對每一領域訂出公認的術語用法。例如上述 Down's syndrome 在國立編譯館的學術名詞資訊網查詢，可得到四、五種不同的譯法：

表 4 Down's syndrome 國立編譯館學術名詞資訊網查詢結果

編號	名詞領域	按英文名排列	按中文名排列
1	動物學名詞	Down's syndrome	唐氏症侯群；21- 三體蒙古癡呆 先天性疾病
2	內分泌學名詞	Down's syndrome	唐氏症

3	實驗動物及比較醫學名詞	Down's syndrome	蒙古癡呆症；唐氏症候群
4	畜牧學	Down's syndrome	唐氏徵候群

由此可見，網路查詢對於術語的翻譯有一定的功能。而專業作家包括學術界研究者，在網路搜索可以觸及的著作中，可以考慮多多使用術語中英並列的型態，如此在專業術語由英入中的過程中，經由網路為本的翻譯流程推動，對術語標準化的建立將有莫大的功效。

四、結構的合法性

結構的問題應分為文法和 phraseology（姑譯為“套詞學”）兩種類別來討論。在學生的譯文中經常會發現讀起來怪怪的句子，俗稱為翻譯腔（translationese）。翻譯腔的型態有很多，前文所提的語意、搭配、詞性、語用等問題等都有可能產生奇怪譯文。但是造成譯文閱讀障礙的最大原因應該是句法或套語結構的違反常態。以下先談文法再談套詞。

“文法”是很抽象的一個概念。貝克（1992）一書對於較抽象的文法理論（如杭士基的變形語法或「移動 α 」等）並沒有加以關注，只著眼於具體可見的文法層級，如單複數、性別、人稱標記和動詞時式時貌等。這些問題與中文相關不大，因為中文沒有這一類的文法標記。依照筆者的觀察，學生中譯文在文法方面出現的問題可籠統分為兩大類：

- 結構不合法
- 結構合法但詞彙套用不當

以下各舉一例子說明這兩種文法乖謬情形：

(15)

- a. Not enough is being done to support the growing number of people with Down's syndrome who have dementia

對於已經開展的幫助越來越多的先天愚型癡呆的行動是不夠的

(453285)

b. Nikki is convinced the problem is not recognised at government level

尼克相信這個問題卻沒有在政府層面上承認 (517656)

和 (15a) 的語病相同的還有另一位同學的另一個句子：“他們現在可以活得更長久要感謝醫療上的不斷的新發展” (447303) ——一個句子當中有三個“的”，而且每個“的”之前的結構特性各不相同（有介詞片語“醫療上”，也有副詞“不斷”和形容詞“新”），造成閱讀速度減緩和詮釋困難。而 (15a) 句中“的”連續結構更有極端複雜的包孕情形，造成可讀性差和閱讀障礙，令人完全無法瞭解句意。(15a) 的不良結構可以簡易樹狀圖分析如下：

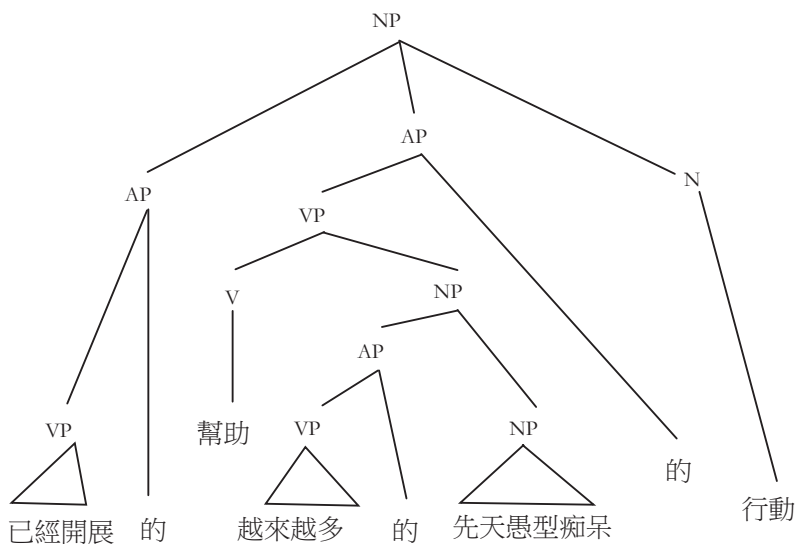


圖 4 不良句中“的”結構的重複與包孕示意圖

關於“的”結構的合法性問題，我們可以在 Google 輸入“*的*的*的”（*代表萬用字元，Google 可以接受萬用字元的查詢），觀察所得的網頁使用案例。我們發現在撰稿此時有 25,270,000,000 筆含有此種結構的網頁片段，顯示三個“的”以上的中文結構並不少見。但仔細觀察可以發現，這些案例中大部分是條列式的簡短形容詞串連結構，例如“環境考驗是長期的、複雜的、嚴峻的”，或是同一形容詞重複數次，例如“我是真的真的真的很愛你”。也有以逐漸加大名詞片語方式構成複雜片語，但是結構很清楚，例如“曹操的長子的妹妹的老公的老爸是誰？”。其中最複雜的案例大概是類似“‘能被 3 整除的數的特徵’的教學與思考”這樣的包孕結構，它的簡易樹狀分析圖如下所示：

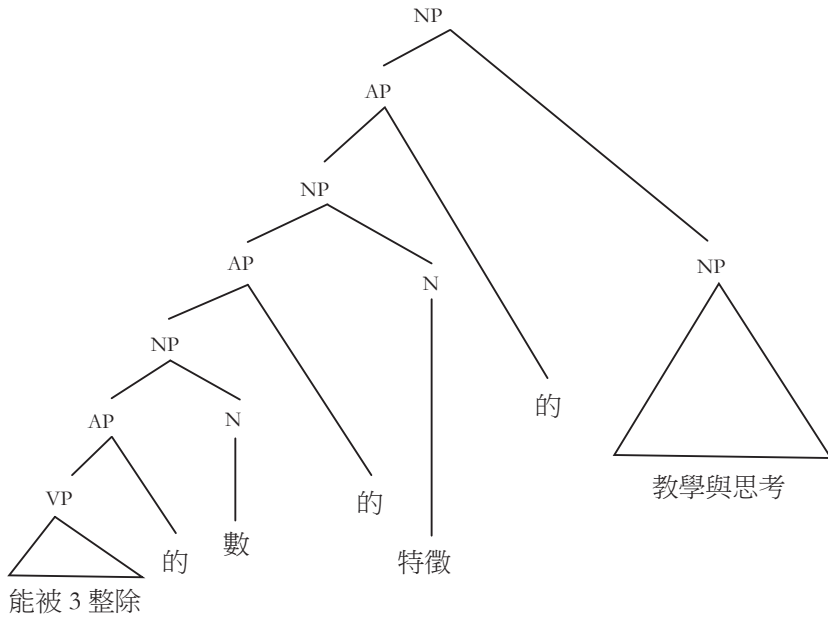


圖 5 網路搜索較複雜“的”結構分析圖

圖 5 顯示，網路上可見的“*的*的*的”複雜結構，只是形容詞片語和名詞片語交互循環逐漸加大的形式，在語意詮釋上雖然不直接，但

也不困難。和圖 5 學生所翻譯出來的動詞、名詞、形容詞片語反覆包孕的複雜結構相差很多：學生根據原文杜撰出來的文法結構令人完全無法理解。利用網路語料庫來觀察例子進而整理出合理結構，是一個可行的校正方法。雖然學生不見得能用樹狀圖來分析結構，但是仿照上述利用萬用字元來查詢 Google 並觀察例子，多少能夠令人一窺該結構的主流用法，察覺自己想用結構的缺點。

接下來討論第二種常見英翻中不妥翻譯的結構疏失，即學生所寫句子雖然符合在地人使用母語的習慣，但是某些欄位填入了不適當的詞彙，相關例句是學生的翻譯（15b），句中“這個問題卻沒有在政府層面上承認”這個句子讀來繞口，原因是該句型雖然符合文法習慣，但卻填入了不恰當的詞彙。這種現象，可以用語料庫語言學的“句型文法”（pattern grammar）學說來解釋（參見 Hunston and Francis, 2000）。根據句型文法的說法，每一個單詞都會有它習慣出現的句型，反之，每一種句型也都有它慣常適用的字彙群。雖然韓士登和法蘭西斯是以英文為研究對象，但是句型和詞彙交互生義的現象應該每種語言都有。例如我們以（15b）中的可能句型“在政府層面上”查詢 Google，結果有 26,800 筆，包括以下使用案例：

(16)

- a. 所以在**政府層面上**，雙方一直保持著良好的關係
- b. 粵港澳在**政府層面上**建立了一個合作的機制
- c. 不過雙方的談判並沒有在**政府層面上**進行

觀察（16）的句子可以發現，以“在政府層面上”為中心的句子，可以將它當成主位（theme）放在句子最前面（16a），也可以當成述位（rheme）放在主詞後面（16b 和 16c）。不過和學生譯文（15b）最有關係的結構是（16c）。試將兩句並列分析如下：

（15b）問題卻**沒有在政府層面上**承認（學生）

(16c) 談判並沒有在政府層面上進行(網路)

(15b) 和 (16c) 兩句的結構其實是一樣的：都是無主句，主要述語也都是動賓結構（“進行談判”，“承認問題”），但是“承認問題”這個結構就是不適合這種句型。這情況和“殺很大”被許多中文老師指為不合文法類似：雖然有“賭很大”、“虧很大”的用法，但是基本上“V 很大”句型裡面的動詞 V 必須和形容詞“大”相容。“賭”和“虧”的金額可以很大，但是“殺”的動作就很難和“大”聯想在一起。然而，“殺很大”在獲得大多數民眾接受後，無形中修正了社區的語感（參考下文將提到的霍依“詞彙打底”理論），使“V 很大”當中的 V 變得更有彈性，“騙很大”、“燒很大”、“瞎很大”、“跑很大”等前所未有的用法相繼出籠，充分顯示社區語感的可塑性與“由外而內”力量的可觀，也讓杭士基完全忽略實際語料的方法更加可議。但本文雖一再引用網路語料庫為例說明語言規範由使用者決定，有一點卻不可忽略：譯者本身的責任主要是傳達，而不是創新。如果一位譯者完全不顧當前社會使用語言的習慣，一篇翻譯作品裡有一百種獨創用法，那是很不負責任的作法。媒體或社會大眾可以創新，但是翻譯家的主要任務應該是觀察和整理。

最後在結構方面要討論的是套詞問題。本文所稱的套詞學（phraseology）是一種套用固定或半固定語言結構的說話方式。即是辛客萊（Sinclair, 1991）所謂的「成語原則」（idiom principle）。近 20 年來很多研究實際語料的學者都發現，人類說話的方式主要不是運用文法規則來創造全新的句子，而是利用一些熟諳的半成品結構來拼湊成句。研究者所提出的固定或模糊結構很多，例如前文所提的辛客萊和史大博提倡的延伸詞彙單位、聶庭傑和狄卡瑞可的「詞彙短語」（lexical phrases）（見 Nattinger & Dicarrico, 1992）、瑞伊所闡述的「公式語言」（formulaic language）（見 Wray, 2002）等。而霍伊（Hoey, 2005）的 lexical priming（姑譯為“詞彙打底”）理論更能從語言習得的角度來解釋語料學的發現。他主張所有

的詞彙都有其特定的使用環境，包括常用搭配、並列文法結構、語意偏好、語用特性等等。而我們身為在地人生活在特定社區天天耳濡目染下，習慣了一些以詞彙為首的常用套句，每次一想到某一詞彙，就連帶引出一些相關結構，讓我們能夠快速、流暢而達意的使用語言。

套詞學的考慮並不是譯句合不合文法？而是譯句是否採用常見的詞彙組合方式，讀來順暢？試比較（17）中兩位學生對同一英文句子的翻譯：

(17)

However, ministers said action was being taken to reform dementia care (BBC, 2009)

- a. 然而部長們表示，正在採取行動改善癡呆症護理
- b. 可是部長們卻說政府已經採取了癡呆症護理的改革方案

“正在採取行動改善癡呆症護理”和“已經採取了癡呆症護理的改革方案”這兩種表達法，哪個較符合中文使用者的習慣呢？讓我們先分離出該二句的主要結構，分別輸入 Google 觀察相關的使用案例。也就是說，我們可以去掉時式記號和特定名詞，只比較“採取行動改善**”和“採取**的改革方案”這兩個結構在網路上的使用頻率。在“採取行動改善”方面，我們查詢 Google 得到 57,200 筆數，包括（18）中的一些例子：

(18)

加拿大**採取行動改善**食品安全

奧巴馬承諾迅速**採取行動改善**經濟

香港廣場鳩山認為北韓應先**採取行動改善**日朝關係

要鼓勵學生**採取行動改善**自己的健康

必須**採取行動改善**美台之間的密切關係

而在“採取**的改革方案”方面，查詢 Google 實際只有 41 個部份

重複的例子，(19) 列出其中幾個：

(19)

解決股權分置問題必須結合上市公司的不同特點**採取多樣化的改革方案**

針對不同險種基金運行特點，**採取綜合性的改革方案**

社會經濟相對發達地區和經濟效益相對好的單位一般**採取貨幣化的改革方案**

由於俄羅斯**採取的激進的改革方案**並沒有完全達到預期目的為確保實現“十五”期間既定建設目標，宜**採取漸進式的改革方案**

由(19)的例子可以看出，大部分“採取**的改革方案”當中萬用字元所代表的元素，在網路的例子中大多是形容詞，如“多樣化”、“激進”、“漸進式”等。(17b)學生所拼湊出的結構“採取了癡呆症護理的改革方案”顯然不是很符合在地人習慣的用法。反過來說，另一個學生(17a)“採取行動改善癡呆症護理”的用法在網路上可以找到很多類似結構，是一個接受度較高的用法。

網路查詢除了提供頻率和案例讓我們比較各種候選結構的優劣，也可令我們發現原先不在候選池中的結構。例如我們只輸入“採取行動”可以找到比用“改善”來直譯 reform 更通用的其他用法，比如“解決”：將“採取行動解決”輸入 Google 可得 141,000 筆資料（約為“採取行動改善”的三倍），標準使用案例如下：

(20)

英首相籲 G8 **採取行動解決**糧食危機

潘基文敦促希、土兩族**採取行動解決**賽普勒斯問題

“世界水周”呼籲各國**採取行動解決**水資源問題

金融時報澳大利亞央行**採取行動解決**流動性問題

穆沙拉夫呼籲以色列繼續**採取行動解決**巴以爭端

由(20)中相當一致的案例可以看出這個結構的紮實性：它以“解決問題(或危機)”這個搭配為中心，適用於政治金融民生等場合，表達刻不容緩的情緒。這個集語言多層次意義於一體的結構，很適合用來表達(17)的原文——例如“政府正在採取行動解決癡呆症護理的問題”。另外，在量化和圖形的表達上，“採取行動解決”也明顯優於原來學生的直譯結構“採取行動改善”，請參見表5和圖6。

表5 “採取行動改善”與“採取行動解決”頻率比較表

字串	Google 原始頻率	詞數	log ₂ 頻率
採取	9,840,000	1	23
採取行動	9,040,000	2	23
採取行動改善	57,200	3	16

字串	Google 原始頻率	詞數	log ₂ 頻率
採取	9,840,000	1	23
採取行動	9,040,000	2	23
採取行動解決	141,000	3	17

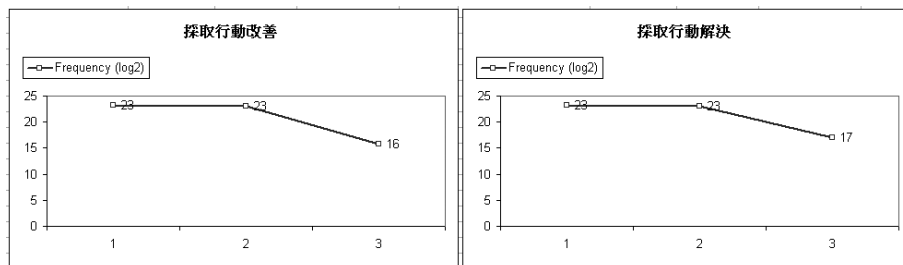


圖6 “採取行動改善”與“採取行動解決”頻率對數曲線比較圖

從表5我們可以看出，在“採取行動”之後接“解決”的案例大約

有“改善”的三倍之多，圖 6 則用視覺效果反映了“採取行動解決”在套詞學上略優於“採取行動改善”的情形。本例所提供的數據對比雖不是很強烈，但主要目的在說明：適當的網路查詢技巧，不只可以讓我們比較候選結構的優劣，更可以讓我們無意中發現原先未想到的更佳表達法。

四、網路語料庫導向的翻譯研究

語料庫語言學最初以固定建成的語料庫 (corpus) 為底子，對所蒐集的語言做各種方式的處理，例如 tagging (詞類標示) 和 parsing (句法分析) 等等，然後以特別設計的電腦程式 concordancer 來搜索語料庫，做成各種統計和觀察。雖然有名的英語庫諸如 British National Corpus 和 Bank of English 等都號稱有一億以上的字，但這和網路浩瀚的文字比起來只是九牛一毛而已，很多較大單位的字串在這些固定語料庫中都找不到例子 (例如本文的 *it can be a struggle* 字串，這兩大語料庫都沒有例子，而 *the right kind of help* 在 Cobuild 只找到一個例子，BNC 只有四個例子)。因此，以網路作為語料庫是一個必然的趨勢——歐洲計算語言學會 2006 年在義大利 Trento 舉辦了 Web as Corpus 研討會，2007 年 Hundt、Nesselhauf 和 Biewer 出版了 *Corpus linguistics and the Web* 一書，同年 Guo 和 Zhang 出版了一篇很有意義的以 Google 為工具從網路蒐集英語搭配的文章。語料庫語言學大師辛克萊偏好使用純文字無電腦程式標記的語料做研究，並且主張語料庫愈大愈能看出語言規則 (Sinclair, 2004)，網路語料庫剛好符合這兩個條件，大量的語言使用案例讓我們可以看到語言全貌，也可以找出最微小而有趣的例子。

網路語料庫很難定義。2006 Trento 研討會主席 Kilgarriff 和 Baroni 認為網路基本上是一個語言物件 (a linguistic object) 可以看成是“一個很大的袋子，裡面有一袋又一袋的語言” (a bag of bags of words)。這個大袋子可以讓我們更了解語言。網路語料庫的優點包括大堆頭字數、免費和使用方便 (不必另行購買和安裝軟體)、熟悉的使用者介面、不斷自

動更新等。但是網路語料庫的研究也有它的誤區存在，例如商用網路搜索引擎的黑盒子 (black box) 科技令人無法得知原理和改進 (見 Renouf, Kehoe, & Banerjee, 2007; Ripple, 2006; Leech, 2007)，Google 所聲稱的字串出現頻率和重複性令人起疑 (Fletcher, 2007; Lüdeling, Evert, & Baroni, 2007; Leech, 2007)，網路文件在文體類型上是否有代表性 (Thelwall, 2005; Leech, 2007)，以及網路多變性所導致的搜索結果不可重複性 (Lüdeling et al., 2007; Fletcher, 2007) 等等，都令人在使用網路語料庫作為研究基礎時感覺憂心。但是以上這些作者也都承認網路和搜索引擎有以上所提的具體研究功能，特別是在提供夠細的語言規則 (Thelwall, 2005) 和新的語言用例這些方面 (Leech, 2007)。另外，本文作者也一再證明，Google 和其他搜索引擎所得可以用來進行字串的慣用強度比較。

綜合以上所述，本文作者提出一個以網路為本的翻譯理論模式，在此模式當中，翻譯家扮演了推波助瀾的角色，適當的辨認和採用一般詞彙或術語的標準型態。而一般寫作者和專業作家和研究者在無形中或特意將術語中英並列的情況下，也肩負了造就一般語言和術語標準的責任。

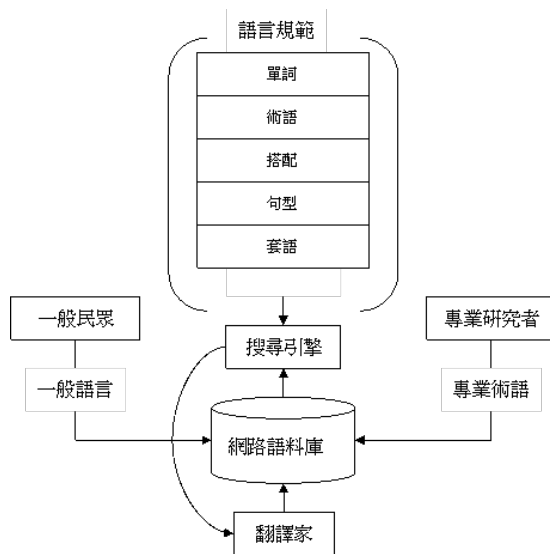


圖 7 以網路為本的翻譯與網路語料庫關係圖

在翻譯學界廣受認可的何姆斯翻譯理論分類法 (Holmes, 1988 請見 Munday, 2008, p. 10)，先將翻譯理論分成純翻譯學 (pure) 與應用翻譯學 (applied) 兩大類，在純理論部分先分成理論性 (theoretical) 和描述性 (descriptive) 兩種，前者又分為一般性的 (general) 和片面的 (partial)——例如只著眼某種媒體、某個地區、某種文體、某個時代的翻譯，後者則可分為產品的描述、過程的描述、功能的描述等等。至於應用翻譯學的研究，則可分為翻譯教學、翻譯輔助工具、翻譯批評等等。整個翻譯理論體系有如表 6 所示：

表 6 翻譯學理論分類表

翻譯學	純翻譯學	理論性	一般理論	
			片面理論	媒體
				區域
				級別
				文體
				時代
	問題			
	描述性	作品		
		過程		
		功能		
應用翻譯學	翻譯教學			
	翻譯輔助工具			
	翻譯批評			

資料來源：研究者根據 Munday (2008, p. 10) 整理

根據表 6 的分類，本文所提出的以圖 7 為代表的翻譯理論，在純翻譯學的理论 (以實際語言使用為規範) 和描述 (以網路查詢為基準的翻譯過程) 方面均有著墨。在應用翻譯學方面，除了講解網路搜索工具的

某些使用方式外，也為翻譯教學指出了一些方向。換句話說，本文所提出的理論，在某些層面上，將何姆斯分類法上的多種翻譯理論做了有益的融合，對某些翻譯議題提供了跨層次的綜合看法。

伍、結論

本文從兩種截然不同的語言學理論出發，主張根據理性論（rationalism）所發展出來的語言學缺少實用價值，離著重實務的翻譯學遠矣。而基於經驗論（empiricism）哲學，採用實際語料研發理論的語言學，例如語料庫語言學，似乎更適合翻譯理論和實務的發展。筆者認為各種層次的語言規範應由大眾使用的實際情形歸納而成，而數量最多又容易取得的語料便是網路。使用搜尋引擎所獲得的語言使用案例，可以讓我們觀察單詞和字串的實際用法用量，而搜尋引擎所提供的數據，則可以讓我們就某些類似結構進行比較，決定最佔優勢的文法結構和套詞用法。凡此，都能讓我們以最貼近真相的方式觀察語言、討論語言。

無論何種翻譯理論都離不開譯文，因為翻譯的目標就是排除萬難以獲得最符所需的譯文（參見 Shei, 2005a, 2005b）。雖然不同的理論對於翻譯過程有不同的界定和推論，但是對於譯文在語言和文化方面都有一定程度的要求。本文認為大部分譯文必須符合當地的語言規範，在用詞、文法，和套語各方面參照語言規範而設定最適合該翻譯目標的用語。為了達成此一目標，譯者必須學會如何善用搜尋引擎去開發網路這個巨大的語料庫。在解說此一理念的過程當中，本文作者示範了幾種 Google 查詢的技巧，例如基本的雙引號使用、特定語言或網域的設定、萬用字元的用法、用英文查得術語的中文翻譯等等。

本文最後所提出的翻譯模式，以網路語料庫為中心，點出語言規範的地位和出處，以及一般民眾、專業作家，和譯者在語料庫建立和翻譯過程中所扮演的角色。這一翻譯模型有執行多種功能的潛力：促進譯者

和研究者對網路語料庫功能的認識、啟動網路語料查詢技術的研究和教學機制、激發語言學者對網路社區和語言規範的研究等等。而在純理論的成就方面，本理論結合了何姆斯翻譯學分類表裡面的多種元素，提供了跨層面的因式融合，令人更瞭解整個翻譯理論體系的運作方式，也為未來的翻譯理論點出一個網路查詢為本的過程模式和理想的譯文標竿。

本文所指出的網路搜索方式有其先天性缺陷：商業製搜尋引擎如 Google 等往往膨脹所得，誇大搜索結果筆數，雖然所得案例極有參考價值，但是數據則不可盡信。另外，商業搜尋引擎只能提供最基本的語言使用案例搜索，無法提供較進階的語言查詢功能，例如搭配的偵測、關鍵字的演算等等。尤其是中文的搜索，到底是以詞為本或是以字元為搜索單位，或是某種混合的形式，商業搜尋引擎並沒有提供相關文獻和技術，這就使得中文的查詢不能充分發揮語料功能，數據的誤差也比英文查詢多。近年來雖然有國外學者陸續開發語言研究用的搜尋引擎（例如 Webcorp <http://www.webcorp.org.uk/>），但是很遺憾並沒有查詢中文的機制。希望本文能引起中文計算語言學者的注意，開始研發精確和具備高階功能的中文網路搜尋引擎，則未來將有更多以網路語料庫為本的理论與實務誕生，將網路語料庫的支援翻譯功能發揮得更淋漓盡致。

最後，在翻譯教學的應用上，本文所討論的網路運用技巧，大部分需要成熟的語言技能才能從雜亂的網路搜索結果中看出端倪。因此，最適當的用法就是將網路查詢規劃為一種練習，搭配一個語言學為本的翻譯教程，在教師的指導下進行。例如，翻譯教程可分為詞彙翻譯、搭配、文法、銜接、訊息結構、背景知識等，在每個章節的語言學理論和翻譯學應用講述後，讓學生進行網路驗證，一方面藉由網路實例觀察所學到的語言規則，一方面練習老師所給予的翻譯題目。如此，學生可以對語言學和翻譯理論得到更具體的概念，也可以養成有效和正確使用網路來協助翻譯的技能。

感謝詞

感謝三位審查者的指正與建議，使本文減少一些錯誤和補強了一些弱點。其他不足或錯處，均由本人承擔。

參考文獻

- Baker, M. (1992). *In other words: A coursebook on translation*. London: Routledge.
- Chang, Y-C., Chang, J., Chen, H-J., & Liou, H-C. (2008). An automatic collocation writing assistant for Taiwanese EFL learners: A case of corpus-based NLP technology. *Computer Assisted Language Learning*, 21(2), 283-299.
- Chesterman, A. (1993). From “is” to “ought”: Laws, norms and strategies in translation. *Target*, 5, 1-20.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.
- Crawford, A. (2009, October 4). Down’s dementia risk ‘overlooked’. *BBC News*. Retrieved October 21, 2009, from <http://news.bbc.co.uk/1/hi/health/8287385.stm>.
- Fasold, R., Yamada, H., Robinson, D., & Steven, B. (1990). The language-planning effect of newspaper editorial policy: Gender differences in The Washington Post. *Language in Society*, 19, 521-539.
- Fletcher, W. H. (2007). Concordancing the Web: Promise and problems, tools and techniques. In M. Hundt, N. Nesselhauf, & C. Biewer (Eds.), *Corpus linguistics and the Web* (pp. 25-45). Amsterdam: Rodopi.
- Guo, S. & Zhang, G. (2007). Building a customised Google-based collocation collector to enhance language learning. *British Journal of Educational Technology*, 38(4), 747-750.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London/New York: Routledge.
- Holmes, J. (1988). *Translated! Papers on literary translation and translation studies*. Amsterdam: Rodopi.
- Hundt, M., Nesselhauf, N., & Biewer, C. (Eds.). (2007). *Corpus linguistics and the Web*. Amsterdam: Rodopi.
- Hunston, S., & Francis, G. (2000). *Pattern grammar—A corpus-driven approach to the lexical grammar of English*. Amsterdam/ Philadelphia: John Benjamins.
- Leech, G. (2007). New resources, or just better old ones? The Holy Grail of representa-

- tiveness. In M. Hundt, N. Nesselhauf, & C. Biewer (Eds.), *Corpus linguistics and the Web* (pp. 133-149). Amsterdam: Rodopi.
- Lüdeling, A., Evert, S., & Baroni, M. (2007). Using Web data for linguistic purposes. In M. Hundt, N. Nesselhauf, & C. Biewer (Eds.), *Corpus linguistics and the Web* (pp. 7-24). Amsterdam: Rodopi.
- Maia, B. (2005). Terminology and translation—Bringing research and professional training together through technology. *Meta*, 50 (4), from <http://www.erudit.org/revue/meta/2005/v50/n4/019921ar.pdf>.
- Munday, J. (2008). *Introducing translation studies: Theories and applications*. New York / London: Routledge.
- Nattinger, J. R., & Dicarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Newmark, P. (1988). *A textbook of translation*. London: Longman.
- Nida, E. A. (2001). *Context in translating*. Amsterdam: John Benjamins.
- Quah, C. K. (2006). *Translation and technology*. Basingstoke: Palgrave Macmillan
- Renouf, A., Kehoe, A., & Banerjee, J. (2007). WebCorp: An integrated system for web text search. In M. Hundt, N. Nesselhauf, & C. Biewer (Eds.), *Corpus linguistics and the Web* (pp. 47-68). Amsterdam: Rodopi.
- Ripple, A. S. (2006). Expert googling: Best practices and advanced strategies for using Google in health sciences libraries. *Medical Reference Services Quarterly*, 25(2), 97-107.
- Robins, R. H. (1990). *A short history of linguistics*. (3rd ed.). London: Longman.
- Schäffner, C. (1999). The concept of norms in translation studies. In C. Schäffner (Ed.), *Translation and norms* (pp. 1-8). Clevedon: Multilingual Matters.
- Shei, C.-C., & Pain, H. (2000). An ESL writer's collocational aid. *Computer Assisted Language Learning*, 13(2), 167-82.
- Shei, C.-C. (2005a). Integrating content learning and esl writing in a translation commentary writing aid. *Computer Assisted Language Learning*, 18(1), 33-48.
- Shei, C.-C. (2005b). Translation commentary: A happy medium between translation curriculum and EAP. *System*, 33(2), 309-325.
- Shei, C.-C. (2008). Discovering the hidden treasure on the Internet: Using Google to uncover the veil of phraseology. *Computer Assisted Language Learning*, 21(1), 67-85.
- Sinclair, J. (1991). *Corpus concordance collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2004). *Trust the text: Language corpus and discourse*. London: Routledge.
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell Publishers.
- Stubbs, M. (2009). Memorial article: John Sinclair (1933–2007): The search for units of

- meaning; Sinclair on empirical semantics. *Applied Linguistics*, 30, 115-137
- Sun, Y-C., & Wang, L-Y. (2003) Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, 16(1), 83-94.
- ten Hacken, P. (2005). Unpublished lecture notes on translation theory. Swansea University, UK.
- Thelwall, M. (2005). Creating and using Web corpora. *International Journal of Corpus Linguistics*, 10(4), 517-541.
- Timmis, I. (2002). Native-speaker norms and international English: A classroom view. *ELT Journal*, 56, 240-249.
- Venuti, L. (1995). *The translator's invisibility: A history of translation*. London: Routledge.
- Vermeer, H. (1989). Skopos and commission in translational action. In A. Chesterman (Ed.), *Readings in translation theory* (pp. 173-187). Helsinki: Finn Lectura.
- Vito, N. O. (1991), Incorporating native speaker norms in second language materials. *Applied Linguistics*, 12, 383-396.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wright, S. E., & Budin, G. (1997). *Handbook of terminology management* (Vols. 1-2). Amsterdam: John Benjamins.

附錄：學生英翻中練習之 BBC 新聞報導原文

BBC NEWS

Down's dementia risk 'overlooked'

By Angus Crawford
BBC News reporter

Not enough is being done to support the growing number of people with Down's syndrome who have dementia, say campaigners.

As many as 50% of all people with Down's in their 50s may now have Alzheimer's disease, as they often get the condition much earlier than others.

But critics say the issue was barely mentioned in a recent green paper on dementia care in England.

However, ministers said action was being taken to reform dementia care.

The Department of Health in England says it wants to make the system "fairer, simpler and more affordable - all of which should help people with Down's syndrome".

There are approximately 40,000 people with Down's syndrome in the UK, and thanks to medical advances they are living longer.

However, they appear to be vulnerable to Alzheimer's disease, because a protein thought to play a key role in the condition called amyloid tends to build up in their brains more quickly than in the rest of the population.

Professor Tony Holland, an expert in the psychiatry of learning disabilities at the University of Cambridge, said the issue had effectively been ignored by the government.

"We would like to see it acknowledged at a governmental level and then clear strategies put in place, which may vary across the country, that ensures they have access

to the right services.”

Prof Holland said some areas already had very high levels of care - but others lagged behind.

Carers claim it can be a struggle getting the right kind of help.

‘Very distressing’

Nikki Lewis describes her brother Andrew as “warm hearted, loving and gregarious”.

He had Down’s syndrome, but lived an independent life.

Then, in his 40s, his behaviour changed and his memory began to fail.

“It was very distressing, he hated not being able to find his words, he hated not being able to do all the things he used to do,” she said.

Nikki had to fight, first to get an accurate diagnosis, and then for the care Andrew needed. He died aged 52.

Seven years on she says the system has begun to improve.

“I think the levels of care now are becoming much more professional, it’s very gradual, some areas are far more progressive than others.”

Nikki is convinced the problem is not recognised at government level.

Some good care

Dr Karen Dodd, from the Surrey and Borders NHS trust, said there were examples of excellent care, including a Surrey County Council-run project known as The Cottage.

Doors are painted red to help clients who, because of their illness, lose the ability to distinguish colours.

There are pictures on each door to show what is inside the room, a toilet on one, a television on another, a bubbling pan on the kitchen.

But Dr Dodd said a lack of funding can mean people getting the wrong kind of care.

“You might have someone who is only 40 or 50 going into a nursing home with people in their 80s and 90s,” she said.

Care services minister Phil Hope said the green paper set out plans to redesign the care system so that it works better for everyone.

He said: “The green paper is based on making the system fairer, simpler and more affordable for everyone.

“Doing this will benefit people with Down’s syndrome as much as any other group.”

Mr Hope also said there were plans to improve healthcare for people with learning disabilities, for instance by offering annual health checks through the NHS.

Story from BBC NEWS:

<http://news.bbc.co.uk/go/pr/fr/-/1/hi/health/8287385.stm>

Published: 2009/10/04 23:02:26 GMT

© BBC MMIX