

利用專門可比語料庫結合機器翻譯 自動提取雙語對譯 N 連詞：以合約文類為例

陳碧珠

摘 要

在筆譯工作中，合約翻譯屬於高度專門的領域，合格譯者供不應求。合約文體迥異於一般文件，但同時又具有制式重複的特徵，極適合運用翻譯記憶系統（Translation Memory，簡稱 TM）。近年來 TM 系統在臺灣筆譯市場上的應用日益普遍，其原理是將既有的翻譯建立為雙語平行資料庫（parallel corpus，通常為逐句對齊），從中比對找出重複或近似的文句，而將譯文「回收利用」。因此 TM 系統須有適當的雙語資料庫，才能發揮效益，但這正是合約翻譯運用 TM 系統的限制所在，雙語對譯的合約語料取得困難，而人工翻譯所累積的翻譯資料庫，一方面費時費力，另方面還有譯文品質的疑慮。解決之道可以從不同語言的單語專門語料庫著手，亦即可比語料庫（comparable corpora）；專門領域的可比語料庫彼此雖無對譯關係，但所涵蓋的領域概念、術語及慣用表達，仍有許多交集且互為翻譯。本研究目的在於探討一個可行的方法，從中文與英文合約可比語料庫提取對譯的 N 連詞（N-Grams）。

研究方法是利用統計式機器翻譯與字串相似度比對技術，自動擷取中、英合約當中互為翻譯的詞串。首先建置中文與英文合約可比語料庫，並以語料檢索工具提取核心主題詞，據以建立 N 連詞；其次利用 Google 譯者工具包，產生中文與英文 N 連詞的機器譯文。最後，借用 TM 系統的相似度比對功能，以英文 N 連詞與「中文 N 連詞機器英譯」進行字串相似度比對，兩者若完全相同或高度相符，即表示該英文 N 連詞與對應的中文 N 連詞極可能互為翻譯。最後所得到的中英對譯 N 連詞經由專家評估，發現高度相符的 2 連詞至 8 連詞，對應正確的比率可達到 80% 以上。這些中英對譯的 N 連詞，可以直接匯入 TM 工具中做為術語表，或藉由 N 連詞檢索全句，找出合約譯文的適當句型。

關鍵詞：可比語料庫、機器翻譯、翻譯記憶系統、合約翻譯、N 連詞

Using Comparable Specialized Corpora with Machine Translation for Extracting N-Gram Translation Equivalents: A Case Study of Chinese and English Contracts

Pearl Chen

Abstract

This study is motivated by the need of contract translation. Business contracts belong to a highly specialized genre, characterized by specific vocabularies, domain terms, formulaic expressions and repetitive standard clauses. These features make contract texts an ideal candidate for applying a Translation Memory System, or TM, which searches and retrieves past translations from a database of source texts and their equivalents in the target language, usually aligned at the sentence level. A TM system thus requires a large database of past translations (a parallel corpus) in order to get the best result. And there lies the difficulty in using a TM system for English-Chinese contract translation, as parallel corpora of English and Chinese contracts are scarcely available.

To overcome the limitations of parallel corpora, this study turns to comparable corpora, i.e. monolingual corpora of similar design in two or more languages. Comparable corpora of a specialized domain, though not direct translations of each other, contain domain terms, concepts and fixed expressions that are mutual translations. This study aims to explore a simple yet effective method for extracting such translation equivalents from a comparable corpus of Chinese and English contracts.

The proposed method involves statistical machine translation and string similarity comparison. First, a comparable corpus of Chinese and English contracts is built from texts mined from the Internet. Keyword and key keyword lists are built with concordancer tools, on which Chinese and English N-grams are then built. Third, the N-grams are machine-translated into English and Chinese respectively with Google Translator Toolkit. And finally, the English N-grams are compared with the Google-translated English, using the built-in similarity comparison function of a TM system. English N-grams that exceed a pre-defined match value are automatically mapped to the corresponding Chinese N-grams to establish a list of English-Chinese N-gram pairs. These N-gram pairs are evaluated by experienced contract translators, and the results show that 2-word to 8-word N-grams with a match value of 95% and above have an accuracy rate of 83%. These correctly mapped N-gram pairs can be imported to a TM system as a translation resource, or they can be used as concordance search keywords to retrieve examples of adequate sentence patterns to aid the translator.

Keywords: comparable corpora, machine translation, translation memory system, contract translation, N-grams

壹、前言

資訊科技的進步加上電腦產品日益普及，改革了今日譯者的工作模式，翻譯使用的工具從過去的紙筆、參考工具書，進步到文書處理軟體與電子化參考資源，現在更新增了各式各樣的電腦輔助翻譯工具。善用工具能夠提高翻譯品質與效率，不懂工具則可能連接件的機會都會流失。筆譯職場工作模式的改變，進一步也影響了翻譯的教學與研究，在文學翻譯之外，實用文類漸受重視，有關電腦輔助翻譯工具的應用、其優缺點的探討，亦納入翻譯課程，語料庫翻譯研究及專門文體的研究論文，更快速增加。

在電腦輔助翻譯工具當中，效益最為直接的是翻譯記憶系統（Translation Memory，簡稱 TM）。其原理是將翻譯過的文件儲存為雙語對譯的資料庫，通常以句子為單位，在翻譯新的文件時，TM 工具會自動搜尋資料庫，只要比對得到完全相同或近似的句子，就會自動帶出該句的對應譯文。因此，譯者可以加快翻譯速度，維持一致的翻譯內容；客戶與業者則可節省時間和成本，相同內容不需要重複翻譯。

TM 系統並不能無中生有自動產生譯文，而是將既有的翻譯回收利用，所以文件的內容重複愈多，TM 的效益愈明顯；其次，資料庫中累積的翻譯句對要夠多、譯文品質良好，工具才能夠檢索出有用的參考譯文。但愈是專門的領域，這個條件愈難以達到，職場上使用的中英雙語 TM 資料庫，主要來自人工翻譯，其譯文品質可以維持一定水準，但限制在於費時費力，又需要投入大量的資源（包括人力與經費），難以在短時間內建立足夠規模的資料庫。

在專門技術文件的翻譯當中，合約翻譯可說是最困難的領域之一，合格的譯者供不應求。合約文體特殊，術語眾多、好用長句、句構複雜，不論中文或英文合約，其書寫風格都迥異於其他文類。但另一方面，合約條款多有重複，表達句式相對固定，非常適合利用 TM 系統做為輔助工具。實務上的困難在於雙語對譯的合約語料有限，取得不易。

除了 TM 系統之外，近年來語料庫工具在國外的翻譯界也漸受重視，不過在臺灣的筆譯職場還不普遍。我們日常使用的網際網路，可以視為當今規模最大、語種最多、範圍形式最廣的語料庫，但語料庫語言學所定義的語料庫，是指依特定條件而抽樣選取的語言資料，具有特定的代表性，並且能夠利用語料庫檢索工具查詢不同層級的語言資訊。經過加工整理的語料庫，就語種而言可以分為單語、雙語或多語，就內容題材而言可分為通用（general purpose）語料庫或專門（specialized 或 special purpose）語料庫，就形式而言可分為書面、口語或影音多模態語料庫，就語料年代而言可分為貫時或共時。

涉及兩種語言的雙語語料庫，通常又有平行語料庫（parallel corpus）及可比語料庫（comparable corpus）之分。本研究採用 Zanettine（1998）、McEnergy 與 Xiao（2005）的定義，將雙語對譯的「原文 + 譯文」語料庫稱為平行語料庫；相對的，可比語料庫

則是指兩種語言各依相同的取樣條件所建置的語料庫，兩者的文類、主題領域、年代、規模、溝通功能等條件相當，但彼此並非互為翻譯的關係。相較於平行語料庫，雙語可比語料庫的建置要容易得多，規劃適當的取樣條件之後，即可迅速蒐集大量的可比文本。可比語料庫包含豐富的領域知識、術語、搭配詞、語用、句型等資訊，而且原則上沒有文字品質方面的疑慮，兩個語料庫都不是翻譯的文字，不會出現所謂的誤譯或翻譯腔。

有鑑於平行對譯語料並非隨處可得，尤其是資源較少的語種或高度專門的領域，而單語的語料則為數眾多，因此近十年來學者開始轉向不同語言的可比語料庫，探討如何從中自動提取互為翻譯的術語、片語、對譯句或近義句。這方面的研究絕大多數來自計算語言學或資訊科技領域，須具備計算語言學的專業能力，建立演算模型、編寫電腦程式，通常也需要龐大的軟硬體資源，因此相關研究多半由機構中的研究團隊合作進行。

對於中小型翻譯業者和自由譯者，以及翻譯課程的教師和學生，要等待專業工具上市滿足翻譯上的需求，往往緩不濟急；若要自行開發工具，通常又欠缺電腦資工的专业能力。因此本研究從合約領域的雙語可比語料庫著手，探討一個相對簡單可行的方法，以擷取可比語料庫中所包含的對譯資料，做為合約翻譯的參考資源。

貳、文獻回顧

一、語料庫語言學

語料庫語言學是以語料庫為本的語言學研究，其研究內容為大量甚至海量的語言樣本，研究工具則是各種檢索軟體程式。利用語料庫檢索工具所做的量化分析，基本的統計包括詞表（wordlist）、詞頻（frequencies）、形符（token）總數、類符（type）總數、相異詞比率（type/token ratio，簡稱 TTR）、句數、平均句長等。關於形符的定義，英文每個單字皆視為一個形符，縮寫字如 don't 則視為兩個形符，需要斷開為 do 及 n't，稱為詞形還原。中文的形符單位則是「單詞」而非個別的方塊字，必須經過「分詞」，否則檢索工具便無法處理，若分詞不當也會影響後續的分析統計與解讀。與形符相對的類符指的是出現多少個不同的形符，相同的形符不重複計算。類符數目反映文本真正使用的詞彙量，所有的類符彙整後即為詞表，依各類符出現的頻數可以看到高頻詞、中頻詞或低頻詞的分佈情形，類符與形符總數相除，即為 TTR，*WordSmith Tools*（Scott, 2012）進一步計算標準化相異詞比率（STTR），便於比較不同文本或語料庫的詞彙密度。專門領域的語料庫還可統計出主題詞（keywords）與關鍵主題詞（key keywords），亦即利用統計檢定方法（例如對數似然率〔log likelihood〕或卡方檢定〔x2〕），找出語料庫中明顯異於一般通用語言或文本的用詞特徵，包括出現頻率超高及顯著少用的詞彙。

除此之外，愈來愈多的研究指出，最小的意義單位其實不在單詞，而是多詞結構（multi-word expressions/units）¹。對於重複出現的多詞結構，不同學者提出的術

語繁多，具體定義也不盡相同。根據梁茂成等人（2010, p. 84）所整理的資料，學界常見的用語包括 chunk／詞塊（Sinclair, 1991）、lexical phrase／詞彙短語（Nattinger 與 DeCarrico, 1992）、phraseology／短語結構（Renouf, 1992; Stubbs, 2002）、lexical bundle／詞束（Biber 等人, 1999）、formulaic sequence／制式詞串（Wray, 2000）、cluster／詞簇（Scott, 2008）等，計算語言學領域則稱為 N 元組或 N 連詞（N-grams）。

語料庫依其語種、內容主題、形式、年代、是否標注等考量因素，可有各種不同的分類。其中平行語料庫和可比語料庫兩個術語，不同學者各有不同的用法，需要明確界定。除了 Zanettin（1998）、McEnery 與 Xiao（2005）的定義之外，Lee（2010, pp.119-120）援用 Aston 的觀點，將可比語料庫定義為兩種（以上）語言的原文語料庫，且其領域內容、文類、出版時間等條件相當；單向平行語料庫是指原文語料庫加上其譯文語料庫，有時也稱為翻譯語料庫，而且通常原文與譯文逐句或逐段對齊；雙向平行語料庫則是雙語可比語料庫加上兩個語言的譯文語料庫。可比語料庫和平行語料庫的關係，如表 1 所示；至於計算語言學界對於平行語料庫及可比語料庫，普遍也採用相同的定義。

表 1 可比語料庫及平行語料庫定義

	A 語言		B 語言
可比語料庫（多語）	專門語料庫	≈	相同設計之專門語料庫
單向平行語料庫	專門語料庫	⇒	語料庫 A 之翻譯文本
雙向平行語料庫	A1：專門語料庫	≈	B1：與 A1 相同設計之專門語料庫
	A2：B1 之翻譯文本	↔	B2：A1 之翻譯文本

資料來源：引用翻譯自 Lee（2010）。

雙語語料庫和翻譯的密切關聯不言可喻，平行語料庫可以進行翻譯研究，探討翻譯策略、比較兩種語言的異同，或做為翻譯的參考資源，檢索既有的譯句或術語的對應譯文。Baker（1993, 1995）、Laviosa（1995）、Xiao（2010）等學者則透過翻譯文本語料庫與非翻譯文本的比較，探討翻譯文體的普遍共性、與母語創作文體的差異，或是個別譯者、文類的翻譯文體特徵。

在計算語言學界，從雙語語料庫中提取對譯的詞彙、片語或句子，一直是學者研究的焦點，研究目的則是應用於機器翻譯系統，以提升機器翻譯的效能。雙語對應技術長期以來依賴語言結構分析及人工建置的翻譯模組，但整體效果始終沒有太大突破；1990 年代發展方向轉為從龐大的雙語語句對應資料庫中，搜尋與待譯句相關聯的雙語翻譯句對，再從中以自動或半自動的方式得到翻譯知識與規則（參見林語君、高照明，2009，pp. 168-169）。目前最新的研究趨勢又從平行語料庫再轉向可比語料庫，例如語言資源與評估（LREC）國際研討會自 2008 年開始，每年舉辦「建置與使用可比語料庫工作

1 例如參見 Stubbs（2002）、Sinclair（2005b）、Fitzpatrick（2006）、Granger & Meunier（2008）、Stubbs（2009）、Hunston（2011）。

坊」並出版論文集，第一屆工作坊論文集的序言指出：

Research in comparable corpora is motivated by the scarcity of parallel corpora. Parallel corpora are a key resource to mine translations for statistical machine translation or for building or extending bilingual lexicons and terminologies. However, beyond a few language pairs such as English-French or English-Chinese and a few contexts such as parliamentary debates or legal texts, they remain a scarce resource.... A more fundamental limitation is that translated texts, whatever the skills of translators, are generally influenced by the very translation process and by the language of source texts.

可比語料庫的研究動機來自平行語料庫之不足。平行語料庫對於擷取對譯內容用於統計式機器翻譯，或建立雙語詞典及術語，都是關鍵的資源。但除了部分語種如英／法或英／中，以及少數文本類型如議會議事錄或法律文本，平行語料庫並不常見。……一個更根本的限制是，翻譯的文本無論譯者技巧再好，仍難免受翻譯過程及原文的影響。（Zweigenbaum, Gaussier, & Fung, 2008, p. v，筆者自譯）

Sharoff 等人（2006）亦認為：

[I]t is possible to use comparable corpora to find linguistic constructions that are used for similar purposes in source language (SL) and target language (TL). Even if the equivalence between constructions in comparable corpora can not be complete, there is sufficient similarity between linguistic resources available in the two languages....

利用可比語料庫，應可找出源語和譯語中用於相同目的的語言結構，即使兩者不是完全對應，仍然可以呈現兩個語言之間許多相似的語言資源……。（筆者自譯）

應用可比語料庫以擷取對譯內容，困難在於兩個語言之間沒有明顯的「橋樑」（Sharoff, Babych, & Hartley, 2009），不像平行語料庫能夠在句對應的基礎上，進一步擷取更小單位的對譯內容。因此，相較於平行語料庫，使用可比語料庫擷取對譯資訊之前，必須先確立兩個語料庫的「可比性」，其中的文本內容是否真的相關且有可能存在互為翻譯的句對或詞串，以提高後續資訊擷取的成效。例如 Su 與 Babych（2012）提出以詞彙對應、基於主題詞、基於機器翻譯等三個方式，以量化計算文本的可比性。

不論是文本對應、句群對應、單句對應、詞串對應或術語對應，擷取雙語對譯內容的技術皆不外乎三個主要途徑：基於語言學規則、基於統計分析，以及併用兩者的混合

式方法（參見 Loginova et al., 2012; Pazienza, Pennacchiotti, & Zanzotto, 2005; Thurmair, 2003; Zhang, Iria, Brewster, & Ciravegna, 2008; 季培培、鄒小燕、岑咏華，2010；康小麗、章成志、王惠臨，2009）。基於語言學規則的方法，主要是利用既有的術語詞典，並歸納術語或詞串特定的語言結構，以搜尋比對可能的對譯詞，例如術語最常見的詞類結構為「名詞 + 名詞」、「形容詞 + 名詞」。這個方法的優點是「實現比較簡單，在語言學知識與文本匹配的情況下準確率高，尤其是對低頻術語的抽取效果較好」，但編寫語言規則費時費力且難以窮盡，同時也取決於具體的語種、主題領域，不容易移植應用（季培培、鄒小燕、岑咏華，2010）。

季培培等人將統計式分析分為兩類，第一類是基於統計量度的方法，依據詞彙頻率、分佈等數據，計算「單元性」（*unithood*，亦即詞串能否結合形成結構穩定的詞彙單元，而不僅是搭配詞）以及「術語性」（*termhood*，亦即詞彙單元與特定領域概念的相關程度）。第二類是基於機器學習的方法，主要步驟為：建構訓練語料，根據某種演算法對訓練語料學習生成模型，以測試語料進行擷詞實驗，最後評估該演算法的成效。這樣的方式不需要語言學知識或專家領域知識，但限制在於訓練語料的建構成本，以及訓練語料與測試語料的可比性。

目前的趨勢是合併使用語言學規則及統計法，但如 Zhang 等人（2008）所指出，儘管學者提出了形形色色的術語擷取技術，卻沒有任何一個方法能解決所有的問題，尤其是單詞術語和低頻術語容易因設定的統計閥值而被忽略；另外，各研究的測試語料領域及規模互異，評估方法及範圍也不相同，因此擷詞技術的成效難以直接比較。

二、機器翻譯

如前所述，計算語言學界對於平行語料庫的研究，一大目的是應用於機器翻譯。雖然機器譯文的品質至今仍無法取代人工翻譯，但機器翻譯最大的優勢在於速度與數量，在全球化的今天，翻譯需求有增無減，確實難以依賴人工翻譯滿足所有的翻譯需求。

Hutchins（2010）將機器翻譯的發展分為五個階段：萌芽草創（1933-1956）、過度樂觀至期望落空（1956-1966）、低潮期（1966-1976）、復甦期（1976-1989）、當前發展（1989 年迄今）。在 1956 年到 1966 年的樂觀期，機器翻譯是根據雙語詞典及語言學規則對原文「解碼」，將分析結果轉換為譯文的對應詞彙及語法結構，最後重新「編碼」組合成譯文，稱為基於規則的機器翻譯（*rule-based machine translation; RBMT*）。1980 年代末 IBM 發佈以統計方法進行機器翻譯（*statistical machine translation; SMT*）的實驗結果，日本則有研究團隊開始使用語料庫的翻譯例句，亦即今日所謂基於實例的機器翻譯（*example-based machine translation; EBMT*）。統計式機器翻譯與基於實例的機器翻譯，都借助語料庫大量的語料，馮志偉（2010）指出兩者的區別如下：

- 在基於統計的機器翻譯方法中，知識的表示是統計數據，而不是語料庫本身；翻

譯知識的獲取是在翻譯之前完成，翻譯的過程中不再使用語料庫。

- 在基於實例的機器翻譯方法中，雙語語料庫本身就是翻譯知識的一種表現形式（不一定是唯一的），翻譯知識的獲取在翻譯之前沒有全部完成，在翻譯的過程中還要查詢並利用語料庫。（馮志偉，2010）

統計式機器翻譯除了需要規模夠大的實例（雙語句對）資料庫，做到自動對齊，另外還要建立有效的實例匹配檢索機制和相似度計算準則，以找出有效的相似句或詞串。計算相似度包括從字詞、詞串、句子、段落、文本到文本集等不同層級的比對，Islam 與 Inkpen（2008）將相似度演算法分為四類：基於字詞共現／空間向量的文件模型、基於語料庫統計、混合式方法，以及基於描述特徵的方法。

兩個詞（亦即字串）是否相似，通常採用最短編輯距離（minimum edit distance）演算法計算字串距離（string distance），亦即以插入、刪除、取代等編輯動作更改一個字串，所需要的最少編輯次數。另一個常見方法是計算最長共同子序列（longest common subsequene），亦即尋找兩個字串中相同的最長子序列。計算句子或段落的相似度，常用的公式是 Dice 係數 = $2c / (a + b)$ ，其中 a、b 各為兩個句子的字串長度，c 為兩句共有的字串長度。另外也可能考量詞與詞之間的互現訊息（mutual information; MI）。文本的相似度則常用向量空間模型來計算，亦即以文件中包含的詞彙所組成的向量空間來表示文件集，以用於比較（曾守正等，2006）。

基於向量空間模型的方法，是將句子、段落看成詞的線性序列，語句相似度衡量機制只能利用句中的詞頻、詞性等表層資訊。學者因此提出其他方法，例如進行語句與語義的深層結構分析，基於語義的依存關係以計算相似度（例如 Islam & Inkpen, 2008; Maynard & Ananiadou, 1999；李彬、劉挺、秦兵、李生，2003）。林語君、高照明（2009, pp. 186-187）結合統計與語言訊息，提出混合式雙語句對應演算法，句子的相似度計算參數包括雙語詞典翻譯詞及停用詞表，句中重要標點符號序列，以共同數字詞、時間詞、原文詞為對應錨。

一般使用者最為熟悉的機器翻譯工具，應屬 2005 年 Google 推出的線上自動翻譯服務，2009 年又推出譯者工具包，結合了自訂 TM 與詞彙表的功能。關於 Google 翻譯及譯者工具包的翻譯品質，學者與業界的研究發現確實有其效益（Garcia, 2010; Ramos, 2010），產出的譯文雖不能直接做為發佈之用，但經過人工編輯可達交件品質，並縮短整體翻譯時間，節省成本。

三、合約文體與合約翻譯

不論英文或中文合約，獨立的文體研究極其有限，絕大部分是包含在法律文體之下。有關於法律文本的分類，Sarcevic 依據文本的功能，將法律文本區分為「規範性」與「描述性」兩大類，以及兼具規範與描述的混合類，而法律、法規、條例、公約、合約即屬於「規範」功能為主的法律文書，法律乃是社會整體的行為準則或規範，合約則

是針對簽約當事人的行為規範（1997, pp. 10-11）。宋雷從文體學的角度，將法律文本區分為「規範性法律語言」與「法學著述」，其中規範性法律語言同樣包含立法、司法文本，以及公民在日常事務和商事活動中所訂立的各種合約（2010, p. 14）。

至於法律語言之翻譯，學界與實務界也多所探討，例如 Morris（1995）、Sarcevic（1997）、Varo 與 Hughes（2002）、陸文慧（2002）、孫萬彪（2002）、王道庚（2006）、Cao（2007）、盧敏（2008）、宋雷（2010）等，著重於比較雙語甚至多語的法律文本之間有何語言特徵、異同，以及翻譯的策略技巧與困難所在。王道庚的《法律翻譯—理論與實踐》雖以法律翻譯為題，但所列舉的實務範例，絕大多數皆為合約條款，孫萬彪的《法律翻譯教程》更完全以合約條款為例。從這些論述中可以看到，「立法司法語言」與「合約語言」具有共通的特點，學界與專業人士一般皆將之歸類為同一文體。

關於合約的語篇結構，劉承愚《如何閱讀英文合約》一書指出：

一份完整的英文合約通常可以分為標題、序文、主文條款及結尾辭四大部分。「標題」開宗明義地顯示合約的性質；「序文」是用最簡單的說明，大略介紹合約訂立的背景；「主文條款」裡包括依各種合約性質的不同而約定的特殊條款，以及不論何種類型合約都會出現的一般條款；最後「結尾辭」則是當事人簽名前的一段文字，為整份合約畫下句點。（2007, p. 30）

標題、序文、主文條款及結尾辭，依據 Biber 有關語域、文類、風格的討論（2009, pp. 53-55），可以視為合約的「文類標記」，少見於其他文體。尤其是，合約撰寫並不講求「創意」，正如 Triebel（2009）所指出：

Common law lawyers seldom draft a contract from scratch, but avail themselves of form books, standard forms, precedents and model contracts. The common law world is drowning in such precedents.

普通法律師草擬合約絕少完全從零開始，而是參照各種合約格式書籍、標準範本、前例及合約範例。在普通法的世界中，這樣的前例多不勝數。（2009, p. 152，筆者自譯）

中文合約的結構也大同小異，具有相同的文類標記。更由於政府積極推廣定型化契約，因此在一般條款（或稱標準條款）部分，直接援用範本的情況更加普遍。在網際網路上搜尋中文合約文件，常可見到許多條款一字不差完全相同。這些特徵最適合使用 TM 工具，可以有效減少翻譯工作量，提高效率與譯文品質。

在德國翻譯功能學派的文本分類中（參見 Nord, 1997）²，合約屬於內容導向的訊息類型文本，Snell-Hornby（1988）認為：

the more “specialized” or “pragmatic” the source text, the more closely it is bound to a single, specific situation, and the easier it is to define the function of its translation; the more specific the situation and the more clearly defined the function, the more target-oriented the translation is likely to be;

原文文本愈是「專門」或「實用」，就愈靠近單一特定的語境，也愈容易界定譯文的功能；語境愈是特定、功能愈是明確，譯文也就愈趨向於譯語導向。（1988, p. 114，筆者自譯）

Nord（1997）的翻譯目的論與 Snell-Hornby 主張相仿，同樣認為訊息類文本在翻譯時要求內容正確完整，風格則應符合譯語的語言及文化成規。朱定初（2001, 2004）引述法國比較法學者 Constantinesco 所謂法律翻譯是一種法律轉換和語言轉換同時進行的雙重工作，提出了「雙重功能對等原則」，一是語言傳達功能之對等，二是法律功能之對等，「法律譯者的首要任務是運用讀者可接受的語言盡可能精確重現原文所表達的意思，使譯文在法律上具有與原文接近對等的效果。」（2004, p. 62）此處的讀者，朱定初特別指明並非一般大眾，而是法律專業人士，因此對於「通順」的要求標準不同於一般讀者。Sarcevic（1997, pp. 57-61）也區分法律譯文的直接接受者與間接受者，認為合約當事人只是合約文本的間接受者，真正撰寫、詮釋、使用合約的其實是律師及訴訟時的司法人員。宋雷則提出法律翻譯必須做到「法人法語」，或如荷蘭比較法學家 de Groot 所謂「確保將一種法律語言翻譯成另一種法律語言，而不是將一種法律語言翻譯成另一種語言。」（引自宋雷 2010, p. 23）

在翻譯實務中，從事合約翻譯的譯者多數並非律師，「合約語言」也不是一般人能夠下筆成章的日常語言。在這個方面，電腦輔助翻譯工具可以發揮很好的效益，例如 Kubler 與 Aston（2010）即認為，查詢資料可以借助字典、百科全書、網路搜尋甚至請教專家，但語料庫的優勢在於：

[T]hey can provide data which is not pre-digested but comes in the shape of samples of actual text.... This is particularly true for specialised translations, where the translator may need to get acquainted with not only the terms and concepts of the domain, but also the rhetorical conventions of the genre.

提供未經事先取捨的資料，而且呈現真實的文本……。這一點對專門翻譯尤其重要，因為譯者不但要熟悉領域術語、概念，往往也需要掌握文類的修辭風格。（2010, p.503，筆者自譯）

2 Reiss 依據德國心理學家 Bühler 的語言符號功能論，從語言功能將文本分為三種類型：內容導向（*inhaltsbetont*）、形式導向（*formbetont*）、訴求導向（*appellbetont*），Nord 在 *Translating as a Purposeful Activity* 一書中對於這三個文本類型的介紹則是使用 *informative*, *expressive*, *operative* 三個英文詞彙。

可惜的是歐美地區雖有英文合約語料庫，但目前尚未見到可以公開取得中文合約語料庫或中英對譯合約 TM 或語料庫，因此本研究的主要目的有三：

1. 建置適當的中文與英文合約可比語料庫。
2. 分析比較中文與英文合約的詞彙特徵。
3. 利用目前已經成熟且容易取得的語料庫工具和機器翻譯技術，自動擷取中英文合約可比語料庫當中互為翻譯的 N 連詞。

參、研究方法

本研究著重於中文與英文合約的詞彙特徵，比較兩者之間的異同與對應關係，進而從合約可比語料庫當中自動提取互為翻譯的 N 連詞。提取方法是利用機器翻譯和翻譯記憶系統的自動比對功能，這些技術已有現成的開源或商用軟體可以直接應用，並不需要另外撰寫程式。

研究架構如圖 1 所示，分為語料庫建置分析和自動提取對譯 N 連詞兩個部分。第一部分採取語料庫語言學的途徑，建置中文與英文合約專門可比語料庫，其次利用語料庫檢索工具，探討中英合約的詞彙特徵，並與參照語料庫相比較，提取合約文類的主題

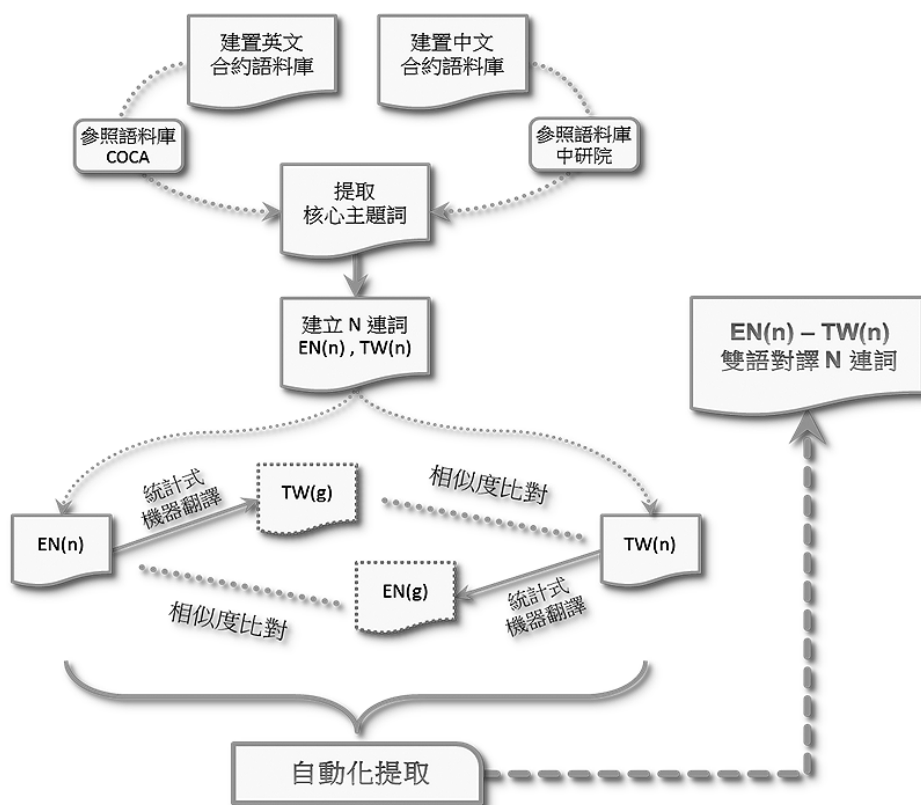


圖 1 研究架構

詞和關鍵主題詞，據此建立核心 N 連詞。第二部分利用機器翻譯結合字串相似度比對技術，自動提取雙語對譯的 N 連詞組，最後評估對譯詞組的準確率。研究流程分為七個主要步驟：

- 首先以網際網路為語料來源，規劃相同的合約主題及取樣條件，以建置中文及英文合約可比語料庫。
- 使用語料庫檢索工具進行語料庫量化分析，自動提取詞表、主題詞表及關鍵主題詞表。
- 考量詞彙頻數、文本頻數、主題詞顯著值、參照語料庫高頻詞等條件，手動篩選詞表，以決定核心主題詞。
- 根據核心主題詞，使用語料庫檢索工具自動建立中文及英文合約 N 連詞。
- 中、英文兩組 N 連詞，利用 Google 統計式機器翻譯產生平行語料。
- 借用 SDLX 翻譯記憶系統的翻譯單元相似度自動比對功能，比對兩組平行 N 連詞的字串相似度，從中提取互為翻譯的內容。
- 由專家評估所得到的 N 連詞對譯詞組是否正確可用。

一、建置合約可比語料庫

建置語料庫所需的工具，包括語料蒐集、語料整理、語料處理三個部分。考量語料庫的代表性、平衡、規模、同質性等要求（Sinclair, 2005a），本研究的中英合約文本採樣條件總結如下：

表 2 中英合約語料庫取樣條件

	中文	英文
取樣來源	網際網路	網際網路
取樣語言	臺灣正體中文合約	美國及英國之英文合約
取樣時間	1990 年迄今	1990 年迄今
主題類別	平衡分佈，共 11 類	平衡分佈，共 11 類
取樣規模	每篇 900 字以上，總計 90 萬字	每篇 600 字以上，總計 60 萬字

1. 文本來源：由於合約的特殊功能，涉及簽約各方重大權益及機密資料，因此語料取得比其他文類更加困難。本研究採用目前日益普遍，以網路為語料來源（Web for Corpus）的做法，利用搜尋引擎檢索合約文體明顯的「文類標記」，找出合約文本後再篩選下載適當的檔案，檔案格式主要包括 DOC、PDF、HTML。需要注意的是，網路上公開的合約多為依法須申報揭露的類型，容易造成特定合約類型採樣不平衡，在選用文本時需要手動分析篩選。
2. 語言：英文限於美國及英國合約，中文限於臺灣合約。

3. 時間：設定為 1990 年以後之合約文件。

本研究除了學術上的分析探討之外，也希望研究結果能夠實際應用於合約翻譯與教學，發揮參考效益，因此語言及時間這兩項條件，著眼於臺灣合約筆譯市場最常見的文本，並以近廿年的當代合約為主。

4. 合約主題類別：「合約」看似一個範圍明確的文類，但可能涵蓋的主題形形色色、篇幅長短不一，語言特徵也非完全相同。本研究主要基於「方便取樣」原則，在網路上容易搜尋到中、英文本且篇幅相當的類別優先收錄。不過網路上數量最多的本文，也反映了合約最常見的類別，依主題區分包括資訊科技、工程、金融、醫學，依功能區分包括租賃、授權、採購、聘雇、保密、合作案、商務服務，總計十一類。只不過，這些類別並無法嚴格劃分，同一份合約可能涵蓋多項功能，歸類時只能依據最主要的合約標的。另外，簽約當事人涵蓋政府、企業、機構團體、個人等。

5. 文本篇幅及語料庫規模：語料庫的規模並沒有固定標準，取決於研究目的，以及資源、人力、時間限制等因素。例如 SUSANNE 句法樹語料庫僅 13 萬字，Biber (1990) 亦指出，自大型語料庫中擷取 120 篇文本做為子庫，即足以反映大語料庫的語言特徵（引述自 Semino & Short, 2004, p. 25）。本研究的合約語料庫設定為 150 萬字，中文語料在蒐集之初尚未分詞，只有文書處理軟體所統計的「字數」；根據筆者的翻譯審稿工作經驗，中英兩種語言相同內容對譯時，字數比例平均約為 3：2，所以中文合約預計收集 90 萬字，英文約 60 萬詞。每一子類的中英文本儘可能字數、篇數相當，每篇文本為中文 900 字、英文 600 字以上。如果中英文篇數無法相當，則維持子類的總字數相當。至於文本是否全文收錄，學界並沒有定論，鑑於本研究著重在合約的標準條款，而非合約全文的語篇結構，因此省略太過詳盡的序文、特定條款及結尾部分，並未全文收錄。

二、語料庫量化分析

本研究採用檢索軟體 WordSmith Tools 6 (Scott, 2012) 及 AntConc (Anthony, 2011) 進行基本的檢索統計，包括詞表及描述統計、主題詞表及關鍵主題詞、N 連詞及搭配詞。統計主題詞所採用的參照語料庫，英文為 Corpus of Contemporary American English (Davies, 2008-，簡稱 COCA) 所提供的 70 萬字詞表，中文則使用中研院《現代漢語平衡語料庫》4.0 版 (陳克健、黃居仁，2007)。

N 連詞通常分析兩詞至四詞結構，但在合約文類中，一些標準條款常有完全相同的敘述，或只是略做變動。本研究經過初步測試後，決定選取 N-2 至 N-8 的詞串。至於 N 連詞所需的最低頻數，學者尚無定論 (參見 Ari, 2006)。

與 N 連詞相關的另一個詞彙特徵是「搭配詞」(collocations)，亦即與檢索詞最常共同出現的搭配詞彙。軟體工具對於搭配詞的檢索，通常設定為左側或右側第一個詞至第三個詞，搭配詞之間不一定連續，這一點與 N 連詞不同。檢索軟體還可計算兩詞

之間的「搭配強度」，數值愈高表示兩個詞愈傾向於在一定的語境中共同出現，常用的統計方法包括互見訊息（mutual information）、T 值（T-score）、Z 值（Z-score）、對數似然率（LL）等。

三、機器翻譯及字串相似度比對

本研究不採複雜的計算語言學技術，而借用現有翻譯記憶系統最基本的翻譯單元比對功能，「參照 TM」則來自可比語料庫機器翻譯的版本。這樣的方式可以快速產生大量的平行語料，克服人工翻譯平行語料不足的困難。研究中使用的機器翻譯工具為 Google 譯者工具包免費線上自動翻譯服務，平行語料的自動比對則是利用 SDLX 2007 翻譯記憶系統。下圖說明以英文做為相似度比對語言的具體操作步驟。

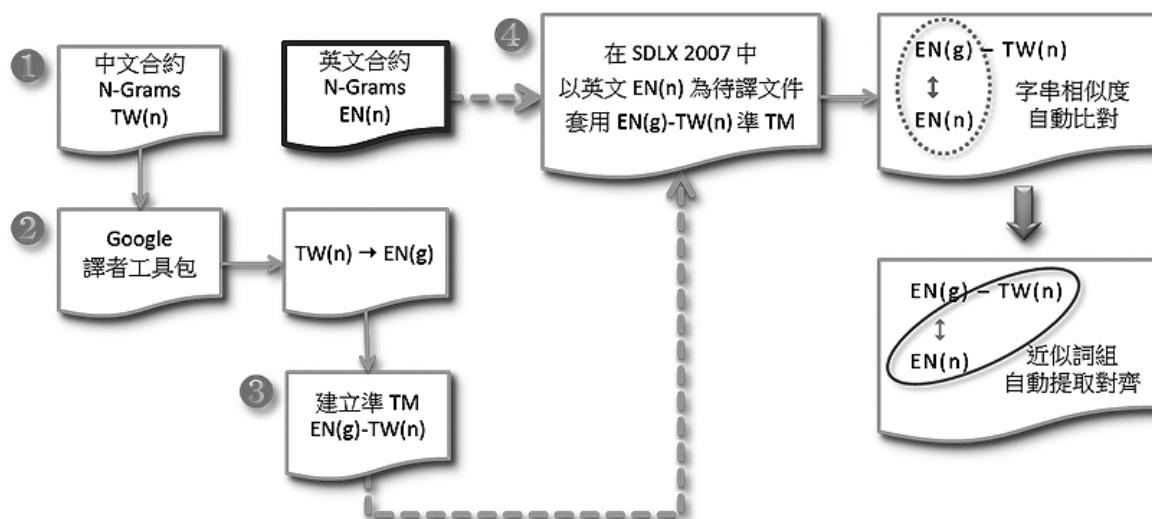


圖 2 合約雙語對譯 N 連詞提取步驟

1. 以語料庫檢索工具所建立的中文合約 N 連詞，稱為 TW (n)，英文合約 N 連詞稱為 EN (n)，分別儲存為純文字檔案。
2. 以英文 N 連詞 EN (n) 做為待譯的新文件，找出其對應的中文，可以視為「英譯中」。此時是以英文做為比對語言，因此須將 TW (n) 透過 Google 機器翻譯產生其英文平行語料，下載儲存為純文字檔案 EN (g)。Google 譯者工具包的使用方法，可參見其網頁說明，翻譯結果可儲存下載。
3. 將機器翻譯的 EN (g) 與中文 N 連詞 TW (n) 貼入到 Excel 工作表左右兩欄，另存為 tab 分隔的純文字檔案，用以匯入翻譯記憶系統，建立「機器英譯 + 中文原文」的 TM 檔案 EN (g) _TW (n)。
4. 在 SDLX 翻譯記憶系統中，以英文 N 連詞 EN (n) 做為待譯的新文件，經過軟體工具轉換格式產生 .itd 檔案。其次，開啟 TM 套用功能，以 EN (g) _TW

(n) 做為參照 TM 套用到 EN (n) .itd。軟體工具會自動將待譯的新單元與參照 TM 相比對，從中搜尋相符或高度相似的內容並擷取 TM 中的對應譯文。TM 套用功能可以自訂相似度下限，達到設定值以上才會顯示比對結果。最後將比對相符的結果匯出，即成為「英文—中文」互為翻譯的 N 連詞對譯表。

至於從中文 N 連詞尋找其對譯的英文，可以視為「中譯英」。此時可將中文做為比對語言，亦即以中文 N 連詞做為待譯的新文件，與英文 N 連詞的機器中譯相比對；另外也可以反向操作，將中文 N 連詞以機器翻譯後，與英文 N 連詞相比對。初步測試的結果，發現第二個方式能夠比對出較多相符的對譯 N 連詞，因此在「中譯英」部分同樣以英文做為比對中介語言。N 連詞英譯中、中譯英皆完成後，將配對的對譯詞組匯出至 Excel 工作表，再由專家手動評估 N 連詞對譯詞組的品質。

肆、結果與討論

根據預定的採樣條件，尤其合約類別儘可能多元平衡，以反映美國及臺灣當代合約語言的普遍特徵。最後收錄的類別、篇數及字數，如表 3、圖 3、圖 4 所示。

表 3 中英文合約收錄類別、總字數及每篇平均字數

合約類別	英文			中文		
	篇數	字數	每篇平均	篇數	字數	每篇平均
資訊科技	18	69539	3863	25	91212	3648
工程	14	74344	5310	12	111295	9275
合作	12	42019	3502	13	60491	4653
服務	11	35773	3252	19	58784	3094
金融	24	127165	5299	24	157453	6561
保密	13	22231	1710	16	29230	1827
租賃	24	58695	2446	29	79433	2739
授權	22	101562	4616	40	142792	3570
採購	14	82564	5897	20	121554	6078
聘雇	9	23089	2565	15	27931	1862
醫學	6	29566	4928	16	42792	2675
總計	167	666547	3991	229	918684	4012

英文合約共計收錄 167 篇 66 萬多詞，中文合約 229 篇近 92 萬字，英文每篇平均長度大於中文，不同子類的篇幅也有極大差異，工程及融資類單篇甚至可達數萬字，而保密及聘雇合約較短，約在兩千字以內。如前所述，本研究著重在合約的標準條款，並未全部全文收錄。

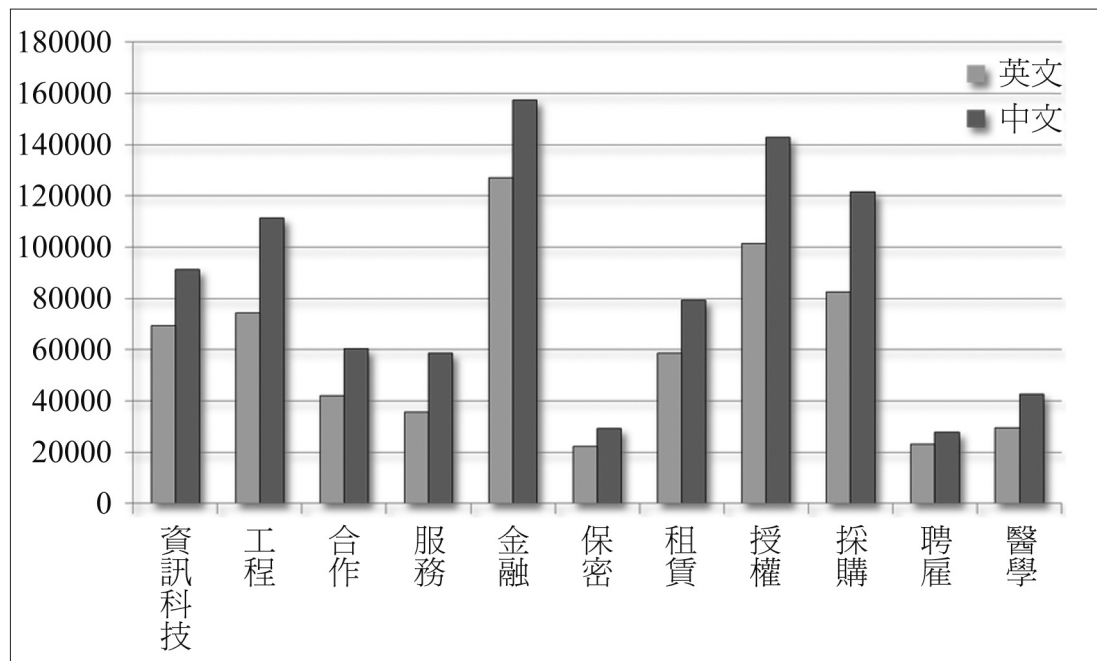


圖 3 中英文合約收錄類別及總字數

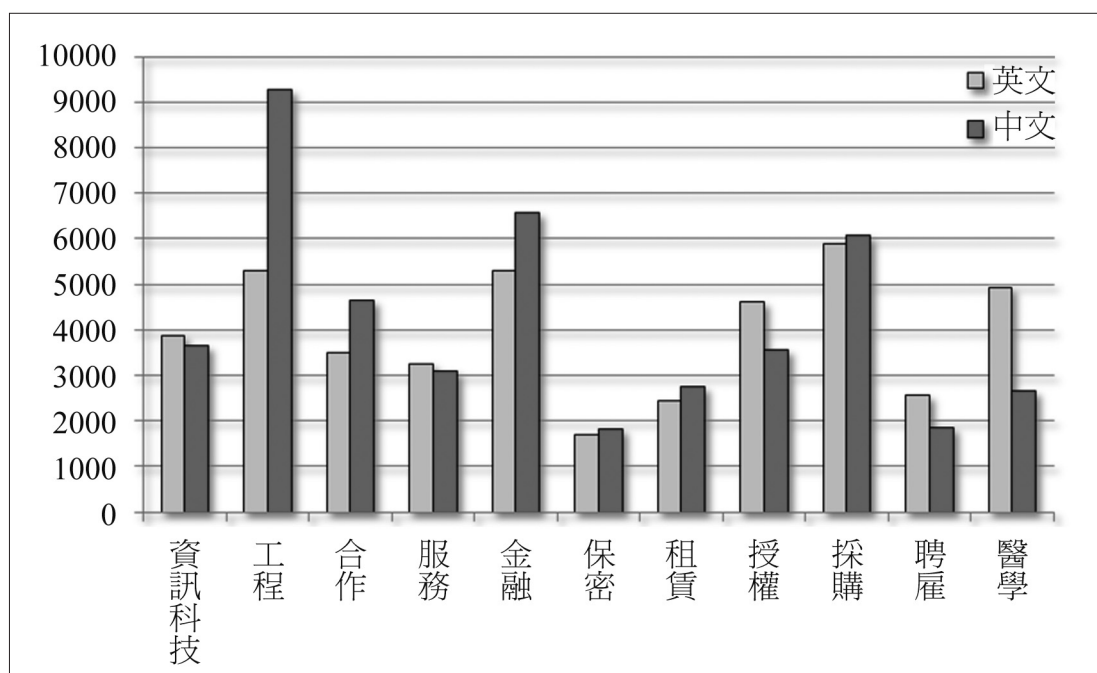


圖 4 中英合約單篇平均字數

一、語料庫描述統計

英文合約語料共 167 篇，經過 WordSmith 分析後，描述統計結果如表 4 所示，納入詞表統計的形符總計 660,218 詞，單篇最長超過 15,000 詞，最短為 843 詞，每篇平均 3953 詞。在平均句長方面，根據 Flesch 提出的可讀性計算公式（參見 DuBay, 2004, p. 23），平均句長 29 字以上即屬於七級難度標準中最困難的一級，而合約語料庫中各篇的句長最大值達到每句平均 60 字以上，67 萬詞的句長總平均為每句 35.58 字。這個數值事實上還屬於低估，因為合約大量使用條款標題及逐條列舉項目，兩三字獨立成行也視為一句；因此，排除標題及條列項目之後所得的句數重新計算平均句長，即達到 42.53 字，更加凸顯合約文本多用長句的語言特徵。

表 4 英文合約語料描述統計

	詞表統計 所用形符	類符	標準化 相異詞比例	平均句長	平均字長
最大值	15029	1620	38.58	67.67	5.56
最小值	843	283	26.30	17.00	4.52
全距	14186	1337	12.28	50.67	1.03
平均數	3953	776	31.93	35.47	5.06
標準差	2459	296	2.27	10.28	0.18
篇數	167	167	167	167	167

反映文本詞彙變化程度的標準化相異詞比例（STTR），最大值為 38.58%，最小值僅有 26.30%，整體平均為 31.93%。學者研究不同文類的 STTR 變化很大，例如 FLOB 英國英語書面語語料庫的平均 STTR 為 43.46%，其中的科學文本子類平均 STTR 則高達 68.99%；FROWN 美國英語書面語語料庫平均 STTR 為 43.71%，科學文本子類為 68.67%（參見 Holtz, 2011, p. 86; Holtz & Teich, 2009; Teich, 2009）。另外，Kang 與 Yu（2011）分析旅遊英文的平均 STTR 為 44.96%，與 FLOB 相近；Pollach（2006）比較產品使用者意見評論、《經濟學人》雜誌、線上隱私權政策三種文類，平均 STTR 分別為 39.90%、49.74%、33.41%。相對而言，英文合約的平均 STTR 較低，顯示詞彙變化較小，原因在於大量使用制式條款的重複句式及表達，然而其平均句長卻遠高於多數文類，閱讀難度最高。

在中文合約部分，語料文本共 229 篇，描述統計結果如表 5 所示。中文合約字數總計 92 萬字，經過分詞處理後納入詞表統計的形符數目約為 50 萬詞。單篇最長篇幅為 13,158 詞，最短為 400 詞，每篇平均 2175 詞，顯示整體而言中文合約的篇幅不如英文合約那麼長（英文平均 3966 詞）。這個數據也反映出英文與中文合約的法系差異，英美合約依據普通法系，注重判例，因此契約文件必須「鉅細靡遺」，詳述一切可能的條件狀況，好用同義詞、近義詞、對義詞，以求「萬無一失」。相較而言，臺灣中文契約依據的是民法法系，法有明文，契約效力以既有的法律條文為準，因此在合約文件中無

表 5 中文合約語料描述統計

	詞表統計 所用形符	類符	標準化 相異詞比例	平均 小句長度	平均 單詞長度
最大值	13158	1661	50.00	15.26	1.80
最小值	400	200	25.90	6.87	1.59
全距	12758	1461	24.10	8.39	0.21
平均數	2175	546.95	36.41	9.69	1.68
標準差	1866	247.28	4.18	1.50	0.04
篇數	229	229	229	229	229

需逐一重述。

至於平均句長方面，由於中文的「句」沒有嚴格固定的定義，也沒有明確的語法界限或標記，因此中文的句長概念與計算，無法與英文直接對應比較。正如第 3.1-1 節所指出，中研院《領域詞典工具》對於中文句子的處理是以「小句」為單位，從表 6 可以看到，合約文本的小句平均長度，介於 6.87 詞至 15.26 詞之間，整體平均為 9.69 詞，若乘以平均單詞長度（1.68 字），即表示小句的平均長度為 16.28 字。相較於《現代漢語平衡語料庫》，平衡語料庫的平均小句長度為 6.86 詞，平均單詞長度為 1.67 字，相當於小句平均字數為 11.46 字，可見中文合約句長顯著高於一般文本。

中文合約的 STTR 標準化相異詞比例介於 25.9% 至 50.0% 之間，229 篇文本總平均為 36.41%，各文本的變異程度大於英文合約；漢語平衡語料庫的 STTR 總平均則為 50.16%。由此可見中文合約的句長及詞彙變化傾向，與英文合約相同，皆為詞彙重複較多，而句子偏長。

二、主題詞、關鍵主題詞與 N 連詞

英文合約出現頻數最高的 9 個實詞，依序為 agreement, party, company, parties, information, use, section, time, date，這些也是合約文本的主題詞及關鍵主題詞，涉及契約文件本身的指稱（this Agreement）、簽約當事人（party, parties, company），條款章節指稱（section），以及合約通常涉及的内容，例如簽約日期時間、合約資訊、合約標的之使用等（date, time, information, use）。

至於最常見的 41 個功能詞，其中有 27 字與 COCA 參照語料庫高頻 50 詞重疊，特別值得注意的是其餘 14 個超出 COCA 高頻 50 詞範圍的功能詞，包括情態助動詞 shall, may, will，表示範圍或關係的 any, each, no, such, other, 's, its, which，以及引介條件或列舉項目的 under, including, provided。其中 shall, will 主要用法並非表示未來時態，而是權利義務，provided 也非動詞「提供」的過去式，而是做為連接詞表示條件。同樣的，這 14 個功能詞也是英文合約文類的主題詞及關鍵主題詞，大量重複出現在各種合約文本。

WordSmith Tools 的主題詞統計，得到顯著多用的主題詞 2243 詞，顯著少用有 1381 詞，另有關鍵主題詞 1421 詞。匯入 Excel 工作表依主題詞顯著值（keyness）、詞頻篇數、關鍵主題詞篇數、是否屬於參照語料庫高頻詞 3000 詞等條件，進一步篩選出核心主題詞 438 詞，並據以建立 N 連詞。

中文合約的最高頻實詞，與英文合約類似，同樣包括括契約文件本身的指稱（契約、合約），簽約當事人（廠商、機關），與合約事項有關的動詞（規定、同意、使用、通知、終止、履約、辦理），狀態不及物動詞「相關」，用於定義或界定句型的分類動詞「為」，以及與合約事項有關的名詞（服務、資料、費用、書面）。而在中文合約最常用的 32 個功能詞當中，多達 20 詞超出參照語料庫高頻 50 詞的範圍，其中乙方、甲方、不得更超出了參照語料庫高頻 1000 詞。這些功能詞同樣也是中文合約文類的主題詞兼關鍵主題詞，用於引介合約最常用的句型。

以中研院現代漢語平衡語料庫做為參照，統計所得的顯著多用主題詞有 2227 詞，顯著少用主題詞 1024 詞，關鍵主題詞 1135 詞。同樣採用英文合約的篩選程序，進一步得到核心主題詞 375 詞，同時據以建立核心 N 連詞。

三、N 連詞機器翻譯及提取對譯詞組

N 連詞與機器譯文經 TM 工具轉檔及比對匹配之後，將相符比例 75% 以上的結果匯出，兩個語言方向各有一千多條，由筆者及另一位資深合約譯者進行專家評估。兩位評分人皆任職於翻譯公司，翻譯審校的合約文件已有數百萬字，熟悉合約文體及領域術語。評分量表採四級分，最高 3 分，表示對應正確可用，2 分表示需要小幅修改，若需要大幅修改（訊息對應有增減或歧義）為 1 分，對應錯誤不可使用為 0 分。

圖 5 為各級分的評分實例，詞數從 N-2 到 N-8，相符比例從 75% 到 100%，都有 0 分到 3 分，分數與相符比例或詞數並沒有絕對關聯。評分完成後以 Pearson 積差相關檢定評分人間信度，結果如表 6 所示。兩位評分人的英譯中平均分數皆低於中譯英，且評分人 A 給分較為嚴格。Pearson 相關係數分別是 .79 與 .80，達到高度相關。

英文合約 N 連詞 EN (n) 與機器英譯 EN (g)，共有 458 條完全相符（100% match），總計 1399 字，75%~84% 相符有 946 條 5000 多字，65%~74% 相符 3833 條 16000 多字。評分結果中，95%~100% 相符的 N 連詞平均分數為 2.5，相當於準確率 83%，全部 1446 條平均分數為 2.0，相當於準確率 67%（見表 7、圖 6）。

合約 N 連詞中譯英部分，比對中介語言仍然採用英文，匹配結果 95%~100% 相符有 727 條 3368 字，85%~94% 所佔比例最低僅 100 多字，75%~84% 有 510 條 3500 多字。評分結果如表 8、圖 7 所示，95% 以上相符的 N 連詞，平均分數 2.5，相當於準確率 83%，與英譯中的結果相同；全部 1250 條平均 2.2 分。

以上的專家評估結果，排除英譯中、中譯英兩邊重複的內容後，得到評分 3 分、可直接使用的雙語對譯 N 連詞合計 600 多條，另有 300 多條評分為 2.5 分，小幅修改後可以使用。這些 N 連詞包含了術語（terms）、固定片語（phrases），以及大於片語的定

N-	TF	DF	EN(n)	match	TW(n)	average
2	49	36	without notice	100%	不 經 通 知	3.0
3	19	7	force majeure event	100%	不可抗力 情事	3.0
4	76	45	contained in this agreement	100%	本 合 約 所 載	3.0
8	4	2	unless in writing and signed by the parties	85%	並 作 成 書 面 且 經 雙 方 簽 名	3.0
7	3	2	damages caused by fire or other perils	75%	火 災 等 事 由 所 致 之 損 害	3.0
7	6	3	respective rights and obligations of the parties	75%	甲 乙 雙 方 之 權 利 義 務	3.0
3	18	8	unless otherwise specified	100%	除 法 律 另 有 規 定	2.0
5	14	4	hereby represents and warrants that	75%	機 構 茲 聲 明 並 保 證	2.0
7	4	2	unless the parties agree in writing to	75%	雙 方 同 意 以 書 面 修 改	2.0
6	3	2	shall be borne by the contractor	75%	概 由 廠 商 自 行 負 擔	2.0
5	4	3	willful misconduct or gross negligence	75%	故 意 或 重 大 過 失 所 致	2.0
4	7	7	shall have the following	100%	不 得 有 下 列	1.0
5	4	3	written consent of the landlord	75%	以 書 面 同 意 延 長	1.0
6	14	6	survive the expiration or termination of	75%	租 賃 期 滿 或 終 止	1.0
5	9	7	deemed to have been given	75%	視 為 已 合 法	1.0
6	3	3	for the benefit of the parties	75%	為 該 合 約 之 利 益	1.0
2	28	16	or default	100%	或 過 失 所 致	0.0
4	6	5	a of this agreement	75%	本 合 約 一 式 兩 份	0.0
5	4	2	this contract shall be governed	75%	本 契 約 所 生 爭 議	0.0
5	4	3	exhibit a of this agreement	75%	甲 方 於 本 契 約	0.0
5	6	3	this agreement shall be submitted	75%	本 合 約 自	0.0

圖 5 N 連詞 0~3 級分的評分實例

表 6 N 連詞專家評估評分人間信度

	英譯中		中譯英	
	評分人 A	評分人 B	評分人 A	評分人 B
平均數	1.820	2.15	2.0	2.4
變異數	0.978	0.702	0.871	0.541
觀察值個數	1446	1446	1249	1249
皮耳森相關係數	.79		.80	

表 7 N 連詞英譯中評分結果—依相符比例

平均分數	100% 相符	95% 相符	85% 相符	75% 相符	小計
3.0 分	223	3	7	147	380 (26%)
2.5 分	104	2	3	140	249 (17%)
2.0 分	64	1	8	202	275 (19%)
1.5 分	35	1	1	217	254 (16%)
1.0 分	21	1		87	109 (16%)
0.5 分	5		1	138	144 (8%)
0.0 分	6			30	36 (10%)
小計	458 (32%)	8 (1%)	20 (1%)	961 (66%)	1447
平均	2.5	2.3	2.3	1.7	2.0

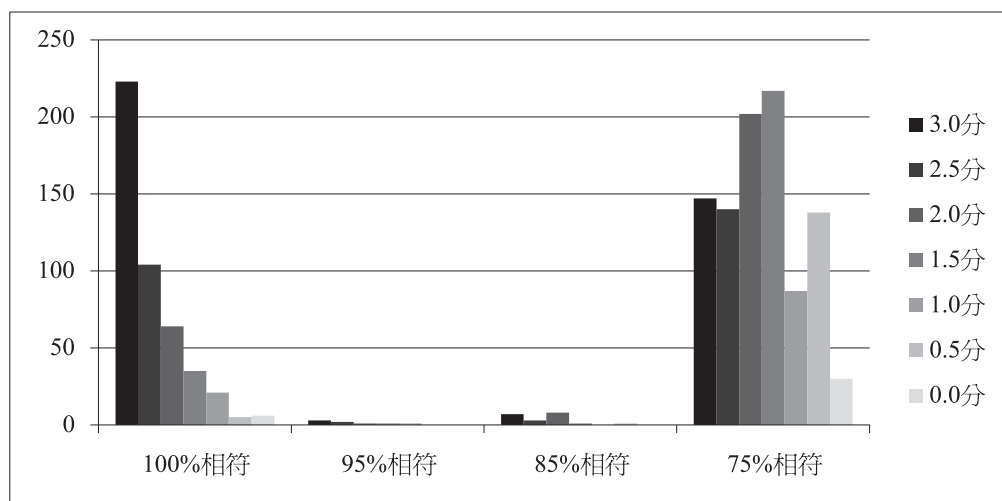


圖 6 N 連詞英譯中評分結果—依相符比例

表 8 N 連詞中譯英評分結果—依相符比例

平均分數	100% 相符	95% 相符	85% 相符	75% 相符	小計
3.0 分	359	5	5	82	451 (36%)
2.5 分	176	1	3	76	256 (20%)
2.0 分	73	1	3	127	204 (16%)
1.5 分	58	2	1	141	202 (16%)
1.0 分	17			27	44 (04%)
0.5 分	26		1	48	75 (02%)
0.0 分	9			9	18 (01%)
小計	718 (57%)	9 (1%)	13 (1%)	510 (41%)	1250
平均	2.5	2.5	2.3	1.9	2.2

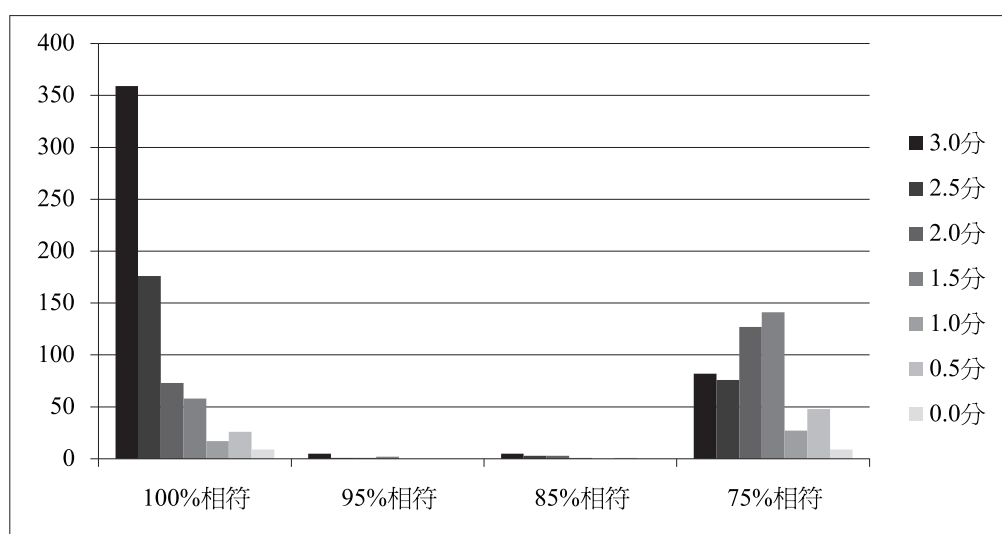


圖 7 N 連詞中譯英評分結果—依相符比例

型化句段，對於合約翻譯實務、教學上都有具體的用途，也可做為其他語言應用研究的資料。評分 3 分的中英對譯 N 連詞組，彙整於附錄。

四、完全相符或高度相符（95%~100% match）

「英譯中」95%~100% 相符的對譯 N 連詞有 400 多條，平均分數為 2.5，亦即準確率在 80% 以上，「中譯英」更有 700 多條，匹配準確率同樣達 80% 以上。這個結果顯示，Google 統計式機器翻譯背後使用的語料庫，已存在這些 N 連詞的中英雙語對譯單元，利用 Google 譯文做為檢索中介，確實能夠從可比合約語料庫中提取出這些互為翻譯的詞串。

「中譯英」高度相符的 N 連詞乍看之下比「英譯中」多出近一倍，但是仔細分析則會發現多所重複，許多 N 連詞其實訊息內容相同，只是措詞略有變化。若依對應的英文 N 連詞重新排序去除重複部分，則剩下 400 多條，顯示中文合約相同條款的用詞細微變動大於英文合約。這個現象也可以說明，為什麼以中文為比對的中介語言時，所得到的相符結果較少，原因就在於一字兩字的差別（例如「權利及義務」、「權利與義務」）即足以干擾短詞串的相似度比對結果。

圖 8、9 是評分後依相符比例、分數及中文 N 連詞排序的結果，不論英譯中或中譯英，正確對應的都屬於合約標準條款，包括權利義務關係、履約相關規定、保密、通知、合約條款效力等。圖中第二欄的詞彙頻數（TF）和第三欄的文本頻數（DF），也可看出這些 N 連詞屬於高頻且出現於多篇文本的常用詞串。

N-	TF	DF	EN(n)	match	TW(n)	average
3	24	15	disclose confidential information	100%	揭露 機密 資料	3.0
2	226	58	intellectual property	100%	智慧 財產	3.0
3	66	39	the expiration of	100%	期限 屆滿	3.0
4	12	10	after the expiration of	100%	期限 屆滿 後	3.0
6	68	15	expiration or termination of this agreement	100%	期滿 或 契約 終止	3.0
4	43	20	expiration or termination of	100%	期滿 或 終止	3.0
4	8	7	upon the expiration of	100%	期滿 時	3.0
3	27	22	no obligation to	100%	無 義務	3.0
5	6	5	the event of a dispute	100%	發生 爭議 時	3.0
3	8	6	copyright or other	100%	著作權 或 其他	3.0
3	72	56	deemed to have	100%	視為 已	3.0
2	398	133	limited to	100%	僅 限於	3.0
4	7	7	the risk of loss	100%	損失 之 風險	3.0
3	78	46	loss or damage	100%	滅失 或 損害	3.0
4	72	36	by the other party	100%	經 他方	3.0
5	5	4	by the other party to	100%	經 他方 以	3.0
3	46	36	upon termination of	100%	經 終止	3.0
6	5	5	in writing signed by the parties	100%	經 雙方 簽署 之 書面	3.0
3	10	5	terminate this contract	100%	解除 本 契約	3.0
2	60	22	the account	100%	該 帳戶	3.0

圖 8 N 連詞英譯中 100% 相符評分結果

N-	TF	DF	TW(n)	MATCH	EN(n)	Average
2	104	59	應遵守	100%	shall comply with	3.0
3	11	10	應立即以	100%	shall immediately	3.0
2	27	21	不得於	100%	shall not	3.0
2	61	48	不得為	100%	shall not be	3.0
2	9	9	不得視為	100%	shall not be considered	3.0
2	12	11	不得洩露	100%	shall not disclose	3.0
2	75	44	不得超過	100%	shall not exceed	3.0
2	12	10	不得逾	100%	shall not exceed	3.0
2	11	10	不得利用	100%	shall not use	3.0
2	12	10	所簽訂	100%	signed by	3.0
5	18	6	經甲乙雙方簽署	100%	signed by the parties	3.0
3	26	12	經雙方簽署	100%	signed by the parties	3.0
2	71	47	自行負責	100%	solely responsible for	3.0
2	13	4	該等契約	100%	such contract	3.0
2	13	5	租賃期限	100%	term of the lease	3.0
3	80	33	終止本合約	100%	terminate this agreement	3.0
3	229	73	終止本契約	100%	terminate this agreement	3.0
4	12	6	隨時終止本契約	100%	terminate this agreement at any time	3.0
4	6	3	立即終止本合約	100%	terminate this agreement immediately	3.0
3	8	6	解除本合約	100%	terminate this contract	3.0
3	51	24	解除本契約	100%	terminate this contract	3.0

圖 9 N 連詞中譯英 100% 相符評分結果

表 9 「部分或全部」檢索行

甲方得沒收履約保證金	一部或全部	，
且足以影響契約	一部或全部	之履行者，
得為	一部終止或全部	終止。
	一部無效及全部	無效
(一) 本工程保固期限自	全部完工或部分	完工，
十四、承租人對於租賃物之	全部或一部分	不繼續使用時，
雙方均同意不轉讓本契約之	全部或一部分	之權益。
指將原契約中應自行履行之	全部或主要部分	，
指將契約中應自行履行之	全部或其主要部分	，
指示受託人就信託財產投資標的之	全部或部分	，
並得追回	全部或部分	已撥付之補助款，
於乙方資產移轉完成後解除其	該部或全部	之履約保證責任。
四、廠商不得將契約或債權之	部分或全部	轉讓予他人。

這樣的對譯 N 連詞雖非全是嚴格定義的術語，但若加入 TM 系統的術語庫，對譯者而言仍有極大的助益。例如 “company reserves the right” 對應「公司有權」，可以提示譯者 “reserves the right” 未必都要譯為「保留權利」，減少原文的干擾。再如英文合約極其常用的片語 “any and/or all”，對應於中文合約，慣用的表達為「部分或全部」及其各種變體，如表 9 所示；「部分」、「全部」的順序可以互換，但英文片語的順序

則固定不可更動。在語料庫中並未發現以「任何」搭配「全部」、「所有」的實例，顯見依中文合約的使用習慣，“any and/or all”不宜翻譯為「任何」。

五、部分相符（75%~94% match）

85%、75% 部分相符的 N 連詞也有許多屬於正確對應，評分 3 分的英譯中 N 連詞有 154 條，中譯英有 87 條。訊息對應而相符比例未達 100%，一個原因是 N 連詞的斷詞未必都正確，含有雜訊，加上比對單元字數太少，只要一字之差或甚至標點、大小寫不同，相符比例很容易就會降低到 75% 以下。

以終止條款為例，“any termination or expiration of this agreement”比對的 Google 英譯“termination or expiration of this agreement”，一字之差相符比例落入 75% fuzzy match，但對應的中文「於本契約終止或期滿」，其實可以直接套用。至於“subsequent to the termination of this agreement”比對 Google 英譯“due to the termination of this agreement”，同樣是一字之差，訊息內容卻是錯誤的對應；這個錯誤同時也源自 Google 英譯已經不正確，因果關係正好顛倒，而機器之所以誤譯，又和中文的表述有關，「所致之本契約終止」是非常文言的句式，在 Google 的統計語料庫中應該找不到太多實例。至於“survive expiration or termination of this agreement”，75% 相符對應為「期滿或契約終止」，但缺漏了最重要的訊息「仍然有效」，顯示“survive”在合約中的語義異於一般用法。這類對應錯誤的例子，可說是從反面印證合約詞彙的特殊用法及近乎「術語性」的特徵。

六、對應不當或對應錯誤

(一) 詞串太短，語境不明

這是預期中的錯誤。N-2 若不是明確的術語，或包含了功能詞，通常訊息就不完整（如圖 10），機器翻譯難以適當處理，比對也就隨之出錯。但若排除 N-2 不列入比對，又可能遺漏確實可用的對譯詞組，「名詞＋名詞」是術語常見的結構。從評分結果也可看到，N-2 約有 300 條評分為 2.5 分或 3 分，排除重複內容後仍有近百條是正確可用的對譯 2 連詞。這個情況是資料檢索中固有的矛盾，準確率（precision）及檢出率（recall）難以兼顧，只能視情況取捨。

(二) 被動及否定句型

包含被動或否定訊息的 N 連詞，對 Google 翻譯會造成困難，進而影響相似度比對結果，如圖 11 所示。在專家評估過程中發現，含有「不」字的中文 N 連詞，Google 英譯時容易缺漏否定的訊息。

N	TF	DF	EN(n)	mat	TW(n)	average
2	51	34	party for	100%	一方 對於	1.0
2	120	65	copy of	100%	文件 影本	1.0
2	590	140	the date	100%	日 起 生效	1.0
2	40	24	any damages	100%	任何 損害 賠償 責任	1.0
2	74	43	ownership of	100%	成果 歸屬	1.0
2	27	20	a dispute	100%	而 生 爭議 後	1.0
2	105	63	hereunder shall	100%	依 本 契約 應	1.0
2	22	16	insurance shall	100%	保險 金額 應	1.0
2	34	14	contract is	100%	契約 一 式	1.0
2	22	14	contract to	100%	契約 以	1.0
2	54	42	any liability	100%	不 負 任何 賠償 責任	0.5
2	25	21	agreed by	100%	有 約定 者 外	0.5
2	25	15	contract for	100%	契約 如 需	0.5

圖 10 英文 N-2 訊息不完整的例子

N	TF	DF	TW(n)	MATCH	EN(n)	Average
5	14	6	不 負 損害 賠償 責任	100%	liable for any damages	0.5
3	8	8	不得 有 下列	100%	shall have the following	0.5
4	7	4	不得 因此 而 要求	100%	shall not be required	0.5
2	53	42	不得 要求	100%	shall not be required	0.5
5	18	9	不可 撤銷 擔保 信用狀 繳納	75%	irrevocable letter of credit	1.5
4	5	4	甲方 不得 將 其	75%	party shall not be	1.5
3	6	6	不得 用以 拘束	75%	shall not be used for any	1.5
6	4	2	事 前 書面 之 同意 不得	75%	prior written consent shall be	1.0
4	6	3	任 一方 不得 為	75%	either party shall be	0.5
6	4	2	不 履行 本 合約 約定 時	75%	the performance of this contract	0.5
3	8	4	本 契約 不得	75%	this contract shall be	0.5
4	6	3	不得 洩漏 予 任何	75%	shall not apply to any	0.0
3	11	11	不得 轉讓 予 任何	75%	shall not be subject to any	0.0

圖 11 否定句型對機器翻譯造成困難

(三) 特有的表達用法

1. 限制條款：非經……不得

「非經……不得」對 Google 翻譯造成困難，「非經」兩字完全省略或直譯為“non-”（見表 10）。亦即這種句式並非一般文本的常見用法，在 Google 語料庫中實例太少。同義詞「未經……不得」，Google 翻譯就處理得較好（見表 11），在 EN(g) - EN(n) 比對中準確率也較高，表示 Google 比較「認識」這個詞。然而在中文合約語料庫中，「未經……不得」出現 90 次，遠低於「非經……不得」131 次，兩者對應的英文表達皆為“shall not … without”，句型主要用於限制條款，如保密義務（機密資訊

不得揭露）、所有權（合約標的不得使用的情況）、合約權利義務（不得轉讓）等。

非經 未經	甲方 乙方 他方 甲乙双方 機關	事前 書面	同意 許可 授權 通知	……	不得
shall not ... may not ... will not ...	without	prior written express (our) (Company's)	consent permission approval authorization notice	of the other party of Lessor of Landlord	

表 10 「非經」句型 Google 英譯錯誤

TW (n)	EN (g)
非經甲方之同意	non-agreed to by Party A of
非經所有者正式開放或授權	non-authorized by the owner of the formal open or
非經甲、乙雙方合意	non-desired by the parties
非經甲方事前書面之同意	non-Party A prior written consent
非經雙方當事人以書面協議	non-written agreement by the parties
非經機關及廠商雙方合意	Party A and manufacturers of the parties desired
非經甲方書面同意不得	Party A written consent shall not
非經甲方事先書面同意	Party A's prior written consent
非經甲方及乙方雙方合意	when by Party A and Party B the parties desired

表 11 「未經」句型 Google 英譯處理較好

TW (n)	EN (g)
未經乙方同意	without the consent of Party B
未經甲方授權	without Party A authorized
未經書面告知	has not been informed in writing
未經主管機關核准	without the competent authority for approval
未經他方書面同意	without the other party written consent of
未經他方書面同意前	without the other party prior written consent
未經甲方之書面同意	without the consent of Party A written
未經法定代理人書面同意	without the legal representative written consent of
未經同意收集他人電子郵件位址	e-mail addresses of others without consent to the collection

2. 終止及存續條款

終止條款的 N 連詞比對，中文與英文皆對應良好，與「終止」一詞搭配強度最高、詞頻 50 次以上的實詞是「解除」，最常見的 N 連詞為「解除及／或終止」、「終

止或解除」。英文合約中與 terminate/termination 搭配強度最高、詞頻 50 次以上的實詞包括：

survive

expiration, employment, employee, cause, notice, event, term, rights, right, determination

immediately, provided, effective, written

agreement, lease, article, section, contract

date, days, time, party, company

其中最值得注意的是 survive 一字，常見 N 連詞為 “survive (the) termination of”，對應的中文表達為「仍具效力」、「仍然有效」，屬於合約的存續條款。存續條款 N 連詞的中英相似度比對極不理想，反映出這個概念在兩種語言的表述方式差異很大。中文慣用表達有正面及反面兩種說法（表 12），正面表達「仍具效力」、「仍然有效」可以對應英文 “remain in (full) effect/force”（語料庫中出現 81 次）。反面表達「不因……而失效」則出現困難，無法對應到 “survive”（語料庫中有 152 次，遠高於 “remain in force”），一方面是 survive 的用法異於一般文件，造成 Google 翻譯不當，其次是中文「不因……而」句型偏長，取 N-2 至 N-8 時詞串被切斷而造成訊息不完整。

表 12 中文合約存續條款常見句型

下列條款於契約終止後仍具效力
於本協議書終止或解除後二年內仍具有效力
於本契約終止或期滿後仍然繼續有效力
第 24 條於本契約終止或期滿後仍然繼續有效力與拘束力
本項之約定於本合約解除或終止後仍繼續有效
下列條款於契約終止後仍繼續有效
於本合約終止後仍然存續生效
前揭保密義務不因委外契約終止而失其效力
不因契約屆滿終止或解除而失其效力
本條款所約定事宜不因本合約之終止或屆滿而失效
保密條款不因本約無效、解除、終止或期限屆滿而失效
條款中之權義不因本契約終止、解除或有效期間屆滿而消滅
保密義務不因本約定條款終止或因其他原因之失效而免除
乙方接到信函與否不影響終止效力之生效

3. 賠償條款

賠償條款 N 連詞的比對結果未盡理想，原因是條款涉及的術語較多，句型也偏長，加上中文與英文合約對於賠償的規定，有明顯不同的表述方式。以 “damages” 及「賠償」兩詞檢索語料庫中的強搭配詞，頻數 10 次以上的前 15 個實詞，中英兩邊重疊的詞彙只有三分之一：

DAMAGES			賠償		
Freq.	MI	Collocate	Freq.	MI	Collocate
94	10.1327	consequential	11	9.0725	請求權人
28	10.0933	punitive	18	8.2912	賠償額
20	10.0553	exemplary	678	7.9973	損害
40	10.0229	liquidated	30	7.9053	請求權
65	9.7395	incidental	17	7.7937	傷亡
13	9.5009	finest	20	7.7651	理賠
17	9.4729	possibility	48	7.5116	責
26	9.4014	profits	379	7.4209	負
19	9.0832	suffered	28	7.3330	受有
66	9.0181	losses	489	7.2125	責任
40	8.9912	indirect	154	7.1899	損失
20	8.9602	penalties	18	7.1106	相等
18	8.9033	judgments	203	7.0234	請求
12	8.8709	injuries	19	6.9542	到達
53	8.7211	liabilities	26	6.8217	遭受

再對照 N 連詞比對結果，不論中譯英或英譯中，除了極少數的部分相符之外，完全沒有交集。英文合約的賠償條款大抵可分為四個部分，慣用的表述方式如下：

Party B shall/agrees to	Indemnify, defend, and hold harmless from/against	Party A[list 1]	
		any and all	losses[list 2]
		including(without limitation)	expenses[list 3]
		arising out of resulting from relating to	this Agreement. [list 4]

整個長句中可以包含多個「清單」，羅列一切可能的情況，例如所稱甲方 [list 1] 可包含 subsidiaries, affiliates, officers, agents, employees, partners, licensors；一切損失 [list 2] 可包含 liabilities, damages, claims, fines, penalties, injuries；各種費用 [list 3] 可包含 costs, expenses, reasonable attorneys fees, court costs and other amounts；與合約有關行為

[list 4] 可詳述一切損失成因。就連主要動詞、介詞也採用同義詞或近義詞並列的寫法。因此典型的英文賠償條款可以長達一兩百字，相對而言，中文合約的賠償條款則言簡意賅，通常不超過 30 字，慣用表述方式如下：

(乙方)	違反	致	甲方	(受有)		
(如)因	違約	所致	他方	損失	(時)	(應)負賠償責任。
	不履行	造成		損害		
	故意					
	過失					

其中表示關係的连接詞及副詞「如」、「因」、「時」、「應」等又可省略，充分顯現中文以意合而非形合的語言特徵。由於兩種語言對於賠償責任的措詞差異太大，因此在 N 連詞比對中完全無法對應。

4. 合約份數

合約份數是中文合約篇尾的固定條款，常用的表達如下：

本	合約	計正本	一	式	OO 份	副本	OO 式 OO 份	
	契約		壹					
	同意書		乙					
由	甲乙	各執	正本		OO 份	為憑		
	雙方					存照		

這個句型在 N 連詞比對中完全找不到英文對應，「OO 份」的 Google 翻譯為 duplicates, copies 甚至 triplicates，但檢索英文合約語料庫可知這些用法並不存在。然而實際上英文合約也有份數條款，其固定表述方式如下。

This Agreement	is	signed	one or more	
	may be	executed	in several	counterparts,
			any unnumber of	
all of which	shall	together	constitute	an original.
each of such counterparts	will		be deemed	

同樣的，這個句型在 N 連詞比對時找不到中文對應，Google 機器翻譯也多有錯誤。這是一個有趣的現象，「一式兩份」和 “in counterparts” 雖是相近的概念，但都對 Google 翻譯造成了困難，兩個語言之間也就搭不上線了。從這個比對「缺項」再次凸顯這兩個詞彙是中英文合約的術語，應該收錄於合約術語庫。

(5) 準據法及管轄法院

中文和英文的準據法條款 N 連詞的相符比例偏低，往往落在 75% 以下，但對應的詞彙及句型其實很清楚，「本契約以……法律為準據法」，對應 “This Agreement shall

be governed (and construed) by the laws of...”。「管轄法院」與“a court of competent jurisdiction”的情形也相同，概念及表達完全對等，但 N 連詞比對結果卻不理想，可能原因有三：首先，準據／governed、法院／court、無效／invalid 這幾組詞彙構成的 N 連詞都偏長，N-5 甚至 N-8 以上而被切斷，且中英文詞序明顯有別；第二，有明顯的本土用詞（依中華民國法律規定無效時），兩個語言不可能對應；第三，表述習慣不同，中文說「依法認定無效」（不提法院）、「經法院認定無效」（不提管轄權），因而減低了與英文“held by a/any court of competent jurisdiction to be invalid”的相符比例。

伍、結論

本研究建置了 150 萬字的中英合約可比語料庫，並藉由檢索工具分析合約詞彙的特徵，除了印證既有的單語合約研究成果之外，更著重於翻譯的需求，探討中文與英文合約詞彙互為翻譯的對應關係。研究結果顯示，在合約的標準條款中，例如違約、賠償、爭議解決、準據法、通知、效期、終止等，中英文有許多概念皆為共通，無須另行翻譯，只要能夠從可比語料庫中「找出」對應的表達，即可直接套用。本研究藉由機器翻譯技術和翻譯記憶系統的詞句相似度比對功能，自動提取互為翻譯的中英合約詞彙，以相對簡單的原理及技術，進一步發揮可比語料庫的效益。

本研究所建立的中英合約雙語對譯詞組，具有實用的效益，可直接應用於電腦輔助翻譯工具，匯入為術語庫資料，以供自動查詢、自動顯示、預先套用，或做為手動檢索的關鍵詞，發揮句級以下的對應、查詢及套用效益。學界與 TM 系統業者目前積極研發句級以下的擷取及對應技術，要進一步提升翻譯記憶系統的回收使用率，N 連詞對是很重要的資源。

合約可比語料庫一方面可以利用語料庫工具直接檢索使用，另方面也可視為大量的真實平行文本。平行文本的效益不言可喻，學界已有許多探討，例如 Snell-Hornby (1988)、Zanettin (1998)、Bowker (2000)、Sharoff (2006)、Nord (2007) 等人，都在教學和研究中應用平行文本。中國學者近年來也日漸重視這個議題，林克難 (2006, 2007) 針對實用文件與法律文件的翻譯，提出了「看易寫」原則，特別強調翻譯的第一步要先大量閱讀譯語的真實材料，在翻譯時仿照同類文本的風格和特徵，使譯文達到同等效果。

然而，譯者在日常工作中或教師學生在翻譯課堂上，不太可能隨時查看數十篇平行文本，語料庫工具可以加快檢索效率，但仍然需要在檢索工具（例如 AntConc）和文書軟體（例如 Office Word）之間反覆切換。藉由本研究提出的方法，以機器翻譯產生可比語料庫的譯文，建立「機譯—原文」假 TM，就能夠在翻譯記憶系統中更有效使用平行文本和可比語料庫。例如圖 12 是「中文原文 + Google 英譯」的合約語料，匯入翻譯記憶系統中做為「英譯中」的參考 TM，以機器英譯做為檢索的橋樑，最終可以查出相關的中文條款，以模仿其術語、詞彙和表達句型。從圖中可以看到，待譯的新句與 TM

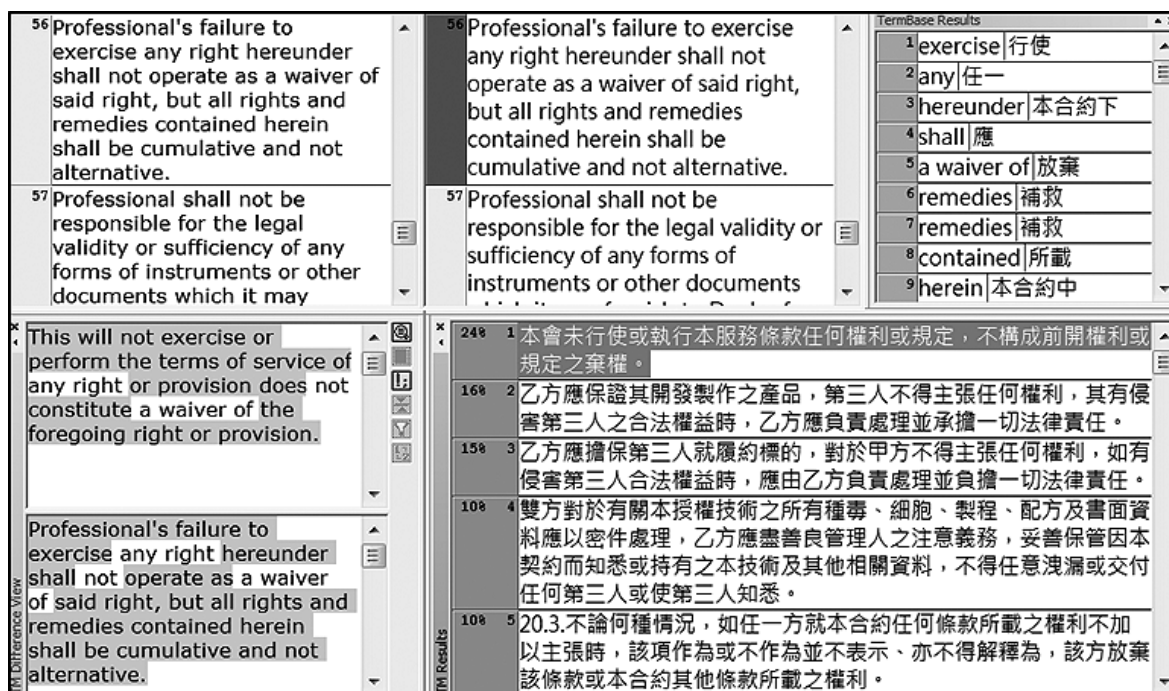


圖 12 原文結合機器翻譯做為參考 TM

的機器英譯相似度只有 24%，但對應的中文條款「本會未行使或執行本服務條款任何權利或規定，不構成前開權利或規定之棄權。」對於譯者卻有很大的參考價值，前半句可以完全依據中文參考條款而改寫。

陸、研究限制

本文研究過程中遭遇的困難包括兩大方面，一是語料來源，二是工具及技術。合約文件由於內容涉及敏感機密資料，外界難以取得，因此本研究依循以網路為資源自建語料庫的趨勢，從網際網路搜尋可用的文本。不過公開在網路上的合約以定型化範本為主，內容較容易雷同甚至重複。其次，自建的合約語料庫若要公開做為共享資源，使用授權問題也不易解決。這一點是所有基於網路的語料庫共同的困難，迄今學界並沒有定論。

在工具及技術方面，蒐集語料階段筆者採用手動檢索下載的方式。雖然利用 spider/crawler 程式將蒐集語料、篩選整理文本的流程自動化，可以節省可觀的人力和時間，但完全自動擷取的語料，需要再搭配其他檢查機制，才能確保所得的語料符合預設的抽樣條件，能夠針對研究問題提供有效的資料。第二個工具困難是中文語料的前置處理，中文合約分詞的錯誤比例高達 10% 以上，修正極其費時費力，合約詞類標注也有困難及錯誤。再其次是現有的語料檢索工具仍然不足，WordSmith Tools 提供的 ConcGram 功能，筆者在中文作業系統中測試時屢次當機，而香港理工大學 Winnie Cheng 開發的

ConcGram 軟體，處理中文語料的效果也不理想。因此本研究未能針對 ConcGram 多做探討。

本研究礙於人力時間的限制，對於專家評估 1.5 分以下的 N 連詞並未深入探討低分的原因及代表的意涵，相符比例 75% 以下的 N 連詞未進行專家評估，相符比例 65% 以下的無對應 N 連詞也未做討論，這些都是研究的不足之處。此外，最後提取的對譯 N 連詞，實際使用效益尚未提供實證數據，還有待日後補足。

柒、未來研究方向

本研究建置的合約可比語料庫，以及使用的雙語資料提取方法和流程，仍有許多值得進一步探討的議題。首先是合約語料並未充分利用，合約類別及詞類標注資料在研究中皆未討論。未來可以針對不同類別的合約做細類分析，例如軟體使用條款、銀行貸款合約、土地開發合約等，各自的詞彙、術語及風格仍然有別。建立 N 連詞時可以結合詞類標注資訊，應可過濾許多無效的雜訊資料。利用詞類標注資訊也能夠進一步探討中、英合約的語法和句型結構。

第二，以可比語料庫為基礎，結合語料庫工具、機器翻譯和翻譯記憶系統，可以建立平行文本翻譯教學模式，設計有效的教材與教法。功能學派目的論的翻譯準則，受到許多學者、教師支持，但是徒有準則尚不能達到特定譯文文本所需的具體功能。資深譯者馮國扶主張（個人通訊，2011 年 9 月 18 日）：

〔法律經貿〕翻譯教學需要文本為核心的教材與教學方法。按文本類型分科，例如課目是「合同」而非商事法，「財務報表」而非會計學，「信用證」而非國貿實務。

以文本為核心的教學內容恰是翻譯所需，學生所學專科知識可達到專業翻譯的要求。不但教材要出自專業文本，翻譯練習的材料也必須出自專業文本。

未來可以研究可比語料庫結合機器翻譯，建立機譯平行語料庫做為翻譯記憶系統的參考 TM。以大量的單語專門文本或雙語平行文本做為翻譯記憶系統的參考 TM，利用其自動查詢顯示功能，譯者可以在瞬間存取資料庫數百篇文本當中最相關有用的例句，立即套用，快速累積領域所需的各種知識及翻譯技巧。

第三，本研究所建立的中英合約對譯 N 連詞，可以往下擷取術語及片語，編纂合約雙語詞典，或是以 N 連詞為種子往上擴展找出更大語境的詞串、句段，甚至擴大提取互為翻譯的對應句。Hartmann（1994, 1996）認為使用可比語料庫編纂雙語詞典，比起傳統的「原文→譯文」翻譯式編纂方法，一方面可以得到更自然道地、符合領域真實語言的詞典，另方面收錄的詞目能夠更完整，避免傳統雙語詞典收詞不足的缺點。利用

可比語料庫所擷取的對譯內容，也可用於開發專門領域文件的機器翻譯系統；或是應用在電腦輔助翻譯工具，提供翻譯「預測」功能，根據合約語境自動提示常用的詞串、句段、句型。此外也可建立合約技術文件的索引，進而應用於其他的語言處理模型或技術，例如從網際網路搜尋特定合約主題的雙語文本，以建立更大規模的專門領域可比語料庫。

第四，在研究方法方面，相同的步驟可以應用於其他專門領域，尤其是平行語料資源嚴重不足的專門文件，例如醫學期刊研究論文、金融技術文件（衍生商品公開說明書、壽險產險保單等），透過兩種語言的可比語料庫與機器翻譯技術，個別譯者或研究人員也能夠更有效擷取可比語料庫當中豐富的語言資訊。

捌、參考文獻

- 王道庚（2006）。《法律翻譯——理論與實踐》。香港：香港城市大學出版社。
- 朱定初（2001）。〈談英語法律專門術語之翻譯〉。《翻譯學研究集刊》6，27-52。
- 朱定初（2004）。〈談法律專門術語翻譯之雙重功能對等原則〉。《國立編譯館館刊》32(1)，60-66。
- 宋雷、張紹全（編）（2010）。《英漢對比法律語言學》。北京：北京大學出版社。
- 李彬、劉挺、秦兵、李生（2003）。〈基於語義依存的漢語句子相似度計算〉。《哈爾濱工業大學資訊檢索研究室論文集》(1)
- 季培培、鄢小燕、岑咏華（2010）。〈面向領域中文文本信息處理的術語識別與抽取研究綜述〉。《圖書情報工作》（北京），54(16)，124-129。
- 林克難（2006）。〈法律文件宜先看後譯〉。《上海翻譯》（上海），4，40-42。
- 林克難（2007）。〈從信達雅、看易寫到模仿一借用一創新一必須重視實用翻譯理論建設〉。《上海翻譯》（上海），3，5-8。
- 孫萬彪（2002）。《法律翻譯教程》。上海：上海外語教育出版社。
- 康小麗、章成志、王惠臨（2009）。〈基於可比語料庫的雙語術語抽取研究述評〉。《現代圖書情報技術》(10)，7-13。
- 梁茂成、李文中、許家金（2010）。《語料庫應用教程》。北京：外語教學與研究出版社。
- 陳克健、黃居仁（2007）。《現代漢語平衡語料庫》（4.0 版）。臺北：中央研究院。
網址：http://www.aclclp.org.tw/use_asbc_c.php。
- 陸文慧（編）（2002）。《法律翻譯：從實踐出發》。香港：中華書局。
- 曾守正等（2006）。《基於文件倉儲之中文文件探勘模式、平行處理架構與其相關推廣應用之研究》。國立高雄第一科技大學資訊管理系。
- 馮志偉（2010）。〈基於語料庫的機器翻譯系統〉。《術語標準化與信息技術》(1)，28-35。

- 劉承愚（2007）。《如何閱讀英文合約》。臺北：益思科技法律事務所。
- 盧敏（編）（2008）。《英語法律文本的語言特點與翻譯》。上海：上海交通大學出版社。
- Anthony, L. (2011). *AntConc* (Version 3.3.5). Tokyo: Waseda University. Available at <http://www.antlab.sci.waseda.ac.jp/>
- Ari, O. (2006). Review of three software programs designed to identify lexical bundles. *Language Learning & Technology*, 10(1), 30-37. Retrieved from llt.msu.edu
- Baker, M. (1993). Corpus Linguistics and Translation Studies: Implications and Applications. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 233-250). Amsterdam/Philadelphia: John Benjamins.
- Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. *Target*, 7(2), 223-243. doi:10.1075/target.7.2.03bak
- Biber, D., & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge: Cambridge University Press.
- Bowker, L. (2000). Towards a methodology for exploiting specialized target language corpora as translation resources. *International Journal of Corpus Linguistics*, 5(1), 17-52. doi:10.1075/ijcl.5.1.03bow
- Cao, D. (2007). *Translating Law*. Clevedon/Buffalo/Toronto: Multilingual Matters.
- Davies, M. (2008-). *The Corpus of Contemporary American English: 425 million words, 1990-present*. Available at <http://corpus.byu.edu/coca/>
- DuBay, W. H. (2004). *The Principles of Readability*. Costa Mesa, California: Impact Information.
- Fitzpatrick, E. (Ed.). (2006). *Corpus Linguistics Beyond the Word: Corpus Research from Phrase to Discourse*. Amsterdam/New York: Rodopi.
- Garcia, I. (2010). Is machine translation ready yet? *Target*, 22(1), 7-21. doi:10.1075/target.22.1.02gar
- Granger, S., & Meunier, F. (Eds.). (2008). *Phraseology: An interdisciplinary perspective*. Amsterdam/Philadelphia: John Benjamins.
- Hartmann, R. R. K. (1994). The use of parallel text corpora in the generation of translation equivalents for bilingual lexicography. In *Proceedings of the Euralex 1994* (291-297), Amsterdam.
- Hartmann, R. R. K. (1996). Contrastive textology and corpus linguistics: On the value of parallel texts. *Language Sciences*, 18(3-4), 947-957.
- Holtz, M. (2011). *Lexico-grammatical properties of abstracts and research articles* (Darmstadt). Available from Technische Universität Darmstadt. Retrieved from <http://tuprints.ulb.tu-darmstadt.de/2638/1/PhD-Thesis-Monica-Holtz.pdf>

- Holtz, M., & Teich, E. (2009). *Design of the Darmstadt Scientific Text Corpus*. Technische Universität Darmstadt. Retrieved from <http://www.linglit.tu-darmstadt.de/fileadmin/linglit/holtz/DaSciTex/dfg-corpus-design.pdf>
- Hunston, S. (2011). *Corpus approaches to evaluation: Phraseology and evaluative language*. New York/London: Routledge.
- Hutchins, J. (2010). Machine translation: a concise history. *Journal of Translation Studies*, 13(1-2), 29-70.
- Islam, A., & Inkpen, D. (2008). Semantic text similarity using corpus-based word similarity and string similarity. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 2(2), Article 10. doi:10.1145/1376815.1376819
- Kang, N., & Yu, Q. (2011). Corpus-based stylistic analysis of tourism English. *Journal of Language Teaching and Research*, 2(1), 129-136. Retrieved from <http://www.academypublisher.com/jltr/vol02/no01/jltr0201.pdf> doi:10.4304
- Kubler, N., & Aston, G. (2010). Using corpora in translation. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 501-515). London/New York: Routledge.
- Laviosa, S. (1995). The design and analysis of a comparable corpus of English newspaper articles. *Outros Temas*, 307-314. Retrieved from ler.letras.up.pt
- Loginova, E., Gojun, A., Blancafort, H., Guegan, M., Gornostay, T., & Heid, U. (2012). Reference lists for the evaluation of term extraction tools. In *Proceedings of the Terminology and Knowledge Engineering Conference (TKE 2012)*, Madrid.
- Maynard, D., & Ananiadou, S. (1999). Term extraction using a similarity-based approach. In D. Bourigault, C. Jacquemin & M.-C. Lhomme (Eds.), *Recent Advances in Computational Terminology* (pp. 261-278). Amsterdam: John Benjamins.
- McEnery, T., & Xiao, R. (2005). Parallel and comparable corpora: What are they up to? In G. James (Ed.), *Corpus Linguistics and Translation Studies*. Clevedon: Multilingual Matters.
- Morris, M. (1995). *Translation and the Law*. Amsterdam/Philadelphia: John Benjamins.
- Nord, C. (2007). Looking for help in the translation process-the role of auxiliary texts in translator training and translation practice. *Chinese Translators Journal*, 1, 17-26.
- Pazienza, M., Pennacchiotti, M., & Zanzotto, F. (2005). Terminology extraction: an analysis of linguistic and statistical approaches. *Knowledge Mining*(185), 255-279. Retrieved from http://ai-nlp.info.uniroma2.it/pennacchiotti/publications/SFSC_2005.pdf
- Pollach, I. (2006). Electronic word of mouth: A genre analysis of product reviews on consumer opinion web sites. In *Proceedings of the 39th Hawaii International Conference on System Sciences*, Hawaii.
- Ramos, L. C. (2010). Post-editing free machine translation: from a language vendor' s

- perspective. In *Proceedings of the Ninth Conference of the Association for Machine Translation in the Americas (AMTA-2010)*, Denver, Colorado.
- Sarcevic, S. (1997). *New Approach to Legal Translation*. The Hague/London/New York: Kluwer Law International.
- Scott, M. (2012). *WordSmith Tools* (Version 6.0). Liverpool: Lexical Analysis Software. Available at http://www.lexically.net/publications/citing_wordsmith.htm
- Semino, E., & Short, M. (2004). *Corpus Stylistics: Speech, writing and thought presentation in a corpus of English writing*. London/New York: Routledge.
- Sharoff, S. (2006). Translation as problem solving: uses of comparable corpora. In *Proceedings of the International Workshop on Language Resources for Translation Work, Research and Training*, Genoa, Italy.
- Sharoff, S., Babych, B., & Hartley, A. (2006). Using collocations from comparable corpora to find translation equivalents. In *Proceedings of the 2006 International Conference on Language Resources and Evaluation* (465-470), Genoa, Italy.
- Sharoff, S., Babych, B., & Hartley, A. (2009). Irrefragable answers: Using comparable corpora to retrieve translation equivalents. *Lang Resources and Evaluation*, 43(1), 15-25. Retrieved from <http://www.springerlink.com/content/8k6631431pl3538l/> doi:10.1007/s10579-007-9046-4
- Sinclair, J. M. (2005a). Corpus and text: Basic principles. In M. Wynne (Ed.), *Developing Linguistic Corpora: A Guide to Good Practice* (pp. 1-16). Oxford: Oxbow Books.
- Sinclair, J. M. (2005b). The phrase, the whole phrase and nothing but the phrase (Phraseology 2005 plenary lecture). In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 407-410). Amsterdam/Philadelphia: John Benjamins.
- Snell-Hornby, M. (1988). *Translation Studies: An Integrated Approach*. Amsterdam/Philadelphia: John Benjamins.
- Stubbs, M. (2002). *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.
- Stubbs, M. (2009). The search for units of meaning: Sinclair on empirical semantics. *Applied Linguistics*, 30(1), 115-137. doi:10.1093/applin/amn052
- Su, F., & Babych, B. (2012). Measuring comparability of documents in non-parallel corpora for efficient extraction of (semi-) parallel translation equivalents. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics* (10-19), Avignon, France.
- Teich, E. (2009). Exploring a corpus of scientific texts using data mining. *Language and Computers*, 71(1), 233-247. Retrieved from <http://www.l3s.de/web/upload/documents/1/teich-fankhauser-final09.pdf>

- Thurmair, G. (2003). Making term extraction tools usable. In *Proceedings of the EAMT-CLAW 03*, Dublin.
- Triebel, V. (2009). Pitfalls of English as a Contract Language. In F. Olsen, A. Lorz & D. Stein (Eds.), *Translation Issues in Language and Law* (pp. 147-181). New York: Palgrave Macmillan.
- Varo, E. A., & Hughes, B. (2002). *Legal translation explained*. Manchester: St Jerome.
- Xiao, R. (2010). How different is translated Chinese from native Chinese? A corpus-based study of translation universals. *International Journal of Corpus Linguistics*, 15(1), 5-35. doi:doi:10.1075/ijcl.15.1.01xia
- Zanettin, F. (1998). Bilingual comparable corpora and the training of translators. *Meta*, 43(4), 616-630. Retrieved from <http://www.erudit.org/revue/meta/>
- Zhang, Z., Iria, J., Brewster, C., & Ciravegna, F. (2008). A comparative evaluation of term recognition algorithms. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC '08)* (2108-2113), Marrakech, Morocco.
- Zweigenbaum, P., Gaussier, E., & Fung, P. (Eds.). (2008). *Proceedings of LREC 2008 Workshop on Comparable Corpora*. Las Palmas, Spain: LREC.

附錄：中英文合約對譯 N 連詞

act on behalf of 代為 處理
additional terms 附加 條款
after the expiration of 期限 屆滿 後
after the expiration of 期滿 後
after the termination of this agreement 本 契
約 終止 後
agree to comply with 同意 遵守
agree to the following 同意 下列
agree to the following 同意 遵守 下列
agreed by the parties 經 甲 乙 雙方 同
意
agreed by the parties 經 甲 乙 雙方 議
定
agreed by the parties 經 雙方 同意
agreed by the parties 經 雙方 議定
agreed in writing 書面 約定
agreed to 同意 以
agreed to 同意 於
agrees not to 同意 不得
all damages 一切 損害
all losses 一切 損失
all losses 所有 損失
any compensation 任何 補償
any damage 任何 損害
any damages 任何 損害 賠償
any dispute 任何 爭議
any dispute 任何 糾紛
any fees 任何 費用
any liability 任何 責任
any liability 任何 賠償 責任
any litigation 任何 訴訟
any loss 任何 損失
any notice 任何 通知
any provision of this agreement 本 合約 任
何 條款
any third party 任何 第三人
any violation of 任何 違反
at any time by written notice 隨時 以 書面
通知
at its own expense 自行 負擔 費用
attorneys' fees 律師 費用

authorized representative 授權 之 代表
authorized representative 授權 代表
be terminated 被 終止
breach or default 違反 或 不 履行
but are not limited to 但 不 限於
by the other party 經 他方
by the parties 由 甲 乙 雙方
by the parties 由 雙方
by the parties 經 甲、乙 雙方
by the parties 經 甲 乙 雙方
by the parties 經 雙方
by the parties in 由 甲 乙 雙方 於
company reserves the right 公司 有 權
completion of the project 工程 完工
computer software 電腦 軟體
confidential information 機密 資料
confidential information 機密 資訊
contained in 所 載
contained in this agreement 本 合約 所
載
contract price 契約 價金
copyright or other 著作權 或 其他
court of competent jurisdiction 管轄 法院
damage caused by 所 造成 之 損害
damage caused by 造成 之 損害
damage to property 財物 損失
days prior written notice 日 前 以 書面
通知
days prior written notice to 天 前 以 書面
通知
deemed to have 視為 已
deems necessary 認為 必要
delivery of any 交付 任何
disclose confidential information 揭露 機密
資料
disclosed to 洩露 予
disclosed to 揭露 予
dispute resolution 爭議 解決
does not apply to 不 適用 於
duration of this agreement 本 契約 之 存
續 期間

effective date of 生效 日期
 either party 任 一方
 either party 任 一方 均
 either party 任 一方 當事人
 either party may 任 一方 均 得
 either party may 任 一方 得
 either party may terminate 任 一方 均 得
 終止
 electronic data 電子 資料
 e-mail 電子 郵件
 except for those 者 不在此限
 exclusive license 專屬 授權
 expenses incurred in 所 生 之 費用
 expiration date 有效 期限
 expiration or termination of 期滿 或 終止
 expiration or termination of this agreement 期
 滿 或 契約 終止
 fails to perform 未 履行
 failure to comply with 未 遵守
 failure to perform 不能 履約
 failure to perform 未 能 履行
 fees and other 費用 及 其他
 financial statements 財務 報表
 for breach of 因 違反
 for the completion of 為 完成
 force majeure 不可抗力 之 因素
 force majeure 不可抗力 之 事由
 force majeure 不可抗力 因素
 force majeure event 不可抗力 情事
 from the date of 日期 起
 fully responsible for 應 負 完全 責任
 governed by the laws 法律 為 準據法
 gross negligence 重大 過失
 gross negligence 重大 過失 行為
 have agreed 有 約定
 have agreed to 約定 以
 have been delivered 已 交付
 hereby agrees to 特此 同意
 hereinafter collectively referred to 以下 合稱
 hereinafter referred to as 下 稱
 hereinafter referred to as 以下 稱
 hereinafter referred to as 以下 簡稱
 if the parties 如 雙方
 if the parties 若 雙方
 if this agreement 若 本 合約
 immediately notify the other party 立即 通知

他方
 in accordance with the following 依 下列
 方式
 in the contract 在 合約
 in this agreement 在 本 合約 中
 in this contract 在 本 合約
 in this contract 在 本 契約 中
 in writing signed by the parties 經 雙方 簽
 署 之 書面
 including but not limited 包含 但 不 限
 於
 including but not limited 包括 但 不 限
 於
 including but not limited to 包括 且 不
 限於
 insurance company 保險 公司
 insurance coverage 保險 範圍
 intellectual property 智慧 財產
 is not liable 均 不 負 任何 責任
 laws and regulations 法令 之 規定
 laws and regulations 法令 規定
 laws and regulations 法令 規章
 liability insurance 責任 保險
 license agreement 授權 合約
 limitation of liability 責任 限制
 limited to 僅 限於
 loss or damage 損失 或 損害
 loss or damage 滅失 或 損害
 material adverse effect 重大 不利 之 影
 響
 material adverse effect on 重大 不利 影響
 material breach 重大 違約
 may request 得 請求
 may suspend 得 暫停
 may terminate this agreement 得 終止 本
 合約
 mutual agreement 共同 協議
 neither party may 任 一方 不得
 neither party shall 任何 一方 均 不得
 no party shall 任何 一方 不得
 non-exclusive 非 專屬
 non-exclusive license 非 專屬 授權
 nor shall 亦 不得
 not in accordance with 未 依照
 not to exceed 以 不 超過
 notice of termination 通知 終止

notice of termination of this agreement 通知
終止 本 合約
notify the other party in writing 以 書面 方
式 通知 他方
notify the other party in writing 以 書面 通
知 另 一方
obligation of confidentiality 保密 義務
obligations of confidentiality 應 負 之 保
密 義務
or agreement 或 約定
or any third party 或 任何 第三人
or approval 或 核准
or arbitration 或 仲裁
or breach of 或 違反
or content 或 內容
or damage 或 損害
or damage 或 損壞
or damage to property 或 財物 損失
or damages 或 損害 賠償
or delay 或 遲延
or delivery 或 交付
or disclosure 或 揭露
or equipment 或 設備
or loss 或 損失
or loss 或 滅失
or modify 或 修改
or obligations 或 義務
or other intellectual property 或 其他 智慧
財產權
or payment 或 給付
or perform 或 執行
or representative 或 代表人
or request 或 請求
or suspend 或 暫停
or termination 或 終止
or this agreement 或 本 合約
or transfer 或 移轉
or unenforceable 或 無法 執行
oral or written 口頭 或 書面
other documents 其他 文件
other expenses 其他 費用
other intellectual property rights 其他 智慧
財產權
other provisions 其他 規定
other third party 其他 第三人
otherwise agreed 另行 議定

otherwise agreed by the parties 由 甲 乙
雙方 另行 協議
otherwise agreed by the parties 由 雙方 另
行 協議
otherwise agreed in writing 另 有 書面
約定
otherwise provided in 另 有 規定
otherwise provided in 另 有 規定 外
part of the contract 本 契約 之 一 部份
part of the contract 契約 之 一 部分
part of the contract 契約 之 部分
part of the contract 部分 之 契約
party may terminate this agreement 一方 得
終止 本 契約
payment of rent 繳 租金
prior notice 事 前 通知
prior notice 事先 通知
prior to the termination 終止 前 已
prior to the termination of this agreement 於
本 契約 終止 前
prior written consent of 事 前 書面 之
同意
prior written consent of 事 前 書面 同意
prior written consent of 事先 書面 同意
representations and warranties 聲明 及 保
證
representations and warranties 聲明 與 保
證
represents and warrants that 聲明 並 保證
reproduce, modify 重製、修改
required to 須 於
required to be 規定 須
rights and obligations 權利 及 責任
rights and obligations 權利 及 義務
rights and obligations 權利 義務
rights and obligations 權利 義務 關係
rights and obligations 權利 與 義務
rights and obligations of the parties 雙方 之
權利 義務 關係
rights or obligations 權利 或 義務
shall apply to 應 適用
shall be conducted 得 辦理
shall be deemed 即 視為
shall be in accordance 應 依照
shall be in accordance with 均 應 依
shall be in writing 均 應 以 書面

shall be in writing 均應以書面為
 shall be in writing 須以書面
 shall be in writing 應以書面
 shall be in writing 應即以書面
 shall be liable for 應負責賠償
 shall be responsible for 應負責
 shall be returned 應退還
 shall be solely responsible 應自行負責
 shall comply with 均應遵守
 shall comply with 應遵守
 shall immediately 應立即以
 shall not 不得於
 shall not be 不得為
 shall not be considered 不得視為
 shall not disclose 不得洩露
 shall not exceed 不得超過
 shall not exceed 不得逾
 shall not use 不得利用
 signed by 所簽訂
 signed by the parties 經甲乙雙方簽署
 signed by the parties 經雙方簽署
 solely responsible for 自行負責
 such contract 該等契約
 term of the lease 租賃期限
 terminate this agreement 終止本合約
 terminate this agreement 終止本契約
 terminate this agreement at any time 隨時終止本契約
 terminate this agreement immediately 立即終止本合約
 terminate this contract 解除本合約
 terminate this contract 解除本契約
 terminated at any time 隨時終止
 termination of the contract 合約之終止
 termination of the contract 合約終止
 termination of the contract 契約之終止
 termination of the contract 終止合約
 termination of this agreement 本合約終止
 termination of this agreement 本契約終止
 termination or expiration of 終止或期滿
 terms of the contract 合約內容
 terms of the contract 合約條款
 terms of this agreement 本合約之條款
 the account 該帳戶
 the agreement 該約定

the damage 其損害
 the date of termination 終止之日
 the effective date of this agreement 本合約生效之日
 the expiration of this agreement 本合約屆滿
 the facilities 各項設施
 the invalid 該無效
 the maximum extent 最大範圍
 the maximum extent permitted 允許之最大範圍
 the maximum extent permitted by 允許之最大範圍內
 the maximum extent permitted by law 法律許可之最大範圍
 the notice 該通知
 the other party 他方當事人
 the other party in 他方於
 the parties 甲、乙雙方
 the parties 甲、乙方
 the parties 甲乙雙方
 the parties 甲乙雙方當事人
 the parties 雙方當事人
 the parties agree 甲、乙雙方同意
 the parties agree 甲乙雙方同意
 the parties agree 當事人同意
 the parties agree 雙方同意
 the parties agree that this agreement 雙方同意本合約
 the parties are 甲乙雙方為
 the parties hereby agree 甲乙雙方特此同意
 the parties hereby agree to 甲乙雙方特此同意以
 the parties pursuant to 雙方各依
 the parties shall 甲乙雙方應即
 the parties shall 甲乙雙方應依
 the parties shall 當事人應
 the parties shall 雙方應即
 the parties shall 雙方應依
 the party 該當事人
 the party 該方當事人
 the performance of this contract 本契約之履行
 the performance of this contract 就本契約之履行

the performance of this contract	履行 本 合 約	upon the expiration of	期滿 時
the performance of this contract	履行 本 契 約	use the software	使用 本 軟體
the provisions of this contract	本 契約 之 規定	violation of this agreement	違反 本 合約
the reasonable control of	合理 控制	will immediately	將 立即
the rights and obligations of the parties	甲 乙 雙方 之 權利 義務	without notice	不 經 通知
the rights and obligations of the parties	雙方 之 權利 義務	written agreement	書面 協議
the risk of loss	損失 之 風險	written consent of	書面 同意
the terms of the contract	本 合約 條款	written notice	書面 通知
the terms of the contract	合約 所 定 之 條款	written notice to	以 書面 通知
the terms of the contract	契約 之 條款	written notice to	書面 方式 通知
the terms of this contract	本 契約 之 條 款	written notice to the other party	以 書面 通 知 他方 當事人
the third party	該 第三人	written notice to the other party	書面 通知 另 一方
this agreement	本 合約	written or oral	書面 或 口頭
this agreement	本 協議	you agree that	您 同意
this agreement	本 約定	not limited to	不 限於
this clause	本 條款	unless otherwise agreed	除 另 有 約定
this contract	本 契約	unless otherwise agreed	除 另 有 約定 外
this contract	本 契約書	unless otherwise specified	除 另 有 規定
this lease	本 租賃	unless otherwise specified	除 另 有 規定 外
this license	本 授權	made in writing and signed by both parties	作 成 書面 且 經 雙方 簽名
to the extent permitted	許可 範圍 內	made in writing and signed by both parties	並 作成 書面 且 經 雙方 簽名
transferred to	移轉 予	the prior written consent of the other party	經 他方 事 前 書面 同意
unable to perform	不能 履行	this agreement may be terminated at any time	本 契約 得 隨時 終止
unable to perform	無法 執行	upon the expiration or termination of this agreement	期滿 或 契約 終止 後
unable to perform	無法 履行	agree to the following	同意 下列 事項
under this agreement	依 本 協議書	agreed to by the parties in writing	經 雙方 書面 同意
under this contract	本 契約 下	all or part of the contract	本 契約 之 全 部 或 一部
understand and agree that	了解 並 同意	any of the terms of this agreement	本 契約 任何 條款
understand and agree that	瞭解 並 同意	any party to this agreement	本 契約 任何 一方
unless otherwise agreed by the parties	除 當 事人 另 有 約定 外	as a result of this agreement	如 因 本 契 約
unless otherwise provided	除 另 有 規定 者 外	confidential or proprietary information	特定 機密 或 專有 資訊
unless the parties	除 當事人		
upon delivery	於 交付		
upon termination of this agreement	本 合約 終止 後		
upon termination of this agreement	於 本 契約 終止 後		

construed as a waiver of 視同 放棄
construed as a waiver of 視為 放棄
damages caused by fire or other perils 火災
等事由所致之損害
dispute arising under this agreement 因本
契約所生爭議
disputes arising under this agreement 因本
契約所生之爭議
during the term of this contract 本契約執
行期間
either party may terminate this agreement 任
一方得終止本契約
has the right to suspend or terminate 有權
暫停或終止
in accordance with the arbitration rules 依仲
裁法
in accordance with the contract 依契約
in accordance with the contract 依照契約
in accordance with the following 依下列
in accordance with the provisions 按規定
in accordance with this agreement 依本合
約
in writing and signed by the parties 書面且
經雙方簽名
in writing by the parties 書面且經雙方
notice of termination of this agreement 通知
終止本契約
notify the other party in writing 經書面告
知他方
obligations hereunder, in whole or in part 義務
之全部或一部
or in any other manner 或以其他任何
or in any other manner 或以其他任何
方式
or in any other manner 或以其他任何
方式使
other applicable laws and regulations 其他有
關法令規定
other applicable laws and regulations 其他法
令規定
other applicable laws and regulations 其他相
關法令之規定
otherwise agreed by the parties 除雙方另
有約定
performance of the contract 履行契約應
prior to the termination of this agreement 本

契約終止前
prior to the termination of this agreement 係
於本契約終止前
rights and obligations under this agreement 在
本契約中之權利及義務
shall be deemed to have been made 即視為
已送達
shall not assign this agreement 不得轉讓
本契約
shall not be deemed to constitute 不得視為
或構成
shall not in any way 不得以任何方式
shall promptly notify the other party 應立即
通知他方
shall, at its own expense 並自行負擔費用
signed by the parties 並經甲乙雙方
簽署
signed by the parties 並經雙方簽署
specified in the contract documents 本契約
文件所載
specified in the contract documents 契約文
件所載
the effective date of this agreement 本契約
生效之日
the effective date of this agreement 本契約
生效日
the expiration or termination of this agree-
ment 本合約屆滿或終止
the expiration or termination of this agree-
ment 本契約期滿或終止
the non-breaching party may 未違約之
一方
the occurrence of the event 事故之發生
the parties agree that this agreement 甲、乙
雙方同意本契約
the parties agree that this agreement 雙方同
意本契約
the parties are unable to 甲乙雙方無法
於
the provisions of this article 本條規定
the provisions of this contract 合約之規定
the termination or expiration of this agree-
ment 本契約終止或期滿
the terms and conditions hereof 下列條款
the terms and conditions set forth 所定之

條款 與 條件
the terms of this contract 本 合約 之 約
定
the terms of this contract 本 合約 所 定
之 條款
the validity of this agreement 本 契約 之
效力
this agreement, in whole or in part 本 契約
之 全部 或 部分
to the other party 向 他方
unless otherwise provided in this agreement 除
本 合約 另 有 約定 外
unless the parties agree otherwise 除 當事人
另 有 約定
upon expiration or termination of 契約 期滿
或 終止 後
upon expiration or termination of 期滿 或
終止 時
upon receipt of such notice 於 接獲 通知
後
upon termination or expiration of 契約 終止
或 期滿 後
upon termination or expiration of this agree-
ment 於 本 契約 終止 或 期滿
後
with or without notice 於 通知 或 未 通
知 之 情形 下
with or without notice 通知 或 未 通知
之 情形 下
without the other party's prior written con-
sent 未 經 他方 書面 同意 前
written consent of the company 經 本 公司
書面 同意
written consent of the other party 經 他方
書面 同意
written notice to the other party 書面 告知
他方
written notice to the other party if 書面 通知
他方 於
you or any third party for 您 或 任何 第
三人 均
you understand and agree that 您 了解 並
同意
not liable 不 負 賠償 責任
is not responsible for 不 負責
not directly or indirectly 不得 直接 或 間

接
as follows 內容 如下
the other party in writing 他方 以 書面
to comply with 以 符合
the extent necessary 必要 範圍 內
without the other party 未 經 他方
the provisions of this agreement 本 合約 規
定
any provision of 任一 條款
by either party 任何 一方 於
the provisions of 各 項 條款
within a reasonable period of 合理 期間 內
the maximum extent permitted by law 在 法
律 許可 之 最 大 範圍 內
as a result of this agreement 如 因 本 合
約
any dispute 有 任何 糾紛
obligation of confidentiality 有 保密 義務
have a material adverse 有 重大 不利
entitled to 有 權 於
prior written 事 前 書面
other laws 其他 法令
other material 其他 重大
or expenses 或 支出 費用
or limit 或 限制
or expiration 或 期滿
or maintenance 或 維護
referred to 所 稱
termination or expiration of this agreement 於
本 契約 終止 或 期滿
a reasonable period of time 於 合理 期間
receipt of 於 接獲
whichever is 者 為準
court of competent jurisdiction 為 管轄 法
院
patent application 專利 申請
after the termination of 終止 後
the extent permitted 許可 範圍
the expiration of 期限 屆滿
no obligation to 無 義務
the event of a dispute 發生 爭議 時
by the other party to 經 他方 以
upon termination of 經 終止
the parties shall be 雙方 均 應
in writing signed 簽署 之 書面
including, but not limited to 包括 且 不

限於
the rights and obligations of the parties here-
to 甲 乙 雙方 之 權利 義務
unless in writing and signed by the parties 並
作成 書面 且 經 雙方 簽名
in writing and signed by the parties 並 作成
書面 且 經 雙方 簽名
by written notice to the other party 一方 得
以 書面 通知 他方
by written notice to the other 一方 得 以
書面 通知 他方
party may terminate this agreement by 一方
得 終止 本 契約
party may terminate this agreement in 一方
得 終止 本 契約
party may terminate this agreement upon 一方
得 終止 本 契約
party may terminate this agreement with 一方
得 終止 本 契約
shall not, without the prior written consent
of 不得 於 未 經 他方 書面 同意
前
shall not directly or indirectly 不得 直接 或
間接
to the maximum extent permitted 允許 之
最 大 範圍
to the maximum extent permitted by 允許 之
最 大 範圍 內
the fullest extent permitted by 允許 之 最
大 範圍 內
maximum extent permitted by 允許 之 最
大 範圍 內
company reserves the right to 公司 有 權
the company reserves the right 公司 有 權
days prior written notice of 天 前 以 書面
通知
days' prior written notice to 天 前 以 書
面 通知
days written notice prior to 天 前 以 書面
通知
days written notice to 天 前 以 書面 通
知
for damages caused by fire or other 火災 等
事由 所致 之 損害
the other party in writing of 他方 以 書面
the other party may 他方 並 得

notify the other party in writing of 以 書面
方式 通知 他方
giving written notice to the other party 以 書
面 通知 他方 當事人
providing written notice to the other party 以
書面 通知 他方 當事人
written notice thereof to the other party 以 書
面 通知 他方 當事人
notify the company in writing 以 書面 通
知 廠商
hereinafter collectively referred to as 以下 合
稱
including but not limited to any 包括 且 不
限於
including, but not limited to, all 包括 且 不
限於
including, but not limited to, those 包括 且
不 限於
the non-breaching party 未 違約 之 一方
terms of this agreement to 本 合約 之 條
款
of the effective date of this agreement 本 合
約 生效 之 日
of the date of this agreement 本 合約 生效
之 日
the effective date of the agreement 本 合約
生效 之 日
effective date of this agreement 本 合約 生
效 之 日
the date of this agreement 本 合約 生效
之 日
any provision of this agreement is 本 合約
任何 條款
that any provision of this agreement 本 合約
任何 條款
any provision of the agreement 本 合約 任
何 條款
expiration of this agreement shall 本 合約
屆滿
expiration of this agreement 本 合約 屆滿
contained in this agreement to 本 合約 所
載
this agreement in whole or in part 本 契約
之 全部 或 部分
the duration of this agreement 本 契約 之
存續 期間

the interpretation of this agreement 本 契約
之 解釋
after termination of this agreement 本 契約
終止 後
otherwise agreed to by the parties 由 甲 乙
雙方 另行 協議
otherwise mutually agreed by the parties 由
甲 乙 雙方 另行 協議
respective rights and obligations of the parties 甲 乙 雙方 之 權利 義務
the rights and duties of the parties 甲 乙 雙
方 之 權利 義務
the parties hereto hereby agree 甲 乙 雙方
特此 同意
the parties hereby agree that 甲 乙 雙方
特此 同意 以
the parties hereby consent to 甲 乙 雙方
特此 同意 以
parties hereby agree to 甲 乙 雙方 特此
同意 以
terminate this agreement immediately upon 立
即 終止 本 合約
to terminate this agreement immediately 立即
終止 本 合約
promptly notify the other party 立即 通知
他方
either party to terminate this agreement 任 一
方 片面 終止 本 契約
any kind, express or implied 任何 明示 或
默示
within a reasonable period of time 合理 期間
內
agree to comply with all 同意 遵守
out of or in connection with this agreement 因
本 契約
the maximum extent permitted by applicable
law 在 法律 許可 之 最 大 範圍
內
maximum extent permitted by applicable
law 在 法律 許可 之 最 大 範圍
內
the fullest extent permitted by law 在 法律
許可 之 最 大 範圍 內
or not in accordance with 如 未 依照
a result of this agreement 如 因 本 合約
any breach of this agreement 如 有 違反

本 契約
to have a material adverse 有 重大 不利
but is not limited to 但 不 限於
are not limited to 但 不 限於
but are not limited 但 不 限於
upon written notice to the other 即 以 書面
通知 他方
shall be deemed to have been given 即 視為
已 送達
shall be given in writing 均 應 以 書面
為
in writing and signed by the parties hereto 並
作成 書面 且 經 雙方 簽名
the prior written consent of 事 前 書面 之
同意
in accordance with applicable laws 依 相關
法律
in accordance with the laws 依 相關 法律
terms and conditions set forth 所 定 之 條
款 與 條件
upon any termination of this agreement 於 本
契約 終止 後
upon the termination of this agreement 於 本
契約 終止 後
be governed by the laws 法律 為 準據法
governed by the laws of 法律 為 準據法
a material adverse effect on 重大 不利 影
響
the term of the lease 租賃 期限
unless otherwise mutually agreed by the parties 除 當事人 另有 約定 外
may terminate this agreement at 得 終止 本
合約
may terminate this agreement if 得 終止 本
合約
may terminate this agreement under 得 終止
本 合約
may terminate this agreement upon 得 終止
本 合約
may terminate this agreement with 得 終止
本 合約
to the extent permitted under 許可 範圍 內
notice thereof to the other 通知 對方
performance of this contract 就 本 契約
之 履行
expiration or termination of this agreement

for 期滿 或 契約 終止
upon expiration or termination of this agree-
ment 期滿 或 契約 終止
expiration or termination of the agreement 期
滿 或 契約 終止
upon the expiration or termination 期滿 或
終止 時
in the event of a dispute 發生 爭議 時
event of a dispute 發生 爭議 時
a court of competent jurisdiction 訴訟 管轄
法院
any court of competent jurisdiction 訴訟 管
轄 法院
the other party's prior written 經 他方 事
前 以 書面
the other party's prior written consent 經 他
方 事 前 以 書面 同意
agreed by the parties hereto 經 甲 乙 雙
方 同意
in writing signed by the parties hereto 經 雙
方 簽署 之 書面
at any time by giving written notice 隨時 以
書面 通知

at any time by written notice to 隨時 以 書
面 通知
any time by giving written notice 隨時 以
書面 通知
be terminated at any time 隨時 終止
terminated at any time by 隨時 終止
terminate this agreement at any time by 隨時
終止 本 契約
terminate this agreement at any time for 隨時
終止 本 契約
to terminate this agreement at any time 隨時
終止 本 契約
or after the termination of this agreement 應
於 本 契約 終止 後
represents and warrants that it 聲明 並 保
證
rights and obligations of the parties hereto 雙方
之 權利 義務 關係
rights and duties of the parties 雙方 之 權
利 義務 關係
parties agree that this agreement 雙方 同意
本 合約