

建構語意與句型擷取之服務系統

羅靖華

國立高雄第一科技大學管理學院

黃文楨

國立高雄第一科技大學資訊管理系

陳欣蓉

國立高雄第一科技大學應用德語系

摘要

在學習語言的過程中，除了學習單字與單辭之外，能夠有完整無錯的句子更是重要。但是目前所有字典辭典系統的例句，依然侷限在編寫者依照查詢的單字或單辭來編寫的。這種作法比較傳統而且通常不能切入使用者學習上的需要。本系統的目的就是要解決這樣的語言學習困難，以目前較少人開發的德文學習這個區塊，使用「以句找句」的方式，輸入中文詞句就會找出結構最相近的德文詞句，對於語言的學習很有幫助。本研究突破傳統詞彙比對的作法，以樹狀結構表示句型結構，提供有序與高擴展性的句型結構框架，建立以句構與辭性為基礎的索引機制。由全句的句子組成結構與其中各單詞的辭性進行快速索引，並以本體論(Ontology)結合主題地圖(Topic Maps)概念建構不同情境的應用領域路徑，克服迥異情境下辭彙取用問題，實作系統後透過問卷調查得知，服務機制獲得使用者認同，反映出本研究提出的創新方式對於語言學習將確實有所助益。

關鍵字：機器翻譯、本體論、主題地圖

Construction of Semantic and Sentence Patterns Retrieval Service System

Ching-Hua Lo
College of Management,
National Kaohsiung First University of Science and Technology

Wen-Chen Huang
Department of Information Management,
National Kaohsiung First University of Science and Technology

Hsing-Lung Chen
Department of Applied German,
National Kaohsiung First University of Science and Technology

Abstract

In a language learning process, it is important to express complete and correct sentences in addition to learning vocabulary and idioms. However, the sample sentences in current dictionaries are based on user's querying specific vocabulary, and do not usually meet user's wider learning requirements. The purpose of the proposed system is to solve such a language learning difficulty by using a "sentence to sentence" mechanism for Chinese to German sentence query and translation. It allows inputting a Chinese sentence and the system then displays several German sentences with similar structure/meaning by descending similarity. The proposed system, based on a sentence parsing tree structure, provides an ordered and highly extensible sentence structure framework to construct a semantic and grammar-based indexing utility. Fast indexing is processed by parsing a sentence structure and extracting the syntactical function of each word. It is supported by Ontology and Topic Maps to construct various application domain paths for vocabulary extraction while in different scenarios. Questionnaire survey was informed that this innovative study for language learning would be helpful.

Key words: Machine Translation, Ontology, Topic Maps

壹、緒論

「全球化」時代的來臨，語言的力量更顯得無遠弗屆，除了英文以外，第二外國語的學習已經蔚為風潮；再加上，目前語言學習軟體仍然侷限在以字翻字的階段，對於需要完整應用的句子卻沒有更多的著墨。傳統學習外語方式大都被動接受教材提供的句型，但是所學的外文句子常常不合乎學習者溝通時所真實需求的句子。目前學習外語表達時學習者常使用電子辭典以及使用機器翻譯軟體：電子詞典雖然提供了豐富的外文詞彙，然而經由各種外文詞彙常常無法組成道地的外文句子。機器翻譯由直譯中文為外文句子，但是並未以句型與句構提供語意與句意的正確的概念與全貌，經常產出錯誤的外文詞句，並無法有效幫助學習(陳欣蓉 2008)。

繼續與連結是人類思想與學習的基本過程(De Bono 1992)，知識管理(Knowledge Management)的本體論、語意網主題地圖等等理論與技術常用於概念與行動有序的結構化與關聯化，可以協助人們進行繼續與連結的思考活動，以及提供知識全貌。另外，雖然單字辭是構成句子的最小單位，但是單句卻是表達意念與進行有效溝通的基礎。

本研究提出一個超脫「詞彙」範疇，建構將句子視為外語學習的最小單位。基於本體論、語意網與主題地圖等知識管理技術的語文學習系統，可以在「任何時間」，「任何地點」找到「所需的資訊」。面對用外語溝通情境時，使用本系統可以在「關鍵時間」，「關鍵地點」尋查達到溝通目的的句子。藉由開發中文對照外語(德語)的溝通句型之學習系統，提供學習者由中文尋找所需的外文表達，並積極促進主動學習外語之機能。經過系統效益評估，本研究發現以句子結構為基礎，結合知識管理技術為基礎的語文學習服務，能夠有效的透過知識關聯將類句有脈絡的擴展呈現，協助學習者進行知識深化與記憶。系統主要特色有「以句構與辭性為基礎的索引機制」、「整合知識管理技術的智慧型學習導引」、「具有延展能力的句型評比媒合機制」等項目，分述如下：

一、以句構與辭性為基礎的索引機制

目前語言學習軟體仍然侷限在以字翻字的階段，對於需要完整應用的句子卻沒有更多的著墨，以字翻字也無法掌握句構全貌。本研究突破傳統詞彙比對的作法，提出以句構與辭性為基礎，由全句的句子組成結構與其中各單詞的辭性進行快速索引的機制。

二、整合知識管理技術的智慧型學習導引

以本體論(Ontology)結合主題地圖(Topic Maps)架構，建構不同情境的應用領域路徑，克服迥異情境下(如購物，問路等等)辭彙取用的精準度問題。本體論物件樹中提供特定領域中物件彼此的關聯與相對位階，很適合用以自動推論物件之間的垂直或水平延展關係。一般本體論的應用方式將操作物件依本體論給予特定屬性定義後儲存，應用時再依此進行搜尋。本系統則是透過建立本體論對映表的方式，在資料儲存時，只記錄使用者對此資料定義與描述之後設資料。待應用需求發生時才透過映射找出操作關鍵資

料父、兄、子代物件與本體論的脈絡關係進行延伸搜尋，並限制查尋有限迭代，避免動態查詢時的效能耗損。本系統提供在資源檢索時透過相關類別本體論定義資料庫延展映對以符合本體論中資源描述關係定義的衍生關鍵字，對於使用者定義的資源描述資料進行綜合即時查詢，進行後設式本體論資料映射的強化型的檢索服務機能。

三、具有延展能力的句型評比媒合機制

傳統學習外語方式大都被動接受教材提供的句型，但是所學到的外文句子常常不合乎學習者溝通時所需的句子。然而經由各種外文詞彙的拼湊往往無法組成道地的外文句子。而藉由機器翻譯由中文尋找外文句子，又經常產出錯誤的外文詞句。因此若只提供固定的句型，並無法適性適境的滿足學習所需。本研究突破傳統詞彙比對的作法，以樹狀結構表示句型結構，提供有序與高擴展性的句型結構框架，以及基於句構與辭性為基礎的索引機制。由全句的句子組成結構與其中各單詞的辭性進行快速索引，能夠針對前述傳統語言學習系統進行語言學習時的問題提供完整的改善。而且系統模組功能獨立，有利於彈性應用與擴充機能，並能有助於因應服務需求動態提供整合服務。

本文分為五個章節，在緒論中我們闡述本研究的動機，目的與貢獻，第二節透過文獻探討研究發展背景，第三節中說明本研究之方法與系統架構與實作，第四節進行系統效益分析討論，第五節中提出結論、研究限制與未來研究方向。

貳、文獻探討

語言本身就是一種知識。具體而微的，我們可以將語言的學習視為以語言為標的之一聯串學習深化與自我知識管理的過程。語文的學習強調與學習者的生活經驗相互結合，以利透過學習者進行語文知識內化。無所不在的學習服務對此提供了良好的機制。近年來廣泛被討論應用的服務導向架構則很適合作為無所不在的學習服務的基礎。以下針對語文處理知識管理與數位網路教學應用服務的相關文獻進行探討，並對目前常見語文學習系統的問題進行分析。

一、中文斷詞與知識管理

語言知識表示的最小單位為具有語意的單辭或單字。例如在中文語言處理的作業過程中，當應用程式接收使用者輸入的中文查詢句子之後，必需進行中文斷詞 (Yeh & Chen 2004)，不過由個別服務開發者自行處理中文句斷詞，不但甚難處理，也欠缺統一的後設資料 (metadata) 辭性映對，更不易保障其處理的正確度。中研院的CKIP (Chinese Knowledge Information Processing) 服務，可將使用者的查詢字串拆解成不同的詞性之組合，以利中文斷詞相關應用的推展。CKIP平衡語料提供178個詞類 (中央研究院資訊科學所中文組實驗室中文詞知識庫小組 2007a)，43個簡化標記，3個特殊標記，並提供三類詞類標記服務(邱智銘 2003)，CKIP 辭性之間能很便利的使用如樹結構等資料結構表示，極

易於操作應用(中央研究院資訊科學所中文組實驗室中文詞知識庫小組 2007b)。相關應用例如發揮在「中文動詞自動分類」(曾慧馨等 2002)、「應用混合式機率模型於新聞資訊檢索之研究」(李建志 民91)等等,使用極廣。本研究中將利用 CKIP 服務處理中文斷詞,以減低系統建置成本與提昇服務穩定度。

在知識的表示法上,常以概念(Concept)為最小的表示單元。多個概念之間,有可包羅繼承,從屬或邏輯關係,藉以組成為知識體系。常見的知識與概念的表示法,約有三類:(1) 樹狀:各概念通常組織成一個樹型結構以表示相互之間的從屬次第關係(Su & Iiebrekke 2002) 例如本體論的樹狀表示法,(2) 網狀:以網絡的方式表示其相互交錯的連結狀態,例如語意網表達語意關聯的交互參考網狀表示法,以及(3) 地圖:針對特定領域的多個概念,以空間定位方式表示其相互關係,例如以主題地圖表示法表示出特定領域的概念映對(Sigel 2000)。在處理知識的表示時,三種表示法可依序使用,或綜合的透過內隱或外顯的方式交互應用。例如使用 Ontology 建構本體知識,應用於語意網或主題地圖的建構(Su & Gulla 2006),就可藉知識本體論定義的關鍵詞,進行邏輯推理。本研究中將以樹狀結構表示以句為單位的句型結構,作為句意知識表示核心架構,並利用本體論,與意網與主題地圖的理念建構相關索引服務。

二、本體論、主題地圖與資訊檢索

本體論是對某一個概念的詳細描繪,其中包含了對於概念、關聯、實體的表述,並清楚的定義了要表達的概念,主要用於知識的分享與再利用(Lee et al. 2007; Thomas 1995)。本體論也是將通識概念正式詳盡闡明與抽象化塑模的方式(Thomas et al. 2007)。它可為特定領域的概念或知識提供一般性的定義。本體論在哲學上廣泛被使用來闡釋萬物本然存在的屬性,在電腦科學方面的應用上,由於本體論具有層層相依,階組性與關聯性等特色,能夠表達特定領域的物件特性與物件之間的關聯(Chandrasekaran et al. 1999) 特別適合整合展現特定領域辭彙,建構知識庫的骨幹(Swartout et al. 1997)。利用本體論,可以將知識轉化為概念規則(姜琇森等 2007)用於知識的表達與處理(Bunge 1977)。應用本體論將特定領域的知識建立通用的解釋或表示架構,也常用以建立描述概念及概念之間關係的概念模型(Hwang 2003)。實務上,可以針對不同領域物件特性依據本體論的建置方法論建構專屬的本體論(Fernández et al. 1997)。本體論可以用來延展查詢(Query)的語意,藉由特定辭彙可能隱含的語意,從而推求與查詢真正相關的資訊(Erdmann & Studer 2001)。在資訊檢索和特別領域使用本體論,可以提供結構化和形式化的描述來表達和分享個人概念。目前在語意網的處理上,常藉由本體論作為基礎(Lee et al. 2007)。另外在很多領域也常用本體論來解決,如辭彙上的語意模糊(Stevens et al. 2002)、知識管理、電子商務或數位化教學問題(Kim et al. 2007)。本體論依所描述領域的不同呈現極富變化的內涵,其建立過程一般而言是將概念或知識切分為小單元,並以特定辭彙去描述它。然後藉由一些條件或規範界定這些單元的關係與定義,而形成在這專業領域(domain)中可以解釋其知識的架構。透過本體論建立各元素之間正確的關聯,藉以在特定領域中延展搜尋的深度與廣度,以提昇檢索成效。

主題地圖是2000年由國際標準組織(International Organization for Standardization; ISO)和國際電機技術委員會(International Electro-technical Commission; IEC)聯合制訂的國際標準(ISO/IEC 13250:2000 2002)，可用以表達與重現知識的方法。主題地圖中包含了一整群以概念相互連結的知識網絡的型態呈現的物件(Fisher et al. 2000)。主題地圖又猶如書後所附的關鍵詞索引(ISO/IEC 13250:2000 2002; 林光龍 & 歐陽彥正 民91)，包含主題(topics)、關聯(associations)與事件(occurrences)等三大類型資料。在資源的分享與檢索部份，使用資源物件本身的後設資料描述資料結合本體論機制可以提供自動推演的服務機制，以提昇檢索命中率與效能。本體論是對某一個概念的詳細描繪，其中包含了對於概念、關聯、實體的表述，並清楚的定義了要表達的概念，主要用於知識的分享與再利用。本體論也是將通識概念正式詳盡闡明與抽象化塑模的方式。它可為特定領域的概念或知識提供一般性的定義。應用本體論將特定領域的知識建立通用的解釋或表示架構，也常用以建立描述概念及概念之間關係的概念模型。一個具有良好結構的本體論也可以讓我們對巨量資料複雜的查詢與自動化推論新知識(Bada & Altman 2000; Russell & Norvig 2002)。主題地圖由於能表達多個物件的配置與從屬關係，可以整合眾多分散於網路上的知識庫，進而建構出一個強大的知識管理系統(林信成等 2004)，目前常見其應用於表達本體論與語意網絡的骨幹結構。

三、中文斷詞與知識管理

機器翻譯(Machine Translation; MT) 是使用電腦將一種自然語言自動翻譯成另一種的自然語言。機器翻譯藉由使用語料庫(corpus)的技術可更佳的處理不同的文法結構、辭彙辨識、慣用語的對應等，也能提供自然語言處理(Natural Language Processing; NLP)的教材(Jurafsky & Martin 2008)。一般的機器翻譯分為規則法和統計法兩種。規則法是用文法結構以及大量字詞規則所建構的。其主要的問題在於無法提供足夠的資訊來滿足不同語言規則之間的轉換。而且用這種方法所翻譯出來的作品，往往有太多的例外無法用規則來有效規範。其中有很多是只用於「字對字」或「辭對辭」之間的翻譯，較適合於專有名詞之互相翻譯等特定領域的應用，如產品型錄或使用手冊的翻譯。而統計機器翻譯(Statistical Machine Translation)較適合非特定領域的應用，透過對大量的平行資料庫作統計分析進而建構起翻譯模型(Lopez 2008)，其缺點在於只適合於較短句子的翻譯。本研究的目地在於融合以上兩種方法的優點，既可以產生完整句子對句子的翻譯，也可以用於非特定領域。大部分的搜尋技術是根基於「關鍵字」的搜尋，比須要完全符合所輸入的查詢字或詞，才會找到精確的結果。但是一般使用者往往無法使用和想找出的資料完全一樣的關鍵詞。於是模糊搜尋(Fuzzy search)方法的出現。所謂的模糊搜尋是用字串的pattern互相計算之後找出其相似值的一種搜尋技術(Gonzalo 2001; Navarro et al. 2001)。大部分的文字搜尋是利用字串轉成「一般表示法」(regular expression)之後，計算兩者之間的距離來決定其相似程度。距離愈小表示兩個字串愈相似。這種比對方式較適合用於詞彙之間的轉換。對於整句的搜尋幫助不大。舉例來說，若你想要找「我要去火車站」的句子。利用模糊搜尋的機制，可能會找到語料庫中所有和「火車站」有關的句子，例如

說「請問火車站怎麼走？」或是「這火車跑得很快」之類的句子。找到的其實都不是原來使用者所想要表達的意思。但是透過本研究的句構比對法，即使語料庫中沒有原來欲找尋的句子，卻會比對出如「我要去機場」或「我要去學校」之類的有類似句型結構的句子來。就學習外語來講，只須要將翻譯出來之外語的「機場」或「學校」改成「火車站」，就可以完成你想要表達的完整意思。因此很適合初學者或出國時使用。如果搭配手持式的行動裝置，如手機或PDA等，更能發揮其行動支援的功能。

四、目前常見語文學習系統的問題：

目前常見的語文學習系統，經分析計有「以字翻字無法掌握句構全貌」、「固定的句型無法適性適境滿足學習所需」、「缺乏知識管理的基礎支持」及「對於無所不在的學習支援不足」等待解問題，分述如下：

（一）以字翻字無法精確掌握句子中的意義

目前語言學習軟體仍然侷限在以字翻字的階段，對於需要完整應用的句子卻沒有更多的著墨，也無法掌握全句的真正意義。是目前主要翻譯軟體的缺失。

（二）固定的句型無法適性適境滿足學習所需

傳統學習外語方式大都被動接受教材提供的句型，但是所學的外文句子常常不合乎學習者溝通時所需的句子，所以利用中文尋找所需的外文表達更合乎學習外語者的需求。目前學習外語表達時學習者常使用電子辭典以及使用機器翻譯軟體。然而電子詞典提供外文詞彙，而經由各種外文詞彙常常無法組成道地的外文句子。以機器翻譯由中文尋找外文句子，經常會產出錯誤的外文詞句。

（三）缺乏知識管理的基礎支持

語言本身就是一種知識。知識管理是發現、存儲、維護、傳遞、應用與創造知識的動態過程。在語言學習活動中，導入知識管理的理論與技術，例如以本體論、語意網主題地圖與等知識管理模型作為智慧型分類，將可有助於學習者組織語文學習的整體結構，藉此提高學習成效。然而目前常見的語文學習或對譯服務中卻缺乏上述知識管理的基礎支持。僅能提供給學習者零散與片段的語詞資料，並無法有效幫助學習者建構與擴展語言知識。

（四）對於無所不在的學習支援不足

語言的學習強調與學習者的生活經驗相互結合，以利透過學習者進行語言知識內化，無所不在的學習服務對此提供了良好的機制。然而近年來廣泛被討論應用的服務導向架構則很適合作為無所不在的學習服務的基礎，目前常見的語文學習或對譯服務卻未提供使用者這些便捷易用的機制，對於無所不在的學習支援更明顯有所不足。

參、系統架構與實作

在語言學習上，尤其是外語的學習。初學者往往無法用很精確的外語句子來表達其真正的意含。(Chen 2007)發現語言學習策略與語言能力呈現正相關。也就是說，使用多種的語言學習策略，其學習的效果也更好(Wenden 1986)。我們發現可以運用我們的系統在語言學習策略中的「補償策略」(compensation strategies)來輔助學習者在說與寫的訓練(Oxford 1989)。這套語言學習策略系統簡稱為SILL (The Strategy Inventory for Language Learning)，而其中的「補償策略」即鼓勵用外語的片語或單字來組合成完整的句子，並善用母語為溝通的媒介來尋找援助（提問問題或確認外語用法的正確性）。本研究在提出於輸入母語後，列出與其句型結構相似的德語的完整句子，並依其句型結構間之相似程度排序，正好可以提供語言教學者或學習者很多可選擇的語意辭彙(linguistic clues)。有了這些新辭彙的輔助，配合句型結構，對於外語句子的組合也就愈有信心。同時對於外語的學習也就更能得心應手。目前常見語文學習系統多侷限於以字找字及以固定句型提供基本資訊。本研究突破傳統詞彙比對的作法，以樹狀結構表示句型結構，提供有序與高擴展性的句型結構框架，由全句的句型元素結構與其中各單詞的辭性進行快速索引，並以本體論(Ontology)結合主題地圖(Topic Maps)概念建構不同情境的應用領域路徑，以克服迥異情境下辭彙取用問題。以下分別針對句型元素表示法與系統架構與設計的重要概念加以說明，最後提出實作展示：

一、句型元素表示法

本句子結構中的句型可透過句型元素加以組合與表示。句型元素的資料表示法可透過如環結構，星結構或樹結構等方式建構。本研究發現，環結構面對複雜的句子結構時，其表示式將倍趨繁複，也不易有效的維護與檢索；星結構中則具有以單一源起為中心的特色，但卻不能簡單表示句型的複雜度。透過樹結構表示句型元素結構，則能夠提供有序與高擴展性的句型結構框架，不但建構彈性高，也易於映射使用者輸入查詢的句型片段，更具有高度的容錯能力。故本研究中提出系統將以樹結構作為句型元素的動態表示方法。

當接收使用者輸入的中文查詢句子斷詞後，進行相關後續處理：(1) 以單詞詞性進行句型分析：完整句子經斷辭為多個單詞之後，能透過序列組合，由句型檢索機制在建構的合法句構樹結構中找出其全句或子句的定位與關聯，或與合法句構樹的既有結構進行比對，找出差異。(2) 所屬領域定位，找出生詞與類同詞句：由本體論相關處理功能找出單詞本身的定位與關聯，進行迭代式生詞與類同詞句的檢出。(3) 單詞解釋：針對個別單詞進行詞意分析，提供多語對譯。(4) 延展服務：保留提供以單辭為基礎的延伸查詢服務，或配合TTS或VoiceML建構語音服務的能力。正確的中文斷詞的具有極高度的重要性，本系統的中文斷詞系統是使用中研院所開發的中文斷辭系統CKIP (<http://ckipsvr.iis>).

sinica.edu.tw/) 服務將使用者的查詢字串拆解成不同的詞性之組合，如「她具有溫和友善的天性。」使用CKIP斷詞結果可得到「她 (Nh) 具有 (VJ) 溫和 (VH) 友善 (VH) 的 (DE) 天性 (Na) 。」。

(一) 本體論與樹結構：

本研究中的本體論結合樹結構表示句型結構，將經由CKIP 斷詞的單詞，分層各自映對到本研究設定的本體論結構中。例如主句 “他乘船”，斷詞為表示移動動詞 “乘”，在本研究設定的本體論結構中可得知其同輩動詞可為 “跳、走、跑、架”，乘與架的相關受詞可為同為交通工具之 “車、船、機”，因此可以提出：(1) 他乘船 (2) 他乘車 (3) 他乘機 (4) 他架船 (5) 他架車 (6) 他架機等衍生領域類句，如圖1。

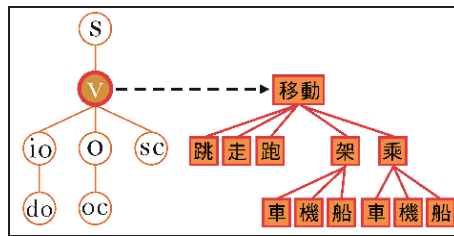


圖1：在樹結構下以本體論結構衍生領域類句

(二) 本體論與主題地圖：

為擴展以上的架構，本研究應用本體論與主題地圖的概念於建構具有知識推演潛力的句庫結構，有效的建立適用於特定領域與環境使用，有序列關聯意涵的句子集合—情境句序：特定情境下彼此相關的句子可以組成情境句序，各自獨立的情境句序結構又可歸屬於同一個主題情境，如圖2，繼而構成情境句庫，期能於進行句庫資料搜尋時，可以透過使用者用句的即時定位解析，找出它的關聯句序，提供建議參考句序資料。

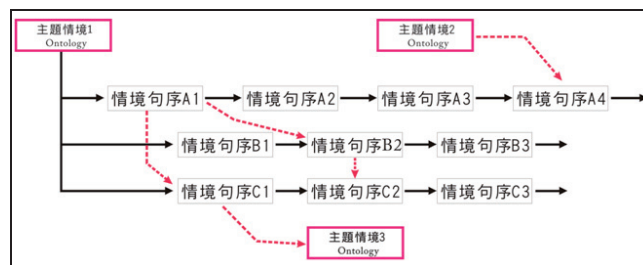


圖2：建立主題情境本體，並導出跨主題情境句序的關聯結構

根情境句序實例如：車、船、車站、港埠皆屬於交通領域本體物件，透過主體地圖，可以依場所情境建立不同的句序集合，如在車站、港埠內，都可能具有購票、用餐問路等等主題情境，將在各個情境之中的句子串聯，就成為個別的情境句序。各個情境

句序之間有可能有所關聯，如車站內的購票與問路主題可能有類同現象，甚至跨主題情境的句序也可能會有所關聯，如車站內的問路與港埠內的問路句序可能有類同現象，如圖3。

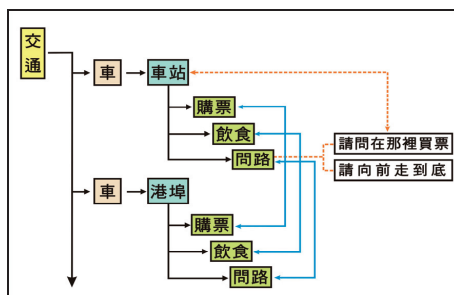


圖3：情境句序示意圖

二、系統架構

本系統由「基礎資料」、「前置處理模組」、「基礎媒合搜尋模組」與「進階媒合搜尋模組」等三大部份，系統架構圖如圖4，說明如後。

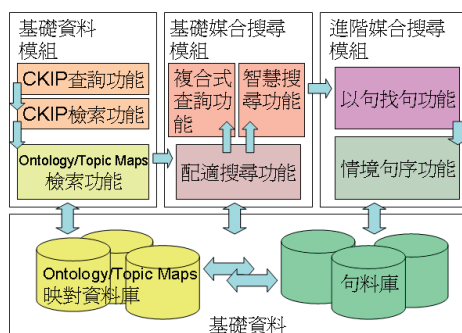


圖4：系統架構圖

在「基礎資料」部份，包含「句料庫」與「Ontology/Topic Maps映對資料庫」：

(一) 句料庫

目大量閱讀雙語類句可以有效的幫助外國籍學生進行第二語言的學習(Krashen 2004)。然而，對於需要藉由大量閱讀吸收學習外國語文的語言學習者而言，學習的素材內容如果沒有百分之百的正確度，勢將難以避免在其語言內化的過程中發生誤入誤出的問題。因此，為了提供學習者正確的學習範句，使用平行中-德語料庫為操作基礎。並且為配合後續處理需求，語料庫中的中文句子皆以CKIP查詢功能及CKIP索引功能進行前置處理，完成句子的斷辭標註與相關索引的建立。

(二) Ontology/Topic Maps映對資料庫

對於語料庫中的句子，透過前小節說明的樹結構Ontology辭性解析，與其多接點Topic Maps 串聯定義方法，以Ontology/Topic Maps檢索功能進行預前處理。依Ontology從屬關聯與Topic Maps句序關聯對每一個句子建構供檢索使用的 metadata，作為本系統Ontology/Topic Maps檢索功能、以句找句功能與情境句序功能的處理基礎。

在「前置處理模組」部份，可分為「CKIP查詢功能」、「CKIP索引功能」與「Ontology/Topic Maps檢索功能」：

(一) CKIP查詢功能

提供以即時線上查詢，或是批次查詢的方式，將使用者提供之中文文句資料自動送入CKIP查詢服務進行中文斷詞查詢，取回CKIP斷詞結果的功能。例如，當其中設有的CKIP查詢功能收到使用者提供的中文句資料後，將以單句為單位，逐句向CKIP查詢中文斷詞結果，如果查詢失敗，將再次嘗試送出查詢3次，如果查詢成功，將進行格式解碼，如果正確無誤，將記錄該CKIP斷詞於斷詞中繼資料檔中，之後接續後筆資料查詢時，系統內部將採複合式語句進行條件過濾檢索直至所有CKIP斷詞需求處理完畢。

(二) CKIP索引功能

將由CKIP查詢功能取回的CKIP斷詞結果，依單字詞與詞性關係，以六種不同的標註方式建構該句子的字詞與辭性的描述資訊(metadata)索引，以利提供在不同檢索需求下進行最佳化利用。例如，其中設有之CKIP索引功能主要在於將由CKIP查詢功能取回的CKIP斷詞結果，由簡至繁，依單一或混合索引的差異建立六類句構辭性索引，以利依不同需求進行最佳化利用。CKIP索引功能取得CKIP查詢功能逐句向CKIP查詢中文斷詞結果，以進行六類辭性分類，並將結果儲於資料庫，直至所有CKIP斷詞結果處理完畢。

(三) Ontology/Topic Maps檢索功能

使用者輸入待查句子時，由前述CKIP查詢功能透過即時方式，對CKIP服務要求斷辭，取回待查句之CKIP斷詞結果，之後透過CKIP索引功能依CKIP斷句與辭性對句中單字與單辭元素進行 metadata 的建構，轉換為符合以本系統Ontology/Topic Maps映對資料庫索引的狀態，成為「待檢句」。Ontology/Topic Maps檢索功能，繼此提供了以該「待檢句」資料對Ontology/Topic Maps映對資料庫的內容進行特徵值比對，找出符合預設門檻值的檢索用 metadata 與對應句料集。例如，其中設有Ontology處理功能，將由CKIP查詢功能取回的CKIP斷詞結果，進行Ontology詞性定位與詞意關聯分析，以迭代方式推搜找出所屬領域相關關鍵辭句分類。Ontology處理功能逐句向CKIP查詢中文斷詞結果，以進行Ontology領域辭性分類，並將結果儲於資料庫，直至所有CKIP斷詞結果處理完畢。

在「基礎媒合搜尋模組」部份，可分為「配適搜尋功能」與其下「複合式查詢功能」與「智慧搜尋功能」兩個子功能：

(一) 配適搜尋功能

本功能針對Ontology/Topic Maps檢索功能所檢出之符合預設門檻值的檢索用 metadata 與對應句料集，分別依照：

- (1) 單辭媒合率：斷句後全句中，單辭相符的比率。
- (2) 句構媒合率：斷句後全句中，辭性句構相符的比率。
- (3) 單一單辭與句構：斷句後全句中，每個單一的單辭與句構相符的比率。
- (4) 複合單辭與句構：斷句後全句中，各個以複合型態出現的單辭與句構相符的比率。
- (5) 單字媒合率：未全句中，單辭相符的比率。

等不同特徵媒合比率進行適配權值運算，最後產生依適配權值高低排序的前處理資料集，提供複合式查詢功能或智慧搜尋功能進行進階查詢處理，期能提高使用者搜尋句庫資料的準確性，分述如下：

1. 複合式查詢模組

本系統對於使用者提出的多樣化查詢需求，透過正規的資料庫SQL查詢語句進行資料查詢，以保障查詢服務的穩定度，其中主查詢語句與可能存在的各子查詢語句以鬆耦合方式結繫，並依句首、句中、句尾等方式分類，動態建構複合式查詢服務，以增加實作上的操作彈性。

2. 智慧搜尋模組

為提供使用者便利的搜尋經驗，本系統融合 Ontology 與 Topic Maps 等知識管理方法，建構智慧搜尋服務機能，使用者僅需輸入單一關鍵句。本系統內建之智慧搜尋功能將透過「辭性及辭意媒合」與「句序媒合」等句構媒合機制自動分析輸入句子，透過內建句庫提出正確類句建議，如圖5。

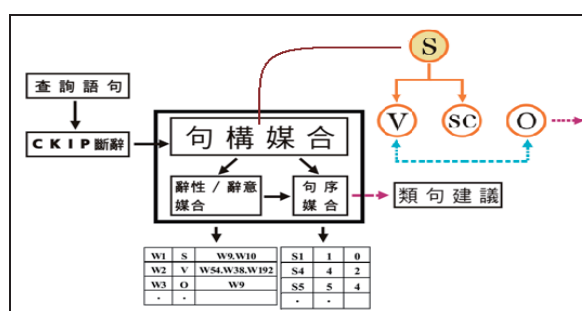


圖5：智慧搜尋功能架構圖

其中，句構媒合機能透過「辭性及辭意媒合」與「句序媒合」完成：

(1) 辭性及辭意媒合

Ontology/Topic Maps映對資料庫提供依預設的主題地圖與領域本體論賦於從屬與關聯特徵屬性。當接到使用者的查詢語句之後，系統首先將查詢語句透過CKIP斷辭服務取

的斷詞結果，此時原查詢句將被分析為由單字辭與其句構辭性組成的序列資料。句構媒合經由原始句庫內建的本體論句構樹，找出與斷辭結果的句構相符或相近的句構句型。句構如S-V，S-V-O或S-SC等等。經由內建的語意網關聯，找出與斷辭結果的單辭與單句的相符或相近的單辭與單句。圖中左下表中W1.. Wn 表示單字辭，S，V，O.. 為該單字辭於不同句構下可能的辭性。表中依這兩組資料為分類鍵值，藉以索引其同辭性不同語意單字辭的關聯資料，如 W1 單字辭為 S辭性時，與W9，W10有關聯。

(2) 句序媒合

經由內建的語意網關聯與主題地圖，找出與斷辭結果的句序相符或相近的單句。如圖左下表中左邊欄位S1.. Sn 表句子，其後兩欄位分別為本身索引鍵值與前句索引鍵值。透過索引鍵值之間的關係，就能提供相關句序的句子。使用者輸入的查詢句能直接反映使用者真實的語意需求，透過上述智慧搜尋，可以自內建句庫中提供直接媒合與近似類句。

智慧搜尋服務流程圖如圖6：

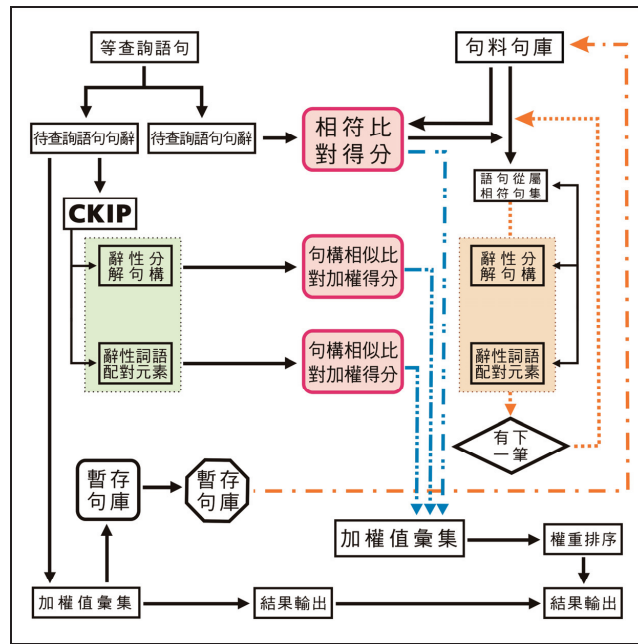


圖6：智慧搜尋功能流程圖

如圖6所示，系統首先比對句庫中相同學習單元及具有相同源頭學習單元之子句庫進行「相符比對」，並給予得分（如2.00000）。另外，將該句進行CKIP斷辭，取回斷辭結果之後，轉換為「辭性分解句構」與「辭性詞語配對元素」兩類特徵值結構。例如，查詢句「請問在那裡可以兌換外幣」的CKIP斷辭結果為「請問(Vt)在(P)那裡(N)可以(ADV)兌換(Vt)外幣(N)」，其辭性分解句構為：

VT，P，N，ADV，與VT，P，N，ADV，VT，N 序列。

辭性詞語配對元素為：

請問-Vt，在-P，那裡-N，可以-ADV，兌換-Vt，外幣-N。

系統將上述兩資料與句庫內進行比對，並分別給予加權得分（如0.60000，0.33333），再將這三類得分予以加總，最後將全部搜彙得出的結果依加總後權重進行排序。

在「進階媒合搜尋模組」部份，可分為「以句找句功能」與「情境句序功能」：

（一）以句找句功能

承襲本系統架構特點，本系統進行中德句庫以句找句運作時，系統內部將採複合式搜尋語句進行條件過濾檢索，分別是：

- （1）依使用者需求建立主搜尋語句(S)。
- （2）以主搜尋語句整合數個來自CKIP斷詞後的領域搜尋子語句。
- （3）依單一或混合句構詞性索引，建立詞性搜尋子語句。
- （4）依使用者特別指定的單詞或子句建特定字詞搜尋子語句建立複合式搜尋語句之後，即可順利找出使用者需要的對譯或類句資料。

（二）情境句序功能

系統基於本體論與主題地圖原理建立句庫的骨幹結構（圖6），將用以建構情境句序功能。系統實作部份，情境句序功能分「情境句序前置處理」與「情境句序搜尋」兩階段完成情境句序的檢出服務。首先，在「情境句序前置處理」部份，系統預將句庫建立者定義具有父-子從屬關係的單句間關聯資料，建立句對，之後進行多句對間之索引化處理，建立句序索引集儲於Ontology/Topic Maps映對資料庫。在「情境句序搜尋」部份，當系統以句找句服務檢出使用者查詢句的類句之後，系統自動將每一個查詢結果句子與上述句序索引集中的句序索引進行比對檢索，找出該句的句序建議，產出情境句序。

肆、系統實作

一、開發環境與工具

本系統開發依循前述系統分析，在伺服器方面，以Microsoft Windows Server 2003 為基礎平台，Microsoft IIS6 作為網頁伺服器、Microsoft SQL server 2000作為資料庫伺服器、Microsoft DNS Server 作為名稱解析伺服器、Microsoft UDDI Service 作為 UDDI伺服器與 Microsoft dot net platform 進行協同運作；在開發工具方面，則並以 Borland Delphi 6 與2005 以及 Microsoft VS.net 2005 透過 Microsoft ASP.NET 技術開發動態網頁，並編寫對本體論關聯法則。

二、系統介面與實例

本系統提供Web Page方式呈現的使用者介面，使得學習者無需安裝與依賴專屬服務

程式，僅需利用通用的網頁瀏覽器進行學習。一般外語學習者，常慣用母語為溝通的媒介來提問問題或確認外語用法的正確性，並以用外語的片語或單字來組成完整的句子(Oxford 1989)。本服務提供彈性查詢機制，學習者登入服務網頁後，可以針對章節進度，對句構相似程度與辭性與字詞配適度等權重與待查詢中文語句分別彈性指定，當送出查詢後，系統自動依指定權重提供最適類句。由於本系統提供以句構為核心的類句搜尋，故可以在完全沒有匹配關鍵字辭下提供跳躍式檢索，查詢輸出結果則提供有各類句的正確發音，句構資訊與各單辭進階查詢等等延伸學習資源，提供學習者進行自主式學習活動，作業各階段處理程序與系統模組、功能之間的相關對映關係如表1。

表1：系統各模組使用系統功能、處理程序與主要配伍資料項別之間的相關對映關係表

作業模組	使用功能	處理程序	主要配伍資料項別
基礎資料	CKIP查詢功能 CKIP索引功能	(前置處理，對句料庫中全部資料以六種不同的標註方式建構該句子的字詞與辭性的描述資訊(metadata。)	句料庫 Metadata
	CKIP查詢功能	接收使用者輸入「待檢句」字串，即時轉送CKIP查詢並獲得斷辭結果。	
前置處理模組	CKIP索引功能	對斷辭結果以六種不同的標註方式建構該句子的字詞與辭性的描述資訊(metadata)索引。	Metadata
	Ontology/Topic Maps檢索功能	以句找句前置處理： 對「待檢句」資料對Ontology/Topic Maps映對資料庫的內容進行特徵值比對，找出符合預設門檻值的檢索用metadata與對應句料集。 情境句序前置處理： 將句庫建立者定義具有父—子從屬關係的單句間關聯資料，建立句對，之後進行多句對間之索引化處理，建立句序索引集。	Ontology/ Topic Maps 映對資料庫
基礎媒合搜尋模組	配適搜尋功能	依不同特徵分析媒合比率，計算適配權值最後產生依適配權值高低排序的前處理資料集。	
	複合式查詢功能	剖析「待檢句」斷辭，依句首、句中、句尾等方式分類透過正規的資料庫SQL查詢語句進行資料查詢類句建議。	句料庫
	智慧搜尋功能	透過「辭性及辭意媒合」與「句序媒合」等句構媒合機制自動分析輸入句子，並透過內建句庫提出類句建議。	句料庫 Ontology/ Topic Maps 映對資料庫
進階媒合搜尋模組	以句找句功能	提供外顯「以句找句」服務。 依類句建議進一步依單一或混合句構詞性索引，建立詞性搜尋子語句，或依使用者特別指定的單詞或子句建特定字詞搜尋子語句建立複合式搜尋語句之後，找出使用者需要的對譯或類句資料。	句料庫 Ontology/ Topic Maps 映對資料庫
	情境句序功能	提供外顯「情境句序」服務。 由系統以句找句服務檢出使用者查詢句的類句查詢結果為基礎，將每一個查詢結果句子與Ontology/Topic Maps映對資料庫metadata資料進行媒合比對，檢出該句的相關句序建議。	句料庫 Ontology/ Topic Maps 映對資料庫

以下透過情境模擬說明本系統「以句找句」與「情境句序」服務：

本情境模擬以中文為母語的學習者欲透過本系統學習於搭乘交通工具時常用的德語相關會話，實務程序為：學習者在輸入其母語「請問在哪裡可以兌換外幣」句子進行查詢，經本系統服務，使用者將獲得「以句找句」與「情境句序」等服務。

以句找句，例如：

原待檢句「請問在哪裡可以兌換外幣」經CKIP查詢功能與CKIP索引功能處理，結果依單詞辭性區分的句構為：請問(Vt)在(P)哪裡(N)可以(ADV)兌換(Vt)外幣(N)。接下來，系統由Ontology/Topic Maps檢索功能與配適搜尋功能檢索句庫中相同句構的類句。此時優先檢出了與原查詢句構一致的句子：「請問在哪裡可以買到車票」（句構：請問(Vt)在(P)哪裡(N)可以(ADV)買到(Vt)車票(N)）。之後再透過複合式查詢功能或智慧檢索功能協同以句找句功能檢出其他類句，如子句包括與原查詢句構一致的句子，如：

「是否一個學術論文的發表在國內外都具品質」（句構：是否(ADV)-(DET)個(M)學術(N)論文(N)的(T)發表(Vt)在(P)國內外(N)都(ADV)具(Vt)品質(N)）以及

「請問我在哪裡可以兌換外幣」（句構：請問(Vt)我(N)在(P)哪裡(N)可以(ADV)兌換(Vt)外幣(N)），「我在哪裡可以買到航空郵簡呢」（句構：我(N)在(P)哪裡(N)可以(ADV)買到(Vt)航空(N)郵簡(N)呢(T)）等等類句。

情境句序，例如：

由情境句序功能對以句找句服務檢出的類句查詢結果，對Ontology /Topic Maps映對資料庫中metadata資料進行媒合比對，檢出該句的相關句序建議，例如：

使用者查詢「請問我在哪裡可以兌換外幣」之情境句序，將可以得到：

- 「請問我在哪裡可以兌換外幣」
 - >>你可以至銀行櫃台兌換外幣
 - >>請問銀行櫃台在哪裡
 - >>直走後右轉就可以看到銀行
 - >>請問換美金的兌換最大限額是多少
 - >>兌換美金的最大限額是一萬歐元
 - >>請問今日兌換率多少
- 等等具有序列意涵的相關句子。

以句找句與情境句序查詢結果如圖7。

相似度	資料編號	中文語句	句序
2.66666	4903	請問在哪裡可以買到車票 請問(Vt)在(P)哪裡(N)可以(ADV)買到(Vt)車票(N) [單辭查詢] [句構查詢] [檢索功能] [檢索功能]	
1.16666	11013	是否一個學術論文的發表在國內外都具品質 是否(ADV)-(DET)個(M)學術(N)論文(N)的(T)發表(Vt)在(P)國內外(N)都(ADV)具(Vt)品質(N) [單辭查詢] [句構查詢] [檢索功能] [檢索功能]	
1.00000	15867	請問我在哪裡可以兌換外幣? 請問(Vt)我(N)在(P)哪裡(N)可以(ADV)兌換(Vt)外幣(N)? [單辭查詢] [句構查詢] [檢索功能] [檢索功能]	1
0.66666	15839	請問可以在哪裡換匯票? 請問(Vt)可以在(P)哪裡(N)換(Vt)匯票(N)? [單辭查詢] [句構查詢] [檢索功能] [檢索功能]	2

相關句序！
>>請問我在哪裡可以兌換外幣?>>你可以至銀行櫃台兌換外幣>>請問銀行櫃台在哪裡?>>直走後右轉就可以看到銀行>>請問換美金的兌換最大限額是多少?>>兌換美金的最大限額是一萬歐元>>請問今日兌換率多少?

圖7：以句找句與情境句序查詢結果

伍、討論

本節針對本研究提出之系統架構對傳統作業方式的主要問題改善的益處，並由系統運作原理與架構效益評估提出說明：

一、傳統作業方式的主要問題

以傳統語言學習系統進行語言學習時的主要困境為：

1. 侷限在以字翻字的階段，學習者被動接受教材提供的句型，但是所學的外文句子常常不合乎學習者溝通時所需的句子。
2. 機器翻譯由直譯中文尋找外文句子，並未以句型與句構提供語意與句意的正確的概念與全貌。機器翻譯由中文尋找外文句子，又經常產出錯誤的外文詞句。
3. 缺乏知識管理的基礎支持，僅能提供給學習者零散與片段的語詞資料，並無法有效幫助學習者建構與擴展語言知識。
4. 未具無所不在的學習機能，學習無法適性適境的與生活經驗相互結合。

二、本研究相對優勢評估

(一) 本研究提出之系統架構對傳統服務困境的突破

本研究突破傳統詞彙比對的作法，以樹狀結構表示句型結構，提供有序與高擴展性的句型結構框架。由全句的句子組成結構與其中各單詞的辭性進行快速索引，並以知識管理之本體論、語意網與主題地圖建構句意關聯與不同情境的應用領域路徑，克服迥異情境下辭彙取用問題，能夠針對前述傳統語言學習系統進行語言學習時的問題提供完整的改善。

(二) 系統架構相對優勢評估

1. 以句找句的效益：

雖然單字辭是構成句子的最小單位，但是單句卻是表達意念與進行有效溝通的基礎，因此本系統針對傳統語言學習系統以單字辭為主的設計模式產生以字找字與以字找句的侷限性，突破性的建構以句找句的方式直接提供學習者利於透過類句語意串聯，強化語言學習成效。

2. 以句構與句型索引提供類句的效益：

在以句找句的系統實作方面，本系統透過樹結構表示句型元素結構，可以提供有序與高擴展性的句型結構框架，提供了以句構與句型為基礎的類句索引機制，不但建構彈性高，也易於映射使用者輸入查詢的句型片段，且具有高度的容錯能力。

3. 情境句序導入的效益：

本系統情境句序服務，提供使用者以本體為綱，情境主題為目的學習體驗，透過具有關聯意義的句序導引，可有效幫助學習者強化記憶聯想，熟悉語文應用的能力。

4. 知識管理的導入的效益：

本系統以知識管理的理論與技術作為建構語言學習的基石。透過本體論、語意網主題地圖與等知識管理模型建立的智慧學習功能，可有效幫助學習者組織語文學習的整體結構，提高學習效果。

三、系統機能與成效評估

為瞭解本系統機能與服務是否適切，對使用者學習是否有所助益，以及使用者的使用滿意程度，本研究系統開發與測試分別以「專家評鑑」、「使用者訪談」與「教師與使用者問卷調查」等方式進行評估，實施操作分述如下：

(一) 專家評鑑

為了確認系統服務的有效性，強化系統服務機能，本研究進行中特別邀請資訊科學領域，語文教學領域與知識管理領域之資深教育學者，為本服務進行專家評鑑本系統服務機能，並依評鑑建議修改系統服務設計。

(二) 使用者訪談

為了使得本系統之服務內容與使用者需求目標一致，瞭解使用者真實需求，增進使用者使用效能，本研究進行中特別以立意抽樣(purposive sampling)方式，針對具有網路自我學習經驗的使用者進行訪談，並依評鑑建議修改系統服務機能。

(三) 學習者滿意度問卷調查

1. 問卷設計

科技接受模型(Technology Acceptance Model (Davis et al. 1989))與系統成功模型(D&M IS Success Model (DeLone & McLean 1992; DeLone & McLean 2003))廣泛的應用於評估使用者對新科技導入的接受度與系統成效評估(Lee et al. 2003; Rai et al. 2002)本問卷主要參考上述兩模型的衡量變項，針對本研究的情境加以修改，分別就：

- (1) 資訊/知識品質(Learning Material)：系統輸出對使用者有價值的程度。
- (2) 系統品質(System Function)：本系統本身的品質水準。
- (3) 易用性(Easy of use)：使用者認知本系統易於使用操作的程度。
- (4) 有用性(Usefulness)：使用者認知本系統能夠提供幫助的程度。
- (5) 使用者滿意度(Satisfaction)：使用者對本系統服務的滿意程度。

等五個構面發展李克特量表(Likert Scale)之五點評量尺度問卷(1：非常不滿意 .. 5：非常滿意)

2. 問卷測度

在問卷的信度評估方面，本研究以Cronbach α 係數驗證本問卷測驗結果的一致性與穩定性，各構面標準化Cronbach α 值均大於 0.72 以上，表示有很好的信度。在問卷的效度評估方面，本研究採用專家評鑑與前測(pretest)的方法測量問卷有效性，延請 3 位領域

專家學者進行問卷評估，與 3 名具有二年以上線上語言學習經驗的學生進行問卷前測，根據他們的意見進行問卷的修改。

3. 問卷調查結果分析

本問卷調查由使用本系統的大學部學生與研究生計149人進行填答，回收計114卷，回收率約為 77%，表2 列示問卷中各評估項目的結果，依統計平均值進行相關討論 (Suhong & Binshan 2006)說明如後：

表2：使用者滿意度問卷調查統計結果

問卷題目		平均值	總平均值
資訊知識品質	01·我覺得本學習系統提供的學習教材內容是具有實用性的	4.965	3.013
	02·我覺得本學習系統提供的學習教材內容份量是豐富的	3.596	
	03·我覺得本學習系統提供的學習教材內容是難易度是適中的	3.456	
	04·整體而言，我覺得本學習系統提供的學習教材內容對我是有幫助的	3.035	
系統品質	05·我覺得本學習系統的運作是穩定的	3.807	3.184
	06·我覺得本學習系統的使用介面設計是良好的	3.421	
	07·我覺得本學習系統的操作提示是足夠的	3.842	
	08·整體而言，我覺得本學習系統提供的系統機能是令我滿意	3.667	
易用性	09·我覺得本學習系統的操作方式是清楚以及容易了解的	3.211	3.890
	10·我覺得可以輕鬆學習就能使用本學習系統	3.912	
	11·我覺得我能輕易的使用系統的各種功能	3.561	
	12·整體而言，我覺得本學習系統的操作方式是容易的	4.877	
有用性	13·我覺得本系統可以增加我的學習效率	3.632	3.267
	14·我覺得本系統讓我能夠更輕鬆的進行學習	3.825	
	15·我覺得本系統可以改善我的學習成效	3.456	
	16·整體而言，我覺得本學習系統對我而言是有用的	3.158	
使用者滿意度	17·我對於本學習系統提供的教材內容感到滿意	3.895	3.690
	18·我對於本學習系統提供的系統功能感到滿意	3.070	
	19·我對於使用本學習系統的使用經驗感到滿意	3.281	
	20·整體而言，我對使用本學習系統獲得的學習成效感到滿意	3.193	

問卷系統提供資訊/知識品質、系統品質、易用性、有用性與使用者滿意度各構面得分總平均值介於3.184~3.890之間，各整體評估問項(04、08、12、16、20)總平均值3.386，可知本系統的創新服務表現皆能獲得多數使用者認可。其中，在資訊知識品質部份，總平均值3.013，表示透過本系統內部檢索機制的檢出結果能夠符合使用者期望。問項01。「我覺得本學習系統提供的學習教材內容是具有實用性的」獲分平均得分最高(4.965)，反映出本系統提出的以句找句與情境句序檢索機制，能夠滿足使用者的實際需求，具有實務上的效益；在系統服務的易用性與有用性方面，總平均得分在3.000以上，表示本創新系統將是易於被使用者所接受的；在系統品質部份，問項05。「我覺得本學

習系統的運作是穩定的」得分最高(4.877)，反映出使用者認同本系統的服務執行能力；在使用者滿意度部份，問項18·「對於本學習系統提供的系統功能感到滿意」所獲得分最高(3.8070)，反映出使用者對於本系統以句找句與情境句序檢索創新的服務方式感到滿意；另外，由於本系統較著重於機能面設計，在視覺介面則以極簡設計原則提供使用者簡易於使用的操作環境，這個特性已明顯的反映在問項06·「我覺得本學習系統的使用介面設計是良好的」與問項12·「整體而言，我覺得本學習系統的操作方式是容易的」兩部份得分的差距上。

陸、結論、研究限制與未來研究方向

本研究提出以句為基礎的語言學習架構，利用樹狀結構表示句型結構，提供有序與高擴展性的句型結構框架。並以知識管理理論與技術克服迥異情境下辭彙取用問題，對於傳統語言學習系統進行語言學習時的問題提供完整的改善，透過問卷調查，得知本系統的創新服務表現皆能獲得多數使用者認可，其具體創新成果為：

一、以句構與辭性為基礎的索引機制

突破傳統詞彙比對的作法，提出句構與辭性為基礎，由全句的句子組成結構與其中各單詞的辭性進行快速索引的機制。

二、整合知識管理技術的智慧型學習導引

以本體論(Ontology)結合主題地圖(Topic Maps)架構，建構不同情境的應用領域路徑，克服迥異情境下（如購物，問路等等）辭彙取用的問題。

三、具有延展能力的句型評比媒合機制

突破傳統詞彙比對的作法，以樹狀結構表示句型結構，提供有序與高擴展性的句型結構框架。由全句的句子組成結構與其中各單詞的辭性進行快速索引，能夠針對前述傳統語言學習系統進行語言學習時的問題提供完整的改善。

以下對於本研究限制與遭遇之困難，及解決途徑的作法提出3項要點，提供未來後續研究者參考：

（一）語文知識處理方面：

本研究以本體論為核心結合主體地圖處理語文知識，然而整合建構與維護的操作難度極高，建議先獨立發展，再整合應用。

（二）搜尋效能方面：

本研究的智慧型搜尋服務，搜索階層與範圍如不加以限制，可能造成無窮關聯的狀態，動態線上媒合大量關聯資料也將造成對服效能的低落。因應作為是定立有限的搜索階層與範圍，並於前置作業建立既有句庫資料的各種關聯資訊，以增進線上媒合的速度。

(三) 情境句序前置處理方面：

本研究的情境句序服務，目前以句庫建構者自行定意領域內主題的分類與句間從屬關係，透過前處理程序建構句序庫，以增進查詢回應效能。因此服務規模有所侷限，因應作為是強化領域本體自我的推理機能，減低人為介入操作的範圍，以提昇句序服務的處理效能。

在未來研究方向方面，由於目前後設資料僅應用於進行句辭屬性的基本管理，未來期能繼續擴展後設資料對知識管理的處理架構，以多階層抽象萃集的方式管理與應用後設資料，提昇系統服務的擴展彈性。另外，本研究中強調系統架構的建立，並未針對於本體論等知識管理資源的自動化與最佳化調校進行分析，期望在未來研究能對此進行更深入的探討。

致謝

感謝國科會專題研究計畫97-2631-S-327-001之經費贊助。

參考文獻

1. 中央研究院資訊科學所中文組實驗室中文詞知識庫小組 “中研院平衡語料庫詞類標記集,” 中央研究院資訊科學所中文組實驗室, 2007a (available online at <http://godel.iis.sinica.edu.tw/CKIP/paper/poslist.pdf>).
2. 中央研究院資訊科學所中文組實驗室中文詞知識庫小組 “詞庫授權資源,” 中央研究院資訊科學所中文組實驗室, 2007b (available online at <http://godel.iis.sinica.edu.tw/CKIP/publication.htm>).
3. 李建志, 民91, 應用混合式機率模型於新聞資訊檢索之研究, 國立成功大學資訊工程學系碩士論文。
4. 林光龍、歐陽彥正, 民91, 『應用主題地圖於知識管理系統』, 第四屆數位典藏技術研討會。
5. 林信成、歐陽慧、歐陽崇榮, 2004, 『以主題地圖建構索引典之語意網路模型』, 圖書與資訊學刊, 第四十八卷: 35~56頁。
6. 邱智銘 “詞類與詞類標記原則: 現代漢語,” 中央研究院, 2003 (available online at <http://linganchor.sinica.edu.tw/data/result/workshop/20030909workshop.html>).

7. 姜琇森、施東河、黃信銓，2007，『以本體論為基礎之惡意郵件偵測』，資訊管理學報，第十四卷·專刊：1~28頁。
8. 陳欣蓉，2008，『中文對照德文的溝通句型學習系統之設計』，歐洲語文學報，44~73頁。
9. 曾慧馨、劉昭麟、高照明、陳克健，2002，『以構詞律與相似法為本的中文動詞自動分類研究』，中文計算語言學期刊，第七卷·第一期：1~28頁。
10. Bada, M. A., and Altman, R. B. "Computational Modeling of Structural Experimental Data," *Methods in Enzymology* (317), January 2000, pp. 470-491.
11. Bunge, M. *Treatise on Basic Philosophy: Volume 3: Ontology I: The Furniture of the World*, Springer, 1977.
12. Chandrasekaran, B., Josephson, J. R., and Benjamins, V.R. "What Are Ontologies, and Why Do We Need Them?" *IEEE Intelligent Systems* (14:1), 1999, pp. 20-26.
13. Chen, M.H. "Language Learning Strategies and English Proficiency of Taiwan Language College Students," *Languages, Literary Studies and International Studies: An International Journal* (4), December 2007, pp. 49-72.
14. Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Management Science* (35:8), 1989, pp. 982-1003.
15. De Bono, E. *Practical Thinking: 44 Ways to be Right; 5 Ways to be Wrong; 5 Ways to Understand*, Baker & Taylor Books, 1992.
16. DeLone, W. D., and McLean, E. R. "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update," *Journal of Management Information Systems* (19:4), 2003, pp. 9-30.
17. DeLone, W. H., and McLean, E. R. "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research* (3:1), March 1992, pp. 60-95.
18. Erdmann, M., and Studer, R. "How to Structure and Access XML Documents with Ontologies," *Data Knowledge Engineering* (36:3), 2001, pp. 317-335.
19. Fernández, M., Gómez-Pérez, A., and Juristo, N. "Methontology: From Ontological Art towards Ontological Engineering," in *Proceedings of the AAAI97 Spring Symposium Series on Ontological Engineering*, 1997.
20. Fisher, K. M., Wandersee, J. H., and Moody, D. E. *Mapping Biology Knowledge (Science & Technology Education Library Volume II)*, Kluwer Academic Publishers, 2000.
21. Gonzalo, N. "A Guided Tour to Approximate String Matching," *ACM Computing Surveys* (33:1), 2001, pp. 31-88.
22. Gruber, T. R. "Toward Principles for the Design of Ontologies Used for Knowledge Sharing," *International Journal of Human-Computer Studies* (43:5-6), 1995, pp. 907-928.

23. Hwang, G.J." A Conceptual Map Model for Developing Intelligent Tutoring Systems," *Computer & Education* (40:3), 2003, pp. 217-235.
24. ISO/IEC 13250:2000" ISO/IEC 13250:2000 Topic Maps: Information Technology -- Document Description and Markup Languages," The International Organization for Standardization and the International Electrotechnical Commission, 2002 (available online at http://www1.y12.doe.gov/capabilities/sgml/sc34/document/0322_files/iso13250-2nd-ed-v2.pdf).
25. Jurafsky, D., and Martin, J. H. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition* (2nd ed.), Prentice Hall, 2008.
26. Kim, J.M., Choi, B.I., Shin, H.P., and Kim, H.J." A Methodology for Constructing of Philosophy Ontology Based on Philosophical Texts," *Computer Standards & Interfaces* (29:3), 2007, pp. 302-315.
27. Krashen, S. D. *The Power of Reading* (2nd ed.), Heinemann, 2004.
28. Lee, C.S., Kao, Y.F., Kao, Y.H., and Wang, M.H." Automated Ontology Construction for Unstructured Text Documents," *Data & Knowledge Engineering* (60:3), 2007, pp. 547-566.
29. Lee, Y., Kozar, K. A., and Larsen, K. R. T." The Technology Acceptance Model: Past, Present, and Future," *Communications of the Association for Information Systems* (12), 2003, pp. 752-780.
30. Li, S., and Lin, B." Accessing Information Sharing and Information Quality in Supply Chain Management," *Decision Support Systems* (42:3), 2006, pp. 1641-1656.
31. Lopez, A." Statistical Machine Translation," *ACM Computing Surveys* (40:3), 2008, pp. 1-49.
32. Navarro, G., Baeza-Yates, R., Sutinen, E., and Tarhio, J." Indexing Methods for Approximate String Matching," *IEEE Data Engineering Bulletin* (24:4), 2001, pp. 19-27.
33. Oxford, R. L. *Language Learning Strategies: What Every Teacher Should Know*, Heinle ELT, 1989.
34. Rai, A., Lang, S. S., and Welker, R. B." Assessing the Validity of IS Success Models: An Empirical Test and Theoretical Analysis," *Information Systems Research* (13:1), March 2002, pp. 50-69.
35. Russell, S. J., and Norvig, P. *Artificial Intelligence: A Modern Approach* (2nd ed.), Prentice Hall, 2002.
36. Sigel, A." Towards Knowledge Organization with Topic Maps," in *Proceedings of XML Europe*, 2000.
37. Stevens, R., Goble, C., Horrocks, I., and Bechhofer, S." Building a Bioinformatics Ontology Using OIL," *IEEE Transactions on Information Technology in Biomedicine* (6:2), 2002, pp. 135-141.

38. Su, X., and Gulla, J. A." An Information Retrieval Approach to Ontology Mapping," *Data & Knowledge Engineering* (58:1), 2006, pp. 47-69.
39. Su, X., and Ilebrikke, L." A Comparative Study of Ontology Languages and Tools," in *Proceedings of the Fourteenth International Conference on Advanced Information Systems Engineering*, Toronto, 2002.
40. Swartout, B., Patil, R., Knight, K., and Russ, T." Toward Distributed Use of Large-Scale Ontologies," in *Proceedings of the AAAI97 Symposium on Ontological Engineering*, 1997.
41. Thomas, P. D., Mi, H., and Lewis, S." Ontology Annotation: Mapping Genomic Regions to Biological Function," *Current Opinion in Chemical Biology* (11:1), 2007, pp. 4-11.
42. Wenden, A. L." Incorporating Learner Training in the Classroom," *System* (14:3), 1986, pp. 315-325.
43. Yeh, C.L., and Chen, Y.C." Creation of Topic Map by Identifying Topic Chain in Chinese," in *Proceedings of the 2004 ACM Symposium on Document Engineering*, Milwaukee, 2004, pp. 112-114.