

漢語語料庫與語法資訊

國立清華大學語言學研究所·致遠科技

湯廷池

一、前言

近幾年來，國內語言學家與資訊科學家對於漢語語料庫的建立表示關心與重視。建立漢語語料庫的目的，在於提出一定的畫分詞語與詞類的標準，並研究各種詞類的內部結構與外部功能，然後根據這些語法資訊廣泛蒐集漢語語料加以分析、分類並賦以各種詞性〔註1〕標誌，以便將來能利用電腦搜尋程式自動歸納漢語詞法與句法的規律性以及二者之間的相關性。在這一篇文章裡，我們準備從語法分析的觀點，討論漢語詞語的分斷、詞類的畫分與功能、複合詞內部結構的分析與分類以及與外部功能的關係等問題，以供語言學家與資訊科學家在建立漢語語料庫與決定詞性標誌時的參考之用。

二、詞語的分斷：詞、語(素)與詞組的畫分

在建立漢語語料庫的時候，第一個遭遇到的問題是如何從一串詞群中把個別獨立的詞分斷出來（叫做「斷詞」或「分詞」）；也就是說，如何界定「詞」（word）與「非詞」（non-word）的畫分。所謂「非詞」，包括小於「詞」的「語」（morph）或「語素」（morpheme），以及大於「詞」的「詞組」（phrase）。而所謂「詞」，則包括由單獨一個語素成立的「單純詞」（simple word）與由兩個或兩個以上的語素合成的「複合詞」（compound word）〔註2〕。單純詞只包含一個語素，但這個語素可能是單音（節）詞（如「人、書、門、雷、鬼」）、雙音（節）詞（如「玻璃、葡萄、蟋蟀、蝙蝠、鞦韆、檸檬、咖啡、菩薩、巴黎、納粹」）或多音（節）詞（如「凡士林、維他命、三溫暖、馬拉松、紐西蘭、蓋世太保、尼加拉瓜」）；而複合詞則依其內部結構可以分為「主謂式」（如「地震、肉麻」）、「述賓式」（如「主席、出席、順耳」）、「述補式」（如「跳高、分明、搖動」）、「偏正式」（如「電車、回憶、雪白」）、「並列式」（如「語言、廣大、動搖」）、「重疊式」（如「猩猩、爺爺、某某、往往」）等〔註3〕。

「語」或「語素」，依其能否單獨出現或獨立活動而可以分為「自由語(素)」(free morph(eme))與「黏著語(素)」(bound morph(eme))。但是，有時候，漢語裡自由語與黏著語的界限並不十分清楚，無法對每一個語或語素做很明確的區分。例如，有些語(素)（如「雲、時」）在白話或口語裡（如「雲彩、時候」）不單用而屬於黏著語，但在文言或書面語裡（如「雲、時」）卻可以單用而屬於自由語；有些語(素)（如「虎、言、語」）在一般說話裡（如「老虎、言語」）不單用而屬於黏著語，但在成語或熟語裡（如「前怕狼、後怕虎；你一言、我一語」）卻可以單用而屬於自由語；有些語(素)（如「氧、葉」）在一般說話裡（如「氧氣、樹葉」）不單用而屬於黏著語，但在專業文獻裡（如「氧(化學)、葉(植物學)」）卻可以單用而屬於自由語；有些語(素)在某一個方言裡（如北京話的「老虎、大象」）不單用而屬於黏著語，但在另一個方言裡（如閩南語的「虎、象」）卻可以單用而屬於自由語；有些語(素)（如「工」）在「工人、工藝、工業」這些意義上不能單用而屬於黏

著語，但在“工作”（如“作工、上工”）與“工程”（如“開工、動工”）這些意義上卻可以單用而屬於自由語〔註4〕。如果說自由語(素)是詞，而黏著語(素)則不是詞，那麼就這些詞而言，有關「口語」、「書面語」、「成語」、「熟語」、「專業術語」、「方言類別」等資訊就需要儲存於語料庫裡。另外，有關語素“工”因其含義的不同而有黏著與自由之分，則不容易單憑詞義來判斷，卻可以從這些複合詞的內部結構來做相當客觀而明確的區別：即屬於「偏正式」的“工{人／藝／業}”等的第一個語素的“工”是黏著語，因而在第一個語素與第二個語素之間不允許其他語素的介入；而屬於「述賓式」的“{作／上／開／動}工”等的第二個語素的“工”則可能是自由語，因而有時候在第一個語素與第二個語素之間允許其他語素（如動貌標誌的“了、著、起(來)”與動相標誌的“完、好”等）的介入〔註5〕。他如，述賓式的“幫忙”可以在前後兩個語素之間嵌入動貌標誌或修飾成分，但不常在複合動詞後面帶上賓語名詞組（如“幫了我的大忙”與“??幫忙你的朋友”），而並列式的“幫助”卻不能在前後兩個語素之間嵌入任何語素，但可以在複合動詞後面帶上賓語名詞組（如“*幫了我的大助”與“幫助你的朋友”）；述補式的“搖動”可以在前後兩個語素之間嵌入表示可能或不可能的“得、不”（如“搖{得／不}動”），而並列式的“動搖”則不能在前後兩個語素之間嵌入表示可能或不可能的“得、不”（如“動{得／不}搖”）；並列式的“嘲笑、譏笑”只能做不及物而不能做及物動詞使用（如“{嘲／譏}笑別人”與“*對別人{嘲／譏}笑”），而偏正式的“微笑、大笑、竊笑”等則只能做不及物而不能做及物動詞使用（如“*{微／大／竊}笑別人”與“對別人{微／大／竊}笑”）；口語並列式複合形容詞最容易重疊（比較“漂漂亮亮、老老實實、熱熱烈烈”與“*美美麗麗、*忠忠厚厚”），而其他如偏正式（如“*小小康康、*熱熱心心”〔註6〕）、述補式（如“*鎮鎮定定、*充充實實”）、述賓式（如“*順順耳耳、*醒醒目目”）與主謂式（如“*面面熟熟、*膽膽小小”）等複合形容詞則不易重疊〔註7〕；一般形容詞都可以受程度副詞“很、太、非常”等的修飾（如“很{漂亮／美麗／狂熱／順耳／面熟}”），但由名詞語素修飾形容詞語素而成的偏正式形容詞則不受程度副詞的修飾（如“(*很){雪白／漆黑／冰涼／火熱／筆直／水平}”〔註8〕）；一般雙音複合形容詞卻只能以“A A B B”的形式重疊（如“乾(乾)淨(淨)、簡(簡)單(單)、秀(秀)氣(氣)、和(和)氣(氣)”），但是表示貶義的雙音複合形容詞卻可以用“A A B B”或“A 裡 A B”的形式重疊（如“{骯骯髒髒／骯裡髒髒}、{囉囉唆唆／囉裡囉唆}”）甚或非用“A 裡 A B”的形式重疊不可（如“{*俗俗氣氣／俗裡俗氣}、{*土土氣氣／土裡土氣}”）。以上的觀察顯示：漢語複合詞的內部結構（如主謂式、述賓式、述補式、偏正式、並列式等）與其外部功能（如能否嵌入動貌、動相標誌或其他修飾成分、能否做及物或不及物動詞使用、能否重疊或受程度副詞修飾等）之間存在著相當密切的關係；連雙音形容詞的含有貶義與否也與其重疊形式具有密切的關連。因此，漢語複合詞的內部結構、貶義與褒義的區別以及與這些詞法與詞義屬性有關的句法功能都屬於應該儲入語料庫的重要語法資訊。

詞語的分斷，一方面牽涉到語(素)與詞(語)的分辨；另一方面則牽涉到詞(語)與詞組的區別。下面從「直觀推斷」(heuristics)的觀點，對於如何辨別「詞」與「非詞」（包括「語」與「詞組」）提出一些「經驗規律」(rule of thumb)做為檢驗詞與非詞的標準。不過，這裡應該注意：這些標準至多只能視為「充足條件」(sufficient condition)而

不能視為「必需條件」(necessary condition)；即滿足下列標準之一的句子成分就可以判定為詞、(黏著)語或詞組，但並不是所有的詞、語或詞組都非滿足所有有關的標準不可。而且，有些是肯定的認可條件，而有些則是否定的排斥條件。

(一) 凡是「譯音詞」(transliterated word)〔註9〕，不管音節多寡(如"沙發、伏特加、哀的美敦、布爾什維克")，而且包括「音義兼用詞」(hybrid loan-word；如"綠卡、卡片、吉普車、卡賓槍、迷你裙、蘋果派、哀的美敦書")，都一律視為獨立的詞，不能從中分斷〔註10〕。

(二) 凡是北京話裡讀「輕聲」(neutral tone)〔註11〕的語素，都不能獨立成詞，而必需與前面的語素合成詞。

(三) 如果在兩個或兩個以上的語素之間無法插入任何語素或表示停頓的語助詞"啊、呀、嚶"，而在這些語素的前後卻可以插入其他語素或表示停頓的語助詞，那麼這些語素就應該合成獨立的詞，不應該從中分斷。

(四) 在複合詞裡只能出現於詞尾位置的「後定位語素」(end-position morpheme)或「後綴」，如出現於名詞、形容詞、動詞等語素後面形成名詞的名詞語素(如"({樂/矮/亂/耳刮})子、({木/甜/看/上})頭(兒)、({桃/蓋/活/今/水餃})兒、({作/來/強/演奏})者、({畫/親/東/科學})家、({詩/情/心中})人、({醫/廚/工程})師、({演/店/觀察})員、({反/左/主流})派、({彈/黨/可能})性、({限/高/靈敏})度等)、出現於名詞、形容詞、動詞等語素後面形成「非謂形容詞」(non-predicative adjective)的名詞語素(如"({男/舊/美國/自由/填鴨})式、({大/重/流線})型、({大/小/四十})號、({上/特/劣})等、({高/超/各})級、({綠/茶/卡其})色、({良/假/陽/世界/建設})性"等)、出現於形容詞與名詞語素後面而形成「作格動詞」(ergative verb)〔註12〕的動詞語素"化"(如"({美/軟/簡單/理想/電氣/大眾})化")、出現於形容詞、動詞、名詞等語素後面形成副詞的副詞語素(如"({突/勃/慘/居/依/全})然、({偶/率/莞})爾、({似/近/斷/無須})乎、({幸/俄/反/時/繼})而、({尤/更/何})其"等)以及出現於動詞等語素後面形成介詞或連詞的動貌或動詞語素(如"({憑/本/爲/順/接})著、({憑/靠/爲})了、({關/至/基/由/甚至})於)〔註13〕等，都不是獨立的詞，而必須與出現於前面的語素合成詞。

(五) 在複合詞裡只能出現於詞首位置的「前定位語素」(head-position morpheme)或「前綴」，如出現於表示人或動物的名詞語素等前面而含有親暱、尊敬、輕視等感情色彩的"阿"(如"阿({爸/姆/姨/哥/寶/飛/狗/豬})")、"老"(如"老({爹/公/婆/師/板(娘)/總/鄉/百姓/虎/鼠/鷹/張/李/大/么/表})")與"小"(如"小({姐/娘子/老婆/狗/豬/張/王/強/華})")、出現於動詞與名詞語素等前面加強動詞意味的"打"(如"打({算/開/掃/敗/勝/歌/哈欠/瞌睡/噴嚏})")、出現於書面語單音動詞語素前面表示被動或"對我如何如何"的"見"(如"見({笑/諒/罪/怪})")、出現於單音動詞語素前面形成形容詞的"可"(如"可({愛/憐/惜/恨/惡/悲/恥/靠/取})")、出現於單音或雙音動詞語素前面形成形容詞的"好"與"難"(如"({好/難}){看/

聽／聞／吃／用／愛／處理／商量}〔註14〕、在文言或書面語詞彙裡出現於單音詞素前面形成連接副詞的"而"（如"而{且／況／今／後}"）與形成方位(副)詞、對等連詞、從屬連詞等的"以"（如"以{前／後／上／下／內／外／東／西／南／北／往／及／便／免／至／致}"）、出現於數詞前面表示序數的"第"與出現於數詞"一"到"十"的前面表示日初的"初"以及出現於"大"與數詞"二"到"九"的前面表示排行的"老"、出現於名詞或形容詞前面形成否定「非謂形容詞」（如"(*很)不{法／毛}"）、否定「謂述形容詞」（predicative adjective）（如"(*很)不{軌／力}"）、否定形容詞（如"(*很)不{利／錯／規則／道德／人道／經濟／禮貌}"）與副詞（如"(*很)不{日／時／意}"）的"不"、出現於形容詞或名詞語素前面形成否定非謂形容詞（如"非{凡／分／法／常}"）或名詞（如"非{金屬／導體／黨員／國員／主流}"）的"非"、以及出現於名詞或動詞前面形成非謂形容詞的"反"（如"反{革命／間諜／潛艇／飛彈／國民黨／台獨／核}"）等，都不是獨立的詞，而必須與出現於後面的語素合成詞。

（六）附加於動詞後面的動貌標誌（如"了、著、過"）與動相標誌（如"完、到、掉、好、住、著（ㄙㄠˇ）、了（ㄌㄠˊ）、得"）不宜視為獨立的詞，而方位補語動詞（如"上、下、進、出、過、開"）與趨向補語動詞（如"來、去"）也可以分析為與前面的動詞合成詞語單元〔註15〕。出現於可能補語（如"跳得高"）、描述補語（如"跳得很高"）與結果補語（如"跳得{精疲力盡／氣喘不已／連臉色都蒼白了}"）的"得"也可以分析為與前面的動詞合成詞語單元〔註16〕。又「動態」（actional）單音動詞與雙音動詞分別以"A(一)A"（如"看(一)看"）與"ABAB"（如"觀察觀察"）的形式重疊，並表示「短暫」（tentative）的動貌意義。由於凡是動態動詞都可以重疊，而且動詞的重疊只具動貌的「語法意義」（grammatical reading）而不改變其「詞彙意義」（lexical meaning）；因此，動詞的重疊式不必一一登記於語料庫中，而只要在有關動詞的詞項記載裡登記「動態」（actional）這個語法資訊即可。在分斷詞語的時後，動詞的重疊式"A(一)A"與"ABAB"也可以分斷成一個詞語單元。

（七）附加於形容詞後面的「修飾語標誌」（modification marker）"的"常讀成輕聲，並且在語音上變成前面形容詞的「依附成分」（clitic），所以"的"宜與前面的形容詞分斷成一個詞語單元〔註17〕。附加於形容詞後面的副詞標誌"地、的"（如"{{慢慢／快(快)樂(樂)／飛快／(很)迅速／(很(有))禮貌}{地／的}"與"兒"（限於北京話單音形容詞的重疊式後面，如"慢慢兒"）也讀成輕聲而依附於前一個語素〔註18〕，所以"地、的、兒"也宜與前面的形容詞一併分割成詞語單元。又抽象名詞與動詞"有"連用後可以受程度副詞的修飾（如"很有{禮貌／才華／意思}({的／地})"）。這個時候，"有"與後面的抽象名詞似乎可以分析為詞語單元而同受程度副詞的修飾或共同帶上修飾語標誌"的"與副詞標誌"地"。形容詞與動詞一樣可以重疊：單音形容詞以"AA(兒)"的形式重疊（如"好好(兒)"），而雙音形容詞則以"ABAB"（如"乾乾淨淨、骯髒髒、和和氣氣"）或"A裡AB"（如"骯裡骯髒、流裡流氣"）的形式重疊，並具有「加強程度」（degree intensification）〔註19〕與表示「主觀評價」（subjective evaluation）〔註20〕的作用。形容詞的重疊，與動詞的重疊一樣，屬於句法上的變化而不屬於詞法上的變化；因此，形容詞的重疊式不必一一登記於語料庫中，而只要在有關形容詞的詞項記載裡登記「能

重疊」(reduplicable)與「貶義」(pejorative)這兩種語法資訊即可〔註21〕。在分斷詞語的時候，形容詞的重疊式"A A (兒)"、"A A B B"、"A 裡 A B"、"A B B"也可以分斷成一個詞語單元。

(八)附加於指人名詞後面的複數標誌"們"是黏著語素，需依附於前一個語素出現，所以應該與前面的名詞語素分析為同一個詞語單元。漢語裡有不少表示「主事者」(Agent)或人物的後綴，如"者、家、人、師、員、士、生、手、漢、夫、丁、郎、屬、鬼、棍、迷、犯"等，而且有些後綴的孳生力相當強；這些後綴都應該與前面的動詞、名詞或形容詞語素分析為同一個詞語單元。重疊式複合名詞"星星、猩猩、爺爺、奶奶、太太、哥哥、姊姊、弟弟、妹妹、嫂嫂、伯伯、叔叔、舅舅、姑姑、嬸嬸、姥姥、公公、婆婆、寶寶"等非以重疊式的形式出現不可，而且第二個音節需讀輕聲；所以也以重疊式的形式儲存於語料庫中。"爸(爸)、爹(爹)、媽(媽)"〔註22〕可以重疊，也可以單用，就以附括號的形式儲存於語料庫中。除了以上屬於詞法的重疊以外，少數名詞還可以靠句法上的重疊來表示「全稱量化」(universal quantification)；例如，"人人、事事、家家戶戶"分別表示"每一個人、每一件事、每一家每一戶"。漢語的名詞很少單獨重疊來表示全稱，卻可以與另一個詞義相近、相對或相反的重疊式連用來表示全稱，如"男男女女、子子孫孫、老老少少、大大小小、世世代代、家家戶戶、星星點點、點點滴滴、形形色色、口口聲聲、鬼鬼怪怪、上上下下、層層疊疊、日日夜夜、朝朝夕夕、年年月月、年年歲歲、早早晚晚、時時刻刻"等〔註23〕。這些四字連用的重疊式，大都是已經固定化的成語，而且分開了就不能單獨使用；所以不妨以四字成語的形式儲存於語料庫中〔註24〕。

(九)漢語的副詞是相當「異質」(heterogeneous)的詞類，由形容詞轉類來充當副詞或狀語的，(如"{快(一點)／慢(慢地)／(很)小心(地)／快(快)樂(樂)地／很(有)禮貌地}走過來")似乎不必歸入副詞，而可以利用「詞彙冗贅規律」(lexical redundancy rule)來處理。不過，在不受程度副詞修飾之下，出現於(多半是單音)動詞的前面直接修飾這個(單音)動詞的單音形容詞(如"快跑、慢走、高喊、大叫")都是屬於少數形容詞的特異屬性，因而應該在有關形容詞的詞項記載裡註明〔註25〕。其他，如程度副詞(如"很、太、最、稍微、十分、更加、越發、非常、格外")、情狀副詞(如"私下、親自、自個兒、趕緊、盡量、竭力、大肆")、時間副詞(如"立刻、馬上、已經、曾經、剛(才)、一直、始終、曾經")、範圍副詞(如"都、也、又、再、只、僅(僅)、全(都)、總共")、否定副詞(如"不、沒、未、別")與語氣副詞(如"卻、倒、偏、也許、恐怕、莫非、難道、反正、究竟")等，大都屬於只能出現於句子中固定位置而不能隨意移位的「定位語素」(fixed-position morpheme)〔註26〕，而且很少可以單獨出現〔註27〕。因此，副詞雖然常與名詞、動詞、形容詞並列為「實詞」(content word)，但是在性質上相當接近「虛詞」(function word)；創造「新詞」(neologism)的機會也很少，幾近於「封閉類」(closed class)。形成副詞的後綴或詞尾，有"({忽／竟／當})然、({偶／率／莞})爾、({似／幾／斷})乎、({幸／反／時})而、({暫／姑／苟})且、({尤／更／何})其、({這／怎／多})麼、({特／忽／倏})地、({是／真格／(說)實在／(飛也)似})的"等；但是其中"然、爾、乎、而、且、其"等都由古漢語繼承下來，在現代漢語中已不具有孳生

力；"地"與單音形容詞的合成早見於唐代，而在現代漢語裡卻已失去孳生力〔註28〕；至於"麼"與"的"則雖然屬於現代口語詞彙，但"麼"不再具有孳生力，而"的"則在書面語裡被"地"取代或與"的"合而為一。又，副詞的重疊可以分為三種情形：(i)可重疊，亦可不重疊（如"常(常)、剛(剛)、漸(漸)、僅(僅)、時(時)"〔註29〕）；(ii)必須重疊，不能單用（如"單單、恰恰、足足、整整、通通、頻頻、匆匆、悄悄、徐徐、往往、微微"〔註30〕）；(iii)少數雙音副詞可以以"AAB"的形式重疊（如"茫(茫)然、貿(貿)然、飄(飄)然、渾(渾)然、幾(幾)乎"）〔註31〕。因此，這幾種不同的重疊情形也應該登記於語料庫中。

（十）漢語的代詞只有"你、我、他(她、牠、它、牠)、咱、您"〔註32〕幾個，並可以帶上複數標誌"們"。另外，"某某(人)、誰誰、彼此"以及文言或書面語說法的"余、{吾／汝／彼}"等也可以包括在代詞。代詞是典型的封閉類虛詞，非但數目少得可以列舉，而且創造新詞的可能也微乎其微。

（十一）漢語的介詞大都由動詞演變而來，並依其型態可以分為四類：(i)單音(單純)介詞（如"對、從、到、跟、被、由、把、將、在、照、憑、靠"等）；(ii)含有動貌標誌"了"或"著"的雙音(複合)介詞（如"{憑／靠／爲}了、{憑／靠／照／本／爲／對／朝／向／沿／順／接}著"）；(iii)含有"於"的雙音與三音介詞（如"關／對／至／基／由／瀕／(有)鑒／甚至}於"）；(iv)其它雙音介詞（如"(打)從、(直)到、有關"）。介詞是不能隨意移位或單獨出現的定位語素，而且屬於封閉類虛詞；因而可以在語料庫中列舉。不過，在實際分析裡，有時候介詞與動詞的界限並不十分清楚，特別是含有"了、著"的介詞(組)與連謂結構的第一個動詞組之間的區別相當微妙，必須分別討論介詞與動詞的句法功能並加以比較才能界定二者的區別。

（十二）漢語的連詞可以分為對等連詞（如"跟、和、同、與、及、以及、而、且、(不但...)而且、並(且)、或(者)、還是、但(是)"）與從屬連詞（如"如(果)、假使、即使、除非、既然、雖然、固然、與其(...不如)"），而且還可以包括連接副詞（如"那麼、所以、因而、從而、繼而、進而、然而、(不但...)反而、因此、否則、(與其...)不如、(越...)越...、((一)邊...) (一)邊...、(一面...)一面..."）。對等連詞主要出現於兩個並列詞組的中間；從屬連詞只能連接兩個(或兩個以上的)子句，並且出現於前半句的句首(或主語與謂語之間)擔任起接的作用；而連接副詞則出現於後半句的句首(或主語與謂語之間)發揮承接的作用。連詞也是不能隨意移位或單獨出現的定位語素，而且也屬於封閉類的虛詞，因而可以在語料庫中加以列舉。

（十三）句尾助詞或句終詞是典型的定位語素，而且是典型的封閉類虛詞；只能出現於句尾的位置，不能單獨出現或移位，而且可以在語料庫中全數列舉。句終詞依其所表示的情態或語氣的不同可以分為三類：(i)表示詢問（如"嗎、呢、麼；乎、歟、也、邪、耶、焉"）；(ii)表示祈使（如"罷、吧、啊、呀、哇；也、哉、來、乎"）；(iii)其他（如"啊、呵、呀、哩、呢、哇、哪、啦、咧；也、矣、乎、哉、邪、與(=歟)、來、夫、兮"）。句終詞的主要功能在於表達說話者的情態或語氣，所以常與情態與語氣副詞搭配使用；例如，表示反詰的"難道"只能與表示是非問句的"嗎"連用（如"你難道不去

嗎？")，而表示追究的"到底"則只能與表示正反或選擇問句的"呢"連用(如"你到底去不去呢？")。

(十四) 漢語的數詞包括：(i)「系數詞」(即"零、一、二(兩)、三...九")；(ii)「倍數詞」(即"十、百、千、萬、億、兆...")；(iii)「複數詞」(即由系數詞與倍數詞組合而成的數詞，因為可以無限大，所以可能無限長)；(iv)「概數詞」(如"好些、若干、好幾(下面可接倍數)"，或在倍數後面加上"來、多、好幾")。其中，系數詞、倍數詞與概數詞是有限的，因而可以在語料庫中全數列舉；而複數詞卻是無限的，所以無法也無須在語料庫中一一列舉〔註33〕。至於序數，則只要在基數的前面加"第"即可。從詞語分斷的觀點而言，似乎可以把整個複數詞(尤其是不包含倍數的說法，如"一九九四"與"三二九、五六四")當做一個詞語單元來處理。又，數詞的重疊，除了"一一"表示"逐一、全部"以及"三三兩兩"、"三三五五"與"千千萬萬"等熟語必須儲入語料庫以外，其他數詞都很少重疊。

(十五) 漢語的量詞包括「名量詞」與「動量詞」：名量詞常出現於數詞之後，名詞之前(即"數詞—(名詞)")；而動量詞則多出現於動詞(以及其實語名詞)與數詞之後(即"動詞(名詞)數詞—")〔註34〕。名量詞主要包括(i)「個體量詞」(如"本(書)、張(床)、頭(牛)、匹(馬)")、(ii)「集合量詞」(如"雙(筷子)、對(夫妻)、胎(小貓)、套(家具)、群(人)、批(貨)、系列(問題)")、(iii)「度量詞」(如"尺、寸、斤、兩、斗、升、里、畝、公斤、海里")、(iv)「借用(名)量詞」(如"碗(飯)、口袋(麵)；一桌子(菜)、一屋子(人)")、(v)「準用量詞」(如"(兩)縣(人口)、(三)站(路)、(一)世紀(時間)")與(vi)「不定量詞」(如"點兒、些")；而動量詞則主要包括(i)「動量詞」(如"(洗—)下、(走—)趟、(看兩)次、(叫幾)聲、(念幾)遍")與(ii)「借用動量詞」(如"(放—)槍、(切—)刀、(抽幾)鞭、(看幾)眼、(踢—)腳、(打了他兩個)耳光")。數詞屬於「開放類」(open class)而可能無限多，並且可以單獨出現；但是量詞則屬於封閉類而可以在語料庫中列舉，並且不能單獨出現。量詞的重疊多見於個體與集合量詞(如"個個、條條、張張、頭頭、件件、雙雙、對對")，並且含有"每一(個)"的全稱意義。量詞與名詞之間的連用限制相當「任意而武斷」(arbitrary)，無法利用一般性的規律推斷出來〔註35〕，只得在語料庫中為每一個可數名詞加注。

(十六) 針對各種詞類〔註36〕，討論了如何分斷詞語以及哪一些語法資訊應該在語料庫中登記之後，我們想對各種詞類的詞語可能包含音節數目的多寡提出一些觀察，以供分斷詞語的參考。漢語的名詞在理論上可以無限長，在實際上也可以相當長(如"(那一位)台灣省教育廳廳長辦公室主任秘書助理")。同時，應該注意北京話的「指人名詞」(human noun)，除了"人、爹、娘、爸、媽"等少數是單音節以外，其他都是雙音節或多音節。漢語的動詞絕大多數是單音節與雙音節，三音節的只有含有後綴"化"的三音動詞(如"電氣化、機械化、現代化")與譯音動詞"馬殺雞"〔註37〕以及含有動相標誌"到"的三音動詞(如"意識到、體會到、感覺到")而已。漢語的形容詞絕大多數也是單音節與雙音節，多音節的只有極少數的譯音形容詞(如"阿沙力、羅曼蒂克")，而且似乎以四音節為限。不過，雙音形容詞經過重疊後也變成三音節或四音節(如"甜蜜蜜、氣呼

呼、高高興興、糊裡糊塗")，而三音節與四音節形容詞在加上修飾語標誌"的"後就變成四音節或五音節。因此，漢語形容詞的音節數目似以五音節為限〔註38〕。漢語的副詞絕大多數也是單音節與雙音節，但是經過重疊或帶上"乎、似地、般地"後就可能變成三、四個(甚至更多的)音節(例如"無須爭、貿貿然、幾幾乎、多多少少(地)；匆匆忙忙(地)、(很)有禮貌地(前兩例來自形容詞的轉類)；{飛也／(像)著了魔}似地、{風馳電掣／發了羊瘋癲}般地〔註39〕")。漢語的代詞也以單音節與雙音節為限(如"我(們)、他(們)、某某、誰誰")，唯一能找到的三音節代詞是"某某人"；但是我們也可以說"某某{先生/太太/小姐/老師/學生}"等，所以"某某人"也可能分析為名詞組。介詞與連詞也多屬單音節與雙音節，屬於三音節的只有"甚至於、以{致／至}於"等極少數的例詞。數詞與名詞一樣，理論上可以無限長；量詞限於單音節或雙音節；而句終詞則多屬單音節，但有"罷了、來著"等少數雙音節的例外。

以上討論詞語的分斷，特別是如何區別「詞」與小於詞的「(黏著)語素」，以下談如何分辨「複合詞」與大於詞的「詞組」。

(一) 凡是移位可的句子成分都是詞組，包括名詞組、介詞組與子句等。

(二) 凡是可以疑問詞"誰"(指人名詞組)、“什麼”(名詞組或子句)、“怎麼樣”(動詞組與形容詞組)、“做什麼”(動詞組)、“怎麼”(情狀與理由副詞組，包括介詞組與從屬子句)、“為什麼”(原因介詞組與從屬子句)、“什麼時候”(時間副詞組，包括詞組與從屬子句)、“什麼地方、哪裡”(處所副詞組，包括名詞組與從屬子句)與“哪幾”(數量詞組)等取代的句子成分都是詞組。

(三) 凡是含有修飾語標誌"的"的詞組都屬於形容詞組、領位名詞組、同位子句或關係子句。如果這個詞組已經含有程度副詞或者可以受程度副詞的修飾(如“(很)快樂的(小孩子)”)，那麼這個詞組就是形容詞組。如果"的"的前面是名詞組，而"的"的後面也是名詞(節)(如“那一個小孩子的(父親)”)，那麼"的"與前面的名詞組合成領位名詞組。如果"的"前面的詞組是含有主語與謂語的子句(如“老張去逝的(消息)”)，那麼這個詞組是同位子句；而如果子句裡含有「空缺」(gap；常用符號"e"來代表，如"e在外面等你的(孩子)、你喜歡吃e的(水果)”)，那麼這個詞組是關係子句。

(四) 凡是含有副詞標誌"地"的詞組都屬於廣義的副詞組。如果這個詞組已經含有程度副詞(包括形容詞的重疊)或者可以受程度副詞的修飾，(如“(很)快樂地、快快樂樂地、(很)有禮貌地”)，那麼這個副詞組是來自形容詞的轉類〔註40〕。如果這個詞組含有介詞(如“像瘋子{似／般}地”)，那麼這個副詞組含有介詞組。如果這個詞組含有動詞(如“{飛也／發了瘋}似地”)，那麼這個副詞組含有以空號代詞為主語的從屬子句。其他詞組(如“忽地、不時地”)是一般的副詞組。

(五) 凡是含有限定詞(如“這、那、每、哪”)的詞組都是名詞組或省去主要語名詞的限定數量詞組(如“這一個(孩子)”)；不含有限定詞但含有數量詞的詞組是名詞組或省去主要語名詞的數量詞組(如“一個(孩子)”)。

(六) 凡是含有程度副詞或可以受程度副詞修飾的詞組都是形容詞組。形容詞組可以經

過轉類而充當副詞組或狀語。

(七) 介詞與後面的名詞組合成介詞組，而從屬連詞則與後面的子句合成連詞組或從屬子句。介詞組與連詞組可以概化而合併成同一詞組。

(八) 含有修飾語標誌"的"的詞組(包括形容詞組、領位名詞組、同位子句與關係子句)與後面的名詞(節)合成名詞組。

(九) 句終詞與前面的子句合成句子。

(十) 動詞(包含動相與動貌標誌在內)連同後面的賓語名詞組與補語名詞組或介詞組合成動詞組。

三、詞類的畫分：屬性分析與次類畫分

湯(1990b)〈漢語的詞類：畫分的依據與功用〉曾經針對漢語詞類的分析與界定提出下列幾個觀點。

(一) 詞類是漢語裡客觀存在的東西，必須依據客觀的標準來畫分，個別具體的詞語也必須依照這些客觀的標準來決定其歸類。因此，在建立漢語詞類體系的過程中，必須重視漢語的語言事實與語法特徵，從照實觀察的「描述主義」(descriptism)與憑藉證據的「客觀主義」(objectivism)的觀點來從事詞類分析。

(二) 漢語詞類的畫分非但需要客觀明確的標準，而且還應該在語法、語意或語用上具有相當的「詮釋功效」(explanatory power)，即必須能從分類中獲得某些「在語言學上具有重要意義的條理化」(linguistically significant generalization)。例如，漢語名詞的「次類畫分」(subcategorization)應力求與量詞的「選擇限制」(selection restriction)配合；充當述語的動詞可以有及物與不及物之分，那麼充當述語的形容詞也可以有及物與不及物之分；介詞與名詞組的連用相當於及物動詞與賓語名詞組的關係，而從屬連詞與子句的連用則相當於及物動詞與賓語子句的關係等等。

(三) 漢語詞類的畫分，不但要兼顧語意內涵、形態特徵與句法功能，而且還要參酌傳統漢語語法的研究成果與歐美語法理論的最近發展來檢驗畫分的依據與功用。

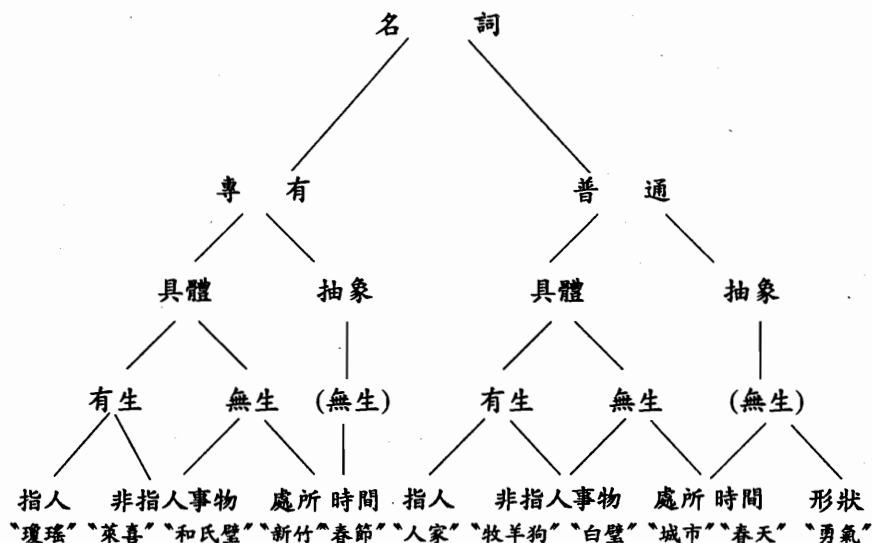
(四) 漢語裡某些詞類之間的界限並不一定「清晰而確定」(clear-cut)，而可能有些「模糊而曖昧」(fuzzy)的地方。歸屬於同一詞類的詞語，有些是比較「典型」(prototypical)或「無標」(unmarked)的詞語；而有些則是比較「周邊」(peripheral)或「有標」(marked)的詞語。因此，各種詞類的畫分標準最好能依其重要性或優先次序由上而往下地予以列舉；越能滿足多項標準或越能符合上面標準的詞語，是越典型或越無標的詞語；反之，只能滿足少數標準或只能符合下面標準的詞語，是比較周邊或比較有標的詞語。例如，一般典型而無標的形容詞都具有陳述功能而可以出現於主語的後面充當述語，也具有限制功能而可以出現於名詞的前面充當定語；並且，可以受程度副詞「很、最、太」等的修飾，也可以照「A A (兒)」或「A A B B」的形式重疊。但是，「真、假、錯」等形容詞，則既不能受程度副詞的修飾，又不能重疊，連充當述語的時候都必須出現於「是...的」句式；

因此，是比較周邊而有標的形容詞。又如，一般典型而無標的動詞都屬於「動態動詞」(actional verb)而可以帶上「了、過、著」等動貌標誌，可以用「A(一)A」或「A B A B」的形式重疊，可以出現於祈使句或充當「兼語動詞」(pivotal verb;如「勸、逼、強迫」等)的補語子句述語。但是，「死亡、去逝、變化、屬於、獲得、得到、喪失」等「靜態動詞」(stative verb)，則不容易帶上「著、起來、下去」等動貌標誌，也無法重疊、或出現於祈使句、或充當謂語動詞的補語子句述語；因此，是比較周邊或比較有標的動詞。再如，一般名詞都表示實體而具有指涉個體的功能，並且可以與數量詞連用。但是，由名詞語素與量詞語素合成的集合名詞「人類、馬匹、船隻、車輛、布匹」等，卻不能與數量詞連用；因此，也可以說是比較特殊或例外的名詞。

(五) 一個詞語可能屬於兩個或兩個以上的詞類，漢語裡這種「兼類」或「跨類」的現象，包括「一詞多用」、「一詞多類」與「詞類活用」三種不同的情形，而且應該做三種不同的處理。例如，「魚是死的」與「魚死了」裡的「死」分別是動詞的「狀態」(stative)與「起動」(inchoative)用法；但是「我們不能把規律定得太死」與「他死緊地抓住了我的手」裡的「死」分別是形容詞與(程度)副詞用法。「死」的狀態用法與起動用法是許多動詞(如「魚是活的；魚活了」、「門是{開/關/鎖/壞}的；門{開/關/鎖/壞}了»)共有的用法，是屬於「一詞多用」的情形，不必在詞項記載裡特別注明。另一方面，「死」的形容詞用法與程度副詞用法是「死」這個詞的特殊用法，屬於「一詞多類」的情形，應該在詞項記載裡特別注明。又如，在「他用一把特別訂製的鎖把珠寶箱牢牢地鎖住了」裡，前面與數量詞連用的「鎖」是名詞用法，而後面帶上動相標誌「住」與動貌標誌「了」的「鎖」是動詞用法。因為只有少數的詞可以同時兼有與數量詞連用的名詞用法與帶上動相以及動貌標誌的動詞用法；所以這是「一詞多類」的情形，應該在「鎖」的詞項記載裡特別注明。另一方面，在「我愛他、疼他」與「打是愛、罵是疼」裡，出現於第一句話的「愛」與「疼」是充當述語的純粹動詞用法，而出現於第二句話的「愛」與「疼」是充當補語的「名物性」(nominal)用法。由於絕大多數的動詞都可以出現於主語(如「打是愛、罵是疼」、「打罵孩子也沒用」)、賓語(如「不聽我的勸，結果挨了(一頓)罵」)或補語的位置來充當類似「體詞」(substantive)的名物性用法〔註41〕；所以這是屬於「一詞多用」的情形，不必在所有動詞的詞項記載裡一一注明。同樣地，幾乎所有的形容詞都有狀態用法(如「她的眼睛很大」)、起動用法(如「她的眼睛大了起來」)、陳述用法(如「她的眼睛太大的」)與副詞用法(如「我們應該太大地獎勵她一番」)；所以可以視為「一詞多用」來處理，不必在所有形容詞的詞項記載裡一一注明。另一方面，只有少數形容詞(如「硬、直、挺、彎、豐富、壯大、充實」)可以有使動(及物)用法(如「硬著頭皮、直著脖子、豐富你的人生、壯大我們的聲勢」)，而且形容詞用法可以受程度副詞的修飾，而使動及物用法則不受程度副詞的修飾；因此，必須視為「一詞多類」來處理，在個別形容詞的詞項記載裡特別加以注明。至於「詞類活用」，指的是基於「比照類推」(analogy)或「比喻引伸」(metaphoric extension)所做的臨時轉用。例如，在「他這個人很黃牛」、「她寫的小說很瓊瑤」與「我是喝黃酒的，可是如果你們一定要喝白干，我也可以白干一下」的例句裡，名詞「黃牛」與專名「瓊瑤」的形容詞用法以及名詞「白干」的動詞用法，並不屬於真正的「一詞多類」，而是臨時起意而做的「詞類活用」。有些詞的「一詞多類」可能

是由「詞類活用」中發展出來的，並且已經為一般大眾所接受（如“孩子是我的心肝寶貝；不要太寶貝孩子；他這個人很寶貝”與“他這個人一點也不懂幽默；他的談吐很幽默；他幽了我們一默”裡“寶貝”與“幽默”的名詞、動詞與形容詞用法），理應在詞項記載裡注明。但是尚未為一般大眾所接受的「詞類活用」，則視為臨時的借用或轉用，不必在詞項記載裡增添新的詞類。

（六）詞類體系不但具有一定的「階層組織」(hierarchical structure)，而且還可能「交叉分類」(cross-classifying)。詞類體系的階層組織可以用「上下支配關係」(dominance)的「樹狀結構」(tree structure)來表示；而交叉分類則可以對詞更進一步做「屬性分析」(feature analysis)來處理。例如，名詞可以分為「專有名詞」與「普通名詞」，也可以分為「具體名詞」與「抽象名詞」，還可以分為「有生名詞」與「無生名詞」，更可以分為「指人名詞」與「非指人名詞」、以及「事物名詞」、「處所名詞」、「時間名詞」、「性狀名詞」等。但是這些名詞並非對等或並立存在的，而是各類名詞之間有一定的上下支配關係。這種上下支配關係可以用下面的「樹狀圖」(tree diagram)來表示：



從上面的樹狀圖可以看出：在名詞的次類之間存在著一定的上下支配關係。例如，專有名詞支配具體與抽象名詞，具體名詞支配有生名詞與無生名詞，抽象名詞支配時間名詞（與性狀名詞），而有生名詞與無生名詞則分別支配指人與非指人名詞以及事物與處所名詞〔註42〕。從這種名詞次類之間的支配關係，可以推斷：有生名詞必然是具體名詞，指人名詞必然是有生而且是具體名詞，而事物、處所與時間名詞必然是無生名詞。另一方面，具體與抽象、有生與無生、指人與非指人的區別，以及事物、處所、時間與性狀〔註43〕等次類，不但出現於專有名詞底下的次類畫分，而且也出現於普通名詞底下的次類畫分。也就是說，這些名詞的次類有一大部分是交叉分類的。這些事實顯示：詞類的次類畫分不能完全依賴階層組織，而必須兼採屬性分析。凡是名詞都具有“[+名詞]”這個句法屬性，並依據其固有的語意內涵含有“[±抽象]”、“[±有生]”、“[±指

人]"、"[±事物]"、"[±時間]"等語意屬性。如此，各種詞類不再視為「無法分析的整體」(unanalyzable entity)，而應該更進一步分析為由許多屬性匯集而成的「屬性匯合體」(feature matrix)。這樣的屬性分析不但可以妥善處理次類之間交叉分類的問題，而且還可以利用屬性之間的上下支配關係與「詞彙冗贅」(lexical redundancy)關係來簡化詞項記載裡面的「屬性標示」(feature specification)。例如，只要擬定"[±指人]→[+有生]"、"[+有生]→[-抽象]"、"[±事物]→[-有生]"、"[±時間]→[-有生]"、"[-有生]→[-指人]"這類「詞彙冗贅規律」，那麼在詞項記載裡"瓊瑤"只需標示"[+專有，+指人]"、"萊喜"只需標示"[+專有，-指人]"、"和氏璧"只需標示"[+專有，+事物]"、"新竹"只需標示"[+專有，-事物]"、"春節"只需標示"[+專有，+時間]"即可，而與這些專有名詞相對應的普通名詞如"作家"、"牧羊狗"、"白璧"、"城市"、"春天"等則只需把有關詞項記載裡的"[+專有]"改為"[-專有]"即可，不必把所有的屬性都一一加以注明。

根據以上觀點，湯(1990b)針對漢語「實詞」與「虛詞」的分辨、「體詞」(substantive)與「謂詞」(predicative)的區別，以及名詞的句法功能與次類畫分，提出相當詳細的討論與分析。湯(1988b)〈為漢語動詞試定界說〉，也從同樣的觀點討論漢語動詞的句法功能與次類畫分。湯(1991a)〈從動詞的「論旨網格」談英漢對比分析〉、(1992a)〈語法理論與機器翻譯：原則參數語法〉與(1994b) "A Minimalist Approach to Contrastive Analysis of English, Chinese and Japanese" 更利用「論旨網格」(theta-grid)的概念把漢語動詞、形容詞、副詞、介詞、連詞等的「論元結構」(argument structure)、「論旨屬性」(thematic property)與「投射條件」等加以登記，並與此相對應的英語與日語的比較對照下，做了極為詳盡的討論與分析。而湯(1993) "On the Relation Between Word-Syntax and Sentence-Syntax: A Case Study in Chinese Compound Verbs" 與(1994a) "More on the Relation Between Word-Syntax and Sentence-Syntax: A Case Study in Chinese Compound Nouns" 則分別就漢語複合動詞與名詞的內部結構、外部功能、內部結構與外部功能的關係、複合動詞與動詞組(以及句子)之間的結構關係、複合名詞與名詞組之間的結構關係提出相當周全而深入的研究。由於篇幅的限制，我們無法把這些文章的要點一一加以介紹。

四、結語

漢語語料庫的建立是重要而有意義的工作。為了建立有意義且有價值的漢語語料庫，我們必須深入研究漢語語法(包括句法與詞法)，才能真正了解理想的語料庫應該提供那些語法資訊。我們也必須承認國內的語法專家至今尚未建立一套妥善而完備的漢語詞類體系〔註44〕，連詞類的名稱都無從統一，畫分的依據都無法提出，更遑論如何決定或登記詞類以外的其他語法資訊。就這一點意義而言，漢語語料庫的建立與研究工作只是剛起步而已，前面還有一段崎嶇而漫長的道路要走。

腳註：

註1：這裡所謂的「詞性」，包括狹義的「詞類」(lexical category)以及其他各種「詞彙屬性」(lexical feature)與「語法資訊」(grammatical information)。

註2：有些語法學家，更進一步區別「複合詞」與「合成詞」(complex word)。二者的區別在於：「複合詞」由「詞根」(root)與詞根(或「詞幹」(stem)與「詞幹」)合成，而「合成詞」則由「詞根」(或「詞幹」)與「詞綴」(affix)合成。在詞法學上，「詞根」與「詞幹」的區別在於：把所有的「詞綴」或附加成分去掉後留下來的成分，叫做「詞根」；而把任何詞綴或附加成分去掉後留下來的成分，叫做「詞幹」(stem)。因此，詞根不可能包含詞綴或其他成分，而詞幹則可能包含詞綴或其他成分。例如，"可塑性"這個複合名詞是以動詞語素"塑"為詞根，加上表示形容詞的「前綴」(prefix；又稱「前加成分」或「詞首」)"可"來形成形容詞詞幹"可塑"後，再加上表示名詞的「後綴」(suffix；又稱「後加成分」或「詞尾」)"性"來形成名詞的。又如"理性化"這個複合動詞是以名詞語素"理"為詞根，加上表示名詞的後綴"性"來形成名詞詞幹"理性"後，再加上表示動詞的後綴"化"來形成動詞的。由於漢語裡「自由語(素)」(free morph(eme))與「粘著語(素)」(bound morph(eme))的界限並不十分清楚，而二者的區別在漢語詞法學上的意義也並不十分重要；所以詞根、詞幹、詞綴的區別，以及複合詞與合成詞的區別，也就不清楚或不重要。因此，在底下的討論裡，我們不區別詞根、詞幹與詞綴而一律稱為「詞幹」，也不區別複合詞與合成詞而一律稱為「複合詞」。

註3：關於漢語單純詞、合成詞與複合詞更詳細的分類與討論，參湯(1989-1992)〈漢語的「字」、「詞」、「語」與「素」〉，收錄於湯(1992：1~57)。

註4：參呂(1979：18-19)。

註5：含有動貌標誌與動相標誌的述賓式複合動詞，究竟應該分析為詞組（即由述語動詞與賓語名詞組合成的動詞組）抑或詞語（即由動詞語素與名詞語素合成的述賓式複合動詞），學者間仍有異論。有些學者（如朱(1984)）主張含有動貌或動相標誌的時候即分析為動詞組，而不含有動貌或動相標誌的時候則分析為複合動詞。另有些學者（參陸等(1975)）則提議把這些複合動詞視為允許特定語素介入的「離合詞」(separable word)。我們的看法是：把「可以離合」(separable)這個句法屬性做為語法資訊登記於語料庫裡。我們的理由是：(i) 可以離合的複合動詞不限於由自由語素與自由語素合成的述賓式複合動詞，而包括由粘著語素與粘著語素合成的述賓式（如"鞠(了)躬"）、並列式（如"游(了)泳"）、偏正式（如"體(了一堂)操"，以及由自由語素與粘著語素合成的"小(好)便"、大(完)便"）複合詞（不限於動詞，但是離合的時候則屬於動詞），甚至包括譯音詞（如"幽(他一)默"）；(ii) 可以介入的語素也不限於動貌或動相標誌，而包括形容詞（如"生(了很大的)氣"）、領位名詞組（如"生(了丈夫的)氣"）、數量詞（如"生(過三次)氣"）等修飾成分；因此，複合詞的能否離合應該視為這些複合詞（包括譯音詞）的「特異屬性」(idiosyncratic property)而記載於語料庫裡。至於含有動貌或動相標誌以及修飾成分的離合詞究應分析為詞語或詞組的問題，則根據下面所討論的檢驗詞語與詞組的標準來決定。

註6：但有"正正式式、大大方方、扎扎實實"等少數例外。

註7：關於漢語形容詞重疊規律的討論，參湯(1982)〈國語形容詞的重疊規律〉。

註8：這是由於這些偏正式複合形容詞的名詞語素具有程度副詞的功能（如"雪一般的白

”)的緣故。同樣的情形也發生於由動詞語素修飾形容詞語素而成的偏正式形容詞（如“(很){飛快／滾熱／刺癢／垂直}”）以及由形容詞語素修飾形容詞語素而成的偏正式複合形容詞（如“(很){暴富／窮酸／酸痛／小康／鮮／暗／朱／紫}紅”）。但是這些偏正式複合形容詞也有可以受程度副詞修飾的少數例外（如“(很){天真／膚淺／沈悶／鎮定／逼真／扎實／狂熱}”）；在這些例外裡，有些修飾語素似乎不具有程度副詞的功能（如“天真、膚淺”），而有些偏正式則似乎亦可分析為並列式（如“沈悶、鎮定、狂熱”）。又注意，不能受程度副詞修飾的複合形容詞都不能重疊。

註9：包括來自古漢語的「聯綿詞」（包括「雙聲詞」(alliterating word)、「疊韻詞」(rhyming word)與「外來詞」(loan word)，如“玻璃、葡萄、葫蘆、蟋蟀、蜘蛛、蝴蝶、蜻蜓、螳螂、蝙蝠、法瑯、檸檬、琵琶、佛陀、菩薩、涅槃、般若、檀那（以上係名詞）；蹉跎、叮嚀、逍遙、躊躇、匍匐、徜徉、彷徨、瀰漫、徘徊、抖擻、嗚咽、嘀咕（以上係動詞）；爛漫、窈窕、蒼茫、堂皇、玲瓏、參差、崎嶇、伶俐、慷慨、忸怩、惆悵、唐突、苗條、朦朧（以上係形容詞）；依稀、彷彿、須臾、剎那、婆娑（以上係副詞）”等。

註10：這些不能從中分斷的譯音詞，包括音譯兼顧的“浪漫、甜不辣、俱樂部”等。參湯(1989a)〈新詞創造與漢語詞法〉。

註11：北京話的輕聲包括「自然輕聲」與「規定輕聲」：前者包括動貌標誌“了、著、過”，複數標誌“們”與名詞後綴“子、頭、兒”等；而後者則指依照北京一帶本地人的讀法並由詞典來規範的輕聲（如“學生、影片、東西”等）。那些語素該讀自然輕聲，本地人之間的意見相當一致；但是那些語素該讀規定輕聲，則本地人之間常因年齡、居住地區、生活背景等的不同而常有出入。

註12：作格動詞可以有「使動及物」(causative-transitive；如(i)句)與「始動不及物」(inchoative-intransitive；如(ii)句)兩種用法：

- (i) 我們要綠化我們的校園。
- (ii) 我們的校園綠化了。

註13：這裡的“著”與“了”在歷史演變上似乎來自動貌標誌，而且是來自「連謂結構」(serial-VP construction)中附加於前項單音動詞的動貌標誌；但是在介詞用法裡動貌意義已經虛化，不再具有完成貌或持續貌意義。

註14：我們把出現於雙音動詞的“{好／難}{處理／商量}”與“{容易／難以}{處理／商量}”兩相對照之下視為複合形容詞。同時注意，並不是所有雙音動詞都可以與“好、難”連用之後受程度副詞的修飾。

註15：例如，“跳上來”與“跳上(車子)”都可以分斷成一個詞語單元。比較有問題的是，方位補語動詞與趨向補語動詞被賓語名詞分割的時候，如“跳上(車子)來”。我們主張這個時候也比照“跳上來”而分析為「離合詞」或「非連串詞」(discontinuous word)；與漢語的“跳上(車子)來”相對應的英語“jump up (into the car)”與日語“(車に)飛び乗る”也分別形成「雙詞動詞」(two-word verb；又稱 verb-particle combination)與複合動詞。又出現於動詞

前面的趨向動詞（如“{來／去}（讀書）”）則比照一般動貌動詞（如“{在／（沒）有／開始／繼續}（讀書）”）與情態動詞（如“{要／想／會／能／夠／應該／可以／願意}（讀書）”）分析為獨立的詞。

註16：在語音形態上這裡的“得”（以及方位與趨向補語動詞）都常讀成輕聲而變成「依前成分」((en)clitic)；而且，在詞法分析上也有證據來支持“得”與前面的動詞合成詞語單元。參湯（1990a）〈漢語動詞組補語的句法結構與語意功能：北平話與閩南語的比較分析〉、（1991a）“The Syntax and Semantics of Resultative Complements in Chinese: A Comparative Study of Mandarin and Southern Min”與（1991b）〈漢語述補式複合動詞的結構、功能與起源〉。

註17：附加於名詞後面充當「領位標誌」(gentive marker)的“的”也屬於廣義的修飾語標誌，所以似乎也應該分析與前面的名詞(組)：形成詞組單元。附加於子句（句括以「空號代詞」(empty pronoun)為主語的動詞組)的修飾語標誌“的”的分斷方法，則比較棘手；因為這個時候的“的”讀成輕聲並依附於前一個語素，但並不與這一個語素形成詞語單元，而是與整個子句形成詞組單元。

註18：“兒”甚至於與前一個語素形成同一個音節而稱為「兒化韻」。

註19：因而形容詞的重疊式不能再受程度副詞的修飾。

註20：關於漢語形容詞的重疊意義與重疊規律，參湯（1982）。同時注意：閩南語與客家語的單音形容詞不單可以「重疊」，還可以「三疊」（即「A A A」；如“金金金、芳芳芳”）。

註21：有些形容詞只有“A B B”的重疊式，卻沒有“A A B B”的重疊式（如“年輕(輕)、火辣(辣)”）；而有些形容詞則只有“A B B”的重疊式，卻沒有相對應的“A B”形式的形容詞，因而也沒有“A A B B”的重疊式（如“酸溜溜、亂哄哄、熱呼呼、白淨淨、綠油油、黑漆漆、嬌滴滴、懶洋洋”）。這些“A B B”的重疊式無法由一般的規律衍生，只好把整個重疊式做為詞語儲存於語料庫中。註22：近年來台灣的孩子在稱呼中也開始使用“哥(哥)”與“姊(姊)”的單音用法。

註23：例詞中包括名詞以外語素（如形容詞的“大、小、早、晚”）的重疊，而這些四字連用的重疊式並不一定充當名詞使用，也有可以充當副詞或狀語使用的。

註24：但是我們仍能模仿這個形式而創造新的重疊連用詞，如“山山嶺嶺、山山谷谷、草草木木”等，而一經使用並且獲得眾人的認可之後即可儲入語料庫中。

註25：這些單音形容詞與單音動詞的有限連用，甚至可以做為詞語單元來處理。不過，從機器翻譯的觀點而言，仍以兩個詞語來處理為妥。

註26：只有語氣副詞中表示「情態」(modality)的少數副詞（如“也許、或許、恐怕、大概、反正、難道”）可以出現於句首與句中（即主語與謂語之間）兩種位置。

註27：單音副詞，除了表示否定的“不”以外，幾乎都不能單獨出現；雙音副詞也除了部

分表程度、時間與範圍的副詞以外很少單獨出現。

註28："地"（又做"的"）與雙音或多音形容詞的連用則在現代漢語裡仍然具有旺盛的孳生力；不過，我們把形容詞與"地"的連用（以及北京話裡單音形容詞的重疊與"兒"的連用）視為形容詞的轉類，用"形容詞+'地'→副詞"（與"重疊單音形容詞+'兒'→副詞"）的詞彙冗贅規律來統一處理。

註29：單用的"漸、僅、時"屬於書面語詞彙。

註30：另有"滔滔(不絕)、咄咄(逼人)、循循(善誘)、歷歷(在目)、紛紛(響應)、唯唯(聽命)、冉冉(上升)、炯炯(有神)"等，但是大都出現於書面語裡，很少在口語裡使用。

註31：只有極少數的雙音副詞（如"多少"）可以疊成"A A B B"的形式（如"多多少少"）。

註32："您"是北京話裡"你"的尊稱，並有"您哪(= 您)"的說法。北京話裡另有"他"的尊稱"您"，但在台灣卻很少使用。另外，曾經有人主張用"伊"來取代"她"，但是現在已不流行。又"您、您、伊"似乎很少與"們"連用。註33：漢語的複數詞似乎可以用下面(i)的「規律母式」(rule schema)來衍生，並用(ii)的規律來處理"零"的出現：

(i) 複數詞→(((系數詞)千)((系數詞)百)(系數詞)兆)

((((系數詞)千)((系數詞)百)(系數詞){億/萬萬}))

((((系數詞)千)((系數詞)百)(系數詞)萬))

((((系數詞)千)((系數詞)百)(系數詞)十))

(系數詞)

(ii) a. 零倍數詞→零

b. 倍數詞零→倍數詞

c. 零ⁿ→零ⁿ

註34：但是也有"進了一趟城"與"一趟城也沒進過"這樣的說法，參朱(1980：51)。

註35：我們只能找出一些相當籠統而無法周全的經驗規律：如(i)細長的東西用"枝"；(ii)有延長的平面用"張"；(iii)小而圓的東西用"顆"；(iv)有把的東西用"把"等。

註36：由於篇幅的限制，我們還有許多問題（如限定詞、疑問詞、感嘆詞等）無法在此討論。

註37：三音動詞在漢語裡是「有標」(marked)或例外而特殊的動詞；因此，"電氣化"與"馬殺雞"都無法出現於賓語名詞組的前面。試比較：

(i) a. ?* 政府準備電氣化農村。

b. 政府準備把農村電氣化。

(ii) a. ?* 她在馬殺雞客人。

b. 她在替客人馬殺雞。

相形之下，含有動相標誌或動貌標誌的三音或四音動詞（事實上是雙音動詞加上動貌標誌）卻可以出現於賓語名詞組的前面。

試比較: (iii) a. 如今他也體會到了親情的可貴。

b. 他的沈默意味著事態的嚴重。

註38：非謂形容詞所包含的音節數目，在理論上不受五音節的限制。但是事實上也以四或五音節為上限（如“非主流派、前蔣經國(時期)、後結構主義(作品)”），加了修飾語標誌“的”後就可能達到六個音節（如“後結構主義的作品”）。

註39：帶上“{似／般}地”的副詞與狀語，可能應該分析為詞組，因為與此相對應的說法在英語裡，都由介詞“like, as”引介的介詞組或由“as {if/ though}”引介的從屬子句來表達，而在日語裡則以“{名詞組/子句}の{如く/ように}”的句式來表達。

註40：「轉類」(conversion)又稱「零附加」(zero-affixation)，是指某一種詞類在不改變其語音形態之下去充當另一種詞類。修飾標誌的“的”與副詞標誌的“地”，是為了書寫文字上識別的方便而採取的權宜措施。許多人仍然把“地”寫成“的”，或讀成“的”。註41：動詞的名物性用法可以保留動貌標誌（如“去了比不去好”、“躺著比較舒服些”），甚至可以保留賓語、情態動詞、情狀副詞與否定副詞等（如“這本書的遲遲不能出版”）。

註42：這樣的分類基本上屬於依據語意畫分的「語意次類」(semantic subcategorization)或「義類」。但是如果把名詞分為「可數名詞」與「非可數名詞」之類，那就成為參酌句法功能的分類了。

註43：在專有抽象名詞底下仍然可以有「三位一體」(the Trinity)這樣的性狀名詞。

註44：因此，國內至今尚未出版一部標明詞類的漢語辭典。

參考文獻

Lu, Zhiwei (陸志韋) et al., 1975, 《漢語的構詞法(修訂本)》，中華書局。

Lu, Shuxiang (呂叔湘), 1979, 《漢語語法分析問題》，商務印書館。

Tang, Tingchi (湯廷池), 1979a, <國語句法中的重疊現象>, 《語文週刊》1559. 1560期, 並收錄於湯(1979b: 113-142)。

Tang, Tingchi (湯廷池), 1979b, 《國語語法研究論集》，台灣學生書局。

Tang, Tingchi (湯廷池), 1982, <國語形容詞的重疊規律>, 《師大學報》27期 279-294頁, 並收錄於湯(1988: 29-57)。

Tang, Tingchi (湯廷池), 1988a, 《漢語詞法句法論集》，台灣學生書局。

Tang, Tingchi (湯廷池), 1988b, 《為漢語動詞試定界說》，《清華學報》新18卷1期, 43-69頁, 並收錄於湯(1989b: 1-42)。

Tang, Tingchi (湯廷池), 1989a, <新詞創造與漢語詞法>, 《大陸雜誌》78卷4期5-91頁. 5期27-34頁, 並收錄於湯(1989b: 93-146)。

Tang, Tingchi (湯廷池), 1989b, 《漢語動法句法續集》，台灣學生書局。

- Tang, Tingchi (湯廷池), 1990a, <漢語動詞組補語的句法結構與語意功能：北平話與閩南話的比較分析>，收錄於湯 (1992c：1-93)。
- Tang, Tingchi (湯廷池), 1990b, <漢語的詞類：畫分的依據與功用>，《第三屆計算語言學研討會論文集》161-183頁，並收錄於湯 (1992b：59-92)。
- Tang, Tingchi (湯廷池), 1991a, "The Syntax and Semantics of Resultative Complement in Chinese：A Comparative Study of Mandarin and Southern Min", an invited paper presented at The Third North American Conference on Chinese Linguistics，並收錄於湯 (1992c：165-204)。
- Tang, Tingchi (湯廷池), 1991b, <漢語述補式複合動詞的結構、功能與起源>，第二屆中國境內語言暨語言學國際研討會發表論文，並收錄於湯 (1992c：95-164)。
- Tang, Tingchi (湯廷池), 1991c, <從動詞的「論旨網格」談英漢對比分析>，第三屆世界華語文教學研討會發表論文，並收錄於湯 (1992c：205-250)。
- Tang, Tingchi (湯廷池), 1991a, "The Syntax and Semantics of Resultative Complements in Chinese：A Comparative Study of Mandarin and Southern Min", an invited paper presented at The Third North American Conference on Chinese Linguistics，並收錄於湯 (1992c：165-204)。
- Tang, Tingchi (湯廷池), 1991b, <漢語述補式複合動詞的結構、功能與起源>，第二屆中國境內語言暨語言學國際研討會發表論文，並收錄於湯 (1992c：95-164)。
- Tang, Tingchi (湯廷池), 1991c, <從動詞的「論旨網格」談英漢對比分析>，第三屆世界華語文教學研討會發表論文，並收錄於湯 (1992c：205-250)。
- Tang, Tingchi (湯廷池), 1992a, <語法理論與機器翻譯：原則參數語法>，《第五屆計算語言學研討會論文集》，53-84頁，並收錄於湯 (1992c：251-342)。
- Tang, Tingchi (湯廷池), 1992b, 《《漢語詞法句法三集》》，台灣學生書局。
- Tang, Tingchi (湯廷池), 1992c, 《漢語詞法句法四集》，台灣學生書局。
- Tang, Tingchi (湯廷池), 1993, "On the Relation Between Word-Syntax and Sentence-Syntax：A Case Study in Chinese Compound Verbs", an invited paper presented at Second International Conference on Chinese Linguistics.
- Tang, Tingchi (湯廷池), 1994a, "More on the Relation Between Word-Syntax and Sentence-Syntax：A Case Study in Chinese Compound Nouns", a paper presented at The Sixth North American Conference on Chinese Linguistics.
- Tang, Tingchi (湯廷池), 1994, "A 'Minimalist' Approach to Contrastive Analysis of English, Chinese and Japanese", an invited paper presented at 4th International Symposium on Chinese Languages and Linguistics.
- Zhu, Dexi (朱德熙), 1984, 《語法講義》，商務印書館。