

詞語分類系統中同級義類中實現關聯性的意義與改進

蘇新春¹ 李安²

¹臺灣中央大學中文系客座教授

廈門大學中文系教授

suxch@xmu.edu.cn

²廈門大學漢語言文字學博士生

jjinaji@gmail.com

摘要：《現代漢語語義分類詞典》收詞 7.8 萬條，8.3 個義項。按五級義類編排，共有一級類 9 個，二級類 62 個，三級類 515 個，四級類 2076 個，五級類 12613 個。科學的分類詞典既要有較好的縱向關係，這主要體現為上下位的覆蓋與被覆蓋的關係，還應該有較好的橫向關係。橫向關係主要體現為義類與義類之間的搭配關係、多義詞之間的義項分佈的對應關係。橫向關係在 TMC 分類詞典中有所體現，但仍值得進一步的挖掘與完善。

關鍵字：語義分類詞典 義類 語義關聯

Significance of achieving correlation for these semantic categories under the same hierarchy and its improvement in the system of words classification

Su Xinchun¹ Li An²

¹Guest Professor in Chinese department of National Central University, Taiwan32001;

Professor in Chinese department of Xiamen University, China 361005.

suxch@xmu.edu.cn

²Doctoral candidate majoring in Chinese Philology in Xiamen University, China 361005.

jjinaji@gmail.com

Abstract: The Thesaurus of Modern Chinese (TMC) involves 78, 000 words and 8.3 semantic items. Laid out by five levels of semantic categories, there got 9 categories under the first level, 62 under the second level, 515 under the third level, 2, 076 under the fourth level and 12,613 under the fifth level in TMC. A scientific thesaurus should contain both rather good longitudinal relationship, which mainly presents the relationship of covering and being covered between hyponym and hyponym, and contain rather good latitudinal relationship, which mainly presents the collocation relationship among semantic categories and the corresponding relationship among semantic distribution of polysemous words. The latitudinal relationship has got some presentations somewhere in TMC and is worthy of further exploration and perfection.

Keywords: Thesaurus of semantics, semantic category, semantic correlation

一、語義系統的建構中應充分考慮語義的橫向關係

《現代漢語語義分類詞典》(A Thesaurus of Modern Chinese, 簡稱 TMC) 是我們正在研製中的一個語義分類詞典。收詞 7.8 萬條, 8.3 個義項。按五級義類編排, 共有一級類 9 個, 二級類 62 個, 三級類 515 個, 四級類 2076 個, 五級類 12613 個。

在研製過程中, 曾對 TMC 的分類原則、性質、功用作過一些探討。如《現代漢語類義詞庫建構中的幾個問題》(第二屆海峽兩岸現代漢語學術研討會, 澳門, 2006.11)、《關於語義分類詞典的若干問題的思考》(第 6 屆中青年辭書學工作者會議, 武漢, 2008.4)。論文正式發表於《〈現代漢語語義分類詞典〉(TMC) 研製中若干問題的思考》(中文資訊學報, 2008.5), 後又寫了《再論義類詞典的分類原則與方法》(世界漢語教學, 2010.2)。那時主要討論的是 TMC 的立類與歸詞, 中心是如何處理上下義類之間的隸屬關係。上下類之間遵循的主要原則是“下一層語義類對上一層語義類的義域作出周遍的切分”、“同級同類的概稱名放在同級類的最前面; 同級同類的分指名放在同級類的最後面”、“同級語義層中各義類按一定的語義關係排列”、“義類類名的確定是把是否具有代表性放在首位”、“三級類以下的義類類名一般都要在本類中出現”、“嚴格控制類名的重出”。

現在看來僅從上下關係來探討義類系統是遠遠不夠的。固然, 從一個認知觀出發來分出 9 個一級類, 9 個一級類下面又分出哪 62 個二級類是重要的, 但那解決的只是上下統屬關係, 它的視角是單向的直線觀察, 在語義系統中解決的是縱向關係, 所反映的只是語義的概括與具體、整體與局部、上位與下位、統屬與分轄的關係。

而無論是現實語義世界, 還是努力實現對現實語義世界作真實再現的語義系統建構來說, 語義關係都不可能是單一的縱向, 而一定是錯綜交織的。這種複雜性, 不僅僅是還會有橫向的, 交叉的, 還會有隔空關聯的現象。常說到的聯想、蒙太奇、跳躍, 揭示的都是語義之間關係的多樣性、非直線性、跳躍性。¹當然, 在平面的再現中來建構語義系統, 希冀於再現現實世界中複雜的網狀關係, 自然是不太現實, 因為紙本的“平面”性, 二維空間的局限, 成為一個天然的自然條件限制, 但努力實現對語義系統中橫向關係的探索卻是應該的。語義系統中的橫向關係包括相關性、聯動性、制約性。

在結構主義語言學理論中, 聚合關係與組合關係是語言系統最基本的兩種關係, 聚合關係其實就是語義系統中的縱向關係, 組合關係就是語義系統中的橫向關係。兩種關係缺一不可。語義系統中如果只有聚合關係, 得到的將只是一個概念分類集, 詞語類聚庫。只有在概念分類集、詞語類聚庫中同時又能反映出橫向的相關性、聯動性、制約性, 才能真正體現並發揮語義系統的性質與作用。

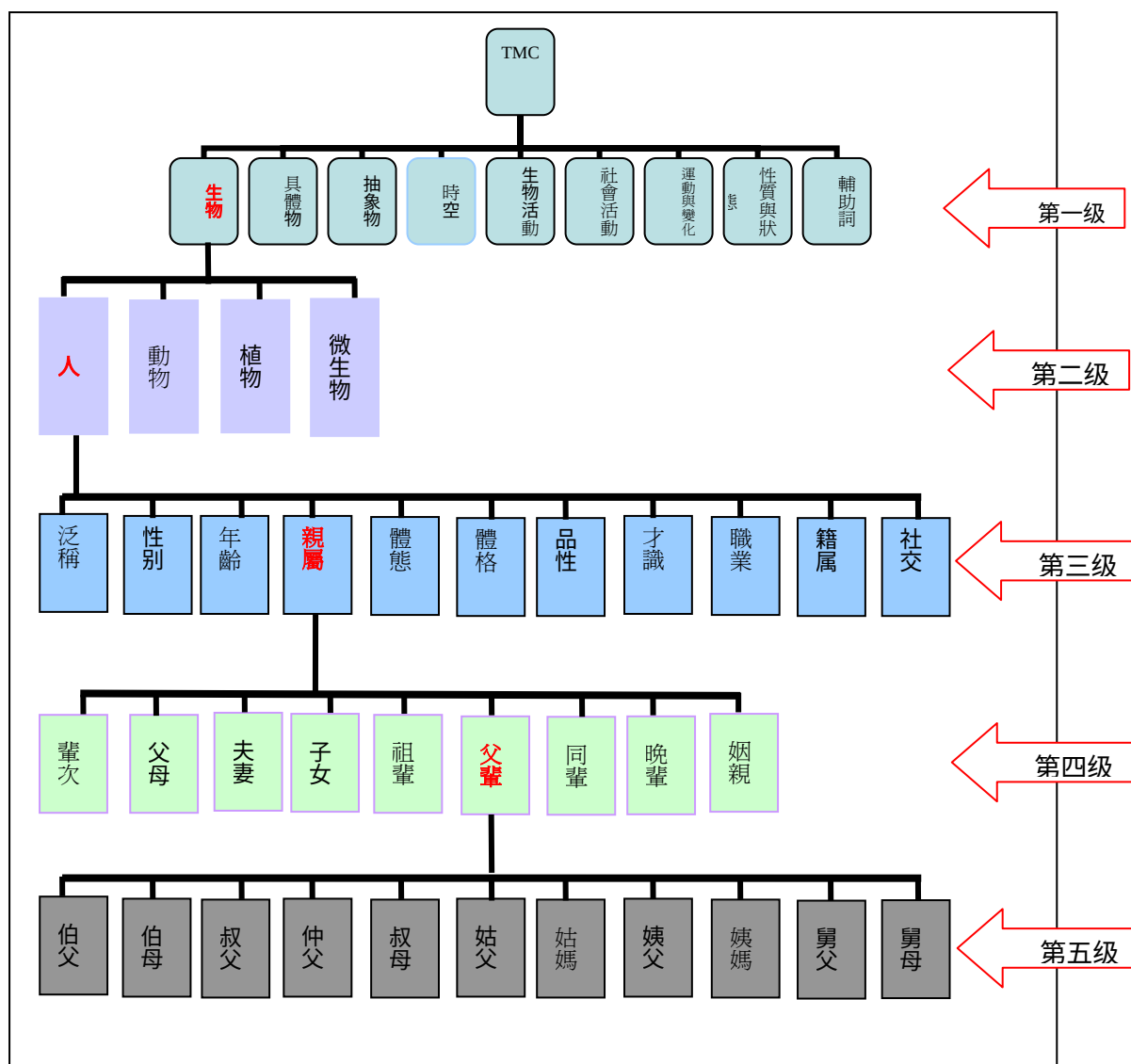
¹ 在這方面, 董振東先生的“知網”較好地實現了這種網狀關係。董先生的重要論文見於“HOWNET”
<http://www.keenage.com>。

二、 TMC 中橫向語義關係實現的基礎與條件

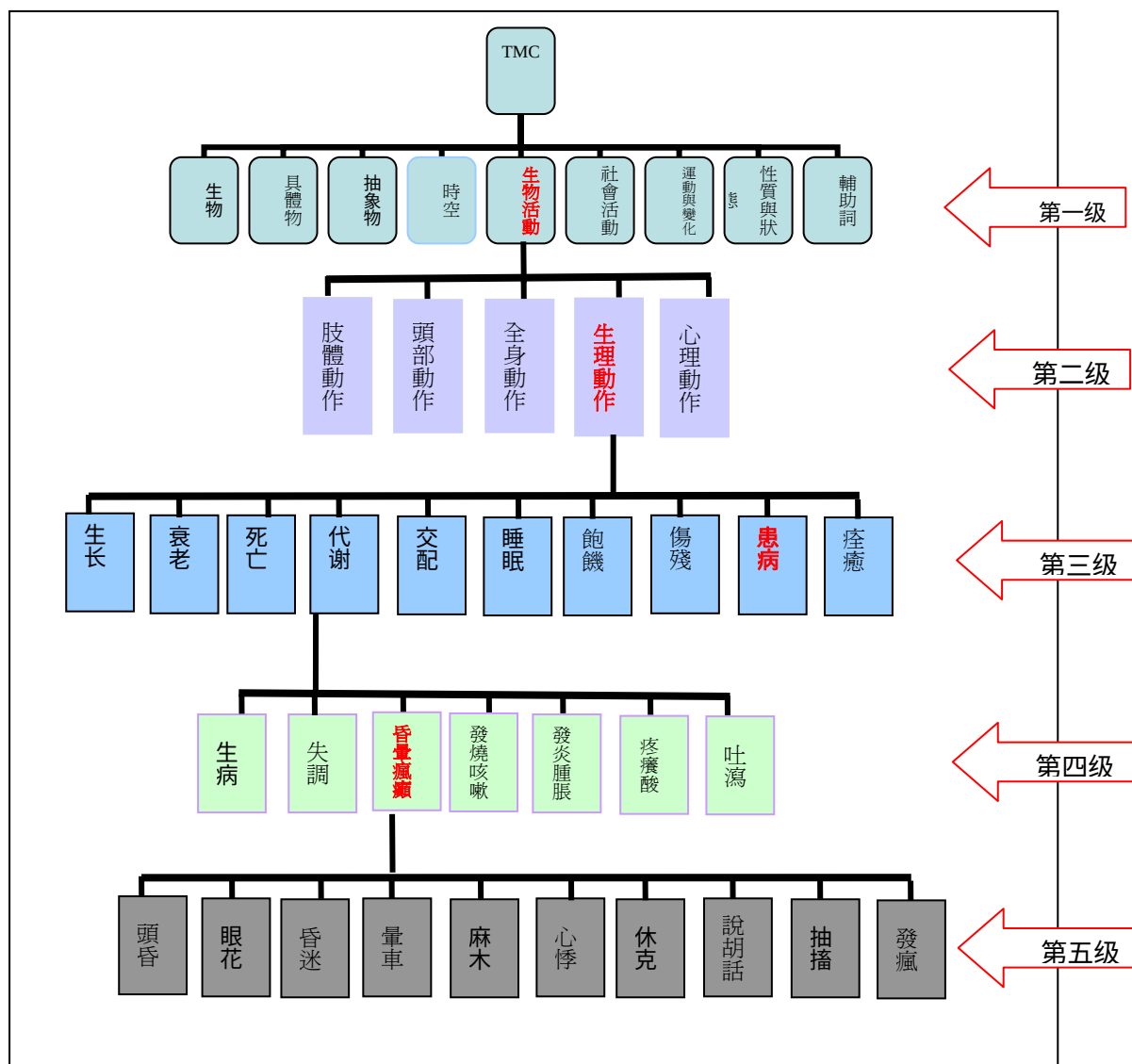
TMC 相當程度上已經蘊含了橫向語義關係存在的基礎和條件。主要表現在以下兩方面：

1) 語義層級數量的一致

TMC 執行的是五級語義層的架構，且在各個大類中都貫徹到底。無論是名詞、動詞、形容詞，或是起關結作用的輔助類，都是五級語義層劃分。這樣最上一級的語義層在語義世界中就具有了“根目錄”的作用，九個一級類成為語義系統中九個最基本的語義類。往下到第五級為限，這就使得各大類內部有著相同的層級，其語義的深淺粗細會表現出大體相近的劃分分佈，而不會出現深淺不一的犬齒狀。儘管中間類、下層類的劃分原則與標準不盡相同，但其確定的空間和層級佈局，卻規約著語義的大致走向，使得語義系統建構理論得以完整而徹底。如下面是關於“父輩”義類的五層分佈。



以上展示的五級義類劃分法，在整個系統中是貫徹始終的。下面再看一個動詞的分類，“昏暈瘋癲”例：



2) 語義層級劃分標準的近似

在層級劃分中起著保障作用的是這樣兩個原則：“對語義域的窮盡覆蓋”與“相鄰義類互不包容”。

“對語義域的窮盡覆蓋”指的是下一個層級的語義小類的義域總和應等同於上一個層級的語義類。“相鄰義類互不包容”指的是分佈於同一個層級語義類相互之間不發生交叉、相容、包含關係。

試以“父輩”為例，第二級是對第一級語義域的窮盡覆蓋，第二級的4個小類是互不包容的；第三級是對第二級語義域的窮盡覆蓋，第三級的11個小類是互不包容的；第四級是對第三級語義域的窮盡覆蓋，第四級的9個小類是互不包容的；第五級是對第四級語義域的窮盡覆蓋，第五級的11個小類是互不包容的。

當然，這兩個原則都是存在於這樣的事實之中，即它所建構的整體語義系統

是二維的，是非立體、非網路、非蒙太奇的，是可以作平面再現的；所遵循的基本分類思路是單向的邏輯劃分，在一個總的語義分類思想下將其繼續地、分步驟地劃分，中間不會出現總的分類思想和原則的轉換。

“相鄰類互不包容”是一種目標、效果的追求，卻是可以通過不同的方式來實現的。如第三級的“品性”類與“才識”類本來是很接近的，分立後則前者只指個人的道德修養類，後者只指才能知識類。又如第四級的“父母”類與“父輩”類本來是很接近的，分立後前者只指“父親”“母親”，後者只指不包括“父親”“母親”在內的父輩“親屬”。“姻親”類指因婚姻而建立起來的“親屬”關係，它裏面包括了長輩，如“岳父、岳母”；也包括了平輩，“大舅子、小姑子”；還包括了長輩眼中的下輩“子婿、半子、兒媳”，但憑著“姻親”這一點而與其他所有的“血親”有了清晰的區割。

這裏映透出一個很重要的分類思想，即“分類是相對的”，“一個義類的存在與否，義域範圍的大小，是相對於其他義類而言的”。這種分類思想在詞彙學中就表現為一個著名論點：“一個詞的詞義範圍大小是由它的同義詞決定的”。

2

在第五級中將“伯父”“叔父”“姨父”“姑父”“舅父”獨立出來，顯然這裏做了血緣的區分；“伯父”與“伯母”、“叔父”與“叔母”、“姑父”與“姑媽”、“姨父”與“姨媽”、“舅父”與“舅母”的區分，又顯示出了性別的區分。可見愈往下級義類，義類劃分愈小，詞義愈接近，已經明顯表現出由義詞到近義詞，由近義詞再到同義詞的演遞性質。

3) 語義類之間的認知關聯

上面兩點談的還是由上到下的類聚關係，這是在語義體系中能夠建立起橫向組合關係的基礎和條件，最主要的是要在橫向的相鄰義類之間要有建立組合關係的可能。要獲得這種組合可能最根本是源於一級類的劃分中。一級類尤如分水嶺，儘管粗線條，卻使得統轄的所有詞語獲得了根本不同的性質。

九個一級類實質上可分為四大類：實體——變化——性質與狀態——輔助詞。

“實體”包括：“生物”、“具體物”、“抽象物”、“空間”四個類，其共性都是一種存在，一種客觀，是與“用”相對的“本”與“體”。

“變化”包括：“生物活動”、“運動與變化”、“社會活動”三個類，其共性都是一種運動，是與“本”“體”相對的“用”。

“性質”包括“性質與狀態”一類，其含義是對“存在”、“客觀”、“體”的性質、狀態的反映。

“輔助詞”包括“輔助詞”一類，主要起到語言關聯、組合等輔助作用。

這四個大類構成了世界客觀世界最基本的差異分類。只是由於“實體”比較複雜，從而根據“生物”與“非生物”，“具體”與“抽象”，“客觀的有形”

² 李友鴻《詞義研究中的一些問題》，刊《西方語文》1958年第1期。

與“客觀的無形”再分出了“生物”、“具體物”、“抽象物”、“時間和空間”四類。“變化”也是如此，根據其變化的主體、性質和範圍，分出了“生物運動變化”、“一般性運動變動”、“社會生活中的運動活動”三類。可以說最上一個層級的九個大類代表著人們對客觀世界認知的九個基本範疇。每一個範疇的詞語都代表著人們對世界某一方面認知的結果，在進行每一個新認知、新推理、新表述時，其實都是在進行著跨大類的關聯。如：

“這是一個球”這句話，表現出的跨大類關聯是“這（**具體物**）——是（**運動與變化**）——（**抽象物**）——一個（**輔助詞**）——球（**具體物**）”。

“哲學是關於世界觀的科學”這句話，表現出的跨大類關聯是“哲學（**抽象物**）——是（**運動與變化**）——關於（**輔助詞**）——世界觀（**抽象物**）——的（**輔助詞**）——科學（**抽象物**）”。

不管語言表達多麼繁複，句子運用多麼冗長，但從最小的詞開始，往上的每一層組合，到短語，到分句，到單句，到複句，其實都是在進行著跨類關聯，只是所跨類的或大或小而已。

跨類關聯，也就使得類與類之間具有了邏輯、認知、語義上的對應、互動、組合。這就使得語義系統中具有了存在著橫向關係的客觀必然。

跨類關聯，這是人們認知活動的基本形式，它有著根本的哲學意義。只是停留在“大類關聯”、“一級類關聯”，在諸多具體的語言應用領域卻未必“管用”、“夠用”。真的要能解決問題還必須深入到語義系統中更細緻的義類，探索它們的之間的種種關聯。現在常說的“語法搭配”，其實從事的就是這項工作。可惜只停留在對語法屬性、功能的認識上，沒有深入到語義。現有的“語義搭配”大都只是就局部而無整體，窮“有知”而未達“未知”。不是人們不想超越這一步，缺乏對語義系統的總體認識和描寫往往成為一個重要的前提缺失。

在TMC中，可以完整地顯示出各種可能與不可能，最大可能與最小可能。例如“買賣”，這個動詞屬“社會活動”（一級類）——“經貿”（二級類）下的一個動詞，由於它的重要，還佔據了下面一個三級類的位置，它的關聯物件人們很自然地想到的是“具體物”（一級類）。這當然是它的常態關聯。但這樣的劃分似乎太粗略，我們還可以進一步確定在具體物中是那些具有商品性質的物品：

當不具備商品性質的物品被“買賣”關聯時，可能就會表現出通用與非通用、合法與非法等的意味，如與“空氣”“人口”的關聯；

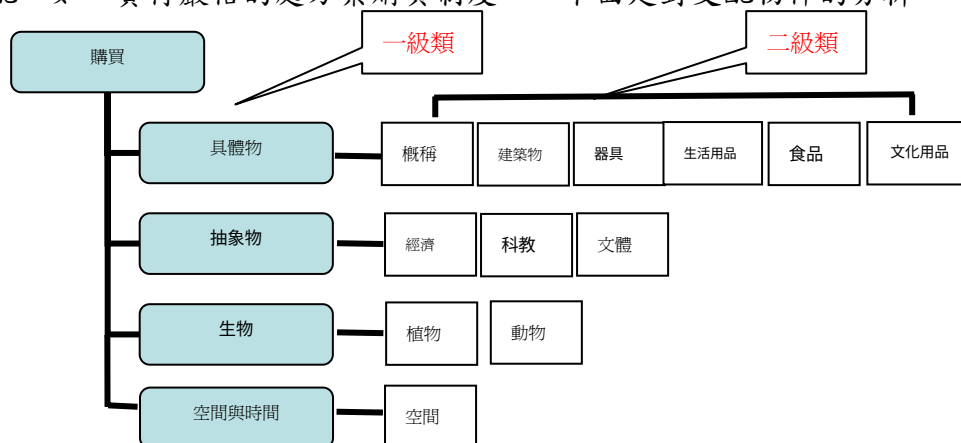
當“買賣”的對象是“燈光”時，這在“具體物”類中是一種不可能，當這種不可能出現時就會表現出“實質語言生活”變化的效果，例如現在新出現的“排氣指標的買賣”；

當“買賣”的物件是“生命”時，這在“抽象物”類中也是一種不可能，當這種不可能出現時會表現出一種“具象”的效果，留給人們文學性的遐想空間。

與“燈光”的關聯是可能中的不可能，與“生命”的關聯是不可能中的可能，語言傳遞效果全然不同。這就是語言。當我們通過語義分類系統把各種最大的可能與最小的可能都描繪出來後，其實每一個詞的語義搭配物件，每一個具體

義類的語義關聯路線圖也就都清晰了。

下面是一個樣例調查：選取了“買賣”類中“購買”一詞進行了語義搭配調查，材料是2000年的“人民日報”，共2500萬字，共用1257例。其中有支配物件的搭配占了多數，如“購買汽車”；還有有支配能力但沒出現支配物件的搭配，如“消費者一定要根據說明書謹慎地購買和使用藥品”；有修飾限定作用的搭配，如“實行嚴格的處方藥購買制度”。下面是對支配物件的分析：



下面把各個二級類中所出現過的詞語例舉出來，見下表：

具體物	(概稱) 材料、產品、設施、設備、商品、禮物、貨物、器材、年貨、燃料、工藝品、用品、物品、製品、物資、原料；(建築物) 土地、房屋、廠房、商品房、用房、住房、住宅、公房；(器具) 汽車、客車、機動車、拖拉機、儀器、機械、治療儀、武器、槍支、手提電話、手機、電話、飛機、直升機；(生活用品) 洗衣粉、首飾、衣物、電視機、家電、手套、影碟機、化妝品、彩電、時鐘、羽絨；(食品) 大米、中藥、肉食、食物、藥片、食品、藥品、白酒、麵粉；(文化用品) 發票、證券、電腦、軟體、債券、手冊、彩票、攝像機、時刻表、車票、光碟、拷貝、卡、門票
抽象物	(經濟) 股份、人民幣、股票、房產、貸款、國債、成本；(科教) 書號、標書、圖書、教材、書籍、收據、材料、書；(文體) 電視連續劇
生物	(動物) 鴨子；(植物) 種苗、人參、花生、小麥、蘋果、鮮花
空間與時間	(空間) 土地

所能與“購買”進行搭配的各種可能與不可能，最大可能與最小可能都清楚地顯示出來了。首先給我們留下清楚印象的是這四個與“購買”形成支配關係的一級類都是屬於語言中的“體”的範圍。

最多的是具體物，有搭配關係的二級類有6個，沒有出現的2個是“材料”類和“生活用品”類。這可能是由於抽樣的原因，實際用例應該是沒有問題的，如“材料”類中的金屬、鋼材、水泥，“生活用品”中的桌子、椅子。

次多的是抽象物，有搭配關係的二級類有3個，沒有出現的7個是“事情”“屬性”“意識”“社會”“政治”“軍事”“數量”。這7個有的是因抽樣的原因而沒有出現，如“數量”；有的是在特定條件上可以，如“軍事”(部隊、兵力)；有的基本上是不可以的，如“事情”“屬性”等其他幾個類。

而對與“體”相對的“用”的類，如“生物動作”“一般變化與運動”“社會活動”；與“體”為表裏關係的“性質與狀態”，則都是“購買”所不能支配

的。

這樣，我們就能由語義庫中各個義類的所屬詞入手，來梳理各種可能與不可能的搭配，窮盡搭配之可能，顯示其不可能。根據其結果，可以給每個詞所能關聯的類作“常用”、“不常用”、“可能”、“不可能”的區割，分別用 1、2、3、4 標示，其結果就得到了每個詞的“語義關係圖”。

三、語義系統中橫向語義關係的作用與意義

TMC 是一個已經“定型”了的語義分類系統，現在再來探討它的橫向關係，顯然是有兩個作用，一是挖掘，就是努力揭示其內部已經蘊含著的橫向語義關係，二是反思，即努力加以改善。而語義系統的改善，除了分類思想與方法外，最重要的影響因素就是對語義系統作用的認識。不同的目的與作用顯然會極大地影響到語義系統的建構。

語義系統的建構有多種作用，如辭彙教學中的同義詞辨析、反義詞對比、辭彙系統的整體學習作用，如提高寫作中的詞語聯想作用，如詞典編纂中提高釋義與比較的作用，如反映一個時期的社會認知體系與認知能力的作用，如反映辭彙系統中詞與義類的距離與關聯度的作用（如上面的“二”所涉及）。

這裏只從提高對多義詞內部語義關係的認識與辨析能力的作用來談談。在漢語詞語義頻的研究中，首要任務是如何認別多義詞之間的差異，要在電腦中實現這個目的，實際就是詞義自動識別，或曰詞義自動標注，或詞義排歧。由於排歧可能包含非正確的用法與表義，故在我們的研究中不太用排歧，而是用多義詞自動識別的說法。以往學術界利用的多是詞的語法資訊，如北大的《現代漢語語法資訊詞典》，我們希望能儘量利用語義關係，觀察多義詞的多個義項在語義系統中的分佈特點，從而將觀察範圍從“詞”擴大到“義類”，利用語義類與類的搭配規律，增加詞的語義區別特徵，從而收到以類馭詞、以簡馭繁、以類概括具體的效果。

1. 多義詞的多個義項之間存在橫向關係

這就有了瞭解多義詞內部多個義項在語義系統中是如何分佈的任務。分佈在哪些義類，義類差異的層級如何，是大類，還是中類，或小類；這些義類有著怎樣的關係，是隸屬關係，還是緊靠的相鄰關係，或是遙遙相隔的相對關係？當然，在談這些問題時，所蘊含的前提是我們的義項是已經整理得比較清晰，是達到了機用義項庫的要求，而不是照搬傳統義項劃分的精華與不足而來的現成義項庫。

³

在多義詞識別研究中，涉及到的雙音節多義詞有 3775 個。義項之間是存在著橫向關係的，即橫向語義類中存在著左右相鄰的關係。

例如“突破”有兩個義項：

詞	義項	例句
突破	集中兵力向一點進攻或反攻，打開缺口。	～封鎖 ～防線 ～敵人陣地。

³ 苏新春、李安、洪桂治，《机用义项库在词义标注中的作用及其完善》，刊《厦门大学学报》2010.3

突破	打破（困難、限制等）	～難關 ～定額 對這個問題的研究又有新的～。
----	------------	------------------------

從組合關係來看，義項①的支配物件具體事物，如封鎖、防線、敵人陣地等等，對應的語義類是分屬於“抽象物”的“封鎖線 3002E0b010”、“防線 3002E0b005”，和屬於“時間空間”大類的“險要地 4002C0d0”、“根據地 4002D0a010”、“白區 4002D0a011”、“軍區 4002D0a012”、“敵區 4002D0a014”、“戰場 4002D0c0”。義項②的支配物件為多為抽象事物，例如難關、定額、數量等，相應的義類為“困難 3001K0a006”、“妨礙 3001Q0c001”、“數量 3010B0d015”等。

再來看看“突破”兩個義項的語義類的橫向關係，義項①位於“6060E0b002”，義項②位於“7050C0g001”，差異始於一級類，“60”為“社會活動”類，“70”為“運動與變化”

又如“出場”，有兩個義項：

詞	義項
出場	演員登臺（表演）。
出場	運動員進運動場（參加表演或競賽）。

這兩義項的主語位置的語義位有差異，義項①指演員，類似詞有“演員、樂隊、歌手”等，義項②指運動員，類似詞有“運動員、球隊、主攻手、守門員、相撲手、隊員”等，其語義類同為“生物——人——職業”，區別只是在“職業”下面的四級類“演職員”與“運動員裁判員”；“樂隊”與“球隊”同屬“抽象事物”下的二級類差別“社會”類與“文體”類。可知兩個義項的距離相當接近。

又如“掃射”，兩個義項如：

詞	義項
掃射	用機關槍、衝鋒槍等左右移動連續射擊。
掃射	指目光或燈光向四周掠過。

義項①是指射擊動作，義項②是指目光的掠過。如在現代漢語語料庫中它的真實語料為：

- 1) 敵人/n 飛機/n 順著/p 窄/a 狹/Ag 的/u 山溝/n [掃射]/v 、 /w 轟炸/v ， /w 想/v 阻止/v 我軍/n 前進/v 。 /w
- 2) 敵人/n 督戰/v 隊/n 用/p 機槍/n [掃射]/v 他們/r 潰逃/v 的/u 士兵/n 。 /w
- 3) 梁朝棟/PER 令/v 兵士/n 用/p 機關槍/n 向直軍/PER [掃射]/v ， /w 直軍/ORG 死傷/v 甚/Dg 多/a 。 /w
- 4) 吳孫甫/PER 走/v 前/f 一/NUM 步/q ， /w 威嚴/a 的/u 眼光/n 在/p 屋子/n 裏/f [掃射]/v ， /w 最後/f 落/v 在/p 阿/ALOC 萱/nr 的/u 身上/s 。 /w
- 5) 淮海/LOC 目光/n 如/v 電/n ， /w [掃射]/v 一/NUM 明/a 全身/n 。 /w
- 6) 拖拉機/n 的/u 車燈/n [掃射]/v 著/u 前面/f 的/u 路/n 和/c

遠處/s 的/u 沙丘/n 。/w

前三例的主題義類是“戰鬥機 2004F0d0”、“槍炮 2004F0b0”，後三例的主題義類是“眼 1005A0d002”、“燈具 2006B0f0”、“燈光 2002D0c008”。義項①的代碼是“6006F0”，即“社會活動——戰爭——射擊”類，義項②的代碼是“5020D0a0”，即“生物活動——頭部活動——看聽嗅”，和“7001C0a0”，即“運動與變化——自然現象——照射發亮”。差異都始於最上面的一類，義項②是對義項①推衍而來，屬類似引申而來。

2. 多義詞義項的遠近與義類分佈有關

在已經進行的調查中，多義詞內部的幾個義項之間是有著距離遠近關係的，會反映在義類的分佈上，會表現出這樣一些規律：

一般說來義項分佈的語義級層愈低，其語義相似度愈高；

義項差異分佈的級層愈低，表明義項之間的距離愈近，差異愈多，是基於相同基本義之上的細緻概念義，或色彩義的差別；

義項之間有詞性差別的，主要表現為義類之間有相對相轉的關係；

義項差別一般以三級類為界，即義項之間的差異多表示為不在同一個三級類中。

下面來看看表現在級層上的義項差別：

1) 一級類的差別：

“著述”有兩個義項：“(1)著作①；編纂：專心～。(2)著作和編纂的成品：先生留下的～不多。”語義分佈前者屬“3008E0a001”，後者屬

“6005E0a002”，差異起於一級類，前者屬“抽象物”，後者屬“社會活動”下的“文教——寫作——寫作——撰著”。

2) 二級類的差別：

“主題”：“(1)主見：大家七嘴八舌的一說，他倒拿不定～了。(2)辦法：出～|餓～|這個～好|人多～多。”語義分佈前者屬“3001C0b001”，後者屬“3003D0a003”，即都屬於“抽象物”，差異起於二級類，前者屬“事情”，後者屬“意識”。

“追求”：“(1)用積極的行動來爭取達到某種目的：～真理|～進步|～名利。(2)特指向異性求愛。”語義分佈前者屬“5005U0c001”，後者屬

“5007K0a004”，即都屬於“生物活動”，差異起於二級類，前者屬“心理活動”，後者屬“社會活動”。

3) 三級類的差別

“狀元”：“(1)科舉時代的一種稱號。唐代稱進士科及第的第一人，有時也泛稱新進士。宋代主要指第一名，有時也用於第二、三名。元代以後限於稱殿試一甲(第一等)第一名。(2)比喻在本行業中成績最好的人：養雞～|行行出～。”語義分佈前者屬“1001H0a009”，後者屬“1001K0j042”。即都屬於“生物——人”，差別起於三級類，前者屬“才識——人才——人傑”，後者屬“社交——君臣——狀元”。

“點名”：“(1)按名冊查點人員時一個個地叫名字。(2)指名：他要求派人支援，～你去。”語義分佈義項(1)屬“6001H0a013”，義項(2)屬“6001G0a009”。即都屬於“社會活動——管理”，差別起於三級類，前者屬“調查——查審——點名”，後者屬“調動——調派——指名”。

4) 四級類的差別

“工友”有兩個義項：“(1)機關、學校的勤雜人員。(2)舊時稱工人，也用於工人之間的互稱。”語義分佈義項(1)屬“1001I0g014”，義項(2)屬“1001I0e001”即都屬於“生物——人——職業”，差別起於四級類，前者屬“服務員——勤雜工”，後者屬“工人——工人”。

5) 五級類的差別：

“主人”有兩個義項：“(1)接待客人的人（跟‘客人’相對）。(2)財物或權力的所有人”。語義分佈前者屬“1001K0e002”，後者屬“1001K0e009”，即都屬於“生物——人——社交——主人客人”。差別起於五級類，前者屬“主人”，後者屬“物主”。

以上顯示了多義詞的義項分佈情況。一級類的區別，往往相互之間有著“體用同稱”的變化，如“採訪：(1)只就某個問題或對某個人進行採訪：接受記者～。(2)這樣採訪寫成的文章：登載了一篇關於他的模範事蹟的～。”有的大類距離，仍會表現出較緊密的關係，特別是在具體物與抽象物之間，如“彩電：(1)彩色電視的簡稱：～中心。(2)指彩色電視機：一台～。”表現為一級類的“具體物”與“抽象物”的差別，但其語義距離卻相當近。或是“生理動作”與“社會運動”之間。

探討義項分佈在不同義類之間的橫向關係，是瞭解詞義距離的遠近、詞義轉化、引申、發展的重要途徑。

四、 橫向語義關係建構的不足與改進

利用 TMC 來瞭解多義詞的義項關係是一個重要的資源，但無論是 TMC 本身，還是由此來達到多義詞自動識別的目標中，都還存在著有待完善的地方。例如 TMC 多義詞的義項收錄是否齊全。如“主宰”“主人”“住家”“轉向”“狀紙”“暗中”“測驗”中多義詞還闕如。

又如對傳統義項的修訂，因為多義詞的自動識別將在電腦中實現，它必須要有符合機用需要的義項。例如“轉借：(1)把借來的東西再借給別人。(2)把自己的證件等借給別人使用：借書證不得～他人”，又如“(1)中藥店按照顧客的藥方取藥，也指醫院的藥房為病人取中藥。(2)拿著藥方到中藥店買藥：抓一服藥”；“專機：(1)在班機之外為某人或某事特別飛行的飛機。(2)某人專用的飛機”；“專人：(1)專門負責某項工作的人。(2)臨時派遣專辦某件事的人”；“鞭炮：(1)大小爆竹的統稱。(2)專指成串的小爆竹”，在機用義項庫中如何處理，都值得商量。

當然，還有的問題是由於平面、二維的語義分類詞典難以克服的問題，有的詞會出現可以多歸的可能，但受制於義項只能一歸不能“重出”的限制，這就會

出現在可能出現的地方不出現的現象。

參考文獻：

- [1] 蘇新春,《〈現代漢語語義分類詞典〉(TMC)研製中若干問題的思考》,《中文資訊學報》,2008.5
- [2] 蘇新春、洪桂治、唐師瑤《再論義類詞典的分類原則與方法》,世界漢語教學,2010.2
- [3] 李友鴻《詞義研究中的一些問題》,刊《西方語文》1958年第1期。
- [4] 蘇新春、李安、洪桂治,《機用義項庫在詞義標注中的作用及其完善》,刊《廈門大學學報》2010.3
- [5] 王惠、詹衛東、劉群,《〈現代漢語語義詞典〉的概要及設計》,刊於《1998 中文資訊處理國際會議論文集》,北京,清華大學出版社,1998 年。
- [6] 董振東、董強,《知網和漢語研究》,《當代語言學》,2001 年第 1 期。
- [7] 董振東、董強,《面向資訊處理的辭彙語義系統中的若干問題》,《語言文字應用》,2001 年第 3 期。