

詞彙系統的存儲結構

楊建軍*

1. 詞彙系統的基本成員單位

任何系統都有成員、結構和功能這三大要素，要討論它的結構和功能，都先要確定它的基本成員，詞彙系統也不例外。本文把詞的意義作為詞彙系統的主體，當然它的基本成員也應該與此一致。“在語言中，語義層的單位最多，單位之間的關係最複雜。語義層最大，最繁雜。”¹通常認為，與詞的意義相關的基本單位有義位和義素這兩個層次，該怎樣選擇呢？

顯然，選擇義素作為詞彙系統的基本成員單位無疑具有很多便利條件，最大的好處就在於能夠對意義進行形式化，便於採用更加科學的形式理論和方法進行深入的研究。但我們認為，義素並不適合作為詞彙這一層次系統的基本成員單位，理由有三。首先，義素理論自上個世紀40年代丹麥語言學家葉姆斯列夫（L.Hjelmslev）提出以來，儘管取得了很大的成就，比較成功地運用在表示親屬，軍銜和坐具等相對封閉的類別詞中，但是還無法實現對所有的詞語進行形式化的描寫。賈彥德對漢語的詞進行了義素分析，但也只是整個詞彙系統的部分成員。而且隨著具體分析的展開，其局限性也逐漸暴露出來。林杏光對此進行了總結：其一、語義是開放性的複雜系統，義素的數目很難確定；其二、義素是純粹的語義單位，沒有任何形式可以依附，因而確定義素具有很大的隨意性；其三、某些詞，甚至是基本詞似乎根本無法分析，比如“紅、黃、藍”，“牛、馬、狗”等。²雖然沒有必要對一個詞彙系統所有的詞義都進行義素分析，但是基本詞則是必須要考察的對象，因此，義素作為基本單位，不能對所有基本的系統成員進行形式描述。其次，若採用義素作為基本成員，與詞彙系統的具體情況不符，違背了系統的層次原則。系統可以具有層次，但是不同層次的單位不發生直接的關係，任何一個確定的系統，它的結構和功能都是以它同層次的直接成員作為基礎的。比如，原子結構不同於物質結構，原子結構的基本要素是質子、中子和電子，物質結構的基本要素是原子或者由原子化合而成的分子，顯然原子結構是物質結構的子結構，但原子結

* 南京大學 中文系 博士研究生，江蘇 南京。

1 賈彥德：漢語語義學[M]。北京：北京大學出版社，1992，18頁、19頁、44頁。

2 林杏光：詞彙語義與計算語言學[M]。北京：語文出版社，1999，305~306頁。

構的基本要素並不是物質結構的直接結構要素。按照義素理論，義素是義位的基本要素，但是它並不是詞彙系統的直接結構要素。要反映詞彙系統的基本面貌，無論是詞，還是語義場，其內部的關係及變化都不是以義素作為基本單位的，而是以義位作為整體表現出來的。換句話說，義素不能直接反映出詞彙系統的基本面貌。最後，義素是純粹的語義單位，不能直接跟語言的外在形式發生關係，不便於討論詞彙系統的演變問題。“詞彙成分必得是個語音和語義的結合物。”³雖然詞彙系統的主要內容是詞的意義，但是也必須兼顧它的外在形式，否則就無法區分意義一致，但分屬不同形式的成員。

詞語的意義是詞彙系統的主要內容，義位能夠反映出它的主要面貌和歷時變化，但是直接把義位作為詞彙系統的基本成員單位也不恰當。在現代語義學中，“義位”一般也是被當成一個純粹的語義單位，因此，通常是把“義素”和“義位”等基本單位放在一起考慮，最終建立起來的是分層次的語義系統，而不是詞彙系統。詞彙系統的基本成員是和外在形式直接聯繫的符號，所以它應當是以“義位”為主體的符號單位。其次，詞彙系統的基本成員之間不僅具有語義上的聯繫，還存在相對的“權重”價值關係。在一個多義詞裡面，中心成員就比其他非中心成員跟它們共同的外在形式具有更緊密的關係。位於同一個語義場中的成員，核心成員就比非核心成員具有更重要的地位。單純的語義系統是不考慮這些價值關係的，但在詞彙系統中，這些關係跟純粹的語義關係是一樣重要的，它們同樣可以決定詞彙系統的基本面貌。最後，詞彙系統的基本成員是語言運用的基本單位，它是可以獨立運用的。詞典中列出的“義位”，有些是作為構詞成分出現的，並不能單獨運用，它們不屬於詞彙系統的基本成員。

由此，我們提出把“詞位”作為詞彙系統的基本成員單位。詞位的構成形式是：

詞位 = 外在形式 + 義位

從本質上講，詞位仍舊是一個符號，是一個能夠獨立運用的最小符號單位。詞位與詞位之間的差別是由形式和義位共同決定的。形式相同，義位不同是不同的詞位，比如一個多義詞的內部，各個詞位具有相同的形式，不同的義位。形式不同，即使義位一致，也是不同的詞位。比如“飲”和“喝”，義位基本一致，但具有不同的形式，是不同的詞位。

3 高名凱：《語言論》[M]，北京：商務印書館，1995，258頁。

2. 詞的結構與功能

在確定了把詞位作為詞彙系統的基本成員之後，我們接下來就要討論詞位與詞之間的關係。以前的詞彙學研究一般把“詞”作為詞彙系統的基本成員，這個位置現在給了詞位，那麼“詞”應該放在何處呢？因此，必須探討詞的結構和功能。

在語法學研究中，對“詞”比較一致的看法是能夠獨立運用或者單獨作句子成分的最小的語言單位。語言單位是語言系統的符號單位，是形式和意義的結合體，排除了純粹的語義單位和語音單位。純粹的語義單位和語音單位的提出，是出於研究的需要，在實際的語言中，它們都不是自然的單位。基本的語言單位包括四級，語素、詞、短語和句子，這些單位都是形式和意義不分離的單位。能夠獨立運用是針對語素提出的，語素是構詞和構形的最小語言單位，但不是造句的最小語言單位。單獨作句子成員是為了把虛詞包括進來，除了一部分助詞，虛詞是不能獨立運用的。可見這個定義用“語言單位”限定範圍，用“獨立運用”排除語素，用“最小”排除短語和句子。之所以要這樣定義，就是要確定研究句法結構的最小單位。顯然，這個定義是從語言運用的角度展開的。但是，這個問題還有討論的餘地。

一般都說，句子是由詞組成的，這其實是表面現象。這可以從句子的生成和理解兩個角度來說。任何一個說話者，要把自己想要表達的意思生成一句話，僅從語言材料上講，說話者選定的任何一個符號，其意義都是確定的，唯一的（除非修辭上的需要，故意模稜兩可），如果這些符號中包括多義詞，那麼它的意義也應該是確定的，唯一的。比如：

花 / 開 / 了 / 。

在這個句子中，我們通常說由3個詞組成。詞“花”有6個義位（《現代漢語詞典》共列有19個義項，包括語素義和可以獨立運用的義位，不包括同形詞“花費”的“花”，這裡所說的義位僅僅是指可以獨立運用的義位）：①指種子植物的有性繁殖器官；②可供觀賞的植物；③顏色或種類錯雜；④（眼睛）模糊迷亂；⑤衣服磨損或要破沒破的樣子；⑥姓。詞“開”有14個義位（《現代漢語詞典》共列有20個義項，不包括另外兩個同形詞）：①打開；②打通，開關；③展開，分離；④（河流）解凍；⑤發動或操縱；⑥（隊伍）開拔；⑦開辦；⑧舉行；⑨寫出；⑩〈方〉開除；⑪（液體）受熱而沸騰；⑫〈方〉吃；⑬用在動詞或形容詞後；⑭姓。詞“了”有2個義位：①用在動詞或形容詞後面，表示動作或變化已經完成；②用在句子的末尾或句中停頓的地方，表示變化或出現新情況。這3個詞有這樣多的義位，對於說話者，並不需要去考慮它們所有的義位，而是直接確定他表達需要

的義位，因為他所需要表達的意義對於自己是明確的、唯一的。即便是出於修辭的需要，對相關的義位進行選擇，但這種選擇並不是一個詞之內的義位的選擇，而是在不同詞的義位之間的選擇。對於聽話者來說，他要按照語言系統約定的規則和意義對語音鏈進行切分，從語言材料的角度看，切分的結果應該是3個意義確定的符號，聽話者最後能夠理解的句子也就是由這3個意義確定的符號組成的。

我們常說，信息從一方傳遞到另一方，這其實是一個比喻的說法。通過聲波或者電波傳遞的內容其實是符號的形式，而不是形式與意義相結合的符號。試想，如果意義真能夠通過物質載體跟形式特徵一起傳遞，那麼各種語言之間的交流就不再需要翻譯了，因為意義已經在裡面了。索緒爾把語音鏈切分下來之後的單位叫做“音響形象”，它本質上是一個心理實體，是已經重新賦予確定意義的符號，顯然是看到在信號的傳遞過程中，傳遞的僅僅是物理上的音響特徵。聽話者是通過認知活動，在語言規則、意義系統和語用環境的基礎上把音響特徵重建為心理實體，這樣才能理解說話者表達的意思。說話者在造句的過程中使用的規則和材料的具體意義，在傳遞的過程中都消失了，僅僅是把按照這些規則和意義建立起來的物理實體傳遞給了聽話者，聽話者需要使用相應的規則和意義來重建相應的心理實體，從而理解整個句子。在聽話者重建的過程中，只有這些規則和材料的意義完全確立之後，這個句子才算是真正建立起來了。有人會說，句子是由詞構成的，詞義的選擇是在語境中最終確定的。這種說法並不符合事實。要從語音鏈上把詞切分出來，切分的依據既有語法規則，也有意義的參與，這個意義不能是一個籠統的詞的意義，而應該是具體的義位。再看上面這個例子，“花”和“開”有這麼多意義，如何確定它們各自的籠統意義呢？事實上，籠統的詞義根本就不存在。“花”和“開”能夠切分開，從語義搭配的角度看，那是因為它們各有可以搭配的確切義位。我們不能說，我們已經把一個句子的基本組成單位劃分開了，卻還沒有確定這些基本單位的意義。即便一定要認為句子是由詞組成的，那也是一個近似抽象的句子。因此，任何一個具體的句子，對於說聽雙方，其意義組成的基本單位是義位，而不是整個詞。

不過，無論是“詞”還是義位，對於句法結構的研究不會有甚麼實際的影響，因為它們在句子中都可以近似地看成是一個個能夠獨立運用的最小語法單位，對句子的結構本身不產生影響。但這對於詞彙系統來說，影響可就大了。事實上，詞彙系統內部的語義關係，價值關係一般都是建立在與義位緊密相關的詞位這一層次上的，“詞”顯然不是詞彙系統的基本成員單位。

但是，“詞”也的確是詞彙系統中存在的層級單位，正如句法結構一樣，除了詞這一最小的基本單位之外，還有短語和句子這兩個層級單位。這就必須重新確立“詞”在詞彙系統中的地位和作用。

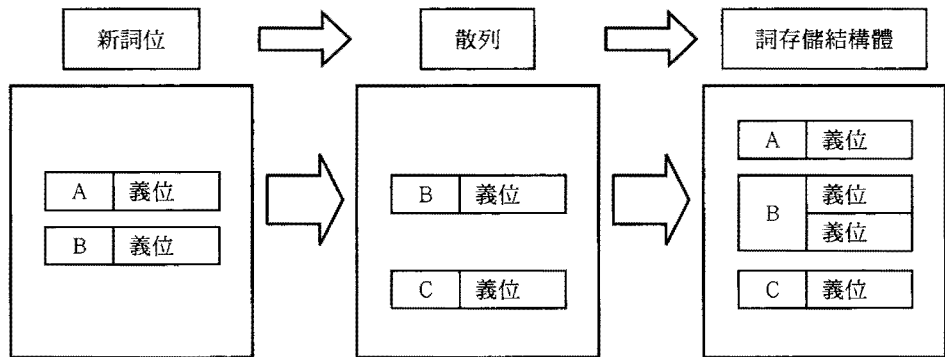
詞彙是語言的建築材料，這是詞彙最本質的屬性。語言是符號系統，任何符號系統本身的自然組成部分當然也是符號集和規則集，對於語言也就是詞彙和語法。語法可以自成系統，詞彙當然也可以自成系統。但無論這個系統的構成怎樣，都不應該偏離它為語言系統提供語言材料這一基本功能。按照“結構決定功能，功能制約結構”這一系統學原理，詞彙的結構是圍繞實現自己作為語言材料倉庫這一功能展開的。其首要的職能便是要能夠採用某種方式把這些語言材料存儲起來，以便取用，也就是欲取先存。

詞彙作為一種系統，要實現語言材料的存儲，就表現為用來組織材料的存儲策略，換句話說，就是採用甚麼樣的數據結構來實現語言材料的存儲，而這樣的數據結構是簡潔而高效的，體現出系統存儲的優化性能。語音和語義是語言符號的兩“極”，存儲的內容當然包括形式和意義這兩個“極”，但還不止這兩項內容。“名無固實，約之以命實，約定俗成，謂之實名。”（《荀子·正名》）一條最基本的語言材料應該包括“名”、“實”和“約”，這“約”就是“名”與“實”的對應關係。詞彙系統自然是需要存儲“名”和“實”這些物質內容，也要存儲它們“約”定的對應關係。一項新的意義產生之後，跟一定的外在形式約定在一起，形成一條語言材料數據，也就是詞彙系統的基本成員詞位。

這個新詞位是如何作為一個整體存儲起來的呢？如果系統中不存在與這個新詞位具有“同名”的存儲項，那麼它就會形成一個新的獨立存儲體。這便是我們所看到的單義詞，其存儲體跟詞位是一致的。假如詞彙系統需要存儲的都是這一類不同名的成員，那麼它的存儲形式就非常單一，存儲體與系統的基本成員詞位之間是一一對應的關係。換句話說，系統有多少個存儲體就等於有多少個基本成員，有多少個基本成員就等於有多少個存儲體。但事實上，語言使用者在根據需要創造新的語言材料的時候，不僅僅採用這種用新“名”約定新“實”的完全創新形式，也在已經存在的系統成員的基礎上加以引申，以舊“名”約定新“實”進行部分創新。詞彙系統如何存儲這種舊“名”新“實”的成員呢？如果詞彙系統繼續採用“一一對應”的存儲策略，那麼其存儲結構既不簡潔也不高效。首先，雖然存儲體的結構簡單化了，但是導致了數量的急劇增加。以《現代漢語詞典》為例，共收詞語6萬多，平均每個詞有4個義項（假設義項大致與義位一致），則總共有20多萬個義項。顯然，6萬與20多萬這兩個數量之間的差距是很顯著的。其次，由於組織形式

的單一，增強了詞彙基本成員的離散性，減弱了相關性，從而使整個存儲機制顯得冗餘並且呆滯。最後，同“名”異“實”的成員之間，在形成之時具有引申關係，在形成之後具有價值關係，存儲在不相關的結構體中不便於反映。我們認為，詞彙系統沒有採用“一一對應”的存儲策略，而是採用了高級的“散列”存儲策略。

甚麼是“散列”呢？簡單地說，對於由名稱和實體組成的數據，這個名稱可以看成是這條數據的“鍵”，其實體可以看成是這條數據的內容。根據“鍵”轉化為哈希代碼，通過“鍵”來存儲和檢索數據的內容。在數據結構中，哈希代碼是由哈希函數 $H(\text{key})$ 產生的，它把一組數據的關鍵字映象到一個有限的連續空間上，並且把這個代碼作為數據的存儲位置。這一映象過程也就是散列。⁴詞彙系統直接把“名”作為系統成員的“鍵”，根據“鍵名”來散列各個成員。更確切地說，就是根據詞位的形式進行散列，“鍵名”相同的存儲在一個結構體之內，不同的存儲在不同的結構體。其結果是有些結構體內只有一個成員，另一些結構體內則有兩個或者兩個以上的成員。這些存儲結構體就是“詞”，用“散列”的術語叫做“桶”。高名凱早就提出詞“是一種複雜結構”⁵，但是沒有指明它到底是一種甚麼樣的結構。我們認為，詞位是詞彙系統的基本成員，詞就是詞彙系統用來儲存這些基本成員的存儲結構，是詞彙系統採用“散列”策略的存儲結果。詞彙系統對新詞位進行“散列”的過程如下圖：



(示意圖說明：新詞位 = 鍵名 + 義位，其中鍵名用英文字母表示。同時，鍵名還是詞位的外在形式，作為整個詞位的代表符號，用來指稱這個詞位，如詞位A，詞位B。由於在同一結構體中的詞位共享一個鍵名，因此也用鍵名來代表整個存儲結構體，如詞A，詞B。在該示意圖中，詞位A和詞位B是新產生的詞位，在存入詞結構體的過程中，按照詞位的鍵名進行散列。鍵名相同的散列在同一詞結構體中，

4 嚴蔚敏、吳偉民：《數據結構》[M]北京：清華大學出版社，1997，252~258頁。

5 高名凱：《語言論》[M]北京：商務印書館，1995，268頁。

如新詞位B散列到詞彙系統中原有的存儲結構體詞B中。如果詞彙系統中不存在與新詞位鍵名相同的存儲結構體，新詞位則獨立形成一個新的詞結構體，如詞位A散列的結果是形成新的存儲結構體詞A。)

接下來就需要進一步討論“詞”的內部結構形式問題。

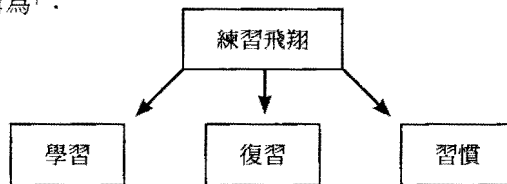
“散列”策略只是把詞位分到了不同的“桶”內，但沒有解決“桶”內部的結構問題。單義詞內只存儲了一個詞位，它的結構當然很簡單，多義詞內部有兩個或者兩個以上的成員，它們之間就存在相對有序的結構問題。探討“詞”的內部結構有三個途徑可以考慮。

其一，從意義的引申關係考慮，這是傳統訓詁學和詞彙學研究詞義演變的主要內容。通常情況下，一個詞內部的意義分成本義和引申義，引申義再分為直接引申義和間接引申義，由此形成一個意義的引申關係結構。根據引申途徑的差異，可以把由此形成的結構分成三種。

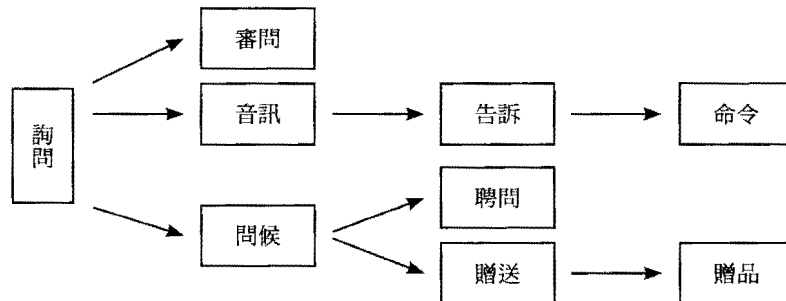
第一種是線性引申(不同的研究者使用不同的術語)，即從本義開始，各個意義逐一引申，比如，“秋”的引申結構為⁶：



第二種是發散引申，即其他意義都以其中的一個意義作為直接引申的對象。比如，“習”引申結構為⁷：



第三種是複合引申，即前面兩種方式結合在一起。比如，“問”的引申結構如下圖⁸：



6 趙克勤：《古代漢語詞彙學》[M]。北京：商務印書館，1994，88頁、89頁、91~92頁。

7 同上注。

8 同上注。

從數據結構上講，這三種引申結構嚴格地應該歸為兩類，第一種是線性結構，第二種和第三種都屬於樹形結構。儘管意義的引申結構表現了引申的先後次序，因為引申義後於被引申義產生，但是這種反映是局部的，相對的。單純的線性引申和樹形結構中的線性分支可以從整體上表現各個意義的絕對有序。比如，“秋”的引申序列表現了各個意義產生的絕對先後關係。在“問”的複合引申結構中，詢問→音訊→告訴→命令，詢問→問候→聘問，詢問→問候→贈送→贈品，這三個線性分支也表現了各個意義產生的先後時間序列。然而，不論是發散引申，還是複合引申，不處於線性分支上的意義之間，其產生的先後關係並不確定。比如“習”的三個引申義都直接從本義引申而來，處於引申義的第一個層次，但很難說它們就同時產生。在“問”的結構中，不能認為表示贈品的意義一定產生在表示音訊的意義之後。

其二，直接從意義產生的先後關係考察，按時間進行排列。詞的意義按照其內部各個詞位形成的時間先後關係累積存儲，於是出現“呈層積狀的詞義現象”，可以稱之為“詞的義層”⁹。比如“寺”幾個詞位及其產生的時間分別是，“奄人”產生在先秦，“官署名”產生於西漢，“佛教廟宇”產生在東漢，“伊斯蘭教徒禮拜、講經的地方”產生於明清。所以“寺”的最終存儲結構就是一個具有方向的線性序列。如果能夠弄清楚每一個存儲結構中所有詞位的具體形成時間，那麼也就相當於確定了所有存儲結構的內部序列，這些序列都是具有方向的線性序列。但是，這樣定義的“義層”不能全面反映存儲結構的內部變化，特別是結構態的轉變。“寺”在先秦是一個簡單結構體，僅僅包含“奄人”一個詞位；到西漢增加了表示“官署名”詞位，變成了複合結構體，結構體的中心詞位也被後起的詞位所取代，結構體本身發生了質變；東漢增加了“佛教廟宇”這個詞位，但還僅僅是結構體內存儲詞位數量的變化，不過表示“奄人”的詞位在結構體內則進一步邊緣化；東漢以後，“佛教廟宇”這個詞位逐漸取得結構體內的中心地位，結構體再次發生質變；明清以後，結構體增加了一個詞位（《明史·西域傳·天方》中稱“禮拜寺”），但沒有導致結構體的質變，這種情況延續至今。可見，“義層”以及前面提到的意義引申結構都不能反映詞的這些結構態的變化，因此也就不宜用作構建詞彙系統的存儲結構的依據。

其三，按照意義的語用價值分層。“詞”作為詞位的存儲結構，的確是層化的結構，但不是以形成時間為基礎的層級關係，而是以結構體內各個詞位的相對語

9 洪成玉：詞的義層[J]，首都師範大學學報，96，1。

用價值作為“層化”的依據，從而在結構體內形成中心區和非中心區，詞位也據此分化為中心詞位和非中心詞位。中心和非中心的對立是存儲結構內部的核心矛盾，也是結構態形成的主要依據。“詞裡有一種語義層級，其中包含有強的意義，也包含有弱的意義。前者不一定是最古的，可是人們一聽到詞就必然會想起這些意義；它們之所以有這種力量是由於它們的使用的重要性。”¹⁰那麼，這種以語用為基礎的“強弱”對立為甚麼是詞內部存儲結構形成的基礎呢？這是因為其對立可以決定整個結構體的“外在形象”，換句話說，對詞結構體的性質具有規定作用。不同的存儲體，在通常情況下，其外在的形式差異就可以把它們相互區別開來，而不管其內部的具體內容如何。因為不同的形式就代表了不同的存儲體，形式的區分就代表了各自的存儲結構，形式就具有規定性，而不需要深入到內部的結構。但這裡所說的是，一個存儲體在其形式沒有發生變化的情況下，其內部結構的性質是如何規定它自身的。作為一個存儲體，“詞”在不斷地發展，其發展也有量變和質變之分。再看上面談到的“寺”。從先秦到漢代，它由單義變成多義，由簡單變成複雜，首先是存儲詞位數量上的變化，但這樣的變化還不能導致整個結構體的質變。不過它在漢代的確發生了質變，那是因為其中心詞位發生了交替，“官署名”取代了“庵人”在存儲體的中心地位。這種情況同樣發生在東漢以後，“佛教廟宇”取代了“官署名”的中心地位。可見，結構體本身存儲詞位的數量變化僅僅是一種量變，而非質變，其內部中心詞位的交替才導致了結構體的質變。這是結構體發生質變的一種情況，另一種情況是中心詞位雖然沒有被其他非中心詞位排擠掉，但是其本身發生了質變，其結果就如同被一個新的詞位取代了原有的中心詞位。傳統詞彙學上講到的詞義的轉移就屬於這種情況，而詞義的擴大或者縮小的結果也往往產生質變。

存儲結構體的層化結構具體情況是這樣的。首先，任何一個詞結構體都至少存在一個中心詞位。單義詞只存儲了一個詞位，這個詞位也是它的中心詞位，因此也只有中心區。比如，“斜”的存儲結構是：

斜	不正
---	----

(示意圖說明：在示意圖中，存儲體的中心區與它的鍵名並列在一起，不與鍵名並列在一起的則表示非中心區。)

多義詞存儲結構通常具有中心區和非中心區，比如，“武”在先秦時期的存儲結構為：

10 約瑟夫·房德里耶斯著，岑麒祥、葉蜚聲譯：《語言[M]》北京：商務印書館，1992，223~224頁。

武	與軍事、戰爭相關的事
	勇猛
	足跡
	繼承
	半步
	周代樂曲

一個存儲結構中，中心詞位可能不止一個，卻不存在非中心詞位，比如，“月”在唐宋以前的存儲結構為：

月	月亮
	計時單位

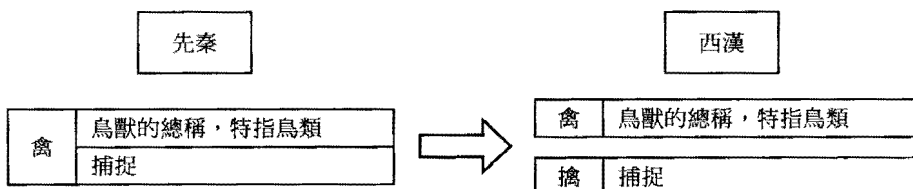
3. 詞結構的發展

各個時代的存儲結構就稱為結構態，存儲結構的發展表現為結構態的轉變。從存儲結構的角度看，可以分成三種，即不變，量變和質變。比如“斜”，從古至今都沒有發生任何變化，這樣的情況是很少的。存儲結構的量變主要表現在存儲詞位數量的變化，或增或減，一般發生在非核心區，比如，“武”在西漢的存儲結構發展為：

武	與軍事、戰爭相關的事
	勇猛
	足跡
	半步
	武士

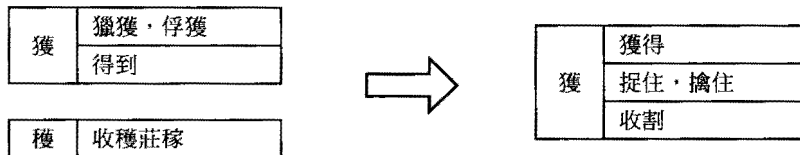
非中心詞位的增減改變了存儲體存儲信息量的多少，但沒有改變該結構體本身的性質，因此這樣的量變對結構體的影響是不大的。

所謂結構體的質變，主要包括四種形式，即結構體中心詞位的交替，結構體的裂變，結構體的聚變和結構體的解體。結構體中心詞位的交替是指原來的中心詞位被其他非中心詞位取代。比如，“寺”的中心詞位在整個發展過程中先後有兩次交替，“官署名”取代“庵人”，“佛教廟宇”取代“官署名”。所謂結構體的裂變是指某些詞位從原來的結構體中分裂出去，形成新的結構體。比如，“禽”結構體從先秦到西漢的發展過程如下圖：



從文字的角度看，“禽”和“擒”是古今字，從存儲的角度看，詞結構“禽”發生了裂變，裂變為兩個存儲結構體。

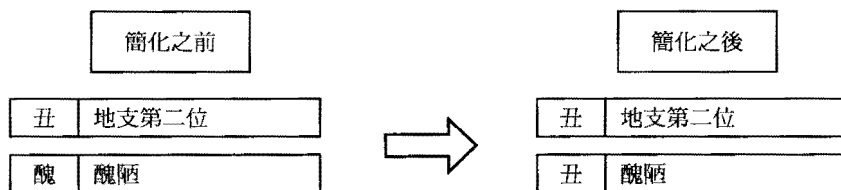
所謂結構體的聚變是指兩個結構體由於語言的發展導致合而為一，並且共享一個“散列”鍵名。比如，“獲”和“穫”本是兩個同源詞，屬於不同的結構體，在後來的發展過程中逐漸聚變成一個詞結構體。



(示意圖說明：聚變後的結構體採用了《現代漢語詞典》的觀點。)

這個聚變過程我們還沒有研究清楚。在漢字簡化之前，雖然“獲”與“穫”在字形上有別，而且也不能完全相互通用，但這樣的區分在聚變過程結束後就應該僅限於文字的範圍之內。漢字在形體上的區別作用超出了語言內在的區分。“漢字的確從字面上區別了很多在語言裡沒有從語音上加以區別的意義。其中不但有彼此沒有親屬關係的一般同音詞，也有關係極其密切的同音同源詞，很可能還有不少屬於同一個詞的不同用法。”¹¹再如，在上古時期介詞“于”跟“於”是用法有別的两个不同音的詞，後來聚變為一個詞。因此，我們認為詞結構的聚變是存在的，它是裂變的逆過程。

聚變是詞結構的質變，它的變化是以語言的變化為基礎的。如果僅僅是由於文字形體的變化導致“同名”，不按照“散列”的原則進行重新散列，因為“散列”是針對新詞位發生作用的，並且任何詞位都只進行一次散列。另外，這些僅僅“同名”的結構體之間也不發生聚變，仍然維持原來的獨立狀態。比如“丑”和“醜”在簡化之前是兩個結構體，簡化之後仍然是兩個結構體：



某些結構體中，由於其存儲的所有詞位都不再單獨使用，從而導致整個結構體的解體。從先秦到現在，許多單音節的存儲結構由於這種原因，逐漸衰變，沒有能

11 裘錫圭：文字學概要[M]。北京：商務印書館，1988，244~245頁。

夠單獨運用的詞位，最終解體而退出詞彙系統。至於它轉變為生成詞彙系統新的基本成員的構成要素，那是另外一個層面的內容。

在存儲體的內部，無論發生怎樣的變化，量變或者質變，其影響都主要在存儲結構中，也就是說，其發展變化規定存儲體的性質。存儲體之間，也就是“詞”與“詞”之間，總體沒有系統性的聯繫。這些存儲體是離散的，相對自由的，在存儲空間內本質上呈現渙散的狀態。因此，從這些角度看，“詞”是一盤散沙，這是不容否認的事實。但是，不能由此否認詞彙的系統性，因為“詞”已經是詞彙系統的高級結構，詞彙系統以“詞”的形式實現了存儲結構，也就實現了它作為語言材料倉庫的存儲功能。