

聯併 (*Unification*): 語法理論與剖析

Chu-Ren Huang

黃居仁

Academia Sinica

中央研究院

Proceedings of ROCLING I (1988)

R.O.C. Computational Linguistics Workshops I pp 29-54

中華民國第一屆計算語言學研討會論文集 29-54 頁

聯併 (*Unification*) : 語法理論與剖析

黃居仁

中央研究院歷史語言研究所

I、前言：自然語言處理與語言學理論

II、聯併：*Unification*

1. 何謂聯併

2. 以聯併爲本 (*Unification-based*) 的一些理論特性

3. 聯併在語法理論與剖析上之重要性

III、簡介以聯併爲本的一些語法理論

1. 概化詞組結構語法 (*GPSG*)

2. 詞彙功能語法 (*LFG*)

3. 以中心語驅動的詞組結構語法 (*HPSG*)，功能聯併語法 (*FUG*) 與其他

IV、結語：聯併及相關語法理論文獻導讀

I、前言：自然語言處理與語言學理論

「巴別魚：銀河系便車客指南定義如下：巴別魚是種外觀像水蛭的黃色小魚，很可能是整個宇宙間最詭異的事物。這種魚依賴寄主所在環境中所產生的腦波能量維生。它吸收腦波能量中各種非意識的頻率，用以供給自己養份。同時，他會把發自同一來源的有意識思維頻率及語言中心的腦神經訊號匯合；然後轉換成一種電脈衝的形式分泌到他寄主的腦中。此種生物行為的實際效應很明顯，任何人只要耳中塞入了一條巴別魚就可以立刻聽懂以任何語言說出的任何話語。你所聽到的語言形式實際上是得自分解你耳中的巴別魚餵入你腦中的脈衝程式。．．．．．

不幸的是，可憐的巴別魚雖消除了所有不同種族的文化間的溝通障礙，也同時間接導致了生物史上從未有過，更多更血腥的戰爭。」

——摘自道格拉斯·亞當斯著銀河便車客三部曲之一

(Douglas Adams, The Hitchhiker's Guide to the Galaxy)

何謂語言，人類如何理解語言，如何以語言溝通，如何習得語言，以至於語言的起源為何？這些都是人們一直想釐清的問題。困難的是，人類嘗試以思維解析語言，卻又無法不以語言表達思維，更無法確切定義出語言與思維的分野。所幸，這一世紀來語言學與資訊科學的發展走出了除了思維辯證外另一種研究語言的途徑。

索緒爾 (F. de Saussure) 通常被稱為近代語言學之父，他對語言研究最大的影響在於他把語言和常識，思維之間的糾葛分離，而把語言視為信號 (sign) 所構成的結構系統 (文學上的結構主義 structuralism 也源自此觀念)。索緒爾的信號 (sign) 並不是一個具體的標記，而是一個抽象的關係；是符號 (signifier) 和訊息 (signified) 之間的關係。在研究語言時運用其理論，我們不再把實際的語音 (或文字) 看成是語言的主體，他們只是符號 (signifier)。我們也不會把語言與語言所要表達的知識或思維混為一談，他們只是訊息 (signified)。語言的定義不在於這些實際存在的事物；因為這些實際的事物龐雜而難掌握。也不在於推究造成符號與訊息之間關係的理由，因為此種關係在自然語言中是任意而無規律可循的。其定義在於由符號與訊息之間關係所組成的有結構的系統。換句話說，理解符號與訊息之間的關係系統便是理解語言。

在語言學方面，索緒爾啓迪了把語言看成了承載訊息的系統的看法，而且把語言學研究的方向轉移成探究這個關係系統的結構。在自然科學方面，善農（Shannon, Claude E.）則在本世紀中發展了訊息理論（Information Theory，亦稱 Communication Theory）；建立了測量所有傳遞訊息系統訊息內容及效率的數學模式。以位元（bits）來計算訊息的模式，對後來電腦的發展有很大的影響。訊息理論雖然並不是針對語言，但可以應用於自然語言（參見 Pierce, 1980，第四、五章），也對語言學家的觀念產生了影響（如趙元任，見 Chao 1968一書）。因此，在索緒爾與善農之後，把自然語言當成是個承載訊息的信號系統的觀念有了，也有數學模式可計算研究信號系統。學者開始意識到研究自然語言也許可踏出傳統人文與社會科學的窠臼。

實際把自然語言和其他傳達訊息的系統放在一個規範模式下討論的當然是杭士基（N. Chomsky, 1957, 1965）。杭士基把人類對語言的知識（語法部份）看成是一套規律，經由此規律系統恰可產生所有合法語句。也就是說，在邏輯上，語言可由此套規律定義。此套規律所能產生的語句即是該語言中的語句；所不能產生的，即不是該語言中的句子。這樣的觀點稱為衍生語法（generative grammar），至今仍為大多數理論語言學家所遵循。更進一步的，杭士基把這些規律規範化之後建立了一個數學模式——杭士基階層（Chomskian Hierarchy）。杭士基階層並不是專為自然語言而設立的，而是為評量各種不同訊息系統（即規範語言，formal languages）的語法（即產生他們的規律系統）的複雜度而設立的。也就是說，在杭士基階層的定義中，及在往後資訊科學與理論中，語言（language）一字指的是任何可以這些設定的規律形式產生的信號系統，因此包括了電腦的程式語言（program language）。有了規範語言學（formal linguistics）以及數理語言學（mathematical linguistics），自然語言不但被當成一個有規律的訊息傳遞系統來研究；產生（或定義）這系統的規律也和人工規範語言的規律相比擬，在同一嚴謹的數學模式下研究。此種研究，在解決何謂語言或人類如何學習理解語言等觸及人類本質的問題或許尚未提供明確的答案，在工程的實際運用上卻立即開了新途徑。

杭士基在提出他的新語法理論（Chomsky, 1957, 1965）時，也正是科學界正在探索電腦應用範圍與威力時。在他的變形語法（transformational Grammar）中，基底結構只是一些免用語境（context-free）改寫律（rewriting rules）；而由基底至表層（surface）的變形（transformation）則是一些成份之間順序調換的描述加上施用環境的限制。任何用在電腦上的人工規範語言幾乎都用到改寫律；而調換順序也是電腦可輕而易舉執行的運算。因此，嘗試在電腦上利用變形語法解析或生成自然語言，幾乎是天經地義。而此類研究最顯見的應用，便是機器翻譯（machine translation, MT）。因此，在五、六十年代間，變形語法理論與機器翻譯蓬勃發展，氣勢似要在短期內解決所有有關語言之千古疑問。

不幸的是，機器翻譯一直沒有突破；更糟的是一九六六年美國國家科學院的一篇報告（ALPAC Report）斷言全自動，高可信度的機器翻譯不可行。機器翻譯因之在六十年代末期及七十年代間逐漸沈寂了下來。而理論語言學界也在經過衍生活語意學（*generative semantics*）的一場大爭辯，與佩德斯與李奇（*Peters and Ritchie, 1973*）證明變形語法的能力太過龐大無法駕御之後，慢慢放棄了變形的路徑。這些挫敗，對語言學在自然語言處理上的用處，以及計算語言學的應用與理論語言學發展的關係提出了疑問。

最早，資訊學者天真地以為把一部完善的字典放入電腦即可完成自動翻譯機，根本就沒想到語言分析或語法理論。語言學家當然也很難想像把複雜的語言現象套入規律嚴謹生硬的計算模式裡。可是語言學者與從事人工智慧（*Artificial Intelligence*）的資訊學者實際上是在探索同一個問題的兩面。語言學家所要描述的，根據杭士基定下的目標，是人類的語言能力（*Linguistic Competence*）。所謂語言能力，包括了對語言的掌握、運用、理解、習得等。假設人們腦中有一套其母語的語法，並藉以使用理解語言，語言學家的目標便在於正確描述這些語法的規律。語言學家如何去驗證其理論呢？唯一的辦法是拿假設語法所預測的結果和實際的語言現象比較。換句話說，語言學理論嘗試描述一個極重要的心智活動，卻無法真正去檢查驗證其描述，只能憑檢查比較理論模式運作的結果。這個方法可能的問題包括了理論架構不夠完全，以及推導預測的過程不夠嚴謹詳備。蓋志達等（*Gazdar et al. 1985*）在概化詞組結構語法（*Generalized Phrase-Structure Grammar*）一書中就曾指出近代語言學理論發展雖日新月異，卻幾乎無學者肯定下心來寫任何語言的完整語法。以自然語言這麼複雜的系統看來，若只定了語法系統的一部份，而無完整語法的其他部份來配合，其預測實難以信賴。另一方面，資訊學者欲賦予電腦類似於人類思維及理解語言的能力。所採用的方式是設計一套機器可以執行的運算方式（*Algorithm*），然後試驗機器是否能執行一些人類思維行動。這個方向的潛在問題在於偏重於機械化的運算模式及電腦處理大量資料的能力，以及只執行定義太狹隘的人類思維。以自然語言處理為例，偏重工程方法的結果往往是以便宜行事（*ad hoc*）的方法解決局部的問題，反而導致他處的局部問題。永遠是見樹不見林。六、七十年代計算語言學的挫敗或多或少皆源自上面兩類缺失。資訊學者雖應用了語言學理論（如變形語法），卻只把它當作是一種運算方式的藍本。並未去探究理論與剖析架構甚或電腦運作所依據的規範語言理論之間的關係。而理論語言學者，雖有杭士基立下了良好的規範理論基礎，卻只把這些理論當成採用變形的藉口，而未深探規範理論和語言學理論之間的關係。

上文中提到資訊科學中的自然語言處理與理論語言學中的語法理論實際上是由不同的方向來探索人類的語言能力。語言學家所求的，是一套嚴謹完備的理論

規範，不但要正確描述語言現象，更要解釋其所以然。即杭士基所說解釋上的完備性 (*explanatory adequacy*)。而資訊學家的目的是在機器上仿效人類的語言能力，並運用仿效成功的能力為人所利用。語言學家的問題在於光憑直覺的語感及少數規律無法驗證語法的全貌；而光靠論證也很難判定理論是否正確反映了黑箱子 (*Black box*)—腦中的活動。資訊科學的難處則在於光靠加快運算速度與記憶容量等工程技術來解決處理自然語言時遭遇的困難。為什麼科學家用這麼強而有力的工具仍解決不了人類輕而易舉，家常便飯的語言現象呢？答案極可能在於人類具有一套特別適合自然語言的解析方式，而這套解析方式與目前電腦上習用的機械式運算法不同。唯有套用此種解析方式，自然語言處理才能成功。這套解析方式很可能就是語言學家所追求的具解釋上完備性的語法。既然兩個研究方向的著眼點都在同一套語法理論（或解析方式），自可互通有無。語法理論欲描述人類的心智活動，但心智活動是黑箱子，無法觀察，只能靠比較驗證實際與預測的結果。但語言系統太龐雜，語言理論往往無法窺其全豹。電腦上的自然語言處理正好提供了模擬測試語言理論的最好工具。資訊科學欲模擬人類心智活動，又發覺純粹由工程觀點出發的解決方式並不可行；故需要真正針對人類語言現象的規範理論作為研究的指針。可是語言學家與資訊科學家長久以來一直使用不同的「語言」。變形語法運用在機器翻譯上的失敗恰可說明此一現象。語言學理論儘可以在抽象的層次上發揮，其理論卻未必可能在任何規範系統內執行。相反的，若缺乏一個基本的架構或共通的基本的運作 (*operation*)；兩套不同理論執行的優劣很可能取決於其工程建構 (*implementation*) 技術的優劣而非取決於理論。因此，把語言學理論應用到自然語言處理上，用以選擇驗證更好的模式，其前提實在於理論與理論之間及理論與實際執行之間是否能有共同的基本運作。這個基本運作必須在觀念上極簡單，又要能在應用上足以執行不同語法理論及不同剖析運算法中要求的複雜規律。近十年來理論與計算語言學的發展逐漸找到了一個中心的運作，那就是聯併 (*Unification*)。

II. 聯併：Unification

1. 何謂聯併

聯併其實是個很簡單的觀念，有點像是集合裡的聯集 (*set union*)。聯集取的是各集合裡所有不同元素所成的集合。如 (1)。

$$\begin{aligned}(1) \quad & \{\text{加州大學、清華}\} \cup \{\text{中研院、清華}\} \\ &= \{\text{加州大學、中研院、清華}\}\end{aligned}$$

聯集裡一個很重要的觀念是集合裡每個元素都是單獨不可分的個體。例如 $\{\{1\}\}$ 和 $\{1\}$ 這兩個集合各自只有一個元素，而且這兩個元素不相同($1 \neq \{1\}$)。因此 $\{1\}$ 聯集 $\{\{1\}\}$ 的結果有兩個元素 1 與 $\{1\}$ ，即 $\{1\} \cup \{\{1\}\} = \{1, \{1\}\}$ 。如果我們把序對(ordered-pair)列入考慮，得到的結果也一樣。

$$(2) \quad A = \{ \langle a, 1 \rangle, \langle b, 2 \rangle, \langle c, 3 \rangle \}$$

$$B = \{ \langle a, 2 \rangle, \langle c, 3 \rangle \}$$

由聯集的觀點看來， A 有三個元素： $\langle a, 1 \rangle$ 、 $\langle b, 2 \rangle$ 、 $\langle c, 3 \rangle$ 。而 B 有兩個元素： $\langle a, 2 \rangle$ 、 $\langle c, 3 \rangle$ 。而 A 與 B 有一個共同的元素 $\langle c, 3 \rangle$ 。若定義 C 為 A 與 B 的聯集，則 C 有四個元素。

$$(3) \quad C = A \cup B = \{ \langle a, 1 \rangle, \langle a, 2 \rangle, \langle b, 2 \rangle, \langle c, 3 \rangle \}$$

但是，很顯然的，序對有其內部結構。邏輯上最常見的用法之一是利用序對表達某種關係(relation)。序對表達的關係可以是數學關係，如 $>$ 、 $<$ 、 $=$ ，也可以是實際生活世界中的關係。比如說，我們可以學術機構代替 a 、 b 、 c ，而以地名代替 1 、 2 、 3 。同時，我們可把序對兩成份之間的關係定義為機構與其所在地之間的關係。如此一來，集合 A 成了某些常識訊息的表達方式 A' 。

$$(4) \quad A' = \{ \langle \text{中研院, 南港} \rangle, \langle \text{加州大學, 柏克萊} \rangle, \langle \text{清華, 新竹} \rangle \}$$

把序對如此解釋之後，我們可清楚看出聯集的限制。聯集把所有集合的元素視為不可分解的原子(atom)，而無法顧及其內部結構。維持上述機構—所在地的定義，以同樣的實際機構名與校名替代 a 、 b 、 c 、 1 、 2 、 3 等符號，可得到 B' 、 C' 如下。

$$(5) \quad B' = \{ \langle \text{中研院, 柏克萊} \rangle, \langle \text{清華, 新竹} \rangle \}$$

$$(6) \quad C' = A' \cup B' = \{ \langle \text{中研院, 南港} \rangle, \langle \text{中研院, 柏克萊} \rangle, \langle \text{加州大學, 柏克萊} \rangle, \langle \text{清華, 新竹} \rangle \}$$

由常識判斷， A' 和 B' 表達的訊息有所抵觸。即中研院不可能同時又在南港，又在柏克萊。聯集雖把 A' 和 B' 兩者的訊息滙合整理了，所得的結果 C' 卻不能代表一個實際有用的訊息體。把 A' 和 B' 看成是人類語言處理的兩項訊息，聯集雖可描述牽涉訊息的總數，卻很顯然無法描述人類如何整合運用訊息。因此，我們需要不同的數學運算來描述語言現象。研究學者們發現聯併同於描述訊息傳遞相當適當。基本上，聯併與聯集最大的不同在於訊息有抵觸時聯集兩者均收而聯併根據定義無法成立，即聯併失敗(fail)。在文獻中，聯併均以等號“=”標示。為避免混淆，本文由此起將等號保留給聯併，而將各種運算的結果以箭號“ $\rightarrow$ ”標示。

當聯併的對象所含元素均為原子 (atom, 即不可分割的個體) 時聯併等於聯集。這是因為原子元素 (atomic element), 要不就相等 (即指同一對象), 要不就不相等 (即指不同對象), 描述內容不可能相抵觸。因此, 聯集和聯併的差異在涉及上面提到的關係 (relation) 或更複雜的層次時才彰顯出來。本文下節中將討論到, 以聯併為本 (unification-based) 的理論重要特性之一是運用複雜特徵 (complex feature)。基本上來說, 複雜特徵是以屬性—特徵值配對 (attribute-value) 的形式出現。而這些特徵值本身未必是原子性的 (atomic), 他們也可能是多層屬性—特徵值組成的特徵構造 (feature structure, 或稱 feature matrix, functional structure 等)。例 (5) 與 (6) 中的關係可改用特徵構造的方式來表現如下。

- (7) a. $A' \rightarrow$

「	中	研	院	南	港	」
	加	州	大	學	柏	克
└	清	華	新	竹	└	┐
- b. $B' \rightarrow$

「	中	研	院	柏	克	萊	」
└	清	華	新	竹	└	┐	┐

特徵構造的左邊是特徵名 (feature name), 或稱屬性 (attribute); 而其右邊則是特徵值 (feature value)。故整個特徵構造可看成是由屬性映至特徵值的一個部份函數 (partial function)。上例中屬性為學術機構, 而特徵值為其所在地。聯併 A 與 B 及聯併 A' 與 B' 的結果分別表示如下:

- (8) a. A

「	加	州	大	學	」
└	清	華	└	┐	┐
- b. B

「	中	研	院	」
└	清	華	└	┐
- c. $A = B \rightarrow$

「	加	州	大	學	」
	中	研	究	」	
└	清	華	└	┐	┐

- (9) $A' = B' \rightarrow fail$ [無結果]

(因 [中研院 南港] 與 [中研院 柏克萊] 無法聯併)

由上例得知 A 聯集 B 的結果等於 A 聯併 B 的結果； A' 聯集 B' 的結果則和 A' 聯併 B' 的結果不相等，主要是因為 A' 和 B' 需有相衝突的訊息，聯併無法完成。因此，聯併可看成是一個單純結合訊息的運算，在無互相衝突的訊息時結合特徵結構中承載的所有訊息。聯併也可看成是對等限制 (*equality constraint*)。因為甲聯併乙的結果並非是一個代表新事物的特徵結構；甲聯併乙的結果既是甲，也是乙；只是我們知道甲與乙這兩個特徵結構均描述同一件事物；而由甲或乙的觀點來看，則是經由聯併增加了訊息。這也是為什麼習慣上聯併用等號表示。因為聯併在應用時就是表示各成份內含的訊息是描述同一事物。比如說，某人可能知道武俠小說「神鵰俠侶」作者筆名是金庸，也知道香港明報社的社長是查良鏞，卻從未把這兩件訊息連起來。可是當他知道了查良鏞的筆名即是金庸之後，他也知道了明報社的社長寫了「神鵰俠侶」，及其他許多訊息。這一現象，我們可以聯併描述。(請注意此處的表達方式只用於說明聯併的功能，而不代表任何表達知識 (*knowledge-representation*) 的模式。) 換句話說，我們把描述查良鏞的特徵結構和描述「神鵰俠侶」作者的特徵結構聯併，得到(11)。

(10) a. 神雕俠侶 = 「分類 武俠小說」
 「作者 [筆名 金庸]」

b. 查良鏞 = 「筆名 金庸」
「職業 明報社社長」

(11) 神雕俠侶 =

分類	武俠小說
作者=查良鏞	筆名 金庸
	職業 明報社社長

由 (10a) 與 (10b) 中我們發覺「神雕俠侶的作者」這個特徵結構(10a 中包在裡面的小結構)和「查良鏞」這個特徵結構描述的是同一個對象，因此與聯併加以結合，聯併之後得到 (11)。注意到聯併雖未增添任何新的屬性—特徵值的配對，卻提供了 (10a) 與 (10b) 中間無法表達的訊息。以聯併描述這類訊息的整合簡潔且合於直覺。若用其他傳統的邏輯作法，如 (12) 中的兩個三段式論證，則顯得相當笨拙。

(12) a. 神雕俠侶作者的筆名是金庸

查良鏞的筆名是金庸

因此 神雕俠侶作者是查良鏞

b. 查良鏞是明報社社長

神雕俠侶作者是查良鏞 → (得自 12a)

因此 神雕俠侶作者是明報社社長

由上面的例子看出聯併用於描述傳載訊息的系統的能力遠高於我們所熟悉的邏輯。自然語言又已知是個效率極高，又能精確表達各種複雜意念的訊息結構。由此看來，聯併當然可能是描述語言的良好工具。另一個不容忽視的理論假設是語言的合成性 (compositionality)。語法 (syntax) 一向被視為是小語法成份 (constituents) 組合成大語法成份，如詞組 (phrases)、句子 (sentences) 的規律。而語意學上的正統假設則是句子的語意也是經同樣的步驟由小成份的語意組合而成。把聯併視為語法與語意上的基本運作恰可描述此先後觀念。聯併同時執行語法與語意訊息的組合。由於語法語意所依賴的規範運作相同，其平行結構也可得到解釋。

介紹了聯併的基本觀念後，本節將以聯併的定義作為結尾。定義 (13b) 取自蓋志達等人 (Gazdar et al. 1988) 的 "Category Structure" 一文。定義原是為詞類範疇 (category) 而設。但我們把原定義中的範疇的部份以特徵結構代入。

(13) a. 定義： α 為一特徵結構，若且唯若 α 可以

$\vdash f_0 \ v_0 \neg$

$| f_1 \ v_1 \ |$

$| \ . \ . \ |$ 的形式表示， $n \geq 0$ 。而且

$| \ . \ . \ |$

$\vdash f_n \ v_n \neg$

甲. 特徵名 $f_0, f_1 \dots f_n$ 為原子。特徵值 $v_0, v_1 \dots v_n$ 為原子或特徵結構。

乙. $\alpha(f_i) = v_i$

b. 定義：聯併

甲. 若 a 與 b 均為原子 (atom), 則 a 聯併 b 為 a
(記為 $a=b \rightarrow a$), 若且唯若 a 等於 b 。

乙. 若 α 與 β 均為特徵結構, 則

(一) 若特徵結構 $[f, v]$ 屬於 α , 但 $\beta(f)$ 的值未經定義, 則 $[f, v]$ 屬於 α 聯併 β 。

(二) 若特徵結構 $[f, v]$ 屬於 β , 但 $\alpha(f)$ 的值未經定義, 則 $[f, v]$ 屬於 α 聯併 β 。

(三) 若特徵結構 $[f, v1]$ 屬於 α , $[f, v2]$ 屬於 β , 則 $[f, v1 = v2]$ 屬於 α 聯併 β , 若 $v1 = v2$ 有定義, 而 α 聯併 β 不能經其他方式定義。

(13) 是一個遞迴定義 (recursive definition)。乙(一)、乙(二)兩項基本上相當於聯集的定義。即若某特徵值 (此特徵值亦可能是一個複雜的特徵結構), 只出現在聯併兩成份中的一個, 則該特徵亦屬於聯併的結果。但若該特徵值在兩成份同時出現, 又非原子, 則必須看此兩特徵值的聯併是否成立 (乙(三))。如此一層層下去, 檢驗兩成份所帶訊息是否有衝突。若皆無衝突, 則聯併成功。

2. 以聯併為本 (unification-based) 的一些理論特性

在所有以聯併為本的語法理論中除了都使用聯併這個運作外, 也同時有些共同的假設。而這些假設所造成的特性, 主要在於使得聯併這個運作能更有效率的運用。施別 (Shieber 1986) 把這些假設稱為設計特色 (design feature), 歸納如下:

- <1> 以實際語句為準 (surface-based): 變形語法盛行造成了衍生理論一定要有多層結構的錯誤印象。可是多層的抽象結構增加了實際剖析的複雜度。更重要的, 除非全盤接受了這些變形理論未經證明的假設, 這些看不見的抽象的結構或成份的存在根本無法證實也無法反駁。由工程與直覺的觀點看來, 除非能證明無任何理論可憑實際語句正確剖析自然語言, 否則並無假設抽象層次之必要。

- <2> 訊息性的 (*informational*): 語法的功能在描述語句如何和訊息結合。上文中提到聯併最大的功能之一在描述訊息的結合, 以聯併為本的理論自然也利用了這個運作的長處。強調訊息性 (如 Kay 1985) 事實上也擴展了杭士基對規範語法較狹隘的定義。根據杭士基的定義, 語法的作用在判定任意語句是否是該語言的合法語句。聯併為本的語法則在定義合法語句時也定義合法語句所代表的訊息。
- <3> 演繹性 (*inductive*): 聚積訊息而非縮減訊息。以遞迴定義決定小語法成份如何組成大語法成份, 也同時規定小成份所承載的訊息如何相加成大成份所承載的訊息。
- <4> 宣示性 (*declarative*): 定義語法成份和訊息之間結合的關係時只規定合法的組合而不規定如何結合。這特色使得聯併為本的語法理論不受限於 (剖析的) 運算方式 (*algorithm-free*)。這將在下文有關剖析的部份再做討論。
- <5> 以複雜特徵為本 (*complex-featured-based*): 上面提到聯併為本的語法重點在描述實際語句如何承載訊息。至於如何以嚴謹的規範系統表示訊息; 各理論所一致採用的是複雜特徵。最主要的特色在於任何屬性的值 (或特徵值 *feature-value*) 本身都可能是另一個特徵結構。依此方式可表達目前需要描述的語法或訊息現象。

另外, 在其他討論聯併的文章裡 (如 Sag et al. 1986), 下列兩個特性也經常被提起。

- <6> 單調性 (*monotonic*): 語法規律只增加訊息, 不減少或改變訊息。這當然是因為聯併不允許相衝突的訊息存在。這個特性也可看成是上面第一與第三特性的組合。因為以實際語句為準, 故不在衍生過程中改變結構也不會消去任何訊息。因為是演繹性, 故訊息只增不減。
- <7> 承載部份訊息 (*partial information*): 由聯併的定義可以看出, 一個語句或其他語法單位所承載的訊息可分佈在各個成份中。故各成份承載了部份訊息。這和施別談到的第二個特性—訊息性—有很大的關係。根據純粹規範語法的定義, 任何字串要是不能剖析成句子 (*S*), 即被語法所排除。可是所有不成句的片段中可能有名詞組 (*NP*), 可能有動詞組 (*VP*), 甚至單字。這些在實際使用語言時都承載了些訊息, 而不是毫無意義。一個只接受或排除句子的語法顯然有所缺失。聯併為本加上使用了雜特徵使語法可以精確代表每個大大小小的語言片斷所傳達的訊息。

總結以上聯併爲本語法理論的特色，事實上除了代表語法理論與自然語言原理兩個研究方向整合的結果外，也代表了部份語言學家對抗士基學派（由變形語法理論 TG 到管制理論 GB）所抱持的假設的檢討。至於聯併對這些理論及他們在剖析上的運用有何重要性，將在下節討論。

3. 聯併在語法理論與剖析上之重要性

聯併在以聯併爲本的語法理論中扮演的角色是各理論共有的基本運作。有了這個基本運作之後，表面上看起來南轅北轍的理論可以作實質上的比較。也就是說，各理論固然有其各具特色的規律與表現方式，如概化詞組結構語法 (GPSG) 的特徵共存限制 (FCR, *Feature Cooccurrence Restriction*) 與詞彙功能語法 (LFG) 的限制性功能等式 (*constrain equations*) 在表面上看起來即大不相同。但若都化解成聯併的基本運作，我們便可討論其實質上的異同。更重要的，往往在某一理論中效果良好的分析可藉聯併的媒介引入其他的理論中。伯樂與薩格 (Pollard and Sag, 1987) 就給了一個相當成功的例子。儘管目前語法理論的派別仍然很多，每個理論也仍各具特色；但各個以聯併爲本理論之間的研究成果已可互相流通。而不是像以前理論一有更新即「各說各話」，很難利用前人或不同理論研究的成果。

另一方面，針對剖析而言，聯併爲本的系統則有下列的優點：

- <1> 不限於剖析進行的先後順序 (*order-free*)：由於聯併爲本的理论是演繹性的，而且也是單調性的；不論聯併由那個方向開始執行及其先後次序爲何，其加成的結果一致。在構建剖析程式及人類理解語言的模式時，我們當然不希望因剖析的起點不同而導致不同的結果。與這相關的是下面一個特性。
- <2> 不限於剖析的運算方式 (*algorithm-free*)：資訊科學家早就知道不同的運算方式效率不同，解決問題的方式也不同。如果用於剖析的語法理論規範性質迥異，很可能影響其所能適用的剖析運算方式，相對的也影響了剖析的結果。類似的，由於自然語言處理的重要目的之一在於比較各語法理論(包括其預測語言現象的結果以及是否爲有效的語言處理模式等)。這種比較只有在各理論能使用相同的運算方式；而且語法表達模式不受選定的運算方式影響時才有意義。聯併不破壞結構也不更改訊息，其執行順序爲不影響結果；故聯併爲本的剖析不受限於運算方式的選擇。

<3> 不限於語法理論 (*theory-free*)：相對於前項特色，聯併為本的剖析規範系統建立之後可執行任何聯併為本的語法理論；不會因理論不同而不適用。我們因而可在同一系統上建構根據不同理論設立的剖析器，並真正深入比較各理論得失異同。

<4> 宣示性 (*declarative*)，上個小節中提到聯併為本的語法理論只規定何種成份與訊息的結合為合法，卻不規定如何去計算執行何者合法。由設計剖析器的觀點，或由習得成理解自然語言的程式看來，宣示性都極重要。一般而言，非宣示性的系統(大部份是程序性的*procedural*)在局部的訊息(如詞彙訊息)更動或增加時至少得決定新訊息處理的方式；或甚至要更改整個系統的處理程序。而宣示性系統則不受局部更動之影響。試以中文量詞 (*classifier*) 為例說明兩類系統 (或語法理論) 之差異。

- (14) a. 一本書
b. * 一個書

- (15) a. * 一本理論
b. 一個理論

我們知道中文量詞與名詞中心語之間有共存限制；某個名詞可用某幾個量詞修飾，不可隨便配對使用。我們在語法中可以用描述性的方法規定那幾類名詞和那幾個量詞可以一起出現，然後把名詞分類。比如說，我們可以說「書」這個名詞是屬於第*N*類，而這類名詞只能用「本」這個量詞。至於在語法理論上或在剖析系統中，較早用的辦法是去檢查配對的情形(如變形語法中的條件*conditions*)。比如說，語法可規定第*N*類名詞出現時必須檢查其量詞是否為「本」，若不是「本」則不合語法。這是程序性的。因為語法不但要指定什麼情況下執行檢查，也要規定如何辨識名詞組中量詞的位置，相反的，宣示的系統，如聯併為本的語法，則只需要詞彙中記載一項訊息就可描述整個現象，不用另行規定。

(16) 詞項記載：

a. 書: $\left[\begin{array}{l} \text{pred 書} \\ \text{mod [class ben]} \end{array} \right]$

b. 本: [class ben]

c. 個: [class ge]

以 (16) 為例,不論聯併的方向細節是由另一套特徵結構決定(如HPSG)還是由詞組結構律上附加的等式決定(如LFG),結果均一致。即當「本」與「書」結合時聯併成功,得到特徵結構 (17)。而「個」和「書」結合時聯併失敗,因CLASS (代表classifier)這個特徵值在兩個詞彙記載項下互相衝突。〔(16).(17)僅用於說明,並未正確遵循任一理論。〕

$$(17) \quad \text{一本書} \quad \left[\begin{array}{c} \text{PRED 書} \\ \text{MOD} \left[\begin{array}{c} \text{CLASSSS BEN} \\ \text{NUM} \quad 1 \end{array} \right] \end{array} \right]$$

由上例看出,聯併為本的語法(或剖析系統),不需規定任何條件或檢查步驟。(14b)和(15a)不合法這個事實是依靠聯併的基本定義來解釋。即所有不合法語法現象均是因為不同成份承載了互不相容的訊息而導致聯併失敗。其他熟知的語法現象,如語意的選擇限制(selectional restriction,如「殺」字的賓語必須有生命。*「張三殺桌子」),人稱與數的對協(agreement,如 * children cries)等等,聯併為本的理論處理方法並無二致。均是靠選擇正確的特徵,然後在詞彙中或規律上作適當的宣示。聯併所控制的結合過程自然會剔除含不相容語法訊息的字串或語句。

上面的說明我們可看出宣示性理論最大的特性在於其解過程不因語法而異,因此在語法現象改變或改變某個分析時,只需把詞彙或規律的內容作適當地調整即可得到效果。不需要考慮到更動剖析程序及各程序之間所可能產生的複雜互動。在描述現象上,宣示性系統較簡單。而且宣示性的系統可以準確顯出每一項規律或特徵更動的影響,從而提供此項更動的正確評估。而程序性系統則常因規律施用順序,與規律之間結果的互相牽制,使得更動某項規律的正確與否往往不會在整個系統運作的結果上明白呈現出來。

- <5> 容許部份剖析 (partial parse): 聯併為本的剖析系統著重在表達字串所傳載的訊息(即上節所提的訊息性)。故而相對於每個(語法所容許的)成份單份均有其承載的訊息。剖析器在解析字串同時也建立了處理字串所表達的特徵結構。因此即使剖析失敗(即在某層次發現衝突訊息)或不能成句;都還保存了正確剖析部份得到的訊息。也就是,相對於任何剖析的片斷均有其特徵結構。這與傳統剖析器的定義只在於判斷字串是否成句,在實際應用上及在理論上均有好處。實際應用上並非所有

剖析的對象均為完整的句字（如標題等）。而且即便剖析失敗我們也希望檢查已剖析的結果。更進一步說，在工程上我們期待不懼失敗 (*fail-soft*) 的系統。不希望因某個小細節錯誤，而使剖析完全白做。故必須保留每一步驟之相對成果。在理論上，人類對於片斷語句均能了解，得到說話者欲表達的信息；甚至對於不合法語句（如外國人說中文）也可理解其部份訊息。因此執行部份剖析的能力對剖析系統而言是一大優點。

綜上所論，運用以聯併為本的剖析系統不限於剖析先後順序，不限於運算法，而且可採用各種聯併為本的語法，不限於理論模式。加上聯併為本的剖析由於無方向性，適於使用威力極大的平行處理 (*parallel processing*)，也就是說，提供了剖析上自由的應用空間，使計算語言學者得以使用各種可能的理論資源。

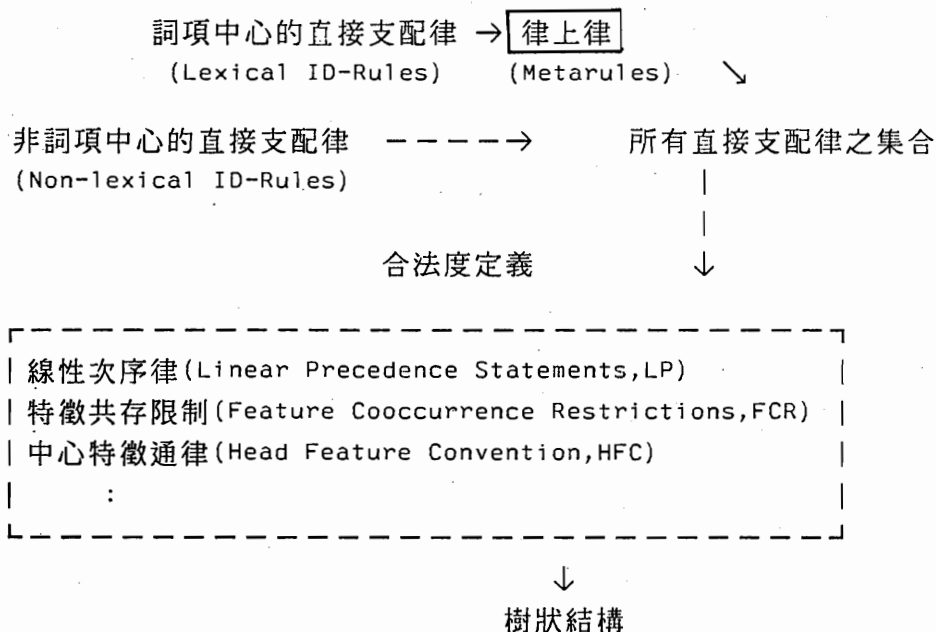
Ⅲ、簡介以聯併為本的一些語法理論

1. 概化詞組結構語法 (*Generalized Phrase Structure Grammar, GPSG*)

概化詞組結構語法（以下簡稱GPSG）最大的特色在於維持免用語境語法 (*Context-free Grammar*)。免用語境語法剖析時難度低，而且有許多有效的運算方式（如 *Earley's Algorithm*）可用。可惜的是語言學家幾乎未經嚴謹證明，便判定自然語言超出了免用語境語法的管轄範圍。除了免用語境語法在規範理論及剖析難度上的好處外，GPSG 運用此類語法的兩個最有力論證分別是蒲倫與蓋志達 (*Pullum and Gazdar 1982*) 證明前人所舉自然語言超過免用語境語言的例子在數學上均不成立；以及蓋志達 (*Gazdar 1981*) 證明如英語疑問詞提前等的無限對應 (*unbounded dependencies*) 不但可以免用語境語法處理，而且解決了一些變形語法無法處理的問題。儘管後來施別 (*Shieber 1985*) 提出了一個自然語言超出免用語境語法的實際例子，即瑞士德文的交互序列對應 (*Swiss German Cross-serial Dependencies*)，黃 (1989) 也利用國語的反複是非問句提供了相似的證明。事實是絕大多數自然語言中的絕大多數現象仍屬於免用語境語言。而 GPSG 所採用的嚴謹分析法仍是已知理論中相當嚴謹有效的。

GPSG 只有單層結構，直接由詞彙詞組律產生實際語句；同時由依據語法律而定的解釋律直接產生類於孟太奇語意學 (*Montague Semantics*) 的邏輯代表形式。理論上，GPSG 的規律是語法樹的合法度限制 (*tree-admissibility conditions*)，而非實際的衍生律。然而在目前實際應用上（包括分析語言現象與剖析），這樣的區分也許並不重要。

(18) 概化詞組結構語法 (GPSG) 的語法：[改編自 Sells 1985]



(18)中顯示 GPSG 只有由直接支配律 (*immediate dominance rules, ID-Rules*) 到語法樹一層結構。律上律管的只是直接支配律之間的內部組織；而合法度定義 (*well-formedness Definitions*) 則只宣示語法准許 (*admit*) 什麼樣的樹狀結構。由於 GPSG 所有的語法訊息都掛在樹狀結構上，包括特徵在節點 (*node*) 之間傳遞。故由聯併為本的理論看來，GPSG 管制可能的樹狀結構即是管制可能的特徵 (訊息) 結構。

除了完全使用免用語境的規律外；GPSG 有幾項特色或為其所獨有，或為其所首用而為其他聯併為本的其他理論循用。筆者認為最重要的三項特色為視範疇為特徵群 (*feature bundle*)，ID/LP 模式，與特徵傳遞的限制 (*feature percolation constraints*)。

GPSG 的描述力強於以前所有的免用語境語法，其關鍵在於善用特徵的觀念。以往把 *S*，*NP*，*VP* 等視為單純的詞類範疇標記；只為分類名稱，而不賦與任何附加意義。但 GPSG 把詞類範疇視為特徵群的縮寫，而每個特徵均有其特別的語法意義，影響語法行為。舉例說明 *NP* (名詞組) 在 GPSG 中是下列特徵群的簡寫：

[<*N*, +>、<*V*, ->、<*BAR* 2>]

<*BAR* 2> 這個特徵表示這個範疇是最高投射 (*maximal projection*)；即在語法上可獨立充當論元。*N* 和 *V* 則是用以定義主要詞類 (*major categories*) 的兩個特徵。形容詞和名詞均有 <*N*, +> 這個特徵，因此可解釋他們有時可在對等位置出現，譬如英文 '*He is a linguist and proud of it*'。把語法訊息細分成特徵後，便可藉特徵的傳遞來描述一些複雜的現象。

管制特徵傳遞的限制在 GPSG 中佔很重要的地位；所管制的大概可分為每個範疇上不同特徵之間的共現關係，母節 (*mother node*) 和子節 (*daughter node*) 之間相同特徵之間的關係，以及姊妹節 (*sister nodes*) 之間特徵的關係等。GPSG 就是靠規定這些局部 (*local*) 的關係，分解了許多非局部的現象；因而免除了語法中似乎是需用語境 (*context-sensitive*) 的部份。如語言學界熟知的 *wh* 問句長距對應 (*long-distance dependency*，如英文 'Who do you think will win the presidency this year?' 就是以 SLASH 這個特徵來描述。SLASH 的特徵值為一範疇，所以可標示其為一 NP 以及其他與 NP 相關的訊息。在 'will win the presidency this year?' 這個子句上以 SLASH 記載所缺的主語名詞組，然後這個特徵循管制特徵的通律一層一層往上傳，直到傳到最高層，找到 'who' 這個詞組恰可滿足其特徵的範疇值為止。特徵通律中最重要的是 '中心語特徵通律' 規定中心語 (*Head*) 的某些特徵必須和其母節相同 (比如說名詞組的中心語必是名詞，可看成 $\langle N, + \rangle$ 、 $\langle V, - \rangle$ 兩個主要詞類特徵同值)。這個通律在其他理論中也採用，儘管其規律式未必相同 (如 LFG 在中心語上標明 ' $\uparrow = \downarrow$ ' 這個等式)。

ID/LP 是直接支配 (*immediate dominance*) 和線性次序 (*linear precedence*) 的簡稱。普通的詞組律，如 (19)，表達了兩種訊息，即 A 直接支配 b、c 兩個成份與 b 出現在 c 之前。

$$(19) \quad A \rightarrow b \ c$$

GPSG 認為詞組律中表達的兩項訊息是可區分的，而且在語法理論中具有不同的意義，故提議把直接支配與線性支配兩項訊息抽離。直接支配律是局部的，管那個母節可支配那些成份；線性次序則是全面的，是範圍及於所有直接支配律的一種合法度限制。對語序自由 (*free word order*) 的語言而言；只要不給任何線性次序律，直接支配律即可產生所有可能詞序，如 (20)a、(20)b。可是對有嚴謹詞序的語言，如中、英文；以特徵入線性次序律可決定許多詞序現象，如 (21)a、(21)b。可見 ID/LP 模式是描述語法共通現象的利器。

$$(20) \quad \text{ID-rule}$$

$$\begin{array}{lll} \text{a. } A \rightarrow a, b, c \Rightarrow \text{產生} & \text{b. } A \rightarrow a \ b \ c & A \rightarrow a \ c \ b \\ & A \rightarrow b \ c \ a & A \rightarrow b \ a \ c \\ & A \rightarrow c \ a \ b & A \rightarrow c \ b \ a \end{array}$$

$$(21) \text{ a. } [\text{SUBCAT}] < [-\text{SUBCAT}] \quad (\text{LP 規律})$$

$$\text{VP} \rightarrow \text{V}, \text{NP}$$

$$\text{PP} \rightarrow \text{P}, \text{NP}$$

規定 $\Rightarrow$

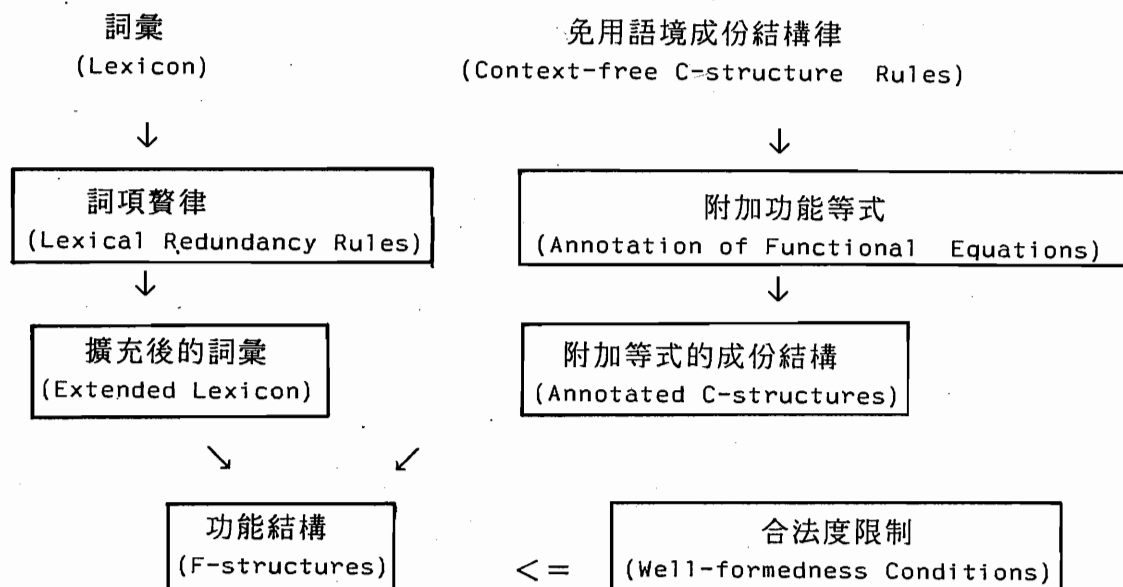
$$\text{b. VP} \rightarrow \text{V NP}$$

$$\text{PP} \rightarrow \text{P NP}$$

2. 詞彙功能語法 (Lexical-functional Grammar, LFG)

詞彙功能語法 (以下簡稱 LFG) 顧名思義有兩大特色：一是以語法功能 (grammatical function, 如主語 SUBJECT, 賓語 OBJECT) 為基本概念, 二為倚重詞彙。它與 GPSG 很大的不同之處在於除了樹狀的成份結構 (constituent structure, C-structure) 外, 另有功能結構 (functional-structure, f-structure)。而且大部份合法度的判斷都在功能結構上進行。功能結構事實上只是上文中提到的特徵結構的一個變體, 用以表達語法訊息。故 LFG, 如 (22) 所示, 比 GPSG 多了一個語法表達階層。

(22). 詞彙功能語法 (LFG)



LFG 理論中值得特別注意的幾個特色包括了詞彙贅律, 附加在成份結構上的功能等式, 以及在功能結構上的合法度限制。

變形語法中以變形解釋的一些相關現象, 如主被動式的關係、主題句等, 在 GPSG 中是以詞組律與詞組律之間的律上律來表示, 卻發現有衍生能力上的潛在問題, 將在下節中討論。LFG 則使用詞項記載之間的詞彙贅律表達。詞彙贅律有兩項主要功能, 改變詞的形態與改變論元結構。以被動詞彙贅律為例, 詞項最大的改變是減少了一個論元; 形態上, 以英文為例, 則是增加了被動詞尾, 通常是 -ed。

(23) 被動詞彙贅律

- a. $SUBJ \rightarrow \phi / BY\ OBJ$
 $OBJ \rightarrow SUBJ$
- b. $V \rightarrow V-ed$

採 (23) 這樣的詞彙贅律的結果便是認為主被動式之間 的關係在於他們有相同的論元結構 (*argument structure*)。而且認為這樣的關係是由詞彙決定的，而不是由樹狀結構或詞組律決定的。因此 LFG 可把主被動式論元改變與動詞形態改變在同一條規律中說明，也可以解釋為什麼少數動詞在主被動式的對應上有例外（語法律應是絕對而無例外的，而詞彙律則往往因詞項變異而有例外）。事實上，把解釋語法現象的能力轉移到詞彙中是自七十年代以來各語法理論共有的大趨勢。也是大多數聯併為本理論所採行的方法。

詞組律上附加的功能等式其功用其實在標定執行聯併的成份。在 LFG 中同時也是由成份結構構建功能結構的依據。

$$(24) \quad VP \rightarrow \quad V \quad \quad NP$$

$$\quad \quad \quad \uparrow = \downarrow \quad \quad \uparrow \text{ OBJ} = \downarrow$$

(24) 中的上箭號 "↑" 代表母節，下箭號 ↓ 代表子節。故 ↑ = ↓ 即表示母節與子節的語法訊息聯併；而 ↑ OBJ = ↓ 則表示子節的語法訊息和母節的一個功能 (LFG 把所有語法功能視為屬性，而其值則是另一個功能結構) 聯併。換句話說，(24) 標示了 V 所承載的訊息是骨幹句 (*matrix*)，即整個大句子的訊息；而 NP 的訊息則屬於大句子所支配的賓語 (OBJ)。↑ = ↓ 即表示子節和母節共有所有特徵；符合上節所討論中心語特徵通律的要求。故我們說 ↑ = ↓ 相當於 GPSG 中的中心語特徵通律。

LFG 的另一特色是語句是否合法完全看功能結是否合法。這也反映了聯併為本語法理論只規定什麼樣的訊息組合才合法這個特色。LFG 規定合法的功能結構必須符合三個條件：完備性 (*completeness*) 要求謂語 (*predicate*) 所規定的必要語法功能均必須在功能結構中出現；而一致性 (*coherence*) 則規定每個在功能結構中出現的語法功能均需有謂語管轄；而雙向唯一性 (*Bi-uniqueness*) 則規定每個屬性均只能有一個特徵值，反之亦然。

綜上所述，LFG 以明確的規律表現了聯併為本理論的特色。其以語法功能為基本觀念之選擇已有心理語言學上的論證支持；而成份與功能結構的區分雖在觀念上相當明確；其在實際剖析上之效用均待更深入研究來證明。

3. 以中心語驅動的詞組結構語法 (*Head-driven Phrase Structure Grammar, HPSG*)，功能聯併語法 (*Functional Unification Grammar, FUG*) 及其他

以中心語驅動的詞組結構語法 (以下簡稱 HPSG) 脫胎於 GPSG。主要著眼點在於保存 GPSG 免用語境語法的優點，並同時擴展其能力以解決一些已知 GPSG

遭遇的困難，如超免用語境的交互序列對應現象。HPSG改進的第一步在廢除律上律(Pollard 1985)，而改用詞彙律。規範語言學家證明律上律的施用若不加限制，其威力相當於杜林機器(Turing Machine)，即是一個毫無限制的改寫系統，可以生成任何語言。在辨識剖析上卻毫無效率可言。故 GPSG 把律上律限制到只能施用於以詞彙為中心的直接支配律。伯樂(Pollard 1985)則指出這樣的限制等於承認這些關係是受詞項控制的，倒不如乾脆把這些規律放入詞彙中。HPSG另一個重要的特色是把次類劃分(subcategorization)及各必要成份之間結合的順序都納入 SUBCAT 這個屬性的值中。這個分析法近於範疇語法(categorical grammar)中以論元結合順序定其語法角色的作法。而與 GPSG 中以特徵值規定選用的直接支配律以決定其次類劃分以及 LFG 中詞彙決定的謂語—論元結構(Predicate-argument Structure)而無關於語法兩種分析迥異。

HPSG另一個重要的特色源自功能聯併語法(以下節稱 FUG, Kay 1985)以及 FUG的前身—功能語法(Functional Grammar, Kay 1979)。這個特色便是把所有特徵通律限制等均以特徵結構表示，而不用文字敘述。也就是說，明確的把聯併定為語法中唯一的運作。

(25) 中心語特徵原則(Head Feature Principle, Sag & Pollard 1987.17)

$$\vdash \text{SYNTAX} \mid \text{LOCAL} \mid \text{LEX} - \neg \Rightarrow$$

$$\vdash \text{DTRS} \mid \text{HEAD-DTR} \quad [] \vdash$$

$$\vdash \text{STNTAX} \mid \text{LOCAL} \mid \text{HEAD} \quad []$$

$$\vdash \text{DTRS} \mid \text{HEAD-DTR} \mid \text{SYNTAX} \mid \text{LOCAL} \mid \text{HEAD} \quad [] \vdash$$

(25)在 HPSG 中指定任何詞組的中心特徵必和其中心語子女(Head-daughter, 規律中為 HEAD-DTR)的中心語特徵相等。這是以特徵結構明確表示出了中心特徵傳遞分佈上的特性。這種表達方式使得詞項和規律的表達方式相同，都是聯併運作下的訊息結構。

FUG 嚴格說來不是語法理論，而是剖析系統中應用的一種語法表達模式。他的特色之一是以成份集合(CSET)與分佈類型(Pattern)兩個屬性來決定實際語中各成份出現的位置。另一個特色，上面已提過的，便是把整個語法，包括詞彙、規律與限制，視為是由一群特徵結構以 $\wedge$ (and) 與 $\vee$ (or) 連接成的一個大訊息體。

其他與聯併為本的語法理論有關的有 PATR-II, DPATR 等在 CSLI 發展為執行聯併為本語法而設計的程式(見 Shieber 1986 及 Karttunen 1986)。這些系統各有其建構上的特色，不在此細述。

IV、結語：聯併及相關語法理論文獻導讀

本文中討論到了聯併爲本的語法理論用在剖析上的許多好處，以及這些語法理論的一些特色。語言學上的文獻提供了以這些理論成功分析自然語言的實例，如 Gunji (1987) 分析日文，Gazdar et al. (1985) 分析英文，Huang (1987) 分析中文等。由實際有效地在電腦上處理中文以及建立一套可驗證的語法理論用以分析中文這兩個觀點考慮，聯併爲本的語法理論該是將來研究的主要方向。更重要的，藉著聯併的媒介，我們可以在剖析自然語言時直接引用語法理論，以剖析驗證語法理論，更根據剖析的結果修改語法理論。換句話說，聯併可成爲理論語言學與自然語言處理工程的橋樑。我們寄望這個橋樑加速計算語言學的發展。若對聯併與相關理論想有進一步了解，請參閱以下簡介的幾篇文章。

施別 (Shieber 1986) 是介紹聯併的經典之作，謝爾師 (Sells 1985) 則是對三種重要語法理論 (GPSG, LFG; 以及通常被認爲非以聯併爲本的 GB) 最簡明切要的入門介紹。以自然語言爲承載訊息的系統，可看索緒爾的經典之作「普通語言學教程」(*A Course in General Linguistics*)。此書有多種英譯本，出版年份也各異。至於規範語言的觀念，杭士基的兩部經典 (Chomsky 1957, 1965) 均未深入。倒是蓋志達 (Gazdar 1987) 深入淺出爲衍生語法理論下了明確的定義。華爾 (Wall 1972) 是給語言學家最通俗的簡介，稍失於簡略，也不夠嚴謹。霍普克勞復與烏爾曼 (Hopcroft & Ullman 1979) 最完整。中文簡介則有黃 (待刊 1989) 一文，同時也淺論了一些漢語的規範性質。電腦和語言研究之間的關係，在王 (Wang 1985) 一書中收集了許多可讀性很高並能探及問題核心的文章。剖析理論與語法理論在剖析中的運用可參見溫若貴 (Winograd 1983)。語法理論方面，中文有黃 (1987) 的簡介，深度卻不夠。各理論的重要文獻，概化詞組結構語法 (GPSG) 是 Gazdar et al. (1985)，詞彙功能語法 (LFG) 是 Bresnan (1982)，中心語驅動的詞組結構語法 (HPSG) 則是 Pollard and Sag (1987)，另有功能聯併語法 (FUG) 見於 Kay (1985)。有一大部份與聯併有關的重要論文 (包括理論與剖析運用) 收集在施別等人所編的論文集 (Shieber et al. 1986) 中。史丹福大學資訊與語言研究中心 (Center for the Study of Language and Information, CSLI) 所出版的報告和書 (Report & Lecture Notes) 是相關資料的主要來源。另外在最近四、五年內 ACL 和 COLING 論文集集中有許多聯併爲本，用到 GPSG、LFG、HPSG 等的剖析論文，在此不一一列舉。蓋志達 (Gazdar 1985) 主編的 *Linguistics* 23 卷 2 期收了許多剖析 GPSG 的重要文章。墨力殊 (Mellish 1988) 則是聯併爲本剖析的最新文獻之一，可在其參考文獻中找到相關資料。本論文集王士元先生的文章爲計算機與語言學合作的歷史與展望作了深入淺出的介紹。伯樂 (Pollard) 爲計算語言學下了新界說。蘇克毅與陳克健分別介紹了剖析策略與這些策略時在中文裡的運用。Gunji 則以日文爲我們舉了一個利用聯併爲本的理論分析一個特定自然語言的最佳實例。

※致謝

本文之研究得國科會專題研究計劃 (NSC-77-0408-E001-01) 及中央研究院與工業技術研究院電子工業研究所合作之「中文詞知識庫第三期及中文語句剖析系統第一期合作研究與開發計劃」之部份補助，特在此申謝。本文承中央研究院計算中心協助以雅墨 (Inkart) 排版系統列印，生色不少。初稿經詞庫計劃同仁特別是陳克健教授提供了許多寶貴意見，改進了許多疏失。文中若仍有遺誤，筆者自負思慮不週之責。

中文參考資料：

1. 黃居仁 1989 (待刊) 試論漢語的數學規範性質 歷史語言研究所集刊第六十本。
2. 黃居仁 1987 現代語法理論與華文語法教學 華文世界第46期 (一九八七年十一月) 1-6 頁。

西文參考資料：

- Bresnan, Joan. 1982. The Mental Representation of Grammatical Relations. Cambridge: MIT Press.
- _____, Ronald Kaplan, Stanley Peters, and Annie Zaenen. 1982. Cross-serial Dependencies in Dutch. Linguistic Inquiry. 13.4.613-23.
- Chao, Yuen-Ren. 1968. Language and Symbolic Systems. Berkeley: University of California Press.
- Chomsky, Noam. 1955. The Logical Structure of Linguistic Theory. New York: Plenum (1975). Cambridge: Harvard University Press (Paper Back, 1985).
- _____. 1957. Syntactic Structures. The Hague: Mouton.
- _____. 1965. Aspects of the Theory of Syntax. Cambridge: MIT Press.
- Gazdar, Gerald. 1981. Unbounded Dependencies and Coordinate Structure. Linguistic Inquiry. 12.2.155-184.
- _____. 1985. (Ed.) Linguistics. vol. 23. no. 2.
- _____. 1987. Generative Grammar. John Lyons, Richard Coates, Margaret Deuchar and Gerald Gazdar. 1987. Eds. New Horizons in Linguistics II. Harmondsworth: Penguin.

- _____, E. Klein, G. Pullum and I. Sag. 1985. Generalized Phrase Structure Grammar. Cambridge: Harvard University Press. And Oxford: Blackwell.
- _____, Geoffrey K. Pullum, Robert Carpenter, Ewan Klein, Thomas Hukari and Robert Levine. 1988. Category Structures. Computational Linguistics. 14.1.1-19.
- Gunji, Takao. 1987. Japanese Phrase Structure Grammar. Dordrecht: Reidel.
- Hopcroft, E. John and Jeffrey D. Ullman. 1979. Introduction to Automata Theory, Languages and Computation. Reading: Addison-Wesley.
- Huang, Chu-Ren. 1987. Mandarin NP de--A Comparative Study of Current Grammatical Theories. Unpublished Cornell University Ph.d. Dissertation.
- _____. 1988. A Unification-based LFG Analysis of Lexical Discontinuity. Presented at the Fourth International Workshop on Asian and Oriental Linguistics. June 15-16. Paris.
- Karttunen, Lauri. 1986. D-PATR: A Development Environment for Unification-based Grammars. CSLI Report no. CSLI-86-68. Stanford: Center for the Study of Language and Information.
- Kay, Martin. 1979. Functional Grammar. Proceedings of the Fifth Annual Meeting of the Berkeley Linguistic Society. 142-158. Berkely: Berkeley Linguistic Society.
- _____. 1985. Parsing Functional Unification Grammar. David Dowty, Lauri Karttunen and Arnold M. Zwicky. Eds.. Natural Language Parsing: Psychological, Computational and Theoretical Perspectives. 251-278. Cambridge: Cambridge University Press.
- Mellish, Christopher S. 1988. Implementing Systemic Classification by Unification. Computational Linguistics. 14.1.40-51
- Perrault, C. Raymond. 1985. On the Mathematical Properties of Linguistic Theories. CSLI Report No. CSLI-85-18. Stanford: Center for the Study of Language and Information.
- Peters, P. Stanley Jr. and R. W. Ritchie. 1973. On the Generative Power of Transformational Grammars. Information and Control. vol.18. 483-501.
- Pierce, John R. 1980. An Introduction to Information Theory. (Revised version). New York: Dover.

- Pollard, Carl.** Phrase Structure Grammar without Metarules. Proceedings of the Fourth West Coast Conference on Formal Linguistics. 246-261. Stanford: Stanford Linguistics Association.
- _____ and Ivan Sag. 1987. Information-based Syntax and Semantics: Volume I. Fundamentals. CSLI Lecture Notes no.13. Stanford: Center for the Study of Language and Information.
- Pullum, Geoffrey K. and Gerald Gazdar.** 1982. Natural Languages and Context-free Languages. Linguistics and Philosophy. 4. 471-504.
- Sag, Ivan, Ronald Kaplan, Lauri Karttunen, Martin Kay, Carl Pollard, Stuart M. Shieber, and Annie Zaenen.** 1986. Unification and Grammatical Theory. Proceedings of the Fifth West Conference on Formal Linguistics. Stanford: Stanford Linguistics Association.
- _____ and Carl Pollard. 1987. Head-Driven Phrase Structure Grammar: An Informal Synopsis. CSLI Report No. CSLI-87-79. Stanford: Center for the Study of Language and Information.
- Saussure, Ferdinand d.** 1915. Course in General Linguistics. New York: McGraw Hill (English Translation, 1966).
- Sells, Peter.** 1985. Lectures on Contemporary Syntactic Theories. CSLI Lecture Notes no. 3. Stanford: Center for the Study of Language and Information.
- Shieber, Stuart M.** 1985. Evidence against the Context-freeness of Natural Language. Linguistics and Philosophy. 8.333-343.
- _____ 1986. Introduction to Unification-based Approaches to Grammar. Stanford: Center for the Study of Language and Information.
- _____, Fernando C.N. Pereira, Lauri Karttunen, and Martin Kay. 1986. A Compilation of Papers on Unification-based Grammar Formalisms. Parts I and II. CSLI Report No. CSLI-86-48. Stanford: Center for the Study of Language and Information.
- Wall, Robert.** 1972. Introduction to Mathematical Linguistics. Englewood Cliffs: Prentice-Hall.
- Wang, William S-Y.** 1985. (Intro.) Language, Writing, and the Computer. Readings from Scientific American. New York: W. H. Freeman.
- Winograd, Terry.** 1983. Language as A Cognitive Process: Vol. I Syntax. Reading: Addison-Wesley.