

中文裡「動-名」述賓結構與「動-名」偏正結構的分析

陳克健 洪偉美

中央研究院資訊科學研究所

摘要

在電腦剖析中文的過程中，V+N的結構究竟是偏正關係(VN)，還是述賓關係(VO)，是一個常遇到的歧義現象。本文藉由電腦的輔助，蒐集及觀察大量的語料，找出兩種結構間的一些語法及語意的差異，並提出一個區別兩者的方法，希望有助中文歧義結構的分析。

首先，利用自動分詞系統及詞類訊息，從二千萬字的語料庫中找出56936筆不同的動名配對。接著，利用修飾語動詞的名物化特性及名詞個體與非個體的區別，排列出動詞及名詞的VO及VN傾向。初步判斷結構的方法，即根據此二種傾向的組合來決定是VO還是VN關係，而其中不易判斷的部份者則由人為決定。

有了初步判斷的資料，接著，我們加以進一步的分析。我們發現除了動詞的名物化特性與名詞的個體與非個體的區別外、動詞對名賓語的選擇限制及其持續性都影響了動詞直接修飾名詞的可能性。以上這些觀察將有助於電腦對V+N結構的剖析。

一、前言

在中文裡[動詞+名詞]的關係基本上有兩種可能：(1)述賓關係[_{VP} V NP] (VO)、(2)偏正關係[_{NP} V N] (VN)[註1]，單音節動詞與其後的名詞，若要表達修飾關係，則必須中插助詞"的"，例如：‘吃的東西、看的書、買的禮物’。如果緊接名詞，則幾乎都是述賓關係[註2]，例如：‘吃東西、看書、買禮物’。而雙音節的動詞則可直接修飾名詞，不需助詞"的"，例如：‘學習方法、使用情形、參觀地點’，而這種結構與一般的述賓結構一

樣，例如：‘學習中文、使用設備、參觀博物館’。因此造成電腦剖析句子時的困擾。我們研究的重點在如何區別[及物雙音節動詞+雙音節名詞]的不同句法功能。例如：

勒索金錢(VO)

控制人數(VO)

勒索金額(VN)

錄取人數(VN)

我們想知道為什麼會有這樣的差別。什麼樣的雙音節動詞可以修飾名詞，什麼樣的動詞不行。什麼樣的名詞容易被動詞所修飾，什麼樣的名詞不容易。是什麼特性造成動詞修飾性的差異，而又是什麼特性造成名詞的差異。我們相信一定有某些語意特性造成這些差異，而且這些差異不是由上下文(discourse)決定的，因為絕大部份的「動詞+名詞」結構，不必由上下文就可判斷其為述賓還是偏正結構。因此，如果我們能把其間的關係清楚地找出來，就可以使電腦的分析系統也可以自動分辨兩者的差別，如此將有助中文歧義結構的分析。

二、研究方法

觀察語料是研究述賓(VO)關係、偏正(VN)關係差異的最好方法，但由於欠缺結構標記的語料庫，因此無法很容易的得到大量「動詞+名詞」(V+N)並標有VO關係或VN關係的配對，所以研究的第一步工作便是如何得到V+N的配對資料。

有了大量V+N的配對資料，我們希望從一些相關句型結構的統計分佈上，反映出VO及VN的不同，或者從一些已知的動詞或名詞的性質協助分辨VO及VN的關係，例如：我們猜測‘V+的+N’的結構如果在語料中出現，則此一V+N的結構很可能是VN關係，而實際上是否如此呢？反過來，‘V+的+N’不出現，是否意味著該‘V+N’為VO結構呢？這些都要進一步觀察研究的。

一旦我們利用一些方法分辨並標記大量‘V+N’配對的VO或VN關係，我們可以進一步分析是什麼特性使得V+N有不同的關係產生，進一步可以導出自動分辨‘V+N’結構的方法。

2.1 V+N配對的蒐集

從詞庫的詞類標記語料庫中，我們發現「單音節及物動詞+名詞」幾乎都是VO關係，而「不及物動詞+名詞」必然是VN關係。因此，我們只在語料中找出所有「雙音節及物動詞+名詞」的配對來研究，從約二千萬字的詞庫語料庫中，利用自動斷詞程式，找出56936個不同的V+N配對，有些配對在語料庫中只出現一次，有些配對則出現多次，其分佈如下表：

出現次數	1	2	3	4	5	6	7	8	9
VN配對個數	33549	9718	4135	2442	1493	1008	745	548	410

10	11	12	13	14	15	16	17	18	19	20
322	252	212	209	169	146	124	128	91	86	85

256	263	282	284	290	300	318	342	390	454	607
1	1	1	1	1	1	1	1	1	1	1

出現10次以上的只有2888對，只佔百分之五。

2.2 句型的統計分佈與結構的關係

我們猜測如果一個'V+N'的配對，在語料庫也同時出現'V+的+N'的結構，是否這個'V+N'為VN結構。事實不然，從實際語料觀察，我們發現'V+N'為VO或VN結構和是否存在'V+的+N'的結構，並無直接關聯。

另一個猜測是，如果一個'V+N'的配對，在語料中也存在著大量'V+...+N'的結構，則這個'V+N'很可能是述賓關係。所謂'V+...+N'的結構，表示在語料中，V和N中間有其他成份出現，而這些成份很可能是N的

修飾語，因此，這個 'V+N' 很可能是述賓關係。我們利用全文檢索的技術，將語料建立字的翻轉索引檔(invert file)，查詢所有 'V+N' 配對、'V+的+N' 及 'V+...+N' 在語料中出現的次數。我們設定 'V+...+N' 句型中，V和N的距離小於等於6。下圖為產生的一些例子：

		'V+N'	'V+的+N'	'V+...+N'	結構
獎勵	金額	7	1	0	VN
輸入	金額	1	1	0	VN
勒贖	金額	1	1	0	VN
損失	金額	12	1	3	VN
投資	金額	212	6	39	VN
回饋	金額	13	2	2	VN
出口	金額	32	0	23	VN
外銷	金額	6	0	11	VN
銷售	金額	32	0	3	VN

根據我們的猜測，以上統計數字中，那些出現 'V+的+N'，而沒出現 'V+...+N' 的 'VN' 配對，應該是VN結構，例如：'獎勵金額、輸入金額、勒贖金額'等，事實上也是如此。但是，反過來地，那些不出現 'V+的+N'，而出現 'V+...+N' 的 'VN' 配對，應該是VO結構，例如：'出口金額、外銷金額、銷售金額'等。結果並不然，這些都還是VN結構。而那些既有 'V+的+N' 及 'V+...+N' 的 'VN' 配對，則可能是VN結構或VO結構，例如：'損失金額、投資金額、回饋金額'等。但卻全都是VN結構，類似這樣的例子很多。因此，我們歸納出，句型的統計分佈與結構並沒有一定的關係，它們結構可能是由動詞或名詞的某些特性來決定的，例如：這些 'V+金額' 的VN配對，全都是VN結構，可能是因為 '金額' 這個名詞的關係。

2.3 從一些既有的詞類性質探討VO、VN關係

首先，我們從一些既有的詞類特性，來探討VO, VN的關係，前面我們已經提到「不及物動詞+名詞」一定是偏正關係。這裡我們再仔細地觀察詞庫的各類動詞後接名詞的情形。

		說明
VA	動作不及物	後接名詞時，必為VN關係，例如：‘謀生能力、違規行為’。
VB	動作類單賓	賓語不直接出現在動詞後，而以介詞引介或賓語前提的方式出現。因此，此類動詞後接名詞，必為VN關係，例如：‘轉手方式’。
VC	動作單賓	後接一名詞組當賓語，與VN關係的‘V+N’在結構上一樣。因此，此類動詞後接名詞時，可能是VO關係或VN關係，例如：‘管理員工(VO)、管理人才(VN)’。
VD	雙賓	後接兩個名詞組當賓語，例如：送他(DO)一本書(IO)，這一類動詞有極少數會形成VN結構，例如：‘餽贈方式’，此時需靠動詞與賓語的選擇限制來判斷。
VE	動作句賓	後接一個句子當論元，例如：‘主席宣布[今天會議到此結束]’，這類動詞後接名詞時，有些會形成VN結構，例如：‘宣布方式’。

VF	動作謂賓	後接動詞組當賓語，或是帶一個屬人名詞組當終點及一動詞組當述語，例如：‘業者強迫員工加班’。這類動詞後接名詞時，有少數會形成VN結構，例如：‘強迫行爲’。可利用賓語的選擇限制來區別。
VG	分類	後接一個名詞當補語，例如：‘姓、名叫...’等。此類動詞數量很少，也很少當修飾語用。
VH	狀態不及物	此類動詞用來描述某種狀態，後面不帶賓語。因此，若後接名詞，必為VN關係，例如：‘動聽歌曲、特別日子’。
VI	狀態類單賓	賓語不直接出現在動詞後，而以介詞引介或賓語前提的方式出現。因此，此類動詞後接名詞，必為VN關係，例如：‘相干人士’。
VJ	狀態單賓	後接一名詞組當賓語，與VN關係的‘V+N’在結構上一樣。因此，此類動詞後接名詞時，可能是VO關係或VN關係，例如：‘適應環境(VO)、適應能力(VN)’。
VK	狀態句賓	後接一個句子當論元，例如：‘我明白[他昨天為何不來]’，這類動詞幾乎不形成VN結構。

VL	狀態謂賓	後接動詞組當賓語。此類動詞不當修飾語用，例如：‘故意、致力、甘願、不禁、足以...’等。
----	------	--

從以上說明可歸納出，不及物動詞及類單賓動詞後接名詞時，一定會形成VN關係。而及物動詞接名詞時，則可能分析為VN關係或VO關係。其中單賓動詞接賓語時，和VN關係的結構完全一樣，造成電腦處理上很大的困擾。因此，我們從區別「單賓動詞+名詞」的不同結構著手。

2.3.1 偏正關係和名物化動詞的關係

由語料觀察得知，動詞是否能直接修飾名詞，與它們能被名物化與否有密切的關係，如下：

[名物化][當修飾語]

投資	+	+	做有效的投資(名物化)，投資途徑(偏正)
處理	+	+	公務的處理(名物化)，處理原則(偏正)
支持	+	+	精神上的支持(名物化)，支持程度(偏正)
零售	+	+	煙酒的零售(名物化)，零售價格(偏正)
找到	-	-	*途徑的找到(名物化)，找到途徑(述賓)
沒有	-	-	*糧食的沒有(名物化)，沒有糧食(述賓)
造福	-	-	*縣民的造福(名物化)，造福縣民(述賓)
加強	+	-	治安的加強(名物化)，加強治安(述賓)
遠離	+	-	颱風的遠離(名物化)，遠離毒品(述賓)
釀成	+	-	災害的釀成(名物化)，釀成災害(述賓)

朱德熙(1990)指出“這類格式中的動詞(如:建築材料、研究方法)已經名物化了，它是以名詞的資格充任定語的，因此凡是不大能被名物化的雙音節動詞也就不能直接作定語”。語料也顯示能直接當修飾語的動詞，一般都

能被名物化。而能被名物化的動詞，則不一定能當修飾語，例如：‘加強、遠離、釀成...等’，在語料庫中並未發現當修飾語的例子。

2.3.2 名詞的語意特性

名詞的語意特性與它的語法功能有密切關係。Hopper & Thompson(1980)指出一個動詞與其後的名詞能否形成述賓關係和該名詞的「個體」(individuated)或「非個體」(non-individuated)有密切關係，個體名詞比非個體名詞容易充當賓語[註3]。專有名詞、具體名詞、指涉性名詞、稱代詞是個體名詞；而普通名詞、抽象名詞、非指涉性名詞是非個體名詞。非個體名詞當賓語時，一般會前接數量詞、代詞、名詞等定語，所以，若緊接在動詞之後時，較容易充當被修飾語，如下：

	述賓關係	偏正關係
專有名詞vs.普通名詞	登陸<科威特>	登陸<地點>
具體名詞vs.抽象名詞	投資<黃金>	投資<項目>
指涉性名詞vs.非指涉性名詞	抨擊<市長>	抨擊<對象>

2.4 初步判別VO，VN結構的方法

首先從兩千萬字的語料庫中找出所有[雙音節單賓動詞+雙音節名詞]的配對，共有56936個不同配對，而初步判斷VO, VN結構的方法如下：(1)由於我們發現修飾名詞的動詞通常為名動詞，所以我們利用名物化句型出現的頻率排列出名動傾向的動詞及非名動傾向的動詞(即無名物化句型出現的動詞)[陳1994]。(2)同樣的，名詞也有V-O傾向的差異，個體名詞多傾向於V-O結構，非個體名詞較可能形成V-N結構。而V-N程度測量的方法是，看每一特定N前面出現的動詞中名動詞所佔的比例，比例愈高，愈為V-N傾向。(3)此兩個動詞與名詞傾向交互組合出其可能的結構，若動詞與名詞的V-O傾向皆強，則此‘V+N’為VO結構；反之，若動詞與名詞的V-N傾向皆強，則此‘V+N’為VN結構。但如果兩個傾向有衝突或二者均無明顯傾

向時，則必須由人爲決定。值得一提的是，動詞的V-O傾向強過名詞的V-N傾向。因此，帶V-O傾向的動詞，即使後接V-N傾向的名詞，仍會組合出V-O結構，例如：‘沒有措施/方法/狀況、面對缺失/問題/紛爭...’等。

2.4.1 名詞的被修飾傾向

如前述，修飾名詞的動詞通常爲名動詞，而辨別名動詞的方法是借重語料庫，利用一些已知的語法現象協助判別[陳 1994]。名詞V-N傾向的測量方法是，看每一特定N前面出現的動詞中名動詞所佔的比例，比例愈高，愈爲V-N傾向；比例愈低者，愈爲V-O傾向。僅列出VN傾向最強的前10個名詞如下：

焦點	題材	途徑	講座	流程	行列	方面	層面	事宜	要點
25/26	19/20	25/27	19/21	27/30	52/58	195/217	24/27	137/154	32/36
96%	95%	93%	90%	90%	90%	90%	89%	89%	89%

而VO傾向最強的前10個名詞如下：

經濟	民意	環保	政治	電腦	人民	傳統	國人	黨員	攤販
0/140	0/98	0/96	0/95	0/93	0/88	0/80	0/70	0/74	0/71
0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

2.4.2 動詞修飾名詞的傾向分佈

我們已經知道不能被名物化的動詞是不會直接修飾名詞的，所以這類動詞的VO傾向很強，例如：‘沒有、引起、從事、具有、面對、引發、達到...’等。而名動詞VN傾向的測量方法是，利用VN傾向強的名詞排列出前面的動詞，出現次數愈多者，VN傾向愈強；出現次數愈少者，VN傾向愈弱，即愈傾向中性(N)。VN傾向最強的前10個動詞如下：

投資	管理	服務	銷售	選舉	改革	出口	運輸	外銷	宣傳
167	102	98	92	73	70	67	52	51	47

(表中數據表示出現的次數)

VN傾向最弱的10個動詞如下：

加強	享受	注重	增進	實現	縮短	凸顯	釀成	偏重	遠離
----	----	----	----	----	----	----	----	----	----

2.4.3 標記V+N配對的VO，VN關係

利用前述動詞及名詞的特性，去判斷一個動名結構為述賓關係(VO)或偏正關係(VN)。表一中V-O代表VO傾向，V-N代表VN傾向，N代表中性。所謂中性的意思是利用前述統計方法無法判斷為VN或VO傾向。

V	N	V+N	例子
V-O	V-O	述賓	掌握水源
V-O	V-N	述賓	掌握狀況
V-O	N	述賓	掌握時效
N	V-O	述賓	支持國建
N	V-N	偏正	支持聲浪
N	N	?	支持政策(VO) 支持決心(VN)
V-N	V-O	?	宣傳佛法(VO) 宣傳布條(VN)
V-N	V-N	偏正	宣傳手法
V-N	N	偏正	宣傳資料

表中的動名關係大部份可以確立，少數帶'?'部份，必須由人為觀察，以判別其為述賓關係或偏正關係。

三、資料的分析結果

由以上動詞與名詞傾向所判斷出的VO與VN結果，我們再仔細研究，發現動詞的一些語意特性：

(1)對其後賓語幾乎沒什麼選擇限制的動詞，大多形成述賓結構。

一個及物動詞對其後的賓語通常有語意上的選擇限制，但有些動詞卻幾乎不限制它賓語的種類，例如：‘沒有、面對、失去、透過、具有...’等。這些動詞即使其後出現的名詞是「非個體名詞」，仍然會形成述賓結構，例如：

沒有策略	面對差異	失去原則
沒有方法	面對缺失	失去作用
沒有狀況	面對問題	失去機會
沒有題材	面對紛爭	失去據點

(2)持續性動詞比非持續性動詞易形成偏正結構。

持續動詞是指一個動詞所指涉的動作沒有明顯的起點、過程和終點，這一類動詞通常會後接"著"。持續動詞比非持續動詞較易與後面名詞形成偏正結構，例如：‘居住(地點/環境)、研究(精神/主題)、批評(焦點/聲浪)...’等。而非持續動詞則較易形成述賓結構，例如：‘接獲(消息/訊號)、放棄(立場/機會)、擺脫(官司/媒體)、變成(事實/習慣)、辦妥(手續/保險)、拿出(名單/成果)、打破(慣例/漏習)...’等。

另外，根據統計結果，以下是常當被修飾語的名詞，例如：

	出現總數	當被修飾語比例	偏正結構例子
項目	91	100%	使用/投資/回收/檢查/展示項目
部門	42	100%	代理/仲介/執行/管理/開發部門
體系	35	100%	教育/加盟/服務/領導/生產體系
手段	31	100%	防治/取締/杯葛/威脅/拐騙手段
因素	28	100%	影響/緊縮/損壞/炒作/干擾因素
狀態	25	100%	使用/準備/經營/虧損/緊縮狀態

從語料庫中也發現，其他名詞如：次數、餘地、機率、架構、容量、頻率、盛況、總值等，前接動詞時，也都當被修飾語用。而這些名詞大多數是抽象的非指涉性名詞，的確符合非個體性名詞的特性。

四、自動分辨的方法

歸納以上的討論，我們在利用電腦剖析V+N結構時，首先會考慮到動詞的詞類。不及物動詞或類單賓動詞後接名詞時，必為VN關係，例如：‘行車路線、分類標準’。而及物動詞後接名詞時則可根據以下的判斷方式來處理。

由資料分析結果發現，動詞的能否名物化(即動詞的VO傾向)、動詞的VN傾向及名詞個體與非個體的區別(即VN,VO傾向)是決定V+N結構的重要因素。因此，電腦在V+N的剖析上，將考慮這些因素：

(1)不能被名物化的動詞幾乎不當修飾語，所以這些動詞後接名詞時，皆形成述賓結構，例如：‘造福縣民、辦妥手續’。

(2)而能被名物化的動詞則有兩種可能：

(A)若動詞的VN傾向強，且後接非個體名詞則形成偏正結構，例如：‘宣傳管道、發生機率’。

(B)若動詞的VN傾向弱或後接個體名詞，則形成述賓結構，例如：誤觸法網、登陸台灣。

五、結論

語料提供語言研究最重要的素材。許多語言特性反應在語料裡，以不同的統計分佈呈現。而這些統計分佈常常需要利用一些已知的語言知識，才能真正的反應出來。在本論文裡，我們利用已知的語言特性，以名物化動詞片語句型的統計分佈，找出名物化動詞。接著利用VN配對中V的名物化動詞比例分佈，判別N的VO或VN傾向。這些過程同時利用了已知的語言知識及統計分佈，為語料庫語言學的典型研究方法。

本論文對動名結構有進一步的了解。動詞能否被名物化、動詞對名賓語的選擇限制及動詞的持續性，影響了動詞直接修飾名詞的可能性。而影響名詞的性質主要為個體和非個體的差異。未來我們希望應用這些特性，協助解決電腦剖析中文時V+N結構造成的歧義分析。

附註

- [1]在語料中出現V+N的結構時，不一定是VO或VN關係，也有可能N為後面中心語名詞的修飾語，或是修飾語中的一部份，例如：提供 道路 用地、取消 土地 所有權 移轉。
- [2]張國憲(1989)指出“有兩類單音節動詞可以直接充當定語，組成偏正短語。第一類，”飛、死、來”等少數單音節動詞，第二類，“烤、炒、燒、煎、燴、炖”等表烹飪的單音節動詞。由於能夠直接充當定語的單音節動詞極少，所以不影響我們的立論。”
- [3]這裡所說的個體名詞與非個體名詞包括名詞及名詞組。

參考書目

- [1]張國憲，“動+名”結構中單雙音節動作動詞功能差異初探，中國語文 1989年第3期
- [2]中文詞庫小組，1993，國語的詞類分析(修訂版)，中央研究院
- [3]朱德熙，1987，“定語和狀語”，漢語知識講話。上海教育出版社
- [4]Hopper, P.J., and Thompson, S.A. 1980 Transitivity in Grammar and Discourse, *Language* 56, 25-299
- [5]陳克健,1994,素材語言學與文本處理,ICCL-3

