

## 「V 到」結構的合分詞及語意區分

### Word segmentation and sense representation for V-*dao* structure in Chinese

黃淑齡、李詩敏、白明弘、吳鑑城、簡盈妮、林慶隆  
國家教育研究院編譯發展中心

Development Center for Compilation and Translation  
National Academy for Educational Research

{slhuang, smli, mhbai, wujc, jyingni, cllin}@mail.naer.edu.tw

#### 摘要

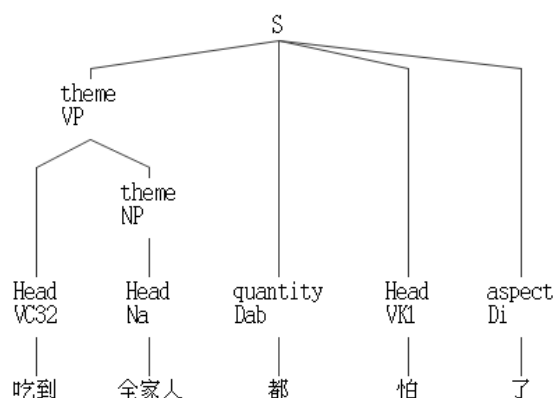
中文複合詞中有很大部份是由動補結構所組合產生，它們在語料中常呈現合分詞不一致或錯誤的情況，本文以動補結構中合分詞類型最複雜的結構「V 到」為例，探討其合分詞及語意區分問題。我們根據「到」是否有「到達」的語意，利用七條原則及簡易判準——即以賓語類型{地點 時點 狀態}為標準——採人工的方式進行標記，並評估合分詞正確率，結果證明「吃到」可以由目前的 70.6% 正確率提升到 94.5%。以「V 到」結構整體來看，在 9 個例詞 500 條隨機選取的例句中，合分詞正確率可達到 93.4%，基於此合分詞結果的語意合成正確率也達到 86%。顯示複雜的動補結構合分詞問題可經由簡易人工規則來改善。未來我們計畫將人工規則轉為自動化處理程式並檢驗其正確性。

關鍵詞：V 到，分詞，中文動補結構，語意表達

#### 一、緒論

中文複合詞中有很大部份是由動補 (Verb-complement) 結構所組合產生，例如：看到、想來等等。為顧及分詞效率，一般系統會將這些高頻的動補結構視為複合詞收詞。然而，在不同的語境下，有時它們需分詞處理，例如：看到 傻了、想來沾點邊等等，也就是說，以收詞的方式來處理動補式複合詞，將難以避免分詞錯誤，事實上，語料庫中動補結構分詞錯誤或不一致的例子頗為常見，如(1)

(1) 吃到全家人都怕了



上例中「吃到」合詞，「全家人」則誤為賓語，正確的斷詞為 吃 到 全家人都怕了，其意義應解釋為全家人 吃 到 都怕了，「到」為結構助詞，作用如「得」字。以「吃到」為例，檢視國家教育研究院建置的華語文語料庫（Corpus of Contemporary Taiwanese Mandarin，簡稱 COCT<sup>1</sup>）1 億 1,220 萬字書面語語料[1]，在 711 個例句中，有 209 個分詞錯誤，<sup>2</sup>分詞正確僅 70.6%，與平均分詞正確率 94%相去甚遠。<sup>3</sup>本文嘗試修正這些分詞錯誤，我們以「V 到」結構為例，因為「到」這個虛化動詞，在語言分析時有時被視為介詞，引介時間和地方，例如：走到巷口、等到明天；有時被視為結構助詞，連接狀態程度，例如：冷到不行、佩服到五體投地；有時被視為動相補語（phase marker），表示動作瞬間完成、實現，例如：碰到怪事、找到鑰匙等等；顯示「V 到」結構包含複雜的合分詞情況，很適合作為範例，來說明應如何處理動補結構的語意區分問題、訂定分詞原則，並修改語料中的分詞錯誤或不一致情況。本文第二節將就「V 到」結構的語意分析作文獻探討，並根據語境，提出合理的分詞原則及語意預測。第三節則探討這些原則的適用性、如何依據這些原則做合分詞判斷，及合分詞後句子語意表達的正確性。第四節討論未來自動合分詞的方法及可能遭遇的困難。最後則提出結論。

<sup>1</sup> 國教院語料索引典系統網址為：<http://coct.naer.edu.tw/cqpweb/>。

<sup>2</sup> 根據語境應分詞，目前的分詞系統卻合詞的例子為：吃到第二天、吃到幾乎反胃、吃到一半、從小吃到大、吃到這種地步、吃到嘴裡、為什麼要吃到一萬多元、吃到最後等等。根據語境應合詞，目前的分詞系統卻分詞的例子為：吃到肥豬肉、吃到一塊珍饈、有時吃到十五粒等等。

<sup>3</sup> COCT 語料庫採自動分詞，分詞結果未經人工檢查；比較經人工檢查修訂的平衡語料庫 4.0 版 (<http://asbc.iis.sinica.edu.tw/>)，其中「吃到」有 127 筆，正確分詞有 104 筆，合分詞正確率僅有 81.8%，亦低於平衡語料庫之分詞正確率平均值 95%。

## 二、「V 到」結構的語意分析

### (一) 文獻探討

《說文》：「到，至也。」一般均認為「到」的原始語意為「到達」，它的賓語為處所，因此「V 到」也被視為連動結構或動補結構，如：走到圖書館。然而，隨著歷史演化，「到」的賓語已不限於處所，可能是時間、抽象概念或狀態，這也使「到」逐漸演化為引介處所、時間、範圍的介詞，如：走到一半；或是動相補語，如：買到車子等等。呂叔湘[2]認為「V 到」有以下五種語意：

1. V 到+名(受事)，表示動作達到目的或有了結果，例如：我今天收到了一封信、這個人好像在哪兒看到過。
2. V 到+名(處所)，表示人或物隨動作到達某地。例如：他回到了家鄉、他一直把我送到村口。
3. V 到+名(時間)，表示動作繼續到什麼時間，例如：等到明年暑假我再來看你、找到天亮還沒有找著李強。
4. V/Adj 到+名，表示動作或性質狀態達到某種程度，例如：他的視力已經減退到零點一了、事情已經發展到十分嚴重的地步。
5. V 到+動/小句，表示狀態達到的程度，例如：聲音高到不能再高了、船上平穩到跟平地上差不多。

趙元任[3]從聲調分析「V 到」；<sup>4</sup> Li & Thompson [4]<sup>5</sup>和劉月華[5]<sup>6</sup>從語意切入，將「V

---

<sup>4</sup> 趙元任 (1994:184-228)：

- (1) V 到+N(處所)，「到」的弱化語式，念 • de，例如：他搬到/得哪兒去了、頭髮掉到/得地下了。
- (2) V 到+V，表示時間、程度或範圍，例如：說到嘴乾 (同時表時間和程度)、累到走不動 (表示程度或範圍)。
- (3) V 到為方向補語 (全聲調)，例如：咱們居然趕到了、這件事情我早料到了。
- (4) V 到為狀態補語 (輕聲調)，例如：我碰到一件怪事。

<sup>5</sup> Li & Thompson (1992:59,286,314)：

- (1) V 到表示動作達成 (結果式動詞複合詞)，例如：看到、找到、想到
- (2) 到為 coverb(動介詞)，與名詞組成處所片語或方向片語，修飾動詞，例如：飛到上海、唸到第三行。

<sup>6</sup> 劉月華 (1996：304-305)：

- (1) 結果意義：
  - A. V 到表示動作達到了目的，例如：你丟的那支鋼筆找到了。

到」分為趨向義及結果義；王錦慧[6]從歷史發展的角度切入，綜合上述各家說法，進一步分析「到」的語法功能及「V到」的語意類型如下：

1. 「到」為趨向補語，「V到」表示時間或空間位移到某一個具體或抽象的終點。
  - (1) V到(+處所賓語)，例如：清華大學五分鐘就可走到、一滴都不會漏到油瓶外面。
  - (2) V到+時間賓語，例如：去年春天最好的春茶，放到今年也要失味。
  - (3) V到+處所賓語（隱含程度），例如：鬧到要離婚的地步。
2. 「到」為動相補語，「V到」表示動作的瞬間完成、實現。

V到(+受事賓語)，例如：八十元的雖然好吃又營養，卻不容易買到、港商在上海搶到不少一級地段的商業地產。
3. 「到」為結構助詞，起連接作用，「V到」隱含狀態所累積達到的程度。

V到+程度補語，例如：店面市場真是冷到不行。

王錦慧 1 下的(1)、(2)、(3)類可分別對應到呂叔湘的 2、3、4 類，王的 2、3 類則分別對應到呂的 1、5 類，趙、Li & Thompson 及劉的分析結果亦不出這五類範圍，因此，可確認「V到」的語意類型為五種。

## (二) 分詞原則及語意預測

參考上述分析，本文以「到」是否有「到達」語意來決定合分詞，凡有「到達」義者，不管是到達地點、時間點或某狀態，都分詞，該地點、時間點或某狀態視為「到」引介的賓語或補語；凡無「到達」義者則合詞，表示「到」已虛化為新詞的一部分，無具體語意。據此，我們共分析出七種合分詞原則，每一原則都用廣義知網的義原及語意角色來表達其語意[7]，如下所列：

- 
- B. V到表示動作持續到什麼時間，例如：昨天晚上我們談到十點半。
  - C. V到表示事情、狀態發展變化所達到的程度，例如：事情已經鬧到不可收拾的地步了。
- (2) 趨向意義：
- A. V到表示通過動作使事物達到某處，賓語一定是表示處所的詞語，例如：這個消息已經傳到外地了。

1. 有「到達」義者：

**I. V 到+地方/ →分詞**

到：引介地方賓語的介詞

語意：{V:LocationFin={地方},theme={}}

例如：走 到 櫃檯

**II. V 到+時間→分詞**

到：引介時間賓語的介詞

語意：{V:TimeFin={時間}}

例如：等 到 婚後

**III. V 到+範圍→分詞**

到：引介物體範圍的介詞

語意：{bel是:means={V},range={範圍},theme={}}

例如：鳥類 佔 到 近 1/2

**變化形式：V 到+範圍→分詞**

到：引介動作範圍的補語標記（complementizer）

語意：{ V:range={範圍}}

例如：想 到 一半、進行 到 第二輪

**IV. V 分類動詞到+範疇→分詞**

到：引介範疇的介詞

語意：{bel是:means={V 分類動詞},whole={範疇}}

例如：歸 到 運動類、分 到 廣西廠

**V. V 到+VP/S→分詞**

到：結構助詞

語意：{V:result={VP/S}}

例如：走 到 腳痠、佩服 到 五體投地

**變化形式：V 狀態動詞到+程度→分詞**

語意：{V 狀態動詞:degree={程度}}

例如：冷到不行、怪到不可置信的地步、高到某種程度

## 2. 無「到達」義者：

### VI. V 到+V 的賓語→合詞

到：虛化為動相補語

語意：{V:aspect={Vachieve|達成},goal={}}

例如：注意到他的成績、吃到好東西

**變化形式：V 到+賓語範圍（省略賓語）→合詞**

例如：分到一半(的蘋果)

### VII. V 到已詞彙化→收詞

到：無義

語意：合成與表面結構不同的新語意

例如：提到經濟、經驗老到

以上七條合分詞原則，與王（2013）的不同處在於，我們將王 1(1)、(2)類「V 到+處所/時間」的「到」改為介詞；(3)類「V 到+處所賓語（隱含程度）」的「到」改為助詞。由於這兩處的「到」都有「到達」義，因此都分詞處理。事實上，「到」也常獨用為處所或時間標記（location/time marker），例如：我從小到大都愛吃魚、他到紐約茱麗亞音樂院就學，因此我們依循 Li & Thompson 的分類，將處所及時間當作介詞賓語，使分詞的一致性更高。另外，王 1(3)類的「程度」與 3 的「程度補語」語意分隔不夠明顯，前者的例子如鬧到要離婚的地步，後者如快到令人眼花繚亂，僅以「V 到」後接成份為名詞或動詞作為區分，然而前者也可省略中心語「的地步」變成動詞，後者也可加上「的地步」變成名詞，因此我們均以「到」為結構助詞來歸納這些用例。

## 三、「V 到」結構的合分詞及語意表達

本節探討上述七條原則是否適用所有的「V 到」結構；其次，如何根據這七條原則來判斷「V 到」結構應合詞或分詞；最後，檢視句子成分的語意合成正確性。

### (一) 合分詞原則的適用性

在七條合分詞原則中，除了需收詞的最後一條外，其餘六條皆與「V 到」結構的動詞語

意相關，因此我們從語料庫中挑出 9 個分別代表「V 到」後接賓語、地方、時間及 V 為狀態動詞的例詞：吃到、湧到、等到、分到、嚇到、燙到、熱到、高到、老到等，並隨機挑選包含這些詞的 500 條例句，以觀察合分詞原則的適用性。部分例子及其適用的合分詞原則如(2)所列：

- (2) a.有時甚至 **吃到** 八點半→原則(II)  
b.一口一口都 **吃到** 蘋果的滋味→原則(VI)  
c.不肯輕易叫他 **吃到** 口→原則(I)  
d.佩吉特 **吃到** 一半就走了→原則(III)變化形式  
e.這些焦慮 **湧到** 梅西嘴邊→原則(I)  
f.河床容納不了上游 **湧到** 的河水→原則(VI)  
g.政府和軍部的車輛特需，源源 **湧到** →原則(VI)  
h.**等到** 這群僵屍消失之後→原則(II)  
i.終於 **等到** 機會出國→原則(VI)  
j.了解男女的差別要 **等到** 五歲以後→原則(II)  
k.行路的人都 **分到** 菜飯→原則(VI)  
l.他被 **分到** 廣西廠→原則(IV)  
m.我也 **分到** 一塊→原則(VI)變化形式  
n.孩子已 **嚇到** 發抖→原則(V)  
o.**嚇到** 了廚房的小狗→原則(VI)  
p.一直 **燙到** 脖子，乃至全身→原則(I)  
q.我很快地喝一口，**燙到** 了嘴→原則(VI)  
r.天氣 **熱到** 像在鐵板上燒烤一樣→原則(V)  
s.票價 **高到** 一百元一張→原則(V)變化形式  
t.她還沒有 **老到** 不能穿紅穿綠→原則(V)變化形式  
u.女人處理這種事非常小心而且手法**老到**→原則(VII)

在 500 個例句中，除了少數如(l)我們會猶豫它是符合原則(I)或原則(IV)、(s)是符合原則(V)的變化形式或原則(VI)的變化形式外，其他的例子都不難判斷，故推論七條合分詞原則應是足夠的。下一個階段則是測試它們在實作上的正確性。

## (二) 評估合分詞的自動化方法及正確率

本文所提出合分詞自動化的方法為：簡單的用賓語（或補語）是否為{地點 時點 狀態}來決定「到」是否有「到達」義，如果是前述三種賓語，表示有「到達」義，「V 到」分詞；反之，則合詞。簡易評估以上合分詞自動化方法的步驟有四：

第一步：依據合分詞原則針對 500 條例句做出標準答案(gold standard)。

第二步：用賓語（或補語）是否為{地點 時點 狀態}來決定「到」是否有「到達」義，如果是前述三種賓語，表示有「到達」義，「V 到」分詞；反之，則合詞。

第三步：在 500 條例句中標示出以上三類賓語（或補語），表示該句必須分詞。

第四步：依據標準答案，檢視第三步驟合分詞的正確率。以(2)的 21 條例句為例，標記(a)(h)(j)屬於時間；(n)(r)(t)屬於狀態；如果把身體部位也視為地方，(c)(e)(l)(p)(q)屬於地點，以上都應分詞，其餘則合詞。

這個簡易判準所得出的合分詞結果與標準答案比對後，發現在 21 條例句中，有 3 個錯誤，分別是(d)(q)(s)。(d)吃到一半的「一半」屬於動作範圍，因為也可能符合原則(I)的變形「分到一半（蘋果）」，所以簡易判準中未標記「範圍」類，故造成錯誤。(q)燙到了嘴的「嘴」是賓語，但身體部位也常當地方詞，如：一路燙到脖子、到全身，因此本文將身體部位一律視為地方，故造成錯誤。(s)高到一百元一張，「一百元一張」不是典型的狀態，而是表示狀態範圍的定量式複合詞，如：八百度、90 錢等等，此類與原則(VI)省略賓語的變化形式很像，故本文暫不標記，造成了錯誤。依此方法，針對 500 條例句合分詞的正確率為 93.4%，詳細分析如表一：

表一、以賓語類型決定「V 到」結構合分詞的正確率

	地點	時點	狀態	其他	總計
標記數	82	40	88	290	500
正確數	75	40	87	265	467
正確率(%)	91.5	100	98.9	91.4	93.4

上表的錯誤類型分析如下：

1. 賓語類型為「地點」的 7 個錯誤可分為兩類型：(1)身體部位應為賓語而不是地點（4 筆），如：燙到手；(2)地點應為賓語而不為地方（3 筆），如：分到一間房。



2. 賓語類型為「時點」者沒有錯誤。
3. 補語類型為「狀態」者只有一個錯誤例子：分到了從第二冊看起，其表面結構是動補，故依判準應分詞，但是深層語義其實是動賓，應合詞。
4. 不屬於上述三種類別者均標記為「其他」，全部合詞，其錯誤類型有三：(1)後接成份省略了時間、地方或狀態（3筆），例如：一家一家吃到 Z 字首的（餐廳）、官位高到國家的『三公』（的程度）。(2)後接成份詞類為普通名詞，但是類似身體部件，可作為地方詞用（1筆），如：把很多的情緒分到快樂系統或痛苦系統。(3)後接成份為定量複合詞，部份屬於（動作）範圍，應分詞，故造成錯誤（21筆），如：溫度高到八百度、吃到一半等等。

同時，本文也發現「分到」這個結構的錯誤率比較高，那是因為「分到」有語意方向性，「把 A 分到 B」與「A 被分到 B」同一方向，A 指的是被分配的事物，應分詞；當語意為「B 分到(了)A」時，B 指的是得到 A 的主語，應合詞。無論如何，依據賓語類型來判斷合分詞雖然有上述錯誤，但是「V 到」結構的合分詞正確率由 70% 提升到 93%，語意細緻度也提高了。

### (三)「V 到」結構的語意表達式

在 2.2 節中曾針對每一合分詞原則給予語意預測，並以廣義知網的語意表達式來呈現，本節要檢視依據 3.2 節的方法合分詞後，是否能根據語意預測合成「V 到」與其賓語（或補語）的語意。茲舉(2)的部份合成範例如(3)：

- (3) a. 有時甚至 **吃到** 八點半

原則(II) → {V:TimeFin={時間}} → {吃:TimeFin={八點半}}

- b. 一口一口都 **吃到** 蘋果的滋味

原則(VI) → {V:aspect={Vachieve|達成},goal={}} → {吃:aspect={Vachieve|達成},goal={蘋果的滋味}}

- d. 佩吉特 **吃到** 一半就走了

原則(III) → {V:range={範圍}} → {吃:range={一半}}

- e. 這些焦慮 **湧到** 梅西嘴邊

原則(I) → {V:LocationFin={地方},theme={}} → {湧:LocationFin={梅西嘴邊},

theme={這些焦慮}}

l.他被分到廣西廠

原則(IV) → {be|是 :means={V 分類動詞},whole={範疇}} → {be|是 :means={分},whole={廣西廠},theme={他}}

n.孩子已嚇到發抖

原則(V) → {V:result={VP/S}} → {嚇:result={發抖},experiencer={孩子}}

u.女人處理這種事非常小心而且手法老到

原則(VII) → {able|能:degree={very|很},theme={手法:owner={女人}}}

依據 3.2 節的簡易判準，將標記有{地點 時點 狀態}的例句分別依原則(I)(II)(V)進行語意合成，{其他}類則依原則(VI)合成。然後以人工檢視 500 條例句的語意合成正確率如表二：

表二、不同賓語類型的語意合成正確率

	地點	時點	狀態	其他	總計
標記數	82	40	88	290	500
語意合成 正確數	72	38	72	248	430
正確率(%)	87.8	95	81.8	85.5	86

原則上，若合分詞錯誤，則語意合成必然也錯誤。因此，依賓語類型來做語意合成時，錯誤類型會比合分詞的錯誤類型多，分析如下：

1. 賓語類型為「地點」的 10 個錯誤中除了原合分詞的 7 個錯誤外，還包括原先應屬於「狀態」類的 3 個例子，例如：價格也不會高到哪裡，「哪裡」被標記為地點，因此分詞，恰與「狀態」類須分詞一致，但是語意合成為{高:LocationFin={哪裡}}是錯誤的，正確應為{高:degree={not({very|很})}}。
2. 賓語類型為「時點」者有 2 個錯誤，屬於同一類，如：(把鹿的全部骨肉平分了，)分到鹿頭時...，依原則(II)語意合成為{分:TimeFin={TimePoint({鹿頭})}}顯然不正確，正確應為 TimePoint={分:goal={鹿頭}}，這屬於語意合成難免遭遇的例外情形。
3. 補語類型為「狀態」者有 16 個錯誤例子，除了原合分詞的 1 個錯誤例子外，另外 15 個錯誤分為兩個類型：(1)從小吃到大 (3 筆)，「大」在此處應表時間，故不能表達為{吃:result={大}}；(2)由於合成時，狀態類沒有細分變化類型，因此屬於

{V:degree={}} 全都標記為{V:result={VP/S}}，故造成錯誤，如：高到某個程度（12筆）合成為{高:result={某個程度}}。此種錯誤應可修正。

4. 賓語類型為「其他」者，除了原合分詞 25 個錯誤外，還有 17 個合成錯誤，分屬兩類：(1)賓語不是被處置的受事者(goal)，而是移動的客體(theme)（4 筆），如：精銳大軍(如潮水般)湧到，應合成為{湧:aspect={Vachieve|達成},theme={精銳大軍}}而非{湧:aspect={Vachieve|達成},goal={精銳大軍}}；(2)應收詞的一類也歸為「其他」，因此造成錯誤（13 筆）。如：手法老到依據原則(I)誤合成為{老:aspect={Vachieve|達成},goal={手法}}。此種錯誤亦可修正。

#### 四、 未來研究：自動合分詞的可能性及困難

提出合分詞原則的最終目標是能依此達到自動化的目的，以下兩部分是達成自動化可以努力的方向：

##### (一) 訂定「動詞語意」搭配「賓語類型」原則來改善合分詞

自動化即是要討論如何判別賓語是否為{地點 時點 狀態}，例如：訂定當動詞是移動動詞、姿態動詞、延伸動詞...等+到+地方詞時，賓語為{地點};分級狀態動詞(Gradable state verbs)...等+到+動賓或補語時，賓語為{狀態}等等可操作的細緻規律。事實上，上一節僅以賓語類型來判斷合分詞時，已從錯誤類型中發現可進一步改善此簡易判準的方法，即是必須結合動詞語意來做判斷，如(4)：

##### (4) 看到學校

上例中「學校」雖是地方詞，但因搭配的動詞不是移動動詞，故「學校」應為受事者，前面提到過身體部件有時作地點，有時作受事者，也可以利用動詞語意輔助判斷。此外，自動化合分詞還有一個技術性問題須克服，即判斷「V 到」結構後接成分中心語的問題，如(5)：

##### (5) a.吃到埃及最有名的料理

##### b.嚇到他們噤若寒蟬

以人工判斷，(5a)的賓語是料理，而不是埃及；雖然在大多數的情況下，嚇到的對象是人，但(5b)的賓語不是他們，而是以他們噤若寒蟬整句作為補語。也就是，同一個「V

到」結構，往往視語境而有複雜的合分詞情況，在自動化時必須用到剖析程式得到句結構，增添了判斷的難度，前面提到的「吃到」即是一例。

## (二) 利用把、被句型判斷合分詞

在中文複合動詞中，動補結構比其他結構的動詞容易與把字句和被字句搭配，因為把、被字句有以下搭配限制：1.動詞常有處置語意(disposal sense)；2.動作或狀態需有終點(end point)，這兩項特徵恰好符合大部份動補結構的語意特徵，即：處置動詞+結果終點或趨向終點，參照(6)所列出的對比例子，可以很容易看出這點。

- |                 |     |               |
|-----------------|-----|---------------|
| (6) 把百姓安撫住 (動補) | vs. | *把百姓安撫 (並列)   |
| 把招牌懸起 (動補)      | vs. | *把招牌高懸 (偏正)   |
| 把被害人殺死 (動補)     | vs. | *把被害人下毒手 (動賓) |

另外，雖然動補結構也可以直接加賓語，但是，將賓語前移(包括使用把、被句或 ergative 句型)的用例比較多，例如，統計動補結構「吹乾」一詞的 27 個例句中，直接加賓語只有 5 例，與把、被字句搭配使用則有 22 例。<sup>7</sup>如(7)：

- (7) 他將頭髮吹乾刷亮 (把字句)  
汗水早已被晚風吹乾 (被字句)  
杉木全部吹乾了 (ergative)

利用動補結構的搭配特徵，只要「V 到」結構的後接成分中心語可以和該動詞的把、被字句搭配，就可推論合詞機率較高；反之，則分詞機率較高。如(8)(9)所示：

- (8) a.走到學校；\*把學校走到；\*學校被走到→「走到」分詞  
b.看到學校；\*把學校看到；學校被看到→「看到」合詞
- (9) a.等到十點；\*把十點等到；\*十點被等到→「等到」分詞  
b.等到機會；?把機會等到；機會被等到→「等到」合詞

據此方法，如果把、被句都不成立，即分詞；有其中之一成立；即合詞。然而，實際應用時並不容易自動化判斷一個句子是否合理，或許可以利用大量語料檢索的方式，協助判斷這些句子是否出現，如：學校被看到、機會被等到等等，作為句子是否合理的依據。

<sup>7</sup> 使用國教院書面語語料庫。

## 五、 結論

本文討論中文動補結構中「V 到」的合分詞及語意區分問題，根據「到」是否有「到達」的語意，利用七條原則及簡易判準——即以賓語類型{地點 時點 狀態}為標準——採人工的方式進行標記，並評估合分詞正確率，結果證明「吃到」可以由目前的 70.6% 正確率提升到 94.5%。<sup>8</sup>以「V 到」結構整體來考量，在 9 個例詞 500 條隨機選取的例句中，合分詞正確率可達到 93.4%，基於此合分詞結果的語意合成正確率也達到 86%。顯示複雜的動補結構合分詞問題可經由簡易人工規則得到改善。本文的下一步計畫是將人工規則轉為自動化處理程式並檢驗其正確性。

## 六、 參考文獻

- [1] 柯華葳等，*華語文八年計畫「建置應用語料庫及標準體系」105 年工作計畫期中報告*，頁 5，臺北：國家教育研究院，2016。
- [2] 呂叔湘，*現代漢語八百詞*，頁 151-152，北京：商務印書館，1999。
- [3] 趙元任，*中國話的文法*，頁 184-228，臺灣：學生書局，1994。
- [4] Li, Charles N. & Sandra A, Thompson, *Mandarin Chinese*, pp.59,286,314, Taipei: The Crane Publishing Co., Ltd., 1992.
- [5] 劉月華，*實用現代漢語語法(繁體版)*，頁 304-305，台北：師大書苑，1996。
- [6] 王錦慧，論「V 到」結構的歷史發展，頁 227-252，*成大中文學報第四十一期*，2013。
- [7] Keh-Jiann Chen, Shu-Ling Huang, Yueh-Yin Shih, Yi-Jun Chen, Su-Chu Lin, You-Shan Chung, Ming-Hong Bai (2009), No. 09-01, *Lexical Semantic Representation and Semantic Composition-- An Introduction to E-HowNet*, Academia Sinica, CKIP group.

---

<sup>8</sup> 在國教院書面語料庫 711 個例句中，經簡易判準標記後重新合分詞，從原先的 209 個錯誤降為 39 個錯誤。