Computational Linguistics and Chinese Language Processing
Vol. 5, No. 1 , February 2000, pp. 1-18
© Computational Linguistics Society of R.O.C.

漢語動詞詞彙語意分析：表達模式與研究方法

A Lexical-Semantic Analysis of Mandarin Chinese Verbs: Representation and Methodology

張麗麗*、陳克健*、黃居仁+

Li-li Chang*, Keh-jiann Chen* and Chu-Ren Huang+

Abstract

在這篇文章中我們將簡單扼要地介紹詞庫小組分析動詞語意的作法，包含了理論部分以及分析的方法和步驟。這套理論架構是從實際分析的過程中逐漸形成的，但尚未完全成熟。截至目前爲止，我們已經分析了四十多組近義動詞和近十組動詞語意場，並初步架構了一套由語意屬性組成的動詞語意表達模式。這套理論和分析方法是建立在詞庫小組十幾年的研究基礎上，結合了眾人的力量逐步完成的。

本文嘗試統合眾人的研究所得，設法銜接理論與語言事實，並詳細介紹我們的研究方法。關於理論部分，詞庫小組已發表了許多篇論文。早期架構請參考 Tsai 等人[1998]、Huang 等人[1998]，最近的架構請參考 Huang 等人[見本期刊]。這些論文側重理論架構在學理上的探討，對語言事實涉及較少。本文則採用較多的語料，設法將理論落實。此外，詞庫小組也發表多篇論文探討了幾個特定近義詞組或語意場，像是 Chang 等人[見本期刊]探討情緒動詞、Liu 等人[1997]探討「建、造、蓋」、Liu 等人[見本期刊]探討「投、擲、丟、扔」、Liu 等人[1999]探討「趕、追」、以及 Chief 等人[見本期刊]探討「方便、便利」。本文也會綜合這些論文的發現以及其他未發表的研究成果在一個整體架構中。

*中央研究院 資訊科學研究所 Institute of Information Science, Academia Sinica
+中央研究院 語言學研究所 Institute of Linguistics, Academia Sinica

在文中我們將先介紹研究背景，緊接著進入主題，介紹理論架構的三個部分，分別是 1)動詞詞彙語意的表達模式; 2)動詞詞彙語意的區分和詞義延伸; 3)動詞詞彙語意和其他詞彙語意的結合情況。接下來，我們會詳列實際分析的方法、步驟和範例。最後作個總結。

這篇文章有兩個目的：一方面是將我們一再修正的理論和方法作個階段性的總結，作爲日後分析的參考依據；另一方面是將我們的理念和作法介紹給大家，希望能得到各界人士的批評指教。

In this paper we will briefly introduce the Module-Attribute Representation of Verbal Semantics (MARVS) and present in detail the methods used to analyze verbal semantics by the CKIP group. The theory and the methodology are based on the analysis of forty synonym pairs of verbs as well as verbs from ten different semantic fields.

This paper will focus on the linguistic data and our research methodology. For more information on the theoretical issues performing to MARVS, please see Huang *et al.* [this volume]. The research results published by the members of CKIP group on a certain synonym pairs or semantic fields will also be discussed in this paper, such as Chang *et al.* [this volume] on mental verbs, Liu *et al.* [1997] on building verbs *JIAN*, *GAI* and *ZAO*, Liu *et al.* [this volume] on throwing verbs *TOU*, *ZHI*, *DIU* and *RENG*, Liu *et al.* [1999] on chasing verbs *ZHUI* and *GAN*, and Chief *et al.* [this volume] on verbs meaning "beneficial", *FANGBIAN* and *BIANLI*.

This paper will be organized in the following way. In section 1 we will first introduce our basic ideas on verbal semantics. In section 2 we will discuss three related research topics, i.e. the MARVS theory, the distinction and extension of verbal meanings, and the co-occurrence of verbs with certain sentence patterns or adjuncts. In section 3 the methodology used for analyzing synonym pairs and verbs in a particular semantic field will be presented. In section 4 we will give an example of the near-synonym verbs KUAILE "happy" and GAOXING "glad" and show precisely what to observe, how to compare and how to explain the differences in detail.

This paper is a record of our research methodology and will be used as a technical guide for the CKIP group. We will keep on modifying our research methods and the theory in the future and we look forward to feedback from readers of this paper.

1. 研究背景：

近來詞彙語意學及詞彙驅動的語法研究都不約而同地得到一個相同的觀點：許多語法現象可以由詞彙語意預測或決定。這個論點與認知學派和功能學派一向的主張也相符。詞彙語意能決定許多重要的語法規律，有許多詞和詞之間、詞和句型之間的搭配關係和焦點轉移，都可以成功地藉由詞彙語意結合的邏輯性來解釋和預測。我們的研究也是建立在這樣的基本論點上。我們研究動詞的詞彙語意，就是希望能從動詞的詞彙語意特性整理出更具普遍性的語言規律，像是動詞的論元結構、功能的分佈、搭配關係、甚至也包含了動詞搭配不同成分時詞彙語意焦點的轉移情形。瞭解語意結合現象以及語意語法互動關係後，我們才能進一步探討動詞詞彙語意的表達模式，釐清動詞詞義的分項和延伸的現象。在整個研究過程中，我們將逐步抽取語意屬性，利用語意屬性間的邏輯關係來表示動詞的各種搭配關係以及詞彙語意焦點轉移情形，並利用語意屬性架構動詞語意。

這套動詞語意表達模式是由語意屬性所架構成。由於這些語意屬性代表具有左右詞彙用法的語意成分，因此我們所提出的表達模式將超越規範式的語法模式，成爲具有解釋性和預測性的語意關係模式。在此模式中我們將清楚界定各語意屬性之間的互容關係、互斥關係、邏輯關係和影響焦點轉移的情況，並利用這套語意屬性來表示和區分每個動詞、狀語和句型的語意特性。因此，理想中每個動詞的每個義項都有各自的一套由語意屬性所組成的事件訊息結構，從這套屬性我們可以推論出這個動詞的使用情形。

在研究方法上，我們要找出這些決定詞彙用法的語意屬性，可以從許多不同的角度下手：從近義動詞下手、從動詞語意場下手、從狀語下手、從句型下手、從論元下手。不同的角度各有不同的著力點，各有不同的觀察優勢，也會帶來不同的研究成果。從近義動詞下手可以觀察到近義動詞的差異處，抽離出區分這組動詞的關鍵語意屬性。從動詞語意場下手可以觀察到整組動詞的共同處，像是一致的論元結構、相同的搭配關係、和共通的語法功能分佈狀況…，以抽離出一組互相牽制的語意屬性。從動詞語意場下手也可以觀察同一語意場中最明顯的對比，並抽離出關鍵語意屬性。從狀語下手可以區分狀語類型，觀察每個狀語所能搭配的動詞類型，並抽離出決定這種搭配關係的語意屬性。從句型和論元下手所能進行的觀察和觀察成果和狀語類似。特別是在事態（Aktionsart or verbal aspect）缺乏構詞標記時，我們可以從動詞所搭配的狀語與句型得到相關的訊息。

2. 研究主題：

正如研究背景的說明，我們的觀察和研究主題包括三方面：1)動詞詞彙語意的表達模式；2)動詞詞彙語意的區分和詞彙語意延伸； 3)動詞詞彙語意和其他詞彙語意或結構語意的結合情況。

2.1 動詞詞彙語意的表達方式

爲了明確掌握詞彙語意的影響力，我們需要一套由語意屬性組成的架構來表達詞彙語

意。動詞表達的概念是「事件」，其內容涉及事件的參與者、事件的執行方式、事件進行的時間長短和地點…等。但是通常一個句子所表達的事件內容，可能只是完整事件的一部份，由動詞及其論元及附加成分所表示的語意成分結合而成。句子所能表達的事件內容及方式主要是受核心動詞限制的，所以我們定出一個動詞語意表達模式“Module-Attribute Representation of Verbal Semantics (MARVS)”，簡稱爲”MARVS Representation”來表達每一個動詞義項的事件訊息結構[Huang and Ahrens, 1999]。

事件訊息結構中所含的訊息共有四種，分成兩個層面：第一個層面的訊息直指事件本身，包含了事件類型和特質屬性(Inherent Attributes)；第二個層面的訊息涉及事件的參與成分，包含了參與角色(Roles)和角色內部屬性(Role-internal Attributes)[1]。事件類型也是由一個個語意單位所組合而成的，我們稱這些語意單位爲事件模組(Event Module)，它和其他語意屬性不同處在於，這些事件模組可以互相結合，組合架構成一個個事件結構(Event Structure)，每個事件結構則代表一個事件類型。動詞的每個意義都有其事件訊息結構，其整體組織架構如下：

- 動詞
  - 意義一：
  - 事件訊息結構：
    - 事件類型
    - 特質屬性
    - 參與角色
    - 角色內部屬性
  - 意義二：
  - 事件訊息結構：
    - 事件類型
    - 特質屬性
    - 參與角色
    - 角色內部屬性

目前我們所整理出的事件類型及語意屬性將一一分述如下：

a. 事件類型(Event Types)及事件結構(Event Structures)

我們認爲一個動詞所能表達的事件類型主要是由五種基本事件模組(event module)組成，即各種不同事件類型都是由這五種基本事件模組組合而成的。這五個模組描述了：

[1] 稱之爲角色的主要原因之一是詞彙語義中能影響預測語法表現的參與者並不一定是句法中受次類劃分的論元。

a)事件的基本類型，包括了過程(process)、狀態(state)及階段(stage)、b)事件的長短、以及c)事件的端點 (Boundary)。

五個基本事件模組：

•	端點 (Boundary)
////////	過程 (Process)
^^^^^	階段 (Stage)
——	狀態 (State)
/	瞬時 (Punctuality)

端點：在我們訂出的事件結構中，是以「端點(boundary)」這個事件模組表示事件的端點，包含「開端」及「終點」，我們用一個黑點「・」表示。在語言學的文章中，慣用 telic 或 endpoint 來指稱有明確「終點」的動詞。但是依據我們的觀察，在現代漢語中當我們指涉事件的端點時，常常是不區分開端或終點的。例如當動詞接上「了」，可能是指事件的開端：「開會了」，也可能是指事件的終端：「搬家了」。又例如副詞「馬上」只能接含有端點事件模組的動詞，不論這動詞是指開端，如「馬上處理」，還是指終點，如「馬上交出」。而不具有端點事件模組的動詞就不能搭配「馬上」，像是「馬上散步」不好，要說成「馬上去散步」。因此我們用端點同時表示事件的開端和終點。若一個事件只有端點，沒有過程，便是所謂的完整(complete)事件。

過程：指能夠持續進行的動作，我們用一連串的斜線「////////」表示。在一般語言學著作中，這種事件類型多半用 activity 來表示。我們用「過程(process)」是因爲它僅是事件類型的組成成分，不表示一個完整的事件類型。爲避免誤解，故用「過程(process)」表示。

階段：指包含狀態改變、具有階段性的事件形態，所以我們用一連串的曲折線「^^^^^」表示。它和過程的差別在於過程是由一連串動作組成，像是「建房子」包含了打地基、安置鋼筋、砌牆…等一連串動作，而階段則包含一連串不同的狀態，像是「凋謝、上昇、瀰漫」包含了一連串相接續卻又有細微差別的狀態。

狀態：指持續不改變的事件，我們用一條延續的橫線「——」表示。

瞬時：指該事件只存在於一個時間點上，我們用單條斜線「/」表示。具有瞬時模組的動詞表現在語法上的特性是，不能搭配持續的時段，例如「打算」屬瞬時，因爲我們不說「我打算去看電影打算了很久。」

一個動詞的事件類型是它所能表達的事件形態(Situation Types)的總合，所以該動詞的事件結構也是一個綜合體。「事件形態」是指在語句中，動詞搭配角色或附加成分後所能表示的個別的事件呈現。因此一個動詞在不同語句中所呈現的事件形態往往不同，也就是說，一個動詞所能表現的事件形態往往是多樣的。例如，「建」可以表達三種事

件形態：可以指一個過程「//////」，如「他在建房子」；也可以同時指出該過程的開端「•//////」，如「這房子建了五年了還沒建好」；也可以同時指出該過程的終點「//////•」，如「這間房子建於民國五十一年」。因此我們就綜合這三個事件形態在一個事件結構中，以「•//////•」這個事件結構表達「建」的事件類型。這個事件結構是「建」能呈現的事件形態的總合：第一個點指過程的開端，第二個點指過程的終點，中間的「//////」指過程本身。必須注意的是，並沒有一個句子同時指「建」的開端、過程和終點。也就是說在真實的語句中，並沒有「•//////•」這樣的事件形態呈現方式。但是如果省去了其中任一個模組，詞彙訊息中對「建」這個動詞的事件類型的描述就不算完整。

剛剛以「建」爲例，探討了動詞的事件形態呈現情況以及我們標示動詞事件類型的基本原則。我們依照同樣的精神檢驗了現代漢語中的各式動詞，依據動詞所能表示的事件形態，找出十三種不同的事件類型。一個動詞所指涉的事件類型可能是單純的，也可能是組合的。在指涉事件的開端或終端時每個動詞各有不同呈現能力，有的能指涉開端和終端，有的只能指涉開端，有的只能指涉終端，有的開端終端通通不指涉。依照這些區別，我們將這十三種事件類型區分成三大類：一、表示核心事件的動詞只能表示單一的基本事件模組；二、表示簡單事件的動詞也只能表示單一的基本事件模組，不過卻可以指涉該事件的端點；三、表示組合事件的動詞能表示兩種基本事件模組。

我們相信在此所提的事件模組具有語言共通性(language universality)，其共通性不在於所提事件類型，而在於模組形式。我們在此列出的十三個類型也許不是所有語言共有的分類，透過更廣泛的分析，或運用到其他語言，很可能還會發現其他新的事件類型，特別是組合事件類型。但是我們認爲模組形式是所有語言在事件類型上的一種共性，所有的語言的事件類型都是以模組的方式形成的，都能以模組方式適切地表達。此架構具有隨時新增的彈性，能夠依據每個語言的特色，適切地描繪該語言所有的事件類型。自然，透過跨語言的研究也可以找出此模組中最常出現的類型，一樣能看出什麼樣的模組形態是最常見的。當然這還需要更多的研究才能達到，不過我們預測至少下列的核心事件應該是所有語言所共有的。

(一) 表示核心事件（nuclear event）的動詞：這一類動詞只能呈現一種單純的事件形態。

////// 散步、旅行　　動作（activity）

——快樂、疲倦、恨…　均質狀態（homogeneous state）

/　打算 瞬時（punctuality）

•　死、斷、醉…　完整（complete）

(二) 表示簡單事件（simplex event）的動詞：這一類動詞只能表示單一的事件模組，但是可以同時指涉事件的端點。

•//////　下雨、開會、追趕、考慮…啓始動作（inchoative activity）

• //////• 建、製造、吃、吃飯… 始終動作（bounded activity）

• ^^^^ 上昇、下降、縮小、衰老…啓始階段（inchoative stage）

• ^^^^• 凋謝 始終階段（bounded stage）

• —— 高興、累、瘦、破、懂… 啓始狀態（inchoative state）

╱• 打死、打破、騎累、答對…結果（resultative）

•╱ 出發、畢業、離開、抵達…瞬間完成（completive punctuality）

(三) 表示複合事件（composite event）的動詞：這類動詞可以指涉過程或過程完成後的狀態，但是這兩種事件模組不能同時呈現。

╱• ——坐、躲、包圍、放、知道…完整結果（completive resultative[2]）

• //////• —— 穿、掛、綁、連接、瞭解…動態雙元（dual process-state）

上面的圖示和分類具有下列幾點重要的意義和突破：

1. 這十三個事件類型主要是依據幾個標準判定。首先，這個動詞所表達的事件概念是否延續一定的時段，若是，則該事件可能爲過程、狀態或階段；若不是，則該事件爲瞬時或完整。其次，該動詞可否指涉事件的端點，若不可，則爲核心事件；若可以，則爲簡單事件。其三，判定該動詞是過程、狀態還是階段，或是兩種以上的複合事件。

2. 以上判定的每個步驟都有一定的方法可以依循。判定一個動詞是否可以延續一段時間，可以從該動詞和某些詞或片語的搭配狀況來判斷，如狀語「繼續、下去、一直、不斷…」，或表示時間長度的片語，像是「…了三個小時、從以前到現在」。因此，從「不斷咒罵、從小到大都很漂亮」這樣的句子中，我們知道「咒罵」和「漂亮」都是可以持續的事件。判定一個動詞是否可以指涉事件的端點，可以從動詞和「了」的搭配狀況來看。例如我們不說「他旅行了」，但說「他去旅行了」，表示「旅行」這個動詞不被用來指「旅行」這個事件的開端，所以它只能表達「旅

[2] 我們區分了「結果」(resultative)和「完整結果」(completive resultative)兩種類型，二者有類似處，但可以清楚區分。在 MARVS 的架構中，前者是簡單事件，後者是複雜事件。「打死」是簡單事件，透過一個動作的方式瞬間達到一個結果，在我們的架構中分類爲「結果」。「坐」則含括了兩個事件類型：在「請坐」中，「坐」指涉一個瞬間就完成的「坐下來」的動作，「台上坐著一個人」則指涉一個持續的「在某個地點坐著」的狀態。在我們的架構中，這種包含一個瞬間動作和一個持續狀態的事件類型稱爲「完整結果」。

行」的核心形態。判定一個動詞是過程、狀態還是階段也有一定的標準。表示階段的事件形態的動詞可以搭配「持續」，同時意義上含有事件的變化，像是「持續上昇、持續下降」。過程和狀態的區分方法，就如同一般區分動作和狀態一樣，大部分表示過程的動詞可以搭配「正、正在、在」，而大部分表示狀態的動詞則不然。以上的討論也說明了句中事態(aspectual)成分和動詞間的搭配關係可以由此一表達方式推導出來。

3. 由於我們的架構是利用基本事件模組組合，所以事件類型隨時可以增加，就能更靈活精確地表達各種可能的事件形態及其變化。在一般的事件表達模式中，事件類型數量是固定的，在對動詞分類時，往往有不同事件特性卻歸入同類的現象發生。依照我們組合精神，就不會產生這種現象。

4. 依據較常被採用的 Vendler[1957, 1967]對事態(aspect, Aktionsart)的分類，共分出四種基本類型，分別是：activity, state, achievement, accomplishment。我們可以看到，這四種類型在我們的架構中都可以表達，並依其組合方式明白區分。而且截至目前爲止，我們區分了十三種事件類型，比起 Vendler 的四類來說，是更加精細了。

5. 從解釋性來看，我們的架構也達到更高的功效。例如，我們區分了兩種狀態：均質狀態和啓始狀態，其差別就在於啓始狀態之前有一個端點，表示該動詞的語意有一個明白的開端。例如，在「他一高興就會手舞足蹈」中，「高興」蘊含了狀態的轉變，指經驗者從一般的心理狀態進入高興的心理狀態，所以「高興」是個啓始狀態，其事件結構是「・——」。但是均質狀態就無法有這樣的用法，日常生活中見不到「他一快樂了就手舞足蹈」這樣的句子，可見快樂僅能表示均質的、持續的狀態，其事件結構就是「——」，不含端點「・」。

b. 特質屬性(Inherent Attributes)

特質屬性是用來描述事件本身特質的屬性。一個動詞的特質屬性非關乎動詞的事件類型或論元，它是和動詞核心概念緊緊相關。我們目前找到的特質屬性不多，茲舉三例：

控制(control)：指主事者可以操控該動作或狀態，帶有此屬性的動詞可以搭配命令句。基本說來，大部分動作動詞帶有控制的屬性，而大部分狀態動詞則不帶。但是也有例外，例如「到達」是動作動詞，卻不可操控，所以我們不說「別到達」。但「高興」雖是狀態動詞，卻可操控，因爲我們看得到「你別太高興！」這樣的句子。因此，在我們的架構中，「到達」不帶控制的屬性，但「高興」會帶有此屬性。

事實(realized)：我們發現有些動詞不能被「不」否定，我們認爲這是由此動作的核心概念決定的，用「事實」此屬性來表示此語意特質。「不」是用來否定主事者從事一個動作的意願，或是否定一個狀態。不能被「不」否定的動詞，表示該動詞的核心概念中已經表示這是一個「事實」，不容我

們去否定執行的意願或其存在的狀態。例如，我們可以說「我不打算去、我不準備去」，卻不能說「我不計畫去」，要否定原先的計畫，只能說「我計畫不去了」。當我說出「我計畫…」，該語言行爲已成立，所以不容再否定。當我們對一件事作猜測時，我們可以說「我不認爲他走了」，卻不能說「我不以爲他走了、我不想[3]他走了」。這也是因爲當我一說出「以爲」、「想」，這個猜測的行爲就成立了，就不容再否定。這就是爲什麼凡是不能被「不」否定的動詞我們都給它「事實」的語意屬性。

處置(disposal)：我們認爲探討漢語把字句所提的處置概念也是一個特質屬性。帶有處置屬性的動詞，如「打破、趕走」，清晰表示賓語被處置後的結果，可直接用於把字句。這類動詞同屬同一個事態類型，但是它們所具有的概念比事態還要複雜，還包含對賓語處置後的狀況，故用「處置」這個屬性來統稱。

c. 參與角色(Roles)

動詞可攜帶的每個角色都視爲其事件訊息結構上的一個語意屬性。詞庫小組原定有十一個基本論元，皆可視爲動詞表達事件中的參與角色[其定義及例子請參見詞庫小組, 1993]：主事者(agent)、受事者(theme)、對象(goal)、經驗者(experiencer)、接收者(recepient)、來源(source)、肇始者(causer)、地點(location)、範圍(range)、目標(target)、比較者(comparison)。此外，根據近來分析我們得到的新角色如下：

事因(cause)：表示事件之肇因，像是「減肥真累、我很高興你能來」。事因和詞庫原就定有的肇始者(causer)並不相同，肇始者必須是肇始事件(causative event)的行爲者，像是「顏色的變換豐富了視覺感受、大家的問候溫暖了我的心」。

漸生題旨(incremental theme)：表示經歷該事件後產生之結果物，像是「烤蛋糕、擺棋譜、綁蝴蝶結、紮辮子」。

填位(locus)：表示被填補之位置，像是「補洞、填海、裝箱」。

d. 角色內部屬性(Role Internal Attributes):

角色內部屬性，顧名思義，是指角色本身的特性。因爲該特性會影響到動詞的使用狀況，所以特別標示出來。我們目前所找到的角色內部屬性也不多。

意志(volition)：指主事者具有行使此動作的意圖。例如「遇見」和「等」的不同在於前者不帶意志，後者帶意志。因此，我們不能說「我一定要遇見他」，

[3] 動詞「想」有好幾個意思，在此指的是表猜測的意思，像是「我想他已經走了」。

只能說「我一定要等他」。[4]

2.2 動詞的詞彙語意區分和詞義延伸

動詞詞義的區分是一個相當複雜難解的問題，大致上可以分成兩個重要議題：意義的區分和意義的延伸。截至目前爲止我們並沒有深入的探討，只能列出一些初步的觀察。

a. 如何判定一個動詞具有幾個詞義(sense)、如何區分一個動詞不同的意義、並找出區分動詞意義（和義面）的原則。目前對這方面的問題初步探索得到的暫訂原則如下：

原則一、一個動詞具有幾個意義是由其所能組成的「事件訊息結構」的數目所決定。

原則二、可以由規律的詞意延伸(meaning extension)所預測的詞彙語意變化應列在同一個意義下。

由原則一得到的判斷方法是把無法統合在一個事件訊息結構中的意義呈現分列在不同的詞義(sense)中。原則二則突顯出詞義延伸研究的重要性。

b. 詞彙語意如何延伸：我們認爲動詞的意義延伸是經由邏輯推導(logical implicature)或是經驗知識結構[qualia structure, Pustejovsky 1995]中的層面轉移而得到的，它和本義的差別在於意義焦點的轉移。這些都應算是同一意義的不同呈現，而不是不同的意義。當動詞搭配不同的論元、狀語或句型時，語意會有所變化，這些都不是動詞詞彙語意完全改變了，只能算是事件焦點轉移的結果。目前這部分的觀察還不夠完整，部分現象羅列如下，其中詞彙語意改變的規律與導因，還需進一步的區分：

(一) 時態改變：正**走**在街上[-端點]／**走**了[+端點]；他正在**穿**衣服/**穿**著一件襯衫。

(二) 論元結構改變：行爲(受事者)**端正**/我們(主事者)要**端正**社會風氣(受事者)；他的打扮(受事者)很**奇怪**/我(經驗者)正**奇怪**你怎麼不在家(對象)。

(三) 論元類型改變：他**跑**操場／他專**跑**議會新聞；**走**小巷子／**走**學術路線；他**送**朋友出門/他**送**一批貨到南部。

---

[4] 特質屬性「控制」和角色內部屬性「意志」很接近，大凡能「控制」的動詞通常也需要主事者有執行的「意志」。不過二者之間並沒有必然的關連，例如有些動詞是可以控制，但主事者並不具有執行的意願，像是「高興」、「傷心」，這可以從句型的搭配看出。「高興」、「傷心」可搭配祈使句，如「你別太高興」、「別傷心了」。這顯示二者具有「控制」的屬性，因爲祈使句是要求他人「控制」其行爲。但「高興」、「傷心」不能搭配表意願的法相詞(modal verbs)或動詞，例如我們不能說「我一定要高興」或「我一定要傷心」，這顯示這兩個動詞不具有「意志」此屬性。

2.3 動詞詞彙語意和其他詞彙語意或結構語意的結合情況：

研究動詞詞彙語意和其他詞彙語意的結合情況就是在瞭解動詞和論元、狀語以及句型之間的互容關係、互斥關係以及焦點轉移情形。我們可以從動詞出發，研究每組動詞所搭配的角色、狀語以及句型，也可以從另一個方向出發，研究各類角色、狀語和句型所能搭配的動詞。以下僅舉一些例子，兼有兩個角度的觀察所得。

a. 角色和動詞

例如情緒動詞，像是「高興、難過」，搭配一個必要論元經驗者和一個非必要論元事因。又例如理解動詞，像是「知道、瞭解」，搭配兩個必要論元：經驗者和對象。

(1) 我們(經驗者)很**高興**創刊號終於發行了(事因)。

(2) 高中女生的自殺(事因)更令人(經驗者)**難過**。

(3) 我(經驗者)**知道**我平常是任性了一些(對象)。

(4) 大家(經驗者)也都**瞭解**這一點(對象)。

又例如主事者搭配過程動詞，不搭配狀態動詞；經驗者只搭配狀態動詞，不搭配過程動詞。

b. 狀語和動詞

例如情緒動詞常搭配程度副詞或程度補語，像是「很快樂、非常快樂、十分快樂、快樂得不得了、快樂得很、快樂無比…」。理解動詞中的「得知」和「體會」不能搭配否定副詞「不」，我們不說「不得知、不體會」。

又例如「馬上」必須搭配事件結構中含有端點動詞，所以不能說「馬上旅行」，要說「馬上去旅行」；相對地，「一直、持續」只能搭配事件結構中不含端點的動詞，所以不說「一直打破了」，只能說「一直打」。

c. 句型和動詞

例如情緒動詞除了直接前接經驗者，像是「他很快樂」，也會前接「覺得、感到」，出現在「他覺得很快樂、他感到十分快樂」這樣的句型中。

又例如把字句搭配帶有處置語意屬性的動詞組。動詞「搬運」不帶處置屬性，所以不能說「把貨物搬運」。但是「搬運到倉庫、搬運出去」就帶有處置屬性，因此我們可以說「把貨物搬運到倉庫、把貨物搬運出去」。又例如地方片語倒置句(locative inversion construction)搭配姿態動詞、放置動詞：「台上站著一個人、桌上放著一本書」。

3. 分析方法和步驟

動詞詞彙語意分析可以從許多不同的角度下手，去掌握各種由詞彙語意決定的語言規律。在過去兩年的研究中，我們主要是從近義動詞和動詞語意場下手，其方法和步驟如下：

1. 從下列幾點觀察整理近義動詞或同一個語意場內動詞：
   a. 意義和詞類分項
   b. 語法功能異同：作謂語、定語、狀語、補語、名物化的分佈情形
   c. 參與角色：所搭配的必要論元和非必要論元
   d. 時態
   e. 搭配情形： 作謂語時所搭配的狀語、補語、句型
      作定語時所搭配的名詞中心語、
      作狀語或補語所配的動詞中心語
2. 找出近義動詞或同一個語意場內動詞共有的語法表現，建構基本的事件訊息結構。例如情緒動詞都可以前接程度副詞「很、非常、十分…」，其事件結構中都帶有狀態的事件模組。又，情緒動詞的論元結構除了作出情緒反應的體驗者外，還應該帶有事因。因爲情緒動詞的前後常常會出現一個子句，表示引發該情緒的原因，雖然它出現的位置不一，見下例，卻表示共同的語意訊息。
   (5) 你那麼晚回來，讓媽媽覺得很難過。
   (6) 更令我感到難過的是他還有一個交往七年的女友。
   (7) 我很難過他這麼年輕就去世了。
   (8) 全家人也爲您出了車禍而感到難過。
3. 找出近義動詞或同一個語意場內動詞的用法差距，並設法以語意屬性解釋並區分之。例如情緒動詞雖然在時態上都屬狀態動詞，其事件結構中都帶有狀態，但是可以再細分成兩大類，並以有無端點來區分。[5]又例如「趕」後接的名詞片語有詞意誘迫(meaning coercion)，但是「追」後接的名詞片語沒有。當我們說「趕公車、趕報告」，是指「趕著『搭』公車、趕著『寫』報告」。這「公車」和「報告」出現在「趕」的後面是指涉一個事件，而非單純的名詞。而從「公車」延伸到「搭公車」是從公車的經驗知識結構中的主事層面（agentive）推導出來的。[6]

[5] 這兩大類動詞可以依據五種語法行爲區分開來，請參考Chang, Li-li等人[1999]。
[6] 請參考Liu, Mei-Chun等人[1999]以及Pustejovsky[1995]。

4. 分析範例：以「快樂、高興」爲例

以下的分析以 Tsai 等人[1998]的分析架構爲主軸，並佐以 Chang 等人[1999]文中提出的補充。表格中的數字和解釋則以 Chang 等人[1999]爲主。

4.1 觀察

a. 意義和詞類分項

快樂：一個語意，一個詞類（不及物狀態動詞）。

高興：一個語意，一個詞類（後接句賓的狀態動詞）。

b. 語法功能分佈：從語料庫中觀察這兩個動詞所扮演的功能，如作謂語、定語、狀語、 補語、名物化的分佈情形

「快樂」以定語功能和名物化的傾向高。

「高興」以謂語功能爲主。

表 1.「快樂」和「高興」句法功能分佈比例

	謂語功能	補語功能	狀語功能	定語功能	名物化
快樂	37.79%	5.20%	5.73%	24.84%	26.43%
高興	85.05%	1.35%	11.96%	1.35%	0.30%

c. 論元結構：該詞有幾個論元，每個論元個扮演什麼語意角色

快樂：經驗者

高興：經驗者、事因

結論：「快樂」只帶有一個論元，「高興」多帶了事因，是個句賓。

表 2.「快樂」和「高興」後接句賓次數

	句賓表事因
快樂 365	--
高興 280	69

d. 事態

快樂：少加時態標記「了」。在語料庫中，接在「快樂」之後的「了」多半不是時態標記，而是作語尾助詞，像是「我實在太快樂了」、「我今天最快樂了」。在下列表格中的數字是「快樂」後接「了」的所有句數，並未將語尾助詞的句數刪除。

高興：較常和時態標記「了」搭配，例如「客人高興了會賞你錢」。

結論：「高興」比「快樂」更常和時態標記「了」搭配。

表 3.「快樂」和「高興」後接「了」的次數

	V - 了
快樂	10
高興	20

e. 搭配情形：根據其不同語法功能，分別觀察每個功能重要搭配成分之類型及分佈，包括作謂語時所搭配的狀語、補語和句型、作狀語或補語所配的動詞中心語、作定語時所搭配的名詞中心語。

(一)「快樂」和「高興」作謂語時所搭配狀語和補語並無明顯差別。

(二)「快樂」和「高興」作謂語時所搭配句型：

快樂：

受事者 ＜V：他很快樂

V＜ 得 ＜ 補語：一聽到披薩，真是快樂得不得了。

祝 ＜ 你 ＜V：祝你快樂！

高興：

受事者 ＜V：他很高興

受事者 ＜V＜ 事因：他很高興張三來了。

V＜ 得 ＜ 補語：我和妹妹都高興得大聲叫好。

別 ＜V：你別高興得太早。

受事者 ＜ 不值得 ＜V：這件事不值得高興。

結論：「快樂」和「高興」搭配句型的差異在於「快樂」可接祈願句，「高興」可接祈使句和價值判斷句。

表 4.「快樂」和「高興」搭配祈願句、祈使句和價值判斷句的次數

	祈願句	祈使句	價值判斷句
快樂	8	--	--
高興	--	6	6

(三)「快樂」和「高興」作定語時所搭配的名詞中心語

高興：高興的時候、高興的事情、高興的神色、高興的神情、高興的心情

快樂：

時間：快樂歲月、快樂的時候、快樂時光、快樂的時刻、快樂的關鍵、快樂的日子、快樂的一天、快樂的夜晚、快樂的節日、快樂的兒童節、快樂的端午節、快樂的耶誕節、快樂的時光、快樂的春天、快樂的季節、快樂的兒時、快樂童年、快樂的晚年、快樂的人生

事件：快樂的事、快樂的事物、快樂的生態之旅、快樂的舞會、快樂的記憶、快樂的婚姻、快樂的性生活、快樂的開始、快樂的結束

心情：快樂的心、快樂的情緒、快樂感覺、快樂的感覺

表情：快樂的臉兒、快樂的笑容、快樂的眼淚

人　：快樂的人、快樂的孩子、快樂的小天使、快樂的創作人、快樂的小男人、快樂的母親、快樂的修女、快樂的員工、快樂的老人、快樂的朋友、快樂的賣鹽人、快樂的上班族、快樂寡婦、快樂郵差、快樂的奴隸、快樂的工程師、快樂神仙、快樂的戰痘勝利者、快樂的自己、快樂的自我、快樂的小青蛙、快樂的小麻雀、快樂的浪花、快樂的靈魂、快樂的家、快樂家庭

地方：快樂的環境、快樂的學習環境、快樂的家庭聚會場所

氣氛：快樂的氣氛、快樂景象、快樂畫面、快樂的氣息

其他：快樂境界、快樂的能力、快樂笛聲、快樂的唱法、快樂的動作、快樂的行爲、快樂的戀情、快樂的部分、快樂的情趣、快樂的慾念、快樂的意義、快樂的觀念、快樂的第一要件、快樂的滋味、快樂哲學、快樂的本質、快樂的來源、快樂的泉源、快樂的成分、快樂翻譯紙、快樂紙、快樂筆、快樂學園、快樂農村、快樂星

結論：「快樂」所能修飾的名詞組比起「高興」要多得多,「高興」做定語的次數很少，所能修飾的名詞總類也少。

(四)「快樂」和「高興」作狀語時所搭配的動詞中心語並無明顯差別。

(五)「快樂」和「高興」作補語時所搭配的動詞中心語

高興: 玩得好高興, 看得很高興, 跳舞跳得這樣高興, 讀得也很高興, 吃得很高興, 唱得高興, 欣賞得很高興, 拆得正高興, 談得正高興

快樂: 活得快樂, 過得快樂, 生活過得快樂, 唸得快樂, 玩得很快樂, 忙得很快樂

結論:「高興」搭配短暫性動詞;而「快樂」兼可搭配短暫性和長期性動詞。

4.2 比較

接下來我們要根據上述觀察歸納出這組近義詞的規律性對比。「快樂」和「高興」主要差異有以下六點:

a. 從語法功能分佈來看,「快樂」名物化和作定語的情形相當普遍,「高興」則少。

b. 「高興」後面可以接句子,「快樂」則無此用法。

c. 「高興」和時態標記「了」搭配的頻率比「快樂」高。

d. 「快樂」可以出現在祈願句,「高興」則不行;「高興」可以出現在祈使句和價值判斷句,「快樂」很少。

e. 「快樂」做定語時,修飾的中心語比「高興」廣泛得多;

f. 「高興」做補語時,只能修飾短暫性動詞;「快樂」做補語時,能修飾短暫性和持久性動詞。

4.3 解釋:

最後我們要根據詞彙語意表達架構對上述規律的對比提出系統性解釋。

a. 「快樂」用來表達持久的狀態,其事件類型是均質狀態,事件結構是「——」,「高興」可以用來表示狀態的改變,其事件類型是啓始狀態,事件結構是「 •——」。

b. 當一個事件用作名物化動詞時,它的指涉功能被強調,而整個事件會被當作一個單位看待,在此情況下,表示均質持久狀態的動詞比較適用於此,所以「快樂」的名物化傾向高。

c. 因爲所接句賓語是指使情緒產生變化的事因,例如「他很高興我來了」中「我來了」導致「他高興」,所以自然只有表示狀態改變的「高興」可以後接事因。

d. 時態標誌「了」可以用來表示事件的完成或新狀態的產生。當「了」和情緒動詞搭配時,是指進入新的情緒狀態,例如「你一逗他,他就高興了」。因此自然是表示狀態改

變的「高興」能夠更常搭配「了」。

e. 當我們祝福別人時，自然是祝福別人能夠長久擁有該祝福，所以自然會挑選僅能表示持久狀態的「快樂」用於祈願句中。

f. 當我們說「你別高興」或「不值得高興」時，是在說服別人改變他的情緒狀態，所以能表示狀態改變的「高興」自然比較適合用於這樣的句型。

g. 當一個動詞用作名詞修飾語時，多半是用來指涉其持久的性質，所以僅能表示持久狀態的「快樂」用做定語的狀況會比「高興」來得多，所能修飾的中心語範圍也會比較大。

h. 因爲「高興」指涉狀態的改變，所以當「高興」用作補語時，我們傾向只用來修飾短暫性的動詞。唯有表示持久狀態的「快樂」才能修飾持續性動詞，像是「過得快樂、活得快樂」。

5. 結論

在本文中，我們粗略地說明了詞庫小組所提出的動詞詞彙語意的理論架構和分析方法。截至目前爲止，我們所作出的研究離理想還有一大段距離。雖然本研究已進行了兩年，對個別動詞的探討也相當深入，但是由於研究的動詞數量還太少、涵蓋範圍和類型還太狹窄，因此我們所找到的語意屬性還不足以充分地表達動詞的語意概念，更無法藉語意屬性來全面掌握語言規律。此外，語意屬性之間的共容、互斥及邏輯推導關係也還有待通盤的研究。我們誠心期盼各方先進的批評指教，也期待這套理論架構能日益完整，確實有效地反映現代漢語的語言現象。

參考書目：

詞庫小組，1993，*中文詞類分析*，技術報告 93-05，中文詞知識庫小組，中央研究院資訊科學研究所、中央研究院歷史語言研究所。

Chang, Li-li, Keh-jiann Chen and Chu-Ren Huang. 1999. “Alternation Across Semantic Fields: A Study of Mandarin Verbs of Emotion.” Proceedings of the 13th Pacific Asia Conference on Language, Information and Computation, pp39-50, Taipei.

Chief, Lian-Cheng, Chu-Ren Huang, Keh-Jiann Chen, Mei-Chih Tsai, and Li-li Chang. 1998. “What Can Near Synonyms Tell Us.” In Proceedings of the 1998 International Lexical-Functional Grammar Conference. Brisbane, Australia.

CKIP. 1995. *A Description to the Sinica Corpus.* Technical Report 95-02. Academia Sinica. Taipei.

Grandy, Richard E. 1992. “Semantic Fields, Prototypes, and the Lexicon.” In Lehrer and Kittay Eds. *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization.* Pp103-122. Hillsdale: Lawrence Erlbaum.

Huang, Chu-Ren. 1998. Classifying Event Structure Attributes: A Verbal Semantic Perspective from Chinese. Invited paper. Chinese Workshop. The 1998 International Lexical-Functional Grammar Conference. Brisbane, Australia.

Huang, Chu-Ren and Kathleen Ahrens. 1999. "The Module-Attribute Representation of Verbal Semantics." In Working Papers on Chinese Verbal Semantics, Vol. I. Corpus Research Group, Institute of Linguistics, Academia Sinica, and Chinese Knowledge Processing Group, Institute of Information Science, Academia Sinica. Taipei.

Huang, Chu-Ren and Mei-Chih Tsai. 1997. From Near Synonyms to Event Structure: Corpus-based Studies of Mandarin Verbal Semantics. Mini-Conference on Lexical Semantics. November. National Chung-cheng University.

Huang, Chu-Ren, Mei-chun Liu, and Mei-chih Tsai. 1998. "From Lexical Meaning to Event Structure Attributes: Across Semantic Classes of Mandarin Verbs." The 6th International Conference on Chinese Linguistics/The 10th North American Conference on Chinese Linguistics. June 26-28. Stanford.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.

Li, Charles & Sandra Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. California: University of California Press.

Liu, Mei-chun. 1997. "Lexical Meaning and Discourse Patterning – the three Mandarin cases of 'build'." Paper presented at the Third Conference on Conceptual Structure, Discourse, and Language. Boulder, Colorado.

Liu Mei-chun, Chu-Ren Huang, and Charles C.L. Lee. 1998. "When Endpoint Meets Endpoint: A Corpus-based Semantic Study of Throwing Verbs." The 6th International Conference on Chinese Linguistics/The 10th North American Conference on Chinese Linguistics. June 26-28. Stanford.

Liu, Mei-Chun, Chu-Ren Huang and Ching-Yi Lee. 1999. "Lexical Information and Beyond: Constructional Inferences in Semantic Representation." Proceedings of the 13th Pacific Asia Conference on Language, Information and Computation, pp27-38, Taipei.

Pustejovsky, James, S. Bergler, and P. Anick. 1993. "Lexical Semantic Techniques for Corpus Analysis." Computational Linguistics, 19.2, pp331-358.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge: MIT Press.

Smith, C. 1991. *The Parameter of Aspect*. Dordrecht: Kluwer.

Tenny, C. 1992. "The Aspectual Interface Hypothesis." In I. Sag and A. Szabolcsi Eds. *Lexical Matte*r. Standford: CSLI.

Tsai, Mei-chi, Chu-Ren Huang, Keh-jiann Chen, and Kathleen Ahrens. 1998. "Towards a Representation of Verbal Semantic: An Approach Based on Near-Synonyms." International Journal of Computational Linguistics & Chinese Language Processing, pp62-74.

Vendler, Zeno 1957. "Verbs and Times." Philosophical Review 56, 143-160. Also in Z. Vendler. 1967. *Linguistics in Philosophy.* Cornell University Press. Ithaca, NY, 97-121.

Computational Linguistics and Chinese Language Processing
Vol. 5, No. 1, February 2000, pp. 19-46

© Computational Linguistics Society of R.O.C.

The Module-Attribute Representation of Verbal Semantics: From Semantics to Argument Structure

Chu-Ren Huang*, Kathleen Ahrens+, Li-Li Chang*, Keh-Jiann Chen*, Mei-Chun Liu ** and Mei-Chi Tsai++

Abstract

In this paper, we set forth a theory of lexical knowledge. We propose two types of modules: event structure modules and role modules, as well as two sets of attributes: event-internal attributes and role-internal attributes, which are linked to the event structure module and role module, respectively. These module-attribute semantic representations have associated grammatical consequences. Our data is drawn from a comprehensive corpus-based study of Mandarin Chinese verbal semantics, and four particular case studies are presented.

1. Background

Generative theories have long assumed that lexical semantics are encoded on each and every lexical entry, and hence represent idiosyncracies of each lexical item. This assumption, however, goes back much farther than generative theories. For example, Levin [1993] pointed out that Bloomfield wrote in 1933: "The lexicon is really an appendix of the language, a list of basic irregularities" [1993: 274]. As a consequence of this assumption, lexical semantics was not intensively studied within the generative framework because it was not expected to offer any interesting generalizations.

The notable exceptions, other than the short period of intense work on the generative semantics paradigm, were studies by Jackendoff [1983] and Wierzbicka [1985]. However, as grammatical theories became more and more lexicon-driven, more in-depth theoretical and empirical studies on the lexicon were carried out, and the above assumption was no longer valid. Levin [1993] in particular sounded the call for in-depth work on a theory of lexical knowledge. She writes that a theory of lexical knowledge:

*Academia Sinica, E-mail: Chu-Ren Huang: hschuren@ccvax.sinica.edu.tw
Li-Li Chang: lili@iis.sinica.edu.tw
Keh-Jiann Chen: kchen@iis.sinica.edu.tw
+National Taiwan University, E-mail: ahrens@ms.cc.ntu.edu.tw
**National Chiaotung University, E-mail: mliu@cc.nctu.edu.tw
++National Sun Yet-Sen University, E-mail: tsmei@mail.nsysu.edu.tw

> ...must provide linguistically motivated lexical entries for verbs which incorporate a representation of verb meaning and which allow the meanings of verbs to be properly associated with the syntactic expressions of their arguments (p.1).

This goal of a theory of lexical knowledge has not yet been attained, for reasons we will discuss in Section 2 below. It is, however, a worthy goal, and is in fact, the goal of this paper - to provide a theory of lexical knowledge based on lexical semantic features that are associated with a verb and predict their associated syntactic expressions.

In what follows, we will first look at why Levin's [1993] proposed use of diathesis alternations to ferret out meaning has fallen short of its goals. We will then propose a different way of looking for relevant syntactic behavior in Section 2. We will next present two underlying assumptions of our theory of lexical knowledge in Section 3, and then present the theory in Section 4. We will give four case studies in which we apply our theory in Section 5. We will summarize our theory in Section 6.

2. Verbal Semantics

Levin (1993) assumed that:

> "....the behavior of a verb, parti cularly with respect to the expression and interpretation of its arguments, is to a large extent, determined by its meaning. Thus, verb behavior can be used effectively to probe for linguistically relevant pertinent aspects of verb meaning" (p.1).

We agree with this assumption. But as we will discuss below, we look at different aspects of verb behavior from Levin [1993].

Levin [1993] concentrated on the range of possible synactic alternations of a single verb (or a single verb class) and extracted semantic information from syntactic behavior. For example, she pointed out that *break* verbs (verbs such as *break, crack, rip, shatter, snap* etc.) all can appear in the middle alternation but cannot appear in the conative or body-part ascension alternation while cut verbs (verbs such as *cut, hack, saw, scratch, slash* etc.) can appear in all three alternations [1993: 7]. After comparing these two verb groups with two others, *touch* and *hit* (and their respective alternations), she concluded that *break* is a pure change of state verb, and that *cut* is "a verb of causing a change of state by moving something into contact with the entity that changes state" (p. 10). The syntactic differences they display, she argued, are a direct result of their semantic differences.

However, there are two reasons why we have not followed Levin in examining the relationship between a verb alternation and its associated semantics. First, although the work done by Levin [1993] in this area is impressive (having determined 50 different types of alternations and over 125 different semantic classes of verbs), the sheer number of possible permutations of alternations makes analysis difficult. In addition, when comparing verbs of very different meanings, as in the *cut* and *break* example given above, it becomes hard to determine the relevant area of semantic difference. For example, in order to attain the generalization concerning *cut* and *break*, Levin had to look at two other verbs (*touch* and *hit*) and their respective diathesis alternations, as well as look at other verbs that could fit into those alternations [cf. 1993, pp. 5-8]. If she had picked different verbs from *touch* and *hit* or different diathesis alternations from the three that she did, she might not have been able to come up with a generalization at all. These factors may have contributed to the fact that there is currently no unified theory of lexical knowledge based on verb alternations because the scope of the undertaking is so vast.

Second, our research group [e.g. Liu 1997] tried a pure-alternation based approach and found that it was not adequate for defining Mandarin verb classes. There are several possible reasons for this. The first is that diathesis alternations have not been extensively studied in Mandarin, unlike English, where as Levin notes, several important studies were done on the verbs *cut, hit, break and touch* prior to her own work. The second reason has to do with the vastness of the enterprise as we mentioned above. How does one decide which verbs to compare? How does one decide which alternations are relevant? The third possibility is that Mandarin differs from English in such a way as to make alternations a non-viable option for prying into a verb's relevant semantics. Liu [1997] argued that that verb alternations are not suitable for extracting semantic generalizations from syntactic behavior in Mandarin Chinese because argument placement is relatively flexible.

If we agree, then, that syntactic behavior can shed light on the relevant semantics of a verb, and that for languages like Mandarin (if not for all other languages), diathesis alternations, while originally promising, can not move us towards a unified theory of lexical knowledge, then what other type of behavior is available?

We will concentrate on delimiting the lexical semantic distinctions between near-synonym pairs that differ slightly in both syntactic behavior and in semantics. Sometimes a semantic difference is apparent at first glance as in the case of *fang4* (put) and *bai3* (set), and sometimes it is not clear and only becomes apparent after we compare the syntactic differences, as in the case of *kuaile* 'happy' and *gaoxing* 'glad'. (We will discuss both examples further in Section 5.)

However, even in cases where there is a difference in meaning, what we are looking for is the relevant differences in both syntax and semantics; that is, along what semantic lines do these two words differ, and how is this difference related to their synactic behavior (and vice versa)?

How do we determine these syntactic and semantic differences? The answer to this question was explained in much more detail by Tsai *et al.* [1998] and Liu *et al.* [1997]. But we will give a very brief sketch here. First, we examined these near synonym pairs by first combing the Sinica Corpus for all relevant examples of the words in question. These examples were then categorized according to their syntactic functions. Third, each instance was classified according to its argument structure type. Fourth, the aspectual type associated with each verb was determined, and fifth, the sentential type for each verb was also determined. We found that near synonyms usually have several cases of complementary distribution of synactic functions. It is often these cases of complementary distribution that allow us to formulate a hypothesis concerning the relevant nature of their semantic differences.

3. Assumptions

We share the following assumptions with some of the recent works on lexical semantic theories. The first assumption is that lexical semantic contents are mapped to the morphosyntactic level and can be used to predict grammatical behavior [e.g. Dowty 1991, Levin 1993, Goldberg 1995]. What is crucial behind this assumption is that a mapping must be rule-governed and regular by definition. Hence, the assumption entails the idea that lexical semantic generalizations are not only worth studying, but that they can also be verified by means of grammatical realizations.

The second assumption is that lexical semantics exists on the grammatical level that mediates conceptual structures with grammatical representations [e.g. Bresnan and Kanerva 1989, Zaenan 1993, Pustejovsky 1995]. In other words, lexical semantics not only can be empirically verified through grammatical predictions, but can also be justified by means of conceptual arguments.

In fact, we will take the second assumption further and make it our premise that lexical semantic representations are the grammaticalization of conceptual information. Based on the above assumptions, we propose that an adequate theory of verbal semantics must have the three following properties: direct representation, conceptual motivation, and representational clues.

First, lexical semantic information must be represented in a way that can be linked directly to grammatical structures. We assume that such a representation in verbal semantics must be based on event structure. Second, lexical semantic information must have conceptual motivation. This justifies the inclusion of such information as qualia structure in lexical semantics [Pustejovsky 1995]. Third, all lexical semantic attributes must be attested by representational clues: either collocating structure, selectional constraints, or distributional patterns. This last premise is especially important because it restricts the type of evidence that may be brought to bear on the question of whether something shares a particular attribute or not, and it limits the possibility of ad-hoc explanations. That is, it strongly focuses analyses in verbal semantics on corpus-based approaches since representational clues are best extracted from corpora.

In particular, in our work on lexical semantics, we have concentrated on exploring the semantic and syntactic differences between near synonyms in the Sinica Corpus. We have examined near synonyms in order to extract the contrasts that dictate their semantic and associated syntactic behaviors [Chief *et al.* 2000, Huang *et al.* 1999, Liu *et al.* 2000, and Tsai *et al.* 1998]. Conceptually, each group of near synonyms that we study forms a contrast set that is a constituent of a semantic field [Grandy 1992]. Our goal is to locate the linguistic relation that defines the contrast. In particular, we look for the semantic relation that can predict the difference in grammatical behaviors of the set. It is our strong hypothesis that syntactic variations, including Levin's [1993] alternations and morpho-semantic variations, can be predicted by logical implicatures of the semantic attributes encoded on the event structure of each verb.

4. Model-Attribute Representation

In the Module-Attribute Representation of Verbal Semantics (MARVS), lexical knowledge is classified into two types: structural information is represented by means of the composition of atomic modules while content information is represented by means of attributes attached to these modules.

First, the overall shape of event structure is defined by the composition of five *Event Modules*. The roles that participate in the event are represented in the *Role Module*s. The semantic attributes pertaining to the whole event are called the *Event-Internal Attributes* and are attached to the event modules. The semantic attributes pertaining to each role are termed *Role-Internal Attributes* and are attached to the appropriate role within the role

module. A sketch of the representation is given in Figure 1. [1]

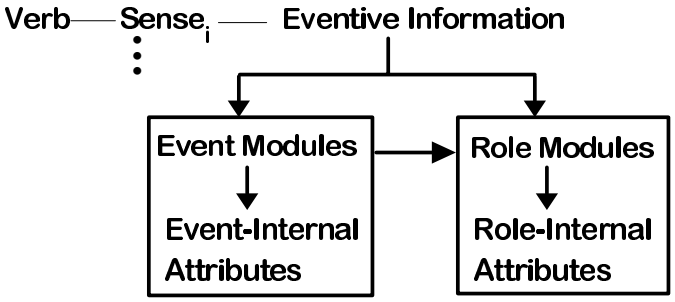

Figure 1 *Module-Attribute Representation*

It is important to note that the eventive information is attached to the sense of a verb. Verbs with different senses will have different eventive information. [2]

The second important hypothesis of this proposal is that the event representation of a verb is the sum of all attested event realizations of a particular verb. In other words, it is possible that a complex lexical event representation is never fully instantiated,

---

[1] In a prior version of the theory, there were only attributes: aspectual attributes, event-internal (inherent) attributes, role attributes, and role-internal attributes. The original definition is given below [Huang and Tsai 1997, Huang 1998].

1) Aspectual attributes: attributes pertaining to the composition of the event(s), such as Telicity, Homogeneity, etc.

2) Event-internal attributes: attributes referring to the semantics of the event itself, such as Control, Effect, etc.

3) Role attributes: attributes referring to the focussed roles of the event, such as Agent, Theme, Instrument, Manner, etc.

4) Role-Internal attributes: attributes referring to the internal semantics of a particular focused role (of the event), such as sentience, volition, affectedness, etc.

However, as the theory progressed, the aspectual attributes became more and more well-defined, and five basic event types were found to occur and reoccur when discussing the semantic differences among verbs. These 'atomic' event structures were then found to combine in certain ways, and as a result of their ability to combine, these aspectual attributes grew or graduated to a 'module' level.

Then the event-internal attributes were surmised to be associated with the event structure of the verb and so, were linked to this module. The close relationship between the role-internal attributes and the role attributes was also noted, and the importance of participant roles in other theories, such as Construction Grammar [Goldberg 1995], led us to postulate a role module. It was also noted that these roles may also be considered atomic roles, which then may combine to create a role module, similar to the way atomic event structures combine, when necessary, to create the event module. We feel, however, that the inventory of role modules still needs to be made more comprehensive and precise, and we will look at this in future studies.

[2] Ahrens *et al.* [1998] gives a working definition and criteria for distinguishing between senses of nouns.

although each component is linguistically attested. This hypothesis is motivated by our desire to maintain the theoretical elegance of one-to-one mapping between verbal sense and event representations. It is also conceptually motivated by the fact that the same verb form is often used in natural languages to refer to different aspects of an extended event. For instance, the activity of 'sitting down' and the state of 'be sitting' share the same verb form. Similarly, in Chinese at least, the activity or 'putting on' and the state of 'wearing' some piece of clothing share the same verb form. Since they have different (logical) event structures, previous theories have had to treat them as homophones. However, the conceptual tie is so salient that we feel it is counterintuitive to assign them to two different senses. We postulate that there will be conceptual/cognitive motivations to encode such complex event structures with one representation. Hence, the contrastive event realization can be understood as different (partial) realizations of the same complex event under a particular event focus, and not as two senses.

The third crucial premise in this representation is that the event modules constitute the basic frame of verbal semantics. By establishing a the two-way distinction between modules and attributes, we assume that modules refer to pre-packaged semantic information while the attached attributes give more a detailed description. The two types of modules also represent the two basic atomic terms in formal semantics: event and individuals. However, individuals are understood in the context (i.e. events) in which they participate. Figure 1 shows clearly that role modules are attached to the event modules. There is strong motivation for such a representation: first, role modules represent the participants of the event; thus, they cannot stand outside of the event representation; second, the participating roles can be partially predicted by the event types; finally, hierarchical constraints can be entailed, as will be discussed later (Section 4.3).

In what follows, we will first discuss event modules, and then the event-internal Attributes that are associated with the event modules (Section 4.1). Then, we will discuss the role modules and the role-internal attributes that are associated with these modules (Section 4.2).

4.1 Event Modules

A central issue in lexical semantics, especially verbal semantics, is the representation of events [e.g. Jackendoff 1983 and Pustejovsky 1991]. A tradition shared by philosophical and linguistic semantics, as well as the cognitive sciences, is that there are only two basic types of entities: events and individuals. Hence, a language must conceptually describe both events and individuals. Individuals are prototypically denoted by the referential properties of nominals while events are denoted by verbs. Thus, an adequate theory of verbal semantics must include a theory of event structures. Of course, all semantic

theories must also account for type-shifting and semantic coercions, such as the telic and agentive structures in Pustejovsky's [1995] nominal semantics.

In this section, we will concentrate on the basic building blocks of our verbal semantic theory. In particular, we will propose a theory in which event structures can be created from a small set of event modules and the backbone of verbal semantics can be taken to be combinations of these event modules. This account is crucially different from the autonomous view of event structure [e.g. Vendler 1967] or the attribute-value view [Jackendoff 1983]. It shares some assumptions with Smith [1991], such as the viewpoint focus interpretation of aspectual facts. However, our modules and rules of combination are different.

4.1.1 An Inventory of Event Modules

Event modules are the building blocks of linguistic event structures. They can be used in combination or alone. When used alone, they are atomic logical event structures. We list five atomic event structures below, along with their associated symbols. A brief explanation follows each event structure.

Atomic Event Structures

(1) • Boundary (includes a Complete Event)

Boundary is an event module that can be identified by means of a temporal point and must be regarded as a whole.

(2) / Punctuality

Punctuality is an event module that represents an single occurrence of an activity that cannot be measured based on duration.

(3) ///// Process

Process is an event module that represents an activity that has a time course, i.e., that can be measured in terms of its temporal duration.

(4) ____ State

State is a homogeneous event module in which the concept of temporal duration is irrelevant; i.e., it is neither punctual nor does it have a time course.

(5) ^^^^^ Stage

Stage is an event module consisting of iterative sub-events.

In sum, we postulate that these five atomic event structures are the only building blocks necessary to capture the range of complex linguistic event structures.

4.1.2 Tests for Event Modules

Since event modules are logically and conceptually primary units, each event module has logical entailments that can be attested based on their grammatical behavior and/or their interpretation. A partial list of their verifiable entailments follows.

First, only boundaries (including stand-alone complete events) can be identified with a temporal point, such as in (6).

(6) *Complete event vs. other events*

a. *Sheme shihou V (le)*
When V ASP
b. *Sheme shihou kaihui (le)*?
When meeting
'When does the meeting (start)?'
c. **Sheme shihou dasuan (le)*?
When plan

Second, since process encodes a time course, a durational phrase naturally measures the length of the time course and can distinguish between process events and boundary/complete events, as (7) and (8) show.

(7) *Process vs. Complete Event/Boundary*

V le Duration
V ASP Duration

(8) a. *(*yizhi si)*
always die
b. *(yizhi pao)*
always run
'(She has been) running continuously'
c. *(*yizhi si) si le san ge xiaoshi*
always die die ASP three CL hours
'(He's) been dead for three hours.'
d. *(yizhi pao) pao le san ge xiaoshi*
always run run ASP three CL hours
'(He has kept on) running for three hours.'

Since complete and boundary events both have a delimiting temporal point (but contain no time course), the durational phrase can only be interpreted as the distance between a reference point in time and that delimiting temporal time (i.e. the death time in (8)a&c). On the other hand, the durational phrase will be interpreted as the time course of a process ((8)b&d). The contrast in interpretation can also be demonstrated by the continuous adjunct *yi1zhi2* 'always, keep on V-ing', which cannot co-occur with complete/boundary events (8).

(9) Stage vs. Activity

a. *ta pao-le san ge zhongtou*
s/he run-LE three CLS hour
'S/He has been running for three hours.'

b. **hua diaoxie-le san tian*
flower wither-LE three day
comp.'These flowers have been withering (on the vine) for the past three days.'

c. **shiqi miman-le san tian*
humidity permeate three day
'*The humid air has been permeating for three days.'

Third, a stage in MARVS refers to an event, which is necessarily understood as the sum of iterative sub-events. In other words, a stage as an event module means that it contains sub-events that can be distinguished conceptually, but can not be represented lexically. In contrast, an activity is holistic and can only be sub-divided with event-external measurements, such as time. Thus, although both event modules can be viewed as taking up temporal duration and can be used with the durative aspect *(zheng4)*zai4 to refer to overlapping time, only an activity can be temporally measured (9a). This is because stage refers to equilibrium (e.g. *mi2man4* 'to permeate') or a constant and irreversible tendency towards a state (*diao1xie4* '(plants) to wither') involving dynamic or iterating sub-events. In other words, this event describes homogeneity achieved through dynamic iteration. Thus, it is predicted to exhibit some activity-like behavior and some state-like behavior, but to also differ crucially from either event types.

4.1.3 Typology of Lexical Event Representations

In this section, we present three different types of event structures that are encoded on Chinese verbs: atomic, bounded, and composite events that are made up of one or more of the five atomic event structures. Note that we propose and follow the strong hypothesis that each sense of a verb form encodes a unique eventive information representation. Hence, each meaning realization can focus on different elements of that encoded event information but cannot refer to a different event representation. This is the

One-Event-Representation-per-Sense hypothesis. Lexical event representations can be classified based on the complexity of their component event modules into three types: Nucleous, Bounded, and Composite event representation.

In this theory, event structure modules are events that cannot be further divided. Our claim is that human linguistic representation of events does not necessarily correspond to these logical and atomic events. We assume that conceptual and cognitive motivations require that certain event module combinations be perceived as a whole, and thus be mentally and linguistically represented by a single event structure with compositional modules. In other words, we are proposing a non-homomorphism between logical event structure and (human) linguistic event structure. We will focus our discussion on the linguistic event structures since they are conceptually more interesting.

4.1.3.1 Atomic Event Representation

The verbs listed below in (10) have stand-alone event modules.

(10) a. *Completion (achievement)*

- *si3* 'to die', *po4* 'to break'

b. *Punctuality*
/ *da3suan4* 'to plan to'

c. *Process*
////// *zou3* 'walk', *pao3* 'run'

d. *Homogeneous State*
____ *kuai4le4* 'to be happy', *pi2juan4* 'to be tired'

We have not found any examples yet of the stage event module standing alone in a verb in Mandarin. However, our hypothesis is that this list of atomic events will not grow past the five listed in Section 4.1.1 for any language.

4.1.3.2 Bounded Event Representations

Bounded events have one atomic event and must be bounded at at least one end (but may be bounded at both ends). The verbs listed in (11) encode both a boundary and an associated non-instantaneous event.

(11) a. *Inchoative Process*

- ///// *xia4yu3* 'to rain', *kai1hui4* 'to convene a meeting'

b. *Bounded process*

• ///// • *gai4* 'to build'

c. *Resultative*

/ • *da3si3* 'to hit and kill'

d. *Completive Punctuality*

' *chu4fa1* 'set forth', *bi4ye* 'graduate', *li2kai1* 'go away'

e. *Inchoative State* (Effect State)

• ____ *gao1xing4* 'to be glad'

f. *Inchoative Stage*

• ^^^^ *shang4sheng1* 'to rise'

g. *Bounded Stage*

• ^^^^ • *diao1xie4* '(flowers) to wither'

We think we have exhaused the combinations of boundary events with the list above for Mandarin Chinese. Other languages may have other combinations.

4.1.3.3 Composite Event Representations

Composite events involve more than one atomic event (and may or may not be bounded). Two examples are given in (12). We expect this partial list of complex events to grow with further study of both Mandarin verbs and verbs in other languages.

(12) a. *Completive Resultative*

/_____ *zuo4* 'to sit', *tang3* 'to lie [down]', *bao1wei2* 'to surround'

b. *Dual Process-State*

• ///// • ___ *chuan1* 'to wear', *dai* 'to wear'

Let us take a closer look at the verb *zuo4*. In (13a), the focus is on punctuality while in (13b), the focus is on state. In (13c), the focus is on the length/duration of state as delimited by the punctual event and a reference point. In (13d), the focus is on the manner of the state, with an implied (controllable) punctual event that could change the state.

(13) a. *zuo*

sit

'Sit [down]!, Be seated!'

b. *ta zuo qianmian*
s/he sit front
'S/He is seated in the front.'
c. *ta zuo le san ge zhongtou*
s/he sit ASP three CLASS hour
'S/He has been sitting for three hours.'
d. *haohao zuo*
well sit
'Sit straight!'

4.1.4 Event-internal Attributes

In our module-attribute representation, Event-internal attributes are linked to the event structure modules (when necessary). Event-internal attributes refer to the semantics of the event itself, such as [control], [effect] etc. Example (14), for example, shows that the two verbs *gaoxing* and *kuaile* differ in terms of the attribute of control [see Tsai *et al.* 1998 & 1999 for more details about this relationship].

(14) *[control]*
*bie gaoxing/*bie kuaile*
NEG happy /NEG happy
'Don't be happy.'

4.2 Role Modules

Role modules contain the focussed roles of an event and typically include all required (i.e., thematic) arguments but can also include optional arguments and adjuncts. The roles that we have considered are the following: Agent, Cause, Causer, Comparison, Experiencer, Goal, Instrument, Incremental Theme, Location, Locus, Manner, Range, Recipient, Source, Target, Theme, etc. We will illustrate how this module works with an optional argument. In example (15a), the focus is on an incremental theme; therefore, the measure phrase describes the resulting number of cuts. However, in (15b), there is no such focus; therefore, the measure attached to the cognate object describes the frequency of the activity

(15)a. *ta ba shoubi gele shijidao yi shi juexin*
s/he BA arm GE-PERF ten-plus-knife so show resolution
'S/He made more than ten cuts on his/her arm to show his/her resolution.'

b. *zai qindi shenshang kanle wushiliu dao*
at love-foe body-top KAN-PERF 56 knife
'[The person] hacked his/her rival in love affair 56 times.'

4.2.1 Role-Internal Attributes

These attributes refer to the internal semantics of a particular focused role (of the event), such as [sentience], [volition], [affectedness], [design] etc.

In (16), we give an example of the role internal attribute of [design]; when attached to the role Loc, it implies that the role can be specified based on orientation.

(16) *Role-Internal Attribute Loc [design]*
a. *na ge taishiyi bai dongbian/zhao dong bai*
that CLS easy-chair set east-side/towards east set
'Put that easychair so that it faces east.'
b. **na ge taishiyi fang dongbian/zhao dong fang*
that CLS easy-chair put east-side/towards east put

Some readers might wonder what the difference is between role-internal attributes and the selectional restrictions placed on lexical items that previous versions of transformation theories postulated. This issue was addressed by Huang *et al.* [1999], who showed that alternative interpretations in a context can be accounted for by means of role-internal attributes but not selectional restrictions. Role-internal attributes interact with (context-induced) meaning to determine the appropriate reading while selectional restrictions are projected from a fixed lexical entry. From an informational point of view, role-internal attributes are information-bearing and declarative (i.e., directly specify knowledge about the semantics of that role). On the other hand, selectional restrictions are passive grammar-checking mechanisms.

4.3 Hierarchial Constraints

All conditions being equal, a higher-level module (i.e., event structure module) or attribute (i.e., event-internal attribute) is preferred for the sake of generality and greater explanatory power. For instance, [control] will be preferred over [volition] if both offer an equally adequate account since [control] is an event-internal attribute belonging to the whole event; on the other hand, [volition] is a role-internal attribute describing a participant of an event. If volition can be predicted by a [control] event-internal attribute (and it usually can), then there is no need to list volition again in the role-internal attribute. The [control] event-internal attribute will predict volition through the semantic relationship of implicature. However, if hypothetically a verb has the attribute [control] but has a

non-volitional subject, then there is a place in the role-internal attribute to indicate that fact, and the usual implicative relationship between [control] and [volition] will be cancelled.

In addition, when a set of near synonyms includes a covering term of a field, then the grammatical contrast is weakened to a marked/unmarked situation. That is, the covering term, as a unmarked element, can substitute for its near syonoym in many cases. It simply has a wider range than its near syonoym. In this case, the lack of clear-cut contrasts does not affect the legitimacy of a defining relation. Another near synonym forming a contrast set should be substituted to verify the claim. For instance, not all predicted grammatical contrasts demonstrate themselves between *ge1* 'to slice' and *qie1* 'to cut [covering term]'. But when *ge1* is contrasted with *ci4* 'to stab', the proposed contrasting relation of [effect] is clearly evident.

5. Research Methodology and Case Studies

In this section, we will show that cross-category generalizations can be captured by delimiting the lexical semantic distinctions between near-synonym pairs. We will illustrate, with four case studies, the correlation between lexical semantic specifications and event-structure attributes.

5.1 Research Methodology

Our research methodology studies on Chinese lexical semantics have produced an approach that is different from traditional approaches. First, it is corpus-based. In other words, we emphasize observations and generalizations based on qualitative and quantitative studies of actual language use. Second, we target near synonym pairs as our initial focus. In targeting near synonyms, we in effect restrict our scope to a semantic field for each study. In addition, near synonym pairs are often (minimal) contrast sets in the theory of semantic fields [Grandy 1992]. Through a comparative study on a contrast set as well as its grammatical consequences, we will be able to identify the critical semantic element(s) that distinguish contrast sets. Since contrast sets are lexical items that differ minimally semantically, the semantic elements identified should be the primary semantic elements that need to be represented in a lexical semantic theory.

Our research methodology involves three consecutive steps: 1. Make generalizations about grammatical relation contrasts based on distributional differences observed and/or extracted from corpora. 2. Deduce event structure elements that would predict the above generalizations (by examining the semantic implicatures of such elements). 3. Verify these elements by applying them to new syntactic/semantic frames.

This last step is the only one that uses linguistic intuition to generate ungrammatical sentences to test our hypothesis. During all these three steps, the following corpus-based distributional information is our primary data:

1. the syntactic functions that a verb can play,
2. the argument number and types that a verb can take,
3. the aspectual types that a verb can associate with,
4. the sentential types that a verb can occur in, and
5. the types of arguments that a verb integrates with in compounds.

How the above information can be used in argumentation will be illustrated in the following subsections.

5.2 Case Study 1: *bai3* vs. *fang4* - Event Structure Focus

Both *bai3* and *fang4* are verbs of putting, and they seem to be synonymous and exchangeable in certain contexts.

(17) a. *bai/fang qizi*
set/put chess-piece
'to put down chess pieces'

b. *bai/fang yizi*
set/put chair
'to put down chair(s)'

However, there are distributional differences between *bai3* and *fang4*: *bai3* can co-occur with progressive *zheng4zai4* to describe a process, but *fang4* cannot (18); *bai3* can take a resultant object, but *fang4* cannot (19); and *bai3* can be modified with an orientational adjunct, but *fang4* cannot (20).

(18) a. *ta zhengzai bai shu*
s/he DUR set book
'S/He is putting down the books now.'

b. *?ta zhengzai fang shu*

(19) a. *mama baichu yi zuo cai*
mother set-out one table dish
'Mother (cooked) and set a tableful of dishes.'

b. **mama fangchu yi zuo cai*
mother put-out one table dish

(20) a. *na ge taishiyi bai dongbian/chao dong bai*
that CLS easy-chair set east-side/towards east set
'Put that easychair so that it faces east.'

b. **na ge taishiyi fang dongbian/chao dong fang*
that CLS easy-chair put east-side/towards east put

The above three contrasts, attested by corpus data, point to a crucial difference between the meanings of *bai3* and *fang4*, which is that *bai3* entails that the act of putting follows a certain plan, and therefore that the orientation of the placed object can be specified while only location can be specified for *fang4*. In addition, since the plan which the putting action follows entails a resultant state to be attained, *bai3* can take a resultant object while *fang4* has no such entailment and cannot take such an object. Third, following a plan implies that *bai3* involves a process that can be broken down into constituent steps while *fang4* is a simple activity. Thus, only *bai3* can be attached with a progressive aspect referring to internal steps being carried out.

Based on the above contrasts and generalizations, we propose that the lexical semantic attribute that differentiates *bai3* and *fang4* is the role internal attribute of [design]. By [design], we mean a plan that the actor is cognizant of when s/he carries it out. This feature not only affects the interpretation of the two verbs. It also entails that only *bai3* can take an incremental theme as an object (the resultant object in (19)) as well as the aspectual and adjunct constraints described above.

One immediate implication of this account is that all idioms or compounds involving a [design] scheme can only be composed using the verb *bai3*, not *fang4*. This is confirmed by the following idioms/compounds involving setting up a scheme or taking on a certain (affected) attitude:

(21)a. *bai jiazi*
set-shelf
'to put on airs'

b. *bai SOMEBODY yi dao*
set - SB - one - CLS
'to set someone up once'

c. *bai ditan*
set-ground-spread
'to set up a street vending position (by spreading a piece of cloth on the ground)'

d. *bai kuo*
set-rich
'to show off one's wealth'

The above idiom/compound evidence not only offers additional support for the [design] attribute, but also strongly suggests the position where this attribute should be attached. *A priori*, the role internal aspect of [design] attribute describes the resultant location. However, since it affects the collocation of aspects, there are also motivations for arguing that it is represented at a higher level. However, an account of the above data makes it necessary for the [design] attribute to be present at the locative object. It is the lexical semantic specification of [design] on the locative object that allows the above compounds and idioms to acquire the 'affected attitude' or 'planned scheme' meaning. To account for its interaction with an aspectual specification, our analysis leads us to propose that the locative object (together with the [design] attribute) receives an Event Structure Focus. Thus, even though the attribute is Role-Internal, it is also 'visible' and can interact at the event structure level. Our account can be shown in MARVS by the following diagram (with irrelevant parts omitted). Take note that the roles are listed within angled brackets while focused roles are indicated by boldface type. Unspecified attributes simply are not represented.

Diagram 1
MARVS for *bai3* and *fang4*

bai3 • ____ <Agent, Theme, **Location**>
|
[design]

fang4 • ____ <Agent, Theme, Location>

In conclusion, we want to point out that the [design] feature is not only useful for accounting for the lexical semantic differences between the members of the current pair, but it can also be applied to other pairs where the notion of a certain design is inherent in the verb. Two additional examples are *dui1* 'to pile' vs. *fang4* again and *hua4* 'to paint, to draw' vs. *tu2* 'to cover with paint, to doodle.'

5.3 Case study 2: *peng4* vs. *mo1* - Motional Path

The next pair of near synonyms *peng4* and *mo1* are verbs of touching. At first glance, they seem to differ mostly in the force used: *peng4* refers to all types of touching while *mo1* seems to be restricted to light touching with fingertips, such as caressing. However, there are additional grammatical contrasts that cannot be explained by this simple difference in degree of force.

First, it is observed that durative *-zhe* can only co-occur with *mo1*, not *peng4* (22). Similarly, only *mo1* can take a durational complement; *peng4* cannot (23).

(22)a. *xiaohai mo-zhe bizi*
child touch-DUR nose
'The child is touching his/her own nose.'
b. **xiaohai peng-zhe bizi*
child touch-DUR nose

(23)a. *Ta mo le bantian, (sheme ye mei mo dao)*
s/he touch PERF half-day what YE NEG touch reach
'S/He groped for a long time but did not touch anything.'
b. **Ta peng le bantian, (sheme ye mei peng dao)*
s/he touch PERF half-day what YE NEG touch reach

The two sets of contrasts suggest that *peng4* denotes an instantaneous activity, and that its motional path ends with a focus on one impact point while *mo1* denotes the activity of touching with a focus on either continuous contact or the motion towards touching. In other words, *mo1* has a time course while *peng4* does not. This generalization can be nicely captured by using two of the proposed event modules: Process for *mo1* and Punctuality for *peng4*. To account for the fact that durational phrases are interpreted as a temporal distance from the beginning point of the *mo1* activity (23a), its event structure will include a beginning boundary. Thus, *mo1* is of the Inchoative Process event type while *peng4* is of the Punctuality type.

Another important piece of semantic information that needs to be encoded is that both verbs involve a (motional) path. Following tradition in this field, path is not explicitly marked. Instead, its presence is implied by either goal or source roles (or both). In agreement with other spatio-temporal expansion of an event, we will treat path as an embedded sub-event. This will allow us to describe path and other spatio-temporal elaboration of an event by using established theories of event structures. In this particular case, *mo1* has a path that is underspecified while *peng4* has a path specification with a focus on its single endpoint. In addition to the above contrasts, this is attested to by the

fact that the goal of *peng4* is more definite and can occur as either an effective object or as the subject of a presentative sentence, as in (24a) and (24b), respectively.

(24) a. *ta (tou) peng-le san-ge bao*
s/he head bump-PERF three-CLS bumps
'S/He bumped three bumps (in the head).'
b. *chezi peng-le yi-ge da dong*
car bump-PERF one-CLS big hole
'There was a big hole in the car as a result of bumping (into something).'

In a formal representation, we will stipulate that the terminus of the Path of *peng4* (i.e., the Goal role) be definite (i.e. role-internal attribute). Thus, formally speaking, the pair of verbs contrast in their lexically specified event contours which are specified at both the event structure and the role-internal levels. The above account can be formally represented as follows:

Diagram 2
MARVS representation of *peng4* and *mo1*

Peng4 / <Agent, Goal>
|
[definite]

mo1 • //// <Agent, Goal>

Our account suggests that the event contour will be necessary to in account for contrasts of other verbs of contact, such as *ji3* 'to squeeze' vs. *ya1* 'to push down', and *an4* 'to press down' vs. *ya1* 'to push down'.

5.4 Case Study 3: *gai3* vs. *bian4/ ban1* vs. *yi2* - The Causative Alternation

The third contrast involves a pair of 'change' verbs: the 'change of state' verbs *gai3* 'to revise' and *bian4* 'to transform' as well as the 'change of position' verbs *ban1* 'to move (something)' and *yi2* '(something) moves'. This contrast is commonly seen with similar verbs in other languages, where theme (i.e. the entity that changes) occurs in the objective position with one set of verbs and in the subjective position with another set of verbs. This alternation is referred to in the literature [e.g., Levin 1993] as causative alternation. Typically, the theme occurs in the subjective position for *bian4* and *yi2* and in the objective position for *gai3* and *ban1*. For instance, 85% of all occurrences of *ban1* in Sinica Corpus have an explicit theme object while 80% of the occurrences of *yi2* have a single theme subject. Illustrative examples are given below.

(25) a. *laoshi gai le san pian zuowen*
teacher revise PERF three CLS writing
'The teacher corrected three writing assignments.'
b. *tianqi bian le*
weather transform PERF
'The weather changed.'

(26) a. *ban-chu liang zhang yizi*
move-out two CLS chair
'[someone] moved two chairs out.'
b. *shitou yi(dong) le*
stone move(move) PERF
'The stone moved.'

Since causative alternation has been thoroughly studied in the literature, we will follow previous works and characterize the contrast as directly involving event-structures. In other words, the causative verb will have a unique (complex) event-contour represented as being composed of two event modules linked by a causative transition. Without such a specification, the non-causative counterpart will be a simple change of state event. Such a specification will predict all observed contrasts of the two pairs of synonyms. Since causation entails a volitional causer, the 'causative' verbs are [+control] and can occur in imperative construction (27). Since simple change of state verbs focus on the transition of changing, they are achievement verbs that do not take durational complements (28). Lastly, since causative verbs are willed by the causer, the direction of change is implied to be for the better (e.g. to correct), but simple 'change' verbs have no such implication (29).

(27)a. *kuaidian ban*
hurry-up move
'Move [the things] faster.'
b. *manman gai*
slowly revise
'Revise/correct slowly (and carefully).'

(28)a. **tianqi bian le san xiaoshi [with the intended interpretation of duration of activity]*
weather transform PERF three hours
b. **taiyang xiang xi yi le yi ge zhongtou*
sun toward west move PERF one CLS hour

(29)a. *qingkuang bianhao/huai le*
situation chang-good/bad PERF
'The situation has improved/worsened.'
b. *Ni zhege maobing yiding yao gai*
you this shortcoming must want change
'You must improve by getting rid of this shortcoming.'

The above contrasts clearly show that the lexical semantic specification of causative event-transition has many more implications then do the simple argument structure changes previously studied. For instance, the current explanation allows lexically specified direction of change-of-state, where *gai3* specifies a change of state for the better, while *bian2* has no such specification. Our study will show again how a lexical semantic attribute can be a powerful explanatory tool.[3]

5.5 Case Study 4: *qie1* vs. *ge1* - Manner

Last, we will look at the verbs of cutting *qie1* 'to cut' and *ge1* 'to slice' again. Huang and Tsai [1997] studied this near synonym pair and claim that the contrast is that *ge1* has the inherent attribute of [effect] and hence will take an incremental theme object while *qie1* cannot. The inherent [effect] attribute also allows prediction of the fact that cognate objects following the verbs are interpreted as results for *ge1* but as measurement for *qie1* (30 & 31).

(30) *ta ge-le yi kuai rou*
s/he GE-PERF 1-CLS meat
'S/He made a slice of meat.'

(31)a. *ta ba shoubi gele shiji-dao yi shi juexin*
s/he BA arm GE-PERF ten-plus-knife so show resolution
'S/He cut more than ten cuts on his/her arm to show his/her resolution.'
vs. b. *zai qing-di shen-shang kanle wu-shi-liu dao*
at love-foe body-top KAN-PERF 56 knife
'[The person] hacked his/her rival in love affair 56 times.'

However, the [effect] account may not offer a complete and sufficient lexical semantic explanation. We also observe that there is a fundamental difference in manner between the two verbs. That is, *ge1* entails a careful, traceable movement that has an inherent time-duration while *qie1* denotes a movement whose manner is not specified;

[3] We will not give a MARVS representation for the verbs in this section. The MARVS representation of linked events, such as causative, purposive etc., are still being developed.

thus, neither is its time-duration. When *ge1* co-occurs with a durational complement, it is more likely to interpret the duration as one single movement of *ge1*; when a durational complement co-occurs with *qie1*, it is interpreted as the cumulated time of consecutiv episodes of *gie1*. The emphasis on manner may help to explain why there are far more descriptive VR compounds with *ge1* then with *qie1*. The above account is translated intoa MARVS representation in Diagram 3. Note that the syntactic realization can have either a complete set or a subset of roles found in the lexical semantic representation (cf. Example 30, which only has two roles).

Diagram 3
MARVS representation of *ge1* and *qie1*

ge1 •//// • <Agent, Theme, Manner>
|
[effect]

qie1 •//// • <Agent, Theme>

In sum, the emphasis on manner, as an event-internal attribute of the verb *ge1*, also contributes to account for the contrast between the two near synonyms. This observation suggests that we should look into how different lexical semantic attributes can conspire to produce the same grammatical entailments. Whether they can also contradict or even cancel each other out will be another interesting topic for future studies.

5.6 Summary

First, the observed grammatical contrasts between *bai3* 'to set' and *fang4* 'to put' show that *bai3* specifies positioning with structural/spatial design' while *fang4* names simple positioning. The feature [design] inherent in the meaning of *bai3* leads to a crucial implication about the aspectual focus of its event structure; namely, *bai3* is focused on the result-state, the durative state resultant of the event of *bai3*. This, in turn, explains the corpus-based collocational patterns associated with *bai3*.

Second, the grammatical contrasts between *peng4* 'to touch, bump' and *mo1* 'to touch, caress' are found to be significant with regard to event contour. To be specific, *peng4* specifies (as a sub-event) a motional path ending with a fixed boundary. Event contour specifications may in turn account for polysemic extensions of many subclasses for verbs of contact.

Third, both "change of state" verbs (*gai3* 'to revise' vs. *bian4* 'to transform') and change of position' verbs (*ban1* 'to move (something) vs. *yi2* '(something) moves') share

a causative event-transition attribute. The attribute explains the differences in argument structure as well as in lexical implications (i.e.,'positive or motivated change' for *gai3*; 'upward detachment' for *ban1*).

Finally, the manner distinction between *qie1* 'to cut' and *ge1* 'to slice' has a consequence for the interpretation of their object-role. The careful, traceable, and time-consuming movement inherent in *ge1* enables it to take an incremental theme as its object, and allows it to be combined in ditransitive VV compounds where the incremental theme is transferable (an alternative explanation to that of Huang and Tsai 1997).

Altogether, the four cases serve to illustrate one point: generalized event structure attributes derived from lexical meaning contrasts can be utilized to categorize and represent verbal information across natural classes, which is crucial for obtaining an explanatory account of Mandarin verbal semantics.

6. Conclusion

In this paper, we have set out the underpinnings of our new representation of lexical knowledge, known as the Module-Attribute Theory of Verbal Semantics (MARVS). This theory incorporates and supercedes the earlier versions proposed by Huang and Tsai [1997] and Huang [1998]. In this theory, the event contour (i.e., the aspectual information) is represented using the composition of five atomic event modules, which can be combined to form a complete event representation. In addition, event participants are represented using role modules. It is worth noting that the range of roles is wider than that which is traditionally covered by theories of predicate-argument structures. Crucially, our role modules represent all participant roles that semantically contribute to the event content and have grammatical consequences. In other words, the roles that are traditionally termed optional arguments or adjuncts can be represented in the role modules as long as there is evidence showing their contribution to the verbal semantics of the verb. In addition, semantic attributes can be added to the backbone of event structure (event-internal attributes) and roles (role-internal attributes) to elaborate the lexical semantic information

Based on the hierarchical inheritance relation, we postulate that some of the attributes can be predicted with logical implicatures based on the modules they attach to. In addition, role modules can be partially predicted with event modules. In other words, only the attributes that are not logically implied need to be lexically specified.

We adopt a strong hypothesis built upon the above premise of the encoding of eventive information: that each sense of a lexical verb is uniquely identified with an

eventive information representation. This follows from the premise and the lexicographic convention of treating a sense as a (prepackaged) information unit. Two theoretically significant consequences of this hypothesis are that there must be a clear-cut criterion to identify verbal senses, and second, that lexical event encoding must allow variations at the realization level. First, the simple criterion is that meanings that cannot be represented by the same event structure must belong to different senses. Second, a verbal semantic puzzle that has not been fully accounted for is that the aspectual type/event classification of a bare verb can often be changed when it co-occurs with certain complements/adjuncts. This poses a dilemma for the lexical representation of eventive information. On one hand, if a sense is not identified with a unique event structure, then this variation seems to force either complex and ambiguous event encodings on each lexical verb or contextual encoding of event information. On the other hand, if the one event structure per sense principle is adopted, then this fact seems to suggest superfluous lexical ambiguity by predicting as many senses as possible event type realizations.

Our study has shown that this dilemma is unnecessary. Human lexical conceptualization does not necessarily stop at a logical event unit (such as Vendler's activity, state, accomplishement, or achievement). In other words, lexical conceptualization can integrate the complex course of an event, possibly by including many elements of the above classifications. The whole integrated event representation will then be the lexical meaning of the verb. However, when grammatically realized, the focus can fall on part of that complex and complete event representation, and the verbal semantics can be projected to one of the typical event classes.

Two crucial assumptions of such explanation are that these focused sub-events do not contradict each other, and that the sum of the realized events can be conceptually motivated and formally represented. In other words, an activity *run* (as in 'he is running') and an accomplishment *run* (as in 'he ran two miles') belong to the same sense and share the same lexical eventive representation. Speakers focus on different aspects of the same event structure, the first one on the activity part but the second on the endpoint.

A similar approach can be taken towards the problems involving so-called optional arguments and obligatory adjuncts. We do not need to worry about the relationship between each predicate argument structure and each sense. Instead, a possible set of roles can be specified based on the complete event structure as described above. Then each realization, with a different event focus, will take a subset of the encoded roles.

In sum, MARVS differs from previous attempts to understand lexical knowledge, especially those based on the interaction of syntactic-semantic information in verbs, because it analyzes near synonym pairs. It also differs in postulating event structure

modules, which may be combined to form a complex representations and may be attached to a verb.

We have examined four sets of near synonym contrasts based on the Module-Attribute Representation of Verbal Semantics. We found that both the composition of event modules and the attested lexical semantic attributes can be generalized across the natural class they belong to. This is a crucial step towards establishing an explanatory account of Mandarin verbal semantics. Our explanations not only offer support for the MARVS theory of lexical semantic representation, but also demonstrates the explanatory power of lexical semantics in a theory of (Chinese) grammar.

Acknowledgements

The authors would like to thank the National Science Council for support under grants NSC89-2420-H-001-002, NSC88-2411-H-001-040-M8, and NSC87-2411-H-110-040-M7, all of which contributed towards supporting the research for and writing of the paper. The authors would like to thank their colleagues at CKIP, Academia Sinica, who worked on the verbal semantics project for the past three years. The two reviewers of CLCLP provided detailed comments that helped to focus our arguments and sharpen our presentation. We would also like to thank Vicky Tzu-yin Lai for her help in preparing the final version of the manuscript. All errors are, of course, our own responsibility.

References

Ahrens, Kathleen, Li-Li Chang, Keh-jiann Chen, and Chu-Ren Huang. 1998. "Meaning Representation and Meaning Instantiation for Chinese Nominals." *Computational Linguistics and Chinese Language Processing*. 3.1:45-60.

Bresnan, Joan and Jonni Kanerva. 1989. "Locative Inversion in Chichewa." *Linguistic Inquiry,* 20:1-50.

Chief, Lian-Cheng, Chu-Ren Huang, Keh-Jiann Chen, Mei-Chih Tsai, and Lili Chang. 2000. "What Can Near Synomyms Tell Us." *Computational Linguistics & Chinese Language Processing*. 5.1:47-60.

Dowty, David. 1991. "Thematic Proto-Roles and Argument Selection." *Language*. 67:547-619.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Grandy, Richard E. 1992. "Semantic Fields, Prototypes, and the Lexicon." In Lehrer and Kittay Eds. 1992. 103-122. Hillsdale: Lawrence Erlbaum.

Huang, Chu-Ren. 1998. "Classifying Event Structure Attributes: A Verbal Semantic Perspective from Chinese." Invited paper. Chinese Workshop. *The 1998 International Lexical-Functional*

Grammar Conference. Brisbane, Australia.

Huang, Chu-Ren and Shen-Min Chang. 1996. Metaphor, Metaphorical Extension, and Grammaticalization: A Study of Mandarin Chinese *-qilai*. A. Goldberg. Ed. *Conceptual Structure, Discourse, and Language.* Stanford: CSLI, and Cambridge: Cambridge U. Press.

Huang, Chu-Ren and Mei-Chih Tsai. 1997. "From Near Synonyms to Event Structure: Corpus-based Studies of Mandarin Verbal Semantics." *Mini-Conference on Lexical Semantics.* November. National Chung-Cheng University.

Huang, Chu-Ren, Chang, Li-Ping Kathleen Ahrens, and Chao-ran Chen. 1999. "The Interaction of Lexical Semantics and Constructional Meanings". (In Chinese) *Chinese Langauges and Linguistics V.* Taipei: Institute of Linguistics, Academia Sinica.

Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge: MIT Press.

Jackendoff, Ray. 1983. *Semantics and Cognition.* Cambridge: MIT Press.

Lehrer, Adrienne and Eva Efder Kittay. 1992. Eds. *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization.* Hillsdale: Lawrence Erlbaum.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation.* Chicago: University of Chicago Press.

Levin, Beth and Steven Pinker (Eds.) 1991. *Lexical and Conceptual Semantics.* A Cognition Special Issue. Cambridge: Blackwell.

Liu, Meichun. 1999. "Lexical Meaning and Discourse Patterning: the Three Mandarin Cases of "build"." In Barbara A. Fox, Dan Jurafsky and Laura A. Michaelis, eds. *Cognition and Function in Language.* 181-199. Stanford: Center for the Study of Language and Information (CSLI).

Liu, Mei-chun, Chu-Ren Huang, and Charles C.L. Lee. 2000. "When Endpoint Meets Endpoint: A Corpus-based Lexical Semantic Study of Mandarin Verbs of Throwing." *Computational Linguistics & Chinese Language Processing.* 5.1:81-96.

Luo Yi-fen and Chu-Ren Huang. 1997. Hanyu Dongmaoci Zai yu Zhe de Cihui Yiyi [The Lexical Semantics of Mandarin Aspects 'Zai' and 'Zhe']. *Proceedings of the Fifth International Conference on World Chinese Language Teaching*. Taipei: World Chinese Association.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge: MIT Press.

Pustejovsky, James. 1991. "The Syntax of Event Structure." *Lexical and Conceptual Semantics.*

Pustejovsky, James, S. Bergler, and P. Anick. 1993. "Lexical Semantic Techniques for Corpus Analysis." *Computational Linguistics.* 19.2:331-358.

Smith, Carlota. 1991. *The Parameter of Aspect.* Dodrecht: Kluwer.

Tsai, Mei-Chih, Chu-Ren Huang, Keh-Jiann Chen, and Kathleen Ahrens. 1998. "Towards a Representation of Verbal Semantics -- An approach based on near synonyms." *Computational Linguistics and Chinese Language Processing*. 3.1:61-74.

Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca: Cornell University Press.

Wierzbica, Anna. 1985. *Lexicography and Conceptual Analysis.* Ann Arbor: Karoma.

Zaenen, Annie. 1993. "Unaccusativity in Dutch: An Integrated Approach." In J. Pustejovsky Ed. *Semantics and the Lexicon*. 129-161. Dordrecht: Kluwer.

Computational Linguistics and Chinese Language Processing
Vol. 5, No. 1, February 2000, pp. 47-60

© Computational Linguistics Society of R.O.C.

What Can Near Synonyms Tell Us? [1]

Lian-Cheng Chief*, Chu-Ren Huang*, Keh-Jiann Chen*, Mei-Chih Tsai+, Li-li Chang*

Abstract

This study examines a near synonym pair *fangbian* and *bianli*, 'to be convenient,' and extracts the contrasts that dictate their semantic and associated syntactic behaviors. Corpus data reveal important but opaque distributional differences between these synonyms that are not readily apparent based on native speaker intuition. In particular, we argue that this synonym pair can be accounted for with a lexical conceptual profile. This study demonstrates how corpus data can serve as a useful tool for probing the interaction between syntax and semantics.

1. Introduction

The aim of this paper is to find the semantic features that determine the relevant syntactic behaviors of the near synonym pair *fangbian* and *bianli*. Tsai *et al.* [1998 & 1999], in their recent comparative studies of near synonymous Chinese verbs, claim that basic semantic components or features can predict the different syntactic behaviors of near synonyms. One example is their comparison of the near synonym pair *gaoxing* and *kuaile* 'happy vs. glad.' Tsai *et al.* [1998] proposed two features, [±effect] and [±control], to account for the different syntactic behaviors of these synonyms. In this study, we use the same methodology to find other semantic features that can predict syntactic patterns. The syntactic patterns of the near synonym pair *fangbian* and *bianli*, which mean 'to be convenient,' are examined to extract relevant semantic features. We demonstrate that the lexical conceptual profile is one semantic feature that determines the relevant syntactic behavior of the near synonym pair. It is hoped that each proposed semantic feature will contribute to our understanding of the interaction between syntax and semantics. This paper is organized as follows. First, we introduce our methodolog in section 2.

* Academia Sinica

+ National Sun Yat Sen University

[1] An earlier version of this paper appeared in the electronically published Proceedings of the LFG98 Conference (Miriam Butt and Tracy Holloway King Eds, http://www-csli.standford.edu/LFG/3/lfg98.html). We would like to thank Kathleen Ahrens for her detailed comments on several versions of this paper. We are also grateful to participants of the conference as well as colleagues at CKIP, Academia Sinica for their helpful comments.

Then, we discuss the syntactic behaviors of and the distributional differences between these synonyms in section 3. The final section summarizes the information that near synonyms can give us.

2. Methodology

Our approach is corpus-aided. In addition to the syntactic variations that can be easily recognized by means of our intuition, implicit or opaque distributional differences in terms of syntactic functions that cannot be discerned simply by means of intuition were extracted from the Sinica Corpus. Specifically, we believe that introspection is incomplete, and that distributional information is important in contrastive studies on near synonyms. Our aim is to try to determine the syntactic and semantic differences between members of near synonym pairs. We follow the approach adopted by Tsai *et al.* [1999]. The first step is to determine distributional differences in syntactic patterns. The second step is to deduce the semantic features from the syntactic phenomena. Finally, we test the semantic features in new syntactic frames.

Through this approach, several semantic features have been discovered. For example, [±effect] can account for the distinctions between *lei* and *pijuan* 'tired,' and *gaoxing* and *kuaile* 'happy or glad.' In the case of *lei* and *pijuan*, [±effect] accounts for why *lei* can be a resultative complement while *pijuan* cannot. In the case of *gaoxing* and *kuaile*, [±effect] explains why *gaoxing* can be associated with the sentential-final particle *le*, whereas *kuaile* cannot. This is because *gaoxing*, with the feature [±effect], represents a change of state triggered by some cause. In addition, [±telic] is used to explain the differences between *quan* and *shuifu* 'persuade.' [±control] distinguishes between *gaoxing* and *kuaile*.[2] Liu [1997] also employs the same methodology to account for the distinctions among three Mandarin verbs of 'build,' *jian*, *zao*, and *gai*. These previous studies demonstrate that semantic components account for the syntactic differences between the members of near synonym pairs. In other words, these studies offer evidence that syntactic behaviors can be predicted based on lexical semantics. This is also the point that the present study aims to support.

3. The Data

The data used in this study were taken from the Sinica Corpus (version 2.0), which contains 3.5 million tagged Chinese words.[3] In this corpus, we found 445 entries of

[2] For details, please refer to Tsai *et al.* [1999].

[3] Sinica Corpus 3.0, which contains 5 million words, was released in June of 1998. It can be found at *http://www.sinica.edu.tw/ftms-bin/kiwi.sh*.

fangbian and 125 entries of *bianli*. We will first present their syntactic behaviors in section 3.1 and then their distributional differences in section 3.2.

3.1 The Near Synonym Pair: *Fangbian* and *Bianli*

The members of the near synonym pair *fangbian* and *bianli* are used to define each other in many dictionaries. In addition to their similarity in meaning, these two verbs seem to be parallel syntactically. For instance, both of them have transitive and intransitive usages, can serve as nominal modifiers, and undergo nominalization. In this section, we will introduce their syntactic behaviors.

3.1.1 The Transitive/Intransitive Alternation

Fangbian and *bianli* both have transitive and intransitive usages. Sentences (1) and (2) show the intransitive usages of these two verbs.

(1) 停車 方便
tingche **fangbian**
parking convenient
'Parking (here) is convenient.'

(2) 交通 便利
jiaotong **bianli**
traffic convenient
'Transportation is convenient.'

In addition to their intransitive usages, they also have transitive usages as shown in sentence (3) and (4).

(3) 設置 辦事處 方便 民眾 出國 觀光
shezhi banshichu **fangbian** minzhong chuguo guanguang
establish office convenient people go-abroad visit
'Establishing an office makes it convenient for people to travel abroad.'

(4) 修改 許多 法規 便利 山民 墾植
xiugai shuduo fagui **bianli** shanmin kenzhi
modify many rule convenient mountain-people cultivate
'Modifying many rules makes it convenient for the aborigines to cultivate [land].'

In their intransitive usages, both *fangbian* and *bianli* take a proposition as a subject. In their transitive usages, they take a propositional object. Usually, the propositional subject or propositional object is represented by a clause, a verb phrase, or a complex nominal element. The proposition describes what is convenient. However, the propo-

sitional object of *fangbian* can undergo inversion as in (5a) and (5b) while *bianli* does not allow such alternation.

(5a) 理想 的 場地 是 鄰近 工作 地點， 方便 員工 參加
lixiang de changdi shi linjin gongzuo didian, **fangbian** yuangong canjia
ideal DE place be near work place convenient worker join
'An ideal location is near the work place and convenient for workers to join (the meeting). '

(5b) 理想 的 場地 是 鄰近 工作 地點，員工 參加 方便
lixiang de changdi shi linjin gongzuo didian, yuangong canjia **fangbian**
ideal DE place be near work place workers join convenient
'An ideal location is near the work place and convenient for workers to join (the meeting). '

(6a) 有 各種 產品， 便利 消費者 選購
you gezhong changpin bianli xiaofeizhe xuan-gou
have various product convenient consumer choose-buy
'The variety of products makes selection convenient for consumers. '

(6b) * 有 各種 產品， 消費者 選購 便利
you gezhong changpin xiaofeizhe xuan-gou bianli
have various product convenient consumer choose-buy

We will account for this phenomenon in section 4.

3.1.2 Other Syntactic Functions of *fangbian* and *bianli*

In addition to verbal predicates, these two near synonyms can also appear as nominal modifiers and undergo nominalization. (7) and (8) illustrate the use of *fangbian* and *bianli* as nominal modifiers.[4]

(7) 方便 的 資訊
fangbian de zixun
convenient de information
easily-accessible information

[4] However, we only found examples of *bianli* (but not *fangbian*) used in nominal compounds in the Sinica Corpus as shown below. We do not account for this difference in this paper.

便利 商店
bianli shangdian
convenient store
convenience store

(8) 便利 的 方式
bianli de fangshi
convenient de way
convenient way

Examples (9) and (10) show that this pair of near synonyms can function as nominal heads.

(9) 聯繫 上 的 方便
lianxi shang de fangbian
communicate in de convenience
'convenience in communicating'

(10) 生活 的 便利
shenghuo de bianli
living de convenience
'convenience in living'

As shown above in this section, it appears that *fangbian* and *bianli* can be used interchangeably. However, the statistics obtained from the corpus demonstrate that these syntactic patterns have different statistical distributions.

3.2 Distributional Differences

In this section, we will examine the distributional differences based on all the examples extracted from the Sinica Corpus. After searching for all the instances of *fangbian* and *bianli* in the corpus, we first classified each occurrence according to its syntactic function, such as nominal verbal predicate, nominal modifier, and verbal modifier. Second, we calculated the number of occurrences of transitive and intransitive alternations of these synonyms as verbal predicates. Third, we classified them in terms of the object types they take. The results demonstrate clear distributional differences.

3.2.1 Distributional Differences in Terms of Syntactic Functions

In this section, we will present the distributional differences in terms of syntactic functions. The range of syntactic functions of this near synonym pair can be illustrated by the previously given examples (1)-(10) as well as (11) below.

(11) 使用者 可以 更 方便 的 處理 事情
shiyungzhe keyi geng fangban de chuli shiqing
user can more convenient de manage thing
'Users can manage things more conveniently.'

Four different functions are identified. First, verbal predicates are exemplified by (1)-(6). Second, nominal modifiers are given in (7) and (8). Third, (9) and (10) are instances of nominalization. Lastly, (11) is an example of a verbal modifier in which *de* is preceded by *fangbian* and followed by a head verb. We cannot find any example in which *bianli* is used in this way in our corpus, which also confirms our intuition.

Table 1 illustrates their distributions in terms of syntactic functions.

Table 1. *Distributional Differences in terms of Syntactic Function*

	Verbal Predicates	Nominal Modifiers	Verbal Modifiers	Nominalization
Fangbian 445	77%	7%	5%	10%
Bianli 125	44%	34%	0%	22%

In Table 1, some differences between *fangbian* and *bianli* can be found. First, *bianli* cannot be used as a verbal modifier, whereas *fangbian* can. Second, when used as a nominal modifier, *bianli* is preferred more than *fangbian*. These two pieces of evidence give rise to two questions. First, why can't *bianli* be used as a verbal modifier? Second, why is *bianli* often selected when people try to express the idea that some event is convenient?

3.2.2 Distributional Differences in terms of the Transitive / Intransitive Alternation

The distributional differences indicated in Table 2 show that *fangbian* more often appears in intransitive form (e.g. (1)) while *bianli* shows no such preference. In addition, when used as a transitive verb, *fangbian* usually takes a sentential object (e.g. (3)).

Table 2. *Distributional Differences in terms of the Transitive/Intransitive Alternation*

	Transitive	Intransitive
Fangbian 342	31%	69%
Bianli 55	53%	47%

Table 3. *Distributional Difference in terms of the Type of Object*

	Sentential or Verbal Object	Complex Nominal Object
Fangbian 107	90%	10%
Bianli 29	62%	37%

3.2.3 Negation

From the corpus, we also find that *bianli* cannot be modified by the negative marker *bu* 'not,' as shown in Table 4.

Table 4. *Co-occurrence with Negative Marker bu 'not'*

	Negation (preceded by *bu* 'not')	Total instances
fangbian	44	445
bianli	0	125

This also gives rise to the second question as to why *bianli* cannot be negated syntactically.

3.3 Summary

The distributional differences extracted from the corpus not only give us a clear picture of their differences in usage, but also show the inadequacy of their present definitions in dictionaries. Though they are used to define each other in many dictionaries, their differences in terms of function and distribution are neither described nor explained.

4. Explanation

To account for the observed differences in syntactic distribution, we propose two semantic factors, (i) beneficial role and (ii) lexical conceptual profile. In other words, we propose that there is a beneficial role in the argument structure of *bianli*. Further, we point out that profiling different perspectives of an event nicely captures the differences between the two verbs. In this paper, the lexical conceptual **profile** refers to the most prominent or salient sub-part of the whole event. Specifically, in a group of verbs that are similar in meaning, there are different focal points in different participants or different levels of verb frames. A similar but not identical idea can be found in Goldberg [1995] and Croft [1998], in which profiling is also used to describe semantic differences among verbs.

4.1 Beneficial Role

From the evidence presented in section 3, there are at least four major differences between *fangbian* and *bianli*. First, *bianli* never appears as a verbal modifier. Second, *bianli* occurs as a transitive verb in most cases. Third, in 90% of the instances in which *fangbian* is used as a transitive verb, it takes either a sentential or a verbal object. Fourth, *bianli* cannot be negated. To account for these variations, we propose that *fangbian* profiles the whole event, whereas *bianli* profiles the beneficial role of the event. The following pair of sentences ((12a) and (12b) repeated from (3) and (4)) illustrates this.

(12a) 設置 辦事處 方便 民眾 出國 觀光
shezhi banshichu **fangbian** minzhong chuguo guanguang
establish office convenient people go-abroad visit
'Establishing an office makes it convenient for people to travel abroad.'

(12b) 修改 許多 法規 便利 山民 墾植
xiugai shuduo fagui **bianli** shanmin kenzhi
modify many rule convenient mountain-people cultivate
'Modifying many rules makes it convenient for the aborigines to cultivate [land].'

In sentence (12a), the main verb is *fangbian*, and the verbal meaning profiles the whole embedded event "people go abroad and visit." The syntactic evidence as shown by the constructed sentences (13a) and (13b) support this argument because in (13a), the post-verbal element, the propositional event, can be inverted to the pre-verbal position, whereas in sentence (13b), such an inversion is not allowed.

(13a) 設置 辦事處 民眾 出國 觀光 方便
shezhi banshichu minzhong chuguo guanguang **fangbian**
establish office people go-abroad visit convenient
'Establishing an office makes it convenient for people to travel abroad.'

(13b)* 修改 許多 法規 山民 墾植 便利
xiugai shuduo fagui shanmin kenzhi **bianli**
modify many rule mountain-people cultivate convenient

Furthermore, in contrast to (12a), in (12b) the main verb is *bianli*, and the verbal meaning profiles the beneficial role (the aborigines) of the embedded event (to cultivate). In other words, the focus of sentence (12b) is on *the aborigines* who cultivate rather than the event "to cultivate" itself. Therefore, we propose a semantic feature which shows the difference between these near synonyms to be [±beneficial role]. Specifically, the beneficial role in the event structure of *bianli* is prominent. In contrast, there is no beneficial role in the event structure of *fangbian*, or its status is trivial. In short, the meaning of this pair of near synonyms is 'to be convenient,' but the concept of convenience is on different levels. For *fangbian*, it means that the way to perform the action is convenient, whereas for *bianli*, it means that for the profiled entity, the action is convenient or beneficial to perform.

4.2 Profile on Event vs. Profile on Beneficial Role

The notion that the lexical conceptual profile focuses on different sub-parts of an event also accounts for the differences between *fangbian* and *bianli*. First, we have observed

that *bianli* cannot function as a verbal modifier. In other words, when people want to describe that a certain event is easily conducted, they will choose *fangbian* to express this concept. Why is this so? Since the lexical conceptual profile of *fangbian* focuses on the propositional event, when *fangbian* modifies a verb, the eventive profile is projected to the sentential level, and semantic composition is preserved. In other words, a profile of the whole propositional event is the inherent meaning of *fangbian*. In contrast, the lexical conceptual profile of *bianli* focuses on the beneficial role of the propositional event; therefore, semantic compositionality cannot be maintained if *bianli* is used to modify a verb.

Second, the data from the corpus show that *bianli* cannot be negated whereas *fangbian* can be negated by the negative marker *bu* 'not.' Our proposed semantic features also properly explain this. First, since the profile of *fangbian* focuses on the whole positional event, it can be negated like any proposition. Therefore, *fangbian* can co-occur with *bu*. In contrast, the profile of *bianli* focuses on the beneficial role rather than the whole sub-event. In order for the profile to focus on the beneficial/causee role, the whole proposition must be presupposed. Also, it is well-known that a presupposition cannot be negated/cancelled. In addition, the semantics of the beneficial role also exclude negation since the semantics of *bianli* denote a positive meaning. It would be semantically anomalous if the predicate were negated.

4.3 Syntactic Patterns

Based on the two semantic features, the beneficial role and the lexical conceptual profile, we propose that *fangbian* and *bianli* have different event structures and argument structure frames.

(14) *fangbian* [AGENT **GOAL (Proposition)**]

<SUBJ XCOMP>

(15) *bianli* [AGENT **BEN** GOAL (Proposition)]

<SUBJ OBJ XCOMP>

(14) and (15) show that *fangbian* has two roles (AGENT and GOAL), whereas *bianli* has three roles (AGENT, BEN, and GOAL). The shadowed bold text indicates the scope of the profile. That is, the profile of the event of *fangbian* focuses on the whole embedded

event, whereas that of *bianli* focuses on the beneficial role. As mentioned previously, this account has two advantages. First, *bianli* cannot be an adjunct of a verb because it does not profile an event. On the contrary, *fangbian* can modify a verbal predicate because its semantics inherently profile an event. Second, *fangbian* rather than *bianli* can be negated because the scope of the negation can cover the whole sub-categorized XCOMP of *fangbian* but cannot cover the XCOMP of *bianli*.

Finally, the difference in lexical conceptual profile also accounts for the syntactic alternation of *fangbian* and the lack of such alternation of *bianli* as shown in (5) and (6), and repeated here for convenient reference.

(5a) 理想 的 場地 是 鄰近 工作 地點， 方便 員工 參加
lixiang de changdi shi linjin gongzuo didian, **fangbian** yuangong canjia
ideal DE place be near work place convenient worker join
'An ideal location is near the working place and convenient for workers to join (the meeting).'

(5b) 理想 的 場地 是 鄰近 工作 地點， 員工 參加 方便
lixiang de changdi shi linjin gongzuo didian, yuangong canjia **fangbian**
ideal DE place be near work place workers join convenient
'An ideal location is near the working place and convenient for workers to join (the meeting).'

(6a) 有 各種 產品， 便利 消費者 選購
you gezhong changpin bianli xiaofeizhe xuan-gou
have various product convenient consumer choose-buy
'The variety of products makes selection convenient for consumers.'

(6b) * 有 各種 產品， 消費者 選購 便利
you gezhong changpin xiaofeizhe xuan-gou bianli
have various product convenient consumer choose-buy

Sentences (5)-(6) demonstrate that post-verbal elements of *fangbian* can undergo inversion whereas those of *bianli* cannot. Since *bianli* has two postverbal elements, one of the grammatical functions cannot be inverted by itself. On the contrary, *fangbian* has only one post-verbal element.[5] In brief, the syntactic profile cannot contradict the lexica conceptual profile.

[5] For the scope of this paper, we do not discuss which pattern (transitive/intransitive) of *fangbian* is the basic pattern nor do we discuss whether *fangbian* has two lexical entries or one lexical entry.

4.4 An Additional Perspective

An additional possibility is that the distinction between this pair of synonyms might have to do with the distinction between the type and token of a certain event. Since *fangbian* profiles the whole proposition event and *bianli* profiles the beneficial role of the event, *fangbian* tends to be used to describe a generic event while *bianli* tends to be used to describe the specific event. The profile of the event of *bianli* focuses on how the event affects the individual who performs the action. In the event of *fangbian*, the status of the individual is trivial. It is important that the manner/way to perform the action/event is convenient. Therefore, *fangbian* comments on the generic event. On the contrary, *bianli* focuses on the individual. It profiles how the individual performs the action in each event, so *bianli* tends to be used to describe a specific event. In conclusion, we suggest that the type and token are also the potential distinctions between *fangbian* and *bianli*. *Fangbian* refers to a group of events, that is, the type of event. *Bianli* refers to a single event, that is, the token of the event.

4.5 Summary

From distributional syntactic differences, we have discovered differences between *fangbian* and *bianli* that are not easily determined solely by means of intuition. We assert that two semantic factors determine the relevant syntactic behaviors of these near synonyms. The lexical conceptual profile accounts for why *bianli* cannot function as an adjunct of verb and why *bianli* cannot be negated. The additional beneficial role of *bianli* explains the lack of syntactic alternation that *fangbian* allows. Finally, the distinction between event type and event token also contributes to the distributions of these synonyms.

5. What Can Near Synonyms Tell Us

The hypothesis that the syntactic behaviors of verbs are semantically determined has been supported by a series of studies which have compared near synonyms. The present study can be viewed as one of the building blocks contributing to the study of Mandarin Chinese lexical semantics, based on the framework proposed by Huang and Tsai [1997]. The semantic features proposed in this paper to distinguish between the relevant syntactic behaviors of the near synonyms *bianli* and *fangbian* are lexical conceptual profile and beneficial role. The lexical conceptual profile determines both the syntactic function that a word can have and also the scope of negation in embedded predicates. The presence or absence of a beneficial role predicts the relevant syntactic alternation.

So far, this series of studies [Tsai *et al.* 1998 & 1999 as well as Huang *et al.* 2000,

Chang *et al.* 2000] has proposed several semantic features that explain syntactic differences and predict syntactic behaviors. If semantics can properly predict syntactic behaviors, then pairs of words that have exactly the same meaning should have exactly the same syntactic behaviors. Therefore, the syntactic differences between near synonyms indicate the existence of subtle semantic differences. However, these syntactic differences are not easily discovered solely by means of intuition. In the present study, we used corpus data to find differences, and we then looked for semantic explanations for the relevant syntactic behaviors. In conclusion, this approach, which is based on comparing synonyms and is aided by corpus studies, provides a new way to understand the interaction between syntax and semantics in Mandarin Chinese.

References

Chang, Li-li, Keh-jiann Chen, Chu-Ren Huang. 2000. "Alternation Across Semantic Fields: A Study on Mandarin Verbs of Emotion." *International Journal of Computational Linguistics & Chinese Language Processing*. Vol.5 no.1 pp.61-80

Croft, William. 1998. "Event Structure in Argument Linking." In *The Projection of Arguments: Lexical and Compositional Factors*. Ed. by Miriam Butt and Wilhelm Geuder. pp. 21-65. Stanford: CSLI Publications.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.

Huang, Chu-Ren, Kathleen Ahrens, Li-li Chang, Keh-jiann Chen, Mei-Chun Liu and Mei-Chi Tsai. 2000. "The Module-Attribute Representation of Verbal Semantics: From Semantics to Argument Structure." *International Journal of Computational Linguistics & Chinese Language Processing*. Vol.5 no.1 pp.19-46

Huang, Chu-Ren and Mei-Chih Tsai. 1997. "From Near Synonyms to Event Structure: Corpus-based Studies of Mandarin Verbal Semantics." Paper presented at *Mini-Conference on Lexical Semantics*. National Chung Cheng University, Nov. 25.

Liu, Meichun. 1997. "Lexical Meaning and Discourse Patterning - The three Mandarin cases of 'build'." Paper presented at the *Third Conference on Conceptual Structure, Discourse, and Language*. Boulder, Colorado.

Liu, Meichun, Chu-Ren Huang, and Ching-Yi Lee. 1998. "When Endpoint Meets Endpoints: A Corpus-based Lexical Semantic Study of Mandarin Verbs of Throwing." Paper presented at the *7th International Conference on Chinese Linguistics/The 10th North American Conference on Chinese Linguistics*. Stanford University, June 16-18.

Tsai, Mei-Chih, Chu-Ren Huang, and Keh-jiann Chen. 1999. "You jinyici bianyi biaozhun kan yuyi jufa zhi hudong (From near-synonyms to the interaction between syntax and semantics)." *Chinese Languages and Linguistics*, 5:439-459.

Tsai, Mei-Chih, Chu-Ren Huang, Keh-jiann Chen, and Kathleen Ahrens. 1998. "Towards a Representation of Verbal Semantics - An Approach Based on Near Synonyms." *Computational Linguistics and Chinese Language Processing*, 3(1), pp. 61-74.

Computational Linguistics and Chinese Language Processing
Vol. 5, No. 1, February 2000, pp. 61-80

© Computational Linguistics Society of R.O.C.

Alternation Across Semantic Fields: A Study on Mandarin Verbs of Emotion

Li-li Chang*, Keh-jiann Chen*, Chu-Ren Huang+[1]

Abstract

This paper explores possible co-relations between lexical semantics and morpho-syntactic structures. We first examine a consistent dichotomy among verbs of emotion, which was first observed for verbs of happiness by Tsai *et al.* [1998]. It is shown that the dichotomy can be determined based on the criterion of whether a verb is a VV compound or not.[2] The linguistic contrasts observed include: the grammatical functions of a verb as well as their distribution, the selectional restrictions the verbs impose as an adjunct, a verb's occurrences in imperative and evaluative constructions, its *aktionsart*, and its transitivity. We will show that the overt morpho-syntactic contrasts are due to lexical event structure properties. The description of a state (of emotion) can focus on how the state comes to be (i.e., the inchoative state) or on the being of the state (i.e., the homogeneous state). Since VV compounding has the semantic function of referring to the generic properties of the set of event tokens, it is natural for VV compounding to be chosen as the morpho-syntactic representation of homogeneity.

1. Introduction

Recent lexical semantic studies, such as those of Levin [1993] and Pustejovsky [1995], have tried to explain how lexical meaning predicts syntactic regularities. One approach is to study the differences between near synonyms to identify the minimal semantic

* Institute of Information Science, Academia Sinica
+ Institute of Linguistics, Academia Sinica

[1] The authors are listed in alphabetical order of their last names. We would like to thank all our colleagues in the CKIP group who contributed to construction of the Sinica Corpus. Without it, it would have been impossible to obtain the findings and statistics that led to the generalizations presented in this paper. We would also like to thank Kathleen Ahrens for her comments on various versions of this paper. An earlier version of this paper was presented at PACLIC13 [1998] and appeared in the proceedings.

[2] What we refer to as "VV compounds" are also referred to as parallel verb compounds in the literature. They are verbs composed of two near synonym verbs or verb stems. For example, *bei1shang1* 悲傷 "to be sad" is a VV compound since both of its components are stative verb stems meaning "to be sad." In this paper, the term "VV compounds" is used in contrast to SV, VO, AV and VR compounds in Chinese.

attributes that create the differences [Tsai *et al.* 1998, Liu 1997, and Liu *et al.* 1998]. In this current study, we extend the range of study to semantic fields which contain more than one synonym pair. Thus, we can attest to the primary status of the semantic attributes proposed in previous studies by showing that similar generalizations can be extended to the other synonym pairs in the same semantic field.

This study attempts to elaborate on Tsai *et al.*'s [1998] work on the differences between the synonyms *kuai4le4* 快樂 "to be happy" and *gao1xing4* 高興 "to be glad." We re-examine the differences over a broader range, i.e., the verbs of emotion. Our study will lead to the following four important generalizations: 1) These differences are not specific to *kuai4le4* and *gao1xing4*, but to the whole semantic field of verbs of emotion. 2) These differences can be more succinctly defined. 3) These differences are motivated by different lexical event types. 4) The source of these differences can account for the choice of compound structure.

Verbs involving seven sub-fields of emotion will be examined in this paper, i.e., Happiness, Depression, Sadness, Regret, Anger, Fear and Worry[3]. This allows us to obtain generalizations about the complete semantic field of emotion as well as to observe if there are any variations among the sub-fields. All the observations and statistics in this paper are based on the "Academia Sinica Balanced Corpus of Modern Mandarin Chinese" (abbreviated as the "Sinica Corpus" in the following text), which is a tagged Mandarin corpus containing a total of five million words [CKIP, 1995]. We take into account only verbs with occurrence frequency rates of over 40 in the Sinica Corpus, so that there will be sufficient empirical evidence for any observed contrasts. The verbs studied here are listed in Table 1 with their frequency of occurrence in the Sinica Corpus listed in parentheses. There are thirty-three (33) verbs in total. Four of them are mono-syllabic, and twenty-nine (29) of them are disyllabic. In this study, we will focus on the disyllabic verbs in order to explore the co-relation between morpho-lexical composition and lexical semantics. In addition, the exclusion of mono-syllabic verbs allows us avoid dealing with potential complications involving polysemy and boundedness.

[3] Note that they have the uniform categorical assignment of stative verbs in Chinese while the translation equivalents may predominantly be adjectives in languages such as English. Our classification of the verbs of emotion largely follows that of the *Tongyici Cilin*.

Table 1. *Verbs of emotion with a frequency of over 40 in the Sinica Corpus*

Subtype	Verbs of Emotion and their Frequency of Occurrence in the Sinica Corpus
Happiness	*kuai4le4* 快樂(942), *gao1xing4* 高興(669), *yu2kuai4* 愉快(271), *le4* 樂(264), *xi3yue4* 喜悅(156), *kai1xin1* 開心(152), *huan1le4* 歡樂(141), *huan1xi3* 歡喜(107), *kuai4huo2* 快活(48), *tong4kuai4* 痛快(40)
Depression	*tong4ku3* 痛苦(443), *nang2guo4* 難過(232), *chen2zhong4* 沈重(83), *ju3sang4* 沮喪(62), *tong4xin1* 痛心(48)
Sadness	*shang1xin1* 傷心(134), *bei1shang1* 悲傷(52)
Regret	*yi2han4* 遺憾(198), *hou4hui3* 後悔(102)
Anger	*sheng1qi4* 生氣(295), *qi4* 氣(126), *fen4nu4* 憤怒(112), *qi4fen4* 氣憤(49)
Fear	*pa4* 怕(548), *hai4pa4* 害怕(261), *kong3ju4* 恐懼(149), *wei4ju4* 畏懼(40)
Worry	*dan1xin1* 擔心(609), *fan2nao3* 煩惱(199), *dan1you1* 擔憂(64), *fan2* 煩(54), *you1xin1* 憂心(46), *ku3nao3* 苦惱(45)

2. Initial observations and theoretical assumptions

2.1 Initial observation

According to Tsai *et al.* [1998], the verbs *gao1xing4* "to be glad" and *kuai4le4* "to be happy" differ in the following four aspects: 1) *gao1xing4* takes sentential objects while *kuai4le4* cannot. 2) *gao1xing4* takes the sentential-final particle *le5* while *kuai4le4* cannot. 3) *gao1xing4* never occurs in wish sentences but admits evaluative sentences while *kuai4le4* occurs in wish sentences but never appears in evaluative sentences.[4] Lastly, 4) *gao1xing4* occurs in imperative sentences while *kuai4le4* cannot.

We noticed that the differences between *gao1xing4* and *kuai4le4* create a clear dichotomy among verbs of Happiness. They are the two most frequently used verbs of the sub-field; moreover, all other verbs of Happiness also fall onto either side of the dichotomy. Hence, it is natural to ask if such generalizations can be carried over to other sub-fields of emotion. In other words, are these contrasts idiosyncratic to verbs of Happiness or do they represent some common conceptual motivation of the verbs of emotion? Our thorough examination of the verbs of emotion has revealed that these differences are repeated in each of the seven sub-fields of emotion. In addition, we have also found additional representational clues as to this dichotomy. Based on our study, five distributional criteria can be used to create a bipartite classification of the verbs of emotion. The first two criteria are newly proposed here, while the following three were proposed by Tsai *et al.* [1998] :

[4] Wish sentences refer to the greeting constructions, such as *zhu4 ni3 kuai4le4* 祝你快樂 'May you be happy!'.

a. the distribution of its grammatical functions;

b. its selectional restrictions when it functions as an adjunct;

c. its occurrence in imperative and evaluative constructions;

d. its verbal aspect or *aktionsart;*

e. its transitivity.

According to the above five criteria, we classify each of the 29 disyllabic verbs as one of two types for each of the seven subclasses of verbs of emotion. Type A includes the verbs similar to *gao1xing4*, and type B includes verbs similar to *kuai4le4*. For each of the seven subclasses, the two most frequent verbs form a contrast pair, i.e., one is type A and one is type B, as shown in Table 2.

Table 2. *Dichotomy of the Verbs of Emotion*

Subtype	Type A	Type B
Happiness	***gao1xing4*** 高興(669) *kai1xin1* 開心(152) *tong4kuai4* 痛快(40)	***kuai4le4*** 快樂(942) *yu2kuai4* 愉快(271) *xi3yue4* 喜悅(156) *huan1le4* 歡樂(141) *huan1xi3* 歡喜(107) *kuai4huo2* 快活(48)
Depression	***nan2guo4*** 難過(232) *tong4xin1* 痛心(48)	***Tong4ku3*** 痛苦(443) *chen2zhong4* 沈重(83) *ju3sang4* 沮喪(62)
Sadness	***shang1xin1*** 傷心(134)	***bei1shang1*** 悲傷(52)
Regret	***hou4hui3*** 後悔(102)	***yi2han4*** 遺憾(198)
Anger	***sheng1qi4*** 生氣(307)	***fen4nu4*** 憤怒(112) *qi4fen4* 氣憤(49)
Fear	***hai4pa4*** 害怕(261)	***kong3ju4*** 恐懼(149) *wei4ju4* 畏懼(40)
Worry	***dan1xin1*** 擔心(609) *dan1you1* 擔憂(64) *you1xin1* 憂心(46)	***fan2nao3*** 煩惱(199) *ku3nao3* 苦惱(45)

2.2 Theoretical Premise: Contrast-based semantic fields

The fact that, in each of the fields of emotions we have examined, the two most frequent and, therefore, most dominant terms form a contrast pair leads us to adopt a revision of Grandy's [1992] definition of a semantic field.[5] Even though Grandy formulated that membership in a semantic field as defined by contrast pair relations, he makes the covering term a crucial exception. This means that a single and unique covering term heads each semantic field and does not enter into contrast relations with other terms. However, since the (possibly transitive) contrast relations comprise the defining relation of membership in a semantic field, the fact that the covering term is not definable by such a relation seems to be an anomaly. On the other hand, if a covering term does enter into contrast pair relations with other terms, how can the primary status of the covering term be distinguished (from all other contrast pair relations)? Since our data clearly show that there are two dominant terms in each semantic field, they suggest an alternative view that there are possibly two covering terms for each field.

Our proposal is that for each semantic field, there are two covering terms that form a Covering Contrast Pair that defines the field. Note that each linguistic term has a set of semantic properties. If a covering term of a semantic field stands alone and, thus, has no contrast relation with any other term in the same field, then the semantic properties defining that field must be independently motivated. However, if there is a Covering Contrast Pair, then the defining semantic properties of the field can be defined by extracting common semantic attributes of the pair and need not be motivated independently. After the Covering Contrast Pair is established, it follows that other terms in the field will contrast with either of the covering terms, similar to Grandy's original formulation. In our definition, we will refer to this privileged contrast set as a **contrast pair**. Hence, there will be a unique contrast pair for each semantic field. Adopting this revised view of the structure of semantic fields, we will thoroughly examine the seven

[5] The definition of a Semantic Field according to Grandy [1992] is as follows:

(i) [**A semantic field**] is a set including one or more contrast sets and possibly also including permutation relations such that:

1. at most one covering term does not occur as an element of a contrast set in the semantic field;
2. except for the covering term, any expression that occurs in a contrast set with an element of the semantic field is also in the field.

In addition, we re-interpret Grandy's [1992] formal definition of a Contrast Set below:

(ii) A **contrast set** will consist minimally of a covering term T, a set of fundamental contrast relations, and a set of linguistic terms such that:

1. there are common linguistic beliefs that each linguistic term in the set is a kind of T (that is, the relation between any term and T can be defined by the is-a relation.);
2. for any two different terms in the set, it is a common linguistic belief that they contrast in terms of a single relation which is defined by the set or is derivable from the relations defined in the set.

contrast pairs in the following sections. Our prediction is that the other verbs, as members of the field, will behave like either covering term in the pair.

3. The representational distinctions

In this section, we will examine the contrasts between the two groups based on the five proposed criteria: their grammatical functions, their co-occurrence restrictions, their appropriateness in the imperative and evaluative construction, their verbal aspect and their transitivity.

3.1 Grammatical functions

Generally speaking, type A (i.e., *gao1xing4*) verbs are predominantly used as predicates while type B (i.e., *kuai4le4*) verbs are much more often used in their nominalized forms as arguments or nominal modifiers. In this section, we will give a qualitative account of such contrasts based on three different quantitative criteria: 1) First is the distribution of all the grammatical functions for each covering term, with the focus on the contrast between the predicative and nominal uses. This study will illustrate how distributional information underlies linguistic generalizations. 2) In addition, there is the ratio between each contrast pair for both nominal and predicative uses, which highlights the preferential status of the dominant term for each function. 3) Last is the distribution threshold demarcation between type A and type B verbs for the entire field, both of which attest to the universal validity of this functional motivation.

3.1.1 Distributional pattern of grammatical functions for the contrast pairs

For each of the seven contrast pairs, the same distributional pattern is found, as shown in Table 3. On one hand, type A verbs exhibits a very high tendency of being used as a predicates, i.e., no less than 76%; type B verbs show a much lower tendency, no more than 41%. On the other hand, type A verbs are seldom used in their nominalized forms, i.e., less than 3.07%; type B verbs are ten-times as likely to be used in their nominalized forms (with a distribution of no less than 26.43%.) Finally, type B verbs are four times as likely as type A verbs to serve as nominal modifiers, i.e., 14.21% to 3.73% on average.

Table 3. *Distribution of Grammatical Functions of the Seven Contrast Pairs*

Type A	Total	Pred.	Nom.	N.M.	Adjunct	Comp.	Else
gao1xing4 高興	669	**85.05%**	**0.30%**	**1.35%**	11.96%	1.35%	0.00%
nan2guo4 難過	232	**86.64%**	**2.16%**	**2.59%**	4.74%	3.88%	0.00%
shang1xin1 傷心	134	**76.12%**	**2.99%**	**11.19%**	5.97%	3.73%	0.00%
hou4hui3 後悔	102	**94.12%**	**0.00%**	**2.94%**	2.94%	0.00%	0.00%
sheng1qi4 生氣	271	**87.82%**	**0.00%**	**4.06%**	7.75%	0.37%	0.00%
hai4pa4 害怕	261	**93.10%**	**3.07%**	**2.68%**	1.15%	0.00%	0.00%
dan1xin1 擔心	609	**96.72%**	**1.97%**	**1.31%**	0.00%	0.00%	0.00%
Average	325	**88.51%**	**1.50%**	**3.73%**	4.93%	1.33%	0.00%

Type B	Total	Pred.	Nom.	N.M.	Adjunct	Comp.	Else
kuai4le4 快樂	942	37.79%	26.43%	24.84%	5.73%	5.20%	0.00%
tong4ku3 痛苦	443	25.73%	45.60%	20.54%	6.09%	2.03%	0.00%
bei1shang1 悲傷	52	40.38%	28.85%	19.23%	9.62%	1.92%	0.00%
yi2han4 遺憾[6]	198	34.85%	33.84%	3.54%	4.04%	0.00%	23.74%
fen4nu4 憤怒	112	28.57%	37.50%	17.86%	16.07%	0.00%	0.00%
kong3ju4 恐懼	149	23.49%	68.46%	7.38%	2.04%	0.00%	0.00%
fan2nao3 煩惱	199	24.12%	69.85%	6.03%	0.00%	0.00%	0.00%
Average	299	30.70%	44.36%	14.21%	6.23%	1.31%	3.39%

3.1.2 Likelihood ratio: measuring the encoding preference

The above data show the clear distributional disparity between type A and type B verbs in terms of predicative and nominal uses. However, in order to obtain a linguistically significant account and its associated implications from the distributional disparity, finer statistical contrasts must be utilized. In particular, from a functional point of view, the two near synonyms in a contrast pair are competing with each other to represent the sam e concept. In other words, when a certain grammatical function is expressed, the choice is between the two verbs of a contrast pair.[7] Given this functional perspective, the

[6] *Yi2han4* can also be used to express a speaker's judgement as shown in (i). In such cases, it functions as an evaluative adjunct.

(i) 這位 藝術家 的 作品， 很 遺憾地， 今年 無法 展出。
zhewei yishujia de zuopin hen yihandi jinnian wufa zhanchu
this artist 's works very regretfully this year couldn't exhibit
"It's regretful that the works of this artist couldn't be exhibited this year."

[7] In this study, the choice of the contrast pairs has important methodological considerations. As explained in our previous discussion, there are usually more than two terms in each semantic field. Hence, for either type A or type B, the covering term in the contrast pair may not be the only choice. However, we can see from the frequency statistics in Table 1 that a covering term has a frequency much higher than even the next most frequent term of the same type. Thus, our study uses a simplified model where only the two dominant terms in the contrast pair are compared, assuming that the less frequent terms cancel themselves out and do not contribute to significant differences if taken into consideration.

quantitative measurement that directly characterizes the contrast between type A and type B verbs is the ratio of how often one type is chosen over another type for a certain grammatical function. To obtain this significant measurement, we compiled two likelihood ratios for each pair with the dominant type as the dividee in either case: for the likelihood ratio in predicative uses, type A term frequency was measured against the type B term frequency and vice versa for nominal uses. For instance, the frequency of predicative use of type A *sheng1qi4* is 238 while that of its near synonymous contrasting term *fen4nu4* is 32. Hence, the type A verb Anger is 7.44 times (238/32) more likely to be chosen to express a predicative meaning than is its type B counterpart. On the other hand, for the same pair of verbs, the type B verb is 5.64 (62/11) times more likely to be chosen to express a nominal meaning than is its type A counterpart. Both likelihood ratios were calculated for each of the seven pairs and given below.

Table 4. *The Likelihood Ratio of Dominant Type over Non-dominant Type in terms of Predicative and Nominal Functions*

Type A/Type B verbs	Predicate Frequency	Ratio of A over B	Nominal Frequency	Ratio of B over A
gao1xing4 高興/ *kuai4le4* 快樂	569/356	1.59	11/483	43.91
nan2guo4 難過/ *tong4ku3* 痛苦	201/114	1.76	11/293	26.64
Shang1xin1 傷心/*bei1shang1* 悲傷	102/21	4.86	19/25	1.32
hou4hui3 後悔/*yi2han4* 遺憾	96/69	1.39	3/74	24.67
sheng1qi4 生氣/*fen4nu4* 憤怒	238/32	7.44	11/62	5.64
hai4pa4 害怕/*kong3ju4* 恐懼	243/35	6.94	15/113	7.53
dan1xin1 擔心/*fan2nao3* 煩惱	589/48	12.27	20/151	7.55
Average ratio		5.62		16.75

Summing up the statistics shown in Table 4, for each contrast pair, type A verbs are more likely to occur in a predicate context while type B verbs are more likely to occur in a nominal one. Even though the likelihood ratio varies from one pair to another, on average, type A verbs are chosen as predicates almost **six times** as often as type B verbs. On the other hand, type B verbs are chosen for nominal uses almost **seventeen times** as often as type A verbs. The above ratios reveal the most significant functional contrast between type A and B verbs: that type A verbs have a dominant predicative function while type B verbs have a predominant nominal function. Since similar bi-directional ratios exist for all seven contrast pairs, they also constitute strong support for a uniform functional motivation and eliminate any possibility that the distributions of grammatical functions are idiosyncratic.

3.1.3 Sorting two types of verbs: verification of the quantitative criterion

One last piece of statistical evidence we want to give for the type A and type B dichotomy

is that there can be an empirical demarcation between the two types of verbs. Following the discussion in the previous section, the two types of nominal uses, i.e., that of a nominalized event and that of a nominal modifier, will be merged. We will refer to the merged frequency as the quantitative index of "being deverbal." In addition, since what is studied here amounts to quantitative criteria used to determine whether a term is a type A or type B verb, all relavent verbs of emotion are taken into consideration.

Table 5. *Verbs of Emotion Sorted According to Deverbal Uses*

Type A Verbs	Nom.	N.M.	deverbal
*tong4kuai4*痛快	0.00%	0.00%	0.00%
*gao1xing4*高興	0.30%	1.35%	1.65%
*hou4hui3*後悔	0.98%	2.94%	2.94%
*dan1xin1*擔心	1.97%	1.31%	3.28%
*sheng1qi4*生氣	0.00%	3.58%	3.58%
*tong4xin1*痛心	2.08%	2.08%	4.17%
*nan2guo4*難過	2.16%	2.59%	4.75%
*hai4pa4*害怕	3.07%	2.68%	5.75%
*you1xin1*憂心	6.52%	0.00%	6.52%
*kai1xin1*開心	1.97%	5.92%	7.89%
*dan1you1*擔憂	9.38%	0.00%	9.38%
*shang1xin1*傷心	2.99%	11.19%	14.18%

Type B Verbs	Nom.	N.M.	deverbal
*qi4fen4*氣憤	20.41%	4.08%	24.49%
*wei4ju4*畏懼	22.50%	2.50%	25.00%
*yu2kuai4*愉快	7.75%	22.14%	29.89%
*huan1xi1*歡喜	21.50%	9.35%	30.84%
*kuai4huo2*快活	6.25%	27.08%	33.33%
*ju3sang4*沮喪	20.97%	12.90%	33.87%
*yi2han4*遺憾	33.84%	3.54%	37.38%
*ku3nao3*苦惱	35.56%	11.11%	46.67%
*bei1shang1*悲傷	28.85%	19.23%	48.08%
*chen1zhong4*沈重	0.00%	48.19%	48.19%
*kuai4le4*快樂	26.43%	24.84%	51.27%
*fen4nu4*憤怒	37.50%	17.86%	55.36%
*tong4ku3*痛苦	45.60%	20.54%	66.14%
*kong3ju4*恐懼	68.46%	7.38%	75.84%
*fan2nao3*煩惱	69.85%	6.03%	75.88%
*xi1yue4*喜悅	90.38%	1.92%	92.20%
*huan1le4*歡樂	31.91%	60.99%	92.91%

Table 5 shows clearly that the 29 verbs of emotion can be correctly classified as type A or type B according to the simple quantitative measurement of the frequency of their deverbal uses.[8] For example, in Table 5, we find that, without exception, all type A verbs have a nominal use frequency of 14.18% or lower, while all type B verbs have a nominal use frequency of 24.49% or higher. Between the two groups, there is an obvious gap in two crucial senses: First, the least frequent deverbal use of a type B verb is more than 10% higher than that of the most frequent type A verb. In other words, the distributions of the two groups of verbs are clearly discrete and not continuous. Second, the contrast in each contrast pair is even more prominent than the above gap. The smallest gap between a pair is 33.90% (*shang1xin1* vs. *bei1shang1*).

[8]There are two complimentary positions for nominal use: as either a referential complement or as a nominal modifier. Hence, we often find that when certain type B verbs show a low tendency to be used as referential complements, they necessarily show a higher tendency to be used as nominal modifiers, see, e.g., *chen2zhong4* 沈重 , *kuai4huo2* 快活 , and *yu2kuai4* 愉快 .

3.1.4 Summary

In this section, we will summarize the quantitative measurements we chose to characterize the distribution of the grammatical functions of the two types of verbs and their interpretations. First, we observed the overall distributions of all the represented grammatical functions of the seven contrast pairs and established that type A verbs are used predominantly for predicative uses, and that type B verbs are used predominantly for nominal uses. Second, the likelihood ratio measurements for both predicative and nominal uses were calculated for each contrast pair. This measurement was used to test the function-driven hypothesis that the two contrast pair members are candidates which compete to represent the same concept in any given context. The quantitative measurements were predicted based on the hypothesis and, thus, supported a functional account. Last, to verify that our quantitative measurements represented a true classificatory criterion instead of a random demarcation point in continuous distribution, we showed that the type A and type B verbs actually form two discrete groups separated by a significant gap according to their frequency of deverbal uses.

3.2 Selectional restrictions the verbs impose as adjuncts

The second important observation regarding the distribution of the two types of verbs of emotion is that, as adjuncts, they impose very different selectional restrictions on their heads. Type A verbs can only modify a very restricted set of nouns or verbs while type B verbs seem to be much freer.

In the Sinica Corpus, type A verbs, such as *gao1xing4*, can only modify six types of nouns, "time when" (e.g. *shi2hou4* 時候 / *shi2* 時), "event/story," "mood," "facial expressions," "person" and "utterance." In contrast, type B verbs, such as *kuai4le4*, can be adjuncts for many additional noun classes. The contrast is shown in (1) and (2).

(1) Type A

? 高興的 童年 /? 高興的 婚姻 /? 高興的 上班族 /? 高興的 環境
gaoxingde tongnian /gaoxingde hunyin /gaoxingde shangbanzu /gaoxingde huanjing
glad childhood /glad marriage/glad workers /glad environment

(2) Type B

快樂的 童年 / 快樂的 婚姻 / 快樂的 上班族 / 快樂的 環境
kuailede tongnian */kuailede* hunyin */kuailede* shangbanzu */kuailede* huanjing
happy childhood /happy marriage /happy workers /happy environment
"happy childhood/ happy marriage/ happy workers/ happy environment"

With regard to post-verbal adjuncts, both groups can modify transient activities, such as *wan2 de hen3 gao1xing4* 玩得很高興 "play happily" and *wan2 de hen3 kuai4le4* 玩得很快樂 "play happily." However, only type B verbs can be adjuncts of non-transient (state-like) activities, such *ashuo2 de kuai4le4* 活得快樂 "live happily," *guo4 de kuai4le4* 過得快樂 "live happily," and *ao2 de hen3 tong4ku3* 熬得很痛苦 "endure terribly."

3.3 The imperative and evaluative constructions

Some verbs of emotion are used in imperative sentences containing deontic modal verbs, as in (3). Many of them can also occur in evaluative sentences which contain the verb *zhi2de2* 值得 "be worthwhile (to)" or the phrase *mei2 she2me5 hao3 … de5* 沒什麼好…的 "be not worthwhile to," as in (4). In either case, they lose the prototypical "command" or "evaluation" meaning. Pragmatically speaking, both constructions with verbs of emotion have the same "dissuading" function.[9]

(3) 別 傷心 / 莫 傷心 / 不要 傷心。
bie shangxin /mo shangxin /buyao shangxin
don't sad /don't sad /don't sad
"Please don't feel sad."

(4) 不 值得 傷心 / 沒 什麼 好 傷心 的。
bu zhide shangxin /mei sheme hao shangxin de
NEG worth sad /without anything worth sad PARTICLE
"It is not worthwhile to feel sad. /There's nothing to be sad about.
(Please don't feel sad.)"

Based on the Sinica Corpus, we find that 1) all type A verbs appear in the imperative or the evaluative construction, and 2) with only one exception, (i.e., *fan2nao3*)[10], type B verbs do not appear in either type of construction, as shown in Table 6.

---

[9] In most cases, verbs of emotion which appear in evaluative constructions do not just express the speaker's judgement, but "dissuade" the listener from the stated emotion. Of course, the dissuading function of the imperative comes from the negative constructions, such as the negative imperative constructions or the negative evaluative constructions.

[10] We assume that the inherent properties of each emotion, such as the perceived degree of controllability, will affect the uses of each class of verb in these two constructions. Hence, it is more accurate to directly contrast the frequency of uses of verbs in the same field. Although *fan2nao3* has 7 occurrences in the two constructions, it is still a relatively small distribution when compared with the 80 occurrences of its contrast set counterpart *dan1xin1*.

Table 6. *Imperative and Evaluative Uses of the Seven Pairs*

Verb Types	A				B			
	Verbs	Imp.	Eva.	Total	Verbs	Imp.	Eva.	Total
Happy	*gao1xing4* 高興	6	6	12	*kuai4le4* 快樂	0	0	0
Depression	*nan2guo4* 難過	10	1	11	*tong4ku3* 痛苦	0	0	0
Sadness	*shang1xin1* 傷心	4	0	4	*bei1shang1* 悲傷	0	0	0
Regret	*hou4hui3* 後悔	3	0	3	*yi2han4* 遺憾	1	0	**1**
Anger	*sheng1qi4* 生氣	12	0	12	*fen4nu4* 憤怒	0	0	0
Fear	*hai4pa4* 害怕	9	0	9	*kong3ju4* 恐懼	0	0	0
Worry	*dan1xin1* 擔心	78	2	80	*fan2nao3* 煩惱	6	1	**7**

3.4 Verbal aspect or aktionsart

Verbs of emotion express mental states. They can represent either a homogeneous state, as in (5), or an inchoative state, as in (6).

(5) 他 爲 此 事 傷心 不已。
ta wei ci shi shangxin buyi
he for this matter sad continuous
"He has been sad about this for a long time."

(6) 他一 想起 妻子 已經 死 了，就 傷心 了 起來。
ta yi xiangqi qizi yijing si le jiu shangxin le qilai
he once think of wife already die LE then sad LE asp.
"He felt sad whenever the thought came into his mind that his wife had died."

The inchoative *le* can be used to differentiate between the two types of verbs.[11] We find in the Sinica Corpus that in each contrast pair, the particle *le* is associated with the type A verb much more frequently than it is associated with the type B verb, as shown in Table 7.

[11] Li & Thompson [1981], among others, characterized the sentential-final particle *le* as marking a new state, and *le* attached to a verb as marking the perfective aspect. However, when *le* co-occurs with a state verb, it always represents a change of the state (thus inchoative), regardless of its position.

Table 7. *Verbs of Emotional Association with the Sentential Final Particle le*

Verb Types	A		B	
	Verbs	Freq.	Verbs	Freq.
Happiness	*gao1xing4* 高興	20	*kuai4le4* 快樂	10
Depression	*nan2guo4* 難過	9	*tong4ku3* 痛苦	0
Sadness	*shang1xin1* 傷心	2	*bei1shang1* 悲傷	1
Regret	*hou4hui3* 後悔	7	*yi2han4* 遺憾	0
Anger	*sheng1qi4* 生氣	14	*fen4nu4* 憤怒	0
Fear	*hai4pa4* 害怕	5	*kong3ju4* 恐懼	2
Worry	*dan1xin1* 擔心	6	*fan2nao3* 煩惱	3

3.5 Transitivity

A verb of emotion takes either a cause or a goal as its direct object.[12] In the previous section, we showed that a verb of emotion can indicate an inchoative state. A new state does not come into being without a cause. Hence, a logical cause can be implied for each inchoative state. Grammatically, however, only the verbs of *Happiness*, *Fear* and *Worry* take a cause as their object.[13] More precisely, only type A verbs of *Happiness*, *Fear* an d *Worry* do so while none of the type B verbs take eventive Cause objects. This is demonstrated in (7) and (8), as well as in Table 8.

(7) 他們 很 高興 張三 沒 走。 [Tsai 1998]
tamen hen gaoxing zhangsan mei zou
they very glad John doesn't go
"They were glad that John didn't go."

---

[12] Adopting Teng's [1975] framework, goal in the Sinica Corpus and CKIP lexicon refers to both a transitivity goal (vs. patient) and a circumstantial goal (vs. source). In this paper, only a transitivity goal is considered. Please also note that since there is a theme but no patient in the CKIP argument role system, some of Teng's patient roles that do not qualify as themes are also classified as (transitivity) goals. See Lin [1992] for more details regarding the role classification system.

[13] For those verbs that do not take a cause event as a direct object, the cause event shows up in other positions, such as an adjunct PP (i) or a topic clause (ii).

(i) 爲了 這 件 事， 我 曾 傷心 了 好 久。
weile zhe jian shi wo ceng shangxin le hao jiu
for this piece matter I ever sad LE quite long time
"I've felt sad about this matter for quite a long time."

(ii) 母 子 竟 不得 見面， 怎麼 能 不 傷心 呢？
mu zi jing bude jianmian zenme neng bu shangxin ne
mother son dare couldn't meet how can not sad NE
"How can they not feel sad that the mother and son can't meet each other."

(8) * 他們 很 快樂 張三 沒 走。 [Tsai 1998]
tamen hen kuaile zhangsan mei zou
they very glad John doesn't go
"They were happy that John didn't go."

As for goal, only the verbs of *Angry*, *Afraid* and *Worried* semantically take this kind of argument and, thus, syntactically take them as direct objects. However, in the Sinica Corpus, only Group A verbs of those types can take a goal as a direct object while Group B verbs as a rule do not take a goal as a direct object, as shown in Table 8.

Table 8. *The Transitive Uses of Four Representative Pairs*

Type A	-Cause Event		-Goal	Type B	-Cause Event		-Goal
	-VP/S	-Event N	-Simple N		-VP/S	-event N	-Simple N
gao1xing4 高興	69	3	0	*kuai4le4* 快樂	0	0	0
sheng1qi4 生氣	0	0	0/12[14]	*fen4nu4* 憤怒	0	0	0
hai4pa4 害怕	68	9	8	*kong3ju4* 恐懼	3	0	2
dan1xin1 擔心	285	17	35	*fan2nao3* 煩惱	2	0	2

4. Semantic explanation

In this section, we will first summarize the contrasts and then propose a lexical semantic explanation for all the contrasts.

4.1 The syntactic contrasts

In the previous section, we presented the syntactic basis for our bipartite classification of the verbs of emotion. There are five distributional syntactic criteria. The two groups differ in terms of tendency. In each of the five schemes of grammatical representations, one of the two types of verbs dominate, as shown below:

Type A verbs:

1. function mostly as predicates and are seldom used deverbally;
2. have strict selectional restrictions on the head when they function as adjuncts;
3. can appear in imperative or evaluative constructions;

[14] Whether *sheng1qi4* can be a transitive verb or not depends on one's definition of transitivity in Chinese. This is because its goal can only be inserted into the so-called 'possessive object' position and never into a canonical object position, such as *shang1 ta1 de5 qi4* 生他的氣 [Huang 1990].

4. can represent inchoative states; and

5. can take causes or goals as their direct objects.

Type B verbs:

1. are the predominant choice in a deverbal context;

2. have looser selectional restrictions on the head when they function as adjuncts;

3. are seldom used in imperative or evaluative constructions;

4. rarely represent inchoative state; and

5. seldom take causes or goals as their direct objects.

4.2 The semantic basis for the bipartite classification

Bear in mind that the 14 verbs we are studying here form seven contrast pairs. While each pair represents a different semantic field, they all belong to the same subordinating semantic field of emotion. If the same five contrasts differentiate all seven pairs, we may assume that there is a fundamental semantic motivation underlying all these contrasts. This semantic motivation may be a design feature of the field of emotion. It is highly unlikely that these exact five contrasts are independently motivated for each contrast pair and yet are identically represented in each of the fields.

It is also important to note that the members of each contrast pair differ minimally in semantics and are mutually substitutable in many contexts. In other words, it is natural for a null hypothesis account to assume that the minimally contrasting feature contributes to the grammatical contrasts. We can understand the behavioral contrasts we have observed better by rephrasing the question as follows in (9):

(9) Why are type A verbs chosen over type B verbs (and vice versa) in construction X?

From a functional perspective on language, this question helps us to directly look for motivations to differentiate between the two types of verbs. Since the contrasts exist regularly across the seven sub-fields of emotion, we expect the motivation to be semantic in nature, and expect that it may involve the fundamental semantic dichotomy of the semantic field of emotion. In anticipation of this interpretation, we summarize and re-organize the contrasts between the two groups as follows:

Table 9. *Contrasts in Linguistic Distribution*

Linguistic Instantiation	TypeA	Type B
Predicative use	**Strong**	Weak
Inchoative states	**Predominate**	Rare
Transitivity	**Strong**	Weak
Imperative or evaluative constructions	**Predominate**	Rare
Adjuncts to non-transient activities	Rare	**Predominate**
Adjuncts to nouns	Weak	**Strong**
Nominalization	Rare	**Predominate**

From the above contrast, we generalize that all the distinctive linguistic instantiations are related to event structure properties. Generally speaking, type A verbs are preferred for indicate **transition** while type B verbs are preferred for **homogeneity**. In particular, when we want to indicate a change of state, such as with the change-of-state *le*, type A verbs are usually used. When an object or cause is present, the event focus naturally shifts to the transition to a new state, and again, type A verbs are preferred. When dissuasion is intended and, thus, the potential for transition is involved, type A verbs are usually used.

On the other hand, type B verbs are preferred for indicate continuous and homogeneous states. This is why only type B verbs are used to modify non-transient verbs and to ascribe attributes to nouns. This is also why type B verbs are preferred as deverbal nouns since a referential entity is regarded as a wholistic unit and, thus, homogeneous composition is implied.

5. A semantic interpretation of the preferred sub-lexical structure

In this section, we will explore and explain the close relationship between the sub-lexical structure of these compound verbs and their bipartite classification.

An interesting observation involving the current set of data is that the distinctions among the internal structures of these compounds seem to correspond to the distinctions between the two groups. We find that 14 of 16 type B verbs are VV compounds while none of the 13 type A verbs are VV compounds, as shown in Table 10.

Type A: *gao1xing4* 高興 (non-VV), *nan2guo4* 難過 (non-VV) , *hou4hui3* 後悔 (non-VV),*tong4ku3* 痛快 (non-VV), *dan1you1* 擔憂 (non-VV),*sheng1qi4* 生氣 (non-VV), *chi1jin1* 吃驚 (non-VV), *dan1xin1* 擔心 (non-VV), *shang1xin1* 傷心 (non-VV) , *kai1xin1* 開心 (non-VV) , *you1xin* 憂心 (non-VV) , *tong4xin1* 痛心 (non-VV), *hai4pa4* 害怕 (non-VV)

Typep B: *kuai4le4* 快樂 (VV), *xi3yue4* 喜悅 (VV), *huan1le4* 歡樂 (VV), *fan2nao3* 煩惱 (VV), *kong3ju4* 恐懼 (VV), *tong4ku3* 痛苦 (VV), *fen4nu4* 憤怒 (VV), *chen2zhong4* 沈重 (VV), *bei1shang1* 悲傷 (VV), *ku3nao3* 苦惱 (VV), *yi2han4* 遺憾 (AN or VO)[15], *ju3sang4* 沮喪 (VV), *kuai4huo2* 快活 (VV or AV)[16], *huan1xi3* 歡喜 (VV), *yu2kuai4* 愉快 (VV), *wei4ju4* 畏懼 (VV)

We have shown that all VV compounds examined here belong to type B. Based on our lexical semantic account offered earlier, this means that VV compounds are preferred to represent homogenous states. We argue that this fact is due to the semantic properties of VV compounds.

VV compounds differ from other compounds (such as SV, VO, AV and VR) in the compounding process. In contrast to the other major verbal compounding processes, VV compounds are double-headed. In all the other constructions, the V root employs one more constituent to elaborate on the event so as to make it either more complete or more specific. For instance, in a SV compound, the subject is added to the event structure; in a VO compound, an object is incorporated into the event structure; in an AV compound, the manner of executing an action is described; and lastly, in a VR compound, the result of the action is explicitly indicated. However, a VV compound does not elaborate. In VV compounding, the concept of an event is "diffused" because two similar events are juxtaposed so as to suggest extraction of the common properties of the pair. It is a common morpho-lexical strategy in Mandarin to concatenate two antonyms or synonyms to form the concept of "kind" or "property." For example, the word *hu1xi1* 呼吸 "to breathe" is the juxtaposition of *hu1* "exhale" and *xi1* "inhale" while *da4xiao3* 大小 "size" is the juxtaposition of *da4* 'big' and *xiao3* 'small.'

Since the concept of an event is diffused or lifted to "kind/property," it is natural for VV compounds to be used to indicate a homogeneous state, but it is difficult to use them to indicate an inchoative state. That is why Mandarin employs VV compounds to indicate more referential contexts, such as a nominalized event or a nominal modifier. It is also a natural consequence that the VV verbs of emotion are seldom used in imperative and evaluative constructions since in both constructions transitional characteristics are highlighted, which is contrary to the nature of a VV compound.

[15] *yi2han4* could be viewed not only as a VO compound verb, but also an AN compound noun because it can be interpreted as an abbreviation of the idiom: *yi2zhu1zhi5han4* 遺珠之憾 "the regret of missing one pearl," and thus be realized as a noun. If this is true, then *yi2han4* was originally a noun. As a verb of emotion, it is a denominal verb formed through abbreviation.

[16] The inner structure of *kuai4huo2* is hard to determine. It might be VV (happy and vivid) or AV (happily live).

6. Conclusion

In this paper, we have illustrated consistent grammatical and distributional contrasts in seven types of verbs of emotion and proposed a semantic explanation of the contrasts. The homogeneous state/transition semantic dichotomy is found to be inherent in the semantic field of emotion. It is natural to suggest, however, that this dichotomy may show up in all semantic fields of states. Our re-interpretation of the structure of semantic fields, where a **contrast pair** (instead of a single cover term) heads a semantic field, has several theoretical implications. First, we need to find out if all semantic fields are actually headed by contrast pairs that are defined by a primary contrast relation. Second, we need to exhaustively list all primary contrast relations and try to develop a theory of classification of semantic fields based on them. Third, it will be important to see if theoretical constraints are placed on the primary contrast relations. For instance, the transition/homogeneous state contrast is clearly central to the event type definition of states.

In addition, we have observed that all VV compounds belong to type B, and we have proposed a morpho-semantic explanation for their distribution. VV compounds undergo a process that involves merging two individual events to create a superset of properties covering both events; hence, it is a more appropriate morpho-lexical process for representing homogeneous states.

In research on regularities between lexical meaning and syntactic behaviors, it is very important to distinguish between the constructional meaning and the core meaning. The explanation we have offered above suggests that the regularities we have extracted from VV compounds in the semantic field of emotion exist in all Mandarin VV compounds. A preliminary study on the Sinica Corpus does confirm that all VV compounds have a higher tendency of being nominalized. In addition to our continuing research on the formal properties of semantic fields, we are also looking into the process whereby morpho-lexical structures encode constructional meanings.

References

Chinese Knowledge Information Processing (CKIP). 1995. *An Introduction to the Sinica Corpus.* [In Chinese]. CKIP Technical Report 95-02. Taipei: Academia Sinica.

Grandy, Richard E. 1992. "Semantic Fields, Prototypes, and the Lexicon." In Lehrer and Kittay Eds. *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization.* Pp.103-122. Hillsdale: Lawrence Erlbaum.

Huang, Chu-Ren, 1998, "Classifying Event Structure Attributes: A Verbal Semantic Perspective from Chinese." Chinese Workshop. The 1998 International Lexical-Functional Grammar Conference. June 30-July3. Brisbane, Australia.

Huang, Chu-Ren. 1990. "A Unification-based LFG Analysis of Lexical Discontinuity." *Linguistics*. 28.263-307.

Huang, Chu-Ren, Mei-chun Liu, and Mei-chih Tsai, 1998, "From Lexical Meaning to Event Structure Attributes: Across Semantic Classes of Mandarin Verbs." The 6th International Conference on Chinese Linguistics/The 10th North American Conference on Chinese Linguistics. June 26-28. Stanford.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.

Li, Charles, and Sandra Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley, Los Angeles: University of California Press.

Lin, Fu-Wen. 1992. *Some Reflections on the Thematic System of Information-based Case Grammar (ICG)*. [In Chinese.] CKIP Technical Report No. 92-01. Nankang: Academia Sinica.

Liu, Mei-chun. 1999. "Lexical Meaning and Discourse Patterning - the three Mandarin cases of 'build'." In Barbara A. Fox, Dan Jurafsky and Laura A. Michaelis, eds. *Cognition and Function in Language*. 181-199. Stanford: Center for the Study of Language and Information (CSLI).

Liu Mei-chun, Chu-Ren Huang, and Charles C.L. Lee. 2000. "When Endpoint Meets Endpoint: A Corpus-based Lexical Semantic Study of Mandarin Verbs of *Throwing*." *Computational Linguistics and Chinese Language Processing*. 5.1:81-96.

Pustejovsky, James, S. Bergler, and P. Anick. 1993. "Lexical Semantic Techniques for Corpus Analysis." *Computational Linguistics*. 19.2:331-358.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge: MIT Press.

Teng, Shou-hsin. 1975. *A Semantic Study of Transitivity Relations in Chinese*. Berkeley and Los Angeles: University of California Press.

Tsai, Mei-chi, Chu-Ren Huang, Keh-jiann Chen, and Kathleen Ahrens. 1998. "Towards a Representation of Verbal Semantics: An Approach Based on Near-Synonyms," *Computational Linguistics and Chinese Language Processing*. 3.1:62-74.

Computational Linguistics and Chinese Language Processing
Vol. 5, No. 1, February 2000, pp. 81-96

© Computational Linguistics Society of R.O.C.

When Endpoint Meets Endpoint: A Corpus-based Lexical Semantic Study of Mandarin Verbs of *Throwing*

Mei-Chun Liu*, Chu-Ren Huang+, Charles Lee*, Ching-Yi Lee**

ABSTRACT

Since verbal semantics began to receive much attention in linguistics research, many interesting findings have been presented regarding the semantic structure or meaning contrasts in the lexicon of Chinese [cf. Tsai, Huang & Chen, 1996; Tsai *et al,* 1997; Liu, 1999, etc]. Adopting a corpus-based approach, this paper aims to further study and fine-tune Mandarin verbal semantics by exploring the lexical information specific to verbs of throwing, with four pivotal near-synonomous members: *TOU* (投), *ZHI* (擲), *DIU* (丟), *RENG* (扔). To account for their semantic differences, two kinds of 'endpoints' are distinguished: the *Path-endpoint* (i.e., the Goal role) vs. the *Event-endpoint* (i.e., the resultative state). These two variables are crucial for cross-categorizing the four verbs. Although the verbs all describe a directed motion with a Path in their event structure, they differ in their lexical specifications on participant roles and aspectual composition. *TOU* and *ZHI* have a specified Path-endpoint while *DIU* and *RENG* do not specify a Path-endpoint. Moreover, *TOU* and *ZHI* can be further contrasted in terms of the spatial character of the Path-endpoint they take: *TOU* selects a spatially bounded Path-endpoint while that of *ZHI* is unspecified in this regard, as manifested by the fact that *TOU* collocates most frequently with a CONTAINER-introducing locative. On the other hand, *DIU* and *RENG* can be further differentiated in terms of event composition: only *DIU*, not *RENG*, allows an aspectual focus on the endpoint of the event contour (the Event-endpoint) since it manifests a resultative use. The observed distinctions are then incorporated into a representational paradigm called the Module-Attribute Representation of Verbal Semantics (MARVS), proposed in Huang & Ahrens [1999].

* Department of Foreign Languages and Literature, National Chiao Tung University, Hsinchu, Taiwan, R.O.C. E-mail: mliu@cc.nctu.edu.tw
+ Institute of Information Science, Academia Sinica, Taiwan, R.O.C.
** Institute of Linguistics, National TsingHua University, Hsinchu, Taiwan, R.O.C.

Finally, conclusions are drawn as to the most effective approach to lexical semantic study of Mandarin as well as theoretical implications in general.

Keywords: Mandarin verbs, Lexical semantics, Verbs of throwing, Event-endpoint, Path-endpoint

1. Introduction

This work presents a corpus-based approach to the lexical semantic study of a particular class of Mandarin verbs - verbs of *throwing*. In order to account for the observed differences in use patterns among the verbs, the notion of 'event focus' with its implication on 'event-structure attributes' is introduced in this paper. It aims to show that a semantically-constrained framework of event structure is needed to make sense of the crucial distributional facts in lexical differentiation.

1.1 Verbal Semantics

A recent focus of linguistic studies has been lexical semantics, especially verb meanings. Being the most essential part of the lexicon, verbs provide the key to studying the nature of lexical knowledge as well as sentence processing. Most lexical semantic studies on verbs share a common assumption that *the syntactic behavior of a verb, especially its argument expression, is determined by the meaning of the verb* [Pustejovsky 1995, Levin 1993, Atkins and Levin 1991, Atkins *et al.* 1988, etc.]. However, two issues still need to be further explored: 1) what exactly makes up verbal semantics? 2) how exactly can the differences in argument expression be attributed to lexical semantic features? Instead of looking for alternation patterns that are class-dependent, this study focuses more on corpus-based morpho-syntactic behavior as an indicator of lexical-semantic differences.

From the perspective of Chinese linguistics, previous studies on the Mandarin verb system have attempted to categorize verbs into classes with respect to general semantic types [e.g. 'active' vs. 'stative', Chao 1968], argument structure [Her 1990, Tsao 1996], or a hybrid of event types and thematic roles [CKIP 1988]. Given the typological and parametric variations among languages, some of the frameworks used for English cannot be readily transferred to Chinese. Liu [1996b] found that a purely alternation-based approach, such as that of Levin 1993, may not be adequate for categorizing and describing Mandarin verbs. A more **semantically constrained** system is indeed needed for natural language processing purposes. This study, thus, aims to provide detailed analysis of finer semantic distinctions as preparation for a complete representation of Mandarin verbal meanings.

1.2 Corpus-based Study of Near-Synonyms

In response to the need of fine-tuning verbal semantics, Tsai, Huang, and Chen [1996] presented an interesting work on differentiating a pair of near-synonyms - *gaoxing* (高興) 'happy, glad' and *kuaile* (快樂) 'happy, joyful'. These two verbs are semantically similar but syntactically distinct in many respects. By examining the correlation between their syntactic behaviors and lexical semantic properties, Tsai et al. showed that the syntactic contrasts can be systematically explained with two semantic features: <±control> and <± change-of-state>. The same account can also be extended to the semantic distinction of near-synonym pairs in English and French.

Adopting a similar approach, Liu [1999] examined another interesting set of near-synonymous verbs - *jian* (建), *gai* (蓋), and *zao* (造), roughly glossed as 'to build'. The three verbs are supposed to be prototypical transitive verbs involving creation of physical entities, but corpus data show that they have very little in common. Their distinct morphosyntactic behaviors provide revealing indications of their distinct lexical properties. Framing their differences based mainly on a cognitive-semantic perspective, the study showed that verbs may share the same cognitive schema but profile different event focus, incorporate various degrees of object specification, and map onto varying constructional frames due to distinct event structures and argument saliency.

As part of a long-term project on the lexical semantic study of Mandarin verbs, the present work extends the research frontier to a new semantic field with four contrastive near-synonyms - *TOU* (投), *ZHI* (擲), *DIU* (丟), and *RENG* (扔), all glossed as 'to throw'. It is believed that only a comprehensive corpus-based study on these verbs can render significant contrasts that help to differentiate their unique meanings.

1.3 Focus of the Paper

The four verbs of throwing are generally viewed as belonging to the same semantic field [Grandy 1992], representing prototypical transitive verbs that 'instantaneously cause ballistic movement by imparting a force' [Levin 1993]. However, as near-synonyms, they are bound to involve certain contrast sets [Grandy 1992], and the verbs have not been adequately examined in terms of their contrastive semantic properties. Adopting a goal similar to that of some lexicographers as well as linguists [e.g., Levin, 1993; Atkins and Levin, 1991; Atkins *et al.*, 1988], this study attempts to establish semantic-syntactic interdependences by observing the morphosyntactic behaviors of the verbs displayed in a large corpus. Their distributional patterns in the corpus help reveal the semantic features inherent in their meanings. For the four verbs of throwing, except for their common transitive use, they display quite different association patterns: *TOU* and *ZHI* form a

subgroup and differ from *DIU* and *RENG* in at least two respects:

- Both *TOU* and *ZHI* may take a Goal as the direct object, but *DIU* and *RENG* do not (e.g. *TOU- lan* (投籃) 'to shoot the basket', *TOU-hu* (投湖) 'to throw (oneself) into the lake'; *ZHI- di-you-sheng* (擲地有聲) 'to throw (something) to the ground with a thump').
- *DIU* and *RENG* form typical V-V compounds with V1 (Manner) or V2 (Result)[1], while *TOU* and *ZHI* do not seem to form these compounds (e.g. *luan DIU/RENG* (亂丟 / 扔) 'to recklessly throw (something)', *DIU/RENG diao* (丟 / 扔掉) 'to throw away).

Moreover, further contrast can be found within the same group. Although both specify a Path-endpoint, *TOU* selects a *spatially bounded* Path-endpoint, but *ZHI* does not. This is evident from the fact that when occurring with a locative phrase, about 76% of the occurrences of *TOU* take *ru* (入) or *jin* (進) 'into' as the locative; that is, *TOU* collocates most frequently with a CONTAINER-goal while the majority (87%) of the occurrences of *ZHI* is followed by *xiang/chou/wang* (向 / 朝 / 往) 'toward', which indicates that the path of *ZHI* is not specified for spatial boundedness. As for the other pair, *DIU* and *RENG* can be further differentiated based on their aspectual specifications: *DIU* may be used to describe the endpoint of an event, i.e., the resultative state of *DIU* , while *RENG* does not have a stative use. The observed distinctions are then represented from the viewpoint of a recently proposed framework that takes event-structure attributes as the primary defining mechanisms for lexical semantic contrasts [Huang and Ahrens 1999, Huang, Liu and Tsai 1999]. It is through the characterization of eventive information that the verbs studied here can be best differentiated (details in Section 4).

1.4 The Data

The data for the analysis presented in this paper come from a Mandarin corpus, the Sinica Corpus, which is the largest balanced corpus of both written and spoken contemporary Mandarin, containing a total of 5 million words and developed by the CKIP group at Academia Sinica, Taiwan. The relevant data were extracted from the corpus by means of a key-word search with 30 additional words on either side. The total number of occurrences of each verb follows:

---

[1] While the grammatical category of the elements expressing Manner and Result may be controversial, we take them as verbs here, assuming that the issue of their grammatical status may not be crucial to the argument.

TOU: 556 *ZHI*: 303
DIU: 268 *RENG*: 77

Following the above background introduction, section 2 in this paper outlines the preliminary contrast that exists among the four verbs. Section 3 then details their distributional differences. Section 4 establishes a systematic representation of the semantic distinctions. Finally, section 5 concludes with a discussion of the significance of this work.

2. Preliminary Observation: *TOU* vs. *DIU*

As members of the near-synonym set pertaining to the action of 'throwing,' the four verbs *TOU, ZHI, DIU,* and *RENG* display quite different morpho-syntactic patterns, despite their semantic class membership. Conceptually and theoretically, each group of near synonyms constitutes a contrast set that is a component of a semantic field [Grandy 1992]. The purpose of comparing their behavior is, then, to locate the linguistic relation that defines the contrast.

2.1 Interpretational Distinction between *TOU* and *DIU*:

By encoding a ballistic movement, the four verbs can potentially be associated with a Path contour which ideally contains a start-point, a trajector, and an endpoint [cf. Lakoff 1987]. The major difference among the verbs lies exactly in their inherent specification of the Path: they highlight various facets of the path. Our initial observation starts with the different interpretations that *TOU* vs. *DIU* may render when followed by the same object-theme, forming a V-O compound. As shown in (1) below, *TOU-qiu* (投球) and *TOU-piao* (投票) may differ completely from *DIU- qiu* (丢球)/*DIU-pia* (丢票) in terms of manner and directionality:

(1) Interpretational Differences between *TOU-QIU* (投球) and *DIU-QIU* (丢球):

	MANNER	DIRECTIONALITY
TOU-QIU	carefully targeting	toward a single and precise direction
DIU-QIU	randomly throwing	no specific direction

2.2 Distinction in Path-Endpoint

The second observation concerns the semantic role of the direct object following *TOU* or *DIU*, which is termed the Path-endpoint. By *Path-endpoint*, we refer specifically to the semantic role generally and loosely termed the Goal, which marks the final point of a trajectory inherent in a directed motion [cf. the case study of English 'over' discussed by Lakoff (1987)]. There are two sets of evidence that show that *TOU* is lexically specified with a Path-endpoint. First, in term of compounding, examples in (2) below illustrate that only *TOU* may take a Path-endpoint as its direct object, not *DIU*:

(2) *TOU* with Path-endpoint:

a. *tou-lan* 投籃	'to shoot a basket'
tou-hu-zhi-jin 投湖自盡	'to throw oneself into a lake'
tou-gong *投共*	'to defect to Communist China'
tou-qi-suo-hao (投其所好)	'to please someone by satisfying his wishes'
b. ***diu-lan*** 丟籃	'to shoot a basket'

The possible compounding of *TOU* with a Path-endpoint indicates that the final point and the direction of the motion plays a more salient and central role in the meaning of *TOU* than in that of *DIU*. The verb *DIU*, on the other hand, is typically modified by manner adverbs or resultatives that highlight the lack of directionality:

(3) Typical Manner-modifier or Resultative-Complement with *DIU*:

a. luan-diu 亂丟 'to mindlessly throw (something somewhere)'

b. diu-diao 丟掉 'to throw (something) away'

Secondly, while Path-endpoint is not marked in the case of *TOU, DIU* tends to take an overt marker introducing a Path-endpoint. *DIU* occurs far more often than TOU (43% vs. 26%) with an additional locative marker (e.g. *ru* (入) 'into,' *xiang* (向) or *wang* (往) 'toward,' *zai* (在) 'at,' *dao* (到) 'to'), thus overtly introducing a Path-endpoint. In other words, if the Path-endpoint in the event of *DIU* is expressed, it tends to be overtly marked with a locative phrase:

(4) Overt Marking of Path-endpoint:

	Occurrence with Post-verbal Locatives
TOU	26% (147 out of 556)
DIU	43% (116 out of 268)

2.3 Tentative Hypothesis

From the above discussion, we may conclude with a tentative hypothesis that *TOU* and *DIU* differ in their lexical specification of a Path-endpoint; that is, *TOU* is inherently specified with a Path-endpoint, but *DIU* is unspecified in that regard.

(5) Major Distinction between *TOU* and *DIU*:

TOU-verbs	Path-Endpoint Specified (+ Path-endpoint)
DIU-verbs	Path-Endpoint Unspecified (- path-endpoint)

In the next section, we will group the other two verbs, *ZHI* and *RENG*, according to the behavior of *TOU* vs. *DIU*.

3. Observation on *ZHI* and *RENG*

Having laid out the major difference between *TOU* and *DIU*, we may proceed to examine the other two verbs: *ZHI* and *RENG*. Basically, it is found that *ZHI* is similar to *TOU* while *RENG* is similar to *DIU*.

3.1 Properties Shared by *ZHI* and *TOU*

Like *TOU*, *ZHI* may also take a Path-endpoint as its direct object:

(6) *ZHI* with a Path-Endpoint:
a. *zhi-di-you-sheng* 擲地有聲
ZHI-ground-have-sound
'throwing (something) to the ground with a thump'
b. *leiqiu-zhi-yuan* 壘球擲遠
softball-ZHI-distant place
'softball-throwing'

In view of the fact that in the case of *ZHI*, the path-endpoint can also serve as the direct object, we assume that *ZHI* can be paired with *TOU* as they both take a Path-endpoint as an essential participant role. As verbs of directed motion, both *TOU* and *ZHI* are inherently specified not just with a trajectory-path, but more specifically, with a Path-endpoint.

3.2 Properties Shared by *DIU* and *RENG*

On the other hand, the verb *RENG* behaves more like *DIU* since *RENG* cannot be compounded with a Path-endpoint but may form a typical V-V compound with V1

(Manner) or V2 (Result), which implies a lack of directionality:

(7) *RENG* with modifiers that lack directionality:
a. luan-reng 亂扔 'to mindlessly throw (something) in all directions'
b. reng-diao 扔掉 'to throw (something) away'.

Furthermore, when followed by a locative, *TOU/ZHI* occur predominantly with *ru/jin/xiang/chao/wang* (入 / 進 / 向 / 朝 / 往), which are strongly direction-oriented, but *DIU/RENG* occur more commonly with *zai/dao* 在 / 到 , which are less specific in directionality. As shown in the highlighted portions in (8) below, taken together, over 90% of the uses of *TOU/ZHI* take a directional locative:

(8) Locative Markers Prefacing the Path in *TOU/ZHI* vs. that of *DIU/RENG*

	Directional Locatives *ru/jin/xiang/chao/wang* (入, 進, 向, 朝, 往)	Non-directional Locatives *zai/yu/dao* (在, 於, 到)
TOU	**89%**	10%
ZHI	**94%**	6%
DIU	42%	**58%**
RENG	43%	**57%**

Therefore, summing up the above discussion, we conclude that *ZHI* belongs to the *TOU*-group since both are [+ Path-endpoint]; *RENG* belongs to the *DIU*-group since both are [- Path-endpoint].

(9) Tentative Conclusion: *TOU/ZHI* vs. *DIU/RENG*

TOU/ZHI	Path-endpoint specified, strongly directional
DIU/RENG	Path-endpoint unspecified, non-directional

Having discussed the shared properties for the two groups of verbs, we will proceed to indicate the finer distinctions between the verbs in the same group.

3.3 Fine Distinctions between *TOU* and *ZHI*

When taking into consideration the spatial character of the Path-endpoint, we find that *TOU* and *ZHI* are associated with different locative markers that characterize different spatial boundaries of the Path-endpoint. In the corpus, we find that *TOU* occurs predominantly (76%) with a container-introducing locative, *ru* (入) or *jin* (進) 'into,' which manifests a bounded, container-type Path-endpoint. The verb ZHI occurs predominantly (87%) with *xiang* (向), *chao* (朝) or *wang* (往), all meaning 'toward,' which simply indicates a directed path with no further specification of the shape of the

endpoint, as shown in (10):

(10) Locative Markers Typically Following *TOU* vs. *ZHI*

	ru/jin (入/進) 'into'	*xiang/chao/wang* (向/朝/往) 'toward'	*zai/yu* (在/於) 'at'	*dao* (到) 'to'
TOU	**76%**	13%	10%	1%
ZHI	6%	**87%**	6%	0%

Another interesting difference between *TOU* and *ZHI* is that *ZHI* often occurs as the second verb in a cognate V-V compound, indicating that the event of *ZHI* is categorially less-marked and lexically less-specified with manner (since the first verb in the cognate V-V compound is more manner-specific), as shown in (11):

(11) *ZHI* as the default V in cognate V-V compounds all meaning 'to throw':

a. *tou-zhi* 投擲
b. *reng-zhi* 扔擲
c. *diu-zhi* 丟擲
d. *pao-zhi* 拋擲

The above observation concerning the morpho-syntactic differences between *TOU* and *ZHI* seems to point to a finer distinction: *TOU* is semantically more loaded, with a further specification of the spatial boundedness of its Path-endpoint, while *ZHI* is lexically less informative, as summarized in (12):

(12) Tentative Conclusion (Lexical-semantic Distinction between *TOU* and *ZHI*):

TOU	+ Path-endpoint; + Spatially bounded
ZHI	+ Path-endpoint

3.4 Fine Distinctions between *DIU* and *RENG*

Although both *DIU* and *RENG* are not lexically specified with a Path-endpoint, they differ significantly in another respect, i.e., the coding of an Event-endpoint. By *Event-endpoint*, we refer to the final state resultative of a given activity-event. The most salient difference in their use patterns is that *DIU*, but not *RENG*, displays a causative-intransitive use, which profiles the endpoint of the event, a resultative state:

(13) Causative-intransitive Use of *DIU*

w*ode gangbi diu/*reng le.* 我的鋼筆丟 /* 扔了 .

'My pen is lost.' (= 我的鋼筆掉了 .)

The possible inclusion of an Event-endpoint in the use of *DIU* gives rise to the potential ambiguity of (14a):

(14) Interpretational Differences:

a. *wo diu le yi-zhi gangbi.* 我丟了一支鋼筆 .

■ 'lost' (inchoative, stative, +result, -control)[2]

■ 'thrown away' (completive, active, -result, +control)

b. *wo reng le yi-zhi gangbi.* 我扔了一支鋼筆 .

■ 'thrown away' (completive, active, -result, +control)

Given its stative use, the verb *DIU* may occcur as the resultative complement in a Verb-Resultative compound:

(15) *wode gangbi gao-DIU/*gao-RENG le.* 我的鋼筆搞丟 /* 搞扔了 .

'My pen got lost.'

We see that *DIU* is polysemic with two meaning facets. Besides its use as an activity verb, it can also be used as an achievement verb. The main reason is that *DIU* lexically specifies an Event-endpoint, thus allowing the focus to be on the ending state of the event. We now draw the conclusion that *DIU* differs from *RENG* in that it allows aspectual emphasis to be placed on the Event-endpoint:

(16) Distinction between *DIU* and *RENG*

DIU	+ Event-endpoint
RENG	- Event-endpoint

So far we have mentioned two types of endpoints: *Path-endpoint* vs. *Event-endpoint*. *Path-endpoint* marks the final point of a trajectory-path in ballistic motion, which coincides with the semantic role Goal. *Event-endpoint,* on the other hand, is relevant to the final point of an event contour, usually indicating a resultative state.

---

[2] Following Smith [1991], the difference between *inchoative* and *completive* is mainly aspectual: *inchoative* refers to a change of state or the starting point of a new event; *completive* describes an event as it is completed. The stative vs. active distinction concerns kenesis in general, as explained by Chao [1968]. The feature *control* concerns volitionality of the subject.

These two types of endpoint are crucial for fine-tuning the lexical semantics of the four verbs studied here.

3.5 Distinctions Based on the Two Types of Endpoint

As a near-synonym set, the four verbs *TOU , ZHI, DIU, RENG* demonstrate a two-way contrast in terms of their specification of Path-endpoint and Event-endpoint:

(17) The Distinction based on Path-Endpoint vs. Event-Endpoint

	TOU	*ZHI*	*DIU*	*RENG*
Path-Endpoint	+; bounded	+	-	-
Event-Endpoint	-	-	+	-

It is clear from (17) that while both *TOU* and *ZHI* is lexically specified with a Path-endpoint, only *TOU* requires a *spatially bounded* path-endpoint. As for *DIU* and *RENG*, their lexical meanings are not sensitive to the encoding of a Path-endpoint; instead, they can be further distinguished in terms of their lexical specification of an Event-endpoint.

4. Verbal Semantics as Eventive Information

The observed differences as outlined in (17) above can be viewed from a more general perspective proposed in Huang and Ahrens [1999], in which verb meanings are described in terms of structural and attributive distinctions. They argue that all grammatical information is encoded in the lexicon, and that verbs express eventive information. Each verbal sense is then taken to be a unique event structure (see 4.2 below for details). The framework makes use of the concept of an event *focus* to identify different event types, as explained and illustrated in 4.1 below.

4.1 Event Focus

A (prototypical) verb is used to describe an event, and its lexical meaning specifies the possible scope of events it can describe. Following Smith's [1991] proposal of viewpoint focus in her account of verbal aspects, an *event focus* is taken to be a conceptual and cognitive profile that allows meaning extensions within the scope of lexical specification. The notion of event focus is as important as that of event components. A typical example can be found in the following case of 'building' verbs: *jian* (建) vs. *gai* (蓋). The two synonomous verbs seem to have the same event components, yet they have different event focuses. The verb *jian* allows an intransitive use with the Theme being the subject, thus highlighting the Event-endpoint [cf. Liu 1999; Huang, Liu and Tsai 1999]. Thus, in

(18b) below, only *jian* can be used:

(18) a. *gongren zhengzai jian/gai fangzi.* 工人正在 建 / 蓋 房子 .
workers DUR JIAN/GAI house
'The workers are building the house.'

b. *Na-dong fangzi jian/*gai yu 1888 nian.* 那棟房子 建 /* 蓋 於 1888 年 .
That-CL house JIAN/*GAI in 1888 year
'That house was built in 1888.'

Given its lexical specification, the verb *jian* is capable of describing an event of building from the perspective of its completion. This is why *jian* is allowed in (18b), where the focus is on the ending state of the event, i.e., the Event-endpoint.

4.2 A Representational Framework: MARVS

As mentioned above, a representational scheme called the Module-Attribute Representation of Verbal Semantics (MARVS) was proposed as the basis for verbal semantic description and representation [Huang and Ahrens, 1999]. It characterizes verb meanings in terms of modular and attributive distinctions: information pertaining to the aspectual composition is represented as the Event Module, and any event-internal specifications are coded as Inherent Attributes; information pertaining to participant roles is coded as the Role Module, and further specifications on a particular role are coded as Role-Internal Attributes. Below is a more detailed explanation of the four components of the model:

- **Event Module:** properties pertaining to the aspectual composition of the event(s). Five atomic event structures are distinguished ; they are Boundary [.], Punctuality [/], Process [////], State [___], and Stage [^^^^]. The combination of these atomic event structures renders 12 different event types.
- **Inherent (Event-internal) Attributes:** attributes referring to the semantics of the event itself, such as Control, Change-of-state, etc.
- **Role Module:** properties referring to focused (though not necessarily obligatory in its predicate argument structure) roles of the event, such as Agent, Theme, Instrument, Manner, Goal, etc.
- **Role-Internal Attributes:** attributes referring to the internal semantics of a particular focused role (of the event), such as Factive, Generic, Volition, Affectedness, etc.

4.3 Lexical Distinctions Redefined as the MARVS Representation

The distinctions among the four verbs, *TOU, ZHI, DIU,* and *RENG* can be re-defined and represented within the proposed MARVS framework:

- In terms of the Event Module, all four verbs describe an activity with a starting point, but only ***DIU*** is specified with an Event-endpoint, represented as a bounded process [• ///// •].
- In terms of the Inherent Attributes, *TOU/ZHI* behave differently from *DIU/RENG* in that the events of *TOU/ZHI* are highly directional while *DIU/RENG* is underspecified in terms of directionality.
- In terms of the Role Module, *TOU/ZHI* can both take a Path-endpoint as the direct object while the role of a Path-endpoint is not salient in the meaning of *DIU/RENG*.
- With regard to Role-internal Attributes, *TOU* casts a further specification on the spatial characteristics of the Path-endpoint: it has to be bounded as a container.

Below is a schematic MARVS representation of the lexical distinctions among the four verbs:

(19) MARVS Representation of the Semantic Differences among Verbs of Throwing

Module/ Attributes	***TOU***	***ZHI***	***DIU***	***RENG***
Event Module	Inchoative Process • /////	Inchoative Process • /////	Bounded Process • ///// •	Inchoative Process • /////
Inherent Attributes	+ Directional	+ Directional	- Directional Endpoint-focused	- Directional
Role Module	+ Path-endpoint	+ Path-endpoint	- Path-endpoint	- Path-endpoint
Role-Internal Attributes	Spatially-bounded			

5. Conclusion

The set of four Mandarin near-synonyms studied here serves to illustrate a newly developed framework for Mandarin lexical semantic studies. It also raises several important questions concerning the proper approach to lexical semantic research:

- While some works on English verbal semantics [e.g. Levin 1993, Atkins and Levin 1991, Atkins *et. al.* 1988] have concluded that diathesis alternations are most useful in identifying crucial semantic-syntactic interdependencies, such an approach may not be adequate when applied to Mandarin, given that Mandarin is relatively flexible in argument placement. The findings of this study seem to indicate that V-O compounding in Mandarin is an important clue for delimiting lexical meanings.

- Given that a Mandarin-specific framework is needed, this study may be taken as a pilot effort in searching for the most suitable and effective approach to studying of the Mandarin verbal system. The model of event-structure information as proposed above help to identify and represent the crucial semantic factors that are syntactically relevant.
- Viewed in a more general context, this work may help to illustrate several theoretical and methodological points. First, corpus data and computation may reveal some important generalizations that might not be available from elicited data only. In other words, semantic distinctions may not be easily captured if corpus-based, discourse-triggered syntactic patterns are ignored. Secondly, semantic distinctions may have various event-structure facets, which can be best understood if event focuses and event types are taken into consideration. Finally, the clustering of morpho-syntactic patterns with lexical-semantic characteristics proves to be fruitful in differentiating near-synonyms as well as in systematically disentangling the complex interaction between syntax and semantics.

References

ATKINS, B.T., and LEVIN, B., "Admitting impediments," in *Lexical Acquisition: Exploiting On-line Resources to Build a Lexicon,* ed. by U. Zernik Hillsdale. NJ: Lawrence Erlbaum Associates, 1991.

ATKINS, B.T., J. KEGL, and B. LEVIN, "Anatomy of a Verb Entry: from Linguistic Theory to Lexicographic Practice," *International Journal of Lexicography 1*, 1988: pp. 84-126.

BIBER, D., "Investigating Language Use Through Corpus-based Analyses of Association Patterns," *International Journal of Corpus Linguistics Vol. 1(2)*,1996: 171-197.

Chao, YUEN-REN, "A Grammar of Spoken Chinese," Berkeley: University of California Press, 1968.

CKIP group (Chen, Keh-jiann, Li-li Chang, Chu-Ren Huang, and Ching-Chun Hsieh). "A classification of Chinese verbs for language parsing," *Proceedings of 1988 International Conference on Computer Processing of Chinese and Oriental Language (ICCPCOL)*, 1988: pp. 414-17. Toronto, Canada.

CROFT, WILLIAM (ed.), *Syntactic categories and grammatical relations*, Chicago: University of Chicago Press, 1991.

CROFT, WILLIAM, "Possible verbs and the structure of events," in *Meanings and Prototypes: Studies in Linguistic Categorization,* ed. by S. Tsohatzidis. London/New York: Routledge, 1990.

DOWTY, DAVID, "Thematic Proto-roles and Argument Selection," *Language 67*,1991: pp. 547-619.

GOLDGERG, ADELE, Constructions: *A Construction Grammar Approach to Argument Structure*, Chicago: The Univ. of Chicago Press, 1995.

GRANDY, RICHARD E., "Semantic Fields, Prototypes, and the Lexicon," in *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization,* ed. by Lehrer and Kittay, pp. 103-122, Hillsdale: Lawrence Erlbaum, 1992.

GIVON, T., *English Grammar: A Function-based Introduction*, Vol. 1, Amsterdam/Philadelphia: John Benjamins Publishing Co, 1993.

HER, ONE SOON, *Grammatical Functions and Verb Subcategorization in Mandarin Chinese,* Taipei: Crane Publishing Co, 1990.

HOPPER, PAUL, and THOMPSON, SANDRA, "Transitivity in Grammar and Discourse," *Language 56*, 1980: pp. 250-299.

HUANG, CHU-REN. and AHRENS, KATHLEEN, "The Module-Attribute Representation of Verbal Semantics," in *Working Papers on Chinese Verbal Semantics, Vol.* 1, pp. 1-14, ed. by K. Ahrens, C-R Huang, and M-C Tsai, Taipei: Academia Sinica, 1999.

HUANG, CHU-REN, Liu, MEI-CHUN, and TSAI, MEI-CHIH, "From Lexical Meaning to Event Structure Attributes: Across Semantic Classes of Mandarin Verbs", in *Working Papers on Chinese Verbal Semantics, Vol. 1,* pp. 15-28, ed. by K. Ahrens, C-R Huang, and M-C Tsai, Taipei: Academia Sinica, 1999.

HUANG, S, "Emergent Lexical Semantics," Paper presented at the Mini-Conference on Lexical semantics, Graduate Institute of Linguistics, National Chung-Cheng University, 1995.

JACKENDOFF, R, *Semantic Structures,* Cambridge: MIT Press, 1990.

LAKOFF, GEORGE, *Women, Fire, and Dangerous Things,* Chicago: Univ. of Chicago Press, 1987.

LANGACKER, R, *Foundations of Cognitive Grammar, Vol. 1*, Stanford: Stanford Univ. Press, 1987.

LANGACKER, R, *Foundations of Cognitive Grammar, Vol. 2*, Stanford U, Press, 1991.

LEVIN, BETH, *Verb Classes and Alternation.* Chicago: Univ. of Chicago Press, 1993.

LIU, MEICHUN, "Lexical Meaning and Discourse Patterning - the Three Mandarin Cases of *'Build*'," in *Cognition and Function in Language,* ed. by B. Fox, D. Jurafsky, and L. Michaelis, pp. 181-199, Stanford: CSLI, 1999.

LIU, MEICHUN, "Conceptual Basis and Categorial Structure: a Study of Mandarin V-R Compounds as a Radial Category," *Chinese Languages and Linguistics 4*: pp. 425-51, Taipei: Academia Sinica, 1997.

LIU, MEICHUN, "Conceptual Manipulation and Semantic Distinction: the Case of Mandarin Postverbal DE Complement," in *Conceptual Structure, Discourse, and Language*, ed. by Adele Goldberg, pp. 333-46. Stanford: CSLI, 1996.

LIU, MEICHUN, "A Pilot Study on Chinese Verb Classes and Alternations," NSC Project Report (NSC85-2418-H-009-003), 1996b.

PUSTEJOYSKY, JAMES, *The Generative Lexicon,* Cambridge: The MIT Press, 1995.

PUSTEJOYSKY, JAMES, "The Syntax of Event Structure," *Cognition 41*, 1991: pp. 47-81.

SMITH, C. S., *The Parameter of Aspect*, Dordrecht: Kluwer Academic Publishers, 1991.

TENNY, C., "The Aspectual Interface Hypothesis," in *Lexical Matters*, ed. by Ivan A. Sag and Anna Szabolcsi, pp. 1-27, Stanford: Stanford University, 1992.

TENG, SHOU-HSIN (ed.), *Chinese synonyms usage dictionary*, Taipei: Crane Publishing Co, 1994.

TSAI, MEI-CHI, CHU-REN HUANG, and KEH-JIANN CHEN, "You jinyici bianyi biaozhun kan yuyi, jufa zhi hudong," (From near-synonyms to the interaction between syntax and semantics), paper presented at IsCLL-5, Taipei, Taiwan, 1996.

TSAI, MEI-CHIH, CHU-REN HUANG, KEH-JIANN CHEN, KATHLEEN AHRENS, "Towards a Representation of Verbal Semantics - an Approach Based on Near Synonyms," *Proceedings of ROCLING X*,pp. 34-48, Hsinchu, Taiwan, 1997.

TSAO, FENG-FU, "On Verb Classification in Chinese," Journal of Chinese Linguistics 24, 1996: pp. 138-191.