

國科會研究計畫期末報告

總計畫：知識表徵—由詞彙出發之基礎研究

Knowledge Representation: A Lexical Approach

子計畫三：事件結構與詞彙知識表徵

Event Structure and Lexical Knowledge Representation

計畫編號：NSC 90-2411-H-009-004-MC

主持人：交大外文系 劉美君

計畫助理：巫宜靜，吳欣達，陳伊貞

本計畫的執行，使中文動詞語意的研究更深入，也更趨完整。在以下報告中，將分三部分呈現研究成果。第一部份對詞彙語意研究的主要方法架構做一分析對照，接著以討論類的動詞為例，說明本研究所採取的方法，及其優點。第二部份則將所研究的其它類型動詞簡略介紹。第三部份則比較現有的動詞詞彙庫分類方法之異同，作為後續研究的參考。

第一部份：總論與範例

Abstract

The importance of deciphering lexical knowledge, especially the eventive information encoded on verbs, has been highly recognized and various approaches have been proposed to represent lexical information that plays a crucial role in grammatical realization. While quantitative NLP approaches might tend to overlook some semantic details, it is shown with detailed illustrations below that a corpus-based, contrastive analysis of near-synonyms can be most useful in extracting lexical information that are critical in differentiating verbs.

1. Introduction

As lexical information is considered to be the key for natural language processing

(NLP), the need for processing a large amount of lexical information has always been an issue in NLP research and applications. From a purely linguistic perspective, the study of lexical semantics has also been a focal area in recent years as linguistics in general is pushing its frontier toward lexicon-driven theories and practices. More and more researchers believe that the lexicon is where most of our linguistic knowledge has been stored. Entries of verbs, in particular, constitute a central part of the lexicon and are crucial for understanding grammatical structures. In the following, I will give a brief overview of the development of verbal semantics in terms of *why* it is studied, *what* is to be studied and *how* it has been studied.

1.1 Why studying verbal semantics?

The semantics of verbs has always been a core concern in linguistic theories. In the early generative paradigm, the meaning of a verb is generalized into and disguised under the so-called subcategorization frame. The formal theories in general try to find ways to link sentence structures with the argument structures of verbs, which, presumably, are lexically-specified. The theory of Case Grammar looked at argument structures from a purely semantic perspective and defines verb meanings with semantic roles (Fillmore 1968). As huge lexical databases are built to respond to the need of NLP applications, linguistic research in recent years has also shifted its orientation from structurally-based to lexically-based approaches. Verbal semantics has thus gained increased significance with regard to linguistic analysis as well as knowledge representation.

Each verb lexicalizes some unique eventive information with a range of possible ‘templates’ for argument expressions. As Levin and Rappoport Hovav (1996) puts it, it is generally assumed that the syntactic realization of arguments – their syntactic type and grammatical function – is predictable to a large extent from the meaning of verbs. The major goal in verbal semantic studies is then to extract the ‘meaning’ of a verb from examining the range of its argument expressions and other collocational associations.

1.2 What is to be studied?

Under the assumption that the meaning of a verb determines its syntactic behavior, various models have been proposed to explain the mapping from lexical semantics to syntax. Central to the issue is the task to *identify and represent semantic ‘determinants’ that shape the syntactic behavior of a verb*. Given that meanings are

multi-faceted, what are exactly the meaning components that are grammatically relevant?

Levin and Rappoport Hovav (1996) provided examples for English to illustrate grammatically relevant aspects of meaning. For most people, ‘loudness of speech’ seems to be a well-defined semantic feature that is also cognitively salient. However, verbs that differ in loudness of speech, such as *whisper vs. shout*, do not show any major differences in argument expression. This shows that the parameter of ‘loudness’ may not be syntactically relevant. On the other hand, the semantic distinction between ‘manner of speaking’ and ‘content of speaking’ proves to be relevant to syntax, as evidenced from their differences in participating in a conative pattern, taking an *at*-phrase:

(1) I whispered/*say at Mary.

It is clear that only certain aspects of meaning will surface in syntactic realizations and constitutes the target of investigation. Lexical semanticists are concerned with ways of delimiting these syntactically relevant semantic components from the syntactically irrelevant information.

1.3 How is verbal semantics studied?

In searching for the grammatically relevant aspects of meaning, a number of approaches have been attempted in the past. In the following, I will briefly introduce some of the major approaches, including frame semantics (Fillmore and Atkins 1992), Alternation-based approach (Levin 1993), conceptual structure and the localist approach (Jackendoff 1990), Generative lexicon (Pustejovsky 1995) and the corpus-based approach (Biber et al. 1998).

1.3.1 Frame Semantics

The semantic frame-based approach argues that a word’s meaning ‘can be understood only with reference to a structured background of experience, constituting a kind of conceptual prerequisite for understanding the meaning’ (Fillmore and Atkins 1992). Take for example the group of commercial transaction verbs - *buy, sell, charge, spend, pay, and cost*. These verbs all require an understanding of property ownership and money economy, a knowledge schema involving four major semantic categories: Buyer, Seller, Goods, and Money. These verbs differ in the ways of expressing these categories (The table below is taken from Fillmore and Atkins 1992: 79):

(2) Semantic and Syntactic Valence for Verbs in the Transaction Frame

| | Buyer | Seller | Goods | Money |
|---------------|----------------|----------------|------------|------------|
| <i>Buy</i> | Subj | (from) | Direct-Obj | (for) |
| <i>Sell</i> | (to) | Subj | Direct-Obj | (for) |
| <i>Charge</i> | (Indirect-Obj) | Subj | (for) | Direct-Obj |
| <i>Spend</i> | Subj | NULL | For/on | Direct-Obj |
| <i>Pay</i> | Subj | (Indirect-Obj) | (for) | Direct-Obj |
| <i>Cost</i> | (Indirect-Obj) | NULL | Subj | Direct-Obj |

In sum, verbs of the same class share the same semantic frame, defined with a set of frame-specific elements. The differences among the verbs can be accounted for with different ‘profiles’ of the frame as well as the various mapping relations between frame elements and grammatical roles.

1.3.2 Diathesis Alternation

Levin (1993) provides a complete study of English verbs and verb classes, based on diathesis alternation patterns, i.e., alternations in the expression of arguments. The basic assumption behind the work is that the behavior of a verb, particularly with respect to the expression of its arguments, is to a large extent determined by its meaning. And the group of verbs exhibiting the same alternation patterns shares the same meaning components and belongs to the same semantic class. For example, the Locative Alternation in English can be used to distinguish the *spray/load* group of verbs from the *fill/cover* group, since only the *spray/load* group can participate in Locative Alternation, as exemplified below:

(3) Locative Alternation in English

- a. *He sprayed water on the plants*
- b. *He sprayed plants with water.*

A more complicated example is found with a set of prototypical transitive verbs: *touch*, *hit*, *cut*, and *break* (Levin 1993:6-10). These verbs vary in terms of their participation across four different transitive alternations:

(4) English Transitive Alternations

| | <i>touch</i> | <i>hit</i> | <i>Cut</i> | <i>break</i> |
|--|--------------|------------|------------|--------------|
| Conative <i>'X hits Y.'</i> <i>'X hits at Y.'</i> | No | Yes | Yes | No |
| Body-part Possessor Ascension <i>'X hit Y's head.'</i> <i>'X hits Y on the head.'</i> | Yes | Yes | Yes | No |
| Middle <i>'X cut Y.'</i> <i>'X cuts easily.'</i> | No | No | Yes | No |

| | | | | |
|---|----|----|----|-----|
| Causative 'X broke Y.' 'Y broke.' | No | No | No | Yes |
|---|----|----|----|-----|

Each of the above alternations is in principle associated with certain meaning components. Verbs that may participate in a given alternation are said to have the associated meaning components. Based on their differences as shown above, the verbs can be categorized into different verb classes, with distinct semantic features:

(5) Four Verb Classes:

- a. *touch*: pure verb of contact (no implication for change of state)
- b. *hit*: a verb of contact by motion
- c. *cut*: a verb of causing a change of state by moving something into contact with the entity that changes state.
- d. *break*: a pure verb of change of state

In order to successfully classify verbs with alternations, the semantic characteristics required by each alternation will have to be clearly defined. And the alternation-based approach faces the problem of identifying all the necessary and sufficient semantic conditions for each alternation.

1.3.3 Conceptual Structure and the Localist Approach

According to Jackendoff (1990), 'meaning' in natural language is an information structure that is mentally and spatially encoded. The semantic primitives that may be utilized to construct the mental representation include the following:

(6) Semantic primitives:

- a. [STATE] BE ([THING], [PLACE])
- Jim is in the pub. (spatial) → The party is on Sat. (non-spatial)
- b. [EVENT] GO ([THING], [PATH])
- Jim went into the pub. (spatial) → The prize went to Kate (non-spatial).
- c. [PLACE] IN ([THING])
- d. [PATH] TO ([PLACE])

This mental structure approach is de-compositional in nature. And it is not clear how complicated information can be represented with the primitives.

1.3.5 Generative Lexicon

As an attempt to deal with both lexically-specified information and contextually-derived meaning coercions, Pustejovsky (1995) proposed some

compositional mechanisms for meaning generation. In his framework, there are multiple levels of semantic representation:

(7) Levels of Semantic Representation:

a. Argument Structure

eg. *Build* ARG1=animate, ARG2 = artifact

b. Event Structure (lexical aspect)

eg. *Build* E1= Process, E2=state

c. Qualia Structure (eg. *book*)

-Formal : type of things

‘This book is thick.’

-Agentive : how to ‘come to being’

‘He wrote a book.’

-Telic : purpose

‘He read a book.’

-Constitutive: parts/content

‘The book is interesting.’

d. Inheritance Structure

1.3.5 The Corpus-based Approach

The corpus-based approach (e.g., Biber et al 1998) believes that the meaning of a verb is projected in its ‘association patterns’ found in a large corpus. Statistic counting of ‘association patterns’ then serves to reveal the semantic distinctions. For example, the difference between the two adjectives *strong vs. powerful* is best illustrated by their collocational tendencies: *strong coffee vs. powerful car*. And the statistical findings on the complement types between the two verbs *begin* and *start* also demonstrate their semantic differences (Biber et al 1998):

(8) Corpus-based Statistics: *begin vs. start*

| | <i>Begin</i> | <i>Start</i> |
|--------|--------------|--------------|
| +To -V | 60% | 17% |

It is shown that *Begin* is followed much more often by an infinitival phrase, indicating that *begin* is a complement-taking verb, while *start* predicates a simple event.

The association patterns observable from a large corpus help to ‘delimit’ the grammatical distinctions that require a semantic account. However, linguistic analyses are still needed to ‘define’ the relevant meaning facets behind the grammatical differences.

1.4 Current Program on Mandarin Verbal Semantics

The above approaches have been successfully applied to the study of English verbs. But are they equally applicable to the study of Mandarin verbs? A team of researchers have worked on Mandarin verbal semantics in recent years and found that a contrastive analysis of near-synonym pairs, based on corpus observations proves to be effective in ‘extracting’ verb meanings that are syntactically relevant. The program can be characterized as follows:

(9) Research Program on Mandarin verbal semantics:

Task: to identify eventive information that is *syntactically* relevant

Data: Sinica Corpus with 5 million words

(<http://www.sinica.edu.tw/ftms-bin/kiwi.sh/>)

Method: Comparison of near-synonym sets, based on *association patterns*

Representation: Module-Attribute Representation of Verbal Semantics, MARVS

(For details, see Huang et al 2000).

Various works have been published that focus on individual sets of Mandarin verbs (Tsai et al. 1998; Chang et al 2000; Liu 1991, Liu 2000; for a collection of results, see Liu 2002). In this paper, I will focus on a sub-set of communication verbs, namely, verbs of discussion (商量 vs. 討論), and demonstrate how a detailed linguistic analysis of near-synonyms can help find the crucial semantic distinction between the pair of verbs, while other currently available resources with a top-down approach to the verb system tend to miss these important distinctions.

範例：討論類動詞（Verbs of Discussion）

2. An illustration: Verbs of dicussion

The class of communication verbs all encode an event that involve the action of verbal communication. The class may be broken down to at least 7 sub-classes:

(10) Sub-classes of Communication Verbs

- a. Verbs of talking: 談, 提
- b. Verbs of speaking: 說, 講
- c. Verbs of telling: 告訴, 告知, 通知
- d. Verbs of discussion: 討論, 商量
- e. Verbs of expressing: 表達, 表示, 表露
- f. Verbs of complaining: 埋怨, 抱怨
- g. Verbs of asking: 問, 詢問

2.1 Frame elements

Each verb class and hence a different frame is clearly defined with its prototypical frame elements. Take verbs of discussion as an example, they belong to the

*Conversation frame and its frame elements include:

(11) Frame Elements in the Conversation frame

- a. Interlocutor 1
- b. Interlocutor 2
- c. Interlocutors
- d. Topic
- e. Medium

According to FrameNet, the verb *discuss* shares the same set of frame elements with other verbs in the Conversation frame. However, in Mandarin, there are two distinct lexical entries that may both be glossed as ‘discuss’ - 討論 and 商量. What are the semantic distinctions between them? It seems that though sharing the same frame, the two verbs must have encoded different semantic details that make them distinct? What are the distinct semantic details and how can we find them?

In the following, I will first examine information in existing resources and see what they can offer.

2.2 Dictionary Definitions - Verb Usage Dictionary (動詞用法詞典)

In Verb Usage Dictionary (Meng et al. 1985), a dictionary focusing solely on verbal knowledge, the two verbs are defined as:

(12) Dictionary definition:

- 討論：就某一問題交換意見或進行辯論
商量：交換意見

The definitions seem to be interchangeable and do not show any critical differences. If we look at their respective usages, again, the information is roughly the same:

(13) Usages for 討論 and 商量

a. 討論

- [名賓] ~ 科研計畫
[動賓] ~ 怎麼處理這件事

[小句賓] ~這個人怎麼安排
 [動時量] ~一下兒
 [了著過] ~了一上午/還~著呢/~過這個問題
 [重疊] ~~計畫生育問題
 【名賓類】[對象]~問題|~計劃|~措施|~方案
 【動結】~//清楚|~//完~得/不了~//成 ~好
 【動趨】~上~//下去~出 ~[不]出來~得/不過來~起

b. 商量

[名賓] ~一件事
 [動賓] ~怎麼辦
 [小句賓] ~此事讓誰處理
 [動時量] ~一下兒
 [了著過] ~了一上午/還~著呢/~過這個問題
 [重疊] ~~計畫生育問題
 【名賓類】[對象]~問題|~計劃|~措施|~方案
 【動結】~//清楚|~//完~得/不了~//成 ~好
 【動趨】~上~//下去~出 ~[不]出來~得/不過來~起

From the above, we see that the ranges of usage for the two verbs are also identical. The information in the dictionary doesn't appear to be helpful.

2.3 The Grammatical Knowledge-Base of Contemporary Chinese: A Complete Specification (現代漢語語法信息辭典詳解)

Next, we look at resources that are designed for NLP applications. In the Grammatical Knowledge-Base of Contemporary Chinese (現代漢語語法信息辭典詳解, Yu et al. 1998), there are two differences found: 1) 討論 can be the object of 'have' as in 有一個討論 'have a discussion', but 商量 cannot; 2) 討論 can be nominalized and preceded by a noun as in 政策討論 'policy discussion'. The two differences are highlighted below:

(14) Information in The Grammatical Knowledge-Base of Contemporary Chinese

| 詞語 | 討論 | 商量 |
|-----|-----------------|------------------|
| 准謂賓 | 准 | 准 |
| 有賓 | 有 | |
| 前名 | 可 | |
| 後名 | 可 | 可 |
| 體謂准 | 體謂 | 體謂 |
| 句賓 | 句 | 句 |
| 動趨 | 趨 | 趨 |
| 著了過 | 著了過 | 著了過 |
| 在 | 在 | 在 |
| 重疊 | ABAB | ABAB |
| 謂 | 可 | 可 |
| 備註 | 時事~/~時間/~這事怎樣處理 | ~對策/~結果/~誰去辦理/這件 |

| | | |
|--|--|------|
| | | 事~過了 |
|--|--|------|

2.4 HowNet

In HowNet (Dong 1998a, b), both verbs are defined with the concept ‘discuss’ (DEF=discuss|商討), and the concept ‘discuss|商討’ is linked with the following hierarchical sense relations:

(15) Hierarchical relation of the concept ‘discuss|商討’ inHowNet

1.2.2.2.2.3.3.2.3.2.,discuss, 商討,1, {agent, partner, content }

- event,事件 – act,行動 – ActSpecific,實動
- AlterSpecific,實變 – AlterState,變狀態
- AlterMental,變精神 – AlterKnowledge,變感
- MakeOthersKnowledge,使他人感知
- communicate,交流 – discuss,商討

Basically, the two verbs share the same concept and are not further differentiated in the framework of HowNet.

2.5 Chinese Synonyms Usage Dictionary (近義詞用法詞典)

In Chinese Synonyms Usage Dictionary (Teng 1994), the two verbs are compared in detail and significant differences can be found. First, their definitions differ:

(16) Definitions in Dictionary of Near-synonyms:

- 討論：to discuss something so as to establish pros and cons; talk over
- 商量：to discuss something so as *to settle an issue or to achieve an aim*

Second, they may collocate with different grammatical patterns. Six differences can be found in the book:

(17) Collocational Variations

| | 討論 | 商量 |
|----------------|-----|-----|
| Modifier 會 | Yes | no |
| Noun 這一次 進行 | Yes | no |
| V 出主意 辦法 | No | yes |
| VR 好 得怎麼樣 | No | Yes |
| Manner 熱烈 | Yes | No |

| | | |
|------|----|-----|
| abab | No | Yes |
|------|----|-----|

As shown in the table, only 討論 may be nominalized to function as a modifier or head noun, and it may take a manner adjunct, while only 商量 may co-occur with certain resultatives. It is clear that by focusing on near-synonyms, Teng (1994) is able to discover some important distinctions between the verbs. However, the above data are based mainly on intuitive judgments. We might wonder if corpus data will provide any new insight or different pictures?

In the next section, findings based on corpus observations will be presented.

3. Corpus-based Contrastive Analysis

Observations of the corpus data (Sinica Corpus) show that the two verbs are associated with quite different patterns in terms of the semantic details of the participant roles (frame elements) and their tendency for nominalization and manner modification.

3.1 Association Patterns: semantic distinction in participant roles

The two verbs commonly take a discussant(s)-agent as the subject and a topic-theme as the direct object, but they display different association patterns in terms of the kinds of topics and discussants they may take. With regard to the topic-theme, when a solution is sought (20a), both 討論 and 商量 can be used, but when the object names predetermined agenda (20b) or a potential outcome (20c), only 討論 can be used:

(18) Association Patterns: Topic-theme

a. when a 'solution' is sought:

討論/商量 回大陸的方式和時間

討論/商量 如何克服目前的困境

b. with predetermined agenda:

討論/*商量 人我關係/語言學/文學

c. with a potential outcome

討論/*商量 戰爭對股市可能的影響

Besides taking different object-themes, the verbs differ in terms of the semantic requirements of the discussant-agent. For 商量, only those discussants who have authority or control over the issue being discussed can occur as the subject:

(19) Association Patterns: Discussant-Agent

- a. 董事開會 討論/商量 公司的經營策略
- b. 工人開會 討論/*商量 公司的經營策略

Although sharing the same frame elements <discussant, topic>, the two verbs differ in the semantic details of these elements.

3.2 Association Patterns: Nominalization and Modification

As also mentioned in Teng (1994), the two verbs show different tendencies in nominalization and manner modification. The verb 討論 can undergo nominalization and function as the head of a possessive phrase as in (20), and it can be modified with a preverbal manner adverb as in (21a) or a postverbal complement (21b):

(20) Association Patterns: distinction in nominalization

他們的 討論/*商量 很熱烈

(21) Association Patterns: Manner Modification:

- a. 他們 討論/*商量 得很熱烈
- b. 他們熱烈(地) 討論/?商量 了三個小時

The most crucial point the above corpus observations added to our understanding of the verbs is that they are associated with different semantic requirements of the participant roles. In term of the semantic characters of the Topic-theme (an NP or a clause), 討論 takes a wider range of topics, which may be a referential entity independent of the event of discussion, but 商量 is quite restricted in taking only those topic-themes that call for a solution, which the discussants have **control and authority** over. The coming about of the solution depends completely on the solution-seeking process of 商量. Therefore, the theme associated with 商量 should be viewed as an Incremental Them since its existence arises from the progress of the event (Dowty 1991) and it may serve to delimit or measure out the event (Tenny 1992). The difference are summarized below:

(22) Differences in the associated Topic-Theme

- a. 討論: - Any topic (human or non-human, declarative or interrogative)
 - referentially independent
- a. 商量: - Issues that call for a solution the discussants have control over
 - The existence of the Theme depends on the realization of the Event.

3.3 Distinction in Event Structure

The above distinctions can be analyzed as deriving from the fundamental difference in their event structures. 討論 is a process verb that encodes a simple event of verbal activity, roughly equivalent to the meaning ‘discuss over a Topic’. Given its simple eventuality, 討論 may be used as a noun to refer to the nominal activity. On the other hand, 商量 encodes a more complex event that involves a process and an endpoint. The verbal activity of 商量 ends up with a verbal product or solution. The meaning of 商量 can be glossed as ‘discuss in order to come up with a Solution’. Their distinction in event types is summarized as follows:

(23) Event Structure

討論: Process
 <Verbal Activity>
 - Def.: Discuss over a THEME
 - ex. 討論[主題]

(24) Event Structure

商量: Process + Endpoint
 <Verbal Activity + Verbal Product>
 - Def.: Discuss in order to come up with a SOLUTION
 - ex. 商量[對策]

3.4 Overt Coding of Verbal Product

The coding of an incremental theme arising from the event of 商量 is not an isolated phenomenon in the class of communication verbs. In another sub-set, i.e., verbs of explaining (解釋 vs. 說明), a similar patterning is found, whereby the incremental theme is overtly coded in the presence of a topic-theme:

(25) Incremental Theme in the Use of 解釋

| | | |
|---------|-----|-------------------|
| 將[他的提議] | 解釋為 | [自私自利的舉動] |
| | | |
| Theme | | Incremental Theme |
| [topic] | | [verbal product] |

The arising of some kind of an end product through the process of verbal communication seems to be characteristic of the group of communication verbs.

3.5 Nominalization and Event Focus

When communication verbs are nominalized, they may refer to different portions of the event structure. With a simple eventuality, such as 討論, nominalization normally refers to the on-going process, a nominal activity. With a complex

eventuality, such as 解釋/說明, nominalization may refer to the verbal product arising from verbal activity, or the activity itself. As shown below, in (26a), the modifier 熱烈 ‘energetically’ can only characterize the process of 討論, focusing on the activity nominal, while in (26b), the modifier 簡單 ‘simple’ applies to the verbal product, and in (26c), either the activity or the product may be the focus of modification:

(26) Nominalization with Different Event Focuses

- a. Nominal Activity
他們的 討論/*商量 很熱烈
- b. Verbal Product
他的 解釋/說明 很簡單
- c. Both
激烈的表達 (Nominal Activity)
簡單的表達 (Verbal Product)

With verbs of communication, the grammatical process of nominalization profiles different ‘event focuses’, that is, the focus of the nominalization, as revealed by the scope of the modifier, may fall upon different portions of the event structure

4. Semantic Representation with MARVS

The representational scheme proposed in Huang et al (2000), namely, the Module-attributed Representation of Verbal Semantics (MARVS), tries to translate all detectable verbal information into two categories: information pertaining to event types (the Event Module) vs. information pertaining to participant roles (the Role Module). And further semantic distinctions within each module can be coded as Event-internal or Role-internal attributes. The above-mentioned semantic distinctions characteristic of the Mandarin verbs of discussion (討論/商量) can be represented within the MARVS framework as follows:

(27) MARVS for 討論/商量

| Verb | Event Module Event Internal Attribute | Role Module Role Internal Attribute |
|------|--|--|
| 討論 | process /// | <DiscussantAgent> <TopicTheme > |
| 商量 | bounded process · ///· | <DiscussantAgent>: [control/authority] <IncrementalTheme>: [solution] |

The two verbs differ in event types, one is a typical process verb; the other is a bounded process with a potential endpoint. As for core participants, they both take a discussant-agent as the subject, but 討論 takes a topic-theme as the object, while 商量 prefers an incremental theme. In addition, 商量 imposes further semantic requirements on both participants. The agent of 商量 is specified with the feature [control/authority], to show that it has to hold authority over the issue being discussed, and the incremental theme is some kind of a ‘solution’, an outcome dependent of the event.

第二部分:其它類型的動詞研究

1. 情緒活動類動詞

先前對於情緒動詞的研究,我們首先觀察“羨慕/嫉妒”與“同情/憐憫”這兩組動詞,詳細分析了它們的語法功能,進而擴及於整組情緒動詞的分析。我們發現情緒動詞皆為引動的心理狀態 (caused state),「肇因」(CAUSE) 這個語意成分是关键,可作為這組動詞的語意區分。此發現與過去針對英文的詞彙語意研究有高度的相關性。根據 Levin & Rappaport (1995)¹,他們認為不及物動詞²有及物使動的用法皆是“externally caused”,如無及物使動的用法則為“internally caused”,我們亦將以 causation 作為此類情緒動詞的區分依據。

■ 語料分析:

語料中 cause 若出現,則明顯為 externally caused,如 cause 不出現,則表示不一定有 external cause。Cause 可能以不同的語法語法角色出現,在以下的例句中‘小明’可視為一 causer,而其表現方式如下:

- (1) 小明讓我很羨慕-- causative use
- (2) 我很羨慕小明—transitive
- (3) 我對小明很羨慕—intransitive

我們將有「令、讓、使、叫」等字的句式視為 causative use 的句子,觀察的結果發現“羨慕”的使動比例約為“嫉妒”的四倍,換句話說,“羨慕”這個情緒是必須要有一個 cause 而導致,且這個 cause 必須在句法上出現,出現的形式就是 causative use。這樣的區分幫助我們了解“羨慕”與“嫉妒”的關鍵差異:

¹ Unaccusativity At Syntax-Lexical Semantics Interface by Beth Levin and Malka Rappaport Hovav

² 我們在此暫不討論此組動詞是否即為 Levin & Rappaport 所指的不及物動詞,亦即動詞其後是否以接 NP 與否作為及物與不及物的區別。

| | 羨慕(112) | 嫉妒(38) |
|--------------------|---------|---------|
| 1. Causative--令讓使叫 | 21(19%) | 2(5%) |
| 2. + NP | 53(47%) | 8(21%) |
| 3. 對+ target | 4(4%) | 3(8%) |
| 總計 | 78(70%) | 13(34%) |

觀察“嫉妒”的語料，我們發現有時 external cause 完全不出現，事件可以是 internally caused, 如例(4)(5)(6)(7), 同時，我們發現我們可以說 [他生來就愛**嫉妒**]，但不可說 [*他生來就愛**羨慕**]

(4) 為什麼要被愚迷.貪婪.嫉妒之心所緊縛,以致失去悅樂呢?

(5) 將嫉妒之心轉化為奮的力量

(6) 深表嫉妒

(7) 於是嫉妒之火在動物心裡燃燒著

我們將語料做統計分析，發現使用“嫉妒”時，55%的句子沒有標記外部成因，而“羨慕”的句子一定清楚標記導因：

| | 羨慕(112) 'envy' | 嫉妒(38) 'jealous' |
|-------------------|-------------------|---------------------|
| Mention (NP or S) | 100% (112) | 45% (17) |
| No Mention | 0% (0) | 55% (21) |

根據上表，我們得到以下的結果：

- 羨慕類：externally caused verb
- 嫉妒類：internally caused verb

將以上結果應用在其他近義詞組，我們有以下發現：

- CAUSE 的分佈

a) Verbs of sympathy

| | 同情(131) 'sympathize with' | 憐憫(19) 'pity' |
|-------------------|------------------------------|------------------|
| Mention (NP or S) | 82% (107) | 16% (3) |
| No Mention | 18% (24) | 84% (16) |

b) Verbs of Anger

| | 生氣(285) | 憤怒(111) |
|------------|-----------|-----------------|
| Mention | 56% (160) | 38% (42) |
| No Mention | 44% (125) | 62% (69) |

c) Verbs of Fear

| | 害怕(259) | 恐懼(141) |
|------------|-----------|-----------------|
| Mention | 73% (188) | 63% (89) |
| No Mention | 27% (71) | 37% (52) |

d) Verbs of Sadness

| | 傷心(134) | 悲傷(51) |
|------------|----------|-----------------|
| Mention | 52% (70) | 31% (16) |
| No Mention | 48% (64) | 69% (35) |

e) Verbs of Depression

| | 難過(223) | 痛苦(436) |
|------------|-----------|------------------|
| Mention | 40% (90) | 17% (73) |
| No Mention | 60% (133) | 83% (363) |

此外，語意較負面的「嫉妒類」動詞有另一特徵，即出現的頻率較低，但名物化的用法比例很高。名物化的用法又和肇因的標記息息相關：21個“嫉妒”的名物化例子中，有20(95%)個沒有標記外部肇因。“憐憫”的16個名物化例子中，全部沒有標記外部肇因。因此，我們得到一結論：語用上出現較少，語意限定較特殊 (pragmatically marked) 的情緒動詞，傾向於缺乏外部肇因的標記，這些動詞比較強調內心生成的情緒活動，為「內在因素」導向，不一定有可以解釋所以然的外因。同樣的分析可適用於“愛/喜歡”的對照。“喜歡”後可接表達肇因的子句，“愛”則不行：

(7) 我 喜歡/*愛 你穿裙子。

我們的結論是：

- 「肇因」是情緒動詞分類的一大語意關鍵。語用上較特殊 (pragmatically marked) 的動詞較常出現名物化，為肇因於內在因素的事件。
- 動詞的語用特性和語意特徵是彼此相關的。語意的界定有助瞭解語用的內涵。

■ 「模組屬性」為準的語意表達模式 (MARVS Representation)

| Verbs | Event Module Event Internal Attribute | Role Module Role Internal Attribute |
|-------------------------------|--|---|
| Pragmatically unmarked 羨慕類 | Inchoative state • _____ | <Experiencer, Target, Cause> [external] |
| Pragmatically marked 嫉妒類 | Homogeneous state _____ | <Experiencer, Target, Cause> [internal] |

2. 空間型態—懸掛類動詞 (Verbs of Spatial Configuration)

「懸、掛、吊」這三個詞意義十分類似。在教育部國語辭典上用「掛、系」來批註「懸」，用「懸吊」來解釋「掛」，用「懸掛」來批註「吊」。在 HowNet 中，「懸、掛、吊」都有用到「懸掛」作為解釋，而且都可以譯為英語的「hang」。如果我們按照蔡美智等(1996)對近義詞所採取的定義方式，那麼這三個字的近義詞地位應該是成立的。

從句法表現來看，這三個詞的相似度仍然很高。首先，他們都有不及物的用法：

- (1) a. 鐘懸在牆上 b. 鐘掛在牆上 c. 鐘吊在牆上

詞 類 除 了 不 及 物 變 用 現 法 之 外 也 這 三 個 (Locative Inversion) 句 中。

- (2) a. 牆上懸了一面鐘 b. 牆上掛了一面鐘 c. 牆上吊了一面鐘

再者，「懸、掛、吊」它們也都可以有把字句的及物用法：

- (3) a. 他懸一面鐘在牆上 b. 他掛一面鐘在牆上 c. 他吊一面鐘在牆上

此外，「起來」這個動向標誌也可以伴隨這三個動詞一起出現：

- (4) a. 我們把畫懸起來 b. 我們把畫掛起來 c. 我們把畫吊起來

這一組動詞的相同處固然很多，但是在下面的句子中，三個動詞的表現卻不盡相同：

- (5) a. 天上 懸 / ?掛 / *吊 著一輪明月

b. 喬登(Jordan)將球 高 *懸 / *掛 / 吊 籃下 (解釋為「將球高傳至籃下」)

- c. 職缺 / 兩人的感情 就這麼 懸 / *掛 / *吊 著

按照句法表現上來看，這三個動詞的論元結構都相同，能夠出現在其中的句型(及物、不及物)有很大部分相重疊。那麼，為什麼這三個詞會有(5)裏面這些不能互換的情況發生呢？顯然句法層次的解釋還不足以將「懸、掛、吊」這三個動詞區分開來。所以本文將試圖在句法之外尋找可以將這三個動詞加以區分的標準。以下是「懸、掛、吊」在語料庫中的出現分佈情形。

根據中研院的平衡語料庫，我們找到了如下表的語料分佈：

| | 懸 | 掛 | 吊 |
|--------------------|--------------------|--------------------------|-----------------------|
| 有 Agent 出現 | 6 (6/35 ÷ 17%) | 73 (73/228 ÷ 32%) | 16 (16/29 ÷ 55) |
| 例： | 這小女孩正式霍元龍懸下重賞要追尋支人 | 也有許多人自動在家門口掛國旗 | 赴國強頻吊阿迪對角球 |
| 後接動貌標誌 | 8/35(約 23%) | 78/228 (約 34%) | 3 筆 3/29(約 10%) |
| 著 | 8 | 63 (63/228 ÷ 28%) | 3 |
| 例： | 遺缺尚懸著 | 腰上掛著一把劍 | 井中吊著點亮的蠟燭 |
| 了 | 0 | 15 | 0 |
| 例： | NA | 樹上掛了彩色的燈 | NA |
| Locative Inversion | 10 (約 29%) | 57 (約 25%) | 2 (約 7%) |
| 例： | 頸下懸著無數條項鍊 | 臉上不時掛著微笑 | 在破破爛爛的藍色牛仔褲上，吊著四支白色口袋 |
| 有 locative 出現 | 21 (75%) | 149 (約 65%) | 9 (約 33%) |
| 例： | 意為懸在天上的湖 | 而且掛在樹上的牛頭掉下來 | 整夜在山上吊蚊帳誘蚊 |
| 後接方向詞 | 0 | 1 | 6 |
| 例： | NA | 右手大拇指可以把一部分的右手重量「掛」到琴筒上。 | 王俊隆將球吊至門前 |

從上面的語料分佈，我們可以發現，雖然如前面第一段所說，這三個動詞有許多相同之處，但是這些相同之處在語料庫中出現的比例卻有所不同。以「懸」來看，它比例最顯著的部分是在「處所倒」與「有處所」則是在「後接動貌標誌」這個標準上最為突出。「吊」比較多的情況是要有 Agent 出現以及後面要帶上方向詞。下面的段落將試圖從前人的研究找出可能用來解釋這種分佈現象的語意成分。

從「懸、掛、吊」可能出現的句法形式上，我們認為湯廷池(2000)中所提到的使動結構可以用來表達這三個動詞的「預設(default)」概念結構：[[x ACT (ON y)] CAUSE [BECOME [y BE AT <state>/<place>]]]。但它們個別的出現狀況則讓我們聯想到劉美君(1999)的「事件焦點」的概念。

就分佈情況來看，「懸」的事件焦點側重於處所狀態，也就是[BECOME [y BE AT <state>/<place>]]的部分，所以它帶的處所詞最多的概念結構則比較自由，可把焦點放在處所狀態或活動任一個，所以在各種句法

表現分佈上跟「懸」與「吊」比起來都較中庸。而「吊」的事件焦點主要在強調活動：[x ACT (ON y)]的部分，所以反映到句法上就是最能帶上 Agent。

另外，我們也將上面的分佈情況，整理成 MARVS 的表現模式，得到了下面的結果：

| 動詞 | 事件模組 事件內部屬性 | 參與角色模組 參與角色內部屬性 |
|----|--|--|
| 懸 | Simplex Event — [- agent control] | <theme, location> |
| 掛 | Complex Event · ///· — [+ agent control] | <theme, location> |
| 吊 | Complex Event · ///· [+ agent control] | <agent, theme, (location)> location can be [+ directional] |

「懸、掛、吊」這一組詞的研究，再一次證明「事件焦點」在近義詞的區辨當中，扮演著重要的角色。我們發現如果套用湯老師的表現模式，則可以抓住一般人的語感：「掛」這一組詞當中比較常用，也比較能代表整組詞，可惜的是這樣的表示模式卻太過廣泛，許多細節沒有辦法包括進來。如果使用 MARVS 則正好相反：許多詳細的資訊可以納入版圖，但是卻沒辦法抓住語感。這是這兩種呈現方式的可惜之處。

第三部分：FrameNet, HowNet, 與 Levin 的動詞分類異同

我們針對英語的詞彙網絡 FrameNet 與 Beth Levin (1993)的動詞分類，以及漢語的詞彙網絡 HowNet，研究其中的分類以及內容架構的異同，並探討建立漢語動詞詞彙語意網絡的問題。

一、分類方式

在分類架構方面，如下表所示，FrameNet 中分有 13 個類別的領域(domains), 60 個框架(frames), 共有 1,905 個詞項(lemma)，總共有三層。HowNet 中分了有 2 個大類，每類有 4 到 13 層，總詞數約有 116,533 個。Beth Levin 的分類中，分了 49 個類別，每類又細分成 1 到 4 層不等，共有 4,199 個詞。

| | FrameNet | Levin | HowNet |
|-------------|----------|-------|---------|
| 層級數 | 3 | 1-4 | 4-13 |
| 最大類別數 | 13 | 49 | 2 |
| 總詞數 (token) | 1,905 | 4,199 | 116,533 |

二、內容

在內容呈現方式上，此三者各有不同，如下表：

| | FrameNet | Levin | HowNet |
|--------|----------|-------|--------|
| 英文詞 | v | v | v |
| 中文詞 | | | v |
| 類別說明 | v | v | |
| 其他同類詞項 | v | v | |
| 語意角色 | v | | v |
| 句型 | v | v | |
| 句式合法度 | | v | |
| 例句 | v | v | |
| 說明 | v | v | |
| 特徵 | | v | v |

FrameNet 和 Levin 皆以英文詞彙為分類的標的，在內容上較為接近，皆有類別的說明，也列舉了其他同類的詞項，並且皆指出句型，標出例句以及說明，Levin 的分類中還標示出句式的合法度，以及句型特徵，但是並未如 FrameNet 和 HowNet 一樣有語意角色的訊息。FrameNet 中並未顯示不合法的句子，也未指出句型特徵；HowNet 中含有中文詞也有英文詞，有語意角色也有句型特徵，但無其餘訊息。

三、問題

FrameNet 的動詞類別較有限且無靜態動詞(static verbs)，例如: *be, belong, compare, include; begin, exist, fail* 等等。狀態動詞(state verbs)在英語中被歸為形容詞(adj.)，例如: *long, lucky, sad*。FrameNet 上的動詞，屬於 HowNet 上的「行動動詞」(act verbs)，且多屬於「實動動詞」(ActSpecific)。HowNet 中相關的詞彙語意以及句式的訊息較少。因此建立漢語動詞語意網絡，除了在既有的詞彙網絡基礎上截長補短之外，還必須注意語言的差異所導致的不同分類等問題。

References

- Baker, Collin F., Fillmore, Charles J., and Lowe, John B. 1998. The Berkeley FrameNet project. Proceedings of the COLING-ACL. Montreal, Canada.
- Biber, Douglas, Susan Conrad and Randi Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Dong, Zhendong. 1988a. Enlightenment and Challenge of Machine Translation.

- Shanghai Journal of Translators for Science and Technology*. 1.9-15.
- Dong, Zhendong. 1988b. Knowledge Description: What, How and Who? *Proceedings of International Symposium on Electronic Dictionary*. Tokyo, Japan.
- Dowty, David. 1991. Thematic Proto-Roles and Argument Selection. *Language* 67(3).547-619.
- Fillmore, Charles J. 1968. The Case for Case. *Universals in Linguistic Theory*, ed. by Emmon Bach and Robert T. Harms. 1-88. New York: Holt, Rinehart and Winston.
- Fillmore, Charles J., and Atkins, Beryl T. 1992. Toward a Frame-Based Lexicon: The Semantics of RISK and Its Neighbors. *Frames, Fields, and Contrasts*, ed. by Adrienne Lehrer and Eva Feder Kittay. 75-102. Hillsdale. New Jersey: Lawrence.
- Huang, Chu-Ren, Kathleen Athens, Li-Li Chang, Keh-Jiann Chen, Mei-Chun Liu, and Mei-Chih Tsai. 2000. The Module-Attribute Representation of Verbal Semantics: From Semantics to Argument Structure. *International Journal of Computational Linguistics and Chinese Language Processing*. 5(1).19-46. Also appeared in *Proceedings of the Symposium on Selected NSC Projects in General Linguistics from 1998-2000*. 119-46. 2001.
- Jackendoff, R. S. 1990. *Semantic Structures*. Cambridge: MIT Press.
- Levin, Beth. 1993. *Verb Classes and Alternation*. Chicago: University of Chicago Press.
- Levin, Beth, and Malka Rappaport Hovav. 1996. From Lexical Semantics to Argument Realization. Unpublished MS. Northwestern University and Bar Ilan University. A revised and expanded version will appear in the Cambridge Research Surveys in Linguistics Series.
- Liu, Mei-Chun. 2000. Categorical Structure and Semantic Representation: Mandarin Verbs of Communication. Paper presented at the 5th Conference on Conceptual Structure, Discourse and Language. University of California, Santa Barbara.
- Liu, Mei-Chun. 2002. Corpus-based Lexical Semantic Study of Verbs of Doubt: *Huayi* and *Cai* in Mandarin. *Concentric*. 28.2.
- Meng, Cong, Huai-De Zheng, Qing-Hai Meng, and Wen-Lan Tsai (eds.) 1985. Verb Usage Dictionary (*Dongci Yongfa Cidian*, 動詞用法詞典). Shanghai: Shanghai

Cishu Chubanshe.

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

Teng, Shou-Hsin (ed.) 1994 *Chinese Synonyms Usage Dictionary*. Taipei: Crane Publishing Co.

Tenny, Carol. 1992. The Aspectual Interface Hypothesis. *Lexical Matters*, ed. by Ivan A. Sag and Anna Szabolcsi. 1-27. Stanford: CSLI.

Tsai, Mei-Chih, Chu-Ren Huang, Keh-Jiann Chen, and Kathleen Ahrens. 1998. Towards a Representation of Verbal Semantics. *International Journal of Computational Linguistics and Chinese Language Processing*. 3(1).61-74.

Yu, Shi-wen, Xue-feng Zhu, Hui Wang, and Yun-yun Chang. 1998. *The Grammatical Knowledge-Base of Contemporary Chinese: A Complete Specification* (現代漢語語法信息詞典詳解). Beijing: Qing Hua University.

Website Resources

FrameNet. <http://www.icsi.berkeley.edu/~framenet/>

HowNet (知網). http://www.keenage.com/html/c_index.html/

Sinica Corpus (中央研究院現代漢語平衡語料庫).
<http://www.sinica.edu.tw/ftms-bin/kiwi.sh/>