

## 1. Abstracts

This work presents a corpus-based approach to the lexical semantic study of the major classes of Mandarin verbs. The distinct morphosyntactic behaviors of verbs provide revealing indications on their distinct lexical properties. As a pilot effort, a number of near-synonym sets which share the same semantic fields were first investigated. In order to account for the observed differences, a newly-developed framework (Huang and Tsai 1997; Liu, to appear) is adopted with the notion of ‘event focus’ and its implication on ‘event-structure attributes’. This research aims to show that a semantically-constrained framework of event-structure attributes is needed to make sense of the crucial distributional facts in lexical differentiation.

中文動詞詞彙語意之研究一直缺乏全面、細微、且有系統之成果。在「詞義影響語法表現」的前提下，本研究嘗試以語料庫為依據，由「近義詞」著手，來界定動詞最基本的、具有語法表徵的詞彙語意因素。

## 2. Background Introduction and Goal of the Research

### 2.1 Verbal Semantics

A recent focus of linguistic studies has been on the area of 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 (cf. Levin 1993, Pustejovsky 1995, 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-based, this study focuses more on corpus-based morpho-syntactic behavior as an indicator of lexical-semantic attributes.

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 1994), or a hybrid of event types and thematic roles (CKIP 1988). Given the typological and parametric variations between languages, some of the frameworks used for English cannot be readily transferred to Chinese. Liu (1996b) found that purely alternation-based approach may not be adequate in categorizing and representing Mandarin verbs. A more **semantically constrained** system is indeed needed for natural language processing purposes. This study thus aims to provide

detailed analyses of finer semantic distinctions to prepare for a complete representation of Mandarin verbal semantics.

## **2.2 Study of Near-Synonyms**

As a response to the need of fine-tuning verbal semantics, Tsai, Huang, and Chen (1996) presents an interesting work on differentiating a pair of near-synonyms - 高興 'happy, glad' and 快樂 'happy, joyful'. These two verbs are semantically similar but syntactically distinct in many aspects. By examining the correlation between their syntactic behavior 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>. Similar accounts can also be extended to the semantic distinction of near-synonym pairs in English and French.

As part of a long-term project of lexical semantic studies of Mandarin verbs, the present work extends the frontier to a new semantic field with other sets of contrastive near-synonyms. It is believed that only a comprehensive corpus-based study of these terms can render significant contrasts that help to differentiate their unique meanings.

The observed distinction among the four verbs is then viewed from a recently proposed framework that takes *event structure attributes* as the primary defining mechanisms for lexical semantic contrast (Haung and Tsai 1997). According to Huang and Tsai, there are basically four types of event structure attributes: Aspectual Attributes (i.e., generalized semantic properties pertaining to aspectual composition); Inherent Attributes (i.e., event-internal semantic focus); Role Attributes (i.e., salient role types); and Role-Internal Attributes (i.e. nominal characteristics of the roles). It is through the characterization of these four types of attributes that the verbs under study can be best differentiated.

## **2.3 The Data**

The data for the analysis of 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 word, developed by the CKIP group in Academia Sinica, Taiwan. The relevant data were extracted from the corpus by a key-word search with 30 additional words on either side.

## **3. Outline of Results**

### **3.1 Preliminary Observation**

**Clustering of Semantic Features and Syntactic Patterns:** +: lexically specified

F1: structure/design	[+design]	[-design]
Interpretational Contrast +Orientation	擺棋子/擺/子 擺東邊/邊東擺	/ 棋子// / 子 */ 東邊/*邊東/
F2: spatial boundary	[-bounded]	[+bounded]
	*擺邊 (0%) *擺在口袋中 (4%)	/ 邊壺內 / 在口袋中 (18%)
F3: process	[+process]	[-process]
	正在擺書	?正在/ 書
F4: result-state	[+result-state]	[-result]
	擺地攤/擺出一桌菜	*/ 地攤/*/ 出一桌菜

	菜	菜
F1: [causative]	[+causative]	[-causative]
	v 出兩張/子	** 出兩張/子
F2:[detach]	[+detach]	[-detach]
+* 來 +Residence	v * 來 v *	** * 來 ** *
Interpretational contrast	石頭v 了	石頭* 了
F3:[control]	[+control]	[-control]
Imperative+agentivity Natural phenomena	快了v ! *物換星v	*快了* ! 物換星* /太陽向西* (*v )
F4: salient role	object-patient	theme
	Object present (85%)	1 argument (80%)

	g	g
F 1: control and cause	[+contro][+causative]	[-contro][ -causative]
Inanimate subject Deliberate	能改就改 *天氣改了 *我不小心改了作文	*趕快變黑 天氣變了
F2: direction of change	[+positive]	
	*毛病改了 情況改好(*好)了	情況變好(好)了
F3: aspect	+process	-process (achievement)
	衣服改了三個小時	*天氣變了三小時

Note: 變2 (變魔術)

[+process]

	術	術
F1: ownership specified?	[+ownership]	[-ownership]
	搶銀行 *搶行	*行銀行 行行
F2: aspect	+process	-process (+achievement)
Progressive	他們正在搶刀子	*他們正在行刀子
Goal obtained	他在比賽中搶金	他在比賽中行金

	術	金
F1: [ownership]	[+ownership]	
	搶銀行	金東西
F2: manner vs. activity	more eventive	manner-focused
	*搶搶地	金金地
	不金搶了(stop the doing)	?不金金了(stop the habit)
	大* 開始「搶」吃餅乾	??大* 開始金吃餅乾

	乾	乾
F1: endpoint	+ result-theme	+result-state
Incremental theme as object	擠果汁	?
+著	擠著 (0%)	汁著 (7%)
F2: force type	two-sided force	one-sided force
	擠牙膏	*汁牙膏

	膏	乾
F1: body part	+ body part	- body part

	膏	膏
F1: instantaneous	+ instantaneous	-instantaneous
+著	*碰著鼻子	鼻著鼻子
single-point contact	碰在爐台上 cf. 上/上在頭上	*鼻在頭上 *上/上/在頭上
F2: impact	+impact	-impact

strong affect	碰上/上	鼻上
---------------	------	----

### 3.2 Distinctions in terms of Event-Structure Attributes

Four types of event-structure attributes are proposed as the basis for verbal semantic description and representation (Huang and Tsai 1997):

**Aspectual Attributes:** attributes pertaining to the composition of the event(s), such as Endpoint(s), Homogeneity, etc.

**Inherent Attributes:** attributes referring to the semantics of the event itself, such as Control, Change-of-state, etc.

**Role Attributes:** attributes referring to focussed (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 focussed role (of the event), such as Factive, Generic, Sentience, Volition, Affectedness, etc.

### 3.3 An example:

#### A. Initial Observation on 上 vs. 上

(1) Interpretational Difference:

- a. 投球、丟球
- b. 投丟、丟丟

(2) Goal as Direct Object:

- a. 投籃、投湖自盡、投共、投其所好
- b. \*丟籃

(3) Typical Manner or Result: 亂丟、丟掉

#### → Tentative Hypothesis:

投 (掉)	+endpoint
丟 (掉)	-endpoint (unspecified)

#### B. Further Contrast between 上/掉

(4) Goal as Direct Object

- a. 掉地有聲、墨球掉遠

(5) Directed motion (+path)遠

	邊 (進)	向 (邊、))	在 ( ))	)
投	76%	13%	10%	)
掉	6%	87%	6%	0%
丟	26%	19%	32%	23%
掉	4%	38%	50%	7%

(6) V-V: 投掉、掉掉、丟掉、)掉

→ Distinction between 上/掉:

上	+endpoint; +bounded
掉	+endpoint

C. Further Contrast between 上/)

(7) Interpretational Difference:

- a. 我丟了一支鋼筆
  - 掉了 (inchoative, stative, +result, -deliberate)
  - 掉了 (completive, active, -result, +deliberate)
- b. 我掉了一支鋼筆 (completive, active, +deliberate)

- (8) a. 我的鋼筆筆丟/筆丟了
- b. \*我的鋼筆筆掉/筆掉了

→ Distinction between 上 1/上 2/):

上 1 = )	-endpoint; +motion; +control
上 2	-endpoint; +state; -control; +result (achievement?)

D. Summary:

The distinction among the four verbs, 投, 掉, 丟, 掉 can be re-defined with the proposed event-structure attributes:

- In terms of Aspectual Attributes, 丟 differs from the others (activity events) in that it may focus on the event-endpoint, thus rendering an intransitive-causative use.
- In terms of Inherent Attributes, 投/掉 behave differently from 丟/掉 in that 投/掉 are highly directional, while 丟/掉 is underspecified in directionality.
- In terms of Role Attributes, 投/掉 can both take a Path-endpoint as the direct object,

while the role of Path-endpoint is not salient in the meaning of 丟/掉.

- With regard to Role-internal Attributes, 投 casts a further specification on the spatial characteristics of the Path-endpoint: it has to be bounded .

(9) Summary of Lexical-Semantic Distinctions among 投, 掉, 丟, 掉 with Event-Structure Attributes

Attributes	上	掉	上	)
Aspectual	- Event-endpoint	- Event-endpoint	+ Event-endpoint	- Event-endpoint
Internal	+ Directional	+ Directional	- Directional	- Directional
Role	+ Path-endpoint	+ Path-endpoint	- Path-endpoint	- Path-endpoint
Role-Internal (Path-Endpoint)	+ Bounded			

### 3.4 Other synonym sets

- 懸、掛、吊

**Aspectual Attributes :**

懸 : Stative 。明鏡高懸 。亭內懸一鐘 。

掛 : Punctual State 。畫一直掛在那兒 。她將衣服掛在這裡 。

吊 : Process 。校警隊正在吊車 。

**Role Attribute :**

掛 : Locus 。掛鐘 vs. 吊燈 。

- 捆、綁

**Aspectual Attributes :**

捆 : Process 。他捆了很久還沒捆好 。

綁 : Process + Endpoint 。她正(在)把蝴蝶結綁到頭上 。

**Role Attributes :**

捆 : theme 。捆柴 。

綁 : a) Incremental Theme 。綁馬尾、綁票 。

b) Locus vs. Locutum 。白布條綁頭 vs. 頭綁白布條 。

**cf. Theme/ Location Alternation :**

I sprayed the paint on the wall.

I sprayed the wall with paint.

● 束、縛、紮

**Aspectual Attributes :**

束：Punctual。\*他一直在束。\*他束了好久都沒束好。

縛：Process。米篩下垂的一邊縛著根小棒。媽媽正在縛粽子。

紮：Process。她正在紮辮子。

**Role Attributes :**

束：Theme but not Incremental Theme。束手就擒。

縛：Theme but not Incremental Theme。縛粽子、手無縛雞之力。

紮：Theme or Incremental Theme。在頭上紮上白緞帶。紮馬尾。

● 擠、壓、按

**Aspectual Attributes:**

擠：Process。他一直擠。她正在擠牛奶。

壓：Punctual State。書壓在桌上。用力壓了下去。

按：Puncture。按了很久都沒人應門。

**Role Attributes :**

擠：Theme or Incremental Theme。人擠人。擠果汁、擠牛奶。

壓：Theme but not Incremental Theme。壓花、\*壓果汁。

按：Theme。按按鈕。

● 呼、喊、叫、喚、召

**Aspectual Attributes :**

呼：Process。他一直呼著我的名。

喊：Process。他喊著喊著就睡著了。

叫：Process。他一直在叫。

喚：Process。媽媽喚了我三次。

召：Puncture。召了岳飛來。\*召了三個鐘頭。

**Role Attributes :**

呼：Theme、Incremental Theme or Goal。呼朋引伴。呼口號。呼天搶地。

喊：Theme or Goal。喊他'老林'。呼天喊地。

叫：Theme、Goal or Locus。叫車。叫門、叫陣。叫床。

喚：Theme。他喚人拿錢來。

召：Theme。沙皇召宗子文回國。

● 捏、夾

**Aspectual Attributes :**

捏：Process。他捏了好久才捏出一隻老虎。他一直捏我。



夾：Puncture State。他一筷子夾了下去。手夾在門縫裡。

**Role Attributes：**

捏：Theme or Incremental Theme。捏臉頰。捏麵人。

夾：Theme。夾頭髮、夾娃娃。

● 容易、簡單

簡單：NP-predicating

容易：VP-predicating

● 追趕、跟隨

**Aspectual Attributes:**

追趕：[stage]

追趕者與被追趕者之間的距離會隨時間逐漸拉近或拉遠。

例：比數形成十一比二，令韓國隊 追趕 無望。

\*跟隨

跟隨：[process]

跟隨者與被跟隨者的距離保持一定。

例：米食 跟隨 我國走過了五千年的歲月。

\*追趕

● 追、趕

**Role Internal Attributes:**

追：goal [moving]

例：追小偷

**趕 1：[Temporal target]**

動詞後賓語會從名詞演變成事件，再引伸為事件發生的時間。

例：趕國科會報告/趕飛機/趕火車/趕集/趕廟會。

**趕 2：[Patient]**

例：趕鴨子/趕牛/趕羊。

**趕 3：[Manner]**

例：我很趕。

**追 vs. 趕**

例：正在 追/趕 公車。

追：[Spatial event] 公車是所追的實體。

趕：[Temporal event] 同趕 1，公車已轉化為趕的原因。

● 清楚、明白

Role Attributes:

清楚：Nominative--[+theme] [+experiencer]

明白：Nominative--[+theme] [+experiencer]

● 跳、躍

Aspectual Attributes:

跳 1. [+process] 在水裡跑，在水裡跳。

孫悟空在她肚子裡又叫又跳。

心跳

跳舞

跳繩

跳 2. [+punctual] 他整個人像觸電一樣跳了起來。

絕大部份選手只懂得拿著撐竿就跳。

跳車

跳水

躍 1. [+process] 雀躍

躍 2. [+punctual] 鯉魚躍龍門

躍牆而過

註：基本上，躍和跳的語意接近，躍多用在較文言的詞組中。

#### 4. Conclusion

The set of the four Mandarin near-synonyms under study raises several important issues concerning the approach for lexical semantic research:

- While some works on English verbal semantics (e.g. Levin 1993, Atkins and Levin 1991, Atkins et. al. 1988) conclude that *diathesis alternations* are most useful in identifying crucial semantic-syntactic interdependencies, such an approach may not be adequate when applying to Mandarin, given that Mandarin is relatively flexible in argument placing.
- Viewed from a language-specific perspective, this study may be taken as part of the pilot efforts in searching for the most suitable and effective framework for the study of the Mandarin verbal system.
- Viewed in a more general context, this work may help to illustrate several theoretical and methodological points. First, 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 focus is taken into consideration and if verbal

meanings are represented in terms of specific categories of attributes. Finally, the clustering of discourse-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.

## 5. Self-Evaluation

**Three papers have been completed and as a direct result of the research:**

To appear. Lexical meaning and discourse patterning - the three Mandarin cases of 'build'. In the 3<sup>rd</sup> Volume of Conceptual Structure, Discourse, and Language. Stanford: CSLI.

1998. When Endpoint Meets Endpoint: A Corpus-based Lexical Semantic Study of Mandarin Verbs of *Throwing*. Presented at ICAL-7/NACCL-10, Stanford University.

1998. Lexical Information and Beyond: Constructional Inferences in Semantic Representation, submitted to Pacling 13, National Cheng-Kung University.

## 6. References

Atkins, B.T. and Levin, B. 1991. Admitting Impediments in Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon, ed. by U. Zernik. Hillsdale, NJ: Lawrence Erlbaum Associates.

Atkins, B.T., J. Kegl, and B. Levin. 1988. Anatomy of a verb entry: from linguistic theory to lexicographic practice. *International Journal of Lexicography* 1: 84-126.

Biber, D. 1996. Investigating language use through corpus-based analyses of association patterns. *International Journal of Corpus Linguistics* Vol. 1(2): 171-197.

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

and Oriental Language (ICCPCOL), 414-17. Toronto, Canada.

Croft, William, ed. 1991. *Syntactic Categories and Grammatical Relations*. Chicago: University of Chicago Press.

----- . 1990. Possible verbs and the structure of events. In *meanings and Prototypes: Studies in Linguistic Categorization*, ed. by S. Tsohatzidis. London/New York: Routledge. 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: The Univ. of Chicago Press.

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

Her, One Soon. 1990. *Grammatical Functions and Verb Subcategorization in Mandarin Chinese*. Taipei: Crane Publishing Co.

Hopper, Paul, and Thompson, Sandra. 1980. Transitivity in grammar and discourse. *Language* 56: 250-299.

Huang, Chu-ren. and Tsai, Mei-chih. 1997. From Near Synonyms to Event Structure: Corpus-based Studies of Mandarin Verbal Semantics. Paper presented at the Mini-Conference on Lexical Semantics. Graduate Institute of Linguistics. National Chung Cheng University.

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

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

Lakoff, George. 1987. *Women, Fire, and Dangerous Things*. Chicago: U of Chicago Press.

Langacker, R. 1987. *Foundations of Cognitive Grammar*, vol 1. Stanford: Stanford U Press.

----- . 1991. *Foundations of Cognitive Grammar*, vol. 2. Stanford U, Press.

- Levin, Beth. 1993. *Verb Classes and Alternation*. Chicago: U of Chicago Press.
- Levin, Beth, Grace Song, and B.T.S. Atkins. 1997. Making sense of corpus data: A case study of verbs of sound, " *International Journal of Corpus Linguistics* 2 (1): 23-64.
- Liu, Meichun. To appear. Lexical meaning and discourse patterning - the three Mandarin cases of 'build'. In the 3<sup>rd</sup> Volume of *Conceptual Structure, Discourse, and Language*. Stanford: CSLI.
- . 1997. Conceptual basis and categorial structure: a study of Mandarin V-R compounds as a radial category. In *Chinese Languages and Linguistics, Vol. IV*: 425-51. Taipei: Academia Sinica.
- . 1996a. Conceptual manipulation and semantic distinction: the case of Mandarin postverbal DE complement. In *Conceptual Structure, Discourse, and Language*, ed. by Adele Goldberg, 333-46. Stanford: CSLI.
- . 1996b. A Pilot Study on Chinese Verb Classes and Alternations. NSC Project Report (NSC85-2418-H-009-003). Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge: The MIT Press.
- . 1991. The syntax of event structure. *Cognition* 41:47-81. Smith, C. S. 1991. *The Parameter of Aspect*. Dordrecht: Kluwer Academic Publishers.
- Tenny, C. 1992. "The Aspectual Interface Hypothesis". *Lexical Matters*, ed. Ivan A. Sag and Anna Szabolcsi, 1-27. Stanford: Stanford University.
- Teng, Shou-hsin, ed. 1994. *Chinese Synonyms Usage Dictionary*. Taipei: Crane Publishing Co.
- Tsai, Mei-chi, Chu-ren Huang, and Keh-jiann Chen. 1996. You jinyici bianyi biao zhun kan yuyi, qufa zhi hudong. (From near-synonyms to the interaction between syntax and semantics), paper presented at IsCLL-5, Taipei, Taiwan.
- Tsai, Mei-chi, Chu-Ren Huang, Keh-jiann Chen, Kathleen Ahrens. 1997. Towards a representation of verbal semantics - an approach based on near synonyms. *Proceedings of ROCLING X*, 34-48, Hsinchu, Taiwan.

Tsao, Feng-fu. 1994. On verb classification in Chinese. Paper presented in the 6th North American Conference on Chinese Linguistics (NACCL VI).