The Construction of Mandarin VerbNet:
A Frame-based Study of Statement Verbs

Mei-chun Liu and Ting-yi Chiang
National Chiao Tung University

This paper introduces the infrastructure and the principles of a semantic framework used for the analysis and classification of verbs, developed with the aim of constructing a lexical database of Mandarin verbal semantics, called the Mandarin VerbNet. Distinct from most existing lexical databases that enumerate word senses without detailed grammatical considerations, the Mandarin VerbNet is designed to *provide lexical semantic information based on grammatical descriptions and anchored in linguistic theories*. It looks for systematic correlations between syntax and semantics and classifies verbs according to these syntax-to-semantics correspondences. The framework adopts the approach of Frame Semantics (Fillmore & Atkins 1992) in defining verb meanings in a semantic frame and building a frame-based verbal lexicon, but it differs from the structure of the English FrameNet in distinguishing different scopes of frames. Evolved and refined from previous works (Liu, Chiang & Chang 2004, Liu & Wu 2003, Liu 2002), this study summarizes the current model of the analytic framework with a detailed illustration from Mandarin statement verbs. It ultimately seeks to identify a theoretically sound and operationally effective representational scheme that bases its semantic analysis on grammatical behaviors and provides linguistic motivations for its semantic classifications.

Key words: lexical semantics, Mandarin VerbNet, frame semantics, statement verbs, syntax-semantic interaction

1. Introduction

1.1 Concerns in lexical semantics

The study of lexical semantics has become a focal area in linguistic research as lexicon is considered to provide key information for grammatical structure as well as language processing. The study of verbs, in particular, has drawn substantial attention in the development of linguistic theories and frameworks since verbs play a crucial role in determining sentence patterns (Levin 1993, Jackendoff 1983). In Mandarin, a series of studies are dedicated to the distinction of verb meanings based on corpus data (Huang et

al. 2000, Chief et al. 2000, Chang et al. 2000a, 2000b, Liu 1999, Liu 2002, Liu & Wu 2003). More recently, a comprehensive program has been launched to capture the semantic granularity of Mandarin verb types (Liu & Wu 2003, Liu, Chang & Chiang 2004, Liu & Chang 2005). It aims to offer a systematic and principled framework for detecting verbal semantics and herewith classifying verbs into semantic categories. In doing so, three questions need to be tackled first: (1) How do we define verb meaning? (2) What are the criteria for differentiating verb meanings? (3) What is a useful and meaningful taxonomy for verb classification?

1.2 A Frame-based approach

To answer the first question, we adopt the perspective in cognitive semantics or more specifically, Frame Semantics, that "a word's meaning can be understood only with reference to a structured background of experience, beliefs, or practices, constituting a kind of conceptual prerequisite for understanding the meaning" (Fillmore & Atkins 1992:76-77). The meaning of verb is therefore defined in relation to a **semantic frame** that provides the conceptual prerequisite for anchoring the verb. Each frame characterizes uniquely the semantic background for one particular sense. Hence verbs with multiple senses belong to multiple frames.

As for the criteria for semantic distinction, we take the mainstream view expressed overtly in Levin (1993) that the meaning of a verb determines the structure of a sentence. In other words, semantic differences give rise to grammatical differences and hence grammatical observations serve as the basis for semantic distinctions. Only the semantic factors that are grammatically relevant will be taken into consideration. Under this principle, the framework is meant to provide a linguistically motivated database that shows syntax-to-semantics correspondences (cf. Levin & Rappaport Hovav 1996).

With regard to the third question of classificational scheme, a frame-based model has been developed, modifying the design of English FrameNet (http://framenet.icsi. berkeley.edu/), as an attempt to construct a comprehensive lexical database called the Mandarin VerbNet. It aims to capture the semantic granularity of distinguishable verb types and represent the relevant information in an analytic framework. Among the current models of lexically based networks, such as WordNet (Miller et al. 1990), EuroWordNet (Alonge et al. 1998), HowNet (Dong & Dong 2001, Dong, Hao & Dong 2004) and ontological hierarchies such as SUMO (Pease et al. 2002, Niles & Pease 2001), the English FrameNet (Baker, Fillmore & Lowe 1998) is the major one that utilizes a linguistic theory in constructing its lexical database, namely, the approach of Frame Semantics (Fillmore & Atkins 1992). And our study also utilizes the same approach in delimiting frame-based lexical information by specifying frame-specific participant

roles, called Frame Elements (FEs), and the grammatical patterns expressing the Frame Elements.

However, a purely frame-based approach may still pose two problems. First, the various frames used to categorize verbs need themselves to be defined. Secondly, the single concept of "frame" may be inadequate to capture the different semantic scopes a verb may be anchored in, as well as the varied hierarchical relations among them. In regard to the first problem, Liu & Wu (2003) attempt to characterize a given frame with a specific image schema (cf. Sweetser 1990), that provides the conceptual link between core frame elements. By providing a cognitive schema as a macro-structure, the distinction of frames is then well motivated, as pointed out by Liu & Wu (2003):

> By proposing such an overarching cognitive model, this study is able to explain independently the motivations for the proposed frames and account for the semantic distinctions and interrelationships among them. It also helps decipher the aspectual differences and collocational patterns associated with each frame.

1.3 Frame-to-frame hierarchy

As for the scope of semantic frames, although frames may be of various scopes and associated with different numbers of lexical entities, no distinction of frame types are made in FrameNet. The interrelations of frames are expressed in a graphical chart showing the links among various frames in eight kinds of frame-to-frame relations. Liu & Chang (2005) proposed that the notion of "subframes" should be adopted to implement a finer classification of verb meanings. Subframes may be cross-framed with multiple inheritance, i.e. inheriting meanings from two or more frames. In our current design, to capture the different scopes of semantic specification and degrees of lexical granularity inherent in different frames, the latter can be further grouped into four types according to their semantic scopes and the relative importance of their associated lemma. An *archi-frame* corresponds to a broad semantic domain, providing the maximal scope of background information for a unique event type (e.g. COMMUNICATION archi-frame for all communication-related senses with the representative lemma 溝通 'communicate'); a *primary frame* represents a major relational subpart of an archi-frame, corresponding to primary categories in cognition with frequently used or generic lemma (e.g. Conversation frame represents a reciprocal directionality in communication, with the frequent and generic lemma 談 'talk'); a *basic frame* highlights a particular participant role or relation within a primary frame (e.g. Discuss frame foregrounds the role of Topic in conversation, with the representative lemma 討論/商量 'discuss, talk over'). A basic frame can be

further divided into *micro-frames* with a finer specification of role-internal features (e.g. 協商 'Discuss-negotiate' vs. 'discuss-explore' 研討 may be two different microframes under the 'Discuss' frame, having different semantic restrictions on topics and purposes).

The differentiations of the four scopes are by no means inherently determined or hierarchically fixed. They simply serve to differentiate and structure the varied scopes of frames within a larger domain. A frame may be primary, basic, or micro, in relation to other distinguished frames in the whole domain. In other words, the possible levels of frames are determined relative to the levels of semantic specificities or degrees of semantic granularity lexicalized in a language. The largest semantic scope, the archi-frame, corresponds to an intuitively broad and independent domain with a relatively uniformed cognitive basis. They are identified as a convenient means to link with a possible upper semantic ontology. Again, all the semantic classifications are drawn on syntactically observable criteria.

The aim of this paper is to introduce the infrastructure and the principles of constructing the Mandarin lexical database of verbs, the Mandarin VerbNet, with a clear illustration from the analysis of statement verbs. It is meant to provide a linguistically motivated database with detailed lexical semantic information on the basis of grammatical and collocational analysis. The details of the framework are introduced below.

2. The classificational framework of Mandarin VerbNet

2.1 Principles for analysis

The Mandarin VerbNet is primarily a lexical database of verb classes with key semantic and syntactic information of each class. It involves the classification of verbs and the distinction of senses. Both are done in a consistent and systematic manner, observing the following principles:

  i.   Frames are of various semantic scopes. A hierarchically-structured, multi-layered working taxonomy is adopted to capture the various scopes of semantic frames: Archi-frame > primary frame > basic frame > micro-frames. The smaller-scoped or lower-layered frames can be considered subframes of the higher-layered ones. Each frame is specified with a descriptive definition, a set of core vs. non-core Frame Elements, Conceptual Schema, Basic Patterns (grammatical expression of core arguments), Collocational Associations (grammatical or lexical collocations of non-core arguments), and Semantic Attributes (distinctions pertaining to event structure, participant roles or discourse context).
  ii. Archi-frames are distinguished along a relatively broad and self-containing conceptual schema with a default set of frame elements. Primary frames are

defined as a subpart of the archi-schema with a unique set of core frame elements. Basic frames are distinguished according to syntactically expressed patterns of foregrounding or backgrounding certain frame elements; Micro-frames are distinguished according to role-internal specifications of frame elements.

iii. It attempts to define frames, not just with a description and a list of frame elements, but with a schematic representation as the conceptual motivation to characterize the cognitive basis for a specific frame and the interrelations between its subparts (Liu & Wu 2003).

iv. All the semantic classifications of verbs are made according to syntactically attested distinctions, including grammatical patterns of argument expression (Basic Patterns, BPs) or aspectual, lexical or adjunct collocations (Collocational Associations, CAs).

v. Sense distinction of polysemy follows the principle: one sense, one frame. Therefore, verbs with multiple senses belong to multiple frames. Semantically similar lemmas, or near-synonyms, are defined as sharing the same semantic frame. Near-synonym sets may be anchored at various levels of the frame hierarchy, depending on their semantic granularity and the frame-to-frame distance between them. The smaller scope their semantic anchor is, the closer they may be semantically related.[1]

vi. Semantic inheritance exists from top to bottom in the hierarchical structure. There are basically two types of frame relations: Inheritance and Using. Frames may have multiple inheritances from other frames, e.g., a given micro-frame may inherit features from a number of basic, larger-scope frames.

As in FrameNet, the term 'frame' is used in a rather generic sense, referring to all kinds of semantic scopes serving as the cognitive prerequisites for lexical senses. As mentioned above, in order to be consistent in the analysis, clearly defined operational principles are applied to distinguishing the different types of frames. An archi-frame is the maximal scope of an eventive background, characterized schematically with a Conceptual Schema (CS), plotted with a set of default participants, the Frame Elements (FEs); a primary frame is characterized as a subpart of the archi-frame with a set of focused FEs, expressed in grammatical templates, the Basic Patterns (BPs). A basic frame is cognitively more salient and associated with basic-level words; it further highlights (or suppresses) a certain FE out of the primary set, accompanied with highly-specified BPs. The smallest scope, a microframe, is the level where role-internal (or FE-internal) specifications or requirements are found, while syntactically it may

---

[1] See §2.4 below for Liu (2003)'s study on Mandarin near-synonyms of discussion, *taolun* 討論 vs. *shangliang* 商量.

Mei-chun Liu and Ting-yi Chiang

inherit the BPs from a basic frame. Below is a detailed account of the possible layers of the hierarchy, using the Communication domain as an example (cf. Liu & Wu 2003).

2.2 Layer 1: Archi-frame (e.g., **COMMUNICATION** frame)

An Archi-frame is at the top of the semantic hierarchy. It corresponds to a relatively large and independent semantic domain, similar to a higher-level node in a semantic ontology. An archi-frame provides an overarching conceptual schema as the semantic prerequisite for the individual frames within a relatively large and independent domain. For instance, the Communication archi-frame, with its representative lemma *communicate* in English or *goutong* 溝通 'communicate' in Mandarin, involves a conceptual schema with a set of default participants pertaining to all communication-related events. The archiframe-associated lemma *goutong* 溝通 'communicate,' is then lexicalized with the least restricted meaning and the widest semantic scope, built upon such a background understanding. Other possible archi-frames include: Cognition archi-frame (representative lemma *sikao* 思考 'cognize, think'); Perception archi-frame (*ganzhi* 感知 'perceive' and Internal-Sensation archi-frame (*juede* 覺得 'feel'). As illustrated below, the Communication archi-frame can be represented in a conceptual schema with a set of frame-specific elements (i.e. prototypical participants) and the default relations between them. The schema serves as the cognitive basis to motivate other more focused communication frames. The frame-selected core elements in the schema may be encoded as core arguments in frequently associated grammatical patterns, as exemplified below:

(1) Conceptual Schema of the Communication Archi-frame:

244
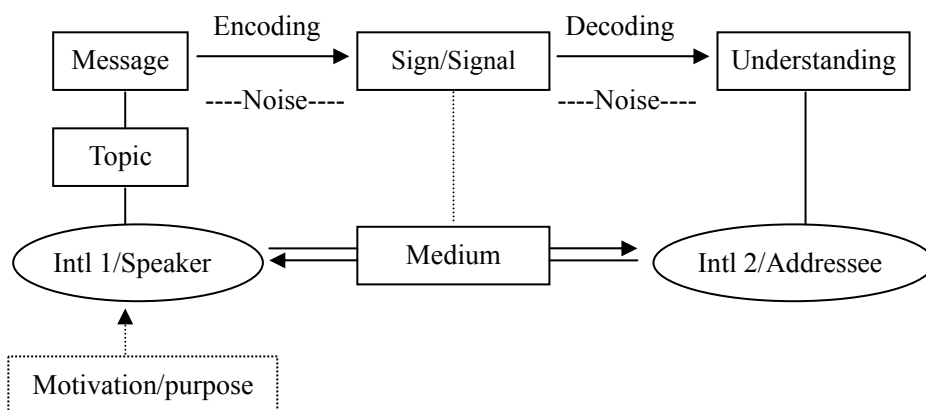
Representative Lemma: *goutong* 溝通 'communicate'

Basic Patterns: a. [Interlocutor1] and [Interlocutor 2] > V > [Topic]

他[Interlocutor1]和同事[Interlocutor2] **溝通** 工作的事[Topic]。

'He communicates with his colleagues about work.'

b. [Speaker] > V > [Topic]

他[Speaker]不習慣 **溝通** 他的想法[Topic]。

'He is not used to communicating his thoughts.'

2.3 Layer 2: Primary Frame (STATEMENT vs. CONVERSATION vs. TELLING)

Under each archi-frame, Primary Frames (PMs) can be distinguished to profile a given portion of the schema. Conceptually, primary frames correspond to a more focused relation between a limited set of frame elements and can be defined with a more restricted set of syntactic expressions. For instance, three primary frames can be readily distinguished under the Communication archi-frame: the Statement frame, with its representative lemma *shuo/chenshu* 說/陳述 'say/state', focuses on the information-giving process where a Speaker gives out a Message on a Topic; the Conversation frame, with its representative lemma *tan/jiaotan* 談/交談 'talk, converse,' profiles the reciprocal relation between two or more Interlocutors as they exchange views on a given Topic; the Telling frame includes both the Addressee and the Message as internal arguments as it highlights the unidirectional communicative process where a Speaker informs an Addressee of a Message on a certain Topic (see the graphs in (2)-(4)).

The three primary frames differ in highlighting different portions of the schema, selecting different sets of core frame elements, assigning different roles, and expressing the different event types with different grammatical templates, the Basic Patterns (BPs). Verbs of a primary frame share the event structure anchored in the same subpart of the schema and thus share the same pattern of expressing the core arguments: the defining BP for the Statement frame is: Speaker > V > Topic/Message (他[Speaker]一再陳述[state]自己的主張[topic]/自己是無辜的[Message]); the defining BP for the Conversation frame is: Interlocutors > V > Topic (他們[Interlocutors]談論[Converse]明天的天氣[Topic]); and the defining BP for Telling frame is: Speaker > V > Addressee > Message (他[Speaker]告訴[tell]法官他是無辜的[Message]). Although all the frame elements in the Communication archi-frame constitute the underlying background for any communication-related event, primary frames only profile a rather focused set of elements out of the default set, depicting a major, distinguishable event type among the selected roles. The schematic anchors for the two primary frames of communication are shown below in the dotted area with profiled frame elements and the unique relation between them:

(2) Conceptual Schema of the Statement Primary Frame:
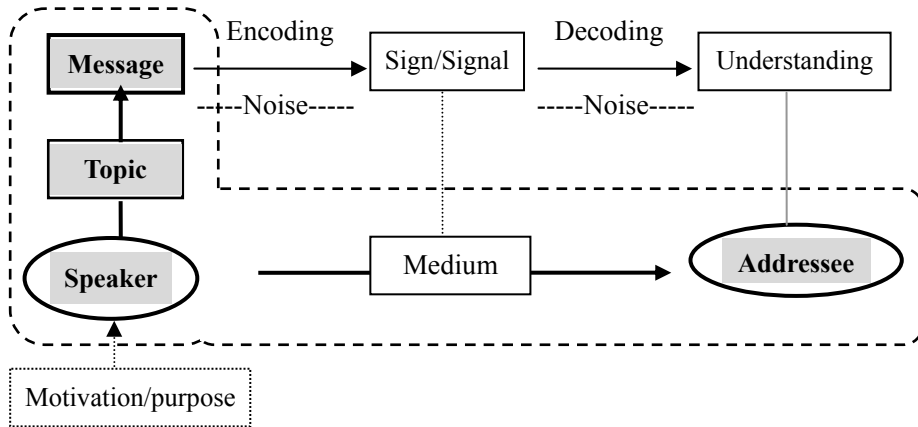Core Frame Elements: *Speaker*, *Topic*, *Message*
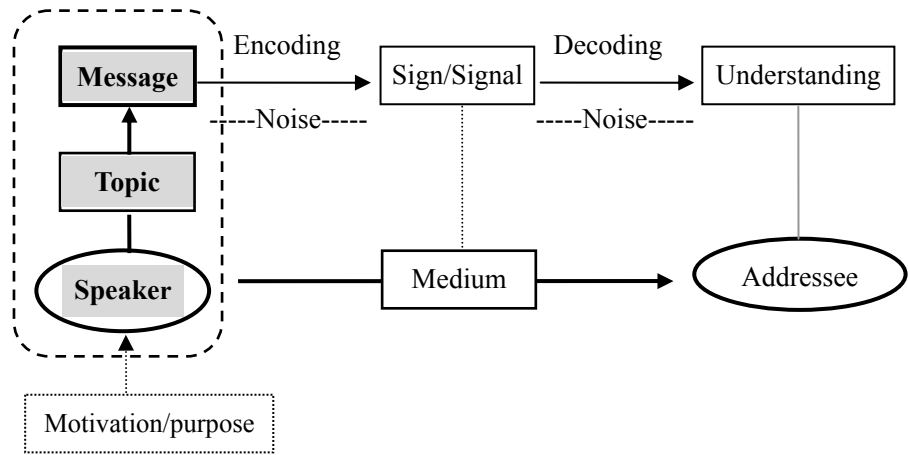
(3) Conceptual Schema of the Conversation primary frame: *Interlocutors*, *Topic*
Core Frame Elements: *Interlocutor1*, *Interlocuter2*, (*or Interlocutors*), *Topic*
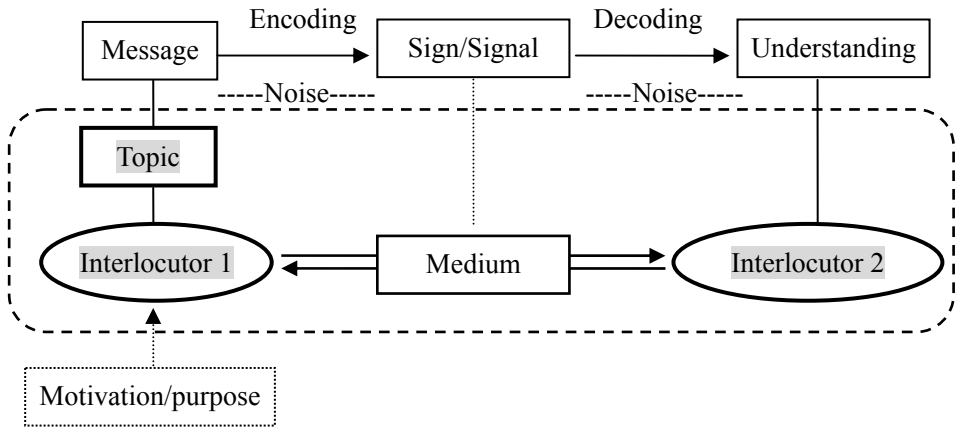
(4) Conceptual schema of the Telling primary frame:
     Core Frame Elements: *Speaker*, *Addressee*, *Topic*, *Message*

2.4 Layer 3: Basic Frame (e.g., Say frame vs. Speak frame)

Under each primary frame, semantically more restricted classes, called Basic Frames (BFs), may be further distinguished as they specify an even narrower scope of meaning. Basic frames correspond to the cognitively salient categories or the so-called basic-level categories (Rosch 1973, 1975a, 1975b). They are semantically more informative, distributionally more frequent and common, and acquisitionally easier and earlier. They are associated with foregrounded or backgrounded frame elements within the set of primary-selected elements. A detailed analysis of basic frames for Conversation verbs was given in Liu & Chang (2005), then called 'subframes' as a preliminary finding. In that study, five subclasses of the Conversation frame are distinguished, according to grammatical and collocational distributions: the 'talk' (or 'converse') frame, *tan*/*tanhua* 談/談話, foregrounds the Interlocutors and the Medium; the 'Discuss' frame, *taolun*/ *shangliang* 討論/商量, foregrounds the Topic role; the 'Negotiate' frame 協調/溝通[2] allows a third Interlocutor, the Moderator, to be present and the Interlocutors can be coded as the direct object. The 'Quarrel' frame *chao*/*chaojia* 吵/吵架 inherits Cause from the Hostile-encounter frame, allowing an additional core frame element. The 'chat' frame, *liaotian*/*xianliao* 聊天/閒聊, backgrounds the element Topic and prefers a human subject due to its informal register. The five 'subframes' of Conversation can now be considered as five basic frames within the Conversation primary frame.

---

[2] The narrower sense of *goutong* 溝通 only refers to the 'negotiate' sense, in contrast to the broader sense, representing the Communication archi-frame.

As for the Statement primary frame, a number of basic frames can be distinguished. Basic frames inherit the same set of Frame Elements from the primary frame, but foreground or background certain elements of the set. A detailed analysis and classification of the Statement verbs will be given in §3 below as a comprehensive illustration of the operational mechanisms of Mandarin VerbNet. Here, only two basic frames will be mentioned to exemplify the defining factors. One major Basic Frame under Statement is the *Say* frame, represented by the commonly occurring lemmas *shuo/jiang* 說/講. Although it inherits the Statement-related FEs as its background, the Say frame foregrounds a clausal Message in its basic pattern: 他[Speaker]說[Say]明天會下雨[Message]. The core frame element Message commonly occurs as a required complement, rendering the Say verbs as typical complement-taking verbs. In contrast, the Speak frame, represented by *shuohua/jianghua* 說話/講話 'speak' in Mandarin, syntactically suppresses the role Message as it is lexically incorporated and backgrounded. The element Message thus rarely occurs with Speak verbs. The difference in the two basic frames can be illustrated below:

(5) Foregrounding and Backgrounding of FEs in Basic Frames
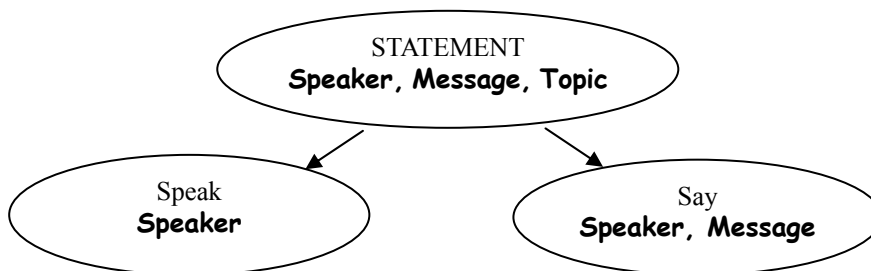
2.5 Layer 4 and below: Microframes (*Marked* vs. *Unmarked Speak* frames)

Within a basic frame, there may be smaller classes of verbs that require further distinctions. For example, Liu (2003) discusses Mandarin near-synonyms of discussion, *taolun* 討論 vs. *shangliang* 商量 and concludes that they differ in the semantic requirements on both the Topic-object (討論/*商量 語言學) and the Interlocutor-subject (董事們/?員工們 商量經營方針). According to the current framework, they belong to two different micro-frames and can be viewed as near-synonyms at the level of basic frames (the Discuss frame) since they share the same set of basic frame elements but differ in terms of role-internal features.

In the aforementioned Speak frame, there are near-synonym sets that can be classified into further distinguished micro-frames. While all the Speak verbs lexically incorporate a backgrounded Message, the more frequently used verbs, such as *shuohua/*

jianghua 說話/講話 'speak', are in contrast with the semantically loaded verbs, 訓話 'give a lecture' / 講道 'preach,' in terms of the relational restriction on the Speaker and Addressee, purpose and manner of speaking, and the range of possible registers. The *Marked* Speak frame is highly specified in all the above features, while the *Unmarked* Speak frame is lexically underspecified in the features. Compared to other scopes of frames, the micro-frames are more language-dependent as each language may vary in the ultimate range of semantic granularity encoded in the lexicon.

(6)   Different Frame-internal Features of Microframes

```
                    ┌──────────────────────────┐
                    │          Speak           │
                    │  Message-incorporated    │
                    └──────────────────────────┘
                       /                    \
```

Unmarked - *shuohua* 說話	Marked - *xunhua* 訓話
Speaker: unspecified	Speaker: higher status
Addressee: unspecified	Addressee: lower status
Purpose: unspecified	Purpose: lesson-giving
Manner: unspecified	Manner: serious
Register: unspecified	Register: formal
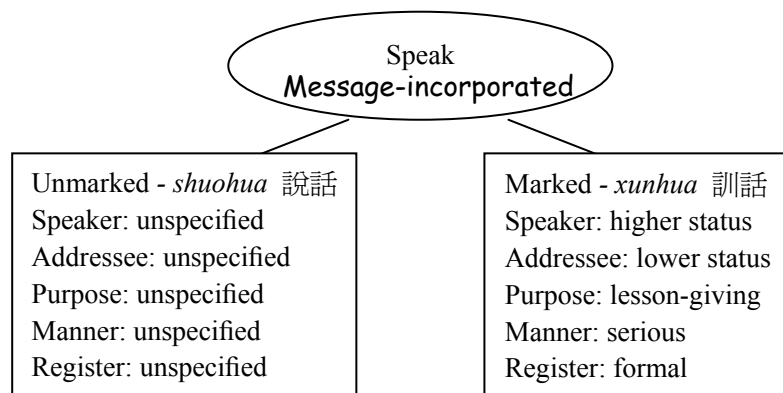
2.6 Inheritance and using relations among hierarchically structured frames

In FrameNet, there are eight different frame-to-frame relations (Ruppenhofer et al. 2006:8, 103-111), two of which are adopted in our model, i.e. inheritance and using relations. Inheritance refers to the part-whole or membership relation: A is a subtype of B. For instance, the Statement Primary Frame is a subtype and subpart of the COMMUNICATION archi-frame. Using is a relation of reference: A presupposes B as background. For instance, Communication_Truthfulness verbs such as 坦白/老實 'speak candidly' and Volubility verbs such as 冗長 'long-dragged' / 簡潔 'terse' all require the Communication frame as their background and semantic reference but do not constitute a direct subclass of the frame. Most stative or adverbial predicates will 'use' an eventive frame as their referential prerequisite.

The operational principles and the layered structure of frames, as explained above, constitute the underlying mechanisms for the ongoing project in constructing a comprehensive lexical semantic database of verb meanings, the Mandarin VerbNet. In the following section, verbs of statement will be used to illustrate the frame-based approach to syntax-to-semantic distinctions of verbs.

3. Frame-based analysis of the statement verbs

3.1 The statement frame

Liu (2003) pointed out that Communication is a cognitively important and independent domain that encodes one of the most fundamental aspects of human activities. The Communication archi-frame encompasses a number of distinct event types, including Statement (陳述類), Conversation (交談類), Telling (告訴類), Encoding (傳達類), Questioning (詢問類), Request (請求類), Responding (回應類), Communication_Means (打電話 'call', 傳真 'fax'), etc. Each event type is anchored in a primary frame that profiles a subpart of the background archi-frame. Among all the primary frames, the Statement frame is probably the largest with a wide variety of verbs describing a message-delivering process: *shuo* 說 'say', *jianghua* 講話 'speak', *biaoshi* 表示 'express', *jieshi* 解釋 'explain', *jianyi* 建議 'suggest', *chengren* 承認 'admit', *chengnuo* 承諾 'promise', *toulu* 透露 'reveal', *baoyuan* 抱怨 'complain', *pinglun* 評論 'comment', *baogao* 報告 'report' and many others. As schematically represented in the Conceptual Schema in (2) above, the statement frame profiles an information-giving process of a Speaker conveying a Message on a Topic, potentially to an Addressee and through a Medium. Verbs of the statement frame share the default set of core Frame Elements (FEs): Speaker, Message, Topic. As attested in the corpus data, a number of basic patterns of argument expressions (Basic Patterns, BPs) are found to encode the core frame elements. The Speaker tends to be the subject, taking a Message or Topic as an internal argument. The Message is typically a clausal complement,[3] or it may be nominalized as a Message-description; the Topic may be coded as the object, or preposed with a Topic-marker (*guanyu* 關於 'about'). Other frequently-collocated participants inherited from the communication schema, such as Addressee or Medium, may also occur with an overt marker:

(7)   Defining the Statement frame:
    (i)   Definition: a unidirectional information-delivering process whereby a Speaker conveys a Message on a Topic, potentially to an Addressee and through a Medium
    (ii)   Conceptual Schema: see (2) above
    (iii)   Core Frame Elements: Speaker, Message, Topic
        Frequently-collocated Elements: Addressee, Medium

---

[3]   In Givón's (1993) classification, statement verbs may be categorized into the Perception-Cognition-Utterance (PCU) category, typically taking a clausal complement.

(iv) Basic Patterns with Core Frame Elements:

    a. 他**[Speaker]** 說/表示 明天會下雨**[Message]**
      'He said that it would rain tomorrow.'

    b. 他[Speaker] 述說/表達 感謝之意**[Message-description]**
      'He expressed his gratitude.'

    c. 他[Speaker] 陳述/說明 事情的經過**[Topic]**
      'He recounted the incident.'

Basic Patterns with non-core Frame Elements:

    d. 他[Speaker] 向/對 法官**[Addressee]** 宣稱自己是無辜的[Message]
      'He claimed to the judge that he is innocent.'

    e. 他[Speaker] 透過/以 書面報告**[Medium]** 說明事情的經過[topic]
      'He explained the incident through a written report.'

(v) Associated Lemma:

chengshu 陳述 'state', *shushuo* 述說 'recount', *shuo* 說 'say', *jianghua* 講話 'speak', *biaoda* 表達 'express', *jieshi* 解釋 'explain', *jianyi* 建議 'suggest', *chengren* 承認 'admit', *chengnuo* 承諾 'promise', *toulu* 透露 'reveal', *baoyuan* 抱怨 'complain', *pinglun* 評論 'comment', *baogao* 報告 'report', etc.

Although all the verbs above share the statement frame with a set of core frame elements, they encode very different eventive information. One fundamental question in our research is: how do these seemingly related verbs differ in semantic details? More specifically, what are the semantic distinctions in their lexical meanings and what are the syntax-to-semantics correlates lexicalized in each verb?

3.2 Database and methodology

The analysis of the present study is mainly based on the corpus data from Academia Sinica Balanced Corpus of Modern Mandarin Chinese (the Sinica Corpus), which contains a total of five million words. The distributional tendencies rather than grammaticality were taken as the important evidence for linguistic analysis. Within each class, high-frequency verbs were chosen as the representative lemmas to start with. The corpus data were used primarily for examining the basic syntactic patterns as well as collocational tendencies. Whenever needed, two popular daily-updated databases 'Google' and 'Yahoo' were used to verify collocational observations.

3.3 Grammatical and collocational distinctions

As introduced previously, verb meanings are anchored in semantic frames. Verbs sharing the same frame will display similar grammatical patterns characteristic of the shared frame. By investigating the grammatical behavior of a number of high-frequency statement verbs in the corpus, we found a list of criteria that are significant in distinguishing the subframes of statement. These criteria involve the distribution of key Frame elements as well as collocational variations: (1) distribution of Message; (2) distribution of Topic; (3) distribution of Speaker; (4) distribution of Addressee; (5) overt marking of Medium/Means; (6) the semantic range of Message; (7) aspectual collocation. Based on the above seven criteria and the frequency of nominalization of the verb, verbs of statement can be further categorized into several basic frames:

3.3.1 Distribution of message

The statement verbs may vary in their expressions of the key element, Message. The high-frequency verbs *shuo/jiang* 說/講 'say' and the public-declaring verbs 宣稱/宣告 'declare, claim' may encode the message as a direct or indirect quotation, but not in a nominal form, while verbs such as *biaoshi/biaoda* 表示/表達 'express' may allow a nominalized message, or a Message_description (Baker, Fillmore & Lowe 1998):

(8)　a.　Message as a direct quotation:
　　　　周先生毫不遲疑地 **說/宣稱/表示**：「我很感謝。」
　　　　'Mr. Chou said/claimed/expressed unwaveringly: "I am thankful."'
　　b.　Clausal, non-quoted Message:
　　　　周先生毫不遲疑地 **說/宣稱/表示** 他很感謝。
　　　　'Mr. Chou said unwaveringly said/claimed/expressed (that) he's thankful.'
　　c.　Message-description in nominalization:
　　　　周先生毫不遲疑地 **表示/*說/*宣稱** 他的感謝。
　　　　'Mr. Chou unwaveringly expressed/*said his thankfulness. '

However, the 'express' and 'declare' verbs are rather restricted in the semantic range of their message, as they prefer a declarative or assertive, but not an interrogative proposition:

(9)　周先生 **說/?宣稱/?表示:**「明天會下雨嗎？」

Other verbs with a more specified purpose, such as 解釋/說明/建議, may also

take a clausal message as in (10a). What is of interest is the fact that the verb *shuo* 說 may be added after the verb as a Message-marker, introducing a quoted or non-quoted message, as in (10b-c):

(10) a. 警方**解釋**並沒有這樣一條規定。
          'The police explained that there isn't such a rule.'
      b. 莊子向他**解釋**說：「你的問題太籠統了！」
          'Zhuangzi explained to him, saying: "Your question is too vague!"'
      c. 她又**解釋**說，跟往昔不同。
          'She then explained saying (that) it's different from before.'

As Message plays an important role in the statement frame, the way Message is coded serves as a key factor in distinguishing the basic frames: The *Say*-group and *Declare*-group verbs profile a propositional message, but differ in propositional assertion; the *Express*-group and *Declare*-type verbs prefer a declarative message, but differ in taking a nominalized message. The commonly-used, mono-syllabic verb *shuo* 說, with its Message-profiling function, may be added freely to almost all the statement verbs, introducing a propositional Message.

In contrast to an overtly coded Message, a unique group of V-O compounds in the statement frame, such as *jianghua* 講話 'speak', *xunhua* 訓話 'admonition', or even *jiangdao* 講道 'preach', tend not to take a grammatically marked Message. Morpho-logically, these lemmas consist of a mono-syllabic statement verb (e.g., *jiang* 講 'speak') followed by a pro-form for message *-hua* 話 or *-dao* 道 'words'. The morphological make-up entails that Message is lexically incorporated and backgrounded. The generic form for message *-hua* 話/*-dao* 道 functions like a pronoun, referring to something said without detailing the content of the message:

(11) 講演的人，開口**講話**，眼睛要看著聽眾。
      'A speaker, when opening the mouth to speak, should look at the audience with the eyes.'

The absence of an overt Message in the above example suggests that verbs with a lexically incorporated message (the Speaking event) focus more on the engagement of articulation rather than on the content of the message being delivered.

3.3.2 Distribution of topic

As a potentially relevant role, the Topic of a statement event can always be added

Mei-chun Liu and Ting-yi Chiang

as an adjunct, with an overt Topic-marker, *youguan* 有關, *guanyu* 關於, or *zhendui* 針對 'about, regarding', at the beginning of a sentence, as in (12a). However, some statement verbs take Topic as a more central role and allow it to be directly coded in the object position (12b), but some do not (12c):

(12) a. 關於教育問題，李教授一再 **表示/解釋/評論/建議** 政府要負責。
'<u>Regarding educational issues</u>, Prof. Lee repeatedly expressed/explained/ commented/suggested that the government should be responsible.'
   b. 李教授一再 **陳述/解釋/說明/評論** 教育問題。
'Prof. Lee repeatedly stated/explained/explicated/commented <u>educational issues</u>.'
   c. 李教授一再 ***說/*表示/*宣稱/*建議** 教育問題。
'Prof. Lee repeatedly *said/*expressed/*claimed <u>educational issues</u>.'

The Topic-predicating verbs in (12b) differ from the performative verbs in (12c) in that they do not perform an instantaneous speech act only, but they can also describe a durative speech process centered and extended on a Topic:

(13) 他對<u>教育問題</u> **陳述/解釋/說明/評論** 了三個小時。
'He made a three-hour statement/explanation/explication/comment on educational issues.'

In general, statement verbs that may profile a Topic are typically non-quotative, non-performative, non-instantaneous, and therefore extendable in time.

3.3.3 Distribution of speaker

As a fundamental element of the statement frame, the Speaker is typically encoded as the subject, an active initiator of the event. The Speaker may be human-animate as in (14a), or non-human or inanimate (14b), as a result of metaphorical or metonymic transfer from a locational entity to the people located there:

(14) a. 孔子/李教授/專家們 解釋中國文化的精義。
'Confucius /Prof. Lee/Experts explained the essence of Chinese culture.'
   b. 北京/聯合國/博物館 表示並未參與。
'Beijing/The UN/The Museum indicated that they are not involved.'

The occurrence of inanimate, organizational subjects has been discussed by Liu &

Chang (2005). They asserted that organizational or locational speakers tend to co-occur with verbs associated with a formal register and a non-personal or official topic/message.

In a statement event, the Speaker is normally present and identified as the source of information. However, with verbs of a marked manner/motivation of expressing information, such as *toulu* 透露 'reveal, divulge' or *yaochuan* 謠傳 'rumor', the source of information may be downplayed, either with an unspecified identity or totally omitted:

(15) a. 據一位金融界權威人士**透露**：今年發行上市股票的公司將增加十家。
'According to an authority in financing, the number of companies issuing stocks will increase by 10.'

b. **據透露/據謠傳**，警方已掌握重要線索。
'**As divulged/as rumored,** the police have found important clues.'

While predicating a propositional message, the expressing verbs with a marked manner allows a partial or complete omission of the identity of the Speaker, in a semi-fixed pattern, headed by the marker of information source ***ju*** 據 'as, according to'. What is worth noting in (15b) is that with the ***ju***-construction, the marked manner of letting out a message can replace or substitute for the source of information. It indicates a special case of conceptual transfer: Manner for Source; the marked manner of expressing information is viewed as the source of information.

3.3.4 Distribution of addressee

Another observation related to the key frame elements is that some statement verbs tend to prefer an overt realization of the Addressee. Although all the statement events might involve an Addressee by default, only the reporting verbs have a higher tendency (36% vs. less than 14%) in taking an Addressee, overtly marked by the directional marker *xiang* 向 'to, towards':

(16) 學生立刻向老師[Addressee]**報告**：「實驗結果出來了！」
'The student reported to the teacher immediately: "The experimental results came out!"'

(17) Frequency of Addressee:

V	說	說話	表示	說明	建議	透露	抱怨	承認	報告
+Addr	2.5% (5/200)	4.9% (14/284)	2.6% (13/494)	13.8% (38/282)	7.1% (11/156)	6.1% (12/197)	8.1% (15/185)	4.7% (14/296)	**35.5%** (39/110)

The high frequency of Addressee with reporting verbs is correlated and may be attributed to the lexically inferred status relation between the Speaker and the Addressee. Verbs like *baogao/bingbao* 報告/稟報 'report' entails an inferior-to-superior relation between the Speaker and the Address, whereby a lower-status reporter presents information to a higher-status Addressee. The reverse relation holds for verbs like *jiaodai/fenfu* 交代/吩咐 'leave words with, order' that encode a superior-to-inferior communication event whereby a higher-status speaker communicates with a lower-status Addressee:

(18)　縣太爺向身旁的衙役 交代 一番。
　　　'The County Governor left words with the subordinates besides him.'

Apparently, the social relation between Speaker and Address may also be lexicalized in the semantics of a subclass of reporting verbs.

3.3.5 Overt marking of Medium/Means

As shown in the conceptual schema (2) above, communication events may be performed via a given Medium or Means of communication. According to Liu & Wu (2003), it may be a language (Chinese vs. English) or a tool (phone vs. fax). In the statement frame, a particular means for conveying information was found: rhetoric devices. Over 16% of the events of *shuoming* 說明 'explicating' and *jieshi* 解釋 would take a rhetoric strategy, an analogy, or a comparison, as the means of expressing a message (*ju lizi* 舉例子 'with an example', *yong gushi* 用故事 'by a story', or *yong piyu* 用譬喻 'by an analogy'):

(19)　關於這一點，我們可以用一個譬喻來說明：譬如向朋友借書，絕不能突
　　　然說：「把書借給我。」
　　　'About this point, we can use an analogy to make it clear: for example, if we borrow a book from a friend, we cannot just say: "Lend me the book."'

The use of a rhetorical Means in the subgroup may correlate with the purpose of the explicating event: to clarify an issue by delivering a better-packaged, easier-to-understand Message.

3.3.6 Frequency of nominalization

In the Statement frame, some members may be used as a nominal, referring to the message-delivering activity (a durative process) or the 'packaged information' produced in the process. But quotative and instantaneous verbs, 說 'say' / 透露 'reveal' / 承認 'admit', cannot be used as a noun:

(20) 他們努力不懈，直到他人接受他們的**建議/解釋/*說/*透露/*承認**。
      'They worked hard until other people accepted their suggestion/explanation *saying/*revealing/*admitting.'

(21) Distribution of predicate vs. nominal uses

V\\Func.	說 200	說話 350	表示 600	說明 400	建議 251	透露 197	抱怨 279	承認 296	報告 239
Predicate	**100%** (200)	81.2% (284)	82.4% (494)	70.5% (282)	61.4% (154)	**100%** (197)	67.0% (187)	**100%** (296)	39.1% (110)
Nominalized	%	**18.9%** (66)	**13.6%** (106)	29.5% (118)	39.6% (118)	0%	**33%** (92)	0%	**60.9%** (129)

The nominal use correlates with the occurrence of the light verb, *zuo* 做/作 'do, make', which is often used to introduce a nominal activity as in (22). The table in (23) shows the frequency of collocating with *zuo* 做/作:

(22) 我們有足夠的時間來做 **說明/*說/*透露/*承認**。
      'We have enough time to make an explication/*saying/*revealing/ *admitting.'

(23) Collocation with the verb *zuo* 做/作 'do'

[+nom]	說	說話	表示	說明	建議	透露	抱怨	承認	報告
做/作	%	**4%** (1/25)	**6.7%** (7/104)	**17.8%** (19/107)	**1.1%** (1/95)	0%	**8.9%** (7/79)	0%	**12.9%** (16/124)

As suggested by Huang et al. (1995), the light verb, *zuo* 做/作, takes a deverbal noun referring to the process or the result of the process. It not only retains the "activity skeleton" but also entails the creation of a theme. The deverbal noun may refer to the durative process of making a statement or the delivered content from the statement (the message conveyed).

Moreover, all the deverbal nouns may be countable as they can be modified by a numeral-classifier or determiner, such as *yige* '一個', *yipian* '一篇', *yifen* '一份', *yixiang* '一項', or *yizhong* '一種', and quantifiers, such as *yixie* '一些', *xuduo* '許多' or *zhuduo* '諸多', as exemplified below:

(24) 部長表示已對石化業未來的發展做了 <u>一個/許多</u> **說明**。
'The Minister indicated that he already made <u>an account/many accounts</u> of future development of the oil-chemical industry.'

In contrast to *zuo* 做/作 'make, do', the light verb *jinxing* 進行 'proceed' may also take a deverbal noun. According to Liu & Chang (2005), *jinxing* 進行 'proceed' collocates with a procedural and durative process normally associated with a serious topic, a formal register and a conventionalized procedure:

(25) 主席在會場<u>進行</u> **說明/解釋/報告/*建議/*表示**。
'The Chair proceeded with explicating/explaining/reporting/*suggesting/ *expressing inside the conference room.'

Another frequently used verb to introduce a nominal form is *tichu* 提出 'present/ offer', which focuses more on the 'packaged information' than on the process of message-delivering or verbatim-quoting, as it cannot be used with the 'say', 'speak', or 'express' verbs:

(26) 我會就我所知的<u>提出</u> **解釋/報告/建議/*說/*說話/*表示**。
'I will offer an explanation/a report/a suggestion/*a saying/*a speaking/*an expression from what I know.'

The collocational tendency of *tichu* 提出 'present/offer' and *shuohua* 說話 'speak' may be attributed to the fact that *tichu* introduces the creation of an 'informational product' derived from, but independent of, the message-delivering process, which is an eventive feature characteristic of the explaining verbs. While other verbs do not produce an information package as they profile something different: the saying verbs focus on the verbatim of the message, the speaking verbs focus on the utterance process itself, and the expressing verbs focus more on the encoded meaning of the message.

3.3.7 Semantic variation of message

As mentioned before, Message is central to the statement frame and it may be realized as a clause or an NP, termed respectively as Message_Content and Message_description in Baker, Fillmore & Lowe (1998). While a message is required, it displays a wide range of semantic variation, since a verb 'subcategorizes' a message syntactically as well as semantically. For verbs of disclosing, secretive matters such as *mimi* 秘密 'secret', *zhenxiang* 真相 'truth', *neimu* 內幕 'inside story' are frequently collocates:

(27) a. 她向我**透露**了<u>她的秘密</u>：她做夢都想當一名解放軍戰士。
          'She revealed <u>her</u> <u>secret</u> to me: she dreamed of being a Red Army soldier.'
      b. 他決定向世人**揭露**<u>真相</u>。
          'He decided to disclose <u>the</u> <u>truth</u> to the world.'

The collocation certainly correlates with the secretive manner and marked purpose unique to the revealing event. A pragmatic presupposition is in play here: the Message is presumably something kept to the speaker and unknown to the public.

With verbs of complaining or judgment, the expressed Message tends to be a negative evaluation/judgment containing negative modifiers; e.g., *buzu* 不足 'insufficient, *chou* 醜 'ugly', or *shao* 少 'less', as illustrated below:

(28) a. 有的人**評論**<u>他的字說渾厚不足</u>。
          'Some commented on his calligraphy, saying it's not solid enough.'
      b. 她們**抱怨**<u>衣服永遠少一件</u>。
          'They complained that their clothes are always one-lacking.'

Positive messages are also found with positive evaluation verbs such as *chengzan* 稱讚 'compliment' or *zanmei* 讚美 'praise'. The judging verbs, including complaining and praising, should be viewed as a cross-frame group that inherits semantics from both the statement and judgment frames.

Thirdly, the propositional message with suggestion verbs are bound to be irrealis and future projecting, collocated frequently with irrealis modals such as *yao* 要 'want, will', *yinggai* 應該 'should', or *bixu* 必須 'must':

(29) a. 楊教授**建議**，身爲輔導工作者，<u>必須</u>確切了解次級文化。
          'Professor Yang suggested that, as a counselor, one must understand subcultures well.'
      b. 早晨，營地的一個同志**提醒<u>要</u>**帶乾糧。
          'In the morning, a comrade reminded that (we) should take dried food with us.'

Contrary to the irrealis modality, the admitting verbs cast an epistemic control over the message proposition, usually rendering a realis assertion to a presuppostional background:

(30) 他承認自己沒有盡到父親的責任。
      'We admitted that he did not fulfill the responsibility of a father.'

In sum, the semantic features of the Message are highly correlated with the semantic features of the verb.

3.3.8 Variation in aspectual collocation

Last but not least, statement verbs differ with respect to their collocation with aspectual modifications. The less-performative, non-punctual verbs (speaking, explaining, complaining, and reporting verbs) differ from the quotative, non-durative verbs (saying, revealing, suggesting, admitting verbs) in that the former group may collocate with a progressive marker, *zheng* 正 or *zhengzai* 正在, a process-boundary marker *wan* 完 or a time duration:

(31) a. 他 <u>正在</u> **解釋/說明/報告/?透露/?建議/?承認** 那件事。
'He is explaining/explicating/reporting/?revealing/?suggesting/?admitting that issue.'
b. 他 **解釋/說明/報告/?建議/?透露/?承認**了<u>三個鐘頭</u>。
'He explained/explicated/reported/?suggested/?revealed/?admitted for three hours.'

From aspectual observations, two kinds of event types can be distinguished for the statement frame: a relatively instantaneous act of uttering a message or a potentially durative process of packaging information. While verbs are lexically specified with aspectual properties as shown above, constructional variables may play a role in constraining the aspectual behavior: when a quotative message is profiled, the utterance can hardly be progressive or durative, even with a potentially durative verb:

(32) a. ?他 正在 解釋/說明/報告：「明天會下雨。」
'?He is explaining/explicating/reporting: "It will rain tomorrow."'
b. ?他 解釋/說明/報告了三個鐘頭：「明天會下雨。」
'?He has explained/explicated/reported for three hours: "It will rain tomorrow."'

As indicated in previous literature (cf. Dowty (1991) and Smith (1991)), lexical aspect interacts with clausal constraints in shaping the aspect of an utterance.

4. Frame-based classification of statement verbs

4.1 Basic frames of statement

Although sharing the same cognitive frame, statement verbs exhibit varied grammatical and collocational properties. The asymmetrical distribution patterns discussed in §3 clearly suggest that statement verbs can be further divided into subgroups. As verbs are anchored in semantic frames, a preliminary classification of the statement frame into subframes is provided below: eight basic frames are distinguished first, each with a definition, foregrounded or backgrounded Frame Elements, Basic Patterns, semantic attributes, and associated lemmas.

(33) Basic frames under the Statement Primary Frame:

a. **Basic frame 1 – *Say*** 說/宣稱
Def.: A message-profiling act whereby a Speaker utters a Message as a direct or indirect quote.
Foregrounded FEs: Speaker, Message
Basic Pattern: (i) Speaker[NP] < V < Message[CL]　他說明天會下雨
Semantic Attributes: [Event type: non-durative activity]
[Speaker: human or non-human]
[Message: propositional]
[Register: distinguished at the level of microframes]
[Purpose: to quote verbatim]
Lemmas: 說, 講, 道, 言, 宣稱, 堅稱, 聲稱

b. **Basic Frame 2 – *Speak*** 說話/講道
Def.: A Speaker is engaged in a message-delivering activity.
Foregrounded FE: Speaker
Backgrounded FE: Message
Basic Pattern: (ii) Speaker[NP] < V　他不停說話
Semantic Attributes: [Event type: process (potentially durative)]
[Register: formal vs. informal; distinguished at micro-frames]
[Speaker: human]
[Message: unspecified and incorporated]
[Purpose: marked vs. unmarked; distinguished at micro-frames]
Lemmas: 說話, 講話, 發言, 傳道, 講道, 訓話, 致詞

c. **Basic frame 3 –** *Express* 表示/透露

Def.: A speaker sends out a Message as a sign of expression to an Addressee.

Foregrounded FE: Speaker, Message_description, Message, Addressee

Basic Pattern:

  (i)   Speaker[NP] < V < Message[VP/CL]　他表示不想參加
  (ii) Speaker[NP] < V < Message_description[NP]　表示謝意/宣布獨立
  (iii) Speaker[NP] < {向/跟/對} Addressee[NP] < V#　他向媒體揭發弊案
  (iv) %Nominal use:[4] 做/作　< V[+nom]　作一個表達

Semantic Attributes: [Event type: activity (inherently non-durative]
                      [Speaker: human or organizational]
                      [Message: further specified at the micro-frame level]
                      [Register: distinguished at micro-frames]
                      [Manner: marked vs. unmarked]
                      [Purpose: to express an idea]

Lemma: 表示, 表達, 傳達, 說出, 說起, 提到, 提及, 轉述, 宣布, 宣揚, 宣告,主張, 透露, 揭露, 吐露, 揭發

d. **Basic frame 4 –** *Explain* 解釋/說明

Def.: A Speaker packages information in a Message to clarify an Issue or Topic.

Foregrounded FEs: Message, Topic, Issue, Means

Basic Pattern:

  (i)   Speaker[NP] < V < Message[VP/CL]　他說明情況很嚴重
  (ii) Speaker[NP] < V< Topic/Issue　他解釋他的決定/他如何作決定
  (iii) 以/舉/用　< Means[NP] < V#　他用數學解釋語言學
  (iv) 對/針對　< Topic[NP/CL] < V#　對語言學進
  (v) {提出/作/做/進行} < V[+nom]　提出一個解釋/進行解釋

Semantic Attributes: [Event type: procedural process with an incremental theme]
                      [Speaker: human or organizational]
                      [Message: packaged information]
                      [Topic: issue to be clarified]
                      [Medium: means of clarification]
                      [Register: non-casual]
                      [Purpose: to clarify an issue]

Lemma: 說明, 解釋, 陳述, 闡述, 陳明, 重申, 澄清

---

[4] **%** represents an optional pattern that may not occur with some class members.

e. **Basic frame 5 – *Suggest* 建議/承諾**
   Def.: A speaker offers an advice for an Actor to take an action in a Message
   　　　on a Topic. The Actor may or may not be the direct Addressee.
   Foregrounded FEs: Actor, Action, Topic
   Basic Pattern:
   (i)   Speaker[NP] < V < Message[VP/CL]　我建議這件事應先處理
   (ii) Speaker[NP] < V < Actor[NP] < Action[VP]　我建議他去看醫生
   (iii) 針對/對 < Topic < V#[5]　針對醫療問提建議政府大力改革
   (iv) {提出/作/做} < V[+nom]　針對醫療問題提出一個建議
   Semantic Attribute:　[Event type: non-durative, performative act]
   　　　　　　　　　　　[Speaker: human or organizational]
   　　　　　　　　　　　[Topic: issue for solution]
   　　　　　　　　　　　[Message: irrealis proposition]
   　　　　　　　　　　　[Actor: may or may not be Addressee]
   　　　　　　　　　　　[Register: formal]
   　　　　　　　　　　　[Purpose: to provide a solution or an advice]
   Lemma: 建議, 提議, 規勸, 警告, 提醒, 保證, 承諾

f. **Basic frame 6 – *Criticize* 評論/批評**
   Def.: A Speaker utters a Message as a judgment on a Topic or towards a
   　　　human Target.
   Foregrounded FEs: Topic, Target
   Basic Pattern:
   (i)   Speaker[NP] < V < Message[CL]　他批評政府不該出兵
   (ii) Speaker[NP] < V < Topic[NP]　他評論時事
   (iii) Speaker[NP] < V < Target[human NP]　他批評小王
   (iv) {提出/作/做/進行} < V [+nom]　作一個評論/進行評論
   Semantic Attribute:　[Event type: durative process]
   　　　　　　　　　　　[Speaker: human or organizational]
   　　　　　　　　　　　[Topic: inanimate]
   　　　　　　　　　　　[Target: human]
   　　　　　　　　　　　[Register: unspecified]
   　　　　　　　　　　　[Purpose: to judge]
   Lemma: 評論, 批評, 批判, 針貶, 埋怨, 抱怨, 稱讚, 讚美

---

[5] # represents a possible extension of accepted patterns.

g. **Basic frame 7 –** *Admit* 承認/保證

Def.: A Speaker utters a Message to verify a presumably known information on a Topic or pre-existing Agenda.

FEs: Speaker, Message, Agenda, Topic

Basic Pattern:

  (i) Speaker[NP] < V < Message[VP/CL]　她承認孩子不是她生的

  (ii) Speaker[NP] < V < Agenda[NP]　他承認警方的指控

  (iii) {關於/有關}Topic[NP]>V#　關於孩子的事，他承認了

Semantic Attribute: [Event type: punctual act]

　　　　　　　　　　[Speaker: human or organizational]

　　　　　　　　　　[Message: typically realis modality]

　　　　　　　　　　[Topic: issue for verification]

　　　　　　　　　　[Purpose: to verify a presupposition]

Lemma: 坦承, 承認, 否認, 保證, 證實, 證明

h. **Basic frame 8 –** *Report* 報告/交代

Def.: A Speaker gives an account of a Topic by delivering a Message to an Addressee.

Foregrounded FEs: Speaker, Topic, Message, Addressee

Basic Pattern:

  (i) Speaker[NP] < V < Message[VP/CL]　氣象報告將有颱風

  (ii) (由) Speaker[NP] < V < Topic[NP]　(由)他報告研究成果

  (iii) 對/針對 ＜ Topic[NP/CL] < V#　針對研究成果作報告

  (iv) {向/爲/對}Addressee[NP] < V#　向委員報告調查結果

  (v) {提出/作/做/進行} ＜ V[+nom]　向委員作一個報告

Semantic Attributes:

  [Event type: procedural process with an incremental theme]

  [Register: formal]

  [Speaker vs. Addressee: a status relation between the two]

  [Purpose: to give an account]

Lemma: 報告, 報導, 簡報, 回報, 稟報, 稟告, 交代, 吩咐

The eight basic frames are distinguished according to the unique set of FEs foregrounded or backgrounded in each frame and the patterns of coding these elements. To go one step further, semantically more fine-grained subclasses can still be found if distinctions are made as to detailed semantic features pertaining to a specific role. Although a full-ranged classification of smaller subclasses awaits another study, an illustration of a micro-frame is given below.

4.2 Examples of micro-frame

The *Express* Basic Frame highlights a message-sending process whereby a Speaker conveys a Message as a sign of expressing something to an Addressee. The associated verbs share the basic schema of information-sending, but may differ in their specifications of the semantic details of the participating elements. A closer examination of the manner and purpose of expressing information will indicate a further classification: three microframes can be differentiated: (1) the '*reveal*' micro-frame with verbs *toulu/tulu/jielu* 透露/吐露/揭露 profiles a secretive manner in disclosing a previously kept message; (2) the 'declare' microframe with *xuanbu/xuanyang/xuangao* 宣布/宣揚/宣告 describes a message-announcing event in a publicly open manner; (3) the 'mention' microframe with *tidao/shuoqi* 提到/說起 profiles a causal manner with an unmarked purpose. The three micro-frames can then serve as the semantic background for analyzing different sets of near-synonyms.

5. Conclusion

To sum up, the *basic frames* correspond to the basic-level categories (Brown 1958, Berlin & Paul 1969, Rosch 1973, 1975a, 1975b) in the Statement frame, pertaining to cognitively salient event types that are lexicalized in the Mandarin lexicon. With the goal of searching for the most effective and comprehensive way of exploring Mandarin verbal semantics, this study adopts a frame-based approach within the theoretical framework of Frame Semantics and worked out a multi-layered hierarchical taxonomy to distinguish the varied scopes of frames for the classification of verbs. A large-scale, in-depth investigation of Mandarin statement verbs was presented in the study as an illustration to the infrastructure for constructing a comprehensive lexical database, the Mandarin VerbNet. It proposes that in the frame-based hierarchy, verbs can be linked to a well-structured semantic network anchored in a well-defined ontology of frames. This work is essential to the distinction and representation of the semantic granularity of the verbal lexicon in Mandarin. Ultimately, this model is designed to capture the fine-grained syntax-to-semantics correlations, referred to by Pustejovsky (1995) as 'the semantic well-formedness', as well as to explore Mandarin verbal semantics in a cross-linguistic perspective.

References

Alonge, Antonietta, Nicoletta Calzolari, Piek Vossen, Laura Bloksma, Irene Castellon, Maria Antonia Marti, and Wim Peters. 1998. The linguistic design of the EuroWordNet database. *Computers and the Humanities* 32.2-3:91-115.

Baker, Collin F., Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. *Proceedings of 36th Annual Meeting of the Association for Computational Linguistics and the 17th International Conference on Computational Linguistics (COLING-ACL '98)*, 86-90. Montreal: Université de Montréal.

Berlin, Brent, and Kay Paul. 1969. *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.

Brown, Roger. 1958. How shall a thing be called? *Psychological Review* 65.1:14-21.

Chang, Li-li, Keh-Jiann Chen, and Chu-Ren Huang. 2000a. A lexical-semantic analysis of mandarin Chinese verbs: representation and methodology. *Computational Linguistics and Chinese Language Processing* 5.1:1-18. (In Chinese)

Chang, Li-li, Keh-Jiann Chen, and Chu-Ren Huang. 2000b. Alternation across semantic fields: a study on mandarin verbs of emotion. *Computational Linguistics and Chinese Language Processing* 5.1:61-80.

Chang, Miao-Hsia. 1998. The discourse functions of Taiwanese *kong* in relation to its grammaticalization. *Selected Papers from the Second International Symposium on Languages in Taiwan*, ed. by Shuanfan Huang, 111-127. Taipei: Crane.

Cheng, Robert L. 1989. *Guoyu Changyong Xuci Ji Qi Taiyu Duiyingci Shili [Mandarin Function Words and Their Taiwanese Equivalents]*. Taipei: Crane.

Chief, Lian-Chen, Chu-Ren Huang, Keh-Jiann Chen, Mei-Chih Tsai, and Li-li Chang. 2000. What can near synonyms tell us? *Computational Linguistics and Chinese Language Processing* 5.1:47-60.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67.3: 547-619.

Fillmore, Charles J., and Beryl T. S. Atkins. 1992. Toward a frame-based lexicon: the semantics of RISK and its neighbors. *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*, ed. by Adrienne Lehrer & Eva Feder Kittay, 75-102. Hillsdale: Lawrence Erlbaum Associates.

Givón, Talmy. 1984. *Syntax: A Functional-typological Introduction*. Amsterdam & Philadelphia: John Benjamins.

Givón, Talmy. 1993. *English Grammar: A Function-based Introduction*. Amsterdam & Philadelphia: John Benjamins.

Huang, Chu-Ren, Meili Yeh, and Li-ping Chang. 1995. A corpus-based study of nominalization and verbal semantics: two light verbs in Mandarin Chinese.

Proceedings of the 6th North American Conference on Chinese Linguistics (*NACCL-6*), Vol. 1, 100-112. Los Angeles: GSIL.

Huang, Chu-Ren, Kathleen Ahrens, 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. *Computational Linguistics and Chinese Language Processing* 5.1:19-46.

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

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

Levin, Beth, and Malka Rappaport Hovav. 1996. Lexical semantics and syntactic structure. *The Handbook of Contemporary Semantic Theory*, ed. by Shalom Lappin, 487-507. Oxford & Cambridge: Blackwell.

Liu, Mei-chun. 1994. Discourse explanations for the choice of *jiu* and *cai* in Mandarin conversation. *Chinese Languages and Linguistics*, Vol. 2: *Historical Linguistics*, ed. by Paul Jen-kuei Li, Chu-Ren Huang & Chih-Chen Jane Tang, 671-709. Taipei: Institute of History and Philology, Academia Sinica.

Liu, Mei-chun. 1999. Lexical meaning and discourse patterning: the three Mandarin cases of 'build'. *Cognition and Function in Language*, ed. by Barbara A. Fox, Dan Jurafsky & Laura A. Michaelis, 181-199. Stanford: CSLI.

Liu, Mei-chun. 2002. *Mandarin Verbal Semantics: A Corpus-based Approach*. Taipei: Crane.

Liu, Mei-chun. 2003. From collocation to event information: the case of Mandarin verbs of discussion. *Language and Linguistics* 4.3:563-585.

Liu, Mei-chun, and Yi-ching Wu. 2003. Beyond frame semantics: insight from Mandarin verbs of communication. Paper presented at the 4th Chinese Lexical Semantics Workshop. Hong Kong: City University of Hong Kong.

Liu, Mei-chun, Ting-yi Chiang, and Chun Edison Chang. 2004. The construction of Mandarin VerbNet: a frame-based corpus approach. Paper presented at the Symposium of Current Trends and Development of Chinese Lexical Semantics. Beijing: Peking University.

Liu, Mei-chun, and Chun Edison Chang. 2005. From frame to subframe: collocational asymmetry in Mandarin verbs of conversation. *International Journal of Computational Linguistics and Chinese Language Processing* 10.4:431-443.

George A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine J. Miller. 1990. Introduction to WordNet: an on-line lexical database. *International Journal of Lexicography* 3.4:235-244.

Niles, Ian, and Adam Pease. 2001. Towards a standard upper ontology. *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems*

(*FOIS-2001*), ed. by Chris Welty & Barry Smith, 2-9. Ogunquit, Maine, USA.

Pease, Adam, Ian Niles, and John Li. 2002. The suggested upper merged ontology: a large ontology for the semantic web and its applications. *Ontologies and the Semantic Web: Papers from the AAAI Workshop*, ed. by Adam Pease, 7-10. Menlo Park: AAAI Press.

Pustejovsky, James. 1991. The syntax of event structure. *Cognition* 41.1-3:47-81.

Pustejovsky, James. 1995. The nature of lexical knowledge. *The Generative Lexicon*, 5-26. Cambridge: MIT Press.

Rosch, Eleanor. 1973. Natural categories. *Cognitive Psychology* 4.3:328-350.

Rosch, Eleanor. 1975a. Cognitive reference points. *Cognitive Psychology* 7.4:532-547.

Rosch, Eleanor. 1975b. Cognitive representation of semantic categories. *Journal of Experimental Psychology: General* 104.3:192-233.

Ruppenhofer, Josef, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, and Jan Scheffczyk. 2006. *FrameNet II: Extended Theory and Practice*. Available at http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper&Itemid=126.

Smith, Carlota S. 1991. *The Parameter of Aspect*. Dordrecht & Boston: Kluwer.

Sweeter, Eve E. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge & New York: Cambridge University Press.

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

董強，郝長伶，董振東. 2004.〈基於知網的中文信息結合抽取〉，http://www. keenage.com。

董振東，董強. 2001.〈知網：中文信息結構庫〉，http://www.keenage.com。

Website Resources

HowNet: http://www.keenage.com

Sinica Corpus (中研院平衡語料庫): http://www.sinica.edu.tw/SinicaCorpus/

Sinica Bow (中研院中英雙語知識本體詞網): http://bow.sinica.edu.tw/

FrameNet: http://framenet.icsi.berkeley.edu/

Google (online archive): http://www.google.com.tw/

Yahoo (online archive): http://tw.yahoo.com/

教育部重編國語辭典修訂版: http://140.111.1.22/mandr/clc/dict/dict/?open

Mei-chun Liu
Department of Foreign Languages and Literature &
Graduate Institute of Linguistics
National Chiao Tung University
1001, Ta Hsueh Road
Hsinchu 300, Taiwan
mliu@mail.nctu.edu.tw

Ting-yi Chiang
Department of Foreign Languages and Literature &
Graduate Institute of Linguistics
National Chiao Tung University
1001, Ta Hsueh Road
Hsinchu 300, Taiwan
tingyi.flg92g@nctu.edu.tw

中文動詞語意網的建構：由陳述動詞出發

劉美君　　江亭儀

國立交通大學

　　本文介紹「中文動詞語意網」的架構及原則。有別於目前國際上已研發出的數個資料庫僅列出詞彙語意卻無語法考量，「中文動詞語意網」提供的語意訊息以語法表現爲根據，並以語言學理論爲基礎，探討影響語法現象的詞彙語意屬性。本語意網著重語法與語意的系統性連結，以語法表現及共現模式爲語意分類之依據。本研究採用框架語意學理論 (Fillmore & Atkins 1992)，將動詞語意在語意框架的基礎上定義，建構一個以框架爲本的動詞詞彙庫。不同於英文框架網路 (FrameNet) 的架構，本詞網區分了語意範圍大小有所不同的四種相對框架，作爲語意範疇間階層結構的依歸。本研究改良自之前的研究 (Liu, Chiang & Chang 2004, Liu & Wu 2003, Liu 2002)，以中文陳述動詞爲例，介紹目前最新的分析架構。本研究最終目標爲建立一個有語言學理論基礎，操作上有方法、有效率，並以語法表現爲根據，訊息完整而深入的語意分析表徵系統。

關鍵詞：詞彙語意，中文動詞，語意網，框架語意學，陳述動詞，語法語意
　　　　互動