

行政院國家科學委員會專題研究計畫 期中進度報告

統整動詞語意：中文「動詞詞網」的強化(2/3)

計畫類別：整合型計畫

計畫編號：NSC92-2411-H-009-020-ME

執行期間：92年08月01日至93年07月31日

執行單位：國立交通大學外國語文學系

計畫主持人：劉美君

計畫參與人員：巫宜靜，張群，賴蔚鍾，江亭儀，黃琬婷

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 93 年 6 月 1 日

行政院國家科學委員會專題研究計畫期末成果報告
統整動詞語意 (II): 中文「動詞詞網」的強化

Integrating Verbal Semantics (II) : The Augmentation of Mandarin VerbNet

計畫編號：NSC 91-2411-H-009-012-ME

執行期限：93 年 8 月 1 日至 94 年 7 月 31 日

主持人：交大外文系 劉美君

計畫助理：巫宜靜，張群，賴蔚鍾，江亭儀，黃琬婷

Abstract

This project examines the collocational patterns of Mandarin verbs of conversation and proposes that a finer classification scheme than the flat structure of ‘frames’ is needed to capture the semantic granularity of verb types. The notion of a ‘subframe’ is introduced and utilized to explain the syntactic-semantic interdependencies among different groups of verbs in the conversation frame. The paper aims to provide detailed linguistic motivations for distinguishing subframes within a frame as a *semantic anchor* for further defining near-synonym sets.

Keywords: Verbs of communication, Frame, Subframe, Conduit Metaphor, Collocation

1. Introduction

As the importance of lexical semantic research grows with the need of representing human knowledge, various lexically-based information networks have been proposed, such as the comprehensive work of differentiating word senses and sense relations in WordNet (Miller et al. 1990), the ontological hierarchy in SUMO (Das et al 2002, Pease et al 2002, Niles and Pease

2003), and the more linguistically-motivated model of FrameNet (Baker, Fillmore and Cronin 2003). While all providing valuable information regarding certain aspects of word meaning, the first two are constructed in a more intuitive manner. FrameNet, on the other hand, is based on the theory of Frame Semantics (Fillmore and Atkins 1992) and attempts to define meaning within a set of shared knowledge or background information, that is, a *frame*. However, as pointed out in Liu and Wu (2003), if meaning is anchored in the notion of ‘frame,’ then we need independent motivations for postulating different ‘frames’. What seems to be lacking in the current framework is a cognitive linguistic explanation as to how the individual ‘frames’ are distinguished and interrelated? In other words, what are the ‘semantic relations’ among all the frames? To answer the question, Liu and Wu (2003) proposed an overarching conceptual schema which incorporates all the core frame elements (FEs) and accounts for the interrelationship among various frames in the communication domain. By providing a cognitive schema as a macro-structure, the distinction of frames is then well-motivated. However, there still

remains another issue at a micro-level, as indicated by Liu and Wu (2003):

Within each frame, a wide range of verbs are found and one would wonder how these verbs differ from each other. For example, English verbs *speaking*, *discuss*, *quarrel*, and *gossip*, are all found in the Conversation Frame, but obviously, these lemmas encode something different. What are the differences? There seem to be frame-internal features that also need to be characterized.

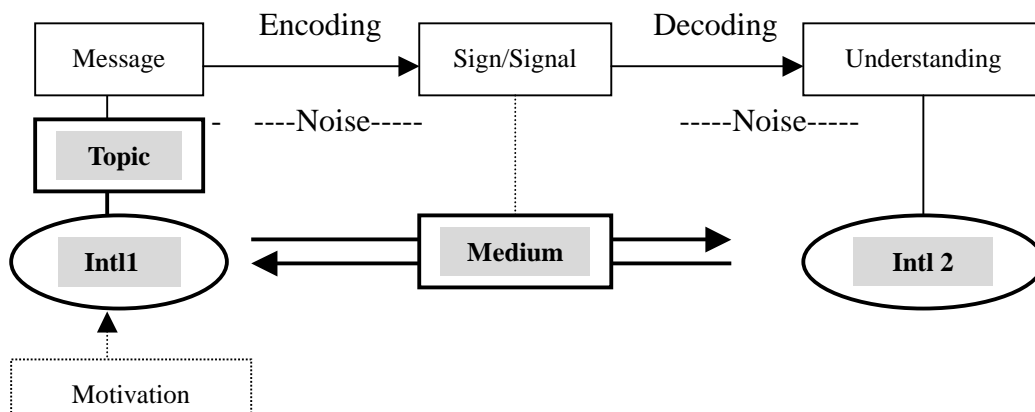
In this paper, we will show that within each frame, a more elaborated classification system is needed to account for a variety of verb behaviors. The notion of ‘subframe’ is introduced and utilized to capture the syntactic-semantic interdependencies observed in the corpus data¹.

2. Motivation for the Conversation Frame

Compared with the other communication frames, the Conversation

Frame is unique in that it profiles the property of *reciprocity* or two-way communication. Verbs in the Conversation frame encode reciprocal events where participants are involved as Interlocutors, such as *tan* 談 ‘talk’, *tanlun* 談論 ‘converse’, *taolun* 討論 ‘discuss’, *shangliang* 商量 ‘discuss’, *xietiao* 協調 ‘negotiate’, *xieyi* 協議 ‘negotiate’, *goutong* 溝通 ‘communicate’, *chaojia* 吵架 ‘quarrel’, *zhenglun* 爭論 ‘argue’, *xianliao* 閒聊 ‘chat’, and *liaotian* 聊天 ‘chat’, etc². These verbal events highlight part of the conceptual schema as discussed in detail in Liu and Wu (2003) and represented in (1) below. The core Frame Elements (with bold fonts and grayed areas) help define the frame as a bidirectional communicative activity conducted by both the Speaker and Addressee as Interlocutor 1 and Interlocutor 2 (or Interlocutors), via a certain Medium, on a given Topic.

(1) Conceptual schema for the Conversation frame:



The Conversation frame proves to be well-motivated in relation to other

communication frames, as most of its verbs share the same conceptual schema and realize similar constructions in coding the core frame elements. There is, however, a fundamental question to be answered, that is, within the Conversation frame, are there semantic subtypes that are also syntactically motivated?

3. Motivation for Distinguishing Subframes

As mentioned above, verbs of conversation involve a set of core Frame Elements: Interlocutor 1, Interlocutor 2 (or combined to Interlocutors), Topic and Medium. In most cases, the default Medium is ‘face-to-face’ when not overtly mentioned, as in the following sentence: 他們在 談/討論/溝通/吵/聊 人生的意義 ‘They are talking/discussing/exchanging views about/argue about/chat about the meaning of life.’ Intuitively, these different lemmas seem to encode differences in manner, formality or purpose, while sharing the same topic. But what are the grammatical correlates to the lexicalized meaning differences? When looking closely at their collocational patterns, we found that there are asymmetrical distributions in five respects: 1) V+V pattern: some may occur with a preceding verb such as *jinxing* 進行 ‘proceed’ or *dacheng* 達成 ‘achieve’; 2) V+NP pattern: the core element (3) Percentage with *Jingxing* (進行)

V1 \ V2	討論	溝通	Other Verbs (談/談論/吵架/爭論/聊天/閒聊)
進行	4% (3/83)	6% (25/419)	0%

‘Interlocutor2’ may sometimes be coded as the direct object; 3) Metonymic subject: the subject of the event may be inanimate entities taking the role of Interlocutors by the principle of metonym; 4) V+ Complement: some verbs take a postverbal complement or adverbial adjunct denoting effect evaluation, such as *chenggong* 成功 ‘succeed’ or *shibai* 失敗 ‘fail’; 5) In terms of distribution of grammatical functions, they show different frequencies of nominalization. Based on the five criteria, verbs of conversation can be further divided into 5 groups with corresponding sets of near-synonyms. We will address the syntactic-semantic interdependencies revealed by each pattern in the following sections.

3.1. V+V Pattern: with the preceding verbs *jinxing* (進行) or *dacheng* (達成)

The use of the light verb ‘進行’ entails a formal register and encodes a process or *atelic* event. It tends to occur with an activity verb compatible with the formal register and involving a durative process, as shown in (2):

- (2) a. 進行 討論/溝通/協調
 b. *進行 談/談論/吵架/爭論/聊天/閒聊

Below is the distributional tendency found in Sinica Corpus:

Another verb *dacheng* (達成) ‘achieve’ is also compatible with some conversation verbs, requiring a formal register but encoding a *telic* event. The verb *dacheng* is only found with the nominalized form of such verbs as 溝通/協調/協議³, i.e., activity verbs entailing a semantic endpoint with an incremental theme, as shown in (4):

- (4) a. 達成 溝通/協調/協議
 b. *達成 談/討論/吵架/爭論/聊天/閒聊

The co-occurrence with the preceding verbs *jinxing* (進行) or *dacheng* (達成) serves to distinguish the conversation events in terms of its pragmatic mode (formal vs. informal) and event types (telic vs. atelic).

3.2 V+NP pattern: Interlocutor2 as the Direct Object

Another pattern that sets the verbs apart

(6) Percentage with Interlocutor2 as DO

DO \ V	溝通/協調	Other Verbs (談/談論/吵架/爭論/聊天/閒聊/討論)
Interlocutor2 as DO	23%(28/123)	0%

3.3 Use of Inanimate Subject

Interlocutors in the conversation events are by default human participants. However, some verbs may take inanimate subjects (place or institute names) as Interlocutors via metonymic extensions from institute/building to human organization:

- (7) a. 台北和北京 談/討論/溝通/協議 了

regards the semantic role of the object NP. While most verbs can only take the Topic as the direct object, some verbs may encode Interlocutor2 as the direct object without adding the associative marker 和/與/跟, as shown in (5):

- (5) a. [執行秘書]_{Int1} 已協調 [相關單位]_{Int2}
 b. 由[他]_{Int1} 負責溝通 [校方]_{Int2}

This suggests that with the verbs 協調/溝通, the co-participant, Interlocutor2, may be viewed as the undergoer or the affected target of the event. Among the sentences of 協調/溝通 followed by an object NP, an average of 23 percent may take an Interlocutor 2 as the direct object in Sinica Corpus:

很久

- b. 台北和北京 *聊天/*閒聊 了很久

The application of metonymy tends to be associated with verbs that comply with the formal register requiring also a formal, non-personal topic (e.g., public affairs). Sinica corpus shows that verbs with marked manners tend not to be used with metonymic

subjects:

(8) Percentage with Inanimate Subject:

V \ Subj. type	談	討論	溝通	Other Verbs (吵架/爭論/聊天/閒聊)
Inanimate Subject	2% (4/191)	6% (5/83)	25% (21/85)	0%

3.4 Postverbal Complement with Effect Evaluation

Among the conversation verbs, only the ‘negotiate’ verbs (e.g., 協調, 溝通) may collocate with effect evaluation complement such as 成功 ‘successfully’ and 失敗

‘failingly’, as shown below with examples and percentage rate from the Sinica Corpus:

- (9) a. 國防部和兩廳院已初步協調[成功]
 b. 在協調[失敗]後, 水公司終於昨天宣布放棄其中一口

(10) Percentage of result evaluation complement:

Comp. \ V	協調	Other Verbs (談/談論/吵架/爭論/聊天/閒聊/討論)
成功/失敗	12.5% (2/16)	0%

The co-occurrence with effect evaluation complement indicates that the semantics of the ‘negotiate’ events encode some kind of an ‘effect’ or ‘result’ that is being sought by the negotiation process. This also implies that the two-way communication in the event of 協調/溝通 is a solution-seeking process which is semantically bounded and may be *evaluated* as to whether the solution or purpose has been achieved.

This pattern also correlates with the use of *dacheng* 達成 ‘achieve’, as mentioned above, in signaling an evaluation of the attainment of the desired result.

3.5 Overt Marking of Cause

As noted above, the frame elements associated with most verbs of conversation include Interlocutors, Topic and Means. However, the ‘quarrel’ verbs tend to take Cause as one important argument, as exemplified below:

- (11) 他們 為[錢] 吵架.

The overt marking of Cause with ‘因為/為’ (6%) seems to indicate that the ‘quarrel’ events, being a highly *marked* way of communication, requires an explanation for its happening. The overt Cause in the above example also fulfills the role of Topic, since what’s being argued about has to do with the

cause of the argument. We choose to use Topic-Cause to represent the unique element in the ‘quarrel’ events.

3.6 Absence of Topic

Another observation related to the marking of frame elements is that with verbs of chatting, there is seldom the occurrence of Topic. Topic can be found frequently in the event of talking, discussing, negotiating, and quarreling (Topic-Cause), but tends to

be absent in the event of chatting. In Sinica Corpus, Topic is simply not found with the verb 聊天.

3.7 Frequency of Nominalization

Some groups of verbs tend to be nominalized more frequently than the others. Comparing the high-frequency verbs and their distributions over grammatical functions, we see clear skewing in nominal uses:

(12) Distribution of Predicate vs. Nominal Uses

Func. \ V	談	討論	溝通	吵架	聊天
Predicate	97% (680/701)	52% (83/161)	55% (415/1013)	76% (123/162)	94% (134/142)
Nominalized	3% (21/701)	48% (78/161)	45% (598/1013)	24% (39/162)	6% (8/142)

Nominalization serves to change verbs to event nominals that may be referred to as a quantifiable entity. Nominalization is also highly correlated with the formal register of written texts.

4. Subframes as an Anchor for Near-synonyms

The asymmetrical distributions of the conversation verbs over different collocational associations clearly suggest that verbs can be further divided into subtypes. Although sharing the same conceptual frame, subclasses of verbs show distinct patterns of syntactic-semantic interdependencies that may serve as the basis to further define near-synonym sets. These subtypes may be viewed as anchored

in different *subframes*. And each subframe may display unique collocation patterns (syntactic properties) that manifest its unique semantic properties.

Among verbs of conversation, we found clear variations with regard to the coding of frame elements. The verbs 談/談論, as the frame-representative members, tend to take all the specified core Frame Elements, including Medium, such as ‘面對面/透過各種場合/在網路上談/談論(3%)’, while other verbs seldom co-occur with Medium. Although 談論 and 討論 are quite similar in argument taking behavior, 討論 may follow the light verb 進行 with a preposed Topic and often occurs with inanimate subject, indicating its formality.

協調/溝通 may also occur with 進行 and inanimate subject, but they differ from 討論 in taking Interlocutor2 as the Direct Object. The verbs 吵架/爭論 with a marked manner may use 因為/為 to specify a Topic-Cause. And verbs 聊天/閒聊 describe

the most casual and entertaining event, is inclined not to specify a Topic, while Topic is common with all the other verbs.

Below is a summary of the collocational patterns associated with the 5 subframes within the conversation frame:

(13) Collocational Patterns associated with the Conversation Subframes⁴

CP Subtypes	進行 +V	達成 +V	Intl2 as DO	Inanimate Subj.	CM 成功/失敗	Topic as Cause	Absence of Topic	[+Nom]
1. Converse 談/談論	No	No	No	Yes	No	No	No	Low
2. Discuss 討論	Yes	No	No	Yes	No	No	No	High
3. Negotiate 協調/溝通	Yes	Yes	Yes	Yes	Yes	No	No	High
4. Quarrel 吵架/爭論	No	No	No	Yes	No	Yes	No	Mid
5. Chat 聊天/閒聊	No	No	No	No	No	No	Yes	Low

Based on the distributional variations of collocation patterns, we can define the five subframes and group all the other conversation lemmas into the subframes:

(14) Conversation Subframes: Definition and Lemmas

a) Subframe_Converse

Interlocutors talk to each other on some Topics, sometimes via Medium, including means or language, to make their action of conversation more explicitly.

Lemmas: 談, 談論, 交談, 講, 談話, 講話, 說話, 會話, 交流, 會

談, 對話, 對談

b) Subframe_Discuss

Interlocutors exchange opinions on a Topic via evaluating pros and cons or seeking for a possible solution in a rather serious manner.

Lemmas: 討論, 商量, 商討, 商談, 會商

c) Subframe_Negotiate

Interlocutors confer and negotiate over a certain [Topic-Purpose] (Topic may overlap with Purpose) with each other so as to arrive at a settlement.

Lemmas: 協調, 協議, 溝通, 協商,

協議, 交涉, 接洽, 洽商

d) Subframe_Qarrel

Interlocutors dispute actively or exchange different opinions over a certain Topic-Cause in a heated manner. Usually, the topic tends to be coded as a cause of the proposition.

Lemmas: 吵架, 爭論, 爭吵, 爭執, 爭辯, 辯論

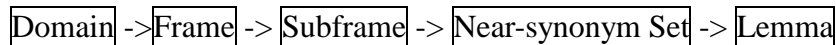
e) Subframe_Chat

Interlocutors are engaged in an informal verbal activity in a rather casual manner.

Lemmas: 閒聊, 聊天, 閒談, 閒聊, 談天

5. Concluding Remarks and Theoretical Implications

With the proposal of subframes within the theoretical construct of Frame Semantics, verb meanings may be defined with finer distinctions that are syntactically motivated. However, further fine-grained semantic distinctions are still needed to differentiate near-synonyms within each subframe, such as 交談 vs. 交流 or 討論 vs. 商量, etc. It is exactly at the subframe level that we may anchor all the near-synonym sets as closely related. In sum, to fully represent the meaning relations among verbs, we'd like to propose the following classificational scheme for representing verb meanings:



The five-layered structures allow verbs to be represented in a frame-based semantic

hierarchy with detailed lexical information to further disambiguate near-synonyms.

Endnotes

¹ A preliminary model of the Mandarin VerbNet (<http://140.114.75.18/verbnet/webform1.aspx/>) has been constructed by researchers from National Chiao Tung University and National Tsing Hua University with the support of a NSC grant.

² The lemmas discussed in this paper are high-frequent words of the conversation frame used mainly in Taiwan. Besides the words mentioned, there are some other lemmas, as an anonymous reviewer mentions, such as *tunghua* 通話

‘interconnect’, *shiangtan* 詳談 ‘descant’, *shenliau* 神聊 ‘tittle-tattle’, *sz yu* 私語, etc. These words can be included as members of one the subframes discussed in the paper.

³ In Sinica Corpus, only ‘達成協議’ is found, but in the Kimo website, we found examples: 1) 具備這三項條件就可以達成溝通; 2) 兩邊的陣營終於達成協調.

⁴ The abbreviate symbols here represent CP as collocational pattern, DO as direct object, CM as complement, [+Nom] as nominalization, respectively.

References

- Baker, Collin F., Charles J. Fillmore and Beau Cronin .2003. The Structure of the Framenet Database. International Journal of Lexicography 16(3).281-296.
- Das, Subrata, Kurt Shuster, and Curt Wu. 2002. Ontologies for Agent-Based Information Retrieval and Sequence Mining. Proceedings of the Workshop on Ontologies in Agent Systems (OAS02), held at the 1st International Joint Conference on Autonomous Agents and Multi-Agent Systems, Italy, July 15-19.
- 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.
- Kennedy, C. and Levin, B. 2002. "Telicity Corresponds to Degree of Change" Handout to Speech at Georgetown University.
- Liu, Mei-Chun and Yiching Wu. 2003. Beyond Frame Semantics: Insight from Mandarin Verbs of Communication. Paper presented at the 4th Chinese Lexical Semantics Workshop. City University of Hong Kong, Hong Kong. June22-July 11. (<http://icl.cityu.edu.hk/conference/4CLSW/BIG5/home.htm>)
- Miller, A., R. Beckwith, C. Fellbaum, D. Gross, K.J. Miller. 1990. Introduction to WordNet: An on-line Lexical Database. International Journal of Lexicography. 3.235-244.
- Niles, Ian and Adam Pease. Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. Proceedings of the 2003 International Conference on Information and Knowledge Engineering (IKE'03), Las Vegas, Nevada, June 23-26, 2003.
- Pease, A., Niles, I., and Li, J. 2002. The Suggested Upper Merged Ontology: A Large Ontology for the Semantic Web and its Applications. Working Notes of the AAIL-2002 Workshop on Ontologies and the Semantic Web, Edmonton, Canada, July 28-August 1, 2002.

Website Addresses:

FrameNet:

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

HowNet (知網):

http://www.keenage.com/html/c_index.html/

Sinica Corpus (平衡語料庫):

<http://www.sinica.edu.tw/SinicaCorpus/>

SUMO Ontology:

<http://ontology.teknowledge.com/>

WordNet:

<http://www.cogsci.princeton.edu/~wn/>

The Academia Sinica Bilingual Ontology

WordNet (BOW): <http://bow.sinica.edu.tw/>