

Semantic Annotation for Mandarin Verbal Lexicon: A frame-based constructional approach

Meichun Liu 劉美君 & Jui-ching Chang 張睿敬

Dept. of Linguistics and Translation

CityU of Hong Kong



Outline

1. Problems of existing works
2. A frame-based constructional approach
3. Conclusion

Major Issues

- ◇ Existing works in semantic annotation
 - ◇ Sinica Treebank (Chen *et al.*, 1999)
 - ◇ Chinese Proposition Bank (Xue and Palmer, 2005)
 - ◇ FrameNet (English) (Fillmore & Baker, 2002)

- ◇ What is the problem of these existing works?

Generic Roles Assignment

◇ Sinica Treebank: a pre-determined, generic tagset of semantic roles.

◇ Overgeneralization and low flexibility.

◇ **AGENT** of the action verb *hit*

1. The boy *hit* her hard. → human Agent



2. The wind *hit* her hard. → non-human physical Agent (?)

3. His words *hit* her hard. → non-human, non-physical Agent (??)

Verb by Verb Role Assignment

- ◇ Chinese Propbank: a predicate-specific frameset
 - ◇ Solves the problem of overgeneralization
- ◇ E.g. Frameset of 搬 (caused to move)
- ◇ Problem:
 - ◇ Simple listing of verbs
 - ◇ Lacks a systematic analysis

A Frame-based Approach

- ◆ FrameNet:

- ◆ Verbs that are similar in meaning are put together

- ◆ E.g. Cause motion frame

- ◆ Problem:

- ◆ Only a list of semantic roles.

- ◆ Ignore syntactic properties.

Syntactic Behaviors Matter

- ◇ Distinctions of frames can be made with clearly defined collo-construction.
- ◇ 搬 vs. 放:

	Agentive BA	Locative Inversion
搬 (<i>move</i>)	我把書搬到桌上	*桌上搬著一本書
放 (<i>put</i>)	我把書放在桌上	桌上放著一本書

- ◇ Locative Inversion is an identifying construction of PL verbs (Liu & Chang 2015).

Collocations Matter

- ◇ Compare the PPs with 搬 vs. 放
(corpus: Chinese GigaWord)

	+在	+到
搬	2	3203
放	24384	2704

- ◇ Caused Motion (搬, 移): typically collocate with 到
 - ◇ Caused to MOVE
- ◇ Placement (放, 擺): typically collocate with 在
 - ◇ Caused to BE-AT

A Frame-Based Constructional Approach

- ◇ A frame is defined with
 - ◇ The collo-constructions
 - ◇ A set of frame-specific elements

- ◇ Frame Semantics (Fillmore 1982) + Constructional Grammar (Goldberg 1995)

Cause-motion Frame

◇ Elements: Mover, Mv.Figure, End.Ground

Collocations: Path (e.g. 到/進)

◇ Lemmas: 搬 (*move*), 移 (*move*), 推 (*push*), etc.

◇ Basic constructional patterns

i. **Mover-prominent transitive BA**

[我們_{Mover}] 把 [兩張桌子_{Mv.Figure}] 搬 [到_{Path}] [教室裡_{End.Ground}]

‘We moved two tables to the classroom.’

ii. **Figure-prominent Inchoative**

[桌子_{Mv.Figure}] 搬 [進_{Path}] [教室裡_{End.Ground}] 了

‘The tables (were) moved to the classroom.’

Placement Frame

◆ Elements: Placer, Pl.Figure, Loc.Ground

Collocation: Locative (在)

◆ Lemmas: 放 (*put*), 擺 (*set*), 掛 (*hang*), 裝 (*load*), etc.

Placement Frame

◇ Basic constructional patterns:

i. **Placer-prominent transitive BA**

[布希 _{Placer}] 把 [紙條 _{Pl.Figure}] 放 [在 _{Locative}] [口袋裡 _{Loc.Ground}]

‘Bush put the note in the pocket.’

ii. **Figure-prominent Inchoative**

[炸彈 _{Pl.Figure}] 放 [在 _{Locative}] [紙盒裡 _{Loc.Ground}] (了)

‘lit. The bomb (was) put in the box.’

iii. **Ground-prominent Locative Inversion**

[房間裡 _{Loc.Ground}] 放著 [一張桌子 _{Pl.Figure}]

‘lit. In the room (was) put a table.’

Sense Disambiguation of 放

◇ 放: Placement vs. Releasing (cf. Liu and Chang 2015)

1. 她把貓放出來 → Releasing

She let the cat out

2. 房間裡放了一隻貓 → Placement

In the room was put a cat

◇ Distinctions of senses can be made with clearly defined collo-construction.

How Verbs are Annotated

Verb	Frame	Annotation
搬, 移	Cause-Motion	[他們 _{Mover}] 把 [書 _{Mv.Figure}] 搬 [到 _{Paht}] [二樓 _{End.Ground}]
放	Placement	[布希 _{Placer}] 把 [紙條 _{Pl.Figure}] 放 [在 _{Locative}] [口袋裡 _{Loc.Ground}]
	Releasing	[他 _{Releaser}] 把 [鳥 _{Rl.Figure}] 從 [籠子裡 _{Source}] 放 [出來 _{Path out}]

Conclusion

- ◇ A frame-based constructional approach
- ◇ Frames are defined by elements, and also identified with syntactic collo-constructions.
 - ◇ Class-specific: unique behaviors of the frames of CM verbs and PL verbs
 - ◇ Verb-specific: analysis of the polysemy of 放
- ◇ Semantic annotation based on frame-specific elements.
- ◇ Information of verbs are clearly provided.

References

- ◆ Chen K.J., et al. 1999. The CKIP Chinese Treebank: Guidelines for Annotation. ATALA Workshop IV Treebanks, Paris. pp 85-96.
- ◆ Fillmore, C. J., & Baker, F. C. 2000. FrameNet: Frame semantics meets the corpus. In Proceedings of the 74th Annual Meeting of the Linguistic Society of America.
- ◆ Goldberg, A. E. 1995. Constructions: A construction grammar approach to argument structure. Chicago: University of Chicago Press.
- ◆ Liu, M., and Chang, J.C. 2015. Redefining Locative Inversion in Mandarin: A Lexical-constructional Approach. in Proceeding of The 27th North American Conference on Chinese Linguistics (NACCL-27), vol 2. 439- 461. H. Tao, Ed. UCLA.
- ◆ Xue, N., and Palmer, M. 2005. Automatic Semantic Role Labeling for Chinese Verbs, in Proceedings of the 19th International Joint Conference on Artificial Intelligence. Edinburgh, Scotland.
- ◆ Sinica Treebank: <http://turing.iis.sinica.edu.tw/treesearch/>
- ◆ Chinese Proposition Bank: <http://www.cs.brandeis.edu/~clp/ctb/cpb/>
- ◆ FrameNet: <https://framenet.icsi.berkeley.edu/>
- ◆ Chinese GigaWord: <https://catalog ldc.upenn.edu/LDC2011T13>

Thank You