Automatic Classification of Verb-Direction Constructions in Mandarin Chinese

Tao Lin

B.A., Central China Normal University, 2011

B.S., Central China Normal University, 2011

M.A., University of Colorado Boulder, 2013

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Linguistics and Institute of Cognitive Science
2019

This thesis entitled:
Automatic Classification of Verb-Direction Constructions in Mandarin Chinese
written by Tao Lin
has been approved for the Department of Linguistics and Institute of Cognitive Science

____

Dr. Laura Michaelis

____

Dr. Martha Palmer

Date ____________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Lin, Tao. (Ph.D., Linguistics and Cognitive Science)

Automatic Classification of Verb-Direction Constructions in Mandarin Chinese

Thesis directed by Dr. Laura Michaelis and Dr. Martha Palmer

Processing Multi-Word Expressions (MWEs) presents a challenge for Natural Language Processing (NLP) systems (Sag et al. 2002). In Mandarin Chinese, there are different kinds of MWEs, such as compounding constructions and serial verb constructions containing multiple predicates (Chao 1968). In this project, I will focus on parsing the semantics of a family of constructions called Verb-Direction Constructions (VDCs) in Chinese. Similar to English Verb-Particle Constructions, VDCs include a verb of precondition followed by a directional verb (e.g., *na chu* (lit 'take exit') 'take out'). VDC functions include Self-Motion, Caused-Motion, Aspect, Discourse-Connective, and Evidential, among others (Liu et al. 1998).

Achieving native speakers' interpretation of a language in machine learning systems can support different applications. Inspired by the framework of Sign-Based Construction Grammar (Sag 2012, Michaelis 2009, 2013) as well as Conceptual Metaphor Theory (Lakoff & Johnson 1980, 1998), I conducted three classification tasks. In the first task, I designed a VDC taxonomy, which categorizes distinct functions of VDCs, such as event structures (both causative and non-causative) that involve the movement of an entity through space to a final location. Two versions of the taxonomy were developed and learned. The annotation guideline was mainly based on an analysis of Frame Semantics (Baker, Fillmore, & Lowe 1998) for different VDC events. In the second task, VDCs were annotated as metaphoric and literal expressions, and metaphor detection was performed. The third task makes preliminary steps aimed at detecting the coerced use of VDCs, in which VDCs alter the canonical argument structures of verbs (Goldberg 1995, 1999).

This research makes two primary contributions. First, it establishes linguistic analyses of the VDC properties in question, including taxonomies of event types, metaphorical mappings, and coercion, most of which directly support the VDC classification tasks. Second, based on the linguistically motivated categories, it develops an automated method for semantic classification of VDC constructions, surpassing the scope of classification resources previously devised within Chinese NLP (Xue et al. 2000; Xue & Palmer 2009; Huang et al. 2010; Lu & Wang 2017). The system developed potentially supports other NLP applications, such as machine translation, event detection, metaphor processing, and word sense disambiguation.

Acknowledgements

I want to first acknowledge my co-advisors for their support on my graduate career. Dr. Laura Michaelis has been advising, training, and encouraging me from my Master's degree to my Ph.D. Her knowledge and insights have led me through my academic adventure. Dr. Martha Palmer opened the world of computational semantics for me. I have learned immeasurably from her since my first class of Semantic Role Labeling.

Dr. Mans Hulden and Dr. Miika Silverberg helped me with my computational skills in a substantial way. Dr. Nianwen Bert Xue and Dr. Jim Martin also provided important input to this research.

I also want to thank Dr. Qingyang Yan, Dr. Bing Liu, and Dr. Yo Matsumoto, for their support and their opinions on the Chinese linguistics portions of this research. My annotators, Ling Liu and Dr. Yingjie Li, also provided valuable suggestions. I could not have finished this work without the support from family and friends, including my parents and Bob and Anita Hueftle, and my colleagues in the Linguistics Department, including Jesus Villalpando, Norielle Adricula, and Sarah Moeller. My final special thank you is given to my middle school Chinese teacher, Lan Zhang, for showing me the importance of the Rosenthal Effect.

The great Chinese philosopher Chuang Tzu once said, "Our lifetime passes like a speedy white horse, which runs past a crevice." It is surprising that my journey to the Ph.D. degree has already come to its conclusion. It is my hope that this work will be nourishing and enjoyable for those who work in and out of the field.

Table of Contents

Tables

Table

Figures

Figure

Chapter 1
Introduction

1.1 Overview

A major challenge in the design of Natural Language Processing (NLP) applications is to get the right interpretations of multi-word expressions (MWEs) (Sag et al. 2002), which are known to cause problems like overgeneralization and unpredictability (Baldwin & Kim 2010). When several words group together, their combinations may have different syntactic or semantic behaviors from the combinatory behaviors of the subparts (Sikos et al. 2008). For example, Verb-Particle Constructions (VPCs) like *drink up* do not express directed motion. Instead, the particle UP indicates completion. Another VPC, *egg on*, which means 'urge someone to do an action that might be a bad idea,' cannot be predicted from the meaning of *egg* or *on*. Similar to English and other languages, Mandarin Chinese is a language with MWEs like compounds and complex predicates (Chao 1968). The Chinese Verb-Direction Construction (VDC) is roughly comparable to the English VPC (Ju 1992). Formally, these constructions share the basic structure "verb (V1) + directional verb(s)[1] (DV/V2)." There are about 28 DVs, including simple DVs (e.g., *shang* 'ascend,' *jin* 'enter,' and *hui* 'return') and compounding DVs (e.g., *shanglai* 'ascend come' and *xiaqu* 'descend go'). Semantically, VDCs can express motion, result, and aspect. In addition, some combinations have been further grammaticalized as discourse and modality markers.

[1] I give arguments in favor of identifying these directional particles as DVs in the first part of my Appendix 1.

Given that VDCs express several dimensions of information about semantics and syntax, this section will first describe three theoretical issues affecting native speakers' interpretations of distinctions in VDCs. If an NLP system needs to achieve a human level in the understanding and use of language, it should be able to distinguish the categories of functions, metaphors, and novelty in VDCs as humans do. In the second part, these theoretical issues that pose challenges in the NLP applications related to these problems will be discussed, such as problems in machine translation and metaphor processing.

This project aims at developing a multi-purpose classification system that can address these problems in VDCs. This classification system mainly serves three main goals: classifying event types, detecting metaphors, and detecting novel uses. This project will help the machine better understand the interpretations of VDCs while supporting several applications. Section 1.1 focuses on the linguistic distinctions required for the machine to correctly interpret VDCs, and Section 1.2 focuses on VDC interpretation problems that arise with several computational applications.

1.2 Several Problems in Machine Interpretation of VDCs

The first challenge is to make the machine understand the multi-functionality of VDCs as native speakers do. The functions of VDCs vary from directed motion to aspect and modality, and each DV has its own distribution within these functions (Liu et al. 1998). The following sentences are examples of the various functions of VDCs:

1. deng shan dui pa shang le zui gao feng
   climb mountain team climb ascend ASP SUP high peak

"The mountaineering party climbed up the highest peak." (cf. Google Search)

2. ta ai shang ta le

   he love ascend she ASP

   "He fell in love with her." (cf. Google Search)

3. hua-ping kan shang qu hen guanghua

   follower-bottle look ascend go very smooth

   "It looks like the vase is smooth." (cf. CCL)

4. cong ɪntərnɛt shuo qi

   from Internet speak rise

   "Speaking of the Internet…" (an article title) (cf. Wang 2006)

In Sentence 1, the VDC *pa shang* means 'climb up/onto,' which encodes directed motion. In Sentence 2, when the same DV combines with the verb *ai* 'love,' the VDC means a state transition, from not-loving to loving. The VDC in Sentence 3 is *kan shangqu* (lit 'look ascend go') 'it looks like...' It expresses the speaker's evidentiary source for the assertion about the vase. In Sentence 4, the meaning of the DV *qi* 'rise' is inchoative. However, the pattern *cong…shuoqi* 'start speaking from…' appears frequently in the news and magazines as headlines. This pattern has the function of evoking a topical referent. By listing VDCs encoding directed motion, resultative, aspect, and discursive functions together with an idiomaticity scale and construction hierarchy, the framework of Construction Grammar (Fillmore 1988; Croft &

Cruse 2004; Goldberg 2004, 2006, etc.) offers an analytical approach to various combinations of "V1+DV." Based on this theoretical insight, the classification system should predict the event structures associated with various VDCs.

The second challenge is the appropriate interpretation by the system of figurative VDCs. There is a distinction between literal and non-literal uses in natural language (Sikos et al. 2008), including metaphors, similes, metonymies, idioms, and so on. Being cross-domain analogies (Lakoff & Johnson 1980 & 1998), metaphor is one of the most important realizations of figurative language. In the case of VDCs, the event they describe is often neither concrete nor literal:

5. zhong-guo renmin zhan qi lai le
middle-country people stand rise come ASP
"Chinese people have stood up!"

The example, a declaration from Chairman Mao Zedong in 1949, concerns the political independence of Chinese people. No action of literally raising one's body was involved. Native speakers can understand that this metaphor is using an ascent schema to describe increasing independence and influence, similar to *The Rise of X* in English (e.g., *The Rise of Theodore Roosevelt*, and *the rise of giant consumer startups*). The third task of the classification system aims at extracting metaphoric VDCs from literal ones.

Apart from multifunctionality and metaphoric usages in VDCs, humans can understand and create novel VDC combinations that are hard for the machine to learn. This means that the argument structure of a verb changes through the addition of arguments not ordinarily licensed by the verb (Goldberg 1995, 1999; Michaelis 2004, 2006), a process known as coercion (De Swart 1998). Here is an example of a coerced VDC:

6. baba yong bing-gun ba ta pian jin li-fa-guan

   dad use/with ice-stick BA he trick enter cut-hair-store

   "Dad tricked him into the barber's shop with an ice cream bar." (cf. Google Search)

As a transitive verb of manipulation, *pian* 'trick' does not have any locative arguments in its valence list. However, in a Caused-Motion VDC, the use of the DV *jin* 'enter' introduces a Goal argument *lifaguan* 'barber's shop.' The cheating action of the father caused the entering motion by the child. As a result, the VDC *pian* (…) *jin 'trick into'* has three arguments: Agent, Theme, and Goal. It is important to adopt a theoretical framework to explain how V1, V2, and the VDC contribute to argument variability. Based on such a framework, the VDC processing system will be able to detect tokens of novel VDCs.

VDCs can be analyzed based on categories event structures, metaphors, and coercion in the usage of native Chinese speakers. The goal of achieving human-like interpretation of VDCs should be considered in the building of NLP systems. Having discussed the interpretation problems involving VDC's, I will now survey how these problems are further related to some challenges in NLP applications, such as machine translation and metaphor processing.

1.3 Challenges in Processing VDCs in NLP Applications

The multifarious interpretations of VDCs not only create obstacles for general machine intelligence, but also pose challenges for a variety of NLP applications. The VDC classification system can offer solutions or outputs that can facilitate progress in these applications.

The first challenge is in applications like Human-Robot-Interaction (HRI) that involves Word Sense Disambiguation. Proposals and experiments have been made to create robotic agents which can understand humans' speech inputs (Fong et al. 2001; Skubic et al. 2004; Zhang et al. 2008). Transparency is an important design feature that can contribute to developing trust and effectiveness in HRI (Nikolaidis 2017), and which should allow a system to arrive at the appropriate interpretation of an ambiguous sentence. Imagine a case when a Chinese user gives a speech command to an automobile robot:

7. cong zhe-li zou xia qu

   from DEM-place walk descend go

   "Go down from here." Or "Continue walking from here."

There are two possible event interpretations of the VDC *zou xiaqu* lit 'walk descend go' because the compounding directional element *xiaqu*, can either mean 'continue' or 'descend and move away.' To increase communicative effectiveness, in, for example, the recognition of human commands, it is necessary for the machine to output the most probable guess and confirm it with the human, i.e., *ni yao wo cong zheli wang xia zou ma* 'Do you want me to move down from here?' In order to achieve such a goal, the function classifier in the VDC processing system needs to distinguish between subtypes of constructions, such as aspect vs. motion. In future work, a link can be built between the semantic representations (of constructions) and execution functions of the automobile robot (Lv et al. 1998; Skubic et al. 2004).

The second challenge is figurative language processing. Metaphor detection is important for information retrieval systems that target creative language (Shutova 2015). The metaphorical VDC in Example 6 is repeated below:

8. zhong-guo renmin zhan qi lai le
   middle-country people stand rise come ASP
   "Chinese people have stood up!"

Based on the common orientational metaphor POLITICAL CONTROL OR FORCE IS UP (Lakoff and Johnson 1980: 19), the natural language reasoning system uses source-target relationships to make metaphoric inferences by identifying source/target/metaphor-related words and answering metaphorical questions like "When/why did Chinese people kneel?" with the help of Synsets from Chinese WordNet (Huang et al. 2010). As for literal VDCs, such as the *boy stood up from his seat*, since it is a pure physical motion event, the semantic roles will be mapped in image or video captioning systems based on Semantic Role Labeling (SRL) (Gupta et al. 2015). Developing a binary classifier for literal and non-literal VDCs can establish the first step toward NLP systems that further process linguistic metaphors or concrete physical motion events. Imagine that during a soccer game the speaker made two comments using the same VDCs *ti jin* 'kick into':

9. A. faguo-dui ti jin jue-sai

France-team kick enter decide-game

"The French team kicked their way into the final."

B. wumudidi xiang qiu-men <u>ti</u> <u>jin</u> yi li jin-qiu

Umtiti toward ball-gate kick enter one CL enter-ball

"Umtiti kicked one ball into the goal."

In Sentence 9A, we understand that the game in which the French team was playing and the final are two different reference points in the time schedule of the World Cup. The metaphor underlying (9A) is STATES IS LOCATIONS (Lakoff 1993: 16), so this sentence cannot be fully mapped to physical elements in an image or video. In (9B), participant roles like Umtiti, the goal, the kicking action, and the ball are all concrete, thus having the potential to be recognized and mapped. Distinguishing literal and non-literal VDCs can distinguish abstract expressions from those with concrete and physical elements. While this dissertation will not address advanced metaphor processing or concrete language directly, it will provide a binary classifier of VDC metaphors as a means to establish metaphor-related processing tasks.

The third challenge is Chinese-English machine translation (MT). Many Chinese VDCs with novel verb and/or non-canonical argument structures have errors in their target English outputs on Google Translate[2]:

10. wo gui ye yao ba ta gui hui jia

I kneel still want BA she kneel return home

[2] https://translate.google.com/

"I would even get on my knees, if it would make her come home (lit. Even if I kneel down, I want to kneel her to return home.)" (cf. CCL)

[Google Translate] I have to pick him up.

11. gong-kuan-chi-he bu neng ba GDP chi shang qu

public-money-eat-drink NEG can BA GDP eat ascend go

"Eating and drinking on tax-payers' money cannot eat GDP to go up." (cf. CCL)

[Google Translate] Public money can not eat GDP

The Google output in Example 10 reflects the process of him being caused by the speaker to be back home, but the causative action *gui* 'kneel' is not translated. In Example 11, the Google output[3] does not include the direction UP, nor does the translation include the notion of caused-Motion. (10) and (11) involve new and rare usage of lexical items but can easily be understood and created by native speakers. Neural machine learning systems in MT, such as encoder-decoder models (Wu et al. 2016), are not good at learning low-frequency data with complicated argument structures. Current linguistically motivated NLP resources do not collect coercion well, such as Chinese PropBank and WordNet. This dissertation aims at providing a linguistic analysis and a detection system of coerced VDCs dealing with this challenge.

1.4 Organization

[3] I provided a collection of VDCs with their problematic Google MT translations in **Appendix 3**.

The goal of this project is to develop a VDC classification system that reveals the syntax and semantics of VDCs in a way that can contribute to problem solving by machines tasked with interpreting natural language in the kinds of efforts mentioned above. This process will entail a series of studies to handle the classification tasks. This dissertation is organized as follows: Chapter 2 gives a description on the basic settings of the classification system, including data collection, annotation, algorithms, and experimental design. Chapter 3 includes the classification task for categories of different VDC event types. Chapter 4 second describes VDC metaphor detection. Section 5 deals with a linguistic analysis of coerced VDCs. Chapter 3, 4, and 5 each include theoretical background, review of the current NLP resources and experiment results of classification tasks. Chapter 6 provides concluding remarks.

Chapter 2
A System Description of the Computational Models

2.1 Overview

Research Questions: How is the data collected and annotated in this project? What are the computational models used in the classification tasks?

In order to further investigate the problems mentioned in Section 1, I utilize a multi-purpose classification system to categorize VDCs. Motivated by research on the construction processing tasks in English, this section introduces the configurations of the classifiers. It has four components: data, annotation, algorithms, and experimental conditions. This section starts with a review on using supervised learning approaches for construction processing and highlights the importance of utilizing a similar approach to processing Chinese constructions like VDCs (Sec. 2.2). Then, system configurations in this research, such as data collection (Sec.2.3), annotation (Sec.2.4), and algorithms and experimental conditions (Sec. 2.5), are described.

2.2 Construction Processing in Supervised Learning

NLP uses three main types of machine learning approaches: unsupervised, supervised, and semi-supervised (Wikipedia). This project adopts a supervised learning approach. Classifying linguistic categories, like constructions and conceptual metaphors, mainly falls within supervised learning. This means that these NLP systems rely on a certain amount of human-annotated data for training, validation, and testing. During training and validation, the systems learn features from examples, and perhaps external resources, and make probabilistic guesses in the testing phase. The guesses are then compared with annotated labels or sequences which function as gold standards for calculation of performance scores.

While approaches like unsupervised or semi-supervised learning are good at handling domain-general tasks, supervised learning excels at capturing idiosyncratic concepts that require a lot of human input. Efforts that have applied supervised learning to the processing of constructions in English include Cook and Stevenson (2006) for VPCs, Chen et al. (2015) for Light Verb Constructions (LVCs), and Bonial et al. (2011) and Hwang (2014) for Resultative Constructions. The shared objective in these projects includes supervised detection of certain constructions and their subcategories. In Chinese NLP, although much work has been done to improve algorithms for MWE extraction (Piao et al. 2006; Wang & Liu 2011; Fu et al. 2012; Liang et al. 2017) and build knowledge bases for certain types of idioms like *chengyu*[4] (Wang & Yu 2010), little work has been done on processing specific constructions. Jiang et al. (2017) apply the PARSEME guidelines[5] to annotate Chinese LVCs. However, other types of Chinese MWEs, such as resultative constructions, including VDCs as a subclass, have not been discovered. Therefore, this project chooses a supervised learning approach to extract meaningful categories of VDCs.

2.3 Data Collection from Corpora

[4] *Chengyu 成语* 'fixed speech' is a type of idiom in Chinese, which usually has four-characters and different syntactic functions (Wang 1987).

[5] http://parsemefr.lif.univ-mrs.fr/parseme-st-guidelines/1.0/?page=040_Annotation_process_-_decision_tree

This section covers the data extraction process and specific considerations behind that process. The supervised learning system data is collected from a number of popular corpora: Chinese TreeBank (CTB) 9.0[6], the Center of Chinese Linguistics (CCL) corpus from Peking University (Zhan et al. 2003) and the Beijing Language and Culture University Modern Chinese Corpus (BCC) (Xun et al. 2016). An additional resource is *The Little Prince* (LP) (De Saint-Exupéry 2015), which has been annotated for Chinese Abstract Meaning Representation (CAMR) (Li et al. 2016) and Semantic Network for Adposition and Case Supersenses (SNACS) for Chinese (Zhu et al. 2018). The total genre coverage includes social, economic, political, and sports news, literature, academic texts, conversations, micro-blogs, recipes, and so on. The VDC data is collected semi-automatically. Since this search approach does not rely on specific syntactic structures, the returned instances would contain VDCs in any syntactic environment, varying from the simple patterns with implicit roles, like "V1+DV," to VDCs in more complicated patterns like embedded clauses and other types of SVCs. Two searching standards are used: the DVs and their Parts-of-Speech (POS). In the POS tagging, each of the 28 DVs is either labeled as "v" (i.e. verb) or "u" (i.e., auxiliary for aspectual DVs). To quickly tag the POS labels for annotated data, Tsinghua University Lexical Analyzer of Chinese (THULAC) (Li & Sun 2009) is used. For example, in order to search VDC tokens with the DV 上 *shang* 'ascend,' there are four regular expressions: "_v 上_v," "_v 上_u," "_a 上_v," "_a 上_v." These expressions mean that the V1 in a VDC can be a verb or predicative adjective (i.e., "a"), while V2 is *shang* labeled as another verb or auxiliary. Then, two rounds of manual selection are done to delete any irrelevant uses of DVs and to make sure that there is at most one VDC in each

[6]http://dla.library.upenn.edu/dla/olac/record.html?id=www_ldc_upenn_edu_LDC2016T13

instance. As a result, using the semi-automatic approach, I collected 15,852 tokens of VDCs from the four corpora.

	CTB 9.0	CCL	BCC	LP	Total
VDC	7,011	8,077	665	100	15,852

Table 1. Current Counts in VDC Tokens

CTB is widely used in different Chinese NLP tasks, and CCL is the major corpus for Chinese linguistic research. They provide most of the data for classification. The distribution of the DVs in these corpora is given in Table 2.

	DV	Literal Meaning	CTB9.0	CCL
1	guo	‘pass’	544	350
2	guolai	‘pass come’	239	350
3	guoqu	‘pass go’	128	350
4	kailai	‘move.away come’	20	350
5	kaiqu	‘move.away go’	0	348
6	kai	‘move.away’	220	32
7	shang	‘ascend’	407	400
8	shangqu	‘ascend go’	104	350
9	shanglai	‘ascend come’	35	340
10	xia	‘descend’	160	375
11	xiaqu	‘descend go’	199	350

12	xialai	‘descend come’	492	350
13	dao/dao...lai/dao...qu	‘reach/reach come/reach go’	369	332
14	jin	‘enter’	238	350
15	jinlai	‘enter come’	54	52
16	jinqu	‘enter go’	100	350
17	chu	‘exit’	468	400
18	chulai	‘exit come’	890	350
19	chuqu	‘exit go’	173	350
20	hui	‘return’	144	350
21	huilai	‘return come’	142	350
22	huiqu	‘return go’	83	295
23	qi	‘rise’	695	350
24	qilai	‘rise come’	508	350
25	mixed	NA	600	252
	Total	NA	7,011	8,076

Table 2. Distribution of DVs in CTB 9.0 and CCL

As shown in Table 2, the two corpora have different distribution patterns. All the CTB data was exhausted, so it shows imbalance: VDCs with the DV *chulai* ‘exit come’ have the most tokens while *kaiqu* ‘move.away go’ has zero. CCL is much bigger than CTB. Only part of the CCL data was collected. Most DVs have a similar amount of tokens (about 300-400) as an effort to achieve more data diversity. CCL data compensate for the lack of certain VDCs in CTB, e.g., the collection of *shanglai* ‘ascend come’ and *huiqu* ‘return go.’

2.4 Annotation Process

After the data is collected, the next step is to develop an annotation guideline, so that the annotators can learn to add conceptual labels to the tokens with regard to the category of interest. (The definitions of these categories will be discussed in the coming sections with specific tasks.) The development of the guideline is an iterative process encompassing drafting, expert review, data testing, and revising. After pre-annotating 500 instances, I drafted the guideline in Chinese, shown in Figure 1. Then the draft was reviewed by three Chinese linguists and tested on more instances. Revision in the next round served to reduce linguistic terminologies so that annotators could understand the guidance intuitively. For example, instead of defining Caused-Motion VDCs as "an Agent causes a Theme to move to a Goal along a Path," its definition was simplified to "a person or object causes another object to move from or to a place." Furthermore, morphosyntactic diagnostics were also added when necessary, e.g., Caused-Motion VDCs can be realized as BA and passive sentences. Additional resources, such as Chinese examples for VDC functions and metaphors, as well as lists of typical verbs of motion and causative verbs (Li 2008), were included in the guideline.

Self-motion VDCs: Class [1] and [2]
Class[1] 和 [2] 为自移型动趋式(self-motion)。这两类动趋式表示移动主体自行进行的移动，如：

> *这时他向我跑来。*Class[1]
> *登山队员爬上顶峰。*Class[1]
> *大伙儿站起来。*Class[1]
> *坏人追了上来。*Class[2]
> *你们都退下吧。*Class[2]

[1]和[2]的区别在于，趋向动词在[1]表示本义，但是[2]中的趋向词的方向发生了改变（通常是垂直方向变为水平方向）。

在句法上，[1]和[2]相似，都可以满足可能式(*V1 得/不 DP*)，一般都会出现或者补出表示地点的论元（如起点、终点和途径点）。

Caused-motion VDCs: Class [3]-[7]
Class [3]-[7]为致使移动型动趋式(caused-motion)。这类动趋式不同于自移型动趋式（[1]和[2])的地方在于：

> 语义上，表现为某物体直接或间接地由于某主体而产生移动或发生变化，如：
>
> > *他们把桌子搬上楼。*（他们使桌子发生移动）
> > *她终于寄回了那封信。*（她使那封信发生移动）

Figure 1. A Sample Page of the VDC Annotation Guideline (in Chinese)

Three annotators were involved in labeling the dataset on all three tasks. I, as the main annotator labeled all 15,852 instances. Two other Chinese native speakers, with background in either Construction Grammar or advanced Chinese language teaching, annotated 3,000 instances randomly selected from the total dataset within two months. To evaluate the possibility of expanding the annotation project to different annotators, the inter-annotator agreement (IAA) scores and ranking after adjudication are used. IAA is calculated by the percentage of the shared labels.

	Task 1	Task 2	Task 3
Annotator1	0.7	0.89	0.31
Annotator2	0.59	0.86	0.44
Mean	0.65	0.88	0.38

Table 3. IAA scores for the Two Annotators

Table 3 shows each annotator's performance compared against the gold standard. Given the time constraints of their work, the results indicate the plausibility of developing these tasks into annotation projects on a larger scale.

Task 1 was a 15-class classification task, including non-VDCs and 14 VDC classes. To accomplish this, annotators were required to build up their own standards of analysis for events evoked by the interaction between V1s and DVs, of which the former is an open class. Task 2, as noted, was the binary classification of metaphors. This task achieved the highest agreement rate (i.e., 0.88) among all tasks. Task 3 (coercion detection) was also involved a binary classification. However, the distribution among the two labels is extremely imbalanced (about 1:100), which renders the agreement score (i.e., 0.38) less reliable. Both annotators reported the challenge of the time constraint in accomplishing all three tasks, particularly Tasks 1 and 3. Also, the two annotators developed their own strategy. For example, the annotator with the Construction Grammar background conceptualized coercion as a low-frequency phenomenon, while the other annotator, whose background was Chinese language teaching, was more sensitive to the creativity of VDC tokens. These different cognitive strategies may have contributed to their labeling differences. In future work, agreement can be improved by developing a more detailed guideline that includes more examples, providing additional annotation time (or more pre-annotation practice with trouble shooting), and unifying individual strategies.

The second annotation evaluation method is ranking after adjudication, which indicates how reliable an annotator's decisions are when there is a complete or partial divergence among the annotators. This reliability is measured by the percentage of matching decisions between the annotator's decisions and the adjudications from group discussion (Pustejovsky & Stubbs 2012: 134).

	Tao Lin	Annotator 1	Annotator 2
Task 1	55%	30%	15%
Task 2	75%	40%	25%
Task 3	60%	60%	35%

Table 4. Percentages of Annotator's Decisions in Disagreement

Table 4 shows how often the annotators were correct when their labels differ in all three tasks. In each task, 20 VDC instances in which annotators disagree were selected[7] (complete disagreements for Task 1, but partial ones for Task 2 & 3). The commonly agreed labels of these instances were created after group discussion. In general, my annotation is always more reliable than the other two annotators. However, in the coercion detection classification (Task 3), given that coercion is dependent on the annotators' intuition on the argument information of V1s, the other two annotators' opinions are also valuable.

2.5 Algorithms and Experimental Design

Two algorithms were used for the implementation of all the classification tasks: deep neural network and Support Vector Machine (SVM). Each of the algorithms presents advantages and disadvantages, as discussed in Section 2.5.2. Settings of experimental design are given in Section 2.5.3.

2.5.1 Deep Neural Network

[7] The instances for adjudication were selected from the beginning of the dataset, where the annotators made more mistakes. The expected adjudication agreement scores for both annotators should be higher than their performance in Table 4.

This section starts with the neural model used and its architecture and then introduces some configurations of the model. Deep neural network is a technique using multiple layers of non-linear processing units to learn representations of different levels of abstraction in data (Deng & Yu 2014). One kind of deep neural network is bidirectional long short-term memory (biLSTM) network (Schuster & Paliwal 1997). This bidirectional network can traverse a character sequence from both left and right and is good at learning long-distance dependencies such as relative clauses or topicalization in Chinese. Every input node receives both left and right representations. The structure of the biLSTM network for binary classification is demonstrated in Figure 2.

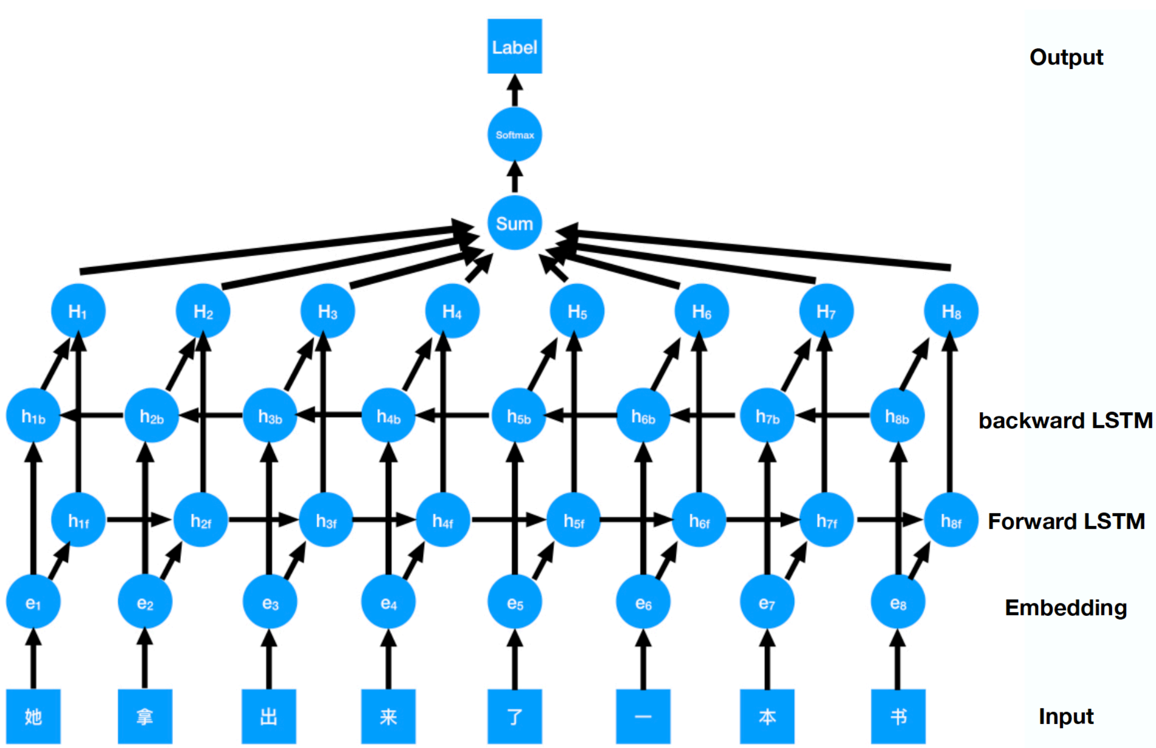

Figure 2. The Architecture of a Binary Classification System with biLSTM

Each input sentence is a sequence of words [w_1, w_2, w_3, …, w_n]. In the embedding layer, a vector represents each word. Then the LSTM layer, including forward and backward components, computes the state sequence with [H_1, H_2, H_3, …, H_n][8]. The forward and backward LSTMs traverse the word sequence and produce vectors h_{if} and h_{ib}. This processing is recurrent because the vector of w_{n-1} is always appended into the representation of w_n. The h_{if} and h_{ib} vectors are concatenated ("sum" in Figure 2), following which the total score is passed to a softmax function[9] to output a final predicted label. This label will be compared against the gold standard from the annotation for accuracy measures.

As mentioned in the architecture in Figure 2, word embedding is a multi-dimensional numerical representation of words. Embedding can be pre-trained in other neural models to introduce additional distributional semantics from external resources (Mikolov et al. 2013; Pennington et al. 2014). The semantics, which are based on word occurrences in large-scale corpora, are distributional because they contain the information of "you shall know a word by the

---

[8] The following matrix manipulations came from the model in Hochreiter & Schmidhuber 1997. W_n is E_n in Figure 2:

$i_t = \sigma(W_w^i w_t + W_h^i h_{t-1} + W_c^i w_{t-1} + b_i)$,
$f_t = \sigma(W_w^f w_t + W_h^f h_{t-1} + W_c^f w_{t-1} + b_f)$,
$g_t = tanh(W_w^c w_t + W_h^c h_{t-1} + b_c)$,
$c_t = f_t \odot c_{t-1} + i_t \odot g_t$,
$o_t = \sigma(W_w^o w_t + W_h^o h_{t-1} + W_c^o w_{t-1} + b_o)$, and
$h_t = \tanh(c_t) \odot o_t$,

in which σ is the non-linear sigmoid function, $\odot$ stands for matrix multiplication, and *f, i, o, c*, and *h* are the vectors of the forget gate, input gate, output gate, memory cell, and hidden state, respectively.

[9] The softmax function (Goldberg 2017) is an exponential function that forces y to be interpretable through:

$$P(y|x) = \frac{e^{w \cdot \varphi(x,y)}}{\Sigma_{\tilde{y}} e^{w \cdot \varphi(x,\tilde{y})}},$$

where $\varphi(x, y)$ is the function in the model, e.g., a linear function *y=Wx+b*, and y and $\tilde{y}$ are the predicted label and the gold standard respectively.

Figure 2 demonstrates binary classification by using one softmax function. In the case of multiclass classification, several softmax functions are required. For example, the code of a 6-class softmax in Keras is realized as below:

```
model.add(Dense(6, activation='softmax'))
```

company it keeps" (Firth 1957: 11). For example, embedding can be used to identify synonyms within a certain word window. The embedding vector of *Hogwarts* is numerically related to those of other Harry Potter words like *Dumbledore*, *Malfoy*, and *half-blood* (Levy & Goldberg 2014). In this research, the Chinese pre-trained word embedding developed by Beijing Normal University and Renmin University (Li et al. 2018) used in this project includes two types of 300-dimension embedding: word-level and word-and-character-level. Different settings of embedding as experimental conditions will be introduced in Section 2.5.3.

The biLSTM model in this project is implemented in Keras[10]. The annotated data is split into training, validation, and testing with ratios of 72%, 18%, and 10%, respectively. Some hyper-parameters are listed in Table 5:

Epoch	50
Batch Size	128
LSTM dimension	128
Vocabulary size	40000
Dropout Rate	0.1
Optimizer	Adam (learning rate = 0.001)
Embedding dimension	300
Loss functions	Categorical Cross-entropy
Activation function	Softmax
Class weight	Based on class frequency

Table 5. Some Shared Hyperparameters in the biLSTM Classifiers

[10] https://keras.io/

2.5.2 Support Vector Machine

The advantage of neural models like biLSTM is their ability to discover distributional features. They can learn various features (whether these features can be recognized by humans or not) and their relationships based on a vast amount of data. A possible weakness of neural models is that they rely too much on the data size and can be sensitive to frequency effects, such as the distribution in imbalanced, multi-class datasets. Since some of our classification tasks involve imbalanced, multi-class detection, SVM was chosen as a comparison against biLSTM. It is expected that the performance of biLSTM models should be close to or better than that of SVM models. This section introduces the basic mechanism, feature selection, and parameter settings of the SVM models.

The goal of SVM is to discover a hyper-plane as a decision boundary between classes (Corte & Vapnik 1995). This hyper-plane needs to be maximally distant from the nearest training examples. Since data is represented as vectors in the decision space built by features, the example vectors which can identify the decision boundary, are called “support vectors.”

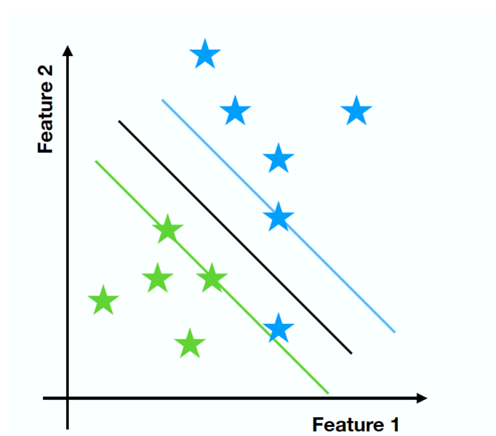

Figure 3. Demonstration of a Binary SVM

The figure above is an example of a linear binary SVM classifier[11]. Imagine that we want this classifier to distinguish two constructions (i.e., the blue and green stars). The vector representations of these constructions fall into two dimensions. The problem in the task is thus to find out the key stars (the blue and green ones on the lines) that identify an ideal classifying boundary (the black line). Notice that one blue star is close to the green line. Its distance from the blue line shows that it is a slack variable. SVM has a device to reduce the weight of outlier cases like this. Sometimes the classes cannot be linearly classified or involve complicated features from multiple dimensions. To handle this situation, we require a new function class, called a "Kernel," to take two examples and return their similarity. Kernels can help us find non-linear decision boundaries (Shawe-Taylor & Cristianini 2004).

One challenge for SVM models is feature selection and engineering. I chose four linguistic features to construct the inputs: raw data, words, word-POS pairs, and dependency trees of the inputs. Compared to my pilot study, the features of POS and dependency trees were added. First, the raw data and their n-grams string are collected. This feature set includes useful information, such as V1s and argument nouns. Second, words paired with their POS labels are included, i.e., in the format of "word_POS." Both segmentation and POS labeling are supported by THULAC. Last, the Chinese dependency trees come from Stanford Universal Dependency (UD) Parsing (Qi et al. 2019) is adapted to capture possible long-distance dependencies. The format of the parse trees for Example 1 *deng shang dui pa shang le zuigaofeng* 'The

---

[11] A mathematical description of support vectors is given below:

$$\min_{w,b,\zeta} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \zeta_i \ ,$$

subject to $y_i(w^T \phi(x_i) + b) \geq 1 - \zeta_i, \zeta_i \geq 0, i = 1, 2, \ldots, n.$

w is the weight vector, $\frac{1}{2} w^T w$ is the margin. ζ_i is the slack variable, and C is a regularization parameter to show the cost of wrongness.

mountaineering party climbed up the highest peak.' is given below. Each line is taken as a sub-feature of a dependency tree.

```
('登山', '2', 'case:suff')
('队', '3', 'nsubj')
('爬上', '0', 'root')
('了', '3', 'case:aspect')
('最高峰', '3', 'xcomp')
('。', '3', 'punct')
```

Figure 4. A Stanford UD Tree for Example 1

Features like n-grams, POS, and dependency trees can bring rich information about linguistic properties to the classifier by highly lengthening the feature vectors. Taking features of n-grams, POS, and dependency nodes into account, the vector length can reach about 4,600. The question is how to make the program computationally less expensive. I used two techniques to reduce the high dimensional vectors to low dimension: hashing vectors (Weinberger 2009) in my pilot models and Principal Component Analysis (PCA) (Jolliffe 2010) in the current models. Compared to hashing vectors, PCA is more capable of preserving more linguistic information in the reduced dimensions in most of the VDC classification tasks. The library used for SVM is scikit-learn[12]. The main parameters are summarized in the table below:

[12] http://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

Penalty parameter C	C=12
Kernel Function	$\exp(\gamma\|x - x'\|)^2$ ("rfr")
Kernel Parameter	1 / number of features
Class Weight	Balanced
Decision Function	one vs. rest ("ovr")
Hashing Features	20
PCA Dimensions	500

Table 6. Some Shared Features in the SVM models

2.5.3 Experimental Conditions, A Comparison with Pilot Results, and Evaluation

In total, I consider 14 experimental conditions in each of the classification tasks. These conditions include different settings of algorithms, embedding, and input lengths, the last of which was newly introduced after my pilot study. First, as mentioned in Sec.2.5.2-2.5.3, I use two algorithms in this study: SVM and biLSTM (repeated in Table 7). The SVM model with the linguistic features is considered as one setting, which only interacts with inputs length.

Algorithm	Description
SVM	Support-Vector Machine
"EM"-initial Models	Bi-directional long short-term memory

Table 7. Model Setting (I): Algorithms

Second, as shown in Table 8, there are three kinds of embedding in biLSTMs: randomized vectors in the embedding layer ("EM0"), pre-trained word-level ("EM1"), and word-and-character-level ("EM2") vectors (Li et al. 2018). The latter two conditions are expected to perform better than the prior because they contain richer distribution semantics from other corpora data.

Embedding (EM)	Description
0	Randomized embedding
1	Word-level embedding
2	Word-and-character-level embedding

Table 8. Model Setting (II): Three Embedding Conditions of biSLTM

Also, Table 9 shows that within these neural models the input data has two kinds of segmentations: word-level and word-and-character-level. Since raw Chinese texts do not exhibit space separation, we can either segment sentences with space (marked as "char") or separate them with word boundaries (marked as "seg") (i.e., using the segmentation model in THULAC).

Segmentation	Description
Word ("seg")	THULAC word segmentation
Character ("char")	Inputs strings are separated by space

Table 9. Model Setting (III): Two Conditions of Input Segmentation

Finally, the length of inputs may affect the predictability in different tasks. As shown in Figure 5, the CCL input collected from the DV-based search contains at least one clause, underlined in blue, which contains the VDC (i.e., the red characters), and other clauses or strings as the surrounding context, underlined in green. The average character length of the latter is 40, while the former is only 15. This condition includes two string inputs: long context (the green underline; marked as "long") vs. short context (the blue underline; marked as "short")[13].

[13] The condition of inputs containing both long and short contexts was not used, because it did not outperform the best results in the models.

Figure 5. Two Conditions of the Input Length

Context	Description
“long”	Full inputs from corpora
“short”	Inputs of clauses that contain VDCs

Table 10. Model Setting (IV): Two Conditions of Input Length

For classification tasks that mainly rely on VDC information at a clausal level, e.g., VDC event types, having the surrounding strings can be misleading because there can be related events from other constructions in the context. However, other tasks like metaphor detection, the full context might be more useful, because certain implicit arguments can be referred in other clauses.

In summary, for each of the five classification tasks, including VDC detection, VDC coarse and fine event type classifications, metaphorical VDC detection, and coerced VDC detection, there are 14 experimental conditions. These conditions and their model settings are summarized in Table 11.

Model Type	Embedding			Segmentation		Input Length		SVM
	Randomized	Word	Word & Character	Word	Character	Long	Short	
EM0+char+long	☑				☑	☑		
EM0+seg+long	☑			☑		☑		
EM1+char+long		☑			☑	☑		
EM1+seg+long		☑		☑		☑		
EM2+char+long			☑		☑	☑		
EM2+seg+long			☑	☑		☑		
SVM+long						☑		☑
EM0+char+short	☑				☑		☑	
EM0+seg+short	☑			☑			☑	
EM1+char+short		☑			☑		☑	
EM1+seg+short		☑		☑			☑	
EM2+char+short			☑		☑		☑	
EM2+seg+short			☑	☑			☑	
SVM+short							☑	☑

Table 11. Model Setting (V): A Summary of Models for Each Classification Task

Also, during my prospectus, pilot studies on VDC detection, event classification, and metaphor detection were conducted on 3,000 CTB instances. Later, I improved the study by revising the models and increasing the instances from 3,000 to 15,852. In each of the task, I compared the best results from the current studies to the pilot ones to evaluate the efficacies of my changes. Both pilot and current models were compared on the pilot testing data. Apart from being trained on more data, the current LSTMs and SVMs were debugged, and certain parameters were changed. The main parameter changes include: In SVMs, I reworked the tokenizer and feature extraction and replaced hashing vectors with PCA. In biLSTMs, I changed the loss functions of “sparse_categorical_crossentropy” into “categorical_crossentropy” in the current models. In most of the cases, these changes improved system performance.

Finally, I chose to use two Scikit-learn evaluation settings for the classification in this project. As for the binary classification tasks, I used “average=‘binary’,” which means only reporting the precision, recall, and F-score for the positive class. In my annotation, the positive label means VDC or metaphor in the two binary tasks. Given that the nature of my binary classification tasks is to detect only one class out of the two, reporting the performance of the key class is conceptually more important than that of the noise information. As for multiclass classifications, I used “average= ‘weighted’,” which means the final precision, recall, and F-score are the sum of the products of weights and measures for each class. When classifying 6 or 14 VDC classes with equal conceptual status, the “weighted” parameter can factor out any inherent performance improvements based solely on frequency, which gives a clearer picture of overall performance.

Chapter 3
Classifying Categories of Event Structures in VDCs

Research Question: How can we develop classification tasks to label event types of VDCs?

This chapter describes the task of detecting distinct event types, each with specified participant roles. Prior to classification of event types, we must first develop taxonomies and model the construction functions in the taxonomies with linguistic theory. As discussed in Chapter 1, one of the major challenges in VDC processing is to distinguish the various VDC functions. This section thus begins with a synthesis of developing taxonomies of VDCs based on coarse-grained functions (Sec.3.1) and fine-grained functions (Sec.3.2), and then applies principles of Sign-Based Construction Grammar to describe the semantic and syntactic properties of VDCs, and in particular the properties that follow from their being types of control predicates (control and raising types) (Sec.3.3). Finally, based upon this understanding of VDC multi-functionality and control properties, as well as the progress of VDC-related parsing in Chinese NLP (Sec.3.4), the chapter concludes with the presentation of a series of classification tasks using the synthesized VDC taxonomies (Sec.3.5).

3.1 Analyzing VDC functions (I): A Coarse-grained VDC Taxonomy

As mentioned in Section 1, the multi-functionality of VDC causes problems for NLP, both in theory and applications. I developed two VDC taxonomies to face the challenges: coarse-grained and fine-grained classes. The former identifies the prototypical types of VDC events, and the latter focuses on richer event semantics, such as subtypes of Self-Motion and different metaphoric interpretations of Caused-Motion events. This section will introduce coarse-grained taxonomy of VDCs. There are six coarse-grained functions. These functions differ not only in their event types, but also in their sharing of Frame Elements by V1 and DV across sentences. Leveraging the argument-sharing feature for typological classification of SVCs (Haspelmath 2016) means that both V1 and DV occurring in a single VDC can act as independent verbs. For example, a VDC such as "Subject V1 DV Object" can be converted into the diagnostics "Subject V1 Object" and "Subject DV Object." **In my VDC research, "argument sharing" is defined as situations of syntactic role sharing between the valence structures of V1s and DVs, not their semantic role sharing**.

3.1.1 Self-Motion

The first coarse-grained function is *Self-Motion*, in which a VDC indicates that an entity moves with a certain orientation relative to some spatial reference:

12. A. deng-shan-dui pa shang le zui-gao-feng
   climb-mountain-team climb ascend ASP SUP-high-peak
   "The mountaineering party climbed up to the highest peak." (cf. Google Search)

   B. deng-shan-dui pa le zui-gao-feng "The mountaineering party climbed the highest peak."

   C. deng-shan-dui shang le zui-gao-feng "The mountaineering party moved up to the highest peak."

In Sentence 12A, the VDC is *pa shang* 'climb onto,' while (12B) and (12C) demonstrate the application of argument-sharing tests. According to FrameNet (Baker, Fillmore, & Lowe 1998; You et al. 2007) (FN), the corresponding frame in (12A) is Motion. The Figure *dengshandui* 'the mountain climbing party' moves to the Goal *zuigaofeng* 'the highest peak' in the Manner of *pa* 'climbing' with an upward Path. (12A) can then be syntactically converted into (12B) and (12C), respectively, in which either the manner verb or the DV can be a stand-alone predicate for the Theme and Goal.

3.1.2 Caused-Motion

Related to Self-Motion, the second coarse-grained function of VDCs is *Caused-Motion*:

13. A. xuesheng-men <u>gan</u> laoshi <u>xia</u> le tai

student-PL force teacher descend ASP podium

"The students drove the teacher off the stage." (cf. Google Search)

B. xuesheng-men <u>gan</u> le laoshi "The students drove the teacher."

C. laoshi <u>xia</u> le tai "The teacher moved off the podium."

The event structure in Sentence 13A can be explained by the Cause_motion Frame: the Agent *xueshengmen* 'the students' cause the Theme *laoshi* 'the teacher' to move down from the

Source *tai* ‘the podium’ in the Manner of *gan* ‘drive.’ In (13A) and (13B), the sentences “the students drove the teacher” and “the teacher moved down the podium” are grammatical sentences. The role that was caused to move is the shared Theme in both the forcing and the descending actions.

3.1.3 Resultative

The third coarse-grained function is *Resultative*, which is related to the completion of an action. It is not purely aspectual because DVs may or may not add DV-specific readings of the result of an action. An example is the resultative use of *xia* ‘descend’:

14. A. shuqi xianggang <u>mai</u> <u>xia</u> hao-zhai

HsuChi Hongkong buy descend expensive-house

“Hsu Chi bought an expensive house in Hong Kong.” (cf. Google Search)

B. shuqi xianggang mai haozhai “Hsu Chi bought the expensive house.”

C. *shuqi xianggang xia haozhai “Hsu Chi finished the expensive house in Hong Kong.”

The basic meaning of Sentence 14A is the accomplishment of Hsu Chi’s purchase action. However, an additional implicature in *xia* ‘descend’ is that it did not take much effort for the speaker to buy. A native speaker can test for this implication by changing the DV into *shang* ‘ascend,’ which instead implies that achieving the result was difficult. (14C) shows that the DV in this function cannot act as an independent verb given its abstract meaning. The DV *xia* can be interpreted as a completive function that takes the state of affairs *shuqi mai haozhai* 'Hsu Chi buys an expensive house' as its argument.

3.1.4 Aspectual

The fourth major function is *Aspectual*. Just like many other languages (Bybee & Dahl 1989), some directional morphemes in VDC developed into temporal modifications of the V1. One example is the use of the inchoative DVs *qi lai* 'rise come':

15. A. wu-li renao qi lai

house-inside cheery rise come

"The room started getting cheerier." (cf. Liu et al. 1998)

B. wuli renao le "The house was cheery."

C.*wuli qi lai le "The house started."

Sentence 15A has two frames: the frame evoked by V1 (i.e., Emotion_directed) and an aspectual frame (i.e., Beginning). The aspectual frame takes the action of V1 as its semantic argument. Similar to the property in Resultative VDCs, the compounding DVs *qi lai* 'rise come' are highly grammaticalized so they cannot act as independent verbs, as demonstrated in (15C).

3.1.5 Discourse-Connective

The fifth major function is *Discourse-Connective*, in which a discourse marker that initiates or links pieces of discourse, like *speaking of* and *next* in English. For example, the VDC *shuo qi* lit ('speak rise') 'speaking of…' is often used to initiate the following discourse:

16. A. shuo qi aomen, ren-men ziran hui xiang dao da-san-ba,

speak rise Maucau, person-PL naturally MOD think reach big-three-eight

mazu-dong-wang dengdeng jingdian

Mazu-east-look and.so.on view-point

"Talking about Macau, people would naturally recall places of interests like Ruins of St. Paul's and A-Ma Temple." (cf. CTB)

B. (wo-men) shuo aomen "We speak of Macau."

C. *(wo-men) qi "We started."

In Sentence 16A, the VDC *shuo qi* evokes the following sentence, namely people's impression of the places to visit in Macau. The original sense of DV *qi* is inchoative, indicating the beginning of the speaking event. Therefore, both Chatting and Beginning Frames can describe the function of the VDC in (16A). (16B) shows that there is an implicit subject for the V1 *shuo* 'speak,' usually the Interlocutor *women* 'we'. Macau serves as the Topic mentioned by the Interlocutor. However, in (16C), this implicit subject is not able to take the inchoative DV.

Similar to our treatment of Aspectual and some Resultative VDCs, the DV is treated as a raising verb that takes the whole state of affairs "(the Interlocutor) speaks TOPIC" as an argument.

3.1.6 Evidential

The last kind of VDC, *Evidential*, encodes evidential modality: the speaker's source of evidence for some state of affairs:

17. A. ta (ting shang qu) bu da-diao

she sound ascend come NEG get-tune

"She sounds like she is going off key." (cf. Liu et al. 1998)

B. wo ting "I listen."

C.*ta shang qu "She seems"

The parentheses show that the use of the VDC *ting shang qu* lit ('listen ascend go') 'it sounds like…' is pragmatically optional. Without the VDC, Sentence 17 becomes a factual rather than inferential statement s*he is off-key*. The addition of this VDC highlights the assessment of the speaker based on auditory evidence. The V1 evokes a Perception_active Frame with two main elements: Perceiver (i.e., the speaker) and Phenomenon (i.e., her singing); the Perceiver, the speaker, is unspecified. The V2 *shangqu* has a similar function to an aspectual raising V2. When modeling the semantics of the DVs *shangqu*, the V2 can be interpreted as taking the whole state of affairs, including *ta* 'she,' *ting* 'sound,' and *buzhaodiao* 'be off key' as arguments.

Table 12 summarizes the semantic and syntactic features of the coarse-grained functions of VDCs. Each function has its unique event type and essential frame. Also, an argument-sharing test (i.e., predicate split) can be used as a diagnostic to distinguish different functions. Assuming the similarity of DVs to control/raising verb lexemes (Huang 1992), the sharing of two arguments can be labeled as "two-arg-sharing," the Agent role sharing as "subject-control," the Theme role sharing as "object-control," and lack of sharing (which entails that each verb has a "state of affair" or "proposition" argument) as "raising" (Kim & Michaelis, to appear: 167).

VDC function	Event Description	Arg-sharing	Core Frame
Self-Motion VDCs	A Self-mover moves in some Path with respect to some Location	Two-arg-sharing or Subject-Control	Motion-fr
Caused-Motion VDCs	An Agent causes a Theme to move to a Goal along a Path	Object-Control	Cause_motion-fr
Resultative VDCs	An action is achieved or accomplished with specific results	Raising	Completing-fr
Aspectual VDCs	The temporal structure of an action or state change is specified	Raising	Beginning/ Continuing-fr
Discourse-Connective VDCs	A clause is introduced with its relation to previous element(s) in discourse	Raising	Communication -fr
Evidential VDCs	The speaker assesses some state of affairs via a Phenomenon	Raising	Perception_active-fr[14]

Table 12. Summary of the Semantic and Syntactic Features of Coarse-grained VDC Taxonomy

[14] An alternative is Appearance-frame (Hsieh et al. 2006).

3.2 Analyzing VDC Functions (II): A Fine-grained VDC Taxonomy

The preceding section (Sec.3.1) provided a general classification of VDCs with several basic event types, such as motion, aspect, result, and discourse modality. However, richer event categories (as shown in the numbered categories in Figure 6), along with their frame semantics, can be established under most of the coarse-grained classes. An English demonstration of this further classification comes from the VPC examples *move up the mountain* and *come up to me.* They both denote concrete agentive motion, but the former event shows literal upward motion while the latter encodes an axis shift, e.g., from the vertical axis to the horizontal. To accommodate similar differences in VDCs for NLP purposes, I create two subtypes of event structures for Self-Motion: motion with and without axis shift. This section introduces a fine-grained VDC taxonomy based on the previously described (Sec.3.1) categories of Self-Motion (Sec.3.2.1), Caused-Motion (Sec.3.2.2), Aspectual (Sec.3.2.3), and Resultative VDCs (Sec.3.2.4). The numbered categories in the figure below presents the taxonomy which will be explained in this Sec.3.2.

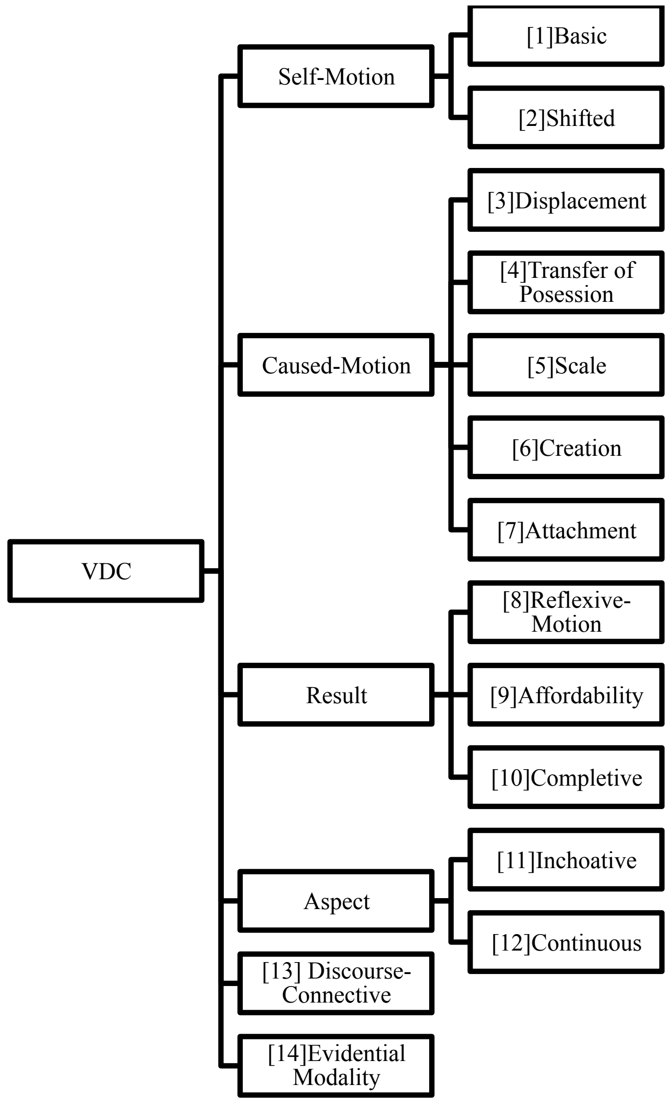

Figure 6. A Taxonomy for VDC Event Classification

3.2.1 Self-Motion: Basic vs. Shifted

Self-Motion means a Figure that propels itself to move along a Path. As mentioned in the English VPC examples (Sec.3.2), the primary distinction[15] in my fine-grained taxonomy of Self-Motion VDCs is the difference between literal and shifted orientations (Liu et al. 1998). *Self-Motion-Basic* VDCs mean that DVs retain their original meanings. Sentence 1 is repeated as (18), in which the DV *shang* does not change its vertical orientation.

18. deng-shan-dui pa shang le zui-gao-feng

climb-mountain-team climb ascend ASP SUP-high-peak

"The mountaineering party climbed up to the highest peak." (cf. Google Search)

In *Self-Motion-Shifted* VDCs, the direction of DVs usually changes from a vertical axis to a horizontal or an unspecified one, as seen in Sentence 19:

19. ni pao de zhen man, wo gan shang ni le

you run DE real slow, I chase ascend you ASP

"You ran so slow; I caught up with you." (cf. CCL)

Compared to Sentence 18, the same DV *shang* means 'approaching (usually on a horizontal axis),' so vertical motion is not always necessary.

[15] Another distinction discussed in Chinese linguistic literature is the use of deictic function of the DV *lai* 'come' and *qu* 'go' (Chen 2007). The verb *lai* indicates that the moving object approaches some reference point (usually the speaker or narrator), while *qu* indicates that it moves away from that point. Given that reference points are often vague in the context and thus difficult to identify, this feature is not considered in my taxonomy.

3.2.2 Caused-Motion

Similar to Self-Motion VDCs, Caused-Motion VDCs (CM-VDCs) also express directed motion events. However, the prototypical meaning of caused-Motion is that an Agent exerts a force on a Theme causing the Theme to move relative to a location (Source/Goal) via a certain Manner and Path (Goldberg 1995). For example, in *she put the book on the desk*, the person caused the object to move from an unspecified location to the top of the desk. However, apart from the literal displacement events, VPCs like *he composed a new series of events into his ever-growing novel* and *Mary bought the coat from the Salvation Army* show different figurative interpretations of caused-motion events, especially in the metaphoric readings of the Location and Theme. Since Hwang's (2014) Caused-Motion event taxonomy in English has been successfully implemented as a machine learning task, I have applied her analyses (the first four categories below) to Chinese VDCs and added a fifth category. My analysis of CM-VDCs is in addition to the classification in the dissertation proposal.

3.2.2.1 CM-Displacement

Displacement is the first fine-grained class of CM-VDCs, as indicated by the prototypical definition and the "put" VPC example discussed above. This class covers a concrete motion process and its metaphoric extensions, applied in the Cause_motion Frame.

20. ren-men ba laji <u>diu</u> <u>jin</u> laji-tong

person-PL BA trash throw enter trash-can

"People throw trash into the trash cans." (cf. Google Search)

Sentence 20 is a concrete motion event in which the throwing action of the people caused the trash to change its physical location. In addition, as in this example, we can also see that CM-VDCs usually appear in an object-fronting/BA construction, which will be taken into account in the section on grammatical modeling.

3.2.2.2 CM-Transfer-of-Possession

With the second CM-VDC class, *Transfer of Possession*, the metaphoric interpretation of the location of either Goal or Source, is reanalyzed as a target of a change of possession event. The semantics of this construction can be described by frames related to changes of possession in FN, such as Borrowing, Commerce_buy, and Sending Frames. In the English VPC *she borrowed $10 from me*, she caused the money to move from one owner (*her*) to another (*me*).

21. lao-lü xiang ba dian pan hui ziji shou-li
    old-lü want BA store buy back REFLX hand-inside
    "Old Lü wanted to buy the store back." (cf. Google Search)

Sentence 21 is a purchase event in which Old Lü had an intention to change the possession of the store from someone other than himself to himself. The noun phrase *ziji shouli* 'the inside of his own hand' in Chinese is a metonym for his possession of some object.

3.2.2.3 CM-Scale

The third CM-VDC class, *Scale*, applies to a process of caused motion in which the Theme moves along a measurable range (realized as a numerical axis or scope), resulting in a value change, e.g., the VPC *The government increased the tax to $500*. In this VPC, the government moved the value of the tax from a lower number to a higher one by increasing the tax. A VDC example is given below:

22. ta zi-ding yao ba ziji de fen-shu <u>ti</u> <u>shang</u> <u>qu</u>

    he Self-determine want BA REFLX ASSOC score-number raise ascend go

    "He determined to raise his score." (cf. CCL)

The difference between the VPC in the English example *The government increased the tax to $500* above and Sentence 22 is that the scale measurements, e.g., score points, are implicit in (22) while the endpoint in the English example is explicit. However, the inference in (22) is still one about upward motion from low scores up to higher scores.

3.2.2.4 CM-Creation

The fourth class of caused motion indicates a process of creating or revealing the Theme, and is thus called *Creation*. Similar to the use of the directional morpheme *out of / from* in English CM events, as in *I grew it out of / from a seed*, this class in CM-VDC is specifically related to the DV *chu* 'exit.'

23. wo-men yijing qiu chu le tong-jie

I-PL already get exit ASP general-answer

"We already figured out the general solution (to the math question)." (cf. CCL)

By using the DV *chu* in Sentence 23, the Resultative VDC does not mean only that the calculation was performed, but also that a Created_entity was successfully produced by the Cause. The orientational metaphor involved here is CREATION IS EXITING, in which the hidden or non-existing state of the solution is mapped to a contained space and the creation of the solution is mapped onto a space that is outside the original space.

3.2.2.5 CM-Attachment

The last CM class is featured by attachment, fixation, and connection, or their reversed processes, called *Attachment*. This category can be exemplified in Chinese by the VDCs of putting on/taking off clothes in Chinese:

24. wo zhengzai chuan shang / tuo xia mao-yi

I ASP (PROG) wear ascend / separate descend fur-clothes

"I am putting on / taking off the sweater." (cf. Google Search)

In Sentence 24, the vertical motion of ascending and descending is mapped onto attachment and detachment, respectively. The locative element (Goal or Source) is not realized in the sentence, e.g., arguments like *yigui* 'the closet' or *shenshang* 'the surface of one's body' are not used in (24).

There are various kinds of metaphorical interpretations of locatives in VDCs. Specifically, the extensions of purchase activities, possession change, creation, and attachment are important metaphoric variants of caused motion events, so they are listed as separate classes of CM-VDCs.

3.2.3 Aspectual VDCs: Inchoative vs. Continuous

Aspectual VDCs indicate the inception or continuation of a state or activity. As mentioned in the coarse-grained taxonomy, since aspectual DVs function like suffixes of V1s, therefore they do not usually support any control patterns (i.e., one shared subject or object). There are understood to be three types of aspect in VDCs: inchoative, continuous, and completive. However, in Section 3.1, I argued completive aspect should not fall into the Aspectual category, but should be considered as another class (as will be further demonstrated in Sec.3.2.4). Here are examples of Inchoative and Continuous VDCs.

25. A. ling-sheng xiang <u>qi</u>

    bell-sound ring ASP(rise)

    "The bell started ringing." (cf. Google Search)

    B. zhong ri liang guo yao shi-shi-dai-dai

    China Japan two country should generation-generation-generation-generation

    <u>you-hao</u> <u>xia</u> <u>qu</u>

friend-good ASP(descend) ASP(go)

"China and Japan should continue being friendly to each other for generations." (cf. CTB)

In Sentence 25A, the bell ringing has begun but has not been completed. The inchoative aspect is usually marked by *qi* (*lai*) 'rise (come)' and sometimes by *shang* 'ascend'. But the DV *xia qu* 'ascend go' in (25B) shows the ongoing nature of friendship between the two countries.

3.2.4 Resultative

Although the Resultative class is one of the most controversial categories in the taxonomy, in Section 3.1 I use either the accomplishment or achievement of events (Vendler 1957) to define this coarse-grained class. Traditionally, Chinese linguists use the difference between "basic" and "non-basic" to describe the Resultative VDCs (Liu et al. 1998; Qi 2002). For example, the results of "wearing clothes" and "closing one's eyes" are natural consequences, which do not have a high cost. Therefore they represent basic results. On the other hand, successful results of "making friends" and "taking an exam" require resources like time, effort, and money, so these events are non-basic. However, this approach to Resultative VDCs relies mainly on the semantics of various V1s and even on the context of the VDC, and there is no effective morphosyntactic standard to differentiate between the two categories. Rather than using the information of V1 to differentiate the basic vs. non-basic, I recognized two results encoded by DV and put the remaining Resultative cases into a third class.

3.2.4.1 Reflexive-Motion

The **first** category of Resultative VDCs denotes the expansion or contraction from one shape or size to another and is referred to as *Reflexive-Motion* (Lindner 1981). It can be described by the Expansion Frame in FN. Particularly, the DV *kai* 'move away' has an extended sense 'be open' (Wang 2013) (similar to *up* in English), and the DV *qi* 'rise' usually co-occurs with shrinking and contracting actions in V1s (similar to *out* in English). To illustrate, Sentence 26 shows a pair of VDCs with opposite meanings.

26. guniang da kai / shou qi le yu-san

girl hit open / shrink rise ASP rain-umbrella

"The girl opened / shut the umbrella." (cf. Google Search)

Additionally, in some cases the metaphoric use of *kai*-related VDCs is conceptually close to inchoative aspect. For example:

27. zheng-ge wu-li luan kai le

whole-CL room-inside messy open ASP

"The whole room started getting messier and messier."

The DV *kai* in Sentence 27 is used to describe the inception of a process in which messiness gradually increases. However, I do not treat it as an Aspectual VDC for two reasons. First, compared to the use of the inchoative DV *qi* (*lai*) 'rise come' (Chao 1968) and continuous

DV *xia qu* 'descend go' (Gong 1995), the productivity of "V1+*kai*" has much more restricted combinatoric potential. V1 classes such as motion (**pao kai* 'run open'), emotion (*ai kai* 'love open'), some stative/predicative adjectives (**hao kai* 'be good open'), and some light verbs (**zuo kai* 'do open') do not allow modification by *kai*. Second, during pre-annotation, the annotators reported that they could semantically visualize the VDC "V1+*kai*," such as *luan kai* 'be.messy open' in (27), as the process of spreading. Therefore, it is reasonable to think this emergent use of inchoative as a metaphoric case of Resultative-Reflexive-Motion VDCs.

3.2.4.2 Affordability

The **second** category of Resultative VDCs evokes the *Affordability* frame. It refers to the agent's ability to bear the cost of action of V1; the V2 is *qi* 'rise.' This kind of VDC usually appears in Potential Constructions with infix-like morphemes *de* (achievable) or *bu* (non-achievable) (Lü 1979).

28. nan-you <u>mai</u> bu <u>qi</u> fang zenme-ban
male-friend buy NEG rise house how-do
"What should I do if my boyfriend cannot not afford a house?" (cf. Google Search)

Sentence 28 shows that when *qi* 'rise' is combined with some V1s, such as the purchase verb *mai* 'buy,' it refers to the buyer's ability to pay the cost of a house. Note that the Affordability class differs from the Transfer of Possession class in CM-VDCs, because it mainly highlights whether the buyer possesses assets or the resources sufficient to cover the cost of V1, not just the change in ownership.

3.2.4.3 Completive

The last type of Resultative VDC is Completive, and refers to the completion of the action expressed by V1. One of the most typical cases of Completive VDCs is "V1+*guo,*" in which *guo* 'pass' encodes only an accomplishment or achievement. The DV *guo* can follow different verbs in Sentence 29, which supports the view that it is already a fully developed aspectual marker in Chinese grammar (Liu 2015).

29. wo-men fu chu guo, qipan guo, huanxiao guo, beishang guo,
    I-PL pay exit ASP, expect ASP, laugh ASP, sad ASP,
    renao guo, chenji guo, zou guo, kan guo, lei guo, ao guo
    bustling ASP, quiet ASP, walk ASP, see ASP, tired ASP, live.through ASP

    "We made efforts. We expected (things). We had happy and sad moments. We had bustling and quiet times. We once walked and saw (things). We were tired. We survived (in difficulties)." (cf. CCL)

However, Completive is not purely aspectual because some idiomatic uses of DVs not only contain temporal information, but also specific results of V1s. These results are usually linked with certain manners in DV, such as finishing an action successfully, easily, or satisfactorily (see in Example 14); in other cases, the results are only applied to objects with certain selectional restrictions. Consider the following examples:

30. A. zongzhi, da-jia hao-xiang dou kan bu qi ta

in.summary, big-family good-resemble all look/evaluate NEG rise he

"In a word, it seems that everyone did not look up to him." (cf. Google Search)

B. ta genben kan bu shang zhe zhong pianyi-huo

she totally look/evaluate NEG ascend DEM kind(CL) cheap-goods

"She did not like this kind of cheap stuff at all." (cf. Google Search)

In Sentences 30A and 30B above, the VDCs *kan bu qi* and *kan bu shang* both mean 'finish evaluating and (as a result) give negative feedback,' but their selectional restrictions on the object are not the same. In both sentences, *kan bu shang* can be used, but *kan bu qi* would not be grammatically correct if used in the second sentence. This is because *kan bu qi* only applies to moral evaluation of persons and opinions; but *kan bu shang* is a more neutral evaluation of goods as well as people. Building up too many separate classes for idiomatic and idiosyncratic may make sentences like (30) too difficult for machines to learn, so I have maintained the Completive class as "all the non-Resultative-Reflexive-Motion/Affordability VDCs."

In Sections 3.2.1-3.2.4, I have outlined a fine-grained taxonomy of VDCs based on the classes for Self-Motion, Caused-Motion, Resultative, and Aspectual, which are summarized in Table 13. The table contains the labels of the VDC events, the control patterns, and the core frame that highlights each event type. The event type of each fine-grained subclass is described by frame semantics, although the syntax (i.e., argument-sharing patterns) does not necessarily differentiate among subclasses at this level.

VDC function	Event Description	Arg-sharing	Core-frame
Self-Motion-Basic	A Self-mover moves in some Path literally indicated by the DV	Two-arg-sharing or Subj-control	Motion-fr
Self-Motion-Shifted	A Self-mover approaches a Goal in some Path not literally indicated by the DV	Raising	Motion-fr
Caused-Motion-Displacement	An Agent causes a Theme to move from a Source to a Goal	Obj-control	Cause-Motion-fr
Caused-Motion-Transfer-of-Possession	A Buyer/Seller/Donor causes a Theme/Goods to change from Owner1 to Owner2	Obj-control	Commerce_buy-fr, Giving-fr, etc.
Caused-Motion-Scale	An Agent causes an Item to change from Value1 to Value2 in along Path	Obj-control	Cause_change_of_position_on_a_scale-fr
Caused-Motion-Creation	A Cause leads to the formation of a Created_entity	Raising	Create-fr
Caused-Motion-Attachment	An Agent causes an Item to be attached/connected to a Goal or detached/disconnected from a Source	Obj-control or Raising	Attach-fr
Result-Reflexive-Motion	An Item changes its physical sizes from Initial_size to Result_size	Obj-control or Raising	Expansion-fr
Result-affordability	A Payer uses an Asset to gain Goods	Raising	Expensiveness-fr
Result-completive	An Action is completed (in a certain Manner)	Raising	Completing-fr

Aspectual-inchoative	An Action starts	Raising	Beginning-fr
Aspectual-continuous	An Action continues but has not terminated	Raising	Continuing-fr
Discourse-Connective	A Clause is introduced by its relation to a Previous_Discourse	Raising	Discourse-connect-fr
Evidential-modality	The speaker assesses some state of affairs via a Phenomenon	Raising	Perception_active-fr

Table 13. Summary of the Semantic and Syntactic Features of Fine-grained VDC Categories

3.3 Modeling VDCs with Sign-Based Construction Grammar

To model the multi-functionality of VDCs, the theoretical approach taken in this project is Sign-Based Construction Grammar (SBCG) (Sag 2012, Michaelis 2009, 2012), which introduces constructions as syntactic rules with associated semantic specifications. SBCG originated from Berkeley Construction Grammar (BCG) (Fillmore & Kay 1993; Michaelis & Lambrecht 1996; Kay & Fillmore 1999; Michaelis & Ruppenhofer 2001) and Head-Driven Phrase-structure Grammar (HPSG) (Ginzburg & Sag 2000) (HPSG). There are four main advantages of SBCG in term of VDC event type modeling: (1) it uses a scale of idiomaticity to describe the semantics of constructions, (2) it captures relations among constructions and lexemes (including MWEs) by means of a type hierarchy that represents semantic and syntactic similarities among the types, (3) it offers an analysis of argument sharing (in which a single expression plays a different argument role in each of two conjoined verbs) that is based on coindexation of arguments rather than an unpronounced syntactic subject, and (4) it represents syntactic and semantic information in parallel, in a construction or word's feature-structure description. This section will start with the theoretical propositions (1) and (2), and then will illustrate (3) and (4) in the practice of SBCG formalism.

3.3.1 The Idiomaticity Continuum

Like other types of Construction Grammar (e.g., Langacker 1987; Goldberg 1995; Croft 2001; Bergen & Chang 2005), SBCG views linguistic patterns as form-meaning pairs (Goldberg 2006). Constructions fall onto a continuum based on degree of lexical fixity (Michaelis, to appear). In VDCs, the lexical fixity can be represented by the V1s' type frequency (including V1s' polysemous senses). According to Liu et al. 's (1998) VDC dictionary, the coarse-grained categories of the VDC pattern "V1 + *qi lai*" 'V1 rise come' display an imbalanced distribution of V1 frequencies in Table 14.

Self-motion	Caused-motion	Resultative	Aspectual	Discourse-Connective	Evidential
17	47	196	582	2	92

Table 14. V1 type frequencies, based on the VDC "V1 + *qi lai*"

These frequencies can be analyzed into three groups. The first group is Aspectual function, namely the continuous use of *qi lai*. This group has the highest V1 productivity. If we take the Resultative-Completive from the Resultative-VDCs into account, Aspectual *qi* can combine with 696 kinds of V1s. This shows that the V1 for Aspectual *qi* is almost an open class. Resultative, Self and Caused Motion fall into the second group, in which V1 is related to one or several verb classes, such as motion, spread, and connection, and therefore compared to the Aspectual class they are intermediately productive. The last group is Discourse-Connective VDC, e.g., *kan qilai* lit ('look rise') 'according to my view.' This type restricts the V1 slot to just two lexemes, and features other idiomatic restrictions that are symptomatic of a high degree of grammaticalization: the subject argument (construed as referring to the speaker or writer) is necessarily null instantiated (Sec.3.1.5). As Table 14 suggests, some VDC constructions are open patterns and some have the status of formulaic multi-word expressions; the SBCG framework provides an analysis of all such patterns, wherever they fall on the gradient from fixed formula to fully productive rule.

3.3.2 A Type Hierarchy of VDCs and DVs

SBCG not only describes the semantics of phrasal patterns on a scale of lexical fixity, but also links their syntactic similarity within a non-transformational "construction inheritance" network (Michaelis, to appear). This means that a highly lexicalized idiom like an MWE is usually licensed by a general construction with certain constraints. In other words, MWEs may have idiomatic interpretations but are syntactically regular, and assembled by the same constructions that assemble non-idiomatic words. Take the idiomatic VDC *fang xia jiazi* lit ('put descend frame') 'be humble.' The manner verb, DV, and object noun are all necessary parts of this MWE.

31. lingdao-ganbu yao <u>fang</u> <u>xia</u> jia-zi
    leader-leader need put descend frame-SUF
    "Government leaders need to be humble." (cf. Google Search)

Although the MWE/VDC in Sentence 31 is based on the metaphor HIERARCHY IS A FRAME, its constituency as a VP (i.e., a serial verb construction) is no different from that of a more general caused-Motion VDC "V *xia* NP" with variables in the manner verb and object, e.g., *fangxia shubao* 'put down a bag.' Therefore, a type hierarchy can be built to include a serial verb phrase "V DV NP," as well as the MWE *fang xia jiazi;* the latter is a subtype of the former. Combinatory properties of lexemes defined by constructions are mirrored in lexical entries. For example, both *fangxia shubao* and *fang xia jiazi* allow the object-control pattern (i.e., they have the entailments *fang* NP 'put NP' and NP *xia le* 'NP descends'), as discussed in Sec.3.2.2; both allow the BA construction (i.e., *ba* NP *fang xia* 'put NP down').

In the construction hierarchy of VDCs (see in Figure 7), the top construction node is Serial Verb Construct type (*sv-cxt*), which is generally defined as a complex predicate with multiple verbs that share one or more arguments (Müller & Lipenkova 2009). Although in Chinese linguistics, a debate exists covering the lexical category of DVs and the constituency of VDCs, I argued in my Appendix 1 that according to several diagnostic standards in Construction Morphology (Booij 2009), it is more reasonable to model VDCs as a kind of SVC than compound verb.

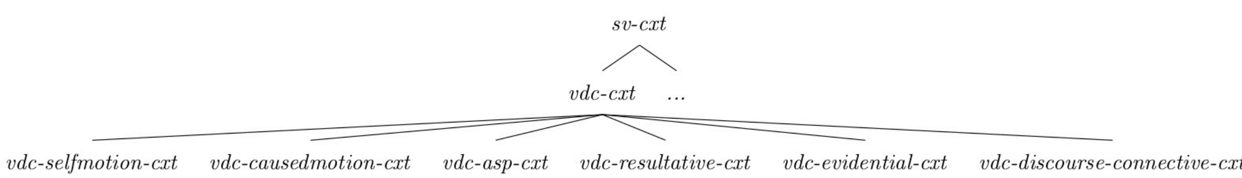

Figure 7. A Type Hierarchy of VDCs

3.3.3 Representing VDCs in the SBCG formalism

Since SBCG is a grammatical theory that associates the syntax and semantics of constructions, the challenge of modeling VDCs in SBCG is to represent morphosyntactic constraints in order to differentiate construction meanings. Based on the findings discussed in Section 3.1 and the common approach taken by different schools of syntax on VDC-related phenomena, coarse-grained VDCs can be identified by argument-sharing patterns, which are captured by the SBCG formalism. In this section, I will propose SBCG treatments of VDCs with several typical argument-sharing patterns: subject-control, object-control, and raising.

Before exploring possible SBCG representations for VDCs, it is necessary to know how serial verb constructions (especially so-called "V-V compounds" or "resultative verb compounds") are traditionally approached in syntactic theory. Both transformational syntax and declarative theories like Head-drive Phrase Grammar (HPSG) have approached Chinese resultative V-V constructions, but few accounts have attempted to distinguish various kinds of VDCs. In transformational syntax, Huang (1992) argues that some "V-V compounds" can be analyzed as a control pattern using a transformation of *de*-resultative clause. Without claiming transformations, Li (1990) applies some assumptions of Government-Binding Theory to model "V-V compounds" as complex verbs with argument-sharing from the argument structures of both V1s and V2s. In declarative models of syntax, especially HPSG, the notion of argument-sharing is also applied. Müller and Lipenkova (2009) use object-sharing as a standard to classify different serial verb constructions. Song et al. (2015) creates two constraints to describe resultative constructions: the number of arguments shared by V1 and V2, and the grammatical function of the shared argument (e.g., subject/object orientation). Although these researchers differ in their assumptions and constructions of interest, they agree on the analysis of argument-sharing. The SBCG analysis I am proposing aligns with Li (1990) and Song et al. (2015). I show that the argument-sharing information is preserved in the syntactic representations of V1s and V2s.

Given that the main constraint on VDCs is argument-sharing, the next step is to introduce the basics of the SBCG formalisms. As in HPSG, each sign, whether a lexeme or phrase, is represented as an attribute-value-matrix (AVM). Within each AVM, there are feature-value pairs known as feature structures (FSs), including phonological forms ("FORM" or "PHON"), syntax ("SYN"), semantics ("SEM"), argument structures ("ARG-ST"), and so on. In the AVMs for verbs or VPs, "SYN" contains FSs like external arguments ("XARG"), used to specify the argument of an argument-taking expression that is visible from outside its local domain, and valence ("VAL"), a list-valued feature that represents the arguments that a predicator must combine with to form a phrase. "ARG-ST" includes all possible arguments of verbs, including those that are extracted and omitted. "SEM" contains the frame semantics and sometimes the semantic scopes of quantifiers. The indices of frame elements that can be **coindexed** in the SYN. Among all FSs, "SYN" and "SEM" are the most relevant to modeling the argument-sharing patterns of VDCs.

The first VDC type is subject control, in which means both V1 and DV shared the same subject, as in the Self-Motion example and Figure 8.

32. A. ta <u>zou</u> le <u>jin</u> wu

s/he walk ASP enter house

"S/he walked into the room"

B. ta zou le "S/he walked."

C. ta jin le wu “S/he entered the house.”

As the derivation tree shows in Figure 8, the VDC (*vdc-selfmotion-cxt*) is the right head VP daughter for the *subject-pred-cxt* (marked as “H”). The VDC *vdc-selfmotion-cxt* has three daughter nodes: V1 (*vdc-manner-lxm*), V2 (*vdc-direction-lxm*), and the object (*noun-lxm*). The VAL features for both V1 and V2 completely reflect their argument structures in the VDC. These structures are the same as those of Sentences (32B) and (32C). The referent of the subject pronoun *ta* is construed as the first argument of both V1 and V2. Also, the SYN feature of *vdc-cxt* is the same as that of V1.[16]

[16] There are divergent views concerning the headedness of Chinese serial-verb constructions: both V1 (Li 1990) and V2 (Cheng & Huang 1994) have been identified as the head of the construction, and headless analyses have been proposed as well (Li 2009). This study assumes a headless analysis of VDCs. In VDCs, although V2/DV is typically viewed as the head (because V1 can be treated as the manner/means/precondition of the central directional predicate), the DV’s syntactic properties are often dependent on V1. For example. in *wo zou jin fang jian* (lit ‘I walk enter room’), ‘I walk into the room,’ the first DV *jin* ‘enter’ qualifies as a subject-control verb, while in *wo ba jiaju ban jin fangjian* lit (‘I move furniture enter room’), ‘I move the furniture into the room,’ the very same DV is an object-control verb. This flexibility suggests that V1 is the head, as it controls the construal of its complement. Because both V1 and V2 have head properties, the VDC construction is plausibly treated an exocentric/headless construction, as per Li 2009. The fact that it is the argument-structure of V1 that determines the argument structure of the whole complex suggests the centrality of V1, but there is no marking of the head (i.e., “H”) of the *vdc-cxt*.

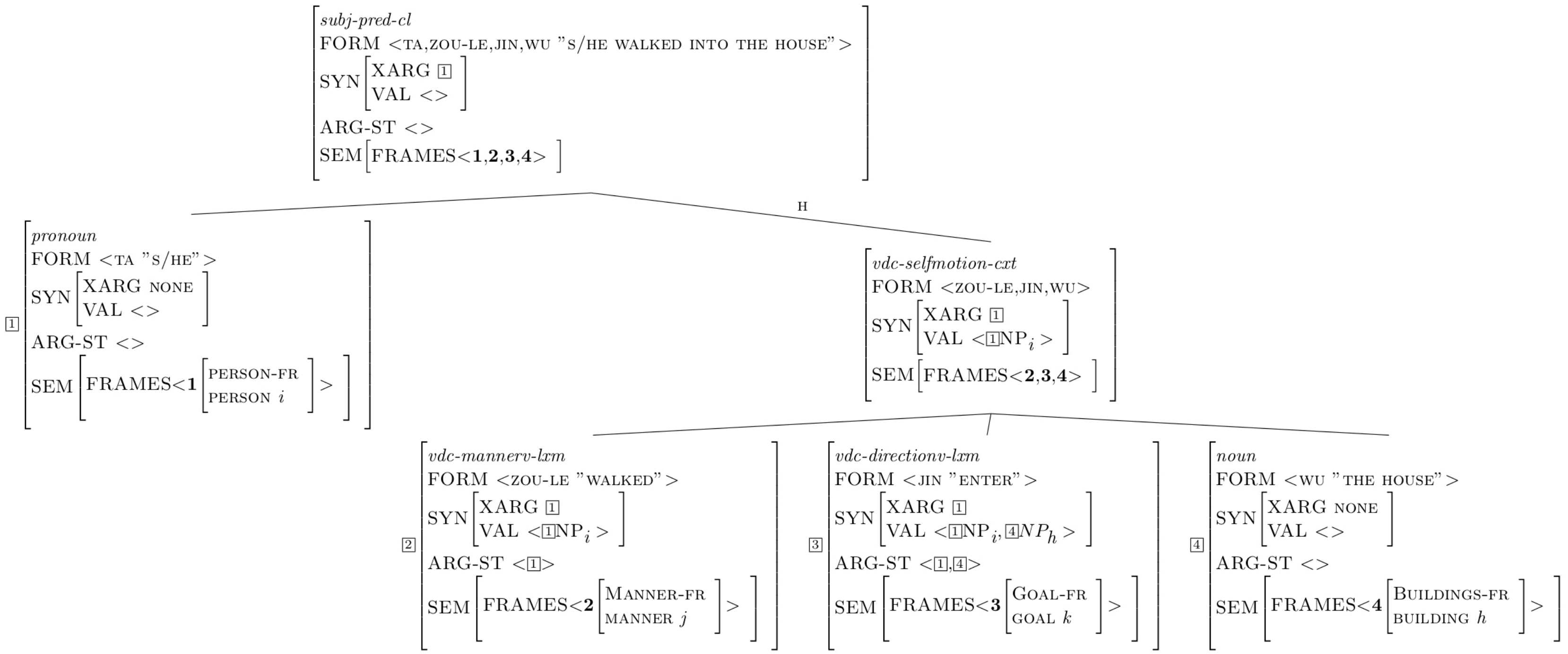
subj-pred-cl
FORM <TA,ZOU-LE,JIN,WU "S/HE WALKED INTO THE HOUSE">
SYN [XARG 1, VAL <>]
ARG-ST <>
SEM [FRAMES<1,2,3,4>]
H
1 pronoun
FORM <TA "S/HE">
SYN [XARG NONE, VAL <>]
ARG-ST <>
SEM [FRAMES<1 [PERSON-FR, PERSON i]>]
vdc-selfmotion-cxt
FORM <ZOU-LE,JIN,WU>
SYN [XARG 1, VAL <1NP i >]
SEM [FRAMES<2,3,4>]
2 vdc-mannerv-lxm
FORM <ZOU-LE "WALKED">
SYN [XARG 1, VAL <1NP i >]
ARG-ST <1>
SEM [FRAMES<2 [MANNER-FR, MANNER j]>]
3 vdc-directionv-lxm
FORM <JIN "ENTER">
SYN [XARG 1, VAL <1NP i, 4NP h >]
ARG-ST <1,4>
SEM [FRAMES<3 [GOAL-FR, GOAL k]>]
4 noun
FORM <WU "THE HOUSE">
SYN [XARG NONE, VAL <>]
ARG-ST <>
SEM [FRAMES<4 [BUILDINGS-FR, BUILDING h]>]

Figure 8. A Derivation of a Self-Motion VDC

The object-control pattern is usually found in Caused-Motion VDCs, in which V1 and DV share the same object. In addition to the challenge of representing the argument-sharing pattern, this analysis has captured the BA-construction. This construction frequently appears in this type of VDCs and augments the interpretation of object manipulation, which means "an agent argument causes the theme to be involved in a certain change" (Cheng 2018: 20).

33. A. wo ba hai-zi han-le jin wu

I BA kid-SUF call-ASP enter house

"I called the kid to move into the house"

B. wo han le haizi "I called the kid."

C. Haizi jin le qu "The kid entered the house"

In Sentence 33, the BA construction indicates causation of change of the location. The diagnostic sentences show that the NP *haizi* 'the kid' is both the patient of V1 and agent of V2.

Despite the debate about the proper syntactic analysis of BA-constructions among different schools of syntax, the SBCG analysis I offer treats the BA construction as a subject control structure: BA is a subject-control verb analogous to the English verb *cause*; it takes NP subject and object, and a VP complement as its arguments. According to this analysis, which is supported in the Chinese syntacticians' work (mostly in HPSG) and some constituency tests (Sybesma 1999; Ding 2000; Gao 2000; Wang et al. 2009; Fong 2015), (33) means something like ‘I caused the kid to be called into the house.’ With respect to the shared object NP, the SBCG analysis does not presume any syntactic gap. The diagram of the VDC in (33) is given below. Similar to the analysis of Self-Motion VDCs, conindexation of arguments in the valence sets of V1 and V2, indicate an object-control pattern.

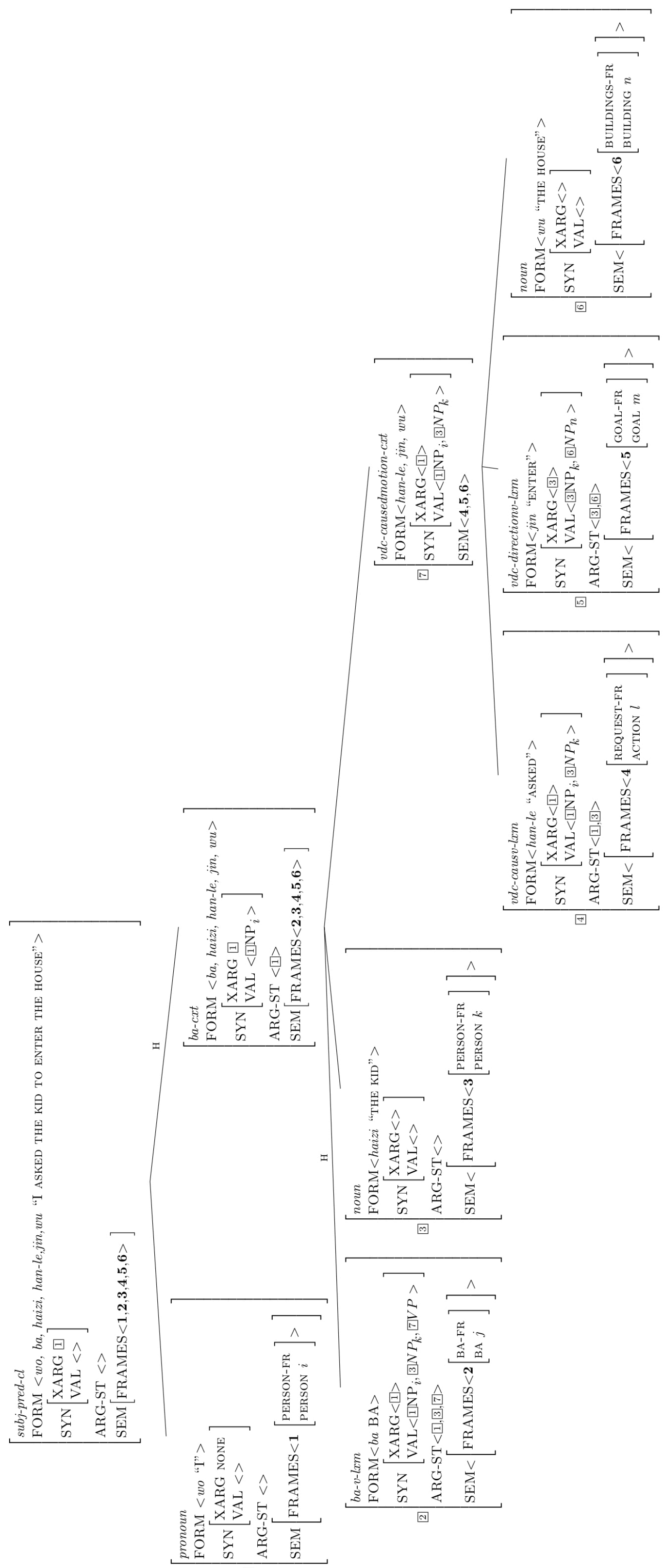
subj-pred-cl
FORM <wo, ba, haizi, han-le,jin,wu "I ASKED THE KID TO ENTER THE HOUSE">
SYN [XARG 1, VAL <>]
ARG-ST <>
SEM [FRAMES<1,2,3,4,5,6>]
H
pronoun
FORM <wo "I">
SYN [XARG NONE, VAL <>]
ARG-ST <>
SEM [FRAMES<1 [PERSON-FR, PERSON i]>]
ba-cxt
FORM <ba, haizi, han-le, jin, wu>
SYN [XARG 1, VAL <1NP i>]
ARG-ST <1>
SEM [FRAMES<2,3,4,5,6>]
H
ba-v-lxm
FORM<ba BA>
SYN [XARG<1>, VAL<1NP i, 3NP k, 7VP>]
ARG-ST<1,3,7>
SEM< [FRAMES<2 [BA-FR, BA j]>]
noun
FORM<haizi "THE KID">
SYN [XARG<>, VAL<>]
ARG-ST<>
SEM< [FRAMES<3 [PERSON-FR, PERSON k]>]
vdc-causedmotion-cxt
FORM<han-le, jin, wu>
SYN [XARG<1>, VAL<1NP i, 3NP k>]
SEM<4,5,6>
vdc-causv-lxm
FORM<han-le "ASKED">
SYN [XARG<1>, VAL<1NP i, 3NP k>]
ARG-ST<1,3>
SEM< [FRAMES<4 [REQUEST-FR, ACTION l]>]
vdc-directionv-lxm
FORM<jin "ENTER">
SYN [XARG<3>, VAL<3NP k, 6NP n>]
ARG-ST<3,6>
SEM< [FRAMES<5 [GOAL-FR, GOAL m]>]
noun
FORM<wu "THE HOUSE">
SYN [XARG<>, VAL<>]
SEM< [FRAMES<6 [BUILDINGS-FR, BUILDING n]>]

Figure 9. A Derivation of a Caused-Motion VDC in a BA construction

The raising pattern is commonly seen in Aspectual, Discourse-Connective, Evidential VDCs, and some Resultative and Self-Motion-Shifted VDCs. In this pattern, **I treat DVs as subject-raising verbs, on the assumption that, like English subject-raising verbs, they feature a mismatch between syntactic valence and semantic argument structure: they have just one semantic argument, the situation or state of affairs denoted by V1, but usually two syntactic valence members: the external argument (subject) of V1 (which is typically a topical entity) and V1 itself.** The argument-sharing pattern seen in raising is similar to subject control: the first argument of a state of affairs predicator (like *wuli renao* 'The house is cheery' below) is also the argument of a verb that describes that state of affairs as evident, salient, etc. As in HPSG, the difference between subject control and subject raising comes down to a semantic distinction: a subject-control verb (like *try*) assigns a semantic role both to its subject and the subject of its verbal complement, while a subject-raising verb (like *begin*) does not assign a semantic role to its subject, but only to its verbal complement (V2), which represents a state of affairs. The raising example below is an Aspectual VDC, which takes the state of affairs denoted by the NP subject and V1, *wuli renao* 'The house is cherry':

34. A. wu-li renao qi lai

house-inside cheery <u>rise</u> <u>come</u>

"The house started getting cheerier." (cf. Liu et al. 1998)

B. wu li renao le "The house was cheery."

C.*wuli qi lai le "The house started."

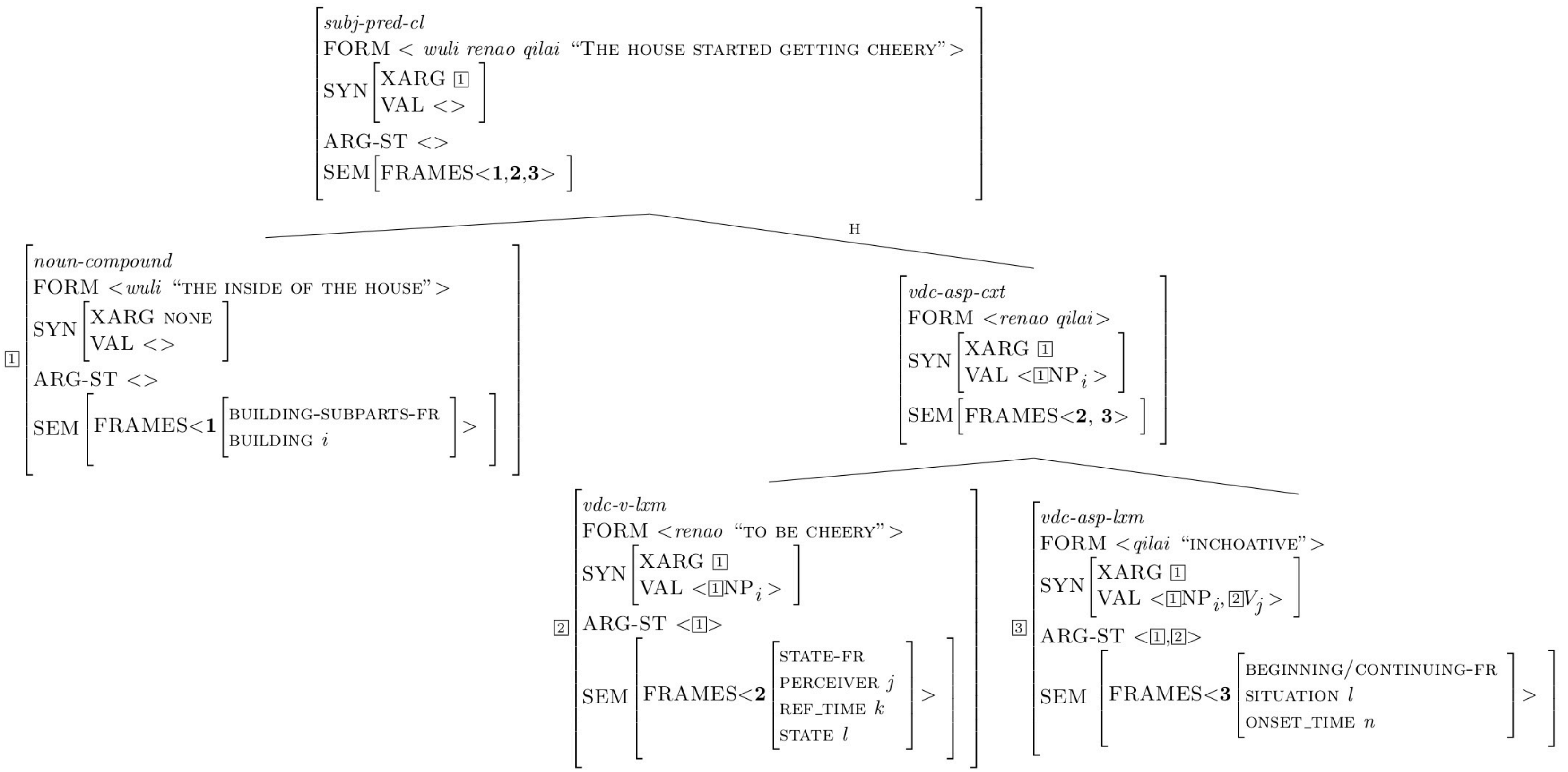
subj-pred-cl
FORM < wuli renao qilai "THE HOUSE STARTED GETTING CHEERY">
SYN [XARG 1, VAL <>]
ARG-ST <>
SEM [FRAMES<1,2,3>]
H
noun-compound
FORM <wuli "THE INSIDE OF THE HOUSE">
SYN [XARG NONE, VAL <>]
ARG-ST <>
SEM [FRAMES<1 [BUILDING-SUBPARTS-FR, BUILDING i]>]
vdc-asp-cxt
FORM <renao qilai>
SYN [XARG 1, VAL <1NP_i>]
SEM [FRAMES<2, 3>]
vdc-v-lxm
FORM <renao "TO BE CHEERY">
SYN [XARG 1, VAL <1NP_i>]
ARG-ST <1>
SEM [FRAMES<2 [STATE-FR, PERCEIVER j, REF_TIME k, STATE l]>]
vdc-asp-lxm
FORM <qilai "INCHOATIVE">
SYN [XARG 1, VAL <1NP_i, 2V_j>]
ARG-ST <1,2>
SEM [FRAMES<3 [BEGINNING/CONTINUING-FR, SITUATION l, ONSET_TIME n]>]

Figure 10. A Derivation of an Aspectual VDC

In an Aspectual VDC like Sentence 34A, the DV describes a property of the state of affairs denoted by V1. This function is indicated by two aspects of the formalism: (1) syntactically, V2's valence set contains both the subject and V1, and (2) semantically, the aspectual frame of the compounding DVs *qi lai* takes the State element of V1 *renao* 'be cheery' as its sole semantic role.

Many **Discourse-Connective** VDCs developed from Aspectual VDCs. For example, the DV *qi* in *cong* 'from' + Topic + Speech V1 + *qi* or Speech V1 + *qi* + Topic 'speaking of the Topic' encodes the inchoative state of the speaking. When an Interlocutor starts speaking about a Topic, the speaking action introduces new discourse. Therefore, I analyze Discourse-Connective VDCs, as in Figure 9, as a variation of Aspectual VDCs in which the "speaker" argument is unrealized, and contextually construed as the speaker or writer of the sentence.

35. cong xijinping chi bao-zi <u>shuo</u> <u>qi</u>

from Xijinping eat bun-SUF speak rise

"Speaking of the fact that President Xi Jinping ate buns." (cf. Google Search)

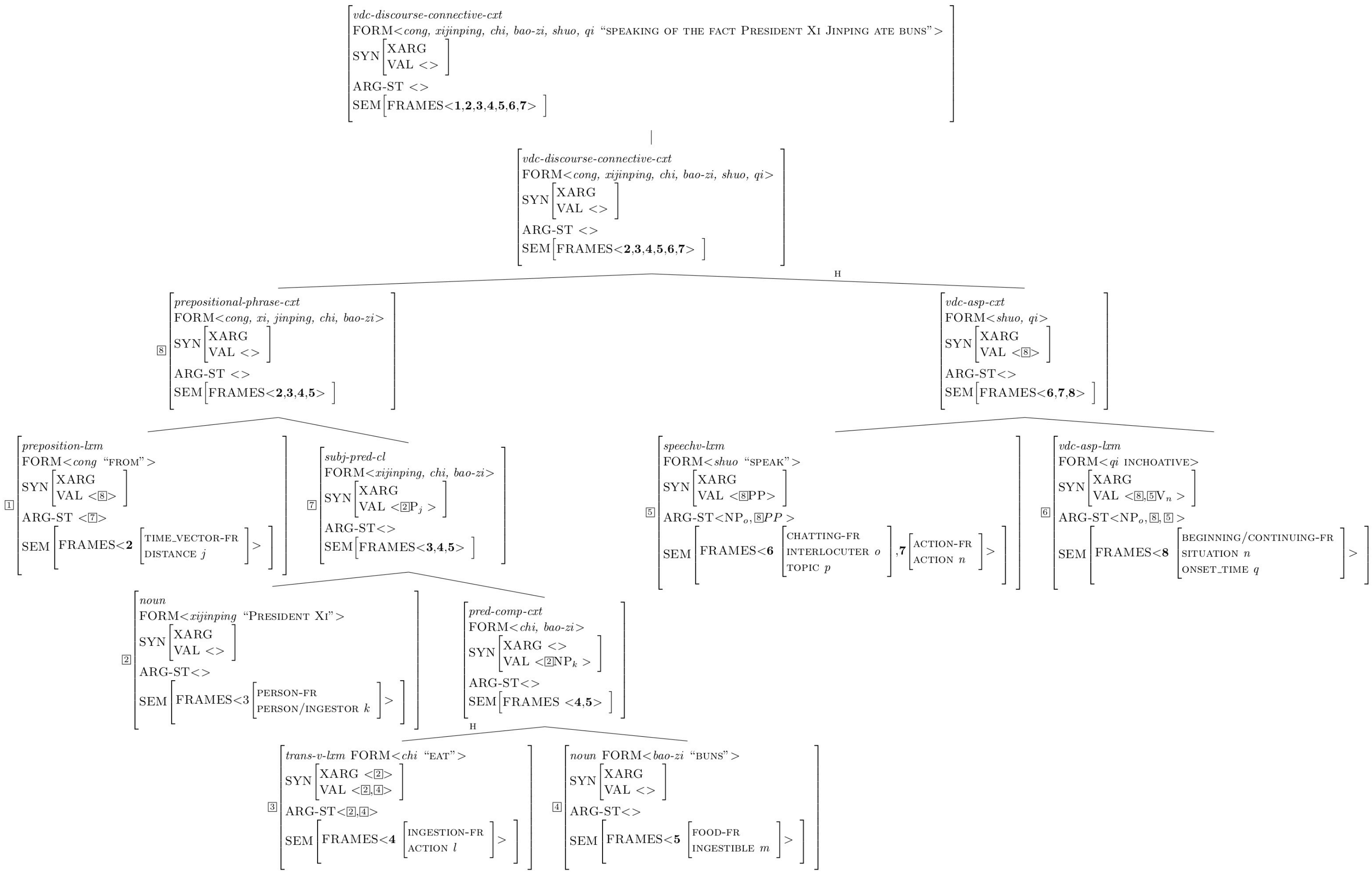
vdc-discourse-connective-cxt
FORM<cong, xijinping, chi, bao-zi, shuo, qi "SPEAKING OF THE FACT PRESIDENT XI JINPING ATE BUNS">
SYN XARG VAL <>
ARG-ST <>
SEM FRAMES<1,2,3,4,5,6,7>
vdc-discourse-connective-cxt
FORM<cong, xijinping, chi, bao-zi, shuo, qi>
SYN XARG VAL <>
ARG-ST <>
SEM FRAMES<2,3,4,5,6,7>
H
8 prepositional-phrase-cxt
FORM<cong, xi, jinping, chi, bao-zi>
SYN XARG VAL <>
ARG-ST <>
SEM FRAMES<2,3,4,5>
vdc-asp-cxt
FORM<shuo, qi>
SYN XARG VAL <8>
ARG-ST<>
SEM FRAMES<6,7,8>
1 preposition-lxm
FORM<cong "FROM">
SYN XARG VAL <8>
ARG-ST <7>
SEM FRAMES<2 TIME_VECTOR-FR DISTANCE j >
7 subj-pred-cl
FORM<xijinping, chi, bao-zi>
SYN XARG VAL <2Pj >
ARG-ST<>
SEM FRAMES<3,4,5>
5 speechv-lxm
FORM<shuo "SPEAK">
SYN XARG VAL <8PP>
ARG-ST<NPo,8PP >
SEM FRAMES<6 CHATTING-FR INTERLOCUTER o TOPIC p ,7 ACTION-FR ACTION n >
6 vdc-asp-lxm
FORM<qi INCHOATIVE>
SYN XARG VAL <8,5Vn >
ARG-ST<NPo,8,5 >
SEM FRAMES<8 BEGINNING/CONTINUING-FR SITUATION n ONSET_TIME q >
2 noun
FORM<xijinping "PRESIDENT XI">
SYN XARG VAL <>
ARG-ST<>
SEM FRAMES<3 PERSON-FR PERSON/INGESTOR k >
pred-comp-cxt
FORM<chi, bao-zi>
SYN XARG <> VAL <2NPk >
ARG-ST<>
SEM FRAMES <4,5>
H
3 trans-v-lxm FORM<chi "EAT">
SYN XARG <2> VAL <2,4>
ARG-ST<2,4>
SEM FRAMES<4 INGESTION-FR ACTION l >
4 noun FORM<bao-zi "BUNS">
SYN XARG VAL <>
ARG-ST<>
SEM FRAMES<5 FOOD-FR INGESTIBLE m >

Figure 11. A Derivation of a Discourse-Connective VDC

As indicated in Figure 11, the Discourse-Connective VDC has two daughters: a *PP-cxt* and an *vdc-asp-cxt*, the later of which acts as the head. In the *PP-cxt*, the preposition *cong* introduces the Topic clause. The element of the Interlocutor is implicit. No empty category is introduced in the derivation, but this missing element is listed in the Chatting frame of the V1. The *vdc-asp-cxt* is similar to the representation in Figure 8, in which the V2 takes the arguments of the implicit Interlocutor subject and the state of affairs, co-expressed by the PP *cong xijinping chi baozi* 'from the fact that President Xi ate buns' and the speech V1 *shuo* 'speak.'

Our last case of raising is the **Evidential** VDC, which encodes the process a perceiver's comments on certain phenomenon. Similar to the Discourse-Connective VDC, the Evidential VDC is a variation of Aspectual VDC because the DV usually means the inception of the state of the affairs.

36. ta ting <u>shang</u> <u>qu</u> bu zhao diao

she listen ascend go NEG get tune

"She sounds like she is going off key." (lit "She started to sound off key.") (cf. Liu et al. 1998)

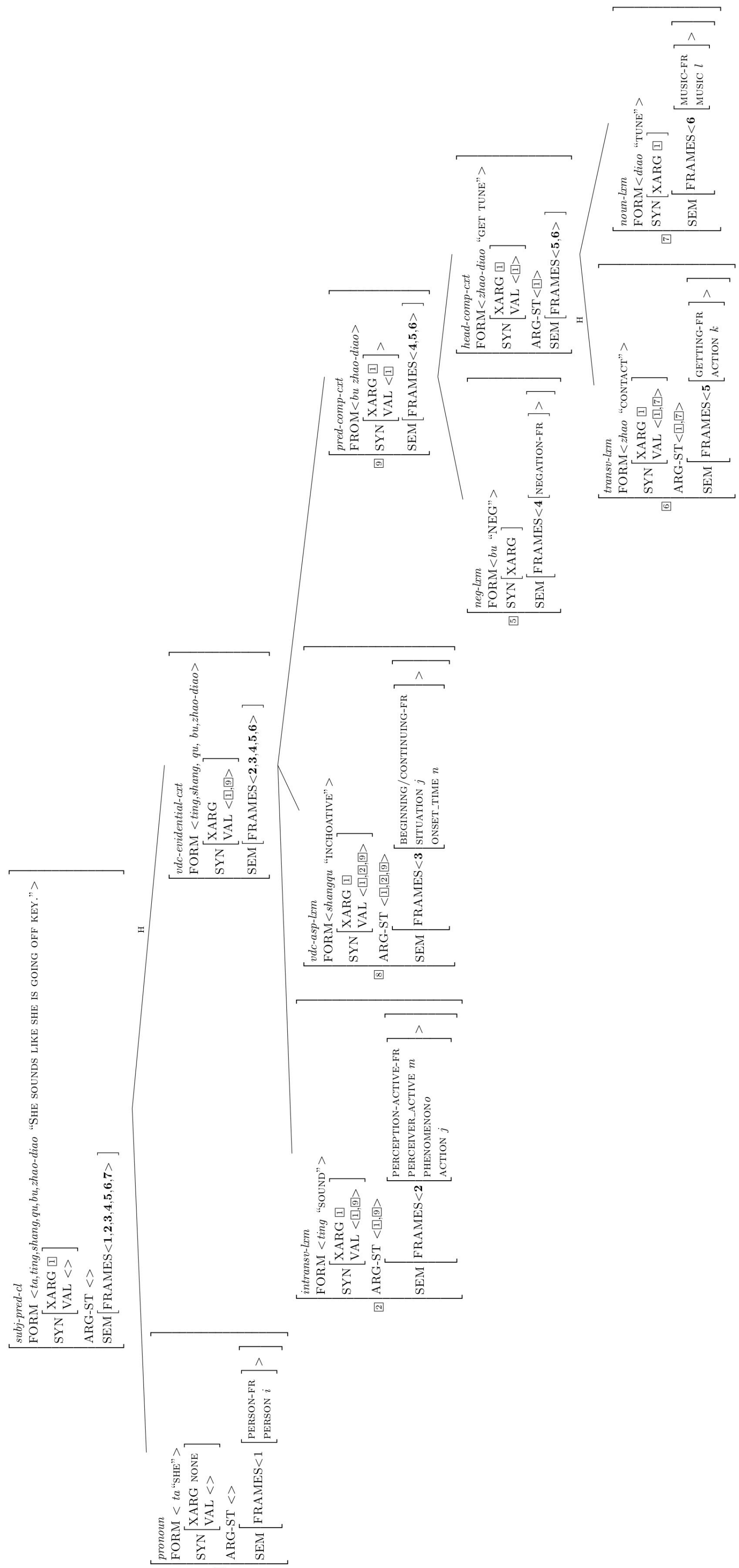
subj-pred-cl
FORM <ta,ting,shang,qu,bu,zhao-diao "She sounds like she is going off key.">
SYN [XARG 1, VAL <>]
ARG-ST <>
SEM [FRAMES<1,2,3,4,5,6,7>]
H
pronoun
FORM < ta "SHE">
SYN [XARG NONE, VAL <>]
ARG-ST <>
SEM [FRAMES<1 [PERSON-FR, PERSON i] >]
vdc-evidential-cxt
FORM <ting,shang, qu, bu,zhao-diao>
SYN [XARG, VAL <1,9>]
SEM [FRAMES<2,3,4,5,6>]
intransv-lxm
FORM <ting "SOUND">
SYN [XARG 1, VAL <1,9>]
ARG-ST <1,9>
SEM [FRAMES<2 [PERCEPTION-ACTIVE-FR, PERCEIVER_ACTIVE m, PHENOMENON o, ACTION j] >]
vdc-asp-lxm
FORM <shangqu "INCHOATIVE">
SYN [XARG 1, VAL <1,2,9>]
ARG-ST <1,2,9>
SEM [FRAMES<3 [BEGINNING/CONTINUING-FR, SITUATION j, ONSET_TIME n] >]
pred-comp-cxt
FROM<bu zhao-diao>
SYN [XARG 1, VAL <1]
SEM [FRAMES<4,5,6>]
neg-lxm
FORM<bu "NEG">
SYN [XARG]
SEM [FRAMES<4 [NEGATION-FR] >]
head-comp-cxt
FORM<zhao-diao "GET TUNE">
SYN [XARG 1, VAL <1>]
ARG-ST<1>
SEM [FRAMES<5,6>]
H
transv-lxm
FORM<zhao "CONTACT">
SYN [XARG 1, VAL <1,7>]
ARG-ST<1,7>
SEM [FRAMES<5 [GETTING-FR, ACTION k] >]
noun-lxm
FORM<diao "TUNE">
SYN [XARG 1]
SEM [FRAMES<6 [MUSIC-FR, MUSIC l] >]

Figure 12. A Derivation of an Evidential VDC

Example 36 is a speaker's negative comment based on their perception of the agent's singing. In Figure 12, I propose a tripartite branching structure for the *vdc-evidential-cxt*. Its three daughter nodes are: the perceptual V1 *ting* 'sound,' the compounding Aspectual DVs *shang qu* 'ascend go,' which indicates inception, and a VP *bu zhao-diao* lit ('not get tune') 'be off key.' **Both V1 and DV are raising verbs that take a state of affairs as their argument.** The V1 *ting* 'sound' takes two valance arguments: the subject (*ta* 'she') and the VP *bu zhaodiao* 'be off key.' The Perceiver argument of V1 is pragmatically inferred, so there is no explicit syntactic node for it. As with my solution to the valence of the Aspectual VDCs, the inchoative DV *shangqu* takes as arguments the subject and the state of affairs jointly expressed by V1 *ting* 'sound' and its complement VP *bu zhaodiao* 'be off key.' As in our previous treatment of Aspectual DVs in Figure 10, *shangqu* takes the V1 *ting* 'sound' as its only semantic argument.

In summary, SBCG as both a theory and a formalism provides four advantages for analyzing different kinds of VDCs: first, VDCs are differentiated according to the degree of lexical fixity; second, distinctive properties of argument-sharing are captured through coindexation of valence members of V1 and V2, respectively; third, SBCG captures the distinction between control and raising in its frame-semantic representations (in particular, these representations show whether V2, the DV, takes an individual or a state of affairs as its first or only argument); and fourth, SBCG models the interaction between the VDC and the BA-construction, where the latter is represented as a kind of subject-control verb. Leaving aside the issues of theoretical framework, in order to develop a VDC event classification system, we must ascertain the state of the art of VDC representations in Chinese NLP resources. This topic is discussed in the following section.

3.4 Current NLP Resources for Representing VDC Event Structures

The state of the art of VDC representations in Chinese NLP resources concerns syntactic and semantic labeling of VDCs/DVs. The syntactic labels come from the Penn Chinese Treebank (PCTB/CTB) (Xue et al. 2000), as well as THULAC. The semantic labels are from Chinese PropBank (CPB) from University of Colorado Boulder and Brandeis University (Xue & Palmer 2009) and Chinese WordNet (CWN) from National Taiwan University (Huang et al. 2010).

Two problems of syntactic labeling of VDCs arise in terms of VDC event representations: labeling inconsistency and the limitation of syntax.

First, both CTB and THULAC have two major ways of syntactically tagging VDCs, only one of which is effective. However, the ineffective way is what is more commonly used in practice. CTB (Xue et al. 2000: 221) uses the label "VRD" ("directional and resultative compound") for most VDCs, and each verb may or may not be labeled as "VV." The directional compound is similar to VDCs defined in this research, while resultative compounds are related to resultative constructions like *he zui* 'drink (oneself) drunk;' and "VRD" is the mother node of the V1 and DV. For example, the verb phrase *pa shang shan* lit ('climb ascend mountain') 'climb onto the mountain' can be represented as "[VP [VRD [VV *pa* 'climb'][VV *shang* 'ascend']] [NP-OBJ [NP [NN *shan* 'mountain']]]]." Similarly, THULAC adopts a "VV" analysis for VDCs from the "863" Tagset in China (Yang et al. 2006). One newly proposed label for VDCs is "VD" ("directional verb"), but this design is not used for all versions of THULAC.

Furthermore, in the practice of labeling in both resources, there are inconsistencies where a VDC is tagged as one or two "VV"s (as in the treatments of "Verb + *shang*" in (21A) and (21B). This inconsistency will cause inaccuracies in event-type-related parsing tasks like SRL (Palmer et al. 2010) and event detection, such as Automatic Content Extraction (Doddington et al. 2004), because one verb usually triggers one argument structure or event in the annotation of these resources. For example, for the same type of VDC “V + *shang* ‘ascend,’” there are two verbs (labeled in “_VV”) in (37A), but only one in (37B):

37. A.押_VV 上_VV 了_AS 警车

ya (VV) shang(VV) le jing-che

escort ascend ASP police-car

“be escorted into the police car.” (cf. CTB 6.0)

B.爬上_VV 了_AS 自由_NN 女神

pa-shang (VV) le ziyou nüshen

climb-ascend(VV) ASP freedom goddess

“climb onto the Statue of Liberty.” (cf. CTB 6.0)

Second, the syntactic labels in CTB and THULAC cannot sufficiently explain the event semantics of VDCs. Take two “idiomatic” VDCs for example: (1) *jie xia lai* lit (‘connect descend come’) ‘next,’ and (2) *kan qi lai* lit (‘look rise come’) ‘it looks like.’ Based on my VDC taxonomy, we know that the Discourse-Connective VDC in (1) and the Evidential VDC in (2) are two types of events. CTB treats both examples as “AD” (adverbial) (thus they do not have argument structures in CPB), and THULAC treats both as “v” (verb). Thus, neither of these resources provides the event semantics, such as participant roles or semantic frames, as these labels merely identify constituency information but lack semantic information.

Unlike syntactic labeling, the semantic labeling in CPB and CWN covers event semantics of VDCs, but has generalization problems. That is, CPB provides semantic annotation of the argument structures of predicates on the top of CTB, and its treatment is consistent for DVs: the simple DVs are labeled as part of "REL" (relation); the compounding DVs, such as *shang qu* 'ascend go' are labeled as an independent "REL." For example, in the CPB frame files, *shang* 'ascend' as a simple DV is not labeled on its own. VDCs including *shang*, such as *deng shang* 'get on' are labeled as "REL"s. Given that there is a lack of coverage of compounding DVs in the frame files, only the simple DVs are evaluated here. By treating VDCs as RELs/verbs, there is no degree of generalization of the event categories of VDCs. Take several VDCs with the pattern "V1+ *jin* 'enter'" for example: *ban jin* lit ('move enter') 'move into,' *chuang jin* lit ('rush enter') 'rush into,' *da jin* lit ('hit enter') 'play into,' and *gong jin* lit ('attack enter') 'attack into.' These Self-Motion VDCs share the same arguments: ARG0 (agent) and ARG1 (goal). CPB does not have an MWE layer of analysis to show how these VDCs are related to each other.

DVs are semantically labeled in CWN, a computational dictionary based on different lexical relationships, like synonymy and hyponymy. CWN offers fine-grained senses for each DV. DVs, whether as part of a VDC, or as independent verbs, are treated together in the same sense group.

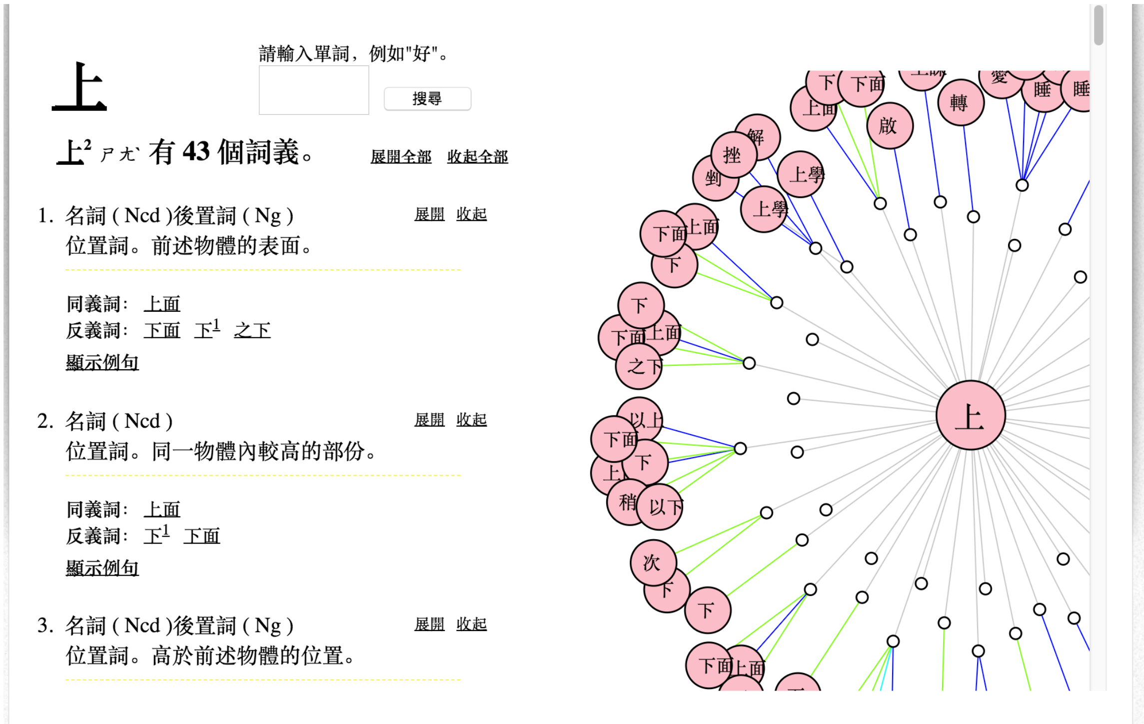

Figure 13. The Search Result of Polysemous 上 *shang* in CWN

For example, in the treatment of the DV *shang* ‘ascend’ in Figure 13, it has senses of aspect, attachment, and upward motion, and these senses work both for stand-alone DVs and DVs in VDCs. These senses are listed together with other polysemous meanings of *shang,* such as its use as a noun, adjective, or postposition. The main labeling problems of VDCs in CWN are (1) there is a lack of generalization above individual meanings of DVs, (2) lack of a level to generalize various “V1+V2/DV” patterns, and (3) insufficient DV sense information to classify participant roles in the events.

After examining different Chinese NLP resources, we see that VDCs are not represented well from either syntactic or semantic labeling. More generally, none of these resources has a construction layer encoding the form-meaning paired patterns. Problems like inconsistency and inadequate generalizations motivate the following study on classifying VDC categories.

3.5 Implementing, Training, Evaluating an Automatic VDC Classifier

This section reports the results of two classification studies: VDC detection and VDC event category classifications. The first study is binary VDC classification, namely detecting whether an expression is a VDC or not; the second study omits non-VDCs and classifies all remaining VDCs into coarse-grained and fine-grained VDC classes. This section starts with result reports and then continues with error analyses of VDC event classifications.

3.5.1 Study 1: Detecting VDCs

In Section 2.3, I introduced a semi-automatic method to retrieve the VDC tokens using an automatic search of DVs and the POS of both V1s and DVs as well as manual checking. Using this method, a total amount of 15,852 tokens was collected and annotated, out of which 14,370 tokens were VDCs. Non-VDCs may include independent uses of DVs, other POS including directional morphemes, and "V-V" forms which do not fit the semantics of VDCs, and so on, as illustrated by Sentences 38A-C.

38. A. ban ge shiji guo qu le

Half CL century pass go ASP

"Half a century has passed." (cf. CCL)

B. bu-guo, ta de gai-kuang feichang qingxi

NEG-pass, it ASSOC general-situation very clear

“However, its general situation is very clear.” (cf. CCL)

C. saluman, ni yuanyi xia lai ma

Saruman, you want descend come QP

“Saruman, do you want to come down?” (cf. CCL)

None of the examples above is a VDC. Example 38A has only the compounding DVs *guo qu* ‘pass go,’ but no preceding V1. Example 38B has the directional character *guo* ‘pass,’ but it is part of the adverb *buguo* ‘however’ to capture a between-sentence transition. Example 38C fits my search standard of “V1 + DV.” The problem is that V1 does not encode the manner or precondition of DV, therefore it does not fit the semantics of VDCs. The purpose of the VDC detection classifier is to sort out these possibilities so that an automatic VDC parsing system can collect clean data for other VDC-related tasks.

Before classifying coarse and fine grained categories of VDCs, it is important to verify that VDCs and non-VDCs are distinguishable from each other. Binary classifiers were trained using both biLSTM and SVM models. Table 15 summarizes the results of the VDC detection models.

Model#	Context	Setting	Precision	Recall	F-Score
1		EM0+Char	0.92	0.93	0.93
2		EM0+Seg	0.93	0.92	0.92
3		EM1+Char	0.93	0.94	0.94
4	Long	EM1+Seg	0.94	0.9	0.92
5		EM2+Char	0.93	0.94	0.94
6		EM2+Seg	0.94	0.91	0.92
7		SVM	0.95	0.77	0.84
8		EM0+Char	0.95	0.92	0.93
9		EM0+Seg	0.93	0.91	0.92
10		EM1+Char	0.95	0.94	0.94
11	Short	EM1+Seg	0.95	0.92	0.94
12		**EM2+Char**	**0.95**	**0.96**	**0.95**
13		EM2+Seg	0.95	0.92	0.93
14		SVM	0.89	0.7	0.8

Table 15. VDC Detection Results

Most conditions have more than 92% accuracy in their F-scores, which are above the percentage of VDCs in the data (90.6%). This indicates that VDC detection is a simple task. Models#3, 5, 10, 11, and 12 all have very close F-Scores. I ran these five models three times for each, and found that #12's average performance (F-Score=0.9367) is slightly better than the other four models. Model#12 uses both character inputs and word-and-character level embedding. The high recalls for different models also show that few false negative decisions were made. As seen in Table 15, there is not much difference across conditions of embedding, segmentation, and context length.

As a follow-up test to analyze the progress of the best models compared to the ones in my pilot study, the best of the current biLSTM and SVM models were tested on the pilot testing data (3,000 CTB instances).

Model	Precision	Recall	F-Score
LSTM-pilot: EM1+seg	0.92	1	0.96
LSTM-Model#12	0.97	0.95	0.96
SVM-pilot	0.86	0.96	0.93
SVM-Model#7	0.9	0.78	0.86

Table 16. A Comparison Between Current Best Models and Best Pilot Models for VDC Detection

Table 16 shows that there is little progress with respect to F-scores for both LSTMs and SVMs. While the SVM-Model#7 shows a 7% decrease. The results suggest that the pilot LSTM and SVM models are good enough to detect VDCs. As for SVM models, the use of the hashing vectors is a better choice than PCA on this task.

3.5.2 Study 2: Classifying VDCs with Coarse and Fine Taxonomies

This section gives the descriptive statistics and classifying results of two kinds of VDC event type classifiers: coarse-grained and fine-grained. After the development of the annotation guideline, I annotated the 14,370 VDC instances into 14 fine-grained categories. As stated in Section 3.1 and 3.2, fine-grained labels can be regrouped into coarse labels, so there is no need to manually label coarse-grained categories.

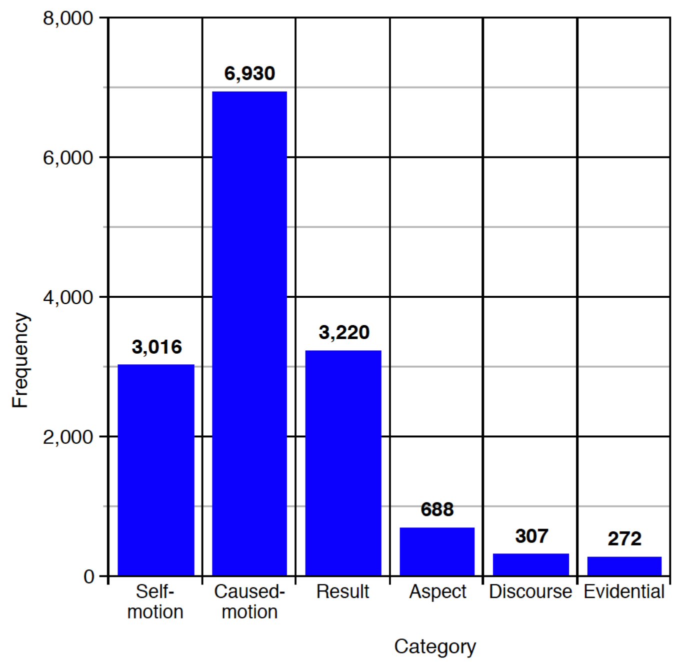

Figure 14. Distribution of Coarse-grained VDC Categories

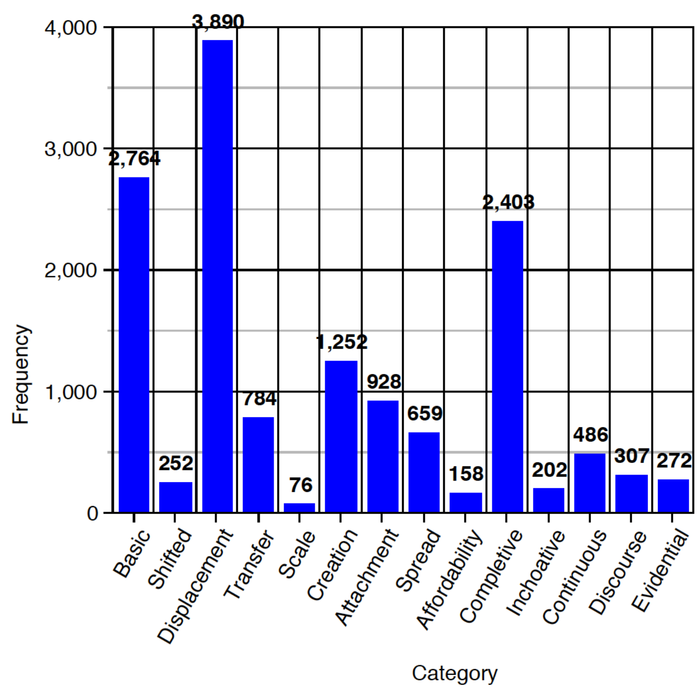

Figure 15. Distribution of Fine-grained VDC Categories

Figures 14 and 15 show the frequency distributions of categories in both versions of the VDC taxonomy. In the coarse-grained version, Caused-Motion VDC is the largest class. Self-Motion and Resultative are roughly the same. The more idiomatic classes, namely Discourse-Connective and Evidential VDCs, are the least frequent. The imbalanced distribution continues in the fine-grained version, with the Self-Motion-Basic, Caused-Motion-Displacement, and Resultative-Completive as the three most frequent classes, and Caused-Motion-Scale, Result-Affordability, and Aspect-Inchoative as the three least frequent ones. It is expected that the fine-grained version is more difficult to learn than the coarse-grained one because it has more classes and a greater degree of between-class imbalance. The results of the different models are given in Table 17:

Model#	Context	Model#	Precision	Recall	F-Score
15		EM0+Char	0.7	0.7	0.7
16		EM0+Seg	0.65	0.65	0.65
17		EM1+Char	0.69	0.67	0.67
18	Long	EM1+Seg	0.65	0.64	0.65
19		EM2+Char	0.69	0.69	0.69
20		EM2+Seg	0.64	0.64	0.64
21		SVM	0.65	0.58	0.62
22		EM0+Char	0.73	0.73	0.73
23		EM0+Seg	0.71	0.7	0.71
24		**EM1+Char**	**0.75**	**0.75**	**0.75**
25	Short	EM1+Seg	0.72	0.72	0.72
26		EM2+Char	0.75	0.74	0.75
27		EM2+Seg	0.72	0.71	0.71
28		SVM	0.68	0.59	0.62

Table 17. Coarse-grained VDC Categories Classification Results

Table 17 shows results of the coarse-grained category classification. There are three findings: First, models with character-level segmentations (Model#15, 17, 19, 21, 23, and 25) perform better than their word-level counterparts while controlling for other factors (Model#16, 18, 20, 22, and 24). Second, the condition of contextual length influences classification. In neural models, short context conditions (i.e., clause containing target VDCs) have better F-scores than the longer ones (the full context retrieved from corpora). Third, the performance of SVMs falls behind biLSTMs, which is truer in the short condition than long condition.

The F-Scores of Model #24 and 26 shared similar F-Scores. After three runs for each model, the average performance of #26 (F-Score=0.767) is higher than that of #24 (F-Score=0.727), therefore #24 is the best current model. To further evaluate the performance of current models, Table 18 compares the new and pilot models on the pilot testing data.

Model	Precision	Recall	F-score
LSTM-pilot: EM2+char	0.54	0.53	0.53
LSTM-Model#24	0.77	0.77	0.77
SVM-pilot	0.08	0.44	0.28
SVM-Model#28	0.68	0.58	0.62

Table 18. A Comparison Between Current Best Models and Best Pilot Models for Coarse-grained VDC Classification

There are significant improvements to both biLSTM and SVMs over pilot studies, especially on the new SVM model. We can see 24% and 34% increases on biLSTM and SVM, respectively. Next, the performances on the fine-grained VDC classification is reported:

Model#	Context	Model	Precision	Recall	F-Score
29		EM0+Char	0.57	0.56	0.56
30		EM0+Seg	0.51	0.51	0.51
31		EM1+Char	0.58	0.56	0.56
32	Long	EM1+Seg	0.53	0.53	0.53
33		EM2+Char	0.59	0.59	0.59
34		EM2+Seg	0.54	0.53	0.53
35		SVM	0.55	0.47	0.48
36		**EM0+Char**	**0.67**	**0.67**	**0.67**
37		EM0+Seg	0.63	0.63	0.63
38		EM1+Char	0.65	0.64	0.64
39	Short	EM1+Seg	0.62	0.62	0.62
40		EM2+Char	0.65	0.64	0.64
41		EM2+Seg	0.59	0.59	0.59
42		SVM	0.58	0.48	0.48

Table 19. Fine-grained VDC Categories Classification Results

Compared to the general performance in coarse event type classification, fine-grained classification has lower scores: the prediction performance ranges from 0.47 to 0.67. Similarly to the findings in Table 17, we can observe the benefit of using character segmentation and short context. Although neural models benefit from using short inputs, the input length effect in SVMs is not obvious. The best performance comes from the character-level neural model with randomized embedding and short inputs (Model#36), whose average performance after five runs (F-Score=0.640) is slightly higher than that of #41 (F-Score=0.632). Another common feature between the two tables is that the role of pre-trained embeddings fails to provide improvements in either the coarse or fine grained classification task. Randomized embedding layers can achieve equal or better performance than the pre-trained embedding layers.

Both new biLSTM and SVM models also show improvement over the pilot models on the pilot testing data in Table 20. The F-scores of both models increased about 40%, which indicates the effectiveness of the revisions to these models after the prospectus.

Model	Precision	Recall	F-score
LSTM-pilot: EM2+char	0.27	0.28	0.27
LSTM-Model#36	0.69	0.68	0.68
SVM-pilot	0.001	0.06	0.03
SVM-short 42	0.53	0.46	0.48

Table 20. A Comparison Between Current Best Models and Best Pilot Models for Fine-grained VDC Classification

To further investigate the classification performance, I represent the prediction of one coarse-grained and one fine-grained neural model in confusion matrices. Figure 16 shows the learning of the six coarse-grained VDC classes in Model#24. The numbers in the diagonal indicate that the model made the correct predictions. The grids adjacent to the diagonal are between-class errors, e.g., the numbers 45 and 59 near 221 show mutual errors between Class #1 and #2. There are different degrees of confusion between Class#1 (Self-Motion) and #2 (Caused-Motion), between Class#2 (Caused-Motion) and #3 (Resultative), and between Class#3 (Resultative) and #4 (Aspect). These confusion results tell us that the boundaries between different categories are not always clear. Although this trend mirrors native speakers' cognition of the grammaticalization of different VDCs, namely a semantic continuum, there is still work to be done to clarify the between-class categorization to improve performance of the system.

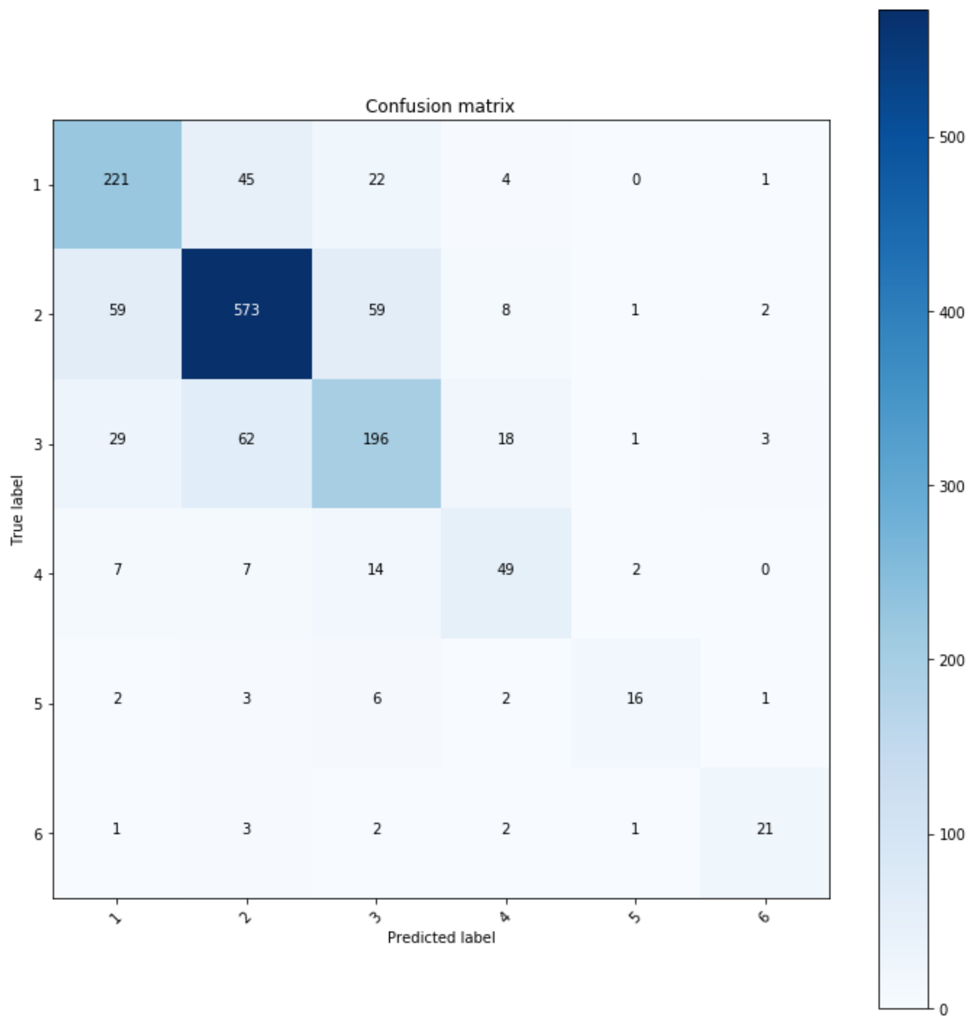

Figure 16. Confusion Matrix of Model#24 (Coarse)

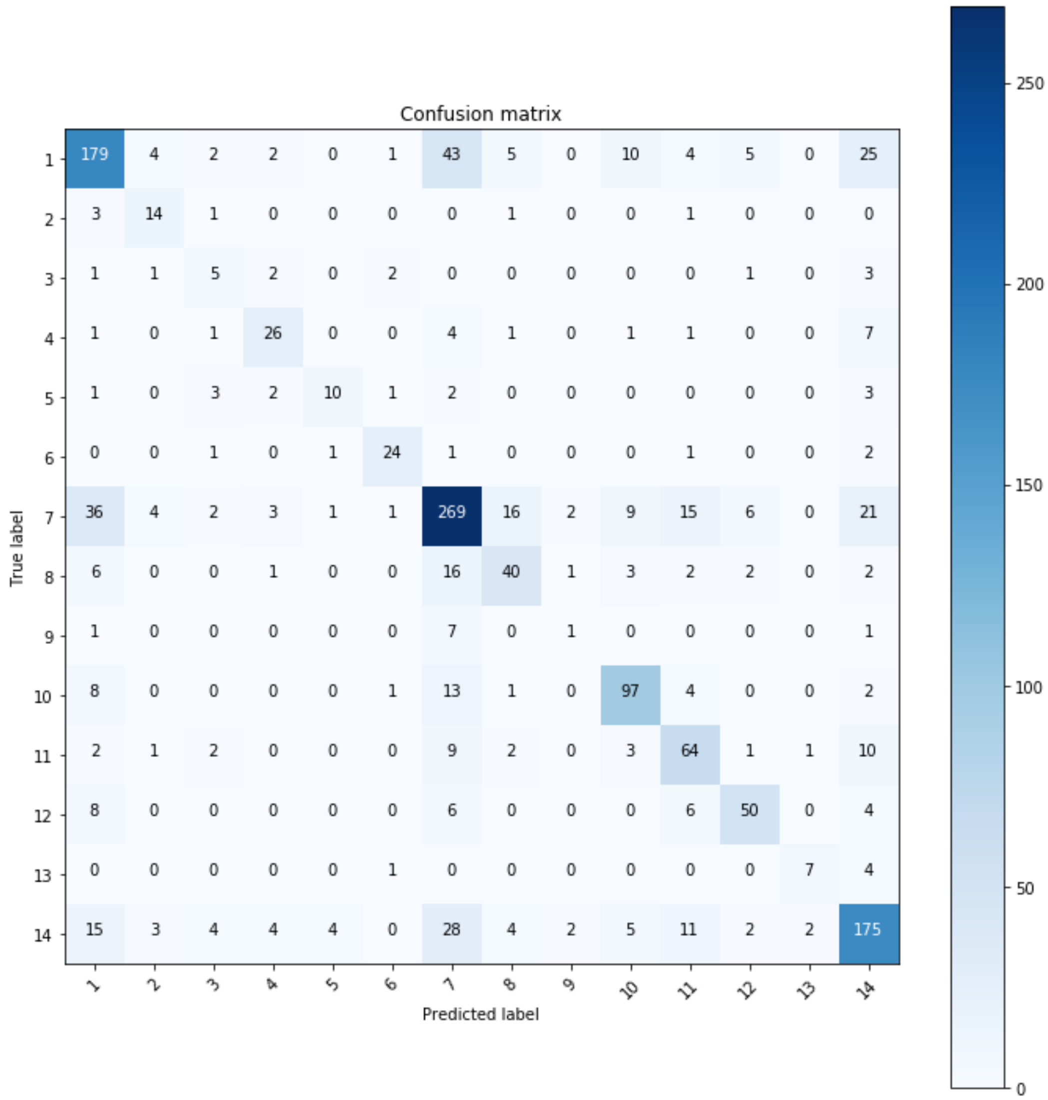

Figure 17. Confusion Matrix of Model#36 (Fine)

In the prediction of a fine-grained model (#36), similar to the main findings in Model#24, there is a diagonal effect with respect to the true positives in Figure 17. At the axis of true labels, it is noticeable that Class#1 (Self-Motion-Basic), Class#7 (Resultative-Attachment), and Class#14 (Discourse-Connective) are predicted with other labels. At the axis of predicted labels, these three classes are also predicted incorrectly. The error suggested by these two observations is that the boundary between these three classes may benefit from more clarification. For example, reconsideration is necessary for certain productive VDCs, such as VDCs with the V1 *jie* 'connect' in Examples 39A-C:

39. A. jie xia lai, women qu nar

link descend come, we-PL go where

"After that, where are we going?" (cf. CCL)

B. zai jie xia lai de yi tian ban shijian li, pujin jiang tong

at connect come descend ASSOC one day half time inside, Putin FUT with

furen zai kaisiteluo de bieshu zhong xiuxi

madam at Castro ASSOC cottage inside rest

"For the following day and a half, Putin and his wife will rest in Castro's cottage." (cf. CTB)

C. na-shi haiku yi-dai cai jie shang dian-xian

DEM-time Haikou one-belt just connect ascend electricity-line

"At that time, the Haikou City and its nearby area just had their electrical wires connected." (cf. CTB)

Usually, the VDC *jie xia lai* 'connect descend come' functions as an expression connecting different pieces of discourse or action in the flow of time, which should be labeled as Class#14. There is no explicit subject in (39A). Example 39B is different from (39A) in that the same VDC is embedded in a relative clause with its Theme argument *shijian* 'time.' This metaphoric case is similar to the wire-connecting example in (39C), annotated as Class#7. Confusion on the part of the models between predicting (39B) as #14 or #7 accords with human understanding of these examples as potentially ambiguous. A practical solution to such ambiguous cases is to consistently treat "*jie* + DV" as examples of #7 in the future annotations.

3.5.3 Future Improvements in the VDC Event Classification Systems

The performance of the VDC detection and fine and coarse grained event type classification show the plausibility of developing a VDC classification system that first automatically collects VDC tokens and then organizes them into different event categories. There are three ways to improve this supervised VDC event type classification system. First, definition boundaries of constructions should be further clarified by reorganizing and reanalyzing some categories. Second, in order to let the taxonomy better serve NLP applications the taxonomy can be revised for specific purposes, i.e., disambiguating Aspectual VDCs vs. Motion VDCs or, VDCs encoding economic vs. non-economic activities. Third, building a VDC classification system is not the ultimate goal of VDC parsing. The next step of the research is to create a Semantic Role Labeling (SRL) system that can tag the entities and relations for different VDC-initiated events, or otherwise integrate the VDC annotation into the training of preexisting semantic parsers, such as CPB, CAMR, and Chinese FN parsers.

Chapter 4
Detecting Metaphors in VDCs

Research Question: How can we develop tasks to distinguish metaphoric from non-metaphoric VDCs?

This section begins with competing theories developed to explain the nature and properties of metaphors. To carry out this analysis, I rely on Conceptual Metaphor Theory (CMT) (Lakoff & Johnson 1980, 1998). CMT is more useful than other theories for distinguishing between metaphoric and non-metaphoric ones because it not only reveals the cognitive nature of metaphors, but also provides an analytic framework for considering the productivity and systematicity of metaphors (Sec.4.1). Next, to reduce the annotation effort required by this task, I apply both CMT and the VDC taxonomy to define the scope of metaphor in this taxonomy (Sec.4.2). After examining the current Chinese NLP resources for metaphor processing (Sec.4.3), I evaluate performance of the metaphoric VDC classification system (Sec.4.4.1) and point to the future improvements (Sec.4.4.2).

4.1 Understanding Metaphors in VDCs with Conceptual Metaphor Theory

This section provides a theoretical foundation for analyzing figurative language in general and VDC metaphors in particular. The development of linguistic theories of metaphors began in 1960s and has resulted three approaches: interaction, comparison, and cognition. The interactive view (Black 1963) evaluates metaphors from the perspective of the speaker's intentions. The speaker's production of metaphors is related not only to the degree of variation in the choices available to the speakers, but also to the expected implicature of the metaphor. This view overlooks common patterns of metaphors abstracted from individual speakers' metaphor variation.

The comparative view (Searle 1993) treats metaphors conceptually by formally defining them as "X is Y." In this view, since the speaker can use the sentence meaning ("S is R") to convey the utterance meaning ("S is P"), there should be a systematic mapping between P and R. The major problem of the comparative view is that it narrows the scope of metaphors to include only similes with a fixed sentence type; sentences like *they ate up the funding* are not considered to be metaphors in this approach.

The cognitive view, CMT (Lakoff & Johnson 1980, 1998), overcomes the limitations of both interactive and comparative views by arguing that metaphors reflect the fundamental analogical basis on which how humans acquire new knowledge through experience. Furthermore, the CMT approach has three advantages in explaining the mechanism of metaphors. First, metaphors are cross-domain conceptual analogies, which are generally shared by different languages (Lakoff 1993; Lakoff 1998). Each metaphor has a Source domain and Target domain. Take *foundations of a theory* for example. The more concrete domain, a building with a *foundation*, is the source. The more abstract domain, *theory*, is the Target. Thus, the conceptual metaphor here belongs to a broader metaphor THEORY IS A BUILDING. Metaphors are not simply driven by languages, but by shared human experience. For example, among various kinds of metaphors, spatial metaphors are commonly seen in different languages (Kövecses 2005). Take the metaphorization of the morpheme denoting upward motion in both Chinese VDCs and English VPCs:

40. A. HAPPY IS UP

ta hen-kuai jiu gaoxing qi lai

he very-fast then happy rise come

"Then he quickly cheered up." (cf. CCL)

B. VISIBLE IS UP

yanjiu-shi de deng liang qi lai

research-room ASSOC light light rise come

"The light in the research room lit up." (cf. CCL)

C. ACCESSIBLE IS UP

ji-zhe yi-shi ye da bu shang lai

record-person one-time also answer NEG ascend come

"The journalist also could not come up with the answer in the moment."

Although the Chinese VDCs in Sentences 40A-C reveal different choices among directional morphemes to express these metaphors (DVs *qi lai* 'rise come' and *shang qu* 'ascend go' instead of the particle *up*), these examples show that certain source-target relationships are shared between Chinese VDC data and their English VPC translations.

Second, metaphors are productive, and have different layers of analyses (Shutova 2015). At the conceptual level, a metaphor is understood as X IS Y. While at the level of expression (usually in sentences), metaphors are known as linguistic metaphors, e.g., *the government is* *driving down the road of disaster*. In linguistic metaphors, some words have the potential to evoke analogical relationships. These words are called metaphor-related words (MRWs). For example, in the linguistic metaphor *the government is driving down the road of disaster*, the verb *drive* is an MRW. At the discourse level, metaphors can connect with each other to construct a coherent envisionment of a scene, as in Example 41:

41. "Bobby Holloway says my imagination is a three-hundred-ring circus. Currently I was in ring two hundred and ninety-nine, with elephants dancing and clowns cart wheeling and tigers leaping through rings of fire. The time had come to step back, leave the main tent, go buy some popcorn and a Coke, bliss out, cool down." (cf. Koontz 2007)

The topical conceptual metaphor here is IMAGINATION IS A CIRCUS, but there are many related metaphoric elaborations in the sentences, such as cart, elephants, popcorn, and so on. If the expression *go buy some popcorn and a Coke* has no context, it is difficult to fully interpret its metaphoric meaning.

These different layers of analyses can also be performed on metaphoric VDCs. Take VDCs with the DV *qi lai* 'rise come' for example. The VDC *gaoxing qi lai* 'cheer up' in Sentence 40A is based on the metaphoric frame HAPPY IS UP. At the same time, this sentence is counted as a linguistic metaphor, in which the MRWs are the compounding DVs *qi lai* 'rise come' (the mechanics of MRW in VDCs will be discussed in the next section). Finally, a linguistic metaphor like *xi wang liang qi lai le* '[Someone] lit up with hope' can co-occur with other metaphors under the superordinate metaphor HAPPINESS IS LIGHT are illustrated in the discourse below:

42. (A) wo xiang ta wanzheng biaoda ziji de shihou,

I to her completely express REFLX ASSOC time

(B) jiushi ba yang-guang <u>da</u> <u>kai</u>,

then BA sun-light hit open

(D) rang guang zhao <u>jin</u> <u>lai</u>, (D) zhaoyao bici, (E) yinmai bu zai,

let light shine enter come, shine RECP, cloud NEG EXIST

(F) xiwang liang <u>qi</u> <u>lai</u>

hope light rise come

"When I fully expressed myself to her, it was like a burst of sunlight. I let the light come in, and we were showered by it. The dark clouds were gone. We lit up with hope." (cf. Google Search)

The above discourse example is a statement from a psychotherapist about a treatment encounter. Clause B shows that the sunlight is hidden in a container (A PERSON IS CONTAINER FOR EMOTIONS). In Clause C, the patient's mental state is analyzed as a contained dark space (DEPRESSION IS DARKNESS). Clause D expands the happiness as a light metaphor to both the patient and the therapist. Clause E treats the psychological problems as dark clouds (DEPRESSION IS DARKNESS). Clause F shows that the speaker and the patient's hope was lit up (HOPE IS EMISSION OF LIGHT[17]). Clause F with the VDC *liang qi lai* 'light up' is embedded in a wider metaphoric context in which all metaphors form dependent relationships. For example, the patient needs hope as light because their psychological state is like a dark space. Also, since giving hope is a mutual process in this context, both the therapist and the patient enlighten each other. While in the circus example in English we see elaboration of one central metaphorical mapping (IMAGINATION IS A CIRCUS), the metaphors in 42B-F demonstrate that the manner in which several different general mappings (based on location of light, containment and light emission) may cohere.

Third, speakers' knowledge of conceptual metaphors is systematic. Taking the orientational metaphor MORE IS UP as an example (Lakoff & Johnson 1980), the same underlying analogy is mapped to many different expressions such as *speak up*, *turn up*, *high temperature*, and *the number of orders rose*. These expressions may be morphosyntactically related or unrelated, but they share the same cognitive foundation. Furthermore, metaphors like MORE IS UP and LESS IS DOWN can be generalized as a more abstract metaphor QUANTITY IS VERTICALITY.

The systematicity of metaphoric mappings can also be found in some VDCs. For example, the three pairs of VDCs below indicate antonymic mappings that can be generalized as a simpler metaphor:

43. A. Self-Motion-Shifted

(i) gan shang lit ('chase ascend') 'chase'

APPROACHING HORIZONTALLY IS ASCENDING

[17] According to my fine-grained VDC taxonomy, the DV belongs to Resultative, which is not based on the semantics of V1s. The conceptual metaphor of the DVs is ACHIEVEMENT IS RISING AND COMING.

(ii) tui xia lit (‘retreat descend’) ‘retreat’

MOVING AWAY HORIZONTALLY IS DESCENDING

B. Resultative-Attachment

(i) chuan shang lit (‘wear ascend’) ‘put on (the clothes)’

ATTACHMENT IS ASCENDING

(ii) tuo xia lit (‘separate descend’) ‘take off (the clothes)’

DETACHMENT IS DESCENDING

C. Caused-Motion-Tranfer-of-Possession

(i) mai jin lit (‘buy enter’) ‘buy in’

GAINING OWNERSHIP IS ENTERING

(ii) mai chu lit (‘sell exit’) ‘sell out’

LOSING OWNERSHIP IS EXITING

The more general metaphor in the VDC pairs in 43A-C can be further analyzed as HORIZONTAL MOTION IS VERTICAL MOTION, ATTACHMENT IS ASCENDING, and POSSESSION IS A LOCATION, respectively.

Given that CMT offers an in-depth understanding of the nature and features of metaphors, this study adopts a CMT perspective to analyze metaphoric VDCs. The next section examines the benefit of applying CMT to the VDC taxonomies in order to detect metaphors.

4.2 Using the Metaphor and VDC Taxonomies to Reduce Annotation Workload

When designing the annotation task for determining whether a given instance of VDC is a metaphor, an important consideration is to reduce the annotation burden. This section demonstrates that the taxonomy for VDC classification, together with a metaphoric taxonomy inspired by CMT, can help reduce the workload in metaphor annotation.

According to CMT, VDC metaphors not only follow general features like being conceptual, productive, and systematic (Dippner 2010; Yin 2011; Yu 2012), but also allow different modes of contribution by MRWs. Sullivan (2013) noted that the main contributors of metaphoric frames in a sentence are predicates. For example, in the sentence *she stole my hear*t, the verb *steal* evokes the metaphoric frame HEART IS AN OBJECT in which the stolen Theme should be perceived as concrete. When analyzing the function of the particle *out*, Morgan (1997: 355) argues that the VPCs for *out* can be classified into four classes depending on whether the verb or the particle is an MRW: “literal container and literal verb” (e.g., *he took the toy out of bag*), “literal container and extended verb” (e.g., *we fished out the ring (from a bowl)*), “literal verb and metaphoric container” (e.g., *I handed out the quizzes*), and “metaphoric container and extended verb” (e.g., *the journal comes out once a year*). Since VDCs and VPCs share much in the semantics of both manner verbs and directional morphemes, a similar analysis can be applied to VDCs:

44. A.V1(literal) + DV(literal)

ta-men pa shang shan

3P-PL climb ascend mountain

"They climbed up the mountain."

B. V1(metaphoric) + DV(literal)

youmei-de xuanlü ba wo huan hui lin-yin xiao-dao

beautiful-ASSOC melody BA I call return forest-shadow small-path

"The beautiful melody called me to come back to the tree-lined path."

V1(metaphor): MUSIC IS A PERSON

C. V1(literal) + DV(metaphoric)

ta jiu zhe-me zou xia qu

she then DEM-SUF walk descend(ASP) go(ASP)

"She continued walking like this."

DV(metaphor): CONTINUING IS DESCENDING AND GOING

D. V1(metaphoric) + DV(metaphoric)

hai-zi-men song shang zhufu

kid-SUF-PL send ascend wish

"The kids sent up their wish (to someone who is respectable)."

V1(metaphor): COMMUNICATION IS SENDING

DV(metaphor): STATUS IS HIGH

These examples illustrate three kinds of metaphors: "metaphoric + literal" in (44B), "literal + metaphoric" in (44C), and "metaphoric + metaphoric" in (44D). Similar to the VPC analysis, VDCs have two sources of metaphors: V1 or DV. Metaphors evoked by V1 usually change concreteness of the argument(s). For example, the metaphor in (44B) MUSIC IS A PERSON shows that the Agent is [-concrete]. Metaphors on DV/V2 can be also viewed as the grammaticalized functions of DVs, which do not necessarily change the features of arguments. For example, there is no metaphor in the arguments in (44C). The metaphor is indicated by the aspectual use of the compounding DVs *xia qu* 'descend go.'

Based on different analyses of MRWs in VDCs, we can reexamine taxonomies introduced in Section 3.2 (repeated as Figure 18 at the end of this section). The general differentiation standard for classes like Self-Motion, Caused-Motion, Resultative, Aspectual, Discourse-Connective, Evidential, as well as their sub-classes, is based on event structure metaphors (Lakoff 1990) of DVs or locatives. DVs developed from the source domain of physical motion to target domains like result, time, discourse, and so on; some VDC classes, such as CM-VDCs, are based on metaphoric extensions of physical locations to abstract domains like possession, values, and creation. The fine/coarse-grained VDC taxonomy is related to metaphoric DVs and locatives. A table including subclasses of VDCs, their VDC-specific metaphorical mappings, and possible metaphor types is given below. Sentence examples of this Table are listed in Appendix 2.

VDC Name	Class#	VDC-specific Metaphor	Metaphor Types
Basic	**1-1**	**Literal**	**NA**
Basic	1-2	Metaphoric	ML
Shifted	2	HORIZONTAL MOTION IS VERTICAL MOTION	ML or MM
Displacement	**3-1**	**Literal**	**NA**
Displacement	3-2	Metaphoric	ML
Transfer	4	POSSESSION IS A LOCATION	LM or MM
Scale	5	VALUES ON A SCALE IS A LOCATION	LM or MM
Creation	6	CREATION IS EXITING	LM or MM
Attachment	7	ATTACHMENT IS ASCENDING	LM or MM
Reflexive-Motion	8	SPREAD/CONTRACTION IS DIRECTED MOTION	LM or MM
Affordability	9	ATTAINMENT IS DIRECTED MOTION	LM or MM
Completive	10	COMPLETION IS DIRECTED MOTION	LM or MM
Inchoative	11	BEGINNING IS DIRECTED MOTION	LM or

			MM
Continuous	12	CONTINUATION IS DIRECTED MOTION	LM or MM
Discourse-Connective	13	(DISCOURSE) TIME IS DIRECTED MOTION	MM
Evidential modality	14	EMERGENCE OF EVIDENCE/COMMENT IS DIRECTED MOTION	LM or MM

Table 21. VDC Subclasses and Their Metaphor Types

As shown in Table 21, most subclasses, except for Self-Motion-Basic (Class #1) and Caused-Motion-Displacement (Class #3), are always metaphoric. Once we split #1 and #3 into "Literal" and "Metaphoric/Figurative," the literal VDCs can be identified in the taxonomy. Therefore, the "non-metaphoric VDCs vs. metaphoric VDCs" distinction is conceptually the same as "Self-Motion-Basic-Literal (#1-1) and Caused-Motion-Displacement-Literal (#3-1) vs. other classes" in the VDC taxonomy. As a result, the mapping between VDCs' fine-grained classes and their metaphor assignments can support the annotation of metaphor detection. Thus, adding a brief annotation based on the fine-grained VDC taxonomy (i.e., annotating Class #1-1 and #3-1), eliminates the need to annotate metaphors for every subclass of VDCs.

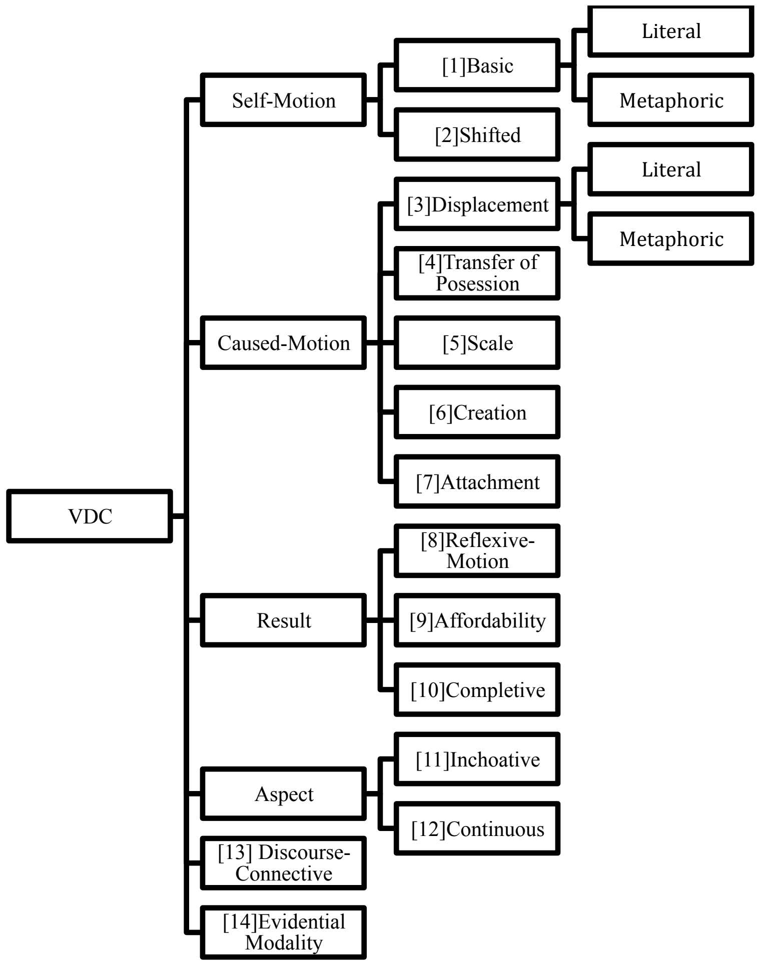

Figure 18. A Taxonomy for VDC Event Classification (Figure 6)

4.3 Current NLP Resources for Types of VDC Event Structures

In addition to theoretical bases of metaphors, it is also necessary to evaluate the current representation of metaphoric VDCs in Chinese NLP resources. Chinese MIPVU (Metaphor Identification Procedure Vrije Universiteit) from Penn State University (Lu & Wang 2017) is a metaphor corpus containing 30,012 words from the Lancaster Corpus of Chinese (McEnery & Xiao 2004). It includes both linguistic metaphor and word-level metaphor annotations.

45. <w POS="r">他</w> (*ta* ‘he’)

<w POS="v">脱</w> (*tuo* ‘take.off’*)*

<w POS="v"><seg function="mrw" type="indir" subtype="conv">下</seg></w> (*xia* ‘descend’)

<w POS="n">皮鞋</w> *(pi-xie* ‘leather-shoes’)

The example above[18] shows the metaphor annotation for the sentence “he took off his leather shoes.” Note that the DV *xia* carries the metaphor DETACHMENT IS DESCENDING, so it is annotated as an MRW. An indirect metaphor means that there is relevance between the contextual or extended meaning and the basic meaning[19]. For example, the basic sense of *xia* is ‘descend.’ Its actual meaning, ‘detach (from the wearer’s body),’ is indirectly related to the basic meaning. One advantage of this corpus is that it treats most functions of VDCs as MRWs, such as VDCs with modality or aspect. However, just like CPB, its coverage of VDCs is limited, with only 106 VDC examples in total (3.8 examples per type of VDC). Since this corpus does not cover enough combinatory diversity between V1s and DVs in VDCs, additional detection of metaphoric VDCs is needed.

4.4 Implementing, Training, and Evaluating an Automatic VDC Metaphor Classifier

4.4.1 Results

[18] The linguistic glosses in “( )” are not in the original MIPVU annotation.

[19] http://www.vismet.org/metcor/documentation/relation_to_metaphor.html

As mentioned in Sec.4.2, the metaphoric VDC annotation was performed on top of the layer of fine-grained VDC annotation. Once the literal Self-Motion and Caused-Motion categories were annotated in terms of metaphors, their opposition with other VDC categories was identical to a non-metaphor vs. metaphor distinction. There are 12,524 metaphoric VDCs, and 3,328 non-metaphoric VDCs.

Model#	Context	Model	Precision	Recall	F-Score
43		EM0+Char	0.85	0.82	0.83
44		EM0+Seg	0.85	0.85	0.85
45		EM1+Char	0.86	0.87	0.86
46	Long	EM1+Seg	0.86	0.87	0.86
47		EM2+Char	0.86	0.88	0.87
48		EM2+Seg	0.86	0.84	0.85
49		SVM	0.8	0.84	0.83
50		EM0+Char	0.86	0.85	0.86
51		EM0+Seg	0.88	0.83	0.86
52		**EM1+Char**	**0.88**	**0.86**	**0.88**
53	Short	EM1+Seg	0.88	0.86	0.87
54		**EM2+Char**	**0.88**	**0.87**	**0.87**
55		EM2+Seg	0.88	0.83	0.85
56		SVM	0.5	0.82	0.71

Table 22. Metaphoric VDC Detection Results

The results are summarized in Table 22. The basic finding is that the neural models have higher F-scores than SVMs. These F-scores generally range from 0.83 to 0.88, higher than the percentage of metaphors in the data (73.4%). Although F-scores for neural models with the word-level embedding were slightly better than the other models, the advantage is not obvious. After five runs, Models #52 and #54 have very close performance (F-core = 0.874 and 0.876, respectively), which is marginally better than #47 (0.867) and #53 (0.857) after three runs. Unlike the findings in VDC event type classifications, there is no segmentation effect. Although the length effect is unclear for biLSTMs, the long context condition is better than the shorter one in SVM models.

As a final evaluation, the best current models are compared to the best pilot ones on the pilot testing data in Table 23. There is a slight increase in the F-score of biLSTM (5%), but a big one in SVM (67%). As for SVM, the use of PCA was the main contributing factor to the improvement.

Model	Precision	Recall	F-score
LSTM- pilot: EM1+seg	0.78	1	0.88
LSTM-Model#52	0.95	0.92	0.93
SVM-pilot	0.02	0.25	0.14
SVM-Model#49	0.85	0.74	0.81

Table 23. A Comparison Between Current Best Models and Best Pilot Models for Metaphoric VDC Detection

4.4.2 Future Improvements

This chapter examines the performance of a detection system for metaphoric VDCs. However, detecting metaphoric VDCs is only the first step toward automating metaphor processing. As described in Section 4.1, humans use conceptual metaphors as a fundamental way to structure and understand the world. Future NLP systems should be able to reproduce the different dimensions of human metaphor processing. To reach this goal, several additional tasks are required. First, after detecting metaphors, the machine should able to label the targets and sources, or output the metaphoricity score for each word; and since literal VDCs are less frequent than metaphoric ones, the system must also be trained to label concrete roles in non-metaphoric VDCs.

Second, the machine must be designed to convert metaphoric VDCs into literal language, especially when the meaning is implicit knowledge in the context. For example, in the VDC *cha jin lai yi ju hua* lit ('insert enter one CL speech') 'add a sentence by interjection,' the space of insertion, namely the conversation, is only implicitly understood by the speakers. To accomplish this conversion task, a sequence-to-sequence framework trained on pairs of metaphoric and literal VDCs will be required.

Third, detecting metaphorical VDCs can help improve machine translation of metaphors. Although orientational metaphors are common in both Chinese and English, each language has specifications or cultural preferences of the actual metaphorical frames. For example, given that the imperial authority in China was not overthrown until the early 20th century, the mapping POLITICAL DEPEDENCE IS KNEELING DOWN sounds more natural to Chinese speakers, while in English, the mapping SUBJUGATION IS DOWN[20] is more natural. Translating *zai 1949 nian qian weishenme zhongguo renmin shi gui xia lai de* as 'why were Chinese people under control before 1949' is more acceptable than the literal expression 'why did Chinese people kneel down before 1949' to English speakers. After we detect metaphorical VDCs, instances with low confidence translation (linguistically and culturally) in MT systems might need to be situated in a culturally relevant context.

Finally, parsing metaphor dependencies in discourse also remains a challenge. This requires linking the metaphor taxonomy of VDCs with MIPVU or MetaNet (David et al. 2014) to capture the representations of metaphoric VDCs as well as other types of metaphors in the VDCs' surrounding context.

---

[20] https://metaphor.icsi.berkeley.edu/pub/en/index.php/Metaphor:SUBJUGATION_IS_DOWN

Chapter 5
Detecting and Analyzing Coerced VDCs

Research question: Can we develop a system to detect innovative uses of the VDC pattern? In particular, what is the relationship of verb meaning to construction meaning when a novel verb is used in the V1 position?

Four prerequisites to implementing a coerced VDC classifier will be examined in this chapter. First, we must find an appropriate theoretical framework for describing the mechanism of coercion in caused-motion predications in English (Sec.5.1). Second, we must find a way to apply the relevant theories, developed for English, to VDCs (Sec.5.2). Third, we must establish a practical definition of coercion for the purpose of annotation (Sec.5.3). Based on the annotation guideline, and in light of the limitations of Chinese NLP resources (Sec.5.4), I implement a detection task (Sec.5.5). Finally, I discuss the improvements of the detection system (Sec.5.6).

5.1 Explaining Coercion in Construction Grammar

"Coercion" is used by theorists of syntax-semantics interface to refer to adjustments of word meanings in context (De Swart 1998; Aduring & Boiij 2016). The term covers many kinds of phenomena, including mismatch between a count or mass nominal and its determiner, as in, e.g., *a water* (Michaelis 2003), mismatch between grammatical and lexical aspect (Michaelis 2004), as in, e.g., *I'm liking it*, construal of nominal complements of aspectual verbs (Pustejovsky & Jezek 2008), as in, e.g., *She started a book*, and so on. Although less commonly recognized as such, coercion occurs when a verb is combined with an otherwise incompatible argument-structure construction, causing a change in the verb's semantic and combinatoric properties. One such phenomenon is the uses of a non-causative verb in the Caused-Motion pattern:

46. A. When a visitor passes through the village, young lamas stop picking up trash to mug for the camera. A gruff 'police monk' <u>barks</u> them back to work. (attested example cited by Michaelis 2004)

    B. Cynthia <u>blinked</u> the snow off her eyelashes. (attested example cited by Hwang 2014)

The manner verbs, *bark* and *blink*, are usually intransitive; but in Examples 46A and 46B, the VPCs allow the prepositional phrases *back to work* and *off her eyelashes*. Two opposing views can be used to describe the Caused-Motion Construction examples above. The first view is a lexical-item explanation. According to the Projection Rule, syntactic structures "must be projected from Lexicon, in that they observe lexical properties of the items they contain" (Radford 1988: 391). In other words, each verb licenses a particular syntactic structure. This rule is a basic assumption in Generative Transformational Grammar (Chomsky 1995), Lexical-Functional Grammar (Kaplan & Bresnan 1982), Head-Drive Phrase Structure Grammar (Pollard & Sag 1994), and Role and Reference Grammar (Van Valin 2009). According to the lexical view, when verbs appear with non-canonical argument arrays, they should be treated as senses distinct from their original ones. Thus the examples in (46) illustrate new caused-motion verbs *bark* and *blink*. This solution is not only counter-intuitive, but also inefficient.

The second view, the context view, is the position that coercion is semantic reinterpretation in context. Theories like CxG (Fillmore 1988; Croft & Cruse 2004; Goldberg

2004 & 2006, etc.) and Generative Lexicon (GL) (Pustejovsky 1991) seek interpretive mechanisms to explain the motivation behind non-canonical interpretations of verb senses. Thus these researchers maintain that the combinatoric properties of a word are not always determined by that word alone but can also be influenced by other elements in the context. That is, there is an interaction between the meaning of the word and the meaning of the construction in which it appears. According to GL, new word meanings emerge in context because of semantic co-composition, not additional sense (Pustejovsky & Jezek 2008). Take for the example the sentences *they believed the books* and *they burned the books*. The salient property of the NP *books* can be physical or informational. When this NP is the object of the verb *burn*, the physical interpretation is required; when it is the object of the verb *believe*, the content reading is foregrounded. However, when we use GL to describe caused-motion VPCs, there are two problems that may arise: first, GL focuses primarily on domain shifting and preserving processes (Pustejovsky 2007), which do not appear applicable to abnormal configurations of arguments in constructions. Second, GL does not relate coercion to a theory of construction meaning (Michaelis 2005).

Unlike the within-lexicon operation perspective of GL, CxG views coercion as by-product of the ordinary signification of constructions (Michaelis 2006a; Audring & Booij 2016). In a non-coerced situation, the verb's argument structure is identical to the construction's; but in a coerced situation, the construction contributes arguments not supplied by the verb (Michaelis 2003). When a "mismatch" verb is embedded in a construction, the construction alters the argument structure of the verb through an integration relationship called argument augmentation (Goldberg & Jakendoff 2004; Michaelis 2006b). When applying the construction approach to the sentences in (46), we can say that the intransitive verb's argument array is augmented up to that of Caused-Motion Construction (i.e., Agent, Theme, Goal), in the service of verb-construction conflict resolution. However, although this approach reasonably describes the Caused-Motion Constructions in English, it must be adjusted to describe how SVCs work in Chinese. The Chinese situation will be discussed in the next section.

5.2 Analyzing Coerced VDCs

As in the foregoing account of Caused-motion VPCs, this study adopts a constructionist view of coercion to analyze VDCs. The main difference between VPCs and VDCs is that the caused-motion construal of a CM-VDCs results from the combination of two verbs' argument structures rather than attaching to a single argument structure[21]. Therefore, instead of modeling the argument structure of a CM-VDC as "Verb (Agent, Theme, Locative)," I will present the general argument array of CM-VDC as "**VDC [Agent, Theme/Patient, Locative] = V1 (Agent, (Theme/Patient)) + DV (Theme, Locative)**" according to the argument-sharing pattern of the particular VDC[22]. The first argument structure in the array encodes the instigation of motion by the Agent potentially manipulating the Theme (or performing any change of state on the Patient), while the second array expresses the Theme's motion trajectory. A critical observation is that the verb lexeme chosen for the V1 slot may not lexically entail causation of motion, although it will

[21] Although Caused-Motion VPC is different from CM-VDC because the English VPC only has one verb, it is still possible to think of the VPC as a construction with two argument structures. This is because the PP in a VPC, as a relational category (Gentner & Kurtz 2005), still takes the Theme of the verb as its argument. Therefore, the fusion in the sentence *They kicked me out of the campus* means the Patient of the verb *kick* is the Theme/XARG of the directed-motion PP *out of*, which is not very different from the fusion process in which the Theme or Patient of V1 is the Theme of V2 in Chinese CM-VDCs.

[22] Because of that the coerced CM-VDC is the most frequent coarse-grained type of coerced VDC in the dataset (135 out of 155 tokens), only CM-VDC coercion modeling is described in this section.

have this entailment in the context of the VDC construction. If V1 is, for example, a causative change of state verb, rather than a causative motion, the VDC imposes a directed-motion reading on V1, as in Example 47, whose two sub-events are captured by the paraphrases in (ii) and (iii):

47. (i) ni gancui zhao ge leng-dong-gui, ba ziji dong jin qu

you simply find CL cold-freeze-closet, BA REFLX freeze enter go

"You can simply find a fridge, and freeze yourself in it."

(ii) ni dong ziji "You freeze yourself."

(iii) ziji jin qu "You yourself enter (the fridge)."

The coerced VDC is the second clause *(ni) ba ziji dong jin qu* in 47(i); we characterize this as a case of argument sharing because the Patient of V1 *dong* 'freeze' is identical to the Theme of the DVs *jin qu* 'enter go,' as represented by 47 (ii) and (iii), respectively. The verb *dong* is a causative change of state verb ('cause to be frozen'); accordingly, clause (ii) shows it to have a transitive pattern. The verb *dong* 'freeze' does not select for a Goal argument; it is the VDC construction that licenses that argument (*lengdonggui* 'the fridge'), via identification of the Patient of V1 with the Theme of the VDC. As noted earlier, coerced VDCs may contain intransitive V1s, in which case there is no "fusion" of Patient and Theme roles. Examples are given in (48), along with Google machine translations that appear to miss the mark for these particular tokens:

48. A (i). wo gui ye yao ba ta gui hui jia

I kneel still want BA she kneel return home

"I would even get on my knees, if it would make her come home (lit. Even if I kneel down, I want to kneel her to return home.)" (cf. CCL)

[Google Translate] I have to pick him up.

(ii)*wo gui ta "I kneeled her."

(iii) ta hui jia "She returned home"

B. (i) gong-kuan-chi-he bu neng ba GDP chi shang qu

Public-money-eat-drink NEG can BA GDP eat ascend go

"Eating and drinking on tax-payers' money cannot eat GDP to go up." (cf. CCL)

[Google Translate] Public money can not eat GDP.

(ii) *gongkuanchihe bu neng chi GDP "Eating and drinking on tax-payers' money cannot eat GDP."

(iii) GDP shang qu le "GDP went up."

Because (48A) and (48B) contain intransitive V1s, the event-structure entailments in 48A(i) and B(ii), respectively, are invalid. These are examples in which no fusion (in the sense of Goldberg 1995) takes place. Because the in the V1 slot has no second argument, there is no argument to identify with the Theme of the VDC. This is similar to the case of coerced English CM predications with *blink* and *bark* in (46) above. In (48A), the V1 *gui* 'kneel' is intransitive, therefore it does not take a Patient argument, as shown by the invalidity of the entailment in

48A(i). The case of the V1 *chi* 'eat' in 48B is similar, although not obviously so because the verb takes two arguments in FN: Ingestor and Ingestable. In (48B), however, we see an intransitive use of the verb, in which the Ingestable argument does not appear. An analogous English CM predication is *He's going to eat me out of house and home*. While the construction supplies a Theme argument (the first argument of V2), it is not selected argument of V1. For this reason again, the entailment presented in 48B(ii) is not valid.

5.3 Constructing an Annotation Standard for Coercion

Now that we have used the CxG framework to describe the general mechanisms of coercion, our next step is to operationalize the definition of coercion for the purpose of annotation. To do so we must answer three questions: (1) How can annotators verify the novelty of V1s? (2) How can they distinguish coerced VDC tokens from other novel uses of VDCs? (3) What VDC types as defined in the coarse-grained taxonomy of VDC uses described in Sec.3.1, potentially display coercion effects?

To answer the first question, I define the novelty of V1s in these CM-VDCs in two ways: using verb lists and intuition. I created a list of 430 typical Caused-Motion verbs in Chinese (using 171 from Li 2008 and 259 from Luo 2015); these verbs are commonly seen in non-coerced Self-Motion and Caused-Motion VDCs. Annotators can combine their intuitions with these lists to decide whether a certain Caused-Motion VDC is coerced or not: if a verb is not on the list it stands a good chance of being a novel verb in the construction. The VDC in (49) below is an example of coercion from conversational data.

49. wanshang shang weibo, hai zhen kan dao ge-bi
evening ascend micro-blog, still real look reach next-wall
na ya de weibo shipin, hai zhen pai shang qu
DEM girl ASSOC micro-blog video, still real shoot ascend go
chuan weibo le
upload microblog ASP

"(I) logged onto microblog last night. Surprisingly (I) saw the microblog video the person next door uploaded. He/she actually took the video and put it onto microblog." (cf. CTB)

In (49) the adverb adverb *hai zhen* (lit 'still real') 'really' shows the speaker's surprise at the unexpected fact that the video had been uploaded onto the microblog. The coerced VDC is the first SVC in the last clause, *pai shang qu* 'shoot ascend go,' which means something like 'shoot the video onto the microblog.' The Theme of V1 *pai* 'shoot/take,' is unrealized, since *shipin* 'the video' is mentioned in the prior context. The Goal argument of the V1, *weibo* 'microblog,' is also unexpressed because it is mentioned in both the previous context and the following SVC *chuan weibo* 'upload to the microblog' in the same clause. Although this coerced VDC does not have all arguments expressed, annotators can still identify it as a case of coercion because the verb *pai* 'shoot' is not a typical caused-motion verb, based on their intuition and this verb's absence from the verb list.

To answer the second question, I instructed the annotators to distinguish between coerced novel V1s with novel compound verbs, as in (50).

50. A. hu-shi mao xia le yao

protect-person cat descend ASP waist

"The nurse bent down her waist like a cat."

B. hu-shi mao le yao "The nurse bent her waist."

C. yao xia le "Her waist went down."

Sentence (50B) shows that the noun has already became a manner of motion verb "to bend one's body parts like a cat" before entering the VDC. Although the use of *mao* 'cat' in (50A) is a creative change from a noun to a causative verb, we cannot regard this as a coerced VDC because it is not clear whether the argument sharing relationship is between V1 and V2: is the cat-like bending of the waist happening while her waist is moving downward, or is the bending the cause of the downward motion? The semantic relation in (50B) appears more typical of a compound verb than a control structure. Therefore, annotators were instructed to exclude VDCs with denominal V1s as potentially coerced VDCs.

As for the third question, since identifying instances of coercion in VDCs requires use to infer argument-sharing between V1 and V2 (whether or not the shared argument is a selected argument of V1), annotators must evaluate only those VDCs in which V2 has the argument structure that it would have outside a VDC, as a directional verb. It is only in such VDCs that we can interpret V1s as the means by which the V2 event occurs. VDCs with suffix-like DVs, such as aspectual DVs, cannot express a "means" relation and therefore cannot trigger coercion. Combining all three answers, the coerced VDCs are defined as "Self-Motion, Caused-Motion, and non-Completive Resultative VDCs with V1s that are primarily verbs in daily use" to the annotators.

5.4 Current NLP Resources for Coercion in VDCs

In order to evaluate the representation of coercion in Chinese NLP resources for a coercion detection task, I examined the coverage and representations in both CPB and CTB.

As a primary SRL corpus as well as a lexicon, the frame files in CPB keep track of the argument structure(s) of the predicates. There are two problems with CPB with regard to coercion: coverage and representation. First, there are not enough VDCs in the frame files of CPB, which can be illustrated by searching DVs and V1s.

Frameset: f1

ARG0: *entity described*
ARG1: *place arg0 climbs onto*

Frame:

```
(IP (NP-SBJ (NN 皱纹) )
 (VP (VV 爬上)
     (LCP-OBJ (NP (NN 眼) )
              (LC 角))))
```

ARG0: 皱纹
ARG1: 眼 角
REL: 爬上

Figure 19. The Frame File for the VDC *pa shang* 'climb onto' in CPB

Figure 19 shows the file of the predicate VDC (i.e., REL) *pa shang* lit ('climb ascend') 'climb onto' The file contains a frameset with ARG0 and ARG1, as well as a fully syntactically parsed example *zhouwen pashang yanjiao* 'the wrinkle climbed onto one's eyes.' I collected VDC predicates by searching some DVs.

DV	*shang* 'ascend'	*xia* 'descend'	*jin* 'enter'	*chu* 'exit'
VDC types	148	96	71	53
DV	*shanglai* 'ascend come'	*shangqu* 'ascend go'	*xialai* 'descend come'	*xiaqu* 'descend go'
VDC types	0	0	1	1
DV	*chulai* 'exit come'	*chuqu* 'exit go'	*qilai* 'rise come'	
VDC types	0	0	0	

Table 24. The Frequencies of VDCs Sorted by Some DVs in the CPB Frame Files

Table 24 shows that many DVs have few or no combinations with verbs, e.g., there is no file for the compounding DVs *chu lai* 'exit come' and *shangqu* 'ascend go.' I also searched for all possible DVs following the two frequent verbs *pao* 'run' and *dai* 'take.'

Main Verb	*shang*/ascend	*xia*/descend	*jin*/enter	*chu*/exit	*hui*/return	*dao*/reach
pao/run	yes	no	yes	no	no	no
dai/take	no	no	no	yes	no	yes

Table 25. The Use of VDCs Containing the Verbs *pao* 'run' and *dai* 'take'
in the CPB Frame Files

As is shown in Table 25, seen VDC patterns like ‘run down/out of/back/to’ and ‘take… onto/down/into/out of/back’ are not included in CPB. It is quite possible that uncommon VDC patterns are not well collected.

Second, current CPB lacks coercion-specific representations to explain the overriding effect of the argument structures of V1s in VDCs. For example, the Caused-Motion VDC in Example 48B is repeated below:

51. gong-kuan-chi-he bu neng ba GDP chi shang qu

Public-money-eat-drink NEG can BA GDP eat ascend go

“Eating and drinking on tax-payers’ money cannot eat GDP to go up.” (cf. CCL)

[Google Translate] Public money can not eat GDP

According to CPB, the simple VDC *chi shang* ‘eat ascend’ should be a compound verb, which requires at least three arguments: ARG0 (cause/agent), ARG1 (theme), and an implicit ARG2 (starting point). The verb *chi* ‘eat’ is profiled as another verb with ARG0 (agent) and ARG1 (theme). By listing these two predicates as separate entries, there is no layer of analysis showing how original transitivity of the V1, *chi* ‘eat,’ is changed in the VDC.

CTB cannot represent how rare the “V1 + DV” combination is in the syntactic structure. The syntactic tree containing the coerced CM-VDC *pai shang qu* lit (‘shoot ascend go’) ‘shooting causes the pictures to be uploaded onto somewhere’ in (49) is given below:

```
( (IP (IP (NP-SBJ (-NONE- *pro*))
          (VP (VP (NP-TMP (NT 晚上) )
                  (VP (VV 上)
                      (NP-OBJ (NN 微薄) ) ) )
              (PU , )
              (VP (ADVP (AD 还) )
                  (ADVP (AD 真) )
                  (VP (VV 看到)
                      (NP-OBJ (NP (NP (NN 隔壁) )
                                  (DP (DT 那) )
                                  (NP (NN 丫的) ) ) )
                              (NP (NN 微博)
                                  (NN 视频) ) ) ) ) ) ) )
      (PU , )
      (IP (NP-SBJ (-NONE- *pro*))
          (VP (ADVP (AD 还) )
              (ADVP (AD 真) )
              (VP (VP (VRD (VV 拍)
                           (VV 上去) ) )
                  (VP (VV 传)
                      (NP-OBJ (NN 微博) ) ) ) ) )
          (SP 了) )
      (PU 。) ) )
```

Figure 20. The CTB Tree Diagram of Example 49

The coerced VDC has the VP structure "[VP [VRD [VV V1 (*pai* 'shoot')][VV *shangqu* 'ascend go']]]," which is the same as a non-coerced VDC, e.g., *pa shangqu* 'climb ascend go.' This indicates that the coercion cannot be captured in the phrase structure representations. Our evaluation of CPB and CTB shows that neither of these resources represents the coercion accurately.

5.5 Implementing, Training, and Evaluating an Automatic VDC Coercion Classifier

Because of the rarity of coerced VDCs, the distribution of coercion vs. non-coercion is extremely imbalanced. After annotation, 10,606 targeted non-coercion VDCs and 155 coerced CM-VDCs were labeled. Only about 1.4% of the targeted VDCs were tagged as coercion.

Also, if we examine the five most frequent verbs for both non-coerced and coerced VDCs (Table 26), the verb frequencies in non-coerced VDCs are much higher than those in the coerced ones. The non-coerced V1s are typical displacement verbs, such as putting and sending, while the coerced V1s come from different verb classes, such as speech, cognition, and non-caused-motion-causative.

	1st	2nd	3rd	4th	5th
Non-Coerced	*dai* 'take' (163)	*na* 'get/take' (140)	*song* 'send' (140)	*ti* 'lift' (108)	*fang* 'put' (98)
Coerced	*pan* 'expect' (5)	*ban* 'do' (5)	*pai* 'take pictures' (4)	*shuo* 'speak' (2)	*da* 'hit' (2)

Table 26. Five Most Frequent Verbs in Non-coerced and Coerced Categories

The performance of detecting the class of coercion on the different models is reported in Table 27. Since there is an imbalance between the two classes, both neural and SVM algorithms tend to output the more frequent class of the two, namely non-coercion. The F-scores for the SVM models are marginally better than the percentage of coercion in the data (2%>1.4%), which is far less than the results of neural models. In biLSTM models, similar to the event type classification task, there is a short context effect, meaning that information within the VDC's clause contains more important representations. However, within the short context group, segmentation increases the models' performance, whereas segmentation slightly decreases the models' performance for the event type classification. This indicates that VDC event types rely more heavily on meanings with richer distributional granularity, as in the word-and-character level embedding. Finally, the best performance comes from Model 69 (F-score=0.34), which is higher than the percentage of coerced VDCs in the dataset.

Model#	Context	Model	Precision	Recall	F-Score
57	Long	EM0+Char	0.04	0.07	0.05
58		EM0+Seg	0.14	0.09	0.11
59		EM1+Char	0.25	0.1	0.14
60		EM1+Seg	0.13	0.15	0.14
61		EM2+Char	0.22	0.1	0.14
62		EM2+Seg	0.09	0.1	0.09
63		SVM	0.01	1	0.02
64	Short	EM0+Char	0.08	0.07	0.07
65		EM0+Seg	0.2	0.29	0.24
66		EM1+Char	0.3	0.15	0.2
67		EM1+Seg	0.25	0.3	0.25
68		EM2+Char	0.22	0.1	0.14
69		**EM2+Seg**	**0.33**	**0.35**	**0.34**
70		SVM	0.01	1	0.02

Table 27. Coercion Detection Results

5.6 Future Improvements

I suggest four ways to improve performance on the coercion detection task. First, the major challenge is the low resource nature of the coercion. A simple way to increase the amount of data would be to annotate more corpora, but this method is inefficient. Given the creative nature of coercion, the data can be augmented by automatically generating coerced VDCs based on V1 classes and letting human subjects judge their grammaticality. Second, the genre features of the corpus can alter the occurrence of coercion. For instance, CTB contains many tokens from the Xinhua Newswire, which may lack variation in its coercion. CCL has plenty of texts in literature and education, while BCC is mostly Sina Microblog data. The data from CCL and BCC might include more coercion. Therefore, diversifying genres will help us better understand coercion. Third, in order to improve the Google MT performance on cocerced VDC translation, the new classifier developed in Sec.5.5 can be used to collect coercion data. The hard instances can be tested on Google Translate. If their MT results are problematic, then their corrected translation by humans can be directly incorporated into the training of the MT systems. Fourth, coercion detection does not specify argument-related change, such as the extension of direction vs. object augmentation of V1s. Future research will link V1s in the coercion data with computational lexicons, such as VerbNet (Schuler 2005; Hwang 2014) and FrameNet, to provide a more informative representation of argument structure change in coerced VDCs.

Chapter 6
Conclusions

6.1 Summary

Much like MWEs in other languages, Chinese VDCs create parsing challenges for NLP systems in both theory and practice, due to their multifarious semantic constraints. In this work, I have developed a multi-purpose classification system to deal with the three most important semantic domains of VDCs, using both biLSTM and SVM models. The first component of this system was designed and implemented to label different kinds of events of VDCs. There are three subtasks: VDC direction, coarse VDC classification, and fine VDC classification. The models that achieved best performance on these three tasks are Model#12 (biLSTM with short character inputs and pre-trained word-and-character-level embedding), Model#24 (biLSTM with short character level inputs and pre-trained word-level embedding), and Model#36 (biLSTM with short character inputs and randomized embedding), respectively. The second component was developed to distinguish metaphoric and literal VDCs. The best models are #52 (biLSTM with short character level inputs and pre-trained word-level embedding) and #54 (biLSTM with short character level inputs and pre-trained word-and-character-level embedding). The third component was the detection of creative uses of V1s licensed by the meaning of the VDC. Our preliminary study shows that the best model is #69 (biLSTM with short segmented word inputs and pre-trained word-and-character-level embedding). Apart from individual tasks, this project also shows positive effects of character-level segmentation and short context on the system performance. The use of the pre-trained embedding is not clear, which is possibly confounded by segmentation and context length.

6.2 Contribution and Future Improvements on All Tasks

This research benefits from NLP and linguistics. My multi-purpose classification system can improve NLP in both theory and application. First, Construction parsing is rarely done in Chinese NLP. Second, from the perspective of building a machine learning agent that can achieve a level of natural language understanding of Chinese, the performance of the three components in my system show that linguistic concepts like event structures, metaphors, and creative use in VDCs are highly attainable in supervised learning. Third, the success of my multi-purpose classification system can help improve downstream NLP applications in three ways: (1) by classifying different event types to enable a robotic agent to distinguish Self-Motion from Aspectual VDCs in ambiguous human voice inputs, (2) by detecting metaphoric VDCs as the first step toward deep semantic parsing of metaphoric VDCs, and by detecting concrete VDCs to help the machine match arguments to images, and (3) by identifying coercion tokens and providing their human translation as data back into the training of MT systems, such as Google Translate.

From a linguistic perspective, I have applied both CxG and CMT to the VDC data and then made adjustments that can, in turn, enrich the theories. In the preparation of the VDC event type classification, the coarse-grained and fine-grained taxonomies include both phrasal and formulaic idiomatic VDCs, which verify the basic assumptions of Construction Grammar; I also defined classes of VDCs in the formalism of SBCG, which revealed the close relationship between VDCs and the class of control predicators. Similar analyses can be applied to other types of SVCs. In the metaphor analysis of VDCs, I classified the metaphors according to the contribution of both verbs in VDCs. This indicated that the Source-Target assignments can be non-linear in Chinese SVCs. Finally, when using Construction Grammar to analyze the argument change in coerced VDCs, I argued that the valence augmentation process is dependent on argument sharing and a particular verb-construction relation (the "means" relation identified by Goldberg 1995 for English Caused Motion predications).

Although this project shows that the NLP system can achieve a significant degree of human understanding in parsing different aspects of VDCs, there is room for improvement. In addition to the improvement suggestions given at the end of each classification chapter (Sec.3.5.3, Sec.4.4.2, and Sec.5.6), in future research I plan to improve the classification system in several directions. First, more neural network techniques require experimentation. The nature of RNN-based architecture in this research is still sequential processing. It is important to focus on other parts of the context instead of on the last element in the input by using attention-based mechanism, such as self-attention (Vaswani et al. 2017). Newly developed architectures and resources that have proved successful in various NLP tasks, such as Bidirectional Encoder Representations from Transformers) (BERT) (Devlin et al. 2018), XLNET (Yang et al. 2019), and ERNIE 2.0 (Sun et al. 2019), may achieve better performance in some of the classification tasks developed in this research by creating more contextualized embeddings. Additionally, more embedding resources can be retrieved from linguistic features or parses, such as dependency parses (Chen & Manning 2014; Levy & Goldberg 2014), WordNet (Saedi et al. 2018), and SRL (Strubell et al. 2018). Second, this research assumes the independence of the three tasks, so no transfer learning has been applied (Ruder et al. 2019). For example, during the annotation process, a conceptual connection was found between the labeling of metaphors and event types. This indicates that high performance in one of the three classification tasks may cause good performance in the other two, such as the role of metaphor processing in coercion detection.

Bibliography

Anderson, S. R. (1992). *A-morphous morphology* (Vol. 62). Cambridge University Press.

Audring, J., & Booij, G. (2016). Cooperation and coercion. *Linguistics*, 54(4), 617-637.

Baker, C. F., Fillmore, C. J., & Lowe, J. B. (1998, August). The Berkeley Framenet project. In Proceedings of the 17th international conference on Computational linguistics-Volume 1 (pp. 86-90). Association for Computational Linguistics.

Baldwin, T., & Kim, S. N. (2010). Multiword expressions. *Handbook of natural language processing*, 2, 267-292.

Bergen, B., & Chang, N. (2005). Embodied construction grammar in simulation-based language understanding. *Construction grammars: Cognitive grounding and theoretical extensions*, 3, 147-190.

Black, M. (1962). *Models and metaphors*: *Studies in language and philosophy*.

Bonial, C., Brown, S. W., Hwang, J. D., Parisien, C., Palmer, M., & Stevenson, S. (2011, June). Incorporating coercive constructions into a verb lexicon. In *Proceedings of the ACL 2011 Workshop on Relational Models of Semantics* (pp. 72-80). Association for Computational Linguistics.

Booij, G. (2009). Lexical integrity as a formal universal: a constructionist view. In Universals of language Today, 83-100. Springer, Dordrecht.

Bybee, J. L., & Dahl, Ö. (1989). *The creation of tense and aspect systems in the languages of the world*. Amsterdam: John Benjamins.

Chao, Y. R. (1948). *Mandarin primer: An intensive course in spoken Chinese*. Harvard University Press.

Chao, Y. R. (1968). A Grammar of Spoken Chinese. University of California Press.

Chen, D., & Manning, C. (2014, October). A fast and accurate dependency parser using neural networks. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 740-750).

Chen, W. T., Bonial, C., & Palmer, M. (2015, January). English Light Verb Construction Identification Using Lexical Knowledge. In *AAAI* (pp. 2368-2374).

Cheng, L., & Huang, C.-T. J. (1994). On the argument structure of resultative compounds. In Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language Change, ed. by Matthew Y. Chen and Ovid J. L. Tzeng (pp. 187-221). Taipei: Pyramid Press.

Cheng, Q. (2018). Family of Chinese BA-construction-A Cognitive Approach (Doctoral dissertation, Northumbria University).

Chomsky, N. (1995). *The minimalist program* (Vol. 1765). Cambridge, MA: MIT press.

Chung, K. S. (2006). *Mandarin compound verbs*. Crane publishing.

Cook, P., & Stevenson, S. (2006, July). Classifying particle semantics in English verb-particle constructions. In *Proceedings of the Workshop on Multiword Expressions: Identifying and Exploiting Underlying Properties* (pp. 45-53). Association for Computational Linguistics.

Cortes, C., & Vapnik, V. (1995). Support vector machine. *Machine learning*, 20(3), 273-297.

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press on Demand.

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge University Press.

David, O., Dodge, E., Hong, J., Stickles, E. L. I. S. E., & Sweetser, E. (2014). Building the MetaNet metaphor repository: The natural symbiosis of metaphor analysis and construction grammar. In *The 8th International Conference on Construction Grammar* (ICCG 8), Osnabrück, Germany.

De Saint-Exupéry, A. (2015). *Le petit prince: avec des aquarelles de l'auteur*. Ernst Klett Sprachen GmbH.

De Swart, H. (1998). Aspect shift and coercion. *Natural Language & Linguistic Theory*, 16(2), 347-385.

Deng, L.; Yu, D. (2014). "Deep Learning: Methods and Applications." *Foundations and Trends in Signal Processing*. 7 (3–4): 1–199.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Ding, P. S. (2000). A Computational Study of the Ba Resultative Construction: Parsing Mandarin Ba Sentences in HPSG. In Proceedings of the 14th Pacific Asia Conference on Language, Information and Computation, 67-78.

Dippner, K. (2010). *Metaphorical use of Mandarin compound directional complements*. (unpublished master thesis). Ohio University.

Doddington, G. R., Mitchell, A., Przybocki, M. A., Ramshaw, L. A., Strassel, S., & Weischedel, R. M. (2004, May). *The Automatic Content Extraction (ACE) Program-Tasks, Data, and Evaluation*. In LREC (Vol. 2, p. 1).

Dodge, E. K., & Petruck, M. R. (2014). Representing caused motion in embodied construction grammar. In *Proceedings of the ACL 2014 workshop on semantic parsing* (pp. 39-44).

Fillmore, C. J., & Kay, P. (1993). *Construction grammar coursebook*: chapters 1 thru 11. University of California.

Fillmore, C. J., Kay, P., & O'connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, 501-538.

Fillmore, C. J., Lee-Goldman, R., & Rhodes, R. (2012). The framenet constructicon. *Sign-based construction grammar*, 309-372.

Fong, R. (2015). The Chinese ba as a Verb: A Constructional-Cognitive Approach. *Open Journal of Modern Linguistics*, 5(01), 48.

Fong, T., Thorpe, C., & Baur, C. (2001). Collaboration, dialogue and human-robot interaction, 10th international sumposium of robotics research (lorne, victoria, australia). In *Proceedings of the 10th International Symposium of Robotics Research*.

Fu, Y., Ge, N., Zheng, Z., Zhang, S., Meng, Y., & Yu, H. (2012, September). Extraction of chinese multiword expressions based on artificial neural network with feedbacks. In *Pacific Rim International Conference on Artificial Intelligence* (pp. 734-739). Springer, Berlin, Heidelberg.

Gao, Q. 2000. Argument Structure, HPSG and Chinese Grammar. PhD thesis, Ohio State University.

Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W.-k. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), APA decade of behavior series. *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin* (pp. 151-175). Washington, DC, US: American Psychological Association.

Ginzburg, J., & Sag, I. (2000). Interrogative investigations. Stanford: CSLI publications.

Goldberg, A. E. (1995). Constructions: *A construction grammar approach to argument structure*. University of Chicago Press.

Goldberg, A. E. (1999). The emergence of argument structure semantics. *The Emergence of Language* (MacWhinney, B., ed.), 197-212.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

Goldberg, A. E., & Jackendoff, R. (2004). The English resultative as a family of constructions. *Language*, 532-568.

Goldberg, Y. (2017). *Neural network methods for natural language processing*. Synthesis Lectures on Human Language Technologies, 10(1), 1-309.

Gupta, S., & Malik, J. (2015). Visual semantic role labeling. *arXiv preprint arXiv:1505.04474*.

Haspelmath, M. (2016). The serial verb construction: Comparative concept and cross-linguistic generalizations. *Language and Linguistics*, 17(3), 291-319.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.

Hsiao, H. S., & Lin, C. (2012). On Mandarin Potential Complement Constructions. *Language and Linguistics*, 13(5), 963.

Hsieh, A. Y. (2008). *Semantic Extension from Perception-Cognition-Utterance to Modality and Manipulation Verbs in Mandarin: a Frame-based Account*. Master's Thesis. National Chiao Tung University.

Huang, C. J. (1989). Pro-drop in Chinese: A generalized control theory. In *The null subject parameter* (pp. 185-214). Springer, Dordrecht.

Huang, C. T. J. (1992). Complex predicates in control. In *Control and grammar* (pp. 109-147). Springer Netherlands.

Huang, J. R. (Ed.). (2016). *A reference grammar of Chinese*. Cambridge University Press.

Huang, J. R., Hsieh, S. K., Hong, J. F., Chen, Y. Z., Su, I. L., Chen, Y. X., & Huang, S. W. (2010). Chinese Wordnet: Design, implementation, and application of an infrastructure for cross-lingual knowledge processing. *Journal of Chinese Information Processing*, 24(2), 14-23.

Hwang, J. (2014). *Identification and Representation of Caused Motion Constructions*. Doctoral Dissertation. University of Colorado Boulder.

Jackendoff, R., & Syntax, X. B. (1977). A study of phrase structure. *Linguistic Inquiry Monographs*, 2.

John R. Firth. A synopsis of linguistic theory 1930–1955. (1957) In *Studies in Linguistic Analysis, Special volume of the Philological Society*, 1–32. Firth, John Rupert, Haas William, Halliday, Michael A. K., Oxford, Blackwell Ed.

Jolliffe, I. (2011). *Principal component analysis*. Springer Berlin Heidelberg.

Kaplan, R. M., & Bresnan, J. (1982). Lexical-functional grammar: A formal system for grammatical representation. Formal Issues in Lexical-Functional Grammar, (47), 29-130.

Kay, P., & Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: the What's X doing Y? construction. *Language*, 1-33.

Kim J-B., Michaelis, L. forthcoming. *Syntactic Constructions in English*. Cambridge: Cambridge Press.

Koontz, D. (2007). *Seize the Night: A Novel* (Vol. 2). Bantam.

Kövecses, Z. (2005). Metaphor in culture: Universality and variation. Cambridge University Press.

Lakoff, G. (1990). The invariance hypothesis: Is abstract reason based on image-schemas?. *Cognitive Linguistics* (includes Cognitive Linguistic Bibliography), 1(1), 39-74.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (pp. 202-251). New York, NY, US: Cambridge University Press.

Lakoff, G. (1993). The syntax of metaphorical semantic roles. In *Semantics and the Lexicon* (pp. 27-36). Springer, Dordrecht.

Lakoff, G. (2008). Women, fire, and dangerous things. University of Chicago press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic books.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford university press.

Levy, O., & Goldberg, Y. (2014, June). Dependency-based word embeddings. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers) (pp. 302-308). Li, S., Zhao, Z., Hu, R., Li, W., Liu, T., & Du, X. (2018). Analogical reasoning on chinese morphological and semantic relations. arXiv preprint arXiv:1805.06504.

Li, B., Wen, Y., Weiguang, Q. U., Bu, L., & Xue, N. (2016). Annotating the little prince with Chinese AMRs. In *Proceedings of the 10th Linguistic Annotation Workshop held in conjunction with ACL 2016 (LAW-X 2016)*, (pp. 7-15).

Li, C. N., & Thompson, S. A. (1973). Serial verb constructions in Mandarin Chinese: Subordination or coordination. *You take the high node and I'll take the low node: Papers from the Comparative Syntax Festival*, 96-103.

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference gr*ammar. University of California Press.

Li, S., Zhao, Z., Hu, R., Li, W., Liu, T., & Du, X. (2018). Analogical Reasoning on Chinese Morphological and Semantic Relations. arXiv preprint arXiv:1805.06504.

Li, Y. (1990). On VV compounds in Chinese. Natural language & linguistic theory, 8(2), 177-207.

Li, Y.A. (2016). Verbs and verb phrases. In Huang, J.R & Shi, D.X (Eds.), *A reference grammar of Chinese*. Cambridge University Press.

Liang, Y., Tan, H., Li, H., Wang, Z., & Gui, W. (2017, May). A language-independent hybrid approach for multi-word expression extraction. In *Neural Networks* (IJCNN), 2017 International Joint Conference on (pp. 3273-3279). IEEE.

Lin, T. (2018). On the semantics and syntax of Verb-Direction Construction in Mandarin Chinese. In *The 10th International Conference on Construction Grammar* (ICCG 10), Paris, France.

Lindner, S. 1983. *A lexico-semantic analysis of English verb particle constructions with OUT and UP*. Doctoral dissertation, University of Indiana, Bloomington.

Liu, M. (2015). Tense and Aspect in Mandarin Chinese. The Oxford Handbook of Chinese Linguistics, 274.

Lu, X., & Wang, B. P. Y. (2017). Towards a metaphor-annotated corpus of Mandarin Chinese. *Language Resources and Evaluation*, *51*(3), 663-694.

Lv, X., Zhang, M., & Li, H. (2008, September). Robot control based on voice command. In *Automation and Logistics, 2008. ICAL 2008. IEEE International Conference on* (pp. 2490-2494). IEEE.

Machine Learning. (n.d.). In Wikipedia. Retrieved December 19, 2018, from https://en.wikipedia.org/wiki/Machine_learning

McEnery, A., & Xiao, Z. (2004). The Lancaster Corpus of Mandarin Chinese: A corpus for monolingual and contrastive language study. *Religion*, 17, 3-4.

Michaelis, L. A. (2003). Headless constructions and coercion by construction. *Mismatch: Form-function incongruity and the architecture of grammar*, 259-310.

Michaelis, L. A. (2004). Type shifting in construction grammar: An integrated approach to aspectual coercion. *Cognitive linguistics*, 15(1), 1-68.

Michaelis, L. A. (2005). Entity and event coercion in a symbolic theory of syntax. *Construction Grammars: Cognitive grounding and theoretical extensions*, 45-88.

Michaelis, L. A. (2006a). Complementation by construction. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 32, No. 1, pp. 529-549).

Michaelis, L. A. (2006b). Construction Grammar. In K. Brown (ed.), *The Encyclopedia of Language and Linguistics*, Second edition, volume 3, 73-84.

Michaelis, L. A. (2009). Sign-Based Construction Grammar. In B. Heine and H. Narrog, (eds.), *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press, 155-176.

Michaelis, L. A. (2013). Sign-based construction grammar. *The Oxford handbook of construction grammar*. Oxford: Oxford University Press, 133-152.

Michaelis, L. A. (2015). Constructions License Verb Frames. In *Perspectives on Complementation* (pp. 7-33). Palgrave Macmillan, London.

Michaelis, L. A., & Lambrecht, K. (1996). Toward a construction-based theory of language function: The case of nominal extraposition. *Language*, 215-247.

Michaelis, L. A., & Ruppenhofer, J. (2001). *Beyond Alternations: A Constructional Model of the German Applicative*. Stanford: CSLI Publications.

Michaelis, L. forthcoming. Constructions are Patterns and so are Fixed Expressions. In B. Busse and R. Moehlig, (eds.), *Patterns in Language and Linguistics*. Berlin: Mouton de Gruyter. (Topics in English Linguistics series)

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. & Dean, J., (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, 3111-3119.

Morgan, Pamela S. 1997. Figuring out figure out. Metaphor and the semantics of the English verb-particle construction. *Cognitive Linguistics (includes Cognitive Linguistic Bibliography)* 8, no. 4: 327-358.

Müller, S. & Lipenkova, J. (2009). Serial verb constructions in Chinese: An HPSG Account. In *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar*, University of Göttingen, Germany, Stanford: CSLI Publications, 234–254.

Nikolaidis, S., Nath, S., Procaccia, A. D., & Srinivasa, S. (2017, March). Game-theoretic modeling of human adaptation in human-robot collaboration. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction* (pp. 323-331). ACM.

Palmer, M., Gildea, D., & Xue, N. (2010). Semantic role labeling. *Synthesis Lectures on Human Language Technologies*, 3(1), 1-103.

Paul, W. (2008). The serial verb construction in Chinese: A tenacious myth and a Gordian knot. *The Linguistic Review*, 25(3-4), 367-411.

Pennington, J., Socher, R., and Manning, C.D. (2014). *GloVe: Global Vectors for Word Representation*.

Piao, S., Sun, G., Rayson, P., & Yuan, Q. (2006). Automatic extraction of Chinese multiword expressions with a statistical tool. In *Workshop on Multi-word-expressions in a Multilingual Context held in conjunction with the 11th Conference of the European Chapter of the Association for Computational Linguistics* (EACL 2006).

Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. University of Chicago Press.

Pustejovsky, J. (1991). The generative lexicon. *Computational linguistics*, 17(4), 409-441.

Pustejovsky, J. (2007). Fundamentals of generative lexicon. Presented at Linguistics Summer Institute, Linguistic Society of America, Stanford. Retrieved from: http://timeml.org/jamesp/LSA-2007-L1.pdf

Pustejovsky, J., & Jezek, E. (2008). Semantic coercion in language: Beyond distributional analysis. *Italian Journal of Linguistics*, 20(1), 175-208.

Pustejovsky, J., & Stubbs, A. (2012). *Natural Language Annotation for Machine Learning: A guide to corpus-building for applications.* O'Reilly Media, Inc.

Qi, P., Dozat, T., Zhang, Y., & Manning, C. D. (2019). Universal Dependency Parsing from Scratch. arXiv preprint arXiv:1901.10457.

Rappaport Hovav, M. and B. Levin (1998) 'Building Verb Meanings' in M. Butt and W. Geuder (eds). *The Projection of Arguments* (Stanford: CSLI Publications), pp.97–134.201

Ross, C., & Ma, J. H. S. (2014). Modern Mandarin Chinese grammar: A practical guide. Routledge Press.

Ruder, S., Peters, M. E., Swayamdipta, S., & Wolf, T. (2019, June). Transfer Learning in Natural Language Processing. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Tutorials* (pp. 15-18).

Ruppenhofer, J., Ellsworth, M., Petruck, M. R., Johnson, C. R., & Scheffczyk, J. (2016). FrameNet II: Extended theory and practice. Institut für Deutsche Sprache, Bibliothek.

Saedi, C., Branco, A., Rodrigues, J. A., & Silva, J. (2018, July). Wordnet embeddings. In *Proceedings of The Third Workshop on Representation Learning for NLP* (pp. 122-131).

Sag, I. A. (2012). Sign-based construction grammar: An informal synopsis. *Sign-based construction grammar*, 193, 69-202.

Sag, I. A., Baldwin, T., Bond, F., Copestake, A., & Flickinger, D. (2002). Multiword expressions: A pain in the neck for NLP. In *Computational Linguistics and Intelligent Text Processing* (pp. 1-15). Springer Berlin Heidelberg.

Schuler, K. K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon*. (Doctoral Dissertation), University of Pennsylvania, Philadelphia, PA, USA

Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing, 45(11), 2673-2681.

Searle, J. R. (1993). Metaphor. In *Metaphor and Thought*, (Ortony, A., ed.), pp. 83-111. Cambridge University Press Cambridge.

Shutova, E. (2015). Design and evaluation of metaphor processing systems. *Computational Linguistics*, 41(4), 579-623.

Sikos, L., Brown, S. W., Kim, A. E., Michaelis, L. A., & Palmer, M. (2008). Figurative Language:" Meaning" is Often More than Just a Sum of the Parts. In *AAAI Fall Symposium: Biologically Inspired Cognitive Architectures* (pp. 180-185).

Skubic, M., Perzanowski, D., Blisard, S., Schultz, A., Adams, W., Bugajska, M., & Brock, D. (2004). Spatial language for human-robot dialogs. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *34*(2), 154-167.

Song, S., Bo, C., Sio, U., & Bond, F. (2015). An HPSG-based Analysis of Resultative Compounds in Chinese *Proceedings of the 22nd International Conference on Head-Driven Phrase Structure Grammar*, 96-109.

Strubell, E., Verga, P., Andor, D., Weiss, D., & McCallum, A. (2018). Linguistically-informed self-attention for semantic role labeling. arXiv preprint arXiv:1804.08199.

Sullivan, K. (2013). *Frames and constructions in metaphoric language* (Vol. 14). John Benjamins Publishing.

Sun, Y., Wang, S., Li, Y., Feng, S., Tian, H., Wu, H., Wang, H. (2019). ERNIE 2.0: A Continual Pre-training Framework for Language Understanding. arXiv preprint arXiv:1907.12412

Sybesma, R. (1999). *The Mandarin VP*. Dordrecht: Kluwer.

Talmy, L. (1972). Semantic structures in English and Atsugewi. Ph.D dissertation, Department of Linguistics, University of California, Berkeley.

Thompson, S. A. (1973). Resultative verb compounds in Mandarin Chinese: A case for lexical rules. *Language*, 361-379.

Van Valin Jr, R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning, and function*. Cambridge University Press.

Van Valin, R. D. (2009). Role and reference grammar. Grammar, *Meaning and Pragmatics*, 5, 239.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998-6008).

Vendler, Z. (1957). Verbs and times. The philosophical review, 66(2), 143-160.

Wang, L., & Liu, R. (2011, September). A rapid method to extract multiword expressions with statistic measures and linguistic rules. In *International Conference on Web Information Systems and Mining* (pp. 234-241). Springer Berlin Heidelberg.

Wang, L., & Yu, S. (2010). Construction of Chinese Idiom Knowledge-base and Its Applications. In *Proceedings of the 2010 Workshop on Multiword Expressions: from Theory to Applications* (pp. 11-18).

Wang, X., Iwasawa, S., Miyao, Y., Matsuzaki, T., Yu, K., & Tsujii, J. I. (2009, December). Design of chinese hpsg framework for data-driven parsing. In *Proceedings of the 23rd Pacific Asia Conference on Language, Information and Computation*, Volume 2 (pp. 835-842).

Weinberger, K., Dasgupta, A., Attenberg, J., Langford, J., & Smola, A. (2009). Feature hashing for large scale multitask learning. arXiv preprint arXiv:0902.2206.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Klingner, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.

Xue, N., & Palmer, M. (2009). Adding semantic roles to the Chinese Treebank. *Natural Language Engineering*, 15(01), 143-172.

Xue, N., Xia, F., Huang, S., & Kroch, A. (2000). *The bracketing guidelines for the Penn Chinese Treebank (3.0).*

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. V. (2019). XLNet: Generalized Autoregressive Pretraining for Language Understanding. arXiv preprint arXiv:1906.08237.

Yin, H. (2011). The cognitive semantics of Chinese motion/directional verbs. *Working Papers of the Linguistics Circle*, 21(1), 118-125.

Yip, P. C., & Rimmington, D. (2009). *Basic Chinese: a grammar and workbook*. Taylor & Francis.

You, L., Liu, T., & Liu, K. (2007, August). Chinese FrameNet and OWL representation. In *Advanced Language Processing and Web Information Technology*, 2007. ALPIT 2007. Sixth International Conference on (pp. 140-145). IEEE.

Yu, N. (2012). The metaphorical orientation of time in Chinese. *Journal of Pragmatics*, 44(10), 1335-1354.

Zhan, W., Guo, R., Chen, Y., (2003). *The CCL Corpus of Chinese Texts: 700 million Chinese Characters, the 11th Century B.C. - present*, Available online at the website of Center for Chinese Linguistics (abbreviated as CCL) of Peking University, http://ccl.pku.edu.cn:8080/ccl_corpus

Zhu, Y., Liu, Y., Peng, S., Blodgett, A., Zhao, Y., & Schneider, N. (2018). Adpositional Supersenses for Mandarin Chinese. arXiv preprint arXiv:1812.02317.

陈昌来. (Chen, C.) (1994). 论动后趋向动词的性质—–兼谈趋向动词研究的方法. 鲁东大学学报: 哲学社会科学版, (4), 64-70. (The properties of postverbal directional verbs)

陈贤. (Chen, X) (2007). 现代汉语动词 "来" ,"去" 的语义研究 (Doctoral Dissertation), 上海: 复旦大学) (On the verbs *lai* and *qu* in Modern Chinese)

范晓. (Fan, X.) (1981). 怎样区别现代汉语的词同短语. 东岳论丛, (4), 104-111. (How to tell the difference between words and phrases in Chinese)

房玉清. (Fang, Y). (1992). "起来" 的分布和语义特征. 世界汉语教学, (1), 23-28. (On the semantics and distribution of qilai)

龚千炎. (Gong, Q). (1995). 汉语的时相时制时体. 北京:商务印书馆. (On the temporal notions in Chinese)

黄伯荣, & 廖序东. (Huang ,B., & Liao, X.) (1991). 现代汉语. (下册) 高等教育出版社. (Modern Chinese II)

江蓝生. (Jiang, L.) (2012). 汉语连-介词的来源及其语法化的路径和类型. 中国语文, (4), 291-308. (Sources, paths, and types of the grammaticalization of Chinese conjunction-prepositions)

居红. (Ju, H). (1992). 汉语趋向动词及动趋短语的语义和语法特点 (On the semantic and grammatical properties of directional verbs and verb-direction phrases in Chinese). 世界汉语教学, (4), 276-282.

黎锦熙, & 刘世儒. (Li, J, & Liu, S) (1957). 汉语语法教材. 北京:商务印书馆. (Chinese grammar)

黎锦熙. (Li, J). (1924). 新著国文法. 北京:商务印书馆. (A new Chinese grammar)

李德津. (Li, D) (2008). 外国人实用汉语语法: 中英文对照. 北京语言大学出版社. (A practical Chinese grammar for foreigners)

栗爽 (Li, S). (2008). 现代汉语位移动词研究. Master's Thesis. 上海: 上海师范大学. (A study on motion verbs in modern Chinese)

刘叔新. (Liu, S.) (1985). 试论趋向范畴. 载< 语法研究和探索 (三)>, 北京大学出版社. (On directional category)

刘月华 (Liu, Y) (Ed.). (1998). 趋向补语通释 (A general explanation on Chinese verb-direction constructions). 北京语言文化大学出版社.

刘月华, 故韡, & 潘文娱. (Liu, Y., Gu, W., Pan, W.) (2001). 实用现代汉语语法. 商务印书馆. (A practical modern Chinese grammar)

陆志韦. (1956). (Lu, Z.) 北京话单音词词汇. 科学出版社. (Monosyllabic lexicons in Beijing Mandarin)

陆宗达, & 俞敏. (Lu, Z., & Yu, M.) (1954). 现代汉语语法. 北京: 群众书店. (Modern Chinese grammar)

骆蓉 (Luo, R). 认知构式语法视阈下的致使移动句研究. (Doctoral Dissertation). 浙江大学, 2015. (A cognitive construction grammar approach to caused-motion sentences)

吕叔湘. (Lü, S.) (1979). 汉语语法分析问题. 商务印书馆. (Questions on grammatical analysis in Chinese)

齐沪扬 (Qi, H).(2002). 现代汉语短语.上海：华东师范大学出版社. (Modern Chinese phrases)

王力. (Wang, L).(1943). 中国现代语法. 中華書局. (Modern Chinese grammar)

王涛 (Wang, T). (1987).中国成语大辞典 上海辞书出版社(Great dictionary of Chinese *chengyu*).

王宜广 (Wang, Y.). (2013). 现代汉语动趋式语义研究述评. 汉语学习, 3. (A review on the semantic studies of VDCs in Modern Chinese)

王小波 (Wang, X). (2006). 王小波全集. (Vol. 7). 云南人民出版社. (Complete work of Xiaobo Wang).

邢福义, 吴振国, & 汪国胜 (Eds.). (Xing, F., Wu, Z., & Wang, G.) (2003). 现代汉语 (下). 高等教育出版社. (Modern Chinese II)

徐静茜. (Xu, J.) (1983). “趋向动词” 应归属何种词类?. 湖州师范学院学报, (2), 52-59. (What lexical category do “directional verbs” belong to)

荀恩东, 饶高琦, 肖晓悦, 臧娇娇 (Xun, E., Rao, G., Xiao, X., Zang, J.). (2016). 大数据背景下 BCC 语料库的研制. 语料库语言学, (1). (The construction of the BCC corpus in the age of big data)

杨尔弘, 方莹, 刘冬明, & 乔羽 (Yang E., Fang, Y. Liu, D., Qiao, Y.). (2006). 汉语自动分词和词性标注评测 中文信息学报, 20(1), 46-51. (The evaluation of Chinese word segmentation and POS tagging).

于辉荣(Yu, R). (2010). “V 上” 及其相关问题研究 (Master's thesis) 上海师范大学. (On the “Verb+*shang*” structure and relevant issues)

岳中奇. (Yue, Z) (1996). 体助词 “去” 的语法意义及其相关动词. 语文学刊: 高等教育版, (3), 38-40. (The grammar of auxiliary *qu* and related verbs)

张志公. (Zhang, Z). (1953). 汉语语法常识. 中国青年出版社. (Common knowledge in Chinese grammar)

张志公. (Zhang, Z). (1957).语法和语法教学: 介绍 “暂拟汉语教学语法系统”. 人民教育出版社. (Introducing the “Temporal Teaching Grammar for Chinese”)

周一民. (Zhou, Y.) (1999). 汉语趋向动词规范谈. 语文建设, (3), 32-34. (On the norms of Chinese directional verbs)

朱德熙. (Zhu, D.) (1982). 语法讲义. 商务印书馆. (A lecture of Chinese grammar).

Appendix 1: Debates on the Constituency Analyses of DPs and VDCs

In Chinese linguistics, two major issues of contention with respect to VDCs are identifying the POS of Directional Particles (DPs) and the constituent type of VDCs. Given the complexity of VDCs, "what are DPs/VDCs" remains a challenging question. These two issues are related: the more functional and bound DPs are, the more likely that VDCs are words. For example, if DPs are thought as suffixes, then VDCs should be treated as words. Although serial verb construction (SVC) is a debatable grammatical category among Chinese linguists (Chao 1948; Li and Thompson 1973, 1981; Paul 2008), like many other languages, Chinese SVC can be briefly described as "a monoclausal construction consisting of multiple independent verbs" with no element linking them and with argument sharing among the verbs (Haspelmath 2006: 292). Treating VDCs as SVCs (Liu et al. 1998) is strongly reasonable because DPs can semantically and syntactically justify themselves as verbs and because most VDCs do not follow the Principle of Lexical Integrity (Anderson 1992).

In the first issue, there are two approaches to analyzing the POS of DPs: one or multiple tags. Furthermore, within each approach there are different opinions. Within the first approach, DPs can be viewed as Auxiliary verbs (Xu 1983; Fang 1992; Yue 1996). This view emphasizes the similarity between the features of the aspectual markers *zhe* (progressive) and *le* (perfect) and those of DPs. For example, both DPs and *zhe/le* (a) are post-verbal, (b) receive a neutral tone, (c) form a closed class, (d) are productive, (e) give additive and non-independent meanings (e.g., state of events) to the main verbs, and (f) have a historical origin as verbs (Xu 1983: 59). Using a between-category comparison, the auxiliary approach highlights the fact that some DPs fall into the category of functional words, but there are several weaknesses. First, (a), (c), (d), and (f) can be shared by other function words in Chinese, such as *gei*-class prepositions 'to/for' (Jiang 2012: 300). Second, many DPs do not have a neutral tone, such as *zou3 lai2* (lit ('walk come') 'walk toward') and *ji4 qi3* (lit ('remember rise') 'come to one's memory') (Zhou 1999: 33). Third, aspectual *zhe* and *le* cannot be independent verbs, but some DPs can, which indicates that the semantics of DPs is not always additive or dependent.

A similar opinion is to label DPs as suffixes (Ross & Ma 2014). The problem is that this analysis ignores DPs' more concrete meanings, such as directional and resultative. For example, one can argue that in the VDC *tuo xia lai yi jian yifu* (lit ('take descend come one CL clothes') 'take off a piece of clothing') there is still a reading that the piece of clothing moves off one's body. Additionally, an object can be placed between V1 and DP (see Example (2b)), which violates the definition of suffixes.

Another viewpoint is to tag DPs as adverbs (Lu & Yu 1954; Lu 1956). The main reason is that DPs modify the main verb just as adverbs do (Lu 1956: 45). However, not only adverbs can modify verbs; aspectual markers function similarly. Also, this claim overlooks the connection between DPs and objects. For example, in the VDC *ta zou jin wu* (lit (he walk enter house) 'he walked into the house'), the DP introduces the locative object, just like the preposition *into* in English.

Yet another solution is to put all DPs in the category of "directional verbs" and explain the non-directional VDCs as special usages of directional verbs (Zhang 1957). This taxonomy is taken as the official Mandarin grammar in middle school education. The advantage is that students can easily start learning from the most basic usage of DPs (directional) and expand it to others. Although some Chinese linguists think that the special usages of DPs are not verbs and that "usage" does not serve as a POS tag, from the perspective of Conceptual Metaphor Theory, those special usages are different metaphorical interpretations of direction. More justification will be given after the introduction of other approaches.

Alternatively, many linguists choose to categorize DPs as a kind of "directional complement" (*quxiang buyu 趋向补语*) in their grammatical analyses (Chao 1948; Zhu 1982; Huang & Liao 1991; Liu et al. 1998; Xing 2003 et al; Li 2008; Li 2016). A "complement" is an umbrella term that roughly means post-verbal non-object elements complementing the meaning of a verb, such as direction, duration, location, and result (Yip & Rimmington 2009: 214). These elements are syntactically mapped to units like adjectives, numerical phrases, and prepositional phrases. The nature of this "complement" is purely semantic[23], and must not be confused with the complement in modern syntax, such as X-bar theory (Jackendoff 1977). This influential choice cannot be a candidate for the POS of DPs.

One common shortcoming among many syntactic claims outlined in the first approach is that neither the tags of function words nor those of content words can include both concrete and abstract DPs. Therefore, it is more comprehensive to give multiple tags to DPs. Some possible solutions include: (a) *jieci* (*介词*) 'introducing words' for directional DPs and auxiliary for aspectual DPs (Li 1924), (b) auxiliary verbs for directional DPs and special usages of auxiliary verbs for other DPs (Zhang 1953), (c) directional verbs, postverbal resultative components, and temporal/aspectual auxiliaries (Chen 1994), and (d) verbs (VA/VC) for directional and resultative DPs, and adverb (Di) for aspectual and metalinguistic DPs in CWN (Huang et al. 2010). (d) has clear advantages over (a), (b), and (c), since it avoids using vague terms like "introducing words," "special usages," or "components." More importantly, all senses of DPs are mapped to certain syntactic categories.

[23] Similarly, Liu (1985) uses *趋向范畴 quxiang fanchou* 'directional category' for DPs, but this term cannot be an appropriate POS tag.

Although the multi-tag approach can give an accurate POS taxonomy for different DPs, this research agrees with the view to treat VDCs as (directional) verbs (Zhang 1957). One underlying assumption of giving multiple tags to DPs is that DPs are homonyms (Chen 1994: 68). However, the senses of DPs are not discrete or unrelated. In the polysemous network of DPs, the abstract meanings form upon the concrete ones as metaphoric extensions (Sec.4.2). For example, the inchoative *qi* 'rise' comes from the metaphor BEGINNING IS RISING. Additionally, in certain contexts or constraints, some abstract DPs can be syntactically independent like the directional DPs given below:

(1) a. youpiao <u>xia</u> <u>lai</u> le [24]

stamp descend come ASP

"The stamp came off (the envelope)."

b. jintian <u>qi</u> feng le

today rise/ASP wind ASP

"The wind started today." (cf. Baidu Search)

In the examples above, the DPs *xia* 'descend' and *qi* 'rise' act as independent detachment verb and inchoative verb, respectively. Although the independent use in abstract DPs is limited and constrained, it shows that some non-directional DPs still maintain their properties as verbs. The grammaticalization of DPs can be seen as the decategorization of directional verbs. Furthermore, directional, resultative, and aspectual DPs share arguments with their preceding

[24] Retrieved from: http://blog.sina.com.cn/s/blog_713f7f690100o5h4.html

verbs differently in VDCs, just like control/raising verbs (Sec.3.1). Treating DPs as verbs with different argument structures can reveal semantic distinctions in different linguistic applications.

Given the disagreement on the POS of DPs, the second issue of debate is on how to analyze the constituent type of VDCs. Although some researchers try to avoid the controversy by using "verb-direction structures/combinations," the main disagreement is on whether VDCs are compound words or phrases. Those who view VDCs as words define VDCs as a syntactic word morphologically made up of two verbal elements (Chao 1948 & 1968; Thompson 1973; Li & Thompson 1981). The major problem for the compound analysis is that there is little evidence based on linguistic standards that distinguishes between compounds and phrases.

An alternative view is phrase analysis (Liu et al. 2001; Huang & Liao 1991; Xing et al. 2003). According to the Principle of Lexical Integrity, if VDCs are compound words, syntax has no influence on their internal structures. That being said, VDCs do not show any **syntactic interruptability or semantic accessibility** (Booij 2009: 1). However, violations of these two properties can be found:

(2) a. Insertion of potential particles

wo chuan de / bu shang yifu

I wear DE / NEG descend clothes

"I can/cannot wear the clothes" (cf. Google Search)

b.Insertion of an object and perfect marker *le*

ta na le yi ben yuyan-xue de shu chu lai

he take ASP one CL language-study ASSOC book exit come

"He took a linguistic book out." (cf. Google Search)

c.Different aspectual markers for V1 and V2

ta pao zhe shang le shan

he run ASP(PROG) ascend ASP(PERF) mountain

"He went running up the mountain." (cf. Google Search)

First, the subparts of VDCs can be split by different morphemes. Sentence (2a) is an insertion of the potential particles *de* and *bu*, which mean 'can' and 'cannot.' Although it is debatable to analyze these particles as bound morphemes like infixes (Chao 1968; Lü 1979), the insertion of an aspectual marker and a long object between *na* 'get' and the DP *chu lai* 'exit come' in Sentence (2b) still supports that VDCs can be syntactically separable (Chung 2006; Fan 1981; Hsiao & Lin 2010). Second, subparts of VDCs can be modified by distinct semantic operators, such as aspect (Van Valin & Lapolla 1997: 456). In Sentence (2c), the progress aspect *zhe* has its scope over the manner verb *pao* 'run', and the perfect aspect *le* has its scope over the DP *shang* 'ascend.' Therefore, the violation of Principle of Lexical Integrity indicates that VDCs are phrases instead of compound words.

In addition to syntactic interruptability and semantic accessibility, other diagnostics on semantics, syntax and frequency are developed (Liu et al. 1998). For example, a compound tends to intransitivization (*sheng chu* lit ('win exit') 'win (intransitive)') and high frequency. Liu (1998: 54) finds that only around 200 out of 3000 types of VDCs are compounds.

Another way to analyze VDCs is to create terms to describe the patterns between words and phrases, similar to phrasal verbs for English VPCs. Wang's *leyu* (*仂語*) 'compounding units'

(1943: 25) and Lü's *duanyuci* (*短语词*) 'phrasal words' (1979: 10) are examples of this type of analysis These terms demonstrate VDCs as dualities between phrases and words, but this is very difficult to apply to linguistic applications like automatic syntactic parsing, dictionary development, and grammar instruction. Therefore, this approach may be theoretically correct but cannot bring enough distinctions to be practical.

Our primary conclusion with respect to the constituent type of VDCs is that a phrase analysis has advantages over a compound or intermediate level analysis. After being tested on multiple standards, the majority of VDCs still justify themselves as phrases. DPs can be viewed as verbs on a continuum of idiomaticity, varying in their argument structures. Given these conclusions for DPs and VDCs, we choose to model VDCs are as SVCs in this research.

Appendix 2: Metaphors Required by the Subclasses of VDCs

Metaphors necessary for the subclasses are underlined.

Class#1-1 Self-Motion-Basic-NonMetaphor

ta congrong-de <u>zou</u> <u>jin</u> wu-zi

he leisure-ADV walk enter house-SUF

"He walked into the house leisurely." (cf. Baidu search)

Class#1-2 Self-Motion-Basic-Metaphor

ta de nei-xin <u>yong</u> <u>shang</u> yi fen shenqie de ci-tong

she ASSOC inner-heart gush ascend one CL deep ASSOC pierce-pain

"A piercing pain gushed out of her heart." (cf. BCC)

PAIN IS LIQUID

HEART IS A LOCATION

Class#2 Self-Motion-Shifted

suoni ganjin chu xin-pin, jianneng yijing <u>zhui</u> <u>shang</u> u le

SONY hurry.up produce new-item, Canon already chase ascend come ASP

"SONY, please make new products! Canon already caught up (with you)!" (cf. BCC)

<u>APPROACHING HORIZONTALLY IS ASCENDING</u>

BRAND IS A MOVING OBJECT

Class#3-1 Caused-Motion-Displacement-NonMetaphor

qing ba zhe ben shu <u>fang</u> <u>hui</u> yuan-chu

please BA DEM CL book put return original-place

“Please put this book to where it was.” (cf. Baidu search)

Class#3-2 Caused-Motion-Displacement-Metaphor

gai ju yunyong shui-shou ganggan, ba zhe-xie qiye <u>qiao</u> <u>shang</u> <u>lai</u>

DEM bureau use tax-collect level, BA DEM-PL company pry ascend come

“This bureau used tax as a level to pry these companies out of trouble.” (cf. BCC)

ORGANIZATION IS A PERSON

TAX IS AN INSTRUMENT

<u>SITUATION IS A LOCATION</u>

Class#4-1 Caused-Motion-Transfer

zhongzuo ba xianggang <u>shou</u> le <u>hui</u> <u>lai</u>

China BA HongKong collect ASP return come

“China took Hong Kong back (from Britain).” (cf. BCC)

NATION IS A PERSON

CITY IS AN OBJECT

<u>POSSESSION IS A LOCATION</u>

Class#5 Caused-Motion-Scale

you-jia die hui yi nian qian

oil-price fall return one year before

“The oil price fell back to what it was a year ago.” (cf. Baidu search)

PRICE IS A MOVING OBJECT

VALUE ON A SCALE IS A LOCATION

Class#6 Caused-Motion-Creation

jintian fubai de cheli-zi, wo zhi xi chu lai yi xiao wan

today rotten ASSOC cherry-SUF, I only wash exit come one small bowl/CL

“Today, I only washed one small bowl of (good) cherries out of all the rotten ones.” (cf. BCC)

CREATING IS EXITING

Class#7 Caused-Motion-Attachment

ta zai che-shen liang-ce pen shang G zi biaoshi

she at car-body two-side spray ascend G character mark

“She sprayed the ‘G’ marks on both sides of the car.” (cf. BCC)

ATTACHMENT IS ASCENDING

Class#8 Result-Reflexive-Motion

ta juran kan dao ziji zai man-man <u>huansan</u> <u>kai</u> <u>lai</u>

she unexpectedly see reach self ASP slow-slow slacked move.away come

"Unexpectedly, she realized that she was beginning getting slacked." (cf. BCC)

EMOTIONAL STATE IS GAS

<u>SPREADING IS DIRECTED MOTION (or SPREADING IS BEGINNING)</u>

Class#9 Result-Affordability

<u>shu</u> de <u>qi</u> caineng <u>ying</u> de <u>qi</u>

lose DE rise then win DE rise

"Those who can afford the cost of losing can be able to win." (cf. BCC)

<u>ATTAINMENT IS DIRECTED MOTION</u>

Class#10 Result-Completive

ta cong xiao jiu yu pingpang-qiu <u>jie</u> <u>xia</u> bu-jie-zhi-yuan

He from small then with pingpong-ball tie descend NEG-dissemble-ASSOC-connection

"He has built an indissoluble bond with Ping-Pong since young. " (cf. BCC)

SPORT IS A PERSON

CONNECTION IS A TIE

COMPLETION IS DIRECTED MOTION

Class#11 Aspect-Inchoative

yu'er shuo zhe ye mo qi le yan-lei

yu'er speak ASP also wipe rise(ASP) ASP eye-tear

"Yu'er also started wiping her tears while speaking." (cf. BCC)

BEGINNING IS DIRECTED MOTION

Class#12 Aspect-Continuous

da-pan queshi zhang le, erqie hai

big-plate certainly increase ASP, and still

you-keneng yizhi zhang xia qu

have-possibility always increase descend(ASP) go(ASP)

"The market price certainly increases, and possibly always continues increasing."

MARKET IS LIQUID

CONTINUATING IS DIRECTED MOTION

Class#13 Discourse-Connective

hua shuo hui lai, wo-men shui you zhenzheng liaojie guo shui ne words speak return come, I-PL who again really understand ASP who QP

"After all, among us who really understood others?" (The VDC is the first clause and its meaning is "after all/anyway")

(DISCOURSE) TIME IS DIRECTED MOTION

Class#14 Evidential-Modality

na ji kuai yuncai kan qi lai zhen xia ren

DEM several CL cloud look rise come really scare people

" (I think) Those clouds look really scary."

EMMERGENCE OF EVIDENCE/COMMENT IS DIRECTED MOTION

Appendix 3: Some Coerced VDCs and Their Google MT Errors

The coerced VDC sentences below were tested on Google Translate (Chinese-English) on November 6, 2019. There are two basic types of MT errors: (1) translation problems with V1s, and (2) translation problems with V2s. Free translation from native English speakers is also provided.

1. youde guo-jia jiu gei ni yi-dian xiao-lian,

   some country-family just give you one-little smile-face,

   jiu ba zhong-guo mei shang tian le

   then BA middle-country beautify ascend heaven ASP

   "Some countries give you some smiley faces, and they compliment China to the heavens." (cf. CCL)

   [Google MT] Some countries will give you a smile and put China to the sky.

   Problem: no translation of V1

2. ta tui zou da guo lai de san

   he push move open pass come ASSOC umbrella

   "He pushed away the umbrella that someone had opened for him." (cf. CCL)

   [Google MT] He pushed away the umbrella that was hit

   Problem: wrong translation of V1

3. yi chu lai, na gu re-lang jiu gei ni gong hui lai le

   once exit come, DEM CL hot-wave then give you push return come ASP

   "When you come out of the house, the heat wave will push you back." (cf. CCL)

   [Google MT] When the heat wave comes out, it will arch you in.

   Problem: wrong translation of V1

4. ta yinwei zai huo-che-zhan fujin dao-mai che-piao bei juliu

   he because at fire-vehicle-station nearby reverse-sell vehicle-ticket PASS detain

   xuexiao pai ren qu ba ta bao hui lai le

   school send person go BA he bail return come ASP

   "He was detained for reselling a ticket near the train station. The school sent someone to bail him out of the police station." (cf. CCL)

   [Google MT] He was detained for reselling a ticket near the train station. The school sent someone to keep him back.

   Problem: wrong translation of V1

5. zhong-ri zhanzheng zhong, ri-jun cong Nanjing qie qi ling-gu

   China-Japan war middle, Japan-army from Nanjing steal his spiritual-bone

   qian xie riben, zhan-hou jiaoshe, de ying hui ling-gu

   dive take Japan, war-behind communicate, MOD welcome return spiritual-bone

"During the Sino-Japanese War, the Japanese army stole the monk's spiritual bones from Nanjing. They secretly took them to Japan. China negotiated with Japan after the war, and was able to welcome the spiritual bones back from Japan." (cf. CCL)

[Google MT] In the Sino-Japanese War, the Japanese army sneaked his spiritual bones from Nanjing, sneaked in Japan, negotiated after the war, and greeted the spirits.

Problem: no translation of V2

6. ta shi xiang ba ta <u>kuang</u> <u>hui</u> jia <u>qu</u>

   he COP want BA she trick return home go

   "He indeed wanted to trick her into coming home."

   [Google MT] He wants to take her home.

   Problem: wrong translation of V1

7. zhe shi wo zhong de shu, wo xihuan <u>kan</u> <u>hui</u> jia <u>qu</u>

   DEM COP I grow ASSOC tree, I like chop return home go

   "This is the tree I grew. I like chopping it down and taking it back home."

   [Google MT] This is my tree, I like to go home

   Problem: no translation of V1

8. jin-nian ben-gang jiang yong kuo-gu fangshi zhu-bu ba

   now-year Ben-steel will use expand-stock manner gradual-step BA

   tie-chang, leng-zha-chang kuo jin qu

   iron-factory, cold-roll-factory expand enter go

   "This year, Benxi Steel will gradually include the iron and cold rolling mills into their plan by means of increasing stock."

   [Google MT] This year, Benxi Steel will gradually expand the iron and cold rolling mills by means of share expansion.

   Problem: wrong arguments of V1

9. ni ba wo de linggun pai jin qu le

   you BA I ASSOC soul shoot enter go ASP

   "Your shooting captured my soul in the pictures."

   [Google MT] You took my soul into it.

   Problem: wrong translation of V1

10. gu-gong-dui ba zhe-xie-ge chou-ni quan wa dao ma-lu bian'er shang

    hire-labor-team BA DEM-PL-CL stinky-mud all dig reach horse-road side top

    "The hired team who were digging the stinky mud flung the excess mud from their shovels to side of the road." (cf. CCL)

[Google MT] The hired team dug all these stinky muds to the side of the road.

Problem: wrong English CM-VPC

11. ruguo bei mou-ren qifu le, bu yao keqi-di <u>qifu</u> <u>hui</u> <u>qu</u>

if PASS some-person bully ASP, NEG MOD polite-ADV bully return go

"If you were bullied by someone, don't bully back politely." (cf. CCL)

[Google MT] If you are bullied by someone, don't be polite to go back.

Problem: wrong translation of V1

12. ta yong zhe zhao ba qian-nü-you <u>zhui</u> <u>hui</u> <u>lai</u>

he use DEM trick BA previous-female-friend chase return come

"He used this trick to win his ex-girlfriend back."

[Google MT] He used this trick to chase his ex-girlfriend back. (cf. CCL)

Problem: wrong English CM-VPC

13. yao-shi bu gen wo jie-hun, ta wanquan keyi

MOD-COP NEG with I connect-marriage, she totally MOD

<u>ban</u> <u>hui</u> shanghai <u>qu</u>

do return Shanghai go

"If it were not that she married me, she could have totally gone back to Shanghai by finishing the documents." (cf. CCL)

[Google MT] If she doesn't marry me, she can go back to Shanghai.

Problem: no translation of V1 (*ban* means 'deal with (business, documents, cases, and so on)')

14. wo-men pan a pan a, hao bu

I-PL expect ahh expect ahh, good NEG

rongyi ba ta-men <u>pan</u> <u>hui</u> <u>lai</u> le

easy BA he-PL expect return come ASP

"We kept expecting. Finally, our expectation brought them back." (cf. CCL)

[Google MT] We are looking forward to it! Finally, they are looking forward to it.

Problem: no translation of V2

15. yi ge jin-fa-bi-yan de da lao-wai

one CL gold-hair-green-eye ASSOC big old-out

bei liumeiping <u>xiu</u> <u>guo</u> <u>lai</u> le

PASS Meiping.Liu sniff pass come ASP

"A tall foreigner with blonde hair and green eyes was enticed to come by Meiping Liu."

[Google MT] A blond big foreigner was sniffed by Liu Meiping

Problem: no translation of V2

16. zai jia deng le ji tian, dou

at home wait ASP some day, still

mei-you ba ge-ge <u>deng</u> <u>hui</u> <u>lai</u>

NEG-have BA older.brother-older.brother wait return come

"I waited home for several days. My waiting didn't bring my older brother back." (cf. CCL)

[Google MT] I waited for a few days at home and didn't return my brother.

Problem: no translation of V1

17. ta-men cha-dian ba lankuang <u>xiao</u> <u>xia</u> <u>lai</u>

she-PL lack-little BA basket laugh descend come

"They almost laughed down the basket." (cf. CCL)

[Google MT] They almost laughed at the basket

Problems: wrong argument of V1 & no translation of V2

18. fuqin ba er-zi <u>bing</u> <u>hui</u> <u>lai</u> le

father BA son-SUF sick return come ASP

"The Father's sickness brought his son back." (cf. Google Search)

[Google MT] Father returned his son to illness

Problem: wrong argument structures

19. ta-men ba wo xu chu le wu-zi

he-PL BA I boo exit ASP house-SUF

"They booed me out of the house." (cf. Google Search)

[Google MT] They took me out of the house

Problem: no translation of V1

20. laowang ba dian kai dao le Zhengzhou

old-Wang BA store open reach ASP Milan

"Old Wang opened a store in Zhengzhou, a place he didn't choose before." (cf. Google Search)

[Google MT] He drove the store to Zhengzhou

Problem: wrong translation of V1

21. kan xiao-shuo neng ba chengji kan shang qu ma

read small-speak MOD BA score read ascend go QP

"Can reading novels increase your grade?" (cf. Google Search)

[Google MT] Can reading a novel look like a grade?

Problem: wrong translation of V1 and V2

Appendix 4: Glossary

ADV the adverb morpheme 地 *de*
ARG argument in PropBank
ASP aspect
ASSOC associative 的 *de*
AVM attribute-value matrix
BA the 把 *ba* morpheme
BCG Berkeley Construction Grammar
biLSTM bidirectional long short-term memory
CAMR Chinese Abstract Meaning Representation
CCL Center of Chinese Linguistics
CL classifier
CM Caused-Motion
CMT Conceptual Metaphor Theory
CPB Chinese PropBank
CTB Chinese TreeBank
CWN Chinese WordNet
CxG Construction Grammar
DE the 得 *de* morpheme
DEM demonstrative
DP Directional Particle
DV Directional Verb (=V2)
EM embedding
EXIST existential
FN FrameNet
FORM the phonological form in SBCG
FS feature structure in SBCG
GL Generative Lexicion
HPSG Head-Driven Phrase-structure Grammar
LVC light-verb construction
MIPVU Metaphor Identification Procedure Vrije Universiteit
MOD modal verb
MRW metaphor-related word
MT machine translation
MWEs multi-word expressions
NEG negation
NLP Natural Language Processing
PCTB Penn Chinese TreeBank
PERF perfect
PL plural
POS part-of-speech
PP prepositional phrase
PROG progressive
QP question particle
RECP reciprocal
REFLX reflexive
REL relation in CPB
SBCG Sign-Based Construction Grammar
SEM the semantic feature structure in SBCG
SNACS Semantic Network for Adposition and Case Supersenses
SRL semantic role labeling
SUF suffix
SUP superlative
SVC serial verb construction
SVM support vector machine
SYN the syntactic feature structure in SBCG
THULAC Tsinghua University Lexical Analyzer of Chinese
UD Universal Dependency
V1 the first verb in a VDC
VA intransitive action verb in CWN
VAL valence in SBCG
VC transitive action verb in CWN
VDC Verb-Direction Construction
VP verb phrase
VPC Verb-Particle Construction
VRD the “verb and ressultative compound” label in CTB
VV the normal verb label in both CTB and the 863 Tagset
XARG the external argument in SBCG