Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.

2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.

3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

A CORPUS BASED COMPUTATIONAL MODEL OF THE LEXICAL ASPECT AND VIEWPOINT ASPECT IN CHINESE

HONGCHAO LIU

PhD

The Hong Kong Polytechnic University

2018

The Hong Kong Polytechnic University

Department of Chinese and Bilingual Studies

A Corpus Based Computational Model of the Lexical Aspect and Viewpoint Aspect in Chinese

Hongchao Liu

A thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

December 2017

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor materials that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____ (Signed)

_____Hongchao Liu_____ (Name of Student)

Abstract

This thesis talks about the lexical aspect and viewpoint aspect in Mandarin Chinese through statistical and computational methodologies.

I firstly show the necessity of studying verbs' situation type and applying statistical methodologies toward linguistics studies. Some of the previous aspectual studies on Mandarin Chinese deny the possibility and appropriateness to classify verbs into different situation types.

However, they are self-contradictory by using at least three strategies. The first one is to assign the whole structure's situation type to the constituting verb. The second one is to use different terminologies to refer to the situation type in lexical level such as aspectual parameter which is the same with situation type in essence. The third one is to explicitly deny the possibility to classify verbs into different situation types but implicitly do the classification in lexical level.

Some other studies are on the right track to admit that situation types are supposed to be differentiated in at least two levels including lexical and sentential level. However, none of them applied statistical validation toward the interaction between situation types and viewpoint aspects.

Because of the problems in lexical aspect and methodologies, verb situation type and statistical validation fall into the focus of this thesis.

Based on our own intuition and previous studies, I construct a hypothesis stating that aspectual markers including ZHE, LE1, LE2, GUO, ZAI and ZHENGZAI are able to classify different situation types.

I also insist that situation type in lexical level is attached to the different senses of a verb instead of the verb *per se* and that situation type system is a prototype category.

With a hypothesis stating that situation type system is a prototype category, the members of the category are supposed to be clustered based on their family resemblance represented by their ability to co-occur with different aspectual markers.

Whether a verb or verb sense is able to co-occur with an aspectual marker is firstly judged by our own intuition and then cross-validated by other annotated resources. A matrix of co-occurrence is constructed including the verbal senses as the rows and the aspectual markers as the columns.

The family resemblance is simulated by the distance of the rows position in the vector space represented by the matrix. Hierarchical clustering is implemented and automatically generates the situation type system based on the distance between members. In this way, three situation types are constructed and annotated to all of the selected verbs' senses.

Since the situation type system is actually based on human intuition, a corpus-based validation is necessary.

All the verb senses are manually linked to Sinica corpus' verbs and a co-occurrence frequency matrix is constructed based on the corpus data. Statistical methodologies such as multinomial logistic regression analysis, are used to validate our situation type system. Aspectual markers' relationship with situation

type's cognitive conceptual features including [Telic], [Durative], [Dynamic] etc. are also constructed in this way.

Finally, we construct a dataset with verb senses and their situation types and make evaluation tests on it. By using word embedding vectors and supporting vector machine classifier, a best accuracy of 72.05% is achieved.

Acknowledgements

The thesis urged and realized my diving into linguistics with a more scientific way and during this process, I should say thanks to my supervisor, friends and family. Without them, there will be no chance for me to finish such a journey that is full of overgrown bushes.

Firstly, my sincere appreciation is given to my supervisor, Prof. Chu-Ren Huang who is the wisest and the most generous scholar in my mind. He is the one who lent me his hands when I have to abandon my research career because of my family problems. His eyes are sharp enough to give light on my direction when I was lost in and slowed down by distortions. He is also the one who leads me to empirical methodologies to linguistics studies. When I was in despair, the encouraging words are always timely from him.

Secondly, thanks are supposed to go to Dr. Hou Renkui, Dr. Lin Yuleng, Jiang Menghan, Sun Jiaming and Wang Shichang. Dr. Hou Renkui's study inspires me a lot and the discussion with him helped me to find new ways in analyzing data. Dr. Lin Yuleng, Jiang Menghan, Sun Jiaming and Wang Shichang kept comforting me when worries came into flood me. I learned a lot from them not only on research but also on how to be a good person.

Finally, my thanks and apologizes go to my family. I feel sorry to my wife who gave birth to our lovely son for my absence from her when she need me and to my parents for my inability to accompany them when they are already sixty years old.

Table of contents

I

List of figures

List of tables

VIII

List of formulas

List of abbreviations

ZHE	Aspectual marker 着
LE1	Post verbal aspectual marker 了
LE2	Sentential final aspectual marker 了
ZAI	Adverb 在
ZHENGZAI	Adverb 正在
GUO	Aspectual marker 过
DE	的
HEN	Degree adverb 很
RVC	Resultative verb compound
POS	Part-of-speech
VV	Verb reduplication
VP	Verb phrase
NP	Noun phrase
PP	Prepositional phrase
TP	Temporal phrase
CLA	Classifier
ASP	Aspectual marker
SYLL	Syllable length
VSM	Vector space model

MNLogit	Multinomial logistic regression classifier
SVM	Supporting vector machine
ANN	Artificial neural network
RNN	Recurrent Neural Network
CNN	Convolutional neural network

Chapter 1

Introduction

This thesis talks about aspect in Mandarin Chinese which is an aspect language. Two components theory developed by Smith (1991), the compositionality and recursiveness of situation type developed by Verkuyl (1972) and Pustejovsky (1991) and the prototype nature of lexical aspect shared by Smith (1991), Zhang (1993) and Guo (1993), are adopted as the theoretical foundation for this thesis.

Statistical methodology runs through the attests of the interaction between verbs situation types including state, activity and transition and viewpoint aspects materialized by six aspectual markers including ZHE, LE1, LE2, GUO, ZAI and ZHENGZAI.

The paradigm shift from intuition to corpus has made a giant leap on the objectivation and scientification of lexical semantics and syntax. This thesis is trying every effort to promote the shift from example-counting method to statistical method to validate the linguistics assumptions, which is a more scientific way.

This part of the thesis is dedicated to several basic questions including the definition, focuses, significance, methodology, organization of the thesis and also a part of literature review.

1.1 What is aspect

This section sets foot on the definition and nature of aspect.

1.1.1 Definition

While aspect becomes a battlefield for scholars to testify their different theories, the term "aspect" have been used to refer to different concepts and "the lack of terminology consistency in the distinction of aspect makes reviewing research on aspect a very challenging task" (Xiao & McEnery, 2004).

The term "aspect", as a French word, originally derives from the Russian word *vid* which refers to "view" (Xiao & McEnery, 2004; Zhang, 1993). While Xiao and McEnery (2004) emphasize its vision or scanning facets, which is similar with Smith (1991)'s viewpoint aspect that only covers grammaticalized forms such as adverbs, Zhang (1993) focuses on verbs by arguing that *vid* actually stands for "type, category" in Russian grammar by claiming that "N.Grec……classifies six *vidy* (group) of verbs according to their morphological and semantic features" (Zhang, 1993). Similar to "aspect" in French and English, German also borrows the meaning of *vid* in Russian grammar to its word *Aktionsart* which refers to "kind of action" (Klein, 1994; Klein, Li, & Hendriks, 2000).

Vendler (1967) and Comrie (1979) stand in different lines in terms of the type of aspect. Vendler (1967), who pays attention to content words especially verbs, quartered the verbs and verb phrases according to three features including

2

[Dynamic], [Durative] and [Telic], which are widely adopted and adapted in other aspect studies. These three features, although with different names, are directly related with the internal temporal structure of a situation which consists of inception, duration and final stage of a situation.

While Vendler (1967) turns his focus on the inherent meaning of verbs, Comrie (1979) seeks to explore the grammticalized forms or mechanisms. A camera metaphor is used to simulate aspectual markers' function. Perfective and imperfective are used to distinguish among different perspectives that aspectual markers take for the internal temporal structure of a situation. While perfectives take a whole or complete picture of a situation, imperfectives scan a part of the temporal structure such as the inception or duration. While all of languages' verbs, verbs phrases, sentences and even discourses can be grouped into different types according to their internal temporal structures, not all of languages grammaticalize abundant forms to view or take pictures of the internal temporal structure of situations expressed by different language units. In this sense, languages are divided into aspect language such as Chinese and non-aspect language such as German (Zhang, 1993).

Smith (1990, 1991, 1994) make a clear cut between these two kinds of aspects by supplying us with the two components theory. Two components theory treats situation type and viewpoint aspect as independent but interacted parts of aspect. While situation type is determined by the internal temporal structure expressed

3

through the inherent meaning of verbs' constellation, viewpoint aspects are usually denoted by grammatical morphemes (Smith, 1991).

Xiao and McEnery (2004) agree on the distinction of these two aspectual components while Zhang (1993) declines to accept it. She maintains that there is no clear boundary between viewpoint aspect and situation type since both of them are concerned with the internal temporal structure of a situation and "are understood against a common domain, involve the same conceptualization of situations and operate on the same underlying semantic structure" (Zhang, 1993). There is no doubt that viewpoint aspect and situation types stay together within one domain which belongs to the internal temporal structure of the situations. However, they are not twins who are not able to be distinguished and their differences are big enough to pull them away from each other.

What's more, Zhang (1993) is self-contradictory by assigning different names to situation type and viewpoint aspect. While perfective and imperfective are used to distinguish between situation types encoding change and no-change, perfectivization and imperfectivization are created to signal the process of adding aspectual information to verbs by attaching grammatical elements. Zhang (1993)'s statement of the relationship between change and perfective, although partially accepted by us, is debatable and will be discussed in detail in chapter 2.

Before diving into the nature of aspect, a definition is supposed to be constructed based on previous studies discussed above. Two components theory

is adopted as the foundation of this study and thus the distinction between situation type and viewpoint aspect is inherited. Aspect is reserved to be the hypernym of situation type and viewpoint aspect.

However, we argue against Smith (1991) and also some other scholars on the assertion that situation type can only be assigned to sentence.

Situation type is classified based on the inherent meaning of a verb or the compositional meaning of a structure bigger than a verb.

Some scholars such as Smith (1991) and Verkuyl (1972) insist that situation type can only be assigned to a verb constellation or sentence, which is quite debatable. It will be shown in the next section that situation type is compositional and recursive in nature which is inherited from the morphology and syntax of language structures and thus can be assigned to language structures in different levels. Among them, situation type of verbs, which is usually named by the German word *Aktionsart*, is called lexical aspect in this thesis. What Smith (1991) and Verkuyl (1972) called situation type are actually the aspectual meaning/information of the whole sentence and aspectual information contains both situation type and viewpoint aspect. However, Sentence is neither the solo nor the smallest unit for bearing aspectual information.

Situation type signals the conceptual features of [Dynamic], [Telic] and [Durative] etc. which are shortcuts of the internal temporal structure of situations. We insist a cognitive prototype model of situation types.

Guo (1993) and Zhang (1993) contribute a lot to the implementation of the model but they are not without any flaws. Semantic distance based on vector space model (Hongchao Liu, Karl, Enrico, & Huang, 2016; Turney & Pantel, 2010) and Hierarchical clustering methodology (Hou, Huang, Do, & Liu, 2017; Hou, Huang, & Liu, 2017) are used to validate and calibrate situation types. Situation types will be discussed in detail in chapter 3.

Viewpoint aspect takes a whole or a part of the internal temporal structure of a situation. Comrie (1979) applies a metaphor between grammaticalized aspectual markers and camera. While a scan of the whole internal temporal structure including the inception, duration and final from outside, is called perfective, a scan or a view of a part of the structure such as the duration from inside, is called imperfective.

Klein (1994) is critical of the name of "viewpoint" which is a metaphorical use and lacks accuracy in describing the function of aspectual markers. However, internal temporal structure, as an abstract concept, has to be simulated by a concrete alternate. Before "viewpoint", "telic" and "bound", which are spatial term, have been widely accepted to depict temporal ends and thus there is no need to argue against the appropriateness of using "viewpoint" aspect. However, Smith (1991)'s statement of the function of viewpoint, although influential, is problematic because she actually denies the recursiveness of aspectual information which will be explored in the next section.

1.1.2 Nature: compositionality and recursiveness

Whether verbs can be classified based on situation types have been debated till now since Vendler (1967)'s quartering classification of verbs. The problem arises, on one hand, from terminology confusion of the so-called situation type and from the nature of compositionality and recursiveness of aspectual meaning, on the other hand. This section argues that verbs are supposed to be assigned with situation types by reviewing the previous studies and explaining the compositionality and recursiveness of aspectual meaning. Let's start with the very first modern study of verbs' situation type.

Vendler (1967)'s verb classification can be illustrated by the following table:

Table 1 Vendler (1967)'s verb class

Verb class	Features			Examples
	Telic	Durative	dynamic	
State	-	+	-	know, love
Activity	-	+	+	run, drive a car
Accomplishment	+	+	+	build, run a mile
Achievement	+	-	+	die, recognize

Although with various problems, Vendler (1967) serves as an excellent initiation for the study of situation type.

The flaw of Vendler (1967)'s theory can be clearly portrayed by *run* and *run a mile*. While *run* itself is an activity, the *run* in *run a mile* implies a situation type of accomplishment because it proceeds "toward a terminus which is logically necessary to their being what they are" (Vendler, 1967). Thus, the

7

process of classification definitely triggers collision of verbs' classes since they may belong to different situation types.

Scholars developed different models to solve this problem (Dowty, 1979; Friedrich, 1974; Moens, 1987; Mourelatos, 1978; Pustejovsky, 1991; Smith, 1991; Verkuyl, 1972). But one of the consensus they share is the compositionality of aspectual meaning. The core of the compositionality is that verb is not the only constituent that contributes to the aspectual meaning of a structure, the arguments, adverbials and adjuncts also effect the aspectual meaning or situation type of the structure. However, compositionality along cannot solve the double entry problem of a verb in Vendler (1967)'s model:

(1) a. He ran.

   b. He ran to the store.

(2) a. He wrote.

   b. He wrote a letter.

While the verb *ran* in (1a) is classified as an activity verb, the VP phrase *ran to the store* gains a situation type of transition because of the existence of *to the store* which serves as "an attainment of the goal" (Pustejovsky, 1991).

(2a) is similar but the constituent, which contributes to the aspectual meaning other than verb, become its internal argument.

It seems no difference with Vendler (1967) so far. The problem of *ran*'s collision on its situation type still exists even by confirming the contribution of

other elements in a structure. Vendler (1967)'s real weakness lies in that he is not aware of the hierarchy of situation type and assign the situation type of the whole structure to its constituting verb. Thus, the real contribution of accepting compositionality is that it positioned verbs, verb phrases and sentences into different level which reflects the recursiveness of aspectual information.

The way begins to bifurcate from here. Smith (1990, 1991, 1994) stand for one group which persists the consistency of aspectual class between the whole structure and its constituting verb while Verkuyl (1972) and Pustejovsky (1991) stand for another branch which maintains the independency of verb and the structure that contains it.

Trying to hold fast to the widely accepted classification rules and also examples of Vendler (1967), Smith (1991) moves her analysis to VP level and sentence level by throwing verbs out from the bearing units of situation types or aspectual class. In this way, the classification of situation types can be proceeded above lexical level and thus the collision of verb's situation type seems to be bypassed. However, it is impossible to veil the classification of verbs since the interaction between verbs, arguments and viewpoint aspects is of vital importance to the study of aspect in all of the languages. While nouns are divided into telic and non-telic groups and viewpoint aspects are divided into perfective and imperfective categories, it seems strange to keep verbs out of the aspectual classification.

Since it is inevitable to classify verbs, Smith (1991) uses aspectual value instead of situation type to refer to the aspectual class of verbs. The aspectual values include [Telic], [Durative] and [stative] which is exactly the same with situation type's conceptual features. Smith (1991) is neither the first nor the last one to use terminology strategy to deny that verbs have situation type attribute.

In Smith's model, "if a verb constellation in a maximally simple sentence is [+Telic], the verb will have the intrinsic feature [+Telic]; and similarly for the other features......Verbs will have the intrinsic features [Stative], [Telic], [Durative]" (Smith, 1991):

Table 2 Examples of verb's aspectual value in Smith (1991)

	Verb	Aspectual value of verb	Sentence
1	walk	[-Telic]	Mary walked
2	breathe	[-Telic]	John breathed
3	draw	[+Telic]	Ellen drew a picture

It is quite clear that there is no fundamental difference between Smith's clustering of verbs based on aspectual value and Vendler's classification of verbs. The double-entry problem of verb's aspectual value is still pendent. Instead of the compositionality, the collision of verbs' situation type serves as one of the most important reasons that she argues against to assign situation types to verbs. Smith (1991) keeps stressing that verbs is not the only element that contributes to the aspectual meaning of a sentence or a situation but there is no relationship

between this statement and dropping verbs out of situation type assigning implementation since compositionality of aspect never blocks the way to classify verbs into different situation types. It is the holding that verb's situation type must keep consistent with the verb phrase or sentence that excludes verbs out of the assigning of situation types.

What's more, it is quite weird to interpret the *draw* in *draw a picture* to be different from *draw per se*. After accepting the compositionality, it is even more strange to assign the whole structure's situation type or aspectual value to its constituting verb since compositionality is able to explain the independency between verb's situation type and the whole structure's situation type well enough and there is no reason to keep the consistency of situation type or aspectual value between verb and the whole structure.

The so-called shifting of verb's aspectual value in a structure makes it unable to analyze the interaction between verb and aspectual markers and many computational studies such as Zarcone and Lenci (2008), W. Li, Wong, Cao, and Yuan (2004) and Xu (2015) applied prediction of situation types only into the sentential level because of their accepting the shifting view.

By accepting a shifting view, these scholars deny the recursiveness of situation type or aspectual value and the compositionality is pseudo in their studies. On one hand, they keep arguing that verb is not the only element that contributes to aspectual meaning of a sentence. On the other hand, they pour all of the aspectual

11

meaning into verb which violates the compositionality of aspectual information. For example, in *run to the store*, the meaning of "attainment of goal" is clearly demonstrated by *to the store*, but the shifting of *run* into accomplishment actually implies that the "attainment of goal" has been transformed from *to the store* to *run*.

Recursiveness is a basic nature for syntactic structure. The language is recursive or hierarchical which means that there are different levels in a sentence. Words combined together to form phrases and phrases combined together to form sentences. While a sentence is pronounced or written linearly, the structure of the sentence is hierarchical.

Verkuyl (1972) makes a comprehensive study on the compositionality of aspect. The recuresivenss of aspectual information has been accepted and implemented in his analysis. Verkuyl (1972) maintains that "the opposition between imperfective and perfective aspect is not a matter settled at the verbal level. I proposed that aspect be 'taken away' from the verb and assigned to higher level of sentential structure……" However, there is no way to bypass the aspectual classification of verbs if compositionality and recursiveness is accepted.

Verkuyl (1972) assigns aspectual parameter to verbs which is followed by Smith (1991). However, there is no fundamental difference between assigning aspectual parameter and situation type as the classification of situation type is also based on aspectual parameter or value.

There are two parameters in Verkuyl (1972)'s compositionality theory. [A] represents the feature of [Change] for verbs while [B] "stands for a specified quantity of the NP " as it shown in Zhang (1993)'s examples:

Table 3 Aspectual parameters of NPs and verbs in Verkuyl (1972)

1	She	[+B]	played	[+A]	that sonata	[+B]	terminative
2	She	[+B]	played	[+A]	sonatas	[-B]	durative
3	She	[+B]	hated	[-A]	that sonata	[-B]	durative
4	She	[+B]	hated	[-A]	sonatas	[-B]	durative
5	Soldiers	[-B]	played	[+A]	that sonata	[+B]	durative
6	Adults	[-B]	hated	[-A]	that sonata	[+B]	durative
7	Nobody	[-B]	hated	[-A]	sonatas	[-B]	durative

The subjects and objects in the above table imply that NP with [+B] feature is telic while NP with [-B] is atelic.

Verkuyl (1972)'s feature of [Change] classifies verbs into dynamic verbs and stative verbs which are with names of "perfective" and "imperfective" verbs in his model, although he denied applying situation type in lexical level. The so-called situation type in Verkuyl (1972)'s study is assigned in VP level and sentence level.

Table 4 Situation type in VP level in Verkuyl (1972)

1	(played	[+A]	that sonata	[+B]))	terminative
2	(played	[+A]	sonatas	[-B]))	durative

13

3	(hated	[-A]	that sonata	[-B]))	durative
4	(hated	[-A]	sonatas	[-B]))	durative
5	(played	[+A]	that sonata	[+B]))	terminative
6	(hated	[-A]	that sonata	[+B]))	durative
7	(hated	[-A]	sonatas	[-B]))	durative

In sentential level, situation type is firstly tri-sectioned into event (terminative) and non-event (durative) and then non-event (durative) is further divided into process and state.

Table 5 Situation type in sentence level in Verkuyl (1972)

1	(She	[+B]	(played	[+A]	that sonata	[+B]))	terminative	event
2	(She	[+B]	(played	[+A]	sonatas	[-B]))	durative	**process**
3	(She	[+B]	(hated	[-A]	that sonata	[-B]))	durative	state
4	(She	[+B]	(hated	[-A]	sonatas	[-B]))	durative	state
5	(Soldiers	[-B]	(played	[+A]	that sonata	[+B]))	durative	**process**
6	(Adults	[-B]	(hated	[-A]	that sonata	[+B]))	durative	state
7	(Nobody	[-B]	(hated	[-A]	sonatas	[-B]))	durative	state

The rules of compositionality in Verkuyl (1972) can be generalized as:

Table 6 Compositionality rules of situation type in Verkuyl (1972)

No.	NP1	Verb	NP2	Situation type	
				Binary	Triparted
1	([+B]	([+A]	[+B]))	terminative	event

14

2	([+B]	([+A]	[-B]))	durative	**process**
5	([-B]	([+A]	[+B]))	durative	**process**
6	([-B]	([-A]	[+B]))	durative	state
7	([-B]	([-A]	[-B]))	durative	state
3	([+B]	([-A]	[-B]))	durative	state

It is able to generalize that terminative appears only if all of NP1, verb and NP2 have plus features. Otherwise, the sentence gains a situation type of durative. For durative sentences, if their verbs have a feature of [-A], their situation type is state while durative sentences are assigned process if they involve verbs featured with [+A]. Generalizations are also applicable in VP level.

However, Verkuyl (1972)'s rules don't always work well which is illustrated in Zhang (1993):

Table 7 Examples contradicting Verkuyl (1972)'s rules

1	She	[+B]	killed	[+A]	that chicken	[+B]	terminative
2	She	[+B]	killed	[+A]	chickens	[-B]	terminative or durative
3	Adults	[-B]	killed	[+A]	that chicken	[+B]	terminative
4	Adults	[-B]	killed	[+A]	chickens	[-B]	terminative or durative

While (1) follows the rule of terminative or event since all of its components are featured with [+A] and [+B], all of the rest examples contradict Verkuyl (1972)'s assertion.

Both (2) and (4) are not processes without question and only iterative interpretation is able to guarantee a process reading. However, both (2) and (4) have no tendency to be iteratively decoded although their verbs are with [+A] aspectual parameter.

Sentence (3) is a validation of the rules. A verb with [+A] implies that only process or terminative is possible for sentence (3)'s situation type and the external argument with [-B] burns terminative out. However, it is clear that sentence (3) is a terminative.

It is the verb instead of the compositionality that makes the rule invalid. With the solo aspectual parameter of [+A], only stative and dynamic verbs are differentiated. However, *kill*, as a terminative or achievement verb, is not marked out in the lexical level but behaves variously from both *play* and *hate*. In order to patch the compositionality rules, another aspectual parameter which illustrates the difference between terminative and non-terminative verb must be created in the lexical level. It turns out to partition verbs into three categories which is consistent with the three situation types in VP level and sentential level in Verkuyl (1972)'s model.

Although insisting that situation type is not a lexical level phenomenon, both Smith (1991) and Verkuyl (1972) contradict themselves by actually categorizing verbs into exactly the same classes with situation types. The only difference is terminology.

However, Smith (1991) and Verkuyl (1972) are fundamentally different on assigning situation type or aspectual value to verbs.

Smith (1991), together with a lot of Chinese aspectual studies such as P. Chen (1988) and Xu (2015) which will be reviewed later, actually deny the compositionality of aspectual information although they stress the contribution of constituents other than verb in a sentence or VP. They either assign the whole structure's aspectual class to verb or assert that the verb shifts its aspectual class in a structure whose situation type is different from the verb *per se*. What they do is to accumulate all of the aspectual information to the verb because they hold that verb's situation type is dependent on the whole structure's and thus verb has to shift its aspectual value and even meaning back and force according to the arguments and the adjuncts that attached to the sentence. Under this perspective, there is no way to study verb's situation type because there is no fixed situation type for a verb. Unsurprisingly, they bypass this problem by claiming that situation type doesn't exist in lexical level.

Xu (2015) criticizes Eric Victor Siegel (1998) that assuming "each verb has a default event type excluding all aspectual operations" is the reason why "he has to rule out the high ambiguous verbs……" But Xu (2015) chooses the wrong target to blame since both of them belong to Smith (1991)'s group who denies the compositionality of aspectual information and also the independence of verb's situation type from the whole structures. The so-called "high ambiguous

17

verb" are generated because of accumulating all of the other elements' aspectual information to shift the verbs' aspectual class but not because of assigning situation types to verbs *per se*.

Verkuyl (1972) is on the right track, although he also denies using situation type to refer to verb's aspectual class and to confirm the independence of lexical aspectual class and compositionality. The recursiveness of aspectual information is deeply analyzed in her model.

Dai (1997) insists that "verb is not the only element expressing aspectual meaning and it is even not the bearing unit of aspectual meaning. The bearing unit of aspectual meaning is sentence……" But it is acknowledged later in his statement that "the study of situation type must be processed hierarchically. At least, the distinction between verb's situation type and sentence's situation type should be constructed……" However, what on earth aspectual meaning refers to is never clear in Dai (1997)'s model. It seems that Dai (1997) maps the term "aspectual meaning" to the biggest unit in syntactic level which is sentence. Although creating another terminology, Dai (1997) accepts the reucursiveness of situation type.

Pustejovsky (1991) follows Verkuyl (1972) and creates "the syntax of event type" to capture the rules of compositionality and recursiveness of aspect. The operation of the compositionality of situation type or event type is called "syntax" which is a metaphorical use of the operation in the syntactic level. We agree with

him and the syntax of event type makes it easier to understand the independency of verb's situation type from the whole structure's because in syntactic level, verb's grammatical category is not decided by the whole structure.

Zhang (1993) and Xiao and McEnery (2004) confirm the recursiveness of situation type and based on this perspective, Xiao and McEnery (2004) constructed a clearer model for Chinese aspect system.

In this section, the interchangeable usage of aspect's various terminologies is clarified. By giving clear statement of these terminologies, it is shown that there is no way to bypass the classification of verbs into different situation types and shifting the verb's aspectual value to the same with the whole structure or assigning the whole the structure's aspectual value to the verb is denying the compositionality and recursiveness of situation type and aspectual information.

1.2 What's the focus of this thesis and its significance

This thesis strives to supply a statistical analysis of the lexical aspect and viewpoint aspect in Mandarin Chinese. Thus, the contribution is about to be made on both the studied object and also the methodology.

1.2.1 Lexical aspect

Lexical aspect concerns the internal temporal structure of verbs or the situation type of verbs. The terminologies which are used to refer to lexical aspect are colorful. Although mixed with some phrases, Vendler (1967) uses aspect to refer

to the situation type of verbs. This is definitely a misleading that aspect only includes situation types of verbs. Comrie (1979) seems to go to another extreme by naming grammaticalized aspectual marker or viewpoint aspect as aspect. The German word *Aktionsart,* which is originally a translation of *vid* in Russian, serves as a specific term for the situation type of verbs in many studies (Brinton, 1988; Guo, 1993; Xiao & McEnery, 2004).

Although denying the possibility of classifying verbs into different situation types, Verkuyl (1972) and Smith (1990, 1991, 1994) apply aspectual parameter/value to partition verbs based on their internal temporal structure. Situation type is strived to be shadowed in lexical level. Lexical aspect is thus to be connected with the aspectual parameter/value in these studies.

Zhang (1993), Dai (1997) and Xiao and McEnery (2004) confirm the existence of situation type in both the lexical level and syntactic level. The situation type in lexical level (basically, verbs, including verb compounds) is "modeled as verb class" (Xiao & McEnery, 2004), while "situation type" is reserved to refer to the situation type in syntactic level specially and thus lexical aspect is concerned with verb class in these studies.

Lexical aspect is depicted by the features of internal temporal structure of verbs and situation type of verbs is also identified based on these features such as [Dynamic], [Durative] and [Telic].

In this thesis, lexical aspect, viewpoint aspect and their interaction are positioned as the studied object. Few studies on Mandarin aspect focus on lexical aspect, especially the ones beyond China mainland. The reason why less attention is paid on lexical aspect for them is the rejection of the independence of verbs' situation type from the whole structure and the insisting that the verb should shift its situation type back and force according to the attached elements. For example, the *push* in *push a cart* is portrayed as an accomplishment while *push* itself stands as an activity.

However, it has been clarified that this view is untenable and verbs' situation type or aspectual value should be isolated and there is no way to generalize or validate any rules for the interaction between verbs and viewpoints aspect under this perspective because verbs are not with set situation type. In this way, situation type of the whole structure, as a matter of fact, is assigned to the verb regardless of the verb's own meaning.

Table 8 Composite schema example in Smith (1991)

1	[Mary walk to school]	I…..F	Accomplishment schema
2	[be + ing]	…	Imperfective schema
3	Mary was walking to school.	I..//////..F	Composite schema

Since Smith (1991) maps the whole structure's situation type or aspectual value back to its constituting verb, the interaction between *be + ing* and *Mary*

walk to school doesn't conflict with the interaction between *be + ing* and *walk* itself because both of them are accomplishment under her perspective.

What's more, Smith (1991) further specifies that the imperfective viewpoint *be + ing* is applied to the whole structure of *Mary walk to school* instead of the verb *walk* by claiming that viewpoint aspect is independent from situation type. However, there is no doubt that the progressive marker is an inflection of the verb *walk* instead of the whole structure of *Mary walk to school*. We agree that viewpoint aspect is independent from situation type but not in this sense. The independency of the viewpoint aspect actually refers to its incapability of shifting the situation type of the verb or structure it is attached to.

Another self-contradictory argumentation is that if situation type only exists in sentential level, the interaction of situation types with other elements such as NPs, PP adjuncts and viewpoint aspects actually cannot be studied because these elements belong to the sentence which bearing situation type themselves. Thus, Smith (1991) and Xu (2015) lower down the bearing unit of situation type to VP level. However, this cannot be the final revision of their model because the interaction also exists between verb, not only VP, and other elements. As it is discussed in the previous section, Smith (1991) has no choice but to assign the so-called aspectual value to verb which is in essence assigning situation type to verb.

Another strategy is to strip off all of the viewpoint aspects, adjuncts or even arguments from the sentence when deciding its situation type like what Teng (1985) does in his study. However, this is a clear violation of the compositionality of aspectual meaning which severs as their foundation to reject the existence of situation type in lexical level.

By confirming the compositionality and revursiveness of situation type, *walk* itself is an activity, while *walk to school*, as a whole, has a situation type of accomplishment and thus their interaction with viewpoint aspect must be studied individually.

It should thus be noticed that lexical aspect is supposed to be studied before the research on aspectual meaning in VP level and sentence level. Verkuyl (1972) has made a detailed demonstration which have been shown in the previous section and the generalization of the composite rules is from bottom to up. Without studying the interaction between verb and other elements, which is positioned in the bottom of the syntactic level, there is no way to generalize the rules above it.

Even for studies like Zhang (1993) and Xiao and McEnery (2004) who adopt the recursiveness and compositionality of situation type, they also suffer to process the lexical aspect. Polysemy and homograph is heavy in Mandarin and different senses of the same form may appear to be with different situation type. To solve this problem, Zhang (1993) "categorizes polysemous verbs according

to their central or basic meaning." Such as the following examples in Zhang (1993):

(3) a. 他在走呢。

　　ta1　zai4　zou3　ne

　　He　ZAI　walk　NE

　　'He is walking.'

　　b. 他走了。

　　ta1　zou3　le

　　He　walk　LE

　　'He has left.'

　　Zhang (1993) assumes that the verb *zou* (walk) in (3a) has the central or basic meaning and the verb *zou* (leave) in (3b) is derived from the meaning of (3a). Thus, the verb *zou* (leave) is supposed to follow the verb *zou* (walk)'s situation type and becomes an activity. However, this implementation is problematic.

　　Firstly, it is hard to decide which sense of the polysemy verb serves as the basic or central meaning. Even it is possible, the process is quite subjective and fuzzy. Taking (3) as an example, Zhang (1993) doesn't supply any evidence that supports her statement that *zou* (*walk*), instead of *zou* (leave), is the basic or central meaning of *zou*.

Secondly, *zou* (leave)'s situation type is entirely different with *zou* (walk) and assigning activity to *zou* (leave) is contradicting the language reality. By claiming *zou* (walk) as the prototypical *zou* and *zou* (leave) as an untypical sense of *zou*, Zhang (1993) strives to applying prototype theory on partitioning the situation type and assigning situation types to verbs at the same time. However, *zou* (leave) as a transition or an achievement, behaves radically different from *zou* (walk) and if they are tied up together to be assigned the situation type of activity, the situation type in lexical level is about to become useless.

Xiao and McEnery (2004) judges verbs' situation type in 'neutral contexts' which is borrowed from Moens (1987). Xu (2015) criticizes that "sometimes it is not clear what context is neutral." However, by the definition of Xiao and McEnery (2004), which is adapted from Moens (1987), the neutral context is clear: "the object is syntactically and semantically a singular countable noun and should only be present if it is obligatory, i.e. with a necessarily transitive verb" and " perfective viewpoint aspect is preferable."

Xiao and McEnery (2004)'s problem doesn't lie in the clarity of neutral context but in the implication that verb's situation type could shift as a result of the existence of complements, arguments and non-arguments for they claim that the context "is deemed neutral when everything has been excluded that might change the aspectual value of a verb" (Xiao & McEnery, 2004). This assertion actually pulls Xiao and McEnery (2004) back to the shifting view in lexical level

25

and exposes that their confirmation of the existence of situation type in lexical level is pseudo and that their criticism of Smith (1991)'s problem is self-contradictory because one of the most important reasons that Smith (1991) rejects to assign situation types to verbs is that verb may shift its situation type because of the existence of other elements and thus verbs have no settled situation types. If Xiao and McEnery (2004) also take the similar view, how could they assign situation type to verbs without escaping this shifting problem?

Although accepting Moens (1987)'s definition, Xiao and McEnery (2004) actually have not realized that they are fundamentally different. While the neutral context denies the independency of verbs' situation type from other elements including the viewpoint aspect, the confirmation of recursiveness and compositionality of situation type or aspectual value is actually holding the independency of verb's situation type from the whole structure and they belong to the latter camp if they intend to maintain situation type in lexical level.

The root that makes Xiao and McEnery (2004) swing back and force within these two camps is that on one hand, they have realized that situation type is supposed to be assigned in lexical level based on the language reality. On the other hand, they are striving to keep examples like *draw a circle*, which is accomplishment in Vendler's model, in the same level with verbs. However, Vendler's accomplishments are in VP level and are "compositional both semantically and grammatically" (Brinton, 1988; Zhang, 1993).

By reviewing previous studies, lexical aspect or situation type/aspectual value of verb is either omitted or processed problematically. Based on the foundation of Zhang (1993) and Guo (1993, 1997), a situation type hierarchy is constructed and validated through statistical and computational methodology. Verbs are assigned situation types according to their senses which are decided in big corpus with the help of verb dictionary (Meng, 1987).

What's more, the prediction and application of lexical situation type has gain much more attention than sentential situation type in computational linguistics including Pustejovsky et al. (2003), Kazantseva and Szpakowicz (2010), Saurí and Pustejovsky (2012), Costa and Branco (2012) and Falk and Martin (2016) to name a few.

Lexical aspect is far from being studied enough. Because of the hardness of annotation of lexical aspect, many studies choose to bypass it which leave space for us to make a detailed research.

1.2.2 Scientific procedures

Other than aspect, quantitative or statistical methodology is also strived to be made prominent through most of the assertions and their argumentations in this thesis. While intuition and counting examples in corpus, as tendency detection methods, are sharp on putting forward hypothesis, they can hardly be used to validate the assumptions. While computational studies are scientific on testing linguistics models, they rarely step further into details on linguistics topics.

The importance and necessity of scientific procedures and methodologies, especially statistical methodology, are going to be materialized in this section.

Although usually mentioned together with literature and philosophy, linguistics *per se* is trying to be made different from them as a more scientific subject.

Just as Eddington (2009)'s statement, "literature, philosophy, pure mathematics, formal logic and art" are typical subjects which needn't to be scientific and they don't intend to follow the steps that taken in scientific subjects such as physics:

(i) Discover an interesting phenomenon that seems to show a tendency;

(ii) Propose a hypothesis based on the observation to match the tendency;

(iii) Validate the hypothesis using experiments and other supporting data;

(iv) Confirm or deny the hypothesis based on the validation results and make further analysis based on the validation.

It seems that quite a lot of linguistics studies, including studies on aspect, are trying their efforts to be on the track of science instead of art. Gabriel Altman, the pioneer of quantitative linguistics, formulated procedures for scientific linguistics studies (Haitao Liu & Lin, 2017):

(i) Come up with hypotheses that are empirically related and able to be validated;

(ii) Formulate the hypotheses using statistical expressions;

(iii) Seek for appropriate statistical methodology to validate the hypotheses;

(iv) Accept or refute the hypotheses according to the statistical test result;

(v) Explain the hypotheses.

Although displaying sharp eyes on coming up with hypotheses in linguistics, quite a few linguistics studies are reliable enough to be included into scientific studies because many of them share characteristics of pseudoscience including astrology, magic and extra-sensory perception etc. For example, their theories are lack of falsifiability; the validity of their theories is realized through intuition, introspection, reasoning and counting examples in corpus instead of statistical methodology.

1.2.2.1 Falsifiability

Some of the linguistics hypotheses are unable to be tested because they are unfalsifiable. Stanovich (2013) tells a story about the yellow fever diffused in Philadelphia in 1793 which shows the essence of unfalsifiability. When the fever broke out, Benjamin Rush, one of the persons who signed the Declaration of Independence, engaged in treating the patients. He applied a therapy of bloodletting to his patients including himself. Many people criticized that his application is more fatal than the fever *per se*. But Rush is confident in his theory even with a bunch of died cases. To sum up, his logic is that if the one, who has accepted bloodletting, got a symptom relief, he counted it as a supporting

29

evidence for his therapy; if the patient got worse or even died, he assumed that "the patient had been too ill for *any* treatment to work" (Stanovich, 2013).

Logically speaking, there is no way to refute Rush's hypothesis. But this doesn't mean his theory is perfect. On the contrary, it sounds ridiculous in now day. Unfalsifiable hypothesis is unable to be validated. In this sense, falsifiability equals to testability. Scientists handle studies on "testable theory" and "the way scientists makes sure they are dealing with testable theories is by ensuring their theories are falsifiable" (Stanovich, 2013).

A theory blocks its own way to make any progresses if it is unfalsifiable and their existence jeopardizes relevant research. The Freudian psychoanalysis always explains observations after they occur which means it predict nothing and thus is unfalsifiable. One example that it do harm to psychology is depicted by Stanovich (2013) for the study of the cause of Gilles de la Tourette syndrome. "Author after author presented unfalsifiable psychoanalytic explanations for the syndrome" (Stanovich, 2013). However, the abstract explanation actually blocks scientists' way to reveal the nature and treatment of the syndrome for three decades from 1921 to 1955 (Stanovich, 2013).

Linguistics studies also propose this kind of theory. The shifting view of lexical aspect serves as a typical example to display it. Smith (1991) formulates an interaction rule for aspectual marker LE2, which is used to express change of state in Mandarin, stating that LE2 only appears with non-statives (dynamics); if

30

it appears with statives, the situation types of state will shift to achievement or accomplishment which are able to co-occur with LE. For example, *bing4* (sick) which is a stative verb "shifts" its situation type to achievement:

(4) 我病了。

    wo3   bing4 le   .

    I   sick   LE   .

    'I got sick.'

When we are laughing at Rush's theory, the shifted view represented by Smith (1991) actually falls into the same circle. There is no way to falsify her rule logically and thus the rule is unable to be validated. Frankly speaking, if a verb's situation type is considered to be not independent from the whole structure, all rules for the shifted situation type are unfalsifiable because all of the counter-examples can be explained by the so-called "shifting".

Although these studies narrow down their situation type to VP level or sentential level, it should be noticed that most of their examples are actually concerned with the interaction between verbs and viewpoint aspects and thus the rules for shifted verbs are not uncommon which means that many of their rules are unfalsifiable.

Unfalsifiable rules or hypothesis is not rare in linguistics studies. On the contrary, it is quite common. Eddington (2009) lists various examples from

traditional phonology to fashionable optimality theory. It should be noticed that falsifiable theory doesn't equal to false theory and actually scientific theories are supposed to be falsifiable but true ones.

To achieve logic self-consistency is one of the requirements that scientific studies should follow. However, logic self-consistency should not be arrived through unfalsifiable assumptions. Neither should it be realized through neglecting contradictory evidences. Next section will cover this problem by showing examples in aspectual studies.

1.2.2.2 Statistical methodology

There is a paradigm shift for linguistics studies from intuition-based methodology to empirical-based methodology while the latter one mainly refers to corpus-based studies.

Before the fashion of corpus, intuition and introspection, as methodologies, almost dominate traditional linguistics. It is not our intention to deny the contribution of intuition and introspection in linguistics studies. On the contrary, their roles in the procedure of proposing hypotheses are confirmed. "With the intuition-based approach, researchers can invent purer examples instantly for analysis because intuition is readily available……" (Xiao & McEnery, 2004). With the help of incisive intuition, linguists are usually with higher possibility to find tendencies to be tested than other ones.

What's more, a lot of incisive explanation models are developed by intuition and introspection methodologies. Smith (1991)'s two components theory serves as one of the most influential works for aspectual studies including this thesis.

However, it would be dangerous if linguists' own intuition and introspection are also applied to validate hypotheses or theories. Their self-intuition and introspection are not observable and what seems grammatical may be refuted by others since intuition may be influenced by linguists' language or dialect background (Xiao & McEnery, 2004). One of the reasons that Smith (1991) and Tai (1984)'s theories are challenged is that some of their self-created examples are not accepted by other studies. Here is an example edited from Tai (1984):

(5) 他写了一封信，但是没写完。

ta1	xie3	le	yi4	feng1	xin4	,	dan2shi4	mei2	xie3	wan2	.
He	write	LE	a	CLA	letter	,	but	not	write	WAN	.

'He wrote a letter but didn't finish it.'

For many studies such as Teng (1985), sentence (5) is not acceptable because *xie2-le yi4-feng2 xin4* (wrote a letter) implies that the letter has been finished and thus is not able to be cancelled by *mei2 xie3-wan2* (not finish it).

The grammaticality of this sentence is not within our concern but this sentence does illustrate the debates on conclusions based on intuition.

What's more, it's not uncommon to bypass or leak contradictory evidences only by intuition. As mentioned in section 1.1, Verkuyl (1972) actually doesn't cover verbs of achievements into his rules when assigning aspectual value.

Although his rules work well for activity and state verbs, achievements such as *kill* refute his rules. The omitted ones only stand for a part of the untouched examples and in an era of big data, more contradictory evidences are leaked in the corpus.

One person's intuition or introspection is not observable and thus is hard to be validated if linguistics doesn't want to be dropped into the same box with philosophy or literature which don't need to follow scientific procedures (Eddington, 2009). Thus, the conclusion based on the validation of intuition or introspection is unreliable.

Tai (1984) insists that ZAI is not able to co-occur with statives based on his self-created examples. However, when tested in Sinica corpus (K.-J. Chen, Huang, Chang, & Hsu, 1996), 15% of stative verbs have counter-examples including *ai* (love), *danxin* (be worried), *xihuan* (like) and *jianchi* (insist) etc.

Guo (1993) merges ZAI and ZHENGZAI to be one aspectual marker based on their meanings. However, statistical validation shows that their distributions are significantly different based on the result of multinomial logistic regression analysis.

Compared with intuition and introspection, the construction of corpus jumps a big leap beyond all questions for the scientific procedures of linguistics because it makes language data observable and thus can be validated. Corpus-based studies begin to become fashionable for linguistics. The methodological

paradigm shifts from intuition to corpus or other observable language data. While accepting Xiao and McEnery (2004)'s statement that corpus-based approach circumvents some disadvantages of intuition-based approach, it should also be noticed that simply counting examples in corpus, as an methodology, is still unreliable.

In corpus-based studies of traditional linguistics, the validation of theories or hypotheses is mainly applied by counting examples in corpus instead of by statistical methodology which is stressed to be a critical part of scientific procedures. Two flaws are inevitable for counting examples methodology.

Firstly, the tendency shown by simply counting examples in corpus may not be tenable. On one hand, manually counting examples sometimes do limit the corpus size one can handle while computational or statistical methodology usually refer to a much bigger corpus because small corpus implies high possibility to mislead linguists to the wrong direction.

Xiao and McEnery (2004)'s work on GUO based on corpus illustrates this problem well. Below follows the distribution of GUO in Xiao and McEnery (2004)'s corpus:

Table 9 The distribution of GUO in Xiao and McEnery (2004)'s corpus

	State		Activity		Transition		Total
	IIS	SLS	ACT	SEM	ACC	ACH	
Frequency	7	1	25	1	36	14	84
		8		26		50	
Percentage		10%		31%		60%	100%

It seems obvious that GUO has a tendency to co-occur with transitions since they account for more than a half of the whole sentences that taking GUO. However, by using an unbalanced and small corpus, Xiao and McEnery (2004)'s tendency is quite untenable.

It will be shown in chapter 4 that the difference between states and transitions on GUO is not significant and the odds ratio calculated through multinomial regression analysis actually shows that GUO has positive relation with activities and negative relation with states and transitions which is quite different from Xiao and McEnery (2004)'s tendency.

The corpus size of Xiao and McEnery (2004)'s study is only about 10,000 Chinese characters while our analysis is on Sinica corpus which has a size of more than 10 million words instead of characters.

On the other hand, the difference shown by simply counting example may not be significant. Compared with corpus size, the lack of validation based on statistical analysis is more dangerous for making decisions or inferring rules from

the frequencies by simply counting examples. High frequency doesn't have to lead to solid tendency. This can be illustrated by the POS distributions of different situation types of verbs in our dataset.
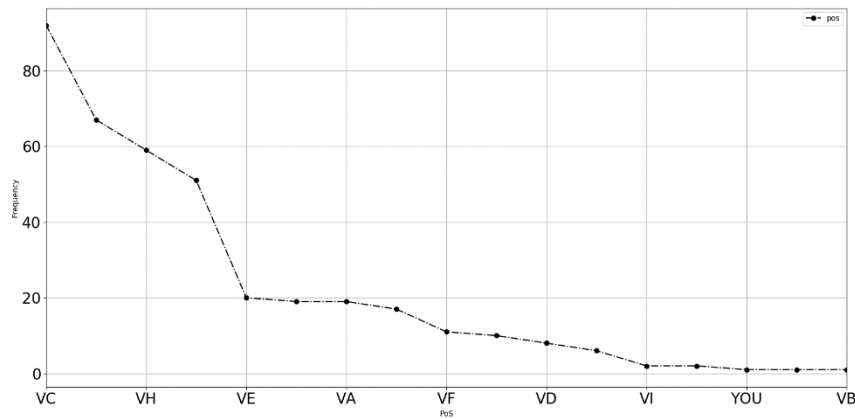
Figure 1 POS distribution of states in Sinica corpus

The falling curve evidently portrays the tendency of POS distribution for states verbs. VC (active transitive verb)[1] accounts for a relatively high portion of all the statives. It naturally follows that VC is supposed to be a feature of statives. It should be noticed that this is not the case because VC actually also accounts for a large part of transition verbs (accomplishments and achievements):

---

[1] Appendix introduces POS information of Sinica corpus in detail.
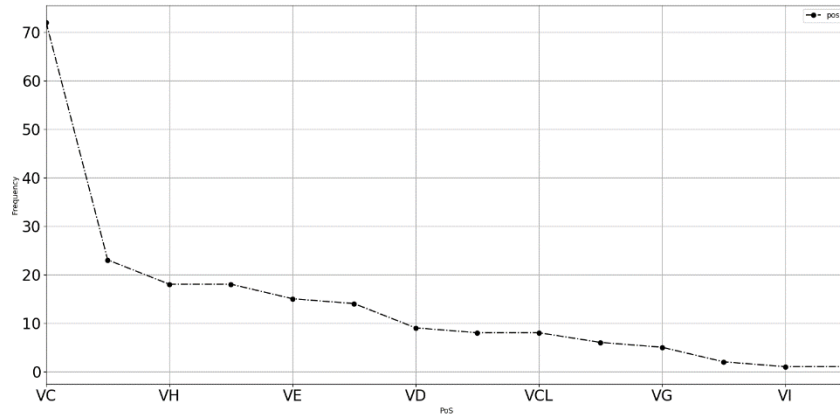
Figure 2 POS distribution of transitions in Sinica corpus

Without statistical validation, simply counting examples may lead us to that VC must be the feature of both states and transitions. However, the distribution difference for transition and state on VC is not significant (z test, p-value = 0.354, coefficients of multinomial regression). On the contrary, the high p-value actually denies that VC can be used to distinguish these two situation types.

Secondly, it's hard to give quantitative or precise analyses based on counting examples. Counting examples is only able to help with the qualitative analysis since only frequencies or percentages can be observed while quantitative analysis requires statistical analysis. Frequencies or percentages, although are numbers, but cannot quantify the relationship between features and categories.

For aspectual studies in Chinese, the relationship between situation types and viewpoint aspect is positioned in the central area. However, most of their descriptions are qualitative with a matrix showing whether a situation type is able

38

to co-occur with a certain viewpoint aspects or not. None of them is able to give quantitative depiction of the relationship. For example, all studies confirm that activities tend to be able to co-occur with LE1 and ZAI but they are not able to answer the following questions:

(i) To what extent does ZAI influence the probability of one situation type to be activity?

(ii) Are LE1 and ZAI in equal position on their ability to distinguish activities from other situation types? If not, how to measure their difference?

Without statistical analysis, there is no way to give answers to these questions. Computational studies are able to give answers but this is not their focus and thus no computational studies concern this problem.

Computational studies are fundamentally different from linguistics studies on their targets. Computational studies focus on practical and applicable task. It is obvious that only several areas of linguistics attract computational studies to persistently develop and improve different algorithms such as name entity recognition, word relations etc. (Baroni, Dinu, & Kruszewski, 2014; Mikolov, Chen, Corrado, & Dean, 2013; Mikolov, Le, & Sutskever, 2013; Mikolov, Sutskever, Chen, Corrado, & Dean, 2013; Mikolov, Yih, & Zweig, 2013; Turney & Pantel, 2010). These topics are highly relevant to information retrieval, artificial intelligence, and machine translation etc. which have application value in engineering. Computational studies may search help from linguistics studies

39

but not with an aim to validate linguistics theories. What they want to do is to achieve higher performance.

Borrowing Frederick Jelinek's popular words, which state that each time he fire a linguist, his performance gets up, to illustrate the embarrassing status of linguistics. However, Haitao Liu and Lin (2017), together with other computational linguists actually imply that linguistics is an appendage of computational linguistics which is strongly opposed in this thesis.

The existence of linguistics studies is not dependent on the application of their theories by computational studies and the value of linguistics studies is not decided by its benefits toward computational applications. According to our knowledge, the goal of linguistics is to probe into the very nature of language structures and to give better explanations for certain language phenomenon. Statistical methodology is stressed within this thesis not because that it can help to achieve computational goals such as automatic prediction and classification but because that it is able to make linguistics conclusions more tenable and reliable.

The difference of goals for computational studies and linguistics studies decides that their focuses are diverse. Taking Mandarin aspectual study as an example, linguistics studies concerns the relationship between [Telic], [Durative], [Dynamic] and situation types but computational studies don't; linguistics studies concerns about the difference of ZAI and ZHENGZAI, but computational studies

don't; linguistics studies concerns about the syntactic and semantic status of verb-reduplication but computational studies don't. Computational studies are not likely to step such further into linguistics topics which seem to be helpless for improving their algorithms' performance. And thus it is natural to find that most of computational studies on aspect, including Mandarin and other languages, are only about the prediction of situation types. Even for prediction and classification studies on aspect, the number is very limited.

The proper way to treat computational methodology for linguistics studies is to position it together with statistical methodology as a scientific procedure to validate linguistics hypotheses. This means linguistics topics and sub-topics, instead of the application or realization of prediction or classification, are supposed to be the principle line and goal. We don't care about the performance of a certain computational algorithm only if it can help to validate a linguistics hypothesis.

1.2.3 Research questions

As is illustrated in the previous sections, lexical aspect, to a large extent, is shadowed in aspectual studies. Although a few papers focus on the interaction between lexical aspect and aspectual markers, they fail to validate and quantify the contribution from aspectual markers toward differentiating distinct situation types. Our research question is raised under this circumstances.

Firstly, are those aspectual markers including LE1, LE2, ZHE, ZAI, ZHENGZAI and GUO, which stands for different viewpoint aspects, able to differentiate between different situation types including state, activity and transition?

Secondly, how do these viewpoint aspects interact with these situation types?

Thirdly, can we quantify the contributions of these viewpoint aspects toward differentiating different situation types?

Fourthly, is it able to predict the situation type of a verb or verb's sense based on aspectual markers? Are there any other features that can help to improve the prediction precision of the situation type of verbs' senses?

This thesis tries to answer the above questions mainly via statistical and computational methodologies.

1.3 Organization of the thesis

Chapter 1 focus on the definition and nature of aspect. Aspect refers to the internal temporal structure of a situation and it contains situation type and viewpoint aspect. While situation type is a categorization based on the internal temporal structure of verbs, phrases and sentences, viewpoint aspect serve as a camera to observe the internal temporal structure of a situation. Situation type is compositional and recursive in nature which is like syntactic units. While verbs have their own situation types, their combinations with other elements as a whole have their own situation type. The constituting elements' situation type or

42

aspectual value is independent from the whole structure just like a verb's syntactic category is independent from the whole structure's syntactic categories.

Chapter 2 reviews previous studies on the classification of situation types and viewpoint aspect in Mandarin Chinese. Other studies such as the definition and nature of aspect are reviewed in the corresponding chapters.

Chapter 3 annotates the interaction between the selected verbs senses and viewpoint aspect by our own intuition and cross-checks the annotation based on other annotated language resources. The situation type system is clustered from bottom to up based on the annotation according to the verbs senses distances with each other in vector space following the prototypical theory.

Chapter 4 validates the interaction between situation type and viewpoint aspect in statistical way. On one hand, the interaction between situation types and viewpoint aspects is validated and quantified by multinomial logistic regression analysis. On the other hand, the relationship between the conceptual features of [Durative], [Telic], [Dynaminc] etc. and viewpoint aspect are constructed based on the statistical results.

Chapter 5 evaluates the dataset of verbs senses' situation types. Since the situation type system is constructed based on supervised features by hierarchical clustering, only features extracted in unsupervised way and semi-unsupervised way are used to evaluate the dataset. The accuracy of 72.05% achieved by supporting vector machine based on word embedding vectors proves that our

dataset can be directly used for the evaluating and training of models for predicting verb or verbs' senses situation type.

Chapter 6 concludes the whole thesis and depicts our future work for and based on the situation type dataset.

Chapter 2

Literature review

This chapter covers a part of literatures on Mandarin aspect. Studies on the definition and nature of aspect have been reviewed in the first chapter and thus they are skipped over in this chapter. Computational studies on aspect is reviewed in chapter 5. There are also literatures on other topics such as vector space model. They will be mainly reviewed in chapter 4 and 5 to keep consistent in content.

This chapter mainly review studies on the situation type's classification and the interaction between situation types and adverbials in Mandarin Chinese including Tai (1984), Teng (1985), P. Chen (1988), Smith (1991), He (1992), Zhang (1993), Guo (1993, 1997), Dai (1997), Xiao and McEnery (2004) and Xu (2015).

Tai (1984) serves as one of the pioneering studies on Mandarin aspect following Vendler (1967)'s classification of English verbs. Quartered situation types collapse into three including state, activity and result. While state and activity are almost equivalent to their English counterparts, accomplishments are different from English equivalents.

Tai (1984) consists that there are no accomplishment verbs or phrases in Mandarin similar with English ones. Accomplishments like *write a letter* contain both the action and also the attainment of the goal in English. But Tai (1984)

holds that phrases like *write a letter* in Mandarin doesn't have to imply the attainment of the goal:

(6) a.我昨天画了一张画，可是没画完。

wo3	zuo2tian1	hua4	le	yi1zhanng1	hua4	,	ke3shi4	mei2	hua4	Wan2	.
I	yesterday	paint	LE	a	picture	,	but	not	paint	WAN	.

'I painted a picture yesterday, but I didn't finish it.'

b.我昨天写了一封信，可是没写完。

wo3	zuo2tian1	xie3	le	yi1feng1	xin4	,	ke3shi4	mei2	xie3	Wan2	.
I	yesterday	write	LE	a	letter	,	but	not	write	WAN	.

'I wrote a letter yesterday, but I didn't finish it.'

Sentence (6a) and (6b) illustrates that the implication of attainment of goal can be canceled and thus the counterparts of English accomplishments don't guarantee the attainment of goal.

Tai (1984) further specifies that RVC is resorted to insure the attainment of goal such as *learn* is mapped to *xue2-hui4* and *kill* is mapped to *sha1-si3*. However, even for RVC, Tai (1984) also denies their equivalence to English accomplishments because RVC is not able to co-occur with progressive aspectual marker such as ZAI:

(7) a.*我在学会中文。

wo3　zai4　xue2　hui4　　zhong1wen2

I　　ZAI　learn　acquire　Chinese

'I am learning Chinese.'

b.*我在杀死张三。

wo3　zai4　sha1　si3　　Zhang1san1

I　　ZAI　kill　　dead　Zhangsan

'I am killing zhangsan.'

While activity verbs such as *xue2* (study) or *sha1* (murder) are able to co-occur with ZAI, RVCs in (7a) and (7b) are not able to.

Tai (1984) is on the right track to merge accomplishments and achievements into result which is renamed as transition in this thesis. However, his argumentation is untenable.

Firstly, (6a) and (6b) with Vendlerain accomplishment phrases are arguably accepted to be a proof that the attainment of goal doesn't have to be guaranteed in Mandarin accomplishments. Teng (1985) points that both (6a) and (6b) are not grammatical in Mandarin Chinese. The controversy on this problem displays the

fragileness of intuition based validation on hypotheses. While some native speakers accept (6a) and (6b), others don't. Since intuition is not observable, there is no way to persuade the opposite side by quarrelling with each other.

Secondly, not all RVCs are not able to co-occur with progressive and durative aspectual markers. For example, *gai3bian4* (change) is able to co-occur with ZAI, ZHENGZAI and ZHE in Sinica corpus:

(8) a.然而这个现象正慢慢在改变。

ran2er	zhe4ge	xian4xiang4	zheng4	man4man4	zai4	gai3bian4	.
however	this	phenomenon	ZHENG	slowly	ZAI	change	.

'However, this phenomenon is changing slowly.'

b.我们的国家正在改变。

wo3men	de	guo2jia1	zheng4zai4	gai3bian4	.
our	DE	country	ZHENGZAI	change	.

'Our country is changing.'

c.因此朋驰不断的改变着汽车界的历史。

yin1ci3	peng2chi2	bu2duan4	de	gai3bian4	zhe	qi4che1jie4	de	li4shi3	.
Thus	Benz	continuously	DE	change	ZHE	motor-dom	DE	history	.

'Thus, Benz is changing the history of motor-dom continuously.'

There are also other accomplishments or achievements which can co-occur with progressive and durative aspectual markers and the interaction with ZAI,

48

ZHENGZAI and ZHE is always not consistently. For example, the co-occurrence frequency for *xing2cheng2* (form) with ZHENGZAI is high while its co-occurrence frequency with ZHE is zero. And even for the same aspectual marker, different words may have different co-occurrence with it which can be ranked in descending order.

By confirming that "resultative verb compounds and resultative simple verbs can be considered to belong to one single category" which is result, Tai (1984) points out that both of them have "a definite point" while states and activities don't. However, verbs like *gai3bian4* (change), *zeng1jia1* (increase) and *xing2cheng2* (form) are all able to co-occur with progressive and durative aspectual markers. There is no question that these verbs refer to change which definitely pushes them into transition or result, but it seems that durative is preferred to be a gradual-change feature instead of a binary one.

Tai (1984) is on the right track to treat phrases and verbs individually and he actually accepts the resursiveness of situation types since the situation type of phrases such as *write a letter* hasn't been assigned to its constituting verb.

The interaction between situation types and aspectual markers and other adverbials can be summarized as below:

Table 10 The interaction of situation types and adverbials in Tai (1984)

	State	Activity	Result
HEN + V	+	-	-
ZAI + V	-	+	-
V + LE	(-)	+	+
TP (in-adverbial) + V	(-)	-	+
V + TP (for-adverbial)	(-)	+	-
hua-le(took) + TP + V	(-)	-	+

Most of aspectual studies use minus or plus to refer to the ability of one situation type to co-occur with a certain aspectual marker or other elements. The value is basically judged through intuition. However, the emergence of large corpus actually shatters these kind of ideal hypotheses. The distribution of situation types is hard to be simulated by such an ideal matrix.

It's easy to find counter-examples in corpus. On one hand, the co-occurrence of activities and HEN can be found in Sinica corpus when the above table denies this possibility:

(9) a. 大医院对这样的事情也很支持。

da4 yi1yuan4 dui4 zhe4yang4 de shi4qing ye3 hen3 zhi1chi2 .

big hospital for such DE thing also very supportive .

'Big hospital is also very supportive for such things.'

b. 她也很注意外表。

ta1 ye3 hen zhu4yi4 wai4biao3 .

she also very pay attention to appearance .

'She also pays a lot of attention to the appearance.'

50

On the other hand, some result or transition verbs, which "must co-occur with the inchoative aspect marker -le" (Tai, 1984) cannot be found in Sinica corpus.

Teng (1985) opposes Tai (1984) on two basic holdings although both of them claim to follow Vendler (1967)'s situation types.

Firstly, he confirms that situation type is the classification of the predicate instead of the verb in a sentence. However, he further points outs that verbs which appear in activities and accomplishments are action verbs, the ones appear in transitions are transition verbs and the rest ones which appears in states are stative verbs. He confirms the equivalence between verbs' classification and situation type but stresses that they are clustered based on different principles although share the similar type names. But what on earth is the difference is not illustrated in his paper.

Tai (1984) is on the right track to implicitly accept the recursiveness or hierarchy of situation types while Teng (1985) denies it. By intangibly insisting that the classification of verbs follows different principles from situation types, Teng (1985) excludes the existence of lexical aspect.

Teng (1985) stresses that the situation type of a sentence cannot be decided only by its constituting verb. However, Tai (1984) never deny this although he is studying the verb's situation type. To confirm the possibility of classifying verbs into different situation types doesn't mean to deny other elements' contribution

51

to the aspectual class of the whole structure. Teng (1985) actually assumes that the compositionality of aspectual information has been rejected by Tai (1984) which is not the truth.

Secondly, Teng (1985) insists to quarter the situation type system which collapse into three in Tai (1984)'s theory. Tai (1984) holds that the counterparts of English accomplishments in Mandarin behave variously and merge them with achievement into result while Teng (1985) keeps consistent with Vendler (1967) and claims that Chinese also has accomplishments.

Shelving the dispute on the ungrammaticality of sentence (5a) and (5b), Teng (1985)'s argumentation of the existence of accomplishments in Mandarin Chinese is untenable.

According to corresponding examples in Teng (1985), he makes a more radical assumption that all the sentences in which RVCs appear are accomplishments.

(10) a.*他洗好衣服洗了半个小时。

    ta1   xi3   hao3   yi1fu4   xi3   le   ban4   ge4   xiao3shi2 .

    he   wash HAO clothes wash LE half CLA hour   .

    'It took him half an hour to finish washing all of these closes.'

    b.他洗好衣服半个小时了。

    ta1   xi3   hao3   yi1fu4 ban4   ge4   xiao3shi2 le   .

he	wash	HAO	closes	half	CLA	hour		LE	.

'He has finished washing these closes for half an hour.'

Teng (1985) maintains that accomplishments cannot be durative and thus the *half an hour* in (10a) cannot be used to specify the durative time of *xi3-hao3* (finish washing) but to be used to express the lapse time after of *xi3-hao3* (finish washing) in (10b).

However, no aspect studies hold that accomplishments are non-durative since the very start of Vendler (1967) and thus the conclusion that accomplishments are not durative of Teng (1985) is unsupported. His analysis of *xi3-hao3* (finish washing) is on the right track but it should be noticed that *xi3-hao3* (finish washing) is achievement instead of accomplishment.

He further illustrates that the ability to co-occur with *yi1xia4* (all of a sudden) and *ma3shang4* (at once) proves that the following sentences are accomplishments:

(11) a.他一下就煮好饭了。

ta1	yi1xia4	jiu4	zhu3	hao3	fan4	le	.
he	all of a sudden	boil	HAO	meal	LE	.

'He finished cooking the meal all of a sudden.'

b.他马上就写好三封信了。

ta1   ma3shang4   jiu4   xie3   hao3   san1   feng1   xin4   le   .

he   at once   write   HAO   three   CLA   letter   LE

'He finished writing three letters at once.'

Regardless of the weirdness of (11a) and (11b), *yi1xia4* (all of a sudden) and *ma3shang4* (at once) illustrate that (11a) and (11b) are punctual instead of durative. Both (10) and (11) actually confirm that these sentences with RVCs are achievements instead of accomplishments which contradicts Teng (1985)'s own assertion.

Teng (1985) also uses *in*-adverbials to test accomplishments:

(12) a.他五分钟吃了二十个饺子。

ta1   wu3   fen1zhong1   chi1   le   er4shi2   ge4   jiao3zi   .

he   five   minutes   eat   LE   twenty   CLA   dumplings   .

'He ate twenty dumplings in five minutes.'

b.*他五分钟写了信。

ta1   wu3   fen1zhong1   xie3   le   xin4   .

he   five   minutes   write   LE   letter   .

'He wrote letters in five minutes.'

It is true that (12) illustrates the difference between accomplishments and activities. However, achievements are also able to co-occur with *in*-adverbials such as *he died in five minutes*. Thus, *in*-adverbial along cannot distinguishes accomplishments from achievements.

Teng (1985) deviates a lot on striving to map all RVCs into accomplishments. Teng (1985)'s assertion that accomplishments are punctual actually makes accomplishments almost the same with achievements on all of the Vendlerian features. To distinguish these two situation types, he specifies that accomplishments stress the attainment of goal while achievements focus on a transition from one state to another. However, accomplishments also imply transition.

What's more, RVCs, although specify the attainment of goal by V2, behave similar to the so-called achievements instead of accomplishments like *draw a circle*. Teng (1985) actually implicitly omits the feature of [Durative] which can be used to distinguish accomplishments and achievements in Vendlerian situation types and uses attainment of goal as a standard to differentiate them, which makes his principles different from Tai (1984). His examples are not able to illustrate the difference between accomplishments and achievements. On the

contrary, they confirm accomplishments' similarity with achievements, which makes his quartered situation types questionable at least in lexical level.

P. Chen (1988) breaks achievement into simple change and complex change and subdivids state into three types including attribute or relation verbs (type Ⅰ) such as *shu3yu2* (belong to), mental or physical state verbs (typeⅡ) such as *xi3huan1* (like) and location verbs (type Ⅲ) such as *zuo4* (be seated):

Table 11 The interaction of situation types and adverbials in P. Chen (1988)

	State			Activity	Accomplishment	Complex Change	Simple Change
	Ⅰ	Ⅱ	Ⅲ				
HEN + V	-	+	-	-	-	-	-
V + ZHE	-	-	+	+	+	-	-
ZAI + V	-	-	-	+	+	+	-
V + LE + TP (action duration)	-	-	-	+	+	-	-
V + LE + TP (state duration)	-	+	+	-	+	+	+

Broadly speaking, state verbs cannot co-occur with HEN and ZHE according to previous studies. But P. Chen (1988) is suspicious of this and moves a further step. He pints out that state can be subsectioned and some of them can co-occur with HEN and ZHE such as *ai4* (love) while some other cannot. He (1992) also shares the same view.

Achivements are further divided into complex change and simple change. Resultative verb compunds such as *gai3liang2* (improve), which is able to co-occur with ZAI or ZHENGZAI, are counted as comlex change in P. Chen (1988).

Although, the co-occrance with ZAI or ZHENGZAI indicates these resultative verb compunds are durative but P. Chen (1988) holds that verbs like *gai3liang2* (improve) or *sheng1qi3* (rise) are harly to keep a relatively stable status.

Resultative verb compunds which cannot co-occur with ZAI such as *da3po4* (break) and punctual verbs like *si3* (die) are treated as simple change.

P. Chen (1988)'s work is excellent and inspiring for Mandarin aspectual studies and his analysis is in details and subdivides the four classical situation types according to their co-occurance with linguistics features and his discovery shows that even for verbs within the same situation type, they behave variously on linguistics features and also that linguistics features don't distribute consistantly within one spercific situation type. For example, boradly speaking, states cannot co-occur with ZHE, but there are stative verbs which can co-occur with ZHE; on ther other hand, resultative verb compunds used to be put together as results in some aspectual studies but actually some of them cannot co-occur with ZAI while others can.

Smith (1991) develops two components theory and insists the independency of situation type and viewpoint aspect with each other. Mandarin Chinese, as a typical aspect language, has been used to test her theory in detail.

There is no doubt that Smith (1991) serves as one of the most influential model for aspect and we also follow her two components theory in this thesis. However, her model is not always without any flaws.

In order to escape the polysemy or double entering collision of verbs' situation type, Smith (1991) hooks situation type on verb constellations which contain verbs and their arguments and insists that situation type cannot be assigned in the lexical level

However, she is self-contradictory because aspectual values, on which situation types are based on, are assigned to verbs. Smith is very clear about the necessity of assigning aspect values to elements in a verb constellation since she claims that "to state compositional rules, we must develop a procedure for giving verbs, nominals etc. aspectual values" and these values "will be entered in the lexicon" (Smith, 1991). As reviewed in previous sections, this is how Verkuyl (1972) does to formulate the compositional rules for aspectual information.

But there is no fundamental difference between aspectual values and situation types. According to Smith's procedures, aspectual values refer to [Telic], [Durative] etc. which are exactly the same features for classifying situation types and she herself also admits that state, activity, accomplishment and achievement actually "function as shorthand for the cluster of properties that distinguish the situation types" with property referring to [Telic], [Durative] etc. (Smith, 1991).

Another self-contradictory statement of Smith (1991) is her denying of the independence of verbs situation type or aspectual value from the verb constellation. In Smith's procedures, the verb constellation's aspectual value is assigned to its constituting verb.

(13) Ellen drew a picture.

The *draw* in (13) has been assigned a value of "$_v$[+Telic]" in Smith (1991). However, she refutes herself on both the lexical level and syntactic level.

(14) a. 他们吃饱了。

　　　ta1men2　chi1　bao3　le　　.

　　　they　　　eat　　full　　LE　.

　　　'They ate their fill.'

　　　b.他写了一封信。

　　　ta1　xie3　le　　yi1　feng1　xin4　　.

　　　he　write　LE　one　CLA　　letter　.

　　　'He wrote a letter.'

According to Smith (1991), in (14a), *chi1bao3* (eat one's fill), as an compound, is composed of an atelic process and its outcome. However, Smith's procedure is clear that verb constellation's aspectual value should be assigned to its constituting verb which means *chi1bao3* (eat one's fill)'s aspectual value should be assigned to *chi1* (eat). And thus, *chi1* should be [+Telic] instead of [-Telic] regardless of her untenable claiming that *chi1bao3* (eat one's fill) is an accomplishment.

Similar problems occur in (14b) which is an accomplishment phrase instead of compound. *xie3* (write) should be assigned [+Telic] value following the whole structure which is an accomplishment but Smith call the verb atelic process.

It can be found in most of Smith's discussion on accomplishments that she tries to use process instead of activity to refer to the main verb. However, there is no fundamental difference between a process and an activity.

It has been argued in the previous chapter that taking situation type away from lexical level is not possible to solve the polysemy problem of verb's situation type because the root of the polysemy is the rejection of the independence of verb's situation type from the whole structure and the recursiveness of situation type or aspectual value.

What's more, the self-contradictory examples in (13) and (14) actually prove that even she herself is hardly persuaded to assign [+Telic] value to activity verb such as *draw* in accomplishments such as *draw a circle*.

We agree with Smith (1991) on assigning aspectual value to verbs, nominals, adverbials etc. no matter in lexicon or by other methods. However, she clearly differentiate situation type from aspectual value which, as far as our concern, is questionable especially for verbs. She is just using different terminologies to refer to the same thing.

One of the most important contributions of Smith (1991), which is shared by Guo (1993), Zhang (1993), Xiao and McEnery (2004) and us, is to explain

situation types by prototype theory which holds that "human categories and concepts have prototypical rather than hierarchical structure" (Smith, 1991). Family resemblance is borrowed to simulate the relationship between typical and atypical members of a category.

As the core of prototype theory, the semantic distance and clustering between members and categories are supposed to be implemented. However, none of Smith (1991), Guo (1993), Zhang (1993) and Xiao and McEnery (2004) really put them into action although Guo (1993) proceeds a lot by using intuition-based methodology. We will discuss it latter in chapter 3.

Although accepting the prototype theory, Smith (1991) doesn't position all of the situation type onto a continuum as Zhang (1993) and Guo (1993) but only stress the transition from typical to atypical within a category. However, by confirming that "it is not always clear whether an event should be classified as an accomplishment or achievement……", Smith (1991) implicitly realize that situation type may form a continuum.

Five situation types are identified in Smith (1991) including state, activity, semelfective, accomplishment and achievement.

For statives, *cun2zai4* (exist), *qian4* (owe), *shu3yu2* (belong to) etc. are all verbs instead of verb constellations. Theoretically, Smith keeps stressing that situation type doesn't fit into lexical level. However, in actual operation, she

61

implicitly applying situation type into lexical level, which makes her analysis untenable, inconsistent and confusing:

(15) 他们知道了什么叫科学。

ta1men2　zhi1dao4　le　shen2me　jiao4　ke1xue2　.

they　　　know　　LE　what　　call　science　.

'They have learnt what is called science.'

Smith (1991) uses (15) to illustrate that "Stative constellations in Mandarin have stative verbs……" However she is contradicting herself by using this example. Sentence (15) is definitely not a stative sentence even in Smith's own standards. *zhi1dao4*-LE1 (have learnt), as a whole, is an inchoative which "belongs to achievement or accomplishment" (Smith, 1991).

The only reason that she calls (15) stative is that *zhi1dao4* (know) is a stative verb. However, she denies the possibility of assigning situation type to verbs. It then follows that *zhi1dao4* (know) can only be assigned aspectual value. According to Smith (1991)'s procedure, if the verb constellation gains a situation type with [Telic] feature, the constituting verb will received the same feature with the whole structure. Since (15), as a whole, is an accomplishment or achievement, *zhi1dao4* (know) should be assigned [+Telic] value and thus cannot be called a stative.

Sentence (15) is neither the first nor the last one in Smith (1991) showing the chaos of herself for the situation types' bearing unit. It is quite clear that there is no way to bypass the classification of verbs into situation types if the interaction between verbs and other elements are about to depict. To assign the whole structure's situation type to its constituting verb is not applicable and also violates the language intuition of even the linguists themselves.

For accomplishment and achievement, Smith (1991) introduces natural end point and arbitrary endpoint to distinguish them from activity and state. While telic situation types have natural end point, atelic situation types have arbitrary end point. Telic situation types are toward a goal and natural end point is used to "escape the agentive connotations of goals" (Smith, 1991). Contrary to the "intrinsic goal" of telic situation types, atelic situations is able to stop at any time and thus is assumed to have an arbitrary end point (Smith, 1991).

It then follows that goal plays a key role in distinguishing telic situations from atelic ones. Regardless of the properness of "intrinsic goal" for achievement which expresses a transition from one state to another, the relationship between RVC and telic situation types is untenable for Smith (1991).

Accomplishments are portrayed as "situations consisting of a process and an associated outcome" (Smith, 1991). RVC, which contains two verbal elements with the former one serves as an action and latter one as an outcome, matches this standard properly:

(16) 他们吃饱了。

ta1men2   chi1   bao3   le   .

they   eat   full   LE   .

'They ate their fill.'

With *chi1* (eat) as an process and *bao3* (full) as an outcome or goal, (16) is created in Smith (1991) as an example of accomplishment. However, accomplishments are durative while RVCs like *chi1bao2* (eat one's fill) are not. *chi1bao2* (eat one's fill) is hard to be durative since it cannot co-occur with durative adverbials no matter in corpus or based on intuition. *chi1* (eat) is durative is not able to be detached from *chi1* (eat) from *chi1bao2* (eat one's fill) which is a compound. *chi1bao2* (eat one's fill) is instantaneous instead of durative.

Similar problem exists in Teng (1985) who deviates even further to includes all of RVCs into accomplishments while Smith (1991) includes part of them. The root that results in the collision is to apply double standard to identify accomplishments. On one hand, features of [+Telic] and [+Durative] are used to distinguish accomplishments from other situation types; on the other hand, structure with a process and its outcome is also considered to be accomplishments. However, these two standards are not consistent with each other all the time. RVC plays as a violation of them. While containing a process and an outcome,

some RVCs are durative such as *kuo4da4* (enlarge) while others are instantaneous such as *chi1bao2* (eat one's fill).

RVC itself serves as an proof that the distinction between accomplishment and achievement is not clear which is also shared by Smith (1991) and that's one of the reasons why Tai (1984) merges them into result.

He (1992), although claims that situation types are supposed to be realized in sentential level, distinguish lexical situation type and sentential situation type consistently with pertaining the independence of verb's situation type from the whole structure's which is very similar to Verkuyl (1972).

Situation types collapse into three but each situation type are subdivided into more branches than previous studies in He (1992). In his study, He (1992) merges activity and accomplishment into durative while both RVC and achievement are dropped into punctual or change of state situation type:
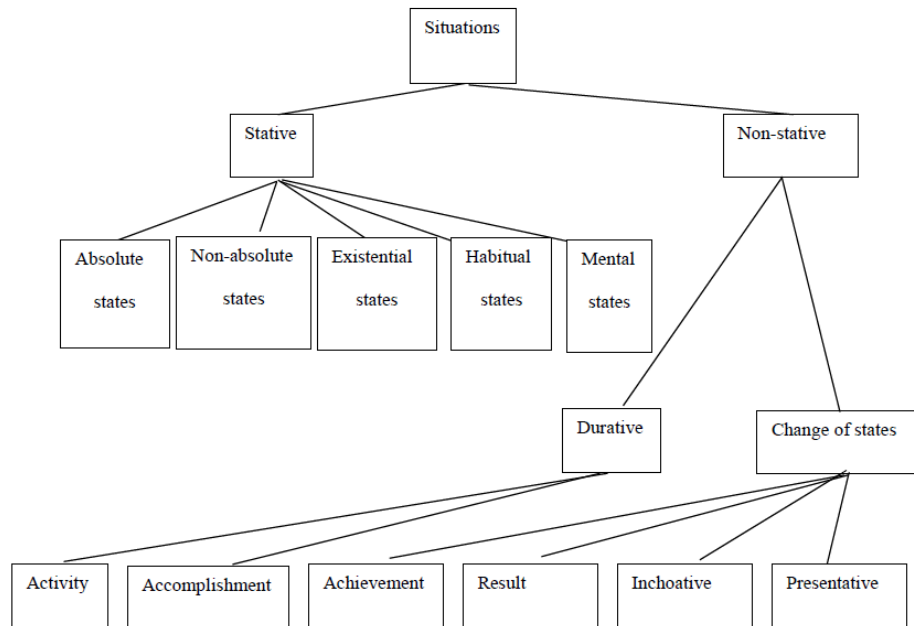
Figure 3 He (1992)'s situation type hierarchy

For states, five categories are distinguished. Absolute states and non-absolute states partition states based on that whether a state can co-occur with HEN or not. While Tai (1984) and Smith (1991) simply list the co-occurrence with HEN as an distinguished feature for states, He (1992) and P. Chen (1988) step further since they found that not all states are able to co-occur with HEN and formulize rules for this phenomenon. However, their intuition, although incisive, is challenged by corpus data.

Firstly, absolute states are able to co-occur with HEN. Several typical absolute states can co-occur with HEN in a very high frequency such as *you3* (have), *xiang4* (resemble) and *shi4* (be):

66

(17) 这一刻很是矛盾。

zhe4   yi1   ke4   hen3   shi4   mao2dun4   .

this   one   moment   very   is   ambivalent

'He is very ambivalent at this moment.'

Their rankings of co-occurrence frequency are of the highest ones compared with other non-absolute states. Among them, *you3* (have) ranks in the first place with a co-occurrence frequency of 394.

Secondly, most non-absolute states cannot be found to co-occur with HEN. Most states cannot be found to co-occur with HEN in Sinica corpus and only 31% of the stative verbs' co-occurrence frequency is higher than 1.

As we have stressed in the previous chapter, counting examples and their percentages are not reliable to validate rules or hypotheses.

Another problem for state's classification is that the subtypes are not mutually exclusive. Mental states overlaps a lot with non-absolute states such as *xi3huan1* (like).

What's more they are not even in the same level. While absolute states, non-absolute states and mental states refer to verb's classification, existential states and habitual states refers to sentences or phrases:

(18) a.床上躺着一个病人。

chuang2　shang4　tang3　zhe　yi1　ge4　bing4ren2　.

bed　　　top　　lie　　ZHE　one　CLA　patient　　.

'On the bed liess a patient.'

b.他抽烟。

ta1　chou1　yan1　　　.

he　smoke　cigarette　.

'He smokes.'

Similar problems also exist in the situation type of change of state. While achievement and result operate in lexical level, inchoative and presentative work for sentential level.

He (1992) merges activity and accomplishment into durative. Accomplishment is a transitive type between activity (durative or process) and achievement (result or change of states) in most of the aspectual studies. On one hand, they are durative, which is shared by activities and on the other hand, they imply transition or change, which is shared by achievements.

In addition, He (1992) also notices that mental state verbs are vague in their belongings. He argues that "comparatively speaking, the stativity of this group is

68

less strong than the stative verbs…Some of them have a very strong flavor of activity" (He, 1992). Accomplishment and mental state serve as proofs that situation types are not mutually exclusive but form a continuum which is shared by Smith (1991), Zhang (1993) and Guo (1993).

Activities are further divided according to their syllable length and transitivity in He (1992). He (1992) is on the right track to include syllable length into consideration but he doesn't find any tendencies from the classification. Actually, syllable length plays an important role in distinguishing situation types which will be stated in chapter 5.

RVCs are all included into the type of change of state in He (1992) which is shared by Tai (1984) and P. Chen (1988). Both Teng (1985) and Smith (1991) include some RVCs into accomplishment such as *chi1bao3* (eat one's fill). However, it should be noticed that *chi1bao3* (eat one's fill) is judged as an accomplishment not because of its internal temporal structure but because of its "process + outcome" structure which is very similar to *xie2 yi1feng1 xin4* (write a letter). However, they are fundamentally different. On one hand, *chi1bao3* (eat one's fill) is not durative, while *xie2 yi1feng1 xin4* (write a letter) is; on the other hand, *chi1bao3* (eat one's fill) is a compound which is in lexical level while *xie2 yi1feng1 xin4* (write a letter) is in phrase level.

Zhang (1993) makes a contrastive study of English, German and Chinese on aspect. Situation types are not treated as isolated categories in Zhang (1993) and

she insists that they are positioned on a continuum from perfective to imperfective. Trying to maintain their generality in human conceptualization, Zhang (1993) applies perfective and imperfective to both situation type and viewpoint aspect. However, it actually brings in confusion. The importing of terminologies of perfectivation and imperfectivation actually isolate viewpoint aspect from situation type. This is still in consistent with two components theory although by using different terminologies.

Although questioning her terminologies, the continuum of situation type and prototype theory applied in Zhang (1993) are shared by this thesis:

Perfective	A	B	C	D_i	D_{ii}	E	F	G	**Imperfective**
	dao4	shang4	dou3	ben1	dui1	dai4	fa1chou2	jue2de	
	arrive	go up	shake	run	stack	carry	worry	feel	
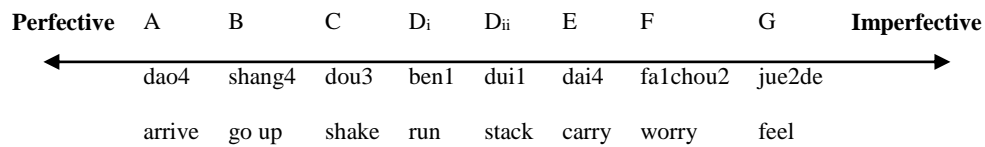
Figure 4 Situation type continuum in Zhang (1993)

According to Zhang (1993), perfective (achievement) and imperfective (sate) are positioned on the two poles of the continuum while activity is between them. Change and boundedness serve as the two basic features to distinguish them. While imperfectives hold their properties without interruption, perfectives experience a transition from one state to another; while boundary is easy to be constructed for a perfective, no boundary exists for a state.

70

Zhang (1993) is neither the first nor the last to realize that situation types are better to be simulated by prototype theory and continuum. Before her, L. M.-j. Huang (1987) proposes that Mandarin verbs can be positioned on a continuum of periodicity:

Short	si3	tiao4	pao3	chuan1	zhi1dao4	Long
Periodicity	die	jump	run	wear	know	Periodicity
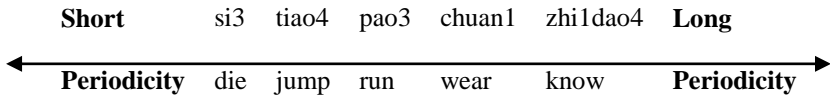
Figure 5 Situation type continuum in L. M.-j. Huang (1987)

Guo (1993) also shares the similar view with L. M.-j. Huang (1987) and Zhang (1993). His study will be reviewed later.

Zhang (1993) also generalizes the interaction between several aspectual markers:

Table 12 The interaction of situation types and aspectual markers in Zhang (1993)

	A	B	C	D	E	F	G
ZHE			+	+	+		
ZAI		+	+	+			
LE	+	+	+	+	+	+	

Guo (1993)'s contribution lies in that he tags all of the verbs in the dictionary of Meng (1987) with situation types.

71

Table 13 The interaction of situation types and aspectual markers in Guo (1993)

Type	Temporal structure	Sub-class	V+LE	V+TP	V+ZHE	ZAI+V	V+GUO	Example
Va	Unbounded		−	−	−	−	−	*shi4*(is)
Vb	I bounded		I	I	−	−	−	*zhi1dao4*(know)
		1	I	I	−	−	+	*xiang1xin4*(believe)
		2	I	I	+	−	+	*you3*(have)
Vc	I&F bounded	3	I,F	I	+	−	+	*zui4*(be drunk)
		4	I,F	I	+	+	+	*deng3*(wait)
		5	I,F	I,F	+	+	+	*chi1*(eat)
Vd	F bounded	1	F	F	+	+	+	*ti2gao1*(enhance)
		2	F	F	−	+	+	*li2kai1*(leave)
Ve	Punctual		F	F	−	−	+	*lai2*(come)

Note: I refers to inception; F refers to final.

The sub classes of verbs are positioned on a continuum of situation types:

State   Activity   Transition

← Va1   Vb   Vc1   Vc2   Vc3   Vc4   Vc5   Vd1   Vd2   Ve →
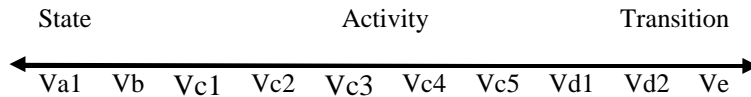
Figure 6 Situation type continuum in Guo (1993)

In Guo (1993), four classical event types collapse into three by merging accomplishments and achievements. Vd1 verbs such as *zeng1jia1* (increase) or *xiao1shi1* (disappear) are much like accomplishments since they are able to co-occur with ZHE and ZAI/ZHENGZAI which indicates that they are durative while Ve verbs are much like achievements because of the lack of ability to co-occur with progressive aspectual markers. Vd2 verbs are between accomplishments and achievements.

Merging accomplishments and achievements into transition is like what Tai (1984) does but with much more solid ground since the situation type and sub

situation type system are constructed from bottom to up by merging verbs into groups in a dictionary one by one. Pustejovsky (1991) also merges accomplishments and achievements into transition.

But latter in Guo (1997), he switches the continuum into a hierarchy:

```
                        ┌─────────────────┐
                        │ Situation types │
                        └─────────────────┘
               ┌───────────────┴───────────────┐
      ┌───────────────┐                 ┌───────────────┐
      │ Stative verbs │                 │ Dynamic verbs │
      └───────────────┘                 └───────────────┘
    ┌────┬────┬────┬────┐           ┌───────────────┬───────────────┐
  ┌────┐┌────┐┌────┐┌────┐    ┌──────────────┐ ┌──────────────────┐
  │ Va ││ Vb ││Vc1 ││Vc2 │    │ Action verbs │ │ Transition verbs │
  └────┘└────┘└────┘└────┘    └──────────────┘ └──────────────────┘
                           ┌─────┬─────┬─────┐ ┌─────┬─────┬─────┐
                        ┌─────┐┌─────┐┌─────┐┌─────┐┌─────┐┌─────┐
                        │ Vc3 ││ Vc4 ││ Vc5 ││ Vd1 ││ Vd2 ││ Ve  │
                        └─────┘└─────┘└─────┘└─────┘└─────┘└─────┘
```
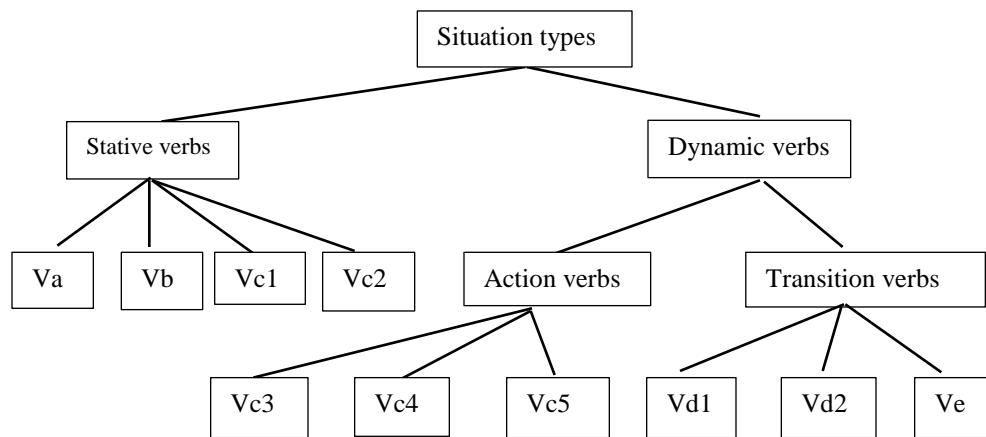
Figure 7 Guo (1997)'s situation type hierarchy

Dai (1997) stresses situation type should be studied on different levels including, at least, sentential level and lexial level. Although these two levels shared the same distinguishing feature, different situation types are idendified. In lexical level, stative, activity and result are firstly classified and statives are further divided into propterty, mental or feeling, posture and position verbs individually:

Table 14 Situation types of verbs in Dai (1997)

		Property	*shi4* (be), *deng3yu2* (eaqul to)
		Mental or	*zhi1dao4* (know), *pa4* (be afraid)
	Stative	Feeling	
Situation types in lexical		Posture	*zhan4* (stand), *zuo4* (sit)
level		Position	*gua4* (hang), *na2* (take)
		Activity	*ti1* (kick), *chi1* (eat)
	Dynamic	Result	*si3* (die), *bian4hao3* (become better)

In setential level, accomplishment is added to quarter sitution type system.

The classification of statives of Dai (1997) is the same with P. Chen (1988). While there are few debates on property and mental verbs, posture and position verbs are not without questions. While Smith (1991), He (1992) and Dai (1997) identify them as stative verbs, Guo (1993) and Xiao and McEnery (2004) tag their stative meaning and activity meaning as different situation types. We agree with the latter one and assign different situation types to their different senses.

Xiao and McEnery (2004) also support to process situation types in different levels which is similar to He (1992), Zhang (1993) and Dai (1997). Xiao and

McEnery (2004) serve as the pioneers to totally apply corpus data to validate aspectual hypotheses although mainly by the methodology of counting examples. However, because of the small size of corpus and the absence of statistical methodology, the validation is still untenable.

Six situation types are constructed in Xiao and McEnery (2004) including individual level states, stage-level states, acitivities, semelfectives and achievements. Xiao and McEnery (2004) are different from other studies by breaking states into two different situation types. While the former one is not able to co-occur with progressive or durative aspectual markers, the latter one is. Other situation types are similar with Smith (1991).

Xu (2015) denies the exsitence of situation types in lexical level and insists that verb constellation is the basic unit that can be assigned situation type. Following Smith (1991), he also idendifies five situation types in sentential level inclduing state, activities, semelfectives and achievements.

In addition, activities occur with progressives are treated as "dynamic state":

(19) He is running.

He creates linguistics event which is the combination of situation type and viewpoint aspect and identifies 18 linguistics situation types.

His situation type classification, although innovative, is questionable because viewpoint aspect and situation type are actually mixed up which makes the aspectual system complex and unclear.

The biggest weekness for linguistics studies on Mandarin aspect is the methodological one. Most of linguistics studies seek to construct the situation type system and depict the interaction rule between situation type and aspectual markers via intuition or introspection, which is not a sound way because of two reasons.

On one hand, the self-created examples are often debatable. As is mentioned in example (6), the grammaticality of these examples can not be shared even among native speakers and thus Tai (1984)'s hypothesis that the attaiment of goal can be cancled for Mandarin accomplishments is not tenable. It is quite common for aspectual studies to quarrel with each other on the accpetability of their supporting or violating examples. The lack of observability of the data makes it almost impossible to judge their reliability by simply asking several other native speakers.

On the other hand, the tendencies they reason out are mainly based on several or a dozen of examples. The sample size is so small that almost no statistical tests can be implemented on it and without any statistical validation, the tendency shown by these seveal examples maybe just coincidence. Many of the rules are easy to be broken when we expand our invesitigation in corpus with larger size. In previous studies, ZAI and ZHENGZAI are always put together to represent the same aspectual marker. However, statistical test in this thesis shows their

significant difference when combined with different situation types of verb senses.

Linguistics studies including Tai (1984),Teng (1985), He (1992), Zhang (1993) and Guo (1993), although developed promising model, all suffered from the problems triggered by their intuition methodology.

Realizing the unreliability of intuition or introspection, Xiao and McEnery (2004) begin to seek for help from corpus data to validate their hypotheses. By searching for examples in a corpus, the assumptions or tendencies reasoned out from their intuitions are able to be validated by uttered sentences appearing in the corpus. However, simply counting examples is still not reliable since the difference between frequencies are easily to be exagerated without any validation through statistical test.

Xu (2015) performed his model by using computational methodology which is quite reliable. However, his study, together with other aspectual studies in computational method such as W. Li et al. (2004), Cao, Li, Yuan, and Wong (2006) and Zarcone and Lenci (2008), are not predicting the lexical aspect. Although claiming to do classification of verbs, they are in fact predicting the situation type of sentences. However, it is the lexical aspect that needs further study since situation type in sentential level has gained too much atttion. What's more, the interaction between situation type and viewpoint aspect is supposed to be initiated from lexial level. The prediction of situation types of sentences are

77

actually processing the interaction of phrases and aspectual markers. But no rules on the interaction between verbs and aspectual markers can be generalized.

The lack of statistical validation in linguistics studies and the ommition of lexical aspect in computational studies spare space for this thesis to further explore the interaction between verbs' situation type and viewpoint aspects in statistical way and computational method.

In this chapter, we mainly reviewed and criticize various classification of situation types in different studies on Mandarin Chinese aspect. Based on these studies, our situation types system is contructed and further validated in next chapter.

Chapter 3

The classification of lexical aspect

This chapter models the situation type system in lexical level which we called lexical aspect. As is stated in previous chapters, a prototype nature of situation type system is accepted in this thesis and the members of each situation type form a claque from typical ones to atypical ones, which can be explained by prototype theory. We will firstly introduce prototype theory and then demonstrate the prototypical nature of situation type system and their members. The situation type system is further constructed through hierarchical clustering. Finally, the annotation of situation type for selected verbs is introduced.

The construction of situation type system is based on three hypotheses.

The first hypothesis is that viewpoint aspects including ZHE, LE1, LE2, GUO, ZAI and ZHENGZAI are able to be used to differentiate between different situation types, which has become a common consensus in previous aspectual studies. We have no intention to debate on how many situation types are there in Mandarin Chinese by analyzing concrete examples and intuitions. The situation type hierarchy is automatically generated though hierarchical clustering on the manually annotated dataset featured with verb and their ability to co-occur with aspectual markers.

It should be noticed that by accepting these aspectual markers as the classifying features, we are not refusing that other elements, such as adjuncts or

temporal adverbials, are also able to identify different situation types. There are dozens of features which can be used to classify different situation types but it doesn't mean that situation types' differentiation must include all of these features. Based on previous studies, those aspectual markers are able to differentiate between different situation types and we mainly focus on the interaction between these aspectual markers and situation types in this thesis. Other features such as verb reduplication and syllable length are also concerned. But they are not used to construct our situation type system but are used in the evaluation of our situation type system. The improvement of classification performance actually validate their capacity to differentiate between different situation types.

The second hypothesis we made is that situation type system is a prototype concept system. Verbs are grouped together based on their family resemblance. Consequently, the situation type system can be constructed according to their family resemblance with each other. The family resemblance is simulated by the distance between each verbs in the vector space constructed by their ability to co-occur with aspectual markers.

The third hypothesis for situation type in lexical level is that situation type information is attached to verb senses instead of verb *per se*.

The ability to co-occur with an aspectual marker for a verbal sense is judged by our own intuition and then cross-validated in Meng (1987), Guo (1993) and Yu, Zhu, Wang, and Zhang (1996) to guarantee the reliability.

The situation type system can be constructed based on the intuition judgement in this way and thus all of the verbs are assigned a situation type.

The validation of the intuition-based situation type system is implemented on Sinica corpus through logistic regression in next chapter.

3.1 The prototype model of situation type

This thesis supports a prototype explanation of verbs' situation type system and their members which is explicitly shared by L. M.-j. Huang (1987), Smith (1991), Zhang (1993) and Guo (1993). The corpus data and vector matrix based on it also reflect the nature of prototype of verbs' situation type in this thesis.

3.1.1 Prototype theory

The prototype model of categories insists that "human categories and concepts have prototypical rather than hierarchical structure" (Smith, 1991). With family resemblance, the prototype model is much clearer.

According to prototype theory, members in a category are organized around the most typical one and other members become less and less typical, which is illustrated by their distance from the center of the cluster which are the most

typical ones. The relationship between members in a category is measured by family resemblance (E. H. Rosch, 1973; Wittgenstein, 2010).

The frequently cited example to illustrate prototype and family resemblance is the category of *bird* and its members. While sparrow is a typical bird with almost all of the characteristics of a bird, ostrich is atypical. When subjects are asked to describe the category of a bird, sparrow, instead of ostrich, is usually picked out as the reference point and the same is for circle and ellipse (E. Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; E. H. Rosch, 1973).

According to Eysenck and Keane (2000), there are several conclusions for prototype theory.

Firstly, category or concept has a prototype structure which is illustrated by either a set of features or the most typical example of the category or concept. While robin or sparrow can be seen as the best example of bird, features of wings, beak etc. can also be seen the prototype of bird.

Secondly, there is no deciding set of "necessary and sufficient attributes" for one category and some attributes may be necessary but "they are not jointly sufficient" (Eysenck & Keane, 2000). Wings may be one of the necessary features of bird but they are not sufficient ones because some animals with wings are not bird such as bat.

Thirdly, the borderline is vague or fuzzy between categories or concepts. Eysenck and Keane (2000) pick out tomato as an example illustrating the vague boundary between fruit and vegetable.

Fourthly, members of a category or concept are positioned according to their typicality. There is a continuum of typicality illustrating the members' similarity or distance with the prototype.

Finally, whether one example belongs to a certain category or not is decided by "the similarity of an object's attributes to the category's prototype" (Eysenck & Keane, 2000).

Situation types are considered to be with a prototype structure because the conceptual features including [Durative], [Dynamic], [Telic] etc. are prototypical features which will be introduced in the next section.

3.1.2 Prototypical conceptual features

In this section, we borrowed conceptual features of situation types to illustrate the prototypical nature of situation types.

Conceptual features serve as shortcuts for the internal temporal structure of situations (Smith, 1991). The nature of prototype and the contrasting between stative-dynamic, durative-punctual and telic-atelic are considered to be rooted in our cognitive system. While E. H. Rosch (1973) illustrates the prototype nature of natural categories with a lot of cognitive experiments, Smith (1991) and Zhang (1993) provide the cognitive foundation of the conceptual features of situation

types which are [Stative], [Telic] and [Durative]. This thesis will not focus on the validation of prototype category and the nature of situation types' conceptual features through cognitive experiments. They are accepted directly from the previous studies as the foundation to explain the interaction between situation types and viewpoint aspects.

It should be noticed that we have no intention to construct our situation type system according to these conceptual features including [Durative], [Dynamic], [Telic] etc. since they are not observable and hard to be annotated. The constructing of situation type system is based on the verbs' ability to co-occur with different aspectual markers which are much easier to be annotated and there are also existing language resources for us to cross-check the annotation of the verbs' ability to co-occur with aspectual markers.

Vendler (1967)'s classification features of [Durative], [Dynamic] and [Telic] has been widely used as the conceptual features of situation type. Although with debates on the definition of [Telic], they are consistently accepted as features for identifying different situation types. However, all of L. M.-j. Huang (1987), P. Chen (1988), Smith (1991), He (1992), Guo (1993) and Dai (1997) have noticed their prototypical instead of hierarchical nature.

3.1.2.1 Durative

The feature of [Durative] stands for whether one situation is able to last for a certain period of time or not. Vendler (1967) treats [Durative] as a binary feature which is not the case no matter between different categories or within one category.

Both Tai (1984) and Teng (1985) treat accomplishments and achievements as non-durative or instantaneous situations, which can be proved by their lack of ability to co-occur with progressive aspectual marker ZHE while durative situation types including states and activities are able to. However, this is not the case. While most of transitions cannot be found in the corpus to co-occur with ZHE, some are able to.

L. M.-j. Huang (1987) has already noticed the durativeness of verbs, which she uses periodicity to refer to. The continuum of periodicity has been illustrated in chapter 2 which is copied as follows:

Short	si3	tiao4	pao3	chuan1	zhi1dao4	**Long**
Periodicity	die	jump	run	wear	know	**Periodicity**
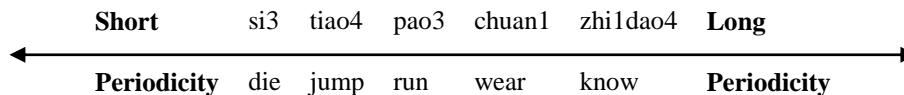
Figure 8 Periodicity continuum in L. M.-j. Huang (1987)

Zhang (1993) followed L. M.-j. Huang (1987), but gives continuum with more grains:

Perfective	A	B	C	D$_i$	D$_{ii}$	E	F		G	Imperfective
	dao4	shang4	dou3	ben1	dui1	dai4	fa1chou2	jue2de		
	arrive	go up	shake	run	stack	carry	worry		feel	
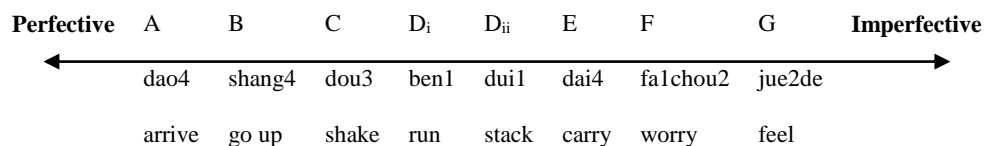
Figure 9 Zhang (1993)'s situation type continuum

While typical perfectives are instantaneous without durative period, typical imperfectives are durative enough to be a state or attribute, which are without any interruptions during their existing period.

Durative aspectual marker ZHE and adverbials such as post-verbal temporal phrases are usually used to test the durativeness of one certain verb. ZAI and ZHENGZAI are also used to test the durative of a situation. While ZAI is confirmed as a measurement of [Durative], ZHENGZAI is totally different and it is related with [Dynamic] according to our statistical analysis in the next chapter.

Not all durative situations are able to co-occur with progressive aspectual marker and adverbial and also not all punctual situations are not able to co-occur with them according to P. Chen (1988), He (1992), Guo (1993) and Xiao and McEnery (2004).

For states, their ability to co-occur with ZHE varies. While some absolute states such as *you3* (have), position verbs such as *dai4* (carry) and posture verbs such as *tang3* (lie) are able to co-occur with ZHE, most of absolute states verbs

86

and mental verbs are rarely to be found to co-occur with ZHE. Transitions are hard to be found to co-occur with ZHE with several exceptions such as *chan3sheng1* (come into being) which is with very low co-occurrence frequency in the corpus. Activities are supposed to be able to co-occur with ZHE.

3.1.2.2 Dynamic and Telic

While [Dynamic] is widely accepted in previous studies, [Telic] is almost always without consensus. But it should be noticed that both of them are concerned with the comparison between non-change and change.

[Dynamic] is the opposite of [Stative]. Smith (1991) points out that "the distinction between stasis and motion is fundamental……The distinction between states and events is commonly reflected in languages……in English, an event *occurs*, *happens*, *take place* while a state *holds* or *obtains*."

Prototypical non-dynamic situations are absolute states such as *deng3yu2* (be equal to), *xing4* (be surnamed as). There are no "internal phrases" and "no change" for them. Typical statives "don't happen nor can be done" but are obtained or held during their persisting (Xiao & McEnery, 2004). They are non-decomposable:

(20) 我姓刘。

     wo3   xing4   liu2   .

     I   surname   LIU   .

'I am surnamed as LIU.'

Once was surnamed as LIU, the state of that person is held and during its persisting, there is no phases that the state is interrupted.

Mental verbs such as *xi1wang4* (hope) are more dynamic than absolute states on the sense that they represent mental activity which is not that stable.

Activity, as a dynamic situation such as *pao3* (run), is decomposable and contains phases during their happening. While states are with homogenous internal structure, typical activities are with heterogeneous internal structure. For *pao3* (run), if taking pictures for every moment, different positions of arms and legs will be observed. Langacker (1987) points out that dynamic situations is composed of "a series of states representing different phases of the process……" and that's the reason why Zhang (1993) and Xiao and McEnery (2004) also treat heterogeneous internal structure as a kind of change.

However, it should be noticed that transition verbs such as *si3* (die) also imply change which is different from the change illustrating by the heterogeneous internal structure of *pao3* (run). Dowty (1979) creates indefinite change and definite change to refer to these two different kinds of change. But other scholars such as Smith (1991) insists to uses [Telic] to refer to the definite change.

Xiao and McEnery (2004) insist that ZAI serves as the test aspectual marker for [+Dynamic] or [-Stative]. However, our validation shows that there is no significant different between the ZAI and ZHE since both of them show positive

relation with [+Durative] situations. It is ZHENGZAI that serves as the indicator for [Dynamic] according to the statistical results.

[Telic] is originally used to distinguish between "Aristotle's *kinesis* (accomplishments) and *ergergerai* (achievements)" and refers to "toward a goal" (Xiao & McEnery, 2004). To avoid the hint of human agency, Smith (1991) maps [Telic] to natural final point which includes accomplishments into telic situations. While atelic situations such as activities have an arbitrary final point, telic situations such as achievements and accomplishments have a natural end point. This definition actually links [Telic] to [Bounded].

We have no intention to create new terminologies and explanations to further roil the water. We agree with Zhang (1993), Xiao and McEnery (2004) and Dai (1997) that both heterogeneous internal structure and change of state are included into the category of change. Dowty (1979)'s indefinite change and definite change can be borrowed to imply the obviousness of changing. Smith (1991) illustrate this through *widen*:

(21) They are widening the road.

Sentence (21) is treated as an activity but she insists that it is vague on *widen*'s situation type because different states are included in the process of widening which is similar to transitions in which a change of state takes place. However, the absence of a "natural final point" pushes it out from transitions. Although

89

arguing against the belonging of *widen*, the vagueness of *widen* is accepted by us
to illustrate the continuum of change which is shown in Guo (1993):

Va	Vb	Vc1	Vc2	Vc3	Vc4	Vc5	Vd1	Vd2	Ve
shi4	ren4shi2	xiang1xin4	xi1wang4	ai4	gong1zuo4	chi1	ti2gao1	li2kai1	wang4
be	know	believe	hope	love	work	eat	enhance	leave	forget
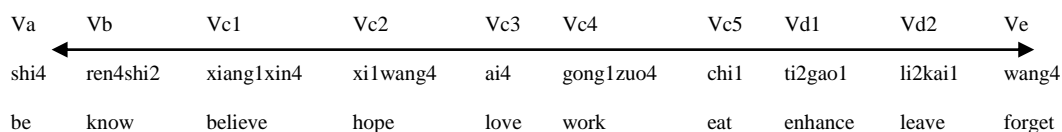
Figure 10 Guo (1993)'s continuum of situation types

LE is accepted as the main indicator to illustrate [Dynamic] and [Telic]. It is
widely accepted in Mandarin studies such as C. N. Li and Thompson (1989) that
there are two LEs in Chinese. While detailed analysis is proceeded in chapter 4,
two LEs perspective is held in this thesis. Following C. N. Li and Thompson
(1989), sentential final LE and post verbal LE are used to refer to them
respectively.

Post verbal LE, which we called LE1, is a perfective aspectual marker which
illustrates occurrence. Although occurrence is related with completedness. Both
Smith (1991) and Xiao and McEnery (2004) stress that LE1 tends to co-occur
with telic situations. This is validated in our latter analysis.

Guo (1993) notices that the meaning of "V+LE2" actually exposes a transition
from inchoative to completive for different situation types of verbs:

(22) a. 病了。

     bing4   le   .

     sick   LE   .

'become sick.'

b.吃了。

chi1　　　　le　　　　.

eat　　　　LE　　　　.

'have eaten. 'or 'begin to eat.'

c.到了。

dao4　le　　.

arrive　LE　.

'have arrived.'

While states tend to choose an inchative meaning with LE2, transitions tend to choose a completive meaning with LE2. Activities are between them and thus are vague in their meanings when co-occurred with LE2.

This is further illustrated by "V+LE1+TP+LE2":

(23) a.病了三天了。

ta1　bing4　le　　san1　tian1　le　　　.

he　sick　　LE1　three　day　　LE2　.

91

'He has been sick for three days.'

b.我吃了三个小时了。

wo3   chi1   le   san1   ge4   xiao3shi2   le   .

I   eat   LE1   three   CLA   hour   LE2   .

'I have been eating for three hours.' or 'it has been three hours since I finished my meal.'

c.她到了三个小时了。

ta1   dao4   le   san1   ge4   xiao3shi2   le   .

She   arrive   LE1   three   CLA   hours   LE2   .

'It has been three hours since she arrived.'

In (23a), the temporal phrase illustrates the durative time from the beginning of *bing4* (sick) which is an state; in (23c), the temporal phrase illustrates lapsed time after the compeltion of *dao4* (arrive) which is an transition; However, in (23b), the temporal phrase's meaning is vague.

However, LE2 is different from LE1 and our statistical validation will shows that it is complementary with GUO which is relatated with [Repeatable].

3.1.2.3 Repeatability

[Repeatability] is related with a situation's ability to be carried out repeatedly or iteratively. Many studies including Smith (1991), Zhang (1993), Dai (1997) and Xiao and McEnery (2004) noticed this feature of situation types in Chinese.

Broadly speaking, activities tend to be easier to be repeated compared with states and transitions because of the internal temporal structure which is with successive interval stages. For example, *pao3* (run) implies a process of lifting and laying down legs etc.; semelfectives such as *ke2sou4* (cough) are typical activities that are usually interpreted in an iterative way.

The continuum of repeatability exists not only between situation types but also within the same situation type. While absolute states such as *deng3yu2* (equal to) are hard to be repeated, non-absolute states such as *zuo4* (serve as)'s ability to be repeated is a little stronger; compared with achievement verbs such as *si3* (die), other transitions such as *kuo4da4* (enlarge)'s ability to be repeated is better.

The frequently used aspectual marker to measure repeatability is GUO. Zhang (1993) insists that only repeatable situations are able to co-occur with GUO. This statement, although too assertive, is basically on the right track and it is further edited by us that GUO tends to co-occur with repeatable situations. Verb reduplication is also able to measure repeatability. Although it is not included into our situation type classification system, the ability to be reduplicated of verbs

is tested through classification experiments and the improved performance confirms its' function in distinguishing situation types.

In this section, the conceptual features of situation types and their tests are illustrated in detail. However, the relationship between conceptual features and the tests materialized by aspectual markers are supported only by intuition or partial corpus examples. The relationship is not validated in statistical way.

For now, the only hypothesis that we can make is that aspectual markers including ZHE, LE1, LE2, GUO, ZAI and ZHENGZAI are related with the conceptual features and thus they are supposed be used as indicators for classifying verbs into different situation types. Accordingly, our later application of situation type clustering are actually based on these aspectual markers instead of these aspectual values including [Durative], [Dynamic], [Telic] and [Repeatability].

The validation of the relationship between aspectual markers and aspectual values is implemented in chapter 4.

3.2 The construction and annotation of situation type system

It has been proposed that situation type system is with a prototype structure in previous sections. According to Eysenck and Keane (2000), the prototype structure which is either a set of features or the most typical example of the category or concept. We insist that the porotype structure of situation type system is a set of features of [Durative], [Dynamic], [Telic] and [Repeatable] which are

94

measured by several aspectual markers. None of these features is decisive on identifying a verb's situation type. All of the categories' members are clustered according to their family resemblance with each other. In this section, a bottom-up process is implemented to construct the situation type system and the situation type system is applied to the extracted verbs from the corpus.

The foundation for situation type system is the feature of the members, which is illustrated by their ability to co-occur with the several aspectual markers proposed in the previous section. It should be noticed that the judgement of the co-occurrence ability is mainly based on Meng (1987), Guo (1993) and our intuition.

3.2.1 Annotation of verbs' ability to co-occur with aspectual markers

The verbs should be selected before the annotation of these verbs' corresponding features and Meng (1987) is chosen as the source for selecting target words because of the following reasons.

Firstly, it is a verb dictionary. Meng (1987) selects more than 1200 verbs from Lv and Ding (1978). Since this thesis is about verbs' lexical aspect, a dictionary focus on Chinese verbs is the best choice. Lv and Ding (1978), as an authoritative dictionary, is strict in selecting words, partitioning senses and explaining meanings and thus the quality and representativeness of the selected verbs can be guaranteed.

Secondly, verbs are organized by senses. One of the most reasons that Zhang (1993) is argued against in this thesis is that she chooses the so-called "typical" meaning of a verb as the basic unit to assign situation type. Regardless of the debatable definition of typical meaning of a verb, different verbs may share the same form:

(24) a.阿扁做总统。

A1 Bian3   zuo4   zong3tong3   .

A Bian   act as   president   .

'Abian acts as the president.'

b.李逸洋正在做最后评估。

Li3 Yi4yang2   zheng4zai4   zuo4   zui4hou4   ping2gu1   .

Li Yiyang   ZHENGZAI   make   last   evaluation   .

'Li Yiyang is making the last evaluation.'

While *zuo4* (act as) in (24a) is a state, *zuo4* (make) is an activity.

What's more, posture and positon verbs are able to show different internal structures:

(25) a.他穿着一件非洲的大褂。

ta1   chuan1   zhe   yi1   jian4   fei1zhou1   de   da4   gua4   .

he   wear   ZHE   one   CLA   African   DE   big   gown   .

'He wears a big African gown.'

b.小敏让我下去穿件外套。

xiao3min3   rang4   wo3   xia4   qu4   chuan1   jian4   wai4tao4   .

Xiaomin   let   I   down   go   wear   CLA   coat   .

'Xiaomin told me to go downstairs to wear a coat.'

Thus, following C.-R. Huang et al. (2000)'s basic hypothesis in their Module-Attribute Representation of Verbal semantics (MARVS), we also maintain that the situation type (ST) or eventive information is attached to the sense of a verb instead of the verb *per se*.
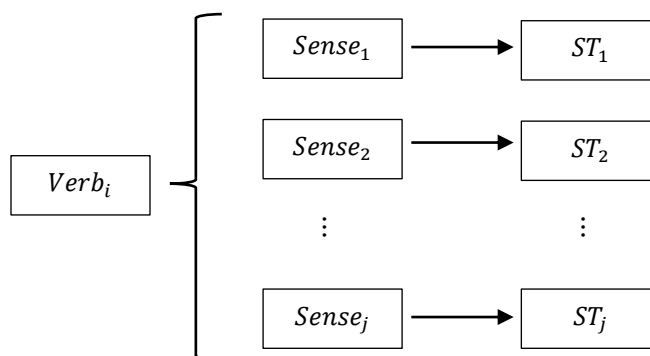
Figure 11 Verbs senses and their situation type

97

Meng (1987) and Lv and Ding (1978) make further differentiation between homographs and different senses of the same verb forms and thus can be used as a base to make a more precise annotation of verbs' interaction with aspectual markers. Their generalization of verbs' senses, although is not thorough, is based on both verb's syntactic and semantic information which serves as an excellent bridge to be linked to the corpus words which are not tagged based on senses but on syntactic functions and some semantic features.

Thirdly, it contains co-occurrence information of verbs with aspectual markers. Meng (1987) lists 13 syntactic functions including their ability to co-occur with ZHE, LE and GUO which serve as one of the benchmarks for checking our own judgement.

Based on the analysis in the last section, the following aspectual markers are annotated to be the set of prototypical features:

(i) ZHE;

(ii) ZAI and ZHENGZAI;

(iii) LE1 and LE2;

(iv) GUO

These aspectual markers are chosen not only because they are able to measure [Durative], [Dynamic], [Telic] and [Repeatable] but also because they are also tagged in Meng (1987),Guo (1993) and Yu et al. (1996) which can help us to make cross-checking.

Verbs with the same distribution are merged into small groups and typical words are picked out as the representatives of these groups. These types are type *shi* (be), type *ren4shi* (know), type *xi3huan1* (like), type *xin4ren4* (trust), type *ai4* (love), type *gong1zuo4* (work), type *chi1* (eat), type *chan3sheng1* (generate), type *li2kai1* (leave) and type *si3* (die) respectively.

Both Guo (1993) and Yu et al. (1996) also annotate some of the co-occurrence information of verbs with aspectual markers. A triple-checking is implemented on these dictionaries to gurantee the quality of the annotation. Since our goal is to acquire verbs with annotated features as more as possible, consistency test is not appropriate for our study since all of the inconsistent annotations are further checked and modified instead of keeping them there without any changes.

3.2.2 Hierarchical clustering of verbs

In prototype theory, typicality is mainly measured by family resemblance which is created by Wittgenstein (2010). It is quite common to calculate the family resemblance score by "noting all the attributes that that member has in common with all the other members of the category" (Eysenck & Keane, 2000) just as what E. H. Rosch (1973), E. Rosch et al. (1976) and E. Rosch (1999) did in their studies.

Guo (1993) also applies the counting method to measure the aspectual resemblance between different verbs. However, this is not a good way to simulate

family resemblance or family similarity which can be shown by the following figure:
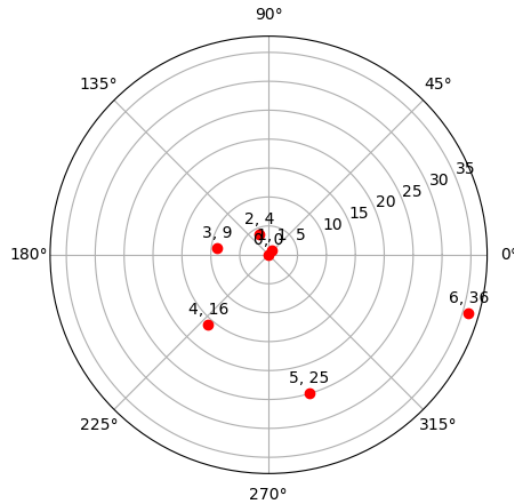
Figure 12 Family resemblance sample

If the circle in the outermost layer is identified as the boundary of a category and the dots are distributed as the members within this category, their family resemblances are better simulated by the Euclidean distance between them while features' values of them are treated as their coordinates in vector space.

Measuring the similarity between different verbs in vector space has been widely accepted in computational linguistics and also other scientific subjects. In this section, the aspectual resemblance is also measured based on members' distance within the vector space. Literature review on vector space models is unfolded in chapter 5.

Before the distances' calculation between different members, it's better to visualize them in the coordinate system. While it is possible to calculate the distance between vectors with six dimensions, it is not possible for human to portray a subject in a space with more than three dimensions. Dimensionality reduction has to be implemented before the visualization of a high-dimensioned vector matrix.

In this section, PCA (Principle Components Analysis) is implemented to reduce the six dimensions represented by the verbs' ability to co-occur with the six aspectual markers. PCA recombines all of the features into one new feature by giving different weights to the original features and thus it is inevitable to lose some information of the original matrix. Broadly speaking, the more the components are, the less the loss is. Thus explained variance ratio is imported to measure the explanation of the new matrix for the original matrix. Before the dimensionality reduction, a curve of the explained variance ratio is calculated and depicted for the matrix in the following figure:
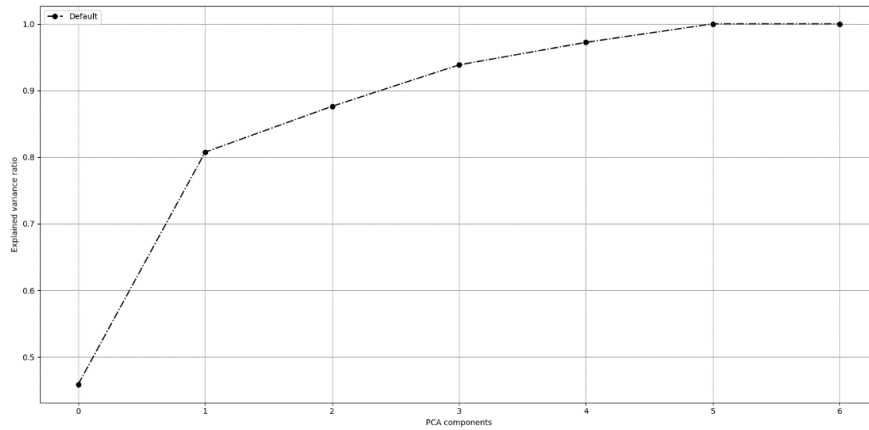
Figure 13 The explained variance ratio of PCA components

It is quite clear that the explained variance has reached 94% when three components are created while two components explained about 87% of the original matrix.

Calculating the family resemblance by counting the matched features of a verb is actually mapping the multi-dimensional matrix to a one dimensional matrix. It is very similar to only keep one PCA component but it remains far less information than one PCA component does because PCA gives different weights to different features to keep as much information as it can. One PCA component only keeps 80% information of the original matrix. It then follows that simply counting features only keeps less than 80% information of the original matrix, which makes it unreliable.

To make the visualization more reliable, three components are set and thus a three-dimension vector space is created with the representative verbs' position on it:
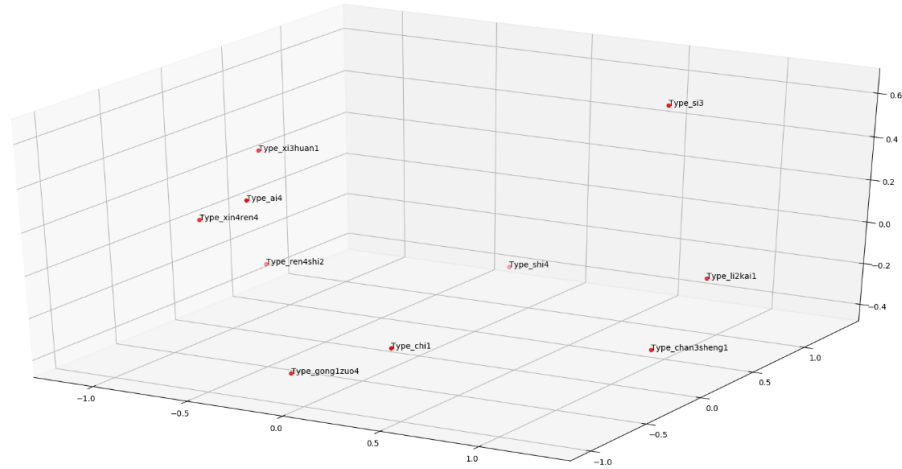
Figure 14 Positions of different types of verbs in aspectual vector space

Several types of verbs are very close in the three dimension figure. Type *xi3huan1* (like), type *ai4* (love), type *xin4ren4* (trust) and type *ren4shi2* (know) can be definitely clustered together; type *chi1* (eat) and type *gong1zuo4* (work) are also supposed to be grouped into one kind; type *chan3shengq* (generate), type *li2kai1* (leave) and *si3* (die) are close enough to be distinguished from other type of verbs. It should be noticed that *si3* (die) seems to be isolated because of the angle of the projection but actually it is close to type *chan3shengq* (generate) and type *li2kai1* (leave) which is proven in the hierarchical clustering dendrogram.

Type *shi4* (be) is isolated from other clusters and thus is vague in its belongings from the visualized figure.

Clustering as a methodology to partition subjects is different from classification. While classification predicts a members' category based on the trained model or tagged data, clustering grouped members together only based on vector distance. This also makes clustering unreliable since clusters can also be found even in random data (Hou & Jiang, 2014).

Different coefficients are calculated to measure the reliability of the clustering. While silhouette coefficient is used to validate the reliability of clusters generated through partitioning clustering such as *k*-means algorithm, cophenetic correlation coefficient (CPCC) is used to validate the hierarchical clustering.

Hierarchical clustering generates a tree by continuously merging members from bottom to up which is like a dendrogram in biology. It is not sensitive to sample size and "are more efficient in handling noise and outliers than partition clustering" (Hou, Yang, & Jiang, 2014). Thus hierarchical clustering is targeted as the method to construct the situation type system from bottom to up.

CPCC is "a measure of how faithfully a dendrogram preserves the pairwise distances between the original unmolded data points " (Sokal & Rohlf, 1962). If we use $\{X_i\}$ to represent the original matrix and use $\{Y_j\}$ to represent the dendrogram clustered by hierarchical clustering, CPCC between the original data and clustered data can be depicted as:

$$CPCC = \frac{\sum_{i<j}(x(i,j) - \bar{x})(y(i,j) - \bar{y})}{\sqrt{\left[\sum_{i<j}(x(i,j) - \bar{x})^2\right]\left[\sum_{i<j}(y(i,j) - \bar{y})^2\right]}}$$

Formula 1 CPCC

$x(i,j)$ and $y(i,j)$ are defined as the Euclidean distance between the i-th member and the j-th member while $\bar{x}$ and $\bar{y}$ are the mean of them respectively.

Different linkage method can be applied to calculate the distance between two clusters and thus various dendrograms can be constructed. All of them are with a value of CPCC and the one with the highest score of CPCC is chosen as the best dendrogram.

The process of targeting linkage method will not be unfolded here. With a CPCC of 0.77, the method of *centroid* gains the highest CPCC. If c_s and c_t are the centroids of cluster s and clustert, the distance between these two clusters can be depicted as followsing:

$$d(s,t) = \|c_s - c_t\|_2$$

Formula 2 Centroid linkage method of hierarchical clustering

The result of hierarchical clustering for different types of verbs is generated as follows:
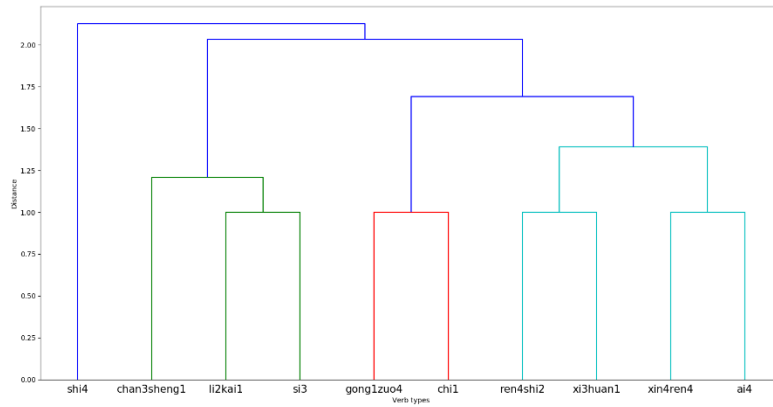
Figure 15 Hierarchical clustering of verbs

Four clusters are identified which is in consistent with the three dimensional figure from right to left:

(i) Type *ai4* (love), *xin4ren4* (trust), *xi3huan1* (like) and *ren4shi2* (know) group together into the first situation type. Among them, type *ai4* (love) and *xin4ren4* (trust) are close enough to be merged into one small group while *xi3huan1* (like) and *ren4shi2* (know) are close enough to form another small group. Most of them are mental verbs which express a kind of mental state or mental activity.

(ii) Type *chi1* (eat) and type *gong1zuo4* (work) are clustered into the second situation type. These are typical physical actions.

(iii) Type *si3* (die), *li2kai1* (leave) and *chan3sheng1* (generate) are merged into the third situation type. These types of verbs imply a change of state.

(iv) Type *shi4* (be) is isolated to be an individual group because the distance between it and other types are almost the same. There are only a dozen of verbs in this type such as *shi4* (be) and *deng3yu2* (eaqul to) which illustrate a kind of attribute. Small sample size is not appropriate to implement statistical validation and thus it is merged into the first situation type which is state.

It then follows that situation type system is tri-sectioned into state, activity and transition in this thesis:

Table 15 Distribution of verbs in Meng (1987)

	State	Activity	Transition	Total
Number	387	978	245	1610

 In this section, based on the annotated features which are materialized by the ability to co-occur with several aspectual markers, hierarchical clustering is implemented on the selected verbs in Meng (1987). The situation type system is thus constructed.

All of the annotation of the verb's ability to co-occur with aspectual marks are cross-checked between our own annotation, Meng (1987)'s annotation, Guo (1993) and Yu et al. (1996)'s annotation to ensure the reliability of intuitions.

The situation type system should be validated statistically on a corpus which is independent from the intuition-based data for constructing the verbs' situation type system.

Before the validation, annotation of situation types of the verbs in the corpus is supposed to be applied firstly.

3.3 The annotation of verbs' situation type

Although verbs in Meng (1987) are organized by senses, however, none of corpuses is tagged by verbs senses. Since situation type of verbs are based on senses in Meng (1987), a mapping between the words tags in the corpus and the situation type of words' senses is supposed to be confirmed. Before this, the corpus that are about to be used in this thesis is firstly introduced.

3.3.1 Corpus

Two corpuses are used in this thesis including the Sinica corpus and Chinese Gigaword corpus. While the former one is used to extract selected features' co-occurrence frequency, the latter one is used to train the neural network models to extract word embedding vectors.

The Taiwan Academia Sinica Corpus (Sinica corpus) is the first balanced Chinese tagged corpus (K.-J. Chen et al., 1996). Version 4.0 is accepted in this thesis to construct the matrix filled with co-occurrence frequency between aspectual markers and verbs (senses). Sinica corpus is targeted because of the following reasons.

Firstly, it is with high quality. The POS tagging for Sinica corpus is firstly completed automatically. Post manual checking and editing is applied on the automatic tagged corpus and thus the quality is guaranteed.

Secondly, the tagging information is abundant. On one hand, the segmentation principle of Sinica corpus is consistent with Meng (1987) because they all insist that both semantic and syntactic function should be considered and thus words are not only tagged with the tags of verb, adjective and noun etc. but also with more detailed semantic and syntactic information. For example, VG is the tag of classification verb such as *zuo4wei2* (serve as). These information are very useful in identifying different senses of a verb. On the other hand, most of the aspectual markers are clearly tagged in the corpus with the tag of Di which make it possible to extract the precise co-occurrence information of verbs with these aspectual markers. LE1 and LE2 are also tagged respectively in the corpus. The POS tags are listed in appendix.

Thirdly, the corpus is balanced. Sinica corpus is composed of different genres, styles, modes, topics and mediums. Balanced corpus make the data much more various and thus the tendencies based on balanced corpus are less possible to be influenced by the factors other than the language *per se* and thus is much reliable.

Finally, the corpus size is big. Version 4.0 of Sinica corpus is with more than 10 million words. Among manually tagged corpus, it is one of the biggest ones.

Since machine learning methodologies are also accepted in the last chapter to evaluate the situation type dataset, Sinica corpus is not big enough and thus Chinese Gigaword Corpus (CGW) is also used in this corpus to predict the neural word embedding vectors through *word2vec*.

CGW is one of the biggest Chinese corpuses with tagged information. The size is about 1.12 billion words distributed from Mandarin in mainland China, Taiwan and Singapore which guarantee the balancing of language variation and make the validation reliable for the Chinese in China mainland, Taiwan and Singapore.

The construction of matrix based on these two corpuses is introduced in chapter 4 and 5 respectively.

3.3.2 Annotating situation type of the verbs in the corpus

Since all of the words in Sinica corpus are tagged with POS and the segmentation and tagging of words follow both semantic and syntactic principles, the lexical situation type is linked with words' tags. The annotation of situation types for verbs in Sinica corpus is implemented as follows.

The first step is to construct the dictionary of Sinica corpus and match verbs forms in our annotated dataset. The dictionary of Sinica corpus is organized around word forms and their POS tags. All of the words' forms that exist in the dataset of Table 15 are persisted as the candidates for situation type annotating.

The second step is to map verbs in Sinica corpus with verbs' senses in Table 15. A word-POS pair is treated as a lexical item. For example, 長 has two POS

tags including VC and VH. While the former one is mapped to *zhang3* (grow), the latter one is mapped to *chang2* (long).

The third step is to match each lexical item's meaning with the corresponding sense of that verb in Table 15 which is based on Meng (1987). The meaning of the lexical item is identified in the extracted bigrams and sentences. For example, 除 (*chu2*) has two tags including VC and VJ. By checking their bigrams, 除_VC is identified as "get rid of" while 除_VJ is identified as "divided by".

The final step is to map all of the verbs' senses' situation types in Table 15 to lexical items extracted from Sinica corpus.

Table 16 Distribution of Sinica corpus's verbs' situation types

	State	Activity	Transition	Total
Number	386	634	200	1220

It should be noticed that not all of verbs in Meng (1987) can be found in Sinica corpus' dictionary and some senses of verbs also cannot be matched in the corpus while some other senses are merged into one sense. Thus, the total number of verbs in Sinica corpus is less than that in Meng (1987).

Chapter 4

Statistical validation of situation type system

This chapter focuses on validating the situation type system in statistical methodologies. The interaction between the verbs and aspectual markers is validated through the statistical tests on the co-occurrence frequencies extracted from Sinica corpus.

What's more, we are also trying to quantify the contribution of each aspectual markers in differentiating different situation types when they are all combined together to form a classifying model.

4.1 Validation of the whole classifying features

In this section, multinominal logistic regression is applied to validate the whole classification system materialized by ZHE, LE1, LE2, GUO, ZAI and ZHENGZAI.

4.1.1 Introduction of Multinominal logistic regression

Although various classifiers such as supporting vector machine and artificial neural network are widely used in processing linguistic data and achieve excellent performance in a lot of tasks, logistic regression still occupies a very important position since it is much easier and more straightforward to explain the variables.

Linear dependent variable can be predicted through linear regression but categorical or nominal variable is not suitable for linear regression:
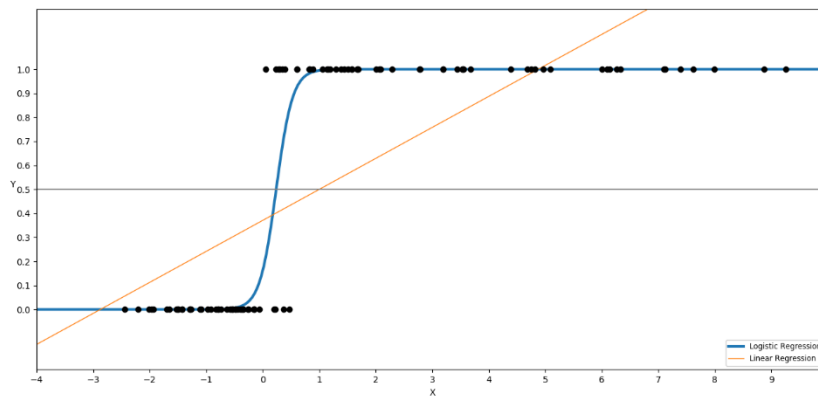
Figure 16 Logistic regression

For data with dummy dependent variables, their scatter diagram is like above which is featured with dots distributed along two parallel lines. Linear regression model is hard to match them while logistic regression curve is able to. As is shown in the figure, the logistic regression curve with an "S" shape covers most of the dots.

A lot of non-linear function is able to display a curve with an "S" shape but logistic transformation of the probability is much easier compared with them and thus is widely used (Pampel, 2000).

Two steps are included in the logit transformation of the probability.

113

The first step is to take the odds of the occurred event. If P_i is used to represent the probability that an individual has a certain feature, the odds of P_i can be defined as:

$$O_i = {P_i}/{(1 - P_i)}$$

Formula 3 Odds

The above formula illustrates the odds of P_i, which stands for the percentage between the occurrence and nonoccurrence of one event. In this way, the upper limit of 1.0 of the probability has been erased and the upper curve of the "S" shape has been straightened.

The second step is to take the logarithm of the odds.

$$L_i = In\left[{P_i}/{(1 - P_i)}\right]$$

Formula 4 Logits of odds

In this way, the lower limit of 0 of the probability is taken away and thus lower curve of the "S" shape is straightened.

Finally, the logit of the probability can be expressed by a linear equation:

$$L_i = In\left[{P_i}/{(1 - P_i)}\right] = \beta_0 + \beta_1 X_i$$

Formula 5 Linearization of logits

The whole transformation process is for binary classification. Multiple classification is very similar. While binary classification uses non-occurred type as the reference, multinominal logistic regression usually picks one of the categories as the reference. In this paper, state is chosen as the reference type.

The estimation of logistic regression is different from linear regression. Linear regression converges the model through OLS (least square method) while logistic regression uses MLE (maximum likelihood estimation). The whole model is validated through likelihood ratio test (LLR)

4.1.2 Multinominal logistic regression analysis

Before logistic regression, multicollinearity test should be implemented. For linear regression, multicollinearity may lead to model distortion and multicollinearity test is implemented here to avoid the similar problem. Since multicollinearity is implemented through correlation analysis, whether the data follows normal distribution or not should be confirmed. Pearson correlation is chosen if the data follows normal distribution. Otherwise, Spearman correlation is supposed to be used.

Shapiro-Wilk test is specially designed for testing the null hypothesis that a sample is extracted from a normally distributed population (Shaphiro & Wilk, 1965). If $x_{(i)}$ are the ordered values in the data and a_i are the constants, the W statistic can be calculated as follows:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x^i - \overline{x})^2}$$

Formula 6 W statistic of Shapiro-Wilk test

Table 17 Shapiro-Wilk test on aspectual markers' co-occurrence frequencies

	W	$p-\text{value}$
ZHE	0.1118	0.0000
LE1	0.2165	0.0000
LE2	0.1097	0.0000
GUO	0.1049	0.0000
ZAI	0.1014	0.0000
ZHENGZAI	0.08760	0.0000

It could be concluded from the above table that none of the co-occurrence frequencies of verbs and aspectual markers follow normal distribution and thus Spearman correlation analysis is implemented to test their multicollinearity.

Table 18 Spearman correlation coefficients between aspectual markers

	ZHE	LE1	LE2	GUO	ZAI	ZHENGZAI
ZHE	1	-0.03636	0.120529	0.127262	0.133219	0.045576
LE1	-0.03636	1	0.027713	0.019172	-0.04071	-0.02603
LE2	0.120529	0.027713	1	0.094726	0.084482	0.010214
GUO	0.127262	0.019172	0.094726	1	0.143605	0.162479
ZAI	0.133219	-0.04071	0.084482	0.143605	1	0.177979
ZHENGZAI	0.045576	-0.02603	0.010214	0.162479	0.177979	1

Table 19 p-values of the coefficients

	ZHE	LE1	LE2	GUO	ZAI	ZHENGZAI
ZHE	0.000000	0.204379	0.000024	0.000008	0.000003	0.111585
LE1	0.204379	0.000000	0.333467	0.503474	0.155270	0.363621
LE2	0.000024	0.333467	0.000000	0.000924	0.003146	0.721551
GUO	0.000008	0.503474	0.000924	0.000000	0.000000	0.000000
ZAI	0.000003	0.155270	0.003146	0.000000	0.000000	0.000000
ZHENGZAI	0.111585	0.363621	0.721551	0.000000	0.000000	0.000000

Usually speaking, only high correlation ($\rho > 0.7$) denotes collinearity. None of the coefficients is bigger than 0.7. Multicollinearity test is passed and thus all the variables can be used for logistic regression.

To make the co-occurrence frequencies of aspectual markers comparable, the data should be normalized firstly. Based on McKinney (2012), the following normalization is applied:

$$x'_i = \frac{x_i - \overline{X}}{s}$$

Figure 17 Normalization

The classification system of situation type we modeled in chapter 3 passes the LLR test according to the following result:

Table 20 LLR test for the classification model

Log likelihood	Log likelihood null	LLR p-value
-1118.7	-1220.8	0.0000

The LLR p-value, which is almost zero, validates the model's difference compared with the unmolded data. Multinominal logistic regression analysis compares between different nominals and thus only $k - 1$ comparison pairs are given with one of the k types as a reference. In our model, state is set as the reference type.

Table 21 Activity vs. state

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
Const.	0.3917	0.111	3.533	**0.000**	1.479	0.174	0.609
ZHE	-0.1125	0.081	-1.383	0.167	0.894	-0.272	0.047
LE1	0.2989	0.172	1.738	0.082	1.348	-0.038	0.636
LE2	-2.9244	0.617	-4.740	**0.000**	0.054	-4.134	-1.715
GUO	0.7654	0.331	2.313	**0.021**	2.150	0.117	1.414
ZAI	0.0215	0.089	0.241	0.810	1.022	-0.153	0.196
ZHENGZAI	1.1802	0.460	2.566	**0.010**	3.255	0.279	2.082

It should be noticed that the values in the table are used to describe activity while state is used as the reference. Taking ZHE as an example, if a verb were to increase its co-occurrence frequency with ZHE by one point, the multinomial log-odds for activity with respect to state would be expected to decrease by 0.1125 unit while holding all other variables in the model constant. Log odds is hardly to be explained and thus odds ratio is attached to the table.

There is a direct relationship between coefficient and odds ratio:

$$OR = e^c$$

Formula 7 The relationship between odds ratio and coefficient

Table 22 Transition vs. state

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
Const.	-3.1972	0.769	-4.157	**0.000**	0.041	-4.705	-1.69
ZHE	-20.564	5.634	-3.65	**0.000**	0.000	-31.607	-9.521
LE1	0.8799	0.193	4.567	**0.000**	2.411	0.502	1.257
LE2	0.0301	0.06	0.499	0.618	1.031	-0.088	0.148
GUO	0.3204	0.359	0.893	0.372	1.378	-0.383	1.024
ZAI	-0.8504	0.432	-1.967	**0.049**	0.427	-1.698	-0.003
ZHENGZAI	1.219	0.474	2.571	**0.010**	3.384	0.29	2.148

Further analysis will be based on these parameters in the following sections in detail.

Chinese is an aspect language with abundant grammaticalized aspectual markers to view the whole or different part of a situation. Perfective aspects treat situation as a whole and they observe a situation from outside while imperfective aspects focus on the inception, duration or final of a situation.

While LE1 and GUO are treated as perfective viewpoint aspect, ZHE and ZAI/ZHENGZAI are dropped into imperfective viewpoint. However, their viewpoint focuses vary even within the same category. They are introduced based on the interaction results with situation types of multinominal logistic regression analysis in the following sections.

4.2 LE1

This section focuses on LE in Mandarin Chinese. We firstly differentiate LE1 and LE2 and then further illustrate their interactions with different situation types.

4.2.1 Two LEs or one LE

C. N. Li and Thompson (1989) illustrate the difference between LE1 and LE2 clearly by their names. While LE1 is called the verb-final suffix, LE2 gains the name of sentential final particle. Sinica corpus is consistent with C. N. Li and Thompson (1989).

However, not all of the aspect studies hold the same view. L. M.-j. Huang (1987) insists to merge LE1 and LE2 into one morpheme because they all express a boundary between two contrasting situations by using the following examples:

(26) a.李四跳了绳了。

Li3si4   tiao4   le   sheng2   le   .

Lisi   jump   LE1   rope   LE2   .

'Lisi has jumped the rope.'

b.李四跳绳了。

Li3si4   tiao4   sheng2   le   .

Lisi   jump   rope   LE2   .

'Lisi begins to jump the rope.'

c.李四高了三寸。

Li3si4   gao1   le   san1   cun4   .

Lisi   high   LE1   three   inche   .

'Lisi grows three inches higher.'

While agreed with L. M.-j. Huang (1987) that both LE1 and LE2 are related with a boundary that illustrates a kind of change, it is argued against that her definition of boundary is so vague that all kinds of changes are included.

Zhang (1993) also holds the one-morpheme view of LE by using the following reasons:

Firstly, she insists that the fact that LE1 and LE2 are able to co-occur in the same sentences proves their semantic similarity. The distributional semantics hypothesis holds that words that have the similar context tend to have similar meaning (Harris, 1954). However, no inferences about the semantic similarity between the co-occurred words can be made from it. Zhang (1993)'s reasoning is untenable.

Even regardless of the untenable reasoning that semantic similarity is related with their co-occurrence, if two LEs are exactly the same, they should not be able to co-occur because of the redundancy of the same morpheme.

Secondly, she insists that the hapology of LE1 and LE2 in the sentential final position, which is proposed by Chao (1968), shows that they are the same because if they are different, the hapology should not occur. However, phonology level is independent from semantic and syntactic level. The difference between LE1 and LE2 is in their semantic content instead of pronunciation.

Thirdly, she argues that both LE1 and LE2 express a change of state by using the following examples:

(27) a.猪皮箱自己跑了。

zhu1pi2　xiang1　zi4ji3　pao3　　　le　　．

pigskin　suitcase　self　run away　LE2　．

'The pigskin suitcase run away by itself.'

b.火车到了。

huo3che1　dao4　le　　　．

train　　　arrive　LE2　．

'The train arrived.'

Zhang (1993) claims that both the post-verbal LE in (27a) and the sentential final LE in (27b) illustrate a change of state while *pao3* (run away) is treated as an activity verb and *dao4* (arrive) is treated as an achievement verb. However, she made at least two mistakes. On one hand, the LEs in (27a) and (27b) are both sentential final particle LE which definitely express a change of state; on the other

hand, both the verbs in (27a) and (27b) are transition verbs which express a change of state *per se*. Thus, examples in (27) cannot support her hypotheses.

Other aspectual studies like Smith (1991), Dai (1997) and Xiao and McEnery (2004) insist the two-LEs view in Mandarin Chinese. Xiao and McEnery (2004) lists the following reasons to show the differences between LE1 and LE2:

(i) From the perspective of syntax, LE1 distributes in the post verbal position while LE2 appears in sentential final position;

(ii) From the perspective of semantics, LE1 expresses the actualization of a situation while LE2 illustrates a change of state;

(iii) From the perspective of etymology, LE1 derived from LE2 which has been proved by Bybee, Perkins, and Pagliuca (1994).

(iv) From the perspective of empirical data, LE1 is more productive than LE2.

While accepting the former three reasons, the last one is questionable according to our data.

Table 23 The distribution of LEs in Xiao and McEnery (2004) and Christensen (1994)

Corpus	LE1	LE2	Double-role	Total
Weekly Xiao and McEnery (2004)'s corpus	1138	175	27	1340
Christensen (1994)'s corpus	75	11	0	86

Xiao and McEnery (2004) actually don't give any definition of productivity and simply assume that the one with higher frequency is more productive. However, token frequency of the word along is not able to reflect its productivity.

123

Himmelmann (2004)'s grammaticalization theory measure the productivity of a grammaticalizing item by the type frequency of its co-occurred words (Peng, 2017). The productivity of LE1 and LE2, which are grammaticalized aspectual markers, can be measured by the type frequency of the co-occurred verbs. However, this point of view neglects the contribution of token frequency in measuring the productivity of a grammaticalizing or grammaticalized item.

The productivity of a grammaticalizing or grammaticalized item reflects the lexical variety of the co-occurred words with it. Type-token ratio (TTR) is a widely accepted measurement of lexical variety in corpus linguistics and quantitative linguistics (C. R. Huang & Liu, 2017). TTR is thus borrowed here to measure the productivity of LE1 and LE2 in our corpus data.

Table 24 TTR of the co-occurred verbs of LE1 and LE2

	Type frequency	Token frequency	TTR
LE1	504	10409	0.048
LE2	121	697	0.174

From the view of lexical variety, LE2 is much higher than LE1, which indicates that LE2 is more productive than LE1.

Now, we turn our scope to measure the difference between LE1 and LE2 in statistical methodologies instead of descriptive ones.

The basic hypothesis is that if LE1 and LE2 are the same and should be merged, there would be no significant difference between their distributions with

124

verbs and the two samples that represent LE1 and LE2's distribution are supposed to be from the same population. If not, the distributions should be statistically different and they are from different populations.

It has been shown that both the co-occurrence frequencies of LE1 and LE2 with verbs don't follow normal distribution and thus only non-parametric tests can be chosen to test their difference.

Kruskal-Wallis H-test is chosen as the non-parametric test here. The null hypothesis of Kruskal-Wallis H-test is that different groups of the same population shows the same median. It is suitable for two or more independent samples (Kruskal & Wallis, 1952). The H statistic can be calculated as follows:

$$H = (N-1) \frac{\sum_i^g n_i (\bar{r}_{i.} - \bar{r})^2}{\sum_{i=1}^g \sum_{j=1}^{n_i} (r_{ij} - \bar{r})^2}$$

Formula 8 H statistic of Kruskal-Wallis test

n_i stands for the amount of the observations in group i; r_{ij} represents the rank that individual j in group i is positioned; N equals to the total number in all of the groups. $\bar{r}_{i.}$ equals to the average rank of the whole individuals of group i; $\bar{r}$ equals to the mean of the whole r_{ij}.

Table 25 Statistics of LE1 and LE2's co-occurrence frequency with all of the verbs

	LE1	LE2
count	1220	1220
mean	8.531967	0.571311

std	36.55057	4.273158
min	0	0
25%	0	0
50%	0	0
75%	4	0
max	884	109
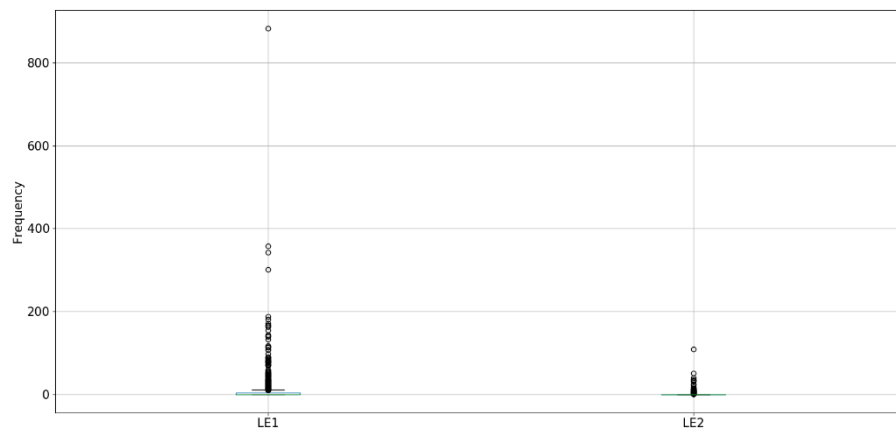
Figure 18 Boxplot of LE1 and LE2's co-occurrence frequency with all verbs

It is clearly observed from the table and figure that most of the data are zero although LE1's mean is obviously bigger than LE2's. The rejection of normal distribution of the co-occurrence frequencies denies the possibility to use mean as a standard to measure their difference. Here follows the results of Kruskal-Wallis H-test:

Table 26 Kruskal-Wallis H-test between LE1 and LE2

H	$p-\text{value}$
339.4425	0.0000

The almost zero p-value (8.44×10^{-76}) denies the original hypothesis that LE1 and LE2's samples are the same and from the same population.

In this section, Kruskal-Wallis H-test is implemented to validate whether LE1 and LE2 should be merged. It is concluded that their distributions are significantly different and thus should be isolated from each other. Their difference is further demonstrated by their co-occurrences with various situation types.

More differentiations of LE1 and LE2 are about to be discussed in the corresponding sections. The following section only concerns LE1.

LE2 will be pushed off to the next section because of its close relationship with GUO.

4.2.2 LE1's interaction with situation types

Most of aspectual studies are with consensus that LE1 signals completeness or prefectivity of a situation. However, Smith (1991)'s differentiation of perfective viewpoint aspect and imperfective viewpoint aspect actually lifts the term of perfectivity or completeness onto the common feature of perfective aspectual markers such as LE1, GUO, verb reduplication etc. Dai (1997) narrows down LE1's function to mark the actuality of a situation and Xiao and McEnery (2004) make it much clearer by claiming that "actual –*le* signals the actualization of a situation with respect to a past, present or future reference time."

We agree with Xiao and McEnery (2004) but escape to use *actualization* because of the agentive hint and insist that LE1 marks the occurrence of a situation relative to a reference time which could be a point in the past, present and future compared with the speech time.

LE1's interaction with situation types' differentiation is isolated as below and we have no intention to imply that LE1 along achieves this influence. It should be noticed that different features interact with each other in a model and thus their contributions in the model to differentiate situation types are totally different from their contribution by themselves.

Table 27 Activity vs. state by LE1

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
LE1	0.2989	0.172	1.738	0.082	1.348	-0.038	0.636

Since p-value is bigger than 0.05, there is no significant difference between activity verbs and state verbs on their co-occurrence frequencies with LE1 and thus their ability to co-occur with LE1 is the same.

Table 28 Transition vs. state by LE1

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
LE1	0.8799	0.193	4.567	0.000	2.411	0.502	1.257

The zero p-value shows that the difference between transition verbs and state verbs is significantly different.

128

The positive coefficient shows that transition verbs has positive relationship with LE1's co-occurrence frequency, which implies that state verbs have negative relationship with LE1's co-occurrence frequency.

It then follows that the higher is the co-occurrence frequency, the higher is the possibility that a verb is to be predicated as transition; the lower the co-occurrence frequency is, the higher is the possibility that a verb is to be predicted as a state or activity.
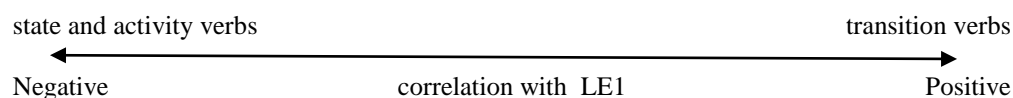
Figure 19 The relationship between LE1 and situation types

According to odds ratio, if a verb were to increase its co-occurrence frequency with LE1 by one unit, the odds for predicting it as a transition with respect to predicting it as a state would be expected to increase by 141% (2.411-1) while holding all other variables in the model constant.

Since state and activity verbs are [-Telic] verbs and transition verbs are [+Telic] verbs, LE1 is better to be explained by the feature of [+Telic].

4.3 LE2 and GUO

While LE1 is thoroughly studied in aspectual studies, LE2 gains far less attention except Xiao and McEnery (2004). LE2 is very different from LE1 syntactically

and semantically. The distribution difference has been validated in the last section, now we turn to the semantic difference.

(28) a.她去了学校。

ta1   qu4   le   xue2xiao4   .

she   go   LE1   school   .

'She went to school.'

b.她去学校了。

ta1   qu4   xue2xiao4   le   .

she   go   school   LE2   .

'She has gone to school.'

c.他去过学校。

ta1   qu4   guo4   xue2xiao4   .

he   go   GUO   shcool   .

'He has been to school before.'

In (28a), LE1 only marks the occurrence of *qu4* (go) but the LE2 in (28b) supply us with a contrasting between the current state and the previous state. In

130

the previous state, she has not gone to school but now, which is the speech time, she has gone there.

This is just opposite to GUO which also supply us with a contrasting between the previous sate and current state, which is illustrated in (28c). In the previous state, he is in school, but now, which is the speech time, he is not there anymore.

Both of LE2 and GUO imply a contrasting between the previous state and current state but with an opposite implication. This is validated in our data. LE2 is firstly depicted.

Table 29 Activity vs. state by LE2

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
LE2	-2.9244	0.617	-4.740	0.000	0.054	-4.134	-1.715

The zero p-value demonstrates that LE2 is able to be used to differentiate between activity and state.

The negative coefficient shows that activity is with negative correlation with LE2's co-occurrence frequency while state is with positive correlation with LE2's co-occurrence frequency.

According to odds ratio, if a verb were to increase its co-occurrence frequency with LE2 by one unit , the odds for predicting it as an activity with respect to predicting it as a state would be expected to decrease by 94.6% (0.054 -1) while holding all the other variables in the model constant.

131

Table 30 Transition vs. state by LE2

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
LE2	0.0301	0.06	0.499	0.618	1.031	-0.088	0.148

The p-value, which is bigger than 0.05, shows that there is no significant difference between the co-occurrence frequency of transition with LE2 and that of state with LE2.

The relationship between LE2 and situation types can be depicted as follows:

activity verbs   state and transition verbs

Negative   correlation with LE2   Positive
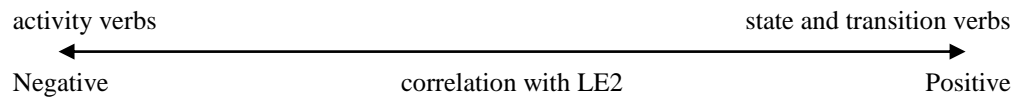
Figure 20 The relationship between LE2 and situation types

C. N. Li and Thompson (1989) points out that LE2 regards a situation, no matter stative or non-stative, to be state and build a contrasting between this state and the previous state which is just the opposite of the speech time. This makes state verbs quite compatible with LE2.

Transition verbs *per se* indicate a change of state, which is also compatible with LE2 (Moens, 1987; Xiao & McEnery, 2004).

Now, let's turn to GUO which is just in the opposite side of LE2.

Table 31 Activity vs. state by GUO

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
GUO	0.7654	0.331	2.313	0.021	2.150	0.117	1.414

The p-value which is smaller than 0.05 illustrates that the difference between activity verbs' co-occurrence frequency with GUO and state verbs' co-occurrence frequency with GUO is significant.

The positive coefficient shows that activity is with positive correlation with GUO while state is with negative correlation with GUO.

According to odds ratio, if a verb were to increase its co-occurrence frequency with GUO by one unit, the odds for predicting it as an activity with respect to predicting it as a state would be expected to increase by 115% (2.150 -1) while holding all the other variables in the model constant.

Table 32 Transition vs. state by GUO

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
GUO	0.3204	0.359	0.893	0.372	1.378	-0.383	1.024

The high p-value shows that there is no significant difference between transition verbs' co-occurrence frequency with GUO and state verbs' co-occurrence frequency with GUO.

It then follows that the relationship between LE2 and situation types can be depicted as follows:

state and transition verbs   activity verbs

$\longleftrightarrow$

Negative   correlation with GUO   Positive
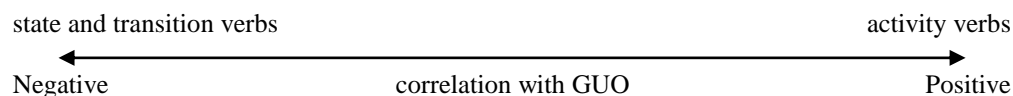
Figure 21 The relationship between GUO and situation types

It surprises us that the tendencies of LE2 and GUO are complementary with each other.

Both Chao (1968) and C. N. Li and Thompson (1989) maintain that GUO implies that an event has occurred or experienced at least one time and stress the repeatability or iterativeness of a situation. Zhang (1993) has noticed that only repeatable situation types are able to co-occur with GUO while Xiao and McEnery (2004) maintain that [+Repeatable] is a sufficient instead of necessary condition for a situation to take GUO.

They are on the right track on the interaction between GUO and situation types. [Repeatable] or [Iterative] is further proposed to be a conceptual feature of situation type.

Since activity verbs tend to be repeatable while state and transition verbs are not, LE2 is better explained by the feature of [-Repeatable] because its negative correlation with activity and positive correlation with state and transition. GUO is just on the contrary and better explained to be with a positive correlation with [+Repeatable] or [+Iterative].

134

4.4 ZHE and ZAI

ZHE and ZAI are two typical imperfective aspectual markers which take a view of the duration of a situation and thus they are supposed to be related with the feature of [+Durative].

Xiao and McEnery (2004) insists that ZAI is different from ZHE and that ZAI is related with [+Dynamic] while ZHE is related with [+Durative]. However, these two aspectual markers show no significant difference on their co-occurrence with situation types based on our data. Let's start from ZHE.

Table 33 Activity vs. state on ZHE

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
ZHE	-0.1125	0.081	-1.383	0.167	0.894	-0.272	0.047

The high p-value denies the significant difference between activity verbs' co-occurrence frequency with ZHE and state verbs' co-occurrence frequency with ZHE.

Table 34 Transition vs. state on ZHE

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
ZHE	-20.564	5.634	-3.65	0.000	0.000	-31.607	-9.521

The zero p-value shows that the difference between transition verbs' co-occurrence frequency with ZHE and state verbs' co-occurrence frequency with ZHE is significantly different.

According to odds ratio, if a verb were to increase its co-occurrence frequency with ZHE by one unit, the odds for predicting it as an transition with respect to predicting it as a state would be expected to decrease by 100% (0.000 -1) while holding all other variables in the model constant.

The coefficient shows a negative correlation between transition and ZHE and a positive correlation between state and ZHE.
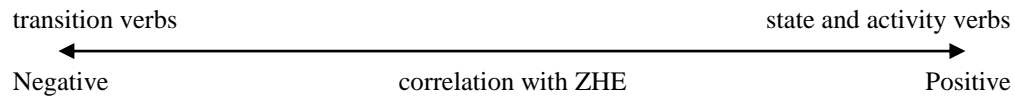
Figure 22 The relationship between ZHE and situation types

ZAI's co-occurrence frequency with verbs also shows the similar tendency.

Table 35 Activity vs. state on ZAI

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
ZAI	0.0215	0.089	0.241	0.810	1.022	-0.153	0.196

The high p-value rejects the significant difference between activity verbs co-occurrence frequency with ZAI and state verbs' co-occurrence frequency with ZAI.

Table 36 Transition vs. state on ZAI

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
ZAI	-0.8504	0.432	-1.967	0.049	0.427	-1.698	-0.003

The p-value, although is close to 0.05, confirms the significant difference between transition verbs co-occurrence frequency with ZAI and state verbs' co-occurrence frequency with ZAI.

The coefficient shows that transition holds a negative correlation with ZAI while state holds a positive correlation with ZAI.

According to odds ratio, if a verb were to increase its co-occurrence frequency with ZHE by one unit, the odds for predicting it as an transition with respect to predicting it as a state would be expected to decrease by 57.3% (0.427 -1) while holding all other variables in the model constant.

The tendency of the interaction between ZAI and situation types is in consistent with ZHE:

transition verbs   state and activity verbs

←——————————————————————————————————————→

Negative   correlation with ZAI   Positive

Figure 23 The relationship between ZAI and situation types

It then follows that both ZHE and ZAI hold positive correlation with [+Durative].

4.5 ZHENGZAI

ZHENGZAI gains little attention in previous aspectual studies. It is either ignored or merged with ZAI such as in Guo (1993) and Yu et al. (1996). However, our data shows that ZHENGZAI is different from ZAI.

137

Xiao and McEnery (2004) stresses that ZAI instead of ZHE is the best aspectual marker that is related with [+Dynamic]. However, the statistics in last chapter has validated that there is no significant different between ZHE and ZAI on their interaction with different situation types and thus both of them are actually related with [+Durative] instead of [+Dynamic].

In this section, it will be shown that ZHENGZAI, instead of ZAI, is supposed be related with [+Dynamic].

Table 37 Activity vs. state by ZHENGZAI

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
ZHENGZAI	1.1802	0.460	2.566	0.010	3.255	0.279	2.082

The p-value shows that there is significant difference between activity verbs' co-occurrence frequency with ZHENGZAI and state verbs' co-occurrence frequency with ZHENGZAI.

The positive coefficient illustrates that activity holds a positive correlation with ZHENZAI while state holds a negative relation with ZHENGZAI.

According to odds ratio, if a verb were to increase its co-occurrence frequency with ZHENGZAI by one unit, the odds for predicting it as an activity with respect to predicting it as a state would be expected to increase by 225.5% (3.255 -1) while holding all other variables in the model constant.

Table 38 Transition vs. state by ZHENGZAI

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
ZHENGZAI	1.219	0.474	2.571	0.010	3.384	0.29	2.148

The p-value shows that there is a significant difference between transition verbs' co-occurrence frequency with ZHENGZAI and state verbs' co-occurrence frequency with ZHENGZAI.

The positive coefficient illustrates that transition holds a positive correlation with ZHENZAI while state holds a negative relation with ZHENGZAI

According to odds ratio, if a verb were to increase its co-occurrence frequency with ZHENGZAI by one point, the odds for predicting it as an transition with respect to predicting it as a state would be expected to increase by 238.4% (3.384 -1) while holding all other variables in the model constant.

Since the difference between activity and transition on their co-occurrence with ZHENGZAI is not clear by setting state as the reference type, we further displays the result isolated from the analysis by setting transition as the reference type:

Table 39 Activity vs. transition by ZHENGZAI

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
ZHENGZAI	-0.036	0.139	-0.26	0.795	0.960	0.73	1.27

The high p-value shows that there is no significant difference between activity verbs' co-occurrence frequency with ZHENGZAI and transition verbs' co-occurrence frequency with ZHENGZAI.

The tendency of the interaction between ZHENGZAI and situation types can be inferred as:

state verbs		activity and transition verbs
Negative	correlation with ZHENGZAI	Positive
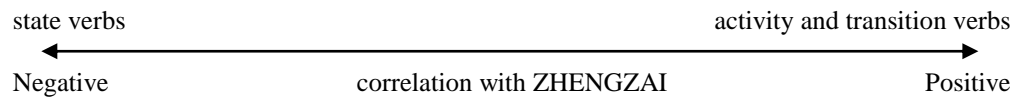
Figure 24 The relationship between ZHENGZAI and situation types

Since activity and transition tend to be related with [+Dynamic] and state is related with [-Dynamic], ZHENGZAI can be inferred to be related with [Dynamic].

4.6 Conclusion

In this chapter, the classifying indicators and features are validated through statistical methodologies.

On one hand, each aspectual makers' effect on classifying situation type is validated based on the result of multinomial logistic regression analysis. Our analysis confirms whether one aspectual marker is able to differentiate between different situation types or not. In addition, we also quantify their effect on differentiating different situation types by interpreting the odds ratios.

On the other hand, the relationship between aspectual markers and situation type features are confirmed.

Table 40 Relationship between aspectual markers and classifying features

Aspectual markers	Situation type feature
LE1	[Telic]
LE2, GUO	[Repeatable] or [Iterative]
ZHE, ZAI	[Durative]
ZHENGZAI	[Dynamic]

Our model for classifying situation type are confirmed and the dataset of verbs and their situation types has been constructed which can be directly used to make further studies.

The next chapter is dedicated to the evaluation of the dataset. The dataset is applied to train classification models to realize the prediction of verbs situation types which have rarely been done in previous studies.

Chapter 5

Evaluation of Mandarin lexical situation type dataset

After the validation of our situation type system, this chapter is dedicated to the Evaluation of the annotated dataset which contains more than 1000' verbs (senses) and their situation types. It should be noticed that our goal in this chapter is not to achieve high performance but to evaluate the efficiency of verb senses' situation type dataset although comparison is implemented between baseline and other similar studies.

In this chapter, we firstly review previous computational studies to illustrate the importance of situation type in temporal reasoning and in "language interpretation, generation, summarization, information retrieval and machine translation tasks" (Eric V. Siegel & Mckeownt, 2002) and also to supply comparison baselines for our predicting results.

Since the training of the prediction model is based on vector space models, an introduction on it is unfolded in section 5.2.

Finally, the prediction of verbs' situation type is implemented in section 5.3 and 5.4. Both selected features and automatic word embedding vectors are used.

It should be noticed that the hierarchical clustering in chapter 3 is based on manually annotation data and thus serves as a supervised way to classify verbs into different types. Since it has already been used to construct our situation type

system, there is no need to implement evaluation based on supervised features. This chapter only concerns about features extracted in an unsupervised way or semi-unsupervised way.

5.1 Previous computational studies on aspect

In this section, we mainly review computational studies of the application of situation type data and the prediction of situation type covering languages of English, Chinese, French, Italian, and Portuguese.

The former one covers studies of Kazantseva and Szpakowicz (2010), Costa and Branco (2012) and Meyer, Grisot, and Popescu-Belis (2013).

The latter one covers studies of Eric V. Siegel and Mckeownt (2002),Cao et al. (2006), Zarcone and Lenci (2008), Meyer et al. (2013), Xu (2015) and Falk and Martin (2016).

5.1.1 The application of situation type

Kazantseva and Szpakowicz (2010) strive to achieve the automatic summarizing of short stories. Lexical aspect and sentential aspect are used to select expressions or sentences which set out the background of a story.

Kazantseva and Szpakowicz (2010) extracts summarizing information candidate based on the following hypotheses:

(i) The candidate should discuss key things such as the characters and locations;

(ii) The candidate should contain background descriptions rather than events.

The rules in Kazantseva and Szpakowicz (2010) are designed as follows:

Table 41 Manual rules of Kazantseva and Szpakowicz (2010)

(i) Rule 1:

 if a clause contains a character mention as subject **and** a temporal expression of type enactment (e.g. *ever*, *never*, *always*):

   return True;

(ii) Rule 2:

if a clause contains a character mention as subject or object **and** a stative verb:

   return True;

(iii) Rule3:

if a clause is in progressive tense:

   return False.

In "Rule2", the situation type of a verb should be identified firstly. Thus in Kazantseva and Szpakowicz (2010), they maintain that "the model states that verbs are categorized into aspectual classes based on……dynamicity, durativity and telicity."

Costa and Branco (2012) resort to aspectual information in Portuguese to improve the performance of temporal relation reasoning tasks of TempEval developed by Verhagen et al. (2007) and Verhagen, Sauri, Caselli, and Pustejovsky (2010).

TempEval follows TimeML scheme which is set by Pustejovsky et al. (2003) as followsing:

144

Table 42 Sample of TimeEval

```
<s>
In Washington
<TIMEX3 tid="t53" type="DATE" value="1998-01-14">
today
</TIMEX3>,
the Federal Aviation Administration
<EVENT eid="e1" class="OCCURRENCE" stem="release" aspect="NONE"
tense="PAST" polarity="POS" POS="VERB">
released
</EVENT>
air traffic control tapes from
<TIMEX3 tid="t54" type="TIME" value="1998-XX-XXTNI">
the night
</TIMEX3>
the TWA Flight eight hundred
<EVENT eid="e2" class="OCCURRENCE" stem="go" aspect="NONE"
tense="PAST" polarity="POS" POS="VERB">
went
</EVENT>
down.
</s>
<TLINK lid="l1" relType="BEFORE" eventID="e2"
relatedToTime="t53"/>
<TLINK lid="l2" relType="OVERLAP" eventID="e2"
relatedToTime="t54"/>
```

The TempEval task is to identify the `relType` in each sentence (Costa & Branco, 2012). Costa and Branco (2012) insist that the aspectual type information of verbs is useful to improve the performance on this task. TimeML annotates the situation type information. For example, *released* in the above sample is tagged as `class="OCCURRENCE"`.

However, TimeML only differentiates between state and occurrence (dynamic situations), which is far from enough and it is quite hard to annotate situation type of verbs according to Costa and Branco (2012).

145

Thus, they resort to the co-occurred indicators such as aspectual markers and time adverbials to express aspectual information. The co-occurrence information is extracted from the web.

The improved performance, compared with the baselines, shows that the aspectual indicator co-occurrence information do help to identify temporal relation.

Table 43 Result of Costa and Branco (2012)

Classifier		Task		
		A	B	C
trees.J48	Baseline	0.57	0.77	0.53
	With best indicator			0.55
rules.JRip	Baseline	0.60	0.76	0.51
	With best indicator	0.61		0.54
lazy.KStar	Baseline	0.54	0.70	0.52
	With best indicator		0.73	0.53
bayes.NaiveBayes	Baseline	0.50	0.76	0.53
	With best indicator	0.53		0.54
functions.SMO	Baseline	0.55	0.79	0.54
	With best indicator	0.56		0.55

Meyer et al. (2013) apply verbs' situation type to help to improve the machine translation of English verbs in simple past tense into French. The automatic prediction of English verbs' situation type is introduced in next section.

146

Verbs are divided into narratives (non-stative) and non-narratives (stative) in their study and are tagged accordingly. They create three different systems with or without situation type's tags:

Table 44 Annotation forms of Meyer et al. (2013)

1. BASELINE SMT: on wednesday the ˘cssd declared the approval of next year's budget to be a success. the people's party was also satisfied.

2. TAGGED SMT: on wednesday the ˘cssd declared**-Narrative** the approval of next year's budget to be a success. the people's party was**-Non-narrative** also satisfied.

3. FACTORED SMT: on wednesday the ˘cssd declared|**Narrative** the approval of next year's budget to be a success. the people's party was|**Non-narrative** also satisfied.

The results show that FACTORED SMT improves the performance significantly compared with the BASELINE.

Table 45 Result of Meyer et al. (2013)

Translation model	BLEU	TER
BASELINE	21.4	61.9
TAGGED	21.3	61.8
FACTORED	**21.6**	**61.7**

5.1.2 The prediction of situation type

Studies on the prediction of situation types is dichotomous based on their basic hypotheses.

For Eric V. Siegel and Mckeownt (2002), Cao et al. (2006) and Xu (2015), they actually map the sentences' situation type to the constituting verb and deny

147

the compositionality and recursiveness of situation type. We have stated the problem of this assumption from a linguistics point of view in the first chapter. They are actually predicting the sentence's situation type and the accuracy is predictably high as some of the indicators are decisive on sentential situation types such as LE2, which is not the case in lexical level:

(29) 他知道了。

    ta1   zhi1dao4   le   .

    he   know   LE2   .

    'He has known it.'

LE2, which implies a change of state, is decisive on the sentential situation type of (29) which is a transition but it is not able to predict the verb's situation type which is a state verb. We intend to call their implementation as sentential situation type prediction.

For Meyer et al. (2013) and Falk and Martin (2016), they pertain the independence of the constituting verb's situation type from the whole sentence, which is shared by us. We intend to call their implementation as verb situation type prediction.

Both sentential situation type and verb situation type can be predicted with the similar features but their difficulties are totally different which can be shown by

the accuracy difference between sentential situation type prediction and verb situation type prediction.

5.1.2.1 Sentential situation type prediction

Eric V. Siegel and Mckeownt (2002) become one of the earlier birds to do the prediction of situation types based on computational methods. They splits situation type into state and event (dynamic) based on the feature of [Stative]. Events are quartered by two features including [Telic] and [Durative] naming extended-culminated event (accomplishment), extended-non-culminated event (activity), punctual event (achievement) and point event (semelfactive). Semelfactives are verbs like *cough*, *knock*, *wink* and *hiccup* which happen once and only once which is introduced in Comrie (1979).

Eric V. Siegel and Mckeownt (2002) pick fourteen features which he called linguistics indicators to distinguish different situation types.

Table 46 Eric Victor Siegel (1998)'s linguistics features

Linguistics indicator	Example
frequency	-
not or *never*	She can **not** explain why.
temporal adverb	I saw to it **then**.
no subject	He was admitted to the hospital.
past/pres. particle	…blood pressure going up.
duration *in*-PP	She built it **in an hour**.
perfect	They **have** landed.
present tense	I am happy.
progressive	I am behaving myself.
manner adverb	She studied **diligently**.
evaluation adverb	They performed **horribly**.
past tense	I was happy.
duration *for*-PP	I sang **for ten minutes**.
continuous adverb	She will live **indefinitely**.

It should be noticed that although partitioning situation types into five categories, Eric V. Siegel and Mckeownt (2002)'s prediction actually predict only three of them. They implement two binary experiments instead of one multi-class classification experiment to do the prediction. State and event are differentiated from each other in the first experiment while culminated event

150

(accomplishment and achievement) and non-culminated event (semelfactive and activity) are distinguished from each other in the second experiment.

Decision tree, binary logistic regression and generic programming are implemented to do the classification task:

Table 47 Results of Eric Victor Siegel (1998)

	Overall accuracy	State Recall	State Precision	Event Recall	Event Precision
Decision tree	93.9%	74.2%	86.4%	97.7%	95.1%
Genetic programming	91.2%	47.4%	97.3%	99.7%	90.7%
Logistic regression	86.7%	34.2%	68.3%	96.9%	88.4%
Baseline	83.8%	0.0%	100.0%	100.0%	83.8%

Taking overall accuracy as the standard, decision tree reaches the best performance when classifying verbs into state and event. When applying these classification methods to distinguish between culminated and non- culminated events, CART (a kind of decision tree) achieves an overall accuracy of 74% with a baseline of 63.3%. But just as Eric V. Siegel and Mckeownt (2002) said, the performance of the classification method depends on tasks.

Eric V. Siegel and Mckeownt (2002) calculate the linguistics features' coefficients and their significance on distinguishing one event from another and conclude that manner adverb and duration *in*-PP rank the highest in deciding the situation type of a verb in both of the two experiments.

151

However, he fails to notice the p-value of the t test for the significance of these linguistics features. Manner adverb is not significant in the experiment of differentiating event from state because the p-value of 0.1681 is even bigger than 0.1; duration *in*-PP is not significant in the experiment of differentiating culminated event and non-culminated event with a p-value of 0.0626 which is bigger than 0.05 although its weight is -0.1207664 which is very big compared with other features.

Another problem of Eric V. Siegel and Mckeownt (2002)'s experiments lies in that the classification of situation type is implemented by two individual experiments instead of one complete classification experiment, which makes the overall performance of his model unreliable (Zarcone & Lenci, 2008).

Cao et al. (2006) predict sentential situation type in Chinese. By combining 49 linguistics indicators, they achieve an accuracy of 94.92% in close evaluation and 93.41% in open evaluation.

Zarcone and Lenci (2008) predict sentential situation type in Italian using both supervised way and unsupervised way. By sufficiently selecting linguistics features that are relevant to situation types, they construct the supervised matrix. Linguistics features include adverbial features, morphological features, syntactic and argument features. Maximum entropy model is implemented to carry on the experiment of classifying verbs into four classical types instead of three and achieves an overall accuracy of 85.4%.

Xu (2015) implements experiments not only on situation type but also on modalities etc. For situation type including static, dynamic, achievement and accomplishment in his study, Xu (2015) achieves an accuracy of 77.95%.

5.1.2.2 Verbe situation type prediction

Meyer et al. (2013) use lexical aspect to help to improve the performance of machine translation. Only narrative (non-stative) and non-narrative (stative) are differentiated in their study.

Meyer et al. (2013)'s dataset is depicted as follows:

Table 48 The distribution of verbs in Meyer et al. (2013)'s dataset

	Narrative	Non-narrative
Training	230	228
Test	75	43

POS tag, syntactic information and 66 temporal markers are used as features to train and predict the verb's situation type. The situation type information generated from TimeML parser (Verhagen & Pustejovsky, 2008) is also treated as a feature. MaxEnt and CRF classifiers are used for the classification test and the result is as follows:

Table 49 The situation type classification result of Meyer et al. (2013)

Model	Recall	Prec.	F1
MaxEnt	0.71	0.72	0.71
CRF	0.30	0.44	0.36
Baseline		0.6356	

Falk and Martin (2016) attach situation type to verb's senses which is shared by us. They extracted 167 frequent verbs with 1199 senses from LVF which is a French lexical resource (Falk & Martin, 2016) and each sense of the verb is manually tagged with situation type through Vendlerian 4-way categorization. They construct a 'telicity scale' of eight values and senses of verbs are tagged accordingly. For example, as stative verbs, *know* ranks higher than *think*.

All the morph-syntactic and semantic features are also extracted from LVF to train and predict the situation types of the senses of a verb. The result is as follows:

Table 50 Result of Falk and Martin (2016)

Algorithm	Complete features	Selected features
trees.j48	61.80	63.00
rules.jrip	63.89	61.56
lazy.kstar	62.89	**67.47**
functions.libsvm	62.72	61.13
bayes.naivebayes	60.22	65.80
baseline	48.37	48.37

5.2 Vector space models

Using vector to represent a word has been a tradition for a long time in computer science area. For instance, we use hash table to record all of the words in a corpus and it's called one-hot representation. But there will be no other information except the position information for a word in the dictionary. Vector is able to overcome this problem to some extent as it is a two-order tensor. Since it is two-order, we can construct a matrix. The row refer to words while the column refer to context ranging from bigram to n-gram.

VSM (vector space model) is developed for the SMART information retrieval system since Salton (1971) and Salton, Wong, and Yang (1975). It has been quickly and widely used for the measurement of word similarity and document similarity and latter be used for pattern or relation similarity (Turney & Pantel, 2010). Other approaches on measuring the similarity between words depend on language resources like WordNet and usually their calculations are based on the number of nodes that between two words in a hyponymy hierarchy constructed according to the language resource such as Miller, Leacock, Tengi, and Bunker (1993) for English and Q. Liu and Li (2002) for Chinese. However, VSMs don't need database like WordNet and the vector matrices are constructed based on corpus (with or without POS) and the result is better (Turney & Pantel, 2010).

155

5.2.1 The semantic hypotheses behind VSMs

VSM is also called distributional semantic model (DSM) since its basic hypothsis is about words' distribution. The distributional hypothesis holds that words have similar context tend to be more similar semanticly (Harris, 1954). This makes vector matrix an excellent two-order tensor that can be used to reprsent words under different contexts window size ranging from construction to sentence and even document.

Distributional semantics hypothsis insists that words that have similar context tend to have simialr meaning (Harris, 1954). In other words, words that have similar vectors in a row in the vector matrix tend to have similar meaning.

And there are several inferences from this basic hypothesis:

(i) Bag of words hypothesis: if words vectors in a document (d1) are simmiar with words vectors in anther document (d2), two documents are simialr with each other. (Salton et al., 1975). In other words, if two columns of vectors is similar with each other, they tend to be similar with each other semanticlly.

Lin and Pantel (2001) and Turney and Littman (2003) latter propose another two hypotheses aiming to solving different appilcation tasks:

(ii) Extended distributional hypothesis: patterns that co-occur with similar pairs of words are similar with each other. In other words, in a word pair-pattern matrix, if a column of vectors (hence a particular pattern) is simmialr to another

column of vectors (hence another pattern), they tend to represent similar meaning.

(iii) Latent relation hypothesis: pairs of words co-occur in similar patterns tend to hold similar relations between them. In other words, if words pairs have similar row vectors in a pair-pattern matrix, they are supposed to hold similar semantic relations.

One of the obvious problemes for these hypotheses is that all of them ignore the word-order's role in deciding the semantics of a language structure compared to the ditributional anaylysis in linguistics (Turney & Pantel, 2010).

Some of the VSMs optimize the model to be sensitive to word order such as Mitchell and Lapata (2008). However, the results seem to be no better or even worser than VSMs without considering the word order.

5.2.2 Different VSM matrices

Turney and Pantel (2010) classify the matrices of VSM into three different types based on the difference of the items in rows and columns in the matrices including word-document, word-context and pair-pattern. Document and context are just different window sizes for computational linguistics and sometimes, a window of 5 words (that is five words before the target word and five words after the target word) may exceed a document if the document only contains one sentence with less than ten words to be extrmely. And thus the three different

types of matrices can be generalized by two inclduing word-context and pair-pattern.

The VSM was first used by Salton et al. (1975) to measure the similarity between two documents. They constructe a word-document matrix. Based on the bag of words hypothesis which states that documents with similar word vectors tend to be simialar with each other. Thus it can be meaured by comparing the vectors in one column which represent one document.

Deerwester, Dumais, Furnas, Landauer, and Harshman (1990) noticed that the same matrices can be used for the measurement of word similarity based on word-context matrix. The matrix is the same with Salton et al. (1975)'s, which is with words as its rows and columns as its context or documents. However, Salton et al. (1975) use the matrix to measure the similarity between columns (hence the documents or contexts) while Deerwester et al. (1990) use it to measure the similarity between rows (hence the words). They are based on different hypothesis. While Salton et al. (1975)'s method is based on bag of words hypothesis, Deerwester et al. (1990) is based on distributional hypothesis.

But latter when evaluating the model with different length of context on TOEFL test (Landauer & Dumais, 1997), the performance tends to decrease with the expanding of the window size (Turney & Pantel, 2010). In other words, for the measurement of word similarity, only the instantly near context of the target

words contribute to their semantics while words far way from the target word is not able to help to measure the simialrity between words.

Lin and Pantel (2001) use pair-pattern matrix to measure the similarity between patterns. In the pair-pattern matrix, the words represented by rows are replaced by word pairs with certain relations such as "*carpernter:wood*" and "*mason:stone*" while the documents represented by columns are replaced by patterns like "*X cuts Y*" and "*X works with Y*". They use this matrix to measure the similarity between columns (hence the patterns). They propose the extended distributional hypothesis to be verified by this task.

Following the same steps as in Deerwester et al. (1990), we can measure the similarity between rows (hence the word pairs) in pair-pattern matrix. Turney (2008) proposes that word pairs with similar row vectors in a pair-pattern matrix tend to have similar semantic relations which forms the latent relation hypothsis that can be used to predict the semantic relation between word pairs.

All of matrices constructed above are based on the frequncy of words or word pairs that occur in contexts or patterns. Baroni et al. (2014) call these matrices count-matrices.

However, we have to weight the vectors according to their informativeness. In information theory, a surpring event has higher information than expected one and thus *tf-idf* (term frequncy-inversed document frequncy) or pointwise mutual information (PMI) is introduced to weight the vectors' informativeness.

What's more, since there is data sparcity, we have to smooth the model. And because of noise vector and zero vector, we need to decrease the dimmenstionality to decrease the algorithem complexity.

Instead of counting the words or word pairs occurance in a context or pattern, Mikolov, Chen, et al. (2013) predict a word's vector in a context based on neural network.

They propose two arcitechtures to predict the word's vectors including Skip-gram and CBOW (continuous bag of words).

CBOW architecture predicts the current word based on context while the Skip-gram use the current word to predict the surrounding words (hence the context). No matter which one, the architecture is able to construct a word-context matrix efficiently. These two models learn the word vectors based on a simple neural network and thus the complexity for computing is much lower and "can learn a model from billions of words in hours"(Mikolov, Sutskever, et al., 2013).

Baroni et al. (2014) make a comprehensive comparison between count models and predict models. The comparison includes their performances on the task of semantic relatedness, synonym detection, concept categorization, selection preferences and analogy. Baroni et al. (2014) Following are the results:

Table 51 Comparison between count and predict Models in Baroni et al. (2014)

	rg	ws	wss	wsr	men	toefl	ap	esslli	battig	up	mcrae	an	ansyn	ansem
						best setup on each task								
cnt	74	62	70	59	72	76	66	84	98	41	27	49	43	60
pre	84	75	80	70	80	91	75	86	99	41	28	68	71	66
						best setup across tasks								
cnt	70	62	70	57	72	76	64	84	98	37	27	43	41	44
pre	83	73	78	68	80	86	71	77	98	41	26	67	69	64
						worst setup across tasks								
cnt	11	16	23	4	21	49	24	43	38	-6	-10	1	0	1
pre	74	60	73	48	68	71	65	82	88	33	20	27	40	10
						best setup on rg								
cnt	(74)	59	66	52	71	64	64	84	98	37	20	35	42	26
pre	(84)	71	76	64	79	85	72	84	98	39	25	66	70	61
						other models								
soa	86	81	77	62	76	100	79	91	96	60	32	61	64	61
dm	82	35	60	13	42	77	76	84	94	51	29	NA	NA	NA
cw	48	48	61	38	57	56	58	61	70	28	15	11	12	9

Among the table, *rg*, *ws*, *wss*, *wsr* and *men* are datasets used for the evaluation of semantic relatedness; *toefl* is a dataset used for the task of synonym detection; *ap*, *essli*, *battig* are datasets used for the task of categorization; *up* and *macrae* are datasets used for the task of selectional preferences; *an*, *ansyn* and *ansem* are used for the task of analogy.

It is clear that for most of the tasks and datasets, predict model (word2vec) performs much better than the count one and on seven datasets, its performance exceeds the state-of-art one.

5.2.3 Application of different VSM matrices

Different VSMs can be used to achieve different tasks. While the word-context matrices are usually used to measure word similarity and cluster words into different groups, the pair-pattern matrices are usually used to measure the relation similarity and implement relational clustering or classification.

The basic application for vectors is to measure the similarity and thus we will introduce how to measure similarity between vectors firstly. It can be used to measure the similarity between words, patterns and documents. The most popular way for measuring similarity is to take the vectors' *cosine* (Turney & Pantel, 2010).

Instead of considering the length of vectors, *cosine* only measures the angle between vectors:

Cosine similarity illustrated. $\text{sim}(d_1, d_2) = \cos\theta$.
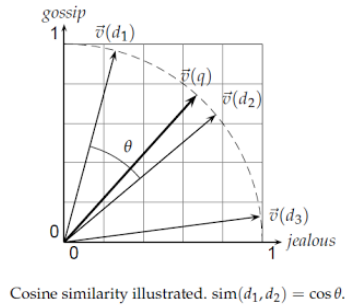
Figure 25 Vector cosine

5.2.3.1 Application of Word-Context Matrices

Deerwester et al. (1990) use the word-context matrix to measure the similarity of words. But there is no evaluation for their model and thus Landauer and Dumais (1997) introduce TOEFL test, which is designed for human, to evaluate the result of the measurement of word similarity based on word-context matrix. The TOEFL dataset contains 80 multiple-choice questions, each of which is with one target word and four synonym candidates. Subjects are required to choose the correct synonyms from the four candidates. Such as for *levied*, there are four candidates: *imposed*, *believed*, *requested* and *correlated* while *imposed* is the correct answer. Deerwester et al. (1990) achieve 64.4% precision compared to the 64.5% precision of human. This model considers the context with a length of 151 words.

Lund and Burgess (1996) measure the word similarity based on a word-context with shorter context containing ten words while Rapp (2003) only consider context with four words. The latter achieves 92.5% precision.

Above are all count models. Baroni et al. (2014) completely compare different models with the predict model (CBOW) and achieve 91% precision.

Other applications include word classification and clustering. Turney and Littman (2003) used a word-context matrix to classify words into positive (*honest*, *intrepid*) and negative (*disturbing*, *superfluous*).

5.2.3.2 Application of Pair-Pattern Matrices

Lin and Pantel (2001) serve as one of the earlist studies that introduce the pair-pattern matrix for the purpose of measuring the similarity between patterns. It is a simulation of the document similarity because both of them are comparing the similarity between vectors in different clolumns. Based on this matrix, their algorithm is able to find similar patterns of "X solves Y" such as "Y is solved by X", "Y is resolved in X" and "X resolves Y".

Inversely, Turney and Littman (2003) introduce the pair-pattern matrix to measure the pair similarity based on the row vectors and Turney (2006) evaluate this approach with 374 multiple-choice analogy questions from SAT college entrance test. The algorithm achieves 56% precision, compared to the 57% precision of human.

We tend to introduce Mikolov, Chen, et al. (2013)'s method which also concerns the analogy task. Mikolov, Chen, et al. (2013) predict word vectors instead of word pair vectors and thus it is impossible for it to apply the same steps as Turney and Littman (2003). They transform the analogy task into a process of adding and subtracting vectors: *vector* (*mason*) − *vector* (*stone*) + *vector* (*carpenter*) ≈ *vector* (*wood*).

In this way, their algorithm achieves 66% performance while the state-of-art algorithm is 61% (Baroni et al., 2014).

It is safe to say that vectors predicted by Mikolov, Chen, et al. (2013) have improved a lot compared with the current count VSMs no matter in performance or in training speed (Baroni et al., 2014).

In the next two sections, VSM is applied to classify situation type based on both counting vectors and predicted vectors.

5.3 Classification based on selected vectors

In this section, we evaluate our verb situation type dataset based on selected feature vectors. We firstly introduce our selected features together with the construction of the vector matrix and then implement the experiments and finally analyze the results.

5.3.1 Features selection

The features selection is based on the study in previous chapters and also other aspectual studies.

5.3.1.1 Aspectual markers

All the aspectual markers discussed in the precious chapters are targeted including ZHE, LE1, LE2, GUO, ZAI, ZHENGZAI. However, it should be noticed that almost all of the aspectual markers have homographs. For example, there are at least three GUO:

(30) a.土地要价过高。

tu3di4   yao4   jia4   guo4   gao1   .

land   ask for   price   GUO   high   .

'The land is overcharged.'

 b.他拍过多部武侠片。

ta1   pai1   guo4   duo1   bu4   wu3xia2   pian1   .

he   shoot   GUO   many   CLA   swordsman   film   .

'He has ever filmed a lot of swordsman movies.'

 c.他们也过端午节。

ta1men2   ye3   guo4   duan1wu3   jie   .

they   also   celebrate   Duanwu   festival   .

'They also celebrate Duanwu festival.'

    Sinica corpus has annotated most of the aspectual markers with the tag of Di
or D in a high quality but some are still mis-tagged and thus a semi-supersized
way is applied to extract aspectual markers' co-occurrence information.

    We use ASP as an abbreviation for aspectual marker in the latter experiments.

5.3.1.2 Verb reduplication and syllable length

Verb reduplication has been mentioned in several aspectual studies including Smith (1991), Dai (1997) and Xiao and McEnery (2004) as an perfective viewpoint aspect. Since it is a mechanism instead of grammaticalized aspectual maker, we didn't include it into our classification system. However, it should be noticed that verb reduplication is a typical measurement of the feature of [Repeatable] which is in consistent with GUO, according to our supplementing multinominal logistic regression analysis. Thus, it is also chosen as a feature to evaluate the dataset.

Sinica corpus doesn't annotate verb reduplication and thus another way must be found. All of the reduplicated words and bigrams are extracted from the corpus together with the sentences they appear, and they are all manually checked by us since some are not real reduplicated examples:

(31) 爱的奇迹已打中打中我的眼睛。

ai4	de	qi2ji4	yi3	da3zhong4	da3zhong4	wo3	de	yan3jing	.
love	de	miracle	already	hit	hit	I	DE	eye	.

'The miracle of love has hit my eyes.'

After checking all of the examples, their frequency is attached to the corresponding verbs' vectors.

We use VV as an aberration for verb reduplication in the latter experiments.

The number of syllables is noticed by He (1992) and applied in his sub-classification of verbal situation types. This feature is simulated by the word length of the verbs. SYLL is chosen as an abbreviation for syllable length in the latter experiments.

5.3.1.3 POS tag

The part-of-speech tags in Sinica corpus are abundant in syntactic and semantic information. For example, VA represents "active intransitive verb" and VG represents classificatory verb. While the former one has clues for activity verb, the latter one is related with state verb although it is not always the case. Let's take *zuo4* (serve as), which is tagged with VG, as an example:

(32) 李天羽做了空军总司令。

Li3tian1yu3   zuo4   le   kong1jun1   zong3   si1ling4   .

Li Tianyu   serve as   LE1   air force   chief   commander   .

'Li Tianyu has been the commander-in-chief of the air force.'

We use POS as an abbreviation for in the latter experiments.

5.3.2 Classification settings and experiments

Three classifiers are applied to do the classification evaluation of our dataset including multinomial logistic regression (MNLogit), supporting vector machine (SVM) and artificial neural network (ANN). The configurations are as follows:

(i) For SVM, the penalty parameter is tested between 0-100 and finally it is set to 5; the kernel method is set to "rbf";

(ii) For ANN, the layers and knots on each layer are tested with different numbers and finally a five-layer sequential artificial neural network model is constructed. All layers are forward dense layer. RNN and CNN layers are not considered in this thesis. The activation function is "relu".

(iii) Multinomial logistic regression uses the default configuration in *sklearn* package.

Tenfold cross-validation is applied to guarantee the reliability of our evaluation.

Accuracy is used to compare between different classifiers and models while other performance scores including precision, recall and F1 are also reported for the best system.

Table 52 Performance of different classifiers based on selected vectors

	MNLogit	SVM	ANN
ASP	57.46%	55.74%	**58.36%**
ASP+POS	**67.54%**	62.79%	65.66%
ASP+POS+VV	**67.79%**	63.28%	67.38%
ASP+POS+VV+SYLL	**68.44%**	64.43%	67.54%
Basline		51.97%	

The baseline is set by dropping all of the verbs' senses into one situation type which account for the most part of the dataset.

Each kind of feature is able to improve the performance of the classifiers, which proves that these features are useful for predicting situation types in our dataset.

The best system achieves 68.44% accuracy with a 0.17 improvement from the baseline. Although the experiments results are not comparable since they are implemented on different dataset, we intend to repeat Meyer et al. (2013) and Falk and Martin (2016)'s result as follows because they are also predicting verb or verb senses situation type.

Table 53 The situation type classification result of Meyer et al. (2013)

Model	Recall	Prec.	F1
MaxEnt	0.71	0.72	0.71
CRF	0.30	0.44	0.36
Baseline		0.6356	

Table 54 Result of Falk and Martin (2016)

Algorithm	Complete features	Selected features
trees.j48	61.80	63.00
rules.jrip	63.89	61.56
lazy.kstar	62.89	**67.47**
functions.libsvm	62.72	61.13
bayes.naivebayes	60.22	65.80
baseline	48.37	48.37

The detailed scores of the best system, which is MNLogit based on all of the selected features, are as follows:

Table 55 Detailed result of the best system based on selected vectors

	Precision	Recall	F1	Number
State	**0.70**	0.58	0.63	386
Activity	0.68	**0.90**	**0.78**	634
Transition	0.64	0.20	0.30	200
Average/total	0.68	0.68	0.65	1220

From the scores above, it could be concluded that while activity verbs achieve the best performance based on all of the selected features, transition verbs' performance ranks the lowest.

5.4 Classification based on word embedding vectors

This section sets foot on evaluating our verb senses situation type dataset based on unsupervised features. We firstly introduce the training and extraction of unsupervised features and then report our results.

There are few studies using unsupervised way to implement the prediction of verb situation type as far as we know.

5.4.1 Training of word embedding model

Word embedding vectors which are also called predicted vectors has been introduced in section 5.2 and it shows more powerful predicting ability compared with the counting vectors.

We accept *word2vec* which is re-optimized by Radim Řehůřek and Petr Sojka (2010) to train our context-vector predicting model.

The model is trained on CGW with a size of 1.12 billion words serving as one of the biggest POS-tagged Mandarin corpus.

CGW contains corpora from China Mainland, Taiwan and also Singapore. The balancing of the corpus makes it possible to train a balanced predicting model for context vectors.

The widow size which represents the maximum distance between the current word and predicted word within a sentence is set to 5.

We tested vectors matrix with 100, 300 and 500 dimensions and finally choose the 500-dimension vectors as the verb's context vectors.

172

5.4.2 Classification settings and experiments

The classifiers used for the evaluation based on unsupervised word embedding vectors is the same with ones based on selected features' co-occurrence frequency including MNLogit, SVM and ANN. The configuration is almost the same while ANN's knots on each layers are adjusted to fit the 500 dimensions.

Table 56 Accuracy of different classifiers based on unsupervised features

	MNLogit	SVM	ANN
accuracy	67.54%	**72.05%**	71.48%
Best supvervised		68.44%	
Basline		51.97%	

It could be concluded from the table that SVM achieves the best performance and there is a 4% improvement from the best result achieved based on selected features.

The detailed result of the best system, which is SVM based on unsupervised features, is depicted as follows:

Table 57 Detailed result of the best system based on unsupervised vectors

	Precision	Recall	F1	Number
State	0.69	0.59	0.64	386
Activity	0.72	**0.87**	**0.78**	634
Transition	**0.81**	0.51	0.63	200
Average/total	0.72	0.72	0.71	1220

As a comparison, the best result based on selected features achieved by MNLogit is copied as below:

Table 58 Detailed result of the best system based on selected vectors

	Precision	Recall	F1	Number
State	**0.70**	0.58	0.63	386
Activity	0.68	**0.90**	**0.78**	634
Transition	0.64	0.20	0.30	200
Average/total	0.68	0.68	0.65	1220

It could be concluded from the above two tables that the precision of transition verbs senses' situation type is enormously improved by 20% and the recall of transition improved by 31% while the performance of state verbs and activity verbs change little.

The weak performance of the selected features are mainly caused by the corpus size. Sinica corpus is not big enough to extract co-occurrence frequencies between the verbs and the aspectual markers. We tried CGW to extract selected features' co-occurrence frequency. However, the performance becomes even worse because of the low quality of the POS tagging. A large portion of the POS tagging is wrong, which result in the decreasing of the performance.

Chapter 6

Conclusion

In this chapter, we conclude previous chapters by summarizing the controversial issues which leave space for this study, the significance and limitation in this thesis.

6.1 The existence of lexical aspect: the basic controversial issue

Situation type, as a branch of aspectual studies, has triggered many controversial issues since Vendler (1967) while studies on grammaticalized aspectual markers brought far less debates. Whether a verb can be assigned situation type remains to be the most basic dispute.

Although Vendler (1967)'s quartering of situation types is widely adopted and revised by latter studies, the bearing unit of situation type in his study is not always shared and inherited. Judging from the listed instances of each situation type, both verb and verb phrase are able to be assigned situation types. The existence of double-entry problem of verbs' situation type, however, is conflict with this view. Among Vendler (1967)'s examples, *run per se* is treated as an activity while the *run* in *run a mile* shifts to be an transition or accomplishment to be precise. It is not possible to investigate the nature of lexical aspect providing

175

this shifting view of situation type because theoretically speaking, a verb can be any one of the four situation types.

In order to tackle the ambiguity of verb's situation type, different hypotheses are developed and thus controversy appears. Some studies insist to take situation type away from lexical items while others don't and turn to the independence of lexical aspect from the whole structure for help. Both of them have their reasons.

For studies who accept a shifting view of lexical aspect, they have no choice but to refuse the existence of situation type in lexical level. The situation type of a verb is considered to change back and force according to the attached arguments, adjuncts and other elements in the sentence in these studies. The reason why they hold a shifting view is that verb is not supposed to be the only constituent that contributes to the aspectual meaning of a structure and that other elements also impose influence on the aspectual meaning. In this way, situation type can only be studied in VP or sentence level, which is able to be assigned a set situation type.

The argumentation, however, is problematic because of the following reasons.

Firstly, they fail to target the root causes of the double-entry or shifting problem in Vendler (1967). On one hand, Vendler (1967) drops units including verb (*run*), verb phrase (*run a mile*) and even sentence (*He runs a mile every day.*) into one single level and tries to classify them all at once. On the other hand, Vendler (1967) actually denies the independence of verb's situation type from

the whole structure and impose the whole structure's aspectual meaning on verb. Both of these two facets are inherited by Smith (1991) and other studies who hold the shifting view. Therefore, it is natural that they discard situation type in lexical level.

Secondly, they fail to realize that adopting the contribution of constituents other than verb in a sentence doesn't have to entail a shifting nature of verb's situation type. It should be noticed that almost no study denies the other elements' contribution toward aspectual meaning in a structure, which is also shared and named as the compositionality by us. However, not all studies insist the shifting view. It is the holding that verb's situation type have to be consistent with the whole structure, instead of the compositionality, that leads to the shifting of verb's situation type. Under this perspective, they are actually attach all of the aspectual meaning to verb, which violates the language reality. What's more, it's hard to understand that the *run* in *run to the store* is different from *run* itself.

Thirdly, simply removing situation type from lexical level is not able to solve the ambiguity problem. Smith (1991) and Xu (2015) insist that the smallest unit for situation type is verb phrase. However, VP's situation type is not able to be fixed either since the attachment of other elements to the verbs phase also triggers shifting of situation type. Thus Teng (1985) only accepts sentence as the bearing unit of situation type.

It is the above reasons that result in the lack of study on lexical aspect especially for Mandarin Chinese and spare space for this thesis to make further research.

6.2 Significance of this thesis

In this section, the significance of this thesis in terms of theory, methodology and application contributions is listed.

Aspect, as a comparatively hot topic, attracts both linguistics and computational attention. While linguistics studies try to explain the interaction between aspectual markers and situation type by supplying us with different theoretical models, the lack of statistical validation of their assumption usually leads to unreliable conclusions. Most linguistics studies on Mandarin aspect including Tai (1984),Teng (1985), He (1992), Zhang (1993) and Guo (1993) generalize the interaction rules between situation type and viewpoint aspects simply based on self-created examples. Xiao and McEnery (2004), although is incisive enough to apply corpus to validate their hypotheses, is not able to reason out solid tendency by simply counting examples in corpus.

Computational studies, although applying statistical methodologies, tend to ignore the linguistics explanations and theoretical contributions of aspect. What's more, most computational studies on Mandarin aspect including W. Li et al. (2004), Cao et al. (2006) and Xu (2015) only focus on situation type of sentences and ignore the situation types of verbs or verbs' senses.

178

Focusing on linguistics explanation of the interaction between situation type of verbs and aspectual markers in Mandarin Chinese, this thesis strives to apply statistical validation for linguistics hypotheses and thus filled the gap of previous linguistics studies and computational studies on Mandarin aspect.

6.2.1 Theoretical contributions

This thesis seeks to explore the linguistics rules and explanations on the interaction between verbs' situation type and aspectual markers. The statistical and computational results are treated as evidences to support our linguistics hypotheses.

Firstly, a situation type system is constructed from bottom to up based on the human annotation of verb senses' interaction with aspectual markers including ZHE, LE1, LE2, GUO, ZAI and ZHENGZAI. The hierarchy is generated through hierarchical clustering by measuring the similarity between verbs based on their ability to co-occurrence with aspectual markers.

It should be noticed that the situation type hierarchy is constructed completely based on the verbs' semantic distance in vector space automatically. We have no intention to argue with other studies on why there are only three situation types because the situation type system is not constructed subjectively but based on verbs' semantic distance. There are debates on the number of situation type ranging from two to more than ten and each one of them has their reasons. Our

computational model supports and is consistent with Tai (1984) and Guo (1993) etc. but with more solid argumentation.

Secondly, the concrete relationship between situation types and aspectual markers are measured and quantified through multinominal logistic regression. On one hand, previous linguistics studies on Mandarin aspect are only able to depict qualitative relationship between situation types and aspectual markers based on intuition or limited examples in the corpus. On the other hand, previous computational studies, focusing on the improvement of their model's prediction performance, have no interest to further investigate and explain the relationship between situation types and aspectual markers using statistical or computational methodologies. The absence of statistical validation and quantification in linguistics studies leaves us space for further study. Aiming at better prediction accuracy, most of computational studies are just trying to use linguistics theories to help to improve their performances in different application tasks. They have no intention to make theoretical contributions toward aspect. On the contrary, although stressing the importance of statistical and computational methodologies, this thesis is applying them to help to testify our linguistics hypotheses rather than to realize better prediction models.

The almost zero p-value of LLR test validate that aspectual markers, including ZHE, LE1, LE2, GUO, ZAI and ZHENGZAI, have significant relationship with verbs' situation types.

Following are results of multinominal logistics regression analysis which confirms and quantify each aspectual marker's influence on the differentiation between different situation types.

Table 59 Activity vs. state

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
Const.	0.3917	0.111	3.533	**0.000**	1.479	0.174	0.609
ZHE	-0.1125	0.081	-1.383	0.167	0.894	-0.272	0.047
LE1	0.2989	0.172	1.738	0.082	1.348	-0.038	0.636
LE2	-2.9244	0.617	-4.740	**0.000**	0.054	-4.134	-1.715
GUO	0.7654	0.331	2.313	**0.021**	2.150	0.117	1.414
ZAI	0.0215	0.089	0.241	0.810	1.022	-0.153	0.196
ZHENGZAI	1.1802	0.460	2.566	**0.010**	3.255	0.279	2.082

Table 60 Transition vs. state

	Coef.	Std.	z	P>|z|	Odds ratio	0.025	0.975
Const.	-3.1972	0.769	-4.157	**0.000**	0.041	-4.705	-1.69
ZHE	-20.564	5.634	-3.65	**0.000**	0.000	-31.607	-9.521
LE1	0.8799	0.193	4.567	**0.000**	2.411	0.502	1.257
LE2	0.0301	0.06	0.499	0.618	1.031	-0.088	0.148
GUO	0.3204	0.359	0.893	0.372	1.378	-0.383	1.024
ZAI	-0.8504	0.432	-1.967	**0.049**	0.427	-1.698	-0.003
ZHENGZAI	1.219	0.474	2.571	**0.010**	3.384	0.29	2.148

We also support that there should be two LEs in Mandarin Chinese by supporting statistical evidence.

Table 61 TTR of co-occurred verbs of LE1 and LE2

	Type frequency	Token frequency	TTR
LE1	504	10409	0.048
LE2	121	697	0.174

Table 62 Statistics of LE1 and LE2's co-occurrence frequency with all of the verbs

	LE1	LE2
count	1220	1220
mean	8.531967	0.571311
std	36.55057	4.273158
min	0	0
25%	0	0
50%	0	0
75%	4	0
max	884	109

Table 63 Kruskal-Wallis H-test between LE1 and LE2

H	$p - $ value
339.4425	0.0000

The almost zero p-value denies that two LEs have same distribution and thus they are supposed to be separated from each other.

6.2.2 Methodological contributions

The situation type system constructed based on the intuition annotations of linguists is validated and confirmed in Sinica corpus which is with high tagging

quality. The validation result is only one of the contributions that I intend to show, the process of validation through statistics is another facet that I focus.

Intuition based methodologies have dominated linguistics study almost since it's occurring especially in the era that language data is hardly observable. No one can deny the contribution of the intuition-based linguistics studies which ground the foundation for latter studies and supply sharp views for new language phenomenon and hypotheses.

The appearance of corpus enables us to observe language data in a more efficient way and provides possibility to validate the hypotheses in a statistical way. However, counting examples instead of statistics is counted as the main way to validate linguistics hypotheses. It is quite unreliable by simply counting supporting evidences and contradicting examples.

In this thesis, our hypotheses are mostly supported by statistical tests and computational tests which actually illustrate the way to do linguistics studies in a more reliable way compared with counting examples.

6.2.3 Computational contributions

Computational contributions of this thesis are illustrated below in terms of language resource and the prediction of lexical aspect.

Firstly, a dataset featured with verbs' senses and their situation type is constructed with high reliability. On one hand, the situation type system is constructed based on the manual annotation of verbs' ability to co-occur with

different aspectual markers, which is cross-validated among three different raters. On the other hand, the dataset is evaluated on SINICA corpus and CGW by different classifiers, which proves its validity.

Secondly, a comparatively high performance is achieved for the prediction of situation type in lexical level. Trained and evaluated on our dataset, our model achieved an accuracy of 72.05% using word embedding vectors as features. The detailed result is as follows.

Table 64 Detailed result of the best system based on unsupervised vectors

	Precision	Recall	F1	Number
State	0.69	0.59	0.64	386
Activity	0.72	**0.87**	**0.78**	634
Transition	**0.81**	0.51	0.63	200
Average/total	0.72	0.72	0.71	1220

6.2.4 Implications of this thesis

The study of aspect in Mandarin Chinese in this thesis is able to facilitate the understanding of English grammar for English learners. Aspect is not a familiar terminology for Chinese students. They are usually taught to merge tense and aspect together when they are leaning English grammar. However, the difference between the so-called "progressive tense" and "present tense" is clear enough

184

that triggers puzzle among the students. According to my teaching experience in schools, English learners tend to have better idea of English grammar system when they are taught the concept of aspect and shown the comparison between Chinese aspect and English tense, which are developed in this thesis and also previous studies.

For Chinese learners, having a better idea of how to express aspectual information in a sentence is of vital importance for them to acquire Chinese.

6.3 **Some limitations and future work of this thesis**

There are, however, some limitations of this study.

Firstly, this thesis only concerns about lexical aspect. We insist a recursive and compositional view of situation type and thus all of predicates including verbs (and adjectives), verb phrases and sentences are supposed to have their situation types. The smallest unit that can be assigned situation type is verb or a sense of a verb, to be precise. When arguments and other elements are attached to a verb, they together form a phrase which has its own situation type which is decided by not only the constituting verb but also other constituents. Sentences have the same nature. Situation types in other levels are not included because of several reasons. On one hand, other aspectual studies such as Xu (2015) have done a lot of work on Mandarin aspect in sentential level. On the other hand, the annotation of situation type is a time-consuming task which limits our study within part of

them. The annotation and validation of lexical aspect along takes us about half a year which leave no time for further study on situation type in other levels.

Secondly, compositionality is another topic which is rarely discussed in this thesis. The basic hypothesis for compositionality is that verb is not the only element that decide the situation type a structure as a whole. The situation type of a structure is a result of the composition of verb and other elements such as the arguments and adjuncts. Although the nature of compositionality is shared in this thesis, the interaction of verb and its arguments and adjuncts and how they generate a bigger structure which has its situation type are not touched. In other words, the syntax of situation type is not generalized in this thesis. What's more, the encoding of compositionality in computational model is not tested either.

Thirdly, semantic nuances between different aspectual markers are neglected in this thesis. Although supplying solid validation of interactions between aspectual markers and situation types, the semantic difference between different aspectual markers are, however, not included in this thesis. One of the most attractive phenomenon is the difference between ZHE, ZAI and ZHENGZAI. Syntactically speaking, ZHE and ZAI are almost the same in terms of their ability to co-occur with different situation types while ZHENGZAI is completely different from ZHE and ZAI. It then follows that what on earth the semantic nuances between ZHENGZAI and ZAI is. What's more, there must be some semantic or pragmatic difference between ZHE and ZAI. Otherwise, there is no

186

need for Mandarin to grammaticalize two aspectual markers with same function and meaning.

I intend to focus on these three issues in the future study.

For the firt two questions, with the same goal that investigates the interaction between the situation type of a structure and different constituents within the structure, compositionality and the generation of situation type of a verb phrase or sentence are actually two sides of one problem. The relationship between the situation type of a structure and its constituents is going to be tackled in latter studies. The annotation of situation type of verb phrase and sentence, however, should be completed before the validation of the relationship.

For the third question, a detailed exploration of the semantic nuances between different aspectual markers are supposed to be implemented in the future study.

I also intend to build the classification model by including other aspectual mechanisms such as verb reduplication and resultative verb compound to see whether there is adjustments of situation type system or not. If yes, how to explain the new situation type system.

On the other hand, the verbs situation type will also be used to do other temporal relation reasoning application to see its contribution towards other tasks.

Appendix: POS tag in Sinica corpus

Simple POS	Abbreviation	Interpretation
Adjective	A	Non-predicative adjective
Conjunction	Caa	Conjunctive conjunction (e.g.和、跟)
Conjunction	Cab	Conjunction (e.g. 等等)
Conjunction	Cba	Conjunction (e.g. 的话)
Conjunction	Cbb	Correlative Conjunction
Adverb	D	Adverb
Adverb	Da	Quantitative adverb
的-Construction	DE	的, 之, 得, 地
Adverb	Dfa	Pre-verbal adverb of degree
Adverb	Dfb	Post-verbal adverb of degree
Adverb	Di	Aspectual adverb
Adverb	Dk	Sentential adverb
Foreign word	FW	Foreign word
Interjection	I	Interjection
Noun	Na	Common noun
Noun	Nb	Proper noun
Noun	Nc	Place noun
Noun	Ncd	Localizer
Noun	Nd	Time noun
Determiner	Nep	Demonstrative determinatives
Determiner	Neqa	Quantitative determinatives
Determiner	Neqb	Post-quantitative determinatives

Determiner	Nes	Specific determinatives
Determiner	Neu	Numeral determinatives
Measure	Nf	Measure
Postposition	Ng	Postposition
Noun	Nh	Pronoun
Preposition	P	Preposition
Verb	SHI	是
Particle	T	Particle
Verb	VA	Active intransitive verb
Verb	VAC	Active causative verb
Verb	VB	Active pseudo-transitive verb
Verb	VC	Active transitive verb
Verb	VCL	Active verb with a locative object
Verb	VD	Ditransitive verb
Verb	VE	Active verb with a sentential object
Verb	VF	Active verb with a verbal object
Verb	VG	Classificatory verb
Verb	VH	Stative intransitive verb
Verb	VHC	Stative causative verb

Verb	VI	Stative pseudo-transitive verb
Verb	VJ	Stative transitive verb
Verb	VK	Stative verb with a sentential object
Verb	VL	Stative verb with a verbal object
Verb	V_2	有

References

Baroni, M., Dinu, G., & Kruszewski, G. (2014). *Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors.* Paper presented at the ACL (1).

Brinton, L. J. (1988). *The development of English aspectual systems: Aspectualizers and post-verbal particles*: Cambridge University Press.

Bybee, J. L., Perkins, R. D., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world* (Vol. 196): University of Chicago Press Chicago.

Cao, D., Li, W., Yuan, C., & Wong, K.-F. (2006). Automatic chinese aspectual classification using linguistic indicators. *International Journal of Information Technology, 12*(4), 99-109.

Chao, Y.-R. (1968). *A Grammar of Spoken Chinese*. Berkeley: The University of California Press.

Chen, K.-J., Huang, C.-R., Chang, L.-P., & Hsu, H.-L. (1996). Sinica corpus: Design methodology for balanced corpora. *Language, 167*, 176.

Chen, P. (1988). Lun Xiandai Hanyu Shijian Xitong de Sanyuan Jiegou. *Chinese Language, 6*, 22.

Christensen, M. B. (1994). *Variation in spoken and written Mandarin narrative discourse.* The Ohio State University,

Comrie, B. (1979). Aspect: An Introduction to the Study of Verbal Aspect and Related Problems. *Modern Language Review, 74*(1), 1964-1968.

Costa, F., & Branco, A. (2012). *Aspectual type and temporal relation classification.* Paper presented at the Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics.

Dai, Y. (1997). *Xiandai hanyu shiti xitong yanjiu* (Vol. 5): Zhejiang education press.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American society for information science, 41*(6), 391.

Dowty, D. R. (1979). *Word Meaning and Montague Grammar*: Springer Netherlands.

Eddington, D. (2009). Linguistics and the scientific method. *Southwest Journal of Linguistics, 27*(3), 229–230.

Eysenck, M. W., & Keane, M. T. (2000). *Cognitive psychology: A student's handbook*: Taylor & Francis.

Falk, I., & Martin, F. (2016). *Automatic Identification of Aspectual Classes across Verbal Readings.* Paper presented at the * Sem 2016 THE FIFTH JOINT CONFERENCE ON LEXICAL AND COMPUTATIONAL SEMANTICS.

Friedrich, P. (1974). On aspect theory and Homeric aspect. *International Journal of American Linguistics, 40*(4, Part 2), S1-S44.

Guo, R. (1993). Hanyu dongci de guocheng jiegou. *Chinese Language, 6*, 410--419.

Guo, R. (1997). Guocheng he feiguocheng-hanyu weicixing chengfen de liangzhong waizai shijian leixing. *Chinese Language*(3), 162-175.

Harris, Z. S. (1954). Distributional structure. *Word, 10*(2-3), 146--162.

He, B. (1992). *Situation types and aspectual classes of verbs in Mandarin Chinese*.

Himmelmann, N. P. (2004). Lexicalization and grammaticization: Opposite or orthogonal. *What makes grammaticalization*, 21-42.

Hou, R., Huang, C.-R., Do, H. S., & Liu, H. (2017). A Study on Correlation between Chinese Sentence and Constituting Clauses Based on the Menzerath-Altmann Law. *Journal of Quantitative Linguistics*, 1-17.

Hou, R., Huang, C.-R., & Liu, H. (2017). A study on Chinese register characteristics based on regression analysis and text clustering. *Corpus Linguistics and Linguistic Theory*.

Hou, R., & Jiang, M. (2014). Analysis on Chinese quantitative stylistic features based on text mining. *Digital Scholarship in the Humanities*, fqu067.

Hou, R., Yang, J., & Jiang, M. (2014). A Study on Chinese Quantitative Stylistic Features and Relation Among Different Styles Based on Text Clustering. *Journal of Quantitative Linguistics, 21*(3), 246-280.

Huang, C.-R., Ahrens, K., Chang, L.-l., Chen, K.-j., Liu, M.-c., & Tsai, M.-c. (2000). The Module-Attribute Representation of Verbal Semantics.

Huang, C. R., & Liu, H. (2017). Corpus-based automatic extraction and analysis of Mandarin alphabetic words. *Journal of Yunnan Normal University (Humanities and Social Sciences), 3*, 10-21.

Huang, L. M.-j. (1987). *Aspect: A general system and its manifestation in Mandarin Chinese.* Rice University,

Kazantseva, A., & Szpakowicz, S. (2010). Summarizing short stories. *Computational Linguistics, 36*(1), 71-109.

Klein, W. (1994). *Time in language*: Psychology Press.

Klein, W., Li, P., & Hendriks, H. (2000). Aspect and assertion in Mandarin Chinese. *Natural language & linguistic theory, 18*(4), 723-770.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association, 47*(260), 583-621.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review, 104*(2), 211.

Langacker, R. W. (1987). *Foundations of cognitive grammar*. Stanford, California: Stanford University Press.

Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*: Univ of California Press.

Li, W., Wong, K.-F., Cao, G., & Yuan, C. (2004). *Applying machine learning to Chinese temporal relation resolution.* Paper presented at the Proceedings of the 42nd annual meeting on association for computational linguistics.

Lin, D., & Pantel, P. (2001). *DIRT@ SBT@ discovery of inference rules from text.* Paper presented at the Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining.

Liu, H., Karl, N., Enrico, S., & Huang, C.-R. (2016). *EVALution-MAN: A Chinese Dataset for the Training and Evaluation of DSMs*. Paper presented at the Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), Paris, France.

Liu, H., & Lin, Y. (2017). The methodology and trend of linguistics study in the era of big dara. *Journal of Xinjiang Normal University(Edition of Philosophy and Social Sciences)*(1).

Liu, Q., & Li, S. (2002). Word Similarity Computing Based on How-net. *Computational Linguistics and Chinese Language Processing, 7*(2), 59--76.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28*(2), 203-208.

Lv, S., & Ding, S. (1978). *Contemporary Chinese Dictionary*: The Commertial Press.

McKinney, W. (2012). *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*: " O'Reilly Media, Inc.".

Meng, C. (1987). *Chinese verb usage dictionary* (Vol. 612): The Commercial Press.

Meyer, T., Grisot, C., & Popescu-Belis, A. (2013). *Detecting narrativity to improve English to French translation of simple past verbs.* Paper presented at the Proceedings of the 1st DiscoMT Workshop at ACL 2013 (51st Annual Meeting of the Association for Computational Linguistics).

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mikolov, T., Le, Q. V., & Sutskever, I. (2013). Exploiting similarities among languages for machine translation. *arXiv preprint arXiv:1309.4168*.

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

Mikolov, T., Yih, W.-t., & Zweig, G. (2013). *Linguistic Regularities in Continuous Space Word Representations.* Paper presented at the HLT-NAACL.

Miller, G. A., Leacock, C., Tengi, R., & Bunker, R. T. (1993). *A semantic concordance.* Paper presented at the Proceedings of the workshop on Human Language Technology.

Mitchell, J., & Lapata, M. (2008). *Vector-based Models of Semantic Composition.* Paper presented at the ACL.

Moens, M. (1987). Tense, aspect and temporal reference.

Mourelatos, A. P. (1978). Events, processes, and states. *Linguistics and Philosophy, 2*(3), 415-434.

Pampel, F. C. (2000). *Logistic regression: A primer* (Vol. 132): Sage Publications.

Peng, R. (2017). Host-Class Changes in Grammaticalization: Modes and Roles. *Linguistic Sciences, 16*(2), 142-157.

Pustejovsky, J. (1991). The syntax of event structure. *Cognition, 41*(1), 47--81.

Pustejovsky, J., Castano, J. M., Ingria, R., Sauri, R., Gaizauskas, R. J., Setzer, A., . . . Radev, D. R. (2003). TimeML: Robust specification of event and temporal expressions in text. *New directions in question answering, 3*, 28-34.

Radim Řehůřek, & Petr Sojka. (2010). *Software Framework for Topic Modelling with Large Corpora.*

Rapp, R. (2003). *Word sense discovery based on sense descriptor dissimilarity.* Paper presented at the Proceedings of the ninth machine translation summit.

Rosch, E. (1999). Principles of categorization. *Concepts: core readings, 189.*

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive psychology, 8*(3), 382-439.

Rosch, E. H. (1973). Natural categories. *Cognitive psychology, 4*(3), 328-350.

Salton, G. (1971). The SMART retrieval system—experiments in automatic document processing.

Salton, G., Wong, A., & Yang, C.-S. (1975). A vector space model for automatic indexing. *Communications of the ACM, 18*(11), 613-620.

Saurí, R., & Pustejovsky, J. (2012). Are you sure that this happened? assessing the factuality degree of events in text. *Computational Linguistics, 38*(2), 261-299.

Shaphiro, S., & Wilk, M. (1965). An analysis of variance test for normality. *Biometrika, 52*(3), 591-611.

Siegel, E. V. (1998). *Linguistic indicators for language understanding: using machine learning methods to combine corpus-based indicators for aspectual classification of clauses.* (Ph.D), Columbia University,

Siegel, E. V., & Mckeownt, K. R. (2002). Learning Methods to Combine Linguistic.

Smith, C. S. (1990). Event types in Mandarin. *Linguistics, 28*(2), 309-336.

Smith, C. S. (1991). *The parameter of aspect* (Vol. 43): Springer Science & Business Media.

Smith, C. S. (1994). Aspectual viewpoint and situation type in Mandarian Chinese. *Journal of East Asian Linguistics, 3*(2), 107-146.

Smith, C. S., & Erbaugh, M. S. (2005). Temporal interpretation in Mandarin Chinese. *Linguistics, 43*(4), 713-756.

Sokal, R. R., & Rohlf, F. J. (1962). The comparison of dendrograms by objective methods. *Taxon, 11*(2), 33-40.

Stanovich, K. E. (2013). How To Think Straight About Psychology: Pearson New International Edition, 10/E.

Tai, J. (1984). Verbs and times in Chinese: Vendler's four categories. *CLS 20/2. Papers from the parasession on lexical semantics, Chicago, IL: Chicago Linguistic Society. (289-296) Tenny, Carol 1992. « The Aspectual Interface Hypothesis », in: I. Sag & A, 92*(5), 75–78.

Teng, S. (1985). The temporal structure of Verbs in Chinese. *Language Teaching and Linguistic Studies*(4), 7-17.

Turney, P. D. (2006). Similarity of semantic relations. *Computational Linguistics, 32*(3), 379-416.

Turney, P. D. (2008). The latent relation mapping engine: Algorithm and experiments. *Journal of artificial intelligence research, 33*, 615-655.

Turney, P. D., & Littman, M. L. (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS), 21*(4), 315-346.

Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research, 37*(1), 141-188.

Vendler, Z. (1967). *Linguistics and philosophy*: Cornell University Press.

Verhagen, M., Gaizauskas, R., Schilder, F., Hepple, M., Katz, G., & Pustejovsky, J. (2007). *Semeval-2007 task 15: Tempeval temporal relation identification.* Paper presented at the Proceedings of the 4th International Workshop on Semantic Evaluations.

Verhagen, M., & Pustejovsky, J. (2008). *Temporal processing with the TARSQI toolkit.* Paper presented at the 22nd International Conference on on Computational Linguistics: Demonstration Papers.

Verhagen, M., Sauri, R., Caselli, T., & Pustejovsky, J. (2010). *SemEval-2010 task 13: TempEval-2.* Paper presented at the Proceedings of the 5th international workshop on semantic evaluation.

Verkuyl, H. J. (1972). *On the compositional nature of the aspects* (Vol. 15). Dordrecht: Reidel.

Wittgenstein, L. (2010). *Philosophical investigations*: John Wiley & Sons.

Xiao, R., & McEnery, T. (2004). *Aspect in Mandarin Chinese: A corpus-based study* (Vol. 73): John Benjamins Publishing.

Xu, H. (2015). *The Chinese aspectual system.* The Hong Kong Polytechnic University,

Yu, S., Zhu, X., Wang, H., & Zhang, Y. (1996). Specification for contemporary Chinese syntactic information dictionary. *Journal of Chinese Information Processing, 10*(2), 1-22.

Zarcone, A., & Lenci, A. (2008). *Computational Models for Event Type Classification in Context.* Paper presented at the LREC.

Zhang, L. (1993). *A contrastive study of aspectuality in German, English, and Chinese*: University of California, Berkeley.