Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.
Cleft Constructions in Chinese: An Information Structure Analysis

Danyang Zheng

Abstract

This thesis is the first systematic and comprehensive study of cleft constructions in Chinese. By the cleft construction in Chinese is meant the syntactic structure mainly holding the form of “shi...(de)”, and realizing the function of contrastive focus by putting the focus in a syntactically more prominent position (i.e. the following position of the focus marker “shi”), and by “cleaving” the canonical proposition into two segments (i.e. presupposition and focus) from the perspective of information structure.

In this thesis, the criteria of cleft constructions are provided from the prospectives of semantics and information structure aiming to distinguish Chinese cleft constructions from the non-qualifying sentences with similar surface structures. And then the typological classification of Chinese “shi...de” sentences is presented. Considering the information structure of cleft constructions in Chinese, their properties are explored by providing answers to the following questions: firstly, how are the components (i.e. focus and topic) of the information structure delivered through their mental representations as well as linguistic forms; secondly, what are the types of cleft constructions in Chinese in accordance with the properties of their information structure; and thirdly how do the components of the information structure (i.e. focus, topic, presupposition) distribute with each other?

In order to get some more authentic and objective results, the data of the present study is selected from the online corpus designed by the Center for Chinese Linguistics of Peking University (abbreviated as CCL hereafter).

The theoretical foundation of the present study is Lambrecht’s (1994) theories of information structure (i.e. the types of the focus structure, the Topic Acceptability Scale, the model of identifiability etc.). In addition, other important concepts are also
borrowed in the present study, for example, “exhaustivity” (or “exhaustive” in the terminology of Kiss (1998)) proposed by Buring & Kriz (2013) is considered to be a very important semantic criterion of cleft constructions in Chinese. Chen’s (2004) study of the relationship between identifiability and definiteness of Chinese is also an important theoretical background of the present study.
Contents

Abstract .................................................................................................................................................. ii
Acknowledgements .............................................................................................................................. ix
List of Tables ......................................................................................................................................... x
List of Figures ...................................................................................................................................... xii
List of Abbreviations ......................................................................................................................... xiv

Chapter 1 ............................................................................................................................................... 1
Introduction ........................................................................................................................................... 1
  1.1 The topic of the thesis ............................................................................................................. 1
  1.2 Why study cleft constructions in Chinese? ........................................................................... 11
  1.3 Aims of the study .................................................................................................................. 14
  1.4 Research questions ................................................................................................................ 16
  1.5 Research hypotheses ............................................................................................................. 17
  1.6 Organization .......................................................................................................................... 18

Chapter 2 ............................................................................................................................................. 20
Literature Review ............................................................................................................................... 20
  2.1 Introduction ........................................................................................................................... 20
  2.2 Cleft constructions in languages other than Chinese ............................................................ 21
    2.2.1 Cleft constructions in English .......................................................................................... 22
    2.2.1.1 The definitions of cleft constructions ........................................................................ 22
    2.2.1.2 Seminal studies ......................................................................................................... 24
    2.2.1.3 The classifications of cleft constructions and syntactic structures of cleft constructions ............................................................................................................................. 27
    2.2.1.4 The semantic features of cleft constructions ............................................................. 31
    2.2.1.5 Functions of cleft constructions ................................................................................ 34
    2.2.1.6 The information structure of cleft constructions in English ...................................... 35
    2.2.2 Cleft constructions in other languages .......................................................................... 37
    2.2.2.1 Cleft constructions in French .................................................................................... 38
    2.2.2.2 Cleft constructions in Italian ..................................................................................... 41
    2.2.2.3 Cleft constructions in German .................................................................................. 44
    2.2.2.4 Cleft constructions in Japanese ................................................................................. 46
  2.3 Cleft constructions in Chinese .............................................................................................. 49
    2.3.1 The seminal study ............................................................................................................. 50
    2.3.2 Forms and classifications of cleft constructions in Chinese ............................................ 52
      2.3.2.1 Forms of cleft constructions in Chinese ................................................................. 52
5.6 Summary ............................................................................................................................. 195
  5.6.1 The prototypical structure of the cleft construction and its variants ......................... 195
  5.6.2 The distribution of cleft constructions in formal and informal texts ....................... 195
  5.6.3 The classification of “shi...de” constructions ............................................................ 195
  5.6.4 The word classes of “shi” and “de” and the omission of them in “shi...de” constructions 196

Chapter 6 ........................................................................................................................................... 200

Analysis of the Information Structure of Chinese Cleft Constructions ........................................ 200

  6.1 Introduction ......................................................................................................................... 200
  6.2 The focus domain and types of focus structures ................................................................. 202
     6.2.1 Argument-focus cleft constructions ............................................................................. 203
     6.2.2 Adjunct-focus cleft constructions ................................................................................. 212
  6.3 Syntactic categories of focus constituents ........................................................................ 215
     6.3.1 “Shi...de” cleft constructions ...................................................................................... 216
     6.3.2 Bare “shi” cleft constructions ...................................................................................... 216
     6.3.3 Bare “de” cleft constructions ...................................................................................... 217
  6.4 Thematic relations of focus constituents of adjunct-focus cleft constructions .............. 218
     6.4.1 Syntactic categories of the thematic relations .............................................................. 221
     6.4.2 The thematic relations represented by different syntactic categories ...................... 223
  6.5 The identifiability of the “focus” and “topic” constituent of the cleft construction ...... 226
     6.5.1 The distribution of focus and topic constituents in the three identifiable statuses ...... 226
         6.5.1.1 The identifiability of focus/topic constituents in argument-focus cleft constructions 230
         6.5.1.2 The identifiability of focus/topic constituents in adjunct-cleft constructions ...... 235
     6.5.2 The grammatical expressions and their identifiability .................................................. 236
         6.5.2.1 The word classes and the identifiability of topic and focus ..................................... 236
         6.5.2.2 The degrees of identifiability of different word classes .......................................... 238
     6.5.3 The identifiability of the thematic relations ................................................................. 244
         6.5.3.1 The identifiability of the thematic relations of focus constituents of Chinese cleft constructions .................................................................................. 244
         6.5.3.2 The identifiability of the thematic relations of topic constituents of Chinese cleft constructions .................................................................................. 246
  6.6 The configuration of the Chinese cleft construction from the perspectives of the information and syntactic structure ................................................................................................................................. 247
  6.7 The markedness of Chinese cleft constructions from the perspective of focus typology ... 259
  6.8 Summary ............................................................................................................................. 261
Chapter 7 ........................................................................................................................................... 264
Conclusion ......................................................................................................................................... 264

7.1 The structures of cleft constructions in Chinese ................................................................. 264
7.2 The configuration of the information structure ................................................................. 266
7.3 The thematic relations of focus/topic constituents .......................................................... 267
7.4 The identifiability of focus/topic constituents ................................................................ 268
7.5 Significance ....................................................................................................................... 268
7.6 Future directions of the study .......................................................................................... 270

Appendices ......................................................................................................................................... 271
References .......................................................................................................................................... 292
Acknowledgements

My deepest gratitude goes to my main supervisor, Professor Yan Huang, for his consistent encouragement and continual support throughout my whole study. I benefited a lot from the academic discussions with him and from attending his classes. His passion for academic study and his attitudes towards it impressed me deeply. He devoted a large amount of time to reading my drafts and gave me lots of comments that really enlightened me in the whole procedure of writing. Without him, this dissertation would not have been possible.

I am also indebted to my co-supervisor, Dr. Helen Charters, for her efficient comments, feedback on my drafts, supported with insightful ideas. Her brilliant opinions inspired me to consider problems in different ways.

I would also like to thank the staff and my fellow PhD students in the School of Cultures, Languages and Linguistics for their support and suggestions.

I sincerely appreciate the financial support provided by Chinese Scholarship Council for my study in New Zealand.

Finally, I express my thanks to my parents, for their unconditional love, support and understanding throughout the years.
List of Tables

Table 2-1 Referential Givenness Categories (Hedberg & Fadden, 2007, p. 5) .......................36
Table 2-2 The classifications of cleft constructions and the distribution of “new” and “given”
information (Calude, 2007) ......................................................................................................36
Table 4-1 The distribution of sentence length by Zong et al. (1999) ........................................112
Table 4-2 Distributions of Chinese sentences in different genres by Li et al. (2014) ............112
Table 4-3 Information structure of sentence (4-5) ..................................................................123
Table 4-4 Information structure of sentence (4-6) ..................................................................124
Table 4-5 Information structure of sentence (4-7) .................................................................125
Table 4-6 Information structure of sentence (4-8) .................................................................127
Table 4-7 Information structure of sentence (4-11) ...............................................................128
Table 4-8 Types of focus structures and their function of contrastivity ................................143
Table 5-1 Classifications of copular sentences of the present study .....................................156
Table 5-2 Classification of “shi...de” sentences in the present study ..................................159
Table 6-1 Configuration of the information structure of the example sentence (6-9) ........208
Table 6-2 Configuration of the information structure of the example sentence (6-12) ....210
Table 6-3 Configuration of the information structure of the example sentence (6-14) .......213
Table 6-4 The identifiability of the focus constituent in argument-focus cleft constructions.
................................................................................................................................................233
Table 6-5 The identifiability of the topic constituent in argument-focus cleft constructions
................................................................................................................................................233
Table 6-6 Demonstration of examples from the perspectives of information, syntactic
structures, word classes, identifiability and thematic relations (1) ......................................248
Table 6-7 Demonstration of examples from the perspectives of information, syntactic
structures, word classes, identifiability and thematic relations (2) .......................................249
Table 6-8. Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (3) ................................. 250
Table 6-9 Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic (4) ........................................ 252
Table 6-10. Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (5) ................................. 254
Table 6-11. Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (6) ................................. 255
Table 6-12. Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (7) ................................. 256
Table 6-13. Demonstration of examples from the perspectives of information, syntactic structure, word classes, identifiability and thematic relations (8) ................................. 257
Table 6-14. Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (9) ................................. 258
List of Figures

Figure 2-1. Structure of the English cleft construction (Prince, 1978) .........................28
Figure 2-2. Structure of the pseudo-cleft construction (Prince, 1978) ...........................28
Figure 3-1. Model of Identifiablity (Lambrecht 1994, p. 109). .....................................97
Figure 3-2. The Topic Acceptability Scale (Lambrecht 1994, p.165). ............................101
Figure 4-1. Some of the results of “shi $15 de” constructions in the primary search. ....114
Figure 4-2. Some of the results of “shi $ 15 ” constructions in the primary search. ....115
Figure 4-3. Some of the results of “$15 de” constructions in the primary search. ........116
Figure 5-1. The relationship between “shi...de” sentences, copular sentences, cleft sentences and emphatic sentences ..................................................................................162
Figure 5-2. The distributions of cleft constructions of different types in the two corpora. ...163
Figure 6-1. Types of focus structure of cleft constructions in Chinese ..........................202
Figure 6-2. Syntactic categories of focus constituents of Chinese cleft constructions .......215
Figure 6-3. Syntactic categories of focus constituents in “shi...de” cleft constructions......216
Figure 6-4. Syntactic categories of focus constituents of bare “shi” cleft constructions.....217
Figure 6-5. Syntactic categories of focus constituents of bare “de” cleft constructions.....218
Figure 6-6. Thematic relations of focus constituents .......................................................219
Figure 6-7. Syntactic categories of focus constituents whose thematic relations are “time”, “means” and “manners” .........................................................................................222
Figure 6-8. The thematic relations designated by prepositional phrases. .......................223
Figure 6-9. Model of identifiablity applied in the present study. .....................................228
Figure 6-10. Identiifiability of the focus and topic constituents .......................................229
Figure 6-11. Word classes vs. the focus/topic nominal constituents. .............................237
Figure 6-12. Word classes and identifiabilities ...............................................................238
Figure 6-13. The distributions of proper nouns having “inactive”, “active” and “accessible”
statuses. .................................................................................................................................. 240

Figure 6-14. The distributions of personal pronouns having “inactive”, “active” and
“accessible” statuses. ............................................................................................................. 240

Figure 6-15. The distributions of demonstrative nouns having “inactive”, “active” and
“accessible” statuses. ............................................................................................................. 241

Figure 6-16. The distributions of bare nouns having “inactive”, “active”, “accessible”
statuses. .................................................................................................................................. 241

Figure 6-17. The distributions of modified nouns having “inactive”, “active” and “accessible”
statuses. .................................................................................................................................. 242

Figure 6-18. Degrees of identifiability of referents represented by different grammatical
expressions. .................................................................................................................................. 243

Figure 6-19. Thematic relations of focus constituents and the statuses of "unidentifiable",
“inactive”, “accessible” and “active”..................................................................................... 245

Figure 6-20. Thematic relations of the topic constituent and the status of “active” and
“accessible”.................................................................................................................................. 247
List of Abbreviations

Shi$_{cop}$—copular verb “shi”

Shi$_{aux}$—auxiliary verb “shi”

De$_{poss}$—possessive marker “de”

DE$_{ptcl}$—particle “de”

De$_{nomi}$—nominalizer “de”

De$_{rel}$—relative pronoun “de”

De$_{modi}$—modifier marker “de”

NP—nominal phrase

VP—verb phrase

AP—adjective phrase

AdvP—adverbial phrase

PP—prepositional phrase

OBJ—object

SUBJ—subject

N—noun

V—verb

SC—small clause
ZHE—imperfective aspectual marker

NE—mood particle
Chapter 1

Introduction

This thesis is concerned with cleft constructions in Chinese. The aim of the study is to explore the Chinese cleft construction in a way that is seldom examined (i.e., information structure), and to establish a new and more comprehensive framework for understanding Chinese cleft constructions. The tasks of the study are firstly to ascertain the syntactic forms of Chinese cleft constructions based on the observation and analysis of empirical data, and secondly to investigate the information structure (i.e., the focus and topic) of cleft constructions in Chinese from the identifiability of the components and the thematic relations expressed by them. The study is a qualitative one, with its main part being the analysis of data collected from the online corpus established by the Center for Chinese Linguistics of Peking University. The principal theoretical framework I will use is that of Lambrecht’s theories of information structure, based on which alterations are made to adjust to the characteristics of Chinese.

1.1 The topic of the thesis

Cleft constructions are widely-used syntactic constructions of contrastive focus that serve to give prominence to the constituent (considered as the “focus”) the speaker aims to highlight in conversation or discourse. Normally, cleft constructions are used to indicate contrast so as to highlight the “focused” constituent. To give prominence to the “focused” constituent, the sentence or proposition is divided into two parts from an informative perspective if the syntactic partition is not significantly noticeable in some languages.
The concept of the cleft construction was put forward and named with respect to the study of English (Jespersen, 1927). Since then, the concept has been gradually accepted and applied to many languages (cf. Huber, 2002; Kizu, 2005; Lambrecht, 1988; Lee, 2005; Paul & Whitman, 2008). How does the same concept adjust to different languages despite the distinctions between languages (especially the syntactic differences)? Considering the similar function of cleft constructions in different languages, we assume that there must be some common criteria that are not influenced by language differences that we can resort to when judging cleft constructions.

From the perspectives of both function and syntax, we can define cleft constructions in English as sentence structures that put a focusing influence on the proposition by syntactically separating it into two parts.

English is very rich in cleft constructions. The main form of the English cleft sentence¹ is “It BE ... that/WH-”, with the constituent after BE as the “focus” of the sentence, and the “that-” or “WH-clause” expressing the old-information known by both the speaker and listener called “presupposition” in pragmatics (Collins, 1994). According to Collins’ (1994) corpus-driven study, besides cleft constructions, there are two other main variants in English. They are pseudo-cleft constructions and reversed pseudo-cleft constructions (which are represented by “what…BE (that/WH-)” and “(that/WH-) BE what…” respectively), with the focused constituent as the “that” or “wh-” clause.

¹English cleft constructions are classified into different types. The most accepted classification is given by Collins (1994)—cleft constructions (equal to IT-cleft); pseudo-cleft constructions, which are subdivided into wh-clefs and th-clefs; and reversed pseudo-cleft constructions. In this study, the term “cleft constructions” is used to refer to the first type (i.e, IT-cleft constructions) in all languages (i.e, English, Chinese, and other mentioned languages). In order to avoid any misunderstanding caused by the dummy subject “it”, I avoid the term “IT-cleft” construction.
The coining of the term “cleft constructions” in English made scholars of other languages to examine whether cleft constructions existed in their own languages or not. Based on native speakers’ intuition, English cleft constructions are translated into corresponding structures in other languages that are syntactically similar to their English counterparts. For example, the structures “Er sein..., de...” in German and “C'est...qui...” in French are syntactically like the English cleft construction. Intuitively, native speakers will translate English clefts into these two forms in German and French respectively. But doubts are proposed by scholars considering the relationship between function and the syntactic form of the translated version. Functionally and semantically speaking, English cleft constructions are structures aiming to place “focus” syntactically in a highlighted position of a sentence beyond the “presupposed” proposition. In English, there is more than one way to show focus function, which is counted as a noncanonical function within a sentence domain; for example, noncanonical meaning can be delivered by methods of information-packaging such as preposing, clefts, inversions and so on. (Prince, 1981). So, from the perspective of information delivery, the cleft construction is a marked syntactic way to show the focus function of a proposition. Taking the functional aspect into consideration, scholars of different languages began to consider whether the translated counterparts in their languages should be counted as cleft constructions in nature; in other words, whether or not the function of the translated structures was the same as that of the English cleft construction. Positive and negative opinions merged: to take one example, Lambrecht (1994) proposes that French cleft constructions are legitimate and nearly equivalent to those of English; however, Clech-Daebon et al. (1999) take the opposite idea—they state that there are in fact no real cleft sentences per se in French.
There is also both consensus and dispute in the literature of Chinese cleft constructions. To start with the consensus, generally speaking, Chinese cleft constructions are said to include “shi” (BE) introducing an NP argument, NP/AP/AdvP/PP\(^2\) predicate in a simple clause, and “de” in the final position of the sentence. The widely-accepted form of the cleft construction in Chinese usually is “shi...(de)” where “de” is optional, and the constituent after “shi” (usually represented by nominal and adverbial phrases) is normally the focus of the sentence which often gets stressed prosodically; in the case of pseudo-clefts, the structure “…de shi…” (in which “de” represents a nominal clause with the element before “de” usually taking the form of verb phrase) is widely accepted by Chinese scholars (cf. Fang, 1995; Huang & Fawcett, 1996; Shang, 2002). The following sentence is a typical example of a Chinese cleft construction with both “shi” and “de”.

(1-1) 我是昨天来北京的。

Wo SHI\(_{aux}\) zuotian lai Beijing DE\(_{ptcl}\).

I SHI\(_{aux}\) yesterday come to Beijing DE\(_{ptcl}\).

It was yesterday that I came to Beijing.

On the problem of the syntactic form of the Chinese cleft construction, although many scholars reach certain agreement, there are still some related problems staying unsolved into further detail. For example, to begin with, under what circumstances can “de” and “shi” be omitted, and is there any semantic or pragmatic alternation generated by the omission? The solution to this problem leads to a better understanding and summary of the grammatical and semantic functions of “shi” and “de”. In this study,

\(^2\)From here, the abbreviations “NP”, “AP”, “AdvP” and “PP” are short for nominal phrase, adjective phrase, adverbial phrase and prepositional phrase respectively. There are other abbreviations in the following text, for example, “VP” is short for verb phrase, and “OBJ” is for object.
as one of the results, the classifications of “shi” and “de” will be shown in accordance with different types of “shi...de” sentences. Secondly, are all “shi...(de)” sentences cleft constructions? Since there are still no convincing explanations for the above two problems, before the analysis of the information structures of cleft constructions in Chinese, a discussion of the above two aspects is also included in the present study.

One thing I shall mention in particular is that in the present study, pseudo-cleft constructions in Chinese are not discussed in detail. The reasons are as follows:

(i) Unlike the cleft construction in Chinese, whose focus domain cannot be the object, the focus domain of the pseudo-cleft construction can only be the object. Here is one example of the pseudo-cleft construction in Chinese:

(1-2) 打小华的是小明。
Da Xiaohua DErel SHIcop³ Xiaoming.
Hit Xiaohua DErel SHIcop Xiaoming.

The person who hit Xiaohua is Xiaoming.

The different focus domain in the above two structures can be explained by the nature of “de” and “shi” in the structure. “De” in pseudo-cleft constructions is widely accepted as a relativization which establishes the connection between the VP and the omitted subject “人 (ren; person)”. The

---

³ “Shi” is classified, in the present study, into two types: one is a copula in most of the “shi” copular sentences; the other type acts as an auxiliary in Chinese cleft constructions and emphatic sentences. A detailed discussion of the differences between these two types and how do they work in each type of the sentence can be found in chapter 5, but for the sake of the consistency of the format, I show the classes of “shi” in all the examples. In the above example, “shi” is a copula. “De”s are divided into 5 classes (i.e. a particle; a nominalizer; a relative pronoun; a modifier marker, and a possessive marker). In this example (a pseudo-cleft sentence), “de” is a relative pronoun being the head of the relative clause “the person who hit Xiaohua”.

5
relative clause “打小华的 (人) (da Xiaohua de (ren); the person who hit Xiaohua)” is the old information or the topic of the sentence, and the constituent following “是 (shi)” which is the copula of the sentence, is the new information or the focus of the sentence. So compared with the various patterns of the information structure of Chinese cleft constructions, the model of the information delivery in pseudo-cleft constructions is relatively unitary.

(ii) Regardless of the particular context, pseudo-cleft constructions function similarly to equivalent sentences linked by copular “shi” in Chinese (see Hashimoto 1969; Makino 1968; Muraki 1970 and Paris 1979). The most important functional feature of cleft constructions is the contrastive focus meaning, and the only reading of the cleft construction is the contrastive focus reading. However, pseudo-cleft constructions usually have two readings; one is a broad reading (reading 1) and the other is a narrow focus reading or contrastive reading (reading 2). To take two examples:

(1-3) 是小明打小华的。

\[\text{SHI}_{\text{aux}}^4 \text{ Xiaoming da Xiaohua DE}_{\text{ptcl}}.\]

\[\text{SHI}_{\text{aux}} \text{ Xiaoming hit Zhangsan DE}_{\text{ptcl}}.\]

It was Xiaoming who hit Xiaohua.

---

4 In Chinese cleft constructions, “shi” acts syntactically as an auxiliary. It shows different syntactic properties compared with “shi” in other similar sentences, such as “shi” copular sentences in Chinese. The detailed analysis of the syntactic and functional classes and properties are shown in Chapter 5. The auxiliary “shi”, as a distinguishing element, is shown as “\text{SHI}_{\text{aux}}” in the examples of Chinese cleft constructions. “\text{DE}_{\text{ptcl}}” signals that “de” acts as a particle in Chinese cleft constructions.
Focus reading: Xiaoming was the particular person who hit Xiaohua, but not the others.

(1-2) 打小华的是小明。

Da Xiaohua DE_{rel} SHI_{cop} Xiaoming.

Hit Xiaohua DE_{rel} SHI_{cop} Xiaoming.

The person who hit Xiaohua was Xiaoming.

Reading 1 (broad reading): The one who hit Xiaohua was Xiaoming (Note: the answer semantically acts as a brief introduction).

Reading 2 (focus reading): Xiaoming was the particular person who hit Xiaohua, but not the others.

Considering that the aim of the present study is to explore the properties of the information structure (i.e., presupposition, focus, topic, etc.), I prefer not to take pseudo-cleft constructions into consideration in the present study.

Most of the previous studies of Chinese cleft constructions are about the structures themselves. They mainly deal with the following questions, and I summarize the results of the discussions here (the detailed illustration of the previous studies will be shown in the next chapter):

(I) The syntactic forms of cleft constructions in Chinese

Until now, scholars concur with the long tradition of identifying “shi (NP/PP/AP) … (VP) (de)” as a cleft construction (Fang, 1995; Huang & Fawcett, 1996; Paul & Whitman, 2008; Shang, 2002; Teng, 1979). However, the prototypical and variant forms of Chinese cleft constructions should be confirmed. On one hand, not all sentences with the structure
“shi…(de)” can be taken as cleft constructions; and on the other hand, some sentences with the same function as clefts only hold the forms of bare “shi” or bare “de”. So can they be counted as cleft constructions too? If they can, what is the prototypical form of the cleft construction in Chinese—“shi…de”, bare “shi” or bare “de”? If not, where can we put the non-qualifying sentences? In other words, what are their sentence types? To conclude, the question is to figure out under what circumstances “shi” or “de” can be omitted but still keep the nature (mainly functionally) of the cleft construction? If the sentences changed in any way, either syntactically, semantically or pragmatically, how can they be distinguished from cleft constructions?

In the present study, I have taken “shi…de” as the assumptive prototypical form of Chinese cleft constructions for the sake of convenience when searching the corpus. I have also included sentences with bare “shi” and bare “de” in the search process. By analyzing the semantic and informative relations involved in the sentence, and by comparing the different distributions of these structures in the corpus, the present study aims to affirm the prototypical and variant forms of cleft constructions in Chinese.

(II) The syntactic nature of “shi” and “de”

Chinese cleft constructions realize their contrastive focus function by the use of “shi” and “de”. “Shi” is widely accepted as the focus marker (Shi, 2006; Teng, 1979; Zhan & Feng, 1995), following which the focus constituent is located. “Focus marker” is the functional classification of “shi”, which makes its syntactic classification the center of the discussion. The syntactic classification of “shi” also helps with distinguishing different sentence types.

The various syntactic explanations of “shi” in Chinese cleft constructions are arranged below:
(i) “shi” as a main verb (Hashimoto, 1969; Li & Thompson, 1981; Ross, 1982; Tang, 1983; Xu & Li, 1993; Zhu, 1978)

(ii) “shi” as an adverb (Huang, 1982)

(iii) “shi” as a raising auxiliary (Huang, 1982; 1988)

(iv) “shi” as a modal verb (Shi, 1994; Tsao, 1990)

(v) “shi” as a copular (Hu & Wen, 1990; Wang, 1985; Zhan & Sun, 2013)

In the present study, “shi” is presented functionally as a focus marker. To explain why some constituents of the sentence cannot be focused by putting “shi” in front of them, I treat “shi” as a raising auxiliary. The detailed discussion will be conducted in Chapter 5.

Like the studies on “shi”, the studies of “de” in “shi...de” vary among scholars. The previous approaches to “de” should be summarized according to the following criteria:

(i) “de” as a modification-marker approach (Simpson & Wu, 1999; Ross, 1983; Zhu, 1961, 1978)

(ii) “de” as a particle approach (Chiu, 1993; Chu, 1979; Tang, 1983; Teng, 1979)

In the present study, “de” is treated as a sentence-final modal particle whose main function is to trigger the modality of the sentence. Other properties, such as the expression of the past tense or perfect aspect, depend on different conditions which were covered in detail later.

Because there is still no standard conclusion about the forms of Chinese cleft constructions, one of the tasks of the present study is to provide a more comprehensive and systematic illustration of these constructions. The assumptive definition of Chinese cleft constructions at the early stage of the present study is stated as:
Chinese cleft constructions are marked constructions expressing contrastive focus meanings by deliberately partitioning unmarked sentence structures into two segments from the perspective of information structure. In Chinese, the “shi…(de)” structure is treated as the prototypical structure of the cleft sentence, based on which some variants are produced by the omission of “shi” or “de” respectively. By uttering the cleft constructions, speakers establish the relationship between given and new information by providing the possible presupposition shared by both the speaker and the hearer, and also providing the exhaustive set of possible focus items. Pragmatically, the application of the cleft constructions reflects how the speaker speculates on the knowledge and perception of the hearer. After all the analysis, a more accurate system of the judgment of cleft constructions was postulated as a result and part of the conclusions.

The main research object of the present study is the information structure of Chinese cleft constructions. Information structure intervenes at all meaning-bearing levels of the grammatical systems (Lambrecht, 1994, p. 6). It is formally manifested from perspectives of particular grammatical markers, of the form of syntactic constituents, of the word order, of the sentence arrangement, of the formations of complex grammatical constructions, of certain choices between lexical items and of prosody. On information-structure analysis, Lambrecht states:

Information-structure analysis is centered on the comparison of semantically equivalent but formally and pragmatically divergent sentence pairs, such as active vs. passive, canonical vs. topicalized, canonical vs. cleft or dislocated, subject-accented vs. predicate-accented sentences, etc. (Lambrecht, 1994, p. 6)

Besides analyzing the information structure of Chinese clefts itself, this study aims to explore the relationship between the syntactic structure of cleft constructions (i.e.,
syntactic categories) and their information structure (i.e., focus/topic) by observing the different distributions of syntactic categories as focus or topic constituents, and by checking the identifiability of the focus and topic constituents.

1.2 Why study cleft constructions in Chinese?

When looking at Chinese cleft constructions in Chinese, it is clear that they are not as thoroughly studied as those of English. The studies of the information structure of Chinese cleft constructions are even fewer. While English cleft constructions have been studied from perspectives of forms and functions, the relevant Chinese studies are mainly lingering in the discussions concerning the syntactic forms and properties of cleft constructions.

It is important to note here that, although this study is not a comparative study between Chinese and any other languages, English is still considered as the point of reference because of its prominent influence in this field.

Moreover, inevitably, the information structure of Chinese cleft constructions has not been studied before. Being an important means of information-packaging, cleft constructions need to be investigated pragmatically in discourse.

Li & Thompson (1976) propose that the syntactic notions like subject and object are not grammaticalized in Chinese. LaPolla (1990) supports their study by showing that Chinese is a language in which there is no grammaticalization of the syntactic notions “subject” and “object”. For them, compared with the syntactic structure, the information structures are more reasonably used in Chinese to convey information. Although in my opinion, LaPolla’s (1990) ideas are a little too extreme if we take some language facts into consideration, I do accept some of Li & Thompson’s (1976) postulations, such as that they do not deny the existence of “subject” and “object” in Chinese. So studying the Chinese
cleft constructions from the perspective of the information structure seems to be a reasonable method as well as a meaningful task.

Some studies of Chinese cleft constructions are conducted from the perspective of contrastive analysis between English and Chinese. Indeed, the concept of Chinese cleft constructions is derived and translated directly from that of English. While doing the Chinese study, the researchers can hardly avoid comparing Chinese with English. But the fact is that, as two different language systems, Chinese and English in many aspects simply are not comparable. For example, syntactically, the English clefts and pseudo-clefts are formed by syntactic movements, such as wh-movement (Akamjian, 1970); but for Chinese clefts, there is no such movement in the process (Waltraud & John, 2008). Considering the differences in the syntactic structures between English and Chinese constituents, cross-linguistic studies conducted from the perspective of syntax will be less practical. However, if we take information structure as the starting point, what we really are concerned with is how the information (which is a pragmatic and functional unit) is delivered and accepted successfully. This study is an information structure analysis conducted on the syntactic term—Chinese cleft construction. Despite the fact that the point of departure is the information structure, the premise of the analysis is still to begin by confirming the syntactic forms of Chinese cleft constructions, because obviously they have not been settled yet.

The data used in this study are collected from the online Chinese corpus—CCL, established by the Center for Chinese Linguistics of Peking University. It is a large corpus encompassing both ancient and modern Chinese texts, with a total capacity of 581,794,456 characters. The data in the present study are chosen from modern Chinese in the genres of modern plays, TV talk shows, and essays such as legal papers and social science essays,
among which the former two are included in the “informal” sub-corpus, while the latter two make up the “formal” sub-corpus. The reason why the above genres are taken into consideration is that there is still no study about the frequency of Chinese cleft constructions in informal or formal discourses, which means we cannot neglect the possible difference caused by different genres. To avoid bias, in the present study, I have included different genres as representatives of both informal and formal texts. But two things have to be noted: firstly, the distributional differences of cleft constructions in different genres are not the key point of the present study, so different texts are only chosen as the source of data that guarantee balance and objectivity; secondly the reason why no oral data (i.e., audio corpus) are included is because the aim of the present study is to explore the properties (from more than one perspective) of Chinese cleft constructions, so the premise of the study is to make sure that there are enough approachable data for the study. But in oral communications, the main function of cleft constructions—that is, contrastive focus, could be realized through other methods that would cost less effort, such as by putting prosodic prominence on the potential focus constituent. In that case, the potential data should be much fewer in number. This possible phenomenon can be explained by the economy principle of language. Information structures realize their pragmatic functions by various means, i.e., morphologically, syntactically, phonologically, and so on. If the same function can be reached through the phonological method in the process of communication (especially for languages in which meaning is greatly influenced by phonetic factors), the usage of syntactic methods would be optionally

---

5 As a tenet or tendency shared by all living organisms, the concept of economy may be referred to as “the principle of least effort”. It explains humans’ tendency to achieve the maximum result by the minimum amount of effort that is necessary. Being a biological principle at the very beginning stage, this phenomenon has been found in the linguistic field by scholars such as Whitney (1877), Sweet (1888), Leopold (1930), and Zipf (1949). Martinet (1955) provides a comprehensive definition of the linguistic economy as “the unstable balance between the needs of communication—which are always changing—and natural human inertia, two essential forces contributing to the optimization of the linguistic system” (Vicentini, 2003, p.39).
reduced depending on the context. But in written contexts, authors cannot resort to the phonetic method; they have to try to realize the pragmatic function through syntactic means, thus excluding the influence of one variable—the phonological features of sentence constituents. Conducted in the written corpus, this study cannot and also does not have to take prosodic factors as distinguishing features of Chinese cleft constructions, although it is true that prosodic factors are important methods in the fulfillment of the information structure of sentences and they do perform differently (syntactically or semantically) in different types of sentences.

In the process of data collection and selection, some criteria must be used to retrieve the relevant data and exclude the superficially similar but non-qualifying data. Because the exact forms (not the assumptive one) of the “target sentences” (i.e., cleft constructions in Chinese) of this study are not determined before the data-selection procedure, and the criteria cannot be explicitly described by regular expressions, the qualifying or non-qualifying data have to be searched manually on the basis of the results of the automatic (i.e., primary) search. After the initial automatic search, all the data were thoroughly examined according to the criteria listed and outlined in Chapter 4. The ones that do not satisfy any of the criteria will be excluded in the selection procedure (i.e. secondary selection). The detailed criteria and the indexed forms for the corpus are stated in detail in the following chapters.

1.3 Aims of the study

Information structure is still a controversial concept, with the main disagreement falling on the definition of its nature (cf. Carlson, 1983; Chafe, 1976; Halliday, 1967; Roberts, 1998; Vallduvi, 1992, 1993). Here in this study, Lambrecht’s point of view is involved and applied. For him, the information structure of a sentence is the formal expression of
the pragmatic structuring of a proposition in a discourse. So in this investigation, to make a study of information structure is to pragmatically interpret the organization and distribution of sentences from aspects of syntax, semantics, and function. To interpret the title, this study explores the process of information delivery in the semantically and pragmatically defined Chinese language phenomenon of Chinese cleft constructions (the establishment of the semantic/pragmatic criteria is also one objective of the present study) by the analysis of their information structure. Considering the disagreement over Chinese studies, during the analysis, I try to establish a more convincing and comprehensive interpretation of the Chinese cleft constructions. The detailed tasks of the present study are shown below:

(i) To identify the appropriate syntactic forms (i.e., both prototypical and variant forms) of Chinese cleft constructions, and to distinguish them from the nonqualifying “shi…de” sentences.

(ii) To classify cleft constructions in Chinese from the perspective of information structure.

(iii) To check the features of the components (i.e., presupposition, focus, and topic) of the information structure of Chinese cleft constructions in different types.

(iv) To explore the identifiability of the components of the information structure of Chinese cleft constructions.

This study, being a qualitative and mostly corpus-based one, is a continuation and development of previous research. Unlike previous studies, which are mainly based on introspective examples, the application of a corpus in the present study provides more attested data. The conclusions of the study firstly, and empirically, give the systematic description and
explanation of cleft constructions in Chinese from the perspectives of syntax, semantics and information structure. Secondly, and typologically, this study shows some characteristics of Chinese cleft constructions.

1.4 Research questions

For many scholars in the field of pragmatics (e.g., Roberts, 1998, etc.), information structure is a topic covered by both sentence-level constructions and by discourse-level structures.

It is generally characterized as a variation of sentential structure along certain parameters to modulate the presentation of the information imparted by the sentence in such a way as to relate that information to prior context. The factors in that relationship are characterized in terms of primitive functional roles such as theme/rheme, focus/background, topic/comment, old/new, etc. These primitives and the correlated information structure of sentences are used to explain the roles of intonational focus and of particular structure focus constructions, of topicalization and other displacement transformations, and of a fairly wide range of other phenomena, including, e.g. word order and the functions of distinguished structural positions in a relatively grammar-free language such as Chinese. (Roberts, 1998:1)

Therefore, this study starts from the sentence-level construction—the Chinese cleft construction; it investigates how the information structure works in the discourse by observing and analyzing its distributions and configurations. To make the goals more specific, the following questions will be discussed:

(i) What are the semantic/functional criteria of cleft constructions in Chinese?
(ii) What are the prototypical form of the Chinese cleft construction and its variants syntactically; and what are the syntactic properties of “shi” and “de”?

(iii) What is the organization of the information structure of cleft constructions selected from the Chinese corpus, and how can we analyze the delivery of information of cleft constructions in context? How are the focus and topic constituents accessed by the listener? And is there a degree of accessibility for different constituents? If so, does the accessibility of constituents differ from one another?

(iv) What are the syntactic categories of focus constituents?

(v) Is there any semantic difference between the focus and topic constituents?

1.5 Research hypotheses

Considering the objectives of the present study, I list the following hypotheses here as the tentative interpretations as well as guidances of the study.

(i) The syntactic properties of English clefts cannot be applied directly to Chinese. Instead, some semantic criteria have to be established for Chinese cleft constructions.

(ii) The traditionally-accepted prototype of the form of Chinese cleft construction ---“shi...de” is agreed by the present study; however, the “shi...” (with “de” being omitted) sentence is not the main variant of cleft constructions in Chinese.

(iii) In English, the focused constituents in cleft constructions will normally not be the verb phrase in the sentence (e.g., *6It is broke the window that I did.);

---

6 The symbol “*” signifies that this sentence is ungrammatical or grammatically ill-formed.
however, in Chinese, the verb in the proposition can also be focused (我是骑车的/ wo SHI qiche lai DE/ I came here by bike).

(iv) The argument or subject of the sentence being the focus constituent is a common phenomenon in Chinese cleft constructions.

(v) There is also a corresponding relationship between the syntactic categories of cleft components and their information structure.

1.6 Organization

The remainder of this thesis is structured as follows:

In Chapter 2, the previous research based on cleft constructions is elaborated across languages. Among the different languages, English clefts, as the initiation of the cleft sentence studies, and Chinese clefts, as the objectives of the present study, are discussed in detail.

Chapter 3 discusses the theoretical background and the succession as well as the development of the theories in this study. The involved concepts are presupposition, assertion, focus, etc., which realize the process of meaning-delivery as a combination of pieces of information; and identifiability and definiteness, which encode how the syntactic constituents are valued on aspects of (cognitive) perception (of both speaker and listener) and grammatical manifestation. Cleft constructions in Chinese are examined from these two levels in order to get some conclusions regarding their particular features of information delivery in clefts. Chapter 4 introduces the main methods and the steps of the research. Besides that, specific ways of selecting the data and excluding the non-qualifying ones are illustrated here.
Chapter 5 displays the data collected and selected in the corpus. Based on the data, the prototypical structure of Chinese cleft constructions is derived according to the distributions of the possible structures. The nature of “shi” and “de” is reanalyzed by taking the conditions of omission into consideration. Also, another important part of the chapter is an overall summary of the classification of the syntactically similar “shi...de” sentences by illustrating their different semantic features. In the summary, the role of the cleft construction in Chinese and the relationship between it and other similar sentences are shown.

Chapter 6 furthers the analysis of the Chinese cleft constructions from the perspective of information structure combined with the types of focus structures. To understand how information is delivered by clefts, the connections between syntactic constituents of the clefts (i.e., of their focus and topic constituents) and their identifiability, as well as their thematic relations, are established. The data collected from the corpus are checked from the perspectives of the syntactic structure of their focus and topic constituents, their thematic relations, and their identifiability according to the context. And the correlations and connections are established between focus and topic constituents in and across the above different perspectives.

In Chapter 7, a summary of the findings and significances of the study is presented. The conclusion consists of the following parts: firstly, the syntactic structure of Chinese cleft constructions; secondly, the configuration of the information structure of Chinese cleft constructions; thirdly, the thematic relations of the focus and topic constituents; and fourthly, the identifiability of the focus and topic. The significance and the future ideas of the study are also stated in this chapter.
Chapter 2

Literature Review

2.1 Introduction

This chapter consists of two parts, firstly, a summary of the previous literature of cleft constructions across languages; secondly, a summary of the literature of Chinese cleft constructions.

Following the studies of English, the studies of cleft constructions extend from language to language. However, the cases of the studies of cleft constructions in different languages vary among themselves. Studies of cleft constructions in English are well-formed and consist of diverse opinions from various perspectives. Studies of cleft constructions in other languages mainly pay attention to the problem of whether cleft constructions are preferred in realizing the same function as English cleft constructions do. Studies of cleft constructions in Chinese discuss the following four problems: firstly, the forms of cleft constructions in Chinese (Cheng, 2008; Fang, 1995; Gao, 1986; Hashimoto, 1969; Huang & Fawcett, 1996; Shang, 2002; Li, 1981; Lü, 1944; Paul & Whitman, 2008; Song, 1978; Teng, 1979; Wang, 1944, etc.); secondly, the syntactic classes of “shi” and “de” in Chinese cleft constructions (Chiu, 1993; Chu, 1979; Hashimoto, 1969; Huang, 1982; Huang, 1999; Li & Thompson, 1981; Ross, 1983; Shi, 1994; Shi, 2006; Simpson & Wu, 1999; Tang, 1983; Tang & Tang, 1997; Teng, 1979; Xu & Li, 1993; Zhang & Feng, 1995; Zhu, 1961, 1978, etc.); thirdly, the functions of Chinese cleft constructions (Cheng, 1998; Hole 2011); and fourthly, although not very thoroughly, the information structure of cleft
constructions in Chinese is studied as a subcategory of focusing on emphatic sentences (Hole, 2012).  

As a linguistic phenomenon, the cleft construction has been gaining much attention from linguists for its contrastive focus function in the sentence. The discussion of the cleft construction not only stimulates the development of the theoretical principles of this construction, but also offers a typological topic for studies of different languages. In this chapter, I will review the previous studies in order to pave the way for the present study.

2.2 Cleft constructions in languages other than Chinese

Cleft constructions are firstly discussed with respect to English, and then the counterparts in other languages are observed and studied thereafter. Based on the theoretical foundation of English cleft constructions, scholars of other languages began to explore the cleft constructions in their own languages. From the perspective of translation, English cleft constructions are grammatically interpreted into their counterparts in different languages. However, considering the perspectives of function and information structure, native speakers claim that cleft constructions sometimes are not preferred in their language (cf. Asya, 2004; Belletti, 2005; Engel, 1988, etc.). In other words, cleft constructions in other languages may not be the best choice to realize the functions of cleft constructions of English. This section involves two parts: firstly, an introduction to the structure and nature of cleft constructions (in English) and the previous literature about them; secondly, a summary of studies of cleft constructions in other languages (in this thesis, notably, cleft constructions in French, Italian, German, and Japanese are included).

---

7 The prerequisite of the four aspects of Chinese previous literature is that scholars agree on the existence of the counterpart of English cleft constructions in Chinese.
2.2.1 Cleft constructions in English

2.2.1.1 The definitions of cleft constructions

The cleft construction, like many other linguistic concepts, has been defined and redefined time and time again. Here are some samples defining the *it*-cleft, quoted from Calude (2007):

(i). Quirk et al. (1985, p. 951): “(the cleft construction is) a special construction which gives both thematic and focal prominence to a particular element of the clause […] so called [ cleft sentences] because it divides a single clause into two separate sections, each with its own verb”

(ii). Sornicola (1994, p. 4638): (cleft sentences are) defined by the “clefting pattern”“be+NP+S, where the NP is focused and in English, one also finds the dummy it”.

(iii). Bussmann (2000, p.76): Cleft/ pseudo-cleft constructions are syntactic constructions where a single clause has been divided into two clauses. The term “clefting” refers to the transformation in generative transformational grammar which derives the cleft construction from the basic construction. Cleft/ pseudo-cleft sentences serve to mark the constituents that are the focus of the sentence and are especially used to indicate contrast.

(iv). Lambrecht (2001, p.467): cleft construction is defined more comprehensively, from a structural as well as pragmatic standpoint as “a complex sentence structure consisting of a matrix clause headed by a copula and a relative or relative-like clause whose relativized argument is coindexed with the predicate argument of the copula. Taken together, the matrix and the relative express a logically simple proposition, which can also be expressed in the form of a single clause without a change in truth conditions,”
(v). Huddleston and Pullum (2002, p.1417): (cleft constructions are) defined structurally as the formula It/ DEMONSTRATIVE PRONOUN +BE+foregrounded bit+(REL CL),

(vi). Pavey (2003, p.1): “the cleft construction is a marked syntactic bi-clausal option which expresses a simple semantic proposition; in terms of information structure, the construction places an element in focus position, within a copular matrix clause”

Delin and Oberlander (2005, pp.2-4): although they do not give a strict definition as such, they list four main characteristics which capture the essence of cleft constructions: uniqueness, presuppositional nature of the construction, separateness between information structure and presupposition, and stativeness. (pp. 58-59).

Most of these definitions or descriptions are syntax-oriented, and they are applicable to languages such as English. The syntactic partition of the sentence into a relative clause and a matrix clause is not applicable to Chinese counterparts. The focus marker “shi” and the particle “de” are only inserted elements, which do not influence the syntactic structure of the original basic sentence. So the definitions from the syntactic perspective are not borrowed in the case of Chinese. Comparatively, Delin and Oberlander’s (2005) description of cleft constructions is a cross-linguistic one that expresses the semantic and functional features of cleft constructions. The present study is also a functional oriented one, so semantic and functional criteria of cleft constructions are preferred.

---

8 E.L.Keenan (1975) distinguishes the notion “basic sentence” by the following two criteria: (i) a sentence A is more basic than a sentence B if, and only if, the syntactic form and the meaning of B are understood as a function of those A. (e.g., the form of B is some modification [possibly addition to] that of A, and the meaning of B is some modification of that of A.) (ii) a sentence is a basic sentence in L if and only if no other sentence of L is more basic than it.
2.2.1.2 Seminal studies

Jespersen’s studies can be considered as the most original and theoretically neutral ones. His latter studies explore cleft constructions in the field of semantics as well as pragmatics.

Jespersen (1927) makes the following observation:

…(restrictive clauses introduced by it is ) are interesting from a logical point of view because it is not really the antecedent (or what looks like the antecedent) that is restricted by a relative clause. When we say “it is the wife that decides” or “it was the Colonel I was looking for” what we mean is really “the wife is the deciding person” and “the Colonel was the man I was looking for”: the relative clause thus might be said to belong rather to “it” than to the predicative following after “it is”. (1927, p. 88f)

In his later works (1937:83-89), Jespersen criticizes his own earlier account, which by him was called the “transposition theory”. Instead, he considers this type of relative as a special kind of “parenthetic clause,” by which he means cleft sentence. From then on, the concept of “cleft sentence” officially came into being. The following items are the syntactic properties of cleft constructions proposed by Jespersen and summarized by Lambrecht (2000, p. 464):

(i) The relative clause and the preceding predicative phrase must be adjacent to each other;

(ii) They cannot be separated by a pause;

(iii) They are “intonationally coherent”;

(iv) The relative pronoun or marker may be absent in some language;

---

9 “Transposition” relates to a structural “movement” analysis; while “parenthetical” is also structural, but it implies the “insertion” of a clause into another and also implies a functional implication (i.e. an afterthought) or information at a different level.
(v) In English, “that” is used preferably to “who” or “which”;

(vi) In English and French, no comma is used before the relative clause;

(vii) Almost universally, the relative-clause verb agrees in person and number with the immediate antecedent;

(viii) There are languages, like Italian, where no pronoun corresponding to “it” is used;

(ix) There is no possible substitute for “it” when the element following “it” is an adverb or a similar word;

(x) Unlike ordinary antecedents of relative clauses, “it” cannot be stressed. Lambrecht (2000, p. 464)

In order to replace the transposition analysis, Jespersen makes the following interesting syntactic proposal afterwards. He takes the sequence it is, together with the “connective word” (the relative pronoun or marker), as a kind of “extraposition,” symbolized by “[]”; and he treats the rest of the sentence as if the extraposed words were absent. Thus in the sentence It is the wife that (or who) decides (cf. Lambrecht, 2000) wife is not a P (predicative) but an S (subject), and the words it and is are what Jespersen calls the “lesser subject and verb,” symbolized by lower-case s and v. Thus:

(2-1). It is the wife that decides: [sv] S [3c] V

(2-2). It is the wife who decides: [sv] S [sc] V

A cleft sentence without a connective word such as It was John we saw is represented as follow:

\[10]^{10}[3c] and [sc] stand for the connectives that and who, respectively.
(2-3). It was John we saw: [sv] O S V

With this representation, Jespersen observes the intuition of native speakers which tells that the cleft constructions are semantically equivalent to the canonical noncopular sentences, so sentences (2-2) and (2-3) are semantically equivalent to the sentences *The wife decides* and *We saw John*; he also observes that the speakers’ pragmatic intuition tells that (2-1) and (2-2) are related to the subject-accented *The WIFE decides*, while (2-3) is related to the object-accented “focus-movement” structure *JOHN we saw.*

Typologically, Jespersen also offers the following explanation for the use of cleft structures across languages:

In some, though not in all cases, this construction may be considered one of the means by which the disadvantages of having a comparatively rigid grammatical word-order (SVO) can be obviated. This explains why it is that similar constructions are not found, or are not used extensively, in languages in which the word order is considerably less rigid than in English, French, or Scandinavian languages, thus German, Spanish and Slavic. (1937, p.85)

Finally, Jespersen (1949) proposes the following interpretations of the functional use of cleft sentences:

A cleaving of a sentence by means of it is (often followed by a relative pronoun or connective) serves to single out one particular element of the sentence and very often, by directing attention to it and bringing it, as it were, into focus, to mark a contrast (p.147f.)

To sum up, Jespersen (1927; 1937; 1949) coins the name of the “cleft” construction in English; and then defines the syntactic properties and functions of English cleft
constructions. By relating functions with syntactic forms, he concludes, typologically, that languages with less rigid word-order may use cleft constructions more seldom.

### 2.2.1.3 The classifications of cleft constructions and syntactic structures of cleft constructions

As with most scholars, Collins (1991) classifies cleft constructions into three main types—*it*-clefs; basic *wh*-clefs and reversed *wh*-clefs. The last three examples here are the representatives of the three types following a canonical sentence.

(2-4) They were discussing pragmatics.

(2-5) It was pragmatics that they were discussing.

(2-6) What they were discussing was pragmatics.

(2-7) Pragmatics was what they were discussing.

Besides the three types, there are also several special subtypes of cleft constructions. They are *if-because*-clefs, *all*-clefs, *since*-clefs and *only*-clefs (Collins, 1991). In his study, the variants of clefts are not discussed in particular.

By observing and summarizing the phenomenon of clefts and pseudo-clefts, Collins puts forward his models of the structure of the two constructions on the basis of Prince’s models. According to Prince (1978), the structure of English cleft is: *It is/was C which/who(m) / that S-C* (S=sentence, C=constituent). Collins (1991, p.36) subsequently completes the structure as follows:
Prince’s (1978) pseudo-clefts model is expressed as *What S-C is/was C*, while Collins’ modified version is:

```
What/ the (adv) (the/ thing) (that/ (prep) which/ \(\emptyset\))
Who/ that (adv) one/ that/ who/ (prep) who/ (prep) which/ \(\emptyset\)
Where/ the (adv) place/ (that/ where/ (prep) which/ \(\emptyset\))
When/ the (adv) time/ (that/ when/ (prep) which/ \(\emptyset\))
Why/ the (adv) reason/ (that/ why/ (prep) which/ \(\emptyset\))
How/ the (adv) way/ (that/ (prep) which/ \(\emptyset\))
```

Figure 2-1. Structure of the English cleft construction (Prince, 1978).

Figure 2-2. Structure of the pseudo-cleft construction (Prince, 1978).
The main research object of the present study is the cleft construction in Chinese (whose counterpart in English is the “It-cleft” construction) instead of other variant types of cleft constructions (i.e. pseudo-cleft contructions; wh-reversed-cleft constructions). So other types of cleft constructions are not discussed in the present study.

There are heated discussions about the structures of cleft constructions in English. On the structures of cleft constructions, studies can be divided into two schools according to the different aspects they are concerned with: traditional school and transformational school.

(i) Traditional school:

Traditional studies on cleft constructions mainly discuss the formal structures of cleft constructions. It is Poutsma (1928) who firstly observed the existence of the construction (although the term “cleft” is coined by Jespersen (1937) later). In his book, Poutsma holds the view that the words string “it be...that” is semantically empty and that it is merely used for structural purposes. Poutsma observes that the structure of the cleft sentence is the artificial expedient of a normal sentence to get an effect of giving prominence to “the word (group) which indicates the foremost notion in the speaker’s thoughts” (1928, p.140).

The objectives of the studies of Quirk, et al. (1985) are mainly concerned with the use of “it”, the restrictions of the focus, the relative clause, and the information processing. Considering the use of “it”, Quirk et al. state that “it is the most neutral and semantically unmarked one among the personal pronouns, and it is used as an ‘empty’ or ‘prop’ subject, especially in expressions denoting time, distance or atmospheric conditions.”(1985, p.348) Besides, it is worth mentioning that the authors made great contributions to the
restrictions to the “focus”. From their point of view, subject, direct object and adjunct are free to enter the scope of focus in the constructions; whereas indirect object and object complement can only marginally act as the initial focus of a cleft sentence. And more usually, the indirect object would be replaced by a prepositional phrase. There are strict restrictions on the use of subject complement in its function of focus, especially with the verb “be” at the end of the second clause and when the subject complement is realized by an adjectival phrase.

(ii) Transformational school:

With the ultimate aim of figuring out the deep or underlying structure of the construction, transformational studies of cleft constructions mainly focus on the origin of cleft constructions. The scholars in this group propose at least four answers to explain the question of where the cleft sentence comes from.

Lee (1963) is one of the earliest scholars who analyzes cleft sentences from a transformational-generative point of view. He believes that the cleft construction is a modification of its corresponding simple sentence. By bringing the selected pivotal expression to the front of the sentence, we can change a simple sentence into a cleft one. He also proposes that the cleft sentences may be followed by a prefixed WH-morpheme which is also generated by the derivation of the pivotal expression (1963, p.377).

Akmajian (1970) claims that “the cleft sentence is syntactically derived from the pseudo-cleft sentence with the rule of “extraposition”, by which the initial clause of the pseudo-cleft sentence is moved from the end of the sentence” (1970, p.149).

The third version of the source of clefts is given by Gundel (1988). She also assumes that the cleft sentence is derived from the pseudo-cleft, but the method of derivation is
“right-dislocation”. Considering the two classifications of pseudo-cleft constructions, Gundel claims that only identificational equative pseudo-clefts can generate clefts while attributive pseudo-clefts can’t.

The last version is different from the previous ones. Delahunty and Knowles (1986) draw a similar conclusion that cleft sentence is universally “base-generated” regardless of the language differences (1986, pp. 295–317); and the cleft sentence, pseudo-cleft, and their corresponding simple sentences are synonymous at the level of logical structure, so it is inappropriate to say that any of them is derived from the others.

The relationship between cleft constructions and pseudo-cleft constructions is not the key point of the present study. Besides, as far as I am concerned, there is no syntactic movement from the original basic sentence to the cleft construction in Chinese. So the transformational approach is not applicable to the present study.

2.2.1.4 The semantic features of cleft constructions

Semantically, there are some features of cleft constructions which are agreed by some scholars (i.e., Hole, 2011; Huang, 2010; Kiss, 1998, etc.).

(I) Dichotomization of meaning

As proposed by Huang (2010, p.14): “a cleft sentence has the universal semantic property of dichotomizing a sentence into focus and presupposition…”. The distribution of the cleft sentences mainly consists of “presupposition” and “focus” (terms may vary in different studies). In Collins’ (1991) study, he summarizes the nature of the relationship between the two opposite constituents by covering the identifying instead of the attributive function of cleft sentences. He concludes that cleft constructions are identifying constructions, expressing a relationship of identity between the elements realized as the
highlighted element and the relative clause. One thing that should be taken into special consideration is that, as identifying constructions, clefts and pseudo-clefts need to be distinguished from other superficially similar attributive constructions. Collins interprets the differences between attributive constructions and identifying constructions by describing the nature of both of them:

Attributive construction expresses a relationship between an entity and some attribute that is ascribed to it, be it an indication of class membership, a quality, role or other such characteristic”; identification is a relationship between two entities, characteristic, the one serving to define the identity of the other. While identifying constructions are typically reversible, attributive constructions are not. (1991, p.15)

Hole (2011) proposes his understanding of the criteria and features towards cleft constructions:

i. partition—there is a syntactic partition between the cleft constituent and an open sentence (i.e. presupposition).

ii. cleft focus—the cleft constituent often contains focal material.

iii. cleft presupposition—the open sentence is presupposed.

iv. clefts are never necessarily additive—in the absence of contradicting material, the cleft focus is never restricted to an additive reading.

v. topic/ frame-setter—cleft constructions may depend on overt topic or frame-setters. (p. 1709)
The five properties express the relationship between the “presuppositions”, “focus” and “topic”. The detailed illustration of these concepts and the relationship between them are explained in detail in Chapter 3.

(II) Exhaustivity of the focus constituent

Kiss (1998) proposes two semantic properties of cleft constructions—exhaustiveness and exclusiveness. “The first property describes that the constituent under contrastive focus in a sentence is a complete list of the entities that make the truth value of the proposition true. The second property says that those and only those entities under contrastive focus will make the truth value of the proposition true, excluding other entities. (Li, 2008, p.760).”

Buring and Kriz (2003) define the “exhaustivity claim” as:

A cleft of the form ‘It is x that p’ not only expresses that x has property p, but also that x is the only individual to have p, i.e. that x exhaustively identifies p (in the relevant contextual domain) (2013, p.1).

To explain these properties, I display one English example here:

(2-8) It was pragmatics that they were discussing.

In the above sentence, the set “They were discussing X” with “X” representing the variables suitable for the expression of the set “What they are discussing”. The set may be \{pragmatics, semantics, syntax, math, the election, etc.\}. To assign the variable “pragmatics” to “X” makes the proposition a completed one which is called “assertion”, and the focused constituent “pragmatics” is an exhaustive focus.

Delin & Oberlander (2005) express this similar property of cleft construction with a different term. In their study, cleft constructions convey “uniqueness”, which determines that the set of the cleft constituents is an exhaustive listing of the elements presupposed by the cleft clause.
2.2.1.5 Functions of cleft constructions

Considering the discourse perspective, the functions of cleft constructions are widely discussed by scholars and they reach the consensus that the main function of cleft constructions is the function of focusing (Chomsky, 1972; Lambrecht, 2001; Prince, 1978).

Besides, scholars (e.g. Collins, 2001; Delisle, 1975; Li, 2008, etc.) get consensus on the contrastive emphasis of cleft constructions by the special syntactic layout derived from a canonical sentence. To quote from Collins (2001):

The key to why they have evolved as a resource in the English language, I shall assume, is a consideration of a speaker’s communicative needs—the need to select, at a given point in a discourse, a form that is appropriate in the light of assumptions about what information an addressee already possesses, the need to select a form which appropriately emphasizes or focuses upon particular parts of the message, the need to draw a certain contrast, and so forth. (p. 4)

Delisle (1975) explains that the nature of any contrastively emphasized constructions should be traced back to cleft constructions:

It is argued that all contrastively emphasized constructions have underlying cleft sentences, independent of whether the surface structure is an equational or a non-equational one. (p. 419)

However, one thing should be taken into consideration. Contrastiveness is not the particular function possessed only by cleft constructions. In any language, be it English, Chinese, or some other languages, there are other ways to realize the function of contrastive emphasis morphologically, syntactically or pragmatically. Just to name a few, the English “even” sentences; the object-preposing in both English and Chinese; the usage of accent in English or the change of tone in Chinese are all methods of the expression of contrastive emphasis. The differences between other syntactic structures and cleft
constructions will be explained as the distinguishing properties (and also as criteria) of cleft constructions later.

Other scholars also propose that their opinion of the discourse function of the IT-cleft is “contrastive” (cf. Hedberg, 1988; and Biber et al., 1999), and serves for the “unambiguous marking of the focus of information in written English, where the cue of intonation is absent” (cf. Quirk et al., 1985, p.951, and also Sornicola, 1988, p. 372). Its being “oriented towards newness” (Collins, 1987, p.16) leads to a “higher communicative dynamism” (Collins, 1991b, p. 514) than in other cleft types. Aside from this more general function, the it-cleft can be used for “expressing a connection to a preceding text” (Biber et al., 1999, p. 962), and also as a “remind me” toll (Miller, 1994).

2.2.1.6 The information structure of cleft constructions in English

Studies on the information structure of clefts are mainly focused on the distribution of “new” and “given” information in different variations of clefts constructions. Generally speaking, the disagreement between scholars is due to the different classifications of “given” and “new” information. The studies of Collins (1991), and Hedberg & Fadden (2007) are all worth noticing for their elaborate stratification about the “newness” of the information. The information structure of an it-cleft is ascertained by Collins on the basis of whether the cleft constituent and the cleft clause are each “new”, “given”, “contrastive” or “inferable”. The category “new” is further divided into “fresh” and “contrastive”, and the category “given” is further divided into “inferable” and “stable”.

In the study conducted by Hedberg and Fadden (2007), they measure various cleft constructions in terms of referential newness/givenness by evaluating their respective cleft constituents and cleft clauses on a scale containing ten values, as given below (Hedberg & Fadden, 2007, p.5).
Table 2-1 Referential Givenness Categories (Hedberg & Fadden, 2007, p. 5)

<table>
<thead>
<tr>
<th>Discourse Old</th>
<th>Activated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recently activated</td>
</tr>
<tr>
<td></td>
<td>Inferable from activated situation</td>
</tr>
<tr>
<td></td>
<td>Inferable from activated proposition</td>
</tr>
<tr>
<td></td>
<td>Inferable from recently activated proposition</td>
</tr>
<tr>
<td>Discourse New</td>
<td>Familiar</td>
</tr>
<tr>
<td></td>
<td>Inferable from familiar proposition</td>
</tr>
<tr>
<td></td>
<td>Informative</td>
</tr>
<tr>
<td></td>
<td>Cataphoric</td>
</tr>
<tr>
<td></td>
<td>Question-word</td>
</tr>
</tbody>
</table>

The following table is a summary given by Calude (2007, p. 66) detailing the general patterns observed in literature with regard to the information structure of cleft constructions, which only represents the consensus of most scholars.

Table 2-2 The classifications of cleft constructions and the distribution of “new” and “given” information (Calude, 2007)

<table>
<thead>
<tr>
<th>Cleft type</th>
<th>Cleft constituent</th>
<th>Cleft clause</th>
</tr>
</thead>
<tbody>
<tr>
<td>It-cleft</td>
<td>new in informative-presupposition</td>
<td>new or given</td>
</tr>
<tr>
<td></td>
<td>given in stressed-focus</td>
<td>new or given</td>
</tr>
<tr>
<td>Wh-cleft</td>
<td>new</td>
<td>given</td>
</tr>
<tr>
<td>reversed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wh-cleft</td>
<td>given</td>
<td>new</td>
</tr>
</tbody>
</table>
In this study, although we are dealing with the information structure of Chinese cleft constructions; more factors are taken into consideration besides the “giveness” and “newness” of the information, and more criteria are introduced in Chapter 3. The criteria are mainly from Lambrecht's model of “identifiability” (1994, p. 109) and some categories representing “the ways in which the identification of the referents is established” (Chen 2004, p.1164) are borrowed from Chen (2004). These criteria aim to establish the relationship between the grammatical/semantic referents of focus/topic and their mental representations. Chapter 3 gives a detailed introduction to the factors and variables applied in this study.

2.2.2 Cleft constructions in other languages

For studies of cleft constructions in other languages, especially in European languages, the method of “translation” is applied (Delin & Oberlander, 2005). This approach takes a cleft in context in a source language and examines its translation and thus makes a contrastive study from the comparison. The translation approach is widely practiced in European and particularly Scandinavian studies (cf. Aijmer, Altenberg and Hohansson, 1996).

The studies of cleft constructions in other languages are not as sophisticated as those of English. Here I try to state the important and classic studies.

In this section, the literature of cleft constructions in French, Italian, German and Japanese are introduced. Some contrastive studies are conducted in order to compare the acceptability of cleft constructions in different languages. It is generally assumed that clefts (in particular cleft sentences) are a diagnostic feature of neo-standard Italian. This is a new variety of contemporary Italian (Berruto, 1987; Sabatini, 1985) in which it is also believed that the frequency of cleft constructions are on the rise, while being considered absolutely standard in French, i.e. as being a “vital construction found with great frequency
in spoken and written French” (Katz 2000, p. 253), and restricted to certain registers in
German (Altmann, 2009) or even in English, where it is associated with formal and higher
conduct one contrastive study of form and frequency of cleft constructions in corpora.
They state that cleft sentences proper (i.e. the it-cleft construction in English and the
relevant counterparts in other languages) are thus often claimed to occur more frequently
in French than in most of the other main Romance languages (cf. Dufter, 2008) and English
(Katz, 2000). By contrast, the typological literature gives us a somewhat different picture.
Here, cleft constructions seem to enjoy a special status in English, in particular as opposed
to Italian or even to French, because it is claimed that only English has and frequently uses
three different forms of clefts (see Miller 2006, p. 205).

Miller (2006) proposes his typological differences on cleft constructions in English,
German and French as:

English is the most striking in having three cleft constructions which are not only
described in grammars of English, but are in frequent use and occur in the English map
task dialogues (and, of course, in conversation and writing). Other languages in the west of
Europe have one cleft construction or two, but not three. French has IT-cleft and WH-cleft
[…but no Reverse WH-cleft]. German has clefts, which occur far less frequently than in
English, but it has frequently-occurring particles. (Miller 2006, pp. 203-204)

In the following sections, the main problems of the studies of cleft constructions in
French, Italian, German and Japanese are introduced.

2.2.2.1 Cleft constructions in French

In French, the c’est cleft is used as a way to mark the focus. Talking about ways of marking
focus across languages, there are various methods. English, for example, relies heavily on
prosody to mark focus, while French makes comparatively greater use of syntactic constructions to communicate the same focus relations. (Katz, 2000; Lambrecht, 2010). The C’est cleft is a widely-used method to express focus in French.

Example: (2-9) C’est Madeleine qui chantait dans le couloir.

It’s Madeleine who was singing in the hall.

Lambrecht (1994) claims that the usage of the c’est cleft is to be pragmatically motivated to mark focus on elements that occur in positions where French disallows prosodic marking (i.e. French categorically bans prosodic marking from the sentence-initial position). So he concludes that c’est-clefts are used to mark focus on arguments, in other words, in accordance with the focus types\textsuperscript{11} proposed by Lambrecht himself, the focus type of c’est-clefts can only be the argument focus.

The subsequent studies provide different opinions of the cleft’s occurrence. One representative study is conducted by Clech-Darbon et al. (1999). He redefines the syntactic structure of the cleft. While the cleft had been previously discussed as a single CP by Belletti (2005) or as a construction by Lambrecht (1994), Clech-Darbon and his colleagues argue that there are in fact no real cleft sentences per se in French. Instead, a cleft is simply analyzed as the combination of an identificational TP in which the focused constituent is merged as a complement and to which a CP is right-adjoined. The CP is a classical relative clause in which a relative operator moves from SpecTP to SpecCP.

Hamlaoui (2008) agrees with this idea. She proposes that there is one advantage of analyzing the cleft in this way, namely that it makes correct predictions regarding prosody:

\textsuperscript{11}Departing from the focusing function of structures, Lambrecht (1994) divides the types of focus structures into three criteria according to the syntactic classes of the focused constituents—predicate-focus structure; argument-focus structure and sentence-focus structure. The classification is an important concept that will be discussed in detail in chapter 3.
main stress falls on the rightmost edge of an intonational phrase in French, and by creating two separate IPs, the cleft allows the focus element to receive main stress and to fulfil the rightmost preference of the language for accent placement. She also argues that a cleft construction is preferred over a canonical sentence under two circumstances: one is to answer the subject-constituent questions and one is to present contrastive/corrective contexts. Her study challenges Lambrecht’s in the sense that she postulates no need for the focused constituent or the main stress to move a dedicated focus position. Instead, she argues that the mapping of phonology and syntax allow the focused constituent to be directly merged in the position where grammar assigns main stress (rightmost). Grammatical subjects are realized in a cleft to receive rightmost stress when focused, and grammatical objects remain in situ.

Belletti (2005) concludes that when answering a question on the subject, it is more acceptable to use a cleft sentence; the reason is due to the prominent status of clefts in answering questions on the subject. However, when answering a question on the object, because of the computational economy, cleft constructions are not preferred; instead, an in-situ parallelism answer is welcomed.

Concerning the semantic or formal analysis of French clefts, there are only a couple of studies conducted from the perspective of the exhaustivity of French clefts. The representatives are Clech-Darbon et al. (1999) and Doetjes et al. (2004). In the former study, the authors test the exhaustivity under the truth conditions of the sentence and propose that the focus is under a scope of an exhaustive operator. The latter study claims that when the focused element is referential, the exhaustivity of clefts arises. The exhaustivity is also studied as one of the discourse functions of \textit{c’est} cleft by Lambrecht (1994), Katz (1997) and deCat (2007). They all express the exhaustivity in the way of
“identifying the X as having the property P, carrying the inference that nothing else in the context displays the property P.” (Destruel, 2012, p. 98).

To sum up, the currently available literature of French cleft constructions pays great attention to the syntactic form and especially focuses on the discussion of the eligibility and necessity of “cleft constructions” in French. Besides, the exhaustivity gives rise to discussions about whether it is one of the functions or semantic properties of cleft constructions in French.

### 2.2.2.2 Cleft constructions in Italian

The widely-accepted form of the cleft construction in Italian is a null subject sentence, shown in the following example:

(2-10) a. Chi e partito/ha parlato? (quoted from Belletti, 2005)

Who has left/who has spoken?

b. È partito/ha parlato Gianni.

has left/has spoken Gianni.

Belletti (2005) illustrates:

In a null subject language like Italian, a sentence containing a postverbal subject, focus on new information should correspond to a representation along the line as:

[CP…[TP pro…e…partito/ ha parlato…[TopP [FocP Gianni [TopP [VP…]]]]]] (p. 3)

Beninca (1978), Frison (1982, 2001), Graffi (1978), and Grewendorf & Poletto (1989) explore cleft sentences in Italian from a generative perspective; and cleft sentences have been investigated diachronically by Metzeltin (1989), D’Achille et al. (2005), and Sornicola (1991). Cleft sentences in special diamesic variants or genres have also been

However, above all these extensive studies, whether cleft constructions are normal or accepted by native speakers or in the system of Italian itself are the basic questions which are still unsettled.

Asya (2004) proposes that because topic and focus are manifested mainly through linear positions of the constituents, other structures will be preferred in Italian when expressing the same function realized by cleft or pseudo-clefts in English. She claims there are no straightforward counterparts of English cleft and pseudo-cleft constructions.

Belletti (2005) claims that, for reasons of economy, inversion derivation is directly available in a null subject language like Italian, so it is adopted as it involves less structure and, consequently, less computation, than a cleft. However, the structure of the cleft construction is not denied by any grammatical reason, and in fact, cleft constructions seem natural if they are prompted by questions.

He also concludes that:

All other things being equal, null subject languages, as standard Italian, can exploit the inversion strategy marking use of the dedicated focus position in the VP periphery, with a silent preverbal subject. In languages where the null subject parameter is set negatively, one of two different strategies appears to be utilized to realize new information subject: i. creating a non-null subject compatible “inversion”structure, exploiting the informational content of the VP periphery; ii. Adopting a (DP internal) focus-in-situ strategy. French and English illustrate the two options. The quite typical cleft answer, widely adopted in
French, has been interpreted here as a kind of ‘inversion in disguise’ sharing important properties with subject inversion/ VS structures of Italian type. (Belletti 2005, p.18)

From the perspective of information structure (e.g. topic and focus, etc.), other methods of expression are discovered to construct focus in the sentence. Besides subject-verb inversion and cleft formation, Italian also offers focus fronting. Although such structures as “focus fronting” and cleft constructions realize a similar function, they are not totally equivalent to one another. For cleft constructions, exhaustive interpretation is embedded; however, there is no exhaustive interpretation in focus fronting constructions (cf. Brunetti, 2004; Delin & Oberlander, 2005; Frascarelli & Ramaglia 2013; Kiss, 1999). Obviously, exhaustivity is also considered as a property and a criterion of cleft constructions in Italian.

Taking cleft constructions as a legitimate existence of Italian, contrastive studies are conducted by some scholars. From a typological point of view, Lambrecht (1994) points out, the cleft formation is a possible alternative in English and Italian, but is obligatory in French; while for German, he says that cleft formation is a highly unnatural construction used in some given context.

To sum up, different opinions still exist on problems such as the authenticity of cleft constructions in Italian. Nevertheless some consensus has still been worked out. At least, scholars agree that the general classifications of cleft constructions based on English do not fit the cases of Romance languages such as Italian. First of all, the label IT-cleft is not good for pro-drop languages such as Italian, because in these languages, sentences are not started with any form of the pronoun. Secondly, the label WH-cleft is also problematic since in Romance languages it is also supposed to refer to constructions that are not started with a wh-word (i.e. by a free relative pronoun of the form chi/qui/quien “who” or dove/ou/donde “where”, etc. in Italian, French and Spanish, respectively) but by a complex
pronoun, such as *Quello che* in Italian, *Ce que* in French, *Lo que* in Spanish, all representing ‘what’ (Anna-Maria De Cesare, 2014). Thirdly, exhaustivity/exhaustiveness is the semantic property of cleft constructions in both French and Italian.

### 2.2.2.3 Cleft constructions in German

Despite the fact that there are still some unsolved syntactic problems of German clefts, Fischer (2009) applies the following structure as the prototypical layout of German cleft constructions (or “Spaltsatz” in German).

*es/it* + form of *sein/be* + X [stressed] + subordinate (relative) clause

To take one example:

(2-11) Es ist Plato, der unseres Wissens als erster explizit zwischen Nomina und Verben unterscheidet.

It is Plato, who our knowledge as first explicit in between Noun and Verb distinguishes.

It is Plato, who, to our knowledge, first explicitly distinguishes between nouns and verbs.

As mentioned, in the above sections, the syntactic problems are mainly the discussion of whether or not the cleft constructions are legally accepted considering the characteristics of the information structure of German. Durrell (2002, p. 479) observes that unlike German, which can easily move clause elements into the initial position to form the topic of the sentence, English requires a range of complex constructions that actually correspond to simple sentences in German. Although German and English cleft have a similar syntactic configuration, their stylistic value and information structure are very different. The following items are the facts of German cleft constructions (Durrell, 2002).
(i) Clefts appear much less frequently in German than they do in English.

(ii) The reason of the first claim is due to the different fronting opportunities in the two languages.

(iii) Clefts sound unnatural in German.

(iv) The more natural reading of cleft constructions in German is, in fact, an expansion of another construction.

Before him, Doherty (1999, p. 293) remarks that “there are hardly any studies of German clefts alone” and that “even German linguists seem to find the English side more rewarding”. Ahlemeyer & Kohlhof (2000) conduct a corpus-based study by exploring the translation of English cleft constructions in German. They conclude that only about one-third of English it-cleft constructions are translated into German equivalents. Instead, other mono-clausal sentences highlighting a focus XP are preferred to realize the same function of the German cleft sentences. Even earlier, Engel (1988) “warns” learners of German not to overuse clefts, although he considers subject clefting, case complement clefting, and prepositional complement clefts to be grammatically acceptable.

Targeting the field of translation, and to solve the problem of the differing frequency of cleft constructions in English and German, Fischer (2009) concludes that besides the structural differences, the problem can be explained more accurately if stylistic factors are involved. English is a far more verbal language, often favoring subordinate clause constructions over non-clausal alternatives. Clefts are two-clause constructions, their non-cleft German counterparts will usually feature one clause fewer. In a number of respects, the English verbal constructions are semantically more explicit than their German non-clausal counterparts (Fischer 2007, pp.392-398).
In conclusion, although cleft constructions are introduced in German grammar books (e.g. Zifonun et al, 1997, pp.231, 528-529; Dudenredaktion, 2006, pp.1044, 1056-1057), most of the scholars take negative opinions towards the issue of whether cleft constructions are natural or preferred constructions in German or not.

2.2.2.4 Cleft constructions in Japanese

According to Hasegawa (1997), the cleft construction is characterized by the structure [...no wa/ga] X da]. Yoshimura (1994) analyses the two structural patterns of cleft constructions by paying particular attention to the particle “wa” and “ga”. The following two are examples in his study:

(2-12)John-ga kinoo eki de atta no-wa [Focus Mary] da.

John-NOM yesterday station at met person-TOP Mary Copula

It was Mary who John met at the station yesterday.

(2-13) [Focus John-ga kinoo eki de atta NO]-ga Mary da.

John-NOM yesterday station at met person-NOM Mary Copula

The person whom John met at the station yesterday was Mary.

Yoshimura (1994) agrees that the pronominal form “no” can be marked with either the topic particle “wa” or the nominative particle “ga” (cf. Kuno, 1973; Noda, 1996). The two types are relatively corresponding to cleft and pseudo-cleft constructions in English from the perspective of translation. The particles “wa” and “ga” make a crucial difference in the focus interpretation between the two clefts. In the “wa-cleft”, the focus is on the noun phrase Mary in the predicate; while in the “ga-cleft” the focus is on the clause John ga kinoo eki de atta no in the subject position, which is a preposed focus structure (Amano, 1996; Sunakawa, 1995).
Like English cleft constructions, Japanese counterparts have been widely discussed from the perspective of generative grammar (Hasegawa, 1997; Hiraiwa & Ishihara, 2002; Hoji, 1987, 1990; Kizu, 2005; Koizumi, 1995; Kuwabara, 1995, 1996; Takahashi, 2006; Hoji & Ueyama, 1998 among others). On the generative syntactic structure of cleft constructions in Japanese, there are two different results. The first explains that the focus phrase is base-generated in its surface position with an operator movement inside the presuppositional clause, which is very similar to the analysis of the (it-) cleft in English. The other involves the movement of the focus phrase and the subsequent remnant movement of the presuppositional clause.

Followed the terminology in Hoji (1987), cleft sentences in Japanese are divided into two types. The sentences with the case-marker on the focus NP are CM\(^{12}\)-cleft and those without the case-marker are Non-CM-cleft. I quote one example of each type from Ueyama & Hoji (1998):

(i) CM-Cleft:

\[
(2-14) \ [\text{Kokuren-ga} / \text{kibisiku} / \text{[so-re-o]}_1 / \text{hihansita}/ \text{no}] - \text{WA} / \text{[amerika-no/ booei-seisaku-o]}_1 / \text{DA}.
\]

The United Nations-NOM / harshly/ that-thing-ACC/ criticized/ COMP-TOP/ USA-GEN/ defense-policy-ACC/ COPULA.

It was [the USA’s defense policy-ACC]\(_1\) that the United Nations harshly criticized it\(_1\).

(ii) Non-CM-CLEFT

\(^{12}\)CM stands for “case-marker”.

47
(2-15) [Syuukansi-ga/ [neta-ni/ tumaru/ to] /yoku /[so-re-o]/ tokusyuusuru/ no]-wa/ [daietto /to /onsen]/ da.
Magazine-NOM/ topic-DAT/ stuck/ if/ often/ that-thing-ACC/ feature/ COMP-TOP/ diet/ and/ spa/ COPULA.
It is [diet and hot spa], that magazines often feature [it/them], [when they ARE stuck FOR topics].

Based on observation, they claim that there is an empty operator movement in the CM-cleft sentences, but there is no movement in Non-CM-cleft sentences.

Kizu (2005) analyses the cleft construction in Japanese from its topicalization, nominalization, dependency, ellipsis, etc. On the problem of whether there is a movement in cleft constructions of Japanese, he proposes that there does exist operator movement in Japanese cleft construction, but the question of “how the movement takes place” differs dramatically.

In English, it is generally assumed that the null operator moves successive-cyclically; however, this is not what we have found to hold in Japanese. Instead, the null operator of Japanese long-distance clefts is base-generated adjoined to the highest embedded clause in the presuppositional clause and undergoes short movement to the spec of CP. Short A’-movement is also observed in local clefts; the operator at first undergoes A-movement to sentence-initial position, and then further A’-movement takes place. (Kizu, 2005, pp. 205-206).

On the whole, the term “cleft constructions” is coined in English, so studies of English cleft constructions developed rather vigorously. Contrastively speaking, the studies of cleft constructions in other languages still focus on the grammaticality of the structure. As stated, whether cleft constructions are preferred to express the same function as English
cleft constructions do is questioned in studies of French/Italian/German cleft constructions. For studies of Japanese, scholars mostly accept the Japanese cleft construction as a natural way to express focus and contrastive emphasis.

### 2.3 Cleft constructions in Chinese

Until now, scholars have concurred with the long tradition of identifying “shi (NP/PP/AdvP)…(VP) (de)” as a cleft construction. (cf. Teng, 1979; Fang, 1995; Huang & Fawcett, 1996; Shang, 2002; Paul & Whitman, 2008, etc.). However, they also admit the fact that the structure “shi (NP/PP/AdvP)…(VP) (de)” (shortened as shi...(de) clefts hereafter) seems to be extremely different from its counterpart of English (e.g. it is…that).

Paul & Whitman (2008) propose that the distinguishing syntactic property of “shi…de” clefts is that they involve no A’movement. This is due to the fact that the presupposition in the “shi…de” clefts involves a projection smaller than CP. Chinese “shi…de” clefts lack a landing site for A’movement and fail to host material such as negation, normally compatible with full CP clefts in languages such as English. The difference can be traced back to the semantic/pragmatic prominence phenomenon in Chinese (cf. Huang, 1982, 1984; Huang, 2007; LaPolla, 1995; Li & Thompson, 1975), so the function of cleft constructions is not revealed by the aspect of syntactic form, but by cleaving the proposition into two parts from the perspective of information structure.

---

13The name “Chinese cleft constructions” is borrowed from some available studies of this type of structure. Considering that the name of Chinese cleft constructions is derived from English, in which cleft constructions are structurally cleaving constructions, it is not very accurate to name Chinese shi…de structures in the present study as “cleft constructions” in Chinese. However, firstly, there is no better term for this construction available in the previous studies, so following the traditions of the studies of Chinese cleft constructions, I keep using this name to refer to the objectives of the present study; secondly, instead of being syntactically cleft into a bipartition, Chinese cleft constructions do have a bipartition consisting of the focused element and the presupposition (Paul & Whitman, 2008) functionally and semantically. Thirdly, compared with syntactic properties, the functional similarities between English cleft constructions and Chinese cleft constructions are more obvious and worthy of being studied.
Teng (1979) is one of the pioneers of the studies of Chinese clefts, and his study is elaborated here. He made a great contribution to the syntactic forms and characteristics of Chinese cleft constructions, but until now, the problems of the criteria and properties of cleft constructions in Chinese are still unsolved and are still the subject of heated discussion, and happened to be part of our goals for the present study.

2.3.1 The seminal study

Teng’s (1979) study can be counted as one of the preliminary researches on Chinese cleft constructions. In his study, he firstly explains the process of “cleaving” in general; and then gives the criteria he uses when identifying Chinese cleft constructions; finally, he gives his conclusion that he does not agree with the prototype of cleft construction (“shi…de”) in Chinese proposed by Hashimoto (1969) and other scholars although he did mention that the discussion about the form of the cleft construction is still open.

In Teng’s (1979) opinion, “cleaving” is a syntactic description for marking a constituent of a sentence for the purpose of focus, contrast or emphasis by syntactic devices. Although focus, contrast or emphasis can also be indicated phonologically, cleft constructions realize the function of emphasis/contrast by syntactically putting a specific constituent in a prominent position.

According to the general features of cleft constructions, Teng (1979) proposes that there should be three criteria for identifying Chinese cleft constructions: (i) the presence of a syntactic marker (i.e. “shi”) to single out the focused constituent; (ii) the absence in the main clause of an NP which is co-referent to the focused constituent; and (iii) the distribution of “asserted” and “assumed” in the entire sentence such that the focused constituent always bears the property of “asserted”.
Teng states his opinion like the following while displaying his disagreement with Hashimoto (1969) on some minor aspects:

We can establish that, structurally, a cleft sentence in Chinese is one indicated by the occurrence of the non-equative copula verb alone, contrary to the traditional assumption that the shi…deconstruction defines a (pseudo-) cleft sentence in Chinese, cf. Hashimoto (1969). …the selection of the final particle de is dependent on the feature of the entire sentence, cleft or non-cleft. (Teng, 1979, p. 105)

To conclude, he states that, structurally, the prototype Chinese cleft construction is “shi… (de)” instead of “shi…de”, which is traditionally considered to be the prototype by scholars such as Hashimoto (1969). To make his opinion clearer, I will cite two examples from his study below and clarify his idea that “de” is not necessarily needed in cleft-constructions:

(2-16) 我是明天要到纽约去。

Wo SHIaux mingtian yao dao Niuyue qu.
I SHIaux tomorrow will go to New York.
It is tomorrow that I will go to New York.

(2-17) 我是昨天到纽约去的。

Wo SHIaux zuotian dao Niuyue qu DEptcl.
I SHIaux yesterday go to New YorkDEptcl.
It was yesterday that I went to New York.
Both of the two sentences are considered as clefts. However, “de” is just an incidental particle whose appearance is due to the feature of the entire sentence, which is what Teng (1979) insists in his study. So he disagrees with Hashimoto’s (1969) proposal of the structure of the cleft construction.

But two problems have still not been solved in Teng’s study. Firstly, he does not give any analysis of the absence or presence of “de”. In other words, he does not solve the problem of under which condition “de” can be deleted legally without any altering of the sentence meaning and function. Sometimes, the object and de can switch their positions in the sentence, but only under certain conditions (i.e. the action indicated by the embedded sentence must be in the past time, and in addition, its VP must contain an adverbial). Secondly, he disregards the differences between cleft sentences he claimed and other “shi” sentences.

2.3.2 Forms and classifications of cleft constructions in Chinese

2.3.2.1 Forms of cleft constructions in Chinese

Cleft sentences in Chinese\textsuperscript{14} usually take the form of "shi... (de)" where "de" is optional. However, we cannot say that all sentences with this form are cleft constructions. Until now, some scholars have studied the criteria of cleft constructions in Chinese from the syntactic point of view.

\textsuperscript{14} The construction resembles the “shi” emphatic construction in Chinese. As to whether the “shi” emphatic construction in Chinese is the cleft sentence or not, there are different views in the Chinese linguistic literature. Some linguists do not admit there is a cleft sentence in Chinese (cf. shi, 1994). To figure out whether cleft sentences exist in Chinese or not, we have to analyze the prototypical form from syntactic and functional perspectives respectively. It is true that there is no neat correspondence in different languages. But the function of the “shi” emphatic constructions parallels that of the cleft sentence in English. And crucially, the differences between cleft constructions and “shi” emphatic sentences are discussed in the following study.
Early Chinese researchers (e.g. Gao, 1986; Lü, 1944; Song 1978; Wang 1944, etc.) notice that there is a “shi...de...” construction existing in the Chinese language to express emphasis. The consensus on this language phenomenon is that there is a basic sentence beneath the cleft construction. Moreover, the basic proposition is broken up by “shi”.

e.g. (2-18) Basic form: 他昨天到了。

Ta zuo tian dao LE.

He arrived here yesterday.

(2-19) Cleft sentence: 他是昨天到的。

Ta SHI$_{aux}$ zuo tian dao DE$_{ptcl}$.

It was yesterday that he arrived.

(2-20) Basic form: 他第一个写完作业。

Ta di yi ge xie wan zuo ye.

He is the first person who finished his homework.

(2-21) Cleft sentence: 是他第一个写完作业的。

SHI$_{aux}$ ta di yi ge xie wan zuo ye DE$_{ptcl}$.

It is him who finished the homework first.

Sentence (2-19) and (2-21) are all cleft sentences, however, (2-19) puts focus on the adjunct of the sentence (being the following constituent of the focus marker “shi”) while sentence (2-20) puts focus on the subject of the sentence.
Some scholars have given their descriptions of this construction. Zhao (1979) held that in the construction of “shi…de…”, “shi” is not the primary verb; instead it is placed before the verb or the adjunct to denote the emphasis; “de” is placed behind the verb, indicating the tense/aspectual feature of the primary verb. Such construction is used on occasions when the action or situation has finished or has been realized. Liu (1983) assumes that what fills the gap between “shi” and “de” is mainly the verb, verb phrases or the subject-predicate phrases with the verb as the predicate. “de” is also a tense marker of the clause, illustrating that the action indicated by the structure of “shi…de…” has finished or has been realized in the past. The focus of the cleft sentence is not the action itself, but aspects related to the action, such as time, place, the manner, condition, purpose, object or agent, etc.

Comparatively, some scholars (e.g. Cao, 1990; Hashimoto, 1969; Tang, 1983; Zhang & Feng, 1996, etc.) relate the Chinese “shi…de…” construction to the English cleft sentence. Tang (1983) considers the Chinese cleft as a construction with a judgment verb “shi” and particle “de” expressing assertion while separating the whole sentence into two: the first part as the presupposition and the second part as the focus. Zhang and Feng (1996) held that “shi…de…” is the focus-marking construction in Chinese, and “shi” is the focus marker followed by the emphatic part.

**2.3.2.2 Classifications**

Paul & Whitman (2008) propose several forms of the cleft constructions derived form “shi…de”. They subcategorize the classifications of cleft constructions into: cleft focus patterns, represented by the sentence-intial bare shi (subject focus only) and the “Shi…de” proper (subject and adjunct focus); association with focus pattern, represented by the...
sentence-medial bare shi; and the propositional assertion which involves the copula shi and a marker of clausal subordination.

In Cheng’s study (2008), besides defining “shi” as the copula and “de” as the spell-out of null operators, she still re-analyzes sentences with shi and de, including cases with only shi or only de. All the cases involving “shi” share the copular structure, but the subject or predicate in the small clause selected by the copula can differ. In her study, the outlines of the different structures are displayed:

Canonical Predication: XPi COP [sc15 [ti][de-clause]]
Inverse Predication: [de-clause] COP [sc [XP][ti]]
Broad shi-de: proi COP [sc[de-clause][ti]]
Bare-shi: proi COP [sc[CP][ti]]
Bare-de: [OP[IP]-de] (2008, pp. 27-32)

Under her observation, Lisa proposes her conclusion as:

There is no shi…de construction. Shi has no particular affinity with de. We have seen that the shi…de combinations can come about due to different structures. These sentences appear to be related because of the presence of the copula. However, abstracting away from the copula, the different shi-de sentences have different syntactic properties. (Cheng, 2008, p. 39)

Paul & Whitman (2008) analyze that “shi…de” and bare “shi” as two distinct constructions with completely different properties. They consider that constructions with bare “shi” are actually associated with focus constructions instead of cleft constructions. According to them, the distinctions are as follows:

15 “sc” refers to the small clause.
(…the shi…de) construction shows a bipartitioning consisting of the focused element and the presupposition, which is always outside the scope of negation; it is subject to the exclusiveness condition, and focus is positionally determined. In the bare shi construction, by contrast, any constituent to the right of shi that is marked by intonational prominence may be associated with focus, and the exclusiveness condition does not hold. (Paul & Whitman, 2008, p. 420)

To explain the difference, Paul & Whitman’s (2008) refine the classifications of cleft constructions into: the cleft focus pattern which includes “sentence-initial bare shi” construction and “shi…de proper” constructions; the association with focus pattern that refers to medial “shi” sentences, and propositional assertions. To illustrate by several examples:

(2-22) 我是负责这件事。

SHI aux wo fuze zhejian shi.

SHI aux SHI me in charge of this matter.

It is me who is in charge of this matter.

This sentence is a “sentence-initial bare shi” one putting the subject as the focus.

(2-23) 我是昨天到北京的。

Wo SHI aux zuotian dao Beijing DE pcl.

I SHI aux yesterday arrive Beijing DE pcl.

It is yesterday when I arrived Beijing.

(2-24) 是我负责这件事的。
SHI_{aux} wo fuze zhejian shi DE_{ptcl}.

SHI_{aux} me in charge of this matter DE_{ptcl}.

It is me who is in charge of this matter.

Sentence (2-23) is the “shi...de” proper one with the focus being the adjunct of the sentence; while sentence (2-24) is of a similar syntactic type but with the subject/argument being the focus.

The structure “sentence-medial shi” is not considered to be a cleft construction by Paul and Whitman (2008). For them, any constituent in any position in the clause to the right of “shi” can be focused by prosodic prominence. The example in their study is:

(2-25) 老魏是明天去。

Laowei SHI_{aux} mingtian qu.

Laowei SHI_{aux} tomorrow go.

The available translation will be one of the following, according to the actual intonation in the conversation:

It is tomorrow when Laowei is going (somewhere). / Laowei is going somewhere TOMORROW.\(^\text{16}\)

It is Laowei who is going (to somewhere). / LAOWEI is going to somewhere tomorrow.

Laowei is GOING TO somewhere tomorrow, not GOING BACK.

(2-26) 他是跟你开玩笑的。

\(^{16}\) The capitalization refers to the focus constituent of these sentences.
Ta SHIaux gen ni kaiwanxiao DEptcl.

(It is the case that) he was joking with you.

This example is proposed by Chao (1968, p. 296) which is named by him an example of “propositional assertion pattern”. The aim of the sentence is to convey the speaker’s certainty that the proposition holds in a given situation. In this structure, no specific constituent is focused, even though as in cleft constructions, “shi” and “de” are both present.

Li’s (2008) study categorizes Chinese cleft constructions as the contrastive focus structure while classifying them into two types--one is marked by “shi”, and the other by “shi…de”. So Li’s study seems the compromise between Teng (1979) and Hashimoto’s (1969) ideas. Tang (1983) subdivides Chinese cleft constructions into the names of “cleft constructions”, “the derivation of cleft constructions” and “pseudo-cleft constructions”. His classifications are similar to those of Collins’ (1991). One more thing worth mentioning in Teng’s study is that he notices the there is one derivational version of the Chinese cleft construction---represented as “shi…de OBJ” with an objective following “de. In order to examine whether or not the combination of “shi” and “de” really yield a special functional construction; and whether “shi…de” combinations provide a different interpretations from “shi” plus “de”, a better and more important way is to include bare “shi” and bare “de” into investigation.

To sum up, the forms of Chinese clefts have been a controversial topic for a long time. And until now, there is still no agreement on that issue. The following are examples of the discussion.

(I) “shi…de” vs. bare “shi” and the classifications of cleft constructions
In their studies, Teng (1979), Huang (1982, 1988) and Chiu (1993) do not give a distinguishing difference between shi...de and bare shi and propose the same analysis for both. They mainly regard the two constructions as possessing similar functions, and “shi...” is just the brief version of “shi...de” with “de” omitted for the tense reason.

In this study, we are going to include all possible syntactic structures of Chinese cleft constructions in the data-collection procedure, and the classification afterwards is conducted from an informative point of view.

(II) “shi...(de)” cleft sentences vs. other “shi” sentences (identification and class membership)

Generally speaking, there are four types of “shi” sentences in Chinese:

(a) Equational sentences that express the identifications, classification, etc. of entities in the sentence (e.g. 我是老师; wo shi laoshi; I am a teacher);

(b) Attributive sentences which describe or explain the attribution of the entities or events in the sentence (e.g. 花是红的，草是绿的; hua shi hong de, cao shi lü de; The flower is red, and the grass is green.);

(c) Existential sentences (e.g. 楼下是一群小孩; louxia shi yiqun xiaohai; there is a group of kids downstairs). “shi” sometimes acts as the copula of the existential sentences in Chinese.

(d) Emphatic sentences whose function is to emphasize part or the whole of the sentence. The differences between cleft constructions and other contrastive sentences will be elaborated in the analysis section.

The following examples are the illustrations of the above classifications.
Examples (2-27) and (2-28) all belong to the first type. However, the difference is that in the sentence expressing the identification of some constituent of the sentence, the two constituents on the two sides of the copula “shi” can change their positions without any alternation of meaning or any syntactic deformation. But in the case of the sentence in (2-28) whose semantic function is to describe the classification or the attribution of one constituent, the exchange of positions is not allowed. This phenomenon can be used as one important method to distinguish cleft constructions from some of the identifying ones.

In contrast to copular sentences, the function of “shi” in cleft sentences is not to describe or to classify, instead is to focus the constituent directly following the focus marker “shi”. It is not hard to find that “shi” plays a rather important role in distinguishing the types of sentences in Chinese “shi” sentences.

To avoid ambiguity and confusion, the classifications of “shi...do” sentences will be refined in chapter 5 by showing different properties of Chinese cleft constructions from the superficially similar but functional different constructions.
2.3.3 Semantic and functional properties of cleft constructions in Chinese

There is only a limited-number of studies exploring the semantic and functional properties of Chinese cleft constructions.

Hole (2011), in his study, proposes several properties of cleft constructions cross-linguistically. He finds that cleft constructions display the following characteristics: there is a partition between the focus constituent and the rest; there is a focus and a presupposition existing in the sentence; clefts are never necessarily additive and there is a topic/frame setter in the cleft construction.17

To deny the “additive” property of clefts, Hole (2011) is actually talking about the “exhaustivity” discussed by other scholars (e.g. Buring and Kriz, 2013; Kiss, 1998, etc.).

2.3.4 Syntactic/functional status of “shi” in cleft constructions

Different linguists have different opinions about the syntactic status and functions of "shi". The opinions should be categorized into two kinds: considering “shi” as a main verb/copula and considering “shi” as a focus marker.

(I) “Shi” as a main verb/ copular of the sentence

This category claims that “shi” is the main verb in the cleft sentence. Scholars who hold this view are Hashimoto (1969), Li & Thompson (1981), Ross (1983), Tang (1983), and Zhu (1978).

Hashimoto (1969) suggests that “shi” should be treated as a verb on the ground of its special distribution. She said that “shi” can be negated just like any other verbs. In Chinese,

---

17 The detailed explanation of these properties is shown in 2.2.1.4.
only verbs can be negated, hence she didn’t exclude “shi” from the verb category. She treated “shi” as the main verb and clefts/pseudo-clefts as the copulative sentence.

Ross (1983) treats “shi” as a main verb, and believes that the cleft sentence is a special case of the equational sentence because “shi” denotes an indirect equational relation. The clause after “shi”, to some extent, determines the elements before it.

The problems of this approach lie in the following aspects: (i) Scholars holding this approach prefer to treat cleft sentences as equational sentences whose function is different from that of cleft sentences. (ii) Suppose “shi” is a verb as this approach insists, how can’t it appear before object? This problem cannot be solved by this approach.

(II) “Shi” as a Focus Marker

The second solution treats “shi” as a focus marker functionally in the cleft sentence (Shi, 2006; Teng, 1979; Zhan & Feng, 1995). “Shi” itself has no substantial meaning and the cleft sentence is a simple sentence. The scholars thought that “shi” was free to be placed before any constituent before the predicate and did not affect the structure of the original proposition. Hence, the cleft sentence remained a simple sentence.

However, there are four subcategories of the word category of “shi”. The first one displays that “shi” is an independent focus marker in the cleft sentence; the other three consider “shi” as a focus marker belonging to different syntactic categories.

(i) “Shi” as a specialized focus marker

Fang (1995) proposes that “shi” is originally a surplus element in the spoken language, but it has a great effect in the written language. “Shi” is not an essential element in the sequential syntactic structure of a sentence having some substantial meaning, so it can be left out. Teng (1979) holds the view that “shi” in the cleft sentence
is a marker specialized for indicating the focus. By inserting “shi” before the focus of a sentence, the original sentence becomes a cleft sentence.

(ii) “Shi” as a focus marker having the status of an adverb on a par with negation and modals

This view is proposed by Huang (1982). He claims that “shi” is a focus operator which has the attribute of an adverb on a par with negation and modals.” (Huang 1982, p. 213). Its status as an adverb explains why it cannot directly focalize the element behind a verb, but at the same time, it can be inserted before any constituent in front of the predicate and then focalized without any change of the sentence order. For example:

(2-29) Basic sentence: 他打我。

Ta da wo.

He hit me.

(2-30) cleft construction: 他是打我。

Ta SHIaux da wo.

(i) It is me who was hit by him.

(ii) He HIT me.

(2-31) cleft construction: 是他打我。

SHIaux ta da wo.

It is he who hit me.
(2-32) *他打是我。

*Ta da SHI wo.

(iii) “Shi” as a focus marker having the status of a verb

Xu & Li (1993) assume that “shi” is a verb with the function of a focus marker. In other words, the function of "shi" is to denote the focus. At the same time, since "shi" is a verb, it must obey the rule of being a verb.

(iv) “Shi” as a focus marker and a raising auxiliary

Huang has written two papers about the status of “shi” (Huang 1982, 1988). Differing with his earlier statement, Huang (1988) considers “shi” in the cleft sentence to be a subject raising auxiliary like “可能, keneng” (may), “应该, yinggai” (should or ought to), “会, hui” (may) in Chinese, or as “seem” in English.

“Shi” in the clefts is allowed to appear before the subject if the subject is not raised; or between the subject and predicate, if the subject is raised. The subject-raising in the clefts which contain an auxiliary “shi” is not obligatory, but optional; whereas, “会, hui” (may) in the sentence “这本书会涨价” (zhe ben shu hui zhang jia /The price of this book is to raised.) has to be raised (Huang, 1988).

The advantage of treating “shi” as an auxiliary is that the rule for the distribution of “shi” is thereafter restricted. The auxiliary can be positioned before the predicate, but not behind the predicate or prepositions. “Shi” is required to obey these restrictions. The preposition-object structure belongs to the predicate; hence, “shi” and other auxiliaries cannot appear behind these preposition-object structures. That “shi” can be put before or behind the words “可能 (ke neng; maybe)”, “应该 (ying gai; should)” is because these
auxiliaries are similar to the auxiliary “shi”, which can be distributed as the head word of
the parent clause or the subclause.

In the present study, “shi” is analyzed as a focus marker with its syntactic class as a
raising auxiliary, it can be put at the initial position of the sentence; and functionally, as
with other auxiliaries of Chinese, it acts as a modal verb by, for example, enhancing the
confirmative mood of the sentence, etc. This distinguishes clefts from the equational
sentences in which “shi” is acting as the copula; and explains the function of the assertive
mood of the clefts; thirdly, it explains why “shi” cannot be placed between verb and object
to highlight the object of the sentence.

2.3.5 Syntactic status of “de” in cleft constructions
Similar to the studies on “shi”, the studies on “de” in “shi…de” cleft constructions vary
from scholars to scholars. The previous approaches to “de” should be summarized into the
following criteria:

i. Modification-marker approach

Representatives for this idea are Zhu (1961, 1978); Ross (1983) and Simpson & Wu
(1999).

The modifier-marker approach tries to assume “de” in “shi…de” construction as a
modification marker. Ross (1983) includes possessive “de”, adjective “de”, relative clause
“de”, and appositive clause “de” into the same classification of “modification-marker”.
Simpson & Wu (1999) include “nominalizer” and “determiner” in this approach. In
addition, they reanalyze “de” from category D to category T and make it a past tense
morpheme base-generated as the head of TP.

ii. Particle approach:
The relevant representatives are Teng (1979); Tang (1983); Chiu (1993); Chu (1979).

Unlike Simpson & Wu, Chu (1979) does not suggest that “de” can be a past tense marker. He proposes that “de” is a particle and it can contain the presupposition. The presupposition of a cleft sentence is defined as the string preceding “de” except the focus. Teng (1979) and Tang (1983) regard “de” as a particle and it occurs independently from the subject and the predicate.

In the present study, “de” is treated as an ending particle that may present the tense or mood of the sentence. A more detailed classification of the syntactic nature of “shi” and “de” will be presented in the analysis section (Chapter 5) in accordance with the different sentences structures in which they are involved.

2.3.6 Information structure of cleft constructions in Chinese


Hole (2012) summarizes the information structure of Chinese in accordance with aspects of “focus”, “topic” and “given & new” information. From the perspective of “focus”, firstly, just as that Chinese is a wh-in-situ language, it is also a focus-in-situ language (cf. Huang (1982) or Soh (2006) for discussion of the overt and covert wh-syntax of Chinese); secondly, Chinese is an SVO language with circumstantial adjuncts normally following subjects and preceding verbs.
Taking “given” and “new” information into consideration, scholars of Chinese agree that two issues determine the given/new dichotomy in Mandarin Chinese: one is the definiteness effects depending on syntactic function/position and the other is the array of anaphoric expression types in Chinese as well as restrictions on their (non-)use (Hole, 2012, p. 60). Chen (2004) studies the relationship between identifiability and definiteness of Chinese morphosyntactic constituents and the study becomes a very important supporting background of the present study.

Hole (2012, p. 62) claims that unlike English, where subject merely tends to be definite, subject DPs in Chinese must be interpreted as definite. So from this perspective, he proposes that Chinese links closer relationships between the discourse relation of topicality and the grammatical relation of subjecthood.

As an important means of focus expression, cleft constructions may establish the correlation between information structures such as topic and focus and their syntactic representations as well as their identifiability18 to the speaker. There are some conclusion already made in languages as English, such as the scale of the acceptability of the topic provided by different scholars (e.g. Lambrecht, 1994; Gundel, 1993); however, this rule is not testified either as a necessary or universal principle for all languages. In this study, the identifiability of referents of topic and focus constituents; the definiteness of their grammatical constituents; and the distributive feature of the information structure of Chinese cleft constructions are all discussed.

The previous studies of information structure of Chinese cleft constructions are to explore different means of focus delivery, and cleft constructions in Chinese are only one method among them. This study treats the Chinese cleft construction as the starting point

18Identifiability is the nature of the acceptability of a referent. Detailed information see Chapter 3.
in the first place, and then studies the inner distribution or configuration of the information structure of the construction from both the sentential and discourse perspectives.
Chapter 3

Theoretical Background

3.1 Information structure

Studies of information structure can be traced back to the 19th century. Other labels used by twentieth-century linguists are “information structure” or “theme” (Halliday, 1967), “information packaging” (Chafe 1976; Prince 1981a), and “informatics” (Vallduvi, 1990b).

3.1.1 Introduction to information structure

Prince (1981a) provides the concept of “information packaging” as the following:

(Information packaging is) the tailoring of an utterance by a sender to meet the particular assumed needs of the intended receiver. That is, information packaging in natural language reflects the sender’s hypotheses about the receiver’s assumptions and beliefs and strategies. (Prince 1981a, p. 224)

Several points may be taken care of in this definition, which together produce the whole process of information delivery.

Firstly, information structure bridges the addressee and the addressee.

Secondly, “hypotheses about the receiver’s assumptions” are hypotheses about the states of the mental representations of the referents or concepts of linguistic expressions in the mind of the receiver at the moment of utterance (c.f. Chafe 1976, p. 27).

Thirdly, there are strategies during the delivery of information. So it is crucial to notice here that the study of information structure is not concerned with the lexical and propositional content of the sentence constituent, but with the way such content is transmitted. The
psychological and mental thinking between addresser and addressee is transmitted by particular grammatical strategies.

The studies of information explore the mutual relationship between the interpretation of words or sentences in given conversational contexts and the discourse circumstances under which given pieces of propositional information are expressed.

Lambrecht proposes the definition of “information structure” in his book as:

INFORMATION STRUCTURE: That component of sentence grammar in which propositions as conceptual representations of states of affairs are paired with lexicogrammatical structures in accordance with the mental states of interlocutors who use and interpret these structures as units of information in given discourse contexts. (1994, p. 5)

Lambrecht’s theories of information structure are applied and discussed accompanying the Chinese empirical and introspective data in Chapter 6.

3.1.2 Important concepts of information structure

The most important concepts of information structure proposed by Lambrecht (1994) are:

(i) Presupposition and assertion, which have to do with the structuring of propositions into portions which a speaker assumes an addressee already knows or does not yet know, respectively.

(ii) Identifiability and activation, which have to do with a speaker’s assumptions about the statuses of the mental representations of discourse referents in the addressee’s mind at the time of an utterance; although identifiability and activation are independent cognitive categories, identifiability having to do with knowledge while
activation has to do with consciousness, the two correlate with each other in certain predictable ways. A referent being activated entails that it is identifiable. The detailed differences and categories of activation are discussed in the following sections.

(iii) Topic and focus, which have to do with a speaker’s assessment of the relative predictability vs. unpredictability of the relations between propositions and their elements in given discourse situations. (Lambrecht 1994, p. 6)

Two problems must be clarified here. Firstly, the “presupposition” mentioned here by Lambrecht (1994) is a pragmatic concept which has to be distinguished from the other type of presupposition.

(a) Semantic presupposition:

In Huang’s (2012) “The Oxford Dictionary of Pragmatics”, semantic presupposition is described as a relation between sentences or statements, and the semantic presupposition is much like an entailment.19

Similarly, Keenan (1972, p. 45) explains the semantic (in his words, logic) presupposition as: “A sentence S logically presupposes a sentence P just in case S logically implies P, and the negation of S also logically implies P.”

The semantic presupposition is dependent on the structure of an utterance. The truth value of the semantic presupposition is not influenced by the negation of the sentence or utterance. For example, A and B are discussing their mutual friend C. A said:

(3-1) A: “He knows that the capital of New Zealand is Auckland.”

---

19 Lyons (1977, p. 85) points out that entailment is “a relation that holds between P and Q where P and Q are variables standing for propositions such that if the truth of Q necessarily follows from the truth of P (and the falsity of Q necessarily follows from the falsity of P), then P entails Q.”
A’s presupposition “The capital of New Zealand is Auckland” is a presupposition failure for it is in fact false. This negation of the sentence “He doesn't know that the capital of New Zealand is Auckland” does not make A’s presupposition a different one. The presupposition is a semantic one here.

(b) Pragmatic presupposition:

Jackendoff (1972, p. 230) considers the pragmatic presupposition to be the information of one sentence that is assumed by the speaker to be shared by him/her and the hearer; and the “focus of a sentence” as the information in the sentence that is assumed by the speaker not to be shared by him/her and the hearer. Karttunen (1971) defines the pragmatic presupposition as follows: “To presuppose a sentence in the pragmatic sense is to take its truth for granted and to assume that the audience does the same” (also see Stalnaker, 1973 and Shanon, 1976).

Huang (2014) distinguishes the two types of presupposition in his book. According to him, the pragmatic presupposition refers to the given information which is “(the) part of the (relevant) context (or common ground), and in particular, the speaker’s commitment slate in which the sentence is uttered (Huang, 2014, p. 85)”.

In his study, he calls the cleft sentence as one of the presupposition triggers that engender presuppositions by their structural forms. To quote one example from Huang (2014):

(3–2) It was Baird who invented television.

>>\textsuperscript{20} Someone invented television.

In the present study, presuppositions triggered by cleft constructions in Chinese differ in each type of Chinese cleft construction. The classification of the Chinese cleft construction in

\textsuperscript{20}This symbol means a proposition “presupposes” the other proposition.
the present study is a focus-structure-oriented one, and in different types of cleft constructions, presuppositions have different configurations.

Keenan (1972) gives a series of example to illustrate that there are indeed two kinds of presuppositions (semantic and pragmatic presuppositions) in natural language. His examples for each kind of presupposition are rather different. All the examples for the logical/semantic presuppositions are drawn from English, however all the examples for the pragmatic presuppositions are drawn from exotic languages. This leads to the conclusion that pragmatic presuppositions require “certain culturally defined conditions” (Keenan, 1972, p. 46) and are context-driven.

Stalnaker (1973) distinguishes between two ways of treating presuppositions. From a semantic point of view being a presupposition is a relation between a sentence and a proposition; from a pragmatic point of view being a presupposition is a relation between the speaker and a proposition. Stalnaker argues in favor of the second analysis, which he finds more general and intuitive than the first. More specifically, he claims that “a proposition presupposed by a sentence in a technical semantic sense provides a reason for requiring it be presupposed in the pragmatic sense whenever the sentence is used” (p. 452). The reason why he takes pragmatic presupposition as the more intuitive form is that pragmatic presupposition explains the presupposition failure caused by the semantic presupposition. Presupposition failure occurs when the proposition assumed to be true is in fact false. In Stalnaker’s pragmatic account of presupposition, the example (3-1) is not problematic. He treats the presupposition failure as a non-catastrophic one21 (i.e., presupposition does not necessarily have to make sense semantically). The pragmatic presupposition of the sentence is that the

21 Failure is catastrophic if it prevents a thing from performing its primary task—in the semantic case making an evaluable claim. Non-catastrophic presupposition failure then becomes the phenomenon of a sentence still making an evaluable claim despite presupposing a falsehood.
knowledge “The capital of New Zealand is Wellington” is not in the common ground; if it is, the speaker would know this, so the speaker’s claim would be false. There is a close relationship between semantic and pragmatic presuppositions. An utterance of a sentence which has a semantic presupposition is associated with a pragmatic presupposition (Sudo, 2014).

Let us return to the important concepts of information structure. As mentioned at the beginning of this section, the important concepts of information structure proposed by Lambrecht (1994) are applied in this study. However, according to the previous literature, various terms are given to these concepts by different schools. The terminology for information dichotomy includes:

(I) psychological subject-psychological predicate (Paul, 1880; von der Gabelentz, 1869);

(II) theme-rheme (Ammann, 1928; Halliday, 1967b; Mathesius, 1929);

(III) topic-comment (Reinhart, 1981; von der Gabelentz, 1869);

(IV) topic-focus (Sgall & Hajicová & Benešová, 1973);

(V) presupposition-focus (Chomsky, 1971; Lambrecht, 1994; Jackendoff, 1972);

(VI) background-focus (Chafe, 1976; Jacobs, 1982);

(VII) old/ given-new information (Chafe, 1976; Halliday, 1967b);

(VIII) open proposition-focus (Prince, 1981);

(IX) notional subject-notional predicate (E. Kiss, 1995).
Regardless of the various terminologies, most of the researchers agree that the defining criteria of the partition are the contrastivity/contrastiveness and informativeness, and the terminologies just depart from different perspectives.

For example, Chafe (1976) pays attention to the givenness of constituents of a sentence. As a type of activation in the consciousness of the speaker and hearer, givenness initiates the discussion of a scale of givenness. Several givenness and activation hierarchies are proposed by scholars, such as the Scale of Assumed Familiarity (Prince, 1981), the Givenness Hierarchy (Gundel & Hedberg & Zacharski, 1993) and Lambrecht’s (1994) system of identifiability and activation. Lambrecht's system of identifiability and activation is based on Prince's scale. They categories the familiarity (in Prince, 1981) and unidentifiable (in Lambrecht, 1994) of the referents in the sentence according to their newness or givenness to both the communicators and the discourse. For example, the referent known by the speaker but not by the hearer is a brand-new one (in Prince, 1981) and inactive (in Lambrecht, 1994). Lambrecht's (1994) model is the basis of the analysis of the present study, which we will illustrate in detail later.

Gundel et al. (1993) proposed six cognitive statuses of referents that evoked by relevant determiners and pronouns. This scale relates the givenness and the grammatical constituent (i.e., demonstratives, pronouns, etc.) together directly. One of the tasks of the present study is to check the identifiability (i.e., “familiarity” in Prince (1981)) of focus and topic constituents of Chinese cleft constructions. The identifiability model of the present study is mostly based on Lambrecht's system of identifiability and activation.

Lambrecht (1994) states that partitions such as “theme-rheme”, “topic-comment” and “topic-focus” ignore one categorical problem, which is that the information is not made up by segmental sentence constituents. Under the above categories, sentences or propositions are
divided into theme & rheme; topic & comment; and topic & focus in a clear demarcated and binary way. He proposes the two concepts “presupposition” and “focus”, corresponding to the partitions mentioned above. Because the syntactic constituents do not align with distinctions in information structure, presupposition and focus are propositions instead of sentence constituents. Besides the two concepts (presupposition and focus), assertion and topic are also important concepts.

The difference between “meaning” and “information” is important for the understanding of the relationship between syntactic constituents and information structure. “Meaning” is a semantic concept expressed either by individual words or the relations between words, while information is expressed via propositions. So, the information of a proposition is realized by establishing the relationships between informative units instead of the semantic ones in the proposition. The guideline we resort to must be something information-oriented and not limited by the syntactic configuration or word order of the sentence because these methods are what we should be basing on. Lambrecht’s “Presupposition versus Focus” is employed as an important partition of information structure. In his study, he considers the “old information” to be the pragmatic presupposition, while the “new information” is the pragmatic assertion. He sees information in terms of propositions, and propositions must contain both a predicate and at least one argument. His ideas are summarized in general as the following:

(i) Both the presupposition and the new information are “propositional in nature” and the new information generally involves an assertion about a familiar referent, (i.e., the topic);

---

22 This statement does not deny the relationship between word-order and information structure. Word-order is a very important parameter in the domain of information structure. Here, I just want to emphasize that word-order is not the only criterion of information structure.
(ii) The topic is then excluded from the focus (the focal part of the assertion), because
(the relevance of) the topic is part of what is presupposed.

(iii) The focus is that part of the assertion that is not presupposed.

(iv) This also allows for information that is neither topical (not what the assertion is
about) nor focal to appear in an assertion: Assertion = [[[TOPIC] (other presupposed
material)] [FOCUS]]. The Topic is typically a syntactic constituent; the focus need
not be, but it is generally contained in a syntactic constituent that normally excludes
the topic. That constituent is called the “focal domain”.

In the present study, the above parameters (i.e., presupposition, assertion, focus, topic,
identifiability, and activation, whose definitions will be presented in the following sections)
are borrowed to explore the configuration of cleft constructions. To explain explicitly, firstly,
what is the layout of the presupposition, focus and assertion of cleft sentences and how is this
layout determined by the identifiability (accessibility) of constituents of the sentence
(especially the focus)? Secondly, for cleft constructions in Chinese, what are their types of
focus structures? Thirdly, is there any difference between cleft constructions and unmarked
Chinese sentences from the perspective of information flow on the issue of accessibility?

3.1.2.1 Presupposition
Strawson (1964) postulates the condition on which the transmission of information may be
realized successfully, named the “Principle of the Presumption of Ignorance”. This principle
is to guarantee that information will be delivered successfully only if the information the
speaker is trying to convey is not already stored in the hearer’s mind. However, considering
that the state of ignorance of a hearer is never totally new information added to a zero
foundation, this principle must be complemented by another principle, which Strawson calls
the “Principle of the Presupposition of Knowledge”.

77
There are two issues we have to state here, firstly, old information and new information cannot be equated with the old and new referents, and the information of a sentence is not “old information” plus “new information” in a clearly demarcated way. In Lambrecht’s terminological practice, he restricts the use of the terms “old information” and “new information” to aspects of information associated with propositions. To avoid the confusion, he replaces the “old” and “new” information with “presupposition” and “assertion”, which have been mentioned previously. The definitions of presupposition and assertion (shown in 3.1.2.2) are given in his book:

Pragmatic presupposition:\textsuperscript{23} the set of propositions lexicogrammatically evoked in a sentence which the speaker assumes the hearer already knows or is ready to take for granted at the time the sentence is uttered. (Lambrecht 1994, p. 52)

Lambrecht (1994) also points out two other types of presuppositions. The first one is the conscious-presupposition. The conscious-presupposition is the presupposition evoked by a sentence which concerns the assumed knowledge state of the addressee. It represents, to some extent, the addressee’s judgment or assumption about the state of consciousness or awareness of the addressee at the time of utterance. The other one is called relevance or topicality presupposition. It is the presupposition evoked by a sentence in which the assumptions a speaker has are concerned with the contextual relevance or topicality of a referent in the discourse; that is, the degree to which a referent can be taken to be a center of current interest with respect to which a proposition is interpreted as constituting relevant information.

\textsuperscript{23}For the sake of convenience, the term pragmatic presupposition is called “presupposition” hereafter.
For example, the pragmatic presuppositions lexicogrammatically evoked by the utterance “I finally met the woman who moved in downstairs” (1994, p. 51) may be loosely stated as the following set of propositions:

(1) The addressee can expect and accept the existence of the female individual designated by the definite noun phrase;

(2) Someone moved in downstairs from the speaker;

(3) One would have expected the speaker to have met that individual at some earlier point in time.

(4) The addressee is aware of the referents of the pronouns I and who at the time these pronouns are uttered.

(5) The proposition expressed by the sentence is construable as relevant information about the referent of I; the proposition expressed by the relative clause is construable as relevant information about the referent of who.

(Lambrecht 1994, p. 55)

The first presupposition is evoked by a grammatical morpheme, the definite article the; the second is evoked by a grammatical construction, the relative clause who moved in downstairs; and the third is evoked by a lexical item, the adverb finally. The fourth is the consciousness presupposition. The fifth is the relevance presupposition expressed via the two unaccented pronouns. The five presuppositions are all equally important, however, the last two have no relation to the formal structures. Because the present study is an analysis of information structure based on the syntactic structure of cleft constructions, the consciousness presupposition and the relevance presupposition are not included in the analysis of the present study.
3.1.2.2 Assertion

Together with the pragmatic presupposition, the definition of pragmatic assertion is postulated by Lambrecht (1994).

“Pragmatic assertion: the proposition expressed by a sentence which the hearer is expected to know or take for granted as a result of hearing the sentence uttered.” (1994, p. 52)

So, unlike the presupposition, the assertion is a proposition established after the utterance is finished but not before. It is established based on the presupposition and focus. The assertion is a combination of two sets of propositions—one is the set of the presupposed proposition established before the utterance of the sentence (i.e., presupposition); the other is the non-presupposed proposition which is added to, or superimposed on, the former (i.e., focus). We can formulate the relationship between presupposition, assertion and focus as: assertion = presupposition + focus.

Lambrecht (1994) states that because the speaker expects to know something “informative” by hearing a sentence, the assertion must differ from the presupposed set. It is possible, according to Lambrecht (1994), to have a null presupposition instead of a null assertion. For example, in the sentence-focus structures (see section 3.2.3) proposed by Lambrecht (1994), there is no presupposition, and the assertion coincides with the focus of the sentence.

3.1.2.3 Focus

Traditionally, the focus is considered to be the complement of the topic. All sentences convey new information, so all sentences must have a focus; however, not all sentences have a topic (Lambrecht 1994, p. 206). For this reason, it is not accurate to define “focus” as the complement of the “topic”. Just as a topic is included in the presupposition without being identical to it, a focus is part of an assertion without necessarily coinciding with it. Within the
framework developed by Lambrecht (1994), the focus of a sentence, or rather, the focus of the proposition is seen as the element of information whereby the presupposition and the assertion differ from each other. The focus is the portion of an assertion that cannot be taken for granted at the time of speech. It is the unpredictable or pragmatically non-recoverable element in an utterance. The focus makes an utterance into an assertion.

Previous definitions of “focus” are given by scholars like Bolinger (1954), Halliday (1967), Chomsky (1970) and Jackendoff (1972). Lambrecht’s concept of presupposition is pragmatics-oriented, thus making a subtle difference between the two “presuppositions”, as does the term “focus”. The “presupposition” in the Chomsky-Jackendoff tradition is in fact only one particular subtype of pragmatic presupposition, and the accent rules proposed by these authors are insufficient to account for the focus-presupposition relation in general, for to explain the focus and presupposition only from the perspective of prosody ignores the distinction between focus and activation. Just as Lambrecht (1994, p. 270) says: “The absence of pitch prominence on a clause or portion of a clause merely has the effect of marking a propositional denotatum as ACTIVE.”

Prosodic factors are important parameters for the delivery of information structure in both English and Chinese, which is the fact I do not deny. However, prosodic factors are not analyzed in detail in the present study for the following two reasons: firstly, cleft constructions are marked sentence forms which deliver special pragmatic functions through the special arrangement and configuration of the syntactic forms, so prosodic factors do not take as important a role as the syntactic method does. Actually, in spoken communication, according to the Principle of Least Effort, speakers prefer to realize the functions they plan

---

24 Zip (1949) firstly proposes the Principle of Least Effort (and the Principle of Sufficient Effort) which are followed by wide discussions. The principle is an overarching principle not only for languages, but also everything else in the human universe. Zip (1949, p. 20ff) acknowledged two basic and competing forces—the force of unification (i.e. Speaker’s Economy) and the Force of Diversification (i.e. Auditor’s Economy). Based on the Principle of Least Effort, many principles for different linguistic branches are proposed, for example the maxims of conversation (Grice, 1975).
to express by using prosodic methods instead of the syntactic ones, since prosodic prominence is much more “economical” than the syntactically unmarked configuration. So the functional structures as cleft constructions that we are looking for are based on written language. Secondly, and also because of the first reason, all the data collected in the present study are from the written corpus, so a prosodic analysis is not conducted.

Coming back to the defining of “focus”, we said that focus is not a complement of the topic, so the segmentation view does not fit in defining “focus”. Like topic, focus is a relational pragmatic category. Lambrecht (1994, p. 207) makes one example to illustrate the nature of “focus”:

(3-3) Q: Where did you go last night?
A: I went to the MOVIES.

The expression “the movies” is absolutely informative, being an element of the proposition made up by the entire sentence. However, the new information of the sentence is not the constituent “the movies” or its designatum, but “its role as the second argument of the predicate ‘go-to’ in the pragmatically presupposed open proposition ‘speaker went to x’” (Lambrecht 1994, p. 209-210).

Lambrecht gives his definition of focus as “the semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition (1994, p. 213)”.

To sum up, the descriptions of focus are displayed:

(i) The focus is an element of a pragmatically structured proposition which makes the utterance of the sentence into an expression of information.

(ii) It is the remaining part when all the presupposed components are taken from an assertion.

(iii) Sometimes, focus coincides with assertion, supposing that there is no presupposition evoked in the sentence.
(iv) The relation between the focus and the proposition is assumed to be unpredictable or non-recoverable for the addressee at the time of the utterance. The focus relation creates a new state of information (assertion) in the mind of the addressee.

(v) The focus of a proposition may be marked prosodically, morphologically, syntactically, or via a combination of the above methods.

3.1.2.4 Topic

The notion of topic provided by Lambrecht (1994) is different from that proposed by systemic functional linguists. In Lambrecht’s study, there is one important feature of the topic that distinguishes it from the views of linguists mentioned above—topic is no longer the “element which comes first in a sentence”, which is the idea proposed by systemic functional linguistics. In English, the sentence-initial position can be taken by either topic or focus, so it cannot be treated as one of the criteria for either of the two terms. Contrarily, the notion of topic/theme as the first element in the sentence is extensively discussed in Prague School research (cf. Firbas, 1966; Fries, 1983; Halliday, 1967).

Topic has sometimes been defined as a “scene-setting” expression, or as an element which sets “a spatial, temporal or individual framework within which the main predication holds” (Chafe, 1976). There are three properties of topic:

(i) It must have a referent;

(ii) The referent must be identifiable;

(iii) It has a certain degree of pragmatic salience in the discourse.

In his study, Lambrecht (1994) defines “topic” as “the thing which the proposition expressed by the sentence is about”. To define the topic in terms of the pragmatic concepts of amounts and relevance, Lambrecht (1994) illustrates his stand that to determine a topic by the syntactic structure of the sentence alone is not appropriate and accurate. In order to determine whether an entity is a topic in a sentence or not, it is often necessary to take into account the
discourse context in which the sentence is embedded. Which means “being a discourse referent” is a criterion of being a topic. Lambrecht (1994) proposes one example to illustrate the relationship between the topic and the discourse referent.

Speaker A has dialed a wrong number and is asking to speak to a person unknown to speaker B, who receives the phone call:

(3-4) A: Is Alice there?
   B: a. There is no Alice here.
      b. #Alice isn’t here.
      c. # She isn’t here.
      d. #No.

Even though the noun Alice is a referring expression, it cannot serve as a topic for B (i.e. the hearer) because it does not designate a specific discourse referent in B's mind. By the utterance of A, a presupposition about the existence of a woman called Alice is already evoked in the mind of hearer B. However, B cannot identify the specific Alice intended by A (i.e. the referent of Alice in not actively evoked by the hearer B). So the discourse referent of “Alice” cannot be a topic positioned initially in the sentence. The NP “Alice” must appear in postverbal focus position, as the subject of a thetic sentence.

3.1.2.5 The connections between these concepts

To illustrate the connection between the presupposition, assertion, and focus, we will take two examples of cleft constructions in both English and Chinese. They are of the same type (i.e. the argument-focus structure) from the perspective of focus structure (discussed in the following section).

(3-5) English: It is a watch that Susan gave Mike.

Presupposition: Susan gave Mike something(=X).
Focus: a watch

Assertion: X=a watch

(3-6) 是小明打破窗户的。

ShIaux Xiaoming dapo chuanghu DEpct.

ShIaux Xiaoming break window DEpct.

It is Xiaoming who broke the window.

Presupposition: 有人(=X)打破窗户。

Youren (=X) dapo chuanghu.

Someone break window.

There is someone (X) who broke the window.

Focus: 小明

Xiaoming

Assertion: X=小明

X=Xiaoming

The presupposition (if there is any) is always a proposition; sometimes the focus constituent is more than a referent (although not shown by the previous two examples), and assertion means to establish the relationship between presupposition and focus. The “x” in (3-5) and (3-6) is added by the hearer automatically after the sentence is uttered.
3.2 Types of focus structures and the configuration of the information structure

As Dik (1980, 1997) makes clear, languages not only express information packaging in different ways, they also express different focus meanings, or “focus types”. The division of focus types is a long-lasting subject being discussed by many scholars. Scholars (e.g. Halliday, 1967; Kiss, 1998; Rochemont, 1986; Selkirk, 2002 etc.) consider that there is a bipartition of the types of focus structures. The difference between the two types lies in the fact that one of the two types aims to provide the new information, and the other aims to establish the contrastive relationship by highlighting the focus. The distinction between the two types has been widely agreed in the linguistic literature for a long time (cf. Halliday, 1967; Rochemont, 1986); however, different terms are proposed by different scholars. For example, Kiss (1998) distinguishes the two types as identificational focus and information focus; Gundel’s partition is between the information focus and the contrastive focus. Selkirk (2002) classifies the contrastive focus and the presentational focus. In this study, my prerequisites are those—firstly, the acceptance of the distinction between the two types; and secondly, cleft constructions are treated as sentence structures which express contrastive focus. However, in order to conduct a detailed research on the relationship between the syntactic representation and the focus function of the sentence, I’d like to follow some semantic-syntactic based criteria of the focus types. So here, I borrow the criteria proposed by Lambrecht (1994). 25

Lambrecht claims the advantage of his criteria in the following way:

In combining the semantico-syntactic terms “predicate”, “argument”, and “sentence” with the pragmatic term “focus”, my intention is to capture the correlation between certain

---

25 Lambrecht’s criteria of the types of focus structure are an overall classification of the focus structures of all types of sentences, i.e. they are not for cleft constructions exclusively.
formal and semantic categories and certain types of communicative functions, such as the
d function of commenting on a given topic of conversation (predicate focus), of identifying
a referent (argument focus), or of reporting an event or presenting a new discourse referent
(sentence focus). There is thus a correlation between types of focus structure and type of
communicative situation. (1994, p. 222)

The three types of focus structure are predicate-focus structure, argument-focus structure
and sentence-focus structure. The communicative functions of the three types are: to
comment on a given topic of conversation (i.e. predicate focus); to identify a referent (i.e. 
argument focus); or to report an event or present a new discourse referent (i.e. sentence
focus) respectively.

3.2.1 Predicate-focus structure

The predicate-focus structure is also called the “topic-comment” structure in. Obviously,
the focus of the predicate-focus sentence is the predicate (i.e. normally, the sentence
constituents other than the subject). In the present study, “shi” copular sentences are
examples of this type. The information structure of the predicate-focus structure is shown by
this example from Lambrecht (1994, p. 226):

(3-7) Sentence: My car broke DOWN. 27

Presupposition: “speaker’s car is a topic for comment x”

Assertion: “x=broke down”

Focus: “broke down”

Focus domain: VP

26Predicate here is a semantic concept with its counterpart being the argument. It refers to the event or state the
sentence is dealing with; while arguments are the things that refer to the participants in the event/state.
Normally, predicates are expressed by verbs. In Chinese, there are nominal predicates and adjective predicates.
27 The capital words in the sentence (here and below) are the ones with prosodic prominence in the sentence to
present focus.
Focus domain is the smallest constituent that contains the focus, and in this sentence the focus domain is a VP. The function of the sentence is to provide the nonpresupposed information as a complement of the subject (i.e., something happened to the speaker’s car).

In the present study, “shi” copular sentences confuse us in judging the qualifying cleft constructions due to their similar superficial sentence structures. However, from the perspective of information structure, “shi” copular sentences are predicate-focus structures with the focus falling on the predicate (the whole predicate, not any part of it) of the sentence. The function of the “shi” copular sentence is to provide more information about the subject from the perspectives of identities, classifications, attributions, or other explanations. To take one example:

(3-8) 我是种田的。

Wo SHI\textsubscript{cop} zhongtian DE\textsubscript{nomi}.

I SHI\textsubscript{cop} farm DE\textsubscript{nomi}.

I am a farmer.

Presupposition: “我 (wo; I) is the topic of comment x” (in the following analysis, for the sake of convenience, such presupposition is expressed as “我 ( x ) ”).

“Shi” is included in the focus domain of the sentence which is a copula-predicate.

Focus: 是种田的

SHI\textsubscript{cop} zhongtian DE\textsubscript{nomi}.

SHI\textsubscript{cop} zhongtian DE\textsubscript{nomi}.

is a farmer.

Assertion: X=是种田的

X=SHI\textsubscript{cop} zhongtian DE\textsubscript{nomi}.

X is a farmer.
Focus domain: VP

The focus following the copula “shi” conveys no contrastive information. The focus, semantically, is the predicate of the sentence. Out of context, the predicate can be any constituent expressing attribution, classification, etc., and there is no contrastive focus expressed by the predicate-focus structure, so “shi” copular sentences in Chinese are predicate-focus structures with no contrastive focus expressed.

3.2.2 Argument-focus structure

In the argument-focus structure, the focus identifies the missing argument in a presupposed open proposition. The focal argument in an argument-focus structure is usually, but not necessarily, the subject of the sentence. The function of the argument-focus structure is to make the argument the focus by prosodic, syntactic or any other means. Unlike the other two types, the identifying function of the focus constituent in the argument-focus structure (i.e. argument), to some extent, leads to the exhaustivity of the focus.

(3-9) Sentence: My CAR broke down.

    Presupposition: “speaker’s X broke down”
    Assertion: “X=car”
    Focus: “car”
    Focus domain: NP

The sentence can be considered as a means of correction in the following scenario:

A: I heard that your motorcycle broke down.

B: No, my CAR broke down.

Lambrecht (2001) puts forward that cleft formation is a formal strategy for expressing argument-focus. So to put it in a cleft construction, the unmarked-form of argument (3-9) sentence can be changed into:

(3-10) It is my car that broke down.
Due to the particularities of the syntactic configuration of English cleft constructions, scholars (e.g. Lambrecht 2000) consider the argument-focus structure as the major focus group of English cleft constructions. However, by observing their Chinese counterparts, it is not difficult to find out that the argument-focus structure is not the only type of focus structure they express, instead, adjunct-focus structures are also accepted in Chinese cleft constructions.

3.2.3 Sentence-focus structure

There is one prominent characteristic of the sentence-focus structure—the pragmatic presupposition does not exist in its information structure. In other words, assertion and focus coincide in these structures. The function of this type of sentences is to give rise to the eventive interpretation of the proposition.

(3-11) Sentence: My CAR\textsuperscript{28} broke down.

Presupposition: ---

Assertion: “speaker’s car broke down”

Focus: “speaker’s car broke down”

Focus domain: S

The focus domain of this type is the whole sentence.

In the present study, some of the superficially similar “shi” sentences are the sentence-focus structure without any assignment of presupposition. They do not satisfy the criteria given in the present study of cleft constructions (see chapter 4), so they are excluded in the selection procedure.

\textsuperscript{28}Interestingly, scholars such as Lambrecht find that in both argument-focus structure and the sentence-focus structure in English, the prosodic prominence of the sentence all falls on the argument of the sentence. However, the unmarked position of prosodic prominence of an unmarked English sentence should be the final position in the sentence.
3.2.4 Types of focus structure of Chinese cleft constructions

I accept the criteria proposed by Lambrecht (1994), and the criteria go well with Chinese sentences in general. We can find the corresponding sentences for each type, and I even find they are convincing criteria by classifying the data in hand after the primary search of the present study. For Chinese cleft construction, the argument focus structure is truly one of the types, however other non-argument types do exist. But the semantic terms of “argument” and “predicate” are not enough for the types of focus in the present study, to take one example:

(3-12) 我是昨天到北京的。

Wo SHI<sub>aux</sub> zuotian dao Beijing DE<sub>ptcl</sub>.
I SHI<sub>aux</sub> yesterday arrive Beijing DE<sub>ptcl</sub>.
It was yesterday that I arrived in Beijing.

Presupposition: 我到北京。
Wo dao Beijing.
I arrive Beijing.
I arrived in Beijing.

Focus: 昨天 (zuotian; yesterday)

Focus domain: NP

The focus, as an adjunct of the sentence, is inserted into the presupposition “我到北京 (I arrived in Beijing)” which is a canonical sentence before the insertion. The focus “昨天 (zuotian; yesterday)” is neither the argument nor the predicate of the sentence. Being the smallest constituent that contains the focus, another type of focus domain has to be established for this kind of structure. Because the focus is an adjunct of the sentence, in this present study, structures like this are considered to be adjunct focus structures.

The types of focus structures in Chinese cleft constructions in actual use are investigated in the present study through the corpus-search and the subsequent analysis.
3.3 Identifiability and activation

While the pragmatic relation between propositions is established between “presupposition” and “focus”, cognitive statuses of the mental representations of discourse referents are discussed by “identifiability” or “activation”.

3.3.1 Identifiability

Identifiability is a cognitive category being firstly proposed by Chafe (1976). He uses the term “identifiable” to express the status of the designated referents in the mind of the speaker instead of “known” or “familiar”. An identifiable referent is one for which a shared representation already exists in the speaker’s and the hearer’s mind at the time of utterance, while an unidentifiable referent is one for which a representation exists only in the speaker’s mind. The relationship between the identifiable referent and the unidentifiable referent and the relationship between presupposed and asserted propositions are alike. A presupposed proposition is the proposition shared by both the speaker and the hearer by means of either internal or external knowledge. The difference between presupposed proposition and an identifiable referent is that the latter may not be presupposed to exist; instead, it has a certain representation in the mind of the hearer and the representation can be evoked in a given discourse. For example, in the sentence “The earthquake yesterday caused the big blackout”. The presupposed proposition is “There was an earthquake yesterday”, and the speaker uses the two identifiable referents (“the earthquake” and “the big blackout”) of the sentence to suggest to the hearer, via the form of the expression (definite article in this case), that the speaker assumes that the hearer has some mental representations of the individual and the event designated by that expressions.

How to judge whether a referent is identifiable or not:
(i) Some grammatical factors like articles in languages such as English are the most unmarked way of expressing the identifiability of a referent.

(ii) The presence of the referent in the context of discourse. If a referent can be traced back into the previous context, it is identifiable to the speaker but not necessarily to the hearer. In the previous example:

(3-4) A: Is Alice there?

B: a. There is no Alice here.
   b: #Alice isn’t here.
   c. # She isn’t here.
   d. #No.

For hearer B, the referent of “Alice” can be traced back to A. So it is identifiable for B, though it is not necessarily a topic. The identifiability of referents is categorized by different degrees, and the specific categories are illustrated in 3.3.2. The identifiability of the topic is varied in degrees in different sentences. Lambrecht (1994) proposes the Topic Acceptability Scale, summarizing the acceptability of the identifiability of the topic (see 3.3.4). Checking the identifiability of the topic in the Chinese cleft construction is one of the tasks of the present study.

(iii) Shared background knowledge

Chen (2004) explains that some referents can be inferred from other activities and entities in the discourse through logical reasoning due to general knowledge of the interrelationship between the entities or activities involved. This shared background information triggers the mental representation of other activities or entities automatically in the consciousness of the participants during the utterance. It is often captured by the theoretical constructs in cognitive sciences such as “frame”, “schema”, “script”, etc.
Based on these three criteria, there are complicated mechanisms in languages to show the identifiability of a constituent. The detailed methods may differ from language to language. By investigating the identifiability of the information constituents, especially the focus and the topic of the cleft construction, the study aims to show the trend of the identifiability of focus and topic constituents in Chinese cleft constructions.

### 3.3.2 Activation

With the idea “that our minds contain very large amounts of knowledge or information, and that only a very small amount of this information can be focused on, or be “active” at any one time,” (Chafe 1987, p.22ff), Chafe argues that a particular “concept” may be in any one of three activation states, which he calls active, semi-active (or accessible) and inactive respectively. An active concept is one “that is currently lit up, a concept in a person’s focus of consciousness at a particular moment”. An accessible/semi-active concept is one “that is in a person’s peripheral consciousness, a concept of which a person has a background awareness, but one that is not being directly focused on”. An inactive concept is one “that is currently in a person’s long-term memory, neither focally nor peripherally active” (Chafe 1987, p.25).

About the degree of accessibility, Huang (2007) explains the different degrees of accessibility by the long- or short-term memory of the hearer as well as the context in which the referents are involved. The entities that are retrieved from encyclopaedic knowledge are stored in long-term memory and are represented by low accessibility markers. Entities or activities which are present in the physical surroundings are in the short-term memory and therefore are encoded with intermediate accessibility markers. Entities or activities mentioned in the immediate preceding context are stored in the short-term memory, and could be perceived with highest accessibility.
There are two formal ways to express the active status of a referent: firstly, by a lack of pitch prominence; secondly by the pronominal coding of the relevant linguistic expression. As for the formal expression of the inactive status of a referent, the conditions are the opposite of those of active referents. The semi-activeness is much more complicated than the other two states, which is why it has aroused interest among scholars for so many years. The defining of the semi-active discourse referent seems trickier. According to Chafe, semi-active (accessible) referents can be of two kinds. A referent (“concept”) may become semi-active either by the deactivation from an earlier state, normally by having been active earlier in the previous discourse, or it can become semi-active because it is one member of the set of expectations associated with a “schema”. A schema is defined by Chafe as “a cluster of interrelated expectations (1987, p.29)”.

Chafe’s notion of a schema and its associated expectations, which he takes from cognitive psychology, are closely related to the Fillmorean notion of a semantic frame. Clark (1977) and Prince (1981a) also discuss the types of inferences that can cause a referent to be accessible.

Besides the two parameters, Lambrecht (1994) adds another kind of accessible referent. This type of referent is accessible due to its presence in the external world beyond the text. He presumes one scenario: sitting in an office with a friend, one might say “Those pictures sure are ugly” with reference to some photographs on the wall which the speaker assumes his/her addressee can easily access the referent of, although it may be that the hearer cannot do so.

Accessibility (semi-activeness) of a referent can thus be ascribed to three factors: deactivation from an earlier state, inference from a cognitive schema or frame, or presence in the text-external world. The first factor is called as “textually accessible” by Lambrecht; the
second is “inferentially accessible” and the third is “situationally accessible”. The two categories “textually accessible” and “situationally accessible” correspond to the text-internal and text-external world respectively, while the category “inferentially accessible” is neutral with respect to this distinction: a referent can be inferred from an element in the linguistic as well as in the extra-linguistic context.

From a strictly grammatical point of view, a binary distinction is justified and enough, which refers to the distinction between referents that are marked as being active (attenuated pronunciation and/or pronominal coding), and those which are not so marked. But this is not to say that the postulation of an intermediary category “accessible” has no grammatical reality. The difference between accessible and inactive referents can have syntactic consequences; in particular, it can influence the position of a constituent in the sentence or the choice of one rather than another grammatical construction. Different syntactic constraints on the coding of inactive and of accessible referents have been observed by Prince (1981a) and Chafe (1987), who both conclude that subjects of English are active and accessible referents instead of inactive ones. Semi-activeness does have grammatical realization in Chinese. Chen (2004) treats bare NPs and cardinality expressions as “indeterminate” lexical encodings. This type is neutral in respect of the interpretation of identifiability. So the identifiability of these expressions is highly decided by other syntactic factors, i.e. the word order of the sentence or the occurrence of their referents in previous discourse, etc.

The relationship between identifiability and activation and other terms involved are illustrated in the following diagram borrowed from Lambrecht (1994, p.109)
The unidentifiable referents are referents which are represented by the brand-new information and do not present in the mind of the hearer by any means. The status of unidentifiable is divided into two sub types: unanchored and anchored. The anchored referent is the item being part of the previous discourse or part of the situation. For example, the first example below delivers the unanchored unidentifiable referent and the second delivers an anchored one:

(3-13) I saw a cat in the park yesterday.

(3-14) A student I taught before went abroad recently.

The status of activation is divided into three types: inactive, accessible and active. The status of inactive is called “unused” by Prince (1981) which means the referent is a discourse-new but hearer-old one. For example:
Your mother just came.

The identifiable active status of a referent suggests that it is both a piece of discourse-old and hearer-old information which can be evoked by the hearer most easily. Grammatical methods such as pronouns are used to designate the activeness of the referent.

I bought a book yesterday. It is one of the best sellers this year.

The subtypes of the accessible category can be explained by the following examples:

(i) Textually-accessible:

I met one girl and one boy today. They are probably lovers, and the boy started to hit the girl suddenly, but the girl just stood there and did nothing to stop him.

The referent “the girl” is grammatically preceded by the definite article “the”, which means it is a grammatically identifiable concept. However, instead of being an identifiable active one, it is actually a textually accessible referent. A textually accessible expression requires an explicit antecedent “which has not been mentioned in the last two or three clauses and is thus only semi-active” (Baumann 2006, p.157).

(ii) Situationally-accessible:

Situational accessibility means that the referent can be accessed through extra-textual context. The extra-textual context, for example, could be the spatial-temporal conditions shared by both the speaker and the hearer.

Those pictures sure are ugly.

This example is quoted from Lambrecht (1994). When the speaker and the hearer are in the same place at the same time, without particular indication, the hearer would retrieve the referents the speaker refers to.

(iii) Inferentially-accessible:
An inferential accessible referring expression does not have explicit antecedents. They are (semi-)activated by bringing inference (cf. Clark, 1977) from another entity already present in the hearer’s discourse model.

(3-19) A: Are you ok?/ Why do you have a stomachache?/ What happened to you?

   B: I bought a combo from KFC, the ice Coke made my stomach ache.

The referent represented by the grammatical constituent “the ice Coke” (definite article+NP) does not appear in the previous discourse, and cannot be retrieved situationally. However, any one who is familiar with fast food combos such as those sold at KFC knows that ice Coke may be included in the combo, so the information is easily accessed by the hearer.

The investigation of the identifiability of the information structure (i.e., focus) of Chinese cleft constructions is based on the above criteria proposed by Lambrecht (1994). However, a modified version is proposed in the present study to adjust to the Chinese language based on these criteria. Some ideas from Chen (2004) are borrowed.

Chen (2004) summarizes that the uses of the English definite articles fall into four categories. Definite articles are to be situational, anaphoric, shared specific or general knowledge, and associative. By using these categories, Chen (2004) also tests the identifiability of referents in languages without articles, such as Chinese. Some of the categories are borrowed as types of accessibility in this present study. The detailed model is presented in Chapter 6.

3.3.3 Identifiability and definiteness

There is more than one way to realize the identifiability of a referent in a discourse. Identifiability is a cognitive concept, while definiteness is a grammatical one. Definiteness is one of the ways to realize identifiability. As Lambrecht (1994) puts it, “the grammatical
category of definiteness is a formal feature associated with nominal expressions which signals whether or not the referent of a phrase is assumed by the speaker to be identifiable to the addressee. (Lambrecht, 1994, p.79)".

In languages such as English, definiteness is regularly expressed by the grammatical category of articles. However, in languages without articles, definiteness is marked by other grammatical means, such as word order, case-marking particles, or numeral classifiers.

Chen (2004) makes elaborate studies of the definiteness of Chinese. He proposes that besides the identifiable referring expressions such as personal nouns, proper nouns, and definite noun phrases, there are some definiteness-inclined expressions in Chinese, however, definiteness is still on the path of grammaticalization.

As is discussed in more detail later, noun phrases (NPs) in Chinese fall into three major formal categories in terms of how the pragmatic property of identifiability is encoded: definite, indefinite, and indeterminate (cf. Chen 2004).

Although there is no articles in Chinese, demonstratives are developing the uses of a definite article, and yi “one+classifier” has developed the uses of an indefinite article and the indeterminate lexical codings include bare NPs and cardinality expressions (Chen, 2004). The identifiability of the indeterminate expressions is closely related to their positions in the sentence.

In the present study, an investigation is made based on the data to find out the identifiability of the focus as well as the topic constituents.

3.3.4 Identifiability/activation and the topic

Topics are relationally given. “The association of topics with definiteness across languages suggests that topics must be familiar (or at least uniquely identifiable)” (Gundel & Fretheim, 2006: p.181).

Lambrecht (1994, p.165) proposes the Topic Acceptability Scale as:
The acceptability can be, to some extent, tested by the effort consumed by the hearer in order to retrieve the referent. The cognitive efforts made to interpret the referents increase from “active” topic to “brand-new unanchored” topic. The active topic refers to the one with obvious markers such as unaccented pronominals. Accessible topics are the ones that can be retrieved by inferring other logical relationships, so much time is needed for the processing task.

3.3.5 Identifiability/activation and the focus

Unlike a topic, a focus constituent is in principle free with respect to the question of identifiability and activation. Although the focus is always treated as the “new” information of a sentence, what causes the “newness” of it is not the status of its denotatum in the discourse but the newness of its relation to the asserted proposition at the time of utterance. So it is the new relationship, instead of the new denotatum (or reference). So focus does not equal “inactiveness”; instead, the status of the identifiability of the focus may be checked according to individual circumstance with scrutiny.

Being a marked structure, the information structure of cleft constructions differs from unmarked ones. One of the differences is reflected on the identifiability of focus constituents. H.W. Cowles (2012) states that focus, particularly when encoded using clefts, appears to increase the identifiability of referents in memory.
In the present study, the discussion of the identifiability of the focus constituent is treated as an important task.

3.4 Frame semantics

Fillmore (1982) describes frame semantics as a research program in empirical semantics that emphasizes the continuities between language and experience and provides a framework for presenting the results of that research. The basic idea is that one cannot understand the meaning of a single word without accessing all the essential knowledge that relates to that word. Fillmore's most famous example of frames is the frame of “commercial transaction” that involves a seller, a buyer, the goods and the money. The four are core-elements of the frame; besides them, there are relations between the seller and the goods and the money; the relationship between the buyer and the goods, and the money and so on. They are the non-core elements.

3.4.1 Frame

The concept of the “frame” is derived originally from cognition. The cognitive frame is used by people to interpret their experience. Frame semantics is established on the basis of the cognitive theory of frame and aims to establish the relationship between human cognitive experiences and language expressions.

A semantic frame is a collection of facts that specify “characteristic features, attributes, and functions of a denotatum, and its characteristic interactions with things necessarily or typically associated with it” (K. Alan 2001, p. 251).

The notion of the frame is taken as “any system of concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits” (Fillmore 1982, p.111).
3.4.2 Introduction to FrameNet

FrameNet is a project hosted by the International Computer Science Institute in Berkeley, California. It is established on the theory of frame semantics. A semantic frame is a conceptual structure describing an event, relation, or object and the participants in it. Based on the sentence, the database includes the frames of the sentence, the frame elements the sentences consist of, and the lexical units of the sentence. The FrameNet lexical database contains around 1,200 semantic frames, 13,000 lexical units and over 190,000 example sentences. Fillmore makes a substantial contribution to the establishment and development of this project. To quote one example from FrameNet:

Verb: put

The verb “put” belongs to the frame of “Placing”.

The definition of the frame:

“Generally without overall (translational) motion, an Agent places a Theme at a location, the Goal, which is profiled. In this frame, the Theme is under the control of the Agent/Cause at the time of its arrival at the Goal.”

Example: David put his briefcase on the floor.

Frame elements:

Core: Agent: the Agent is the person (or other force) that causes the Theme to move.

    Cause [Cause]

    Goal: the Goal is the location where the Theme ends up.

    Theme: the Theme is the object that changes location during the Placing.

Non-core: Area: the Area is the setting into which the Theme is placed.

---

29https://framenet.icsi.berkeley.edu/fndrupal/framenet_search
Beneficiary: This extrathematic FE applies to participants that derive a benefit from the occurrence of the event specified by the target predicate. Further, the target predicate should involve some sort of Agent that intends that the benefit goes to the Beneficiary.

Cotheme: This is the second moving object, expressed as a direct object or an oblique.

Degree: The Degree indicates the degree to which the placing occurs.

Depictive: This FE identifies any Depictive phrase describing the Agent or Theme of the placing.

Distance: This FE is any expression which characterizes the extent of motion.

Duration: The Duration is the amount of time for which the Theme is to stay in the Goal location.

Explanation: The Explanation indicates why the Placing occurs.

Manner: Any expression which describes a property of the placing action which is not directly related to the trajectory of the Theme's motion expresses the frame element Manner. Descriptions of speed, steadiness, grace, means of motion, and other things count as Manner expressions.

Means: The Means is an act whereby the Agent achieves the placing of the Theme.

Path: Any description of a trajectory of motion which is neither a Source nor a Goal expresses the frame element Path. In this frame, Path expressions almost always have a via-sense.

Place: The Place indicates where the Placing occurs.

Result: This FE identifies the Result of the Placing.

Source: The Source is the initial location of the Theme before it changes location.

Speed: The Speed is the rate at which Placing occurs.
Time: Time is when the Placing occurs

As a counterpart of the English FrameNet project, Chinese FrameNet\(^{30}\) has also been established based on Fillmore’s frame semantics. By taking the Berkeley FrameNet project as the reference and using authentic texts from Chinese corpora, Chinese FrameNet is structured similarly to the English one.

The frames are used to retrieve the relationship between the frame elements and the focus and topic constituents in the adjunct-focus cleft constructions in Chinese.

### 3.4.3 Frame elements and identifiability/accessibility

Thematic relations can be featured by their different identifiabilities/accessibilities. Scholars (c.f. Aissen, 1997, 2001; Artstein, 1999; Baker, 1989; Fillmore, 1968; Grimshaw, 1990; Gruber, 1965; Jackendoff, 1972, 1987; Larson, 1988; MacWhinney, 1977; Speas, 1990) usually relate the identifiability of thematic relations to the structure of sentences. For example, De Smedt (1990) suggests that accessibility is an important factor in determining word order. In his account, accessible words are assigned to the sentence-initial position independently of their grammatical function. However, Ferreira (1994) cannot distinguish between the effect of role accessibility on grammatical function assignment and on word order, because the two are highly correlated in English, whose subject is usually in the sentence-initial position.

Following Chafe (1976) and Xu & Langendoen (1985), Huang (2007) considers Chinese-style topic constructions as “those whose comment clause is not syntactically but semantically and/or pragmatically related to topic” (2007, p.266). Although Chinese-style topic constructions are not the only type of topic constructions, they are widespread in Chinese and other East Asian and South East Asian languages (Huang 1994, p.160). In the

\(^{30}\)http://sccfn.sxu.edu.cn/portal-en/home.aspx
present study, I assume that, in a similar vein, in Chinese, the relationship between the information structure and the identifiability of its thematic relation is not influenced by the syntactic structure of the sentence as much as that of English. However, I do not deny the influence of the sentence structure on the identifiability of sentence constituents, especially the focus/topic constituent.

A referent’s accessibility refers to its activation in a language user’s mental model (Arnold, 2010). Many factors influence a referent’s accessibility. For example, syntactically, the subjecthood or the first mention of the referent is the most noticeable one; semantically, factors such as the semantic biases of events also influence the referent's accessibility.

In the present study, accessibility is also considered as a term eligible to the thematic relations. This is because most of the structures we will be looking for are not all argument-focus structures; instead, they are adjunct focus structures with the focus constituent being AP or PP, etc. Regardless of the syntactic classes of the focus constituents, the focus of the non-argument focus structure is still unidentifiable/identifiable when we treat them as thematic relations.

The detailed situation for the usage of thematic relations and their accessibility in Chinese cleft constructions are discussed in Chapter 6.

3.5 Summary

To establish the relationship between the information structure and the syntactic/semantic form as well as the mental representation of the information constituent, the theories of information structure and frame semantics are combined. Guided by these theories, the following problems are checked and discussed based on the data in chapter 6.

(i) The focus types of cleft constructions in Chinese according to their different focus domains.
(ii) The thematic relations of the topic and focus constituents.

(iii) The identifiability of the focus and topic constituents, their grammatical referents, and their thematic relations.
Chapter 4

Methodology of the Present Study

4.1 Introduction

This is a qualitative study paying attention to the analysis of the corpus-based data collected from the on-line Chinese corpus (i.e. CCL). This is a comprehensive corpus holding more than 500 million Chinese characters with no part-of-speech tagging. Due to the limitation of the auto-retrieval of corpus data, most of the work is done manually, and every qualifying sentence is annotated by myself. A clear illustration of the criteria of the manual searching becomes rather important.

In this chapter, the following tasks are fulfilled:

(i) To introduce the main method of the data collection, as well as the choice of corpora and genres.

(ii) To explain how the data are retrieved, and how many steps there are in the collection and selection procedures. The strings I used to search the corpus and the criteria I used to exclude the non-qualifying data are introduced. Examples are given to illustrate the operation.

(iii) To illustrate how the data is being used in the following two analysis.
4.2 Where is the data from?

The language discussed in this study is Modern Standard Chinese, which is the official language used in the People’s Republic of China and other regions such as Taiwan. It is considered as the standard form of oral communication and the modern written norm.

The present study aims to show some details about Chinese cleft constructions, especially their information structures, which means primarily the configuration and distribution of the information structure. To get a more comprehensive and accurate result, I prefer to conduct the research in the environment of the corpus. In this study, the on-line corpus CCL, which is established by the Center of Chinese Linguistics of Peking University, is used. Besides the searched data from the corpus, some introspective examples are also used to explain some phenomena. The main function of the corpus is to provide the empirical data. The previous literature mainly uses introspective evidence to support the view about cleft constructions in Chinese. This study resorts to the corpus to select data from the natural use of the language; and more importantly, the real context is involved in the analysis as well.

The corpus contains a large volume of texts written in Chinese characters in both Ancient and Modern Chinese. It consists of different genres, such as journals, literature, transcripts of TV shows, plays, etc. Due to the large capacity of the full corpus, and considering the manual procedure of the data selection, I do not take the whole CCL corpus as the searching basis, instead part of the corpus is chosen.

Three principles are followed on the issue of how to use the corpus. Firstly, the volume of the corpus has to be large enough to include the possible data as much as it

---

31 Modern Standard Chinese is called “Putonghua” (common language) in mainland China and “Guoyu” (national language) in the Taiwan area.

32 “Possible data” here is different from the “actual data”. The former refers to all the data that includes the “actual” cleft sentences and that are actually not but still hold the similar syntactic structures. There are two
can, but small enough to ease the selecting procedure as much as possible. There are huge numbers of sentences with similar structured strings, so it is not practical to manually collect the data one by one among the possible data. Therefore parts of the corpus are separated as qualified sub-corpora for the secondary selection procedure. One thing has to be clarified in particular here—because the primary search is automated, and results from the primary search are not the final data needed in the present study, the scale of the corpora is not controlled. All the data collected from the primary search have to be sorted manually in the secondary selection and then be generated as the result of the corpus searching procedure, so the sets encompassed by the data collected from the primary search are better controlled regarding capacity. Secondly, the sub-corpora are not the result of mere random separation of the original corpora; instead, they are the sub-corpora belonging to two different styles of texts. There is no relevant literature discussing the different distribution of Chinese cleft constructions in different genres or registers, so it is not appropriate to pay all the attention to a single genre.

In the present study, instead of subcategorizing genres into detailed types, I prefer to categorize texts in the corpus into two styles: formal and informal. For informal ones, I choose the data from modern plays and TV talk shows to be one sub-corpus; while for the formal ones, I choose essays such as the legal paper works and social science essays.

Thirdly, to get a balanced result, the similar scale of the sets chosen from the results of the primary search is necessary. In the present study, after the primary search, 5000 sentences\textsuperscript{33} for each corpus are chosen to form the sets only for the secondary selection.

\textsuperscript{33} The 5000 sentences are the first 5000 ones chosen from the searching results of the primary auto-search.
To sum up, two sub-corpora are selected according to their different styles. For the sake of convenience, I name them FORMAL corpus and INFORMAL corpus respectively. The frequencies of the three different strings which will be explained in 4.3.1 are shown in numbers, although the numbers are not the direct evidences for the result of the analysis of the information structure.

4.3 Data retrieval

The data-collection procedure of the present study consists of two parts—the primary search procedure and the secondary selection procedure. The primary search is the preliminary procedure aiming to automatically retrieve data holding structures of the searching strings by the online corpus itself. The secondary selection means to exclude non-qualifying ones in accordance with the criteria of Chinese cleft constructions used in the present study. This part of the search cannot be finished in the auto searching procedure by simply inputting syntactic searching strings, so the secondary selection procedure has to be done manually.

4.3.1 Primary search

The main goal of the primary search is to collect as much as possible in accordance with some strings defined syntactically in the two sub-corpora. In the primary search, the following strings are the indexes used in the corpus searching procedure.

(1) 是$15 的 (shi $15 de)

(2) 是$15。34 (shi $15。)

34 The type (2) string represents both structures of “initial bare shi” and “medial bare shi”; while the type (3) string represents both “final bare de” and “…de O” structures.
The wildcard “$15” is the searching language of the CCL corpus representing that there are 15 characters between the two characters “是” and “的”. The reason why I chose the length of 15 characters is based on some relevant surveys about the length of Chinese simple sentences.

The length of Chinese sentences is studied mostly by scholars from computational linguistics. The aim of their study is to innovate software or corpus tools such as corpus parsing. The following two tables suggest that firstly, generally speaking, sentences with the length of ten Chinese characters overwhelmingly exist in communications; secondly, different genres have different distributions of simple and complex sentences. For example, the length of simple sentences in literature is about 12 characters long.

| Table 4-1 The distribution of sentence length by Zong et al. (1999) |
|-------------------|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Length (characters) | 1      | 2         | 3         | 4         | 5         | 6         | 7         | 8         | 9         | 10        | 11-67     |
| Percentage (%)      | 15.12  | 8.34      | 9.28      | 8.54      | 7.68      | 6.78      | 5.27      | 5.27      | 4.78      | 4.09      | 24.85     |

| Table 4-2 Distributions of Chinese sentences in different genres by Li et al. (2014) |
|-------------------|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|

35The punctuation “。 ” is the full stop of Chinese, and the marker “$” refers to the distance between the character preceding and following it. For example, type (1) is read as: there are 15 characters between “是” and “的”. The second is read as: there are 15 characters between “是” and the end of the sentence. The third type goes as there are 15 characters between the end of the previous sentence and the particle “的” of the present study.
<table>
<thead>
<tr>
<th>Genres</th>
<th>Number of passages</th>
<th>Percentage of the genre</th>
<th>Number of sentences (simple/complex)</th>
<th>Percentage of complex sentences in the passage</th>
<th>Average length of sentences (simple sentence/complex sentence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature</td>
<td>225</td>
<td>37.25</td>
<td>24799 (10614/14185)</td>
<td>57.2</td>
<td>21.4 (12.0/28.4)</td>
</tr>
<tr>
<td>News</td>
<td>156</td>
<td>25.83</td>
<td>6773 (2882/3951)</td>
<td>58.3</td>
<td>25.0 (14.3/32.7)</td>
</tr>
<tr>
<td>Academic Writing</td>
<td>28</td>
<td>4.64</td>
<td>9395 (4387/5008)</td>
<td>53.3</td>
<td>28.1 (18.0/3.6)</td>
</tr>
<tr>
<td>Essays</td>
<td>195</td>
<td>32.28</td>
<td>3169 (1869/1300)</td>
<td>41.0</td>
<td>21.0 (12.9/32.6)</td>
</tr>
<tr>
<td>Total</td>
<td>604</td>
<td>100</td>
<td>44136 (19692/24444)</td>
<td>55.4</td>
<td>23 (13.73/31.1)</td>
</tr>
</tbody>
</table>

Regardless of genre differences, the average length of the Chinese simple sentence ranks from 12 to 18 characters, so I set 15 characters as an appropriate distance between “shi” and “de” (or between “shi”/“de” and the end of the sentence).

The results of the primary search include all sentences with the strings. There are huge numbers of sentences searched by the strings in the primary search procedure. To cite three examples as results of the primary search:
Figure 4-1. Some of the results of “shi $15 de” constructions in the primary search.
Figure 4-2. Some of the results of “shi $ 15.” constructions in the primary search.
Figure 4-3. Some of the results of “$15 de” constructions in the primary search.

I do not take all of them into the second selection procedure, instead, I select the first 5000 sentences from each corpus to be the sets of the secondary selection.

4.3.2 The secondary selection procedure

The aim of the secondary selection procedure is to clarify from these data collected in the primary search and find out which are real cleft constructions and which are not.

The criteria used in the secondary selection procedure are for the exclusion of non-qualifying data from the primary search.
Based on the functional and semantic observation of the data collected from the primary search, the criteria applied to the secondary selection procedure are:

(I) There is a bipartition of “presupposition vs. focus” in the sentence. The presupposition, under most circumstances, is a proposition. And the focus has to be a constituent inserted into the “presupposition”. The removal of the “focus constituent” does not make the sentence an ungrammatical or an incomplete one syntactically and semantically. The boundary between the “presupposition” and “focus” is not a clearly demarcated one. “Shi” and “de” cannot obligatorily be included into either “presupposition” or “focus”.

(II) The focus should be exhaustive, which means if the focus consists of more than one constituent linked by the conjunction as “也 (ye; also)”, the sentence is not a cleft construction.

(III) The focus should be contrastive.

Any sentence not meeting any of the above four properties is excluded from the set of cleft constructions. In the next section, the three criteria are explained and how the non-qualifying sentences were excluded is discussed with examples.

4.4 The three criteria

How these three criteria become the defining features of cleft constructions in Chinese is discussed in this section. Examples are given to illustrate the differences between non-qualifying sentences and cleft constructions from the three aspects.

---

36Exhaustiveness of cleft constructions is a very important concept, a brief introduction has already been introduced in chapter 2 (P14-15) and how to test the exhaustiveness of a sentence will be discussed in Chapter 5.
4.4.1 The distribution of “presupposition” and “focus”

Most sentences have a focus and a presupposition; however, the distribution of presupposition and focus varies in different types of sentences. Considering the distribution of “presupposition” and “focus”, it is necessary to talk about the types of focus structures proposed by Lambrecht (1994) first.

According to Lambrecht (1994), focus types are divided into three kinds: predicate focus structure, argument focus structure and sentence focus structure due to the different focus domains. Let’s start with the predicate focus structure in Chinese:

(I) Predicate focus structure:

(4-1) A: 你的脚怎么了?

Ni DEposs jiao zenme LE38?

You DEposs foot how LE?

What happened to your foot?

B: 我的脚扭了。

Wo DEposs jiao niu LE.

My foot was sprained.

Presupposition: “我的脚 (x)”

My foot (x).

---

37 By focus domain is meant the syntactic domain in a sentence which expresses the focus component of the pragmatically structured proposition. (Lambrecht, 1994, p. 214)
38 “Le” is an aspectual marker of Chinese expressing the perfect aspect of the verb.
Focus: 扭了

was sprained.

Assertion: x=扭了。

x= sprained.

Focus domain: VP

The relationship between the presupposition and focus in a predicate focus structure is that of “topic-comment”. In this sentence, the presupposition coincides with the topic, and the presupposition and focus are clearly partitioned into two parts. The function of the sentence is to provide new information about the topic. If we represent presupposition as “P” and focus as “F”, then the whole proposition of the sentence is configured as “P+F”.

(II) Argument focus structure:

(4-2) A: 听老李说你的脖子扭到了。

Ting Lao Li shuo ni DEposs bozi niu dao LE.

Hear Lao Li say your neck sprain.

I heard from Lao Li that you hurt your neck.

B: 不，是我脚踝扭到了。

Bu, SHI wo jiaohuai niu dao LE.

No, SHI my ankle hurt LE.

No, I hurt my ankle.
Presupposition: “我 (x) 扭到了”(Wo niu dao le; I sprained.)

Assertion： “x=脚踝”(x=jiaohuai; ankle)

Focus： 脚踝（jiaohuai; ankle)

Focus domain: NP

Again, if by representing presupposition as “P” and focus as “F”, we can formulate the proposition as “P_i+F+P_i”. 39 The two “P”s do not mean that there are two presuppositions; instead, they mean that the focus is inserted into the middle of the presupposition, and the presupposition is already a completed proposition before the insertion (i.e. without the insertion of the focus, the presupposition itself is still a meaningful proposition). The configuration of this type represents a typical configuration of the information structure of cleft constructions in both English and Chinese.

(III) Sentence-focus structure:

(4-3) A: 你今天怎么迟到了？

Ni jintian zenme chidao LE.

You today how late LE?

Why are you late today?

B: 我脚扭了。

39 The footnote “i” indicates that the presuppositions (shortened as “P”) are the same and continuous ones rather than two different and separated presuppositions.
Wo jiao niu LE.
My ankle sprain LE.
My ankle was sprained.

Presupposition: --

Assertion: 我脚扭了。 (wo jiao niu le; my foot sprained.)

Focus：我脚扭了。 (wo jiao niu le; my foot sprained.)

Focus domain：sentence

A sentence focus structure is called an “event-reporting” structure by Lambrecht (1994). The main function is to report an event by providing the new information based on no “presupposition”, while the focus extends over both the subject and the predicate. So from the perspective of the partition, a sentence-focus structure does not show the bipartition between “presupposition” and “focus”, because the prerequisite of the bipartition does not exist (i.e. there is no presupposition). Similarly, the “shi” emphatic sentence, which emphasizes the truth-value or the modality of the whole sentence, is a type of sentence-focus structure, which means it also lacks presupposition. This is the other difference between “shi” emphatic sentences and cleft sentences in Chinese.

(4-4) 这的确是我的问题，我是该为这事儿负责任的。

Zhe dique SHI\textsubscript{cop} wo DE\textsubscript{poss} wenti, wo SHI\textsubscript{aux} gai wei zhejiashier fuzheren DE\textsubscript{ptcl}.

This indeed SHI\textsubscript{cop} my problem, I SHI\textsubscript{aux}should take this responsibility DE\textsubscript{ptcl}.

This is my problem, indeed, I should take responsibility for this.
In “shi” emphatic sentences, the proposition (formed by deleting “shi” and “de”) is emphasized from the perspective of either its truth-value or its modality. In the sentence (4-4), the insertion of the auxiliary "shi" confirms the subjective opinion of the speaker towards the proposition that "I am the person who should be responsible for this problem". The focus is the whole proposition "我该为这事儿负责 (I should take responsibility for this)", and there is no presupposition. However, in cleft constructions, there must be a presupposition that makes the propositional limitation for the exhaustive focus.

One thing has to be clarified in particular here—Lambrecht’s classifications of focus structures are borrowed to explain the configurations of presupposition and focus from a general perspective. However, due to the language differences, the detailed focus types of Chinese cleft constructions are classified according to the focus domains of data collected in the present study. The detailed discussion is conducted in Chapter 6.

Despite the different types of focus structures, what I want to emphasize here is that the configuration of “presupposition” and “focus” in the argument-focus structure is typical of the configuration of presupposition and focus in both English and Chinese. In cleft constructions, the presupposition is a completed proposition, and the focus is the inserted (new, but not necessarily new) information that provides detailed information, and the deletion of the focus does not make the sentence an incompleted one. To be specific, the properties of the configuration of the presupposition and the focus of cleft constructions in Chinese are shown by the following aspects.

4.4.1.1 The boundary between the presupposition and the focus is not a clearly demarcated bipartition

A predicate-focus sentence expresses a “topic-comment” relationship between the subject and the predicate and, in Chinese copular sentences, the presupposition very often coincides
with the topic of the sentence, while the focus represents the comment. So the boundary
between the presupposition and focus in a predicate-focus sentence is a clear-cut one, and the
presupposition itself does not form a completed proposition. To take one example of a
copular “shi” sentence:

(4-5) 他是小学生。

Ta $SHI_{cop}$ xiaoxuesheng.

He $SHI_{cop}$ pupil.

He is a pupil.

Presupposition: 他(x) (ta; he)

Focus: 是小学生 (shi xiaoxuesheng; is a pupil)

Assertion: x=是小学生。 (x=$SHI_{cop}$ xiaoxuesheng; x= a pupil)

The assertion is “the establishment of an aboutness relation between the topic referent and
the event denoted by the predicate.” (Lambrecht, 1994, p.226). The topic as well as the
presupposition of the sentence is the subject “他 (ta; he)” and the focus is the copular
predicate “是小学生 ($SHI_{cop}$ xiaoxuesheng; is a pupil)” with the copula being part of the
verbal predicate. The assertion shows the relationship established between the presupposition
and the focus. To show the relationship, the following table is used:

**Table 4-3 Information structure of sentence (4-5)**

<table>
<thead>
<tr>
<th>Predicate-focus sentence</th>
<th>sentence</th>
<th>他</th>
<th>是小学生</th>
</tr>
</thead>
<tbody>
<tr>
<td>He</td>
<td>$SHI_{cop}$ a pupil.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The boundary between presupposition and focus is clearly demarcated. Let’s have a look at the configuration of presupposition and focus in the cleft construction.

(4-6) 我是坐车到北京的。

Wo $SHI_{aux}$ zuoche dao Beijing $DE_{ptcl}$.

I $SHI_{aux}$ by car arrive Beijing $DE_{ptcl}$.

I arrived in Beijing by car.

Presupposition: 我（x）到北京。（wo dao beijing; I arrived in Beijing.）

Focus: 坐车（zuoché; by car）

Assertion: x=坐车(x=by car)

Table 4-4 Information structure of sentence (4-6)

<table>
<thead>
<tr>
<th>Adjunct-focus sentence</th>
<th>sentence</th>
<th>我</th>
<th>是</th>
<th>坐车</th>
<th>到北京</th>
<th>的。</th>
<th>$DE_{ptcl}$</th>
<th>Information structure</th>
<th>Presupposition</th>
<th>focus</th>
<th>Presupposition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>$SHI_{aux}$</td>
<td>By car</td>
<td>arrive Beijing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

124
The presupposition of the adjunct-focus\textsuperscript{40} sentence is a completed proposition, and the focus is the information inserted into the proposition, and the boundary between the presupposition and the focus is not a clearly demarcated one.

Besides adjunct-focus cleft constructions, the argument-focus cleft construction is the other type of focus structure and the configuration of the presupposition and the focus of argument-focus cleft sentence differ from that of the adjunct-focus cleft construction.

\begin{verbatim}
(4-7) 是我打破窗户的。

SHI\textsubscript{aux} wo dapo chuanghu DE\textsubscript{ptcl}.

SHI\textsubscript{aux} me break window DE\textsubscript{ptcl}.

It was me who broke the window.

Presupposition: \((x) \text{打破了。} ((x) \text{ da po boli le; broke the window})

Focus: 我 (wo; me)

Assertion: x=我 (wo; me)
\end{verbatim}

<table>
<thead>
<tr>
<th>Argument-focus cleft construction</th>
<th>Sentence</th>
<th>是</th>
<th>我</th>
<th>打破</th>
<th>的。</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SHI\textsubscript{aux}</td>
<td>I</td>
<td></td>
<td>window</td>
<td>DE\textsubscript{ptcl}</td>
</tr>
</tbody>
</table>

\begin{table}[h]
\centering
\caption{Information structure of sentence (4-7).}
\end{table}

\textsuperscript{40}The adjunct-focus cleft construction is the construction with the adjunct constituent as the focus domain.
<table>
<thead>
<tr>
<th>Assertion</th>
<th>Presupposition</th>
<th>Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>打破窗户</td>
<td>我 (wo; me)</td>
</tr>
<tr>
<td>There is somebody</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The presupposition of the sentence is actually a presupposition of the existing fact that “there is someone who broke the window”.

The presupposition of the argument-focus structure is much like an existential presupposition proposed by Huang (2014). For example, the presupposition is not “打破窗户 (dapo chuanghu; broke the window)” instead of the proposition that “there is someone who broke the window”, and by providing the focus as “我 (wo; me)”, the assertion is formed as “there is someone who broke the window, and that person is me”. The predicate “打破窗户 (dapo chuanghu; broke the window)” is only part of the presupposition.

4.4.1.2 “Shi” and “de” are not obligatorily included in either “presupposition” or “focus” in cleft constructions

“Shi” is an auxiliary verb of the sentence and “de” is the particle of the sentence, and they are only functional words that do not dependently adhere to any constituents, so they
do not have to be involved in the presupposition and focus as the copular sentence does.

For example:

(4-8) 我是昨天到北京的。

Wo SHIaux zuotian dao Beijing DEptcl.

I SHIaux yesterday arrive Beijing DEptcl.

It is yesterday that I arrived in Beijing.

<table>
<thead>
<tr>
<th>我</th>
<th>是</th>
<th>昨天</th>
<th>到北京</th>
<th>的</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>SHIaux</td>
<td>yesterday</td>
<td>arrive Beijing</td>
<td>DEptcl</td>
</tr>
<tr>
<td>presupposition</td>
<td>focus</td>
<td>presupposition</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4-6 Information structure of sentence (4-8)

This criterion distinguishes cleft constructions from “shi” copular sentences in which “shi” acting as the copular verb and “de” being the nominalizer, relative pronoun, or the modifier marker.

(4-9) 我是种田的。

Wo SHIcop zhongtian DEnomi.

I SHIcop farm DEnomi.

I am a farmer.

(4-10) 花是红色的。

Wo SHIcop hua shi hongse de.
Hua $\text{SHI}_{\text{cop}}$ hongse $\text{DE}_{\text{modi}}$.

Flower $\text{SHI}_{\text{cop}}$ red $\text{DE}_{\text{modi}}$.

The flower is red.

In both examples, the presuppositions are the topics of the sentence—“我 (wo)” and “花 (hua)” respectively, and the copular phrases “是种田的 (shi zhongtian de)” and “是红色的 (shi hongse de)” are the verbal predicates of the subjects. The copula and the nominalizer/modifier “de” are involved in the focus of the sentence.

As shown by the examples in the last section, “shi” and “de” are not necessarily involved in either the presupposition or the focus.

However, this sub-criterion does not distinguish cleft constructions from emphatic sentences:

(4-11) 我是该为这件事负责的。

$\text{Wo} \text{SHI}_{\text{aux}} \text{gai wei zhejiashi fuze} \text{DE}_{\text{ptcl}}$.

$I \text{SHI}_{\text{aux}} \text{for this issue take responsibility} \text{DE}_{\text{ptcl}}$.

I sure should take responsibility for this issue.

In emphatic sentences, there is no presupposition and the focus is the whole proposition. “Shi” and “de” are not necessarily involved in “focus” or “presupposition”, just as in the case of cleft constructions.

**Table 4-7 Information structure of sentence (4-11)**

| 我 | 是 | 该为这件事负责 | 的。 |

128
I

SHI$_{aux}$

Should take responsibility for this issue

DE$_{ptcl}$

focus

focus

The word classes of “shi” and “de” in “shi” emphatic sentences are alike those in cleft constructions, and neither “shi” nor “de” is involved in the focus constituent.

Being an auxiliary focus marker of cleft constructions, the main function of “shi” is to put the following constituent in the focused position; and being an ending modal particle, the functions of “de” may be enhancing the modality of the sentence or expressing the tense/aspect property of the action involved in the sentence under special context. “Shi” and “de” are not necessarily the components of either presupposition or focus. If they are, they cannot be omitted from the sentence. Such as in the example “我是种田的 (wo shi zhongtian de; I am a farmer)”, “de” is the nominalizer of the verb phrase ahead of it. Compared with the obligation of “de”, in the copular sentence, “shi” is not bound to the focus constituent as rigidly as “de” is, because nominal predicates are accepted in Chinese copular sentences.

Nominal predicates were brought to light as early by Chao (1968), Zhu (1982), and many others. Chinese linguists (Chao, 1958; Lü, 1980; Tang, 1979; Zhu, 1982, et al.) have noticed that some nominals can be taken as predicates via an omission of the copular verb, “shi”. Tang (1998) claims that the parametrization of categorical features among languages
such as English, Chinese and Japanese is reflected by the “bareness” discrepancy between Chinese small clauses and English small clauses. Unlike Chao (1958) and Hashimoto (1959), who claim that the nominal predicate sentence is derived from the omission of the copulative verb "shi", Tang (1998) claims that the combined feature of nouns from two categories, substantive [N] and predicative [V], is subject to parametric variation. For instance, Chinese nouns contain a primary categorical feature, [N], along with a secondary primary categorical feature, [V]. Thus, Chinese nominals can be predicative in the bare SC. If the copula is omitted, the constituent “zhongtian de” is considered as the nominal predicate, but if not, the predicate of the sentence is still the verbal predicate with the structure of “copula+ nominalized phrase”. To get a unified configurative model of “presupposition” and “focus”, “shi” is not deleted from the focus constituent of “shi” copular sentences, which means the focus domain of the copular sentence is VP, headed by the copula “shi”.

(4-12) 我种田的。

Wo zhongtian DE_{nomi}.

I farm DE_{nomi}.

I am a farmer.

(4-13) 桌上吃的。

41 Stowell (1989), Longobardi (1994), Szabolcsi (1994), Tang (2001b), and Li (1998, 1999) agree that bare-NP can be a nominal predicate. However, examples are found against this agreement:

?他是人 (ta shi ren./ he shi person/ he is a human being).

42 Bare SC (small clause): bare SC means there is no functional projection inside the construction. Tang (1998) claims that unlike English SC, the Chinese SC is structurally “bare” which means there is only the lexical projection in the structure.
This discussion illustrates that normally, neither “shi” nor “de” is obligatorily included in the presupposition or the focus in cleft constructions; the same situation happens to “shi” emphatic sentences; in copular sentences, “shi” and “de” (especially the modification “de”) has to be included in the focus constituent.

4.4.2 The exhaustivity of the focus

The feature of exhaustivity distinguishes cleft constructions from nonqualifying sentences (i.e. “shi” emphatic sentences, copular sentences, etc.), so it is the most crucial criterion.

4.4.2.1 What is exhaustivity?

Büring and Kriz (2013, p. 1) define the “exhaustivity claim” as: “A cleft of the form It is x that p not only expresses that x has property p, but also that x is the only individual to have p, i.e. that x exhaustively identifies p (in the relevant contextual domain)”. 

Kiss (1998) proposes two functional properties of cleft constructions—exhaustiveness and exclusiveness. The first property shows that the constituent under contrastive focus in a sentence is a complete list of the entities that make the truth value of the proposition true. The second property claims that those and only those entities under contrastive focus will make the truth value of the proposition true, excluding other entities. These two terms express a concept similar to Büring and Kriz’s term “exhaustivity”. For the sake of convenience, I use the term “exhaustivity” in the present study.

To cite one example from Li (2008, p.750) for further explanation:
(4-14) It’s [John] FOC that stole the cookie.

The function of the sentence is to stress that no one else besides John stole the cookie.

Hedberg (1990, 2000) claims that the exhaustive semantics is compatible with the exclusive focus adverb “only”, but not with the additive focus particle like “also” or scalar focus particles like “even”. Just as Horn (1969, p. 106) says “clefting, like ‘only’, specifies uniqueness, while ‘even’ and ‘also’ presuppose non-uniqueness and thus cannot be clefted”.

e.g. (4-15) It’s only Muriel who voted for Hubert.

(4-16) *It’s also Muriel who voted for Hubert.

(4-17) *It’s even Muriel who voted for Hubert.

The compositional semantic implementation of Chinese clefts derives the exhaustivity associated with this pattern from a presuppositional uniqueness condition on events. This renders the exhaustivity tied to Chinese clefts maximally similar to the uniqueness presupposition of definite determiners. I agree with Hole’s (2011) opinion that: “shi…de” clefts in Chinese have the uniqueness and familiarity presupposition typical of definite determiners, applied to events. “This renders its semantic similar to that of definite determiners, but without leading to the definite reference to particulars typical of definite determiners” (Hole 2011, p, 1708). One condition has to be added here for this opinion: this comparison between exhaustivity and “definite determiners” is only tenable under the condition of argument-focus cleft constructions.

To take a Chinese example:

(4-18) 是小明打破窗户的。
SHI\textsubscript{aux} Xiaoming dapo chuanghu DE\textsubscript{ptcl}.

SHI\textsubscript{aux} Xiaoming broke window DE\textsubscript{ptcl}.

It is Xiaoming who broke the window.

The exhaustive reading of the cleft sentence is—the focus “小明 (Xiaoming)” is the only possible variable of the value “the person who broke the window”, which means there is no one besides 小明 (Xiaoming) who broke the window.

4.4.2.2 How to test the exhaustivity of a sentence

The claim that a non-cleft constituent with a pitch accent does not express exhaustive identification in English is not based merely on intuition; it can also be confirmed by tests of exhaustive identification devised by Szabolcsi (1981) and Farkas (1992). Szabolcsi’s test involves a pair of sentences in which the first sentence contains a focus consisting of two coordinate NPs and the second sentence differs from the first one only in that one of the coordinate NPs has been dropped. If the first sentence does not entail the second one, in other words, if the second sentence is not among the logical consequence of the first one, the focus of the second sentence expresses exhaustive identification. This method is called a “failing entailment” test. For example:

(4-19) 是小强和小明要求的。

SHI\textsubscript{aux} Xiaoqiang he Xiaoming yaoqiu DE\textsubscript{ptcl}.

SHI\textsubscript{aux} Xiaoqiang and Xiaoming require DE\textsubscript{ptcl}.

It is required by Xiaoqiang and Xiaoming.

(4-20) 是小明要求的。

SHI\textsubscript{aux} Xiaoming yaoqiu DE\textsubscript{ptcl}.

SHI\textsubscript{aux} Xiaoming require DE\textsubscript{ptcl}.
SHIaux Xiaoming yaoqiu DEptcl.

SHIaux Xiaoming require DEptcl.

It is Xiaoming who required that.

(4-21) 老李是老板以及项目负责人。

Laoli SHIcop laoban yiji xiangmu fuzeren.

Laoli SHIcop boss and project manager.

Laoli is the boss and the project manager.

(4-22) 老李是老板。

LaoLi SHIcoplaoban.

LaoLi SHIcop boss.

LaoLi is the boss.

By the test, sentence (4-20) is not entailed by (4-19), which means sentence (4-20) is a sentence with exhaustive focus. But sentence (4-22) can be logically entailed from the sentence (4-21), that means sentence (4-22) is not a sentence with exhaustive focus. Sentences (4-19) and (4-20) are Chinese cleft constructions in the form of “shi...de”; while sentences (4-21) and (4-22) are copular sentences in the form of “bare shi”.

Paul & Whitman’s (2008) test is more convenient and practical (i.e., just by simply adding one clause to the original sentence) for checking the exhaustivity of focus with similar gist as Szabolcsi (1981). The fundamental idea of the test is that under the condition of exhaustivity, asserting that the property denoted by the presupposition also holds of an entity distinct from the focus of the cleft leads to a contradiction. To quote from his study:
It is Mary that I gave the book to.

It is Mary that I gave the book to, and it is also John that I gave the book to.

The same situation is for Chinese cleft constructions:

(4-24) #是小明打破窗户的，也是小华打破的。

? SHI\textsubscript{aux} Xiaoming dapo chuanghu DE\textsubscript{ptcl}, yeshi xiaohua dapo DE\textsubscript{ptcl}.

? SHI\textsubscript{aux} xiaoming break window DE\textsubscript{ptcl}, also xiaohua broke DE\textsubscript{ptcl}.

? It is Xiaoming who broke the window, it is also Xiaohua who did that.

This approach is used in the present study when excluding the non-qualifying sentences.

The compound sentence linked by the coordinating conjunction “也 (ye; and)” is pragmatically unacceptable.

If the modified sentence is unacceptable, that means the original sentence expresses exhaustive meanings or focus.

Exhaustivity distinguishes cleft constructions from “shi” copular sentences and “shi” emphatic sentences. “Shi” copular sentences are shown as the predicate-focus structure from the perspective of information structure, with the whole predicate as the focus, which means, the focus is normally the VP of the sentence with the copula “shi” as the head of the VP. However, the predicate focus could be replaced by any other VPs or even NPs/APs, without being limited by any propositional conditions. For example:

(4-25) 小明是高中生。

43 The symbol “?” means the sentence is grammatically accepted, but pragmatically anomalous.
Xiaoming SHI\textsubscript{cop} gaozhongsheng.

Xiaoming SHI\textsubscript{cop} secondary school student.

Xiaoming is a secondary school student.

The information structure of the sentence is:

Presupposition: 小明(x)

Xiaoming (x)

Focus: 是高中生。

SHI\textsubscript{cop} gaozhongsheng.

is a secondary school student.

Assertion: x=是高中生。

x=SHI\textsubscript{cop} gaozhongsheng

x=SHI\textsubscript{cop} a secondary school student

In sentence (4-25), a copular sentence, the focus is not restricted to a verbal predicate, such as the one initiated by a copula in the example, but could be any other type of predicate, for example the nominal predicate\textsuperscript{44} and the adjective predicate.

The following example is a copula-less sentence:

(4-26) 今天星期一。 (Tang, 2001)

\textsuperscript{44} Copular-less sentences in Chinese are in the form of “subject + nominal predicate”. They “are subject to ‘Generalized Anchoring Principle’, which requires that every clause be anchored at the interface for LF convergence. To satisfy the principle, clauses may be either tense or focused. It is shown that copula-less sentences in Chinese are subject to focus anchoring.” (Tang, 2001).
Jintian xingqiyi.

Today Monday.

Today is Monday.

“Xingqiyi” is the nominal predicate of the example sentence, being an identifying copula-less sentence. The predicate nominals in the copula-less sentences denote the characteristic and quality of the subject. They can be common nouns, proper nouns, and numerals.

The example (4-26) could be reduced into a copula-less identifying sentence too:

(4-27) 小明高中生。

Xiaoming gaozhong sheng.

Xiaoming secondary school student.

Xiaoming is a secondary school student.

While the function of the sentence is to show the identity of the subject, there is no limitation for the focus (i.e., predicate) of the sentence; the identity of “Xiaoming” could be a secondary school student, an officer, a teacher, etc. Besides demonstrating the identity of the subject, the copula-less sentence, could also show the characteristics of the subject, as the following sentence:

(4-28) 小明 92 届的。

Xiaoming 92 jie DE_{rel}.

Xiaoming class of 92 DE_{rel}. 

“92 jie” is the nominal predicate showing the characteristic of the subject as an agent who graduated in 1992. The subject of this sentence is the same as the one in the last example, and the focus is also a nominal predicate, but the focus realizes a different semantic function. We can draw the conclusion that the focus of copular sentences (with or without the copular verb) is not exhaustive.

“Shi” emphatic sentences are distinguished from cleft constructions by, firstly, the exhaustivity and contrastivity of the focus, and secondly, the distribution of “presupposition” and “focus” from the perspective of information structure, which is discussed later. Hole (2012) illustrates that the copula “shi” added to the predicate without “de” licenses various foci, however, no exhaust effect is observed. In his study, Hole (2011) does not distinguish the bare “shi” cleft constructions from “shi” emphatic sentences. In the present study, I do collect the emphatic sentences both in the form of bare “shi” and “shi...de” and distinguish sentences from the semantic perspective. To cite one example collected in the present study:

(4-29) 我想，当当是必须要上市的。

Wo xiang, Dangdang SHIaux bixu yao shangshi DEptcl.

I think, Dangdang SHIaux have to list on market DEptcl.

I think that Dangdang has to be listed on the market.

This sentence, although with both “shi” and “de”, is only an emphatic sentence because, firstly, there is no exhaustive limitation for the focus constituent. The function of this sentence is to emphasize the confirmative mood of it by adding auxiliary “shi” in the
preceding position of the modal verb “必须 (bixu; have to)”. The focus of the sentence is the proposition “当当必须要上市 (dangdang bixu yao shangshi; Dangdang has to be listed on the market)”, and there is no presupposition in the sentence. The function of this emphatic sentence is to highlight the confirmative modality of the sentence. Secondly the distribution of the focus and the presupposition is another criterion reflecting one of the distinguished features of cleft constructions both in English and in Chinese, and it is closely related to the exhaustivity of focus, for only if the focus is a constituent (instead of the proposition), could it be exhaustive. This type of emphatic sentence is found in my corpus, and besides it, there is another type of emphatic sentence. We can call the former type “shi emphatic sentence—type 1”, whose function is to emphasize the sentence modality by being put in front of the modal word of the sentence. The other one becomes “shi emphatic sentence—type 2”, with the function of emphasizing the truth value of the whole sentence. To take one example of type 2:

(4-30) A: 太阳是从东边升起吗？

A: Taiyang SHIaux dongbian shengqi ma?

Sun SHIaux east rise ma?

Does the sun rise from the east?

B: 太阳是从东边出来。

B: Taiyang SHIaux cong dongbian chulai.

Sun SHIaux from east up.

The sun, indeed, rises from the east.
Like the previous example, this sentence is also an emphatic sentence with the whole sentence as the focus; however, the differences are: firstly, “shi” partners with prosodic prominence (without which, the sentence is only an explanatory copular sentence without any emphatic function) in the sentence intuitively when the sentence is uttered by native speakers; secondly, instead of emphasizing the confirmative mood of the sentence, the function of this type of emphatic sentence is to affirm the truth-value of the sentence—“it is totally true that the sun rises from the east”.

Despite the slightly different function, the emphatic sentences of both type 1 and type 2 have similar distributions of the “presupposition” (i.e., a null one) and the “focus” (i.e. the whole sentence); and neither of their foci is exhaustive. These two features distinguish “shi” emphatic sentences from Chinese cleft constructions. 45

So, despite the fact that many scholars (e.g., Lee, 2005; Shi, 1994) classify “shi…de” sentences as emphatic sentences, in this section, I want to show the difference from the perspective of exhaustivity. The difference in exhaustivity is actually derived from the distributive difference of presupposition and focus.

We find examples both of the two types of “shi” emphatic sentences in the present corpus, and their functions are either to emphasize the truth value of the sentence or to emphasize the modality of the sentence by highlighting the modal words (e.g. modal adverbs).

To quote one example of the “shi” emphatic sentence from the corpora:

(4-31) 刘春华是知道弟弟刘招华的再次被通缉的，而且是全国范围的悬赏通缉。

45 Also, there is another feature that is not held by emphatic sentence—the contrastivity of the focus (see 5.2.3).
(Liu Chunhua did know that her younger brother Liu Zhaohua was wanted again; and he was wanted with reward nationwide.)

刘春华是知道弟弟刘招华的再次被通缉的。

Liu Chunhua SHIiaux zhidao didi LiuZhaohua DEposs 46 zaici bei tongji DEptcl.

Liu Chunhua SHIaux know younger brother Liu Zhaohua DEposs again being wanted DEptcl.

Liu Chunhua did know that her younger brother Liu Zhaohua was wanted again.

刘春华知道弟弟刘招华的再次被通缉。

Liu Chunhua zhidao didi Liu Zhaohua DEposs zaici bei tongji.

Liu Chunhua knew that her younger brother Liu Zhaohua was wanted again.

Presupposition: ----

Focus: 刘春华知道弟弟刘招华的再次被通缉。

Liu Chunhua zhidao didi Liu Zhaohua DEposs zaici bei tongji.

Liu Chunhua knew that her younger brother Liu Zhaohua was wanted again.

There is no presupposition in the sentence-emphatic sentence, and the focus constituent is the whole sentence, so it is impossible for the focus to be exhaustive or contrastive. Just as mentioned before, the distribution of the “presupposition” and “focus” and the exhaustivity of the focus are closely related, because exhaustivity relates presupposition and focus together.

46“De” here is different from the one involved in the present study. It is a possessive marker to illustrate the subordinating relationship between the entity following “de” and the one preceding it.
For the copular sentences and the “shi” emphatic sentences, there is no exhaustive reading. That distinguishes them from cleft constructions.

4.4.3 The contrastivity of the focus

Contrastive emphasis is used by a speaker to mark a constituent as being in contrast with another structurally identical constituent. To test the contrastive meaning, Harries-Delisle (1978) uses the negative clause of the sentence such as:

(4-32) John bought a camel, not a donkey.

So the test for cleft constructions of English could be:

(4-33) It is a watch that Susan gave Mike, not a belt.

The same situation explains the Chinese counterparts. Cleft sentences can undergo the test like this:

(4-34) 是小明打破窗户的，不是小强。

SHIaux Xiaoming dapo chuanghu DEptcl, bushi Xiaoqiang.

SHIaux Xiaoming break window DEptcl, not Xiaoqiang.

It is Xiaoming who broke the window, not Xiaoqiang.

According to Tomioka (2007), the notion of contrastivity is connected to various linguistic phenomena, such as exhaustive answers in question-answer pairs, contrastive statements, or instances of corrective focus. The contrastive focus may be realized by different means in different languages. For example, in languages such as Hungarian or Finnish, “contrastive” elements are realized in a particular syntactic position suggesting the oppositeness (Kiss, 1998; Valduvi & Vilkuna, 1998).
Interestingly, Harries-Delisle (1978) argues that all contrastively-emphasized constructions have underlying cleft sentences, independent of whether the surface structure is an equational or non-equational one. This statement is evidence to support the idea that contrastivity is a distinguished feature of cleft constructions.

There are two pairs of concepts relating to the contrastivity of focus. The first two concepts are contrastive focus structure and informational focus structure. The second pair of concepts are the narrow and wide focus structure.

The partitions of the focus structure are proposed by many scholars with different names (Information focus and identificational focus in Kiss (1998), informational rhematicity and quantificational “contrast” in Vallduví and Vilkuna (1998) and informational focus and operational focus in Roberts (1998), etc.). In the present study, contrastive focus and informational focus are used. To borrow from Kiss (1998), the function of contrastive focus structures is “to represent a subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set for which the predicate actually holds”. On the contrary, if the function of a structure is to provide the nonpresupposed information to the original proposition but not to express any exhaustive identification performed on a set of contextually or situationally given entities, it is an informational focus structure. Considering the focus domains, the contrastive focus structure is often treated as narrow focus structure while the informational focus structure is the wide focus structure.

The relationship between the types of focus structures and their function of contrastivity is presented in the table:

**Table 4-8 Types of focus structures and their function of contrastivity**
<table>
<thead>
<tr>
<th>Broad focus structure</th>
<th>Narrow focus structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-contrastive focus</td>
<td>Contrastive focus</td>
</tr>
<tr>
<td></td>
<td>(cleft is one of them)</td>
</tr>
<tr>
<td></td>
<td>Non-contrastive focus</td>
</tr>
</tbody>
</table>

According to the previous studies, there are four prerequisites of a contrastive focus:

(i) There is a focus. Generally, every sentence has a focus, and different languages have their different unmarked distribution of the focus in a sentence, regardless of the types of focus structure. For example, the sentence-final position, also the most deeply embedded position on the recursive side of branching, is the default position for informational focus in Chinese.

(ii) The focus has to be a narrow one. Under some circumstances, a whole sentence or some parts of the speech are needed only for the goal of providing new information, that structure is a wide focus structure. Wide focus, which we shall refer to as the neutral focus, is typically associated with utterances in out-of-the-blue contexts, and is marked by ambiguity as to which constituent is singled out as the focus of the utterance.

These are some examples of the wide focus in Chinese:

(4-35) A: 你怎么迟到了？

A: Ni zemne chidao LE?

You why late LE?

Why are you late?

B: 路上堵车。
Lu shang duche.

On the road traffic jam.

There was a traffic jam on the road.

The answer B is an example of wide focus. The whole sentence is the focus, which means there is no presupposition existing in the answer.

(4-36) 我是很穷。

Wo SHI\textsubscript{aux} hen qiong.

I SHI\textsubscript{aux} very poor.

I am very poor indeed.

Focus: 我很穷。(wo hen qiong; I am very poor)

Assertion: 我很穷。(wo hen qiong; I am very poor)

The function of the sentence is to emphasize the truth-value of the presupposition, or to ascertain the correctness of the presupposition. Prosodically, “shi” is stressed in this “shi” emphatic sentence.

On the contrary, the cleft can only have a narrow focus interpretation (cf. Lambrecht 1994, p.17). As De Jong (1981, p. 99) notices, the pragmatic goal of a speaker uttering a cleft is the reduction of ambiguity as to which items are in focus.

(iii) The focus has to express the contrastivity

Just as discussed before, if the sentence can have a clause added that includes the negation of another entity (should be listed in the exhaustive set), it does express the contrastivity. If not, there is no contrastive function of the sentence, and the focus of the sentence is not a contrastive one.

The following two examples explain how contrastive focus is tested:
(4-37) 是我打破窗户的。

SHIaux wo dapo chuanghu DEptcl.

SHIauxI break window DEptcl.

It is me who broke the window.

Presupposition: 有人(x) 打破窗户。 (youren(x) dapo chuanghu; there is someone who broke the window.)

Focus: 我(wo; me)

The exhaustive set of the focus of the cleft sentence (4-37) could be {我 (me), 你 (you), 他 (him), 小明 (Xiaoming), 小强 (Xiaoqiang), 爸爸 (dad), 叔叔 (uncle), etc.}. By adding another clause to the original sentence, the sentence becomes a compound sentence with the latter clause negating another variable from the set.

是我打破窗户的，不是他。

It is me who broke the window, not him.

In this cleft construction, after adding the negative clause negating the possibility of “ta” being the focus of the presupposition “youren (x) dapo chuanghu. (there is someone (x) who broke the window)”, the compound sentence is still grammatical. This proves that the original sentence delivers a contrastive meaning.

(4-38) 我知道发生了什么了。

Wo zhidao fasheng LE shenme LE.

I know happen LE what LE.
I knew what happened.

Presupposition: 我 (wo; I)

Focus: 知道发生了什么了。 (zhidao fasheng le shenme; knew what happened)

*我知道发生了什么了，不知道没发生什么。

On the contrastivity of (4-38), there is no such exhaustive set for its focus; so it is hard for us to target any other element from the set and negate it. Being the predicate of the subject “wo”, the focus may be in various kinds. For example, the focus could be “是学生 (shi xuesheng; is a student)” and “爱游泳 (ai youyong; love swimming)” etc. If we negate them, and make the negating clause as the following clause of the original sentence:

*我知道发生了什么了，不是学生。

Wo zhidao fasheng LE shenem LE, bushi xuesheng.

I know happen LE what LE, not SHI student.

*I knew what happened, not a student.

The sentence makes no sense. The fact tells us that the original sentence “wo zhidao fasheng le shenme le/I knew what happened” does not express contrastive meaning.

(iv) There is an exhaustive set for the focus constituents.

Focus is contrastive only when it belongs to an exhaustive set of all possible variables satisfying the presupposition of the sentence.

(4-39) 是我打破玻璃的。

SHIaux wo dapo boli DEptcl.
It is me who broke the window.

Presupposition: (有人=x) 打破玻璃。((youren=x) dapo chuanghu; there is someone who broke the window.)

Focus: 我 (wo; me)

For the second one, the set is {我 (wo; me), 你 (ni; you), 他 (ta; him), 小明 (Xiaoming), 小强 (Xiaoqiang), etc.}, which may involve anyone who is the possible person “who broke the glass”. However, no such limited set can be established for copular sentences, because the focus can be any identification of the subject or the attribution of the sentence. The presupposition is an open proposition that cannot limit the set of the focus constituent.

To sum up, a cleft focus in Chinese, like its English counterpart discussed in Rochemont (1986, p.133), must receive a contrastive interpretation. It may, but need not, represent new information at the same time. To compare with English cleft constructions, Chinese counterparts do not involve syntactic reordering. An expression conveying contrastive focus remains in situ.

The criteria of cleft constructions in Chinese consist of three aspects—the exhaustivity of focus; the configurative properties of “presupposition” and “focus”; and the contrastivity of focus. They distinguish cleft sentences from other similar syntactic sentences. For example, the criterion of “exhaustivity” shows the differences between cleft sentences and the sentence-focus structures like the sentence-emphatic "shi” sentences in Chinese, as well as existential sentences in Chinese; the criterion of the configuration of “presupposition” and “focus” distinguishes cleft constructions from copular sentences; the criterion of
“contrastivity” tells the difference between cleft constructions and identifying and attributive copular sentences holding the informational focus. More generally, “contrastivity” also distinguishes cleft constructions from Chinese “even” sentences (i.e. 连...都/lian...dou).47 The detailed classifications of the similar-structured sentences are presented in the next chapter.

4.5 How is the data being used in the analysis sections?

This section is a brief introduction to the following two analysis chapters. What are the analyses for, and how is the data being used to support the analysis?

Chapter 5 deals with the syntactic structure of cleft constructions. There are several goals:

(i) In accordance with the different frequencies of the different types of strings, to conclude which structure is the most frequently used one, in other words, which is the prototypical or unmarked structure of cleft constructions;48 also, to figure out whether structures represented by other strings are variants of cleft constructions or merely other different types of sentences.

(ii) Based on the corpus, to classify the other types of “shi…de” sentences, and to distinguish cleft constructions from them (i.e. copular sentences and emphatic

47 Chinese “even” sentences (i.e. “lian…dou” sentences) are considered as emphatic sentences that without generating any contrastivity. Hole (2004) classifies Chinese “lian…dou” sentences simply as emphatic sentences, and Badan & Del Gobbo (2010) treat them as focus construction instead of contrastive focus constructions.

48“Generally the unmarked form is the more frequent option and also the one that has the most neutral meaning.”(Leech, 2006). Following Leech, the present study treats high frequency as a criterion of the unmarked or prototypical structure of the cleft construction in Chinese. That means the structure being used most frequently is the unmarked prototypical one.
sentences) on aspects such as: whether “shi” or “de” can be omitted or not; the word class of “shi” and “de”; the partition between presupposition and focus; the expression of exhaustivity and contrastive focus meaning; the focus domain and the relationship between presupposition and focus, etc.

Chapter 6 pays attention to the information structure of cleft constructions in Chinese. The following topics are discussed:

(i) The types of the focus structure of Chinese cleft constructions.

(ii) The identifiability and activation of referents of topic and focus (i.e. the status of mental representation in the addressee’s mind, from active to unanchored unidentifiable) are analyzed.

(iii) The relationship between presupposition and focus in different types of Chinese cleft constructions.

(iv) The thematic relations of focus and topic constituents.
Chapter 5

Forms and Properties of Chinese Cleft Constructions

5.1 Introduction

Unlike English cleft constructions, the Chinese counterparts cannot barely distinguished by their syntactic structures. In other words, if we treat all sentences with the syntactic structure “shi…de”, “bare shi” or “bare de” as cleft constructions in Chinese, there will be many non-qualifying sentences which are by nature non-cleft constructions, but resemble “true” cleft constructions syntactically. So in this chapter, the question of which are cleft constructions and which are not is discussed. To be specific, in the following discussions, firstly, the criteria that make a sentence a cleft construction are involved; secondly, the classification of “shi…de” sentences as a whole (more than just cleft constructions) is summarized.

In Chapter 4, the properties and distinguished features of Chinese cleft constructions were introduced; in the present chapter, we discuss how they reflect on the real data and how they apply to the judgement of cleft constructions are discussed.

This chapter aims to explore more comprehensively the structure of cleft constructions in Chinese. Three sentence types of cleft constructions (“shi…de”; bare “shi”, and bare “de”) are studied. The following topics are discussed in this Chapter:

(i) The prototypical structure and variants of cleft constructions of Chinese are confirmed by checking how commonly they are used in the discourse. They are confirmed by the frequencies of the three structures in the two corpora.

(ii) For each type, there exist some non-qualifying structures that interfere with the judgment of cleft constructions. They are the ones excluded by the secondary data
selection procedure. In this chapter, they are classified into different types in accordance with: firstly, the types of focus structure; secondly, the relationship between subject and the predicate; thirdly, the classifications of "shi" and "de", and fourthly, the relationship between presupposition and focus.

(iii) Under what circumstances “shi” or “de” can be omitted in cleft constructions and other non-cleft constructions.

5.2 What are cleft constructions and what are not?

The secondary procedure solves the problem of excluding the non-qualifying data from the set of the qualifying ones. As mentioned in Chapter 4, sentences that do not satisfy any of the exclusion criteria mentioned again here are non-qualifying sentences for the present study:

(i) There is a bipartition of “presupposition vs. focus” in the sentence, although “presupposition” and “focus” are not distributed in a clearly demarcated way. And the focus has to be a constituent inserted into the “presupposition”. The removal of the “focus constituent” does not make the sentence an ungrammatical or an incomplete one syntactically and semantically. “Shi” and “de” are not obligatorily included into either “presupposition” or “focus”.

(ii) The focus should be exhaustive.

(iii) The focus should be contrastive.

Each criterion itself may not make the construction a cleft one, but the combination of them decides whether the construction is a cleft. I have already introduced the criteria in Chapter 4. The following section is the result of the secondary selection procedure conducted on the basis of these criteria.
5.3 The classification of the data

The results collected from the primary search are composed of different types of sentences, and the cleft construction is only one type of them. In this section, I will put forward my classification of these types to help to distinguish cleft constructions from superficially similar structures; and to prove that although “shi...de” is the prototypical form of Chinese cleft constructions, not all “shi...de” sentences are cleft constructions.

Sentences are divided into two types in general considering the narrative status of the verbal constituent. A sentence headed by an eventive verb is called an eventive sentence; if the verb of the sentence is not an eventive one, normally it is called a state-of-affairs sentence. The Chinese "shi" sentences have usually considered to be copular sentences in previous studies. Treating "shi” as the main verb of the sentence, copular sentences are normally treated as state-of-affairs sentences. Yuan (2003) claims that sentences ended by “de” are also state-of-affairs sentences. He says, in his study, that state-of-affairs sentences have the function of changing the focus of the sentence. However, he does not consider “bare de” sentences to be a type of cleft construction having a contrastive reading of focus. In this present study, cleft constructions are also considered to be state-of-affair sentences.

Different types of “shi...de” sentences are given in the present study in accordance with the information structure and the relationship between subject and predicate of the sentences.

5.3.1 Copular sentences in Chinese and their sub-types

Copular sentences are normally defined as those sentences whose main verb is "be" (the copular verb) and its equivalents across languages (Zhan, 2012). Correspondingly, Chinese copular sentences are those in which the copular verb “shi” is involved. In the literature treating “shi” as the copula and “de” as the modification, scholars (e.g. Zhan, 2012) consider
cleft constructions to be a type of copular sentence, however the present study does not agree with this opinion.

Semantically, English copular sentences are divided into two types: predicational and specificational; in other words, copular sentences function to signal either a predicational or a specificational meaning (Declerck, 1988; Higgins, 1979; Mikkelsen, 2005, etc.). Declerck (1988, p. 2) defines a specificational sentence as one whose semantic function is to specify a value for a variable. The crucial characteristic for creating specificational meaning is that a copular sentence denotes a universally quantified restricted or existentially presupposed set, which is inherent in the semantics of a definite noun phrase, and a referential member that specifies the set. By contrast, a predicational sentence is “about” something, and usually explains or describes the nature, attribution or characteristics of the subject. To cite two examples from Niimura (2007):

(5-1) John is a philosopher.

(5-2) The bank robber is John Smith.

Sentence (5-1) is a predicational copular sentence with the main function being to describe the subject; sentence (5-2) is a specificational sentence whose function is to specify the value (i.e., John Smith) for a variable (i.e., the bank robber) by establishing the one-to-one relationship between the subject and the predicate.

In the present study, some of the “bare-shi” sentences and “shi...de” sentences are copular sentences with different specificational and predicational subtypes, according to the semantic relationship between the subject and the predicate of the sentence.

More specifically, predicational “shi” sentences are sentences with the post-copular expression describing a property of the subject referent. The post-copular expression may be
an NP, AP or PP. The copula serves as a linking word between the subject term and the predicative.

Specificational sentences perform a different function: the pre-copular constituent provides a variable, and the post-copular constituent provides a value for that variable or vice versa (Declerck, 1988; den Dikken, 2001; Higgins, 1979; Mikkelsen, 2005). Specificational sentences are claimed to obligatorily focus the post-copular constituent (Declerck, 1988; Higgins, 1979; Mikkelsen, 2005, etc.), while predicational sentences carry no such requirement.

Zhan (2012) classifies Chinese copular sentences into two semantic copulative categories as well: specificational and predicational. She also treats Chinese cleft sentences as specificational sentences, so in other words, she presupposes clefts to be copular sentences. In her study, she finds 2750 tokens of the cleft-SUBJ sentences in CCL Modern Chinese Corpus, of which 2122 (75.9%) share the form [PRO shì(ADV/TP/PP) VP de], and 538 (23.1%) are those in which the final nominalizer “de” is not present.

There are two points on which I do not agree with Zhan (2012). Firstly, I do not treat cleft constructions as a type of copular sentence. There are three reasons why Chinese cleft constructions are not copular sentences: (1) semantically, cleft sentences do not indicate equation, identification, inclusion or predicative property; (2) syntactically, one important restriction on the cleft formation is that no post-verbal phrase may be clefted, which means that “shì” cannot be directly located in the preceding position of the post-verbal constituent (i.e. object). That conflicts with the syntactic property of copular “shì”. (3) “De” cannot be part of the predicate. So when the corpus is searched according to the criteria of Chinese cleft sentences applied in the present study, we find the number of cleft constructions greatly
reduced. In the present study, most of the “bare-shi” sentences are analyzed as copular sentences; and some of the “shi...de” sentences are also copular sentences.

To sum up, in the present study, I do not consider Chinese cleft constructions to be one type of copular sentences. However, I do agree that specificational and predicational sentences are two types of copular sentences with different functions and I use them as a classification of copular sentences. The detailed classifications of each type (specificational and predicational sentences) are given according to the semantic relationship between the subject and the predicate.

Table 5-1 Classifications of copular sentences of the present study

<table>
<thead>
<tr>
<th>Copular sentences</th>
<th>specificational</th>
<th>Identifying</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Identifying</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classifying</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Explanatory</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Attributive</td>
</tr>
</tbody>
</table>

Identifying sentences are specificational sentences with both the subject and the nominal predicate being referential NPs. The function of the identifying sentences, obviously, is to express the identification of the subject by the complement. The structure of the identifying sentence is “SUBJ+shi+nominal predicate (VP+de)”. Classifying sentences aim to express the classification of the subject by the complement. The difference between the identifying sentence and the classifying sentence is that the former shows the relationship between two
referential NPs, while the latter shows the relationship between a referential NP and a non-
referential one. For example:

(5-3) 我是小明。

Wo SHI\textsuperscript{cop} Xiaoming.

I SHI\textsuperscript{cop} Xiaoming.

I am Xiaoming.

(5-4) 我是小学生。

Wo SHI\textsuperscript{cop} xiaoxuesheng.

I SHI\textsuperscript{cop} a pupil.

I am a pupil.

Sentence (5-3) is an identifying sentence with both subject and complement “我 (wo)” and
“小明 (Xiaoming)” mutually referential with each other. Sentence (5-4) is a classifying
sentence as well as a predicational sentence, with only the subject being a referential noun.
The semantic function of the sentence is to show the subject’s occupation. The structure of
the classifying sentence is similar to that of the identifying sentence—“SUBJ+shi+NP”.

The other two types of predicational sentences are explanatory and attributive sentences.
The explanatory sentences aim to give more explanatory information about the subject by
means of the complement. The attributive sentences aim to provide modifying information
about the subject from the perspective of its attribution by the complement. For example:

E.g. (5-5) Explanatory sentence:
5.3.2 The classifications of “shi…de” sentences

Lü (1980) classifies “shi…de” sentences according to the different constituents between “shi” and “de”:

(i) SUBJ+shi+NP+de—the “genus and material” of the subject
(ii) SUBJ+shi+VP/ADVP+de—the classification of the subject
Semantically, Lü (1980) includes four semantic interpretations within “shi…de” sentences. They are to express the geneity and material of the subject; the classification of the subject; the illustration or description of the subject (with stressed mood or not); and to highlight the subject of the clause.

Under the division of “specificational” and “predicational” sentences, following the criteria of Lü (1980) and by observing the collected data, in the present study, “shi…de” sentences are classified into several types by semantic functions—explanatory; classifying; identifying; attributive; emphatic and of course cleft sentences. Some of the types overlap with those of the copular sentences discussed in the previous section. To explain these types, I use some examples from the collected data.

**Table 5-2 Classification of “shi…de” sentences in the present study**

<table>
<thead>
<tr>
<th>Copular sentence</th>
<th>Specificational</th>
<th>Identifying</th>
<th>他  中国 第 一个 会溜冰 的。 (Ta SHI&lt;sub&gt;cop&lt;/sub&gt; zhongguo diyige hui liubing DE&lt;sub&gt;rel.&lt;/sub&gt;/ he SHI&lt;sub&gt;cop&lt;/sub&gt; China first can skate DE&lt;sub&gt;rel.&lt;/sub&gt;/ He is the first person who can skate in China.)</th>
</tr>
</thead>
</table>
| Classifying | 宝三儿是耍幡的。
(Baosaner SHI\text{cop} shuafan DE\text{nomi.} Bansaner SHI\text{cop} play juggling DE\text{nomi.} Bansaner juggles for a living.) |
| --- | --- |
| Predicational | 我是做报纸出身的。
(Wo SHI\text{cop} zuobaozhi chushen DE\text{ptel.} I SHI\text{cop} make papers DE\text{ptel.} My occupation starts from newspaper industry.) |
| Attributive | 大海、落日、沙滩是不需要用钱买的。
(dahai, luori, shatan SHI\text{cop} buxuyao yong qian mai DE\text{modi.} Sea, sunset, beach SHI\text{cop} no necessary need money buy DE\text{modi.} The sea, sunset, and beach are priceless.) |
The emphatic sentence, as discussed previously, is divided into two types—one aims to emphasise the truth value of the sentence; the other aims to emphasise the modality of the sentence. The detailed classifications and properties are shown in Table A in the Appendix.

5.3.3 The relationship between “shi...de” sentences, copular sentences and cleft sentences

To show the relationship better, I use a diagram to make the relationship as simple as possible:
The three constructions are independent of each other. Although they may overlap syntactically, they are totally different sentence structures delivering different functions. The reason why I do not trichotomize the whole set of “shi…de” sentences is that, besides these three constructions, there are some other constructions, for example, the existential sentence as “楼下是卖煎饼的 (Louxia SHI\textsubscript{cop} mai jianbing DE\textsubscript{nomi.} under the building SHI\textsubscript{cop} sell pancakes DE\textsubscript{nomi.} There is a stall selling pancakes under the building.)”, are also formed in the structure of “shi…de”. I do not deny the existence of these examples; however, the “shi…de” existential sentence does not appear in my corpus, so I did not include any data in the research.
5.4 Results of the corpora search

After the two steps of corpus-search, the final results include all the sentences satisfying the criteria mentioned in Chapter 4. The frequencies of the three strings (i.e. “shi...de”; bare “shi” and bare “de”) in the two corpora are presented evidence of their typicality in discourses.

![Bar chart showing the distribution of cleft constructions in informal and formal discourse.]

**Figure 5-2. The distributions of cleft constructions of different types in the two corpora.**

All the data selected by the secondary procedure satisfy the criteria of cleft constructions used in the present study. The general results show as:

(i) The overall frequency of cleft constructions is very low. The highest ratio, which is the frequency of the “shi...de” construction in informal discourse, reaches 0.4%. This fact shows the markedness of cleft constructions in Chinese.

(ii) The prototypical structure of cleft constructions is, just as discussed by most of the scholars, the “shi...de” construction.
Previous scholars (i.e., Fang, 1995; Huang & Fawcett, 1996; Shang, 2002) agree that “de” can be omitted, and the sentences are still cleft ones. In other words, compared with the omission of “shi”, the omission of “de” does not make the cleft construction a non-cleft one; but under most circumstances, the omission of “shi” does. However, this study shows that comparatively speaking, the frequency of bare “de” sentences is slightly higher than that of bare “shi”; in other words, in the present study, comparatively speaking, the bare “shi” construction (with “de” omitted) is not as “common” as represented in the previous literature. Instead, the structure of bare “de” is more frequently presented as cleft constructions. Although the difference of frequency is not very great, the result shows that instead of ignoring the bare “de” variant, we should start to consider bare “de” sentences as one of the variants and most importantly, as the main variant of Chinese cleft constructions.

Considering the genre differences of texts, normally, cleft constructions are more commonly used in informal discourses (e.g. TV talks shows, etc.) than they are in formal discourses (e.g. essays, etc.).

“Bare de” cleft constructions are not found in the formal corpus. This tells us that the omission of only “shi” in Chinese cleft constructions often happens in an informal context.

---

49 The two corpora are composed of genres of two different types (informal and formal texts).
5.4.1 Illustration of different types of “shi...de” sentences

5.4.1.1 Identifying sentences:

(5-7) 他是中国第一个会溜冰的。

Ta SHI<sub>cop</sub> zhongguo diyige hui liubing DE<sub>rel</sub>.

He SHI<sub>cop</sub> China first can skate DE<sub>rel</sub>.

He is the first person who can skate in China.

The subject of the sentence is a personal pronoun that holds a high degree of identifiability in the sentence. And the complement of the sentence is a nominal phrase with the head modified by the ordinal number “the first” and being omitted because of the repetition of the subject. The NP modified by the definite determiners is also considered to have a high degree of identifiability of reference. Identifying sentences are the ones with the subject as a definite NP and the complement as another definite NP linked by the copula. The identity of the subject is presented by a definite complement, so the relationship between the subject and the complement is a one-to-one relationship. In this sentence, the subject is normally the topic, as Chen (2004, p.1167) says: “A highly identifiable referent, which is by definition referential, is more likely to serve as the topic of the utterance than a referent with a lower degree of identifiability.” The verbal predicate is the focus of the sentence. The structure of the identifying “shi...de” sentence is usually configured as the subject+shi+nominal predicate (VP+de). “Shi” is the copular verb and “de” is the relative pronoun of the appositive clause “zhongguo diyige hui liubing de ren (the first person who can skate in China)”.

5.4.1.2 Classifying sentences

The difference between the identifying sentence and the classifying sentence falls on the referentiality of the complement. The function of the classifying sentence is to classify the
subject from a perspective of material, occupation, etc. The subject is usually a definite one, in the form of the personal pronoun, definite noun phrases, or proper nouns. In the sentence--宝三儿是耍幡的 (Baosaner shuafan de; Baosaner is an acrobat), the subject is a proper noun representing the person named “宝三儿 (Baosaner)”, and the complement is a nominal phrase with the particle “de” as a nominalizer which nominalizes the VP “耍幡 (shuafan; do acrobats)” into the nominal phrase “耍幡的 (人 (shuafan de (ren); (the people who) do acrobats)”. The construction “VP+de” is equal to the “VP+de+人 (ren; people)” and allows the omission of the head “人”. The configuration of the classifying “shi” sentence is: SUBJ+shi+nominal predicate (VP+de/ NP/NP+de). If “de” exists at the end of the “shi” identifying and classifying sentences, it is usually the modification (i.e. relative pronoun; nominalizer, etc.) of the constituent preceding it.

5.4.1.3 Explanatory sentences

Explanatory sentences aim to explain or describe the subject by providing more detailed information about it. The sentence is normally arranged as: “SUBJ+shi+nominal predicate (VP/ AP/ PP (+de)”. “De” in this sentence is an ending particle enhancing the (affirmative) mood of the sentence. So theoretically, “de” can be omitted without any grammatical inappropriateness.

e.g. (5-8) 我是做报纸出身的。

Wo SHI<sub>cop</sub> zuobaozhi chushen DE<sub>ptcl</sub>.

I SHI<sub>cop</sub> make paper start DE<sub>ptcl</sub>.

My occupation starts from the newspaper industry.
There is one thing worth mentioning here. In explanatory sentences (and attributive sentences), although the constituents following “shi” may be varied from VP+de, AP+de, PP+de, the post-copular constituents are, in nature, adjective. “De” in explanatory sentences is actually the modal particle.

5.4.1.4  **Attributive sentences**

The attributive sentence illustrates that the relationship between the subject and the predicate is an attributive one. The function of the predicate is to show more information about the subject from the perspective of its attribution. The structure of the sentence is mainly: SUBJ+shi+VP/AP/PP+de. The “de” in these sentences is the modifier of an omitted object. “Shi” is the copula linking the subject and the nominal predicate. The omission of “shi” is acceptable, as the structure “SUBJ+nominal predicate” is acceptable in Chinese. However, the omission of only “de” is not acceptable when the nominal predicate is composed of “VP+de”.

For example:

(5-10) 大海、落日、沙滩是不需要用钱买的。

Dahai, luori, shatan SHI_{cop} buxuyao yong qian mai DE_{modi}.

Sea, sunset, beach SHI_{cop} no necessary need money buy DE_{modi}.

The sea, sunset, and the beach are priceless.
大海、落日、沙滩不需要用钱买的。  
Dahai, luori, shatan buxuyao yong qian mai DE\text{modi}.

Sea, sunset, beach no necessary need money buy DE\text{modi}.

The sea, sunset, and the beach are priceless.

大海、落日、沙滩不需要用钱买。

Dahai, luori, shatan buxuyao yong qian mai.

Sea, sunset, beach no necessary need money buy.

The sea, sunset, and the beach are priceless.

*大海、落日、沙滩是不需要用钱买。

*Dahai, luori, shatan SHI\text{cop} buxuyao yong qian mai.

*Sea, sunset, beach SHI\text{cop} no necessary need money buy.

(5-11) 我刚刚讲的坚持是非常难做到的。

Wo ganggang jiangde jianchi SHI\text{cop} feichang nan zuodao DE\text{modi}.

I just mention de persistence SHI\text{cop}very hard to reach DE\text{modi}.

The persistence I just mentioned is very hard to reach.

我刚刚讲的坚持非常难做到的。

Wo ganggang jiangde jianchi feichang nan zuodao DE\text{modi}.

I just mention de persistence hard to reach DE\text{modi}.
The persistence I just mentioned is very hard to reach.

我刚刚讲的坚持非常难做到。

Wo ganggang jiang de jianchi feichang nan zuodao.

I just mention de persistence very hard to reach.

The persistence I just mentioned is very hard to reach.

*我刚刚讲的坚持是非常难做到。

*Wo ganggang jiangde jianchi SHI\textsubscript{cop}feichang nan zuodao.

I just mention de persistence SHI\textsubscript{cop} very hard to reach.

The persistence I just mentioned is very hard to reach.

Among the four types of copular sentence, the omission of “shi” (but not along with “de”) is unacceptable only in “identifying” copular sentences\textsuperscript{50}. This may be because the subject and complement of the identifying copular sentence are in a one-to-one referential relationship.

5.4.1.5 Emphatic sentences

Emphatic sentences aim to emphasize the whole sentence or some part of the sentence. As mentioned before, there are two types of “shi” emphatic sentences—type 1 functions to stress the modality of the sentence by directly emphasizing the modal words of the sentence; type 2 functions to emphasize the truth-value of the sentence. As a whole, the differences between the emphatic sentences and cleft sentences show from three perspectives: firstly, there is no

\textsuperscript{50} The omission of “shi” is unacceptable in identifying “shi” sentences, especially in written contexts. In oral context, sometimes, “shi” can be replaced by a pause.
contrastive focus in emphatic sentences; secondly, the focus is not exhaustive; thirdly, the
distribution of “presupposition” and “focus” is different.

To take one sentence from the corpus as an example:

(5-12) （但事发之前，）乡人是不知刘招华制冰的营生的。

But the incident happen before, fellow-villagers do not know Liuzhaohua poison-making business.

Before the incident happened, fellow-villagers did not know about the poison-making business of Liuzhaohua.

Presupposition: --

Focus: 乡人不知刘招华制冰的营生。(xiangren buzhi Liu Zhaohua zhibing de yingsheng; fellow-villagers do not know Liuzhaohua poison-making business)

Assertion: =focus

The focus is the whole sentence, so of course, there is no exhaustive choice of the focus
nor contrastive reading of the focus.

Talking about the word class of “shi” and “de”, we can tell from the variation of the
example sentence with “shi” and “de” omitted respectively that both “shi” and “de” can be
omitted at the same time without any ungrammatical or infelicitous readings. That means,
“shi” and “de” are auxiliary and modal words whose main functions are to express focus and
the certain mood of the sentence, and do not have any semantic functions.

(5-13) （但事发之前，）乡人不知刘招华制冰的营生。
(Dan shifa zhiqian, ) xiangren buzhi Liu Zhaohua zhibing de yingsheng.

But the incident happen before, fellow-villagers do not know Liu Zhaohua poison-making business.

Before the incident happened, fellow-villagers did not know about the poison-making business of Liu Zhaohua.

When “shi” is used in an emphatic sentence, “de” usually appears together with it. In the type 1 “shi” emphatic sentences, “shi” cannot exist in the sentence alone without the appearance of “de”:

(5-14) *(但事发之前，)乡人是不知刘招华制冰的营生。

But the incident happen before, fellow villagers SHIaux do not know Liu Zhaohua poison-making business.

Before the incident happened, fellow villagers did not know about the poison-making business of Liu Zhaohua.

On the contrary, “de” can appear alone in the sentence:

(5-15) （但事发之前，）乡人不知刘招华制冰的营生的。

But the incident happen before, fellow-villagers do not know Liu Zhaohua poison-making business DEptcl.

Before the incident happened, fellow-villagers did not know about the poison-making business of Liu Zhaohua.
In the type 2 “shi” emphatic sentences, the omission of “shi” and “de” is slightly different from that of in type 1:

(5-16) 太阳是从东边升起的。

Taiyang $\text{SHI}_{aux}$ cong dongbian shengqi $\text{DE}_{ptcl}$.

Sun $\text{SHI}_{aux}$ from east up$\text{DE}_{ptcl}$.

The sun, indeed, rises from the east.

(5-17) 太阳从东边升起。

Taiyang cong dongbian shengqi.

Sun from east up.

(5-18) 太阳是从东边升起。

Taiyang $\text{SHI}_{aux}$ cong dongbian shengqi.

Sun $\text{SHI}_{aux}$ from east up.

To sum up, “shi” and “de” are syntactically similar with “shi” and “de” in cleft constructions. The omission of both “shi” and “de” at the same time is acceptable in both types of emphatic sentences. The omission of only “de” in type 1 emphatic sentences is not accepted; but for type 2, this omission is acceptable.

5.4.1.6 Cleft constructions

For cleft constructions, firstly, let’s explain their properties in accordance with the criteria of cleft constructions proposed in the present study.

(5-19) 我是申请退的。
Wo SHIaux shenqing tui DEptcl.

I SHIaux apply for retire DEptcl.

I did apply for retirement.

(i) The relationship between presupposition and focus:

Presupposition: 我（x）退（了）。(wo (x) tui le; I retired.)

Focus: 申请 (shenqing; apply for)

Firstly, the presupposition is a completed proposition; secondly, neither “shi” nor "de" is obligatorily included in presupposition or focus; thirdly, the focus is a part of the sentence instead of a whole sentence; fourthly, the boundary between presupposition and focus is not clear-cut.

(ii) Exhaustivity of the focus:

(5-20) 我是申请退的，也是被迫退的。

? I did apply for retirement, and I was forced to do that too.

The focus is exhaustive, the exhaustive set of the focus would be {申请，被迫 etc.}.
The focus has to be the one and only variable from the set.

(iii) Contrastive focus:

(5-21) 我是申请退的，不是被迫退的。

Wo SHIaux shenqing tui de, bushi beipo tui DEptcl.

I SHIaux apply for retirement, not force to retire DEptcl.
I applied for a retirement; I was not forced to retire.

The focus is contrastive, which means to accept one variable and reject any other variable in the same sentence is acceptable and logical just as shown in the above example.

Secondly, on the syntactic categories of “shi” and “de”, “de” in cleft constructions is the ending particle and “shi” is the auxiliary and they can be omitted from the sentence. I’d like to show the syntactic structure of the cleft construction by the following tree diagram:

```
  S
 /\     /\    \\
 /      /      \\
 NP     VP     Particle
  \     /      /  \\
   \  /      /   \\
    Wo  Aux  V  \\
           /\    \\
          /  \   \\
         SHI  Adv  V
             /\     \\
            /  \     \\
           shenqing tui DE
```

To sum up, the structures of the types of “shi...de” sentences are:

(i) SUBJ+SHI+nominal predicate (referential—VP+DE_{rel})—identifying sentence

(ii) SUBJ+SHI+nominal predicate (nonreferential—VP+DE_{rel}/ De_{nomi})—classifying sentence

(iii) SUBJ+SHI+VP/AP/PP+DE_{ptcl}—explanatory sentence
5.4.2 Bare “shi” cleft constructions

There are two things worth mentioning based on the data of bare “shi” cleft constructions. Firstly, the omission of “de” in these sentences is fairly acceptable, which means “de” here, theoretically being a particle, does not have any tense/aspectual function in these sentences. In this section, I will exemplify the data from the perspective of the omission of “shi” or “de” or both; the information structure, and the exhaustivity as well as the contrastivity of the focus.

(5-22)答：那个官司根本不判，那个官司告到地方检察处，检察机关根本就不敢碰他，所以，你告他，他们不敢碰他。我不是只告计程车司机的。我是从那个国民党伪总统一路告起。

(That case couldn’t be judged at all. The case was appealed to the District Procuratorate, and they did not dare to offend the defendant. So although I sued him, the procuratorate did not dare to deal with it. I do not only sue taxi drivers. I sued people all the way up to the puppet government of Kuomintang.)

(i) 我是从国民党伪总统一路告起。

Wo SHIaux cong guomindang weizhengfu yilu gaoqi.

I SHIaux from the puppet government of Kuomintang along sue up.

I sued people all the way up to the puppet government of Kuomintang.
我从国民党伪总统一路告起。

Wo cong guomindang weizhengfu yilu gaoqi.

I from the puppet government of Kuomintang along sue up.

I sued people all the way up to the puppet government of Kuomintang.

(ii) Presupposition: 我(x)一路告起。

Focus: 从那个国民党伪总统。

Assertion: x = 从那个国民党伪总统。

(iii) Exhaustive: yes

(iv) Contrastive: yes

In this example, according to the context, there is another clause before this target cleft construction negating the other element from the exhaustive set of the focus. The combination of the positive and the negative clauses is to narrow down the focus domain of the “shi” clause. In “shi…de” cleft constructions, the focus is more fixed, which is the constituent closely following the focus marker “shi”; however, in bare “shi” cleft constructions, the focus may be any constituent following “shi”:

(5-23) 阿光是在刘招华跟他聊到这一节的时候对刘招华感到了非常的愤怒和恼火。他说，刘招华，你的老婆要生小孩，你他妈的怎么能够忍心将她一个人丢下不管呢?

(A Guang felt great anger and rage when Liuzhaohua told this to him. He said: Liu Zhaohua, your wife is going to deliver your baby, how could you leave her alone like this?)

(i) 阿光在刘招华跟他聊到这一节的时候对刘招华感到了非常的愤怒和恼火.
A Guang zai liuzhaohua genta liaodao zheyijie de shihou dui Liu Zhaohua gandao LE feichang de fennu he naohuo.

A Guang at Liu Zhaohua and him tell this part to Liu zhaohua feel LE great anger and rage.

A Guang felt great anger and rage towards Liu Zhaohua when Liu zhaohua told this to him.

(ii) Presupposition: 阿光（x）对刘招华感到了非常的愤怒和恼火。（A Guang dui Liu Zhaohua gandao LE feichang de fennu he naohuo; A Guang felt great anger and rage towards Liu Zhaohua.）

Focus: 在刘招华跟他聊到这一节的时候（zai Liu Zhaohua gen ta liaodao zheyijie de shihou; when Liuzhaohua told him that.）

Assertion: x=在刘招华跟他聊到这一节的时候（x= when Liuzhaohua told him that）

(iii) Exhaustive: yes

? 阿光是在刘招华跟他聊到这一节的时候对刘招华感到了非常的愤怒和恼火，也是在刘招华说到其他事情时对刘招华感到了非常的愤怒和恼火。

? A Guang felt great anger and rage when Liuzhaohua told this to him, and felt great anger and rage when he told him other things.

(iv) Contrastive: yes

阿光是在刘招华跟他聊到这一节的时候对刘招华感到了非常的愤怒和恼火，而不是在刘招华说到其他事情时对刘招华感到了非常的愤怒和
恼火。（A Guang felt great anger and rage when Liuzhaohua told this to him, instead of when they are talking about other things.）

The omission of “de” is acceptable in both sentence (5-22) and (5-23), because (5-22) is in present tense instead of past. The sentence (5-23) is a past-tensed sentence, however, there is already a past tense marker “le” following the main verb of the sentence—“感到 (gandao; felt)”, so the existence of the past tense marker “de” is not necessary.

(5-24) 李胜对司马懿说：“这次蒙皇上恩典，派我担任本州刺史（李胜是荆州人，所以说是本州），特地来向太傅告辞。”司马懿喘着气说：“哦，这真委屈您啦，并州在北方，接近胡人，您要好好防备啊。我病得这样，只怕以后见不到您啦！”

李胜说：“太傅听错了，我是回荆州去，不是到并州。”

(Li Sheng said to Sima Yi: “Thanks to the grace of the Emperor, I was nominated as the governor of this county” (Li Sheng is a local of Jingzhou, so he called Jingzhou “this” county). I am here to bid farewell to you.” Sima Yi panted out: “Sorry to trouble you. Bingzhou is in the north and close to Hu tribes, please take precautions against them. I am badly ill like this, I am afraid I won’t have another chance to meet you again.”)

李胜说：“太傅听错了，我是回荆州去，不是到并州。”

Li Sheng shuo: “Taifu tingcuo LE, wo SHIaux hui jingzhou qu, bushi dao bingzhou.”

Li sheng said: “Grand Preceptor heard wrong, I SHIaux back to Jingzhou go, not to Bingzhou”

Li sheng said: “Grand Preceptor, you heard me wrong. I am going back to Jingzhou, not to Dingzhou”
(i) 我回荆州去，不是到并州。

Wo hui Jingzhou qu, bushi dao Bingzhou.

I back to Jingzhou go, not to Bingzhou.

I am going back to Jingzhou, not to Bingzhou.

(ii) Presupposition：我回(x)去 wo hui (x) qu; I go back.)

Focus: 荆州 (Jingzhou)

(iii) Exhaustive: yes

？我是回荆州，也是回并州去。

？I am going back to Jingzhou, and also to Bingzhou.

(iv) Contrastive: yes

我是回荆州去，不是到并州去。

I am going back to Jingzhou, not to Bingzhou.

This is the only example among all the data that is an argument-focus cleft construction (to be discussed in Chapter 6 in detail). Strictly speaking, the focus domain of this sentence is not the “place”; however, the argument is indeed put into the focus position by that particular contrastive focus pattern “shi...bushi...”. Just as in example (5-22), the function of the compound sentence is to narrow down the focus domain. The fact of the low frequency of argument-focus cleft constructions in the corpus reveals that the argument-focus cleft construction is quite rare and marked compared with the other type of cleft constructions (i.e. adjunct-focus cleft constructions).
To sum up, bare “shi” cleft constructions are not as typical as they have previously been claimed to be. The omission of “de”, which is an ending particle and may function as a modal particle and an aspect/tense marker if necessary, is applicable in bare “shi” cleft constructions because there is no strict requirement of the tense/aspect property of “de”, but if there is, the omission of “de” will be unacceptable.

5.4.3 Bare “de” cleft constructions

As I mentioned previously, bare “de” cleft constructions appear more frequently in the corpus than bare “shi” cleft constructions do. This result is an unexpected one and against my hypothesis. Still, the data are examined from the same perspectives:

(5-25). 老王亲嘴告诉我的，窗子都炸得直响，他们谁也不动。
Lao Wang qinzui gaosu wo, chuangzi dou zha de zhi xiang, tamen shei ye budong.
Lao Wang told me himself that the window was exploded, and the window sound terribly, but they all did not move.

(i) 老王亲嘴告诉我，窗子都炸得直响，他们谁也不动。
Lao Wang qinzui gaosu wo, chuangzi dou zha de zhi xiang, tamen shei ye budong.
Lao Wang told me himself that the window was exploded, and the window sound terribly, but they did not move.

With the particle “de”, the focus of the sentence is the way Laowang told me—in person; if being omitted, the focus of the sentence is the content following the first clause—“窗子都炸得直响，他们谁也不动 (and the window sound terribly, but they did not move)”. So grammatically, the sentence after the omission of “de” is acceptable, however, it highlights different focus.
(ii) Presupposition: 老王（x）告诉我 (Laowang (x) told me.)

Focus：亲嘴 (himself)

(iii) Exhaustive: yes

? 老王亲嘴告诉我的，也是老王让别人转告我的。

? Laowang told me himself, also by other people.

(iv) Contrastive: yes

老王亲嘴告诉我的，不是老王让别人转告我的。

Laowang told me himself, not other people.

(5-26). 方珍珠 爸爸就给了我买烧饼的钱。刚才我走着回来的，连车都没雇！

Fang Zhenzhu: Dad just gave me the money for bread. I just came back on foot, and did not even hire a car.

刚才我走着回来的，连车都没雇。

Gangcai wo zou ZHE huilai DEptcl, lian che dou mei gu.

Just I go ZHE back DEptcl, even car not hire.

I just came back on foot, and did not even hire a car.

(i)  *刚才我走着回来（连车都没雇）。

* Gangcai wo zou ZHE huilai, lian che dou mei gu.

Just I go ZHE back, even che not hire.

I just came back on foot, and not even hire a car.

The action happened in the past, so an aspect marker “de” is needed.

(ii) Presupposition: 刚才我（x）回来。

Gang cai wo (x) huilai.

I just came (x) back.
Focus: 走着 (on foot)

(iii) Exhaustivity: yes

刚才我走着回来的，也是坐车回来的。

I just came back on foot, and also by car.

(iv) Contrastivity: yes

刚才我走着回来的，不是坐车回来的。

I just came back on foot, not by car.

（5-27）杨柱国：好，我马上回去。那什么，平亦奇来了没有？唐石青来了。（指卧室的门）在里边睡觉呢。他夜里一点才赶到的。

Yang Zhuguo: Ok, I’ll be back soon.Well, is Ping Yiqi here?

Tang Shiqing: Yes. (pointing to the bedroom door). He is sleeping inside.

他夜里一点才赶到的。

Ta yeli yidian cai gandao DEptcl.

He morning 1am not until arrive DEptcl.

He did not arrive until 1 in the morning.

(i) 他夜里一点才赶到。

Ta yeli yidian cai gandao.

He morning 1am not until arrive.

He did not arrive until 1 in the morning.
The sentence is in past tense, though, it is still accepted not only grammatically, but also semantically, if we omit “de”.

(ii) Presupposition：他（x）才赶到。(He has arrived (x))

Focus: 夜里一点（not until 1am in the morning)

(iii) Exhaustivity: yes

? 他夜里一点才赶到的，也是早晨10点才赶到。

? He did not arrive until 1am in the morning, also at 10 am in the morning.

(iv) Contrastivity: yes

他夜里一点才赶到的，不是早晨10点。

He did not arrive until 1 in the morning, not 10 in the morning.

（5-28）那会儿，她文化大革命前死的，那会儿八十岁。

Nahuier, ta wenhuadageming qian si DE, nahuier bashi sui.

At that time, she Culture Revolution before die DE, then 80 years old.

At that time, she died before the Cultural Revolution, when she was 80 years old.

*那会儿，她文化大革命前死，那会儿八十岁。

*Nahuier, ta wenhuadageming qian si, nahuier bashi sui.

At that time, she Cultural Revolution before die, then 80 years old.

At that time, she died before the Great Culture Revolution, when she was 80 years old.

“死（si; die)” is a momentary result verb, this sentence delivers the past tense and perfect aspect of the verb “si”, and the omission of “de” is not acceptable.
(i) Presupposition: 她 (x) 死了。 (she (x) died.)

Focus: 文化大革命前 (before the Great Culture Revolution.)

(ii) Exhaustivity: yes

? 她文化大革命前死的，也是文化大革命后死的。

? At that time, she died before the Great Culture Revolution, she also died after that.

(iii) Contrastive: yes

她文化大革命前死的，不是文化大革命后死的。

At that time, she died before the Great Culture Revolution, not after that.

(5-29) 这条大街，解放以后才修好的。

Zhetiao dajie, jiefang yihou cai xiu hao DEptcl.

This street, liberation after not until build well DEptcl.

This street had not been built until after the liberation.

(i) 这条大街，解放以后才修好。

Zhetiao dajie, jiefang yihou cai xiu hao.

This street, liberation after not until build well.

This street had not been built until after the liberation.

The temporal adverb “才 (cai; not until)” is usually in the structure of “SUBJ+Time+cai+VP” to express a finished action. The time expression before the adverb “才 (cai; not until)” is the time the action finished. The adverb is usually translated into
English as “not until”, so the sentence (5-29) is translated as “the street had not been built until the liberation”. Similar case is for sentence (5-27), the sentence is translated as “he did not arrive until 1am”. “Cai” has already has the function of aspect marking, so it is not necessary for the tense/aspect marker “de” to appear in the same sentence, and the omission of “de” is accepted in sentences.

(ii) Presupposition: 这条大街(x)修好了。(This street was built up.)

Focus：解放以后 (before the Liberation)

(iii) Exhaustivity: yes

? 这条大街，解放以后才修好，也是解放前就修好了。

? This street was built up after the Liberation, also before the Liberation.

(iv) Contrastivity: yes

这条大街，解放以后才修好，不是解放前就修好了。

This street was built up after the Liberation, not before it.

(5-30) 父母在这儿，我们打寿长街儿搬来的。

Fumu zai zher, women da Shouchang street banlai DEptcl.

Parents at here, we from Shouchang street move here DEptcl.

Our parents are here; we moved from Shouchang street to here.

(i) 父母在这儿，我们打寿长街儿搬来。

The main verb of the sentence (5-30) (i.e.搬过来) is a directional verb. Zhang (2000) categorizes “起来 (qilai; get up)” “过来 (guolai; come over)” “来 (lai; come)” “上来
(shanglai; come up), and “下来 (xialai; come down)” as the directional verb phrases signifying aspect, and the tense of the sentence is present. So the sentence is a present perfect one. In this sentence, “de” is not obligatory.

(ii) Presupposition: 我们(x)搬来。(We moved (x) here.)

Focus: 打寿长街儿。(from Shouchang street)

(iii) Exhaustivity: yes

我们打寿长街儿搬来，也是打东经路搬来。

We moved from Shouchang street, also from Dongjinglu street.

(iv) Contrastivity: yes

我们打寿长街儿搬来，不是打东经路搬来。

We moved from Shouchang street, not from Dongjinglu street.

(5-31).我由那儿爬过来的。

Wo you naer pa guolai DEptcl.

I from there crawl hereDEptcl.

I crawled from there to here.

(i) 我由那儿爬过来。

Wo you naer pa guolai.

I from there crawl here.

I crawled from there to here.
Similar to (5-30), the sentence is a present-perfect sentence.

(ii) Presupposition: 我(x)爬过来。(I crawled here (x).)

Focus: 由那儿(from there)

(iii) Exhaustivity: yes

? 我由那儿爬过来的，也是由那儿爬过来的。

? I crawled from there to here, also from there.

(iv) contrastivity: yes

我由那儿爬过来的，不是由那儿爬过来的。

I crawled from there to here, not from there.

(5-31) 我那个五零年出去工作去了。四八，四八年结的婚，那才十八岁。生我们大学生就，就二十岁才。二十岁生的。现在大学生三十多了，三十多了。那会儿也是，够累的慌反正。

我那个五零年出去工作去了，四八，四八年结的婚。

I went out for work in 1950, in the year 1948, in 1948 I got married.

(i) *我那个五零年出去工作去了，四八，四八年结婚。

* Wo nage wulingnian chuqu gongzuo qule, siba, sibanian jie DEptcl hun.

I (particle) 50th out work go le, year 1948, year 1948 marry DEptcl.
I went out for work in the 50th, at the year 1948, at the year 1948 I got married.

(ii) Presupposition: 我（x）结婚了。(I got married (x))

Focus：四八年（year 1948）

(iii) Exhaustivity: yes

? (我)四八年结的婚，也是五六年结的。

? I got married in the year 1948, also in 1955.

(iv) Contrativity: yes

(我)四八年结的婚，不是五六年结的。

I got married in the year 1948, not in 1955.

An interesting phenomenon is worth mentioning here. In this sentence, the particle “de” is preposed to the preceding position of the object. It is a phenomenon of defocalization.

Theoretically, the focus can fall on any constituent between “shi” and “de”, and under most circumstances on the constituent immediately following the focus marker “shi”, the preposing of “de” is functionally accepted, for the preposing does not interfere with the reading of the focus. In the structure of “shi…de” cleft, it is impossible to make the object as the focus, so the preposing of “de” is informatively legitimate.

From the perspective of Chinese morphology, this preposing phenomenon of “de” in this sentence is also considered as a typical characteristic of Chinese separable words. Chinese separable words are divided into different types according to different scholars. Li (1983)

---

51 Chinese separable words refers to the words which can be grammatically detached or separated by the other word, although the separable word itself before the detachment is already a completed word instead of two or more separated characters.
classifies them based on the syntactic constructions the separable word holds. The above example “结的婚 (jie de hun)” expresses a “verb-object” relationship. He concludes that under most circumstances, there are several properties of the separable words themselves:

(i) The object is often a nominal or an adjective.

(ii) The separable word can act as the main verb of the whole sentence.

(iii) The separable word can be modified by adverbials.

(iv) The separable word can be put in the negative form such as “A 不 (bu/not) A” and “A 没 (mei/no) A”

All the properties are satisfied by the verb “结婚 (marry)”, so the separate action by “de” is legitimate.

On the function of the insertion of “de”, Wang (2008) postulates that “de” stresses the confirmative mood of the speaker towards the whole or part of the proposition he/she made. In his corpus-based study, he finds that under most circumstances, the separable words which are separated by “de” always follow the focus marker “shi” to highlight the following constituents being the focus of the sentence. This fact explains the above example besides the reason of defocalization: the preposing of “de” to the preceding position of the object does not interfere with the focal reading of the present sentence and the preposing is applicable, although I assume that this kind of “de” preposing in Chinese cleft constructions is only applicable in some limited cases. I do not have enough data to prove that assumption in the present study, but it could be an interesting topic for future studies.

To sum up, the bare “de” cleft sentences is a variation of “shi…de” cleft sentences, with a similar information structure; “de” can be preposed to the preceding position of the object to
defocalize the object; in the bare “de” cleft sentences, “shi” is omitted legitimately; and “de” can be omitted (still keeping the sentence grammatical, but the focus may vary) if it is not the tense/aspect maker of the sentence. In a sentence with past tense and perfect aspect, the omission of “de” is not accepted. Compared with “shi…de” cleft constructions and bare “shi” clefts, bare “de” sentences still have some eventiveness, in which the tense/aspect of the main verb got more attention.

5.5 The word classes of “shi” and “de” in cleft constructions
As discussed in the previous chapter, in the present study, “shi” is considered as an auxiliary and “de” as a sentence-ending modal particle. In this section, they are analyzed into details based on the data collected in the present study.

5.5.1 The word class of “shi”
Verbs can be divided into two types, one being raising verbs and the other being control verbs. The control verb is a two-place predicate whose thematic structure includes two arguments: one is the subject, which is called the external argument (i.e. agent) and the other is a complement which is called the internal argument (i.e. theme). Raising verbs are one-place predicates which only have one internal argument acting as a complement.

Huang (1988) considers “shi” in cleft constructions as a raising auxiliary that allows the raising of the subject. According to Huang (1988), the deep structure of the sentence: 我打了他 (shi wo dale ta./ SHIaux I hit LE him. It was me who hit him.) is:
The subject “我” (wo; me) in the clause “我打了他 (wo da LE ta./ I hit le him. It was me who hit him.)” is raised to the spec of the IP, “shi” here is the auxiliary which allows the movement of the subject, hence it is called the raising auxiliary. The omission of the auxiliary is acceptable in these sentences. But one thing has to be mentioned in particular: although the subject of the clause can be raised before the raising auxiliary, the movement does change the focus of the sentence. Before movement, the focus of the sentence “是我打了他 (SHIaux wo da le ta./ SHIauxI hit LE him. It was me who hit him.)” is the subject “我(wo; me)”; the focus of the moved sentence “我是打了他 (wo SHIaux da LE ta; I SHIaux hit LE ta.)” is the whole clause “我打了他(wo da LE ta; I hit LE him.)”. In other words, the raising potential of the subject determines the possible location of “shi” both before the subject and after it. In Chinese “shi” sentences, if “shi” is located in the preceding position of the main verb of the sentence, very likely, it is a sentence-emphatic sentence and the auxiliary “shi” is often stressed prosodically. For the example sentence “SHIaux wo da LE ta.”, if the subject of the clause is moved, the sentence is an emphatic one; if not, it is a cleft sentence. To check this opinion against the data collected in the present study:

(5-32) 她要不接班儿还不让我退呢，我是申请退的。
Ta yao bu jieban hai burang wo tui NE, wo SHI shenqing tui DEptcl.

She if not take duty still not allow me retire NE, I SHI apply retire DEptcl.

I am not allowed to retire if she does not take my duty; I applied for retirement.

The subject is already raised in front of “shi”, and the sentence becomes an adjunct-focus cleft construction with the focus constituent being “申请” (shenqing; apply); and if the subject is not raised and remains in the following position of “shi”, the sentence is still grammatical:

(5-33) 是我申请退的。

SHIauxwo shenqing tui DEptcl.

SHIaux I apply for retire DEptcl.

It is me who applied for retirement.

Although the sentence is still a grammatical one, it does focus differently. Sentence (5-33) is an argument-focus cleft construction with the subject of the sentence being the focus.

As a cleft construction, “shi” in the emphatic sentence is also a raising auxiliary. The differences between the two sentence structures rest on the distribution of the information structure; the exhaustivity and the contrastivity of the focus constituent.

(5-34) 但事发之前，乡人是不知刘招华制冰的营生的。

(Dan shifa zhiqian, ) xianger SHIaux buzhi Liuzhaohua zhibing de yingsheng DEptcl.

But the incident happen before, fellow villagers SHI do not know Liuzhaohua poison-making business DE.
But before the incident happened, fellow villagers did not know about the poison-making business of Liuzhaohua.

(5-35) 但事发之前，是乡人不知刘招华制冰的营生的。

(Dan shifa zhiqian, ) xiangren SHIaux buzhi Liuzhaohua zhibing de yingsheng DEptcl.

But the incident happen before,SHI fellow villagers do not know Liuzaohua poison-making business DE.

But before the incident happened, it was the fellow villagers who did not know about the poison-making business.

Before the raising of the subject “乡人 (xiangren; fellow-villagers)”, the sentence (5-35) is an argument-focus cleft construction; but after the raising, the sentence becomes an emphatic sentence. Although (5-35) and (5-34) are different functional structures, both of them are grammatically and semantically acceptable. The differences (i.e. presupposition, focus, exhaustivity, etc.) between cleft constructions and “shi” emphatic sentences still exist (discussed in 5.4.1.5).

5.5.2  The word class of “de”

Unlike “shi...de” copular sentences, “de” in cleft constructions is not a modifier, relative pronoun or nominalizer, etc., but an ending particle. Its main function is to enhance the modality of the sentence at the very end of the sentence. And under certain circumstances, it also expresses the tense/aspect of the sentence.

Many linguists claim that the “shi...de” construction is frequently associated with a past-time implication to the occurrence of the particle “de”. Simpson & Wu (1999) even point out
that “de” is a past tense maker. However, according to the data of the present study, we can
tell that cleft constructions are not necessarily interpreted as past events.

Hsieh (1998:132) points out that “shi...de” and “shi...le” present the same distribution, so
he considers “de” as an aspectual marker. Lee (2005) supports Teng's assumption and regards
“de” as a particle which refers to the speaker's assertive attitude. This is agreed on by the
present study. However, her opinion that “the possibility of the omission of ’de’ may be
because the function of ’de’ overlaps with the function of ’shi’” (Lee 2005,p. 157) is not
supported by the present study. If their functions do redundantly overlap, how do we explain
the prototypical form of the cleft construction—“shi...de”?

In the present study, “de” is treated as an ending particle that does function to enhance the
modality of the sentence; however, I do not deny the tense/aspect characteristic it may still
hold. The characteristic is found by the omission “de” test I conducted on the qualifying cleft
data collected in the present study. I found that:

(i) Under most circumstances, the omission of “de” is grammatically acceptable. This
phenomenon proves that “de” is not necessarily the tense/aspect marker of Chinese
cleft constructions.

(ii) Under other circumstances, the omission is not acceptable. This usually happens in
sentences in which the perfect aspect must be realized by “de”.

So according to the cases of “de” omission, I would like to conclude that, although we
cannot say that “de” must be a tense/aspect marker in cleft constructions, it must have
some tense/aspectual characteristics left as an ending particle of the Chinese cleft
construction.
5.6 Summary

5.6.1 The prototypical structure of the cleft construction and its variants

The prototypical structure of the cleft construction in Chinese is in the structure of "shi...de" with "shi" as a raising auxiliary and "de" as the ending particle enhancing the sentential modality, and may show the past/aspect characteristic of the sentence under some circumstances.

Unlike the previous literature, which pays more attention to the bare "shi" structure and considers it as the main variant of cleft constructions, the present study finds the frequency of the bare “shi” construction is slightly lower than the frequency of the bare “de” construction. This result differs from the view widely held by many scholars (Fang, 1995; Huang & Fawcett, 1996; Shang, 2002) that bare “shi” cleft constructions are commonly accepted as the main variant of Chinese cleft constructions. On the contrary, the data prove that bare “de” cleft constructions are slightly more frequent in the corpus than bare “shi” cleft constructions.

5.6.2 The distribution of cleft constructions in formal and informal texts

In the two types of texts, the formal and the informal ones, the distributions of the cleft construction are quite different. Regardless of the detailed structures, generally speaking, cleft constructions are more frequently used in informal texts.

5.6.3 The classification of “shi...de” constructions

Although they are the prototypical structure of cleft constructions, "shi...de" constructions have to be subcategorized in order to avoid non-qualifying constructions. According to the data collected in the present study, “shi...de” sentences are divided into three types—copular sentences, emphatic sentences and cleft sentences. They differ in various perspectives, for instance, the exhaustivity and contrastivity of the focus, the semantic functions, the
relationship between “presupposition” and “focus”, etc. The focus of cleft constructions shows exhaustivity and contrastiveness; and the semantic function of the cleft construction is to make some part of the sentence the contrastive focus of the sentence; the presupposition of the cleft construction is normally a completed proposition and the boundary between presupposition and focus is not a clearly demarcated one.

Copular sentences are subdivided into “identifying”, “classifying”, “attributive” and “explanatory” sentences; emphatic sentences are divided into type 1 (which emphasizes the modality of the sentence); and type 2 (which emphasizes the whole sentence). The detailed classifications and their respective properties are shown in TableA in the Appendix.

5.6.4 The word classes of “shi” and “de” and the omission of them in “shi...de” constructions

Based on the data collected in the study, it is shown that in cleft constructions, “shi” can be omitted, but comparatively speaking “de” cannot be deleted as deliberately as “shi” can. My conclusion is at variance with some of the previous literature which concludes that it is more normal to omit “de” and treat the bare “shi” instead of bare “de” sentences as a short version of “shi...de” cleft constructions. However, like the previous studies, I do agree that the omission of both “shi” and “de” at the same time is acceptable grammatically.

In copular sentences, the omission of “shi” is rather free, except in identifying copular sentences. This is because it is grammatical in Chinese to have a structure of "subject+predicate" without a link in the form of the copula "shi", regardless of the types of the predicate (i.e. nominal predicate/verbal predicate/adjective predicate). However, to state a referential relationship between the subject and the predicate, as in identifying sentences, it is highly recommended to have a copula between the subject and the predicate.
In copular sentences, “shi” is the copula, and “de” is the head of the functional category (maybe a nominalizer, relative pronoun, etc.). As a functional head, “de” cannot be deleted alone in the three types (except for explanatory sentences in which “de” is a modal particle).

(5-36) 宝三儿是耍幡的。

Baosaner SHI\textsubscript{cop} shuafan DE\textsubscript{nomi}.

Bansaner SHI\textsubscript{cop} play acrobatism DE\textsubscript{nomi}.

Bansaner does acrobatics for a living.

“De” here is a nominalizer which nominalizes the VP “耍幡 (do acrobatics)” into a NP “耍幡的人 (the person doing acrobatics)”.

In cleft constructions, “shi” is the raising auxiliary, “de” is the particle of the sentence in cleft constructions, and the omission of both “shi” and “de” is acceptable. “De” cannot be
omitted alone when it reflects the aspectual/tense information of the sentence. As mentioned, the tree diagram of sentence (5-33) is shown as:

In emphatic sentences, the classifications of “shi” and “de” and the omission issue are all identical with those of cleft constructions; however, differences between them are not reflected in their syntactic structure. The syntactic analysis of sentence (5-35) is shown:
To sum up, the omission of “shi” is not rigidly limited, whether it be in copular sentences, emphatic sentences or cleft sentences. In other words, under most circumstances, “shi” can be omitted in these three types of sentences, even though it functions differently in the three types. However, the omission of “de” is relatively more rigid. Being a relative pronoun, or nominalizer, “de” cannot be omitted in copular sentences; and in emphatic and cleft sentences, if “de” is in a sentence with past tense and perfect aspect and there is no other tense/aspect marker in the sentence, it cannot be omitted.

The detailed differences between cleft constructions, copular sentences and emphatic sentences and the subtypes of these constructions are shown in Table A in the Appendix. In the next chapter, I will examine Chinese cleft constructions (in the form of “shi...de”, bare “shi” and bare “de” respectively) from the perspective of information structure.
Chapter 6

Analysis of the Information Structure of Chinese Cleft Constructions

6.1 Introduction

In Chapter 5, we showed the procedure of data selection based on the criteria of Chinese cleft constructions that distinguish them from other superficially similar structures. After the data selection procedure, all the qualifying data are prepared to be analyzed from the perspective of information structure. In this chapter, I will mainly discuss the two important components (i.e. the focus and topic) of the information structure of the Chinese cleft construction in all three structures (i.e. “shi...de”, bare “shi” and bare “de”) in Chinese. The topics discussed in this chapter are:

(i) The types of Chinese cleft constructions from the perspective of focus structure;

(ii) The configuration of the information structure of the collected data (i.e. the distribution of “presupposition” and “focus” in different types of Chinese cleft constructions);

(iii) The syntactic categories of focus constituents;

(iv) The identifiability of focus/topic constituents in argument/adjunct\textsuperscript{52}-focus cleft constructions.

(v) The thematic relations of focus and topic constituents in adjunct-focus cleft constructions.

\textsuperscript{52}Argument and adjunct focus cleft constructions are two types of cleft constructions in accordance with the focus domains of cleft constructions.
(vi) The identifiability of the thematic relations of the focus and topic constituents in adjunct-focus cleft constructions by checking the nominal constituents involved in the focus or topic.

(vii) The typological explanation of Chinese cleft constructions from the perspective of focus typology.

In this chapter, section 6.2 illustrates the types of Chinese cleft constructions by investigating the focus domain of each cleft sentence. The types are illustrated by providing some examples from the collected data, or introspective observation where the example of that type is insufficient. The configurations of the information structure of these two types are also checked. Section 6.3 pays attention to the syntactic classes of the focus constituents. Section 6.4 explores the thematic relations that the focus and topic constituents have in the frame of the sentence by establishing correlation between the syntactic categories of focus/topic constituents and their thematic relations. Section 6.5 checks the identifiability of the nominal phrases included in the “focus” and “topic” constituents of cleft constructions. To investigate the relationship between identifiability and the focus and topic constituents, we establish correlations from three aspects: firstly, the distribution of focus and topic constituents in the three identifiability statuses respectively; secondly, the grammatical expressions (of focus and topic constituents) and their identifiability; the third part deals with the relationship between the identifiability of the focus and topic constituents and the thematic relations they realize in the sentence frame. Section 6.6 shows the configuration of the information structure of Chinese cleft constructions by providing and analyzing the data collected from the present study. Section 6.7 discusses the typological characteristics of Chinese cleft constructions by checking firstly the order of topic and focus, and secondly the domain of focus constituents. The last section is a brief summary of this chapter.
6.2 The focus domain and types of focus structures

In the first place, I have to state that, on the whole, cleft constructions are such marked constructions that they are seldom used in discourse. Generally speaking, in the study I find 40 sentences in total as the qualifying cleft constructions in Chinese. And secondly, against my hypothesis that the argument-cleft construction would be the major group of the cleft constructions in Chinese, I only find one instance of an argument-focus cleft construction, so cleft constructions other than the argument-focus cleft construction exist overwhelmingly in the present corpus. To show the distributions of the argument-focus cleft construction and the non-argument-focus cleft constructions, the following chart is presented:

![Figure 6-1. Types of focus structure of cleft constructions in Chinese.](chart)

In the previous literature, there is no relevant study of the types of focus structure of Chinese cleft constructions, and the English literature does not discuss the difference between the argument-focus cleft constructions and any other type either.

Among the previous literature regarding Chinese cleft constructions, we can find many examples that can be categorized as the argument-focus cleft construction. However, by
searching the corpus, we confirm that in actual use, non-argument focus cleft constructions are more frequently used than the argument-focus cleft constructions.

Here, as in the chart, the term “non-argument” focus cleft constructions is only used as a general concept opposite to the “argument” cleft construction. It is not the specific name of the type of focus structure. A detailed illustration of the property of the type is shown in section 6.2.2.

6.2.1 Argument-focus cleft constructions

As stated, only one argument-focus cleft sentence is found:

(6-1) (李胜对司马懿说: “这次蒙皇上恩典，派我担任本州刺史（李胜是荆州人，所以说本州），特地来向太傅告辞。”司马懿喘着气说: “哦，这真委屈您啦，
在北方，接近胡人，您要好好防备啊。我病得这样，只怕以后见不到您啦！")

(Li Sheng said to Sima Yi: “Thanks to the grace of the Emperor, I was nominated as the governor of this county” (Li Sheng is a local of Jingzhou, so he called Jingzhou “this” county). I am here to bid farewell to you.” Sima Yi panted out: “Sorry to trouble you. Bingzhou is in the north and close to Hu tribes, please take precautions against them. I am badly ill like this, I am afraid I won’t have another chance to meet you again.”)

李胜说: “太傅听错了，我是回荆州去，不是并州”。

Li sheng shuo: “Taifu tingcuo LE, wo SHIaux hui Jingzhou qu, bushi Bingzhou.”

Li sheng say: “Taifu hear wrong LE, I SHIauxback Jingzhou to, not Bingzhou.”

Li sheng said: “Taifu heard it wrong. I am going back to Jingzhou, not Bingzhou.”

Semantically, it is an argument-focus cleft construction, and syntactically, it is a bare “shi” cleft sentence. This sentence is a quite special one, for the focus is not directly caused by the cleft construction; instead, it is caused by the combination of the cleft construction and the
sub-clause “bushi...” (a similar situation is discussed in Chapter 5) which delivers the contrastive meaning.

Regardless of the syntactic types (i.e. “shi…de”, bare “shi” or bare “de”) of Chinese cleft constructions, the object of the sentence cannot be focused in either of these types. This is an important feature of Chinese cleft constructions. The feature leads to the conclusion that regardless of the special context, argument-focus cleft constructions in Chinese are mainly subject-focus cleft constructions. The following example (6-4) explains the unacceptability of the object-focused situation:

(6-2) 小明打破窗户了。

Xiaoming dapo chuanghu LE.

Xiaoming break window LE.

Xiaoming broke the window.

(6-3) 是小明打破窗户的。

SHIaux Xiaoming dapo chuanghu DEptcl.

SHIaux Xiaoming break window DEptcl.

It is Xiaoming who broke the window.

(6-4) *小明打破是窗户的。

*Xiaoming dapo SHIaux chuanghu DEptcl.

*Xiaoming broke SHIaux window DEptcl.

(6-5) 小明打破的是窗户。
Xiaoming d apo DE_rel SHIcop chuanghu.

Xiaoming break DE_rel SHIcop window.

What Xiaoming broke is the window.

Sentence (6-2) is an unmarked Chinese sentence with indicative mood, aspect marker “了” ("le") and it satisfies the temporal sequence and the topic-prominence of Modern Standard Chinese. Sentence (6-3) is the argument-focus cleft version of the canonical sentence (6-2). The focus domain of (6-3) acts as the argument of the sentence which, at the same time, is the subject of the sentence represented by the proper noun “小明 (Xiaoming)”. Sentence (6-4) is ungrammatical, which means that informationally, the object of a sentence cannot be focused by the structure of “shi...de” cleft construction or syntactically, “shi”, as a focus marker, can only be put preverbally. However, sentence (6-5) is both grammatically and informationally acceptable. A form of (6-5) is called the pseudo-cleft construction in Chinese. Pseudo-cleft constructions in Chinese put the object of the sentence as the focus of the sentence. Unlike cleft constructions, whose focus constituents may vary from one type to another according to different focus domains (except for the object), pseudo-cleft constructions only have one type—object cleft constructions. Besides, the classifications of “shi” and “de” are different in pseudo-cleft construction of Chinese compared with those in cleft constructions. Chinese pseudo-cleft constructions act as a type of copular sentence aiming to assign one and the only entity/identity represented by the constituent following “shi” to the subject, which is often modified by the relative pronoun “de”.

However, I am not saying that example (6-1) is a non-cleft construction. It is really a bare “shi” construction, and as expected, the focus domain of the clause “我是回荆州去 (I am going back to Jingzhou)” is not the object “荆州 (Jing Zhou)”, instead it is the VP “回荆州
去 (going back to Jingzhou). But, it seems that it is the object “Jing Zhou” that is being contrastively focused. The object is being focused by the sentence pattern “shi…bushi…” in which the second clause supports the first clause by providing a negative proposition. Due to the contrastivity established by this sentence pattern, the focus domain of this sentence is comparatively narrowed down from VP to NP (i.e. the object of the clause). So although the original focus of this bare “shi” cleft sentence is not the object of the sentence, the object is accepted as a focus by hearers due to the particular clause pattern it evolved into.

As I said before, this is not a typical example of the argument-focus cleft construction in Chinese. To present a more general picture of the argument-focus cleft construction, I also cite two examples from other sources (i.e., “《莎菲女士的日记》/Sufei de riji/Miss Sophies Diary” by Ding Ling, 2008; “《暴风骤雨》/Bao feng zhou yu/ The Hurrican” by Zhou Libo, 2009).

(6-6) 我有如此一个美的梦想，这梦想是凌吉士给我的。

Wo you ruci yige mei de mengxiang, zhe mengxiao SHIaux Ling Jishi gei wo DEptcl.

I have such a beautiful dream, this dream SHIaux Ling Jshi give me DEptcl.

I have such a beautiful dream. It is Ling Jishi who gave me this dream.

This sentence is a topicalized argument-focus cleft sentence, with the canonical sentence as “是凌吉士给我这梦想的（SHIaux Ling Jishi gei wo zhe mengxiang DEptcl; SHIaux Lingjishi give me this dream DEptcl; It is Lingjishi who gave me this dream.）”. The object “梦想 (dream)” is preposed to the initial position of the sentence, and this is a phenomenon called topicalization. Topicalization is defined mainly syntactically as the movement of an element other than subject to the left edge of the sentence (cf. Prince 1986, p. 218). However, the
syntactic movement may generate two different functions. It may encode the topichood of the fronted element, or express contrastive focality. Whichever way this is done, the “prominent” meaning is expressed. To cite one example of topicalization from Ross (1967, p. 168):

(6-7) BEANS I don’t LIKE.

The unmarked order is: “I don’t like beans”. By preposing the object of the sentence, the aim of the sentence is to put the object in a more prominent position and thus realize the function of the sentence—to show the topichood of the object “beans” and make the listener aware that it is the specific part the speaker wants to highlight in particular by putting it at a prominent position. However, focus contrast may also be topicalized as in the following example:

(6-8) FIDO they named their dog. (Prince 1981, p. 259)

Semantically, contrastive focus is always generated by the following mechanism: firstly, there is a set of alternatives of focus constituents; secondly, a variable or an element in the set is chosen as the exact variable suitable for the value expressed by the sentence. In this example, the alternative set is the set of dog names such as {Fido, Jessica, Jack, Cooper, Lola, Harley…}; and “Fido” is the one and the only variable chosen as the focus of the sentence.

To distinguish these two types of “topicalization”, Prince (1981) names the first one as “topicalization” and the latter “focus movement”. However, Badan & Del Gobbo (2011) state that Chinese does not have a “topicalized” focus movement like the sentence (6-8). They test and conclude that the contrastive focus moved to the left periphery in Chinese is not accepted by native speakers. The sequence of the occurrence of the moved topic and contrastive focus is:
Moved topic > Focus (with contrastive stress) moved to the left periphery

Following their opinion, the topicalization of sentence (6-6) is a phenomenon of topic movement, which means that the initial nominal phrase “这梦想 (zhemengxiang)” is the topic (instead of the focus) of the sentence. The topicalized movement does not influence the focus of the sentence which is still the constituent directly following the focus marker “shi”; besides, just as in the case of “shi...OBJ de”, the above topicalization also partly explains the phenomenon of “defocalization”. The information structure of the sentence is shown in table 6-1.

(6-9) 这梦想是凌吉士给我的。

Zhe mengxiang SHIaux Ling Jishi gei wo DEptcl.

It is Ling Jishi who gave me this dream.

**Table 6-1 Configuration of the information structure of the example sentence (6-9)**

<table>
<thead>
<tr>
<th>Topic</th>
<th>Focus</th>
<th>presuppositioni</th>
<th>Assertion</th>
</tr>
</thead>
<tbody>
<tr>
<td>这梦想</td>
<td>是</td>
<td>凌吉士</td>
<td>给我</td>
</tr>
<tr>
<td>This dream</td>
<td>SHIaux</td>
<td>Ling Jishi</td>
<td>gave me</td>
</tr>
<tr>
<td>presuppositioni</td>
<td>(有人=X; there is someone=X)</td>
<td>presuppositioni</td>
<td>Assertion</td>
</tr>
</tbody>
</table>
In a topic-prominent language like Chinese or Japanese, object topicalization should be related to the passive\(^{53}\) because it can perform the following two functions:

(i) it foregrounds the patient

(ii) it de-topicizes the agent

In this passive sentence, the ending particle “de” is a past tense\(^{54}\) marker representing the completed status of the action shown by the main verb. In fact, regardless of the specific context, it is a well-known fact that subject-focus clefts give rise to a past tense reading only (cf. Dragunov, 1952: 116; Paris, 1979; Teng, 1979 among others). If “de” acts as the tense marker of the sentence, the omission of it causes the ungrammatical reading of the original sentence.

\(6-10\) *这梦想是凌吉士给我。

* Zhe mengxiang SHI\(_{aux}\) Ling Jishi gei wo.

* It is Ling Jishi who give me this dream.

The canonical sentence without the focus marking and topicalization is:

\(6-11\) 凌吉士给我梦想。

Lingjishi gei wo mengxiang.

\(^{53}\)It is generally accepted that Chinese does not have the passive voice from the perspective of syntax defined by traditional scholars (i.e. Bresnan, 1987; Bresnan & Kanerva, 1988, etc.). “Only if we look at passives from the point of view of pragmatics and define passives as constructions which defocus the agent and emphasize the affectedness of the patient (Shibatani, 1985) can Chinese be said to have passives” (Lapolla, 1998). Noonan & Woock (1978) name this type of passive the functional passive. The functional passive is realized by changing the word order, certain referential properties and the orientation of a sentence, but does not change grammatical functions.

\(^{54}\)It is widely-accepted that Chinese is a language without tense morphology. However Chinese does have tense-aspect particles, and some pragmatic reasoning to determine the temporal interpretation of sentences (Li, 1999; Lin, 2003, 2005; Smith & Erbaugh, 2005).
Lingjishi give me dream.

Lingjishi gave me the dream.

It is obvious that the focused constituent is the subject of the canonical sentence.

(6-12) 斗争韩老六时，悄悄溜号的刘德山也从山边的小窝棚里，回到家来了。老孙赶着老杜家的大车，常对人们说：“工作队长是我接来的。”

Douzheng Han Laoliu shi, qiaoqiao liuhao de Liun Deshan ye cong shanbian de xiao wopeng li, hui dao jia lai le. Lao Sun ganzhe Lao Du jia de dache, chang dui renmen shuo: “gongzuo duizhang SHIauxwo jie lai DEptcl.”

While fighting against Han Laoliu, Liu Deshan stopped sneaking away and arrived home from the shack at the foot of the mountain. Driving Lao Du’s cart, Lao Sun often said to the people that: “The working team leader was picked up by me”.

The same topicalized phenomenon happens to this sentence too. Very similarly, the information structure of the sentence goes like:

<table>
<thead>
<tr>
<th>Topic</th>
<th>Focus</th>
<th>presupposition₁</th>
<th>X</th>
<th>presupposition₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>work leader</td>
<td>SHIaux</td>
<td>me</td>
<td>picked up</td>
<td>DEptcl.</td>
</tr>
</tbody>
</table>

Table 6-2 Configuration of the information structure of the example sentence (6-12)
Presupposition: 工作队长（有人=x）接来。 (There is a person who picked up the working team leader.)

Focus: 我 (wo/me)

Assertion: x=我

Focus domain: personal pronoun

Topic: 工作队长 (the working team leader)

The canonical sentence is:

（6-13）我接来工作队长。

Wo jie lai gongzuo duizhang.

I pick up working team leader.

I picked up the working team leader.

Here too, the focused constituent is the subject of the sentence.

From the above examples, we can conclude that:

Firstly, the frequency of the argument-focus cleft construction is quite low. There is only one example that puts the argument as the focus of the cleft sentence, although the sentence is quite a marked argument-focus cleft construction, because in the sentence, the argument is
focused not only by the cleft construction, but also by the disjunctive clause “shi... bushi...” in the same sentence.

Secondly, based on the other data collected from other sources\textsuperscript{55} (i.e. novels and essays), topicalization is a common phenomenon in argument-focus cleft constructions (i.e. (6-9) and (6-12)). So combined with the first conclusion, the argument-focus cleft construction in Chinese is always arranged in a marked form.

Thirdly, the presupposition of the argument-focus cleft construction is a proposition, however it is an existential proposition with the subject of the existential proposition being an unknown person/entity (which is often supplemented by “x”), as in the analysis shown in the examples (6-9) and (6-12). Still, compared with other types of focus structures, such as predicate-focus structure whose presupposition is often not a completed proposition, the presupposition of the argument-focus cleft construction is a whole one.

\textbf{6.2.2 Adjunct-focus cleft constructions}

The prerequisite of Lambrecht’s (1994) categories of focus structures is the semantic division between argument and predicate of the sentence. This division is agreed on and accepted by the present study. However, such division is not comprehensive enough to deal with the types of cleft constructions in Chinese. As mentioned in Chapter 3, the adjunct-focus structure is also considered as one important type of Chinese cleft constructions.

Among the qualifying data, there is only one argument cleft construction. By analyzing the grammatical roles of the focus constituents, we conclude that all the non-argument cleft constructions involved in this present study are adjunct cleft constructions. Their functions are to elaborate the manner/purpose/place/time, etc. of the action by treating them as the

\textsuperscript{55} The data collected from other sources and the introspective data are not counted as part of the results of the present study. They are used as supplements when analyzing the data.
focus of the sentence. Unlike their English counterparts, there is no syntactic movement in Chinese cleft constructions; Chinese cleft constructions get their focus constituents announced mainly by the insertion of the focus marker “shi” in front of them.

The pragmatic function of the adjunct-focus cleft construction is to focus the place, time, purpose, etc. of the action of the sentence. The adjunct can be in various syntactic categories such as APs, VPs, PPs, etc. Here I take one example from the collected data to show that the differences between adjunct-focus cleft constructions and argument-focus cleft constructions mainly fall on two aspects: firstly, the different focus domains; and secondly, different configurative relationships between presupposition and focus.

(6-14) 陈存仪是在抓住刘招华后的第一时间由北京飞抵福建的。

Chen Cunyi SHIaux zai zhuazhu Liu Zhaohua DEmodi diyi shijian you Beijing feidi Fujian DEptcl.

Chen Cunyi SHIaux at catch Liu Zhaohua DEmodi first time from Beijing fly to Fujian DEptcl.

Chen cunyi flew back to Fujian from Beijing immediately after Liuzhaohua being caught.

“Zai Zhua zhu Liu Zhaohua de diyi shijian” is the focus constituent of this cleft sentence. The head noun of this constituent or clause is actually “diyi shijian”, and “zhua zhu Liu Zhaohua” is the preceding modifying constituent linked to the head noun by the modifier “de”. The focus is self-contained information in which some constituent is evoked dependently on the other one.

Table 6-3 Configuration of the information structure of the example sentence (6-14)
<table>
<thead>
<tr>
<th>Topic</th>
<th>Focus</th>
<th>Assertion</th>
</tr>
</thead>
<tbody>
<tr>
<td>presupposition&lt;sub&gt;i&lt;/sub&gt;</td>
<td>X</td>
<td>presupposition&lt;sub&gt;i&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

The presupposition of the sentence is a completed proposition already:

陈存仪由北京飞抵福建。

Chen Cunyi you Beijing fei di Fujian.

Chen Cunyi arrived Beijing from Fujian.

Because the focus of this type of cleft constructions acts as the adjunct of the sentence, the omission of the focus from the sentence does not make the sentence an ungrammatical one.

Relatively speaking, due to the large portion of adjunct-focus clefts compared with argument-focus clefts in the data, a more detailed analysis of adjunct-focus clefts is conducted.
6.3 Syntactic categories of focus constituents

The only example of the argument-focus cleft construction is a bare “shi” cleft construction with the word class of the focus constituent being a noun, to be specific, a proper noun semantically acting as a “place” in the sentence.

The syntactic categories of focus constituents in Chinese cleft constructions include verb phrases, noun phrases, adverbial phrases, and prepositional phrases, etc. The distributions of the categories are presented in the following chart:

![Figure 6-2: Syntactic categories of focus constituents of Chinese cleft constructions.](image)

Being adjuncts of cleft constructions, prepositions place 50% of the total number of the word classes of focus constituents. Noun phrases rank in second place with 17%; adverbial phrases and verb phrases hold 15% equally, and the category of focus constituent with the lowest percentage is the adjectival phrase.
In the following section, the syntactic categories of the adjunct cleft constructions are discussed in the three structures—“shi…de” cleft constructions, bare “shi” cleft constructions and bare “de” cleft constructions respectively.

### 6.3.1 “Shi…de” cleft constructions

In the “shi…de” cleft constructions, the rankings of the syntactic categories of focus constituents from high to low are prepositional phrases (44%); verb phrases (24%); noun phrases and adverbial phrases (with the same ratio), with the lowest being the adjective phrases (4%). Because the data for “shi…de” cleft constructions occupy a majority of the total number, the distribution of the word classes (i.e. chart 6-3) resembles the overall situation shown in chart 6-2.

![Figure 6-3. Syntactic categories of focus constituents in “shi…de” cleft constructions](chart)

### 6.3.2 Bare “shi” cleft constructions

There are only seven examples of bare “shi” cleft constructions, and one of them is the argument cleft construction. The situation of word classes of focus constituents appears as:
Figure 6-4. Syntactic categories of focus constituents of bare “shi” cleft constructions

There are only three categories of focus constituents in bare “shi” adjunct cleft constructions. This result does not show comprehensive conditions due to the limitation of the number of data. The percentage of bare “shi” sentences is too low to search a large quantity of data examples. We find that prepositions dominate (57%) the syntactic categories of focus constituents in this type of cleft construction. Besides, verb phrases are not found as the focus in bare “shi” cleft constructions.

6.3.3 Bare “de” cleft constructions

Based on the limited data, there are four types of syntactic categories of focus constituents of bare “de” cleft constructions. The rankings of the syntactic classes are PPs (50%), NPs (25%) and APs (12.5%) and VPs (12.5%). Similar to the other two types, prepositions are the dominant word classes of focus constituents. The distributions are shown by the chart:
Besides the most predominant syntactic category (i.e. the prepositional phrase) of the focus constituent of bare “de” cleft constructions, the second frequent syntactic category of bare “de” cleft constructions is the noun phrase. And because none of these bare “de” cleft constructions in the present study is an argument-focus cleft construction, and no cleft construction can focus the object of the sentence.

Due to the limitation of data we have in hand, especially the data for bare “shi” and bare “de” sentences, in the present study, I cannot make a conclusion about the difference between bare “shi” and bare “de” cleft constructions from the syntactic categories of their focus constituents. However I assume that if more data were collected, there might be some critical conclusions about this issue.

6.4 Thematic relations of focus constituents of adjunct-focus cleft constructions

As discussed in Chapter 3, the thematic relations involved in the FrameNet project are more comprehensive than those proposed traditionally. The thematic relations are called frame elements in FrameNet. The non-argument frame elements are named as adjunct arguments.
Every argument of the sentence, regardless of the type, is represented by one frame element. The same frame element may have various syntactic categories. To take a general look, the thematic relations of the focus constituents of adjunct-focus cleft constructions in the present study are shown by the following chart:

![Thematic relations of focus constituents chart]

**Figure 6-6. Thematic relations of focus constituents**

The frequencies of the thematic relations rank as:

Time (28%) > means (18%) = manner (18%) > place (8%) = source (8%) > cause (5%) > explanation (2.7%) = wrongdoer (2.7%) = degree (2.7%) = depictive (2.7%) = money (2.7%) = circumstance (2.7%)

A detailed demonstration of the data from three perspectives: the “name of the frame”, the “lexical unit” and the “frame element of the focus constituent” is shown in Table B of the Appendix.
Five frames are repeatedly used by the data. They are “Quitting”, “Intentionally-act”, “Arriving”, “Departing” and “Death”. The model of “arriving” appears five times in the target sentences, thus making this model the most frequently used one among all other models.

The frame elements represented by the focus constituents are mainly the non-core element. However, there are three exceptions:

(6-15) 我是从那个国民党伪总统一路告起。

Wo SHI AUX cong nage Guomindang weizhengfu yilu gaoqi.
I SHI AUX from that Guomindang puppet government away sue.
I sued along from the puppet government of Kuomintang.

(6-16) 我由那儿爬过来的。

Wo you naer pa guolai DE ptcl.
I from there crawl DE ptcl.
I crawled from there.

(6-17) 父母在这儿，我们打寿长街儿搬来的。

Fumu zai zher, women da Shouchangjier ban lai DE ptcl.
Parents at here, we from Shouchang street move DE ptcl.
My parents are here, and we moved from Shouchang street.

The frame represented by the verb of sentence (6-15) is “reporting”, and the thematic relation of the focus constituent is the “wrongdoer” which is the core element of the frame;

---

56Non-core frame element refers to the frame elements providing additional information to the semantic structure of a sentence. Each frame has a number of core and non-core frame elements which can be thought of as semantic roles. Core frame elements are essential to the meaning of the frame while non-core frame elements are generally descriptive (i.e. time, plane, manner, etc.) (https://framenet.icsi.berkeley.edu/fndrupal/glossary#core).
the focus constituent of sentence (6-16) is the “source” of the frame “Self_motion”; the focus constituent of sentence (6-17) is also acting as the “source” in its frame—“Bringing”.

In the following sections, I check the syntactic categories of the thematic relations and show the thematic relations of some of the focus constituents by taking examples from the collected data.

6.4.1 Syntactic categories of the thematic relations

This section aims to check the mutual relationship between the syntactic categories of the focus constituents and their thematic relations realized by these syntactic categories. In this section I will, on one hand, check the syntactic categories of the different thematic relations; on the other hand, I will show the distribution of thematic relations for each category.

The detailed illustration of the Frame element and the syntactic categories is shown in Table C of the Appendix.

“Time”, “manner” and “means” are the three thematic relations that are most frequently expressed by the focus of Chinese cleft constructions in the present study. The syntactic categories of the focus constituents expressing the three thematic relations are illustrated below.
“Time” is the thematic relation of focus with highest frequency in the present study.

Prepositional and nominal phrases are used to deliver the thematic relation “time”, and the prepositional phrases take the predominant position. The results, on one hand, are understandable because time phrases in Chinese are always headed by a preposition; on the other hand, they tell us that to deliver the “time” of an action in the sentence, it is acceptable to omit the head of the temporal prepositional phrase (preposition), and just leave the nominal phrase.

“Means” takes the second position among all the thematic relations. Surprisingly, VP has slightly higher frequency (67%) according to this investigation. That means that in Chinese cleft constructions, verbal phrases are often used to express the thematic relation “means”.

**Figure 6-7. Syntactic categories of focus constituents whose thematic relations are “time”, “means” and “manners”.

“Time” is the thematic relation of focus with highest frequency in the present study.
“Manner” is the third frequent thematic relation expressed by focus constituents of Chinese cleft constructions. The thematic relation “manner” is realized by VP and AdvP focus constituents. And AdvP occupies 67%.

In the following subsections, the respective distributions of thematic relations by PPs, AdvPs, NPs, and VPs are illustrated with examples.

6.4.2 The thematic relations represented by different syntactic categories

In this section, the thematic relations represented by prepositional phrases, adverbal phrases, nominal phrases and verbal phrases are checked.

![Figure 6-8](image)

**Figure 6-8. The thematic relations designated by prepositional phrases.**

Prepositions are used to express the time, the means and the place, etc. of the sentence. To take one example in which the thematic relation of the focus is “time”:
Chen Cunyi flew back to Fujian from Beijing **immediately** after Liu Zhaohua being caught. He had to meet the guy named “Liu Zhaohua” as an “opponent” of the police, who was the only person in the country using chemical synthesis to produce methamphetamine. He wanted to figure out what kind of person he was.

陈存仪是在抓住刘招华后的第一时间由北京飞抵福建的。他要亲眼看看这个可以算得上是警方的“对手”的、国内惟一用化学合成的方法创造了世界制造冰毒’奇迹’的刘招华到底是怎样的一个人……

Adverbs are used in two categories of thematic relations. They are “manner” and “degree”, and the thematic relation “manner” holds a predominant position.

“Zai zhuazhu Liu Zhaohua de diyi shijian (immediately after Liu Zhaohua being caught)” realizes the thematic relation of “time” of the main verb of the sentence “feidi (flew back)”.

NPs are used as time, explanation, and place from the perspective of thematic relations. Among the three thematic relations, “time” has the highest frequency.
我前天看了《珍珠衬》嗬，评剧嗬，当时呢我们是晚上看的。

Wo qiantian kanle shi “zhenzhu chen” he, pingju he, dangshi ne women

“I the day before yesterday watch le “Pearl Liner”, Pingju he, at that time ne we

SHIauxwanshangkan DEptcl.

I watched “Pearl Liner” the day before yesterday. It is a Pingju. We watched it at night.

“Wanshang” is the time of the action “watch”, its thematic relation is marked as “time” by FrameNet.

VPs are only used to express the manner and means of the main verb of the sentence.

中国的宗教情绪跟外国是有点不同。但中国也有宗教战争，可是在借口上不把宗教搬出来。太平天国是打着宗教旗帜起来的，而曾国藩打太平天国的理由就是维护名教。

Chinese religious sentiment is a little bit different from the foreign ones. But there is also religious war in China, although religion is not named as the reason for the war. The Taiping Heavenly Kingdom broke out under the cover of religion, and Zeng Guofan responded with the justification of guarding the principles.

“打着宗教旗帜 (Da zhe zongjiao qizhi)” is the adverbial of the main verb “qilai”, and its thematic relation is the “means” of the frame “Activity_start”.

(6-20) 我前天看了是《珍珠衬》嗬，评剧嗬，当时呢我们是晚上看的。

(6-21) 中国的宗教情绪跟外国是有点不同。但中国也有宗教战争，可是在借口上不把宗教搬出来。太平天国是打着宗教旗帜起来的，而曾国藩打太平天国的理由就是维护名教。
6.5 The identifiability of the “focus” and “topic” constituent of the cleft construction

This section is divided into three subsections. The first subsection aims to check the identifiability of the focus and topic constituents; the second subsection explores the grammatical expressions and their degree of identifiability; the third subsection aims to explain the identifiability of the thematic relations represented by focus/topic constituents. The results of this discussion relate the information structure constituents (i.e. focus and topic) to the identifiable status of their referents.

6.5.1 The distribution of focus and topic constituents in the three identifiable statuses

The identifiability of the grammatical form of a sentence constituent does not have a direct relationship with the identifiability of the referent of it. The latter is decided by the cognitive status of the hearer. Although the cognitive status of the hearer cannot be detected easily, the identifiability of the referent is not impossible. The identifiability is reflected by some grammatical and syntactic methods of a sentence or even a discourse.

As discussed in Chapter 3, the “identifiability” model proposed by Lambrecht (1994) is adjusted according to the data I collected in the present study. The model of identifiability used in the present study is a combined one based on Lambrecht’s (1994) and Chen’s (2004) studies. Lambrecht’s model categorizes the status of identifiability into “unidentifiable”, “inactive”, “accessible” and “active”; and “accessible” is further subcategorized into “textually”, “situationally” and “inferentially” according to the way the constituent is accessed.

Chen (2004) proposes the cognitive basis of identifiability in his study.
The status of being identifiable can be assumed by the speaker to have been established for an entity between him and the addressee by virtue of a variety of identificatory resources. Roughly speaking, they fall into two major categories. In the first category, the identifiability is directly evoked from its presence in the context of discourse, which is composed of the physical situation of utterance, and the linguistic text. In the second category, the identifiability of the entity in question is established on the basis of shared background knowledge between speaker and addressee, or inferable from other entities in discourse by virtue of the knowledge shared by participants of the speech event about the associations between the former and the latter. (pp. 1136-1137)

He summarizes that the uses of the English definite article fall into four major categories and some subcategories, and among all these categories, “shared general”, “specific knowledge”, “frame-based association”, and “self-containing association” are the subtypes of the second cognitive basis mentioned in the last paragraph, which is quite similar to the “inferentially-accessible” category proposed by Lambrecht (1994).

By analyzing the data, I find that Lambrecht's model cannot cover all cases in the present study, so I have combined some of the subcategories proposed by Chen (2004). The new model is as follows:
The focus and topic constituents of adjunct-focus cleft constructions are checked from the perspective of their identifiability according to the context in which they are involved, and the detailed results are shown in Table D in the Appendix.

To take a general perspective, and to check the identifiability of focus and topic constituents respectively, the distribution of focus and topic constituents in all identifiable statuses is shown as:

---

57 30 sentences among the 39 adjunct-focus cleft constructions are chosen according to the focus domain of each sentence. Because the identifiability refers to the cognitive status of a “referent”, the premise of “identifiability” should be a nominal or should be any constituent that projects a nominal (i.e. PP=preposition + noun etc.) from the perspective of grammar. That means if the focus of a sentence is neither a NP nor any construction projecting a noun, the sentence is excluded from the analysis of this section.
Figure 6-10. Identifiability of the focus and topic constituents

All the possible identifiabilities of the focus and topic constituents of the tested examples show that the “unidentifiable” and “inactive” statuses are not eligible statuses for topic constituents of Chinese cleft constructions in the present study. However, the focus can be expressed in all the four statuses of identifiability and the tendency of usage decreases a little from “unidentifiable” to “active”. The detailed results are shown as follows:

(1) The topic is never an unidentifiable or an inactive referent.

(2) This chart is evidence for the Topic Acceptability Scale proposed by Lambrecht (1994) (see Chapter 3, Figure 3-2). In the present study, the identifiability and activation of focus/topic are explored; however, the further criteria of “accessibility” are not discussed on purpose. To see the detailed information on the types of identifiability/activation of the focus and topic, please see Figure E and F in the Appendix.

(3) There are still no rules about the identifiability/activation of the focus constituent, because in principle the focus constituent is free with respect to the question of
identifiability and activation. The constituent of the focus may be new or given information, but the relationship between the focus and the asserted proposition must be new. Generally speaking, there is no significant numeral distinction between the different categories of the identifiability of the referents of focus constituents. However, there is still a slight difference between the four statuses.

In the following sub-sections, argument-focus cleft constructions and adjunct-focus cleft constructions are illustrated in details.

6.5.1.1 The identifiability of focus/topic constituents in argument-focus cleft constructions

Besides the argument cleft construction collected from the corpus, I state two other examples from other texts beyond the corpora. The ratio of the cleft construction in discourse is rather low, that is, only 40 of the 10,000 sentences are cleft constructions in the present corpus. The ratio of the argument-focus construction is even lower. So it is not very common to get the qualifying argument-focus cleft constructions in the present corpus, although it is quite common and natural to make introspective examples. For example:

(6-22) 是小明打破玻璃的。

SHI aux Xiaoming dapo boli DEptcl.

SHI aux Xiaoming broke the window DEptcl.

It is Xiaoming who broke the window.

(6-23) 是我对这件事负责。

SHI aux wo dui zhejianshi fuze.

SHI aux I to this issue take responsibility.
It is me who should take responsibility for this issue.

However, in order to find out the relationship between the information structure and the syntactic form of the referent, other examples are also stated in this section.

Firstly, let us consider the data collected from the present study:

(6-24) 李胜对司马懿说：“这次蒙皇上恩典，派我担任本州刺史（李胜是荆州人，所以说是本州），特地来向太傅告辞。”司马懿喘着气说：“哦，这真委屈您啦，并州在北方，接近胡人，您要好好防备啊。我病得这样，只怕以后见不到您啦！”

(Li Sheng talked to Sima Yi: “Thanks to the grace of Emperor, I was nominated as the governor of this county” (Li Sheng is the local people of Jingzhou, so he called Jingzhou “this” county). I am here to bade farewell to you.” Sima Yi panted out: “Sorry to trouble you. Bingzhou is in the north and close to Hu tribes, please take precautions against them. I am badly sick like this, I am afraid I don't have another chance to meet you again.”)

我是回荆州去，不是并州。

Wo SHIauxhui Jingzhou qu, bushi Bingzhou.

I SHIaux back Jingzhou go, not Bingzhou.

I am going back to Jingzhou, not Bingzhou.

Jingzhou is the focus of the sentence represented by a proper noun.
“Jingzhou” is not mentioned within the three clauses, so it is not an active referent; instead, it is a textually accessible\textsuperscript{58} one.

(6-25) 我有如此一个美的梦想，这梦想是凌吉士给我的。

Wo you rucí yíge meí de mèngxiǎng, zhè mèngxiǎng SHIaux Ling Jishi geiwo DEptcl.

I have such a beautiful dream, this dream SHIaux Ling Jishi give me DEptcl.

I have such a beautiful dream. It is Ling Jishi who gave me this dream.

The proper noun is a person’s name and the name is discourse-old as well as hearer-old information. However, the distance between the proper noun “Ling Jishi” and the former reference is more than three clauses. So it is textually accessible.

(6-26) 斗争韩老六时，悄悄溜号的刘德山也从山边的小窝棚里，回到家来了。老孙赶着老杜家的大车，常对人们说：“工作队长是我接来的。”

Douzheng Han Laoliu shi, qiaoqiao liuhao de Liu Deshan ye cong shanbian de xiaowopeng li, huidao jialaile. Lao Sun ganzhe Lao Du jia de da che, chang dui renmen shuo: “gōngzuò duizhang SHIauxwo jie lai DEptcl”.

Fight against Han Laoliu when, secretly sneak away de Liu Deshan too from hillside de shanty in, go back home le. Lao Sun drive Lao Du’s cart, often to people said: “working team leader SHIaux I pick up DEptcl.”

\textsuperscript{58}The classifications of identifiability were introduced in Chapter 3, and the criteria for each type were also illustrated.
While fighting against Han Laoliu, Liu Deshan stopped snuck away and arrived home from the shack at the foot of the mountain. Driving Lao Du’s cart, Lao Sun often said to the people that: “It is me who picked up the working team leader”.

The focus of the sentence is the subject and the object is preposed to the initial position of the sentence. The subject, which is at the same time the focus of the sentence, is a personal pronoun which is, according to the binding theory of anaphora, not bound in its binding domain. In other words, the meaning of the pronoun is referred by some word beyond its binding domain. In the example sentence, “我 (wo; me)” is traced back to the argument of the previous clause—“老孙 (Lao Sun)”. The focus constituent "我 (wo; me)" is active.

To summarize, the foci of these sentences are all active or accessible with the following word classes:

Table 6-4 The identifiability of the focus constituent in argument-focus cleft constructions.

<table>
<thead>
<tr>
<th></th>
<th>identifiability</th>
</tr>
</thead>
<tbody>
<tr>
<td>荆州 (proper noun)</td>
<td>active</td>
</tr>
<tr>
<td>凌吉士 (proper noun)</td>
<td>textually accessible</td>
</tr>
<tr>
<td>我 (personal pronoun)</td>
<td>active</td>
</tr>
</tbody>
</table>

The identifiability of the topic of these sentences is shown below:

Table 6-5 The identifiability of the topic constituent in argument-focus cleft constructions
In Chinese argument-focus cleft constructions, there is no significant distinction in identifiability between the topic and the focus. Proper nouns, personal pronouns, and demonstrative nouns are typical grammatical methods of identifiability.

Chen (2004) proposes that proper nouns and personal pronouns in Chinese are the morphological methods used to express the definiteness of the referents. Demonstratives are the definite determiners; bare NPs are the indeterminate expressions whose identifiability of referents is highly limited by their positions in the sentence. And definiteness is one of the ways to realize identifiability. According to the two tables, we can tell that in argument-focus cleft constructions, personal pronouns and the proper nouns deliver “active” or “accessible” status; the noun modified by the definite determiner (i.e., the Chinese demonstrative) has “active” status too. The modified noun is a textually-accessible constituent.

---

59Modified nouns in the present study are the nouns modified by the relative clauses in Chinese.
6.5.1.2 The identifiability of focus/topic constituents in adjunct-cleft constructions

The syntactic classes of the focus constituents in adjunct cleft constructions are mainly four types: VP, AdvP, PP, and NP. We are looking for the accessibility of the nominal phrases involved in these constituents, so all the nominal phrases are analyzed. All the nominals from the adjunct arguments are collected and stated in Table E in Appendix by showing their identifiability and their word classes.

The subcategories of “self-containing” and “frame-based” are merged into the category of “inferentially-accessible” as shown in Figure 6-14. I’d like to take two examples to make an introduction.

(6-27) 陈存仪是在抓住刘招华后的第一时间由北京飞抵福建的。

Chen Cunyi SHIaux zai zhuazhu Liu Zhaohua de diyi shijian you Beijing feidi Fujian DEptcl.

Chen Cunyi flew back to Fujian from Beijing immediately after Liu zhaohua being caught.

“Zhua zhu Liu Zhaohua de diyi shijian” is the focus constituent of this cleft sentence. The head noun of this constituent or clause is actually “diyi shijian”, and “zhua zhu Liu Zhaohua” is the preceding modifying constituent linked to the head noun by the modifier “de”. The focus is a self-contained one in which a constituent is evoked dependently of the other one.

(6-28) 继昨天进行两轮磋商之后，今天美方代表团三进三出中国外经贸部，最后的一次磋商是从下午三点四十分开始的，只进行了两小时二十分钟。

Ji zuotian jinxing liang lun cuoshang zhihou, jintian meifang daibiantuan san jin san chu Zhongguo waijingmaobu, zuihou yici cuoshang SHIauxxiawu sandiansishifen kaishi DEptcl, zhi jinxing le liangxiaoshi ershi fenzhong.
After two-rounds of negotiation yesterday, the American delegation went to the Ministry of Commerce three times today. The last negotiation started at 3.40 pm, and it only lasted for 2 hours and 20 minutes.

The topic of the clause “zuihou yici cuoshang (the last negotiation)” is a frame-based constituent that can be triggered by the previous mentioned constituent “lianglun cuoshang (two-round negotiations)”.

Being the two criteria of “inferentially-accessible”, “self-containing” and “frame-based” are considered as two methods of the realization of accessibility; however, for the sake of simplicity and convenience, the “self-contained” referent and the “frame-based” referent are generally counted as “inferentially-accessible” ones. In other words, although I use the two concepts to judge “inferentially-accessible” referents, there is no need, in the present study, to divide the referents into “self-contained” or “frame-based” ones.

6.5.2 The grammatical expressions and their identifiability
The identifiability of the sentence constituent is a concept only relevant to nominal constituents, so the whole of section 6.5 is talking about the nominal constituent involved in the focus and topic of the sentence. In this subsection, I check the grammatical expressions of these constituents and explore the relationship between their grammatical classes and their statuses of identifiability.

6.5.2.1 The word classes and the identifiability of topic and focus
In each cleft sentence, there is one topic and one focus. I classify the topic and focus according to their word classes, and by doing that, the result shows the correlation between some specific word classes (i.e. bare nouns) and their informative roles (i.e. focus/topic). The result is shown by the chart:
All the bare nouns are used as the focus constituents of Chinese cleft constructions; all personal pronouns are acting as topics of the sentences; proper nouns and demonstrative nouns are used both as topic and focus, and the frequencies of the proper nouns being the topic and the focus constituent are the same; most of the modified nouns express the focus of the sentence.

According to Chen (2004), definite and indefinite as well as indeterminate expressions in Chinese are varied in degrees of identifiability. In Chinese, there are some grammatical means to express definite and indefinite, although there are no articles, which are the most unmarked way to express definiteness in languages such as English. Therefore in the next section, I check the degrees of identifiability expressed by the above word classes in the sentences collected from the present corpus. And in this section, at least, we can get two conclusions: firstly, all bare nouns of the collected sentences act as the focus of the sentence; secondly, all personal nouns are the topic constituents of the sentence.
6.5.2.2 The degrees of identifiability of different word classes

Ariel (1990) holds the idea that, in Chinese proper nouns, personal pronouns, and nouns modified by definite determiners, the referent of a personal pronoun normally registers a higher degree of identifiability than that of the other two types (cf. Ariel, 1990). This postulation is checked by the data collected in the present study:

![Figure 6-12. Word classes and identifiabilities](image)

Figure 6-12. Word classes and identifiabilities

Studying the relationship between the types of identifiability and word classes, I do not distinguish among the specific types of accessibility; instead, I take the four sub-types (i.e. inferentially-accessible; shared knowledge-accessible; textually-accessible; situationally-accessible) as one big type –accessible to get a gradual hierarchy with the other three criteria (i.e. “unidentifiable”, “inactive” and “active”).

The distributions of the word classes in each type of identifiability are shown as:

Only bare nouns and modified nouns express the unidentifiable status of referents.
There are four types of word classes to express the accessible state of referents, and they are proper nouns, demonstrative nouns, bare nouns and modified nouns. Among the four classes, the modified noun takes the prominent place, and demonstrative nouns, due to their limited number, only express the accessible state of referents but not the other three categories.

The “active” data outnumbers the other three categories. There are two word classes that mainly express the activeness of the referent. They are proper nouns and, most outstandingly, the personal pronoun. All of the personal pronouns are used as the method of expressing the active state of referents. Most of the proper nouns are also used to express the active state of a referent.

Only by checking the distribution of each word class in all four statuses, we cannot figure out the degree of identifiability of each word class of the focus/topic in Chinese cleft constructions. To calculate and compare the degrees of identifiability of these grammatical expressions, we assign the respective integer to each identifiable status. The different degrees are due to the different identifiability of the referents being represented by these grammatical methods. I assign the “unidentifiable” referent the score “1” which means the lowest degree of identifiability among these four types; the “inactive” referent “2”, the “accessible” referent “3”, and the “active” referent the score “4”, representing the highest degree of identifiability. Due to the different total numbers of different word classes, here I check the ratio of each word class in the statuses of unidentifiable, inactive, accessible and active respectively, on the basis of which the degree of identifiability is calculated. The basis of the calculation is the ratio of each word class in the four types of identifiability statuses shown by the following charts:
Figure 6-13. The distributions of proper nouns having “inactive”, “active” and “accessible” statuses.

12.5% of the proper nouns are inactive, 12.5% of them are accessible, and 75% are active.

Figure 6-14. The distributions of personal pronouns having “inactive”, “active” and “accessible” statuses.

All of the personal pronouns are “actively” identifiable.
Figure 6-15. The distributions of demonstrative nouns having “inactive”, “active” and “accessible” statuses.

All of the “demonstrative+nouns” are “accessibly” identifiable.

Figure 6-16. The distributions of bare nouns having “inactive”, “active”, “accessible” statuses.
25% of the bare nouns are accessible; 25% of them are unidentifiable, and 50% of them are inactive.

Figure 6-17. The distributions of modified nouns having “inactive”, “active” and “accessible” statuses.

41.2% of modified nouns are unidentifiable; 11.8% of them are inactive and 47% of them are accessible.

The calculations of the degrees of identifiability of the referents represented by the grammatical word classes are shown by the following computations:

Proper nouns: $12.5\% \times 2 + 12.5\% \times 3 + 75\% \times 4 = 3.625$

Personal pronouns: $100\% \times 4 = 4$

Demonstrative nouns: $100\% \times 3 = 3$

Bare nouns: $25\% \times 1 + 50\% \times 2 + 25\% \times 3 = 2$
Modified nouns: $41.2\% \times 1 + 11.8\% \times 2 + 47\% \times 3 = 2.058$

The degrees of identifiability of proper nouns, personal pronouns, demonstrative nouns and bare nouns are displayed in the following chart:

Figure 6-18. Degrees of identifiability of referents represented by different grammatical expressions.

The types with the highest degree of identifiability of referents are personal pronouns. And proper nouns hold a very close degree of identifiability. The identifiability of the bare nouns is the lowest among all the possible word classes. This calculation partially agrees with Ariel’s (1990) study and the proper nouns and the personal pronouns have similar degrees of identifiability.

To sum up, among all the five word classes, personal nouns and proper nouns are the two types having the highest degree of identifiability, while bare nouns have the lowest degree of identifiability.
6.5.3 The identifiability of the thematic relations

In this section, I check the relationship between thematic relations and the identifiability of the referents represented by focus and topic constituents. What I check are the four statuses of identifiability—“unidentifiable”, "inactive", "active" and "accessible", and their thematic relations.

6.5.3.1 The identifiability of the thematic relations of focus constituents of Chinese cleft constructions

Thematic relations are divided into different groups according to their statuses of identifiability. The following chart is a summary of thematic relations with unidentifiable, inactive, accessible and active status respectively.
“Means” is the most significant thematic relation expressing the unidentifiable referents of focus constituents.

“Time” occupies a prominent position among these thematic relations. Interestingly, in the following two statuses, “accessible” and “active”, the thematic relation “time” also takes the first place compared with the “unidentifiable” status. This tells us that, at least in the present study, the thematic relation “time” is frequently used in expressing all three statuses of referents: “inactive”, “accessible” and “active”. The reason why it does not appear in the chart of “unidentifiable” is probably that “time” is a common state-of-affairs in daily life, and therefore it is quite hard to be counted as brand-new information for the listener.

There are only three types of thematic relations (i.e. “wrongdoer”, “source” and “time”) that represent “active” focus constituents, and just as in the case of “accessible” and “inactive” statuses, "time" takes the biggest portion in the category of “active”.

From these four charts, we can draw some conclusions for the present study about the relationship between the identifiability of the focus constituent and its thematic relations being part of the sentence:

(i) The thematic relation “time” is most likely to show the “active” status being the focus of the Chinese cleft construction. Besides, it is also an important thematic relation that expresses “accessible” and “inactive” statuses of the focus constituent.

(ii) “Means” is the most frequently expressed thematic relation to display an unidentifiable focus.
6.5.3.2 The identifiability of the thematic relations of topic constituents of Chinese cleft constructions

The topic constituents of Chinese cleft constructions are checked by the thematic relations they show in sentences. The detailed results are shown in Table F in the Appendix.

It is obvious that there is no such identifiable status of “inactive” or “unidentifiable” for topic constituents of Chinese cleft constructions, and thematic relations holding the active status are much more varied than those holding accessible status. The distributions of thematic relations in the statuses of “active” and “accessible” respectively are displayed below:
Figure 6-20. Thematic relations of the topic constituent and the status of “active” and “accessible”

The topic constituents realizing thematic relations of “theme” are “actively” identifiable. Besides “theme”, the second place is taken by “agent”.

The statuses of “theme”, “created-entity”, “activity”, “experiencer”, “event” and “competitors” have the same percentage with regard to their accessibility, which means their tendency of being “accessible” is similar.

Table F in the Appendix shows the thematic relation each topic constituent realizes, and compared with Table E, which gives the thematic relations of focus constituents, we find that, semantically, the thematic relations taken by topic constituents are the core elements of the frame. So we can conclude that, in Chinese cleft constructions, topic constituents are presented as core-elements of the sentence frames; while most of the focus constituents are the non-core elements of the sentence frames.

6.6 The configuration of the Chinese cleft construction from the perspectives of the information and syntactic structure

In this section, examples are given to illustrate the configurations of the information structure, the syntactic structure and the thematic relations.

(6-29) 我前天看了是《珍珠衬》嗬，评剧嗬，当时呢我们是晚上看的。

Wo qiantian kanle shi “zhenzhu chen”he, pingju he, dangshi ne women SHIaux wanshang kan DEptcl.
I the day before yesterday watch le “Pearl Liner” HE^{60}, Pingju HE, at that time NE we SHIauxnight watch De_{ptcl}.

I watched “Pear Liner” the day before yesterday. It is a Pingju. We watched it at night.

“Ne” is a mood particle that is used to mark a pause in an utterance. It does not have any syntactic or informational meanings. Here, “ne” is got rid of for the sake of simplicity.

**Table 6-6 Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (1).**

<table>
<thead>
<tr>
<th>Sentence</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>当时我们</td>
<td>是</td>
<td>晚上</td>
<td>看</td>
<td>的。</td>
<td></td>
</tr>
<tr>
<td>At that time, we</td>
<td>SHIaux</td>
<td>at night</td>
<td>watched</td>
<td>De_{ptcl}</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Information structure</th>
<th>Topic</th>
<th>Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presupposition_i</td>
<td>Focus</td>
<td>Presupposition_i</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Assertion</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>Adjunct</td>
<td>verb</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Syntactic categories</th>
<th>Personal pronoun</th>
<th>Auxiliary</th>
<th>Bare Noun</th>
<th>verb</th>
<th>Particle</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Identifiability</th>
<th>Active</th>
<th>Inactive</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Thematic relations</th>
<th>Perceiver</th>
<th>Time</th>
</tr>
</thead>
</table>

The focus expresses the inactiveness of its referent by the grammatical form of a bare noun. On one hand, it is discourse new; on the other hand, the referent of the time “晚上” is not specifically expressed or inferentially evoked by the hearer. The topic is the subject of the sentence whose meaning is actively referential to the hearer.

^{60}“HE” and “NE” are mood particles of Chinese.
(6-30)父母在这儿，我们打寿长街儿搬来的，寿长街儿就西半拉，福长街儿，寿长街儿，反正，不是那边儿都盖成大楼了吗？

Our parents are there, and we moved from Shouchang St. Shouchang St is a half street, while the other part is Fuchang St. Anyway, hasn’t that area been renovated into buildings?

我们打寿长街儿搬来的。

Women da Shouchangjier ban lai DEptcl.

We from Shouchangjier move DEptcl.

We moved from Shouchangjier.

Table 6-7 Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (2).

<table>
<thead>
<tr>
<th>Sentence</th>
<th>我们</th>
<th>打寿长街</th>
<th>搬来</th>
<th>的。</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>We</td>
<td>from Shouchang street</td>
<td>moved</td>
<td>DEptcl.</td>
</tr>
<tr>
<td>Information structure</td>
<td>Topic</td>
<td>Focus</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Presupposition</td>
<td>Focus</td>
<td>Presupposition</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assertion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic structures</td>
<td>Subject</td>
<td>Adjunct</td>
<td>Verb</td>
<td></td>
</tr>
<tr>
<td>Syntactic categories</td>
<td>Personal pronoun</td>
<td>Prepositional phrase (preposition + proper NP)</td>
<td>Verb</td>
<td>Particle</td>
</tr>
<tr>
<td>Identifiability</td>
<td>Active</td>
<td>Inactive</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Similar to the last example, the personal pronoun is also an identifiable grammatical form. “寿长街(Shouchang Jie)” is neither situationally, inferentially or textually evoked. However, under some specific circumstances, the referent of the nominal “寿长街(Shouchang Jie)” may be retrieved by the speaker if it is shared knowledge for the hearer, in which case, it would be accessible inferentially. For example, the hearer is also a local, or is very familiar with places in Beijing. If not, it is new information provided by the speaker.

(6-31)我知道他所说的‘公安部缉毒局的领导’所指是禁毒局的副局长、7.28专案指挥陈存仪。陈存仪是在抓住刘招华后的第一时间由北京飞抵福建的。

I know “the head of Ministry of public security Narcotics Bureau” mentioned by him refers to Chen Cunyi who is the associate director of Narcotics Bureau and the director of the “7.28 case”. Chen Cunyi flew back to Fujian from Beijing immediately after Liuzhaohua was caught.

Table 6-8. Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (3).

| Sentence | 陈存仪 | 是 | 在抓住刘招华后的第一时间 | 由北京飞抵福建 | 的 
|----------|--------|----|------------------------|----------------|-------
| Chen Cunyi | SHIaux | in | Immediately after Liu Zhaohua being caught | Flew back to Fujian from Beijing | 

| Information structure | Topic | Focus | Presupposition
|-----------------------|-------|-------|------------------
| Presupposition
| Focus | Presupposition
| 250 |
The proper noun “陈存仪 (Chen Cunyi)” is not only the main participant of the discourse but also mentioned and explained clearly by the previous closest clause, so it is active. The focus constituent modified by a relative clause “抓住刘招华的 (zhuazhu liuzhaohua de)”, so relatively, the time is specified by the restricted information provided by the modification. This case is called “containing inferrable” by Prince (1981, pp. 236-237), and the relevant issue is discussed also by Lyons (1999).

Du Bois (1980) discusses three main types of definite first mentions: (i) ones marked by an unstressed demonstrative; (ii) ones containing identifying information in a postmodifying relative clause, and (iii) ones due to association with a frame. NPs containing specific and new information in the presupposed format of a restrictive relative clause may also make a definite initial mention “acceptable”, as Du Bois (1980, p. 223) puts it, e.g. *she knocks the hat that he’s wearing off on the ground* (Du Bois 1980, p. 222). That the identifying information provided by defining relative clauses may motivate the use of a definite article for the whole

<table>
<thead>
<tr>
<th>Syntactic categories</th>
<th>Subject</th>
<th>Adjunct</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proper noun</td>
<td>Auxiliary</td>
<td>Prepositional phrase (preposition+modified noun)</td>
<td>Verb</td>
</tr>
<tr>
<td>Active</td>
<td>Inferentially accessible</td>
<td>Time</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Thematic relations</th>
<th>Theme</th>
</tr>
</thead>
</table>
NP is, again, well-established. About the third type, Du Bois (1980) states that a larger whole or a specific activity may serve as the ‘frame’ enabling the definite initial mention of referents typically associated with them, as with living room – the wall (Du Bois 1980, p. 233), sell – the money (Du Bois 1980, p. 215).

The modified nouns in the above sentence are just as the second case mentioned by Du Bois (1980), and the whole modified noun is partially evoked by part of the constituent itself.

(6-32) 继昨天进行两轮磋商之后，今天美方代表团三进三出中国外经贸部，最后的一次磋商是从下午三点四十分开始的，只进行了两小时二十分钟。

After the two rounds of negotiations yesterday, American representatives went to MOFTEC three times today. The last negotiation started from 3:40 pm and only lasted for 2 hours and 20 minutes.

**Table 6-9 Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic (4).**

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Information structure</th>
<th>Syntactic structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>最后的 一次 磋商 (the last negotiation)</td>
<td>Topic</td>
<td>Subject</td>
</tr>
<tr>
<td>是</td>
<td>Focus</td>
<td>Adjunct</td>
</tr>
<tr>
<td>从下午三点四十分</td>
<td>Presupposition</td>
<td>Verb</td>
</tr>
<tr>
<td>start</td>
<td>from 3.40 pm</td>
<td></td>
</tr>
<tr>
<td>Syntactic categories</td>
<td>Modified noun</td>
<td>Auxiliary</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Identifiability</td>
<td>Inferentially accessible</td>
<td>Shared knowledge accessible</td>
</tr>
<tr>
<td>Thematic relations</td>
<td>Event</td>
<td>Time</td>
</tr>
</tbody>
</table>

Between the communicators, the referent of the time is accessible because they share the mutual knowledge. Although “磋商 (negation)” is mentioned previously, however, the modifiers before it are different, so the two nouns are not identical instead, the second one can be inferentially evoked by the hearer.

(6-33) Li Xiaoyue said, after they got married, Liuzhaohua came back home every new year and every one or two months. However, the total time they spent together was less than two months. And they just got together “for real” when they started to escape.

两个人真正到一起还是从 1999 年逃亡生涯开始……

她记得刘招华是在 11 月 5 日凌晨回到武夷山别墅的。李小月问刘招华究竟出了什么事，刘招华只告诉李小月是由于税务方面出了事，让她别多问。

Li Xiaoyue said, after they got married, Liuzhaohua came back home every new year and every one or two months. However, the total time they spent together was less than two months. And they just got together “for real” when they started to escape.
She just remembered that Liuzhaohua came back to their house on Wuyi Mountain in the early morning of 5th of November. Li Xiaoyue asked Liu Zhaohua what happened, Liu just told her that something went wrong with his tax, and stopped her asking.

**Table 6-10. Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (5)**

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Information structure</th>
<th>Syntactic structures</th>
<th>Syntactic categories</th>
<th>Identifiability</th>
<th>Thematic relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>刘招华 (Liu Zhao hua)</td>
<td>是 S H I</td>
<td>在 11 月 5 日凌晨 in the early morning of 5th of November.</td>
<td>回到武夷山别墅 Went back to the house of Wuyi Mount.</td>
<td>的</td>
<td>的</td>
</tr>
<tr>
<td></td>
<td>Topic</td>
<td>Focus</td>
<td></td>
<td>D E</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Presupposition;</td>
<td>Focus</td>
<td>Presupposition;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Assertion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subject</td>
<td>Adjunct</td>
<td>Verb</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Proper noun</td>
<td>Auxiliary</td>
<td>Prepositional phrase (preposition+bare noun)</td>
<td>Verb</td>
<td>Particle</td>
</tr>
<tr>
<td></td>
<td>Textually accessible</td>
<td>Inferentially accessible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Theme</td>
<td>Time</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The topic is textually accessible because the previous referent does not appear within 2-3 clauses; like the last example, the focus here does not describe a specific time, but the hearer can retrieve the referent by inference in accordance with the whole passage.

(6-34) 我说，我三十来岁时候，国民党要完还没完时候，粮食也是一天长好几回价儿啊，还买不着。挣、这、挣这工资，追不上他、追不上他那物价，追不上他那、他那粮价啊！我现在这么说的都录下来了吧？我说这就是。是实际情况。哎，我、我本身是由那儿、由那儿、从…可以说是爬过来的，我由那儿爬过来的。

I said, when I was in my thirties, Kuomintang was still in power. At that time, the price of food increased several times each day, and the supply of food was too limited to get. The income could not cover the commodity price and the food price! Did you record what I said? What I said is fact. I, myself, was crawling from there, I was crawling from there.

Table 6-11. Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (6)

<table>
<thead>
<tr>
<th>Sentence</th>
<th>我</th>
<th>由那儿</th>
<th>爬过来</th>
<th>的</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information structure</td>
<td>Topic</td>
<td>Focus</td>
<td>Presupposition</td>
<td>DE.</td>
</tr>
<tr>
<td>Presupposition</td>
<td>Focus</td>
<td>Presupposition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assertion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic structures</td>
<td>Subject</td>
<td>Adjunct</td>
<td>Verb</td>
<td></td>
</tr>
<tr>
<td>Syntactic categories</td>
<td>Personal pronoun</td>
<td>Prepositional phrase (preposition + demonstrative NP)</td>
<td>Verb</td>
<td>Particle</td>
</tr>
</tbody>
</table>
“那儿 (near; there)” refers to the time governed by the nationalist party of China, which is not mentioned but inferred from the observation of the previous discourse.

(6-35) 哎，后来他也学着请安，那阵儿要像我，我大大他们活着的时候，她活着现在 都将近九十岁了。那会儿，她文化大革命前死的，那会儿八十岁。

Then he learned to bob a curtsey when my aunt was still alive. If she were still alive now, she’d be almost 90 years old. But back to that time, she died before the Culture Revolution.

Table 6-12. Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (7).

<table>
<thead>
<tr>
<th>Sentence</th>
<th>她</th>
<th>文化大革命前</th>
<th>死</th>
<th>的</th>
<th>DE</th>
</tr>
</thead>
<tbody>
<tr>
<td>她 (She) 文化大革命前 (before the Culture revolution) 死 (died) 的 (DE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information structure</td>
<td>Topic</td>
<td>Focus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presupposition</td>
<td>Focus</td>
<td>Presupposition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assertion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic structures</td>
<td>Subject</td>
<td>Adjunct</td>
<td>Verb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic categories</td>
<td>Personal pronoun</td>
<td>Prepositional phrase (preposition+ proper noun)</td>
<td>Verb</td>
<td>Particle</td>
<td></td>
</tr>
</tbody>
</table>
“她 (ta; she)” refers to “我大大 (wo dada; my aunt)” in the previous clause (within a 2-clause distance), so it is an active referent.

(6-36) 这是个死胡同，N，是个，或者以前是活胡同，后来堵死的。这条大街，解放以后才修好的，以前没这么好。

This is a “Hutong” with a dead end, or maybe was not a dead end before and has since been blocked later. This street was finished after the Liberation, and it was not this well-built before.

Table 6-13. Demonstration of examples from the perspectives of information, syntactic structure, word classes, identifiability and thematic relations (8).

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Information structure</th>
<th>Syntactic structures</th>
<th>Syntactic categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>这条大街，解放以后才修好的</td>
<td>Topic</td>
<td>Subject</td>
<td>Demonstrative NP</td>
</tr>
<tr>
<td></td>
<td>Focus</td>
<td>Adjunct</td>
<td>Prepositional phrase (preposition+proper noun)</td>
</tr>
<tr>
<td>这条大街，解放以后才修好的的</td>
<td>Focus</td>
<td>Verb</td>
<td>Verb</td>
</tr>
<tr>
<td></td>
<td>Presuppositioni</td>
<td></td>
<td>Prepositional phrase (preposition+proper noun)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Particle</td>
</tr>
</tbody>
</table>
The topic “这条大街 (zhetiao dajie; this street)” is situationally accessible with the exact location where the communication happens and is shared by both the speaker and the listener. And the focus “jiefang yihou (after the Liberation)” is a prepositional phrase including a proper noun which is an active referent for the listener.

(6-37) 我在那儿住不少，了十几年了，搬到东城区又十几年了，我是七一年退休的，七一年从干校回来的。

I lived there for more than 10 years, and moved to Dongcheng District and lived there for more than another 10 years. I retired in 1971, and I was back from the cadre school in 1971.

**Table 6-14. Demonstration of examples from the perspectives of information, syntactic structures, word classes, identifiability and thematic relations (9).**

<table>
<thead>
<tr>
<th>Sentence Structure</th>
<th>Information structure</th>
<th>Syntactic structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>我</td>
<td></td>
<td>Subject</td>
</tr>
<tr>
<td>是</td>
<td>SHI</td>
<td>Adjunct</td>
</tr>
<tr>
<td>七一年</td>
<td>YEAR 1971</td>
<td>Verb</td>
</tr>
<tr>
<td>退休</td>
<td>retired</td>
<td></td>
</tr>
<tr>
<td>的</td>
<td>DE</td>
<td></td>
</tr>
<tr>
<td>Information structure</td>
<td>Topic</td>
<td>Focus</td>
</tr>
<tr>
<td></td>
<td>Presupposition</td>
<td>Focus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Presupposition</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assertion</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Identifyability</th>
<th>Situationally accessible</th>
<th>Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thematic relations</td>
<td>Protagonist</td>
<td>Time</td>
</tr>
</tbody>
</table>
Both the topic and the focus constituent are active from the perspective of identifiability.

6.7 The markedness of Chinese cleft constructions from the perspective of focus typology

Inspired by the Prague School’s works on communicative dynamism (e.g. Firbas, 1964, 1966; Mathesius, 1928, 1929; Sgall, et al., 1986), there emerged in the late 1990s an interest in classifying the way languages grammatically organize the expression of the various information statuses that elements within a sentence may have, e.g. Dik (1989), Bearth (1992), Vallduví (1992), Lambrecht (1994). For example, Kiss (1987, 1994) and Vallduví (1992) have noted some typological differences among languages with respect to focus structure.

According to Van Valin (1999), one distinction is whether or not the language has a rigid syntax, and the other is whether or not it has a rigid focus structure: “languages in which the potential focus domain is the entire main clause in simple sentences will be considered to have flexible focus structure, whereas those in which the potential focus domain is restricted to a subpart of the main clause will be considered to have rigid focus structure” (pp, 3-4).

As far as I am concerned, there is still no relevant study talking about the Chinese focus from a typological point of view. From a general point of view, Chinese is a language with flexible
focus structure, in which the potential focus domain should be any constituent or the whole main clause of the simple sentence.

For example, in the following sentences, focus falls on different constituents ranging from the whole VP or the whole clause to the subject, or the direct NP alone.

(6-38) 我是该走了。（emphatic sentence—focus domain: whole sentence）

Wo SHI aux gai zou LE.
I SHI aux have to go LE.
I have to go indeed.

(6-39) 我是昨天到北京的。（the cleft construction—adjunct phrase）

Wo SHI zuotian dao Beijing de.
I SHI yesterday arrive Beijing DE.
It was yesterday that I arrived in Beijing.

(6-37) 苹果在桌上。（topicalized sentence—focus domain: prepositional predicate）

Pingguo zai zhuoshang.
Apple on table.
The apple is on the table.

(6-40) 桌子上有苹果。（existential sentence—focus domain: complement）

Zhuozi shang you pingguo.
Table on have apple.
There is an apple on the table.

(6-41) 我捡起来的是钢笔。（pseudo-cleft construction—focus domain: object）

Wo jian qilai DE rel SHI cop gangbi.
I pick up DE rel SHI cop pen.
What I picked up is a pen.
Weile zheci yuehui, wo shenzhi qu mai LE pixie.
For this date, I even go buy LE leather shoes.
For this date, I even went to buy leather shoes.

We can simply assume that Chinese is a flexible focus structure and cleft constructions, just as topicalized sentences, existential sentences, emphatic sentences, pseudo-cleft sentences, contrast clauses are all marked constructions.

Although the potential focus domain of Chinese sentences is quite randomly assigned, the focus domain of Chinese cleft constructions is not. The cleft constructions collected in the present study show that all the focus constituents of the “shi...de” cleft sentences are the constituents directly following the focus marker “shi”, although it is widely accepted that the constituents between “shi” and “de” all hold the potential to be the focus constituent of Chinese cleft constructions. All of the the constituents directly following “shi” are acting as the focus of the cleft constructions in the present study.

According to the two criteria propsed by Keenan (1975) (see footnote 9 on page 23), we can compare the Chinese cleft construction with the sentence B mentioned by Keenan, whereas sentence A is the canonical or unmarked sentence that is obtained by deleting the focus constituent as well as the auxiliary “shi” and particle “de” from the cleft construction.

6.8 Summary
Among all the qualifying cleft constructions in Chinese, there is only one argument-focus structure taking the argument of the sentence as the focus constituent, and the others (97.5%) are all adjunct-focus cleft constructions taking the adjunct of the sentence as the focus constituent. Unlike the English counterparts, the object of the sentence cannot be treated as the focus of the cleft construction. The result also shows that unlike English, most of the
Chinese cleft constructions are adjunct-focus structures, which is a brand new category established for Chinese cleft constructions in the present study.

Considering the syntactic categories of the focused constituents of Chinese cleft constructions, prepositional phrases take the dominant position; and then noun phrases take the second place; adverbial and verb phrases have a similar ratio, and adjective phrases hold the lowest percentage.

The thematic relations of focus and topic constituents vary. Among all these thematic relations, the top three thematic relations of focus constituents with the most frequency are “time” and “means”. “Means” are frequently used as “unidentifiable” focus; “time” is both accessibly and actively as well as inactively identifiable. The thematic relations of focus constituents are mainly the non-core elements of the frame of the sentence; however the thematic relations realized by the topic constituents are mainly the core elements of the sentence frames. There are only two identifiable statuses for topic constituents—“active” and “accessible”, and “theme” plays the dominant position among the “active” thematic relations.

Considering the identifiability of the focus and topic constituent, one result of the present study shows that the topic is never an inactive or an unidentifiable referent; and the gradual degree of identifiability of the topic constituent proves the Topic Acceptability Scale proposed by Lambrecht (1994). For the focus constituents, there is no significant difference between the four statuses—“unidentifiable”, “active”, “inactive” and “accessible”. However, the tendency is still reflected by the slight numerical difference between them. From “unidentifiable” to “active”, the frequency of “topic” reduces gradually.

So the present study shows the acceptability of the topic as:
Active                                                           most acceptable
Accessible
Inactive
Unidentifiable                                                least acceptable

The identifiability of the focus constituent, in the present study, shows in a converse order:

Unidentifiable                                             most acceptable
Inactive
Accessible
Active                                                         least acceptable

Considering the relationship between the degree of identifiability and the word classes, the present study partially proves the conclusion proposed by Ariel (1990) that personal pronouns have a higher degree of identifiability than other word classes, for example, bare NPs or demonstrative NPs. However, in the present study, proper nouns also have similar (i.e. slightly lower) degree of identifiability to personal pronouns. And bare nouns have the lowest degree of identifiability as part of the focus constituent of the Chinese cleft construction.

Finally, from a focus typological point of view, cleft constructions in Chinese are considered as marked constructions of Chinese that hold more rigid focus structures.
Chapter 7

Conclusion

This thesis has been concerned with cleft constructions in Chinese. It deals with two main problems: the structural and semantic properties of cleft constructions in Chinese; and the properties of their information structure.

Based on the corpus established by the Center for Chinese Linguistics of Peking University, the qualifying data are collected firstly by indexing the specific sentence strings “shi…de”, “shi…” and “…de…”; and then the non-qualifying data are excluded on the basis of the data collected from the first step in accordance with the criteria proposed by the present study. All the qualifying data are further analyzed from the perspective of their syntactic forms and their information structure. The conclusions are stated in the following sections.

7.1 The structures of cleft constructions in Chinese

To distinguish cleft constructions from other superficially similar constructions, the semantic and informative properties of cleft constructions are treated as the criteria of Chinese cleft constructions. The following criteria are the feedback of the first research hypothesis in section 1.5, and the first research question in section 1.4:

(i) Just as with most of the general sentence, there is a bipartition of “presupposition” and “focus” in Chinese cleft constructions; however, the configuration and the properties of the presupposition and focus are different from other sentences. Firstly, the presupposition of the Chinese cleft construction is a proposition (although the forms of the presupposition of argument-focus cleft constructions and adjunct-focus cleft constructions vary slightly). Secondly, the boundary between the presupposition and the focus is not a clearly demarcated one. And
thirdly, the focus is not the obligatory element of the sentence; in other words, the omission of the focus constituent does not make the sentence ungrammatical. Finally, “shi” and “de” are not obligatorily involved in either the presupposition or focus.

(ii) The focus constituent should be exhaustive.

(iii) The focus constituent should be contrastive.

Following these criteria, the cleft constructions can be distinguished from copular sentences of Chinese, which do not hold any of the three properties above, and from emphatic “shi” sentences, which do not hold the exhaustive and contrastive focus.

In response to the second research hypothesis of section 1.5 (also the second research question of section 1.4), the present study supports the widely-accepted idea that the prototypical structure of the Chinese cleft construction is the sentence form of “shi…de”. However, unlike some of the previous studies (e.g. Fang, 1995; Huang & Fawcett, 1996; Shang, 2002) which take bare “shi” structure as the main variant of Chinese cleft constructions, the result of the present study shows that, in fact, bare “de” structure is also one variant of Chinese cleft constructions, and it is the main variant of Chinese cleft constructions with a slightly higher frequency of appearance.

By observing the relationship between the argument and predicate, and the cases of omission of “shi” and “de”, I have concluded that in Chinese cleft constructions, “shi” is a raising auxiliary that accepts the subject-auxiliary inversion; “de” is the sentence-ending particle whose main function is to enhance the modality of the sentence. Under certain circumstances (but not necessarily), it also reflects the “tense/aspect” information of the sentence. In cleft constructions in Chinese, “shi” can usually be omitted rather freely;
however, if “de” shows the tense/aspect property of a sentence holding the past-tense and past perfect aspect, it cannot be omitted. “Shi” and “de” act differently in copular sentences (which are subcategorized into identifying sentences, classifying sentences, explanatory sentences and attributive sentences) but “shi” and “de” in emphatic sentences hold the same syntactic categories as in cleft constructions. “Shi” in copular sentences is the copula of the sentence and “de” is the modification or the modal particle (i.e. in explanatory sentences).

With regard to the third hypothesis (and the fourth research question), this study concludes, from the empirical data, that VPs can be focused in Chinese (being the adjunct of the sentence). And among all the syntactic categories of focus constituents (i.e. AP, AdvP, NP, PP, VP), the PP takes the prominent position with the highest frequency of occurrence.

On the distribution of cleft constructions in different types of contexts, it is shown by the present study that Chinese cleft constructions have higher frequency in informal texts than in the formal ones. But generally speaking, the cleft construction is such a marked sentence structure that its frequency of occurrence is only about 0.4%.

7.2 The configuration of the information structure

Considering the types of focus structures, based on the corpus, Chinese cleft constructions are divided into two types—the argument-focus cleft construction and the adjunct-focus cleft construction. Most Chinese cleft constructions are adjunct-focus cleft constructions. The adjunct-focused constructions highlight the adjunct of the sentence (i.e. the time and manner, etc. of the action expressed by the main verb). This result is against my fourth hypothesis, because by observation we can find numerous examples of argument-focus cleft constructions; however in fact, at least in the present corpus, the numbers of argument-focus cleft constructions are quite limited. Instead, adjunct-focus cleft constructions appear more frequently.
The differences between argument-focus cleft constructions and adjunct-focus cleft constructions mainly fall into two aspects: firstly, the focus domain of the cleft construction is different; and secondly, the configuration of the presupposition is also different. In adjunct-focus cleft constructions, the presupposition is the proposition established by removing the focus constituent as well as “shi” and “de” from the sentence; whereas in argument-focus cleft constructions, the presupposition is an existential proposition established partially from the original sentence and partially by the automatic supplementation of the missing subject by the listener. For example:

(7-1) 是小明打破窗户的。

SHIaux Xiaoming dapo chuanghu DEptcl.

SHIaux Xiaoming break window DEptcl.

It is Xiaoming who broke the window.

The presupposition of the sentence is: “有人打破窗户.”(There is a person (or X) who broke the window.)

The presupposition is made by making the predicate of the original sentence the subclause of the sentence, and the addition of the missing subject (X) by the listener.

7.3 The thematic relations of focus/topic constituents

The research question of the semantic difference between focus and topic constituents is explained by their thematic relations. The thematic relations realized by focus and topic constituents are checked to find some underlying rules of the focus and topic constituents from the semantic perspective. “Time” is the thematic relation most frequently focused and topicalized; “means” ranks second in the focus constituent. And the thematic relations of
focus constituents are mostly the non-core elements of the frame set by the sentence; whereas the thematic relations of topic constituents are mainly the core-elements of the sentence frame.

**7.4 The identifiability of focus/topic constituents**

To answer the third research question in section 1.4, the identifiability of the focus constituents in adjunct-focus cleft constructions is conducted on the basis of the NPs involved in the focus constituents. The results include two aspects. Firstly, there is a different tendency of the identifiability of the focus constituent and that of the topic constituent. The result confirms Lambrecht’s (1994) scale of topic acceptability which states that “active” is the most acceptable status for a topic, while “brand-new unanchored” is the least acceptable status. The result does not show the trend to support the scale of focus acceptability; however, the result does show the opposite tendency in the case of focus. In other words, the “unidentifiable” status is the most acceptable status for focus constituents, and “active” is the least acceptable status.

Secondly, grammatical expressions and their identifiability are also discussed. I have collected all the topic and focus constituents, and classified them by their degrees and types of identifiability according to the context in which they are involved. The results show that personal nouns and proper nouns are the two types having the highest degree of identifiability, and bare nouns have the lowest degree of identifiability.

**7.5 Significance**

This study is the first systematic and comprehensive study of cleft constructions in Chinese from the perspectives of their forms, semantic properties and information structure. It examines the prototypical structure of the Chinese cleft construction and its variants. The study proposes the criteria of Chinese cleft constructions from the semantic and information-
structure points of view for the first time. The criteria avoid the typological confusion caused by directly applying the criteria of English cleft constructions to the Chinese counterparts by method of translation; and at the same time, the criteria further distinguish Chinese cleft constructions from other syntactically similar but functionally different structures (i.e. copular sentences and “shi” emphatic sentences). I believe that the criteria can be used as the new judgement of Chinese cleft construction, and by using them, we can avoid many ambiguous cases due the incomplete or misleading understandings (for example, the confusion of cleft constructions and copular sentences, etc.) that may be provided by previous studies or just by intuition. The criteria can be considered as general properties of cleft constructions cross-linguistically, and can be treated as general judgmental criteria of cleft constructions for languages like Chinese, in which cleft constructions can hardly be judged by syntactic structures.

Chinese cleft constructions are categorized, according to their focus domains, into argument-focus cleft constructions and adjunct-focus cleft constructions, and adjunct-focus cleft constructions are used more frequently than argument-focus cleft constructions. This study represents the first attempt, so far as I know, to classify Chinese cleft constructions from focus structures and to check the actual usage of these types in corpus. Besides, this study classifies “shi...de” constructions in a rather systematic way based on the empirical data, and the result of the classification explains the relationship between the copular sentence, the emphatic sentence and the cleft sentence in Chinese.

This dissertation has established the relationship between different components of information structure by checking their identifiability. Instead of only checking the argument focus whose form is a NP, I have a check the nominal phrases projected from the focus constituents acting as the adjunct of the adjunct-focus cleft construction. Interestingly, besides finding underlying semantic rules about the topic/focus constituents, I have found a
way to check the identifiability of different grammatical/lexical constituents (i.e. bare nouns, personal nouns, etc.). So, the results also show that the degree of identifiability of different word classes is also eligible for the focus/topic constituent of Chinese cleft constructions.

7.6 Future directions of the study

Phonological factors are to be included in the future study. Phonological means are important parameters of the information structure. But because this study is an analysis of information structure on the basis of the syntactic structure of cleft constructions in Chinese, the phonological element is avoided. In the future studies, phonological parameters are to be checked, provided that there is enough scientific equipment (e.g. the acoustics laboratory) or sufficient data sources (e.g. the oral corpus). Taking phonological parameters into consideration leads us to a new research direction of Chinese cleft constructions—the information structure of Chinese cleft constructions in daily communications.

More data, especially the data of bare "shi" and bare "de" cleft constructions are to be involved in order to present a clearer picture of the variants of Chinese cleft constructions. With the help of more empirical data, a more refined classification of "shi...de" sentences can be proposed; the underlying rules of the syntactic categories of focus constituents and the omission of "shi" or "de" could be explained more comprehensively.
Appendices
Table A. The classification of “shi...de” sentences collected from the present corpus

<table>
<thead>
<tr>
<th>Subtypes</th>
<th>Word class of “shi”</th>
<th>Word class of “de”</th>
<th>Omission of “shi” and “de” at the same time</th>
<th>Omission of “shi” only</th>
<th>Omission of “de” only</th>
<th>exhaustivity</th>
<th>Contrastiv- eness</th>
<th>Semantic functions</th>
<th>Sentence structures</th>
<th>presupposition</th>
<th>Focus</th>
<th>Focus type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copular sentences</td>
<td>Identifyi ng</td>
<td>Copular</td>
<td>Relative pronoun</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>the referential nominal predicate identifies the referential subject by establishing the identical relationship between them linked by the copula</td>
<td>subject+shi+(referential) nominal predicate (VP +de)</td>
<td>Presupposition=topic</td>
<td>Focus=nominal predicate (with “de”involved, and adding the omitted object if necessary)</td>
<td>Predicat-focus structure</td>
</tr>
<tr>
<td>Classifying</td>
<td>Copular</td>
<td>Relative pronoun or nominalizer</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>the nonreferential nominal predicate classifies the referential subject by establishing</td>
<td>subject+shi+(nonreferential)nominal predicate( VP +de)</td>
<td>Presupposition=topic</td>
<td>Focus=nominal predicate (with “de”involved, and adding the omitted object if necessary)</td>
<td>Predicat-focus structure</td>
</tr>
<tr>
<td>Clause Type</td>
<td>Copular</td>
<td>Modal Particle</td>
<td>Explanatory Information</td>
<td>Attributional Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>---------</td>
<td>----------------</td>
<td>-------------------------</td>
<td>---------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attributive</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- g the categorical / material etc. relationship between them linked by the copula
- the predicate expresses the explanatory information about the subject from aspects of reasons, purpose etc.
- the predicate expresses the attributive properties about the subject
<p>| Empathic sentences | Sentence - emphatic sentence | (Raising) auxiliary | (Modal) particle | Yes | Yes | No | No | No | The whole sentence or the proposition it delivers is emphasized by highlighting the mood and modality of the sentence | Subject+sh i+clause/ VP+de; Shi+subjec t+clause/ VP+de | Presupposition= Ø | Focus= The whole sentence | Sentence-focus structure |
|------------------|----------------------------|--------------------|----------------|-----|-----|----|----|----|---------------------------------------------|-----------------------------------------------|-----------------|------------------|-----------------|-------------------|
| Modal-emphatic sentence | (Raising) auxiliary | (Modal) particle | Yes | Yes | No | No | No | The whole sentence or the proposition it delivers is emphasized by highlighting the mood and modality of the sentence | Subject+sh i+clause/ VP+de; Shi+subjec t+clause/ VP+de | Presupposition= Ø | Focus= The whole sentence | Sentence-focus structure |</p>
<table>
<thead>
<tr>
<th>Cleft sentences</th>
<th>(Raising) auxiliary</th>
<th>(Modal) particle</th>
<th>Yes</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>The sentence puts some part of the sentence as the contrastive focus of the sentence</td>
<td>Subject+sh i+vp/ advp/ pp/ clause+de; Shi+subject+clause+de.</td>
<td>Presupposition=a presupposed proposition=assertion-focus</td>
<td>Focus=some part of the sentence: argument, adjunct etc.</td>
<td>Argument-focus structure; adjunct; vp etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name of the frame</td>
<td>Lexical unit</td>
<td>Frame element of the focus constituent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------------</td>
<td>----------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quitting</td>
<td>退休(tuixiu; retire)</td>
<td>Manner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity_finish</td>
<td>毕业(biye; graduate)</td>
<td>Place</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception_active</td>
<td>看(kan; watch)</td>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Certainty</td>
<td>相信(xiangxin; believe)</td>
<td>Manner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intentionally_act</td>
<td>做(zuo; do)</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arriving</td>
<td>回来(huilai; come back)</td>
<td>Source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>赶到(gandao; arrive)</td>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>回到(huidao; arrive)</td>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Departing</td>
<td>走(zou; walk)</td>
<td>Circumstance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>撤回(chehui; go back)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motion</td>
<td>飞抵(feidi; fly to)</td>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bringing</td>
<td>带过去(dai huiqu; bring back)</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giving</td>
<td>给(gei; give)</td>
<td>Place</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>Chinese Character</td>
<td>English Translation</td>
<td>Part of Speech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>-------------------</td>
<td>---------------------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education teaching</td>
<td>教育(jiaoyu; educate)</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success_or_failure</td>
<td>失败(shibai; fail)</td>
<td>Explanation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research</td>
<td>探底(tandi; search)</td>
<td>Manner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity_start</td>
<td>起来(qilai; start)</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death</td>
<td>死(si; die)</td>
<td>Cause</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renting</td>
<td>承包(chengbao; hire)</td>
<td>Money</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity_ongoing</td>
<td>进行(jinxing; operate)</td>
<td>Depictive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seeking</td>
<td>找(zhao; find)</td>
<td>Manner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questioning</td>
<td>请教(qingjiao; ask)</td>
<td>Manner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Process_start</td>
<td>开始(kaishi; start)</td>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finish_competition</td>
<td>取胜(qusheng; win)</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tolerating</td>
<td>忍(ren; bear)</td>
<td>Degree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reporting</td>
<td>告起(gaoqi; sue from)</td>
<td>Wrongdoer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling</td>
<td>感到(gandao; feel)</td>
<td>Cause</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forming_relationship</td>
<td>结婚(jiehun; marry to)</td>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manipulate_into_doing</td>
<td>骗取(pianqu; cheat)</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Make_acquaintance</td>
<td>碰到(pengdao; meet)</td>
<td>Place</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telling</td>
<td>告诉(gaosu; tell)</td>
<td>Manner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Role</td>
<td>Event</td>
<td>Time</td>
<td>Source</td>
<td>Manner</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>----------------------</td>
<td>----------</td>
<td>--------</td>
<td>--------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Building</td>
<td>修好 (xiuhao; built)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bringing</td>
<td>搬来 (banlai; move)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self_motion</td>
<td>爬过来 (paguolai; crawl to)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telling</td>
<td>告诉 (tell)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focus constituents</td>
<td>Frame element (thematic relations) of the focus</td>
<td>Syntactic category</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>申请（shenqing; apply）</td>
<td>Manner</td>
<td>V^{61}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>七一年(71 nian, year (1971)</td>
<td>Time</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>六十五中(65 Zhong; No. 65 middle school)</td>
<td>Place</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>晚上(wanshang; night)</td>
<td>Time</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>绝对(juedui; absolutely)</td>
<td>Manner</td>
<td>Adv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>有意(youyi; deliberately)</td>
<td>Purpose</td>
<td>Adv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>从英国(cong Yingguo; from Britain)</td>
<td>Source</td>
<td>PP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>在非正常状态下(zai feichang zhuangtai xia; under special circumstance)</td>
<td>Circumstance</td>
<td>PP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>在抓住刘招华后的第一时间(zai zhuazhu Liuzhaohua de diyi shijian; as soon as Liuzhaohua being caught)</td>
<td>Time</td>
<td>PP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>装在一个方便面的纸箱内(zhuang zai yige fangbianmian de zhixiang nei; put it in a carton of instant noodles)</td>
<td>Means</td>
<td>VP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>在结婚的当日(zai jiehun de dangri; at the day they got married)</td>
<td>Time</td>
<td>PP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^{61} The lexical category and the phrasal category are shown separately in the table, such as “V” and “VP”, “N” and “NP”. However, for the sake of convenience, in the figures, I include the all “V” into the category “VP” and “N” into “NP”.

279
<p>| 在事发的前一天(zai shifa de qianyitian; at the day before the incident happened) | Time | PP |
| 在刘招华包租的广州总统大酒店 818 总统套房(zai Liuzhaohua baozu de Guangzhou zongtong dajiudian 818 zongtong taofang; at the 818 suite of the Presidential Hotel which was hired by Liuzhaohua ) | Place | PP |
| 在 11 月 5 日凌晨(zai 11 yue 5 ri lingchen; in the early morning of 5th of November) | Time | PP |
| 用理念(yong linian; by faith) | Means | VP |
| 集体(jiti; group) | Explanation | N |
| 亲自(qinzi; by oneself) | Manner | Adv |
| 打着宗教旗帜 (dazhe zongjia qizhi; under the banner of religion) | Means | VP |
| 老(lao; old) | Cause | Adj |
| 以最低造价 30 万元(yi zuidi zaojia 30 wan yuan; by the lowest cost of 300,000 yuan) | Money | PP |
| 在 WTO 框架之外(zai WTO kuangjia zhi wai; beyond the frame of WTO) | Depictive | PP |
| 有心装扮作陪嫁奴隶 (youxin zhuangban zuo peijia nuli; dressing as a slave deliberatly) | Manner | VP |</p>
<table>
<thead>
<tr>
<th>专程(zhuancheng; on purpose)</th>
<th>Manner</th>
<th>Adv</th>
</tr>
</thead>
<tbody>
<tr>
<td>从下午三点四十分(cong xiawu sandiansishifen; from 3:40pm)</td>
<td>Time</td>
<td>PP</td>
</tr>
<tr>
<td>靠无可比拟的国籍知名度(kao wukebini de guoji zimingdu; by the uncomparable international reputation)</td>
<td>Means</td>
<td>PP</td>
</tr>
<tr>
<td>很(hen; quite)</td>
<td>Degree</td>
<td>Adv</td>
</tr>
<tr>
<td>从那个国民党伪总统(cong nage Guomindang Wei zheng fu; from the president of Kuomintang puppet government)</td>
<td>Wrongdoer</td>
<td>PP</td>
</tr>
<tr>
<td>在刘招华跟他聊到这一节的时候(zai Liuzhaohua gen ta liao dao zhe yi jie de shihou; when Liuzhaohua told this problem to him)</td>
<td>Cause</td>
<td>PP</td>
</tr>
<tr>
<td>四八年(48 year; year (19)48)</td>
<td>Time</td>
<td>N</td>
</tr>
<tr>
<td>在觉悟状态(zai juewu zhuangtai; under awareness)</td>
<td>Means</td>
<td>PP</td>
</tr>
<tr>
<td>在明知供销社委托他提运的120箱白酒已全部提运完毕的情况下(zai mingzhi gongxiaoshe weituo ta tiyun de 120 xiang baijiu yi quanbu tiyun wanbi de qingkuang xia; under the circumstance that he knew that the all the 120 cartons of alcohol have already being transported.)</td>
<td>Means</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>刚才在路上(gangcai zai lushang; just now on the way)</td>
<td>Place</td>
<td>PP</td>
</tr>
<tr>
<td>亲嘴(Laowang qinzui; Laowang by himself)</td>
<td>Manner</td>
<td>AdvP</td>
</tr>
<tr>
<td>走着(zouzhe; by walk)</td>
<td>Means</td>
<td>VP</td>
</tr>
<tr>
<td>夜里一点(yeli yidian; 1 am at night)</td>
<td>Time</td>
<td>NP</td>
</tr>
<tr>
<td>文化大革命前(Wenhuadageming qian; before the Great Culture Revolution)</td>
<td>Time</td>
<td>PP</td>
</tr>
<tr>
<td>解放以后(jiefang yihou; after the Liberation)</td>
<td>Time</td>
<td>PP</td>
</tr>
<tr>
<td>打寿长街(da Shouchang jie; on the Shouchang Street)</td>
<td>Source</td>
<td>PP</td>
</tr>
<tr>
<td>由那儿(you naer; from there)</td>
<td>Source</td>
<td>PP</td>
</tr>
<tr>
<td>word classes</td>
<td>Bare nouns</td>
<td>Proper nouns</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Unanchored unidentifiable</td>
<td>Focus (以（最低造价 30万）/ zui di zaojia 30 wan; the lowest price—300,000)</td>
<td>Focus (用(理念); yong linian; by the faith)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anchored unidentifiable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inactive/ nonidentifiable</td>
<td>Focus (晚上/ wanshang; night)</td>
<td>Focus (六十五中/ liushiwuzhong; No. 65 middle school)</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Focus (集体/ jiti; group)</td>
<td>Focus ((打)寿长街; Shouchangjie; Shouchang st.)</td>
<td>Focus (靠(无可比拟的国际知名度)/ incomparable international reputation)</td>
</tr>
<tr>
<td>Focus (在(觉悟状态)/ zai juewu zhuangtai; under wareness)</td>
<td>Focus (在(路上)/ zai lushang; on the way)</td>
<td></td>
</tr>
<tr>
<td>Inferentially accessible (self-containing)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
association and frame-containing association

62 Self-containing association means the identifiability of the normal expressions is established on the basis of the information that is contained in the nominal itself. It is considered as a type of inferential accessibility. It is often used to show the relationship between the head noun phrase and the modifier. To quote two examples from Ina Rosiger (2014: 47): “The structure of the protein”; “the thoracic circuit state in HK mutants.”

63 The concept of the “frame-based” association is also borrowed from Chen (2004). It is a kind of association that makes the mention of some constituents “triggers the identifiability of all the things that are typically associated with it” (Chen, 2004, p.1144).

64 Topic 1 is the external topic of the sentence, the identifiability of the external topic is not discussed relating to their word classes. Chafe (1976, p.50-51) refers to the topic1 and topic2 by "external" and "internal" topics. According to Chafe (1976) and Li & Thompson (1976), clause-external topics are characterized as expressions "a spatial, temporal or individual framework within which the main predication holds". So they are normally represented by adverbials. On the contrary, clause-internal topics are typical "aboutness" topics in the sense mentioned above. The referent of the "aboutness" is indispensably a thing or an event whose corresponding NP plays an argument role in the proposition. Only the clause-internal topic is discussed to show the relationship between the word classes of referents and the identifiability of them within a minimal clause. However, the clause-external and internal topic of a sentence are discussed from the perspective of their degrees of identifiability aiming to find out the underlying law of the configuration of the multi-topic sentence.

<table>
<thead>
<tr>
<th>association and frame-containing association</th>
<th>the negotiation of airline agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Topic (最后一次磋商/ zuihou yici cuoshang; the last negotiation)</td>
</tr>
<tr>
<td></td>
<td>Focus (（在）WTO 框架之外) / WTO kuangjie; the frame of WTO)</td>
</tr>
<tr>
<td>Shared knowledge-accessible</td>
<td>Focus (（从）下午三点四十分/ (cong) xiawu sandian sishi fen; at 3:40 pm)</td>
</tr>
<tr>
<td></td>
<td>Topic (哈佛与耶鲁等美国长春藤名校/ hafo yu yelu deng meiguo changchunteng mingxiao; the famous university of Ivy League such as Harvard and Yale.)</td>
</tr>
<tr>
<td>Focus (夜里一点/ yeli yidian; 1 am)</td>
<td></td>
</tr>
<tr>
<td>Textually accessible</td>
<td>Topic 1 (刘招华；Liuzhaohua)</td>
</tr>
<tr>
<td></td>
<td>Focus (由那儿/ you near; from there)</td>
</tr>
<tr>
<td></td>
<td>Topic 2 (刘招华（11.5）; Liuzhaohua)</td>
</tr>
<tr>
<td></td>
<td>Topic 1 (她--liuzhaohua)</td>
</tr>
<tr>
<td></td>
<td>Focus (打着（宗教旗帜）/dazhe zongjia qizhi;</td>
</tr>
<tr>
<td>Topic</td>
<td>Situationally accessible</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Topic（阿光; A guang）</td>
<td>Topic（这条大街/ zhetiao dajie; this street ）</td>
</tr>
<tr>
<td><strong>Situationally accessible</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Active/identifiable</strong></td>
<td></td>
</tr>
<tr>
<td>Focus (七一年/ qiyi nian; year 1971)</td>
<td>Topic1（这 2700 万元现金/ zhe 2700 wan yuan xianjin; this 27,000,000 yuan cash）</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Focus（解放以后/ jiefang yihou; after the Liberation）</td>
<td></td>
</tr>
<tr>
<td>Focus (从（英国）/ (cong) Yingguo; (from) Britain)</td>
<td></td>
</tr>
<tr>
<td>Focus (四八年/ 48 nian; year 1948)</td>
<td></td>
</tr>
<tr>
<td>Topic2 (陈炳锡/ Chen bingxi)</td>
<td></td>
</tr>
<tr>
<td>Focus(文化大革命前/ wenhua dageming qian; after The Great Culture Revolution)</td>
<td></td>
</tr>
<tr>
<td>Focus（(从)那个国民党伪政府）</td>
<td></td>
</tr>
<tr>
<td>Topic（陈存仪/ Chen cunyi）</td>
<td></td>
</tr>
<tr>
<td>Topic（刘招华/ Liu zhaohua）</td>
<td>Topic（他/ ta; he）</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Topic（刘招华/ Liu zhaohua）</td>
<td>Topic（我/ wo; I）</td>
</tr>
<tr>
<td>Topic1（陈文印/ Chen wenyin）</td>
<td>Topic（你/ ni; you）</td>
</tr>
<tr>
<td>Topic（陈/ Chen）</td>
<td>Topic（他/ ta; he）</td>
</tr>
<tr>
<td>Topic（太平天国/ Taipingtianguo）</td>
<td>Topic（我们/ women; we）</td>
</tr>
<tr>
<td></td>
<td>Topic（我/ wo; I）</td>
</tr>
<tr>
<td></td>
<td>Topic（他/ ta; he）</td>
</tr>
<tr>
<td></td>
<td>Topic（他/ ta; he）</td>
</tr>
<tr>
<td></td>
<td>Topic（我/ wo; I）</td>
</tr>
<tr>
<td>Name of the frame</td>
<td>Lexical unit</td>
</tr>
<tr>
<td>------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>Quitting</td>
<td>退休(tuixiu; retire)</td>
</tr>
<tr>
<td>Activity_finish</td>
<td>毕业(biye; graduate)</td>
</tr>
<tr>
<td>Perception_active</td>
<td>看(kan; watch)</td>
</tr>
<tr>
<td>Intentionally_act</td>
<td>做(zuo; do)</td>
</tr>
<tr>
<td>Arriving</td>
<td>回来(huilai; come back)</td>
</tr>
<tr>
<td></td>
<td>回到（huidao, go back）</td>
</tr>
<tr>
<td></td>
<td>赶到(gandao; arrive)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Departing</td>
<td>走(zou; walk)</td>
</tr>
<tr>
<td></td>
<td>撤回(chehui; go back)</td>
</tr>
<tr>
<td>Motion</td>
<td>飞抵(feidi; fly to)</td>
</tr>
<tr>
<td>Bringing</td>
<td>带过去(dai huiqu; bring back)</td>
</tr>
<tr>
<td>Giving</td>
<td>给(gei; give)</td>
</tr>
<tr>
<td>Education_teaching</td>
<td>教育(jiaoyu; educate)</td>
</tr>
<tr>
<td>Event Type</td>
<td>Chinese (pinyin)</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Success_or_failure</td>
<td>失败 (shibai; fail)</td>
</tr>
<tr>
<td>Activity_start</td>
<td>起来 (qilai; start)</td>
</tr>
<tr>
<td>Death</td>
<td>死 (si; die)</td>
</tr>
<tr>
<td>Renting</td>
<td>承包 (chengbao; hire)</td>
</tr>
<tr>
<td>Activity_ongoing</td>
<td>进行 (jinxing; operate)</td>
</tr>
<tr>
<td>Seeking</td>
<td>找 (zhao; find)</td>
</tr>
<tr>
<td>Process_start</td>
<td>开始 (kaishi; start)</td>
</tr>
<tr>
<td>Finish_competition</td>
<td>取胜 (qusheng; win)</td>
</tr>
<tr>
<td>Reporting</td>
<td>告起 (gaoqi; sue from)</td>
</tr>
<tr>
<td>Feeling</td>
<td>感到 (gandao; feel)</td>
</tr>
<tr>
<td>Forming_relationship</td>
<td>结婚 (jiehun; marry to)</td>
</tr>
<tr>
<td>Manipulate_into_doing</td>
<td>骗取 (pianqu; cheat)</td>
</tr>
<tr>
<td>Make_acquaintance</td>
<td>碰见 (pengjian; meet)</td>
</tr>
<tr>
<td>Building</td>
<td>修好 (xiuhao; built)</td>
</tr>
<tr>
<td>Bringing</td>
<td>搬来 (banlai; move)</td>
</tr>
<tr>
<td>Self_motion</td>
<td>爬过来 (paguolai; crawl to)</td>
</tr>
<tr>
<td>Name of the frame</td>
<td>Lexical unit</td>
</tr>
<tr>
<td>------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Quitting</td>
<td>退休(tuixiu; retire)</td>
</tr>
<tr>
<td>Activity_finish</td>
<td>毕业(biye; graduate)</td>
</tr>
<tr>
<td>Perception_active</td>
<td>看(kan; watch)</td>
</tr>
<tr>
<td>Intentionally_act</td>
<td>做(zuo; do)</td>
</tr>
<tr>
<td>Arriving</td>
<td>回来(huilai; come back)</td>
</tr>
<tr>
<td></td>
<td>回到(huidao; come back)</td>
</tr>
<tr>
<td></td>
<td>赶到(gandao; arrive)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Departing</td>
<td>走(zou; walk)</td>
</tr>
<tr>
<td></td>
<td>撤回(chehui; go back)</td>
</tr>
<tr>
<td>Motion</td>
<td>飞抵(feidi; fly to)</td>
</tr>
<tr>
<td>Bringing</td>
<td>带过去(dai huiqu; bring back)</td>
</tr>
<tr>
<td>Giving</td>
<td>给(gei; give)</td>
</tr>
<tr>
<td>Education_teaching</td>
<td>教育(jiaoyu; educate)</td>
</tr>
<tr>
<td>Success_or_failure</td>
<td>失败(shibai; fail)</td>
</tr>
<tr>
<td>Activity_start</td>
<td>起来(qilai; start)</td>
</tr>
<tr>
<td>Action</td>
<td>Chinese</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Death</td>
<td>死(si; die)</td>
</tr>
<tr>
<td>Renting</td>
<td>承包(chengbao; hire)</td>
</tr>
<tr>
<td>Activity_ongoing</td>
<td>进行(jinxing; operate)</td>
</tr>
<tr>
<td>Seeking</td>
<td>找(zhao; find)</td>
</tr>
<tr>
<td>Process_start</td>
<td>开始(kaishi; start)</td>
</tr>
<tr>
<td>Finish_competition</td>
<td>取胜(qusheng; win)</td>
</tr>
<tr>
<td>Reporting</td>
<td>告起(gaoqi; sue from)</td>
</tr>
<tr>
<td>Feeling</td>
<td>感到(gandao; feel)</td>
</tr>
<tr>
<td>Formingrelationship</td>
<td>结婚(jiehun; marry to)</td>
</tr>
<tr>
<td>Manipulate_into_doing</td>
<td>贪取(pianqu; cheat)</td>
</tr>
<tr>
<td>Make_acquaintance</td>
<td>碰见(pengjian; meet)</td>
</tr>
<tr>
<td>Building</td>
<td>修好(xiuhao; built)</td>
</tr>
<tr>
<td>Bringing</td>
<td>搬来(banlai; move)</td>
</tr>
<tr>
<td>Self_motion</td>
<td>爬过来(paguolai; crawl to)</td>
</tr>
</tbody>
</table>
References


Coherence and grounding in discourse, 21-51. Amsterdam: Benjamins.


Press.


Bulzoni.


Journal of Tianjin Normal University, 4: 75-76.


305


Conference on Computational Linguistics (JSCL) .115-122.