An LFG Account of Empty Pronouns in Mandarin Chinese

Wei-Tai Liao

Wolfson College, Oxford University

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

Michaelmas Term 2010

An LFG Account of Empty Pronouns in Mandarin Chinese

Wei-Tai Liao
Wolfson College, Oxford University
A thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy, Michaelmas Term 2010

Abstract

This thesis incorporates a study of the syntactic distribution and pragmatic properties of empty pronouns in Mandarin Chinese (Chinese in short) from the point of view of Lexical-Functional Grammar (LFG). With reference to various accounts of Chinese empty pronouns by previous researchers (e.g., J. Huang 1984; Y. Huang 1995; Pan and Hu 2000) and the problems with their analyses, we offer a conjunctive approach to capture how Chinese zero anaphora is constrained at the level of syntax and the level of pragmatics. A series of syntactic rules are proposed and every occurrence of Chinese empty pronouns must strictly obey these absolute, inviolable syntactic constraints. Intersententially, discourse-dependent pragmatic conditions based on Lambrecht's (1994) information-structure theory form the second part of our two-way characterization of Chinese empty pronouns. Lastly, we present a comparison between the case study of empty pronouns in Japanese by Kameyama (1985) and our current study of empty pronouns in Chinese, and demonstrate two distinctive *pro*-drop phenomena.

Chapter 1 is an introduction to the notion of *pro*-drop, types of empty pronouns identified in Chinese, and the theoretical frameworks assumed in this study. Chapter 2 reviews three different previous analyses of Chinese empty pronouns, discusses the problems each account suffers, and promotes an integrating account which takes both syntax and pragmatics into consideration. Chapter 3 investigates the syntactic distribution of Chinese zero anaphora, whilst Chapter 4 provides a contrasting pragmatic characterization, followed by an examination of our proposal by testing it in naturalistic data and then a comparative study of Chinese and Japanese empty pronouns in Chapter 5. Finally, Chapter 6 summarizes the previous chapters, points out relevant issues that are not covered in this thesis but which could be research topics for future investigation, and reiterates how LFG is useful in this present work.

To my parents –

Liao, Kuang-Feng & Chang, Fu-Chu.

CONTENTS

1 Introduction 1

1.1 The purpose of the study and thesis outline……. 1

1.2 Preliminaries…… 4

1.2.1 Identifying empty pronouns – a verb-valency-based approach 4

1.3 An overview of the licensing of *pro*…… 5

1.3.1 The *pro*-drop parameter…… 5

1.3.1.1 *Pro*-drop in Italian…….. 8

1.3.1.2 Taraldsen's Generalization…….. 9

1.3.2 *Pro*-drop in Chinese…….. 11

1.3.2.1 Chinese as a topic-prominent language…….. 12

1.3.2.2 Sentencehood: full vs minor sentences…….. 15

1.3.2.3 Text-dependent vs text-independent reference…….. 18

1.4 Two types of empty pronouns identified in Chinese…… 20

1.4.1 Empty subject pronouns…… 20

1.4.2 Empty object pronouns……. 22

1.5 Theoretical frameworks…… 23

1.5.1 Formal representation: Lexical-Functional Grammar…… 23

1.5.1.1 C(onstituent) structure……. 24

1.5.1.2 F(unctional) structure…….. 26

1.5.1.3 A(rgument) structure…… 32

1.5.1.4 I(nformation) structure……. 33

1.5.2 Lambrecht's information-structure theory…… 35

1.6 Summary……. 38

2 A syntactic investigation of empty pronouns in Chinese 40

2.1 Topichood …….. 40

2.1.1 Topichood in LFG ……... 40

2.1.2 Topichood in Chinese …….. 46

2.2 A review of previous solutions and their problems…… 52

2.2.1 A syntax-oriented account by J. Huang…… 52

2.2.2 A pragmatics-oriented account by Y. Huang…… 60

2.2.3 A syntactic-pragmatic account by Pan and Hu…… 68

2.2.4 The need for an integrating account……. 75

2.3 Topic-comment structure: a syntactic or semantic device? 77
2.3.1 A pro-syntax view by Shi (2000)…….. 77
2.3.2 A pro-semantics view by Pan and Hu (2002)…… 94
2.3.3 An LFG solution for Chinese dangling topics…….. 101
2.4 Summary…….. 113

3 Syntactic constraints on Chinese zero anaphora 114
3.1 An introduction ……... 114
3.2 Syntactic constraints…… 115
3.2.1 No coindexation between an empty pronoun and its coargument…….. 115
3.2.2 No empty pronouns allowed after preposition or in pivotal position…… 116
3.2.3 Empty subordinate subjects must be anteceded by arguments 119
3.2.4 Empty object pronouns must be anteceded by topics……. 126
3.2.5 Extended binding…… 142
3.2.6 Local binding of empty possessive pronoun by the closest antecedent…….. 146
3.2.7 Arbitrary empty pronouns……. 151
3.3 Summary…….. 154

4 A pragmatic characterization of Chinese zero anaphora 155
4.1 Pragmatic generalizations for intersentential zero anaphora…….. 155
4.1.1 Empty pronouns must have actively or accessibly identifiable antecedents…… 155
4.1.2 A pronoun may be empty if it bears the discourse function as a topic, primary or secondary…….. 163
4.1.3 A discourse topic may extend its domain to the following clauses to form a topic chain…… 171
4.1.4 Summary…….. 174
4.2 An examination of our analysis …….. 175
4.2.1 An examination of zero anaphora in naturalistic data…….. 176
4.2.2 Summary……182

5 A comparative study of Chinese and Japanese zero anaphora 183
5.1 A comparison with Kameyama's (1985) study of zero anaphora in Japanese…… 183
5.1.1 Similarities……. 184
5.1.1.1 Both languages allow subjects and objects to be empty…… 185
5.1.1.2 Both languages allow arbitrary empty pronouns….. 186
5.1.1.3 Dummy subjects are not pronounced ……... 187
5.1.1.4 No coindexation between an empty pronoun and its coargument…… 188
5.1.1.5 The interpretation of zero anaphora relies heavily on discourse…… 189
5.1.1.6 Empty pronouns must be anteceded by discourse-prominent entities…….. 191
5.1.1.7 Both languages show subject-subject obviation….. 194
5.1.2 Differences…….. 197
5.1.2.1 Japanese empty subordinate subjects need not be anteceded by arguments…… 197
5.1.2.2 Only empty pronouns in Japanese are subject to Kameyama's *General Identification Constraint*….. 199
5.1.2.3 Chinese zero anaphora escapes Kameyama's Centering rules…….. 204
5.1.2.4 Japanese zero anaphora allows indefinite antecedence and antecedence in focus……. 207
5.2 An examination of empty pronouns in Japanese and Chinese texts.. 209
5.3 Summary…… 222

6 Conclusions and further research 223

References 225

ABBREVIATIONS

1	first person
2	second person
3	third person
ACC	accusative
ACTV	activation
ADJ	adjunct
ADJV	adjectival marker
ASP	aspectual marker
ASN	assertion marker
AVM	attribute-value matrix
CL	classifier
CMP	complementizer
COMP	complement
CONTRAS	contrastive
COP	copular verb
CP	complement phrase
DAT	dative
DEF	definiteness
DES	desiderative
DF	discourse function
DIR	directional
DRef	discourse referent
EC	empty category
EC	empty category
F	feminine
FOC	focus
GEN	genitive
GEND	gender
INFL	inflectional
LF	logical form
LOC	locative
M	masculine
NEG	negation
NLZ	nominalizer
NOM	nominative
OBL	oblique
OBJ	object
PART	particle
PASS	passive
PERF	perfective
PERS	person
PF	phonetic form
PL	plural
POL	politeness marker
PRED	predicator
PRES	present tense
PROG	progressive aspect
PST	past tense
Q	question particle
RECIP	reciprocity
RLTV	relativizer
RPT	reportive particle
SPEC	specifier
SUBJ	subject
SG	singular
TOP	topic

Chapter 1: Introduction

1.1 The purpose of the study and thesis outline

It has been widely recognized that Mandarin Chinese[1] (henceforth, Chinese) exhibits common use of empty categories in subject as well as in object positions, and such usages are instantiated in the following sentences in (1a) and (1b), in which the empty category is represented by Ø.

(1) a. John$_i$ shuo [Ø$_i$/ta$_i$ renshi Mark].
say he know
'John$_i$ said that (he$_i$)/he$_i$ knew Mark.'
b. John$_i$ shuo [ta$_i$ bu renshi Ø$_j$/ta$_j$].
say he not know him
'John$_i$ said that he$_i$ didn't know (him$_j$)/him$_j$.'

Due to the possibility of replacing these empty categories with overt pronouns, we will call them, for their pronominal nature, *empty pronouns*, and, collectively speaking, *zero anaphora*. In Mandarin Chinese, both finite and non-finite clauses allow the use of empty subjects and/or empty objects. Empty pronouns occurring in finite clauses are generally called *pro*, while those in subject position of non-finite clauses are termed PRO in the Principles and Parameters (henceforth, P&P) Theory. In this thesis, we will only be interested in unexpressed pronouns in finite clauses (*pro*), while the occurrences of PRO will not be discussed as an instance of genuine empty pronouns. Moreover, in this thesis, we will only be interested in the sort of empty pronouns which occur as missing arguments in full sentences,[2] while *argument* here is defined as 'a constituent in a sentence which is assigned a theta role by a predicate' (Adger 2003: 81). In sum, we will treat missing arguments occurring in a finite, full sentence as empty pronouns.

[1] All Mandarin Chinese examples provided in this thesis will be given in *Hanyu Pinyin* Romanization.
[2] Based on Chao's (1968: 60) classification, a full sentence is understood to comprise a subject and a predicate. More discussion on Chao's classification of *full* versus *minor* sentences will be provided in 1.3.2.2.

Chapter 1. Introduction

The study of empty categories in Chinese (whether it is in Mandarin Chinese or in Chinese dialects, such as Cantonese) has intrigued many Chinese grammarians of various theoretical backgrounds because of the complete lack of agreement morphology or inflectional features on Chinese verbs. The distribution of empty pronouns and how they are conditioned remain an exciting challenge for modern linguists. The literature regarding the resolution of Chinese zero anaphora is abundant; there have been transformational approaches as well as non-transformational attempts. Among them, some opt for a purely syntactic account, while some prefer to tackle the problem by means of semantics or pragmatics. We will show in Chapter 2 the disadvantages of adopting syntax exclusively or solely semantics/pragmatics to resolve empty pronouns, and, therefore, we will opt for an integrating analysis which aims to account for zero anaphora in Chinese syntactically *and* pragmatically. We will propose syntactic constraints in Chapter 3 and pragmatic conditions in Chapter 4 to which empty pronouns in Chinese are subject.

The motivation for this study initially comes from (i) the significant lack of extensive literature on the resolution of empty pronouns in Mandarin Chinese carried out in the theoretical framework of Lexical-Functional Grammar (henceforth, LFG); and (ii) the absence of a satisfying analysis which discusses different usages of empty pronouns at different linguistic levels, e.g., syntactically, semantically, at information structure, intersententially, or intrasententially. The study is also motivated by the theoretical interest in dispelling the common misconception that Chinese and Japanese are very similar to each other in that both allow free, or almost free, distributions of zero anaphora. In view of this, Chapter 5, which is built on our proposal and Kameyama's (1985) research on Japanese zero anaphora, will set out to compare the two languages first theoretically, followed by an examination of the same text taken from one chapter of the first *Harry Potter* novel by J. K. Rowling in 1997 (i.e., *Harry Potter and the Philosopher's Stone*) translated into Chinese and Japanese.

The main purpose of this study is to investigate how zero anaphora behaves in Chinese at intersentential and intrasentential levels, and how it is constrained by syntax and conditioned by pragmatics in the theoretical frameworks of LFG, Lambrecht's (1994)

theory of information structure, and Dalrymple and Nikolaeva's (to appear) information-structure formation. Our present work also sets out to compare how Chinese and Japanese differ from each other, given that they both permit empty pronouns in finite clauses and discourse plays a crucial role in both languages.

Chapter 1 is an introduction to the *pro-drop parameter* (or the *null subject parameter*), the types of empty categories identified in Chinese, the method we use, and the theoretical frameworks of LFG, Lambrecht's information-structure theory, and the information-structure representation assumed by Dalrymple and Nikolaeva (to appear). Chapter 2 reviews and compares various analyses tackling empty categories in Chinese from different angles, before reaching the conclusion that an integrating approach which takes both syntax and pragmatics into account is preferred. It also discusses the discourse role played by topic in Chinese, topichood, and whether the topic-comment structure is a syntactic or semantic device.

A detailed discussion of how zero anaphora is constrained by syntax is the focus of Chapter 3, and a list of syntactic constraints will be provided. After we complete the investigation of how zero anaphora is constrained syntactically in Chapter 3, we will present a pragmatic characterization of zero anaphora in Chapter 4. To put our proposal to the test, the last section of Chapter 4 presents an examination of Chinese intersentential empty pronouns occurring in naturalistic data. Chapter 5 offers a comparative view of Chinese zero anaphora with Kameyama's (1985) study of zero anaphora in Japanese, followed by an examination of the same text taken from one chapter of the book *Harry Potter and the Philosopher's Stone* – the first novel of the *Harry Potter* series – translated into Chinese and Japanese.

Finally, as the closing chapter, Chapter 6 summarizes the findings of the thesis and highlights how the theoretical framework of LFG is helpful to our analysis. It also touches on some remaining issues as well as pointing out potential research topics for any future work.

1.2 Preliminaries

This section provides an overview of how empty pronouns in finite clauses, namely, *pro*s, are treated in a transformational theory such as the Principles and Parameters (P&P) approach, and explains the *pro*-drop phenomenon in the Italian-type and Chinese-type of *pro*-drop languages.

1.2.1 Identifying empty pronouns – a verb-valency-based approach

Before starting to examine the *pro*-drop phenomenon in Chinese and reviewing the different *pro*-drop treatments to be discussed in this thesis, let us first determine how we identify occurrences of empty pronouns in a non-inflectional language such as Chinese. Chinese, being one of the *pro*-drop languages, which, by definition, allow the appearance of an empty pronoun in a tensed clause, makes use of empty pronouns to avoid the repetition of a previous noun phrase and to achieve discourse coherence. With no inflectional morphology on verbs to identify the occurrences of empty pronouns in Chinese, we will rely on verb valency for the identification of zero anaphora. For example, the Chinese verb *zhuo* 'capture' in (2) and (3) is a two-place predicate, as, factually speaking, *zhuo* 'capture' can appear with two overt arguments in Chinese and expresses a relationship between two entities in the world, where one captures the other. Treating *zhuo* 'capture' as a one-place predicate is not only syntactically inaccurate but also semantically impossible to convey a complete proposition in Chinese. In order to satisfy the verb valency and form a proposition, *zhuo* 'capture', intuitively, must have two (and only two) arguments. Clearly, neither (2) nor (3) has met this requirement, yet both sentences are well-formed in Chinese, which suggests that a phonologically null element is at work, illustrated by Ø to denote the presence of an empty pronoun. The unexpressed subject in the embedded clause in (2) and the unexpressed object in (3) are realized by an empty pronoun.

(2) Zhangsan$_i$ shuo [Ø$_i$ mei zhuo dao Lisi].
say NOM not capture ASP
'Zhangsan$_i$ said that (he$_i$) did not capture Lisi.'

(3) Ni zhuo dao Ø le ma?
you capture ASP ACC ASP Q
'Have you captured (something or someone)?'

In comparison, a verb such as *chuxian* 'appear, as in becoming noticeable or becoming present' intuitively involves only one entity in the world, i.e., the entity that appears (sometime or somewhere). The example in (4) demonstrates such a usage. Since *chuxian* 'appear' is a one-place predicate which requires a subject, we assume that the subject position in (4) is filled by an empty pronoun in order to form a complete proposition.

(4) Ø chuxian le!
NOM appear ASP
'(Someone or something) appeared!'

From the above demonstration, we show a verb-valency-based approach to identify occurrences of empty pronouns in a non-inflectional language such as Chinese, and now we move on to an overview of the licensing of *pro*s.

1.3 An overview of the licensing of *pro*

In this section, we will look into the *pro*-drop parameter, how it is applied in Italian, how it runs into difficulty in Chinese, and the difference between a subject-prominent and a topic-prominent language.

1.3.1 The *pro*-drop parameter

Let us first take a look at examples (5) to (10), illustrating how Chinese, Italian, and English allow or disallow occurrences of empty pronouns. (All (a) sentences are Chinese,[3] (b) Italian, and (c) English).

[3] In spoken Chinese, the third person singular pronoun *ta* is ambiguous between 'he' and 'she', as a subject, or 'him' and 'her', as an object. The gender difference is only reflected in each pronoun's own Chinese character when written down, but is indistinguishable phonetically. In the Chinese examples from (5) to (10)

(5) a. Ø kanjian Mark le. (Chinese)
NOM see ASP
'(Someone) saw Mark.'
b. Ø vide Marco. (Italian)
NOM see.PST.3.SG
'(He/she) saw Marco.'
c. *Ø saw Mark. (English)

(6) a. Ta kanjian Ø le.
he see ACC ASP
'He saw (someone).'
b. *Egli ha visto Ø.
he have.PRES.3.SG seen ACC
'He has seen (someone).'
c. *He saw Ø.

(7) a. Ø kanjian Ø le.
NOM see ACC ASP
'(Someone) saw (someone else).'
b. *Ø ha visto Ø.
NOM have.PRES.3.SG seen ACC
'(He/she) has seen (someone).'
c. *Ø saw Ø.

(8) a. $John_i$ shuo [$Ø_{i/*j}$ kanjian Mark le].
say NOM see ASP
'$John_i$ said that ($he_{i/*j}$) saw Mark.'
b. $John_i$ ha detto [che $Ø_{i/j}$ ha visto Marco].
have.PRES.3.SG said that NOM have.PRES.3.SG seen
'John said that (he/she) saw Marco.'
c. *John said [Ø saw Mark].

(9) a. $John_i$ shuo [$ta_{i/j}$ kanjian Ø le].
say he see NOM ASP
'$John_i$ said that $he_{i/j}$ saw (someone).'
b. *$John_i$ ha detto [che lui_i ha visto Ø].
have.PRES.3.SG said that he have.PRES.3.SG seen ACC
'$John_i$ said that he_i saw (someone).'
c. *John said [he saw Ø].

---

and in the rest of the thesis, unless it is under a context which explicitly indicates a feminine reading of the third person singular pronoun *ta* or where the dual reading of both possibilities matters, the translation for the said pronoun will be simplified as 'he' or 'him' for the ease of illustration.

(10) a. John$_i$ shuo [Ø$_{i/*j}$ kanjian Ø le].
say NOM see ACC ASP
'John$_i$ said that (he$_{i/*j}$) saw (someone).'
b. *John ha detto [che Ø ha visto Ø].
have.PRES.3.SG said that NOM have.PRES.3.SG seen ACC
'John said that (he/she) saw (someone).'
c. *John said [Ø saw Ø].

A more detailed discussion of how empty pronouns behave syntactically and semantically will be given in later chapters. So far, we can see from the above *pro*-drop sentences that Chinese and Italian are different from English in that the former permit the appearance of an empty category (EC) in a tensed, or finite, clause, whereas the latter forbids any EC in a finite/tensed clause. Chinese, on the other hand, also behaves distinctly from Italian in the sense that Chinese seems to allow an EC in both subject and object position, while only subject ECs are permitted in Italian.

In the P&P approach (Chomsky and Lasnik 1993; Carnie 2007), the categories of non-overt nominal positions recognized are trace, PRO, and *pro*. Both (11) and (12) below are grammatical sentences.

(11) John$_i$ shuo [Ø$_i$ renshi Mark]. (Chinese)
say NOM know
'John$_i$ said that (he$_i$) knew Mark.'
(12) a. Stefano parla. (Italian)
speak.PRES.3.SG
'Stefano speaks.'
b. Ø parla.
NOM speak.PRES.3.SG
'(He/she) speaks.'

According to the Extended Projection Principle (EPP) (Chomsky 1981, 1982), the syntactic representation of a subject, overt or not, must be projected, and sentences must have subjects. Following the EPP and assuming that the subject position, i.e., the specifier of IP, must be filled, we postulate that the two arguments of *renshi* 'know' in (11) are to be realized by the empty category Ø and *Mark*. The external argument assigned by *parla*

'speak' in (12b) is realized by Ø, too. In accordance with the Empty Category Principle (ECP),[4] the syntactically present but phonetically null element (Ø) which saturates the subject position in (11) and (12b) should have a formal licenser. Given the PRO theorem, which states that PRO occurs only in an ungoverned position, Ø cannot be PRO, as the definite reference *John* in (11) antecedent-governs Ø. As the ECP rules that traces must be properly governed and neither theta-government nor antecedent-government can be found in (12b), Ø cannot be a trace, either. Therefore, little *pro* is assumed to be what the missing subject Ø of the finite clauses in (11) and (12b) is.

The possibility of using an empty pronoun in the subject position of a tensed clause without degrading the well-formedness of the sentence is known as the *pro*-drop phenomenon. The *pro*-drop parameter is formulated to parameterize the properties and the distribution of the empty pronoun *pro* that occurs in finite clauses cross-linguistically. The *pro*-drop parameter is also used to distinguish languages that allow empty subjects in finite clauses from those which do not. In Chapter 2 we will review three different analyses regarding the distribution and the regulation of empty pronouns in Chinese: a purely syntactic account by J. Huang (1984, 1987, 1989, 1991); a purely pragmatic account by Y. Huang (1995); and a syntactic-pragmatic account by Pan and Hu (2000). We will first begin with the discussion of the *pro*-drop parameter in the next subsection and introduce how *pro*-drop develops differently in two language groups: Italian-like and Chinese-like.

1.3.1.1 *Pro*-drop in Italian

Traditional analyses of the syntactic typology of NPs classify them according to the two binary features [± anaphor, ± pronominal] (Chomsky 1981). In the light of this classification, the empty category *pro* is identified as [– anaphor, + pronominal], and is seen as the counterpart of an overt pronoun. This empty pronoun *pro* is not witnessed in every language; English and French, for example, disallow *pro*-drop, and subjects of

[4] The Empty Category Principle (ECP) requires all traces to be properly governed, i.e., theta-governed or antecedent-governed (Haegeman 1994: 442).

finite clauses are required to be syntactically and phonetically present in these two languages. On the other hand, languages like Spanish and Italian, which, as shown in (12b), allow a pronominal subject in finite clause to be null, are the so-called *pro*-drop languages.

As demonstrated, subjects of finite clauses in Italian can be left unexpressed on the surface, and considerable attention has been devoted to the study of what ensures that *pro* is interpreted without any confusion. Rizzi (1986) proposes that (i) *pro* in *pro*-drop languages is formally licensed under case assignment by a designated head, i.e., a Case-assigning INFL, and (ii) the features on *pro* are recovered from the licensing head through a binding relationship. The subject-verb agreement morphology in Italian is very rich, and the content of *pro* can be easily recovered through the verbal inflections. English, however, does not possess such richness in agreement morphology. The meager inflectional system in English is insufficient to recover the reference of *pro*. Therefore, recoverability is assumed to be the principle determining whether *pro* can be dropped or not. Cited in J. Huang (1984), Jaeggli's (1982) 'identification hypothesis', which follows Taraldsen's (1978) 'principle of recoverability', indeed separates *pro*-drop languages like Italian and Spanish from non-*pro*-drop languages like English. However, when it comes to Chinese, which exhibits no agreement morphology but still permits *pro*-drop, the identification hypothesis and the principle of recoverability both run into difficulty.

1.3.1.2 Taraldsen's Generalization

Taraldsen's Generalization, similar to Jaeggli's identification hypothesis, proposes that the possibility of *pro*-drop in a language is based on recoverability. If verbal agreement is rich enough, as in Italian and Spanish, to recover the content of *pro*, be it in subject or object position, such an EC is licensed and thus well-formed. Taraldsen's Generalization correctly explains how a subject EC is permissible in Italian while an object EC is not, due to the presence of subject-verb agreement and unavailability of object-verb agreement. Further evidence which shows that recoverability seems to be the key to

licensing *pro*-drop and thus supports Taraldsen's Generalization comes from Pashto.[5] According to J. Huang (1984: 535), subject-verb agreement in Pashto is observed in sentences expressing present events regardless of the verb being transitive or intransitive. In contrast, in the corresponding past-tense sentences, Pashto exhibits ergative verbal agreement: an intransitive verb agrees with the subject, and a transitive verb with the object. Examples (13) and (14) (adapted from J. Huang (1984: 536)) illustrate the ability of a subject pronoun or an object pronoun to be dropped according to whether there is subject-verb agreement or object-verb agreement, respectively. The ill-formed sentences in (15) and (16) can be accounted for along the same lines.

(13) (J. Huang 1984, (13))
a. Ø ra-z-i.
NOM DIR-come-3.M.SG
'(He) comes.'
b. Ø mana xwr-əm.
NOM apple eat-1.M.SG
'(I) eat the apple.'

(14) (J. Huang 1984, (14))
a. Ø ra-g-ay.
NOM DIR-come-3.M.SG
'(He) came.'
b. Ma Ø wə-xwar-a.
I ACC PERF-eat-3.F.SG
'I ate (it, with *it* being grammatically feminine).'

(15) (J. Huang 1984, (15))
*Zə Ø xwr-əm.
I ACC eat-1.M.SG
'I eat (something).'

(16) (J. Huang 1984, (16))
*Ø mana wə-xwar-a.
NOM apple PERF-eat-3.F.SG
'(Someone) ate the apple.'

[5] Pashto, also known as Pushto, Pakhto, or Afghan, is a split-ergative language with very rich inflection. It is spoken by twenty to thirty million people in Afghanistan (J. Huang 1984).

Taraldsen's Generalization is supported by languages like Italian and Pashto, in which the possibility of *pro*-drop neatly accords with subject/object-verb agreement. The ungrammaticality of (15) and (16) shows that a subject or an object can only be null when it agrees with the verb carrying features that can recover the content of the unexpressed subject or object; otherwise, the pronoun must be explicit and cannot be null.

1.3.2 *Pro*-drop in Chinese

The *pro*-drop parameter predicts that the possibility of *pro*-drop depends on the richness of the inflectional morphology in a language. Rizzi (1986) suggests that *pro*-drop is regulated according to two requirements: (i) *pro* is governed by X^0 (head-government), and (ii) *pro* has the same grammatical features as whatever *pro*'s licensing head is coindexed with. The important aspects of *pro* can therefore be recovered by virtue of the agreement on the tensed verb or auxiliary. However, Chinese is known for its zero inflectional morphology and, interestingly, as we already showed earlier, empty categories are not only found in subject position but also in object position, whereas in Italian, only subject *pro* is allowed. (17) is a Chinese discourse in which all of Speaker B's alternative replies ranging from (a) to (e) are grammatical. Note that the available combinations of an overt/empty subject and object inside the embedded clause in (e) parallel the patterns seen in (a) to (d).

(17) A: John lianluo Bill le ma?
contact ASP Q
'Has John contacted Bill?'
B: a. Ta lianluo ta le.
he contact him ASP
'He has contacted him.'
b. Ø lianluo ta le.
NOM contact him ASP
'(He) has contacted him.'
c. Ta lianluo Ø le.
he contact ACC ASP
'He has contacted (him).'

d. Ø lianluo Ø le.
NOM contact ACC ASP
'(He) has contacted (him).'
e. John shuo [ta/Ø lianluo ta/Ø le].
say he contact him ASP
'John said that he has contacted him.'

The *pro*-drop phenomenon in Italian has been exhaustively explained as attributable to the very rich inflectional system of subject-verb morphological agreement. In line with the principle of recoverability and the identification hypothesis, this explanation predicts correctly that the use of an empty object will be unacceptable, for there is no object-verb agreement in Italian and the like. Yet, it is hard to see how Chinese can be similarly accommodated.[6] An alternative way of resolving *pro*s in languages with no agreement morphology must be called for. Another important difference between the Italian-type of *pro*-drop languages and the Chinese-type of *pro*-drop languages is that the former group can be characterized as subject-prominent languages, and the latter topic-prominent languages. Let us now look at what distinguishes a topic-prominent language from a subject-prominent language in the following subsection.

1.3.2.1 Chinese as a topic-prominent language

Chinese is known as a topic-prominent language (Chao 1968; Li and Thompson 1976), in which topic plays a significant role, in contrast to a subject-prominent[7] language like English or Italian. Generally speaking, subject-prominence is characterized by the grammatical relation between subject and predicate in a subject-predicate construction, as in '$[\text{Fish}]_{\text{Subject}}$ $[\text{swim}]_{\text{Predicate}}$'. Topic-prominence, on the other hand, is characterized by the relation of 'aboutness' holding between topic and comment in a topic-comment structure, as in 'As for $[\text{fish}]_{\text{Topic}}$, $[\text{they swim}]_{\text{Comment}}$'. Subject-predicate construction is treated as the basic, or standard, sentence structure in a subject-prominent language,

[6] The kind of *pro*-drop observed in Chinese, where there is a total lack of agreement morphology, prompts the following generalization by Jaeggli and Safir (1989): null subjects occur in the context of either very rich agreement or no agreement at all.
[7] The term 'subject-prominent' is first introduced by Keenan (1976).

whereas topic-comment construction is the basic sentence type in a topic-comment language.

The subject-predicate construction has always been considered basic, or standard, in most linguistic theories, and sentences tend to be analyzed in terms of subject and predicate. However, Li and Thompson (1976) point out that all the languages they have investigated have the topic-comment construction, and not all of them have the subject-predicate structure. According to Li and Thompson's typology, languages can be divided into the following four groups: (i) subject-prominent languages, such as English; (ii) topic-prominent languages, such as Chinese; (iii) languages that are both subject-prominent and topic-prominent, such as Japanese; and (iv) languages that are neither subject-prominent nor topic-prominent, such as Tagalog. Claiming one language is subject-prominent (or topic-prominent) does not imply that topic (or subject) is not in use in this language. Instead, a language being typologically subject-prominent only amounts to a claim that this language 'can be more insightfully described by taking the notion of subject as basic' (Li and Thompson 1976). It is Chao (1968) who first comes up with the suggestion that 'the grammatical meaning of subject and predicate in a Chinese sentence is topic and comment' (Chao 1968: 69), even though his association of all subjects to topics is by no means exceptionless (since there are many cases in which the subject and the topic differ, or the topic is not necessarily the subject and could be the object or other syntactic categories). Li and Thompson (1976) elaborate and modify Chao's intuition, taking it one step further to argue for topic-prominence in Chinese and treat the topic-comment structure as basic in Chinese.

Generally speaking, a subject is most likely to be considered a topic, if not otherwise indicated by discourse information. Following Li and Thompson's (1976) insight, Chinese is characterized by a number of properties of topic-prominent languages. To name some of the properties, for instance, one of the differences between topic-prominent and subject-prominent languages is that, according to Li and Thompson (1976), the passive construction is used far more widely in subject-prominent languages than in topic-prominent languages. Moreover, passivization does not even exist in some topic-prominent languages (e.g., Lahu, Lisu), and if it does occur, as in the case of

Chinese, the passive construction is used marginally or usually carries some negative connotation, e.g., describing unfortunate or adverse events.

As we will see in more detail in subsequent chapters, the licensing and constraining of Chinese empty pronouns largely surrounds the notion of topic, unlike empty pronouns in subject-prominent *pro*-drop languages such as Italian. For instance, the empty subject in Speaker B's utterance from the following Chinese example (18) refers back to Speaker A's daughter who is identifiable and active in reference, and it serves as a topic in this mini discourse.

(18) A: Qiao. Zhe shi wo nur_i de zhaopian.
look this is my daughter GEN picture
'Look. This is a picture of my $daughter_i$.'
B: $Ø_i$/ta_i zhangde zhen ke-ai.
she grow really cute
'(She_i) looks really cute.'

According to Lambrecht's (1994) information-structure theory (which follows Chafe's (1976, 1987) terminology), *identifiability* is used as a cognitive category to be a speaker's estimation of whether the discourse representation of a certain referent is *identifiable* or *unidentifiable* to the hearer's mind. An identifiable referent can be further distinguished based on its activation state. *Activation* is another cognitive category and it concerns the speaker's evaluation of the status of the representation of an identifiable referent as already *activated*/*active*, *accessible*, or simply *inactive*. Speaker A's daughter in (18) is identifiable because her representation is present in the speaker's and hearer's mind at the time of Speaker B's utterance, and has become the topic of B's comment. Such discourse prominence licenses the use of the empty subject in (18), which has a lot to do with Chinese being a topic-prominent language.

1.3.2.2 Sentencehood: full vs minor sentences

Having made the claim earlier in 1.2.1 that the identification of empty pronouns in a non-inflectional language such as Chinese is derived from verb valency, it is important to further distinguish a full sentence from a minor sentence because the way we employ verb valency to identify an empty pronoun only makes sense with regard to a full sentence. The terminology and the distinction between full versus minor sentences is first introduced by Chao (1968). We will now look at their definitions and how they are related to our identification of empty pronouns.

The pioneering work in Chinese linguistics by Chao (1968) classifies Chinese sentences into full sentences and minor ones according to their structures. Verb valency, as was suggested earlier, serves as a reliable device to indicate the occurrence of an empty pronoun in a non-inflectional language; we believe that this test should only be applied to full sentences but not minor sentences. In this thesis, we will only be interested in the sort of empty pronouns which occur as missing arguments in full sentences. Any missing argument (including subjects) of a predicate in a minor sentence is irrelevant to our study here. Let us begin with Chao's definition of full and minor sentences.

Chao (1968: 60) states that 'a full sentence consists of two parts, a subject and a predicate, and is the commonest type in connected discourse'. Minor sentences, on the other hand, do not satisfy this subject-predicate requirement in form, but they are equally grammatical in Chinese. It follows that since minor sentences are not regulated by the strict subject-predicate, or argument-predicate, requirement, verb valency is of less relevance.

Minor sentences, either being verbal expressions or nominal predicates themselves, are not of the canonical subject-predicate structure, and they are usually used as commands, vocatives, responses, and exclamations (Chao 1968: 60). The followings are a few common types of minor sentences listed in Chao (1968), some of which might be regarded as containing empty categories, but attention will be drawn only to empty pronouns appearing in full sentences and whose references are text-dependent. Based on

Chao's (1968) classification, we begin with the first type of minor sentence: nominal predicates. Chao argues that in Chinese, deictic or demonstrative phrases like *zhe shi* 'this is' and *na shi* 'that is' are often left unexpressed, leading to minor sentences such as *Xiao niao!* 'Little bird!', if uttered by a child, for example, when he saw a flying animal in the sky. In cases like this, the nominal expression, say, *Xiao niao!* 'Little bird!', functions as a nominal predicate, and is thus viewed as a minor sentence by Chao. One could also use nominal predicate minor sentences to attract attention, e.g., *Shui!* 'Water!', when it is used to mean 'Watch out for the water!' and to call for attention. Nominal predicates can also be combined with modifiers. One of Chao's examples is *Du yao!* 'Poisonous chemical (literally), or poison!' uttered, possibly by an adult, to a child playing with a bottle of pills. Another type of nominal predicate minor sentence functions as 'a command in the form of things desired' (Chao 1968: 66). As an example, one might say *Guozhi!* 'Juice!' to mean that he wants to order some juice, or *Zhangdan!* 'Bill!' to demand the bill when in a restaurant. This usage of nominal expressions is also observed in English. For instance, 'Two tickets' or 'One latte to go' can be the English equivalent of nominal predicates used as commands in the form of things desired.

The second kind of minor sentence consists of verbal expressions, which can be used to command (as imperatives), as in (19), or to form questions (as interrogatives) which one addresses directly to the hearer, as in (20).

(19) a. Ø qing mang yong.
NOM please slowly eat
'Please take your time eating.'

b. Ø zou!
NOM go
'Go/leave!'

(20) a. Ø qu nar le?
NOM go where ASP
'Where have (you) been?'

b. Ø zou bu zou?
NOM walk not walk
'Do (you) want to go or not?'

Minor sentences in (19) act as orders, and it is very common, cross-linguistically, for imperatives to have no overt subjects, so we will not include in our analysis any missing subjects found in imperatives. The examples in (20) are minor sentences used as questions, and it is normally a second person subject, i.e., the hearer or addressee, that is left unexpressed under such circumstances. That a second person matrix subject as the hearer/addressee is absent from an interrogative sentence is not uncommon, and it can even be compared to some interrogative sentences in English, a non-*pro*-drop language, when it is used informally or conversationally, e.g., '(Do you) Fancy some coffee?', '(Are you) Ready to go now?', '(Do you) Wanna bet?', etc. When forming a question in informal speech, the second person pronoun, singular or plural, is very likely to serve as the subject of the question as the addressee, and it is often left unpronounced in such an informal context. Therefore, the absence of a second person matrix subject in a verbal expression used as an interrogative in the fashion of (20) will not be considered a linguistically typical feature of *pro*-drop languages, since this also occurs in informal speech in non-*pro*-drop languages like English. In comparison to the omitted second person pronouns as the addressees seen in (20), if we replace the second person pronoun with a third person pronoun *ta* 'he', demonstrated in the corresponding examples in (21) below, the sentences become unacceptable when an empty subject occurs and the intended reading is unavailable. Therefore, we establish that imperative and interrogative minor sentences which lack an overt second person matrix subject in the manner observed in (19) and (20), although containing a missing argument, will not be considered genuine occurrences of empty pronouns in this study, as a non-*pro*-drop language such as English allows the same usage.

(21) a. Ta_i/*$Ø_i$ qu nar le?
he go where ASP
'Where has he/(he) gone?'
b. Ta_i/*$Ø_i$ zou bu zou?
he go not go
'Does he/(he) want to go or not?'

Thirdly, meteorological sentences are often subjectless in Chinese, as illustrated by (22). Unlike English, which must employ a dummy subject such as 'it', the subject of such meteorological sentences is often left unexpressed in Chinese.

(22) a. Ø xiayu le.
NOM rain ASP
'(It) is raining.'
b. Waitou Ø xiaxue le ma?
outside NOM snow ASP Q
'Is (it) snowing outside?'
c. Yinwei Ø tai re le.
because NOM too hot ASP
'Because (it/the weather) is too hot.'

Likewise, an empty subject referring to time, exemplified by (23) below, is similar to the unpronounced subject in meteorological sentences in that both function like a dummy subject in English and do not have real semantic substance. Therefore, empty subjects referring to time or the weather will not be further discussed here. In conclusion, all those empty pronouns found in minor sentences plus the empty subject denoting weather or time will not concern us in this study of empty pronouns in Chinese.

(23) a. Ø san dianzhong le.
NOM three o'clock ASP
'(It) is three o'clock.'
b. Ø yijing zhongwu le.
NOM already noon ASP
'(It) is already noon.'

1.3.2.3 Text-dependent vs text-independent reference

In addition to the aforementioned minor sentences which may contain missing arguments but will not be studied in this thesis, any empty pronoun whose reference is *text-independent* is also irrelevant to our analysis. A text-independent reference takes a

non-linguistic antecedent, instead of a linguistic one, for the interpretation of the pronoun, and such non-linguistic antecedence will not concern us in characterizing the conditions an empty pronoun is subject to.

As the source of any pronoun's reference should be either *text-independent*, i.e., taking a non-linguistic antecedent, or *text-dependent*, i.e., taking a linguistic antecedent, empty pronouns are no exception. (Here the definition of text can be of written or of spoken form). To illustrate a text-independent type of antecedent, the empty object in (24a) refers to the beautiful dress in the scenario, and the empty subject in (24b) refers to some contestant appearing on TV at the time of the utterance.

(24) a. [Two girls are standing outside a shop's window display, staring at a beautiful dress$_i$, and one says the following to the other …]
Wo hao xiang mai Ø$_i$ o!
I well want buy ACC PART
'I want to buy (it/this dress) so much!'
b. [Two women are watching the beauty pageant on TV, and one makes the following comment …]
Ø zhangde zhen piaoliang!
NOM look really pretty
'(She/a certain contestant on TV) looks really pretty!'

The antecedents in both examples are of non-linguistic form, and the reference can only be recovered from the extra-linguistic material. This type of text-independent empty pronoun is called *zero deictics* by Kameyama (1985), and they refer to empty pronouns whose antecedents can only be deictically recovered. Due to their situationally-evoked antecedence, text-independent empty pronouns will be excluded from our study of zero anaphora in this thesis.

In contrast to text-independent empty pronouns, the sentences in (25) below instantiate the opposite – text-dependent empty pronouns, which take a linguistic antecedent either from the sentences itself or from the previous linguistic discourse. For example, (25a) could be a reply to a question 'Where is John?', and the empty subject in (25a) would unmistakably refer to the previously mentioned topic, *John*. Likewise, (25b)

could also be a reply to a question 'What is John doing?', and the empty subject, again, would refer to the topic in the discourse, *John*.

(25) a. [Where is John$_i$?]
Ø$_i$ qu tushuguan le.
NOM go library ASP
'(He) went to the library.'
b. [What is John$_i$ doing?]
Ø$_i$ zai ting yinyue.
NOM PROG listen music
'(He) is listening to music.'

It is evident that text-dependent empty pronouns differ from text-independent ones in their source of reference. Our analysis of zero anaphora will focus on text-dependent empty pronouns and not text-independent ones, as the former present far more interesting linguistic issues and also raise more theoretical problems.

1.4 Two types of empty pronouns identified in Chinese

This section presents the types of empty pronouns identified in Chinese, and how they are classified according to their structural position and grammatical function.

1.4.1 Empty subject pronouns

It is uncontroversial to claim that Chinese allows empty subjects. Note that this does not imply conversely that any subject can be null in Chinese even given the most appropriate and relevant context, because the distribution of empty pronouns is by no means entirely free and is subject to syntactic as well as pragmatic constraints which we will investigate in subsequent chapters. An empty subject pronoun can occur in a matrix clause, as exemplified in (26) (and also in the previous example (25)); in an embedded complement

clause, as in (27); in a relative clause,[8] as in (28); and in an adjunct clause, as in (29).

(26) [Two friends are chatting about another friend John$_i$, and one asks the other whether John has gone to New York recently. The other answers with the following reply.]
Shi a. Ø$_i$ qu le san-ge yue le.
yes PART NOM go ASP three-CL months ASP
'Yes. (He) has gone for three months.'

(27) John$_i$ shuo [Ø$_i$ mei kanjian Bill].
say NOM not see
'John$_i$ said that (he$_i$) didn't see Bill.'

(28) Wo$_i$ hen xihuan [jintian Ø$_i$ kandao __$_j$ de] na-jian yangzhuang$_j$.
I very like today NOM see ACC RLTV that-CL dress
'I like the dress which (I) saw today very much.'

(29) [Mei-tian zaoshang Ø$_i$ yao shangban zhiqian], mama$_i$ dou
every-day morning NOM is.going.to go.to.work before mother always
dei xian song John qu xuexiao.
have.to first send to school
'Every morning before (she/Mother) is going to work, Mother always has to take John to school first.'

Cross-linguistically speaking, a subject can have various structural forms, e.g., in the form of verbal inflectional morphology, as in Ulster Irish (Bresnan 2003), or in the form of discontinuous nominals carrying a specific case, as in Warlpiri (Bresnan 2001), but the overt subject in Chinese can only be either a non-discontinuous nominal or a sentential constituent. Instances of nominal subjects, overt or not, can be observed in all the above examples. (30a), on the other hand, exemplifies a sentential subject whereas (30b) illustrates the possibility of using a null sentential subject.

[8] In Chinese, a relativized construction has the word order where the relative clause, e.g., '(which) I saw today' in example (28), precedes the head noun modified by the relative phrase, e.g., 'the dress' in (28). This is the opposite of the word order in English relativization. The underscore '__' seen in the relative clause in (28) marks a gap, which is filled by a resumptive pronoun in some languages such as Hebrew (Sharvit 1999). Importantly, *de*, identified in the relative clause in (28), acts as the relative clause marker in Chinese, which unifies the usage of various relative markers such as 'who(m)', 'which', or 'where' in English, and its position precedes that of the head noun the relative clause modifies.

(30) a. [John changchang qiao-ke] rang mama feichang nanguo.
often cut-class make mother very upset
'That John often cut classes upset Mother very much.'
b. [John$_i$ changchang qiao-ke]$_j$, Ø$_{i/j}$ rang mama feichang nanguo.
often cut-class NOM make mother very upset
'That John often cut classes upset Mother very much. Or John$_i$ often cut classes; (he$_i$) upset Mother very much.'

Note that the empty subject found in (30b) can be coindexed with *John* or the fact that he often cut classes. (30b) differs from (30a) in placing a pause between the two separate clauses. Without this pause, (30b) would sound identical to (30a), with the same meaning of (30a) as the only available reading. Since this thesis is concerned with empty pronouns only, it will focus more on nominal than sentential ones. To sum up this subsection, empty subject pronouns in Chinese may appear in matrix clauses, complement clauses, relative clauses, or modifying adjunct clauses.

1.4.2 Empty object pronouns

Unlike empty subject pronouns, the existence of empty object pronouns in Chinese is not accepted by all linguists. For example, J. Huang (1984, 1987, 1989, 1991) disputes the status of an object EC in Chinese as pronominal; instead, he claims that object ECs are, in fact, variables bound by empty topics which functions as operators. Let us assume, for now, that empty object pronouns exist in Chinese. A detailed investigation will be presented in subsequent chapters.

The distribution of an empty object pronoun is not free but conditioned by syntax as well as discourse factors to be investigated later. Uncontroversially, Chinese has its canonical word order as subject-verb-object (SVO), and in the unmarked case, a nominal expression or a complement clause following a verb serves as the object of the sentence. The following examples in (31) to (34) demonstrate the possibilities for an empty object to be in a simple sentence, a complement, a relative clause, and an adjunct subordinate clause, respectively.

(31) [Does $Bill_i$ like $dogs_j$?]
$Ø_i$ xihuan $Ø_j$.
NOM like ACC
'(He) like (dogs).'

(32) [$Bill_i$ likes dogs very much …]
Dan $John_j$ shuo [$ta_{i/j}$ bu xihuan Ø].
but say he not like ACC
'But $John_j$ said that $he_{i/j}$ doesn't like (them/dogs).'

(33) [Is he the guy who miraculously found your stolen car_i?]
Bu, ta shi [$___j$ tou $Ø_i$ de] na-ge ren_j!
no he is NOM steal ACC RLTV that-CL person
'No, he is the guy who stole (it)!'

(34) [A $girl_i$ saw a nice $dress_j$ in the shop but didn't buy it immediately. She is thinking about going back and getting it as a present for her $sister_k$, but also worries that it might not fit. Her friend replies …]
Qu mai $Ø_j$ ba. Ruguo ta_k bu xihuan $Ø_j$, ni_i haishi keyi na
Go buy ACC PART if she not like ACC you still can take
$Ø_j$ hui dian-li tui-qian.
ACC back shop-inside return-money
'Go buy (it). If she doesn't like (it), you can still take (it) back to the shop for a refund.'

1.5 Theoretical frameworks

We aim to tackle the issues of zero anaphora resolution through the theoretical frameworks of Lexical-Functional Grammar, Lambrecht's (1994) theory of information structure, and Dalrymple and Nikolaeva's (to appear) information-structure representation.

1.5.1 Formal representation: Lexical-Functional Grammar

This subsection outlines the basic theoretical assumptions of LFG (Bresnan & Kaplan 1982; Dalrymple 2001), which is a constraint-based, non-configurational, non-derivational, mathematically well-defined, formal linguistic theory. Intended for the

analysis of typologically diverse languages, LFG is designed to accommodate, descriptively and theoretically, configurational (e.g., Chinese, English) as well as non-configurational (e.g., Warlpiri) languages. Our outline only focuses on the syntactic aspect of LFG that is relevant to our present work with deliberate omission of work in other linguistic fields, such as semantic research within the LFG framework (see Dalrymple 1999) and Optimality Theoretic syntax based on LFG (OT-LFG, in short).

The grammatical representation in LFG consists of the following four primary parallel structures: constituent structure (*c*-structure), functional structure (*f*-structure), argument structure (*a*-structure), and information structure (*i*-structure). These structures are associated with one another by principles of functional correspondence also known as *projection functions*; no derivation is assumed. Importantly, LFG separates different levels of linguistic information, or components, and keeps them at different structures because LFG views language as consisting of multiple dimensions and each dimension is represented individually within its own structure. In a nutshell, the subcategorization of a predicate and its semantic roles are listed at *a*-structure, structural relationships such as word order and constituency are displayed at *c*-structure, and grammatical functions[9] belong to the level of *f*-structure. There is also an *i*-structure that stores information related to discourse functions and provides links to pragmatics. Below we present the basic characterization of these four primary structures assumed in LFG: *a*-structure, *c*-structure, *f*-structure, and *i*-structure.

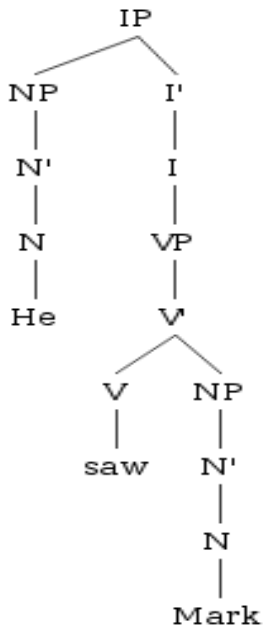
1.5.1.1 C(onstituent) structure

Constituent structure, or *c*-structure, encodes the actual words that make up a sentence in constituent units, and represents surface precedence and dominance relations. (35) demonstrates the *c*-structure of a simple sentence 'He saw Mark'. Notably, all entities at *c*-structure are base-generated, and they involve no movement. Moreover, (35) also tells

[9] The word *function* or *functional* may have a different meaning to different people in linguistics, but in LFG, grammatical functions denote notions like subject, object, complement, etc. Such grammatical relations are often defined in a tree-structural configuration in transformational theories, but in LFG, they are encoded in an attribute-value matrix at functional structure.

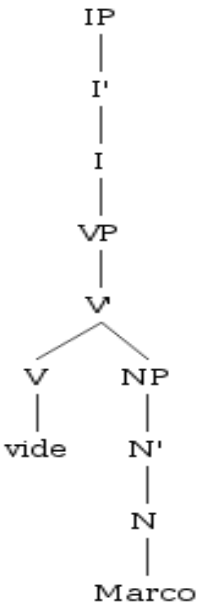
us that the mother node IP dominates the node of NP, I', I, and VP; and the subject precedes the object.

(35) He saw Mark.

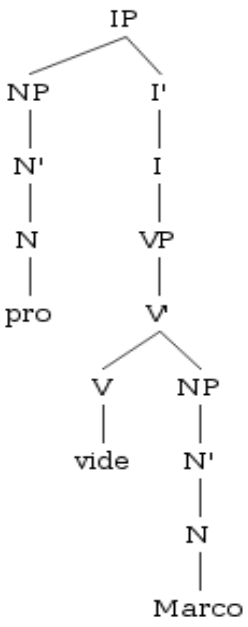


C-structure resembles closely the familiar tree structure from transformational theories in form, except that only *overt* expressions, i.e., expressions which are syntactically as well as phonetically realized, will appear at *c*-structure, while non-overt expressions such as *pro* may be present in a transformationalist tree diagram. Take the *pro*-drop Italian sentence in (36) for example. (36a) is the *c*-structure for this *pro*-drop sentence, and because the sentence lacks an overt subject, such a syntactic role is accordingly absent from the *c*-structure. In contrast, a transformationalist tree structure might look like the one in (36b), whose subject position is structurally filled by a hypothetical *pro*. That is to say, *c*-structure only reflects structurally overt information. As far as LFG is concerned, the employment of an empty pronoun is the interplay between *c*-structure and *f*-structure; an empty category's feature will be recorded at *f*-structure, which we will now turn to.

(36) a. Ø vide Marco.
NOM see.PST.3.SG
'(He/she) saw Marco.'



b.



1.5.1.2 F(unctional) structure

Functional structure (*f*-structure) is annotated as an attribute-value matrix (AVM), instantiated in (37) below for the sentence 'I saw a rare butterfly'. An attribute is a feature

or function name, including, in the case of (37), SUBJ (the grammatical notion of subject), TENSE, PRED (the grammatical notion of predicator), OBJ (the grammatical notion of object), ADJ (the grammatical notion of adjunct), DEF (definiteness), PERS (person feature of a noun phrase), NUM (number of a noun phrase), and SPEC (specifier). Each attribute has its value to the right. In the *f*-structure of (37), the attribute TENSE has a simple value, namely, it does not further contain a smaller *f*-structure (or an AVM), whereas other attributes like SUBJ and OBJ are grammatical functions because the values for SUBJ or OBJ are smaller *f*-structures. Let us consider the contents one by one.

(37) I saw a rare butterfly.

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{FORM} & \text{I} \\ \text{DEF} & + \\ \text{NUM} & \text{SG} \\ \text{PERS} & 1 \end{bmatrix} \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{'see} < \text{SUBJ, OBJ} > \text{'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'butterfly'} \\ \text{SPEC} & \text{a} \\ \text{ADJ} & \{[\text{PRED} \ \text{'rare'}]\} \end{bmatrix} \end{bmatrix}$$

Symbol-valued attributes such as TENSE, PERS(on), NUM(ber), CASE, and GEND(er) always have one simple and unique value because, for example, a sentence cannot be present tense and future tense at the same time, just like a pronoun cannot be nominative or accusative at the same time. The feature PRED(icator) represents the syntactically relevant aspects of the meaning of a lexical item. In contrast, dummy subjects or expletives, due to their lack of semantic content, will not have PRED features, but just FORM, exemplified by (38) below. Notice that in (37) above, the value for the PRED feature of SUBJ is written as PRO, which abstracts away the grammatical notion of pronoun in LFG. This convention is not to be confused with the PRO assumed in transformational grammar as the empty subject of a non-finite clause. The subject of the

sentence in (37) is predicated as PRO at *f*-structure, with first person, singular, definite features and the structural FORM as 'I'.

(38) It rained heavily yesterday.

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{FORM} & \text{it} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{'rain} < \ > \text{SUBJ'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'heavily'} \end{bmatrix} \right. \\ & \left. \begin{bmatrix} \text{PRED} & \text{'yesterday'} \end{bmatrix} \right\} \end{bmatrix}$$

The value for PRED is conventionally represented as the lexical item itself in single quotation marks, followed by the argument(s) selected by the predicate to show subcategorization. If the selected arguments are assigned theta roles by the predicate, such as SUBJ (as AGENT) and OBJ (as THEME) in the subcategorization of the predicate 'see' in (37) above, the said arguments are contained *within* a pair of angled brackets. However, if the selected argument does not receive any theta role from the predicate, as in the case of SUBJ subcategorized by 'rain' in (38) above, the said argument will appear *outside* the angled brackets. Also in the *f*-structure in (38), the value of the ADJ(unct) feature is encased by a pair of curly brackets that contain two *f*-structures inside. That is because, in theory, a sentence could consist of an indefinite string of words, mostly adjunct modifiers. (Imagine a sentence 'John sharpens his knife slowly, quietly, unhappily, thoughtfully … at midnight …'). Thus, the use of curly brackets indicates the possibility of having an indefinite set of smaller *f*-structures inside. An *f*-structure-valued attribute-value pair is known as a (grammatical) function. AVMs as functions, such as the values of SUBJ(ect), OBJ(ect), OBL(ique), and COMP(lement), are *f*-structures in their own right. Noteworthily, the term *f*-structure may describe the entire representation of the sentence (indicated by the largest, most-outward pair of brackets) or any smaller AVM within.

All *f*-structures are subject to the following three well-formedness conditions in (39): completeness, coherence, and uniqueness.

(39) (Falk 2001: 63-4)

a. **Completeness Condition**
   All argument functions specified in the value of the PRED feature must be present in the local *f*-structure. All functions that receive a thematic role must have a PRED feature.

b. **Coherence Condition**
   All argument functions in an *f*-structure must be selected by their local PRED. Any argument function that has its own PRED feature must be assigned a thematic role.

c. **Uniqueness Condition** (also known as **Consistency Condition**)
   Every attribute has a unique value.

The Completeness Condition rules out the ungrammatical English sentence in (40) because in this *f*-structure, two argument functions (SUBJ and OBJ) are subcategorized by the PRED feature, but only one is present. The sentence is incomplete.

(40) *Ø forgot the homework.

$$\begin{bmatrix} \text{PRED} & \text{'forget } < \text{SUBJ, OBJ} > \text{'} \\ \text{TENSE} & \text{PAST} \\ \text{OBJ} & \begin{bmatrix} \text{SPEC} & \text{the} \\ \text{PRED} & \text{'homework'} \end{bmatrix} \end{bmatrix}$$

The example in (41) is ill-formed because it fails to meet the Coherence Condition in having an additional OBJ function, which is not an argument function subcategorized by the PRED feature.

(41) *The MPs rebelled the prime minster.

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{SPEC} & \text{the} \\ \text{PRED} & \text{'MP (Member of Paliament)'} \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{PRED} & \text{'rebel} < \text{SUBJ} > \text{'} \\ \text{TENSE} & \text{PAST} \\ \text{OBJ} & \begin{bmatrix} \text{SPEC} & \text{the} \\ \text{PRED} & \text{'prime minister'} \end{bmatrix} \end{bmatrix}$$

Lastly, the sentence in (42) is ungrammatical, and it leads to an unacceptable *f*-structure because one of its attributes within the OBJ function, i.e., NUM, has been doubly valued as PL(ural) and SG (singular), violating the Uniqueness Condition. Every attribute must have a unique value.

(42) *The fishermen caught a valuable prawns today.

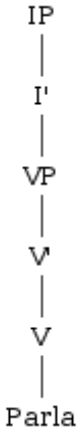
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{SPEC} & \text{the} \\ \text{PRED} & \text{'fisherman'} \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{'catch} < \text{SUBJ, OBJ} > \text{'} \\ \text{OBJ} & \begin{bmatrix} \text{ADJ} & \{[\text{PRED} \quad \text{'valuable'}]\} \\ \text{PRED} & \text{'prawn'} \\ \text{SPEC} & \text{a} \\ \text{NUM} & \text{PL, SG} \end{bmatrix} \\ \text{ADJ} & \{[\text{PRED} \quad \text{'today'}]\} \end{bmatrix}$$

The features and functions inside an *f*-structure representation can be ordered freely and do not have to follow the actual word order of the sentence, unlike the constituents presented at *c*-structure. To sum up, we repeat the example in (9b) in the following (43a) to demonstrate how an empty category is represented. Owing to the absence of a phonetic realization, the empty subject plays no role in the *c*-structure, as shown in (43b). The

f-structure in (43c), however, houses the grammatical functions whether they are overly present or not and, therefore, the empty subject is still characterized as carrying the supposed features, such as being third person singular with an unspecified gender either being female or male. This representation only differs from that of an overt pronoun in having a minus value for the attribute OVERT.[10]

(43) a. Ø parla.
NOM speak.PRES.3.SG
'(He/she) speaks.'

b.



c.

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{NUM} & \text{SG} \\ \text{GEN} & \text{F/M} \\ \text{PERS} & 3 \\ \text{OVERT} & - \end{bmatrix} \\ \text{PRED} & \text{'}\textit{parla}\text{''speak''} < \text{SUBJ} > \text{'} \\ \text{TENSE} & \text{PRES} \end{bmatrix}$$

In terms of the classification of grammatical functions housed at *f*-structure, SUBJ and OBJ are said to be core functions. Non-core functions include obliques, complements, and adjuncts. In addition to the contrast between core and non-core functions, there are

[10] 'OVERT' is a simple way of expressing the constraint that the inverse φ relation is empty and that there are no *c*-structure nodes corresponding to the *f*-structure.

also discourse functions and non-discourse functions. The former group contains TOP (the abstract notion of topic), FOC (the abstract notion of focus), and SUBJ, with the rest (OBJ, OBL, COMP, and ADJ) being non-discourse functions. TOP, FOC, and ADJ also form a group belonging to non-argument functions, while the rest are all argument functions. Though being discourse functions and playing essential roles at *i*-structure, TOP and FOC do appear at *f*-structure, too, as '*grammaticalized discourse functions* whose synchronic function is purely syntactic, related to but different from the information structure roles of topic and focus' (Dalrymple 2001; Dalrymple and Nikolaeva to appear). For instance, if a certain NP is a topic at *i*-structure, it will also be present at *f*-structure as TOP, and vice versa.

1.5.1.3 A(rgument) structure

Argument structure, or simply *a*-structure, provides subcategorization information for a predicator and the semantic roles involved. For instance, the *a*-structure for a two-place predicate like 'attack' in English would minimally contain the following two pieces of information – (i) grammatical functions (such as SUBJ and OBJ), and (ii) semantic roles (AGENT, THEME, GOAL, EXPERIENCER, etc.).

(44) $\text{'ATTACK} < \frac{\text{SUBJ}}{\text{AGENT}} , \frac{\text{OBJ}}{\text{THEME}} >$'

Different linguists have proposed different annotations for the *a*-structure, with some more complicated than others. The *a*-structure in (44) is one of the more uncomplicated versions that only contains minimal information in a simple format, but this is sufficient for our purpose intended here as a schematic guide to the basic architecture of LFG. (An extensive discussion of argument structure and various proposals of the constitution can be found in Chapter 8 of Dalrymple 2001).

1.5.1.4 I(nformation) structure

In early LFG literature, TOP and FOC were commonly treated as *f*-structure attributes. King (1997), Choi (1999), and Butt and King (2000) argue for an independent representation of *i*(nformation) structure [11] separate from *f*-structure. Following Dalrymple and Nikolaeva (to appear), *i*-structure is defined as 'the level of sentence organization which represents how the speaker structures the utterance in context in order to facilitate information exchange' and 'how the propositional content of an utterance fits the addressee's state of knowledge at the time of utterance'. We will adopt the representation of *i*-structure and discourse roles involved at *i*-structure assumed by Dalrymple and Nikolaeva (to appear) and Lambrecht (1994).

Information structure represents discourse entities and their relationship with the proposition communicated in the form similar to that of an *f*-structure. Information structure is commonly associated with the notions of old/given information and new information, and the contrast between the two. Lambrecht (1994) indicates the inappropriateness of equating old/new information with old/new sentence constituents (such as discourse referents), and we will follow his insight in rejecting such a segmentation view of analyzing information structure. Instead, he draws a distinction between 'the presupposed' and 'the asserted', and argues that *pragmatic presupposition* (or simply *presupposition*) and *pragmatic assertion* (or simply *assertion*) are the superior cover terms for old and new information, respectively. Lambrecht's definition of these two concepts is given below in (45). The two can coexist within the same sentence.

(45) (Lambrecht 1994: 52)

a. Presupposition: 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.

b. 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.

[11] Cited in Dalrymple and Nikolaeva (to appear), the analyses of Bresnan and Mchombo (1987), King (1995), Butt and King (1996, 2000), Choi (1999), King and Zaenen (2004), and O'Connor (2006) offer a useful overview of the development of *i*-structure in LFG.

As TOP and FOC are the discourse functions at *i*-structure, the topic of a sentence is defined by Lambrecht as 'the entity which the proposition expressed by the sentence is ABOUT', whereas the focus of a sentence is defined as 'the semantic component of a pragmatically-structured proposition whereby the assertion differs from the presupposition' (Lambrecht 1994: 213). The stance on defining topic in this way is also taken by various linguists, including Kuno (1972), Gundel (1976), Dik (1978), and Reinhart (1982). In Chinese, topicality can be explicitly marked by phrases like *as for* … or *speaking of* …, but topicality is not overtly indicated by any compulsory topic marker, such as *–wa* in Japanese. Nor is it marked by any special structural position, since the sentence-initial[12] position is just as available for focus or subject and is not exclusive for topics only. Thus viewed, the determination of topicality according to syntactic structure solely might not be reliable in Chinese, and the pragmatic aboutness relationship we are holding onto should require further discourse context in order to decide whether or not an entity is really a topic of the sentence. Presupposition and assertion are both propositions. The former is closely related to topic, while the latter is closely related to focus. Lambrecht (1994: 206) summarizes that 'a topic is included in the presupposition without being identical to it, and a focus is part of an assertion without coinciding with it'.

Following Dalrymple and Nikolaeva's (to appear) representation of *i*-structure, which is derived from Butt and King's (2000) proposal, we will assume that *i*-structure includes the following four information structure roles: topic is [–NEW] and [+PROMINENT], focus is [+NEW] and [+PROMINENT], completive information is [+NEW] and [–PROMINENT], and background information is [–NEW] and [–PROMINENT]. In particular, background information is like topic in containing old/known information (i.e., presupposition), but its information value is not prominent. Completive information, on the other hand, is pragmatic assertion that is not prominent in the discourse. For example, in (46) below, (46b) summarizes the *i*-structure roles available in Speaker B's reply in (46a), and (46c) is the *i*-structure. In (46c), the use of the term 'discourse referent' (DRef) at *i*-structure is similar to the use of the grammatical

[12] Structurally speaking, even though the sentence-initial position is equally available for subject or topic, Chu (1998: 249) adds that when a topic and a subject of the same sentence differ, the former comes before the latter.

term ‘argument’; DRef is the preferred term at *i*-structure because it gives a more information-structure-related and less syntactic impression.

(46) a. A: What is John writing?
B: He is writing a report in his room.
b. He is writing a report in his bedroom.
TOPIC BACKGROUND FOCUS COMPLETIVE

c.
$$\begin{bmatrix} \text{TOP} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{FORM} & \text{he} \end{bmatrix} \\ \text{BACKGROUND} & [\text{PRED 'is writing'}] \\ \text{FOC} & \begin{bmatrix} \text{PRED} & \text{'report'} \\ \text{SPEC} & \text{a} \end{bmatrix} \\ \text{COMPLETIVE} & \begin{bmatrix} \text{PRED} & \text{'in < DRef >'} \\ \text{DRef} & [\text{PRED 'his room'}] \end{bmatrix} \end{bmatrix}$$

Like Dalrymple and Nikolaeva (to appear), we also prefer a scale of prominence in distinguishing the gradable topicality, rather than a simple binary feature [± prominence]. We will assume Lambrecht’s (1994) proposal (to be discussed next) in differentiating various types of identifiable referents (active, accessible, or inactive) according to their activation status.

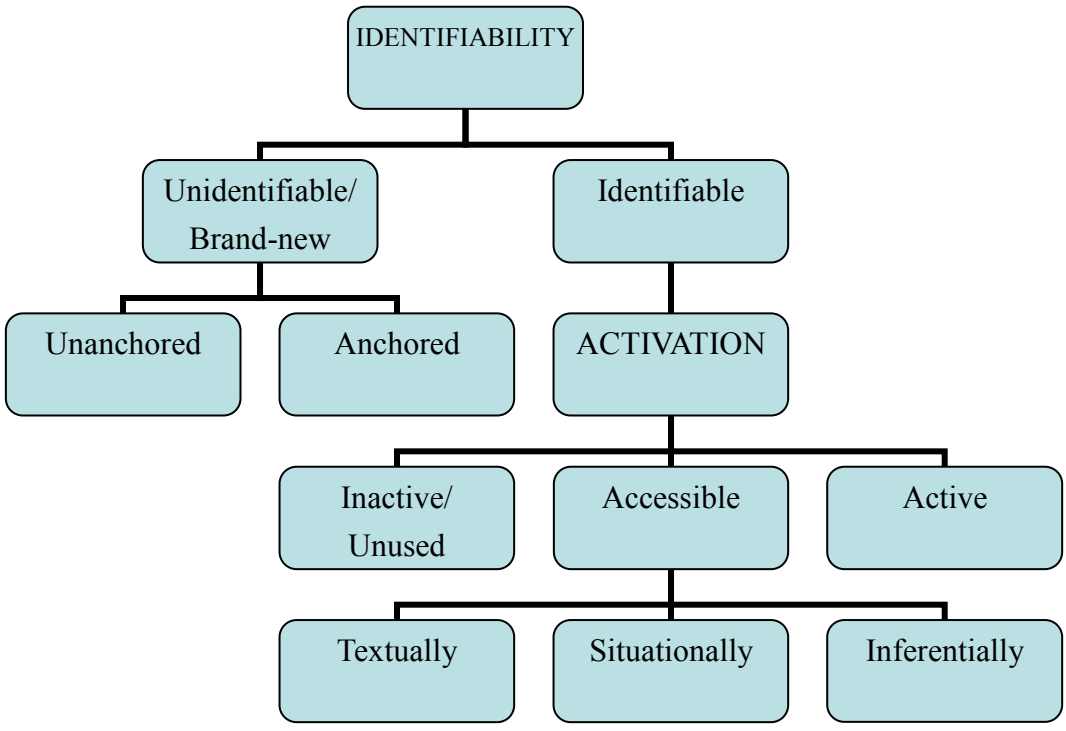
1.5.2 Lambrecht’s information-structure theory

In addition to LFG as the theoretical groundwork assumed in this thesis, we will also rely on Lambrecht’s (1994) theory of information structure to characterize our pragmatic conditions for intersentential zero anaphora. Lambrecht (1994) proposes a theory of the relationship between the form of a sentence and the context, linguistic or extra-linguistic, in which the sentence is used to convey a certain proposition. To be more concrete, Lambrecht treats information structure as a component of sentence grammar and argues that the use of a certain grammatical form, e.g., full lexical NP, unstressed or stressed

reduced form, or even empty pronoun in a sentence, could reflect the speaker's mental representation at the time of the utterance.

Derived from Chafe's (1976, 1987) analysis, Lambrecht distinguishes two information-structure categories: *identifiability* and *activation*. The former concerns a speaker's estimation of whether the discourse representation of a certain referent is *identifiable* or *unidentifiable* to the hearer's mind, whereas the latter is linked to the speaker's evaluation of the status of the representation of an identifiable referent. A cognitively identifiable referent is characterized by its representation being present in the speaker's and the hearer's mind at the time of utterance, while conversely, a cognitively unidentifiable referent is characterized by its representation being absent in the hearer's mind but only existing in the speaker's. Such identifiability is fluid and gradable, and the gradability of identifiability ranges from unidentifiable/brand-new, unidentifiable anchored/brand-new anchored, inactive/unused, accessible, to active/given, as illustrated by the diagram in (47) below.

(47) (Lambrecht 1994, (3.25))



Unidentifiable, or brand-new, entities can be further divided into two groups, depending on whether they are anchored or unanchored. A discourse entity is *anchored* if this noun phrase representing it can be linked to another discourse entity (i.e., the anchor), while a brand-new unanchored entity does not have this association. For example, *a boy* (as in 'a boy won the lottery') is brand-new unanchored, but *a boy in my class* (as in 'a boy in my class won the lottery') is brand-new anchored. As Lambrecht notes (1994: 105), in languages which differentiate definiteness and indefiniteness through grammatical means, brand-new unanchored categories are typically in the form of indefinite NPs, e.g., *a man*, *a house*, whereas brand-new anchored categories usually combine indefinite and definite elements, e.g., *a man I met randomly*, *a friend of mine*. Lambrecht defines an active referent as one that is present at the forefront of the interlocutor's consciousness and is commonly expressed in an unstressed, reduced form like pronominals, inflection, or zero coding. An inactive referent is one whose cognitive status is in the interlocutor's long-term memory and has prosodic and morphological correlates marked by full lexical coding with stress. An accessible referent is one of which the interlocutor has background awareness and is peripherally active in the interlocutor's consciousness. *Accessibility* of a referent can be (i) textually derived (therefore, *textually accessible*) if it is active at an earlier point in the discourse; (ii) inferentially derived (therefore, *inferentially accessible*) if inference can be made via other active or accessible elements in the discourse register; or (iii) situationally derived (therefore, *situationally accessible*) if the accessibility comes from the presence in the text-external world. For example, a textually accessible referent could be the second mention of a certain person who first appeared in the discourse briefly but did not get talked about again till later. An inferentially accessible referent could be any mention of *food*, *service*, *prices*, *location*, etc. when a certain *restaurant* is under discussion, because the former are the associated expectations of the latter. A situationally accessible referent could be any mention of a certain painting in an art gallery to which the hearer might not have paid any attention but could easily locate. All these different mental representations of discourse referents, summarized above by the diagram in (47), have grammatical correlates in the structural forms. The distinction is of great importance in helping us characterize how empty pronouns are pragmatically

conditioned and identifying their relationship with discourse.

Having established the different activation status, let us revisit the example in (46) in 1.5.1.4. We propose to add an ACTV (short for activation) feature inside TOP, which can be valued as active, accessible or inactive. (An active topic can sometimes become inactive if it has not been mentioned in the discourse for a while). We revise the *i*-structure in (46c) as (48) below. We will see more examples of *i*-structures based on Lambrecht's theory in Chapter 4.

(48)
$$\begin{bmatrix} \text{TOP} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{FORM} & \text{he} \\ \text{ACTV} & \text{active} \end{bmatrix} \\ \text{BACKGROUND} & \begin{bmatrix} \text{PRED} & \text{'is writing'} \end{bmatrix} \\ \text{FOC} & \begin{bmatrix} \text{PRED} & \text{'report'} \\ \text{SPEC} & \text{a} \end{bmatrix} \\ \text{COMPLETIVE} & \begin{bmatrix} \text{PRED} & \text{'in > DRef >'} \\ \text{DRef} & \begin{bmatrix} \text{PRED} & \text{'his room'} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

1.6 Summary

Chapter 1 began with the motivation and the purpose of this study, followed by the organization of the thesis. The subsequent preliminaries explained the *pro-drop parameter* before applying it to the Italian-type of *pro*-drop languages and the Chinese-type of *pro*-drop languages; the contrast between the two suggested that resorting to verbal agreement morphology, such as Taraldsen's Generalization and Jaeggli's identification hypothesis, as the licensor of zero anaphora would run into difficulty when accounting for empty categories in Chinese. Next, the distinctions between a full sentence and a minor sentence and between two types of linguistic antecedents (i.e., text-dependent and text-independent) were made in order to establish how an empty category as a genuine occurrence of an empty pronoun is identified in our study and what will be discounted. Structurally speaking, only subjects and objects are allowed to be empty

pronouns in Chinese, which amounts to a generalization that at *f*-structure, only core functions, namely SUBJ and OBJ, can become empty pronouns. We will touch lightly on the issue of licensing empty possessors when discussing the syntactic constraints of zero anaphora in Chapter 3, but the focus of our analysis essentially remains on empty subjects and objects. Finally, two theoretical frameworks assumed in this thesis, LFG and Lambrecht's (1994) information-structure theory, were introduced in 1.5; they provide the foundation upon which we will build our syntactic and pragmatic characterizations of zero anaphora in Chinese. We have also showed how the two theories can be combined. Such integration advances our analysis of the pragmatic conditions Chinese zero anaphora is subject to. We have chosen LFG as our theoretical architecture because the way LFG depicts a grammar, as consisting of different structures, or dimensions, parallel to one another, helps us present our conjunctive view of Chinese empty pronouns neatly. That is, the syntactic properties of empty pronouns can be elegantly represented at *f*-structure, while the pragmatic features of empty pronouns can be precisely captured at *i*-structure. LFG provides a well-defined theoretical ground for us to illustrate the relations and interactions between different dimensions of the grammar when characterizing Chinese zero anaphora. Another utility of LFG is its integratability with Lambrecht's information-structure theory.

Chapter 2: A syntactic investigation of empty pronouns in Chinese

2.1 Topichood

Before we review a selection of previous approaches to Chinese empty pronouns, let us first discuss topichood in Chinese. We will follow Lambrecht's (1994) definition: the topic of a sentence is defined as 'the entity which the proposition expressed by the sentence is ABOUT', whereas the focus of a sentence is 'an element of information which is ADDED TO, rather than superimposed on, the pragmatic presupposition'.

2.1.1 Topichood in LFG

In LFG, the TOP function is understood as the mapping of a class of topical, *c*-structure expressions to the *f*-structure or *i*-structure. The FOC function is realized in the same way, by mapping a class of focal, *c*-structure expressions to the *f*-structure or *i*-structure. TOP and FOC are both known as non-argument, discourse functions. At the syntax-oriented *f*-structure, TOP and FOC are the non-argument, grammaticalized discourse functions, whose presence need not be selected[13] by the predicate, as opposed to argument functions, such as SUBJ and OBJ, whose *c*-structure realization must refer to an argument at *a*-structure. At the pragmatics-motivated *i*-structure, TOP and FOC feature different informative values as discourse functions. Importantly, the values of the information-structural functions TOP and FOC must meet the following Extended Coherence Condition in (49), taken from Bresnan and Mchombo (1987).

(49) Extended Coherence Condition (ECC):
FOCUS and TOPIC must be linked to the semantic predicate argument structure of the sentence in which they occur, either by functionally or by anaphorically binding an argument.

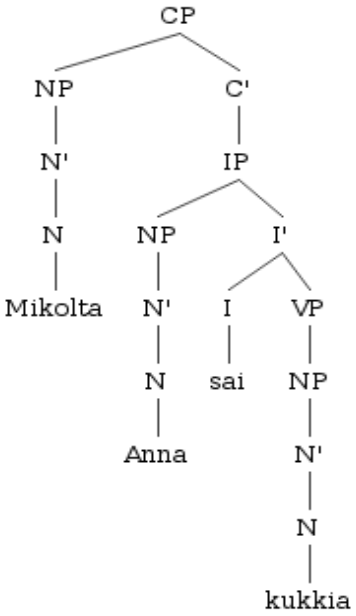
[13] Dalrymple and Nikolaeva (to appear) point out that while verbs commonly subcategorize for argument functions (such as SUBJ, OBJ, and OBL), subcategorization for TOP and FOC is much less common, and may not occur at all. C.-R. Huang (1989), Her (1991), and Culy (1994) are among the very few researchers who argue differently.

For example, cited in Dalrymple (2006), in Finnish, the specifier of IP is associated with the TOP function, whereas the specifier of CP is associated with the FOC function. (50b) and (50c) are the *c*-structure and *f*-structure for the Finnish example in (50a). In (50c),[14] TOP also bears the SUBJ function and FOC also bears the OBL function; the same *f*-structure is used to fill both functions as an instance of *functional binding*. Therefore, TOP and FOC in (50) are licensed in accordance with the ECC.

(50) (= Dalrymple 2006, (34))

a. Mikolta Anna sai kukkia.
Mikko Anna got flowers
'From Mikko, Anna got flowers.'

b.



c.
$$\begin{bmatrix} \text{PRED} & \text{'}sai\text{''get''} < \text{SUBJ, OBJ, OBL} > \text{'} \\ \text{TOP} & {}_1\left[\text{PRED} \quad \text{'Anna'}\right] \\ \text{SUBJ} & {}_1\left[\quad\right] \\ \text{FOC} & {}_2\left[\text{PRED} \quad \text{'Mikolta'}\right] \\ \text{OBL} & {}_2\left[\quad\right] \\ \text{OBJ} & \left[\text{PRED} \quad \textit{kukkia} \quad \text{'flowers'}\right] \end{bmatrix}$$

[14] In the *f*-structure in (50c), the index 1 and 2 indicates that the numbered *f*-structures are identical. In other words, TOP and SUBJ share the same *f*-structure; so do FOC and OBL.

On the other hand, the bracketed NP in (51a) below is a topic anaphorically linked to the resumptive object pronoun *ta* 'him', illustrated by the *f*-structure in (51b).[15] The *anaphoric binding* relationship between the topic and the object pronoun is indicated by coindexation between the two functions.

(51) a. $[\text{Zhangsan}]_i$ wo zui bu xihuan ta_i.
I most not like him
'(Speaking of) Zhangsan_i, I dislike him_i the most.'

b. *f*-structure of (51a)

$$\begin{bmatrix} \text{PRED} & \text{'}bu\ xihuan\text{ 'dislike'} < \text{SUBJ, OBJ} > \text{'} \\ \text{TOP} & \begin{bmatrix} \text{PRED} & \text{'Zhangsan'} \\ \text{INDEX} & i \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{PERS} & 1 \\ \text{NUM} & \text{SG} \\ \text{FORM} & wo\text{ 'I'} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{FORM} & ta\text{ 'him} \\ \text{INDEX} & i \end{bmatrix} \\ \text{ADJ} & \{\begin{bmatrix} \text{PRED} & \text{'}zui\text{ 'the most'} \end{bmatrix}\} \end{bmatrix}$$

Observing the ECC, the discourse functions in (50) and (51) do follow the pattern of saturating a non-argument position and binding an argument either functionally or anaphorically at *f*-structure. In addition to this kind of functional or anaphorical linkage, there is a third possibility for a discourse function to fulfill the coherence condition by being associated with an argument at *i*-structure. Take (52) for example, in which the main verb *mafan* 'to trouble' is typically a two-place[16] predicate. At first sight, the topic

[15] In the *f*-structure in (51b), TOP and OBJ are coreferential. Following Bresnan's (2001) practice, coreference is indicated by sharing the same index (e.g., *i* in the case (51b)). Being coreferential is not the same as sharing the same *f*-structure (as in (50c)), so the two notions are annotated differently.

[16] Shi (2000) analyzes the sentence in (52) as involving a topic NP unsubcategorized by the main verb

NP is not related to any structural position inside the comment clause, either functionally or anaphorically, and the comment clause looks complete on its own. Therefore, the topic appears to be a dangling topic and violates the ECC.

(52) (Shi 2000, (32))

[Zhe-jian shiqing]$_{\text{Topic}}$ [ni bu neng guang mafan yige ren]$_{\text{Comment}}$.
this-CL matter you not can only punish one person
'(As for) this matter, you cannot just trouble one person, (meaning that this matter requires more than one person's efforts).'

While it is true that *mafan* 'to trouble' is typically a two-place predicate in Chinese (just like in English), we argue that *mafan* 'to trouble' can also be ditransitive by taking an extra infinitival complement, as the bracketed clause in (53a) indicates; or an NP complement, as in (53b). When this complement is not present and *mafan* 'to trouble' acts as a two-place predicate, as in (53c), the sentence is also well-formed.

(53) a. Bu-yao yizhi mafan bie-ren [bang ni jiejue wenti]$_{\text{Complement}}$.
don't constantly trouble other-people help you solve problems
'Stop troubling other people constantly to help you solve problems.'

b. Bu-yao mafan renjia [zheme duo wenti]$_{\text{NP}}$.
don't trouble other.people so many questions
'Stop troubling other people with so many questions.'

c. Bu-yao mafan renjia.
don't trouble other.people
'Stop troubling other people.'

Despite the syntactic transitivity on the surface in the case of (53c), examples like those in (53a) and (53b) validate the hypothesis of treating *mafan* 'to trouble' as a ditransitive verb. Such ditransitivity could be explained by the following reasoning: conceptually speaking, it would be difficult to imagine how one could manage to trouble someone else without involving any (inconvenient) activity or behavior. That is, the action of 'to trouble' should conceptually consist of (i) a troubler, one who troubles

mafan 'to trouble'; the comment clause contains no gap or resumptive pronoun for the said NP to relate to.

someone else; (ii) a troublee, someone who is troubled; and (iii) something with which the troubler troubles the troublee. When this third element is left out in a sentence, as in (53c), the unmentioned argument is normally something salient in the discourse/context and is easily recoverable, e.g., *I saw John asking his friends to do him a favor when they were all clearly very busy with work*, so I uttered (53c) to *John*. It would be pragmatically odd if (53c) is uttered without any sort of behavior going on that is known to the speaker at least in order to prompt the speaker's order on the hearer. Following this logic, we propose that the example in (52) can be reanalyzed as that in (54), in which the problematic non-gapping topic is treated as an object subcategorized by *mafan* 'to trouble'; by placing it in its original unmarked position inside the comment clause, as (54) exemplifies, the apparent violation of the ECC is resolved.

(54) Ni bu neng guang mafan yige ren [zhe-jian shiqing]$_{NP}$.
you not can only punish one person this-CL matter
'(As for) this matter, you cannot just trouble one person.'

(55) is a variant of (53a). Crucially, if we compare the two examples, the ungrammaticality of (55), in which there is an additional topic *zhe-jian shiqing* 'this matter' – the hypothesized argument subcategorized by *mafan* 'trouble', shows that the main predicate *mafan* 'trouble' is unable to take on this extra, topical argument. If *zhe-jian shiqing* 'this matter' in (55) were a mere unsubcategorized topical element, the resulting sentence should have been grammatical, but it is not. This strongly supports our view of treating the sentence-initial NP in (52) as a topicalized argument of the main verb *mafan* 'trouble.

(55) *[Zhe-jian shiqing], bu-yao yizhi mafan [bie-ren] [bang ni
this-CL matter don't constantly trouble other-people help you
jiejue wenti]$_{Complement}$.
solve problems
'As for this matter, stop troubling other people constantly to help you solve problems.'

Analyzed this way, the seemingly dangling topic in (52) can be either functionally licensed at *f*-structure, as (56a) demonstrates, or semantically licensed at *i*-structure, as (56b) shows.

(56) a. *f*-structure of (52)

$$\begin{bmatrix}
\text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{PERS} & 2 \\ \text{NUM} & \text{SG} \\ \text{FORM} & \textit{ni}\text{'you'} \end{bmatrix} \\
\text{TOP} & {}_1\begin{bmatrix} \text{PRED} & \text{'}\textit{shiqing}\text{ "matter"'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}\textit{zhe}\text{ "this"'} \end{bmatrix} \right. \\ & \left. \begin{bmatrix} \text{PRED} & \text{'}\textit{jian}\text{ "CL"'} \end{bmatrix} \right\} \end{bmatrix} \\
\text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'}\textit{ren}\text{ "person"'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}\textit{yi}\text{ "one"'} \end{bmatrix} \right. \\ & \left. \begin{bmatrix} \text{PRED} & \text{'}\textit{ge}\text{ "CL"'} \end{bmatrix} \right\} \end{bmatrix} \\
\text{PRED} & \left[\text{'}\textit{mafan}\text{ "trouble" } < \text{SUBJ, OBJ, OBJ} >\text{'}\right] \\
\text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}\textit{neng}\text{ "can"'} \end{bmatrix} \right. \\ & \left. \begin{bmatrix} \text{PRED} & \text{'}\textit{bu}\text{ "not"'} \end{bmatrix} \right. \\ & \left. \begin{bmatrix} \text{PRED} & \text{'}\textit{guang}\text{ "only"'} \end{bmatrix} \right\} \\
\text{OBJ} & {}_1[\quad]
\end{bmatrix}$$

b. *i*-structure of (52)

$$\begin{bmatrix}
\text{TOP} & {}_1\begin{bmatrix} \text{PRED} & \text{'}\textit{shiqing}\text{ "matter"'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}\textit{zhe}\text{ "this"'} \end{bmatrix} \right. \\ & \left. \begin{bmatrix} \text{PRED} & \text{'}\textit{jian}\text{ "CL"'} \end{bmatrix} \right\} \end{bmatrix} \\
\text{FOCUS} & \begin{bmatrix}
\text{PRED} & \text{'}\textit{mafan}\text{ "trouble"} < \text{DRef1, DRef2, DRef3} >\text{'} \\
\text{DRef1} & \begin{bmatrix} \text{PRED} & \text{'}\textit{ni}\text{ "you"'} \end{bmatrix} \\
\text{DRef2} & \begin{bmatrix} \text{PRED} & \text{'}\textit{ren}\text{ "person"'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}\textit{yi}\text{ "one"'} \end{bmatrix} \right. \\ & \left. \begin{bmatrix} \text{PRED} & \text{'}\textit{ge}\text{ "CL";} \end{bmatrix} \right\} \end{bmatrix} \\
\text{DRef3} & {}_1[\quad] \\
\text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}\textit{neng}\text{ "can"'} \end{bmatrix} \right. \\ & \left. \begin{bmatrix} \text{PRED} & \text{'}\textit{bu}\text{ "not"'} \end{bmatrix} \right. \\ & \left. \begin{bmatrix} \text{PRED} & \text{'}\textit{guang}\text{ "only"'} \end{bmatrix} \right\}
\end{bmatrix}
\end{bmatrix}$$

In the above discussion, we argued for the ditransitivity of the verb *mafan* 'to trouble' in taking a third argument on the assumption that at the level of *f*-structure and *i*-structure, *mafan* 'to trouble' is ditransitive in order to complete its proposition. The apparent transitive use of *mafan* 'to trouble' (e.g., A troubled B) is possible because the reason behind A's troubling B might already be salient. If it is unclear, it is very likely for the hearer to ask for the cause of such troubling immediately in the subsequent discourse. Importantly, coindexation of DRef3 and the topic in the *i*-structure of (56b) represents a semantically licensed topic and exemplifies how it is linked to an argument at *i*-structure. The ECC is thus met. A semantically licensed topic associated with an argument at *i*-structure manifests an alternative path of linking a discourse function (TOP or FOC) to an argument of the sentence in which the topic or focus appears. This amounts to saying that the binding relationship holding between a discourse function and an argument in the same sentence can also occur at *i*-structure, which may be added to the original claim in the ECC. To conclude, a discourse function like TOP can not only be syntactically licensed through functional or anaphorical linkage to an argument at *f*-structure, but can also be semantically identified through the role it plays at *i*-structure.

2.1.2 Topichood in Chinese

Since Chinese, as a topic-prominent language, is characterized by its extensive use of topics (Li and Thompson 1976, 1989), we feel the necessity to elaborate on the notion of topic. We start with the identification of a topic, that is, how a topic is encoded in Chinese.

King and Zaenen (2004) observe three major ways to encode discourse functions: (i) by structurally occupying a specific position, e.g., pre-verbal, post-verbal, or sentence-initial; (ii) by the use of discourse markers or particles to indicate the topical or focal status of a constituent; or (iii) by certain intonation patterns. It is common for languages to combine these encoding techniques; for example, a focal element may saturate a post-verbal position while also bearing a special intonation pattern or stress. Chinese is one of the languages which employ a mixture of the above three methods to

mark a discourse function. For instance, in terms of structure, Chinese topics indeed occur sentence-initially, even though this sentence-initial position is not reserved exclusively for topics, because under some circumstances focus can also be fronted to the sentence-initial position. Secondly, regarding discourse markers, even though there is no topic marker to be used obligatorily to indicate the topical status in Chinese, unlike Japanese, in which the morpheme *wa* is invariably suffixed to a topical constituent, the topical status of an NP in Chinese can be identified by the use of one of the following four pause particles: *a* (and its variant *ya*), *ne*, *me*, and *ba*, attached to the end of a topical NP, separating the topic from the following comment clause (Tsao 1990; Chu 1998). The example in (57) below illustrates how a pause particle separates the topic from the comment clause.

(57) [Zhangsan]$_{\text{Topic}}$ a/ya, [zongshi shuohua bu laoshi]$_{\text{Comment}}$.
PART always speak not honestly
'(Speaking of) Zhangsan, he never speaks honestly.'

Note that suffixation to mark a topic is not compulsory in Chinese; the use of the pause particle is only optional. As for the third method of encoding a discourse function via a special intonation pattern, it is very common for a focused constituent to receive stress. Being in focus in Chinese can be expressed on a par with the English example below, taken from King and Zaenen (2004). The capitalized word indicates that this word is focally used and receives stress.

(58) A: Did you see Mary or John?
B: I saw JOHN.

After showing how Chinese topic can be encoded via the three aforementioned methods, now we move on to the characterization of topichood in Chinese, and how it differs from subjecthood, despite their many similarities. It is generally agreed that topic is a discourse notion (hence the discourse function of TOP assumed in LFG), whereas subject is a syntactic notion (hence the grammatical function of SUBJ in LFG). The two do not exclude each other, as it is very common for an NP to carry both roles in a sentence, especially when subjects in English are usually treated as topics (but not vice

versa) unless otherwise indicated. This accounts for the fact that SUBJ is the only function that can be grammatical as well as discoursal (Bresnan 2001). For a long time in the literature, the definition of topic and subject in Chinese and their difference have been fuzzy and controversial. One of the most recognized attempts comes from Tsao's (1990) functional definition of topic and subject. According to him, Chinese subjects can be characterized as having the following properties listed in (59), and Chinese topic can be identified by those in (60). A number of similarities shared by topic and subject can be spotted easily.

(59) Properties characteristic of Chinese subject (Tsao 1990: 170):

a. Subject is always unmarked by preposition.
b. By position, subject can be identified as the animate NP to the left of the verb; otherwise, the NP immediately before the verb.
c. Subject always bears some selectional relation to the main verb of the sentence.
d. Subject tends to have a specific reference.
e. Subject plays an important role in the following coreferential pronominalization or deletion processes: reflexivization, imperativization, and Equi-NP deletion.

(60) Properties characteristic of Chinese topic (Tsao 1990: 170):

a. Topic invariably occupies the sentence-initial position of the first clause in a topic chain.
b. Topic can optionally be separated from the rest of the clause in which it overtly occurs by one of the four pause particles: *a*(*ya*), *ne*, *me*, and *ba*.
c. Topic is always definite[17] or generic.

[17] Definiteness in Chinese can be encoded by the use of proper nouns or pronominals (including possessives), or by the use of a demonstrative, such as *zhe* 'this' or *na* 'that'. However, not all definite nouns need to be explicitly marked as such, because nouns in Chinese, including singular forms, are not always required to have determiners. The subjects in (i) and (ii) below, both in their bare form, are definite in reference, despite the apparent lack of definite marking. Both sentences are also ambiguous between a singular and plural reading because neither NP is marked for number whilst being countable. The indefinite determiner '*yi* "one" + classifier' is the most approximate to an indefinite article, such as *a/an* in English (Chen 2003). Indefinite NPs cannot be topics, unless they are generic, as in (iii). In Chinese, when an indefinite NP becomes a subject, it must be prefixed by the existential marker *you* 'existential there', as in (iv) below. Without it, the remaining sentence would be grammatically marked. Objects are not subject to this requirement, as illustrated by (v).

(i) Shu zai zhuo-shang.
book at table-on
'The book/books is/are on the table.'
(ii) Keren lai le.
guest come ASP
'The guest/guests arrived.'

d. Topic is a supraclausal notion; it may, and often does, extend its semantic domain to more than one clause.
e. Topic is in control of the pronominalization or deletion of all the coreferential NPs in a topic chain.
f. Topic, except in clauses where it is also a subject, plays no role in such processes as reflexivization, Equi-NP deletion, or imperativization.

The above characterization has been challenged by Chu (1998) as loosely defined in that some of the proposed properties are not differentiating or exclusive enough to truly define topic or subject. To name a few examples, even though subject is indeed always unmarked by a preposition in Chinese, as (59a) suggests, topic can also be commonly unmarked by a preposition, as the marking on topic is only optional in Chinese. This argument is supported, because we can easily refer to most examples involving a topic cited in this thesis and find them unmarked by any preposition. Chu (1998: 249) also adds that the sentence-initial position (see (60a)) and the pre-verbal position (see (59b)) are reserved typically for both topic and subject in Chinese. The structural difference suggested by Tsao, i.e., the sentence-initial position for topic and pre-verbal position for subject, is only at work when a topic and a subject stand separately in the same clause. More detailed evaluation of Tsao's work by Chu can be found in Chu (1998). Three remaining properties have survived from Chu's examination as the most essential and characteristic of topic, listed in (61) below.

(61) (Chu 1998, (7))
a. Topic can be optionally separated from the rest of the clause by a pause or one of the four pause particles: *a*(*ya*), *ne*, *me* and *ba*.
b. Topic is a supraclausal notion: it extends its semantic domain to more than one clause.

(iii) Yi-tai dian-nao$_i$ hen gui. Ni buyao nong-huai Ø$_i$ le.
one-CL computer very expensive you don't make-spoiled ACC ASP
'A computer is very expensive. Don't break (it).'
(iv) You/*(you) yi-ge ying-er gang chusheng le.
there.be one-CL baby just.now born ASP
'A baby was born just now.'
(v) Wo xiangyao zhe-ben/yi-ben shu.
I want this-CL/one-CL book
'I want this book/a book.'

c. Topic is in control of the pronominalization or deletion of all the coreferential NPs in a topic chain.

Among the above three essential properties of topic in Chinese, (61b) and (61c) are closely related to the Topic-NP Deletion Rule proposed by Tsao (1977). The rule states that the topic of a sentence can be deleted if this topic is the same as the one already established in the discourse, and the deleting process forms a *topic chain*. Tsao's Topic-NP Deletion Rule can be illustrated by the often cited example in (62) below, in which *Zhongguo* 'China' is the supraclausal topic, and its occurrences in the subsequent clauses are deleted under coreference. (Note that in the last bracketed clause of (62), the empty pronoun is a fronted object that has been topicalized).

(62) (J. Huang 1984, (55))

[Zhongguo$_i$, difang hen da], [Ø$_i$ renkou hen duo], [Ø$_i$ tudi hen feiwo],
China place very big NOM population very many NOM land very fertile
[Ø$_i$ qihou ye hen hao], [Ø$_i$ women dou hen xihuan].
NOM climate too very good ACC we all very like
'(As for) China, (its) land area is very large; (its) population is very big; (its) climate is also very good; we all like (it).'

Although the above three properties in (61) seem to serve as reliable indicators of what Chinese topic may be like, Chu (1993, 1998) continues to argue for a more general definition of Chinese topic, and proposes an approach relying on the prototypicality of topic in Chinese. In Chu's view, topic can have various degrees of topicality; that is to say, some NPs will be more typically usable as topics if they exhibit more properties (or *attributes*, in Chu's term) of topichood, whilst NPs with fewer attributes of topichood will be less typically usable as topics (Chu 1998: 253). This definition flexibly allows an NP to be more (or less) topical and prototypical than another and, therefore, can include the possibility of a marginally used topic, or a less prototypical one, when it only possesses very few attributes of topichood. Chu proposes the following attributes of topichood in Chinese.

(63) (Chu 1998: 253, (11))

Attributes of a Mandarin Chinese topic:

A. Primary Attributes:
   a. Being a nominal.
   b. Serving as an interclausal link.
B. Secondary Attributes:
   a. Being specific/referential.[18]
   b. Occupying the sentence-initial/pre-verbal position.
C. Non-essential Quality: bearing no selectional relations to the predicate verb.

Chu admits that the above attributes are essentially elicited from predecessors' work, but he claims that these attributes have been validated through many tests on different kinds of topics and surveys. The new contribution also lies in these *hierarchical* attributes because they reflect that some topical features are closer to the core of topichood than others, which accords with his claim that the criteria for Chinese topic are hierarchical in nature, and that any given Chinese topic need not exhibit every attribute on this list (Chu 1998: 254).

The conclusion we would like to draw from Tsao's insightful observations of the properties of subject and topic and Chu's hierarchical perspective on the attributes of topichood is that subject and topic do share similar features and can overlap sometimes. A subject may appear to cause the deletion of the same subjects occurring in subsequent clauses, as if there were a Subject Deletion Rule, if this subject also happens to be the topic. Conversely, a topic could seem to hold a selectional relation to the verb predicate if

[18] Chu (1998) does not define his notion of *specificity* and *referentiality* as attributes of Chinese topics, but we suspect that at least proper names, pronouns, definite NPs, and generic NPs would be included in his classification. For a more detailed discussion of *specificity*, refer to Enç (1991), in which she defines specificity as involving 'being a subset of or standing in some recoverable relation to a familiar object' (Enç 1991: 24).

(i) (= Chu 1998: 254, (12))
Qiang-shang pazhe hen duo bihu.
wall-on crawl very many salamanders
'The wall has many salamanders on it/there are many salamanders crawling on the wall.'

In the example above, Chu considers the sentence-initial NP *qiang* 'wall' definite and, therefore, specific, and he considers the other NP *bihu* 'salamanders' specific. The former has one primary attribute (being nominal) and two secondary attributes (being sentence-initial and specific), whereas the latter also has the same primary attribute but only one secondary attribute (being specific). Therefore, *qiang* 'wall', as the stronger candidate, is more likely to become a topic.

this topic happens to be the subject, too. What Tsao and Chu propose as properties or attributes of topichood are both correct essentially, and the two do not diverge much in spirit. Crucially, they both help us distinguish between subject and topic. Topic, the most salient referent in the discourse, plays the role commented on by the rest of the sentence in which it occurs, and can have a supraclausal domain at *i*-structure. Subject, the syntactically required argument by a predicate, is closely tied to syntactic mechanisms such as reflexivization, control, and formation of imperatives. In this section we see that a topic can be licensed anaphorically, functionally, or at *i*-structure. It shares many properties with a subject. We will see in the subsequent parts of the thesis how being a topic sanctions the use of an empty pronoun in Chinese.

2.2 A review of previous solutions and their problems

This section reviews a selection of previous approaches to account for empty pronouns in Chinese, from a purely syntactic point of view by J. Huang (1984, 1987, 1989, 1991), a purely pragmatic one by Y. Huang (1995), and a syntactic-pragmatic one by Pan and Hu (2000). These analyses aim to tackle empty pronouns in Chinese, but they are not without problems. We will discuss their weakness before arguing for the necessity for an integrating account in which both syntax and pragmatics need to be taken into consideration to yield a more complete view of empty pronouns in Chinese.

2.2.1 A syntax-oriented account by J. Huang

The renowned work by J. Huang (1984) is one of the most representative, syntax-driven analyses of Chinese empty pronouns, and this line of analysis is also further extended in later papers (e.g., J. Huang 1987, 1989, 1991), which, along with the previous paper, all maintain the classical P&P Theory (Chomsky 1981; Chomsky and Lasnik 1993). Despite the common misconception that Chinese-type of *pro*-drop seems to behave freely with hardly any constraint due to the well-known absence of verbal agreement morphology, J.

Huang (1984) has insightfully pointed out that the distribution of empty pronouns in Chinese is *not* entirely free. Furthermore, he claims that empty subject pronouns and empty object pronouns behave distinctly, and the case of *pro*-drop in Chinese only challenges the agreement-based theory (e.g., the aforementioned Taraldsen's (1978) Generalization and Jaeggli's (1982) identification hypothesis) insomuch as Chinese, as a non-inflectional language, allows empty subject pronouns but *not* empty object pronouns. In J. Huang's opinion, Chinese does allow subject pronouns to be null, but he rejects the existence of object *pro*.[19] Instead, he treats the seemingly null object pronominals as variables. To support his argument that the nature of an empty category (EC) in subject position is different from that in object position, J. Huang proposes that the subject-object asymmetry, in which a subject EC and an object EC are found to behave distinctly, is to be conditioned by different constraints. Pursuing J. Huang (1984), let us consider the following sentences containing ECs of different status in Chinese.

(64) a. Ø bu-jian le.
NOM not-see ASP
'(Someone/something) is lost/missing.'
b. John$_i$ kanjian Ø$_{j/*i}$ le.
see ACC ASP
'John saw (someone).'
c. John$_i$ shuo [Ø$_{i/j}$ hui qu baifang Mark].
say NOM will go visit
'John$_i$ said that (he$_i$/someone) would go visit Mark.'
d. John$_i$ shuo [Mark$_j$ ye zhidao Ø$_{*i/*j/k}$].
say also know ACC
'John said that Mark also knew (someone).'

The EC in (64a) and (64b) both refer to someone whose reference is recovered from the discourse outside the sentence itself. In (64c), it is J. Huang's grammatical judgment that the reference of the embedded subject EC could be the matrix subject *John*, or

[19] In the classical P&P Theory, adopted by J. Huang (1984, 1987, 1989, 1991), the categories of non-overt nominal positions recognized are NP-trace, *wh*-trace, PRO, and *pro*. An empty category (EC) occurring in object position appears to be a *pro* in Chinese, but J. Huang analyses it otherwise as a bound variable.

someone else who is salient or has been mentioned in the discourse. We find the empty subject EC in (64c) unambiguously coreferential with the matrix subject *John* and do not agree with J. Huang's judgment in allowing an alternative reading. We will fully discuss this issue in Chapter 3 when talking about the syntactic constraint for empty subordinate subject pronouns. In contrast, the object EC in (64d) can never be the matrix subject *John*; the referent must be someone else in the discourse. At this point, we have observed that there seem to be more restrictions on an object EC than on a subject EC, because if either of them occurs in an embedded complement clause, only the embedded subject EC is allowed to be coindexed with the matrix subject, while the object EC is forbidden from being so. Instead, an overt pronoun has to replace the object EC in question in order to be coreferential with the matrix subject. This is the subject-object asymmetry proposed by J. Huang, which can be further exemplified by (65).

(65) a. John$_i$ shuo [ta$_i$/Ø$_i$ cong-mei tingshuo guo Mark].
say he ever-not hear.of ASP
'John$_i$ said that he$_i$/(he$_i$) never heard of Mark.'
b. John$_i$ shuo [Mark cong-mei tingshuo guo ta$_i$/*Ø$_i$.
say ever-not hear.of ASP him
'John$_i$ said that Mark never heard of him$_i$/*(him$_i$).'

Similar subject-object asymmetries can also be found in other East-Asian languages including Japanese (see Kuroda (1965) for discussion). According to J. Huang, the interpretation of an object EC is limited to a reference introduced in the previous discourse, while coindexation with the matrix subject is illicit. J. Huang (1984), therefore, suggests that the discourse topic, i.e., the most prominent person or thing in a given discourse, plays an important role in identifying the content of empty objects in Chinese. In view of this asymmetrical distribution, J. Huang proposes that the status of an empty object is different from that of an empty subject. He treats an empty subject as a *pro* and excludes an empty object from being considered an empty pronominal, treating it, instead, as a variable (bound by a discourse topic). Hence, an empty object is not a *pro* in his analysis. J. Huang (1991) gives two other pieces of evidence in support of his

empty-object-as-variable hypothesis. The first piece of evidence comes from the observation that an overt pronoun cannot be associated with a quantificational NP as a bound variable in a language allowing *pro*-drop, which J. Huang takes from Montalbetti's (1984) generalization known as the Overt Pronoun Constraint (OPC). The OPC states that 'overt pronouns cannot be linked to formal variables, if and only if the alternation of overt/empty pronouns obtains'. This constraint is witnessed in *pro*-drop languages such as Japanese (Saito and Hoji (1983)), Chinese (J. Huang (1991), Italian and Spanish (Montalbetti (1984)), in relation to non-*pro*-drop language like English. The contrast between Chinese and English is illustrated below in (66) and (67).

(66) (J. Huang 1991, (46, 48))

a. *Shei$_i$ xiwang [ta$_i$ neng kanjian Lisi]?
who hope he can see
'Who$_i$ hopes that he$_i$ can see Lisi?'

b. Shei$_i$ xiwang [Ø$_i$ neng kanjian Lisi]?
who hope NOM can see
'Who$_i$ hopes that (he$_i$) can see Lisi?'

(67) Who$_i$ hopes that he$_i$ can meet Lisi?

However, as J. Huang argues, Chinese, as a *pro*-drop language, only falls under the OPC insomuch as the alternation of overt/empty pronouns in subject position holds, as seen in (66a) and (66b). When the alternation of overt/empty pronouns fails, shown in the ungrammatical use of an empty object in (68), the interpretation between the quantificational NP *meiyou-ren* 'nobody' and the embedded overt object pronoun *ta* 'him' becomes acceptable in (68).

(68) (J. Huang 1991, (51))

Meiyou-ren$_i$ shuo [Lisi hen xihuan ta$_i$/*Ø$_i$].
no-person say very like him
'Nobody$_i$ says that Lisi likes him$_i$/*(him$_i$) very much.'

Following the generalization of the OPC and the grammaticality of coindexation between

an overt object pronoun and a quantificational NP, J. Huang concludes that there is no alternation of overt/empty pronouns in object position in Chinese, and *pro* is excluded from object position.

The second piece of evidence comes from the marked status of an emphatically used reflexive (e.g., *ta* + *ziji* 'he/she + self, translated as himself or herself') in object position and its null counterpart, i.e., Ø + *ziji*. As (69a) and (69b) show, the emphatic construction occurring in subject position is well-formed. Either form, 'overt pronoun + *ziji*' or 'Ø + *ziji*', is permitted. According to J. Huang, this implies the grammaticality of the alternation of overt/empty pronouns in subject position.

(69) (Adapted from J. Huang 1991, (63))

a. Zhangsan$_i$ shuo [ta+ziji$_i$ hui hui jia].
say he+self can return home
'Zhangsan$_i$ said that he$_i$ himself can go home.'

b. Zhangsan$_i$ shuo [Ø+ziji$_i$ hui hui jia].
say NOM+self can return home
'Zhangsan$_i$ said that he$_i$ himself can go home.'

On the other hand, the emphatic construction occurring in object position, as indicated in (70), only has the intended reading when the form 'overt pronoun + *ziji*' is used, as in (70a), and this intended reading (emphatically coreferential with *Zhangsan*) becomes unavailable when 'Ø + *ziji*' is used, shown in (70b). In other words, the emphatic use of 'Ø + *ziji*' is unmarked if it appears in subject position, but marked if in object position. Following this line, J. Huang excludes *pro* from object position, yielding another example of the subject-object asymmetry.

(70) (Adapted from J. Huang 1991, (64))

a. Zhangsan$_i$ shuo [wo zhi piping ta+ziji$_i$].
say I only criticize him+self
'Zhangsan$_i$ said that I only criticized him$_i$ himself.'

b. Zhangsan$_i$ shuo [wo$_j$ zhi piping Ø +ziji$_{*i/j}$].
say I only criticize ACC+self
'Zhangsan$_i$ said that I only criticize myself/*him$_i$ himself.'

Based on the above observation, J. Huang concludes that an object EC should not be a pronominal. Instead, he proposes an EC-as-variable analysis, that is, an empty category in Chinese may be most naturally bound by a discourse topic as a variable. Specifically, the empty category is coindexed with an overt topic in (71a); mirroring the structure of (71a), (71b) is hypothesized to have a sentence-initial empty topic which functions like an operator according to J. Huang's analysis. The EC in (71b) is bound by the operator (in this case, the topic) that is phonetically null itself.

(71) a. [Neige ren_i]$_{Topic}$, John shuo Mark ye zhidao [$Ø_i$]$_{EC}$.
that person say also know ACC
'That person$_i$, John said that Mark also knew $Ø_i$.'
b. [$Ø_i$]$_{Operator}$, John shuo Mark ye zhidao [$Ø_i$]$_{EC}$.
TOP say also know ACC
'(Someone$_i$), John said that Mark also knew $Ø_i$.'

Because the object EC in (71b) is assumed to be bound by the empty topic operator, which is a non-argument, this line of analysis is also known as the null-topic analysis, while the kind of empty category bound by a discourse topic, overt or empty, is *not* a pronominal but a bound variable in J. Huang's hypothesis. Additionally, as far as object ECs are concerned, J. Huang provides a second explanation in his later work, in which an empty object in Chinese is analyzed as a null epithet because empty objects seem to share various properties with epithets, e.g., (i) it may not be A-bound, (ii) it may be A'-bound, and (iii) it need not be A'-bound (J. Huang 1991: 61). (72a), (72b) and (72c) demonstrate these three possibilities, respectively.

(72) a. *$John_i$ mei kandao [zhege bendan]$_i$/$Ø_i$.
not see this idiot/ACC
'$John_i$ didn't see this $idiot_i$/($himself_i$).'
b. $John_i$, hen duo ren dou taoyan [zhege pianzi]$_i$/$Ø_i$.
very many people all dislike this liar/ACC
'(Speaking of) $John_i$, very many people all dislike this $liar_i$/(him_i).'
c. Ni mei kanjian [nage pianzi]/Ø ma?
you not see that liar/ACC Q
'Didn't you see that liar/(someone salient in the previous discourse)?'

We find this null-epithet account by J. Huang similar, in spirit, to his EC-as-variable analysis in the sense that they both reject the pronominal status of an empty object, and both null epithets and null variables exhibit properties shared by referential expressions. To conclude, an empty object in Chinese is analyzed by J. Huang either as a variable bound by a topic operator in his EC-as-variable hypothesis, or a null epithet in his null-epithet hypothesis.[20]

As for the status of an empty subject, J. Huang treats it as an unambiguously genuine *pro* regulated by the Generalized Control Rule (GCR) and the Disjoint Reference Principle, which rules that a pronoun must be free in its governing category. The GCR, first formulated in J. Huang (1984) and then elaborated in J. Huang (1989), states that an empty pronominal is controlled in its control domain (if it has one). His definition of a control domain is as follows: α is the control domain for β if and only if it is the minimal category that satisfies both (i) α is the lowest S(entence) or NP that contains β, or the minimal maximal category containing β; and (ii) α contains a SUBJECT[21] accessible to β. This revised definition of the GCR allows flexibility for an empty pronoun to possibly have no control domain. In other words, an empty pronoun need not always be controlled and may have a long-distance antecedent or an arbitrary reference when it does not have a control domain and is free. For instance, (73a) below demonstrates two available readings, in J. Huang's grammatical judgment, not ours, for the same sentence due to the ambiguous reference of the empty subject. In the first reading (indicated by the

[20] On the status of null object in Chinese, in addition to the EC-as-variable analysis and the null-epithet analysis, J. Huang (1991) offers a third perspective, the VP-ellipsis account, in which an object gap is analyzed as a VP gap. For example, the missing object in (i) is in fact an empty VP, not just an empty object, and the sentence on the whole has dichotomous (strict and sloppy) readings. The repetition of the verb in the second clause, *kanjian* 'see', may seem lexical, but Huang analyzes it as an auxiliary functioning as '*do*-support'. (See J. Huang 1991 for further discussion).

(i) (J. Huang 1991, (31))
[John kanjian le tade mama]. [Mary ye kanjian le Ø].
see ASP his mother also see ASP ACC
'John saw his mother, and Mary did, too.'

However, since the nature of this kind of VP-ellipsis structure crucially involves an empty VP rather than an empty pronoun, we will not look into it in this thesis and leave it for future research.

[21] SUBJECT, the big subject, is defined as the abstract agreement feature (Agr) of a finite clause (Haegeman 1994: 218). For instance, there will be no SUBJECT in non-finite clauses such as infinitivals. On the other hand, to say that A is an accessible SUBJECT for B, it means that the coindexation of A and B does not violate any grammatical principles (Haegeman 1994: 220).

subscripted *i* index), the embedded empty subject follows the prediction of the GCR and is a controlled empty pronominal. The minimal clause that contains this *pro* and a SUBJECT accessible to the empty pronoun in question is the matrix clause. In an attempt to maintain the agreement-based theory on resolving empty pronouns with the help of the GCR, J. Huang argues that even though there is no verbal agreement (Agr) in (73a), there is still an accessible SUBJECT, *John*, to control the reference of the *pro*. The *pro* has the matrix clause as its control domain, so this *pro* must be controlled unambiguously by *John*.

(73) a. $John_i$ shuo [$Ø_{i/j}$ mingtian yao qu pa shan].
say NOM tomorrow will go climb mountain
'$John_i$ said that (he_i/$someone_j$) was going mountain-climbing tomorrow.'
b. [$Ø_j$]$_{Operator}$ John shuo [$Ø_j$] mingtian yao qu pa shan.
TOP say NOM tomorrow will go climb mountain
'John said that (someone) would go mountain-climbing tomorrow.'

In the second reading of (73a), indicated by the *j* index, the empty subject refers to someone else salient in the discourse, presumably a topic referent. J. Huang's GCR seems to fail to predict this interpretation, where the empty subject is not controlled despite clearly having a control domain. Following J. Huang's grammatical judgment on (73a), the sentence itself is ambiguous between two readings, and a context is required to determine who the antecedent is. Although we see the possibility of having two readings in (73a) as a problem undermining his GCR analysis, J. Huang views the second reading as unproblematic, as 'nothing seems to prevent this [the coindexation of the empty subject and an empty topic]' (J. Huang 1984: 556). (73b) illustrates the structure hypothesized for the second reading of (73a), in which the empty subject is bound by an empty topic operator.

To conclude, J. Huang presents a purely syntactic account in which different constraints are proposed for empty subjects and empty objects. An empty subject in Chinese is treated as a genuine *pro* subject to the Disjoint Reference Principle and the Generalized Control Rule. This empty pronoun can be controlled if it has a control

domain, or uncontrolled when it does not have one. An empty object, on the other hand, is excluded from being a pronominal (hence, the subject-object asymmetry), and it can be analyzed as a null epithet or a variable bound by an empty discourse topic. Another importance of the GCR is that it further helps J. Huang exclude an empty object from being a *pro* because in his account, if an empty object were a *pro*, it would be subject to both the Disjoint Reference Principle and the GCR, and this leads to contradiction. Take (64b), repeated below as (74), for example. If the empty object were a pronominal, it would have to be coindexed with *John*, the accessible SUBJECT, but at the same time, this coindexation would violate the Disjoint Reference Principle in that the empty object, as a pronominal, would not be free in its governing category. Therefore, according to J. Huang's hypothesis, an empty object cannot be a pronominal due the interaction between the Disjoint Reference Principle and the General Control Rule.

(74) John$_i$ kanjian $Ø_{j/*i}$ le.
see ACC ASP
'John saw (someone).'

J. Huang's analysis emphasizes that Chinese only admits *pro*s in subject position and all empty objects are variables bound by empty topics. Also by proposing the GCR, he has attempted to maintain the validity of the *pro-drop parameter* and the agreement-based theory, although we have shown in this subsection that his analysis fails to make a coherent or correct prediction in some cases.

2.2.2 A pragmatics-oriented account by Y. Huang

In contrast to the syntax-driven account of empty pronouns, Y. Huang (1995) proposes a pragmatics-oriented analysis to resolve empty pronouns in Chinese. Based on the observation of a wide range of genetically unrelated languages, Y. Huang claims that 'the conditions that license and identify null subjects and null objects remain to be fully isolated, and a single or a few syntactic parameters may never be adequate in accounting for them' (Y. Huang 1995: 1082). He first classifies empty subjects into three types

according to referentiality and argumenthood: (i) referential argumental empty subjects, as in (75a); (ii) non-referential argumental or quasi-argumental empty subjects, as in (75b); and (iii) non-referential non-argumental, or simply expletive, empty subjects, as in (75c).

(75) a. John$_i$ shuo [Ø$_i$ cong-mei jian guo xiong].
say NOM ever-not see ASP bear
'John$_i$ said that (he$_i$) has never seen any bear.'
b. Ø xia da xue le!
NOM fall big snow ASP
'(It) is snowing heavily!'
c. Ø sihu [bu xie zuoye ye mei guanxi].
NOM seem not write homework also not matter
'(It) seems that not writing any homework is also okay.'

Languages can be accordingly classified as those which permit all three types of empty subjects, e.g., Chinese and Rumanian; those which disallow all, e.g., English and French; and those which allow a partial use of empty subjects, e.g., German and Malagasy (Y. Huang 1995: 1084). Then he points out that the richness or meagerness of verbal inflections in a language is not a key factor to the typology of languages allowing or disallowing *pro*-drop. Arguing against the agreement-based theory such as J. Huang (1984), Y. Huang uses the Kuki-Chin-Lushai[22] languages as a counterexample. The Kuki-Chin-Lushai languages possess both verb-subject and verb-indirect-object agreement, and do not have any verb-direct-object agreement. However, they nevertheless permit direct objects to be empty, which is an unexpected result for any agreement-based theory. Moreover, an agreement-based theory would have difficulty in explaining the occurrence of an expletive empty subject, such as the one in (75c).

Apart from disagreeing with the agreement-based approach to tackling empty pronouns, Y. Huang (1995) also challenges the following four syntactic analyses of the licensing and identifying of empty pronouns: Rizzi's (1986) *pro* module; J. Huang's

[22] The Kuki-Chin-Lushai languages are under the Kuki-Chin branch of Tibeto-Burman and are spoken in the border areas between India and Burma.

(1984, 1989) Generalized Control Rule; Jaeggli and Safir's (1989) morphological uniformity hypothesis; and Platzack's (1994) minimalist account. Rizzi's *pro* module relies heavily on Agr, and therefore it cannot be extended to the Chinese-type of *pro*-drop languages. J. Huang's GCR not only overlooks some empty subjects which are not controlled when they have a control domain (for instance, see the second reading indexed as *j* in (73a)), but it also fails to license the occurrence of a non-referential empty subject seen in (75b) and (75c). Jaeggli and Safir's morphological uniformity hypothesis assumes that if empty subjects are permitted in a language, the paradigms in the language in question must be morphologically uniform, either uniformly underived, such as in Chinese, or uniformly derived, as in Italian. As Y. Huang argues, in addition to the intuitive question as to why morphological uniformity would be a crucial factor in licensing empty subjects, Jaeggli and Safir's generalization does not hold because there are still languages like Old French, which does not have a morphologically uniform inflectional paradigm but nevertheless allows referential empty subjects in V2 configurations. Lastly, as cited in Y. Huang (1995), Platzack's minimalist account distinguishes between morphologically and syntactically strong/weak $AgrS^0$ (head of the agreement phrase of the subject) and proposes that an empty subject in the Specifier of $AgrS^0$ is allowed only if the specifier feature of $AgrS^0$, known as the N feature, is syntactically weak, and the checking may take place in covert syntax. It follows that there may be a DP in [Spec, $AgrS^0$] at the interface level LF, allowing [Spec, $AgrS^0$] to be empty at PF and to be realized by a *pro* later. However, if the specifier feature of $AgrS^0$ is strong, it will be checked and then deleted before the interface PF, resulting in a DP in [Spec, $AgrS^0$] in overt syntax before Spell-out and preventing the occurrence of an empty subject. Y. Huang argues that even though this minimalist view succeeds in licensing empty subjects, it seems unfeasible for this line of analysis to be able to identify the content of empty subjects in languages like Chinese. The total lack of any independent evidence for the concept of morphologically and syntactically strong/weak specifier feature of $AgrS^0$, on which Platzack's account heavily relies, also weakens this hypothesis. (For more detailed discussions of the above four syntactic views on empty subjects, please refer to the authors' original work).

Y. Huang goes on to criticize two generative analyses of empty objects: one is the aforementioned EC-as-variable hypothesis proposed by J. Huang, and the other is the *pro* analysis, which, as the name suggests, treats an empty object as *pro*, a genuine pronominal. J. Huang's EC-as-variable analysis assumes that an EC in object position is not a *pro* but a variable locally A'-bound by an empty topic operator in [Spec, CP]. Coindexation between the empty object and the hypothesized topic operator is made possible by movement. Y. Huang challenges J. Huang's EC-as-variable proposal with the following three questions. Firstly, he claims that J. Huang's empty-topic hypothesis is not coherent because topic, overt or not, is a notion at discourse level rather than at sentence level. Secondly, if the EC-as-variable analysis is correct, it will also imply that an empty topic can occur in the absence of its chain-initial topic, which, in turn, has the undesirable impact of 'forcing one to analyze every sentence as having an empty topic, and, more absurdly, even to analyze every sentence as containing an indefinite number of empty topics' (Y. Huang 1995: 1110). The reasoning behind J. Huang's EC-as-variable analysis is that, given the traditional classification of non-overt nominal categories, this empty topic cannot be PRO because it is governed and Case-marked; it cannot be an NP-trace because it has no antecedent; it cannot be *pro* because it is not locally identified. So the last alternative, a variable, is assumed to be the answer to the status of the empty topic. The next question arises as how this empty topic, as a variable, is locally A'-bound. Y. Huang supposes that an empty topic is locally A'-bound by another topic, which could in turn be A'-bound by another topic, and so on and so forth, leading to a highly unlikely situation where every sentence contains an indefinite number of empty topics.

According to Y. Huang (1995), the occurrence of a genuine object *pro* has been observed, e.g., in French by Authier (1989), as in (76); in Italian by Chierchia (1989), as in (77); and in Brazilian Portuguese by Chao (1992), as in (78). The object *pro*'s existence is confirmed by, for instance, its ability of controlling in (76) when the empty object acts as the controller of the subject of *commettre* 'commit'. Object-verb agreement is lacking in these three languages, which undermines any agreement-based analysis of empty pronouns. Sentences like (78), in which the empty object is indeed A-bound either by the matrix subject or object, particularly pose a serious problem for J. Huang's

EC-as-variable approach.

(76) (Y. Huang 1995, (28b))
L'ambition amène Ø à commettre des erreurs.
the.ambition leads ACC to commit the mistakes
'Ambition leads (one) to make mistakes.'

(77) (Y. Huang 1995, (30))
L'atteggiamento di Giolitti in quell'occasione lasciò Ø davvero perplessi.
the.attitude of on that occasion let ACC really perplexed
'Giolitti's attitude on that occasion left (one/someone) really perplexed.'

(78) (Y. Huang 1995, (32c))
Pedro$_i$ disse pra Maria$_j$ que a polícia veio interrogar Ø$_{i/j}$.
told that the police came interrogate ACC
'Pedro$_i$ told Maria$_j$ that the police came to interrogate (him$_i$/her$_j$).'

Y. Huang also lists Farrell's (1990) identification condition as one of the representative analyses which treats empty objects as genuine pronominals, *pro*s. Farrell's analysis assumes that 'in the absence of identifying Phi-features on its licensing head, *pro* is intrinsically specified as [+3 person]' (Y. Huang 1995: 1109). Y. Huang disagrees with this view by arguing that Farrell's analysis cannot be extended, at least, to Chinese because the embedded empty object in (79) below could also be first person or second person.

(79) (Adapted from Y. Huang 1995, (34))
John$_i$ yiwei [laoshi you yao chufa Ø$_{i/j}$ le].
think teacher again will punish ACC ASP
'John$_i$ thought that the teacher was going to punish (him$_i$ or someone else$_j$, including you or me) again.'

In view of the inconsistent nature of empty objects, which seem to be variables in some sentences and genuine pronominals in others, Y. Huang concludes that empty objects are realized by syntactically heterogeneous gaps and not by specific empty syntactic categories. Moreover, he argues that the coindexation between an empty pronoun and its antecedent is pragmatically motivated. Empty objects may have heterogeneous syntactic

properties, but they do form a pragmatic class obeying a certain set of rules. Following this reasoning and the above observations contradicting various syntactically-derived accounts of empty pronouns, Y. Huang resorts to a pragmatics-oriented analysis, a neo-Gricean pragmatic approach.

The pragmatic approach advocated by Y. Huang hinges on the familiar Gricean conversational implicatures and the following neo-Gricean pragmatic principles of Levinson (1987, 1991), listed in (80).

(80) (Y. Huang 1995, (38))

a. *The Q(uantity)-principle*:
   Speaker's maxim: do not say less than is required (given *I* below).
   Recipient's corollary: take it that the speaker made the strongest statement he was in a position to make.

b. *The I(nformativeness)-principle*:
   Speaker's maxim: do not say more than is required (given *Q* above).
   Recipient's corollary: amplify the information content of the speaker's utterance, by finding the most specific interpretation, up to what you judge to be the speaker's intended point.

c. *The M(anner)-principle*:
   Speaker's maxim: do not use a marked expression without reason.
   Recipient's corollary: if the speaker used a marked expression, he did not mean the same as he would have, had he used an unmarked alternative.

Y. Huang opts for the means of pragmatic inference to determine possible interpretations of an empty subject or object, and he claims that the principles we see above should not run counter to a more general picture of the world, e.g., background knowledge, contextual information, or semantic entailment. This pragmatic approach is also constrained by a hierarchy of information saliency (e.g., topic > subject > object …), and a disjoint reference presumption (DRP) by Farmer and Harnish (1987), which states that 'the coarguments of a predicate are intended to be disjoint, unless marked otherwise' (Y. Huang 1995: 1119). The I-principle is particularly active in that unless there is a salient discourse topic (to be the antecedent), the I-principle always assigns a local antecedent to be coindexed with an empty subject or object, as long as such coindexation does not

conflict with the DRP and other general background assumptions. The search for a suitable antecedent could move across clauses from local domains to higher clauses, and an arbitrary reference will be assigned if no topic referent or intrasentential antecedent is found. On the other hand, the M-principle, Y. Huang adds, accords well with the widely attested observation that if a language permits overt/empty pronoun alternations, an overt pronoun is typically dropped unless it is required for contrastive, emphatic or focal reasons. The major advantage of a pragmatic account comes from its mechanism which allows for flexibility as well as defeasibility. While favoring a pragmatics-driven analysis, Y. Huang also acknowledges the role syntax plays in a linguistic theory explaining empty pronouns. The independence of syntax is presumed, Y. Huang believes, in a pragmatic approach, and the calculation of pragmatic inference is built upon a level of independent syntactic structure and semantic representation (Y. Huang 1995: 1112).

The employment of the pragmatic principles seen in (80) indeed benefits from not having to determine an absolute syntactic category for empty pronouns, e.g., deciding whether an empty object is a variable or a *pro*, and the interpretation induced this way also avoids being restricted to inflexible syntactic constraints, such as J. Huang's Generalized Control Rule, or to other strict syntactic constraints which limit the referent of an empty pronoun to be third person only. The only syntax-relevant requirement in Y. Huang's pragmatic proposal is the disjoint reference presumption (DRP), which favors the coarguments of a predicate to be disjoint, unless marked otherwise. However, without resorting to the help of syntax, Y. Huang's analysis, as flexible as it is, seems unable to explain why the following sentences in (81) and (82) are ill-formed, given that the salient discourse topic, *John*, is the antecedent in both cases. Y. Huang's pragmatic analysis would predict the discourse topic to be selected as the antecedent, but is unable to detect that both sentences are not even well-formed to start with.

(81) [Someone mentioned that it was John$_i$'s birthday yesterday.]

*Wo	song	le	yi-fen	liwu	gei	Ø$_i$.
I	give	ASP	one-CL	present	to	ACC

'I gave a present to (him$_i$).'

(82) [Someone asked if anyone had seen John$_i$ lately.]
*Wo gang kandao Ø$_i$ zai gong-yuan paobu.
I just.now see ACC in park jog
'I saw (him$_i$) jogging in the park just now.'

There are two syntactic constraints at work here, both proposed by C. Li (1985). (81) is bad because an empty pronoun can never occur as the object of a coverb,[23] i.e., *gei* 'to' in this case. (82) is equally unacceptable because an empty pronoun can never occur in the pivotal position, in which the object of the first verb, *kandao* 'see', simultaneously acts as the subject of the second verb, *paobu* 'jog'. Empty pronouns are still further conditioned by other syntactic constraints, the exact details of which along with the two aforementioned constraints will be the focus of the next chapter. At present, the examples in (81) and (82) have illustrated that a purely pragmatic account of empty pronouns is not rigorous enough to rule out various ill-formed sentences. The I-principle and the M-principle, being heuristic in nature, allow for flexibility and defeasibility, but at the same time, they might not be precise enough in determining the reference of an empty pronoun. Take (83) for instance, a variant of J. Huang's (1984) example (38). (83a) and (83b) are the same utterance and hypothetically, each utterance contains one empty pronoun embedded in the relative clause either as the subject or the object. However, only the reading in (83a) is available, whereas the reading in (83b) is impossible to get.

(83) a. Li Xiaojie$_i$ zongyu zhaodao yige [Ø$_i$ xinzhong xihuan __$_j$ de] nanren$_j$.
Miss finally find one NOM in.heart like ACC RLTV man
'Miss Li$_i$ finally found a man whom (she$_i$) likes in her heart.'
b. *Li Xiaojie$_i$ zongyu zhaodao yige [__$_j$ xinzhong xihuan Ø$_i$ de] nanren$_j$.
Miss finally find one NOM in.heart like ACC RLTV man
'Miss Li$_i$ finally found a man who likes (her$_i$) in his heart.'

Neither the I-principle nor the M-principle would be able to explain the ill-formedness of (83b) or predict the unavailable reading. They could blindly identify

[23] The term 'coverb' refers to a class of words that appear to function as prepositions in Mandarin Chinese. They behave partly like verbs and partly like prepositions, so the name 'coverb' is used to describe its dual status and to avoid labeling it as either a verb or a preposition (Li and Thompson 1989: 360). There will be more discussion of coverbs later in 3.2.2.

Miss Li as the antecedent of the empty object. Moreover, suppose (83a) is uttered right after a comment that 'Miss Li's sister has been single for many years, and Miss Li has been trying to be a matchmaker for her sister'. In this context, *Miss Li's sister* is clearly the topic, and the I-principle would predict incorrectly that the embedded empty subject in (83a) is anteceded by *Miss Li's sister*, the topic, which is in fact an impossible interpretation. There does not appear to be anything that would prevent this unavailable reading if we closely follow Y. Huang's pragmatic account. Therefore, even though Y. Huang's pragmatics-oriented theory escapes the problems surrounding J. Huang's syntactic account (e.g., the inviolable GCR being too strict and yielding incorrect predictions), it is still not a desirable analysis, as it suffers from the inability of determining sentence well-formedness and fails to give a precise, clear-cut interpretation without the help of syntax.

2.2.3 A syntactic-pragmatic account by Pan and Hu

We have reviewed one syntactic account and one pragmatic account of empty pronouns in Chinese so far. The former suffers from the Generalized Control Rule being too strict, due to the inviolability of syntax, to yield the full range of possible interpretations of an empty pronoun. The latter is undermined by the lack of rigid syntactic regulations of the distribution of empty pronouns, and that the heuristic I-principle and M-principle do not always predict correct interpretations. In view of this, it seems not only wise but also necessary to propose an analysis which takes both sides into consideration, at least in the case of non-inflectional *pro*-drop languages like Chinese. The syntactic integrity must be maintained to the extent that it also allows flexibility in permitting a full range of empty pronoun interpretations available in the discourse. Pan and Hu's (2000) syntactic-pragmatic account is such an attempt. In contrast to J. Huang's syntactic proposal based on the GCR and Y. Huang's neo-Gricean account, Pan and Hu (2000) tackle the problem of empty pronouns in Chinese with the Relativized Locality Condition (RLC), defined in (84), which appears to be a loosely constrained version of J. Huang's GCR, and the Compatibility Condition (CC), listed in (85), which also has a hint of Y.

Huang's pragmatic checking of whether a sentence interpretation predicted by the pragmatic rules in (80) runs counter to any syntactic condition and background knowledge. In addition, Pan and Hu add the Topic Interpretation Rule (TIR), shown in (86), as an independent condition in their analysis.

(84) Relativized Locality Condition (RLC) (Pan and Hu 2000, (23)):
Coindex the null element with the closest compatible NP.
(85) Compatibility Condition (CC) (Pan and Hu 2000, (24)):
A and B are compatible if and only if
a. A and B have the same animacy features; and
b. the coindexation of A and B does not violate any syntactic conditions, lexical/semantic restrictions, or pragmatic knowledge.
(86) Topic Interpretation Rule (TIR) (Pan and Hu 2000, (25)):
Coindex the topic NP with an adjacent open predicate.[24]

To show how the above three conditions are integrated to account for *pro*-drop sentences and also how their proposal differs from J. Huang's, Pan and Hu adopt J. Huang's (1984) sentences such as the following. (The square brackets indicate the boundary of a modifying clause, as in (87), or the boundary of a complement clause, as in (88)).

(87) (J. Huang 1984, (78))
a. Zhangsan$_i$, [Ø$_i$ changge de] shengyin hen haoting.
NOM sing RLTV voice very good.to.hear
'(Speaking of) Zhangsan, the voice with which (he) sings is good.'
b. *Zhangsan$_i$, wo hen xihuan [Ø$_i$ changge de] shengyin.
I very like NOM sing RLTV voice
'(Speaking of) Zhangsan, I like the voice with which (he) sings a lot.'

As predicted, the empty subject in (87a) is indeed coindexed with the closest compatible

[24] Pan and Hu propose this Topic Interpretation Rule (TIR) based on the Principle of Full Interpretation, which, we suppose, refers to the Chomskyan principle that requires every element of PF and LF to receive an appropriate interpretation in order to be licensed. Pan and Hu also do not give any concrete definition of what an 'open predicate' refers to. According to their given examples, an open predicate could consist of (i) a predicate and an unfulfilled argument, e.g., subject or object; or (ii) a predicate and an argument coreferential with the discourse topic. In the latter case, the argument coreferential with the topic is seen as a resumptive pronoun in Pan and Hu's view.

NP, that is, the topic referent *Zhangsan*. Also obeying the TIR, the topic referent is fully interpreted by being coindexed with the empty argument in an adjacent open predicate, i.e., the empty subject in the relative clause. The conditions in (84) to (86) are all satisfied, and a correct interpretation is predicted. On the other hand, (87b) is bad because firstly, according to Pan and Hu's RLC and the Compatibility Condition, the matrix subject *wo* 'I', not the topic *Zhangsan*, would be the closest compatible NP to be coindexed with the empty object; and secondly, the requirement of the TIR is not met here, as the topic NP is not adjacent to the open predicate containing the missing argument. So far, so good. Let us look at Pan and Hu's next example, also taken from J. Huang (1984).

(88) (J. Huang 1984, (73))

a. Zhangsan$_i$, ta$_i$ shuo [Ø$_i$ mei kanjian Lisi].
he say NOM not see
'(Speaking of) Zhangsan$_i$, he$_i$ said that (he$_i$) didn't see Lisi.'

b. ?Zhangsan$_i$, ta$_i$ shuo [Lisi mei kanjian Ø$_i$].
he say not see ACC
'(Speaking of) Zhangsan$_i$, he$_i$ said that Lisi didn't see (him$_i$).'

In Pan and Hu's view, the grammatical (88a) follows their prediction in that the embedded empty subject is indeed coindexed with the closest compatible NP, the matrix subject *ta* 'he'. When the topic NP *Zhangsan* is coindexed with the matrix subject, this matrix subject pronoun has then become a resumptive pronoun, and then the whole clause (beginning with *ta* 'he') containing this resumptive pronoun in turn becomes an open predicate. The TIR is thus respected, and the topic is fully interpreted. Pan and Hu stress the necessity of the TIR: if it were not for the application of the TIR, the topic NP in (88a) would be left uninterpreted, leading to ungrammaticality. Now let us turn to the more complicated (88b), which is originally J. Huang's example and has been analyzed as ill-formed by him. The question mark in front of (88b) is Pan and Hu's judgment; it indicates the marginality of the interpretation, meaning that it is not a preferred reading, but by no means impossible, and is still syntactically well-formed. In Pan and Hu's analysis, the empty object is analyzed as a null pronominal and should be coindexed with the closest compatible NP, *ta* 'he', according to the RLC and the Compatibility Condition,

the interaction of which rules out the coargument *Lisi* to be syntactically compatible. The topic NP receives its full interpretation in the same fashion as the one in (88a). The reason resulting in the lower acceptability of (88b) comes from another hypothesis of Pan and Hu's: the Generalized Principle of Obviation (GPO), stated in (89).

(89) (Pan and Hu 2000, (40))
Generalized Principle of Obviation (GPO):
When an empty pronoun and an overt pronoun occupy the same position,
a. an overt pronoun tends to be obviative from the antecedent of the empty pronoun;
b. an empty pronoun tends to be obviative from the antecedent of the overt pronoun.

The Generalized Principle of Obviation (GPO) predicts the embedded empty object in (88b) to be obviative from the matrix subject because the matrix subject would have been the preferred antecedent for the overt pronoun *ta* 'him' sitting in the same embedded object position. Corresponding to (88b), the example in (90) illustrates the replacement of an overt pronoun for the empty pronoun in the embedded object position, and under this circumstance, the coreference between the matrix subject and embedded object is the salient reading, prompting Pan and Hu's expectation of the dispreferred coreference between the matrix subject and the embedded empty object seen in (88b).

(90) Zhangsan$_i$, ta$_i$ shuo [Lisi mei kanjian $\underline{\underline{\text{ta}_i}}$].
he say not see him
'(Speaking of) Zhangsan$_i$, he$_i$ said that Lisi didn't see him$_i$.'

Pan and Hu explain that the overt pronoun in the embedded object position in (90), underlined, prefers to be coreferential with the matrix subject, which is the closest compatible NP, and this also accords well with their Relativized Locality Condition in (84). Such coindexation leads to an empty pronoun occurring in the same embedded object position being obviative from the phrase its overt counterpart is proximate to. Yet, Pan and Hu consider this proximateness to, or obviation from, a certain antecedent to be just a case of a preferred or dispreferred reading at best. It does not rule out the dispreferred interpretation as ungrammatical. This is the reason that (88b) is simply

marked with a question mark by Pan and Hu, instead of an asterisk, as J. Huang would have done (for the reason that J. Huang analyzes an empty object as a variable, and it is not allowed to be coindexed with an argument, such as the matrix subject). Although Pan and Hu predict coindexation between the matrix subject and the embedded object pronoun in (88b) to be dispreferred, they do not consider (88b) ungrammatical. For Pan and Hu, such coindexation is still possible in some contexts, since the Generalized Principle of Obviation is not an absolute syntactic constraint in essence, and it simply offers some sort of 'tendency' at best, which could still be overridden. However, we do question the validity of their GPO, which does not provide an explanation why the references of the overt/empty pronouns should obviate from each other. Although the contrast between the marginally acceptable reading in (88b) and the preferred reading in (90) appears to follow Pan and Hu's theory, such obviation in reference owing to whether a pronoun is phonetically realized or not, does not seem fully supported. For instance, the sentence in (91) below allows the overt/empty pronoun alternation in the embedded subject position, and there is *no* obviation in reference between the alternative usages. There is no preferred or dispreferred interpretation; both readings are equally acceptable.

(91) Zhangsan$_i$, ta$_i$ shuo [ta$_i$/Ø$_{i/*j}$ hen gaoxing [ta$_i$/Ø$_{i/*j}$ mingtian jiuyao
he say he very happy he tomorrow be.going.to
shangxue le]].
go.to.school ASP
'(Speaking of) Zhangsan$_i$, he$_i$ said that he$_i$ was very happy that he$_i$ was going to school tomorrow.'

So far we have briefly outlined Pan and Hu's analysis incorporating both syntax and pragmatics, and like J. Huang's and Y. Huang's analyses, Pan and Hu's syntactic-pragmatic account is not without problems. Pan and Hu's approach seems to attempt to restrict empty pronouns in Chinese pragmatically to allow for contextual flexibility with the minimum of syntactic constraints. This strategy of incorporating both syntax and pragmatics is on the right track, given the inadequacy previously discussed in J. Huang's syntactic approach and Y. Huang's pragmatic method. However, Pan and Hu's

hypothesis suffers the following problems. Let us begin with the Topic Interpretation Rule (TIR), as listed in (86). The TIR plays a crucial part in deciding the grammaticality of a sentence containing a topic NP in their analysis, and the TIR is also regarded as not only a necessary, but also an independent condition. However, we question the necessity, or even the validity, of the TIR. It is not controversial to say that a topic indeed needs to be identified with, or anaphorically linked to, another syntactic argument in the comment clause, but is the TIR really a necessary requirement? If a topic needed to be fully interpreted in the fashion the TIR describes, then the following well-formed English sentences of topic-comment structure in (92) would both be ruled out.

(92) a. Speaking of John, I like the way [he tells stories].
b. Speaking of John, I think [the way he tells stories is fascinating].

John is the topic NP in (92a) and (92b), and is coreferential with the underlined pronoun, *he*, in the comment clause. The square brackets indicate the clause containing the NP with which the topic is identified. Following Pan and Hu's reasoning, both bracketed clauses may be considered open predicates if we treat 'he' as a resumptive pronoun, but it is clear that neither of the bracketed clauses is adjacent to the topic *John*. The TIR requires a topic NP to be coindexed with an adjacent open predicate – we can either opt to treat (92a) and (92b) as some special cases which remain grammatical despite violating the TIR, or argue that the TIR is in fact a superfluous, if not incorrect, requirement.

The second problem concerning Pan and Hu's analysis comes from the circularity of their Generalized Principle of Obviation (GPO), shown in (89). Such a principle is used to account for the ungrammaticality of (93b). In (93a) below, suppose that an overt pronoun *ta* 'her' occurs in object position inside the relative clause, as the enclosed brackets indicate, and this overt pronoun, underlined, will be coindexed with the closest compatible NP, *Miss Li*. Everything else being the same, the GPO would predict that an empty pronoun occurring in the very same position, exemplified in (93b), should be obviative from this antecedent, *Miss Li*.

(93) (J. Huang 1984, (38))

a. Li Xiaojie$_i$ zongyu zhaodao yige [__$_j$ xinzhong xihuan ta$_i$ de] nanren$_j$.
Miss finally find one NOM in.heart like her RLTV man
'Miss Li$_i$ finally found a man who likes her$_i$ in his heart.'

b. *Li Xiaojie$_i$ zongyu zhaodao yige [__$_j$ xinzhong xihuan Ø$_i$ de] nanren$_j$.
Miss finally find one NOM in.heart like ACC RLTV man
'Miss Li$_i$ finally found a man who likes (her$_i$) in his heart.'

The GPO correctly explains the contrast between (93a) and (93b). However, let us turn to (94a) and (94b) below, which are variants of (93a) and (93b) and are to be analyzed in the same manner.

(94) a. Li Xiaojie$_i$ zongyu zhaodao yige [ta$_i$ xinzhong xihuan __$_j$ de] nanren$_j$.
Miss finally find one she in.heart like ACC RLTV man
'Miss Li$_i$ finally found a man whom she$_i$ likes in her heart.'

b. Li Xiaojie$_i$ zongyu zhaodao yige [Ø$_i$ xinzhong xihuan __$_j$ de] nanren$_j$.
Miss finally find one NOM in.heart like ACC RLTV man
'Miss Li$_i$ finally found a man whom (she$_i$) likes in her heart.'

Because the overt subject position of the relative clause in (94a) is coindexed with the closest compatible NP, *Miss Li*, the GPO would predict an empty subject pronoun occurring in the same position, as (94b) shows, to be obviative from this antecedent *Miss Li*, rendering the coindexation between *Miss Li* and the empty subject unavailable, or dispreferred. But in fact, (94b) is perfectly acceptable, and the reading is readily available. Additionally, if the checking procedure of the GPO reverses, namely, if we examine (94b) first, and then proceed to (94a), the prediction yielded for (94a) by the GPO remains incorrect for the same reason. That is, if we examine (94b) first and conclude that the coreference between *Miss Li* and the empty subject is the salient reading, the GPO would then predict the coreference between *Miss Li* and the overt subject pronoun inside the relative clause of (94a) to be a dispreferred reading. This prediction is grossly incorrect, and it shows that the GPO itself is theoretically circular. Instead of solely relying on a few chosen examples that appear to fit the condition, Pan and Hu's GPO must be defined independently in order to avoid the problem of circularity, but they have not done so.

Therefore, in addition to being theoretically circular, Pan and Hu's GPO also fails to yield correct interpretations. In view of this, we believe that Chinese does not exhibit salient obviation in reference between overt/empty pronouns the way Pan and Hu have portrayed.

The above discussion shows the problems with Pan and Hu's syntactic-pragmatic analysis. The proposed GPO is not a sound generalization because there are ample examples, at least in Chinese, where the alternation of overt/empty pronouns in a given sentence makes no significant difference at all, and the resulting sentences, whether both well-formed or both ill-formed, convey pretty much the same meaning. These sentences constitute a major setback for the GPO and, of course, for Pan and Hu's account, too. Even though the GPO, which is not syntactic in nature, only claims to offer a tendency and is not an absolute constraint itself, the seemingly advantageous flexibility aimed for and favored by Pan and Hu's analysis appears to become one of its downsides as well.

2.2.4 The need for an integrating account

After reviewing J. Huang's syntax-oriented analysis, Y. Huang's pragmatics-based analysis, and Pan and Hu's attempt to accommodate the problems of empty pronouns by combining both syntax and pragmatics, we are now in a position to decide what would be the best way to resolve empty pronouns in Chinese. It is clear by now that it would be inadequate to rely on syntax or pragmatics exclusively, and this kind of single-tracked argument is unable to cover the full range of interpretations of empty pronouns, which may be subject to various syntactic constraints and contextual factors. Failure to take both syntax and pragmatics into consideration results in incompleteness. Therefore, in this thesis, we argue for an integrating account which caters for both syntax and pragmatics. The distribution of Chinese empty pronouns is by no means free, cannot be oversimplified, and must be constrained by syntactic rules which are not to be overridden by any pragmatic condition and contextual reasons. This restriction by syntax is self-evident, due to the clear impossibility of dropping any random Chinese pronoun as one would like to, even if given a carefully designed context to trigger the intended

interpretation. If the syntactic regulation were not at work in Chinese, then we would expect that a pronoun could be omitted whenever a suitable context is given. However, this expectation is not met, as we can see from the example in (95) below, in which the contextual information strongly suggests the omitted object pronoun in the bracketed relative clause to refer to the topic referent, *Jane*, and yet, the use of the empty pronoun in the sentence is unacceptable. An overt object pronoun must be used instead in order to maintain grammaticality.

(95) [Jane$_i$ is the least attractive girl in the world with a terrible personality.]
Ta$_i$ jingran zhaodao le yige [__$_j$ zhen-xin xihuan ta$_i$/*Ø$_i$ de]
she surprisingly find ASP one NOM genuine-heart like her RLTV
da-shuai-ge$_j$.
big-handsome-guy
'Surprisingly, she$_i$ found a handsome guy who loves her$_i$/*(her$_i$) genuinely.'

Conversely, the same thing can be said about the necessity of including pragmatic factors. If there were no pragmatics at work with regard to interpreting Chinese empty pronouns, syntax alone would not be able to explain, e.g., why the embedded empty object in (96) can only refer to *Mary's watch* but not the matrix subject, since the latter coindexation does not violate any syntactic rule.

(96) [Mary's watch$_i$ got stolen …]
Dan John$_j$ shuo [Mary benlai jiu bu xiang-yao Ø$_{i/*j}$].
but say originally just not want ACC
'But John said that Mary never wanted (it$_i$) anyway.'

These examples illustrate that empty pronouns in Chinese are indeed constrained both syntactically and pragmatically. Absolute syntactic rules which are not to be overridden by context, and heuristic pragmatic tendencies which rely very much on the discourse will both be addressed in this study in order to resolve empty pronouns in Chinese. Within the LFG framework assumed in this study, syntactic constraints will be discussed at the level of *f*-structure; pragmatic restrictions, on the other hand, will be reflected at the level of *i*-structure, which houses contextual information and discourse

properties. What is missing at *c*-structure or what is not present at *f*-structure could well have a salient role and prominent function at *i*-structure. This mechanism of different linguistic tiers integrates syntax and pragmatics, and serves as the ideal platform on which we build an integrating analysis of zero anaphora in Chinese.

2.3 Topic-comment structure: a syntactic or semantic device?

As previously discussed, subject-prominence is characterized by the grammatical relation between subject and predicate in a subject-predicate construction, as in (97), whereas topic-prominence is characterized by the relation of 'aboutness' holding between topic and comment in a topic-comment structure, as (98) shows.

(97) [Dogs]$_{\text{Subject}}$ [bark]$_{\text{Predicate}}$.
(98) As for [dogs]$_{\text{Topic}}$, [they bark]$_{\text{Comment}}$.

If the assumption that Chinese is a topic-prominent language is on the right track and the topic-comment construction is the basic sentence type in Chinese, what concerns us next is (i) whether the employment of topic-comment sentences is a syntactic or a semantic device, as many instances of empty pronouns can be found in the so-called topic-comment sentences; and (ii) whether the possibility that a pronoun can be empty has something to do with this pronoun being a topic or not. We will first compare a pro-syntax point of view by Shi (2000) and a pro-semantics analysis by Pan and Hu (2002), before concluding that an LFG formation of topic-comment construction can take both syntax-oriented topics and semantically-licensed ones into account.

2.3.1 A pro-syntax view by Shi (2000)

That Chinese is a topic-prominent language is widely accepted by the majority of Chinese linguists; most of them see topic and subject as two separate grammatical notions, and admit the coexistence of the topic-comment structure and the subject-predicate structure in Chinese, e.g., Li and Thompson (1976) and Tsao (1979, 1990). Despite having this

assumption in common, Chinese grammarians do not seem to see eye to eye when it comes to the definition of subject and topic, and the usage of these two terminologies is not always consistent. Particularly unsatisfied with the most common way of defining these two crucial notions only by 'listing the distribution and properties of topic and subject, and then to use these properties as criteria for determining whether a given pre-verbal NP is a topic or a subject' (Shi 2000: 383), Shi (2000) proposes a clearer definition of topic, shown in (99) below.

(99) (Shi 2000, (5))

> A TOPIC is an unmarked NP (or its equivalent) that precedes a clause and is related to a position inside the clause; a topic represents any entity that has been mentioned in the previous discourse, and is being discussed again in the current sentence; namely, topic is what the current sentence is set up to add new information to. The clause related to the topic in such a way is the comment.

Shi's definition of topic aims to be structurally precise to the extent that it also allows this topic to function pragmatically in discourse. He believes that the topic-comment structure in Chinese is a grammatical means to achieve certain discourse functions, and the topic-comment construction is derived from the basic sentence structure, e.g. subject and predicate, via syntactic operations. Shi rejects the common method of listing properties of a topic (e.g., topic always occupies the sentence-initial position, as discussed in 2.1), instead of defining it clearly, because the listed properties are often quite general and could incorrectly include non-topics. For Shi's definition of topic to be valid, he must first reject the existence of the so-called Chinese-style topics, which are characteristically unsubcategorized by the verb of the comment clause. Chinese-style topics, also known as dangling topics or non-gapping topics, are generally acknowledged by Chinese grammarians, as long as an aboutness relation holds between the topic and the comment clause. The occurrence of Chinese-style topics not only seriously undermines Shi's proposal, but its existence also violates LFG's assumptions about topic. To integrate this special kind of topic-comment construction into the LFG architecture, we will sketch an LFG resolution in 2.3.3 and establish the appropriate

correspondence between a dangling topic and the function it links to at *f*-structure and *i*-structure. In defending his definition of topic and his claim that the topic-comment construction in Chinese, as a grammatical device, is syntactically derived from the basic 'subject-predicate' sentence structure, Shi argues that the seemingly non-gapping topics, which, by definition, are not related to any position in the comment clause, are in fact *not* topics at all and, therefore, they do not pose a problem for his analysis.

Shi distinguishes six types of dangling topics gathered from the literature of Chinese linguistics. He aims to eliminate them by reanalyzing the sentences syntactically, and proposes that the problematic dangling topics, which actually belong to different kinds of syntactic categories (e.g., subject or adverbial) in disguise, can be related to some position inside the comment clause. We will examine these six types one by one. Unlike those who simply assume these topics as dangling, as a special kind, with no further query, Shi rejects these pre-verbal NPs as topics because they cannot be structurally linked to any gap in the comment clause. No matter how strong the aboutness relation is that holds between the dangling topics and their comment clauses, this semantic sense of aboutness should not override syntax. We agree with Shi in this aspect, as in LFG, topic should always be identified with, or anaphorically linked to, another syntactic function. Shi's first kind of dangling topic looks like the one in (100) and (101).

(100) (Shi 2000, (14))

[Tamen]$_{\text{Topic}}$ [wo kan ni ni kan wo]$_{\text{Comment}}$.
they I look you you look me
'They look at each other.'

(101) (Shi 2000, (15))

[Tamen]$_{\text{Topic}}$ [da-yu chi xiao-yu]$_{\text{Comment}}$.
they big-fish eat small-fish
'They act according to the law of the jungle.'

A traditional analysis would treat the first NP in both sentences, *tamen* 'they', as the sentence topic, and the clauses after the topic NP are the comment clauses, leaving the topic NP, which is unrelated to any gap inside the comment clause, *dangling*. Shi offers a different explanation. He argues that because the bracketed string of words traditionally

analyzed as the comment is unable to alter its internal order, or be replaced by other lexical items, the whole string of words is seen as an idiom chunk, which is not to be further analyzable and, in turn, functions as a predicate. This leads Shi to the speculation that the sentence-initial NP *tamen* 'they' is, in fact, a licensed subject, and not a dangling topic. However, treating the comment clause as an idiom chunk seems to be oversimplifying the matter. For example, the lexical items inside the bracketed comment clause could actually change their word order in an appropriate context, as (102) exemplifies, in contrast to (101). (100) would still be well-formed syntactically and semantically, even if the bracketed string of words *wo kan ni ni kan wo* 'I look at you; you look at me' is changed to *ni kan wo wo kan ni* 'you look at me; I look at you'.

(102) Tamen$_i$ jingran [xiao-yu chi da-yu].
they unexpectedly small-fish eat big-fish
Ø$_i$ ba IBM na-zhong da gongsi da-kua le.
NOM BA that-kind big company hit-collapse ASP
'They$_i$ unexpected fought back as the underdog. (They$_i$) defeated a huge company, IBM.'

In (102), the coreference between *tamen* 'they' and the empty subject of the second clause evidences the topical status of the former, because the subject of 'defeating a huge company, IBM' is left unpronounced under Tsao's (1979) Topic-NP Deletion Rule, which permits a topic NP (and the successive ones in the following clauses) to be empty if it is the same as the initial topic, e.g., *tamen* 'they', in this case. Moreover, another inadequacy of Shi's idiom-chunk analysis, as noted by Pan and Hu (2002), is that sometimes, the comment clause can become so big that it seems unreasonable to treat it as an idiom chunk, e.g., the long string of words in the bracketed comment clause in (103).

(103) [Tamen]$_{Topic}$ [ni taoyan wo wo taoyan ni, ni piping wo
they you dislike me I dislike you you criticize me
wo piping ni]$_{Comment}$.
I criticize you
'They dislike each other and criticize each other.'

We would like to add that even though the long string of words in the comment clause in (103) might look like coordinated predicates, as a repeated idiom chunk, and might not be a counterexample to Shi's analysis, as Pan and Hu have proposed, the fact that different verbs can appear inside Shi's so-called idiom chunk, from *kan* 'look at' in Shi's original example in (100) to *taoyan* 'dislike' or *piping* 'criticize' in (103), already weakens Shi's idiom-chunk hypothesis. The irreplaceability of any lexical item inside the idiom chunk assumed in Shi's analysis is not supported.

The sentence in (104) below demonstrates Shi's second type of dangling topic, in which the *wh*-word, *shei* 'who', is traditionally analyzed as the subject, and the initial NP is the non-gapping topic, unrelated to any position inside the comment clause.

(104) (Shi 2000, (18))
[Tamen]$_{\text{Topic}}$ [shei dou bu lai]$_{\text{Comment}}$.
they who all not come
'(Speaking of) them, none of them are coming.'

Shi argues that when the *wh*-word *shei* 'who' in (104) appears on the left of the emphatic adverbial *dou* 'all', which modifies everything to its left (Lee 1986; Huang 1996), this *wh*-word should be interpreted as equivalent to a universal quantifier, amounting to an interpretation in which *shei* means 'anyone', instead of 'who/which person'. Without the emphatic adverbial *dou* 'all' on its right, the *wh*-word *shei* loses its quantifier interpretation (as 'anyone') and can only mean 'who'. This usage is demonstrated by (105) below, in contrast to the quantifier reading of the *wh*-word shown in (104).

(105) Tamen shei bu lai?
They who not come
'(Speaking of) them, who will not come?'

After establishing the meaning of *shei* as 'anyone' in (104), Shi argues that *shei*'s status is not a subject (but *tamen* 'they' is), and *shei* functions more like an adverbial. According to Shi, subject in Chinese may be characterized as being able to precede a certain modal verbs such as *keyi* 'may', *hui* 'will', and *neng* 'can' (Shi 2000: 390); the opposite order

(modal verb before subject) is unacceptable. (106a) is the original sentence that involves a non-gapping topic, copied from (104), while (106b) is Shi's example revised from (106a). In (106b), the subject status of *tamen* 'they' is indicated by *tamen*'s preceding the underlined modal verb, *keyi* 'may'. Along the same line, because the *wh*-word *shei* in (106b) follows the modal verb, it is not a subject.

(106) a. [Tamen]$_{\text{Topic}}$ [shei dou bu lai]$_{\text{Comment}}$.
they who all not come
'(Speaking of) them, none of them are coming.'
b. Tamen keyi shei dou bu lai.
they may who all not come
'None of them may come.'

Following this reasoning, Shi concludes that the sentence-initial NP *tamen* 'they' in (106a) is in fact the subject, not a (non-gapping) topic. Then he proceeds with the claim that the *wh*-word in (106a) is actually optional, because if we remove *shei* 'who' from (106a), the resulting sentence, exemplified by (107), still means pretty much the same as (106a) does.

(107) (Shi 2000, (21))
Tamen dou bu lai.
they all not come
'None of them are coming.'

Moreover, Shi uses the example in (108) to show that the *wh*-word *shei* can alternate with a universal quantifier such as *quan* 'all', but cannot alternate with a referential NP, such as *John and Bill*. He claims that because *quan* 'all' in (108) functions more like an adverbial than a subject, an alternative replacement in the same position, *shei*, is also likely to be a quantificational adverbial, instead of a referential NP.

(108) (Shi 2000, (22, 23))
Tamen shei/quan/*John he Bill dou bu lai.
they who/all/ and all not come
'None of them are coming/*Neither John nor Bill is coming.'

According to Shi's reanalysis shown above, the sentence in (106a) does not contain any dangling topic and is not problematic, because *tamen* 'they' is the subject and the *wh*-word *shei* is an optional, quantificational adverbial. The sentence structure has been reanalyzed as of subject and predicate. This line of argument is not unproblematic. First of all, it is possible, contra Shi's claim, for the *wh*-word *shei* in (106a), being quantificational, to be replaced by a referential NP. The counterexamples in (109) and (110) below demonstrate such a possibility, as *fu-qi* 'husband and wife' in (109) and *xiong-di* 'brothers' or *jie-mei* 'sisters' in (110a) are both (collective) referential NPs, and both sentences are well-formed. Likewise, *quan-jia-ren* 'the whole family' in (110b) can also be used in the same fashion, in contrast to Shi's claim about the impossibility of replacing the *wh*-word in question with a referential NP.

(109) [Tamen]$_i$ [fu-qi$_i$ dou bu lai].
they husband-wife all not come
'(Speaking of) them$_i$, neither member of the couple$_i$ (consisting of a husband and his wife) is coming.'

(110) a. [Tamen]$_i$ [xiong-di$_i$/jie-mei$_i$ dou bu lai].
they brothers/sisters all not come
'(Speaking of) them$_i$, neither member of the brothers$_i$/sisters$_i$ is coming.'

b. [Tamen]$_i$ [quan-jia-ren$_i$ dou bu lai].
they all-family-people all not come
'(Speaking of) them$_i$, none of the family members$_i$ are coming.'

While we agree with Shi on *not* interpreting the *wh*-word *shei* in (104) as an interrogative pronoun such as 'who/which person', we disagree with him on treating *shei* as a quantifying adverbial that cannot alternate with any referential NP. The unresolved status of *tamen* 'they' as a problematic dangling topic also undermines LFG's assumptions. A different solution under the LFG framework will soon be proposed in 2.3.3.

The third type of dangling topic according to Shi (2000) comes from one of the most cited Chinese-style topic sentences by Chao (1968), exemplified by (111) below, in which the sentence-initial NP, *na-chang huo* 'that fire', is the non-gapping topic, unrelated to any position inside the comment clause. Semantically, the comment clause implies that

the fire might have had caused serious damage or casualties, if the fire brigade had not come quickly.

(111) (Shi 2000, (29))

[Na-chang huo]$_{Topic}$ [xingkui xiaofang-dui lai-de-kuai]$_{Comment}$.
that-CL fire fortunately fire.brigade come-PART-fast
'As for the fire, fortunately the fire brigade came quickly.'

The comment clause is headed by the adverb *xingkui* 'fortunately', and Shi proposes that it is one of those connective adverbs which always work in pairs. To be specific, connective adverbs like *xingkui* 'fortunately' usually pair with *buran* 'otherwise', as in 'fortunately … otherwise …', or with *cai* 'therefore', as in 'fortunately … therefore …', in the form of two successive clauses, the first of which is more important than the second. He explains that the first clause introduced by *xingkui* 'fortunately' is the cause, and the second clause introduced by *buran* 'otherwise', or *cai* 'therefore', is the hypothetical effect or the consequence (Shi 2000: 391-2). He then argues that it is common for the second adverb in a connective adverb pair to be omitted together with the clause it introduces, and hearers are usually able to induce the implied effect caused by the first clause. In the case of (111), the presumed result, if the fire brigade had not come quickly, can be recovered by the hearer, even though the second part of the clause ushered by *buran* 'otherwise' is not expressed. To relate the above explanation to his proposal of eliminating dangling topics, Shi hypothesizes that (111) may be ambiguous between two possible readings, depending on what the unexpressed second clause is which indicates the consequence of the first clause. That is to say, the sentence in (111) may be interpreted as (112a) or (112b). In (112a), the sentence-initial NP, 'that fire', also functions as the subject of the second, bracketed clause. In (112b), the second, bracketed clause has its own subject, and 'that fire' is treated by Shi simply as a temporal adverb.

(112) (Shi 2000, (30))

a. Na-chang huo$_i$, xingkui xiaofang-dui lai-de-kuai, [buran Ø$_i$
that-CL fire fortunately fire.brigade come-PART-fast otherwise NOM
jiu hui shao-si bu-shao ren].
really will burn-die not-few people
'As for that fire, fortunately the fire brigade came quickly; otherwise, (the fire) would have killed many people.'

b. Na-chang huo$_i$, xingkui xiaofang-dui lai-de-kuai, [buran <u>na-ci$_i$</u>
that-CL fire fortunately fire.brigade come-PART-fast otherwise that-time
women dou hui shao-si].
we all will burn-die
'As for that fire, fortunately the fire brigade came quickly; otherwise, we would all have been burned to death at that time.'

In Shi's account, the sentence-initial NP, *na-chang huo* 'that fire', in (112a) is the subject of the bracketed clause and could have killed many people. Alternatively, in (112b), Shi claims that 'that fire' could also function as a temporal adverbial related to the position between the second connective adverb *buran* 'otherwise' and the subject *women* 'we', and therefore 'that fire' is not a problematic non-gapping topic. This view is supported by that the position in question is resumptively filled by the underlined temporal adverbial, *na-ci* 'that time'. In both readings, the sentence-initial NP, 'that fire', is linked to a position inside the clauses headed by *buran* 'otherwise', leading to Shi's conclusion that the sentence-initial NP, 'that fire', in (111) is licensed in the same manner by being associated with a position inside the unexpressed, hypothetical clause headed by the connective adverb *buran* 'otherwise'.

Shi's attempt to explain away the non-topical status of *na-chang huo* 'that fire' in (112) is well elaborated, but it suffers from oversimplification. Firstly, in (112b), the only evidence Shi has in order to analyze the sentence-initial NP, 'that fire', as a temporal adverbial comes from the possibility of resumptively employing a temporal adverbial *na-ci* 'that time' in the very position to which Shi assumes 'that fire' is linked. Thus, 'that fire' is also a temporal adverbial, not a non-gapping topic. However, we find this test unreliable. Note that the use of *na-ci* 'that time' in (112b) is only optional; without this temporal adverbial, the very same position could also be filled by other types of

adverbials such as *hen keneng* 'very likely', or *shi-zhi-ba-jiu* 'eight or nine out of ten (literally), meaning eighty or ninety percent of the chance', to indicate the high probability of the presumed casualties, as the example in (113) shows.

(113) Na-chang huo$_i$, xingkui xiaofang-dui lai-de-kuai, [buran
that-CL fire fortunately fire.brigade come-PART-fast otherwise
hen kenengi/shi-zhi-ba-jiu women dou hui shao-si].
very possible/ten-of-eight-nine we all will burn-die
'As for that fire, fortunately the fire brigade came quickly; otherwise, we all would have been very likely to be burned to death.'

Following Shi's line of reasoning, one could also claim that 'that fire' functions as an adverb of possibility, which is by no means correct. The hypothetical relation between the sentence-initial *na-chang huo* 'that fire' and a position inside the clause ushered by *buran* 'otherwise' cannot be established simply by saying a temporal adverbial also occurs *resumptively* in that position, as this assumption would wrongly imply that in theory, *na-chang huo* 'that fire' could also become various sorts of categories, e.g., as an adverb of possibility, as long as the substitution of the adverbial in (112b) from *na-ci* 'that time' to *hen keneng* 'very likely' (or *shi-zhi-ba-jiu* 'eighty or ninety percent of the chance') works. This is an undesirable result. Therefore, we argue that Shi's treatment of a dangling topic as a temporal adverb is weak, and we will offer a different solution to dangling topics of this kind in 2.3.3. The second problem with Shi's analysis with regard to his third type of non-gapping topic is pointed out by Xu and Liu (1998) (cited in Pan and Hu (2002)). Suppose that Shi's account of a dangling topic like the one in (111), which could be in turn elaborated as (112a) or (112b), is all correct. There are still other ways of elaborating (111) such that the clause introduced by the second connective adverb of the pair, e.g., *buran* 'otherwise', does not contain any syntactic gap or pronoun to which the sentence-initial NP *na-chang huo* 'that fire' can possibly relate. (114) below is such an example.

(114) Na-chang huo, xingkui xiaofang-dui lai-de-kuai, [buran shei
that-CL fire fortunately fire-brigade come-PART-fast otherwise who
zhidao women jintian hui zai nar].
know we today will at where
'As for that fire, fortunately the fire brigade came quickly; otherwise, who knows where we would have been today.'

In this well-formed (114), the topic 'that fire' cannot be related to any position within the clause introduced by *buran* 'otherwise' and it remains dangling. Examples like this undermine Shi's argument we have reviewed so far.

The fourth kind of dangling topic in Shi's analysis is illustrated below in (115) by the sentence-initial NP, *zhe-jian shiqing* 'this matter'. According to him, the said NP in (115) is not subcategorized by the main verb *mafan* 'to trouble', and the comment clause contains no gap or resumptive pronoun for *zhe-jian shiqing* 'this matter' to relate to. Therefore, the sentence-initial NP has appeared to be a dangling topic.

(115) (= (52))
[Zhe-jian shiqing]$_{\text{Topic}}$ [ni bu neng guang mafan yige ren]$_{\text{Comment}}$.
this-CL matter you not can only trouble one person
'(As for) this matter, you cannot just trouble one person,' (meaning that this matter requires more than one person's efforts).'

We first came across the example in (115) in 2.1.1. Shi's argument against the dangling status of the topic in question is to hypothesize that the NP *zhe-jian shiqing* 'this matter' is actually an oblique object of an adjunct adverbial preposition phrase (PP), with the preposition head *wei* 'for' omitted, as (116) illustrates, which means the same as (115).

(116) Wei zhe-jian shiqing, [ni bu neng guang mafan yige ren].
for this-CL matter you not can only trouble one person
'As for this matter, you cannot just trouble one person,' (meaning that this matter requires more than one person's efforts).'

It is true that in Chinese, some prepositions can often be dropped from an adverbial PP and the remaining elements can still mean the same as the original PP with an overt

preposition does. Shi argues that because the NP *zhe-jian shiqing* 'this matter' in (115) is derived from the PP 'for this matter' with the preposition *wei* 'for' omitted and the remaining NP itself definite, this remaining NP in question functions as a topic covertly marked by the preposition *wei* 'for', as in (116). In other words, the NP *zhe-jian shiqing* 'this matter' in (115) is syntactically licensed by an omitted preposition *wei* 'for' and, therefore, it is not a dangling topic. This line of analysis seems plausible for examples like (115) and (116) but, as Pan and Hu (2002) point out, it would run into trouble when it comes to a sentence like (117), revised from Pan and Hu's (2002) example (20).

(117) a. *Zhe-jian shiqing, Zhangsan gen ren dajia le.
this-CL matter with people fight ASP
'Zhangsan fought with people for this matter.'
b. Wei zhe-jian shiqing, Zhangsan gen ren dajia le.
for this-CL matter with people fight ASP
'Zhangsan fought with people for this matter.'

(117a) is structurally similar to (115), and so is (117b) to (116). If Shi's preposition-omitting view is correct, Pan and Hu (2002) question why the underlined sentence-initial NP in (117a) cannot drop the prepositional head *wei* 'for', since the underlined PP in (117b) parallels that in (116). Pan and Hu's observation suggests that Shi's preposition-omitting analysis is flawed, and the status of the sentence-initial NP in (115) should be resolved differently. Moreover, when discussing how a topic should be licensed in LFG in 2.1.1, we also offered an alternative view of reanalyzing the sentence in (115) and treating the sentence-initial NP as a fronted object subcategorized by the ditransitively-used main predicate, *mafan* 'to trouble'.

The fifth type of dangling topic in Shi's categorization is one like the sentence-initial NP in (118) below.

(118) (Shi 2000, (38))
[Na-zhong douzi]$_{Topic}$ [yi-jin sanshi-kuai quian]$_{Comment}$.
that-kind beans one-catty thirty-CL money
'That kind of beans, one catty is thirty dollars.'

Shi describes a sentence like (118) as characteristically consisting of two sets of NPs with no verb as the predicate. As the brackets indicate, (118) is traditionally analyzed as a topic-comment structure, and the aboutness relation holds between the two. Not related to any structural position inside the comment clause, the topic here is regarded as dangling. Arguing against this dangling topic analysis, Shi proposes to consider the relationship between the two bracketed elements as subject and predicate, rather than topic and comment, by hypothesizing that sometimes an NP can also function as a predicate (i.e., nominal predicate) in modern Chinese, and it takes another NP as a subject. In other words, Shi proposes that in (118), the seemingly dangling topic, *na-zhong douzi* 'that kind of beans', is the subject of the nominal predicate, *yi-jin sanshi-kuai qian* 'one catty for thirty dollars'. Arguing for the subject status of the sentence-initial NP, 'that kind of beans', in (118), Shi claims that this NP does not have to be definite, which accords with the general consensus that while a topic must be definite, a subject need not be. Despite claiming so, Shi does not give any example of how the sentence-initial NP in (118) can be indefinite, and it is indeed difficult for us to imagine how it can possibly be so. Instead, Shi provides the following example (119), a well-known Chinese nursery rhyme which parallels structurally to that of (118), and then claims that in the same fashion, the two bracketed constituents, both being indefinite, are in fact the subjects, not the topics, of the nominal predicates, i.e. (having) four legs and (having) eight legs, respectively.

(119) (Shi 2000, (39))

[Yi-zhi	qingwa]	si-tiao	tui,	[liang-zhi	qingwa]	ba-tiao	tui.
one-CL	frog	four-CL	legs	two-CL	frogs	eight-CL	legs

'A frog (has) four legs, and two frogs (have) eight legs.'

However, there are at least two problems with this analysis. First of all, regardless of whether the bracketed constituents in (119) are subjects or topics, although their structural forms suggest that they are indefinite, they are an instance of indefinite generic usage. Such indefinite genericity refers to the general fact that any one frog will have four legs, and any two frogs will have eight legs, which is *not* the same as a non-generic indefinite reading. A non-generic indefinite reading of (119) would mean that a certain frog has four

legs, which could imply that other frogs might not have four legs, or that the speaker believes that the hearer does not already know this fact. To treat indefinite genericity (e.g., any one frog has four legs) on a par with non-generic indefiniteness (e.g., a certain frog has four legs) as having the same kind of unspecified nature would be overlooking the difference between the two. We believe that there is still some sense of definiteness in the generically used NPs in (119), despite their indefinite forms. Furthermore, an indefinite subject in Chinese is generally marked (or prefixed) by *you* 'existential there' (Li and Thompson 1989), while a definite subject is not marked at all, as the contrast between (120a) and (120b) shows. (120c), on the other hand, is well-formed because *yi-tai diannao* 'one computer' here is used generically as a class of entities. Such indefinite genericity cannot be prefixed by *you* 'existential there', which suggests that, at least in Chinese, indefinite generic nouns carry definite meaning to some extent. The prefixed version of (120c) is demonstrated in (120d) with a completely different, non-generic meaning. The above observation is in agreement with our view of the definite nature of the two bracketed constituents in (119), contra Shi's speculation.

(120) a. *Yi-tai diannao huai le.
one-CL computer break ASP
'A computer broke down.'
b. You yi-tai diannao huai le.
there.be one-CL computer break ASP
'A computer broke down.'
c. Yi-tai diannao hen gui.
one-CL computer very expensive
'A computer is very expensive.'
d. You yi-tai diannao hen gui.
there.be one-CL computer very expensive
'A certain computer is very expensive.'

Secondly, suppose that Shi's claim that the initial NP in (118) does not have to be definite is valid even though he does not explain how. Intuitively, as Pan and Hu (2002) also discuss, such a sentence might look like the one in (121), which is ill-formed due to

the requirement for an indefinite subject in Chinese to be prefixed by *you* 'existential there', and (121) does not allow for a generic reading, either.

(121) (Pan and Hu 2002, (13))
*Yi-zhong douzi, yi-jin sanshi-kuai qian.
one-kind beans one-catty thirty-CL money
'One catty of a kind of beans is thirty dollars.'

Shi claims that an important feature of the initial NP 'that kind of beans' in (118) is that it does not have to be definite, which, to the best of our knowledge as a native Chinese speaker, might be understood as licensing a sentence like (121). However, (121) is clearly ungrammatical and can only be paraphrased by (122).

(122) You yi-zhong douzi, yi-jin sanshi-kuai qian.
there.be one-kind beans one-catty thirty-CL money
'One catty of a kind of beans is thirty dollars.'

Arguing against Shi's analysis, Pan and Hu (2002) notice that the inability of the sentence-initial NP, *na-zhong douzi* 'that kind of beans', in (118) to be indefinite, as the unacceptable (121) demonstrates, turns out to be evidence for its topical status, as topic is always definite in reference. As a result, we argue that Shi's treatment of his fifth kind of dangling topic, as in (118) and the like, to be a subject is not convincing, as its topical status still persists in being definite or generic.

The last type of dangling topic presented in Shi's categorization is instantiated by the sentence-initial NP *wu-jia* 'price of things (i.e., living expenses)' in (123) below, in which, in traditional views, *Niuyue* 'New York' is the subject of the comment clause predicated by the adjectival phrase *zui gui* 'most expensive', leaving the topic, *wu-jia* 'price of things', unrelated to any position inside the comment clause.

(123) (Shi 2000, (42))
[Wu-jia]$_{\text{Topic}}$ [Niuyue zui gui]$_{\text{Comment}}$.
thing-price New York most expensive
'Speaking of the price of things, New York is the most expensive.'

In an attempt to syntactically resolve all the problematic, dangling topics, Shi proposes to treat *Niuyue* 'New York' as a locative adverbial modifying the subject, *wu-jia* 'price of things'. He claims that this reading, in which *wu-jia* 'price of things' is the subject and *Niuyue* 'New York' is the locative adverbial, is more natural than the other reading, in which *Niuyue* 'New York' is the subject and *wu-jia* 'price of things' is the non-gapping topic, although this preference is only Shi's own judgment, and we find both readings equally natural. Shi's second argument for the subject status of the sentence-initial NP *wu-jia* 'price of things' stems from the possibility of inserting a modal verb like *hui* 'will', as (124) instantiates, which cannot precede but must follow a subject; or of inserting an adverbial like *yiding* 'certainly', as in (125), which can only occur between a subject and a predicate, but never before a subject.

(124) (Shi 2000, (44))

Wu-jia hui Niuyue zui gui.
thing-price will New York most expensive
'The price of things will be the highest in New York.'

(125) (Shi 2000, (43))

Wu-jia yiding Niuyue zui gui.
thing-price certainly New York most expensive
'The price of things is certainly the highest in New York.'

In other words, according to Shi, the availability of inserting the modal verb *hui* 'will' or the adverbial *yiding* 'certainly' between the two NPs, as (124) and (125) demonstrate, shows that the first NP should really be treated as the subject of the sentence, and the second NP is simply a locative adverbial modifying this subject. Thus analyzed, the sentence-initial NP, *wu-jia* 'price of things', in (123) is not a dangling topic, but the subject of the sentence in Shi's account.

In reply to Shi's analysis on sentences like (123), we agree that it is indeed possible to treat the initial NP as the subject and the second NP as a locative adverbial modifying the subject, but we disagree with Shi that this reading is necessarily more natural than the other reading, in which the first NP is the non-gapping topic and the second NP is the subject. Given Shi's treatment of inserting a modal verb, or an adverbial, between the two

NPs in (123), as (124) and (125) exemplify, to argue for the subject status of the first NP, we believe that this line of analysis can also be used to favor the subject status of the *second* NP in (123), that is, *Niuyue* 'New York', leaving the first NP, *wu-jia* 'price of things', to be the dangling topic of the sentence. The following (126) and (127) correspond to (124) and (125), respectively. The optional discourse particle following the initial NP in (126) and (127) is one of those pause particles in Chinese, e.g., *ya*, *ne*, *me*, and *ba*, which overtly indicate a pause and can optionally follow a topic NP to separate it from the comment clause. In (126) and (127), the topical status of *wu-jia* 'price of things' is marked by the optional pause particle, and the second NP *Niuyue* 'New York' is then undoubtedly the subject, followed by the same modal verb or adverbial Shi has used to argue for the subject status.

(126) Wu-jia (a), Niuyue hui zui gui.
thing-price PART New York will most expensive
'(Speaking of) the price of things, New York will be the most expensive.'
(127) Wu-jia (a), Niuyue yiding zui gui.
thing-price PART New York certainly most expensive
'(Speaking of) the price of things, New York is certainly the most expensive.'

Viewed this way, (123) can have two equally natural readings, one of which poses a problem to Shi's syntactic account. The unproblematic reading is favored by Shi with *wu-jia* 'price of things' as the subject and *Niuyue* 'New York' as the locative adverbial, whereas the other reading has *Niuyue* 'New York' as the subject and *wu-jia* 'price of things' as the topic, which appears unrelated to any position inside the comment clause. A solution different from Shi's will be provided in 2.3.3.

In this section we have seen the six types of dangling topic constructions categorized by Shi, and his attempts to account for them from a purely syntactic point of view by assuming those dangling topic constructions should be reanalyzed variously as instances of arguments. After the reanalysis, the so-called non-gapping topics are not dangling at all, but are related, one way or another, to some position in the sentences. In the discussion of each type of dangling topic constructions, we showed how we disagree with Shi and the

problems his account suffers. In contrast to Shi's pro-syntax analysis, there is a pro-semantics hypothesis offered by Pan and Hu (2002) towards explaining away the six types of dangling topics. This will be addressed in the next subsection.

2.3.2 A pro-semantics view by Pan and Hu (2002)

In comparison to the syntactic analysis offered by Shi (2000) of dangling topic constructions introduced in 2.3.1, Pan and Hu (2002) propose a semantics-oriented analysis of dangling topics. Like Shi, Pan and Hu also believe that a topic should be related to some position inside the comment clause, and the occurrence of a dangling topic should be resolved somehow. In Shi's view, a topic should be associated with a syntactic position inside the comment clause either as an empty category or a resumptive pronoun, and many of the seemingly dangling topics may be reanalyzed as subjects instead. Alternatively, Pan and Hu argue that in addition to being licensed by an empty category or a resumptive pronoun, a topic can be semantically linked to a *semantic variable* in the comment clause as the third option. This semantic variable hypothetically saturates somewhere in a comment clause which licenses the occurrence of a dangling topic, but this semantic variable need not have any syntactic realization. In Pan and Hu's analysis, those dangling topics previously reanalyzed syntactically by Shi (2000) as subjects or adverbials are legitimate topics *semantically licensed* (by semantic variables) in the way we are going to examine below. (100), repeated here as (128), is the first type of dangling topic construction according to Shi's classification, and Shi analyzes the initial NP as the subject modified by the rest of the sentence as an idiom chunk.

(128) (= (100))

[Tamen]$_{\text{Topic}}$ [wo kan ni ni kan wo]$_{\text{Comment}}$.
they I look you you look me
'They look at each other.'

The questionability of Shi's analysis has been discussed in 2.3.1. Pan and Hu offer a different point of view: the initial NP in (128) is a semantically licensed topic by the set-member relation holding between the initial NP *tamen* 'they' and the subject pronouns

inside the comment clause, i.e., *wo* 'I' and *ni* 'you'. To be specific, the topic NP *tamen* 'they' delimits a set, and the subject pronouns inside the comment clause (*wo* 'I' and *ni* 'you') pertain to the members of this set. Notice that the subject pronouns in the comment clause are not to be taken literally to mean the speaker and the hearer, which would have been an impossible reading for (128). The fact that the exact referents to whom *wo* 'I' and *ni* 'you' in (128) refer are unclear is immaterial, as long as they do refer to the members of the set provided by the topic NP, *tamen* 'they'. Thus viewed, the pronouns in the comment clause, *wo* 'I' and *ni* 'you', can be seen as semantic variables with their values constrained by the set, *tamen* 'they'. Formally, Pan and Hu propose the following semantic licensing condition in (129) to resolve topic-comment sentences in Chinese. In (129), P stands for a predicate while x and y are individual-level variables.

(129) (Pan and Hu 2002, (22))

a. $\lambda P\, \lambda Z\, \lambda x\, \lambda y\, [x \in Z \,\&\, y \in Z \,\&\, P(x, y)]$ (for a two-place predicate P)

b. $\lambda P\, \lambda Z\, \lambda x\, [x \in Z \,\&\, P(x)]$ (for a one-place predicate P)

c. $\lambda P\, \lambda Z\, \lambda x_1 \ldots \lambda x_n\, [x_1 \in Z \,\ldots\, \&\, x_n \in Z \,\&\, P(x_1, \ldots x_n)]$ (for an n-place predicate P)

For instance, a dangling topic construction like (128) can be semantically resolved by employing (129a), as *kan* 'look at' in (128) is a two-place predicate. Applying (129a) to the comment clause in (128), in which x equals *wo* 'I' and y equals *ni* 'you' with the predicate *kan* 'look at' as P, yields (130) below. Note that the two occurrences of the predicate 'look at' in (130) is derived from the fact that the comment clause in (128) consists of a coordination of two clauses (i.e., I look at you and you look at me).

(130) (Pan and Hu 2002, (24))

$\lambda Z\, [I \in Z \,\&\, you \in Z \,\&\, \text{Look-at}\, (I, you) \,\&\, \text{Look-at}\, (you, me)]$

Pan and Hu's analysis continues: now (130) is an *open* semantic predicate, so it can take another argument by being applied to the topic NP *tamen* 'they' in (128), which in turn results in (131) below.

(131) λZ [I∈Z & you∈Z & Look-at (I, you) & Look-at (you, me)] (they)
= [I∈they & you∈they & Look-at (I, you) & Look-at (you, me)]

Pan and Hu believe that (131) correctly translates the meaning of (128) as two people who, together, form a third person plural topic NP (*tamen* 'they') looking at each other. In accordance with (131), Pan and Hu conclude that the seemingly dangling topic *tamen* 'they' in (128) is, in fact, not dangling at all, but licensed by the open semantic predicate presented in (131). This line of analysis seems novel at first sight, but under close scrutiny, we find out that the meaning of (131) does not truly reflect the meaning of (128), and the method Pan and Hu propose to analyze topic-comment sentences also suffers the following problems. Firstly, it is unclear to us how the semantic formulae of lambda calculus proposed in (129) to resolve topic-comment sentences are grounded either theoretically or empirically. How the formulae in (129) are derived and why they are purposely designed for Chinese topic-comment constructions is not mentioned in Pan and Hu (2002), so we could only suppose that the motivation of the proposed formulae would hinge upon the absolute accuracy and truthfulness when the formulae, as a semantic licensor, are applied to sentences to represent their formal meaning. However, this expectation is not met, which will be explained as we proceed to our second argument. If the formulae do not yield correct meanings, the validity of Pan and Hu's proposal is questionable. Secondly, we agree with Shi (2000) and Pan and Hu (2002) that under the circumstances of (128), the two pronouns are not to be taken literally to mean the speaker and the hearer, but only metaphorically, as the sentence is unable to describe a staring situation between the speaker and the hearer. In our view, *wo* 'I/me' and *ni* 'you' in (128) should be analyzed as two idiomatic personal pronouns whose actual reference is unspecified and also immaterial. This prevalent Chinese phrase *wo kan ni ni kan wo* '(literally) I look at you and you look at me' simply means 'look at each other' metaphorically without actually referring to any person, especially not the speaker or hearer. The role that the first and the second person pronouns play in this phrase is neither referential nor should be taken literally unless in the context of (132) below, in which the personal pronouns inside the comment clause just happen to refer to the speaker and the

hearer by accident. Shi's analysis treats the whole comment clause in (128) as a fossilized idiom chunk and allows no substitution or change of order of the lexical items inside, whereas we propose to treat only the personal pronouns inside this phrase as metaphorical, idiomatic, or non-referential.

(132) [Women]$_{\text{Topic}}$ [wo kan ni ni kan wo]$_{\text{Comment}}$.
we I look you you look me
'We look at each other.'

Next, if we suppose that Pan and Hu's proposal and the semantic licensing condition in (129) are without problems, (131), repeated below as (133), implies that in the meaning of the sentence in (128), the participants consist of only two people, namely the x and y in (129a) and no more. It would, at best, describe a scenario given by (134) below, which is elaborated from (128) for the sake of contextual clarity and consists of two topical referents as members of *tamen* 'they'.

(133) λZ [I$\in$Z & you$\in$Z & Look-at (I, you) & Look-at (you, me)] (they)
= [I$\in$they & you$\in$they & Look-at (I, you) & Look-at (you, me)]

(134) [Speaking of Alex and Bill, they broke an expensive vase in an antique shop.]
Tamen$_i$, wo kan ni ni kan wo, Ø$_i$ yi-ju hua ye mei you shuo.
they I look you you look me NOM one-CL sentence also not have speak
'They (Alex and Bill) looked at each other; (they) didn't even say a word.'

The first part of (134), underlined, is the same sentence as in (128), with an additional second clause describing that the referents did not even say a word (after breaking the expensive vase). The syntactic and semantic significance of the original (128) remains intact in (134). Suppose Pan and Hu's representation in (133) is correct and does translate the underlined sentence in (134). Pan and Hu's proposal in (130) runs into trouble when the sentence in (128) is used in a context like (135), in which the topic consists of more than two persons. (135) is identical to (134) except for the number of referents involved in the topic.

(135) [Speaking of Alex, Bill, Chris, Dave and Eric, they broke an expensive vase in an antique shop.]

Tamen$_i$, wo kan ni ni kan wo, Ø$_i$ yi-ju hua ye mei you
they I look you you look me NOM one-CL sentence also not have
shuo.
speak
'They (the five people) looked at each other; (they) didn't even say a word.'

Pan and Hu's analysis can at best translate the meaning of (128) when it appears in a context like (134), where there are two and only two referents forming the topic, say x and y, and x and y indeed look at each other. However, it poses a challenge to Pan and Hu's analysis when there are more than two persons referred to by the topic, as in (135). In this reciprocal statement, these five people in (135) form an antecedent group, and the most natural reading of such reciprocity would simply require every member in the antecedent group to at least look at another one, and being only passively looked at does not count (Dalrymple et al. 1998). (133) may not reflect the meaning of *tamen* 'they' as denoting a group of more than two people. Nor can (133) capture the more plausible way of 'five people looking at each other' (that is, each member need to look at another one at least). Instead, (133) may only describe an unnatural situation in which each member must be looked at by the other four equally. If Pan and Hu's semantic formulae are correct, it would simply translate the first clause in (135) as any random two referents out of the five (delimited by the topic) looking at each other under the constraint that there are only two individual variables, i.e., x and y, for the two-place predicate *kan* 'look at'. This line of analysis cannot translate the correct group reading of five people looking at each other, and it also fails to capture the more plausible manner of looking at each other between the five referents. Therefore, our second argument against Pan and Hu's semantic account comes from the inability of their proposed semantic formulae, shown in (129), to translate the meaning of (128) correctly, especially when (128) occurs in a scenario like (135).

Next, the second dangling topic structure discussed by Shi (2000) is exemplified by (104), repeated below as (136). Pan and Hu propose that the comment clause in (136) can be translated by the lambda calculus expression in (137a). This formula forms a

semantically open predicate which can then take the topic NP, *tamen* 'they', to be the argument Z. After the application of (137a) to the topic NP, the formal meaning of (136), shown in (137b), is translated as 'for every individual x, if x is a member of (the unspecified topic) them, x is not coming'.

(136) (= (104))
[Tamen]$_{Topic}$ [shei dou bu lai]$_{Comment}$.
they who all not come
'(Speaking of) them, none of them are coming.'

(137) a. $\lambda Z\ \forall x\ [x \in Z \rightarrow \sim Come(x)]$
b. $\lambda Z\ \forall x\ [x \in Z \rightarrow \sim Come(x)]$ (they)
$= \forall x\ [x \in they \rightarrow \sim Come(x)]$

The first problem with this analysis is, again, to do with how and why Pan and Hu's proposed lambda calculus expression, as seen in (137), is able to formalize the meaning of (136), and then license the topic NP *tamen* 'they'. It is not yet clear to us on what ground, theoretically or empirically, (137) can be a formal licenser. Secondly, even if we assume Pan and Hu's approach to resolving the dangling topic in (136) to be well-established and without problems, their analysis would result in an incorrect translation if the verb in (136), *lai* 'come', is replaced by a collective predicate (e.g., *forming a line* or *looking alike*), which, by definition, always requires a plural subject. To illustrate our point, let us look at (138), whose sentence structure is identical to that of (136) except that the verbs differ.

(138) Tamen [shei dou bu he].
they who all not compatible
'(Speaking of) them, none of them are compatible with each other; (or alternatively, "they are not compatible with anybody", as this sentence is ambiguous between the two readings, with the latter one irrelevant to our discussion here).'

Suppose that Pan and Hu's analysis of (136) is on the right track and without any aforementioned problems. The formalized meaning of (138), under Pan and Hu's lambda calculus, would look like (139) according to their hypothesis, and this gives rise to an

undesirable result: it translates (138) as 'for every individual x, if x is a member of (the unspecified topic) them, x is compatible', which is clearly ungrammatical because the collective predicate *he* 'being compatible' in Chinese must take a plural subject just like its English equivalent. *He is compatible*[25] is not a well-formed sentence in English. Therefore, based on this incorrect translation in (139) yielded by their semantic formula, we argue that Pan and Hu's analysis does not seem to hold.

(139) $\lambda Z \forall x [x \in Z \rightarrow \sim \text{Compatible}(x)]$ (they)
$= \forall x [x \in \text{they} \rightarrow \sim \text{Compatible}(x)]$

In the discussion above, we have illustrated how Pan and Hu's semantic approach to tackling dangling topics in Chinese does not accommodate the problems they aim to address. Neither is their hypothesis general enough to yield a correct prediction, as evidenced by the examples in (135) and in (138). We acknowledge Pan and Hu's contribution in attempting to explain dangling topics at the semantic level; however, their novel proposal of the semantic licensing conditions seen in (129) does not seem to rest on any solid theoretical or empirical evidence, while these dubious semantic formulae form the very core of their semantic account of Chinese non-gapping topics. These concerns have undermined the validity of Pan and Hu's semantic approach. In this section, we only show a few examples of how Pan and Hu have analyzed dangling topics based on Shi's (2000) classification, and our argument that their view oversimplifies the matter and suffers some fundamental problems. The remaining examples of dangling topics discussed in Pan and Hu (2002) are the same ones that we have already seen from Shi's examples in 2.3.1. Though not covered here, Pan and Hu's analysis of those remaining examples is still driven from their lambda calculus seen in (129) and suffers similar problems in one way or another. We have demonstrated clearly a number of the examples as representatives and shown that their core semantic licensing condition is questionable.

[25] One may say 'A is compatible with B' in English, but if the preposition phrase 'with B' is removed and given no specific context, the remaining sentence 'A is compatible' becomes unacceptable. There is some narrow context in which 'A is compatible' acts as a legitimate sentence, e.g., as an elliptical reply to the question 'Is A compatible with B?'. It is the same in Chinese: if the subject of 'being compatible' is singular, a prepositional phrase similar to 'with B' must follow. If the subject is plural, the prepositional phrase becomes optional, as in the case of (138).

We will not continue to review how Pan and Hu have analyzed every example of dangling topics taken from Shi's classification introduced in 2.3.1, because going through each of them one by one will be less relevant to the focus of our study, which is to introduce an alternative semantic analysis. Readers who are interested in the full-length discussion of Pan and Hu's semantic method can refer to their original work.

2.3.3 An LFG solution for Chinese dangling topics

The previous two subsections provide a comparative view on tackling Chinese dangling topics via syntactic and semantic methods, and how such absolute angles (e.g., absolutely syntactic or absolutely semantic) are of restricted value because they exclude themselves from the benefit of combining tools from syntax and semantics. We believe that an account that takes both syntax and semantics into consideration would be ideal, and what follows is such an attempt at explaining dangling topics in Chinese under the framework of LFG. In this account, the problematic dangling topics may be resolved (i) by means of syntactic reanalysis, the same approach taken by Shi (2000) but argued differently here; or (ii) by resorting to semantics or pragmatics and exploiting the role a given dangling topic plays at the level of *i*-structure. However, since the goal of this thesis is to tackle empty pronouns in Chinese, instead of addressing the full-fledged problems that surround topics of all sorts in Chinese, our discussion of dangling topics so far simply serves more as a preliminary which aims to review two contrastive analyses before proceeding to offer an alternative solution balancing the previous two views, than a thorough, exhaustive exploration of all types of problematic topics in Chinese. In other words, this subsection will present an LFG proposal for three kinds of dangling topics among the six aforementioned types in Shi (2000), which are also reviewed by Pan and Hu (2002). Our discussion below is set as an example of how a certain kind of dangling topic can be resolved by syntactic reanalysis, while some can be integrated at *i*-structure through a pragmatic approach. This proposal is by no means a complete account of every type of dangling topics in Chinese. We seek to sketch an outline that opens up a new possibility of resolving dangling topics and also invites further research into the problems observed.

Let us begin with the familiar example (100), repeated in the previous subsection as (128) and here as (140) below.

(140) [Tamen]$_{Topic}$ [wo kan ni ni kan wo]$_{Comment}$.
they I look you you look me
'They look at each other.'

The problems with Shi's (2000) and Pan and Hu's (2002) analyses have been addressed in the previous two subsections. Here we argue that the bracketed comment clause in (140) is an instance of a special sentence structure in Chinese, *wo* + V_i(erb) + *ni* + *ni* + V_i + *wo* 'I V_i you and you V_i me' with both verbs being identical (hence the shared index on V). The order of the personal pronouns *wo* 'I' and *ni* 'you' in the said structure can also be altered as *ni* + V_i + *wo* + *wo* + V_i + *ni* 'you V_i me and I V_i you', and the meaning remains essentially the same. In addition to the restriction that both verbs in this structure must be identical, the referent(s) modified by this sentence structure (e.g., *tamen* 'they' in (140)) must be plural. This restriction on plurality is to be explained by the following. The said sentence structure acquires its unique status because it conveys an idiomatic, reciprocal reading as 'whichever the plural referents described by this very sentence structure are, the referents are carrying out a certain activity (denoted by the V in question) with respect to each other'. Therefore, a singular referent will be incompatible with such a reciprocal reading. (Imagine a sentence like **The boy is hitting each other*). Take the verb *kan* 'look at' for example: *wo kan ni ni kan wo* 'I look at you and you look at me' will simply mean 'to look at each other', and similarly, *wo da ni ni da wo* ' I hit you and you hit me' will mean 'to hit each other', with *da* being the verb 'to hit'. The first and second person pronouns in these examples are idiomatic and have no semantic contribution. In other words, the first and second person pronouns that appear in this structure have lost their semantic content completely, and must not be taken literally, as also argued by Shi (2000) and Pan and Hu (2002). This amounts to saying that in (140), the string of words *wo kan ni ni kan wo* simply means 'to look at each other', and we propose to treat it as an instance of COMP. The reason we do not see it as an open COMP, i.e., XCOMP, is that despite the idiomatic interpretation, the said construction does appear to be structurally

complete (but semantically incomplete). A plural subject is needed. Thus viewed, *tamen* 'they' in (140) is not a dangling topic anymore, but a legitimate subject. This syntactic reanalysis is similar to that claimed in Shi (2000), but we have pointed out in 2.3.1 that Shi's view is flawed in analyzing the whole bracketed comment clause in (140) as one big idiom chunk which allows no further analysis or internal alternation. His statement incorrectly implies that the elements inside the idiom chunk are fixed, neither movable nor replaceable, contra the possibilities we have shown so far. Instead, we believe that what is idiomatic is the use of the first and second person pronouns, as they are truly fixed and cannot be replaced in order to preserve the meaning of reciprocality. Note that this construction, *wo* + V_i + *ni* + *ni* + V_i + *wo* 'I V_i you and you V_i me' (or the alternative order of the pronouns), yields the idiomatic reading of carrying out some activity with respect to each other, *only* when the first and second person pronouns are employed. Replacing any of the first or second person pronouns in this unique construction, say, with a third person pronoun, will cause the pronouns inside this structure to lose their idiomatic status, and the whole phrase after the replacement, if it still remains grammatical at all, will only express a literal meaning. The example in (141) illustrates the possibility of replacing the second person pronoun, *ni* 'you', with the third person pronoun, *ta* 'he/she', and the resultant sentence is semantically incomprehensible, allowing neither idiomatic interpretation nor literal reading.

(141) *Tamen [wo kan ta ta kan wo].
they I look him/her he/she look me

Despite our claim that the first and second person pronouns have lost their literal meaning in this unique construction and become purely idiomatic, as in (140), if the plural referent required by this unusual structure happens to be *women* 'we', as in (142) below, the literal meaning of these pronouns does appear to be kept, as an accidental coreference.

(142) Women [wo kan ni ni kan wo].
we I look you you look me
'We look at each other.'

When standing alone on its own, the structure *wo* + V_i + *ni* + *ni* + V_i + *wo* 'I V_i you and you V_i me' maintains its literal meaning. The example in (143) demonstrates such a usage.

(143) Wo kan ni ni kan wo.
I look you you look me
'I look at you and you look at me.'

In the construction of *wo* + V_i + *ni* + *ni* + V_i + *wo* 'I V_i you and you V_i me', given the idiomatic status of the pronouns, we propose that whenever this sentence structure is used to modify a plural NP or pronoun, the following idiom rule in (144), which is based on the analysis of Dalrymple et al. (1998), applies and a reciprocal reading will be prescribed.

(144) P (*wo* 'I', *ni* 'you') & P (*ni* 'you', *wo* 'I') → λ x. RECIP (x, P), in which P stands for a predicate, x stands for an NP (or a pronoun) modified by the sentence structure, and RECIP is an operator that conveys the reciprocal relationship between x and P.

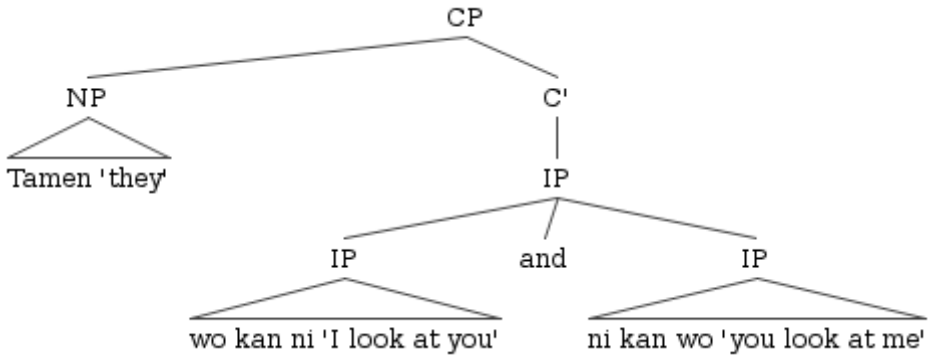
For instance, given the idiom rule in (144), the formalized meaning of the bracketed comment clause in (140) can be captured by (145a). When we apply the meaning of (145a) to the meaning of the subject, *tamen* 'they', the complete meaning of (140) can be formally represented in (145b).

(145) a. *kan* 'look at' (*wo* 'I', *ni* 'you') & *kan* 'look at' (*ni* 'you', *wo* 'I') →
λ x. RECIP (x, *kan* 'look at')
b. [*kan* 'look at' (*wo* 'I', *ni* 'you') & *kan* 'look at' (*ni* 'you', *wo* 'I') →
λ x. RECIP (x, *kan* 'look at')] (*tamen* 'they')
= RECIP *(tamen* 'they', *kan* 'look at')

(146) and (147) below are the *c*-structure and *f*-structure of the sentence in (140). As the *f*-structure in (147) shows, we are treating the clause *wo kan ni ni kan wo* 'I look at you and you look at me' as an instance of COMP, consisting of two conjuncts. The

c-structure in (146)[26] reflects such a conjoined structure. At the *f*-structure in (147), the COMP and the SUBJ required by PRED are the two arguments connected by the abstract predication RECIP, yielding a reciprocal meaning. *Tamen* 'they' has a dual status as SUBJ and TOP at the *f*-structure in (147); its topical status comes from its role at *i*-structure sketched in (148).

(146) *c*-structure of (140)



(147) *f*-structure of (140)

$$\begin{bmatrix} \text{SUBJ} & {}_1\begin{bmatrix} \text{FORM} & \textit{tamen}\ \text{'they'} \\ \text{PRED} & \text{'PRO'} \end{bmatrix} \\ \text{TOP} & {}_1\begin{bmatrix} \end{bmatrix} \\ \text{COMP} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}\textit{kan}\ \text{"look at"}\langle \text{SUBJ, OBJ}\rangle\text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{FORM} & \text{'}\textit{wo}\ \text{"I"'} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{FORM} & \text{'}\textit{ni}\ \text{"you"'} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \text{PRED} & \text{'}\textit{kan}\ \text{"look at"}\langle \text{SUBJ, OBJ}\rangle\text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{FORM} & \text{'}\textit{ni}\ \text{"you"'} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{FORM} & \text{'}\textit{wo}\ \text{"me"'} \end{bmatrix} \end{bmatrix} \right\} \\ \text{PRED} & \text{'RECIP}\ \langle \text{SUBJ, COMP}\rangle\text{'} \end{bmatrix}$$

[26] A topic is generally analyzed as occupying the Specifier of CP in Chinese (Chen 1995).

(148) *i*-structure of (140)

$$\begin{bmatrix} \text{TOP} & {}_1\begin{bmatrix} \text{FORM} & \textit{tamen}\,\text{'they'} \\ \text{PRED} & \text{'PRO'} \end{bmatrix} \\ \text{FOC} & \begin{bmatrix} \text{PRED} & \text{'}\textit{wo kan ni ni kan wo}\ \text{"look at each other"}\langle \text{DRef} \rangle\text{'} \\ \text{DRef} & {}_1[\quad\quad] \end{bmatrix} \end{bmatrix}$$

Notice that since the personal pronouns *ni* 'you' and *wo* 'I/me' in (140) do not have any literal semantic content, they do not play any role at *i*-structure (illustrated in (148) above), even though their structural forms are still present at the *c*-structure and *f*-structure. For the same reason, *wo kan ni ni kan wo* is glossed as 'I look at you and you look at me' at *f*-structure because, syntactically speaking, that is the literal meaning conveyed. However, at *i*-structure, which houses the true communicative information, the same string of words is adequately translated as 'look at each other', in accordance with the view that the personal pronouns inside this phrase have no semantic contribution at the level of information structure. Based on the idiomatic rule in (144) and (145) specifically for the string of words *wo kan ni ni kan wo*, we employ a RECIP predicator to reflect the sentence's reciprocal interpretation at the *f*-structure in (147). There is no need to use RECIP if the reciprocality of a sentence is straightforward and literal (as opposed to idiomatic, non-literal). The *f*-structure for the sentence 'We love each other' in (149) below is such an example.

(149) We love each other.

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{FORM} & \text{we} \\ \text{PRED} & \text{'PRO'} \end{bmatrix} \\ \text{PRED} & \text{'love} < \text{SUBJ, OBJ} > \text{'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'each other'} \end{bmatrix} \end{bmatrix}$$

Proceeding to our second dangling topic example, we have come across (150) below, repeated from (136) in 2.3.2, in which the first NP *tamen* 'they' is traditionally analyzed as the topic, and the second NP *shei* 'who' is the subject. In this view, the initial NP would

be a problematic dangling topic because it is not related to any position inside the comment clause. (151) is a variant example of (150), in which the second NP *fu-qi* 'husband-wife' is a collective noun referring to a married couple coreferential with *tamen* 'they'. The sentence in (151), also involving a dangling topic, suffers the same problem as (150) does.

(150) [Tamen]$_{Topic}$ [shei dou bu lai]$_{Comment}$.
they who all not come
'(Speaking of) them, none of them are coming.'

(151) (= (109))
[Tamen$_i$]$_{Topic}$ [fu-qi$_i$ dou bu lai]$_{Comment}$.
they husband-wife all not come
'(Speaking of) them$_i$, neither member of the couple$_i$ (consisting of a husband and his wife) is coming.'

First of all, we need to clarify different usages of the *wh*-word in (150), *shei* 'who', in Chinese. *Shei* is an interrogative pronoun meaning 'who/which person' when it appears in a direct question, as in (152a), or in an indirect question, as in (152b).

(152) a. Ta shi shei?
he is who
'Who is he?'
b. Wo zhidao shei hui qu.
I know who will go
'I know who will go.'

In a declarative sentence, *shei* has the meaning of 'who(m)ever', which links one clause to another clause, as demonstrated in (153). As (153) shows, the *wh*-word *shei* appears as an argument selected by the preceding verb, *yaoqing* 'invite', and the shorter clause where *shei* occurs acts as the fronted object of the longer clause, as the two pairs of brackets indicate.

(153) [[Ni xiang yaoqing shei] wo dou keyi jieshou].
you want invite who I all can accept
'I can accept whomever you want to invite.'

Moreover, it is crucial to point out that where in a declarative or an interrogative sentence, *shei* will appear as an indefinite pronoun meaning 'anyone/everyone' (or 'no one' if the sentence is negative), if *shei* immediately precedes and is then modified by the quantifying adverbial *dou* 'all' (A. Li 1992; J. Li 1995; Shi 2000). This usage can be seen in the examples in (150) and (151). This indefinite use of *shei*, e.g., as 'anyone/everyone/no one', does not connect two clauses like *shei* in (153) does. This rule regarding the precedence of *dou* 'all' can be generalized and be applied to other *wh*-words, too. Take *shenme* 'what' for instance. *Shenme* in (154) is translated as 'whatever', whereas *shenme* becomes an indefinite pronoun to mean 'everything' in (155) because it is immediately followed by the quantifier *dou* 'all'.

(154) [[Ni shuo shenme] wo dou bu xiangxin].
you say what I all not believe
'Whatever you say I will not believe.'
(155) Wo shenme dou xiang-yao.
I what all want
'I want everything.'

Now let us return to (150) and (151), both of which begin the sentence with two NPs in a row and are structurally parallel. To be specific, the two sequential NPs in (150) consist of a personal pronoun *tamen* 'they' plus an indefinite pronoun *shei* 'anyone', whilst those in (151) are made up of the same personal pronoun *tamen* 'they' plus a referential NP *fu-qi* 'husband-wife'. Furthermore, we assume that these two sequential NPs can be viewed on a par to the first two NPs in an English sentence like *My friend John lives in the mountains*, in which the initial NP, *my friend*, is the subject, whereas the second NP, *John*, simply modifies the subject as an appositive and is omissible, as (156) shows. The second NPs of the sentences in (150) and (151) can also be omitted; the resulting sentences after the reanalysis, (157) and (158), are both well-formed but just less

informative. Notice that the two NPs in (157) are coindexed, but they are not coreferential, unlike their counterparts in (156) and (158).

(156) [My friend$_i$ (John$_i$)]$_{Subject}$ lives in the mountains.
(157) [Tamen$_i$ (shei$_i$)]$_{Subject}$ dou bu lai.
they anyone all not come
'None of them are coming.'
(158) [Tamen$_i$ (fu-qi$_i$)]$_{Subject}$ dou bu lai.
they husband-wife all not come
'Neither of the couple (consisting of a husband and his wife) is coming.'

In this analysis, (150) and (151) can be reanalyzed as (157) and (158), respectively. The first NPs in (157) and (158) are topical subjects, and the second NPs are simply nominal modifiers, either to strengthen the sentence meaning, as in the case of (157), or to add information, as in the case of (158). There is no dangling topic in our syntactic reanalysis. Our analysis differs from Shi's in treating the second, optional NPs in (157) and (158) as appositive NPs, each of which forms a subject constituent with the initial NP. The brackets in (156) to (158) indicate that the two serial NPs form a single constituent, and their constituency can be tested as follows. Take (157) and (158) for example. If a group of words can undergo movement, they form a constituent, as evidenced by (159), in which the bracketed elements act as the object, in contrast to its subject usage in (157). Also, the insertion of an adverb such as *weiweisuosuo-de* 'sheepishly' between *tamen* 'they' and *fu-qi* 'husband-wife', as shown by (160a), leads to ungrammaticality, which is consistent with the observation that no adverb can be positioned inside an NP constituent. (160b) is the correct way of conveying the meaning intended by (160a).

(159) Wo gen [tamen$_i$ (shei$_i$)]$_{Object}$ dou bu shou.
I with them anyone all not familiar
'I am not familiar with anyone of them.'
(160) a. *[Tamen$_i$ weiweisuosuo-de (fu-qi$_i$)] dou bu lai.
they sheepishly husband-wife all not come
'Sheepishly, neither of the couple is coming.'

b. [Tamen$_i$ (fu-qi$_i$)] weiweisuosuo-de dou bu lai.
they husband-wife sheepishly all not come
'Sheepishly, neither of the couple is coming.'

Lastly, we provide the *f*-structure and *i*-structure[27] for (157) and (158) below.

(161) a. *f*-structure of (157)

$$\begin{bmatrix} \text{PRED} & \text{'}lai\text{ "come"} < \text{SUBJ} > \text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{FORM} & tamen\text{ 'they'} \\ \text{PRED} & \text{'PRO'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}shei\text{ "anyone"'} \end{bmatrix} \right\} \end{bmatrix} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}dou\text{ "all"'} \end{bmatrix} \\ \begin{bmatrix} \text{PRED} & \text{'}bu\text{ "not"'} \end{bmatrix} \right\} \end{bmatrix}$$

b. *i*-structure of (157)

$$\begin{bmatrix} \text{TOP} & \begin{bmatrix} \text{FORM} & tamen\text{ 'they'} \\ \text{PRED} & \text{'PRO'} \end{bmatrix} \\ \text{FOC} & \begin{bmatrix} \text{PRED} & \text{'}lai\text{ "come"} < \text{DRef} > \text{'} \\ \text{DRef} & \begin{bmatrix} \text{PRED} & \text{'}shei\text{ "anyone"'} \end{bmatrix} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}dou\text{ "all"'} \end{bmatrix} \right. \\ & \left. \begin{bmatrix} \text{PRED} & \text{'}bu\text{ "not"'} \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix}$$

(162) a. *f*-structure of (158)

$$\begin{bmatrix} \text{PRED} & \text{'}lai\text{ "come"} < \text{SUBJ} > \text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{FORM} & tamen\text{ 'they'} \\ \text{PRED} & \text{'PRO'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}fu\text{ - }qi\text{ "husband - wife"'} \end{bmatrix} \right\} \end{bmatrix} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}dou\text{ "all"'} \end{bmatrix} \\ \begin{bmatrix} \text{PRED} & \text{'}bu\text{ "not"'} \end{bmatrix} \right\} \end{bmatrix}$$

[27] The type of *i*-structure assumed in our study excludes any issue regarding the relative scopes of a certain words (e.g., quantifiers and negation words).

b. *i*-structure of (158)

$$\begin{bmatrix} \text{TOP} & \begin{bmatrix} \text{FORM} & \textit{tamen}\ \text{'they'} \\ \text{PRED} & \text{'PRO'} \end{bmatrix} \\ \text{FOC} & \begin{bmatrix} \text{PRED} & \text{'}\textit{lai}\ \text{"come"} < \text{DRef} > \text{'} \\ \text{DRef} & [\text{PRED}\ \text{'}\textit{fu}\text{-}\textit{qi}\ \text{"husband-wife"'}] \\ \text{ADJ} & \left\{ \begin{matrix} [\text{PRED}\ \text{'}\textit{dou}\ \text{"all"'}] \\ [\text{PRED}\ \text{'}\textit{bu}\ \text{"not"'}] \end{matrix} \right\} \end{bmatrix} \end{bmatrix}$$

The last dangling topic sentence we will address in this study is in (163), repeated from (115) in 2.3.1. According to Shi (2000), the initial NP looks like a non-gapping topic unrelated to any structural position inside the comment clause and therefore poses a problem.

(163) (= (115))

[Zhe-jian shiqing]$_{\text{Topic}}$ [ni bu neng guang mafan yige ren]$_{\text{Comment}}$.
this-CL matter you not can only trouble one person
'(As for) this matter, you cannot just trouble one person,' (meaning that this matter requires more than one person's efforts).'

We showed in 2.1.1 that it is possible to treat the main verb *mafan* 'to trouble', normally a two-place predicate, as ditransitive, and the sentence-initial topic in (163) can be reanalyzed as a preposed object, the third argument of *mafan* 'to trouble', exemplified by (164) below.

(164) (= (54))

Ni bu neng guang mafan yige ren [zhe-jian shiqing]$_{\text{NP}}$.
you not can only punish one person this-CL matter
'(As for) this matter, you cannot just trouble one person.'

Other ditransitive use of *mafan* 'to trouble' can be viewed below in (165a) and (165b), repeated from our previous discussion in 2.1.1. The ungrammaticality of (165c), repeated from (55), supports our analysis of treating *zhe-jian shiqing* 'this matter' as a

subcategorized argument of *mafan* 'to trouble', because if it were just an unsubcategorized topic, (164c) should have been well-formed.

(165) (= (53), (55))

a. Bu-yao yizhi mafan bie-ren [bang ni jiejue wenti]$_{Complement}$.
don't constantly trouble other-people help you solve problems
'Stop troubling other people constantly to help you solve problems.'

b. Bu-yao mafan renjia [zheme duo wenti]$_{NP}$.
don't trouble other.people so many questions
'Stop troubling other people with so many questions.'

c. *[Zhe-jian shiqing], bu-yao yizhi mafan [bie-ren] [bang ni jiejue wenti]$_{Complement}$.
this-CL matter don't constantly trouble other-people help you solve problems
'As for this matter, stop troubling other people constantly to help you solve problems.'

As it is possible to rearrange the word order of (163) as that in (164) below, in which the sentence-initial topic is moved to the sentence-final position, the original non-gapping topic is no longer dangling in this new sentence structure. Though with a slightly different emphasis, (163) and (164) mean essentially the same. The fact that we are able to place *zhe-jian shiqing* 'this matter' in the sentence-final position, as in (164), confirms its non-dangling status when fronted as a sentence-initial topic, as in (163). Such a usage is not only syntactically licensed but also obeys the Extended Coherence Condition in (166) below.

(166) (= (49))

Extended Coherence Condition (ECC):
FOCUS and TOPIC must be linked to the semantic predicate argument structure of the sentence in which they occur, either by functionally or by anaphorically binding an argument.

2.4 Summary

Chapter 2 began with how topichood is defined in LFG and in Chinese, followed by a literature review of various previous attempts to tackle empty pronouns in Chinese and how each of them suffers different problems, before we reached the conclusion that an integrating analysis which takes both syntax and semantics into account would be a superior one. Then we proceeded to discuss the problems of dangling topics and the question of whether dangling topics should be syntactically or semantically resolved. After discussing Shi's and Pan and Hu's analyses of dangling topics and the weakness of their hypotheses, we presented an LFG account of how to tackle dangling topics by incorporating them at *f*-structure or at *i*-structure.

Chapter 3: Syntactic constraints on Chinese zero anaphora

Following the syntactic investigation of various issues surrounding Chinese zero anaphora and the topic-comment structure, and discussion of previous attempts to resolve Chinese empty pronouns in the previous chapter, Chapter 3 will look into how exactly Chinese empty pronouns are constrained by syntax at the intrasentential level (with minimum context if required). As we deal with intrasentential empty pronouns, we aim to reduce the effect that context might have on licensing zero anaphora. However, it may be impossible under some circumstances not to provide any context due to the oddity of using an empty pronoun completely out of the blue. In such cases, a minimum discourse will be given, but it will not affect how syntactic constraints are not to be overridden by discourse when licensing zero anaphora in Chinese.

3.1 An introduction

After the previous discussion in Chapter 2 about the notion of topic, we can now return to our central problem: the occurrences of empty pronouns in Chinese, their distribution, and how they are constrained. We will begin with the syntactic constraints, and then we will explore pragmatic conditions in Chapter 4. Because of the more flexible nature of discourse, the pragmatic constraints we will look at serve more like *tendencies* than inviolable *rules*. This way of constraining empty pronouns is modeled on Iida's (1996) analysis of the Japanese reflexive *zibun* 'self', in which she advocates a *conjunctive* binding theory which requires any instance of *zibun* 'self' to meet both syntactic conditions and discourse conditions. Adopting this conjunctive view, we suggest that empty pronouns be conditioned by both syntactic and pragmatic factors, with the former being rigorous and context-independent, and the latter flexible and context-dependent. The syntactic constraints to be introduced below incorporate a few other scholars' previous work (e.g., Chu 1993, 1998; J. Huang 1984, 1987, 1989, 1991; Iida 1996; C. Li 1985; Pan and Hu 2000) and our new observations. Before we start to discuss any syntactic constraints in detail, there is a prerequisite requirement to be laid out first: prior

to any proposed constraints, it is fundamental for an empty pronoun and its antecedent to be semantically compatible. That is to say, coindexation between an empty pronoun and its antecedent should obey semantic restrictions and accord with common sense or general world knowledge. For example, coindexation between an animate referent and an inanimate one will violate the semantic restriction on animacy, and similarly, an incompatible person or number feature between an empty pronoun and its antecedent will immediately result in ungrammaticality. If this prerequisite condition is not met first, the sentence in question will just fail grammatically without having to look into any other syntactic constraint. The following discussion of the syntactic behavior of empty pronouns will respect the aforementioned ground rule of semantic compatibility.

3.2 Syntactic constraints

The following subsections discuss various syntactic conditions of zero anaphora in Chinese, none of which are to be overridden by discourse or contextual factors.

3.2.1 No coindexation between an empty pronoun and its coargument

Iida (1996) proposes a constraint on empty pronouns in Japanese, which states that an empty pronoun cannot be coindexed with any of its coarguments (Iida 1996: 303). This constraint, we believe, could be seen as extended from Principle B of the Binding Theory (Chomsky 1980), also known as Disjoint Reference: a pronoun must be free in its governing domain. Iida's constraint concerning the forbidden coreference between an empty pronoun and its coargument can be indisputably generalized to Chinese as well, as evidenced by the following examples.

(167) a. *John$_i$ hen xihuan piping bieren dan [ta$_i$ conlai bu piping Ø$_i$].
very like criticize others but he ever not criticize ACC
'John$_i$ likes to criticize others a lot, but he$_i$ never criticizes (himself$_i$).'
b. *John$_i$ hen xihuan piping bieren dan [Ø$_i$ conlai bu piping Ø$_i$].
very like criticize others but NOM ever not criticize ACC
'John$_i$ likes to criticize others a lot, but (he$_i$) never criticizes (himself$_i$).'

c. John$_i$ zuo guo hen duo huai shi dan [tade$_i$ gege$_j$ conlai bu
do ASP very many bad things but his brother ever not
piping Ø$_i$].
criticize ACC
'John$_i$ has done many bad things, but his$_i$ brother never criticizes (him$_i$).'

From (167a) and (167b), we see that the empty object of the embedded verb, *piping* 'criticize', cannot be coindexed with its coargument, i.e., the embedded subject, whether it is overtly expressed or not. The ill-formedness of (167b) further indicates that an empty pronoun cannot be coindexed with a non-c-commanding coargument, either. The contrast between (167a, b) and (167c) illustrates that coindexation between an empty pronoun and its coargument is disallowed in a similar fashion to Principle B's regulation of overt pronouns. Iida's constraint on Japanese empty pronouns also holds in Chinese.

3.2.2 No empty pronoun allowed after preposition or in pivotal position

C. Li's (1985) study of anaphoric forms lists a series of conditions with regard to the use of full nouns, overt pronouns and empty pronouns, and she suggests that an empty pronoun can neither occur as the object of a *coverb*[28] (which is also commonly recognized as a mere preposition by many other scholars) nor in the *pivotal*[29] position.

[28] The term 'coverb' refers to a class of words that appear to function as prepositions in Mandarin Chinese. Typical coverbs include *gei* 'to/for', *gen* 'with', *cong* 'from', *dui* 'to/toward', and so on. Because coverbs seem to behave partly like verbs and partly like prepositions, the name 'coverb' was created to avoid classifying them as either verbs or prepositions (Li and Thompson 1989: 360). As Li and Thompson (1989: 360) point out, the reason coverbs have this binary status is that most modern-day coverbs, though functioning as prepositions nowadays, can be historically traced back to earlier stages of the language as verbs. Owing to the historical transition from verbs to prepositions, some of the coverbs are more like prepositions than others. For more discussion related to this issue, see Li and Thompson (1989) and Chang (1977).

[29] The term 'pivotal construction' was first introduced in pioneering work by Chao (1968). A pivotal construction consists of a succession of a verbal expression V_1, a nominal expression, and another verbal expression V_2, with the nominal expression serving simultaneously as the object of V_1 and subject of V_2, such as the pivotal status of *you* and *him*, respectively, in the following 'he asked you to leave' and 'I saw him write a letter'. Chao (1968: 125-6) notes that pre-pivotal verbs are of the 'cause to' type, e.g., request, let, advise, send, blame, prohibit, and so on. A direct object is always required to be present for a pivotal verb in Chinese (Li and Thompson 1989: 454). The Chinese pivotal construction is similar, but not limited, to object control observed in English. The former does not need infinitival 'to' to license the use of the second verb (V_2) in the sequence of verbs, and object-raising sentences can also be instances of a Chinese pivotal construction.

For example, in (168), the underlined *gei* 'give/to' may serve as a coverb, with a dual status as the verb 'give' and the preposition 'to', and the object of *gei* 'give/to' must be overtly expressed. Otherwise, the use of an empty pronoun after *gei* renders the sentence ungrammatical, as (168b) shows.

(168) a. John$_i$, wo song yi-yang liwu [gei ta$_i$].
I give one-CL present give/to him
'John$_i$, I gave a present to him$_i$.'
b. *John$_i$, wo song yi-yang liwu [gei Ø$_i$].
I give one-CL present give/to ACC
'John$_i$, I gave a present to (him$_i$).'

Similarly, in (169), the object following the underlined, patient-promoting coverb *ba*[30] must not be left unexpressed no matter how clear it is for the omitted object to naturally refer to the overt topic, *zhe-ba yizi* 'this chair', in the context. (The debate between the two terminologies, coverb versus preposition, is immaterial to our study here. However, since a coverb is partly a preposition, following Occam's razor, we will adopt the latter terminology for the sake of simplicity. Otherwise, we would need to impose two

[30] The study of the coverb (or simply preposition for some researchers) *ba* has been one of the most extensively discussed topics in the literature of Chinese grammar. The very complex properties of *ba* have led many grammarians to tackle the *ba* construction from various angles, and no consensus has yet been reached. For purposes of the discussion of empty pronouns in this study, we will not present any detailed discussion of *ba* structures. We will simply assume that the *ba* construction in Chinese is conventionally considered a patient-topicalization or patient-promotion device in that it has the effect of moving forward (or promoting) a post-verbal patient argument (e.g., a direct object) and placing it after *ba*, which is always pre-verbal. For instance, (ia) and (ib) convey essentially the same meaning but with a slightly different focus. The direct object of *he* 'drink', i.e., *jiu* 'wine', is placed immediately after *ba* and before the verb in (ia), while (ib) shows the canonical SVO word order in Chinese.

(i) a. Ta ba jiu he le.
he BA wine drink ASP
'He drank the wine.'
b. Ta he le jiu.
he drink ASP wine
'He drank the wine.'

This oversimplified view of the *ba* construction is by no means complete, but it should be enough to serve the purpose of our study on empty pronouns here. For further discussion of the complicated *ba* construction, see Hsueh (1987) for a topic-comment analysis, Li (2001) for a syntax-oriented view, or Szeto (1988) and Yong (1993) for an aspectual treatment. Alternatively, Bender (2000) presents an LFG account of treating *ba* as a verb.

separate rules, one disallowing empty objects following a coverb and the other forbidding them following a preposition. The distinction of coverbs and prepositions will not be addressed here, and the latter term is preferred in this thesis for the purpose of our study on empty pronouns). From (168) and (169), we see that the use of an empty pronoun following a preposition results in ungrammaticality.

(169) a. Zhe-ba yizi$_i$, John [ba ta_i] nong-huai le.
this-CL chair BA it make-spoil ASP
'(As for) this chair$_i$, John broke it$_i$.'
b. *Zhe-ba yizi$_i$, John [ba $Ø_i$] nong-huai le.
this-CL chair BA ACC make-spoil ASP
'(As for) this chair$_i$, John broke (it$_i$).'

Chinese pivotal constructions, similar to object control observed in English, involve a noun phrase that is simultaneously the object of the first verb and the subject of the second verb. For instance, the pronoun *ta* 'him' in (170) serves as the object of the first verb *jiao* 'ask' and the subject of the second verb *xie* 'write', acting like a pivot connecting two sides (hence the name 'pivotal' construction). The same analysis applies to *ta* 'him' in (171), which is the object of the first sense verb[31] *kandao* 'see' and the subject of *manpao* 'jog'. As C. Li (1985) points out, a pivotally used pronoun must not be empty, as the ill-formed alternatives with an empty pronoun in (170) and (171) exemplify. The syntactic constraint proposed by C. Li (1985) on empty pronouns is valid – an empty pronoun is forbidden from following a preposition and from the pivotal position.

(170) John$_i$, wo jiao ta_i/*$Ø_i$ xie gongke.
I ask him write homework
'(As for) John$_i$, I asked him$_i$/*(him$_i$) to write the homework.'
(171) John$_i$, wo gang-gang kandao ta_i/*$Ø_i$ zai gongyuan manpao.
I just.now see him in park jog
'(As for) John$_i$, I saw him$_i$/*(him$_i$) jogging in the park just now.'

[31] Other sense verbs include to *hear*, to *feel*, to *smell*, to *watch*, etc.

Having shown that the constraint proposed by C. Li (1985) indeed applies to empty pronouns in Chinese, we would like to add that the aforementioned pivotal construction is not unique to the Chinese language, and a more commonly known construction – object-control – can be seen as a type of pivotal structure, as in (170). (Some of the object-control verbs in Chinese are listed in Footnote 29). As the following example in (172) demonstrates, C. Li's original constraint on empty pronouns in the pivotal construction also includes sentences involving object-raising verbs. That is, an empty pronoun cannot occur as the object of an object-raising verb, e.g., *xiangxin* 'believe' in (172).

(172) John$_i$, wo xiangxin ta$_i$/*Ø$_i$ yijing dao meiguo le.
I believe him already arrive America ASP
'(As for) John$_i$, I believe him$_i$/*(him$_i$) to have already arrived in America.'

Based on C. Li's observation, we conclude that the second syntactic constraint for Chinese empty pronouns is revised as follows: an empty pronoun cannot appear after a preposition (or a coverb, in C. Li's (1985) terms) or occur in a pivotal position (e.g., after a control verb, a raising verb, or a sense verb).

3.2.3 Empty subordinate subjects must be anteceded by arguments

The third syntactic constraint concerns the restriction of empty subordinate subject pronouns, e.g., those occurring in relative clauses, complement clauses, or adverbial adjuncts. Before we start, we should stress that since empty pronouns have no realization at the level of *c*-structure but do play a role at *f*-structure (and *i*-structure), we will assume that any syntactic operation regarding empty pronouns, such as binding, will be at work at *f*-structure. Therefore, for example, while the traditional notion of binding is associated with coindexation and c-commanding in a tree structure, the kind of binding involving empty pronouns will involve coindexation and f-commanding relations at *f*-structure instead, because empty pronouns are absent at *c*-structure and c-commanding cannot

apply. *F-command*, on the other hand, is 'a relation between *f*-structures analogous to the c-command relation defined on trees' (Dalrymple 2001: 159). The definition of f-command[32] is given as follows: an *f*-structure α f-commands an *f*-structure β if and only if α does not contain β and every *f*-structure that contains α also contains β (Falk 2001).

Returning to our third syntactic constraint on empty pronouns, we propose that an empty subordinate subject pronoun must be bound by an argument function, i.e., coindexed with and f-commanded by an argument function like SUBJ or OBJ (or simply A-bound, in the terminology of P&P Theory). In other words, this rule forbids the use of an empty subject in a subordinate clause if it is bound by a non-argument function at *f*-structure, which is most likely to be TOP especially when TOP and SUBJ differ. Nevertheless, this constraint does not rule out accidental binding between an empty subordinate subject and a non-argument topic if the topic happens to be identical to the subject. In this case, what licenses this empty pronoun is the subject argument, not the topic. To illustrate the effect of this constraint, we can see that the use of an empty subject pronoun in the bracketed relative clause in (173a) is well-formed, as it is bound by an argument, i.e., the matrix subject. (173b) further shows that the embedded argument, that is, the underlined *ta* 'he', can also antecede an empty subordinate subject, as long as the former is an f-commanding argument regardless of whether it appears in the matrix clause or not. Note that because the underlined subject *ta* 'him' in (173b) can ambiguously refer to either *John* or *John's brother*, the empty subject in question is accordingly anteceded by either referent. (173b) and (173c) are the two possible readings of the same sentence.

[32] To illustrate f-command, we adopt an example from Dalrymple (2001: 159). In (i) and (ii) below, the *f*-structure labeled *f* f-commands the *f*-structure labeled *g*. Conversely, *g* also f-commands *f* in (i), but not in (ii).

(i) $\begin{bmatrix} \text{SUBJ} & f[\ ] \\ \text{OBJ} & g[\ ] \end{bmatrix}$

(ii) $\begin{bmatrix} \text{SUBJ} & f[\ ] \\ \text{COMP} & [\text{SUBJ} \quad g[\ ]\ ] \end{bmatrix}$

(173) a. John$_i$, ta$_i$ zhongyu zhaodao yi-jian [Ø$_i$ keyi nian __$_j$ de]
he finally find one-CL NOM can study ACC RLTV
xuexiao$_j$.
school
'John$_i$, he$_i$ finally found a school where (he$_i$) can go.'

b. John$_i$ de gege$_j$ shuo [ta$_i$ zhongyu zhaodao yi-jian [Ø$_i$ keyi nian
GEN brother say he finally find one-CL NOM can study
__$_k$ de] xuexiao$_k$].
ACC RLTV school
'John$_i$'s brother$_j$ said that he$_i$ finally found a school where (he$_i$) can go.'

c. John$_i$ de gege$_j$ shuo [ta$_j$ zhongyu zhaodao yi-jian [Ø$_j$ keyi nian
GEN brother say he finally find one-CL NOM can study
__$_k$ de] xuexiao$_k$].
ACC RLTV school
'John$_i$'s brother$_j$ said that he$_j$ finally found a school where (he$_j$) can go.'

(174a) and (174b) illustrate that in Chinese, an empty subject pronoun in an adjunct subordinate clause has to refer to the matrix subject, in the same manner English restricts an omitted subject in an adjunct clause to be coindexed with the matrix subject in a sentence like *John loves to sing while taking a shower*.

(174) a. [Yinwei zuo-wan Ø$_i$ mei shui hao], John$_i$ jintian kanqilai
because last-night NOM not sleep well today look
hen lei.
very tired
'Because (he$_i$) didn't sleep well last night, John$_i$ looks very tired today.'

b. [The teacher$_i$ told John$_j$ that he$_i$ would inform his$_j$ parents$_k$ of his$_j$ bad behavior at school …]
[Yi Ø$_{j/*i/*k}$ huidao jia], John$_j$ mashang jiu kaishi xie gongke.
once NOM return home immediately just start write homework
'As soon as (he$_{j/*i/*k}$) returned home, John$_j$ just started to do the homework immediately.'

Given the above rule, an overt pronoun must be used if an empty subordinate subject is bound by a non-argument, e.g., by a topic that differs from the subject in the same

sentence. For instance, the empty subject pronoun in the bracketed relative clause of (175) is disallowed because it is bound by a non-argument, that is, the topic *John*. The position in question must be filled by an overt pronoun instead in order to maintain the grammaticality of the sentence.

(175) John$_i$, wo nian le bu shao [ta$_i$/*Ø$_i$ xie __$_j$ de] shu$_j$.
I read ASP not few he write ACC RLTV books
'(Speaking of) John$_i$, I have read quite a few books that he$_i$/*(he$_i$) wrote.'

The reason behind this syntactic constraint could be that the syntactic notion of an argument (subject or object) is hierarchically prior to the discourse notion of a topic in being a proper antecedent of an empty subject pronoun in a subordinate construction. The fact that the f-command relation must hold between an empty subordinate subject pronoun and its antecedent can be further evidenced by the following example in (176). The context of (176) is purposely designed so that the reference of the empty subject pronoun should have, most naturally, been the non-f-commanding topic *John*, but the syntactic restriction that we have proposed forbids this would-have-been-pragmatically-natural reading. Instead, the only available reading is marginal and pragmatically odd, that *John's mother bought many toys that she dreams to own* (while it is John's birthday tomorrow), because *his mother* is the subject, an argument, and the only possible antecedent for the empty subordinate pronoun. Such coindexation is syntactically acceptable but semantically odd, as indicated by the # sign.

(176) [It is John$_i$'s birthday tomorrow …]
Tade$_i$ mama$_j$ mai le hen duo [Ø$_{*i/\#j}$ meng-mei-yi-qiu __$_k$ de]
his mother buy ASP very many NOM dream.to.have ACC RLTV
wanju$_k$.
toys
'His mother$_j$ bought many toys which (he$_{*i}$/she$_{\#j}$) dreams to own.'

From the above discussion, we have seen how an empty subject of a relative clause is subject to this binding constraint. Now let us turn to empty subjects in complement clauses, which, according to our proposal, must be anteceded by an f-commanding argument function; otherwise, ungrammaticality arises and an overt pronoun should be used instead. Therefore in (177a), the non-f-commanding topic, *John*, cannot serve as the antecedent of the empty subject, even though this reading would have been natural, given *John*'s topical status and the subsequent comment clause. In this case, it is syntax that rules out this possibility. The only available reading in (177a) is for the empty subject to be anteceded by the SUBJ argument, *John's sister*. In contrast, the sentence in (177b) is ambiguous between two readings because the underlined overt subject pronoun is not restricted by the aforementioned constraint and can be coindexed with either antecedent.

(177) a. John$_i$, tade$_i$ jiejie$_j$ shuo [Ø$_{j/*i}$ xihuanshang le yige gaozhong laoshi].
his sister say NOM like ASP one high.school teacher
'(Speaking of) John$_i$, his sister$_j$ said that (she$_j$/*he$_i$) fell in love with a high school teacher.'

b. John$_i$, tade$_i$ jiejie$_j$ shuo [ta$_{i/j}$ xihuanshang le yige gaozhong laoshi].
his sister say he/she like ASP one high.school teacher
'(Speaking of) John$_i$, his sister$_j$ said that he$_i$/she$_j$ fell in love with a high school teacher.'

The scenario could become complicated when there is more than one occurrence of an f-commanding argument for an empty subordinate subject to choose its antecedent from. The unmarked reading of (178a) is for the empty subject pronoun in the bracketed relative clause to refer to *Bill*; coindexation with *John* might be syntactically plausible but semantically ruled out. In (178b), the empty pronoun can refer to either of the f-commanding arguments, *the speaker* or *John*, so (178b) is two-ways ambiguous.

(178) a. John$_i$ gaosu Bill$_j$ yi-jian [Ø$_{j/\#i}$ juedui bu neng xielou __$_k$ de]
tell one-CL NOM absolutely not can reveal ACC RLTV
xiaoxi$_k$.
news
'John$_i$ told Bill$_j$ some news which (he$_{j/\#i}$) could never reveal.'

b. Wo$_i$ pian John$_j$ [Ø$_{i/j}$ mingtian yao qu meiguo].
I lie NOM tomorrow will go America
'I lied to John$_i$ that (I/he$_i$) would be going to America tomorrow.'

Specifically, because *John* and *Bill* in (178a) are both syntactically legitimate antecedents to license the use of the empty subordinate subject, it is semantics that rules out the marked meaning in which the empty pronoun refers to *John*. In contrast, the two available readings in (178b) are both very natural and highly plausible. The examples above show that when there is more than one f-commanding argument to license an empty subordinate subject pronoun, the choice of the antecedent would then be determined by semantics/pragmatics or real-world knowledge, but not syntax, as long as the syntactic condition has been satisfied. As far as (178) is concerned, it is general knowledge and semantic factors that decide which f-commanding argument(s) is to antecede the empty subject pronoun in question, once the syntactic requirement has been met. Note that the matrix verb *pian* 'lie' in (178b) is transitive in Chinese, which renders *John* the object argument of the matrix verb rather than the object in an oblique phrase as it would be in English, e.g., *I lied to John*. To sum up, the above discussion has led us to propose that an empty subordinate subject pronoun must be bound by an f-commanding argument, and it may be semantics or general knowledge that determines what potential interpretations are when there is more than one f-commanding argument antecedents in the sentence. Importantly, note that the proposed constraint does *not* have the converse implication, namely, as long as a subordinate subject pronoun has a legitimate argument to be its potential antecedent, it can therefore be null. This converse statement is a false assumption not implicated by our constraint. As we can see from the contrast of the two alternatives in (179) below, even though the subject pronoun in the relative clause has the matrix subject to be its legitimate antecedent, the use of an empty form is not permitted.

(179) Wo$_i$ du guo hen duo [wo/*Ø$_i$ xie __$_j$ de] shu$_j$.
I read ASP very many I write ACC RLTV books
'I have read very many books that I/*(I) wrote.'

The example in (179) demonstrates that there may be micro-linguistic factors that rule out the use of an empty pronoun even though the sentence does not violate any syntactic, semantic, and pragmatic constraints. There does not seem to be any regular pattern, and to capture these micro-linguistic factors will be an objective for future research. Our current proposal describes (i) the rules zero anaphora must obey in order to be permitted (e.g., an empty subordinate subject being anteceded by an f-commanding argument), and (ii) under what circumstances the use of zero anaphora must be ruled out (e.g., when it appears after a preposition). But it cannot account for the reason every ill-formed sentence involving zero anaphora is bad, such as the ungrammatical version in (179), whose structure parallels that of its grammatical counterpart with an overt pronoun and in which the null pronoun is anteceded by a legitimate argument. We can only assume, at this stage, that some micro-linguistic factors are at work which render the use of the empty pronoun in (179) unacceptable.

Given the constraint on empty subordinate subjects, the same requirement of being bound by an argument does not seem to apply to empty subordinate object pronouns, which, in fact, behave in quite an opposite manner. To illustrate, it is clear that the empty object pronoun in the bracketed relative clause in (180) refers to the topic, which is a non-argument antecedent, and the sentence in (180) remains well-formed.

(180) Zhe-bu dianying$_{i,}$, wo mei jian guo [__$_j$ bu tuijian Ø$_i$ de] ren$_j$.
this-CL movie I not see ASP NOM not recommend ACC RLTV person
'(Speaking of) this movie$_i$, I haven't seen anyone who wouldn't recommend (it$_i$).'

This suggests that empty object pronouns in relative clauses seem to escape the aforementioned binding condition. So do empty objects in complement clauses, as the contrast in the overt/empty alternation in (181) shows. In (181), coindexation between the empty object and the subject argument is disallowed, but if the empty object is bound by a topic non-argument (i.e., *John*, as indicated by the subscript *i*), the sentence would be acceptable.

(181) Shuodao $John_{i,}$, $Bill_j$ zhidao [ni hen jiu mei jian dao $ta_i/Ø_{i/*j/*k}$ speaking.of know you very long not see ASP him le].
ASP
'Speaking of $John_i$, $Bill_j$ knows that you haven't seen him_i/($him_{i/*j/*k}$) for a very long time.'

Viewed this way, we then observe a subject-object asymmetry with regard to how subject and object are permitted to be null in subordinate clauses: while empty subordinate subject pronouns must be bound (i.e., coindexed with and f-commanded) by an argument, empty subordinate object pronouns need not be so. The example in (182) below shows that the empty subordinate object refers to the matrix subject, which simultaneously fills the SUB and TOP functions. Given the contrast between (180-1) and (182), we conclude that an empty subordinate object need not be bound by an argument; what crucially licenses such use is a topical element (the TOP function) which may or may not be an argument, such as SUBJ, at the same time. Whether or not the antecedent is an argument is irrelevant to the licensing of empty objects. We will discuss more in the next part.

(182) Zhe-zhi gou_i zhongyu zhaodao [$___j$ yuanyi lingyang $Ø_i$ de] $renjia_j$.
this-CL dog finally find NOM willing adopt ACC RLTV family
'This dog_i finally found a family who is willing to keep (it_i).'

3.2.4 Empty object pronouns must be anteceded by topics

In view of the previous examples in (180) and (181), empty object pronouns do not appear to be restricted in the sense that they need not be bound by arguments. Some empty objects are bound by a topical non-argument, as in (180) and (181), whilst some by a topical matrix subject, as in (182) and (183).

(183) Zhe za-cao-cong-sheng de huayuan$_i$ daibiao [yijing hen jiu mei you ren zhengli Ø$_i$ le].
this wild-grass-bush-grow ADJV garden mean already very long not have people tend ACC ASP
'This unweeded garden$_i$ suggests that no one has looked after (it$_i$) for a very long time.'

Moreover, the following examples in (184) and (185) indicate that an object pronoun may be empty if its antecedent is topical. In (184) and (185), the simple sentence (a) and the complex sentence (b) are two possible utterances to follow immediately the initial, bracketed sentences, in which the topics, *this DVD box set* and *Elvis*, respectively, are introduced.

(184) [This DVD box set$_i$ is a must-have!]
a. Danshi John que bu xiang shouji Ø$_i$.
but however not want collect ACC
'But John does not want to collect (it$_i$).'
b. Danshi John shuo [Bill bu xiang shouji Ø$_i$].
but say not want collect ACC
'But John said that Bill did not want to collect (it$_i$).'

(185) [Elvis$_i$ is *so* famous …]
a. Ni zenme neng mei ting-shuo guo Ø$_i$?
you how can not hear-say ASP ACC
'How could you never hear of (him$_i$)?'
b. Danshi John shuo [Bill jingran mei ting-shuo guo Ø$_i$].
but say surprisingly not hear-say ASP ACC
'But John said that surprisingly Bill never heard of (him$_i$).'

Our examples in (182) and (183) indicate that an object pronoun may be empty even when it is bound by an argument in the same sentence, as long as this antecedent is topical. This line of analysis contradicts J. Huang's (1984, 1987, 1989, 1991) empty-topic hypothesis, in which he argues that empty objects are not genuine pronouns (*pro*s) but variables, because they can only be bound by topics but *never* by arguments. A context-less sentence like the one in (186) has long been J. Huang's crucial evidence

against the pronominal status of an empty object, owing to the unavailable coindexation between the empty object and the matrix subject, as (186a) demonstrates. The employment of an overt pronoun in the same position, as in (186b), permits the embedded object to be coindexed with the matrix subject or some other topic previously mentioned in the discourse.

(186) a. John$_i$ shuo [Bill mei kanjian Ø$_{*i/j}$].
say not see ACC
'John said that Bill didn't see (someone who cannot be John).'
b. John$_i$ shuo [Bill mei kanjian ta$_{i/j}$].
say not see him
'John$_i$ said that Bill didn't see him$_i$ or someone else in the previous discourse.'

While we agree with J. Huang's judgment on the possible interpretations of (186a) and (186b), we disagree with his approach of providing no context for a sentence like (186a) in order to maintain purposely a 'pragmatically neutral environment'. The definition of a pragmatically neutral environment is not clearly given by Huang, but from his point of view, our example in (182) would be regarded as *not* occurring in a pragmatically neutral environment due to the use of the embedded verb, *lingyang* 'adopt', which, according to him, enables the embedded empty object to be 'pragmatically inferred' to refer to the matrix subject and allows such coindexation. Therefore, a sentence like (182) should not serve as a counterexample to his empty-object-as-variable analysis. We would like to dispute Huang's view for the following two reasons. Firstly, his definition of 'pragmatic neutrality' is unclear; does a pragmatically neutral context mean that a sentence should be free of any context? When it comes to the use of empty pronouns, a certain amount of context is required in order for hearers to process the information. J. Huang's stance of assuming no context whatsoever is also a special kind of context itself, and this is not necessarily a natural one. Secondly, we disagree that the lexical verb *lingyang* 'adopt' strongly favors or disfavors a certain type of referent and therefore leads to a pragmatically non-neutral environment. Neither did J. Huang explain

how this particular verb is pragmatically biased. We would like to stress that if a constraint is syntactically sound, it should be absolute and not be overridden by context or lexical choices anyway, unlike the problem with J. Huang's empty-object-as-variable analysis, which denies the pronominality of empty objects in Chinese but faces counterexamples.

As discussed, there are at least two problems with a context-less analysis of (186a) if we assume J. Huang's hypothesis. First of all, having no context at all is itself a type of context. It is not necessarily a natural or pragmatically neutral one, let alone a superior one. Secondly, it is true that some sentences involving the use of empty pronouns might not require any context and can still be interpreted without confusion. However, while we understand Huang's intention of providing as little context as possible in order to eliminate any effect context might have on the interpretation of a sentence, for a sentence like (186a) to occur naturally, a context should be obligatorily present, as it would make little sense and confuse hearers if (186a) is uttered out of the blue. A response like 'John said Bill didn't see what?' would be highly expected to follow immediately the context-less usage of the utterance in (186a). Recognizing this extreme unnaturalness and implausibility of (186a) used without any context, we argue against Huang's empty-object-as-variable hypothesis and instead, propose that the use of an object pronoun may be, but need not be, empty if its antecedent, as an argument or not, is topical. While J. Huang rejects the pronominality of empty objects, we stress that what syntactically licenses an empty object pronoun is the topicality of the antecedent, be it an argument or not. Returning to the discussion of the sentence in (186a), we demonstrate below in (187) to (190) the possible interpretations of the embedded empty object in question. As discussed, it is highly unnatural, if not impossible, for the sentence in (186a) to be uttered with no context, and it should not be considered evidence against the pronominal status of an empty object. Therefore, by giving the sentence in question a minimal appropriate context, the embedded object can refer to various sorts of topical referents from animate to situational, and, crucially, coindexation between the embedded empty object and the matrix subject is definitely feasible, as exemplified by the sentence in (190), contra J. Huang's (1984, 1987, 1989, 1991) hypothesis that denies the

pronominal status of an empty object. To sum up, our fourth syntactic constraint states that an object pronoun may be, but need not be, empty as long as its antecedent is topical (either as an argument or a non-argument).

(187) [There was a huge snake$_i$ on the street …]
Danshi John shuo [Bill mei kanjian Ø$_i$].
but say not see ACC
'But John said that Bill didn't see (it$_i$).'

(188) [A famous band$_i$ came to school today …]
Danshi John shuo [Bill mei kanjian Ø$_i$].
but say not see ACC
'But John said that Bill didn't see (them$_i$).'

(189) [John$_i$ just fell$_j$ over on the street …]
Danshi ta$_i$ shuo [Bill mei kanjian Ø$_j$].
but he say not see ACC
'But he$_i$ said that Bill didn't see (it, i.e., John's falling over).'

(190) [John$_i$ was waving$_j$ excitedly at Bill across the street …]
Danshi John$_i$ shuo [Bill haishi mei kanjian Ø$_{i/j}$].
but say still not see ACC
'But John$_i$ said that Bill still didn't see (him$_i$/the waving$_j$).'

After examining empty objects in simple and complement clauses, we predict that empty objects appearing in adverbial adjunct clauses would behave in the same manner: they must have topical antecedents. Take (191) for example. In (191a), the empty object of *xihaun* 'like' in the adjunct clause refers back to the topic under discussion, *Taiwan*. In (191b), the empty objects of *ma* 'scold' and *da* 'hit' appear before their antecedent, *women* 'we', which is topical. Both examples are taken from the Sinica Corpus.[33]

(191) a. Taiwan$_i$ jiu zheme xiao. Ni bu xihuan Ø$_i$, jiu yimin dao guowai qu.
just this small you not like ACC just migrate to abroad go
'Taiwan$_i$ is just very small. If you don't like (it$_i$), just move abroad.'

[33] *Sinica Corpus* is shortened from *Academia Sinica Balanced Corpus of Modern Chinese*, maintained by the Institute of Information Science and CKIP group in Academia Sinica, Taiwan. The size of the corpus is up to five million words and the Sinica Corpus is commonly used in analyses of modern Chinese language.

b. Dang fumu hen yenlide ma $Ø_i$ huozhe da $Ø_i$ de shihou,
when parents very strictly scold ACC or hit ACC of time
$women_i$ keneng-hui hen haipa.
we might-would very scared
'When parents strictly scold (us) or hit (us)[34], we might be very scared.'

(Sinica Corpus)

So far our examples concerning the fourth syntactic constraint only consider the distribution of empty objects occurring in simple clauses, complement clauses, and adjunct adverbials. We have not yet demonstrated how empty objects in relative clauses behave, because they seem to be more puzzling than other empty subordinate object pronouns. Adapted from J. Huang's (1984) example, (192a) and (192b) are two hypotheses of how the utterance in (192c) can be analyzed with regard to the position of an empty category inside the relative clause. The differences between (192a) and (192b) only lie in the grammaticality and the hypothetical position of the empty pronoun inside the relative clause: the former is in subject position and the latter in object position. As the examples show, (192c) is unambiguous and only has one interpretation as (192a); the interpretation of (192b) is unavailable. The only way to keep the intended reading in (192b) is to replace the embedded empty object with the overt pronoun *ta* 'him'. The brackets indicate the boundary of the relative clause.

(192) a. $John_i$ zhongyu zhao dao yige [$Ø_i$ hen xihuan $___j$ de] ren_j.
finally find ASP one NOM very like ACC RLTV person
'$John_i$ finally found a person whom (he_i) likes very much.'

b. *$John_i$ zhongyu zhao dao yige [$___j$ hen xihuan $Ø_i$ de] ren_j.
finally find ASP one NOM very like ACC RLTV person
'$John_i$ finally found a person who likes (him_i) very much.'

c. John zhongyu zhao dao yige [hen xihuan de] ren.
finally find ASP one very like RLTV person
'$John_i$ finally found a person whom (he_i) likes very much.'

[34] As the majority of empty pronouns exemplified in this study tend to be third person, here we see an instance of a first person empty pronoun in use. Erteschik-Shir (2007: 17-8) points out that first person and second person pronouns, as speakers and hearers, are 'permanently available topics' in a conversation even without previous mention.

As the intricate examples in (192) have illustrated, the question arises as to why the use of an empty object pronoun in the relative clause is forbidden as such. There does not seem to be any restriction as to why an empty object pronoun cannot occur in the relative clause of (192b) like it did previously in (180), (182) and (183). In particular, we would also expect the sentences in (192) to pattern with those in (193) below, in which an empty pronoun can appear in the embedded subject position, shown in (193a), as well as in the embedded object position, shown in (193b), resulting in two-way ambiguity of the same utterance.

(193) a. John$_i$ zhao-bu-dao yige [Ø$_i$ keyi zhaogu ___$_j$ de] ren$_j$.
find-not-ASP one NOM can look.after ACC RLTV person
'John$_i$ can't find a person whom (he$_i$) can look after.'
b. John$_i$ zhao-bu-dao yige [___$_j$ keyi zhaogu Ø$_{i/k}$ de] ren$_j$.
find-not-ASP one NOM can look.after ACC RLTV person
'John$_i$ can't find a person who can look after (him$_i$ or someone else).'

The use of the empty subject pronoun in the relative clause of (193a) accords with our previous constraint which states that an empty subordinate subject must be bound by an argument. The use of the empty object pronoun in (193b) also follows our proposal by being anteceded by a topical subject, or an aforementioned topic that is not present. Then why is the option of employing an empty object in (192b) blocked, while its sentence structure parallels that of (193b) and only the embedded verbs differ? In addition, the unavailable coindexation in (192b) also provides J. Huang's empty-object-as-variable analysis another piece of evidence in rejecting the pronominal status of an empty object, which, in his account, must not be bound by an argument, but by a topic non-argument as a variable only. Here again we disagree with J. Huang's approach of treating an empty object as a variable bound by a topic that cannot be an argument, as the sentence in (193b) clearly serves as a counterexample. However, we are still left with a puzzle as to why the use of an empty object in (192b) is disallowed, leaving (192b) to be ungrammatical, especially when (192b) is, in fact, syntactically parallel to (193b). To account for this problem, we propose the bracketed elements in sentences in (192) to be *reanalyzed* as not

involving a relative construction, but merely a noun-modifying adjectival phrase that contains no gap. That is, (192c) can be reanalyzed as (194) below, in which the same string of words, bracketed, is treated as an adjectival phrase (AdjP) modifying the following head noun *ren* 'person'. After the reanalysis, the relative clause marker *de* in (192) has now become an adjectival marker in (194), which is responsible for the adjectival status of the bracketed elements. The sentence meanings of the two sentences (before and after the reanalysis) are perceived as identical in Chinese, despite their slightly different English translations in order to reflect their structural difference.

(194) John zhongyu zhao dao yige [hen xihuan de]$_{AdjP}$ ren.
finally find ASP one very like ADJV person
'John finally found a liked-very-much-by-him person (= John finally found a person that he likes very much).'

This view of turning a relative clause into an adjectival phrase is by no means ad hoc due to the polysemous, multifunctional use of the particle *de* in Chinese, which has in turn led to the long-term debate about the relative construction in Chinese linguistics because its existence is not always recognized by grammarians (e.g., Chao 1968; Paul 2005). As cited in Wong (2005), some grammarians (e.g., Li and Thompson 1989; Aoun and Li 1993; Chiu 1993; Wu 2000) assume that a nominalization consisting of a verb phrase, as in (195), or a clause, as in (196), immediately followed by the particle *de* as a relativizer (RLTV), is to be analyzed as a relative construction if the head noun being modified by the verb phrase or the clause corefers with a gap (indicated by the underscore) inside the modifying verb phrase or clause. The examples in (195) and (196) demonstrate such coreference between the gap and the modified head noun.

(195) [__$_i$ chuan hong yi de] nushi$_i$.
NOM wear red clothes RLTV lady
'The lady who wears red clothes.'

(196) Ta jiushi [wo shang-ci ti $___i$ de] nage ren_i.
He is I last-time mention ACC RLTV that person
'He is the person whom I mentioned last time.'

On the other hand, those scholars who do not recognize the existence of the relative construction in Chinese reject it for at least the following two reasons. First of all, cited from Wong (2005), Chao (1968: 285) argues that the particle *de* is simply a marker of explicit modification that occurs between the modifier and the modified (e.g., a noun phrase). *De* can be multifunctional in being a genitive marker, cf. (197a); an adjectival marker, cf. (197b); or a nominalizer (NLZ), cf. (197c), in which *de* turns the verb phrase *mai cai* 'sell vegetables' into a noun phrase – someone who sells vegetables. The multiple purposes of *de* render it not a proper counterpart of the unifunctional relative pronouns in English; it is not always easy to determine the function of *de* in *de* constructions.

(197) a. [Wo de]$_{Genitive}$ shu.
I GEN book
'My book.'
b. Zhe-ben [hong-se de]$_{AdjP}$ shu.
this-CL red-color ADJV book
'This red book.'
c. [Mai cai de]$_{NP}$ lai le ma?
sell vegetable NLZ come ASP Q
'Has the vegetable-selling person come yet?'

The second reason to reject the existence of the relative construction in Chinese is that according to Chao (1968), some clauses that appear to be analyzed as a relative structure in Chinese contain no gap to which the head noun can appropriately corefer. For instance, in (198) below, nothing in the bracketed clause seems to be relativized because both arguments of the two-place predicate, *shang* 'appreciate/watch', are present. The bracketed clause is gapless.

(198) [Wo shang niao de] wang-yuan-jing.
I appreciate bird RLTV binoculars
'The binoculars (with) which I watch birds.'

We will not go into detail with regard to how different grammarians argue for the nature of the relative construction in Chinese, especially when no consensus has been reached. This brief introduction simply serves the purpose of presenting the unsettled, controversial status of relative constructions in Chinese as well as showing that our view of analyzing some of them as adjectival modifiers is certainly not ad hoc but well-motivated. Having demonstrated the contrastive views of analyzing relative constructions in Chinese, we can now start to analyze how the bracketed adjectival modifier in (194) can be derived from its relativized counterpart in (192). According to Ross (1983), even though the particle *de* seems to appear in various structures and denotes various functions (e.g., nominal, adjective, etc.), in most cases, the relationship between the phrase (or clause) where *de* occurs and the head noun immediately following *de* is one of modification. Namely, the *de* phrase (or clause) modifies the head noun. Ross (1983) also crucially points out that when *de* appears before a noun phrase and after a stative verb,[35] as in (199), the English translation for such constructions involves adjectival modification. On the other hand, if *de* appears before a noun phrase and after a non-stative verb, as (200) shows, the English translation involves a relative clause instead.

[35] Even though Ross (1983) does not explicitly state her definition of stative verbs, according to her example, we believe that by stative verbs she means adjectival verbs, as adjectival verbs are one common type of verb phrase in Mandarin Chinese (Li and Thompson 1989: 141-2). Namely, most adjectives in Chinese may act as verbal predicates, as (i) instantiates. *Tian* 'sweet' in (ia) and *hao* 'good' in (ib) are adjectives which behave like verbal predicates.

(i) a. Zhe ping-guo hen tian.
this apple very sweet
'This apple is very sweet.'
b. Tianqi zhen hao.
weather really good
'The weather is really good.'

(199) [Gao de]$_{AdjP}$ shu.
tall ADJV trees
'Tall trees'

(200) [__$_i$ zai zuihou guantou fangqi de]$_{Relative\ Clause}$ ren$_i$.
NOM at last time give.up RLTV person
'People who gave up at the last minute.'

As stated in Footnote 35, Ross seems to treat adjectival predicates as stative verbs. Cheng (1986) also remarks that 'the distinction between adjectives and verbs in Mandarin is not clear cut' (Cheng 1986: 317), and the construction of AdjP/VP + *de* can be viewed as a modifier modifying the noun to appear after *de*. Returning to our previous example (192), repeated below as (201), and (194), repeated below as (202), (201) is the original sentence, analyzed as involving a relative structure and an empty pronoun, while (202) is the later revised sentence hypothesized to be associated with an adjectival modifier and containing no empty category. The *de*s in these two examples are homophonous but functionally distinctive.

(201) a. John$_i$ zhongyu zhao dao yige [Ø$_i$ hen xihuan __$_j$ de] ren$_j$.
finally find ASP one NOM very like ACC RLTV person
'John$_i$ finally found a person whom (he$_i$) likes very much.'

b. *John$_i$ zhongyu zhao dao yige [__$_j$ hen xihuan Ø$_i$ de] ren$_j$.
finally find ASP one NOM very like ACC RLTV person
'John$_i$ finally found a person who likes (him$_i$) very much.'

(202) John zhongyu zhao dao yige [hen xihuan de]$_{AdjP}$ ren.
finally find ASP one very like ADJV person
'John finally found a liked-very-much-by-him person (= John finally found a person that he likes very much).'

There is no doubt that *xihuan* 'to like' is a verb, just like in English, but we would like to postulate that in the case of (202), in which no argument of *xihuan* 'like' is present and *xihuan* 'like' precedes the particle *de*, this sequence of words – *xihuan de* – actually *can* function as an adjectival modifier to mean that the following head noun has the properties of 'being liked', though lacking an appropriate English translation. *Hen*

'very/very much' appearing in (201) and (202) can modify a verb, an adjective, or an adverb, which conforms to our treatment of *xihuan de* in (202) as an adjectival modifier. In the same vein, the following stative verbs such as *taoyen* 'resent' and *zhenxi* 'cherish' in (203) and (204) can be analyzed alike. For concreteness, if we look at the sentences in (203a) and (204), which are the same utterances but analyzed differently, the bracketed string of words in (203a) is analyzed as a relative clause, as some other researchers might do, whereas the same string of words in (204) is analyzed as an adjective modifier in our hypothesis.

(203) a. John$_i$ zhongyu zhao dao yige [Ø$_i$ bu taoyen ___$_j$ de] laoban$_j$.
finally find ASP one NOM not dislike ACC RLTV boss
'John$_i$ finally found a boss whom (he$_i$) doesn't dislike.'
b. *John$_i$ zhongyu zhao dao yige [__$_j$ bu taoyen Ø$_i$ de] laoban$_j$.
finally find ASP one NOM not dislike ACC RLTV boss
'John$_i$ finally found a boss who doesn't dislike (him$_i$).'

(204) John zhongyu zhao dao yige [bu taoyen de]$_{AdjP}$ laoban.
finally find ASP one not dislike ADJV boss
'John finally found a not-disliked boss (= John finally found a boss who (he) doesn't dislike).'

Only the former approach has to account for the puzzle why the reading in (203b) is not available as that in (193b), repeated below as (205b), since both examples (203b) and (205b) are structurally parallel, context-less, and contain a gap. In our analysis, *zhaogu* 'look after' in (205) is a non-stative verb. Therefore, the clause where this verb occurs is indeed a relative clause and there is no room for reanalyzing it as an adjectival phrase.

(205) a. John$_i$ zhao-bu-dao yige [Ø$_i$ keyi zhaogu ___$_j$ de] ren$_j$.
find-not-ASP one NOM can look.after ACC RLTV person
'John$_i$ can't find a person whom (he$_i$) can look after.'
b. John$_i$ zhao-bu-dao yige [___$_j$ keyi zhaogu Ø$_{i/k}$ de] ren$_j$.
find-not-ASP one NOM can look.after ACC RLTV person
'John$_i$ can't find a person who can look after (him$_i$ or someone else$_k$).'

Similarly, the contrast between (206) and (207) below is on a par with that between (203) and (204). We prefer the line of analysis assumed in (204) and (207) because it does not face the aforementioned problem of having to account for the bizarre ungrammaticality in sentences like (203b) and (206b), and therefore this preferred approach is neater.

(206) a. John shi wo$_i$ [yige Ø$_i$ hen zhenxi ___$_j$ de] pengyou$_j$.
is I one NOM very cherish ACC RLTV friend
'John is a friend of mine whom (I) cherish very much.'
b. *John shi wo$_i$ [yige ___$_j$ hen zhenxi Ø$_i$ de] pengyou$_j$.
is I one NOM very cherish ACC RLTV friend
'John is a friend of mine who cherishes (me) very much.'

(207) John shi wo yige [hen zhenxi de]$_{AdjP}$ pengyou.
is I one very cherish ADJV friend
'John is a very cherished friend of mine.'

In order to show that the bracketed strings of words in (202), (204) and (207) do possess the properties of adjectives, we demonstrate below in (208) that the bracketed elements form a comparative adjectival phrase, in which *bu taoyen de* 'not disliked' is an adjective prefixed by a comparative word in Chinese, *bijiao* 'more' (as in *more beautiful*, *more historical*), and belongs to the class of comparative adjectives.

(208) Biqi Bill, John shi yige [bijiao bu taoyen de]$_{AdjP}$ tongshi.
compare.with is one more not dislike ADJV colleague
'Compared to Bill, John is a less disliked colleague.'

In addition to verbs like *xihuan* 'like', *taoyen* 'resent', or *zhenxi* 'cherish', other verbs that behave alike (in the way that can be further reanalyzed to be adjectival phrases) include *tonghen* 'hate', *ai* 'love', which is shown in (210), and *renshi* 'known', as in (212). Importantly, all of these verbs are stative verbs. We prefer the analysis assumed in (210) and (212), in comparison with the less ideal, relative-clause approach exemplified in (209) and (211) because the latter analysis has to account for the odd impossibility of

coindexation between the embedded empty object and the matrix subject displayed in the (209b) and (211b).

(209) a. John$_i$ zhao-bu-dao yige [Ø$_i$ keyi ai ___$_j$ de] ren$_j$.
find-not-ASP one NOM can love ACC RLTV person
'John$_i$ can't find a person whom (he$_i$) can love.'
b. *John$_i$ zhao-bu-dao yige [___$_j$ keyi ai Ø$_i$ de] ren$_j$.
find-not-ASP one NOM can love ACC RLTV person
'John$_i$ can't find a person who can love (him$_i$).'

(210) John$_i$ zhao-bu-dao yige [keyi ai de]$_{AdjP}$ ren.
find-not-ASP one can love ADJV person
'John can't find a loveable person (for himself to love).'

(211) a. Wei-he ni$_i$ yao bangzhu yige [Ø$_i$ genben bu renshi ___$_j$ de] ren$_j$?
for-what you want help one NOM at.all not know ACC RLTV person
'Why would you want to help a person whom (you) don't know at all?'
b. *Wei-he ni$_i$ yao bangzhu yige [__$_j$ genben bu renshi Ø$_i$ de] ren$_j$?
for-what you want help one NOM at.all not know ACC RLTV person
'Why would you want to help a person who doesn't know (you) at all?'

(212) Wei-he ni$_i$ yao bangzhu yige [genben bu renshi de]$_{AdjP}$ ren?
for-what you want help one at.all not know ADJV person
'Why would you want to help a totally unknown person?'

The combination of the multifunctional *de* and stative verbs involved has motivated our alternative reanalysis of treating some relative structures as adjectival modifiers. The above examples and Ross's (1983) analysis[36] of *de* have backed up our hypothesis, and we have shown how this approach avoids the problem a traditional view that treats the same construction as relativization would face. Note that we do not reject the possibility of treating the sequence of words in question as a relative structure. We simply propose and opt for an alternative reanalysis, as some verbs (but not all) under these circumstances do trigger an adjectival reading which means essentially the same in

[36] Namely, when *de* appears between a noun phrase and a stative verb, the English translation for such constructions involves adjectival modification. On the other hand, if *de* appears between a noun phrase and a non-stative verb, the English translation involves a relative clause instead.

Chinese as the reading assumed by a relative-clause analysis. Between the two conflicting solutions, our choice of the adjectival-modifier hypothesis over the relative-clause one is superior for the reason that the former is neater, does not have to assume the occurrence of an empty pronoun, does not have to explain the unavailable readings in (201b), (203b), (206b), (209b), and (211b). Moreover, it also accords better with the economy principle stated in (213):

(213) (Bresnan 2001: 91)

> Economy of Expression: all syntactic phrase structure nodes are optional and are not used unless required by independent principles (completeness, coherence, semantic expressivity).

Compared with the adjectival-modifier solution, the empty pronoun assumed by the relative-clause hypothesis is like 'an empty category dominating no terminal element and providing only redundant *f*-structure information' (Bresnan 2001: 91) that could have been conveyed equally clearly under the adjectival-modifier approach. There is potential competition between these two analyses, but it is the adjectival-modifier approach that wins out because it is structurally simpler and might be empirically more intuitive, too, since it does not involve the use of empty categories. As for verbs that pattern with *zhaogu* 'look after' in (205) in having two possible interpretations with regard to the position of the empty category and cannot be reanalyzed as adjectival modifiers, these include *dabai* 'defeat', shown in (214), *jieshou* 'accept', shown in (215), or *zhuiqiu* 'pursue in the romantic sense', as in (216), all of which are non-stative verbs.

(214) a. John$_i$ zhao-bu-dao yige [Ø$_i$ keyi dabai ___$_j$ de] ren$_j$.
find-not-ASP one NOM can defeat ACC RLTV person
'John$_i$ can't find a person whom (he$_i$) can defeat.'

b. John$_i$ zhao-bu-dao yige [___$_j$ keyi dabai Ø$_i$ de] ren$_j$.
find-not-ASP one NOM can defeat ACC RLTV person
'John$_i$ can't find a person who can defeat (him$_i$).'

(215) a. John$_i$ zhao-bu-dao yige [Ø$_i$ keyi jieshou __$_j$ de] ren$_j$.
find-not-ASP one NOM can accept ACC RLTV person
'John$_i$ can't find a person whom (he$_i$) can accept.'

b. John$_i$ zhao-bu-dao yige [__$_j$ keyi jieshou Ø$_{i/k}$ de] ren$_j$.
find-not-ASP one NOM can accept ACC RLTV person
'John$_i$ can't find a person who can accept (him$_i$ or someone/something else).'

(216) a. John$_i$ you hen duo [Ø$_i$ zhuiqiu ___$_j$ de] ren$_j$.
has very many NOM pursue ACC RLTV people
'There are very many people whom (John) is pursuing.'

b. John$_i$ you hen duo [__$_j$ zhuiqiu Ø$_i$ de] ren$_j$.
has very many NOM pursue ACC RLTV people
'There are very many people who are pursuing (John).'

Finally, we would like to add that the adjectival-modifier approach preferred by us in (202), (204), (207), (208), (210) and (212) should not be discredited simply because some of their English translations might seem odd and not as natural as their counterpart English translations assumed by the relative-clause approach. For an example like that in (217), in which *xianggan* 'relate' is an intransitive verb and the bracketed elements are, again, structurally ambiguous in Chinese just like other examples we have previously seen, most people might prefer to treat *de* as an adjectival marker rather than a relative clause marker because the English translation for the former treatment sounds better, even though in Chinese, the sentence itself is perceived to convey essentially the same meaning by either approach.

(217) John zhi shi ge [bu xianggan de]$_{AdjP}$/[__$_i$ bu xianggan de]$_{Relative\ Clause}$ ren$_i$.
only is CL not relate ADJV NOM not relate RLTV person
'John is just an unrelated person/John is just a person who does not relate (to the speaker or someone else).'

In conclusion, in this rather complicated discussion of syntactic constraints on empty objects, we have demonstrated that empty object pronouns must be bound (coindexed with and f-commanded) by topical antecedents. Or alternatively put, an object pronoun may be, but need not be, empty if it is anteceded by a topical element whether it is an argument or not.

3.2.5 Extended binding

Previously, we established that (i) an empty subordinate subject must be bound (coindexed with and f-commanded) by an argument, and (ii) an object may be, but need not be, empty if it is bound by a topic. As complicated as the derivations for these two constraints might already seem, they still do not quite capture the full picture with regard to the syntactic distribution of empty subjects and objects, because there are cases involving split antecedents that have not yet been taken care of. For instance, in the following three examples from (218) to (220), the embedded subjects all have split antecedents. Intransitive verbs *jianmian* 'meet' in (218) and (220) and *renshi*[37] 'know' in (219) both require a plural subject in order to be semantically plausible. In these three examples, the empty subject is coindexed with the combination of the matrix subject and another argument explicitly mentioned in the same sentence. Clearly, our current constraint on empty subordinate subjects does not cover cases with split antecedents like these.

(218) Wang xiansheng$_i$ dai Wang taitai$_j$ dao [dang-nian $Ø_{i+j}$ diyi-ci jianmian
Mr take Mrs to then-year NOM first-time meet
de] difang.
RLTV place
'Mr$_i$ Wang took Mrs$_j$ Wang to the place where (they$_{i+j}$) met for the first time back then.'

[37] The verb *renshi* 'know', exemplified in (219), can be transitive or intransitive in Chinese. One of its intransitive meanings requires a plural subject and has the same meaning as in 'John and Bill knew each other in high school', or 'We know each other very well', except that no object (e.g., each other) needs to be present in Chinese. Therefore, to convey the meaning of 'we know each other' in Chinese, one can simply utter the following sentence in (i). Note that the sentence in (i), unlike its English counterpart, is a complete sentence without any empty category and it never means 'we know'. A different verb must be used in order to convey the meaning of 'we know'.

(i) Women renshi.
we know
'We know each other.'

(219) Wang xiansheng$_i$ juede xianzaide Wang taitai$_j$ [bi yiqian Ø$_{i+j}$ gang
Mr feel present Mrs compare before NOM just
renshi shi geng you meili].
know time more have charm
'Mr Wang$_i$ feels that the present Mrs Wang$_j$ is more charming than when (they$_{i+j}$) first knew each other.'

(220) (J. Huang 1989, (42d))
Wo$_i$ xiang [Ø$_{i+j}$ jianmian yihou, ta$_j$ jiuhui huan qian].
I think NOM meet after he will return money
'I$_i$ think that after (we$_{i+j}$) meet, he$_j$ will return the money.'

Our current constraint on empty objects suffers from the same problem, as the following (221) exemplifies. In (221), the object of the verb *saorao* 'harass' is ambiguous: it could refer to the matrix subject *John*, *his girlfriend*, or the combination of the two, depending on the context. The last reading involves a split antecedent (in a likely scenario in which John and his girlfriend live together and the landlord has harassed them).

(221) John$_i$ gaosu tade$_i$ nu-pengyou$_j$ [ruguo fangdong zai zheyang saorao
tell his female-friend if landlord again like.this harass
Ø$_{i/j/i+j}$, ta$_i$ jiuyao tichu gaosu].
ACC he will raise lawsuit
'John$_i$ told his girlfriend$_j$ that if the landlord keeps harassing (him$_i$/her$_j$/them$_{i+j}$) like this, he$_i$ will just file a lawsuit.'

The problem of split antecedents becomes more complicated when some of the referents from the split antecedents come from outside the sentence itself, i.e., not explicitly mentioned within the sentence. For instance, given no context, in (222) below, the empty subject who got married could be (i) *John*, (ii) *Mary*, (iii) *John* and *Mary*, (iv) *John* and *his partner* (indexed as *k*), or (v) *Mary* and *her partner* (indexed as *l*). In the last two readings, the partial antecedent (John's or Mary's own partner) is *implied* and does not actually appear in the sentence. Similarly in (223), without any context, the object of *saorao* 'harass' can be (i) *Mr Wang*, (ii) *Mrs Wang*, (iii) *Mr* and *Mrs Wang*, or (iv) the two plus others (indexed as *k*) who also live in the same house, presumably the rest of the

Wang family. The ambiguity of these examples we have seen so far comes from general world knowledge about the scenario described by the sentence; these sentences do not require any particular context to give rise to the ambiguity.

(222) John$_i$ ji gei Mary$_j$ yi-zhang [duo nian qian Ø$_{i/j/i+j/i+k/j+l}$ jiehun shi paide] zhaopian.
send to one-CL many years ago NOM marry time taken photo
'John$_i$ sent Mary$_j$ a photo taken many years ago whey (he$_i$/she$_j$/they$_{i+j/i+k/j+l}$) got married.'

(223) Wang xiansheng$_i$ xiang Wang taitai$_j$ baoyuan [ruguo linju zai zheyang saorao Ø$_{i/j/i+j/i+j+k}$, ta$_i$ jiuyao tichu gaosu].
Mr to Mrs complain if neighbor again like.this harass ACC he will raise lawsuit
'Mr Wang$_i$ complained to Mrs Wang$_j$ that if the neighbor kept harassing (him$_i$/her$_j$/them$_{i+j/i+j+k}$) like this, he would just file a lawsuit.'

In view of the above examples of split antecedents of empty subjects and objects, whether the split antecedents are explicitly mentioned within the sentence or simply implied in the given scenario, our previous syntactic constraints need to be revised in order to accommodate such cases. To solve this problem, we suggest that the standard sense of *binding* which we have assumed so far be extended. The standard binding theory considered in the P&P Theory essentially involves coindexation and c-commanding between two NPs, while in our LFG account, binding is defined by the coindexation and f-commanding relationship between two NPs. In *Extended Binding*,[38] the normally strict coindexation and f-commanding relationship is loosened up as follows:

(224) Extended Binding: A binds B if A f-commands B and also satisfies at least one of the following two conditions:
(i) A is coindexed with B, or
(ii) A is included in the reference of B.

[38] The notion of Extending Binding stems from Lasnik's (1976, 1981) discussion of the binding conditions for coreference, disjoint reference, and split antecedence. Refer to Büring (2005) for a complete overview.

We argue that the definition of 'binding' when applied to empty pronouns in Chinese may be extended as the above (224), and the aforementioned problems of split antecedents can be resolved. Given the introduction of Extended Binding, our previous third and fourth constraints established in 3.2.3 and 3.2.4 should be revised as the following (225) and (226), respectively.

(225) An empty subordinate subject in Chinese must be bound in the extended sense by an argument.
(226) An object may be (but need not be) empty in Chinese if it is bound in the extended sense by a topical antecedent (whether this antecedent is an argument or not).

To see how the revised proposal works, let us return to the earlier examples (218) and (221), repeated below as (227) and (228), respectively. In (227), the antecedent of the empty subject Ø includes *Mr* and *Mrs Wang*, which correctly falls under the prediction of Extended Binding because *Mr Wang* f-commands the embedded empty subject, and *Mr Wang* is also included in the reference of Ø. The alternative view is to say that *Mrs Wang* f-commands the empty subject in question and *Mrs Wang* is also included in the reference of Ø.

(227) Wang xiansheng$_i$ dai Wang taitai$_j$ dao [dang-nian Ø$_{i+j}$ diyi-ci jianmian
Mr take Mrs to then-year NOM first-time meet
de] difang.
RLTV place
'Mr$_i$ Wang took Mrs$_j$ Wang o the place where (they$_{i+j}$) met for the first time back then.'

As for (228), the embedded empty object following *saorao* 'harass' ambiguously refers to *John* himself, *his girlfriend*, or the combination of the two.

(228) John$_i$ gaosu tade$_i$ nu-pengyou$_j$ [ruguo fangdong zai zheyang saorao
tell his female-friend if landlord again like.this harass
Ø$_{i/j/i+j}$, ta$_i$ jiuyao tichu gaosu].
ACC he will raise lawsuit
'John$_i$ told his girlfriend$_j$ that if the landlord keeps harassing (him$_i$/her$_j$/them$_{i+j}$) like this, he$_i$ will just file a lawsuit.'

In the last reading, the split antecedent is licensed by Extended Binding because *John* f-commands the empty object Ø and is also included in the reference of Ø. Alternatively analyzed, we could also say that *John's girlfriend* f-commands Ø and is included in the reference of Ø. Therefore, the use of an empty pronoun under these circumstances is permitted by Extended Binding and it accords with our revised constraints in (225) and (226). Summing up, our earlier proposal was challenged by problematic sentences which involve split antecedents, and these have prompted us to revise the constraints by applying the Extended Binding to the conditioning of Chinese empty pronouns.

3.2.6 Local binding of empty possessive pronoun by the closest antecedent

Our next constraint regards the use of empty possessive pronouns in Chinese. J. Huang (1984) observes that the extraction of a possessive pronoun is possible if the NP to which the possessive pronoun is attached is a subject, even though such extraction violates Ross's (1967) Left Branch Condition. However, J. Huang also notes that the same extraction is unavailable if the NP from which the possessive pronoun is extracted is an object. (229a) and (230a) exemplify the possible extraction of an empty possessive pronoun from a subject NP, while extracting an empty possessive pronoun from an object NP is disallowed in (229b) and (230b), and Ross's Left Branch Condition is thus not violated.

(229) (J. Huang 1984, (83))
a. Zhangsan$_i$, [Ø$_i$ baba]$_{Subject}$ hen youqian.
GEN father very rich
'(Speaking of) Zhangsan$_i$, (his$_i$) father is very rich.'

b. *Zhangsan$_i$, wo kanjian [Ø$_i$ baba]$_{Object}$ le.
I see GEN father ASP
'(Speaking of) Zhangsan$_i$, I saw (his$_i$) father.'

(230) (J. Huang 1984, (84))

a. [[[Ø$_i$ baba]$_{Subject}$ hen youqian de] neige xuesheng$_i$] lai le.
GEN father very rich RLTV that student come ASP
'The student (whose) father is rich came.'

b. *[[Wo kanjian [Ø$_i$ baba]$_{Object}$ de] neige xuesheng$_i$] lai le.
I see GEN father RLTV that student come ASP
'The student (whose) father I saw came.'

We agree with J. Huang (1984) about this subject-object asymmetry with regard to the extraction of possessive pronouns, but we disagree with his derivational approach to treating the empty possessive pronoun in (229a) and (230a) as a base-generated *pro* which then becomes a variable later when it is bound by a topic, as in (229a), or by the head noun of a relative clause, as in (230a), in order to be anteceded. In Huang's hypothesis, the reason that (229b) and (230b) cannot be analyzed in the same manner is that the empty possessive pronouns in these two examples, treated as *pro*s, would have been identified with the local antecedent first, i.e., *wo* 'I' in both cases, and cannot reach the non-local NP. Therefore, the intended readings (for the empty possessive pronoun to be bound by the non-local antecedent) are unavailable in (229b) and (230b). There is an alternative view that also targets the licensing of empty possessive pronouns in Chinese. Pan and Hu (2000) propose a strict Locality Constraint similar to J. Huang's but without assuming any necessary derivation: an empty possessive pronoun must be coindexed with the closest NP. This accounts for the impossible reading in (229b) and (230b) because an intervening local NP blocks the intended coindexation. This is also true of the example in (231a) below, because coindexation between the empty possessive pronoun and the non-local antecedent is not attainable. However, examples like (231b) not only seriously undermine J. Huang's (1984) analysis but also violate the Locality Constraint proposed by Pan and Hu (2000), because the empty possessive pronoun in (231b) refers to the farther away topic, *Zhangsan*, despite the intervention of the local NP, *wo* 'I'.

(231) (Pan and Hu 2000, (37))

a. *Zhangsan$_i$, wo renwei [[Ø$_i$ baba] hen youqian].
I think GEN father very rich
'(Speaking of) Zhangsan$_i$, I think that (his$_i$) father is very rich.'

b. Zhangsan$_i$, wo renwei [[Ø$_i$ yanjing] zhangde hen piaoliang].
I think GEN eyes grow very pretty
'(Speaking of) Zhangsan$_i$, I think that (his$_i$) eyes are very pretty.'

In view of this, Pan and Hu (2000) formulate a rule to take care of the apparent violation. Based on the assumption that all topicalized NPs must have proper licensers, which may be seen as parallel to the notion that a topic must be associated, one way or another, with an argument, Pan and Hu propose that a topical NP can be licensed by an empty possessor across an intervening NP if and only if the empty possessor is inalienable from the possessed NP. That is to say, only an *inalienable* empty possessor can properly license a topicalized NP across an intervening local NP. Even though Pan and Hu did not address how an inalienable relation is defined in their theory, given the contrast in grammaticality between (231a) and (231b) (along with other examples in their paper), we believe it to be a physical kind of inalienability, e.g., one and one's eyes, one and one's skin, one and one's voice, etc., and their definition of inalienability does not include kinship, which therefore rules out (231a).

Pan and Hu's two rules for licensing empty possessive pronouns look very plausible, but one question arises as to why the initial Locality Constraint does not apply to (232a) and (232b), revised from their own example seen in (231). Namely, if the Locality Constraint rules that an empty possessive pronoun must be coindexed with the closest NP, which will be the matrix subject *wo* 'I' in (232a) and (232b), why is the reading of the empty possessor bound by the local NP *wo* 'I' impossible in both examples? A syntactic condition is supposed to be inviolable and be rigorously applied, say, to every occurrence of empty pronouns. The implausible coindexation between the empty possessor and the local antecedent in (232) seems to escape the prediction of the Locality Constraint, which is unaccounted for in Pan and Hu's analysis.

(232) a. Shuodao Zhangsan$_i$, wo$_j$ renwei [[Ø$_{*i/*j}$ taitai] zhangde hen piaoliang].
speaking.of I think GEN wife grow very pretty
'Speaking of Zhangsan, I think that (*his/*my) wife is very pretty.'
b. Shuodao Zhangsan$_i$, wo$_j$ renwei [[Ø$_{i/*j}$ yanjing] zhangde hen piaoliang].
speaking.of I think GEN eyes grow very pretty
'(Speaking of) Zhangsan, I think that (his/*my) eyes are very pretty.'

Given the unavailable interpretation of the empty possessor bound by its closest local antecedent, *wo* 'I', in (232a) and (232b), we feel that Pan and Hu's Locality Condition is too strong. There is no obvious syntactic factor that prohibits the coindexation of the empty possessor and the local antecedent *wo* 'I' in these two cases, as the Locality Constraint would have predicted. However, it is pragmatic considerations that prevent this coreference. According to Reinhart (1982) (cited in Lambrecht 1994: 127), 'the relation "topic-of" expresses the pragmatic relation of aboutness which holds between a referent and a proposition with respect to a particular discourse'. Suppose that the prerequisite 'aboutness' relation must hold between a topic and its comment clause to maintain the pragmatic well-formedness of a topic-comment sentence. Such a requirement would not be fulfilled in (232) if coindexation of the empty pronoun and the local NP *wo* 'I' were allowed. Imagine a scenario where a speaker concerns the person *Zhangsan* as the sentence topic, but somehow this speaker's following comment about *Zhangsan* is that *I think that my wife is very pretty*, or *I think that my eyes are very pretty*, neither of which says anything about the topic *Zhangsan*. Such utterances might not necessarily be ruled out by syntax, but they are clearly very odd pragmatically and against our general world knowledge. Therefore, it is our general understanding of the world that rules out coindexation of the empty possessor and its local antecedent in (232a) and (232b). Given (232), Pan and Hu's Locality Constraint does not seem to hold. In view of this and also to combine their Inalienability Condition, we suggest that Pan and Hu's Locality Constraint can be modified to a weaker claim: an empty possessive pronoun (i) must be bound by the closest local antecedent without violating general world knowledge, and/or (ii) it may also be anteceded by the topic across an intervening local NP if and only if the relation between the possessor and the possessed is physically inalienable.

There is another concern regarding the above revised constraint over empty possessive pronouns. Let us revisit (232b), repeated as (233) below. In (233), the empty possessor is bound by *Zhangsan*, the closest local antecedent without violating our general understanding of the world, as the coindexation with *wo* 'I' would be a marginal reading. Along the same line, we would expect the empty possessor in (234) to be also bound by *Zhangsan*, since the comment clause should be *about* the topic. However, this expectation is not met, and the only available reading is for the empty possessor to be coindexed with *Lisi*, despite the large degree of pragmatic unnaturalness. The sentence in (234) seems to escape the proposed constraint because the empty possessor in question does not refer to the sentence topic, whereas coreference with the closest local antecedent (*Lisi*) is semantically/pragmatically odd (indicated by the # symbol).

(233) Shuodao Zhangsan$_i$, wo$_j$ renwei [[Ø$_{i/*j}$ yanjing] zhangde hen piaoliang].
speaking.of I think GEN eyes grow very pretty
'(Speaking of) Zhangsan, I think that (his/*my) eyes are very pretty.'

(234) Shuodao Zhangsan$_i$, <u>wo$_j$ gaosu Lisi$_k$</u> [[Ø$_{*i/*j/\#k}$ yanjing] zhangde hen piaoliang].
speaking.of I tell GEN eyes grow very pretty
'(Speaking of) Zhangsan$_i$, I told Lisi$_k$ that (*my/his$_{*i/\#k}$) eyes are very pretty.'

The contrast between (233) and (234), we believe, lies in whether the underlined elements, intervening between the empty possessor and a non-local NP, are parenthetical[39] or not. Non-parenthetical elements may be considered a proper barrier that blocks non-local antecedence (e.g., coindexation between Ø and the topic, *Zhangsan*, in (234)), whereas parenthetical elements may not count as a barrier and allow long-distance antecedence

[39] Dehé and Kavalova (2007) define *parentheticals* as 'expressions that are linearly represented in a given string of utterance (a host sentence), but seem structurally independent at the same time. They have been argued to interrupt the prosodic flow of an utterance, introducing intonational breaks and featuring prosodic properties different from those of their host. They are … usually associated with non-truth-conditional meaning. Parentheticals typically function as modifiers, additions to or comments on the current talk. They often convey the attitude of the speaker towards the content of the utterance, and/or the degree of speaker endorsement.' Parentheticals have been argued to be external to the syntactic structure of their host sentence (Burton-Roberts 1999).

such as coindexation between Ø and *Zhangsan* in (233). The contrast between (235) and (236) can also be accounted for in the same manner. *I guess* or *I heard* are often parenthetical in nature when used in a situation like (235). However, *I emphasized to Lisi*, as shown in (236), is non-parenthetical and therefore acts as a barrier that blocks long-distance binding, even though the local reading is marginal and pragmatically odd (owing to the lack of comment about the sentence topic).

(235) Zhangsan$_i$, wo$_j$ cai/tingshuo [[Ø$_{i/*j}$ yanjing] zhangde hen piaoliang].
I guess/hear GEN eyes grow very pretty
'(Speaking of) Zhangsan$_i$, I guess/heard, (his$_i$) eyes are very pretty.'

(236) Zhangsan$_i$, wo$_j$ xiang Lisi$_k$ qiangdiao [[Ø$_{*i/*j/\#k}$ yanjing] zhangde hen piaoliang].
I to emphasize GEN eyes grow very pretty
'(Speaking of) Zhangsan$_i$, I emphasized to Lisi$_k$ that (*my/his$_{*i/\#k}$) eyes are very pretty.'

This has shown that when the relationship between an empty possessor and the possessed is inalienable, the potential long-distance binding (e.g., coindexation between *Zhangsan* and Ø in (234) and (236)) may still be blocked by intervening NPs if the intervening NPs are not parenthetical and form a barrier. In conclusion, in view of the above examples, the original constraint on empty possessors should be revised again as follows: an empty possessor (i) must be bound by the closest local antecedent without violating the general world knowledge, and/or (ii) it may also be anteceded by a non-local topic across intervening local NPs if and only if the relation between the possessor and the possessed is physically inalienable and the intervening NPs are parenthetical.

3.2.7 Arbitrary empty pronouns

Lastly, it has been observed that there are sentences which involve empty pronouns whose references are arbitrary. Arbitrary reference tends to be indefinite or generic, and sentences like those indeed pose a problem to our analysis so far because our previous constraints have not been able to account for the arbitrary reference of any empty pronoun.

The study of arbitrarily used empty pronouns is far less extensive in the literature, compared to the study of referentially used empty pronouns. The examples in (237) and (238) below illustrate the arbitrary use of empty pronouns (indicated as $Ø_{arb}$) in Chinese. In both examples, the two occurrences of empty subjects must be coreferential and share the same index. However, in (238), the shared reference can be either *John*, which is a referential interpretation predicted by our syntactic constraint, or an indefinite, generic pronoun, which leads to an arbitrary, undetermined interpretation.

(237) [Da $Ø_{arb}$] shi qing, [ma $Ø_{arb}$] shi ai.
hit ACC is affection scold ACC is love
'Hitting (one) shows affection. Scolding (one) shows love. (This well-known Chinese proverb refers to the positive meaning hidden behind physical punishment, e.g., the kind imposed by parents to their children).'

(238) John faxian [zheli meiyouren$_i$ zhidao [ruguo $Ø_{arb/i}$ you he jiu
discover here no.one know if NOM have drunk alcohol
$Ø_{arb/i}$ jiu bu neng kai che]].
NOM just not can drive car
'John discovered that no one$_i$ here knows that if (they$_i$/one) drank any alcohol, (they$_i$/one) must not drive.'

To account for cases like those in (237) and (238), we will follow Lebeaux's (1984) analysis of PRO and extend his treatment of PRO_{arb} in English to our $Ø_{arb}$ in Chinese. Let us first look at his examples in (239) below. (239a) and (239d) can be suitably translated into Chinese in (240).

(239) (Lebeaux 1984, (17))
a. PRO_{arb} making a large profit (as a slum landlord) requires PRO_{arb} exploiting the tenants.
b. PRO_{arb} becoming a movie star entails/involves PRO_{arb} being recognized by everyone.
c. PRO_{arb} winning this game requires PRO_{arb} mastering the end game.
d. PRO_{arb} to know him is PRO_{arb} to love him.

(240) a. [As a slum landlord …]
Ø$_{arb}$ yao zhuan da qian dehua, Ø$_{arb}$ jiu dei yazha fangke.
NOM want.to earn big money if NOM just must exploit tenants
'If (one, as a slum landlord) wants to make a lot of money, (one) must exploit the tenants.'
b. Ø$_{arb}$ yao liaojie ta jiu shi Ø$_{arb}$ qu ai ta.
NOM want.to know him just is NOM go love him
'If (one) wants to know him, (one) must go and love him.'

Lebeaux observes that each PRO$_{arb}$ in (239) (and our Ø$_{arb}$ in (240)) 'ranges over some universal set of entities (the domain restricted by context) and the elements picked out by the two PROs$_{arb}$ must be identical' (Lebeaux 1984: 260). He hypothesizes that there is a universal quantifier binding two open positions (e.g., the empty subjects of *making a large profit* and *exploiting the tenants* in (239a) and in (240a)). Lebeaux terms this sort of obligatory arbitrary reference between two PROs (or in our case, Øs) as 'linked reference'. Opposite to the linked reference, there are examples in which two arbitrarily used empty pronouns do not share the same reference, as exemplified by (241) and the Chinese sentence in (242) below. Note that in (242), the two adjectives, *fuyou* 'rich' and *jidu* 'jealous', function as predicates; adjectival predicates are very common in Chinese.

(241) (Lebeaux 1984, (27b), (28b))
a. PRO$_{arb}$ having ancestors in the Old World doesn't make stories about PRO$_{arb}$ winning the West any less interesting.
b. PRO$_{arb}$ winning the trust of the populace requires that PRO$_{arb}$ having to serve in the army be abolished.
(242) Ø$_{arb}$ fuyou rongyi yinqi Ø$_{arb}$ jidu.
NOM rich easily cause NOM jealous
'Wealth breeds envy.'

Following Lebeaux's analysis, we will treat such arbitrarily used empty pronouns (linked in reference or not) as anaphoric indefinites which are bound by an implicit universal operator in a generic context, giving rise to the indefinite reading of such pronouns. When two arbitrarily used empty pronouns are linked in reference, as in (239) and (240), there is

only one universal quantifier binding both empty categories simultaneously. When they are not linked in reference, as in (241) and (242), there are two different universal quantifiers at work.

3.3 Summary

This chapter focuses on how zero anaphora in Chinese is constrained by syntax. Incorporating the predecessors' efforts into our own analysis, including the suggestion of *extended binding*, we present the following five syntactic constraints in (243)[40] for empty pronouns in Chinese.

(243) a. No coindexation between an empty pronoun and its coargument is allowed.
b. No empty pronoun is allowed to appear after a preposition or in a pivotal position (e.g., after a control verb, a raising verb, or a sense verb).
c. An empty subordinate subject must be bound in the extended sense by an argument.
d. An empty object must be bound in the extended sense by a topic.
e. An empty possessive pronoun (i) must be bound by the closest local antecedent without going against the general world knowledge, and/or (ii) it may also be anteceded by a non-local topic across intervening local NPs *if and only if* the relation between the possessor and the possessed is physically inalienable and the intervening NPs are parenthetical.

Our syntactic constraints in Chapter 3 work in conjunction with the next chapter, which deals with how intersentential zero anaphora in Chinese is conditioned pragmatically at the level of information structure. Combined together, the two chapters present an interdisciplinary analysis that characterizes zero anaphora in Chinese.

[40] Within the scope of our current study in this thesis, we can so far claim that these constraints are Chinese-specific, and whether or not they have any wider typological or universal relevance requires further research in the future.

Chapter 4: A pragmatic characterization of Chinese zero anaphora

Chapter 4 focuses on how zero anaphora in Chinese is conditioned pragmatically, in contrast to the syntactic constraints introduced in the previous chapter. These pragmatic conditions, based on Lambrecht's (1994) information-structure theory, apply to intrasentential and intersentential empty pronouns alike, and are related to what role an empty pronoun plays at information structure, suggesting the significance of context and discourse in such pragmatic conditions. Owing to the importance of context and discourse in characterizing pragmatic generalizations from an information-structure-oriented perspective, it may be more difficult to examine pragmatic conditions in intrasentential empty pronouns with only minimum or zero context available. Therefore, there will be more examples using intersentential empty pronouns at the connected-discourse level in this chapter in order to observe the pragmatic generalizations at information structure. The representation of *i*-structure and Lambrecht's (1994) information-structure theory were introduced in Chapter 1. Now let us start our pragmatic characterization of empty pronouns in Chinese.

4.1 Pragmatic generalizations for intersentential zero anaphora

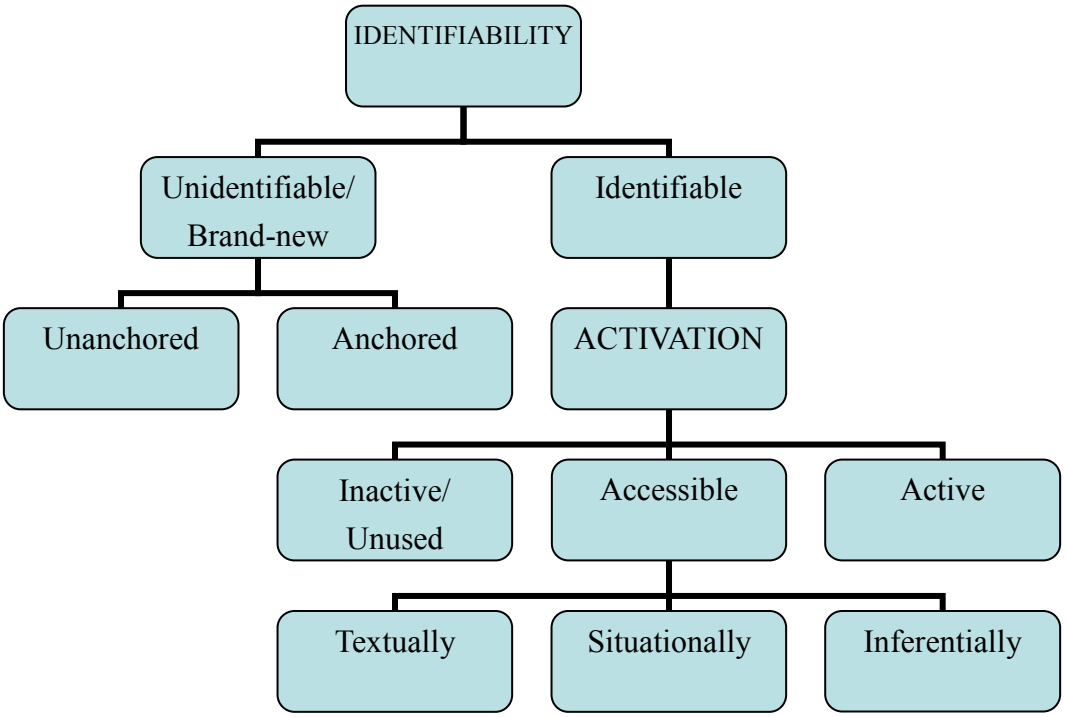
This section centers on how intersentential empty pronouns in Chinese are pragmatically conditioned according to their status at information structure.

4.1.1 Empty pronouns must have actively or accessibly identifiable antecedents

Based on Lambrecht's diagram of identifiability and activation previously seen in (47), repeated below in (244), we suggest that the antecedent of an empty pronoun must be actively or accessibly identifiable. Following Chafe's (1976, 1987) terminology, Lambrecht uses *identifiability*, as a cognitive category, to be a speaker's estimation of whether the discourse representation of a certain referent is identifiable or unidentifiable to the hearer's mind. What distinguishes identifiable and unidentifiable referents may be

viewed as on a par with the distinction between presupposition and assertion (Lambrecht 1994: 77). *Activation*, on the other hand, is also a cognitive category and it concerns the speaker's evaluation of the status of the representation of an identifiable referent as already activated/active, accessible, or simply inactive. The definitions of these different cognitive subcategories such as anchored/unanchored, inactive, accessible, etc. were discussed previously in 1.5.2; we will provide further examples below.

(244) (Lambrecht 1994, (3.25))



Our first pragmatic condition, which states that the antecedent of an empty pronoun must be actively or accessibly identifiable, might, at first glance, seem similar to Taraldsen's (1978) 'principle of recoverability' and Jaeggli's (1982) 'identification hypothesis', both of which emphasize that whether an empty pronoun can be used depends on its recoverability. If the referent cannot be recovered, using an empty pronoun in this position results in ungrammaticality. We argue that our version of identifiability is different from Taraldsen's recoverability and Jaeggli's identification in that the identifiability we assume here is a cognitive property, and being identifiable does not

entail any morphological marking, e.g., verbal inflections. Instead, what an identifiable referent adopted in our theory entails is that a shared representation of such a referent exists in the speaker's and the hearer's mind at the time of the utterance, while a referent is unidentifiable if the representation of the referent is not shared but only available in the speaker's mind (Lambrecht 1994: 77-8). One next question might be: why don't we just claim that the referent of an empty pronoun must be identifiable, rather than that it has to be actively or accessibly identifiable? Recall that we mentioned the three kinds of gradable activation states in 1.5.2, ranging from active, accessible, to inactive. Summarized from Chafe's (1976, 1987) notion and Lambrecht's (1994) elaboration, an active referent is currently focused at the forefront of the interlocutor's consciousness at a certain time of the speech, and it is often expressed as a reduced form. An accessible referent, which can be triggered in three different manners (textually, situationally, or inferentially), is one of which the interlocutor has background awareness and is in his peripheral consciousness, not directly focused. Active and accessible types of referents are stored in the interlocutor's short-term memory; their representation may be more vivid but such active or accessible mental states can cease as soon as another referent comes into the speech and the attention of the interlocutors will have shifted. In contrast, an inactive referent is one whose cognitive status is in the interlocutor's long-term memory, neither focally nor peripherally active, and has prosodic and morphological correlates marked by full lexical coding with stress. Its cognitive status stays for a much longer period of time than that of active or accessible referents. The three activation statuses have correlates in prosody and morphology, leading to various linguistic expressions from accented to unaccented, from full lexical NPs, pronominals, and inflections to zero anaphora. Having refreshed the terminology assumed by Lambrecht, we can now say that our first pragmatic condition for intersentential zero anaphora bears a certain degree of resemblance to Taraldsen's and Jaeggli's views in that we all acknowledge the importance that an empty pronoun should be used only when the hearer can identify (or recover) the referent without difficulty. We differ from Taraldsen and Jaeggli by emphasizing that the hearer's ability to retrieve the referent of an empty pronoun in discourse stems from its active or accessible cognitive status. An inactive referent does not qualify as a felicitous

use of zero anaphora despite its identifiable status because it lacks pragmatic saliency; an inactive cognitive status entails accentuation of the referential expression and full lexical coding. Lambrecht (1994: 96) concludes that an inactive referent cannot be represented in a pronominal form (except for deictic pronouns) and, therefore, a null form of an inactive referent is simply out of the question.

To formally represent the above tendency at information structure, we propose that a given referent has a binary value either as identifiable or unidentifiable for the feature STATUS (for cognitive status). If the referent is valued as unidentifiable, its *i*-structure goes on to include the feature ANCHORED, which has a plus value if it is cognitively anchored, or a minus value if it is unanchored. In contrast, if the referent has an identifiable value for the STATUS feature, its *i*-structure continues with the referent's activation state and will have an ACTV feature (short for activation), which can be valued as inactive, accessible, or active. The example in (245) is uttered as an opening sentence to introduce some story either about the speaker or the dog, and is of sentence-focus structure. It illustrates the identifiable status and active state of the speaker[41] *I*, and the unidentifiable, brand-new unanchored status of the referent, *a huge dog*.

(245) I saw a huge dog yesterday.

$$\begin{bmatrix} \text{FOCUS} & \begin{bmatrix} \text{PRED} & \text{'see} < \text{DRef1}, \text{DRef2} > \text{'} \\ \text{DRef1} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{FORM} & \text{I} \\ \text{STATUS} & \text{identifiable} \\ \text{ACTV} & \text{active} \end{bmatrix} \\ \text{DRef2} & \begin{bmatrix} \text{PRED} & \text{'dog'} \\ \text{SPEC} & \text{a} \\ \text{ADJ} & \{[\text{PRED} \ \text{'huge'}]\} \\ \text{STATUS} & \text{unidentifiable} \\ \text{ANCHORED} & - \end{bmatrix} \\ \text{ADJ} & \{[\text{PRED} \ \text{'yesterday'}]\} \end{bmatrix} \end{bmatrix}$$

[41] That the speaker *I* is identifiable and active in reference is consistent with Erteschik-Shir's (2007) view of 'permanently available topics'. She believes that speakers and hearers are always readily available as topics in a conversation.

Even though the example in (246) below is also uttered as an opening sentence and is of sentence-focus structure just like the one in (245), the subject, *a friend of mine*, is brand-new anchored because it can be related to the speaker. *His leg*, on the other hand, is textually accessible, given the previous mention of the speaker's friend.

(246) A friend of mine broke his leg.

$$\begin{bmatrix} \text{FOCUS} & \begin{bmatrix} \text{PRED} & \text{'break} < \text{DRef1, DRef2} > \text{'} \\ \text{DRef1} & \begin{bmatrix} \text{PRED} & \text{'a friend of mine'} \\ \text{STATUS} & \text{unidentifiable} \\ \text{ANCHORED} & + \end{bmatrix} \\ \text{DRef2} & \begin{bmatrix} \text{PRED} & \text{'his leg'} \\ \text{STATUS} & \text{identifiable} \\ \text{ACTV} & \text{accessible} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The status of the subject *he* in Speaker B's reply in (247) below is a topic under discussion, hence its ACTV status as active in its *i*-structure shown in (248). Speaker B's sentence itself is of topic-comment structure, in which *buying an island* is the comment about the lit up discourse topic, *he*/*John*. As *he* in Speaker's reply is not a contrastive topic, the value for the CONTRAS(tive) attribute is minus in (248).

(247) A: Do you know what John did yesterday?
B: He bought an island.

(248) *i*-structure of B's reply in (247)

$$\begin{bmatrix} \text{TOP} & {}_1\begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{FORM} & \text{he} \\ \text{CONTRAS} & - \\ \text{STATUS} & \text{identifiable} \\ \text{ACTV} & \text{active} \end{bmatrix} \\ \text{FOC} & \begin{bmatrix} \text{PRED} & \text{'buy} < \text{DRef1}, \text{DRef2} > \text{'} \\ \text{DRef1} & {}_1[\quad] \\ \text{DRef2} & \begin{bmatrix} \text{PRED} & \text{'island'} \\ \text{SPEC} & \text{an} \\ \text{STATUS} & \text{unidentifiable} \\ \text{ANCHORED} & - \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

We did not show the inactive use of a discourse referent and its *i*-structure because it may require a much lengthier context or even several paragraphs in order to illustrate the decrescendo of activation status. Presumably, it might have started as active and then became inactive because the referent has not come into the interlocutors' attention again since its first mention. The principle behind the *i*-structure of an inactive referent remains consistent with the rest. That is, an inactive referent will be valued as identifiable for the STATUS feature and inactive for the ACTV feature.

Returning to our first pragmatic generalization, which states that the referent of an empty pronoun must be actively or accessibly identifiable, according to Lambrecht's (1994) topic acceptability scale in (249) below, we can see that the more cognitively active a referent is (and therefore the easier to be interpreted), the more acceptable for it to become a topic. Such a pragmatic condition strengthens the link between an empty pronoun and a topical antecedent, and the correlation between the two is captured at *i*-structure as we have presented here. The cognitive preference determines how likely a pronoun is to be dropped in discourse: the lower the effort required to process a referent, the more likely it is for the referent to be dropped.

(249) (Lambrecht 1994, (4.34))

THE TOPIC ACCEPTABILITY SCALE

active	most acceptable
accessible	↑
inactive	↑
brand-new anchored	↑
brand-new unanchored	least acceptable

In the Chinese example in (250)[42] below, Speaker B's reply involves an empty subject that is identifiable and active and also serves as a discourse topic. The choice of using an empty subject here is therefore well-formed, as predicted by our pragmatic generalization, although the use of an overt pronoun in the same position is equally plausible and natural. (251)[43] is the *i*-structure of Speaker B's reply that involves an empty subject.

(250) A: Qiao. Zhe shi wo nur_i de zhaopian.
look this is my daughter GEN picture
'Look. This is a picture of my $daughter_i$.'
B: $Ø_i/ta_i$ zhangde zhen ke-ai.
she grow really cute
'(She_i) looks really cute.'

(251) *i*-structure of B's reply in (250)

$$\begin{bmatrix} \text{TOP} & {}_1\begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{OVERT} & - \\ \text{CONTRAS} & - \\ \text{STATUS} & \text{identifiable} \\ \text{ACTV} & \text{active} \end{bmatrix} \\ \text{FOC} & \begin{bmatrix} \text{PRED} & \text{'}zhangde\ ke-ai\text{ ''look cute''} < \text{DRef} > \text{'} \\ \text{ADJ} & \{[\text{PRED '}zhen\text{ ''really''}]\} \\ \text{DRef} & {}_1[\quad] \end{bmatrix} \end{bmatrix}$$

[42] In (250), Speaker A's comment can be ambiguous, as the sentence *zhe shi wo nur de zhaopian* can either mean 'this is a picture of my daughter' or 'this is my daughter's picture'. Given the context here, it is undoubtedly the former.

[43] In order not to overcomplicate the matter for the purpose intended in this section, we treat *zhangde ke-ai* 'look cute' from Speaker B's reply as one single verbal predicate that only requires one discourse referent to fulfill its argument structure.

The sentence in (252), on the other hand, is uttered in a scenario where John is surprised that the movie Bill has been watching is not finished yet. The empty subject in John's utterance is not active because the referent is not explicitly mentioned anywhere in the discourse; neither is it deictic. However, the legitimate use of the empty subject under this circumstance comes from the accessibility of the referent. In other words, the possibility of using an empty pronoun in the well-formed (252) is licensed by the situationally accessible referent, the movie Bill has been watching. The intended interpretation is salient without any confusion because the identifiable and situationally accessible referent (i.e., the movie) can be easily processed with a 'low cost'[44] cognitive effort. (253) demonstrates the *i*-structure for the sentence in (252).

(252) [John entered the living room, found Bill still watching the movie$_i$, and said …]
Zeme Ø$_i$ hai mei yan-wan?
how NOM still not perform-finish
'How come (the movie) is still not finished?'

(253) *i*-structure of (252)

$$\begin{bmatrix} \text{FOCUS} & [\text{PRED} \ '\textit{zeme}\text{"how come"'}] \\ \text{COMPLETIVE} & \begin{bmatrix} \text{PRED} & '\textit{yan-wan}\ \text{"finish"}<\text{DRef}>' \\ \text{ADJ} & \left\{ \begin{matrix} [\text{PRED} \ '\textit{hai}\ \text{"still"'}] \\ [\text{PRED} \ '\textit{mei}\ \text{"not"'}] \end{matrix} \right\} \\ \text{DRef} & {}_1[\ \ ] \end{bmatrix} \\ \text{TOP} & {}_1\begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{OVERT} & - \\ \text{STATUS} & \text{identifiable} \\ \text{ACTV} & \text{accessible} \\ \text{CONTRAS} & - \end{bmatrix} \end{bmatrix}$$

To conclude, our discussion in the previous chapter of the syntactic constraints of empty pronouns has argued that being anteceded by a topic alone does not sanction a

[44] Chafe (1994: 71-5) discusses 'activation cost': a given item will be least costly cognitively when it has already been active, while a new item will require more cognitive efforts to mentally process such information and is therefore more costly.

grammatical use of empty subordinate subject pronouns; a structural requirement such as being an argument has more relevance in allowing a legitimate use of empty subordinate subjects. However, at the discourse level, how cognitively salient a referent is at information structure is fundamentally important for permitting any use of zero anaphora. Only those that require very 'low cost' cognitive efforts to process the meaning may prompt empty pronouns to be used in discourse. We have shown that for any well-formed intersentential empty pronouns, discourse prominence and cognitive saliency are prerequisite, and this property can be nicely captured at *i*-structure under the framework of LFG by using features like STATUS, ACTV, and so on.

4.1.2 A pronoun may be empty if it bears the discourse function as a topic, primary or secondary

Derived from the first pragmatic generalization that we just introduced, our second pragmatic generalization for intersentential zero anaphora concerns the topical status of an empty pronoun, which may appear to be similar to our first claim because being topical necessarily implies that the referent is actively identifiable. The second pragmatic condition we are proposing states specifically that an intersentential pronoun may be empty if it bears the discourse function of topic, primary[45] or secondary, at information structure. It is uncontroversial to assume that some sentences employ more than one topic with one being primary and the other being secondary. All the examples we have come across so far are straightforwardly single-topic sentences. If we look at the previous example in (250), repeated below as (254), it is easy to see that a topical referent in

[45] Tsao (1988, 1990) adopts the following surface structure of the Chinese clause: [COMP [Topic 1 [Topic 2 …]]], in which Topic 1 is not an argument of the maximal clause, whereas Topic 2 is clause-internal but has a non-canonical position (e.g., as a fronted object, the second nominal in a double subject construction, or a pre-verbal adverbial). As primary and secondary topics in Chinese are decided according to their linear order, here we can view Topic 1 as the primary topic and Topic 2 the secondary topic.

(i) John$_i$, naxie dongxi, ta$_i$ lian yi-yang dou bu chi.
those things he even one-CL all not eat
'Speaking of John, of those things, he does not even eat a single one.'

For instance, in the multiple-topical (i) above, *John* serves as Topic 1 (or the primary topic) and 'those things' is Topic 2 (or the secondary topic).

discourse, e.g., the subject in B's reply, may be (but need not be) empty. The *i*-structure in (255) represents B's reply that involves an empty subject, as repeated from (251).

(254) A: Qiao. Zhe shi wo nur_i de zhaopian.
look this is my daughter GEN picture
'Look, this is a picture of my $daughter_i$.'
B: $Ø_i/ta_i$ zhangde zhen ke-ai.
she grow really cute
'(She_i) looks really cute.'

(255) *i*-structure of B's reply in (254)

$$\begin{bmatrix} \text{TOP} & {}_1\begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{OVERT} & - \\ \text{CONTRAS} & - \\ \text{STATUS} & \text{identifiable} \\ \text{ACTV} & \text{active} \end{bmatrix} \\ \text{FOC} & \begin{bmatrix} \text{PRED} & \text{'}\textit{zhangde ke}-\textit{ai}\text{ "look cute"} < \text{DRef} > \text{'} \\ \text{ADJ} & \{[\text{PRED} \text{ '}\textit{zhen}\text{ "really"'}]\} \\ \text{DRef} & {}_1[\quad] \end{bmatrix} \end{bmatrix}$$

On the other hand, according to Tsao's (1990) classification, there are Chinese sentences that involve different kinds of topics – primary versus secondary ones – and he distinguishes the following four[46] types of sentence constructions involving secondary topics: (i) double subject constructions, (ii) non-sentence-initial pre-verbal locatives and temporals as topics, (iii) the *ba* construction, and (iv) object fronting constructions. Among these four types of multiple-topic constructions, only the double subject construction and the object fronting construction are relevant to our analysis of empty pronouns for two reasons. Firstly, the occurrence of empty locatives and empty temporals

[46] Tsao (1990) considers the first two types (double subject constructions and non-sentence-initial pre-verbal locatives and temporals as topics) as involving local secondary topics, while the latter two types (the *ba* construction and object fronting constructions) as involving non-local secondary topics. The difference between local and non-local comes from a transformational view: whether the secondary topic in question shows any sign on the surface of having been moved. Since LFG is a non-transformational grammar, we will leave out Tsao's original distinction between local and non-local secondary topics in this study.

is not considered (in our study) the standard use of empty pronouns. Secondly, the NP following the morpheme *ba* in the *ba* construction (on which we will soon elaborate) can never be empty due to the prepositional characteristic of *ba*, and one of our proposed syntactic constraints (in Chapter 3) forbids pronouns following a preposition from being empty. Let us begin the discussion of multiple-topic sentences with the double subject construction. As the name suggests, the so-called 'double subject construction' involves two consecutive occurrences of NPs and a predicate in the form of NP1 + NP2 + Predicate, in which NP1 and NP2 can be further construed as a complex NP. For example, the sentences in (256a) to (256c) are identical when spoken, but structurally and semantically, they can be three-ways ambiguous. (256a) is a single-topic sentence and the relationship between the two NPs is of topic and comment. The genitive marker *de* in (256b) is optional, and the two consecutive NPs, in fact, form a single subject. In (256c), however, the second NP *mama* 'mother' is treated as a secondary topic, while the initial NP *ta* 'he' is the primary topic. It is sentences like (256c) that will be the subject of our discussion of the double subject construction.

(256) a. [Ta]$_{\text{Topic}}$ [mama]$_{\text{Subject}}$ guoshi le.
he mother pass.away ASP
'Speaking of him, his mother passed away.'
b. [Ta-(de) mama]$_{\text{Subject}}$ guoshi le.
he-GEN mother pass.away ASP
'His mother passed away.'
c. [Ta]$_{\text{Primary Topic}}$ [mama]$_{\text{Secondary Topic}}$ guoshi le.
he mother pass.away ASP
'Speaking of him, as for his mother, (she) passed away.'

Many Chinese grammarians have debated over the controversial status of the double subject construction (Teng 1974; Xu & Langendoen 1985; Tsao 1990; Shi 2000), and it is Tsao's (1990) viewpoint of treating the second NP in the double subject construction as a secondary topic that is the most relevant to our analysis. For instance, given the dual-topic analysis of the sentence in (256c), revised below as (257a), a comment such as the one in (257b) or (257c) can follow (257a) naturally, and the empty subject in question can be

coreferential with the primary or secondary topic, depending on the context.

(257) a. [Ta$_i$]$_{\text{Primary Topic}}$ [mama$_j$]$_{\text{Secondary Topic}}$ shang-ge-yue guoshi le.
he mother last-CL-month pass.away ASP
'Speaking of him, as for his mother, (she) passed away last month.'
b. Suoyi Ø$_i$ zuijin kanqilai hen shiluo.
so NOM recently look very lost
'So (he) looks lost recently.'
c. Ju-shuo Ø$_j$ liuxia le hen duo caichan.
according.to-say NOM leave ASP very much fortune
'It is said that (she) left a huge fortune behind.'

The second type of multiple-topic construction in Tsao's (1990) classification involves temporals and locatives as secondary topics, as the second bracketed NPs in (258) and (259) show. The fact that an optional pause particle *ya* can follow the said NPs suggests that they truly are topics of some sort, because a topic in Chinese can be optionally segmented from the rest of the clause by a pause or one of the four pause particles: *a*(*ya*), *ne*, *me* and *ba* (Tsao 1990; Chu 1998). Because temporals and locatives are adverbials, not pronominals, the possibility for them to be overtly used or not in clauses is immaterial to our scope of study in this thesis.

(258) (Tsao 1990: 156, (88))
[Ta]$_{\text{Primary Topic}}$ [zuotian]$_{\text{Secondary Topic}}$ (ya) meiyou lai.
he yesterday PART not come
'Speaking of him, yesterday (he) didn't come.'
(259) (Tsao 1990: 156, (89))
[Ta]$_{\text{Primary Topic}}$ [zai meiguo]$_{\text{Secondary Topic}}$ (ya) you hen duo pengyou.
he in America PART have very many friends
'Speaking of him, in America (he) has many friends.'

The third type of multiple-topic construction according to Tsao's (1990) classification is the *ba* construction, which is a very unusual structure and has been extensively discussed in Chinese linguistics literature. We first came across the *ba*

construction in the previous chapter when we talked about coverbs. We mentioned at the time that *ba* is commonly analyzed as a coverb or a preposition, and it has the effect of promoting the direct object of the verb and placing it in a higher position right after *ba*. In a nutshell, if we assume that the canonical word order in Chinese is SVO as the word order of the sentence in (260a), then (260b) represents a standard example of the *ba* construction. Semantically, the two sentences are largely the same with subtle differences in what is being emphasized, e.g., who broke the vase or what was broken.

(260) a. Subject + Verb + Object
Wo da-puo huaping le.
I hit-break vase ASP
'I broke the vase.'
b. Subject + *ba* + Object + Verb
Wo ba huaping da-puo le.
I BA vase hit-break ASP
'I broke the vase.'

Tsao (1990) argues convincingly that the *ba* construction is multi-topical, in which the first NP is the primary topic and the second NP, the one following *ba*, is the secondary topic. To cite some of his evidence, a pause particle can be inserted between the NP after *ba* and the rest of the clause, which then explicitly marks the topical status of the NP. The *ba* NP also has to be definite or generic[47] in reference, just like any topical NP should be. Therefore, the schematic structure in (260b) for the *ba* construction should be revised as 'Primary Topic + *ba* + Secondary Topic + Verb' under Tsao's analysis. He (along with Mei (1978, 1980)) specifically points out the inadequacy of associating the *ba* NP with the main verb as an object, because despite the fact that many *ba* NPs are indeed the direct object of the main verb, robust counterexamples have been observed. As Mei (1980)

[47] For instance, the following sentences are examples of generics, all of which express a type of general property.

(i) Whales are mammals.
(ii) The Siberian tiger is an endangered species.
(iii) A cat has four legs.

argues, the *ba* NP, *bilu* 'fireplace', in (261) is not the object of the verb but a locative NP, which does not enter the argument structure of the main verb. This view further supports Tsao's revised schema of the *ba* construction: 'Primary Topic + *ba* + Secondary Topic + Verb'. Since *ba* is more likely to be analyzed either as a coverb or a preposition, and if we assume that one of our syntactic constraints for zero anaphora which states that pronouns following a coverb or a preposition must not be empty is on the right track, the *ba* NP simply cannot be empty even though it may act as a secondary topic.

(261) (Tsao 1990: 172, (18))

[Ta]$_{\text{Primary Topic}}$	ba	[bilu]$_{\text{Secondary Topic}}$	sheng	le	huo.
he	BA	fireplace	build	ASP	fire

'He fired up the fireplace.'

Tsao's fourth type of multiple-topic construction involves an object fronted from its canonical position, that is, after the predicate. Take (262a) for example, which exhibits the canonical SVO word order, and then take a look at (262b), in which the object *zuoye* 'assignment' is fronted to follow the subject immediately and becomes pre-verbal. Semantically, the two sentences are essentially the same but with slight differences in which argument is being emphasized: the subject or the object.

(262) a. [Ta] xie-wan [zuoye] le.
he write-finish assignment ASP
'He finished writing the assignment.'

b. [Ta]$_{\text{Primary Topic}}$ [zuoye]$_{\text{Secondary Topic}}$ xie-wan le.
he assignment write-finish ASP
'He finished writing the assignment.'

In Tsao's analysis, the first NP, *ta* 'he', is the primary topic and the fronted object NP is the secondary topic in (262b). Two pieces of evidence supporting the topical status of the fronted object come from (i) the possibility of inserting a pause particle right after the fronted object, as in (263), and (ii) the fronted object must be definite or generic[48] in

[48] See the previous footnote for examples of generics.

reference. As (264b) shows, ungrammaticality arises when an indefinite object is fronted from the canonical, post-verbal position when compared with (264a). If the NP in question is replaced by a definite one, as (264c) shows, the sentence immediately becomes acceptable.

(263) [Ta]$_{\text{Primary Topic}}$ [zuoye]$_{\text{Secondary Topic}}$ a, zao-jiu xie-wan le.
he assignment PART early-just write-finish ASP
'He already finished writing the assignment a long time ago.'

(264) a. [Ta] kan le [yi-chu xi].
he watch ASP one-CL play
'He watched a play.'

b. *[Ta]$_{\text{Primary Topic}}$ [yi-chu xi]$_{\text{Secondary Topic}}$ kan le.
he one-CL play watch ASP
'He watched a play.'

c. [Ta]$_{\text{Primary Topic}}$ [zhe-chu xi]$_{\text{Secondary Topic}}$ kan le.
he this-CL play watch ASP
'He watched this play.'

The example in (265) below shows the possibility of dropping a secondary topic. Following Tsao's (1990) view of treating the fronted object as a secondary topic and also given our proposal that a pronoun may be empty if it bears the discourse function as a topic (primary or secondary), we can now see why the object in the second clause of the following example in (265) is empty. The empty pronoun in question is the direct object of the verb *jieshao* 'introduce', and it is anteceded by the secondary topic, *mei-nu* 'beautiful girls' from the previous clause, in a scenario where someone claims that he does not like beautiful girls so his friend would not bother to introduce them.

(265) [Ta]$_{\text{Primary Topic}}$ [mei-nu$_i$]$_{\text{Secondary Topic}}$ bu xihuan, na wo jiu bu jieshao
he beautiful-girls not like then I just not introduce
Ø$_i$ la.
ACC PART
'He doesn't like beautiful girls; then I won't introduce (them).'

Having established that a Chinese pronoun can be empty if it bears the discourse function as a topic, it might be tempting to conclude that *all* topical pronouns can be empty. However, this overgeneralization is not borne out because topics that are *contrastively* used (e.g., the underlined, accented pronouns in (266) and (267)) must not be empty.

(266) A: Will he and she be arriving soon?
B: HE will be.
(267) (Lambrecht 1994: 291, (3.20b))
I saw Mary and John yesterday. SHE says HELLO, but HE's still ANGRY at you.

Contrastive topic expressions are defined by Lambrecht (1994: 291) as 'accented pronouns and lexical constituents with discourse-active referents which are in focus'. We must differentiate contrastive topics from non-contrastive topics, and it is essential to point out that the previous three pragmatic tendencies apply to non-contrastive topical pronouns only. A contrastive topic expression must not be null despite its topical status. For instance, the emphasized constituent following *lian* 'even' (which functions as an emphatic adverb) in the *lian ... dou/ye* 'even …' construction in Chinese, shown in (268) below, bears a contrastive meaning, and it cannot be null despite being topical at the same time. (The topical status of the NP right after *lian* 'even' is signified by the optional pause particle *ya*).

(268) [John$_i$ is one of the kindest people in the world …]

Dan	lian	ta$_i$/*Ø$_i$	(ya),	dou/ye	dui	ni	de	xingwei	gandao
but	even	he	PART	all/also	about	you	GEN	behavior	feel

shiwang.
disappointed
'But even *he* felt disappointed by your behavior.'

The above discussion summarizes Tsao's (1990) classification of multiple-topic constructions in Chinese, among which the double subject construction and the object fronting construction are closely related to our analysis because they support our proposal

that a secondary topical pronoun as well as a primary one may be empty in discourse. The examples also demonstrate that multiple-topic sentences are common in Mandarin Chinese, and that it is the discourse function, not the syntactic argument, of an intersentential pronoun at information structure that plays a crucial part in licensing whether or not the pronoun may be empty. A non-contrastively used pronoun may be empty if it bears the discourse function as a topic, primary or secondary.

4.1.3 A discourse topic may extend its domain to the following clauses to form a topic chain

The third pragmatic condition is based upon Tsao's (1979, 1990) insightful work on topics in Chinese. Our view of analyzing empty pronouns on the basis of their topical status at intersentential level stems from Tsao's (1979, 1990) Topic-NP Deletion Rule, which was first introduced in Chapter 2 and is characterized by a chaining function of the topic construction in Chinese. Tsao's Topic-NP Deletion Rule states that when several sentences share the same topic and form a topic chain, the topic in each of the subsequent sentences may be deleted, and the topic in the first sentence remains present. That is to say, the topic of a sentence can be deleted if the topic is the same as the one in a preceding sentence, and the deleting process forms a topic chain. For example, in (269) below, with *zhongguo* 'China' being the topic of the opening sentence and then forming a topic chain with the sequence of sentences that follow, *zhongguo* 'China' is deleted from the rest of the discourse from the topic position and only the initial topic survives. What determines a topical expression to be deleted or not is relevant to its discourse role (or topicality) and has nothing to do with being a subject, which accords with the general assumption that Chinese-like languages (which include Japanese and Korean) are *discourse-oriented* in contrast to *sentence-oriented* languages like English. The example in (269) shows that a topic may extend its domain, semantically and syntactically, over several clauses and form a topic chain. A topical pronoun in discourse may, therefore, be empty in this sense.

(269) (Huang 1984, (55))

[Zhongguo$_i$, difang hen da]. [Ø$_i$, renkou hen duo]. [Ø$_i$, tudi hen
China place very big NOM population very many NOM land very
feiwo]. [Ø$_i$, qihou ye hen hao]. [Ø$_i$, women dou hen xihuan].
fertile NOM climate also very good ACC we all very like
'(As for) China, (its) land area is very large. (Its) population is very big. (Its) land is very fertile. (Its) climate is also very good. We all like (it).'

Note that the chaining function of a sequence of topics under this deletion rule is not restricted to clause-initial position only, as in (269). The example in (270) demonstrates that the initial topic, *zhei-bian* 'this side (of the house, which is implicitly derived but semantically salient in the original text), is in control of the empty pronouns indexed as *i* and is responsible for the subsequent topic NPs' deletion. The fact that the empty pronoun indexed as *j* is coindexed with the matrix subject *ta* 'he' also supports the syntactic constraint we proposed in Chapter 2 that an empty subordinate subject must be anteceded by an argument.

(270) (Tsao 1990: 43, (59))

Zhei-bian$_i$, yinwei limian jiu-jiu-zang-zang, ta$_j$ shuo [Ø$_j$
this-side (of the house) because inside old-and-dirty he say NOM
xiang youqi-hao Ø$_i$ yihou zai mai Ø$_i$].
want paint-finish ACC after then sell ACC
'(Speaking of) the house on this side, because the inside is old and dirty, he said (he) wants to paint (it) and then sell (it).'

The sentences in (269) and (270) demonstrate that a discourse topic may extend its domain to the following clauses to form a topic chain, and the subsequent topics may be empty under the rule of Topic-NP Deletion. However, we showed previously that multiple-topic sentences occur frequently in Chinese, as it is known to be a topic-prominent and discourse-oriented language. This has led to another observation by Tsao (1990) that a secondary topic is also capable of forming a topic chain of its own. Take (271) for instance, in which the primary topic *zhe-ge nuhai* 'this girl' occupies the

sentence-initial position and the secondary topics are subsequently regarding her eyes and her ears. In (271), there are two sets of topic chains. The two clauses in the first set have the topic *yanjing* 'eyes' in common, and the two clauses in the second set have the topic *erduo* 'ears' in common. The two secondary topics *yanjin* 'eyes' and *erduo* 'ears' in (271)[49] are both followed by the pause particle *ya*, which marks the topical status of the two NPs. Even though the primary topic *zhe-ge nuhai* 'this girl' is not explicitly marked by a pause particle, the comma after it shows that a pause after the NP in question is also expected.

(271) (Tsao 1990: 133, (36b))

Zhe-ge nuhai, [yanjing$_i$ ya hen da], [Ø$_i$ hen piaoliang], [erduo$_j$ ya que hen xiao], [Ø$_j$ hen nankan].
this-CL girl eyes PART very big NOM very beautiful ears PART on.the.contrary very small NOM very ugly
'(Speaking of) this girl, her eyes are large and (they) are beautiful; but her ears are small and (they) are ugly.'

Assuming that the proposed pragmatic condition that a discourse topic, primary or secondary, can extend its domain to the subsequent clauses and that empty topic pronouns occur due to the Topic-NP Deletion Rule is on the right track, we can now explain that the ambiguity in the example in (272) falls under the prediction of this pragmatic condition. With the sentence-initial NP *zhe-ge nuhai* 'this girl' being the primary topic and *yanjing*

[49] While we agree with Tsao on treating the example in (271) as involving two sets of topics and forming two topic chains, alternatively, the internal structure of (271) can also be analyzed differently as involving no topic chain but as an instance of VP coordination, as (i) below demonstrates. In this analysis, the two predicates *hen da* 'being very big' and *hen piaoliang* 'being very pretty' in the first part of the example are coordinated, and there is no empty subject. The same analysis applies to the second part of the example.

(i) Zhe-ge nuhai, [yanjing ya [hen da] [hen piaoliang]], [erduo ya que [hen xiao] [hen nankan]].
this-CL girl eyes PART very big very pretty ears PART on.the.contrary very small very ugly
'(Speaking of) this girl, her eyes are large and beautiful, but her ears are small and ugly.'

The examples in (i) and (271) are essentially identical, phonologically and semantically, and it is indeed the case that such a sentence is structurally ambiguous between the two constructions, although, since the version in (i) involves VP-coordination and no empty pronoun, we will only pay attention to the other equally strong, natural interpretation shown in (271).

'eyes' being the secondary one, the empty subject pronoun may be anteceded by either of them and then form a topic chain accordingly.

(272) (Tsao 1990: 133, (36a))

Zhe-ge	nuhai$_i$ ya,	yanjing$_j$	hen	da,	Ø$_{i/j}$	hen	piaoliang.
this- CL	girl PART	eyes	very	big	NOM	very	beautiful

'(Speaking of) this girl$_i$, her eyes$_j$ are large, and (she$_i$/they$_j$) is/are beautiful.'

We have adopted Tsao's (1979, 1990) analysis and integrated his Topic-NP Deletion Rule to our pragmatic characterization of intersentential empty pronouns. From the above discussion, we can conclude that a discourse topic, be it primary or secondary, may extend its domain to the following clauses to form a topic chain, in which the initial topic is overtly expressed and the subsequent topics may be empty under Tsao's Topic-NP Deletion Rule. It is not subject to constraints on the syntactic position, e.g., subject or object, because a topical pronoun can appear in the subject position as well as the object position.

4.1.4 Summary

In 4.1 we have provided a pragmatic characterization of intersentential empty pronouns based on Lambrecht's (1994) and Tsao's (1979, 1990) theories. We began the first pragmatic condition with Lambrecht's diagram of identifiability and activation, and proposed that the referent of an empty pronoun in discourse must be actively or accessibly identifiable. Different kinds of activation and identification are captured at information structure, and by looking at the components at *i*-structure, we observe that the use of an empty pronoun whose referent is not actively or accessibly identifiable is disallowed. We then moved on to the second pragmatic condition, which states that a pronoun may be empty if it bears the discourse function of topic, primary or secondary, at information structure. Here we introduced Tsao's (1979, 1990) classification of multiple-topic constructions in Chinese and incorporated it in our characterization. With Chinese being a topic-prominent and discourse-oriented language, multiple-topic

sentences are by no means uncommon and they occur frequently. The third pragmatic condition stems from Tsao's Topic-NP Deletion Rule. Not only primary topics can extend their domain to the following clauses to form topic chains; secondary topics can equally do so as well. To avoid an overgeneralization that suggests all topical expressions in discourse may be empty in Chinese, we distinguished contrastive topics from non-contrastive topics and maintained that the previous three pragmatic conditions only apply to non-contrastive topic expressions. Contrastive topics must not be null despite their topical status. A contrastive topic expression will have the plus (+) value for the feature CONTRAS (short for contrastive) inside the *i*-structure. We used the *lian ... dou/ye* 'even …' construction to instantiate contrastive topic expressions in Chinese, but this is not an exhaustive list. In conclusion, we have shown how intersentential empty pronouns are pragmatically conditioned in discourse in this section, and all these constraints can be nicely captured at *i*-structure under the framework of LFG. Coupled with the syntactic characterization of context-independent constraints for Chinese zero anaphor introduced in Chapter 3, we have now completed the conjunctive analysis of how zero anaphora is conditioned syntactically *and* pragmatically. The next section will serve as an examination of these proposed constraints by looking at empty pronouns in Chinese narrative. This will allow us to test how our proposal and rules for pronoun interpretation are realized in actual texts.

4.2 An examination of our analysis

After presenting a series of pragmatic generalizations for intersentential empty pronouns at connected-discourse level in the previous section, our characterization of zero anaphora in Chinese and how it is constrained pragmatically and syntactically is now nearly complete. In this section, we extend the discussion by testing how intrasentential and intersentential zero anaphora that occurs naturally is used from an article we have found from an online corpus. This section serves to examine the extent to which our proposed constraints are realized in this article we have taken from a Mandarin Chinese website that provides articles of various topics and of different lengths. We picked one

medium-length text written by a contemporary unknown author who intends to pass on the wisdom of how one should approach conflicts in a non-confrontational way.

4.2.1 An examination of zero anaphora in naturalistic data

Let us begin with our chosen article in (273) below. For ease of reading and analysis, we divide the original text in eight subparts from (273a) to (273h). We will discuss different types of empty pronouns used in this text one by one.

(273) (from http://ibook.idv.tw/enews/enews61-90/enews77.html)

a. I used.to ever live ASP one-CL Japanese-style flat-house
wuzi houmian you yi-pian kong-di, suoyi $Ø_i$ zai hou-yuan
house behind have one-CL empty-land so NOM at back-garden
kaipi le cai-yuan, $Ø_i$ zhong le xie cai.
develop ASP vegetable-garden NOM grow ASP some vegetable
'I used to live in a Japanese-style bungalow. There was a piece of unused land in the back of the house. So (I) developed a vegetable garden. (I) grew some vegetables.'

b. Dan gebi $linju_j$ zai tamen hou-yuan yang le shi-ji-zhi
but next.door neighbor in their back-garden raise ASP ten-or.so-CL
sheng-dan $muji_k$, $Ø_k$ jingchang pao-guolai, $Ø_k$ ba
lay-egg hens NOM often run-across NOM BA
cai-yuan nong de yitahutu.
vegetable-garden make as spoiled
'But the next-door neighbor kept over ten egg-laying hens in their back garden. (The $hens_k$) often came around. ($They_k$) completely spoiled the vegetable garden.'

c. Wo_i xiang $tamen_j$ ti guo hen duo ci, $Ø_i$ xiwang $tamen_j$ ba
I to them mention ASP very many times NOM hope they BA
muji guanqilai. $Tamen_j$ hen keqi daoqian bing daying,
hens cage they very politely apologize and promise,
keshi conglaibu xingdon. Wo_i you bu neng yinwei zhexie
but never take.action I also not can because.of these
xiao shi he $tamen_j$ zheng-lun shang heqi.
small matters with them argue damage harmony

'I mentioned to them many times that (I) hope they would cage the hens. They apologized politely and promised, but never took any action. I also couldn't argue with them over such small matters and damage the harmony.'

d. Houlai, wo_i zhunbei-hao yi-dai $jidan_l$, chen $linju\text{-}xiaohai_m$
afterwards I prepare one-bag eggs while neighbor-kid
guolai jia-lie wan shi, $Ø_i$ song ta_m $Ø_l$ dai hui jia.
come.over home-inside play time NOM give him ACC take back home
'Afterwards, I prepared a bag of eggs, taking the chance when the neighbor's kid_m came over to play in my house, (I) gave him_m (the eggs) to take back home.'

e. Di-er tian $linju\text{-}taitai_n$ guolai dao-xie. Wo_i shuo 'bu yong keqi.
second day neighbor-wife come.over say-thanks I say no need polite
Women jia_o $jidan_p$ laoshi chi-bu-wan. $Ø_{i/o}$ hai zhen yao ganxie
our home eggs always eat-not-finish NOM still really need thank
ning bangmang chi $Ø_l$, miande $Ø_l$ bai-huai le.'
you help eat ACC avoiding NOM put-rotten ASP
'The next day the neighbor's wife came over to say thank you. I said, "No need to be this polite. We can never finish all the eggs in our household. (I/we) actually need to thank you instead, for helping eat (the eggs) in order to avoid that (the eggs) get rotten after being left unused for too long."'

f. $Lao\text{-}taitai_n$ hen jingqide wen, 'Wo_n mei kandao ni jia_o you yang
old-lady very surprised ask I not see your home have raise
muji ah. Zeme $jidan_p$ hui chi-bu-wan?'
hens PART how.come eggs will eat-not-finish
'The old lady asked with surprise, "I never see your house have any hens. How come there are so many eggs that can't be finished?"'

g. Wo_i shuo, '$Ø_{i/o}$ mei-tian zai cai-yuan-li suibian jiu
I say NOM every-day in vegetable-garden-inside randomly just
keyi jiandao haoji-li $jidan_q$. $Ø_o$ zeme chi $Ø_q$ ye
can pick.up quite.a.few-CL eggs NOM no.matter.how eat ACC still
chi-bu-wan."'
eat-not-finish
'I said, "(I/we) can just pick up quite a few eggs randomly every day from our vegetable garden. No matter how (we) eat (the eggs), they just cannot be finished."'

h. Di-er tian qi, $tamen_j$-jia de muji jiu zaiyemei pao-guolai le.
second day from their-house GEN hens just never.again run-across ASP
'From the next day on, their (the neighbor's) hens never came over here ever

again.'

From the above text, we find several types of empty pronouns, all of which are referential: (i) an empty matrix subject, (ii) an empty subordinate subjects, and (iii) an empty objects. The sentences in (274) below involve two occurrences of empty subjects. The first EC is an empty matrix subject. The second EC can be an empty matrix subject, too, if one wishes to view (274) as consisting of two separate sentences; alternatively, it can be treated as an intrasentential subordinate subject if one wishes to view (274) as one long sentence. As the segmentation of Chinese sentences is not always clear (even though the whole sentence meaning is unambiguous) and how to segment them can often be up to personal interpretations, either way of segmenting (274) is acceptable and natural.

(274) *an empty matrix subject*

[But the next-door neighbor kept over ten hens$_i$ in their back garden …]

Ø$_i$	jingchang	pao-guolai,	Ø$_i$	ba	cai-yuan	nong	de
NOM	often	run-across	NOM	BA	vegetable-garden	make	as

yitahutu.
spoiled

'(The hens$_i$) often came around. (They$_i$) completely spoiled the (speaker's) vegetable garden.'

The empty subject pronouns in (274) both refer to the newly-established topic within this mini-discourse, *hens*, which is shifted from the previous topic in the prior discourse, i.e., the speaker *wo* 'I'. The empty subjects in (274) are within the topic domain extended from the clause where the NP 'hens' is initially introduced. Because these clauses share the same topic, a topic chain is formed and the subsequent topics become empty under the Topic-NP Deletion Rule. *Hens* may serve as the primary topic within this mini discourse in (274), but there is still a secondary topic, which is the speaker's vegetable garden, because *cai-yuan* 'vegetable garden' that appears in the last clause of (274) is directly following the morpheme *ba* in the special *ba* construction we already came across earlier. The *ba* construction is characteristically multi-topical. This type of secondary topic

following *ba* in the *ba* construction may also form a topic chain of its own, as in the example in (275a) below. Depending on the context, *muji* 'hens' can also form its own topic chain, as demonstrated by (275b).

(275) a. Muji$_i$ ba cai-yuan$_j$ nong de yitahutu, Ø$_{j/*i}$ nankan-si-le.
hens BA vegetable-garden make as spoiled NOM dead.ugly
'The hens spoiled the vegetable garden$_j$ completely. (It$_j$) is dead ugly.'
b. Muji$_i$ ba cai-yuan$_j$ nong de yitahutu, Ø$_{i/*j}$ zhen taoyen.
hens BA vegetable-garden make as spoiled NOM really annoying
'The hens$_i$ spoiled the vegetable garden completely. (They$_i$) are really annoying.'

Next, we have an empty subordinate subject in (276) below. The embedded subject of the predicate *xiwang* 'hope' is unambiguously coindexed with the matrix subject *wo* 'I', an argument, as predicated by one of our syntactic constraints.

(276) *an empty subordinate subject*
Wo$_i$ xiang tamen$_j$ ti guo hen duo ci, Ø$_i$ xiwang tamen$_j$ ba
I to them mention ASP very many times NOM hope they BA
muji guanqilai.
hens cage
'I mentioned to them many times that (I) hope they would cage the hens.'

There are two occurrences of empty pronouns in (277) below: one subject and the other object, both appearing in the last clause whose main predicate is the ditransitive verb, *song* 'give'. As we can see, the subject of *song* 'give' is an empty pronoun, and so is the object (i.e., the eggs) being given to the neighbor's kid. The empty subject pronoun refers to the speaker, *wo* 'I', which is the matrix subject of the initial clause, hence an argument, as expected by our syntactic constraint. In this mini discourse of (277), *yi-dai jidan* 'a bag of eggs' is the newly-introduced topic, followed by a comment clause which says that these eggs are given to the neighbor's kid to take home by the speaker. The empty object in (277) is, therefore, coindexed with this topic, and this pattern follows our syntactic rule for empty object pronouns that they should be anteceded by a topic,

whether it is an argument or not.

(277) *an empty subject*

Houlai, wo_i zhunbei-hao yi-dai $jidan_j$, chen linju-$xiaohai_k$
afterwards I prepare one-bag eggs while neighbor-kid
guolai jia-lie wan shi, $Ø_i$ song ta_k $Ø_j$ dai hui jia.
come.over home-inside play time NOM give him ACC take back home
'Afterwards, I prepared a bag of eggs, taking the chance when the neighbor's kid_m came over to play in my house, (I) gave him_m (the eggs) to take back home.'

From a different angle, we now discuss cases where an empty pronoun is disallowed. For instance, in a modified version of (276) shown in (278) below, it is not possible to use an empty pronoun after the preposition *xiang* 'to' and after *ba* in the *ba* construction. Both follow our syntactic constraints and fall under our prediction.

(278) [Wo_i xiang $tamen_j$/*$Ø_j$ ti guo hen duo ci [$Ø_i$ xiwang $tamen_j$ ba
I to them mention ASP very many times NOM hope they BA
muji/*Ø guanqilai]].
hens cage
'I mentioned to them/*(them) many times that (I) hope they would cage the hens/*(hens).'

In (277), to use an unexpressed object following the ditransitive predicate *song* 'give' would lead to ungrammaticality, as demonstrated by a revised version of (277) in (279) below. At first glance, it does not seem obvious why using an empty object as the recipient would be impossible, considering that the antecedent of this empty object is very likely to be topical, perhaps as the secondary topic in the mini discourse of (279).

(279) Houlai, wo_i zhunbei-hao yi-dai $jidan_j$, chen linju-$xiaohai_k$ guolai
afterwards I prepare one-bag eggs while neighbor-kid come.over
jia-li wan shi, [$Ø_i$ song ta_k/*$Ø_k$ $Ø_i$ dai hui jia].
home-inside play time NOM give him ACC take back home

'Afterwards, I prepared a bag of eggs, taking the chance when the neighbor's kid_k came over to play in my house, (I) gave him_k/*(him_k) (the eggs) to take back home.'

Before we proceed to explain the ungrammaticality of using an empty object following *song* 'give' in (279), let us first introduce how the ditransitive verb *song* 'give' can be used in Chinese in the following (280). As far as the ditransitive predicate *song* 'give' is concerned, there can be two structural orders regarding the recipient, *ta* 'him', and what is given, *yi-zhi bi* 'a pen', demonstrated by (280a) and (280b).

(280) a. Wo song [yi-zhi bi]$_{\text{What is given}}$ gei [ta]$_{\text{Recipient}}$/*Ø.
I give one-CL pen to him
'I gave a pen to him.'
b. Wo song gei/(gei) [ta]$_{\text{Recipient}}$/*Ø [yi-zhi bi]$_{\text{What is given}}$.
I give to him one-CL pen
'I gave to/(to) him a pen.'

According to our syntactic rule, whichever NP follows the preposition *gei* 'to/give' must not be empty. The ditransitive structure in the bracketed part in (279), underlined, mirrors the structure of (280b), in which the recipient precedes what is given with an optional preposition *gei* 'to' omitted in (279). In other words, the bracketed elements in (279) can be revised as having the structure of (281) below. After this reanalysis in (281), we now know that using an empty object pronoun to replace *ta* 'him' in (279) is disallowed because Ø is following a preposition, *gei* 'to', which is not overtly expressed.

(281) [$Ø_i$ song (gei) ta_k/*$Ø_k$ $Ø_j$ dai hui jia].
NOM give to him ACC take back home
'(I) gave him/*(him) (the eggs) to take back home.'

4.2.2 Summary

The discussion of the above article aims to provide a non-quantitative sample analysis that serves the purpose of examining how zero anaphora is realized in naturalistic data. It is not supposed to be projected to represent most Chinese texts indicatively. Neither does it reflect the full range of zero anaphora usages found to be available in Chinese. The analysis of the article stands as an extended example of our characterization of empty pronouns. As far as the above article in (273) is concerned, the pragmatic generalizations prescribed by the Topic-NP Deletion Rule and the yielded result – topic chains – are very common. We observe that the second mention (and/or the subsequent ones) of the previous topical NP within the same or even the next discourse tends to employ empty pronouns, and the discourse domain of one topic may easily extend further until the topical status is shifted to another NP. The omission of an overt pronoun when it is used to refer to the previous topic may enhance the flow as well as help indicate the boundary of each mini discourse under the same topic. We also find that at the intrasentential level, while an empty subordinate subject pronoun must be bound (in the extended sense) by an argument, the corresponding object pronoun need not be as long as its antecedent is topical. A relevant issue stemming from analyzing this article is how to organize clauses into Chinese sentences. We have pointed out the difficulty of segmenting Chinese clauses, especially when a coordinated structure may not require an explicit conjunction word such as *he*, *bing*, *erchie*, all of which mean 'and'. Even though cohesion certainly plays a crucial part in sentence or paragraph segmentation, some clauses can be regarded as independent ones by some people, but as modifying adjuncts by some others, just because an explicit use of connection words or conjunction words in Chinese is not always necessary. As it is beyond the current scope of the thesis, we will not dwell on the difficulty and the solution to how Chinese sentences should be organized and segmented, but leave such issues for future research.

Chapter 5: A comparative study of Chinese and Japanese zero anaphora

After investigating how zero anaphora in Chinese is syntactically and pragmatically conditioned, we now turn our focus to a comparative study of how zero anaphora behaves in Japanese, which, like Chinese, is also a topic-prominent language that exhibits poor agreement morphology but, nevertheless, allows prevalent uses of empty pronouns. The comparative study can be divided in two. The first part, 5.1, presents the similarities and differences between empty pronouns in Chinese and in Japanese, based on our analysis and Kameyama's (1985) work on zero anaphora in Japanese. The second part of this chapter, 5.2, will center on the Japanese and Chinese translations of Chapter 13 of the first *Harry Potter* book (*The Philosopher's Stone* by J. K. Rowling 1997), and it comprises a comparison of how pronouns (overt versus empty) have been chosen in each version. Since the original text is written in English, which does not permit *pro*-drop, while the omission of pronouns is a prominent feature in both Japanese and Chinese, it is interesting to compare how the officially translated chapter is done in each language, which, we assume, would be a well-written piece of work that reflects a natural use of zero anaphora.

5.1 A comparison with Kameyama's (1985) study of zero anaphora in Japanese

Up till now we have studied the syntactic constraints and the pragmatic conditions for Chinese zero anaphora, followed by an examination of how empty pronouns at connected-discourse level are used in naturalistic data. Our discussion of zero anaphora continues and will now compare how empty pronouns behave and are conditioned in two seemingly similar languages, Chinese and Japanese, both being topic-prominent languages and allowing *pro*-drop. Based on our current analysis of Chinese zero anaphora presented in this thesis and on Kameyama's (1985) detailed study of zero anaphora in Japanese, the discussion below will address the misconception that the two languages seem to behave very similarly in allowing almost rule-free empty pronouns, owing to the

lack of prominent inflectional agreement systems. We aim to show that (i) Chinese and Japanese have very different ways of constraining zero anaphora even though several similarities do exist, and (ii) that zero anaphora in either language is by no means unconstrained.

We will first begin with a brief introduction of Japanese grammar and then continue to compare its extensive usage of zero anaphora with that in Chinese. Like Chinese, Japanese is also classified as a topic-prominent, *pro*-drop language, in which topic, instead of subject, plays an essential role. The basic sentence structure in Japanese consists of a topic and a comment clause. The canonical word order in Japanese is SOV, in contrast to SVO in Chinese. Both Chinese and Japanese allow pronouns to be omitted from the subject and/or the object position, but such *pro*-drop is not licensed by rich inflectional systems. Chinese is not an inflectional language and verbs do not conjugate. Japanese, on the other hand, is inflected (though not highly) when it comes to conjugating verbs in order to distinguish the past and the present/non-past tenses, but such verbal inflections do not mark any grammatical features such as number or gender of nouns. Therefore, we can safely conclude that the reference recoverability of a dropped pronoun in Chinese and in Japanese is not determined by verbal inflections. Instead, discourse plays an important part in retrieving the reference, as we are about to show in the following discussion. Japanese differs from Chinese in having a prominent grammatical system in order to express politeness and formality. Japanese is characterized by a complex system of honorifics: honorific and humble speech reflect the hierarchical nature of Japanese society, whereas such emphasis on politeness and formality has much less grammatical encoding in Chinese.

5.1.1 Similarities

5.1.1 discusses the similarities shared by Chinese and Japanese when it comes to using zero anaphora. The majority of the Japanese data in 5.1.1 will be taken from Kameyama's (1985) work on Japanese zero anaphora. Grammatical judgments on the Japanese data will also be hers, except where noted. Kameyama (1985) investigates in detail the

syntactic properties as well as the discourse properties of empty pronouns in Japanese, and her work serves as a good contrasting ground for us to compare Chinese empty pronouns with.

5.1.1.1 Both languages allow subjects and objects to be empty

First of all, unlike some *pro*-drop languages which only permit subjects to be omitted (due to subject-verb agreement), Chinese and Japanese both allow subject and object pronouns to be empty, even though a finer-grained constraint is called for in order to account for different occurrences of null subjects or objects. The Chinese dialogue in (282), in which (282a) asks the question and (282b) to (282d) are alternative answers, and the Japanese dialogue in (283) below demonstrate the possibility of dropping subject and/or object pronouns when there is no verbal agreement.

(282) a. Ta_i jian dao xiao-zhang$_j$ le ma?
he meet ASP school-master ASP Q
'Did he meet the Principal?'
b. $Ø_i$ jian dao $Ø_j$ le.
NOM meet ASP ACC ASP
'(He_i) met (the Principal).'
c. $Ø_i$ jian dao xiao-zhang/ta_j le.
NOM meet ASP school-master/him ASP
'(He_i) met the Principal/him_j.'
d. Ta_i jian dao $Ø_j$ le.
he meet ASP ACC ASP
'He_i met (the Principal).'

(283) (Kameyama 1985: 6, (6))
a. $Marcie_i$ wa $Lucy_j$ o yobimasita ka?
TOP ACC call.POL.PST Q
'Did Marcie call Lucy?'
b. Hai, $Ø_i$ $Ø_j$ yobimasita.
yes TOP ACC call.POL.PST
'Yes, (she_i) called (her_j).'

c. Hai, $Ø_i$ Lucy o yobimasita.
yes TOP ACC call.POL.PST
'Yes, (she$_i$) called Lucy.'
d. Hai, Marcie wa $Ø_j$ yobimasita.
yes TOP ACC call.POL.PST
'Yes, Marcie called (her$_j$).'

5.1.1.2 Both languages allow arbitrary empty pronouns

The second property Chinese and Japanese empty pronouns have in common is that they both allow arbitrary empty pronouns (or *indefinite zero anaphor* in Kameyama's (1985) terminology) to refer to people in general, such as *one*, *you*, *we*, or *they*. We first came across Chinese empty pronouns whose references are unspecified in 3.2.7 and a Chinese example can be seen in (284) below. The sentence in (285) demonstrates a very similar usage in Japanese. Both examples show a linked reference between two arbitrary empty subjects.

(284) Ruguo $Ø_{arb}$ bu he shui $Ø_{arb}$ hui sidiao de.
if NOM not drink water NOM will die PART
'If (one) does not drink water, (one) will die.'
(285) $Ø_{arb}$ huro ni hair-eba, $Ø_{arb}$ yoku ne-rare-ru.
NOM bath in enter-if NOM well sleep-can-PRES
'If (one) takes a bath, (one) can sleep well.'

As we already saw Chinese examples that involve an unlinked reference between two arbitrary pronouns in 3.2.7, the Japanese example in (286) also demonstrates such an unlinked reference.

(286) (Kameyama 1985: 55, (52))
$Ø_{arb}$ ni to ni o tas-u to, $Ø_{arb}$ yon ni nar-u.
NOM two and two ACC add-PRES when NOM four to become-PRES
'When (you) add two and two, (they, i.e., two and two) become four.'

5.1.1.3 Dummy subjects are not pronounced

The third similarity comes from the observation that in both languages, dummy subject (e.g., referring to time or weather) are not pronounced. As the Chinese examples in (287) and the Japanese examples in (288) demonstrate, the subject of a time-telling or a weather-describing sentence is not overtly expressed in either language. The replacement of an overt pronoun such as *it* in the same position will result in unnaturalness, if not ungrammaticality, because a third person singular pronoun is generally animate[50] in both languages, contra the dummy subject, *it*, in English.

(287) a. Xianzai Ø yijing shi-dian le.
now NOM already ten-o'clock ASP
'(It) is already ten o'clock now.'
b. Waimian Ø zai da-lei.
outside NOM PROG hit-thunder
'(It) is thundering outside.'

(288) (Kameyama 1985: 55, (53))
a. Ø zyuu-zi da yo.
NOM two-o'clock is RPT
'(It) is ten o'clock.'
b. Ø ame datta.
NOM rain was
'(It) was raining.'

[50] Third person singular pronouns in Chinese and Japanese do tend to be animate, but an inanimate use of such pronouns can still be found, as shown by the Chinese sentence in (i) and the Japanese sentence in (ii).

(i) Wo bu xihuan na jian yangzhuang$_i$ yinwei Ø$_i$/ta$_i$ shi hongsede.
I not like that CL dress because it is red
'I don't like that dress because it is red.'

(ii) Sono keki$_i$ wa totemo oishikatta desu kara, Ø$_i$/sore$_i$ wo itsuka zibun de tsukutte mitai to omoimasu.
that cake TOP very tasty was because it ACC sometime self by make want.to.try PART think
'Because the cake was very tasty, I think I would like to try making it myself sometime.'

5.1.1.4 No coindexation between an empty pronoun and its coargument

In addition to the above general properties shared by Japanese and Chinese empty pronouns, syntax also plays a role in conditioning empty pronouns in both languages. Even though Kameyama (1985: 31-2) strongly argues that Japanese zero anaphora is pragmatically constrained, rather than syntactically, because discourse factors and pragmatics are essential in Japanese grammar, there are nevertheless some syntactic rules that are obeyed by Japanese empty pronouns. As proposed by Iida (1996: 303), an empty pronoun in Japanese cannot be coindexed with any of its coarguments. This syntactic binding rule is also respected by every occurrence of Chinese empty pronouns, and it will not be overridden for any contextual factor. The Japanese sentences in (289) and the Chinese ones in (290) below are unacceptable when an empty pronoun is coindexed with its coargument.

(289) (Iida 1996: 303, (39))

a. *Taroo$_i$ ga Ø$_i$ hihansita.
NOM ACC criticized
'Taroo$_i$ criticized himself$_i$.'

b. Taroo$_i$ no ani ga Ø$_i$ hihansita.
GEN brother NOM ACC criticized
'Taroo$_i$'s brother criticized him$_i$.'

(290) (= (167))

a. *John$_i$ hen xihuan piping bieren dan [ta/Ø$_i$ conlai bu piping Ø$_i$].
very like criticize others but he ever not criticize ACC
'John$_i$ likes to criticize others a lot, but he$_i$/(he$_i$) never criticizes (himself$_i$).'

b. John$_i$ zuo guo hen duo huai shi dan [tade$_i$ gege$_j$ conlai bu piping Ø$_i$].
do ASP very many bad things but his brother ever not criticize ACC
'John$_i$ has done many bad things, but his$_i$ brother never criticizes (him$_i$).'

The above binding rule seems to be the only syntactic restriction that Japanese empty pronouns must follow, whereas there are other syntactic rules constraining Chinese

empty pronouns that we already discussed in detail in the previous chapters.

5.1.1.5 The interpretation of zero anaphora relies heavily on discourse

As Kameyama (1985) argues and is also supported by Iida (1996), the interpretation of zero anaphora in Japanese largely relies on discourse and pragmatic factors, and the reference of an unexpressed argument is generally recoverable through context and pragmatics. Chinese can also be characterized similarly because the interpretation of an empty pronoun must be facilitated by requiring context to some extent, such as what or who the topic is. An empty pronoun with no context whatsoever, exemplified by the use of the Chinese empty subject in (291) below, will confuse hearers tremendously, and it is expected for hearers to ask immediately 'Who went to see the movie yesterday?'. Even though the sentence itself on the whole is structurally well-formed, there is just no obvious way of knowing who the subject refers to, and the communicative purpose of this utterance simply fails.

(291) Ø zuotian qu kan dianying.
NOM yesterday go see movie
'(Somebody) went to see the movie yesterday.'

In view of this, we would add that another similarity shared by Chinese and Japanese empty pronouns is the necessity of having a context in order to process the meaning of an unexpressed argument, unlike other *pro*-drop languages that depend on verbal inflections to retrieve pronoun reference. Interpreting Japanese empty pronouns is almost impossible without a discourse context in the same way that a Chinese empty pronoun uttered out of the blue would be considered odd with little communicative value. Take the Japanese sentences in (292) for example. (292a) involves three nominals and one verb in the unmarked word order (Subject-Locative-Object-Verb), and the alternative sentences from (292b) to (292h) demonstrate the range of possibilities for omitting the three nominals (*Mary*, *shelf*, and *vase*). The pattern appears to be syntactically unconstrained, and the

acceptability of each sentence depends on the discourse context. For instance, (292e) would be a very natural reply to the question 'How did Mary use the shelf she bought yesterday?' (Kameyama 1985: 45).

(292) (Kameyama 1985: 44, (35-36))

a. Mary ga tana ni kabin o oki-masi-ta.
NOM shelf LOC vase ACC put-POL-PST
'Mary put a vase on a shelf.'

b. Ø tana ni kabin o okimasita.
NOM shelf LOC vase ACC put
'(Mary) put a vase on a shelf.'

c. Mary ga Ø kabin o okimasita.
NOM LOC vase ACC put
'Mary put a vase (on a shelf).'

d. Mary ga tana ni Ø okimasita.
NOM shelf LOC ACC put
'Mary put (a vase) on a shelf.'

e. Ø Ø kabin o okimasita.
NOM LOC vase ACC put
'(Mary) put a vase (on a shelf).'

f. Mary ga Ø Ø okimasita.
NOM LOC ACC put
'Mary put (a vase) (on a shelf).'

g. Ø tana ni Ø okimasita.
NOM shelf LOC ACC put
'(Mary) put (a vase) on a shelf.'

h. Ø Ø Ø okimasita.
NOM LOC ACC put
'(Mary) put (a vase) (on a shelf).'

The same analysis can also be applied to Chinese, exemplified by the sentences in (293), which allows for an identical range of possible omissions we have just seen in the Japanese data in (292). Therefore, we conclude that discourse and context are essential in resolving zero anaphora in Japanese and Chinese. This statement may be further generalized as that pragmatics and contextual factors are of tremendous importance in

recovering the reference of an unexpressed argument amongst languages that allow *pro*-drop but exhibit meager, if not zero, agreement system, such as Japanese and Chinese.

(293) a. John bai le ge huaping zai zhuo-shang.
place ASP CL vase LOC table-on
'John placed a vase on the table.'
b. Ø bai le ge huaping zai zhuo-shang.
NOM place ASP CL vase LOC table-on
'(John) placed a vase on the table.'
c. John bai le ge huaping Ø.
place ASP CL vase on.the.table
'John placed a vase (on the table).'
d. John bai le Ø zai zhuo-shang.
place ASP ACC LOC table-on
'John placed (a vase) on the table.'
e. Ø bai le ge huaping Ø.
NOM place ASP CL vase on.the.table
'(John) placed a vase (on the table).'
f. John bai le Ø Ø.
place ASP vase on.the.table
'John placed (a vase) (on the table).'
g. Ø bai le Ø zai zhuo-shang.
NOM place ASP ACC LOC table-on
'(John) placed (a vase) on the table.'
h. Ø bai le Ø Ø.
NOM place ASP vase on.the.table
'(John) placed (a vase) (on the table).'

5.1.1.6 Empty pronouns must be anteceded by discourse-prominent entities

We have seen that the referent of an empty pronoun in Chinese must be actively or accessibly identifiable (according to Lambrecht's (1994) classification) in 4.1.1, and this pragmatic generalization is also observed in an analogous manner in the case of Japanese

empty pronouns. Kameyama (1985) does not adopt Lambrecht's (1994) framework to differentiate various stages of activation or identifiability, but a similar, though more simplified, point has been raised in her work. To be exact, Kameyama distinguishes between *contextually-bound* and *contextually-unbound* entities. In a nutshell, a contextually-bound entity is 'old/given' in discourse by virtue of (i) a prior mention, (ii) being perceptually salient in the given context, or (iii) being a generic concept. In contrast, a contextually-unbound entity is informatively 'new' in the sense that (i) it has not been introduced in the discourse, (ii) it is not perceptually salient, or (iii) it is not a generic concept (Kameyama 1985: 122). Her notion of contextually-bound and contextually-unbound entities does not correspond to the difference between topic and focus, which has more to do with discourse functions. For example, Speaker B's replies in (294) and (295) below are two identical sentences, but the two entities involved (i.e., *John* and *Bill*) have opposite discourse statuses. *John* is contextually-bound, or [+ bound], in B's reply in (294), whereas the same entity *John* becomes contextually-unbound, or [- bound], in B's answer in (295). *Bill* in (294) is [- bound], and it becomes [+ bound] in (295).

(294) A: Who is John competing with?
B: John is competing with Bill.
(295) A: Who is competing with Bill?
B: John is competing with Bill.

Kameyama (1985: 126) points out that the distinction between contextually-bound and contextually-unbound entities has syntactic correlations in that an NP encoded as contextually-unbound must be neither a topic nor an empty pronoun, while a contextually-bound NP may be either a topic or an empty category. The dialogue in (296) below involves two NPs, *Rosa* and *Max*. The former is contextually-bound and is allowed to be null in the reply in (296b), whereas the latter is contextually-unbound and cannot be left unexpressed, or ungrammaticality arises.

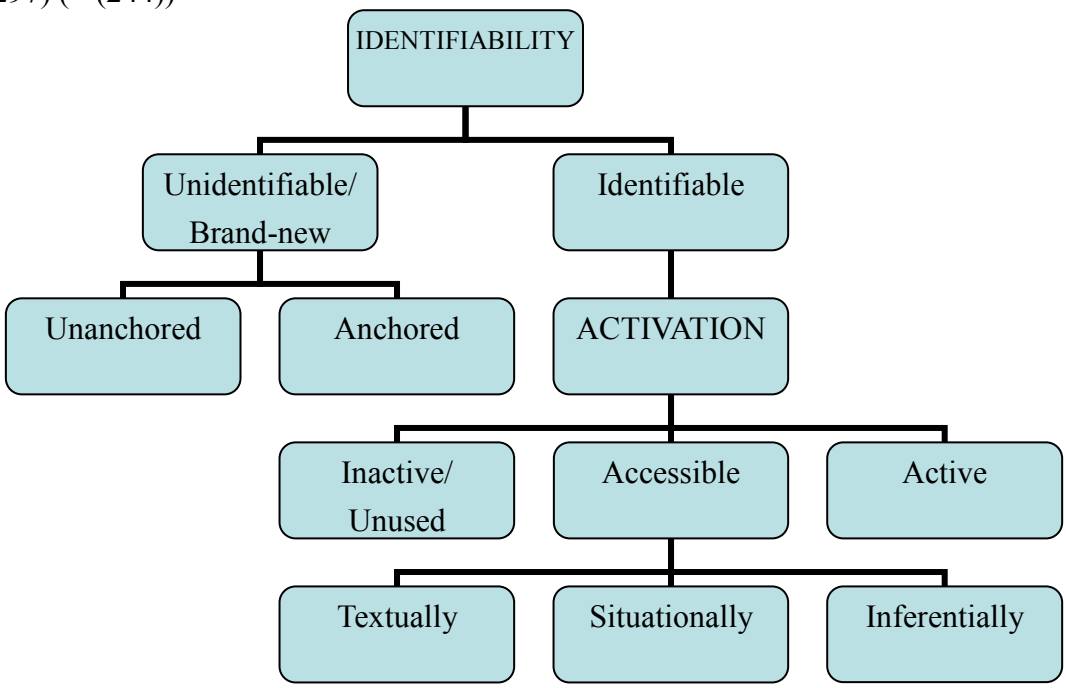
(296) (Kameyama 1985: 123, (47))

a. Dare ga/*wa Rosa o/*wa sagasite-iru no?
who NOM/TOP ACC/TOP is.looking.for Q
'Who is looking for Rosa?'

b. Max ga/*Ø Rosa o/Ø sagasite-iru n-da yo.
NOM ACC is.looking.for ASN RPT
'Max/*(Max) is looking for Rosa/(Rosa).'

Kameyama's observation is reminiscent of our proposed pragmatic generalization which states that the antecedent of an empty pronoun must be actively or accessibly identifiable. When it comes to formalizing different discourse properties at information structure, even though Lambrecht's (1994) diagram, shown below in (297), of identifiability and activation on which our pragmatic condition relies is a finer-grained proposal than Kameyama's (1985) two-way distinction of contextual boundness, the underlying affinity shared by the two views should still not be overlooked. To conclude, we previously established the pragmatic generalization for Chinese empty pronouns which states that the referent of an empty pronoun must be actively or accessibly identifiable, and the use of an empty pronoun whose referent is not actively or accessibly identifiable (or contextually-bound in Kameyama's terms) is disallowed. This contextual property is shared by both Chinese and Japanese empty pronouns. In other words, both languages require empty pronouns to be anteceded by discourse-prominent entities.

(297) (= (244))



5.1.1.7 Both languages show subject-subject obviation

Another pragmatic generalization shared by Japanese and Chinese zero anaphora is obviative reference between a matrix subject and an empty subordinate subject, which is a linguistic phenomenon termed as SUBJ-SUBJ obviation (Bresnan 1982). Cited in Kameyama (1985: 176), Bresnan (1982) proposes the Obviation Principle in (299) and uses SUBJ-SUBJ obviation to account for the referential difference in (298a) and (298b), in which *for to* is analyzed as the obviative clause in English.

(298) a. John$_i$ wished to Ø$_{i/*j}$ win the competition.
b. John$_i$ wished for him$_{*i/j}$ to win the competition.

(299) (Kameyama 1985: 176, (9))
Obviation Principle: if P is the pronominal SUBJ of an obviative clause C, and A is a potential antecedent of P and is the SUBJ of the minimal clause nucleus that properly contains C, P is or is not bound to A according to whether P is expressed or unexpressed.

Kameyama (1985: 182) observes that there is a general referential preference involving a matrix subject and a subordinate subject. In a context of coreference between the two, the subordinate subject is preferred to be null, while disjoint reference between the two prefers the subordinate subject to be overt, as the contrast between (300a) and (300b) exemplifies. Note that coreference between the matrix subject and the embedded overt subject in (300b) may be acceptable only when the latter carries a contrastive meaning, indicated by the exclamation mark on the index The question mark on the index in (300a) means that the reading, though no preferred, is marginally acceptable and not completely ruled out. The two examples in (301) behave just alike and can be analyzed in the same manner.

(300) (Kameyama 1985: 180, (18))

a. Mary$_i$ ga [Ø$_{i/?j}$ toohyoosuru koto] o nozonda (koto).
NOM NOM votes NOM ACC wished NOM
'(That) Mary$_i$ wished that (she$_i$/someone else) would vote.'

b. Mary$_i$ ga [kanozyo$_{!i/j}$ toohyoosuru koto] o nozonda (koto).
NOM she votes NOM ACC wished NOM
'(That) Mary$_i$ wished that she$_i$ would vote.'

(301) (Kameyama 1985: 181, (19a))

a. Ano otoko$_i$ ga [Ø$_{i/?j}$ byooki da] to omotteiru (koto).
that man NOM NOM ill is CMP thinks NOM
'(That) that man$_i$ thinks that (he$_i$/someone else) is ill.'

b. Ano otoko$_i$ ga [kare$_{!i/j}$ ga byooki da] to omotteiru (koto).
that man NOM he NOM ill is CMP thinks NOM
'(That) that man$_i$ thinks that (he$_i$) is ill.'

A similar scenario can be found in the following Chinese data, in which coreference between the matrix subject and the embedded empty subject is strongly favored over disjoint reference, which is, though, not completely impossible in (302a). On the other hand, in (302b), coreference between the matrix subject and the embedded overt pronoun can be obtained if the latter is contrastively used. Otherwise, disjoint reference would be the preferred reading, and because there is no gender difference in the spoken form of the

third person pronoun in Chinese, *ta* 'he/she', the gender of the overt pronoun in this reading of disjoint reference can be either.

(302) a. Mary$_i$ hen xiwang [Ø$_{i/?j}$ neng qu toupiao].
very wish NOM can go vote
'Mary$_i$ wished that (she$_i$/someone else) could vote.'
b. Mary$_i$ hen xiwang [ta$_{!i/j}$ neng qu toupiao].
very wish he/she can go vote
'Mary$_i$ wished that (he/she$_i$) could vote.'

Kameyama (1985: 189) proposes the following generalization regarding the above SUBJ-SUBJ obviation.

(303) (Kameyama 1985: 199)
Given two coreferential referring expressions R1 and R2, where R1 is in the clause immediately superordinate to the clause that contains R2, if R2 is an unexpressed subject, it is a tendency that R1 is also a subject, expressed or unexpressed. This may be schematized as: R2 [+U, +SUBJ] → R1 [+SUBJ], where (i) R1 corefers with R2, (ii) R2 is in the clause immediately subordinate to the clause that contains R1, and (iii) → indicates a preferred tendency.

Even though we did witness SUBJ-SUBJ obviation in the Japanese and Chinese examples, the proposed generalization in (303) still, at best, serves as a non-absolute tendency that may be overridden by syntax. For the same reason, the marked readings (indicated by ? or !) may still be available and not ruled out completely, unlike the English sentences in (298), which exhibits an example of true obviation. For instance, given a context as the one in (304), the SUBJ-SUBJ obviative tendency between the Chinese empty/overt subordinate pronouns is no longer attainable, because one of the syntactic constraints in our proposal requires an empty subordinate subject pronoun, such as the one in (304a), to be bound (in the extended sense) by an argument. Although coreference between *John* and Ø in (304a) is by all means possible and syntactically well-formed, such a reading is semantically odd, as indicated by the # mark.

(304) [John$_i$'s mother$_j$ has been ill for a long time ...]

a. John$_i$ hen xiwang [Ø$_{\#i/*j}$ gan-kwai hai-qilai].
very wish NOM quickly get.well
'John$_i$ wished that (he$_{\#i}$/she$_{*j}$) would get well soon.'

b. John$_i$ hen xiwang [ta$_{\#i/j}$ gan-kwai hai-qilai].
very wish he/she quickly get.well
'John$_i$ wished that he$_{\#i}$/she$_j$ would get well soon.'

The intended coreference between *John's mother* and the unexpressed pronoun in (304a) is unavailable because syntax does not allow it. The overt subordinate subject in (304b) is not subject to the aforementioned syntactic requirement, and therefore it can be coindexed with either antecedent, although coreference with *John's mother* will be the prominent reading. This exemplifies that although the proposed SUBJ-SUBJ obviative tendency in (303) is observed in the Japanese and Chinese data, it is not an absolute rule, as its effect can be overridden by other factors (e.g., syntax).

5.1.2 Differences

The above discussion presents what Chinese and Japanese zero anaphora have in common from the syntactic, pragmatic, and general aspects. Now we will look at the major differences that set the empty pronouns in these two languages apart.

5.1.2.1 Japanese empty subordinate subjects need not be anteceded by arguments

As a reminder, one of the syntactic requirements in our proposal for an empty subordinate subject pronoun in Chinese is that it must be bound (in the extended sense) by an argument function (e.g., SUBJ or OBJ). In other words, this syntactic rule forbids the use of a Chinese empty subordinate subject if it is bound by a non-argument function at *f*-structure, which is most likely to be TOP, especially when TOP and SUBJ differ. The example in (305), repeated from (176) in Chapter 3, explains how the empty subordinate subject cannot be coindexed with the topic *John*, even though coindexation with the

matrix subject, *John's mother*, is semantically odd (but syntactically acceptable).

(305) [It is John$_i$'s birthday tomorrow …]

Tade$_i$ mama$_j$ mai le hen duo [Ø$_{*i/\#j}$ mengmei-yi-qiu __$_k$ de] wanju$_k$.

his mother buy ASP very many NOM dream.to.have ACC RLTV toys

'His mother$_j$ bought many toys which (he$_{*i}$/she$_{\#j}$) dreams to own.'

On the other hand, an empty subordinate subject pronoun in Japanese appears to escape from this restriction. Let us turn to the following Japanese example in (306), which involves two occurrences of empty subordinate subjects. From the array of possible readings glossed below the full translation, we can see that neither empty subordinate subject is restricted by any argument or non-argument binding. Furthermore, the antecedents for the two empty pronouns may not even be identical in the Japanese example in (306), which is unusual, because normally the subject of an adjunct clause should be coindexed with the matrix subject, as in the English sentence and the Chinese sentence in (307) and (308), respectively.

(306) (Kameyama 1985: 48, (40))

Ø ganbareba Ø wakaru-yoo ni naru to kare wa itta.

NOM try.hard.if NOM understand to become CMP he TOP said

'He said that if Ø try/tries hard, Ø would understand.'

he	I'd	I
he	you'd	you
he	she'd	she
he	they'd	they
he$_i$	he$_i$	he$_i$
he$_i$	he$_i$	they
he$_i$	he$_j$	they
he$_i$	he$_j$	he$_j$

(307) John$_i$ always seems energetic even when Ø$_{i/*j}$ getting up early in the morning.

(308) Mei-tian zaoshang $Ø_{i/*j}$ qi-chuang shi, $John_i$ zongshi hen bu kaixin.
every-day morning NOM rise-bed time always very not happy
'John is always unhappy when (he) gets out of bed every morning.'

Now if we turn to exactly the same sentence in Chinese in (309) below, there is only one reading available, in which both empty categories must share an arbitrary referent as the unspecific *one*, or the generic *we/you* (as opposed to the anaphoric kind of *we/you*). Notice that none of these readings is referential and the reference between the two empty subjects must not be disjoint. *We* can be inclusively or exclusively used, depending on the context; the difference between the two usages is immaterial here.

(309) Ta_i shuo ruguo $Ø_{arb}$ nuli shi, $Ø_{arb}$ jiu hui liaojie.
he say if NOM diligently try NOM just would understand
'He said if (one/we/you) tries hard, (one/we/you) would understand.'

Given the contrast between (306) and (309) and the observation that the Chinese version cannot have the same range of readings as the Japanese version does, we can conclude that empty subordinate subject pronouns in Japanese are not subject to the same syntactic requirement, namely, bound (in the extended sense) by an argument as their Chinese counterparts are.

5.1.2.2 Only empty pronouns in Japanese are subject to Kameyama's *General Identification Constraint*

So far we have observed that Japanese zero anaphora appears to be only minimally syntactically restricted, especially when compared to its Chinese counterpart, which accords with Kameyama's (1985) view that Japanese zero anaphora is essentially constrained by pragmatics and discourse and less so by syntax. What really matters in resolving empty pronouns in Japanese, according to Kameyama's (1985) analysis, comes from where the *speaker identification* lies. Kameyama's *speaker identification* (or simply

identification) is derived from Kuno and Kaburagi's (1977) *empathy*, which is defined as 'the speaker's identification, with varying degrees, with a person who participates in the event that he describes in a sentence' (Kuno and Kaburagi 1977: 628). Based on Kuno and Kaburagi (1977), Kameyama (1985: 141) characterizes each predicate as selecting one of its own argument(s) to be its identification locus (or annotated as Ident), a position whose viewpoint the speaker automatically identifies with (to varying degrees). Since a speaker often adopts the viewpoint of an animate being, it follows naturally, as Kuno and Kaburagi (1977) assume and Kameyama (1985) agrees, that only those predicates with animate, human in particular, arguments may have identification loci. Some predicates place their identification loci on subjects and some on objects, although the unmarked position for an identification locus is the subject, and the object is a marked position.

Summarized from Kameyama (1985: 141), in addition to those identification loci associated with different Ident-specific predicates (e.g., Ident on its subject or object), some nominal expressions such as the first person pronoun in Japanese, *boku* 'I', and the reflexive pronoun, *zibun* 'self', also automatically carry the speaker's point of view and, therefore, they are in identification locus. There are also identification-neutral sentences, annotated as [Ø Ident], which involve inanimate arguments or an 'objective' tone of speech allowing for an 'identification-neutral', [51] or 'empathy-free', reference. Importantly, if there is a disagreement between an identification-coding phrase (e.g., I or a

[51] Cited from Kameyama (1985: 193), two instances of identification-neutral, or empathy-free, sentences are given in (i) and (ii) below.

(i) (Ohso 1976: 37, (62))
Kireina onnanoko$_i$ ga __$_j$ Ø$_i$ kisusi-yooto-sita otoko$_j$ ni kamituita.
pretty girl NOM NOM ACC kiss-try-PST man to bit
'A pretty girl$_i$ bit a man who tried to kiss (her$_i$).'

(ii) (Ohso 1976: 37, (63))
Sono onna$_i$ wa __$_j$ Ø$_i$ bikoosite-ita otoko$_j$ ni koros-are-masi-ta.
that woman TOP NOM ACC was-shadowing man by kill-PASS-POL-PST
'That woman$_i$ was killed by the man who was shadowing (her$_i$).'

The above two sentences seem to pose a problem for Ohso's (1976) Empathy Constraint, which states that 'when the speaker identifies with the viewpoint of the referent of two coreferential NPs, the second NP can be zero' (Kameyama 1985: 190), because the omitted objects in both instances appear to be [- Ident], if one assumes the predicates 'to kiss' and 'to shadow' have an unmarked identification locus on the subject. Ohso tries to accommodate this problem by arguing that these two sentences present an objective point of view, which then leads to an identification-neutral or empathy-free reference. But as Kameyama correctly points out, it is extremely unclear how one can draw the line between an objective and non-objective viewpoint when describing something. Ohso's explanation is unprincipled and unsatisfying.

subject) and the identification locus assigned by the predicate, the sentence immediately becomes ill-formed because now this sentence is loaded with two identification loci, when there should only be one per sentence. Take the Japanese data in (310) and (311) for instance. *Ageta* and *kureta* are two verbs that both have the meaning of 'giving', but the former has an unmarked Ident on its subject whereas the latter has a marked Ident on the secondary object.

(310) (Kameyama 1985: 142, (79a))

Boku wa Taro ni hon o ageta/*kureta.
I TOP DAT book ACC gave gave
'I gave Taro a book.'
boku: [+ Ident]
ageta: Ident = Subject
kureta: Ident = Secondary Object

(311) (Kameyama 1985: 142, (79b))

Taro wa boku ni hon o *ageta/kureta.
TOP I DAT book ACC gave gave
'Taro gave me a book.'
boku: [+ Ident]
ageta: Ident = Subject; kureta: Ident = Secondary Object

In (310), the ungrammaticality of choosing *kureta* as the predicate comes from the identification conflict between the secondary object *Taro* and the Ident-coding nominal *boku* 'I', annotated as [+ Ident]. Namely, the speaker's identification should not be on *Taro* as well as on *boku* 'I'. On the other hand, choosing *ageta* as the predicate in (310) will not result in any identification conflict because the predicate places its Ident on the subject, which itself is also an Ident-loaded phrase. Likewise, the same analysis applies to (311).

The speaker identification affects how an empty pronoun in Japanese is permitted, but such effect is not witnessed in Chinese. According to Kameyama (1985), Japanese empty pronouns are subject to the *General Identification Constraint* given below.

(312) (Kameyama 1985: 200, (52))

General Identification Constraint on zero pronominals:

Given two coreferential referring expressions R1 and R2 in a sentence, where R1 is in the clause immediately superordinate to the clause that contains R2, and if either (a) R2 is within the nucleus of main predicate of the superordinate clause containing R1, or (b) R1 refers to an individual who is 'strongly identified with' (i.e., a logophoric individual or the speaker of the utterance), then R2 can be unexpressed only if R1 and R2 have the same identification property. An identification property may be (i) identified [+ Ident], (ii) non-identified [- Ident] in an identification-loaded sentence, or (iii) occurs in an identification-neutral [Ø Ident] sentence or is inanimate, which is then irrelevant to identification. The constraint can be schematized as follows:

R2 [+ Unexpressed] → R1 [α Ident] and R2 [α Ident], where α is +, -, or Ø.

The General Identification Constraint correctly predicts why the following Japanese examples in (313) and (314) permit the use of an empty pronoun and why only a certain pattern of coreference is available. Particularly in (314), where the two sentences are structurally parallel to each other but with very different readings, the reason the empty subject in (314a) is coreferential with the (matrix) subject, and the empty subject in (314b) with the object, can only come from a non-syntactic explanation such as where the identification locus lies. In (313), *John* bears the unmarked identification locus, and so does the empty subordinate object pronoun, because the embedded predicate, *tazunete-kuru* 'come visit', has a marked Ident on its object. Similar to the contrast we saw in (310) and (311), the only difference between (314a) and (314b) is the embedded predicate – *yonda* and *yonde-ita*, both meaning 'to read' but with a different identification locus.

(313) (Kameyama 1985: 191, (36a))

John$_i$	wa	Bill	ga	Ø$_i$	ituka	tazunete-kuru	to	omotte-iru.
	TOP		NOM	ACC	one.day	visit-come	CMP	thinks

'John$_i$ thinks that Bill will one day come visit (him$_i$).'

John: [+ Ident]

tazunete-kuru: Ident = Object

Ø: [+ Ident]

(314) (Kameyama 1985: 195, (43))

a. John$_i$ wa Bill ni Ø$_i$ kinoo __$_j$ yonda hon$_j$ no koto o kiita.
TOP DAT NOM yesterday ACC read book of matter ACC asked
'John$_i$ asked Bill about the book (he$_i$) read yesterday.'
John: [+ Ident]
yonda: Ident = Subject
Ø: [+ Ident]

b. John wa Bill$_i$ ni Ø$_i$ kinoo __$_j$ yonde-ita hon$_j$ no koto o kiita.
TOP DAT NOM yesterday ACC read book of matter ACC asked
'John asked Bill$_i$ about the book (he$_i$) read yesterday.'
John: [+ Ident]
Bill: [- Ident]
yonde-ita: Ident = Object
Ø: [- Ident]

Kameyama's General Identification Constraint neatly explains why the empty pronouns in (313) and (314) are allowed and how the possible readings are arrived, and this rule can be applied to all intrasentential empty pronouns in Japanese. However, such a generalization is not observed in Chinese. First of all, we can not find equivalent examples between two forms of verbs which essentially mean the same but with different identification loci, such as the contrast between *ageta* and *kureta* in (310) and (311), which both have the meaning of 'giving'. A Chinese version of the Japanese example in (314), demonstrated in (315) below, shows that the empty subordinate subject can be coindexed with either *John* or *Bill*.

(315) John$_i$ wen Bill$_j$ guanyu zuotian Ø$_{i/j}$ du __$_k$ de na ben shu$_k$
ask about yesterday NOM read ACC RLTV that CL book
de shiqing.
of matter
'John$_i$ asked Bill$_j$ about that book (he$_{i/j}$) read yesterday.'

Our syntactic rule predicts that both *John* and *Bill* in (315) are qualified to be the possible antecedent for the empty subordinate subject since they are both arguments, and either

coreference is pragmatically plausible. The contrast between (314a) and (314b) cannot be obtained in the case of Chinese, and the speaker's identification appears unrelated when we try to interpret Chinese empty pronouns. Therefore, we conclude that while Japanese intrasentential empty pronouns may follow the General Identification Constraint neatly, as proposed by Kameyama, Chinese empty pronouns are free from such a condition.

5.1.2.3 Chinese zero anaphora escapes Kameyama's Centering rules

The above establishment of the speaker's identification, or the identification locus, has another implication in resolving Japanese zero anaphora, and it has to do with *Centering*. To begin with, a 'center' is defined as an entity, a referring expression that can be identified. A center may either be a backward-looking center (Cb) or a forward-looking center (Cf). Because the Cb is the most significant entity under discussion and can be the linkage across sentences, and the Cf can be an array of discourse referents, any given utterance should be associated with zero or one Cb and any number of Cfs (Kameyama 1985: 93). The term *centering* is then understood to be a transitional process within texts, based upon the way discourse referents are introduced and discussed. It can also determine which entity (or entities) bears discourse saliency during the transition (Poesio et al. 2000). Now how is centering related to the resolution of zero anaphora? Kameyama (1985: 192) proposes the following Centering rules for intersentential empty pronouns in Japanese.[52]

(316) The Center-establishment rule (Kameyama 1985: 164):
When one of the Cfs in the previous discourse is made into the Cb of the current utterance, a zero pronominal is used. There is a basic preference order among Cfs in an utterance as to which is most likely to be the Cb in the next utterance. The order is TOPIC, SUBJECT, OBJECT, and OTHERS (e.g., adjunct).

(317) The Center-retention rule (Kameyama 1985: 165):
If the Cb of the current utterance is the same as the Cb of the previous utterance, a zero pronominal is used.

[52] See Walker et al. (1994) for a thorough discussion of Centering in Japanese.

(318) (Kameyama 1985: 192, (88))

When an utterance contains more than one zero pronominal in the middle of a discourse, the one that shares one or more of the following properties (in the order given) with the previous Cb-encoder is preferred to retain the same Cb:

1. Ident, 2. SUBJ, and 3. nonIdent-nonSUBJ

The Cb-establishment rule in (316), demonstrated by the Japanese dialogue between Speaker A and B in (319), is reminiscent of one of our pragmatic generalizations introduced in 4.1.2, which states that an intersentential pronoun may be empty if it bears the discourse function as a topic, primary or secondary. However, our pragmatic condition does not prescribe that a topical pronoun *must* be in the form of an unexpressed argument in Chinese, unlike the stricter requirement held by Kameyama's Cb-establishment rule in (316).

(319) (Kameyama 1985: 132, (60))

A: Rosa_i wa dare o matte-iru no?
TOP who ACC is-waiting Q
'Who is Rosa waiting for?'

B: Ø_i Mary_j o matte-iru no yo.
NOM ACC is-waiting ASN RPT
'(She) is waiting for Mary.'
[Cb = Rosa, Cf = Mary]

Next, the Center-retention rule in (317) also appears to be very similar to the Topic-NP Deletion Rule (Tsao 1979, 1990), which states that when several sentences share the same topic and form a topic chain, the topic in each of the subsequent clauses may be deleted and the topic in the first clause remains. The generalization proposed in (318) concerns the resolution of Japanese zero anaphora when there are multiple occurrences of empty pronouns. The effect of (318) comes into play when there are multiple empty pronouns and we are not sure of their references regarding the previous discourse. Imagine that the Japanese sentence in the next example (320) is uttered right after A and B's conversation in (319), and now we are left to decide which of the two unexpressed argument is the Cb

in this current discourse in (320).

(320) $Ø_i$ $Ø_j$ yuusyoku ni syootaisita no yo.
NOM ACC supper to invited ASN RPT
'(Rosa) invited (Mary) to dinner.'

The previous Cb-encoder in (319) is *Rosa*, which itself is a subject, and according to Kameyama's generalization in (318), between an empty subject and an empty object, we know that the one sharing one or more properties (in the order of Ident, SUBJ, and nonIdent-nonSUBJ) with the previous Cb-encoder is preferred to retain the same Cb status. Therefore, we arrive at the solution that it is the empty subject that retains *Rosa*'s Cb status, and the strong reading of (320) is for *Rosa* to be the subject, and *Mary* the object. Kameyama's generalization in (318) provides a neat solution to how multiple empty pronouns' references are decided in a connected discourse. However, since the speaker's identification or the identification locus of a predicate is irrelevant in accounting for the interpretation of Chinese empty pronouns, zero anaphora in Chinese is not constrained by the aforementioned generalization, either. This claim is supported by the Chinese examples in (321) and (322) below, the latter of which is an utterance that immediately follows the previous dialogue in (321), analogous to the Japanese examples seen in (319) and (320).

(321) A: $John_i$ zai deng shei?
PROG wait who
'Who is John waiting for?'
B: $Ta_i/Ø_i$ zai deng $Bill_j$.
he PROG wait
'He is waiting for Bill.'

(322) $Ø_{i/j}$ yue le $Ø_{j/i}$ yiqi chi wanfan.
NOM invite ASP ACC together eat dinner
'(He (John/Bill)) invited (him (Bill/John)) to have dinner together.'

The subject of *deng* 'wait' in B's speech in (321) can be overt or empty; both usages are equally natural. The interpretations of the two empty categories in (322), however, differ from those in the Japanese sentence in (320) in that (322) can be two-ways ambiguous, with no strong preference as to who should be the subject of *yue* 'invite'. The Cb status of *John* from B's speech in (321) need not necessarily be retained in (322), contra the Japanese example illustrated in (320).

5.1.2.4 Japanese zero anaphora allows indefinite antecedence and antecedence in focus

The last difference between Japanese and Chinese zero anaphora concerns a certain unusual antecedence of empty pronouns that is allowed in Japanese. Generally speaking, being definite or generic in reference is a prerequisite condition in order to antecede an unexpressed argument, because a non-generic, indefinite referent is typically not identifiable whether it is anchored or unanchored. Although a definite antecedent does not necessarily sanction the use of an empty category because there are still various stages of activation at information structure (e.g., inactive, accessible, or active), a non-generic, indefinite referent is certainly an undesirable antecedent as far as Chinese empty pronouns are concerned, especially when taking into consideration our pragmatic generalizations which require an actively, accessibly identifiable, or a topical antecedent to license the use of an empty pronoun. However, according to our observation of the Japanese data taken from Kameyama (1985) in (323) and (324) below, it *is* possible for an empty pronoun to be bound by an indefinite antecedent in Japanese. Coindexation as such cannot be witnessed in the case of Chinese zero anaphora, as the ungrammatical occurrence of the empty subject in (325) shows.

(323) (Kameyama 1985: 224, (81b))

Aru otoko$_i$ ga [Ø$_{i/?j}$ oosama ni manekareta] koto o hanasite-i-ta.
a.certain man NOM NOM king by was.invited NOM ACC was.telling
'A certain man$_i$ was telling that (he$_i$) was invited by the king.'

(324) (Kameyama 1985: 224, (81c))
Aru otoko$_i$ ga [Ø$_{i/?j}$ natuite-iru] raion o uti-korosite-simat-ta.
a.certain man NOM DAT is.taking.to lion ACC had-shot-to-death
'A certain man$_i$ had shot the lion that was taking (him$_i$) to death.'

(325) You ge ren$_i$ gaosu wo [ta$_i$/*Ø$_i$ gang kandao nuhuang le].
there.is CL person tell me he just.now see Queen ASP
'A certain person$_i$ just told me that he$_i$/*(he$_i$) just saw the Queen.'

Not only does Japanese zero anaphora allow indefinite antecedence but, unusually, also antecedence in focus, e.g., *wh*-NPs in questions. As the Japanese sentences in (326) and (327) illustrate, it is clearly unproblematic for the empty subordinate pronoun to be coindexed with the *wh*-NP, which bears the focus of the sentence.

(326) (Kameyama 1985: 229, (91a))
Dono otoko$_i$ ga [Ø$_{i/?j}$ siawase-da] to omotteiru ka?
which man NOM NOM happy-is CMP thinks Q
'Which man$_i$ thinks that (he$_i$) is happy?'

(327) (Kameyama 1985: 229, (91b))
Dono otoko$_i$ ga [Mary ga Ø$_{i/j}$ aisiteiru] to omotteiru ka?
which man NOM NOM ACC loves CMP thinks Q
'Which man$_i$ thinks that Mary loves (him$_i$)?'

In contrast, a focal referent such as a *wh*-NP in a question can never antecede an unexpressed argument in Chinese, as the unacceptable use of the empty pronoun in (328) demonstrates.

(328) Na ge ren$_i$ juede [ziji$_i$/*Ø$_i$ yiding hui ying]?
which CL person feel self/NOM definitely will win
'Who$_i$ feels that he$_i$/*(he$_i$) will definitely win?'

In virtue of the above discussion, we come to the conclusion that an antecedent for a Japanese empty pronoun may be permitted to be indefinite in reference or even be a focus. On the other hand, neither indefinite nor focal referents are allowed to antecede empty

pronouns in Chinese because the former are unidentifiable at information structure while the latter are simply unpredictable, which is typical of foci. This contrast is another difference between Japanese and Chinese zero anaphora. To summarize, we have shown that Japanese and Chinese empty pronouns are subject to syntactic as well as pragmatic constraints; they are by no means rule-free. While Japanese and Chinese empty pronouns do share quite a few (pragmatics and syntactic) similarities, they are also under different language-specific constraints, and the distributions of zero anaphora in both languages are not parallel.

5.2 An examination of empty pronouns in Japanese and Chinese texts

After presenting the similarities and differences of how empty pronouns are constrained in Japanese and Chinese, this section continues to examine how Japanese and Chinese empty pronouns are actually used, coherently, in the same texts and what kind of empty pronoun is selected so that we can draw further comparison.

In order to scrutinize how a pronoun is realized phonetically or not in the same text in the two languages we seek to compare, we will first take one chapter of the first novel in the renowned *Harry Potter* series written by J. K. Rowling – *Harry Potter and the Philosopher's Stone* (1997) – to be the source text, and then through this English medium, we set out to explore how the English text is translated into Japanese and Chinese by studying the same chapters adopted from the official translated books by Matsuoka (2003) in Japanese, and by Peng (2000) in Mandarin Chinese. The choice of the *Harry Potter* book is for its huge popularity and global circulation and, therefore, a sound translated version of high standard into different languages should be expected. Since English does not permit *pro*-drop while Japanese and Chinese, both being topic-prominent languages, allow empty subjects and objects, we expect that there will be plenty of empty pronouns in the translated versions of Japanese and Chinese. To be exact, we have chosen Chapter 13, titled *Nicolas Flamel*. The choice of Chapter 13 is for no particular reason other than economy and efficiency. Due to the limited space allowed in the current thesis and also to avoid an overload of similar examples which could shift the focus of the comparison,

Chapter 13, being the shortest chapter in the book, is selected as our token text, and it already carries ample instances of empty pronouns of different kinds.

For *pro*-drop languages which employ empty pronouns extensively such as Japanese and Chinese, it would appear to be odd and unnatural if no empty pronouns were used in a continuous text even if this text is translated from a *non*-pro-drop language. The texts we study involve short dialogues as well as narratives, and they serve the purpose for us to observe the usage pattern of both intersentential and intrasentential zero anaphora. It is possible for the translated versions in Japanese and Chinese to contain slight language-specific stylistic differences and individual translating preferences which might affect the decision of whether to apply an overt pronoun or an empty one. We will assume (i) that such variants in the two target languages to be compared are not significant enough to distort any bona fide instance of zero anaphora, and (ii) that all occurrences of zero anaphora observed in the translated texts of both languages are intended to be there, legitimately, as a means of organizing a continuous discourse, creating a natural flow, and maintaining a coherent tone of sentences without losing the spirit of the original English version. Finally, the types of empty categories we seek to examine are only empty pronouns appearing as finite subjects and objects, which exclude controlled PRO. To simplify the matter, possessives and reflexives are also excluded for their infrequency and less relevance to our focus on finite pronouns.

We have shown in 5.1.2 that there are a number of differences in how Japanese and Chinese empty pronouns behave. Japanese zero anaphora appears to be freer, less syntactically constrained, and more pragmatically conditioned. Based on this assumption, we would predict the percentage of zero anaphora employment to be much higher in the translated Japanese chapter than that in the Chinese chapter, because fewer constraints would imply a freer distribution and higher frequency (if we set the factor influenced by individual translators' preferences aside). Such a prediction is indeed borne out. In the official Chinese version of Chapter 13 of *Harry Potter and the Philosopher's Stone*, we find 50 occurrences of empty pronouns; while in the official Japanese version of the same chapter, there are 125 occurrences of empty pronouns, 2.5 times the figure in the Chinese version. Among the 50 occurrences of empty pronouns in the Chinese version, 39 of them

are empty subjects and 11 of them are empty objects. The more frequent use of an empty subject than an empty object is also witnessed in the Japanese version, in which there are 94 empty subjects and 31 empty objects. The ratio of empty subjects to empty objects is roughly three to one in both Japanese (94 : 31) and Chinese (39 : 11) translations. Among the 39 occurrences of empty subjects in Chinese, 25 of them are matrix subjects and 14 of them are subordinate subjects. In contrast, there are 78 matrix subjects and 16 subordinate subjects in the Japanese data. All these figures tell us that overall in both languages, the use of empty subjects is more common than that of empty objects, and that matrix subjects occur more frequently than subordinate subjects. The fact that more empty subjects are used than empty objects in both sets of data may suggest that the role played by a subject is generally more prominent than the object role, and, therefore, as far as reference recoverability is concerned, a subject could be more easily and readily omitted, without sacrificing readers' understanding of the passage. Because the majority of the omitted references, whether subject or object, are topical in both sets of data, the strong association between an empty pronoun and topicality is again confirmed. Topicality indeed plays a crucial part in licensing the use of empty pronouns in Japanese and in Chinese. The higher frequency of an empty subject occurrence than an empty object occurrence also suggests that subject tends to be more likely to be topical than object. On the other hand, overt pronouns are still in frequent use in both translations. With the figure of empty pronouns much higher in Japanese than in Chinese, one would expect the figure of overt pronouns to be higher in Chinese than in Japanese. Such an expectation is met. We have found 40 occurrences of overt pronouns in the Japanese data, 28 of which are subjects and 12 of which are objects. In contrast, 235 overt pronouns are found in the Chinese data, made up of 190 subjects and 45 objects. The total figure of pronouns (combining overt and empty pronouns) is 165 in Japanese and 285 in Chinese.

(329) Summary table of the number and the type of pronoun found in the Chinese and Japanese data:

	Empty pronouns	Empty subjects	Empty matrix subjects	Empty subordinate subjects	Empty objects	Overt pronouns	Overt subjects	Overt objects
Chinese	50	39	25	14	11	235	190	45
Japanese	125	94	78	16	31	40	28	12

Now let us turn to the individual data collected from the authorized translation published in Japan and in Taiwan. All examples we are about to review in the following discussion will be the translated sentences taken directly from the books after our Romanization. In the Mandarin Chinese version, we observe the following six types of empty pronouns: (i) referential matrix subjects; (ii) referential subordinate subjects; (iii) referential objects; (iv) situationally-evoked deictic subjects; (v) non-referential, indefinite objects; and (vi) dummy subjects. In the Japanese version of the same chapter, the number of empty pronouns used is 2.5 times greater than that in the Chinese version, even though the types of empty pronouns found in the Japanese text resemble those found in the Chinese text. They include (i) referential matrix subjects; (ii) referential subordinate subjects; (iii) referential objects; (iv) non-referential, indefinite subjects; and (v) dummy subjects. (330) demonstrates the use of an empty referential matrix subject in Chinese, in which the empty subject refers to the speaker himself when he apologizes ironically to another person he deliberately pokes in the back of the head. The Japanese sentence in (331) illustrates two occurrences of empty matrix subjects. They share the same referent who has been introduced in the prior discourse and readers have been familiar with. The two consecutive empty subjects appear to have formed a topic chain, a phenomenon also observed in Chinese.

(330) *a referential matrix subject*

Ø gangcai mei zhuyidao ni.
NOM just.now not pay.attention.to you
'(I) didn't see you just now.'

(331) *referential matrix subjects*

Ø maiban akumu ni unas-are-ta. Ø yume o
NOM every.night nightmare from suffer-PASS-PST NOM dream ACC
nandomo kurikaeshi mi-ta …
over.and.over repeatedly see-PST
'(He/Harry) suffered from nightmares every night. Over and over again (he) dreamed about …'

Secondly, an empty referential subordinate subject in Chinese is shown in (332) below, in which the empty subject is anteceded by the argument NP, *Harry* himself, within the same sentence, in accordance with our proposed syntactic constraint which requires an empty subordinate subject in Chinese to be bound in the extended sense by an argument function (e.g., SUBJ or OBJ).

(332) *a referential subordinate subject*

Hali$_i$ ye faxian, [ta$_i$ lian qiu lian de yue lei, Ø$_i$ wanshang jiu
Harry also find.out he practice ball practice until more tired NOM at.night just
yue bu rongyi zuo emeng].
more not likely do nightmare
'Harry also found that the more tired he was after training, the less likely (he) had nightmares at night.'

Along the same line, the Japanese example in (333) includes two empty subordinate subjects, both of which are referentially used and share the same antecedent. In comparison with the Chinese version in (332), it is worth noting that replacing one or both empty subjects with overt pronouns in (333) would lead to unnaturalness. However, in the case of the Chinese sentence in (332), not only is it possible for the subject of *lian* 'to practice' to be null; a replacement of an overt pronoun in either or both subordinate subject positions would also be unproblematic.

(333) *referential subordinate subjects*

Ø$_i$ renshu de tsukare-ta ato wa, Ø$_i$ amari akumu o
NOM practice with tire-PST after TOP NOM fewer nightmare ACC
minakunarutoiu-nomo Harry$_i$ wa ishikisite-i-ta.
see-NEG TOP be.conscious-PROG-PST
'Harry$_i$ found that (he$_i$) had fewer nightmares when (he$_i$) was tired out after training.'

It has been noted in the comparative study in 5.1.2.1 that an empty subordinate subject in Japanese may escape the restriction imposed on Chinese empty subordinate subjects, which rules that they must be bound (in the extended sense) by an argument function. Even though the Japanese empty subordinate subject in (333) may seem to follow such a syntactic constraint, the empty subordinate subject of the predicate *tsuketeiru* 'following' in (334) obviously violates this rule, because the antecedent (i.e., *Snape*) of the said empty subordinate subject is not an argument. *Snape*, though mentioned briefly in the previous context and familiar to readers, is clearly a non-argument antecedent as far as the empty subordinate subject in (334) is concerned.

(334) Harry$_i$ ga hitoribochi-ni-nat-ta tokini Ø$_i$ tsukamaeyoto
NOM on.his.own-to-become-PST when ACC trying.to.catch
Ø$_{Snape}$ Ø$_i$ ato o tsukete-i-ru nodewanaikato Ø$_i$
NOM ACC after ACC follow-PROG-PRES wonder NOM
omoukotoga tokidoki-atta.
think sometimes-was
'At times, Harry$_i$ even wondered whether (Snape) was following (him$_i$), trying to catch (him$_i$) on his own.'

Thirdly, an empty referential object in Chinese is exemplified in (335). The object of *kandao* 'see/read' in (335) refers a few clauses back to the name *Flamel* (or the person it represents), which then becomes a topic under discussion after being introduced. In other words, *Flamel*'s status at information structure is not only actively identifiable, but also topical by the time the empty object is used. Such intersentential coindexation between the empty object and a topical antecedent is well-formed and falls under the prediction of

one of our pragmatic conditions, which states that a pronoun may be empty if it bears the discourse function of topic, primary or secondary. Here it clearly acts as a primary topic that connects the three clauses in (335).

(335) *a referential object*

[I've found Flamel$_i$! I *told* you I'd read the name$_i$ somewhere before …]

Wo shi zai zuo huoche dao zhe lai de shihou kandao Ø$_i$ de.
I COP at sit train to here come of time see ACC PART
'It was when on the train coming here I read (it/him).'

Below we have a Japanese example involving an empty referential object in (336), in which the object of the predicate *tell* is not present, but can be recovered from the prior discourse as the two people *Harry* speaks to, *Ron* and *Hermione*, whose referential status at information structure would at least be accessibly or actively identifiable. Note that even though the character *Snape* is an identifiable discourse referent with whom readers are familiar, the subject position occupied by *Snape* cannot be filled by an empty pronoun.

(336) a *referential object*

Hokanohito ni kik-are-nai yoni kogoe de Harry wa Ø
other.people by hear-PASS-NEG so.that low.voice with TOP ACC
Snape ga totsuzen Quidditch no shinpan o yaritaito iidashi-ta toiu
NOM suddenly GEN referee ACC desire.to start.say-PST about
hukitsuna news o tsutae-ta.
sinister news ACC tell-PST
'Speaking quietly so that no one else would hear, Harry told (them/the other two) about the sinister news that Snape suddenly wanted to be a Quidditch referee.'

Subsequently, an instance of an empty deictic pronoun in Chinese is illustrated in (337) below, in which the empty subject has a non-linguistic, text-independent antecedent – a certain passage in the book the speaker of (337) wants her hearers to read so she points out where this passage is to them. However, this sort of situationally-evoked

empty deictic pronoun, which often takes a non-linguistic antecedent, is not the focus of our study in the thesis because the distribution of deictic pronouns, whether they are empty or not, has less to do with the syntax of a language and more with extralinguistic context. We have little to say about non-linguistic antecedence. There is no empty deictic pronoun found in the Japanese text, but we believe that this usage is available in Japanese, just not in this particular text we have chosen.

(337) *a deictic subject*

[Oh, honestly, don't you two read? Look – read that …]
Ø jiu zai zher.
NOM just in here
'(A passage in the book) is just here.'

Our next example in (338) below concerns the use of a non-referential, indefinite empty object in Chinese. The reference of such pronouns is undetermined and does not have a referential antecedent, reflected by the index *arb* (for arbitrary) on the pronoun in the example. We discussed in 3.2.7 that according to Lebeaux (1982), such indefinite or arbitrary empty pronouns are different from their referential counterparts, and $Ø_{arb}$ is bound by a universal operator that gives rise to an arbitrary or a generic interpretation. The exact string of words *wo jiu zhidao* in (338)[53] has a very prominent idiomatic meaning as *I knew it!* in English, but the object of *zhidao* 'know' (used in the exactly same sense as in *I knew it!* in English) is never expressed in Chinese. If an overt object

[53] Notice that because of the presence of the emphatic adverb *jiu* 'just', it is not possible to analyze the sequence of words in (338) as not involving any empty object, i.e., treating the main verb *zhidao* 'know' as intransitive, with a syntactic structure similar to that in (ib) below. Both (ia) and (ib) are suitable answers to the question in the square brackets. Even though (338) and (ib) below are indeed structurally parallel to each other and the selection of words is almost identical, the two sentences are not on a par, but distinctively different as far as their semantic meanings are concerned.

(i) [Do you know this person$_i$?]
a. Wo zhidao ta$_i$/Ø$_i$.
I know him
'I know him/(him).'
b. Wo zhidao.
I know
'I know.'

pronoun is used or the subject is changed to anything but a first person singular pronoun, the idiomatic meaning of 'I knew it' would be lost.[54]

(338) *a non-referential, indefinite object*

[At last she found what she was looking for, and then she said the following.]
Wo jiu zhidao $Ø_{arb}$! Wo jiu zhidao $Ø_{arb}$!
I just know ACC I just know ACC
'I knew (it)! I knew (it)!'

(338) is Hermione's excited utterance right after she discovered something from a book. The reference of this empty object is neither definite nor referential to her hearers. This exact sequence of words *wo jiu zhidao* cannot to be taken to mean 'I knew', 'I know', or 'I know it'. None of these alternative replies would work as a suitable comment in the context of (338), where she feels that she already knew about something all along, likely just as a hunch.

The fourth type of empty pronoun found in the Japanese data – non-referential, indefinite subject – is shown in (339) below, which is the first occurrence among the three. The said subject has an arbitrary reading and the reference is uncontrolled. Empty pronouns of arbitrary or indefinite references are also observed in Chinese, although it is not found in the Chinese counterpart of this exact sentence.

[54] (i) has an idiomatic meaning of 'I knew it', just like its English counterpart. If the subject is changed to anything but a first person singular pronoun, as in (ii), this idiomatic reading will be lost and the sentence becomes unacceptable.

(i) Wo jiu zhidao Ø.
I just know ACC
'I just knew (it)!'
(ii) *Ta/tamen/ni/women jiu zhidao Ø.
he/they/you/we just know ACC
'(He/they/you) just know(s).'

(339) *non-referential*, *indefinite subjects*

[And no wonder (we) couldn't find (Flamel$_i$) in that *Study of Recent Developments in Wizardry* …]

Datte $Ø_{arb}$ 665-saija genmitsuni wa $Ø_i$ saikin to $Ø_i$ ienaiyona.
because NOM 665-years.old strictly TOP NOM recent and NOM can't.say
'Because (one) can't say (he$_i$) is exactly recent if (he$_i$) is 665 years old.'

Lastly, we observe the occurrence of dummy subjects in both sets of Chinese and Japanese data, such as those in (340), and in (341), respectively.

(340) *dummy subjects*

a. Ø shi Mafen.
NOM COP Malfoy
'It was Malfoy.'

b. Ø xian shi Hali Pote$_i$, ta$_i$ mei you fu-mu.
NOM first COP Harry Potter he not have father-mother.
'First (it) is Harry Potter; he has no parents.'

(341) *dummy subjects*

a. Ø bokuno sei janai.
NOM my fault is.not
'(It) is not my fault.'

b. Ø mo kurakunari-kakatteita-ga …
NOM already got.dark-almost-but
'(It) was almost dark …'

Generally speaking, neither Chinese nor Japanese employ dummy subjects since they permit subjects to be empty, and the structural need to fill a subject position explicitly with words with little semantic content like *it*, as in English, is lacking in both languages. When it comes to describing weather or time, neither language requires the presence of a dummy subject such as *it*. Meteorological and time-telling sentences in both languages are most likely to be subjectless, although this usage of lacking a dummy subject is not limited to weather or time, as (340) and (341a) demonstrate. We will have very little to say about such apparent absence of a dummy subject in this study.

To summarize, despite the absence of genuine *pro*-drop instances in the original English text of the *Harry Potter* book, the Chinese and Japanese translations clearly are

not short of *pro*-drop examples, ranging from matrix subjects, subordinate subjects, to referential and non-referential objects. All these occurrences of Chinese empty pronouns can be accounted for by our syntactic constraints and pragmatic conditions. In the above discussion, we have reviewed five kinds of empty pronouns encountered in the Japanese version, in contrast to the six kinds observed in the Chinese version. This is by no means an exhaustive list, as there can still be other types of empty pronouns that just happen not to be adopted within this one chapter. For example, even though no indefinite empty object is found in the Japanese data, the possibility of its occurrence elsewhere is certain.

Another piece of evidence supporting the view that Japanese zero anaphora is more pragmatically conditioned and less syntactically so than the Chinese counterpart can be seen from the contrast between the sets of sentences below. The original English text uttered by the speaker, Neville, is provided in (342) for reference, whereas the Japanese translation and the Chinese translation are listed in (343) and (344), respectively. The context for Neville's speech is that a bullied Neville is telling other students what has happened to him when they see his legs being stuck together, apparently as a result of the Leg-Locker Curse someone cast on him.

(342) ['Malfoy$_i$,' said Neville shakily …] 'I met him$_i$ outside the library. He$_i$ said he$_i$'d been looking for someone to practice that (a certain curse called Leg-Locker) on.'

(343) Ø tosyokan no soto de Ø deat-ta-no.
NOM library GEN outside in ACC meet-PST-PART
'(I) met (him) outside the library.'
Ø Ø darekani jumon o tamesite mi-ta-kata-tte.
NOM NOM someone curse ACC practice do-DES-PST-CMP
'(He) said (he) wanted to try the curse on someone.'

(344) Wo zai tushuguan waimian pengdao ta.
I at library outside bump.into him
'I bumped into him outside the library.'
Ta shuo ta zheng xiang zhao ge ren lai lianxi suo-twei-zhou.
He say he PROG want find CL person in.order.to practice lock-leg-curse
'He said he was just looking for someone to practice the Leg-Locker Curse on.'

There are four occurrences of empty pronouns in the Japanese sentences in (343), while no empty pronoun is used at all in the Chinese counterpart in (344). The first occurrence among the four in (343) refers to the speaker himself as a first person singular pronoun, which is a common tactic also observed in non-*pro*-drop languages, such as in informal or conversational English. The second occurrence is coreferential with Malfoy, who has just been briefly mentioned in the beginning of this short discourse. (Although every character in this chapter already knows who Malfoy is, the relevance regarding him in this mini discourse is not obvious until Neville's reply). In the second sentence of (343), neither the matrix subject nor the subordinate one has any phonetic realization, which is in stark contrast to the choice of full pronouns in the corresponding Chinese sentence in (344). Regarding Neville's second utterance in (343), for anyone who has been reading the book, it almost goes without saying that the person who said he wanted to practice this offensive curse on someone else would most likely be Malfoy, given his usual evil streak and notorious bullying habits. Probably no reader would think that this person could possibly refer to Neville himself. Therefore, the omission of both subjects in the Japanese sentence is acceptable and unambiguous for any reader who has been following the storyline. However, in the second part of the Chinese sentences in (344), even given the same strong context which indicates that the speaker who said he wanted to practice the offensive curse on someone else is almost certainly Malfoy, and any other reading would seem highly implausible, if the matrix subject were left unexpressed, as in (345) below, the available reading would actually be coreference with the speaker, *wo* 'I/Neville', despite its extreme unlikelihood as the storyline has developed.

(345) Wo$_i$ zai tushuguan waimian pengdao ta$_j$.
I at library outside bump.into him
'I bumped into him outside the library.'
Ø$_{\#i/*j}$ shuo Ø$_{\#i/*j}$ zheng xiang zhao ge ren lai lianxi.
NOM say NOM PROG want find CL person in.order.to practice
suo-twei-zhou.
lock-leg-curse
'(I) said (I) was just looking for someone to practice the Leg-Locker Curse on.'

Therefore, an overt pronoun must be adopted in that position, as in the original Chinese example in (344), in order to avoid such a misleading interpretation. As for the empty subordinate subject in (345), it is less relevant here because it would have to share the same reference with whoever the matrix subject is, following the syntactic constraint which requires an empty subordinate subject in Chinese to be bound (in the extended sense) by an argument. In this case, the matrix subject is the only qualified antecedent as an argument. The reason that the matrix empty subject in (345) favors coindexation with the pragmatically unlikely *wo* 'I/Neville' over the more plausible *ta* 'him/Malfoy' may be that in this mini discourse of (344) or (345), the speech is about the speaker himself, Neville, telling other people what has happened to him, and therefore the topic under discussion has been organized to be no other but himself, while the discourse role played by Malfoy is a non-topical one. Since an empty matrix subject in Chinese must be topical and with the topic being the speaker, Neville himself, this syntactic requirement must be satisfied first and it overrides the pragmatic preference of being coreferential with Malfoy. The constraining of empty pronouns in Japanese appears to be more relaxed and more context-dependent, and, therefore, Japanese zero anaphora also has a wider distribution compared with Chinese zero anaphora, which has been shown by the statistic figure comparison provided in the beginning of this section.

To conclude, in the above discussion we have compared different usages of zero anaphora in Japanese and Chinese by studying the official translations of the two languages from the original English chapter of the first *Harry Potter* Book. There is a much higher percentage of zero anaphora found in the Japanese data than in the Chinese version, even though in both languages, empty subjects have a much wider distribution than empty objects. With the evidence presented in this section and in the previous one, we have shown that Japanese zero anaphora is distinctively constrained from Chinese zero anaphora, with the former being less strictly restricted, resulting in more frequent use of empty pronouns in Japanese than in Chinese. Despite some similarities (discussed in 5.1.1), we argue that it would be oversimplifying, if not downright misleading, to assume that the *pro*-drop phenomena in the two languages are very similar just because the two languages are both topic-prominent and exhibit zero subject-verb agreement.

5.3 Summary

Chapter 5 provides a comparative view on the distribution of Chinese and Japanese zero anaphora. We compared our analysis of Chinese empty pronouns with Kameyama's (1985) study of Japanese zero anaphora and found similarities and differences, followed by an examination of comparing how empty pronouns are distributed in Japanese and Chinese by analyzing the same texts translated into each language from one chapter of the first *Harry Potter* book. Our findings have confirmed that Japanese zero anaphora is more pragmatically conditioned and less syntactically restricted, agreed by Kameyama's study. The fact that Japanese zero anaphora appears to be less strictly constrained than its Chinese counterpart is also supported by the statistic figures we found in 5.2, which showed a much higher frequency and wider distribution of empty pronouns in Japanese than in Chinese from the same texts translated into each target language.

Chapter 6: Conclusions and Further Research

In this thesis we have explored a new perspective of analyzing Chinese zero anaphora from the point of view of Lexical-Functional Grammar. Whilst the study of Chinese *pro*-drop has been extensive in the linguistic literature, none have tackled the issues surrounding Chinese zero anaphora based on the theoretical framework of LFG and also Lambrecht's (1994) information-structure theory. In doing so, we have also demonstrated the compatibility of the two theories and the merits of incorporating the two together. Adopting LFG as the theoretical groundwork has been particularly useful to our account because it allows us to analyze language at different linguistic levels that might seem intertwined and cannot be separated easily in other theories. It allows us to concentrate on different dimensions of linguistic information, e.g., at *f*-structure and at *i*-structure. Such formal grammatical representations are not only clear and neat, but also help us focus only on the relevant information. For instance, when comparing the behavior of Chinese and Japanese empty pronouns, we can divert our attention from certain structural differences (e.g., word order) presented at *c*-structure that make no impact, but simply focus on the related *f*-structure or *i*-structure information.

In this thesis, we first showed the problems and inadequacies of three previous analyses attempting to resolve the issues of Chinese zero anaphora, before reaching the conclusion that a solely syntactic or a solely pragmatic analysis would be insufficient to solve the problems. A selection of syntactic constraints and pragmatic conditions must both be considered. The subsequent discussion of whether the prevalent topic-comment structure, characteristic of Mandarin Chinese, is a syntactic or semantic device also shed light on and introduced a new approach to this debate by resolving the problem with dangling topics in LFG. We have pursued a conjunctive view of combining syntax and pragmatics and established a series of syntactic constraints and pragmatic conditions to which Chinese empty pronouns are subject. The present work has introduced a new way of analyzing Chinese empty pronouns and offered syntactic and pragmatic solutions to account for the *pro*-drop phenomenon in Chinese, which exhibits zero verbal agreement

system. On the completion of the syntactic and pragmatic characterizations of Chinese empty pronouns, we then offered a comparative analysis between the set of constraints for Chinese zero anaphora and those for Japanese zero anaphora based on Kameyama's (1985) study. Further, an original comparative work between the same chapters from the first *Harry Potter* book translated into Chinese and Japanese confirms the conclusion drawn from the similarities and differences between Chinese and Japanese empty pronouns – the former is more constrained, especially syntactically, than the latter, and the latter is more pragmatically conditioned.

Adopting Dalrymple and Nikolaeva's (to appear) *i*-structure representation, we have also preliminarily incorporated Lambrecht's information-structure concepts to our *i*-structure and touched on the issue only to the extent that allows us to accommodate our present scope – resolving Chinese empty pronouns. The kind of *i*-structure we have presented so far is not a thorough, fully-fledged one. There are possibilities to be explored in terms of a finer-grained formation of *i*-structure that can be applied beyond the resolution of Chinese zero anaphora and extended to further linguistic needs. As for the types of empty category included in our study, only empty subjects and empty objects are properly addressed, leaving empty reflexive pronouns (i.e., *ziji* in Chinese and *zibun* in Japanese) unaccounted for. A broader study that includes more types of empty categories such as reflexive and possessive pronouns in Chinese awaits further investigation, especially one based on the LFG framework. A corpus-based study of under what circumstances speakers or writers choose to use empty pronouns the most and least often and whether the frequency of using empty pronouns is affected by different social occasions and types of audience are other potentially interesting avenues of research.

REFERENCES

Aoun, Joseph & Audrey Y.-H. Li (1993). *Syntax of scope*. Cambridge, MA: MIT Press.

Adger, David (2003). *Core syntax: a minimalist approach*. Oxford: Oxford University Press.

Authier, J.-Marc (1989). Arbitrary null objects and unselective binding. In Osvaldo Jaeggli & Kenneth Safir (eds.), *The null subject parameter*. Dordrecht: Kluwer Academic Publishers. 45-67.

Bender, Emily (2000). The syntax of Mandarin BA: reconsidering the verbal analysis. *Journal of East Asian Linguistics* **9**. 105-145.

Bresnan, Joan (1982). Control and complementation. *Linguistic Inquiry* **13**. 343-434.

Bresnan, Joan (2001). *Lexical-functional syntax*. Oxford: Blackwell Publishing.

Bresnan, Joan (2003). Explaining morphosyntactic competition. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*. Oxford: Blackwell Publishing. 11-44.

Bresnan, Joan & Ronald Kaplan (1982). Lexical-Functional Grammar: a formal system for grammatical representations. In Joan Bresnan (ed.), *The mental representation of grammatical relations*. Cambridge, MA: MIT Press. 173-281.

Bresnan, Joan & Sam Mchombo (1987). Topic, pronoun, and agreement in Chichewa. *Language* **63**. 741-782.

Büring, Daniel (2005). *Binding theory*. Cambridge: Cambridge University Press.

Burton-Roberts, Noel (1999). Language, linear precedence and parentheticals. In Peter Collins & David Lee (eds.), *The clause in English*. Amsterdam: John Benjamins. 33-52.

Butt, Miriam & Tracy Holloway King (1996). Structural topic and focus without movement. In Miriam Butt & Tracy Holloway King (eds.), *The proceedings of the LFG'96 conference*. [Available at: http://csli-publications.stanford.edu/LFG/1/lfg1.html]

Butt, Miriam & Tracy Holloway King (2000). Null elements in discourse structure. In K. V. Subbarao (ed.), *Papers from the NULLS seminar*. Delhi: Moti Lal Banarsi Das. [Available at: http://ling.uni-konstanz.de/pages/home/butt/]

Carnie, Andrew (2007). *Syntax: a generative introduction* (2nd edition). Oxford: Blackwell Publishing.

Chafe, Wallace L. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles N. Li (ed.), *Subject and topic*. New York: Academic Press. 25-55.

Chafe, Wallace L. (1987). Cognitive constraints on information flow. In Russell S. Tomlin (ed.), *Coherence and grounding in discourse*. Amsterdam: John Benjamins. 21-52.

Chafe, Wallace L. (1994). *Discourse, consciousness, and time*. Chicago: University of Chicago Press.

Chang, Roland (1977). *Co-verbs in spoken Chinese*. Taipei: Jeng Chung Book Company.

Chao, Wynn (1992). Null objects in Portuguese. Paper presented at the Oxford Romance Linguistics Seminar.

Chao, Yuen-Ren (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.

Chen, Chun-Yin Doris (1995). Topics in Chinese: A or A'? *Studies in English Literature and Linguistics* **21**. 79-106.

Chen, Ping (2003). Indefinite determiner introducing definite referent: a special use of '*yi* 'one' + classifier' in Chinese. *Lingua* **113**. 1169-1184.

Cheng, Lisa L.-S. (1986). *De* in Mandarin. *Canadian Journal of Linguistics* **31**. 313-326.

Chierchia, Gennaro (1989). Structural meanings, thematic roles and control. In Gennaro Chierchia, Barbara Partee & Raymond Turner (eds.), *Properties, types and meaning, volume II: semantic issues*. Dordrecht: Kluwer Academic Publishers. 131-166.

Chiu, Bonnie (1993). *The inflectional structure of Mandarin Chinese*. Ph.D. dissertation. University of Los Angeles.

Choi, Hye-Won (1999). *Optimizing structure in context: scrambling and information structure*. Dissertations in Linguistics. Stanford: CSLI Publications. Revised and corrected version of 1996 Stanford University dissertation.

Chomsky, Noam (1980). On binding. *Linguistic Inquiry* **11**. 1-46.

Chomsky, Noam (1981). *Lectures on government and binding*. Dordrecht: Foris Publications.

Chomsky, Noam (1982). *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.

Chomsky, Noam & Howard Lasnik (1993). Principles and parameters theory. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: an international handbook of contemporary research*. Berlin: de Gruyter. 506-569.

Chu, Chauncey C. (1993). The prototypicality of topic in Mandarin Chinese. *Journal of the Chinese Language Teachers Association* **28**. 25-48.

Chu, Chauncey C. (1998). *A discourse grammar of Mandarin Chinese*. New York: Peter Lang Publishing.

Culy, Christopher D. (1994). Extraposition in English and grammatical theory. Manuscript. University of Iowa.

Dalrymple, Mary, Makoto Kanazawa, Yookyung Kim, Sam Mchombo & Stanley Peters (1998). Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* **21**. 159-220.

Dalrymple, Mary (ed.) (1999). *Semantics and syntax in Lexical Functional Grammar: the resource logic approach*. Cambridge, MA: MIT Press.

Dalrymple, Mary (2001). *Syntax and semantics*: *Lexical Functional Grammar*. San Diego: Academic Press.

Dalrymple, Mary (2006). Lexical Functional Grammar. In Keith Brown (ed.), *Encyclopedia of language and linguistics* (2nd edition). Amsterdam: Elsevier.

Dalrymple, Mary & Irina Nikolaeva (to appear). *Object and information structure*.

Dehé, Nicole & Yordanka Kavalova (2007). Parenthetical: an introduction. In Nicole Dehé & Yordanka Kavalova (eds.), *Parentheticals*. Amsterdam: John Benjamins. 1-22.

Dik, Simon C. (1978). *Functional grammar*. Amsterdam: North-Holland Publishing Company.

Enç, Mürvet (1991). The semantics of specificity. *Linguistic Inquiry* **22**. 1-25.

Erteschik-Shir, Nomi (2007). *Information structure: the syntax-discourse interface*. Oxford: Oxford University Press.

Falk, Yehuda (2001). *Lexical-Functional Grammar: an introduction to parallel constraint-based syntax*. Stanford: CSLI Publications.

Farmer, Ann K. & Robert M. Harnish (1987). Communicative reference with pronouns. In Jef Verschueren & Marcella Bertuccelli-Papi (eds.), *The pragmatic perspective.* Amsterdam: John Benjamins. 547-565.

Farrell, Patrick (1990). Null objects in Brazilian Portuguese. *Natural Language and Linguistic Theory* **8**. 325-346.

Gundel, Jeanette (1976). *The role of topic and comment in linguistic theory*. Ph.D. dissertation. University of Texas at Austin.

Haegeman, Liliane (1994). *Introduction to Government & Binding Theory* (2nd edition). Oxford: Blackwell Publishing.

Her, One-Soon (1991). Topic as a grammatical function in Chinese. *Lingua* **84**. 1-23.

Huang, Chu-Ren (1989). Subcategorized TOPICs in Mandarin Chinese. Manuscript. Academia Sinica. [Available at: http://cwn.ling.sinica.edu.tw/churen/SubcategorizedTOPICsinMandarinChinese.pdf]

Huang, James C.-T. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry* **15**. 531-574.

Huang, James C.-T. (1987). Remarks on empty categories in Chinese. *Linguistic Inquiry* **18**. 321-337.

Huang, James C.-T. (1989). *Pro*-drop in Chinese: a generalized control theory. In Osvaldo Jaeggli & Kenneth Safir (eds.), *The null subject parameter*. Dordrecht: Kluwer Academic Publishers. 185-214.

Huang, James C.-T. (1991). Remarks on the status of the null object. In Robert Freidin (ed.), *Principles and parameters in comparative grammar*. Cambridge, MA: MIT Press. 56-76.

Huang, Shizhe (1996). *Quantification and predication in Mandarin Chinese: a case study of dou*. Ph.D. dissertation. Philadelphia University.

Huang, Yan (1995). On null subjects and null objects in generative grammar. *Linguistic*s **33**. 1081-1123.

Hsueh, Feng-Sheng (1987). The *ba* construction in Mandarin: an integrated interpretation. *Language Teaching and Linguistics Studie*s **1**. 4-22.

Iida, Masayo (1996). *Context and binding in Japanese*. Stanford: CSLI Publications.

Jaeggli, Osvaldo (1982). *Topics in Romance syntax*. Dordrecht: Foris Publications.

Jaeggli, Osvaldo & Kenneth Safir (1989). The null subject parameter and parametric theory. In Osvaldo Jaeggli & Kenneth Safir (eds.), *The null subject parameter*. Dordrecht: Kluwer Academic Publishers. 1-44.

Kameyama, Megumi (1985). *Zero anaphora: the case of Japanese*. Ph.D. dissertation. Stanford University.

Keenan, Edward L. (1976). Towards a universal definition of 'subject'. In Charles N. Li (ed.), *Subject and topic*. New York: Academic Press. 303-333.

King, Tracy Holloway (1995). *Configuring topic and focus in Russian*. Dissertations in Linguistics. Stanford: CSLI Publications. Revised and corrected version of 1993 Stanford University dissertation.

King, Tracy Holloway (1997). Focus domains and information structure. In Miriam Butt & Tracy Holloway King (eds.), *The proceedings of the LFG'97 conference*. Stanford: CSLI Publications. [Available at: http://csli-publications.stanford.edu/LFG/2/lfg97.html]

King, Tracy Holloway & Annie Zaenen (2004). F-structures, information structure, and discourse structure. In Miriam Butt & Tracy Holloway King (eds.), *The proceedings of the LFG'04 conference*. Stanford: CSLI Publications. [Available at: http://csli-publications.stanford.edu/LFG/9/lfg04-toc.html]

Kuno, Susumo (1972). Functional sentence perspective: a case study from Japanese and English. *Linguistic Inquiry* **3**. 269-320.

Kuno, Susumu & Etsuko Kaburagi (1977). Empathy and syntax. *Linguistic Inquiry* **8**. 627-672.

Kuroda, Shige-Yuki (1965). *Generative grammatical studies in the Japanese language*. Ph.D. dissertation. MIT.

Lambrecht, Knud (1994). *Information structure and sentence form: topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Lasnik, Howard (1976). Remarks on coreference. *Linguistic Analysis* **2**. 1-22.

Lasnik, Howard (1981). On two recent treatments of disjoint reference. *Journal of Linguistic Research* **1**. 48-58.

Lebeaux, David (1984). Anaphoric binding and the definition of PRO. *The proceedings of the North East Linguistic Society (NELS)* **14**. 253-274.

Lee, Thomas H.-T. (1986). *Studies on quantificational in Chinese*. Ph.D. dissertation. University of California.

Levinson, Stephen C. (1987). Pragmatics and the grammar of anaphora: a partial pragmatic reduction of binding and control phenomena. *Journal of Linguistics* **23**.

379-434.

Levinson, Stephen C. (1991). Pragmatic reduction of the binding conditions revisited. *Journal of Linguistics* **27**. 107-161.

Li, Audrey Y.-H. (1985). *Abstract case in Chinese*. Ph.D. dissertation. University of Southern California.

Li, Audrey Y.-H. (1992). Indefinite *wh* in Mandarin Chinese. *Journal of East Asian Linguistics* **1**. 125-156.

Li, Audrey Y.-H. (2001). The *ba* construction. Manuscript. University of Southern California.

Li, Charles N. & Sandra A. Thompson (1976). Subject and topic: a new typology of language. In Charles N. Li (ed.), *Subject and topic*. New York: Academic Press. 459-489.

Li, Charles N. & Sandra A. Thompson (1989). *Mandarin Chinese: a functional reference grammar*. Berkeley: University of California Press.

Li, Cherry I. (1985). *Participant anaphora in Mandarin Chinese*. Ph.D. dissertation. University of Florida.

Li, Jie (1995). *Dou* and *wh*-questions in Mandarin Chinese. *Journal of East Asian Linguistics* **4**. 313-323.

Matsuoka, Yuko (2003). *Hari- Potta- to kenja no ishi*. (Official Japanese Translation of *Harry Potter and the philosopher's stone* by J. K. Rowling). Tokyo: Say-zan-sha Publication.

Montalbetti, Mario (1984). *After binding*. Ph.D. dissertation. MIT.

Mei, Kuang (1978). The *ba* construction in modern Chinese. *Bulletin of the College of Liberal Arts*, *National Taiwan University* **27**. 145-180.

Mei, Kuang (1980). Is modern Chinese really a SOV language? *Cashiers de Linguistique Asie Orientale* **7**. 23-45.

O'Connor, Robert (2006). *Information structure in Lexical-Functional Grammar: the discourse-prosody correspondence*. Ph.D. dissertation. University of Manchester.

Ohso, Mieko (1976). *A study of zero pronominalization in Japanese*. Ph.D. dissertation. Ohio State University.

Pan, Haihua & Jianhua Hu (2000). Deriving the subject-object asymmetry in topicalization. Paper presented at the International Symposium on Topic and Focus in Chinese, Hong Kong Polytechnic University.

Pan, Haihua & Jianhua Hu (2002). Representing topic-comment structure in Chinese. In *Proceedings of the 16th Pacific-Asia conference on language, information and computation*. 382-390.

Paul, Waltraud (2005). Adjectival modification in Mandarin Chinese and related issues. *Linguistics* **43**. 757-793.

Peng, Qian-Wen (2000). *Hali Pote – shen mi de mo fa shi*. (Official Chinese Translation of *Harry Potter and the philosopher's stone* by J. K. Rowling). Taipei: Crown Publishing Company Ltd.

Platzack, Christer (1994). Null subjects, weak Agr and syntactic differences in Scandinavian. *Working Papers in Scandinavian Syntax* **53**. 85-106.

Poesio, Massimo, Hua Cheng, Renate Henschel, Janet Hitzeman, Rodger Kibble & Rosemary Stevenson (2000). Specifying the parameters of Centering Theory: a corpus-based evaluation using text from application-oriented domains. In *Proceedings of the 38th annual meeting of Association for Computational Linguistics*. Morristown, NJ: Association for Computational Linguistics.

Reinhart, Tanya (1982). Pragmatics and linguistics: an analysis of sentence topics. Distributed by Indiana University Linguistics Club, Bloomington, Indiana.

Rizzi, Luigi (1986). Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* **17**. 501-557.

Ross, Claudia (1983). On the function of Mandarin DE. *Journal of Chinese Linguistics* **11**. 214-246.

Ross, John Robert (1967). *Constraints on variables in syntax*. Ph.D. dissertation. MIT.

Rowling, Joanne Kathleen (1997). *Harry Potter and the philosopher's stone*. London: Bloomsbury Publishing.

Saito, Mamoru & Hajime Hoji (1983). Weak crossover and move alpha in Japanese. *Natural Language and Linguistic Theory* **1**. 245-259.

Sharvit, Yael (1999). Resumptive pronouns in relative clauses. *Natural Language and Linguistic Theory* **17**. 587-612.

Shi, Dingxu (2000). Topic and topic-comment constructions in Mandarin Chinese. *Language* **76**. 383-408.

Szeto, Yee-Kit (1988). *A semantic description of aspectual and temporal reference in Chinese*. Ph.D. dissertation. University of Ottawa.

Taraldsen, Knut (1978). On the NIC, vacuous application, and the *that*-trace filter. Manuscript. MIT.

Teng, Shousin (1974). Double nominatives in Chinese. *Language* **50**. 455-473.

Tsao, Feng-Fu (1977). *A functional study of topics in Chinese: the first step towards discourse analysis*. Ph.D. dissertation. University of Southern California.

Tsao, Feng-Fu (1979). *A functional study of topic in Chinese: the first step towards discourse analysis*. Taipei: Student Book Co.

Tsao, Feng-Fu (1988). Topics and clause connectives in Chinese. *Bulletin of the Institute of History and Philology Academia Sinica* **59**. 696-737.

Tsao, Feng-Fu (1990). *Sentence and clause structure in Chinese: a functional perspective*. Taipei: Student Book Co.

Vallduví, Enric (1992). *The informational component*. New York: Garland.

Walker, Marilyn, Masayo Iida & Sharon Cote (1994). Japanese discourse and the process of centering. *Computational Linguistics* **20**. 193-232.

Wong, May L.-Y. (2005). *Adverbial clauses in Mandarin Chinese: a corpus-based study*. Ph.D. dissertation. Lancaster University.

Wu, Zoe (2000). *Grammaticalization and the development of functional categories in Chinese*. Ph.D. dissertation. University of Southern California.

Xu, Liejiong & D. Terence Langendoen (1985). Topic structures in Chinese. *Language* **61**. 1-27.

Xu, Liejiong & Liu Danqing (1998). *Topic: structural and functional analyses*. Shanghai: Shanghai Education Press.

Yong, Shin (1993). *The aspectual phenomena of the ba construction*. Ph.D. dissertation. University of Wisconsin-Madison.