

把

To *ba* or not to *ba*

Differential Object Marking in Chinese

Geertje van Bergen

To *ba* or not to *ba*

Differential Object Marking in Chinese

Geertje van Bergen

PIONIER Project *Case Cross-linguistically*

Department of Linguistics

Radboud University Nijmegen

P.O. Box 9103

6500 HD Nijmegen

the Netherlands

www.ru.nl/pionier

g.vanbergen@let.ru.nl

To *ba* or not to *ba*

Differential Object Marking in Chinese

Master's Thesis

General Linguistics

Faculty of Arts

Radboud University Nijmegen

November 2006

Geertje van Bergen

0136433

First supervisor: Dr. Helen de Hoop

Second supervisor: Prof. Dr. Pieter Muysken

Acknowledgements

I would like to thank Lotte Hogeweg and the members of the PIONIER-project Case Cross-linguistically for the nice cooperation during the past year; many thanks go to Monique Lamers for the pleasant teamwork. I am especially grateful to Yangning for the close cooperation in Beijing and Nijmegen, which has laid the foundation of this thesis. 谢谢! Many thanks also go to Sander Lestrade for clarifying conversations over countless cups of coffee. I would like to thank Pieter Muysken for his willingness to be my second supervisor and for his useful comments on an earlier version. Also, I gratefully acknowledge the Netherlands Organisation of Scientific Research (NWO) for financial support, grant 220-70-003, principal investigator Helen de Hoop (PIONIER-project 'Case cross-linguistically').

Especially, I would like to thank Helen de Hoop for her great support and supervision, for keeping me from losing courage and keeping me on schedule, for her indispensable comments and for all the opportunities she offered me to develop my research skills.

Table of Contents

Acknowledgements	V
Table of Contents	VII
List of Abbreviations	XI
1. Introduction	1
2. Differential Object Marking in Chinese	5
2.1 Object scrambling and object marking	5
2.2 Differential Object Marking	11
2.2.1 DOM cross-linguistically	11
2.2.2 DOM in Chinese	14
2.3 A functional analysis of DOM	22
2.4 Chinese DOM vs. cross-linguistic DOM	28
2.5 Conclusion	31

3. Word Order	33
3.1 The syntactic role of word order	33
3.2 The semantics of word order	37
3.3 Word order and Chinese DOM	46
3.4 Conclusion	50
4. An OT account of Chinese DOM	51
4.1 An introduction to OT	51
4.2 Aissen (2003): an OT account of cross-linguistic DOM	54
4.3 An OT account of Chinese DOM	63
4.3.1 Animacy and Chinese DOM	63
4.3.2 Definiteness and Chinese DOM	67
4.3.3 Chinese DOM: a full account	73
4.4 Conclusion	77
5. The function of <i>ba</i>	79
5.1 Two-dimensional DOM: animacy, definiteness and the function of the object marker	80
5.2 Chinese DOM: word order and <i>ba</i>	85
5.3 Chinese DOM: animacy and <i>ba</i>	90
5.4 Chinese DOM: definiteness and <i>ba</i>	91
5.5 Object scrambling and the markedness principle	95
5.6 Conclusion	101
6. Conclusion	103
References	107

List of Abbreviations

ACC	accusative case	NSPEC	non-specific
ANIM	animate	O	object
CL	classifier	OBJ	object
COP	copula	PAST	past tense
DEF	definite	PRE	preverbal
ERG	ergative case	POST	postverbal
HUM	human	PL	plural
IMP	imperfective aspect	PRT	particle
INAN	inanimate	PN	pronoun
INDEF	indefinite	S	subject
LOC	locative	SUBJ	subject
NOM	nominative case	SPEC	specific
NOUN	proper noun	V	verb
NP	noun phrase		

1

Introduction¹

In a number of languages in the world that exhibit overt case marking on their direct objects, some objects are marked, but others are not. Consider the following sentences from Pitjantjatjara, an Australian language (cf. Bowe 1990):

(1) a. *Minyma-ngku ngayu-nya pu-ngu.*
 woman. ERG I.ACC hit-PAST
 'The woman hit me.'

b. *Ngayula minyma pu-ngu.*
 I. NOM woman hit-PAST
 'I hit the woman.'

¹ This thesis is based on joint work with Yang Ning on object scrambling and object marking in Chinese (to appear).

We see that the object *ngayu* 'I' in (1a) is followed by the accusative marker *-nya*, whereas the object *minyma* 'woman' in (1b) is not. This is an instance of what is called *Differential Object Marking* (DOM), and this phenomenon takes different forms across languages. Cross-linguistically, DOM varies with respect to exactly which objects can be case marked, and to whether object marking is obligatory or optional.

In this thesis, I will investigate how DOM is realized in Chinese. The Chinese language consists of a group of related languages or dialects that belong to the Sino-Tibetan language family. With almost one billion speakers, it is the most commonly spoken language in the world. I will focus on the main and official dialect of the Chinese language, which is known by a number of names, among which Mandarin or 普通话 *putonghua* 'common speech'. I will henceforth refer to this dialect as Chinese.

In linguistics, Chinese is one of the best known examples of an analytic or isolating language. In analytic languages, there is little to no morphological change or inflection in words. For example, plurality in Chinese is indicated by a word like 一些 *yixie* 'some' or 多 *duo* 'many' instead of by plural inflection like the English affix *-s*. The ratio of words to morphemes in analytic languages is nearly one-to-one. Each individual morpheme has a general meaning and corresponds to a single character, and nuances are expressed by other morphemes. For instance, the Chinese word 电 *dian* means 'electricity'. When it is used together with 脑 *nao* 'brain' it means 'computer', whereas in combination with 话 *hua* 'speech' it denotes 'telephone'. For the Chinese examples I discuss in this thesis, I will make use of 拼音 *pinyin* instead of the Chinese character script. *Pinyin*, in which *pin* means

'spell' and *yin* means 'sound', is a system of phonemic notation and transcription of the Chinese character script into Roman script.

Despite the fact that Chinese is an analytic language, it does exhibit a system of Differential Object Marking: some direct objects are obligatorily preceded by the morpheme *ba*, for some objects *ba* is optional, and for a third set of objects *ba* is prohibited. In this thesis, I will investigate how exactly DOM is realized in Chinese. In Chapter 2, I will examine how animacy and definiteness of the direct object play a role in the realization of Chinese DOM, and I will show that the influence of these features on Chinese DOM deviates from their influence on DOM systems cross-linguistically. I will show that, contrary to cross-linguistic DOM systems that are determined by these semantic properties only, DOM in Chinese is first of all syntactically driven. I will argue that the dimension of word order must be taken into account when determining Chinese DOM. In Chapter 3, I will further investigate the role of word order in languages in general, and its influence on Chinese DOM in particular. I will discuss Aissen's (2003) formal OT account of cross-linguistic instances of DOM in Chapter 4, and I will show how her OT Syntactic model should be adapted to account for the Chinese DOM system as well. In Chapter 5, I will examine the function of *ba* in Chinese in comparison with the function of object markers in languages with two-dimensional DOM.

4 Geertje van Bergen

2

Differential Object Marking in Chinese

In this chapter, I will study the phenomenon of Differential Object Marking in general and the Chinese DOM system in particular. I will show in Section 2.1 that object marking in Chinese is sometimes obligatory, sometimes optional and sometimes prohibited. In Section 2.2, I will illustrate how semantic features of the direct object are involved in cross-linguistic realizations of DOM and I will show how they play a role in Chinese DOM, too. I will discuss a functional explanation of cross-linguistic instances of DOM in Section 2.3, and I will examine whether the Chinese DOM system can be explained accordingly in Section 2.4. I will give a brief conclusion in Section 2.5.

2.1 Object scrambling and object marking

Since words in Chinese are generally not marked by any morphology showing their role in the sentence, word order carries a lot of importance. The elementary SVO word order is the main indicator for

the grammatical functions of the arguments in the sentence. Consider the example in (2)²:

(2) *Ta chi le na-ge pingguo.*
 he eat PRT that-CL apple
 'He ate that apple.'

The subject *ta* 'he' is located on the left side of the verb *chi le* 'ate' and the object *na-ge pingguo* 'that apple' is placed on the right side. Although word order in Chinese is rather strict, direct objects can or sometimes even must move to a preverbal position. For instance, the language restricts the number of elements that can occur after the verb (the postverbal constraint): generally, only one constituent is allowed postverbally (e.g., Travis 1984, Sybesma 1992, Po-Ching and Rimmington 2004). In the example in (3) below, we see that if the verb *fang* 'put' is followed by two constituents, the sentence is ungrammatical:

(3) **Wo fang qiu jin lanzi li le.*
 I put ball into basket inside PRT
 'I put the ball/the balls into the basket.'

In cases like this, the direct object must move to the left side of the verb, where it is preceded by the morpheme *ba*, as shown in (4):

² Many thanks to Yang Ning for all the Chinese data I use in this thesis. Most of the examples are taken from Yang and van Bergen (to appear); some are taken from Yang (in prep.) and others are obtained through personal communication.

(4) *Wo ba qiu fang jin lanzi li le.*
 I BA ball put into basket inside PRT
 'I put the ball(s) into the basket.'

In other cases, direct objects can be placed both in postverbal and in preverbal position, objects in preverbal position also being preceded by *ba*. For instance, the sentence in (2) can also be expressed as follows:

(5) *Ta ba na- ge pingguo chi le.*
 he BA that-CL apple eat PRT
 'He ate that apple.'

Both the sentences in (2) and (5) express the same meaning 'he ate that apple.' Compare also the sentences below, in which the object *wo* 'I' is placed postverbally in (6) and preverbally in (7):

(6) *Ta da le wo.*
 he hit PRT I
 'He hit me.'

(7) *Ta ba wo da le.*
 he BA I hit PRT
 'He hit me.'

These examples show that direct objects that are, either obligatorily or optionally, scrambled to a preverbal position can be preceded by *ba*. In

Classical Chinese, *ba* was a verb, meaning ‘to hold’ or ‘to take’. However, in the sentences above, *ba* does not have any meaning. The exact function of *ba* in modern Chinese is a widely discussed topic among linguists. It is treated either as a verb (Hashimoto 1971), a preposition (Travis 1984, Li 2001) or as a case marker (Huang 1982, Huang 1990, Goodall 1987, Yang in prep.). For now, I will consider *ba* to be a case marker for direct objects, but I will come back to the discussion on the function of *ba* in the last chapter of this thesis.

Case marking of direct objects in Chinese is limited to scrambled objects. Direct objects in postverbal position are never case marked:

(8) *Ta chi le (*ba) na-ge pingguo.*
 he eat PRT BA that-CL apple
 ‘He ate that apple.’

In most cases, scrambled objects are obligatorily marked with *ba*. Sometimes, however, the case marker can be omitted. Compare the example in (5) above with the sentence in (9):

(9) *Ta na-ge pingguo chi le.*
 he that-CL apple eat PRT
 ‘He ate that apple.’

The sentences in (5) and (9) are both grammatical ways to express the intended meaning ‘he ate that apple’. This, however, does not hold for every direct object in preverbal position. For instance, when *na-ge pingguo* ‘that apple’ is changed into *yi-ge pingguo* ‘an apple’, the use of

ba becomes obligatory, as shown in (10). The sentence in (11) illustrates that omitting the case marker is also impossible for the pronoun object *wo* 'I' in preverbal position:

(10) *Ta* *(*ba*) *yi-ge* *pingguo* *chi* *le.*

he BA one-CL apple eat PRT

'He ate an apple.'

(11) *Ta* *(*ba*) *wo* *da* *le.*

he BA I hit PRT

'He hit me.'

I should note here that a pronoun can occur preverbally without being preceded by *ba*. However, it will then be interpreted as the subject of the sentence. Consider the following example:

(12) *Laohu* *wo* *chi* *le.*

tiger I eat PRT

'The tiger, I ate it.'

The object *laohu* 'tiger' precedes the subject *wo* 'I' in the sentence in (12). This is due to the fact that Chinese sentences have a topic-comment structure, in which the topic is the element being talked about or predicated and the comment is what is said about the topic. Chinese is a topic-prominent language, meaning that the topic is always mentioned first in the sentence. This topic-prominent structure is independent of the syntactic ordering of subject, verb and object. If we

compare the sentence in (12) with the sentence in (13), we see that topic, subject and object (the anaphor *ta* 'it', referring to the antecedent *laohu* 'lion') can all precede the verb:

(13) *Laohu_i* *wo* *ba* *ta_i* *chi* *le*.

tiger I BA it eat PRT

'The tiger, I ate it.'

We can infer from this sentence that Chinese has three preverbal positions at least. In this thesis, I will not elaborate on the function of topic, the topic position and the topic-comment structure in Chinese. When I talk about 'scrambled objects' or 'objects in preverbal position', I refer to objects that scramble to the left side of the verb, but to the right side of the subject. I want to make clear that the constituents' order in the sentence changes when the object is scrambled, but it is not my aim to determine where exactly this scrambled object position should be placed in the syntactic structure.

The examples in this section illustrate that case marking of direct objects in Chinese is sometimes obligatory, sometimes optional and sometimes prohibited. This is an instance of what has been called Differential Object Marking or DOM (Aissen 2003). In the next section, I will show how DOM manifests itself cross-linguistically and in Chinese.

2.2 Differential Object Marking

2.2.1 DOM cross-linguistically

Cross-linguistically, Differential Object Marking is not a rarity. For many languages which exhibit overt case marking on direct objects, it is common to mark some objects but not others. In Chinese DOM, there are three sets of direct objects: one set of objects for which case marking is obligatory, a second set for which case marking is optional, and a third set for which case marking is prohibited. Similarly, objects in Spanish can be divided into three categories with respect to object marking. Consider the following examples (Rodríguez-Mondoñedo 2006):

(14) *Juan mató *(a) María.*

Juan killed A María

'Juan killed Maria.'

(15) *María quiere (a) un abogado.*

Maria wants A a lawyer

'Maria wants a lawyer.'

(16) *Juan destruyó (*a) la ciudad.*

John destroyed A the city

'John destroyed the city.'

Some objects in Spanish, like *María* in (14), are obligatorily marked with the preposition *a*, whereas for some objects, such as *la ciudad* ‘the city’ in (16), this preposition is prohibited. A third set of objects can optionally be marked with *a*, like *un abogado* ‘a lawyer’ in (15).

Although DOM is very common, the way in which it is realized differs across languages. Cross-linguistically, DOM varies with respect to exactly which objects can be case marked, and to whether object marking is obligatory or optional. Cross-linguistic variation in DOM is considered to be determined by the semantic properties of the object. Case marking can for instance be triggered by animacy features. Consider the following sentences from Malayalam (cf. Asher and Kumari 1997):

(17) a. *Avan kuttaye aṭiccu.*
 he child. ACC beat-PAST
 ‘He beat the child.’

b. *Avan oru paſuvine vajji.*
 he a cow. ACC buy- PAST
 ‘He bought a cow.’

c. *jaan teejna vajji.*
 I coconut.NOM buy- PAST
 ‘I bought some coconut.’

These examples illustrate that human and animate objects in Malayalam are case marked, whereas inanimate objects are not.

In other languages, the definiteness or specificity of the direct object determines whether or not it is marked with case. In Hebrew, for instance, indefinite objects are not case marked, while definite objects are, as the examples in (18) (Aissen 2003):

(18) a. *Ha-seret her ɻa ɻet-ha-milxama.*
 the-movie showed ACC-the-war
 'The movie showed the war.'

b. *Ha-seret her ɻa (*ɻet-) milxama.*
 the-movie showed (ACC -)war
 'The movie showed a war.'

Languages can also use more than one semantic feature to establish which objects will be case marked and which objects will lack case marking. One of those languages is Hindi, in which the DOM system is based on both animacy and definiteness. Consider the following examples (cf. Mohanan 1990):

(19) a. *Ilaa-ne bacce- ko / *baccaa uthaayaa.*
 Ila- ERG child-ACC/child.NOM lift-PAST
 'Ila lifted the/a child.'

b. *Ilaa-ne haar uthaayaa.*
 Ila- ERG necklace. NOM lift-PAST
 'Ila lifted the/a necklace.'

c.	<i>Ilaa-ne</i>	<i>haar-ko</i>	<i>uthaayaa.</i>
	Ila- ERG	necklace-ACC	lift-PAST
'Ila lifted the/*a necklace.'			

The sentence in (19a) illustrates that human objects are obligatorily marked with case in Hindi. For inanimate objects, the case marker is optional, as shown in (19b)-(19c). There is, however, a difference in interpretation between *haar* 'necklace' in (19b) and in (19c): when the inanimate direct object is not marked with accusative case, it can get both a definite and an indefinite reading. However, when the object is case marked, it can only be interpreted as definite. This shows that in Hindi, three sets of direct objects can be distinguished. Case marking is obligatory for all animate objects; it is optional for inanimate definite objects and for inanimate indefinite objects, case marking is prohibited.

We have seen here that languages with DOM differ in exactly which objects are marked and in whether case marking is obligatory or optional, and that these cross-linguistic differences are based on semantic features of the object, the features animacy and definiteness in particular. In the following section, I will examine how these semantic properties are involved in Chinese DOM.

2.2.2 DOM in Chinese

In this section, I will investigate whether and how animacy and definiteness influence the Chinese DOM system. Nouns can be ranked by animacy or definiteness on a continuum ranging from most to least animate or definite, and various levels of animacy and definiteness can

be distinguished in a language. Both animacy and definiteness are semantic properties that contribute to the prominence of an NP, that is, the high ranking of a noun on a salience scale. In (20) and (21), the most commonly distinguished categories of both features are represented in the form of universal prominence scales, on which $a > b$ means that 'a is more prominent than b':

(20) *Animacy scale:* Human > Animate > Inanimate

(21) *Definiteness scale:* Pronoun > Proper Noun > Definite NP > Indefinite Specific NP > Indefinite Non-specific NP

In general, human individuals are more sentient and therefore ontologically more salient in the discourse than animate and inanimate entities, respectively. As for the definiteness scale, pronouns, proper nouns and definite nouns are generally more specific or better identifiable in a context than indefinite noun phrases, making definite NPs more salient than indefinite NPs.

Let us now consider whether the above prominence scales are of influence on the realization of DOM in Chinese. Starting with animacy, we can see from the following sentences that this dimension plays a role in Chinese DOM:

(22) *Ta * (ba) laoshi tuidao le.*
 he BA teacher push.over PRT
 'He pushed over the teacher.'

(23) *Ta *(ba) she dasi le.*
 he BA snake hit.dead PRT

'He killed the snake.'

(24) *Ta (ba) pingguo chi le.*
 he BA apple eat PRT

'He ate the apple/the apples.'

The dimension of animacy determines whether the case marker in Chinese is obligatory or optional: the human and animate scrambled objects in (22) and (23) are obligatorily marked with *ba*, whereas the case marker for the inanimate scrambled object in (24) can be omitted.

There is an exception to the obligatory marking of animate objects. Consider the following example:

(25) *Ta (ba) ji chi le.*
 he BA chicken eat PRT

'He ate (the) chicken.'

For the scrambled animate object *ji* 'chicken' in this sentence, case marking is optional. This is due to the fact that the object in this sentence is considered a meal and therefore inanimate.

The examples above illustrate that for Chinese DOM, the cut-off point on the animacy scale lies between animate and inanimate objects: human and animate objects are obligatorily marked with *ba* when they are scrambled, whereas for inanimate scrambled objects, the case

marker can be omitted. However, case marking is not optional for every inanimate scrambled object. Consider the following examples:

(26) *Ta* *(*ba*) *yi-ge* *pingguo* *chi* *le.*
 he BA one-CL apple eat PRT
 'He ate an apple.'

(27) *Ta* *(*ba*) *yi-bei* *shui* *he* *le.*
 he BA one-CL water drink PRT
 'He drank a glass of water.'

Even though the objects in (26)-(27) are inanimate, *ba* cannot be omitted in these sentences. But if we now change the predicate *chi-le* 'ate' in (26) into an accomplishment predicate *chi-wan le* 'ate up', or add *dou* 'all' to the object, case marking is optional again:

(28) *Ta* (*ba*) *yi-ge* *pingguo* *chi-wan* *le.*
 he BA one-CL apple eat-finish PRT
 'He ate up an apple.'

(29) *Ta* (*ba*) *yi-ge* *pingguo* *dou* *chi* *le.*
 he BA one-CL apple all eat PRT
 'He ate a whole apple.'

The same holds for the sentence in (27). When *he le* 'drank' is changed into *he-wan le* 'drank up, finished', or when the object *yi-bei shui* 'a cup of water' is followed by *dou*, *ba* can be omitted:

(30) *Ta (ba) yi-bei shui he-wan le.*
 he BA one-CL water drink-finish PRT

‘He drank up a cup of water.’

(31) *Ta (ba) yi-bei shui dou he le.*
 he BA one-CL water all drink PRT

‘He drank a whole cup of water.’

These differences between obligatory and optional case marking of inanimate objects can be explained when we take into account the dimension of definiteness. Chinese does not have (in)definite articles, but the object NPs in the sentences above are all lexically marked with *yi* ‘one’, by which the objects get an indefinite reading. An accomplishment predicate like *chi-wan le* makes that its indefinite object *yi-ge pingguo* ‘an apple’ is interpreted as a specific apple (e.g. Sybesma 1992). A specific reading can also be evoked by adding the universal quantifier *dou* ‘all’ to the indefinite object (e.g., Lee 1986, Liu 1997). Cheng (2006) even argues that *dou* in fact has the same function as a definite determiner, by which the object would be lexically marked for definiteness and consequently get a definite reading.

If an object NP marked with *yi* is neither part of an accomplishment predicate nor followed by *dou*, it gets a non-specific reading (Yang in prep.). In that case, the case marker is obligatory. For lexically indefinite objects which get a specific interpretation, on the other hand, the case marker can be omitted in Chinese, as shown in (28)

– (31). Case marking of definite objects is also optional, as the following sentences illustrate:

(32) *Ta (ba) na-ge pingguo chi le.*
 he BA that-CL apple eat PRT
 'He ate that apple.'

(33) *Wo (ba) na-ge qiu fang jin lanzi li*
 I BA that-CL ball put into basket inside PRT
 'I put that ball into the basket.'

Just like the dimension of animacy, the dimension of definiteness is of influence on the obligatory case marking of certain direct objects in Chinese, and the optional case marking of others. The cut-off point on the definiteness scale should be placed between non-specific and specific indefinite NPs: non-specific indefinite scrambled objects are obligatorily marked with *ba*, whereas the case-marker is optional for specific and definite objects in preverbal position. This would imply that for all objects that are ranked higher on the definiteness scale, case marking would also be optional. However, for sentences in which the scrambled object is a proper noun or a pronoun, *ba* is obligatory:

(34) *Ta *(ba) Zhangsan da le.*
 he BA Zhangsan hit PRT
 'He hit Zhangsan.'

(35) *Ta *(ba) wo da le.*
 he BA I hit PRT
 'He hit me.'

This can be explained by the fact that the proper name *Zhangsan* in (34) refers to a human being, just like the pronoun *wo* in (35). We saw above that case marking for human scrambled objects is obligatory. Even though pronouns and proper nouns have high definiteness, which would permit optional case marking, they have high animacy at the same time, which makes case marking of the direct object obligatory. The same holds for other human and animate objects: if in a Chinese sentence the scrambled object is human or animate, it is obligatorily case marked, regardless of the definiteness of the object. Consider the following examples:

(36) *Ta *(ba) zhe-tiao she dasi le.*
 he BA this-CL snake hit.dead PRT
 'He killed this snake.'

(37) *Ta *(ba) na-ge laoshi piping le.*
 he BA that-CL teacher criticise PRT
 'He criticised that teacher.'

Even though *zhe-tiao she* 'this snake' and *na-ge laoshi* 'that teacher' are definite NPs, case marking is obligatory because the direct objects are animate and human, respectively. The influence of definiteness becomes visible only when the object is inanimate. If an inanimate

scrambled object is non-specific, the case marker is obligatory; if an inanimate scrambled object is specific or definite, case marking is optional. The illustrating sentences are repeated below for convenience's sake:

(38) *Ta (ba) na-ge pingguo chi le.*
 he BA that-CL apple eat PRT
 'He ate that apple.'

(39) *Ta (ba) yi-ge pingguo dou chi le.*
 he BA one-CL apple all eat PRT
 'He ate a whole apple.'

(40) *Ta *(ba) yi-ge pingguo chi le.*
 he BA one-CL apple eat PRT
 'He ate an apple.'

From the sentences in this section, we can conclude that the dimensions of animacy and definiteness together determine whether object marking in Chinese is obligatory or optional. The only scrambled objects for which case marking is optional are both inanimate and specific or definite. The case marker is obligatory for animate and human scrambled objects, as well as for non-specific indefinite scrambled objects.

In the next section, I will discuss the functional explanation for cross-linguistic DOM systems as given by Aissen (2003).

2.3 A functional analysis of DOM

We saw in the previous sections that the way in which DOM manifests itself may differ per language, but that DOM is cross-linguistically determined by the same semantic properties, that is, animacy and definiteness, of the direct object. Both these features contribute to the prominence of an NP. The universal prominence scales of animacy and definiteness are repeated in (41) and (42):

(41) *Definiteness scale:* Pronoun > Proper Noun > Definite NP >
 Indefinite Specific NP > Indefinite Non-specific NP

(42) *Animacy scale:* Human >Animate > Inanimate

From the results of a great amount of theoretical research in functional and typological syntax, as well as from descriptive works of individual languages, Aissen deduced the following general thought behind all manifestations of DOM: ‘the higher in prominence a direct object, the more likely it is to be overtly case marked’ (Aissen 2003: 435). This generalization expresses that in languages with DOM, if a direct object at some point on the prominence scales in (41) and (42) can be case marked, then objects that are more prominent or higher ranked can also receive a case marker, but not necessarily less prominent objects. At the same time, if a direct object must be case marked in a language with DOM, then necessarily all objects that are more prominent must be case marked as well, while lower ranked objects do not have to

receive case marking. Cross-linguistically, languages vary in which dimensions are of influence on their DOM pattern. In some languages, like in Malayalam, only the scale of animacy plays a role, whereas other languages, such as Hebrew, only use the dimension of definiteness for DOM. In languages such as Romanian and Hindi, both animacy and specificity influence the DOM pattern. Furthermore, languages differ cross-linguistically in where exactly on the scale(s) they make the split between case marking and no marking. In some language, only human objects may be case marked, whereas non-human (animate and inanimate) objects do not receive case. In another language, it may be all animate objects that are case marked, while only inanimate objects are not. Wherever languages may exactly place the cut-off point, it is universally the higher prominent objects in terms of definiteness and/or animacy that will receive case marking, and the lower prominent ones which will lack a case marker.

In the literature on DOM, we find the general assumption that case marking is used to distinguish subjects from objects. If objects have high prominence, they are more difficult to distinguish from subjects: for reasons of disambiguation, a case marker is used. This indeed appears to be the reason for differential object marking in a number of languages. However, we also find languages that exhibit DOM, even though the case marker is not necessary for disambiguation. Aissen chooses a weaker formulation of the intuition on DOM. She explains why high prominent objects are more susceptible to case marking in terms of markedness:

(43) 'The high prominence which motivates DOM for objects is exactly the prominence which is unmarked for subjects.'

Markedness is a relative notion: which elements are marked and which are unmarked can only be determined in comparison with other elements. When we want to explain the general thought behind DOM in terms of markedness, the dimensions of animacy and definiteness should be understood in connection with the prominence scale of grammatical function or the relational scale, which expresses that subjects are more prominent than objects (cf. Aissen 2003):

(44) *Relational scale:* Subject > Object

Subjects and objects are generally associated with a number of prototypical properties (e.g., Silverstein 1976, Comrie 1979, Comrie 1989). Comrie lists these properties in the following way:

"[...] in natural languages, certain grammatical relations tend to be characterized by certain features, in particular [that] subjects tend to be definite, animate, and topic (thematic); while direct objects tend to be indefinite, inanimate, and rhematic" (Comrie 1979: 19).

Because of the association of subjects with high animacy and definiteness, NPs that are on the upper end of the prominence scales in (41) and (42) are more natural, or unmarked subjects (Keenan 1976),

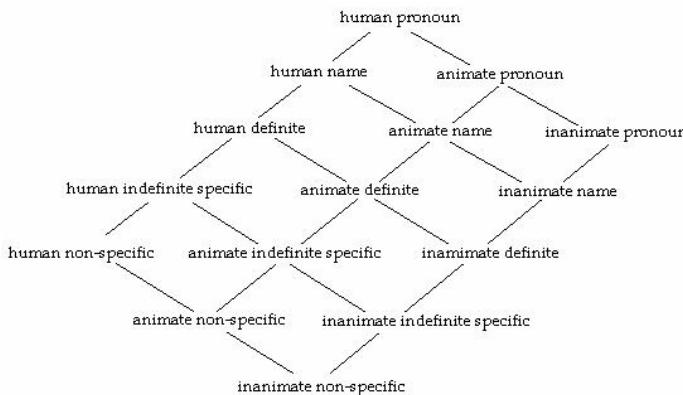
whereas NPs that are on the lower end are more natural or unmarked objects. It follows from these prototypical features that what is marked for objects is unmarked for subjects and vice versa. This is an example of what has been termed *markedness reversal* (e.g. Croft 1990, Battistella 1996). In languages with DOM, it will be those direct objects which have the most marked properties in animacy and/or in definiteness and therefore most typically resemble subjects, that will receive case marking.

We must consider two different types of markedness here (Comrie 1989, de Swart 2003). On the one hand, we are dealing with the semantic markedness of objects: semantic properties like animacy and definiteness determine whether an object is marked or unmarked. On the other hand, we discuss the morphological markedness of objects, i.e., case marking. DOM is based on the relation between semantic and morphological markedness: if an object is semantically marked, it is likely to be morphologically marked (i.e., case marked) as well. This relation between semantic and morphological markedness is motivated by two general principles that underlie language organization, that is, principles of economy and iconicity. By economy principles, it is undesirable to morphologically mark objects, since morphological complexity is costly to process. The influence of iconicity principles, on the other hand, can be seen as follows: the complexity of an object at one level should be reflected in its complexity at some other level. So, in order to be iconic, the semantic markedness of an object should be reflected in its morphological form which, in the case of DOM, means that it will be case marked. In languages with DOM, the interaction of principles of economy and iconicity make that case marking is most

forcefully compelled at those objects which are semantically the most marked, whereas objects that are semantically unmarked will not be marked with case (de Swart 2003).

If overt case marking of direct objects depends on both animacy and definiteness, this is called two-dimensional DOM (Aissen 2003). By crossing both prominence scales, Aissen illustrates her cross-linguistic account of two-dimensional DOM in the form of a lattice, as shown in Figure 1:

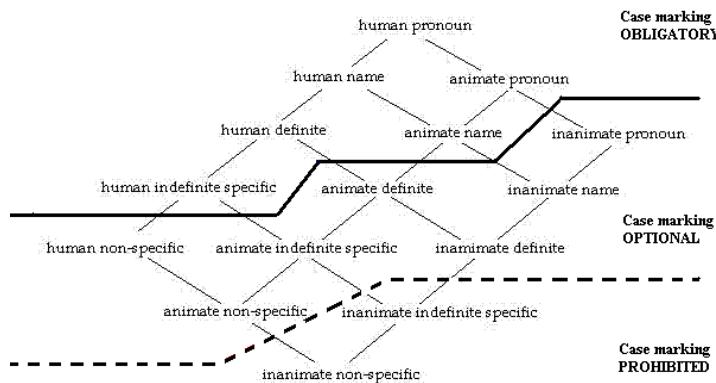
Figure 1: Two-dimensional DOM (adopted from Aissen 2003)



This lattice should be read from the top down: human pronouns outrank all other elements in prominence and should therefore be the most susceptible to DOM; inanimate non-specifics are outranked by all other elements, and should be least susceptible (Aissen 2003). Languages with two-dimensional DOM generally make a three-way distinction between their direct objects: objects for which case marking is obligatory, objects for which it is optional, and objects for which case marking is prohibited. These three sets of objects all cover a part of the

lattice in Figure 1. The exact part that each set takes in the lattice may vary per language, but when a language makes a three-way split, those objects that are on the upper side of highest split will obligatorily receive case marking, and those below the lower split will always lack a case marker. For those objects in between the two splits, the case marker will be optional. In the following figure, I placed the cut-off points for the DOM system in Hindi in the lattice³:

Figure 2: Two-dimensional DOM in Hindi



This figure shows that DOM in Hindi perfectly follows the cross-linguistic predictions: the least prominent objects are never case marked, whereas the most prominent objects are obligatorily case marked. For the category of objects in between, the case marker is optional.

I have shown in the previous section that both animacy and definiteness influence DOM in Chinese as well: hence, we are dealing

³ I adopted the exact cut-off points for Hindi from Aissen (2003).

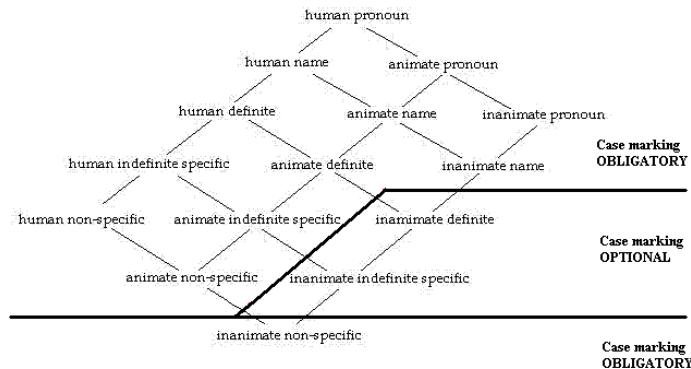
with an instance of two-dimensional DOM. With Aissen's (2003) cross-linguistic DOM account in the back of our minds, I will examine whether the Chinese DOM pattern fits her predictions.

2.4 Chinese DOM vs cross-linguistic DOM

In this section, I will investigate whether the influence of animacy and definiteness on Chinese DOM is comparable to their influence on DOM systems cross-linguistically. If we first consider the influence of animacy, we can see that it corresponds to Aissen's cross-linguistic predictions: if scrambled objects have high prominence in terms of animacy (i.e., if they are human or animate), they are obligatorily marked with *ba*, and if they have low prominence (i.e., if they are inanimate), *ba* is optional.

However, the influence of definiteness on Chinese DOM does not correspond to the cross-linguistic tendency: in Chinese, non-specific scrambled objects are obligatorily marked with *ba*, whereas for specific and definite scrambled objects, *ba* is optional. As for definiteness, it is thus not the *most* marked, but the *least* marked objects that are obligatorily case marked. The dimensions of animacy and definiteness seem to work in opposite directions. This can be illustrated more clearly in the lattice for two-dimensional DOM. When we place the cut-off points for Chinese DOM in this lattice, we see that the Chinese DOM system does not flow from the top down like other DOM systems do:

Figure 3: DOM in Chinese



Let us reconsider Aissen's cross-linguistic prediction that if a direct object must be case marked in a language with DOM, then all objects that are more prominent must be case marked as well. The Chinese data conflict with this prediction: both the least prominent and the most prominent scrambled objects are obligatorily case-marked, but not every type of object in between.

Furthermore, we saw that languages with two-dimensional DOM generally distinguish three sets of objects, which all cover a part of the lattice: one set of objects for which case marking is obligatory, one set for which it is optional, and a third set for which case marking is prohibited. For Chinese, however, we can only find back two categories of objects in the lattice: those objects for which case marking is obligatory and those for which case marking is optional. As I mentioned in the beginning of this chapter, the third category of direct objects, for which case marking is prohibited, also exists in Chinese. This category of objects, however, cannot be placed anywhere in the lattice.

This is due to the fact that the set of objects for which case marking is prohibited in Chinese is not defined on a semantic basis. The reason why certain direct objects cannot be case marked is of syntactic nature. I have shown that direct objects in canonical, postverbal position can never be case marked. Reconsider the following sentence:

(45) *Ta chi le (*ba) na-ge pingguo.*
 he eat PRT BA that-CL apple
 'He ate that apple.'

The set of objects for which case marking is prohibited is determined by word order: in a canonical SVO sentence, the object is never case marked, regardless of its semantic features. Only when direct objects are scrambled to the preverbal position, the case marker can be used. Once the direct object is in preverbal position, semantic features of animacy and definiteness determine whether case marking is obligatory or optional, but it is the syntactic position of the direct object that determines whether a case marker can be used at all.

In her analysis, Aissen does not discuss any syntactic features that influence DOM. Therefore, Aissen's analysis is not sufficient to account for all DOM patterns cross-linguistically: in order to explain the Chinese DOM system, the influence of word order on DOM must be taken into account as well.

2.5 Conclusion

I have shown in this chapter that Chinese exhibits a pattern of Differential Object Marking. I illustrated that this pattern deviates from other cross-linguistic instances of two-dimensional DOM, which are based on semantic properties, that is, animacy and definiteness of the direct object only. Even though these semantic factors do play a role in Chinese DOM as well, I showed that object marking in Chinese is first of all syntactically driven: direct objects can only receive *ba* if they are scrambled preverbally. In the next chapter, I will investigate how exactly the dimension of word order influences the Chinese DOM system.

3

Word Order

We saw in the previous chapter that, besides the semantic dimensions of animacy and definiteness, word order also influences Chinese DOM: direct objects in postverbal position can never be case marked, whereas objects in preverbal position can. In this chapter, I will take a closer look at the dimension of word order. I will examine the syntactic role of word order in Section 3.1, and I will investigate how word order is involved in semantics in Section 3.2. In the next section, I will show how word order influences the Chinese DOM pattern, and I will come to conclusions in Section 3.4.

3.1 The syntactic function of word order

A well-known characteristic of analytic languages is that, because of the general lack of morphology, word order is of utmost importance. For instance, word order in analytic languages generally marks syntactic relationships. If in a language the subject and the object are

placed on opposite sides of the verb, their grammatical functions are distinguished by their syntactic position. This way of argument discrimination is just as effectively as when either or both the subject and the object would be marked with case. If word order has this distinguishing function, case marking of either the subject or the object with the purpose of discrimination is considered to be redundant (Siewierska and Bakker to appear). For this reason, SVO is considered the most economic word order for isolating languages (Sinnemäki 2006). As we already saw in the previous chapter, word order in Chinese is SVO indeed: in a canonical Chinese sentence, the subject is placed left from the verb and the object is placed on the right side. Consider the following sentences:

(46) *Wo da le ta.*
 I hit PRT he
 'I hit him.'

(47) *Ta da le wo.*
 he hit PRT I
 'He hit me.'

In (46), *wo* 'I' is the one who hit, whereas in (47), *wo* 'I' is the one who was hit. The only difference between the sentences is the order of the constituents: the syntactic position of the arguments determines their grammatical function.

Although Chinese is an isolating language, we saw in the previous chapter that the language does have case marking. Chinese

exhibits a pattern of differential object marking, in which some objects are obligatorily case marked, some are optionally marked with *ba* and for some objects case marking is prohibited. We saw in Chapter 2 that the category of direct objects for which case marking is prohibited is the set of objects in postverbal position. This is illustrated again in (48):

(48) **Ta da le ba wo.*
 he hit PRT BA I
 'He hit me.'

Taking the distinguishing function of word order into account, it is not surprising that case marking is prohibited for this particular set of objects. Since the postverbal position in Chinese is the prototypical syntactic position for objects, word order already provides the necessary information to discriminate the grammatical roles of the arguments, by which case marking would be superfluous.

Direct objects in Chinese can only be marked with *ba* when they are scrambled preverbally, as shown in the sentence below:

(49) *Ta ba wo da le.*
 he BA I hit PRT
 'He hit me.'

When an object is scrambled, both the subject and the object are on the same side of the verb: the word order changes from SVO into SOV. In an SOV sentence, the subject and the object can no longer be distinguished from each other by means of their position relative to the

verb. This can be solved by marking either or both the subject and the object with case. If the subject or the object receives a case marker, it can be identified regardless of its position in the sentence: the case marker takes over the distinguishing function of the verb.

Whereas case marking can thus be considered superfluous in SVO languages to distinguish subjects from objects, its function appears to be very effective in SOV languages. This relationship between word order and case marking was already noted by Greenberg in 1963. Based on his vast typological research, he established a number of linguistic universals, including the implicational universal 41:

‘if in a language the verb follows both the nominal subject and nominal object as the dominant order, the language almost always has a case system’ (Greenberg 1963: 96).

It follows from this universal that if a language would change its word order from SVO into SOV, it is to be expected that it develops case marking (Sinnemäki 2006). This is indeed attested in Kamti Tai, an isolating language spoken in Myanmar, which changed its word order from SVO to SOV (probably due to language contact). As a consequence, the particles that were used to mark definiteness in Kamti Tai have developed into object marking particles (Khanitannan 1986, in Sinnemäki 2006). In Chinese, however, the dominant word order still is SVO. In a canonical Chinese sentence, the arguments are distinguished by their syntactic position in relation to the verb, and *ba* is never used for reasons of argument discrimination. Only when the object is

scrambled by which the order of the constituents changes into SOV, word order no longer suffices in distinguishing subjects from objects. In these cases, an object marker can be used to determine the grammatical roles of the arguments.

I have shown here that the SVO word order has an important syntactic function in Chinese. The position of the arguments with respect to the verb generally determines their grammatical role in the sentence. However, word order does not only provide syntactic information; it also contains semantic information. I will examine the semantics of word order in the following section.

3.2 The semantics of word order

Generally, subjects and objects are associated with a number of prototypical properties, which I already mentioned in the previous chapter. These prototypical features of subjects and objects are schematically represented in (50):

(50) SUBJECT	OBJECT
Agent	Patient
Animate	Inanimate
Definite	Indefinite
Specific	Non-specific
Thematic	Rhematic

The properties that are associated with subjects prototypically contribute to high prominence: high animate and high definite nouns

are generally more salient than inanimate and indefinite nouns. Consequently, prototypical subjects are semantically more prominent than prototypical objects. This is indeed attested by Dahl (2000), who found in a corpus of spoken Swedish that about two third of all subjects and 92% of the transitive subjects are animate. Furthermore, in a corpus study on simple transitive sentences in Norwegian, Øvrelid (2004) found that in 97.6% of the cases the subject is higher than or equal to the object in animacy, and in 82.3% of the cases the subject is higher than or equal to the object in definiteness. In some languages, like Japanese, there is even an absolute restriction on the animacy of transitive subjects: the subject of a transitive sentence may not be inanimate (Jacobsen 1992).

Cross-linguistically, the high prominence of subjects tends to be reflected in the syntactic structure of the sentence as well: in over 80% of the world's languages, subjects are placed in front of objects (Hawkins 1983). This tendency is functionally explained by the iconic motivation that linguistic structure to some extent reflects information structure. The tendency of subjects to precede objects can be seen an instance of temporal iconicity, also known as the sequential order principle: the sequential order of events described is mirrored in the speech chain. Jakobson (1971) illustrated this by Caesar's famous dictum *veni, vidi, vici* 'I came, I saw, I conquered': the linear order of these three words reflects the chronological order of the events that they describe. The same principle can be said to underlie the tendency for subjects to precede objects. Consider a sentence like *Tom hit John*. In this sentence, the hitting starts with Tom: he is the agent, who initiates the action. The hitting ends up having consequences for John: John is

the patient, who is affected by the action. This is linguistically reflected in the fact that Tom occurs earlier in the sentence than John.

In Chinese, the subject generally precedes the verb, whereas the object normally follows the verb. In some cases, however, it is possible to place subjects on the other side of the verb, as the following examples show (cf. Li and Thompson 1981):

(51) *Ren* *lai* *le.*
 person come PRT

'The person(s) came.'

(52) *Lai* *le* *ren* *le.*
 come PRT person PRT

'A person/some persons came.'

In both (51) and (52), the bare noun *ren* 'person' is the subject, but there is a difference in meaning between the two sentences. When the subject is in sentence-initial position, like in (51), it is interpreted as definite, whereas the same subject gets an indefinite reading if it is in postverbal position, as in (52).

This relation between structural position and the definiteness of an element is not unique for the Chinese language. Consider, for instance, the following examples from Finnish in (53)-(54) (Karlsson 1983):

(53) *Auto on kadulla.*
 car is in.street
 'The car is in the street.'

(54) *Kadulla on auto.*
 in.street is car
 'There is a car in the street.'

These sentences show that the interpretation of the subject depends on its syntactic position: when *auto* 'car' is in sentence-initial position, as in (53), it gets a definite reading, but it is interpreted as indefinite when it is postverbal and follows the locative, as shown in (54).

Something similar can be seen in Russian. A canonical locative sentence in Russian consists of a subject NP that is followed by a locative phrase, as in (55). In an existential-locative sentence, on the other hand, the canonical order is reversed and the locative phrase precedes the subject, as in (56) (cf. Beaver *et al.* to appear):

(55) *Kniga na stole.*
 book on table
 'The book is on the table.'

(56) *Na stole jest kniga.*
 on table COP book
 'There is a book on the table.'

In (55), where *kniga* 'book' is in sentence-initial position, it gets a definite reading. However, in the existential-locative construction in (56), *kniga* gets an indefinite reading. The position in the sentence determines how the bare noun should be interpreted.

In Dutch, there is a restriction on the definiteness of subjects in existential sentences. Consider the following sentences:

(57) *Er lagen koekjes op tafel.*
 there laid cookies on table
 'There were cookies on the table.'

(58) **Er lag koekje op tafel.*
 there laid cookie on table.

The bare plural *koekjes* 'cookies' in (57) is interpreted as indefinite. When referring to a single cookie, one cannot use a bare singular noun in this sentence, contrary to the Russian and Finnish examples above. To refer to a single cookie, a singular noun must be lexically marked for number, for instance by the indefinite article *een* 'a', which gives the subject in (59) an indefinite reading:

(59) *Er lag een koekje op tafel.*
 there laid a cookie on table
 'There was a cookie on the table.'

The definite article *het* 'the', on the other hand, makes the entity to which a noun refers specific and identifiable in the context. The subject

NP *het koekje* in (60) can no longer refer to any cookie, and the subject can only have a definite reading. We see that when the subject is a definite NP, the existential sentence is ungrammatical:

(60) **Er lag het koekje op tafel.*
 there laid the cookie on table
 'There was the cookie on the table.'

Coming back to the Chinese examples, we see that when the bare noun *ren* 'person' is changed into a lexically marked indefinite NP *yi-ge ren* 'a person', it can no longer be in the canonical, sentence-initial subject position:

(61) **Yi-ge ren lai le.*
 one-CL person come PRT
 'A person came.'

Unlike the bare noun, the subject *yi-ge ren* 'a person' is marked for indefiniteness by the numeral determiner *yi*, by which a definite interpretation is blocked. Because the subject in (61) can only have an indefinite reading, it cannot be placed in the sentence-initial position.

Clearly, the position of the subject influences the interpretation of the subject. The standard subject-position requires a definite reading, whereas the postverbal position requires an indefinite reading. How can this definiteness effect of word order on subjects be explained? Let us sum up what we have established so far. We have seen that elements that are definite or specific are prototypically associated with

subjecthood. We furthermore saw that subjects generally precede objects in the languages of the world. There is thus a similarity between the preverbal position and subjects in Chinese: both of them are generally associated with properties that contribute to high prominence. The semantic properties that are prototypical for subjects can be considered as prototypical properties for the preverbal position, and atypical subject features (that is, the prototypical features for objects) as prototypical properties for the postverbal position:

(62)	PREVERBAL	POSTVERBAL
	Agent	Patient
	Animate	Inanimate
	Definite	Indefinite
	Specific	Non-specific
	Thematic	Rhematic

Just like grammatical function, word order can be regarded as a dimension of prominence: because of the association of the preverbal position with subject-like properties, the preverbal position is generally more prominent than the postverbal position. In (63), this is expressed in the form of a prominence scale:

(63) *Word order scale*: Preverbal > postverbal

We have seen here that the syntactic position of the subject influences its interpretation in terms of definiteness. However, the definiteness effect of word order is not restricted to subjects. To illustrate this, I will

give some examples of how word order affects the definiteness of direct objects in Chinese.

As we already saw in the previous chapter, direct objects in Chinese can sometimes occur either pre- or postverbally. The examples below show that a change in syntactic position of the object yields a shift in interpretation. Whereas the bare noun object *pingguo* 'apple' can get either a definite or an indefinite reading when it is in postverbal position, as in (64), it can only get a definite reading when placed in preverbal position, as in (65):

(64) *Ta chi le pingguo.*
 he eat PRT apple
 'He ate an apple/apples' / 'He ate the apple(s).'

(65) *Ta pingguo chi le.*
 he apple eat PRT
 'He ate the apple(s).'

This effect of word order on the interpretation of the object can be explained in terms of the prominence scale of word order. We saw that preverbal positions in a Chinese sentence are associated with properties that contribute to high prominence. If an element is placed preverbally, it can only get a definite reading because of the association of the preverbal position with high prominence. If the bare noun object is replaced by an indefinite noun phrase *yi-ge pingguo* 'an apple', the sentence becomes ungrammatical:

(66) **Ta yi-ge pingguo chi le.*
 he one-CL apple eat PRT
 'He ate an apple.'

The numeral determiner *yi* 'one' marks the object NP for indefiniteness, by which it can no longer get a definite interpretation and it cannot be placed preverbally (Yang in prep.). However, there are some features that can give *yi-ge pingguo* 'an apple' a more specific reading, as we saw in the previous chapter. When the predicate *chi-le* is changed into an accomplishment predicate like *chi-wan le*, the indefinite object *yi-ge pingguo* is interpreted as a specific apple (e.g., Sybesma 1992). A specific or definite reading can also be evoked by *dou* (e.g., Lee 1986, Liu 1997, Cheng 2006). When a lexically indefinite object NP gets a specific or definite interpretation because of these features, it becomes more prominent and as a consequence, it can occur preverbally, as shown in the following examples:

(67) *Ta yi-ge pingguo chi-wan le.*
 he one-CL apple eat-finish PRT
 'He ate up an apple.'

(68) *Ta yi-ge pingguo dou chi le.*
 he one-CL apple all eat PRT
 'He ate a whole apple.'

We have seen here that word order plays a role in the interpretation of subjects and objects. I have shown that preverbal positions in Chinese

are associated with subjecthood and therefore with high prominence. As a consequence, only definite and specific arguments can occur in preverbal position, regardless of their grammatical function in the sentence.

In the following section, I will show how the definiteness effect of word order can be incorporated in accounting for the Chinese DOM system.

3.3 Word order and Chinese DOM

In Section 3.1, we saw that the use of case marking in Chinese is licensed by a shift in word order: objects in postverbal position cannot be marked with case, whereas objects in preverbal position can. This can be explained as follows. We saw that subjects are prototypically in preverbal position, while objects are generally in postverbal position. This is the most economic order for the arguments: their grammatical relation can be determined by their position relative to the verb. When objects scramble, they end up in an atypical position for objects, as a result of which they can no longer be identified as objects merely by means of their position relative to the verb. In these cases, a case marker can be used to take over the distinguishing function of the verb: adding *ba* helps discriminating preverbal objects from subjects.

However, we saw in the previous chapter that case marking of preverbal objects is not always obligatory; under certain circumstances, *ba* can be omitted. Consider the following examples:

(69) *Ta* *(*ba*) *wo* *da* *le.*
 he BA I hit PRT
 'He hit me.'

(70) *Ta* (*ba*) *qiu* *da* *le.*
 he BA ball hit PRT
 'He hit the ball.'

These sentences illustrate that while animate scrambled objects in Chinese are obligatorily marked with *ba*, the case marker can be omitted if the object in preverbal position is inanimate. Furthermore, if a scrambled object is specific or definite, *ba* is optional, whereas *ba* is obligatory for non-specific scrambled objects, as shown again in (71)-(72):

(71) *Ta* *(*ba*) *yi-ge* *pingguo* *chi* *le.*
 He BA one-CL apple eat PRT
 'He ate an apple.'

(72) *Ta* (*ba*) *na-* *ge* *pingguo* *chi* *le.*
 He BA that-CL apple eat PRT
 'He ate that apple.'

The omission of the case marker on scrambled objects in Chinese is triggered by the prominence dimensions of animacy and definiteness, as already shown in Chapter 2. We also saw that the influence of animacy and definiteness on DOM is cross-linguistically explained in

terms of markedness: case marking is most forcefully compelled at those objects which are semantically the most marked, whereas objects that are semantically unmarked will not be marked with case (Aissen 2003). However, I showed that the Chinese DOM pattern cannot be accounted for in this way. In Chinese DOM, animacy and definiteness seem to work in opposite directions. On the one hand, the *most* marked, i.e., human and animate, objects are obligatorily case marked when scrambled, whereas *ba* is optional for inanimate scrambled objects. On the other hand, the *least* marked, i.e., non-specific, objects are obligatorily case marked when scrambled, whereas specific and definite scrambled objects are optionally marked with *ba*.

The way in which animacy influences Chinese DOM perfectly fits Aissen's (2003) cross-linguistic predictions: high prominent objects are obligatorily marked, and low prominent objects are optionally marked with *ba*. However, the way in which definiteness influences Chinese DOM can not be compared with how definiteness influences other DOM systems. The way in which the dimension of definiteness influences Chinese DOM seems to conflict with the way in which definiteness influences DOM cross-linguistically.

The problem lies in the fact that DOM is cross-linguistically determined by only one instance of markedness reversal. The markedness of animacy and definiteness features is established by their grammatical function: properties that are marked for objects are unmarked properties for subjects and vice versa. In Chinese DOM, however, we are dealing with the extra dimension of word order, which is not part of this instance of markedness reversal. As I explained above, word order affects the definiteness of NPs, in the

sense that arguments get a more specific or definite reading when they are in preverbal position than when they occur postverbally. This can be translated into markedness terms as well. Because of the association of the preverbal position with subjecthood, and therefore with definiteness, we can say that specificity and definiteness are unmarked properties for NPs in the preverbal position. At the same time, non-specificity is a marked property for an NP in this position. This is another instance of markedness reversal: what is marked for the preverbal position is unmarked for the postverbal position and vice versa. If we take this instance of markedness reversal into account, we can explain the influence of definiteness on Chinese DOM as follows: it is the most marked objects, i.e., non-specific NPs in preverbal position, that are obligatorily case-marked, and the least marked objects, i.e., specific and definite NPs in preverbal position, for which case marking is optional.

Hence, animacy and definiteness do not go hand in hand in determining the Chinese DOM system. The two dimensions influence DOM, but they are of influence in different domains. The role of animacy in Chinese DOM can be summarized as follows: when objects are animate or human, they have high prominence, which is a marked property concerning their grammatical function. As a consequence, the case marker is obligatory for human and animate scrambled objects. Inanimate objects have low prominence, which is an unmarked feature for objects. Consequently, the case marker for inanimate scrambled objects is optional. By contrast, the role of definiteness in Chinese DOM should be explained as follows. When objects are definite or specific, they have high prominence. Although this is a marked property for

objects, it is an unmarked property for the preverbal position. Therefore, *ba* is optional for definite and specific objects. Non-specific objects, on the other hand, have low prominence, which is a marked property for NPs in the preverbal position. For that reason, non-specific objects in preverbal position are obligatorily marked with *ba*. The prominence scale of definiteness should thus be understood in connection with the scale of word order to account for its influence on Chinese DOM.

3.4 Conclusion

In this chapter, I have shown how word order influences the Chinese DOM system. When both the subject and the object precede the verb, word order can no longer be used as a cue to distinguish between the grammatical roles of subject and object. As a result of this a case marker is added when objects are scrambled. Semantically, word order affects the definiteness of arguments: the preverbal position triggers a specific or definite interpretation. If the scrambled object has a specific or definite reading, it has the right, unmarked properties that are needed to occupy the preverbal position. Hence, the case marker can be omitted. If the scrambled object is non-specific, however, it has marked properties for this position in the sentence. Because of this, non-specific indefinite objects in preverbal position are obligatorily marked with *ba*.

Now that we have taken this dimension of word order taken into consideration, we can give a full OT account of the Chinese DOM system. I will do this in the next chapter.

4

An OT account of Chinese DOM

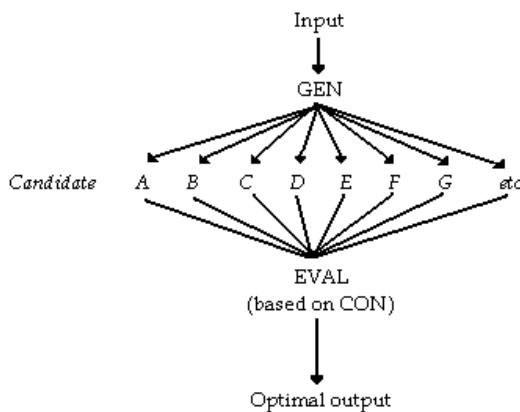
In the previous chapters, I investigated how animacy, definiteness and word order influence the Chinese Differential Object Marking pattern. In this chapter, I will give a formal account of Chinese DOM within an Optimality Theoretic framework. In Section 4.1, I will give an introduction to Optimality Theory (OT) and its application in the syntactic domain. I will explain Aissen's OT formalization of her cross-linguistic DOM findings in Section 4.2 and I will adapt this formalization to account for Chinese DOM pattern in Section 4.3.

4.1 An introduction to OT

Optimality Theory is a model of the system of linguistic knowledge a speaker of a language possesses (cf. Prince and Smolensky 1993/2004). The rules in this grammar are constraints expressing general statements, which can be in conflict with each other. In contrast with traditional models, the constraints in OT are violable: one constraint can be

violated in order to satisfy another, stronger constraint. Constraints are universal, which means that all languages share the same set of constraints. The relative weight of the constraints, however, differs per language. Some constraint that is ranked very high in one language may be overruled by a great number of constraints in another language. Language-particular constraint rankings lead to cross-linguistic variation. The basic architecture of OT is illustrated in the following figure (adopted from Blutner *et al.* 2006):

Figure 4: OT basic architecture



The Generator (GEN) provides an in principle infinite number of possible output candidates based on some input. The Evaluator (EVAL) uses the language-particular ranking of all universal constraints (CON) to evaluate these output candidates. The candidate that best satisfies the constraints, i.e., that violates less and/or lower ranked constraints than its alternatives, is selected by the Evaluator as the optimal output.

The constraints that play a role in OT can be phonological, syntactic, pragmatic or semantic in nature. Dependent on the direction of optimization - that is from meaning to form, or from form to meaning - they become of importance in the evaluation of a candidate (Hendriks and de Hoop 2001). An OT evaluation process is represented in a so-called tableau. In such a tableau, the constraints are placed in the top row from left to right, the leftmost constraint being the highest ranked. Input candidates are listed in the left column. The violations of constraints that a candidate makes are marked by an asterisk in the corresponding cell. I will show how exactly a tableau representation works in the following section.

Aissen has developed an analysis to account for DOM within an OT Syntactic framework. Optimality Theoretic Syntax is an application of OT in the syntactic domain. In OT Syntax, the direction of optimization is from meaning to form. The input is a semantic structure, that is, an intended meaning. Possible syntactic representations of this intended meaning are evaluated by well-formedness constraints, after which the optimal candidate is selected. An OT Syntactic analysis makes it possible to express the universal motivation for DOM by formulating universal constraints, while cross-linguistic variation in DOM systems can be accounted for at the same time by language-particular constraint rankings. In the next section, I will describe how Aissen accounts for cross-linguistic instances of DOM within OT.

4.2 Aissen (2003): an OT account of cross-linguistic DOM

In Chapter 2, I described Aissen's generalization of DOM: the higher in prominence a direct object, the more likely it will be overtly marked (Aissen 2003). The level of prominence is established by the dimensions of animacy and definiteness. Even though the general thought behind DOM seems to be universal, languages differ in exactly which objects will receive overt marking and whether object marking is obligatory or optional. Aissen captures this universality on the one hand and cross-linguistic variation on the other within an OT Syntactic framework of DOM.

Aissen assumes that the central notion underlying DOM is markedness reversal: the high prominence that is marked for objects and motivates object marking is exactly the prominence that is unmarked for subjects. In order to incorporate the notion of markedness reversal in her OT analysis and express the relative markedness of objects with respect to animacy and definiteness in universal constraints, Aissen applies the operation of Harmonic Alignment, as proposed by Prince and Smolensky (1993/2004).

The basic idea of Harmonic Alignment is that an element which is structurally prominent attracts elements which are prominent along some relevant dimension, whereas an element which is low in prominence attracts elements which are low prominent along a relevant dimension. Harmonic Alignment is applied on pairs of scales, connecting the high-ranked element on a binary scale X to the elements on another scale Y from left to right. In the same way, the low-ranked

component on the binary scale is connected to the components on the other scale from right to left. This generates two harmony scales which express the relative markedness of each such connection.

In the case of DOM, the relative markedness of objects is determined by features of animacy and definiteness: the more animate or definite a direct object is, the more marked it is. Therefore, the relevant dimensions for DOM are the dimension of grammatical function on the one hand, and the dimensions of animacy and definiteness on the other. These dimensions can be expressed as prominence scales, as I have shown in Chapter 2. The prominence scales are repeated in (73) – (75):

(73) *Relational scale:* Subject > Object

(74) *Animacy scale:* Human >Animate > Inanimate

(75) *Definiteness scale:* Pronoun > Proper Noun > Definite NP >
 Indefinite Specific NP > Indefinite Non-Specific NP

When Harmonic Alignment is applied to the prominence scale of animacy on the one hand, and the binary relational scale on the other, this yields the harmony scales in (76a) and (76b). The harmony scales express the relative markedness of subjects (76a) and objects (76b) in terms of animacy. The most harmonic or least marked combinations are presented on the leftmost side of the scales:

(76) a. Subj/Hum > Subj/Anim > Subj/Inan
 b. Obj/Inan > Obj/Anim > Obj/Hum

The harmony scale in (76b) expresses that inanimate objects are less marked than animate objects, which on their turn are less marked than human ones. A constraint hierarchy can be derived from this harmony scale by inverting the ranking in (76b) and by prefixing the AVOID operator ‘*’:

(77) *OBJ/HUM >> *OBJ/ANIM >> *OBJ/INAN

This constraint hierarchy expresses the economic motivation that marked configurations should be avoided. It qualifies human objects as the most marked ones; they are to be avoided more than animate and inanimate objects, respectively. However, in languages with DOM, the most marked objects are *not* avoided. Rather, the marked class of objects is overtly marked with case, and the unmarked class does not receive any morphological mark (Bossong 1985). In order to express the overt marking of marked objects in constraints, the constraint hierarchy in (77) should be combined with a constraint which expresses the iconic motivation that the semantic markedness of an object should be reflected in its morphological form as well. Aissen introduces the iconicity constraint ‘Star Zero’, expressing that morphological case should be expressed. This constraint is presented in (78):

(78) * \emptyset_C : penalize the absence of case.

This iconicity constraint can be combined with the hierarchy in (77) by another OT technique, which is called Local Conjunction. Local Conjunction is an operation that is based on the idea that violating two constraints is worse when both constraints are in the same location than when they are separately violated. Local Conjunction ties together two separate constraints, or a constraint and a constraint hierarchy, in this way creating a new constraint. The local conjunction of C1 and C2 in domain D, represented as C1 & C2, is violated when there is some domain of type D in which *both* C1 and C2 are violated. The local conjunction of C1 & C2 is universally ranked above the two separate constraints C1 and C2.

In the case of DOM, we can say that the presence of a semantically marked object is bad, but that it is even worse not to reflect this semantic markedness in its morphological form. Aissen translates this into OT terms by applying Local Conjunction to the subhierarchy in (77) on the one hand and $^* \emptyset_C$ on the other, by which case marking is most forcefully compelled on the most marked objects. This local conjunction yields the following subhierarchy of complex constraints:

$$(79) \quad ^* \text{OBJ/HUM} \& ^* \emptyset_C >> ^* \text{OBJ/ANIM} \& ^* \emptyset_C >> ^* \text{OBJ/INAN} \& ^* \emptyset_C$$

The hierarchy in (79) expresses that it is worse to leave human objects unmarked than it is to not mark animate and inanimate objects, respectively.

In the same vein, the prominence scale of definiteness is aligned with the relational scale, yielding the harmony scales in (80) and the corresponding constraint hierarchies in (81):

- (80) a. Subj/Pn > Subj/Noun > Subj/Def > Subj/Spec >
Subj/Nspec
- b. Obj/Nspec > Obj/Spec > Obj/Def > Obj/Noun > Obj/Pn

- (81) a. *SUBJ/NSPEC >> *SUBJ/SPEC >> *SUBJ/DEF >>
*SUBJ/NOUN >> SUBJ/PN
- b. *OBJ/PN >> *OBJ/NOUN >> *OBJ/DEF >> *OBJ/SPEC >>
*OBJ/NSPEC

The hierarchy in (81b) must be understood as follows: pronoun objects are the most marked configurations; they should be avoided more than proper noun, definite, specific indefinite and non-specific indefinite objects, respectively.

Again, in languages with DOM, objects that are semantically the most marked are not avoided; they receive a morphological marker instead. Therefore, Local Conjunction is also applied to the constraint hierarchy in (81b) and $* \emptyset_C$, in this way creating the following constraint hierarchy:

- (82) *OBJ/PN & $* \emptyset_C$ >> *OBJ/NOUN & $* \emptyset_C$ >> *OBJ/DEF & $* \emptyset_C$ >>
OBJ/SPEC & $ \emptyset_C$ >> *OBJ/NSPEC & $* \emptyset_C$

The complex constraint hierarchies in (79) and (82) describe that if a language with DOM marks any objects, human and pronoun objects are the first ones that will receive case marking. The constraints in both hierarchies are iconicity constraints: they link complexity in meaning to complexity in structure. Remember that constraints in OT are universal: the constraint hierarchies in (79) and (82) must apply to every language. The iconicity constraints, however, compel case on *all* objects. This would mean that all languages would mark all of their objects with case, and languages with differential object marking could not be accounted for. Therefore, Aissen proposes the following economy constraint, which penalizes overt case marking.

(83) *STRUC_C: penalizes a value for the morphological category CASE

The constraint *STRUC_C can be considered an economy constraint, because using morphological marking is costly. This economy constraint can be in conflict with either or both the constraints hierarchies in (79) and (82): the relative strength of the constraints differs per language. The economy constraint may be ranked higher than all iconicity constraints of the hierarchies in (79) and (82). In languages with such a constraint ranking, no case marking will be used at all. In other languages, all iconicity constraints may overrule *STRUC_C: in languages like these, all objects will be case marked. In some cases, *STRUC_C may be overruled by some iconicity constraints, but not by others. If this is the case, a differential object marking system arises. Let us for instance reconsider the DOM pattern of Malayalam, as discussed in Chapter 2. In Malayalam, human and animate objects do receive a

case marker, whereas inanimate objects do not. This can be expressed in OT terms in the following way. In the constraint ranking of Malayalam, ${}^*{\text{STRUC}_C}$ intervenes in the constraint hierarchy in (82): it is ranked higher than ${}^*{\text{OBJ}/\text{INAN}} \ \& \ {}^* \emptyset_C$, but it gets overruled by ${}^*{\text{OBJ}/\text{HUM}} \ \& \ {}^* \emptyset_C$ and ${}^*{\text{OBJ}/\text{ANIM}} \ \& \ {}^* \emptyset_C$. I will illustrate the interaction of the relevant constraints in Tableaux 1 and 2 below. In Tableau 1, the evaluation of a human object is presented, and the evaluation of an inanimate object is given in Tableau 2.

Tableau 1. A human object in Malayalam

Input: object = child	${}^*{\text{OBJ}/\text{HUM}} \ \& \ {}^* \emptyset_C$	${}^*{\text{OBJ}/\text{ANIM}} \ \& \ {}^* \emptyset_C$	${}^*{\text{STRUC}_C}$	${}^*{\text{OBJ}/\text{INAN}} \ \& \ {}^* \emptyset_C$
$\text{Child}.\text{ACC}$			*	
Child	!*!			

In Tableau 1, we see that both candidates⁴ violate one constraint, indicated by the asterisks. The constraint ${}^*{\text{STRUC}_C}$ that is violated by the first candidate (the case marked object) is ranked lower than ${}^*{\text{OBJ}/\text{HUM}} \ \& \ {}^* \emptyset_C$, which is violated by its competitor. Because of the difference in strength of the violated constraints, the violation of the highest ranked constraint is crucial, indicated by the exclamation mark. Therefore, the second candidate (the caseless object) loses the competition. This yields the case marked object as the optimal expression of the intended meaning, denoted by the pointing finger.

⁴ Note that the number of possible candidates is in principle infinite, but I only listed the two relevant candidates for the expression of the intended meaning.

Tableau 2. An inanimate object in Malayalam

Input: object = coconut	*OBJ/HUM & * \emptyset_C	*OBJ/ANIM & * \emptyset_C	*STRUCC	*OBJ/INAN & * \emptyset_C
Coconut. ACC			*!	
☞ Coconut				*

In Tableau 2, both candidates violate one constraint as well. However, the unmarked object now violates the constraint *OBJ/INAN & * \emptyset_C , which is ranked lower than *STRUCC. The violation of *STRUCC by the first candidate is now fatal, by which the optimal output for the syntactic representation of the inanimate object is the unmarked object.

We saw in Chapter 2 that in Hebrew, it is the dimension of definiteness that influences DOM: definite objects are case marked, whereas indefinite objects are not. The interaction of the iconicity constraints concerning definiteness and the economy constraint *STRUCC is illustrated in the Tableaux5 below.

Tableau 3. A definite object in Hebrew

Input: object = the war	*OBJ/PN & * \emptyset_C	*OBJ/NOUN & * \emptyset_C	*OBJ/DEF & * \emptyset_C	*STRUCC	*OBJ/INDEF & * \emptyset_C
☞ War. ACC				*	
War			*!		

In Tableau 3, the unmarked object violates *OBJ/DEF & * \emptyset_C , which is ranked higher than *STRUCC. The second candidate therefore loses the

⁵ For the sake of clarity, I replaced *OBJ/SPEC & * \emptyset_C and *OBJ/SPEC & * \emptyset_C by the covering constraint *OBJ/INDEF & * \emptyset_C .

competition, and the case marked object is the optimal way of expressing the intended meaning ‘the movie showed the war’.

Tableau 4. An indefinite object in Hebrew

Input: object = a war	*OBJ/PN & * Ø _C	*OBJ/NOUN & * Ø _C	*OBJ/DEF & * Ø _C	*STRUC _C	*OBJ/INDEF & * Ø _C
War. ACC				*!	
∅ War					*!

If the input is an indefinite object, as in Tableau 4, we see that the case marked candidate again violates *STRUC_C. However, the unmarked object now violates *OBJ/INDEF & * Ø_C, which is ranked lower than *STRUC_C in the constraint hierarchy of Hebrew. The violation of the economy constraint is now crucial, by which the unmarked object is evaluated as the optimal syntactic representation of the intended meaning.

In Aissen’s view, the constraints discussed in this section are sufficient to give an OT account of all DOM systems cross-linguistically: the general motivation for DOM is captured by the universal character of the constraints, and cross-linguistic differences can be accounted for by different language-particular constraint rankings. However, we saw that differential object marking in Chinese is determined by an additional factor, which Aissen does not take into consideration. The constraints she formulates do not describe the influence of word order on DOM. In the following section, I will take Aissen’s constraints as a starting point and I will reformulate them in such a way that the

dimension of word order is included, by which the Chinese DOM system can also be accounted for within an OT Syntactic framework.

4.3 An OT account of Chinese DOM

4.3.1 Animacy and Chinese DOM

In Chapter 2, I showed that to a certain degree, animacy influences Chinese DOM in the same way as it influences DOM in other languages: the most marked, i.e., human and animate, direct objects are obligatorily marked, whereas for unmarked (inanimate) objects, the case marker is optional. This can be seen from the sentences I presented in Chapter 2, repeated in (84)-(86):

(84) *Ta* *(*ba*) *laoshi* *tuidao* *le.*
 he BA teacher push.over PRT
 'He pushed over the teacher.'

(85) *Ta* *(*ba*) *she* *dasi* *le.*
 he BA snake hit.dead PRT
 'He killed the snake.'

(86) *Ta* (*ba*) *pingguo* *chi* *le.*
 he BA apple eat PRT
 'He ate the apple(s).'

The iconicity constraints as formulated by Aissen seem thus applicable to Chinese DOM as well. The only difference with Aissen's DOM

analysis is that the influence of animacy on Chinese DOM is restricted to scrambled objects. Objects in prototypical postverbal position never receive a morphological marker. In order to account for this additional factor in OT-terms, I will subdivide the economy constraint ${}^*\text{STRUCC}$ into the subconstraints ${}^*\text{STRUCC}_C/\text{UNSCR}$ and ${}^*\text{STRUCC}_C/\text{SCRAM}$, the former being ranked higher than the latter:

(87) ${}^*\text{STRUCC}_C/\text{UNSCR} \gg {}^*\text{STRUCC}_C/\text{SCRAM}$

The constraint hierarchy in (87) describes that it is worse to use case marking for unscrambled elements than it is to mark scrambled element with case. This can again be motivated by principles of economy. We saw that in SVO word order, which is the canonical order in Chinese, the grammatical functions of the constituents can be determined by their position relative to the verb and case marking is redundant. If the case marker does not add any syntactic information, it is preferably not expressed for reasons of economy. However, if the word order is changed into SOV, the verb loses its distinguishing function. In this case, the case marker is no longer superfluous: it helps discriminating the arguments of the sentence.

In Chinese, ${}^*\text{STRUCC}_C/\text{UNSCR}$ outranks all constraints I discussed above, by which an OT evaluation will never yield case marked objects in postverbal position as optimal candidates. Since human and animate scrambled objects are obligatorily case marked, the constraints ${}^*\text{OBJ/HUM} \& {}^*\emptyset_C$ and ${}^*\text{OBJ/ANIM} \& {}^*\emptyset_C$ must outrank ${}^*\text{STRUCC}_C/\text{SCRAM}$ in the Chinese constraint ranking. On the other hand, inanimate objects

are optionally case marked, which means that $^*\text{OBJ}/\text{INAN} \ \& \ ^*\emptyset_C$ and $^*\text{STRUCC}_C/\text{SCRAM}$ should be ranked equally high.

The interaction of the constraints in (77) and (87) is shown in the OT Tableaux below. For clarification, I merged the conjoint constraints $^*\text{OBJ}/\text{HUM} \ \& \ ^*\emptyset_C >> ^*\text{OBJ}/\text{ANIM} \ \& \ ^*\emptyset_C$ into one simplified constraint $^*\text{OBJ}/[\text{HUM}/\text{ANIM}] \ \& \ ^*\emptyset_C$, expressing that both human and animate objects must be marked with case. The constraint interaction for a human scrambled object is illustrated in Tableau 5, for an animate scrambled object in Tableau 6 and for an inanimate scrambled object in Tableau 7⁶.

Tableau 5. A human scrambled object

Input: 'He hit the teacher'	$^*\text{OBJ}/[\text{HUM}/\text{ANIM}] \ \& \ ^*\emptyset_C$	$^*\text{STRUCC}_C/\text{SCRAM}$	$^*\text{OBJ}/\text{INAN} \ \& \ ^*\emptyset_C$
<i>Ta laoshi da le.</i> He teacher hit PRT	*!		
<i>Ta ba laoshi da le.</i> He BA teacher hit PRT		*	

Tableau 6. An animate scrambled object

Input: 'He killed the snake'	$^*\text{OBJ}/[\text{HUM}/\text{ANIM}] \ \& \ ^*\emptyset_C$	$^*\text{STRUCC}_C/\text{SCRAM}$	$^*\text{OBJ}/\text{INAN} \ \& \ ^*\emptyset_C$
<i>Ta she dasi le.</i> He snake hit.dead PRT	*!		
<i>Ta ba she dasi le.</i> He BA snake hit.dead PRT		*	

⁶The Tableaux in this chapter give a simplified representation of the evaluation processes. For the sake of clarity, the higher ranked constraint $^*\text{STRUCC}_C/\text{UNSCR}$ is not expressed in the Tableaux. Furthermore, only scrambled objects are listed as possible candidates. For simplicity, the constraints making scrambled objects preferred to unscrambled objects are not discussed.

In Tableaux 5 and 6, both candidates violate one constraint. The constraint ${}^*\text{STRUCC}_C / \text{SCRAM}$ that is violated by the second candidates (the marked scrambled objects) is ranked lower than ${}^*\text{OBJ}/[\text{HUM}/\text{ANIM}] \ \& \ {}^*\emptyset_C$, which is violated by their competitors. The violation of the highest ranked constraint is thus crucial and the first candidates (the unmarked scrambled objects) lose the competition. This makes case marked human and animate scrambled objects the optimal expressions of the intended meanings.

Tableau 7. An inanimate scrambled object

Input: 'He ate the apple'	${}^*\text{OBJ}/[\text{HUM}/\text{ANIM}] \ \& \ {}^*\emptyset_C$	${}^*\text{STRUCC}_C / \text{SCRAM}$	${}^*\text{OBJ}/\text{INAN} \ \& \ {}^*\emptyset_C$
☞ <i>Ta pingguo chi le.</i> He apple eat PRT			*
☞ <i>Ta ba pingguo chi le.</i> He BA apple eat PRT		*	

In Tableau 7, each candidate violates one of the constraints again. However, the dotted line indicates that there is no difference in strength between ${}^*\text{OBJ}/\text{INAN} \ \& \ {}^*\emptyset_C$ and ${}^*\text{STRUCC}_C / \text{SCRAM}$. For that reason, both the case marked and the unmarked inanimate scrambled object are optimal syntactic representations of the input meaning 'he ate the apple'.

In this subsection, I showed that the interaction of Aissen's OT constraints concerning cross-linguistic varieties of DOM can account for the influence of animacy on Chinese DOM as well, if her proposed economy constraint ${}^*\text{STRUCC}_C$ is subdivided into ${}^*\text{STRUCC}_C/\text{UNSCR}$ and ${}^*\text{STRUCC}_C/\text{SCRAM}$. I will explain how the influence of definiteness on Chinese DOM can be formalized in OT terms in the following subsection.

4.3.2 Definiteness and Chinese DOM

Just like animacy, definiteness in Chinese DOM only plays a role in the differential marking of scrambled objects. This can again be accounted for via the two subconstraints ${}^*\text{STRUC}_C/\text{UNSCR}$ and ${}^*\text{STRUC}_C/\text{SCRAM}$ that I proposed in the previous section. However, there is an additional difference between the influence of definiteness on DOM patterns cross-linguistically and its role in Chinese DOM.

We saw that non-specific scrambled objects are obligatorily marked with *ba*, whereas the case marker can be omitted for specific and definite objects in preverbal position. Contrary to Aissen's cross-linguistic predictions, the case marker in Chinese is obligatory for the least marked objects, whereas *ba* is not obligatory for all objects that are more marked. I argued in the previous chapter that this can be accounted for if the dimension of word order is taken into consideration. High definite scrambled objects are marked nouns concerning their grammatical function, but they are unmarked nouns with respect to the preverbal position. We are dealing with another variation on markedness reversal here: in Chinese, the semantic markedness of objects in terms of definiteness is determined on the basis of their position in the sentence instead of their grammatical function. I will therefore not follow Aissen by applying Harmonic Alignment to the prominence scale of definiteness and the relational scale. Alternatively, I will align the scale of definiteness with the prominence scale of word order. This yields the harmony scales in (88), which in turn results in the universal constraint hierarchies in (89):

(88) a. Pre/Pn > Pre/Noun > Pre/Def > Pre/Spec > Pre/NSpec

b. Post/NSpec > Post/Spec > Post/Def > Post/Noun >

Post/Pn

(89) a. *PRE/NSPEC >> *PRE/SPEC >> *PRE/DEF >> *PRE/NOUN >>

*PRE/PN

b. *POST/PN >> *POST/NOUN >> *POST/DEF >>

*POST/SPEC >> *POST/NSPEC

The harmony scale in (88a) expresses that in pronouns in preverbal position are less marked than preverbal proper nouns, which on their turn are less marked than respectively definite, specific indefinite and non-specific indefinite NPs in preverbal position. The constraint hierarchy in (89a) is to be read as follows: non-specific indefinite NPs in preverbal position are the most marked configurations; they should be avoided more than specific indefinite NPs, definite NPs, proper nouns and pronouns in preverbal position, respectively.

It should be noted here again that in languages with DOM, the most marked combinations are not avoided; they receive a case marker instead. I will express this in OT constraints by applying Local Conjunction to the subhierarchy in (89a) on the one hand and the iconicity constraint $*\emptyset_C$ on the other. This results in the following subhierarchy of complex constraints:

(90) *PRE/NSPEC & $*\emptyset_C$ >> *PRE/SPEC & $*\emptyset_C$ >> *PRE/DEF & $*\emptyset_C$ >>

PRE/NOUN & \emptyset_C >> *PRE/PN & $*\emptyset_C$

This complex constraint hierarchy expresses that it is worse to leave non-specific NPs in preverbal position unmarked than it is not to mark specific NPs, definite NPs, proper nouns and pronouns in preverbal position, respectively. We saw that in Chinese, the cut-off point between obligatory and optional case-marking lies between non-specific and specific indefinite objects in preverbal position: scrambled non-specific indefinite preverbal objects are obligatorily case marked, while specific indefinite and definite objects can occur in preverbal position without *ba*. This can be translated into OT terms as follows: the economy constraint ${}^*\text{STRUCC}/\text{SCRAM}$ intervenes between the first and the second element on the hierarchy. ${}^*\text{PRE}/\text{NSPEC} \ \& \ {}^*\emptyset_c$ outranks ${}^*\text{STRUCC}/\text{SCRAM}$, since non-specific indefinite objects in preverbal position are obligatorily marked. The other types of objects in preverbal position are all optionally marked with *ba*, which means that ${}^*\text{STRUCC}/\text{SCRAM}$ and the remaining constraint conjunctions of the hierarchy in (90) are equally strong in Chinese. For the sake of clarity, I will merge these remaining conjoint constraints into one constraint, which requires that pronoun, proper noun, definite and specific indefinite objects in preverbal position must be marked with case. The result is the following, simplified constraint hierarchy:

$$(91) \quad {}^*\text{PRE}/\text{NSPEC} \ \& \ {}^*\emptyset_c >> {}^*\text{PRE}/[\text{SPEC}/\text{DEF}/\text{NOUN}/\text{PN}] \ \& \ {}^*\emptyset_c$$

Let us now look at the interaction among the constraints in (91) and ${}^*\text{STRUCC}/\text{SCRAM}$. The evaluation of possible expressions of the intended meaning 'he ate an apple' is schematically represented in Tableau 8:

Tableau 8. A scrambled non-specific indefinite object

Input: 'He ate an apple'	*PRE/NSPEC & *Ø _C	*STRUC _C / SCRAM	*PRE/[SPEC/ DEF/NOUN/PN] & *Ø _C
<i>Ta yi- ge pingguo chi le.</i> He one-CL apple eat PRT	*!		
☞ <i>Ta ba yi- ge pingguo chi le.</i> He BA one-CL apple eat PRT		*	

Tableau 8 shows that the first candidate crucially violates the higher ranked constraint *PRE/NSPEC & *Ø_C, which yields the second candidate (the marked scrambled object) as the optimal one. In Tableau 9, the evaluation of possible expressions of the intended meaning 'he ate that apple' is illustrated.

Tableau 9. A scrambled definite object

Input: 'He ate that apple'	*PRE/NSPEC & *Ø _C	*STRUC _C / SCRAM	*PRE/[SPEC/ DEF/NOUN/PN] & *Ø _C
☞ <i>Ta na-ge pingguo chi le.</i> He that-CL apple eat PRT			*
☞ <i>Ta ba na-ge pingguo chi le.</i> He BA that-CLapple eat PRT		*	

We see that in Tableau 9, both candidates violate one of the equally strong constraints *PRE/[SPEC/DEF/NOUN/PN] & *Ø_C and *STRUC_C/SCRAM. Since there is no difference in strength between the violated constraints, both the case marked and an unmarked inanimate scrambled object are optimal expressions of the intended meaning 'he ate that apple'.

I have shown here how the influence of definiteness on Chinese DOM can be accounted for within an OT framework, if the influence of word order is taken into consideration. By applying Harmonic

Alignment to the prominence scales of word order and definiteness, I have constructed the relevant OT constraints to account for the role of definiteness in the Chinese DOM system. The Chinese constraint ranking I presented in this section makes case marking obligatory for non-specific indefinite objects in preverbal position, and optional for all scrambled objects that are more specific or definite. However, I showed in Chapter 2 that for pronoun and proper noun objects in preverbal position, the case marker cannot be omitted. Reconsider the following examples:

(92) *Ta* *(*ba*) *wo* *da* *le.*
 he BA I hit PRT
 'He hit me.'

(93) *Ta* *(*ba*) *Zhangsan* *da* *le.*
 he BA Zhangsan hit PRT
 'He hit Zhangsan.'

In (92) and (93), the case marker cannot be omitted, as both the pronoun and proper noun object refer to human beings. The same holds for definite object NPs that refer to human or animate entities. For instance, the scrambled objects in (94) and (95) are also obligatorily marked with *ba*:

(94) *Ta* *(*ba*) *zhe-tiao* *she* *dasi* *le.*
 He BA this-CL snake hit.dead PRT
 'He killed this snake.'

(95) *Ta* *(*ba*) *na-ge* *laoshi* *piping* *le.*
 He BA that-CL teacher criticise PRT
 'He criticised that teacher.'

The scrambled objects in these sentences are not only highly definite, they are also highly animate, by which case marking is obligatory. The influence of definiteness becomes visible only when the object is inanimate. Recall the following examples:

(96) *Ta* *(*ba*) *yi-ge* *pingguo* *chi* *le.*
 He BA one-CL apple eat PRT
 'He ate an apple.'

(97) *Ta* (*ba*) *zhe-ge* *pingguo* *chi* *le.*
 He BA that-CL apple eat PRT
 'He ate an apple.'

If an inanimate scrambled object is non-specific, the case marker is obligatory; if an inanimate scrambled object is specific or definite, case marking is optional. Both animacy and definiteness constraints thus apply to one and the same object; the two dimensions cannot be treated separately. In the next section, I will show how all the constraints I formulated in this chapter interact and give a full OT account of Chinese DOM.

4.3.3 Chinese DOM: a full account

In this section, I will explain in OT terms how animacy, definiteness, grammatical function and word order together determine Chinese DOM. For convenience, all the constraints on Chinese DOM I formulated in this chapter are listed in (98) – (100):

$$(98) \quad *{\text{STRUCC}}_C/{\text{UNSCR}} >> *{\text{STRUCC}}_C/{\text{SCRAM}}$$

$$(99) \quad *{\text{OBJ}}/[\text{HUM}/\text{ANIM}] \& *{\emptyset}_C >> *{\text{OBJ}}/\text{INAN} \& *{\emptyset}_C$$

$$(100) \quad *{\text{PRE}}/\text{NSPEC} \& *{\emptyset}_C >> *{\text{PRE}}/[\text{SPEC}/\text{DEF}/\text{NOUN}/\text{PN}] \& *{\emptyset}_C$$

We have seen that definiteness features are only of influence on Chinese DOM if the scrambled object is inanimate. This can be formalized as follows: the definiteness constraints in (100) are only decisive if all relevant candidates satisfy $*{\text{OBJ}}/[\text{HUM}/\text{ANIM}] \& *{\emptyset}_C$. For that reason, I will take the lowest element from the constraint hierarchy in (99) and I will locally conjoin this constraint with the elements in the constraint hierarchy in (100). This results in the subhierarchy in (101):

$$(101) \quad \begin{aligned} & *{\text{PRE}}/\text{NSPEC} \& *{\text{OBJ}}/\text{INAN} \& *{\emptyset}_C >> \\ & *{\text{PRE}}/[\text{SPEC}/\text{DEF}/\text{NOUN}/\text{PN}] \& *{\text{OBJ}}/\text{INAN} \& *{\emptyset}_C \end{aligned}$$

This complex hierarchy describes that case marking on non-specific inanimate objects in preverbal position is preferred to case marking on specific, definite, proper noun and pronoun inanimate objects in

preverbal position. Since pronouns and proper nouns are used to indicate human or animate entities, they are considered high animate, as a result of which they are obligatorily case marked. I will therefore remove PN and NOUN from the latter constraint for the sake of clarity:

(102) *PRE/NSPEC & *OBJ /INAN & * \emptyset_C >>
 *PRE/[SPEC/DEF] & *OBJ /INAN & * \emptyset_C

The Chinese constraint ranking should be as follows. Since unscrambled objects are never case marked, *STRUCC/UNSCR is ranked highest in the hierarchy. Human and animate scrambled objects are obligatorily marked with *ba*. Hence, the constraint *OBJ/[HUM/ANIM] & * \emptyset_C is ranked higher than *STRUCC/SCRAM. The first constraint of the hierarchy in (102) also outranks *STRUCC/SCRAM, by which case marking is compelled on scrambled objects that are inanimate and non-specific. The mutual ranking of *OBJ/[HUM/ANIM] & * \emptyset_C and *PRE/NSPEC & *OBJ /INAN & * \emptyset_C is irrelevant. The second constraint of the hierarchy in (102) is ranked equally high with *STRUCC/SCRAM, making case marking optional for inanimate scrambled objects that are specific or definite.

Tableaux 10 to 15 below illustrate that this constraint ranking will yield the correct optimal outputs for every type of scrambled object in Chinese⁷.

⁷ Again, as I only discuss scrambled objects here, I left out *STRUCC/UNSCR for the sake of convenience. Furthermore, for simplicity, I do not discuss the constraints requiring objects to scramble.

Tableau 10. An inanimate, non-specific indefinite scrambled object

Input: 'He ate an apple'	*OBJ/ [HUM/ ANIM] & *Ø _C	*PRE/NSPEC & *OBJ/INAN & *Ø _C	*STRU _C / SCRAM	*PRE/ [SPEC/DEF] & *OBJ/INAN & *Ø _C
<i>Ta yi- ge pingguo</i> He one-CL apple <i>chi le.</i> eat PRT		*!		
☞ <i>Ta ba yi- ge pingguo</i> He BA one-CL apple <i>chi le.</i> eat PRT			*	

In Tableau 10, it is the first candidate that crucially violates the complex constraint *PRE/NSPEC & *OBJ/INAN & *Ø_C, as the unmarked object is non-specific as well as inanimate. The second candidate only violates the lower ranked constraint *STRU_C/SCRAM, by which the case marked scrambled object is the optimal outcome of the evaluation.

Tableau 11. An inanimate, specific indefinite scrambled object

Input: 'he ate a (particular) apple'	*OBJ/ [HUM/ ANIM] & *Ø _C	*PRE/NSPEC & *OBJ/INAN & *Ø _C	*STRU _C / SCRAM	*PRE/ [SPEC/DEF] & *OBJ/INAN & *Ø _C
☞ <i>Ta yi-ge pingguo</i> He one-CL apple <i>chi-wan le.</i> eat-finish PRT				*
☞ <i>Ta ba yi-ge pingguo</i> He BA one-CL apple <i>chi-wan le.</i> eat-finish PRT			*	

Tableau 12. An inanimate, definite scrambled object

Input: 'he ate that apple'	*OBJ/ [HUM/ ANIM] & *Ø _C	*PRE/NSPEC & *OBJ/INAN & *Ø _C	*STRUCC/ SCRAM	*PRE/ [SPEC/DEF] & *OBJ/INAN & *Ø _C
☞ <i>Ta na-ge pingguo</i> he that-CL apple <i>chi le.</i> eat PRT				*
☞ <i>Ta ba na-ge pingguo</i> He BA that-CL apple <i>chi le.</i> eat PRT			*	

Tableaux 11 and 12 show that only for inanimate definite and specific indefinite objects in preverbal position, the case marker can optionally be omitted. Both the case-marked and the unmarked object come out as optimal syntactic representations of the intended meaning.

In Tableaux 13 – 15 below, there is only one candidate that best satisfies the constraints: the remaining types of objects are all obligatorily marked with *ba* when they are scrambled.

Tableau 13. An animate definite scrambled object

Input: 'he killed this snake'	*OBJ/ [HUM/ ANIM] & *Ø _C	*PRE/NSPEC & *OBJ/INAN & *Ø _C	*STRUCC/ SCRAM	*PRE/ [SPEC/DEF] & *OBJ/INAN & *Ø _C
<i>Ta zhe-tiao she</i> He this-CL snake <i>dasi le.</i> hit.dead PRT	*			
☞ <i>Ta ba zhe-tiao she</i> He BA this-CL snake <i>dasi le.</i> hit.dead PRT			*	

Tableau 14. A human definite scrambled object

Input: 'he hit this teacher'	*OBJ/ [HUM/ ANIM] & *Ø _C	*PRE/NSPEC & *OBJ/INAN & *Ø _C	*STRUCC/ SCRAM	*PRE/ [SPEC/DEF] & *OBJ/INAN & *Ø _C
<i>Ta zhe-ge laoshi</i> he this-CL teacher <i>da le.</i> hit PRT	*!			
☞ <i>Ta ba zhe- ge laoshi</i> he BA this-CL teacher <i>da le.</i> hit PRT			*	

Tableau 15. A human pronoun scrambled object

Input: 'he hit me'	*OBJ/ [HUM/ ANIM] & *Ø _C	*PRE/NSPEC & OBJ/INAN & *Ø _C	*STRUCC/ SCRAM	*PRE/ [SPEC/DEF] & *OBJ/INAN & *Ø _C
<i>Ta wo da le.</i> He I hit PRT	*!			
☞ <i>Ta ba wo da le.</i> He BA wo hit PRT			*	

I have shown here how the influence of animacy, definiteness, grammatical function and word order on DOM can be translated into OT constraints, and how these constraints should be ranked in order to account for Chinese DOM.

4.4 Conclusion

In this chapter I gave a formal account of Chinese DOM within an OT Syntactic framework. I introduced Aissen's formalization of DOM, and discussed a shortcoming in her approach that makes her model unsuitable to account for Chinese DOM. By including the dimension of word order in Aissen's model, I demonstrated how an Optimality

Theoretic framework can also be used to explain the DOM system in Chinese.

5

The function of *ba*

In this thesis, I have investigated the pattern of differential object marking in Chinese: I have shown when *ba* is obligatory, when it is optional and when *ba* cannot be used to mark direct objects in Chinese. I have shown that, besides animacy and definiteness, word order also plays an important role in determining Chinese DOM. I have accounted for the differential object marking system in Chinese with an OT Syntactic model, derived from Aissen's (2003) OT account of cross-linguistic DOM.

Direct objects in Chinese are thus differentially marked with *ba*. But what exactly is *ba*? I already mentioned that this is a hot topic among Chinese linguists, and that there is a lot of debate on the status of *ba*. In this chapter, I will examine the function of *ba*. First, I will investigate the function of object marking in languages with two-dimensional DOM. I will compare this function with the function of *ba* in Chinese DOM, and I will investigate how the function of *ba* relates to word order in Section 5.2, to animacy in Section 5.3 and to definiteness

in Section 5.4. In Section 5.5, I will deal with the apparent exceptional interpretation of bare noun objects when they are preceded by *ba*, after which a will give a conclusion.

5.1 Two-dimensional DOM: animacy, definiteness and the function of the object marker

We have seen previously that both animacy and definiteness play a role in Chinese DOM. Contrary to other languages with two-dimensional DOM, animacy and definiteness do not go hand in hand in Chinese. I have shown that the dimensions work in different directions, because of the additional influence of word order on Chinese DOM. However, in those languages where animacy and definiteness do seem to behave similarly in determining the DOM system, there is an important difference between the two dimensions as well. Let us reconsider the sentences from Hindi as presented in Chapter 2 (cf. Mohanan 1990):

(103) a. *Ilaa-ne bacce- ko / *baccara uthaayaa.*
 Ila- ERG child-ACC / child.NOM lift-PAST
 'Ila lifted the/a child.'

b. *Ilaa-ne haar uthaayaa.*
 Ila- ERG necklace. NOM lift-PAST
 'Ila lifted the/a necklace.'

c. *Ilaa-ne* *haar-ko* *uthaayaa*.
 Ila- ERG necklace-ACC lift-PAST
 'Ila lifted the/*a necklace.'

What we see here is that in Hindi, the accusative case marker is obligatory when the object is human in (103a), and it is optional for inanimate objects, as in (103b) and (103c). However, a difference in meaning exists between the sentences in (103b) and (103c). If the case marker is absent, the object can get both a definite and an indefinite reading; if the object is case marked, it can only be interpreted as definite. Note that the sentences in (103b) and (103c) are exactly the same, except for the presence of the case marker. It seems to be the presence of the case marker in (103c) that prevents the object from being interpreted as indefinite. Something similar can be seen in Spanish. Consider the following examples (cf. Rodríguez-Mondoñedo 2006):

(104) a. *María* *quiere* *un* *abogado*.
 Maria wants a lawyer
 'Maria wants a lawyer (any lawyer).'

b. *María* *quiere* *a* *un* *abogado*.
 Maria wants A a lawyer
 'Maria wants a (specific) lawyer.'

When the preposition *a* is expressed, as in (104b), the indefinite human object *un abogado* 'a lawyer' is to be interpreted as a specific lawyer.

Contrastively, when the preposition is not used, the object NP does not refer to a lawyer in particular: Maria wants a lawyer in general. The presence of the preposition gives the object a more specific interpretation. Apparently, the object marker is not used to *indicate* the definiteness of objects: rather, it *marks* an object *as* definite or specific. The case alternation should be regarded as causing the difference in meaning, instead of merely reflecting it. By contrast, a case alternation cannot change the animacy of the object: if the case marker on the human object in (103a) is absent, we do not see a change in interpretation, but a change in grammaticality. This also holds for proper names in Spanish, as can be seen in (105) (Rodríguez-Mondoñedo 2006):

(105) *Juan mató *(a) María.*
 Juan killed A María
 'Juan killed Maria.'

The preposition in this sentence is obligatory: if the proper name *Maria* is not preceded by *a*, this yields an ungrammatical sentence. Whereas the case marker can thus mark an object as definite, it cannot mark an object as animate (de Hoop 2006). Object marking is triggered by the animacy of the object, but object marking is not triggered by its definiteness. Instead, it is the object marker itself that influences the definiteness of the object. In languages with two-dimensional DOM, the case marker is thus used for different reasons. It has both a passive and an active role: an object marker can mark objects that are more

prominent in terms of animacy, but it can make objects more prominent in terms of definiteness.

These different functions of the object marker in languages with DOM can be seen in coherence with the difference between animacy and definiteness in general. Animacy and definiteness can both be considered as contributors to the prominence of a noun phrase, as we saw in the previous chapters. An important difference between the two dimensions, however, is that definiteness is a linguistic category, whereas animacy is not. The definiteness of a noun phrase is determined by linguistic properties, for instance by a determiner. Animacy of a noun phrase, on the other hand, provides information about the individual that the NP refers to, but not about the NP itself (de Hoop 2006). Consider for instance a noun phrase like *this apple*. This noun phrase refers to an inanimate entity, regardless of its linguistic form. But the noun phrase also refers to a specific apple and this interpretation is evoked by the linguistic form of the noun phrase, namely, by the demonstrative *this*. If we now change *this apple* into *an apple*, the noun phrase still refers to an inanimate entity. However, the change of the linguistic form does cause a shift in definiteness: the noun phrase no longer refers to a particular apple. By replacing the demonstrative by an indefinite determiner, the noun phrase gets an indefinite reading.

Definiteness can also be triggered by other factors than lexical markers of the noun phrase, such as determiners. For instance, we saw above that the presence of a morphological marker can cause a more specific or definite interpretation, such as the case marker in Hindi and the preposition *a* in Spanish. Furthermore, we saw in Chapter 3 that a

change in word order can cause a shift in definiteness as well. In Chinese, the preverbal position triggers a specific or definite specific interpretation of the noun phrase that takes this position. Recall the following examples:

(106) *Ta chi le pingguo.*
 he eat PRT apple
 'He ate an apple/apples' /
 'He ate the apple(s).'

(107) *Ta pingguo chi le.*
 he apple eat PRT
 'He ate the apple(s).'

The only difference between the sentence in (106) and the sentence in (107) is the position of the object. Whereas the postverbal object can get both a definite and an indefinite reading, the object can only get a definite interpretation in preverbal position. The function of word order here is comparable to the function of the object marker in Hindi and Spanish. However, it does not influence the animacy of the noun phrase in any way: the noun phrase is inanimate, regardless of its position in the sentence. Whereas animacy of a noun phrase is thus independent of its linguistic representation, the definiteness of a noun phrase is determined by linguistic factors.

We have seen here that animacy and definiteness crucially differ from each other. Therefore, in languages with two-dimensional DOM, the function of the object marker with respect to animacy is different

from its function concerning definiteness. Object marking not only has a semantic *cause*; it has a semantic *effect* as well. In terms of animacy, the presence of an object marker is a *consequence* of the high prominence of an object. At the same time, the presence of the object marker can increase the level of prominence of the object in terms of definiteness: its presence *causes* high prominence of the object.

In languages with two-dimensional DOM, the object marker thus has different semantic functions with regard to animacy and definiteness. However, I have shown in Chapter 3 that object marking in Chinese is first of all syntactically driven: the use of *ba* is licensed by a change in word order. I will therefore investigate how the function of *ba* is related to word order in the next section.

5.2 Chinese DOM: word order and *ba*

In this section, I will investigate the function of *ba* in Chinese. We have seen in Chapter 3 that SVO word order plays a very important role in isolating languages like Chinese: the position of the arguments relative to the verb determines their grammatical function. The argument that follows the verb is interpreted as the object, and the argument in front of the verb as the subject of the sentence. When the word order is changed into SOV, the verb loses this discriminating function. In these cases, *ba* can be used to distinguish subjects from objects.

In Classical Chinese, *ba* was a verb, meaning ‘to hold’ or ‘to take’. An example of how *ba* was formerly used as a verb is given in (108) (cf. Feng 2002):

(108) *Chen zuo shou ba qi xiu.*
 I left hand hold his sleeve
 'I (will) hold his sleeve with my left hand.'

Let us now look at the examples in (109) and (110). The sentence in (109) is ungrammatical, as we already saw in the previous chapters: lexically indefinite objects like *yi-ge pingguo* 'an apple' cannot occur in preverbal position without *ba*. When we consider *ba* a verb which is placed in front of the preverbal object, as in (110), the word order changes into SVO. As a consequence, the indefinite object is in postverbal position and the sentence becomes grammatical again:

	S	O		V
(109)	<i>*Ta yi-ge</i>	<i>pingguo</i>		<i>chi le.</i>
	he	one-CL	apple	eat PRT
	'He ate an apple.'			

	S	V	O		V
(110)	<i>Ta ba yi-ge</i>	<i>pingguo</i>		<i>chi le.</i>	
	he	take	one-CL	apple	eat PRT
	(*'He took an apple and ate it.')				
	'He ate an apple.'				

These sentences give us reason to treat *ba* as a verb in modern Chinese as well. One could argue that the *ba*-sentence in (110) is a serial verb construction, that is, a complex predicate containing at least two (main or independent) verbs in what appears to be a single clause (Veenstra

1996). Compare the use of *ba* in Chinese with the use of *teki* 'take' in the following sentence from Sranan, a Creole language spoken in Suriname (cf. Sebba 1987):

	S	V1	O		V2
(111)	<i>Yu</i>	<i>e</i>	<i>teki</i>	<i>den</i>	<i>krosi</i>
	you	IMP	take	the-PL	clothes
	'You hid the clothes.'				

This sentence is an example of a serial verb construction: both verbs *teki* 'take' and *kibri* 'hide' are within the same clause, and the verbs together express a single event. The same can be said about *ba* 'take' and *chi le* 'ate' in (110). However, one of the criteria for a serial verb construction is that both verbs 'must be lexical verbs, i.e., must be capable of appearing as the only verb in a simple sentence' (Sebba 1987: 39). The following sentences show that both *teki* and *kibri* are lexical verbs indeed, as they can appear as the only verb in a sentence⁸:

(112)	<i>Suma</i>	<i>teki</i>	<i>den</i>	<i>sani</i>	<i>dya?</i>
	who	take	the-PL	thing	here
	'Who took the things from here?'				

(113)	<i>Fred</i>	<i>kibri</i>	<i>en</i>	<i>moni</i>	<i>na</i>	<i>ini</i>	<i>wan</i>	<i>tomati</i>	<i>blik.</i>
	Fred	hide	his	money	LOC	in	a	tomato	can
	'Fred hid his money in a tomato can.'								

⁸ Thanks to Margot van den Berg for these examples.

However, *ba* has lost its semantic contents in modern Chinese: it no longer carries any meaning and it can no longer be used as a lexical verb. This can be seen from the sentences in (114a) and (114b):

(114) a. *Ta na le yi-ge pingguo.*
 he take PRT one-CL apple
 'He took an apple.'

b. **Ta ba le yi-ge pingguo.*
 he BA PRT one-CL apple
 'He took an apple.'

These sentences illustrate that *ba* cannot be used in the same way as regular verbs. The grammaticality contrast between (114a) and (114b) shows that a verb like *na* 'take' followed by the aspect marker *-le* can be used to express the intended meaning 'he took an apple', whereas *ba* cannot. Furthermore, *ba* cannot form a V-not-V question or serve as an answer to a question, like regular verbs in Chinese can (cf. Li and Thompson 1981):

(115) a. *Ni he bu he cha ?*
 you drink not drink tea
 'Will you drink tea?'

b. *He.*
 drink
 'Yes.'

(116) a. **Ni ba bu ba pingguo chi-wan ?*
 you BA not BA apple eat-finish
 'Will you finish the apple?'
 b. **Ba.*
 BA
 'Yes.'

These examples show that *ba* cannot be treated as a lexical verb, and a *ba*-sentence cannot be considered a serial verb construction. Since *ba* is semantically empty, it does not share any semantic properties with other verbs. But even though *ba* has lost its semantic properties, it does behave like a verb on a syntactic level: the presence of *ba* discriminates between the subject and the object. The argument to the left of *ba* is the subject, and the one to the right is the object. In this sense, *ba* functions as a verb in an SVO configuration.

Syntactically, the use of *ba* is thus not only the consequence of a shift in word order; it causes a change in word order as well, that is, from SOV to SVO. Semantically, on the other hand, *ba* seems to have no function at all. However, we saw that semantic features do play a role in the Chinese DOM system. I have shown that *ba* is not obligatory for all scrambled objects, depending on their animacy and definiteness. I will investigate the relation between the function of *ba* and animacy in Section 5.3, and between *ba* and definiteness in Section 5.4.

5.3 Chinese DOM: animacy and *ba*

We saw in the previous chapters that animacy features of the direct object determine whether *ba* is obligatory or optional. Reconsider the following sentences:

(117) *Ta *(ba) laoshi tuidao le.*
 he BA teacher push.over PRT
 'He pushed over the teacher.'

(118) *Ta *(ba) she dasi le.*
 he BA snake hit.dead PRT
 'He killed the snake.'

(119) *Ta (ba) pingguo chi le.*
 he BA apple eat PRT
 'He ate the apple/the apples.'

In (117) and (118), the scrambled objects are human and animate, respectively. In these cases, *ba* cannot be omitted. When the preverbal object is inanimate, as in (119), *ba* is optional. The obligatory presence of *ba* indicates the high prominence or the semantic markedness of the noun phrase in relation to its grammatical function.

Yet, *ba* can only occur in front of preverbal objects, where it takes over the discriminating function of the verb by establishing an SVO word order again. If the preverbal object is inanimate, *ba* is optional.

This can be explained as follows: if subjects and objects can be distinguished on the basis of their semantic properties, it would be redundant to make a syntactic distinction as well. So if animacy features of the arguments provide enough information to determine which argument will be selected as the subject and which argument will be the object, the syntactic function of the verb becomes redundant. Of course, regular verbs do not only provide syntactic information: they contain semantic information as well, which contributes to the meaning of the sentence. However, as shown above, *ba* is semantically empty, so no semantic information gets lost when *ba* is not expressed. Since *ba* only plays a syntactic role, it can be omitted when the preverbal object is inanimate, without affecting the meaning or grammaticality of the sentence in any way.

Whereas the optional omission of *ba* can thus be seen as a consequence of animacy features of the object, the presence of *ba* in itself is not semantically driven; it is a consequence of the syntactic structure of the sentence. In the following section, I will investigate the function of *ba* in relation with definiteness.

5.4 Chinese DOM: definiteness and *ba*

With regard to definiteness and two-dimensional DOM, we saw in Section 5.1 that cross-linguistically, it is in fact not definiteness that influences case marking, but rather the other way around: the presence of an object marker causes high prominence of the object in terms of definiteness. However, in Chinese DOM, the presence of *ba* does *not* give a more prominent reading to preverbal objects:

(120) *Ta * (ba) yi-ge pingguo chi le.*
 he BA one-CL apple eat PRT
 'He ate an apple.'

(121) *Ta * (ba) yixie pingguo chi le.*
 he BA some apple eat PRT
 'He ate some apples.'

Ba is obligatory in both sentences; yet, the preverbal objects are interpreted as non-specific (Yang in prep.). In addition, the omission of *ba* does not trigger an indefinite or non-specific reading. The scrambled object in (120) has a definite reading, whether or not *ba* is present:

(122) *Ta (ba) na-ge pingguo chi le.*
 he BA that-CL apple eat PRT
 'He ate that apple.'

The above sentences show that the presence of *ba* does not directly influence the definiteness of the scrambled object in any way, contrary to object markers in languages with two-dimensional DOM. This can again be explained by the fact that *ba* is used for syntactic reasons only: the use of *ba* is triggered by a change in word order from SVO to SOV, and its presence creates an SVO word order again. We saw in the previous sections that the syntactic function of *ba* becomes redundant when the preverbal object is inanimate, since the semantics already provides the necessary information to discriminate the arguments in

the sentence. If *ba* is omitted, the word order is SOV again. I showed in Chapter 3 that this change in word order has a semantic influence on the object. We saw that an NP in preverbal position must have a specific or definite reading. So if *ba* is omitted, the inanimate object gets a specific or definite interpretation because of its preverbal position. That is, although *ba* does not have any direct semantic effect on the object, the omission of *ba* does give rise to a clear-cut SOV word order, which triggers high prominence of the preverbal object. The object in (122) is marked with the demonstrative *na* 'that', by which it has a definite reading. In this case, *ba* can be omitted without affecting the grammaticality of the sentence. However, NPs that are lexically marked for indefiniteness, like the objects in (120) and (121), cannot get a definite reading and they can therefore not take a preverbal position, as shown in Chapter 3. As a consequence, *ba* is obligatory to induce an SVO word order, by which indefinite objects are syntactically licensed to precede the main verb.

By treating *ba* as a syntactic tool that causes a change in word order, we can thus explain why indefinite objects can only occur in preverbal position when they are preceded by *ba*, whereas for definite objects in preverbal position, *ba* is optional. However, a problem arises when we consider bare noun objects. Recall that for inanimate bare noun objects in preverbal position, *ba* is optional. This is illustrated once more in (123):

(123) *Ta* (*ba*) *pingguo* *chi* *le*.
 he BA apple eat PRT
 'He ate the apple(s).'

Because the object in this sentence is inanimate, the subject can be distinguished from the object by semantic features, and *ba* is optional. The alternation between the presence and the absence of *ba* corresponds to a shift in word order, that is, from SVO to SOV. I showed that bare noun objects can have either a definite or an indefinite reading in postverbal position, but only a definite reading when they are scrambled, because of the requirement of the preverbal position. The examples I gave to illustrate this are repeated below for convenience:

(124) *Ta chi le pingguo.*
 he eat PRT apple
 'He ate an apple/apples' /
 'He ate the apple(s).'

(125) *Ta pingguo chi le.*
 he apple eat PRT
 'He ate the apple(s).'

If the presence of *ba* would indeed cause a change in word order from SOV into SVO, we would expect the bare noun object to get either a definite or an indefinite reading when it is preceded by *ba*. Yet, this is not the case: the bare noun object in (123) can only get a definite reading, whether or not *ba* is present.

How can we explain that a scrambled bare noun object cannot get an indefinite reading when it is preceded by *ba*? I will explain this in the following section.

5.5 Object scrambling and the markedness principle

We saw above that even though bare nouns can get both a definite and an indefinite reading in postverbal position, they can only get a definite interpretation when they are scrambled and preceded by *ba*. This can be explained if we consider the notion of markedness again. There is overwhelming evidence across languages that marked forms are used for marked meanings, whereas unmarked forms are used for unmarked meanings. This general tendency is better known as the markedness principle (Horn 1984). The underlying thought is that both speaker and hearer want to minimize their effort. Unmarked forms are easier for the speaker to produce; unmarked meanings are easier for the hearer to understand. Hence, unmarked forms are preferred to marked forms, and unmarked interpretations are preferred to marked interpretations. If a speaker does use a marked form, this implies that he wants to convey a marked meaning: if he wanted to convey an unmarked meaning, he would have chosen the easier form. This is something that both the speaker and the hearer take into account.

The OT Syntactic account of Chinese DOM I gave in Chapter 4 is an analysis from the speaker's point of view. I have shown how a scrambled object in Chinese is optimally expressed, that is, with or without *ba*. However, we have ignored the fact that there is a less marked way to express a direct object, namely, in a canonical SVO

sentence. An SVO sentence is not only easier to produce for the speaker; it is also easier to interpret for the listener. If the speaker deviates from the canonical word order and produces an SOV sentence, he chooses a marked syntactic structure, by which he implicitly conveys a marked reading: otherwise, he would have used an unmarked SVO order.

The speaker does not always have a choice between an unmarked and a marked form. For instance, we saw in Chapter 2 that an object in Chinese is sometimes scrambled for syntactic reasons: the postverbal constraint only allows for one constituent to follow the main verb, by which the direct object must move to the other side of the verb (e.g., Travis 1984, Sybesma 1992, Po-Ching and Rimmington 2004). But even if the preverbal position may syntactically be the optimal form of expressing a direct object in Chinese, this does not alter the fact that it is a marked form.

In the unmarked, postverbal position, objects thus get an unmarked interpretation. When an object is scrambled, it is in a marked syntactic position, where it gets a marked reading. In Chinese, the interpretation of a bare noun can not be determined until it is placed in a sentence, where features such as context or the nature of the verb determine their reading (Li and Thompson 1981). The unmarked reading for a bare noun object *per se* is neutral: bare noun objects in postverbal position can get either a definite or an indefinite reading. However, bare nouns can only have a definite reading in preverbal position. The marked interpretation for a bare noun object is thus the merely definite reading: the indefinite reading cannot be associated with a marked form. This is perfectly in accordance with the general

link between the preverbal position and high prominence, which I discussed in Chapter 3.

Bare noun objects in Chinese thus obey the principle of markedness: a marked form goes with a marked meaning. This relation between markedness of form and markedness of meaning is captured in Bidirectional Optimality Theory (Blutner *et al.* 2006). In Bidirectional OT, both the speaker's and the hearer's perspective are taken into account. Unlike OT Syntax or OT Semantics, in which either optimal forms or optimal meanings are selected, Bidirectional OT evaluates form-meaning *pairs* (Blutner *et al.* 2006). Such a form-meaning pair is recursively defined as *super-optimal* if and only if there is no other super-optimal form-meaning pair with a different form that expresses the same meaning better, and there is no other super-optimal form-meaning pair with a different meaning that is a better interpretation for the same form. This yields two super-optimal form-meaning pairs which are in accordance with the markedness principle, namely, the unmarked form with the unmarked meaning and the marked form with the marked meaning. Bidirectional OT is ideal to account for pair evaluation when there are two closely related meanings and two closely related forms for one and the same noun phrase in the same linguistic context.

I will now give a bidirectional OT account of bare noun objects in Chinese. The two constraints that are relevant to the relation between form and meaning for bare noun objects in Chinese are given in (126) and (127).

(126) STAY: objects do not scramble.

This syntactic constraint is a reformulation of a constraint introduced by Grimshaw (1997), expressing that words preferably appear in the position given by the basic word order in the sentence. This constraint in fact expresses the general economy principle that movement is costly. The semantic constraint in (127), on the other hand, accounts for the fact that bare noun objects get a neutral interpretation:

(127) MEANING BARE NOUN OBJECT (MBO): a bare noun object is underspecified for definiteness.

Note that the constraints I use here are quite limited and superficial. However, for the sake of simplicity, I will not go any deeper into the underlying motivations; the constraints as formulated here will suffice in accounting for the data I want to explain.

Tableau 16 below shows the two super-optimal form-meaning pairs for bare noun objects that are obtained by using these constraints. These two pairs are indicated by the super-optimality sign '◊'.

Tableau 16. Super-optimal form-meaning pairs for bare noun objects

Bare noun object	MBO	STAY
◊ <postverbal, ±def>		
<postverbal, +def>	*	
<preverbal, ±def>		*
◊ <preverbal, +def>	*	*

We can see from this tableau that the first candidate does not violate either of the relevant constraints: the unmarked, postverbal bare noun with an unmarked, neutral meaning is selected as the first super-optimal form-meaning pair. The latter candidate violates both constraints; yet, it is evaluated as the second super-optimal form-meaning pair. This results from the fact that the other candidates are both blocked by the first super-optimal meaning pair. The pair <postverbal, +def> is blocked because the postverbal position already goes with a less marked, that is, neutral meaning. The pair <preverbal, \pm def > is also blocked, because the neutral meaning is already expressed by a less marked form, that is, the postverbal position. The last form-meaning pair <preverbal, +def> is neither blocked because of its form, nor due to its meaning: it is therefore selected as the second super-optimal pair.

I showed here that using Bidirectional OT makes it possible to explain why bare noun objects can only have a definite reading in preverbal position. The two super-optimal meaning pairs we found perfectly fit the markedness principle: bare noun objects in unmarked, postverbal position are associated with an unmarked, neutral reading, whereas the marked, preverbal position goes with the marked, definite reading.

Lexically indefinite NPs can only have one interpretation: they always get an indefinite reading. If an indefinite object would be placed in a more marked position, it could not get a more marked meaning, as no other meaning is available for indefinite NPs in Chinese. Similarly, lexically definite NPs can also get just one reading: both in SVO and in SOV order, a lexically definite NP can only have a definite

interpretation. As there is no meaning difference, it is unnecessary to use a Bidirectional OT analysis to account for scrambling behaviour of lexically indefinite and definite NPs.

When *ba* is placed in front of a preverbal object, it could be argued that the unmarked SVO structure is recovered. However, the syntactic structure of the whole sentence is not SVO, but SVOV. Even when *ba* is present, the object is still in front of the main verb, which is a marked syntactic position. Consequently, this marked position goes with a marked meaning, that is, only a definite interpretation is possible for a scrambled bare noun object. This definite reading makes that the object obeys the requirements of the preverbal position, and for this reason, *ba* is optional for scrambled bare noun objects. Since *ba* does not have any semantic function like regular verbs, it cannot cause a shift in interpretation and is therefore not necessarily expressed.

If a lexically definite object NP is scrambled, it obeys the high prominence requirement of the preverbal position, for it can only have a definite reading. This is why *ba* is optional for scrambled definite objects as well. An indefinite object, on the other hand, cannot be placed preverbally: this yields an ungrammatical sentence. The indefinite NP in this position can only be licensed by adding *ba* to the sentence and thereby obtaining an SVO structure (albeit a marked one). For indefinite NPs in preverbal position, *ba* is therefore obligatory: the presence of *ba* makes it possible for an indefinite NP to keep this indefinite reading, but it does again not cause a shift in interpretation.

5.6 Conclusion

I have shown in this chapter that an object marker can have various functions: I showed that in languages with two-dimensional DOM, object marking is a consequence of high animacy, whereas it causes high definiteness of the direct object. I illustrated that this is not exactly the same for *ba* in Chinese: the use of *ba* is first of all syntactically driven, and its function is comparable to the distinguishing function of the verb in an SVO construction. The presence of *ba* is a consequence of a shift in word order, and it causes a change in word order at the same time. However, since *ba* is semantically empty, its presence does not influence the interpretation of the object in any way.

The possible omission of *ba*, on the other hand, is a consequence of the semantic features of the object. First, if subjects can be distinguished from objects by animacy, a syntactic distinction becomes redundant and *ba* can be omitted. Next, *ba* is optional if the preverbal object is lexically definite, as it obeys the definiteness requirement of the preverbal position. Lexically indefinite objects, however, cannot occur in preverbal position. As a consequence, *ba* is obligatory to induce an SVO word order, by which non-specific indefinite NPs are licensed to occur in front of the main verb. I showed that a bidirectional OT approach can account for the fact that bare noun objects in a marked (preverbal) position get a marked (definite) reading, by which *ba* is optional for scrambled bare noun objects as well.

6

Conclusion

To *ba* or not to *ba*: that is the question I have answered in this thesis. I showed that Chinese has a system of Differential Object Marking, and I explained why some direct objects in Chinese are preceded by *ba*, while others are not. The differential use of *ba* depends on both syntactic and semantic factors, that is, word order on the one hand, and animacy and definiteness on the other. Chinese DOM conflicts with Aissen's (2003) prediction that cross-linguistically, direct objects with high prominence (i.e., high animacy and definiteness) are more susceptible to overt case marking than objects with low prominence. I showed that the Chinese DOM pattern only partly follows this prediction. High animate objects in Chinese are indeed obligatorily marked with *ba*, whereas for low animate objects, *ba* is optional. Contrastively, however, *ba* can be omitted for high definite objects, while for low definite objects *ba* is obligatory.

I have accounted for the difference between Chinese and cross-linguistic DOM by taking the dimension of word order into account.

While cross-linguistically DOM systems are determined by semantic features only, object marking in Chinese is first of all syntactically driven. Direct objects in their canonical, postverbal position can never be marked with *ba*, as word order already provides enough information to distinguish between subjects and objects. When both the subject and the object precede the verb, word order can no longer be used as a cue to discriminate the grammatical roles of the arguments. As a result of this, objects are preceded by *ba* when they are scrambled. Word order also affects the definiteness of arguments: the preverbal position triggers a specific or definite interpretation. Word order can be considered a dimension of prominence: the preverbal position is associated with high prominence, whereas the postverbal position is associated with low prominence.

I argued that Chinese DOM is based on two instances of markedness reversal. First, if scrambled objects are inanimate, they have unmarked properties with respect to their grammatical function, and *ba* is optional. Animate and human scrambled objects, on the other hand, have marked properties for their grammatical function, by which *ba* becomes obligatory. Second, if scrambled objects have a specific or definite reading, they have unmarked properties with regard to the preverbal position. As a consequence, *ba* can be omitted. If scrambled objects are non-specific, however, they have marked properties for this position in the sentence. Therefore, non-specific indefinite objects in preverbal position are obligatorily marked with *ba*. On the basis of these two instances of markedness reversal, I gave a formal account of Chinese DOM within an OT Syntactic framework. I introduced Aissen's (2003) OT formalization of DOM, and by including the

dimension of word order in her model, I demonstrated how an OT Syntactic model can perfectly be used to explain the DOM system in Chinese as well.

In languages with two-dimensional DOM, object marking is a consequence of high animacy, whereas it causes high definiteness of the direct object. I showed that this is not the case for *ba*: the use of *ba* is first of all syntactically driven, and its function corresponds to the discriminating function of the verb in an SVO configuration. The use of *ba* is a consequence of a shift in word order from SVO to SOV, and its presence provides an SVO word order at the same time. As *ba* is semantically empty, its presence does not influence the meaning of the sentence in any way. The possible omission of *ba* is a consequence of the semantic features of the object. If objects can be distinguished from subjects by their animacy, a syntactic distinction becomes redundant and *ba* can be omitted. Furthermore, *ba* is optional when preverbal objects are lexically definite, as they meet the definiteness requirement of the preverbal position. Bare noun objects are lexically underspecified for definiteness, but if they are in a marked, preverbal position, they get a marked, definite reading. I have shown that this relation between markedness of form and markedness of meaning can be accounted for with a Bidirectional OT approach (Blutner *et al.* 2006). Scrambled bare nouns thus fulfil the definiteness requirement of the preverbal position as well, by which *ba* is optional for preverbal bare noun objects. Only if objects are lexically indefinite, they cannot occur in preverbal position. In order for lexically indefinite objects to scramble, *ba* is obligatorily used to induce an SVO word order, albeit a marked one, which licences indefinite objects to precede the main verb.

References

AISSEN, J. (2003). 'Differential Object Marking: Iconicity vs. Economy'. *Natural Language and Linguistic Theory* 21, 435-483.

ASHER, R.E., and T.C. KUMARI (1997). *Malayalam*. London/New York: Routledge.

BATTISTELLA, E. (1996). *The logic of markedness*. Oxford: Oxford University Press.

BEAVER, D., I. FRANCEZ, and D. LEVINSON (to appear). 'Bad Subject: (Non-)Canonicality and NP Distribution in Existentials'. To appear in: E. Georgala and J. Howell (eds.), *Proceedings of Semantics and Linguistic Theory XV*. Ithaca, New York: CLC Publications.

BLUTNER, R., H. DE HOOP, and P. HENDRIKS (2006). *Optimal Communication*. Stanford: CSLI.

BOSSONG, G. (1985). *Diffentielle Objektmarkierung in den neuiranischen Sorachen*. Tübingen: Gunter Narr Verlag.

BOWE, H. (1990). *Categories, constituents and constituent order in Pitjantjatjara*. London: Routledge.

CHENG, L. (2006). 'Dou as an external contextual domain restrictor'. Presentation at: *Semantics Colloquium*, May 16. Nijmegen: Nijmegen Centre for Semantics.

COMRIE, B. (1989). *Language universals and linguistic typology*. Chicago: University of Chicago Press.

COMRIE, B. (1979). 'Definite and animate direct objects: A natural class'. *Linguistica silesiana* 3: 13-21.

CROFT, W. (1990). *Typology and universals*. Cambridge: Cambridge University Press.

DAHL, Ö. (2000). 'Egophoricity in discourse and syntax'. *Functions of language* 7, 37-77.

FENG, S. (2002). 'Hanyu dongbu jiegou laiyuan de jufa fenxi [A formal analysis on the origin of verb-resultative constructions in Chinese]'. *Yuyanxue Luncong* 26, 178-208.

GOODALL, G. (1987). 'On the Argument Structures and L-marking with Mandarin Chinese ba'. In: J. McDonough and B. Plunkett (eds.), *Proceedings of NELS 17*. Amherst: CSLA.

GREENBERG, J. H. (1963). 'Some universals of grammar with particular reference to the order of meaningful elements'. In: J.H. Greenberg (ed.), *Universals of Language*, 73-113. Cambridge: MIT Press.

GRIMSHAW, J. (1997). 'Projection, heads and optimality'. *Linguistic Inquiry* 21, 187-222.

HASHIMOTO, A. (1971). 'Mandarin Syntactic Structures'. *Unicorn* 8, 1-149.

HAWKINS, J.A. (1983). *Word Order Universals*. New York: Academic Press.

HENDRIKS, P. and H. DE HOOP (2001). 'Optimality Theoretic Semantics'. *Linguistics and Philosophy* 24: 1-32.

DE HOOP, H., 2006. 'On (in)animate noun phrases'. Manuscript, Radboud University Nijmegen.

HORN L.R. (1984). 'Toward a new taxonomy for pragmatic inference: Q-based and R-based implicatures'. In D. Schiffrin (ed.), *Meaning, Form and Use in Context*, 11-42. Washington: Georgetown University Press.

HUANG, C.T.J. (1990). *Complex predicates in generalized control*. Irvine: Ms. University of California.

HUANG, C.T.J. (1982). *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. dissertation, MIT, Cambridge.

JACOBSEN, W.M. (1992). *The transitive structure of events in Japanese*. Tokyo.

JAKOBSON, R. (1971). *Word and language. Selected writings, Volume II*. The Hague: Mouton.

KARLSSON, F. (1983). *Finnish grammar*. Juva: WSOY.

KEENAN, E. (1976). 'Towards a universal definition of 'subject''. In C. Li, (ed.), *Subject and topic*, 303-333. New York: Academic Press.

LEE, T. H. (1986). *Studies on Quantification in Chinese*. PhD Dissertation, UCLA.

LI, Y.H.A. (2001). 'The ba-construction.' Manuscript. Available at:
<http://www.usc.edu/dept/LAS/ealc/chinling/audreyli.htm>.

LI, C., and S. THOMPSON (1981). *Mandarin Chinese: a functional reference grammar*. Berkeley: University of California Press.

LIU, F. (1997). 'An Aspectual Analysis of *ba*'. *Journal of East Asian Linguistics* 6, 51-99.

MOHANAN, T. (1990). *Arguments in Hindi*. Stanford Dissertation.

ØVRELID, L. (2004). 'Disambiguation of syntactic functions in Norwegian: modelling variation in word order interpretations conditioned by animacy and definiteness'. In: F. Karlsson (ed.), *Proceedings of the 20th Scandinavian Conference of Linguistics*, Helsinki.

PO-CHING, Y. and D. RIMMINGTON (2004). *Chinese: a comprehensive grammar*. London: Routledge.

PRINCE, A., and P. SMOLENSKY (1993/2004). 'Optimality Theory: Constraint interaction in generative grammar'. Rutgers University Center for Cognitive Science Technical Report 2. Revised version published by Blackwell, 2004.

RODRÍGUEZ-MONDOÑEDO, M. (2006). 'The Acquisition of Differential Object Marking in Spanish'. Manuscript, University of Connecticut.

SEBBA, M. (1987). *The syntax of serial verbs*. Creole Language Library 2. Amsterdam: John Benjamins.

SIEWIERSKA, A. and D. BAKKER (to appear). 'Case and alternative strategies; word order and agreement marking'. To appear in: A. Malchukov (Ed.) *Handbook of Case*. Oxford: Oxford University Press.

SILVERSTEIN, M. (1976). 'Hierarchy of features and ergativity'. In: R. M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112-171. Canberra: Australian Institute of Aboriginal Studies.

SINNEMÄKI, K. (2006). 'Languages with SOV word order and no morphological marking of core arguments'. Presentation at: *RaRa-Rarissima: Collecting and interpreting unusual characteristics of*

human languages, March 29–April 1. Leipzig: Max Planck Gesellschaft.

DE SWART, P. (2003). *The Case Mirror*. MA Thesis, Radboud University Nijmegen.

SYBESMA, R. (1992). *Causatives and Accomplishments: The case of Chinese ba*. PhD. dissertation, Leiden University.

TRAVIS, L. (1984). *Parameters and Effects of Word Order Variation*. Ph.D. dissertation, MIT, Cambridge.

VEENSTRA, T. (1996). *Serial verbs in Saramaccan. Predication and Creole Genesis*. PhD. dissertation, University of Amsterdam.

YANG, N. (in prep.). *Indefinite objects in Mandarin Chinese*. PhD Dissertation, Radboud University Nijmegen.

YANG, N. AND G. VAN BERGEN (to appear). 'Scrambled Objects and Case Marking in Mandarin Chinese'. To appear in: *Lingua*.