

**The structure of nouns in Old Xining and Modern Standard Chinese**

Qi Wang

Submitted to the School of English Literature, Language and Linguistics

for the degree of

Doctor of Philosophy (Integrated)

at

Newcastle University

October 2018



## Abstract

The goal of this thesis is to articulate a theory which can account for the structure of content words in Modern Standard Chinese and Old Xining Chinese, the traditional dialect of Xining in North West China. Various types of nominal compounds in Modern Standard Chinese and Old Xining Chinese will be described and analyzed, including endocentric attributive compounds and several types of coordinative compounds. Word formation by derivational affixation in the two languages will also be described and discussed. Reduplicated nouns in Old Xining Chinese will be described and analyzed, and shown to have particularly interesting consequences for morphological and syntactic theory.

The initial question concerns the number of components in the structure of a word in Modern Standard Chinese. Following much recent work in generative syntax and morphology, the Root is defined as a primitive without word category, which can therefore never form a word on its own. A Root can form a Minimal Word only together with a categorizer, which is typically a phonologically null morpheme. A condition is proposed that there are always at least two components in the structure of a free content word in Modern Standard Chinese and Old Xining Chinese. This implies a new analysis of the items, typical of Chinese, often called Bound Roots in the literature. They are not Bound Roots (all Roots are bound), but Bound Words, a primitive with inherent syntactic category, and therefore unable to occur with a categorizer. To meet the two-component condition they have to be merged with another item, such as another word or derivational affix.

It is characteristic of Old Xining Chinese that simple nouns are always reduplicated. Therefore, a stricter condition is proposed for Old Xining Chinese nouns than for Modern Standard Chinese nouns: they must always have at least two *pronounced* components. It is argued that the reduplicated simple nouns in Old Xining Chinese have a structure where the base is the acategorial Root and the reduplicant is the null Nominalizer which copies the phonological content of the base Root. Old Xining Chinese thereby provides strong empirical evidence that words, particularly nouns, have minimally two components, a Root and a Nominalizer. It also provides evidence that a nominal

attributive compound is made up of a Root and a noun, not by two nouns. This hypothesis can explain, among others, the following facts: First, Bound Words cannot be reduplicated, which follows if they have inherent category. Second, the base of an affixed noun cannot be reduplicated if the affix is the head, but can be if the affix is not the head. Third, the head of a nominal compound can be reduplicated, the modifier cannot. The absence of recursive nominal compounds in Old Xining Chinese will also be discussed.

## **Acknowledgment**

I would like to express my sincere gratitude to my supervisor Professor Anders Holmberg who had faith in me, encouraged me and inspired me so much. I am truly grateful for his clear guidance and great patience. I am also very thankful to Professor Maggie Tallerman, my other supervisor, who has provided me with valuable guidance and advice.

My sincere appreciation also goes to Professor Redouane Djamouri (CNRS) and Dr William Van Der Wurff (Newcastle University) for their valuable comments and suggestions on the improvement of the thesis.

I am also very grateful for the support I received from my family, who is always there for me.



## Table of contents

<b>Abstract</b> .....	iii
<b>Acknowledgment</b> .....	v
<b>Table of contents</b> .....	vii
<b>List of Tables</b> .....	x
<b>Abbreviations</b> .....	xi
<b>Chapter 1 Introduction</b> .....	1
1.1 Linguistic background of Modern Standard Chinese and Old Xining Chinese.....	4
1.2 Word-formation in Modern Standard Chinese and Old Xining Chinese.....	6
1.3 Contribution of the thesis to Chinese morphology and morphology in general.....	11
1.4 Organization of the thesis.....	12
<b>Chapter 2 Affixed words and compounds in Modern Standard Chinese and Old Xining Chinese: a description</b> .....	15
2.1 Derivational Affixation in MSC and OXC.....	15
2.1.1 Introduction.....	15
2.1.2 MSC prefix.....	15
2.1.3 MSC suffixes.....	16
2.1.4 OXC prefixes.....	19
2.1.5 OXC suffixes.....	20
2.2 Compounds in MSC and OXC.....	24
2.2.1 Definition and classification.....	24
2.2.2 Attributive phrases and compounds.....	26
2.2.3 Coordinate compounds.....	31
2.2.4 Remaining compounds.....	35
2.2.4.1 Verb-object compounds.....	35
2.2.4.2 Subject-predicate compounds.....	35
2.2.5 Recursive compounds in MSC.....	36
2.3 Summary.....	43
<b>Chapter 3 Theoretical background: previous analyses of word structure</b> .....	45
3.1 Basic concepts: morpheme, Root and stem.....	45
3.2 Selkirk (1982).....	46
3.3 Bound items in MSC.....	48

3.3.1 Dai (1992).....	49
3.3.2 Sproat and Shih (1996).....	50
3.4 Packard (2000): a detailed theory of MSC word structure.....	51
3.5 Zhang (2007).....	62
3.6 Summary.....	64
<b>Chapter 4 Modern standard Chinese word formation: proposed theory.....</b>	<b>67</b>
4.1 Primitives of the structure of MSC words.....	67
4.1.1 Harris (1991): Spanish.....	69
4.1.2 Holmberg (1992): Swedish.....	71
4.1.3 Modern Standard Chinese.....	72
4.2 Different types of heads.....	83
4.3 Minimal Words in MSC: structure and head location.....	88
4.4 Affixed words in MSC: structures and head locations.....	90
4.4.1 Suffixation.....	91
4.4.2 Prefixation.....	101
4.5 Compounding in MSC: structures and head locations.....	104
4.5.1 Attributive compound nouns.....	104
4.5.2 Coordinate compounds.....	112
4.5.2.1 Synonymy-compound nouns.....	113
4.5.2.2 Parallel compound nouns.....	117
4.5.2.3 Synonymy-compound adjectives.....	122
4.5.2.4 Parallel compound adjectives.....	123
4.5.3 Remaining compounds: Antonymy-compounds, verb-object compounds and subject-predicate compounds.....	125
4.5.4 Structures of MSC recursive compounds: an unexpected constraint.....	129
4.6 Summary.....	145
<b>Chapter 5 The structure of Old Xining Chinese words.....</b>	<b>149</b>
5.1 Primitives in OXC common nouns.....	149
5.2 Primitives in OXC adjectives and verbs.....	160
5.3 Reduplicated Nouns, Minimal Adjectives and Minimal Verbs in OXC: structure and head locations.....	161
5.4 Simple affixed words in OXC: structure and head location.....	163
5.4.1 Suffixation.....	164



5.4.2 Prefixation.....	178
5.5 Compounding in OXC: the structure and head locations.....	180
5.5.1 Simple attributive compound nouns in OXC.....	180
5.5.2 On reduplication as a phonologically motivated operation.....	189
5.5.3 Coordinate compounds.....	190
5.5.3.1 Synonymy-compound nouns.....	190
5.5.3.2 Parallel compound nouns.....	196
5.5.3.3 Synonymy-compound adjectives and parallel compound adjectives.....	203
5.5.3.3.1 Synonymy-compound adjectives in OXC.....	204
5.5.3.3.2 Parallel compound adjectives in OXC.....	205
5.5.4 Remaining compounds: Antonymy-compounds, verb-object compounds and subject -predicate compounds.....	207
5.5.5 Absence of recursive nominal compounds in OXC: some explanations.....	210
5.6 Reduplicated nouns in OXC: a summary.....	216
5.7 Summary.....	218
<b>Chapter 6 Conclusions.....</b>	<b>221</b>
<b>References.....</b>	<b>227</b>

**List of Tables**

Table 1.....82  
Table 2..... 159  
Table 3.....160  
Table 4.....162

## Abbreviations

Abbreviations which are applied in this thesis are given below:

AFF	affix
BW	Bound Word
LINK	linking morpheme
MN	Minimal Noun
MSC	Modern Standard Chinese
MW	Minimal Word
OXC	Old Xining Chinese
R	Root
RED	Reduplication
RW	recursive words
WM	word marker



## Chapter 1 Introduction

The relationship between morphology and syntax has been at the center of the debate within theoretical linguistics at least since the 1980s. Works include Lieber (1980), Williams (1981), Selkirk (1982), Di Sciullo and Williams (1987), Baker (1988) Marantz (1997), Embick and Noyer (2007) and Harley (2011). It has been argued in the literature that morphology is closely related to syntax. More specifically, Selkirk (1982) claims that formal properties of X-bar theory of syntax can be extended to the structure of a word. That is to say the category names and category levels and the rewrite rules (i.e.  $X^n \rightarrow \dots X^{n-1} \dots$ ) of X-bar theory of syntax also account for word structure. Modern Standard Chinese word structure, too, has been claimed to be analyzable in terms of X-bar theory (Packard 2000).

The notion of head is also applicable to the word. The Righthand Head Rule (RHR) has been proposed for the word, according to which the head is the rightmost constituent (Williams 1981:248). The linear order is the key to determine the head here. However, the right-hand constituent may not always be the head and scholars including Lieber (1980) and Selkirk (1982) all consider the possibility that the left-hand constituent may as well be the head. Packard (2000:39) follows this line of argument and proposes a Headedness Principle for Modern Standard Chinese simple compound nouns and verbs, which states that the right-hand constituent is the head in a nominal compound and the left-hand constituent is the head in a compound verb. The location of the head here is clearly identified based on its position and the linear order is the determining factor and also according to Packard (2000) the head is the item which provides the categorial information for the compound as a whole. In the present work, the determining factors for locating the head and the notion of the head will be reviewed and it will be argued that instead the head is identified based on feature percolation (Holmberg 1992; Lieber 1980; Scalise and Fábregas 2010) and the head provides categorial and semantic features for the word as a whole (Di Sciullo and Williams 1987; Liao 2014; Scalise and Fábregas 2010) in both Modern Standard Chinese and Old Xining Chinese.

Most scholars working on Chinese morphology would probably agree that a Modern Standard Chinese word typically consists of only one constituent, a root word, as Packard (2000:165) calls it, for instance, *shou* ‘hand’, *mei* ‘beautiful’.<sup>1</sup> However, Modern Standard Chinese also has a type of word component which is uncommon or even absent from many other languages, that is, Bound Roots (to use Packard’s (2000) term) such as *mao* ‘looks’, *shi* ‘teacher’, *qi* ‘wife’ *zhuo* ‘table’, *liang* ‘good’, *yi* ‘sweet’, *xiao* ‘disappear’. Bound Roots, as the name suggests, cannot stand alone as independent words (Packard 2000:166). Rather they can only form independent words with other Roots, words or derivational affixes (Packard 2000:166-168), such as, *liang shi* ‘good teacher’ where both *liang* ‘good’ and *shi* ‘teacher’ are Bound Roots and *zhuo tui* ‘table leg: leg of table’ where *zhuo* ‘table’ is a Bound Root and *tui* ‘leg’ is a Root Word.<sup>2</sup> The example *liang shi* ‘good teacher’ shows that a Bound Root, in this case *shi* ‘teacher’, can be the head of a word, and the example *zhuo tui* ‘table leg: leg of table’ shows that the Bound Root *zhuo* ‘table’ can be a modifier. The nature of these bound items is a source of controversy in Chinese morphological theory. So according to Packard (2000:165-166, 168), a content word of Modern Standard Chinese can contain either one, two, or more constituents. The question whether this is the case will be discussed in the present work. One of the claims here is that no content word in Modern Standard Chinese or Old Xining Chinese consists of less than two constituents, an acategorial Root and a categorizer morpheme.

This thesis focuses on the formation of content words, particularly various types of nominal compounds and words generated through derivational affixes in Modern Standard Chinese and Old Xining Chinese. In particular, I examine attributive compound nouns in which one component modifies the other component (Ceccagno and Basciano 2007, 2011), for instance, *xian hua* ‘fresh flower: fresh flower’, and several kinds of coordinated compound nouns, including what will be called synonymy-compound nouns, made up of two synonyms (Chao 1968), such as *shu cai* ‘vegetable vegetable: vegetable’, and what will be called parallel compound nouns (Chao 1968), made up of two components with associated meaning, but where one provides the semantic content

---

<sup>1</sup> Throughout the thesis, Chinese *pinyin* is used to transcribe Modern Standard Chinese data and Old Xining Chinese data.

<sup>2</sup> In terms of notation, the gloss is followed by the translation of the word. The gloss and the translation are separated by a semicolon. This kind of notation will be used throughout the thesis.

for the compound as a whole, for instance, *chuang hu* ‘window house: window’.

Reduplication of common nouns in Old Xining Chinese is also described and discussed, for example, *mo mo* ‘steamed bun’.

The following research questions are addressed in the thesis. (i) What primitives are needed to account for the structure of content words in Modern Standard Chinese as well as Old Xining Chinese? (ii) What is the nature of Roots in word formation? (iii) How many components are there minimally in a free content word of Modern Standard Chinese and Old Xining Chinese? (iv) What is the nature of reduplication of nouns in Old Xining Chinese? (v) What are the rules needed to derive content words in Modern Standard Chinese and Old Xining Chinese?

In the present work, the definition of Root, familiar from Distributed Morphology, will be adopted which says that a Root has no categorial feature, but merges with a categorizer to form a word (Embick and Noyer 2007,2008; Harley and Noyer 1999; Marantz 1997, 2001). Also, I will not assume X-bar theory for word structure of Modern Standard Chinese and Old Xining Chinese. Instead I will follow the notion that words are formed by the operation Merge, constructing words from the bottom up (Chomsky 1995:243-246, 2015:223-227; Embick and Noyer 2007; Harley and Noyer 1999; Marantz 2001) in both languages. It will be proposed that there is a condition for both Modern Standard Chinese and Old Xining Chinese, that a free content word must consist of at least two constituents, which may also be a universal condition. All this will necessitate a reanalysis of the so-called Bound Root (Packard 2000). It will be proposed that they are content words which have an inherent category, which means that they cannot merge with a categorizer. The reason why they are bound is the condition that a free content word must contain at least two constituents in both languages.

A stronger condition applies to nouns of Old Xining Chinese, which says that a noun must consist of minimally two pronounced constituents. This explains the obligatory reduplication of nouns in Old Xining Chinese, for instance, *fu fu* ‘book’, *fei fei* ‘water’, *mo mo* ‘cat’. The reduplicated nouns in Old Xining Chinese are argued to have a structure where there is a category-less Root and an overt Nominalizer. This structure of reduplicated nouns in Old Xining Chinese supports the proposal that a Root and a noun

may form an attributive nominal compound and two nouns cannot do so. Further, the fact that Bound Words in Old Xining Chinese cannot be reduplicated can also be explained. Furthermore, the fact that the host of a derivational affix which is a head in Old Xining Chinese cannot undergo reduplication and the host of a derivational affix which is a non-head can do so can also be understood. Moreover, this proposal can also offer explanations for the fact that the head of a nominal compound in Old Xining Chinese can be in a reduplicated form, but not the modifier.

### **1.1 Linguistic background of Modern Standard Chinese and Old Xining Chinese**

The varieties of Xining dialect used in the Xining area are Old Xining Chinese and New Xining Chinese (Dede 2006). There are two varieties of Old Xining Chinese (Dede 2003:335). One is spoken by the Han speakers, the other is spoken by the Hui (Muslim) speakers (Dede 2003:335). The Chinese languages discussed in this thesis are Old Xining Chinese which is spoken by the Han speakers and Modern Standard Chinese. From now on, I will use the abbreviation MSC for Modern Standard Chinese and OXC for Old Xining Chinese which is spoken by the Han speakers. Now MSC and OXC are introduced.

MSC (*Putonghua*) was promoted in the whole country of China by the Chinese central government in the early 1950s (Dede 2003:337). Today, it is the official language of mainland China, which is spoken by the Han people alongside their local varieties of Chinese (Po-Ching and Rimmington 2016:3; Sun 2006:10). The pronunciation and grammar of MSC is based on the Northern Mandarin used in the Beijing area (Huang and Shi 2016:2; Po-Ching and Rimmington 2016:3). Northern Mandarin is a subgroup of Mandarin, one of main groups of Chinese used by the Han people in China (Huang and Shi 2016:4). The other groups are Wu, Xiang, Yue, Min, Hakka and Gan (Huang and Shi 2016:4).<sup>3</sup>

---

<sup>3</sup> It is quite common in the literature to use the term Mandarin for MSC, but in this thesis I will not do so in order to avoid any confusion between the language and the dialect group.



OXC is a term adapted from Dede's (2006) term 'old Xining dialect', which is classified as a subgroup of Central Plains Mandarin (Dede 2003:330-331; Wang & Dede 2016:407; Zhang 2001:3). Although OXC and MSC both belong to the Mandarin group, they are mutually incomprehensible. OXC refers to the variety of Chinese traditionally spoken in and around Xining. More specifically, it is the mother tongue of the Han people, particularly older generations, who reside in the city of Xining and its surrounding villages Huangyuan, Huangzhong, Pingan, Menyuan, Huzhu, Guide and Hualong (Zhang & Zhu 1987:1-4; Zhang 2001:1-3) and it is also spoken as a second language by older speakers of the other ethnic groups in the area.

Xining is today the provincial capital of Qinghai (Zhang & Zhu 1987:1). The city has a known history of more than 2,000 years (Zhang & Zhu 1987:1). This is a history with great language diversity. During this long period of time, the Qinghai Region including the Xining city area has been occupied by people with different languages and cultural backgrounds (Zhang & Zhu 1987:1-4; Zhang 2001:1-3). Tribes of Qiang, Xianbei, Tuyuhun, Tubo, Mongols, Monguors, Salars and Hui have all been residents of this area and apart from the Hui people, they all spoke their own languages, different from ancient Chinese spoken by the Han people (Dede 1999:72; Dede 2003:328; Zhang & Zhu 1987:2; Zhang 2001:1-3). There are also periods of time in history during which Han people migrated to Qinghai, including the Xining area. Those periods are the Han Dynasty (206BC-220AD), the Sui Dynasty (581-618), the Tang Dynasty (618-907), the Northern Song Dynasty (960-1127), the Ming Dynasty (1368-1644) and post-1949 (Zhang & Zhu 1987:1-4; Zhang 2001:1-3).

Dede (1999:71-72, 2003:328) has argued that the Chinese language which is used in the Qinghai region at the present day descends from the Chinese used by the population of the area since Ming Dynasty (1368-1644). Firstly, historical accounts including family history have shown that the ancestors of local Qinghai people were from the area of lower Yangtze during Ming Dynasty (Dede 1999:71-72; Dede 2003:327). Also since Ming, the administration of the region of Qinghai has become relatively stable and the ethnic/linguistic situation in Qinghai has not changed to this present day (Dede 1999:71-72; Dede 2003:327-328). More specifically, the Han, Tibetans, Mongols, Monguors, Salars and Hui are still the largest ethnic groups in Qinghai (Dede 1999:71-72, 2003:328).

OXC has been affected and shaped by contact with the non-Sinitic languages used by those largest ethnic groups over the years (Dede 2006:320-321, 327; cf. Bell 2016). Syntactically speaking, SOV order is prominent in OXC (Dede 2006:327; Bell 2016:50), and the language has a number of other features which distinguishes it clearly from MSC, for example, a case suffix on the preverbal object e.g. *-ha*, is often required (Dede 1993:72). Also complementizers follow subordinate clauses in OXC (Dede 1999:51; cf. Bell 2016:4, 170ff). It is still debatable whether OXC is underlyingly an SOV language. SVO word order still occurs not infrequently, and the language has certain other characteristics that are typical of SVO languages. For instance noun incorporation, where it occurs, yields VO order. As for morphology, a striking property, which will be important in this thesis, is that common nouns are reduplicated compulsorily in OXC.<sup>4</sup>

In addition, since the early 1950s, with the promotion of Modern Standard Chinese in the Xining city area, another variety of Xining dialect has emerged, which is called ‘New Xining dialect’ by Dede (2006), and is the result of the interaction between Modern Standard Chinese and Old Xining Chinese (Dede 2006). I will call it New Xining Chinese here in contrast to Old Xining Chinese. Dede (2006:326-327) has pointed out that New Xining Chinese morphology behaves quite similarly to its counterpart in Modern Standard Chinese. More specifically, while in OXC reduplication of simple nouns is obligatory, it is not obligatory (although still common) in New Xining Chinese (Dede 2006:325-326).

The data used in this thesis is mainly from the author, who is a native speaker of the language, familiar with both OXC and MSC. The data from OXC have also frequently been checked with older speakers.

## **1.2 Word-formation in Modern Standard Chinese and Old Xining Chinese**

MSC is known to be a highly analytic type of language in which words are rarely inflected for the purpose of grammatical functions, such as person and voice (Lin 2001;

---

<sup>4</sup> Note that the reduplication is a widespread phenomenon among North Western Varieties of Mandarin Chinese.

Packard 2000). Instead, separate words (particles) are employed to denote grammatical functions (Lin 2001; Packard 2000). Compounding is a common way of forming new words (Packard 2000). Simple compounds which are made up of two items are very common, and recursive compounding is also productive in MSC:

#### Simple compounds

- (1) a. shu dan (MSC)  
book list  
'book list'

#### Recursive compound nouns

- (1) b. ling kou se (MSC)  
collar rim colour  
'colour of rim of collar'

- c. bo chen yi (MSC)  
thin inner shirt clothing  
'thin inner shirt'

Affixation is another way of deriving new words in MSC (Lin 2001; Packard 2000). The prefix *chong-* 'again' and the suffix *-zhe* 'one who does/is X' are illustrated below:

- (2) a. chong -du (MSC)  
again- read  
'read again:reread'

- b. ji -zhe (MSC)  
report -one who does/is X  
'one who reports:reporter'

Reduplication of nouns can also be observed in MSC, but in a very limited fashion (Lin 2001:69-71):

(3) a. gu gu (MSC)

aunt aunt

‘aunt’

(Lin 2001:69-71)

b. yue yue (MSC)

month month

‘every month’

(Lin 2001:69-71)

For kinship terms, as seen above in (3a), the non-reduplicated form and the reduplicated form are synonymous; but for other nouns, as seen in (3b), the meanings are different between the non-reduplicated and reduplicated forms, where the reduplicated form typically has repetitive meaning (Lin 2001:69,71). Reduplicated adjectives and verbs can also be seen in MSC (Lin 1990:129; Lin 2001:72-73). Reduplicated adjectives have intensified meaning compared with their non-reduplicated forms (Lin 1990:129; Lin 2001:72):

(4) ai ai (MSC)

short short

‘very short’

Reduplicated verbs normally bear attenuative meanings, compared with their non-reduplicated forms (Lin 1990:129; Lin 2001:72):

(5) dong dong (MSC)

move move

‘move (it) a bit’

Word-formation in OXC is similar to MSC in certain respects. More specifically, OXC adjectives and verbs are like their counterparts in MSC in just about every way: they can be simple or undergo affixation and compounding in similar ways. Reduplication can

also be found in OXC adjectives and verbs, following similar rules as in MSC. With adjectives, the meaning is intensified with reduplication (cf. Ren 2006):

- (6) wan wan (OXC)  
soft soft  
'very soft'

With verbs, reduplication adds a sense of repetition and/or continuation to the verb meaning (cf. Wang 2009):

- (7) hu hu (OXC)  
drink drink  
'drink again or drink repeatedly for a period of time'

Common nouns in OXC are different from their counterparts in MSC in certain respects, though. In contrast to their counterpart in MSC, nouns in OXC always have at least two overt constituents. They can be derived by affixation, as in (8), by the prefix *a-* and the suffix *-e* (neither of which affect the meaning of the noun).

- (8) a. a-ma (OXC)  
A-mother  
'mother'
- b. fei<sup>5</sup>-e (OXC)  
water-E  
'water'

Compounding is also a way of deriving common nouns in OXC:

- (9) a. gueng fei (OXC)  
boiling water  
'boiling water'

---

<sup>5</sup> IPA transcription of the *pinyin* [fei] is [fi].

b. kou<sup>6</sup> lian (OXC)  
mouth face  
'face'

Common nouns which are not made up of two morphemes as a result of affixation or compounding undergo reduplication:

(10) a. can can (OXC)  
scoop scoop  
'scoop'

b. hai hai<sup>7</sup> (OXC)  
shoe shoe  
'shoe'

c. nian nian (OXC)  
hole hole  
'hole'

Simple compounds in OXC may also undergo reduplication. For simple compounds, it is the item on the right that undergoes reduplication, as demonstrated in (11a) and (11b):

(11) a. hai di di (OXC)  
shoe bottom bottom  
'bottom of the shoe'

b. yi jia jia (OXC)  
clothing shelf shelf  
'clothing shelf'

---

<sup>6</sup> IPA transcription of the *pinyin* [kou] is [ku].

<sup>7</sup> IPA transcription of the *pinyin* [hai] is [xE].

Another interesting difference between MSC and OXC is that OXC does not have recursive compounding. The following ungrammatical example (12a) in OXC corresponds to the previous recursive compound noun (1b) of MSC and (12b) below in OXC corresponds to the previous recursive compound noun (1c) in MSC:

(12) a. \*lieng hu sei (OXC)  
collar rim colour  
'??'

b. \*dan han ta (OXC)  
thin sweat inner shirt  
'??'

This difference between the languages will be discussed in section 5.5.5 of chapter 5. It is likely that this difference is due to influence from one or several of the non-Sinitic languages in the region. The restriction against recursive compounding is not entirely strange, as many languages do not allow recursive compounding (Mukai 2008).

### 1.3 Contribution of the thesis to Chinese morphology and morphology in general

As mentioned earlier, in the post-1949 era, MSC was promoted in the whole country of China including the Xining region, and since then, it has gained great popularity among Han immigrants who came to this region after 1949 as well as among local people (Dede 2003:336-337, 342, 2006:323). As a result, the usage of OXC has been greatly reduced (Dede 2006; Zhang 2001). This change in the linguistic situation threatens the status of OXC and puts it in danger of extinction. Further to that, OXC word formation is poorly documented. Thus, this thesis will provide a comprehensive description of word formation in OXC and thereby hopefully, will contribute to enhancing its status as a variety of Chinese, alongside other languages used in this region. Moreover, in terms of theory, the comparison of MSC and OXC in the thesis will shed new light also on the morphology of MSC. More specifically, by looking at the reduplicated data in OXC, it will become clear that there are minimally two items in a word of MSC, which are the

Root and the categorizer. This will bring new insights into the nature of the primitives involved in word formation in MSC. For example, the Root will be considered to be always bound, never able to form an independent word in MSC. One consequence of this is that the ‘Bound Root’ (Packard 2000) will not be a valid primitive of MSC word formation, so a set of new primitives and rules are proposed.

#### **1.4 Organization of the thesis**

Chapter 2 describes all kinds of affixed words and compound words in MSC and OXC. In the section on compounding, modifier-head compound nouns, coordinate compounds, verb-object compounds and subject-predicate compounds are described as well as collocations which can be analyzed as either compounds or phrases. Recursive compound nouns in MSC are also looked at in this chapter. Affixed words and various types of compounds in MSC and OXC described in this chapter will be analyzed in chapter 4 and chapter 5 respectively based on the newly developed theory.

Chapter 3 introduces some of the basic concepts and the theoretical background for the theory developed in the thesis. In particular, it will review the word-syntax theory of Selkirk (1982) and Sproat and Shih (1996) regarding the word components of the MSC compounds, Packard’s (2000) theory of word structure in MSC, and Zhang’s (2007) Root Merger analysis of MSC compounds.

Chapter 4 presents the proposed theory of word formation in MSC. Firstly, new primitives for word structure in MSC are introduced. Then the structure of MSC Minimal Words, affixed words and simple compounds will be analyzed within the theory developed. Then the structure of MSC recursive compounds will be discussed.

Chapter 5 looks at the word structure of OXC. The reduplicated simple nouns in OXC are looked at firstly. Next the primitives in OXC nouns, adjectives and verbs are discussed with some comparisons to their counterparts in MSC. Then the attention will turn to the structure and the head position of OXC Reduplicated Nouns and Minimal Adjectives and Verbs. Affixed words in OXC will then be discussed. The following sections will concern



the head positions and structure of OXC simple compounds. Then the absence of recursive nominal compounds in OXC will be looked at. The last section of this chapter is a summary of the structure and head locations of various types of reduplicated nouns in OXC.

Chapter 6 is the summary of the thesis in which the research questions raised in chapter 1 will be answered.



## Chapter 2 Affixed words and compounds in Modern Standard Chinese and Old Xining Chinese: a description

### 2.1 Derivational Affixation in MSC and OXC

#### 2.1.1 Introduction

Derivational affixation is a word formation procedure in which a larger unit is generated by adding an affix to a bound item or a word (Li and Thompson 1981:36; Packard 2016:73). Derivational affixation can be found in both MSC and OXC. In the following section, productive derivational prefixes and suffixes in both languages are described. It will be seen that in both languages, in some cases, the affixation contributes to the meaning of the resultant word, but the word component the affix is merged with and the resultant word may also have a similar or identical meaning in other cases. Further, in OXC, in some cases, affixation can bring a sense of endearment or pejoration to the resulting word. Traditionally speaking, the item the affix is attached to is a noun or a verb or an adjective, and the result of the affixation either is or is not a different category. But later in chapter 4, based on the newly developed theory, it will become clear that this is not always the case. More specifically, it will be seen there are cases where the item which is merged with the affix can be a Root without word category. To account for such cases and accommodate the theory developed in chapter 4, I will therefore avoid calling the item the affix is merged with noun, verb, or adjective. Instead I will use semantic descriptions.

#### 2.1.2 MSC prefix

##### **-chong**

*Chong-* ‘again’ is a prefix which can merge with an event-denoting item and the resultant word is a verb.

- (1) a. chong- du  
again- read  
'read again: reread'
- b. chong-chao  
again- copy  
'copy (something) again'
- c. chong-zuo  
again- do  
'do (something) again'

### 2.1.3 MSC suffixes

#### **-hua**

*-Hua* is an affix in MSC which means 'make' (Packard 2000:70). It normally merges with a property-denoting item and the resultant word is a verb:

- (2) a. liang -hua  
bright-make  
'make bright'
- b. ruan-hua  
soft -make  
'make soft'
- c. qiang-hua  
strong -make  
'make strong'

#### **-tou**

*-Tou* is meaningless (Lin 2001:82) and normally merges with an object-denoting item. The resulting word is a noun (Lin 2001:58):

(3) a. e -tou

forehead -TOU

‘forehead’

b. quan-tou

fist-TOU

‘fist’

*-Tou* may also merge with a property-denoting item. In this case, *-tou* may have a general meaning which is ‘thing’ and the meaning of the resultant word, for example, whether it denotes a result or condition, depends on the item which *-tou* is merged with. The new word is a noun (Packard 2000:70, 2016:73):

(4) tian -tou

sweet -thing

‘sweetness’

(Lin 2001:59)

*-zi*

*-Zi* commonly merges with a property-denoting item and refers to ‘person who is X’, where X is the property denoted by the item the suffix is merged with. The resulting word belongs to the word category of noun. In other words, the suffix *-zi* is a nominalizer (Packard 2000:70, 2016:73):

(5) a. que-zi

limp -person

‘limp person’

b. shou-zi

thin -person

‘thin person’

- c. xia-zi  
 blind -person  
 ‘blind person’

-Zi can also merge with an object-denoting item and the resultant word is a noun. In this case, -zi is a meaningless suffix (Lin 2001:82; Sun 2006:58):

- (6) a. bei -zi  
 duvet -ZI  
 ‘duvet’

- b. chuang -zi  
 window -ZI  
 ‘window’

**-zhe**

-Zhe is a nominal suffix in MSC (Packard 2000:70) which can merge with a property-denoting item or an event-denoting item.<sup>8</sup> It has a meaning of ‘one who does/is X’ (Packard 2000:72):

- (7) a. ren -zhe  
 kind-one who does/is X  
 ‘one who is kind: kind person’

- b. yi -zhe  
 translate -one who does/is X  
 ‘one who translates: translator’

---

<sup>8</sup> The nominal suffix -zhe ‘one who does/is X’ here should be distinguished from the suffix -zhe which is used in Classical Chinese. These two are different suffixes. Examples of the suffix -zhe which are used in Classical Chinese are as follows: *ri-zhe* ‘day-ZHE: recently, nowadays’, *xi-zhe* ‘formerly-ZHE: erstwhile, in the past’. In the thesis, only the nominalizing suffix -zhe which is used in Modern Standard Chinese will be described and discussed.

- c. zuo -zhe  
write -one who does/is X  
'one who writes: writer'

#### 2.1.4 OXC prefixes

##### *a-*

*A-* is a purely formal prefix, adding no meaning to the new word. It is normally placed before kinship terms:

- (8) a-jiou  
A-uncle  
'uncle'

- a-ji  
A-sister  
'sister'

- a-yi  
A-grandfather  
'grandfather'

The prefix is found in some other varieties of Chinese as well. For example, in Cantonese, prefix *a-* can merge with monosyllabic surnames to denote familiarity (Chao 1968:216):

- (9) a-Liu  
A-Liu  
'Liu'

- a-Sun  
A-Sun  
'Sun'

a-Wang

A-Wang

‘Wang’

(Adapted from Chao 1968:216)

### ***ga-***

*Ga-* is an endearment term, which can merge with an object-denoting component. It can also be used alone meaning ‘small’. By using this prefix, the speaker is showing affection for an entity, for instance:

(10) ga-gou

GA-gou

‘dog’

ga-mo

GA-cat

‘cat’

ga-chei

GA-bike

‘bike’

It can be seen that the resulting word is a noun. The effect of the prefix is to imply that the speaker likes the dog, the cat or the bike.

### ***2.1.5 OXC suffixes***

#### ***-bong***

*-Bong* can be a suffix to an object-denoting or a property-denoting item. It designates ‘person who is associated with X’, where X is the entity or property denoted by the item the suffix is merged with. Furthermore, *-bong* denotes a pejorative notion, the speaker having contempt for the person who is being described. Some examples are shown below:



(11) a. xiong-bong  
countryside-person  
'country bumpkin'

b. nong-bong  
stupid-person  
'idiot, blockhead'

By using these words, the speaker indicates his low opinion of the person who is from the countryside (11a) or stupid (11b). Furthermore, the new words are nouns in those cases.

***-dan***

The suffix *-dan* can merge with a property-denoting item or an object-denoting item. It refers to 'person who is associated with X', where X is the entity or property denoted by the item the suffix is merged with. It is a pejorative term, which adds a sense of contempt. The new word which contains the suffix is a noun:

(12) jian-dan  
sly -person  
'sly person'

lan-dan  
lazy -person  
'lazy person'

rou-dan  
meat-person  
'stupid person:blockhead'

In these cases, the speaker disrespects the person who is described as sly, lazy or stupid.<sup>9</sup>

*-e*

*-e* is another suffix used in OXC, which is meaningless. When it merges with an object-denoting item, the resultant word belongs to the word category of noun:

(13) mo-e

cat- E

‘cat’

hai hai -e

shoe -E

‘shoe’

ca ji -e

tea table -E

‘tea table’

In addition, the suffix *-e* can also merge with a property-denoting item and the resultant word is an adjective<sup>10</sup>:

(14) jieng -e

clever-E

‘clever’

zhan-e

flat -E

‘flat’

---

<sup>9</sup> Note here *-bong* and *-dan* are analysed as suffixes and not as nouns. The criteria I adopted are from Sproat and Shih (1996) and whether an item is an affix or a word depends on its position in the resultant word. More specifically, if this item has a flexible position which means it can be on the right or left side of a resultant word, then it is not an affix and it is a word (Sproat and Shih 1996). But if this item has a fixed position, for instance, it can only be on the right-hand side of a resultant word, then this item must be an affix (Sproat and Shih 1996). On this basis, for instance *-bong* and *-dan* are suffixes, not nouns.

<sup>10</sup> Note here we do not have a suffixation or a vowel lengthening due to the predicative use of the item in OXC.

***-gou***

*-Gou* is a suffix to an event-denoting item, deriving an attenuative verb form. Examples are shown below:

(15) chong-gou

taste -GOU

‘taste a little bit (of something)’

fo -gou

talk -GOU

‘talk a little bit (about something)’

ju-gou

do-GOU

‘do (something) a little bit’

***-gei***

*-Gei* is a suffix which can merge with event-denoting items. The new word belongs to the word category of verb:

(16) hu-gei

fold-GEI

‘fold’

gueng-gei

roll-GEI

‘roll’

ju-gei

do-GEI

‘do’

*-zi*

The suffix *-zi* can merge with an object-denoting item in OXC without contributing any meaning to the resultant word. The resultant word belongs to the word category of noun:

(17) *gou-zi*

bottom-ZI

‘bottom’

*tu-zi*

apprentice-ZI

‘apprentice’

*za za-zi*

powder-ZI

‘powder’

In this section affixation in MSC and OXC have been described. The resultant words may share meanings with items the affixes are merged with. Another way of forming words in MSC and OXC is compounding. The following section will concern simple compounds in MSC and OXC and recursive compound nouns in MSC.

## **2.2 Compounds in MSC and OXC**

### ***2.2.1 Definition and classification***

A broad definition of the compound in Chinese is followed in this thesis. According to this definition, a compound is a word which is composed of free words or bound morphemes which are not affixes or clitics (Chao 1968:359; Li and Thompson 1981:46; Liao 2014:8-9). OXC has similar types of compounds to MSC. The following types of compounds in MSC and OXC are discussed in this section:

- Attributive compounds

- Coordinate compounds
- Verb-object compounds
- Subject-predicate compounds

Attributive compounds are compounds where one of the components is a modifier of the other component (Chao 1968; Ceccagno and Basciano 2007, 2011). Such compounds can be further classified into two groups which are:

- Attributive adjective-noun compounds
- Attributive noun-noun compounds

Furthermore, there are attributive collocations which may be analyzed as phrases and compounds under different circumstances (Chao 1968; Duanmu 1997; Sun 2015). This is also true in OXC.

Coordinate compounds are further categorized into compounds of synonyms, or, as I shall call them, synonymy-compounds, compounds of antonyms, or antonymy-compounds, and parallel compounds (Chao 1968:374-377). The following two kinds of synonymy-compounds are discussed in this section:

- synonymy-compound nouns
- synonymy-compound adjectives

Further, parallel compound nouns and parallel compound adjectives, in which the subparts are not synonyms and antonyms (Chao 1968:377), will also be described, followed by a description of antonymy-compounds (Chao 1968).

There are also verb-object compounds and subject-predicate compounds in MSC (Chao 1968; Huang 1984; Li and Thompson 1981; Liao 2014; Packard 2000). Those two types of compounds are also used in OXC. Those two types of compounds in MSC and OXC will be briefly discussed in chapter 4 and chapter 5. In what follows, those types of compounds/phrases mentioned above in both MSC and OXC will be illustrated.

## 2.2.2 Attributive phrases and compounds

### 2.2.2.1 Attributive adjective-noun collocations

There are attributive adjective-noun collocations in MSC in which the object-denoting component, on the right, is modified by the property-denoting component, on the left, and they can be expanded optionally (Chao 1968:285, 381; Dai 1992:99; Duanmu 1997: 159,162; Liao 2014:9; Sun 2015: 367-368). They can occur with or without the attributive particle *de*.

Insertion of the grammatical/phrasal marker *de*:

- (18) a. gao de                      lou    (MSC)  
        tall PHRASAL MARKER building  
        ‘a tall building’

The grammatical/phrasal marker is not inserted:

- (18) b. gao lou    (MSC)  
        tall building  
        ‘tall building’

This attributive particle *de* is considered as a grammatical/phrasal marker in a noun phrase in MSC which denotes a modification relation between a head noun or NP and its modifier (Chao 1968:285, 287, 381-382; Sun 2015: 366-368). An adjective-noun collocation with the grammatical/ phrasal marker *de* in MSC is generally considered as a phrase, while the corresponding collocation without *de* is a compound (Chao 1968: 285,381; Dai 1992:99; Duanmu 1997: 159,162; Sun 2015:367-368). The following is an argument in favor of this analysis from Duanmu (1997:151-152): first, assume that a phrase can only occur inside another phrase, not inside a word. Second, the collocation of a degree adverb and an adjective is an adjectival phrase. The phrase *hen gao* ‘very tall’ is such a phrase. Third, the adjective-noun collocation with insertion of *de*, as in (18a), is seen to be able to contain the adjectival phrase:

(19) *hen gao de lou* (MSC)  
 very tall PHRASAL MARKER building  
 ‘a very tall building’

It follows that the collocation (18a) must be a phrase. That is to say, an adjective-noun collocation with the grammatical/ phrasal marker *de* in MSC must be considered as a phrase (Chao 1968: 285,381; Duanmu 1997:159; Sun 2015:367-368). On the other hand, an adjective-noun collocation without the grammatical/ phrasal marker *de* must be considered as a compound in MSC (Chao 1968:285, 381; Dai 1992:99; Duanmu 1997:162; Sun 2015:367-368). (18b) is such a compound where there is no insertion of the phrasal marker *de*. Given that a compound, being a word, cannot contain a phrase (Duanmu 1997: 151-152), the prediction is that the phrasal structure [adverb, adjective], for instance *hen gao*, cannot occur inside the collocation (18b). As demonstrated below, the prediction is right.

(20) (\**hen*) *gao lou* (MSC)  
 very tall building  
 ‘??’

A similar situation can be found in some OXC adjective-noun collocations, where the attributive particle *zi* is also optional:

(21) a. *hueng zi bu* (OXC)  
 red ZI fabric  
 ‘a red fabric’  
 b. *hueng bu*  
 red fabric  
 ‘a red fabric’

I will argue that this attributive particle *zi* is a grammatical/phrasal marker in OXC. Duanmu’s (1997) analysis is also adopted here. First, a phrase is only allowed inside another phrase (Duanmu 1997:151-152). Second, the combination of the degree adverb

and adjective is an adjectival phrase (Duanmu 1997:151). When the attributive particle *zi* is inserted into an adjective-noun collocation in OXC, the collocation can include this adjectival phrase. More specifically, the adjectival component of (21a) can be modified by a degree adverb *hudu* ‘very’, as shown below. This follows if (21a) is a phrase:

(22) *hudu hueng zi bv* (OXC)  
 very red ZI fabric  
 ‘a very red fabric’

As for (21b), it is classified as a compound. The argument for this classification is that when the attributive particle *zi* is not inserted into the adjective-noun collocation, the collocation cannot contain an adjectival phrase. Below it is demonstrated that the property-denoting component of (21b) cannot be modified by the degree adverb *hudu* ‘very’. This follows if the collocation (21b) is a compound noun, not a phrase.

(23) (\**hudu*) *hueng bv* (OXC)  
 very red fabric  
 Intended reading: ‘a very red fabric’

It has been demonstrated that the attributive adjective-noun collocation with the attributive particle *zi* is a phrase in OXC; without the attributive particle *zi*, it is a compound.

#### 2.2.2.2. *Attributive noun-noun collocations*

There are also attributive noun-noun collocations in MSC in which the grammatical/ phrasal marker *de* can be inserted optionally (Sun 2015:366, 373).

The grammatical/ phrasal marker *de* is inserted:

(24) a. *pi de xie* (MSC)  
 leather PHRASAL MARKER shoe  
 ‘leather shoe’





optionally inserted in modifier noun-noun collocation in OXC:

The attributive particle *zi* is inserted:

(26) a. *yong zi tuei* (OXC)  
sheep ZI leg  
'sheep leg'

The attributive particle *zi* is not inserted:

(26) b. *yong tuei* (OXC)  
sheep leg  
'sheep leg'

With the insertion of the attributive particle *zi*, the attributive noun-noun collocation, that is, (26a) can include a phrase, e.g. *si go tuei* 'four CLASSIFIER leg: four legs':

(27) *yong zi si go tuei* (OXC)  
sheep PHRASAL MARKER four CLASSIFIER leg  
'four sheep legs'

So (26a) must be a phrase. On the other hand, for (26b), which is an attributive noun-noun collocation without the insertion of the attributive particle *zi*. It is expected that inclusion of the phrase *si go tuei* 'four CLASSIFIER leg: four legs' would cause ungrammaticality, as is indeed the case:

(28)\* *yong si go tuei* (OXC)  
sheep four CLASSIFIER leg  
(cf. (20), (23), (25)) Intended reading: 'four sheep legs'

(26b) has to be a compound, otherwise (28) should be grammatical. From these examples, it is seen that for the attributive noun-noun collocation in OXC, with the attributive particle *zi*, the collocation is a phrase, and without it, it is a compound. The behavior of (21a), (21b), (26a) and (26b) has shown that the attributive particle *zi* is a grammatical/phrasal marker in OXC.

In short, some attributive adjective-noun collocations and attributive noun-noun collocations in MSC and OXC are shown to be phrases and compounds under different conditions. More specifically, such collocations are phrases if they are expanded using the grammatical/phrasal marker *de* or *zi*; and without this marker they are compounds (Chao 1968: 285,381-382; Dai 1992:99; Duanmu 1997: 159,162; Sun 2015:366-368). Further, there are also MSC attributive collocations which can only be compounds, as they cannot have the phrasal marker *de* inserted (Chao 1968:382; Sun 2015:372). The situation is the same in OXC where the phrasal marker *zi* cannot be inserted in some attributive collocations. Some attributive collocations in both languages are shown below:

#### Attributive adjective-noun compounds

(29) a. e (\*de) ren (MSC)  
 bad PHRASAL MARKER person  
 Intended reading: 'bad person'

b. jieng (\*zi) du (OXC)  
 bare PHRASAL MARKER belly  
 Intended reading : 'bare belly'

#### Attributive noun-noun compounds

(30) a. shu (\*de) gui (MSC)  
 book PHRASAL MARKER cupboard  
 Intended reading : 'book cupboard'

b. han (\*zi) ta (OXC)  
 sweat PHRASAL MARKER inner shirt  
 Intended reading : 'sweat inner-shirt'

### 2.2.3 Coordinate compounds

There are also coordinate compounds in MSC (Chao 1968:372) consisting of object

denoting items or property denoting items; other combinations are very rare. Depending on the meanings of constituents, coordinate compounds in MSC can be further classified into three groups: compounds of synonyms, henceforth called synonymy-compounds, parallel compounds and antonymy-compounds (Chao 1968:374-375,377). Synonymy-compounds are compounds where the constituents are synonyms (Chao 1968:374-375, Ceccagno and Basciano 2011:481). Here, compounds in both MSC and OXC where components have identical meanings will be termed synonymy-compounds. Parallel compounds are compounds in which constituents are neither close enough to be considered as synonyms nor opposite enough to be considered as antonyms, according to Chao (1968:377). Here, I will apply the term parallel compound to refer to compounds in MSC and OXC where components neither have identical meanings nor opposite meanings and one of components shares the meaning with the resultant compound. Antonymy-compounds, finally, are compounds where the constituents are antonyms (Chao 1968:375). Those three types of coordinate compounds are also found in OXC. Synonymy-compound nouns and parallel compound nouns in MSC and OXC are demonstrated below:

#### Synonymy-compound nouns

- (31) a. shu      cai (MSC)  
vegetable vegetable  
'vegetable'

#### Parallel compound nouns

- (31) b. er    duo (MSC)  
ear earlobe  
'ear'

- c. chuang   hu (MSC)  
window house  
'window'

Synonymy-compound nouns

- (32) a. beng geng (OXC)  
root root  
'root'

Parallel compound nouns

- (32) b. lian no (OXC)  
face brain  
'face'

- c. kou lian (OXC)  
mouth face  
'face'

Those two types of coordinate nominal compounds are different from conjunction phrases in the sense that the order of the parts is irreversible (Chao 1968:372; Li and Thompson 1981:54). For instance, the above example *shu cai* 'vegetable' from MSC does not have a reversed form *\*cai shu*, and in OXC, the above example *kou lian* 'mouth face: face' cannot be reversed as *\*lian kou*. Synonymy-compounds and parallel compounds which are adjectives in MSC and OXC are shown below:

Synonymy-compound adjectives

- (33) a. mei li (MSC)  
beautiful beautiful  
'beautiful'

Parallel compound adjectives

- (33) b. leng mo (MSC)  
cold indifferent  
'cold'

c. wei ruo (MSC)  
small weak  
'weak'

Synonymy-compound adjectives

(34) a. nong zong (OXC)  
dirty dirty  
'dirty'

Parallel compound adjectives

(34) b. lan gan (OXC)  
lazy dry  
'lazy'

c. fu tan (OXC)  
comfortable flat  
'comfortable'

Again, the order is fixed for such compound adjectives in MSC (Chao 1968:372; Li and Thompson 1981:54), as well as in OXC. For instance, *mei li* 'beautiful beautiful: beautiful' cannot be reversed as *\*li mei*, and in OXC *nong zong* 'dirty dirty: dirty' cannot be *\*zong nong*.

In the antonymy-compounds, the meanings of the constituents are opposite to each other (Chao 1968:375) and both constituents are property-denoting items. The resulting compounds have idiomatic meaning, whose word category is most likely to be noun (Li and Thompson 1981:81). For instance, *mei chou* 'beautiful ugly: appearance' is an example from MSC. In OXC, antonymy-compounds can also be observed, for example, *fei sou* 'fat thin: size'. Similar to the previous types of coordinate compounds, constituents of antonymy-compounds are also irreversible, in MSC (Chao 1968:372) and OXC. That is to say, the above example *mei chou* 'beautiful ugly: appearance' cannot be *\*chou mei* in MSC and *fei sou* 'fat thin: size' cannot be reversed as *\*sou fei* in OXC.

## 2.2.4 Remaining compounds

### 2.2.4.1 Verb-object compounds

Verb-object compounds and subject-predicate compounds in MSC and OXC will not be discussed in great detail in this thesis. The descriptions of those two types of compounds are seen below. We know that some verb-object constructions are compounds since the components cannot be separated, so for example, the insertion of aspect markers is not allowed (Li and Thompson 1981:75; Liao 2014:18). Aspect markers being functional heads which denote properties of predicates in MSC (Lin 2010; Shen 2004) cannot occur inside a word. They also typically bear idiomatic meanings (Li and Thompson 1981:75). It is a similar picture in OXC. Below the verb-object compounds are demonstrated:

(35) a. dan xin (MSC)  
carry heart  
'worry'

b. tiou zuei (OXC)  
adjust mouth  
'flirt'

### 2.2.4.2 Subject-predicate compounds

Subject-predicate style compounds are common in MSC, consisting of an object-denoting item and an event-denoting item which appear to be in a subject-predicate relation. The subject-predicate compounds can be compound verbs and adjectives, which can take modification of adverbs and negation (Chao 1968:369-370; Li and Thompson 1981:70-71). But on the other hand, modification of adverbs or negation of the event-denoting item alone is not allowed (Li and Thompson 1981:70). Further, the subject-predicate compounds in MSC may have idiomatic meanings (Li and Thompson 1981:70). The situation is the same in OXC. The following subject-predicate compounds, which are adjectives, from both MSC and OXC are demonstrated below.

(36) a. xin sui (MSC)  
heart break  
'heartbroken'

b. nian guei (OXC)  
word precious  
'reserved'

### 2.2.5 *Recursive compounds in MSC*

So far, simple compounds in MSC and OXC have been described, compounds which are a result of merge of two items. In addition to simple compounds, recursive compound nouns can also be found in MSC, which is briefly discussed by Packard (2000).

Compared with MSC, recursive compound nouns are rarely used in OXC. The rest of this section concerns recursive compound nouns in MSC only.

For recursive compounds in MSC, I will follow Mukai (2008) and claim that, just like simple compounds, they are derived by the operation Merge (Chomsky 1995:243-245, 2015:223-227). Merge of a simple compound noun with an item on the right would produce a left-branching compound and merge of a compound noun with an item on the left would produce a right-branching compound. In contrast to Bisetto (2010), I will classify both left-branching and right-branching compounds as recursive compounds in MSC, following, in this regard, Roeper, Snyder and Hiramatsu (2002), Booij (2011) and Mukai (2015).

The recursive compound noun containing an attributive compound noun and an item is demonstrated first, followed by the recursive compound containing a synonymy-compound noun and an item, and finally the recursive compound containing a parallel compound noun and an item.



Attributive compound noun

(37) a. *fa xing*

hair style

‘hair style’

b. *shu chu*

book cupboard

‘book cupboard’

The attributive compound noun *fa xing* ‘hair style’ is seen below merged with an item *wang* ‘website’ on the right-hand side, which is an object-denoting item in MSC and together they form a left-branching recursive compound:

(38) [[*fa xing*] *wang*]

hair style website

‘hair style website’

On other hand, the attributive compound noun can also merge with an item on the left, and a right-branching recursive compound is generated as a result. *Shu chu* ‘book cupboard: book cupboard’ is such an attributive compound noun which is enlarged in (39) by means of the merger of a left-hand side object-denoting item *mu* ‘wood’.

(39) [*mu* [*shu chu*]]

wood book cupboard

‘book cupboard made of wood’

In addition, the merged item can also be a property-denoting item:

(40) [*xin* [*shu chu*]]

new book cupboard

‘new book cupboard’

Similar to the previously discussed attributive nominal collocations (see section 2.2.2), (38) and (40) may be compounds or phrases (Dai 1992), as the phrasal marker *de* can be

inserted optionally. This demonstrates that (38) and (40) in which the phrasal marker *de* is not inserted are compounds. But when (38) and (40) are inserted with the phrasal marker *de*, as shown below, the resultant collocations are phrases:

(41) a. [[fa xing] de wang]  
hair style PHRASAL MARKER website  
'hair style website'

b. [xin de [shu chu]]  
new PHRASAL MARKER book cupboard  
'new book cupboard'

As for the collocation (39), it can only be a compound, as it cannot take the phrasal marker *de*:

(42) [mu (\*de) [shu chu]]  
wood PHRASAL MARKER book cupboard  
Intended reading: 'book cupboard made of wood'

The recursive compound nouns may also contain a synonymy-compound noun such as (43a) or a parallel compound noun, such as (43b):

Synonymy-compound noun

(43) a. yi fu  
clothing clothing  
'clothing'

Parallel compound nouns

(43) b. chuang hu  
window house  
'window'

Below, it can be seen that the synonymy-compound noun *yi fu* 'clothing' merges with the

right-hand side item *se* ‘colour’, while the parallel compound noun *chuang hu* ‘window house: window’ merges with *tu* ‘picture’ on the right. In both cases, the resultant compounds are left-branching recursive compounds:

(44) a. [yi fu se]  
 clothing clothing colour  
 ‘colour of clothing’

b. [[chuang hu] tu]  
 window house picture  
 ‘picture of window’

Another possibility is that the synonymy-compound noun and the parallel compound noun merge with an item on the left. The synonymy-compound noun *yi fu* ‘clothing’ is seen below with the item *mian* ‘cotton’, which is an object-denoting item. For the parallel compound noun *chuang hu* ‘window house: window’, it is the object-denoting item *tie* ‘iron’ that *chuang hu* ‘window house: window’ is merged with. On both occasions, right-branching recursive compounds are generated:

(45) a. [mian [yi fu]]  
 cotton clothing clothing  
 ‘cotton clothing’

b. [tie [chuang hu]]  
 iron window house  
 ‘window made of iron’

Furthermore, the synonymy-compound noun and the parallel compound noun can also merge with an item which is a property-denoting item:

(46) a. [hao [yi fu]]  
 good clothing clothing  
 ‘good clothing’

- b. [zang [chuang hu]]  
dirty window house  
'dirty window'

Here, *hao* 'good' and *zang* 'dirty' are merged with the synonymy-compound noun *yi fu* 'clothing' and *chuang hu* 'window house: window' respectively. The above compounds (44a), (44b), (45a), (45b), (46a) and (46b) can be compounds or phrases (Dai 1992), as insertion of *de* is optional in them.

It has been demonstrated that a compound noun, that is, an attributive compound noun, a synonymy-compound noun or a parallel compound noun can merge with an item either on the right or the left. As a result, either a left-branching or a right-branching recursive compound noun is generated.

Note that in Bare Phrase Structure theory (Chomsky 1995: 241-249), which is the basis for the theory assumed in this thesis, there is no 'merge on the left or right'. These terms are used here merely for ease of exposition. There is merge of items A and B which forms a set {A, B}. The question then is, which of the two items merged is the head? This, then determines the grammatical properties of the set and what the linear order is, i.e. how the set is phonetically spelled out (Chomsky 1995:243-246, 2015:223-227,308). MSC nominal compounds are generally head-final (Packard 2000). If the recursive compound has a structure which is [<sub>C</sub> C [<sub>A</sub> B A]], then the spell out form is B>A>C, and a so-called left-branching compound is produced. If the recursive compound has the structure [<sub>A</sub> C [<sub>A</sub> B A]], then the spell-out is C>B>A and a so called right-branching compound is produced as a result.

Furthermore, the recursion can also involve two simple compound nouns (Packard 2000:193). Attributive compound noun, synonymy-compound noun and parallel compound noun are three types of compound nouns in MSC which can merge with each other to form complex recursive compound nouns. The two merged simple compounds may be of the same type, but may also be of different types, in any combination. More specifically, an attributive compound noun can merge with a synonymy-compound noun

or a parallel compound noun.

An attributive compound noun merging with a synonymy-compound noun:

(47) a. [[you zhi] [yi fu]]  
high quality clothing clothing  
'high quality clothing'

b. [[shu cai] [bi zhi]]  
vegetable vegetable wall paper  
'vegetable wallpaper'

(47a) is a case where the attributive compound noun *you zhi* 'high quality' is merged with the synonymy-compound noun *yi fu* 'clothing'. (47b) shows the synonymy-compound noun *shu cai* 'vegetable' is merged with the attributive compound noun *bi zhi* 'wallpaper'.

An attributive compound noun merging with a parallel compound noun:

(48) a. [[shui bei] [zhi liang]]  
water cup quality quantity  
'water cup quality'

b. [[chuang hu] [mei tu]]  
window house beautiful picture  
'beautiful picture of a window'

Here (48a) contains the attributive compound noun *shui bei* 'water cup' and the parallel compound noun *zhi liang* 'quality'; in (48b), there is the parallel compound noun *chuang hu* 'window' and the attributive compound noun *mei tu* 'beautiful picture'.

A synonymy-compound noun merging with a parallel compound noun:

(49) a. [shu cai] [tu pian]  
vegetable vegetable picture film  
'vegetable picture'

b. [wen xue] [yi shu]  
literature study art art  
'art of literature'

Here (49a) is a result of merge of a synonymy-compound noun *shu cai* 'vegetable' and a parallel compound noun *tu pian* 'picture film: picture'; and (49b) is a result of a merge of a parallel compound noun *wen xue* 'literature study: literature' and a synonymy-compound noun *yi shu* 'art'.

Two attributive compounds nouns can also merge together to form a recursive compound noun:

(50) [hu qu] [bi zhi]  
lake district wall paper  
'wall paper of lake district'

Moreover, a complex recursive compound noun may also be formed of two synonymy-compound nouns:

(51) [bo li] [yi shu]  
glass glass art art  
'glass art'

Two parallel compound nouns can also merge together to form a complex recursive compound noun in MSC:

(52) [fa lv] [wen shu]  
law rule document book  
'law book'

Furthermore, the recursive collocations (47a), (47b), (48a), (48b), (49a), (49b), (50) (51) and (52) can be compounds as well as phrases (Dai 1992), depending on if the phrasal marker *de* is inserted or not.

The merger of the various types of compounds has been discussed above and there is nothing unexpected about the facts. If a simple compound composed of two nouns is a noun, then it is predicted that this noun can merge with another noun, to form a bigger compound noun. The synonymy compound nouns and parallel compound nouns are not immediately obviously seen to be nouns just like other nouns, as they do not seem to have a designated head. But the multiple compounding facts above indicate that they are indeed like other nouns.

### **2.3 Summary**

This chapter focuses on the description of word-formation in MSC and OXC. Particularly, affixation and compounding in both languages are described. As for affixation, the resultant word may share the meaning of the item the affix is merged with, or the resultant word may have different meaning. Furthermore, there are clear cases of simple compounds in both languages. But there are also some collocations in both languages, which behave like compounds in one condition and phrases in another. A description of recursive compound nouns in MSC is also provided and it has shown that the recursive compound noun can be the result of a merge of an item with a simple compound noun or a simple compound noun with another simple compound noun. In chapter 4 and chapter 5, affixed words and various kinds of compounds described in this chapter will be discussed based on the newly developed theory.





## Chapter 3 Theoretical background: previous analyses of word structure

In this chapter, basic concepts of morphology in general and in MSC in particular will be introduced first. Then the word-syntax theory of Selkirk (1982) is looked at. The following section will concern bound primitives with lexical meaning in MSC word structure in particular as well as the structure of words of MSC in general. The works Dai (1992), Sproat and Shih (1996), Packard (2000) and Zhang (2007) are discussed in this section.

### 3.1 Basic concepts: morpheme, Root and stem

Morphemes are minimal units in terms of semantic content or grammatical function. Morphemes can be free or bound. A morpheme is free if it can be used as a word on its own. A bound morpheme is always part of a word (Spencer 1991:5).

In a word, after removing all kinds of affixes, that is, the derivational affixes and inflectional affixes, what remains in that word is labelled as a Root (Bauer 1983:20; Lieber 2016:38; 229). A Root is the smallest item in the sense it cannot be further analysed (Bauer 1983:20; Lieber 2016:38; Plag 2003:11). For instance, in the word *dislikeable*, *like* is the Root after removing the derivational prefix *dis-* and derivational suffix *-able*. A stem is related to inflectional morphology mainly and is what remains in a word after the removal of all the inflectional affixes (Lieber 2016:39, 230; Plag 2003:11). For example, in the word *adulthoods*, *adulthood* is the stem and the inflectional suffix is *-s*. The stem can be further divided into the Root *adult* and the derivational suffix *-hood*. A stem may be composed of two Roots (Bauer 1983:20), as in the word *schoolteacher*, which is also the stem *schoolteacher*, which can be further derived from two Roots, namely, *school* and *teacher*. The Root *school* and the Root *teacher* can both stand alone as a free word (Plag 2003:10). There are also Bound Roots, which cannot appear alone as free words, but only together with other types of morphemes (Plag 2003:10). For example, *hap-* is a Bound Root in the word *hapless* (Plag 2003:10).

This is traditional morphology. The theory in the present work follows works in Distributed morphology and proposes that the Root as a primitive of word structure of MSC is always bound and has no word category. It follows that a Root alone cannot be a free word and it can only form a free word with another item. This indicates that the items which are labelled as Bound Roots in traditional morphology are not Bound Roots. Instead, they are Bound Words. The Bound Word is another primitive proposed in the present work, and it is bound and always has a word category. It cannot occur alone as a free word, but instead a Bound Word can make a free word out of another item. In addition, stems are not assumed in the present work.

The following sections review some works concerning the structure of words in English and MSC. Reviewing the word-syntax theory developed by Selkirk (1982) is motivated in that her theory for English words represents a break with traditional morphology, not assuming any stems, just Roots, words, and affixes. She also applies the X-bar theory familiar from syntax to morphology. What follows are details of her theory and discussion of it.

### **3.2 Selkirk (1982)**

Selkirk (1982:2) recognizes the connection between morphology and syntax. She proposes an X-bar framework of morphology in which certain formal properties and rewrite rules of X-bar theory of syntax are applied to explain the structure of words (Selkirk 1982:2). In terms of formal properties, morphological categories are taken to share many characteristics with syntactic categories (Selkirk 1982:7). More specifically, words are structured in the fashion of the X-bar hierarchy of syntax (Selkirk 1982:7). A word is symbolized as  $X^0$ , which is the 'highest' level of word structure (Selkirk 1982:7). Dominated by  $X^0$  is the Root, which is represented as  $X^r$  (Selkirk 1982:95, 98). A morphological word and a syntactic word in Selkirk's framework have the same bar value [0] (Selkirk 1982:6-7). This indicates that in her framework, the syntactic component is naturally connected to the morphological component, as a complete morphological word is able to fill in the syntactic form class slot (Packard 2000:138). The word components are made up of features which include the syntactic category features, such as [ $\pm$ Noun]

and diacritic features (Selkirk 1982:7). Selkirk (1982:6-7) proposes that word structure can be generated by the rewrite rule following the X-bar format, which is shown below:

$$(1) X^n \rightarrow \dots X^{n-1} \dots$$

(Selkirk 1982:6)

Further, the notion head used in syntax is also applicable to word structure, as words have heads (Selkirk 1982:9). As for English words, Selkirk (1982:98) proposes that the Root, the word and the affix are primitives, while the stem is not needed in her theory (Selkirk 1982:98). The Root has the designation  $X^r$  and the prefix and suffix are designated as  $Y^{af}$  and  $X^{af}$  respectively (Selkirk 1982:95). According to Selkirk (1982:98), the Root can be bound or free, which means it can appear alone as a free word. Any monomorphemic nonaffixal item is said to be a Root as well as a word (Selkirk 1982:98). An English word, such as *dad* which only contains a monomorphemic Root (Selkirk 1982:98), is derived by the rule (2). This is the rule which allows a Root to form a word.

$$(2) X^0 \rightarrow X^r$$

(Selkirk 1982:95)

According to Selkirk's word structure framework, word generation is no longer a matter of idiosyncratic knowledge which is simply stored in the lexicon. Rather, there is a generative system in place to account for how words are generated and this system is similar to that of syntax in many respects (Packard 2000:149-150). More specifically, the category levels/names familiar from syntax and context-free rewriting rules are shown to be able to generate the word structure. Those similarities and the obvious connection would cast doubt on whether the two components, that is, the morphological component and syntactic component, are indeed separate components, raising the possibility that there is a single component which is responsible for generating both morphological objects and syntactic objects. This general approach has received support in more recent work on morphology, and will be developed further in the present work as well.

One problem is this: if a Root has the same content/features as the corresponding word, then why can the Root not occur on its own, like a word? This is handled in Selkirk's

theory by rule (2), which stipulates that a Root which is not merged with any other word component is a word. It would be preferable if the distinction between Root and word would follow from independently motivated assumptions. This is what I will propose in the next chapter, where, following work in Distributed Morphology (Embick and Noyer 2007,2008; Marantz 1997,2001; Harley and Noyer 1999; Halle and Marantz 1993), I will assume that what distinguishes Roots from words is that they have no categorial feature, and that their structural behaviour follows from this. Within this theory, Selkirk's rule (2) is not needed, and the morphological rule system is correspondingly simplified.

### 3.3 Bound items in MSC

MSC is considered as an analytic language, with little usage of inflectional affixation (Lin 2001; Packard 2000). Instead, grammatical functions such as tense in MSC are usually indicated by separate words (particles) (Lin 2001:122; Packard 2000:1). Derivational affixation is also sparse (Lin 2001:56). But compound words are common in MSC (Packard 2000:2). A characteristic feature of MSC is that there are a number of items in MSC which only occur bound, either as part of a compound or merged with a derivational affix. Some of them are listed below:

- (3) ping lv jin ju min mao shi lu jian  
 'apple' 'rule' 'mind' 'tool' 'vessel' 'appearance' 'oath' 'land' 'simple'
- jie xing ju liang ci lang jing shang  
 'clean' 'prosperous' 'enormous' 'fine' 'kind' 'bright' 'respect' 'worship'

These items are bound which means that they cannot stand alone as free words in phrases. Below, it is shown that the bound item *ping* 'apple' cannot stand alone as a noun in a sentence:

- (4) \* zhe ge ping hen da.  
 this CLASSIFIER apple very big  
 Intended reading: 'This apple is very big'.

On the other hand, items illustrated in (3) can form words with other items. In the following example, it is demonstrated that the bound item *ping* ‘apple’ forms a compound together with another bound item *guo* ‘fruit’:

- (5) ping guo  
apple fruit  
‘apple’

These bound items with lexical content have been discussed in the literature on Chinese morphology. Some of this literature is reviewed here.

### 3.3.1 *Dai (1992)*

Dai (1992:64) assumes that only free words can form compounds in MSC. The collocations that involve a bound item with lexical content, what he refers to as a bound stem, are considered as bimorphemic lexemes, which are not analysable in terms of syntactic or morphological rules, but instead are listed in the lexicon (Dai 1992:40,75-76). Dai’s idea is not without problems. First, to regard the bound items in compound words as bound stems is rather confusing, as there are hardly any inflections, so stems are almost invariably identical to Roots (possibly apart from compounds, if they are analyzed as stems). As mentioned, Selkirk (1982:98) makes no use of the notion stem in her account of English morphology as it is deemed unnecessary. In MSC, because of its analytic character, it would seem to be even less necessary.

Further, the nominal compounds made up of bound items with lexical content in MSC are in fact highly productive and could therefore not be listed in the lexicon (Sproat and Shih 1996:51,66). That is to say, in addition to the free word, a bound item can also be the component of MSC compounds. This is the view of Sproat and Shih on compound nouns and their constituents in MSC, which will be expanded on in the following section.

### 3.3.2 Sproat and Shih (1996)

In contrast to Dai, Sproat and Shih (1996:51) claim that bound items may also participate in the formation of productive compound nouns in MSC. They discuss whether the bound items are affixes or Bound Roots. The bound items cannot stand alone as free words; in this respect they are like affixes, but the same item can occur in different positions in compounds (Sproat and Shih 1996:50). For instance, the bound item *xiang* ‘elephant’ occurs on the right hand side, hence as a head, in the nominal compound *da-xiang* ‘big-elephant: elephant’. This bound item may also occur on the left-hand side, hence as a modifier, of another nominal compound, such as *xiang-qun* ‘elephant group’. This makes them different from affixes, which have a fixed position in the word (Sproat and Shih 1996:50). Furthermore, the bound items have lexical, referential content whereas affixes do not have such meanings (Packard 2000:72-73). Thus, the bound items are classified as Bound Roots (Sproat and Shih 1996:50).

Sproat and Shih (1996:50) observe that a Bound Root can form a compound noun with another Bound Root in MSC. For instance, the compound noun *mu zhou* ‘wood boat: boat made of wood’ is formed of the Bound Root *mu* ‘wood’ and the Bound Root *zhou* ‘boat’. A free word can occur with a Bound Root and together they form a free word in MSC (Sproat and Shih 1996:50). For example, in the nominal compound *xue xing* ‘blood type’, there is a free word *xue* ‘blood’ and a Bound Root *xing* ‘type’. In addition, a Root can also be free in the sense that it alone forms a monomorphemic word in MSC (Sproat and Shih 1996:59): for example, *ya* ‘duck’ may be a Root as well as a word. This implies that Selkirk’s rule (2) is still valid in Sproat and Shih’s theory. In short, according to Sproat and Shih (1996), simple nominal compounds which contain Roots in MSC may have the following structures:

- (6) a.[Bound Root    Bound Root]
- b.[Bound Root    free word]
- c.[free word     Bound Root]

(Adapted from Sproat and Shih 1996:50)

Following the standard assumption that the head is located on the right in a compound noun in MSC and the category of the head determines the category of the resultant compound, Sproat and Shih (1996:50ff) claim that a Bound Root in a compound noun may also bear the word category, which is noun. A free Root forms a noun alone, so it is assumed that it will have word category too. Similarly, Pirani (2008) also considers that those Bound Roots which are constituents of compound nouns in MSC have word category.

Compared with Dai, there are improvements in Sproat and Shih's account of compound structure in the sense that different types of items in MSC are recognized, some bound, some (potentially) free. Bound Roots are taken to have word class features, which accounts for the word category of MSC words containing two bound items.

Sproat and Shih's theory (1996) concerning the Root is similar to Selkirk (1982) in the sense that a Root is bound except when it is the only item of the word. So the theory of Sproat and Shih has the same problem as that of Selkirk's, which concerns the independence of a Root. A Root should be distinguished from a word by its own properties rather than by a rule, as will be elaborated on in chapter 4. In addition to Dai (1992) and Sproat and Shih (1996), Packard (2000) also discusses the notion of Bound Root, but in a X-bar theory of MSC morphology. The following section considers his notion of the Bound Root as well as this X-bar theory of MSC morphology.

### **3.4 Packard (2000): a detailed theory of MSC word structure**

Similar to Selkirk (1982), Packard (2000:163ff) believes that X-bar theory is extendable to word structure.<sup>11</sup> He particularly argues that formal properties of X-bar theory, including category names and category levels, can also be applied to the structure of MSC words. According to Packard (2000:165-166), the primitives of MSC word structure are

---

<sup>11</sup> In addition to Packard, Tang (1993) is another scholar who extends the X-bar theory to the word structure of MSC compounds. Please see Packard (2000:158-163) for a critical review of Tang's work.

Root Word, Bound Root, word-forming affix and grammatical affix. Here the Bound Root is a primitive of word structure of MSC words, which is consistent with what Sproat and Shih (1996) suggest.

A Root Word has the same bar value as that of the syntactic word, that is, [0], and it is represented as  $X^0$  (Packard 2000:165-166). According to Packard (2000:165-166), a Root Word may be an independent word, which means it can stand alone as a free word in a phrase in this respect, it is not different from the syntactic word designated as  $X^0$ , or it can form a word with other types of primitives, as head or non-head (Packard 2000:165-166). For the Bound Root, like Sproat and Shih (1996), Packard (2000:166) also notes that it cannot be used alone as a free word but can only form a word with another primitive in MSC, in which it may be the head. The Bound Root is therefore assigned a bar value lower than that of the Root Word, and is represented as  $X^{-1}$  (Packard 2000:166). Below in (7a), it is shown that the Root Word *men* ‘door’ is able to occur alone as a noun in a ‘Numeral-classifier-N’ phrasal construction (Duanmu 1997), and in (7b), it is shown that the Bound Root *shi* ‘room’ cannot do so.

(7) a. *yi shan men*  
 one CLASSIFIER door  
 ‘a door’

b. \* *yi jian shi*  
 one CLASSIFIER room  
 ‘??’

In (8), it is demonstrated that a Bound Root can form a word with a Root word or a Bound Root. In (8a), the Root Word *mei* ‘good’ forms the free word *mei mao* ‘good looks’ with the Bound Root *mao* ‘looks’. The head is the right-hand side Bound Root *mao* ‘looks’. In (8b), the Bound Root *mu* ‘wood’ makes a free word *mu he* ‘wood box: box made of wood’ out of the Root Word *he* ‘box’, and the Root Word is the head. In (8c), two Root Words which are *mei* ‘beautiful’ and *hua* ‘picture’ are seen to form a free word *mei hua* ‘beautiful picture’ and the head is the right-hand Root word *hua* ‘picture’. In (8d), two Bound Roots, which are *shi* ‘stone’ and *yi* ‘chair’ respectively form the free



word *shi yi* ‘stone chair: chair made of stone’ and the right-hand side Bound Root *yi* ‘chair’ is the head.

(8) a. *mei mao*

good looks

‘good looks’

b. *mu he*

wood box

‘box made of wood’

c. *mei hua*

beautiful picture

‘beautiful picture’

d. *shi yi*

stone chair

‘chair made of stone’

In short, similarly to a Root Word, a Bound Root also has lexical content and can be head of a word (Packard 2000:166). But unlike a root word, a Bound Root cannot be used independently as a free word (Packard 2000:166).

In addition to the Root Word and the Bound Root, affixes, including the word-forming affix and the grammatical affix, are primitives of word structure of MSC, in Packard’s theory. The word-forming affix resembles the traditional derivational affix and is represented as  $X^W$  (Packard 2000:165-166). Like the Root Word and Bound Root, a word-forming affix can be head of a word (Packard 2000:166). However, in contrast to the Root Word and the Bound Root, the word-forming affix cannot be the base for an affixed word; this is shown by the letter superscript (Packard 2000:165). The new word containing a word-forming affix may end up having a different word class from that of the base of the word (Packard 2000:70). The word-forming suffix *-zi* which is meaningless and the word-forming suffix *-zhe* ‘one who does/is X’ (Packard 2000:70) are

illustrated below:

(9) a. dao-zi

knife-ZI

‘knife’

b. du-zhe

read-one who does/is X

‘reader’

The grammatical affix is similar to the traditional inflectional affix (Packard 2000:71,165). Like the word-forming affix, the grammatical affix cannot be the base of a word in MSC (Packard 2000:166). But unlike the word-forming affix, the grammatical affix cannot function as a head of an MSC word; this is shown by the letter G (Packard 2000:165-166). In other words, the grammatical affix does not have the power to determine the word category of the resulting word (Packard 2000:165-166). Below the plural marker *-men* in MSC is demonstrated:

(10) ren-men

person-PLURAL

‘people’

In short, the Root Words ( $X^0$ ), the Bound Root ( $X^{-1}$ ), the word-forming affix ( $X^W$ ) and the grammatical affix (G) are primitives in Packard’s X-bar word structure for MSC nouns and verbs (Packard 2000:165). Among those primitives, only a Root Word can stand alone as an independent word, filling the syntactic word slot in a phrase (Packard 2000:166). Apart from the grammatical affix, the rest of the primitives, that is, the Root Word, the Bound Root and the word-forming affix can be the head of a MSC word (Packard 2000:166). Packard (2000:168) sets up two rules of word formation for MSC nouns and verbs along with the four primitives based on the universal schema of Selkirk (1982) shown in (1):

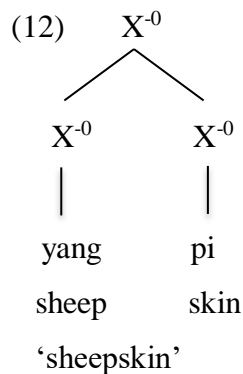
(11) Rule 1  $X^{-0} \rightarrow X^{-0-1\{w\}}, X^{-0-1\{w\}}$

Rule 2  $X^{-0} \rightarrow X^{-0}, G$

(Packard 2000:168)

According to Rule 1, a resulting word  $X^{-0}$  which is on the left side of the arrow is formed by two items which are on the right side of the arrow (Packard 2000:168).<sup>12</sup> These two items which are on the right side of the arrow can be both Root Words ( $X^{-0}$ ) or Bound Roots ( $X^{-1}$ ), or one of them is a Root Word ( $X^{-0}$ ) and the other is a Bound Root ( $X^{-1}$ ) (Packard 2000:168). The two items on the right side of the arrow cannot both be word-forming affixes  $X^W$ , which is designated by the curly brackets (Packard 2000:168). But for the Root Word and the Bound Root, there are no such restrictions (Packard 2000:168). That is to say, in a resulting word both components can be Root words or Bound Roots (Packard 2000:168). Words produced by Rule 1 are illustrated below.

By following Rule 1, a Root Word  $X^{-0}$  and a Root Word  $X^{-0}$  can form a word  $X^{-0}$ , that is,  $X^{-0} \rightarrow X^{-0}, X^{-0}$  (Packard 2000:170), which is shown in the following example *yang pi* ‘sheep skin: sheepskin’:



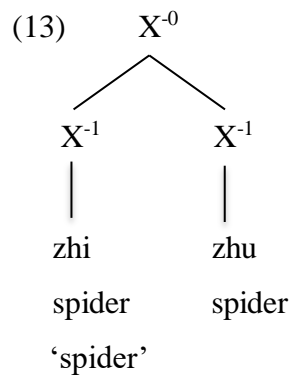
Also according to Rule 1, a Bound Root  $X^{-1}$  can merge with another Bound Root  $X^{-1}$  to form a resulting word  $X^{-0}$ , that is,  $X^{-0} \rightarrow X^{-1}, X^{-1}$  (Packard 2000:170). The example is *zhi*

---

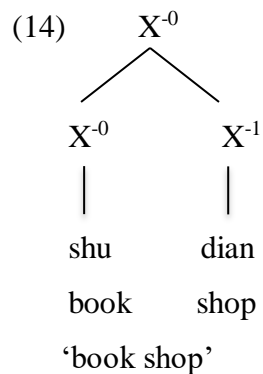
<sup>12</sup> Here the resulting word is a morphological word (Packard 2000:12). A morphological word and a syntactic word are represented differently in Packard's theory. A morphological word is a word which is the result of the application of morphological rules and it is represented as  $X^{-0}$  (Packard 2000:12,138). A syntactic word is a word which occupies a terminal node in a syntactic tree and it is represented as  $X^0$  (Packard 2000:12,138). A morphological word and a syntactic word have the same bar value in Packard's theory and this means that a morphological word can fill the word slot in a phrase (Packard 2000:138,152).

*zhu* ‘spider’, which is a synonymy-compound noun (see section 2.2.3 for a description).

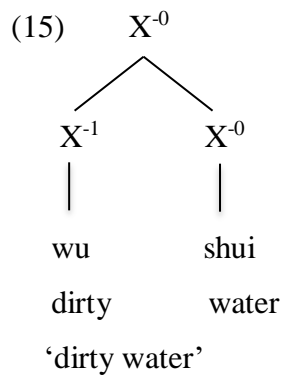
The structure would be this:



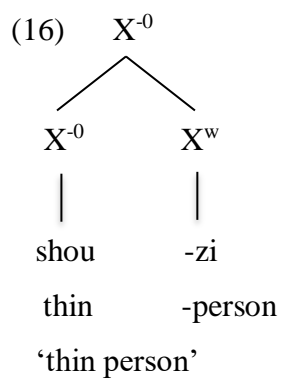
A derived word  $X^{-0}$  can also contain a Root Word  $X^{-0}$  and a Bound Root  $X^{-1}$  based on Rule 1 (Packard 2000:170). An example is *shu dian* ‘book shop’, which is an attributive compound noun (see section 2.2.2 for a description):



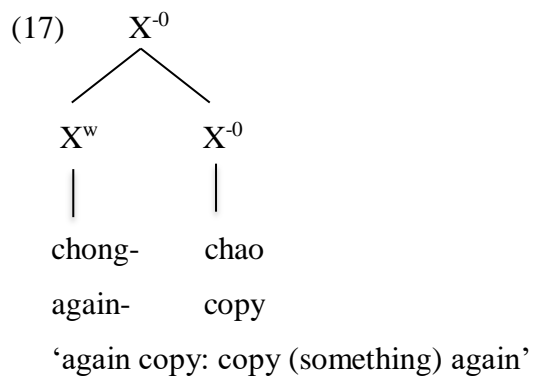
It is also possible that the first item of the resulting word is a Bound Root  $X^{-1}$  and the second item is a Root Word  $X^{-0}$ , according to Rule 1 (Packard 2000:170), as demonstrated in the example *wu shui* ‘dirty water’, which is an attributive compound noun (see section 2.2.2 for a description).



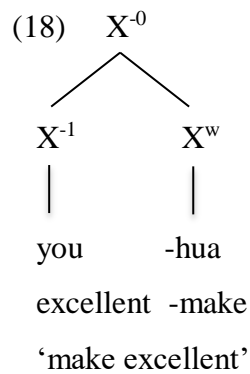
A Root Word  $X^0$  can also merge with a word-forming suffix  $X^w$ , that is,  $X^0 \rightarrow X^0, X^w$  based on Rule 1 (Packard 2000:170). This structure can be seen in the example *shou-zi* 'thin person' in which the suffix *-zi* is combined with the Root Word *shou* 'thin' to form a free word:



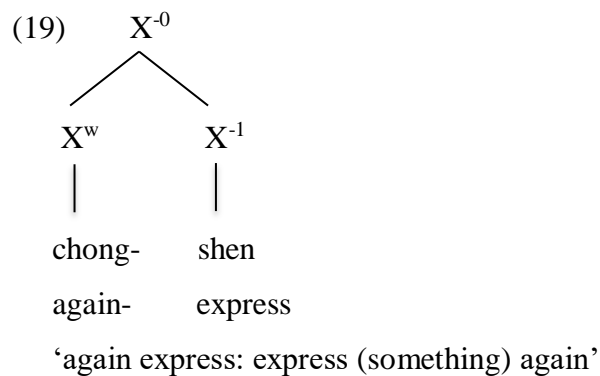
A prefix  $X^w$  and a Root Word  $X^0$  can also form a word  $X^0$ , that is,  $X^0 \rightarrow X^w, X^0$  according to Rule 1 (Packard 2000:170), which is demonstrated in the following example *chong- chao* 'again-copy: copy (something) again' which is composed of the prefix *chong-* 'again' and the Root Word *chao* 'copy':



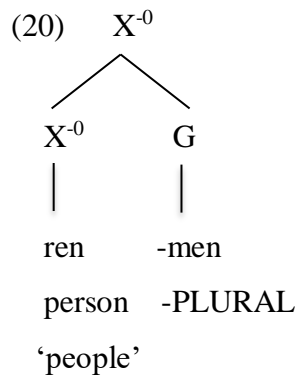
A Bound Root  $X^{-1}$  can also occur with a word-forming affix  $X^w$  according to Rule 1 (Packard 2000:170). The schema is  $X^{-0} \rightarrow X^{-1}, X^w$  (Packard 2000:170). The word below *you-hua* ‘make excellent’ shows such a word formation in which there is the Bound Root *you* ‘excellent’ and the suffix *-hua* ‘make’:



By following Rule 1, a prefix  $X^w$  can also accompany a Bound Root  $X^{-1}$  and the template is  $X^{-0} \rightarrow X^w, X^{-1}$  (Packard 2000:170). The example *chong-shen* ‘again- express: express again’ demonstrates this kind of word formation in which the prefix *chong-* ‘again’ forms a free word with a Bound Root *shen* ‘express’:



Rule 2 means that a Root Word  $X^{-0}$  and a grammatical affix  $G$  can be combined i.e.  $X^{-0} \rightarrow X^{-0}, G$  (Packard 2000:168), which is demonstrated in the example *ren-men* ‘people’ where the plural marker *-men* is the grammatical affix  $G$ :



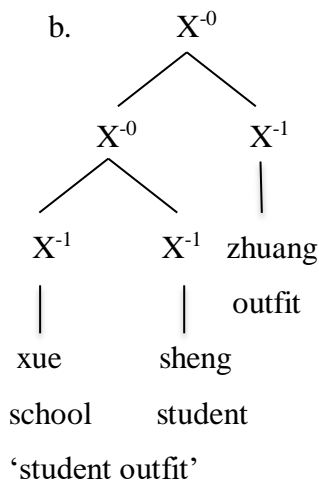
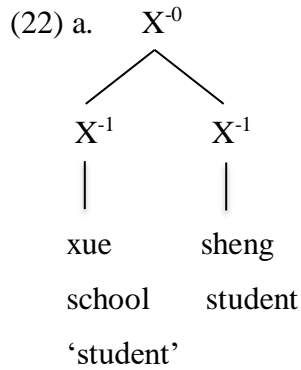
If the theory developed in chapter 4 is applied here, then this Rule 2 can be eliminated for the grammatical affix. More specifically, in chapter 4, it is proposed that Root and Bound Word are two lexical primitives for MSC word structure. A Root is always bound, never used alone as a free word. In other words, there is no such item as a Root Word. Based on this, it will not be possible to combine a Root with a grammatical affix, as the grammatical affix enters a derivation only at the phrasal level, and therefore merges with a word or a phrase, and forms a phrase (assuming X-bar theory, it can only merge with a phrase). A Bound Word cannot be used alone as a free word. Thus it cannot merge, directly, with the grammatical affix either. This means that there is no Rule 2 forming words in MSC. Note here the word structure developed in this thesis in chapter 4 will only concern the derivational morphology of MSC, which means that the grammatical affix will not be considered.

In addition to Rule 1 and Rule 2, recall that Packard (2000:165) also allows a Root Word to be a free word on its own. This indicates that Selkirk's (1982:95) rule  $X^0 \rightarrow X^r$  is still valid in Packard's X-bar theory of MSC. For instance, a word *wan* 'bowl' which is formed of a Root Word *wan* 'bowl' can be used alone as a word in the following sentence:

- (21) zhe ge wan zhen xiao.  
 this CLASSIFIER bowl really small  
 'This bowl is really small.'

Furthermore, Packard (2000:168-169) points out that recursion is also possible in word

formation in MSC, as shown in the following recursive compound *xue sheng zhuang* ‘student outfit’ in which there is a parallel compound *xue sheng* ‘student’ merged with the Bound Root *zhuang* ‘outfit’:



It can be seen in the above example that the recursive node is  $X^{-0}$ , which means it is the free word that is being recursively expanded. This is always the case in MSC (according to Packard 2000:167-168). The other nodes  $X^{-1}$ ,  $X^W$  and  $G$  cannot be recursive nodes (Packard 2000:167-168).

Compared with Dai (1992) and Sproat and Shih (1996), Packard has made improvements to the theory of word structure of MSC and a detailed classification of word components of MSC is offered. Nevertheless, some aspects of the typology of word components of MSC are still not satisfactory. For one thing, the distinction between Root and word is still not clear. It relies on a set of stipulated word structure rules. In particular, Packard (2000) agrees with Selkirk’s rule (2) by admitting that a Root can be a free word. Then he



goes even further to claim that the free Root is a Root Word which has a same bar value as that of a free word, that is,  $X^{-0}$ . Moreover, it is unclear what the relationship is between a Root Word and a Bound Root and there is not a specific rule to indicate such a relationship.

If, on the other hand, a Root has no categorial feature it follows that it cannot occur as a free word, but must be combined with an item with a categorial feature. There would be no rule such as Selkirk's rule (2). It also follows that the notion 'Bound Root' does not make sense, since all Roots are bound. The items called Bound Roots in Sproat and Shih (1996) and Packard (2000) cannot be Roots, as they can function as the head of words, hence must have word category. Instead they are words that must be bound, hence I will refer to them as Bound Words. Like functional categories they are not composed of a Root and a categorizer, but have a category inherently. They are semi-functional categories, in that way.

In addition to compounds with transparent meanings and predictable word category, there are also MSC compounds which have non-compositional meanings and whose syntactic categories cannot be obtained from their constituents. They are exocentric compounds, therefore (Zhang 2007:172). In the literature, exocentric compounds refer to compounds which lack a head in terms of category, semantics or morphology (Bloomfield 1933:235-236). More precisely, in an exocentric compound [A, B] neither A nor B is the head (Zhang 2007). Instead, as will be discussed, when it occurs as a free word, the exocentric compound has a head which is external to [A,B] (Zhang 2007). Such compounds can be a noun, an adjective or a verb, which are demonstrated below:

(23) a. mei      chou  
         beautiful ugly  
         'appearance'

      b. chi    cu  
         eat    vinegar  
         'jealous'

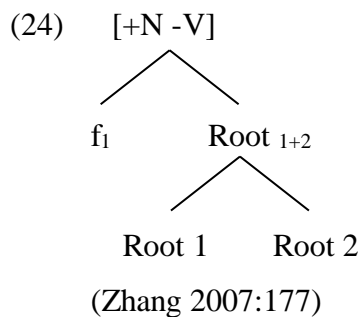
- c. qiang diao
- strong accent
- ‘emphasize’

Exocentricity is branded as an abnormality in compounds of MSC by Zhang (2007), and she claims the Root Merger analysis can account for it, which is introduced now.

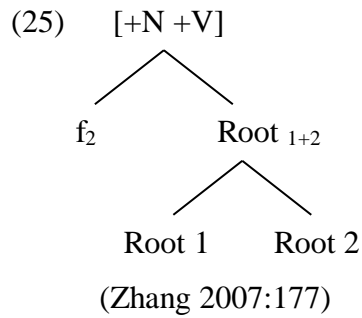
### 3.5 Zhang (2007)

There are abnormal compounds in MSC (Zhang 2007). Zhang (2007:171,177) follows the works of Distributed Morphology and assumes that Roots are without syntactic features and according to her, this assumption can account for these abnormal MSC compounds.

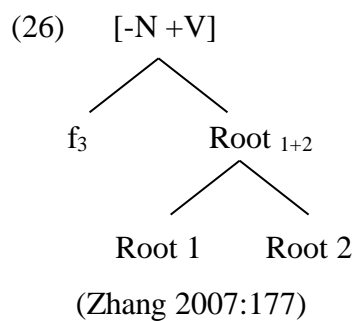
More specifically, according to Zhang (2007:177), the abnormality in MSC compounds can be explained if those compounds are result of the merge of two category-less Roots and a category-defining functional head. The function of the category-defining functional head is to assign word category to the compound (Zhang 2007:177). As for compounds which are nouns, their structures are as follows, where  $f_1$  represents the nominal functional head (Zhang 2007:177):



For compounds which are adjectives, they may have the following structure where  $f_2$  is the adjectival functional head (Zhang 2007:177):



The structure of compounds which are verbs is shown below in which the verbal functional head is represented as  $f_3$  (Zhang 2007:177):



In MSC exocentric compounds, there is a discrepancy between the word category of their constituents and the word category of the compounds themselves, and this is an abnormality according to Zhang (2007). This abnormality can be accounted for if the constituents in the exocentric compounds are Roots without syntactic features (Zhang 2007). This means that when two such constituents merge, there will be no syntactic features, and in particular no categorial features projected from those two constituents to the resultant compound (Zhang 2007:177). Instead, the category of the compound is the result of merge of a category-less Root compound with a category-defining functional head (Zhang 2007:178). Thus the discrepancy that occurs between the category of a compound and the category of its components can be avoided (Zhang 2007:178). For instance, the compound *mei chou* ‘beautiful ugly: appearance’, where both constituents are property-denoting items, and yet the compound is a noun, will have a structure where the category is determined by the third element, that is, the nominal category-defining functional head. In other words, *mei chou* ‘beautiful ugly: appearance’ will have structure (24).

Roots without word category may also explain another abnormality in MSC, which is free projectivity (Zhang 2007:178). Free projectivity is observed in some MSC compounds. For example, the verb-object compound can be a verb, as seen in the case *zhi dao* ‘know road: know’, which is a verb, but the verb-object compound can also be a noun, as seen in the case *zhi ji* ‘know self: close friend’, which is a noun (Zhang 2007:173). The issue of free projectivity can be understood if constituents of such compounds are Roots which are devoid of syntactic features (Zhang 2007:178). Since constituents of such a compound are without syntactic features, there will be no features projecting from the constituents to the resultant compound and the category of the resultant compound will not be determined by that of its constituents (Zhang 2007:178). Instead, whatever word category the compound has is determined by the third element, the functional head, and so the compound may have any word category (Zhang 2007:178). So *zhi dao* ‘know road: know’, which is a verb, has the structure (26), while *zhi ji* ‘know myself: close friend’, which is a noun will have the structure (24). In the cases above, *mei chou* ‘beautiful ugly: appearance’, *zhi dao* ‘know road: know’ and *zhi ji* ‘know myself: close friend’, the meanings of the compounds are not derived compositionally from the constituents, but come directly from the lexicon (see Bauke 2014 on Root-Root compounds in German).

It may be true for compounds in which the word category cannot be derived from their components that a third item is the categorial head, as Zhang (2007) proposes. But for the majority of compounds in MSC, which have transparent word category and meaning, I still think they are best understood in the theory developed in chapter 4.

### 3.6 Summary

In brief, Selkirk and Packard have demonstrated that the syntactic X-bar theory can account for word structure in English and MSC. A set of primitives and rules are proposed for the structure of words in both languages. Selkirk’s rule (2), that is,  $X^0 \rightarrow X^r$  where a Root is the only component of a word is recognized, at least in effect, by Sproat and Shih (1996) and Packard (2000). By admitting this rule, a Root can be a word, but in any other cases, the Root is bound, and called a Bound Root (Packard 2000; Selkirk 1982; Sproat and Shih 1996). Thus the distinction between a Root and a word entirely relies on

Selkirk's rule (2). But this rule can be eliminated, which means the system will become simpler, if the derivation of a simple lexical word in MSC is a result of the operation Merge, merging a category-less Root and a categorizer (Chomsky 1995:243-245, 2015:223-227; Embick and Noyer 2007,2008; Harley and Noyer 1999; Marantz 1997, 2001). As a result, the Root-Word distinction will be seen in a different light. One consequence is that the so-called Bound Root may be ruled out as a primitive for the structure of a word in MSC, as will be discussed in the next chapter.



## Chapter 4 Modern standard Chinese word formation: proposed theory

This chapter concerns derivational morphology and compounding in MSC. Particularly, a set of primitives for structures of MSC words will be proposed and discussed, followed by discussion of the internal structure of MSC words, including their head positions.

### 4.1 Primitives of the structure of MSC words

The primitives of Packard's X-bar word system for structures of MSC words, as previously introduced, are the Root word, the Bound Root, the word-forming affix and the grammatical affix (Packard 2000:165). Packard (2000:168) proposes Rule 1 (i.e.  $X^{-0} \rightarrow X^{-0-1\{w\}}, X^{-0-1\{w\}}$ ). By following this rule, for instance, a Root which is bound (i.e.  $X^{-1}$ ) forms an actual free word (i.e.  $X^{-0}$ ) with other types of primitives. But Packard (2000:165) also recognizes that a Root word (i.e.  $X^{-0}$ ) can be an actual free word (i.e.  $X^{-0}$ ), which means a word may be made up of one single Root. So, Selkirk's (1982:95) rule  $X \rightarrow X^r$  is still assumed in Packard's theory and there is minimally one item in a MSC word according to his theory. In Packard's theory, as previously discussed, the relation between the Root Word and the Bound Root remains unclear.

Instead of following the X-bar theory of word structure of Packard's, in this chapter, I follow Distributed morphology and propose that words in MSC are derived by the operation Merge, which always involves two items (Chomsky 1995:243, 2015:223). In other words, only binary structures are allowed by this operation and unary structure is not possible. Thus Selkirk's (1982:95) rule  $X \rightarrow X^r$ , which is also assumed in Packard's theory, will not be a possible rule for MSC word structure and will be eliminated. It follows that a Root alone cannot form a Word. This means the current theory will in this sense be simpler than Packard's theory.

The nature of the Root has been discussed in the literature. After observing the Swedish simple compounds, for example, *skolflicka* 'school girl' and derived suffixed words, for example, *skolning* 'schooling', Holmberg (1992:28-29,33) concluded that the item in the

non-head position in the compounds or the suffixed words is a Root. This finding indicates that the Root is a maximal category; in other words, it does not project (Holmberg 1992:38). Holmberg's generalization on the nature of the Root seems to be in line with the definition of the Root proposed in Distributed morphology (Embick and Noyer 2007; Harley and Noyer 1999; Marantz 2001).

In Distributed Morphology, morphemes are defined as the terminal nodes of syntactic trees, and Roots are defined as lexical morphemes with phonological features and semantic content (Embick and Noyer 2007:295). They do not possess grammatical features, and therefore always merge with categorizing heads, for example, *v*, *a*, or *n*, to form the lexical categories Verbs, Adjectives and Nouns (Embick and Noyer 2007:296; Harley and Noyer 1999:4; Marantz 2001:10). Therefore, the Root cannot be the head of a word, and in a lexical category there is minimally a Root and a categorizing head (Embick and Noyer 2007:296). For instance, for a noun *apple*, it will contain the Root  $\sqrt{\text{APPLE}}$  and the categorizing head *n*, and has the structure  $[_n \sqrt{\text{APPLE}}, n]$ . Here, the categorizing head has a null exponent (Embick and Noyer 2008:6; Harley 2011:132). Alternatively, the categorizing head can have an overt exponent (Harley 2011:132; Marantz 2001:13). For example, in the word *painful*, there is the Root  $\sqrt{\text{PAIN}}$  and the categorizing head *a* which has the overt exponent *-ful*; the structure of the adjective *painful* is  $[_a \sqrt{\text{PAIN}}, a]$  (Harley 2011:132; Marantz 2001:13). See also Josefsson (1997) who came to essentially the same conclusion based on Swedish data.

I will follow Holmberg (1992), Josefsson (1997), Harley and Noyer (1999), Marantz (2001), Embick and Noyer (2007, 2008) and Harley (2011) and assume that the Root in MSC is always bound and does not bear the word category.<sup>13</sup> In addition to the notion of Root discussed here, the theory also adopts the operation Merge (Chomsky 1995:243, 2015:223), so the following condition is assumed in MSC:

(1) There is always a minimum of two morphemes in a MSC free content word.

---

<sup>13</sup> Following the convention used in Distributed morphology, a Root in MSC is represented as the combination of ' $\sqrt{\text{ }}$ ' and the *pinyin* in capitals. For instance,  $\sqrt{\text{SHU}}$  represents a Root in MSC. The same convention will be used to represent Roots in OXC.



As the Root lacks a category, there must be a morpheme which accompanies the Root and provides the category of the word containing the Root, and together they can form an actual free word of MSC. I will propose that, in the default case, this is a morpheme  $\emptyset$ , which has a null exponent and is meaningless. The grammatical function of this null morpheme is to make a free word out of a Root. Such a morpheme was discussed under the name of word marker in the work of Harris (1991) and Holmberg (1992), as we will now see.

#### **4.1.1 Harris (1991): Spanish**

Harris (1991:27-28) observes that the suffixes *-a* and *-o* correspond to gender in a large class of nouns in Spanish:

- (2) nieta ‘granddaughter’  
    nieto ‘grandson’  
    (Harris 1991:27)

Here *-a* here marks feminine and *-o* marks masculine. But it is not always this clear-cut. The suffixes *-a* and *-o* can be neutral between masculine and feminine for some nouns (Harris 1991:38-39):

- (3) a. colega ‘colleague’  
    (Harris 1991:38)
- b. testigo ‘witness’  
    (Harris 1991:39)

So, it seems that the correlation among the suffixes *-a* and *-o* and the gender features in nouns in Spanish is rather arbitrary and redundant (Harris 1991:36,39). This kind of redundancy of *-a* and *-o* can also be seen in Spanish adverbs which do not have grammatical gender and yet are marked with *-a* and *-o* regardless (Harris 1991:33):

(4) fuera ‘outside’

dentro ‘inside’

(Harris 1991:33-34)

Clearly, the suffixes *-a* and *-o* do not indicate gender here (Harris 1991:34). This can also be seen in Spanish adjectives, in which the suffixes *-a* and *-o* are applied for the purpose of matching the adjectives to nouns with respect to gender (Harris 1991:34):

(5) a. Mi madre (feminine) es alta

my mother is tall

‘My mother is tall’

(Harris 1991:34)

b. Mi padre (masculine) es alto

my father is tall

‘My father is tall.’

(Harris 1991:34)

On the surface, the suffixes *-a* and *-o* seem to indicate the gender of the adjectives used above, where *-a* indicates female gender and *-o* indicates male gender and. But, in fact, the gender of adjective corresponds to the noun it modifies (Harris 1991:34). More specifically, in (5a), the gender of *alta* ‘tall’ matches the gender of *madre* ‘mother’; in (5b), the gender of *alto* ‘tall’ matches the gender of *padre* ‘father’ (Harris 1991: 34). It can be seen that the words in various word classes with the suffix *-a* and *-o* are not systematically restricted as regards gender (Harris 1991:59). Thus, the suffix *-a* and *-o* cannot be gender-marking suffixes, rather they are word markers whose primary grammatical function is to make a complete word out of a Root or a stem (Harris 1991:28, 30-31, 59).

#### 4.1.2 Holmberg (1992): Swedish

There is a large class of nouns in Swedish made up of a Root and a vowel *-a* and *-e*. The suffixes are related to sex (natural gender) for some nouns, the suffix *-a* often referring to females and *-e* often referring to males (Holmberg 1992:36):

(6) *majsa* ‘girl’

*kille* ‘boy’/‘man’

(Holmberg 1992:36)

But the majority of nouns ending in the suffixes *-a* and *-e* are sexless objects (Holmberg 1992:36):

(7) *matta* ‘mat’

*galge* ‘hanger’

(Holmberg 1992:36)

Therefore, the suffix *-a* and *-e* are not direct markers for sex in Swedish (Holmberg 1992:36). Following Harris (1991), Holmberg (1992:36-37) classifies the Swedish suffix *-a* and *-e* as word markers (WM), and which encode number, specifically, [-PLURAL]. Their primary grammatical function is to form a word with a Root (Holmberg 1992:37):

(8) *flick-a*

girl WM

‘girl’

(Holmberg 1992:37)

For other Swedish nouns, which do not have the overt suffix *-a* and *-e*, Holmberg (1992:46) proposes that there is also a word marker present in the structure, which is null (represented as  $\emptyset$ ) and similarly encodes number, specifically [-PLURAL]:

(9) bild- Ø

picture WM

‘picture’

(Holmberg 1992:46)

#### ***4.1.3 Modern Standard Chinese***

In the preceding section, the word markers in Spanish and Swedish are introduced and are shown to be able to complete an actual free word by merging with a Root. Further, those word markers sit in word-peripheral position. The null morpheme in MSC word structure could be classified as a word marker, as it makes a free word out of a Root. But another important grammatical function of this null morpheme in MSC is to provide the categorial feature for the resultant word when the word is formed of a category-less Root and a null morpheme. Thus I will follow Harley and Noyer (1999), Marantz (2001), Embick and Noyer (2007,2008) and Harley (2011) and assume that this null morpheme has a categorial feature and so it can provide the categorial feature for the free word which is formed of a Root and the null morpheme. In other words, the null morpheme is classified as a categorizing head, that is, a null categorizer in the present work.

Further, following Packard, I will assume that the word-forming/derivational affix, which is either a suffix or a prefix, is a primitive in MSC word structure. As will be seen in section 4.4, some of the word-forming/derivational affixes in MSC have a similar grammatical function as that of the null morpheme in the sense that they merge with a Root and provide the categorial feature for the resultant word. In other words, these kinds of word-forming /derivational affixes are categorizers. On the other hand, there are also word-forming/derivational affixes which do not have such grammatical functions but are grammatically redundant, which will also be seen in section 4.4. In addition to word-forming/derivation affixes, grammatical affixes such as the plural suffix *-men*, are also primitives for the structure of MSC words according to Packard (2000:165). But in this chapter, I consider them to be part of phrasal syntax, so they will not play a part here, as this chapter only concerns derivational morphology and compounding in MSC.

The null categorizer (represented as  $\emptyset$ ) and the Root (represented as R) together form a Minimal Word (represented as MW) (i.e.  $MW \rightarrow [R, \emptyset]$ ) in MSC. A Minimal Word can be an actual free word on its own or it can merge with a Root or an affix to form an actual free word, where the word category of the resultant word is obtainable from its components. The Root and the Minimal Word may share the phonological spell-out form, but there are differences too. More specifically, in contrast with the Root, the Minimal Word cannot merge with the null categorizer. That this is not allowed may be due to the theoretical reason that superfluous items are not tolerated in the grammar of natural language (Chomsky 2015: 24, 118,137). Later, in chapter 5, reduplication of common nouns in OXC will be shown to provide the empirical evidence for this claim. Furthermore, a Minimal Word can merge with a Root forming a compound with the Root as a non-head, as is discussed now.

It has long been an accepted idea that the head in a compound word is identifiable based on the notion of feature percolation, also called feature projection (Holmberg 1992; Lieber 1980; Scalise and Fábregas 2010). Percolation (or projection) is a mechanism through which linguistic features of the head item can be transmitted to the dominating word (in morphology/word syntax) or dominating phrase (in phrasal syntax) (Holmberg 1992; Lieber 1980, 1992; Scalise and Fábregas 2010). The principle of feature percolation is that if the linguistic features of a node  $\alpha$  percolate to the node  $\beta$  which immediately dominates  $\alpha$ , then  $\alpha$  is the head of  $\beta$  (Holmberg 1992:38). On the other hand, if the linguistic features of  $\alpha$  do not percolate to  $\beta$ , then  $\alpha$  is the non-head component of  $\beta$  (Holmberg 1992:38). Linguistic features may include categorial features, semantic features and morphological features (Di Sciullo and Williams 1987:24; Liao 2014:11; Scalise and Fábregas 2010:112). I assume, following Chomsky (1995: 243-246, 2015: 223-227) and subsequent work in Minimalist syntax and works in Distributed morphology that the mechanism for constructing syntactic structure, both at word and phrase level, is binary Merge: Merge  $\alpha$  and  $\beta$  forming  $\gamma$ , where  $\gamma$  is the set  $\{\alpha, \beta\}$  and a label which is = either  $\alpha$  or  $\beta$ , depending on which one is the head, i.e. which one's features project onto  $\gamma$  (Chomsky 1995: 243-246, 2015:223-227). The question is what determines whether  $\alpha$  or  $\beta$  is the head. Unlike earlier work in morphology (Di Sciullo and Williams 1987; Packard 2000; Selkirk 1982; Williams 1981), I will not assume that the linear order of  $\alpha$  and  $\beta$  determines headedness. Instead, linear order is a consequence of

headedness, according to language particular rules (Chomsky 2015:308). What determines headedness is feature percolation/ projection (Chomsky 1995:243-246, 2015:223-225; Holmberg 1992:38)

In some cases this is straightforward. In particular, if  $\alpha$  is a Root and  $\beta$  is, say, a Minimal Word, then  $\beta$  is the head. A Root has no categorial features, and therefore cannot be head, so the categorial features, and thereby, I assume, the semantic features, of the Minimal Word percolate to the dominating node, i.e. the resultant word. A default rule for percolation of semantic features is assumed here:

(10) The semantic features of  $\alpha$  percolate along with the categorial feature, except in the case where  $\alpha$  has no semantic features. In that case, the semantic features of  $\beta$  percolate.

This default rule will be applied in nouns, adjectives and verbs in both MSC and OXC. In other cases, it is not so clear what determines whether  $\alpha$  or  $\beta$  is the head. To be specific, a compound consisting of two words poses a potential problem as either word can be the head. Unless some mechanism is present which prevents it, the result is a double-headed word. For synonymy-compound nouns/adjectives which are formed of two Minimal Words, this could be the case. But a condition is assumed in this chapter that the components in synonymy-compound nouns/adjectives and parallel compound nouns/adjectives cannot be Minimal Words. By assuming this, two Minimal Words cannot merge to form synonymy-compound nouns/adjectives in MSC. This assumption will be discussed more in section 5.5.3 in chapter 5. It will be seen that this condition is empirically supported by reduplication of nouns in OXC. The structure of synonymy-compound nouns and synonymy compound adjectives in MSC will be discussed in section 4.5.2.

But for single-headed compounds formed by two words, a mechanism must exist which blocks feature percolation from one of them. The head location in words formed by two words has been discussed in the literature. Holmberg (1992) proposes that the Case-marking prevents feature percolation from the marked word. Mukai (2008) follows this line of argument and assumes that the non-head in a recursive compound can be distinguished from the head by a linking morpheme, and this linking morpheme may be

covert. Below Holmberg's observations will be shown first, followed by Mukai's notion of linking morpheme.

In Swedish, Holmberg (1992:29) observes that some nouns are composed of Roots, shown by their absence of a final vowel and a suffix, as discussed in the previous section.

(11) a. flick-a

girl-WM

'girl'

(Holmberg 1992:28)

b. skol-a

school-WM

'school'

(Holmberg 1992:28)

Here, in (11a) and (11b), *flick* 'girl', *skol* 'school' are Roots, merged with the word marker *-a* to form nouns. Also it is observed that the nominal compounds which are non-heads in complex compound nouns in Swedish are marked by a suffix *-s*, analyzed as a Structural Case marker by Holmberg (1992:30, 34).

(12) skol-flick-s-vän

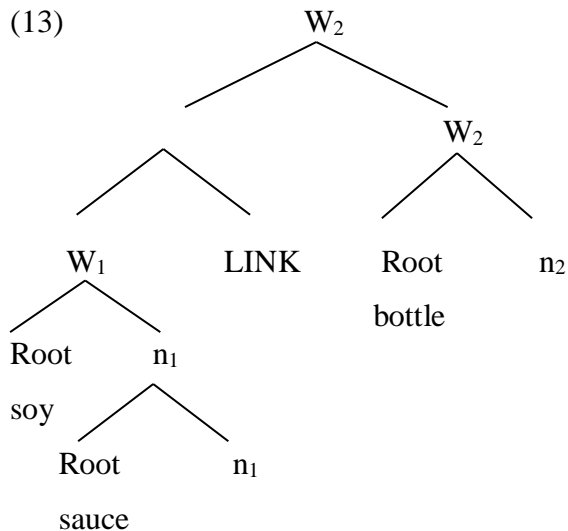
school-girl-CASE-friend

'school girl friend, friend of a school girl'

(Adapted from Holmberg 1992:28,30)

Here, *skol* 'school' and *flick* 'girl' are Roots and *vän* 'friend' is a word. *Skol-flick* 'school girl' which is the non-head is marked by *-s* in (12) and the head is the word *vän* 'friend'. In other words, the branching non-head here is distinguished from the head by Case marking (Holmberg 1992:33-34). More specifically, the Case marking prevents feature percolation from the marked noun (Holmberg 1992:41). Following this line of argument, Mukai (2008:187-188) argues that there needs to be a linking morpheme inside left-branching recursive compounds to check features on the non-head which is a nominal

compound, and thereby allows the features of the head to percolate. In some languages, such as Swedish, the linking morpheme is overt, but in other languages, such as English, the linking morpheme is covert (Mukai 2008:187-188). For instance, an English left-branching recursive compound *soy sauce bottle* may have the following structure:



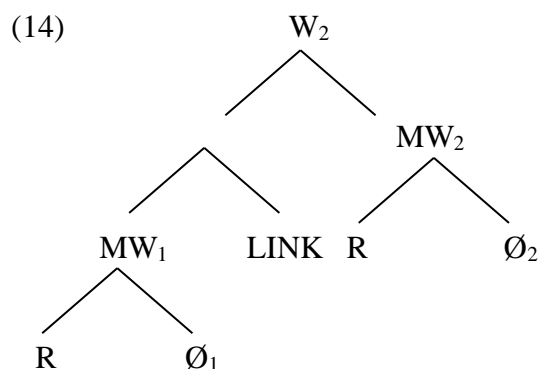
(Adapted from Mukai 2008:187)

The function of this linking morpheme (represented here as LINK) is to check features, including the categorial *n*-feature<sup>14</sup> on the inner nominal compound *soy sauce* of the left-branching recursive compound *soy sauce bottle* (Mukai 2008:187-188). So, the features, including the *n*-feature, on the inner nominal compound *soy sauce* will not project to the recursive compound *soy sauce bottle* but instead the features of the noun *bottle*, including the *n*-feature, which are unchecked, will percolate to the recursive compound *soy sauce bottle* and so the noun *bottle* is the head of the resultant compound (Mukai 2008:187-188). It can also be seen from this English example that the linking element is assumed here even though it is not pronounced (Mukai 2008:188). On the other hand, there are languages that do not have such a linking morpheme, and therefore do not have left-branching recursive compounds, such as Latin, and other Romance languages and Hebrew (Mukai 2008:194).

<sup>14</sup> In Mukai's (2008:181) theory, the linking morpheme also checks what she calls a theta-role feature, based on the idea that all lexical nouns can assign a theta-role and the D (determiner) which heads DP is a binder of the theta-role discharged from the lexical nouns (following an idea by Higginbotham 1985). I will ignore this feature, because it makes no difference to the theory here.



Following Holmberg (1992) and Mukai (2008), I assume that there is an abstract linking morpheme in MSC for the single-headed compounds which are composed of two words (rather than a Root and a word). It must be abstract because there is no overt linking morpheme observed in MSC. A single-headed MSC compound which is composed of two Minimal Words may thus have the following structure:



The function of the abstract linking morpheme is to check the features of one of the Minimal Words, so the other Minimal Word can percolate and be the head in terms of category and semantics. Here, we assume the Minimal Word (MW<sub>2</sub>) which is on the right is the head, so the abstract linking morpheme will check the features of the left-hand Minimal Word (MW<sub>1</sub>). The abstract linking morpheme seems to have solved the issues with the single-headed compound words in MSC consisting of two Minimal Words. However, there is always a simpler solution: merging a Root and a Minimal Word. Roots and Minimal Words are phonologically indistinguishable in MSC, but behave quite differently in word formation. Because a Root has no categorial features, the categorial features of the Minimal Word will always percolate to the dominating node, then following the default rule (10), the semantic features will also percolate from the Minimal Word to the dominating node. The Minimal Word will thus be interpreted as the head. This solution uses fewer features, categories (no linker), and operations, so it will always be preferred (by the grammar) when there is a choice. Following this, in this chapter, it will be assumed to be true that a Minimal Word cannot be the non-head in terms of category and semantics in a compound. The discussion here is in relation to compounds. However, I will assume that it holds more generally. In particular, it also applies to affixed words. This hypothesis will be discussed more in chapter 5 where OXC

reduplicated common nouns are taken into account and will be shown to support this assumption empirically. By contrast, for the left-branching recursive compound nouns in MSC, where the non-heads are compounds, and therefore must be words, not Roots, the presence of the abstract linking morpheme is needed to prevent feature percolation from the word to the dominating node. Since MSC has left-branching recursive compounds, it must have the linking morpheme, following Mukai (2008). This will be expanded more in section 5.5.5 of the next chapter.

A Minimal Word can also merge with another primitive, which is labelled Bound Word in the present work. The definition and features of Bound Words are introduced now.

As discussed in section 3.4, a Bound Root according to Packard's theory of the word structure of MSC is a Root which cannot stand alone. I will keep the insight that there is a class of morphemes that look like Roots but which always merge with another morpheme, or word in MSC. I will, however, change the label to Bound Word. The notion 'Bound Root' does not make sense as a class in the theory developed so far, as all Roots are bound in the current theory. The Bound Word is an item which is inherently categorial (for instance nominal). In other words, Bound Words have a category and a semantics. However, unlike other content words, Bound Words do not have internal structure. As a Bound Word has a word category, it can be a categorial head. But it can also be a non-head in terms of category, as it will be shown in section 4.5.1 that two Bound Words can form single-headed attributive compound nouns. Therefore, it is assumed that the Bound Word projects optionally. The logic behind this assumption is that the features from the categorizer percolate unless there is a mechanism to stop it, but the features of the inherently categorial Bound Word percolate optionally. The categorial feature of the Bound Word will project when it is necessary. Later in chapter 5, the reduplication of common nouns and attributive compound nouns in OXC will provide the empirical evidence for this assumption.

A Bound Word in MSC can merge with any item which is overt. Together they form an actual free word. This overt item may be a Root, an affix, a Minimal Word, or it may be another Bound Word. A Bound Word can merge with a Minimal Word only when the Minimal Word is the head. A Bound Word may merge with a Root, in which case the

Bound Word is the head. This could be the case for synonymy-compounds/parallel compounds and attributive compound nouns in MSC. Here, in those types of compounds, if the Bound Word is allowed to project its categorial features, and based on the default rule (10) of percolation of semantic features which states that semantic features and categorial features percolate together, then semantic features of the Bound Word will also percolate. This means that the Bound Word in the synonymy-compound/parallel compound and the Bound Word in the attributive compound noun would be the heads in terms of category and semantics. This, in turn, would mean that the synonymy-compound/parallel compound and the attributive compound noun are single-headed and have the same structure. It would follow that synonymy-compounds/parallel compounds are attributive compounds, which is clearly not true. Thus an assumption is made here that in synonymy-compounds and parallel compounds, categorial features of Bound Words do not project to the resultant words, so they cannot be categorial heads. Consequently, the default rule (10) concerning semantic feature percolation does not apply to those two types of compounds in MSC. Thus for synonymy-compounds and parallel compounds which are formed of two Bound Words, none of the Bound Words is the categorial head. Recall it is assumed that synonymy-compounds/parallel compounds cannot be formed of Minimal Words either. So both types of compounds cannot be double-headed. More details on synonymy-compounds/parallel compounds in MSC will be offered in section 4.5.2. For attributive compound nouns, which are considered to be single-headed in the literature, categorial features of the Bound Word are allowed to project when it is the head.

On the other hand, like the Minimal Word, the Bound Word cannot merge with a null categorizer as it has inherent word category. This will be an assumption here, which will be supported by empirical evidence from reduplicated common nouns in OXC in chapter 5. Because of this incompatibility between the Bound Word and the null categorizer, the Bound Word cannot meet condition (1) stated in this chapter which says that there are minimally two items in a free content word of MSC. Thus the Bound Word, in spite of having semantic and categorial features must be bound and cannot occur as an actual free word on its own. Also in this respect concerning the null categorizer, the Bound Word is quite different from the Root, as the Root can merge with the null categorizer.

As for the Root, in addition to the null categorizer, it can also merge with other types of word items to form an actual free word. The items are Minimal Words, Bound Words, affixes and compounds. It would seem that that it cannot merge with another Root, since the Root is devoid of word category, hence the resultant word containing only two Roots would have no category. However, later it will be seen that following Zhang (2007), two merged Roots and a categorizer will account for the structure of antonymy-compounds, verb-object compounds and subject-predicate compounds in MSC and OXC. The function of the categorizer is to assign the word category to the resultant words (Zhang 2007).

It can be seen then, the Bound Word and the Root in this developed theory of word structure may have similarities. Firstly, they both bear lexical meanings. And secondly, they are both bound; in other words, neither of them can be used alone in a MSC word but need to merge with another word item in order to form an actual free word, and those items are the Minimal Word, the Bound Word, the affix and the compound.

On the other hand, there are also differences between the Bound Word and the Root concerning the word items they can merge with. For a Bound Word in MSC, in order to make a free content word, it has to merge with an item, which must be overt, that is, it must have phonological features. That is to say, a Bound Word cannot merge with a null categorizer. But for a Root, in order to form a free content word, the word item it is merged with can be without phonological features. In other words, a Root can merge with a null categorizer.

Neither the Bound Word nor the Minimal Word can merge with the null categorizer. Also both the Bound Word and the Minimal Word bear word category. In addition, later in section 4.4, it will be discussed that the affix cannot merge with a null categorizer to form a free content word either. Thus, only a Root can merge with a null categorizer to form a free content word. So, the following condition is proposed concerning the null categorizer:

(15) The null categorizer can only merge with a Root.

This condition is also true in OXC words as will be discussed in chapter 5. The difference between a Bound Word and a Minimal Word is that the Bound Word has word category inherently. In a Minimal Word, the word category comes from the null categorizer. Another difference is that the Bound Word cannot occur alone as a free word in a phrase, but the Minimal Word can do so.

Summarizing, the Root has no categorial features and therefore cannot be a free word or be the head of a word. The Minimal Word is a Root plus a null categorizer. The Root and the Minimal Word are phonologically equivalent in MSC (but not in all other languages), but structurally different. Bound Words are like Roots in that they cannot form a free word on their own. They are different from Minimal Words in that they are not made up of a Root and a categorizer, but combine semantic and categorial features in one item.

In the following table, the items which can or cannot merge with a Root, a Bound Word, a Minimal Word and the null categorizer are illustrated. If item A is allowed to merge with the item B to form a free content word, the corresponding box is ticked with '✓', otherwise it is left blank:

Item A \ Item B	Root	Bound Word	Minimal Word	Null categorizer
Root	√ (for the situation where category of the resulting word is unobtainable from its components)	√	√	√
Bound Word	√	√	√ (for the situation where the Minimal Word is the head)	
Minimal Word	√	√ (for the situation where the Minimal Word is the head)		
Null categorizer	√			
Affix	√	√	√	

**Table 1 Possible combinations of MSC words**

So far, it has been established that a Root does not bear word category, but the other types of items in a MSC word, that is, the Bound Word, the null categorizer and the affix may all bear word category. So, all items apart from the Root, are candidates for the categorial head in MSC words. In addition to the categorial head, there are other types of heads in MSC words, which are discussed now.

## 4.2 Different types of heads

Headedness in a word has been a matter of debate in the literature. Williams (1981), Lieber (1980) and Selkirk (1982) all express their views regarding the head location in a word, as will be discussed first. Then I will look into more details about the head position in MSC compounds in particular. More specifically, the works of Packard (2000), Scalise and Fábregas (2010) and Liao (2014) will be discussed.

According to the Righthand Head Rule (RHR), the head is the item which is on the right-hand side in a word (Williams 1981:248). In other words, the linear position is the key factor for determining the head. There are some English words which have their heads on the left, though, as Williams (1981:249) also notices, for instance words with the prefix *en-*. He treats such words with the left-hand member as the head as exceptions to RHR (Williams 1981:249-250). However, left-hand heads of words might not be an exception in other languages. Lieber (1980:100) points out that compounds in Vietnamese seem to have the heads on the left-hand side uniformly, from which categorial features percolate to the resultant word. She therefore proposes a symmetrical system for head position which allows both a right-hand member and a left-hand member to be the head of a word subject to cross-linguistic variation, where the head here provides the syntactic category information and diacritic information for the resultant word (Lieber 1980:84-88,93,95-96,100). Similarly, Selkirk (1982:20-21) suggests that the position of the head in a word which provides syntactic category information and diacritic information for the word is a parameter, on the right-hand side in some languages, but on the left-hand side in others.

After looking at a series of MSC bimorphemic compound nouns and verbs, Packard (2000:39) also recognizes variation in the head location in MSC compounds, and he proposes the Headedness Principle as follows:

(16) '(Bisyllabic) noun words have nominal constituents on the right and verb words have verbal constituents on the left' (Packard 2000:39).

Some of Packard's examples are demonstrated below to illustrate this principle:

(17) a. zhi bi

paper currency

‘paper currency’

(Packard 2000:38)

b. zhi hua

paper flower

‘paper flower’

(Packard 2000:38)

(18) a. tao zou

flee walk

‘flee’

(Packard 2000:40)

b. zou dong

walk move

‘walk about’

(Packard 2000:40)

Following Packard’s Headedness Principle, in (17a), the righthand side noun *bi* ‘currency’ is the head; in (17b), the noun *hua* ‘flower’ which is on the right is the head. In (18a), the verb *tao* ‘flee’ which is on the left is the head and in (18b), *zou* ‘walk’ which is the verb on the left is the head. This Headedness principle relies on the position of the item involved. In other words, according to Packard (2000), the linear order is the clue for identification of the categorial head position in MSC bimorphemic compound nouns and compound verbs. More specifically, what he calls the canonical head of a bimorphemic MSC compound noun is the component which is on the right; the canonical head for a bimorphemic compound verb is the component which is on the left (Packard 2000:194). Many MSC compounds do follow Packard’s Headedness Principle (Liao 2014:13). However, there are some exceptional cases concerning the Headedness Principle (Packard 2000:233). For instance, in some compound verbs, as seen below, the component which



has a matching word category with the resulting compound, is in a position other than the canonical head position denoted in the Headedness Principle (Packard 2000:233):

(19) a. lang du  
bright read  
'read aloud'  
(Packard 2000:228)

b. cai pai  
colour perform  
'dress rehearse'  
(Packard 2000:228)

Here, the head in (19a) is the right-hand constituent *du* 'read' and the head in (19b) is the right-hand constituent *pai* 'perform' (Packard 2000:233). Liao (2014:13) follows Duanmu (2002) and points out that such compounds cannot be accounted for by Packard's Headedness Principle.

Liao (2014:13-14) also points out that Packard's Headedness Principle purely relies on observations of the category features of the components of the compound, while the semantic content of both constituents and the compound are ignored, which leads to wrong predictions for the heads of some compounds. The example Liao (2014:14) uses to support this idea is a subject-predicate compound, which is a noun:

(20) ri chu  
sun out  
'sunrise'  
(Liao 2014:14)

Looking only at the category of constituents, the left component *ri* 'sun' seems to be the head, as the left-hand constituent and the compound are nouns (Liao 2014:14). However, this judgement may not be accurate if the semantic content is taken into account. Specifically, the left-hand constituent cannot be the semantic head, as the whole

compound does not describe a type of sun (Liao 2014:14). In this case, the compound is exocentric in terms of category and semantics (Liao 2014:14). This example shows that semantic features as well as categorial features should be taken into consideration when it comes to locating the head in a compound word in MSC (Liao 2014).

In fact, for MSC compounds, Liao (2014:11) follows Di Sciullo and Williams (1987), and considers that in a compound, the head bears the linguistic features that agree with the linguistic features of the compound. The linguistic features include the categorial features and semantic features (Liao 2014:11). This position is not so different from that of Scalise and Fábregas (2010:112,114), who also observe that the head in a compound in general provides categorial, semantic and morphological features for the resultant compound. It may be the case that one of the components of the compound is the head which is responsible for all the features of the compound (Scalise and Fábregas 2010:120). It is also possible that those features are transmitted from different components of the compound, so that those components may be identified as different types of heads based on the types of features they define for the resultant compound (Scalise and Fábregas 2010:124). Specifically, the heads are categorized into three different types, which are the categorial head, the semantic head and the morphological head (Scalise and Fábregas 2010:124). The categorial head determines the word category of the resultant compound (Scalise and Fábregas 2010:124). The semantic head determines the semantic content which the resultant compound denotes (Scalise and Fábregas 2010:121, 124). The morphological head determines the morphological properties of the resultant compound (Scalise and Fábregas 2010:124).

Liao (2014:14-15) also makes observations on the connection between compounds and their syntactic counterparts, that is, phrases, with respect to the head location. He observes that the position of the head in MSC compounds and phrases, is isomorphic (Liao 2014:14). In other words, MSC compounds have their heads in terms of category and semantics located on the same side as that of their syntactic counterparts (Liao 2014:14). For instance, in the attributive compound noun *bai xie* ‘white shoe: white shoes’, the right-hand item *xie* ‘shoe’ is the head in terms of category and semantics (Liao 2014:14-15). In its corresponding syntactic counterpart noun phrase *bai de xie* ‘white PHRASAL MARKER shoe: white shoes’, the head in terms of category and semantic is

also on the right-hand side (Liao 2014:14-15). However, this close connection concerning the head position among compounds and their syntactic counterparts may not always be tenable. For instance, a subject-predicate compound and its syntactic counterpart do not have such a connection. The syntactic counterpart of a subject-predicate compound in MSC is a simple declarative sentence containing a subject and a verbal item (Li and Thompson 1981:70). A subject-predicate compound and its syntactic counterpart are demonstrated below:

Subject-predicate compounds

(21) a. yan hong  
      eye red  
      ‘jealous’

Subject predicate sentence

(21) b. Yan hong  
      eye red  
      ‘The eye is red’

In a subject-predicate sentence, the head is Tense (Lin 2015; Sybesma 2007), while in a subject-predicate compound, there is no such head, and therefore, the proposed close connection between heads in those two structures is untenable.

In sum, Packard’s Headedness Principle plays an important role in identifying heads in terms of category in MSC bimorphemic compound nouns and verbs, and it will be seen in section 4.5 that it is true that attributive compound nouns in MSC are right-headed. However, linear order is not the cause but an effect of headedness (Chomsky 2015:308). Following Chomsky (1995:243-246,2015:223-225,308) and much subsequent work within the Minimalist Program, I assume that the syntax, including word syntax, does not ‘see’ linear order, but only hierarchical structure. Once the structure is completed, linear order is then determined by headedness, according to language particular rules (Chomsky 2015:308).

The notion of head in a compound in general proposed by Scalise and Fábregas (2010)

can be extended to MSC compounds. To be specific, there may be a distinction between a categorial head and a semantic head in an MSC compound. The reason for leaving out the possibility of a MSC compound having a morphological head is that MSC is a language which has little morphological marking. Following Lieber (1980), Holmberg (1992) and Scalise & Fábregas (2010), I will assume feature percolation is the mechanism which determines the categorial head and the semantic head of a compound in MSC. More specifically, percolation of categorial features as well as semantic features in compounds will be looked into in order to determine the head location in terms of category and semantics. Above, we have discussed the notion of head and the ways of identifying the head in a compound in MSC. I propose that these same principles are also applicable to other types of non-function words in MSC. That is to say there is a categorial head and a semantic head in all kinds of non-function words in MSC: Minimal Words, affixed words and compound words.

### 4.3 Minimal Words in MSC: structure and head location

A Minimal Word in MSC contains a Root and a null categorizer. The null categorizer can be a null Nominalizer (n); a null adjectival categorizer (a) or a null verbal categorizer (v). The structure of, for instance, the noun *shu* ‘book’ is (22a), the structure of the adjective *mei* ‘beautiful’ is (22b) and the structure of the verb *tiao* ‘jump’ is (22c):

(22) a. [<sub>n</sub> √SHU n]

book n

‘book’

b. [<sub>a</sub> √MEI a]

beautiful a

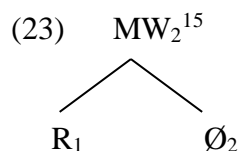
‘beautiful’

c. [<sub>v</sub> √TIAO v]

jump v

‘jump’

Those Minimal Words will have the following structure where MW represents Minimal Word, R represents Root and  $\emptyset$  represents the null categorizer:



Here the Root  $\sqrt{\text{SHU}}$  ‘book’ in (22a) completes a Minimal word *shu* ‘book’ by merging with a null categorizer, more specifically, a null Nominalizer. Following the head notion previously discussed, in (22a), the Root does not have categorial feature. But the null Nominalizer bears word category and its categorial feature percolates to the Minimal Word which dominates it, thus, the categorial head is the null Nominalizer. In terms of the semantic head, the null Nominalizer does not have semantic feature and based on the default rule (10) concerning percolation of semantic feature which states that semantic features of  $\alpha$  will percolate along with the categorial features of  $\alpha$  unless  $\alpha$  has no semantic features, in which case then semantic features of  $\beta$  will percolate, the semantic feature of the Root percolates to the resulting word *shu* ‘book’ which dominates the Root  $\sqrt{\text{SHU}}$  ‘book’. Thus the semantic head is the Root  $\sqrt{\text{SHU}}$  ‘book’.

It is the same case in (22b) and (22c), where the null categorizer is the categorial head which provides the categorial information for the Minimal Word (22b) and (22c); and the Root  $\sqrt{\text{MEI}}$  ‘beautiful’ and the Root  $\sqrt{\text{TIAO}}$  ‘jump’ are semantic heads of (22b) and (22c), which provide the semantic information for the Minimal Word (22b) and (22c).

In addition to the Minimal Words, there are affixed words in MSC. The following section concerns the structure of affixed words with suffixes and prefixes.

---

<sup>15</sup> Here  $\emptyset$  is the categorial head of MW, which is indicated by the same subscript number on  $\emptyset$  and MW. Throughout the thesis, the subscript number will be applied to indicate the categorial head in a structure in MSC and OXC.

#### 4.4 Affixed words in MSC: structures and head locations

There are suffixed words and prefixed words in MSC which are made up of an affix and a meaningful word item (please see section 2.1.2 and section 2.1.3 for a description). The null categorizer ( $\emptyset$ ) is meaningless, therefore cannot be the item that the affix (AFF) is merged with in MSC. Thus, the following will not be a possible structure for affixed words in MSC:

$$(24) *W \rightarrow [\emptyset, \text{AFF}]$$

The other types of items, including the Root (R) and the Bound Word (BW), are candidates for meaningful items that the affix (AFF) can be merged with in affixed words in MSC. In addition, the Minimal Word which contains a Root and a null categorizer ( $\emptyset$ ) can be a component of an affixed word. The possible structures of suffixed words and prefixed words are illustrated below. They will be discussed in detail in the next section and it will be seen that apart from the structure (25e), the rest of the structures are all legitimate:

The structures of suffixed words:

- (25) a.  $W \rightarrow [\text{BW}, \text{AFF}]$   
b.  $W \rightarrow [\text{R}, \text{AFF}]$   
c.  $W \rightarrow [[\text{R}, \emptyset], \text{AFF}]$

The structures of prefixed words:

- (25) d.  $W \rightarrow [\text{AFF}, \text{BW}]$   
e.  $W \rightarrow [\text{AFF}, \text{R}]$   
f.  $W \rightarrow [\text{AFF} [\text{R}, \emptyset]]$

#### 4.4.1 Suffixation

This section concerns suffixes in MSC, please see section 2.1.3 for a description. The suffix *-zhe* ‘one who does/is X’ is a nominal suffix in MSC (Packard 2000).<sup>16</sup> In other words, the resultant affixed words are all nouns referring to agents. *-Zhe* ‘one who does/is X’ is thus a nominalizer. This interpretation, by hypothesis, is the effect of feature percolation. More specifically, it can be understood if the categorial and semantic features of the suffix *-zhe* ‘one who/does/is X’ have percolated from the suffix to the dominating node, creating a word. Thus the suffix *-zhe* is a categorial and semantic head.

*Du* ‘read’ is an item which can merge with the suffix *-zhe*:

- (26) *du* - *zhe*  
read -one who does/is X  
‘reader’  
(Packard 2000:72)

If *du* ‘read’ can be a Root, then it should be able to form a Minimal Word. This is indeed the case. Below it is shown that *du* ‘read’ can occur alone as a free word in a sentence:

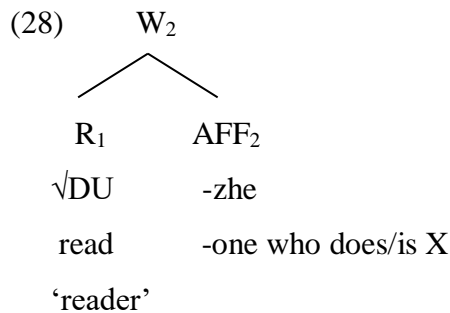
- (27) *wo jintian du le yi ben shu.*  
I today read PERFECTIVE one CLASSIFIER book  
‘I have read one book today.’

The suffix *-zhe* is a head in terms of category and semantics. The non-head which it is merged with may at first blush be a Root or a Minimal Word. However, as discussed in section 4.1.3, when there is a choice, the grammar will always pick a Root over a word, by the principle of economy: A Root has fewer features and less structure than a word. Following this, it is assumed that the non-head cannot be a Minimal Word. The discussion there was in relation to compounds. However, we may assume that it holds here in affixed words too. So here *du* ‘read’ in *du-zhe* ‘reader’ is a Root. Thus (26) will

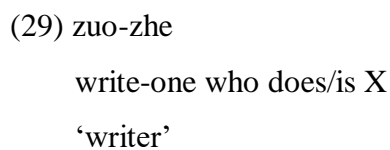
---

<sup>16</sup> Note that the suffix *-zhe* ‘one who does/is X’ used in MSC and the suffix *-zhe* used in Classical Chinese are two different suffixes. Only the suffix *-zhe* which is used in MSC will be discussed in this section.

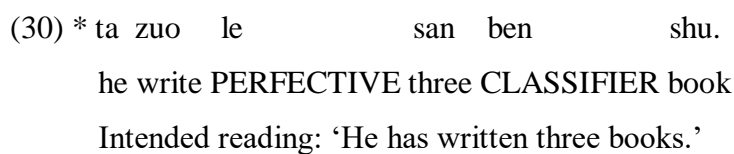
have the following structure  $W \rightarrow [_n R [_n AFF]]$  where AFF represents the affix and its tree structure is shown below:



The suffix *-zhe* ‘one who does/is X’ being the head indicates that the item it is merged with, which is a non-head, can be a Bound Word. This is possible, as has been discussed earlier in section 4.1.3, where it was shown that the Bound Word can be a non-head. *Zuo* ‘write’ is such an item which can merge with the suffix *-zhe*:



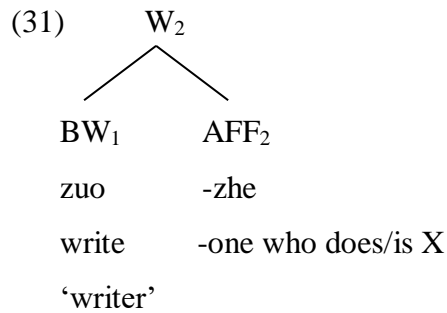
The identification of *zuo* ‘write’ in (29) as a Bound Word here is because *zuo* ‘write’ cannot be used alone as a free verb in a sentence:



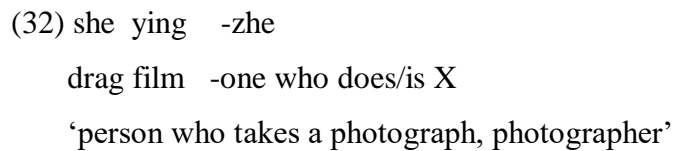
This can be explained if *zuo* ‘write’ has inherent category and as a result, it cannot merge with a null categorizer and therefore violates condition (1) which states that there are minimally two items in a free content word of MSC. And *zuo* ‘write’ is always bound.

The above example (29) has the following structure:  $W \rightarrow [_n [_v BW] [_n AFF]]$ , and the corresponding tree structure is shown below where BW represents the Bound Word:

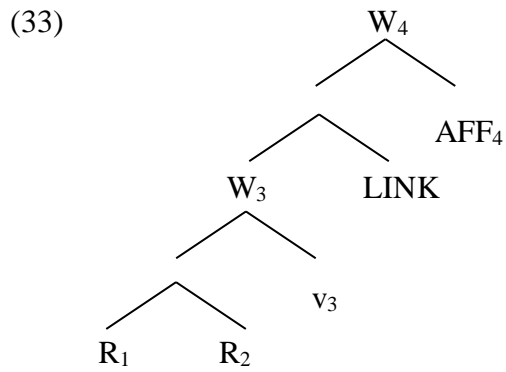




In addition, the suffix *-zhe* is also observed to be able to merge with a compound (Packard 2000:72):



Above it is demonstrated that the verb-object compound *she ying* ‘drag film: photograph’ can merge with the suffix *-zhe* to form a noun. This means that the suffix *-zhe* can merge with a compound verb and be a head in terms of category and semantics. So, there must be an item in the structure which blocks the features percolation from the compound to the dominating node, allowing *-zhe* to have its features percolate. Following Mukai (2008), a linking morpheme is assumed to be in the structure of (32) in order to determine the head. Note here verb-object compounds (see 2.2.4.1 for a description) will have the structure where there are two merged category-less Roots and a categorizer, which is adopted from Zhang (2007). The structure of verb-object compounds will be discussed more in section 4.5.3. So (32) where there is a verb-object compound and a suffix *-zhe* which is a head will have the following structure where LINK represents the linking morpheme:



Following Mukai (2008), the linking morpheme here will check the categorial features and semantic features of the verb-object compound *she ying* ‘drag film: photograph’, so those features will not percolate to the resultant word. Only the categorial features and semantic features of the affix *-zhe*, which are unchecked will percolate to the resultant word and thus the affix *-zhe* is the head in terms of category and semantics in (32).

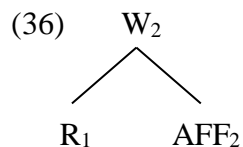
In addition to the suffix *-zhe*, another frequently used suffix in MSC is *-zi* ‘person’. The suffix *-zi* ‘person’ is classified as a nominalizer (Lin 2001:58). This is true for some cases in which the head status of the suffix *-zi* ‘person’ is shown by the interpretation and explained in terms of feature percolation. More specifically, the resulting word containing the suffix *-zi* denotes a type of person and is a noun. This can be explained if the categorial and semantic features percolate from the suffix *-zi* to the resulting word which immediately dominates the suffix. Thus, the suffix *-zi* ‘person’ is a head in terms of category and semantics and *que* ‘lame’ and *pang* ‘fat’ can merge with it, as shown below:

- (34) a. *que -zi*  
 lame -person  
 ‘lame person’
- b. *pang -zi*  
 fat -person  
 ‘fat person’

If *que* ‘lame’ and *pang* ‘fat’ can be Roots, then they should be able to form Minimal Words. This is indeed true. Both *que* ‘lame’ and *pang* ‘fat’ can be used alone as free words in attributive adjective-noun phrases:

- (35) a. *que* de tui  
 lame PHRASAL MARKER leg  
 ‘a lame leg’
- b. *pang* de ren  
 fat PHRASAL MARKER person  
 ‘a fat person’

The suffix *-zi* ‘person’ here is a head in terms of category and semantics and the item it is merged with can be a Root or a Minimal Word, but as before, due to the principle of economy the choice is always the Root, thus it follows that a Minimal Word cannot be the non-head. In (34a) and (34b), both *que* ‘lame’ and *pang* ‘fat’ are non-heads and they are thus Roots. The structure of (34a) and (34b) is  $W \rightarrow [_n R [_n AFF]]$ , and the corresponding tree structure is demonstrated below:



In addition, a different suffix *-zi* is also observed which has no features other than phonological features. The item the suffix *-zi* is merged with can be a Bound Word, for instance *jiao* ‘dumpling’:

- (37) *jiao*-*zi*  
 dumpling-ZI  
 ‘dumpling’

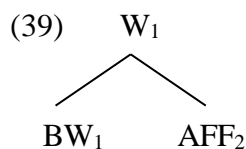
Here *jiao* ‘dumpling’ is categorized as a Bound Word as it cannot be used alone as a free word in an attributive adjective-noun phrase:

(38) \* xian de                                  jiao  
        salty PHRASAL MARKER dumpling  
        ‘??’

This can be understood if *jiao* ‘dumpling’ has inherent category, being a noun, therefore cannot merge with a null categorizer, and thus violates condition (1). Hence *jiao* ‘dumpling’ always has to be bound.

In (37), semantic and categorial features of the Bound Word *jiao* ‘dumpling’ percolate to *jiao-zi* ‘dumpling’ which immediately dominates the Bound Word *jiao* ‘dumpling’. Thus the head in terms of category (N) and semantics is the Bound Word *jiao* ‘dumpling’. The suffix *-zi* on the other hand is the non-head. According to condition (1) which states that there is always a minimum of two morphemes in a MSC free content word, I would claim that the presence of the non-head suffix *-zi* is merely to fulfill that condition. (37)

therefore has the following structure  $W \rightarrow [{}_n [{}_n \text{BW}] \text{AFF}]$ :



Above, it has been seen that the suffix *-zi* can merge with a Bound Word, as a non-head. In addition to this, it can also merge with items other than Bound Words:

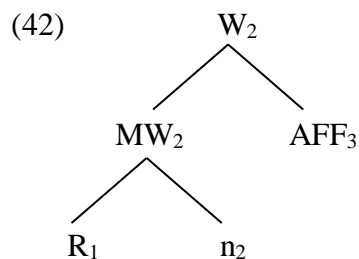
(40) fang-zi  
        house-ZI  
        ‘house’

*Fang* ‘house’ can be used alone as a free word in an attributive adjective-noun phrase:

- (41) hao de                                 fang  
       good PHRASAL MARKER house  
       ‘a good house’

This suggests that *fang* ‘house’ in (40) could be a Minimal Word or a Root. *Fang* ‘house’ cannot be a Root, however. The reason is that the suffix *-zi* has no categorial feature to project, so if it merges with a Root, a combination without word category would be produced, which would thus be ungrammatical. This is clearly not the case: *fang-zi* ‘house’ is a grammatical noun. Thus *fang* ‘house’ can only be a Minimal Word here. This predicts that the suffix *-zi* here is optional, as is indeed demonstrated in (41).

Inside the Minimal Word *fang* ‘house’, the categorial head is the null categorizer, more specifically, the null Nominalizer (n), as its categorial features percolate to the Minimal Word *fang* ‘house’ which immediately dominates the null Nominalizer. And the Root  $\sqrt{\text{FANG}}$  ‘house’, which merges with the null Nominalizer is the semantic head, as its semantic features percolate to the Minimal Word *fang* ‘house’. The structure of (40) is  $W \rightarrow [_n [_n R \ n] \text{AFF}]$  and the corresponding tree structure of is (42):



We have observed two different suffixes here. More specifically, the suffix *-zi* in (34) has meaning and category and affixes to adjectives, deriving nouns. In this case, it is a head and can merge with a Root. The other suffix *-zi*, which is illustrated in (37) and (40), has no meaning and category and affixes to nouns, deriving nouns. In this case, it is a non-head and cannot merge with a Root.

The suffixes *-tion* and *-ure* in English behave like the suffix *-zi* in MSC. Traditionally, *-tion* and *-ure* are considered as nominalizers (De Belder 2011:166):

- (43) a. action  
b. pressure

However, in some cases, *-tion* and *-ure* may be non-heads in terms of syntactic category, being essentially meaningless overt affixes (De Belder 2011:166):

- (44) a. proposition (verb)  
b. picture (verb)  
(De Belder 2011:165-166)

Those two examples are verbs, and yet they end in *-tion* and *-ure*. It is apparent that *-tion* and *-ure* are not categorial heads here (De Belder 2011:166).

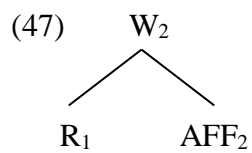
*-Tou* is another suffix which is commonly used in MSC (Lin 2001:59). The meaning of *-tou* is totally general, something like ‘thing’. Then whether the word ends up denoting a result or condition and so on depends on the meaning of the Root. *-Tou* is classified as a nominalizer in MSC (Lin 2001:59). This is indeed the case, as categorial and semantic features of the suffix *-tou* percolate to the word which immediately dominates the suffix *-tou*. Therefore, the suffix *-tou* is the categorial and semantic head and the item that the suffix *-tou* merges with is the non-head in terms of category and semantics. Thus the word which contains the suffix *-tou* denotes a result or condition and so on and is a noun. The non-head which merges with the suffix *-tou* can only be an event-denoting item or a property-denoting item and it can be a Root.

- (45) a. xiang -tou  
think -thing  
‘thought’  
  
b. tian -tou  
sweet -thing  
‘sweetness’  
(Lin 2001:59)

If *xiang* ‘think’ and *tian* ‘sweet’ can be Roots, then they in principle can form Minimal Words. Indeed *xiang* ‘think’ and *tian* ‘sweet’ can be used alone as free words in phrases:

- (46) a. *xiang* *de* *kuai*  
 think PHRASAL MARKER quickly  
 ‘think quickly’
- b. *hen* *tian*  
 very sweet  
 ‘very sweet’

The non-head in (45) could thus, on the face of it, be a Root or a Minimal Word. But as discussed in section 4.1.3, a Root is always chosen due to the principle of economy. So *xiang* ‘think’ and *tian* ‘sweet’ are Roots here. The structure of (45a) and (45b) is  $W \rightarrow [_n R [_n \text{AFF}]]$ , and the corresponding tree structure is as follows:



In addition, a different suffix *-tou* is observed which has no features other than phonological features. It can only merge with an object-denoting item, as shown below:

- (48) *she* *-tou*  
 tongue -TOU  
 ‘tongue’

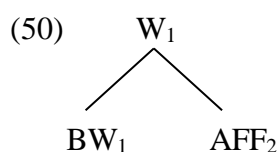
*She* ‘tongue’ is classified as a Bound Word, as it cannot be used alone as a free word in an attributive adjective-noun phrase:

- (49) \* *jiankang* *de* *she*  
 healthy PHRASAL MARKER tongue  
 ‘??’

This ungrammaticality can be explained if *she* ‘tongue’ has inherent category and cannot merge with the null categorizer. Consequently it will not meet condition (1), and must always be bound.

In (48), categorial features and semantic features of the Bound Word *she* ‘tongue’ percolate to *she-tou* ‘tongue’ which immediately dominates the Bound Word *she* ‘tongue’. Thus the Bound Word *she* ‘tongue’ is a categorial and semantic head and the suffix *-tou* is the non-head. The structure of (48) is as follows:

$W \rightarrow [_n [_n \text{ BW}] \text{ AFF}]$  and the corresponding tree structure is shown in (50):



The findings in this section concerning head location in suffixed words in MSC seem to be partially in line with what Williams (1981) and Packard (2000) propose. As mentioned earlier, Williams (1981:248) proposes the Righthand Head Rule for words, which states that the right-hand item of a word is the categorial head. The language he discussed was English. He particularly observes that in English suffixed words, the suffix determines the word category (Williams 1981:248-249). Packard (2000:166) also assumes that the suffix can be a categorial head. It has been shown above that in suffixed word in MSC, the suffix can indeed be the head in terms of category and semantics. This is not the whole picture: it is also possible that the item that the suffix merges with is the categorial and semantic head and the suffix is the non-head in terms of category and semantics. MSC suffixes behave similarly to their counterparts in many other languages, for example German and Italian. Typically suffixes in German determine the word category of the resultant word, for instance the nominalizing suffix *-ung* in *gründung* ‘foundation’ (Clahsen et al. 2003:11). Suffixes are also found in German which do not change the word category of the resultant word, for instance the diminutive suffix *-chen* in *Rädchen* ‘small wheel’ (Clahsen et al. 2003:15). Similarly, there are also diminutive suffixes in Italian, e.g. *-etto* in *orsetto* ‘little bear’, while most suffixes have effect on the category,



e.g. the nominal suffix *-iere* in *barbiere* ‘barber’ (Talamo and Celata 2011:4-5).

Based on the above discussion of the structure of suffixed words in MSC, it is apparent that (25a), (25b) and (25c) are all legitimate structures for suffixed words in MSC.

#### 4.4.2 Prefixation

In the last section, we discussed the structure and the head location in suffixed words in MSC. Now the attention will turn to prefixed words in MSC (please see section 2.1.2 for a description). Prefixes are not so much used in MSC compared with suffixes. The commonly used prefix *chong-* ‘again’ is illustrated below. It can make a complete word out of a Bound Word:

- (51) *chong-xian*  
again-appear  
‘appear again, re-appear’

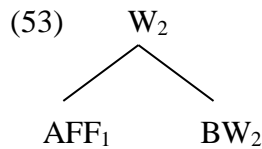
The identification of *xian* ‘appear’ as a Bound Word is because *xian* ‘appear’ cannot be used alone as a free word in a sentence:

- (52) \* *zhe ge wenti xian le.*  
this CLASSIFIER problem appear PERFECTIVE  
‘??’

This ungrammaticality can be explained if *xian* ‘appear’ has inherent word category and cannot merge with a null categorizer, thus condition (1) is violated unless *xian* ‘appear’ is bound.

If the prefix itself has no categorial feature, it is predicted that the features of the verbal Bound Word will percolate. The prefix has a meaning, but one which can be characterized as a functional meaning, not a referential/descriptive meaning. This is the situation here. Recall that the default rule (10) concerning percolation of semantic features is that categorial and semantic features percolate together. So here in addition to

categorial features, semantic features of *xian* ‘appear’ will also percolate to resultant word *chong-xian* ‘appear again’. Thus the Bound Word *xian* ‘appear’ is the categorial and semantic head of (51), as it refers to an activity of appearing and is a verb. The structure of (51) is as follows  $W \rightarrow [_v [AFF[_v BW]]]$  where AFF represents the prefix and the tree structure is shown below :



The prefix *chong-* ‘again’ can also make a word out of an item other than the Bound Word. *Shen* ‘check’ is such an item:

- (54) *chong-shen*  
 again-check  
 ‘recheck’

*Shen* ‘check’ may be a Root, which can form a Minimal Word. Indeed, in (55), it is shown that *shen* ‘check’ can be used alone as a free word in a phrase:

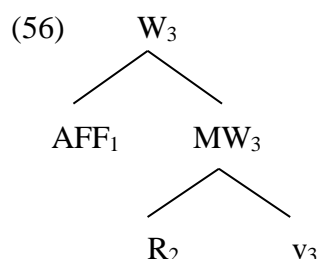
- (55) *manman de shen*  
 slowly DE check  
 ‘check slowly’

*Shen* ‘check’ cannot be a Root, and as such is category-less. If it is a Root and merges with the prefix *chong-* ‘again’, which is also category-less, then the resultant word would have no categorial feature and thus would be ungrammatical, clearly a false prediction. Thus *shen* ‘check’ in (54) is a Minimal Word and is the categorial and semantic head, as (54) denotes an action of checking and is a verb.

And inside the Minimal Word *shen* ‘check’, categorial features of the null verbal categorizer percolate to the Minimal Word which immediately dominates it, so the null verbal categorizer is the categorial head of the Minimal Word *shen* ‘check’. But in terms

of semantics, semantic features of the Root  $\sqrt{\text{SHEN}}$  percolate to the Minimal Word *shen* ‘check’ which immediately dominates it. Thus the Root  $\sqrt{\text{SHEN}}$  ‘check’ is the semantic head of the Minimal Word *shen* ‘check’. The structure for (54) is as follows

$W \rightarrow [_v \text{AFF} [_v [\text{R } v]]]$  and its tree structure is as follows:



In short, in a word of MSC which contains a prefix, the item that the prefix is merged with is the categorial head as well as the semantic head. The prefix, on the other hand, is devoid of category and any referential or descriptive meaning and thereby can only be the non-head in terms of category and semantics. This judgment is consistent with the Righthand Head Rule (Williams 1981). Following this rule, Williams (1981:248-249) observes that prefixes in English words typically do not determine the category of the word (*un-happy* is an adjective while *un-tie* is a verb; *re-surge* is a verb while *re-surgence* is a noun etc).<sup>17</sup> Similarly, Spanish prefixes are not category-changing affixes, for instance, *des-* ‘un-’ in *desplegar* ‘unfold’ (Miguel 2017:38). The same observation has also been made for the Italian prefixes, for instance the iterative prefix *ri-* in *rivendere* ‘to resell’ (Talamo and Celata 2011:3). Furthermore, compared to the prefixes, there are more suffixes in MSC (Lin 2001:61). This is in line with the observation made in other languages in the world by Bybee et al (1990). After looking at suffixes and prefixes in 71 languages distributed across the world, it is clear that there is a much larger number of suffixes than that of prefixes (Bybee et al. 1990:4).

Given that the prefix itself has no categorial feature in OXC, the item it is merged with cannot be a bare Root. Thus, we may conclude that among the structures listed for prefixed words in MSC, only (25e) is illicit and (25d) and (25f) are both licit structures for the prefixed words in MSC.

---

<sup>17</sup> There are a few exceptions, though: The prefix *en-* in *enlarge* and *enrich* appears to be a verbal categorial head (Williams 1981:249).

## 4.5 Compounding in MSC: structures and head locations

In the following sections, compounds in MSC are under analysis; please see section 2.2 for a description. These are attributive compound nouns, synonymy-compounds, parallel compounds, antonymy-compounds, verb-object compounds and subject-predicate compounds. The analyses of simple compounds in MSC are followed by discussion of complex, recursive compound nouns in MSC. Previously in section 4.1, various types of items are proposed to form a MSC word, namely the Root, the Bound Word, the Minimal Word, the null categorizer and the affix. So for MSC compounds which are formed of overt items, the Root, the Minimal Word and the Bound Word can be the components. Now, the attributive compound nouns are analysed.

### 4.5.1 *Attributive compound nouns*

Attributive compound nouns are analyzed in this section (please see section 2.2.2 for a description). Recall as discussed in section 4.1, the Root does not bear a word category, so it cannot be a head in terms of category. In addition to the Root, the Bound Word may also be non-head with respect to category if it does not project. But if it does project, then it can be a head in terms of category and semantics. Further it is argued in section 4.1.3 that the Minimal Word cannot be the non-head in terms of category and semantics. More specifically, for a Minimal Word to be such a non-head, an additional linking morpheme is required to block feature percolation from the Minimal Word. For a Root, no such mechanism is needed. Hence, derivational economy dictates that the Root will always be preferred to be the non-head. Later in chapter 5, this assumption will be supported by the behaviour of reduplicated nouns in OXC. Thus, the Minimal Word can only be a head in terms of category and semantics.

The hypothesis is that the following ordered structures are structures for attributive compound nouns in MSC. In the following sections, it will be seen that those ordered structures are not all valid.

(57) a.  $W \rightarrow [BW, BW]$

- b.  $W \rightarrow [R, BW]$
- c.  $W \rightarrow [[R, \emptyset] BW]$
- d.  $W \rightarrow [BW [R, \emptyset ]]$
- e.  $W \rightarrow [BW, R]$
- f.  $W \rightarrow [R [R, \emptyset]]$
- g.  $W \rightarrow [[R, \emptyset] R]$

To determine the structure of a simple attributive compound, the types of the items involved will be looked at first. Then the head in that compound will be examined. As has been discussed earlier, the head notion which is applied in a MSC compound is that there is a categorial head as well as a semantic head in a MSC compound. So when it comes to identifying where the head is in a MSC attributive compound, both categorial features and semantic features of the constituents will be considered.

Recall from chapter 2 that there are attributive adjective-noun compounds and attributive noun-noun compounds in MSC. As we saw, those two types of attributive compounds may be formed by two Bound Words, as in (58a, b).

Attributive adjective-noun compounds

- (58) a. zhu    jue  
       main role  
       ‘main role’

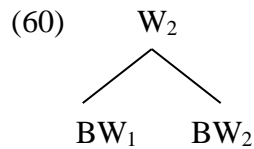
Attributive noun-noun compounds

- (58) b. yue    gan  
       music sense  
       ‘music sense:sense of music’

The components in these two compounds are Bound Words, as neither of them can occur alone as a free word in a phrase:



interpretation. This is allowed for the Bound Word *yue* ‘music’, since as is assumed in section 4.1.3, a Bound Word projects optionally. Thus, the Bound Word *gan* ‘sense’ is the head in terms of category and semantics and is therefore spelled out as the rightmost element. So  $W \rightarrow [{}_{n2}[{}_{n1}BW] [{}_{n2}BW]]$  is the structure for (58b). The tree structure for both (58a) and (58b) is as follows:



Alternatively, the two types of attributive compounds may contain a Root and a Bound word:

Attributive adjective-noun compounds

- (61) a. *mei jing*  
 beautiful scenery  
 ‘beautiful scenery’

Attributive noun-noun compounds

- (61) b. *shui zhi*  
 water quality  
 ‘quality of water’

Above, if the left-hand items, *mei* ‘beautiful’ and *shui* ‘water’ are Roots, then they can form Minimal Words. Indeed this is the case, as they can be used alone as free words in phrases:

- (62) a. *hen mei*  
 very beautiful  
 ‘very beautiful’

b. re de shui  
 hot PHRASAL MARKER water  
 ‘hot water’

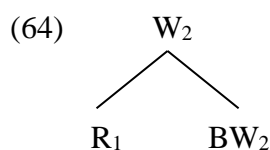
The right-hand items, which are *jing* ‘scenery’ and *zhi* ‘quality’ are Bound Words, as neither of them can be used alone as a free word in an attributive adjective-noun phrase:

(63) a. \* hao de jing  
 good PHRASAL MARKER scenery  
 ‘??’

b. \* cha de zhi  
 bad PHRASAL MARKER quality  
 ‘??’

(61a) cannot be a combination of a Minimal Word *mei* ‘beautiful’ and a Bound Word *jing* ‘scenery’. If it was then the Minimal Word *mei* ‘beautiful’ would be the head in terms of category and semantics, and (61a) would be an adjective, denoting a (modified) property *mei* ‘beautiful’. This is not the case, as (61a) is a noun denoting a kind of scenery. Thus *mei* ‘beautiful’ in (61a) can only be a Root. It is predicted that the categorial and semantic features of the Bound Word *jing* ‘scenery’ percolate to the compound which immediately dominates the Bound Word. Thus the Bound Word *jing* ‘scenery’ is the categorial and semantic head, and  $W \rightarrow [_n R \ [_n BW]]$  is the structure for (61a).

(61b) denotes a kind of quality and is a noun, which can be explained if categorial features and semantic features of the Bound Word *zhi* ‘quality’ percolate to *shui zhi* ‘water quality’ which immediately dominates the Bound Word *zhi* ‘quality’. Thus the non-head item *shui* ‘water’ is a Root. (61b) will have the following structure  $W \rightarrow [_n R \ [_n BW]]$ . The tree structure is demonstrated below for both (61a) and (61b):





Some other attributive compound nouns may be combinations of Bound Words and Minimal Words:

Attributive adjective-noun compounds

- (65) a. *wu* *shui*  
dirty water  
'dirty water'

Attributive noun-noun compounds

- (65) b. *bi* *zhi*  
wall paper  
'wallpaper'

The left item *wu* 'dirty' of (65a) cannot stand alone as a free word, as shown in (66a). The same is true of *bi* 'wall' of (65b) as shown in (66b):

- (66) a. \* *hen* *wu*  
very dirty  
'??'

- b. \* *ruan* *de* *bi*  
soft PHRASAL MARKER wall  
'??'

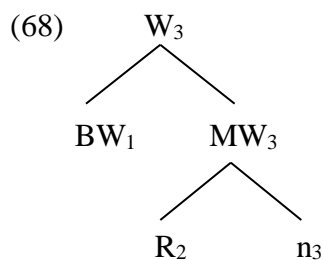
*Shui* 'water' in (65a) and *zhi* 'paper' in (65b) can be used alone as free words in the following phrases:

- (67) a. *re* *de* *shui*  
hot PHRASAL MARKER water  
'hot water'

b. hou de zhi  
 thick PHRASAL MARKER paper  
 ‘thick paper’

This means that *shui* ‘water’ in (65a) and *zhi* ‘paper’ in (65b) can, on the face of it, be Roots or Minimal Words. (65a) denotes a type of water and this can only be explained if categorial features and semantic features of *shui* ‘water’ percolate to the resultant word *wu shui* ‘dirty water’. So *shui* ‘water’ is the categorial and semantic head of *wu shui* ‘dirty water’. Thus *shui* ‘water’ can only be a Minimal Word, as a Root is category-less and cannot be a categorial head.  $W \rightarrow [{}_n [{}_a BW] [{}_n R n]]$  is the structure for (65a).

(65b) denotes a kind of paper and is a noun, which can only be explained if the categorial features and semantic features of *zhi* ‘paper’ percolate to the resultant word *bi zhi* ‘wallpaper’. Thus *zhi* ‘paper’ is the categorial and semantic head of *bi zhi* ‘wallpaper’ in (65b) and it can only be a Minimal Word. Thus (65b) will have the structure  $W \rightarrow [{}_n [{}_n [{}_n BW] [{}_n R n]]$ . Here features of the Bound Word *bi* ‘wall’ do not project, which is permitted, as Bound Words project optionally. The corresponding tree structure for both examples is shown below:



The following are examples of attributive compounds made up of a Root and a Minimal Word.

Attributive adjective-noun compounds

(69) a. xian hua  
 fresh flower  
 ‘fresh flower’

Attributive noun-noun compounds

- (69) b. niu      tui  
         cow      leg  
         ‘cow leg’

Below it is shown that the components of these compounds can all be used alone as free words in phrases:

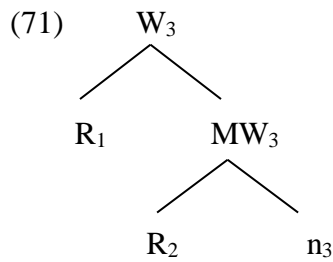
- (70) a. hen    xian  
         very fresh  
         ‘very fresh’

- b. mei      de                                  hua  
         beautiful PHRASAL MARKER flower  
         ‘beautiful flower’

- c. xiao      de                                  niu  
         small PHRASAL MARKER cow  
         ‘small cow’

- d. chang      de                                  tui  
         long PHRASAL MARKER leg  
         ‘long leg’

In both cases, the head status is shown by the interpretation. (69a) refers to a kind of flower and it is a noun. The interpretation can be explained if in (69a), the categorial and semantic features of *hua* ‘flower’ percolate. (69b) denotes a kind of leg, it is also a noun. This can be understood if categorial and semantic features of *tui* ‘leg’ percolate. Thus *hua* ‘flower’ in (69a) and *tui* ‘leg’ in (69b) are heads in terms of category and semantics, which means they can only be Minimal Words. As we have established, in section 4.1.3, that a Minimal Word cannot be a non-head, *xian* ‘fresh’ in (69a) and *niu* ‘cow’ in (69b) can only be Roots. Thus  $W \rightarrow [{}_n R [{}_n R n]]$  is the structure for both (69a) and (69b). The corresponding tree structure is shown below.



As can be seen above attributive compound nouns in MSC are all right-headed in terms of category. This is in line with the Headedness Principle of Packard's which states that in a simple nominal compound, the component which is on the right is the head (Packard 2000:39). But the difference is that the proposal here concerning the location of the head does not rely upon the linear order, rather it is determined by feature percolation (Chomsky 1995:243-246; 2015:223-225; Holmberg 1992:38). The linear order is a consequence of the headedness (Chomsky 2015:308). As shown here, attributive compound nouns in MSC are all right-headed, so among the hypothetical ordered structures illustrated earlier, the left-headed structures, that is, (57c), (57e) and (57g) are not licit structures for MSC attributive compound nouns.

#### ***4.5.2 Coordinate compounds***

This section concerns coordinate compounds in MSC; please see section 2.2.3 for a description. Recall in section 4.1.3, it is assumed that the constituents of synonymy-compounds and parallel compounds cannot be Minimal Words. Reduplicated nouns of OXC in the next chapter will provide direct empirical evidence for this assumption. Also it has been argued in section 4.1.3 that categorial features of Bound Words do not project in synonymy-compounds and parallel compounds, so they cannot be categorial heads. Instead here following Zhang (2007), a null categorizer which is external to the components of the synonymy-compounds and parallel compounds is assumed to be the categorial head.

#### 4.5.2.1 Synonymy-compound nouns

(72) is an example of a synonymy-compound noun. See section 2.2.3 in chapter 2 for a description of such compounds.

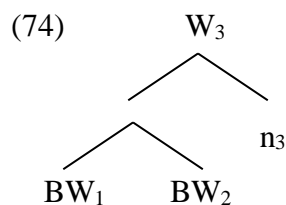
(72) shen ti  
body body  
'body'

Both constituents of the above compound are Bound Words, as neither of them can be used alone as a free word in a phrase:

(73) a. \* zhuang de shen  
strong PHRASAL MARKER body  
'??'

b. \* zhuang de ti  
strong PHRASAL MARKER body  
'??'

(72) consists of two Bound Words. It is discussed in section 4.1.3 that categorial features of Bound Words do not project in synonymy-compounds/parallel compounds. In other words, they cannot be categorial heads. Instead, following Zhang (2007), a third item which is a null categorizer will be assumed to be responsible for the category of the compound in (72). Here the null categorizer is *n*, as the resultant word in (72) is a noun. As for the semantic head, both Bound Words are semantic heads, as semantic features of both Bound Words percolate to the resultant compound in (72); it denotes a body. (72) thus has the following structure:



The two items in a synonymy-compound noun may not always be of the same type.

(75) *hun*     *yin*  
 marriage marriage  
 ‘marriage’

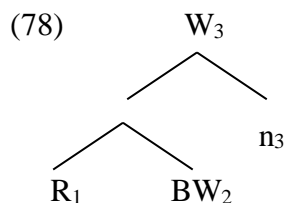
Here, the left-hand component *hun* ‘marriage’ can be used alone as a free word in a phrase, which means it can be a Minimal Word:

(76) *zhe*   *ge*             *hun*  
 this CLASSIFIER marriage  
 ‘this marriage’

Here it will be assumed that *hun* ‘marriage’ in (75) is only a Root, though. On the other hand, the right-hand item *yin* ‘marriage’ is a Bound Word, as it cannot be used alone as a free word in a phrase:

(77) \* *zhe*   *ge*             *yin*  
 this CLASSIFIER marriage  
 ‘??’

Again, the Bound Word *yin* ‘marriage’ will not project its categorial features to the compound. Rather, following Zhang (2007), a null Nominalizer is assumed to be the categorial head. The semantic heads are the Root *hun* ‘marriage’ and the Bound Word *yin* ‘marriage’, as semantic features of both items percolate to the resultant compound, which hence denotes a marriage.<sup>18</sup> The structure of (75) is shown below:



<sup>18</sup> Note here one of the components in a synonymy compound cannot be considered as a head, as there is not a modification or specification relationship among components of a synonymy compound.

Alternatively, the synonymy-compound noun may have a Bound Word on the left and a Root on the right:

- (79) *shu*      *cai*  
      vegetable vegetable  
      ‘vegetable’

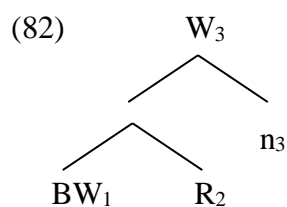
*Shu* ‘vegetable’ is a Bound Word, as it cannot be used alone as a free word:

- (80) \* *hao*   *de*                      *shu*  
      good PHRASAL MARKER vegetable  
      ‘??’

*Cai* ‘vegetable’ can occur alone as a free word in an adjective-noun phrase:

- (81) *xiang*   *de*                      *cai*  
      delicious PHRASAL MARKER vegetable  
      ‘delicious vegetable’

In (79), *cai* ‘vegetable’ is assumed to be a Root only. I assume that the Bound Word *shu* ‘vegetable’ does not project its categorial feature either and thus following Zhang (2007), a null Nominalizer is assumed as the categorial head of *shu cai* ‘vegetable’. The semantic heads are the Bound Word *shu* ‘vegetable’ and the Root *cai* ‘vegetable’. The structure of (79) is demonstrated below:



As synonymy-compound noun may also consist of two Roots:

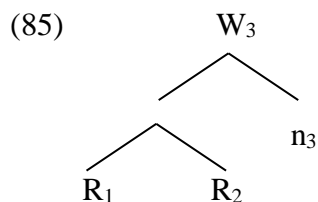
(83) hao ma  
number number  
'number'

Both *hao* 'number' and *ma* 'number' can occur alone as free words in phrases:

(84) a. chang de hao  
long PHRASAL MARKER number  
'long number'

b. chang de ma  
long PHRASAL MARKER number  
'long number'

Here both *hao* 'number' and *ma* 'number' are assumed to be Roots only. Thus Zhang's (2007) Root Merger analysis is adopted where two Roots can merge, forming a compound which is then merged with a categorial head. (83) is a noun, so (83) will have the following structure where the head is *n*:



The semantic features of both *hao* 'number' and *ma* 'number' percolate to the resultant compound, so they are both semantic heads of (83).<sup>19</sup>

---

<sup>19</sup> Here the semantic heads are Roots, as the meanings of the resultant compound nouns are compositional. If the meaning of the resultant compound where there are two merged Roots and a head is not compositional, then the meaning comes directly from the lexicon, as will be seen later in section 4.5.3 in which the Roots cannot be semantic heads. In short, the Root may or may not contribute the meaning to the resultant compound, which has also been discussed by Marantz (2007).

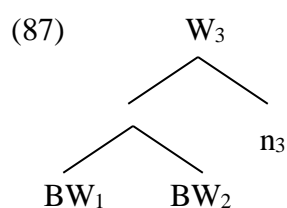


#### 4.5.2.2 Parallel compound nouns

A parallel compound noun (see section 2.2.3 for a description) may consist of two Bound Words:

- (86) *yin* *yue*  
sound music  
'music'

Both *yin* 'sound' and *yue* 'music' are classified as Bound Words, as neither of them can be used alone as free words in phrases. Categorial features of Bound Words in parallel compound nouns are assumed not to percolate to the resultant compound as discussed in section 4.1.3, thus a null Nominalizer is assumed to be the categorial head following Zhang (2007). (86) refers to music which can be understood if semantic features of the Bound Word *yue* 'music' percolate to the compound. Thus the Bound Word *yue* 'music' is the semantic head. (86) will have the following structure:

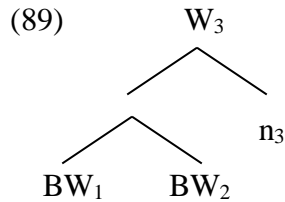


Alternatively, a parallel compound noun may be formed by two Bound Words with the semantic head on the left-hand side:

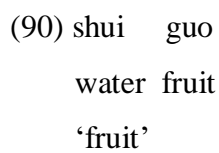
- (88) *zhi* *liang*  
quality quantity  
'quality'

Neither *zhi* 'quality' nor *liang* 'quantity' can be used alone as free words in phrases, so they are Bound Words. (88) refers to a quality and is a noun. This interpretation is because the semantic features of the Bound Word *zhi* 'quality' percolate to the compound in (88) which immediately dominates it. Thus the Bound Word *zhi* 'quality' is the

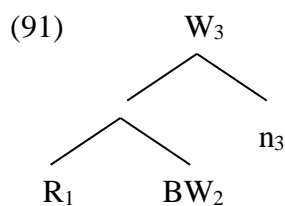
semantic head of (88). In terms of the categorial head, following Zhang (2007), the third item which is the null Nominalizer here is the categorial head. The structure of (88) is demonstrated below:



A parallel compound noun may be formed of a Root and a Bound Word in that order:



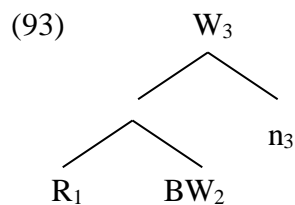
*Shui* ‘water’ can occur alone as a free word in a phrase. Thus on the face of it, it can be a Root or a Minimal Word in (90). But here it is assumed that *shui* ‘water’ is only a Root. This will be supported by empirical evidence in chapter 5. As for *guo* ‘fruit’, it is a Bound Word, as it cannot be used alone as a free word. (90) denotes a fruit. This interpretation can be understood if semantic features of the Bound Word *guo* ‘fruit’ percolate to the compound which immediately dominates it, so the Bound Word *guo* ‘fruit’ is the head in terms of the semantics. Following Zhang (2007), the categorial head is the null Nominalizer, so the structure of (90) is as follows:



The parallel compound noun may contain a Root and a Bound Word, in that order, where the semantic head is the Root:

(92) yan jing  
 eye eyeball  
 ‘eye’

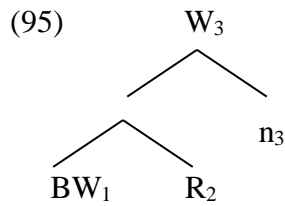
The left item *yan* ‘eye’ can occur alone as a free word in a phrase. So it can, on the face of it, be a Root or a Minimal Word in (92). Here *yan* ‘eye’ is assumed to be a Root only. The right item *jing* ‘eyeball’ is a Bound Word, as it cannot be used alone as a free word in a phrase. (92) denotes eye, because the semantic features of the Root *yan* ‘eye’ percolate to the compound (92) which immediately dominates it. Thus, the Root *yan* ‘eye’ is the semantic head. The null Nominalizer is the categorial head following Zhang (2007), so the structure for (92) is demonstrated below.



The parallel compound nouns may be composed of a Bound Word and a Root, in that order:

(94) xiang bao  
 case bag  
 ‘bag’

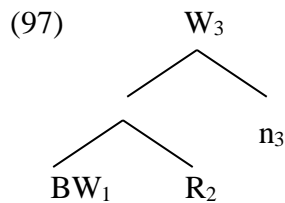
The left-hand item which is *xiang* ‘case’ is a Bound Word as it cannot be used alone as a free word. *Bao* ‘bag’ can occur alone as a free word in a phrase, so it can be a Root or a Minimal Word. But here *bao* ‘bag’ can, by assumption, only be a Root. (94) refers to a bag which means that semantic features percolate from the Root *bao* ‘bag’ to *xiang bao* ‘bag’ which immediately dominates it. Hence the Root *bao* ‘bag’ is the semantic head. Following Zhang (2007), the null Nominalizer is the categorial head, the structure of (94) is demonstrated below:



Alternatively, a Bound Word and a Root may also form a parallel compound where the Bound Word is the semantic head.

(96) *jie*    *mao*  
 eyelash fur  
 ‘eyelash’

*Jie* ‘eyelash’ is a Bound Word as it cannot be used alone as a free word. *Mao* ‘fur’ can stand alone as a free word in a phrase, so it can be a Minimal Word or a Root. Here however, *mao* ‘fur’ is only assumed to be a Root. By looking at the semantic percolation facts in (96) from constituents to the compound, it is apparent that the Bound Word *jie* ‘eyelash’ is the semantic head. The null Nominalizer is the categorial head following Zhang (2007), so (96) will have the following structure:

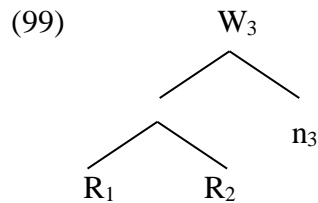


Finally, parallel compound nouns may be formed of Roots only.

(98) *ma*    *lu*  
 horse road  
 ‘road’

Both *ma* ‘horse’ and *lu* ‘road’ can occur alone as free words in phrases, so they can, on the face of it, be Roots or Minimal Words in (98). But here *ma* ‘horse’ and *lu* ‘road’ are assumed to be Roots only. There are thus two category-less Roots in (98), so Zhang’s

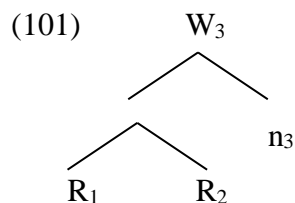
(2007) Root Merger analysis is suitable here. (98) denotes a road, because the semantic features of the Root *lu* ‘road’ percolate to the compound, thus the Root *lu* ‘road’ is the semantic head of (98). (98) is a noun, so the null Nominalizer is the categorial head. Below it is the structure for (98).



It is also possible that the left Root is the semantic head in a parallel compound noun in which case there are two Roots.

(100) *mi fan*  
*rice dinner*  
 ‘rice’

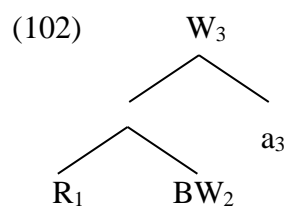
*Mi* ‘rice’ and *fan* ‘dinner’ can occur alone as free words in phrases, thus they can be Roots or Minimal Words. Here it is assumed that both *mi* ‘rice’ and *fan* ‘dinner’ are only Roots. Thus Zhang’s (2007) Root Merger analysis is appropriate here. The categorial head in (100) is the null Nominalizer, as (100) is a noun. (100) denotes rice, which can be explained if semantic features of the Root *mi* ‘rice’ percolate to the compound which immediately dominates it, so the Root *mi* ‘rice’ is the semantic head of (100). Thus (100) will have the following structure:



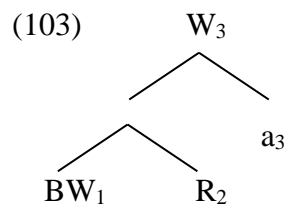
#### 4.5.2.3 Synonymy-compound adjectives

This section concerns synonymy-compound adjectives, which are described in section 2.2.3. They are compounds where both components are semantic heads and following Zhang (2007), the categorial head is the null adjectival categorizer. As shown below, any combination of Bound Words and Roots is possible.

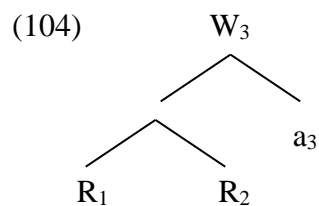
*Mei li* ‘beautiful-beautiful: beautiful’ will have the following structure:



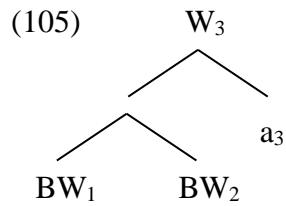
*Xia zhai* ‘narrow-narrow: narrow’ will have the following structure:



*Chun ben* ‘stupid-stupid:stupid’ will have the following structure:



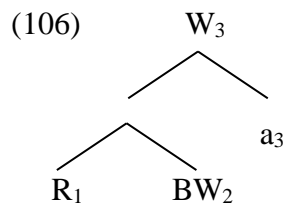
*Cong ying* ‘clever-clever: clever’ will have the following structure:



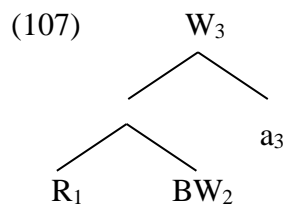
#### 4.5.2.4 Parallel compound adjectives

This section concerns parallel compound adjectives, which are described in section 2.2.3. They are compounds where one of the components is the semantic head. And following Zhang (2007), the null adjectival categorizer is the categorial head. Again, any logically possible combination of Bound Words and Roots can be found.

*Chi dun* ‘slow stupid: stupid’ will have the following structure where the semantic head is the Bound Word *dun* ‘stupid’:



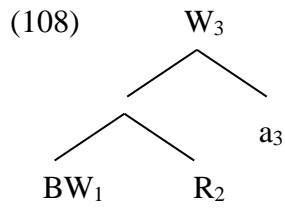
*Leng mo* ‘cold indifferent: cold’ will have the following structure where the semantic head is (arguably) the Root *leng* ‘cold’<sup>20</sup>:



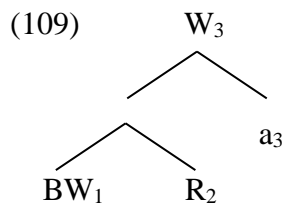
*Wei ruo* ‘small weak: weak’ will have the following structure where the semantic head is the Root *ruo* ‘weak’:

---

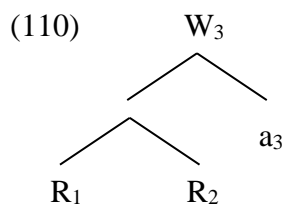
<sup>20</sup> This is one of many cases where determination of semantic head is debatable.



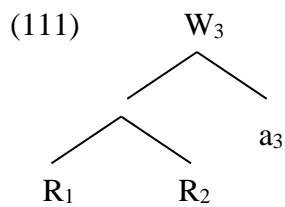
*You hao* ‘friendly fine: friendly’ will have the following structure where the semantic head is the Bound Word *you* ‘friendly’:



*Bei can* ‘sad tragic: tragic’ will have the following structure where the semantic head is the Root *can* ‘tragic’:

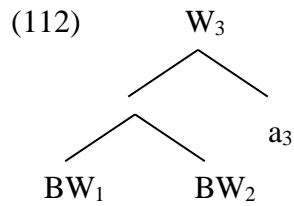


*Mei hao* ‘beautiful good: beautiful’ will have the following structure where the semantic head is the Root *mei* ‘beautiful’:

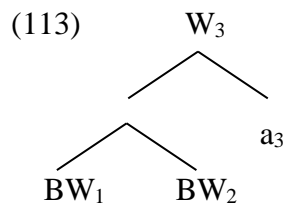


*Bei bi* ‘low mean: mean’ will have the following structure where the semantic head is the Bound Word *bi* ‘mean’:





*Bei wei* ‘low small: low’ will have the following structure where the semantic head is the Bound Word *bei* ‘low’:



The compound words discussed in the previous sections show what we expect to see when a compound is symmetrical (Lieber 1980), the way synonymy and parallel compounds are: we find any possible ordering of semantic head and non-head. Furthermore, we find no overall rule to project categorial features from one of the components. Neither of these is the case with attributive compounds: They follow strict ordering and projection rules. This can be understood only if they are not symmetrical.

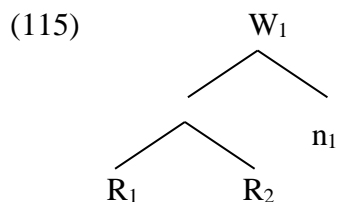
#### ***4.5.3 Remaining compounds: Antonymy-compounds, verb-object compounds and subject-predicate compounds***

This section concerns antonymy-compounds, verb-object compounds and subject-predicate compounds in MSC; please see section 2.2.4 for a description. Antonymy-compounds are observed to have non-compositional meanings and their word category may not be obtainable from their constituents (Zhang 2007). What Zhang (2007) calls Free projectivity has also been observed in verb-object compounds and subject-predicate compounds in MSC (Zhang 2007). As mentioned in section 3.5, according to Zhang (2007), the exocentricity and free projectivity are abnormalities. She proposes that they are cases of Root Merger. This idea is based on the assumption that Roots have no categorial feature, and is compatible with the theory articulated in this chapter.

Below is an example of antonymy-compound; see section 2.2.3 for a description of such compounds:

(114) mei      chou  
         beautiful ugly  
         ‘appearance’

Here, if both components are assumed to have a word category, then it is apparent that the category of the components does not match that of the resultant compound (Zhang 2007:172-173). In other words, there are discrepancies concerning word category among the constituents and the resultant compound (Zhang 2007:172-173). As discussed in section 3.5, this can be understood if this kind of compound is exocentric, formed of category-less Roots and a functional head (Zhang 2007:172-173,178). Roots do not have a category, which means there are no projection issues. Instead, a third element, which is a functional head, will be responsible for the word category of the compound (Zhang 2007:177-178). According to Zhang (2007:177), the functional head is the null categorizer, *n*, *a*, or *v*, which we have also assumed in this chapter. (114) is a noun, thus the categorizer is *n*. The structure of (114) is shown below<sup>21</sup>:



Discussing verb-object compounds in MSC (see section 2.2.4 for a description), as mentioned in section 3.5, Zhang (2007:173) observes that they can be in various categories, such as verb, noun or adjective:

(116) a. dan      xin  
         carry    heart  
         ‘worry’

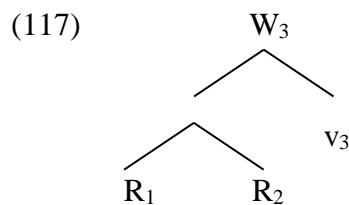
---

<sup>21</sup> Note here, according to Zhang (2007), the set {R<sub>1</sub>,R<sub>2</sub>} is labelled as R<sub>1</sub>+R<sub>2</sub>. This is not necessary, as R<sub>1</sub> and R<sub>2</sub> do not project. Thus in this section, the set {R<sub>1</sub>,R<sub>2</sub>} is not labelled.

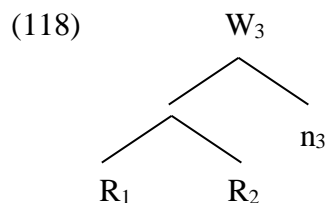
b. sheng yi  
 grow meaning  
 ‘business’

c. dong ren  
 act person  
 ‘moving’

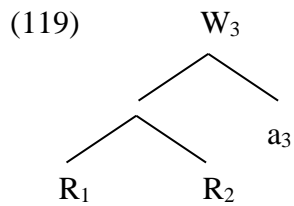
Freedom of projectivity is observed: (116a) is a verb, while (116b) is a noun and (116c) is an adjective. As discussed in section 3.5, the idea that both constituents are Roots without syntactic features can provide explanations for such an abnormality (Zhang 2007:178). More specifically, according to Zhang (2007:178), the word category of such compound will not be able to be projected from their constituents but rather will be determined by the functional head, that is, null categorizer *n*, *a*, or *v*, that we have also assumed in this chapter. More specifically, the category of such compounds depends on which null categorizer they have (Zhang 2007:178). (116a) is a verb, thus the head is null verbal categorizer *v* and the structure for (116a) is as follows:



The structure for verb-object compound (116b) which is a noun is shown below where the head is the null Nominalizer *n*:



(116c) is an adjective. The structure is as follows where the head is the null adjectival categorizer *a*:



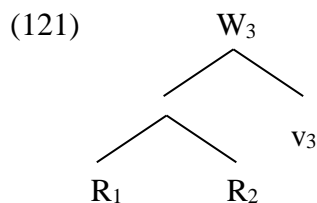
Similar to verb-object compounds, subject-predicate compounds (see section 2.2.4 for a description) can be in more than one category (Chao 1968: 368-370; Zhang 2007:174). More specifically, they can be verbs, nouns and adjectives and have non-compositional meanings (Chao 1968:368):

(120) a. xin tong  
 heart painful  
 ‘hurt’

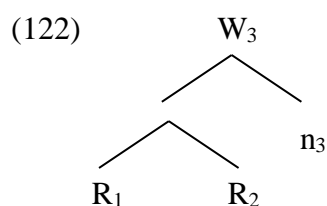
b. hai xiao  
 sea roar  
 ‘tsunami’

c. xin sui  
 heart smash  
 ‘heartbroken’

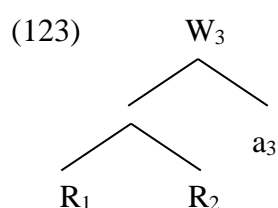
It is apparent that the subject-predicate compound is another case of freedom of projectivity (Zhang 2007). Thus, this type of compound will have a structure where there are two merged category-less Roots and a functional head, that is, a null categorizer (Zhang 2007). The following structure is for the verbal subject-predicate compound (120a):



The subject-predicate compound (120b) which is a noun has the following structure:



The structure for (120c) which is an adjective is demonstrated below:



The meanings of those three types of compounds (i.e. antonymy compounds, verb-object compounds and subject-predicate compounds) are not derivable from their components and instead they are unpredictable, thus lexically derived (see Bauke 2014 on Root-Root compounds in German).

#### ***4.5.4 Structures of MSC recursive compounds: an unexpected constraint***

This section concerns recursive compounds in MSC; please see 2.2.5 for a description. As discussed in section 4.5, the Root, the Bound Word and the Minimal Word are proposed to be the elements of compounds in MSC. So for recursive compound nouns in MSC, the null hypothesis is that:

(124) The Root, the Bound Word and the Minimal Word can merge freely with each other to form recursive compounds in MSC.

However, this is false, as the following examples of recursive compounds are found to be ungrammatical:

(125) a. \*[mu [ yi gui]]

wood clothing cupboard

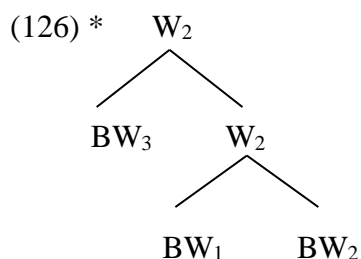
Intended reading: ‘clothing cupboard which is made of wood’

b.\* [wei [wen lv]]

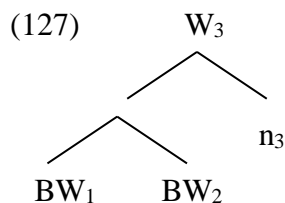
fake character law

Intended reading: ‘fake law’

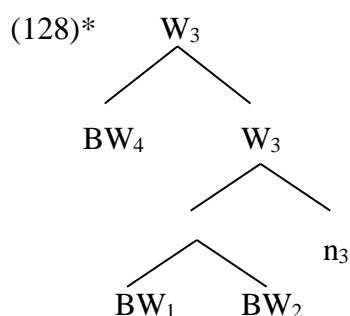
Here, the inner compound *yi gui* ‘clothing cupboard’ in (125a) is an attributive compound noun and *wen lv* ‘character law: law’ is a parallel compound noun. *Mu* ‘wood’, *yi* ‘clothing’ and *gui* ‘cupboard’, *wen* ‘character’ and *lv* ‘law’ are all Bound Words. In *yi gui* ‘clothing cupboard’ of (125a), *gui* ‘cupboard’ is the head in terms of category and semantics. So the structure of *yi gui* ‘clothing cupboard’ is  $W \rightarrow [{}_{n2}[{}_{n1}BW][{}_{n2}BW]]$ . The ungrammaticality of (125a) means that  $W \rightarrow [{}_{n2}[{}_{n3}BW][{}_{n2}[{}_{n1}BW][{}_{n2}BW]]$  is not a legitimate structure of recursive compounds in MSC. The hypothetical tree structure is shown below:



In *wen lv* ‘character law: law’ of (125b), *lv* ‘law’ is the semantic head and the null Nominalizer is the categorial head, and so *wen lv* ‘law’ will have the following structure:



So, (125b) will have the structure below. The fact that (125b) is ungrammatical means that this structure below is not licit for a recursive compound noun in MSC:

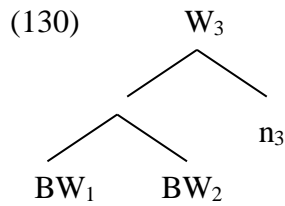


In addition, the following examples are also not grammatical:

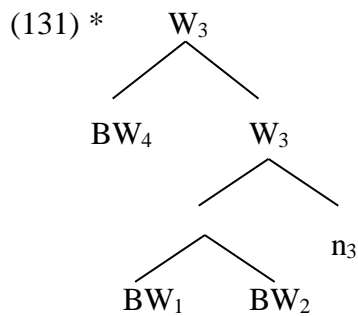
(129) a. \* [gui [zhi liang]]  
 cupboard quality quantity  
 Intended reading: ‘quality of cupboard’

b. \*[mu [cai liao]]  
 wood material material  
 Intended reading: ‘material which is wood’

Above, the inner compound in (129a) which is *zhi liang* ‘quality quantity: quality’ is a parallel compound noun; and *cai liao* ‘material’ which is the inner compound in (129b) is a synonymy-compound noun. *Zhi* ‘quality’, *liang* ‘quantity’, *cai* ‘material’ and *liao* ‘material’ are Bound Words. *Zhi* ‘quality’ is the semantic head of the inner compound *zhi liang* ‘quality’ and *cai* ‘material’ and *liao* ‘material’ are both semantic heads of the inner synonymy-compound noun *cai liao* ‘material’. In both inner compounds the null Nominalizer is the categorical head. Thus inner compound *zhi liang* ‘quality quantity: quality’ of (129a) and the inner compound *cai liao* ‘material’ in (129b) will have the following structure:



The ungrammaticality of (129a) and (129b) means that the following structure cannot be the correct structure for recursive compound nouns:



Further, the following examples are also ill-formed in MSC:

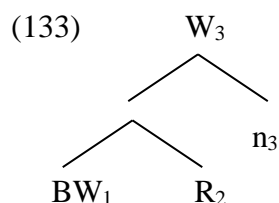
(132) a. \*[mu [jing kuang]]  
 wood mirror frame  
 Intended reading: 'mirror frame made of wood'

b. \*[ge [xiang bao]]  
 leather case bag  
 Intended reading: 'leather bag'

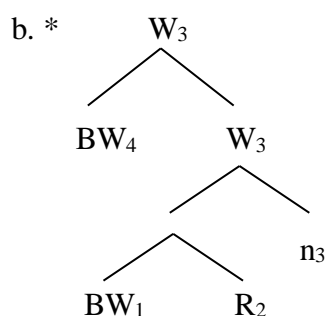
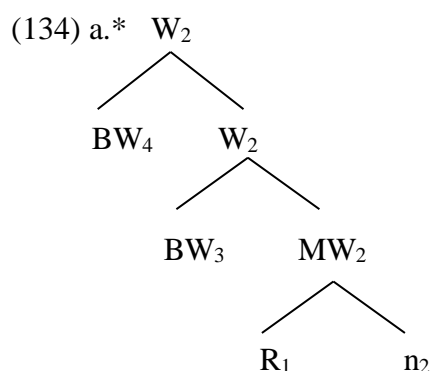
Above, the inner compounds in both examples, that is, *jing kuang* 'mirror frame' and *xiang bao* 'case bag: bag' are an attributive compound noun and a parallel compound noun respectively. *Mu* 'wood' and *jing* 'mirror' in (132a) and *ge* 'leather' and *xiang* 'case' in (132b) are all Bound Words. As for *kuang* 'frame' of (132a), although out of context it can be a Root or a Minimal Word, here it is a Minimal Word as it is a categorial and semantic head of *jing kuang* 'mirror frame' of (132a). Thus *jing kuang* 'mirror frame' of (132a) has the structure  $W \rightarrow [n_2 [n_1 BW] [n_2 R n]]$ .



As for *bao* ‘bag’ of (132b), as it is assumed that the components in a synonymy-compound noun or a parallel compound noun cannot be a Minimal Word, thus it is assumed here that it is a Root. The semantic head of (132b) is the Root *bao* ‘bag’ and the categorial head of *xiang bao* ‘bag’ is the null Nominalizer. The structure of *xiang bao* ‘case bag: bag’ of (132b) is as follows:

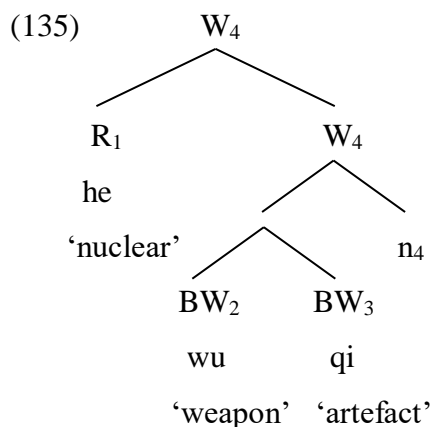


The fact that (132a) and (132b) are not grammatical indicates that the following structures are not possible structures of MSC recursive compounds:



This finding seems to be in line with Packard (2000:180). According to him, there are no examples observed of a right-branching structure where the leftmost and the second elements are Bound Roots (in his terms) and the rightmost element is a root word (in his

terms). But Packard (2000:180) also observes one example of a right-branching recursive compound where all three constituents are Bound Roots (in his terms). In other words, such a structure according to him is legitimate. However, after close examination, the single example he mentions, which is *he wu qi* ‘nuclear weapon artefact: nuclear weapon’, is not a combination of three Bound Roots (in his terms), as *he* ‘nuclear’ is a root word (in his terms). Thus *he wu qi* ‘nuclear weapon artefact: nuclear weapon’ will have the following tree structure:



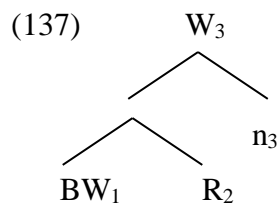
As this is the only example Packard has, I would suggest that the structure where there are two Bound Words structured as in (134) is still not a legitimate structure. In addition, Packard (2000:181) also observes no examples of a right-branching recursive compound where there is a leftmost bound content item (which he terms a Bound Root) and a second (left) bound function item (which he terms a word-forming affix).<sup>22</sup> This finding of his and my findings imply that it is the case that the structure of right-branching recursive compound in MSC in which both the leftmost item and second left item are bound words is not a legitimate structure. This is indeed the case, as apart from the examples above, the examples below are also ungrammatical:

(136) a.\* [you [chu fang]]  
 fine kitchen house  
 Intended reading: ‘fine kitchen’

<sup>22</sup> This structure is not a legitimate structure for recursive compound nouns in MSC may be due to the reason that there are no prefixes in MSC which can be applied to form common nouns.

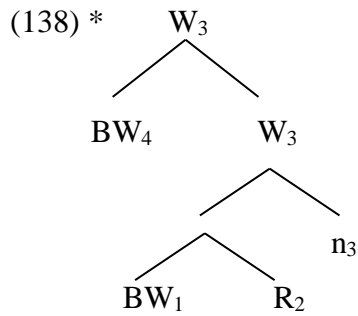
- b. \* [bao [sheng yin]]  
 hailstone sound sound  
 Intended reading: ‘sound of hailstone’

Here, the inner compound in (136a) which is *chu fang* ‘kitchen house: kitchen’ is a parallel compound noun and the inner compound in (136b), which is *sheng yin* ‘sound’ is a synonymy compound noun. *You* ‘fine’ in (136a) and *bao* ‘hailstone’ in (136b) are both Bound Words. In the inner compound of (136a), the head in terms of semantics is *chu* ‘kitchen’. *Chu* ‘kitchen’ is a Bound Word and *fang* ‘house’ is a Root. In the inner synonymy-compound noun *sheng yin* ‘sound’ of (136b), *sheng* ‘sound’ is a Bound Word and *yin* ‘sound’ is assumed to be a Root only. The Bound Word *sheng* ‘sound’ and the Root *yin* ‘sound’ are the semantic heads of *sheng yin* ‘sound’ in (136b). In both inner compounds of (136a) and (136b), the categorial head is the null Nominalizer. So *chu fang* ‘kitchen’ in (136a) and *sheng yin* ‘sound’ in (136b) will have the following structure:



This structure is different from the previous structure (133) in the sense that the semantic head is different. More specifically, (133) is the structure of the inner compound in (132b) and the semantic head is the Root. Here, (137) is the structure of the inner compound in (136a) and the inner compound in (136b). In the inner compound in (136a), the semantic head is the Bound Word and in the inner compound in (136b), the semantic heads are the Bound Word and the Root.

(136a) and (136b) are ungrammatical, which means that recursive compound nouns in MSC cannot have the following structure:



The ungrammaticality shown in (125a) and (125b) is not the result of meaning incompatibility, as the compounds are perfectly meaningful. It is due to the types of word elements. Specifically, in (125a) and (125b), the leftmost item and second left item, that is, *mu* ‘wood’ and *yi* ‘clothing’ in (125a) and *wei* ‘fake’ and *wen* ‘character’ in (125b), are all Bound Words. This is the source of the ungrammaticality, shown by the fact that when the leftmost item or the second item in both examples are substituted by different types of items, for example, a Root, the examples become grammatical. In (139a), the second item *yi* ‘clothing’ in (125a) is changed to a Root, that is, *shu* ‘book’, and in (139b), the leftmost item *wei* ‘fake’ in (125b) is changed to a (synonymous) Root, that is, *jia* ‘fake’:

- (139) a. *mu* [ *shu* *gui* ]  
 wood book cupboard  
 ‘book cupboard made of wood’
- b. *jia* [ *wen* *lv* ]  
 fake character law  
 ‘fake law’

Both (139a) and (139b) are grammatical. This finding can also be observed in examples (129a,b), (132 a,b), (136a,b) where the replacements of the leftmost or the second left item with items other than the Bound Word will produce grammatical examples :

- (140) a. [ *shu* [ *zhi* *liang* ] ] (compare 129a)  
 book quality quantity  
 ‘quality of the book’

b. [lv            [cai      liao]] (compare 129b)

aluminium material material

‘material which is aluminium’

c. [mu [men    kuang]] (Compare 132a)

wood door    frame

‘door frame made of wood’

d. [pi      [xiang    bao]] (compare 132b)

leather case    bag

‘leather bag’

e. [hao [chu    fang]] (compare 136a)

fine    kitchen house

‘fine kitchen’

f. [shui [sheng    yin]] (compare 136b)

water    sound sound

‘sound of water’

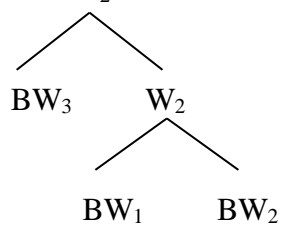
The leftmost Bound Words in each of the example (129a) and (129b), that is, *gui* ‘cupboard’ and *mu* ‘wood’ are changed to Roots *shu* ‘book’ and *lv* ‘aluminium’, as shown in (140a) and (140b). (140d) demonstrates the case where the leftmost Bound Word *ge* ‘leather’ in (132b) is replaced by the item *pi* ‘leather’, which is a Root. For examples (136a) and (136b), the situation is the same, once the leftmost Bound Words *you* ‘fine’ and *bao* ‘hailstone’ in both cases are substituted by Roots *hao* ‘fine’ and *shui* ‘water’, the examples are grammatical, as shown in (140e) and (140f). As for the example (132a), it is the second Bound Word, that is *jing* ‘mirror’ that is replaced by the Root *men* ‘door’ and this replacement also makes a grammatical example, as shown in (140c).

It can be seen above that an attributive compound noun, a synonymy-compound noun or a parallel compound noun can be the inner compound in the ungrammatical right-branching recursive compounds. And in those ungrammatical right-branching recursive compounds,

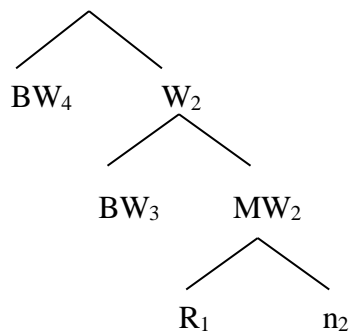
both the leftmost item and the second left item are Bound Words. So the ungrammaticality of the right-branching recursive compound is not caused by the type of inner compound it contains. Rather, it is caused by the configuration of the leftmost Bound Word and the second left Bound Word in the ungrammatical right-branching recursive compound.

In summary the following structures are observed not to be legitimate structures for recursive compound nouns in MSC.

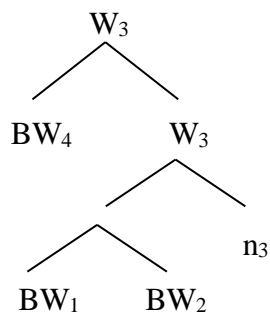
(141) a.\*  $W_2$

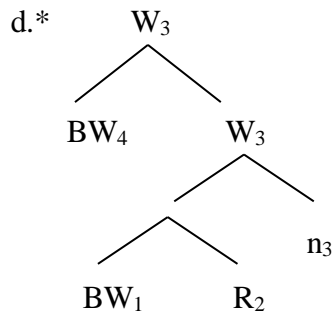


b.\*  $W_2$



c.\*





These structures can be summarized into a structure:

$$(142)^* W \rightarrow [BW / BW X$$

In the above structure, the italic bracket means one or more left-brackets. *X* can be a Bound Word or a Minimal Word. Based on the above empirical evidence, (142) is demonstrated not to be a structure of MSC right-branching recursive compound nouns. Note here, *X* *BW* / *BW* ] is a legitimate structure, as shown by the following examples, here the italic bracket means one or more right-brackets:

(143) a. [[yin yue] xue]  
 sound music study  
 ‘study of music’

b. [[ying guo] ji]  
 England country citizenship  
 ‘citizenship for England: British citizenship’

In other words, the exception is sensitive to linear order. Based on those observations, the following hypothesis is proposed:

(144) An unattached nominal Bound Word will attach to an adjacent nominal Bound Word on its right.

This will give the grammatical structure only in the case when the two Bound Words are sisters. (144) would be a processing rule not a rule of grammar (Hawkins 1990). This can

explain why the rule is sensitive to linear order rather than structure. Hawkins (1990) proposes parsing rules concerning syntactic phrasal constructions. I will adapt his theory here to complex words. Hawkins (1990) does not consider processing of complex words, but there is no *a priori* reason why the theory could not be relevant for the processing of the internal structure of complex words.

According to Hawkins (1990:227-228), a head is said to form a mother node, and is thereby an Immediate Constituent of the mother node. There are constituents which are attached to the mother node but do not form a mother node, and they are also Immediate Constituents of the mother node (Hawkins 1990:227-228). For example, NP which is attached to PP is an Immediate Constituent of the mother node PP and the head N forms the Immediate Constituent NP (Hawkins 1990:227). To recognize and process all Immediate Constituents of a mother node from left to right, a human parser has to parse a group of words and this group of words forms a Constituent Recognition Domain (Hawkins 1990:229). In other words, a Constituent Recognition Domain is identified as a group of words which are required for a human parser to recognize all Immediate Constituents of a mother node (Hawkins 1990:229). More specifically, a Constituent Recognition Domain will include the word which forms the first Immediate Constituent of a mother node, the word which forms the last Immediate Constituent of a mother node and all intervening words (Hawkins 1990:228-229).

So the fewer the words that are involved in processing Immediate Constituents, the shorter the Constituent Recognition Domain is (Hawkins 1990). According to Hawkins (1990:229-230), when there is a choice, the parser favors the syntactic constructions with shorter Constituent Recognition Domain. Here, I adapt one of Hawkins' (1990:228) examples to illustrate his idea. The human parser will opt for *introduce to me the books which are good for my son* rather than *introduce the books which are good for my son to me* (Hawkins 1990). The reasons are as follows. In [<sub>VP</sub> *introduce* [<sub>PP</sub> *to me*] [<sub>DP</sub> *the books which are good for my son*]], the Immediate constituents are V, PP and DP. In order to process all of them, the human parser needs to process words which form the Constituent Recognition Domain, and they are *introduce*, *to* and *the* (Hawkins 1990).<sup>23</sup> So the

---

<sup>23</sup> Processing *the* is sufficient to form a mother node DP, labelled NP in Hawkins (1990).

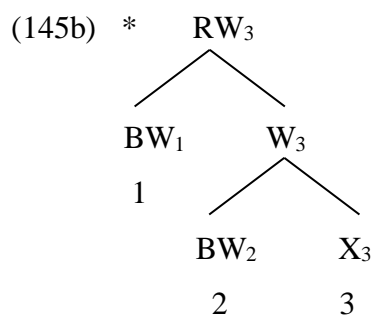


Constituent Recognition Domain contains three words in this case (Hawkins 1990). In comparison, in  $[_{VP} \textit{introduce} [_{DP} \textit{the books which are good for my son} [_{PP} \textit{to me}]]]$ , the human parser needs to process ten words in order to recognize all three Immediate constituents; in other words, the Constituent Recognition Domain here contains ten words, and those ten words are *introduce, the, books, which, are, good, for, my, son, to* (Hawkins 1990). Clearly the Constituent Recognition Domain of  $[_{VP} \textit{introduce} [_{PP} \textit{to me} [_{DP} \textit{the books which are good for my son}]]]$  is shorter, containing less words than that of  $[_{VP} \textit{introduce} [_{DP} \textit{the books which are good for my son} [_{PP} \textit{to me}]]]$  (Hawkins 1990). The VP with the shorter Constituent Recognition Domain is preferred by the human parser (Hawkins 1990).

I will assume that this method of comparing structures and explaining preferences is also applicable to compounds in MSC. Potentially the ungrammaticality of (142) could be a processing effect of Hawkins' type. Note here, (142) may have more than one left bracket, and similarly, with its corresponding grammatical construction, there may be more than one right bracket. Here it is sufficient only to consider the simplest case. That is to say, the ungrammatical structure with one left bracket, that is (145a) and the corresponding grammatical structure with one right bracket, that is (146a), are considered. Now the ungrammatical case is discussed first.

(145a) \*  $W \rightarrow [BW_1[BW_2 X_3]]$

(145a) is demonstrated in the following tree (145b) where RW represents recursive words:

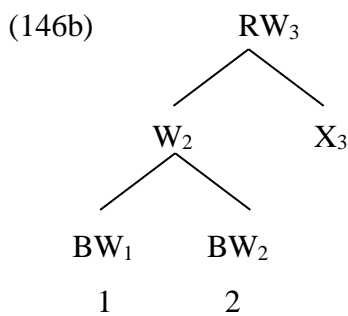


In (145b),  $BW_1$  and  $W_3$  are Immediate Constituents of the mother node  $RW_3$  as  $BW_1$  is

attached to  $RW_3$  and  $W_3$  is the head of  $RW_3$ .<sup>24</sup>  $RW_3$  is the first and only mother node constructed in this construction. In order to recognize those two Immediate Constituents, the human parser has to parse  $BW_1$ ,  $BW_2$  and  $X_3$ . So there are three words in the Constituent Recognition Domain of the mother node  $RW_3$ . In contrast, (146a) is grammatical:

(146a)  $W \rightarrow [[BW_1 BW_2] X_3]$

(146a) is demonstrated in the following tree (146b):



In (146b),  $BW_1$  and  $BW_2$  are two Immediate Constituents of the mother node  $W_2$  as  $BW_1$  is attached to  $W_2$  and  $BW_2$  is the head of  $W_2$ .  $W_2$  is the first mother node which is constructed in (146b). In order to recognize the mother node  $W_2$ , the human parser needs to parse  $BW_1$  and  $BW_2$ . In other words, there are two words in the Constituent Recognition Domain of the mother node  $W_2$ . In order to recognize the mother node  $RW_3$ , only one more word needs to be parsed. Thus the parsing of (146b) needs no Constituent Recognition domain longer than two words, while parsing (145b) needs three words.

This could provide potential insight into what is happening in the Bound Word-Bound Word cases. We can also explain why  $[R [R X]]$  is a grammatical word structure in these terms. This would be because the alternative structure, which would be easier to parse,  $[[R, R] X]$ , is ruled out independently, as two Roots cannot form an attributive compound. We cannot, however, explain why  $[R [BW X]]$  is a possible word structure, alongside  $[[R, BW] X]$ . The explanation in terms of Hawkins' parsing principle is not yet sufficiently fine-grained to distinguish the Bound Word-Bound Word cases from the

---

<sup>24</sup> MSC recursive compounds are attributive compounds and they are thus right-headed.

Root-Bound Word cases. It is, however, the only explanation proposed, to date.

It has been demonstrated that the null hypothesis (124) is false, so a new hypothesis is proposed which states:

(147) Any combination of nominal content items except  $W \rightarrow [BW/ BW X]$  is permitted.

Previously it has been discussed that in MSC, attributive compound nouns, synonymy-compound nouns and parallel compound nouns can merge with different types of items under certain conditions. In addition to this, those types of simple compound nouns can merge with each other to form recursive compound nouns. (147) predicts that combinations of compounds will all be allowed and indeed no ungrammatical examples are observed. That is to say, a Root, a Bound Word, a Minimal Word can combine with each other to form a recursive compound made up of two simple compounds. Some of the combinations are demonstrated in the following examples.

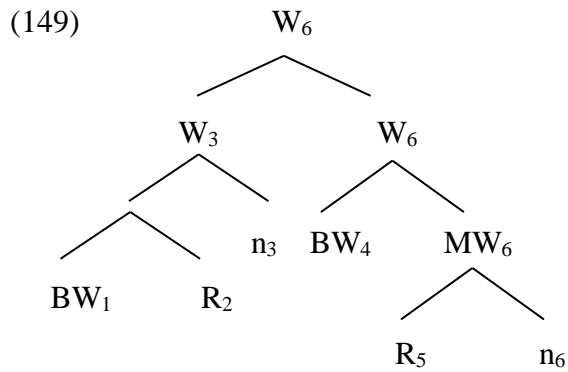
(148) a. [chu fang ] [bi zhi]  
kitchen house wall paper  
'kitchen wallpaper'

b. [shu cai] [bi zhi]  
vegetable vegetable wall paper  
'vegetable wallpaper'

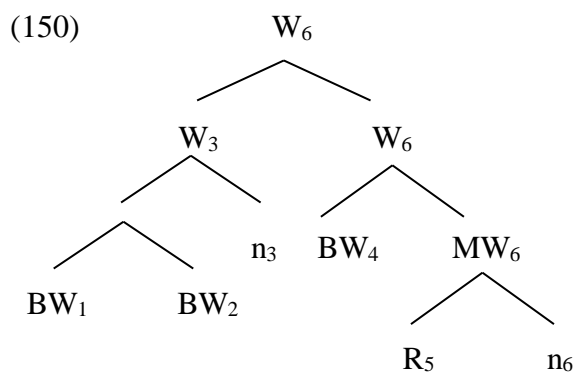
c. [hai yang] [feng jing ]  
ocean ocean wind scenery  
'ocean scenery'

(148a) contains a parallel compound noun *chu fang* 'kitchen house: kitchen' and an attributive compound noun *bi zhi* 'wallpaper'. *Chu* 'kitchen' and *bi* 'wall' are Bound Words and *fang* 'house' is a Root. Inside the parallel compound noun *chu fang* 'kitchen house: kitchen', the categorial head is the null Nominalizer and the semantic head is *chu* 'kitchen'. The head in terms of category and semantics in the attributive compound *bi zhi*

‘wallpaper’ is *zhi* ‘paper’ and *bi* ‘wall’ is clearly the modifier. So *zhi* ‘paper’ here is a Minimal Word. *Bi zhi* ‘wallpaper’ is the head in terms of category and semantics of the recursive compound (148a) *chu fang bi zhi* ‘kitchen wallpaper’. The tree structure of (148a) is as follows:

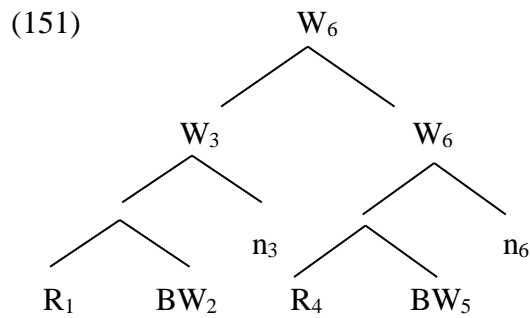


(148b) is a combination of a synonymy compound noun *shu cai* ‘vegetable’ on the left and an attributive compound *bi zhi* ‘wallpaper’ on the right. *Shu* ‘vegetable’, *cai* ‘vegetable’ and *bi* ‘wall’ are Bound Words. For the synonymy compound noun, both *shu* ‘vegetable’ and *cai* ‘vegetable’ are semantic heads. The categorial head is the null Nominalizer. For the attributive compound on the right, *zhi* ‘paper’ is the head in terms of category and semantics and *bi* ‘wall’ is the non-head. So *zhi* ‘paper’ is a Minimal Word. And in (148b), *bi zhi* ‘wallpaper’ is the categorial head as well as the semantic head. (148b) will have the following structure:



In (148c), there is a synonymy compound noun *hai yang* ‘ocean’ and a parallel compound noun *feng jing* ‘wind scenery: scenery’. *Yang* ‘ocean’ and *jing* ‘scenery’ are Bound Words and *hai* ‘sea’ and *feng* ‘wind’ are Roots only. In the synonymy compound noun

*hai yang* ‘ocean’, *hai* ‘ocean’ and *yang* ‘ocean’ are the semantic heads and the null Nominalizer is the categorial head. For the parallel compound noun, *jing* ‘scenery’ is the semantic head and the categorial head is the null Nominalizer. In (148c), *feng jing* ‘scenery’ is the semantic and categorial head. Based the classifications of the items and head locations, the structure of (148c) is as follows:



Above, we have seen that recursion in MSC may involve four items. It is possible that there are more than four items in a recursive compound noun in MSC, the examples are shown below:

(152) a. [chu fang ] [bi zhi] [jia ge]  
 kitchen house wall paper price system  
 ‘kitchen wallpaper price’

b. [hai yang] [feng jing ] [tu pian]  
 ocean ocean wind scenery picture film  
 ‘ocean scenery picture’

The recursive process in MSC is a fully productive option and in principle, recursive compound nouns in MSC can be infinitely large.

#### 4.6 Summary

Following much recent work in generative syntax and morphology, it has been proposed in this chapter that MSC words are derived through the operation Merge (Chomsky 1995:243-246, 2015:223-227) and the Root is category-less (Embick and Noyer

2007,2008; Harley and Noyer 1999; Marantz 1997, 2001). Consequently, Selkirk's rule  $X \rightarrow X^r$  (1982:95) is no longer possible in the current theory, i.e. a Root cannot form a word on its own. Instead, a Minimal (free) Word consists of a Root and a categorizer. It is proposed, therefore, that there are at least two items, two constituents, in a free content word in MSC. It follows that Bound Words in MSC are a class of words which are always bound. The reasons are as follows. Bound Words have inherent category and so cannot merge with the null categorizer, and then the two-constituent condition will not be satisfied and thus they have to be bound. Bound Words can be heads as well as non-heads. It is therefore proposed here that they project optionally, and only as a last resort. A Root cannot be the categorial head, as it lacks a categorial feature. It can in some circumstances be the semantic head, including when it merges with a null categorizer. The null categorizer can only merge with a Root and together they form a Minimal Word in which the null categorizer is necessarily the categorial head. It is proposed that the Minimal Word, consisting of a Root and a null categorizer, can only be a head in terms of category and semantics. It would inherit the obligatory projection property from the categorizer (which has no other function than projecting its categorial feature). In other words, a Minimal Word cannot be a non-head in terms of category and semantics.<sup>25</sup> Some of the suffixes in MSC are shown to be heads in terms of category and semantics, but there are also cases where the suffixes are non-heads. Prefixes on the other hand are all non-heads in terms of category and semantics.

Headedness in a word does not depend on the information of linear order, as traditionally thought. Rather, based on the current theory, headedness determines the linear order,

---

<sup>25</sup> Note here it is possible that in an attributive collocation in MSC, an affixed word (W1) is a non-head. The example is shown below:

(1) *jiao*      *-zi pi*  
       dumpling-ZI wrapper  
       'dumpling wrapper'

*Pi* 'wrapper' can occur alone as a free word in a phrase. Thus on the face of it, it can be a Root or a Minimal Word. (1) denotes a wrapper and is a noun, so categorial features and semantic features of *pi* 'wrapper' percolate to the resultant word *jiao-zi pi* 'dumpling-ZI wrapper: dumpling wrapper'. So *pi* 'wrapper' is a categorial and semantic head. Thus *pi* 'wrapper' can only be a Minimal Word, as a Root is category-less and cannot be a categorial head. This means that the Minimal Word *pi* 'wrapper' can merge with an affixed word and be a head in terms of category and semantics. So, there must be an item in the structure which blocks the features percolation from the affixed word to the dominating node, allowing the Minimal Word *pi* 'wrapper' to have its features percolate. Following Mukai (2008), a linking morpheme is assumed to be in the structure of (1) in order to determine the head.

according to language particular parameter setting (Chomsky 2015:308). Feature percolations determine Headedness: in a word made up of two word components [a,b], if features of *a* percolate to the dominating node, then *a* is the head, the same applies to *b*. The spell-out rule in MSC is that nouns generally have the head spelled out on the right (Packard 2000). One exception to this rule are nouns with certain suffixes which do not affect the category or the interpretation, and which therefore cannot be heads. Another class of exceptions is made up of synonymy-compounds and parallel compounds. Both types of compounds are shown to be symmetrical. More specifically, the categorial head in both types of compounds is a third element which is a null categorizer. The semantic head varies. A synonymy-compound has two semantic heads. One of the components in a parallel compound is the semantic head. Both types of compounds can be any combinations of Root and Bound Word. The antonymy-compounds, the verb-object compounds and the subject-predicate compounds which behave abnormally are analyzed as having a structure where there are two merged category-less Roots and a null categorizer. The meaning of these compounds is typically not compositional, but derives directly from the lexicon. This could be a necessary feature of symmetrical compounds (Bauke 2014; Zhang 2007).

Recursive compound nouns in MSC have also been investigated and there is a rule on attachment on Bound Words in simple recursive compound nouns. More specifically, the unattached Bound Word can only attach to another Bound Word on the right which is its sister. The Bound Word which is not in the sister position cannot be attached to this unattached Bound Word. So there is a linear order preference for Bound Words. The Constituent Recognition Domain principle (Hawkins 1990) may shed some lights on why this is the case.

On the other hand, for complex recursive compound nouns where there are two simple compounds, there are no restrictions. The next chapter concerns the structures of words in OXC; and it will be seen that OXC reduplicated data provide direct empirical support for the current theory.





## Chapter 5 The structure of Old Xining Chinese words

### 5.1 Primitives in OXC common nouns

In chapter 4, I proposed that MSC words are formed through the operation Merge (Chomsky 1995:243-246, 2015:223-227) and proposed a condition for free content words in MSC:

(1) A free content word must consist of at least two items.

A Root (R) can form a smallest free content word, that is, a Minimal Word, with a null categorizer ( $\emptyset$ ) in MSC:

$W \rightarrow [R, \emptyset]$

(2)<sub>[n]</sub>  $\sqrt{\text{SHOU}}$   $n$  (MSC)

hand  $n$

‘hand’

In addition to the null categorizer, the Root in MSC can also merge with other types of items, that is, a Bound Word, an affix or a Minimal Word to form a free word. In addition to the Roots, there are also Bound Words in MSC. As previously discussed in chapter 4, a Bound Word can only form a free word with another item other than the null categorizer. Therefore the possible items which can merge with a Bound Word to form a free word comprise another Bound Word, a Root, a Minimal Word or an affix. The example below shows a Root merged with a Bound Word:

(3) mei mao (MSC)

good looks

‘good looks’

Recall that a Bound Word has an inherent category, which means it can be a categorial head itself and a categorizer is not necessary. In the case of (3), which denotes looks, the

categorial head is the right-hand side item, which is the nominal Bound Word *mao* ‘looks’.

Compared with MSC nouns, reduplication is compulsory in OXC nouns<sup>26</sup>, as shown below:

- (4) a. *ca ca* (OXC)  
‘tea’
- b. *mo mo* (OXC)  
‘steamed bun’
- c. *fei fei* (OXC)  
‘water’

Those reduplicated words can occur alone as free words in phrases. Below it is shown that *fei fei* ‘water’ for instance can be used alone as a noun in an adjective-noun phrase in OXC:

- (5) *being zi feifei* (OXC)  
cold PHRASAL MARKER water  
‘cold water’

Non-reduplicated versions of (4a), (4b) and (4c), are shown below:

- (6) a. *ca* (OXC)  
‘tea’

---

<sup>26</sup> OXC adjectives and verbs can also be reduplicated. The reduplication of adjectives and verbs in OXC has semantic effect and it is not mandatory (please see section 1.2 of chapter 1 for a description of the reduplicated adjectives and verbs in OXC). This is different from reduplicated nouns in OXC. More specifically, the reduplication of nouns does not have semantic effect and it is mandatory.

b. *mo* (OXC)  
'steamed bun'

c. *fei* (OXC)  
'water'

The non-reduplicated versions of (4a,b,c), that is, (6a,b,c) must be bound in OXC, as they cannot occur alone as free words in phrases.<sup>27</sup> For instance, *mo* 'steamed bun' cannot be used as a free noun in the following attributive adjective-noun phrase:

(7) \* *ho zi mo* (OXC)  
good PHRASAL MARKER steamed bun  
'??'

Ren (2006:29) does not entirely agree with this characterization of non-reduplicated forms in Xining Chinese, but claims that non-reduplicated forms do occur as free words. I believe this is due to not keeping OXC and New Xining Chinese distinct. Most speakers are familiar with both varieties, and therefore accept both reduplicated and non-reduplicated forms of nouns as free words. I have consulted people from the old generation. My data have been checked with speakers who are over 70, and who are very clear about the mandatory nature of noun reduplication.<sup>28</sup>

By looking at Minimal Nouns in (2) in MSC and (4a,b,c) in OXC, I hypothesize that the items being reduplicated in (4a,b,c) in OXC are Roots and the reduplicants in (4a,b,c) are null categorizers which have no phonological feature of their own, but copy all the phonological features of sister Roots that they are merged with. That is to say in each case, the null categorizer becomes overt by copying the phonological features of the

---

<sup>27</sup> Later it will be discussed that examples in (6) are Minimal Words. For instance *mo* 'steamed bun' in (6b) is a Minimal Word, which has structure where there is a Root and a null Nominalizer. *Mo* 'steamed bun' which is not reduplicated can merge with an affix, e.g. *-e* and the resultant word is *mo-e* 'steamed bun-E: steamed bun' (please see section 5.4.1 for the discussion of suffixation in OXC). Below it will be discussed that the Minimal Word *mo* 'steamed bun' can undergo reduplication. The resultant word is the Reduplicated Word *mo mo* 'steamed bun n.RED: steamed bun' which contains two constituents.

<sup>28</sup> Furthermore, native speakers were very clear about the reduplication in various types of nouns in chapter 2 and chapter 5. Roots of nouns are all monosyllabic in OXC. The question of reduplication of dissyllabic Roots does not arise. The semantics of nouns makes no difference when it comes to reduplication.

single sister Root that it is merged with. For instance, in (4c), the reduplicant is the null categorizer which copies phonological features of the single sister Root *fei* ‘water’. (4a,b,c) are demonstrated to be able to occur alone as free words in phrases. By contrast, (6a,b,c) also have a structure where there is a Root and null categorizer, but the categorizer does not copy the phonological features of the single sister Root, and (6a,b,c) cannot occur alone as free words in phrases. Based on the hypotheses and those observations, I would claim that OXC nouns follow a stricter version of condition (1), that is:

(8) A free noun must consist of at least two overt items in OXC.

It should be noted here that condition (8) is not a phonological condition on the minimal size of a word (McCarthy and Prince 1990; Hall 1999) in OXC, as adjectives and verbs do not obligatorily reduplicate. Later, evidence from Bound Words in OXC will also show that condition (8) is not a phonological one.

In (6a,b,c), there is only one overt item, thus they are not free nouns, but must be bound. Such words are labelled as Minimal Nouns in OXC in the present work, due to their identical structure to that of the Minimal Nouns in MSC. But the difference is that Minimal Nouns cannot occur alone as free nouns in phrases in OXC, while they can do so in MSC. (4a) (4b) and (4c) are free nouns in OXC and will be called Reduplicated Nouns. Comparing Minimal Nouns of MSC with Reduplicated Nouns of OXC, they all have the same structure where there is a Root and the null Nominalizer, that is,  $W \rightarrow [R, n]$ . In MSC, this null Nominalizer has a null exponent, as one alternative. In OXC, the null Nominalizer copies all phonological features of the single sister Root including the tone. OXC has a system of four lexical tones (Zhang and Zhu 1987). The tone is always copied along with the segments in Reduplicated Nouns, as demonstrated below:

- (9) a. féi féi (OXC)  
       ‘water’
- b. mó mó (OXC)  
       ‘steamed bun’

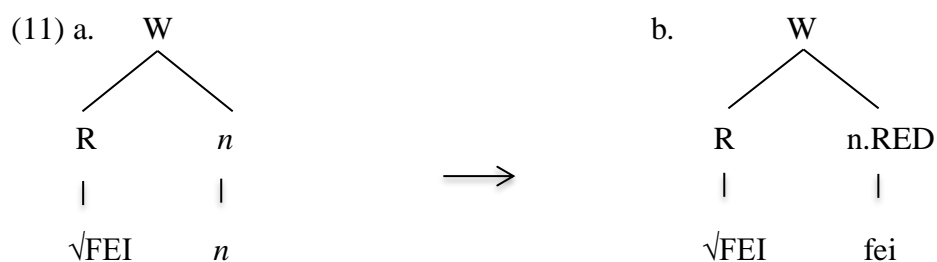
- c. kōng kōng (OXC)  
 ‘chest’

This null Nominalizer with the reduplicated phonological features of the single sister Root in Reduplicated Nouns in OXC is named as a Reduplicated Nominalizer in present theory, represented as n.RED.

Summarizing, a Reduplicated noun in OXC consists of a Root and a Reduplicated Nominalizer and a Minimal Noun in OXC consists of a Root and a null Nominalizer. Comparing both types of nouns, it is clear that the following morphological rule applies in OXC:

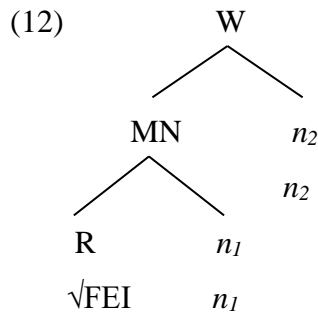
- (10) A null Nominalizer can copy i.e. reduplicate the phonological features of its sister Root.

A Root can merge with the null categorizer in MSC and OXC, so what about the Minimal Word in MSC and OXC? Can they merge with the null categorizer as well? The answer is that the Minimal Word in MSC and OXC cannot merge with the null categorizer. The evidence for such an answer is drawn from Reduplicated Nouns in OXC. We saw that in the reduplication procedure in OXC nouns the null Nominalizer copies the phonological features of the single sister Root. So for example, the Reduplicated Noun *fei fei* ‘water n.RED: water’ in (9a) has the following derivation:



The Nominalizer in (11a) copies all the phonological features of the single sister Root  $\sqrt{\text{FEI}}$  ‘water’ and becomes a Reduplicated Nominalizer (n.RED). Together with the Root  $\sqrt{\text{FEI}}$  ‘water’, they form a Reduplicated Noun as shown in (11b). If a Minimal Word can merge with a null categorizer, then here the left-hand item in *fei* ‘water’ in (11a) could

hypothetically be a Minimal Noun (MN) which is composed of a Root (R)  $\sqrt{\text{FEI}}$  ‘water’ and a null Nominalizer (n):



In MSC we have ruled this structure out on theoretical grounds: the second null categorizer would be superfluous. If the grammar of natural language does not tolerate superfluous items (Chomsky 2015: 24, 118,137), then (12) is ruled out. In OXC we can rule it out on empirical grounds as well. Given condition (10), the structure (12) would produce the word *\*fei fei fei*, with two instances of reduplication. However this is an ungrammatical word. Even granting that reduplication could be optional in a word which already satisfies condition (8), we would predict that *\*fei fei fei* would be an alternative spellout of *fei fei*, but it is not. We thus have empirical evidence from one language for a condition which we have taken to hold in another, related language, where there is arguably never any overt evidence.

This nature of the null categorizer may be language-particular. But it is unknown how a language learner can come to have this knowledge for a particular language. Could the impossibility of the structure (12) be acquired under language acquisition? Conceivably it could be in OXC, if the absence of double reduplication in the input data of the child can count as sufficient indirect negative data (Chomsky 1981:8-9). In MSC, there would be no comparable negative data, so ruling out (12) would presumably not be learnable. So, it is plausible to say that this knowledge concerning the nature of the null categorizer comes from UG and is applicable universally in languages. In other words, condition (13) below which is set for the null categorizer would hold true in both MSC and OXC, and universally. In OXC, there is direct evidence for it.

(13) The null categorizer cannot merge with a Minimal Word, either bound or free.

Recall from section 4.1.3, it has been discussed that the null categorizer cannot merge with the Bound Word either. So the item the null categorizer can merge with can only be a Root in MSC. In other words, in MSC, the following condition would apply:

(14) The null categorizer can only merge with a Root.

Recall from chapter 4 that in MSC, a Bound Word cannot merge with the null categorizer. Are there any Bound Words in OXC common nouns as well? If there are nominal Bound Words, can they merge with the null categorizer, that is, null Nominalizer? It will be discussed below that condition (14) also holds true for OXC nouns. More specifically, a nominal Bound Word in OXC cannot merge with the null Nominalizer.

There are also nominal items in OXC, such as *yi* ‘clothing’, *bo* ‘arm’ and *ta* ‘inner shirt’, which cannot undergo reduplication, as demonstrated below:

- (15) a. \* *yi yi* (OXC)  
           clothing n.RED  
           ‘??’
- b. \* *bo bo* (OXC)  
           arm n.RED  
           ‘??’
- c. \* *ta ta* (OXC)  
           inner shirt n.RED  
           ‘??’

The question is why can reduplication not satisfy (8) with the items in (15) in OXC? This is explained under the analysis shown in (11). Reduplication is copying of the phonological features of a Root by a null Nominalizer in OXC nouns. The ungrammaticality of (15) can be explained if none of the items *yi* ‘clothing’, *bo* ‘arm’ and





another Bound Word, a Root or a Minimal Noun:

(17) a. fei            zi                                    yi            shong                                    (OXC)  
big PHRASAL MARKER clothing clothing  
'big clothing'

b. chong            zi                                    gei            bo                                    (OXC)  
long PHRASAL MARKER armpit arm  
'long arm'

c. fei            zi                                    han            ta                                    (OXC)  
big PHRASAL MARKER sweat inner shirt  
'big inner shirt'

(17a) and (17b) are cases where two nominal Bound Words are merged together to form a compound noun; and in (17c) the Root *han* 'sweat' merges with the nominal Bound Word *ta* 'inner shirt' to form a compound noun. Thus, it is clear that condition (8) does not concern the phonological units, but the number of grammatical components involved in forming a smallest noun in OXC.

Further, it has been argued in chapter 4 that a Minimal Word cannot be a non-head in terms of category and semantics, which is based on the principle of economy. It follows that the non-head of a compound cannot be reduplicated in OXC. So, it is predicted that the non-head can be a Bound Word in OXC. This is indeed the case and it will be shown that the Bound Word can be a non-head in an affixed word or a compound in section 5.4 and 5.5 in this chapter. This supports the assumption that the Bound Word projects optionally (see discussion in section 4.1.3 in chapter 4). Further, In OXC synonymy-compounds and parallel compounds, Bound Words can only be non-heads in terms of category. In other words, categorial features of Bound Words in OXC synonymy-compounds and parallel compounds do not project, because if they did, then it is possible that synonymy-compounds and parallel compounds would end up having the same structure and interpretation as that of the attributive compounds in OXC. Therefore, similar to their counterparts in MSC, Bound Words cannot be categorial heads in synonymy-compounds

and parallel compounds in OXC.

On the other hand, Bound Words can project in OXC attributive compound nouns. More specifically, in such a compound noun, categorial features of the Bound Word could project and following the default rule (10) in chapter 4 which concerns the percolation of semantic features, semantic features of the Bound Word will also percolate, and so the Bound Word can be the head of an attributive compound noun in terms of category and semantics. In short, a Minimal Word cannot be a non-head in terms of category and semantics and a Bound Word can be a head or a non-head in terms of category in OXC. It follows that a Bound Word can merge with a Minimal Word in OXC when the Minimal Word is the categorial and semantic head. A Bound Word can merge with a Root when the Bound Word is the categorial and semantic head in OXC.

So far, we have discussed the nature of the Root, the Minimal Noun, the Reduplicated Noun, the null Nominalizer and the nominal Bound Word in OXC. In addition, affixation is another way of forming nouns in OXC. As in MSC, the grammatical function of affixes in OXC is to form a word by combining with another item. The affix is proposed to be a primitive in OXC word syntax. It will be seen later that in some cases, the affixes are categorizers, but in other cases, they are not.

In short, the Root, the nominal Bound Word, the null Nominalizer and the affix are all proposed to be primitives of OXC nouns. A Root forms a Minimal Noun with the null Nominalizer. There are two connected main differences between a Root and a Minimal Noun in OXC common nouns. First, a Root is category-less and a Minimal Noun bears a word category. Secondly, for this reason a Root can merge with the null Nominalizer and an Minimal Noun cannot do so. In other words, the double reduplication would be ill-formed (*\*fei fei fei*). Further, being a Root means that an item cannot undergo reduplication, but if this item is a Minimal Noun, then it can do so. The difference between a Root and a nominal Bound Word in OXC nouns concerns the null Nominalizer. More specifically, a Root can merge with the null Nominalizer, and consequently it can merge with a null Nominalizer with copied phonological content of the single sister Root. But a nominal Bound Word cannot merge with a null Nominalizer and consequently it cannot merge with a null Nominalizer with copied phonological

content of the Root. A nominal Bound Word is different from a Minimal Noun in the sense that the former cannot have reduplication, but the latter can. These differences will be applied in the following sections to investigate whether an item is a Root, a Minimal Noun or a nominal Bound Word in OXC nouns.

There are also similarities among a Root, a nominal Bound Word and a Minimal Noun, in OXC nouns. First of all, the Root, the Minimal Noun and the nominal Bound Word are bound in the sense that they cannot occur alone as free words in phrases. Further, neither the Minimal Noun nor the nominal Bound Word can merge with the null Nominalizer and consequently neither of them can merge with the null Nominalizer with copied phonological features of the Root. Moreover, neither the Root nor the nominal Bound Word can undergo reduplication. Also, both the Root and the nominal Bound Word can merge with a nominal Bound Word, a Minimal Noun and an affix. Those different types of combinations will all be considered in the later sections concerning the word structure of the affixed nouns and the word structure of compound nouns in OXC.

The differences and similarities concerning the Root, the nominal Bound Word and the Minimal Noun discussed here are summarized in the following two tables. Table 2 concerns the item which a Root, a nominal Bound Word and a Minimal Noun can or cannot merge with. If the combination is allowed then the relevant box is ticked with ‘√’, if the combination is not allowed then the relevant box is left blank:

	Null Nominalizer	Null Nominalizer with copied phonological features of the single sister Root
Root	√	√
Nominal Bound Word		
Minimal noun		

**Table 2 Differences concerning null Nominalizer among Root, Nominal Bound Word and Minimal Noun**

Table 3 demonstrates whether the listed items can be reduplicated and occur alone as free words. If they can undergo reduplication or occur alone as free words, then the relevant box is ticked with ‘√’, if they cannot undergo reduplication or occur alone as free words, then the relevant box is left blank.

	Reduplication	Free words
Root		
Nominal Bound Word		
Minimal Noun	√	

**Table 3 Differences concerning reduplication and the freedom of the constituent among Root, Nominal Bound Word and Minimal Noun**

## 5.2 Primitives in OXC adjectives and verbs

As for the adjectives and verbs in OXC, they are quite similar to their counterparts in MSC. More specifically, condition (1) in chapter 4 which states that there are minimally two items in a free content word in MSC also applies to OXC adjectives and verbs as well, while the counterpart of condition (8) need not be met. That is to say, unlike the smallest OXC nouns in which the null Nominalizer has to be pronounced, the null adjectival categorizer and the null verbal categorizer are not pronounced in the smallest OXC adjectives and verbs. The word types of OXC adjectives and verbs are also the same as their counterparts in MSC: they are affixed words, Minimal Words and compounds. Hence, I would propose that the primitives for MSC words are also valid for OXC adjectives and verbs, which means that the Root, the Bound Word, the null categorizer and the affix are primitives for adjectives and verbs in both MSC and OXC. The previously discussed difference between a Minimal Word and a Bound Word in a MSC also holds in OXC adjectives and verbs, which is that a Minimal Word can occur alone as a free word in a phrase and a Bound Word cannot do so, because a bare Minimal Word satisfies condition (1) stated in chapter 4, but a bare Bound Word does not, because it has inherent category. The Bound Word can be a head or a non-head and the Minimal Word cannot be a head; this is also assumed to hold in OXC adjectives and verbs. Condition



form a word, as it has been established in section 4.1.3 of the last chapter, that based on the principle of economy, a Minimal Noun cannot be a non-head. There are also similarities between a Reduplicated Noun and a Minimal Noun. The categorial heads in both types of nouns are null Nominalizers. The semantic head is the Root in both Reduplicated Nouns and Minimal Nouns in OXC.

Those differences and similarities between Minimal Nouns and Reduplicated Nouns in OXC are summarized in table 4, showing what the Minimal Noun and the Reduplicated Noun may contain (either the null Nominalizer with the null exponent or the null Nominalizer with copied phonological content of the sister Root), whether the Minimal Noun and the Reduplicated Noun can occur alone as free words, and what the heads are in the Minimal Noun and the Reduplicated Noun. The relevant box is ticked with ‘√’, otherwise the box is left blank:

	Containing null Nominalizer with the null exponent	Containing null Nominalizer with copied phonological features of the single sister Root	Free word	The null Nominalizer is the categorial head and the Root is the semantic head
Minimal Noun	√			√
Reduplicated Noun		√	√	√

**Table 4 Differences and similarities between Minimal Nouns and Reduplicated Nouns**

For a Minimal Adjective and Verb in OXC, their structures are identical to their counterparts in MSC, consisting of a Root (R) and a null categorizer (Ø), that is, a null adjectival categorizer and a null verbal categorizer respectively. The structure of, for example, the adjective *jieng* ‘clever’ is (19a) and the structure of the verb *sao* ‘disturb’ is (19b).

$W \rightarrow [{}_c R, \emptyset]$

(19) a. [<sub>a</sub> √JIENG a] (OXC)  
clever a  
'clever'

b. [<sub>v</sub> √SAO v] (OXC)  
disturb v  
'disturb'

In terms of the head locations, similar to MSC, the null categorizers are the categorial heads of Minimal Adjectives and Verbs in OXC, as their categorial features percolate to Minimal Adjectives and Minimal Verbs in OXC. The semantic heads are the Roots as their semantic features percolate to Minimal Adjectives and Minimal Verbs in OXC.

#### 5.4 Simple affixed words in OXC: structure and head location

Affixed words in OXC are analyzed in this section and they are described in section 2.1.4 and 2.1.5. Similarly to their counterparts in MSC, affixed words in OXC consist of a content item plus the affix. Thus the null categorizer and the affix cannot merge to form an affixed word in OXC, nor can two affixes. The content item can be a Root (R), a Bound Word (BW), or a Minimal Word (MW) in OXC affixed words. This indicates that affixed words in OXC may have the following possible structures:

The structures of suffixed words:

- (20) a.  $W \rightarrow [BW, AFF]$   
b.  $W \rightarrow [R, AFF]$   
c.  $W \rightarrow [[R, \emptyset] AFF]$

The structures of prefixed words:

- (20) d.  $W \rightarrow [AFF, BW]$   
e.  $W \rightarrow [AFF, R]$   
f.  $W \rightarrow [AFF [R, \emptyset]]$

In the following section, affixed words in OXC will be discussed to see whether they have those structures. Suffixed words are discussed first.

### 5.4.1 Suffixation

This section concerns suffixes in OXC, please see section 2.1.5 for a description. The suffix *-zi* is one of the commonly used suffixes in OXC to form nouns. It is meaningless:

(21) *za -zi* (OXC)  
powder- ZI  
'powder'

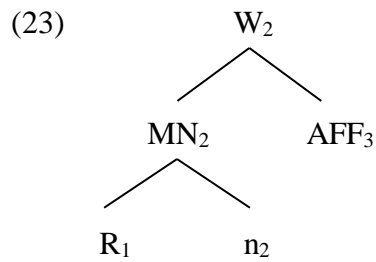
*Za* 'powder' in (21) can be reduplicated, as shown below:

(22)  $\sqrt{ZA}$  *za -zi* (OXC)  
powder n.RED -ZI  
'powder'

This means that *za* 'powder' must be an item where the null Nominalizer is present; in other words, it cannot be a Bound Word or a Root. It can only be a Minimal Noun in which there is a Root  $\sqrt{ZA}$  and the null Nominalizer. Otherwise (22) should be ungrammatical.

*Za* 'powder' is a Minimal Noun, so it is the categorial head and semantic head of (21). This prediction is based on our discussion in chapter 4. More specifically, it has been discussed that the Minimal Word cannot be a non-head, based on the principle of economy: if there is a choice between Root and Minimal Word, the Root is always chosen. Hence a Minimal Word is only found in position where a Root is impossible, that is where a head is needed. Thus, the Minimal Noun *za* 'powder' is the categorial head and semantic head of (21) and the suffix *-zi* is the non-head, so (21) will have the structure  $W \rightarrow [_n [_n R n]AFF]$  and the corresponding tree structure is demonstrated below, in which MN represents the Minimal Noun:





This also means that *-zi* is not just semantically empty, but also lacks a categorial feature, excluding the possibility that it would merge with a Root. If it did, the result would be a category-less construct. Further, by comparing (21) and (22), it can be seen that the reduplication is optional when there are already two overt items in an OXC noun and one of the items is a Minimal Noun. In other words, when an OXC noun meets condition (8), the noun can still be grammatical without being reduplicated. Note here the principle of economy used in the thesis does not rule out the optional reduplication. This is because economy is defined in terms of the presence or absence of structure, not of phonological material.

The non-head suffix *-zi* can also merge with a Bound Word. *Nei* ‘milk’ in (24) is such an item.

- (24) *nei* -zi (OXC)  
 milk -ZI  
 ‘milk’

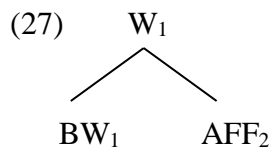
Classifying *nei* ‘milk’ as a nominal Bound Word is justified because firstly, it cannot occur alone as a free noun in a phrase; note that *zi* in (25) is the phrasal marker *zi*, not the suffix.

- (25) \* *xiong* zi nei (OXC)  
 delicious PHRASAL MARKER milk  
 ‘??’

Secondly, *nei* ‘milk’ cannot be reduplicated:

- (26) \* *nei nei* (OXC)  
 milk n.RED  
 ‘??’

This property of *nei* ‘milk’ can be explained if it has inherent word category and cannot merge with a null Nominalizer. As a result, it cannot have its phonological content copied. Thus condition (8) in this chapter which states that there are at least two overt items in a free noun of OXC will not be met, and *nei* ‘milk’ must always be bound. This classification of *nei* ‘milk’ predicts that (24) has no corresponding reduplicated form, which is true, as the reduplicated word \**nei nei -zi* is ungrammatical. The suffix *-zi* is a non-head in terms of category and semantics, which means that its categorial features and semantic features will not project, so it is predicted that categorial features and semantic features of the Bound Word *nei* ‘milk’ will percolate. The head of (24) in terms of category and semantics is the Bound Word *nei* ‘milk’. Thus (24) has the structure of  $W \rightarrow [_n [_n BW] AFF]$  and below its tree structure is shown:



It has been seen that the non-head suffix *-zi* can merge with a Minimal Noun and a Bound Word. The prediction is that it could also merge with compound nouns. This is indeed the case. (28a) is a synonymy-compound noun<sup>29</sup> and in (28b), the suffix *-zi* is seen to be able to merge with such a compound noun:

- (28) a. *yi shong* (OXC)  
 clothing clothing  
 ‘clothing’

<sup>29</sup> The structure of synonymy-compound nouns in OXC will be expanded more in section 5.5.3.1 of this chapter.

b. *yi shong -zi* (OXC)  
 clothing clothing-ZI  
 ‘clothing’

Both *yi* ‘clothing’ and *shong* ‘clothing’ are Bound Words, as firstly neither of them can occur alone as free words in phrases:

(29) a. \* *fei zi yi* (OXC)  
 big PHRASAL MARKER clothing  
 ‘??’

b. \* *sou zi shong* (OXC)  
 small PHRASAL MARKER clothing  
 ‘??’

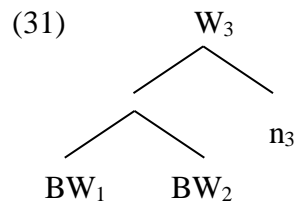
Further, both *yi* ‘clothing’ and *shong* ‘clothing’ cannot be reduplicated:

(30) a. \* *yi yi* (OXC)  
 clothing n.RED  
 ‘??’

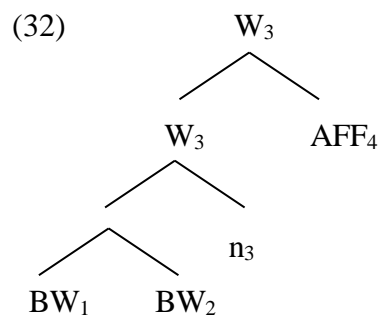
b. \* *shong shong* (OXC)  
 clothing n.RED  
 ‘??’

Those facts can be understood if both *yi* ‘clothing’ and *shong* ‘clothing’ have inherent category, so neither of them can merge with the null Nominalizer. So neither of them can have their phonological features copied. Thus condition (8) is violated. And both *yi* ‘clothing’ and *shong* ‘clothing’ are always bound. *Yi shong* ‘clothing’ in (28a) is a synonymy-compound noun. Recall in section 5.1, it has been assumed that categorial features of Bound Words in synonymy-compound nouns in OXC do not project, so Bound Words cannot be categorial heads in synonymy-compound nouns in OXC. Thus

here, following Zhang (2007), a null categorizer is assumed to be the categorial head for the synonymy-compound noun *yi shong* ‘clothing’, here *yi shong* ‘clothing’ is a noun, so the categorial head is the null Nominalizer. In terms of semantic heads, semantic features of both the Bound Word *yi* ‘clothing’ and the Bound Word *shong* ‘clothing’ percolate to *yi shong* ‘clothing’, thus both Bound Words are semantic heads of *yi shong* ‘clothing’ in (28a). Thus *yi shong* ‘clothing’ in (28a) will have the following structure:



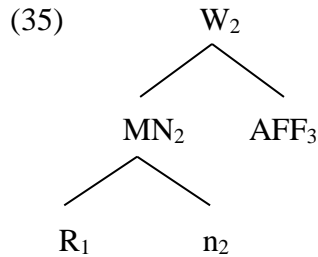
In (28b), the suffix *-zi* is a non-head in terms of category and semantics, and the synonymy-compound noun *yi shong* ‘clothing’ is the head in terms of category and semantics. Thus, the structure of (28b) is as follows:



Above it has been shown that the suffix *-zi* can merge with various kinds of nouns. The suffix *-zi* is demonstrated to be the non-head. Its redundancy is clearly shown in (28a) and (28b) where it merges with a compound noun. More specifically, it can be seen that without the suffix *-zi*, the compound noun *yi shong* ‘clothing’ is still grammatical. In other words, the suffix *-zi* is optional in this case. But when *-zi* is merged with a Minimal Noun and a Bound Word as shown in (21) and (24) respectively, its presence is to fulfill the requirement set in condition (8) and in that respect, the suffix *-zi* is not redundant in those two cases.



Word cannot be a non-head in terms of category and semantics. Thus *mo* ‘steamed bun’ is the head in terms of category and semantics, so (33a) will have the structure  $W \rightarrow [{}_n [{}_n R n] AFF]$ , and the tree structure is:



Comparing (33a) and (34), the reduplication is optional in (33a). This is because condition (8) is met in (33a).

In (33b), the non-head suffix *-e* is merged with the item *sou* ‘minced meat’. The item *sou* ‘minced meat’ is identified as a Bound Word because first, it is ungrammatical as a free noun in an adjective-noun phrase:

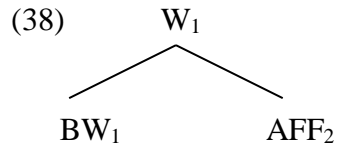
(36) \* guei            zi                                    sou                                    (OXC)  
 expensive PHRASAL MARKER minced meat  
 ‘??’

Further, *sou* ‘minced meat’ cannot be reduplicated:

(37) \* sou                                    sou                                    (OXC)  
 minced meat n.RED  
 ‘??’

This behaviour of *sou* ‘minced meat’ can be understood if it cannot merge with the null Nominalizer due to the reason that it has word category inherently. As a result, *sou* ‘minced meat’ is unable to have its phonological content copied. Therefore, condition (8) is not satisfied and *sou* ‘minced meat’ is bound. The suffix *-e* is a non-head and so its features will not project, thus categorial features and semantic features of the Bound

Word *sou* ‘minced meat’ will project to (33b) and thus it is the categorial and semantic head of (33b). Therefore, the structure of (33b) is  $W \rightarrow [_n [_n BW] AFF]$  and its corresponding tree structure is:



(33c) is a case where an attributive compound noun<sup>30</sup> *ca-ji* ‘tea table’ is merged with the non-head suffix *-e*. *Ji* ‘table’ is a Bound Word, as firstly, it cannot be used as a free noun in a phrase:

(39) \**guei*      *zi*                      *ji*    (OXC)  
 expensive PHRASAL MARKER table  
 ‘??’

Further, *ji* ‘table’ cannot be reduplicated:

(40) \* *ji*      *ji*    (OXC)  
 table    n.RED  
 ‘??’

This observation can be understood if *ji* ‘table’ has inherent word category and thus cannot merge with the null Nominalizer. As a result, *ji* ‘table’ cannot be reduplicated. Thus *ji* ‘table’ violates condition (8) and is always bound. *Ca* ‘tea’ also cannot undergo reduplication in *ca ji* ‘tea table’ in (33c), as shown below:

(41) \* *ca* *ca*      *ji*    (OXC)  
 tea n.RED table  
 ‘??’

---

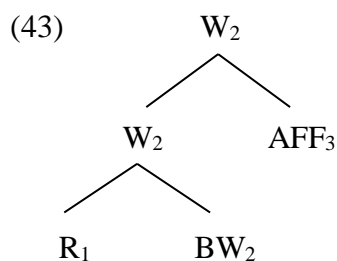
<sup>30</sup> Attributive compound nouns are analyzed in section 5.5.1.

This ungrammaticality is not caused by the non-reduplicated *ji* ‘table’, as it is a Bound Word which cannot undergo reduplication, so it can only be caused by the reduplication of *ca* ‘tea’ in (41). In other words, *ca* ‘tea’ cannot be reduplicated in (41) and so it cannot be a Minimal Noun, since otherwise, (41) should be grammatical. This leaves two possible identifications for *ca* ‘tea’: one is the Bound Word and the other is the Root. *Ca* ‘tea’ can be reduplicated alone:

(42) *ca ca* (OXC)  
*tea n.RED*  
‘tea’

The fact that (42) is grammatical means *ca* ‘tea’ can only be a Root, which can merge with a null Nominalizer and consequently undergo reduplication. *Ca* ‘tea’ cannot be a Bound Word, as Bound Words cannot merge with the null categorizer and consequently cannot be reduplicated.

The Root *ca* ‘tea’ cannot be the categorial head in *ca ji* ‘tea table’, as a Root is category-less, so the Bound Word *ji* ‘table’ is the categorial head of *ca-ji* ‘tea table’. Based on the default rule of percolations of semantic features in section 4.1.3, the semantic features of the Bound Word *ji* ‘table’ will also percolate, so *ji* ‘table’ is the semantic head. Thus *ji* ‘table’ is the categorial and semantic head of *ca-ji* ‘tea table’. The structure of (33c) will be  $W \rightarrow [{}_n [{}_n [R [{}_n BW] AFF]]$  and the tree structure is shown below:



Like the suffix *-zi*, the suffix *-e* in OXC can be redundant, especially in the case where it





- b. \* rou rou -dan (OXC)  
 meat n.RED -person  
 ‘??’

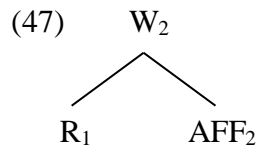
This would mean that *xiong* ‘countryside’ and *rou* ‘meat’ cannot be Minimal Nouns, otherwise, (45a) and (45b) should be grammatical. This leaves two options for *xiong* ‘countryside’ and *rou* ‘meat’, which are the Bound Word and the Root. Both *xiong* ‘countryside’ and *rou* ‘meat’ can merge with the null Nominalizer with copied phonological features of the single sister Root:

- (46) a. *xiong* xiong (OXC)  
 countryside n.RED  
 ‘countryside’

- b. rou rou (OXC)  
 meat n.RED  
 ‘meat’

(46a) and (46b) being grammatical implies that *xiong* ‘countryside’ and *rou* ‘meat’ can only be Roots. They are Roots, so they are able to merge with the null Nominalizer and consequently are able to have their phonological features copied. They cannot be nominal Bound Words, as nominal Bound Words cannot merge with the null Nominalizer and as a result they cannot have their phonological contents copied.

The fact that *xiong* ‘countryside’ and *rou* ‘meat’ here can only be Roots means that they cannot be categorial heads. It is predicted that the suffixes *-bong* ‘person’ and *-dan* ‘person’ are categorial heads in (44a) and (44b) respectively. Based on the default rule of percolation of semantic features in section 4.1.3 in the last chapter, the semantic features of the suffixes *-bong* ‘person’ and *-dan* ‘person’ will also percolate respectively, so they are semantic heads in (44a) and (44b) as well: the words refer to a kind of person. Thus, the suffix *-bong* ‘person’ and *-dan* ‘person’ are categorial and semantic heads in (44a) and (44b) respectively. Thus, the structure of (44a) and (44b) is  $W \rightarrow [{}_n[R[{}_n \text{AFF}]]]$  and the tree structure is shown below:



Comparing the cases above involving the redundant suffixes *-e* and *-zi* with the cases here involving the suffix *-bong* ‘person’ and *-dan* ‘person’, it is clear that in the suffixed nouns in OXC, the item can be reduplicated if the suffix it is merged with is the non-head; but the item cannot be reduplicated if the suffix it is merged with is the head.

In addition to object-denoting items, the suffix *-bong* ‘person’ and *-dan* ‘person’ which are the heads in terms of category and semantics can also merge with property-denoting items, and together they form nouns:

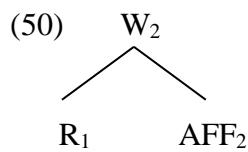
- (48) a. *nong-bong* (OXC)  
 stupid-person  
 ‘idiot, blockhead’
- b. *jian-dan* (OXC)  
 sly- person  
 ‘a sly person’

As discussed earlier, reduplication is not obligatory in OXC adjectives. They are thereby similar to MSC adjectives. Thus condition (1) in chapter 4 which states that there are at least two items, not necessarily overt, in a free content word in MSC is also valid in OXC adjectives. The difference between a Minimal Word and a Bound Word in MSC also applies to OXC adjectives: A Minimal Word can occur alone as a free word in a phrase, but a Bound Word cannot do so.

In (48a), if *nong* ‘stupid’ can be a Root, then it can form a Minimal word. This is indeed true. Below, it is shown that it can occur alone as an adjective in an adjectival phrase:

- (49) hudu nong (OXC)  
 very stupid  
 ‘very stupid’

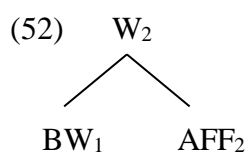
In (48a), the categorial and the semantic head is the suffix *-bong*, thus *nong* ‘stupid’ is the non-head. This means that in (48a), *nong* ‘stupid’ can only be a Root, as it is established in chapter 4 that a Minimal Word cannot be the non-head. The structure of (48a) is thus  $W \rightarrow [{}_n [R[{}_n \text{AFF}]]]$  and the following structure is the tree structure:



*Jian* ‘sly’ in (48b) is a Bound Word as it cannot occur as a free word in a phrase:

- (51) \* hudu jian (OXC)  
 very sly  
 ‘??’

This fact about *jian* ‘sly’ can be explained if it has inherent word category and cannot merge with the null categorizer, more specifically, here a null adjectival categorizer. Thus condition (1) in chapter 4 is not met and it is always bound. The categorial head and the semantic head of (48b) is the suffix *-dan* ‘person’. So it is expected that features of the Bound Word *jian* ‘sly’ do not project and thus the Bound Word *jian* ‘sly’ is a non-head. It is permitted for Bound Words, as discussed in section 4.1.3, Bound Words project optionally. The structure of (48b) is thus  $W \rightarrow [{}_n [{}_a \text{BW}][{}_n \text{AFF}]]$  and the tree structure is demonstrated below:



There are also suffixes which are commonly used to form verbs in OXC. *-Gei* is such a suffix, exemplified below:

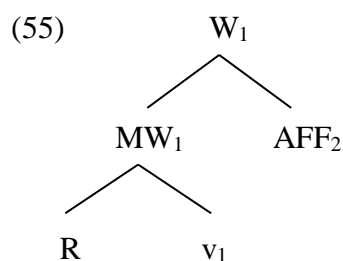
- (53) *ju-gei* (OXC)  
 make-GEI  
 ‘make’

Verbs are not reduplicated obligatorily in OXC. They are thereby similar to MSC verbs. Condition (1) of MSC which states that there are two items in a free content word of MSC is also applicable to verbs in OXC. The difference between a Minimal Word and a Bound Word in MSC also applies to OXC verbs.

As shown below, *ju* ‘make’ can be used alone as a verb in a sentence:

- (54) *Jia tong -ha ju liao.* (OXC)  
 she soup CASE make PERFECTIVE  
 ‘She has made soup.’

Based on feature percolation facts, the categorial head and the semantic head of (53) is the item *ju* ‘make’ and it can only be a Minimal Word to be such a head. Thus, the structure of (53) is  $W \rightarrow [{}_v [{}_v R v] AFF]$  and below is the tree structure:<sup>32</sup>



<sup>32</sup> In addition, a different suffix *-gei* is observed in OXC. It is a causative suffix, which merges with a compound adjective, for instance, *fu tan -gei* ‘comfortable flat -make: make somebody to feel comfortable’.

### 5.4.2 Prefixation

There are also prefixes which are used to form nouns in OXC (please see section 2.1.4 for a description). Some examples of such nouns are given below.

(56) a-ji (OXC)  
A-sister  
'sister'

(57) a-jiou (OXC)  
A-uncle  
'uncle'

*Ji* 'sister' can be reduplicated in (56), so it must be a Minimal Noun:

(58) a-ji ji (OXC)  
A-sister n.RED  
'sister'

As a Minimal Noun, *ji* 'sister' will be the categorial head and semantic head of (56), since as discussed in section 4.1.3, a Minimal Word cannot be a non-head. So the prefix is the non-head which has no effect on the syntactic or semantic properties of the word. So (56) will have the following structure  $W \rightarrow [_n [AFF [_n R n]]]$ .

In (57), the non-head prefix *a-* is merged with *jiou* 'uncle'. *Jiou* 'uncle' is a Minimal Word because it can be reduplicated in (57):

(59) a-jiou jiou (OXC)  
A-uncle n.RED  
'uncle'

It has been discussed in section 4.1.3 that a Minimal Word cannot be a non-head. Thus in (57) *jiou* ‘uncle’ which is a Minimal Word is the categorial and semantic head. The structure of (57) is  $W \rightarrow [_n [AFF [_n R n]]]$ .

Yet another prefix is *ga-*. Like *a-*, it has no effect on the syntactic or semantic properties of the resulting word, the example is shown below:

(60) *ga-hu* (OXC)  
       GA-box  
       ‘box’

The prefix, therefore, cannot be the head of the word, so *hu* ‘box’ must be the head. As such it must be either a nominal Bound Word or a Minimal Noun. It is not a Bound Word, as it can be reduplicated in (60):

(61) *ga- hu hu* (OXC)  
       GA- box n.RED  
       ‘box’

This indicates that *hu* ‘box’ can only be a Minimal Noun. The structure for (60) is thus  $W \rightarrow [_n [AFF[_n R n]]]$ .

As seen above, similar to previously discussed prefixes in MSC, the prefix in OXC is not a head in terms of category and semantics. It is interesting to compare this with English, where, as observed by Williams (1981: 248-249), prefixes are not heads (*un-happy* is an adjective while *un-tie* is a verb; *re-surge* is a verb while *re-surgence* is a noun etc.).<sup>33</sup> As mentioned in section 4.4.2, this is also the case for Italian where the prefixes do not determine the word category of the resultant word, for instance, the iterative prefix *ri-* in *rivendere* ‘to resell’ (Talamo and Celata 2011:3). Similarly, Spanish prefixes do not affect the word category, for instance, *des-* ‘un-’ in *desplegar* ‘unfold’ (Miguel et al. 2017:38).

---

<sup>33</sup> There are a few exceptions, though: The prefix *en-* in *enlarge* and *enrich* appears to be a verbal categorial head (Williams 1981:249).

The suffixes may or may not determine the word category of the resultant word in OXC, as discussed in section 5.4.1. Suffixes in OXC behave similarly to suffixes in MSC and many other languages, for example, German and Italian (for details, see section 4.4.1). Further, there are more suffixes in comparison with prefixes in OXC. This is consistent with what has been observed in MSC (Lin 2001:61) and other languages in the world (Bybee 1990).

In sum, it has been shown that for affixed nouns, structures (20a), (20b), (20c), (20d) and (20f) are possibilities, the reason for ruling out (20e) is that a Root cannot merge with a prefix which is a non-head. As for affixed verbs, (20b) is eliminated, as the suffix in the affixed verb is a non-head, so it cannot merge with a Root. Further, I have not found any example of a verb which is a Bound Word. The theory does not rule out such a word, though, so this could be an accidental gap. So (20a) could be a structure for affixed verbs in OXC. (20c) is a possible structure for affixed verbs in OXC.

## **5.5 Compounding in OXC: the structure and head locations**

Below, the simple compound nouns in OXC are analyzed, which include attributive compound nouns, synonymy-compounds, parallel compounds, antonymy-compounds, verb-object compounds and subject-predicate compounds. For a description of these types of compounds, please see section 2.2.

### ***5.5.1 Simple attributive compound nouns in OXC***

This section concerns OXC attributive compounds, which are described in section 2.2.2. Similar to MSC, simple attributive compound nouns in OXC can be formed of two overt items. Given that, among the items proposed for OXC nouns above, the Root, the Minimal Word and the Bound Word can form such compound nouns in OXC. The expectation is that the following structures would all be possible ordered structures for simple attributive compound nouns in OXC, if there are no constraints.

(62) a.  $W \rightarrow [BW, BW]$







Minimal Word cannot be a non-head, based on the principle of economy. Thus the structure of (63a) is  $W \rightarrow [n_2 [n_1 BW] [n_2 R n]]$ . (63b) will have the following structure  $W \rightarrow [n [a BW] [n R n]]$

There are also Attributive compounds in OXC formed of a Root and a Bound Word in that order:

attributive noun-noun compound

- (68) a. *shou jieng* (OXC)  
 hand towel  
 ‘hand towel’

attributive adjective-noun compound

- (68) b. *dan yi* (OXC)  
 thin clothing  
 ‘thin clothing’

*Jieng* ‘towel’ in (68a) is a Bound Word, as firstly it cannot occur as a noun in an attributive adjective-noun phrase:

- (69) \* *wan zi* *jieng* (OXC)  
 soft PHRASAL MARKER towel  
 ‘??’

Further, *jieng* ‘towel’ cannot be reduplicated:

- (70) \* *jieng jieng* (OXC)  
 towel n.RED  
 ‘??’

This observation concerning *jieng* ‘towel’ can be explained if *jieng* ‘towel’ cannot occur with the null Nominalizer. As a result, *jieng* ‘towel’ cannot be reduplicated. So condition (8) is not fulfilled and *jieng* ‘towel’ is bound. In (68a), *shou* ‘hand’ cannot undergo

reduplication:

(71) \*shou shou jieng (OXC)  
hand n.RED towel  
'??'

This ungrammaticality is not caused by the non-reduplicated form *jieng* 'towel', as it is a Bound Word and cannot be reduplicated. That (71) is ungrammatical implies that *shou* 'hand' cannot be a Minimal Noun, as otherwise, (71) should be grammatical, which is not the case. So two options are left for *shou* 'hand', which are Root or Bound Word. *Shou* 'hand' can merge with the null Nominalizer with copied phonological content of the sister Root:

(72) shou shou (OXC)  
hand n.RED  
'hand'

This compatibility indicates that *shou* 'hand' can only be a Root, as the Bound Word cannot merge with the null Nominalizer and as a result, cannot be reduplicated.

So (68a) is composed of a Root and a Bound Word. As a Root does not bear word category, it cannot be a categorial head. Thus it is predicted that categorial features of the Bound Word *jieng* 'towel' will percolate, and it is the categorial head. Based on the default rule (10) of percolation of semantic features in chapter 4, which states that semantic features and categorial feature percolate together, semantic features of the Bound Word *jieng* 'towel' in (68a) will also percolate to the compound noun in (68a) which immediately dominates it. Thus the Bound Word *jieng* 'towel' is also the semantic head as well as being the categorial head. So the structure is  $W \rightarrow [_n R [_n BW]]$  for (68a).

Below it is shown that *dan* 'thin' in (68b) can occur alone as a free adjective in an adjectival phrase, so it can be a Minimal Word made up of a Root and a null categorizer:

(73) hudu dan (OXC)  
 very thin  
 ‘very thin’

On the other hand, *yi* ‘clothing’ is a Bound Word, as firstly, it cannot stand alone as a noun in an adjective-noun phrase:

(74) \* xieng zi yi (OXC)  
 new PHRASAL MARKER clothing  
 ‘??’

Secondly, *yi* ‘clothing’ cannot be reduplicated:

(75) \*yi yi (OXC)  
 clothing n.RED  
 ‘??’

This can be explained if *yi* ‘clothing’ has word category inherently and thus cannot merge with the null Nominalizer. As a result, it cannot have its phonological feature copied. Thus *yi* ‘clothing’ cannot fulfill condition (8) and is bound. As a Minimal Word, if *dan* ‘thin’ merges with the Bound Word *yi* ‘clothing’, then the Minimal Word *dan* ‘thin’ would be the head in terms of category and semantics. It is expected that (68b) should be an adjective, denoting a property. However, (68b) is a noun and denotes a kind of clothing. Thus *dan* ‘thin’ cannot be a Minimal Word and it cannot be the head in terms of category and semantics of (68b), but can only be a Root. It is predicted that categorial and semantic features of the Bound Word *yi* ‘clothing’ percolate to the node which immediately dominates *yi* ‘clothing’ and *dan* ‘thin’. Thus *yi* ‘clothing’ is the categorial and semantic head in (68b). The structure for (68b) is  $W \rightarrow [_n R [_n BW]]$ .

An OXC attributive compound noun may consist of a Root and a Minimal Noun in that order:

attributive noun-noun compound

- (76) a. mo ku (OXC)  
wool trousers  
'wool trousers'

attributive adjective-noun compound

- (76) b. fei jiou (OXC)  
big foot  
'big foot'

*Ku* 'trousers' can be reduplicated in (76a):

- (77) mo ku ku (OXC)  
wool trousers n.RED  
'wool trousers'

Thus, *ku* 'trousers' can only be a Minimal Noun. It cannot be a Root or a Bound Word, since otherwise, the grammaticality of (77) is left unexplained. On the other hand, (76a) does not have the following reduplicated form:

- (78) \*mo mo ku ku (OXC)  
wool n.RED trousers n.RED  
'??'

*Ku* 'trousers' is a Minimal Noun, so it is expected that it can be reduplicated. So the ungrammaticality in (78) is not caused by the reduplication of *ku* 'trousers'. The ungrammaticality is due to the reason that *mo* 'wool' is reduplicated. In other words, the ungrammaticality in (78) shows that *mo* 'wool' cannot be reduplicated. Thus *mo* 'wool' cannot be a Minimal Noun, since Minimal Nouns can undergo reduplication. Thus there are two possible identifications for *mo* 'wool': one is a Root, the other is a Bound Word. *Mo* 'wool' can occur reduplicated as a free word in a phrase. Thus it is not a Bound Word. It follows that it must be a Root in (76a) and (77).









condition, one would expect it to apply to words in OXC in general ('noun' is not a phonological category). The second reason concerns the reduplication of nominal Bound Words. To be specific, if condition (8) were a phonological condition, and reduplication were a means to satisfy this phonological condition, then reduplication of nominal Bound Words would be predicted to be grammatical in OXC. However, Bound Words cannot undergo reduplication.

In addition to those two reasons, the distribution of reduplication in OXC attributive compound nouns also supports the argument that condition (8) is not a phonological condition. As we have seen, only the item on the right-hand side of an attributive compound noun in OXC can be reduplicated, and the item on the left-hand side cannot do so. This can be understood if condition (8) is a morphosyntactic condition. In other words, the distribution of reduplication of nouns in OXC can be explained in terms of morphosyntax. More specifically, the right-hand side item is a head, which means it has to be an item which can project. So this item can be a Minimal Noun. This explains why reduplication for the right-hand item is possible. But for the left-hand item, which is a non-head, which means it can only be an item which does not project, and so it can only be a Root or a Bound Word. This explains the non-reduplication of the left-hand side item. It is unclear how this kind of distribution of reduplication could be explained in phonological terms.

### *5.5.3 Coordinate compounds*

#### *5.5.3.1 Synonymy-compound nouns*

Similar to MSC, there are also synonymy-compound nouns and parallel compound nouns in OXC. The synonymy-compound nouns are discussed first (see 2.2.3 for a description).

(87) beng    geng (OXC)  
       root    root  
       'root'

*Beng* ‘root’ is a Bound Word, confirmed by the fact that firstly it cannot occur as a noun in a phrase:

(88) \* wengti zi beng (OXC)  
problem PHRASAL MARKER root  
‘??’

Further, *beng* ‘root’ cannot be reduplicated:

(89) \* beng beng (OXC)  
root n.RED  
‘??’

This can be understood if *beng* ‘root’ has word category inherently and cannot merge with the null Nominalizer, and so it cannot have its phonological content copied. Thus *beng* ‘root’ fails to meet condition (8) and is bound. In the case of *geng* ‘root’, on the other hand, it cannot be reduplicated in (87), so it cannot be a Minimal Noun:

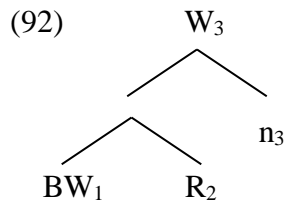
(90) \* beng geng geng (OXC)  
root root n.RED  
‘??’

*Geng* ‘root’ can only be a Root or a Bound Word. The fact that *Geng* ‘root’ can merge with the null Nominalizer with the copied phonological content of the sister Root means that it cannot be a Bound Word and it can only be a Root:

(91) geng geng (OXC)  
root n.RED  
‘root’

Semantic features of the Bound Word *beng* ‘root’ and the Root *geng* ‘root’ all percolate to the compound, thus they are both semantic heads in (87). It is discussed in section 5.1 that categorial features of Bound Words do not project in synonymy-compounds, so the Bound Word *beng* ‘root’ cannot be the categorial head. (87) is a noun, following Zhang

(2007), a null Nominalizer is assumed here to be the categorial head of (87) and the structure is as follows:



A synonymy-compound noun may be composed of a Root and a Bound Word, in that order.

(93) *bo geng* (OXC)  
 neck neck  
 ‘neck’

*Geng* ‘neck’ is a Bound Word, as firstly it cannot occur alone as a free word in a phrase, such as an adjective-noun phrase:

(94) \* *chong zi* *geng* (OXC)  
 long PHRASAL MARKER neck  
 ‘??’

Secondly, *geng* ‘neck’ cannot be reduplicated:

(95) \* *geng geng* (OXC)  
 neck n.RED  
 ‘??’

This can be understood if *geng* ‘neck’ cannot merge with the null Nominalizer, so *geng* ‘neck’ cannot have its phonological content copied. Thus condition (8) is violated, and *geng* ‘neck’ will always be bound. As for *bo* ‘neck’, it cannot be reduplicated in (93), as shown in (96). Note here, the ungrammaticality of (96) is not caused by the non-

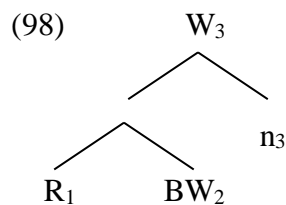
reduplicated form *geng* ‘neck’, as *geng* ‘neck’ is a Bound Word, which cannot be reduplicated. Thus, *bo* ‘neck’ cannot be a Minimal Noun:

- (96) \* *bo bo geng* (OXC)  
 neck n.RED neck  
 ‘??’

There are two possibilities left for *bo* ‘neck’: one is the Root and the other is the Bound Word. Since *bo* ‘neck’ can merge with the null Nominalizer with copied phonological content of the sister Root, it can only be a Root:

- (97) *bo bo* (OXC)  
 neck n.RED  
 ‘neck’

Based on semantic feature percolation facts, both the Root *bo* ‘neck’ and the Bound Word *geng* ‘neck’ are semantic heads. The Bound Word *geng* ‘neck’ cannot be the categorial head, as it is discussed in section 5.1 that the Bound Word does not project its categorial features in synonymy-compound nouns in OXC. Thus following Zhang (2007), I assume that there is a null Nominalizer in the structure which is the categorial head:



There are also synonymy-compound nouns in which both items may be category-less Roots:

- (99) *gueng bong* (OXC)  
 stick stick  
 ‘stick’

Neither *gueng* ‘stick’ nor *bong* ‘stick’ can be reduplicated in (99), which can be explained if neither *gueng* ‘stick’ nor *bong* ‘stick’ are Minimal Nouns.

(100) a. \**gueng gueng bong* (OXC)  
 stick n.RED stick  
 ‘??’

b. \**gueng bong bong* (OXC)  
 stick stick n.RED  
 ‘??’

Taking into consideration the fact that both of them can merge with the null Nominalizer with copied phonological content of the sister Root, *gueng* ‘stick’ and *bong* ‘stick’ in (99) are thus Roots, which are category-less:

(101) a. *gueng gueng* (OXC)  
 stick n.RED  
 ‘stick’

b. *bong bong* (OXC)  
 stick n.RED  
 ‘stick’

Zhang’s (2007) analysis, which involves two merged acategorial Roots merged with a head (a null Nominalizer *n*) is appropriate in this case. The Nominalizer is the categorial head, while the two merged Roots are the semantic heads.<sup>34</sup> The structure is shown below:

---

<sup>34</sup> Here the semantic heads are Roots, as the meaning of the resultant compound nouns are compositional. If the meaning of the resultant compound where there are two merged Roots and a head is not compositional, then the meaning may come from the lexicon, as seen earlier in the cases like verb-object compounds in MSC in chapter 4 in which the Roots cannot be semantic heads. In short, the Root may or may not contribute the meaning to the resultant compound, which has also been discussed by Marantz (2007).







b. *no* *gua* (OXC)  
 brain melon  
 ‘brain’

*Kou* ‘mouth’ in (107a) and *no* ‘brain’ in (107b) are Bound Words as they cannot occur as free nouns in phrases, and neither of them can be reduplicated:

(108) a. \* *kou* *kou* (OXC)  
 mouth n.RED  
 ‘??’

b. \* *no* *no* (OXC)  
 brain n.RED  
 ‘??’

This can be understood if neither *kou* ‘mouth’ and *no* ‘brain’ have word category and thus cannot merge with the null categorizer, and as a result, neither can have their phonological content copied, thus condition (8) is violated and *kou* ‘mouth’ and *no* ‘brain’ are bound always.

*Lian* ‘face’ in (107a) on the other hand cannot be reduplicated in (107a), neither can *gua* ‘melon’ in (107b). So they cannot be Minimal Nouns, which leaves two options for them: one is the Root and the other is the Bound Word. Since both of them can merge with the null Nominalizer with copied phonological features of the sister Root, they are only Roots, not Bound Words:

(109) a. *lian* *lian* (OXC)  
 face n.RED  
 ‘face’

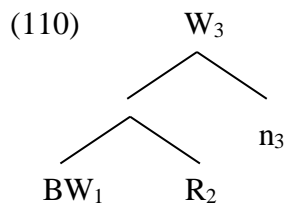
b. gua gua

(OXC)

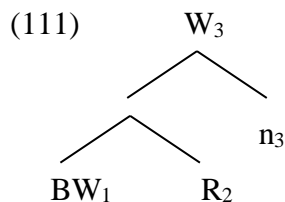
melon n.RED

‘melon’

In (107a), the semantic features of the Root *lian* ‘face’ percolate to the resultant compound which immediately dominates it. Therefore, the Root *lian* ‘face’ is the semantic head. Categorical features of the Bound Word do not project in parallel compounds, as discussed in section 5.1, so following Zhang (2007), a null Nominalizer is assumed here to be the categorial head and the structure for (107a) is as follows:



As for (107b), the Bound Word *no* ‘brain’ is the semantic head as its semantic features percolate to the resultant compound in (107b) which immediately dominates it. Following Zhang (2007), a null Nominalizer is assumed to be the categorial head and thus the structure for (107b) is demonstrated below.



Parallel compound nouns may be composed of a Root and a Bound Word:

(112) a. *lian no*

(OXC)

face brain

‘face’

b. hueng yan (OXC)  
 marriage fate  
 ‘fate’

*No* ‘brain’ in (112a) and *yan* ‘fate’ in (112b) are Bound Words, as they cannot occur alone as free nouns and neither can they be reduplicated:

(113) a.\* no no (OXC)  
 brain n.RED  
 ‘??’

b.\* yan yan (OXC)  
 fate n.RED  
 ‘??’

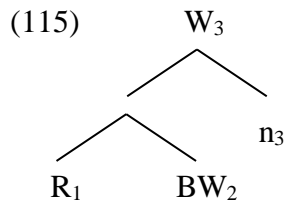
This can be understood if *no* ‘brain’ and *yan* ‘fate’ have inherent category and cannot merge with the null Nominalizer, and as a result, they cannot have their phonological features copied. So condition (8) is not met and both items can only be bound. On the other hand, *lian* ‘face’ cannot be reduplicated in (112a), neither can *hueng* ‘marriage’ be reduplicated in (112b), thus they cannot be Minimal Nouns. They may be Roots or Bound Words. As they can merge with the null Nominalizer with copied phonological content of the sister Root, they are only Roots:

(114) a. lian lian (OXC)  
 face n.RED  
 ‘face’

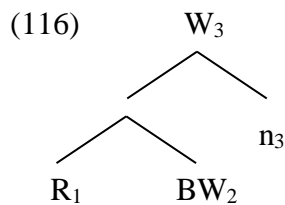
b. hueng hueng (OXC)  
 marriage n.RED  
 ‘marriage’

In (112a), the semantic features of the Root *lian* ‘face’ percolate to the resultant compound which immediately dominates it. Therefore, the Root *lian* ‘face’ is the semantic

head. The structure for (112a) is as follows: following Zhang (2007), the null Nominalizer is the categorial head in the structure:



In (112b), the semantic head is the Bound Word *yan* ‘fate’, as its semantic feature percolate to the compound (112b) which immediately dominates it. So the structure of (112b) is as follows: following Zhang (2007), the categorial head is the null Nominalizer in the structure:



Parallel compound nouns may also be formed of Roots:

(117) a. *nian wu* (OXC)  
           eye pit  
           ‘eye’

      b. *pi mo* (OXC)  
           leather fur  
           ‘fur’

*Nian* ‘eye’, *wu* ‘pit’ cannot be reduplicated in (117a) and *pi* ‘leather’ and *mo* ‘fur’ cannot undergo reduplication in (117b). Thus, they cannot be Minimal Nouns. They are Roots, as they can merge with the null Nominalizer with the copied phonological content of the sister Root:

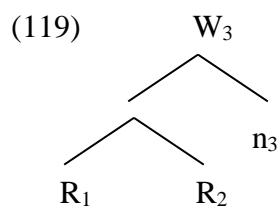
(118) a. nian nian (OXC)  
 eye n.RED  
 ‘eye’

b. wu wu (OXC)  
 pit n.RED  
 ‘pit’

c. pi pi (OXC)  
 leather n.RED  
 ‘leather’

d. mo mo (OXC)  
 fur n.RED  
 ‘fur’

(117a) and (117b) thus contain two acategorial Roots. Therefore, Zhang’s (2007) root merger analysis may be applied to (117a) and (117b). (117a) and (117b) are nouns, so a Nominalizer *n* is assumed to be the categorial head. In terms of the semantic head of (117a), it is clear that the Root *nian* ‘eye’ is the semantic head; for (117b), the Root *mo* ‘fur’ is the semantic head. The structure for (117a) and (117b) is demonstrated below:



There are parallel compounds which contain two Bound Words:

(120) a. gou jia (OXC)  
 dirt scab  
 ‘dirt’

- b. *ban* *jieng* (OXC)  
 board neck  
 ‘neck’

The above items cannot occur alone as free nouns and be reduplicated:

- (121) a. \* *gou* *gou* (OXC)  
 dirt n.RED  
 ‘??’

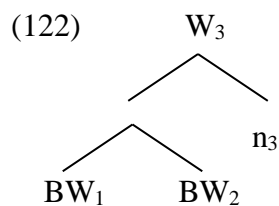
- b. \* *jia* *jia* (OXC)  
 scab n.RED  
 ‘??’

- c. \* *ban* *ban* (OXC)  
 board n.RED  
 ‘??’

- d. \* *jieng* *jieng* (OXC)  
 neck n.RED  
 ‘??’

The ungrammaticalities can be understood if *gou* ‘dirt’, *jia* ‘scab’, *ban* ‘board’ and *jieng* ‘neck’ all have inherent word category and therefore cannot merge with the null Nominalizer. As a result, they cannot have their phonological content copied, thus condition (8) is violated. And those items can only be bound.

In (120a), semantic features of the Bound Word *gou* ‘dirt’ percolate to the compound which immediately dominates it. Thus, the Bound Word *gou* ‘dirt’ is the semantic head. In (120b), it is that semantic features of the Bound Word *jieng* ‘neck’ percolate to the compound. So the Bound Word *jieng* ‘neck’ is the semantic head. In both cases, the null Nominalizer is the categorial head, following Zhang (2007), so they will have the following structure:



Above it has been observed that in synonymy-compound nouns and parallel compound nouns in OXC, the items involved cannot be Minimal Nouns. This provides direct empirical evidence for the following condition which is assumed in chapter 4:

(123) Items in synonymy-compound nouns and parallel compound nouns cannot be Minimal Nouns.

This condition is proposed to be valid in synonymy-compound adjectives and parallel compound adjectives as well. This condition is based on the assumption that reduplication for both synonymy-compound nouns and parallel compound nouns in OXC is possible only to meet condition (8) which states that there are minimum two overt items in a free noun of OXC.<sup>36</sup>

### 5.5.3.3 *Synonymy-compound adjectives and parallel compound adjectives*

This section concerns the compounding structures in adjectives in OXC (please see section 2.2.3 for a description), which are quite similar to their counterparts in MSC. The principles are the same: the components of a synonymy-compound adjective or a parallel compound adjective cannot be the Minimal Word and the categorial features of the Bound Word do not project. Following Zhang (2007), the null adjectival categorizer is the categorial head in both types of compounds. The semantic heads are both components in

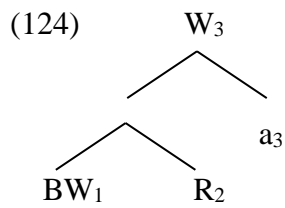
---

<sup>36</sup> This is different from previously discussed attributive compound nouns in OXC where reduplication of the head is optional even though condition (8) is met. This difference between those two types of compounds are not explained in the current theory. Further, there are exceptional cases where parallel compound nouns in OXC are formed of nouns. For example, the parallel compound *bo bo dan dan* 'bag n.RE.DAN n.RE.DAN: bag' is formed of the Reduplicated Noun *bo bo* 'bag n.RE.DAN: bag' and the Reduplicated Noun *dan dan* 'egg n.RE.DAN: egg'.

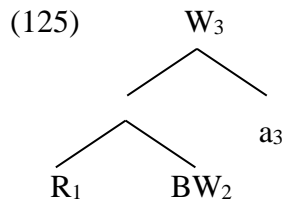
synonymy-compound adjectives and in parallel compound adjectives, only one of the component is the semantic head. Below, examples of synonymy-compound adjectives and parallel compound adjectives in OXC and their structure are demonstrated:

5.5.3.3.1 *Synonymy-compound adjectives in OXC*

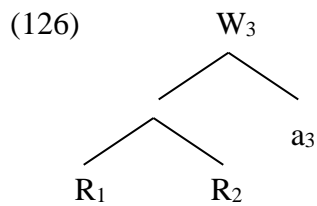
*Nong zong* ‘dirty’ will have the following structure:



*Zuen qio* ‘beautiful’ will have the following structure:

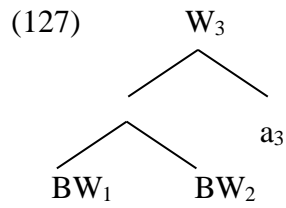


*Mian wan* ‘soft’ will have the following structure:



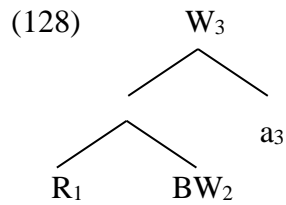
*Ji si* ‘solid’ will have the following structure:



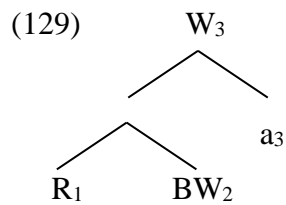


5.5.3.3.2 *Parallel compound adjectives in OXC*

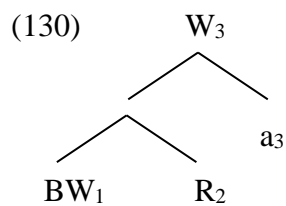
*Da meng* ‘big fierce: fierce’ will have the following structure where the semantic head is the Bound Word *meng* ‘fierce’:



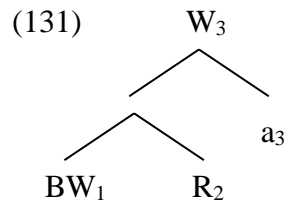
*Ran ma* ‘slow confused: slow’ will have the following structure where the semantic head is the Root *ran* ‘slow’:



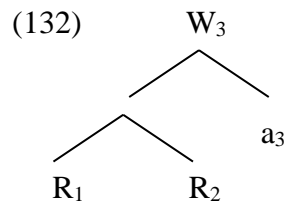
*Dian huong* ‘crazy panic: panic’ will have the following structure where the semantic head is the Root *huong* ‘panic’:



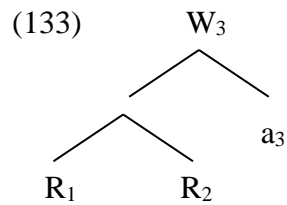
*Li kou* ‘untidy mean: untidy’ will have the following structure where the semantic head is the Bound Word *li* ‘untidy’:



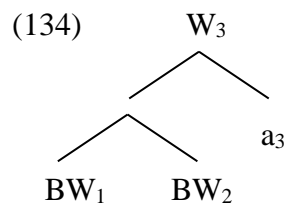
*Lan gan* ‘lazy dry: lazy’ will have the following structure where the semantic head is the Root *lan* ‘lazy’:



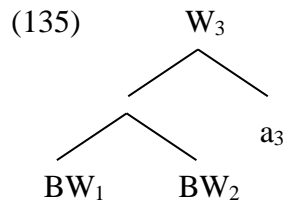
*Huong zo* ‘confused restless: restless’ will have the following structure where the semantic head is the Root *zo* ‘restless’:



*Fu tan* ‘comfortable flat: comfortable’ will have the following structure where the semantic head is the Bound Word *fu* ‘comfortable’:



*Ma li* ‘numerous nimble :nimble’ will have the following structure where the semantic head is the Bound Word *li* ‘nimble’:

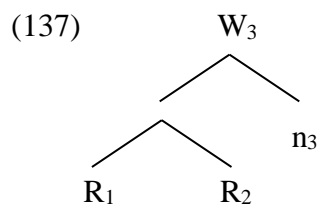


#### 5.5.4 Remaining compounds: Antonymy-compounds, verb-object compounds and subject -predicate compounds

Antonymy-compounds, verb-object compounds and subject -predicate compounds are discussed in this section (please see section 2.2.4 for a description). Similar to antonymy-compounds in MSC, OXC antonymy-compounds have idiomatic meanings and their category is not derivable from that of their components:

(136) fei sou (OXC)  
 fat thin  
 ‘size’

(136) is clearly an exocentric compound, following Zhang (2007), such a compound will be accounted for by Root Merger analysis. More specifically, (136) is formed of two merged Roots<sup>37</sup> and a head, here (136) is a noun, so the head is *n*:




---

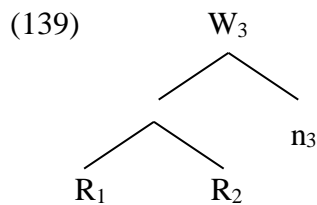
<sup>37</sup> As mentioned in section 3.5, in Zhang’s (2007) analysis, the set {R<sub>1</sub>, R<sub>2</sub>} is labelled as R<sub>1</sub>+R<sub>2</sub>. But this label is not necessary, as R<sub>1</sub> and R<sub>2</sub> do not project. So in this section, the set {R<sub>1</sub>, R<sub>2</sub>} is not labelled.

Similar to verb-object compounds and subject-predicate compounds in MSC, verb-object compounds and subject-predicate compounds in OXC are also in various word category. Verb-object compounds in OXC are demonstrated below:

(138) a. wei bo (OXC)  
 enclose neck  
 ‘scarf’

b. guong zuei (OXC)  
 wander mouth  
 ‘freeload’

(138a) is a noun and (138b) is a verb. Free projectivity is clearly shown in verb-object compounds of OXC. Thus Zhang’s (2007) Root Merger analysis is followed and both examples have a structure where there are two merged Roots and a head. (138a), which is a noun, will have the following structure where the head is  $n$ <sup>38</sup>:

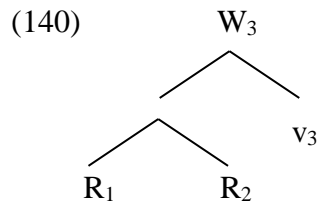


(138b) is a verb, so it will have the following structure where the head is  $v$ :

---

<sup>38</sup> Note here the verb-object compound in OXC is the result of object-denoting item incorporation in OXC. But the incorporated object-denoting item in OXC can only be a non-reduplicated item. The reasons are as follows. It has been discussed in this section that verb-object compounds in OXC have a structure where there are two merged Roots and an external categorizer. In other words, both components in a verb-object compound are Roots in OXC. This predicts that neither of the components in a verb-object compound in OXC can be reduplicated. This is indeed the case. Below it is shown that neither of the components in (138a) can be reduplicated:

- (1) a. \*wei wei bo  
 enclose n.RED neck  
 ‘??’  
 b. \*wei bo bo  
 enclose neck n.RED  
 ‘??’

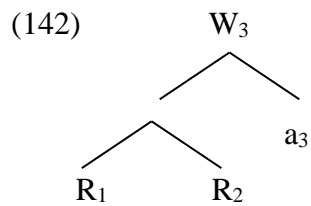


The subject-predicate style compounds in OXC are illustrated below:

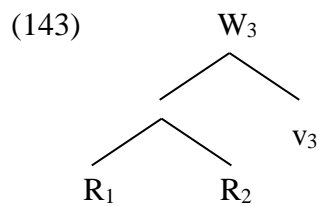
(141) a. nian guei (OXC)  
 word precious  
 ‘reserved’

b. nian hei (OXC)  
 eye black  
 ‘hate’

(141a) is an adjective and (141b) is a verb. Free projectivity is demonstrated in both examples. Thus Zhang’s (2007) analysis will apply here. (141a) will have the following structure where the head is *a*:



(141b) will have the structure where the head is *v*:



### 5.5.5 *Absence of recursive nominal compounds in OXC: some explanations*

The following is an important difference between MSC and OXC: MSC allows either left-branching or right-branching recursive compounding as seen in section 2.2.5, OXC does not. MSC grammatical recursive compounds and their corresponding ungrammatical recursive examples in OXC are demonstrated below:

MSC grammatical recursive left-branching compound nouns

- (144) a. [[ling kou] se] (MSC)  
collar rim colour  
'colour of rim of collar'
- b. [[yan jing] se] (MSC)  
eye eyeball colour  
'eye colour'
- c. [[yi fu] tu] (MSC)  
clothing clothing picture  
'clothing picture'

MSC grammatical recursive right-branching compound nouns

- (145) a. [bo [chen yi]] (MSC)  
thin inner shirt clothing  
'thin inner shirt'
- b. [xiao [yan jing]] (MSC)  
small eye eyeball  
'small eye'
- c. [da [yi fu]] (MSC)  
big clothing clothing  
'big clothing'

OXC corresponding ungrammatical recursive left-branching examples:

- (146) a. \* [[lieng hu] sei] (OXC)  
collar rim colour  
'??'
- b. \* [[nian wu] sei] (OXC)  
eye pit colour  
'??'
- c. \* [[yi shong] tu] (OXC)  
clothing clothing picture  
'??'

OXC corresponding ungrammatical recursive right-branching examples:

- (147) a. \* [dan [han ta]] (OXC)  
thin sweat inner shirt  
'??'
- b. \* [ga [nian wu]] (OXC)  
small eye pit  
'??'
- c. \* [kuan [yi shong]] (OXC)  
big clothing clothing  
'??'

This difference concerning recursion is not unique to Chinese, it is cross-linguistic.<sup>39</sup> For example, according to Ralli (2013:95), it is rare to find left-branching recursive

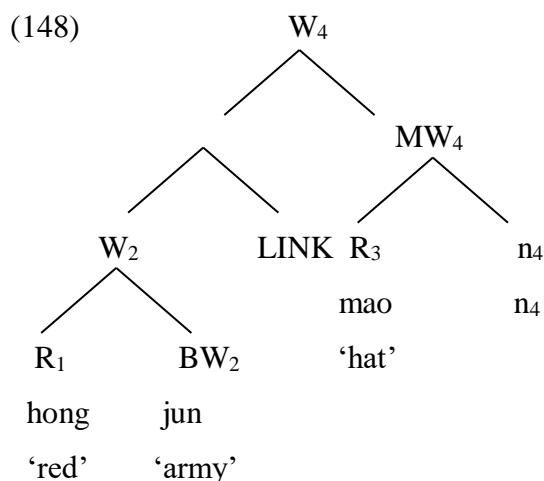
---

<sup>39</sup> See Mukai's (2008) paper for more details.

compound nouns in Greek.

In MSC, allowing left-branching recursion presupposes that nouns can be non-heads of recursive compounds. But the head will not be unambiguously determined if two nouns merge. This indicates that there would be an abstract linking morpheme in MSC. The idea of abstract linking morpheme is adopted from the theory of linking morpheme proposed by Mukai (2008). Recall earlier in chapter 4, according to Mukai (2008:187,193), for the left-branching recursive compounds, there is a linking morpheme which ensures that the heads of the left-branching recursive compounds are located. For languages that do not allow left-branching recursive compounding, such as Latin, she (2008:193ff) proposes that they do not have the linking morpheme.

Mukai's linking morpheme may explain why there are left-branching recursive compound nouns in MSC. More specifically, it is assumed here that there is an abstract linking morpheme in MSC to distinguish the head from the non-head in the situation where a word is composed of two words. The linking morpheme assumed here would have to be abstract because there are no overt linking morphemes observed in MSC compounds. The MSC left-branching recursive nominal compound *hong jun mao* 'red army hat' is demonstrated in the following structure where there is an abstract linking morpheme:



*Jun* 'army' in the inner attributive compound noun *hong jun* 'red army' cannot be used alone as a free word in a phrase, so it is a Bound Word. The Bound Word *jun* 'army' is the categorial and semantic head of the inner attributive compound noun *hong jun* 'red



army'. *Hong* 'red' can be used alone as a free word in a phrase. This means *hong* 'red' can on the face of it be a Minimal Word or a Root. As *hong* 'red' is a non-head, it can only be a Root. *Mao* 'hat' can be used alone as a free word in a phrase. This means *mao* 'hat' can on the face of it be a Minimal Word or a Root. *Hong jun mao* 'red army hat' denotes a kind of hat and it is a noun, which can be explained if categorial features and semantic features of *mao* 'hat' percolate to *hong jun mao* 'red army hat' which immediately dominates *mao* 'hat'. Thus *mao* 'hat' is the head in terms of category and semantics and it can only be a Minimal Word.

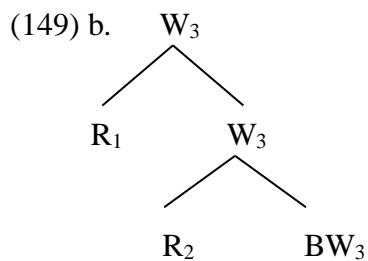
Thus *hong jun mao* 'red army hat' is a combination of a compound noun *hong jun* 'red army' and a Minimal Word *mao* 'hat'. In the structure of *hong jun mao* 'red army hat', the compound noun *hong jun* 'red army' and the Minimal Word *mao* 'hat' will percolate their features if there is not a mechanism to stop them. If they both percolate their features, then *hong jun mao* 'red army hat' will be double-headed. This is not the case: *hong jun mao* 'red army hat' is single-headed. Thus an abstract linking morpheme (LINK) is assumed here to check features of the non-head and so only the features of the head can percolate. More specifically, the abstract linking morpheme will check categorial features and semantic features on the compound word *hong jun* 'red army' ( $W_2$ ), so categorial features and semantic features of the compound word *hong jun* 'red army' ( $W_2$ ) will not percolate to the resultant compound word *hong jun mao* 'red army hat'. Only unchecked categorial features and semantic features of the Minimal Word *mao* 'hat' ( $MW_4$ ) will percolate to the resultant compound word *hong jun mao* 'red army hat' (see Mukai (2008) for details of the checking operation).

For OXC, where no recursive nominal compounds can be observed, one would assume that the reason for such an absence is because there is no abstract linking morpheme in OXC. Indeed, the abstract linking morpheme may explain the absence of left-branching recursive compound nouns where the non-heads are compound nouns in OXC. No abstract linking morpheme in OXC means that in left-branching recursive nominal compounds in OXC, the heads are not determined, thus left-branching recursive nominal compounds do not occur in OXC. Note here, in this chapter it has been observed that the non-heads in compound nouns can only be bound items, that is, Roots or Bound Words. In other words, non-heads cannot be free nouns. This observation is supported by the

finding that there are no left-branching recursive compound nouns in OXC. We have now discussed left-branching compound nouns in both MSC and OXC, and attention will turn to the right-branching compound nouns in both languages.

We have already seen that there are also right-branching recursive compound nouns where the non-heads are Roots in MSC. The example and its structure are demonstrated below:

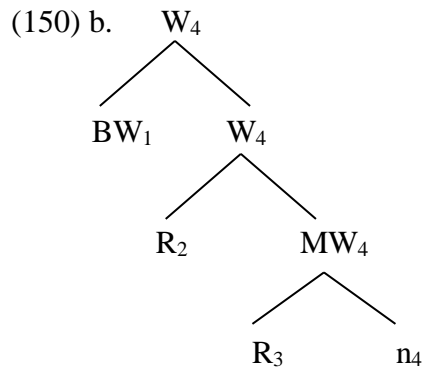
- (149) a. [zang [shui guan ]](MSC)  
 dirty water pipe  
 ‘dirty water pipe’



The fact that this kind of recursive compound noun is allowed in MSC would not be because of an abstract linking morpheme blocking feature percolation. Roots have no category and thus have no categorial features to project. In other words, the heads can be determined in right-branching recursive compounds without assuming the abstract linking morpheme.

Right-branching recursive compounds where the recursive non-head is a Bound Word are also observed in MSC. An example and its structure are shown below:

- (150) a. mu [men kuang](MSC)  
 wood door frame  
 ‘door frame made of wood’



Similarly to the above right-branching recursive compounds where the recursive non-head is a Root, the grammaticality of such recursive compounds are not due to the presence of an abstract linking morpheme. The head is determined in (150b) without the abstract linking morpheme. More specifically, as discussed in section 4.1.3 in chapter 4, Bound Words percolate optionally and only when necessary. Roots are acategorial with no categorial features to percolate. In other words, neither features of the Bound Word nor the features of the Root will percolate to the resultant compound word in (150b). This means they are non-heads in (150b).

As seen above, those two types of right-branching recursive compounds in MSC do not need the abstract linking morpheme to help to identify the head location, as the non-heads are Roots or Bound Words and the heads are determined in each case. Thus for OXC in which the abstract linking morpheme is absent, it is expected that those two types of right-branching compounds should be licit in OXC. However, they are not observed in OXC. In other words, the absence of right-branching nominal compounds where non-heads are Roots or Bound Words in OXC is not due to an undetermined head location, and assuming the absence of the abstract linking morpheme in their structures cannot explain this.

However, right-branching recursive compound nouns where non-heads are Bound Words are not permitted in MSC, as discussed in section 4.5.4. One of the examples and its structure are given below:

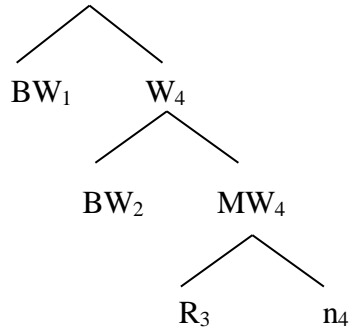
(151) a. \* mu [jing kuang]

(MSC)

wood mirror frame

Intended reading: 'mirror frame made of wood'

(151) b.\* W<sub>4</sub>



In MSC this kind of recursive compound noun is not allowed, so it is expected that they are banned in OXC too.

## 5.6 Reduplicated nouns in OXC: a summary

As seen in the previous sections, components in simple nouns, affixes nouns and compound nouns in OXC can be reduplicated. In section 5.1, it has been discussed that the structure of the Reduplicated Noun is as follows:

(152) [R, n.RED]

It has been seen in the previous sections that the Minimal Noun is the head, which can merge with the affix which is the non-head in OXC. The Minimal Noun can undergo reduplication and become a Reduplicated Noun. Thus a Reduplicated Noun can also merge with the affix which is the non-head. The following structures are predicted for affixed words where there are Reduplicated Nouns:

(153) a. [R, n.RED; AFF]

b. [AFF; R, n.RED]

There are affixes which are non-heads in terms of category and semantics in OXC as shown in section 5.4, for instance, the redundant suffix *-zi* and the redundant prefix *ga-*. They are illustrated in the following examples. So (153) is indeed valid.

(154) a. *xia xia -zi* (OXC)  
           *box n.RED -ZI*  
           ‘box’

b. *ga- xia xia* (OXC)  
       *GA- box n.RED*  
       ‘box’

The Minimal Noun is the head in OXC as demonstrated in previous sections, so it can merge with Root or a Bound Word to form nouns. The Minimal Noun can be reduplicated and become a Reduplicated Noun, so the Reduplicated Noun can merge with a Root or a Bound Word to form nouns too. (155) are predicted structures for compound nouns in OXC in which there are Reduplicated Nouns.

(155) a. [ R; R, n.RED]  
           b.[ BW; R, n.RED]

Those two structures are indeed valid, as demonstrated in the following examples:

(156) a. *ti hu hu* (OXC)  
           *iron box n.RED*  
           ‘iron box’

b. *yi jia jia* (OXC)  
       *clothing shelf n.RED*  
       ‘clothing shelf: shelf for clothing’

*Hu* ‘box’ can be reduplicated in (156a), so *hu* ‘box’ is a Minimal Noun, and *hu hu* ‘box’ is a Reduplicated Noun. *Ti* ‘iron’ cannot be reduplicated in (156a), so it can be a Root or

a Bound Word, as it can merge with the null Nominalizer and as a result, it can have its phonological features copied and it can only be a Root. Thus the structure for (156a) is  $W \rightarrow [{}_n R [{}_n R \text{ n.RED}]]$ . So (156a) is valid.

*Jia* ‘shelf’ can undergo reduplication in (156b), so *jia* ‘shelf’ is a Minimal Noun and *jia jia* ‘shelf’ is a Reduplicated Noun. *Yi* ‘clothing’ in (156b) cannot be used alone as a free word in a phrase, nor can it undergo reduplication. This can be explained if *yi* ‘clothing’ has inherent word category and cannot merge with the null Nominalizer, as a result, it cannot have its phonological feature copied, thus condition (8) is not met and *yi* ‘clothing’ is a Bound Word. Thus the structure for (156b) is  $W \rightarrow [{}_{n2} [{}_{n1} \text{BW}] [{}_{n2} R \text{ n.RED}]]$ . So (156b) is valid.

## 5.7 Summary

In this chapter, the composition of nouns in OXC is discussed, which is different from that of MSC nouns. I previously proposed that MSC content words including nouns must minimally consist of two items. For OXC nouns, a stricter condition is proposed which is that a free noun must consist of at least two pronounced items. There are Minimal Nouns in both languages and in both cases, the Minimal Noun consists of a Root and a null Nominalizer. The difference is that in MSC the Minimal Noun is a free word, but in OXC, the Minimal Noun is bound. The Reduplicated Nouns are free words in OXC and they are formed of a Root and null Nominalizer with copied phonological content from the single sister Root. A nominal Bound Word cannot merge with the null Nominalizer. As a result, it cannot have its phonological content copied. In other words, it cannot be reduplicated alone. For the adjectives and verbs in both MSC and OXC, the same set of primitives are used, as they are in principle the same. In the affixed nouns in OXC, if the item which the affix is merged with is a head, then it can be reduplicated. If the item which the affix is merged with is a non-head, then it cannot be reduplicated. The reduplication of nouns in attributive nominal compounds in OXC supports the assumption that the Minimal Word cannot be a non-head, it is always the head in terms of category and semantics. The reduplication in OXC affixed words and attributive compound nouns is not obligatory because the two-component condition is met independently. The

reduplicated variant is not ruled out on the ground of the principle of economy used in the thesis, as the economy is defined in terms of the presence or absence of the structure, not of the phonological material. In synonymy-compounds and parallel compounds, neither of the components can be reduplicated, which supports the assumption that components in synonymy-compounds and parallel compounds cannot be Minimal Words. Bound Words do not project their categorial features in those two types of compounds and the null categorizer is assumed to be the categorial head following Zhang (2007). The absence of abstract linking morpheme may only account for the absence of left-branching recursive compound nouns in OXC. For the reduplication in common nouns in OXC, all the structures predicted have turned out to be correct.





## Chapter 6 Conclusions

This thesis has looked into word-formation in both MSC and OXC, in particular, the structure of nouns in both languages. Affixed words, attributive compound nouns, synonymy-compound nouns/adjectives and parallel compound nouns/adjectives in both languages and reduplicated nouns in OXC are described and analysed in this thesis. The research questions raised in chapter 1 are repeated here as follows: (i) What primitives are needed to account for the structure of content words in MSC as well as OXC? (ii) What is the nature of Roots in word formation? (iii) How many components are there minimally in a free content word of MSC and OXC? (iv) What is the nature of reduplication of nouns in OXC? (v) What are the rules needed to derive content words in MSC and OXC? The answers to those questions are summarized below.

The Root, the Bound Word, the null categorizer and the affix are proposed to be primitives in the structure of MSC words and OXC words in the present work, which answers the first research question. It is proposed that the Root is without word category in MSC and OXC, so it cannot be a categorial head in both languages. It follows that a Root alone cannot stand alone as a free word in a phrase in both languages. In other words, it is always bound in both languages. It has to merge with another item which has word category to form a free word in both languages. This answers the second research question.

If a Root has to merge with another item to derive a free word, it means that a free word based on a Root has minimally two components. The Minimal Word will be a Root merged with a categorizer, that is a morpheme whose only feature is a categorial feature. A Minimal Noun, for example, will have the structure [Root, n]. It is proposed that there is a more general condition which holds in MSC and OXC, namely that a content word must have minimally two components. The Minimal Word would be a special case of this. In MSC the Minimal Word can always be spelled out as a form identical with the Root form. But for OXC free, simple nouns, the situation is different: They are reduplicated at all times compulsorily. Therefore, it is proposed that OXC nouns obey a stricter condition than MSC nouns: there are minimally two overt components in a noun in OXC. In short, a free content word, noun, verb, or adjective, contains minimally two items in MSC and

OXC, but for OXC nouns there is an additional condition: a free noun contains minimally two pronounced items. This answers the third research question.

By comparing Minimal Nouns in MSC and Reduplicated Nouns in OXC, it is proposed that a Reduplicated noun in OXC is formed of a Root and a null Nominalizer as well, but the null Nominalizer copies the phonological content of its single sister Root. So, the reduplicant is the Nominalizer and the presence of the null Nominalizer is the key for the reduplication to occur. In other words, the reduplication can happen in nouns formed with a null nominalizer in OXC, otherwise, the reduplication cannot occur. An alternative analysis of reduplication in OXC nouns is that it is motivated by a phonological condition on the minimal size of words (McCarthy and Prince 1990; Hall 1999). This alternative analysis can safely be rejected. For one thing, a phonological condition would not be restricted to nouns. Second, the various exceptions to the reduplication which can be explained under the word syntactic analysis proposed here, do not have any easy explanation in phonological terms. This answers the fourth research question concerning the nature of the Reduplicated Nouns in OXC. The reduplication in OXC nouns clearly demonstrates that there are at least two items, which are a Root and a categorizer, in nouns, which is consistent with the claim that it holds for words more generally. This in turn supports the assumption that the Root is category-less.

This theory casts doubt on the traditional view of the lexical items in MSC called Bound Roots in the literature. Roots are all bound in the present work; instead, the lexical items are Bound Words. Bound Words bear both categorial features and semantic features but do not have internal structure, which makes them different from other content words. Those bound items discussed in the literature are Bound Words because they do not meet the two-component condition set for the smallest word for MSC and OXC. That is to say, a free Bound Word is always ruled out because it violates the two-component condition. At the same time the fact that two Bound Words can merge provides crucial support for the two-components condition. It is a word which does not contain a Root or a categorizer, still it is a licit word, by hypothesis because it consists of two components. Moreover, the fact that two Bound Words can merge and form an attributive compound means that Bound Words can be heads as well as non-heads of compounds. It is therefore postulated that Bound Words project their categorial feature optionally on the dominating

node. More precisely, they project their categorial feature only if they have to. If a Bound Word merges with a Minimal Word, it does not have to, since the Minimal Word can project its categorial feature, so the Minimal Word will always be the head. If the Bound Word merges with a Root, it will have to project its categorial feature, so it does. Further, if Bound Words are items with inherent category, the prediction is that they cannot be reduplicated in OXC. This is indeed the case, as Bound Words are with inherent word category, which means that they cannot merge with the null Nominalizer, so they cannot have their phonological features copied.

Further, the nature of the Reduplicated Nouns in OXC can help to identify the structure of nouns, because it clearly shows the difference between a Root and a noun. To be specific, in a Reduplicated Noun in OXC, the reduplicant is a pronounced null Nominalizer. This means that if an item can be reduplicated in a resulting word in which it is a component, it must consist of a Root and a nominalizer. If an item cannot do so, then it is either a Root (without a nominalizer) or a Bound Word. But if an item cannot be reduplicated alone then it can only be a Bound Word. Apart from the Bound Word, based on the stricter condition set in OXC nouns, it is clear that a Minimal Noun, (that is [R, n]) in OXC also has to be bound, as there is only one overt item in a Minimal Noun, and that is the Root.

Moreover, the reduplication in affixed nouns and attributive compound nouns can also be explained by the nature of reduplication in OXC. It is observed that in affixed nouns, when the affix is the non-head, then the item it is merged with, which is the head, can undergo reduplication; but when the affix is the head, then the item it is merged with, which is the non-head, cannot undergo reduplication. In attributive compound nouns, it is always the case that the head can undergo reduplication, not the modifier. In both cases, only the head item can undergo reduplication, but the non-head item cannot do so. The nature of reduplication in OXC can explain why this is the case. A Minimal Noun is formed of a Root and a null Nominalizer. So for the affixed nouns or attributive compound nouns in which the head is a Minimal Noun, the head can be reduplicated. This is because a Minimal Noun contains a null Nominalizer and the presence of the null Nominalizer means that the reduplication can occur. More specifically, the null Nominalizer in the Minimal Noun, can copy the phonological content of its single sister Root and consequently the head gets reduplicated. On the other hand, the non-heads in the

affixed nouns or attributive compound nouns in OXC, can only be Roots or nominal Bound Words. Neither Roots nor Bound Words contain null Nominalizers. It follows that they cannot be reduplicated, which explains why non-heads cannot be reduplicated.

In short, an attributive compound noun may be a combination of a Root and a Reduplicated Noun or a Bound Word and a Minimal Noun. This challenges the traditional point of view regarding the composition of the attributive noun in the literature: It is not a composition of two free nouns, but a bound item and a free noun or two bound items.

The nature of the reduplicated nouns in OXC also empirically support the assumption that synonymy-compounds/parallel compounds in MSC and OXC cannot contain the Minimal Words. Synonymy-compounds/parallel compounds in MSC and OXC can be any combinations of Root and Bound Word. Synonymy-compounds/parallel compounds in both languages are symmetrical. More specifically, Bound Words do not project in synonymy-compounds/parallel compounds in both languages. This means that both components in synonymy-compounds/parallel compounds cannot be the categorial head. Following Zhang (2007), a third item which is categorizer is the categorial head in synonymy-compounds/parallel compounds in both languages. In terms of semantic head, one of the components can be the semantic head or both components are semantic heads. In addition to synonymy-compounds/parallel compounds, other exocentric compounds including the antonymy-compounds, verb-object compounds and subject-predicate compounds in MSC and OXC are also symmetrical. More specifically, following Zhang (2007), they are a combination of two merged category-less Roots and a third item, which is a categorizer. None of the Roots can be the categorial head and the semantic head in those types of compounds in MSC and OXC. Instead, the categorizer is the categorial head. The meanings of those types of compounds are unobtainable from their components, instead they are derived from the lexicon. Here the word-formation rules of synonymy-compounds/parallel compounds, antonymy-compounds, verb-object compounds and subject-predicate compounds in MSC and OXC are summarized, which partially answers the last research question.

Obligatory noun reduplication only occurs in OXC, not in MSC. We thus do not have the same direct empirical evidence of word structure in MSC as in OXC. However, as the

primitives and the word formation rules apart from reduplication are almost exactly the same in the two varieties of Chinese, it is proposed that the rules are the same. For example, MSC has a class of Bound Words just like OXC, with exactly the same properties. In Chapter 4 arguments for this theory of word structure in MSC were presented, based on language-internal facts and theoretical considerations. This theory has then been further supported by empirical evidence from reduplication, in Chapter 5.

There are differences between MSC and OXC concerning recursive compounds. More specifically, MSC has recursive compounding, but OXC does not have recursive compounding. This has been, at least in part, accounted for by adopting Mukai's (2008) hypothesis that languages with recursive right-branching compounding have a linking morpheme which blocks feature percolation from the modifier of a recursive compound. MSC would have such a linking morpheme, OXC would not. However, this issue needs more research.

The last research question concerns the word-formation rules applied in deriving free content words in MSC and OXC. In general, the rule is the operation Merge as characterized by Chomsky (1995: 243-246) and subsequent work by Chomsky, merging two items *a* and *b*, forming a set {*a*,*b*} with a label which is either *a* or *b*, depending on which one is the head. This requires that the derived word is asymmetric, or else the head cannot be determined. A typical content word consists of a (category-less) Root and an item with category, either a null categorizer or a category-bearing affix or a word, which may be a Minimal Word or a compound. The two items cannot both be null categorizers and affixes; in other words, only one of the items may be a null categorizer or an affix. As for the Minimal Word, based on the economy principle, it cannot be a non-head in terms of category and semantics. And the reduplication facts in OXC have supported the hypothesis that the Minimal Word cannot be a component of synonymy-compounds/parallel compounds in either language. Thus two Minimal Words cannot merge to form a free content word in either language. Instead a Minimal Word can only merge with an item which can be a non-head in terms of category, that is, a Root, a Bound Word or an affix. There is a special condition ruling out right-branching compounds with two adjacent non-head Bound Words.

As for attributive compound nouns in both MSC and OXC, the rule is that a Root and a Bound Word can merge freely with each other to form this type of word. Two Minimal Words cannot merge together to form an attributive compound nouns, given that a Minimal Noun cannot be a non-head in terms of category and semantics. The attributive compound nouns are all right-headed in terms of category and semantics, so they cannot be the result of two Roots merging, as the Root is category-less. But of the theory has to allow for two Bound Words merging, so Bound Words can be heads as well as non-heads in terms of category and semantics.

This thesis has provided a comprehensive description of word-formation in two varieties of Chinese, MSC and OXC, with theoretical consequences for the theory of word structure in general. I also hope that the description of OXC here can make at least a small contribution towards raising awareness of this traditional local language, which is already undergoing changes under the influence of MSC.

## References

- Bauer, L. 1983. *English word-formation*. Cambridge: Cambridge University Press.
- Baker, M.C. 1988. *Incorporation: theory of grammatical function changing*. Chicago: University of Chicago Press.
- Bauke, L.S. 2014. *Symmetry breaking in syntax and the lexicon*. Amsterdam Philadelphia: John Benjamins Publishing Company.
- Bell, D. 2016. 'Syntactic change in Xining Mandarin'. PhD dissertation, Newcastle University.
- Bloomfield, L. 1933. *Language*. London: Allen & Unwin.
- Bybee, J.L., W.Pagliuca and R.D. Perkins. 1990. 'On the asymmetries of affixation of grammatical material'. In W.Croft., K.M. Denning and S.Kemmer (eds) *Studies in Typology and Diachrony: Papers Presented to Joseph H. Greenberg on his 75<sup>th</sup> birthday*. Amsterdam: John Benjamins Publishing Company. 1-42.
- Bisetto, A. 2010. 'Recursiveness and Italian compounds'. *SKASE Journal of Theoretical Linguistics* 7:14-35.
- Booij, Geert. 2011. 'Compounding and construction morphology'. In R. Lieber and P. Štekauer (eds) *The Oxford handbook of compounding*. Oxford: Oxford University Press. 201-216.
- Ceccagno, A. and B. Bascinano. 2007. 'Compound headedness in Chinese: An analysis of neologisms'. *Morphology* 17:217-231.
- Ceccagno, A. and B. Bascinano. 2011. 'Sino-Tibetan: Mandarin Chinese'. In R. Lieber and P. Štekauer (eds) *The Oxford handbook of compounding*. Oxford: Oxford University

Press. 477-490.

Chao, Y. R. 1968. *The grammar of spoken Chinese*. Berkeley: University of California Press.

Chomsky, N. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Berlin: Mouton de Gruyter.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2015. *The Minimalist Program: 20<sup>th</sup> Anniversary Edition*. Cambridge, MA: MIT Press.

Clahsen, H., I. Sonnenstuhl, J. P. Blevins. 2003. 'Derivational morphology in the German mental lexicon: A Dual Mechanism account'. In Baayen, H and R. Schreuder (eds) *Morphological structure in language processing*. Berlin: Mouton de Gruyter. 125–155.

Dai, J.X-L. 1992. 'Chinese morphology and its interface with the syntax'. PhD dissertation, Ohio State University.

Dede, K. 1999. 'Language contact, variation and change: the locative in Xining, Qinghai'. PhD dissertation, University of Washington.

Dede, K. 2003. 'The Chinese Language in Qinghai'. *Studia Orientalia* 95: 321-346

Dede, K. 2006. 'Standard Chinese and the Xining dialect: the rise of an interdialectal standard'. *Journal of Asian Pacific Communication* 16:319-334.

De Belder, M. 2011. 'Roots and affixes: eliminating lexical categories from syntax'. PhD dissertation, Utrecht University.

Di Sciullo, A.M and E. Williams. 1987. *On the definition of word*. Cambridge: MIT Press.



Duanmu, S. 1997. 'Wordhood in Chinese'. In J.L. Packard (ed.) *New approaches to Chinese word formation: Morphology, phonology and the lexicon in modern and ancient Chinese*. New York: Mouton de Gruyter. 135-196.

Duanmu, S. 2002. 'Review of the morphology of Chinese: A linguistic and cognitive approach'. *Diachronica* 19.1: 190-200.

Embick, D and R. Noyer. 2007. 'Distributed morphology and the syntax/morphology interface'. In G. Ramchand and C. Reiss (eds) *The Oxford Handbook of Linguistic Interfaces*. Oxford University Press. 289-324.

Embick, D and R. Noyer. 2008. 'Architecture and blocking'. *Linguistic Inquiry* 39:1-53.

Hall, A. 1999. 'The phonological word: a review'. In T.A. Hall and U. Kleinhenz (eds) *Studies on the phonological word*. Amsterdam: John Benjamins Publishing Company 1-22.

Halle, M. and A. Marantz. 1993. 'Distributed morphology and the pieces of inflection'. In K. Hale and S. Keyser (eds) *The view from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge: MIT Press. 111-176.

Harley, H and Noyer, R. 1999. 'Distributed morphology'. *Glott International* 4:3-9.

Harley, H. 2011. 'Compounding in distributed morphology'. In R. Lieber and P. Štekauer (eds) *The Oxford handbook of compounding*. Oxford: Oxford University Press. 129-144.

Harris, J. 1991. 'The exponence of gender in Spanish'. *Linguistic Inquiry* 22: 27-62.

Hawkins, J.A. 1990. 'A parsing theory of word order universals'. *Linguistic Inquiry* 21:223-261.

Higginbotham 1985. 'On semantics'. *Linguistic Inquiry* 16:547-593.

- Holmberg, A. 1992. 'Properties of non-heads in compounds: a case study'. *Working paper in Scandinavian syntax* 49:27-58.
- Huang, J.C.T. 1984. 'Phrase structure, lexical integrity, and Chinese compounds'. *Journal of the Chinese Language Teachers Association* 19.2: 53-78.
- Huang, C.R and D.X. Shi. 2016. 'Preliminaries'. In C.R. Huang and D.X. Shi (eds) *A Reference Grammar of Chinese*. Cambridge: Cambridge University Press. 1-14.
- Josefsson, G. 1997. *On the principles of word formation in Swedish*. Kent: Chartwel-Bratt Ltd.
- Li, C.N. and S.A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. London: University of California Press.
- Liao, W.W.R. 2014. 'Morphology'. In C.T. James, Y. A. Li, A. Simpson (eds) *The handbook of Chinese linguistics*. Malden, Massachusetts: Wiley-Blackwell. 3-25.
- Lieber, R. 1980. 'On the organization of the lexicon'. PhD dissertation, MIT.
- Lieber, R. 1992. *Deconstructing morphology: word formation in syntactic theory*. Chicago: The University of Chicago Press.
- Lieber, R. 2016. *Introducing morphology*. New York: Cambridge University Press.
- Lin, H. 1990. 'Towards an analysis of Mandarin reduplication'. *Working papers of the linguistics circle* 9:129-141.
- Lin, H. 2001. *A grammar of Mandarin Chinese*. Muenchen: Lincom Eupora.
- Lin, T-H.J. 2010. 'Structures and functional categories in Mandarin sentences'. *UST Working Papers in Linguistics* 6:41-79.

- Lin, T-H.J. 2015. 'Tenses in Mandarin Chinese'. *Syntax* 18:320-342.
- McCarthy, J. and A. Prince. 1990. 'Foot and Word in Prosodic Morphology: The Arabic Broken Plurals'. *Natural Language and Linguistic Theory* 8:209-282.
- Marantz, A. 1997. 'No escape from syntax: Do not try morphological analysis in the privacy of your own lexicon.' *University of Pennsylvania Working Papers in Linguistics* 4.2: 201-225.
- Marantz, A. 2001. 'Words'. Paper presented at the West Coast Conference on Formal Linguistics, University of Southern California Los Angeles.
- Marantz, A. 2007. 'Phrases and words'. In S.H. Choe (ed) *Phrases in the theory of grammar*. Seoul: Doing-In Publishing Co. 191-222.
- Miguel, M.M. 2017. 'Instruction in derivational morphology in the Spanish L2 Classroom: What do teachers believe and do' *Konin Language Studies* 1:37-60.
- Mukai, M. 2008. 'Recursive Compounds'. *Word Structure* 1: 178-198.
- Mukai, M. 2015. 'Word formation in phrase theory'. *Newcastle and Northumbria Working Papers in Linguistics* 21.1:84-95.
- Nemat, K. 2016. 'Derivational Morphology in French'. *Journal of Language Sciences and Linguistics* 4:229-234.
- Packard, J.L. 2000. *The morphology of Chinese: a linguistic and cognitive approach*. Cambridge: Cambridge University Press.
- Packard, J.L. 2015. 'Morphology: morphemes in Chinese'. In W. S-Y. William and C.F. Sun (eds) *The Oxford handbook of Chinese linguistics*. New York: Oxford University Press. 263-273.

- Packard, J.L. 2016. 'Lexical word formation'. In C.R. Huang and D.X. Shi (eds) *A Reference Grammar of Chinese*. Cambridge: Cambridge University Press. 67-80.
- Pirani, L. 2008. 'Bound roots in Mandarin Chinese and comparison with European semi-words'. *Proceedings of the 20<sup>th</sup> North American Conference on Chinese Linguistics*. 1:261-277.
- Plag, I. 2003. *Word formation in English*. Cambridge: Cambridge University Press.
- Po-Ching, Y. and D. Rimmington. 2016. *Chinese: a comprehensive grammar*. New York: Routledge.
- Ralli, A. 2013. *Compounding in modern Greek*. New York: Springer.
- Ren, B.S. 2006. *Qinghai fangyan yufa zhuanji yanjiu [The grammar of the Qinghai dialect]*. Xining: Qinghai renmin chubanshe
- Roeper, T., W. Snyder and K. Hiramatsu. 2002. 'Learnability in a Minimalist Framework: Root Compounds, Merger, and the Syntax-Morphology Interface'. In I. Lasser (ed.) *The Process of Language Acquisition*. Frankfurt: Peter Lang Verlag. 25-37.
- Scalise, S. and A. Fábregas. 2010. 'The head in compounding'. In S. Scalise and I. Vogel (eds) *Cross-disciplinary issues in compounding* Philadelphia: John Benjamins Publishing Company. 109-127.
- Selkirk, E. 1982. *The syntax of words*. Cambridge: MIT Press.
- Shen, L. 2004. 'Aspect agreement and light verbs in Chinese: a comparison with Japanese'. *Journal of Eastern Asian Linguistics* 13:141-179.
- Sybesma, R. 2007. 'Whether we Tense-agree overtly or not'. *Linguistic Inquiry* 38:580-587.

- Spencer, A. 1991. *Morphological theory*. Oxford: Basil Blackwell.
- Sproat, R. and C. Shih. 1996. 'A corpus-based analysis of Mandarin nominal root compound'. *Journal of East Asian Linguistics* 5: 49-71.
- Sun, C.F. 2006. *Chinese: a linguistic introduction*. Cambridge: Cambridge University Press.
- Sun, C.F. 2015. 'The uses of *de* 的 as a noun phrase marker'. In S. Y. Wang and C.F. Sun (eds) *The Oxford handbook of Chinese linguistics*. New York: Oxford University Press. 362-379.
- Tang, T.-C. 1993. 'More on the relation between word-syntax and sentence-syntax in Chinese: case study in compound nouns'. In J. Camacho and L. Choueiri (eds) *Proceedings of the sixth North American Conference on Chinese Linguistics*. Los Angeles: GSIL, University of Southern California, 195-248.
- Talamo, L. and Celata, C. 2011. Toward a morphological analysis of the Italian lexicon: developing tools for a corpus-based approach. *Quaderni del Laboratorio di Linguistica*, 1:1-16.
- Wang, S.C. 2009. 'Xining fangyan chongdieshi' [Reduplication in Xining Chinese]. *Qinghai shifan daxue minzu shifan xueyuan xuebao* 5: 1-4.
- Wang, S.C. and K. Dede. 2016. 'Negation in the Xining dialect'. *Language and Linguistics* 17:403-429.
- Williams, E, 1981. 'On the notions lexically related and head of a word'. *Linguistic Inquiry* 12:245-274.
- Zhang, C.C. and S.K. Zhu. 1987. *Xining fangyanzhi [Xining dialect annal]*. Xining: Qinghai renmin chubanshe.

Zhang, C.C. 2001. *Qinghai shengzhi (fangyan zhi) [Qinghai annals (dialect annals)]*.  
Qinghai:Huangshan.

Zhang, N.N. 2007. 'Root merger in Chinese compounds'. *Studia Linguistica* 2: 170-184.