ORIGINAL PAPER

# Compound headedness in Chinese: an analysis of neologisms

**Antonella Ceccagno · Bianca Basciano**

**Abstract**  Compounding is the most significant morphological phenomenon of Chinese. In fact, it has been shown that up to 80% of Chinese words are compounds (Zhou et al. Language and Cognitive Processes 14(5/6):525–565, 1999; Xing Teaching and learning Chinese as a foreign language: A pedagogical grammar, 2006). This paper is based on a corpus of Chinese neologisms. It aims at discussing current theoretical approaches to compounding and presents a novel analysis of Chinese compounds. We first discuss the issue of the classification of Chinese compounds, modifying the proposal put forth by Bisetto and Scalise (Lingue e Linguaggio 2:319–332, 2005), which has first been adopted for Chinese compounding by Ceccagno and Scalise (Lingue e Linguaggio 2:233–260, 2006). After presenting a new classification for Chinese compounding, we show the variety of structures existing in Chinese compounds. We then discuss the much debated notion of headedness in compounding, challenging the widespread idea that any language must be either right-headed or left-headed (as partially highlighted by Packard The morphology of Chinese, 2000). Crucially, Chinese seems to display three different productive patterns: right-headedness, left-headedness and two-headedness. These conclusions represent a challenge for the cross-linguistic study of compounding.

**Keywords**  Morphology · Chinese morphology · Compounding · Chinese compounding · Headedness · Headedness in Chinese compounding · Compound classification

A. Ceccagno
Dipartimento di Lingue e Letterature Straniere, University of Bologna, Bologna 40124, Italy
e-mail: aceccagno@lingue.unibo.it

B. Basciano (✉)
Dipartimento di Germanistica e Slavistica, University of Verona, Verona 37129, Italy
e-mail: bianca.basciano@gmail.com

## 1 Introduction

Compounding is the most productive means of word formation in Chinese. In fact, it has been shown that compound words are over 70% of all words used in Chinese (Institute of Language Teaching and Research, cited in Zhou et al. 1999), or even 80% (Xing 2006).

In Chinese compounding seems to be the rule in the formation of new words. This emerges clearly from the corpus of neologisms developed over the last thirty years analysed in this article. In the corpus, proposed by *The Contemporary Chinese Dictionary* (2002), out of 709 new words (those with no more than two syllables) almost 95% (672) are compounds while only a little more than 2% (16) are derived words. This peculiarity seems to be ascribable to the basically isolating characteristics of Chinese, which set Chinese apart from the majority of Indo-European languages—in which compounding is not as pervasive (Ceccagno 2008, in press). Therefore studying Chinese compounds means studying the most significant morphological phenomenon of that language.

This article, developed within the Morbo/comp project on compounds,[1] presents a new classification of Chinese compounds and illustrates the variety of structures existing in Chinese compounds. The data will also form the empirical basis for discussing the existing theories on headedness in Chinese and formulating a new one. In fact, this article intends to challenge the question of the head in Chinese compounding in relation to the widespread notion that in its core compounding a language must be either right headed or left headed. With regard to headedness in Chinese compounding, we suggest a new approach based on evidence from recently formed new words. The identification of a phenomenon never analysed before, i.e. the formation of new compounds from underlying compounds or 'metacompounding', helps to tackle some problems that could arise in the determination of the position of the head.

## 2 Methodology

In word formation studies, traditional linguistic analysis used to look at established and lexicalized words, which however often showed all kinds of idiosyncrasies that are not indicative of the speakers' knowledge of the structure of possible words. This happened also for Chinese compounding, where traditional forms together with new ones have been analysed (cf. Huang 1998; Packard 2000 among others). In the last decade, however, there has been a trend in word-formation studies to use neologisms as useful data (Plag 1998, 2003, among others) because of the importance of these forms in the description

---

[1] The Morbo/Comp project based in Bologna is led by Sergio Scalise. It concerns the analysis of compounds in some thirty languages, distributed in such a way as to cover the six macro-areas that have been identified by recent typology (cf. Dryer 1992). To present, fifteen languages have been studied, for a total in excess of 70.000 compounds. Cf. http://morbocomp.sslmit.unibo.it/index.php?section = home.

of the morphological competence of today's speakers. Moreover, neologisms help to show what are the patterns in use in a given period.

In this article we present the results of a research on Chinese compounds, which is only based on neologisms. In fact, the data we make use of consist of 1077 neologisms, presented as a separate section in *The Contemporary Chinese Dictionary* (2002). The dictionary does not state the year in which the neologisms were formed. However, we believe that some of the neologisms were already formed in the early 1980s, while the largest part seems to have been formed in the most recent years.

For reasons of simplicity we have chosen to focus on dysillabic words in our analysis, excluding words with more than two syllables (N = 368). We have also excluded from the corpus all non-compound words (N = 37).

## 3 Headedness in Chinese compounds: the existing generalizations

The issue of the identification of the head in compounding is of particular interest. Since the 1980s there has been a debate on this issue, with Williams (1981), Selkirk (1982), Trommelen and Zonneveld (1986), Di Sciullo and Williams (1987) assuming that the position of the head was invariable on the right, and Scalise (1983), Corbin (1987) and Varela (1990) stating that in compounds in Romance languages heads are consistently on the left.[2] The different positions of morphological heads in the different languages studied induced Bisetto and Scalise (2002) to conclude that the position of the head in all the languages of the world is a parameter that needs to be set.

As far as Chinese compounding is concerned, back in 1968, Chao distinguished between subordinate and coordinate compounds in Chinese, saying that "coordinate compounds differ from subordinate compounds in that in coordinate compounds each constituent is a centre while in a subordinate compound only the second constituent is the centre" (we assume 'centre' as the equivalent of head).

In his analysis of Chinese compounding Huang (1998) concluded that: "[…] Chinese is a headless language in its morphology since neither the rightmost nor the leftmost member of a compound uniquely determines the category type of a compound".[3] Huang reached this conclusion after analysing the output categories of a substantial corpus of Chinese disyllabic compounds. He observed that compounds with the same structures can have different output categories. For example, [V + N] compounds can be nouns as well as verbs, and sometimes adjectives. What Huang failed to consider is the relation between the constituents. In fact, the position of the head also depends on the relation running between the two constituents. For example in a compound such as 卖场 màichǎng 'to sell + place = big marketplace for selling commodities', with a

---

[2] For a discussion on the development of the theory of headedness in compounding see Scalise et al. (2005).

[3] For a critical analysis of Huang's positions see Ceccagno and Scalise (2007).

$[V + N]_N$ structure, the head is the noun constituent and the verb acts as a modifier, i.e. 'the place in which one sells'. On the other hand, in a $[V + N]_V$ compound, such as 投资 tóuzī 'to put + money = to invest', the head is the left constituent, the verb, while the noun acts as its internal argument. Finally, a $[V + N]_N$ compound such as 监事 jiānshì 'supervise + matter/responsibility = supervisor' has the same structure and the same output category as the compound in the first example, but the relation between the constituents is completely different: it is an exocentric compound where the constituents have a verb-argument relation. The examples highlight that not only input and output categories are important but also the relation between the constituents contributes to determine the position of the head in a compound.

Starosta et al. (1998) claimed that Chinese compounds are right-headed and "[…] it may be that many of the words which are not right-headed are also not compound words". Packard (2000, p. 42) proposed a generalization (Headedness Principle) according to which "All verbs have a verb on the left; and all nouns have a noun on the right". As we will see, this generalization only applies to subordinate compounds, but it does not take into account the fact that verbs can be right-headed in attributive compounds. Further, Packard fails to acknowledge the existence of coordinate two-headed compounds and exocentric compounds, which are quite widespread in Chinese compounding.

These limitations in Packard's generalization have been highlighted by Ceccagno and Scalise (2006). Their data showed the prevalence of right-headed compounds. Left-headed compounds did exist in their corpus (of the verb-object and resultative types), however the fact that these left-headed compounds fall in between morphology and syntax—in that it is not always clear if they are compounds or phrases—led them to assume that "the canonical position of the head in Chinese compounds is on the right and that the so-called verb-object compounds and resultative compounds are an exception to this principle".

While following in the footsteps of the theoretical and analytical framework of the above mentioned work, our analysis of a larger database of Chinese neologisms enables us to reach different conclusions. After presenting our analysis of the Chinese neologisms we will discuss our generalization on the position of the head in Chinese compounds.

## 4 Classifying compounds

In the analysis of Chinese compounds we adopt the methodology put forth by Ceccagno and Scalise (2007) who propose that the whole set of category, functional and semantic levels should be taken into account for an exhaustive analysis of compounds, i.e. "an analysis capable of identifying: the lexical category of the constituents, their grammatical relationship and therefore the classification of the whole compound, the semantics of the constituents, the semantics of the compound and the position of the head (if any)". They posit that in the absence of one of these aspects the analysis is incomplete, if not misleading.

With regard to classification, we put forth a new classification scheme that is based on the one argued for by Bisetto and Scalise (2005), who identify three macro-types in compounding (subordinate, attributive and coordinate). Each type may be endocentric (with a lexical head) or exocentric (without a lexical head).[4] For a meaningful description of Chinese data, we propose a refinement of Bisetto and Scalise (2005) classification through a new definition of the macro-types.

Subordinate compounds (SUB) are those in which constituents have an argument-head (or head-argument) relation. A first type shows either a verbal or deverbal head which projects an argument satisfied by the non-head constituent. Moreover, we consider subordinate compounds also those compounds that show a verbal head that takes a verb as its complement. Examples are presented below:

(1) 毒贩 *dúfàn* [N + N]ₙ 'drug+vendor = drug trafficker'

(2) 禁毒 *jìndú* [V + N]ᵥ 'prohibit+poison = ban the production, sale and abuse of drugs'

(3) 攀高 *pāngā*o [V + A]ᵥ 'climb+high = climb up; rise'
入住 *rùzhù* [V + V]ᵥ 'come into/enter+live/stop = move into'

(4) 拒载 *jùzài* [V + V]ᵥ 'refuse+carry = (of a tax driver) refuse to take a passenger'

毒贩 *dúfàn* (1) is a compound with a deverbal head, where the noun constituent acts as its argument. 禁毒 *jìndú* (2) is a verbal compound of the verb-object type, where the leftmost constituent is the head of the compound and the noun acts as the internal argument of the verb. 攀高 *pāngāo* and 入住 *rùzhù* (3) are verbal compounds of the resultative type, in which the non-head constituent is in a complement relation with the head constituent, specifying the result produced by the event of the head (the left constituent). 拒载 *jùzài* (4) is a compound of the serial verbs type, in which the event expressed by the verb on the right depends on that expressed by the verb on the left. It can be anticipated here that subordinate nominal compounds are right-headed while subordinate verbal compounds are left-headed.

A second type of subordinate compound shows a relational noun as head, where the non-head acts as a semantic argument saturating the noun head, as in the following examples:

---

[4] Besides, Scalise and Guevara (2006) distinguish between categorial and semantic head. The categorial head is "the constituent which shares with—and percolates to—the whole compound all of its formal features: lexical category and subcategorization frame". The whole compound, thus, is expected to have the same distributional properties of its categorial head. The semantic head "is the constituent which shares with—and percolates to—the whole compound all of its lexical/conceptual information. The whole compound, thus, is expected to be a hyponym of its semantic head". According to the authors, an endocentric compound has at least one categorial head and at least one semantic head; the two must coincide.

(5) 警嫂 *jǐngsǎo* [N + N]ₙ 'police+sister = respectful term for a policeman's wife'

(6) 价差 *jiàchā* [N + N]ₙ 'price+difference = price difference'.

Attributive compounds (ATT) are those in which the constituents have a modifier-head[5] relation. These can be compounds where:

(1) the non-head is an adjective or a noun which expresses a property of the head, as in 黑金 *hēijīn* [A + N]ₙ 'black/illegal + money = money used for bribery and other illegal activities within official circles' or 天价 *tiānjià* [N + N]ₙ 'sky + price = prohibitive price';
(2) the non-head constituent acts as an adjunct modifying the head, as in 口算 *kǒusuàn* [N + V]ᵥ 'mouth + (to) do a sum = (to) do a sum orally';
(3) a verbal non-head acts as a modifier of the head, as in 卖场 *màichǎng* [V + N]ₙ 'sell + large place where people gather for a specific purpose = big marketplace for selling commodities'.

Coordinate compounds (CRD) are those which show a logical coordination between the constituents ('and'), as in 新锐 *xīnruì* [A + A]ₐ 'new + sharp = new and sharp', or a semantic relation of sinonymy, as in 胜绩 *shèngjì* [N + N]ₙ 'victory + achievement = win/victory'; antonymy, as in 呼吸 *hūxī* [V + V]ᵥ 'exhale + inhale = breath'; redundancy, as in 松树 *sōngshù* [N + N]ₙ 'pine + tree = pine tree'; reduplication, as in 天天 *tiāntiān* [N + N]ₐdᵥ 'day + day = every day'.

Figure 1 shows the proposed classification scheme.


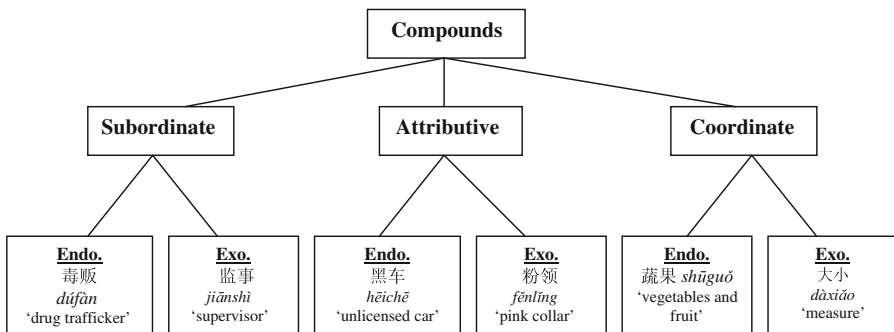
**Fig. 1** Classification of Chinese Compounds

---

[5] A similar distinction between 'argument-head' (or 'argument-predicate') and modifier-head has been proposed for English by Giegerich (2004) and Plag et al. (2007) among others.

## 5 Types of compound and productivity[6]

### 5.1 A discussion of the data

Our analysis of the neologism corpus shows that all three types of compound are productive (Table 1): attributive with 294 instances (43.9%) and subordinate with 241 instances (35.7%) are the most common macro-types. However, coordinates are also productive (137 instances, equal to 20.4%).

#### 5.1.1 Subordinate compounds

Subordinate compounds show an almost equal noun vs. verb output, with the former being right-headed or exocentric and the latter left-headed or exocentric. Table 2 presents endo- and exocentric subordinate compounds with $[N + N]_N$ structure.

There also exist exocentric $[V + N]_N$ subordinate compounds in which the semantic relation between the constituents is of the verb-object type, where the noun constituent acts as the internal argument of the verb.

These compounds behave in the same manner as $[V + N]_N$ subordinate compounds in other languages, i.e. *pickpocket* in English or *tagliacarte* 'paper knife' in Italian, and can be interpreted as agentive, instrumental, locative or eventive (Table 3).

**Table 1** Productivity of Chinese compounds

|  | SUB | ATT | CRD |
|---|---|---|---|
| TOT | 240 | 295 | 137 |
| % | 35.7 | 43.9 | 20.4 |

**Table 2** Subordinate $[N + N]_N$ endocentric and exocentric[a] compounds A = adjective, V = verb, N = noun, Adv = adverb

| Compound | Pinyin | Class. | Structure | CAT. | Head | Gloss |
|---|---|---|---|---|---|---|
| 毒贩 | dúfàn | SUB | [N + N] | N | Right | Drug + vendor = drug trafficker |
| 警嫂 | jǐngsǎo | SUB | [N + N] | N | Right | Police + sister (term of address for a married woman about one's own age) = respectful term for a policeman's wife |
| 市道 | shìdào | SUB | [N + N] | N | Exocentric | Market + way = market prices |
| 网虫 | wǎngchóng | SUB | [N + N] | N | Exocentric | Net + insect = web enthusiast |

[a] On the basis of Scalise and Guevara (2006) statement that in endocentric compounds categorical head and semantic head must coincide, compounds such as 网虫 *wǎngchóng* 'net+insect=web enthusiast' are considered exocentric. Since 虫 *chóng* is used in a metaphorical sense, it thus does not share with the whole compound all of its lexical-conceptual information, i.e. it cannot be consider the semantic head. The compound is right-headed from the categorical point of view, but semantically it is exocentric; therefore, since the semantic and the categorical head do not coincide, we consider this compound as exocentric

[6] In this article we do not intend productivity as a value determined by means of statistical measurements, but rather in the informal sense of wider or lesser presence in our corpus.

**Table 3** Subordinate [V + N]$_N$ exocentric compounds

| 监事 | jiānshì | SUB | [V + N] | N | Exocentric | Supervise + matter = supervisor | Agentive interpretation |
|---|---|---|---|---|---|---|---|
| 文胸 | wénxiōng | SUB | [V + N] | N | Exocentric | Conceal/hide/cover up + breast = bra | Instrumental interpretation |
| 扶手 | fúshǒu | SUB | [V + N] | N | Exocentric | 'Support oneself + hand = handrail' | Locative interpretation |
| 攀岩 | pānyán | SUB | [V + N] | N | Exocentric | 'Climb + rock = = rock climbing' | Eventive interpretation |

**Table 4** Subordinate [V + N]$_V$ compounds

| 待岗 | dàigǎng | SUB | [V + N] | V | Left | Wait for + post = wait for a job |
|---|---|---|---|---|---|---|
| 禁毒 | jìndú | SUB | [V + N] | V | Left | Prohibit + poison = ban the production, sale and abuse of drugs |
| 流标 | liúbiāo | SUB | [V + N] | V | Exocentric | Flow + label/standard/prize = fail to sell at auction because no bids |
| 割肉 | gēròu | SUB | [V + N] | V | Exocentric | Cut with a knife + meat = sell sth. at a price lower than its original price |

Most of the subordinate compounds with verb output in our corpus presents the [V + N]$_V$ structure. Table 4 illustrates examples of endocentric and exocentric subordinate compounds having [V + N]$_V$ structure of the verb-object type.

Chinese verbal compounds of the verb-object type represent a much discussed issue since they may fall in between morphology and syntax. Chao (1968), Li and Thompson (1981), Huang (1984), Chi (1985) and Packard (2000) proposed different criteria to distinguish verb-object and phrases: lexicality or specialization of meaning; inability of the verb and object to be moved or separated; one of the constituents is a bound root, i.e. it cannot occupy a syntactic slot; exocentricity of the construction; the construction allowing an additional object (for a discussion of the criteria, see Packard 2000).[7] Basically, we use three criteria to determine a verb-object construction as word: non-compositionality of meaning; the construction allowing an additional object; the status of one of the constituent as a bound root. Some constructions, however, even when following these criteria, can act both as words and as phrases. In these instances, the degree of lexicalization helps distinguishing the verb-object compound from the corresponding phrase. An in-depth investigation of verb-objects constructions is beyond the scope of this article.

Another type of subordinate compound is the one in which the relation between constituents is of the 'serial verbs' type, modelled on the syntactic structure of Chinese, where the event expressed by the verb on the right depends on that

---

[7] A special case is that of those verbs, such as 看 *kàn* 'read' or 吃 *chī* 'eat', which requires an overt object when used intransitively. In these cases the meaning is not compositional, because 看书 *kànshū* 'read + books', for example, does not mean 'to read books', but simply 'to read'. Ross (1998) calls these obligatory, semantically weak objects, as 'cognate objects'. As highlighted by Packard (2000), two syntactic words that occur in a syntactic verb-object phrase paradoxically are categorizable as the single verb 'read' on semantic ground, since the meaning of 书 *shū* 'book' is not realized.

**Table 5** Subordinate [V + V]$_V$ serial verb type compounds

| 示爱 | shì'ài | SUB | [V + V] | V | Left | Show + love = show one's love |
|---|---|---|---|---|---|---|
| 告破 | gàopò | SUB | [V + V] | V | Left | Inform + crack into pieces = make known that a mystery has been cracked |

**Table 6** Subordinate resultative compounds

| 攀高 | pāngāo | SUB | [V + A] | V | Left | Climb + high = climb up; rise |
|---|---|---|---|---|---|---|
| 入住 | rùzhù | SUB | [V + V] | V | Left | Come into/enter + live/ stop = move into |

expressed by the verb on the left, so that they are considered left-headed compounds (Table 5). As far as we know, this kind of compound has never been discussed in literature, which has focused instead on resultative compounds (often referred to as resultative serial verbs constructions). What we call serial verb compounds show the same relation between the constituents as that shown by syntactic serial verb purpose clause and complement clause constructions (Li and Thompson 1981).[8]

The remaining cases of verbal subordinate compounds, in which the second constituent comes as a verb or an adjective, are resultative verbs that show a verb-result complement relation between constituents (Table 6).

Resultative constructions have been widely studied in literature. Lin (1990), Li (1990, 2005), Cheng (1997), Packard (2000) and Sun (2006) all seem to assume that these constructions exhibit a verb-complement relation, where the resultative verb (on the right) acts as a complement of the verb on the left. As highlighted by Li (2005), Levin and Rappaport Hovan (1995) provide evidence that English resultatives consist of a verb taking a complement XP: in a resultative construction the result phrase is in the complement position. This supports the analysis of resultative compounds as left-headed.[9] A different analysis of Chinese resultative constructions has been put forth by Starosta et al. (1998), who consider resultative verbs as derived words headed by the resultative suffix on the right.[10]

---

[8] In Chinese, serial verb constructions do not indicate a single construction with its own properties but a number of different constructions with different properties. The only common feature is that they are a sequence of two verbs without any overt coordinating or subordinating markers between them. Serial verb purpose clause constructions are those constructions where the first event is done for the purpose of achieving the second. In a complement clause construction the second verb is subcategorized for by the first verb.

[9] The issue of resultative as lexical compounds (e.g. Li 1990) vs. syntactic compounds (e.g. Huang 1992) has been widely discussed. For a detailed view on the topic, see Cheng (1997).

[10] For some Chinese resultative compounds, such as 攀高 pāngāo 'climb + high = climb up/rise' it would seem possible an analysis similar to that proposed for resultative compounds in English or for prefixed verbs in Russian (see, for instance, Spencer and Zaretskaya 1998), where the element correspondent to 高 gāo in these constructions, although is syntactically a secondary predicator, expresses the core semantic predicator, so that semantically the verb stem is subordinated. However, there are a number of differences between these constructions and Chinese resultatives, for example the fact that, unlike languages such as English, in Chinese secondary predication is possible over external argument. Moreover, even though semantically secondary, the verb stem does play an important role in the change of state. Finally, syntactically it is still the primary predicator. A detailed analysis of resultative compounds is, however, beyond the scope of this article.

**Table 7** Subordinate compounds

|  | V | N |  |
|---|---|---|---|
| a. Output categories |  |  |  |
| TOT | 131 | 108 |  |
| % | 53.9 | 44.8 |  |

|  | Exocentric | Left-headed | Right-headed |
|---|---|---|---|
| b. Headedness |  |  |  |
| TOT | 38 | 120 | 83 |
| % | 15.8 | 49.8 | 34.4 |

To sum up, the analysis of Chinese subordinate compounds shows that endocentric compounds exhibit a polarization of verb output—all with head on the left (128 in our corpus) and noun outputs—all with head on the right (108 in our corpus). On the other hand, exocentricity is more limited (38) and mainly associated with compounds having noun output (27)[11] (Table 7).

Anticipating our discussion on the head, to be presented in 5, the above data illustrate that, if we are to exclude exocentric compounds, the behaviour of Chinese subordinate compounds appears to be in line with the generalization in Packard (2000, p. 39), according to which nominal compounds have to be right-headed and verbal compounds left-headed. This generalization however does not apply to attributive and coordinate compounds (and to exocentric compounds as well), as we will see below.

### 5.1.2 Attributive compounds

As stated by Ceccagno and Scalise (2006, p. 255), Chinese endocentric attributive compounds behave in a straightforward manner: the head is always to be found on the right and the non-head acts as a modifier. Furthermore, in attributive compounds all lexical categories—nouns, verbs and adjectives—can act either as the head or the non-head of the compound.

Our analysis of neologisms in Chinese shows that the lion's share of attributives are endocentric (77.9%) and that the most productive ones are those with noun output (233 out of 294, i.e. 79.2%), followed by those with verbal output (44 instances, amounting to 15%), whereas those with adjectival output are only modestly productive (14 instances, amounting to 4.8%)[12] (Table 8).

In turn, attributive compounds with noun output present four structures with varying productivity levels: $[A + N]_N$ is the most common structure (87 cases in our database), followed by $[N + N]_N$ (74) and $[V + N]_N$ (51). Table 9 shows examples of attributive compounds with noun output.

Attributives with verb output have the verb as the head and can have, as the non-head, nouns, adjectives and verbs acting as modifiers of the head ($[N + V]_V$,

---

[11] The corpus also presents 2 adjectives.

[12] The corpus also presents 2 adverbs.

**Table 8** Attributive compound

| | N | A | V |
|---|---|---|---|
| a. Output categories | | | |
| TOT | 233 | 14 | 45 |
| % | 79.2 | 4.8 | 15 |

| | Exocentric | Right-headed |
|---|---|---|
| b. Headedness | | |
| TOT | 61 | 233 |
| % | 20.7 | 79.2 |

**Table 9** Attributive compounds (nouns)

| 猫步 | māobù | ATT | [N + N] | N | Right | Cat + walk = cat's walk |
|---|---|---|---|---|---|---|
| 联网 | liánwǎng | ATT | [V + N] | N | Right | Unite + net = connected network |
| 主页 | zhǔyè | ATT | [A + N] | N | Right | Main + page = homepage |

**Table 10** Attributive compounds (verbs)

| 公示 | gōngshì | ATT | [A + V] | V | Right | Public + show = make known to the public and seek opinions |
|---|---|---|---|---|---|---|
| 飙升 | biāoshēng | ATT | [N + V] | V | Right | Whirlwind + rise = (of price, quantities) soar |
| 品读 | pǐndú | ATT | [V + V] | V | Right | Decide with discrimination + read = read carefully/ponder on |

**Table 11** Attributive compounds (adjectives)

| 统合 | tǒnghé | ATT | [Adv + A] | A | Right | All/totally + whole = uniform/overall |
|---|---|---|---|---|---|---|
| 常销 | chángxiāo | ATT | [A + V] | A | Exocentric | Constant + sell/market = (of merchandise) in constant demand |
| 花心 | huāxīn | ATT | [A + N] | A | Exocentric | Attractive but unreal or insincere + heart = unfaithful |

$[A + V]_V$ and $[V + V]_V$). Among these, the most productive are those having an $[A + V]_V$ structure (Table 10).[13]

Attributive compounds with adjective output are less productive (Table 11).

Finally, analytical issues are raised by $[A + N]_A$ compounds, such as 空置 *kōngzhì* 'empty + place = empty/unoccupied' (previously discussed in Ceccagno and Scalise 2006, cf. 狠心 *hěnxīn* 'ruthless/harden + heart = cruel'), in which the left constituent actually appears as the categorial and semantic head of the compound. Pan et al. (2004), among others, consider these words as derived, where the right constituent acts as an affix/affixoid.

---

[13] Some compounds as 突审 *tūshěn*, 突显 *tūxiǎn*, and 突现 *tūxiàn* seem to have an $[Adv + V]_V$ structure. It should be borne in mind that 突 *tū* can be a verb ('dash forward'), an adjective ('prominent'), and an adverb ('suddenly').

**Table 12** Chinese coordinate compounds

| 节庆 | jiéqìng | CRD | [N + N] | N | Two-headed | Festival + celebration = festival/ holiday |
|------|---------|-----|---------|---|------------|---------------------------------------------|
| 联动 | liándòng | CRD | [V + V] | N | Exocentric | Unite + move = chain effect/reaction |
| 打压 | dǎyā | CRD | [V + V] | V | Two-headed | Hit + suppress = combat and suppress |
| 亮丽 | liànglì | CRD | [A + A] | A | Two-headed | Bright + beautiful = brilliant/bright and beautiful |
| 峰位 | fēngwèi | CRD | [N + N] | N | Right | Peak + place = peak |

### 5.1.3 Coordinate compounds

Our corpus includes coordinate compounds that manifest a logical conjunction between constituents ('and') and those that, from a semantic point of view, are synonymous or redundant coordinate compounds; on the other hand, no instances were found of antonymous or reduplicate coordinate compounds.

Furthermore, our corpus shows that Chinese is productive with regards to the formation of endocentric coordinate compounds with $[V + V]_V$ (75), $[N + N]_N$ (21) and, albeit to a lesser extent, $[A + A]_A$ (8) structures (Table 12). In endocentric coordinate compounds both constituents act as head. As stated by Sun (2006) the constituents of coordinate compounds are coordinate in nature or are parallel with each other within a semantic domain. According to Chao (1968), in Chinese coordinate compounds each constituent is a 'center'. Anderson (1985) argues that neither constituent can exclusively be considered as the 'centre'.

Different criteria contribute to show that Chinese coordinate compounds are two-headed. First, in coordinate compounds both constituents semantically contribute to the interpretation of the whole compound. Second, criteria such as inflection of the head constituent, gender, etc.—which in other languages help to establish which constituent formally acts as the head[14]—do not exist in Chinese. Therefore, it cannot be established which constituent provides the morpho-syntactic properties to the whole compound. Third, the fact that—as will be formally stated later—Chinese compounds seem not to have a canonical position for the head, in that Chinese forms productively both right-headed and left-headed compounds, also strengthens the conclusion that coordinate endocentric compounds are two-headed.

A somehow different group of coordinate compounds is represented by redundant compounds, where the first constituent represents a subclass of the second constituent. On the model of compounds of this type in other languages, such as *palm tree* in English, we can assume, on the basis of semantic criteria, that this kind of coordinate compound is right-headed. Examples are presented below (Table 13).

---

[14] According to different criteria, coordinate compounds in Germanic languages are considered right-headed, taking on the morpho-syntactic properties of the rightmost constituent. For a discussion of copulative compounds in German and English, see Olsen (2001).

**Table 13** Redundant compounds

| 峰位 | fēngwèi | CRD | [N + N] | N | Right | Peak + place = peak |
| 饮品 | yǐnpǐn | CRD | [N + N] | N | Right | Drinks + article = drink |

**Table 14** Coordinate compounds

|  | V | N | A |
|---|---|---|---|
| **a. Output categories** | | | |
| TOT | 75 | 54 | 8 |
| % | 54.7 | 39.4 | 5.8 |

|  | Right-headed | Two-headed | Exocentric |
|---|---|---|---|
| **b. Headedness** | | | |
| TOT | 9 | 102 | 26 |
| % | 6.6 | 74.4 | 19 |

Exocentricity does exist in coordinate compounds; however in our corpus it can be found almost exclusively in coordinate compounds with $[V + V]_N$ structure (23).

Our database shows that Chinese coordinate compounds mainly form verbs (75 out of 137, equal to 54.7%), followed by nouns (54 instances, equal to 39.4%) and lastly adjectives (8 instances, equal to 5.8%). As for input categories, Chinese coordinate compounds show a preference for verbs (98 have a [V + V] structure, 75 of which having verbal output and 23 having noun output). Endocentric compounds make up 81% of all coordinate compounds: two-headed compounds are equal to 74.4% of all coordinate compounds, while redundant right-headed compounds are 6.6% of all coordinate compounds. Exocentricity is equal to 19% (Table 14).

### 5.2 Structures of Chinese compounds

In order to have a better picture of Chinese compounding and of the variety of structures it presents, we put together the compounds previously gathered and analysed by Ceccagno and Scalise (2006) and the neologisms in our corpus. The resulting set of structures is now the database for Chinese in the Morbo/comp project. Table 15 presents an example for each available type of compound. The table shows the compound, its transcription in *pinyin*, output category, internal structure, classification, position of the head, glosses and translation.

As expected considering the dimension of the corpus, the neologisms reveal a greater number of structures as compared with those in Ceccagno and Scalise (2006). Among these, the most significant structure is the right-headed $[V + V]_V$ attributive, as in 品读 *pǐndú* 'decide with discrimination + read = read carefully', and redundant right-headed compounds as well, as in 峰位 *fēngwèi* 'peak + place = peak'.

Our corpus also presents three other structures, all of the endocentric type: SUB $[N + V]_N$, as in 楼花 *lóuhuā* 'floor + spend/use = building that is put up for sale before it is completed'; ATT $[N + V]_N$, as in 婚介 *hūnjiè* 'wedding +

**Table 15** Types of compound

| Compound | Pinyin | Class | Struct | CAT | Head | Gloss |
|---|---|---|---|---|---|---|
| 房型 | fángxíng | SUB | [N+N] | N | Right | House+model = layout of a house |
| 市道 | shìdào | SUB | [N+N] | N | Exocentric | Market+way = market prices |
| 楼花 | lóuhuā | SUB | [N+V] | N | Exocentric | Floor+spend/use = building that is put up for sale before it is completed |
| 监事 | jiānshì | SUB | [V+N] | N | Exocentric | Supervise+matter/responsibility = supervisor |
| 待岗 | dàigǎng | SUB | [V+N] | V | Left | Wait for+post = wait for a job |
| 割肉 | gēròu | SUB | [V+N] | V | Exocentric | Cut with a knife+meat = sell sth at a price lower than its original price |
| 攀高 | pāngāo | SUB | [V+A] | V | Left | Climb+high = climb up; rise |
| 入住 | rùzhù | SUB | [V+V] | V | Left | Come into/enter+live/stop = move into |
| 拒载 | jùzài | SUB | [V+V] | V | Left | Refuse+carry = (of a tax driver) refuse to take a passenger |
| *胆小* | *dǎnxiǎo* | *SUB* | *[N+A]* | *A* | *Exocentric* | *Guts/ courage+small = coward* |
| 失范 | shīfàn | SUB | [V+N] | A | Exocentric | Lose/deviate from the norm+model = irregular |
| 天价 | tiānjià | ATT | [N+N] | N | Right | Sky+price = prohibitive price |
| 色狼 | sèláng | ATT | [N+N] | N | Exocentric | Lust+wolf = sex maniac |
| 婚介 | hūnjiè | ATT | [N+V] | N | Exocentric | Wedding+introduce = matchmaking |
| 飘尘 | piāochén | ATT | [V+N] | N | Right | Float+dust = floating dust |
| 蹦床 | bèngchuáng | ATT | [V+N] | N | Exocentric | Jump+bed = trampoline |
| 速递 | sùdì | ATT | [A+V] | N | Exocentric | Fast+hand over/pass = express delivery |
| 黑车 | hēichē | ATT | [A+N] | N | Right | Black/illegal+vehicle = unlicensed car |
| 黄毒 | huángdú | ATT | [A+N] | N | Exocentric | Yellow/pornographic+poison = pornographic books |
| 互动 | hùdòng | ATT | [Adv+V] | N | Exocentric | Mutually+move = interaction |
| 飙升 | biāoshēng | ATT | [N+V] | V | Right | Whirlwind+rise = (of price, quantities) soar |
| 品读 | pǐndú | ATT | [V+V] | V | Right | Decide with discrimination+read = read carefully; ponder on |
| 航拍 | hángpāi | ATT | [V+V] | N | Exocentric | Navigate (by water or air)+take a photograph = aerial photograph |
| 完胜 | wánshèng | ATT | [A+V] | V | Right | Whole+win victory = win a complete victory |

**Table 15** continued

| Compound | Pinyin | Class | Struct | CAT | Head | Gloss |
|---|---|---|---|---|---|---|
| 突审 | tūshěn | ATT | [Adv+V] | V | Right | Unexpectedly+interrogate = interrogate sb. by surprise |
| 利淡 | lìdàn | ATT | [A+A] | N | Exocentric | Favourable+weak = unfavourable information for the market which may lead to a fall in prices |
| 雪白 | *xuěbái* | *ATT* | *[N+A]* | *A* | *Right* | *Snow+white = white like the snow* |
| 高发 | gāofā | ATT | [A+V] | A | Exocentric | High/above the average+deliver = frequent |
| 花心 | huāxīn | ATT | [A+N] | A | Exocentric | Attractive but unreal or insincere+heart = unfaithful |
| 统合 | tǒnghé | ATT | [Adv+A] | A | Right | All/totally+whole = uniform/overall |
| 频密 | pínmì | ATT | [Adv+A] | A | Exocentric | Frequently+close/dense = frequent |
| 梯次 | tīcì | ATT | [N+N] | Adv | Exocentric | Steps+order = by echelon or by group/ in order of age, size, etc. |
| 蔬果 | shūguǒ | CRD | [N+N] | N | Two-headed | Vegetables+fruit = vegetables and fruits |
| 东西 | *dōngxi* | *CRD* | *[N+N]* | *N* | *Exocentric* | *East+west = thing* |
| 峰位 | fēngwèi | CRD | [N+N] | N | Right | Peak+place = peak |
| 警示 | jǐngshì | CRD | [V+V] | N | Exocentric | Warm+show = warning |
| 高矮 | *gāoǎi* | *CRD* | *[A+A]* | *N* | *Exocentric* | *High+low = height* |
| 研发 | yánfā | CRD | [V+V] | V | Two-headed | Research/study+develop = research and develop |
| 疏离 | shūlí | CRD | [V+V] | V | Exocentric | Scatter/not familiar+be away from = be estranged and keep apart |
| 亮丽 | liànglì | CRD | [A+A] | A | Two-headed | Bright+beautiful = brilliant/bright and beautiful |

A = adjective, V = verb, N = noun, Adv = adverb; Highlighted italic cells = compounds from Ceccagno and Scalise (2006) not available in our corpus of neologisms (non productive)

introduce = matchmaking'; ATT [A + V]$_A$, as in 高发 *gāofā* 'high + deliver = frequent'.[15] Further, a number of endocentric structures in Ceccagno and Scalise (2006) appear in our corpus also in an exocentric form.

It shall be noted that the greater part of the types of compound presented in Ceccagno and Scalise (2006) appear also in the neologism corpus and are thus to be deemed productive. On the other hand, we were not able to find instances of compounds of the following types: ATT [N + A]$_A$, as in 雪白 *xuěbái* 'snow + white = white like the snow'; SUB [N + A]$_A$, as in 胆小 *dǎnxiǎo* 'guts/courage + small = coward'; exocentric CRD [N + N]$_N$, as in 东西 *dōngxi* 'east + west = thing'; exocentric CRD [A + A]$_N$, as in 高矮 *gāoǎi* 'high + low = height'. Therefore these types of compound do not appear to be productive.

The new table of Chinese compounds hence shows 30 types of compound (some of the types can be both endocentric and exocentric).[16] All in all, Chinese neologism compounds exhibit a wide variety of structures, significantly greater than those that can be found in English or Italian in the Morbo/comp database. English exhibits 15 types of structures (4 subordinate, 8 attributive and 3 coordinate); Italian exhibits 11 types of structures (3 subordinate, 5 attributive and 3 coordinate) whereas Chinese exhibits 30 types of structures (8 subordinate, 17 attributive and 5 coordinate).

Moreover, our data reveal that in Chinese, [V + V]$_V$ compounds can be classified in any of the three macro-types according to the relationship running between the two constituents: they can be coordinate compounds, in which the two constituents are linked by a coordinating relation, being it of conjunction as in 推展 *tuīzhǎn* 'recommend + exhibit = recommend and exhibit for sale', or synonymy as in 操控 *cāokòng* 'control + control = operate and control'; furthermore, there can exist right-headed attributive compounds, in which the non-head verb constituent acts as a modifier of the head, as in 跃增 *yuèzēng* 'leap/jump + increase/add = grow by leaps'; and there can also exist left-headed subordinate compounds, both of the resultative type, as in 入住 *rùzhù* 'come into/enter + live/stop = move into', and of the 'serial verbs' type, as in 竞卖 *jìngmài* 'compete + sell = compete to sell'.

The prediction of the syntactic and semantic relation between the two constituents of [V + V] compounds represents a very challenging topic of research, also because compound verbs lack morphological markings (cf. Chang et al. 2000; Chang and Chen 1999, among others).[17] For this reason, we have to focus on the interpretation of the whole compound. For example, in 拒载 *jùzài* 'refuse + carry = (of a tax driver) refuse to take a passenger', the meaning of the whole compounds makes clear that the event expressed by the verb on the right

---

[15] Resultative verbs and 'serial verbs', which are classified as coordinates in Ceccagno and Scalise (2006), are classified as subordinates in this work. Note also that Ceccagno and Scalise (2006) did not take account of compounds having constituents or output in adverb, that are viceversa present in our corpus.

[16] Here as with the subsequent cross-linguistic comparison we only considered nouns, adjectives, verbs and adverbs.

[17] The same difficult of interpretation is found in syntax too, where a sequence of two verbs without any overt subordinative or coordinative marker is often ambiguous and can generate different interpretations.

**Table 16** Output categories of Chinese compounds

|      | N    | V   | A   |
| ---- | ---- | --- | --- |
| TOT  | 395  | 249 | 23  |
| %    | 58.8 | 37  | 3.4 |

depends on that expressed by the verb on the left, so that the compound is a subordinate verb of the serial verbs type. If we were just to consider the two constituents the compound could also be interpreted as a coordinate compound, 'refuse and take'.

## 5.3 Output categories in Chinese compounds

According to our analysis, the most frequent lexical output category in Chinese is noun: in fact, there are 395 nouns in the corpus (58.8% of the total). In Chinese the productivity of verbal outputs is also high: verbs represent in fact 37% of the total (249) whereas the productivity of adjectives is modest: 3.4% (23) (Table 16).

With these data we are now able to formulate a new generalization on headedness in Chinese compounds.

## 6 Headedness in Chinese compounds

### 6.1 Exocentric compounds

In our corpus, 19.2% of Chinese compounds are exocentric (Table 17, Panel a), out of which 29.4% are subordinates, 50.4% attributives, 20.2% coordinate (Table 17, Panel b). Moreover, exocentricity affects all lexical categories.

Our corpus shows that exocentricity in Chinese compounding does not show strong peculiarities compared to other languages; in particular the great semantic opacity (as in 东西 *dōngxi* 'east + west = thing') highlighted by Scalise and Guevara (2006) does not appear to be productive. The output category most affected by exocentricity is noun (76% of exocentric compounds),[18]

**Table 17** Exocentricity

|                                  | Endocentric | Exocentric |      |
| -------------------------------- | ----------- | ---------- | ---- |
| a. Endocentric vs. Exocentric    |             |            |      |
| TOT                              | 543         | 129        |      |
| %                                | 80.8        | 19.2       |      |
|                                  | SUB         | ATT        | CRD  |
| b. Exocentricity among macro-types |           |            |      |
| TOT                              | 38          | 65         | 26   |
| %                                | 29.4        | 50.4       | 20.2 |

---

[18] 98 compounds, equal to 13.8% of the whole corpus.

| Table 18 Exocentric compounds | | N | A | V |
|---|---|---|---|---|
| | TOT | 98 | 15 | 14 |
| | % | 76 | 11.6 | 10.8 |

| Table 19 Endocentric compounds | | Right-headed | Left-headed | Two-headed |
|---|---|---|---|---|
| | TOT | 321 | 120 | 102 |
| | % | 59.1 | 22 | 18.8 |

| Table 20 Right-headed compounds | | N | V |
|---|---|---|---|
| | TOT | 275 | 45 |
| | % | 85.7 | 14 |

| Table 21 Left-headed compounds | | V |
|---|---|---|
| | TOT | 120 |
| | % | 100 |

whereas exocentric compounds with verbal (10.8% of exocentric compounds)[19] and adjectival output (11.6% of exocentric compounds)[20] are far less productive (Table 18).[21] Among exocentric compounds with noun output, roughly a quarter presents a $[V + V]_N$ structure, followed by $[V + N]_N$ structure.

### 6.2 Endocentric compounds

Endocentric compounds make up 80.8% of the corpus and are therefore the vast majority of the neologisms under examination. Among these, 59.1% are right-headed whereas the remaining part split between left-headed compounds (22%) and two-headed compounds (18.8%) (Table 19).

Right-headed compounds are for the majority nouns, 85.7% of the total. Verbs amount to 14% of right-headed compounds and are all attributive (Table 20).

Left-headed compounds represent 22% of endocentric compounds in our corpus and are all subordinate verbs (Table 21). They are usually compounds with a $[V + N]$ structure that present a verb-object relation between constituents, resultative and serial verb subordinate compounds.

Endocentric two-headed compounds are mostly verbs (70.6% of two-headed compounds), followed by nouns (21.6%); adjectives can be considered as marginal (7.8% of two-headed compounds). We can therefore conclude that two-headed compounds are productive and that this productivity applies in particular to compounds with verbal output (Table 22).

---

[19] 14 compounds, equal to 2% of the whole corpus.

[20] 15 compounds, equal to 2.1% of the whole corpus.

[21] In the corpus there are two exocentric adverbs too.

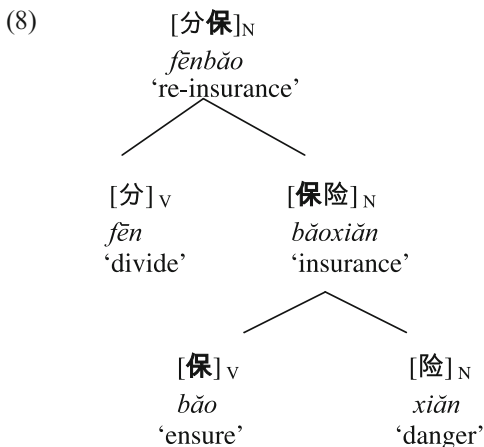| **Table 22** Two-headed compounds | V | N | A |
|---|---|---|---|
| TOT | 72 | 22 | 8 |
| % | 70.6 | 21.6 | 7.8 |

## 6.3 Metacompounding

The analysis of Chinese neologisms leads us to the identification of a phenomenon never singled out before that we call 'metacompounding' or compounds formed from compounded constituents (for further details, see Ceccagno and Basciano 2007). We should shed light on this phenomenon because the identification of the head at first appears to be problematic. In fact, these compounds initially appear to be exocentric, but they are endocentric indeed.

Metacompounding is a mode of compound formation in which at least one of the constituents refers to an underlying compound that does not appear on the surface. The constituent on the surface can therefore be regarded as a truncated form of the underlying compound. In order to clarify the semantics, the lexical category or the head of the compound, an analysis of the underlying compounds is also required. Let us consider a compound like 分保 *fēnbǎo*; the analysis of the surface compound would be as in (7):
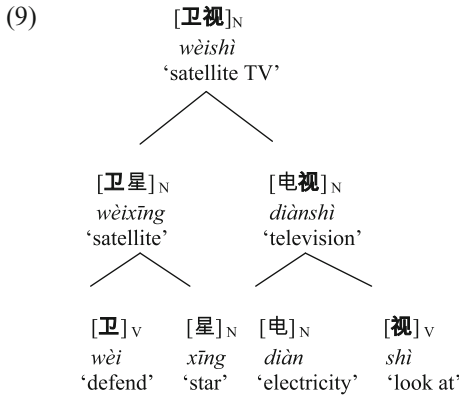
(7) 分保　fēnbǎo　CRD　V+V　N　? divide + protect/ensure = re-insurance

Considering both the structural and semantic point of view, the analysis of the compound appears to be problematic. This would not be the case if we were to analyse the second constituent of the compound, 保 *bǎo*, as a truncated form of the underlying compound, 保险 *bǎoxiǎn* ('insurance'). The analysis of the compound will thus be as follows:

(8)

[分保]ₙ
*fēnbǎo*
're-insurance'

[分]ᵥ          [保险]ₙ
*fēn*          *bǎoxiǎn*
'divide'       'insurance'

[保]ᵥ          [险]ₙ
*bǎo*          *xiǎn*
'ensure'       'danger'

Therefore, in itself the compound 分保 *fēnbǎo* presents both semantic and categorial opacity, and it is not possible to identify the head. However if we consider the underlying compound, the structure of the surface compound and its meaning become clear and the head can be identified.

The compound in (9) refers to two underlying compounds:

(9)

[卫视]ₙ
*wèishì*
'satellite TV'

[卫星]ₙ          [电视]ₙ
*wèixīng*         *diànshì*
'satellite'       'television'

[卫]ᵥ    [星]ₙ    [电]ₙ    [视]ᵥ
*wèi*    *xīng*   *diàn*   *shì*
'defend' 'star'   'electricity' 'look at'

What could appear as a [V + V]ₙ exocentric coordinate compound would be unfathomable from the semantic point of view if we did not refer to the underlying compounds 卫星 *wèixīng* and 电视 *diànshì* 'satellite + TV = satellite TV'. Furthermore, if we consider the two underlying compounds, the input lexical categories are not [V + V], but rather [N + N], and the compound is not exocentric but rather endocentric with head on the right; lastly, it is no longer a coordinate but rather an attributive. Therefore the whole structure of the compound changes when we refer to the underlying compounds; in this way these compounds can be analysed as regular compounds.

Summing up, we define as metacompounds those compounds that require reference to their underlying compounds in order to ascertain semantics, lexical category or head of the compound. A metacompound could be depicted as an iceberg, only the tip of which can be seen, but whose morphology can only be explained by reference to the submerged part.[22] Metacompounding deserves a more in-depth analysis, which is beyond the scope of this article.

---

[22] Metacompounds are in a way similar to the so-called blends (*motel* from 'motor hotel') and truncated compounds in Russian (Molinsky 1973; Billings 1997), as in *partorg* from 'partTIJnyj + organiZAtor' ('(political) party organizer'). The main difference lies in the nature of the truncated constituents, which in metacompounds are compound words. The phenomenon of metacompounding can be somehow related to Chinese phenomena such as abbreviations (cf. Ling 2000), or 'stump compounds' (cf. Li and Thompson 1978). Tibetan exhibits a phenomenon that is similar to metacompounding, called 'clipping' (Wälchli 2005). Further cross-linguistic investigations could show whether similar phenomena do exist in other languages of the Sino-Tibetan family.

# 7 Headedness in Chinese compounds

7.1 Headedness in Chinese compounds: a Cerberus

In Sect. 3 we discussed different generalizations on headedness in Chinese compounds. While sharing the theoretical and analytical framework of Ceccagno and Scalise's works (2006, 2007), our analysis of a larger database of Chinese neologisms enables us to reach different conclusions. Our database shows a significant number of left-headed compounds. In fact, verbal constructions of the verb-object and resultative types are a productive way of forming new compounds. Besides, another group of verbal compounds has emerged from our corpus: the serial verbs type. This leads us to state that the formation of left-headed compounds cannot be ascribed the status of a peripheral phenomenon—as would be the case with, e.g., right-headed compounds in Romance languages—but is rather a central feature of Chinese compounding. Furthermore, a distinctive element of Chinese is the productivity of two-headed endocentric coordinate compounds.

Table 23 shows the productivity of the three types of head features and of exocentric compounds: in our corpus of Chinese compounds 47.8% are right-headed; 17.8% are left-headed; 15.2% are two-headed while 19.2% are exocentric.

This prompts us to propose a new *Canonical Head Position Principle*. Unlike other languages for which the position of the morphological head has been studied and where the compound's canonical head has been identified to be either on the right (Germanic languages) or on the left (Romance languages), Chinese exhibits a Cerberus-like head in that within endocentric compounds three head positions—on the right, on the left and with both constituents as heads—do exist and are productive.

This Cerberus-like feature of the compounds in Chinese can have an impact on the theory of the head(s) in compounding in general. First, the hypothesis according to which the location of the head in compounds is a binary parameter (for further reading on parametric linguistics cf. Longobardi 2003; Gianollo et al. 2008, in press) is clearly challenged by Chinese. Second, our generalization that in Chinese compounding multiple locations of the head exist could also apply to other languages. In fact the generalization according to which the head has to be found invariably on the left for some languages and invariably on the right for some others could well only apply to nominal compounds. In Chinese nominal compounds, for example, the head is invariably to the right, in line with the modifier-modified order of the nominal phrase.

**Table 23** Endocentric compounds

|      | Right-headed | Left-headed | Two-headed | Exocentric |
|------|--------------|-------------|------------|------------|
| TOT  | 321          | 120         | 102        | 129        |
| %    | 47.8         | 17.8        | 15.2       | 19.2       |

## 7.2 Compounding and syntax

The existence of left-headed verbal compounds and of right-headed nomi-nal, verbal and adjectival compounds is in line with the syntax of Chinese, which is a (S)VO language but is left-branching as far as the nominal phrase is concerned. In the discussion of the classification of Chinese compounds we observed a number of similarities between the order of constituents in compounds and in syntactic constructions. Moreover, often syntactic constructions as well seem to show a high level of ambiguity as compounds do.

Anderson (1985) states that "several classes of compounds in Mandarin are structurally quite similar to syntactically created phrases. Sometimes the same sequence of formatives, in fact, may correspond either to a phrase or to a compound […]". Huang (1998) claims that in Chinese heads are ambiguous in phrasal syntax in ways that have not been sufficiently appreciated, even if he observes that word syntax in Chinese is not wholly predictable from phrasal syntax.

In fact, the order of compound constituents seems to follow the order in syntactic phrases. Subordinate [N + N] compounds are in accordance with the left-branching structure of the nominal phrase in Chinese, which exhibit a modifier-modified order (head to the right). Also subordinate [V + N] compounds, in which constituents have a verb-argument relation, follow the syntactic order (S)VO. In the case of subordinate $[V + V]_V$ compounds, the compounds show the same linear order as the corresponding syntactic con-structions.

With regard to attributive compounds, they follow the AN syntactic order. Furthermore, attributive $[V + N]_N$ compounds exhibit the same order as noun phrases in which a relative phrase precedes the head noun. Moreover, attrib-utive verbs with a [N + V] or [V + V] structure follow the syntactic order of verb phrases. In fact, as for verbal phrases, Chinese exhibits a (left-headed) VO order with direct internal argument, but constructs adjuncts to the left (adjunct-verb order, head on the right). These parallels seem to suggest a strong correlation between morphology and syntax.

## 8 Conclusions

Based on an analysis of Chinese neologisms this article presents an exhaustive picture of the variety of structures existing in Chinese compounds. In the analysis of Chinese compounds we have adopted the classification scheme put forth by Bisetto and Scalise (2005) and refined it proposing a new definition of the three macro-types. We were thus able to classify the corpus of compounds, to investigate the productivity of the different types of compound and to give insights into the distribution of compounds across macro-types. We were also able to identify further types of compound with respect to those proposed in previous studies.

Our corpus of neologisms has enabled us to challenge the existing approaches to headedness. We have shown that Chinese is a language with three head position in compounding, since it consistently forms right-headed compounds (noun, verbs and adjectives), left-headed compounds (subordinate verbs) and two-headed compounds (coordinate compounds). This led us to a generalization according to which in Chinese it is not possible to identify the head according to an either/or position criterion (right *or* left). Like a Cerberus, Chinese exhibits three positions for the head, all of which are productive and not peripheral. The identification of a language not exclusively right-headed or left-headed in its compounding morphology seems to rule out the possibility to consider the head as a binary parameter to be fixed in the languages of the world, either on the left or on the right. This issue represents a challenge for the cross-linguistic study of compounding.

In addition, we formulated the hypothesis that the rule according to which in all languages of the world the head should invariably be found on the left or on the right only applies to nominal compounds. If this were the case, it would then be possible to verify whether the position of the head in compounding in different languages is somehow linked to different lexical categories, as it is in Chinese.

# References

Anderson, S. R. (1985). Typological distinctions in word formation. In S. Timothy (Ed.), *Language typology and syntactic description* (Vol. 3, pp. 150–201). Cambridge: Cambridge University Press.

Baker, M. C. (1986). Object sharing and projection in serial verb constructions. *Linguistic Inquiry, 20*(1), 513–553.

Billings, L. A. (1997). Morphology and syntax: Delimiting stump compounds in Russian. In *Proceedings of the First Mediterranean Morphology Meeting (MMM1)*, Mytilene, University of Patras.

Bisetto A., & Scalise, S. (2002). Trasferimento di costrutti sintattici in morfologia. *Lingue e Linguaggio, 2*, 237–261.

Bisetto, A., & Scalise, S. (2005). Classification of compounds. *Lingue e Linguaggio, 2*, 319–332.

Ceccagno, A. (2008). Regolarità e peculiarità dei composti del cinese. Venezia, Cafoscarina (in press).

Ceccagno, A., & Basciano, B. (2007). Complessità della morfologia del cinese. *Atti dell'XI convegno AISC,* Roma 22–24 febbraio 2007 (in press).

Ceccagno, A., & Scalise, S. (2006). Classification, structure and headedness of Chinese compounds. *Lingue e Linguaggio, 2*, 233–260.

Ceccagno, A., & Scalise, S. (2007). Composti del cinese: analisi delle strutture e identificazione della testa. In A. Palermo (Ed.), "La Cina e l'altro", Napoli, Università degli studi di Napoli "L'orientale", pp. 503–543.

Chang, L.-L., & Chen, K.-J. (1999). 動詞詞構與語法功能互動初探 *Dòngcí cí gòu yǔ yǔfǎ gōngnéng hùdòng chū tàn. Proceedings of the 12th Research on Computational Linguistics* (pp. 67–85). Tsinchu, Taiwan.

Chang, L.-L., Chen, K.-J., & Huang, C.-R. (2000). A lexical-semantic analysis of Mandarin Chinese verbs: Representation and methodology. *Computational Linguistics and Chinese Language Processing, 5*(1), 1–18.

Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.

Cheng, L. L. (1997). Resultative compounds and lexical relational structures. *Chinese Languages and Linguistics III: Morphology and Lexicon,* 167–197.

Chi, T. R. (1985). *A lexical analysis of verb-noun compounds in Mandarin Chinese.* Taipei: Crane.

Corbin, D. (1987). *Morphologie dérivationelle et structuration du lexique*. Tuebingen: Niemeyer.

Di Sciullo, A. M., & Williams, E. (1987). *On the definition of word*, Cambridge MA: MIT Press.

Dryer, M. S. (1992). The greenbergian word order correlations. *Language, 68*, 81–138.

Gianollo, C., Guardiano, C., & Longobardi, G. (2008). Three fundamental issues in parametric linguistics. In T. Biberauer (Ed.), *The limits of syntactic variation*. Amsterdam: Benjamins (in press).

Giegerich, H. (2004). Compound or phrase? English noun-plus-noun constructions and the stress. criterion. *English Language and Linguistics, 8*, 1–24.

Huang, C. T. J. (1984). Phrase structure, lexical integrity, and Chinese compounds. *Journal of Chinese Teacher's Association, 19*(2), 53–78.

Huang, C.-T. J. (1992). Complex predicates. In R. Larson, S. Iatroudou, U. Lahiri, & J. Higginbotham (Eds.), *Control and grammar* (pp. 109–147). Boston: Kluwer Academic Publisher.

Huang, S. (1998). Chinese as a headless language. In J. Packard (Ed.), *Compounding morphology: new approaches to Chinese word formation* (pp. 261–283). Berlin and New York: Mouton de Gruyter.

Levin, B., & Rappaport Hovan, M. (1995). *Unaccusativity: At the syntax-lexical semantics interface.* Cambridge, MA: MIT Press.

Li, Y. (1990). On V-V compounds in Chinese. *Natural Language and Linguistic Theory, 8*, 177–207.

Li, Y. (2005). *A theory of the morphology-syntax interface*. Cambridge (MA): MIT Press.

Li, C. N., & Thompson, S. A. (1978). An explortion of Mandarin Chinese. In W. P. Lehmann (Ed.), *Syntactic typology: Studies in the phenomenology of language* (pp. 233–266). University of Texas Press.

Li, C. N., & Thompson, S. A. (1981). *A grammar of spoken Chinese: A functional reference grammar*. Berkeley: University of California Press.

Lin, F. (1990). *The verb-complement (V-R) compounds in Mandarin*. M.A. Thesis, National Tsing Hua University.

Ling, Y. (2000). 现代汉语缩略语 *Xiàndài Hàn yǔ suōlüèyǔ* (Abbreviations in Contemporary Chinese). Beijing: Yuwen chubanshe.

Longobardi, G. (2003). La linguistica parametrica: un nuovo programma di ricerca tra presente futuro delle scienze del linguaggio. *Lingue e Linguaggio, 1*, 3–30.

Masini, F. (1993). *The formation of modern Chinese Lexicon and its evolution toward a National Language: The period from 1840 to 1898*. Monograph No. 6 of the "Journal of Chinese Linguistics". Berkeley: University of California Press.

Molinsky, S. J. (1973). *Patterns of ellipsis in Russian compound noun formations.* Mouton: The Hague.

Morbo/Comp, http://morbocomp.sslmit.unibo.it/index.php?section=home.

Olsen, S. (2001). Copulative compounds: A closer look at the interface between morphology and syntax. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 2000* (pp. 279–320). Dordrecht: Kluwer.

Packard, J. (2000). *The morphology of Chinese*. Cambridge: Cambridge University Press.

Pan, W., Ye. B., & Han, Y. (2004). 汉语的构词法研究 *Hànyǔ de gòucífǎ yánjiū* (Research on Word Formation in Chinese). Shanghai: Huadong Shifan Daxue Chubanshe.

Plag, I. (1998). The polysemy of -ize derivatives. The role of semantics in word formation. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1997* (pp. 219–242). Dordrecht: Foris.

Plag, I. (2003). *Word-formation in English*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Plag, I., Kunter, G., & Lappe, S. (2007). Testing hypotheses about compound stress assignment in English: A corpus-based investigation. *Corpus Linguistics and Linguistic Theory, 3*(2), 199–233.

Ross, C. (1998). Cognate objects and the realization of thematic structures in Mandarin Chinese. In J. Packard (Ed.), *Compounding morphology*: *New approaches to Chinese word formation* (pp. 205–253), Berlin and New York: Mouton de Gruyter.

Scalise S. (1983). *Generative morphology*. Dordrecht: Foris.

Scalise, S. (1994). *Morfologia*. Il Mulino: Bologna.

Scalise, S., Bisetto, A., & Guevara, E. (2005). Selection in compounding and derivation. In Dressler, et al. (Eds.), *Morphology and its demarcations* (pp. 133–150). Amsterdam: Benjamins.

Scalise, S., & Guevara, E. (2006). Exocentric compounding in a typological framework. *Lingue e Linguaggio, 2*, 185–206.

Selkirk, E. (1982). *The syntax of words*. Cambridge, MA: MIT Press.

Spencer, A., & Zaretskaya, M. D. (1998). Verb prefixation in Russian as lexical subordination. *Linguistics, 36*, 1–40.

Starosta, S., Kuiper, K., Ng, S., & Wu, Z. (1998). On defining the Chinese compound word: Headedness in Chinese compounding and Chinese VR compounds. In J. Packard (Ed.), *New Approaches to Chinese word formation* (pp. 346–369). Berlin and New York: Mouton de Gruyter.

Sun, C. (2006). *Chinese: A linguistic introduction*. Cambridge University Press.

*The Contemporary Chinese Dictionary [Chinese-English Edition]* 现代汉语词典 [汉英双语] *Xiàndài Hànyǔ cídiǎn [Hàn Yīng shuāng yǔ]* (2000). Beijing: Foreign Language Teaching and Research Press.

Trommelen, M., & Zonneveld, W. (1986). Dutch morphology, evidence for the right-hand head Rule. *Linguistic Inquiry, 17*, 147–169.

Varela, S. (1990). *Fundamentos de morfologia*. Madrid: Sintesis.

Wälchli, B. (2005). *Co-compounds and natural coordination*. Oxford: Oxford University Press.

Williams, E. (1981). On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry, 12*, 245–274.

Xing, J. Z. (2006). *Teaching and learning Chinese as a foreign language: A pedagogical grammar*. Hong Kong: Hong Kong University Press.

Zhou, X., Marslen-Wilson, W., & Taft, M. (1999). Morphology, ortography and phonology in reading Chinese compound words. *Language and Cognitive Processes, 14*(5/6), 525–565.