Left is right, right is not

On the constituency of the classifier phrase in Chinese

One-Soon Her and Hui-Chin Tsai
National Chengchi University

This paper argues for the left-branching constituency of the Chinese classifier phrase and demonstrates that the right-branching approach assumed by the majority of current syntactic works is not viable. The rejection of the right-branching approach entails the rejection of the 'split' approach, where both left- and right-branching structures are required. In this debate, we offer a vital fresh perspective from the syntax and mathematics of complex numerals. We examine the right-branching argumentation in A. Li (2014), which, crucially, extends Ionin & Matushansky's (2006) non-constituent account of complex numerals, e.g. *two hundred*, in non-classifier languages like English to Chinese and must rely on ellipsis and a silent element YIDIAR 'a bit'. Yet, complex numerals in Chinese, e.g. *liang bai* '200', are in fact constituents (He 2015), and the alleged YIDIAR 'a bit' does affect the semantics of the noun phrase and is thus by definition illicit (Her & Tsai 2014; 2015). Other evidence comes from Chinese synchronic and diachronic syntax as well as the typology of classifier word orders. While the overall argumentation centers on Chinese, it has significant cross-linguistic implications.

Key words: numeral classifier, complex numerals, left-branching, right-branching

1. Introduction

Nominal expressions in classifier languages that consist of a numeral (Num), a classifier (C) or a measure word (M), and a noun (N), mathematically, can have six word orders, as shown in (1).[1] Yet, only four are attested cross-linguistically (Greenberg 1990[1972]: 185, Aikhenvald 2000: 104–105).

[1] Abbreviations used in this paper are as follows: C: classifier; M: measure word; DE: marker of modifying phrases, e.g. genitive phrases and relative clauses; ASP: aspect marker; Num: numeral; N: noun; Q: question particle.

(1) a. √ [Num C/M N] (many languages, e.g. Chinese)
b. √ [N Num C/M] (many languages, e.g. Thai)
c. √ [C/M Num N] (few languages, e.g. Ibibio (Niger-Congo))
d. √ [N C/M Num] (few languages e.g. Jingpho (Tibeto-Burman))
e. * [C/M N Num] (no languages)[2]
f. * [Num N C/M] (no languages)

Furthermore, orders where [Num > C/M], i.e. (1a) and (1b), are much more frequent than those with [C/M > Num], i.e. (1c) and (1d) (Greenberg 1990[1972]; Her *et al.* 2016). In Mandarin, for example, Num precedes C/M, whereas in Jingpho, C/M precedes Num, as in (2a) and (2b), respectively.

(2) a. *san wei laoshi*
3 C teacher
'3 teachers'
b. *sara marai masum*
teacher C 3
'3 teachers'

Inspired by Greenberg's Universal 20 on the word order typology of Dem, Num, Adj, and N, Her (2017b) proposes a generalization, dubbed 'Greenberg's Universal 20A', for C/M's word order typology.

(3) Greenberg's Universal 20A
a. Among Num, C/M, and N, any order is possible as long as N does not intervene between Num and C/M.
b. Between Num and C/M, either order is possible, but [Num C/M] is more common.

An N-parameter thus obtains, as in (4a), given that N appears either on the left or right edge of the classifier phrase. Between Num and C/M, a C/M-parameter also obtains, as in (4b). The interaction of the two parameters produces a taxonomy of the four attested orders, as in Table 1 (Her 2017b: 55).

(4) N-parameter: N-initial [N…] vs. N-final […N]
C/M-parameter: C/M-final [Num C/M] vs. C/M-initial [C/M Num]

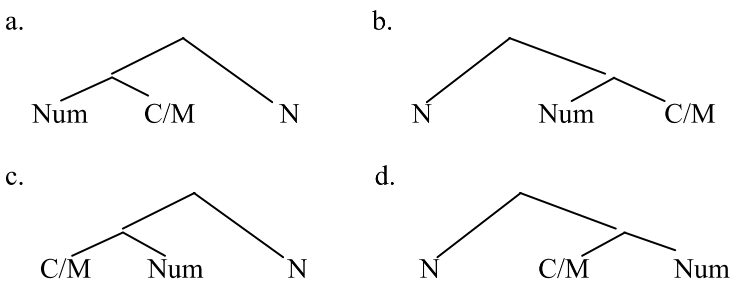
[2] It is well-known that there are alleged exceptions to this word order. However, as demonstrated in Her (2017a,b), in the case of Ejagham, the putative classifiers are in fact nouns, and in the cases of certain Tibeto-Burman and Tai-Kadai languages, the alleged numeral *one* that allows this word order is in fact an indefinite article.

Table 1. Taxonomy of C/M Word Orders (Her 2017b:55)

	N-final	**N-initial**
C/M-final	**(A)** [[Num **C/M**] **N**] e.g. Chinese, Miao	**(B)** [**N** [Num **C/M**]] e.g. Thai, Japanese
C/M-initial	**(C)** [[**C/M** Num] **N**] e.g. Ibibio, Kiriwina	**(D)** [**N** [**C/M** Num]] e.g. Jingpho, Bodo

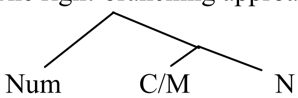
The most fundamental issue regarding the structure of the classifier phrase is whether C/M forms a constituent with Num or with N first. Greenberg (1990[1972]: 227) assumes that Num first forms a unit with C/M. An advantage of this traditional view is that it offers a simple and straightforward account of C/M's word order typology, shown in (5).

(5) A parameter account of C/M word order typology
(Her 2017b: 53–54)



However, in the more recent syntactic works other views have emerged and there are now three competing approaches (e.g. Zhang 2011; Her 2012a; A. Li 2014). In one approach, C/M and N form a constituent first, as shown in (6) (e.g. Cheng & Sybesma 1998, 1999; Borer 2005; Watanabe 2006; Huang *et al.* 2009, among others). This right-branching approach has become the dominant view often assumed in the more recent formalist works.

(6) The right-branching approach to C/M phrases



The traditional view in (5), known as the left-branching approach, though less popular recently, has by no means gone away; e.g. Muromatsu (1998) and Bhattacharya (1999; 2001) regard [Num C/M] as a complex functional head, while others consider it phrasal (e.g. Li & Thompson 1981; Huang 1982; Croft 1994; Lin 1997; Fukui & Takano 2000; Hsieh 2008; He 2016).

Still, a few other syntacticians contend that both structures, the left-branching and the right-branching, are required (e.g. Jenks 2010; Huang & Ochi 2011; X. Li 2011; Liu 2013; Zhang 2011, 2013). This is thus known as the split approach.

Note that the above three-way classification of syntactic approaches to the constituency of the [Num C/M N] phrase is solely based on the quantity-denoting reading of this structure, and thus does not concern the particular property-denoting reading that is also available in a [Num C/M *de* N] phrase, e.g. *san bang de xigua* (three pound DE watermelon) has the usual quantity-reading 'three pounds of watermelons' as well as a property-denoting reading 'three-pound watermelons', meaning watermelons that weigh three pounds. In the latter reading the [Num C/M *de*] portion functions as a modifier of N. This property-denoting reading will be discussed in §4, where A. Li's (2014) proposal is reviewed.

In this paper, we defend the unified left-branching structure of the classifier phrase in Chinese. The paper is organized as follows. Section 2 first offers a taxonomy of the various positions taken up by the many important works in this debate and draws a crucial logical conclusion regarding the split approach, i.e. that the rejection of either the left- or right-branching structure must entail the rejection of the split approach, where both left- and right-branching structures are required. Then, §3 starts deliberating the 'left-or-right' question, and, crucially, brings in a fresh perspective, i.e. complex numerals such as *san-bai* 'three hundred', and discusses whether they are constituents or not. A. Li (2014) is the first to consider complex numerals in defending the right-branching approach and assumes Ionin & Matushansky's (2006) view that complex numerals are not constituents. We support He's (2015) view that they are constituents, a fact that favors the left-branching approach, not only for Chinese but also for the C/M word order typology described above. Section 4 revisits A. Li's (2014) account of the post-C/M *ban/duo* 'half/more' in relation to complex numerals, the property-denoting use of the [Num C/M] phrase, and the phonologically inserted *de* in the quantity-denoting C/M phrase. Section 5 concludes the paper.

2. Two Camps, Four Positions

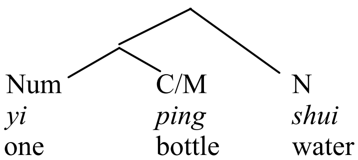
Within Chinese linguistics, there has been considerable research on the constituency of the [Num C/M N] construction. Various views can be categorized as two camps and four positions. The most fundamental disagreement is whether C/M has a unified structure in Chinese. In the unified camp, Cs and Ms belong to the same syntactic category and occupy the same structural position; however, there is no consensus as to whether this structure is left-branching or right-branching.

Those in the unified camp disagree as to whether the structure is left-branching or right-branching. Li & Thompson (1981), Huang (1982), Lin (1997), Hsieh (2008), Her (2012a), and R. Yang (2001) opt for the left, while the unified right-branching analysis appears to be more prominent in recent times, e.g. most notably Simpson (2005), Borer (2005), Huang *et al.* (2009), A. Li (1999; 2014), among many others.

A dissenting view also exists, that both structures are necessary for [Num C/M N] in Chinese. However, those in this split camp also disagree. For Zhang (2011; 2013), there is a split between Cs and Ms in that Cs involve a right-branching structure, and Ms, left-branching. Yet, X. Li (2011) and Li and Rothstein (2012) contend that the spilt resides in C/M's two readings: measure/quantity vs. counting, encoded respectively by a left-branching and right-branching structure.

(7) *yi* *ping* *shui*
1 M-bottle water
'one bottle of water'

(8) Measure reading



(9) *yi* *ke* *pingguo*
1 C apple
'one apple'

(10) Counting reading

Num	C/M	N
yi	*ke*	*pingguo*
one	C	apple

X. Li (2011) claims that Num is required as part of the complex classifier in a measure reading, as in (8). On a counting reading, when Num is 1 it is a modifier of the [C/M N] constituent, and can thus be omitted, as in (10). Yet, Her *et al.* (2015:207) demonstrate that Num 1 in the measure reading can be omitted, as in (11) and (12).

(11)

Ta-de	*jiuliang*	*buxing,*	*zuiduo*	*zhi*	*neng*
his	drinking-ability	not-good	at-most	only	can
he	ping	pijiu.			
drink	M-bottle	beer			

'He's not a good drinker, so he can only drink one bottle of beer at most.'

(12)

Wo	*xuyao*	*mai*	*jin*	*mi*.
I	need	buy	catty	rice

'I need to buy a catty of rice.'

Zhang (2011; 2013), on the other hand, presents arguments from four aspects for a split analysis: the scope of a left-peripheral modifier, the effect of modifier association, the semantic selection relation between a classifier and a noun, and the order of shape and size modifier. However, Her (2012a), Liu (2013), and A. Li (2014) demonstrate convincingly that there is no difference in how left-peripheral modifiers scope over different types of C/Ms. A. Li (2014) further indicates that the semantic selection between a classifier and a noun will not be a tenable test under the notion of extended projection (see Grimshaw (2000) and A. Li's earlier works). Specifically, nouns and their related functional heads, e.g. D, Num, C, can form extended projections. It is therefore possible for a verb to bear selection restrictions on a noun inside extended projection such as DP.

To summarize, there are two camps as to the structure of [Num C/M N]: unified vs. split. In the unified camp, there are two positions: left- vs. right-branching. In the split camp, there are also two positions: left-branching for M and right-branching for C vs. left-branching for measuring and right-branching for counting. In this paper, we shall not review the various arguments put forth for the four positions individually;

rather, our main strategy is to demonstrate that, once the complex numerals are taken into consideration, it will be clear that the unified left-branching position is the only viable option and the other three positions will be naturally ruled out.

Again, the focus of this paper is on the constituency of the quantity-denoting classifier phrase [Num C/M N] in Chinese. The most prominent left-branching advocate is Her (2012a), while the most prominent right-branching defender is A. Li (2014), where the arguments offered by Her (2012a) are disputed. We shall demonstrate that the right-branching approach defended by A. Li (2014) and assumed by the majority of current syntactic works is problematic and that a unified left-branching structure is viable. Note that, logically, the rejection of the right-branching structure logically entails the rejection of the split approach, where both left- and right-branching structures are required.

3. Constituency of Complex Numerals

It is well-accepted that the internal structure of complex numerals involving multiplication and addition are composed linguistically and follow standard syntactic and semantic composition (e.g. Hurford 1975, 1987, 2001; Ionin & Matushansky 2006; He 2015). In §3.1, we shall first explain the significance of multiplicative numeral bases in Her's (2012b) multiplicative theory of C/M. Then, in §3.2, we review Ionin & Matushansky's (2006) non-constituent account of complex numerals in English, and in §3.3, we demonstrate He's (2015) constituency account of Chinese complex numerals. In §3.4, we demonstrate that He's constituency account can be logically extended to the constituency of [Num C/M].

3.1 A multiplicative theory of C/M and its implications on C/M constituency

The numeral system in Chinese is a textbook example of this wide-spread pattern in languages: [(*n* × *base*) + *m*, *where m* < *base*] (e.g. Comrie 2006; 2016). A multiplicative complex numeral in Chinese thus employs simple multiplication [(*n* × *base*)], e.g. *jiu-qian* [*jiu* '9' × *qian* '1000'] and *liu-shi* [*liu* '6' × *shi* '10']. Au Yeung (2005; 2007) and Her (2012b) propose that the underlying function between Num and C/M in the [Num C/M N] construction can likewise be viewed as multiplication, meaning that [Num C/M] and [*n* × *base*] are both [*multiplier* × *multiplicand*]. C and M are thus unified under the concept of multiplicand; however, they can also be accurately distinguished in terms of the mathematical value they encode (Her

2012a), which explains C and M's semantic differences observed by Her & Hsieh (2010).

(13) Convergence and Divergence of C and M (Her 2012b:1679)
[Num K N] = [[n × x] N], where K = C iff x =1, otherwise K = M.

The idea that all Cs function as a multiplicand *1* is first proposed by Greenberg (1990[1972]:172), i.e. 'all the classifiers are … merely so many ways of saying 'one' or, more accurately, 'times one.' In (14), in spite of the three different Cs, the truth value of the phrase remains the same: *five eggs*. In (13), Her (2012b) extends the same concept of multiplicand to Ms, which thus have a value that is not necessarily *1*. Thus, in (15), *dui* has a numerical value of *2*, *da* has a numerical value of *12*, and *pai* 'row' has a variable numeral value that is larger than *2*. Thus, each of the three Ms has a different truth value.

(14) *wu* *ge/ke/li* *jidan* ([*5×1*]=[*5×1*]=[*5×1*])
five C egg
'five eggs'

(15) *wu* *dui/da/pai* *jidan* ([*5×2*]≠[*5×12*] ≠[*5×row*])
five M-pair/dozen/row egg
'five pairs/dozens/rows of eggs'

From a mathematical point of view, the [Num C/M] constituent and multiplicative complex numerals thus share the internal multiplication-based structure, e.g. *san bai* '300' and *san da* '3 dozen' are both analyzable as [3× *bai/da*], or [3× *100/12*]. Thus, C/M functions as a multiplicand, and Num, a multiplier. Her (2012a) thus argues that, given the [multiplier-multiplicand] function of [Num C/M], the latter must form a constituent, and thus the left-branching constituency should be adopted.

Nonetheless, A. Li's (2014) specific arguments against Her's (2012a) [Num C/M] constituency account are crucially based on Ionin & Matushansky's (2006) non-constituent approach to complex numerals in non-classifier languages such as English or Russian. It is therefore critical to examine whether this non-constituent account of complex numerals can be extended to Chinese or not, a classifier language.

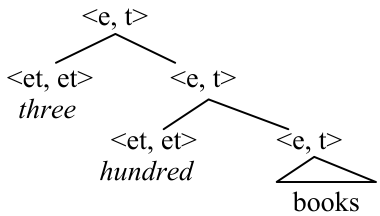
3.2 Ionin & Matushansky (2006): Non-constituent account of English numerals

According to Ionin & Matushansky (2006) (I&M hereafter), cardinals in languages such as Russian and Inari Sami are responsible for case assignment to their sister nouns, as shown in (16).

(16) *čiččâm* *čyeti* *pärnid* (Inari Sami)
seven hundred child-$_{\text{PART.SG}}$
'seven hundred children'

I&M thus contend that numerals, e.g. *two, hundred,* and *thousand* must be nominal heads selecting lexical nouns or other numeral-noun combinations as complements. A complex numeral such as *three hundred books* thus projects a complementative structure where the multiplier-multiplicand relation is captured by a head-complement structure and the complex numeral does not form a constituent, as shown in (17), which receives this interpretation: *three groups of one hundred books*.

(17) Nominal phrases in non-classifier languages
(Ionin & Matushansky 2006)



Consequently, additive complex numerals must be derived via nominal conjunction with NP deletion or right-node raising. For example, in an expression like *three hundred and twenty books*, each coordinated cardinal contains an instance of the NP *books* in its underlying source form, as in (18a), where the head noun is either PF-deleted, as in (18b), or right-node-raised, as in (18c).

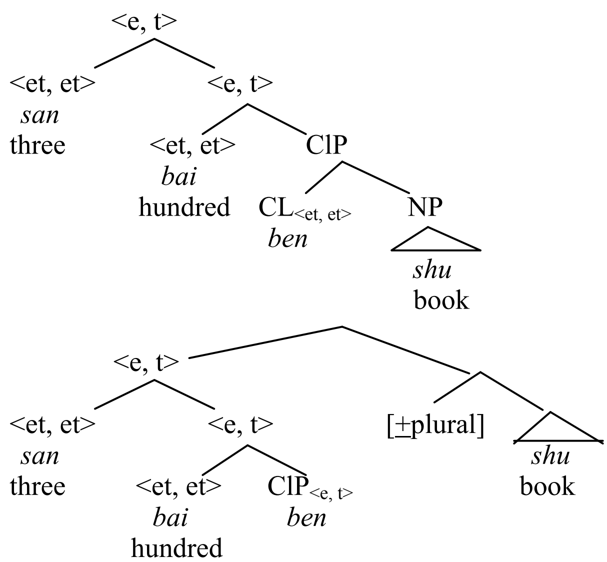
(18) five hundred and twenty books
a. [five hundred books] and [twenty books]
b. [five hundred ~~books~~] and [twenty books]
c. [[[five hundred t_i] and [twenty t_i]] books$_i$]

In short, the multiplier-multiplicand relation should be captured by a head-complement structure, namely, complex numerals like *five hundred and twenty* do not form a constituent. Bare numerals thus also involve ellipsis.

However, A. Li's (2014) extending I&M's right-branching complex numerals to Chinese is a bit hasty, as I&M suggest two possible accounts for classifier languages like Chinese, as in (19a), which has a right-branching [Num [C/M N]], and (19b), which has a left-branching [[Num C/M] N], precisely the two approaches debated between A. Li (2014) and Her (2012a).

(19) Nominal phrases in classifier languages (I&M 2006:328 (22a) & (22b))[3]

a. b.



Yet, the complex numeral *san-bai* does not form a constituent in either (19a) or (19b). We shall demonstrate that the left-branching [[Num C] N] of (19b) is on the right track, except that when Num is a complex numeral, it does form a constituent on its own before it merges with C/M.

---

3 In Ionin & Matushansky (2006), the left-branching structure of (19b) is labeled NumP with [±plural] as the head Num^0. Not to confuse their label Num, which refers to number distinction of singular and plural, and our use of Num, which refers to numerical quantifiers, I have left out some of the labels in (19b).

3.3 He's (2015): constituency account of Chinese numerals

Contra I&M's non-constituent approach to complex numerals, He (2015) follows the traditional approach (e.g. Jackendoff 1977; Selkirk 1977; Borer 2005; Corver & Zwarts 2006; Kayne 2010, among others) and argues that complex numerals are phrasal constituents at least in Chinese, a classifier language. Among the various kinds of support that He (2015) offers from Chinese syntax, semantics, and morpho-phonology, three pieces of evidence in our view stand out in particular.

First, it is the behavior of approximant numerals which denote a small portion of the quantity denoted by an adjacent numeral base, e.g. *lai* denotes a number generally within the range of ±10% of the left adjacent base (He 2015).[4] In (20), *lai* is adjacent to *bai* 'hundred', so the understood value must be obtained in relation to *bai*, thus between 90 and 110. Under I&M's (2006) right-branching non-constituent analysis, (20) must have (21a) as its underlying form, which undergoes either the deletion (21b) or right node raising (21c) of the constituent [*ge xuesheng*]. Under I&M's left-branching non-constituent analysis, (20) must have (22a) as its underlying form, which undergoes either the deletion (22b) or right node raising (22c) of the C/M [*ge*]. Note that the capitalized AND indicates a silent conjunction.

(20)	*yi*	*bai*	*lai*	*ge*	*xuesheng*
	one	hundred	approach	C	student
	'around one hundred students'				

(21) a. [*yi bai ge xuesheng*] AND [*lai ge xuesheng*]
b. [*yi bai ~~ge xuesheng~~*] AND [*lai ge xuesheng*]
c. [*yi bai* t_i] AND [*lai* t_i] <u>*ge xuesheng*</u>$_i$

(22) a. [[*yi bai ge*] AND [*lai ge*]] *xuesheng*
b. [[*yi bai ~~ge~~*] AND [*lai ge*]] *xuesheng*
c. [[*yi bai* t_i] AND [*lai* t_i]] <u>*ge*</u>$_i$ *xuesheng*

---

[4] An anonymous reviewer states that a number of native speakers have confirmed that *(yi) bai lai ge* in (20) cannot refer to a number below 100, and more generally in fact, Num-*lai* cannot refer to a quantity lower than or equal to Num, and *lai* is thus interpreted similarly as *duo* 'more'. For us, this is only true for base *shi* 'ten', but not true for bases above ten, e.g. *bai* 'hundred' and *qian* 'thousand'. Our judgment is likewise confirmed by a number of native speakers of Taiwan Mandarin. He's (2015) judgment is largely based on Putonghua in Mainland China. Suffice to say that there are dialectal and personal variations here. Yet, this fact does not affect the argumentation we put forth.

In all four proposed derivations in (21) and (22), the expression **lai ge* (*xuesheng*) 'approach C student' is incorrectly predicted to be well-formed without being adjacent to a numeral base such as *bai* 'hundred'. Following I&M (2006), in a non-constituent appraoch to complex numerals, *lai* and the preceding numeral base are interrupted by a C/M or a C/M and an N. Logically, the string [*yi-bai lai*] either is or is not a constituent. The simple fact is that *lai* is uninterpretable without an *immediately* preceding numeral base. The failure of the non-constituent approach thus indicates that [*yi-bai lai*] must be a constituent.

As an anonymous reviewer points out, *lai* in (23) below can only be interpreted in accordance with the immediately preceding base, *shi* 'ten', not *bai* 'hundred', nor *shi* 'ten' and *bai* 'hundred' together. This fact can be accounted for straightforwardly by the general pattern of complex numerals, repeated in (24).

(23) *yi bai er shi lai ge xuesheng*
one hundred two ten approach C student
'around one hundred and twenty students'

(24) General Pattern of Number Systems in Languages (Comrie 2006)
For base *b*: $(n \times b) + m$ (where $m < b$)

Thus, in (23), *lai*, which occupies *m* in the above pattern, must be interpreted in accordance with *shi* 'ten', which occupies *b* in the above pattern, and not *bai* 'hundred', which is embedded in the internal structure of *n* and is thus inaccessible to *lai*. Given $m < b$, *lai* thus must receive an interpretation that is smaller than *shi* 'ten'. This general pattern also explains the rigid word order [*yi bai*] > [*er shi*] > [*lai*]; no flipping around is allowed.

I&M's non-constituent approach also encounters a serious semantic problem with conjoined NPs. Examine example (25). Again, in their right-branching options of derivation, an underlying source form (26) is assumed, while the left-branching option assumes (27) as source form.

(25) *shi wu ge nanren he nuren*
ten five C man and woman
'fifteen men and women'

(26) a. [*shi ge nanren he nuren*] AND [*wu ge nanren he nuren*]
b. [*shi ~~ge nanren he nuren~~*] AND [*wu ge nanren he nuren*]
c. [*shi* t_i] AND [*wu* t_i] *ge nanren he nuren*$_i$

(27) a. [[*shi ge*] AND [*wu ge*]] *nanren he nuren*
b. [[*shi ~~ge~~*] AND [*wu ge*]] *nanren he nuren*
c. [[*shi* t_i] AND [*wu* t_i]] *ge*$_i$ *nanren he nuren*

Thus, (25) is expected to have the same meaning as (26) or (27). However, this is not true. For (25), a reading with either 0 men or 0 women is impossible; a well-formed reading minimally requires 1 woman and 1 man. The reading of a group of 14 men and 1 woman or 1 man and 14 women is therefore allowed, but such a reading is not possible for (26) or (27), where the reading of the first conjunct *shi ge* (*nanren he nuren*) minimally requires 1 woman and 1 man, and, likewise, the reading of the second conjunct *wu ge* (*nanren he nuren*) minimally requires 1 woman and 1 man. Therefore, any well-formed reading of (26) and (27) minimally requires 2 men and 2 women, thus rendering impossible the reading of a group of 14 men and 1 woman or 1 man and 14 women, a reading allowed in (25).

We shall add another case of our own, which is even more problematic to I&M's non-constituent approach. Consider (28a); its underlying form in (28b) has *yi ge nanren he nuren* 'one man and woman' as the second conjunct, which is an impossible expression.

(28) a. *shi yi ge nanren he nuren*
ten one C man and woman
'eleven men and women'
b. [*shi ge nanren he nuren*] AND [*yi ge nanren he nuren*]

The third piece of evidence comes from the tone sandhi phenomena of the numeral *yi* 'one'. Within the boundary of a word, it stays in its citation tone, T1 (*yī*) if it is the final syllable; otherwise, it changes to T2 (*yí*) when immediately preceding a T4 syllable, and to T4 (*yì*) elsewhere. Note that these tone sandhi rules apply in the examples of [Num C/M N] in (29), where Num is a simple numeral *yi* 'one'.

(29) a.	*yī*	*zhāng*	*zhǐ*	⇒	*yì*	*zhāng*	*zhǐ*
	one	C	paper				
b.	*yī*	*tiáo*	*yú*	⇒	*yì*	*tiáo*	*yú*
	one	C	fish				
c.	*yī*	*běn*	*shū*	⇒	*yì*	*běn*	*shū*
	'one	C	book'				
d.	*yī*	*jià*	*fēijī*	⇒	*yí*	*jià*	*fēijī*
	'one	C	plane'				

Now examine the cases of [Num C/M N] in (30), where Num is an additive complex numeral ending with *yi* ‘one’. Note that all the [*yi* C/M] sequences are identical in (30) and (29) with the same C/Ms, the only difference being that *yi* in (29) is a simple stand-alone numeral and in (30) it is part of an additive numeral.

(30)	a.	*shí yī*	*zhāng*	*zhǐ*	≠>	**shí yì*	*zhāng*	*zhǐ*
		ten one	C	paper’				
	b.	*shí*	*yī*	*tiáo*	*yú*	≠>	**shí*	*yì*	*tiáo*	*yú*
		ten	one	C	fish’					
	c.	*shí*	*yī*	*běn*	*shū*	≠>	**shí*	*yì běn*	*shū*
		ten	one	C	book’				
	d.	*shí*	*yī*	*jià*	*fēijī*	≠>	**shí*	*yí jià*	*fēijī*
		ten	one	C	plane’				

The fact that the tone sandhi of *yī* ⇒ *yì* does not occur in (29) suggests that expressions such as *shí yī zhāng zhǐ* [ten one C paper] cannot be derived from *shí zhāng* (*zhǐ*) *yī zhāng zhǐ*.[5]

Based on the evidence discussed above and also a host of other evidence, He (2015) is able to conclude that numerals are constituents in Chinese, including simple numerals, e.g. *san* ‘three’, multiplicative numerals, e.g. *san-qian* ‘three thousand’, and additive numerals, e.g. *san-qian san-bai* ‘three thousand three hundred’.[6]

We shall now demonstrate that the same mathematics in a multiplicative complex numeral, which is a constituent, applies equally to [Num C/M], which is thus likewise a constituent.

[5] An anonymous reviewer points out that the fact that *yi* remains its citation tone in *shi-yi* may be motivated by the information status. For example, in the expression *shi-yi* (ten-one) ‘eleven’ both *shi* ‘ten’ and *yi* ‘one’ are important for the notion of cardinality, so both receive its citation tone. We agree that information status may indeed be a factor. However, the evidence derived from (29) and (30) does *indirectly* support complex numerals as constituents, even with information status taken into consideration. Here are the steps taken in our argumentation. First, a complex numeral, e.g. *shi-yi*, is either a constituent or not a constituent; there isn’t a third possibility. Second, the syntactic relation between *shi* and *yi* is coordination, thus *shi*-AND-*yi*, where AND is silent. Third, if *shi-yi* is not a constituent, then its underlying structure must be [*shi* ~~C/M~~ AND *yi* C/M]. Fourth, tone sandhi facts in (29) and (30) demonstrate that the non-constituent structure [*shi* ~~C/M~~ AND *yi* C/M] is unworkable, with information status taken into consideration to explain why *yi* remains its citation tone in (30). Fifth, if *shi-yi* is not a non-constituent, it is a constituent.

[6] He *et al.* (2017) also argue that numerals in a host of languages in southern China are constituents.

3.4 Extending the constituency of complex numerals to [Num C/M] in Chinese

Under Her's (2012b) analysis, multiplicative complex numerals and the [Num C/M] expressions share the same internal multiplicative structure, i.e. C/M and base both function as a multiplicand, and the underlying function [Num C/M] is the same as [*n*×*base*] or [*multiplier*× *multiplicand*]. Compare (31a) and (31b).

(31) a. *wu* *shi/bai* *yuan* ([*5*×*10/100*] dollar)
five ten/hundred dollar
'fifty/five hundred dollars'
b. *wu* *ke/da* *jidan* ([*5*×*1/12*] egg)
five C/M-dozen egg
'five eggs/five dozen eggs'

In (31a), the constituency, as argued successfully in He (2015), is [[*n*×*base*] N], or specifically [[*5*×*10/100*] dollar], as shown in (32a). In (31b), the mathematics is exactly the same, [[Num × C/M] N] = [[*5*×*1/12*] egg], as shown in (32b).

(32) a. [[*wu shi/bai*] *yuan*] ([*5*×*10/100*] dollar = *50/500* dollars)
b. [[*wu ke/da*] *jidan*] ([*5*×*1/12*] egg = *5/60* eggs)

(33) *wu* *bai* *da* *jidan*
five hundred M-dozen egg
'five hundred dozen eggs'

(34) [[[*wu bai*] *da*] *jidan*] ([[*5*×*100*] × *12*] egg = *600* eggs)

Example (33) demonstrates a multiplicative complex numeral and a C/M. As shown in (34), [*5*×*100*] functions as a multiplier, and *12*, a multiplicand. A left-branching [[Num C/M] N] nicely captures the constituency of the multiplicative operation [[*multiplier*× *multiplicand*] × *multiplicand*]], while such parallel in mathematics and constituency are lost in a right-branching structure [Num [C/M N]]. Note also that in the history of Chinese C/Ms the [N Num C/M] order, e.g. *bi shi zhi* (pen ten C) 'ten pens', appeared earlier than the now-dominant order [Num C/M N] (e.g. Peyraube 1998) and some argue that it is still used today in certain restricted contexts such as taking inventories (e.g. Tang 1996). The right-branching [C/M N] approach fares even worse with the [N Num C/M] order, as C/M and N are not adjacent and thus additional machinery of movement or ellipsis becomes necessary. We shall now demonstrate that when the entire range of C/M word order typology is considered, the left-branching [Num C/M] approach enjoys even greater advantages.

3.5 Complex numerals and C/M word order typology

As mentioned earlier, the most common pattern of complex numerals is [(*n* × *base*) + *m*, *where m* < *base*] (e.g. Comrie 2006; 2016). Given the multiplicative nature of (*n* × *base*), the reverse order (*base* × *n*) should also be attested, a fact fully verified in the database of Chan (2016). Greenberg (1990[1978]: 292) is the first to observe that the order between Num and C/M and the order between *n* and *base* 'harmonize'. This important generalization has long been overlooked until Her *et al.* (2015) and Her (2017a,b), where the generalization is formalized more accurately as the synchronization between two parameters, stated in (35) and (36).[7]

(35) Synchronization between C/M-parameter & Base-parameter
 a. C/M-final order ⇒ base-final numerals
 b. C/M-initial order ⇒ base-initial numerals

(36) C/M-parameter & Base-parameter
 a. C/M-parameter:
  C/M-final [Num C/M] vs. C/M-initial [C/M Num]
 b. Base-parameter:
  base-final [*n base*] vs. base-initial [*base n*]

The synchronization between C/M and base makes perfect sense under Her's (2012b) multiplicative theory, where C/M and base are both multiplicands and thus naturally follow the same order in relation to the multiplier. Her's multiplicative theory of C/M thus in turn provides an insight to Greenberg's generalization.

[7] An anonymous reviewer questions the universality of the base-C/M synchronization, citing the fact that, in English, numeral bases and C/Ms do not behave the same, e.g. C/Ms can bear the plural marker *-s*, but numeral bases like *hundred* do not, e.g. *three kilos/dozens*, but **three hundreds*. Note that the base-C/M synchronization, by definition, applies to classifier languages only. In non-classifier languages such as English, there are indeed terms of measure, e.g. *kilo* and *box*, that are semantically similar to mensural classifiers, *aka* measure words, in Chinese. Yet, crucially, such terms in English are NOUNS syntactically, but such measure words (Ms) are a subcategory of an independent syntactic category C/M, or NUMERAL CLASSIFIERS. In English, numeral bases, which function as multiplicands, belong to the syntactic category of NUMERALS, and terms of measure belong to the syntactic category of NOUNS. The two naturally do not always behave the same.

The C/M-base synchronization has significant implications on the debate between the left-branching approach and the right-branching approach. Recall that there are only four orders attested out of the six possible orders among Num, C/M, and N, repeated in (37). As Her (2017b) observes, with Num expanded to either [*n base*] or [*base n*], there are twelve possible orders among *n*, *base*, C/M, and N, and still only four are attested in languages, shown in (38).

(37) Six word orders of Num, C/M, N
a. √ [Num C/M N]
b. √ [N Num C/M]
c. √ [C/M Num N]
d. √ [N C/M Num]
e. * [C/M N Num]
f. * [Num N C/M]

(38) Twelve word orders of *n*, *base*, C/M, and N
a. √ [*n base* C/M N]
a' * [*base n* C/M N]
b. √ [N *n base* C/M]
b' * [N *base n* C/M]
c. √ [C/M *base n* N]
c' * [C/M *n base* N]
d. √ [N C/M *base n*]
d' * [N C/M *n base*]
e. * [C/M N *base n*]
e' * [C/M N *n base*]
f. * [*n base* N C/M]
f' * [*base n* N C/M]

Given the three parameters, N-parameter, C/M-parameter, and *base*-parameter, and the C/M-base synchronization, the four attested C/M orders in (38) naturally fall out, as shown in (39), summarized in (40). All the unattested orders are ruled out for violating either the [Num C/M] constituency or the C/M-base synchronization.

(39) Twelve word orders of *n*, *base*, C/M, and N
a. √ [*n base* C/M N] C/M-base synchronization
a' * [*base n* C/M N] *anti-C/M-base synchronization
b. √ [N *n base* C/M] C/M-base synchronization
b' * [N *base n* C/M] *anti-C/M-base synchronization

c. √ [C/M *base n* N] C/M-base synchronization
c’ * [C/M *n base* N] *anti-C/M-base synchronization
d. √ [N C/M *base n*] C/M-base synchronization
d’ * [N C/M *n base*] *anti-C/M-base synchronization
e. * [C/M N *base n*] *[Num C/M] not a constituent
e’ * [C/M N *n base*] *[Num C/M] not a constituent
f. * [*n base* N C/M] *[Num C/M] not a constituent
f’ * [*base n* N C/M] *[Num C/M] not a constituent

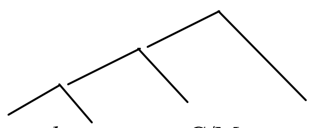
(40) Four attested word orders of *n*, *base*, C/M, and N under [Num C/M]

a. √ [[*n* ***base* C/M**] N] C/M-base synchronization, e.g. Chinese
b. √ [N [*n* ***base* C/M**]] C/M-base synchronization, e.g. Thai
c. √ [[**C/M *base*** *n*] N] C/M-base synchronization, e.g. Ibibio
d. √ [N [**C/M *base*** *n*]] C/M-base synchronization, e.g. Jingpo

The C/M word order typology now finds an insightful motivation in the mathematics within the [Num C/M] constituent, i.e. the chain of multiplicands cannot be interrupted. Consequently, the two orders allowed within the constituent are mirror images, i.e. [[*n* × *base*] × C/M] and [C/M × [*base* × *n*]], or [[*multiplier* × *multiplicand*] × *multiplicand*] and [*multiplicand* × [*multiplicand* × *multiplier*]].

Under the [Num C/M] constituency, C/M-base synchronization is straightforwardly expressed as the alignment of the C/M-parameter and the base-parameter. The functional motivation of this alignment is mathematics. The formal motivation of this alignment is the general setting of head-final of the head-parameter within the Chinese nominal phrase. As shown in (41), *base* is the head within Num [*n base*], C/M is the head of [Num C/M], and N is the head of the whole phrase, all consistently head-final.

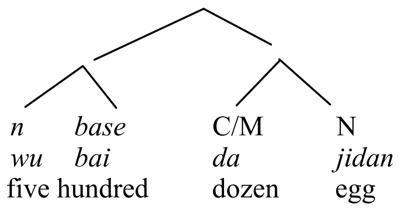
(41)



n *base* C/M N
wu *bai* *da* *jidan*
five hundred dozen egg

Under the right-branching approach, [*n* × *base*] and C/M do not form a constituent; rather, [C/M N] form a constituent, as shown in (42). The immediate problem is how to account for the C/M-base synchronization. As shown in (43), *base* as the head within numeral constituent [*n base*] is head-final, but C/M as the head of [C/M N] is head-initial. The C/M-base synchronization must be achieved by coordinating the conflicting internal orders of two separate constituents, [*n* × *base*] and [C/M N], thus requiring much more powerful context-sensitive operations. The further problem of such an *ad hoc* operation is that it entirely misses the parallel mathematics between [*n base*] and [Num C/M].

(42)



(43) Four attested word orders of *n*, *base*, C/M, and N under [C/M Num]

a.	√	[*n* ***base*** [**C/M** N]]	[C/M N] a constituent
b.	√	[N *n* ***base*** **C/M**]	[C/M N] not a constituent
c.	√	[**C/M** ***base*** *n* N]	[C/M N] not a constituent
d.	√	[[N **C/M**] ***base*** *n*]	[C/M N] a constituent

Furthermore, in order to derive (43b) and (43c), where C/M and N are not adjacent, *move* is required, a more costly operation than *merge*, in the current minimalist framework. An additional costly aspect of such movement-based accounts is how to constrain the required movements so that the unattested orders are not over-generated. Her (2017b) demonstrates that over-generation is precisely a problem with Cinque's (2005) right-branching antisymmetric structure of D, A, Num, C/M, and N. A much more satisfactory solution is found in Abels & Neeleman's (2012) symmetric framework. Though both Cinque (2005) and Abels & Neeleman (2012) adequately account for Greenberg's Universal 20, the latter is simpler and more constrained, and can be straightforwardly extended to account for the C/M word order typology, as long as [Num C/M] remains a constituent (Her 2017b).

To summarize, with the complex numeral and its constituent-hood taken into consideration, the left-branching approach of [Num C/M] stands out as the only workable option for the Chinese C/M phrase.

4. Revisiting A. Li (2014)

A. Li (2014) proposes that the structure of [Num C/M (*de*) N] depends on the interpretation of the phrase. If *de* is present and the [Num C/M *de*] portion functions as a property-denoting modifier of N, similar to a pre-nominal adjective, the entire phrase then has a left-branching structure. In this case, the modification marker *de* is obligatory and base-generated.

Yet, under the quantity-denoting reading of a [Num C/M (*de*) N] phrase, where *de* is not base-generated and is optionally inserted phonologically instead as a phonological phrasing strategy to encode focus on the [Num C/M] modifier, a right-branching structure should be assumed. Thus, contra Zhang (2011; 2013) and Li & Rothstein (2012), A. Li (2014) argues convincingly that all quantity-denoting phrases in Chinese, regardless of C/M subtypes, share the same structure and the optional *de* under this reading is not a factor at all in terms of syntactic structure.

A. Li (2014)'s observations provide important insights into the nature of the classifier phrase, but her right-branching account based on the non-constituent analysis of complex numeral is problematic, due to the fact that complex numerals in Chinese are constituents.

In §4.1, we examine a crucial point raised by A. Li (2014) in relation to complex numerals and demonstrate that the right-branching alternative is not viable. Following the same spirit, in §4.2 we revisit Li's account of the property-denoting use of the [Num C/M] phrase and its implications. Finally, in §4.3 we discuss her phonological account of *de*-insertion and again demonstrate that a left-branching account fares much better.

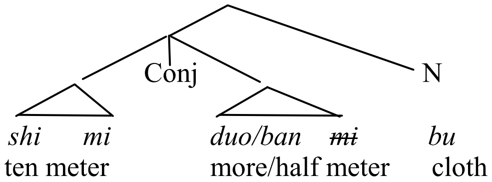
4.1 Evidence from *ban/duo* 'half/more'

One of the arguments for the [Num C/M] constituent is the post-C/M *ban/duo* 'half/more' and fractions (e.g. Hsieh 2008; Zhang 2013; He 2016). Her (2012a) argues specifically that in a left-branching structure, the [Num C/M] constituent merges with *ban/duo* via conjunction, as shown in (44a-b), where *ban/duo* is interpreted in relation to the C/M that immediately precedes it.[8]

[8] As an anonymous reviewer pointed out, this conjunctive relation exists pervasively in additive complex numerals like *san bai wu shi* (three hundred five ten) 'three hundred and fifty', which has the underlying conjunctive structure [[*san bai*] AND [*wu shi*]].

(44) a. *shi* *mi* *duo/ban* *bu*
ten meter more/half cloth
'ten meters of cloth and a bit more/ten and a half meters of cloth'

b.

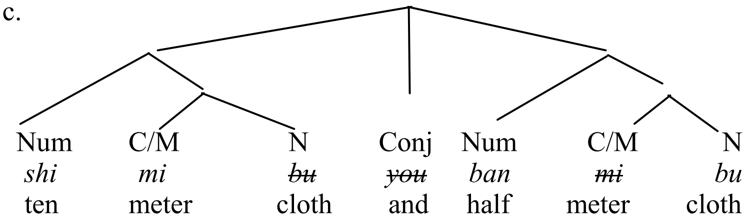


Extending He's (2016) analysis on complex numerals with *duo*, we propose that the two conjuncts in (44b) both contain an instance of M *mi* 'meter', with the second instance PF-deleted. Under the view that complex numerals are constituents, this analysis reflects the mathematics of the conjunctive phrase, i.e. [[Num × C/M] + [*ban/duo* × C/M], and thus enjoys the advantage of being parallel to that of complex numerals such as *san bai wu shi* (three hundred five ten) [[3×100] + [5×10]]. Adopting I&M's non-constituent account of complex numerals, A. Li (2014) contends that a right-branching account is likewise workable; (45a) has (45b) as its underlying form, and (45c) shows its structure and derivation.

(45) a. *shi-mi* *ban* *bu*
ten-meter half cloth
'ten and half meters of cloth'

b. *shi-mi* *bu* *you* *ban-mi* *bu*
ten-meter cloth and half-meter cloth
'ten meters of cloth and half a meter of cloth'

c.



(46) a. *shi mi bu you ban mi bu*
b.**shi mi bu you ban ~~mi~~ bu*
c.**shi mi bu ~~you~~ ban mi bu*
d.**shi mi ~~bu~~ you ban mi bu*
e.**shi mi bu ~~you~~ ban ~~mi~~ bu*
f.**shi mi ~~bu~~ ~~you~~ ban mi bu*

g.**shi mi* ~~*bu*~~ *you ban* ~~*mi*~~ *bu*

The problem with the underlying form proposed in (45b) is that it seriously over-generates, as the three deletable items must either be all deleted, as in (45a), or all show up, as in (46a). All other options are ill-formed. This means that the deletion of three items in unison must be stipulated. Li's account of post-C/M *duo* 'more' likewise requires stipulations.

(47) a. *shi mi duo bu*
ten meter more cloth
b. [*shi-mi* ~~*bu*~~] ~~*(you)*~~ [*duo* ~~*(yidiar)*~~ *bu*]
ten-meter cloth and more slightly cloth
'ten meters of cloth and a bit more (of a meter of cloth)'

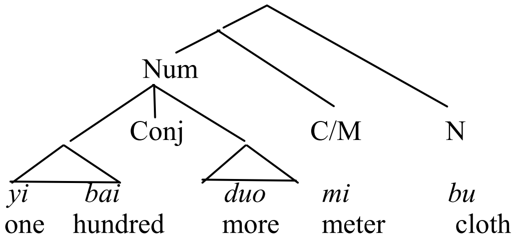
The reason for positing *yidiar* 'a bit' in the source form is because [*duo bu*] 'more cloth' is ill-formed but *duo yidiar bu* is good. There are two ways to view *yidiar* in (47b), either it is a regular lexical item that is deleted during syntactic derivation or it is a base-generated silent YIDIAR (capital letters indicate base-generated silence) in the spirit of Kayne (2006). Either way, it is problematic, because *yidiar* is semantically substantive. Consider the three sentences in (48).

(48) a. *A-Lin bi A-Mei gao*.
A-Lin compare A-Mei tall
'A-Lin is taller than A-Mei.'
b. *A-Lin bi A-Mei gao* YIDIAR
A-Lin compare A-Mei tall slightly
'A-Lin is slightly taller than A-Mei.'
c. *A-Lin bi A-Mei gao* ~~*yidiar*~~.
A-Lin compare A-Mei tall slightly
'A-Lin is slightly taller than A-Mei.'

As demonstrated by Her & Tsai (2014; 2015), a valid source form must be semantically equivalent to the surface form. The surface form in (48a) is clearly not semantically equivalent to either (48b) or (48c); thus neither (48b) or (48c) can serve as the source form of (48a).

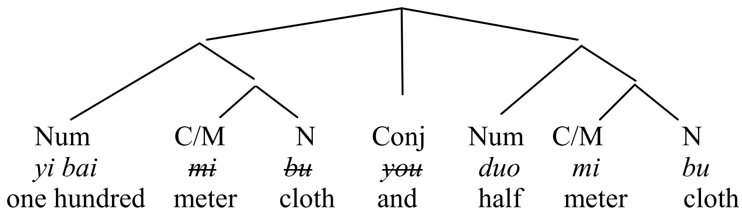
An additional problem of this account of post-C/M *duo* is that it cannot be generalized to post-numeral *duo*, as in (49a). Under the left-branching approach, post-numeral *duo* and post-C/M *duo* behave exactly the same. Compare (49b) with (44b). Recall that both bases and C/Ms function as multiplicands. *Duo* is interpreted in relation to the numeral base *bai* 'hundred' in (49b) and the C/M *mi* 'meter' in (44b). A simple generalization is thus achieved, i.e. *duo* is interpreted in terms of the multiplicand that immediately precedes it.

(49) a. *yi bai duo mi bu*
one hundred more meter cloth
'one hundred plus meters of cloth'
b.



However, under Li's right-branching account, (50a) would be the source form of the surface form of (49a). The problem is that the source form itself is ill-formed; also, as shown in (51a) and (51b), whether overt or covert, *yidiar* is ill-formed following the post-numeral *duo*. It thus must be stipulated that all the deletions proposed in the derivation are all obligatory in this particular context.

(50) a.**yi bai mi bu you duo yidiar mi bu*
b.



(51) a.**duo yidiar mi bu*
b.**duo mi bu*

4.2 Property-denoting [Num C/M] phrases

An expression of the form [Num C/M *de* NP], as in (52), has two distinct readings, which can be disambiguated by contexts.

(52) *san bang de xigua*
three pound DE watermelon

a. 'three-pound watermelons'
b. 'three pounds of watermelons'

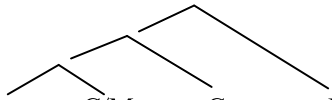
(53) a. *zhe yi ge san bang *(de) xigua zui tian.*
this one C three pound DE watermelon most sweet
'This three-pound watermelon is the sweetest.'
b. *Tamen yigong cai chi-le san bang* (*de*) *xigua.*
they in-total only eat-Asp three pound DE watermelon
'They only ate three pounds of watermelon in total.'

A. Li (2014) establishes very convincingly that the two readings involve two different structures and two different *de*'s. The property reading of [*san bang de*] in (53a) is adjectival and thus similar to [*bu da bu xiao de*] 'not too big and not too small', where *de* is a full-fledged lexical item that is base-generated and undergoes syntactic derivation. It is thus required. Yet, the optional *de* in (53b), which involves a quantity reading, is not base-generated; rather, it is optionally phonologically inserted as a post-syntactic operation to place emphasis on the preceding quantity expression. This crucial distinction nicely explains why the property-denoting lexical *de* cannot be deleted when the head noun is deleted, as in (54a), but the phonologically-inserted *de* cannot appear without a following noun, as in (54b).

(54) a. *Wo yao san bang* *(*de*) *xigua, bushi wu bang* *(*de*).
I want three pound DE watermelon not five pound DE
'I want three-pound watermelons, not five-pound ones.'
b. *Wo cai chi-le san bang* (*de*) *xigua,*
I only eat-ASP three pound DE watermelon
bushi wu bang (**de*).
not five pound DE
'I only ate three pounds of watermelon, not five pounds.'

Under the property-denoting of [Num C/M *de*], crucially, as A. Li (2014) insists, with *de* functioning as a complementizer, [Num C/M] must be seen as a constituent here, e.g. *san bang* 'three pounds' in (55). This fact is more obvious with the property-denoting phrase serving as a predicate in (56). A. Li (2014) assumes that [Num C/M] here has a straightforward left-branching structure with no silent elements. Li & Rothstein (2012) also discuss this property-denoting reading extensively and likewise propose a left-branching structure.

(55)



Num C/M Comp N
san *bang* *de* *xigua*
three pound DE watermelon

(56) *Zhe ge xigua de zhongliang shi san bang.*
this C watermelon DE weight be three pound
'The weight of this watermelon is three pounds.'

Under the right-branching approach, there are two possible views on this property-denoting [Num C/M] constituent. The first option is that each C/M only *optionally* take a nominal complement, and this [Num C/M] string is a complete constituent. Or, one may claim that the property-denoting [Num C/M] phrase in fact has the underlying form [Num [C/M ~~N~~]], where N is silent either by base-generation or by deletion. In all these options under the right-branching approach, Num and C/M never merge directly. See He (2016) for arguments against these options. We shall therefore not engage in a deliberation as to which of the options is less costly under the right-branching approach. We simply point out that they are both costly and the most economical solution is to adopt the left-branching approach, where Num and C/M always form an immediate constituent.

4.3 Phonologically-inserted *de* in the C/M phrase

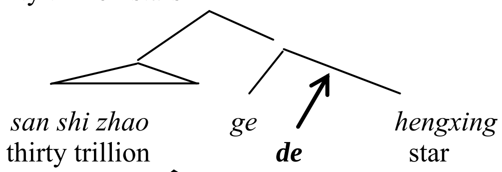
We now focus on the quantity-denoting *de* in a C/M phrase. Consider (57), where the only difference *de* makes is information focus (e.g. Tang 2005; Jiang 2008; Li & Rothstein 2012). A. Li (2014) argues that when the information focus is on Num and C/M in the form of [Num [C/M [NP]]], *de* is phonologically inserted to create proper phonological phrasing reflecting the focus.

(57) *Wo he-le yi da-kou* (*de*) *jiu.*
I drink-ASP one big-mouthful DE wine
'I drank a big mouthful of wine.'

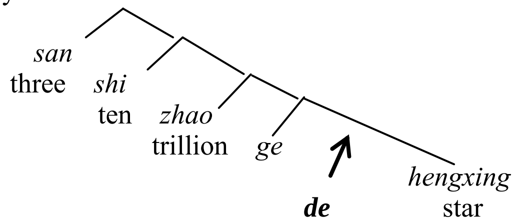
We accept A. Li's phonological account, as it is convincingly argued for; however, we contend that this account in fact makes much better sense in a left-branching structure, i.e. [[Num C/M] *de* N]. Consider (58a) and the right-branching approach first. There are two possible structures, depending on whether numerals are recognized as constituents or not, as shown in (58b) and (58c), respectively. In either structure, the classifier *ge* is a head taking the following NP as its complement, and yet, the insertion of *de* comes right in between the head and its complement and has to place the information focus on the string [*san shi zhao ge*], which is *not* a constituent.

(58) a. *san shi zhao ge de hengxing*
three ten trillion C DE star
'thirty trillion stars'

b.



c.

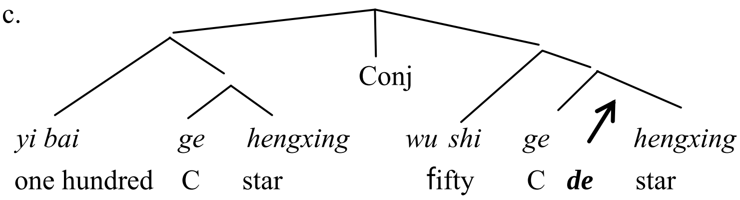


The right-branching account fares even worse with additive complex numerals taken into account. (59a) has (59b) as its source form, which is ill-formed. The insertion of *de* happens in the second conjunct and the information focus falls on the combination of the entire first conjunct and part of the second conjunct.

(59) a. *yi bai wu shi ge de hengxing*
one hundred five ten C DE star
'one hundred and five stars'

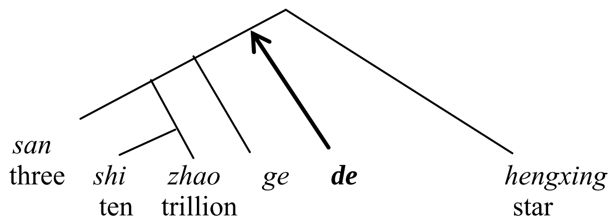
b.**yi bai ge hengxing wu shi ge de hengxing*
one hundred C star five ten C DE star

c.



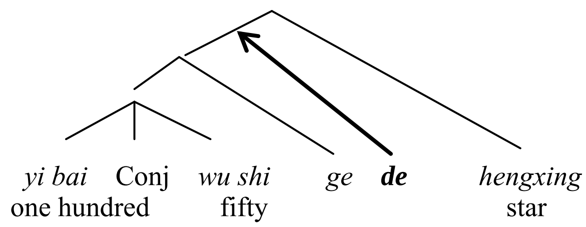
Now consider (58a) and its left-branching structure in (60). The tonal group {Num + Cl + NP} is a constituent, and *de* comes in between the NP and its quantity-denoting phrase; what receives the information focus is thus a coherent constituent reflecting a coherent multiplicative operation [[*3*×*ten*]×*trillion*]×*1*].

(60)



A left-branching account of (59a) is given in (61); again, a complete constituent reflecting a multiplicative operation, i.e. [[[*1*×*hundred*] + [*5*×*ten*]] ×*1*], receives the information focus.

(61)



To summarize, in this section we have thus far examined A. Li's (2014) account of the post-C/M *ban/duo* 'half/more' in relation to complex numerals, the property-denoting use of the [Num C/M] phrase, and the phonologically inserted *de* in the quantity-denoting C/M phrase and demonstrated the grammatical and mathematical insights missed in the right-branching approach but revealed in the left-branching approach.

4.4 The discrepancy between Cs and Ms

As an anonymous reviewer points out, even though tests involving the approximant *duo*, the property-denoting Num-C/M, and the insertion of *de* all favor a left-branching structure over a right structure, one cannot ignore the fact that these tests do not apply to all the subtypes of C/M. Specifically, most of the examples used are restricted to Ms and may not work with Cs.

We agree, but such discrepancy between Cs and Ms is not a problem at all, in our account.

First of all, it has been well etstablished in both camps, e.g. Hsieh (2008) for the left-branching camp and A. Li (2014) for the right-branching camp, that syntactically, Cs and Ms are two subcategories of a single formal category, which we call C/M. Thus, if Ms have a left-branching structure, then Cs must have the same structure, and vice versa. Naturally, Cs and Ms, like any two subcategories within a syntactic category, e.g. count/mass nouns, (in)transitive verbs, stage/individual adjectives, etc., may behave differently in certain respects. Her (2012a), recognizing the C/M category, offers an explanation for such differences based on the fact that Cs function as a multiplicand *1*, an identity item in multiplication, and is thus semantically transparent, as argued for by Her & Hsieh (2010).

Specific to the test with *duo* 'more', even though (62a) is bad, (62b) is perfectly good with *ban* 'half'. In fact, (62a) *can* also be good in certain contexts. For example, one may ask the cook about the ingredients of a certain dish and the cook may reply that one ingredient involves eggs and in this particular dish he used *liang ke duo jidan* (62a) or *liang ke ban jidan* (62b).[9]

(62) a. **liang* *ke* *duo* *jidan*
two C more egg
'two eggs and a bit more'
b. *liang* *ke* *ban* *jidan*
two C half egg
'two and half eggs'

Specific to the insertion of *de*, again (63a) is indeed relatively far worse than (63b), and (63c) is even better. This issue and previous accounts have been thoroughly reviewed and discussed in Her & Hsieh (2010:538-541), where they propose that the computational complexity of Num and the semantic overlap between C and N are two factors that affect the acceptability of [Num C-*de* N]. Briefly, given the fact that C and N overlap in semantic content, the C-N sequence is much more resistant of -*de* intervention, and the more computationally complex the Num is, the more easily -*de* can intervene between [Num C] and [N].[10]

---

[9] In Taiwan Mandarin the most common C for eggs is *ke*, while in Mainland China's Putonghua, it is the general C *ge*.

[10] An anonymous reviewer thus finds (59) less natural than (58). Note that (59) involves 150, while (58) involves 30,000,000,000,000. (58) is computationally far more complex than (59)

(63) a. **liang* *ke* *de* *jidan*
two C DE egg
'two eggs'

b. ?*liang* *bai* *ba* *shi liu* *ke* *de* *jidan*
two hundred eight ten six C DE egg
'two hundred and eighty-six eggs'

c. *ba* *fen* *zhi* *yi* *ke* *de* *jidan*
eight share POSS one C DE egg
'one eighth of an egg'

In short, we can safely conclude that the discrepancy between Cs and Ms is real, but it is due to the fact they are two subcategories of a single syntactic category, and more importantly, such discrepancy can be easily accounted for under a unified left-branching structure.

As indicated earlier, the fact that the unified left-branching approach holds logically falsifies the 'split', or non-unitary, approach. For instance, Li & Rothstein (2012) contend that the counting reading of C/M has a right-branching structure, while the measure reading has a left-branching structure. Similarly, the split analysis provided by Zhang (2011; 2013) is not viable, where she assumes a right-branching analysis for Cs like *tiao.* Given the simple fact that complex numerals are constituents in Chinese and that only a left-branching approach is workable with numerals as constituents, a right-branching structure must be problematic under either a unified right-branching approach or a 'split' approach.

Conclusion

Whether the classifier phrase [Num C/M N] in Chinese has a right-branching or left-branching structure is a fundamental issue that has significant consequences beyond the nominal phrase. The issue is not only critical for Chinese but also has important implications for other classifier languages in the world.

Following A. Li (2014), this paper examines the crucial role of complex numerals in this debate and introduces He's (2015) rigorous arguments for treating complex numerals as constituents, contra Ionin & Matushansky (2006), whose non-constituent analysis provides the critical basis for A. Li's (2014) right-branching account. However, the non-constituent analysis is not applicable to Chinese. If the non-constituent analysis of complex numerals is assumed, extra stipulations are required. For instance, the silent element YIDIAR 'a bit' is needed, which is nonetheless illicit by definition, as discussed in §4. Note that, logically, the rejection of the right-branching structure logically entails the rejection of the split approach, where both left- and right-branching structures are required.

We also demonstrate that with complex numerals analyzed properly as constituents, the [Num C/M] constituency and its internal multiplicative function, i.e. [Num×C/M], are entirely consistent with those of multiplicative complex numerals, i.e. [*n*×*base*]. The paper offers evidence from Chinese as well as cross-linguistic word order typology.

We therefore conclude that the unified left-branching approach to the classifier phrase [Num C/M N] not only fares far better in Chinese than the right-branching approach, but may also be universally applicable to other classifier languages.

Acknowledgements

We are very grateful to the anonymous reviewers of *Language and Linguistics* for their constructive comments, which led to significant improvements of the paper. We have benefited greatly from the extensive discussions with the following students in seminars on Chinese syntax and classifiers offered by the first author: Chia-Chi Chen, Yu-Ting Hsu, Ping-Ho Huang, Tsung-Chia Huang, Meng-Tsiong Lee, Kun-Han Lin, Marc Tang, Wen-Chi Yang, Yu-min Huang, and Chu-Hsien Yeh. However, all remaining errors are our own. The two authors contributed equally, while O.-S. Her is the corresponding author. This study was funded by the following grants by Taiwan's Ministry of Science and Technology (MOST):102-2811-H-004-023, 103-2811-H-004-003, 103-2633-H-004-001, 104-2410-H-004-164-MY3, and 104-2633-H-004-001.

References

Abels, Klaus and Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15(1). 25–74.

Aikhenvald, Alexandra Y. 2000. *Classifiers: A typology of noun categorization devices*. (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.

Au Yeung, Wai Hoo Ben. 2005. *An interface program for parameterization of classifiers in Chinese*. Hong Kong: Hong Kong University of Science and Technology. (Doctoral dissertation.)

Au Yeung, Wai Hoo Ben. 2007. Multiplication basis of emergence of classifiers. *Language and Linguistics* 8(4). 835–861.

Bhattacharya, Tanmoy. 1999. *The structure of the Bangla DP.* London: University College London. (Doctoral dissertation.)

Bhattacharya, Tanmoy. 2001. Numeral/quantifier-classifier as a complex head. In van Riemsdijk, Henk & Corver, Norbert (eds.), *Semi-lexical heads*. 191–221. Berlin: Mouton de Gruyter.

Borer, Hagit. 2005. *Structuring sense*, *vol. 1: In name only*. Oxford: Oxford University Press.

Chan, Eugene S. L. 2016. *Numeral systems of the world's languages*. (https://mpi-lingweb.shh.mpg.de/numeral/) (Accessed 2018-07-17).

Cheng, Lisa L-S. & Sybesma, Rint. 1998. *Yi-wan tang, yi-ge tang*: Classifiers and massifiers. *Tsing Hua Journal of Chinese Studies* (New Series) 28(3). 385–412.

Cheng, Lisa L-S. & Sybesma, Rint. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4). 509–542.

Cinque, Guglielmo. 2005. Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36(3). 315–332.

Comrie, Bernard. 2006. Numbers, language, and culture. (Paper presented at the 16th Jyväskylä Summer School, Jyväskylä, 24 July-11 August 2006.)

Comrie, Bernard. 2016. Numeral bases. In Dryer, Mathew S. & Haspelmath, Martin (eds.), *The world atlas of language structures online*, chapter 131. (https://wals.info/chapter/131) (Accessed 2018-07-17).

Corver, Nobert & Zwarts, Joost. 2006. Prepositional numerals. *Lingua* 116(6). 811–835.

Croft, William. 1994. Semantic universals in classifier systems. *Word* 45(2). 145–171.

Fukui, Naoki & Yuji Takano. 2000. Nominal structure: An extension of the symmetry principle. In Svenonius, P. (ed.), *The derivation of VO and OV*, 219–254. Amsterdam: John Benjamins.

Greenberg, Joseph H. 1990b[1972]. Numeral classifiers and substantival number: Problems in the genesis of a linguistic type. In Denning, K. & Kemmer, Suzanne (eds.), *On language: Selected writings of Joseph H. Greenberg*, 166–193. Stanford: Stanford University Press. (First published 1972 in *Working Papers on Language Universals* 9. 1–39.)

Greenberg, Joseph H. 1990a [1978]. Generalizations about numeral systems. In Denning, K. & Kemmer, Suzanne (eds.), *On language: Selected writings of Joseph H. Greenberg*, 271–309. Stanford: Stanford University Press. (First published 1978 in *Universals of human language*, *vol. 3: Word Structure*, 249–295. Stanford: Stanford University Press.)

Grimshaw, Jane. 2000. Locality and extended projection. In Coopmans, Peter & Everaert, Martin B. H. & Grimshaw, Jane (eds.), *Lexical specification and insertion*, 115–133. Amsterdam: John Benjamins.

He, Chuansheng. 2015. Complex numerals in Mandarin Chinese are constituents. *Lingua* 164. 189–214.

He, Chuansheng. 2016. Lun hanyu shu-liang zuhe de chengfen wanzhengxing. [On the constituenthood of the numeral-classifier combinations in Chinese]. *Dangdai Yuyanxue* [Contemporary Linguistics] 18(1). 1–18.

He, Chuansheng & Her, One-Soon & Hu, Xiaoshi & Zhu, Weijing. 2017. Overt coordination in additive numerals of minority languages in South China. *Syntax* 20(3). 292–316.

Her, One-Soon. 2012a. Distinguishing classifiers and measure words: A mathematical perspective and implications. *Lingua* 122(14). 1668–1691.

Her, One-Soon. 2012b. Structure of classifiers and measure words: A lexical functional account. *Language and Linguistics* 13(6). 1211–1251.

Her, One-Soon. 2017a. Deriving classifier word order typology, or Greenberg's Universal 20A, and universal 20. *Linguistics* 55(2). 265–303.

Her, One-Soon. 2017b. Structure of numerals and numeral classifiers in Chinese: Historical and typological perspectives and cross-linguistic implications. *Language and Linguistics* 18(1). 26–71.

Her, One-Soon & Hsieh, Chen-Tien. 2010. On the semantic distinction between classifiers and measure words in Chinese. *Language and Linguistics* 11(3). 527–551.

Her, One-Soon & Tsai, Hui-Chin. 2014. *Color* isn't silent, *shallow* isn't deep: Two case studies of evaluating silent elements. *Language and Linguistics* 15(6). 775–800.

Her, One-Soon & Tsai, Hui-Chin. 2015. On silent elements: A case study of *grand* and its silent entourage. *Natural Language and Linguistic Theory* 33(2). 575–605.

Her, One-Soon & Chen, Jing-Perng & Tsai, Hui-Chin. 2015. Justifying silent elements in syntax: The case of a silent numeral, a silent classifier, and two silent nouns in Mandarin Chinese. *International Journal of Chinese Linguistics* 2(2). 193–226.

Her, One-Soon & Tsai, Hui-Chin & Lin, Kun-Han & Tang, Marc & Lee, Meng-Chang. 2016. Numeral bases and numeral classifiers in SMATTI: Word order variation in Tibeto-Burman. (Paper presented at the conference of New Ways of Analyzing Variation–Asia Pacific 4 (NWAV AP4), Chiayi, 22-24 April 2016.)

Hsieh, Miao-Ling. 2008. In Her, One-Soon & Chui, Kawai (eds.), *The internal structure of noun phrases in Chinese*. (Taiwan Journal of Linguistics: Book Series in Chinese Linguistics. No. 2.) Taipei: Crane Publishing.

Huang, C.-T. James. 1982/1998. *Logical relations in Chinese and the theory of grammar*. Cambridge: MIT (Doctoral dissertation.) (Published in 1988 by Garland Publishing.)

Huang, C.-T. James & Li, Yen-Hui Audrey & Li, Yafei. 2009. *The syntax of Chinese*. Cambridge: Cambridge University Press.

Huang, C.-T. James & Ochi, Masao. 2011. Classifiers and nominal structure: A parametric approach and its consequences. In Gao, Ming-le (ed.), *Universals and variation* (Proceedings of GLOW in Asia VIII, 2010), 191–197. Beijing: Beijing Language and Culture University Press.

Hurford, James. 1975. *The linguistic theory of numerals*. Cambridge: Cambridge University Press.

Hurford, James. 1987. *Language and number: The emergence of a cognitive system*. Oxford: Blackwell.

Hurford, James. 2001. Numeral systems. In Smelser, Neil. J. & Baltes, Paul. B. (eds.), *International encyclopedia of the social and behavioral sciences*, 10756–10761. Amsterdam: Elsevier.

Ionin, Tania & Matushansky, Ora. 2006. The composition of complex cardinals. *Journal of Semantics* 23(4). 315–360.

Jackendoff, Ray 1977. *X-Bar syntax: A study of phrase structure*. Cambridge: The MIT Press.

Jenks, Peter. 2010. Evidence for the syntactic diversity of numeral classifiers. Paper presented at the 84th Annual Meeting of the LSA, Baltimore, 7-10 January 2010.

Jiang, L. Julie. 2008. Monotonicity and measure phrases in Chinese. (Paper presented at the 11th International Symposium on Chinese Languages and Linguistics (IsCLL-11), Hsinchu, 23-25 May, 2008.)

Kayne, Richard S. 2005. On parameters and on principles of pronunciation. In Hans Broekhuis & Corver, Norbert & Huybregts, Riny & Kleinhenz, Ursula & Koster, Jan (eds.), *Organizing grammar: linguistic studies in honor of Henk van Riemsdijk*, 289–299. Berlin: Mouton de Gruyter.

Kayne, Richard S. 2010. A note on the syntax of numerical bases. *Comparisons and Contrasts*. Oxford University Press. 57–72.

Li, Charles & Thompson, Sandra. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.

Li, Xu-Ping. 2011. *On the semantics of classifiers in Chinese*. Israel: Bar-Ilan University. (Doctoral dissertation.)

Li, Xu-Ping & Rothstein, Susan. 2012. Measure readings of Mandarin classifier phrases and the particle de. *Language and Linguistics* 13(4). 693–741.

Li, Yen-Hui Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics* 8(1). 75–99.

Li, Yen-Hui Audrey. 2014. Structure of noun phrases: Left or right? *Taiwan Journal of Linguistics* 12(2). 1–32.
Lin, Jo-Wang. 1997. Noun phrase structure in Mandarin Chinese: DP or NP? *Chinese Languages and Linguistics* 3. 401–434.
Liu, Yi-Hsien. 2013. *A unified syntactic account of Mandarin subject nominals*. Los Angles: University of Southern California. (Doctoral Dissertation.)
Muromatsu, Keiko. 1998. *On the syntax of classifiers*. Maryland: University of Maryland. (Doctoral dissertation.)
Peyraube, Alain. 1998. On the history of classifiers in archaic and medieval Chinese. In T'sou, Benjamin K. (ed.), *Studia linguistica serica*, 39–68. Hong Kong: Language Information Sciences Research Centre, City University of Hong Kong.
Selkirk, Elisabeth. 1977. Some remarks on noun phrase structure. In Culicover, P. W. & Wasow, T. & Akmajian, A. (eds.), *Formal syntax*, 285–316. New York: Academic Press.
Simpson, Andrew. 2005. Classifiers and DP structure in Southeast Asia. In Cinque, Guglielmo & Kayne, Richard (eds.), *The oxford handbook of comparative syntax*, 806–838. Oxford: Oxford University Press.
Tang, Chih-Chen Jane. 1996. *Ta maile bi shizhi* and Chinese phrase structure. *Bulletin of the Institute of History and Philology* 67(3). 445–502.
Tang, Chih-Chen Jane. 2005. Nouns or classifiers: A non-movement analysis of classifiers in Chinese. *Language and Linguistics* 6(3). 431–472.
Watanabe, Akira. 2006. Functional projections of nominals in Japanese: Syntax of classifiers. *Natural Language and Linguistic Theory* 24(1). 241–306.
Yang, Rong. 2001. *Common nouns, classifiers, and quantification in Chinese*. New Jersey: The State University of New Jersey. (Doctoral dissertation.)
Zhang, Niina Ning. 2011. The constituency of classifier constructions in Mandarin Chinese. *Taiwan Journal of Linguistics* 9(1). 1–50.
Zhang, Niina Ning. 2013. *Classifier structures in Mandarin Chinese*. Berlin: Mouton de Gruyter.

Authors' Addresses

One-Soon Her (corresponding author)
Graduate Institute of Linguistics &
Research Center for Mind, Brain, and Learning
National Chengchi University
64, Sec. 2, ZhiNan Road

Taipei 116, Taiwan
onesoon@gmail.com

Publication History

Date received: 15 March 2017
Date accepted: 3 August 2017