Overt coordination in additive numerals of minority languages in South China

Chuansheng He[a,b], One-Soon Her[c], Xiaoshi Hu[d], Weijing Zhu[a]

[a]Department of English & Institute of Language and Cognition, College of Foreign Languages, Hunan University, Changsha City, Hunan Province, China
[b]Corresponding author
[c]National Chengchi University, Graduate Institute of Linguistics & Research Center for Mind, Brain, and Learning, Taipei, Taiwan
[d]Laboratoire de Linguistique Formelle, Université Paris 7, 30 rue Pascal, Villejuif, 94800, France

Abstract: This paper explores a typology of overt coordination in additive numerals in a number of minority languages in South China. Among about 100 minority languages surveyed, 26 feature different coordinators for additive numerals and for noun phrases, and we demonstrate that these two types of coordinators are neither semantically, etymologically, or morphologically related. This phenomenon presents a serious challenge to the proposal that additive numerals do not form constituents and that numeral coordination is derived from underlying nominal coordination in such languages.

1. Introduction

Since Hurford's (1975, 1987) pioneering studies of the linguistic theory of numerals, the standard syntactic analysis assumed that complex numerals like *five hundred* (multiplicative) and *fifty two* (additive) were phrasal constituents (Corver and Zwarts 2006; among others), until it was challenged by Ionin and Matushansky (2006) (IM hereafter), who proposed a non-constituency analysis for complex numerals. IM assmued that numerals like *two*, *hundred*, and *thousand* are predicate modifiers of type $<<et><et>>$ selecting lexical nouns or other base-noun combinations as complements. A multiplicative numeral expression like *one hundred languages* thus projects a

complementative structure as bracketed in [one [hundred [languages]]]. Consequently, an additive numeral expression like *one hundred and two languages* is derived from a full NP coordination where the head noun is either right-node-raised or PF-deleted, as illustrated in [[one hundred t_i] and [two t_i] languages$_i$] or [[one hundred ~~languages~~] and [two languages]]. IM argued that the complementative analysis for numerals receives empirical support from the Case marking data from Russian.

Though highly influential, this proposal has met some challenges in both its semantics and syntax. IM's semantic assumption is challenged in Rothstein (2013) and He (2015a)[1]. Their syntactic analysis (a consequence of semantics) is challenged in Kayne (2010) and more recently in Meinunger (2015) and He (2015b). Meinunger (2015) presented several syntactic, semantic, and pragmatic arguments against the non-constituency analysis, and proposed a graft (constituent) structure for complex numerals, which can deal with the Russian Case data equally well. Based on Mandarin data, He (2015b) put forward a series of syntactic, semantic, and morpho-phonological arguments against the non-constituency analysis and proposed a more traditional Hurfordian structure for complex numerals. He also investigated a small number of minority languages in South China and found that these minority languages feature morpho-phonological processes similar to those found in Mandarin Chinese, and his findings also support a constituency analysis for complex numerals in these languages.

---

[1] The <<et><et>> semantics requires that any cardinal must first be satisfied with a lexical NP argument and that the lexical NP should be a singular one denoting a set of atoms though superficially it may appear in the plural form. IM (2006: 329) attribute this singularity requirement to a pragmatic constraint: "only individuals of the same (known) cardinality can be counted." In a recent survey paper on the syntax-semantics interface of numerals, He (2015a) points out several problems with IM's semantics. One concerns the pragmatic constraint. It seems wrong to assume that counting can only be possible when the objects to be counted must have the same cardinality. This can be shown in the following phrase (i), which directly tells us that the individuals have different cardinalities. If the constraint really works, (i) should be an inappropriate form. And in (ii), we can clearly count the sets as two even though they have different cardinalities.
    (i) two sets of atoms of different numbers
    (ii) A is the set of all natural numbers and B is the set of all real numbers. These two sets have different cardinalities, according to Cantor.

Considering that many languages use overt coordinators to link numerals, e.g., English *and* and French *et*, and many languages, e.g., Chinese (see section 3.1), use different coordinators to link different syntactic categories, we speculate that there may be languages that use different coordinators for linking numerals and for linking nouns. If we can prove that such different coordinators are indeed different morphemes, and thus not allomorphs of the same morpheme, we will have strong reason to believe that numeral phrases cannot be derived from nominal phrases.

With this logic in mind and inspired by He's initial investigation of minority languages in South China, we conducted an extensive survey of more than 100 minority languages in the region, firstly based on documented resources[2], covering five language families or groups: Tibeto-Burman, Austroasiatic, Austronesian, Tai-Kadai, and Miao-Yao. The results of this extensive literature survey show that 26 languages feature different coordinators for numerals and for nominals, which stand in strict complementary distribution and thus cannot be used interchangeably. Then we did field work. Out of these 26 languages, we have consulted with native speakers and experts in 11 languages and obtained firsthand verification of *core data*, which are critical to our argumentation and are not provided in the documented resources. Among the 11 verified languages, a thorough investigation of Tibetan was made thanks to the large number of Tibetan students in Hunan University.[3] Our purpose is to draw attention to a set of languages spoken in South China which IM's structural analysis cannot account for and thus their proposal cannot be universally correct.

This paper is organized as follows. Section 2 reports the data of different coordinators for

---

[2] Including 57 books in the *Series of Grammar Sketches of Minority Languages in China*; 50 books in the *Series of Newly Discovered Languages in China*; *Zhongguo de Yuyan* 'The Languages of China', an encyclopedic reference grammar book of 129 minority languages in China; and other relevant literature.

[3] We did not obtain verification for the other 15 due to practical difficulties, e.g., the small size of the populations and the remoteness of the areas where these languages are spoken. Nonetheless, we believe the verified observations are sufficiently sound and general for our purpose. While the verification of the other 15 languages would add more data to our observation, it would not strengthen the force of our argument in any qualitative aspect, and, by the same token, a failure to verify them would not significantly weaken the force of our argument.

nominals and for numerals in 26 languages. Based on this data, section 3 argues against the proposal that numeral coordination is derived from underlying nominal coordination, and refutes an alternative solution that treats the numeral coordinators and the nominal coordinators as allomorphs of a single morpheme. Section 4 rejects another alternative solution (for some of these languages) that treats numeral coordinators as numerals similar to the Chinese *ling* 'zero'. Section 5 provides a sketch of the syntax and semantics of numeral coordination and outlines a typology of numeral coordination. Section 6 is the conclusion.

2. The data

2.1 Tibeto-Burman languages

In Tibetan (mainly spoken in Tibet, Qinghai and Sichuan with a population of five million, figure estimated by the end of 2000, see Sun et al 2007, the same below for other languages), noun phrases are conjoined by *ta* or *daŋ* (used in different Tibetan dialects but interchangeable among these dialects, *daŋ* is more frequently used in written Tibetan).

(1)   meto^ʔ mapo **ta/daŋ** loma tɕaŋkhu

      flower red   **CONJ** leaf   green

      'red flowers and green leaves'

The picture for numerals is different. Additive numerals between 10 and 20 are formed by juxtaposition of teens and digits (coordinators are not allowed). Usually there are phonological processes taking place between *ten* and the digits.

(2)   a. mi   tɕu^ʔ tɕi^ʔ b. mi   tɕø: ŋa   c. mi   tɕu:r ku

        person ten one   person ten   five   person ten   nine

4

'11 people'   '15 people'   '19 people'

Numerals above 20 and under 100 employ different coordinators: *tsak* for 20, *so* for 30, *ɕe* for 40, *ŋa* for 50, *re* for 60, *tỹ* for 70, *ca* for 80, and *go* for 90 (Zhou 1998). The following data is based on Lhasa Dialect.

(3) a. mi   ɲi   tɕu   **tsak** tɕiʔ   b. mi   sum   tɕu   **so**   ŋa

      person two ten **CONJ** one   person three ten **CONJ** five

      '21 people'   '35 people'

   c. mi   ɕip tɕu   **ɕe**   tɕiʔ   d. mi   ŋa   tɕu   **ŋa**   ɲi

      person four ten **CONJ** one   person five ten   **CONJ** two

      '41 people'   '52 people'

   e. mi   tʂʼuk tɕu   **re**   tɕiʔ   f. mi   tỹ   tɕu   **tỹ**   tɕiʔ

      person six   ten **CONJ** one   person seven ten   **CONJ** one

      '61 people'   '71 people'

   g. mi   cɛʔ   tɕu   **ca**   ku   h. mi   ku   tɕu   **go**   ku

      person eight ten **CONJ** nine   person nine ten   **CONJ** nine

      '89 people'   '99 people'

The above pattern is the same across all Tibetan dialects, including U-Tsang (Lhasa), Amdo, and Khams. It is easily observable that the numeral coordinators are morphologically related to the multiplier numerals, except for 20 and 60. According to Zhou (1998: 55-56), these numeral coordinators are derived from the multiplier numerals by certain phonological harmony rules (mostly having to do with the same consonants, with ablaut in the vowels). The numeral coordinator

tsak (*rtsa* in written Tibetan) is etymologically derived from *brtsegs* 'add', which was used to conjoin numerals in ancient Tibetan (before the eighth and ninth century). The numeral coordinators *so*, *ɕe*, *ŋa*, *re*, *tỹ*, *ca*, and *go* are later developments.

Interestingly, the numeral coordinators in (3) can occur without a preceding numeral as shown in (4) (mainly occurring in spoken Tibetan, less frequent in written Tibetan), fully indicating that these numeral coordinators are numerals in origin. The relevant morphemes are still coordinators, not numerals. Except for *ŋa* and *tỹ* which happen to be of the same phonetic forms as *ŋa* 'five' and *tỹ* 'seven', all of the others are phonetically different from, though clearly related to, their relevant numerals, e.g., *so*/*sum*, *ɕip*/*ɕe*, *tʂ'uk*/*re*, *cɛʔ*/*ca*, *ku*/*go* (see Zhou 1998: 56-57 for an explanation of *tʂ'uk*/*re*). Thus, the preceding teens, when omitted, can be easily deduced from the relevant numeral coordinators.

(4)   a. mi   **tsak**   tɕiʔ   b. mi   **so**   ŋa

        person **CONJ**   one   person **CONJ**   five

        '21 people'   '35 people'

For numerals above 100, however, the coordinator conjoining hundreds and teens is the nominal coordinator *ta* or *daŋ*. So, in a numeral expression over 100, there may be two different coordinators, as shown in (5). Baima, a close relative of Tibetan, is similar in this regard (see Table 1 below).

(5)   mi   ca   **daŋ**   ɳi   ɕu   **tsak**   cɛʔ

       person hundred **CONJ**   two   ten   **CONJ**   eight

       '128 people'

When there is a missing power between two numerals, Tibetan employs a special method by inserting *tɕu mɛ^ʔ* ('ten no')[4] between the two numerals, indicating that the teens are missing, as in (6a). Where there are two missing powers, there are two occurrences of the *mɛ^ʔ*-structure, indicating that both the hundreds and the teens are missing, as in (6b).

(6) a. mi   sum   ca   **tɕu**   **mɛ^ʔ**   tʂ'uk

person   three hundred   **ten   not-have**   six

'306 people'

b. mi   sum   toŋ   **ca**   **mɛ^ʔ**   **tɕu**   **mɛ^ʔ**   tʂ'uk

person   three thousand   **hundred not-have   ten   not-have**   six

'3006 people'

In Tibetan, *tsak*, *so*, *ɕe*, *ŋa*, *re*, *tỹ*, *ca*, and *ɟo* are exclusively used for conjoining numerals and do not seem to have any other functions. They can be called 'exclusive numeral coordinators'. Among the 45 Tibeto-Burman languages surveyed, in addition to Tibetan we found nine other languages that feature different coordinators for numerals and for nominals, as shown in Table 1. All of the languages feature an exclusive numeral coordinator.

Table 1: Nominal coordinators and numeral coordinators in nine other Tibeto-Burman languages

Languages	Nominal coordination	Numeral coordination	Sources
Achang	sǝh zu^ʔ **lɔ^ʔ** sum zu^ʔ two Cl **CONJ** three Cl	tɑ pak **mɔ** tɑ sau one hundred **CONJ** one twenty	Dai and Cui (1985: 36, 62)

---

[4] *Mɛ^ʔ* is the negated existential verb in Tibetan, literally meaning 'not-have' (the existential verb is *jø* 'have'), for example:

(i) Q. Cherã la ŋy   jø   pɛ?   A: Mɛ^ʔ.

you   money have   PRT   not-have

'Do you have money?'   'No.'

	'two people and three people'	'120'	
Anuŋ (Anong)	ɑhiŋ **sɿ** ɑsuŋ bowl **CONJ** chopstick 'bowls and chopsticks'	phã ɕɑ **i** kuŋ tshai ɑŋi five hundred **CONJ** six ten-**CONJ** two '562'	Sun and Liu (2005: 70-71, 117)
Baima	kɑmɑ **re** dzɑʃe star **CONJ** moon 'stars and the moon'	ŋɑ dʑɑ **re** ŋi ʃo **tsɑ** ŋi five hundred **CONJ** two ten **CONJ** two '522'	Sun et al (2007: 63, 124)
Jingpo	ʃăta **theʔ** ʃăkan moon **CONJ** star 'moon and stars'	khjiŋ măŋa **e** tʃăkhu tsa thousand five **CONJ** nine hundred '5900'	Liu (1984: 48, 60)
Nusu	liɑbɹasu **le** lɑmoṃsu farmers **CONJ** workers 'farmers and workers'	thi ɕhɑ **i** ŋɑ[5] one hundred **CONJ** five '105'	Sun and Liu (1986: 38, 94)
Pɣnru (Bengru)[6]	anai apu gəɣai **rəu** mother father child **CONJ** 'mother, father, and children'	lɣŋ **gałie** rɣŋ waŋ **gałie** akhɣ hundred **CONJ** ten PRT **CONJ** one '111'	Sun et al (2007: 712, 719)
Suloŋ (Sulong)	ahai **da** adʑiaŋ adults **CONJ** children 'adults and children'	sua **na** çun ten **CONJ** one '11'	Li (2004: 64, 67)
Xiandao	tʂu **nɣk** xo chopsticks **CONJ** bowls 'chopsticks and bowls'	ta pak **mua** sum tshi **mua** ta one hundred **CONJ** three ten **CONJ** one '131'	Dai et al (2005: 38, 82)
Zaiwa	tsaŋ lă khɔʔ **əʔ** tshun lă khɔʔ rice one bowl **CONJ** meat one bowl 'one bowl of rice and one bowl of meat'	sum ʃo **kɔm** ʃit three hundred **CONJ** eight '308'	Zhu (2011: 62, 227)

---

[5] In Nusu, the numeral coordinator *i* is used to link missing powers but does not appear between adjacent powers. The same is true for Zaiwa (*kɔm*) (see section 4 for more discussions).

   (i) a. vɹi ɕha sɔ tshe ŋa   b. lă tshə lă
       four hundred three ten five   one ten one
       Nusu: '435'   Zaiwa: '11'

[6] The numeral coordinator *gałie* in Pɣnru is a verb meaning 'exceed'. It is noteworthy that in Pɣnru the nominal coordinator *rəu* is adjoined to the right of the last conjunct, but the numeral coordinator *gałie* occurs between numerals.

2.2 Tai-Kadai languages and Miao-Yao languages

Among the Tai-Kadai languages, e.g., Dai, Zhuang, Kam, Buyi, Shui, Maonan, Mo, Lakkja, Mulao, T'en, etc., only Dai (spoken in Yunnan with a population of about one million) has different coordinators for nominals and for numerals. Among the Miao-Yao languages, e.g., Miao, Yao, Mjen, Bunu, She, etc., only Miao (mainly spoken in Hunan and Guizhou with a population of about eight million) has different coordinators for nominals and for numerals.

Table 2: Nominal coordinators and numeral coordinators in Dai and Miao

Languages	Nominal coordination	Numeral coordination	Sources
Dai	phak **lɛ** man vegetable **CONJ** oil 'vegetables and cooking oil'	a. hok pan **pa:i** sa:u b. sip et six thousand **CONJ** twenty ten one '6020' '11'	Yu and Luo (1980: 41, 66)
Miao[7]	pji qwa **kɔ** pji ʐɑ fruit peach **CONJ** fruit pear 'peaches and pears'	a. ɯ tshɛ **qɑ** zi b. ɑ ku pʑei two thousand **CONJ** eight one ten four '2008' '14'	Xiang (1999: 49, 53, 82)

In both languages, the numeral coordinators are used to conjoin non-adjacent powers only, not adjacent powers. It is noteworthy that the numeral coordinator *pa:i* in Dai means 'more' and can appear in other environments as well, as shown in (7) (see section 3.2 and section 4 for the importance of this fact in relation to our argumentation).

(7) a. sip **pa:i**   b. mi tset sip kun **pa:i.**

   ten **more**   have seven ten person **more**

   'a little more than ten'   'There are a little more than 70 people.'

2.3 Austroasiatic languages and Austronesian languages

---

[7] Miao, with its three major dialects, is an important language in the region. The data here is based on the Hunan West dialect (the Xiangxi dialect) spoken in Huayuan County (Jiwei Township) and Baojing County in the west of Hunan Province.

Many of the Austroasiatic languages spoken in Yunnan province of southwest China are heavily influenced by Dai, a Tai-Kadai language which is more dominant in the region, and use the borrowed morpheme *pa:i* from Dai (in slightly different phonetic forms) to conjoin numerals, while employing native coordinators for nominals, as reported in Table 3.

Table 3: Nominal coordinators and numeral coordinators in five Austroasiatic languages

Languages	Nominal coordination	Numeral coordination		Sources
		Adjacent numerals	**Non-adjacent numerals**	
Blang (Bulang)	pap **kap** piʔ book **CONJ** pen 'books and pens'	sip ɛt ten one '11'	sɔŋ hɔi **pai** phuan two hundred **CONJ** five '205'	Li et al (1986: 36, 46)
Kəmuʔ (Kemu)	bɛʔ **pɔʔ** trak sheep **CONJ** cow 'sheep and cows'	sĭp ɛt ten one '11'	sɔŋ rɔi **blai** ha three hundred **CONJ** five '305'	Chen (2002: 161, 178)
Khɣmet (Kemie)	mɔi **hai** păŋ cow **CONJ** horse 'cows and horses'	sam sip et three ten one '31'	sam ɣɔiʔ **pai** kau three hundred **CONJ** nine '309'	Chen (2005: 88, 112)
Puɕiŋ (Buxing)	ʔua **pɔʔ** tuăih monkey **CONJ** locust 'monkeys and locusts'	ʃĕp sɔŋ ten two '12'	rai **pai** nəŋ hundred **CONJ** one '101'	Gao (2004: 99, 105)
Taʔaŋ (De'ang)	tau **ka:i** jɔŋ vegetable **CONJ** meat 'vegetables and meat'	ʔu kɣːr **lu** ʔu one ten **CONJ** one '11'	(i) ʔu jah **loi** ʔu one hundred **CONJ** one '101' (ii) ʔu hɛŋ **pa:i** ʔu one thousand **CONJ** one '1001'	Chen et al (1986: 46, 65, 73)

Among the ten Austronesian languages spoken in Taiwan that we surveyed, only Rukai has the same form *la* to conjoin numerals, nouns, and verbs. The other nine all feature different coordinators for numerals and for nouns, as reported in Table 4.

Table 4: Nominal coordinators and numeral coordinators in nine Austronesian languages

Languages	Nominal coordination	Numeral coordination	Sources
Amis	(i) k-u futing **atu** hmay NOM-NCM fish **CONJ** rice 'The fish and rice' (ii) ci aki-an **aci** panay-an NCM Aki-DAT **CONJ** Panay-DAT 'Aki and Panay'	a tosa polo' **ira** ko cecay[8] PRT twenty **CONJ** Article one '21'	Wu (2000); Council of Indigenous Peoples 2014 (CIP 2014)
Atayal	caxaʻ ku qumisuwan **ruʾ** caxaʻ ku suwayiʻ mu na kanayril one Nom older-sibling **CONJ** one Nom younger-sister 'an older sister and a younger sister'	magalpuw **cu** qutux ten **CONJ** one '11'	Huang (2000); CIP 2014
Bunun	Bunun **mas** asu hai, masial tu kaviaz. person **CONJ** dog TOPIC good COMP friend 'People and dogs are good friends.'	mapus-an **han** tasa[9] twenty **CONJ** one '21'	Zeitoun (2000); He et al (1986: 98, 101); CIP 2014
Kavalan	sunis **tu** baqian child **CONJ** old man 'children and old men'	Rabetin **yau** usiq ten **CONJ** one '11'	Chang (2000); CIP 2014

---

[8] He et al (1986: 56) report another coordinator for numerals in Amis: *tʃiʃafaw*, which is also a verb 'remain', as in (i). He et al (1986: 57) report a less commonly-used method to form additive numerals in this language: the number 11 and 999 can be expressed as (ii), which is called overcounting (Menninger 1969; Hurford 1975: 235-239; Meinunger 2015).

    (i) a. tuʃa   a   puluʔ **tʃiʃafaw** tu   tuʃa
        two PRT   ten   **CONJ**   PRT   two
        '22'
    (ii) a. tʃətʃaj (ku)   ʃakatuʃa   b. ʃiwa   a   ʃuʔut   ʃiwa ʃakamuətəp
        one (PRT) second   nine PRT hundred nine   tenth
        'second one=11'   'nine hundred and tenth nine=999'

[9] In He et al (1986: 61), the numeral coordinator is *tuhan*, which is described as a fused form of *tu han* 'again count'.

Paiwan	miŋɬayaɬayap a tʃaynan ʔa ta tʃutʃu. Progressive fly Article bee **CONJ** Article butterfly 'Bees and butterflies are flying.'	tapuʎuq **saʔa** ita ten **CONJ** one '11'	Chen and Ma (1986: 56, 73)
Puyuma	ɬatu giŋgiŋ **aw** asiru mango longan **CONJ** orange 'mangos, longans and oranges'	makapat **miasma** ɖa ɬuata forty **CONJ** PRT five '45'	Sun et al (2007: 2321, 2319)
Sakizaya	Dungi **aci** Maya Dungi **CONJ** Maya 'Dungi and Maya'	tusa a bataan **izaw** ku cacay two PRT ten **CONJ** PRT one '21'	CIP (2014)
Thao	Naak a ripnu **numa** taun My LINKER rice-paddy **CONJ** house 'My rice paddy and house'	makthin **ianan** tata ten **CONJ** one '11'	Sun et al (2007: 2227); CIP 2014
Tsou	o- puutu **ho** tsou Nom-Chinese **CONJ** Tsou 'Chinese people and Tsou people'	maskɨ-**veia** ucni ten **CONJ** one '11'	Sun et al (2007: 2254, 2260); CIP 2014

Except for Atayal, all of the Austronesian languages discussed above employ verbs to conjoin numerals. Examples include *ira* 'have' and *tʃiʃafaw* 'remain' in Amis, *(tu)han* '(again) count' in Bunun, *yau* 'have' in Kavalan, *saʔa* 'remain' in Paiwan, *miasma* 'remain' in Puyuma, *izaw* 'exist, have' in Sakizaya, *ianan* 'exist, have' in Thao, and *veia* 'return (to take something)' in Tsou.[10] They have main predicate uses, as shown below.

---

[10] We note that several Austronesian languages use the existential verb to conjoin numerals. This is reminiscent of Archaic Chinese in which numerals are obligatorily conjoined by *you* 'have' (or *you* 'again'). During the Spring and Autumn and the Warring Period (770-221 BC), this rule was relaxed and the existential verb is no longer used in the spoken language (Wang 1957: 256-257). However, the use of *you* within numerals can still be seen today in numerals expressing age with a sense of archaicity (see He 2015b: 192).

(i) 肇 十 有 二 州 ， 封 十 有 二 山。 《尚书》

Zhou shi you er zhou, Feng shi you er shan. *Shangshu*

Zhou ten have two prefecture, Feng ten have two mountain

'Zhou has 12 prefectures, and Feng has 12 mountains.'

Interestingly, the existential verb *ianan* in Thao is optional in numerals like 11, probably indicating that the rule is also weakening as happened in Archaic Chinese (*ianan* is not reported in Council of Indigenous Peoples 2014).

(8) a. Itia   hu   **ira** ka matʃahiaj a lumaˀ.   (Amis)

   that-time still **have** that poor   PRT family

   'There was a poor family once.'

   b. **Yau** uzusa lazum   na   tunek.   (Kavalan)

   have two hand Article clock

   'A clock has two hands.'

   c. Os-'o   **yuovei-a**   'o macucuma.   (Tsou)
   TR.RL-1S.ERG   **return-TR**   ABS   something

   'I returned to take something.'

3. Morphemes or allomorphs?

Table 5 is a summary of the major findings reported in the previous section.

Table 5: A summary of major findings

Languages	Nominal coordinators	Numeral coordinators	
		The word	Etymology
Tibeto-Burman languages			
Achang	*lɔˀ*	*mɔ*	Unclear
Anuŋ	*s ʔ*	*i*	Unclear
Baima	*re*	*tsɑ, re*	*tsɑ* is borrowed from the Tibetan *tsak* 'add'.
Jingpo	*theˀ*	*e*	Unclear
Nusu	*le*	*i*	Unclear
Pɤnru	*rəu*	*ɡaɬie*	*ɡaɬie*, a verb meaning 'exceed'
Suloŋ	*da*	*na*	Unclear
Tibetan	*ta* or *daŋ*	*tsak, so, ɕe, ŋa, re, tỹ, ca,* and *ko*	*tsak* (*rtsa* in written Tibetan) is etymologically from *brtsegs*, meaning 'add'. The others are associated with multiplier numerals of the teens.
Xiandao	*nɤk*	*maˀ* or *mua*	Unclear

Zaiwa	ə̟ʔ	kɔ̠m	kɔ̠m, an adjective meaning 'empty' (likely borrowed from the Chinese *kong* 'empty')
Tai-Kadai languages			
Dai	*lɛ (taŋ)*	*pa:i*	*pa:i*, an adjective meaning 'more'
Miao-Yao languages			
Miao	*kɔ*	*qa*	Unclear
Austroasiatic languages			
Blang	*kap*	*pai*	*pai*, borrowed from Dai, an adjective meaning 'more'
Kəmuʔ	*pɔʔ*	*blai*	*blai*, borrowed from Dai, an adjective meaning 'more'
Khɣmet	*hai*	*pai*	*pai*, borrowed from Dai, an adjective meaning 'more'
Puɕiŋ	*pɔʔ*	*pai*	*pai*, borrowed from Dai, an adjective meaning 'more'
Taʔaŋ	*ka:i*	*lu, loi, pa:i*	*loi, pa:i*, adjectives meaning 'more'
Austronesian languages			
Amis	*aci, atu*	*ira, tʃiʃafaw*	*ira*, a verb meaning 'have'; *tʃiʃafaw*, a verb meaning 'remain'
Atayal	*ruʔ*	*cu*	Unclear
Bunun	*mas*	*(tu)han*	*tu han*, a verb meaning 'again count'
Kavalan	*tu*	*yau*	*yau*, a verb meaning 'exist, have'
Paiwan	*ʔa*	*saʔa*	*saʔa*, a verb meaning 'remain'
Puyuma	*aw*	*miasma*	*miasma*, a verb meaning 'remain'
Sakizaya	*aci*	*izaw*	*izaw*, a verb meaning 'exist, have'
Thao	*numa*	*ianan*	*ianan*, a verb meaning 'exist, have'
Tsou	*ho*	*-veia*	*-veia,* from a verb *yuoveia* meaning 'return (to take something)'

All of the languages in Table 5 employ different coordinators for nominals and for numerals (some further use different coordinators for other categories, e.g., adjectives and verbs). It is important to note that the numeral coordinators in some of these languages are etymologically related to verbs or adjectives, and may have other uses, though the etymology of the numeral coordinators in the other languages is unclear, due to insufficient description in the cited

publications.

3.1 Tibetan

In this subsection, we focus on Tibetan. Under IM's proposal, (9a) is derived from (9b), which is, however, ungrammatical. The nominal coordinator must be either *ta* or *daŋ*, as in (9c). The same is true for other examples in (3). All of the Tibetan data in this paper have been confirmed firsthand by our Tibetan informants.

(9) a. mi   sum   tɕu   **so**   ŋa

   person three ten   **CONJ**   five

   '35 people'

  b. mi   sum   tɕu   ***so**   mi   ŋa

   person   three   ten   **CONJ**   person   five

   Intended meaning: '30 people and 5 people'

  c. mi   sum   tɕu   **ta/daŋ**   mi   ŋa

   person   three   ten   **CONJ**   person   five

   '30 people and 5 people'

This fact poses serious difficulty to the proposal that surface numeral coordination has an underlying nominal coordination structure, because a grammatical form cannot be derived from an ungrammatical base form. However, one may defend IM's proposal and treat the various numeral coordinators and nominal coordinators as allomorphs of a single morpheme. In other words, complex numerals have the syntactic structure outlined in IM, the default morphological form of coordination is the one that appears with overt nominal coordination, but coordination takes on

other morphological forms (i.e., phonological spell-out) in the context of certain types of deletion/movement and relative to the numerals being coordinated.[11]

Note first that under IM's semantics numerical bases are <<et>,<et>> type, requiring NP complements as their <et> arguments, as shown in [one [hundred [languages]]]. However, in Tibetan, the surface order of numeral expressions is head nouns + numerals. Therefore, a Tibetan multiplicative numeral expression *mi sum tɕu* 'people three ten' should have an underlying form of *sum tɕu mi*. The noun *mi* 'people' should be base-generated as the complement of the numerical base *tɕu* 'ten' and raise to the pre-numeral position, leaving a trace *t* behind, as illustrated in the top trees in (10a) and (10b). Therefore, in deriving (9a) from (9c), the noun *mi* should first undergo syntactic operations followed by the morphological change of *ta*/*daŋ* to *so*, as shown in (10a) for the deletion approach and (10b) for the raising approach. It seems that the reason triggering the morphological change of *ta*/*daŋ* to *so* is that *ta*/*daŋ* are *preceded and followed* by an empty trace. [12]

    (10)   a.   b.

---

[11] A reviewer reminds us that this type of allomorphy analysis is commonplace in Generative Semantics, where certain phonological forms can only appear after a movement and deletion operations (e.g., "kill" can only appear as a replacement for "cause to die" after the object of "die" is deleted and/or moved out of the complex verb phrase). However, it should be noted that the days when these types of Generative Semantics accounts which allow "kill" to be derived from "cause-to-die" in syntax were utilized within the field have long passed. Such accounts are too powerful to be entertained in the current generative syntactic theory.

[12] In Tibetan (and many other languages), modifiers including numerals and adjectives uniformly follow head nouns (see example (1) for noun+adjective structures, Sun et al (2007: 172)). A reviewer correctly points out that if the order is [Num N], and then the coordinator does not precede a trace, as illustrated in [Num ~~N~~ and Num N] or [[Num t_i and Num t_i] N_i]. In this case, the coordinator follows a trace.
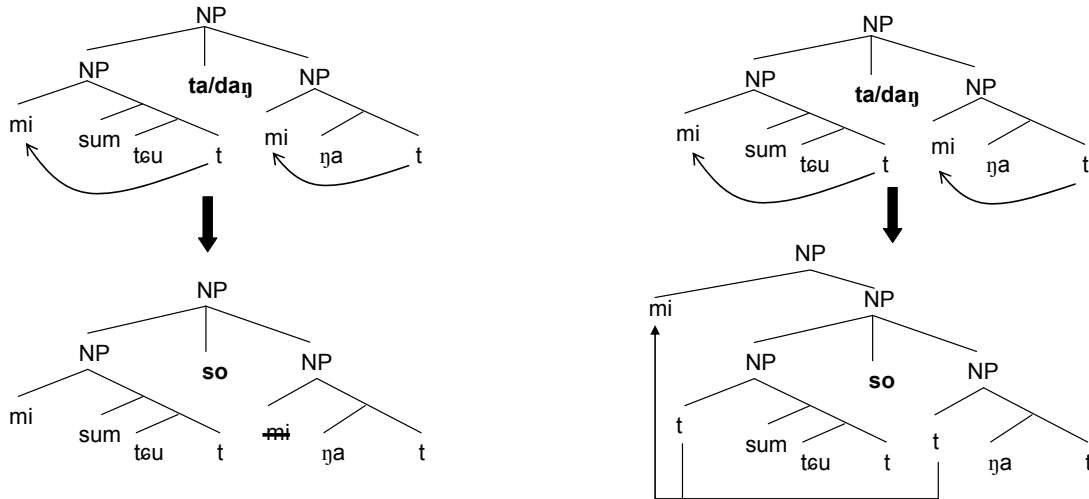
Such an account is problematic from various grammatical perspectives: morphological, etymological, syntactic, and semantic. First, in morphology, allomorphs are usually bound to a root and phonetically similar, e.g., *im*possible, *in*correct, *ir*regular, and *il*legal. In the Tibetan case, *so* and the other numeral coordinators are phonetically dissimilar from the nominal coordinator *ta/daŋ* and, as far as we know, an allomorph attaching in between two empty traces is not attested. Furthermore, such an account cannot explain why there are so many different allomorphs of *ta/daŋ* appearing in exactly the same environment, i.e., between two empty traces. The only solution is to stipulate a different *ad hoc* morpho-phonological rule for each and every phonological form, i.e., *tsak, so, ɕe, ŋa, re, tỹ, ca*, and *go*. Such stipulations must access the internal structure of the preceding NumP conjunct and be able to 'see' the value of the multiplier as well as the value of its complement, i.e., the 10^1 base. However, this would require extraordinary context-sensitive power, which thus renders such treatment unfeasible.

Second, the Tibetan numeral coordinators and nominal coordinators are different etymologically, indicating that they are also different semantically. According to Zhou (1998: 55-56), while *tsak* is derived etymologically from *brtsegs* 'add', the other numeral coordinators are derived from their multiplier numerals by certain phonological harmony rules, and these numeral

coordinators can occur without preceding numerals, fully indicating that such numeral coordinators are numerals in origin, not allomorphs of the nominal coordinators.

Third, in Tibetan the numeral coordinator conjoining hundreds and teens is the same one as the nominal coordinator, but the numeral coordinators conjoining teens (above 20) and digits are the special ones as discussed above. Example (11a) is supposedly derived from (11b) by means of deletion or raising, as shown in (12a) and (12b), respectively. Both instances of *ta/daŋ* in the two derivations are in the same syntactic environment, i.e., immediately preceded and followed by an empty trace. It is a mystery as to why *ta/daŋ* is realized as *tsak* and other allomorphs between teens and digits but unchanged between hundreds and teens, though they occur in exactly the same syntactic environment. If we treat complex numerals as constituents, the use of the nominal coordinator in larger Tibetan numerals can be explained by accounts which assume that *hundred/thousand* are of a different type from other smaller numerals, and more 'nominal'. See e.g. Rothstein (2013).

(11) a. mi   ca   **ta/daŋ**   ŋi   ɕu   **tsak**   cɛ²

       person hundred   **CONJ**   two ten   **CONJ**   eight

       '128 people'

  b. mi   ca   **ta/daŋ**   mi ŋi ɕu   **ta/daŋ**   mi   cɛ²

       person hundred   **CONJ**   person two   ten **CONJ**   person eight

       '100 people and 20 people and 8 people'

(12) a. mi   ca   t   **ta/daŋ** ~~mi~~ ŋi   ɕu   t   **tsak**   ~~mi~~ cɛ² t

  b. mi   t   ca   t   **ta/daŋ** t   ŋi   ɕu   t   **tsak**   t   cɛ²   t

Fourth, independent evidence shows that nominal coordinators do not take on other morphological forms in the context of certain types of deletion/movement. This is illustrated in (13a) (confirmed by all of our Tibetan informants, felicitous when there are two groups of people, 30 in one group and 5 in the other), in which the noun *mi* 'person' is topicalized. In (13a), *daŋ* is still used because after raising or deletion the two conjuncts are still noun phrases, not numeral phrases. Some informants had difficulty in understanding (13a) at first, but nevertheless accepted it as grammatical in its intended reading after our explanation. They readily understood (13b), which involves 35 people in one group.

(13) a. mi   ni   thakī   sum   tɕu   **daŋ**   ŋa   lɛʔsoːŋ.

person Topic-marker   just-now three   ten   **CONJ**   five   come

'People, there came thirty (in one group) and five (in another group) just now.'

b. mi   ni   thakī   sum   tɕu   **so**   ŋa   lɛʔsoːŋ.

person Topic-marker   just-now three   ten   **CONJ**   five   come

'People, there came thirty-five (in one group) just now.'

As mentioned earlier, if there are missing powers between two numerals, Tibetan employs a special method not found in the other languages in our survey, which entails inserting *tɕu mɛʔ* (ten no) or *ca mɛʔ tɕu mɛʔ* 'hundred no ten no', as shown in (6a/b). According to the non-constituency analysis, (6a) should be derived from (14a), and (6b) should be derived from (14b). This implies that in order to derive the surface forms, the nominal coordinator *ta/daŋ* 'and' needs to change to *tɕu mɛʔ* 'ten no' and *ca mɛʔ tɕu mɛʔ* 'hundred no ten no', which seems entirely unworkable.

(14) a. mi   sum   ca   **ta/daŋ**   mi   tʂ'uk

      person   three hundred   **CONJ**   person   six

      '300 people and 6 people'

    b. mi   sum   toŋ   **ta/daŋ**   mi   tʂ'uk

      person   three thousand   **CONJ**   person   six

      '3000 people and 6 people'

A proposal that the alternative underlying form of (6a) be (15) does not work either, as this would require that the two instances of the nominal coordinators *ta/daŋ* should disappear after transformation. Our Tibetan informants stated that *mi tɕu mɛ*ˀ 'person ten no' is grammatical in syntax and interpretable in semantics ('not have ten people' or 'less than ten people'), but that the structure as a whole is uninterpretable. When pressed further, some could obtain a reading of '296 people in total', i.e., [300 people － 10 people ＋ 6 people]. When there is a lack of semantic equivalence between two structures, it is untenable that the two forms can be derivationally related (see Her and Tsai 2015 for a recent discussion).

(15) *mi   sum   ca   **ta/daŋ** mi   tɕu   mɛˀ   **ta/daŋ**   mi   tʂ'uk

    person three hundred **CONJ** person ten   not-have **CONJ**   person   six

Thus far, we have demonstrated that the proposal that numeral coordination is derived from underlying nominal coordination encounters serious difficulty in Tibetan and that the two kinds of coordinators are different morphemes altogether, and thus not allomorphs of the same morpheme. There is also reason to believe that it is quite a common phenomenon across languages that different coordinators are used to conjoin different syntactic categories and that these coordinators are

different morphemes and not allomorphs of the same morpheme. One such typical language is Mandarin Chinese, in which nominal phrases are conjoined by *he*, *gen*, *ji*, *yu*, *yiji*, and *jian*; predicative phrases (VP and AP) and clauses are conjoined by *erqie* or *bingqie*, though a null coordinator may be preferable in certain contexts. Particularly within coordinated noun phrases, the coordinator *he* 'and' only allows a split reading and the coordinator *jian* 'and' is used to express the joint reading (Aoun & Li 2003: 141-143). The sentence (16) below means some are linguists and some are philosophers if *he* is used; but if *jian* is used, it means they are both linguists and philosophers. This clearly indicates that *he* and *jian* are two different morphemes, because they appear in the same syntactic position, not in complementary distribution, which is a distinctive feature of allomorphy. The two forms are not phonetically similar and, more importantly, they are not semantically equivalent.

(16) Tamen shi yuyanxuejia **he/jian**   zhexuejia.

   they are linguist   **CONJ** philosopher

   'They are linguists and philosophers.'

3.2 Other languages

We obtained firsthand verification of core data for ten other languages. All of the data presented here have been confirmed by at least two native speakers and/or experts in the relevant languages.

Table 6: Verification of the grammaticality of numeral expressions for other languages

Austronesian languages		
Amis	Numeral expression	a tosa polo' **ira** ko cecay a tamdaw
		PRT twenty **CONJ** Article one PRT person

		'21 people'
	Alleged source form/Correct form	a tosa polo' a tamdaw ***ira/atu** cecay a tamdaw PRT twenty PRT person **CONJ** one PRT person 'twenty persons and one person'
Atayal	Numeral expression	magalpug **cu** qun cuquliq ten **CONJ** one person 'eleven people'
	Alleged source form/Correct form	magalpug cuquliq ***cu/ru'** taxa cuquliq[13] ten person **CONJ** one person 'ten persons and one person'
Bunun	Numeral expression	mapus-an **han** tasa tu asu twenty **CONJ** one ATTR dog '21 dogs'
	Alleged source form/Correct form	mapus-an tu asu ***han/mas** tasa tu asu twenty ATTR dog **CONJ** one ATTR dog 'twenty dogs and one dog'
Tsou	Numeral expression	maskɨ-**veia** ucni ci sensi ten-**CONJ** one ATTR teacher '11 teachers'
	Alleged source form/Correct form	maskɨ ci sensi ***veia/ho** ucni ci sensi ten ATTR teacher **CONJ** one ATTR teacher 'ten teachers and one teacher'
Tibeto-Burman languages		
Achang	Numeral expression	tʂo tɑ pak **mɔ** tɑ sau ʐuʔ people one hundred **CONJ** one twenty Cl '120 people'

---

[13] Careful readers may have noticed that the numeral for 'one' in this expression is *taxa*, not *qun*. In Atayal, there are two morphemes for *one* and *two* each: *qutux*/*caxa'* (1) and *usayng*/*rarusa* (2). *Qun* (*qutux*) is used for compound numerals while *caxa* (*taxa*) is used to modify nouns. This phenomenon, which is fairly common among minority languages in South China, causes further difficulty to the derivational non-constituency analysis (See He 2015b: 198-199, 211-212 for discussion).

	Alleged source form/Correct form	tʂo tɑ pak ʐuʔ ***mɔ/lɔʔ** tʂo tɑ sau ʐuʔ people one hundred Cl **CONJ** people one twenty Cl '100 people and 20 people'
Jingpo	Numeral expression	tʃum kjin khjiŋ mǎŋa **e** tʃăkhu tsa salt kilo thousand five **CONJ** nine hundred '5900 kilos of salt'
	Alleged source form/Correct form	tʃum kjin khjiŋ mǎŋa ***e/theʔ** tʃum kjin tʃăkhu tsa salt kilo thousand five **CONJ** salt kilo nine hundred '5000 kilos of salt and 900 kilos of salt'
Zaiwa	Numeral expression	səkkam sum ʃo **kɔm** ʃit kam tree three hundred **CONJ** eight Cl '308 trees'
	Alleged source form/Correct form	səkkam sum ʃo kam ***kɔm/əʔ** səkkam ʃit kam tree three hundred Cl **CONJ** tree eight Cl '300 trees and 8 trees'
Tai-Kadai languages		
Dai	Numeral expression	hok pan **pa:i** sa:u kun six thousand **CONJ** twenty people '6020 people'
	Alleged source form/Correct form	hok pan kun ***pa:i/lɛ** sa:u kun six thousand people **CONJ** twenty people '6000 people and 20 people'
Miao-Yao languages		
Miao	Numeral expression	ɯ tshɛ **qɑ** ʑi le ne two thousand **CONJ** eight Cl person '2008 people'
	Alleged source form/Correct form	ɯ tshɛ le ne ***qɑ/kɔ** ʑi le ne two thousand Cl person **CONJ** eight Cl person '2000 people and 8 people'

Austroasiatic languages		
Taʔaŋ	Numeral expression	(i) ʔu kɣːr **lu** ʔu luʔ (ii) ʔu jah **loi** ʔu luʔ one ten Cl **CONJ** one Cl one hundred **CONJ** one Cl '11 Cl' '101 Cl' (iii) ʔu hɛŋ **paːi** ʔu luʔ one thousand **CONJ** one Cl '1001 Cl'
	Alleged source form/Correct form	(i) ʔu kɣːr luʔ ***lu/kaːi** ʔu luʔ (ii) ʔu jah luʔ ***loi/kaːi** ʔu luʔ one ten Cl **CONJ** one Cl one hundred Cl **CONJ** one Cl '10 Cl and 1 Cl' '100 Cl and 1 Cl' (iii) ʔu hɛŋ luʔ ***paːi/kaːi** ʔu luʔ one thousand Cl **CONJ** one Cl '1000 Cl and 1 Cl' (Chen et al 1986: 46, 65, 73)

Similar to the case of Tibetan, there are good reasons here too why the numeral coordinators and nominal coordinators in these languages are not allomorphs of the same morpheme and are not derivationally related. Firstly, these coordinators are phonetically dissimilar. Secondly, the numeral coordinators in some languages are semantically different from the nominal coordinators. For example, most numeral coordinators in Austronesian languages are related to verbs. The numeral coordinator in Zaiwa is an adjective meaning 'empty', and the numeral coordinator in Dai (and several Austroasiatic languages influenced by Dai) is also an adjective *pa:i* meaning 'more', which have other uses. In Taʔaŋ, there are several numeral coordinators but only one nominal coordinator, and the choice of the numeral coordinator depends on the numerals: *lu* between two adjacent powers, *loi* between two powers with one missing power, and *pa:i* between two powers with two or more missing powers. If they were allomorphs, it would be difficult to explain why the same morpheme should take on different forms according to the numerals.

4. Numerals or coordinators?

Among the languages that we surveyed, the numeral coordinators in some can occur between adjacent powers, including in Achang, Anuŋ, Baima, Jingpo, Pɣnru, Suloŋ, Xiandao, Taˀaŋ, Atayal, Amis, Bunun, Kavalan, Paiwan, Puyuma, Sakizaya, Thao, and Tsou, and are thus unmistakably coordinators. However, other languages use numeral coordinators only between non-adjacent powers, but not between adjacent powers, e.g., Zaiwa, Nusu, Dai, Miao, Blang, Buxing, Khɣmet and Kəmuˀ. This is reminiscent of the Chinese morpheme *ling* 'zero', which is used only to conjoin non-adjacent powers, and its loan incarnations occuring in many other languages influenced by Chinese (see section 5.2). He (2015b: 192) argued that the Chinese *ling* 'zero' within additive numerals should be treated as a numeral not as a coordinator because the form of multiple instances of *ling* is acceptable in Modern Chinese (particularly in very formal context), though the form with only one instance of *ling* is far more popular no matter how many missing powers there are between two numerals[14]. This indicates that *ling* is a numeral for zero not a coordinator, because the stacking of coordinators seems not attested. One may thus be tempted to argue that the so-called numeral coordinators in Zaiwa, Nusu, Dai, Miao, Blang, Buxing, Khɣmet and Kəmuˀ are not coordinators and are in fact numerals expressing the meaning of zero like the Chinese *ling* 'zero'. If their so-called numeral coordinators turn out not to be coordinators, then the data of these languages which we have provided here should not be considered in this paper at all to support our position. It is therefore important for us to demonstrate that the numeral coordinators of Zaiwa (*kɔm*), Nusu (*i*), Blang (*pai*), Buxing (*pai*), Khɣmet (*pai*) and Kəmuˀ (*blai*) do not function as numerals in a way

---

[14] Dialectal variation does exist in terms of the stacking of *ling* for missing powers. In Taiwan Mandarin, for example, it is generally not allowed but does appear in special registers such as court rulings and formal contracts. It is thus arguable that, for speakers that do not allow *ling*-stacking at all, *ling*, besides being a numeral for zero, is now undergoing reanalysis to be a numeral coordinator as well.

similar to the Chinese *ling* 'zero'.

First of all, the numeral coordinators in these languages do not express the meaning of 'zero' and cannot occur in modifier and argument positions like the Chinese *ling*. To express the number zero, Zaiwa borrowed the Chinese *ling*, but Dai has an indigenous morpheme *sun*, as shown in (17) and (18). The Dai 'zero' morpheme *sun* cannot be used to link non-adjacent powers, as shown in (19).

(17) a. pju ***kɔm/ling** ju^ʔ b. ***pa:i/sun** kun

person **zero** Cl   **zero** people

Zaiwa: 'zero persons'   Dai: 'zero persons'

(18) a. ʒa jɔm ʒa mjit ***kɔm/ling**.

one subtract one remain **zero**

Zaiwa: 'One minus one is zero.'

b. sɔŋ lup sɔŋ jaŋ ***pa:i/sun**.

two minus two remain **zero**

Dai: 'Two minus two is zero.'

(19) *hok pan **sun** sa:u kun

six thousand **CONJ** twenty people

Intended meaning: '6020 people'

The numeral coordinators in some languages are verbs or adjectives. In Zaiwa, *kɔm* is an adjective meaning 'empty' and can thus be used as a typical adjectival modifier, as in *lo^ʔ akɔm*

'hand empty' (Note that *kɔm* is instantiated as *akɔm* in this structure). And in Dai, *pa:i* means 'more' (thus clearly not 'zero'), which has other uses, as shown in (7) in Section 2.2. The numeral coordinators in several Austroasiatic languages can also have other functions, as evidenced below.

(20) a. kui   kul **pai**   pɣi.   (Blang, Li et al (1986: 36))

   have ten **more** person

   'There are a little more than ten people.'

 b. Ai   ɔ^ʔ pen   hok sip nɣm **pai**.   (Khɣmet, Chen (2005: 90))

   father my   PRT six ten year **more**

   'My father is a little more than 60 years old.'

Second, unlike Modern Chinese that still accepts the stacking of *ling* for multiple missing powers, the above numeral coordinators in Dai and Zaiwa do not allow such stacking at all. They can only occur once no matter how many powers are missing, as shown in (21).

(21) a. *pan   **pa:i   pa:i**   sɔŋ   to

   thousand **CONJ CONJ** two   Cl

   Dai: '1002 Cl'

 b. *səkkam sum   khjiŋ   **kɔm   kɔm**   ʃit   kam

   tree   three   hundred   **CONJ CONJ**   eight Cl

   Zaiwa: '3008 trees'

Miao is different. In Jiwei Township (a typical Miao residential area in Huayuan County in the west of Hunan Province, where people speak Miao within themselves) where we conducted our

field work, the two forms in (22a/b) are both well-formed, but younger people prefer *ljin* (clearly a Chinese borrowing) and older people prefer *qa*. All reject (22c). So it is unclear whether *qa* within additive numerals is a numeral like the Chinese *ling*. But even though *qa* is a real numeral expressing the numerical value of zero, it still causes problems for the non-constituency analysis (see He 2015b: 204-205 for discussion).

(22) a. **qɑ/ljin**   le   ne

      zero   Cl   person

      'zero persons'

   b. A   tɕiʑɔ   a   tɔ   **qɑ/ljin**.

      one subtract one   get   **zero**

      'One minus one is zero.'

   c. *ɯ   tshɛ   **qɑ**   **qɑ**   ʑi   le

      two thousand   **CONJ CONJ** eight Cl

      '2008 Cl'

5. Theoretical implications

5.1 Syntax and semantics of numeral coordination

In the previous sections, we have argued that at least for some languages, additive numerals cannot be derived from nominal coordination. The theoretical implication is that nominal coordination and numeral coordination must be distinguished in natural languages. The following is a proposal of the phrase structure for additive numerals, illustrated with a Chinese example (see He 2015b: 202 section 4.3 'The '&' head' for reasons why a covert coordinator is needed).

(23)

NumeralP

NumeralP   CONJ   NumeralP

san   shi   san
'three'   'ten'   'three'
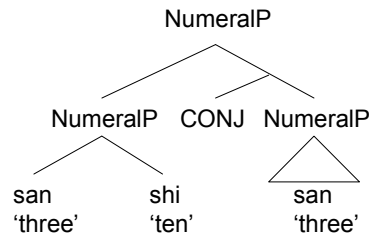
It is commonly agreed that nominal coordinators generate a set or a sum of different entities or just many objects at once (see McKay 2006 for an overview); numeral coordinators, though morphologically identical to nominal coordinators in some languages, encode different semantic relations in that they generate a bigger number out of smaller ones. Therefore, the CONJ morpheme (overt or covert) in (23) contributes a different semantic value from that of nominal coordinators, though they may be homophones. They should be distinguished as different morphemes, and should thus be defined differently. If we assume the Fregean view that numerals denote numbers with semantic type *e*, the semantic relation between coordinated numerals is that of *arithmetic addition*, and the numeral coordinator can be defined as *arithmetic addition* '+' as shown in (24a). If we assume that numerals denote sets (type $\langle et \rangle$) following Rothstein (2013), then numerals are adjectives defined probably as $\|numeral\| = \lambda X[|X|=n]$, then possibly the numeral coordinator can be defined as in (24b). Note that there should be a restriction for the X and Y such that they belong to the same category of objects required for counting (Rothstein 2010), and the equation '$Z = X \cup Y$ & $|Z|=|X|+|Y|$' can guarantee that there is no overlapping among the objects.

(24) a. $\|CONJ_{numeral}\| = \lambda m \lambda n [m+n]$

      b. $\|CONJ_{numeral}\| = \lambda P \lambda Q \lambda Z \exists XY[X \in P \ \& \ Y \in P \ \& \ Z = X \cup Y \ \& \ |Z|=|X|+|Y|]$

In both treatments, the semantics of $CONJ_{numeral}$ contains *arithmetic addition*, which

distinguishes numeral coordinators from nominal coordinators. This is best illustrated in the fact that the numeral coordinators in some languages we have discussed are etymologically verbs or adjectives, conveying a meaning like 'add, supplement, exceed, or count'. Given the semantic relation of arithmetic addition between coordinated numerals, numeral coordination may be subject to certain arithmetically-driven syntactic restrictions that nominal coordination is immune to. For example, nominal coordination does not seem to formally impose an order restriction on the conjuncts, i.e., the two expressions *Jack and Jill* and *Jill and Jack* are semantically equivalent. Though addition of numbers is commutative (10+3=3+10), additive numerals are subject to a sequential order, as shown in the contrast between *one hundred and two* and **two and one hundred*. According to Hurford (2007), this order restriction in numerals is a consequence of a counting principle "Go as far as you can with the resources you have". We can thus further interpret this word order restriction in additive numerals as an example of the grammaticalization of a functional principle.

5.2 A typology of numeral coordination

We cautiously propose that the syntax-semantics for numeral coordination sketched above is universal and applicable to numeral coordination in all languages (except those that may not have surface numeral coordination like Biblical Welsh and Kalabari), including languages that use the same phonetic forms for both numeral and nominal coordinators, as well as languages that use covert coordinators. Through our survey of more than 100 minority languages, we found that both of the latter types of languages are widely attested. In some languages, the numeral coordinators are of the same phonetic forms as the nominal coordinators, including Bai (*liɯ*), Guiqiong (*lɛ*), Lavrung (*ræ*), Monba (*daŋ*), Muya (*rə*), Namuyi (*na*), Naxi (*ne*), Nosu (*sini*), Primi (*nə*), Qiang (*na*), Then

(*thim*), Tosu (*la*), Zhaba (*nə*). Covert coordination is also common, in languages such as Bola, Bugan, Derung, Hani, Jinuo, Kuman, Lhao Vo, Sangkong, etc.[15] These languages are all Tibeto-Burman languages, spoken mainly in the remote Himalaya hinterlands and Yunnan Province in Southwest China.

Some of the languages that do not have overt coordinators within additive numerals employ a special 'zero' morpheme to fill in missing powers only. This is due to the influence of the Chinese *ling* 'zero'. These 'zero' morphemes are phonetically adapted as loan words. Examples of languages in which this occurs include Biao (*liaŋ*), Bouyei (*liŋ*), Bunu (*liŋ*), Buyang (*lan*), Chadong (*ləŋ*), Cun (*ləŋ*), Gelao (*liŋ*), Huihui (*lin*), Jiamao (*leŋ*), Jing (*len*), Kam (*ljən*), Khatso (*li*), Lai (*le*), Lakkia (*lɛŋ*), Lashi (*lɔ*), Lin'gao (*leŋ*), Mang (*liŋ*), Maonan (*li:ŋ*), Mjen (*leŋ*), Mulam (*liŋ*), Sui (*ljen*), Taliu (*ɳi*), Tujia (*lin*), Younuo (*lən*), etc.[16] These languages belong to Miao-Yao and Tai-Kadai, spoken in the provinces of Hunan, Guangxi, Guizhou, Yunnan, and Hainan, where the cultural and linguistic influence from Chinese is great. It is not clear at this point whether the 'zero' morphemes in these Miao-Yao and Tai-Kadai languages are numerals or coordinators due to a lack of confirmation as to whether they allow consecutive 'zero' morphemes in numerals.

A typology of numeral coordination thus emerges, with different languages employing different morphological strategies to conjoin numerals, as shown in Table 7. The conjoining morpheme may be overt or covert. In the case of overt coordination, the coordinators need not be of the same

---

[15] Example (i) is from Qiang, the former type; example (ii) is from Lhao Vo and Sangkong, the latter type.
    (i) a. χma   **na**   dzuə   b. a   tʂhi   **na**   χɑ dy **na** a
        millet **CONJ** chestnut   one hundred **CONJ** one   ten **CONJ** one
        'millets and chestnuts'   '111'   (Sun et al 2007: 852, 861)
    (ii) a. tă   jɔ   ta   b. ɕa   ɕe̱
        one hundred one   hundred eight
        Lhao Vo: '101'   Sangkong: '108'
[16] The following examples are from Chadong and Taliu.
    (i) a. ji   pek **ləŋ** jit   b. tshɿ xõ   **ɳi** ŋo
        one hundred **zero** one   one hundred **zero** five
        Chadong, '101'   Taliu, '105'

phonetic forms as the nominal coordinators. In these situations, they may have diverse etymologies, with many numeral coordinators etymologically derived from verbs and adjectives. This is the case for the 26 languages discussed in this paper.

Table 7: A typology of numeral coordinations of languages in South China

Typology of numeral coordination	Numeral coordinators		Examples
	Surface form	**Homophony with nominal coordinators?**	
Type 1	Covert	NA	Lhao Vo, Sangkong, etc
Type 2	Overt	Yes	Qiang, Guiqiong, etc
Type 3		No	Tibetan, Amis, etc
Type 4	*Ling* morpheme 'zero'	*Ling* is treated as a numeral.	Chinese
		Ling is treated as a coordinator.	Perhaps Miao

We have argued that numeral coordination in Type 3 languages cannot be derived from nominal coordination. For Type 4 languages, if *ling* is treated as a numeral, the arguments concerning the Chinese *ling* against the non-constituency analysis also apply to these languages; if *ling* is treated as a coordinator, the arguments put forth in this paper also apply to these languages because these languages have different nominal coordinators, as shown in Table 8.

Table 8 Nominal coordinators in Type 4 languages (surveyed from Sun et al 2007)

Language	Nominal coordinator	Example	Language	Nominal coordinator	Example
Biao	*kuŋ*	tsai **kuŋ** tsy bowl **CONJ** chopstick 'bowls and chopsticks'	Bouyei	*tiam*	tu ni **tiam** tu ti Cl this **CONJ** Cl that 'this and that'
Bunu	*pu*	aŋ **pu** to water **CONJ** fire 'water and fire'	Buyang	*qha*	ʐuk **qha** lava:k palm tree **CONJ** camphor tree 'palm trees and camphor trees'

Cun	*nam*	hɔ lət azɔ **nam** tɯək na that Cl old man **CONJ** son his 'that old man and his son'	Gelao	*tshɑ*	tsɯ tsaŋ **tshɑ** mpɑŋ pen **CONJ** ink 'pens and ink'
Huihui	*ŋan*	ha **ŋan** thaimai you **CONJ** sister 'you and sister'	Jing	*vəi*	oŋthəi **vəi** hɔktɔ teacher **CONJ** student 'teachers and students'
Kam	*taŋ*	tu na:i **taŋ** tu ta Cl this **CONJ** Cl that 'this and that'	Khatso	*kɯ*	koko **kɯ** titi **kɯ** older brother **CONJ** younger brother **CONJ** 'older brothers and younger brothers'
Lai	*le*	ʔa:u **le** mi I **CONJ** you 'I and you'	Lakkia	*kap*	tsi **kap** ma I **CONJ** you 'I and you'
Lashi	*jɔ*	apho **jɔ** amji father **CONJ** mother 'father and mother'	Lin'gao	*hem*	ma **hem** mo dog **CONJ** pig 'dogs and pigs'
Mang	*ʑua*	vantɕy **ʑua** tətɕɯa man **CONJ** woman 'men and women'	Maonan	*ɗam*	ɦe **ɗam** man I **CONJ** he 'I and he'
Mjen	*tshin*	noku **tshin** noŋu owl **CONJ** dove 'owls and doves'	Mulam	*wən*	ljem **wən** tsui sickle **CONJ** hammer 'sickles and hammers'
Sui	*kɐp*	pɑkjiu **kɐp** mɐifənli scissor **CONJ** ruler 'scissors and rulers'	Taliu	*ŋe*	ŋaʐo **ŋe** ŋamo brother **CONJ** sister 'brothers and sisters'
Tujia	*ne*	lapu **ne** tɕhipu salt **CONJ** bean 'salt and beans'	Younuo	*la*	pje tuŋ ŋu **la** pi tuŋ ʑaŋ three Cl cow **CONJ** five Cl sheep 'three cows and five sheep'

For Type 1 and Type 2 languages, our argument does not apply. It is possible that additive

numerals are derived from NP coordination in these languages, as Hurford (1987: 226-238) proposed quite many years ago, which has a much more detailed discussion of the type of theory proposed by IM (but see He 2015b for possible problems based on other arguments). However, Hurford (1987: 237) cautioned that "[i]t is not claimed that this is the *only* way in which complex numerals could arise." He cited Fijian numerals as an example in which numerals are conjoined by *a*, which also conjoins clauses and phrases (the noun coordinator *'ei* cannot be used). He further commented that "[i]n the light of such facts, it must be admitted that additive constructions do not always arise from conjunctions of NPs, although this may well be their most typical evolutionary source…It would be interesting to investigate cases like Fijian further, rare though they are." Hurford's remarks are robustly demonstrated to be true in this paper, and cases like Fijian are not rare. Among over 100 languages surveyed in this paper, nearly one third belong to this category.

6. Conclusion

In this paper, we investigated a special typology of overt coordination in additive numerals of minority languages spoken in South China. It is found that among about 100 minority languages, 26 feature different coordinators for additive numerals from noun phrases and that these two types of coordinators are not *semantically, etymologically,* or *morphologically* related. We demonstrated that this phenomenon strongly indicates that additive numerals are not syntactically derived from nominal coordination. We made some specific suggestions to formally distinguish the syntax and semantics of numeral coordination from that of nominal coordination, and outlined a typology of numeral coordination, of which Type 3 languages have particular value to the study of the syntax of natural language numerals. We believe Type 3 languages can be found in other parts of the world.

Acknowledgments

References

Aoun, J. and Y-H A. Li. 2003. *Essays on the Representational and Derivational Nature of Grammar*. Cambridge, Mass.: MIT Press.

Chang, H. Y. 2000. *A Reference Grammar of Kavalan*. Taipei: Yuan-Liou Publishing.

Chen, K. & R.-S. Ma. 1986. *Gaoshanzu Yuyan Jianzhi (Paiwanyu)* 'A Grammar Sketch of Gaoshan Ethnic People's Languages (Paiwan Language)'. Beijing: Minzu (Nationality) Press.

Chen, G.-Q. 2002. *Kemuyu Yanjiu* 'Studies on Kemu Language'. Beijing: Minzu (Nationality) Press.

Chen, G.-Q. 2005. *Kemieyu Yanjiu* 'Studies on Kemie Language'. Beijing: Minzu (Nationality) Press.

Chen, X.-M., J.-L. Wang & Y.-L. Lai. 1986. *De'angyu Jianzhi* 'A Grammar Sketch of De'ang Language'. Beijing: Minzu (Nationality) Press.

Council of Indigenous Peoples. 2014. Indigenous Peoples' Supplementary Language Learning Materials on Sentence Patterns: Mayrinax Atayal, Coastal Amis, Mantauran Rukai, Sakizaya, Saisiyat, Thao. Taipei: Council of Indigenous Peoples. Available online at: http://lokahsu.org.tw/resource/book/senior/.

Corver, N. & J. Zwarts. 2006. Prepositional numerals. *Lingua* 116: 811–35.

Dai, Q.-S. & Z.-C. Cui. 1985. *Achangyu Jianzhi* 'A Grammar Sketch of Achang Language'. Beijing: Minzu (Nationality) Press.

Dai, Q.-S. et al. 2005. *Xiandaoyu Yanjiu* 'Studies on Xiandao Language'. Beijing: Central Minzu University Press.

Gao, Y.-Q. 2004. *Buxingyu Yanjiu* 'Studies on Buxing Language'. Beijing: Minzu (Nationality) Press.

He, C.-S. 2015a. Ziran yuyan shuci xitong jufa yuyi jiekou lilun de zuixin yanjiu jinzhan 'Recent developments in the studies of the syntax-semantics interface of numerals'. *Dangdai Yuyanxue* 'Contemporary Linguistics' 17: 253-272.

He, C.-S. 2015b. Complex numerals in Mandarin Chinese are constituents. *Lingua* 164: 189-215.

He, R.-F., S.-Q. Zeng, W.-S. Li, & Q.-C. Lin. 1986. *Gaoshanzu Yuyan Jianzhi (Bunenyu)* 'A Grammar Sketch of Gaoshan Ethnic People's Languages (Bunun Language)'. Beijing: Minzu (Nationality) Press.

He, R.-F., S.-Q. Zeng, Z.-S. Tian & D.-S. Lin. 1986. *Gaoshanzu Yuyan Jianzhi (Ameisiyu)* 'A Grammar Sketch of Gaoshan Ethnic People's Languages (Amis Language)'. Beijing: Minzu (Nationality) Press.

Her, O.-S. & H.-C. Tsai. 2015. On silent elements: A case study of grand and its silent entourage. *Natural Language and Linguistic Theory* 33: 575-605.

Huang, L. M. 2000. *A Reference Grammar of Atayal*. Taipei: Yuan-Liou Publishing.

Hurford, J. 1975. *The Linguistic Theory of Numerals*. Cambridge: Cambridge University Press.

Hurford, J. 1987. *Language and Number*. Oxford: Blackwell.

Hurford, J. 2007. A performed practice explains a linguistic universal: counting gives the Packing Strategy. *Lingua* 117: 773-83.

Ionin, T. &, O. Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23: 315-360.

Kayne, R. 2010. A note on the syntax of numerical bases. In *Comparisons and Contrasts*, 57-72. New York: Oxford University Press.

Li, D.-Q. 2004. *Sulongyu Yanjiu* 'Studies on Sulong Language'. Beijing: Minzu (Nationality) Press.

Li, D.-Y., X.-Z. Nie, & E.-F. Qiu. 1986. *Bulangyu Jianzhi* 'A Grammar Sketch of Bulang Language'. Beijing: Minzu (Nationality) Press.

Liu, L. 1984. *Jingpozu Yuyan Jianzhi (Jingpoyu)* 'A Grammar Sketch of Jingpo Ethnic People's Languages (Jingpo Language)'. Beijing: Minzu (Nationality) Press.

McKay, T. 2006. *Plural Predication*. Oxford: Oxford University Press.

Menninger, K. 1969. *Number Words and Number Symbols*. Trans. P. Bronner. Cambridge, MA: MIT Press.

Meinunger, A. 2015. Complex numeral expressions: A plea for grafts. *Syntax* 18: 103–123.

Rothstein, S. 2010. Counting and the mass count distinction. *Journal of Semantics* 27: 343-397.

Rothstein, S. 2013. A Fregean semantics for number words. In *Proceedings of the 19th Amsterdam Colloquium,* eds. M. Aloni, M. Franke & F. Roelofsen, 179-186.

Sun, H.-K., Z.-Y. Hu & X. Huang (eds). 2007. *Zhongguo de Yuyan* 'The Languages of China'. Beijing: Commercial Press.

Sun, H.-K. & G.-K. Liu. 2005. *Anongyu Yanjiu* 'Studies on Anong Language'. Beijing: Minzu (Nationality) Press.

Sun, H.-K. & L. Liu. 1986. *Nusuzu Yuyan Jianzhi (Nusuyu)* 'A Grammar Sketch of Nusu Ethnic People's Languages (Nusu Language)'. Beijing: Minzu (Nationality) Press.

Sun, H.-K., K.-J. Qi & G.-K. Liu. 2007. *Baimayu Yanjiu* 'Studies on Baima Language'. Beijing: Minzu (Nationality) Press.

Wang, L. 1957. *Hanyu Shigao* 'A History of Chinese'. Beijing: Commercial Press.

Wu, J-L. J. 2000. *A Reference Grammar of Amis*. Taipei: Yuan-Liou Publishing.

Xiang, R.-Z. 1999. *Jiwei Miaoyu Yanjiu* 'Studies on Miao Language in Jiwei'. Chengdu: Sichuan People Press.

Yu, C.-R. & M.-Z. Luo. 1980. *Daiyu Jianzhi* 'A Grammar Sketch of Dai Language'. Beijing: Minzu (Nationality) Press.

Zeitoun, E. 2000. *A Reference Grammar of Bunun*. Taipei: Yuan-Liou Publishing.

Zhou, M.-C. 1998. Zangyu fuhe shuci zhong de lianjie chengfen 'Linking elements within compound numerals in Tibetan'. *Minzu Yuwen* 'Minority Languages', issue 2: 53-58.

Zhu, Y.-H. 2011. *Zaiwayu Cankao Yufa* 'A Reference Grammar of Zaiwa Language'. PhD Dissertation, Central Minzu University, Beijing.