

Chenlei Zhou*

Towards a new typology of comparative constructions in East Asian languages

<https://doi.org/10.1515/lingty-2023-0028>

Received April 1, 2023; accepted November 26, 2023; published online January 22, 2024

Abstract: This paper presents a typology that sheds light on the diversity and complexity of comparative constructions across languages. Specifically, this study focuses on topic-prominent comparatives, a newly discovered type of comparative construction commonly found in East Asian languages, providing a comprehensive analysis of their features and subtypes. Drawing from a wide range of East Asian languages, this study delineates three subtypes of topic-prominent comparatives: double-comparatum comparatives (‘Hair she is longer than me’), comparee-standard mismatched comparatives (‘Her hair is longer than me’), and standard-topic comparatives (‘My hair, her hair is longer’). Additionally, this study introduces a pair of new parameters: concrete measurement and abstract measurement of comparative degrees. Concrete measurement involves the use of precise numerical values or quantifiable units to make comparisons, such as *three years* in *I am three years older than you*. Conversely, abstract measurement involves the use of non-specific, relative terms to establish comparisons, such as *much* in *I am much older than you*. These parameters are positioned differently in some East Asian languages.

Keywords: topic-prominent comparatives; double-comparatum comparatives; comparee-standard mismatched comparatives; standard-topic comparatives; concrete/abstract measurement

1 Introduction

In this paper, I, as many scholars have done, confine the term ‘comparative constructions’ to comparisons of inequality, where at least two objects¹ are compared

¹ The compared “objects” dealt with in this paper are restricted to NPs, without considering VPs or clauses.

***Corresponding author: Chenlei Zhou** [tʃou⁵⁵ tʃhən³⁵ lei²¹⁴], Institute of Linguistics, Chinese Academy of Social Sciences, Dongcheng District, Beijing, China, E-mail: zhouchenlei@126.com. <https://orcid.org/0000-0002-4067-9105>

with respect to a certain property. For example, in (1), the height of *John* versus *Peter* is compared:

- (1) *John is taller than Peter.*

Cross-linguistically, comparative constructions typically involve four elements: the comparee (COMPEE), the standard (ST), the marker of standard (M), and the parameter (P) that encodes the compared property.² In (1), COMPEE = *John*, ST = *Peter*, M = *than*, and P = *tall*.³

There are two remarkable aspects in a prototypical comparative construction like (1). Firstly, a comparative construction typically involves two distinct compared objects: a COMPEE and a ST. This is clearly demonstrated by Stassen's (1985: 24) definition of comparative constructions:

[A comparative] construction has the semantic function of assigning a graded (i.e. non-identical) position on a predicative scale to two (possibly complex) objects.

Stassen further elaborates:

Thus, a case of comparison of inequality minimally involves three things: a gradable predicative scale, ... and two concepts, one of which represents the standard against which the other is measured and found to be unequal.

In Dixon (2012: 343), similar expressions are also found:

The prototypical comparative scheme in a grammar involves comparing two participants in terms of the degree of some gradable property relating to them ...

In East Asian languages, however, a third compared object commonly occurs. See (2) from Modern Standard Chinese (MSM):

- (2) 大象鼻子比其他动物长。
dàxiàng bízi bǐ qítā dòngwù cháng.
 elephant nose_M other animal long
 NP₁ NP₂ [NP₃]
 'The elephant's nose is longer than other animals.'

There are three NPs in (2),⁴ and two of them, 大象 *dàxiàng* 'elephant' and 鼻子 *bízi* 'nose' are comparee-like. This phenomenon has not received attention within the

² The P is usually realized as adjectives and sometimes as verbs (Dixon 2012).

³ Some comparative constructions include the "degree marker" (Heine 1997) or "index" (Dixon 2012), which corresponds to the *-er* in *taller* in (1) (and other instances of *more* as in e.g., *more expensive*). This kind of marker falls outside the scope of this paper.

⁴ Note that *dàxiàng* and *bízi* do not form a single possessive NP due to the absence of the possessive marker 的 *de*. Admittedly, *de* can be omitted in certain possessive constructions, such as 我爸爸 *wǒ*

existing typological literature. In this paper, the NPs analogous to *dàxiàng* and *bízi* are referred to as “comparatum” (COMPM) and “attribute comparatum” (ATTR-COMPM), respectively.⁵ And the comparatives like (2) are termed as “double-comparatum comparatives” (double-COMPM comparatives). See the definition of COMPM and ATTR-COMPM in Section 3.1 and the discussion on double-COMPM comparatives in Sections 3.2–3.4.

The second aspect that is noticeable in a typical comparative construction pertains to the relationship between the COMPEE and the ST. The COMPEE and the ST form a subject-predicate structure with P, as exemplified in *John/Peter is tall* when considering (1). Accordingly, we can claim that the COMPEE and the ST are aligned in the sense that both bear a direct semantic or argument relationship with the P. It is worth noting that many languages do not permit the “COMPEE-ST mismatched” comparatives. For instance, the sentence *John’s hands are bigger than Peter* is not permitted in English when the intended meaning is ‘John’s hands are bigger than Peter’s hands.’. While COMPEE-P is acceptable, as in *John’s hands are big*, ST-P is deemed invalid in this context, as in **Peter is big*. This generalization, however, does not hold when examining East Asian languages, as they exhibit different patterns and allow for more flexibility. Consider (3) from MSM:

- (3) 她的头发比我长。
tā de tóufa bǐ wǒ cháng.
 3 POSS hair M 1 long
 [COMPEE] M ST P
 ‘Her hair is longer than mine.’ (lit. ‘Her hair is longer than me.’)

In (3), while COMPEE-P is acceptable (她的头发长 *tā de tóufa cháng* ‘her hair is long’), the ST fails to establish the same subject-predicate relation with the P (我长 *wǒ cháng* ‘I am long’). Thus, a mismatch occurs between the COMPEE and the ST, as only the former maintains a direct semantic relation with the P. The semantic mismatch between COMPEE and ST is manifested syntactically: while one of them functions as a possessive NP, the other functions as a non-possessive NP, as evidenced by (3).⁶

bàba 1 father ‘my father’. However, in (2), *dàxiàng bízi* is not an NP that omits *de*. This is evidenced by two points, as claimed by Liu (2012): first, a pause and a topic marker (such as 啊 *a*) can be inserted after *dàxiàng*, as in 大象啊, 鼻子 ... *dàxiàng a, bízi* ... ‘as for elephants, the nose ...’; second, *bízi* can be swapped with *dàxiàng*, as in 鼻子 (啊,) 大象 ... *bízi (a) dàxiàng* ... ‘As for nose, elephants ...’. These indicate that *dàxiàng* and *bízi* do not constitute a NP.

5 The term “comparatum” is employed as a synonym for “comparee” in certain literature. However, as these terms never appear together within the same literature, as noted by a reviewer, this paper adopts this synonymous pair and establishes a technical differentiation between them.

6 It can also be the case where COMPEE is a non-possessive NP and ST is a possessive NP. See Section 4.

Section 4 will focus on examining comparatives like (3), which are referred to as “COMPEE-ST mismatched comparatives”.

Based on comprehensive observations of both double-COMPM comparatives and COMPEE-ST mismatched comparatives, which do not fit into any established typological framework (see Section 2), as well as the exploration of “ST-topic comparatives” discussed in Section 5 (in which ST is the topic in the construction “ST, COMPEE P”), this paper presents a novel typology of comparative constructions termed “topic-prominent comparatives.” Furthermore, in Section 6, I present a new pair of parameters in comparative constructions, known as the “concrete measurement” and the “abstract measurement” of comparative degrees. Our findings indicate that these two types of measurement exhibit divergent patterns in their respective locations in some East Asian languages.

2 Previous typological framework for comparative constructions

This section introduces previous typological frameworks for comparative constructions, showing that they do not encompass the topic-prominent comparatives proposed in this paper. I therefore believe that this study can contribute to theoretical frameworks, cross-linguistic comparisons and language description in the realm of comparative constructions.

In his influential study, Stassen classifies comparative constructions into two main types: derived-case comparatives and fixed-case comparatives. Derived-case comparatives involve the marking of the ST being determined or derived from the marking of the COMPEE. In contrast, fixed-case comparatives mark the ST with a specific case form, irrespective of the marking of the COMPEE. Furthermore, Stassen establishes several subtypes of derived-case and fixed-case comparatives, based mainly on the particular type of case form employed (see Table 1).

Table 1: The primary classification of comparative constructions in Stassen (1985).^a

	Subtype	Type	Main performance
Type 1	Separative	Fixed-case	ST encoded by an element indicating source ('from')
Type 2	Allative	Fixed-case	ST encoded by an element indicating goal ('to', 'for')
Type 3	Locative	Fixed-case	ST encoded by an element indicating contact ('on')
Type 4	Exceed	Fixed-case	ST encoded by a verb meaning 'to exceed/surpass'
Type 5	Conjoined	Derived-case	Two clauses, one has COMPEE and the other ST

^aIn addition to the primary (sub)types, Stassen (1985) also addresses some marginal comparative types, such as “mixed cases” and “particle comparatives”. These types are classified based on the same criteria, although they present challenges in fitting neatly into the primary categories.

Stassen’s typological classification of comparative constructions is primarily based on, as he specifically notes, “the encoding of the standard NP in comparatives” (p. 28).

Heine’s (1997: 111) framework posits that “comparative markers tend to be derived from other, more concrete entities.” However, Heine’s focus extends beyond the mere *M*, or, the encoding of the *ST*. Instead, he considers the overall event schema manifested in comparative constructions. He classifies comparative constructions into eight schema categories, as depicted in Table 2.

Table 2: Comparative construction schemas in Heine (1997: 112).

Source schema ^a	Label of schema
<i>X is Y surpasses Z</i>	Action
<i>X is Y at Z</i>	Location
<i>X is Y from Z</i>	Source
<i>X is Y to Z</i>	Goal
<i>X is Y, Z is not Y</i>	Polarity
<i>X is Y, then Z</i>	Sequence
<i>X is Y (like) Z</i>	Similarity
<i>X and Z, X is Y</i>	Topic

^aX = COMPEE, Y = P, Z = ST.

Note that the “topic schema” in Heine’s framework bears resemblance to the *ST*-topic comparatives discussed in this paper. If *ST*-topic comparatives are incorporated into the topic schema using Heine’s symbolic system, they can be expressed as “Z, X is Y.” This has previously been explored by scholars such as Chappell (2015). However, the other two types of comparatives examined in this paper, double-COMPM comparatives and COMPEE-ST mismatched comparatives, cannot be subsumed into existing schemas.

Dixon (2012) classifies comparative constructions based primarily on the grammatical function of the *P*, as shown in Table 3.

Table 3: Types of comparative constructions in Dixon (2012).

Type	Main performance
Type A	P as copula complement or as verbless clause complement
Type B	P as adjective or as intransitive verb
Type C	P as post-predicate constituent
Type D	P as head of both A and O NPs in a transitive clause
Type E	P as head of the predicate

This paper presents a novel approach to classifying comparative constructions. The classification primarily concerns around the relationships between the *COMPEE*, *COMPM*, *ATTR-COMPM*, *ST*, and the *P*, along with specific characteristics of those elements. See Table 4.

Table 4: Types of topic prominent comparatives.

Type	Main performance
Double-comparatum comparatives	<i>COMPEE</i> is divided into <i>COMPM</i> and <i>ATTR-COMPM</i> . <i>COMPM</i> is compared to <i>ST</i> , while <i>ATTR-COMPM</i> directly relates to <i>P</i> in terms of semantics or argument structure. <i>COMPM</i> and <i>ST</i> , on the other hand, do not share such a relationship with <i>P</i> .
Comparee-standard mismatched comparatives	Either <i>COMPEE</i> or <i>ST</i> functions as a possessive NP, while the other serves as a non-possessive NP; only the possessive NP has a direct semantic or argument relationship with <i>P</i> .
Standard-topic comparatives	<i>ST</i> appears at the topic/initial position of the clause.

It is worth noting that the frameworks mentioned, including the one adopted in this paper, exhibit overlapping characteristics rather than being mutually exclusive. That is, a given comparative clause can be simultaneously categorized as, for instance, Stassen’s Type 1, Heine’s source schema, Dixon’s Type B, and the double-*COMPM* comparatives proposed within this study.

3 Double-comparatum comparatives

3.1 Definition

A double-comparatum (double-*COMPM*) comparative construction involves two *COMPEE*-like NPs, namely the comparatum (*COMPM*) and the attribute comparatum (*ATTR-COMPM*). The *COMPM* refers to the NP that is compared to the *ST*. In contrast, the *ATTR-COMPM* is the NP that possesses a specific attribute, often described by an adjective or more specifically, by *P* in the comparative construction. In many languages, the two concepts are expressed by the same NP, the *COMPEE*. For example, in the English sentence *The price of this book is higher than that of another*, the *COMPM* is *the price of this book*, which is compared to the *ST* *that of another*. Meanwhile, *The price of this book* is also the *ATTR-COMPM* since it is being described by the adjective *high*. In this case, the *COMPEE* can be treated as the combination of the *COMPM* and the *ATTR-COMPM*. In MSM, however,

the two concepts are distinguishable,⁷ as illustrated by (2) above. In essence, by considering the degree of relatedness, the COMP_{PM}, ATTR-COMP_{PM}, ST, and P can be grouped into two pairs: COMP_{PM}-ST and [ATTR-COMP_{PM}]-P.

It should be emphasized that the conceptual distinction between COMP_{PM} and ATTR-COMP_{PM} is not an original contribution of this paper. When discussing “比”字差比句 ‘*bǐ-comparatives*’ in MSM,⁸ Liu (2012) identifies two types of COMPEE-like NPs, which he labels as 比较主体 *bǐjiào zhǔtǐ* ‘compared subject’ and 属性主体 *shǔxìng zhǔtǐ* ‘attribute subject’. These two types of *zhǔtǐ* can co-occur in what he terms as 双主体差比句 *shuāng zhǔtǐ chābǐjù* ‘double-subject comparatives.’ However, the use of “subject” to translate both 主体 *zhǔtǐ* ‘main body’ and 主语 *zhǔyǔ* ‘(grammatical) subject’ can lead to ambiguity in English. While *zhǔtǐ* ‘main body’ refers to the central or principal entity, it may not necessarily correspond to the grammatical subject of a sentence. In order to avoid this confusion, this study adopts the terms “comparatum” and “attribute comparatum” to represent Liu’s ‘compared subject’ and ‘attribute subject’, respectively. Consequently, Liu’s ‘double-subject comparatives’ are denoted as “double-comparatum comparatives” in this research.

In double-COMP_{PM} comparatives, the position of COMP_{PM} and ATTR-COMP_{PM} can vary. As shown in (2), the COMP_{PM} can appear before the ATTR-COMP_{PM}. Alternatively, the ATTR-COMP_{PM} may precede the COMP_{PM} (4a) or immediately precede the P (4b). Example (4c) is an atypical case in which the COMP_{PM} and ST have a mismatch: The former is a non-possessive NP, while the latter is a possessive NP. These “COMP_{PM}-ST mismatched comparatives” are considered not only a subtype of double-COMP_{PM} comparative constructions, but also a subtype of the “COMPEE-ST mismatched comparatives”, which will be discussed in Section 4.

- (4) a. 价格这家店比那家店贵。
 jiàgé zhè jiā diàn bǐ nà jiā diàn guì.
 price this CL store M that CL store expensive
 ATTR-COMP_{PM} [COMP_{PM}] M [ST] P
 (lit. ‘As for price, this store is higher than that store.’)
- b. 这家店比那家店价格贵。
 zhè jiā diàn bǐ nà jiā diàn jiàgé guì.
 this CL store M that CL store price expensive
 [COMP_{PM}] M [ST] ATTR-COMP_{PM} P
 (lit. ‘This store, price is higher than that store’)

7 Although they do not necessarily separate, as in 张三比李四高 *Zhāngsān bǐ Lǐsì gāo* (Zhangsan M Lisi tall) ‘Zhangsan is taller than Lisi.’

8 *Bǐ*-comparative construction is the principle comparative constructions of inequality in MSM. Its basic structure is “COMPEE *bǐ*-ST P” where COMPEE can be further divided into COMP_{PM} and ATTR-COMP_{PM}.

- c. 价格这家店比那家店的贵。

jiàgé zhè jiā diàn bǐ nà jiā diàn de guì.
 price this CL store M that CL store POSS expensive
 ATTR-COMPM [COMPM] M [ST] P
 (lit. ‘As for price, this store is higher than that store’s.’)
 ‘The price in this store is higher than the price in that store.’

It should be noted that the constructions like (5) are not double-COMPM comparatives.

- (5) 价格这家店的比那家店的贵。

jiàgé zhè jiā diàn de bǐ nà jiā diàn de guì.
 price this CL store POSS M that CL store POSS expensive
 ‘This store’s price is higher than that store’s.’

Example (5) is a typical MSM *bǐ*-comparative construction (with a headless possessive NP). In this sentence, 这家店的 *zhè jiā diàn de* ‘this store’s’ is a headless possessive NP, referring exactly to “(this store’s) price”. Accordingly, *zhè jiā diàn de* and *jiàgé* are not differentiated as two types of comparatums, and the sentence is thus not a double-COMPM comparative construction discussed in this paper.

COMPM and ATTR-COMPM typically exhibit a semantic possessive relationship in which the COMPM is the possessor and the ATTR-COMPM is the possessee. This can be illustrated via adding the possessive marker (e.g., *de* in MSM) between COMPM and ATTR-COMPM. However, as indicated in footnote 4, it is important to note that they do not form a syntactic possessive relationship nor do COMPM-[ATTR-COMPM] function as a single NP.

3.2 The topicality of double-COMPM comparatives

“Topic” is a pragmatic concept that pertains to the theme or subject matter of a sentence or discourse. Certain languages, referred to as “topic-prominent” languages (Li and Thompson 1976, 1981), have morpho-syntactically marked grammatical categories for sentence topics. For instance, MSM (as well as other Sinitic varieties,⁹ especially Wu) is commonly recognized as topic-prominent, where topic is considered a grammatical notion (see, among others, Shi 2000; Xu and Liu 2007); Similarly, Japanese and Korean employ dedicated topic markers that are distinct from subject markers. The definition of the topic as a descriptive category (Haspelmath 2010) varies across languages, with its status as a grammatical notion not being the focus of this paper. Instead, this study employs “topic” as a comparative concept (Haspelmath

⁹ Most typologists consider these Sinitic varieties to be sufficiently distinct to be called languages, but within China there is a specific tradition using the label ‘Chinese dialects’.

2010) which is defined as: A topic of a sentence is what the sentence is ABOUT or “set up to add new information to” (Shi 2000: 386).

From the definition, two notable characteristics of topic emerge that are relevant to the discussion in this paper. Firstly, a sentence can have more than one topic. Secondly, a topic does not necessarily correspond to the (grammatical) subject (Lambrecht 1994). These characteristics can be illustrated through example (6) from MSM.

- (6) a. 大象鼻子长。
dàxiàng bízi cháng.
 elephant nose long
 (i) ‘As for elephants, their noses are long.’
 (ii) ‘As for elephants, speaking of nose, it is long.’
 b. 那场火幸亏消防队来得快。(Li and Thompson 1981: 96)
nèi chǎng huǒ xìngkuī xiāofángduì lái de kuài.
 that CL fire fortunate fire.brigade come COMP fast
 ‘Speaking of that fire, fortunately the fire brigade came quickly.’

In terms of the number of topics, (6a) has two interpretations. Reading (i) contains one topic, 大象 *dàxiàng* ‘elephant’, and the remainder of the sentence constitutes the comment that adds new information to the topic; In reading (ii), there are two topics: *dàxiàng* ‘elephant’ serves as the primary topic, and 鼻子 *bízi* ‘nose’ the secondary topic (Xu and Liu 2007). As noted above, a sentence can contain more than two topics; however, this aspect is beyond the scope of this paper. Example (6b) shows that a topic does not necessarily align with the grammatical subject of a sentence. In this example, the topic 那场火 *nèi chǎng huǒ* ‘that fire’ lacks a clear thematic role in relation to the predicate 来得快 *lái de kuài* ‘come quickly’, as their semantic relationship is distant. Rather, a subtle “pragmatic relation” (Lambrecht 1994) exists between the topic ‘that fire’ and the comment 幸亏消防队来得快 ‘fortunately the fire brigade came quickly’.

On the one hand, it is noteworthy that an NP can fulfill dual roles as both the topic and the subject. For instance, in the sentence *John went to school*, the NP *John* serves as the subject and can also assume the role of the topic when the sentence is about *John*. On the other hand, as a sentence typically has only one subject, in sentences with two subject-like NPs preceding the predicate, often only one of the NP functions as the subject, while the other assumes the role of a topic. Xu and Liu (2007) present an operational definition that delineates this distinction, positing that within sentences featuring two subject-like NPs, other things being equal, the NP that bears a direct semantic or argument relationship with the predicate tends to be deemed as the subject, whereas the NP lacking such a relationship tends to be analyzed as the topic. Accordingly, *nèi chǎng huǒ* ‘that fire’ in (6b) is analyzed as the topic of the

sentence, while *xiāofāngduì* ‘fire brigade’ is the subject of the predicate *lái de kuài* ‘come quickly’.¹⁰ Considering that there are additional factors influencing topic recognition, such as position (initial or non-initial) and definiteness (definite or indefinite), it is justifiable to argue that, within the scope of this study, in a sentence containing two or more subject-like NPs, the NP(s) lacking a direct semantic or argument relationship with the predicate demonstrate(s) a significant level of topicality.

Liu (2012) provides evidence that in double-COMPM *bǐ*-comparatives in MSM, the ATTR-COMPM has a direct semantic or argument relationship with the P, and the COMPM commonly has no such relationship. In (4a), for example, [ATTR-COMPM]-P is acceptable as in 价格贵 *jiàgé guì* ‘the price is expensive’, whereas the COMPM, 这家店 *zhè jiā diàn* ‘this store’, lacks a direct semantic or argument relationship with *guì* ‘expensive’. The COMPM thus exhibits a considerable degree of topicality. Furthermore, if the COMPM appears in the initial position of the sentence, as in the case of *zhè jiā diàn* ‘this store’ in (4b), its topicality is further enhanced.

Notably, Liu (2012) observes that in *bǐ*-comparatives, the ST may not necessarily exhibit a direct semantic or argument relationship with the P, as seen in examples like 其他动物 *qítā dòngwù* ‘other animals’ and 长 *cháng* ‘long’ in (2), as well as 那家店 *nà jiā diàn* ‘that store’ and 贵 *guì* ‘expensive’ in (4). This is because the ST is used to establish comparisons with the COMPM. When the COMPM lacks a direct semantic or argument relationship with the P, the ST performs the same. The fact that the ST can function independently of a semantic or argument relationship with the P reveals its topicality, despite occurring in a non-typical position for a topic. This is because the ST only has a pragmatic relation to the core predicates in such cases.

3.3 Double-COMPM comparatives in Sinitic varieties

The Chinese linguistic community has long held the belief that while there are significant phonological differences between Sinitic varieties, there is only one universal Chinese grammar (Chao 1968). However, recent research has demonstrated that the grammatical differences between Sinitic varieties are greater than previously thought (Chappell et al. 2007; Dryer 2003; Hashimoto 1976; Szeto and Yurayong 2021; Szeto et al. 2018).

¹⁰ The situation in (6a) is more complex. When the sentence is interpreted as “‘elephant’_{topic}+ [‘nose’-‘long’]_{comment}”, ‘elephant’ assumes the role of both the topic and the subject of the entire sentence, while ‘nose’ is viewed as a “small” subject within the comment. However, in an alternative interpretation as “‘elephant’_{topic1}+‘nose’_{topic2}+‘long’_{predicate}”, ‘nose’ tends to be analyzed as the subject of the sentence (despite also serving as the secondary topic), and ‘elephant’ is the (primary) topic.

Comparative constructions, as a part of the Chinese grammatical system, exhibit a rich diversity. Some studies have been conducted on this topic. Dryer (2003) points out that the word order of comparative constructions in MSM blends the characteristics of both SOV and SVO languages. He provides examples of the *bǐ*-comparatives. See (7).

- (7) Modern Standard Mandarin (Dryer 2003: 49)

Zhāngsān bǐ tā pàng.

Zhangsan _M 3 fat

‘Zhangsan is fatter than him/her.’

MSM has SVO as its basic word order, for which Predicate-Standard NP is predicted according to word order universals (Dryer 2003), but it adopts the Standard-Predicate order typical of SOV languages.¹¹

Chappell and Peyraube (2015) note, however, that in addition to *bǐ*-comparatives, there is another type of comparative construction known as the “surpass” comparatives in Sinitic varieties, which, with the _{P-ST} order, perfectly matches the word order of VO languages. For example:

- (8) Cantonese (Hong Kong): Yue (Chappell and Peyraube 2015: 138)

ngóh pa lóuhsyú dō gwo néih.

1 fear mice more _M 2

‘I’m even more afraid of mice than you.’

The marker of standard *gwo* in (8) has the meaning of “surpass,” thus this type of comparative construction is referred to as the “surpass” comparatives in contrast to the “compare” comparatives represented by *bǐ*-comparatives.

The core difference between the “compare” and “surpass” comparatives lies in their word order and the marking pattern of the _M (dependent-marking vs. head-marking), and the marker of the standard is not limited to *bǐ* and *gwo* and their counterparts in various varieties. See (9).

- (9) a. The “compare” comparatives: CMPEE-[_{M-ST}]_{dependent-marking-P}
(where _M=比 *bǐ*, 赶 *gǎn*, 跟 *gēn*, 照 *zhào* ...)
b. The “surpass” comparatives: CMPEE-[_{P-M}]_{head-marking-ST}
(where _M=过 *guò*, 咖 *kā*, 啊 *a*, 起 *qǐ* ...)

Further insights into the usage, distribution, and diachronic origins of the markers of standard in Sinitic varieties are also found in Cao (2008) and Zheng (2012).

The presence of another comparative construction in Sinitic varieties, in which the _{MS} are post-positioned case markers, has received less attention. This type of

¹¹ Thanks to the reviewer who suggested this expression.

comparative construction, termed as “case-marking” comparatives for convenience,¹² primarily exists in the Gansu-Qinghai linguistic area in northwest China,¹³ where long-term and intensive contact has occurred between Mandarin Chinese, Tibetan and Altaic languages. See (10).

(10) a. Zhoutun (my fieldwork)

ŋɿ tʂaɕi xa kɔ tʂɿ kɿ.

1 Zhaxi M tall PROG PART

‘I am taller than Zhaxi.’

b. Tangwang (Xu 2017: 88)

ma ly ɕiɛ khue.

horse donkey M fast

‘The horse runs faster than the donkey.’

The comparative marker *xa* in (10a) is a widely distributed dative-accusative marker in the Gansu-Qinghai linguistic area (Zhou 2020), which serves as a comparative marker in this context. It is likely derived from the Mandarin Chinese locative particle 下 *ɕia* ‘down’ (Zhou 2019a, 2019b). The M *ɕiɛ* in (10b) is an ablative marker meaning ‘from’. The use of case markers in the Gansu-Qinghai varieties¹⁴ to mark the ST is a result of language contact with nearby Amdo Tibetan and Altaic languages. Under the influence of language contact, some Gansu-Qinghai varieties have also developed the rare dual comparative marking of dative-ablative (Zhou 2020).

Previous studies have explored the diversity of comparative constructions in Sinitic varieties, but a universal feature has been overlooked: the prevalence of the double-COMP_{PM} comparative construction. This construction is found in multiple groups within Chinese and used in different types of comparative constructions, including the “compare”, “surpass”, and “case-marking” comparatives.

The use of double-COMP_{PM} comparatives is explicitly mentioned in the work of Xia (2020), Sheng (2021), Wang (2022), and Zhou (2022, 2023a).

¹² Based on the source of the case markers, the “case-marking” comparatives in this paper can be grouped into different categories in e.g., Stassen (1985) and Heine (1997).

¹³ This linguistic area is situated in the west of Gansu, the east of Qinghai and the bordering areas between the two provinces. It is an area where several Sinitic varieties, Amdo Tibetan and Altaic languages (including Mongolic and Turkic groups) have undergone prolonged and intensive contact. Scholars have referred to this linguistic area using various terms, such as “Qinghai-Gansu Sprachbund” and “Amdo Sprachbund” (Zhou 2020).

¹⁴ In this paper, Gansu-Qinghai varieties refer to the Sinitic varieties in the Gansu-Qinghai linguistic area that have changed their basic word order to SOV, and developed a case marking system.

- (11) a. Yiyang: Xiang (Xia 2020: 551; pers. comm.)

阿价钱我比尔便宜些。

*a*⁴⁵ *tɕia*⁴⁵*lie*³⁴ *ŋo*⁴² *pi*²¹ *n*²¹ *pie*²¹³*ni*³⁴ *ɕia*³⁴.

DEF price 1 M 2 cheap a.little

[ATTR-COMPM] COMPM M ST P

‘The price of mine is a little cheaper than yours.’

- b. 年龄我大咖尔一些。

*nie*¹³*lin*³³ *ŋo*⁴² *tɕæ*²¹ *ka*⁴⁵ *n*²¹ *i*⁴⁵*ɕia*³³.

age 1 big M 2 a.little

ATTR-COMPM COMPM P M ST

‘I am a little older than you.’

- (12) Shaoxing: Wu (Sheng 2021: 377)

钞票渠比我阿哥多。

chaopiao *qu* *bi* *wo* *age* *duo*.¹⁵

money 3 M 1 elder.brother more

ATTR-COMPM COMPM M [ST] P

‘His/her has more money than my elder brother.’

- (13) a. Shenyang: Mandarin (Wang 2022: 232–3)

价钱这种水果比那种水果便宜。

jiaqian *zhe* *zhong* *shuiguo* *bi* *na* *zhong* *shuiguo* *pianyi*.

price this CL fruit M that CL fruit cheap

ATTR-COMPM [COMPM] M [ST] P

‘The price of this kind of fruit is cheaper than that kind.’

- b. 市内药都照郊区便宜。

shinei *yao* *dou* *zhao* *jiaoqu* *pianyi*.

urban medicine all M suburban cheap

COMPM ATTR-COMPM M ST P

‘The medicine is cheaper in the city than in the suburbs.’

- (14) Zhoutun: Unclassified (Zhou 2022: 123)

我你啊岁数大着多。

ŋɿ *nia* *suifu* *ta* *tɕɿ* *tur*.

1 2:M age big COMP much

COMPM ST-M ATTR-COMPM P

‘I am much older than you.’

15 The examples only containing Chinese characters are presented in Pinyin transcription without tone diacritics in this paper.

These data suggest that double-COMPM comparatives are not limited to any specific group or type of comparative construction (i.e., “compare,” “surpass,” and “case-marking” comparatives).

To provide further evidence for the prevalence of double-COMPM comparatives in Sinitic varieties, this study conducted a corpus analysis of 25 varieties from the Corpus of Chinese Dialect Grammar (CCDG) (Liu and Xia 2023).¹⁶ The CCDG consists of 711 test sentences from 22 grammatical categories, including comparative constructions, for each variety, and covers the ten major variety groups of Chinese as well as several unclassified varieties. The double-COMPM comparatives test sentence used in this study is (15), where the COMPM is 你 *nǐ* ‘you’ and 我 *wǒ* ‘I’ in the two clauses, respectively, and the ATTR-COMPM is 样式 *yàngshì* ‘style’ and 价钱 *jiàqián* ‘price’, respectively.¹⁷ The test sentences were recorded by CCDG contributors in the local variety to capture the most natural expressions of each variety.

- (15) 衣服样式你比我好看些, 价钱我比你便宜些。
yīfu yàngshì nǐ bǐ wǒ hǎokàn xie,
 cloth style 2 M 1 good-looking a.little
 ATTR-COMPM COMPM M ST P
jiàqián wǒ bǐ nǐ piányi xie.
 price 1 M 2 cheap a.little
 ATTR-COMPM COMPM M ST P
 ‘The design of your clothes is better than mine, and the price of mine is cheaper than yours.’

The results are presented in Table 5.

If the test sentence (15) is documented in the CCDG as a double-COMPM comparative construction, it is categorized as “tested” in Table 5. Conversely, if (15) is recorded as a “non-double-COMPM” (i.e., “single-COMPEE”) comparative construction, it is “not tested”. Of the 25 varieties examined in Table 5, the double-COMPM comparative construction was tested in 14, which accounts for 56 % of the sample. These varieties represent the ten major groups within Chinese, with seven of them (Mandarin, Yue, Wu, Gan, Xiang, Jin, and Pinghua) exhibiting evidence of double-COMPM comparatives.

¹⁶ The website of the database is <http://www.dialectgrammar.com>, last accessed on 3/20/2023. Note that the database is continuously updated, and more data may be added in the future.

¹⁷ Note that in (15), 衣服 *yīfu* ‘cloth’ is the topic of the whole sentence and not the COMPM or ATTR-COMPM of each comparative clause.

Table 5: The usage of the double-COMPM comparatives in varieties in the CCDG.

	Variety	Double-COMPM comp	M	Type	Contributor
1	Junxian: Mandarin	Not tested	<i>pi</i>	Compare	Xin (2023)
2	Lishu: Mandarin	Not tested	<i>pi</i>	Compare	F. Cao (2023)
3	Zichuan: Mandarin	Tested	<i>pi</i>	Compare	Sun (2023)
4	Deyang: Mandarin	Not tested	<i>pi</i>	Compare	Yang (2023)
5	Lanzhou: Mandarin	Tested	<i>pi</i>	Compare	Yi (2023)
6	Shenyang: Mandarin	Tested	<i>pi/tʂau</i>	Compare	Wang (2023)
7	Guangzhou: Yue	Tested	<i>kuɔ</i>	Surpass	Zheng (2023)
8	Lishui: Wu	Tested	<i>pi</i>	Compare	Sheng (2023a)
9	Shanghai: Wu	Tested	<i>pi</i>	Compare	Zhu (2023)
10	Linhai: Wu	Tested	<i>pi</i>	Compare	Lu (2023)
11	Rui'an: Wu	Tested	<i>de</i>	Compare	Y. Wu (2023)
12	Shaoxing: Wu	Tested	<i>pi</i>	Compare	Sheng (2023b)
13	Duchang: Gan	Tested	<i>pi</i>	Compare	L. Cao (2023)
14	Ji'an: Gan	Not tested	<i>pi</i>	Compare	Chang (2023)
15	Shanyin: Jin	Tested	<i>pi</i>	Compare	Guo (2023)
16	Linfen: Jin	Not tested	<i>pi</i>	Compare	Yan (2023)
17	Binyang: Pinghua	Tested	<i>ko</i>	Surpass	Qin (2023)
18	Dong'an: Xiang	Tested	<i>bie</i>	Compare	Hu (2023)
19	Shaoyang: Xiang	Not tested	<i>pi</i>	Compare	Jiang (2023)
20	Meixian: Hakka	Not tested	<i>pi ... kuo</i>	Mixed	Hou (2023)
21	Ningde: Min	Not tested	<i>pi</i>	Compare	L. Chen (2023)
22	Zhangzhou: Min	Not tested	<i>pi</i>	Compare	Lin (2023)
23	Qimen: Hui	Not tested	<i>pi</i>	Compare	Y. Chen (2023)
24	Zhanmi: Unclassified	Not tested	<i>pi</i>	Compare	F. Wu (2023)
25	Zhoutun: Unclassified	Tested	<i>xa</i>	Compare	Zhou (2023b)

The CCDG provides only a snapshot of the grammar of each variety, and the natural expressions recorded in the database may not exhaustively represent all grammatical structures of each variety. Therefore, the absence of double-COMPM comparative construction in the CCDG for a particular variety does not necessarily imply that it is absent in that variety. For instance, even though the CCDG data do not verify the use of double-COMPM comparatives in Qimen, the respondent has reported their occurrence in (16).

- (16) Qimen: Hui (Yao Chen, pers. comm.)
- 我手比尔大。
- a^{42} sen^{42} pi^{33} n^{11} tha^{33} .
- 1 hand M 2 big
- COMPM ATTR-COMPM M ST P
- ‘My hands are bigger than yours.’

Example (16) provides evidence for the existence of double-COMPM comparatives in the Hui group. Further investigation reveals that the other two major groups, i.e., Hakka and Min, also use double-COMPM comparatives, see (17)–(18), respectively.

- (17) Ninghua: Hakka (Zhang 2004: 259)

佢今年比旧年老掉好多。

kə⁴⁴ kiŋ⁴⁴ŋieŋ²⁴ pik³¹ əu42ŋieŋ²⁴ lau³¹ thiaŋ⁵ hau³¹to⁴⁴.

3 this.year M last.year old COMP much

ATTR-COMPM COMPM M ST P

‘S/he is much older this year than last year’

- (18) Hui’an: Min (Chen 2020: 317)

今年比舊年風恰透。

ken¹-lin² pi³⁻² ku⁵⁻⁴-lin² huaŋ¹ khaŋ⁷⁻⁸ thiaŋ⁵.

this.year M old.year wind more strong

COMPM M ST ATTR-COMPM P

‘The wind of this year is stronger than that of last year.’

The current findings present compelling evidence for the extensive distribution of the double-COMPM comparatives in Sinitic varieties. The construction has been identified in all ten major groups within Chinese, along with several unclassified varieties, and is utilized in a variety of comparative types, namely the “compare,” “surpass,” and “case-marking” comparatives.

3.4 Double-COMPM comparatives in East Asian languages

The double-COMPM comparatives are also found in numerous East Asian languages.¹⁸ This includes Tibeto-Burman languages (19–24), Zhuang-Dong languages (25–28), Turkic languages (29), as well as Japanese (30) and Korean (31).

- (19) Amdo Tibetan (Shao 2012: 38)

ŋigormo cç^ho-a-wtina ŋa noŋ-ngu,

money 2-DAT-M 1 less-AUX

ATTR-COMPM ST-M COMPM P

x^hiwea cç^ho-a-wtina ŋa maŋ-ngu.

knowledge 2-DAT-M 1 more-AUX

ATTR-COMPM ST-M COMPM P

‘My money is less than yours; my knowledge is more than yours.’

¹⁸ The term “East Asian languages” as used in this paper refers to the languages spoken in China, Japan, and Korea.

- (20) Yidu: Lhoba (Li 2008: 56)
e³¹tia⁵⁵nu⁵⁵ e⁵⁵ja⁵⁵nu⁵⁵ mi³³ a³¹tha⁵⁵pra⁵⁵ pa⁵⁵ tsa⁵⁵.
 this.year last.year M grain more PART
 COMPM ST M ATTR-COMPM P
 ‘There is more grain this year than last.’
- (21) Luobozhai: Qiangic (Wang and Dong 2022: 43)
pu⁵⁵su⁵⁵(-mu⁵⁵) na³¹(-mu⁵⁵) qa³¹-so⁵⁵ko⁵⁵ a³¹-pu³¹ ba³¹(-na³¹).
 age(-TOP) 2(-TOP) 1-M one-year big(-2:IMPF)
 ATTR-COMPM COMPM ST-M P
 ‘You are one year older than me.’
- (22) Mebzang nDrapa: Qiangic (Yang Huang, pers. comm.)
ε^ha³³mu⁵⁵=mε⁵⁵hki³³, so⁵⁵ηo⁵⁵=ma³³ε^hy³³ stui⁵⁵ηo⁵⁵ xo⁵⁵ tεa⁵⁵-tεθ³³-zε³³.
 matsutake=TOP Yinshan=M Yangshan more have-IMPV-GNO
 ATTR-COMPM ST-M COMPM P
 ‘Yinshan has more matsutake than Yangshan.’
- (23) Jino: Yi (Zheng 2012: 61)
tshə⁴⁴mjə⁴⁴ jə³³ e⁴⁴mjə⁴⁴ e⁴²fui⁴² la⁵⁵pha⁴² ʃə⁴² ja⁴⁴ a.
 this.year M last.year rainfall less COMP more PART
 COMPM M ST ATTR-COMPM P
 ‘This year’s rainfall is much lower than last year’s.’
- (24) Bangduo Lahu: Yi (Li 2012: 365)
jə⁵³ ηa³¹ ke³⁵ ɔ³³ya⁵³ ɣ³¹.
 3 1 more strength big
 COMPM ST ATTR-COMPM P
 ‘He is stronger than me.’ (lit. ‘He has more strength than me.’)
- (25) Lachi: Gelao (Li 2008: 199)
pi³⁵ne⁴⁴ ze³⁵ a³⁵ vua³⁵ na⁴⁴nu³⁵.
 this.year rice good M last.year
 COMPM ATTR-COMPM P M ST
 ‘This year’s rice is better than last year’s.’
- (26) Pubiao: Gelao (Li 2008: 200)
pje⁵³ mje¹³na⁵⁵ pi⁵⁵ mje¹³kun⁵³ ai⁵³.
 rice this.year M last.year good
 ATTR-COMPM COMPM M ST P
 ‘This year’s rice is better than last year’s.’

- (27) Tailue: Zhuang-Tai (Yan 2018: 357)

ʔa⁴¹ju³³ man⁴¹ nɔi¹¹ lə⁵⁵ ta:n³³.

age 3 little M 2

ATTR-COMPM COMPM P M ST

‘He is younger than you.’

- (28) Xia’ao: Zhuang-Tai (Wei 2012: 187)

nəŋ⁴² tu⁴² ɛiaŋ¹³ reɪ²³¹ wa:i⁴² tu¹³ mu⁴².

nose CL elephant long M CL pig

ATTR-COMPM [COMPM] P M [ST]

‘The elephant’s nose is longer than the pig’s.’

- (29) Salar: Turkic (Zheng 2012: 62)

asman gedze-dən buyun jaxfia.

sky yesterday-M today good

ATTR-COMPM ST-M COMPM P

‘Today’s weather is better than yesterday’s.’

- (30) Japanese (from Hiroyuki Suzuki, pers. comm.)

zou=wa hana=ga hoka=no doobutsu=yori naga-i.

elephant=TOP nose=NOM other=POSS animal=M long-AUX

COMPM ATTR-COMPM [ST-M] P

‘The nose of an elephant is longer than that of any other animal.’

Note that (30) represents an exceptional case of double-COMPM comparatives where a mismatch occurs between the COMPM and the ST. In this case, the COMPM is a non-possessive NP, while the ST is a possessive NP. Consequently, similar to example (4c), (30) can also be classified as a COMPM-ST mismatched comparative construction, which falls under the subtype of COMPEE-ST mismatched comparatives discussed in Section 4.

- (31) Korean (from Ian Joo, pers. comm.)

kokkiri-ka tal-un tongmwul-pota ko-ka te kil-ta.

elephant=NOM different-ADJ animal-M nose-NOM more long-DEC

COMPM [ST-M] ATTR-COMPM P

‘The nose of elephant is longer than that of any other animal.’

4 Comparee-standard mismatched comparatives

4.1 Definition

A comparee-standard (COMPEE-ST) mismatched comparative refers to a comparative construction that exhibits a “seemingly logical inconsistency” (Zhu 1982: 214) due to a

mismatch between the COMPEE and the ST. This construction usually involves two NPs, one of which is a possessive NP while the other is a non-possessive NP. The key feature of this construction is the syntactic mismatch that arises when the non-possessive NP semantically refers to a possessive NP. The non-possessive NP lacks a direct semantic or argument relation with the P. An example of this construction is given in (32) from MSM:

- (32) a. 大象的鼻子比其他动物长。
dàxiàng de bízi bǐ qítā dòngwu cháng.
 elephant POSS nose M other animal long
 [COMPEE] M [ST] P
 ‘The nose of elephant is longer than that of any other animal.’
 (lit. ‘THE ELEPHANT’S NOSE is longer than ANY OTHER ANIMAL.’)
- b. 他比我的玩具多。
tā bǐ wǒ de wánjù duō.
 3 M 1 POSS toy more
 COMPEE M [ST] P
 ‘He has more toys than me.’
 (lit. ‘HE is more than MY TOYS.’)

In (32a), the possessive NP 大象的鼻子 *dàxiàng de bízi* ‘the elephant’s nose’ is syntactically mismatched with the non-possessive ST 其他动物 *qítā dòngwu* ‘other animals.’ Similarly, in (32b), the COMPEE becomes a non-possessive NP 他 *tā* ‘he’, while the ST 我的玩具 *wǒ de wánjù* ‘my toys’ is possessive. Although the non-possessive NPs *qítā dòngwu* ‘other animals’ and *tā* ‘he’ formally correspond to the possessors *dàxiàng* ‘elephant’ and *wǒ* ‘I’, respectively, they semantically refer to the nose of ‘other animals’ and the toys of ‘him’.

These non-possessive NPs lack a direct semantic or argument relation with P, as evidenced by the unacceptability of constructions like *其他动物长 *qítā dòngwu cháng* ‘other animals are long’¹⁹ and *他多 *tā duō* ‘he is more’. This feature highlights the topicality of the COMPEE-ST mismatched comparatives.

Owing to the scarcity of data, this paper refrains from delving into the intricacies of the COMPM-ST mismatched comparatives referenced above, which are a subtype of COMPEE-ST mismatched comparatives where the COMPEE is divided into a COMPM and an ATTR-COMPM.

4.2 COMPEE-ST mismatched comparatives in Sinitic varieties

Zhou (2023a) mentions the existence of COMPEE-ST mismatched comparatives in Zhoutun, a Sinitic variety which is greatly influenced by Amdo Tibetan in the Gansu-Qinghai linguistic area:

¹⁹ The sentence itself is grammatically correct, but it is unacceptable in this particular context.

- (33) Zhoutun: Unclassified (Zhou 2023a: 4)

你的鞋我啊一号大着哩。

nɪ tʃ xɛ ŋa i xɔ ta tʃɿ li.
 2 POSS shoe 1:M one size big PROG PART
 [COMPEE] ST-M P

‘Your shoes are one size bigger than mine.’

(lit. ‘YOUR SHOES are one size bigger than ME.’)

This construction is also found in several other varieties.

- (34) Shenyang: Mandarin (Wang 2022: 233)

你的水平照人家比差远了。

nɪ de shuiping zhao renjia bi cha yuan le.
 2 POSS level M other.people compare low far PART
 [COMPEE] M ST P

‘Your level is far lower than others.’

(lit. YOUR LEVEL is far lower than OTHER PEOPLE.)

- (35) Qilian: Northwest Mandarin (Zhang 2016: 5)

李四的个子张三哈不到啊。

Lisi de gezi Zhangsan ha bu dao a.
 Lisi POSS height Zhangsan M NEG reach PART
 [COMPEE] ST M [P]

‘Lisi is not as tall as Zhangsan.’

(lit. LISI’S HEIGHT is not as tall as ZHANGSAN)

- (36) Liancheng: Hakka (Xiang 1997: 425)

二班个成绩比一班较差。

ŋ¹¹ pa³³ ə³⁵ ʂaŋ⁵⁵ tʂi³⁵ pɿ⁵¹ i³⁵ pa³³ kuə³³ tʂu³³.
 two class POSS score M one class more bad
 [COMPEE] M [ST] P

‘The performance of the second class is even worse than that of the first class.’

(lit. THE SECOND CLASS’S SCORE is even worse than THE FIRST CLASS.)

There has been little attention given to COMPEE-ST mismatched comparatives in grammatical descriptions.²⁰ As a result, explicit descriptions of this construction in

²⁰ Numerous studies have focused on the “asymmetry” phenomenon in comparatives of MSM, with notable discussion by Cui (2014). However, when examining descriptive grammars of Sinitic varieties and other East Asian languages, there is a noticeable scarcity of information on this particular topic.

literature are scarce, with isolated instances found only sporadically. To systematically investigate this issue, the present study turns to the CCDG, which includes a test sentence (37) featuring an inferior comparative construction with a COMPEE-ST mismatched interpretation.

- (37) 我的头发没你 (的) 长。
wǒ de tóufa méi nǐ (de) cháng.
 1 POSS hair NEG 2 POSS long
 [COMPEE] [ST] P
 ‘My hair is not as long as yours.’

Note that whether or not (37) is a COMPEE-ST mismatched construction is determined by the presence of the possessive marker 的 *de* following 你 *nǐ* ‘you’: in cases where *de* is present, the sentence does not belong to the category of COMPEE-ST mismatched comparatives. This is due to the syntactic-semantic match between the ST 你的 (头发) *nǐ de (tóufa)* ‘your (hair)’ and the COMPEE 我的头发 *wǒ de tóufa* ‘my hair’. Conversely, when the sentence lacks the possessive marker *de* after *nǐ*, it is considered a COMPEE-ST mismatched comparative construction: the COMPEE *wǒ de tóufa* ‘my hair’ is syntactically incongruent with the ST *nǐ* ‘you’.

Table 6 shows that 12 out of 25 varieties in the CCDG have been confirmed to use COMPEE-ST mismatched comparatives, making up 48 % of the sample. This construction has been observed in seven of the ten major groups, namely Mandarin, Yue, Wu, Jin, Xiang, Hakka and Hui. While there is currently no confirmation of this construction in the remaining three major groups (Gan, Pinghua and Min), it should be noted that the lack of evidence in the CCDG does not necessarily indicate its absence. For instance, although the CCDG data does not confirm the use of COMPEE-ST mismatched comparatives in the Duchang variety, the contributor has reported its presence in (38).

- (38) Duchang: Gan (Linlin Cao, pers. comm.)
 渠格书多似我。
gɛ¹¹³ kɔ⁰ su³³² tɕ³³² sɿ⁰ ɲɔ³⁵².
 3 POSS book many M 1
 [COMPEE] P M ST
 ‘S/he has more books than I do.’
 (lit. ‘HIS/HER BOOKS are more than ME.’)

Table 6: The usage of the COMPEE-ST mismatched comparatives in varieties in the CCDG.

	Variety	COMPEE-ST comp	Contributor
1	Junxian: Mandarin	Not tested	Xin (2023)
2	Lishu: Mandarin	Not tested	F. Cao (2023)
3	Zichuan: Mandarin	Not tested	Sun (2023)
4	Deyang: Mandarin	Not tested	Yang (2023)
5	Lanzhou: Mandarin	Not tested	Yi (2023)
6	Shenyang: Mandarin	Tested	Wang (2023)
7	Guangzhou: Yue	Tested	Zheng (2023)
8	Lishui: Wu	Tested	Sheng (2023a)
9	Shanghai: Wu	Not tested	Zhu (2023)
10	Linhai: Wu	Tested	Lu (2023)
11	Rui'an: Wu	Tested	Y. Wu (2023)
12	Shaoxing: Wu	Not tested	Sheng (2023b)
13	Duchang: Gan	Not tested	L. Cao (2023)
14	Ji'an: Gan	Not tested	Chang (2023)
15	Shanyin: Jin	Tested	Guo (2023)
16	Linfen: Jin	Tested	Yan (2023)
17	Binyang: Pinghua	Not tested	Qin (2023)
18	Dong'an: Xiang	Tested	Hu (2023)
19	Shaoyang: Xiang	Not tested	Jiang (2023)
20	Meixian: Hakka	Tested	Hou (2023)
21	Ningde: Min	Not tested	L. Chen (2023)
22	Zhangzhou: Min	Not tested	Lin (2023)
23	Qimen: Hui	Tested	Y. Chen (2023)
24	Zhanmi: Unclassified	Tested	F. Wu (2023)
25	Zhoutun: Unclassified	Tested	Zhou (2023b)

The COMPEE-ST mismatched comparatives have also been observed in two other groups, namely Pinghua (39) and Min (40).

- (39) Liangjiang: Pinghua (Xiao 2020: 112)
今日的天色比昨日婁。
kiæ³¹nie⁵ ti³³ thin³¹fo⁵ pæ³³ tshiu³¹nie⁵ wæ³⁵.
today POSS weather M yesterday good
[COMPEE] M ST P
‘The weather is better today than it was yesterday.’
(lit. TODAY’S WEATHER is better than YESTERDAY.)

- (40) Jian'ou: Min (Jiang 2015: 65)
 你的年龄邀我妹仔差不多。
ni de nianling yao wo meizi cha bu duo.
 2 POSS age M 1 younger.sister different NEG much
 [COMPEE] M [ST] [P]
 'You are about the same age as my younger sister.'
 (lit. 'YOUR AGE is about the same as MY YOUNGER SISTER.)

4.3 COMPEE-ST comparatives in East Asian languages

Although the concept of COMPEE-ST mismatched comparatives is not yet widely recognized, and no systematic discussion of such constructions can be found in the literature, examples of this phenomenon have nevertheless been identified in a number of East Asian languages or linguistic groups, including Tibeto-Burman (41–44), Zhuang-Dong (45–47), Korean (48), and Japanese (49).

- (41) Lhasa Tibetan (Li 2008: 55)
khoŋ⁵⁵ ki lɛ⁷¹²ka⁵⁴ miɛ̃¹²ta¹² lɛ ka jo⁷¹² re⁷¹².
 3 POSS work other.people M comparatively good COP
 [COMPEE] ST M P
 'His work is better than others.'
 (lit. HIS WORK is better than OTHER PEOPLE.)
- (42) Bangduo Lahu: Yi (Li 2012: 362)
ɛi³³sua³³pa⁵³na³¹ ve³³ mu⁵³je³¹ ja⁵³ ɲa³¹xu³³ mu⁵³mi³¹
 Xishuangbanna POSS summer M 1:PL place
 [COMPEE] M [ST]
kɛ³⁵ xo³³ a³¹.
 much hot PART
 P
 'Summer in Xishuangbanna is hotter than our hometown.'
 (lit. XISHUANGBANNA'S SUMMER is hotter than OUR HOMETOWN.)
- (43) Khato: Yi (Zhao 2009: 197)
za³¹ni³³ ku³³ zi³¹ni³³ tha³³ yu³¹.
 today POSS yesterday M big
 [COMPEE] ST M P
 'Today's is bigger than yesterday's.'
 (lit. TODAY'S is bigger than YESTERDAY.)

Note that in (43) the COMPEE is a headless possessive NP.

- (44) Kucong: Yi (Chang 2009: 171)
tɕhi³¹ ni³³ khɔ³¹ yu³³ tsa³¹ a³¹ mi³¹ khɔ³¹ tha³¹ na³³.
 this.year POSS rice last.year M good
 [COMPEE] ST M P
 ‘This year’s rice is better than last year’s.’
 (lit. THIS YEAR’S RICE is better than LAST YEAR.)

- (45) Xia’ao: Zhuang-Tai (Wei 2012: 191)
mə:ŋ²³¹ ti³³ ko:ŋ³³ fən³³ pi⁴² te⁴² la:i⁴².
 2 POSS work.point M 3 more
 [COMPEE] M ST P
 ‘Your work points are more than his/hers.’
 (lit. YOUR WORK POINTS are more than s/HE.)

- (46) Maonan: Dong-Sui (Li 2008: 124)
m¹be¹ta⁶ ti⁰ fu⁴ pi³ m¹be¹kun⁵ coŋ².
 last.year POSS rice M year.before.last more
 [COMPEE] M ST P
 ‘There is more rice last year than the year before.’
 (lit. LAST YEAR’S RICE is more than THE YEAR BEFORE.)

- (47) Lachi: Gelao (Li 2008: 199)
vu⁴⁴ ki⁵⁵ to³¹ sei³⁵ laŋ⁵⁵ la⁴⁴ m⁵⁵.
 bowl 1 big M CL POSS 2
 [ATTR-COMPM COMPM]_{COMPEE} P M [ST]
 ‘My bowl is bigger than yours.’
 (lit. As for bowl, I am bigger than YOUR BOWL.)

Example (47) is a COMPM-ST mismatched comparative construction, a subtype of COMPEE-ST mismatched comparatives; it is also a subtype of double-COMPM comparatives, since there is COMPM and ATTR-COMPM.

- (48) Korean (Lianhua Bai, pers. comm.)
kokkiri-ui ko-ka tal-un tongmwultul-pota kil-ta.
 elephant-POSS nose-NOM different-ADJ animal:PL-M long-DEC
 [COMPEE] [ST-M] P
 ‘The nose of an elephant is longer than that of other animals.’
 (lit. The ELEPHANT’S NOSE is longer than OTHER ANIMALS.)

- (49) Japanese (Akiyama and Akiyama 2002: 177)
Amerika yori, Nihon-no hō ga chiisai desu.
 America M Japan-POSS side NOM small is
 ST M [COMPEE] P
 ‘Japan is smaller than America.’
 (lit. COMPARED WITH AMERICA, JAPAN’S SIDE is small.)

Note that in Japanese, a typical comparative construction (COMPEE-ST-M-P) does not seem to allow for a mismatch between COMPEE and ST. Hiroyuki Suzuki (pers. comm.) explains that the sentence *zou-no hana-ga hoka-no doobutsu-yori naga-i* (elephant-POSS nose-NOM other-POSS animal-M long-AUX) (lit. “elephant’s nose [is] other animal(s)’ than long”) can only be interpreted as a comparison between “the length of the elephant’s nose” and “the length of other animals”. That is, it can only mean ‘the length of the elephant’s nose is longer than the length of other animals’. The mismatched reading, ‘the length of the elephant’s nose is longer than the length of other animals’ nose’, is not allowed. The COMPEE-ST mismatched construction presented in (49) (ST-M, COMPEE-P), is atypical. A reviewer highlighted that (49) cannot be readily analyzed as a case of COMPEE-ST mismatched comparative construction, as the role of *hō ga* is merely to emphasize the comparee. Notably, it is not possible to substitute *Amerika-yori* in (49) with **Amerika-no hō-yori* (‘than America’s side’). From a formal standpoint, given that the COMPEE is a non-possessive NP while the ST is a possessive NP, I tentatively classify (49) as an instance of an atypical COMPEE-ST mismatched comparative construction. However, the definitive categorization of (49) within this framework remains an open question.

5 Standard-topic comparatives

5.1 Definition

Standard-topic (ST-topic) comparatives are structured as “topic-comment,” where the ST serves as the topic in the sentence-initial position, providing a reference point for the sentence. The comment portion then compares the ST to the COMPEE. For example, in Shaoxing Wu (Yimin Sheng, pers. comm.), one might say:

- (50) 小张 (么,) 还是小王高。
ɕio³³ tsaj³³ (me?), huɛ⁷¹ ze⁵⁵ ɕio³³ huɔŋ⁵³ kɔ⁵⁵.
 Little.Z TOP still be Little.W tall
 ST COMPEE P
 ‘Little Wang is taller than Little Zhang.’

Here, ‘Little Zhang’ serves as the *ST* in the topic position (it can be followed by a topic marker and a pause), with the comment portion stating that ‘Little Wang is tall’, indicating that ‘Little Wang’ is taller than ‘Little Zhang’.

ST-topic comparatives are cross-linguistically uncommon due to the prevailing placement of *COMPEE* in the subject position (Dixon 2012): It is the *COMPEE* that is more inclined to take on a topical role, while the *ST* is typically positioned after the *COMPEE* and displays relatively lower topicality. Heine (1997) proposed a “topic schema” within his framework (see Section 2), which follows the structure “*COMPEE* and *ST*, *COMPEE* is *P*” (referred to as ‘*X* and *Z*, *X* is *Y*’). According to this schema, *COMPEE* possesses at least an equal level of topicality as *ST*. However, the *ST*-topic schema introduced in this paper is “*ST*, *COMPEE* *P*”, where *ST* assumes a significant topical role.

Note that expressions of the following kind, which resemble *ST*-topic comparatives, do not actually qualify as *ST*-topic comparatives.

(51) *As for the price, this book is more expensive.*

In this sentence, ‘price’ serves as the topic but not the *ST*, since it cannot be placed in the *ST* position of a regular comparative construction (**This book is more expensive than the price*). Thus, such structures do not qualify as *ST*-topic comparatives.

Depending on how the *ST* is marked, *ST*-topic comparatives in East Asian languages can be classified into two major types: “topic comparatives” and “‘look’ comparatives.” Topic comparatives typically indicate the *ST* through the use of a pause and/or a topic marker, or by placing the *ST* at the beginning of the sentence. Topic comparatives are primarily observed in Wu and Hui Chinese. On the other hand, ‘look’ comparatives introduce *ST* with a *M* that conveys the meaning of ‘look’ (‘look’-*M*), and are primarily found in Northwest Chinese and neighboring languages.

5.2 Topic comparatives

In a topic comparative construction, the *ST* is placed in the topic position and can be followed by a pause and/or a topic marker, as shown in (50). According to Li (2003) and Chappell (2015), this type of comparative construction is a result of conflating two conjoined clauses: “Comparing *COMPEE* with *ST*: it is still the case that *COMPEE* is *P*.” > ‘As for *ST*, *COMPEE* is *P*.’ The *ST* in this construction exhibits strong topicality.

Topic comparatives are mainly found in a small number of Wu and Hui varieties (Chappell 2015; Li 2003). The data in (52) provide several examples with subtle differences.

- (52) a. Fuyang: Wu (Chappell 2015: 42)
 法国, 还是中国大。
faʔ⁵¹ kuoʔ⁵¹, huaʔ²³ zɿ²¹³ tɕion⁵³ kuoʔ⁵¹ du³¹³.
 France still be China big
 ST 'still' 'be' COMPEE P
 'Compared with France, China is bigger.'
- b. Tiantai: Wu (Dai 1999: 258)
 小王是小李长。
ɕiau³¹ hũɔ¹³ zɿ ɕiau³¹ li³¹ dzian³⁵.
 Little.W be Little.L tall
 ST 'be' COMPEE P
 'Little Li is taller than Little Wang.'
- c. Wuyi: Wu (Li 2003: 225)
 小姑还大姑好。
xiaogu hai dagu hao.
 younger. aunt still elder. aunt good
 ST 'still' COMPEE P
 'The elder aunt is better than the younger aunt.'
- d. Jixi: Hui (Li 2003: 225)
 渠尔高些。
qu er gao xie.
 3 2 tall a.little
 ST COMPEE P
 'You are a little taller than him.'

Based on the occurrence of 'still' and 'be', *ST*-topic comparatives in Wu and Hui Chinese can be categorized into four subtypes: (i) *ST* (TOP) 'still' 'be' COMPEE P; (ii) *ST* 'be' COMPEE P; (iii) *ST* 'still' COMPEE P; and (iv) *ST* COMPEE P, exemplified in (52a–d) respectively. The synchronic distribution of these four subtypes across varieties provides valuable insights into the developmental trace of topic comparatives in Wu and Hui varieties. While lacking a marker of standard,²¹ *ST*-topic comparatives are grammatical comparative constructions due to their fixed nature, conveying the exclusive meaning of “COMPEE is more P than *ST*” but not the “*ST* is more P than COMPEE”.

²¹ L. Li (2003) categorizes the usage of 'be' in comparatives such as (52a–b) as the “comparative connective marker”. However, Li does not provide any analysis on the function of 'still' in (52a) and (52c). Regardless of the specific terminologies employed, it is unlikely that 'be' or 'still' can be analyzed as the marker of standard (M), because (50) and (52a) clearly show that 'still', 'be' do not attach to the *ST*.

Wang et al. (2018) report that the use of topic comparatives is also attested in other Wu varieties, including Hangzhou, Wenzhou, Tiantai, Jinhua, Suichang, Xuancheng, Shanghai, Danyang, and Ningbo. Topic comparatives can also be found in some Hui varieties such as Jixi, Shexian, Tunxi, Xiuning, Qimen, and Yixian (Li 2003).

Wang and Dong (2022: 46) assert that Luobozhai Qiang has topic comparatives, as in (53).

- (53) *nəi* *lipa-soko-mu* *qa* *thau*.
2:POSS hand-M-TOP 1:POSS long:EVIDENTIAL
[ST-M] COMPEE P
‘My hands are longer than yours.’
(lit. ‘Your hands, mine are longer.’)

According to their analysis, the strong topicality of the *st* is indicated by the mandatory use of the topic marker when it is placed at the beginning of the sentence.

5.3 'Look' comparatives

The 'look' comparative construction utilizes a marker of standard that conveys the meaning of 'look' ('look'-M). The structure of this construction is 'Looking at _{ST}, COMPEE is P' > 'As for _{ST}, COMPEE is P', where the _{ST} exhibits topicality.

'Look' comparatives are commonly found in the northwest region of China. Such structures can be observed in Amdo Tibetan (54).

- (54) Amdo Tibetan (Sandman and Simon 2016: 112)
- | | | | |
|-----------------|---------------|--------------|----------------|
| <i>lhasa-‘a</i> | <i>ϕti-na</i> | <i>səlan</i> | <i>ʈʰe-gi.</i> |
| Lhasa-DAT | look-COND | Xining | big-TEST |
| ST | ‘look’-M | COMPEE | P |
- ‘Xining is bigger than Lhasa.’

The comparative meaning of (54) comes from the literal reading ‘If looking at Lhasa, Xining is big’ > ‘As for Lhasa, Xining is big’. Shao (2012) demonstrates the occurrence of ‘look’ comparatives in classic Tibetan, as exemplified by (55).

- (55) Amdo Tibetan (Shao 2012: 30)
da ro-las-bltasna ni namtshun,
 DEM corpse-M-look TOP weak
 [ST+‘look’-M] P
rma-las-bltasna phogsna tehe-zing-mtshis-na.
 wound-M-look scar big-AUX-AUX-have
 ST+‘look’-M COMPEE P
 ‘(I am) weaker than a corpse; the scar is bigger than the wound.’ (*Dunhuang Historical Documents of Tubo*, P.T. 1287 P1.566)

Shao (2012) argues that the prevalent use of ‘look’ comparatives in modern Amdo Tibetan is a remnant of classic Tibetan. The early context in which the ‘look’-_M is formed highlights the topicality of the _{ST}, which appears at the clause-initial position (and can be marked by a topic marker), setting the scene for the following comment.

Zhoutun also exhibits ‘look’ comparatives, as in (56).

- (56) Zhoutun: Unclassified (Zhou 2023a: 6)

你们的房子看时, 我们的房子大。

ni mɤ tɤ fãtsi khã ʃi ŋɤ mɤ tɤ fãtsi ta.

2 PL POSS house look COND 1 PL POSS house big

[ST] ‘look’-_M [COMPEE] P

‘Our house is bigger than yours.’ (lit. ‘If looking at your house, our house is big’).

‘Look’ comparatives are also seen in the Linxia variety (Sinitic; Zhang 2016: 5) and Kyrgyz (Turkic; Zheng 2012) in northwest China.

Sandman and Simon (2016: 112) have pointed out that both Salar and Wutun, which are spoken in the Gansu-Qinghai linguistic area, have ‘look’-_M comparatives similar to those found in Amdo Tibetan. However, the examples provided by the authors (57) do not show the _{ST} in the topic position. Due to a lack of data, it is currently unknown whether _{ST} can be placed in the topic position in the synchronic or diachronic data of these languages. Based solely on example (57), these constructions cannot be classified as ‘look’ comparatives termed in this paper.

- (57) a. Salar

biqirox jiguo elige ɟʒan-aŋ vaq-sə da aɣər-a ro.

cloth all that.way life-2POSS look-COND too heavy-TEST INT

[COMPEE] ST ‘look’ P

‘Are all such clothes weightier (i.e. more important) than your life?’

- b. Wutun

je-ge jjhakai zhungo kan-la xaige ga-li.

this-REF country China look-COND very small-SEN-INF

[COMPEE] ST ‘look’ P

‘This country is much smaller than China.’

Likewise, the comparatives in Korean with *pota* (*po-* ‘look’) as the _M do not qualify as ‘look’ comparatives, as the _{ST} does not occupy the topic position (which is taken by the COMPEE). See (58).

- (58) Korean (Rhee 2022: 3)

John-un Bob-pota (te) ttokttokha-ta.

John-TOP Bob-M more be.smart-DEC

COMPEE ST-‘look’ P

‘John is smarter than Bob.’

(lit. ‘John, after seeing Bob, is smarter.’)

Contrary to Salar and Wutun, there is ample diachronic evidence that the *pota*-comparatives in Korean have consistently had a COMPEE-ST-*pota* word order during their grammaticalization process (Rhee 2022). Therefore, it can be determined that *pota*-comparatives do not qualify as ‘look’ comparatives.²²

Sections 3 and 4 of this paper introduced double-COMPM comparatives, COMPEE-ST mismatched comparatives, and ST-topic comparatives. All of these types can be classified as topic-prominent comparatives due to their display of certain topic-related properties. In double-COMPM comparatives and COMPEE-ST mismatched comparatives, there is no direct semantic or argument relation between COMPEE (COMPM) and/or ST and P, but rather a pragmatic connection. In ST-topic comparatives, the ST serves as the topic, forming a “topic-comment” type of topic structure. These characteristics contrast with “subject-prominent comparatives”, as represented by English comparative constructions, where both COMPEE and ST have direct semantic relations with P, COMPEE does not distinguish between COMPM and ATTR-COMPM, and ST does not serve as a topic element.

6 A pair of novel parameters: concrete measurement and abstract measurement

This section introduces a novel pair of parameters in comparative constructions, namely, concrete measurement and abstract measurement of comparative degrees. Concrete measurement involves the use of precise numerical values or quantifiable units to make comparisons, as exemplified by *three years* in *I am three years older than you*. In contrast, abstract measurement involves the application of non-specific, relative terms to establish comparisons, as illustrated by *much* in *I am much older than you*.

In most languages, both concrete and abstract measurements of comparative degrees are located on the same side of P, as exemplified by the use of ‘three years’ and ‘much’ in the previous examples. This may account for the lack of differentiation between these two parameters in prior research. Nonetheless, this study reveals that in certain East Asian languages, these two types of measurements may be positioned on different sides of the P.

²² A reviewer pointed out that, despite the ST in (58) not occupying the topic position, its marker and the ‘look’-M in what we refer to as ‘look’ comparatives are both derived from ‘look,’ suggesting a potential generalization between them. While I find the comparison between these two types of comparatives utilizing the ‘look’-M construction intriguing, it is important to note that, for the scope of this paper, (58) cannot be classified as a ST-topic comparative. This is due to the fact that the topic position is occupied by COMPEE rather than ST.

Let us first take a look at MSM. While in most cases both concrete measurement (59a) and abstract measurement (59b) are placed to the right of *P*, the adverb 远 *yuǎn* ‘far’ or 略 *lüè* ‘slightly’ denoting abstract measurement is located to the left of the *P* (59c).

- (59) a. 他比我高一米。
tā bǐ wǒ gāo yī mǐ.
 3 M 1 tall one meter
 COMPEE M ST P [concrete measurement]
 ‘He is one meter taller than me.’
- b. 他比我高很多。
tā bǐ wǒ gāo hěn duō.
 3 M 1 tall very much
 COMPEE M ST P [abstract measurement]
 ‘He is much taller than me.’
- c. 他远/略高于我。
tā yuǎn / lüè gāo yú wǒ.
 3 far slightly tall M 1
 COMPEE [abstract measurement] P M ST
 ‘He is far/slightly taller than me.’

The positioning of abstract measurement on the left side of the *P* in MSM is a rare occurrence, and the expression “*yuǎn/lüè*-*P*-*M_{yu}*-*ST*” could be considered archaic in style. However, in the varieties examined below, it is consistently observed that abstract measurement and concrete measurement are placed on opposite sides of *P*.

According to Zhou (2023a), the placement of concrete measurement and abstract measurement in Zhoutun varies depending on the type of comparative constructions being used. In “case-marking” comparatives, concrete measurement appears before *P* (60a), whereas abstract measurement is positioned after *P* (60b). On the other hand, in “compare” comparatives, abstract measurement comes after *P* (61a), while concrete measurement can either appear after (61b) or before *P* (61c).

- (60) Zhoutun: Unclassified (Zhou 2023a: 4)
- a. 安文栋连珺哈三岁大着哩。
ānwēntǒng liǎncǐhǎ sā suì ta tǝʂ li.
 A L M three year old PROG PART
 COMPEE ST M [concrete measurement] P
 ‘Anwendong is three years older than Lianjun.’
- b. 我你啊岁数大着多。
ŋɿ nia suifu ta tǝʂ tux.
 1 2:M age old COMP much
 COMPM ST-M ATTR-COMPM P abstract measurement
 ‘I am much older than you.’

(61) Zhoutun: Unclassified (Zhou 2023a: 5)

- a. 我比你大着多。

ŋɤ pi ni ta tɕɤ tux.

1 M 2 big COMP much

COMPEE M ST P abstract measurement

‘I am much bigger (taller) than you.’

- b. 我比你大一岁。

ŋɤ pi ni ta i sui.

1 M 2 old one year

COMPEE M ST P [concrete measurement]

‘I am one year older than you.’

- c. 老王比老张头一个高着个。

luã pi lɔtã thu i kɤ kɔ tɕɤ kɤ.

old.W M old.Z head one CL tall PROG PART

COMPEE M ST [concrete measurement] P

‘Old Wang is one head taller than Old Zhang.’

Li (2012) observes that in Bangduo Lahu: Yi, concrete measurement is positioned before P (62a) while abstract measurement appears after P (62b).

(62) Bangduo Lahu: Yi (Li 2012: 363–4)

- a.
- jɔ⁵³ ŋa³¹ tha³¹ te⁵³ qhɔ³¹ ur³¹ ta³¹.*

3 1 ACC one year old PROG

COMPEE ST [concrete measurement] P

‘He is one year older than me.’

- b.
- ɔ³¹vi³⁵pa11 ɔ³¹ni³³pa¹¹ ke³⁵ tshu³³ dza⁵³.*

elder.brother younger.brother more fat very

COMPEE ST P abstract measurement

‘The elder brother is much fatter than the younger one.’

In Luobozhai: Qiangnic, concrete measurement is placed before P (63a), and abstract measurement “mainly appears before the P (63b), and in some instances, it can be placed after the P” (63c). (Wang and Dong 2022: 46).

(63) Luobozhai: Qiangnic (Wang and Dong 2022: 46)

- a.
- nə³¹ tha³¹-so⁵⁵ko⁵⁵ ɕe⁵⁵-pu³¹ ɕtɕau³¹.*

2 3-M three-year young: EVIDENTIAL

COMPEE ST-M concrete measurement P

‘You are three years younger than him.’

- b.
- tha³¹ qa³¹-so⁵⁵ko⁵⁵ tɕe³¹ qə³¹mi⁵⁵ ja³¹.*

3 1-M much thoughtful PART

COMPEE ST-M abstract measurement P

‘He is much thoughtful than me.’

- c. *tu³¹bza³¹* *thə³¹-so⁵⁵ko⁵⁵* *pe³¹εa⁵⁵* *χa³¹so⁵⁵*.
 elder.brother 3-M capable much
 COMPEE ST-M P abstract measurement
 ‘The elder brother is much more capable than he is.’

7 Conclusion

This paper proposes a new type of comparative construction, called “topic-prominent comparatives,” based on data from East Asian languages. These comparatives are characterized by prominent topicality and can be further classified into three subtypes: double-COMPM comparatives, COMPEE-ST mismatched comparatives, and ST-topic comparatives. The paper also introduces a new pair of parameters for analyzing comparative constructions: concrete measurement and abstract measurement of comparative degrees.

The concept of topic-prominent comparatives, as well as the differentiation between COMPM and ATTR-COMPM, and concrete measurement and abstract measurement, offer new perspectives for the typological investigation of comparative constructions.

Overall, the proposed framework sheds new light on the study of comparative constructions and opens up new avenues for future research. Some potential research topics in this field include, but are not limited to the following.

Firstly, it is worth investigating which other languages exhibit topic-prominent comparatives. Based on current observations, topic-prominent comparatives seem to be more prevalent in East Asian languages,²³ which may be attributed to their inherent topic-prominent nature. Further, the distribution of topic-prominent comparatives in East Asian languages and their presence in other linguistic regions, can be examined, thus building on the findings presented in this paper.²⁴

²³ While I have not conducted a large-scale survey, based on the initial review of the data and communications with linguists, few cases of topic-prominent comparatives are found in other parts of the world.

²⁴ According to Denis Creissels (pers. comm.), COMPEE-ST mismatched comparatives are seen in Mande languages, as in the following example from Mandika (Mande, Senegal):

- (i) *Kàlāa lè lá tíñáaróo wàráta tàmbóo tí.*
 pen-FOC-POSS-damage-is.big-spear-in.comparison.with
 ‘The damage caused by a pen is bigger than that caused by a spear.’
 (lit. ‘THE DAMAGE OF A PEN is bigger than A SPEAR.’)

Secondly, a more detailed investigation of topic-prominent comparatives is needed. Given that the concept of topic-prominent comparatives is relatively new, existing literature on this type of comparative construction remains limited in scope, leaving numerous questions. For example, what are the syntactic and semantic properties of topic-prominent comparatives, as well as their discourse-pragmatic functions and diachronic evolution across different languages? Can different subtypes of topic-prominent comparatives co-occur, as shown in (4c), (30) and (47) by the co-occurrence of double-COMPM comparatives and COMPEE (COMPM)-ST mismatched comparatives?

Furthermore, are there any implicational universals in relation to topic-prominent comparatives? As current observations suggest, the existence of subject-prominent comparatives can be inferred from the presence of topic-prominent comparatives. Specifically, if a language has double-COMPM comparatives, then it also has single-COMPEE comparatives. If a language has COMPEE-ST mismatched comparatives, then it also has COMPEE-ST matched comparatives. However, it remains to be further investigated whether there is an implicational relationship between double-COMPM comparatives and COMPEE-ST mismatched comparatives.

Finally, what is the reason for the different positions of concrete measurement and abstract measurement? If concrete measurement and abstract measurement in most languages are located on the same side of *P*, then it needs to be explained why they are located on opposite sides of *P* in some languages. Language contact may be one important reason. Zhou (2023a) suggests that in Zhoutun, the fronting of concrete measurement before *P* is due to the influence of Tibetan, while the postposition of abstract measurement after *P* retains the feature of Mandarin Chinese. Yet further study is needed.

Acknowledgments: My thanks go to Danqing Liu, Zhengda Tang, Liping Xia, Yimin Sheng, Yao Chen, Yang Huang, Hiroyuki Suzuki, Lianhua Bai, Ian Joo, Linlin Cao, and Denis Creissels. I would like to extend my appreciation to the anonymous reviewers for their valuable suggestions and insights, and to the editorial team of *Linguistic Typology*, especially Angela Terrill. Heartfelt thanks to Wenxuan Zhang and Tunxin Zhou for their unwavering support and encouragement. Any shortcomings in the paper are solely my responsibility.

Research funding: This research is supported by the key project of Chinese National Social Science Fund (19AYY004).

Abbreviations

- | | |
|---|------------|
| 1 | 1st person |
| 2 | 2nd person |
| 3 | 3rd person |

ADJ	adjective
AUX	auxiliary
CL	classifier
COMP	complement marker
COND	conditional
COP	copular
DAT	dative
DEC	declension
DEF	definitive
EVIDENTIAL	evidential
FOC	focus
GNO	gnomic
IMPF	imperfective
IMPV	imperfect
INF	inferential
INT	interrogative
M	marker of standard
NEG	negative
NOM	nominal
PART	particle
PL	plural
POSS	possessive
PROG	progressive
REF	referential
SEN	sensory evidence
TEST	testimonial
TOP	topic marker

References

- Akiyama, Nobuo & Carol Akiyama. 2002. *Japanese grammar*, 2nd edn. New York: Barron's Educational Series.
- Cao, Zhiyun 曹志耘 (ed.). 2008. 汉语方言地图集 [*Linguistic atlas of Chinese dialects*]. Beijing: The Commercial Press.
- Cao, Fengxia 曹凤霞. 2023. A grammatical corpus of Lishu. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Cao, Linlin 曹琳琳. 2023. A grammatical corpus of Duchang. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Chang, Junzhi 常俊之. 2009. 元江苦聪话参考语法 [*A reference grammar of Kucong*]. Beijing: Minzu University of China PhD dissertation.
- Chang, Meixiang 昌梅香. 2023. A grammatical corpus of Ji'an. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Chao, Yuen Ren. 1968. *A grammar of spoken Chinese*. Boston: Harvard University Press.

- Chappell, Hilary M. 2015. Linguistic areas in China for differential object marking, passive, and comparative constructions. In Hilary Chappell (ed.), *Diversity in Sinitic languages*, 13–52. Oxford: Oxford University Press.
- Chappell, Hilary, Li Ming & Alain Peyraube. 2007. Chinese linguistics and typology: The state of the art. *Linguistic Typology* 11(1). 187–211.
- Chappell, Hilary M. & Alain Peyraube. 2015. The comparative construction in Sinitic languages: Synchronic and diachronic variation. In Hilary Chappell (ed.), *Diversity in Sinitic languages*, 134–154. Oxford: Oxford University Press.
- Chen, Weirong. 2020. *A grammar of Southern Min: The Hui'an dialect*. Berlin: De Gruyter.
- Chen, Libing 陈丽冰. 2023. A grammatical corpus of Ningde. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Chen, Yao. 2023. A grammatical corpus of Qimen. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Cui, Weizhen 崔维真. 2014. 差比句中短语比较项隐含的不对称 [On Asymmetry of “Whole-Part” bi-sentences]. *华文教学与研究 [TCSOL Studies]* 4. 87–94.
- Dai, Zhaoming 戴昭铭. 1999. 天台话的几种语法现象 [On several grammatical phenomena in the Tiantai dialect]. *方言 [Dialects]* 4. 249–258.
- Dixon, R. M. W. 2012. *Basic linguistic theory. Further grammatical topics*. Oxford: Oxford University Press.
- Dryer, Matthew S. 2003. Word order in Sino-Tibetan languages from a typological and geographical perspective. In Thurgood Graham & Randy J. LaPolla (eds.), *The Sino-Tibetan languages*, 43–55. London: Routledge.
- Guo, Lixia 郭利霞. 2023. A grammatical corpus of Shanyin. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Hashimoto, Mantaro. 1976. Language diffusion on the Asian continent: Problems of typological diversity in Sino-Tibetan. *Computational Analysis of Asian and African Languages* 3. 49–63.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687.
- Heine, Bernd. 1997. *Cognitive foundations of grammar*. Oxford: Oxford University Press.
- Hou, Xiaoying 侯小英. 2023. A grammatical corpus of Meixian. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Hu, Chengling 胡乘玲. 2023. A grammatical corpus of Dong'an. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Jiang, Jie 江洁. 2015. 建瓯方言语法专题研究 [A Research on the Grammar of Jian'ou Dialect]. Fuzhou: Fujian Normal University MA dissertation.
- Jiang, Xiezhong 蒋协众. 2023. A grammatical corpus of Shaoyang. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Li, Charles N. & Sandra A. Thompson. 1976. Subject and topic: A new typology. In Charles Li (ed.), *Subject and topic*, 457–489. New York: Academic Press.
- Li, Charles N. & Sandra A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Li, Chunfeng 李春风. 2012. 邦朵拉枯语参考语法 [A reference grammar of Bangguo Lahu]. Beijing: Minzu University of China PhD dissertation.
- Li, Lan 李蓝. 2003. 现代汉语方言差比句的语序类型 [Word order typology of comparative constructions in Modern Chinese dialects]. *方言 [Dialects]* 3. 214–232.

- Li, Yunbing 李云兵. 2008. 中国南方民族语言语序类型研究 [*A cross-linguistic typology on word orders of minority languages in southern China*]. Beijing: Peking University Press.
- Liu, Danqing 刘丹青. 2012. 汉语差比句和话题结构的同构性: 显赫范畴的扩张力一例 [The homogeneity of comparison and topic structure in Chinese: An illustration of the expansive power of mighty categories]. *语言研究 [Studies in Language and Linguistics]* 32(4). 1–12.
- Liu, Danqing & Liping Xia 刘丹青、夏俐萍 (eds.). 2023. *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Lin, Songyu 林颂育. 2023. A grammatical corpus of Zhangzhou. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Lu, Xiaoyu 卢笑予. 2023. A grammatical corpus of Linhai. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Qin, Dongsheng 覃东生. 2023. A grammatical corpus of Binyang. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Rhee, Seongha. 2022. Comparatives in Korean: A grammaticalization perspective. *Lingua* 265. 103163.
- Sandman, Erika & Camille Simon. 2016. Tibetan as a “model language” in the Amdo Sprachbund: Evidence from Salar and Wutun. *Journal of South Asian Languages and Linguistics* 3. 85–122.
- Shao, Mingyuan 邵明园. 2012. 安多藏语的差比句 [The comparative constructions in Amdo Tibetan]. *东方语言学 [Oriental Linguistics]* 12. 27–41.
- Sheng, Yimin 盛益民. 2021. 吴语绍兴柯桥话参考语法 [*A reference grammar of Shaoxing (Keqiao)*]. Beijing: The Commercial Press.
- Sheng, Yimin 盛益民. 2023a. A grammatical corpus of Lishui. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Sheng, Yimin 盛益民. 2023b. A grammatical corpus of Shaoxing. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Shi, Dingxu. 2000. Topic and topic-comment constructions in Mandarin Chinese. *Language* 76(2). 383–408.
- Stassen, Leon. 1985. *Comparison and universal grammar*. Oxford: B. Blackwell.
- Sun, Kemin 孙克敏. 2023. A grammatical corpus of Zichuan. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Szeto, Pui Yiu, Umberto Ansaldi & Stephen Matthews. 2018. Typological variation across Mandarin dialects: An areal perspective with a quantitative approach. *Linguistic Typology* 22(2). 233–275.
- Szeto, Pui Yiu & Chingduang Yurayong. 2021. Sinitic as a typological sandwich: Revisiting the notions of Altaicization and Taicization. *Linguistic Typology* 25(3). 551–599.
- Wang, Baofeng & Yao Dong 王保锋、董瑶. 2022. 萝卜寨羌语的差比句 [Comparative constructions in the Luobozhai Qiang language]. *民族语文 [Minority Languages]* 5. 42–54.
- Wang, Huayun, Yaling Jing & Shuzhen Jiang 汪化云、荆亚玲、姜淑珍. 2018. 杭州方言的“还是”差比句 [The *haishi* comparative construction in the Hangzhou dialect]. *中国语文 [Studies of Chinese Languages]* 5. 588–592.
- Wang, Yue 王越. 2022. 沈阳方言参考语法 [*Reference Grammar of Shenyang*]. Beijing: Chinese Academy of Social Sciences. Post-Doctorate Report.
- Wang, Yue 王越. 2023. A grammatical corpus of Shenyang. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Wei, Fanmao 韦繁茂. 2012. 下坳壮语参考语法 [*A reference grammar of the Xia'ao Zhuang Language*]. Shanghai: Shanghai Normal University PhD dissertation.
- Wu, Fang 吴芳. 2023. A grammatical corpus of Zhanmi. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Wu, Yue 吴越. 2023. A grammatical corpus of Rui'an. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.

- Xia, Liping 夏俐萍. 2020. 湘语益阳（泥江口）方言参考语法 [A reference grammar of Yiyang (Nijiangkou)]. Beijing: The Commercial Press.
- Xiang, Mengbing 项梦冰. 1997. 连城客家话语法研究 [A grammatical study on the Liancheng Hakka dialect]. Beijing: Language Press.
- Xiao, Yanyun 肖雁云. 2020. 桂林两江平话语法研究 [Research on the grammar of Liangjiang dialect in Guilin]. Guilin: Guangxi Normal University MA dissertation.
- Xin, Yongfen 辛永芬. 2023. A grammatical corpus of Junxian. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Xu, Dan. 2017. *The Tangwang language*. Cham: Springer International.
- Xu, Liejong & Danqing Liu 徐烈炯、刘丹青. 2007. 话题的结构与功能（增订本）[Topic: Structural and Functional Analysis (revised version)]. Shanghai: Shanghai Educational Press.
- Yan, Wenhan 岩温罕. 2018. 西双版纳傣语参考语法 [A reference grammar of the Tailue language in Xishuangbanna]. Shanghai: Shanghai Normal University PhD dissertation.
- Yan, Yanqun 严艳群. 2023. A grammatical corpus of Linfen. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Yang, Rong 阳蓉. 2023. A grammatical corpus of Deyang. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Yi, Li 衣立. 2023. A grammatical corpus of Lanzhou. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Zhang, Tao 张桃. 2004. 宁化客家方言语法研究 [A grammatical study of Ninghua Hakka]. Xiamen: Xiamen University PhD dissertation.
- Zhang, Ansheng 张安生. 2016. 甘青河湟方言的差比句 – 类型学和接触语言学视角 [The comparative construction of He-Huang dialects in Gansu and Qinghai: Perspectives of typology and contact linguistics]. *中国语文 [Studies of the Chinese Language]* 1. 1–19.
- Zhao, Min 赵敏. 2009. 墨江哈尼族卡多话参考语法 [A reference grammar of Khato]. Beijing: Minzu University of China PhD dissertation.
- Zheng, Hui ren 郑惠仁. 2012. 东北亚语言比较标记的类型学研究 [A Cross-Linguistic Study on Comparative markers of Northeast Asian languages]. Beijing: Peking University PhD dissertation.
- Zheng, Yuan. 郑媛. 2023. A grammatical corpus of Guangzhou. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Zhou, Chenlei. 2019a. A special case marking system in the Sinitic languages of Northwest China. *Journal of Chinese Linguistics* 47(2). 425–452.
- Zhou, Chenlei 周晨磊. 2019b. 甘青方言格标记“哈”的来源 [The origin of the dative-accusative marker *xa* in Gan-Qing dialects]. *Language and Linguistics* 20(3). 493–513.
- Zhou, Chenlei. 2020. Case markers and language contact in the Gansu-Qinghai linguistic area. *Asian Languages and Linguistics* 1(1). 168–203.
- Zhou, Chenlei. 2022. *Zhoutun*. London: Routledge.
- Zhou, Chenlei. 2023a. Comparative constructions in Zhoutun from a language contact perspective. *Languages* 8(1). 66.
- Zhou, Chenlei 周晨磊. 2023b. A grammatical corpus of Zhoutun. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.
- Zhu, Dexi 朱德熙. 1982. 语法讲义 [Grammar lecture notes]. Beijing: The Commercial Press.
- Zhu, Jialei 朱佳蕾. 2023. A grammatical corpus of Shanghai. In Danqing Liu & Liping Xia (eds.), *The corpus of Chinese dialect grammar*. Available at: <http://www.dialectgrammar.com>.