Meili Liu*, Hubert Cuyckens and Fangqiong Zhan

Language change in a constructional network: the emergence of Mandarin [bi N hai N] comparative constructions

Abstract: This paper explores the mechanisms of and motivations for two unconventional comparative constructions in Mandarin: [bi N_i hai N_i] and [bi N_i hai N_j]. They are unconventional in that the item expressing the dimension along which the comparison is made is a noun rather than an adjective. It is shown that [bi N_i hai N_i] emerges (i) by analogy with the conventional comparative construction [bi N hai A] and (ii) by inheriting the nominal feature from an existing construction [Adverb N], which is corroborated by a collexeme analysis. At a more schematic level, the extension A > N observed in [bi N hai A] > [bi N_i hai N_i] may have been modeled on the existing development from [Adverb A] to [Adverb N]. Analogical extension and inheritance also underlie the subsequent development from [bi N_i hai N_i] to [bi N_i hai N_j]. This study not only shows how language changes in a constructional network (including node creation and network reconfiguration), but also sheds light on the nature of horizontal links. It also demonstrates how a synchronically perceived relation between constructions may impact a later, similar language change.

Keywords: comparative constructions; language change; constructional network; construction grammar; collostructional analysis; Mandarin Chinese

*Corresponding author: Meili Liu, School of TCM, Pharmacology Health, and Early Childhood Care, Ningbo College of Health Sciences, Ningbo, China; School of Foreign Languages, Jianghan University, Wuhan, China; and Department of Linguistics, KU Leuven, Leuven, Belgium, E-mail: meililiu@hotmail.com. https://orcid.org/0000-0002-8556-810X
Hubert Cuyckens, Department of Linguistics, KU Leuven, Leuven, Belgium,
E-mail: hubert.cuyckens@kuleuven.be. https://orcid.org/0000-0001-7489-9272
Fangqiong Zhan, National Institute of Education, Nanyang Technological University, Singapore,
E-mail: fangqiong.zhan@nie.edu.sg. https://orcid.org/0000-0002-5791-8205

1 Introduction

Making comparisons is a common function of human languages, and this is not different in Mandarin Chinese. There are two types of comparative clauses in Mandarin: (i) superiority comparatives, which are mainly marked by 比 *bi* 'than', and (ii) equality comparatives marked by 跟 *gen* 'with' (and its variants 像 *xiang*, 如 *ru*, and 又 *you*) (see Paris and Shi 2016: 297).[1] Our focus in the present paper lies with the superiority comparatives, which typically take the following shape: [bi N(oun) (hai) A(djective)] as illustrated in (1):

(1) 橡树的叶子比手掌还大。

Xiangshu	*de*	*yezi*	*bi*	*shouzhang*
Oak.tree	GEN	leaf	BI	palm
		COMPAREE	COMPARATIVE	STANDARD
			MARKER	OF COMPARISON

hai	*da*.
HAI	big
DEGREE	DIMENSION
ADVERB	OF COMPARISON

'The leaves of oak trees are even bigger than the palm of a hand.'[2]

In this [bi N hai A] construction, an entity (橡树的叶子 *xiangshu de yezi* 'the leaves of oak trees') is compared to another entity (the standard of comparison, encoded by the noun 手掌 *shouzhang* 'palm (of a hand)') with respect to a particular dimension (here encoded by the adjective 大 *da* 'big'),[3] which can be thought of as scale without absolute values (see Quirk et al. 1985: 1129). The degree adverb 还 *hai* 'still' is optional:[4] in (1), it conveys the presupposition that the palm of a hand is big, the main assertion being that the leaves of oak trees are even bigger; without *hai*, the comparative construction [bi N A] simply asserts that there is a difference in size between the leaves of oak trees and the palm of a hand.

In Mandarin, there are two unconventional noun comparative constructions: [bi N_i hai N_i] and [bi N_i hai N_j]. They are unconventional in that the dimension of comparison is expressed by a noun rather than an adjective. In the former type, this

---

1 Inferiority, then, is expressed by adding negative markers to the superiority or equality comparatives.
2 The terms 'standard of comparison' and 'dimension of comparison' are taken from Paris and Shi (2016: 297) and Li and Thompson (1981: 564). The part-of-speech category of the comparative markers *bi* and *gen* (and variants) is 'prepositions' (see Paris and Shi 2016: 297).
3 As Paris and Shi (2016: 300–301) point out, the scalar requirement can also be met by a predicate that has a direction scale (e.g., *shangyang* 'go up') or that expresses a continuous event (e.g., 'get thin').
4 A variant of the degree adverb *hai* is *geng* (see also Section 2.2).

noun is formally identical to the noun expressing the standard of comparison; in the latter, the nouns are formally different. Another idiosyncrasy of this construction lies in that the nouns sanctioned in this construction can only take the form of bare nouns and they cannot be modified by adjectives, classifiers, or demonstratives. Semantically, the nouns in the two constructions are different as well: while the first noun – expressing the standard of comparison – denotes an entity, the second noun – expressing the dimension of comparison – metonymically refers to the stereotypical attributes of that entity. Accordingly, sentence (2) conveys that the comparee (my heart) outdoes the standard of comparison (the sun) with regard to the dimension along which the comparison is made (the sun's stereotypical attributes, i.e., sunniness); in other words, my heart has even more sun-like characteristics than the sun. Similarly, in (3), with N_j differing from N_i, the comparee outdoes the standard of comparison (idiots) with respect to genius-like characteristics.

(2)   我这心里<u>比阳光还阳光</u>!
        *Wo   zhe   xinli   **<u>bi</u>** **<u>yangguang</u>** **<u>hai</u>** **<u>yangguang!</u>*** (BCC[5])
        I   this heart.in BI   sunshine   HAI   sunshine
        'My heart is even more sunny than the sun!'

(3)   真是天才, <u>比白痴还天才</u>!
        Zhen   shi   tiancai, <u>bi baichi</u>   <u>hai</u>   <u>tiancai</u>.
        Really   is   genius, BI idiot   HAI   genius
        'What a genius, an even better genius than the idiots.'

Whereas conventional comparative constructions have received considerable attention during the past few decades, so far only a small number of studies have been published on the topic of the non-conventional comparative constructions [bi N_i hai N_i] and [bi N_i hai N_j]. Previous studies (e.g., Cheng 2013; Li 2016; Yang 2011) on these non-conventional constructions have generally put more emphasis on their constructional meaning as well as the semantic features of their constituents, while only touching on their emergence and development (and the mechanisms at work therein). Typically as well, they have only discussed a few, often constructed examples; rarely have earlier studies been corpus-based, using quantitative methods (see Sections 3.2 and 4.3).

---

5 BCC refers to the corpus of the Beijing Language and Culture University Corpus Centre, including nearly 15 billion Chinese characters, and the website is as follows: http://bcc.blcu.edu.cn/. All the examples in the present article are taken from BCC, except historical data which were retrieved from the Centre for Chinese Linguistics Corpus (henceforth CCL, Peking University) (Zhan et al. 2003), which spans about 3,000 years (1000 BCE – present). More detailed information on BCC is provided in Section 3.2.

Couched within a Diachronic Construction Grammar (DCxG) framework, the present study examines the emergence of the Mandarin constructions [bi N_i hai N_i] and [bi N_i hai N_j] and explores their underlying mechanisms and motivations, thereby also drawing on a quantitative analysis of corpus data. We propose that at the origin of [bi N_i hai N_i] lie multiple sources: [bi N hai A], [Adverb N], and the development [Adverb A] > [Adverb N]. More specifically, it is argued that [bi N_i hai N_i] emerges by analogy with the conventional comparative construction [bi N hai A], whereby the adoption of the scalar meaning of A by a N, which is in essence non-scalar, can be viewed as coercion (Lauwers and Willems 2011). What motivates this coercion is the metonymic relation that holds between the entity denoted by the noun and the entity's attributes. What may further have facilitated the development from [bi N hai A] to [bi N_i hai N_i] is that the dimension of comparison in the latter construction inherited its nominal encoding from an already existing construction [Adverb N]; in this inheritance relation, the component "hai $N_{i/j}$"[6] of [bi N_i hai $N_{i/j}$] can be regarded as an instantiation of the more schematic [Adverb N]. A collexeme analysis we carried out indicates that the nouns most significantly attracted to [Adverb N] and to [bi N_i hai N_i] show semantic similarity. At a more schematic level, the extension A > N observed in [bi N hai A] > [bi N_i hai N_i] may have been modeled (formally and semantically) on the existing development from [Adverb A] to [Adverb N] in Mandarin; this is a shift-cum-coercion from scalar Adjective > Noun when preceded by an adverb denoting intensification (e.g., 太 *tai* 'too', 很 *hen* 'very'). Analogical extension and inheritance also underlie the subsequent development from [bi N_i hai N_i] to [bi N_i hai N_j] and indirectly also from [bi N hai A] to [bi N_i hai N_j].

This paper is structured as follows. Section 2 sets the background for the study, in which we first situate our investigation against the backdrop of previous research (Section 2.1); we then introduce the constructions from which [bi N_i hai N_i] and [bi N_i hai N_j] will be argued to derive: [Adverb N] (and its source [Adverb A]) and [bi N_i hai A] (Section 2.2); next is a brief introduction to the theoretical framework, i.e., (Diachronic) Construction Grammar (Section 2.3); finally, we will attend to the distinction between nouns and adjectives in Mandarin (Section 2.4). In Section 3, we examine the mechanisms of and motivations for the emergence of [bi N_i hai N_i], and provide corpus-based quantitative support. Section 4 explores its further development, viz., the emergence of [bi N_i hai N_j]. Section 5 discusses the implications of our study for (Diachronic) CxG. Section 6 presents a number of concluding remarks.

---

6 It is impossible for "hai N" to serve as an analogical model for the emergence of [bi N hai N] as it is not a standalone construction and the meaning of "hai N" in non-comparative constructions is different from that in comparative constructions. We will illustrate this point in Section 2.2 and Section 3.1 (ii).

2 Background to the study

2.1 Previous research

Previous research has mainly centered on three issues. One is the constructional meaning of [bi N_i hai N_i]. In this regard, Yang (2011) notes that [bi N_i hai N_i], as a non-canonical comparative construction, does not so much compare two entities as it profiles an attribute of the comparee (e.g., sunniness, attributed to 'my heart' in (2)). This view is also shared by Ji (2012) and Ma (2014), who hold that [bi N_i hai N_i] serves to profile the characteristics of the comparee. Cheng (2013: 62), then, argues that the constructional meaning of [bi N_i hai N_i] can be characterized as "super or extremely + the stereotypical attributes of the noun", which is a non-compositional meaning that has a coloring of exaggeration.

    A second focus lies with the semantic features of the nominal constituent N sanctioned by the construction. It has been suggested that only nouns of strong descriptive character occur in [bi N_i hai N_i]. The descriptor 'strong descriptive character' was used, for instance, in Luo and Lu (2016), Ma (2014), and Yang (2011), to refer to nouns with salient attributes. For instance, in 他比葛朗台还葛朗台 *Ta bi Grandet hai Grandet* 'He is meaner than Grandet', the noun 葛朗台*Grandet* (referring to the character in the famous novel *Eugénie Grandet*) easily fits into the construction because it is strongly associated with the quality of being mean (see also Cheng 2013; Li 2016). Luo and Lu (2016) note that the nouns sanctioned into the construction have even become bywords for the salient attributes and that the referent of the noun is usually familiar to both the speaker/writer and the interlocutor/reader.

    Thirdly, previous research has also touched on the mechanism underlying the development of [bi N_i hai N_i]. In this respect, Yang (2011) notes that the development of [bi N_i hai N_i] is analogically modeled on [bi N hai A]. Cheng (2013) as well contends that [bi N hai A] is the source of [bi$_i$ N hai N_i], and he further proposes that the development of the latter construction is pragmatically motivated, in that the use of a N(oun) rather than an A(djective) avoids direct expression of the construction's dimension of comparison. That said, Yang's and Cheng's studies provide no in-depth exploration of the development of [bi N_i hai N_i], as it is not their main focus. In contrast, the present study presents a nuanced account of the development of [bi N_i hai N_i] from [bi N hai A], and specifically what motivates the extension from A(djectives) to N(ouns).

2.2 A brief introduction to [Adverb A], [Adverb N], and [bi N hai A]

It will be argued (see Section 3.1) that among the source constructions that are of primary importance in the emergence and development of [bi N$_i$ hai N$_i$] and [bi N$_i$ hai N$_j$] are [Adverb N] and [bi N hai A]. The [Adverb N] construction, which dates back to the Qing dynasty (1644–1912), developed from a previously existing high-frequency construction, [Adverb A], a canonical construction in Mandarin – an example is (4). Like the adverbs in [Adverb A], the adverbs in [Adverb N] are on the whole degree adverbs. There are at least 17 degree adverbs that participate in [Adverb N] (cf. Li 2016: 81), such as 很 *hen* 'very', 最 *zui* 'most', 太 *tai* 'too',非常 *feichang* 'very', exemplified in (5)–(8).

(4)   做到让人人喜欢真的<u>很难</u>。
        Zuo   dao   rang   renren   xihuan zhende <u>hen nan</u>.
        Do   arrive make every.one like   really very hard
        'It's really hard to be liked by everyone.'

(5)   每个班总有一个<u>很爷们儿</u>的女生。
        Meige   ban   zong   you   yi   ge   <u>hen</u>   <u>yemener</u> de   nvsheng.
        Each   class always have one CL   very   menfolk MOD Girl
        'There is always a very masculine girl in each class.'

(6)   忘记带钥匙<u>最悲剧</u> 了。
        Wangji   dai   yaoshi   <u>zui</u>   <u>beiju</u>   le.
        Forget   bring   key   very   tragedy   PERF
        'Forgetting to bring the key is one of the most tragic things.'

(7)   这个动作<u>太爷们</u>了。
        Zhe   ge   dongzuo   <u>tai</u> <u>yemen</u>   le.
        This CL action   too   menfolk PERF
        'This action is too mannish.'

(8)   我妹妹现在大学还没毕业, <u>非常青春</u> !
        Wo   meimei   xianzai daxue   hai mei biye,   <u>feichang</u>
        I   younger.sister now   university yet not graduate, extremely
        <u>*qingchun!*</u>
        youth
        'My sister has not graduated from university yet, and she is very young!'

The other source construction is [bi N hai A], a frequently-used, canonical comparative construction in Mandarin Chinese, as in (1), first attested in the late Yuan dynasty (14th century) in the CCL corpus. As 还 *hai* is also an adverb in Mandarin, this construction is closely related to [Adverb A], in that the component part "hai A" of [bi N hai A] instantiates the more schematic [Adverb A]. A variant of *hai*, namely, 更 *geng* can collocate with either an adjective or a noun constituting constructions [geng A] or [geng N] with the capacity to function as predicates, as exemplified by instances such as 她更美 *ta geng mei* 'she is more beautiful' and 他更绅士 *ta geng shenshi* 'he is more gentleman-like'. At the same time, [geng N] can also work as a component part of a comparative construction, viz., [bi N geng N],[7] as in 他比老板更老板 *ta bi laoban geng laoban* 'he is bossier than a boss', which is a variant of [bi N hai N].

2.3 Theoretical framework: (diachronic) construction grammar

The present study is couched within (Diachronic) CxG, a framework that exploits the "fit between the mechanisms of syntactic change and the basic principles of Construction Grammar" (Barðdal and Gildea 2015: 9). Or, as Perek (2020: 142) puts it, "Diachronic Construction Grammar aims to describe and explain language change by drawing on the idea that the grammar of a language consists of an inventory of form-meaning pairs, called constructions." An important tenet of CxG is its network design, whereby all constructions form a structured inventory or a network of constructions, called a "construct-i-con" by Goldberg (2003: 219). This network comprises vertical (taxonomic and meronymic) as well as horizontal links[8] between constructions. As Smirnova and Sommerer (2020: 3) point out, the incorporation of linguistic changes into CxG entails their reconceptualization as 'network changes', which may involve, among others, node creation (giving rise to constructionalization) and constructional network reconfiguration.[9]

Like CxG, DCxG is usage-based in that linguistic change is shaped by language use (Barlow and Kemmer 2000; Bybee 2010; Diessel 2019; Hopper 1987; Langacker

---

7 It can plausibly be assumed that these *geng*-constructions can partake of a more encompassing network of conventional and unconventional comparative constructions, whereby [bi N_i geng N_i] forms a sister node with [bi N_i hai N_i], together forming the higher-order schema [bi N Adverb N]. As a reviewer further pointed out, the presence of [bi N_i geng N_i] as a sister node of [bi N_i hai N_i] may mutually strengthen these constructions, in that the [bi N_i geng N_i] construction may exhibit semantic similarity with [bi N_i hai N_i] in terms of collocating nouns (on the importance of collocating nouns in the development of [bi N_i hai N_i], see Section 3.2).

8 Horizontal links have, on the whole, received less attention than vertical links.

9 Smirnova and Sommerer (2020: 3) also mention node loss (giving rise to constructional death) and node-internal changes (giving rise to constructional change), but these will be seen to be less relevant for the present paper.

2000; Schmid 2020). Major components of this usage-based approach which will be seen to play a role in the present study are frequency, entrenchment, and general cognitive abilities such as analogical reasoning and schematization (see also Smirnova and Sommerer 2020: 2).

Analogy, apart from being a mechanism of linguistic change, can be viewed as a domain-general cognitive processing mechanism: it is at the basis of speakers' ability to create novel utterances (Bybee 2010: 75). The prerequisite of analogical processing is language users' recognition – conscious or subliminal – of similarities between two structural or functional contexts.

Another central tenet in usage-based linguistics, and a major factor in grammatical change, is frequency of use. Frequent instances of a construction impact its cognitive representation, in that each instance contributes to the representation of the construction as a unit in the minds of language users, thus increasing its entrenchment and automatization (e.g., Diessel 2019: 1; Schmid 2020: 216). Frequency plays an important role in analogical extension: as Sommerer (2015: 11) points out, a speaker can extend a construction analogically only if they become aware of it in the first place and "[s]uch awareness will be favored or triggered by the construction's high frequency".

2.4 Nouns versus adjectives

Just like [Adverb N] has been said to have developed from [Adverb A] (see Section 2.2), the [bi N_i hai N_i] construction will be seen to develop from [bi N hai A] (Section 3.1). In either case, a noun is recruited into a slot where it is modified by an adverb; at the same time, adverbial modification is a typical characteristic of adjectives/ verbs. Does this mean that the nouns recruited into this slot have become adjectives? To answer that question, we draw on Zhu's (1997: 207, 213) insight that there is no correspondence between word class and syntactic function in Chinese: when an adjective or verb fills an argument position (e.g., subject or object) in a sentence, it remains a verb or adjective and does not change into a noun. In the same vein, when a noun fills the adjectival slot in [Adverb A] or [bi N hai A] (resulting in [Adverb N] and [bi N_i hai N_i]), it remains a noun, because it has properties not shared by adjectives. Zhu (1997: 204) notes that in Mandarin, the distinction between nouns and adjectives/verbs is prominent in that adjectives and verbs pertain to the category of 谓词 *weici* 'predicate words'; this means that both verbs and adjectives usually serve as predicates in a sentence, and when they are predicates, they can be followed by aspect markers such as 了 *le* or 着 *zhe*, while nouns are seldom used as predicates.

In general, nouns "characteristically represent entities, real or imagined, in the concrete or virtual world" (Shi 2016: 199) such as 孔子 *Kongzi* 'Confucius', 律师 *lvshi* 'lawyer', 会议 *huiyi* 'conference', while adjectives denote properties of entities (Huang et al. 2016: 276) such as 漂亮 *piaoliang* 'beautiful', 白 *bai* 'white', 安静 *anjing* 'quiet'. Among the (more specific) distinguishing properties of adjectives are their amenability to affixation and reduplication (see Huang et al. 2016: 276–296; Shi 2016: 199–255). Nouns recruited into the [Adverb A] or [bi N hai A] structures, then, will not lose their nominal properties, nor will they assume these adjectival properties.[10]

3 The emergence of [bi N_i hai N_i]

In this section, we first discuss (the mechanisms underlying) the emergence of [bi N_i hai N_i]. In particular, we suggest that [bi N_i hai N_i] has developed from three sources, [bi N hai A], [Adverb N], and the earlier development from [Adverb A] to [Adverb N]. Additional support for our analysis will come from a collostructional analysis.

3.1 The multiple sources of [bi N_i hai N_i]

(i) **[bi N hai A] as a source.** It has been amply argued that language users can create novel expressions on the basis of analogical extension from (i.e., perceived similarity in form and/or meaning with) previously existing conventionalized phrases (Boas 2003; Bybee 2010: 60, 63; Gentner and Markman 1997; Krott et al. 2006). This means that many utterances are actually only partly novel. In this respect, Schmid (2020: 100) defines innovations as "partly licensed utterances", and Traugott and Trousdale (2013: 58) note that "no construction is entirely new (except those that are borrowings and some coinings)".

(9)   苍天啊! 一副眼镜比机票还贵。
        Cangtian a! yi   fu   yanjing bi   jipiao   hai   gui.
        God   FP! One CL   glasses BI   plane.ticket HAI   expensive
        'God! A pair of glasses is more expensive than a plane ticket.'

---

10 In this respect, it is undisputed that the noun 绅士*shenshi* 'gentlemen' in [hen shenshi] 'very gentleman-like' is still a noun, as it is the case for the noun 绅士*shenshi* 'gentlemen' in [*bi shenshi hai shenshi*]. Even if the distinction between adjective and noun may not always be clear-cut, as in 很血腥 *hen xuexing* 'very bloody', there are still many instances of [bi N_i hai N_i] where the nouns are unequivocally nouns, as in 比闹钟还闹钟 *bi naozhong hai naozhong* 'BI alarm clocks HAI alarm clocks', 比警察还警察 *bi jingcha hai jingcha* 'BI policemen HAI policemen'.

Against this background, we argue that it is through analogical extension that the construction [bi N_i hai N_i] developed from the prior, conventionalized [bi N hai A], as in (9). The source construction [bi N hai A] is attested from the 14th century[11] (i.e., the late Yuan Dynasty) in the corpus CCL, while the earliest attested instance of [bi N_i hai N_i] (see example (10)) can only be dated back to 1969. In this process of analogical modeling, frequency has been argued to play an important role (Mańczak 1980; Bybee 2010: 25, 63). As Sommerer (2015: 118) states: "[T]he speaker can extend a specific schema analogically only if s/he becomes aware of it in the first place. Such awareness will be favored … by the high frequency of syntactic patterns which are compatible with such an underlying schematic representation". The attested normalized frequency of occurrences instantiating [bi N hai A] is 7.26 per million characters in the multi-genre[12] sub-corpus of BCC. Note that the normalized frequency of [bi N_i hai N_i] is 0.18 per million characters in the same corpus. Its raw frequency, number of types and type/token ratio are presented in

Table 1: Overall frequency of [bi N_i hai N_i] in the multi-genre sub-corpus of BCC.

[bi N_i hai N_i]	Value
Occurrence/tokens	314
Types	177 (140 hapaxes)
Type/token ratio	0.56

Table 2: Top 10 of [bi N_i hai N_i] types in the multi-genre sub-corpus of BCC (raw frequency).

[bi N_i hai N_i]	Noun	Frequency
[bi *nǚren* hai *nǚren*]	女人 *nǚren* 'woman'	38
[bi *qinshou* hai *qinshou*]	禽兽 *qinshou* 'brute'	25
[bi *nanren* hai *nanren*]	男人 *nanren* 'man'	15
[bi *yemen* hai *yemen*]	爷们 *yemen* 'menfolk'	12
[bi *liumang* hai *liumang*]	流氓 *liumang* 'rogue'	7
[bi *mogui* hai *mogui*]	魔鬼 *mogui* 'devil'	7
[bi *zhu* hai *zhu*]	猪 *zhu* 'pig'	5
[bi *jiqiren* hai *jiqiren*]	机器人 *jiqiren* 'robot'	4
[bi *zibenzhuyi* hai *zibenzhuyi*]	资本主义*zibenzhuyi* 'capitalism'	3
[bi *zibenjia* hai *zibenjia*]	资本家*zibenjia* 'capitalist'	3

---

11 The earliest [bi N hai A] instance is attested in《元代话本选集》'*A Collection of Stories from the Yuan Dynasty*': 我五十岁上发迹, 比甘罗虽迟, 比那两个还早 'I made my fortune at 50, later than Ganluo, earlier than those two'.

12 The data in this sub-corpus are all Contemporary Chinese.

Table 1, and Table 2 shows the raw frequencies of the top 10 types of this construction.

(10)   一位坦桑尼亚朋友看到苏联军队在乌苏里江用高压水龙头向中国渔民袭
         击的情景时, 愤慨地说: 可耻! 比帝国主义还帝国主义 。(《人民日报》)

Yi	*wei*	*Tansanniya*	*pengyou*	*kandao*	*sulian*	*jundui*	*zai*	*Wusuli*
One	CL	Tanzanian	friend	see	Soviet	troops	PREP	Ussuri

Jiang	*yong*	*gaoya*	*shuilongtou*	*xiang*	*zhongguo*	*yumin*
river	use	high.pressure	tap	toward	Chinese	fisherman

xiji	*de*	*qingjing shi,*	*fengaide*	*shuo:*	*kechi!*	<u>*Bi*</u>
attack	MOD	Situationtime,	indignantly	say:	shameful!	BI

<u>*diguozhuyi*</u>	<u>*hai*</u>	<u>*diguozhuyi*</u>
imperialism	HAI	imperialism

‘When a Tanzanian friend saw Soviet troops attacking Chinese
fishermen with high-pressure taps on the Ussuri River, he cried
indignantly: "Shame! They are more brutal than imperialism".’

What does the analogical extension consist in? [bi N hai A] is a conventionalized construction conveying that a comparee takes a position on a standard or scale (expressed by A – the dimension of comparison) that is even higher (*hai*) than the (high) position taken by N (the standard of comparison). By way of example, consider examples (1) and (9). In instances such as these, [bi N hai A] provides an analogical base for [bi N_i hai N_i]. Specifically, in employing [bi N_i hai N_i], speakers retain the meaning of the previously existing conventionalized expression [bi N hai A]. They also retain a good part of its form, but they do modify A into N. In the process, non-scalar N is adjusted or reinterpreted in terms of N's typical, scalar attributes; for instance, 阳光 *yangguang* 'sunlight' in (2) is interpreted as 'sunniness', i.e., in terms of the scalar attributes of the sun; in (3), 天才 *tiancai* 'genius' is reinterpreted as (scalar) 'genius-like' characteristics. We suggest that the interpretation of N as A, and hence the extension of [bi N_i hai N_i] from [bi N hai A], can be attributed to coercion[13]

---

13 As one of the reviewers points out, while the constructional meaning of [bi N hai A] plays a role – through coercion – in the interpretation of the new construction [bi N_i hai N_i], it is also advisable to pay attention to the nature of the words filling the N-slot. A first impression of these words is given in Table 2, listing the most frequently used types in the N-slot. This reviewer further proposes that the constructional meaning of [bi N_i hai N_i] may also result from the interaction between the construction and the semantic nature of the noun (N_i). In other words, not all nouns are sanctioned into the [bi N_i hai N_i] construction. While we have not carried out a detailed analysis of the semantic characteristics of the nouns that are sanctioned into the [bi N_i hai N_i] construction, the existing literature suggests that the N-slot is typically filled by generic nouns (i.e., nouns not modified by definite articles, demonstratives or other modifiers) and some proper nouns featuring idiosyncratic and striking attributes (see also our literature survey in Section 2.1). Furthermore, we explored the

(Lauwers and Willems 2011; Michaelis 2005; Traugott and Trousdale 2013: 204–206). As Perek (2020: 147) points out, coercion captures the fact that "speakers occasionally push the boundaries of conventional usage". In this case, while the noun in the construction does not have a scalar meaning of itself and would therefore make [bi N_i hai N_i] ill-formed, it can adopt the scalar meaning under influence of [bi N hai A]'s constructional/schematic meaning 'higher (than expected) position on a scale than'. What facilitates reinterpreting N as A is likely the metonymic relationship between N and its stereotypical (scalar or N-like) attributes.[14] Perek (2020: 147–148) further highlights that "if similar instances of coercion recur, the adjusted schema can itself become a conventional unit". A possible consequence is that a higher-order schema [bi N hai X] may emerge, covering both the original [bi N hai A] construction and the adjusted/analogized construction [bi N_i hai N_i], as Figure 1 shows.[15]
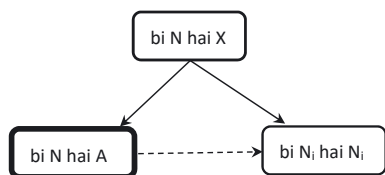

Figure 1: Network of the [bi N hai X] comparative construction family.

In this case, then, we can observe that the criterion for sanctioning an item entering the schematic slot of a construction is not semantic similarity to the central member(s) of the slot (as, for instance, *insane* and *nuts* are sanctioned in [*drive someone X*] via semantic similarity to the prototypical fillers *mad* or *crazy* (see Bybee 2010: 36–37)), but semantic contiguity, i.e., a metonymic relationship.

Within the framework of DCxG, these linguistic changes involve node creation ([bi N_i hai N_i]) and possibly also constructional network reconfiguration (if a higher-order schematic construction [bi N hai X] emerges). The analogical extension from [bi N hai A] to [bi N_i hai N_i] is to be situated at the same level of schematicity;

---

semantics of the N-slot when investigating any correspondence between the semantic categories of nouns significantly attracted to [Adverb N] and to [bi N_i hai N_i] (see Section 3.2).

14 Metonymy has also been proposed by Li (2016) and Liu (2018) as the cognitive motivation for the presence of a N in the construction [Adverb N]. That is, the reason why an adverb can collocate with a noun is that there exists a metonymic relation between the noun and its attributes, so that the noun can metonymically refers to its attributes in the construction (see also below).

15 In this respect, Perek (2020: 148) points to a similar statement by Traugott and Trousdale (2013: 16): "[P]artially sanctioned extensions of an existing conventionalized construction may over time become fully sanctioned instances of a more general, schematic construction".

it thus constitutes a horizontal link (see, e.g., Smirnova and Sommerer 2020: 25–28) symbolizing partial similarity and non-inheritance.[16]

(ii) **[Adverb N]** *as a source.* A second construction motivating the emergence of [bi N_i hai N_i] is the schematic construction [Adverb N],[17] dating back to the Qing dynasty (1644–1912) in the CCL corpus.[18] Unlike the source construction [bi N hai A], which is linked horizontally with [bi N_i hai N_i] (i.e., situated at the same level of schematicity) by analogical extension, [Adverb N] is taxonomically linked up with the constructional constituent "hai N" of [bi N_i hai N_i] through inheritance. In this respect, Goldberg (1995: 98) notes that constituents of constructions can inherit from other, higher-order constructions as well. Hence, the constituent "hai N" of [bi N_i hai N_i] can inherit from [Adverb N]. This notion of inheritance "captures a relation between more abstract constructions, which are situated towards the top of the constructional network, and more specific constructions, which are found in lower levels of the constructional hierarchy" (Hilpert 2014: 57). The type of inheritance link at stake here is an instance link, which is the basic type of inheritance link (see Goldberg 1995: 79–80; Hilpert 2014: 60; Traugott and Trousdale 2013: 50): 还 *hai* is an adverb and when it is used as a component part of a comparative construction, it can denote a degree difference in quality, namely, expressing 'being more (than expected)' (Lü 1999). Thus, 还 *hai* in [bi N hai X] is an instantiation of the general category '(Degree) Adverb'. Summing up so far, we contend that the emergence of [bi N_i hai N_i] is not only the result of analogical extension (see above), but it has also developed through inheritance, that is, by its instantiating the highly frequent schema [Adverb N].[19]

(11)   盗别人作品不是盗窃么? 还学生呢!
        Dao   bieren   zuopin bu   shi   daoqie   me? Hai   xuesheng   ne!
        Steal   others   work   not   COP   steal   FP?   HAI   student   FP
        'Isn't stealing someone else's work stealing? How can he as a student do such things!'

---

16 The fact that constructional networks also contain horizontal links was already pointed out by Langacker (2000: 102), who talks about the "outward growth" of a network.

17 It would actually be more correct to label this construction [Degree Adverb N]; however, we will follow common practice in the Chinese linguistic literature here, and use the more general label [Adverb N].

18 The earliest instance of [Adverb N] is attested in the novel 《八仙得道》 '*The Eight Enlightened Immortals*': 这位春瑛小姐, … 简直还是一位很清高很贞节而又非常近情的好女子 'This Miss Chunying… is also a very virtuous and chaste woman who is very sensible'.

19 The normalized frequency in the multi-genre sub-corpus of BCC of [hen N], the prototypical instance of [Adverb N], is 64.65 per million characters. Hence, that of [Adverb N] must be much higher.

A caveat is in order here. Although [hai N] (as in (11)) is an independent construction in Mandarin, it is impossible for it to be the up-level schema of [bi N_i hai N_i] as the meaning of independent [hai N] is quite different from that of the component "hai N" in [bi N_i hai N_i]. As example (11) shows, in [hai N *ne/na*], the noun is referential while in instances such as (2), the noun following *hai* in [bi N_i hai N_i] is not referential, i.e., it does not denote an entity, but it metonymically refers to the entity's stereotypical attributes. Similarly, the noun in [Adverb N], as in (5)–(8), is not referential either. According to Lü (1999), 还 *hai* is an adverb expressing tones (e.g., of sarcasm, surprise, praise) in Mandarin and only when it is used in comparative constructions can it denote a degree difference in quality, namely, expressing 'being more (than expected)'. Hence, in [hai N *ne/na*], as in (11), 还 *hai* is an adverb denoting tones while in the comparative construction such as [bi N hai A] and [bi N hai N], it expresses a degree difference in quality. In addition, there is a distinction in grammatical stress placement between [hai N ne/na] and [bi N hai N]. In [hai N ne/na], the stress falls on the N, whereas in [bi N hai N], the stress falls on 还 *hai*.

As Zong (1995) notes, in actual usage events, [hai N ne/na] denotes that the subject should have behaved the way Ns usually behave and thus conveying a tone of blame or sarcasm, as expressed by [*hai xuesheng ne*] in (11). Specifically, what the speaker intends to express in (11) is that stealing others' work is not what a student should do, but the student did it anyway. Zong (1995) also compares the semantics of the nouns in [hai N ne] with those in [bi N_i hai N_i] and finds that the semantic categories of nouns in these two constructions are quite different: the nouns in [hai N ne] are predominantly nouns denoting people's occupations, titles/ranks or status, while those in [bi N_i hai N_i] cover a much wider range of categories.

(iii) **The related constructions [Adverb A] and [Adverb N].** Importantly, [Adverb N] can itself be said to have developed from [Adverb A] by analogical extension, whereby – in a way similar to the shift from [bi N hai A] to [bi N_i hai N_i] – the meaning 'A to a still high/higher degree' is retained, but the form gets adjusted (A > N). Moreover, given that the schematic constructions [Adverb A] and [Adverb N] are both present for speakers of Mandarin at a time when the [bi N_i hai N_i] construction had not developed yet,[20] it is likely that these speakers viewed them as related (that is, as horizontally linked nodes in a constructional network). We suggest these linked schematic nodes may have facilitated the shift of the dimension-of-comparison slot from A in [bi N hai A] to N in [bi N_i hai N_i], as illustrated in Figure 2.

20 Recall that [Adverb N] dates back to the Qing dynasty (1644–1912) and had developed from a previously existing high-frequency construction [Adverb A] (see Section 2.2).
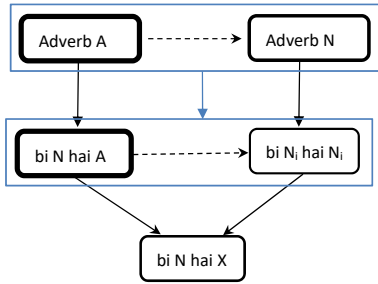
Figure 2: Network of [bi N hai A/N] and [Adverb A/N].

3.2 Finding quantitative evidence for [Adverb N] as a source

If our assumption is correct that [bi N_i hai N_i] emerged via the extension of the A slot in [bi N hai A] to nouns via inheriting the nominal feature from [Adverb N], it would not be unreasonable to expect that the semantic categories of nouns that are significantly attracted to [bi N_i hai N_i] are correspondent to those attracted to [Adverb N]. This type of analogy does not involve identifying a common pattern (as in [bi N_i hai A] > [bi N_i hai N_i] or in [Adverb N] > "hai N" as a component of [bi N_i hai N_i]), but rather consists in similarity of conceptual content (labeled "concrete property matches" in Gentner and Smith 2012: 130, or "physical similarities" in Behrens 2017: 222). To investigate this hypothesis, we conducted a collostructional analysis.

Data for this study were retrieved from the corpus BCC (Xun et al. 2016). BCC has a total size of about 9.5 billion characters and it comprises five genre-based sub-corpora: news reportage (2 billion), literary works (3 billion), dialogues (600 million), multi-genre (1.9 billion), and Literary Chinese (2 billion). The corpus was tagged for part of speech.

Since the [bi N_i hai N_i] construction is a relatively recent development – its earliest attested instance can only be dated back to 1969 and the earliest attested occurrence of [bi N_i hai N_j] dates from 1980,[21] the data were retrieved from the multi-genre sub-corpus of the BCC Contemporary Chinese corpus. This sub-corpus is balanced by genre, including texts from newspapers, fictions, micro-blogs, and academic journal articles of science and technology.

We searched the multi-genre sub-corpus of the BCC corpus with the query "*bi* n *hai* n". In all, 559 concordances of [bi N_i hai N_i] and [bi N_i hai N_j] were retrieved. After

---

21 We attempted to retrieve data of [bi N_i hai N_i] and [bi N_i hai N_j] from the sub-corpus of "Ancient Chinese" in the BCC corpus and also from the corpus of CCL, but we obtained no concordance. All the instances of [bi N_i hai N_i] and [bi N_i hai N_j] pertain to the Contemporary Chinese period, which in the periodization of Chinese (e.g. Wang 2004) starts in 1911.

manual removal of noise, 314 instances of [bi N_i hai N_i] and 149 instances of [bi N_i hai N_j] were retained. To compare the semantic features of nouns in [bi N_i hai N_i] with those in [Adverb N], we searched the multi-genre sub-corpus of the BCC corpus with the query "*hen* n" and obtained 64,648 concordances, from which we randomly sampled 1000 concordances. After discarding noise, we retained 264 instances of [hen N].[22]

To examine the similarities between the lexical preferences expressed by N in [bi N_i hai N_i] and [Adverb N], we conducted a simple collexeme analysis[23] (e.g., Stefanowitsch 2006, 2013; Stefanowitsch and Gries 2003), which can determine which nouns are statistically significantly associated with each of the two constructions when compared against the language as a whole. Albeit subject to some criticism (e.g., Divjak 2008; Schmid 2010; Schmid and Küchenhoff 2013), collostructional analysis is still regarded as a useful method to calculate the association strength between a construction and its slot-fillers.

For every noun in [bi N_i hai N_i] and in [hen N], we calculate its association strength with the construction via computing a 2-by-2 contingency table which contains four different frequencies of occurrence: the frequency of the target noun in the target construction (A), the frequency of the target noun in the corpus in general (B), the frequency of the construction filled with other nouns than the target noun (C), and the frequency of all other constructions with lexemes other than the target noun (D).

On the basis of these frequencies, the expected frequencies and the association measure can be calculated. Based on the contingency table, we also derived the log odds ratio as well as the 95 % confidence interval for every noun (Schmid and Küchenhoff 2013). A significantly positive log odds ratio suggests that the noun is attracted to the construction whereas a significantly negative log odds ratio indicates that the noun is repelled. Figures 3 and 4 visualize the most important part of the results: the top 30 significantly attracted nouns (purple bars) and the top 4 repelled nouns (green bars).

---

22 We are aware that, rather than using [hen N] as a proxy for [Adverb N], it would have been preferable to include at least a selection of degree adverbs combining with N in the collostructional analysis, but operationally, this was difficult to do (many spurious hits). Specifically, we tried to extract data from BCC with the query "d n" ("d" is the tagging for adverbs and "n" is the tagging for nouns in the corpus), but the overwhelming majority of the concordances obtained was not felicitous as in Mandarin sentence structure, an adverb can serve as an adverbial located at the beginning of a sentence, which can then be followed by a noun that serves as the subject as in 终于论文写完了 *zhongyu lunwen xiewan le* '(literally) Finally thesis write finish PERF'.
23 Following Sommerer and Baumann (2021), we made some minor adjustments to overcome statistical shortcomings (see Schmid and Küchenhoff 2013). The simple collexeme analysis was done in R (R Core Team 2020).
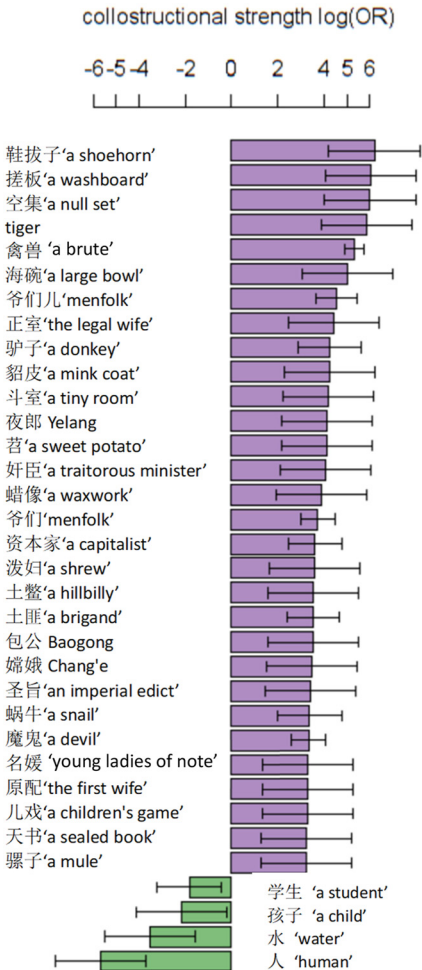
collostructional strength log(OR)

-6-5-4 -2   0   2   4 5 6

	鞋拔子 'a shoehorn'
	搓板 'a washboard'
	空集 'a null set'
	tiger
	禽兽 'a brute'
	海碗 'a large bowl'
	爷们儿 'menfolk'
	正室 'the legal wife'
	驴子 'a donkey'
	貂皮 'a mink coat'
	斗室 'a tiny room'
	夜郎 Yelang
	莒 'a sweet potato'
	奸臣 'a traitorous minister'
	蜡像 'a waxwork'
	爷们 'menfolk'
	资本家 'a capitalist'
	泼妇 'a shrew'
	土鳖 'a hillbilly'
	土匪 'a brigand'
	包公 Baogong
	嫦娥 Chang'e
	圣旨 'an imperial edict'
	蜗牛 'a snail'
	魔鬼 'a devil'
	名媛 'young ladies of note'
	原配 'the first wife'
	儿戏 'a children's game'
	天书 'a sealed book'
	骡子 'a mule'

学生 'a student'
孩子 'a child'
水 'water'
人 'human'

Figure 3: Collostructional strengths of [bi N_i hai N_i] constructions with logOR >3.2 and logOR <−1.8. Strength is measured by log(OR) (horizontal axis). Error bars denote 95 % confidence intervals. Purple denotes attracting nouns ('attractor'; log(OR) significantly positive); green denotes repelling nouns ('repellor'; log(OR) significantly negative).

Based on the results of the collostructional analysis, we classified the top 30 attracted nouns in [bi N_i hai N_i] and those in [hen N] into five categories: human, animal, place, object, and abstract nouns (this categorization was taken from Shao et al. 2019). 'Human' includes nouns such as 男人 *nanren* 'men', 女人 *nüren* 'women', 老板 *laoban* 'boss', 绅士 *shenshi* 'gentlemen', and 警察 *jingcha* "policemen"; "animal" includes 熊猫 *xiongmao* 'panda' and 骡子 *luozi* 'mule'; 'place' involves 宾馆 *binguan* 'hotel' and 地狱 *diyu* 'hell';[24] 'object' encompasses 教科书 *jiaokeshu* 'textbooks' and

---

24 It should be noted that we opted to classify 地狱 'hell' as a place rather than as an 'abstract noun' because people tend to believe that there exists such a place as 地狱 'hell' and that those who are evil will be dispatched to the hell after their deaths.
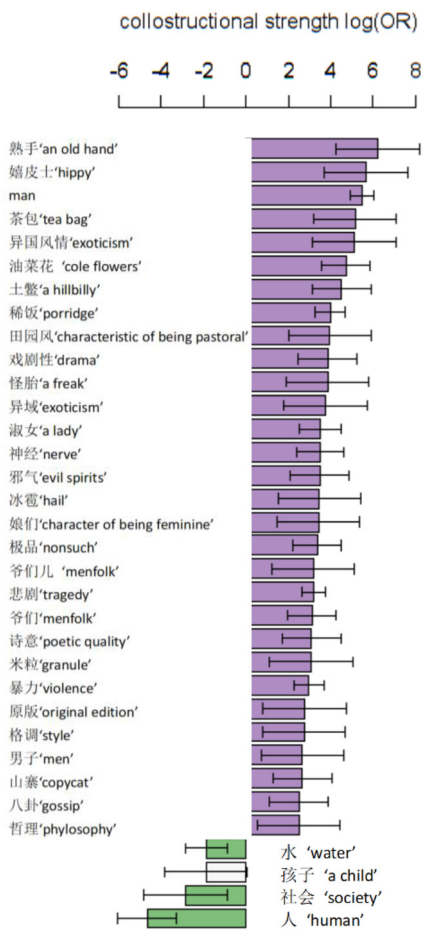
collostructional strength log(OR)

-6 -4 -2 0 2 4 6 8

熟手'an old hand'
婊皮士'hippy'
man
茶包'tea bag'
异国风情'exoticism'
油菜花 'cole flowers'
土鳖'a hillbilly'
稀饭'porridge'
田园风'characteristic of being pastoral'
戏剧性'drama'
怪胎'a freak'
异域'exoticism'
淑女'a lady'
神经'nerve'
邪气'evil spirits'
冰雹'hail'
娘们'character of being feminine'
极品'nonsuch'
爷们儿 'menfolk'
悲剧'tragedy'
爷们'menfolk'
诗意'poetic quality'
米粒'granule'
暴力'violence'
原版'original edition'
格调'style'
男子'men'
山寨'copycat'
八卦'gossip'
哲理'phylosophy'

水 'water'
孩子 'a child'
社会 'society'
人 'human'

Figure 4: Collostructional strengths of [hen N] constructions with logOR >3.1 and logOR <−1.95. Strength is measured by log(OR) (horizontal axis). Error bars denote 95 % confidence intervals. Purple denotes attracting nouns ('attractor'; log(OR) significantly positive); green denotes repelling nouns ('repellor'; log(OR) significantly negative); light gray denotes constructions which are neither repellors nor attractor.

闹钟 *naozhong* 'alarm clocks'. The label 'abstract noun' here refers to nouns that are "non-referrable to concrete objects or entities in the natural or human world" (Po-Ching and Rimmington 2004: 10), and are products of human epistemology, such as 真理 *zhenli* 'truth' and 梦想 *mengxiang* 'dreams', which is in contrast to the other four categories that express concrete concepts. The semantic distribution of the top 30 nouns attracted to [bi N_i hai N_i] and [hen N] is presented in Table 3.

Table 3: Frequency of various semantic categories of top 30 nouns attracted to [bi N_i hai N_i] and [hen N].

	Abstract noun	Animal	Human	Object	Place	Total
[bi N_i hai N_i]	4	5	13	6	2	30
[hen N]	12	1	9	6	2	30

To examine whether the (types of) nouns attracted to the two constructions are similar, we tested the association between the semantic categories of the most attracted nouns in [bi N_i hai N_i] and in [Adverb N], making use of a Fischer's exact test. This test revealed that there is no statistically significant difference (p = 0.1164) between the frequencies of the categories of most attracted nouns in the two constructions, indicating that there is a fair degree of semantic similarity between the nouns attracted to the two constructions.

In summary, we have seen that the emergence of [bi N_i hai N_i] can be attributed to multiple sources: there exist both formal and semantic similarities between the analogical model construction [bi N hai A] and the target [bi N_i hai N_i], such that the latter emerged by analogy with the former; next, [bi N_i hai N_i] also inherits from [Adverb N]; furthermore, the development [Adverb A] > [Adverb N] might have facilitated the shift from [bi N_i hai A] to [bi N_i hai N_i] (Trousdale 2013; Van de Velde et al. 2013).

4 From [bi N_i hai N_i] to [bi N_i hai N_j]

Following the emergence of [bi N_i hai N_i], this construction further developed into [bi N_i hai N_j], where N_j is different from N_i. While the earliest attested instance of [bi N_i hai N_i] in BCC dates back to 1969, the earliest attested occurrence of [bi N_i hai N_j] dates from 1980. In this section, we will discuss the mechanisms of and motivations for the emergence of [bi N_i hai N_j], and the semantic change involved in the dimension slot of [bi N_i hai N_i] and [bi N_i hai N_j] via a quantitative analysis of corpus data.

4.1 The emergence of [bi N_i hai N_j]

We have seen that in [bi N_i hai N_i], the noun following *hai* – expressing the dimension of comparison – does not refer to a specific entity but metonymically refers to the entity's stereotypical attributes and is used to specify the attributes of the comparee. What [bi N_i hai N_j], as exemplified in (12), shares with [bi N_i hai N_i] is that N_j does not refer to a specific entity either. What differentiates the two constructions is that in [bi N_i hai N_j], the noun expressing the dimension of comparison (N_j) differs from the noun expressing the standard of comparison (N_i). Accordingly, in (12), N_j 剧情 *juqing* 'plot' metonymically refers to the quality of being plotty, and the comparee (the case analysis) has more of the stereotypical attributes of a plot (N_j) than a novel (N_i). Since N_j is different from N_i, this construction enjoys relatively more freedom to denote comparison than [bi N_i hai N_i].

(12)   其中案例分析尤其精彩, 分析范文写得**比散文还诗意, 比小说还剧情**。

Qizhong   anli   fenxi   youqi   jingcai,   fenxi
Among.them   case   analysis   particularly   wonderful,   analysis

fanwen   xie   de   bi   sanwen   hai   shiyi,   bi   xiaoshuo
model.essay   write   DE   BI   prose   HAI   poetic.quality,   BI   novel

hai   juqing.
HAI   plot.

'Among them, the case analysis is particularly wonderful. The model essay of analysis is more poetic than a prose and is more plotty than a novel.'

We retrieved (after manual cleanup) 149 tokens of [bi N_i hai N_j] from the multi-genre sub-corpus of the BCC corpus, with 133 types and 101 hapax legomena, as shown in Table 4. The type/token ratio is very high, which suggests that the construction is very productive (Bybee 2010; Dąbrowska 2008, among others).

Table 4: Overall frequency of [bi N_i hai N_j] in the multi-genre sub-corpus of BCC.

[bi N_i hai N_i]	Value
Occurrence/tokens	149
Types/different nouns	133 (101 hapaxes)
Type/token ratio	0.89

We suggest that [bi N_i hai N_i] has served as an analogical base for [bi N_i hai N_j], whereby the latter construction has retained the structural pattern as well as meaning of the former. Structurally, the standard and the dimension of comparison are each expressed by a noun; semantically, the two constructions each convey that the comparee has even more of the stereotypical attributes of the dimension of comparison (N_j/N_i) than the standard of comparison (N_i/N_i). The change that can be observed is that the nouns expressing the standard and dimension of comparison no longer need to be identical. That is, [bi N_i hai N_i] only partially sanctions [bi N_i hai N_j], just as [bi N hai A] only partially sanctions [bi N hai N]. Therefore, in Figure 5, the link between the dimension slot of [bi N_i hai N_i] and that of [bi N_i hai N_j] is indicated by a dotted arrow.
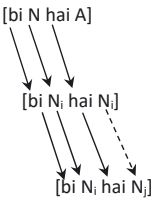
[bi N hai A]

[bi N_i hai N_i]

[bi N_i hai N_j]

Figure 5: The relationship among [bi N hai A], [bi N_i hai N_i] and [bi N_i hai N_j].

4.2 Interim summary

The change from [bi N hai A] to [bi N_i hai N_i] and then to [bi N_i hai N_j] shows an expansion of the dimension-of-comparison slot. In the first stage of the development, from [bi N hai A] to [bi N_i hai N_i], the dimension-of-comparison slot expands from an adjective to a noun which is the same in form as the noun denoting the standard of comparison. We have seen that there are several motivating factors, or sources, of this change: (i) analogy-based extension from [bi N hai A], whereby its scalar meaning is transferred to a formally similar structure, with, in its dimension-of-comparison slot, a noun that is identical to the noun in the standard-of-comparison slot; the noun in the dimension-of-comparison slot metonymically refers to the typical scalar attribute(s) of its referent, thus aligning its semantics with the scalar semantics of adjectives; (ii) [Adverb N], which is taxonomically linked up with the constructional component "hai N" of [bi N_i hai N_i] through inheritance: the category Adverb is instantiated by *hai* and the dimension-of-comparison slot of [bi N hai A] inherits the nominal feature of [Adverb N]; (iii) the shift [Adverb A] > [Adverb N] facilitates the shift to [bi N_i hai N_i]; (iv) conceptual similarity between the type of nouns attracted to [Adverb N] and [bi N_i hai N_i]. The second stage, from [bi N_i hai N_i] to [bi N_i hai N_j], is directly analogous to [bi N_i hai N_i], with [bi N hai A] as a latent analogical supporting pattern. As it did in the first stage, metonymy plays a facilitating role in this stage as well.

4.3 Quantitative description

We have seen that in the [bi N_i hai N_i] construction, the nouns expressing the standard and dimension of comparison are formally identical; in the [bi N_i hai N_j] construction, which developed from [bi N_i hai N_i], this is no longer the case. It therefore seems worthwhile investigating how these two constructions compare; specifically, we will examine the semantic features of the nouns in the dimension-of-comparison slot of the two constructions.[25] To that end, the semantic annotation scheme used to categorize the nouns in [bi N_i hai N_i] and [Adverb N] (in particular [hen N]) (see Section 3.2) is employed here as well; that is, nouns are classified into five categories: 'abstract concept', 'place', 'object', 'animal', and 'human'.

   Figure 6 illustrates the change in the distribution of semantic features of the nouns in the dimension-of-comparison slot of the two constructions, viz., the second N_i in [bi N_i hai N_i] and N_j in [bi N_i hai N_j]. In [bi N_i hai N_i], 53 percent of collocates

---

25 We have also examined pragmatic and lexical differences between the two constructions, but little difference was observed with regard to these factors and hence they are not presented here.

(i.e., 165 instances) consist of nouns referring to humans. Nouns denoting animals rank second, 18 percent (56 instances). Instances boasting the third largest number are those involving nouns encoding objects (40 instances), which is followed by instances with nouns denoting abstract concepts (38 instances). At 5 % (15 instances), nouns expressing places take up the smallest share. In [bi N_i hai N_j], a different picture is observed. The majority of nouns in the N_j slot now denote abstract concepts, accounting for 57 percent (85 instances) of all nouns, or four times more than their share in [bi N_i hai N_i]. Further, the proportion of 'animals' rises by almost 18 percent to 26 percent (39 instances); in contrast, nouns designating people drop to about 12 percent (17 instances) – a big decline from their share in [bi N_i hai N_i]. Finally, there are no nouns denoting places appearing in the N_j slot of [bi N_i hai N_j]. It can be seen, then, that the majority of N_j slots is filled by abstract nouns denoting an abstract quality of the comparee, as in 女人比男人还野心 *nüren bi nanren hai yexin* 'this woman is even more ambitious than men' (literally, this woman BI men HAI ambition). This suggests that the [bi N_i hai N_j] construction can be aligned with (or, is similar to) [bi N hai A], in that abstract concepts are very much in line with the general semantic profile of adjectives, which prototypically denote a(n) (abstract) quality. This semantic alignment of N_j with adjectives, denoting abstract qualities, we argue, is in keeping with Gentner and Smith's (2012: 130) "concrete property matches" (see also Section 3.2; Behrens 2017: 222).
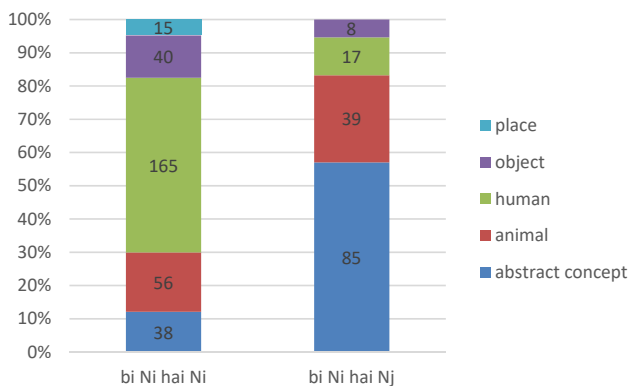

Figure 6: The semantic property of the noun in [bi N_i hai N_i] and [bi N_i hai N_j].

The development from [bi N_i hai N_i] to [bi N_i hai N_j] is in accord with Perek's (2016) observation that diachronically, new types of a construction tend to be semantically close to already existing types. Also related to this fact is the attraction effect proposed by De Smet et al. (2018), who note that as a result of analogy,

"competing forms often show attraction, becoming functionally more (instead of less) alike". This may also be the case for the development from [bi N_i hai N_i] to [bi N_i hai N_j]. That is, the canonical comparative construction [bi N hai A] attracts the noun comparative construction [bi N_i hai N_i] and makes it remove the constraint on the second noun (i.e., the second noun should be formally identical to the first noun), thus giving rise to the emergence of [bi N_i hai N_j]. In this process, the second noun can be extended to more abstract nouns, leading the whole construction to behave more similarly to [bi N hai A].

5 Implications for (diachronic) construction grammar

What has become apparent from the preceding discussion is that [bi N_i hai N_i] has developed from several constructional sources. First, there are the two previously existing, high-frequency constructions [bi N hai A] and [Adverb N]; second, the existing link between [Adverb A] and [Adverb N] may also have facilitated the shift from [bi N hai A] to [bi N_i hai N_i]. The formation of [bi N_i hai N_j] equally results from several sources: it is directly analogous to [bi N_i hai N_i], but it is also facilitated by [bi N hai A] and [Adverb N]. In addition, we have explored the role of conceptual similarity (i) between the types of nouns in [Adverb N] and [bi N_i hai N_i] and (ii) between adjectives and abstract nouns in [bi N hai A] and [bi N_i hai N_j], respectively.
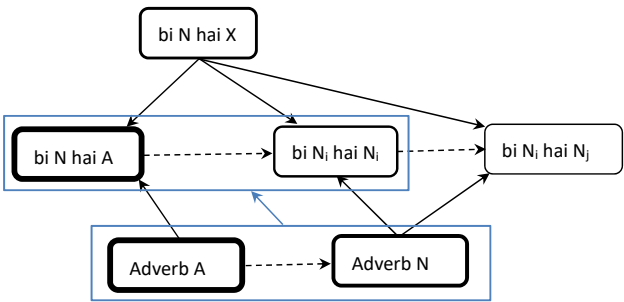

Figure 7: The constructional network of [bi N hai N].

These linked up construction pairs (e.g., [bi N_i hai N_i] and [bi N_i hai N_j] or [Adverb N] and [bi N_i hai N_i]) should not be seen as isolated phenomena; rather, they make up a network of nodes (constructions) and relations representing the organization of our linguistic knowledge (see Smirnova and Sommerer 2020: 2). A representation of this constructional network can be found in Figure 7. Some relations make up

vertical connections in the network, whereby lower-level constructions inherit features from higher-level constructions; for instance, the component "hai N" of [bi N_i hai N_i] instantiates the higher-level construction [Adverb N] in that *hai* is an adverb combining with a noun. Other relations are horizontal, as is the analogy-based development from [bi N_i hai N_i] to [bi N_i hai N_j], where the former construction partially sanctions the latter. In Figure 7, solid arrows are used for a relationship between a schema and a more specific construction instantiating or elaborating it, and a dashed arrow shows a horizontal, extensional relationship, i.e., with some disparity between the source and the target constructions such that the source just partially sanctions the target (Langacker 1987: 69). The thickness of boxes indicates entrenchment. Constructions of higher frequency are more entrenched than those of lower frequencies (Langacker 2000). In this respect, [bi N hai A] and [Adverb A] are the most entrenched ones here in the network. What is also worth pointing out about this figure is the two bigger blue boxes which indicate the emergence of [bi N_i hai N_i] from [bi N hai A] modeled on the emergence of [Adverb N] from [Adverb A].

Obviously, the network as depicted in Figure 7, is the result of diachronic changes. As we discussed in Section 3, what plays an important role in the buildup of this network are "frequency, entrenchment, and general cognitive abilities like analogical reasoning [and] schematization" (Smirnova and Sommerer 2020: 2). With respect to schematization, Figure 7 brings a key issue to the front. That is, is it really necessary to posit an overarching schematic construction such as [bi N hai X] in Figure 7 (cf. Hilpert 2019)? From our preceding discussion, it is obvious that [bi N_i hai $N_{i/j}$] developed from [bi N hai A]. However, rather than vertical, the relation between the constructions is lateral (and the nodes could be called "sister nodes"). This finding also echoes Bloom's (2021) finding that "the transfer of features from multiple source constructions to a target construction not only happens when the source and the target are taxonomically connected, but also takes place between associated constructions at the same level of abstraction". If we assume that there is an over-arching schematic construction, a follow-up question is what is its functionality and when such a schema emerges? These questions still need further research. The present study seems to suggest that it is at least after the emergence of [bi N_i hai N_i] that the schematic construction [bi N hai X] emerges as a generalization over the horizontally-associated constructions [bi N hai A] and [bi N_i hai N_i], and later also [bi N_i hai N_j].

Finally, as Liu (2022: 294) points out, there is no consensus on the conceptual ground or the nature of horizontal links among construction grammarians. Some researchers (e.g., Cappelle 2006; Perek 2015; Van de Velde 2014; Zehentner and Traugott 2020) posit that horizontally linked constructions are semantically connected (similar semantics or distinct/opposite semantics), while others (e.g., Lorenz 2020) assume that constructions sharing formal properties are horizontally linked.

In the present study, however, the horizontally linked constructions [bi N hai A], [bi N_i hai N_i], and [bi N_i hai N_j] have not only semantic similarity but also formal similarity. In this sense, our study contributes to theory building in CxG.

6 Conclusions

The results of the present study show that at the origin of [bi N_i hai N_i] lie multiple sources. First, the construction [bi N_i hai N_i] emerges by analogy with the conventional comparative construction [bi N hai A], whereby the adoption of the scalar meaning of A by a N, which is in essence non-scalar, can be ascribed to coercion. What motivates this coercion is the noun's metonymic shift from 'entity' to 'entity's attributes'. What may further have facilitated the extension from [bi N hai A] to [bi N_i hai N_i] is that the latter may have inherited the nominal feature from an already existent construction [Adverb N]. Corroboration comes from a collexeme analysis we carried out suggesting that the nouns most significantly attracted to [Adverb N] and to [bi N_i hai N_i] are semantically similar. At a more schematic level, the extension A > N observed in [bi N hai A] > [bi N_i hai N_i] may have been modeled (formally and semantically) on the existing development in Mandarin from [Adverb A] to [Adverb N]; this is also a shift-cum-coercion from scalar Adjective > Noun. Analogical extension and inheritance also underlie the subsequent development from [bi N_i hai N_i] (and indirectly also from [bi N hai A]) to [bi N_i hai N_j].

Methodologically, the study shows how collostructional analysis can benefit the study of Mandarin comparative constructions. Theoretically, this study not only showcases how language change (including node creation and network reconfiguration) happens in a constructional network, but also demonstrates that in addition to taxonomical relations between the source and the target constructions, lateral relations also play an important role in language change. Furthermore, this case study has shed light on the nature of horizontal links by revealing that constructions sharing both formal and semantic features can be horizontally linked. Third, this study exemplifies how a synchronically perceived relation between constructions (as between [Adverb A] and [Adverb N]), which results from a diachronic change ([Adverb A] > [Adverb N]), may impact a later, similar language change (namely, [bi N hai A] > [bi N_i hai N_i]). Finally, it also contributes to usage-based Construction Grammar by showing that, in addition to semantic similarity (see Bybee 2010: 36–37), semantic contiguity (as instantiated, for instance, by metonymy) may motivate the sanctioning of an item into the schematic slot of a construction (see the discussion of Figure 1).

presented at the 12th International Conference on Construction Grammar (Charles University, May 2023), and the 13th Chinese Symposium on Cognitive Linguistics (Shanghai Normal University, October 2023); we are grateful to the audience for their valuable feedback. As ever, all remaining errors are our responsibility.

Data availability: The dataset and R code are available at https://osf.io/js6qz/.

Abbreviations:

A	Adjective
BI	Comparative marker 比 *bi*
CL	Classifier
COP	Copular
DE	Clitic 得 *de*
FP	Sentence final particle
GEN	Genitive
HAI	Adverb 还 *hai*
MOD	Modifier marker
N	Noun
PERF	Perfective aspect marker

References

Barlow, Michael & Suzanne Kemmer (eds.). 2000. *Usage-based models of language*. Stanford: CSLI.

Barðdal, Jóhanna & Spike Gildea. 2015. Diachronic construction grammar: Epistemological context, basic assumptions and historical implications. In Jóhanna Barðdal, Elina Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic construction grammar*, 1–50. Amsterdam: John Benjamins.

Behrens, Heike. 2017. The role of analogy in language processing and acquisition. In Marianne Hundt, Sandra Mollin & Simone E. Pfenninger (eds.), *The changing English language: Psycholinguistic perspectives*, 215–239. Cambridge: Cambridge University Press.

Bloom, Barthe. 2021. *Life at the intersection: Two case studies of lateral relations and multiple source constructions*. Paper presented at the 11th International Conference on Construction Grammar. Antwerp.

Boas, Hans. 2003. *A constructional approach to resultatives*. Stanford: CSLI Publications.

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Cappelle, Bert. 2006. Particle placement and the case for 'allostructions'. *Constructions Special* 1. 1–26.

Cheng, Yaheng. 2013. "比X还X"构式的衍生机制与动因 [The derivative mechanisms and motivations of the structure of *bi X hai X*]. 汉语学习 [Chinese Language Learning](1). 59–64.

Dąbrowska, Eva. 2008. The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language* 58. 931–951.

Divjak, Dagmar. 2008. On (in)frequency and (un)acceptability. In Barbara Lewandowska-Tomaszczyk (ed.), *Corpus linguistics, computer tools and applications – State of the art*, 213–233. Frankfurt: Peter Lang.

De Smet, Hendrik, Frauke D'hoedt, Lauren Fonteyn & Kristel Van Goethem. 2018. The changing functions of competing forms: Attraction and differentiation. *Cognitive Linguistics* 29(2). 197–234.

Diessel, Holger. 2019. *The grammar network: How linguistic structure is shaped by language use*. Cambridge: Cambridge University Press.

Gentner, Dedre & Arthur Markman. 1997. Structure mapping in analogy and similarity. *American Psychologist* 52. 45–56.

Gentner, Dedre & Linsey A. Smith. 2012. Analogical reasoning. In Vilayanur Ramachandran (ed.), *Encyclopedia of human behavior*, 130–136. Oxford: Elsevier.

Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: The University of Chicago Press.

Goldberg, Adele. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5). 219–224.

Hilpert, Martin. 2014. *Construction Grammar and its application to English*. Edinburgh: Edinburgh University Press.

Hilpert, Martin. 2019. Higher-order schemas in morphology: What they are, how they work, and where to find them. *Word Structure* 12(3). 261–273.

Hopper, Paul. 1987. Emergent grammar. In Jon Aske, Natasha Berry, Laura Michaelis & Hana Filip (eds.), *Berkeley linguistics society 13: General session and parasession on grammar and cognition*, 139–157. Berkeley: Berkeley Linguistics Society.

Huang, Shi-Zhe, Jing Jin & Dingxu Shi. 2016. Adjectives and adjective phrases. In Chu-Ren Huang & Dingxu Shi (eds.), *A reference grammar of Chinese*, 276–296. Cambridge: Cambridge University Press.

Ji, Yimin. 2012. "X 比 N 还 N" 比较构式的认知考察 [A cognitive approach to the comparative structure *X bi N hai N*]. 语言与翻译 [Language and Translation](1). 27–31.

Krott, Andrea, Harald Baayen & Robert Schreuder. 2006. Analogy in morphology: Modeling the choice of linking morphemes in Dutch. *Linguistics* 39(1). 51–93.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar, Vol. 1: Theoretical prerequisites*. Stanford: Stanford University Press.

Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 1–63. Stanford: CSLI.

Lauwers, Peter & Dominique Willems. 2011. Coercion: Definition and challenges, current approaches, and new trends. *Linguistics* 49(6). 1219–1235.

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

Li, Ying. 2016. 现代汉语"副名"和"比 N 还 N"构式的异同分析 [Differences and similarities of nouns in "adverb + Noun" and "bi N hai N" constructions in Contemporary Chinese]. 外国语文 [Foreign Language and Literature] (5). 77–81.

Liu, Meili. 2018. 语法转喻解读的认知语用机制探究 – 以汉语"很+N"构式为例 [A study of the cognitive-pragmatic mechanism for the interpretation of grammatical metonymy: Taking the Chinese construction "hen + N" for example]. 西安外国语大学学报 [Journal of Xi'an International Studies University](1). 30–34.

Liu, Meili. 2022. Review of Sommerer & Smirnova (2020): Nodes and networks in diachronic construction grammar. *Review of Cognitive Linguistics* 20(1). 290–297.

Lorenz, David. 2020. Converging variations and the emergence of horizontal links: To-contraction in American English. In Lotte Sommerer & Elina Smirnova (eds.), *Nodes and networks in diachronic construction grammar*, 243–274. Amsterdam/Philadelphia: John Benjamins.

Lü, Shuxiang. 1999. 现代汉语八百词 [Eight hundred words in Contemporary Chinese]. Beijing: Shangwu yinshuguan [The Commercial Press].

Luo, Shenglan & Man Lu. 2016. 从构式理论研究"比N还N"结构 [A preliminary study on bi N hai N structure from construction grammar perspective]. 中国高校人文社会科学信息网 [Humanities and Social Sciences Information Network of Chinese Universities]. Available at: http://www.sinoss.net.

Ma, Weizhong. 2014. "比N还N"及其相关句式的句法、语义特点 [On the syntactic and semantic features of *bi N hai N* and its related structures]. 语言教学与研究 [Language Teaching and Linguistic Studies](6). 65–73.

Mańczak, Witold. 1980. Laws of analogy. In Jacek Fisiak (ed.), *Historical morphology*, 283–288. Berlin, New York: De Gruyter Mouton.

Michaelis, Laura. 2005. Entity and event coercion in a symbolic theory of syntax. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction grammars: Cognitive grounding and theoretical extensions*, 45–88. Amsterdam: John Benjamins.

Paris, Marie-Claude & Dingxu Shi. 2016. Comparison. In Chu-Ren Huang & Dingxu Shi (eds.), *A reference grammar of Chinese*, 297–314. Cambridge: Cambridge University Press.

Perek, Florent. 2015. *Argument Structure in usage-based construction grammar: Experimental and corpus-based perspectives*. Amsterdam & Philadelphia: John Benjamins.

Perek, Florent. 2016. Using distributional semantics to study syntactic productivity in diachrony: A case study. *Linguistics* 54(1). 149–188.

Perek, Florent. 2020. Productivity and schematicity in constructional change. In Lotte Sommerer & Elina Smirnova (eds.), *Nodes and Networks in diachronic construction grammar*, 140–166. Amsterdam: John Benjamins.

Po-Ching, Yip & Don Rimmington. 2004. *Chinese: A comprehensive grammar*. London and New York: Routledge.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London & New York: Longman.

R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing Available at: https://www.R-project.org/.

Schmid, Hans-Jörg. 2010. Does frequency in text really instantiate entrenchment in the cognitive system? In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative methods in cognitive semantics: Corpus-driven approaches*, 101–133. Berlin: De Gruyter Mouton.

Schmid, Hans-Jörg. 2020. *The Dynamics of the linguistic system: Usage, conventionalization and entrenchment*. Oxford: Oxford University Press.

Schmid, Hans-Jörg & Helmut Küchenhoff. 2013. Collostructional analysis and other ways of measuring lexicogrammatical attraction: Theoretical premises, practical problems and cognitive underpinnings. *Cognitive Linguistics* 24(3). 531–577.

Shao, Bin, Yingying Cai & Graeme Trousdale. 2019. A multivariate analysis of diachronic variation in a bunch of NOUN: A construction grammar account. *Journal of English Linguistics* 47(2). 150–174.

Sommerer, Lotte. 2015. The influence of constructions in grammaticalization: Revisiting category emergence and the development of the definite article in English. In Jóhanna Barðdal, Elina Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic construction grammar*, 107–138. Amsterdam: John Benjamins.

Sommerer, Lotte & Andreas Baumann. 2021. Of absent mothers, strong sisters and peculiar daughters: The constructional network of English NPN constructions. *Cognitive Linguistics* 32(1). 97–131.

Shi, Dingxu. 2016. Nouns and nominal phrases. In Churen Huang & Dingxu Shi (eds.), *A reference grammar of Chinese*, 199–255. Cambridge: Cambridge University Press.

Smirnova, Elina. & Lotte Sommerer. 2020. Introduction: The nature of the node and the network –Open questions in Diachronic Construction Grammar. In Lotte Sommerer & Elina Smirnova (eds.), *Nodes and networks in diachronic construction grammar*, 1–42. Amsterdam: John Benjamins.

Stefanowitsch, Anatol. 2006. Negative evidence and the raw frequency fallacy. *Corpus Linguistics and Lingistic Theory* 2(1). 61–77.

Stefanowitsch, Anatol. 2013. Collostructional analysis. In Graeme Trousdale & Thomas Hoffmann (eds.), *The Oxford handbook of construction grammar*, 290–306. Oxford: Oxford University Press.

Stefanowitsch, Anatol & Th. Stefan Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.

Traugott, Elizabeth C. & Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.

Trousdale, Graeme. 2013. Multiple inheritance and constructional change. *Studies in Language* 37(3). 491–514.

Van de Velde, Freek. 2014. Degeneracy: The maintenance of constructional networks. In Ronny Boogaart, Timothy Colleman & Gijsbert Rutten (eds.), *Extending the scope of construction grammar*, 141–180. Berlin/Boston: De Gruyter Mouton.

Van de Velde, Freek, De Smet Hendrik & Lobke Ghesquière (eds.). 2013. *On multiple source constructions in language change*. Amsterdam: John Benjamins.

Wang, Li. 2004. 汉语史稿 [Draft on the History of Chinese]. Beijing: Zhonghua shuju [Zhonghua Book Company].

Xun, Endong, Gaoqi Rao, Xiaoyue Xiao & Jiaojiao Zang. 2016. 大数据背景下BCC 语料库的研制[The construction of the BCC Corpus in the age of Big Data]. 语料库语言学 [Corpus Linguistics] 3(1). 93–118.

Yang, Yuling. 2011. "比N还N" 构式探析 [On the Construction *bi N hai N*]. 浙江学刊 [Zhejiang Academic Journal] (4). 105–109.

Zehentner, Eva & Elizabeth C. Traugott. 2020. Constructional networks and the development of benefactive ditransitives in English. In Lotte Sommerer & Elina Smirnova (eds.), *Nodes and networks in diachronic construction grammar*, 167–212. Amsterdam/Philadelphia: John Benjamins.

Zhan, Weidong, Rui Guo & Yirong Chen. 2003. The CCL corpus of Chinese texts: 700 million Chinese characters, the 11th century BC–present. Available online at the website of Center for Chinese Linguistics (abbreviated as CCL) of Peking University http://ccl.pku.edu.cn:8080/ccl_corpus.

Zhu, Dexi. 1997. 现代汉语语法研究 [A Study of Contemporary Chinese grammar]. Beijing: Shangwu Yinshu Guan.

Zong, Shouyun. 1995. "还N呢"与"比N还N"格式试析 [Analysis of "hai N ne" and "bi N hai N"]. 张家口师专学报 [Journal of Zhangjiakou Normal University] (2). 19–23.