

# Exploring the emergence of the postverbal *sin<sup>1</sup>* 先 in Cantonese

Yang Zhou

Shanghai University of Finance and Economics

*Sin<sup>1</sup>* 先 as a function word in contemporary Cantonese encodes a number of grammatical and pragmatic meanings. As its most prominent feature in syntax, it predominantly occurs in the postverbal position while indicating the meaning of ‘first’. This paper explores the emergence of the postverbal *sin<sup>1</sup>* 先 ‘first’ in Cantonese. We first examine the word order typology on the element for ‘first’ in the languages and dialects of southern coastal China. In this linguistic area, the postverbal elements for ‘first’ in Chinese dialects are contact-induced by Tai-Kadai and Hmong-Mien languages; whereas *sin<sup>1</sup>* 先 ‘first’ in the mainstream Cantonese shows a stronger tendency to be placed in the postverbal position than its counterparts in other Chinese dialects. We then discuss the word order and semantic changes of *sin<sup>1</sup>* 先 from 1820s to 1960s based on Cantonese historical materials. Besides the pressure of language contact, the formation of the postverbal *sin<sup>1</sup>* 先 ‘first’ has been further triggered by the semantic motivation to formally differentiate the ‘precedent-subsequent’ polysemy within *sin<sup>1</sup>* 先 itself. In short, the emergence of the postverbal *sin<sup>1</sup>* 先 ‘first’ in Cantonese has been a two-stage process, dually driven by external and internal causes, respectively.

**Keywords:** Cantonese, *sin<sup>1</sup>/xiān* 先, areal linguistics, language contact, word order change, semantic change

## 1. Introduction

The postverbal *sin<sup>1</sup>* 先 ‘first’ is a prominent feature in Cantonese grammar and has received much discussion on its functional variants and historical emergence.<sup>1</sup> It has been observed in previous studies that the postverbal *sin<sup>1</sup>* 先 synchronically has three distinguishable usages (e.g. Matthews & Yip 1994; Tang 2006, 2012, 2015;

---

1. Here we refer specifically to the Guangfu (廣府) subgroup of Cantonese, i.e. mainstream Cantonese, although a wider range of Cantonese dialects will be treated as our discussion goes on.

Cheung 2007; Sybesma 2013): (a) the temporal *sin*<sup>1</sup> 先 (simply *sin*<sub>1</sub> hereafter), with a high-level tone, expressing precedence in time (meaning ‘first’) as shown in (1);<sup>2</sup> (b) the imperative *sin*<sup>1</sup> 先 (*sin*<sub>2</sub> hereafter), with a high-falling tone,<sup>3</sup> expressing deontic modality (indicating speaker’s subjective prioritization of a certain matter and meaning ‘and then we will see’) as shown in (2); and (c) the emphatic *sin*<sup>1</sup> 先 (*sin*<sub>3</sub> hereafter), also with a high-falling tone, occurring only at the end of interrogative sentences as shown in (3). Apart from the tonal difference, the co-occurrence of *sin*<sub>1</sub> with either *sin*<sub>2</sub> or *sin*<sub>3</sub> can also formally identify them to be separable usages as in (4) and (5).

- (1) 我哋行先。

*Ngo*<sup>5</sup>*dei*<sup>6</sup> *haang*<sup>4</sup> *sin*<sup>1</sup>.

1.PL walk first

‘Let’s go first.’

- (2) 不如下次先啦。

*Bat*<sup>1</sup>*jyu*<sup>4</sup> *haa*<sup>6</sup> *ci*<sup>3</sup> *sin*<sup>1</sup> *laa*<sup>1</sup>.

why.not next time *sin*<sub>2</sub> SFP

‘Until the next time (and then we’ll see if we can make it).’

- (3) 邊個最靚仔先？

*Bin*<sup>1</sup>*go*<sup>3</sup> *zeoi*<sup>3</sup> *leng*<sup>3</sup>*zai*<sup>2</sup> *sin*<sup>1</sup>?

who most handsome *sin*<sub>3</sub>

‘Who is the most handsome?’

- (4) 我講先先。

*Ngo*<sup>5</sup> *gong*<sup>2</sup> *sin*<sup>1</sup> *sin*<sup>1</sup>.

1SG speak first *sin*<sub>2</sub>

‘I will speak first.’

(Tang 2015: 281)

- (5) 邊個去先先？

*Bin*<sup>1</sup>*go*<sup>3</sup> *heoi*<sup>3</sup> *sin*<sup>1</sup> *sin*<sup>1</sup>?

who go first *sin*<sub>3</sub>

‘Who will go first?’

(Tang 2006)

2. In this paper *Jyutping*, developed by the Linguistic Society of Hong Kong, is used to romanize all the contemporary and historical examples in mainstream Cantonese for the sake of uniformity even though some historical examples appeared in their own romanizations in the original texts; but the use of Chinese characters in cited examples generally, unless specified otherwise, conforms to the original texts. Besides, tones will be added on Cantonese historical examples based on contemporary reading. *Pinyin* is used for the romanization of Standard Mandarin examples as well as certain Chinese cognates. For the transcription of examples in other Chinese dialects and languages, the International Phonetic Alphabet is adopted. Cantonese examples without citations are data elicited from informants who are native speakers of Hong Kong or Guangzhou Cantonese.

3. The high-falling tone is a variant of the high-level tone in Cantonese and the two tones are not phonologically distinctive (see Gao 1980: 7; Matthews & Yip 1994: 21).

Syntactically, *sin*<sub>1</sub> is an adverb but *sin*<sub>2</sub> and *sin*<sub>3</sub> both serve as sentence-final particles. Yet semantically, the three usages are related since the imperative and emphatic usages are extensions from the temporal meaning 'first'. In other words, *sin*<sub>2</sub> and *sin*<sub>3</sub> are grammaticalized usages of *sin*<sub>1</sub> and the history of the postverbal *sin*<sup>1</sup> 先 began with *sin*<sub>1</sub>. But even considering the observed differences in tone and syntactic status, *sin*<sub>1</sub> and *sin*<sub>2</sub> still cannot be easily disentangled from each other.

Sybesma (2013) has made tremendous efforts in sorting out the similarities and differences between *sin*<sub>1</sub> and *sin*<sub>2</sub>. For one thing, the use of either *sin*<sub>1</sub> or *sin*<sub>2</sub> can be, to some extent, contextually conditioned. In his example *tung*<sup>1</sup>*zi*<sup>1</sup> *keoi*<sup>5</sup> *sin*<sup>1</sup> 通知佢先 inform-3SG-first 'inform him first' and in (1) above, either *sin*<sub>1</sub> or *sin*<sub>2</sub> is acceptable but they confer different contextual or discourse interpretations, i.e. a scenario of sequential events or an imperative meaning like 'and then we'll see'. Nevertheless, he has pointed out that a sentence ending with *sin*<sub>1</sub> does not provide the sentence with enough well-formedness while a sentence ending with *sin*<sub>2</sub> can make the sentence sufficiently well-formed. For another, the use of *sin*<sub>1</sub> or *sin*<sub>2</sub> is also grammatically constrained. By quoting Tang (2006), Sybesma (2013) has concluded that the distribution of *sin*<sub>1</sub>, as an adverb, is determined by the predicate that allows for temporal sequencing whereas the distribution of *sin*<sub>2</sub>, as a sentence-final particle, is sensitive to the type of sentence. For example, in his example *ting*<sup>1</sup>*jat*<sup>6</sup> *sin*<sup>1</sup> 聽日先 tomorrow-*sin*<sub>2</sub> 'wait until tomorrow and then we'll see' and in (2) above, *sin*<sup>1</sup> 先 can only be interpreted as *sin*<sub>2</sub>. If *sin*<sub>2</sub> is replaced with *sin*<sub>1</sub>, those sentences will become ungrammatical.

There have been varied opinions regarding the emergence of *sin*<sub>1</sub>. Cheung (2007: 203–204) has speculated that *sin*<sub>1</sub> emerged as a case of pragmatic reversion and the reversed ordering was later fossilized into a syntactic pattern due to frequent use. L. Feng (2007) has suggested that *sin*<sub>1</sub> was grammaticalized from its earlier usage as a predicate. However, neither Cheung nor L. Feng has proved his hypothesis. D. Liu (2000) regards *sin*<sub>1</sub> as one of the VO word order features in Cantonese, which are in correlation with each other and characterize Cantonese as a strong SVO dialect. But D. Liu has not provided any historical account. Hashimoto (2008[1978]: 25–28) views *sin*<sub>1</sub> as a case of the modified-modifier word order that illustrates earlier stages of the Chinese language, which is believed to have been undergoing a typological change from the Tai type to the Altaic type. In Hashimoto's argument, *sin*<sub>1</sub>, along with other linguistic features, places Cantonese closer to the Tai side on the Altaic-Tai typological continuum. Still, Hashimoto has not made it explicit whether *sin*<sub>1</sub> emerged in Cantonese due to inheritance or language contact. Li (1990) has made a clearer assertion that *sin*<sub>1</sub> is contact-induced by Tai-Kadai languages, rather than an inheritance from Ancient Chinese. More recently, Sybesma (2013) has pointed out that *sin*<sub>1</sub> replicated the word order of *gongq*, which is a post-verbal adverb meaning 'first' in Zhuang, a Tai language mainly spoken in Guangxi.

To conclude, there appears to be stronger evidence to support the view that the emergence of *sin*<sub>1</sub> in Cantonese is contact-induced.

Despite those previous discussions, the word order of *sin*<sup>1</sup> 先 ‘first’ in history remains largely unknown and there are still crucial questions to be clarified. The present paper will explore the emergence of the postverbal *sin*<sup>1</sup> 先 in Cantonese using both synchronic evidence and historical materials. By re-examining the areal linguistic features in southern coastal China, we intend to reinforce Sybesma’s (2013) view on the syntactic diffusion in Cantonese and Tai adverbial word orders and also, based on empirical data, raise a disturbing question on the weaker tendency of *xiān* 先 ‘first’ being used as a postverbal adverb among non-Cantonese dialects in Guangxi, which, however, are expected to have had more intense contact with Tai-Kadai and Hmong-Mien languages. To address this question, a new proposal is forwarded in this paper that the postverbal positioning of *sin*<sup>1</sup> 先 in Cantonese is not only induced by language contact but also syntacticalized due to semantic motivations. Our proposal is evidenced by Cantonese historical materials of the 19th and 20th centuries.

## 2. Postverbal elements for ‘first’ in an areal linguistics view

The linguistic area under current concern encompasses three southernmost coastal provinces of China, namely Guangxi (GX), Guangdong (GD), and Hainan (HN). Located at the northern corner of Mainland Southeast Asia (MSEA), it is a geographic area where Chinese, Tai-Kadai (TK, also known as Zhuàng-Dòng 壯侗 or Dòng-Tái 侗臺), Hmong-Mien (HM, also known as Miáo-Yáo 苗瑤), Tibeto-Burman (TB), Austronesian and Austroasiatic languages have been intermingling with each other. According to the *Language atlas of China: Minority languages* (Huang & Daobu 2012: 28, 216–222), Zhuang (TK), with an estimated total of 13.8 million speakers in Guangxi, is the most populous TK language in southern China. In addition to Zhuang, major non-Sinitic languages in Guangxi also include Mien (HM, 379,000), Bunu (HM, 361,000), and Kam (also known as Dòng 侗語, TK, 259,000), with TK and HM speakers making 33.7% of the overall population of Guangxi. Li (TK, 753,000), Lingao (TK, 600,000), Cun (TK, 80,000), and Mien (HM, 60,000) are the major non-Sinitic languages used in Hainan, with their speakers making 19.8% of the overall population of Hainan. However, in Guangdong speakers of non-Sinitic languages such as Zhuang (TK, 33,200), Mien (HM, 130,000), and Biao (TK, 70,000) are sporadically scattered and only make less than 0.3% of the total provincial population.

Enfield (2005: 181), with a particular reference to the linguistic situation in MSEA, claims that:

When languages are spoken historically in the same location they often show significant parallels in the organization of a wide range of structural domains, whether the languages descend from the same historical source.

Although there is an influential view among scholars in China that TK and HM belong to the Sino-Tibetan language family, in which Chinese and TB are included as core branches (see Sun et al. 2007), the genetic relationships among Chinese, TK, and HM are still much-debated issues (see Matisoff 1991; Diller et al. 2008). Hence, their genetic positioning is certainly beyond our current scope of discussion. Here we shall focus exclusively on the word order parallel of the postverbal elements for ‘first’, specifically Cantonese *sin*<sub>1</sub> and its counterparts in other Chinese dialects and non-Sinitic languages in the above-defined area.

## 2.1 Word order replication among languages of southern coastal China

In TK and HM languages of southern coastal China, most adverbs precede verb phrases (VPs) while a few, including temporal adverbs for ‘first’, usually appear in the postverbal position. Below are examples drawn from TK languages such as Zhuang, Lakkia, Kam, Mulam, Li, Lingao, and Cun, and also from HM languages such as Mien and Bunu, as shown in (6) through (15).

Northern Zhuang (Yanqi dialect in GX)

- (6) *Te*<sup>24</sup> *tup*<sup>33</sup> *ku*<sup>55</sup> *kø:n*<sup>35</sup>.

3SG hit 1SG first

‘He hit me first.’

(Wei et al. 2011: 201)

Southern Zhuang (Jingxi dialect in GX)

- (7) *Ni*<sup>5</sup> *pai*<sup>1</sup> *kø:n*<sup>5</sup> *ŋo*<sup>5</sup> *pai*<sup>1</sup> *tok*<sup>7</sup>*lan*<sup>1</sup>.

2SG go first 1SG go later

‘You go first and I will go later.’

(Zheng 2013: 266)

Lakkia (Jinxu dialect in GX)

- (8) *Lak*<sup>8</sup> *pai*<sup>1</sup> *ba:ŋ*<sup>1</sup>.

2SG go first

‘You go first.’

(Mao et al. 1982: 150)

Kam (Sanjiang dialect in GX)

- (9) *ŋa*<sup>11</sup> *pa:i*<sup>55</sup> *kun*<sup>53</sup>.

2SG go first

‘You go first.’

(Xing 1985: 272–273)

Mulam (Luocheng dialect in GX)

- (10) a. *pai*<sup>1</sup> *kun*<sup>5</sup>  
go first  
'go first'  
b. *kun*<sup>5</sup> *pai*<sup>1</sup>  
first go  
'go first'  
c. *Niu*<sup>2</sup> *tjen*<sup>1</sup> *pai*<sup>1</sup> *kun*<sup>5</sup>.  
1PL first go first  
'We will go first.'

(Wang & Zheng 1980: 53)

Li (Baisha dialect in HN)

- (11) a. *Ho*<sup>33</sup> *vou*<sup>33</sup> *k<sup>h</sup>u:ŋ*<sup>311</sup> *nɔ*<sup>33</sup> *vou*<sup>33</sup> *tui*<sup>53</sup>.  
1SG go first 3SG go later  
'I go first and he goes later.'  
b. *Mɔ*<sup>33</sup> *iu*<sup>55</sup> *p<sup>h</sup>a*<sup>55</sup> *na*<sup>55</sup> *k<sup>h</sup>u:ŋ*<sup>311</sup> *p<sup>h</sup>ɔ*<sup>33</sup>.  
2SG don't alone INTERJ first leave  
'You will not leave alone first.'

(Wang & Qian 1951)

Lingao (Lingao County dialect in HN)

- (12) *Hau*<sup>2</sup> *kɔn*<sup>1</sup> *fia*<sup>4</sup> *kua*<sup>3</sup> *na*<sup>3</sup>.  
1SG eat meal ASP first  
'I had my meal first.'

(Zhang et al. 1985: 151)

Cun (Sigeng dialect in HN)

- (13) *Mɔ*<sup>5</sup> *fo*<sup>1</sup> *k<sup>h</sup>uan*<sup>5</sup>.  
2SG go first  
'You go first.'

(Ouyang 1998: 126)

Mien (Dapingjiang dialect in GX)

- (14) *Tom*<sup>2</sup>*to:i*<sup>5</sup>*tsa:ŋ*<sup>3</sup> *ta:i*<sup>2</sup> *da:ŋ*<sup>6</sup>.  
team.leader come first  
'The team leader will come first.'

(Mao et al. 1982: 57)

Bunu (Meizhu dialect in GX)

- (15) *Kau*<sup>2</sup> *muj*<sup>4</sup> *te*<sup>2</sup>.  
2SG go first  
'You go first.'

(Mao et al. 1982: 97)

Such a word order feature can also be found in Jing, an Austroasiatic language mainly spoken in southern Guangxi and in Huihui, an Austronesian language mainly spoken in southern Hainan, as shown in (16) and (17), respectively.

Jing (Jiangping dialect in GX)

- (16) *Mai*<sup>2</sup> *di*<sup>1</sup> *tuək*<sup>7</sup>.

2SG go first

'You go first.'

(Ouyang et al. 1984: 113)

Huihui (Huihui Village dialect in HN)

- (17) *Ha*<sup>33</sup> *pha*<sup>43</sup> *lau*<sup>11</sup> *kau*<sup>33</sup> *pha*<sup>43</sup> *ku*<sup>24</sup>.

2SG go first 1SG go later

'You go first and I will go later.'

(Zheng 1997: 83)

Building on the data and previous literature, there are some aspects worth noting here.

First of all, in most of the above-listed languages including Zhuang, Lakkia, Kam, Lingao, Cun, Mien, Bunu, Jing, and Huihui, the temporal adverbs for 'first' are claimed by field linguists to appear predominantly, if not strictly, in the postverbal position; whereas there are undeniable cases of variations in some TK languages. For example, the TK adverb for 'first' can occur either preverbally or postverbally in Mulam and Li, signaling an ongoing word order change. Wang & Zheng (1980: 53, 66) note that Mulam not only grants *kun*<sup>5</sup> 'first' two syntactic positions but also uses *tjen*<sup>1</sup> 先 'first', which is a borrowed adverb from Chinese, in the preverbal position as in (10). In TK and HM, the use of the postverbal adverb for 'first' usually does not take the occurrence of another postverbal adverb for 'later' as a prerequisite.<sup>4</sup> However, considering the Li examples provided in Wang & Qian (1951), *k<sup>h</sup>u:ŋ*<sup>311</sup> 'first' tends to occur postverbally when coupled with the postverbal *tui*<sup>53</sup> 'later' as in (11a) while in other contexts without *tui*<sup>5</sup> 'later' the postverbal restriction on *k<sup>h</sup>u:ŋ*<sup>311</sup> 'first' is relaxed as shown in (11b).

Second, in those languages the postverbal adverbs for 'first' usually have their lexical source as a locative noun (or morpheme) for 'front', such as Zhuang *kø:n*<sup>35</sup>/*ko:n*<sup>5</sup>, Kam *kun*<sup>53</sup>, Lingao *na*<sup>3</sup>, Mien *da:ŋ*<sup>6</sup>, Jing *tuək*<sup>7</sup>, Huihui *lau*<sup>11</sup> (pronounced as *nau*<sup>24</sup> to mean 'front'), and so on. Interestingly, it is a striking parallel that the temporal adverbs for 'later' are also frequently observed in the postverbal position, such as southern Zhuang *tok*<sup>7</sup>*lan*<sup>1</sup> 'later' in (7), Lakkia *la*<sup>3</sup> 'later' in *tsi*<sup>1</sup> *tan*<sup>2</sup> *la*<sup>3</sup> 1SG-come-later 'I come later' (Mao et al. 1982: 150), Kam *ta*<sup>33</sup>*lan*<sup>11</sup> 'later' in *ma*<sup>35</sup> *ta*<sup>33</sup>*lan*<sup>11</sup> come-later 'come later' (Xing 1985: 276), Cun *tsuən*<sup>3</sup> 'later' in *kə*<sup>5</sup> *zai*<sup>3</sup> *tsuən*<sup>3</sup> 1SG-go-later 'I go later' (Ouyang 1998: 126), and Huihui *ku*<sup>24</sup> 'later' in (17). Moreover, all of these adverbs for 'later' are traceable to a locative for 'back'. Thus,

4. Although some of our chosen examples show the coupling of two postverbal adverbs for 'first' and 'later', there is not necessarily a co-occurrence relation between the two adverbs. Thanks to a reviewer who suggested a clarification on our choice of examples.

it posits a parallel grammaticalization path of these opposite locatives, which instantiates a more general process of spatial concepts extending to temporal realms.

Lastly, as far as our data allow us to see, the TK and HM postverbal temporal adverbs for ‘first’ do not show the functional extension as reflected by *sin*<sub>1-3</sub> in Cantonese. Nevertheless, those functions may be substantiated by other forms. For example, according to our informants on Nanning Zhuang (a northern Zhuang dialect), a sentence-final particle *ca:i*<sup>35</sup>*ka:ŋ*<sup>55</sup> is used as an imperative marker (e.g. *muŋ*<sup>42</sup> *po:i*<sup>24</sup> *kø:n*<sup>35</sup> *ca:i*<sup>35</sup>*ka:ŋ*<sup>55</sup> you-go-first-IMP ‘just go there first (and we’ll see what to do next)’ in equivalent to *sin*<sub>2</sub>. As another interesting areal feature, the southern Zhuang *ca:i*<sup>35</sup>*ka:ŋ*<sup>55</sup>, which is grammaticalized from a VP consisted of the preverbal adverb *ca:i*<sup>35</sup> meaning ‘again’ and the verb *ka:ŋ*<sup>55</sup> meaning ‘to say’, shares the same internal structure and grammatical meaning as the sentence-final particle *zàishuō* 再說 in Mandarin Chinese (e.g. *huā chū qù zàishuō* 花出去再說 spend-out-go-IMP ‘just spend the money (and we’ll see what to do next)’ further see Peng (2014) for a diachronic account on *zàishuō* 再說).<sup>5</sup>

As for Chinese, according to the *Linguistic atlas of Chinese dialects: Grammar* (Cao 2008)’s Map 84, the occurrence of a postverbal element for ‘first’ is almost ubiquitously seen on Chinese dialect sites under survey in Guangxi, Hainan, and western Guangdong. Therefore, such a word order is not just present in the mainstream Cantonese but also widespread in a range of other Chinese dialects, including Pinghua, Tuhua, Hakka, Southwest Mandarin, and Min, as shown in (18) through (24).

Cantonese (Beihai dialect in GX)

- (18) a. 你行先。  
       *Nei*<sup>13</sup> *haŋ*<sup>21</sup> *tin*<sup>55</sup>.  
       2SG go first  
       ‘You go first.’  
   b. 你去頭先。  
       *Nei*<sup>13</sup> *hui*<sup>33</sup> *tʰeu*<sup>21</sup>*tin*<sup>55</sup>.  
       2SG go first  
       ‘You go first.’

(Chen & Chen 2005: 381)

5. There are basically two sources to derive a sentence-final imperative marker as observed in Chinese and Zhuang, one being a postverbal temporal adverb for ‘first’, which characterizes Cantonese, and the other being a VP meaning ‘to say again’, which is represented by Mandarin Chinese and Zhuang. But temporarily, we shall not go into details on the parallel grammaticalization of the Zhuang *ca:i*<sup>35</sup>*ka:ŋ*<sup>55</sup> and the Mandarin *zàishuō* 再說, which might have involved language contact as a crucial factor. Also due to our limited knowledge of TK and HM, we are currently not able to conclude how TK and HM differ in their expressions of the imperative and emphatic meanings. These two aspects deserve some future attention.

Pinghua (Xinhe dialect in GX)

- (19) 你去先。  
*Noi*<sup>21</sup> *hoi*<sup>35</sup> *tin*<sup>55</sup>.  
 2SG go first  
 ‘You go first.’ (Liang & Lin 2009: 363)

Tuhua (Lianzhou dialect in GD)

- (20) 等渠食先。  
*Tei*<sup>53</sup> *ki*<sup>24</sup> *sek*<sup>22</sup> *seŋ*<sup>22</sup>.  
 wait 3SG eat first  
 ‘Let him eat first.’ (Chang 2004: 360)

Hakka (Bobai dialect in GX)

- (21) 你行先。  
*Ni*<sup>23</sup> *haŋ*<sup>23</sup> *sen*<sup>44</sup>.  
 2SG go first  
 ‘You go first.’ (X. Chen 2004: 351)

Southwest Mandarin (Liuzhou dialect in GX)

- (22) 他來先，你來後，給他站到前頭。  
*T<sup>h</sup>a*<sup>44</sup> *læ*<sup>31</sup> *sẽ*<sup>44</sup>, *ni*<sup>54</sup> *læ*<sup>31</sup> *hɐu*<sup>24</sup>, *kei*<sup>54</sup> *t<sup>h</sup>a*<sup>44</sup> *tsã*<sup>24</sup> *ta*<sup>31</sup> *ts<sup>h</sup>ẽ*<sup>31</sup> *t<sup>h</sup>ɐu*<sup>31</sup>.  
 3SG come first 2SG come later give 3SG stand to front  
 ‘He came first and you came later. So let him stand in the front.’  
 (C. Liu 1995: 272)

Min (Leizhou dialect in GD)

- (23) 我去過前哪。  
*Ba*<sup>52</sup> *k<sup>h</sup>u*<sup>21</sup> *kue*<sup>55</sup> *tsai*<sup>33</sup> *na*<sup>55</sup>.  
 1SG go ASP first SFP  
 ‘I go first.’ (L. Lin 2006: 232)

Min (Qionghai dialect in HN)

- (24) 汝去過前，我農等歇罔才去。  
*Lu*<sup>31</sup> *xu*<sup>55</sup> *kie*<sup>55</sup> *tai*<sup>33</sup>, *gua*<sup>21</sup> *naŋ*<sup>22</sup> *dan*<sup>21</sup> *heŋ*<sup>53</sup> *kia*<sup>31</sup> *na*<sup>33</sup> *xu*<sup>213</sup>.  
 2SG go ASP first 1.PL wait moment then go  
 ‘You go first and we will go after a while.’ (C. Feng 1992)

Also of these Chinese dialects, there are several aspects to be noted in contrast to the neighboring TK and HM languages.

First, there are two types of postverbal adverbs for ‘first’ under discussion: one being cognate with *xiān* 先 ‘first’ and the other being derived from a locative meaning ‘front’. Unlike the widely perceived postverbal restriction on the adverbs for ‘first’ among TK and HM, it is common of *xiān* 先 ‘first’ to flexibly occur either preverbally or postverbally in the above-listed Chinese dialects. As for illustrations, Beihai Cantonese *tin*<sup>55</sup> 先 ‘first’ is also used preverbally in *nei*<sup>13</sup> *tin*<sup>55</sup> *haŋ*<sup>21</sup>

2SG-first-go 'you go first' (Chen & Chen 2005: 381); Pinghua *lin*<sup>55</sup> 先 'first' in *noi*<sup>21</sup> *lin*<sup>55</sup> *hoi*<sup>35</sup> 2SG-first-go 'you go first' (Liang & Lin 2009: 363); Tuhua *seŋ*<sup>22</sup> 先 'first' in *tei*<sup>53</sup> *ki*<sup>24</sup> *seŋ*<sup>22</sup> *sek*<sup>22</sup> wait-3SG-first-eat 'let him eat first' (Chang 2004: 360). On the contrary, those locative-derived adverbs for 'first', such as Beihai Cantonese *t<sup>h</sup>eu*<sup>21</sup>*lin*<sup>55</sup> (頭先) 'front/first' in (18b), Leizhou Min *tsai*<sup>33</sup> (前) 'front/first' in (23) and Qionghai Min *tai*<sup>33</sup>(前) 'front/first' in (24), are used strictly in the postverbal position like their counterparts in TK and HM.

Secondly, unlike TK and HM that are characterized by a pair of postverbal adverbs to express both 'first' and 'later', Chinese dialects of southern coastal China are less commonly observed to show such a similar pattern of postverbal temporal expressions. For example, Liuzhou Mandarin *heu*<sup>24</sup> (後) 'later' can appear in the postverbal position to couple with a postverbal *sẽ*<sup>44</sup> 先 'first' in (22).<sup>6</sup> But it is not a common phenomenon in many other Chinese dialects. Therefore, the postverbalization of temporal adverbs in those Chinese dialects does not show a well-established and coordinated pattern as it does in neighboring TK and HM languages.

Thirdly, postverbal adverbs for 'first' are neither historically attested in Chinese nor geographically wide-spread in its dialects. *Xiān* 先 'first', a pictographic character that originally meant 'to precede, or to go forward' according to *Shuowen jiezi* 說文解字 *Explaining graphs and analyzing characters*,<sup>7</sup> had already been used as a preverbal temporal adverb for 'first' as early as the Pre-Qin period (before 221 BC).<sup>8</sup> But it is a generally acknowledged fact that *xiān* 先 'first' had never been a postverbal adverb in the history of Chinese (Lucas & Xie 1994: 200). The modern Chinese dialects that use postverbal adverbs for 'first' are geographically clustered mainly in, or close to, southern coastal China. As a noteworthy fact, Southwest Mandarin and Min dialects are not only spoken within the area but also used by an even larger population of speakers in the further north. However, postverbal

6. Due to lack of detailed documentation, it remains an intriguing question for further exploration whether the postverbalization of *heu*<sup>24</sup> (後) 'later' is only made possible by the co-occurrence of the postverbal *sẽ*<sup>44</sup> 先 'first' in Liuzhou Mandarin and that to what extent and under what condition the adverb for 'later' can be placed in the postverbal position in other Chinese dialects.

7. It is a dictionary compiled by a Han dynasty scholar named Xu Shen (許慎, 58–149) that analyzes the structure and etymology of each of 9,353 Chinese characters.

8. For example, *xiān* 先 was used as a preverbal adverb for 'first' in the sentence *xiān shēng zhě xiān sǐ* 先生者先死 first-be.born-person-first-die 'those who are born first will die first' in *The Mozi* 墨子, an ancient Chinese philosophy book compiled during the Warring States period (476–211 BC). A reviewer questioned the lexical source of *xiān* 先 'first'. We suggest that it was derived from its verbal meaning of 'to precede' in Pre-Qin Chinese. Since its grammaticalization is not related to our current concern, it may be reserved for discussion under another topic.

adverbs for 'first' are a rare phenomenon in Southwest Mandarin and Min dialects away from this area of concern.<sup>9</sup> On the basis of the comparison made above, the postverbal positioning of adverbs for 'first' cannot be considered a grammatical feature that was rooted in the Chinese language and its presence in the southern Chinese dialects is most likely to have been contact-induced by neighboring TK and HM languages.

To follow the terminology in Heine & Kuteva (2005),<sup>10</sup> what we have in the linguistic area of southern coastal China are recurring cases of grammatical replication of the word order characteristic of postverbal elements for 'first', whereby TK and HM languages serve as the model languages<sup>11</sup> while Chinese dialects are the replica languages. As suggested by sources for 'first', two types of grammatical replication can be easily identified in those Chinese dialects, one type being contact-induced grammaticalization, specifically replica grammaticalization, and the other being restructuring, whose distributions are highly regional.

The first type is featured with the locative-temporal *tau*<sup>4</sup>*sin*<sup>1</sup> (頭先) 'front/first' in the Qin-Lian (欽廉), Gao-Yang (高陽) and Wu-Hua (吳化) subgroups

---

9. Nonetheless, some Hakka, Min, and Wu dialects, although far away from the southern coast of China, can still have the temporal adverb *xiān* 先 'first' being used in the postverbal position. We admit that such a word order in those dialects is also contact-induced except the influence coming from neighboring Chinese dialects rather than any non-Sinitic language. The point will be addressed later in this paper.

10. Heine and Kuteva have been developing a theoretical framework to account for contact-induced grammatical changes in a series of papers and books (e.g. Heine & Kuteva 2003, 2005, 2006; Heine 2008, etc.). In their framework, model language (M) and replica language (R) are the terms for the languages being, respectively, the source and the target in contact-induced transfer. Borrowing and replication are differentiated in a way that the former is a type of contact-induced transfer involving phonetic substance while the latter contrastively not involving phonetic substance of any kind. And by definition, grammatical replication represents a process whereby speakers, by using linguistic material available in language R, create a new grammatical meaning or structure in R on the model of language M. To further understand how grammatical meanings and structures are replicated, a dichotomy is proposed in Heine & Kuteva (2005) to recognize contact-induced grammaticalization and restructuring as two types of processes in grammatical replication. Contact-induced grammaticalization is further divided into ordinary grammaticalization and replica grammaticalization depending on whether or not the whole grammaticalization process is transferred from M to R. Restructuring is a term used to cover all the cases of grammatical replication that are not related to grammaticalization. However, even concerning word order rearrangement, which is usually regarded to represent the diagram case of restructuring, this dichotomy is not clear-cut in that there are cases of restructuring as by-product of grammaticalization, which is notably true in our study.

11. We cannot exclude the possibility that even within TK or HM such postverbal adverbs are contact-induced areal linguistics features. However, this is not the primary concern of this paper.

of Cantonese and *t(s)ai*<sup>33</sup> (前) ‘front/first’<sup>12</sup> in the Leizhou (雷州) and Qionghai (琼海) subgroups of Min. These dialects are geographically concentrated along the two sides of the Qiongzhou Channel. The functional and syntactic resemblance among Cantonese *tau*<sup>4</sup>*sin*<sup>1</sup> (頭先), Min *t(s)ai*<sup>33</sup> (前), Zhuang *kø:n*<sup>35</sup>, Mien *da:ŋ*<sup>6</sup> and their counterparts in many other TK and HM languages is suggestive of a grammaticalization path of ‘front’ > ‘first’ that has been replicated by those Chinese dialects. The general mechanism of replica grammaticalization is sketched in Heine & Kuteva (2003):

- a. Speakers of language R notice that in language M there is a grammatical category Mx.
- b. They develop an equivalent category Rx, using material available in their own language (R).
- c. They replicate a grammaticalization process they assume to have taken place in language M, using an analogical formula of the kind [My > Mx] = [Ry > Rx].
- d. They grammaticalize category Ry to Rx.

In our case, Chinese speakers on both sides of the Qiongzhou Channel notice that in TK and HM (M) a locative for ‘front’ (My) is also used as a postverbal element for ‘first’ (Mx). As a result of intense contact among Chinese, TK and HM, Cantonese and Min (R) have replicated the grammaticalization path [My > Mx] using their own postverbal locatives *tau*<sup>4</sup>*sin*<sup>1</sup> (頭先) ‘front’ and *t(s)ai*<sup>33</sup> (前) ‘front’ [Ry > Rx]. Thus, what has been introduced into those Chinese dialects is not only a new grammaticalized meaning based on ‘front’ but also a new word order for the expression of ‘first’.

The second type, representing a case of restructuring, sees the word order change of an existing adverb *xiān* 先 ‘first’. This word order change has been so extensively practiced as to involve most Chinese dialects in Guangxi and western Guangdong. Given the vast linguistic area with such a dynamic complexity of languages and dialects, areal diffusion within Chinese dialects should be an undeniable issue. It is not on every Chinese dialect site that a process of grammatical replication takes place in the direct model of a TK or HM language, especially considering the fact that TK and HM are much less populous in Guangdong. Thus, there must have been different channels of diffusion in Guangxi and western Guangdong.

In Guangxi, Chinese dialects, such as Cantonese, Pinghua, Tuhua, Hakka, and Southwest Mandarin, are more closely intermingled with TK and HM that serve as model languages. The extensive and prolonged contact among the languages resulted in an attempt of Chinese speakers to assimilate the word order of Chinese

---

12. The morpheme *qián* (前) ‘front/first’ is pronounced differently as *tsai*<sup>33</sup> in Leizhou Min or *tai*<sup>33</sup> in Qionghai Min. Thus, we use the transcription *t(s)ai*<sup>33</sup> to represent the morpheme in these Min dialects.

dialects (R) to that of TK and HM (M). In order to replicate the word order of TK and HM's postverbal adverb for 'first', those Chinese speakers chose to move an existing preverbal equivalent, such as Cantonese *sin<sup>1</sup>* 先 'first' and Pinghua *tin<sup>55</sup>* 先 'first', directly to the postverbal position rather than creating a new postverbal adverb for 'first' by grammaticalization.

Meanwhile, in western Guangdong where Cantonese dialects, especially the Guangfu Cantonese as particularly concerned in this paper, are not equally exposed to TK and HM, this areally diffused feature is therefore most likely to have entered Guangdong Cantonese through the channel of neighboring Chinese dialects, especially Guangxi Cantonese. Given their geographic adjacency and historical relationship, Guangdong Cantonese (R) is believed to have rearranged the word order of *sin<sup>1</sup>* 先 'first' mainly in the model of Guangxi Cantonese (M). In a simplified picture, we regard Guangdong Cantonese as a whole. But we also note that even within Guangdong Cantonese the diffusion has actually gradually made its way eastward to influence most Cantonese dialects in western Guangdong.

However, if areal diffusion is taken into serious consideration, a disturbing question will soon arise.

## 2.2 A disturbing question

In order to put this syntactic diffusion under closer scrutiny, a word order typology of *xiān* 先 'first' in southern coastal Chinese dialects is worth special discussion here. We shall first generalize the possible word order types in terms of an adverb and a VP. Probabilistically, an adverb can either precede or follow a VP to render three types of word order including AdvVP, VPAdv, and AdvVPAdv, the last being a doubling construction often indicative of an ongoing word order change. The three types of word order further predict the following six possible patterns of occurrence in any particular language.<sup>13</sup>

---

13. Here we actually exclude one word order type that predicts an adverb to be placed in the middle of a VP. Some critics may claim that Cantonese *do<sup>1</sup>* 多 'much/many' and *siu<sup>2</sup>* 少 'little/few' may occupy the mid-VP position such as in *se<sup>2</sup> do<sup>1</sup> geoi<sup>3</sup>* 寫多句 write-many-sentence 'write one more sentence' and *gaau<sup>1</sup> siu<sup>2</sup> saam<sup>1</sup> gin<sup>6</sup>* 交少三件 hand.in-few-three-piece 'hand in three pieces less'. But we regard *do<sup>1</sup>* 多 'much/many' and *siu<sup>2</sup>* 少 'little/few' as adjectives and their postverbal usages as verbal complements or particles rather than adverbials (also see Sybesma 2013). They are therefore an untypical case to illustrate the above word order type. Moreover, the above-listed three word order types and six occurrence patterns are not meant to be cross-linguistically exhaustive but mainly used to analyze the word order variations of *xiān* 先 'first'. Since *xiān* 先 'first' has not been observed to occupy a mid-VP position in any Chinese dialect, such a word order type is irrelevant to our current discussion. Thanks to the reviewer who suggested a footnote on this point.

Pattern (i)	VPAdv	Pattern (iv)	AdvVP; AdvVPAdv; VPAdv
Pattern (ii)	VPAdv; AdvVPAdv	Pattern (v)	AdvVP; AdvVPAdv
Pattern (iii)	AdvVP; VPAdv	Pattern (vi)	AdvVP

Based on the patterns above, let's now treat the word order variations of *xiān* 先 'first' in Chinese dialects of Guangxi and Guangdong, by analyzing the data provided in Map 84 of the *Linguistic atlas of Chinese dialects: Grammar* (Cao 2008).<sup>14</sup> Map 84, covering a total of 161 survey sites in Guangxi (66 sites) and Guangdong (95 sites), indicates the presence of a postverbal *xiān* 先 'first' on 108 survey sites (61 in GX and 47 in GD). All the above-listed six patterns can be observed and their typological implications are explained as follows.

Pattern (i). With *xiān* 先 'first' mostly positioned postverbally, VP+*xiān* 先 is the dominant word order observed on 45 dialect sites, the majority being Cantonese sites in Guangdong.

Pattern (ii). The patterning of VP+*xiān* 先 and *xiān* 先+VP+*xiān* 先 exists in four dialect sites in Guangxi, i.e. Cenxi and Tengxian Cantonese, and Pingguo and Tiandong Pinghua. In this pattern the occurrence of a preverbal *xiān* 先 'first' usually presupposes another postverbal *xiān* 先 'first', but not vice versa. Thus, in these dialects *xiān* 先 'first' shows a strong preference to be used postverbally.

Pattern (iii–iv). P(iii) has *xiān* 先+VP and VP+*xiān* 先 while P(iv) further includes *xiān* 先+VP+*xiān* 先. The two patterns are represented by 58 dialect sites, the majority being sites in Guangxi. Qualitatively speaking, *xiān* 先 'first' does not show any preference to precede or follow a VP in either of the two patterns.

Pattern (v). The patterning of *xiān* 先+VP and *xiān* 先+VP+*xiān* 先 is represented by only one dialect site, i.e. Haifeng Min in Guangdong. In this pattern the occurrence of a postverbal *xiān* 先 'first' usually presupposes another preverbal *xiān* 先 'first', but not verse versa. Thus, *xiān* 先 'first' shows a strong preference to be used preverbally.

Pattern (vi). With *xiān* 先 'first' mostly positioned preverbally, *xiān* 先+VP is the dominant word order observed on 40 dialect sites, the majority being Tuhua, Hakka and Min sites spread over northern and eastern Guangdong.

The above-listed patterns establish a word order hierarchy of *xiān* 先 'first' from being strongly postverbal to strongly preverbal. Building on the qualitative features gathered from each individual dialect site, we are able to come up with some quantitative tendencies across the linguistic area.

14. We assume the word order(s) presented in Map 84 as the basic, or at least commonly used, form(s) on each dialect site.

Figure 1, with a scale of Patterns (i–vi), categorizes a total of 148 Chinese dialect sites in Guangxi and Guangdong.<sup>15</sup> Columns above and below the 0-axis represent, respectively, numbers of Cantonese and non-Cantonese sites.

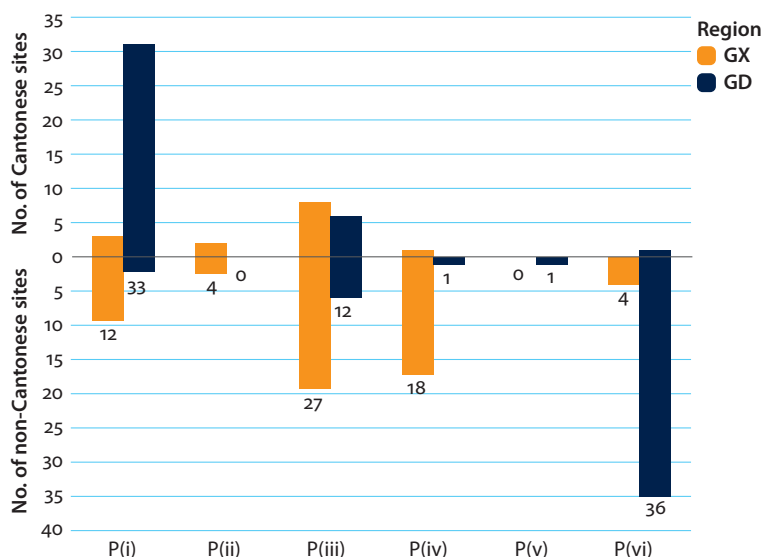


Figure 1. Word order patterns of *xiān* 先 ‘first’ in relation to a VP in Chinese dialects of Guangxi and Guangdong

Since Patterns (i–v) all involve the use of a postverbal *xiān* 先 ‘first’, Figure 1, with a panoramic survey, largely substantiates our previous insight that a postverbal *xiān* 先 ‘first’ is an areal grammatical feature not just extensively represented by Cantonese but also by other Chinese dialects in Guangxi and Guangdong.

However, a conspicuous contrast can be seen between Chinese dialects in the two provinces. Most dialect sites in Guangxi are categorized into Patterns (iii) and (iv), which means that Guangxi dialects as a whole, situated in the middle of the hierarchy, do not exhibit a strong preference to place *xiān* 先 ‘first’ strictly in the postverbal position. In other words, most Guangxi dialects grant *xiān* 先 ‘first’ a relatively free variation of word order between preverbal and postverbal, which can readily be interpreted as an ongoing word order change induced by the contact with the neighboring TK and HM languages.

15. For lack of information on the positioning of *xiān* 先 ‘first’, another 13 dialect sites tagged with postverbal *tsai*<sup>33</sup> 前 ‘first’, *tau*<sup>4</sup>*sin*<sup>1</sup> 頭先 ‘first’ or *jeng*<sup>3</sup> 正 ‘first’ in Guangxi and Guangdong have not been included in Figure 1.

On the contrary, Guangdong dialects do not seem to behave alike with respect to the word order of *xiān* 先 ‘first’ because most dialect sites are densely distributed on the two extremes of the hierarchy, which exhibits a geographical as well as dialectal split: the Cantonese sites in western Guangdong, mainly representing Pattern (i), have the strongest favor of a postverbal *xiān* 先 ‘first’ while the non-Cantonese sites in northern and eastern Guangdong, mainly categorized under Pattern (vi), only use a preverbal *xiān* 先 ‘first’. The split is determined by the intensity of contact with TK and HM. Cantonese dialects in western Guangdong, though not in direct and extensive confrontation with TK and HM, are geographically close to Guangxi and Hainan so as to be prone to this syntactic diffusion, not to mention the direct influence from Guangxi Cantonese. But non-Cantonese dialects in northern and eastern Guangdong are too distant to have so far been involved and affected in this syntactic diffusion.

Once intensity of contact is taken into consideration, a disturbing question soon arises from reading the data in Figure 1. Given the speaker population and geographic distribution of TK and HM, Chinese dialects in Guangxi are supposed to have been involved in more intense contact with TK and HM. However, it is surprisingly Guangdong Cantonese that shows a stronger tendency to place *sin<sup>1</sup>* 先 ‘first’ in the postverbal position. If we regard Guangxi as the epicenter of this syntactic diffusion, it would be extremely difficult to understand why the dialects away from Guangxi have been affected the most. Whereas, if we assume that western Guangdong used to be inhabited by a much greater number of TK and HM speakers and had served as the historical epicenter of this syntactic diffusion, it would also make it very questionable that non-Cantonese dialects in northern and eastern Guangdong have barely assimilated such a word order feature. This disturbing question has therefore put Guangdong Cantonese, especially mainstream Cantonese represented by the Guangfu subgroup, under the spotlight.

Although language contact has induced the word order change of *xiān* 先 ‘first’ in many southern Chinese dialects, it is not a sufficient explanation for the postverbal restriction of *sin<sup>1</sup>* 先 ‘first’ in mainstream Cantonese. Hence, we have to take internal motivation into consideration while revisiting the history of Cantonese *sin<sup>1</sup>* 先 ‘first’. The next section will examine the word order change of *sin<sup>1</sup>* 先 ‘first’ in Cantonese historical materials dated from the early 19th to the mid-20th centuries.

### 3. Word order change of *sin<sup>1</sup>* 先: Preverbal versus postverbal

A less explored aspect of Cantonese *sin<sup>1</sup>* 先 is its word order in history. To reconstruct its historical development, we shall recruit and analyze historical data mainly from three online Cantonese corpuses, i.e. *Early Cantonese colloquial texts*:

A database (1828–1924) (ECCT hereafter),<sup>16</sup> *Early Cantonese tagged database* (1872–1931) (ECTD),<sup>17</sup> and *A linguistic corpus of Mid-20th Century Hong Kong Cantonese* (1952–1966) (CHKC).<sup>18</sup>

Contrary to its frequent use in contemporary Cantonese, *sin<sub>1</sub>* had once been a rare usage in early Cantonese. As the earliest literature included in ECCT, *Vocabulary of the Canton dialect* (Morrison 1828) has eight occurrences of the preverbal *sin<sup>1</sup>* 先 ‘first’ as shown in (25) and (26) but does not contain a single occurrence of *sin<sub>1</sub>*.

- (25) 先學個的。  
*Sin<sup>1</sup> hok<sup>6</sup> go<sup>2</sup> di<sup>1</sup>.*  
 first learn that CL  
 ‘First learn those.’ (Morrison 1828; ECCT)
- (26) 遲來先上岸。  
*Ci<sup>4</sup> lai<sup>4</sup> sin<sup>1</sup> soeng<sup>5</sup> ngon<sup>6</sup>.*  
 late come first get.on shore  
 ‘Late to come, first to get up on the shore.’ (Morrison 1828; ECCT)

The first appearance of *sin<sub>1</sub>* has been found in *A Chinese chrestomathy in the Canton dialect* (Bridgman 1841) with only two occurrences as shown in (27) and (28) while the other 39 occurrences of *sin<sup>1</sup>* 先 ‘first’ all being preverbal such as in (29).

- (27) 唐人點樣初學寫字呢？請先生把手執筆寫紅硃字先咯。  
*Tong<sup>4</sup>jan<sup>4</sup> dim<sup>2</sup>joeng<sup>6</sup> co<sup>1</sup> hok<sup>6</sup> se<sup>2</sup> zi<sup>6</sup> ne<sup>1</sup>? Cing<sup>2</sup>*  
 Chinese how initially learn write character SFP invite  
*sin<sup>1</sup>saang<sup>1</sup> baa<sup>2</sup> sau<sup>2</sup> zap<sup>1</sup> bat<sup>1</sup> se<sup>2</sup> hung<sup>4</sup>zyu<sup>1</sup> zi<sup>6</sup> sin<sup>1</sup> lo<sup>1</sup>.*  
 teacher hold hand hold pen write red character first SFP  
 ‘How do Chinese begin to learn writing characters? They will invite a teacher to guide the child’s hand, holding the pen, to copy red characters.’  
 (Bridgman 1841; ECCT)
- (28) 執住個枝筆每個字唔知寫邊的先呢。  
*Zap<sup>1</sup> zyu<sup>6</sup> go<sup>2</sup> zi<sup>1</sup> bat<sup>1</sup> mui<sup>5</sup> go<sup>3</sup> zi<sup>6</sup> m<sup>4</sup> zi<sup>1</sup> se<sup>2</sup> bin<sup>1</sup>*  
 hold ASP that CL pen each CL character not know write which  
*di<sup>1</sup> sin<sup>1</sup> ne<sup>1</sup>.*  
 CL first SFP  
 ‘Holding the pen, I don’t know which part of the character should be written first.’  
 (Bridgman 1841; ECCT)

16. Website: <http://pvs0001.ust.hk/Candbase>.

17. Website: <http://pvs0001.ust.hk/WTagging>.

18. Website: <http://corpus.ied.edu.hk/hkcc/corpus>.

- (29) 起首點樣學呢？先學講說話咯。

Hei<sup>2</sup>sau<sup>2</sup> dim<sup>2</sup>joeng<sup>6</sup> hok<sup>6</sup> ne<sup>1</sup>? Sin<sup>1</sup> hok<sup>6</sup> gong<sup>2</sup> syut<sup>3</sup>waa<sup>6</sup> lo<sup>1</sup>.

initially how learn SFP first learn speak conversation SFP

'How will I begin to learn? First learn to speak.' (Bridgman 1841; ECCT)

An interesting comparison can be made between two editions of *How to speak Cantonese* (Ball 1902; 1912) to capture the ongoing word order change of *sin<sup>1</sup>* 先 'first'. Compare the positioning of *sin<sup>1</sup>* 先 'first' in (30) and (31).

- (30) 你先抹烟塵 𨋖 𨋖 先掃呢？

Nei<sup>5</sup> sin<sup>1</sup> maat<sup>3</sup> jin<sup>1</sup>can<sup>4</sup> pei<sup>2</sup> gu<sup>1</sup>lei<sup>1</sup> sin<sup>1</sup> sou<sup>3</sup> ne<sup>1</sup>?

2SG first wipe dust or coolie first sweep SFP

'Did you dust the room first, or did the coolie sweep the floor first?'

(Ball 1902; ECTD)

- (31) 你抹烟塵 先 𨋖 𨋖 掃地 先 呢？

Nei<sup>5</sup> maat<sup>3</sup> jin<sup>1</sup>can<sup>4</sup> sin<sup>1</sup> pei<sup>2</sup> gu<sup>1</sup>lei<sup>1</sup> sou<sup>3</sup> dei<sup>6</sup> sin<sup>1</sup> ne<sup>1</sup>?

2SG wipe dust first or coolie sweep floor first SFP

'Did you dust the room first, or did the coolie sweep the floor first?'

(Ball 1912; ECTD)

The occurrences of *sin<sup>1</sup>* 先 'first' were consistently preverbal throughout the 2nd edition (Ball 1902) as exhibited by *sin<sup>1</sup>* 先 'first' preceding the VPs *maat<sup>3</sup> jin<sup>1</sup>can<sup>4</sup>* 抹烟塵 'wipe dust' and *sou<sup>3</sup>* 掃 'sweep' in (30). However, the same sentence had been revised in the 4th edition (Ball 1912), where the two previous occurrences of the preverbal *sin<sup>1</sup>* 先 'first' were amended to appear as *sin<sub>1</sub>* as in (31). This revision, may it be minor, makes itself a sign indicating that *sin<sub>1</sub>* was gaining acceptance among Cantonese speakers at that time. As observed from all the early Cantonese materials in ECCT and ECTD, the preverbal word order of *sin<sup>1</sup>* 先 'first' had remained dominant through 1820s to 1930s.

It was not until 1940s that *sin<sub>1</sub>* had become a major use pattern such as in *First year Cantonese: Part one* (O'Melia 1941), where there were six occurrences of the preverbal *sin<sup>1</sup>* 先 'first' versus eight occurrences of *sin<sub>1</sub>*. O'Melia (1941: 96), being among the first reflections on the word order variations of *sin<sup>1</sup>* 先 'first', suggested that the use of its preverbal or postverbal variations might have been contextually conditioned in his time. Compare (32) and (33).

- (32) 佢先打我。

Keoi<sup>5</sup> sin<sup>1</sup> daa<sup>2</sup> ngo<sup>5</sup>.

3SG first hit 1SG

'He first hit me (and then did something else).'

(O'Melia 1941: 96)

- (33) 佢打我先。

Keoi<sup>5</sup> daa<sup>2</sup> ngo<sup>5</sup> sin<sup>1</sup>.

3SG hit 1SG first

'He hit me first (and then someone did something).' (O'Melia 1941: 96)

According to O'Melia (1941: 96), with *sin*<sup>1</sup> 先 'first' preceding the VP *daa*<sup>2</sup> ngo<sup>5</sup> 打我 'hit me', (32) was most likely to be followed by a subsequent description with the same subject such as *zi*<sup>1</sup> hau<sup>6</sup> jau<sup>6</sup> tek<sup>3</sup> ngo<sup>5</sup> 之後又踢我 afterwards-again-kick-1SG 'then (he) kicked me'. Contrastively, with *sin*<sub>1</sub> following the same VP, (33) could imply a change of subject in the subsequent description such as ngo<sup>5</sup> daa<sup>2</sup> faan<sup>1</sup> keoi<sup>5</sup> 我打返佢 1SG-hit-return-3SG 'then I hit him back'. Despite this intuition-based perception, O'Melia nevertheless admitted that "it is hard to say when 'sin' precedes and when it follows".

The use frequency of *sin*<sub>1</sub> was on the rise in the early 20th century; meanwhile the use frequency of its preverbal counterpart was on the decline. An overwhelming dominance of *sin*<sub>1</sub> appeared in *Cantonese primer* (Chao 1947a) with only one occurrence of the preverbal *sin*<sup>1</sup> 先 'first' versus 24 occurrences of *sin*<sub>1</sub>. Chao (1947a) has not discussed the word order peculiarity of *sin*<sub>1</sub> but its high use frequency in the textbook sufficed to acknowledge its postverbal ordering as a syntactic requirement rather than a pragmatic reversion. Thereafter, the postverbal ordering has remained as a highly dominant word order of *sin*<sup>1</sup> 先 'first' in the mid-20th century's materials of CHKC and through to contemporary Cantonese.

What had accompanied the syntacticization of *sin*<sub>1</sub> was its functional extension into *sin*<sub>2</sub> and *sin*<sub>3</sub>. We shall use Tang's (2012) criteria to tease its imperative usage from its temporal usage. According to Tang (2012), *sin*<sub>2</sub>, in a typical usage scenario, appears in an imperative sentence with a second person subject, which expresses deontic modality marked by modal verbs such as *jiu*<sup>3</sup> 要 'should', causative verbs such as *dang*<sup>2</sup> 等 'let', or negatives such as *mai*<sup>5</sup> 咪 'will not' and *m*<sup>4</sup>hou<sup>2</sup> 唔好 'not good, will not'. Early occurrences of *sin*<sub>2</sub> have been documented in historical materials dated in the 1940s, as shown in (34) to (38).

- (34) 如果你係洗面，要擰水嚟先。

Jyu<sup>3</sup>gwo<sup>2</sup> nei<sup>5</sup> hai<sup>6</sup> sai<sup>2</sup> min<sup>6</sup>, jiu<sup>3</sup> ning<sup>1</sup> sei<sup>2</sup> lai<sup>4</sup> sin<sup>1</sup>.if 2SG COP wash face should bring water come *sin*<sub>2</sub>

'If you are to wash your face, you must first bring here some water.'

(O'Melia 1941: 173)

- (35) 你唔好去住先。

Nei<sup>5</sup> m<sup>4</sup> hou<sup>2</sup> heoi<sup>3</sup> zyu<sup>6</sup> sin<sup>1</sup>.2SG not good go SFP *sin*<sub>2</sub>

'Do not go right away.'

(O'Melia 1941: 260)

- (36) 試過先喇。  
*Si<sup>3</sup> gwo<sup>3</sup> sin<sup>1</sup> laa<sup>3</sup>.*  
 try ASP *sin<sub>2</sub>* SFP  
 ‘(Let me) have a try first.’ (O’Melia 1941: 299)
- (37) 等我問吓個位看護先。  
*Dang<sup>2</sup> ngo<sup>5</sup> man<sup>6</sup> haa<sup>5</sup> go<sup>2</sup> wai<sup>6</sup> hon<sup>1</sup>wu<sup>6</sup> sin<sup>1</sup>.*  
 wait 1SG ask ASP that CL nurse *sin<sub>2</sub>*  
 ‘Just let me ask that nurse first.’ (Chao 1947a; b)
- (38) 你擰右便個個先喇。  
*Nei<sup>5</sup> ning<sup>6</sup> jau<sup>6</sup>bin<sup>1</sup> go<sup>2</sup> go<sup>3</sup> sin<sup>1</sup> laa<sup>3</sup>.*  
 2SG turn right that CL *sin<sub>2</sub>* SFP  
 ‘You turn the switch in the right first.’ (Chao 1947a; b)

Although *sin<sub>2</sub>* had been an established usage in the 1940s, our historical data of the 1960s had not yet confirmed the use of *sin<sub>3</sub>*. Thus, *sin<sub>3</sub>* can be seen as a more recent development of *sin<sup>1</sup>* 先.

Table 1 summarizes our findings on the syntactic distribution of *sin<sup>1</sup>* 先 ‘first’ in selected Cantonese historical materials dated from the 1820s to the 1960s. Additionally, the occurrences of *sin<sub>1</sub>* and *sin<sub>2</sub>* have been counted separately in order to show the semantic development of the postverbal *sin<sup>1</sup>* 先.

As shown in Table 1, *sin<sup>1</sup>* 先 ‘first’ in Cantonese had basically been used as a preverbal adverb before 1930s even though it had occasionally occurred in the postverbal position; and the use of *sin<sub>1</sub>*, in comparison with its preverbal counterpart, had remained at a relatively low frequency in the selected historical materials expanding a century from 1820s into 1930s. The rare use of *sin<sub>1</sub>* before 1930s could be credited to the influence by neighboring Guangxi dialects, which even now do not show a strong preference to use a postverbal *xiān* 先 ‘first’. However, the typological change of *sin<sup>1</sup>* 先 ‘first’ from preverbal to postverbal had been so dramatic and rapid that it took only about a decade’s time around the 1940s. By the late 1940s, *sin<sub>1</sub>* had become a syntacticalized usage with a high use frequency and further grammaticalized into *sin<sub>2</sub>*. In our selected materials of the 1950s and 1960s, the use frequency of *sin<sub>2</sub>* greatly outnumbered that of *sin<sub>1</sub>*; because all those materials are dialogues from Cantonese movies, which heavily involve speakers’ interactions and thus provide more imperative utterances.

To conclude, Table 1 observes a limited but growing use of *sin<sub>1</sub>* from the 1820s to the 1930s before a remarkable replacement took place around the 1940s, when the domination of *sin<sub>1</sub>* and the appearance of *sin<sub>2</sub>* were accompanied by the extreme marginalization of the preverbal *sin<sup>1</sup>* 先 ‘first’. As previously noted, the postverbal positioning of *xiān* 先 ‘first’ in the Chinese dialects of southern coastal China was

**Table 1.** Syntactic distribution of *sin<sup>1</sup>* 先 in selected Cantonese historical materials dated from the 1820s to the 1960s

Year	Material	Preverbal <i>sin<sup>1</sup></i> 'first' (no. of occur.)	Postverbal <i>sin<sup>1</sup></i> (no. of occur.)	
			<i>sin<sub>1</sub></i>	<i>sin<sub>2</sub></i>
1828	<i>Vocabulary of the Canton dialect</i>	8	–	–
1841	<i>A Chinese chrestomathy in the Canton dialect</i>	39	2	–
1872	<i>The gospel according to St. Mark in the Cantonese colloquial</i>	13	–	–
1877	<i>Forty chapters on Cantonese phrases</i>	6	1	–
1877	<i>Easy phrases in the Canton dialect of the Chinese language</i>	3	1	–
1888	<i>Cantonese made easy</i> (2nd edition)	2	–	–
1902	<i>How to speak Cantonese</i> (2nd edition)	8	–	–
1912	<i>How to speak Cantonese</i> (4th edition)	6	2	–
1924	<i>Cantonese made easy</i> (4th edition)	2	–	–
1931	<i>Progressive and idiomatic sentences in Cantonese colloquial</i>	7	2	–
1941	<i>First year Cantonese: Part one</i> (2nd edition)	6	5	3
1947	<i>Cantonese primer</i>	1	20	4
1952	<i>Foster-Daddy's Romantic Affairs</i>	5	12	22
1954	<i>Wrong Wedding</i>	–	8	8
1955	<i>If Only We'd Met When I Was Single</i>	1	4	15
1962	<i>To Capture the God of Wealth</i>	1	2	16
1963	<i>One Queen and Three Kings</i>	1	7	10
1965	<i>The Ideal Husband</i>	–	6	12
1966	<i>Violet Girl</i>	3	7	12

induced by a syntactic diffusion from TK and HM languages, which was assumed to be a prolonged and gradual process. However, what has cast considerable doubt on the emergence of the postverbal *sin<sup>1</sup>* 先 in Cantonese is its radical word order change around the 1940s. Therefore, we shall propose a hypothesis that the rise of postverbal *sin<sup>1</sup>* 先 has also been motivated by internal causes. The next section will deal with the semantic change of *sin<sup>1</sup>* 先, which was responsible for its radical word order change around the 1940s.

#### 4. Semantic change of *sin*<sup>1</sup> 先: Precedent versus subsequent

In previous studies of *sin*<sup>1</sup> 先, the postverbal *sin*<sup>1</sup> 先 has always occupied the center of discussion, but not much talk has been dedicated to the preverbal *sin*<sup>1</sup> 先 and even less attention paid to the syntax-semantic interaction of its functional variants. In this section, we shall first outline the semantic complexity of preverbal *sin*<sup>1</sup> 先, and then reconstruct the semantic change of *sin*<sup>1</sup> 先 based on historical materials.

In the temporal domain, ‘precedent’ and ‘subsequent’ are two contrastive or rather opposite meanings: ‘precedent’ indicates ‘(something) happening prior in time (i.e. before something else)’, whereas ‘subsequent’ indicates ‘(something) happening later in time (i.e. after something else)’. The ‘precedent-subsequent’ contrast in Chinese is usually represented by a set of differentiable forms, such as *xiān* 先 ‘first’ vs. *cái* 才 / *zài* 再 ‘then’ in Mandarin or *sie*<sup>53</sup> 先 ‘first’ vs. *tse*<sup>53</sup> 再 ‘then’ in Shanghai Wu. This contrast is also well differentiated in TK and HM, such as *ba:ŋ*<sup>1</sup> ‘first’ vs. *la*<sup>3</sup> ‘later’ / *hap*<sup>7</sup> ‘then’ in Lakkia (Mao et al. 1982: 150) and *k<sup>h</sup>uan*<sup>5</sup> ‘first’ vs. *tsuən*<sup>3</sup> ‘later’ / *nou*<sup>3</sup> ‘then’ in Cun (Ouyang 1998: 125–126). But unlike those languages and dialects, Cantonese *sin*<sup>1</sup> 先 encodes both meanings in one form.

As shown in § 3, Cantonese *sin*<sup>1</sup> 先 has been found mainly used as a preverbal adverb for ‘first’ before the 1930s, while the contemporary usage of preverbal *sin*<sup>1</sup> 先 ‘first’ therefore characterizes its original word order historically. In addition to encoding ‘precedent’, the preverbal *sin*<sup>1</sup> 先 also represents a variety of other grammatical and pragmatic functions related to the ‘subsequent’ meaning: (a) *sin*<sup>1</sup> 先 to mark temporal subsequence (*sin*<sub>a</sub> hereafter), meaning ‘then’ or ‘afterwards’ as shown in (39); (b) *sin*<sup>1</sup> 先 to mark a conditional apodosis (*sin*<sub>b</sub>), meaning ‘then’ or ‘only after satisfying a certain condition’ as shown in (40); (c) *sin*<sup>1</sup> 先 to express counter-expectation (*sin*<sub>c</sub>), specifically marking an unexpectedly shorter/later time or an unexpectedly smaller amount<sup>19</sup> as shown in (41a–b); and (d) *sin*<sup>1</sup> 先 to express a contrastive emphasis (*sin*<sub>d</sub>), emphasizing a certain property as might not be presupposed about the subject as shown in (42). (See Mai 1993; Sybesma 2013.)

- (39) 食埋啲嘢先走。  
*Sik*<sup>6</sup> *maai*<sup>4</sup> *di*<sup>1</sup> *je*<sup>5</sup> *sin*<sup>1</sup> *zau*<sup>2</sup>.  
 eat approach CL thing then leave  
 ‘Leave after finishing food.’

19. M. Liu (1993: 147–148) regards this function as marking an unexpected “lower-valued end” of a quantitative scale.

- (40) 你要識搵錢，先有女仔鐘意你。  
*Nei<sup>5</sup> jiu<sup>3</sup> sik<sup>1</sup> wan<sup>2</sup> cin<sup>4</sup>, sin<sup>1</sup> jau<sup>5</sup> nei<sup>5</sup>zai<sup>2</sup> zung<sup>1</sup>ji<sup>3</sup> nei<sup>5</sup>.*  
 2SG must know find money *sin<sub>b</sub>* have girl like 2SG  
 ‘Only if you know how to make money will girls like you.’
- (41) a. 佢而家先返嚟。  
*Keoi<sup>5</sup> ji<sup>4</sup>gaa<sup>1</sup> sin<sup>1</sup> faan<sup>1</sup> lai<sup>4</sup>.*  
 3SG now *sin<sub>c</sub>* return come  
 ‘S/he didn’t come back until now.’
- b. 街口個檔先賣十文咋。  
*Gaa<sup>1</sup> hau<sup>2</sup> go<sup>3</sup> dong<sup>3</sup> sin<sup>1</sup> maa<sup>1</sup>i<sup>6</sup> sap<sup>6</sup> man<sup>1</sup> zaa<sup>3</sup>.*  
 street corner CL stall *sin<sub>c</sub>* sell ten CL SFP  
 ‘It’s only priced at ten dollars at the stall around the street corner.’
- (42) 佢先係我老豆。  
*Keoi<sup>5</sup> sin<sup>1</sup> hai<sup>6</sup> ngo<sup>5</sup> lou<sup>5</sup>dau<sup>6</sup>.*  
 3SG *sin<sub>d</sub>* COP 1SG father  
 ‘It’s him that is my father.’

In all the above sentences from (39) to (42), *sin<sup>1</sup>* 先 is interchangeable with *zi<sup>3</sup>* 至 or *sin<sup>1</sup>zi<sup>3</sup>* 先至.<sup>20</sup> Regarding this interchangeability, Mai (1993) proposes that the disyllabic string<sup>21</sup> of *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 ‘first then’ lexicalized into *sin<sup>1</sup>zi<sup>3</sup>* 先至 ‘then’, which was thereafter phonologically reduced to *sin<sub>a</sub>*. That is to say, the historical stages giving rise to *sin<sub>a</sub>* can be reconstructed as follows:

*zi<sup>3</sup>* 至 ‘then’ > *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 ‘first then’ > *sin<sup>1</sup>zi<sup>3</sup>* 先至 ‘then’ > *sin<sup>1</sup>* 先 ‘then’ (*sin<sub>a</sub>*).

Mai’s (1993) proposal has offered us much enlightenment for exploring the semantic change of *sin<sup>1</sup>* 先.

According to our historical materials, the use of *zi<sup>3</sup>* 至 for temporal, conditional, counter-expectation, and emphatic markings is observed at a much earlier period than either *sin<sup>1</sup>* 先 or *sin<sup>1</sup>zi<sup>3</sup>* 先至 for the same functions. Consider the examples in (43) to (46):<sup>22</sup>

20. The above-mentioned functions of *sin<sup>1</sup>* 先, *zi<sup>3</sup>* 至 and *sin<sup>1</sup>zi<sup>3</sup>* 先至 also correspond to the usage pattern of *cái* 才 in Mandarin. (See M. Liu 1993; Lü 1999: 107–108 for *cái* 才.)

21. The term “string”, in some sense equivalent to “sequence”, refers to a linear array of linguistic constituents (see P. H. Matthews 2007: 384). The term is particularly useful in referring to an array of constituents both before and after lexicalization or grammaticalization as it obscures any potential phonological, morphosyntactic, or semantic boundaries within the constituents.

22. The character 致 was sometimes used as an alternative to represent *zi<sup>3</sup>* in Cantonese historical materials. In this paper, only the character 至 will be used to represent *zi<sup>3</sup>* and amendment will be made on cited sentences where necessary.

- (43) 你件衫唔齊整，除紐著好佢至扣番紐啦。  
*Nei<sup>5</sup> gin<sup>6</sup> saam<sup>1</sup> m<sup>4</sup> cai<sup>4</sup>zing<sup>2</sup>, ceoi<sup>4</sup> nau<sup>2</sup> zoek<sup>3</sup> hou<sup>2</sup> keoi<sup>5</sup> zi<sup>3</sup>*  
 2SG CL coat not neat unfasten button wear good 3SG then  
*kau<sup>3</sup> faan<sup>1</sup> nau<sup>2</sup> laa<sup>1</sup>.*  
 fasten again button SFP  
 ‘Your coat is not neatly dressed. Unbutton it and adjust it right before buttoning it up again.’  
 (Bridgman 1841; ECCT)
- (44) 腳大至跂得穩。  
*Geok<sup>3</sup> daai<sup>6</sup> zi<sup>3</sup> kei<sup>5</sup> dak<sup>1</sup> wan<sup>2</sup>.*  
 foot big COND stand PRT steady  
 ‘Only big feet will stand steady.’  
 (Morrison 1828; ECCT)
- (45) 佢聽日至來和。<sup>23</sup>  
*Keoi<sup>5</sup> ting<sup>1</sup>jat<sup>6</sup> zi<sup>3</sup> lai<sup>4</sup> wo<sup>3</sup>.*  
 3SG tomorrow CEX come SFP  
 ‘He will not come until tomorrow.’  
 (Morrison 1828; ECCT)
- (46) 咁至係規矩。  
*Gam<sup>3</sup> zi<sup>3</sup> hai<sup>6</sup> kwai<sup>1</sup>geoi<sup>2</sup>.*  
 such EMP COP rule  
 ‘This is the rule.’  
 (Bruce 1877; ECTD)

Contrary to the commonly witnessed use of *zi<sup>3</sup> 至*, neither *sin<sub>a~d</sub>* nor *sin<sup>1</sup>zi<sup>3</sup> 先至* occurred in our historical materials dated from 1820s to 1930s. The earliest occurrences of the string of *sin<sup>1</sup>+zi<sup>3</sup> 先+至* ‘first then’ were observed in *First year Cantonese: Part one* (O’Melia 1941). The sequencing was made possible by a significantly increased use frequency of *sin<sub>1</sub>* in 1940s as indicated by Table 1. All the three occurrences of the string in the textbook are shown in (47) to (49) with the original translations quoted, which could provide an important clue to the lexicalization of the string.

- (47) 我食飯先至洗面。  
*Ngo<sup>5</sup> sik<sup>6</sup> faan<sup>6</sup> sin<sup>1</sup> zi<sup>3</sup> sai<sup>2</sup> min<sup>6</sup>.*  
 1SG eat meal first then wash face  
 Originally translated as ‘I will eat rice **first before** washing my face.’  
 (O’Melia 1941: 234)
- (48) 你翻嚟先至去佢處。  
*Nei<sup>5</sup> faan<sup>1</sup> lai<sup>4</sup> sin<sup>1</sup> zi<sup>3</sup> heoi<sup>3</sup> keoi<sup>5</sup> syu<sup>3</sup>.*  
 2SG return come first then go 3SG place  
 Originally translated as ‘Come back **first before** going to his place.’  
 (O’Melia 1941: 234)

23. The character 和 was romanized as *wo* in Morrison (1828) but it actually represented the sentence-final particle *wo<sup>3</sup> 喎*.

(49) 你買菜先至食飯。

*Nei<sup>5</sup> maai<sup>5</sup> coi<sup>3</sup> sin<sup>1</sup> zi<sup>3</sup> sik<sup>6</sup> faan<sup>6</sup>.*

2SG buy vegetable first then eat meal

Originally translated as 'Buy vegetables **first before** eating.'

(O'Melia 1941: 234)

In O'Melia (1941), the *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 string was uniformly translated as 'first before', which implied that the two morphemes were analyzed, respectively, as 'first' and 'before'.<sup>24</sup> Therefore, the syntactic context that the string of *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 appeared in was restricted and can be generalized as  $[VP_1 + sin^1 \text{先}] + [zi^3 \text{至} + VP_2]$ . Take (47) for example, *sin<sup>1</sup>* 先 'first' in the postverbal position, i.e. *sin<sub>1</sub>*, modifies the first VP *sik<sup>6</sup> faan<sup>6</sup>* 食飯 'eat the meal', whereas *zi<sup>3</sup>* 至 'then' as a preverbal adverb modifies the second VP *sai<sup>2</sup> min<sup>6</sup>* 洗面 'wash the face'. Despite the linear adjacency that *sin<sup>1</sup>* 先 'first' and *zi<sup>3</sup>* 至 'then' temporarily exemplified as a result of the sequencing of  $[[sik^6 \text{ faan}^6 \text{ 食飯}]_{VP} \text{ sin}^1 \text{ 先}]$  and  $[zi^3 \text{ 至} [sai^2 \text{ min}^6 \text{ 洗面}]_{VP}]$ , there was still a phrasal boundary between the two temporal adverbs. What is noteworthy is that the sequencing of  $VP_1 + sin^1 \text{先} + zi^3 \text{至} + VP_2$  can be both morpho-syntactically and semantically ambiguous. As for examples in (47) to (49), in addition to the separate reading of *sin<sup>1</sup>* 先 and *zi<sup>3</sup>* 至, whereby they are modifiers dependent, respectively, on two sequenced VPs, the *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 string can otherwise be understood as a preverbal temporal adverb meaning 'then' and modifying only the second VP.<sup>25</sup> In those ambiguous contexts, the morpho-syntactical boundaries of  $[VP_1 + sin^1 \text{先}] + [zi^3 \text{至} + VP_2]$  can be reanalyzed as  $VP_1 + [sin^1 zi^3 \text{先至} + VP_2]$ , leading to the lexicalization of *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 'first then' into *sin<sup>1</sup>zi<sup>3</sup>* 先至 'then'.<sup>26</sup> Undeniably,

24. In temporal marking, *zi<sup>3</sup>* 至 neither literally means nor can be glossed as 'before' in a strict sense. In fact, a prepositional phrase like "before + VP-ing" is often used to translate "*zi<sup>3</sup>* 至 + VP" into English, as is the common case in O'Melia (1941). See the example below.

(i) 等我想吓至話你知喇。

*Dang<sup>2</sup> ngo<sup>5</sup> soeng<sup>2</sup> haa<sup>5</sup> zi<sup>3</sup> waa<sup>6</sup> nei<sup>5</sup> zi<sup>1</sup> laa<sup>1</sup>.*

wait 1SG think ASP then tell 2SG know SFP

Originally translated as 'Let me think a while **before** telling you.' (O'Melia 1941: 233)

25. Once the string of *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 appears within two VPs, either of the above interpretations can make it semantically congruous to the context that it appears in. It is hard to identify a string of *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 between two VPs that can only be interpreted with a phrasal boundary. A reviewer noted that it would be more convincing if some examples where the subjects of the two VPs in  $VP_1 + sin^1 \text{先} + zi^3 \text{至} + VP_2$  are different can be raised. Unfortunately, we were not able to collect such examples in our data from the 1940s. In our opinion, a difference in subject will not necessarily confirm a phrasal boundary within *sin<sup>1</sup>+zi<sup>3</sup>* 先+至.

26. In a strict sense, the fusion of the *sin<sup>1</sup>+zi<sup>3</sup>* 先至 string does not tally strictly with either lexicalization or grammaticalization as defined in Brinton & Traugott (2005: 95–102). The process does not characterize a change from "less to more grammatical", which is the defining feature of grammaticalization, whereas the output of *sin<sup>1</sup>zi<sup>3</sup>* 先至 'then' as a functional word may also

there remains another hypothesis that  $[VP_1 + \text{sin}^1 \text{先}] + [\text{zi}^3 \text{至} + VP_2]$  be reanalyzed as  $[VP_1 + \text{sin}^1 \text{zi}^3 \text{先至}] + VP_2$ , with  $\text{sin}^1 \text{zi}^3 \text{先至}$  being pressed into service as a postverbal adverb modifying the preceding VP while taking up the meaning of ‘first’ so as to be in line with  $\text{sin}_1$  rather than  $\text{zi}^3 \text{至}$ . However, this path of reanalysis is typologically unusual and practically unrealized in Cantonese where most adverbs are used preverbally.<sup>27</sup> Notably, O’Melia (1941) represented a preparation stage critical for the lexicalization of  $\text{sin}^1 + \text{zi}^3 \text{先} + \text{至}$ .

Beginning in 1950s, the contextual restriction for the  $\text{sin}^1 + \text{zi}^3 \text{先} + \text{至}$  string was relaxed, as the string did not necessarily occur in two sequenced VPs and moreover developed into the multi-functionality as shown in (50) to (53).

- (50) 日日見醫生！周身都睇過！醫生先至話我大件又精神呀！  
*Jat<sup>6</sup> jat<sup>6</sup> gin<sup>3</sup> ji<sup>1</sup> sang<sup>1</sup>! Zau<sup>1</sup> san<sup>1</sup> dou<sup>1</sup> tai<sup>2</sup> gwo<sup>3</sup>! Ji<sup>1</sup> sang<sup>1</sup> sin<sup>1</sup> zi<sup>3</sup>*  
 day day see doctor whole body all see ASP doctor then  
*waa<sup>6</sup> ngo<sup>5</sup> daai<sup>6</sup> gin<sup>6</sup> jau<sup>6</sup> zing<sup>1</sup> san<sup>4</sup> aa<sup>3</sup>!*  
 say 1SG strong and vigorous SFP  
 ‘During that time I had seen a doctor everyday and had my body examined from tip to toe. Thereafter, the doctor said that I was robustly healthy.’  
 (As Luck Will Have It, 1955; CHKC)

- (51) 噉都唔使著晒褂去啫？呢啲係大場面先至著㗎嘛！  
*Gam<sup>2</sup> dou<sup>1</sup> m<sup>4</sup> sai<sup>2</sup> zoek<sup>3</sup> saai<sup>3</sup> kwaa<sup>2</sup> heoi<sup>3</sup> zek<sup>1</sup>? Ni<sup>1</sup>*  
 such all not need wear ASP Chinese-styled.gown go SFP this  
*di<sup>1</sup> hai<sup>6</sup> daai<sup>6</sup> coeng<sup>4</sup> min<sup>6</sup> sin<sup>1</sup> zi<sup>3</sup> zoek<sup>3</sup> gaa<sup>3</sup> maa<sup>3</sup>!*  
 CL COP big occasion COND wear SFP SFP  
 ‘In this case we don’t need to wear Chinese-styled gowns to go there, do we? Only on important occasions do we need to be dressed in that way.’  
 (Foster-Daddy’s Romantic Affairs, 1952; CHKC)

- (52) 我而家先至知道，原來我有一個咁溫暖嘅家庭。  
*Ngo<sup>5</sup> ji<sup>4</sup> gaa<sup>1</sup> sin<sup>1</sup> zi<sup>3</sup> zi<sup>1</sup> dou<sup>6</sup>, jyun<sup>4</sup> loi<sup>4</sup> ngo<sup>5</sup> jau<sup>5</sup> jat<sup>1</sup> go<sup>3</sup> gam<sup>3</sup>*  
 1SG now CEX know actually 1SG have one CL such  
*wan<sup>1</sup> nyun<sup>5</sup> ge<sup>3</sup> gaa<sup>1</sup> ting<sup>4</sup>.*  
 warm PRT family  
 ‘It’s not until now that I have realized that I have such a warm family.’  
 (If Only We’d Met When I Was Single, 1955; CHKC)

distance itself from the canonical cases of lexicalization, whose outputs are semantically contentful. To overcome this dilemma, we regard the process as a case of lexicalization in the broadest sense based on the fact that a new compound word  $\text{sin}^1 \text{zi}^3 \text{先至}$  ‘then’ is formed and added to the Cantonese inventory.

27. Thanks to a reviewer who reminded us to clarify this hypothesis.

- (53) 有冇廉恥，你哋兩個先至知！

*Jau<sup>5</sup> mou<sup>5</sup> lim<sup>4</sup>ci<sup>2</sup>, nei<sup>5</sup>dei<sup>6</sup> loeng<sup>5</sup> go<sup>3</sup> sin<sup>1</sup>zi<sup>3</sup> zi<sup>1</sup>!*

have not sense.of.shame 2.PL two CL EMP know

'Only the two of you will know whether you have a sense of shame.'

(*If Only We'd Met When I Was Single*, 1955; CHKC)

Syntactically, the string of *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 follows, respectively, a bare noun subject *ji<sup>1</sup>sang<sup>1</sup>* 醫生 'doctor' in (50), a copula construction *hai<sup>6</sup> daai<sup>6</sup> coeng<sup>4</sup>min<sup>6</sup>* 係大場面 '(if) there are important occasions' in (51), an adverb *ji<sup>4</sup>gaa<sup>1</sup>* 而家 'now' in (52), and an NP subject *nei<sup>5</sup>dei<sup>6</sup> loeng<sup>5</sup> go<sup>3</sup>* 你哋兩個 'the two of you' in (53). The original phrasal boundary between *sin<sup>1</sup>* 先 and *zi<sup>3</sup>* 至 has been erased, as *sin<sup>1</sup>* 先 was no longer dependent on a preceding VP. Semantically, the string of *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 marks temporal subsequence in (50), conditional in (51), counter-expectation in (52), and emphasis in (53). In none of these sentences does *sin<sup>1</sup>* 先 contribute its original meaning 'first' to the function of the string. In other words, the compositionality of the string has been lost. Therefore, the *sin<sup>1</sup>+zi<sup>3</sup>* 先+至 string suffices to be identified as a single processing unit, i.e. a lexicalized disyllabic adverb in Examples (50) to (53).

A similar use pattern can also be observed of the preverbal *sin<sup>1</sup>* 先 in the same period as shown in (54) to (57).

- (54) 一陣先講畀你聽！快啲去㗎！

*Jat<sup>1</sup> zan<sup>6</sup> sin<sup>1</sup> gong<sup>2</sup> bei<sup>2</sup> nei<sup>5</sup> teng<sup>1</sup>! Faai<sup>3</sup> di<sup>1</sup> hoei<sup>3</sup> aa<sup>1</sup>!*

one moment then tell give 2SG listen fast CL go SFP

'I will tell you after a moment. Get there quickly.'

(*Foster-Daddy's Romantic Affairs*, 1952; CHKC)

- (55) 你都要買啲嘢，買啲名貴嘅嘢，去孝敬下人啲老人家，噉人哋先對你好印象㗎嘛！

*Nei<sup>5</sup> dou<sup>1</sup> jiu<sup>3</sup> maai<sup>5</sup> di<sup>1</sup> je<sup>5</sup>, maai<sup>5</sup> di<sup>1</sup> ming<sup>4</sup>gwai<sup>3</sup> ge<sup>3</sup> je<sup>5</sup>,*

2SG even should buy CL thing buy CL precious PRT thing

*heoi<sup>3</sup> haau<sup>3</sup>ging<sup>3</sup> haa<sup>5</sup> jan<sup>4</sup> di<sup>1</sup> lou<sup>5</sup>jan<sup>4</sup>gaa<sup>1</sup>, gam<sup>2</sup> jan<sup>4</sup>di<sup>1</sup> sin<sup>1</sup>*

go present.gifts.to ASP person CL elderly so others *sin<sub>b</sub>*

*deoi<sup>3</sup> nei<sup>5</sup> hou<sup>2</sup> jan<sup>3</sup>zoeng<sup>6</sup> gaa<sup>3</sup> maa<sup>3</sup>!*

to 2SG good impression SFP SFP

'You should buy something, like something precious, for the elderly in the girl's family. Only by doing so will her family have a good impression of you.'

(*A Ready Lover*, 1952; CHKC)

- (56) 我𠵼𠵼有𠵼事要做，所以而家先𠵼𠵼𠵼！

*Ngo*<sup>5</sup> *ngaam*<sup>1</sup>*ngaam*<sup>1</sup> *jau*<sup>5</sup> *di*<sup>1</sup> *si*<sup>6</sup> *jiu*<sup>3</sup> *zou*<sup>6</sup>, *so*<sup>2</sup>*ji*<sup>5</sup> *ji*<sup>4</sup>*gaa*<sup>1</sup> *sin*<sup>1</sup>  
 1SG just.now have CL matter need do thus now *sin*<sub>c</sub>  
*lai*<sup>4</sup> *zaa*<sup>3</sup> *maa*<sup>3</sup>!  
 come SFP SFP

'Just now I had some matters to attend to so I came here late.'

(*If Only We'd Met When I Was Single*, 1955; CHKC)

- (57) 死𠵼𠵼先係你女婿呀！

*Sei*<sup>2</sup> *go*<sup>2</sup> *go*<sup>3</sup> *sin*<sup>1</sup> *hai*<sup>6</sup> *nei*<sup>5</sup> *nei*<sup>5</sup>*sai*<sup>3</sup> *aa*<sup>3</sup>!  
 dead that CL *sin*<sub>d</sub> COP 2SG son-in-law SFP

'It's the dead one who is your son-in-law.' (*Wrong Wedding*, 1954; CHKC)

Our historical materials do not provide enough information for us to chronologize the developments of *sin*<sup>1</sup>*zi*<sup>3</sup> 先至 and *sin*<sub>a~d</sub> as illustrated in examples from (50) to (57).<sup>28</sup> Nevertheless, we believe that the emergence of *sin*<sup>1</sup>*zi*<sup>3</sup> 先至 is a prerequisite for the semantic change of *sin*<sup>1</sup> 先.

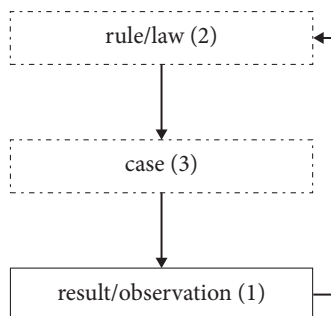
As for *sin*<sup>1</sup>*zi*<sup>3</sup> 先至, two strategies can be postulated for its semantic extension. The first strategy is grammaticalization. The functional variants of *sin*<sup>1</sup>*zi*<sup>3</sup> 先至 are exemplary of a unidirectional cline from (LOCALISTIC >) TEMPORAL > LOGICAL > ILLOCUTIVE/DISCOURSE FUNCTIONAL (see Abraham 1991). As a result of lexicalization, *sin*<sup>1</sup>*zi*<sup>3</sup> 先至 emerged as a temporal adverb to mark a subsequent action as in (50) and then it was metaphorically used to mark a conditional apodosis as in (51). The semantic extension from temporal to logical is cross-linguistically common (see Heine & Kuteva 2002: 293). Its counter-expectation and contrastive-emphatic functions are also related to its 'subsequent' meaning. As a means to express personal attitude, *sin*<sup>1</sup>*zi*<sup>3</sup> 先至 marks a subsequent evaluation, which counters or contrasts with what has been expected or presumed in the precedent situation as shown in (52) and (53). The reconstructed semantic extension of *sin*<sup>1</sup>*zi*<sup>3</sup> 先至 is of a more general path from PROPOSITIONAL > TEXTUAL > EXPRESSIVE (see Traugott 1982). The second strategy is polysemy copy.<sup>29</sup> The semantic pattern of *zi*<sup>3</sup> 至 may serve as the model of polysemy copy. Once the two forms of *sin*<sup>1</sup>*zi*<sup>3</sup> 先至 and *zi*<sup>3</sup> 至 are matched based on their common temporal

28. This can be due to two aspects of reasons: on the one hand, our selected historical materials before the 1950s are focused on written literature (unlike movie dialogues), which could be conservative in reflecting innovations in spoken grammar; on the other, the functional extension of *sin*<sup>1</sup>*zi*<sup>3</sup> 先至 and *sin*<sub>a~d</sub> might have taken place in colloquial Cantonese within such a short time in the 1940s that the process is unlikely to be fully reflected in historical materials.

29. Polysemy copy, referring to the copying of a polysemy pattern, is related to two types of motivations, i.e. one being contact-induced (as discussed in Heine & Kuteva 2005: 100–103) and the other being internally motivated (as discussed here).

meaning ‘then’ (also partially on their formal identicalness), the more grammaticalized meanings of conditionality, counter-expectation and contrastive emphasis can thereafter be encoded into *sin<sup>1</sup>zi<sup>3</sup>* 先至 in the model of *zi<sup>3</sup>* 至. The second strategy indicates a more straightforward way for the semantic extension of *sin<sup>1</sup>zi<sup>3</sup>* 先至.

As for *sin<sub>a~d</sub>*, its use as a reduced form of *sin<sup>1</sup>zi<sup>3</sup>* 先至 may not just be due to erosion, i.e. the loss of phonological or morphemic segments in the continuing process of fusion (Hopper & Traugott 2003: 157), but is more decisively motivated by abductive inference, i.e. an inferential process that proceeds from an observed result, then invokes a rule and finally concludes that something may be the case as shown in Figure 2 (Anderson 1973; Fischer 2001).



The box with continuous lines contains premises/hypotheses that are presupposed as given/true.  
The boxes with dotted lines contain hypotheses that are inferred.

Figure 2. Stages of abductive inference (reproduced from Fischer 2001)

The abductive inference leading to the semantic change of *sin<sup>1</sup>* 先 can be reconstructed as follows. In Stage (1), Cantonese speakers had the observation that *sin<sup>1</sup>zi<sup>3</sup>* 先至 and *zi<sup>3</sup>* 至, as synonyms, were used interchangeably in certain frequently occurring contexts. In Stage (2), a common rule of word formation in Chinese, known as synonym compounding,<sup>30</sup> was invoked and a relevant hypothesis was made that *sin<sup>1</sup>zi<sup>3</sup>* 先至 had been formed by synonym compounding. In Stage (3), a case was thereafter invoked by abduction that *sin<sup>1</sup>* 先 was synonymous with both *zi<sup>3</sup>* 至 and *sin<sup>1</sup>zi<sup>3</sup>* 先至 and could substitute either of the latter two words. Thus,

30. Synonym compounding (*tóngyì fùhé* 同義複合), whereby two synonymous or near-synonymous morphemes form a compound word (e.g. *péngyǒu* 朋友 friend-friend ‘friend’, *tánlùn* 談論 discuss-discuss ‘discuss’, *míngliàng* 明亮 bright-bright ‘bright’ and so on), is one of the common word-formation processes in Chinese morphology (Ge 2014: 83–84). Consequently, a synonym compound (*tóngyì fùcí* 同義複詞) usually carries a synonymous meaning as its components. This means of word formation has been generating a considerable number of Chinese compound words throughout the history of disyllabication, i.e. a shift from monosyllabic to disyllabic words in the Chinese lexical history (Xu 1990; F. Wang 2015).

the ‘subsequent’ meaning was abductively encoded into  $\sin^1$  先 from  $\sin^1 zi^3$  先至 and  $zi^3$  至. Notably, the semantic change of  $\sin^1$  先 was of an irregular kind that did not exemplify the major mechanisms of semantic change, i.e. metaphORIZATION and metonymization (Traugott & Dasher 2001: 27–34), and was only made possible by  $\sin^1 zi^3$  先至 as a bridging form.

The influence between Cantonese  $\sin^1$  先 and  $zi^3$  至 might not just be unidirectional but even bidirectional. As a noteworthy fact, Gao (1980: 138) has pointed out that in Guangzhou Cantonese  $zi^3$  至 was observed to appear in the postverbal position with a usage resembling the imperative  $\sin_b$  as in (58).

(58) 你唔使行得咁快至啦!

*Nei<sup>5</sup> m<sup>4</sup> sai<sup>2</sup> haang<sup>4</sup> dak<sup>1</sup> gam<sup>3</sup> faai<sup>3</sup> zi<sup>3</sup> laa<sup>1</sup>!*

2SG not need walk PRT SO fast IMP SFP

‘You don’t need to walk so fast!’

(Gao 1980: 58)

According to Gao (1980: 138), this imperative usage of  $zi^3$  至 was highly restricted to negative contexts. However, such a usage of  $zi^3$  至 has been confirmed neither in our historical materials or by our Guangfu Cantonese informants, who regard a verbal particle  $zyu^6$  住 ‘stay’ as the appropriate word to be used instead of  $zi^3$  至 in (58). If Gao’s finding did reflect a short-lived usage of  $zi^3$  至 in the 1980s, it is also not totally out of expectation. Based on the existing synonymy of  $zi^3$  至 with  $\sin_{a\sim d}$ ,  $zi^3$  至 had the possibility to be analogically extended to represent other functions of  $\sin^1$  先, leading to more shared semantic and syntactic behaviors of the two words.<sup>31</sup>

To sum up our findings in this section, Table 2 counts the occurrences of  $zi^3$  至,  $\sin^1 + zi^3$  先至 and  $\sin_{a\sim d}$  in their temporal, conditional, counter-expectation, and emphatic usages in selected Cantonese historical materials dating from the 1820s through the 1960s.

As shown in Table 2,  $zi^3$  至 had already demonstrated its full-fledged usages in the 1820s, long before the emergence of either  $\sin^1 zi^3$  先至 or  $\sin_{a\sim d}$ . As previously stated,  $\sin_1$  was a minor usage from the 1820s to the 1930s but its growing use had been paving the way for the possible sequencing of  $\sin^1 + zi^3$  先至. It was not until the 1940s that the string of  $\sin^1 + zi^3$  先至 first appeared in our data. The string thereafter underwent lexicalization before it served as a bridging form to further introduce  $\sin_{a\sim d}$ . In the early and mid-1950s, both  $\sin^1 zi^3$  先至 and  $\sin_{a\sim d}$  had been observed for temporal, conditional, counter-expectation and emphatic markings.

31. Admittedly, the postverbal usage of  $zi^3$  至 is not a widely recognized or well studied phenomenon in Guangfu Cantonese. Although we quoted Gao (1980)’s finding as it is related to our current discussion, more research needs to be done on the possible development of  $zi^3$  至 under the influence of  $\sin^1$  先 in certain Cantonese dialects.

**Table 2.** The usages of *zi<sup>3</sup>* 至, *sin<sup>1</sup>+zi<sup>3</sup>* 先至, and *sin<sub>a-d</sub>* in selected Cantonese historical materials dating from the 1820s to the 1960s

Year	Material	<i>zi<sup>3</sup></i> 至 (no. of occur.)*				<i>sin<sup>1</sup>+zi<sup>3</sup></i> 先至 (no. of occur.)				<i>sin<sup>1</sup></i> 先 (no. of occur.)			
		a	b	c	d	a	b	c	d	a	b	c	d
1828	<i>Vocabulary of the Canton dialect</i>	4	5	2	2	–	–	–	–	–	–	–	–
1841	<i>A Chinese chrestomathy in the Canton dialect</i>	7	10	1	8	–	–	–	–	–	–	–	–
1872	<i>The gospel according to St. Mark in the Cantonese colloquial</i>	1	4	–	4	–	–	–	–	–	–	–	–
1877	<i>Forty chapters on Cantonese phrases</i>	3	3	1	7	–	–	–	–	–	–	–	–
1877	<i>Easy phrases in the Canton dialect of the Chinese language</i>	1	–	4	4	–	–	–	–	–	–	–	–
1888	<i>Cantonese made easy</i> (2nd edition)	–	1	–	1	–	–	–	–	–	–	–	–
1902	<i>How to speak Cantonese</i> (2nd edition)	5	4	1	20	–	–	–	–	–	–	–	–
1912	<i>How to speak Cantonese</i> (4th edition)	8	5	1	21	–	–	–	–	–	–	–	–
1924	<i>Cantonese made easy</i> (4th edition)	–	2	–	1	–	–	–	–	–	–	–	–
1931	<i>Progressive and idiomatic sentences in Cantonese colloquial</i>	15	8	3	10	–	–	–	–	–	–	–	–
1941	<i>First year Cantonese: Part one</i> (2nd edition)	56	45	2	–	3	–	–	–	–	–	–	–
1947	<i>Cantonese primer</i>	27	7	3	3	–	–	–	–	–	–	–	–
1952	<i>Foster-Daddy's Romantic Affairs</i>	4	24	3	2	4	1	–	–	3	8	2	–
1954	<i>Wrong Wedding</i>	2	7	1	6	–	–	–	2	3	6	–	2
1955	<i>If Only We'd Met When I Was Single</i>	2	13	1	4	1	2	1	2	3	3	2	–
1962	<i>To Capture the God of Wealth</i>	3	14	1	5	1	2	1	–	–	3	1	–
1963	<i>One Queen and Three Kings</i>	–	2	2	1	2	2	–	2	4	10	2	2
1965	<i>The Ideal Husband</i>	7	11	1	5	1	2	3	1	3	5	–	–
1966	<i>Violet Girl</i>	–	3	3	2	–	1	1	–	1	8	–	1

\* The occurrences of *zi<sup>3</sup>* 至 are counted, excluding its occurrences in *sin<sup>1</sup>+zi<sup>3</sup>*.  
a. The marking of temporal subsequence.  
b. The marking of a conditional apodosis.  
c. The marking of counter-expectation.  
d. The marking of a contrastive emphasis.

Frankly speaking, there seems to be a gap of data which were supposed to represent an ideal stage when  $\sin^1 zi^3$  先至 had already demonstrated its full-fledged usages while  $\sin_{a-d}$  had yet to appear. Alternatively, there might have been another extreme case whereby the abduction of  $\sin_{a-d}$  from  $zi^3$  至 and  $\sin^1 zi^3$  先至 was tightly synchronized with the emergence of  $\sin^1 zi^3$  先至, which made the changes too abrupt to be captured by written materials. Although these minute changes are temporarily not covered by our data,<sup>32</sup> it is still an obvious fact that the semantic extension of the preverbal  $\sin^1$  先 can be interpreted as closely related to the lexicalization of  $\sin^1 zi^3$  先至.

Unlike the distinct replacement between  $\sin_1$  and its preverbal counterpart as indicated by Table 1, the emergence of  $\sin^1 zi^3$  先至 and  $\sin_{a-d}$  did not entail a simple replacement of  $zi^3$  至 or any one of the three. Although there seemed to be some use preferences, especially in the choice of forms for counter-expectation and emphatic markings where  $\sin^1 zi^3$  先至 and  $\sin_{c-d}$  were less preferred than  $zi^3$  至 as reflected by our data of the mid-20th century, the four usages have until so far been shared within the synonym group. Also to be noted in the mid-20th century, the use of  $\sin_{a-b}$  for temporal and conditional markings was not only preferred over  $\sin^1 zi^3$  先至 but also no less frequently observed than  $zi^3$  至.

To conclude our findings in this section, the semantic extension of the preverbal  $\sin^1$  先 had yielded  $\sin_{a-d}$  for 'subsequent' and related meanings by the early 1950s, which was about the same time that  $\sin^1$  先 'first' completed a radical word order change from preverbal to postverbal as shown in the previous section. The interaction of the syntax-semantic changes of  $\sin^1$  先 will be discussed in the following section.

## 5. Discussion

In historical linguistics, internal and external causes are distinguished to account for language change (Campbell 2013: 325–326). Building on the analyses presented above, the emergence of the postverbal  $\sin^1$  先 in Cantonese is a case exemplifying the interaction of these two types of causes, for which a two-stage process can be reconstructed as follows.

The first stage, which covers a time span of at least a century reaching into the 1930s, sees the emergence of  $\sin_1$  as an optional but uncommon usage in Cantonese. At this stage, a syntactic diffusion, with respect to the postverbal positioning of

---

32. With regard to this, a reviewer reminded us of some narrative texts in colloquial Cantonese composed in the 1930s. Unfortunately, we were aware of their existence but temporarily not able to have access to them. We hope that these materials can be included in our following research.

temporal adverbs for ‘first’, spread from TK and HM languages to influence neighboring Chinese dialects in southern coastal China. As noted above, the emergence of *sin<sub>1</sub>* in Guangdong Cantonese may not be due to the direct influence of TK and HM. Guangxi dialects, which were in the front lines of the diffusion, first rearranged the word order of *xiān* 先 ‘first’ from preverbal to postverbal under the direct and intense contact with TK and HM before passing this grammatical feature eastward to other Chinese dialects. Afterwards, Guangdong Cantonese, which was not deeply surrounded by TK and HM, is believed to have replicated this areally diffused feature under the prolonged influence of neighboring Guangxi Cantonese. The consequence of this intra-Cantonese contact can be attested in our historical materials dating from the 1820s to the 1930s, where *sin<sub>1</sub>* showed sporadic occurrences in mainstream Cantonese. Thus, language contact among TK, HM and Chinese dialects, as an external cause, is responsible for the early emergence of the postverbal *sin<sup>1</sup>* 先 ‘first’ in Cantonese.

The second stage, dated from around the 1940s, is dually featured with a dramatic word order shift of *sin<sup>1</sup>* 先 ‘first’ from preverbal to postverbal and a semantic extension of *sin<sup>1</sup>* 先 into a ‘precedent-subsequent’ polysemy, whereby the former was triggered by the latter. Once *sin<sup>1</sup>* 先 had developed such a polysemy, it yielded a strong motivation to formally differentiate and reassign its opposite meanings. Sybesma (2013) notices that there is a subtle tonal difference between *sin<sup>1</sup>* 先 ‘first’ and *sin<sub>a~d</sub>*, i.e. the former with a high-level tone and the latter with a high-falling tone. However, such a difference is not phonologically distinctive in Cantonese and thus does not suffice as a good formal means to differentiate meanings. Out of the above motivation, Cantonese resorted to word order by reassigning the ‘precedent’ meaning mainly to the existing *sin<sub>1</sub>* and meanwhile using the preverbal *sin<sup>1</sup>* 先 mainly for the ‘subsequent’ meaning. This functional divergence began to surface around the 1940s. *Sin<sub>1</sub>* used to represent a minor and marked word order of *sin<sup>1</sup>* 先 ‘first’, whose use was conditioned by discourse factors.<sup>33</sup> As more and more Cantonese speakers would prefer to use *sin<sub>1</sub>* to express the ‘precedent’ meaning, its use frequency was on the increase and its markedness on the decrease. The preverbal *sin<sup>1</sup>* 先 gradually transferred the ‘precedent’ meaning to *sin<sub>1</sub>* while extending its function into the ‘subsequent’ and other related meanings. As this divergence widened, the usage of *sin<sub>1</sub>* was finally transformed from discourse-prominent to syntax-prominent, representing a process of semantically driven syntacticization. That is to say, the semantic motivation has led to a syntactic

33. For example, Cantonese speakers might have perceived *sin<sub>1</sub>* as an emphatic usage in certain contexts (e.g. to highlight the ordering of sequential events) or as a newly-introduced variant with a certain sociolinguistic connotation (e.g. in an intention to demonstrate a novel and alien usage).

consequence on *sin*<sup>1</sup> 先 ‘first’. The postverbal positioning of *sin*<sup>1</sup> 先 ‘first’ has been established as a major and unmarked word order, which is syntactically required rather than discourse-conditioned. Thus, the second stage is internally motivated by syntax-semantic factors within Cantonese.

Notably, the word order and semantic changes of *sin*<sup>1</sup> 先 are intertwined aspects in its evolution into a multi-functional word. The appearance of *sin*<sub>1</sub> paved the way for its lexicalization with *zi*<sup>3</sup> 至 and semantic extension into *sin*<sub>a~d</sub>; and whereas the formation of *sin*<sub>a~d</sub> in reverse, pushed *sin*<sub>1</sub> to further develop into a syntax-prominent usage.

It is not uncommon across Chinese dialects to use word order as a mechanism for differentiating opposite meanings. Cantonese *sin*<sup>1</sup> 先 is not the only case to exemplify this mechanism. S. Lin (2008) has noted that in some southern Chinese dialects, such as Shadong Xiang, Liancheng Hakka, and Daye Gan, the double object construction (DOC) tends to encode a pair of opposite meanings, i.e. ‘give’ and ‘take’ meanings, in different word orders. Take the DOC in Shadong Xiang for example, the word order of VO<sub>1</sub>O<sub>D</sub> typically encodes the ‘take’ meaning as shown in (59a) whereas VO<sub>D</sub>O<sub>1</sub>, also known as the Inverted DOC, exclusively encodes the ‘give’ meaning as shown in (59b) and (59d). Thus, a ‘give’-type verb like 等 ‘give’ used in VO<sub>1</sub>O<sub>D</sub> will make the construction ungrammatical as shown in (59c).

Shadong Xiang

- (59) a. 你 擔 起 你 只 爺 一 杯 水。 (VO<sub>1</sub>O<sub>D</sub>)  
 2SG carry ASP 2SG CL father one cup water  
 ‘You take a cup of water from your father.’
- b. 你 擔 杯 水 你 只 爺 嚟。 (VO<sub>D</sub>O<sub>1</sub>)  
 2SG carry cup water 2SG CL father SFP  
 ‘You give a cup of water to your father.’
- c. \*你儂 等 渠 滴 東西。 (VO<sub>1</sub>O<sub>D</sub>)  
 2.PL give 3SG CL thing
- d. 你儂 等 滴 東西 渠。 (VO<sub>D</sub>O<sub>1</sub>)  
 2.PL give CL thing 3SG  
 ‘You give something to her.’ (S. Lin 2008)<sup>34</sup>

Although the DOC in the above-mentioned dialects and the Cantonese *sin*<sup>1</sup> 先 in most aspects bear little relevance to each other, they both utilize word order as a formal means to differentiate opposite meanings. Moreover, in S. Lin’s (2008) study as well as ours, semantics is regarded as a crucial perspective in explaining the word order idiosyncrasies in southern Chinese dialects. Meanwhile, the case of DOC can

34. The original source did not indicate pronunciation.

to some extent lend support to the reliability of our reconstruction of the second stage, a history that is yet to be fully attested.<sup>35</sup>

We have to admit that there are grammatical variations among Cantonese dialects in different subgroups and regions (see Kwok et al. 2016) while our study is based on historical materials representing the Guangfu subgroup of Cantonese. Therefore, it leaves open the question for further exploration whether or to what extent our analysis on the emergence of the postverbal *sin<sup>1</sup>* 先 applies to other subgroups Cantonese. For example, Gan (2010: 102) observes that the Siyi (四邑) subgroup of Cantonese uses either *zi<sup>3</sup>* 至 or *sin<sub>a</sub>* to indicate the ‘subsequent’ meaning but does not have the word *sin<sup>1</sup>zi<sup>3</sup>* 先至. Given this fact, Gan speculates that *sin<sup>1</sup>zi<sup>3</sup>* 先至 must be a later development than *sin<sub>a</sub>* in Cantonese. Apparently, Gan’s observation seems to be challenging our reconstruction on the semantic change of *sin<sup>1</sup>* 先. However, there is no evidence to show whether *sin<sup>1</sup>zi<sup>3</sup>* 先至 had never existed in Siyi Cantonese or had fallen out of use. On the other hand, if influence among Cantonese dialects is taken into consideration, the emergence of *sin<sub>a</sub>* in Siyi Cantonese can also be credited to language contact since dialects of lower prestige (e.g. the Siyi subgroup) tend to adopt linguistic features from dialects of higher prestige (e.g. the Guangfu subgroup).

Finally, there is one last question on which we wish to make a few comments. If a broader look is taken at Chinese dialects, we shall also identify a postverbal *xiān* 先 ‘first’ in some dialects of Wu, Hui, Min, Gan, and Xiang (see Map 85 in Cao 2008). For example, there are *ciā<sup>334</sup>* 先 ‘first’ appearing in *ŋie<sup>221</sup>mɔ̃<sup>221</sup>kʰɿ<sup>11</sup>* *ciā<sup>334</sup>* 你莽去先 2.PL-go-first ‘you go first’ in Qingyuan Wu (Cao et al. 2000: 452), *ciē<sup>534</sup>* 先 ‘first’ in *i<sup>52</sup>kʰu<sup>55</sup>kuə ciē<sup>534</sup>* 義去過先 2sg-go-cross-first ‘you go first’ in Sui’an Hui (Cao 1996: 198), and *tɕʰiŋ<sup>53</sup>* 先 ‘first’ in *ni<sup>42</sup>məŋ<sup>21</sup>gian<sup>42</sup>tɕʰiŋ<sup>53</sup>* 你懵行先 2.PL-go-first ‘you go first’ in Shipi Min (Akitani 2008: 382). Do these facts also imply the possibility of non-Sinitic influence?

However, this is not an easy conclusion to reach as those dialects in or near eastern China might not have been in the same linguistic situation as Cantonese or other dialects in southern coastal China. On the one hand, there does not appear to be any historical evidence to prove the existence of a counterpart of the postverbal *xiān* 先 ‘first’ in proto-TK or proto-HM which are believed to have had imposed substratal influences on proto-Wu and proto-Min (see Ballard 1989). On the other hand, synchronic evidence suggests that the postverbal *xiān* 先 ‘first’ in

35. There have not been many studies on word order as a formal means to differentiate opposite meanings in Chinese dialects. S. Lin (2008) is one of the recent few. We cite her study as an example to evidence the interaction between word order and antonymy in Chinese dialects. Thanks to the reviewer who suggested further elaboration on this point.

those dialects can be contact-induced within Chinese dialects. Our comments will focus on the following two geographic areas.

In eastern China, a resultative *qǐ* 起 ‘rise’ is grammaticalized to mark temporal precedence in most Gan and Hui dialects, as well as many of Wu dialects in central Zhejiang. For example, there are *tɕʰi³* 起 ‘first’ in *lan*<sup>6</sup> *kʰɛ² tɕʰi⁷ tɕʰi³* 讓渠喫起 let-3SG-eat-first ‘let him eat first’ in Susong Gan (Li & Chang 1992: 448), *tsʰi²¹³* 起 ‘first’ in *n²¹³ kʰi³⁵ tsʰi²¹³* 爾去起 2SG-go-first ‘you go first’ in Jixi Hui (Hirata 1998: 301) and *kʰi³²⁵* 起 ‘first’ in *nian*<sup>35</sup> *gei*<sup>24</sup> *kʰuən*<sup>55</sup> *kʰi³²⁵* 讓佢睏起 let-3SG-sleep-first ‘let him sleep first’ in Tiantai Wu (Dai 2003: 148). Map 85 of the *Linguistic atlas of Chinese dialects: Grammar* (Cao 2008) has indicated 106 survey sites where *qǐ* 起 ‘rise’ is also used to mean ‘first’. These sites are clustered in the Jiangxi-Zhejiang-Anhui area. Therefore, it is reasonable to suggest that some dialects, especially those Wu and Min dialects located in southern Zhejiang and northern Fujian, have developed a postverbal *xiān* 先 ‘first’ by replicating the word order of *qǐ* 起 ‘first’ from neighboring dialects. Notably, this process of grammatical replication is entirely within Chinese dialects.

In central Hunan, Chinese dialects tend to show a different typological feature by grammaticalizing a locative noun *tóu* 頭 ‘head, front’ into a postverbal element for ‘first’. Map 85 of the *Linguistic atlas of Chinese dialects: Grammar* (Cao 2008) has marked this feature on 10 survey sites, most of which belong to Xiang. For example, there are *tiə¹³* 頭 ‘first’ in *ŋ⁴² koŋ⁴² tiə¹³* 你講頭 2SG-speak-first ‘you say first’ in Lianyuan Xiang (H. Chen 1999: 278–279) and *dʰiə¹* 頭 ‘first’ in *ŋ²¹ yō¹³ dʰiə¹³* 你行頭 2SG-go-first ‘you go first’ in Xinhua Xiang (Luo 1998: 296). Therefore, the emergence of the postverbal *xiān* 先 ‘first’ in other Hunan dialects can be contact-induced by neighboring dialects that use a postverbal element *tóu* 頭 ‘first’. However, there still lurks the question as to whether the grammaticalization path from ‘front’ > ‘first’ in Xiang is internally motivated or contact-induced.<sup>36</sup>

## 6. Conclusion

To recapitulate, this paper explores the emergence of the postverbal *sin¹* 先 in Cantonese. In the previous studies of the postverbal *sin¹* 先, neither the relationship between internal and external causes nor the interaction of syntax and semantics behind its emergence has been put under close scrutiny. In this paper, a two-stage

36. This will certainly be the topic of another article.

process driven by both internal and external causes has been proposed based on Cantonese historical materials to account for the emergence of the postverbal *sin<sup>1</sup>* 先. In the first stage, the postverbal *sin<sup>1</sup>* 先 appeared in Cantonese due to a syntactic diffusion from TK and HM languages, while this word order had remained as a minor type for *sin<sup>1</sup>* 先 ‘first’ in historical materials dated from the 1820s through to the 1930s. In the second stage, *sin<sup>1</sup>* 先 ‘first’ underwent a radical word order change from mainly preverbal to dominantly postverbal around the 1940s, which has been proved to be triggered by the semantic motivation to formally differentiate the newly emerged ‘precedent-subsequent’ polysemy of *sin<sup>1</sup>* 先. The semantically driven reordering of *sin<sup>1</sup>* 先 ‘first’ explains why Cantonese shows a stronger tendency to place the adverb for ‘first’ in the postverbal position than other dialects in southern coastal China.

Due to the limited scope of historical materials and synchronic data under survey in this study, our discussion might have raised more issues than have been resolved. For example, *sin<sup>1</sup>zi<sup>3</sup>* 先至 has been examined in an attempt to recognize it as a bridging form to introduce *sin<sub>a~b</sub>*, but their developments have not found detailed evidence from historical materials. The word order change of *sin<sup>1</sup>* 先 only reflects the syntactic aspect of its evolution into multi-functionality; thus, what interests us most is to what extent this multi-functionality can be observed in languages of southern coastal China and whether *sin<sup>1</sup>* 先 just represents a very isolated case. When historical investigation is combined with cross-linguistic comparison, interesting (or sometimes disturbing) findings may arise to shed new light on debated issues, or even challenge some conventional views. Thus, the diachronic-typological approach taken in this study may also apply to other unique grammatical features in Cantonese, which will be the topics for our future projects.

## Acknowledgements

This work was jointly supported by the Shanghai University of Finance and Economics (SUFU) Research Start-Up Grant and the SUFE International Cultural Exchange School Research Support Scheme. I am grateful to the two anonymous reviewers for their insightful comments and suggestions. Needless to say, any remaining errors are of my own responsibility.

## Abbreviations

1	first person
2	second person
3	third person
Adv	adverb
ASP	aspect marker
CEX	counter-expectation marker
CHKC	<i>A linguistic corpus of Mid-20th Century Hong Kong Cantonese (1952–1966)</i>
CL	classifier
COND	conditional
COP	copular
DOC	double object construction
ECCT	<i>Early Cantonese colloquial texts: A database (1828–1924)</i>
ECTD	<i>Early Cantonese tagged database (1872–1931)</i>
EMP	emphatic marker
GD	Guangdong
GX	Guangxi
HM	Hmong-Mien
HN	Hainan
IMP	imperative marker
INTERJ	interjection
M	model language
MSEA	Mainland Southeast Asia
NP	noun phrase
O	object
O <sub>D</sub>	direct object in a DOC
O <sub>I</sub>	indirect object in a DOC
PL	plural
PRT	particle
R	replica language
S	subject
SFP	sentence-final particle
SG	singular
<i>sin</i> <sub>1</sub>	temporal <i>sin</i> <sup>1</sup> 先
<i>sin</i> <sub>2</sub>	imperative <i>sin</i> <sup>1</sup> 先
<i>sin</i> <sub>3</sub>	emphatic <i>sin</i> <sup>1</sup> 先
<i>sin</i> <sub>a</sub>	<i>sin</i> <sup>1</sup> 先 to mark temporal subsequence
<i>sin</i> <sub>b</sub>	<i>sin</i> <sup>1</sup> 先 to mark a conditional apodosis
<i>sin</i> <sub>c</sub>	<i>sin</i> <sup>1</sup> 先 to express counter-expectation
<i>sin</i> <sub>d</sub>	<i>sin</i> <sup>1</sup> 先 to express a contrastive emphasis
TB	Tibeto-Burman
TK	Tai-Kadai
V	verb
VP	verb phrase

## Appendix. List of selected Cantonese historical materials

1. **Early Cantonese colloquial texts: A database** 早期粵語口語文獻資料庫 (1828–1924)  
 Ball, J. Dyer. 1888. *Cantonese made easy*. 2nd edn. Hong Kong: China Mail Office.  
 Ball, J. Dyer. 1924. *Cantonese made easy*. 4th edn. Hong Kong: Kelly & Walsh.  
 Bridgman, E. C. 1841. *A Chinese chrestomathy in the Canton dialect*. Macao: S. Wells Williams.  
 Morrison, Robert. 1828. *Vocabulary of the Canton dialect* 廣東省土話字彙. Macao: The Honorable East India Company's Press.  
 (<http://pvs0001.ust.hk/Candbase>) (Accessed 2017-03-26.)
2. **Early Cantonese tagged database** 早期粵語標註語料庫 (1872–1931)  
 Anonymous. 1872. *The gospel according to St. Mark in the Cantonese colloquial* 馬可傳福音書 (廣東土白). Shanghai: American Presbyterian Mission Press.  
 Anonymous. 1877. *Forty chapters on Cantonese phrases* 散語四十章. Hong Kong: St. Paul's College.  
 Ball, J. Dyer. 1902. *How to speak Cantonese*. 2nd edn. Hong Kong: Kelly & Walsh.  
 Ball, J. Dyer. 1912. *How to speak Cantonese*. 4th edn. Hong Kong: Kelly & Walsh.  
 Bruce, E. Donald. 1877. *Easy phrases in the Canton dialect of the Chinese language* 英華常語合璧. 2nd edn. San Francisco: Bruce's Printing House.  
 Fulton, A. A. 1931. *Progressive and idiomatic sentences in Cantonese colloquial*. Hong Kong: Kelly & Walsh.  
 (<http://pvs0001.ust.hk/WTagging>) (Accessed 2017-03-26.)
3. **A linguistic corpus of Mid-20th Century Hong Kong Cantonese** 香港二十世紀中期粵語語料庫 (1952–1966)  
*A Ready Lover* 十月芥菜 (1952)  
*As Luck Will Have It* 鴻運喜當頭 (1955)  
*Foster-Daddy's Romantic Affairs* 契爺艷史 (1952)  
*If Only We'd Met When I Was Single* 恨不相逢未嫁時 (1955)  
*One Queen and Three Kings* 一后三王 (1963)  
*The Ideal Husband* 標準丈夫 (1965)  
*To Capture the God of Wealth* 千方百計搶財神 (1962)  
*Violet Girl* 我愛紫羅蘭 (1966)  
*Wrong Wedding* 錯燒龍鳳燭 (1954)  
 (<http://corpus.ied.edu.hk/hkcc/corpus>) (Accessed 2017-03-26.)
4. **Other sources**  
 Chao, Yuen Ren. 1947a. *Cantonese primer* 粵語入門. Cambridge: The Harvard-Yenching Institute, Harvard University Press.  
 Chao, Yuen Ren. 1947b. *Character text for Cantonese primer*. Cambridge: The Harvard-Yenching Institute, Harvard University Press.  
 O'Melia, Thomas A. 1941. *First year Cantonese: Part one*. 2nd edn. 綜合粵語學習法. Hong Kong: Maryknoll House.

## References

- Abraham, Werner. 1991. The grammaticalization of the German modal particles. In Traugott, Elizabeth Closs & Heine, Bernd (eds.), *Approaches to grammaticalization volume II: Types of grammatical markers* (Typological Studies in Language 19:2), 331–380. Amsterdam: John Benjamins. doi:10.1075/tsl.19.2.17abr
- Akitani, Hiroyuki. 2008. *Minbeiqu san xian shi fangyan yanjiu* [Studies on Northern Min dialects: Shipi, Zhenqian, and Dikou] (Language and Linguistics Monograph Series A12-2). Taipei: Institute of Linguistics, Academia Sinica.
- Anderson, Henning. 1973. Abductive and deductive change. *Language* 49(4). 765–793. doi:10.2307/412063
- Ballard, William L. 1989. Pig, tone sandhi and Wumin. *Cahiers de Linguistique Asie Orientale* 18(2). 163–182. doi:10.3406/clao.1989.1302
- Brinton, Laurel J. & Traugott, Elizabeth Closs. 2005. *Lexicalization and language change* (Research Surveys in Linguistics). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511615962
- Campbell, Lyle. 2013. *Historical linguistics: An introduction*. 3rd edn. Edinburgh: Edinburgh University Press.
- Cao, Zhiyun. 1996. *Yanzhou fangyan yanjiu* [Studies on the Yanzhou dialect]. Tokyo: Kohbun.
- Cao, Zhiyun. 2008. *Hanyu fangyan dituji: Yufa juan* [Linguistic atlas of Chinese dialects: Grammar]. Beijing: The Commercial Press.
- Cao, Zhiyun & Akitani, Hiroyuki & Ōta, Itsuku & Zhao, Rixin. 2000. *Wuyu Chuqu fangyan yanjiu* [A study of the Chuqu subgroup of Wu]. Tokyo: Kohbun.
- Chang, Song Hing. 2004. *Lianzhou Tuhua yanjiu* [Studies on Lianzhou Tuhua] (Yuebei Tuhua Yanjiu Congshu [Northern Yue Tuhua Research Series]). Xiamen: Xiamen University Press.
- Chen, Hui. 1999. *Lianyuan fangyan yanjiu* [Studies on the Lianyuan dialect]. Changsha: Hunan Education Publishing House.
- Chen, Xiaojin. 2004. *Guangxi Yulinshi Kejia fangyan diaocha yanjiu* [Investigation and research on Yulin Hakka in Guangxi]. Beijing: China Social Sciences Press.
- Chen, Xiaojin & Chen, Tao. 2005. *Guangxi Beihaiishi Yue fangyan diaocha yanjiu* [Investigation and research on Beihai Cantonese in Guangxi]. Beijing: China Social Sciences Press.
- Cheung, Hung-nin Samuel. 2007. *Xianggang Yueyu yufa de yanjiu (Zengding ban)* [A grammar of Cantonese spoken in Hong Kong]. Revised edn. Hong Kong: The Chinese University Press.
- Dai, Zhaoming. 2003. *Tiantai fangyan chutan* [Exploration of the Tiantai dialect]. Beijing: China Social Sciences Press.
- Department of Linguistics and Modern Language Studies, The Education University of Hong Kong. 2012. *A linguistic corpus of Mid-20th Century Hong Kong Cantonese*. (<http://corpus.ied.edu.hk/hkcc/corpus>) (Accessed 2017-03-26.)
- Diller, Anthony V. N. & Edmondson, Jerold A. & Luo, Yongxian (eds.). 2008. *The Tai-Kadai languages* (Routledge Language Family Series). London: Routledge.
- Enfield, N. J. 2005. Areal linguistics and Mainland Southeast Asia. *Annual Review of Anthropology* 34. 181–206. doi:10.1146/annurev.anthro.34.081804.120406
- Feng, Chengbao. 1992. Hainan Sheng Qionghai fangyan jilue [A study on the Qionghai dialect of Hainan Province]. *Guangdong Minzu Xueyuan Xuebao (Shehui Kexue Ban)* [Journal of Guangdong Institute for Nationalities (Social Sciences Edition)] 1992(2). 41–59, 82.

- Feng, Li. 2007. Cong beibu Wuyu de “V kuai” kan zhongxin weiyu chengfen xuhua wei zhuci de xianxiang [From main predicate to aspect particle: A case study of “V kuai(快)” in the northern Wu dialects]. *Zhongguo Yuwen* [Studies of the Chinese Language] 2007(3). 248–255.
- Fischer, Hans Rudi. 2001. Abductive reasoning as a way of worldmaking. *Foundations of Science* 6(4). 361–383. doi:10.1023/A:1011671106610
- Gan, Yu'en. 2010. *Guangdong Siyi fangyan yufa yanjiu* [A grammatical study of the Siyi dialect in Guangdong] (Nanfang Yuyanxue Congshu [Southern Linguistics Series]). Guangzhou: Jinan University Press.
- Gao, Huanian. 1980. *Guangzhou fangyan yanjiu* [Studies on the Guangzhou dialect]. Hong Kong: The Commercial Press.
- Ge, Benyi. 2014. *Xiandai Hanyu cihuixue* [Modern Chinese lexicology]. 3rd edn. Beijing: The Commercial Press.
- Hashimoto, Mantoro J. 2008[1978]. *Yuyan dili leixingxue* [Linguistic geographical typology]. Beijing: World Books. (Translated by Yu, Zhihong from Hashimoto. (1978) *Gengo ruikai chiriron* [Linguistic geographical typology]. Tokyo: Kobundo.)
- Heine, Bernd. 2008. Contact-induced word order change without word order change. In Siemund, Peter & Kintana, Noemi (eds.), *Language contact and contact languages* (Hamburg Studies on Multilingualism 7), 33–60. Amsterdam: John Benjamins. doi:10.1075/hsn.7.04hei
- Heine, Bernd & Kuteva, Tania. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511613463
- Heine, Bernd & Kuteva, Tania. 2003. On contact-induced grammaticalization. *Studies in Language* 27(3). 529–572. doi:10.1075/sl.27.3.04hei
- Heine, Bernd & Kuteva, Tania. 2005. *Language contact and grammatical change* (Cambridge Approaches to Language Contact). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511614132
- Heine, Bernd & Kuteva, Tania. 2006. *The changing languages of Europe* (Oxford Linguistics). Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199297337.001.0001
- Hirata, Shoji. 1998. *Huizhou fangyan yanjiu* [A study of the Huizhou dialects]. Tokyo: Kohbun.
- Hong Kong University of Science and Technology. 2012. *Early Cantonese colloquial texts: A database*. (<http://pvs0001.ust.hk/Candbase>) (Accessed 2017-03-26.)
- Hong Kong University of Science and Technology. 2012. *Early Cantonese tagged database*. (<http://pvs0001.ust.hk/WTagging>) (Accessed 2017-03-26.)
- Hopper, Paul J. & Traugott, Elizabeth Closs. 2003. *Grammaticalization* (Cambridge Textbooks in Linguistics). 2nd edn. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139165525
- Huang, Xing & Daobu (eds.). 2012. *Zhongguo yuyan ditu ji: Shaoshu minzu yuyan juan* [Language atlas of China: Minority languages], vol. 2. Beijing: The Commercial Press.
- Kwok, Bit-Chee & Chin, Andy C. & Tsou, Benjamin K. 2016. Grammatical diversity across the Yue dialects. *Journal of Chinese Linguistics* 44(1). 109–152.
- Li, Jinfang. 1990. Lun Zhuangtongyu dui Yueyu de yingxiang [The influences of Tai-Kadai languages on Cantonese]. *Guizhou Minzu Yanjiu* [Guizhou Ethnic Studies] 1990(4). 60–65.
- Li, Rulong & Chang, Song Hing. 1992. *Kegan fangyan diaocha baogao* [A survey report of the dialects of Hakka and Gan]. Xiamen: Xiamen University Press.
- Liang, Weihua & Lin, Yi. 2009. *Guangxi Chongzuo Xinhe Zheyuanhua yanjiu* [Studies on the Zheyuan dialect in Xinhe, Chongzuo of Guangxi]. Guilin: Guangxi Normal University Press.
- Lin, Lunlun. 2006. *Yuexi Minyu Leizhouhua yanjiu* [Studies on Leizhou Min in West Guangdong]. Beijing: Zhonghua Book Company.

- Lin, Su'e. 2008. Hanyu nanfang fangyan daozi shuangbin jiegou chutan [A study on inverted double object construction in Chinese southern dialects]. *Yuyan Kexue* [Linguistic Sciences] 2008(3). 308–319.
- Liu, Cunhan. 1995. *Liuzhou fangyan cidian* [The dictionary of the Liuzhou dialect] (Xiandai Hanyu Fangyan Dacidian: Fenjuan [Modern Chinese Dialects Dictionary: Separate Volume]). Nanjing: Jiangsu Education Publishing House.
- Liu, Danqing. 2000. Yueyu jufa de leixingxue tedian [Typological characteristics of Cantonese syntax]. *Yatai Yuwen Jiaoyu Xuebao* [Asian Pacific Journal of Language in Education] 3(2). 1–30.
- Liu, Mei-chun. 1993. *Discourse, grammar, and grammaticalization: Synchronic and diachronic analysis of Mandarin adverbial markers Jiu and Cai*. Boulder: University of Colorado at Boulder. (Doctoral dissertation.)
- Lü, Shuxiang (ed.). 1999. *Xiandai Hanyu babai ci (Zengding ben)* [Eight hundred words in Modern Chinese]. Revised edn. Beijing: The Commercial Press.
- Lucas, Alain & Xie, Honghua. 1994. "There is practically one universal Chinese grammar": A propos de *xian1* 先 en mandarin et *sin1* en cantonais [About *xian1* 先 in Mandarin and *sin1* in Cantonese]. *Cahiers de Linguistique Asie Orientale* 23(1). 189–206. doi:10.3406/clao.1994.1779
- Luo, Xinru. 1998. *Xinhua fangyan yanjiu* [Studies on the Xinhua dialect]. Changsha: Hunan Education Publishing House.
- Mai, Yun. 1993. Guangzhouhua "xian" de zai fenxi [Revisiting "xian" in the Guangzhou dialect]. In Cheng, Ting Au & Zhou, Xiaobin (eds.), *Guangzhouhua yanjiu yu jiaoxue* [Research and teaching of the Guangzhou dialect], 63–73. Guangzhou: Sun Yat-sen University Press.
- Mao, Zongwu & Meng, Chaoji & Zheng, Zongze (eds.). 1982. *Yaozu yuyan jianzhi* [A sketch of the languages of the Yao people]. Beijing: The Ethnic Publishing House.
- Matisoff, James A. 1991. Sino-Tibetan linguistics: Present state and future prospects. *Annual Review of Anthropology* 20. 469–504. doi:10.1146/annurev.an.20.100191.002345
- Matthews, Peter H. 2007. *The concise Oxford dictionary of linguistics* (Oxford Paperback Reference). 2nd edn. Oxford: Oxford University Press.
- Matthews, Stephen & Yip, Virginia. 1994. *Cantonese: A comprehensive grammar*. London: Routledge.
- Ouyang, Jueya. 1998. *Cunyu yanjiu* [A study of Cun]. Shanghai: Shanghai Yuandong Press.
- Ouyang, Jueya & Cheng, Fang & Yu, Cuirong. 1984. *Jingyu jianzhi* [A sketch of Jing]. Beijing: The Ethnic Publishing House.
- Peng, Rui. 2014. The diachronic development of *zaishuo* in Chinese: A case of polygrammaticalization chains. *Journal of Chinese Linguistics* 42(2). 351–387.
- Sun, Hongkai & Hu, Zengyi & Huang, Xing (eds.). 2007. *Zhongguo de yuyan* [The languages of China]. Beijing: The Commercial Press.
- Sybesma, Rint. 2013. Cantonese *sin* 先 and the question of microvariation and macrovariation. In Cao, Guangshun & Chappell, Hilary & Djamouri, Redouane & Wiebusch, Thekla (eds.), *Breaking down the barriers: Interdisciplinary studies in Chinese linguistics and beyond*, vol. 2 (Language and Linguistics Monograph Series 50), 971–994. Taipei: Institute of Linguistics, Academia Sinica.
- Tang, Sze-Wing. 2006. Yueyu yiwen ju mo "xian" zi de jufa tedian [Syntactic properties of *sin* in Cantonese interrogatives]. *Zhongguo Yuwen* [Studies of the Chinese Language] 2006(3). 225–232.

- Tang, Sze-Wing. 2012. Yanyu de jufa fenxi: Yi Yueyu “xian” wei li [A syntactic analysis of the speech act domain (言域): Evidence from *sin* (先) in Cantonese]. *Yuyan Kexue* [Linguistic Sciences] 2012(1). 9–14.
- Tang, Sze-Wing. 2015. *Yueyu yufa jiangyi* [Lectures on Cantonese grammar]. Hong Kong: The Commercial Press.
- Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In Lehmann, Winfred P. & Malkiel, Yakov (eds.), *Perspectives on historical linguistics: Papers from a conference held at the meeting of the Language Theory Division, Modern Language Assn., San Francisco, 27–30 December 1979* (Current Issues in Linguistic Theory 24), 245–272. Amsterdam: John Benjamins. doi:10.1075/cilt.24.09clo
- Traugott, Elizabeth C. & Dasher, Richard B. 2001. *Regularity in semantic change* (Cambridge Studies in Linguistics 97). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511486500
- Wang, Feng. 2015. Multisyllabication and phonological simplification throughout Chinese history. *Journal of Chinese Linguistics* 43(2). 714–718. doi:10.1353/jcl.2015.0025
- Wang, Jun & Zheng, Guoqiao. 1980. *Mulaoyu jianzhi* [A sketch of Mulam] (Zhongguo Shaoshu Minzu Yuyan Jianzhi Congshu [Chinese Minority Languages Outline Series]). Beijing: The Ethnic Publishing House.
- Wang, Li & Qian, Xun. 1951. Hainandao Baisha Liyu chutan [A preliminary study on Baisha Li in Hainan]. *Lingnan Xuebao* [Lingnan Journal of Chinese Studies] 11(2). 253–300.
- Wei, Jingyun & He, Shuang & Luo, Yongxian. 2011. *Yanqi Zhuangyu cankao yufa* [A reference grammar of Yanqi Zhuang]. Beijing: China Social Sciences Press.
- Xing, Gongwan. 1985. *Sanjiang Dongyu* [Sanjiang Kam]. Tianjin: Nankai University Press.
- Xu, Liu. 1990. Lun tongyi fuci [On synonym compound words]. *Guhanyu Yanjiu* [Research in Ancient Chinese Language] 1990(4). 12–17.
- Zhang, Yuansheng & Ma, Jialin & Wen, Mingying & Wei, Xinglang. 1985. *Hainan Lingaohua* [Lingao of Hainan]. Nanning: Guangxi Nationalities Publishing House.
- Zheng, Yiqing. 1997. *Huihuihua yanjiu* [Studies on Huihui] (Zhongguo Xin Faxian Yuyan Yanjiu Congshu [Chinese New Discovery Language Study Series]). Shanghai: Shanghai Yuandong Press.
- Zheng, Yiqing. 2013. *Jingxi Zhuangyu yanjiu* [Studies on Jingxi Zhuang]. Nanning: Guangxi Nationalities Publishing House.

### Author's address

Yang Zhou  
International Cultural Exchange School  
Shanghai University of Finance and Economics  
369 North Zhongshan No. 1 Road  
Hongkou District, Shanghai 200083  
China  
zhou.nus@gmail.com

### Publication history

Date received: 15 June 2016

Date accepted: 29 March 2017