

An intermediate stage of spec-to-head reanalysis: evidence from Mandarin and Cantonese

Shuyan Wang
University of Connecticut

1. Introduction

Mandarin *Yi* ‘one’ has been suggested to have grammaticalized from a numeral to an indefinite article (e.g., Chen 2003; Zhang 2013), but no syntactic analysis of the grammaticalization process has been provided. This paper will go beyond previous analyses and systematically explore the grammaticalization process of Mandarin *yi*. In particular, I will propose that the grammaticalization of Mandarin *yi* is an ongoing process and *yi* is ambiguous between a numeral and an indefinite article, and that the grammaticalization of Mandarin *yi* is an instantiation of a well-known phenomenon: specifier-to-head reanalysis (e.g., Lyons 1999; Wood 2003; van Geldern 2001, 2004). I will further argue that there is an intermediate stage of grammaticalization: a numeral ‘one’ starts as a specifier of a phrase; then, it adjoins to a head before it is reanalyzed as an independent head projecting a separate phrase. In other words, at this intermediate stage, the indefinite article ‘one’ is head-adjoined to another head. Cantonese *jat* ‘one’ and Turkish *bir* ‘one’ will provide evidence for this intermediate stage.

The paper is organized as follows. Section 2 will discuss the unique characteristics of Mandarin *yi* ‘one’. A new account will be proposed in Section 3 to cover the characteristics of Mandarin *yi*. In particular, I will discuss the spec-to-head reanalysis of Mandarin *yi*. Then Section 4 will describe Cantonese *jat* ‘one’ and discuss the intermediate stage of grammaticalization. I will then provide Turkish data to support the proposal of an intermediate stage. Section 5 will conclude this paper.

2. Ambiguity of *yi*

The possibility of the existence of indefinite articles in Mandarin has been discussed in some works (e.g., Wang 1989; Chen 2003). For instance, Wang (1989) suggests that in Old Chinese, when the classifier *ge* occurs in CI-N phrases, it can serve as an indefinite article. In this respect, also relevant is Chen (2003), who discusses a grammaticalization process involving numeral *yi* ‘one’ and classifiers in Mandarin. Chen (2003) suggests that Mandarin has an indefinite article. In particular, he suggests that in *yi*-CI-N phrases, ‘*yi*+CI’ (*yi* ‘one’ and a classifier) together can be used as an indefinite article, while in CI-N phrases the classifier alone is an indefinite article¹.

¹ Note that Chen (2003) simply suggests that ‘*yi*+CI’ semantically function as an indefinite article, in that ‘*yi*+CI’ can be used in the situations where an indefinite article (e.g., *a/an* in English) is usually used. No specific syntactic analysis is, however, given in the paper.

Following Chen (2003), I argue that *yi* has grammaticalized from a numeral to an indefinite article and that *yi* is ambiguous between a numeral and an indefinite article. The ambiguity analysis can cover a range of observations. First, numeral expressions with *yi* ‘one’ in Mandarin can be interpreted as indefinite specific or non-specific, while numeral expressions with other numerals are usually interpreted only as indefinite non-specific (e.g., Huang 1987, Tsai 1996). The Numeral (Num)-Classifier (Cl)-Noun (N) phrases, like *san-ge xuesheng* (three-Cl-student) in (1), are standardly referred to as numeral expressions in Mandarin. Two types of numeral expressions are assumed to exist in Mandarin (e.g., Li 1998): individual-denoting and quantity-denoting expressions. The former refers to some entities/individuals, while the latter expresses the quantity of entities/individuals (e.g., Li 1998). For instance, the numeral expression in (1) refers to some students. The numeral expression in (2) just denotes the quantity and the verb expresses the sufficiency of a certain amount.

- (1). wo kandao san-ge xuesheng.
 I see three-Cl student
 ‘I saw three students.’
- (2). san-ge xuesheng bu gou.
 three-Cl student not enough
 ‘Three students is not enough.’

(Huang et al. 2009)

Individual-denoting numeral expressions in Mandarin are generally considered to be indefinite non-specific expressions (e.g., Huang et al. 2009). They usually do not appear in subject or topic positions (see (3b) and (3c)), since these positions in Mandarin do not allow non-specific readings (e.g., Chao 1968; Li & Thompson 1981; Lee 1986; Li 1996).

- (3). a. wo kandao-le san-ge xuesheng.
 I see-ASP two-Cl student
 ‘I saw two students.’
- b. ??san-ge xuesheng chi-le dangao.
 three-Cl student eat-ASP cake
 ‘Three students ate the cake.’
- c. *san-ge xuesheng, wo yiwei chi-le dangao.
 three-Cl student I think eat-ASP cake
 ‘Three students, I thought (they) ate the cake.’

(Huang et al. 2009)

(Huang et al. 2009)

As a quantity-denoting expression, a numeral phrase can occur in these positions. However, quantity expressions show different structural properties from individual-denoting numeral expressions. For example, quantity-denoting expressions cannot co-refer with a pronoun, as in (4). I will not go into these issues here, since they are not relevant to the main topic of the paper (for relevant discussion, see Li (1998) and Huang et al. (2009)). In this paper, I will focus on individual-denoting numeral expressions. The term ‘numeral expression(s)’ will be used to refer only to individual-denoting numeral expressions.

- (4). *liang-ge daren_i bu ru tamen_i de san-ge xiaohai you liqi.
 two-Cl adult not compare they DE three-Cl children have strength
 ‘Two adults are not as strong as their three children.’ (Huang et al. 2009)

As mentioned above, numeral expressions in Mandarin are indefinite non-specific expressions and are disallowed in subject/topic positions, as in (3). However, *yi* ‘one’-Cl-N phrases are different in this respect. They can be either specific or non-specific. Thus, Tsai (1996) noted that numeral expressions with *yi* ‘one’, but not other numerals like ‘two’ or ‘three’, can be specific (see also Huang 1987). As a result, *yi*-Cl-N phrases, unlike other numeral expressions, can occur in subject positions, as in (5).

- (5). yi-ge xuesheng chi-le dangao.
 one-Cl student eat-ASP cake
 ‘A student ate the cake.’

These data suggest that numeral expressions with *yi* ‘one’ in Mandarin can be interpreted as indefinite specific or non-specific, while numeral expressions with other numerals are usually interpreted only as indefinite non-specific. Then the question is why expressions with *yi* ‘one’ is different from other numeral expressions in this aspect. In fact, the proposal that *yi* is ambiguous between a numeral and an indefinite article can account for these data. I argue that *yi* in (5) actually is an article, not a numeral. As discussed above, numeral expressions in Mandarin are indefinite non-specific, so they cannot appear in subject/topic positions. However, *yi* can function as an indefinite article. Then it is not surprising that expressions with an indefinite article can be interpreted as indefinite specific and can occur as a subject or a topic in Mandarin.

Another unique feature of Mandarin *yi* is that when *yi*-Cl-N phrases stay under the scope of negation, a ‘not any’ interpretation may emerge. As illustrated by (6), a *yi*-Cl-N phrase stays under the scope of negation (i.e., *congmei* ‘never’) and then they can be interpreted as ‘not any’ (i.e., ‘Xiaohong has never had any boyfriend’).

- (6). Xiaohong congmei jiao-guo yi-ge nanpengyou.
 Xiaohong never make-EXP a-Cl boy.friend
 ‘Xiaohong has never had any boyfriend.’

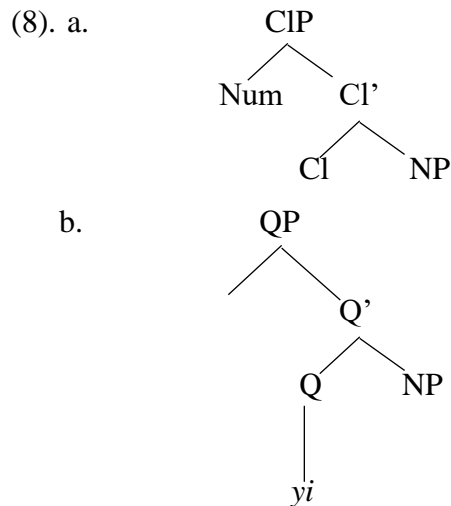
I argue that with the ‘not any’ interpretation, *yi* in (6) is an indefinite article, not a numeral. Similar patterns are found in English. As shown by (7a), when the object contains an indefinite article (i.e., *a boyfriend*), the sentence can be interpreted as ‘Mary has never had any boyfriend’. However, if the object is a numeral expression (i.e., *one boyfriend*), the ‘not any’ interpretation disappears, as in (7b).

- (7). a. Mary has never had a boyfriend.
 ‘Mary has never had any boyfriend.’
 b. *Mary has never had one boyfriend.
 Intended meaning: ‘Mary has never had any boyfriend.’

In summary, Mandarin *yi* demonstrates some unique characteristics, and differs from other numerals in some aspects. Numerals expressions in Mandarin are interpreted as indefinite non-specific. However, expressions with *yi* ‘one’ can be interpreted as specific. Furthermore, *yi*-Cl-N phrases under the scope of negation be interpreted as ‘not any’. This parallels the phrases with an indefinite article in English. I have argued that these unique features of Mandarin *yi* can be explained by its ambiguity status. In particular, I proposed that in these two special cases, *yi* functions as an indefinite article, rather than a numeral. The next section will focus on the grammaticalization process of Mandarin *yi* and I will propose syntactic analyses for the two different *yi*.

3. Specifier-to-head reanalysis

In terms of the diachronic reanalysis over the grammaticalization process, I propose that *yi* ‘one’ as a numeral stays in the specifier of CIP (i.e., Spec CIP), and then the grammaticalization process leads to a reanalysis as an indefinite article, which heads a Quantifier Phrase. As shown in (8a), numerals are located in the Spec CIP (following e.g., Tang (1990), I assume that a classifier heads its own projection above NP), while the indefinite article *yi* ‘one’ is a head projecting a Quantifier Phrase, as in (8b).



It should be noted that we may be dealing here with a more general grammaticalization process. Thus, by examining diachronic variations between Old English, Middle English, and Modern English, van Geldern (2001, 2004) argues that many functional categories developed as a result of a change of a specifier of a phrase to a head that projects a separate phrase. In fact, similar claims have also been made for many elements within CP (like complementizers) and DP, like demonstratives (see also Lyons 1999; Wood 2003; van Geldern 2001; Huddleston & Pullum 2002; Willis 2007, among others).

As discussed above, I argue that *yi* ‘one’ in Mandarin is ambiguous between an indefinite article and a numeral. As shown by (8a), a numeral is in the specifier of CIP, so its existence is closely related to the presence of a classifier. In other words, whenever there is a numeral, a

c. ??I bought a book.

In summary, this section has investigated the grammaticalization process. In particular, I argued that Mandarin *yi* ‘one’, as a numeral, is located in the specifier of a Classifier Phrase; a grammaticalization process leads to its reanalysis as the head of a Quantifier Phrase. This diachronic reanalysis is an instantiation of a well-known phenomenon: specifier-to-head reanalysis (e.g., Lyons 1999; Wood 2003; van Geldern 2001, 2004). The following section will propose an intermediate stage of grammaticalization with evidence from Cantonese.

4. Intermediate stage of grammaticalization

In this section, I will argue that *jat* ‘one’ in Cantonese is also undergoing the grammaticalization process from a numeral to an indefinite article. However, I will further argue that *jat* has not grammaticalized as much as *yi* in Mandarin: *jat* has not been reanalyzed as an independent head projecting a separate phrase; instead, *jat* is head-adjoined to the classifier head.

Numeral expressions in Cantonese are usually interpreted only as indefinite non-specific, while numerals expressions with *jat* ‘one’ can be interpreted as indefinite specific or non-specific. As illustrated by (12a), numeral expressions in Cantonese usually do not appear in subject/topic positions, since these positions usually do not allow indefinite non-specific expressions. However, a *jat*-Cl-N phrase in the subject/topic position is much better, as shown in (12b).

- (12). a. ??saam-go hoksaang mei lei.
three-Cl student not come
‘Three students didn’t show up.’
- b. jat-go hoksaang mei lei.
a- Cl student not come
‘A student didn’t show up.’

I argue that *jat* in (12b) actually is an article, not a numeral. Numeral expressions in Cantonese are indefinite non-specific, so they cannot appear in subject/topic positions. If *jat* can function as an indefinite article, then it is not surprising that *jat*-Cl-N phrases differ from other numeral expressions.

Furthermore, when *jat*-Cl-N phrases stay under the scope of negation, the ‘not any’ interpretation emerges, as in (13). As discussed above, Mandarin *yi* and English *a/an* show similar patterns, as in (6) and (7). This provides more support for the proposal that *jat* in (13) is an indefinite article, not a numeral.

- (13). Siuhung chung-mut gaau jat-go naampanyau.
Siuhung never made a-Cl boy.friend
‘Siuhung has never had any boyfriend.’

Therefore, Cantonese *jat* ‘one’ differs from other numerals in various aspects. Based on its unique features, I argue that *jat* in Cantonese is ambiguous between a numeral and an indefinite article. In these special cases (e.g., in subject/topic positions; interpreted as ‘not any’), *jat* function as an indefinite article, rather than a numeral. Then the next question is what the syntactic structure of *jat* is. If we follow the proposal for Mandarin, the structure should be like (14). In this structure,

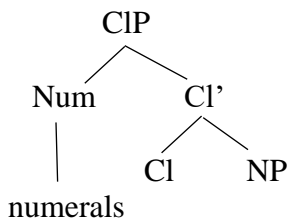
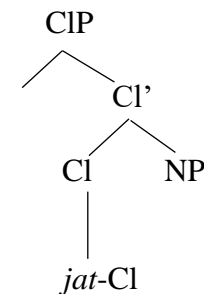
numerals stay in Spec CIP, but the indefinite article *jat* ‘one’ has been diachronically reanalyzed as the head of a QP, as in (14b).

- (14). a. [CIP samm [CI' -go [NP hoksaang]
 three CI student
 ‘three students’
- b. [QP [Q' jat [CIP [CI' -go [NP hoksaang]
 a CI student
 ‘a student’

According to the structures proposed in (14), the article *jat* then would not be dependent on the presence of a classifier and should be directly combinable with a noun. As discussed above, Mandarin *yi* can combine with a noun without a classifier in between. However, it is not the same case for Cantonese *jat*. Cantonese *jat* always needs a classifier, as in (15).

- (15). *jat*-(go) hoksaang
 a/one-CI student
 ‘a student/one student’

One potential explanation for this cross-linguistic variation is that *jat* in Cantonese and *yi* in Mandarin have grammaticalized to different degrees. *Jat* ‘one’ may not have grammaticalized as much as *yi* ‘one’. Then, even when *jat* ‘one’ functions as an indefinite article, it still requires the presence of a classifier. Based on this, I argue that the current structure for Cantonese numeral phrases should be like (16). In (16a), numerals, including the numeral *jat* ‘one’, stay in Spec CIP. On the other hand, the article *jat* is adjoined to the classifier head, as in (16b). The expectation is that in a further stage of grammaticalization, the indefinite article *jat* will be reanalyzed as a head heading its own projection, as in Mandarin. At that time, *jat*-N phrases may be expected to emerge in Cantonese.

- (16). a. 
- b. 

As discussed above, van Geldern (2001, 2004) shows that many functional categories developed from a specifier of a phrase to a head that projects a separate phrase (see also Lyons 1999; Wood 2003; Huddleston & Pullum 2002; Willis 2007). Here I propose that before being reanalyzed as a head that can project by itself, a functional category first adjoins to another head. Applying this more general pattern of grammaticalization, numeral ‘one’ starts as a specifier of a phrase; then, in the grammaticalization process, it adjoins to a head before it is reanalyzed as a head that can project a phrase by itself. Regarding why this intermediate stage exists, it is standardly assumed that when a specifier is merged into the structure, its sister projects. I suggest that when a specifier

5. Conclusion

This paper has systematically discussed the grammaticalization process of Mandarin *yi*. In particular, I proposed that Mandarin *yi* is ambiguous between a numeral and an indefinite article, and that the grammaticalization of Mandarin *yi* is an instantiation of a well-known phenomenon: specifier-to-head reanalysis (e.g., Lyons 1999; Wood 2003; van Geldern 2001, 2004). Then I argued that Cantonese *jat* ‘one’ is also undergoing a grammaticalization process from a numeral to an indefinite article. However, Cantonese *jat* has not grammaticalized as much as Mandarin *yi*. In particular, I argued that Cantonese *jat* is at an intermediate stage of grammaticalization: a numeral ‘one’ starts as a specifier of a phrase; then, it adjoins to a head before it is reanalyzed as an independent head projecting a separate phrase. Turkish *bir* ‘one’ provides more evidence for this intermediate stage.

References

- Bošković, Željko, and Hsieh, I-Ta Chris. 2013. On word order, binding relations, and plurality in Chinese Noun Phrases. *Studies in Polish Linguistics* 8: 173-204.
- Chao, Yuen-Ren. 1968. *A grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Chen, Ping. 2003. Indefinite determiner introducing definite referent: a special use of ‘*yi* ‘one’ + classifier’ in Chinese. *Lingua* 113, 1169-1184.
- Huang, C.-T. James. 1987. Existential sentences in Chinese and (in)definiteness. In *The representation of (in)definiteness*, ed. Eric Reuland and Alice ter Meulen, 226–253. Cambridge, Mass.: MIT Press.
- Huang, C.T. James, Li, Y.H. Audrey, and Li, Yafei. 2009. *The Syntax of Chinese*. Cambridge University Press: Cambridge, UK.
- Huang, Hansheng. 1981. *Xiandai Hanyu* [Modern Chinese], Shumu Wenxian Press.
- Huddleston, Rodney and Pullum, Geoffrey K. 2002. *The Cambridge Grammar of the English Language*, Cambridge: Cambridge University Press.
- Lee, H.-T. Thomas. 1986. Studies on quantification in Chinese. Doctoral dissertation, University of California, Los Angeles.
- Li, Charles, and Sandra Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley and Los Angeles: University of California Press.
- Li, Y.-H. Audrey. 1996. Definite and indefinite existential constructions. *Studies in the Linguistic Sciences* 26: 175–191.
- Li, Y.-H. Audrey. 1998. Argument determiner and number phrases. *Linguistic Inquiry* 29: 693-702.

indefinite article projects its own phrase, and for those who disallow (ii), *yi* as an indefinite article is adjoined to N, hence only adjectives that are adjoined to N (i.e., adjectives without *de*) can intervene between *yi* and N for these speakers. The speaker variation in question could be taken as another argument for an intermediate stage where the indefinite article is head-adjoined, which precedes the full reanalysis where the indefinite article heads its own projection (i.e., the speaker variation here could be a result of an ongoing grammaticalization process from a numeral to a true indefinite article).

- Li, Yen-hui Audrey. 2005. Shengluo yu chengfen qieshi [Ellipsis and missing objects]. *Yuyan Kexue* [Linguistic Sciences] 4 (2): 3–19.
- Li, Yen-hui Audrey. 2008. Duanyu jiegou yu yulei biaoji: ‘de’ shi zhongxinyu? [Phrase structures and categorial labeling: de as a head?]. *Dangdai Yuyanxue* [Contemporary Linguistics] 2: 97–108.
- Lü, Shuxiang. 1990 [1944]. Ge-zi de yingyong fanwei, fulun danweici qian yi-zi de tuoluo [the uses of *ge* and omission of *yi* before classifiers]. In *Lü Shuxiang Wenji* [Collected works of Lü Shuxiang] 2, 144–175. Beijing: Shangwu Press.
- Lü, Shuxiang et al. 1999. *Xiandai Hanyu Babai Ci* [800 Words in Chinese]. Beijing: Shangwu Press (1st edition, 1980).
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Tang, C.-C. Jane. 1990. *Chinese phrase structure and the extended X'-theory*. Doctoral dissertation, Cornell University, Ithaca, N.Y.
- Tsai, W.-T. Dylan. 1996. Subject specificity, raising modals and extended mapping hypothesis. Presentation at the Symposium on the Referential Properties of Noun Phrases, CUHK.
- van Gelderen, Elly. 2001. Phrases, Heads, Grammaticalization, and Economy. International Conference on Historical Linguistics paper, appear as van Gelderen 2004.
- van Gelderen, Elly. 2004. *Grammaticalization as Economy*. John Benjamins Publishing Company.
- Wang, Shaoxin. 1989. Liangci *ge* zai Tangdai qianhou de fazhan [the development of the classifier *ge* around the Tang Dynasty]. *Yuyan Jiaoxue yu Yanjiu* 2: 98–119.
- Willis, David. 2007. Specifier-to-head reanalyses in the complementizer domain: evidence from Welsh. *Transactions of the Philological Society* 105 (3): 432–480.
- Wood, Johanna. 2003. Definiteness and Number. Arizona State University Dissertation.
- Yukseker, Hitay. 2000. Bir ‘one’. *Proceedings of the Tenth International Conference in Turkish Linguistics*. Bogazici University, Istanbul. August 16–18, 2000.
- Zhang, Niina Ning. 2013. *Classifier Structures in Mandarin Chinese*. De Gruyter Mouton.