On Moving Constituents in Chinese Historical Syntax*

Alain Peyraube

CRLAO (CNRS and EHESS)

It is generally assumed that the operation of Moving Constituents (MC) is necessary to account for some important syntactic change in Chinese, especially with respect to problems of word order change. The operation of MC, however, is rarely indisputably argued. It remains, in the great majority of cases, a simple hypothesis that only permits derived structures to be obtained from chronologically prior structures.

It is not then surprising that recently several other hypotheses have been put forward, proposing to identify constraints on the order of constituents, instead of calling upon the well-known operation of MC. Three main problems have been particularly discussed, concerning: (i) quantified expressions with Numbers and Measure Words or Classifiers; (ii) pre-verbal prepositional phrases; (iii) the aspectual marker 'liao'.

The paper discusses the different and competing analyses for syntactic changes that have characterized these three structures. It is shown that discarding the operation of MC raises more problems than it is supposed to solve. The conclusion is that the operation of MC should be upheld in accounting for syntactic change and especially for word order change.

1. Introduction

It is generally assumed that both moving constituents (MC) and reanalysis (including grammaticalization) are absolutely necessary to account for the important syntactic changes in Chinese, especially with respect to problems of word order change. The arguments in favor of the operation of

* This is a revised version of a paper published in French (Peyraube 1999). My thanks, for their comments, to Redouane Djamouri and Wu Fuxiang.
moving constituents are, however, rarely indisputably established. MC remains, in the great majority of cases, a simple hypothesis that only permits derived structures to be obtained from chronological prior structures.

It is not then surprising that recently several other hypotheses have been put forward, proposing to identify constraints on the order of constituents, instead of calling upon the well-known operation of MC. Three main problems have been particularly discussed. These are: (i) quantified expressions with numbers and measure words or classifiers; (ii) pre-verbal prepositional phrases; (iii) the aspectual marker liao.

The paper will discuss the different and competing analyses for syntactic changes that have characterized discussion of these three structures. It will be shown that discarding the operation of MC raises more problems than it is supposed to solve. The conclusion will be that the operation of MC should be upheld in accounting for syntactic change and especially for word order change.

2. Measure words and classifiers

In Pre-Archaic (14th-11th centuries BC) and Early Archaic Chinese (11th-6th c. BC), there is only one structure involving Measure Words (MW): Noun (N) + Number (Num) + MW.1 Examples:

(1) chang wu you (Jiaguwen he ji 25979)  
millet-wine five you  
Five you of millet-wine.2

(2) wang yi jin bai lu (Jin wen: Qin ken)  
king offer bronze hundred lu  
The king offered (him) a hundred lu of bronze.

A new structure appears in Late Archaic Chinese (5th-3rd c. BC): Num + MW + N. Examples:

---

1 At that time, there are yet no classifiers. For differences between MWs and CLs, see Peyraube (1998).
2 You is a common noun meaning ‘vase’, but also a MW used to measure liquids.
(3) you yi yi bei shui jiu yi che xin zhi huo ye (Mengzi: Teng Wen gong, xia)
    be-like with one glass water rescue one cart firewood of fire part.
    It is like fighting a fire of a cart-full of firewood with a glass of water.

(4) sha yi niu qu yi dou rou (Han Feizi: wai chu shuo you, shang)
    kill one ox take one dou meat
    (Whenever he) kills one ox, (he only) takes one dou meat.

Peyraube (1998) made the following hypothesis (5):

(5) N + Num + MW > Num + MW + N

He analyzes the Num + MW phrase as a predicate in the examples (1) and (2).³ He then argues that the change in (5) has been brought into the language through a reanalysis of the Num + MW predicate into a modifier of the following head-noun, as in (3) and (4). As Chinese has always been a Modifier + Head language for the internal structure of the NPs, the Num + MW phrase, after being reanalyzed as a modifier, had to be moved into a pre-nominal position.

Actually, as noticed by Peyraube (1998) himself, it is not possible to interpret all Num + MW post-nominal phrases as predicates and all Num + MW pre-nominal phrases as modifiers. When the Num + MW phrase, for instance, is inside a prepositional phrase (PP), it cannot be analyzed as a predicate, as in the following example:

(6) buruo yi che bai sheng song zhi (Han Feizi: shuo lin, shang)
    best with chariot hundred sheng send him
    The best is to send him (back) with a hundred of charriots.⁴

Yang (1993) challenged the MC hypothesis (5) and proposed instead the following derivations (7a) and (7b):

³ The assumption that the Num + MW is a predicate when it follows a noun is shared by many scholars, especially Ota (1958).
⁴ Such examples are indeed rare. Actually, I only found that one.
(7)  a.   Num + N > Nu + MW + N  
b.   N + Num > N + Nu + MW

In other words, Yang (1993) assumes that the two structures N + Num + MW and Num + MW + N are not related to one another and not derived from each other, but are issued from two different structures already attested in the language, by adding a MW. Her main argument is that the N + Num + MW and Num + MW + N structures display different discourse-pragmatic functions.

The first one (with the MW being post-nominal) expresses a strict quantification and the Num + MW phrase usually conveys new information. It then often appears in enumerative lists of objects, where the Num is important. In the second one (with the MW being pre-nominal), the Num is unimportant and does not express a concrete quantification. This number is always yi ‘one’, and, in most cases, it is used metaphorically as something like the English indefinite article ‘a’.

This analysis is indeed tempting but generates many problems, which render it unlikely.

At the first place, the distinctions made by Yang between the two orders Num + MW + N and N + Num + MW are only applicable, of course, when the orders co-exist. They are not applicable when we only have the Num + MW + N sequence. Is it reasonable to think that the language, at that time, could not express any non-concrete quantification? Probably not.

Actually, if the scenario proposed by Yang is the right one, one cannot understand why the two orders (with a pre-nominal MW and with a post-nominal MW) did not appear at the same time, and why it took so many centuries for Num + MW + N to be derived from Num + N, as it is supposed by Yang. The Num + N form already existed since the Pre-Archaic period, and was even at that time the predominant structure for the quantified NPs.5

One does not understand either why the real classifiers (CLs) did not appear at the same time in both pre-nominal and post-nominal positions, as

5 Djamouri (1988:35) gives the following statistics for the language of the oracle bone inscriptions: 875 occurrences of “Num + N” and 324 occurrences of “N + Num” out of 1209 quantified NPs.
the two orders were already common for the MWs, to express that so-called distinction between concrete quantification and neutral quantification.

Indeed the same time gap is observable some centuries later for the CLs: (i) first, appearance of the N + Num + CL sequence under the Former Han (ca. 2nd c. BC), and (ii) appearance, later, of the Num + CL + N in the Early Medieval period (ca. 6th c. AD).

Is it understandable that the quantified expressions with CLs did not have the need to express a neutral quantification while the quantified expressions involving MWs did have that real need? Assuredly not.

The hypothesis of MC (moving constituent) made above in (5) has to be kept as it does not generate so many problems. Nonetheless, it is not impossible that the coexistence, during many centuries, of the two orders (with pre-nominal and post-nominal MWs and CLs) lead to the pragmatic-discursive differences proposed by Yang (1993). These differences would have been progressively eliminated at the same time when the Num + MW (or CL) + N order became entirely predominant.

3. Locative pre-verbal PPs

In Archaic Chinese and during the Former Han (206 BC-23 AD) period, the locative PPs introduced by the preposition yu are almost exclusively post-verbal. Examples:

(8) xue yu zhong guo (Mengzi: Teng Wen gong, shang)
learn at central state
(He) learned (it) in the Central States.

(9) yi qi min yu he dong (Mengzi: Liang Hui wang, shang)
transfer their people to river east
(They) transfer their people to the eastern side of the river.

There are some occurrences of pre-verbal yu-PPs, but they are few: only five cases in the Zuo zhuan (5th c. BC) out of 1833 yu-PPs, for instance, according to He (1985). Those pre-verbal yu-PPs are then generally interpreted as marked constituents. Example:
(10) jun yu Zhao wei gongzi (Shi ji p.2444)
you at Zhao become noble prince
It is in (the kingdom of) Zhao that you became a noble prince.

Starting in the Late Han (25-220) period, there are suddenly many occurrences of pre-verbal yu-PPs. In the Late Han Buddhist texts, locative pre-verbal yu-PPs are almost as many as the post-verbal ones. Examples:

(11) yu kong zhong zuo yinyue (Taisho 8-224 p.477)
at air in make music
(They) made music in the air.

From then on, the pre-verbal yu-PPs are no longer marked constituents. Most of them are already used in unmarked, neutral sentences. That situation is reinforced for the later periods of Early Medieval (3rd-6th c.) and Late Medieval (7th-13th c.). In different representative texts of these periods, we find the following proportions: 65 pre-verbal yu-PPs and 105 post-verbal in the Shi shuo xin yu (5th c.); 92 pre-verbal and 112 post-verbal in Dunhuang bianwen ji (9th and 10th c.); 123 pre-verbal and 106 post-verbal in Zu tang ji (preface dated 952). Examples:

(12) yu Ben mu qian yanyin (Shi shuo xin yu: fang zheng)
at Ben mother in-front-of feast
(They) feast in front of Ben’s mother.

(13) dizhi zai xibian cun nei juzhu (Dunhuang bianwen ji p.168)
disciple only at west village in reside
(If) I only reside at the western side of the village …

To account for such a situation, Peyraube (1994) made the following hypothesis of MC:

(14) V + [yu + NP] > [yu + NP] + V

The moving could have started under the Late Han (25-220) period,
and only for the *yu*-PPs expressing a general location.\(^6\)

Djamouri and Paul (1997) have challenged this Movement hypothesis. They first stated that the unmarked *yu*-PPs were not only post-verbal in Archaic Chinese, but also pre-verbal, and since the Pre-Archaic period, as shown by the following example taken from the language of oracle bone inscriptions:

(15)   wang yu Gong shi (*Jiaguwen he ji* 7352 recto)
       king  at Gong camp
       The king (will) camp at Gong.

Djamouri and Paul counted 451 occurrences of pre-verbal *yu*-PPs and 2717 post-verbal ones in the language of oracle bone inscriptions, i.e. a proportion of 14.2% vs. 85.8%. They also counted for the next period of Early Archaic 33 pre-verbal *yu*-PPs and 860 post-verbal ones in the *Shu jing* and in the *Shi jing*. Examples:

(16)   ren wu yu shui jian, dang yu min jian (*Shu jing*: jiu gao)
       man negation at water self-reflect have-to at people self-reflect
       The man does not (have to) search his own reflection in water,
       (he) must search it in (his) people.

Djamouri and Paul conclude that the two word orders were needed, all the more that they also consider the morpheme *zai* to be a preposition and not only exclusively a verb.\(^7\)

These two orders were needed because of the existence of a constraint that regulates the word order. The constraint stipulates that the PPs which are arguments of the verb (i.e. sub-categorized [+location] or [+destination]) are post-verbal. They may appear sometimes, in a few cases, in pre-verbal position but they are then marked constituents. PPs that are not arguments of the verb appear in both pre-verbal and post-verbal positions.

---

6 The PPs expressing a resultative location could have remained post-verbal. For a distinction between general location and resultative location, see Peyraube (1994).

7 For a different point of view, stating that *zai* has been only a verb in Archaic Chinese, see Wei (1993), Peyraube (1994).
In any case, as the two orders (pre-verbal and post-verbal) were both attested since the Pre-Archaic period, the Movement hypothesis (in (14)) for locative *yu*-PPs does not stand.

Djamouri and Paul’s analysis is indeed stimulating. However it does not answer all fundamental questions and does not really account for the prevailing situation during the archaic period.

If we put aside the debatable status of *zai*, taking only into consideration the preposition *yu*, the statistics given by Djamouri and Paul for the Early Archaic period do not strengthen their analysis. As there are only 3.69% of locative pre-verbal *yu*-PPs in *Shu jing* and *Shi jing* (33 occurrences vs. 860 occurrences of post-verbal *yu*-PPs), one must conclude that the constraint—if there is any constraint—is not operational. Such a situation would simply mean that the post-verbal position was used for both non-arguments and arguments of the verb. On the other hand, these statistics would be in accordance with a hypothesis stating that the locative pre-verbal *yu*-PPs, whose occurrences were rare, were marked constituents.

Only for the Pre-Archaic period the situation looks quite different. As the pre-verbal *yu*-PPs represent more than 14% of all *yu*-PPs, it could be unreasonable to admit that all those pre-verbal *yu* are marked constituents.

In Late Archaic Chinese, we have approximately the same percentages as in Early Archaic Chinese. No indications are given by Djamouri and Paul for that period, but see above the figures taken from He (1985). They show that if all argument PPs are post-verbal, the non-argument PPs are also post-verbal, for the majority of them.

Thus, it will be quite difficult to reject the MC hypothesis given in (14). Only the Moving hypothesis is able to account for the sudden increase of pre-verbal *yu*-PPs under the Late Han period, where they already represent almost half of all the occurrences of locative PPs introduced by *yu*.

We must indeed suppose that non-argument PPs—that were post-verbal in Archaic Chinese in their great majority—have been moved in a pre-verbal position. The alternative solution put forward by Djamouri and Paul (1997) would lead us to admit that the Chinese language radically changed during the Pre-Medieval period becoming a language where the locative PPs were suddenly used in a non-argument position, that was not the case before. Such big mutations in languages are improbable.
4. The aspectual marker liao

The classical thesis on the origin of the accomplished aspectual marker liao (le in Contemporary Chinese) has been put forward by Mei Tsu-lin (1981, 1994).

There existed in Early Medieval Chinese (ca. 3rd-6th c. AD) a serial-verb construction Verb1 + Object + V2 (V1 + O + V2) where V2 was a verb meaning ‘to finish, to accomplish’. Several synonym verbs could fill that V2 position: bi, qi, yi, jing, and liao around the end of the period. Examples:

(17) wang yin jiu bi… (Shi shuo xin yu: fang zheng)
    king drink wine finish
    Having finished to drink the wine, the king…

(18) ju qi shi qi… (Taisho 2-125, p.665)
    totally beg food finish
    Having finished to beg for food…

(19) fo shuo ci jing yi… (Taisho 9-262, p.2)
    Buddha recite this sutra finish
    Having finished to recite the sutra, Buddha…

(20) (Zhang Jiying) zuo shu qu jing… (Shi shuo xin yu: shang zhe)
    (Zhang Jiying) make several tune finish
    Having finished to play several tunes, Zhang Jiying…

Starting at the beginning of the Late Medieval period (7th-13th c.), liao became predominant and began to replace all the other verbs in filling the V2 position, as in:

(21) zuo ci yu liao, sui ji nan xing (Dunhuang bianwen ji p.8)
    Do this word finish then immediately south move
    Having said these words, (he) immediately moved to the south.

The V2 liao has been first reanalyzed into a phase complement (PC):
V1 + O + V2-liao > V + O + PC-liao.\(^8\) Then, by analogy with other PC like zhuo, que, de, etc. that were used quite commonly at that time in the position between the verb and its object, liao has been moved in the pre-object position (see also Cao 1986):

\[(22)\]  
V + O + PC-liao > V + PC-liao + O

The PC liao has been then grammaticalized again into a simple aspectual marker.

This classical analysis has been recently discussed and challenged by Wu (1997) and by Li and Shi (1997). They both have proposed the following concurrent analysis without any MC, such as the one postulated in (22):

\[(23)\]  
V + AM (que or liao) > V + AM (que or liao) + O  
[V = Verb, AM = Aspectual Marker, O = Object]

In other words, que and liao would have been grammaticalized, first from verbs into PC, and then from PC into AM, when they were directly following the verbs, and when they were used without any object following them. Then, and only then, after the last stage of the grammaticalization process, it would have been possible to add an object. In such a scenario there was not any Movement of constituents.

The main arguments of Li and Shi are as follows:

(i) In the Tang and the Five Dynasties periods (7th-10th c.), 70% of the occurrences of V + O + liao are actually with an adverbial preceding liao. The canonic form is then V + O + Adv. + liao, and the Adv. + V + O + liao form is never found. This shows that liao has to be still considered as a verb and that it has not yet been grammaticalized.

(ii) When there is no object, Li and Shi counted only 2 occurrences of V + Adv. + liao out of 43 occurrences of V + liao without any object.

\(^8\) See Chao (1968:446) for the notion of Phase complement.
Li and Shi conclude that the grammaticalization process of *liao* was on the verge of being completed, but only for the V + *liao* sequence, without object. The origin of the Contemporary Chinese V + *le* + O (*le = liao* is an AM) is then traced back out of the V + *liao* structure where *liao* was already an AM, by addition of an object, and not out of the V + O + *liao* structure, where *liao* was still a verb.

Wu (1997) develops another argument in favor of the same hypothesis of a derivation by addition of a constituent instead of the one consisting of moving a constituent. He first notices that the V + *guo* + O structures involving the experiential aspectual marker *guo* have undoubtedly their origin in the V + *guo* structures (without any object) by adding an object, and not in the V + O + *guo* structures (with an object) by moving a constituent, for the good reason that the V + O + *guo* structure has never been attested in the language. Wu then concludes that the history of *liao* and *guo* should have been the same.

These arguments against the hypothesis of MC for explaining the origin of the AM *liao* also raise several problems.

There are no reasons, first, to think a priori that the origin and the development of the AM *liao* should have been the same as the AM *guo*. The absence of a V + O + *guo* form, from which the V + *guo* + O could have been derived does not justify by itself the following scenario: V + *guo* > V + guo + O. On can well propose a solution which would consist of considering that *guo* has been grammaticalized into an AM in two different structures, namely V1 + V2-*guo* and V1 + V2-*guo* + O.\(^9\)

\[
(24) \begin{align*}
\text{a. } & V1 + V2-\text{guo} > V + \text{AM-guo} \\
\text{b. } & V1 + V2-\text{guo} + O > V + \text{AM-guo} + O
\end{align*}
\]

As for the figures provided by Li and Shi (1997), if they prove that the V + O + Adv. + *liao* structure was still predominant in Late Medieval Chinese, one cannot conclude that *liao* was still certainly a verb. It is not excluded that *liao* could also have been then already a PC. Moreover, nothing indicates that *liao* is already an AM when there is no object in the

\(^9\) On the history of *guo*, see Liu (1999).
V + liao form.

As mentioned above for guo, a hypothesis of a grammaticalization process of liao into an AM in two different structures V + PC-liao and V + PC-liao + O is not invalidated by the data taken from Li and Shi.

Finally, concerning the origin of V + PC-liao + O, there is no way to make a choice between (25) (derivation by adding a constituent) and (26) (derivation by moving a constituent), as both V + PC-liao and V + O + PC-liao structures were attested before the development of the V + PC-liao + O:

(25) V + PC-liao > V + PC-liao + O
(26) V + O + PC-liao > V + PC-liao + O

5. Conclusion

The new hypotheses challenging the classical analyses of Moving constituents are generating several problems certainly worthy of further discussions. Is it then reasonable to reject the operations of MC for the three cases discussed above? Probably not.

There are indeed many other domains of the syntactic structure where MC is needed in diachrony. Just to mention a few, in the following four cases, there are no alternative solutions that would discard the operation of MC.

The first case deals with the development of dative structures (cf. Peyraube 1988:185):

(27) a. V1 + V2-yu + IO + DO > V1 + DO + V2-yu + IO, in Early Medieval (ca. 4th c.);

b. V + DO + Preposition-yu + IO > Prep.-yu + IO + V + DO, in Late Medieval (ca. 9th-10th c.)

[DO = direct object, IO = indirect object, the verb yu = to give, the preposition yu = to]

The second case, quite similar to the problems of liao discussed above, involves constructions with resultative complements (cf. Jiang 1994:195):
(28) \( V + O + bu\ de > V + bu\ de + O \), at the end of the Early Medieval period.

The third case is the evolution of the comparative forms of equality (cf. Peyraube 1989):

(29) \( X + \text{Adjective} + \text{CM} + Y > X + \text{CM} + Y + \text{Adjective} \), in Early Medieval (3rd-6th c.)

[\( X \) and \( Y \) are the two terms of the comparison, and \( \text{CM} \) the comparative morphemes \( ru \) or \( si \)]

The fourth and last case, finally, is well known since several decades. The interrogative pronouns objects and the personal pronouns objects (but in negative sentences), generally pre-verbal in Late Archaic Chinese (5th-3rd c. BC)—the Classical Chinese par excellence—have been moved into a post-verbal position beginning in the Han period, around the 2nd c. BC:

(30) a. Interr. Pronoun-object + V > V + Interr. Pronoun Object
References


漢語歷史語法中的「詞（組）移位」

貝羅貝
法國高等社會科學院

以「詞（組）移位」來解釋漢語詞序改變等重要的句法變化，相當普遍。但是，卻很少有人認真地論證「詞（組）移位」的運作機制。在大多數的情形下，「詞（組）移位」只是一個可以讓某個結構從另一個結構衍生出來的簡單假說。

因此，最近有幾個分析提出「詞序限制」來取代著名的「詞（組）移位」，就不令人意外了。其中有三個問題被特別提出來討論，即 (i) 帶有數量詞或類別詞的量化詞組；(ii) 動前介詞組；(iii) 動貌標誌「了」。

本文討論與這幾種結構句法變化有關的幾個彼此競爭的分析。結果顯示，丟棄「詞（組）移位」的運作機制會帶來更多的問題。結論是，在解釋句法變化，特別是詞序的變化時，「詞（組）移位」的分析較好。