Typology in China
Hilary Chappell, Ming Li, Alain Peyraube

To cite this version:
Hilary Chappell, Ming Li, Alain Peyraube. Typology in China: THE STATE OF THE ART. 2007. halshs-00149611

HAL Id: halshs-00149611
https://halshs.archives-ouvertes.fr/halshs-00149611
Preprint submitted on 28 May 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Typology in China: the State of the Art
(Pre-publication version)

Hilary Chappell, LI Ming and Alain Peyraube*

1. INTRODUCTION
China possesses rich linguistic resources which remain relatively untapped: the 10 main Sinitic languages or dialect groups account for roughly 93% of the population (Mandarin, Jin, Xiang, Gan, Hui, Wu, Min, Kejia, Yue and Pinghua); the remaining 7% comprise the many different ‘minority’ languages in long term contact with Sinitic such as Tibeto-Burman, Mongolian, Hmong and Tai. In an almost unprecedented state of affairs, written records for Chinese extend without a break 3,000 years into the past, furnishing a rich documentation for any kind of historical study.

These factors essentially create an ideal situation for carrying out typology from both synchronic and diachronic viewpoints. Nonetheless, the long-standing tradition in research on language and its relationship to literature in China has meant that very little attention has been directed towards other language families, let alone the dialects of Chinese. This emphasis on ancient versus modern studies, and standard Mandarin versus the other dialect groups in the Sinitic taxon, has led to the situation where the search for linguistic universals on the basis of crosslinguistic work has seen very little development. Even in the many dialect descriptions available, the largest parts of such grammars are devoted to phonology.1

Despite this, during the 1980s, interest in the work of Greenberg and linguistic typology was aroused in linguistic circles in China.2 It is edifying to briefly digress and compare the ways in which the two domains of diachronic linguistics and typology were differently linked in China as opposed to the west – that is, specifically Europe and the USA. While in the west, the new élan in typology set off a revival in the study of diachronic syntax, seen particularly in the renaissance of studies into grammaticalization, the opposite trend took place in China. After the Cultural Revolution (1966-1976) in the late 1970s, scholars began to work in earnest again on historical syntax, notably Liu Jian and Jiang Lansheng, at the Academy of Social Sciences in Beijing. It was thus only in the 1980s that typology began to take off, on the basis of the new research into historical syntax, when linguists such as Zhu Dexi took up the challenge in a decisive way to use dialect material in order to make typological comparisons.

The most important work by Zhu Dexi were his articles of (1980) and (1985): the first examines the use of the highly polysemous subordinating particle de 的 in Beijing Mandarin, and its counterparts in Cantonese (Yue), the Wenshui dialect (Jin),

* The authors, listed in alphabetical order, are affiliated to: (1) Centre de Recherches Linguistiques sur l’Asie Orientale, Ecole des Hautes Etudes en Sciences Sociales (EHESS), France (Chappell); (2) Institute of Linguistics, Chinese Academy of Social Sciences, Beijing (Li); (3) Centre de Recherches Linguistiques sur l’Asie Orientale, EHESS, Paris and Centre National de la Recherche Scientifique, France (Peyraube).
and the Fuzhou dialect (Min), while the second treats the use of the two main types of Yes/No interrogative structures in Sinitic languages with either the form VP-NEG-VP or Adv-VP. This inspired further exceptional scholarship in the same domain by Zhang Min (1990), a comprehensive work on interrogative structures in Chinese languages, and Anne Yue-Hashimoto (1991a) on stratification of the different interrogative structures found in the Min dialects.

In this discussion, we provide an overview of the theoretical developments in Chinese linguistics for typology, mainly for the last 10 years. The subjects we treat are those of word order, topic-prominence and syntactic topics, word categories, verb complementation and verb frames, and the relation of diachrony to typology in terms of grammaticalization studies. We conclude with a brief discussion of future challenges and a description of current projects on typology in Chinese linguistics.

2. **Word Order**

2.1. **Word order change**

The majority of Sinitic languages present a perplexing case for syntactic typology since they display in general head-final characteristics for their NP structure but a mixture of head-initial and head-final ordering for their VPs (though see §2 below on the dominant word order in Wu and Min languages). Thus, standard Mandarin and other Sinitic languages pose somewhat of a challenge: they do not conform to either of the two main alignments, as Dryer (2003), among others, has observed. Table 1 illustrates the incompatibilities in the case of standard Mandarin.

<table>
<thead>
<tr>
<th>Head-Final Structures</th>
<th>Head-Initial Structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consonant with SOV order</td>
<td>Consonant with SVO order</td>
</tr>
<tr>
<td>Adjective – Noun</td>
<td>Verb – Object</td>
</tr>
<tr>
<td>Numeral – Classifier – Noun</td>
<td>Auxiliary – Verb</td>
</tr>
<tr>
<td>Demonstrative – Classifier – Noun</td>
<td>Verb – Modifying adverbal complements of manner, result and degree</td>
</tr>
<tr>
<td>Relative Clause – Noun</td>
<td>Preposition – NP</td>
</tr>
<tr>
<td>Genitive – Noun</td>
<td>Complementizer – S</td>
</tr>
<tr>
<td>Adverb – Verb</td>
<td></td>
</tr>
<tr>
<td>Intensifier – Adjective</td>
<td></td>
</tr>
<tr>
<td>Standard of comparison – Adjective</td>
<td></td>
</tr>
<tr>
<td>Prepositional Phrase – Verb</td>
<td>Exception: A subset of locative constructions take the form Verb – Prepositional Phrase</td>
</tr>
</tbody>
</table>

This problem has also contributed to the controversial debate on the basic order of constituents in Chinese languages which began in the 1970s. Although the main issue has been to determine whether Modern Chinese has basic SOV or basic SVO word order, this was necessarily linked up with the possibility of word order change.
from SVO to SOV. Note that in this context ‘Chinese’ means ‘Mandarin Chinese’. Certain scholars upheld such a diachronic change (Li and Thompson 1975, 1976; Huang 1978; Tai 1973, 1976) while others made the counterclaim that Mandarin, if not Chinese as a whole, has always been, and remains, SVO (Light 1979, Mei 1980, Sun et Givon 1985, Wang 2005 *inter alia*). Furthermore, researchers set out to identify these word order preferences on the basis of the documented history of the language and the use of quantified data for the syntactic constructions in question.

The debate slowly took on larger dimensions, extending to the basic word order of proto-Sinitic and even to that of proto-Sino-Tibetan and its relation to proto-Sinitic (Dryer 2003). First, we take a brief look at studies from early historical periods of Chinese, then at the word order debate for contemporary Chinese.

The controversial issue for reconstruction regards prehistory and the possibility that Sinitic and Tibeto-Burman languages have evolved from an ancestral SOV language. The debate in Chinese linguistics began with the status of Archaic Chinese (11th – 3rd century BCE), shown to be a non-rigid SVO language, given the conditioned appearance of pronominal DOs in preverbal position. This in turn led to the enticing question of whether the SOV pronominal order could be the residue of an even more ancient SOV word order.

Let us first consider the rules allowing SOV order in Archaic Chinese: these are much clearer than in later periods and include (i) interrogative pronouns as DOs preceding the verb in questions (with few exceptions), (ii) a tendency for personal pronouns as DOs to precede the verb in negated sentences (however, there are many exceptions to this rule), (iii) the demonstrative pronoun *shi4* 是, as DO, being preposed before the verb. Similarly, DO nouns taking the focus markers *wei2* 唯- and/or *shi4* -是 are preposed. This is summarized in the Diagram 1 below.

\[
\text{SUBJECT} \quad \text{OBJECT} = \begin{cases} 
\text{interrogative pronoun} \\
\text{personal pronoun} \quad [\negated S] \\
\text{demonstrative pronoun} \\
\text{*wei2* 唯-\text{*shi4* -是 focus noun} 
\end{cases} \quad \text{VERB}
\]

Diagram 1: SOV order in Archaic Chinese

Yu Min 俞敏 (1981), Liu Danqing 刘丹青 (2004) and other scholars believe that this atypical SOV order in Archaic Chinese could be a vestige of the actual word order in proto-Sino-Tibetan. Matisoff (2003), LaPolla (1994) and many other Tibeto-Burman scholars cited in these two works, are all of the view that the common ancestor of Sino-Tibetan must be SOV, given that contemporary Tibeto-Burman
languages remain SOV, apart from Karenic and some of the Bai languages which are SVO. More recently, Dryer (2003) takes up this standpoint with some interesting and subtle additions of argumentation. He maintains too that proto-Sino-Tibetan was OV while proto-Chinese was probably VO. In his approach however, the change from OV to VO has not yet been fully realised since Chinese preserves its head-final characteristics in the noun phrase, a feature regularly correlated with OV ordering.

His explanation for the synchronic situation mainly refers to Mandarin Chinese and is as follows: Mandarin has retained its head-final characteristics from the proto-language, particularly nominal modifiers including ReN and GenN, under the pressure of Altaic languages to the North with which it forms a linguistic area. The correlation between OV and *modifier-modified* orderings appears to be an Asian areal feature, characteristic of the language families found in North Asia but also and interestingly, the Tibeto-Burman languages of the Western Tibetic group in contact with Indo-Iranian. In the case of ReN, this correlation does not hold for all the OV languages in the world, which can be associated with either ordering in fact (Dryer 1992). His explanation would thus appear to support the much-debated Altaicization hypothesis put forward in one of the earliest typological approaches of the Greenberg era concerning the Sinitic taxon, that of Hashimoto (1976, 1986), a hypothesis which still remains to be proven (see also §6.4. below). The important contribution of Dryer is to link up the debate on phylogeny, and diachronic change in word order with areal typology and correlations concerning linguistic universals.

Nonetheless, several aspects still need to be resolved with Dryer’s hypothesis, for example, regarding the existence, and extent of intensive contact between speakers of proto-Sinitic and proto-Altaic in prehistoric times, a precondition for Sinitic to have retained its SOV features. It is well-known that different kinds of language contact situations came into being with the various Altaic-speaking dynasties which ruled China from the North including the Liao or Khitan (907-1125) (Mongolian), the Jin (1115-1234) (Jurchen, Manchurian), the Yuan (1279-1368) (Mongolian) and the Qing (1644-1911) (Manchurian). However, these evidently concern a much later period coinciding with Medieval and Modern Chinese.

These arguments notwithstanding, Djamouri (2001: 146-147) shows that in the oldest extant records of pre-Archaic Chinese, the oracle bone inscriptions of the Shang dynasty (14th – 11th centuries BCE), the dominant word order found for nominal objects is SVO as counted in a corpus of 26,094 complete sentences: 93.8% are SVO and 6.2% are SOV. In fact, the SOV tendency for pronominal objects appears to be slightly *higher* in the following Early Archaic period (11th – 8th BCE), with a greater variety of types possible in this position (Peyraube 1997a, 1997b). This evidence makes it much more difficult to suppose a word order change from SOV to SVO in the case of Sinitic.

To sum up, Peyraube stresses that the SOV order with pronominal objects remains a highly marginal one, and so he concludes, it should not be used as a basis for reconstruction, given the possibility of special positioning and focus rules for pronouns (Comrie 1981: 83). Hence, both Pre-Archaic and Archaic Chinese can be definitively shown to be SVO languages in terms of their dominant word orders. To suppose a prehistoric stage even more ancient than the oracle bone inscriptions of the Pre-Archaic period (14th – 11th BCE), namely, proto-Sinitic with SOV order, is not
therefore empirically grounded, and must remain a pure surmise.

2.2. Word order correlations

Another very important area of research on word order typology in China is the cognitive approach pioneered by Lu Bingfu 陆丙甫. As the foundation of his research, Lu uses the two most important principles proposed by Greenberg (1963) to explain implicit universals - those of dominance and harmony. The explanation and application of the principle of harmony has received wide attention, while research on dominance remains relatively neglected. In order to give an explanation for dominant word order, Lu Bingfu (2005) proposed the following pragmatic factors: other conditions remaining the same, the higher the degree of identifiability of a constituent, the more likely it is to be placed at the head. (The definition of identifiability is based on Lambrecht (1994) and closely resembles the notion of referentiality, but with a more general sense than the latter.)

For example, Universal 1 of Greenberg (1963) (In declarative sentences with nominal subject and object, the dominant order is always the one in which the subject precedes the object.); Universal 5 (If a language has dominant SOV order and the genitive follows the governing noun, then the adjective likewise follows the noun.) 6; Universal 25 (If pronominal object follows the verb, so does the nominal object.) etc can all be explained in the following way: the identifiability of subjects is higher than that for objects, the identifiability of nouns is higher than that for adjectives, and the identifiability of pronouns is higher than that for nouns.

Lu Bingfu subsequently proceeds to apply the principle of 'identifiability attracts elements to the front' to noun phrases. Universals 18 and 20 of Greenberg can thus be explained in terms of the (covert) identifiability of demonstratives and numerals being higher than that of adjectives. (No. 18: When the descriptive adjective precedes the noun, the demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise; No. 20: When all or any of the items (demonstrative, numeral; and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite.) Lu proposes that there is a pragmatic principle interacting with a semantic principle 'the closer the meaning of an element is to that of the head noun, the closer it is placed to the head noun' which explains the fact that when demonstratives, numerals, and descriptive adjectives precede the noun, they always appear in the order of DEM-NUM-ADJ whereas it is impossible to predict any order for their positioning after the noun (Hawkins 1983).

The principle would also explain the following phenomena that the ordering within complex attributive phrases is stable when they precede the noun but free when they follow: preposed, the two principles harmonize, reinforcing one another, with the word order showing crosslinguistic stability; postposed the two principles counteract each other; hence no word order preferences emerge (see Lu Bingfu 1998, 2005).

Recently, Lu Bingfu combined these issues affecting word order to further investigations based on typology. In a crosslinguistic study, Lu (2004) exposes an
important universal, that of the rule of correspondence between distance and marking: the marker may be omitted for constituents whose word order and structural level are close to the head, otherwise it is obligatory. We note that this independent work is reminiscent of Haiman’s iconicity principle (1980). For example, adverbial phrases composed of nouns in standard Mandarin must be placed as close as possible to the verb, or else some kind of marker needs to be added to convert them into this function:

1. We now: more more telephone contact
   ‘From now on let’s keep in contact more often by phone.’

2. We now: *use telephone more-more contact
   ‘From now on, let’s use the phone more often to keep in contact.’

In sentence (1), the noun ‘telephone’ directly precedes the verb ‘to contact’ whereas in sentence (2), it has to be placed into an instrumental phrase with yòng ‘use’, once it has been moved away from the verb it modifies.

3. **Topic Prominence and Syntactic Topics**

Chao (1968: 69) was one of the first to recognize that subject in Chinese languages was different from that found in many European languages: given that the semantic relation between the subject and predicate is typically quite loose, he called it ‘topic and comment’. Inspired by this, Li & Thompson (1976) proposed what became a celebrated four-way typological classification on the basis of the relative importance of subject and topic: languages could be either subject-prominent or topic-prominent, equally topic- and subject- prominent or neither. In other words, topic was not just a discourse notion but could be identified syntactically as such in topic-prominent languages.

As originally discussed by Chao (1968) and Li and Thompson (1976), the construction in question in Mandarin has the form NP<sub>topic</sub>–S. The topic, ‘that dog’ in (3) below, does not enter into any grammatical relation with the main verb of the S found in the comment but is commonly the inalienable possessor of the subject in the comment ‘body’. The syntactic looseness between topic and comment is reflected in the presence of an intonation break or discourse particles in this construction:

3. That dog even body possess unique skill
   ‘As for that dog, it was very talented.’
   (Transcription of conversation with Liu Bing 29: 420; see §6.3 below)
The notion of topic-prominence has been more recently extended to further construction types in Chinese linguistics.

3.1. Syntactic topic

Taking up this earlier research by Li and Thompson, Xu Liejiong 徐烈炯 and Liu Danqing 刘丹青 (1998) proposed that topic is a fundamental syntactic constituent in Chinese languages. They show that topic is a highly grammaticalized syntactic constituent, particularly in Shanghainese and other Wu dialects. In reviewing the typological parameters for discourse configurational languages, Xu Liejiong (2002) shows that Chinese belongs to a subtype of *topic configurational languages* where topic is given a special position in the syntactic structure.

Liu Danqing (2001) also investigates another kind of structure he labels the *split-argument topic structure*. With respect to sentences such as (4), where one of the arguments is split into two components, the preference for the expression of topic in Mandarin is satisfied in terms of his theoretical framework:

(4) 衬衫 他 买了 三件。
*Chénshān tā mǎi-le sānjiàn.*
shirt 3SG buy-PFV three CL.
‘Shirts, he bought three of them.’

Liu argues as follows: The split-argument construction type is more developed in certain of the southern Sinitic languages whose topic-prominence is correspondingly more pronounced. In Wu and Min dialect groups, for example, the status of split topic has reached a higher degree of entrenchment and syntacticization, since the position of subtopic, after the subject, is more common than the main topic position. Liu (2004) has also analysed another kind of topic structure which is not found in subject-prominent languages — the *identical topic structure*:

(5) 他 主任 也 当过 主任。
*Tā zhǔrén yě dāng-guo zhǔrén.*
3SG director also act-EXP head
‘S/he used to be the head.’

According to Liu (2004: 24), these kinds of structures are commonly found in both periods of Archaic and Medieval Chinese; in Shanghainese and also in Tibeto-Burman. It is clear that they seriously violate the principle of economy, in allowing the same constituent to occur twice in the same small clause and in not adding any extra meaning.

However, this structure similarly fulfils the topic requirement, particularly for Wu and Min dialects where such structures are frequent. In Liu’s view, the existence of these kinds of constructions underlines the syntactic nature of topic in Chinese
languages. His conclusion is that if we only accept topic as a discourse notion, the
differences between topic-prominent languages like Chinese and subject-prominent
languages like English will be overlooked.

3.2. Topic and the relative clause accessibility hierarchy

In researching other syntactic problems, the deep-level effects of
topic-prominence can be discovered. According to the noun phrase accessibility
hierarchy set up by Keenan and Comrie (1977), the syntactic constituents extracted by
relative clauses respect the following order of priority:

subject > direct object > indirect object > oblique > genitive > standard in a
comparative

However, in Mandarin the ordering does not fully respect this hierarchy. Obliques
cannot be extracted, for example, while certain constituents which clearly
semantically belong to a genitive, can, for example:

(6) 父亲死了的孩子
fùqín sǐ-le de háizi
father die-PFV REL child
‘the child whose father died’

Nevertheless, háizi ‘child’ and fùqín ‘father’ can only be considered as being in a
possessive relation from a semantic point of view. In subject-prominent languages,
syntactically speaking, they can often only be coded as an attributive-noun
relationship of the genitive type. However in Mandarin, another possibility for coding
the genitive exists, the topic slot, not available in subject-prominent languages. In the
view of Liu (2005a), the probable origin of (7) lies in the topic-comment construction:

(7) 孩子（, ）父亲死了。
Háizi fùqín sǐ-le.
child father die-PFV
‘As for the child, his father died.’

In this way, it can be seen that relative clauses in Mandarin do not necessarily
violate the accessibility hierarchy. However, topic is so far not ranked in this relative
clause accessibility hierarchy. Liu argues that it should be added in to enable
predictions which are correct for topic-prominent languages.

3.3. Topic-comment and double subject clauses in natural discourse data

Using a natural discourse data, Chappell (1995), one of the authors of the present
article, makes use of structural, semantic and prosodic criteria in an analysis of
spoken Mandarin topic constructions. This study reveals that the prototypical ‘double
subject’ type, NP-S, occurs very rarely in spoken discourse. In 3:03 hours of recorded data including Mandarin narratives and conversations, only 30 topic-comment structures were identified in a total of 2801 intonation units, that is, less than 1% at 0.93%. Examples (3) and (8) are from the corpus she used. These data were transcribed according to the methodology of DuBois et al (1993), wherein an intonation unit roughly corresponds to the clause in written data. (CL = classifier, RP = rhetorical particle, // = boundary of intonation unit)

(8) 那个小男孩子呢，
Nēige xiǎo nán hǎizi ng.//
that CL little male child RP

‘The little boy, well, there was no expression on his face whatsoever.’
(Pear II.7/8:77-78, cited from Chappell 1995: 492)

Since the tendency in the natural data found in this corpus is for two fully lexical NPs to occur in different intonation units (or clauses), separated by an intonation break or a rhetorical particle, as in (8), the notion of syntacticized topic at the level of a single clause would be hard to uphold, for lack of any other clear structural criteria.

4. ON WORD CATEGORIES

Developments in typology have led scholars in China to consider the issue of word classes anew, particularly how to define their features, independently of the Indo-European tradition in linguistics.

4.1. Adjectives

Shen Jiaxuan 沈家煊 (1997) importantly shows that adjectives can be established as a separate word class in Mandarin and, furthermore, can be divided into two subclasses: quality and descriptive adjectives. Liu Danqing (2005) remarks that this division into two types is a characteristic of many languages in the Asian area, and one that should be taken into account in typological studies on adjectives.

The unmarked function of quality adjectives is as attributives, while as predicates they need to take the marker de 的 and/or co-occur with the copular verb shì 是. For example:

(9) 红衣服 这件 衣服 *（是）红 *（的）。
hóng yīfu Zhèjiàn yīfu *(shì) hóng *(de).
red clothes this:CL clothes be red DE

‘red clothes’ ‘These clothes are red.’

In contrast to this, the unmarked function of descriptive adjectives is to serve as the predicate, while as attributives they need to take the marker de 的. For example:
(10) 冰冷 *的* 手 他 的 手 冰冷。
bīnglěng *(de)* shǒu Tā de shǒu bīnglěng.

`ice-cold hands' 'Her hands are ice-cold.'

In other words, their syntactic behaviour is different. The subcategory of the stative adjective could thus be added to the marking patterns that are crosslinguistically possible. Shen's research is based on an earlier breakthrough by Zhu Dexi 朱德熙 (1956) (see §1 above).

4.2. Adpositions

Much to its detriment, the study of adpositions in Chinese languages has generally made use of prepositional languages such as English and other European languages as its point of reference. Consequently, this has created difficulties in accounting for postposed particles that display adpositional functions. Chao Yuen Ren (1968: 397-403) and Liu Dancqing (2003) inter alia have shown that, in addition to prepositions with a verbal origin, postpositions can also be found in Mandarin and other Chinese languages.

Postpositions in Mandarin, for example, include not only locative markers such as *shang* 上 ‘on’, *li* 里 ‘in’, *xià* 下 ‘under’, *zhòng* 中 ‘inside’, *within*, *zhīwài* 之外 ‘apart from’, *yīnèi* 以内 ‘within’, *zhījiān* 之间 ‘between’ but also what are commonly known as particles, such as *qǐ* ‘starting’ in *cóng mingtiān qǐ* 从明天起 ‘starting from tomorrow’, *yǐlái* ‘since’ in *zì sān yuè fēn yǐlái* 自三月份以来 ‘since March’, or *wéizhī* ‘up until’ in *dào xīngqītiān wéizhī* 到星期天为止 ‘up until Sunday’.

In Archaic Chinese, the regular position for prepositional phrases was following the verb, for example, *wǔ yù tīng* 舞于庭 dance-LOC-court ‘dance at the court’. In this period, prepositions were in an intermediary position, conforming to the Dik’s relator principle (1997). However, after the Archaic period, prepositional phrases moved into preverbal position and no longer occupied this position of an intermediary. This violation of the principle could be seen as triggering the development of locative markers and other postpositions so as to fill the new gap created between the prepositional phrase and the verb. The co-occurrence of prepositions and postpositions also led to the creation of circumpositions as found in contemporary Mandarin, *zì ... lì* 在...里 ‘at/in...’, *gēn ... shí de* 跟...的 ‘like...’, *yòng ... lái* 用...来 ‘using/taking...to’; see also Peyraube (2003) for a different view of the diachronic development of locative postpositions and the mechanism for the change in ordering preferences.

5. On VERB COMPLEMENTATION AND VERB FRAMES

An area in which the study of Sinitic languages can afford a rich contribution to typology is that of verb complementation. This involves an array of complex V1V2 structures with V1 as the head, in which the V2 is used to express result states caused
by the event of $V_1$, if not its direction and path in the case of motion verbs. For example:

(11) Resultative complement verb
洗干净 $xī gānjìng$ wash-clean $\Rightarrow$ ‘wash (and make it) clean’

Resultative complement verb with infix for potential mode
洗得干净 $xī děi gānjìng$ wash-ABLE-clean $\Rightarrow$ ‘can wash (it) clean’

Directional complement verb
走过去 $zǒu-guòqu$ walk-over-ANDATIVE $\Rightarrow$ ‘walk over to (away from speaker)’

These structures show great diversity across the dialect spectrum in China, yet are a little known outside of Chinese linguistic circles.

Jiang Lansheng (1995) and Lamarre (2001) are among the first to attempt a typology for a large sample of Chinese dialect groups in the case of $V_1<comp>-V_2$, that is, expanded resultative and directional verb complexes which make use a complementizer, such as 得 in example (11) above, to code result, potentiality, manner or direction. This topic is also treated in important articles by Mei Tsu-Lin (1991) and Wu Fuxiang (2001, 2002) concerning the argumentation for the grammaticalization pathway:

$V_1+V_2 \Rightarrow$

$V_1+$ phase complement $\Rightarrow$ $\begin{cases} V_1+$ aspect suffix (+O) \\
V_1+$ manner complement marker + Complement \\
V_1+$ potential complement marker + Complement
\end{cases}$

For the case of motion verbs as $V_1$ and a directional complement as $V_2$, the application of Talmy (2000) and Slobin’s (2004) framework regarding verb-framed and satellite-framed languages to Mandarin by Peyraube (2006) is timely. In his examination of the historical development of directional complements associated with motion verbs, Peyraube (2006) concludes that for the period known as Archaic Chinese, the language was a verb-framed language. It was during the 5th century in the Northern and Southern dynasties that Chinese started to change from a verb-framed language to a satellite-framed language, a state it fully achieved by the 10th century. It should also be noted that Shen Jiaxuan (2003) similarly argues that the resultative construction in contemporary Mandarin primarily belongs to the satellite-framed type.

6. **RESEARCH INTO DIACHRONIC TYPOLOGY AND CHINESE HISTORICAL LINGUISTICS: MORPHOSYNTAX AND GRAMMATICAL CONSTRUCTIONS**

Diachronic typology is the result of meshing typology with historical linguistics,
in particular, grammaticalization processes, and one of its important goals is the
dynamization of synchronic typology. Its basic ideology is that if we can show there
is evidence to suggest a direct connection between linguistic types, then these might
be linked as historical stages, or linked in terms of grammaticalization processes.

Diachronic typology examines the question of the plausibility and likelihood of
morphosyntactic changes. It also considers which changes are common and which,
rare (see Croft 1990, 2003). For Sinitic languages, some of the more important
findings over the last decade are next presented.

6.1. Comitatives

Grammaticalization of the comitative ‘with’ has been the subject of several
typological studies, including Stolz (2001,) and Lakoff and Johnson (1980), also
receiving treatment in the Heine and Kuteva Lexicon (2002). Crosslinguistically, the
pathways can lead to a coordinative conjunction, a marker of manner, an allative, an
ablative or a marker of instrument and the agent.

With respect to Sinitic languages, Liu and Peyraube (1994), Wu Fuxiang 吴福祥
(2003) and Chappell (2000, 2006) all discuss the grammaticalization pathways of
comitatives, using different perspectives. In Liu and Peyraube (1994), the central
hypothesis is as follows: coordinative conjunctions in contemporary Mandarin (tòng
同, hé 和 and gēn 跟, all meaning ‘and’, ‘with’) do not directly evolve out of verbs
but from prepositions, themselves are derived from verbs. This allows us to set up the
grammaticalization chain:

Verb > Adverb > Preposition > Conjunction (for tòng 同 and gòng 共)
Verb > Preposition > Conjunction (hé 和 and gēn 跟, also for ĵi 及, yǔ 与 in
Classical Chinese)

These have all come to be used as conjunctions meaning ‘and’ during different
historical periods.

Wu (2003) is a typological study which focuses on the synchronic situation for
Sinitic languages. For SVO languages in particular, two common grammaticalization
pathways have been crosslinguistically observed for the comitative: (i) comitative >
coordinative conjunction and (ii) comitative > instrumental > manner. Wu’s
contribution is to show how these grammaticalization pathways are determined by the
position of the comitative marker vis-à-vis the verb. The first pathway is found in
languages where the comitative precedes the verb while the second is found in
languages where the comitative follows.

Chappell (2000, 2006) reveals a completely new pathway of
grammaticalization for comitatives: comitative > accusative. None of the studies
mentioned above describes this unusual pathway found in a subset of Sinitic
languages, where the comitative marker has developed into a full-fledged marker of
the direct object. This development is attested particularly in the Min dialects for
kā–kāng 共 and in Hakka dialects for ʈ’un’gh 同 and lau’gh. The first two comitatives
have their source in verbs of accompaniment, while the second is a verb meaning ‘to
mix'. Typologically, the accusative use of comitative markers represents an interesting
development for Sinitic languages, since generally a verb of taking or holding is
employed in this function (Chappell 2007), as too in many West African Benue-Kwa
languages (Lord 1993).

6.2. Object-marking constructions and the beginnings of areal typology
Making use of major typological features in phonology, lexicon and syntax, Mantaro
Hashimoto (1976, 1986) proposed what has become a celebrated north-south
linguistic division in China, supposedly resulting from many centuries of contact
between Altaic and Sinitic languages in the north, and Tai languages in the south. His
early works signalled the beginnings of areal typology in Chinese linguistics. Some of
these tendencies are listed in Table 2 below for the Chinese languages found in the
two zones:

Table 2 : Tendencies in Sinitic languages according to Hashimoto’s North-South
division

<table>
<thead>
<tr>
<th>ALTAICIZATION (North)</th>
<th>TAICIZATION (South)</th>
</tr>
</thead>
<tbody>
<tr>
<td>stress-based and fewer tones</td>
<td>more tones</td>
</tr>
<tr>
<td>higher proportion of polysyllabic words</td>
<td>higher proportion of monosyllabic words</td>
</tr>
<tr>
<td>smaller inventory of classifiers</td>
<td>larger inventory of classifiers</td>
</tr>
<tr>
<td>preponderance of modifier-modified</td>
<td>more instantiations of modified-modifier</td>
</tr>
<tr>
<td>IO-DO word order for ditransitives</td>
<td>DO-IO word order for ditransitives</td>
</tr>
<tr>
<td>preverbal adverbs</td>
<td>possibility of clause-final adverbs</td>
</tr>
</tbody>
</table>

Adopting the same framework, Anne Yue(-Hashimoto) has similarly contributed
many ground-breaking studies, treating the typological differences among Chinese
languages (1991a, 1991b, 2003), including a handbook for typologists carrying out
fieldwork in China which elucidates these very problems (1993). This basic
North-South division was later refined by Norman (1988) who added a third
transitional zone in central China where types intermediate between the two could be
found.

Building on Norman’s foundation, Chappell (2007) explores the possibility of
establishing four major linguistic areas in China in a study on the diversity of
object-marking constructions, combined with other important features such as aspect
marking and comparative constructions. Object-marking constructions are
non-canonical constructions where the direct object is explicitly morphologically
marked and precedes the main verb, for example: S – Object marker – O – Verb.

In terms of grammaticalization pathways, it is shown that verbs of holding and
taking such as bà 把 and nà 拿, well-described for Mandarin, are far from being the
sole source of these object markers in Chinese languages. Apart from comitatives
mentioned in §6.1., another new lexical source is found in verbs of giving and helping.
According to the dominant patterns of use for grammaticalized object markers (OM),
she concludes by tentatively proposing the following four areas:
1. a Northern area (*OM < hold* and *take* verbs: Jin, Mandarin, Northern Wu)
2. a central transitional area (*OM < give* and *help* verbs: Xiang, Gan, Hui, majority of Wu)
3. a Southeastern area (*OM < comitatives: Min, Hakka, certain Southern Wu*)
4. a Southern area (no grammaticalization; use of serial verb constructions with a *take* verb as $V_1$: Yue)

6.3. Negation
Using Croft (1991) as a basis, Zhang Min 张敏 (2002) is able to account for the grammaticalization of negative existential adverbs into general negative adverbs for the negation systems in Sinitic languages. He strikingly shows that the rate of diachronic change may be different, according to the dialect. For example, the Yulin dialect (Pinghua subgroup) of Guangxi 西玉林 has already completed two cycles in this grammaticalization process. This occurs through repeated fusion of general negative markers with verbs of having which creates a new negative existential marker. General negative markers, being the unmarked form, are used to negate events in the present and future, whereas negative existential markers are used to negate events in past contexts. Zhang also discusses similar developments in the Northern dialects and in Archaic and Medieval Chinese.

6.4. Causatives and passives
A crucial area of typological research for Sinitic languages has been the relation between causatives and passives. In traditional Chinese linguistics, the fact that causative verbs historically evolve into passive markers has often been motivated in terms of the lack of any morphological distinction between agent and patient, a characteristic of grammatical relations in the history of Chinese.

More recently, in a different approach, the sources of these markers have been used as a typological criterion to distinguish northern from southern patterns by both Hashimoto (1988) and Norman (1982). They present the argument that for the Modern period (13th century onwards), many Northern Sinitic languages have based their colloquial passives on causative verbs whose source is in verbs meaning ‘to ask’ or ‘to tell’, even while retaining older passives whose markers evolved out of verbs meaning ‘to suffer’. This is similar to many neighbouring Altaic languages which use the same suffix for both causative and passive functions. In contrast to this, Southern Sinitic languages in general base their passives on verbs meaning ‘to give’, thus neatly aligning themselves with many unrelated languages of Southeast Asia, in the so-called ‘sinospheric’ area (Matisoff 1991).

Nonetheless, for this construction type, the Altaicization hypothesis is very much open to debate, since the grammaticalization of causative verbs into passive markers is attested prior to the Modern period. Jiang Lansheng 江蓝生 (1999) was the first to express doubts about the language contact theory as providing the full explanation.

In two crosslinguistic investigations, Zhang Min (2000, 2003) observed that diachronic change from causative verbs to passive markers is very common.
crosslinguistically, occurring in quite diverse (morpho)syntactic construction types. More specifically, in the case of the give verb class as source, the causative stage represents the pivotal link between such verbs and the passive marker in many unrelated languages. Hence, passive markers that have their source in verbs of giving are not directly derived from these verbs of giving, but evolve out of causative verbs (see also Jiang Shaoyu 2002; Chappell and Peyraube forthcoming). The two sets of syncretisms which arise synchronically group together either (i) a verb of giving with a causative verb, a dative and a passive marker, or else (ii) just a causative verb and its related passive marker. No language appears to use the same form for their give verb, dative marker and the passive, without the causative use.

In the case of give verbs, Chappell and Peyraube (2006) propose the further refinement that there are at least two different pathways of grammaticalization, explicable in terms of polygrammaticalization (Craig 1991): 

(i) $V [+\text{give}] > \text{dative marker}$
(ii) $V [+\text{give}] > V [+\text{causative}] > \text{passive marker}$

In support of their analysis, they observe first of all that any proposal for the development from a dative marker directly to causative marker would not be supported by data from either Sinitic or surrounding language families (see Lord et al 2002). In fact, the grammaticalization into these functions takes place at distinct sites in two different kinds of serial verb construction. Second, it is not likely for a dative marker (usually a preposition) to grammaticalize into a causative verb (in Chinese languages, the ‘causative marker’ is still a verb). This would in fact represent the reverse process of de-grammaticalization or lexicalization. They conclude with a further hypothesis: 

As it is completely atypical to find a verb of giving developing directly into a passive marker, but quite common to find that passives are directly grammaticalized from causatives, the following implicational universal is tentatively proposed:

*If a language has a passive marker whose origin is a verb of giving, then it necessarily has a causative verb realised by the same form and having its source in a verb of giving.*

[GIVE > PASSIVE MARKER] $\Rightarrow$ [GIVE > CAUSATIVE]

7. **CHALLENGES AND PROJECTS FOR THE FUTURE:**
Some of the main challenges in typology are the following:

(i) *Language descriptions with a solid comparative, typological basis*
In China, many dozens of reliable and comprehensive descriptive grammars have been published on both Sinitic languages (called ‘Chinese dialects’ within China) and minority languages. These are often very detailed, not to mention timely in the case of the many endangered languages in China which are spoken nowadays only by the
older generations of speakers.  

In terms of theoretical framework, a desideratum would be to make more use of a comparative if not typological perspective in the writing up of such grammars, rather than just a purely descriptive approach. Second, in terms of transcription practices and methodology, for China to meet international standards in linguistic research, linguists should be encouraged to consistently use IPA or the relevant romanization system to transcribe all language examples and texts in their publications, alongside the provision of glossing. Many publications do not follow this convention at all, making it very difficult to cite data and examples. Using just the Chinese characters devised for standard written Chinese to transcribe language data from other dialects and unrelated minority languages is, to say the least, a haphazard and inaccurate way of representing sounds and meaning.

(ii) **Resolving the problem of typological alignment of Sinitic languages**

First of all, it is evident that merely examining basic order for constituents is insufficient. This problem needs to be addressed by examining each Sinitic language and its dialects one-by-one to provide detailed descriptions from a typological perspective, as per (i) above. In addition, this would ideally imply a three-pronged approach: (i) in-depth studies of diachronic change in all major Sinitic languages for which historical documentation is available, combined with (ii) the study of areal typology and (iii) the broad issue of languages in contact. The more reliable and comprehensive grammatical descriptions are with respect to these parameters, the better placed we will be to resolve typological issues of crucial importance in the Asian region.

(iii) **Current projects on the typology of Sinitic languages**

In the last decade, team projects on typology of Sinitic languages, outside of China, have flourished mainly in the European area while they can be characterized more in terms of the individual efforts of scholars in the USA, Japan and Australia. The Centre de Recherches Linguistiques sur l’Asie Orientale (CRLAO) in Paris carries out research in two main domains related to typology and diachronic syntax:

- **Typologie des dialectes et des langues sinitiques** (H. Chappell and A. Peyraube)
- **Syntaxe et sémantique diachroniques du chinois** (A. Peyraube)
- **Le chinois archaïque dans son contexte est-asiatique** (L. Sagart)

Typologists and historical linguists at the Centre de Recherches Linguistiques sur l’Asie Orientale are about to embark upon a four-year project with linguists from the Institute of Linguistics at the Chinese Academy of Social Sciences in Beijing. The project will concentrate on analysis of the language contact situation in Northwest China between Mandarin dialects and Altaic languages and their mutual influence. Historical documents will be examined and fieldwork carried out in several provinces and autonomous regions of central Asia including Xinjiang, Gansu and Ningxia. A
major part of the project is thus to reconsider the question of linguistic areas in China.

The International Institute for Asian Studies at the University of Leiden also hosts a project on typology directed by Rint Sybesma, on the syntax of the Sinitic and non-Sinitic languages of China. This project examines classifier usage, aspect markers and sentence-final particles across six languages found in China, three Sinitic—Mandarin, Wenzhou (Wu) and Cantonese (Yue)—and three unrelated languages from the Tai group, Hmong-Mien and Wa (Austro-Asiatic). Some of the findings have already been published in Sybesma and Cheng (2004, 2005).

In the USA, Anne Yue continues her very fruitful research on a project entitled *Issues in Chinese Dialectology: Taxonomy, Typology, Stratification and Areal Linguistics* at the University of Seattle (see also §6.2. above). Unfortunately, the last decade in the USA has seen a steady reduction in the number of large centres of research into Chinese linguistics, subsequent upon the moving away or retirement of faculty, such as at UC Berkeley where the doyen of Chinese linguistics, Chao Yuen Ren used to be the Agassiz Chair Professor of Oriental Languages and Literatures.

In Hong Kong, Stephen Matthews and Virginia Yip are leading a project on the ‘Min dialect of Chaozhou and comparative Chinese grammar’ to investigate the grammatical features of Chaozhou, a southern Min language. Some of the main topics to be included in the planned reference grammar are question forms, aspect and modality, comparative and passive constructions, and the relationship of tone change to syntax. This will doubtlessly continue the tradition of their earlier very solid and comprehensive research into Cantonese grammar (1994).

(iii) *Language atlases*

Two teams of researchers in China are working on new improved language atlases of China. One is being directed by Prof. Cao Zhiyun at the Languages University of Beijing and will cover over 1000 dialect points according to an inventory of phonological, lexical and syntactic features. The second project involves the revision of the 1987 *Language Atlas of China*, jointly edited by Prof. Stephen Wurm in Australia and Prof. Li Rong in China for the two different language versions. This team is being lead by Prof. Zhang Zhenxiong of the Chinese Academy of Social Sciences in Beijing. Typologists should look forward to these two sets of maps with relish.

The recently published *World Atlas of Language Structures* (Oxford UP, 2005) contains a wealth of data and its digitized version the possibility of crosschecking many different correlations, according to the parameters of the user’s choice. It represents a pioneering effort in typology, given the enormous amount of linguistic material included and should be lauded.

Naturally, the detail and the number of language points for East and Southeast Asia cannot be on the same scale as for the language atlases currently in preparation in China. Nonetheless, we would like to suggest that future revisions might include more than just three languages for China: standard Mandarin, Hakka and Cantonese. Very few conclusions of a typological nature can be made on the basis of just three languages to represent all of Sinitic, a language taxon which accounts for 1/5th of the
world's population. A large amount of diversity is being inadvertently missed.

As authors of this overview article on the state of the art for typology in Chinese linguistics, we believe that the prospect of more collaboration between European research centres, the Chinese Academy of Social Sciences in Beijing and Academia Sinica in Taipei is certain to lead to some very interesting new scientific discoveries in the future that can change the typological landscape substantially, if not dramatically.

REFERENCES:


of Chicago Press.


Wu Fuxiang. 吴福祥 (2001). Nánfāng fāngyán jīgě zhǔángtài bǔyǔ biāoji de láiyuán (yì)《南方方言几个状态补语标记的来源（一）》[On the origin of several staticive complements in the southern dialects (1)]. Fāngyán《方言》4 : 344-354.

Wu Fuxiang. 吴福祥 (2002). Nánfāng fāngyán jīgě zhǔángtài bǔyǔ biāoji de láiyuán (èr)《南方方言几个状态补语标记的来源（二）》[On the origin of several staticive complements in the southern dialects (2)]. Fāngyán《方言》1 : 24-34.


Chinese Linguistics (the 9th Annual Meeting of IACL), Singapore, June 25-27.

1 An overview is given in Chappell (2001) of the history and demography of the 10 main Sinitic languages. The volume represents one of the first attempts to redress the balance in Chinese linguistics of largely carrying out research into standard Mandarin.
2 The first works on typology to be translated into Chinese were (i) the translation of Greenberg (1963) by Lu Bingfu and Lu Zhijie in (1984) and (ii) the translation of Comrie (1981) by Shen Jiaxuan (1989).
3 Bai languages can be both SOV and SVO. Nonetheless, Bai and Sinitic languages have a very close relationship. Hence, Tibeto-Burmanists generally recognize that Bai has undergone heavy structural influence from Chinese. The same explanation is used for Karen, viz, its proximity to Tai and Sinitic groups, both SVO (Dryer 2003).
4 RelN implies the order OV, but not vice versa: OV does not imply RelN.
5 It is certain that word order change did not affect the variety of Mandarin used as the language of the imperial court during these periods. However, particular dialects of Mandarin in contact with Altaic languages in the border regions of China are documented as having undergone word order change to SOV and to having borrowed case marking in form of postpositions, for example, the Lingxia dialect of Zhongyuan Mandarin spoken in Gansu (Huang 1996).
6 Greenberg’s Universal 5 has a precondition: if a language has SOV as the dominant word order, then postposed genitives imply that adjectives will be too. Lu Bingfu finds this precondition is superfluous. For example, the Tai-Kadai languages spoken in China are typically SVO, with attributive phrases following the noun, yet genitives precede.
7 This appears to be the tendency according to a larger crosslinguistic study currently being undertaken by Peyraube, Chappell and Zhang Min on the possibility of the implicational universal described in this section.
8 Such a typical case is Wuxiang, an unclassified Chinese language, spoken in the mountainous regions of western Hunan. Only the older generation have full use of this largely undescribed language, their children use Southwestern Mandarin or standard Mandarin but understand Wuxiang while the current generation of grandchildren in Wuxiang families neither speak nor understand it. Younger generations have fully switched to the local lingua franca and/or standard Mandarin. The situation is similar for the Hmong population who live in this region. Data from one of the authors, H. Chappell,
who carried out fieldwork in the Guzhang district in November, 2006.

9 Chappell brought her ongoing project on ‘Typology of Sinitic languages’, commenced in 1992 at La Trobe University, Melbourne, Australia to the CRLAO in 2005. The main topics are on the morphosyntax of constructions such as passives, causatives, object-marking, datives, comitatives and comparatives.