

TOPIC IN WORD ORDER TYPOLOGY

Xu Liejiong

*Department of Chinese, Translation & Linguistics
City University of Hong Kong*

There has been a heated debate as to whether Chinese is an SOV language or an SVO language in terms of word order typology. This article provides a new perspective on the word order controversy. Linguists that assign Chinese to either type share the view that word order typology must be based on the ordering of subject (S), object (O) and verb (V) relative to one another. However, for topic-prominent languages, of which Chinese is cited as an example in Li and Thompson (1976), a fourth element, topic (T), should be added in typological classification. Without it, it is difficult to pigeonhole Chinese as a whole and it is also difficult to tell the dialects apart. With it, we can easily show that Mandarin is predominantly TSV0 and Shanghainese STVO.

Keywords: topic, word order, typology, Chinese, subject

1. THE PROBLEM

Since Greenberg (1963) cross-linguistic difference in basic word order has been recognized as a major syntactic typological parameter. A language belongs to one of the six types, SOV, SVO, VSO, VOS, OVS, OSV, in terms of the surface ordering of subject, object and verb relative to one another. With the understanding that it is the dominant word order of a language we are considering, disregarding possible variations, most of the languages that have been investigated can be assigned to one type or another. However, typologists are embarrassed to find it difficult to pigeonhole Chinese, the language that has by far the largest population of speakers. Its basic order is considered SVO by some grammarians but SOV by others.

Chinese is not a free order language or a nonconfigurational language, a term used in generative grammar, capable of free permutation and scrambling of constituents. Nor does the difficulty in classification arise from one of the problems familiar to linguists. Unlike German, Chinese shows no difference between the main clause and the subordinate clause in terms of word order. Unlike French, its subject does not split across two elements, typically a full noun phrase preceding the verb and a pronoun following it. Whence the difficulty, then? This article aims to provide a new perspective on the problem.

Although they disagree, typologists that classify Chinese as either an SVO language or an SOV language feel confident of their decision, claiming that they are well supported by linguistic evidence. Take some simple sentences for example.

- (1) Ta maile neizhuang fangzi
he sold that house
- (2) Ta ba neizhuang fangzi maile
he PREP that house sold

The noun phrase *neizhuang fangzi* can occur either preverbally or postverbally. The preposition *ba* is not required, particularly in some contexts and in some dialects.

- (3) Ta neizhuang fangzi maile
he that house sold

The NP before the verb in (3) is known as a subtopic, in contrast to the sentence initial topic in (4).

- (4) Neizhuang fangzi ta maile
that house he sold

For easy exposition, let us use topic as a cover term for the relevant noun phrases in (2), (3) and (4), that is, for preverbal NPs other than those generally recognized as subjects. In fact, whether all of them are topics is another controversial issue, not addressed here. The assignment of Chinese to one or another of the word order types crucially depends on the analysis of such preverbal NPs.

2. Topic Identical To Object?

Since Venneman (1972) and Lehman (1973) many typologists have been convinced that the order of subject is less relevant from a general typological view. So a crucial question is whether the preverbal NPs in the examples cited above, referred to as topics here, are, in fact, objects.

As it is mentioned above, word order typology considers only the basic sentence form without regard to the various derived forms. Those who analyze Chinese as an SOV language take (2) as a basic form in Modern Chinese while those who disagree with them take it as a derivative. Whether topics are derived as a result of movement receives attention in generative grammar. Generative grammarians are divided in their treatment of topic structures illustrated in (4). It is claimed to have undergone movement in Huang (1982) and to be base-generated in Xu and Langendoen (1985).

It is difficult to maintain that a topic in Chinese is always a moved constituent, as it may occur in a sentence where the object remains unmoved.

- (5) Neizhuang fangzi ta maile yi ceng
 that house he sold one floor
 'As for that house, he sold one floor of it (i.e. not the whole building)'
- (6) Ta neizhuang fangzi maile yi ceng
 he that house sold one floor
- (7) Ta ba neizhuang fangzi maile yi ceng
 he PREP that house sold one floor

Such structures are typical of Chinese, a topic-prominent language in terms of Li and Thompson (1976). They have no parallels in English, a subject-prominent language. It stands to reason to regard structures represented by (2)-(4), where the object happens to be empty, as special cases of those in (5)-(7), where another noun phrase coexists with the subject and the object. I will argue that the difficulty in classifying Chinese as any of the six types arises from linguists' failure to recognize this basic constituent property of Chinese, i.e. topic is independent of subject or object.

Attempts have been made to rephrase (5)-(7) as (8) with the intention to show that the topic is merely a part of the object, if not the whole of it.

- (8) Ta maile neizhuang fangzi de yi ceng
 he sold that house 's one floor
 'He sold one floor of the house.'

Such efforts cannot reduce the topic to the object or part of it since one can easily come up with sentences in which they cannot be linked by the modifier marker *de*.

- (9) Neishan men ta tile yige dong
 that door he kicked a hole
 'He kicked a hole in the door.'
- (10) * Ta tile neishan men de yige dong
 he kicked that door 's a hole

The relation between the topic and the comment or a particular element within the comment is semantic. Syntactically, the topic is not always recoverable in the comment. Therefore, that all topic structures are derived is an invalid generalization.

Having shown that the topic is not necessarily the whole or a part of the object, we briefly dismiss the proposal that the topic is part of the subject as is argued in Schlobinski and Schüze-Coburn (1992). The example they use is (11).

- (11) Neike shu yezi hen da
 that tree leaves very large
 'As for that tree, its leaves are large.'

Based on their observation that there exists a possessive relation between the two adjacent NPs, they propose an underlying structure (12), where an associative particle *de* is added to

mark the relation between the modifier and modified. This analysis enables the complex NP to function as subject without appealing to the notion of topic.

- (12) Naike shu de yezi da
 that tree de leaves large
 'The leaves of that tree are large.'

This is obviously an oversimplified solution to the so-called "double-subject" problem in view of the existence of a variety of topic structures such as (5)–(9) where the two related NPs are not adjacent so that the insertion of *de* after the first NP is impossible.

3. Topic Grammaticalized in Chinese?

One may object that topic is a discourse notion, which has no place in syntactic structures even in languages like Chinese. If such were true and if the word order typology is restricted to the relative order of grammatical elements in a clause, then the position of topics would be irrelevant. In what follows I will show that the discourse notion of topic is grammaticalized in Chinese.

A pragmatic topic which can be defined in terms of information structure is not always encoded as a syntactic topic. It may undergo various forms and various degrees of grammaticalization in different languages or even in a single language. The following Chinese sentences show an ascending order of grammaticalization.

- (13) Xuduo ren renwei Caocao shi huan-ren
 many people think Caocao is villain
 'Many people think Caocao is a villain.'
 (14) Shuo dao Caocao, xuduo ren renwei ta shi huai-ren
 speak about Caocao many people think he is villain
 'Talking about Caocao, many people think he is a villain.'
 (15) Zhiyu Caocao, xuduo ren renwei ta shi huai-ren
 about Caocao many people think he is villain
 (16) Caocao me, xuduo ren renwei ta shi huai-ren
 Caocao TOP many people think he is villain

In (13) Caocao is a discourse topic without any formal indication, unless it is read with extra stress or a particular intonation. In (14) a lexical verb in the sense of 'speak' is employed to introduce a topic of conversation. In (15) it is preceded by a preposition, similar to *as for* in English. In (16) a morphological marker *me*, glossed TOP, is used after the topic. In Mandarin Chinese, *me* and other particles may appear at a sentence final position to express modality as well. In some of the Chinese dialects, for instance, in the Wu dialect, represented by Shanghainese, there are special markers used to indicate topics exclusively.

- (17) Zoco zy, jiokue ngeng ngengue i zy uangeng
 Caocao TOP many people think he is villain

Accompanied by this special topic marker, *Zoco* in (17) is a highly grammaticalized topic.

As a topic-prominent language, Chinese has topic markers, but no subject markers corresponding to the agreement inflections in European languages. The fact that not all topics require a marker certainly does not disprove that topics are grammaticalized in Chinese any more than the absence of agreement inflections in most of the sentences in English does not disprove that subjects are grammaticalized. Moreover, insertion of a topic marker after a topic is always possible in Chinese.

Repeated attempts have been made to reduce topic to subject, thus solving the so-called “double subject” problem in structures like (4). The reductionists would argue that the topic markers identified above are, in fact, subject markers. The distinction between the topic and the subject is by no means obvious. As Chinese allows a subject to drop freely, where there is only one noun phrase before the verb, it is not clear whether it is a topic or a subject. Furthermore, since Chinese permits multiple topicalization, even if there are two or more noun phrases, both or all of them may be topics. Some researchers arbitrarily define the one closest to the verb as the subject, e.g. Xu Yulong (1995).

However, the topic and the subject are not always indistinguishable. The following minimal pairs in Shanghainese show the distinction.

- (18) a. Diqgeq xiohutsy buo tozy se ting
 this young-man climb onto hill top
 ‘This young man has climbed to the hill-top.’
 b. Diqgeq xiohutsy me, buo tozy se ting
 this young-man TOP climb onto hill top
 (19) a. Iqgeq xiohutsy buo tozy se ting
 a young-man climb onto hill top
 ‘A young man has climbed to the hill-top.’
 b. *Iqgeq xiohutsy me, buo tozy se ting
 a young-man TOP climb onto hill top

While a subject can be indefinite, a topic must be definite. Of the four sentences only (19b) is ungrammatical, where the relevant noun phrase is indefinite and yet overtly marked by the topic marker *me*.

To summarize, I have shown that topic in Chinese is a grammaticalized element in the sentence, not identical with either the object or subject, though in some cases an NP not followed by an overt marker is ambiguous.

4. WORD ORDER TYPOLOGY

Now we can address the problem raised at the outset: is Chinese an SVO language or an SOV language? On the one side are Light (1979), Mei (1980), Sun and Givon (1985) and on the other are Tai (1973), Li and Thompson (1976, 1978), etc. More recently the polemics has been taken over by linguists working in the Principles-and-Parameters approach. They, too, are divided, with Li (1990) and Mulder and Sybesma (1993) representing the two sides.

The problem lies not so much on a point of disagreement between the two sides as on their shared view that word order typology must be based on the ordering of subject (S), object (O)

and verb (V) but nothing else. In a subject-prominent language, subject is grammaticalized but topic is not grammaticalized or less grammaticalized. So long as we consider subject-prominent languages only, the arrangement of the three elements can successfully describe the typological variation among them. But a problem naturally arises when linguists apply the typology to topic-prominent languages like Chinese, in which what is grammaticalized is topic rather than subject. Chinese does not fit into this pattern simply because there is a fourth element, topic (T), which is as important as, or even more important than, subject and yet its role in word order is neglected.

Since T is not included, researchers have tried to identify it with either O or S or part of one or the other. The argument for the SOV analysis and Schlobinski and Schüze-Coburn (1992)'s proposal represent such attempts.

There is another related question. If the relevant preverbal NPs in (2) and (3) are objects, then the one in (4) should also be analyzed as an object. Why has nobody argued that Chinese is an OSV language based on the order illustrated in (4)? It cannot be because SOV has more frequent occurrences than OSV as we know intuitively and statistically the reverse is true. Those who argue that the basic order of Chinese sentences is SOV have the burden to explain why it is not OSV. Why has it never occurred to them that they owe us an explanation? My speculation is that they believe their disregard for the OSV order is warranted by the well-established analysis of English as an SVO language. The English counterpart of (4) is (20).

(20) That house, he sold

When they classify English as being SVO, linguists abstract away from the existence of topic structures like (20) and interrogative structures like (21).

(21) Which house did he sell?

However, this line of reasoning does not do full justice to the role a topic plays in Chinese constituent structures. The topic construction is a derived form in English, a subject-orientated language. But it is a dominant form in Chinese, which is a topic-prominent language. Even Charles Li and Sandra Thompson, who first introduced the typology to classify languages into topic-prominent and subject-prominent, insisted that Chinese is an SOV language, never considering the possibility of OSV. The importance of their new typology has not yet been fully recognized by its proponents as well as other typologists.

Without T, it is difficult to pigeonhole Chinese as a whole and it is difficult to tell the Chinese dialects apart. With it, we can easily show the difference in word order between Chinese dialects. Mandarin is TSVO whereas Shanghainese is STVO. A recent statistical study cited in Xu and Liu (in press) shows that the ratio of TV to VO is 4.5 to 1 in Shanghainese non-wh-questions and 7.4 to 1 in negative statements. There are more occurrences of TSV and TSVO in Mandarin and more occurrences of STV and STVO in Shanghainese. What follows is a minimal pair.

(22) ?? Ta yiben shu geile Xiao Li
 he one book gave Xiao Li
 'He gave Xiao Li a book.'

- b (23) I iqpeng sy peqtsy Xo Li leq
 he one book gave Xo Li PRT

Whereas the Mandarin sentence (22) is hardly acceptable, its Shanghainese counterpart (23) is, with the subtopic interpreted as a definite NP. Although no large-scaled statistics are available, the intuitions of those who speak both dialects are unanimous.

To summarize, I have argued that T should be added to S, O, V in word order typology in classifying topic-prominent languages.

REFERENCES

- Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In: *Universals of Language* (J. Greenberg (Ed.)), 73-113. MIT Press, Cambridge.
- Huang, C.-T. J. (1982). *Logical Relations in Chinese and the Theory of Grammar*: PhD dissertation. MIT, Cambridge.
- Lehmann, W. (1973). A structure principle of language and its implications. *Language* 49: 47-66.
- Li, A. Y-H. (1990). *Order and Constituency in Mandarin Chinese*. Kluwer, Dordrecht.
- Li, C. N. & S. Thompson. (1976). Subject and topic: a new typology of language. In: *Subject and Topic* (Charles Li (Ed.)), 457-489. Academic Press, New York.
- Li, C. N. & S. Thompson. (1978). An exploration of Mandarin Chinese. In: *Syntactic Typology: Studies in the Phenomenology of Language* (W.P. Lehman (Ed.)), 223-266. University of Texas Press, Austin.
- Light, T. (1979). Word order and word order change in Mandarin. *Journal of Chinese Linguistics* 7: 149-180.
- Mei, K. (1980). Is modern Chinese really an SOV language? *Cahier de linguistique Asie Orientale* 7: 23-45.
- Mulder, R. & R. Sybesma. (1992). Chinese is a VO language. *Natural Language and Linguistic Theory* 10: 439-476.
- Schlobinski, P. & S. Schüze-Coburn. (1992). On the topic of topic and topic continuity. *Linguistics* 30: 89-121.
- Sun, C. F. & T. Givon. (1985). On the so-called SOV word order in Mandarin Chinese. *Language* 61: 329-351.
- Tai, J. H-Y. (1973). Chinese as an SOV language. In: *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society* (C. Corum (Eds)). Chicago Linguistic Society, Chicago.
- Vennemann, T. (1972). Analogy in generative grammar, the origin of word order. *Proceedings of the Eleventh International Congress of Linguistics* 2: 78-83.
- Xu, L. J. & T. Langendoen. (1985). Topic structures in Chinese. *Language* 61: 1-27.
- Xu, L. J. & Liu D. Q. (In press). *Topic: Structural and Functional Analyses*. Shanghai Education Press.