Chinese Word Order: A Look Into SVO and SOV Arguments

Shelley E. Tennison
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/honorstheses

Part of the Typological Linguistics and Linguistic Diversity Commons

Let us know how access to this document benefits you.

Recommended Citation
https://doi.org/10.15760/honors.1221

This Thesis is brought to you for free and open access. It has been accepted for inclusion in University Honors Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.
Chinese Word Order:
A look into SVO and SOV arguments

by
Shelley Tennison

An undergraduate honors thesis submitted in partial fulfillment of the
Requirements for the degree of
Bachelor of Arts
in
University Honors
and
Applied Linguistics

Thesis Advisor
Stephen Wadley

Portland State University
2022
Abstract

In this paper, the issue of Chinese word order is weighed, and findings in the field are presented. First, looking into the historical background and offering a brief introduction to Greenberg's linguistic universals, I will break apart various arguments concerning the word order in Chinese. I will examine many sources, albeit the focus will rest heavily on Sandra Li and Charles Thomson's SOV argument and Timothy Givon and Chaofen Sun's SVO argument. The two sides of this argument are considered regarding X-bar theory, constituent test, morphological limitations, and syntax.

**Keywords:** Chinese, Word Order, Preposition, SVO, SOV, Language Universals
Chinese Word Order:

A look into SVO and SOV arguments

In natural languages, there are observable features. By looking closer at these features, linguistic universals can start to be named. Of the universals, there are two types, absolute and statistical. An absolute universal might be that all languages have nouns and verbs or that all spoken languages consist of vowels and consonants. A statistical universal might be a tendency in certain languages but might not be accurate for all languages. Linguistic typology seeks to uncover generalizations that are made about languages. By naming and categorizing syntactical patterns that are discoverable across thirty natural languages, American linguist Joseph Greenberg has been able to list a set of 45 linguistic universals. A complete list of these universals obtained from the Linguistic Society of America and can be found in Appendix 1. In the following paper, I will be turning to Joseph Greenberg's proposed set of linguistic universals to determine how closely Chinese word order matches these “universals.” In doing so, I hope to understand better common arguments for both the SOV and SVO word orders.

Languages have different word orders when organizing the constituents within sentences. The study of word order primarily focuses on the ordering of subject, verb, and object. Word order is also used to analyze the placement of adverbials and the order of modifiers within a noun phrase, such as adjectives, possessives, demonstratives, numerals, and adjuncts. For some languages, word order is simply a grammatical property used to encode grammatical information. In contrast, some inflectional languages tend to allow for additional flexibility allowing for pragmatic information to be encoded. The tendency or statistical universal remains that languages often stick to a preferred word order.
Some historical Chinese linguists have made claims that there has been an observable drift in Chinese from the SVO (Subject Verb Object) word order to SOV (Subject Object Verb) word order. For example, Li and Thompson believe that the case of Chinese changing from SVO to SOV has just begun and that new pathways of language change should be considered. This process has been slow going, taking some two millennia to get to its current point. Although Li and Thompson note that Chinese still retains instances of S + V + PP, they believe that these instances only occur under particular constraints and pay homage to the dominant monosyllabic verb forms of Archaic Chinese (Li & Thompson, 1974, p. 202). These claims insinuate that Modern Mandarin should be considered an SOV language. However, a quantitative discourse analysis conducted by Chao-fen Sun and Talmy Givón found that 90% of all syntactic objects occur after the head verb (V′) in both written and spoken texts, directly contradicting such suppositions. Li and Thompson's claim that word order in Mandarin Chinese is shifting from an SVO language to an SOV language due to the grammaticalization of serial verb constructions is not supported by such data (Li & Thompson, 1974).

Furthermore, Li and Thompson claim that SOV word order characteristically codes definite objects while SVO word order codes indefinite objects (Li & Thompson, 1975). This can be seen in instances of S + bā + Direct Object + Verb, where the use of bā implies a definiteness to the direct object. However, when referring to their analysis, Chao-fen Sun and Talmy Givón found that "the functional distribution of OV in both texts suggests that it is an emphatic/contrastive discourse device, having little to do with the contrast between definite and indefinite object" (Sun & Givon, 1985). The following paper aims to break down such claims by offering a side-by-side comparison of the two arguments better to understand Chinese syntax and the SOV and SVO arguments.
Historical and Linguistic Background

As Chao-fen Sun (1996) mentions, there are two views regarding language in the Western tradition, conventionalist and naturalist, that stem from Plato's *Cratylus*. In the conventionalist viewpoint, language and the naming conventions for things and objects are arbitrary and simply formed as a matter of convention and mutual agreement. While in the naturalist viewpoint, language is seen as organic and naturally evolving in an effort to encode that which is named with inherent meaning in relation to other known names (Plato & Fowler, 1977). These two views of language were not limited to the West. Existing at roughly the same time as Plato was Chinese philosopher Xunzi, who made similar observations about language. These observations are elaborated upon by S.-Y. Wang in an article remembering the Chinese Linguist, Professor Li Fang-Kuei (Wang, 1989)

The similarity here between Xunzi and Plato then becomes apparent. For both, the relation between the world and that which it names is conventionally (or arbitrarily) formed. However, some relations are better (or more appropriate) than others. Every language has a small stock of words which retain some iconic features, and therefore depart from complete arbitrariness.

When studying syntax, one must seek to understand the arbitrariness and iconicity of language. In doing so, one can understand how words work together on a fundamental level to function within a sentence. In seeking to understand syntactic change noting where the departure from complete arbitrariness occurs is instrumental in understanding how the iconic features are linked to syntactic changes and will help linguists further understand their role in the grammaticalization process. All this culminates in a linguist
being able to understand all elements at play in a given sentence and so they will be better able to determine the overarching sentence structure of the language.

Throughout this paper, when referring to Chinese, I, along with established scholars within the field, are predominately referring to the written language unless otherwise explicitly stated. Despite China's numerous mutually unintelligible spoken dialects, the written language has been shared for 2,000 years. The Chinese writing system is morphemic. Every graph represents a morpheme, and each morpheme is overwhelmingly monosyllabic, corresponding to the representation of each syllable in the individual graphs. Many believe that the written Chinese of the late Zhou dynasty and early Han dynasty was closely related to the spoken language, and vernacular Chinese found in the late Tang dynasty and into the Song dynasty similarly was closely associated with the spoken language. Similar relations have been made between the vernacular written works during the Yuan and Ming dynasties, reflecting the time's spoken language. Today's standard written language is directly related to and developed from the literary vernacular of the Song and Yuan dynasties (Norman, 1988). A further expanded overview of the chronological divisions of Chinese history as laid out by Norman (1988) in Chinese can be referenced in Appendix 2.

**Observations and arguments concerning SVO to SOV drift**

Li and Thompson argue that the SVO to SOV drift has been a slow process, going on for some two millennia. They view Modern Chinese as having a word order of S + PP + V, where PP contains both a preposition and a NP. However, Li and Thompson acknowledge that Modern Chinese allows for some instances of S + V + PP in cases of monosyllabacity or the "the lack of
complex morphological structure of the verb" (Li & Thompson, 1974, p. 202). This point is illustrated through the following author-generated examples:

(1) Zhāng-sān shūi zài chuáng-shàng.

Zhang-san sleep at bed on

Zhan-san sleeps on the bed.

(2) (a) *Zhāng-sān shùijiao zài chuáng-shàng.

Zhang-san sleep at bed on

(b) Zhāng-sān zài chuáng-shàng shùijiao.

Zhang-san at bed on sleep

Zhang-san sleeps on the bed. (Li & Thompson, 1974, p. 202)

Example (5) (a) appears to be ungrammatical due to the polysyllabicity of the verb *shùijiao 'sleep'. In Archaic Chinese, verbs were monosyllabic. While in Modern Mandarin, verbs have become more complex, Li and Thompson believe that S + V + PP constructions are simply remnants of Archaic Chinese. Moreover, Li and Thompson believe that there has not just been a gradual shift from SVO to SOV but rather that there has been a shift from SOV to SVO and then back to SOV. As elaborated in the following quote:

If the existence of OV characteristics is regarded as a causative factor in the shift from VO to OV, one may, then wonder: what brought about such OV characteristics in Archaic Chinese in the first place? A speculation about their source, based on evidence, is possible. Such a speculation would designate pre-Archaic Chinese (before 12th century B.C. as an SOV language. This pre-Archaic Chinese had changed into SVO by the Archaic period (3rd to 10th century B.C.). However, before the SOV stage could fully mature, i.e., before all SOV
characteristics had been supplanted by VO characteristics, the language embarked upon another pathway to shift back to SOV. (Li & Thompson, 1974, p. 208)

Sun and Givón's 1985 study found that "Mandarin is a rather typical SVO language, and that OV is a contrastive/emphatic device with rather restricted text distribution" (Sun & Givón, 1985, p. 330). The study shows that VO word order is overwhelmingly more common at 94% in written language and 92% in spoken conversation, clearly demonstrating that infrequent OV constructions cannot be considered a rising trend.

**Prepositions and their role in SOV and SVO arguments**

The placement of prepositional phrases (PPs) might just be the most frequented argument in favor of the SOV drift hypothesis, "in old Chinese, PPs indicating double-object construction, passive construction, source/locative/goal, and comparative can be or are, postverbal. Yet in Modern Mandarin, all PPs other than those indicating double-object construction and goal must be preverbal" (Sun, 1996, p. 11). This suggests that we should look closer at the parameter setting rather than referring to the PPs as evidence for SOV drift. Looking at the parameter setting of Old Chinese, it is clear that PPs can occur either preverbally or post verbally, while in Modern Mandarin, the adjunct PPs occur in the preverbal position \([\text{XP}_{\text{adjunct}} \ X']\) or post verbally \([X' \text{XP}_{\text{adjunct}}]\). X-bar structure is significant when studying syntax because by looking to it, syntacticians can better understand the differing level of the constituency. In X-bar theory, there are three levels within each phrase, XP is used to represent the phrasal level, X' is used to represent the bar level, and X is used to represent the head or word that assigns the category to the phrase, be it a noun, verb, preposition, adjective, or adverbial. The three levels can be written as shown in Figure 1 below:
The variables X, W, Y, and Z are used to represent the different Phrases. In Figure 2, the kinship relationships can be seen through the 'mother node' or XP, which branches into two 'daughter nodes' YP and X'. YP is 'daughter' to XP and 'sister to X' (Carnie, 2013, p. 199).

Carnie describes an *adjunct* as an XP that is sister to a single bar level (N', V', A', or P') and that is daughter to X'. In Figure 2 above, an adjunct is represented by ZP. An *adjunct* is considered to operate within a sentence in the same way that an adverbial or oblique would, by carrying additional, often optional, information (Carnie, 2013, p. 176). However, an argument in
favor of the SOV drift hypothesis is dependent on PPs being classified as a *complement* or sister to a head (N, V, A, or P) and therefore corresponding to the notion of 'object' in traditional grammar. In Figure 2 above, a complement is represented by WP.

We can conduct a series of tests to determine the complement and adjacency status of a constituent in English. A movement test inserts a like constituent between the tested party and the head verb or moving the tested constituent altogether, such as:

(1) John *in the library* reads a book.

(2) John *in the library* at school reads a book.

Additional tests, such as fronting, can be performed; fronting moves the tested constituent to the beginning of the sentence to test the grammaticality.

(3) *In the library* John reads a book.

In all three of the above examples, the sentence remains grammatical, proving that the relationship between the constituent [*in the library*] and the verb phrase [*read a book*] is an adjunct relationship.

The arguments used by Li and Thompson to substantiate their claims of SOV drift rest heavily on the classification of preverbal PPs as an object of the main verb, and many of their sentences used to prove this fact are self-generated and not pulled from text or corpus data.

Movement tests can also be used to test the relationship between Chinese PPs and Verbal heads. For example:

(4) Tā zài chūfáng lǐ zào jiāozi.

He at kitchen-in make dumplings

He is making dumplings in the kitchen. (Li & Thompson, 1975, p. 266)

(5) Tā zài chūfáng lǐ de zhuōzi shàng zuò jiaozi.
He at kitchen in of table on make dumplings.
He is making dumplings on the table in the kitchen.

If the same constituent is fronted, it remains grammatical:

(6) Zài chūfáng lǐ tā zòu jiǎozi.
At kitchen-in he make dumplings.
In the kitchen, he is making dumplings.

In example (4) Li and Thompson are treating the constituent [zài chūfáng lǐ] as obligatory information. At the same time, the location of the action [zòu jiǎozi] is additional information, and without it, the sentence can still function grammatically. The constituent tests prove that the tested constituent has an adjunct relationship with the VP and should not be treated like an 'object'. Instead, the optional nature of the constituent firmly secures its status as functioning as an adverbial.

**Ba construction and its role in SOV and SVO arguments**

Y.R. Chao's *Grammar of Spoken Chinese* (1968) is quoted in Chaofen Sun's *Word Order and the Grammaticalization in the History of Chinese* (1996), saying that the bā construction is closely associated with the notion of transitivity "[a] special form of the V-V series has a first verb, the pretransitive, and an object which ordinarily would be the object of the pretransitive is often regarded as a form of the inverted object" (52). A verb is said to be transitive when it is affecting two NPs and intransitive if only one NP is affected. Transitivity appears to be a common feature among languages but its use and understanding comes down to semantics and a pragmatical understanding that might not be able to be easily translated from one language to another. In addition to the transitive aspect of the bā construction, it is also said to mark a temporally bound event, something that is not an ongoing action like to love or to miss someone.
Additionally, the object following bā needs to be something definite and concrete. Often we see bā occur in sentences using this, that, these, those etc.

While a different explanation of the bā construction's role in the SVO to SOV shift is described by Li and Thompson as, "the verb in an SVO language can develop into a case marker thus collapsing SVO complex sentences into simple SOV sentences" (Li & Thompson, 1974, p. 209). This shows us that in Modern Chinese prepositions as seen in S + bā + O + V had verbal origins, and should now be considered to function as case markers. If bā is functioning as a case assignor and is responsible for subjects to receive nominative case or (I, you, he, she, they) for objects to receive accusative case (me, you, him, her, they), in many languages, English included, inflection and morphological features can be used to further convey grammatical categories such as case. However, Chinese lacks inflectional morphemes so it could be said that the movement seen within the bā construction is used to convey case to accommodate the morphosyntactic limitations of the language.

Conclusion

Looking again to the typological features of SOV languages, Greenberg argues that SOV languages do not have prepositional features and that when genitive order does indeed deviate that the adjective order is also subject to change (Greenberg, 1963, p.79). This is reflected in the following universals:

*Universal 4:* With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional.

*Universal 5:* If a language has dominant SOV order and the genitive follows the governing noun then the adjective likewise follows the noun.
Universal 7: If in a language with dominant SOV order, there is no alternative basic order, or only OSV as the alternative, then all adverbial modifiers of the verb likewise precede the verb.

Chinese seems to not follow the aforementioned universals further suggests the possibility that such constituents should be considered in other ways. It seems that the underlying structure of Chinese is that of SVO but there is definitive evidence of SOV characteristics such as modifiers preceding their heads.

One explanation of this could be rooted in the fact that Chinese is not easily explained in terms like Subject and Object. As explained by Li and Thompson (1981), “it seems sensible to regard Mandarin as a topic-prominent rather than a subject-prominent language, since the basic structure of sentences can be more insightfully described in terms of topic-comment relation rather than in terms of the subject-predicate relation” (19). Even though Chinese appears to have a predominate SVO word order it could be possible that it simply does not fit into the universals that have been developed, and in order to understand the syntactical structure of the language a word order typography should be developed for topic-comment languages like Mandarin, Japanese, Lakota, and Turkish.

More research into this should occur, looking closely to languages outside of Indo-European language groups could provide greater insight.
References


Appendix 1
Greenberg's linguistic universals obtained from the Linguistic Society of America (Greenberg, 1963).

Word order

1. In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object.

2. In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes.

3. Languages with dominant VSO order are always prepositional.

4. With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional.

5. If a language has dominant SOV order and the genitive follows the governing noun, then the adjective likewise follows the noun.

6. All languages with dominant VSO order have SVO as an alternative or as the only alternative basic order.

Syntax

7. If in a language with dominant SOV order there is no alternative basic order, or only OSV as the alternative, then all adverbial modifiers of the verb likewise precede the verb.

8. When a yes-no question is differentiated from the corresponding assertion by an intonational pattern, the distinctive intonational features of each of these patterns are reckoned from the end of the sentence rather than from the beginning.
9. With well more than chance frequency, when question particles or affixes are specified in position by reference to the sentence as a whole, if initial, such elements are found in prepositional languages, and, if final, in postpositional.

10. Question particles or affixes, when specified in position by reference to a particular word in the sentence, almost always follow that word. Such particles do not occur in languages with dominant order VSO.

11. Inversion of statement order so that verb precedes subject occurs only in languages where the question word or phrase is normally initial. This same inversion occurs in yes-no questions only if it also occurs in interrogative word questions.

12. If a language has dominant order VSO in declarative sentences, it always puts interrogative words or phrases first in interrogative word questions; if it has dominant order SOV in declarative sentences, there is never such an invariant rule.

13. If the nominal object always precedes the verb, then verb forms subordinate to the main verb also precedes it.

14. In conditional statements, the conditional clause precedes the conclusion as the normal order in all languages.

15. In expressions of volition and purpose, a subordinate verbal form always follows the main verb as the normal order except in those languages in which the nominal object always precedes the verb.

16. In languages with dominant order VSO, an inflected auxiliary always precedes the main verb. In languages with dominant order SOV, an inflected auxiliary always follows the main verb.
17. With overwhelmingly more than chance frequency, languages with dominant order VSO have the adjective after the noun.

18. When the descriptive adjective precedes the noun, the demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise.

19. When the general rule is that the descriptive adjective follows, there may be a minority of adjectives which usually precede, but when the general rule is that descriptive adjectives precede, there are no exceptions.

20. When any or all 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.

21. If some or all adverbs follow the adjective they modify, then the language is one in which the qualifying adjective follows the noun and the verb precedes its nominal object as the dominant order.

22. If in comparisons of superiority the only order, or one of the alternative orders, is standard-marker-adjective, then the language is postpositional. With overwhelmingly more than chance frequency if the only order is adjective marker-standard, the language is prepositional.

23. If in apposition the proper noun usually precedes the common noun, then the language is one in which the governing noun precedes its dependent genitive. With much better than chance frequency, if the common noun usually precedes the proper noun, the dependent genitive precedes its governing noun.
24. If the relative expression precedes the noun either as the only construction or as an alternate construction, either the language is postpositional, or the adjective precedes the noun or both.

25. If the pronominal object follows the verb, so does the nominal object.

**Morphology**

26. If a language has discontinuous affixes, it always has either prefixing or suffixing or both.

27. If a language is exclusively suffixing, it is postpositional; if it is exclusively prefixing, it is prepositional.

28. If both the derivation and inflection follow the root, or they both precede the root, the derivation is always between the root and the inflection.

29. If a language has inflection, it always has derivation.

30. If the verb has categories of person-number or if it has categories of gender, it always has tense-mode categories.

31. If either the subject or object noun agrees with the verb in gender, then the adjective always agrees with the noun in gender.

32. Whenever the verb agrees with a nominal subject or nominal object in gender, it also agrees in number.

33. When number agreement between the noun and verb is suspended and the rule is based on order, the case is always one in which the verb precedes and the verb is in the singular.

34. No language has a trial number unless it has a dual. No language has a dual unless it has a plural.
35. There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and the trial are almost never expressed only by zero.

36. If a language has the category of gender, it always has the category of number.

37. A language never has more gender categories in nonsingular numbers than in the singular.

38. Where there is a case system, the only case which ever has only zero allomorphs is the one which includes among its meanings that of the subject of the intransitive verb.

39. Where morphemes of both number and case are present and both follow or both precede the noun base, the expression of number almost always comes between the noun base and the expression of case.

40. When the adjective follows the noun, the adjective expresses all the inflectional categories of the noun. In such cases the noun may lack overt expression of one or all of these categories.

41. If in a language the verb follows both the nominal subject and nominal object as the dominant order, the language almost always has a case system.

42. All languages have pronominal categories involving at least three persons and two numbers.

43. If a language has gender categories in the noun, it has gender categories in the pronoun.

44. If a language has gender distinctions in the first person, it always has gender distinctions in the second or third person, or in both.

45. If there are any gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also.
The sample was, in the spellings that he used: Basque, Berber, Burmese, Burushaski, Chibcha, Finnish, Fulani (Fula), Greek, Guarani, Hebrew, Hindi, Italian, Kannada, Japanese, Loritja (Luritja), Malay, Maori, Masai, Maya, Norwegian, Nubian, Quechua, Serbian, Songhai, Swahili, Thai, Turkish, Welsh, Yoruba, Zapotec.
## Appendix 2

### Major Chronological Divisions of Chinese History as shown in Norman's *Chinese* (Norman, 1988).

<table>
<thead>
<tr>
<th>Dynasty</th>
<th>Time Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xià dynasty</td>
<td>Twenty-first to sixteenth centuries BC</td>
</tr>
<tr>
<td>Shāng dynasty</td>
<td>Sixteenth to eleventh centuries BC</td>
</tr>
<tr>
<td>Western Zhōu dynasty</td>
<td>Eleventh century to 771 BC</td>
</tr>
<tr>
<td>Spring and Autumn period (Chūnqiu)</td>
<td>770 to 476 BC</td>
</tr>
<tr>
<td>Warring States period (Zhànguó)</td>
<td>475 to 221 BC</td>
</tr>
<tr>
<td>Qín dynasty</td>
<td>221 to 207 BC</td>
</tr>
<tr>
<td>Western Hàn dynasty</td>
<td>206 BC to AD 24</td>
</tr>
<tr>
<td>Eastern Hàn dynasty</td>
<td>AD 25 to 220</td>
</tr>
<tr>
<td>Three Kingdoms period (Sānguó)</td>
<td>AD 220 to 265</td>
</tr>
<tr>
<td>Western Jin dynasty</td>
<td>AD 265 to 316</td>
</tr>
<tr>
<td>Eastern Jin dynasty</td>
<td>AD 317 to 420</td>
</tr>
<tr>
<td>Nánbeicháo (Northern and Southern dynasties)</td>
<td>AD 420 to 589</td>
</tr>
<tr>
<td>Súi dynasty</td>
<td>AD 581 to 618</td>
</tr>
<tr>
<td>Táng dynasty</td>
<td>AD 618 to 907</td>
</tr>
<tr>
<td>Five Dynasties period</td>
<td>AD 907 to 960</td>
</tr>
<tr>
<td>Northern Sòng dynasty</td>
<td>AD 960 to 1127</td>
</tr>
<tr>
<td>Southern Sòng dynasty</td>
<td>AD 1127 to 1279</td>
</tr>
<tr>
<td>Liáo dynasty</td>
<td>AD 916 to 1125</td>
</tr>
<tr>
<td>Jīn dynasty</td>
<td>AD 1115 to 1234</td>
</tr>
<tr>
<td>Yuán dynasty</td>
<td>AD 1271 to 1368</td>
</tr>
<tr>
<td>Míng Dynasty</td>
<td>AD 1368 to 1644</td>
</tr>
<tr>
<td>Qīng dynasty</td>
<td>AD 1644 to 1911</td>
</tr>
</tbody>
</table>