

Paddy Fields: A Topological Description of Chinese Word Order¹

Pierre Magistry

CLCLP, TIGP Academia Sinica
Graduate Institute of Linguistics
National Taiwan University
pierre@magistry.fr

Kim Gerdes

LPP (CNRS), Signes (INRIA)
ILPGA
Sorbonne Nouvelle
kim.gerdes@univ-paris3.fr

Abstract

This paper shows how Mandarin Chinese word order can be described in topological terms. After discussing the difficulties in using syntactic dependency and basic word order phenomena of Chinese, we provide the foundations of a formal topological grammar that links a syntactic dependency tree to all of the possible corresponding word orders. We present the formal rules that allow the generation of simple sentences, as well as the more complex *ba* and *bei*-constructions and serial verb constructions.

1 Introduction

This paper is an account of the work in progress regarding a topological description of some basic word order phenomena of Mandarin Chinese. The topological model (Gerdes Kahane, 2001) is a powerful model for linearization. In the Meaning-Text Theory, it can model the interface between surface syntax and the next level, traditionally called morphological representation, that we refer to as the “topological level”, because we not only construct word order and prosodic breaks of different sizes, but a fully-fledged constituent tree. This tree can then provide the basis for the computation of the prosodic groups and pauses. The topological model is a formalization of the traditional analysis of German (Drach 1937) and has been shown to allow for an elegant description of word order phenomena like scrambling in V2 or verb final languages that mix syntactic and communicative constraints. The basic idea is that the sentence is constructed from fixed places, also called positions or fields, which have different constraints on the number of their occupants. All word order constraints are described in this manner: If a word has to precede another word, we don’t use relative placement rules, but these words are placed into different fields. If their mutual order is free (under the given communicative information), they go into the same field.

At first sight, it may seem like an overkill to apply this model to a language like Chinese, sometimes described as having a very restricted word order because of its limited morphology. Alternatively, Her 2003 describe “inversion constructions” in LFG with “simplified Lexical Mapping Theory”. In this approach, a change in argument order is considered as a lexical operation, putting the full burden on a multitude of lexical entries. However, following (LaPolla 95), we will show that Chinese has a fairly complex word order, depending mainly on communicative constraints. Placing this work in the Meaning-Text Framework allows us to have unordered dependency trees at the syntactic level; the linearization process can then be described in topological terms. All the rules and examples we provide have been implemented and tested with the DepLin software (<http://gerdes.fr/soft/deplin/>), which assures that obtrusive interaction between different rules or surplus word orders have not been overseen. The rules have, however, not been tested in parsing, although this is feasible (for example by transcoding the rules in Lexical Functional Grammar (LFG), Clément et al. 2002). Aside from the difficulty involved in

¹ The authors, not being Chinese natives, are deeply indebted to the helpful comments and innumerable grammaticality evaluations by Hsieh Shu-kai, Liu Yeh-hsin, and Jun Miao. We have also benefited greatly from Sylvain Kahane’s and three anonymous reviewers’ comments on our work. Any shortcomings remaining after help from these colleagues are, of course, entirely our own responsibility.

working on written text when we want to include many oral word orders in our account, we would encounter the word separation and ambiguity problem that most rule-based approaches face when parsing Chinese, often obscuring the underlying analysis of word order phenomena (cf. for example the importance that segmentation takes in the development of the Chinese Lexical Functional Grammar in Fang&King 2007). We believe that this is another example of the usefulness of the prevalence of the synthesis direction in the Meaning-Text Theory. It allows concentrating on the non-coincidental properties of language, while keeping in mind the bidirectional character of the rules provided.

2 Adequateness of the Meaning-Text Model

It is nonetheless an important question to ask whether the Meaning-Text Model provides an appropriate framework for a language like Chinese. The pipeline model with semantic, deep syntactic, and surface syntactic representations needs to be discussed for a language where the usual difference between the semantic representation and the surface syntactic representation does not apply easily: Semantemes become full form words. In Chinese, not only do we lack morphologically-based differences between different categories, for example between nouns and verbs, they commonly keep the same valency in whatever syntactic position the words appear. In the following sentences, *ai*, just as its English translation ‘love’, appears as verb or as a head-noun without any morphological change.

- | | | | |
|-----|-----------------|---|-----------------------------------|
| (1) | 你愛她 | / | 你對她的愛有多少 |
| | nǐ ài tā | | nǐ duì tā de ài yǒu duōshao |
| | you love she | | you to her DE love ~have many-few |
| | ‘You love her.’ | | ‘How deep is your love for her?’ |

Contrary to English, however, where we have morphological tests (changing person, time, and number) for a clear distinction between the two categories, in Chinese, the only observable difference between the two ‘ai’ is the syntactic context, for example the appearance of DE, a genitive particle, when *ai* could be called a noun phrase. More generally, the semantic-syntax interface remains the role to provide function words, appearing on the syntactic level (like *de*, *ba*, and *bei* presented in section 4) and to choose pronouns (or, more often, the absence of pronouns) when realizing predicates.

Yet, the main reason for stipulating doubt on the appropriateness of MTT is the central position this model gives to dependency, including the prominent place of syntactic functions. Although categorical borders may be very different in Chinese (see for example Huang 1997), to our knowledge, nobody doubts the existence of categories as a whole. Things are different with syntactic functions: LaPolla 1993 convincingly shows that the usual criteria for subjecthood or objecthood do not exist in Chinese and argues in favor of a completely semantic and pragmatic analysis of the language, meaning that semantic roles such as agent and beneficiary, coupled with communicative values like topic and focus, are sufficient to describe word order constraints in Chinese. At this time, we cannot discuss whether Chinese has truly grammaticalized the subject role, and it is possible that the term “agent”, even in the surface dependency, would be more appropriate. However, we remain with the usual functional terms *subject* and *object* whenever we have the syntactic realization of an agent in a dependency tree. We will nevertheless use semantically tainted terms like *goal* if a common equivalent for the syntactic relation cannot be found among the usual syntactic functions.

In this approach, we follow the common practice in computational and formal description of Chinese such as the work on a Chinese LFG in the Palo Alto Research Center (Fang&King 2007) or the work of Haitao Liu 2007 on syntactic dependency structures for Chinese, using the “European” terms as function names wherever possible. His work on a Chinese dependency treebank has demonstrated that the dependency approach can give important insights into the structure of the Chinese language.

3 Simple Structures and first formalization

We start our description with a simple dependency structure with a transitive verb:

- (2) 我昨天買了書
 Wǒ zuótiān mǎi le shū
 I yesterday buy ASP book
I bought books yesterday.

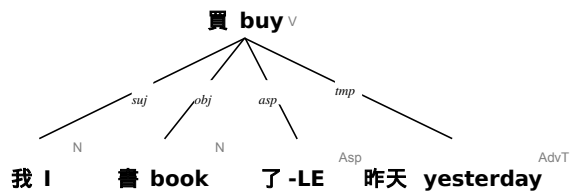


Figure 1: Simple dependency tree

Note that we have the two arguments, the subject and the object, realized as a pronoun and a bare noun. Temporal and spatial relations behave slightly differently than other modifiers and we have to introduce a specific modifier relation, *loc*, which hints further at the close connection between semantic and syntactic relations in Chinese. The aspectual marker LE, marking the accomplishment³ will be treated it as a separate word with a special function: *asp*.

3.1 Topicalization, word order possibilities, and communicative structure

Chinese is said to be an SVO language, which may be misleading considering that S and O functions are potentially irrelevant. LaPolla 95 suggests that that Chinese should be described as “verb medial” language where “Topical or non-focal NPs occur preverbally and focal and or non-topical NPs occur post-verbally.” The typical order given in (2) is in fact the most communicatively neutral, corresponding to Li & Thompson 81 (chapter 4.1.3 D) “sentences with no topic”, i.e. it can constitute an answer to the thetic question: *What is going on?* Topicalization of various dependents of the verb is possible with different communicative structures. Multiple topicalization is possible, too, in particular in spoken language. This can lead to very different word orders for the same dependency structure.

- (3) 這本書我買了 / 昨天書我買了
 Zhè běn shū wǒ mǎi le / zuótiān shū wǒ mǎi le
 this Classifier(Cl) book I buy ASP / yesterday books I buy ASP
This book, I bought (it). / Yesterday, books, I bought (some).

Li & Thompson describe this possibility for *shu* (book) to be in the topic position as in (3). They also remark that the topic position cannot be occupied by indefinite NP and that the interpretation of bare nouns is constrained to be either definite or generic. Interestingly enough, we should add that when in object (post-verbal) position, a bare noun is either generic or indefinite but cannot be interpreted as definite. These differences of possible interpretations seem to be closely related to the communicative value born by the bare noun.

- (4) 我買了書 / 書我買了
 wǒ mǎi le shū / shū wǒ mǎi le
 I buy ASP book / book I buy ASP
 'I bought (a) book(s)' / 'The book, I bought it' or 'Books, I bought'
 *'I bought the book' / *'A book, I bought' or *'I bought a book.'

³ It is generally agreed upon the fact that Chinese has two different markers LE, the other type is called “Current relevant state” (CRS) which always has to go in the last position in the sentence. This place is the last position of our micro domain. LE is sometimes designated as a verbal suffix or as an auxiliary, the lack of segmenting characters making those two explanations plausible.

Note that our analysis differs slightly from Li&Thompson's presentation of Chinese simple declarative sentences. We allow the subject to be placed in a topical position, creating a different constituent structure, whereas Li&Thompson talk about “Sentences in Which the Subject and the Topic are Identical” vs. “Sentences with no subject”. The difference lies in the definition of the subject position, the one they give making it impossible to distinguish those two positions when the topic is an agent. We consider, although we cannot show this here, that the communicative difference also appears prosodically, and we capture this kind of (spoken) word order possibility by allowing more than one element in topic position.

The aspectual marker *le* occupies a position in close proximity to the verb, from which it can only be separated by a verbal resultative (in so called Verb-Resultative compounds) or a specific kind of object in Verb-Object compounds, which are collocational or idiomatic and thus lexically constrained. (5) is an example of a Verb-Object Compound where the bare noun *fàn* can appear before the aspect marker, but it can also be topicalized or appear after *le* as in (6). In the latter cases, *fan* could also have dependents that would specify the meal. This is not possible when *fan* occupies the position between *chi* and *le* where it can only appear as a bare noun.

(5) 我吃飯了

Wǒ chīfàn le
I eat meal ASP
I ate.

(6) 我吃了飯

wǒ chī le fàn
I eat ASP meal
I ate.

/

飯我吃了

fàn wǒ chī le
meal I ate ASP

I've eaten (more like “lunch, I already had”).

For the dependency tree presented above, the topicalization possibilities amount to 8 different word orders (of the 120 theoretically possible orders). They correspond to 16 different communicative structures, which reflect different possibilities for the intonation structure in spoken language.

3.2 Domains and placement rules

Topological grammars can include communicative constraints directly in the rules. In this work, however, we provide a grammar that gives all the possible word orders, independently of the communicative partition, but it is straightforward to specialize the proposed rules with communicative restrictions. The terms we use for the description of these possibilities stem from the syntactic description of oral French (Blanche-Benveniste 1990) where we distinguish the “macrosyntactic” domain providing places for all extraction and topicalization phenomena from a core syntax, called “microsyntax”, with the common order constraints and places for all verbal arguments (used when the arguments are rhematic). Moreover, we consider that Chinese verbs provide places for some of its closer dependents. We call this the “verbal domain”.

The macrosyntactic domain only has two fields: The thema-field and the main field. Note that this macrosyntactic division in two main fields roughly corresponds to Chao 1968's description of the Chinese clause structure as simply topic and comment. The micro domain distinguishes four places to express the ordering constraints: subject field, verbal field, object field, and SVC field. The verbal domain has the following fields: circ(umstantial) field, ba-bei-field, negative field, verbal field, verbal object field, and the field for the aspectual (marker). We obtain the following *domain descriptions* including the placement constraints for each field:

Macro domain:	macro-d = Topic* Micro-field
Micro domain:	micro-d = subject? verbal! object? Svc? CRS?
Verbal domain:	verbal-d = circ* neg? ba-bei? verb! v-obj? Asp?
Nominal domain (simplified):	nd = dem? Num? Cl? atr* noun!

In order to provide places for their (direct or indirect) dependents, words will open these domains under certain conditions, given in the *domain creation rules*: We describe the domain creation rules as a tuple (original field, category, communicative value, domain to be created, final field). A dependant word can occupy an existing position under conditions called the *placement rules*: They can depend on the following values: (the governor’s category, the governor’s communicative value, the governor’s field, the syntactic relation between governor and dependant, the dependent’s category, the dependent’s communicative value, the field where the dependent can go into)⁴.

Below we present the complete rule set needed for the description of the word orders of the examples of section 3. The verb at the root of the dependency tree is placed first and will follow these box creation rules, and the following placement rules may apply to his nominal dependents. (Communicative values would have to be defined at this step):

Initial field	Category	Domain created	Final field	Governor POS	Governor's field	relation	Dependent POS	Dependent field	comment
I	V	macro-d	Micro-field	V	verb	subj	N	subject	Neutral subject
Micro-field	V	micro-d	verbal	V	verb	obj	N	object	Neutral object
verbal	V	verbal-d	verb	V	verb	suj	N	Topic	Topicalized subject
				V	verb	obj	N	Object	Topicalized object

The aspect marker LE will be placed by the first rule, the nominal dependents by the following two rules, and the temporal adverbial by the remaining rules:

Governor POS	Governor's field	relation	Dependent POS	Dependent field	comment
V	verb	asp	ASP	Asp	Aspect marker
N	noun	atr	CL	Cl	Classifier in NP
CL	Cl	qc	Num	num	Numeral in NP
V	verb	loc	AdvT	circ	circumstantial
V	verb	loc	AdvT	Topic	Topicalized circumstantial

Then the nominal dependents may open nominal domains in various positions:

Initial field	Category	Domain created	Final field
Subject	N	nd	noun
Object	N	nd	noun
Topic	N	nd	noun

These rules suffice to compute the different linear structures for the above dependency tree. We show here the topological structure for a simple topicalization of “book”, corresponding to sentence 3 below:

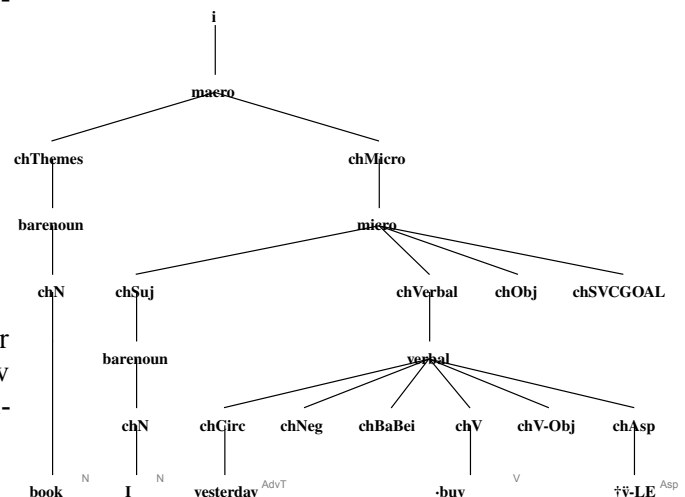


Figure 2: A topological tree = structure 3 below

⁴ Topological grammars can also control extraposition by a¹ simplified extract of the full grammar.

The complete list of all the analyses is given below. Note that we can have different topological structures for the same word order. This means that an analysis of written text using these rules will find for one sentence up to three topological trees, corresponding to the same dependency tree (a *topological ambiguity*). In the generation approach going all the way to sound output, however, these structures are essential for the computation of the prosodic structures.

15. $[macro \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fThemes} 昨天 yesterday \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

Number of structures with the same word order: 1

1. $[macro \textit{fMicro} [micro \textit{fSuj} [barenoun \textit{fN} 我 I] \textit{fVerbal} [verbal \textit{fCirc} 昨天 yesterday \textit{fV} 買 buy \textit{fAsp} 了 -LE] \textit{fObj} [barenoun \textit{fN} 書 book]]]]$

6. $[macro \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fCirc} 昨天 yesterday \textit{fV} 買 buy \textit{fAsp} 了 -LE] \textit{fObj} [barenoun \textit{fN} 書 book]]]]$

8. $[macro \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fThemes} 昨天 yesterday \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE] \textit{fObj} [barenoun \textit{fN} 書 book]]]]$

Number of structures with the same word order: 3

13. $[macro \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fCirc} 昨天 yesterday \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

16. $[macro \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fThemes} 昨天 yesterday \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

Number of structures with the same word order: 2

14. $[macro \textit{fThemes} 昨天 yesterday \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

Number of structures with the same word order: 1

2. $[macro \textit{fThemes} 昨天 yesterday \textit{fMicro} [micro \textit{fSuj} [barenoun \textit{fN} 我 I] \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE] \textit{fObj} [barenoun \textit{fN} 書 book]]]]$

7. $[macro \textit{fThemes} 昨天 yesterday \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE] \textit{fObj} [barenoun \textit{fN} 書 book]]]]$

Number of structures with the same word order: 2

4. $[macro \textit{fThemes} 昨天 yesterday \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fMicro} [micro \textit{fSuj} [barenoun \textit{fN} 我 I] \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

10. $[macro \textit{fThemes} 昨天 yesterday \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

Number of structures with the same word order: 2

3. $[macro \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fMicro} [micro \textit{fSuj} [barenoun \textit{fN} 我 I] \textit{fVerbal} [verbal \textit{fCirc} 昨天 yesterday \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

9. $[macro \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fCirc} 昨天 yesterday \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

12. $[macro \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fThemes} 昨天 yesterday \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

Number of structures with the same word order: 3

5. $[macro \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fThemes} 昨天 yesterday \textit{fMicro} [micro \textit{fSuj} [barenoun \textit{fN} 我 I] \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

11. $[macro \textit{fThemes} [barenoun \textit{fN} 書 book] \textit{fThemes} 昨天 yesterday \textit{fThemes} [barenoun \textit{fN} 我 I] \textit{fMicro} [micro \textit{fVerbal} [verbal \textit{fV} 買 buy \textit{fAsp} 了 -LE]]]]$

Number of structures with the same word order: 2

4 The “Ba” and “Bei” constructions

The Ba and Bei constructions have been widely discussed in Chinese linguistics literature. The latter is sometimes referred to as “passive”. Topologically, they share the same position (since they are in complementary distribution), between the negation marker and the verb, and they behave like a preposition, opening a position for a NP. (7) and (8) provide examples. The difference between these two prepositions is that Ba will take a patient as a complement where Bei will take an agent. As we mentioned in Section 2, we consider them to appear at the syntactic level.

(7) 我把那本書買走了

Wǒ bǎ nà běn shū mǎizǒu le

I BA this Cl. Book buy ASP

I bought this book.

(8) 那本書被我買走了
 nà běn shū bèi wǒ mǎizǒu le
 This Cl. Book BEI I buy ASP
I bought this book / this book was bought by me.

Note that a bare noun would have to be interpreted as definite or generic just like topics. In other words, they cannot introduce new information to the discourse. This confirms the idea that new information has to be postverbal. The position of negation adverbs leads us to locate these constructions inside the verbal domain, just between the verb and the negation adverb:

(9) 我沒把那本書買走了
 Wǒ méi bǎ nà běn shū mǎizǒu le
 I have-not BA this Cl book buy
'I did not buy this book.'

(10) *我把那本書沒買走了
 Wǒ bǎ nà běn shū méi mǎizǒu le
 I BA this Cl book have-not buy

An important point is that Ba and Bei cannot be topicalized, neither can the depending NP:

(11) *書我把買走了 / *把書我買走了
 shū wǒ bǎ mǎizǒu le / bǎ shū wǒ mǎizǒu le
 book I BA buy ASP / BA shu I buy ASP

The position for Ba and Bei is opened by the verb and already included in the rules we have presented in section 3.2. Now we need to define their placement rules and their own domain that will hold the dependent NP. We have two domains: bei-d = bei subject and ba-d = ba object

Domain creation and placement rules:

Initial field	Category	Domain created	Final field	Governor POS	Governor's field	relation	Dependent POS	Dependent field
ba-bei	BA	ba-d	ba	V	verb	pat-obj	BA	ba-bei
ba-bei	BEI	bei-d	bei	V	verb	agt-obj	BEI	ba-bei
				BA	ba	comp	N	object
				BEI	bei	comp	N	subject

These additions to our grammar suffice to generate the more restricted word orders: With a 6 words tree, only two different word orders are possible, corresponding to 5 different topological trees (for 720 theoretical possibilities) :

1. $i_{[macro]} [fMicro_{[micro]} fSuj_{[barenoun]} fN_{[barenoun]} 我 I] fVerbal_{[verbal]} fCirc_{[circ]} 昨天 yesterday fBaBei_{[baBei]} fBa_{[ba]} 把 BA fObj_{[barenoun]} fN_{[barenoun]} 書 book]] fV_{[fv]} 買 buy fAsp_{[fasp]} 了 -LE]]]$
 3. $i_{[macro]} [fThemes_{[barenoun]} fN_{[barenoun]} 我 I] fMicro_{[micro]} fVerbal_{[verbal]} fCirc_{[circ]} 昨天 yesterday fBaBei_{[baBei]} fBa_{[ba]} 把 BA fObj_{[barenoun]} fN_{[barenoun]} 書 book]] fV_{[fv]} 買 buy fAsp_{[fasp]} 了 -LE]]]$
 5. $i_{[macro]} [fThemes_{[barenoun]} fN_{[barenoun]} 我 I] fThemes_{[barenoun]} 昨天 yesterday fMicro_{[micro]} fVerbal_{[verbal]} fBaBei_{[baBei]} fBa_{[ba]} 把 BA fObj_{[barenoun]} fN_{[barenoun]} 書 book]] fV_{[fv]} 買 buy fAsp_{[fasp]} 了 -LE]]]$
Number of structures with the same word order: 3

2. $i_{[macro]} [fThemes_{[barenoun]} 昨天 yesterday fMicro_{[micro]} fSuj_{[barenoun]} fN_{[barenoun]} 我 I] fVerbal_{[verbal]} fBaBei_{[baBei]} fBa_{[ba]} 把 BA fObj_{[barenoun]} fN_{[barenoun]} 書 book]] fV_{[fv]} 買 buy fAsp_{[fasp]} 了 -LE]]]$
 4. $i_{[macro]} [fThemes_{[barenoun]} 昨天 yesterday fThemes_{[barenoun]} fN_{[barenoun]} 我 I] fMicro_{[micro]} fVerbal_{[verbal]} fBaBei_{[baBei]} fBa_{[ba]} 把 BA fObj_{[barenoun]} fN_{[barenoun]} 書 book]] fV_{[fv]} 買 buy fAsp_{[fasp]} 了 -LE]]]$
Number of structures with the same word order: 2

5 Serial Verbs Constructions

Chinese is also known to have Serial Verbs Constructions (SVC) even though it has been convincingly argued that this term in Chinese linguistics subsumes a multitude of different constructions (Paul 2004)⁵.

5.1 Determining the direction of dependency: purpose vs. circumstantial and SVC

It would be impossible to cover all the various phenomena subsumed in the term SVC, therefore, in this paper, we will now focus on the first type described in Li&Thompson 1981, in which the SVC expresses two separate but related events. The same surface form (NP V NP V NP) can lead to four different relations between the two verbs. The relation can be either (i) consecutive, (ii) purpose, (iii) alternating, or (iv) circumstance. Where (i) and (iii) are often ambiguous as well as (ii) and (iv). Paul 2004 argues that not only different interpretations are possible but they should be regarded as different constituent structures with the same surface form. And, following Paul, amongst those different structures, only the purpose relation is a proper SVC. He describes (i) and (iii) as a VP coordination (iv) as an adjunct and (iii) as a proper SVC (ambiguous in surface with (ii). Here we recall Paul's analysis of the SVC:

(12) 我們開會討論這個問題

Wǒmen kāihuì tāolùn zhège wèntí⁶
we hold-meeting discuss this CI problem

- a. Wǒmen [VP [adjunct pro Ø kāihuì] [VP tāolùn zhège wèntí]
'We'll discuss that problem holding a meeting'
- b. Wǒmen [VP kāihuì [purpose clause tāolùn zhège wèntí]]
'We'll hold a meeting to discuss this problem'

5.2 Syntactic structure and additional topological rules

This case can be addressed easily in terms of MTT at a syntactic level. Since the difference is clearly in the relation between the two verbs, we should define two different syntactic relations, in opposing directions, one for the circumstances and one for the purpose, see the two corresponding dependency trees below. We can then define topological rules to account for different linear groupings. By doing this, we pay attention to various constraints on word order that reflect the structural differences:

First, only the matrix verb can take an aspect marker or be negated. In other words, the domains of circumstances and purpose dependents differ and offer a more restricted list of fields than the verbal domain. If the relation between *kaihui* and *taolun* is *circumstance* then we have the word order in (13). If however, the relation is *purpose*, we have (14):

(13) Wǒmen kāihuì tāolùn le zhège wèntí
'We have discussed this problem during the meeting'

*Wǒmen kāihuì le tāolùn zhège wèntí

(14) Wǒmen kāihuì le tāolùn zhège wèntí
'We have held a meeting to discuss this problem'

*Wǒmen kāihuì tāolùn le zhège wèntí

⁵ Among them are some structures that should not be called SVC because they resemble phenomena very common in various languages including languages without SVC, like sentential subjects. Nevertheless, some so-called SVC in Chinese are comparable with structures of African languages well known for their SVC (But even in this case, a close look to characterize structural differences amongst languages is needed, see Wu 2002, Paul 2004)

⁶ We have to note here that when asked about this sentence, some native speakers (of Mandarin spoken in Taiwan) don't even notice the ambiguity (in favor of 17b) or said to have a strong preference for the SVC interpretation.

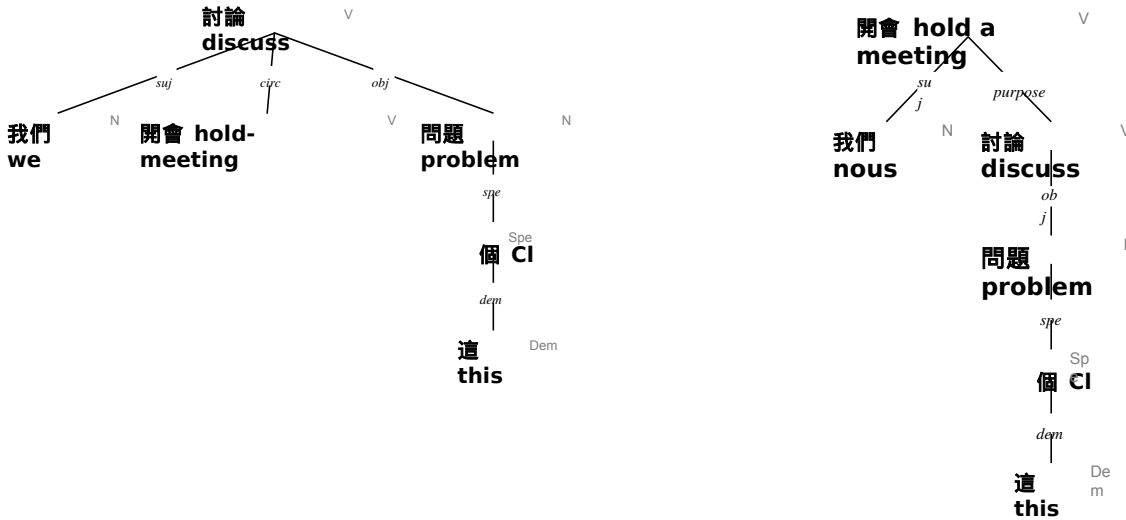


Figure 1: The two dependency trees: circumstantial and SVC

The constraint on the negation is very similar and straightforward.

Second, we have to consider the possible topicalizations that may affect the word order of such structures. In both cases, the NP “zhègè wèntí” ‘this problem’ can be topicalized, but (quite obviously) the matrix verb cannot, neither can the purpose verb (without adding lexical material, like prepositions or verb duplication, and completely changing the syntactic structure). Finally, “kāihui” ‘hold a meeting’ can only be topicalized if it is circumstantial⁷. Our grammar generates all and only these word order possibilities. We only need to introduce the following rules for the SVC reading:

Reduced verbal domain: $r_{vd} = \text{verb! Object?}$ and Domain creation rule $:(\text{SVC}, \text{V}, r_{vd}, \text{verb})$

These additions to our grammar give a different interesting result: Starting with two different dependency structures, we obtain various word orders, some of them common to the two different dependency structures, attesting that the surface form is ambiguous. We also noticed that all the word orders (but none of the topological trees) generated by the SVC dependency tree can be generated from the circumstantial dependency tree, while the contrary does not hold. This observation seems to suit the preferences of our native speaker informants.

Below we show all possible word orders for the first dependency tree (with the circumstantial dependency, 8 word orders):

8. $i_{[...]} [f_{\text{Themes}} [f_{\text{N}} \text{我們 we}] f_{\text{Themes}} [f_{\text{Quantification}} [f_{\text{Dem}} \text{這 this}] f_{\text{Spec}} \text{個}] f_{\text{N}} \text{問題 problem}] f_{\text{Micro}} [f_{\text{Verbal}} [f_{\text{Circ}} \text{開會 hold meeting}] f_{\text{V}} \text{討論 discuss}]]]$
 15. $i_{[...]} [f_{\text{Themes}} [f_{\text{N}} \text{我們 we}] f_{\text{Themes}} [f_{\text{Quantification}} [f_{\text{Dem}} \text{這 this}] f_{\text{Spec}} \text{個}] f_{\text{N}} \text{問題 problem}] f_{\text{Themes}} \text{開會 hold meeting} f_{\text{Micro}} [f_{\text{Verbal}} [f_{\text{V}} \text{討論 discuss}]]]$

1. $i_{[...]} [f_{\text{Micro}} [f_{\text{Suj}} [f_{\text{N}} \text{我們 we}] f_{\text{Verbal}} [f_{\text{Circ}} \text{開會 hold meeting}] f_{\text{V}} \text{討論 discuss}] f_{\text{Obj}} [f_{\text{Quantification}} [f_{\text{Dem}} \text{這 this}] f_{\text{Spec}} \text{個}] f_{\text{N}} \text{問題 problem}]]]$
 6. $i_{[...]} [f_{\text{Themes}} [f_{\text{N}} \text{我們 we}] f_{\text{Micro}} [f_{\text{Verbal}} [f_{\text{Circ}} \text{開會 hold meeting}] f_{\text{V}} \text{討論 discuss}] f_{\text{Obj}} [f_{\text{Quantification}} [f_{\text{Dem}} \text{這 this}] f_{\text{Spec}} \text{個}] f_{\text{N}} \text{問題 problem}]]]$

⁷ Some informants don't accept the topicalization of a bare verb, or find it unnatural. If we add the postposition 時 shí to the verb, however, the verbal topicalization becomes generally acceptable. This particle appears at the syntactic level and can be dealt with a small amendment to our grammar adding a constraint on the topic field. We don't want to stress this point here for clarity reasons.

13. $i_{[micro]} [fThemes_{[base]} fN_{[base]} \text{我們 we}] fThemes_{[base]} \text{開會 hold meeting} fMicro_{[micro]} fVerbal_{[verb]} fV_{[verb]} \text{討論 discuss}] fObj_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}]]]$

Number of structures with the same word order: 3

16. $i_{[micro]} [fThemes_{[base]} fN_{[base]} \text{我們 we}] fThemes_{[base]} \text{開會 hold meeting} fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fMicro_{[micro]} fVerbal_{[verb]} fV_{[verb]} \text{討論 discuss}]]]$

Number of structures with the same word order: 1

2. $i_{[micro]} [fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fMicro_{[micro]} fSuj_{[micro]} fN_{[micro]} \text{我們 we}] fVerbal_{[verb]} fCirc_{[verb]} \text{開會 hold meeting} fV_{[verb]} \text{討論 discuss}]]]$

7. $i_{[micro]} [fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fThemes_{[micro]} fN_{[micro]} \text{我們 we}] fMicro_{[micro]} fVerbal_{[verb]} fCirc_{[verb]} \text{開會 hold meeting} fV_{[verb]} \text{討論 discuss}]]]$

14. $i_{[micro]} [fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fThemes_{[micro]} fN_{[micro]} \text{我們 we}] fThemes_{[base]} \text{開會 hold meeting} fMicro_{[micro]} fVerbal_{[verb]} fV_{[verb]} \text{討論 discuss}]]]$

Number of structures with the same word order: 3

4. $i_{[micro]} [fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fThemes_{[base]} \text{開會 hold meeting} fMicro_{[micro]} fSuj_{[micro]} fN_{[micro]} \text{我們 we}] fVerbal_{[verb]} fV_{[verb]} \text{討論 discuss}]]]$

10. $i_{[micro]} [fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fThemes_{[base]} \text{開會 hold meeting} fThemes_{[micro]} fN_{[micro]} \text{我們 we}] fMicro_{[micro]} fVerbal_{[verb]} fV_{[verb]} \text{討論 discuss}]]]$

Number of structures with the same word order: 2

3. $i_{[micro]} [fThemes_{[base]} \text{開會 hold meeting} fMicro_{[micro]} fSuj_{[micro]} fN_{[micro]} \text{我們 we}] fVerbal_{[verb]} fV_{[verb]} \text{討論 discuss}] fObj_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}]]]$

9. $i_{[micro]} [fThemes_{[base]} \text{開會 hold meeting} fThemes_{[micro]} fN_{[micro]} \text{我們 we}] fMicro_{[micro]} fVerbal_{[verb]} fV_{[verb]} \text{討論 discuss}] fObj_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}]]]$

Number of structures with the same word order: 2

12. $i_{[micro]} [fThemes_{[base]} \text{開會 hold meeting} fThemes_{[micro]} fN_{[micro]} \text{我們 we}] fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fMicro_{[micro]} fVerbal_{[verb]} fV_{[verb]} \text{討論 discuss}]]]$

Number of structures with the same word order: 1

5. $i_{[micro]} [fThemes_{[base]} \text{開會 hold meeting} fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fMicro_{[micro]} fSuj_{[micro]} fN_{[micro]} \text{我們 we}] fVerbal_{[verb]} fV_{[verb]} \text{討論 discuss}]]]$

11. $i_{[micro]} [fThemes_{[base]} \text{開會 hold meeting} fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fThemes_{[micro]} fN_{[micro]} \text{我們 we}] fMicro_{[micro]} fVerbal_{[verb]} fV_{[verb]} \text{討論 discuss}]]]$

Number of structures with the same word order: 2

This is a list of structures obtained for the second dependency tree (with the SVC, 3 word orders):

5. $i_{[micro]} [fThemes_{[micro]} fN_{[micro]} \text{我們 we}] fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fMicro_{[micro]} fVerbal_{[verb]} fV_{[verb]} \text{開會 hold a meeting}] fSVCPURP_{[verb]} fV_{[verb]} \text{討論 discuss}]]]$

Number of structures with the same word order: 1

1. $i_{[micro]} [fMicro_{[micro]} fSuj_{[micro]} fN_{[micro]} \text{我們 we}] fVerbal_{[verb]} fV_{[verb]} \text{開會 hold a meeting}] fSVCPURP_{[verb]} fV_{[verb]} \text{討論 discuss} fObj_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}]]]]$

3. $i_{[micro]} [fThemes_{[micro]} fN_{[micro]} \text{我們 we}] fMicro_{[micro]} fVerbal_{[verb]} fV_{[verb]} \text{開會 hold a meeting}] fSVCPURP_{[verb]} fV_{[verb]} \text{討論 discuss} fObj_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}]]]]$

Number of structures with the same word order: 3

2. $i_{[micro]} [fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fMicro_{[micro]} fSuj_{[micro]} fN_{[micro]} \text{我們 we}] fVerbal_{[verb]} fV_{[verb]} \text{開會 hold a meeting}] fSVCPURP_{[verb]} fV_{[verb]} \text{討論 discuss}]]]]$

4. $i_{[micro]} [fThemes_{[obj]} fQuantification_{[obj]} fDem_{[obj]} \text{這 this} fSpec_{[obj]} \text{個} fN_{[obj]} \text{問題 problem}] fThemes_{[micro]} fN_{[micro]} \text{我們 we}] fMicro_{[micro]} fVerbal_{[verb]} fV_{[verb]} \text{開會 hold a meeting}] fSVCPURP_{[verb]} fV_{[verb]} \text{討論 discuss}]]]]$

Number of structures with the same word order: 2

6 Conclusion

We have shown that various simple and more complex syntactic phenomena of Chinese find a straightforward formalization in terms of dependency and topology, and thus in the framework of MTT. In spite of some doubts on the usefulness of the commonly used syntactic functions, it is possible to translate into this type of topological formalization some analyses of syntactic phenomena stemming from different theoretical frameworks, even from “distant” approaches like generativist theories. Contrary to analysis based reasoning that focuses on ambiguities, we believe that this “synthetic” approach explains naturally the underlying linguistic processes. Our approach differs thus in providing the complete set of paraphrases for a

given dependency tree, a computation that, as soon as we go beyond the simple examples given in this paper, requires the implementation of the grammar in a computer system.

Our grammar includes some more complex phenomena like for example relative phrases, not presented here for lack of space, and we are working on covering further syntactic details. It would be interesting to explore the connection of this grammar with an implementation of a semantic-syntax interface that could provide the input for our system. On the other end of the pipeline, it remains to be shown that the resulting topological structures have a *raison d'être* in providing a smooth basis for the computation of prosodic groups even for tone language like Chinese.

References

- Blanche-Benveniste C. 1990, *Le Français Parlé: Etudes Grammaticales*, CNRS, Paris.
- Chao, Yuen Ren. 1968. *A grammar of spoken Chinese*. Berkeley & Los Angeles: University of California press.
- Clément, L., K. Gerdes and S. Kahane. 2002. "An LFG-type grammar for German based on topological model", in Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG02 Conference*. CSLI Publications, 2002, Stanford
- Drach, Erich. 1937. *Grundgedanken der deutschen Satzlehre*, Diesterweg, Frankfurt/M.
- Fang, Ji and King, T.H. 2007. "An LFG Chinese Grammar for Machine Use". In. T.H. King and E. M. Bender, eds., *Proceedings of the GEAF07 Workshop*.
- Gerdes K., S. Kahane. 2001. "Word Order in German: A Formal Dependency Grammar Using a Topological Hierarchy" in: *Proceedings ACL 2001*, Toulouse
- Gerdes K., S. Kahane. 2007. "Phrasing it differently", in L. Wanner (ed.), *Selected lexical and grammatical issues in the Meaning-Text Theory*, Benja-mins, 297-335.
- Her, One-Soon 2003. Chinese inverted constructions within a simplified LMT. *Journal of Chinese Linguistics*, monograph series 19 Lexical-Functional Grammar Analysis of Chinese
- Huang, Chu-Ren. 1997. Corpus on web: Introducing the first tagged and balanced Chinese corpus. In: *Proceedings of the Annual Conference of the Pacific Neighborhood Consortium*.
- Li, Charles N, & Thompson, Sandra A.1981. *Mandarin Chinese: A functional Reference Grammar*. University of California press.
- Liu, Haitao. 2007. "Dependency Relations and Dependency Distance: a statistical view based on Treebank". *Proceedings of the Third International Conference on Meaning Text Theory (MTT)*, Klagenfurt, Austria.
- LaPolla, Randy J. 1993. "Arguments against 'subject' and 'direct object' as viable concepts in Chinese", in *Bulletin of the Institute of History and Philology* 63.4:759-813.
- Mel'čuk, Igor A. 1988. *Dependency Syntax: Theory and Practice*. SUNY Press, Albany, NY.
- Mel'čuk, Igor A. 2001. *Communicative Organization in Natural Language: The Semantic-Communicative Structure of Sentences*. John Benjamins, Amsterdam.
- Paul, Waltraud 2005. The "serial verb construction" in Chinese: A Gordian knot. In. Oyharçabal B. & Paul W. 2005. *Proceedings of the workshop La notion de « construction verbale en série » est-elle opératoire ?* December 9, 2004, Ehess, Paris.
- Wu, Ching-huei Teresa 2002. Serial Verb Construction and Verbal Compounding. In Sze-Wing Tang & Chen-Sheng Luther Liu, eds., *On the Formal Way to Chinese Languages*. CSLI, Stanford, California.