

# Chinese Empty Category in Generalized Phrase Structure Grammar

Horng-Ming Wu

## Abstract

This paper presents an examination, within the framework of Generalized Phrase Structure Grammar (henceforth GPSG), provided by Gazdar, Klein, Pullum and Sag (1985), of the structure of sentences with zero subject or/and zero object in Chinese and, furthermore, to compare the similarities and differences between the syntactic operation of certain empty categories in Chinese and English. I will argue that GPSG fails to adequately give an explanation to the structure of sentence with zero subject or/and zero object and topicalized sentences in Mandarin, and that to effectively deal with these constructions we need to modify the relevant parts of the grammar.

**Key words:** Empty category, Generalized Phrase Structure Grammar, Topicalization,

# Chinese Empty Category in Generalized Phrase Structure Grammar

## 1. Empty Categories in Mandarin Chinese

In Mandarin, a sentence may have no subject, no object, or no subject and object. In convention, these empty categories are understood. We can often tell what the empty categories refer to from their contexts. For example, in a situation that a son tells his mother that he plans to go to Taipei tomorrow, the sentence may be sounded as in (1) below.

- (1) wo    mingtian    dasuan    qu    taibei.  
I    tomorrow    plan    go    Taipei  
“I plan to go to Taipei tomorrow.”

In response, his mother may have the following possibilities, as shown in (2).

- (2) a. Sentence with no subject  
(ni)    qu    taibei    duojiu?  
(you)    go    Taipei    how long  
“How long do (you) plan to stay there?”
- b. Sentence with no object  
niziji    qu    (taibei)    ma?  
yourself    go    (Taipei)    Q-mark  
“Do you go to (Taipei) alone?”
- c. Sentence with no subject and object  
(ni)    buyao    qu    (taibei).  
(you)    don't    go    (Taipei)  
“(You)’d better not go to (Taipei).”

In general, native speakers of Mandarin Chinese are able to tell what the missing elements are without having a context. Expressions of these kinds are not uncommon in the language.

## (3) a. Sentence with no subject

(youchai) song xin lai le.

(postman) deliver letter come CRS((Currently Relevant State)

“(The postman) delivers letters to us.”

## b. Sentence with no object

wo zhong (dajiang) le.

I win (big prize) PFV(perfective aspect)

“I hit (a jack pot).”

## c. Sentence with no subject and object

(women) shu (qiu) le.

(we) lose (ball game) PFV

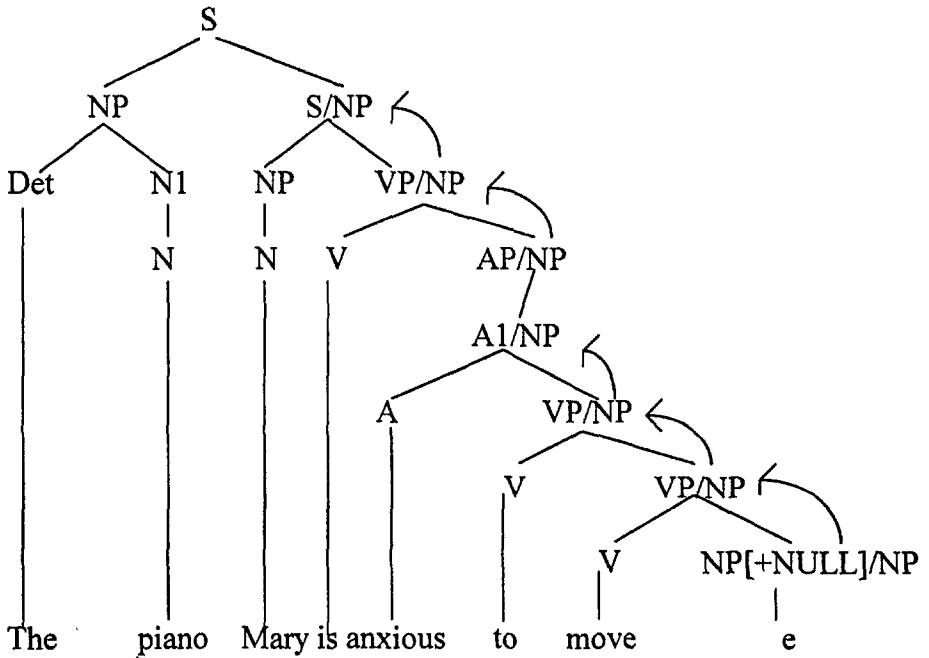
“(We) lost (the game).”

GPSG uses a feature NULL to encode that a constituent is phonologically empty, nevertheless it has to trigger the FOOT feature SLASH by Feature Cooccurrence Restrictions 19. Empty categories like the above Chinese examples do not seem to act the same role as the feature NULL in GPSG framework. Before giving a further characterization to the empty categories in question, next section will first briefly review how GPSG treats empty categories.

## 2. Empty Categories in GPSG

The term NULL in GPSG framework is lexically the same as the term Empty Category in Government-Binding Theory, though they may not function exactly the same way. GPSG, following GKPS (1985), has the use of the feature NULL to put it into a code that an element is phonologically empty. That is to say, there is a one-to-one match relation between a moved phrase and the trace. This idea can be exemplified diagrammatically in (4):

(4)



It is simply the case that the feature NULL has to trigger the FOOT feature, SLASH, by Feature Cooccurrence Restriction 19, as in (5):

(5) FCR 19: [+NULL]  $\supset$  [SLASH]

FCR 19 implicitly says that if something is null, then it has to introduce the FOOT feature SLASH. Thus, the appearance of SLASH, which, being a FOOT feature, will require that the SLASH feature instantiated on a daughter be also instantiated on the mother; as a result, there will be a going up direction in the tree. This 'path' can be seen clearly in (4).

It is necessary to explain some terminology here. In the GPSG framework, features are used to pass information around the tree. There are such features as CASE, COMP, NULL, CONJ, GER, POSS, ...etc. There are also HEAD features and FOOT features, as in (6).

(6) a. HEAD features: AGR, ADV, SLASH, PFORM, AUX,  
SUBCAT, ...etc.

b. FOOT features: RE, SLASH, WH.

If a feature is a HEAD feature, then it must conform the Head Feature Convention (HFC), which is stated as follows (GKPS 1985:26).

(7) HFC: The head features on a mother category are the same as the head features on any daughter which is a head.

Similarly, if a feature is a FOOT feature, it must conform the Foot Feature Principle (FFP), which says that:

(8) FFP: A feature that is instantiated on a daughter is also required to be instantiated on the mother.

The Feature Cooccurrence Restrictions (FCR) is used to define the possible extensions of categories. For example,

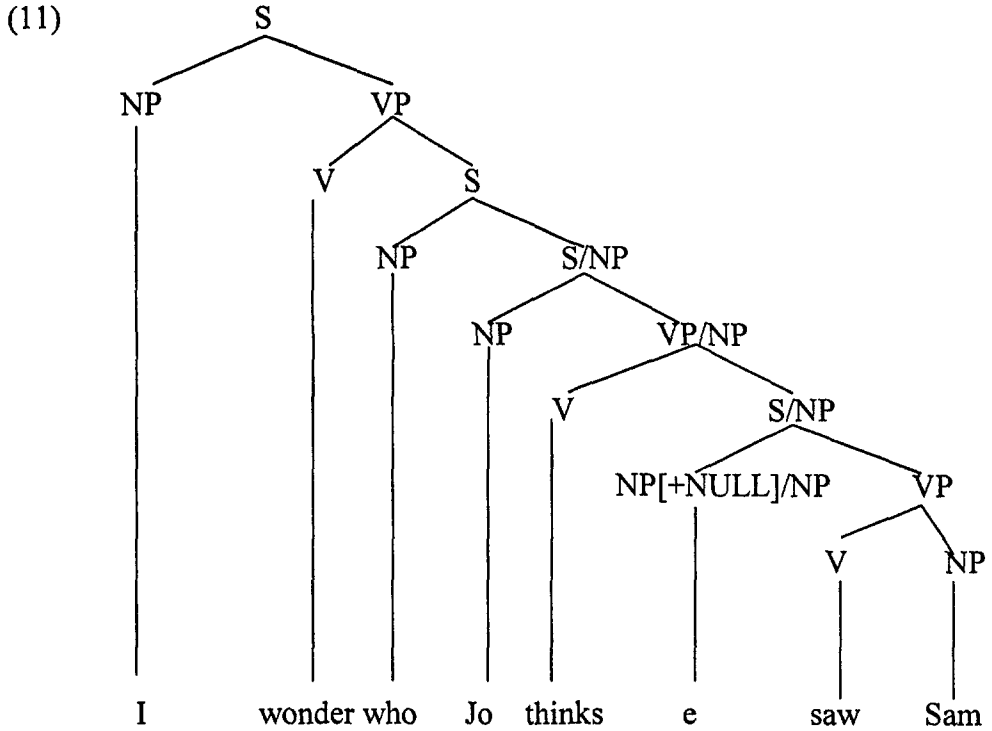
(9) FCR 2: [VFORM]  $\supset$  [+V, -N]

By FCR 2, the VFORM values are limited to verbs only. In other words, any category with a VFORM feature specification must be a verb.

According to GKPS, the feature NULL defaults to being absent, so it can only appear in a tree when sanctioned by a rule. Such a rule will be derived by the application of Slash Termination Metarule 1 (STM 1) to a non-lexical ID rule, if NULL is in subject position. (10) is an example.

(10) I wonder who Jo thinks e saw Sam

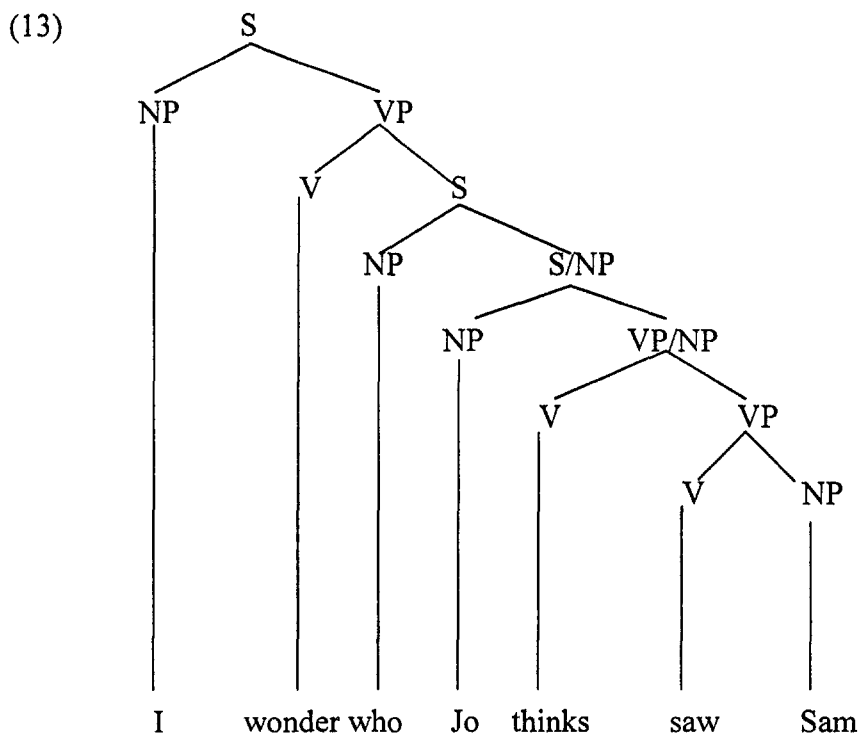
GPSG, however, cannot assign the structure of (10) in (11) below.



Since, in GPSG, a subject can only be introduced by a non-lexical ID rule, a null subject needs to be derived by the application of STM 1. However, there is a constraint on the application of STM 1: the Lexical Head Constraint, which states the following restriction on metarule application (GKPS 1985, P. 59).

(12) Metarules map from lexical ID rules to lexical ID rules.

Therefore, the GPSG analysis of structure (10) should be corrected as follows.



This analysis also shows that a zero pronoun (*ie pro*, in terms of GB) in GPSG is not allowed (or, more straightforwardly, GPSG fails to handle such structure).

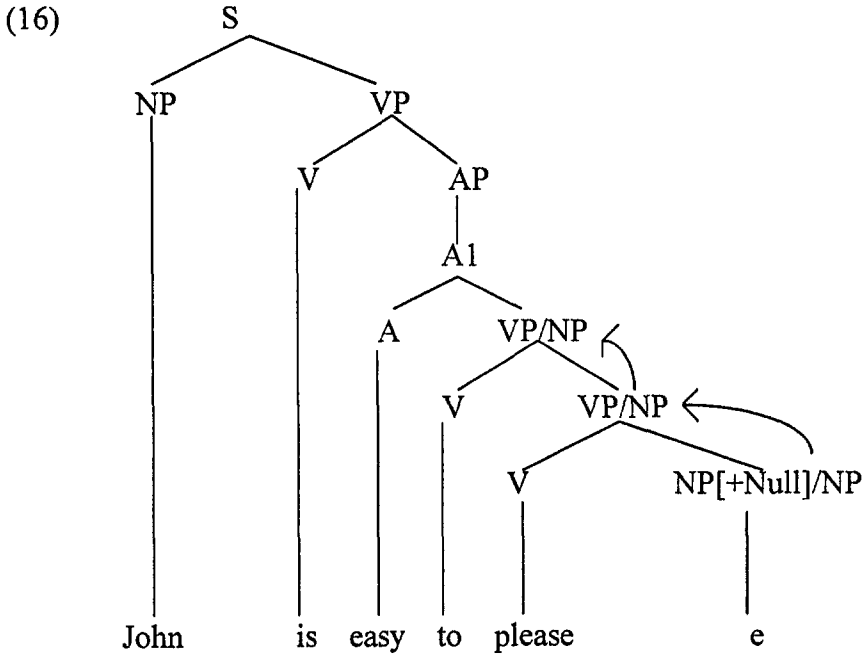
The analysis of missing-object constructions in GPSG is based on the following rule, shown in (14).

(14)  $A1 \rightarrow H[42], V2[INF]/NP[Case\ Acc]$

This rule permits some adjectives to have a slashed infinitive VP or S complement. For example, the rule can account for sentences such as (15).

- (15) a. John is easy to please  
 b. John is easy for us to please  
 c. John is easy for us to make Joe accept

Following GKPS (1985: 151), sentence (15a), for example, has the structure of (16):



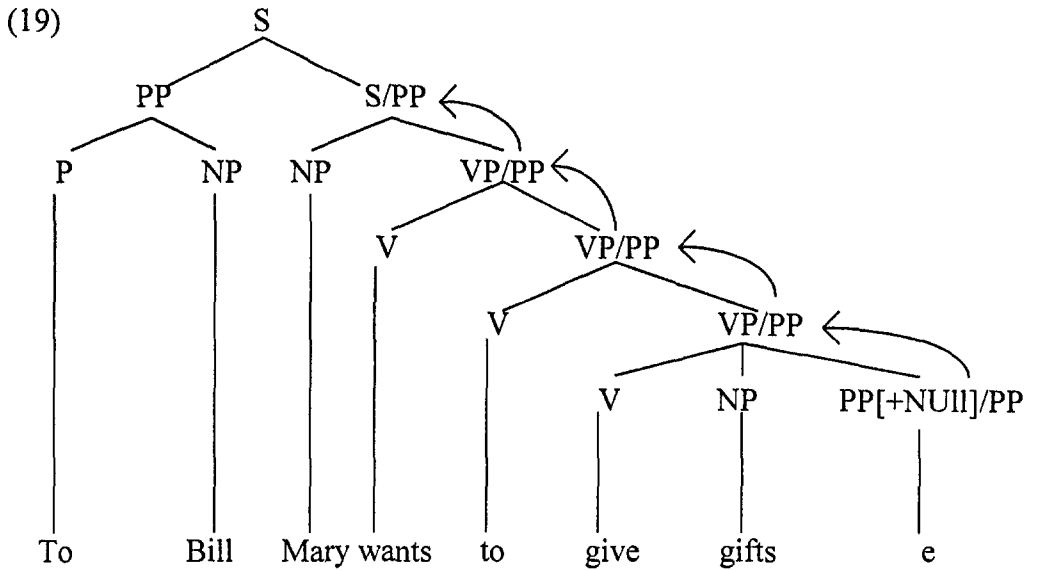
Structure (16) may look like structure (4). Yet, they are two different constructions: the former is a missing-object structure, whereas the latter a topicalization construction. In GPSG, the rule in (17) is responsible for the topicalization.

$$(17) S \rightarrow X_2, H/X_2$$

This rule means that it is possible to have a tree consisting of any Bar<sub>2</sub> category followed by an S which contains a null Bar<sub>2</sub> category. Thus, in a sentence like (18) below, the GPSG analysis of the tree diagram in (19), which represents the structure of (18), may exemplify the difference between the two constructions.

(18) To Bill, Mary wants to give gifts





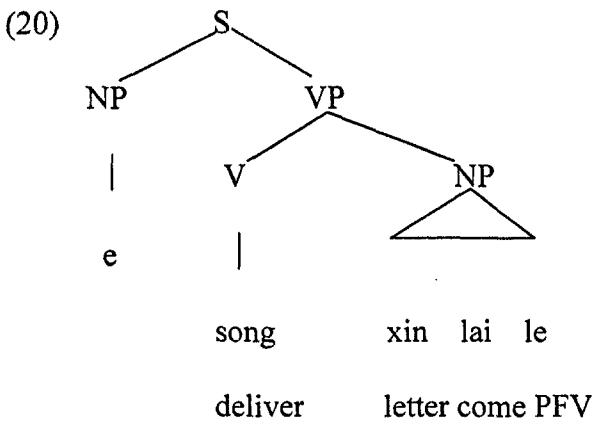
In structure (19), a SLASH feature is projected through the tree from the position of the gap, along the ‘path’ indicated. Basically, each of the two structures requires its own rule to project a tree which is able to reflect each own assumption in GPSG framework. As a result, the projection of a SLASH feature in the tree is different.

To summarize: Apart from using the feature NULL to encode that a constituent is phonologically empty in the structure, the feature SLASH is also required to be taken to be a FOOT feature, and its distribution in trees in accordance with the FFP. These two features play an important role in the analysis of ‘Unbounded Dependencies’ in GPSG framework, since no construction missing an element can be without them. Essentially, the analysis of empty categories in GPSG is always involved in ‘trace’ only.

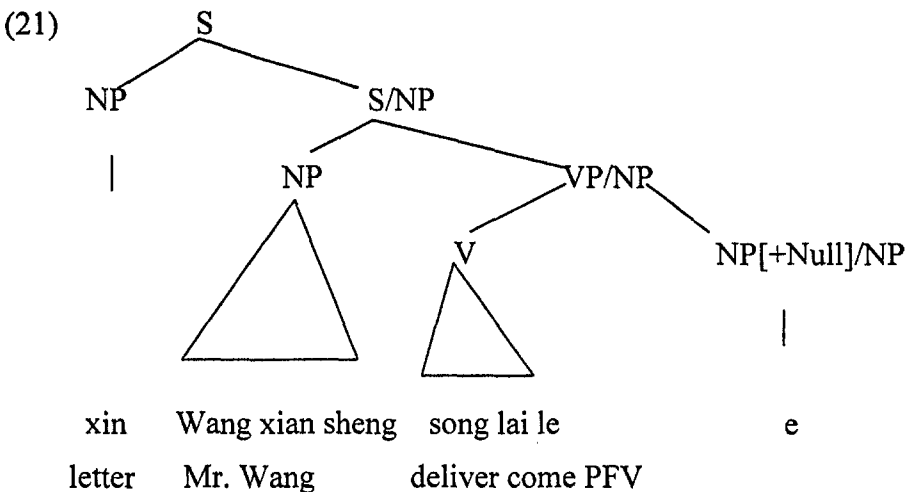
#### 4. Characterizing Empty Category

The term Empty Category in this paper is not the same as the phrase NULL in GPSG framework. As pointed out in the previous section, the feature NULL

encodes that a constituent is phonologically null, often known as a trace. That is to say, there is a one-to-one match relation between an extracted phrase and the trace. However, Empty Category in this work does not refer to a trace: the word simply does not show up. There is no any other extracted phrase to match it. The structures in (20) and (21) exemplify Empty Category and NULL, respectively.



“(The postman) delivers the letters to us.”



“The letter, Mr. Wang delivered to us.”

Thus, the relevant information for defining the empty category is whether or not it is a trace. A definition of empty category in the paper may be given in (22).

(22) In a sentence, if a missing constituent is not a trace, then this empty category has no feature NULL.

What exactly this work assumes is that empty category (not NULL since the empty category is not a trace) does not trigger the appearance of SLASH, which, being a FOOT feature, will require that the SLASH feature instantiated on a daughter be also instantiated on the mother. Therefore, there will be no going up ‘path’ in the tree. The difference can be seen in (20) and (21).

### 5. How Can Empty Category in Chinese be described in GPSG

Since the empty category discussed in this paper is not seen as a trace, it would be interesting to know if GPSG can account for the structure. Recall that there exist three types of empty constructions in Chinese: sentence with no subject, sentence with no object, and sentence with no subject and object, repeated below.

(23) a. Sentence with no subject

(ni) qu taibei duojiu?  
 (you) go Taipei how long  
 “How long do (you) plan to stay there?”

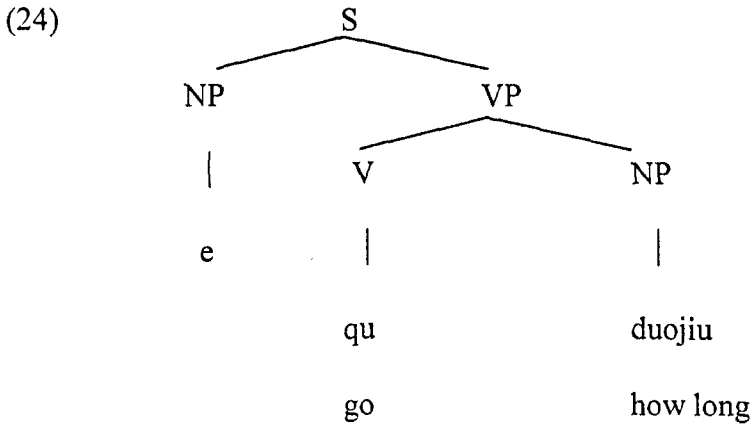
b. Sentence with no object

niziji qu (taibei) ma?  
 yourself go (Taipei) Q-mark  
 “Do you go to (Taipei) alone?”

c. Sentence with no subject and object

(ni) buyao qu (taibei).  
 (you) don’t go (Taipei)  
 “(You)’d better not go to (Taipei).”

Let’s assume that the first type (i.e. (23a)) would have the following structure.



“How long do (you) plan to stay there?”

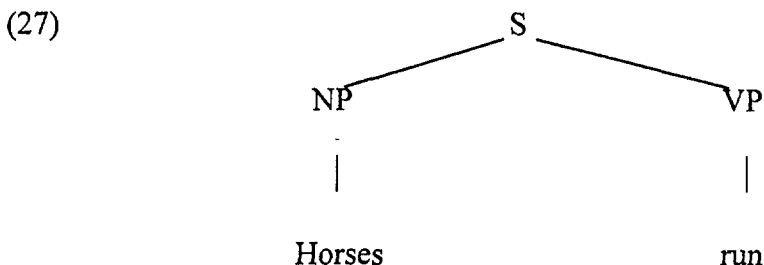
As can be seen, the subject NP position of structure (24) is empty. According to GKPS, in GPSG the feature NULL defaults to being absent; so it can only appear in a tree when sanctioned by a rule, which will be derived by the application of Slash Termination Metarule 1 (henceforth STM1) to a non-lexical ID rule. However, there is a constraint on the application of STM1: Lexical Head Constraint, which states the following restriction on metarule application (GKPS 1985, p. 59).

(25) Metarules map from lexical ID rules to lexical ID rules.

Sine an NP subject in English can only be introduced by a non-lexical ID rule, no ID rules can admit a tree like (24). Hence, it seems to us that none of the existing types of GKPS rules could explain the so-called zero subject sentences in Chinese. By contrast, one may assume that the second type of miss-element sentences, zero object structures, may be accountable, as one of the most striking feature of English structure is that English is a subject-prominent language, a sentence must have a subject, but it is permissible for an English sentence to have no object (without being phonologically realized). Consider example (26).

(26) Horses run.

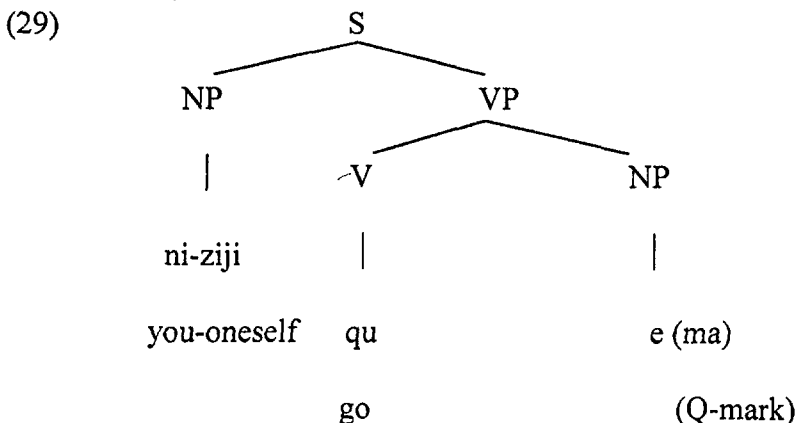
(26) might be diagrammed as follows.



In English, when a verb does not express action that passes over to a receiver (namely, an intransitive verb), a sentence containing the verb can have no object. Sentences with no object in Chinese, however, cannot be treated equally the same as English. Consider sentence (28).

(28) niziji qu (taipei) ma?  
 yourself go (Taipei) Q-mark  
 “Do you go to (Taipei) alone?”

In a like manner, I assume that sentence (28) would have the following tree diagram, as in (29).



“Do you go to (Taipei) alone?”

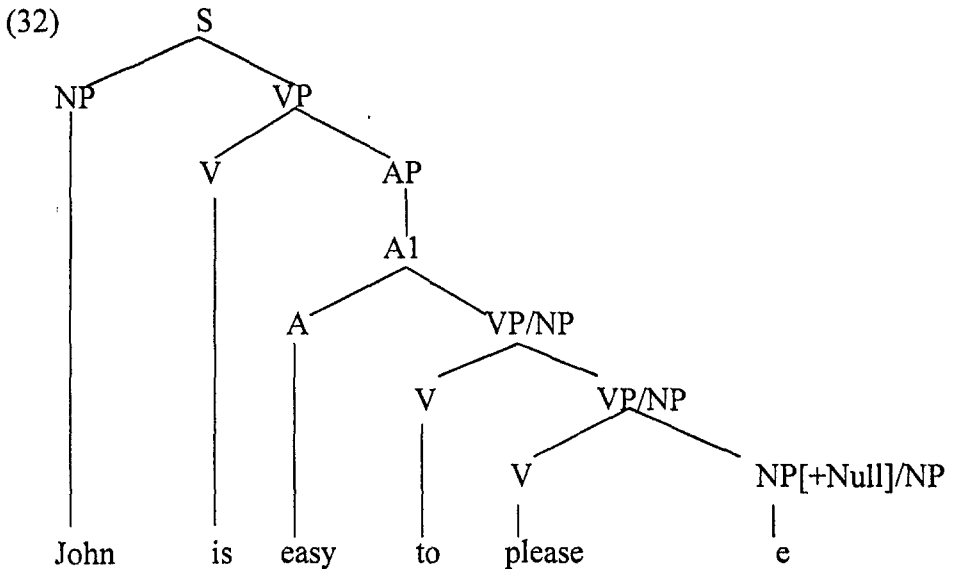
The difference can be seen in (27) and (29). There is an empty category in zero object sentence in the Chinese example, while no such a missing element shown in (27). Nevertheless, this contrast does not imply that English has no missing-object constructions. There exist sentences with no object in English. According to GKPS, their analysis of missing-object constructions is composed of the following rule.

(30)  $A1 \rightarrow H[42], V2[INF]/NP[Case Acc]$

This rule says that it permits some kind of adjectives to have a slashed infinitive VP or S complement. For example, rule (30) can account for a sentence like (31) below.

(31) John is easy to please.

Following GKPS (1985, P. 151), sentence (31) has the structure in (32).



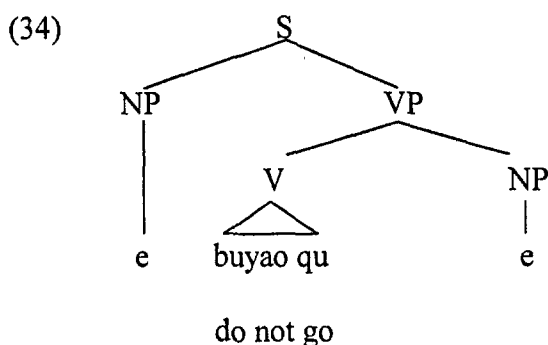
However, in structure (29) we do not see an extracted phrase outside of the clause, whereas within the clause a phrase is correspondingly missing. GKPS's analysis of missing-object constructions, therefore, may not apply to zero object structures

in Chinese, due to the fact that Chinese empty category discussed in this work does not subsume traces.

Turn now to sentences with no subject and object in Chinese, repeated in (33).

- (33) (ni) buyao qu (taibei).  
 (you) don't go (Taipei)  
 “(You)’d better not go to (Taipei).”

Similarly, (33) may be illustrated with structure (34).



“(You)’d better not go to (Taipei).”

Understandably, we would not expect that GKPS’s grammar could present any tenable suggestion to (33), since if they fail to describe sentences with no subject and sentences with no object, respectively, there would be no way for the theory to explain (33).

Consequently, we find that GPSG framework may face one type of problem with missing-something sentences in languages such as Chinese: it would wrongly predict that sentences like 1) *e qu duojiu?* (How long do (you) plan to say there?), 2) *ni ziji qu e ma?* (Do you go to (Taipei) alone?), and 3) *e buyao qu e* ((You)’d better not go to (Taipei).) are ungrammatical. As a matter of fact, these sentences are all perfectly grammatical utterances in Mandarin Chinese, and that they are structurally prominent in Chinese grammar.

## 6. NULL in Mandarin Chinese

We might expect to have a distinct view on the issue. That is, since GPSG cannot provide a favorable account to these empty category cases in Chinese, the framework should be expected to be able to describe the sentences which have the feature NULL in Chinese. Consider examples in (35).

(35) a. \*mali women xiwan e chenggong.

Mary we hope success

“Mary we want to succeed.”

b. \*wotaitai, wo xiang song liwu gei e

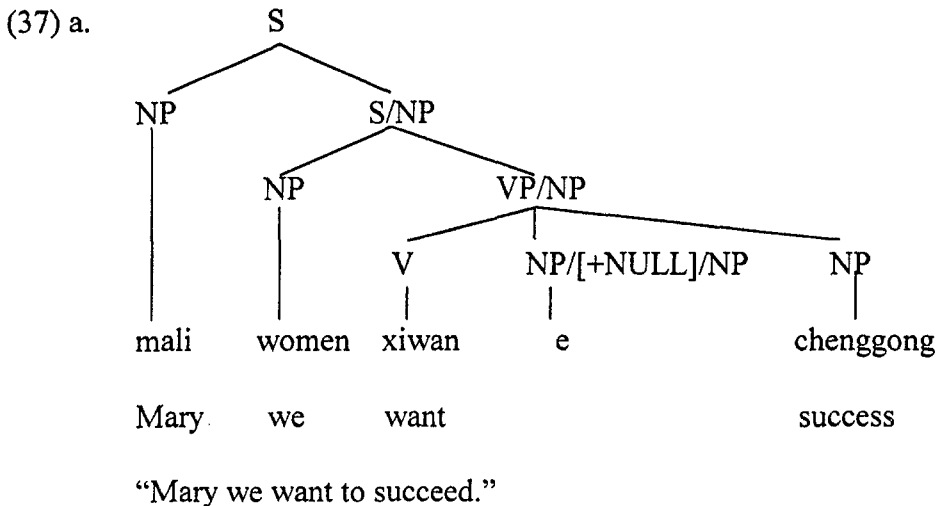
my wife I want give gift to

“My wife, I want to give gifts to.”

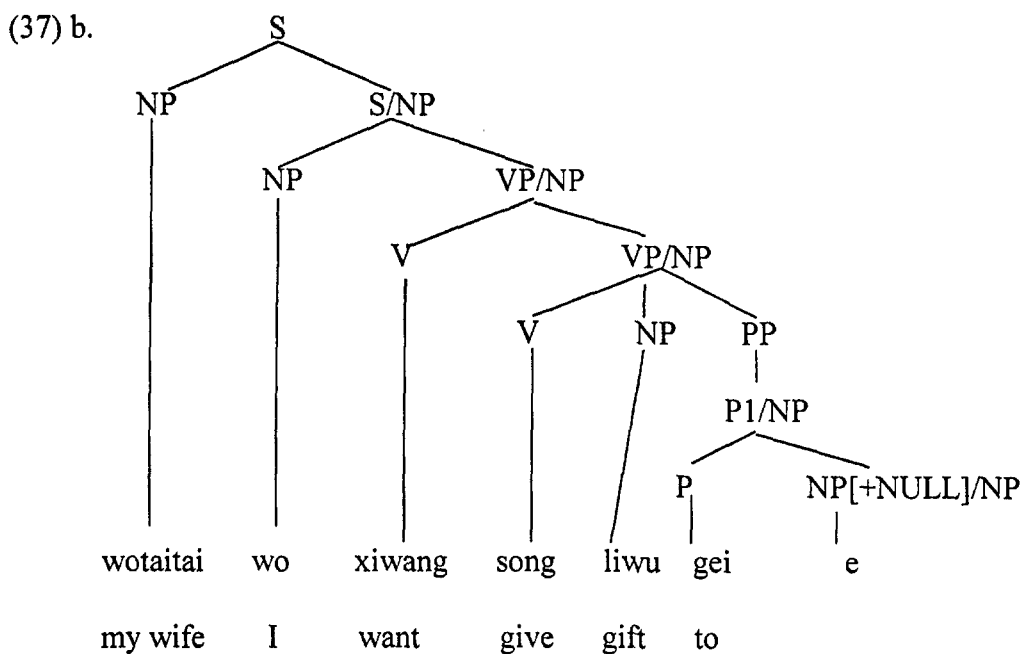
Sentences in (35) are considered topicalization constructions. Notice that in GPSG the rule in (36) is responsible for the topicalization.

(36)  $S \rightarrow X_2, H/X_2$

This rule simply says that it admits a tree that can consist of any Bar2 category followed by an S which contains a null Bar2 category. Thus, following the theory, we may assume that sentences (35a) and (35b) would have the structures in (37a) and (37b), respectively.







“My wife, I want to give gifts to.”

Note that, however, rule (36) can generate not only the sentence structures (37a-b) but also the sentence structures associated with any of the following sentences, as in (38).

(38) a. naben shu, wo chuban le e.  
 that book I publish PFV

“That book, I have published.”

b. chuan, tamen zao hao le e.  
 ship they build finish PFV

“The ship, they have finished building it.”

c. zai yinhang, wo cun le wuqian kuai e.  
 in bank I deposit PFV 5000 dollar

“I deposited five thousand dollars in the bank.”

d. she, ni pa e ma?  
snake you fear Q-mark

“Speaking of snakes, are you scared of them?”

However, it remains to explain why a system of rules can admit both grammatical and ungrammatical sentences, as we may note that sentences (35) are ungrammatical, whereas those in (38) are well-formed.

All of these facts clearly show that the framework of GPSG will incorrectly predict that sentences in (35) are grammatical. More explicitly, the theory does not seem to be able to apply to languages such as Chinese, and this would demand a need to modify its rules. Before taking a step to make any change, let us first see what exactly the problems posed by them.

### 7. Problems Analysis

Recall that we have pointed out that a sentence without a phonologically realized subject or/object is a prominent grammatical feature in Mandarin Chinese. It is the fact that these ‘silent’ elements are understood from contexts (or without contexts, depending on the structure), and that they do not need to be specified. However, this is difficult for speakers of the English language to grasp since there are no rooms for them to be put in the grammar of English. Thus, it is comprehensible that the grammar GKPS have provided may solely apply to English only. In a parallel manner, the problem shown in topicalization constructions (35a-b) is too a salient feature in Chinese. According to the grammar of Mandarin, the ungrammatical sentences (35a-b) should be corrected as in (39), which would then correspond to the respective structures (40a-b).

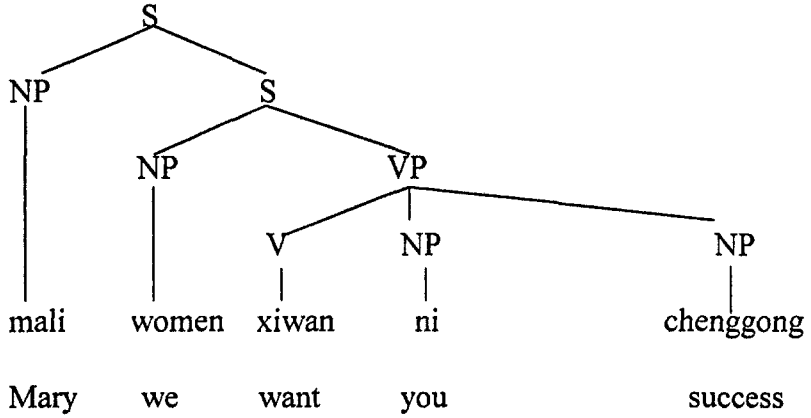
(39) a. mali women xiwan ni chenggong.

Mary we hope you success

“Mary we want you to succeed.”

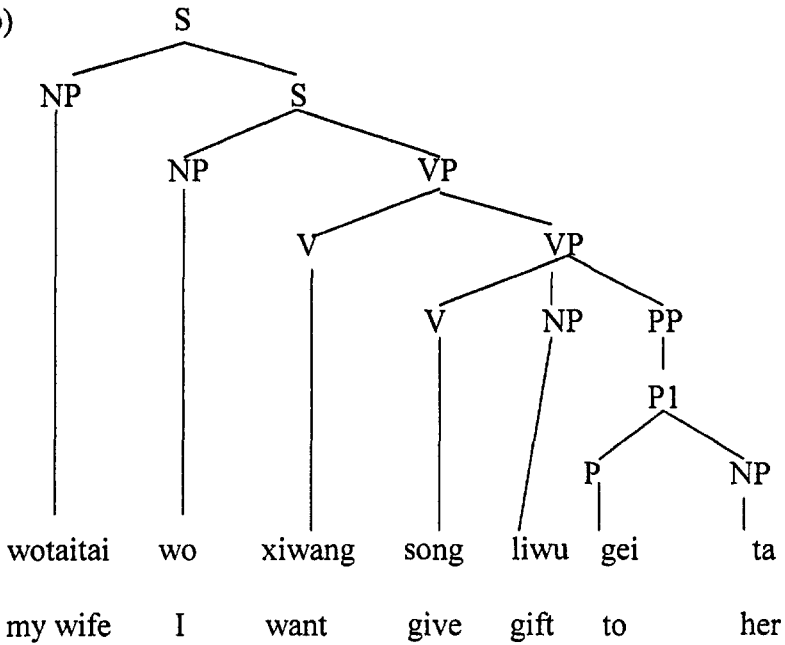
- b. wotaitai, wo xiang song liwu gei ta  
 my wife I want give gift to her  
 “My wife, I want to give gifts to her.”

(40) a.



“Mary we want you to succeed.”

(b)



“My wife, I want to give gifts to her.”

The fact is that in Chinese if a sentence is derived from topicalization, it is needed by the grammar to insert an objective pronoun into that empty site under the condition that the extracted NP and the objective pronoun must match in feature (i.e. number, gender, person, etc.). Sentences (41a-c) explain this operation.

(41) a. *women xiwan mali chenggong.*

*we hope Mary success*

“we want Mary to succeed.”

**Via topicalization:**

b. *mali women xiwan e chenggong.*

*Mary we hope success*

“Mary we want to succeed.”

**Pronoun insertion:**

c. *mali women xiwan ni chenggong.*

*Mary we hope you success*

“Mary we want you to succeed.”

One might argue that sentences in (39) are in fact vocative sentences. But it is justifiable to regard them as being derived from the operations of topicalization and pronoun insertion. The reason is that since *mali*, taking for example (41c), is a topic, it is what the sentence is about. Originally, however, its underlying structure, like many other topicalized constructions, is *women xiwang mali chinggong* (we want Mary to succeed). The only difference is that Chinese grammar will ask that an objective pronoun be inserted into the empty position after the element is transformed to the sentence initial position. One might further argue that if this is the case, then we should expect that the following sentences are grammatical because they are all obtained by topicalization and pronoun insertion.

(42) a. *naben shu, wo chuban le ta.*

*that book I publish PFV it*

“That book, I have published it.”

b. chuan, tamen zao hao le ta.  
 ship they build finish PFV it

“The ship, they have finished building it.”

c. zai yinhang, wo cun le wuqian kuai ta.  
 in bank I deposit PFV 5000 dollar it

“I deposited five thousand dollars in the bank.”

Contrary to what one might presume, the fact, however, is that sentences in (42) are all ungrammatical. They are ill-formed because the rule of objective pronoun insertion requires that the extracted NP be an animate. In Chinese, the pronoun insertion rule can occur only if the condition is met.

In summary, in Chinese after topicalizing an animate NP, a same feature objective pronoun is required to be inserted into the empty position from where the NP moved. Indeed this is one of the most striking features of Mandarin sentence structure, and one that sets Mandarin apart from many other languages, including English. Thus, the main problem with GPSG may be that the theory is unable to give a universal structural account to languages other than English.

## 8. Modification Suggestions

Recall that this study has presented two type of problems for the framework of GPSG: sentences with zero subject or/and zero object and topicalization constructions in Chinese. For giving a favorable account to the former, this study suggests that we need to introduce a new feature in GPSG framework: EMPTY. I will use the feature EMPTY to encode that a constituent is not only phonologically empty but also triggering no the FOOT feature SLASH. Furthermore, this EMPTY feature is assumed to be inserted into the following non-lexical ID rule, as in (43).

(43)  $S \rightarrow X2([+Empty]), H[-Subj]$

That is, this rule admits a tree with any Bar2 category in which the X2 bears an

optional specification [+Empty]. But this will not suffice to solve the whole problem because the revised rule (43) can only deal with zero subject sentences. We therefore need to propose one more rule to explain zero object sentence. (44) is such a suggestion.

## (44) Empty Metarule

$$X \rightarrow W, X_2$$

$$\Downarrow$$

$$X \rightarrow W, X_2[+\text{Empty}]$$

Since the empty category in Chinese can be any phrase (other phrasal categories are also possible), this Empty Metarule will entail that any Bar2 category can be empty. We then can put all the pieces together to see how this revision will work. Those sentences in question will be repeated for convenience. Consider first the zero subject sentence.

(45) e qu taibei duojiu?  
go Taipei how long

“How long do (you) plan to stay there?”

To explain (45), the following rules will take the responsibility.

(46) a.  $S \rightarrow X_2([+\text{Empty}]), H[-\text{Subj}]$

b.  $VP \rightarrow H[2], NP$

Turn now to the zero object sentence.

(47) niziji qu e ma?  
yourself go Q-mark

“Do you go to (Taipei) alone?”

Sentences such as (47) now can be admitted by the rules in (48).

- (48) a.  $S \rightarrow X_2([+Empty]), H[-Subj]$   
 b.  $VP \rightarrow H[2], NP$   
 c. Empty Metarule  
 $VP \rightarrow H[2], NP$   
 $\Downarrow$   
 $VP \rightarrow H[2], NP [+Empty]$

Rules in (48) too admit sentences with no subject and object, as in (49).

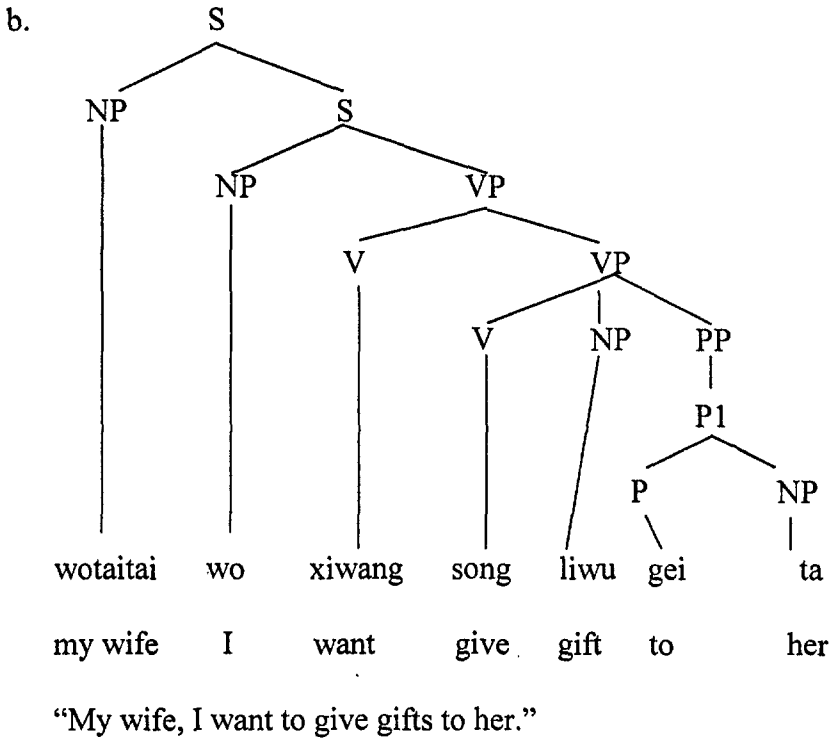
- (49) e buyao qu e.  
 don't go  
 “(You)’d better not go to (Taipei).”

Finally, consider the problem posed by topicalization constructions in Chinese. To solve this problem, I suggest that we revise the non-lexical ID rule in a form shown in (50).

- (50)  $S \rightarrow X_2([+Empty]), H[\alpha Subj]$   
 where  $\alpha$  is in  $\{+, -\}$

This flexible formulation cannot only apply to all the previous discussed cases, it can also explain the problem posed by topicalization in Chinese. For example, sentence (51) can now be described by the revised formulas.

- (51) a. wotaitai, wo xiang song liwu gei ta  
 my wife I want give gift to her  
 “My wife, I want to give gifts to her.”



c. ID rules:

$S \rightarrow X2 ([+Empty]), H[ + Subj]$

$S \rightarrow X2 ([+Empty]), H[ - Subj]$

$VP \rightarrow H[15], VP[INF, +NORM]$

:

:

### 9. Summary

On the basis of facts relating to zero subject sentences, zero object sentences, zero subject and object sentences, and topicalization constructions in Mandarin Chinese, I argue that the grammar GKPS have presented is insufficient to handle them, because these structures are seen as the striking features in Chinese, and that these salient features set Mandarin apart from many other languages, English in



particular. Thus, to expand the grammar's explanation scope, one step to take may be to make slight revision to the grammar so that it can accommodate languages like Chinese. The revised formulas may have the following advantages.

- (52) a. it still keeps its original features  
 b. its explanation capability can now be expanded to include sentences with no subject or/and object and topicalized sentences in Chinese  
 c. it may imply that GPSG is also essential for many other languages  
 d. it reveals that the grammar has great potential to be improved

The revised rules and new features suggested in this work are in (53).

(53) Feature                      Value Range

EMPTY                              { +, - }

Non-lexical ID rule

$S \rightarrow X_2([+Empty]), H[\alpha \text{ Subj}]$

where  $\alpha$  is in { +, - }

Empty Metarule

$VP \rightarrow H[2], NP$

↓

$VP \rightarrow H[2], NP [+Empty]$

## References

- Chao, Yuen-ren (1968) *A Grammar of Spoken Chinese*, Berkeley and Los Angeles: University of California Press.
- Chomsky, Noam (1980) "On Binding," *Linguistic Inquiry* 11: 1-46.
- Chomsky, Noam (1981) *Lectures on Government and Binding*, Dordrecht: Foris.
- Chomsky, Noam (1982) *Some Concepts and Consequences of the Theory of Government and Binding*, Cambridge, Mass.: MIT Press.
- Chomsky, Noam (1986a) *Barriers*, Cambridge, Mass.: MIT Press.
- Chomsky, Noam (1986b) *Knowledge of Language: Its Nature, Origin, and Use*, New York: Praeger.
- Chomsky, Noam (1996) *The Minimalist Program*, Cambridge, Mass.: MIT Press.
- Gazdar, Klein, Pullm and Sag (1985) *Generalized Phrase Structure Grammar*, Cambridge, Massachusetts: Harvard University Press.
- Huang, C-T. James (1982) *Logical Relations in Chinese and the Theory of Grammar*, Ph.D. dissertation, MIT.
- Li, Charles N. and Thompson, Sandra A. (1981) *Mandarin Chinese: A Functional Reference Grammar*, California, University of California Press.
- Soames, Scott & Perlmutter, David (1979) *Syntactic Argumentation and the Structure of English*, Berkeley, University of California Press.

Tie, Hung-Yeh Henry and Lance Donald M. (1990) *A Reference Grammar of Chinese Sentences*, Tucson, The University of Arizona Press.

Y.C. Li, Cheng, L. Robert, Foster Larry, Ho, Shang H., Hou, John Y. Yip, Moira  
*Mandarin Chinese: A Practical Reference Grammar for Students and Teachers (Vol.1)*, Taipei, The Crane Publishing Co.