RESUME


(*) An earlier version of this paper, titled ‘On Coordination Schema and Chinese NP Coordination’, was presented at the 1984 Linguistics Society of America Annual Meeting in Baltimore. I am indebted to Sally McConnell-Ginet for going through all the versions of this paper, and for making most helpful suggestions, to Ivan Sag for introducing the GPSG analysis to me, to Gerald Gazdar, James Huang, Paul Li, Louis Mangion, Marie-Claude Paris, and Annie Zaenen for comments on various versions, and to Carol Rosen and the reviewers of Cahiers de Linguistique Asie Orientale for suggestions which led to improvement of this paper. Responsibility for errors is, of course, mine.

O. INTRODUCTION: CHINESE NP COORDINATION

Three groups of conjunctions used in Chinese NP coordination can be distinguished semantically. The first group marks logical conjunction. It consists of hàn, hé, gēn and other less frequently used variants. The second group marks inclusive disjunction; hùo is the only member. Finally, exclusive disjunction is marked by hūishì. As far as syntactic distribution is concerned, the three groups behave the same. The logical conjunction hàn serves as an example here (1)

1. a. Lǎo-Wáng, Lǎo-Lí, Lǎo-Zhāng
   b. Lǎo-Wáng, Lǎo-Lí, hàn Lǎo-Zhāng
   c. Lǎo-Wáng, hǎn Lǎo-Lí, Lǎo-Zhāng
   d. Lǎo-Wáng, hǎn Lǎo-Lí, hàn Lǎo-Zhāng
   “Lao-Wang, Lao-Li, and Lao-Zhang”

1 shows that conjuncts with no overt marker and conjuncts marked by hàn can occur in almost any order. The only restriction is that no overt marker can precede the first conjunct (2) or follow the last conjunct (3). Coordination with more than three conjuncts also follows this pattern.

There are two possible strategies for analyzing coordination structures. The first is to assign them a flat structure like 2a, the other is to assign them a branching structure such as 2b.

(1) hàn is the most commonly used conjunction for Mandarin Chinese speakers in Taiwan.
(2) Notice that this generalization is only true with Chinese NP coordination. Chinese VP coordination allows overt conjunctions to occur with first conjunct. The NP generalization can be captured in GPSG by a Feature Cooccurrence Restriction. [+N, [CONJ αo]] is part of the feature complex of the first member of an NP coordination, and [CONJ LIL] is part of a feature complex of a conjunct not marked by any overt conjunction. Rule i in effect guarantees that whenever a category is the first member of an NP coordination, it is also a conjunct marked by no overt conjunction.
   (i) [+N [CONJ αo]] ⊃ [CONJ NIL]

(3) This fact was pointed out to me by Paul Li (p. c.). Logically it is possible to have conjunctions following conjuncts, though in fact this phenomenon rarely, if ever, happens in natural languages. I follow the standard assumption in attaching conjunctions to the left of the conjuncts.
In accounting for the Chinese data in 1, flat structures meet with two difficulties. First, the data show that for simple, one level coordination, the same conjunction must be used throughout the string. That is, NPs like 3a should be ruled out as single level coordination. This phenomenon could not be captured if we simply name the node CONJUNCTION, since there is no guarantee that identical lexical items would be inserted every time. One way to solve this problem is to specify idiosyncratically for each conjunction a phrase structure rule, such as 3b.

3  a. *háishi NP bàn NP hē NP
   b. NP → NP bàn NP bàn NP

The second difficulty is that these overt conjunctions can occur before any conjunct except the first one. If we have to capture the occurrence of conjunctions locally in phrase structure rules, as suggested in 3b, we will have to write separate rules for the string marked by an overt conjunction in a second position, and a string not marked there, and so on. For NP coordination structures using the conjunction bàn with four members, there are eight (\(2^{n-1}\), n is the number of members) possible arrangements of the conjunctions. Thus simply to describe the distribution of four-member coordination with bàn adequately, there must be at least eight phrase structure rules in this analysis. This is clearly not a happy solution. The number of phrase structure rules needed increases exponentially with the number of conjuncts, resulting in a rather inelegant grammar.

On the other hand, opting for the branching structure in 2b allows us to reduce the number of phrase structure rules for each conjunction to two, as in 4.
But the branching analysis faces other difficulties. First, this analysis still does not guarantee that the same conjunction will be used throughout the structure. Additional stipulations will be needed to describe the Chinese data adequately. Second, with the readily available distributive reading, there is no semantic grouping to motivate any particular branching structure analysis. Neither is there any evidence for syntactic or phonological sub-groupings. Assuming a homomorphism between syntax and semantics, as many linguists do, there is no reason to follow the branching structure analysis. Third, assigning branching structures to a distributive reading result in unnecessary proliferation of tree structures. All of the eight four-member coordination with hàn, exemplified by 5 here, can receive distributive readings. For each of these eight coordination, there are five possible branching tree structures, exemplified in 6, if we are restricted to the two rules in 4. All these structures should have the identical distributive reading. Five times eight gives us 40, the number of different tree structures for four-member coordination with hàn. It is clearly undesirable to have 40 different tree structures to represent one semantic meaning, and to have to represent one surface syntactic structure with five different constituent trees. Thus the branching structure analysis cannot guarantee using the same conjunction, is not independently motivated, and causes proliferation of tree structures.
c.  
```
NP  
   NP CONJ NP
   NP CONJ NP
   NP CONJ NP
```

Lao-Wang han Lao-Li han Lao-Lin han Lao-Zhang

In this paper, I will try to give a satisfactory account of Chinese NP coordination in Generalized Phrase Structure Grammar (GPSG)\(^{(4)}\), and show that some of the problems just discussed could be solved in this framework given some minor revisions of its formal mechanisms. Another possible GPSG account will also be discussed and shown to have several problems. Therefore, the first analysis will be returned to as the preferred solution. I will start with a brief introduction of the Immediate dominance and Linear Precedence (ID/LP) format used in GPSG, and to the Coordination Schemas developed with the mechanisms of ID/LP format within the framework. This introduction should facilitate my discussion and help readers not familiar with the

\(^{(4)}\) The recent state of GPSG theories and mechanisms is best summarized in Gazdar et al. (1985), to which readers are referred for in-depth discussion of GPSG.
GPSG mechanisms to follow the arguments more easily.

I. **ID/LP Format and The Coordination Schemas**

Instead of the traditional hierarchical phrase structure rules, GPSG employs two sets of statements on well-formedness conditions to admit grammatical strings: Immediate Dominance (ID) statements, and Linear Precedence (LP) statements. ID Statements stipulate which strings of constituents are admissible as the expansions of the designated category, i.e., they specify the mother-daughter relations stipulated in a traditional phrase structure rule. On the other hand, LP statements formalize exhaustively the across-the-board ordering relations among sister constituents. Strings are admissible only if the linear order of sister constituent does not contradict any LP statement. In other words, ID statements list possible expansions of a category while LP statements constrains the possible linear order of the categories in the expansion. An interesting feature of LP statements is that they generalize over the whole language rather than applying just to a specific ID statement as in standard phrase structure rules. For example, if an LP statement in a certain language stipulates that an NP precedes a VP, then a VP can never precede an NP in this language whenever they are sisters. The rules of ID/LP format are illustrated by the mini grammar in 7.

### 7. a. ID Rules
1. \( S \rightarrow A, B, C, \)
2. \( S \rightarrow A, B, \)

### 7. b. LP Statement
1. \( A < C \)

### 8. a. \( S \rightarrow A B C \)
1. \( S \rightarrow A C B \)
2. \( S \rightarrow B A C \)
3. \( S \rightarrow A B \)
4. \( S \rightarrow B A \)

---


(6) Gazdar et al. (1985) use a curly ‘less than’ sign ‘\(<\)’ in LP statements to indicate that it is a somewhat abstract precedence relation. They use the regular ‘less than’ sign ‘\(<\)’ exclusively for the real precedence relation occurring in a string or a tree. For typographic convenience, only the regular ‘less than’ sign ‘\(<\)’ is used in this article. Care will be taken to ensure that there is no confusion.
Sag, Gazdar, Wasow, and Weisler (1984 (SGWW 1984 hereafter) give a detailed analysis of English coordinate structures within the GPSG framework using the ID/LP format. Their analysis relies on two schemas to account for the distribution of conjunctions, the Iterating Coordination Schema (ICS) and the Binary Coordination Schema (BCS). These schemas are proposed as possible language universals. ‘+’ designates the positive Kleene closure of the preceding element. It is interpreted as permitting that element, X[CONJ \( \alpha_1 \)] in this case, to occur one or more times.

9. ICS
   a. \( X \rightarrow H[\text{CONJ } \alpha_0], H[\text{CONJ } \alpha_1]' \)
   b. \( \langle \alpha_0, \alpha_1 \rangle \in \{<\text{and}, \text{NIL}>, <\text{NIL}, \text{and}>, <\text{neither}, \text{nor}>, <\text{or}, \text{NIL}>, <\text{NIL}, \text{or}>\} \)

10. BCS
    a. \( X \rightarrow H[\text{CONJ } \alpha_0], H[\text{CONJ } \alpha_1] \)
    b. \( \langle \alpha_0, \alpha_1 \rangle \in \{<\text{both}, \text{and}>, <\text{either}, \text{or}>, <\text{NIL}, \text{but}>\} \)

11. X[CONJ \( \alpha \)] \( \rightarrow \alpha, H \)
    where \( \alpha \in \{\text{and, but, both, either, neither, nor, or}\} \)

9a and 10a are ID statements specifying the constituents allowed in each structure. 9b and 10b stipulate what values the feature CONJ can take. One should bear in mind that the feature CONJ is given values identical to the set of actually occurring English conjunctions just for mnemonic easiness. The values could as well be a random set of variables \( x, y, z \), etc. The feature values themselves cannot guarantee what conjunctions will be ‘spelled out’. 11 is the ID rule that spells these features out as lexical conjunctions. 11 does not apply to the feature [CONJ LIL] because the value NIL is not specified in the rule\(^{(7)}\). The feature [CONJ NIL] thus guarantees both that no overt conjunction would mark this category and that the category is properly marked as a part of a coordination construction. The dependencies among conjunctions are captured by requiring the two values \( \alpha_0, \alpha_1 \) to be an ordered pair \( \langle \alpha_0, \alpha_1 \rangle \). The ‘member of ‘\( \in \)relation guarantees that only the ordered pairs specified in the sets in 9b and 10b can occur. Coordination with mismatched pairs of conjunctions such as neither and or are not in the sets; therefore they would not be generated by the rules. The schema nicely captures the highly random restrictions on the co-occurrences of different conjunction in a coordination construction. It is able to specify the lexical conjunction to be inserted

\(^{(7)}\) Sag, Gazdar, Wasow, and Weisler (1985), as a revision of SGWW (1984), propose the folling rule to spell out a category with the feature [CONJ NIL].

(i) X [CONJ NIL] \( \rightarrow H \)

What rule I does is to stipulate that any category W with a feature [CONJ NIL] is obligatory rewritten as its head, an identical category of the same bar level in this case.
without proliferating PS rules. More specifically, this schema can guarantee that the same conjunction will be used throughout one Chinese NP coordination. H stands for ‘head’, and X is a minimally specified category. The order is stipulated by LP statements such as 12\(^{(8)}\).

12. \(X[\text{CONJ NIL}] < X[\text{CONJ or}]\)

```
13.                   NP
   NP[\text{CONJ NIL}] NP[\text{CONJ NIL}] NP[\text{CONJ or}]
   \(\triangle\) \(\triangle\) \(\triangle\) \(\triangle\)
   John  Jack  or  NP
   \(\triangle\)
   Jane
```

12 requires the node \(X[\text{CONJ NIL}]\) to precede the node \(X[\text{CONJ or}]\) whenever they are sister constituents. 13 is the flat structure for English or coordination generated by the ICS. Finally, for contrast, 14 is a branching structure generated by the BCS for English either... or constructions. The syntactic structure maps into semantics directly. The interpretation that Jack and Jane form a group in contrast with John is represented in the structure. The whole coordinated NP \textit{jack and Jane} and NP \textit{john} are the two constituents of the binary either... or coordination.

---

\(^{(8)}\) Notice that 12 is written as i in SGWW (1984), where \(\alpha_0\), and \(\alpha_1\), are stipulated differently form the stipulation in ICS and BCS.

(i) \([\text{CONJ, } \alpha_0] < [\text{CONJ } \alpha_1]\)

where \(\alpha_0\) \(\epsilon\) \{NIL, both, either, neither\}, and \(\alpha_1\) \(\epsilon\) \{and, nor, or\}

The notation is a little bit confusing. The order dependencies between pairs of conjunctions, such as \([\text{CONJ both}] < [\text{CONJ and}]\) are not stated explicitly. Most importantly, ii is a problem for this schema.

(ii) (=19a) either John, jack, or Jane

here we have to stipulate both \([\text{CONJ either}] < [\text{CONJ NIL}]\) and \([\text{CONJ NIL}] < [\text{CONJ or}]\). That is, we will have NIL occur as both \(\alpha_0\) and \(\alpha_1\), in the schema i. The result is an obvious contradiction: \([\text{CONJ NIL}] < [\text{CONJ NIL}]\), which prevents \([\text{CONJ NIL}]\) from occurring more than once in the same expansion, contradiction the English data. Therefore, I opt for the more straightforward way in listing all the LP statements as a set of rules.
b. Either John or Jack and Jane are on duty today.

II. APPLYING COORDINATION SCHEMAS TO CHINESE

We now turn to a GPSG account of Chinese NP coordination. The distribution described in 1 could be accounted for by schemas corresponding to 9-11 if we assume branching structures, such as 15b and c.

15.  a. (=1c) Lǎo-Wáng, hán Lǎo-Lǐ, Lǎo-Zhāng
   b.          NP
               NP[CONJ NIL]           NP[CONJ han]
                 Lǎo-Wáng    hán    NP

               NP[CONJ NIL]               NP[CONJ NIL]
                             Lǎo-Lí       Lǎo-Zhāng

   c.          NP
               NP[CONJ NIL]           NP[CONJ NIL]
                 Lǎo-Wáng    hán    NP

               NP[CONJ NIL]
                 Lǎo-Lí
Again, since the distributive reading of three individuals without internal grouping is readily available, a branching structure is neither semantically, nor syntactically, nor phonologically motivated. To account for the distributive reading and to maintain the hypothesis of direct mapping into semantics at the same time, a flat structure has to be assigned.

Two problems arise here. The first is that no LP statement can be written which will both assure that no conjunction occurs before the first conjunct and also allow conjunctions to be spelled out freely elsewhere. The other problem occurs with ICS. ICS allows two types of conjunctions to occur in one coordination structure while restricting the occurrence of one of the types to one. As exemplified in 1, both NIL and han can occur more than once in a single coordination structure, a possibility not accounted for by ICS. Thus, we face a dilemma.

III. SOLUTION 1 AND ITS EXTENSION

One way to solve the dilemma is to loosen the one occurrence constraint. My solution without reformulation the ICS is to allow choices in the values of CONJ.

16. ICS (For Chinese NPs)

    a. NP → H[CONJ \( \alpha_0 \)], H[CONJ \( \alpha_1 \)]
    b. \( <\alpha_0, \alpha_1> \) \( \in \{<\text{NIL}, \{\text{han}\}>, <\text{NIL}, \{\text{NIL}\}>\} \)

The inner curly bracket in 16b allows \( \alpha_1 \) to be a variable whose values alternate between NIL or han\(^{(9)}\). Since there is no LP STATEMENT IN Chinese to order [CONJ

\(^{(9)}\) I am taking a rather loose interpretation of the set and ordered pair notation. Since \( \alpha_1 \) is taken as a variable, instantiation it as either hun or NIL in different occurrence will still allow \( <\alpha_0, \alpha_1> \) to be a member in the defined set.

Strictly speaking, the loose formulation would face a problem when the features are being instantiated. The values are assigned as an ordered pair to the pair of variable \( <\alpha_0, \alpha_1> \) such that neither the value of \( \alpha_0 \), nor the value of \( \alpha_1 \), can vary freely once the assignment is done. What the rule is designed to prevent is the random ungrammatical occurrences of conjunctions in coordination constructions exemplified by i.

(i) *John, nor Jean, or Jane, but Mary

Notice that the schema does allow conjunctions in different pairs to occur in the same string when they are at different levels of a branching structure. the ordered pair \( <\text{NIL}, \{\text{han}/\text{NIL}\}> \), would be interpreted as a shorthand for two ordered pairs \( <\text{NIL}, \text{NIL}> \) and \( <\text{NIL}, \text{han}> \) in a more strict formal system. With this interpretation, \( \alpha_1 \) cannot vary between NIL and hun in the same rule. Such an interpretation defeats my
NIL] against [CONJ han], NIL and han can occur in any order, just as we want them to. Furthermore, with the independently motivated LP statement \(17^{(10)}\) for Chinese, we can guarantee that NP[CONJ \(\alpha_0\)] is always the first conjunct\(^{(11)}\). Since \(\alpha_0\) is stipulated to be NIL by 16b, the fact that no overt conjunctions occur initially is assured. 1c now can be assigned the flat structure 18.

17. \(X[\text{CONJ } \alpha_0] < X[\text{CONJ } \alpha_1]\)

18.

\[
\begin{array}{c}
\text{NP} \\
\text{NP[CONJ NIL]} & \text{NP[CONJ han]} & \text{NP[CONJ NIL]} \\
\text{Lao-Wang} & \text{han} & \text{NP} \\
\text{NP[CONJ NIL]} & \text{Lao-Zhang} & \text{Lao-Li}
\end{array}
\]

solution.

To save my solution under this strict formal requirement, we can assign a variable, instead of values to be instantiated, as a value to \(\alpha_i\). For example, we can assign the variable \(\beta_1\) as the value of \(\alpha_1\), with the stipulation that \(\beta_1\)’s value ranges over the set \{NIL, han\}. The revised rule is presented here as ii.

(ii) a. \(\text{NP} \rightarrow H[\text{CONJ } \alpha_0], H[\text{CONJ } \alpha_1]\)

b. \(\langle \alpha_0, \alpha_1 \rangle \in \{\langle \text{NIL, } \beta_1 \rangle, \langle \text{NIL, } \beta_2 \rangle, \langle \text{NIL, } \beta_3 \rangle, \langle \text{NIL, } \beta_4 \rangle, \langle \text{NIL, } \beta_5 \rangle\} \)

where \(\beta_i \in \{\text{NIL, han}, \beta_2 \in \{\text{NIL, gen}, \beta_3 \in \{\text{NIL, he}, \beta_4 \in \{\text{NIL, huo}, \beta_5 \in \{\text{NIL, haishi}\}\}\}\}\}\}

In instantiating the rule, one of the ordered pairs in iib, such as \(\langle \text{NIL, } \beta_1 \rangle\), is chosen as the value of \(\langle \alpha_0, \alpha_1 \rangle\). So we would have uniformly [CONJ NIL] for [CONJ \(\alpha_1\)] when the rule is applied. \(\beta_1\) would have to be assigned a real value again, because it is a variable. This time we are allowed to assign the value of \(\beta_1\) can be assigned the value of either NIL or han. In this way, we can easily account for the random occurrences of the two conjunctions these two feature values stand for.

\(^{(10)}\) A reviewer of Cahiers de Linguistique Asie Orientale observes that this rule and 28a might be instantiated as [CONJ NIL] < [CONJ NIL] and thus make the incorrect prediction in ruling out more than one occurrence of [CONJ NIL] (cf. Gazdar et al (1985: 55. note 3)). This objection is valid only when \(\alpha_0\) and \(\alpha_1\) are assigned to. When \(\alpha_0\) and \(\alpha_1\) are simply representing two different types, it does not matter what tokens are included in each type. Of course, whether it is desirable to postulate a strict interpretation of LP statements such that \(\alpha_0\) and \(\alpha_1\) are always representing types and whether one should differentiate overtly identical strings by artificially assigning them to two different types are issues with theoretical implications. It suffices to point out here that the analysis I will be giving in section V of this paper, based on a suggestion made by the reviewer, is not subject to this potential objection.

\(^{(11)}\) \(\langle \alpha_1, \alpha_1 \rangle\) is specified in the ICS and BCS for Chinese.
One fact, noted by Ed Keenan as cited by SGWW (1984: 18) is left unaccounted for in SGWW (1984). Some speakers of English use the pair of conjunctions either and or in coordination with more than two members, such as in 19a. 19b is another liberal usage of either/or coordination, noted and accounted for in SGWW (1984).

19.   a. either John, Jack, or Jane
    b. either John, or Jack, or Jane

With my revision of ICS, both 19a and 19b can be accounted for easily. Speakers accepting 19a stipulate the values of CONJ with 20 rather than with 10b.

20. \[ \langle \alpha_0, \alpha_1 \rangle \in \{ \langle \text{neither}, \begin{array}{c} \text{nor} \\ \text{NIL} \end{array} \rangle, \langle \text{either}, \begin{array}{c} \text{or} \\ \text{NIL} \end{array} \rangle, \langle \text{and}, \langle \text{NIL}, \langle \text{and}, <\text{or, NIL}> \rangle \rangle, <\text{NIL, or}> \} \]

For the first two pairs, \( \alpha_1 \) varies between two values, thus allowing any number of NIL to precede or/nor. Of course the linear order will be predicated by LP statements such as those in 21. Both 21a and 21b are independently motivated, which is an indication that some generalization about the language is being captured\(^{(12)}\).

21.   a. (=12) X[CONJ NIL] \( < \) X[CONJ or]
    b. X[CONJ either] \( < \) X[CONJ or]
    c. X[CONJ either] \( < \) X[CONJ NIL]\(^{(13)}\)

\(^{(12)}\) Carol Rosen (p. c.) observes that this schema can generate last non-initial conjuncts are marked by or. She claims that sentences like I are ungrammatical.

\((i)\) either John, Jack, Sam, or Max, or Sue

I checked this phrase out with several other native speakers. Their intuition varies from ruling it out to accepting it as grammatical though a little bit unusual. Speakers I consulted do prefer I to phrases in which conjunctions are scrambled, such as iv. Likewise, some of them prefer ii (cp. 19a), where all non-initial conjuncts are marked with or, to iii (cp. 19b), where only the last conjunct is marked; and iii toi.

\((ii)\) either John, or Jack, or Sam, or Max, or Sue

\((iii)\) either John, Jack, Sam, Max, or Sue

\((iv)\) *either John, or Jack, Sam, Max, or Sue

What we are dealing with here is degrees of acceptance, which is always a tricky topic in linguistics. SGWW (1984) draw a line to exclude all but the grammar book case where the pair of conjunctions either, and or is used in binary coordination only. The advantage of my analysis is that it generates all the possible distribution of either/or coordination, and excludes all the blatantly impossible ones. It generates i, ii and iii regardless of their degrees of acceptance, Which may be deter mined by other grammatical principles. It correctly rules out iv, which is bad by consensus of native speakers. In contrast, SGWW (1984) treat i and iii on a par with iv in failing to generate any of them. This treatment seems to contradict the intuition of native speakers. My revision would be less desirable only if one could prove that the line of grammaticality should be drawn between iii (or ii) and I, with iii (and ii) being grammatical and i being totally out. If this were the case, we would have to com up with a different analysis. Unfortunately, all the analyses I can think of right now to do that are completely ad hoc.

\(^{(13)}\) 21c is highly idiosyncratic since it is solely motivated by 19. This may explain why it is not accepted s a general rule by all English speakers.
One last problem is how to guarantee the selection of at least one *or*. My proposal is to use LP statements to express existential entailment in some cases. The observation is that in English the occurrence of *either* entails the occurrence of *or*, *neither* entails *nor*, *both* entails *and*, etc., but not the other way around. Also bear in mind that a precedence relation is meaningless unless the participating elements co-occur. It would be nice if the two phenomena could be captured in the grammar at the same time. I will use a subscript ‘e’ at the relation side of the precedence marker ‘<’ to indicate that the occurrence of the latter element is predicted by the occurrence of the former element. The same subscript on the lefthand side makes predictions the other way around(14).

22. X[CONJ *either*] <e X[CONJ *or*]

23. *either John, Jack, Jane

24. X[CONJ *budan*] <e X[CONJ *ergie*]

25. a. tā bùdàn da rén, ěrqiē mā rén
   he CONJ hit people CONJ scold people
   ‘He not only hit people, but also called people names.’
   b. * tā bùdàn da rén, mā rén
   he CONJ hit people scold people
   c. * tā ěrqiē da rén, bùdàn mā rén
   he CONJ hit people CONJ scold people

22 guarantees that at last one *or* selected whenever *either* is selected, thus ruling out the ungrammatical 23. 24 and 25 are included here to show that the same mechanism is required in Chinese VP coordination. 24 allows for 25a while ruling out 25b and 25c.

**IV. AN ALTERNATIVE SOLUTION**

Solution 1 as sketched in the last section is subject to an objection. Any addition of a new mechanism to a grammar could be costly according to formal theories of grammar. Among grammars of equivalent descriptive adequacy, the one with the fewest mechanisms (the fewest rules) is often deemed the most economical and therefore the

(14) Example of entailments in the other direction:
    (i) X[CONJ NIL] << [CONJ but]
    (ii) a. Mary is beautiful but arrogant.
        b. *Mary is but beautiful but arrogant.
    I rules out iib while allowing for iia. It is worth mentioning that this formalism reflects the ambiguity in the English sentence ‘A is required to precede B.’ where the interpretations differ in whether the existence of A is presupposed and required.
most desirable. Granting that my revision to mark existential entailment is an addition of a mechanism, one way to withstand such objections is to show that other alternative solutions that do not add new mechanisms are significantly less well-motivated.

An alternative solution, suggested to me by Ivan Sag, is to allow three categories to be specified on the right hand side of the ID rule in a coordination schema. To account for the English example in 19, this revision allows us to take the feature [CONJ NIL] from the conjunct without overt marker as the additional value to be assigned to the third variable.

26. Second Iterating Coordination Schema (ICS2)
   a. X → H[CONJ α₀], H[CONJ α₁], H[CONJ α₂]*
   b. <α₀, α₁, α₂> ∈ \{<neither, nor, NIL>, <either, or, NIL>\}

The Kleene star ‘∗’ here allows the category H[CONJ α₂] to be iterated any number of times, as few as zero and as many as necessary. On the other hand, ‘+’ requires the category H[CONJ α₁] to be chosen at least once while also allowing it to be iterated as many times as possible. When H[CONJ α₂], the value of α₂ being NIL in this case, is chosen at least once, sentences of the same type as 19a are generated. When H[CONJ α₂] is not chosen (iterated zero times) in the expansion of the rule, the more restricted type of coordination accepted by all English speakers is generated. The linear order of the conjuncts is predicated by the well-motivated LP statement 21. This schema generated all and only grammatical strings for English speakers who accept 19a. The mechanism to mark existential entailments is redundant with this schema because the positive Kleene closure, marked by ‘+’, guarantees that at least one [CONJ α₁], instantiated as [CONJ nor] or [CONJ or], would be generated. The dialectal difference between speakers who do and do not accept 19a could be characterized by whether they allow the additional category [CONJ α₂] to occur in the ICS\(^{(15)}\).

Applying ICS2 to Chinese NP coordination, we get the schema in 27. The LP statements in 28 are also required.

27. a. NP → H[CONJ α₀], H[CONJ α₁], H[CONJ α₂]*
   b. <α₀, α₁, α₂> ∈ \{<NIL, NIL, han>, <NIL, NIL gen>, <NIL, huo>, <NIL, haishi>, <NIL, NIL, he>\}

28. a. [CONJ α₀] < [CONJ α₁]
   b. [CONJ α₀] < [CONJ α₂]

\(^{(15)}\) This schema also generates sentences like i in footnote 12. The footnote 12 for a discussion on the consequences of generating such sentences.
The Chinese data given in 1 can be adequately described by 27 and 28(16). The problem with this solution is that it is redundant in always assigning the same value NIL to both $\alpha_0$ and $\alpha_1$. We recall that though conjuncts with overt markers in Chinese coordination structures could be ordered freely with conjuncts marked by conjunctions, one conjunct without overt marking must precede all the other conjuncts as the first conjuncts with the feature [CONJ NIL] are not involved in any linear precedence relation with other conjuncts, and yet one of them precedes all other conjuncts in a coordination construction, including conjuncts without overt markers, just like itself. The solution being discussed here circumvents the dilemma by representing the feature of the first conjunct [CONJ $\alpha_0$] and postulates that [CONJ $\alpha_0$] should precede all other ONCJ features, including [CONJ $\alpha_1$] whose values are also [CONJ NIL]. This is an ad hoc solution in the sense that we are just labeling the same syntactic feature [CONJ NIL] with different names, though represented by variables such as $\alpha_0$, and $\alpha_1$ is not mechanically impossible.

V. A PROPER FORMULATION OF SOLUTION I

Allowing ICS to specify CONJ value for three categories might seem to allow us to describe the English data more naturally, though it is problematic in redundantly assigning the value NIL to $\alpha_1$ in all case. An even more severe problem with the second revision, however, is that it forces us into a seemingly ad hoc analysis of the Chinese data, The Chinese data favor the first solution while the English data apparently favor the second. The first revision is favored for several reasons. First, form an empirical point of view, no language using three different overt conjunctions in a flat coordinate structure has been discovered to motivate the need to specify three different categories in the coordination schema. Anther reason, internal to the GPSG theory, is that the second revision does not go along well with ID/LP format. ICS2 has an inherent problem of having to order two types of [CONJ NIL]s in Chinese while being required to order only one of the types against [CONJ han], etc. If not labeled ad hoc as two different types, the precedence relation between [CONJ NIL] and [CONJ han] cannot be captured with LP statements. The ID/LP format, as a basic mechanism in the theory,

(16) Carol Rosen (p.c.) correctly points out that 27 can only generate two-member coordinations without overt conjunction because two-member coordination can only be generated by selecting H[CONJ $\alpha_0$] and H[CONJ $\alpha_1$], both being assigned the value NIL. This would be an objection to 27 if a BCS to account for two-member coordination with overt conjunction is not independently motivated. Of course we can argue that this analysis is not as satisfactory as my first solution because it is less economical. in deriving coordinations with the same conjunction, it generates two-member coordinations with BCS, and coordinations with more than two members with ICS, while the ICS in my first solution is able to generate both cases.
should be preserved over the more parochial coordination schema. The first revision
does not lead us into such a dilemma. ICS2 would be favored only if we get supporting
evidence from other languages that three different conjunctions are used in the same flat
coordination structure.

Turning back to Solution I, it is superior to both the simplistic flat structure
analysis and the branching structure analyses discussed in section 0. The former
proliferates phrase structure rules, while the latter proliferates tree structures. The only
reservation one may have concerning Solution I is that the proposals to have existential
entailments markers attached to LP statements and to allow alternative values for \( \alpha_1 \)
are descriptions of what an optimal solution should do rather than being well-formalized
analyses\(^{17}\). I will introduce two formal mechanisms to strengthen Solution I: a formal
definition to implement LP statements marking entailments, and incorporation the
alternative values of CONJ into ID statements.

First, a proposal to allow LP statements express existential entailments can be
formalized with a mechanism developed along the same line with Gazdar et al’s (1985: 99) definition of LP-Acceptable\(^{18}\). I will call the condition on the relation where the
existence of the preceded category is entailed LP-Acceptable. LP statements with
entailments on the other direction are called \( e \)-LP-Acceptable\(^{19}\).

\[
\text{LP}_e\text{-Acceptable}
\]

Let \( \Phi_r \) be the set of projections induced by a rule \( r \), where \( r = C_0 \rightarrow C_1, \ldots, C_n \),
and \( <_e \) is the relation of linear precedence (with existential entailment marker)
determined by the grammar, then \( \varphi \Phi_r \) is \( \text{LP}_e\text{-Acceptable} \) if and only if it is
LP-Acceptable and for every category \( \varphi (C_i) \) there are no categories \( C_j \) and \( C_j' \)
such that \( \varphi (C_i) \) extends \( C_j' <_e C_j \) if \( j > n \) (i.e. \( C_j \) is not specified in the rule \( r \)), and
such that \( C_j' <_e C_j \) is an LP-statement.

\(^{17}\) The observation of this inadequacy is due to one of the reviewers of Cahiers de
Linguistique Asie Orientale.

\(^{18}\) Gazdar et al’s (1985) definition is quoted here. Please note the difference between
the curly ‘less than’ sign ‘\( < \)’ and the regular ‘less than’ sign ‘\(<\)’ mentioned in footnote
6. The convention adopted in the definition here and in my definition 29 is that 1, ..., 
n, i, and j are all positive integers.

(i) LP-Acceptable (Gazdar et al 1985: 99. def. 9)

Let \( \Phi_r \) be the set of projections induced by a rule \( r \), where \( r = C_0 \rightarrow C_1, \ldots, 
C_n \), and \( <\) is the relation of linear precedence determined by the grammar, then \( \varphi \Phi_r \) is LP-Acceptable if and only if whenever \( \varphi (C_i) < \varphi (C_j) \) (i.e. precedes in
the tree) there are no categories \( C_j' \) such that \( \varphi (C_j) \) extends \( C_j' \) is an LP-rule.

\(^{19}\) The definition of LP-Acceptable is not given here, but readers should be able to
construct a definition in a way parallel to the definition of LP\(_e\)-Acceptable.
With LP-Acceptable defined, we can now posit that for languages requiring LP statements to express existential entailment, all projections must be LP\(_e\)-Acceptable (and/or LP\(_{e^*}\)-Acceptable) in addition to being LP-Acceptable. Strings such as 23 are LP-Acceptable as far as the coordination schemas proposed here are concerned. They are ruled out by our analysis here because they are not LP\(_e\)-Acceptable.

The proposed markers \(<_e\) and \(<_{e^*}\) and the definition of LP\(_{e}\)-Acceptable and LP\(_{e^*}\)-Acceptable are motivated by data from two typologically very different languages: Chinese and English. This is an encouraging sign for further studies to find out if the notions of LP\(_e\)-Acceptable and LP\(_{e^*}\)-Acceptable are typologically significant. Studies in GPSG argue favorably for the hypothesis that all natural languages can be reduced to ID/LP format and are thus subject to the condition of being LP-Acceptable. It remains to be shown whether all natural languages also employ the notion of LP\(_{e}\)- and/or LP\(_{e^*}\)-Acceptable. If all natural languages do, it is clear that we have an interesting universal which narrows down the set of possible natural languages because languages requiring LP statements to carry existential entailments are only a subset of languages describable with LP statements. If not all languages require LP\(_{e}\)- or LP\(_{e^*}\)-acceptability, then we ought to ask if these notions provide a meaningful partition of the set of natural languages.

Second, the proposal to allow \(\alpha_1\), to be assigned two alternative values, one of which identical to the value assigned to \(\alpha_0\), encounters the kind of problems discussed in footnote 9, i.e. in a strictly formal interpretation, my proposed schema will not be allowed to give \(\alpha_1\) alternating values. The revision discussed there is also rather cumbersome. It is necessary to come up with a formally adequate solution.

The data shows that the first conjunct has to be distinguished form the others because it is the only conjunct which cannot be marked by any overt conjunctions while the others can. This distinction leads to dividing conjuncts in to two groups with one of the groups being a singleton set, which fits the schemata proposed in SGWW (1984) well. Bear in mind the fact that the names of the values of the CONJ features coincide with the lexical conjunctions only conjunct marked with [CONJ \(\alpha_0\)] IS A NATURAL STEP. It is also logical to reduce the redundancy induced by encoding this phenomenon in a general ID statement, such as 31, which allows an NP marked by [CONJ \(\alpha_1\)] to be expanded as either a bare NP or an NP marked by the designated conjunction\(^{(20)}\). 31

\(^{(20)}\) One way to specify what lexical conjunctions are designated by what CONJ values is to stipulate a function mapping these values to conjunctions. The following rule is suggested by one of the reviewers of Cahiers de Linguistique Asie Orientale.

\[(i)\ NP[CONJ x] \rightarrow f(x), H\]

where \(f\) is defined by the following value assignment

\[
x | han | gen | he \ldots
\]
goes with the ICS for Chinese NP 30.

30. ICS (for Chinese NP)
   a. \(X \rightarrow H[\text{CONJ } \alpha_0], H[\text{CONJ } \alpha_1]\)
   where \(\alpha_0 \in \{\text{NIL}\}\), and \(\alpha_1 \in \{\text{han, he, gen, huo, haishi}\}\)

31. \(\text{NP[CONJ } \alpha_1] \rightarrow \left(\alpha_1, H, H\right)\)

31 takes care of the spelling out of all the non-initial conjuncts in a Chinese NP coordination, including all the conjuncts with overt conjunctions. It also nicely generalize the fact that any conjunction marked by an overt conjunction can be replaced by a bare conjunct without changing the grammaticality\(^{(21)}\). Most importantly, we no longer have more than one NP\[CONJ NIL\]s disguised as either NP\[CONJ \alpha_0\], or NP\[CONJ \alpha_1\] to create potential paradoxes. This analysis is very efficient because 30 is the only rule besides the special rule to spell out \(X[\text{CONJ NIL}]\) we need to spell out Chinese NP conjuncts from the Coordination Schema. This is the evidence that the right generalization is being captured. Moreover, the contrast between the distribution of conjunctions in Chinese and English coordination can now be attributed to the difference in the ID statements employed: Chinese uses 31, where an alternative bare head is allowed as an expansion, while English applies 11, where no such option is allowed.

VI. RESIDUAL PROBLEMS

In this paper I have given an analysis for Chinese NP Coordination, and proposed two revised versions of the ICS. I argued for solution I and supplemented it with the necessary formal mechanisms. The only problem remaining is that the revision to allow two possible expansions in ID statements expanding conjuncts marked by NP\[CONJ \alpha_1\], one of them a bare head, does not give the English data discussed in this paper a satisfactory analysis. Strings such as 32 require at least one \(or\) after \(either\) while allowing any number of conjuncts without overt marker to come between. Conjuncts without overt markers are not allowed to follow conjuncts marked by \(or\), such as 32b.

\(^{(21)}\) It is also possible to write 31 as i.

(i) \(\text{NP[CONJ } \alpha_1] \rightarrow \{\alpha_1, H, H[\text{CONJ NIL}]\}\)

i has the advantage of marking all the conjuncts without overt on conjuncts with the same CONJ value NIL. The requirement of LP-acceptability won’t cause any problem because the NP\[CONJ NIL\] thus introduced never share the same local tree with any other node, including the NP\[CONJ NIL\] introduced as NP\[CONJ \alpha_1\]. The only problem with this rule is that it introduced an extra level in the phrase structure tree without any justification.
pair of ID statements similar to 30 cannot guarantee the selection of at least one conjunct marked by *or*, not even with the help of the notion of LP-Acceptability. Nor can the statements guarantee that bare conjuncts precede conjuncts marked by *or*. The only alternative is to stick with the less than perfect solution of allowing two alternating values to be assigned to \([\text{CONJ } \alpha_1]\).

32.  a. either John, Mary, Jean, or Max  
    b. *either John, or Mary, Jean, Max.

One issue arises here. Should one adopt the more satisfactory ID statement solution for Chinese proposed in the last section or should one go for universality and adopt the less than perfect solution I proposed for English? It seems to be desirable to adopt the optimal solution for each language unless we can show convincingly that the alternative solution is universal. In this particular case, we are not sure that the English solution is universal, therefore my suggestion is to adopt the language specific solutions. In any case, how this revision interacts with other languages and how the two mechanisms work as parameters for natural languages should be rewarding fields for future research.

*HUANG Chu-Ren*

Cornell University
REFERENCES


