A NOTE ON ECPO ID/LP GRAMMAR

Ping Xue

Department of Linguistics
University of Victoria

1. INTRODUCTION: ID/LP FORMAT AND THE ECPO PROPERTY

The theory of Generalized Phrase Structure Grammar, as given in Gazdar et al. (1985) -- henceforth GKPS, is defined in Immediate Dominance/Linear Precedence format, which states immediate dominance and linear precedence relations separately. Immediate Dominance (ID) rules specify the constituent relation between the mother and daughter categories but do not imply any particular linear order among sister categories. Linear order is handled by Linear Precedence (LP) rules. The motivation for adopting ID/LP format rather than traditional phrase structure (PS) format is largely that the traditional formulations of phrase structure grammars fail to express generalizations about linear order that natural languages exhibit. In a phrase structure, as is well-known, there are two distinct relations among syntactic categories: the immediate dominance relation between the mother category and the daughter categories, and linear precedence relations among sister categories. Traditionally, these two types of relations among categories were jointly encoded in phrase structure rules, as shown by the following:

(1) a.   S --> NP VP
    b.   VP --> V   NP

The rule in (1a) says that S immediately dominates NP and VP, and NP linearly precedes VP. The rule in (1b) says that VP immediately dominates V and NP, and V linearly precedes NP. However, it has been long observed that linear order is not an idiocyncratic property of individual grammatical rules. Ordering relations among sister categories exhibit regularities in many languages. For instance, in English the lexical head is always initial in verbal phrases as discussed in GKPS. The structures in (2) are some familiar examples.[1]

(2) a.   VP --> V + NP
    b.   VP --> V + NP + PP
    c.   VP --> V + NP + VP
    d.   VP --> V + S'

A grammar that states the ordering in each of these rules clearly fails to express an important generalization about English.

Futhermore, many languages in the world have rather free linear order among categories. For instance, in Latin the major constituents of a simple subject-verb-object sentence can occur in any order, as shown by the structures below:

```
(3) a.  S --> NP_S + V + NP_O
    b.  S --> NP_S + NP_O + V
    c.  S --> V + NP_S + NP_O
    d.  S --> NP_O + V + NP_S
    e.  S --> NP_O + NP_S + V
    f.  S --> V + NP_O + NP_S
```

In the structure above, NP_S stands for the subject NP and NP_O stands for the object NP. Assuming the traditional PS rule format, the structures above require six individual rules, where the immediate dominance relation and linear precedence relations are stated in each of the rules. Sells (1985) notes that if immediate dominance and linear precedence are treated separately in an ID/LP format, the six Latin structures can be expressed in terms of a single ID rule like that in (4), where the commas indicate that the categories are unordered with respect to each other.

```
(4)  S --> V, NP, NP
```

In cases like the Latin one, since the order of constituents does not play any roles in the relevant structures, it is clearly not appropriate to encode such information in grammatical rules. Thus, there are good reasons for assuming ID/LP format (as opposed to PS format) in stating syntactic structures.

In the ID/LP theory adopted in GKPS, the set of expansions of any one category observes an ordering that is also observed by the expansions of all other categories. In other words, if we have a LP rule A < B in the grammar, we will always expect local trees in which a category A precedes a category B and this precedence relation holds for A and B regardless of the category from which A and B are expanded through an ID rule. This property is called Exhaustive Constant Partial Ordering (ECPO). Thus, the theory predicts that there is an ECPO property in a given language, which seems motivated in view of the word order phenomena found in English. As mentioned above, in English the lexical head is always initial in verbal phrases. In fact, it seems that in English lexical categories precede their phrasal sister categories across all constituents, regardless of categorial types, as shown by the structures in (5), in addition to the structures in (2) above.

```
(5) a.  N' --> N + PP
    b.  N' --> N + S
    c.  A' --> A + PP
    d.  A' --> A + S
    e.  A' --> A + VP
    f.  PP --> P + NP
```

In the following sections, we will show that while the assumption of ID/LP format is quite plausible, the claim of the ECPO property for the theory of grammar is too strong cross-linguistically.

2. LINEAR ORDER IN CHINESE

Languages vary in terms of the way in which grammatical functions are encoded. For instance, in lanuages like English inflectional forms of agreement are used to encode subject-predicate relations, while in languages like Japanese, nominal particles are used to mark subjects and objects. In contrast, Chinese employs very few morphological devices; instead, word order plays a crucial role in signaling grammatical functions. The most basic phrase structures in Chinese are shown as follows (cf. Li and Thomspon 1981, and Huang 1982):

```
(6) Non-lexical structures:

     a.  S   --> NP + VP
     b.  S   --> NP + S
     c.  S   --> VP + S
     d.  VP  --> ADVP + VP
     e.  VP  --> PP + VP

     f.  NP  -->  S + NP
     g.  NP  --> VP + NP
     h.  NP  --> AP + NP
     i.  NP  --> PP + NP
     j.  AP  --> ADVP + AP
     k.  PP  --> ADVP + PP

     Lexical structures:

     l.  VP  --> V + XP
     m.  VP  --> V + NP + PP
     n.  VP  --> V + S
     o.  VP  --> V + NP + S
     p.  AP  --> A + VP
     q.  PP  --> P + NP
     r.  NP  --> DET + N
```

Before getting into the discussion, we briefly comment on certain facts about the structures above. (6b) and (6c) refer to topic constructions. Since NP and VP are the most common categories which may occur as topics, we will concentrate on these two cases.[2] (6f), (6g), (6h) and (6i) describe the structures of NPs. NPs in Chinese can consist of a head noun plus other modifying elements, which, immediately followed by a modifier marker **de**, always occur before the head noun, and phrases of almost all categorial types may be used as prenominal modifiers. When a VP is used to modify a head noun, it can be regarded as a relative clause. An NP may also be modified by a clause, which may be a relative clause or an NP complement clause. Furthermore, the head of a Chinese complex NP (containing either a relative clause or an NP complement clause) is a phrasal category rather than a lexical category, since the clause may always precedes a full NP (a noun together with a determiner) as shown by the examples in (7) and (8).

(7)   women hezuo   de   nei-ge wenti
       we cooperate MO that CL question
       'that question concerning our cooperation'

(8)   renmin xunqiu heping de   nei-yi fang'an
       people seek   peace MO that CL plan
       'that plan that people seek peace'

Thus, the (local) constituent structures of Chinese in one way or another include the patterns in which S (or VP) precedes NP (cf. (6f)-(6i)), where we ignore some detailed features which we will come to discuss later.

In view of the structures given in (6), the formal advanteges of ID/LP format seems less apparent for Chinese, compared with languages such as Latin, which allow considerable degrees of word order freedom. An ID/LP grammar needs additional statements to express the generalizations about the linear order but will not drastically reduce the number of grammatical rules for a language like Chinese that relies heavily on word order to convey syntactic information. However, parallel to the situation for English, for Chinese there is at least one point that strongly motivates an ID/LP grammar. An ID/LP grammar, but not a PS grammar, may express the relevant generalization of word order in Chinese, while predicting that no structures like those in (9) will be expected to exist in Chinese:

(9)   a. S --> VP + NP
       b. VP --> NP + V
       c. VP --> NP + V + PP
       d. NP --> NP + S

3. A NON-ECPO PROPERTY AND THE NOTION OF HEAD

Given the structures in (6), however, it seems that the relevant structures in Chinese exhibit a non-ECPO property, since the set of expansions for a given category are closed under an (partial) ordering that is not constant for the expansions of all categories. Consider the following pairs:

(10)   a. S --> NP + VP
       b. S --> NP + S

       c. NP --> VP + NP
       d. NP --> S + NP

The structure in (10a), for instance, seems to motivate an LP rule like (11), but this is contradicted by the structure in (10c).

(11)   NP < VP

In fact, each of the two structures seems to motivate one LP rule, which is in conflict with the other. If so, both of the structures would be impossible since either of the structures would violate one LP rule. Thus, the set of structures can not be directly converted into an ID/LP grammar.

The important generalizations about the word order in Chinese, as noted in Huang (1982), can be characterized straightforwardly: it is head-final at the phrasal level for all the major categorial types, while it is head-initial at the lexical level for all the major categorial types except for the expansion of NP, where the head-final rule applies. These generalizations can be observed clearly from the structures in (6). This seems to suggest that the notion of H(ead) is essential to stating the generalizations of the linear order in Chinese. By incorporating the notion of H, Chinese constituent structures can be expressed in ID/LP format with LP rules like those in (12).

(12) a. X < H[¬SUBCAT]
      b. H[SUBCAT, ¬N] < X

Unfortunately, the LP rules in (12) cannot be formulated under the definitions given in GKPS, where LP rules are statements of linear order among categories (i.e. feature specifications), and the symbol H is not a feature specification. According to GKPS, a constituent structure is LP-acceptable if and only if it contains no daughter categories that are respective extensions of the categories specified in a LP rule and the linear order exhibited by them does not violate the LP rule. Thus, LP rules in GKPS are not sensitive to the notion of H. A potential argument unfavorable for the use of H in LP rules is, as Shieber (1984) points out, that the symbol H as defined in Gazdar and Pullum (1981) and GKPS will extend the expressive power of ID/LP formalism and allow certain non-ECPO grammars to be encoded. Consider the grammar in (13), which is non-ECPO:

(13) a. A' --> B A
      b. C' --> A B C
      c. C' --> B A C
      d. C' --> B C A

By incorporating the notion H, this non-ECPO grammar can be sufficiently stated in an ID/LP grammar as in (14), since the syntactic category represented by H may change from one ID rule to another.

(14) i. a. A' --> B, H
         b. C' --> A, B, H

     ii. a. B < H

As shown above, the same is true of the word order facts in Chinese. Though the relevant Chinese structures show a non-ECPO property, they can be stated in an ID/LP grammar by incorporating the notion H in the formulation.

We now face two options. One is to return to encoding constituents and linear order in PS rules for Chinese, which is certainly undesirable for reasons discussed earlier. The other is to incorporate the notion of H in the formulation of an ID/LP grammar, but this requires redefinition of LP rules in the theory.

One possible way out of the dilemma is to distinguish categories according to the grammatical functions that the categories in question perform in the relevant constituent structures. Let us look, for instance, at the structures in (10) repeated in (15) again, which show conflicting word order.

(15) a. S --> NP + VP
     b. S --> NP + S

     c. NP --> VP + NP
     d. NP --> S + NP

With respect to these structures, there seem to be two ways in which we may characterize the linear order of the daughter categories. One way is to look at the relation between the mother category and the daughter categories. As we indicated earlier, a non-head category precedes a head category, regardless of the categorial types of the constituents in question. The other way is to distinguish categories by marking those which function as modifiers. An important difference between the structures in (15a-b) and those in (15c-d) is that the structures in (15a-b) are expansions of S and the daughter VP and S are head categories functioning as predicates, while the structures in (15c-d) are expansions of NP and thus the daughter VP and S are non-head categories functioning as modifiers. In fact, constituents functioning as modifiers in a structure are syntactically different from those functioning otherwise, regardless of their major categorial types. This is not just that modifying elements in Chinese show a peculiar linear precedence, constantly preceding the head they modify, but more importantly that constituents functioning as nominal modifiers are in general marked by the particle **de**, which denotes various modifying relations of the modifying elements to the head NP, as has been discussed before. For this reason, we may assume that all categories that function as nominal modifiers in ID rules are specified as [+DE]. Thus the structures in (15) can be restated as those in (16).

(16) a. S --> NP + VP
     b. S --> NP + S

     c. NP --> VP[+DE] + NP
     d. NP --> S[+DE] + NP

Accordingly, the relevant ID rules may be stated as in (17).

22

```
(17)  a.  S --> NP, VP
      b.  S --> NP, S

      c. NP --> VP[+DE], NP
      d. NP --> S[+DE], NP
```

Now we are able to state the relevant linear order generalization in the form of a LP rule:

```
(18)  XP[+DE] < NP
```

Clearly, we also need an LP rule like the following to express the linear precedence facts of the structure in (16a-b). Thus, we may formulate LP rules in which the head-final character is not expressed directly, but is gleaned individually from the facts that NP precedes VP and a modifier precedes NP.

```
(19)  NP < VP
```

Note, however, that the LP rules in (18) and (19) have not yet solved the problem of the ordering conflicts. According to GKPS, a constituent structure is LP-acceptable if and only if it contains no daughter categories that are extensions of the categories specified in a LP rule and the linear order exhibited by the daughter categories does not violate the LP rule. The violation of one LP rule is enough to rule out the structure. The problem now is that though the structures in (16a-b) may be LP-acceptable with respect to the LP rule in (17), the structures in (16c-d) will be ruled out by (19) since the daughter categories are extensions of the categories specified in the LP rule (19) and the ordering of the daughter categories violates the ordering specified by the LP rule.

Furthermore, We have said nothing so far that prevents free instantiation of the feature specification [+DE] on non-modifying elements though there is no motivation at all for such an instantiation. This can be accomplished by means of a Feature Specification Default. We could propose that for the feature DE, the default specification is [-DE], which can be stated as:

```
(20)  FSD: [-DE]
```

The feature specification [+DE] will be introduced only through ID rules, i.e. (17c-d). Thus, if nothing is mentioned by any principles or rules, a category must have the feature specification [-DE], according to the approach to defaults adopted in GKPS. With a slight modification, the two LP rules introduced in (18) and (19) can be restated as in (21):

```
(21) a. XP[+DE] < NP
     b. NP < VP[-DE]
```

Given this, the structures associated with the rules in (17c-d) no longer violate the LP rule in (21b), since the daughter categories associated with the rules in (17c-d) will never be unifiable respectively with the categories specified in the LP rules.

Besides the two LP rules in (21), we could add three more LP rules that correspond to the linear order facts reflected in the structures in (6). Thus, without resorting to the notion H, the LP rules necessary for Chinese would be like those in (22).

(22) a. XP[+DE] < NP
     b. NP < VP[-DE]
     c. [SUBCAT, ¬N] < ¬[SUBCAT]
     d. {PP, ADVP} < {VP, AP}
     e. ADVP < PP

This set of LP rules cover the most part of the linear order facts in Chinese. However, the LP rules in (22) are inadequate. One of the potential problems concerns topic constructions. As mentioned earlier, other types of phrases (e.g. VPs) as well as NPs may also occur in sentence initial position. Assuming a general rule like S --> XP, S for the constituents of the relevant structures,[3] nothing so far proposed in the grammar ensures that the XP always precedes the S.[4] Of course, one might be able to invoke a LP rule like [+TOP] < [-TOP]. Here we are not going to discuss the argument against or for such an LP rule and the related features, nor explore further the possibility for this direction. The point is that assuming a set of context-free rules, we can always encode such set of rules in an ECPO ID/LP grammar, as noted in Shieber (1984). However, to encode a set of non-ECPO strucutures in ECPO ID/LP format seems bound to increase the number of syntactic features and LP rules in the grammar.

On the other hand, the correct linear order with respect to topic constructions is straightforwardly ensured by the LP statements in terms of the notion of H(ead) that it is head-final at the phrasal level, as stated in (12) repeated in (23).[5]

(23) a. X < H[¬SUBCAT]
     b. H[SUBCAT, ¬N] < X

4. CONCLUSION

Given the discussion above, it seems clear that though an ECPO ID/LP grammar is possible for Chinese, one can hardly say that the LP rules stated in (22) have captured the relevant generalizations in a most straightforward and natural way, even though they are technically consistant with the definitions of LP statements given in GKPS. This indicates that we have to give up something. We can either insist on the ECPO property in the formulation of an ID/LP grammar by costing more syntactic features and LP rules, or accept the notion of Head in the formulation of an ID/LP grammar without worrying about the ECPO property.

24

If the analysis above is correct, it seems to favor the latter approach. Certain cases from other languages also seem to have the same implication. One of the well-discussed cases is the word order phenomena in German. According to Uszkoreit (1987), the word order in German requires that LP rules be organized disjunctively. Each appropriate pair of daughter categories in a given constituent structure is LP well-formed if the order among them satisfies one of the LP rules. Thus, one LP rule may override another.[6] This virtually gives up the ECPO property in the formulation of an ID/LP grammar for German. The point is that an adequate grammatical theory for a natural language should be not only formally restrictive but also linguistically well-motivated. In short, while ID/LP theory provides a powerful tool for expressing linear order generalizations, for languages such as Chinese which rely heavily on word order to convey syntactic information, the insistence on an ECPO property in the formulation of an ID/LP grammar could be as inefficient and unintuitive as a PS approach to languages with various degrees of linear order freedom.

NOTES

1   Forms like those in (2) are employed in this paper to represent instances of syntactic structures, where the symbol "+" is used, in contrast to phrase structure rules as given in (1) above. Given hierarchical structures, linear order pertains only to cases where the categories are sisters.

2   We will see that the possibility of other categories occuring as topics is not crucial to the analysis.

3   See Xu and Langendoen 1985, and Jiang 1989 for related discussion.

4   Assuming S as an instance of VP, the LP rule NP < VP has the effect of blocking a linear order like S < NP.

5   It should be noted that some word order phenomena in Chinese seem to posit problems for the set of LP rules in (23) as well as the LP rules in (22). Though Chinese is generally head-initial at the lexical level, sentences like the following seem to be counterexamples to this generalization.

(i)   Nei-jian shi [$_{VP}$ [$_{PP}$ yu Zhangsan] wuguan].
      that-CL thing   with Zhangsan have-no-relation
      'That has nothing to do with Zhangsan.'

In this structure, the PP is subcategorized by the verb **wuguan** 'have-no-relation'. This seems to indicate that the relevant local structure is one in which the PP precedes the V (i.e. VP --> PP + V). Note that though it is generally agreed that the PP is subcategorized by the verb, the relation between the PP and the V is looser than the one that we would find between a verb and a subcategorized complement, since we may always insert an adverb between the PP and the verb as the following example shows:

(ii)   Nei-jian shi [$_{VP}$ [$_{PP}$ yu Zhangsan] wanquan wuguan].
       that-CL thing   with Zhangsan at-all have-no-relation
       'That has nothing to do at all with Zhangsan.'

25

This may be an evidence suggesting that the string **wuguan** 'have-no-relation' is not a lexical constituent, but rather a phrasal constituent, maybe a V'. This seems plausible, following the general assumption that adjunction is possible only at the phrasal level. In fact, it is general assumed that adjunction is possible only to a maximal projection (i.e. XP). However, there is also some discussion of adjunction to X', such as Fiengo and Higginbotham (1981), who argue for QR-adjunction to N'. The point here is that the possiblity of adjunction to the string **wuguan** 'have-not-relation' may indicate that it is not lexical category. If this is correct, the fact that the PP precedes the the string **wuguan** 'have-no-relation' follows from the generalization that the linear order is head-final at the the phrasal level or that PP precedes VP or V'. Thus, the example actually poses no problem for either of the formulations of the LP rules we have so far discussed. It should be noted here that if a subcategorized constitutient may not be a sister to the lexical head, i.e. the verb, this will pose potential problems for the GKPS treatment of subcategorization. We are not going to discuss this topic in this thesis.

6   Also see Sag (1987) for discussion of English word order.

REFERENCES

Fiengo, R. and J. Higginbotham. (1981). Opacity in NP. *Linguistic Analysis* 7, 395-422.

Gazdar, G., and G. K. Pullum. (1981). Subcategorization, Constituent Order and the Notion 'Head'. In M.Moortgat, H.v.d. Hulst and T. Hoekstra (eds), *The Scope of Lexical Rules*. For- is, Dordrech, pp. 107-123.

Huang, C.-T. J. (1982). *Logical Relations in Chinese and the Theory of Grammar*. Doctoral disserta- tion, MIT.

Jiang, Z.-X. (1989). A New Typology of Complement Clauses. In *Proceedings of the 25th Annual Meetings of the Chicago Linguistic Society*, pp. 227-242.

Li, C. N., and S. A. Thompson. (1981). *Mandarin Chinese: A Functional Reference Grammar*, Uni- versity of California Press, Berkeley, California.

Sag,   I. (1987). Grammatical Hierarchy and Linear Precedence. In G. J. Huck, and A. E. Ojeda (eds.) *Syntax and Semantics*, Vol. 20, Academic Press, New York, pp. 303-304.

Sells, P. (1985). *Lectures on Contemporary Syntactic Theories*, Center for the Study of Language and Information, Sanford University.

Shieber, S. M. (1984). Direct Parsing of ID/LP Grammars. *Linguistics and Philosophy* 7, 135-154.

Uszkoreit, H. (1987) *Word Order and Constituent Structure in German*, Center for the Study of Language and Information, Stanford University.

Xu, L., and D. T. Langendoen (1985) Topic Structures in Chinese. *Language* 61, 1-27.