A New Approach to Feature Instantiation in GPSG

Erhard W. Hinrichs

I. Introduction

It has been known for quite some time that the instantiation of agreement features on conjoined NPs presents a challenge for any adequate theory of agreement. Unlike other syntactic categories, for which agreement features in coordinate structures distribute from the mother node onto each conjunct, conjoined NPs can require non-identity between features of the mother and the features of one or more conjuncts. A notorious case in English concerns the conjunction and which typically requires plural agreement on the mother, even if all conjuncts have singular agreement. Moreover, the instantiation of agreement features for number is dependent on the following implicit hierarchy of feature values for number: [1 Person] > [2 Person] > [3 Person]. If two or more conjuncts differ in their values for person, the mother node will inherit the feature of that conjunct whose feature value is highest on the person hierarchy. These facts about person and number agreement concerning and are illustrated by the coordinate structure in (1).

(1)

That the subject NP is in fact plural and first person can be derived from the first person plural form of the reflexive pronoun is object position.

Crosslinguistically, non-identity of features in conjoined NPs is not restricted to person and number, but can also involve gender as in the French example in (2).

(2) Un savoir et une adresse merveilleux
'a knowledge (MASC) and a skill (FEM) marvellous (MASC PLURAL)'
a marvellous knowledge and skill

In this paper I do not address the general issue of covariance of agreement features in conjoined NPs, but rather discuss an interesting subcase of this more general problem, namely agreement patterns of
conjoined NPs which depend on the linear order of the conjuncts involved. In particular, I am concerned with agreement between verbs and conjuncts nearest to the verb in a following coordinate NP. Although this phenomenon has been pointed out with respect to English—compare the examples in (3) taken from Bach (1983)—I will mainly concentrate on examples from German and Russian because of the richer morphology and the greater variability of word order in these languages. Moreover, in Russian the conjunction 'i' 'and' can appear in front of every conjunct in a coordinate structure which will prove to be an important detail when we discuss the ramifications of agreement controlled by the conjunct nearest to the verb for the process of feature instantiation in Generalized Phrase Structure Grammar (GPSG). I will assume familiarity with the theory of GPSG as characterized in Gazdar/Pullum (1982). The basic organization of the theory as outlined in that paper can be summarized as in (4).

(3) a. There was a detective and three policemen in the room.
   b. There were three policemen and a detective in the room.
   (Bach 1983, 83)

(4)

```
BASIC ID RULES

Metarules

ID RULES CLOSED UNDER METARULE APPLICATION

Feature Instantiation Principles

FULLY INSTANTIATES ID RULES WITH SEMANTIC TRANSLATIONS

LP Rules

LINEARIZED PS RULES
```

The schema in (4) shows that in a GPSG as defined in Gazdar/Pullum (1982) feature instantiation properly precedes the linearization of syntactic constituents. My main claim in this paper is that such an ordering cannot be maintained if one wants to account for agreement controlled by the first conjunct of a coordinate structure. In order to account for such agreement patterns, one should rather conceive of feature instantiation principles as a set of wellformedness constraints on linearized and semantically translated PS rules. But before I can make this alternative proposal more precise, let me present the relevant data in German and Russian.

II. The Data

According to Drach's Law the finite verb in German declarative clauses occupies the second position in the sentence. Usually the first constituent is the subject, but it can also be fronted prepositional phrase as in (5), an adverbial as in (6), or the dummy es as in (7).
(5) a. Zu der Sitzung kamen (3PL) die neue Professorin und alle Assistenten. `To the meeting came the new professor and all assistants.'

b. Zu der Sitzung kam (3SG) die neue Professorin (3SG) und alle Assistenten.

c. Zu der Sitzung kam (3SG) alle Assistenten (3PL) und die neue Professorin (3SG).

(6) a. Nächste Woche konnt (2PL) Fritz und Du uns besuchen. `Next week may Fritz and you us visit.'

b. Nächste Woche kann (3SG) Fritz (3SG) und Du uns besuchen.

c. Nächste Woche kann (3SG) Du und Fritz (3SG) uns besuchen.

(7) a. Es protestierten (3PL) die Fraktion der SPD und die Grünen. `It protested the faction of the SPD and the Greens.'

b. Es protestierte (3SG) die Fraktion der SPD (3SG) und die Grünen.

c. Es protestierte (3SG) die Grünen (3PL) und die Fraktion der SPD (3SG).

If the subject NP is a conjoined NP as in (5)-(7), finite verb can either agree with the conjunct nearest to the finite verb, or with the conjoined NP as a whole. In the latter case the verb will always be marked as plural, as in (5)a-(7)a. But if the agreement is controlled by the first conjunct and if that conjunct is marked as singular, the finite verb is singular, as in (5)b-(7)b. Agreement with any one conjunct is restricted to the nearest conjunct only, as the ungrammaticality of (5)c-(7)c shows.

The same phenomenon can be found in Russian. If the verb precedes a coordinate structure with the conjunction 'i'meaning 'and', the verb can either agree with the coordinate structure as a whole, which is exemplified in sentences (8)a and (9)a, or the verb can agree with the nearest conjunct, which is the case in sentences (8)b and (9)b.

(8) a. Prepodavalis' (PL) ċerćenie i matematika. `Was taught graphics and mathematics.'

b. Prepodavaloš' (Neut SG) ċerćenie (Neut SG) i matematika.

(9) a. Na sobranie prišli (PL) professor i pjet' studentov. `To the meeting came the professor and five students.'

b. Na sobranie prišel (Masc SG) professor (Masc SG) i pjet' studentov.
For the purposes of this paper I will only analyse coordinate structures which appear as subject NPs following the verb. Corbett (1982) claims that in Russian agreement controlled by the conjunct nearest to the verb is also possible if the subject coordinate structure precedes the verb. Under certain circumstances this is also possible in German, but because of interference by other factors such as semantic salience of the conjuncts involved, these data have to await further study.

III. Theoretical Significance of the Data

If one wants to account for the agreement facts in German and Russian in a GPSG as outlined in (4), the following problem arises. Since agreement features are instantiated on the basis of unordered constituents, there is a priori no guarantee that the daughter constituent matching the agreement features on the mother will be the conjunct nearest to the verb. If one wanted to maintain the overall organization of GPSG outlined in (4) and thus apply feature instantiation to unordered constituents, one would have to distinguish the conjunct controlling agreement from all others by means of some special syntactic feature label. Of the categories used in GPSG, the one that comes to mind, of course, is the head feature, especially since we are dealing with a case of identity between agreement features which are, after all, head features. And identity between head features is commonly handled by the Head Feature Convention. Thus, one might propose a PS rule as in (10) to generate coordination structures like du und deine Freunde in the German sentence in (11).

(10) $N \bar{2} \rightarrow H [\text{CONJ e}], N \bar{2} [\text{CONJ und}]^+$

(11) Natürlich kannst (2SG) Du (2SG) und deine Freunde bleiben. 'Of course can you and your friends stay.'
You and your friends can stay, of course.

The linearization rule in (12) would further guarantee that the conjunct controlling agreement, i.e. the head conjunct, will precede the non-head conjuncts.

(12) $H \bar{2} [+\text{CONJ}] < N \bar{2} [+ \text{CONJ}]$

The first problem for this type of approach, which I will refer to as the "Head Daughter analysis", concerns the number of ID rules and linearization rules that have to be stated separately, if agreement features are instantiated on unordered constituents for a language such as German. In addition to a general coordination schema as in (13), which is modelled after the schema proposed by Gazdar/Klein/Pullum/Sag (1982) for English, the rule in (10) has to be posited along with the linearization statement in (12) to account for coordinate structures controlled by the first conjunct. Moreover, in order to account for coordinate structures with the same agreement pattern, but with und appearing before the last conjunct only, a third PS rule as in (14) would have to be stated.
(13) a. CAT' → CAT'[CONJ a₀], CAT'[CONJ a₁]⁺
   b. $a$ is in $\{\langle $und, e$ \rangle , \langle e, $und$ \rangle , \langle $weder, noch$ \rangle , \langle $oder, e$ \rangle ,
      \langle e, $oder$ \rangle , \langle $sowohl, als auch$ \rangle \}$

(14) $N \ [BAR \ 2] → H \ [BAR \ 2] \ [CONJ e] , \ N \ [BAR \ 2] \ [CONJ e]⁺ ,$
    $N \ [BAR \ 2] \ [CONJ und]$  

Notice that the number of rules necessary to generate all the relevant coordinate structures whose first conjunct agrees with a preceding VP would have to be even greater for a language like Russian. In addition to the ID rules in (10) and (14), in which und would be replaced by i, we would need a rule as in (15), since in Russian all conjuncts may be preceded by i.

(15) $N \ [BAR \ 2] → H \ [BAR \ 2] \ [CONJ i] , \ N \ [BAR \ 2] \ [CONJ i]⁺$

The second, and main objection to the Head Daughter analysis follows from the first one. Because such an analysis forces us in the case of German to use three separate rules in addition to a generalized rule schema for coordinate structures, the resulting grammar misses a number of significant generalizations. Unlike the analysis that I will present below, the head daughter analysis fails to treat coordination as a unified phenomenon by means of one generalized rule schema as in (13), but has to state three separate, and partially redundant PS rules. Moreover, even NP conjunction cannot be treated as a single phenomenon because two distinct ID rules are needed for the distribution of the lexical item und.

Furthermore, by disassociating the LP rule in (12) from the ID rules (10) and (14), it treats the linearization of constituents and the agreement pattern of conjoined NPs as logically independent, when, in fact they are crucially related. Because the rules are independent, the analysis suggests that there might be languages that do have ID rules like (10) and (14), but instead of (12) an LP rule requiring the head to always appear as, say, the third conjunct. But to my knowledge, no such language exists and for perceptual reasons is unlikely to exist.

If, on the other hand, we allow feature instantiation principles to operate on linearized and semantically translated constituents, we can generate all relevant coordinate structures in German by just one generalized schema as in (11). To account for agreement controlled by the first conjunct, we only have to state one additional feature instantiation constraint, regardless of the distribution of the lexical item und in coordinate structures.

Let us briefly outline this alternative approach to feature instantiation. At the heart of my proposal is a one-to-one and onto mapping from constituents of PS rules specifying immediate dominance relations only to nodes of locally ordered trees whose nodes consist of ordered pairs of a syntactic category $B$ and the semantic translation $\varphi$ of the syntactic expression dominated by $B$. 
(16) For each ID Rule
\[ a_0 \rightarrow a_1, \ldots, a_k \]
and ordered tree
\[ n_0 = \langle \beta_0, \varphi_0 \rangle \\
\ldots \\
n_k = \langle \beta_k, \varphi_k \rangle \]
there is a one-to-one and onto mapping
\[ f : \{a\}_i \rightarrow \{n\}_i \text{ for } o < i < k \]
such that \( f(a_0) = n_0 \) and \( \forall a_i [ f(a_i) = \langle \beta_j, \varphi_j \rangle \Rightarrow \beta_j = \text{EXT}(a_i) ] \)

Function \( f \) in (16) maps the mother constituent of the ID rule onto the root of the tree. \( f \) has to be one-to-one so that every constituent of the ID rule is mapped onto a distinct node of the tree, and vice versa. \( f \) also has to be onto so that every node of the tree is linked with some constituent of the corresponding ID rule. Moreover, each syntactic category \( \beta \) has to be an extension of the corresponding constituent of the ID rule.

Feature instantiation principles, under this approach, can be viewed as wellformedness constraints on possible mappings \( f \) from constituents of ID rules to nodes of trees. The Head Feature Convention, for example, can be defined as in (17).

(17) HFC: \[ \forall a_i [ a_i = H [BAR n] \wedge f(a_i) = n_j \Rightarrow \text{HEAD}(\beta_j) = \text{HEAD}(\beta_0) ] \text{ for } i, j \geq 1. \]

Likewise, LP rules can be conceived of as constraints on the set of possible mappings between ID rules and ordered trees. The LP rule for English that requires lexical heads of major syntactic categories to precede sister constituents can be stated as in (18).

(18) \[ \forall a_i, a_j [ a_i = H [BAR 0] \wedge a_j \neq a_i \Rightarrow f(a_i) < f(a_j) ] \text{ for } i, j \geq 1. \]

Presupposing a mapping as in (16), let us turn to an analysis of German coordinate structures controlled by the first conjunct following the verb. I follow Uszkoreit (1982) and Nerbonne (1983) whose GPSG analyses of German account for Drach's Law by treating the first constituent in German declarative clauses as the result of topicalization. Subject NPs following the verb in second position can hence be identified by the features \([-\text{TOPICALIZED}]\) and by \([+\text{NOMINATIVE}]\) as the value for the case feature. In order to generate non-topicalized coordinate subject NPs whose agreement features match that of the leftmost conjunct we have to impose the constraint (19) on possible mappings \( f \) from ID rules to ordered PS markers defined in (16).

(19) \[ \forall a_0, a_i [ a_0 = \text{EXT}(N [BAR 2][-TOP] [+\text{NOM}]) \wedge f(a_i) = n_j \wedge \exists a_j [ a_j = \text{EXT}([\text{CONJ und}]) \Rightarrow \text{AGR}(\beta_j) = \text{AGR}(\beta_0) ], \text{ where } i, j \geq 1. \]
The condition in (19) specifies that for any ID rule in which the 
dominating category is an extension of the feature N" [-TOP] [+NOM], and 
which has at least one daughter with the feature CONJ und, a daughter can 
only be mapped onto the leftmost node in such ordered trees where that 
node has the same agreement features as the root of the tree.

IV. Conclusion

By treating LP rules and feature instantiation principles as 
constraints on possible mappings from ID rules to linearized and 
semantically translated PS rules we arrive at an organization of GPSG in 
which LP rules on the one hand, and feature instantiation principles and 
semantic translation principles on the other hand apply in tandem, rather 
than inseparable components of the grammar. Such an organization of GPSG 
has been independently proposed by Klein and Sag (1982) to capture 
significant generalizations about grammatical relations and word order in 
English. While their work concerns the relationship between LP 
statements and semantic translation principles, the argument presented in 
this paper rests on the interaction between feature instantiation 
principles and LP statements.

FOOTNOTES

*I would like to thank Annie Bissantz, Remo Pereschi, and 
especially Ewan Klein and Arnold Zwicky for their helpful comments and 
suggestions on earlier drafts of this paper. I am grateful to Anelya 
Rugaleva for the Russian examples.

1The tree structure in (1) is modelled after the account of coordination given in Gazdar/Klein/Pullum/Sag (1982).

2The example in (2) is due to Corbett (1983).

BIBLIOGRAPHY


