Coordination as direct process

Augusta Mela and Christophe Fouqueré

LIPN-CNRS URA 1507, University PARIS XIII, 93430 Villetaneuse, France

KEYWORDS : linguistic formalism (HPSG), subcategorization, disjunction

E-MAIL: augusta.mela@ura1507.univ-paris13.fr
christophe.fouquere@ura1507.univ-paris13.fr
FAX NUMBER: (+33) 1 48 26 07 12
PHONE NUMBER: (+33) 1 49 40 35 90

1. Introduction

Coordination has always been a centre of academic interest, be it in linguistic theory or in computational linguistics. The problem is that the assumption according to only the constituents of the same category (1) may be conjoined is false; indeed, coordinations of different categories (2, 3) and of more than one constituent (4, 5, 6, 7) shoud not be dismissed as marginal.

(1) Jean danse la valse et le tango.
(Jean dances the waltz and the tango.)

(2) Je sais son âge et qu'elle est venue ici.
(I know her age and that she came here.)

(3) Un livre intéressant et que j'aurai du plaisir à lire.
(An interesting book and which I will enjoy to read.)

(4) Je demande à Pierre son vélo et à Marie sa canne à pêche.
(I ask Peter for his bike and Mary for her fishing rod.)

(5) Je demande à Pierre son vélo et à Marie d'où elle vient.
(I ask Peter for his bike and Mary where she comes from.)

(6) Pierre vend un vélo et donne une canne à pêche à Marie.
(Peter sells a bike and gives a fishing rod to Mary.)

(7) Pierre a acheté un livre à Marie et un disque à Pierre pour 100F.
(Peter has bought a book for Mary and a CD for Peter for 20 $.)

Consequently, this problem challenges the initial choice of categories of description which is crucial in the conception of the linguistic formalism. We claim here that the constructive power of lexical heads, encoded in the subcategorization feature, explains the previous linguistic facts. In the coordinate structure, conjuncts may be of different categories as well as of more than one constituent, they just have to satisfy the subcategorization constraints imposed by their functor if they are in position of argument or to impose compatible subcategorization constraints if they are functors, where, as in Categorial Grammars (Dowty, 1988), functors categories are those that bear unsatisfied subcategorization requirements.
Our approach which is independent of any framework, is easily and precisely encoded in the formalism of Head Driven Phrase Structure Grammar (Pollard & Sag, 1994), which is based on feature structures and makes available the feature sharing mechanism we need.
The paper is organized as follows. Section 2 is a brief account of previous approaches and section 3 is devoted to our approach. The french coordination with *et* serves throughout the paper as an example.

2. Previous approaches

There exists a classical way to eschew the question "what can be coordinated?" if one assumes a deletion analysis. Indeed, according to this approach (Chomsky, 1957), (Banfield, 1981), only coordination of sentences are basic and other syntagmatic coordinations should be considered as coordinations of reduced sentences, the reduction being performed by deleting repeated elements. This approach comes up against insurmountable obstacles, chiefly with the technical difficulty of reversing the deletion process, in the analysis process (Schachter, 1973), (Mela, 1992).
(Sag et al., 1985) have tried to make a direct description of the conjoined expressions in reducing the constraint from requiring the same category for conjuncts to a weaker constraint of category compatibility. Technically, the compatibility is checked by computing a "generalization" of categories and imposing the generalization comprises all features expected in the given context. For example, the context in (8), that is, the verb *être (to be)*, expects an predicative argument and both categories NP and AP are just predicative categories.

(8) Il est le père de Marie et fier de l'être . (NP + AP)
(He is Mary's father and proud of it.)

However, this solution cannot be applied generally because all coordinations have not such "natural" intersection (cf.(2)). So we claim that we have nothing else to do but explicitly enumerate, within the head subcategorization feature, all the structures allowed as complement for a given lexical head, that is, we will assume that the subcategorization feature has disjunctive values.

3. Our approach

Our proposal involves three stages. We begin by formulating constraints on coordinate structures, then we define how to build the coordinate structures and we end by specifying how the previous constraints filter through such coordinate structures.

3.1. Constraints on coordinate structures

We distinguish the role of functor and that of

argument, where, as in Categorial Grammars, functors categories are those that bear unsatisfied subcategorization requirements. Lexical heads are functors in relation to the arguments which they select and, by composition, any expression that contains an unsaturated functor is a functor (6)-(7) and inherits the unsatisfied requirement from the main functor. Arguments are the complements selected by the head. As suggested by (Miller, 1991), adjuncts could be accorded the same status as arguments by integrating them into the subcategorization requirement through an optional lexical rule. That would enable us to account for coordination of adjuncts of different categories as well as coordination of more than one constituent with adjuncts.

In all cases, an argument may often be realized by different categories. For example, the argument required by *savoir (to know)* may be a NP or a Completive: we say that the requirement is disjunctive and we represent the different alternatives within subcategorization feature disjunctive values. When the lexical head requires several complements (*to ask somebody something*), the requirement is said to be a n-requirement. A n-requirement is a multi-set of simple requirements. We claim then that :

(C_1) A subcategorization 1-requirement is satisfied either by one of the disjuncts or by a coordination of disjuncts.

(C_2) A subcategorization n-requirement is satisfied by m arguments, $0 < m \leq n$, either by a sequence of m arguments such that each argument satisfies one and only one element of the requirement or by a coordination of such sequences. The result has a n-m requirement.

Note that (C_1) and (C_2) should be computed simultaneously in order to account for structures as (4). The notion of partial saturation in (C_2) allows us to account for coordination of sub-series of arguments as in (5). As an entity can be both functor and argument (1), our coordination criterion (necessary condition) is the following one :

The conjuncts must satisfy the same simple or multiple subcategorization requirement and impose compatible subcategorization requirements.

In order to compute this compatibility, we define u^+, an extension of the feature structure unification to disjunctive and set values.

3.2. How coordinate structures are built

As in (Cooper, 1991), if *son âge* is a NP and *qu'elle est venue ici* is a Completive, *son âge et qu'elle est venue ici* is a conjunctive composite categorie NP ^ Compl. We extend the operation "^" to complex categories and we use a new connector in order to represent the expressions of more than one constituent within a tuple. With these two connectors, < > in order to represent the expressions of more than one constituent within a tuple. With these two connectors, a total structuring of complements is possible and all coordinate structures may have a status. Nevertheless, the extension to complex categories is not uniform: coordinate structure features are not necessarily composites or tuples of corresponding features from each conjunct. In fact, features which are allowed to have conflicting values are compounded, whereas other features as Subcat must unify. This structuring is encoded within the following HPSG-like lexical entry of *et*:

$$\left[\begin{array}{l} \text{Phon} \setminus \text{et} \setminus \\ \text{Synsem} \; \langle [1],\ldots,[M] \rangle \wedge \langle [1'],\ldots,[M'] \rangle \end{array} \middle| \; \text{Cat} \left[\begin{array}{l} \text{Part} \; \langle C_1,..,C_M \rangle \wedge \langle C'_1,..,C'_M \rangle \\ \text{Subcat} \left\{ \begin{array}{c} [1]\left[\begin{array}{l}\text{Part } C_1 \\ \text{Subcat } \{\}\end{array}\right],..,[M]\left[\begin{array}{l}\text{Part } C_M \\ \text{Subcat } \Phi_M\end{array}\right] \\ [1']\left[\begin{array}{l}\text{Part } C'_1 \\ \text{Subcat } \{\}\end{array}\right],..,[M']\left[\begin{array}{l}\text{Part } C'_M \\ \text{Subcat } \Phi'_M\end{array}\right] \\ \Phi_M \; u^+ \; \Phi'_M \end{array} \right. \end{array}\right] \right]$$

The following LP-constraint on the lexical entry of *et* ensures the correct order between conjunction and conjuncts :

$$[1] < .. < [M] < conj < [1'] < .. < [M'].$$

The coordination of m-tuples, as well as the coordination of simple conjuncts (M=1) stems from the saturation of the conjunction *et.* As required by linguistic facts, only the last element of the tuple C_M (or C'_M) can be unsaturated and be the source of inheritance.

3.3. Saturation schemata

It remains now to say how a general structured complement satisfies a general requirement, that is to say, where the checking of the conditions C1 and C2 is activated. They are called from a saturation schemata which is intended to replace the HPSG Subcategorization Principle. This saturation schemata allows either a partial ($\Psi \neq \{\}$) or total saturation ($\Psi = \{\}$) by saturated complements ($\Psi' = \{\}$), or total saturation ($\Psi = \{\}$) by complements, the latter being partially ($\Psi' \neq \{\}$) or totally saturated ($\Psi' = \{\}$) :

where Σ satisfies Φ

and Φ is a m-requirement, Ψ n-m requirement, Σ a general structured complement, Ψ or Ψ' is empty.

3.4. Example of analysis

Example of resulting HPSG-like analysis is given in figure 1 for the underlined phrase in (9):

(9) *Jean conseille à son père d'acheter et à sa mère d'utiliser un lave-vaisselle.*

(Jean advises his father to buy and his mother to use a dish washer.)

4. CONCLUSION

This approach based on concept of functor, argument and subcategorization allows us to account for many coordination datas.

Its formalization comprises two parts which are conceptually independants. On one hand, we have extended the feature structure unification to dis-

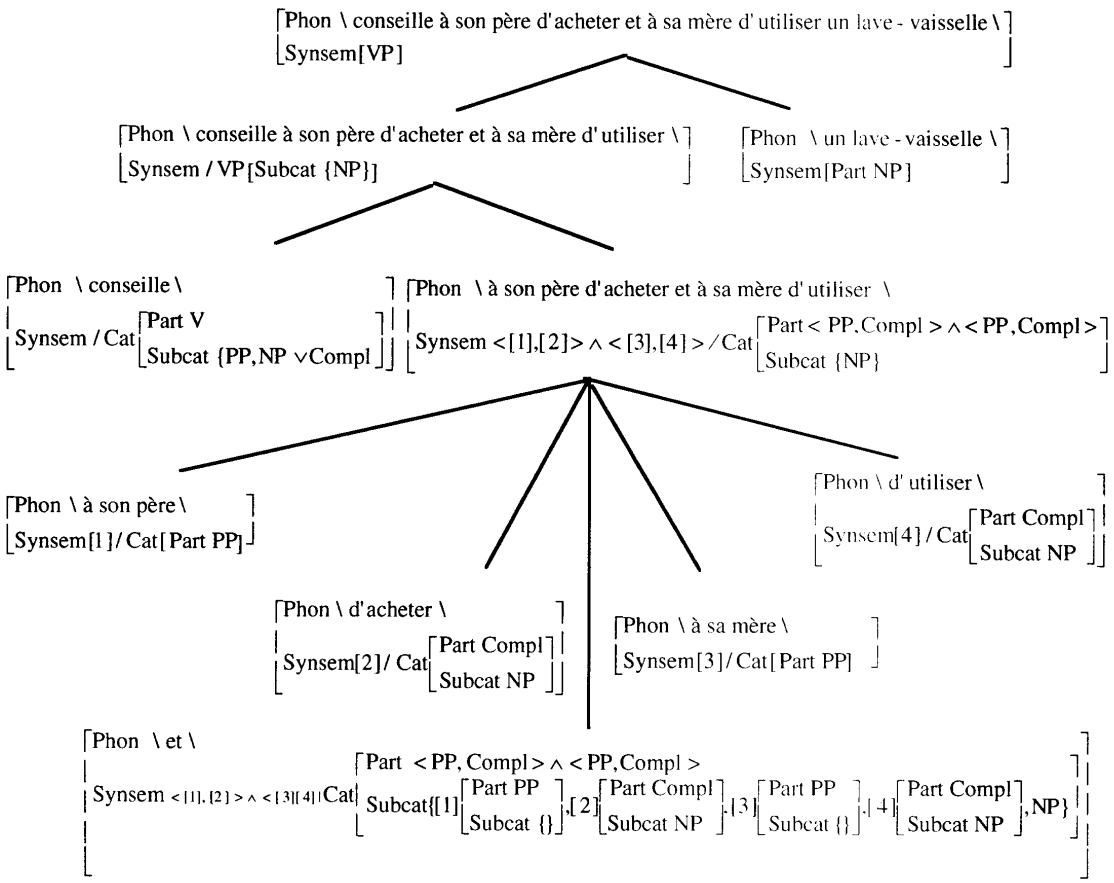

Figure 1. Analysis of *conseille à son père d'acheter et à sa mère d'utiliser un lave-vaisselle*

junctive and set values in order to check the compatibility and the satisfiability of subcategorization requirements by structured complements. On the other hand, we have considered the conjunction *et* as the head of the coordinate structure, so that coordinate structures stem simply from the subcategorization specifications of *et* and the general schemata of the head saturation.

Both parts have been encoded within HPSG using the same resource that is the subcategorization and its principle which we have just extended to account for our disjunctive values of requirements and structured categories.

Bibliography

Cooper, R. P. 1991. Coordination in unification-based grammars. Actes *29° ACL*, 167–172.

Mela, A. 1992.Traitement automatique de la coordination par "et". Thèse de l'université de Paris XIII, Villetaneuse (F).

Mela, A. et Fouqueré, C. 1994. "A coordination criterion based on multiple and disjunctive subcategorization", dans *The Proceeding of the 1994 Kyoto Conference,* Akira Ishikawa et Yoshihito Nitta (éds), Société Logico-Linguistique du Japon, Tokyo, Japon, 31–40.

Mela, A. et Fouqueré, C.1996. "Coordination as a direct process". *Actes 34° ACL*, Santa Cruz. (to appear).

Miller, P. 1991. Clitics and Constituents in Phrase Structure Grammar. Garland (eds), New York. 1992.

Paritong, M. 1992. Constituent Coordination in HPSG. *KONVENS*, Springer Verlag, Berlin, 228–237.

Pollard, C. et Sag I.A. 1994. *Head-Driven Phrase Structure Grammar*. Chicago : UCP and Stanford: CSLI Publications.

Sag I., Gazdar G., Wasow T., Weisler S. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic theory*, n°3, 117–171.

Schachter, P. 1973: "Conjunction", *The Major structures of English*, Chap 6. Holt, Rinehart & Winston. USA.