GRUPO 5: ANÁLISIS PRAGMÁTICO

PRESIDENTE: KEPA SARASOLA

A Coordination Criterion Based on Multiple and Disjunctive Subcategorization

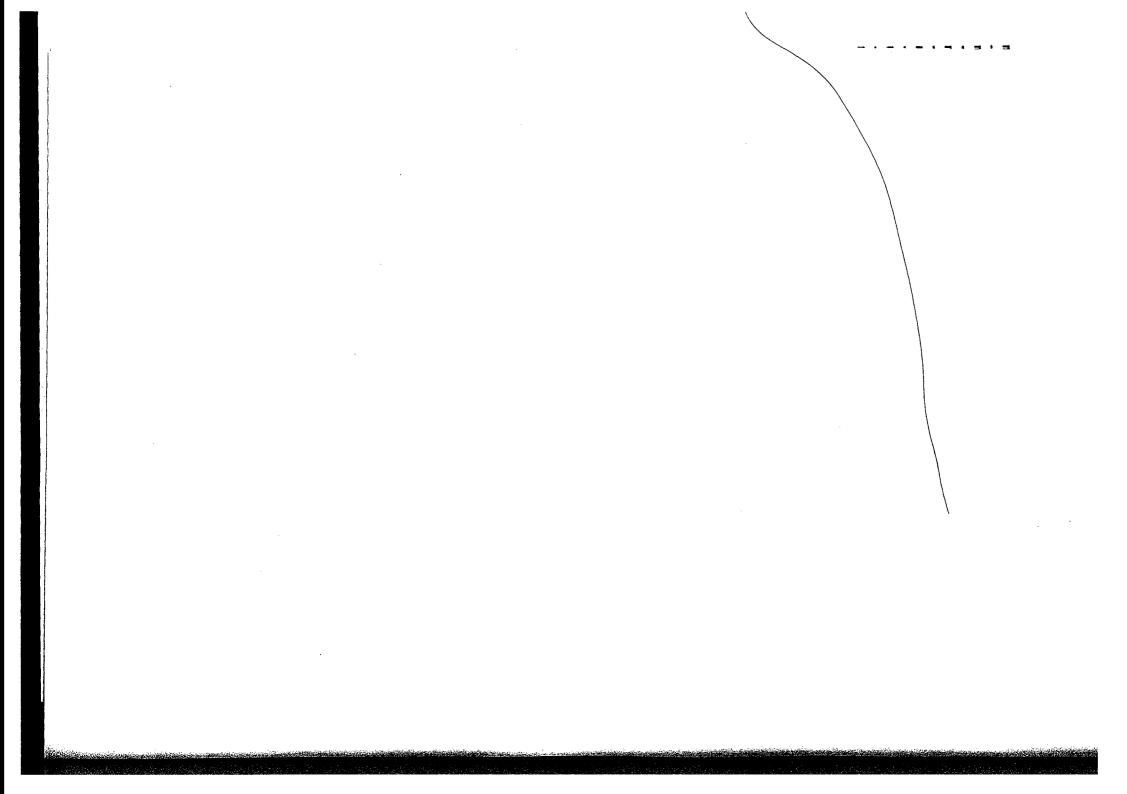
A. Méla, C. Fouqueré

Resolución de la anáfora discursiva mediante una estrategia de inspiración vectorial

C. Rico

Espacios de Minkowski para la representación del tiempo y la modalidad

c. Inchaurralde



A coordination treatment based on multiple and disjunctive subcategorization

Augusta Méla and Christophe Fouqueré

LIPN-CNRS URA 1507

Université Paris-Nord

93430 Villetaneuse France

email: am@ura1507.univ-paris13.fr

Abstract:

This paper proposes a modelization of the coordination within the framework of Head-driven Phrase Structure Grammar (HPSG), based on concepts of functor, arguments and subcategorization requirements. It enables coordination of more than one constituent as well as of different categories. This modelization lies on a lexicalization of the coordination principle and two generalizations of the notion of syntactic category. First, the coordination of different categories requires composite categories: disjunctive for subcategorization requirement, conjunctive for the coordinate structures. Second, the coordination of more than one constituent assumes that a sequence of constituents is represented as a tuple.

Resumen :

Este artículo propone una modelización de la coordinación en Head-driven Phrase Structure (HPSG), fundada en los conceptos de functores, argumentos y demandas de subcategorización. Esta tiende a abarcar la coordinación de mas de un constituyente y tambien la de categorías diferentes. Esta modelización estriba en una lexicalización del principio de la coordinación y sobre dos generalizaciones de la noción de categoría sintaxica. Por una parte, la coordinación de diferentes categorías requiere la noción de categorías compuestas, disyuntivas para las demandas de subcategorización y conjuntivas para las estructuras coordinadas. Por otra parte, la coordinación de mas de un constituyente requiere que una serie de constituyentes sea representada por uno "n-plicado" de categorías.

1

1. Introduction

This paper proposes a modelization of the coordination within the framework of Head-driven Phrase Structure Grammar (HPSG), based on concepts of functor, arguments and subcategorization requirements. It enables coordination of more than one constituent as well as of different categories. This modelization lies on two main ideas. First, the coordination of different categories requires composite categories: disjunctive for subcategorization requirement, conjunctive for the coordinate structures. Second, the coordination of more than one constituent assumes that a sequence of constituents is represented as a tuple of categories.

The french coordination with et serves throughout the paper as an example. An extensive use of lexical subcategorization is made: et subcategorizes two conjuncts that subcategorize complements. All these subcategorization requirements are specified in a declarative way and controlled dynamically during the parsing/generation process. HPSG offers not only such a service through the subcategorization principle, but also integrates linguistic information of all kinds in a single representation and allows to treat in this single principle a range of syntactical and semantical dependencies, lexically determined, e.g. case assignment, government (of particular prepositions) and role assignment. Indeed all this information is necessary to rule out the following sentences:

- (1) * Jean connaît et semble une personne travailleuse. (Jean knows and seems a trustworthy person)
- (2) * Marie promet et ordonne à Luc de partir. (Mary promises and orders Luc to go) 1

The paper is organized as follows. The section 2 is devoted to present the concepts we need through typical examples. We formalize our concepts in section 3. We end by some comments and residual problems.

2. Coordination: subcategorization requirements, functors and arguments

The classical typology of coordination, i.e. coordinations of the same category, of different categories and of more than one constituent, hides the regularity of the phenomenon as it focuses on concepts of constituent and syntactic category. Since this regularity is functional, it would be more relevant to focus on concepts of functor/arguments and subcategorization requirements.

Two roles will therefore be distinguished: the role of functor and that of argument. The unsaturated lexical heads are functors in relation to the complements which they subcategorize and the arguments are the complements specified by the heads. An element can be both functor and argument. In the following sentence:

^{1 (1)} is ruled out because the two verbs don't assign the same case to the shared element "une personne travailleuse". (2) is ruled out because the two verbs don't assign the same control on the subjet of the object clause "de partir".

(15) Je pense offrir et recevoir des cadeaux. (I expect to offer and receive gifts)

offrir and recevoir are arguments with respect to penser (expect) and functors with respect to des cadeaux (gifts). The previous sentence exemplifies a coordination of functors and arguments at the same time.

In this paper, we claim that elements can be coordinated if they satisfy the same (simple or multiple) subcategorization requirement and that they impose compatible subcategorization requirements. The reader may find in [Méla,92] an exhaustive description of coordination problems. We give hereafter only useful examples to understand the paper.

2.1. Arguments satisfying a requirement

A subcategorization requirement constraints the relation between a head and its arguments: it is called simple when the lexical head subcategorizes a single argument (to know something), multiple when the lexical head subcategorizes several arguments (to ask somebody something).

In both cases, the arguments may be realized by different categories. For example the object required by savoir (to know) may be either an NP or a Completive (disjunctive requirement), or the coordination of the two:

(3) Je sais son âge / qu'elle est venue ici/ son âge et qu'elle est venue ici.

(I know her age/ that she came here/ her age and that she came here).

To the extent that disjunction only appears on the level of specification of an argument, a multiple requirement is a set² of simple requirements and is satisfied either by a series of arguments which satisfy respectively that set of simple requirements:

- (4) Je demande à Pierre son vélo / d'où il vient.

 (I ask Peter for his bike / where he comes from)
- (5) Je demande à Pierre son vélo et sa canne à pêche.(1 ask Peter for his bike and his fishing rod)
- (6) Je demande à Pierre son vélo et d'où il vient.

 (I ask Peter for his bike and where he comes from)

or by the coordination of a series of this kind:

(7) 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)

The choice of a set (or more precisely multiset)-value status for SUBCAT rather than list will become clear with examples (7) and (7'). [Gunji, 1987] makes the same choice. However our criterion can be formalized in a theory whose order of arguments obeys to an obliqueness hierarchy.

- (7') Je demande son vélo à Pierre et à Marie sa canne à pêche.
- (8) Je demande à Pierre son vélo et à Marie d'où elle vient.
 (1 ask Peter for his bike and Mary where she comes from)

2.2. Functors: inheritance and compatibility of requirements

The coordination of functors needs that theirs requirements are compatible. As shown in the following coordinations, functors may appear as unsaturated lexical heads:

(9) Jean achète et répare des vélos.

(John buys and repairs bikes)

partially saturated:

(10) Jacques aime et Marie déteste ces beaux spots lumineux.
(Jack loves and Mary hates these beautiful spotlights)

or composed:

(11) Jacques répare et prétend détester ces beaux spots lumineux.

(Jack repairs and claims to hate these beautiful spotlights)

In order to distinguish the different composed functors, we call functor-functor the first functor (claims), functor-argument the second (to hate), etc. The resultant functor may inherit the unsatisfied subcategorization of the functor-argument iff the functor-functor is saturated and the functor-argument is its last argument. Compare (12) with (13-14):

- (12) Marie prétend avoir offert et Jane prétend avoir vendu ces beaux spots lumineux à Paul.

 (Mary claims to have given and Jane claims to have sold these beautiful spotlights to Paul)
- (13) * Marie dit qu'elle déteste à Jacques et Jane dit qu'elle aime à Paul ces beaux spots lumineux.

 (Mary says that she hates to Jack and Jane says that she likes to Paul these beautiful spotlights)
- * Marie dit qu'elle déteste et Jane dit qu'elle aime ces beaux spots lumineux à Paul.

 (Mary says that she hates and Jane says that she likes these beautiful spotlights to Paul)

In both cases, when two functors are coordinated, they share their arguments: there must therefore exist at least one possibility of satisfying them simultaneously; the specification imposed by their coordination is their common specifications which the common arguments must satisfy otherwise the coordination is agrammatical:

* Je dépends et j'obéis à mon frère. (I depend and I obey my brother)

2.3. Satisfying and imposing requirements

As an entity can be both functor and argument (cf. (15), (17) and (21)), our criterion is the following one:

Coordination criterion

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

The conjuncts may be simple:

Je pense offrir et recevoir des cadeaux. (I expect to offer and receive gifts)

composed:

Je pense offrir et espère recevoir des cadeaux. (I expect to give and hope to receive gifts) (16)

unsaturated:

Je pense recevoir de Jean et offrir à Pierre du caviar de Russie. (17)(I expect to receive from John and offer to Peter Russian caviar)

with different structures:

- (I think to offer and that I will receive gifts) Je pense offrir et que je recevrai des cadeaux. (18)
- Je recommande à Pierre la lecture et qu'il s'inspire de la Bible. (19) (I recommend to Peter the lecture and that he inspire himself of the Bible)

with more than one constituent:

Je recommande à Pierre de lire et à Marie d'acheter la Bible. (20)(I recommend to Peter to read and to Mary to buy the Bible)

with more than one constituent and different structures:

Je pense recevoir de Jean et que j'offrirai à Pierre du caviar de Russie. (I think to receive from John and that I will offer to Peter Russian caviar) The notions of functor and argument induce a new typology of coordinations which covers the classical typology:

coord° of:	same categories	different categories	more than one constituent
Argu-	Je demande à Pierre son vélo et sa canne à pêche	Elle dit son nom et qu'on l'appelait Bibi	Je demande à Pierre son vélo et à Marie sa canne à pêche
-ments		Je demande à Pierre son vélo et à Marie d'où elle viens	
Functors	Jean achète et répare des vélos fidèle et dévoué à sa femme il vit par et dans les livres	*offrir et que je recevrais des cadeaux	Jacques aime et Marie déteste ces beaux spots lumineux
	Mon collègue et ami * Kim préfère et promet à Sandy de partir	* lis et par les livres achète et lis des livres	Jacques aime et Marie prétend détester ces beaux spots lumineux
neither functor nor argument		Jean et Marie Jean lit et Marie joue (John read and Mary plays) *Jean lit et une trotinette (John read and a child's scooter)	
Arguments and functors	Je pense donner et recevoir des cadeaux	Je pense offrir et que je recevrai des cadeaux Je recommande à Pierre la lecture et qu'il s'inspire de la Bible	Je pense recevoir de Jean et offrir à Pierre du caviar russe
		Je pense recevoir de Jean et que j'offrirai à Pierre du caviar de Russie	

The formalization in section 3 takes care of all these situations.

3. Formalization in HPSG

Let us recall that in HPSG the feature *Synsem* contains both the semantic and syntactical information and the value of subcategorization feature is a set (or list) of Synsem.

3.1. Disjunctive and multiple satisfiability

Definition

A subcategorization requirement of valence n or n-requirement is a set of n disjunctions of Synsem.

ex:

 $\label{lem:demander} $$ \operatorname{Cat}_{\operatorname{Sysem}}(t) = \operatorname{Cat}_{\operatorname{Sysem}}(t) - \operatorname{Cat}_{\operatorname{Sysem}}(t$

[&]quot;Part" is the abbreviation of "part of speech", i.e. classical categories as NP,PP, etc.

Satisfiability conditions

(CI)

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

(C2)

A subcategorization n-requirement is satisfied either by a sequence of n arguments such that each argument satisfies one and only one element of the requirement or by the coordination of such sequences.

If we want to capture coordinations of more than one constituent like (6,8) we need both (C1) and (C2), hence extending [Cooper,91]. We assume that the following coordinations have the following status:

(3) Je sais son âge et qu'elle est venue ici

: NP ^ Compl 4

(5) Je demande à Pierre son vélo et à Marie sa canne à pêche

: <PP, NP> << PP, NP>

(6) Je demande à Pierre son vélo et d'où il vient

: <PP, NP^Compl>

(8) Je demande à Pierre son vélo et à Marie d'où il vient

: <NP, Compl>^ <PP,NP>

and we propose (as an extension of the present HPSG subcategorization principle) the general constraint (C1 & C2) formalized as follows.

In a functor-arguments structure of the general following form:

where $\Phi = \{\alpha_i / i \in [1,p]\}$ is a p-requirement i.e. $\forall i, \alpha_i = Si_1 \lor ... \lor Si_{n_i}$, Ψ a requirement possibly empty, and $\Sigma = \langle S_1^1, ..., S_p^1 \rangle \land ... \land \langle S_1^q, ..., S_p^q \rangle$, a coordination of q tuples (if q>1) or one tuple (if q=1) of p composite Synsem $S_i^k = \land \{S_{i,m}^k / m \in [1,z_i^k]\}^5$,

Σ satisfies (⊨)Φ iff:

∀ k ∈ [1,q],

(for each tuple of Σ)

 $\forall i / \alpha_i \in \Phi$,

(for each subcategorized argument)

∃ p_k a permutation of [1,p]

(the choice of one synsem $S_{p_k[i]}^k = \Lambda \{S_{p_k[i],m}^k / m \in [1,z_{p_k[i]}^k] \}$

in the tuple is different each time)

 $/ \forall m \in [1, z_{p_k[i]}^k]$,

(for each synsem of the composite synsem)

∃j ∈[1,nj]

(there is one disjunct of α_i)

where NP \(\text{Compl} \) is the abbreviated form of [Sysem=Cat=[Part=NP]] \(\) [Sysem=Cat=[Part=Compl]], etc. if m =1 then the category is "simple" and not composite.

/
$$S_{p_k[i],m}^k \cup S_{ij}$$
 (such as this disjunct legitimates the synsem)

where ' \bigcup ' is the usual unification between categories. The tuple $\langle S_1,...,S_p \rangle$ is defined below.

3.2. Compatibility of subcategorization requirements

The unification of two 1-requirements α and β is defined as follows:

Let $\alpha = \{ \bigvee_{i=1...n} s_i \}$ and $\beta = \{ \bigvee_{j=1...m} t_j \}$, s_i and t_j are Synsem therefore: $\alpha \bigcup \beta = \{ \bigvee_{i=1...n} j_{i=1...m} s_i \bigcup t_j \}$ The unification of two n-requirements α and β is defined as follows:

Let $\Phi = \{\alpha_i \mid i \in [1,n]\}$ and $\Psi = \{\beta_i \mid i \in [1,n]\}$ α_i and β_i are disjunctions of Synsem, therefore: $\Phi \bigcup \Psi = \{\alpha_i \bigcup \beta_i p_{\{i\}} \mid i \in [1,n]\}$ where p is a permutation on [1,n].

The result of the unification of two n-requirements is therefore ambiguous.

Two subcategorization requirements are compatible iff their unification succeeds.

3.3. Lexicalization of the coordination rule

As subcategorization characterizes the individuality of a lexical unit, likewise the fact that the conjunction et requires two conjuncts is independent of the particular utterances in which it appears. It is therefore legitimate that the rule of coordination itself is encoded in the lexical entry of the conjunction et. The originality of [Paritong,92] is to consider the conjunction as the head of the coordinated structure, not a lexical head but a functional head. Indeed he distinguishes among the HEAD features, the substantive features (noun, verb, adjective, preposition, agreement, case, tense,...) which are called MAJOR features and the functional features (determinant, complementizer,...) called MINOR features. This functional head subcategorizes two complements which are the conjuncts.

It is helpful to consider the conjunction as the head of the coordinated construction because the distribution of the conjuncts no longer has to be postulated in the grammar by a special rule of coordination: it stems simply from the specifications of the subcategorization of the conjunction.

In [Paritong,92] the sharing of arguments in the case where the conjuncts are unsaturated lexical heads stems from the "reentrancy" of the lexical entry of the conjunction; the SUBCAT features of the conjuncts and of coordination as a whole share the same value marked [1] in the diagram below:

Following [Pollard & Sag, 1987: 45] et [Miller, 1991: 26], we admit that the requirements are of the same cardinality. This condition "will forbide the conjunction of e.g. verbs with SUBCAT lists of different lengths, but which would have a unification under the alternative interpretation, thus avoiding sentences like *John bought and gave the book to Mary.".

lexical entry of the conjunction "et"

To extend this treatment to coordinations of different categories and those composed of more than one constituent, we propose two extensions:

- 1) since the subcat feature is a set of disjunctions, the sharing of arguments in the case where the conjuncts are unsaturated is modelized not by "reentrancy" but by unification U such as we have defined it above of the subcategorization requirements of each conjunct.
- 2) the feature Part (of speech) of the coordination of two Synsem is the composite of the features Part of the two Synsem⁷,

Then the lexical entry of the conjunction et is:

• Definition of $\langle S_1,...,S_N \rangle$

As shown by the contrast between:

- (20) Je recommande à Pierre de lire et à Marie d'acheter la Bible.
 - (I recommend to Peter to read and to Mary to buy the Bible)

and:

(22) * Je recommande de lire à Pierre et d'acheter à Marie la Bible.

(I recommend to Peter to read and to Mary to buy the Bible)

only S_N , the last complement of the tuple, may be unsaturated; so the expressions of more than one constituent $\langle S_1, ..., S_N \rangle$ has to be constructed by the following rule:

$$\label{eq:cond_problem} \begin{split} & \text{Phon=}\backslash 1 \& ... \& n \backslash \\ & \text{Synsem}<[1],...,[n]>=& \text{Cat=} \begin{bmatrix} \text{Head=}[\text{Part=}],} \\ & \text{Coord=}+, \text{Subcat=}\Phi \end{bmatrix} \\ & \text{Branches} \begin{bmatrix} \text{Phon=}\backslash 1 \backslash \\ \text{Synsem}[1]=& \text{Cat=}[\text{Head=}[\text{Part=}S_1], \text{Subcat=}\{\}] \end{bmatrix} \\ & \text{...} \\ & \text{Phon=}\backslash n \backslash \\ & \text{Synsem}[n]=& \text{Cat=}[\text{Head=}[\text{Part=}S_n], \text{Subcat=}\Phi] \end{bmatrix} \end{bmatrix}$$

Let us recall that $A \wedge A = A$, for instance $NP \wedge NP = NP$.

It remains for us to define how the functors are built by partial saturation, by composition.

 Partial saturation doesn't need any additional rule or modification: the subcategorization principle of HPSG allows unsaturated functors with the following description:

$$\begin{bmatrix} Synsem=Cat=[Head=[1],Subcat=\Psi] & where \Psi \neq \{\} \\ B-Head=Synsem=Cat=[Head=[1],Subcat=\Phi \cup \Psi] \\ B-Comp=<[Synsem=Y_1],...,[Synsem=Y_n] > \end{bmatrix} & and B-Comp \models \Phi \end{bmatrix}$$

However this schema induces a flexible constituency as needed in utterances like:

(10) Jacques aime et Marie déteste ces beaux spots lumineux.

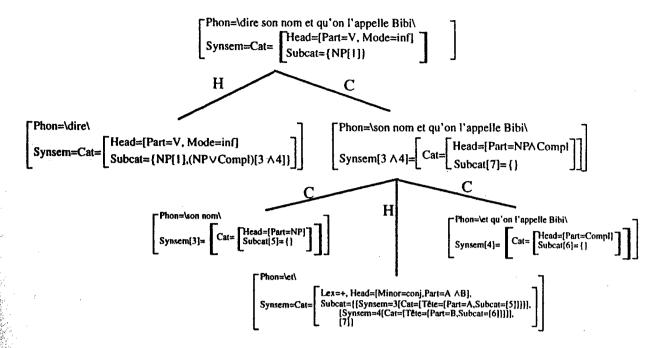
(Jack loves and Mary hates these beautiful spotlights)

In order to overcome the problems of artificial ambiguity which result from it, one can use it only under a coordination: a feature Coord assures this control provided this feature appears both on the resultant category of partial saturation and on the conjuncts in the lexical entry of et.

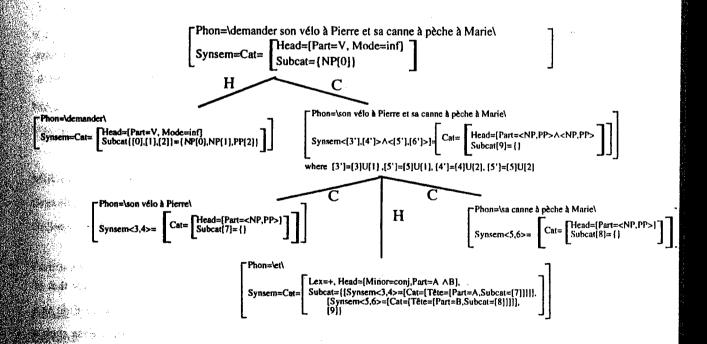
 \bullet Last, in taking account that only S_N , the last complement of the tuple, may be unsaturated, the functors composition is allowed by the following description:

$$\begin{bmatrix} \text{Synsem=Cat=[Subcat=Ψ]} \\ \text{Branches} \begin{bmatrix} \text{B-Head=Synsem=Cat=[Subcat=Φ]} \\ \text{B-Comp=<[Synsem=$Y_1],...,[Synsem=$Y_n[Subcat=Ψ]]} \end{bmatrix} \end{bmatrix} \text{ where B-Comp } \models \Phi$$

The resulting tree for "dire son nom et qu'on l'appelle Bibi" ("to say his name and that he is nicknamed Bibi") is the following one:



The resulting tree for "demander son vélo à Pierre et sa canne à pêche à Marie" is the following one :



4. Other aspects and residual problems

• A status for sequences of more than one constituent is needed for the description of the coordination and likewise for other linguistic phenomenas with symetrical sequences of more than one constituent (comparative constructions, alternative construction, ...):

Paul donne autant de couteaux aux filles que de pièces aux garçons.

(Paul gives as much knifes to the girls as coins to the boys.)

Paul donne soit des couteaux aux filles, soit des pièces aux garçons.

(Paul gives either knifes to the girls or coins to the boys.)

- According to [Miller,91], 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 treat the coordinations of adjuncts of different categories as well as the coordinations of more than one constituent with adjuncts.
- All the modelisations of the coordination of different categories are insufficiently controlled. Let us take a closer look: in GPSG (Generalized Phrase Structure Grammar), the rule of coordination in its simplest version is: X-->H and H', H and H' being minimally specified categories. In addition, by HFC, we have $X/HEAD = \bigcap_{i=1}^{n} Hi/HEAD$, that is to say from the point of view of the head features, X is a generalisation of two conjuncts.

Due to ID rules description X is generally sufficiently specified to prevent unification with any generalisation of different categories. But that is not always the case: nothing will block the generation of utterances such as:

- (22) *Jean lit et une trottinette. (John reads and a child's scooter)
- (23) * offrir et que je recevrais des cadeaux (to offer and that I receive gifts)
- (24) * lis et par les livres (read and by books)

This criticism is also valid for the treatment of the Lexical Functional Grammar (LFG) (at least for that of [Sells,85] and [Cooper, 91]). The latter recognizes "our grammar admits in the garden and chases Fido as a constituent" but he reassures himself by adding: "though there may be no contexts which license such a constituent".

Certainly no linguistic context exists which would be compatible with such sequences of words but they will be generable and analysable in our proposition, as in the others, unless we impose the constraint that any

analysable chain must be of the type (S)entence which is likely to be quite restricting in the analysis of texts. Even if we analyze the categories in terms of the features N, V, it is not clear how to control these coordinations.

However, the formal criterion we propose covers both coordination of arguments and functors introduced in a linguistic context (\neq (23) et (24)). Let us finally recall that this unique criterion is valid in the two cases as well as for the coordination of more than one constituent and of different categories.

5. Bibliography

Abeillé A. 1993. Les nouvelles syntaxes. Armand Colin.

Cooper R. P. 1991. Coordination in unification-based grammars. Proceedings of the ACL, 167-172.

Dowty D. 1988. Type Raising, Functional composition, and non-constituent conjunction in Categorial Grammars and Natural Language Structures, Richard T.Oehrle et al. (eds.), 153-197.

Gazdar G. 1981. Unbounded dependencies and coordinate structures. Linguistic Inquiry, Vol 12.

Gunji T. 1987. Japanese Phrase Structure Grammar. Dordrecht. Reidel.

Kasper, R. T. et Rounds W. C. 1986. A Logical Semantics for Feature Structures. Proceedings of the ACL, 257-265.

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

Miller P. 1991. Clitics and Constituents in Phrase Structure Grammar. Thèse de l'université libre de Bruxelles, Faculté de Philosophie et Lettres en Institut de Recherches en Intelligence Artificielle (IRIDIA). Garland (eds), New York. 1992.

Paritong M.1992. Constituent Coordination in HPSG. KONVENS 92, Springer Verlag, Berlin, 228-237.

Pollard, C. et Sag I.A. 1987. Information-based syntax and semantics, CSLI, Lecture Notes Series, Chicago University Press, Chicago.

Pollard C. 1991. Topics in constraint-based syntactic theory. Universität des Saarlandes, Saarbrücken, August 12-23.

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

Sclls P. 1985. Lectures on contempory syntactic theories. An introduction to government-binding theory, generalized phrase structure grammar and lexical-functional grammar. CSLI, Lecture Notes n°4.

Steedman M. 1985. Dependency and coordination in the grammar of Dutch and English. Language, Vol 61, n°2, 523-568.

Steedman M. 1990. Gapping as constituent coordination. Linguistics and Philosophy 13, 207-263.