

Grammaires d'unification à traits et contrôle des infinitives en français
(Unification Grammars with Features and Control of Infinitives in French)

Karine Baschung

(Université Paris X–Nanterre)

Paris: Éditions Adosa (Langues Naturelles et Traitement de l'Information 2, edited by Gabriel G. Bès), 1991, 401 pp.
 Paperbound, ISBN 2-86639-007-5,
 FF 360.00

Reviewed by

Dominique Estival

Université de Genève

The publication of this (1990) doctoral dissertation, written at the University of Clermont-Ferrand under Gabriel G. Bès, makes it possible to have in French a very useful reference work for syntactic systems based on generalized phrase-structure grammars (GPSGs) and unification categorial grammars (UCGs), as well as a good treatment of French syntax. Not only is the text very precise in its description of the constructions it deals with, but the detailed technical annexes (a quarter of the whole book) present lexical entries, grammar rules, and execution traces that give the reader a concrete idea of the actual implementation. The linguistic analysis, faithfully based on the data given by Gross (1975) for control phenomena, also covers other aspects of syntax (linear order, modifier/argument relations, etc.).

The very good introduction to the theoretical framework of GPSG is based on that of Gazdar et al. (1985) and is followed by a detailed presentation of the way it was adapted for the grammar of French.¹ The introduction to UCG is based on that of Calder et al. (1986) and is illustrated by its implementation in the FDP system, built at Clermont-Ferrand, and by Baschung's own functional grammar (FG).² The explanation of the operations of *composition* and *concatenation* in UCG is so clear that it can serve as a reference for the theoretical framework, and the comparison between the two UCGs reveals those aspects that are due to linguistic choices rather than to the formalism itself.

A thread running through the whole book, and especially emphasized in the section on GPSG, is the argumentation against the *minimal distance principle* (PDM in French) for determining the referent of the controller of an infinitive on a purely structural basis, with a criterion applied to the syntactic (or semantic) representation of the sentence. Baschung's thesis is that, on the contrary, the determination of the controller is based on a set of principles, some universal and some particular to the specific lan-

1 But why wasn't the "Alvey Tools" GDE system used to implement GPSG? The reference to Briscoe et al. (1987) shows that this system was known to the author, yet GPSG is presented as a "non-implemented formalism" in the conclusion (p. 293).

2 See also Baschung (1992) for a shorter presentation of the UCG system and a more general description of FG. Note that the system referred to as PIMPLE from the University of Edinburgh has now become PLEUK; see Calder (1993).

guage (here, French). These principles take into account both the (lexically specified) argumental structure of the matrix verb and the syntactic realization of the arguments (e.g., presence of a content preposition, interrogative structure, raising of an argument, etc.). The starting hypothesis is “the possibility of an infinitive predicate capturing (as subject) any argument available in the matrix clause” (p. 102).³

The chapter introducing control phenomena highlights the fundamentally relational nature of infinitive constructions, which contrast in this respect with finite complement clauses, and underlines the internal aspect of the reconstruction of the referent of their subject. Baschung also introduces here a crucial distinction between obligatory control (PCO principles) and nonobligatory control (PC principles), obligatory control being a particular case of nonobligatory control. The latter is similar to pronominal and anaphoric resolution and remains distinct from arbitrary control (treated as non-obligatory in government-binding (GB) and lexical functional grammar (LFG)).⁴

The author makes a large number of very sharp and pertinent observations—e.g., about prepositional phrases, the fact that the possibility of the noun phrase’s being a controller is conditioned by the nature of the preposition (p. 51); about the semantics of “general condition” predicates such as *il est vrai/faux, il faut* (‘it is true/false/necessary’), that omitting the controller imposes an arbitrary interpretation (p. 89); and the fact that, structurally, a preposition and an interrogative phrase have the same functional value with respect to control (p. 67). However, the general rules are not always convincing; in particular, concerning expletives, the restriction PC2a (p. 38: “An expletive NP (i.e., *il, ce*) or an NP which is necessarily [–HUMAN] is a forbidden controller”) cannot be accepted in the face of examples such as *Il semble se mettre à neiger* (‘It seems to have started to snow’).

Concerning details of the analysis, the claim (p. 60) that Bach’s generalization does not hold in French may be premature. Indeed, if Example (107) is not acceptable in English, neither is its translation into French.

(107) *Mary persuaded to come. (**Mary a persuadé de venir*)

Thus, there is indeed a difference with verbs such as *ordonner*, which accept a phonetically null controller for their infinitive complement. However, Bach’s generalization also has exceptions in English; see Keyser and Roeper (1984) for examples of controllers that are not present at surface structure (e.g., *The boat was sunk to collect the insurance*).

On the other hand, as the correlations observed by Seelbach (1978) do not obtain in a manner regular enough to allow deduction of rules from them, it does seem more appropriate to follow the approach advocated by Baschung and to base the correlation between the position of the noun phrase controlling the infinitive and the mood of the complement upon the pragmatic interpretation of the use of the infinitive as information-providing and of the subjunctive as order-giving.

I will not spend too much time on the presentation of GPSG nor on the grammar that is proposed for it. It is, however, worth noting that the section on infinitive complements highlights some of the unsolved problems in GPSG concerning adjectival constructions of the type *easy to please*. In addition, the implementation in the GPSG of the principles proposed in the beginning of the book, in particular that of PCO2 (“The controller of an infinitive may be any argument which is not explicitly forbidden. The prohibition is lexically specified on the predicate of the matrix clause”), shows

³ All translations of the quotations into English are mine.

⁴ The comparison between these two theories is not very deep, but as it gives rise to some very pertinent remarks, it makes one regret that a comparison with HPSG was not yet possible.

clearly the limitations imposed by the fact that in this framework the only syntactic information available at the level of the lexical entries is the subcategorization number.

One of the general principles of this model is that of functional realization: the form of the semantic rules is automatically determined from the semantic type of the syntactic constituents. In this French GPSG, two operators for semantic combination introduce the schemata for the semantics of Raising and Equi predicates. These predicates are treated through the use of meaning postulates, and the grammatical hierarchy is preserved in the semantic type. (This is motivated by the extension of the treatment of control to subject infinitives.) The two combinatory operators also allow Baschung to account for the control of object-raising adjectives and of (infinitive) indirect questions. The grammar can also treat instances of multiple control, which poses problems not only for building the semantic interpretation, but for syntactic agreement.

In UCG, building the semantic representation is one particular aspect of the general process of functional application. The grammar specifies which argument slot on the argument list is to be filled through the active sign of the categorial list of the functor. Infinitives are not represented as properties, but as open propositions whose subject argument is saturated by the controlling noun phrase.⁵

UCG does not permit an easy distinction between verb phrase modifier and sentential modifier because the indices for verb phrase and subject are the same in simple sentences. This is one of the motivations for the treatment of infinitives that Baschung proposes, and indeed the most important point about the FDP grammar is that an infinitive is treated as a verb phrase whereas its subject remains syntactically accessible, since in categorial grammar a verb phrase is necessarily of category *s/np*. UCG can account directly for PCO1 ("An infinitive must be controlled in its matrix clause") because the mechanism is based on the unification of variables, thus enforcing the strict locality of obligatory control.⁶ The distinction between a nonexistent and an unknown argument, which is introduced at the start of the book and which accounts for the ambiguity of some predicates between their two-place and three-place readings, is elegantly put to use in UCG to treat arbitrary control when the unknown argument is optional.

FG is an "alternative UCG grammar" for French that is largely based on FDP but presents a number of differences, most importantly concerning the treatment of valency in lexical entries and that of the linear order of syntactic constituents. The differences between FDP and FG concerning the treatment of infinitives are mainly theoretical, but have some implications for the implementation. In FDP, verbs are arguments with respect to noun phrases and functors with respect to infinitives. In FG, on the other hand, infinitives as well as noun phrases are type-raised categories, i.e., functors with respect to verbs. Thus in FG, the lexical entry of every verb turns it into a functor and embedded infinitives are treated as regular noun phrases. However, their valency list is not empty, and the subject of the infinitive remains accessible from the lexical entry of the matrix verb. An interesting generalization (p. 67) is that, since embedding (not merely of infinitives) is systematically correlated with type raising, the latter can be seen as a kind of nominalization operator in the grammar. Combinatory rules are also defined differently in FG: the features of the resulting sign are determined partly from the features of the active sign of the functor and partly from the features of the argument.

⁵ In fact it would not be possible to build a propositional function in FDP, because no argument selector is available.

⁶ The nonobligatory control principles (PC2a and PC2b) and the obligatory control principles (PCO2a, PCO2c, and PCO2d) are expressed in FDP only by means of the lexical entries, and not explicitly.

The main concept in FG is that the French linguistic data require that valency be treated not as an ordered list, but as a set. This ties in with another observation, that information leading to concatenation in the grammar is never linked to information concerning linear ordering. Unfortunately, the whole section about erasing in non-adjacency phenomena and about constraint checking in the FG erasing “rule” is rather obscure and deserves clarification. However, it seems to be only a mechanism whose details have no direct relevance except for a Prolog implementation of this particular approach. On the other hand, the remark (p. 258) that erasing should not be treated as a rule, since that would be equivalent to the hidden creation of lexical entries, but rather as a filter to check the well-formedness of signs, deserves further study.

The issue of accessing features independently of the categorial list constitutes one of the motivations for the differences between the FDP grammar—where it is not possible, which leads to false lexical ambiguities—and FG—where it is systematic. FDP and FG also diverge in their treatment of lexical ambiguity: whereas FDP (thereby following GPSG) gives several lexical entries for different subcategorization lists with the same semantics, FG gives different lexical entries for the same word only if different features or different subcategorization lists give rise to different semantic interpretations. In fact, having only one lexical entry for each word is a fundamental principle of FG. Thus the grammar must contain some way of expressing the relations between different valencies. This is the MBU procedure, which is the specification of the result of the unification operation over sets (the valencies) for given items (the functor and the argument) and which plays a crucial role in the determination of control in the grammar. Since the unification operation over sets is not defined in Prolog, this procedure gives a definition for it that allows the expression of the linguistic phenomena to be accounted for. However, this specification cannot be taken as a general principle, a fact that the conclusion of the book underlines. Indeed the hope is expressed that an extension of FG will make it possible to define MBU operations that belong to universal grammar and to distinguish them from MBU operations that are language specific. Nevertheless, the choice of the maximally lexicalist hypothesis (i.e., that of specifying the necessary information in each lexical entry rather than through lexical rules) might result in a more fine-grained description, but as Baschung notes, it also excludes the possibility of stating generalizations in the lexicon, and this can be challenged on methodological grounds.

To conclude, this book not only provides an excellent overview of the UCG framework, but will also prove very useful to researchers working on the treatment of French. The lack of an index is thus to be regretted, since the rather difficult and detailed material being treated requires frequent backtracking. As for the few negative remarks I have to make, I would deplore the number of typographical mistakes. Some are not very damaging (e.g., *Cherchia* instead of *Chierchia* in the bibliography, or example (284), which does not correspond lexically to its extension in (285)), but some impede comprehension of the text. For instance, example (226.b), *Il est triste de partir* (‘It is sad to leave’) with impersonal *il* is given as unacceptable on p. 86, which must be a typo, since in Annex 2, the adjective *triste* correctly allows this construction, as do the GPSG rules (85) and (86) on p. 128. The lack of a translation for some quotations may be a problem, if not for those in English, at least for those in German. Moreover, some ambiguities in the terminology could easily have been clarified (e.g., “raised object” in PCO2d refers to the original position of the argument, not to its landing site).

Acknowledgements

A French version of this review appeared in *Traitement automatique des langues*, 33, 1992. This English version appears by kind permission of Jacques Anis, editor of that journal.

References

- Baschung, Karine (1992). "Un analyseur UCG du français." *T.A. informations*, 32(2):79–88.
- Briscoe, Ted; Grover, Claire; Boguraev, Bran; and Carroll, John (1987). "A formalism and environment for the development of a large grammar of English." In *Proceedings, 10th Joint Conference on Artificial Intelligence (IJCAI-87)*, Karlsruhe, 1987, 703–708.
- Calder, Jo (1993). "Graphical interaction with constraint-based grammars." In *Proceedings, First Conference of the Pacific Association for Computational Linguistics (PACLING 93)*, Vancouver, April 1993, 160–168.
- Calder, Jo; Klein, Ewan; Moens, Marc; and Zeevat, Henk (1986). "Problems of dialogue parsing." In *Syntactic Parser for Dialogue*, ACORD deliverable T2.1. Edinburgh: Centre for Cognitive Science.
- Gazdar, Gerald; Klein, Ewan; Pullum, Geoffrey K.; and Sag, Ivan A. (1985). *Generalized phrase structure grammar*. Cambridge: Harvard University Press.
- Gross, Gaston (1975). *Méthodes en syntaxe. Régime des constructions complétives*. Paris: Hermann.
- Keyser, Samuel Jay, and Roeper, Thomas (1984). "On the middle and ergative constructions in English." *Linguistic Inquiry* 15:381–417.
- Seelbach, Dieter (1978). "De-Infinitiv oder reiner Infinitiv im Französischen—ein Beispiel für die Verwertung von Ergebnissen der linguistischen Grammatik in einer didaktischen Grammatik." *Linguistik und Didaktik* 9:268–282.

Dominique Estival is a research scientist at ISSCO, where, after working on reversible grammars as part of the development of ELU (a linguistic development environment based on unification), her involvement in several European projects now includes work on reusable grammars and on test-suites for natural language processing. Estival's address is ISSCO, Université de Genève, 54 route des Acacias, CH-1227 Genève, Switzerland; E-mail: estival@divsun.unige.ch