

## **ANALYSE ET THÉORIES SYNTAXIQUES**

**Christian TOURATIER, 2005, Coll. Langues et langage, 13, Publication de l'Université de Provence, Aix-en-Provence, France — 328 p. — 30 € — ISBN 2-85399-620-4**

*Note de lecture de Jean-Yves ANTOINE*  
*Mél : jean-yves.antoine@univ-tours.fr*

---

Comme il le rappelle sur la quatrième de couverture, l'objectif de Christian Touratier dans ce livre est de proposer une méthode d'analyse syntaxique qui peut être qualifiée d'objective en ce sens qu'elle ne se base que sur le fonctionnement effectif des langues et en mettant en jeu une réflexion purement syntaxique.

Cette dernière affirmation peut étonner l'étudiant(e) en sciences du langage. Elle est tout de suite clarifiée dans le premier chapitre introductif de l'ouvrage, qui est appelé opportunément « Qu'est-ce que la syntaxe ? ». L'auteur s'y intéresse aux grammaires scolaires, dont on sait qu'elles influencent fortement notre vision commune de la langue. À la suite d'autres auteurs, Christian Touratier rappelle sur des exemples éclairants que là où le grammairien voit une réalité, le linguiste ne retrouve en fait qu'approximation et confusion des genres. En particulier, l'auteur montre que les assertions des grammaires scolaires relèvent bien souvent d'une confusion non assumée entre dimensions syntaxiques et sémantiques du langage. Il en appelle donc à une étude purement syntaxique de la langue qui se limiterait strictement à l'analyse de la structuration des énoncés, en dehors de toute considération de sens.

Le second chapitre présente le cadre méthodologique dans lequel sera mené ce projet : l'analyse en constituants immédiats, adaptée des travaux originaux de Bloomfield (1933). En ce sens, cet ouvrage constitue une curiosité pour le lecteur peu averti, en lui proposant une vision pré-chomskyenne de l'analyse syntaxique du langage. Ce point est d'importance, tant il est vrai que les notions de constituant ou encore de syntagme ont été détournées de leur portée initiale par les générativistes. En particulier, Chomsky a, au fil de ses théories, participé pour beaucoup à l'émergence d'une approche syntaxico-sémantique de la structure des énoncés. Je reviendrai ultérieurement sur la raison de cette confusion. Il est en tous cas étonnant que l'auteur ne se soit pas attardé sur ce point. Sans doute souhaitait-il faire preuve de pédagogie en concentrant son argumentation sur les grammaires scolaires, connues de tous.

Le second chapitre s'essaie donc à une présentation des différentes méthodes d'analyse basées sur les constituants immédiats. En visant l'exhaustivité et le détail des explications, le texte est assurément très lisible pour le néophyte. Il perd par contre en synthèse et en concision pour le lecteur plus averti. Les exemples qui sont étudiés persuadent en tous cas aisément le lecteur qu'il est possible d'établir une représentation structurale en ne se basant que sur des critères purement syntaxiques. L'auteur n'évacue pas les difficultés liées à une telle approche. Il montre

par exemple qu'il est possible de construire plusieurs tableaux de commutations<sup>1</sup> alternatifs pour une phrase donnée. De mon point de vue, il ne s'agit pas d'une limitation de l'analyse en constituants immédiats : c'est précisément en cherchant à réduire ces ambiguïtés syntaxiques que les chercheurs en linguistique computationnelle (pour ne prendre que cet exemple) en sont venus à mélanger dans une seule analyse syntaxe, sémantique voire pragmatique. Les développements récents du TAL robuste (Aït-Mokhtar, Chanod, Roux, 2003) montrent au contraire tout l'intérêt d'une analyse automatique incrémentale où chaque étape d'analyse correspond à une réalité linguistique bien identifiée. Le TAL robuste, qui a permis des avancées significatives en ingénierie des langues, repose entre autres sur la définition de structures sous-spécifiées, c'est-à-dire laissant la place à des alternatives. L'idée est qu'une ambiguïté peut se présenter à un niveau d'analyse particulier sans que cela soit gênant : un niveau supérieur permettra de lever le cas échéant (par exemple par une interprétation sémantique de structures syntaxiques concurrentes). Niveaux d'analyse bien identifiés et cohérents, analyses sous-spécifiées, on voit qu'on retrouve dans le TAL robuste des idées avancées par l'auteur dans une démarche purement linguistique. J'y reviendrai en conclusion.

Ce second chapitre est également l'occasion de faire le point entre ordonnancement linéaire et dimension structurelle des énoncés. Ce problème est souvent ignoré en linguistique mais surtout en TALN, où les approches sont le plus souvent d'origine anglo-saxonne. Ici, l'auteur (latiniste reconnu) donne des exemples sur des langues à ordre variable (latin, russe) qui rappellent opportunément les propos de Covington pour qui l'anglais est l'idiome le moins général qui soit, du fait de son extrême rigidité d'ordonnancement linéaire. Là encore, le TALN est arrivé au même constat, ce qui explique que les chercheurs coréens, finnois ou russes aient accordé une place bien plus importante à Tesnière qu'à Chomsky dans leurs travaux. L'auteur rappelle ainsi que la confusion entre ordre linéaire et ordre structurel opérée par les générativistes a conduit à une dérive connotationnelle du mot syntagmatique qui est devenu au cours du temps un quasi-synonyme du qualificatif syntaxique.

Avec le chapitre 3 (Exercices d'analyse en « constituants immédiats »), l'auteur propose une approche méthodologique et pratique de l'approche présentée. Ces exercices permettent de mieux préciser la démarche utilisée. Citons par exemple le *distinguo* fait entre l'usage du pronom personnel « il » et d'un syntagme nominal tel que « le petit prince ». En étudiant leurs possibilités de commutation, l'auteur montre que ces deux éléments n'appartiennent pas au même paradigme. Comme l'avait noté M. Gross, le pronom personnel doit être vu comme une particule préverbiale qui est liée comme constituant immédiat (CI) au verbe, et ne forme pas un CI de phrase. Cette distinction n'est pas évidente pour le lecteur francophone, notre culture grammaticale nous ayant conduit à mettre sur le même plan les deux structures, car elles partagent les mêmes fonctions dans la phrase. L'auteur rappelle alors que cette identité de fonction est trompeuse, car de nature sémantique et non syntaxique. Des exemples latins achèvent de nous en persuader : le latin exprime lui, explicitement, la divergence de rôle relevée.

Les chapitres suivants sont consacrés à l'étude détaillée de la syntaxe de la langue dans l'approche en constituants minimaux. On aborde tout d'abord la construction

---

<sup>1</sup> Le tableau de commutation représente la succession des regroupements qui sont opérés au cours de l'analyse en constituants immédiats. *In fine*, il représente la structure syntaxique de l'énoncé analysé.

des syntagmes nominaux, verbaux et prépositionnels (chapitre 4), on revient sur les fonctions syntaxiques (chapitre 5) puis on s'attaque dans les deux chapitres suivants à la phrase complexe avec propositions subordonnées. Un dernier chapitre est enfin consacré à la coordination, objet d'étude en linguistique s'il en est ! On regrettera à ce sujet qu'un parallèle ne soit pas fait entre la coordination et les phénomènes d'entassement paradigmatique que l'on rencontre à l'oral spontané (Guénot, 2006).

Au final, cet ouvrage est l'œuvre d'un puriste qui, méthodologie à l'appui, nous montre (ou nous rappelle) qu'il est possible d'analyser la structure des énoncés sur des critères purement syntaxiques. Sans tenir lieu de manifeste, l'ouvrage se veut au contraire suffisamment précis et pédagogique pour intéresser l'étudiant en sciences du langage. Si je regrette parfois que le niveau de détail de l'explication nuise au fil de l'exposé, il est certain que le lecteur trouvera dans ces pages les réponses à ses questions.

Le reproche principal que je ferais à l'auteur est de s'être limité à une perspective purement linguistique. Comme je l'ai déjà fait remarquer, ses propositions raisonnent avec une certaine pertinence dans les évolutions les plus récentes du TALN. Ce qui, au passage, est assez surprenant pour un ouvrage dont la plupart des références bibliographiques semblent être arrêtées aux années 60-70 ! Il est donc regrettable que l'auteur ne se soit pas plus intéressé aux conséquences de ses propositions en ingénierie des langues. Je sais que ce n'est pas son propos. Il se refuse ainsi explicitement à intégrer « toute théorie venue de l'extérieur » (de la logique ou des mathématiques par exemple) ». Ces « théories » ne sont pourtant que des moyens de formalisation au service de l'opérationnalisation de l'étude de la langue. Et si Chomsky s'est très tôt intéressé à cette formalisation, celle-ci n'est pas, ou si peu, la cause des « dérives » de sa théorie linguistique. Je regrette donc que l'auteur, comme nombre de ses collègues linguistes, ignore la force de validation des théories que représente le TALN. En délaissant une démarche expérimentale pour s'en remettre à la seule introspection linguistique, l'auteur risque de voir, une fois de plus, ses propositions égarées dans les affligeantes querelles de chapelles qui rythment l'évolution ou plutôt les sur-places<sup>2</sup> de la pensée linguistique. Assurément, ces travaux méritent mieux.

## Références bibliographiques

Aït-Mokhtar S., Chanod J. P., Roux C. (2003) Robustness beyond shallowness: incremental deep parsing. *Natural Language Engineering*. 8 (2-3), 121-144.

Bloomfield L. (1933) Le langage. Traduction de Janick Gazio (1961), Payot, Paris, France.

Covington. M. A. (1990) A dependency parser for variable-word-order languages. *Technical Report AI-1990-01*, University of Georgia, Athens, GA.

---

<sup>2</sup> La lecture de l'ouvrage est à ce titre illustratif de ces sur-places : l'auteur prend en effet un soin méticuleux à lister l'ensemble des dénominations qu'ont pu trouver certains linguistes pour décrire la même notion. Certes, il y a souvent nuance entre telle ou telle acceptation d'un terme. Mais ces nuances méritent-elles une telle considération ? En s'y attardant, est-on encore dans la démarche scientifique, ou simplement dans le plaisir dialectique. Mes propos, qui sont ceux d'un « traicteur de langue », sont sans doute facilement réducteurs. J'y vois tout de même une part de vérité...

Guénot M. L. (2006) La coordination considérée comme un entassement paradigmatique : description, représentation et intégration. In *Actes de TALN2006*. Presses Universitaires de Louvain, Belgique, 10-13 avril 2006

---

## L'auteur de la revue critique



Jean-Yves Antoine a obtenu un doctorat en informatique en 1994 (Institut National Polytechnique de Grenoble, INPG, France) ainsi qu'une Habilitation à Diriger les Recherche en informatique soutenue en 2003 (Université de Bretagne Sud, UBS, Vannes, France). De 1996 à 2004, il a été Maître de Conférences à l'Université de Bretagne Sud, où il dirigeait l'équipe de recherche CORAIL en ingénierie des langues au sein du laboratoire VALORIA. Il est depuis 2004 Professeur des Universités de l'Université François-Rabelais (Tours-Blois, France) où il poursuit ses recherches au sein du Laboratoire d'Informatique (LI). Ses travaux concernent le dialogue oral homme-machine, les systèmes d'aide à la communication pour personnes handicapées, l'extraction de terminologie et la linguistique de corpus. Il est également membre associé de l'équipe de recherche Langue et Représentation de l'Université François-Rabelais.