

nard Alane laisse tomber ce côté un peu emprunté», KJD voit la construction à cohésion forte, alors que je pense qu'on ne peut pas se décider entre les deux constructions. (3) KJD suppose que dans la construction à cohésion forte, *rien*, sujet logique de l'infinitif, suit toujours l'infinitif, de sorte qu'une proposition où *rien* précède l'infinitif représente toujours la construction à cohésion faible (la grammaire citée n'est pas explicite sur ce point, ses implications semblent se contredire). Mais même dans la construction à cohésion forte, *rien* précède l'infinitif à moins d'être emphatique. Ainsi, dans la phrase «Heinrich Böll était un observateur lucide qui ne laissait rien passer», KJD voit la construction à cohésion faible, alors que je pense qu'on ne peut pas se décider entre les deux constructions. KJD cite d'ailleurs dans sa bibliographie un travail où j'avais exprimé les avis que je viens de reprendre. La répartition adoptée par KJD l'amène à conclure que la grammaire citée a tort de dire que la construction à cohésion faible est beaucoup plus fréquente que la construction à cohésion forte si le verbe régissant est *laisser*. En effet, cette remarque n'est justifiée que pour les cas où le sujet logique de l'infinitif est exprimé. KJD ajoute que le résultat de sa petite enquête «nous apprend à nous méfier des autorités». On ne saurait qu'approuver cette leçon, en regrettant que KJD ne se soit pas mieux méfié de la même autorité.

C'est également en se méfiant trop peu de cette autorité que KJD dit que dans «Je n'ai pas de vin», *de* est l'article partitif réduit, bien que cette idée ait été contestée par David Gaatone, que KJD cite dans sa bibliographie, et d'autres, avec des conclusions différentes, que je ne reprends pas ici. KJD ne dit pas qu'il ne considère que des substantifs qui sont soit du singulier non-nombrable soit du pluriel, en négligeant ceux qui sont du singulier nombrable. Il y en a quelques-uns parmi les exemples de *pas de* qu'il a extraits d'un corpus informatisé: «actuellement il n'a pas de position précise». L'exposé de KJD implique et présuppose que sans la négation, même ces exemples-là auraient l'article partitif, ce qui n'est pas exact.

On voit par ce qui précède qu'une des vertus du livre de KJD est d'inviter à la discussion et à la contestation. Y a-t-il une meilleure preuve des qualités pédagogiques d'un livre?

Povl Skårup
Université d'Århus

Eveline Martin : *Reconnaissance de contextes thématiques dans un corpus textuel. Éléments de lexico-sémantique.* Collection 'Études de sémantique lexicale'. Didier Érudition, Paris, 1993. 283 p.

Avec le développement explosif de bases de textes auquel on a assisté ces dernières années, l'élaboration de logiciels pour leur exploration intelligente devient urgente. En effet, si l'on désire faire de cette masse de textes informatisés un outil intéressant pour les recherches, il faut développer des méthodes de dépouillement automatique qui dépassent la simple reconnaissance des formes. Le travail effectué par Eveline Martin constitue un progrès considérable dans cette direction. Ingénieur de recherches à l'INaLf, elle a accès à la BASE du TLF (le Trésor de la Langue Française), qui est la plus grande base textuelle de ce genre qui existe actuellement, ce qui lui a permis de vérifier constamment ses idées et l'efficacité de son système. Dans son

ouvrage, qui est une version abrégée de sa thèse, elle expose d'abord les problèmes théoriques et pratiques qu'implique la création d'un système de reconnaissance structurée et automatique de contextes thématiques, puis elle présente un exemple d'utilisation pour finir par une discussion du potentiel de la Base.

Quelques principes fondamentaux doivent sous-tendre l'élaboration du système en question : il doit être assez souple et dynamique pour permettre une mise à jour permanente, il doit être facile à appliquer pour l'utilisateur et l'expert doit toujours intervenir pour garantir et contrôler le perfectionnement des structures lexico-sémantiques qui constituent le cœur de la Base. Tout d'abord, il faut concevoir cette structure. Inspirée par les recherches cognitives ainsi que par les systèmes d'experts, l'auteur opte pour un modèle qui «sera constitué de champs lexicaux, ou champs thématiques, désignés chacun par un intitulé ou mot-pôle, et constitué par une collection structurée de corrélats» (p. 25). Le champ sémantique est organisé comme un graphe où le mot-pôle et ses corrélats constituent les nœuds qui sont joints par des arcs symbolisant divers types de relations logico-sémantiques. Le problème est dès lors de choisir pour chaque mot-pôle ses corrélats et de déterminer la nature des relations reliant ceux-ci à celui-là. C'est l'expert qui est chargé de ce travail, pour lequel il aura recours à deux sortes de sources : les données secondaires constituées par les divers dictionnaires (y compris les dictionnaires thématiques – il est, en fait, intéressant de voir comment des travaux philologiques conçus pour tout autre chose seront remis en honneur pour ce travail en informatique) et les données primaires qui résultent de l'exploitation des textes eux-mêmes. Éveline Martin aborde de nombreux problèmes associés à chacune de ces deux méthodes : dans quelle mesure les dictionnaires font-ils autorité? faut-il inclure les antonymes? que faire des métaphores? combien de niveaux inclure dans les structures? comment choisir les corrélats? est-il possible d'appliquer des critères qui ne soient pas purement quantitatifs? etc. On peut ressentir parfois une certaine hésitation devant ces questions. D'une part, l'auteur fait valoir que «inclure dans la Base l'antonymie au même titre que l'équivalence surchargerait inutilement le Système» (p. 59), et d'autre part, elle reconnaît l'importance que peuvent avoir les antonymes pour l'étude d'un champ thématique donné. D'une manière générale, il est évident qu'il reste un certain nombre de problèmes à résoudre, notamment en ce qui concerne la précision de la nature des relations qui relient les corrélats au mot-pôle. L'auteur en discute longuement et se décide pour un inventaire incluant les relations d'équivalence, de synecdoque et de métonymie, notions qui subsument les types de relations que l'on «rencontre le plus souvent en logique, dans la grammaire des cas et plus généralement lors de l'élaboration des langues formalisées» (p. 44). Dans le système, ces relateurs de concepts seront spécifiés davantage, ce qui permet à l'utilisateur d'accomplir des explorations plus subtiles.

Étant donné que la Base du TLF comprend surtout des textes littéraires, le système développé par Éveline Martin se prête d'abord aux études d'aspect littéraire : en effet, pour une étude d'une certaine thématique chez un certain auteur ou pendant une certaine période, la Base sera un outil très adapté ; l'analyse exemplaire que propose l'auteur du champ de l'*ennui* (qui n'est certainement pas ennuyeux tel qu'il est présenté ici) le montre bien. Or l'auteur mentionne bien d'autres emplois potentiels de la Base. Ainsi elle esquisse la manière dont on peut faire une critique métalxicographique du dictionnaire le *Grand Robert*, elle indique comment est facilitée l'étude des collocations si importante pour l'apprentissage du français langue étran-

gère, et elle effleure les applications imaginables à la traduction automatique, où la Base pourrait fonctionner comme une instance de «désambiguïsation». A cela s'ajoutent des exploitations évidentes de la Base en lexicographie informatisée, pour le traitement semi-automatique de l'information et peut-être même, à plus long terme, pour la génération automatique de textes en langue naturelle. Les utilisations virtuelles semblent, en fait, être légion.

Ce livre soulève beaucoup de questions qui incitent à la réflexion. Grâce à un style sobre et à un bel équilibre dans la présentation entre théorie et pratique, il constitue une excellente introduction à cette nouvelle branche de la linguistique computationnelle qui est à cheval sur la lexicographie, l'informatique, l'intelligence artificielle et, même, les études littéraires. Parfois, on peut regretter que la prudence de l'auteur laisse certaines questions un peu dans l'ombre. Ainsi, j'aurais aimé une discussion plus approfondie du problème de la représentativité du corpus, question extrêmement difficile et pourtant cruciale pour l'évaluation de l'utilité de la Base. En effet, quelles sont les conséquences, par exemple, de l'intégration prévue d'une plus grande quantité de textes non littéraires? De même, de quelle sorte seraient les dispositifs de contrôle qui permettraient à la Base de s'auto-enrichir (p. 190)? Mais on comprend : le traitement intelligent automatique d'un corpus textuel est un sujet en pointe et en rapide évolution ; seul l'avenir pourra nous montrer quelle direction il est susceptible de prendre. Si l'on désire prendre le train en marche, le livre d'Éveline Martin constitue une excellente introduction sans prétention, mais qui arrive néanmoins à indiquer de multiples utilisations virtuelles et de nombreuses pistes de recherche. Si l'on peut conclure, me semble-t-il, que cette œuvre soulève plus de questions qu'elle n'en résout, il faut souligner que c'est là un aspect très positif du travail. En effet, la science n'avance qu'en se posant sans cesse de nouvelles questions.

Henning Nølke

Ecole des Hautes Etudes Commerciales de Århus

Robert Martin : *Pour une logique du sens*. (2^e édition revue et augmentée.) 'Linguistique nouvelle'. Presses Universitaires de France, Paris, 1992. 319 p.

Rares sont les linguistes qui se voient sollicités pour republier leurs travaux majeurs. Avec la parution de la deuxième édition revue et augmentée de son *Pour une logique du sens*, Robert Martin entre à ce club des illustres. L'auteur a profité de cette occasion unique pour tenir compte de l'évolution qu'a connue la sémantique depuis 1983, date de la première édition (dont j'ai déjà eu le plaisir de faire un compte rendu dans la *Revue Romane* 19,2, 1984 : 328-332). La deuxième édition contient, en effet, notamment une mise à jour de l'appareil théorique. Profitant de l'élaboration de sa propre théorie, telle qu'il l'a présentée dans son œuvre *Langage et croyance* (datant de 1987 ; voir le compte rendu dans *Revue Romane* 23,2, 1988 : 283-287), Martin a réécrit le premier chapitre qui expose les notions fondamentales. Le résultat en est que le cadre des études présentées dans les autres chapitres, qui n'ont été soumises qu'à quelques rares modifications pour le fond, est devenu encore plus homogène : les analyses y ont gagné en clarté. On peut cependant regretter que Martin n'ait pas allégé en même temps son emploi des symboles. Malgré son affirmation (nouvelle par rapport à la première version) selon laquelle son travail est sans prétentions formali-