

Une analyse cognitive des interrogations d'une banque de données

Éric GILLET

OSMONT (Béatrice) : 1995, *Dynamiques cognitives et stratégies d'utilisateurs* (Paris : Masson, Collection « Sciences cognitives »), 186 pp. ISBN 2-225-84688-X.

Béatrice Osmont nous offre un petit livre très intéressant et original à plus d'un titre. D'abord, il s'agit d'une étude, à l'intersection de la linguistique et des sciences cognitives, prenant comme point de départ un corpus totalement authentique et vierge de biais méthodologiques. Il s'agit en effet de traces laissées par les interrogations d'une banque de données sur Minitel destinée aux travailleurs sociaux; les utilisateurs sont anonymes et l'analyse n'est aucunement influencée par leurs intentions par ailleurs inconnues. Ensuite, l'auteur se propose d'analyser ce corpus sans en éliminer la moindre donnée : les répétitions, les corrections, les hésitations, les annulations, rien n'est écarté ni considéré comme négligeable. Cette approche se révèle à la fois exigeante quant à la méthode et fructueuse dans l'élucidation des mécanismes de sens. Enfin, et plus fondamentalement, cette méthode débouche sur une double analyse de la configuration du lexique et de la dynamique des interrogations qui conduit l'auteur à proposer des outils théoriques de représentation neufs dans une perspective plus large d'économie cognitive.

La première partie de l'ouvrage situe le corpus et les contraintes liées à l'utilisation de la banque de données; une première analyse des données brutes conduit à un découpage du corpus en unités significatives sur lesquelles porteront les analyses ultérieures. Elle se conclut par un chapitre théorique qui critique la notion même de dialogue homme-machine (à cette expression, Béatrice Osmont réagit par un triple rejet (p. 31) : ni dialogue, ni homme, ni machine, elle nous propose de parler de relation utilisateur-ordinateur) et fixe le canevas théorique qui sous-tend le reste de l'étude.

✉ Université de Liège; Séminaire de Logique et d'Épistémologie; place du 20-Août, 32; B-4000 Liège (Belgique).
Fax : +32 41 66 55 59 E-mail : egillet@vm1.ulg.ac.be

D'une part, s'appuyant sur une distinction de Vygotski, l'auteur remplace l'illusion d'un dialogue par la présence, face aux données de la machine, d'éléments de surface émergeant du travail dans la pensée d'un *langage intérieur*¹. Et c'est, bien sûr, de la dynamique cognitive à l'œuvre dans ce langage intérieur qu'il faut rendre compte à partir des traces laissées par le pseudo-dialogue. D'autre part, tout en situant son approche dans le cadre de l'étude de la fonction de recherche du langage, elle reprend la position de Grunig² pour tenter, non la prévision de la formulation d'une requête, mais l'explication de la complexité de la démarche d'interrogation. À cet effet, elle s'attachera à décrire les logiques procédurales des interrogations, les interactions avec les données visualisées et les processus cognitifs sous-jacents organisés en sous-systèmes.

La seconde partie de l'ouvrage, introduite par une présentation des outils utilisés, est constituée par deux longues et rigoureuses analyses portant l'une sur le lexique, l'autre sur la dynamique des interrogations. Le postulat de départ est que chaque terme du lexique est une trace émergeant d'un processus cognitif, différent du processus à l'œuvre lors de la constitution de la banque de données et que chaque terme est *produit* de façon non-aléatoire au sein de la structure de connaissance particulière engagée dans l'interrogation. L'auteur postule que la théorie que Vygotski propose pour la formation des concepts s'applique ici pour l'étude de la dynamique d'une recherche. Elle lui reprend, notamment, le concept de « Zone Proximale de Développement » pour expliquer comment un utilisateur, manifestement dépourvu d'un *script* complet de recherche, peut néanmoins construire une interaction. Munie de ces concepts, elle propose une méthode de représentation figurative des structures de connaissances qui fait droit à la dynamique des interrogations ainsi qu'aux traits cognitifs et socio-linguistiques particuliers de son corpus. Appliquée au lexique, cette méthode lui permet de dégager d'une part, une catégorisation basée sur le niveau de langage, séparant les traits de sens commun, les traits de problématique et les traits désignant une structure sociale, et d'autre part, une schématisation où elle distingue, entre autres, l'action et ses différents arguments (suivant en cela la classification de Fillmore³), les personnes, les lieux, les traits modifiant le champ d'autres traits, comme la particularisation, etc. L'analyse porte ensuite sur la dynamique de l'interrogation et Béatrice Osmont développe la notion de pré-représentation pour faire droit aux schématisations partielles et ponctuelles, en évolution suite aux interactions avec la banque de données. Les opérations entre les termes appartenant à une pré-représentation se répartissent en opérations de révision (ou *palinérèse*), de constriction (ou *synérèse*) et d'expansion (ou *diarèse*); chacune de ces opérations, selon qu'elle porte sur des termes de différentes catégories, nous renseigne sur les pré-représentations de l'utilisateur.

¹ VYGOTSKI (L.S.) : 1985 [édit. originale : 1934, *Myslenie i rec* (Moscou-Leningrad, Sozekgiz)], *Pensée et langage*, Trad. de Françoise Sève, Paris : Éditions Sociales.

² GRUNIG (B.-N.) et GRUNIG (R.) : 1985, *La fuite du sens, la construction du sens dans l'interlocution* (Paris, Hatier-Crédif).

³ FILLMORE (Ch.) : 1968, "The Case for Case", in BACH (E.) & HARMS (R.T.) eds (N.Y. Holt, Rinehart & Winston) pp. 1-88.

Les conclusions de l'auteur insistent sur l'intérêt à la fois méthodologique et théorique de son travail. Les conclusions pratiques mettent en évidence les différences cognitives entre la phase d'encodage et d'indexation des données et la phase de recherche. Une meilleure connaissance de ces différences devrait permettre une organisation plus adéquate des données et une prise en compte des heuristiques lors des visualisations intermédiaires. D'un point de vue théorique, Béatrice Osmont conclut à l'avantage d'une représentation dynamique et propose des pistes de réflexion pour dépasser les difficultés de figuration des pré-représentations. Ainsi envisage-t-elle l'ébauche d'une « logique naturelle » reposant sur des principes simples conformes aux fondements neurobiologiques.

Ce petit livre intéressera très certainement tout qui se penche sur la question des banques de données et plus particulièrement sur le problème de la recherche d'informations pertinentes. L'apport de Béatrice Osmont est à ce titre remarquable : son analyse fine et rigoureuse d'un corpus d'interrogations révèle des aspects linguistiques et cognitifs déterminants pour la constitution des banques de données et pour une heuristique de consultation. Les moyens théoriques qu'elle se donne remplissent parfaitement leur rôle, à savoir la mise en évidence et l'explicitation des représentations cognitives sous-jacentes à la recherche d'informations. On regrettera cependant une présentation peu soignée, surtout au niveau des figures : qu'une figure soit annoncée à une page suivante quand elle est sous nos yeux n'est pas gênant, mais le lecteur reste perplexe quand une légende n'est qu'à moitié imprimée ou que le nom d'un axe de coordonnée sur un graphique ne correspond pas à celui dans le texte; ajoutons à cela que les illustrations sont souvent moins explicites que le texte qui, lui, reste toujours parfaitement clair. Une critique plus fondamentale nous conduit à résister lorsque le texte passe de l'analyse à une perspective exclusivement théorique : Béatrice Osmont maîtrise parfaitement son ouvrage, mais sa rigueur se relâche quand elle s'avance et conclut sur le plan théorique; d'ailleurs, rien dans la structure de son texte ne laisse entendre qu'elle s'attache à construire de nouveaux concepts théoriques à portée générale; elle précise toujours ses sources et si elle les questionne, c'est exclusivement au service de son analyse. Nous devons considérer que dans ses propositions théoriques, elle se laisse emporter et présente comme des avancées ce que seule une analyse comparable à la sienne pouvait mettre en évidence, à savoir des structures incontournables. Précisons que cette impétuosité théorique ne nuit jamais à son analyse, si ce n'est peut-être lors du dernier chapitre consacré à la représentation de la dynamique des interrogations où la place accordée à la justification du concept de pré-représentation l'empêche de mettre suffisamment en évidence les conclusions que l'on peut tirer de son analyse. Bref, il s'agit d'une analyse d'une grande rigueur qui met en lumière tant les processus cognitifs à l'œuvre dans une recherche d'information que les stratégies ergonomiques possibles lors de la création d'une banque de données.