

Présentation du livre Quantitative Semiotic Analysis (Compagno D., dir., Springer, 2018) Dario Compagno

Bonjour et bienvenue à cette présentation du livre collectif *Quantitative Semiotic Analysis* paru chez Springer. Je vais commencer par introduire rapidement sa genèse et finalité, pour ensuite passer la parole à la table ronde avec certains des contributeurs à l'ouvrage. A la fin de la table ronde, j'espère que vous resterez pour boire un petit verre de vin juste à l'extérieur de cette salle et échanger de manière plus informelle.

L'idée de ce livre naît à l'occasion d'une journée d'études sur les méthodes quantitatives en sémiotique, qui eut lieu il y a 2 ans à l'Institut des sciences de la communication ici à Paris. La journée avait été financée par le projet ANR Eneid dirigé par Fanny Georges. A la fin de la journée Alessandro Sarti m'a proposé de diriger un ouvrage sur le même sujet pour sa collection *Lecture notes in Morphogenesis*, proposition que j'ai accepté avec beaucoup d'enthousiasme. Fanny était aussi contente de l'idée et le livre a en fait été en partie financé par l'ANR Eneid. J'en profite alors pour remercier d'or et déjà Fanny et Alessandro qui ont rendu possible cet ouvrage collectif.

Mais quelle était la raison pour organiser une journée d'études sur les méthodes quantitatives dans une discipline comme la sémiotique qui est éminemment qualitative ? En fait, depuis que je suis arrivé en France, j'ai commencé à explorer la possibilité d'utiliser des outils computationnels pour réaliser des analyses sémiotiques. Pour mes expérimentations, un peu hérétiques, j'avais trouvé un accueil exceptionnel au sein du laboratoire IRMECCEN de l'Université Sorbonne Nouvelle, dans une ambiance d'authentique ouverture et interdisciplinarité. Or je savais qu'il y avait d'autres personnes qui faisaient de la sémiotique quantitative, mais j'avais du mal à voir clairement dans ce nouveau domaine. La journée d'études visait alors à identifier des questions de recherche de pertinence sémiotique qui sont effectivement abordées avec des outils quantitatifs, et à comparer les expériences pour voir plus loin, vers des directions collectives de recherche. Le livre est le résultat plus structuré du même besoin.

La sémiotique, si l'on la prend du bas, est une discipline empirique fondée sur des méthodes hérités des sciences du langage, de la critique d'art et littéraire et de l'anthropologie ; si l'on la prend au contraire du haut, la sémiotique est un domaine bien défini de la philosophie, proche à la logique. Dans les deux cas, son objet est le sens, qui est traditionnellement vu comme coextensif au qualitatif. Le sens est ce que la statistique ne devrait pas pouvoir saisir, ce qui ne se laisse pas attraper par des chiffres. Cependant, de plus en plus de chercheurs commencent à mettre en question cette idée. Si c'est vrai que le sens a besoin d'être constitué du haut et du bas, comme concept et comme réalité sociale, rien n'empêche alors de s'ancrer à l'empirie avec des dispositifs scientifiques à jour et performants – en incluant donc des méthodes quantitatives ou quali-quantitatives.

La préférence de la sémiotique traditionnelle pour les méthodes qualitatives s'explique aussi à partir de sa manière de concevoir le sens. Dès son origine, la sémiotique ancienne comme moderne s'est en fait fondée sur une conception autonomiste du langage. C'est-à-dire que l'on voit le langage, et plus en général le sens, comme un phénomène unique qui se distingue nettement du milieu physique et psychologique duquel il naît, et qui pourrait donc être étudié de manière indépendante.

J'aime voir dans les mots d'un sicilien comme moi la première définition du sémiotique comme domaine autonome. Gorgias, le sophiste et inventeur de la rhétorique, est connu surtout pour une suite de trois arguments sur l'inaccessibilité du réel à l'homme. Premièrement, il prouvait, rien n'existe, parce que l'être est en soi contradictoire, et déjà ce premier obstacle *objectif* devrait résoudre le problème à la racine. Mais supposons pour un moment que quelque chose existait de toute manière, elle ne serait quand même pas connaissable par l'homme, parce que notre perception nous trompe. Ce deuxième obstacle *subjectif* invoquant les limites de la connaissance était déjà assez original à l'époque. Mais c'est dans le troisième argument de Gorgias qu'on peut retrouver la naissance de la sémiotique, de la réflexion philosophique donc qui accompagne la conscience du pouvoir du langage. Pour Gorgias, même si quelque chose existait, et même si quelqu'un pouvait y accéder, juste pour un instant, bien, il ne pourrait de toute manière pas communiquer sa découverte aux autres. Un prophète pourrait bien saisir l'être, à travers une révélation directe, mais il serait obligé d'en parler à des non prophètes, avec un langage ordinaire, qui n'est pas fait pour se référer ni à la réalité objective ni aux expériences subjectives privées des individus. Pour Gorgias le langage *est* le monde que nous vivons, nous vivons dans l'immanence d'un monde culturel, public et partagé, et nous ne pouvons pas retourner aux choses mêmes, telles qu'elles seraient avant toute société.

Voici donc la première définition de trois mondes – la nature, la perception individuelle et la communication – qui structurent pour le meilleur et pour le pire encore aujourd'hui notre manière d'analyser l'expérience humaine. Seulement, on ne croit plus comme le faisait Gorgias qu'on ne puisse pas accéder aux deux premiers mondes "primitifs". Bien au contraire, ce que semblait à Gorgias de plus difficile accès est aujourd'hui ce que l'on comprend le mieux et vice versa. Les sciences dures ont eu leur constitution paradigmatique avec Galileo et Descartes ; les sciences psychologiques voient aujourd'hui des importants progrès grâce à la neurophysiologie. Finalement, ce sont les sciences humaines et sociales, fondées sur la parole et donc réellement gorgéennes, qui fatiguent à sortir du préparadigmatique. Mais s'agit-il d'une limite insurmontable ou tout simplement d'un problème de complexité ?

Le philosophe Karl Popper a repris de Gorgias la différenciation de l'expérience humaine en trois mondes, et pour Popper rien n'empêche de constituer une véritable science du troisième monde, de toutes les formes d'expression humaine et de leurs interprétations. Les entités sémiotiques, bien que potentiellement fictives, peuvent être étudiées objectivement. En effet, nous utilisons le langage pour entrer en négociation avec les choses et nos perceptions d'elles. De cette manière, les mots acquièrent une réalité, soit elle autonome ou au contraire dépendante des perceptions et des choses qui la précèdent.

Voici alors la grande option de recherche qui structure les études sémiotiques actuelles, au sens large. Cette même option structure ainsi notre livre en deux parties. En fait, en admettant le potentiel heuristique du quantitatif pour l'étude du sens, devons nous nous limiter à traiter des données culturelles, dimension propre et pour ainsi dire "interne" du sémiotique, ou au contraire il est nécessaire d'élargir le domaine de l'enquête aux seuils de l'appareil biologique et cognitif qui est la condition pour ainsi dire "externe" d'existence du sens ?

Le quantitatif est quant à cela absolument agnostique. Ses potentialités pour l'étude du sens valent dans les deux cas, que l'on veuille perfectionner l'analyse culturelle et immanente ou étudier plutôt le lien du langage avec le reste de l'homme, son corps. D'abord, le quantitatif promet de surpasser la plus grande limite de l'analyse sémiotique qualitative : l'introspection. En fait, l'analyse sémiotique risque toujours de n'être qu'une description éduquée du sens tel qu'il est saisi par un sujet idéal. Sujet idéal comme le veut Roland Barthes, pure variation insoumise à toute autorité, donc le meilleur des critiques possibles ; ou encore sujet idéal comme le veut Algirdas Greimas, donc instance transcendantale qui constitue le sens en-deçà de toute variation et interprétation ; sujet idéal ou quasi-idéal enfin comme le veut Umberto Eco, compétence pragmatique nécessaire pour bien interpréter un certain signe ou texte. Le sujet de la sémiotique n'est jamais natif et naïf comme celui de la linguistique chomskienne, au contraire, mais il reste toujours trop proche de l'analyste. Une statistique du sens et de l'interprétation pourrait alors nous permettre de faire une fois pour toute du sujet sémiotique l'interface entre notre lecture d'un texte et sa vie sociale.

Les autres apports du quantitatif à la sémiotique sont également importants. La possibilité de falsifier nos modèles, et par conséquent de formuler des hypothèses de plus en plus précises et circonstanciées. La possibilité de généraliser et de ne pas donc chercher une langue différente pour chaque texte. La possibilité de travailler par induction, sur des grandes bases de données intraitables autrement. La possibilité enfin de produire des technologies sémiotiques et sociales fondées sur nos découvertes, avant que ça soit Facebook ou le gouvernement chinois à le faire à notre place. N'oublions pas que si Gorgias enseignait la rhétorique pour faire gagner des procès, déjà Aristote y voyait plutôt une arme de défense du citoyen éduqué. Personnellement, il ne me semble pas envisageable d'entreprendre une nouvelle révolte des boxers vis-à-vis du big data, en fait, je ne crois pas que la seule réflexion critique nous rende invulnérable à la prédiction.

Or, si le quantitatif est agnostique à propos de l'immanence, c'est à dire de l'autonomie du langage par rapport à la psychologie et à l'ontologie, les chercheurs qui ont contribué à l'ouvrage ne le sont pas forcément. Bien au contraire cet échantillon des études existantes montre un clair seuil sémiotique. D'un côté de ce seuil on trouve Jean Petitot et Alessandro Sarti du laboratoire de mathématiques sociales de l'EHESS. Petitot et Sarti visent à naturaliser le sens. Naturaliser ne veut pas dire banaliser ni réduire au physique, il veut dire accepter l'idée que le sens est un phénomène naturel en continuité avec ceux étudiés par la psychologie, la biologie et la physique – disciplines avec qui la sémiotique devrait dialoguer. De l'autre côté du seuil on trouve François Rastier et Mathieu Valette de l'équipe textes, informatique et multilinguisme de l'INALCO. Rastier et Valette défendent l'autonomie disciplinaire des sciences du langage et l'adoption d'outils computationnels pour renforcer leur prise empirique. Ils promeuvent eux aussi un dialogue interdisciplinaire, mais avec les autres sciences de la culture et l'informatique, plutôt qu'avec les sciences de la nature.

Il n'est peut-être pas inutile de remarquer que Petitot et Rastier ont les deux participé au séminaire de sémiotique générative de Greimas à Paris et au séminaire de sémiotique interprétative d'Eco à Bologne. De ce fait, leurs travaux ont contribué à établir les fortes connexions qui caractérisent la koinè sémiotique contemporaine. Il n'est donc pas étonnant de voir dans les autres contributions au volume une influence croisée et complexe de ces deux approches au sens.

Céline Cholet travaille sur l'interprétation des images scientifiques à partir de données obtenues avec l'eye-tracking, un instrument qui permet de mesurer les mouvements des yeux. On pense donc toute suite à une naturalisation des parcours du regard, mais Céline cite plutôt la théorie interprétative de Rastier pour comprendre les étapes de la lecture.

Sara Spinelli utilise la sémiotique pour concevoir, administrer et ensuite pour interpréter des questionnaires sur le goût. Elle se sert de la statistique textuelle et travaille donc avec des mots, dans l'immanence, mais à la fin de son chapitre elle introduit une variable physiologique pour mieux interpréter ses résultats : le nombre de papilles fongiformes présentes sur la langue des sujets.

Jimmy Bordarie est un psychologue sociale qui étudie les représentations collectives. Mais pour faire cela il fait recours aux instruments de la théorie des blocs sémantiques développée par Marion Carel, que je remercie beaucoup de sa présence, résolument éloignée de toute préoccupation naturalisante.

Marta Severo enfin a exploré avec Laurent Beaugitte un corpus d'aperçus de news internationales. Dans leur démarche il y a une composante d'analyse textuelle, qui vise à définir le contrat de lecture de ces aperçus, et une deuxième composante qui prend en compte la dimension géographique "externe" de leur production.

Vous voyez donc que si un seuil entre immanence et naturalisation peut sans doute se fonder sur un plan théorique, il est aussi vrai que les chercheurs sont désireux de le remettre en question dans leurs pratiques. Ces contributions témoignent toutes d'une recherche de médiation avec le réel. Ce dernier a été pendant trop longtemps un mot tabou pour la sémiotique européenne, comme si nos mots avaient d'autres buts que d'épaissir notre expérience de l'existant. Nous étudions le sens, mais comme l'écrivait Umberto Eco, cela "ne revient pas à affirmer, avec Gorgias, qu'on ne peut rien dire de l'être : on en parle énormément, et même trop". Eco, disparu quelques jours après la journée d'études à l'origine de ce livre, écrivait encore que "même si l'être apparaît comme un effet de langage, il n'en est pas un au sens où le langage le construirait librement". Cette vision donne à la sémiotique la conscience de ne pas tourner en rond, dans une dimension sui-référentielle sans limites et donc sans substance.

Nous espérons que ce livre puisse contribuer à montrer la richesse des questions sémiotiques que l'on peut aborder avec des outils quantitatifs et qu'il puisse aussi stimuler la discussion sur l'épistémologie de la discipline.

Je laisse maintenant la parole à Matteo Treleani, qui va animer la table ronde à laquelle j'invite tout le monde à participer, et je vous remercie de votre attention.