Sens [public]
Revue internationale
International Web Journal
www.sens-public.org

Hermeneutica, une expérience numérique de l'interprétation

Ariane Mayer

Résumé : Avec *Hermeneutica. Computer-assisted interpretation in the humanities* (MIT Press, 2016), Geoffrey Rockwell et Stéfan Sinclair s'interrogent sur les transformations de l'interprétation de textes dans le milieu numérique. En particulier, au travers d'une méthodologie hybride faisant dialoguer réflexions et exemples, théorie et pratique de l'interprétation, ils réfléchissent à ce que les outils d'analyse textuelle assistée par ordinateur révèlent et infléchissent dans l'activité herméneutique. Au cœur de l'essai se trouve l'outil Voyant, espace numérique de quantification et de visualisation textuelle développé par les coauteurs, qui leur sert de matériau pour aborder les mutations contemporaines de la lecture dans les sciences humaines.

Mots clés : herméneutique, interprétation, dialogue, texte, lecture, numérique

Hermeneutica, une expérience numérique de l'interprétation

Ariane Mayer

Hybridations

Alors qu'il entreprenait une thèse sur la terminologie de l'intériorité chez saint Thomas d'Aquin, le père Roberto Busa observa que la doctrine de la présence était liée à la préposition « dans », et que ce petit mot en apparence trivial pouvait manifester la structure philosophique profonde d'un discours. Il commença donc ses recherches en listant à la main l'ensemble des occurrences du mot « dans » au sein du corpus thomiste, pour rattacher chacune d'entre elles à la phrase qui la contient et élucider, ce faisant, le mystère de l'intériorité chez l'auteur de la *Somme théologique*.

Bien plus tard, à la fin des années 1970, Roberto Busa s'associa avec IBM pour développer le projet *Index Thomisticus.* Publié d'abord en cinquante-six volumes imprimés, transféré sur CD-ROM à partir de 1989 et dorénavant accessible sur le web, cet *Index* gigantesque propose une lemmatisation des millions de termes latins présents dans les œuvres de saint Thomas et d'auteurs connexes. Par cette immense concordance (c'est-à-dire cette recension des occurrences d'éléments lexicaux), Busa a été l'un des pionniers de l'utilisation des techniques informatiques pour l'analyse linguistique et littéraire.

Si le projet du père Busa peut sembler fou ou démesuré, s'il peut nous paraître étrange d'adosser la lecture du sens profond d'une œuvre à la table de ses mots les plus fréquents, il a aujourd'hui toute une lignée de descendants. Le milieu numérique, qu'il s'agisse des *big data* ou des dispositifs d'analyse textuelle par ordinateur, pose des défis particulièrement sensibles à la question de l'interprétation. Les outils et les méthodes de lecture inédits qu'il engendre sont au principe d'une nouvelle série d'interrogations, toute contemporaines : quel est le rôle de la technique de lecture sur le travail interprétatif qu'elle peut façonner ? Et du même coup, quelle est la nature véritable de l'herméneutique d'un texte, si elle peut être secondée par les outils aveugles d'une machine qui le lit sans comprendre ? Enfin, quel devenir pour le travail du chercheur en sciences humaines, s'il doit composer ses pratiques classiques avec celles que lui apporte une condition technologique inédite ?

C'est à ces questions que s'attelle un ouvrage paru en 2016 aux Presses du MIT, *Hermeneutica*, co-écrit par Geoffrey Rockwell (Université d'Alberta) et Stéfan Sinclair (Université McGill). Partant du constat d'une mutation profonde dans les outils d'analyse textuelle, et dans la matérialité même du texte sous

Publication de l'article en ligne : 2017/03
http://www.sens-public.org/spip.php?article1237

sa forme électronique, les deux chercheurs nous livrent une enquête protéiforme et vivante sur le devenir de l'interprétation à l'heure numérique.

Protéiforme d'abord, car le projet *Hermeneutica* se décline sur une pluralité d'interfaces reliées les unes aux autres. Ce n'est pas seulement un livre papier de facture traditionnelle, mais aussi une version électronique de ce dernier (http://hermeneuti.ca), qui offre des compléments, guides et ressources hypertextuelles annotées, et surtout un dispositif en libre accès sur le web, Voyant (http://voyant-tools.org), conçu par les coauteurs, qui propose un environnement de lecture et d'analyse de textes numériques. Traduit en neuf langues dont le français, Voyant comprend un ensemble d'outils de quantification et de visualisation de corpus linguistiques (table de concordances de termes, nuages de mots-clés) à la disposition de l'herméneute. Indissociable du livre, il est à la base de toutes les analyses d'exemples qui jalonnent *Hermeneutica*.

Hybride par ses interfaces de lecture, *Hermeneutica* l'est également par sa méthode, dynamique et collaborative. Œuvre à quatre mains de deux chercheurs et designers, le projet ne se contente pas de penser l'interprétation : il la donne à vivre, il la fait vivre. La structure de l'essai imprimé reflète un parti-pris méthodologique fort : celui de ne jamais dissocier la théorie herméneutique de sa mise en acte pratique. Ainsi la table des matières voit-elle s'alterner des chapitres réflexifs sur l'interprétation, ses procédés informatiques (chapitre 2), ses outils (chapitre 3), ses modèles (chapitre 9) et ses enjeux à l'heure numérique (chapitre 11), et des chapitres consacrés à l'étude d'exemples, nommés « Interludes » (chapitres 4, 6, 8 et 10). Ces derniers mobilisent des corpus de textes très variés, allant des discours d'Obama sur la notion de race à l'analyse de la liste de diffusion en humanités *Humanist*, dont chacun fait l'objet d'une lecture sur Voyant couplée aux interprétations des chercheurs.

Hybride, *Hermeneutica* l'est donc par sa structure autant que par sa question de recherche, l'une et l'autre en étroite symbiose. S'il s'agit de dessiner les contours d'une recherche composite en sciences humaines, qui fait dialoguer ses outils classiques avec les procédés algorithmiques les plus récents, Geoffrey Rockwell et Stéfan Sinclair conçoivent aussi leur méthode de recherche sous le signe du dialogue. Interfaces, disciplines et approches dialoguent pour faire jouer différentes vues sur l'interprétation, en théorie et en acte, en questions et en exemples. Cousin de la recherche-action et de la recherche-création, *Hermeneutica* paraît ainsi inaugurer un genre intellectuel nouveau, que l'on pourrait qualifier de recherche-design.

Publication de l'article en ligne : 2017/03
http://www.sens-public.org/spip.php?article1237

Lectures

L'ouvrage s'ouvre sur un moment de solitude : celui que décrit Descartes au début de la seconde partie du *Discours de la méthode* (1637).

> « J'étais alors en Allemagne, où l'occasion des guerres qui n'y sont pas encore finies m'avait appelé ; et comme je retournais du couronnement de l'empereur vers l'armée, le commencement de l'hiver m'arrêta en un quartier où, ne trouvant aucune conversation qui me divertît, et n'ayant d'ailleurs, par bonheur, aucuns soins ni passions qui me troublassent, je demeurais tout le jour enfermé seul dans un poêle, où j'avais tout loisir de m'entretenir avec mes pensées. »

Ces quelques mots résument le paradigme qui a longuement pu définir la recherche en sciences humaines. Celle-ci, d'après Descartes, exigerait calme et silence, un espace de réflexion solitaire à l'abri de toute compagnie qui risquerait de distraire l'attention concentrée du chercheur. S'il y a dialogue, il s'accomplit de manière toute platonicienne comme un silencieux entretien de la pensée avec elle-même. L'ambitieux objectif d'*Hermeneutica* peut être décrit, en quelque sorte, de la manière suivante : réécrire ce *Discours de la méthode,* en 2016, et à plusieurs. A plusieurs, c'est-à-dire non seulement en équipe, compte tenu du travail de pair à pair entre chercheurs que favorisent de plus en plus les plateformes en réseaux, mais aussi face à une autre forme d'altérité, méconnue jusqu'alors : celle des outils qui mêlent leurs protocoles à nos regards pour multiplier les voies d'approche d'un même texte, et qui glissent dans l'analyse, subrepticement, leur étrangeté et leur opacité.

Il existe en effet deux modes de lecture. Ou plus exactement, si l'on suppose que chaque lecture est singulière et qu'il y en a autant que de lecteurs, deux *sens* du mot « lecture ». D'une part, le nôtre, celui des lecteurs humains doués d'une conscience et pénétrés de la signification d'un texte, l'abordant avant tout pour le comprendre, et d'autre part celui des machines, dont on dit aussi qu'elles « lisent » un texte même si toute la notion de compréhension est évacuée. L'analyse informatique d'un texte, comme la décrit le chapitre 2, identifie celui-ci à un code pourvu d'un ensemble de chaînes de caractères, sur lesquelles mener certaines opératoires automatiques : le repérage de motifs (« *patterns* »), les concordances et KWICs (« *Key Words in Context* »), qui dénombrent les occurrences de termes et les visualisent sous forme de listes, la *tokenisation*, qui permet de résumer un texte à l'aide des mots qui le caractérisent le mieux, ou encore la comparaison, qui confronte la concordance d'un mot et celle d'un autre pour comparer les fréquences respectives.

Un exemple : les deux auteurs soumettent le texte de *Frankenstein*, de Mary Shelley, à l'analyse informatique de l'outil Voyant. Ce dispositif permet de visualiser les termes les plus fréquents sous forme de nuages de mots-clés, mais aussi sous forme de graphes qui traduisent l'évolution de leur fréquence à travers le temps (à travers les pages). Un des principaux résultats est que le mot le plus présent dans *Frankenstein* est « *man* » (« homme »), avec de 131 occurrences, donnée qui permet aux

Publication de l'article en ligne : 2017/03
http://www.sens-public.org/spip.php?article1237

chercheurs d'émettre une piste interprétative sur l'œuvre et sa signification : quel est le véritable thème, la grande question de *Frankenstein* ? C'est « la relation de l'homme avec ses créations technologiques », centrée essentiellement sur l'homme.

Comment interpréter ces deux modes de « lecture », et que dire de leur combinaison ? Geoffrey Rockwell et Stéfan Sinclair construisent une réponse nuancée, hybride encore une fois. « Nous développons des outils interactifs – des *hermeneutica* – qui peuvent augmenter et étendre notre lecture, et non nous remplacer », affirment-ils après l'analyse informatique de *Frankenstein*, en guise de conclusion. Dans cette vision toute proche des pionniers de la cybernétique – Vannevar Bush, Douglas Engelbart, J. C. R. Licklider –, où l'on peut reconnaître le thème cher à Engelbart d'une extension des capacités humaines au moyen de la machine (*Augmenting Human Intellect*, 1962), l'analyse algorithmique et l'interprétation humaine nouent donc une relation de complémentarité, non de substitution. La lecture machinique doit toujours, humblement, rester au service de la lecture herméneutique.

Nous retrouvons cette même conclusion, avec le rapport de subordination qu'elle dessine, bien plus loin dans l'ouvrage, au chapitre 9, qui est certainement l'un des plus riches philosophiquement. Dans ce chapitre, à partir d'une prémisse intéressante qui définit l'interprétation de façon polysémique, *à la fois* comme pratique et outils de cette pratique – les *hermeneutica* justement –, les auteurs se penchent sur la question : dans quelle mesure peut-on penser à travers ces outils analytiques ? Quel cadre heuristique offrent-ils à l'intelligence humaine ? C'est ici que les deux dimensions de notre objet semblent entrer en tension : l'interprétation comme œuvre de l'esprit, libre production d'une idéalité, et l'interprétation comme chose, adjuvant technique dans toute sa matérialité empirique. Comment peut-on penser par ces outils, avec ces outils, dans ces outils ? Tout au long du chapitre et en commençant avec Roberto Busa, les coauteurs interrogent le discours théorique des inventeurs de technologies interprétatives afin d'éclaircir cette tension inhérente à l'activité herméneutique.

Depuis l'*Index Thomisticus* du père Busa jusqu'au logiciel Voyant, en passant par des initiatives aussi passionnantes que méconnues comme le « *computer criticism* » de John B. Smith (1978), on retiendra surtout de ce parcours un apport-clé : les instruments, s'ils ont toujours vocation à amplifier l'intelligence humaine, sont aussi à leur tour des fragments cristallisés d'intelligence humaine. Dans une perspective très proche de celle de Bachelard, qui décrivait les instruments, dans Le Nouvel esprit scientifique, comme des « théories matérialisées », les auteurs cherchent à porter notre attention sur la manière dont « les outils portent des théories ». Ainsi mettent-ils en relation, par exemple, le dispositif informatique de John B. Smith – ARRAS – avec la pensée qui sous-tend son projet de critique littéraire assistée par ordinateur, dont le principe semble ici encore faire écho à Engelbart : « L'ordinateur ne fait qu'amplifier les pouvoirs de perception et de mémoire du critique ». En un mot, ce que Geoffrey

Rockwell et Stéfan Sinclair appellent les *hermeneutica* ressemblent beaucoup à ce que l'anthropologue Jack Goody, dans *La Raison graphique*, appelait les « technologies intellectuelles » : de simples ustensiles au service de l'esprit, certes, mais des ustensiles façonnés et informés par l'esprit.

Interludes

L'une des composantes les plus originales de l'expérience *Hermeneutica* est, justement, la place de choix qu'elle offre à l'expérience interprétative de textes. Comme s'ils conviaient les lecteurs à se joindre à leur travail, les auteurs soumettent à l'analyse des textes libres d'accès, examinés par le dispositif Voyant, qui est également en libre accès et en open-source. Les quatre corpus textuels analysés, qualifiés d'« interludes », sont autant d'occasions pour mettre à l'épreuve, et éventuellement en péril, les théories brièvement résumées ci-dessus.

Les interludes d'*Hermeneutica* peuvent nous étonner, comme nous étonnerait peut-être la mécanique du rire chez Bergson, consistant à plaquer du mécanique sur du vivant. Les analyses de textes proposées construisent presque toutes des interprétations qualitatives autour du thème ou du sens d'une œuvre à partir de données quantitatives, recensant les occurrences des termes les plus fréquents et donnant à voir leur répartition à travers le temps.

Au chapitre 4, l'analyse sous Voyant de la liste de diffusion *The Humanist* (vaste conversation en ligne et entre pairs sur les humanités) permet aux auteurs d'identifier les années, entre 1987 et 2008, où les expressions « *humanities computing* », « *computing in the humanities* », et « *digital humanities* », ont été respectivement les plus fréquentes, pour en inférer un certain nombre de conclusions sur la genèse et l'évolution de la notion d'humanités numériques (*digital humanities*) dans le monde académique. D'une façon fidèle à l'esprit de symbiose entre la structure et l'objet du questionnement qui anime *Hermeneutica*, il s'agit de relire *The Humanist* avec les outils des humanités numériques, et de le comprendre, également, par le prisme de ces outils.

Au chapitre 6, les auteurs procèdent à l'analyse de deux discours sous Voyant, prononcés l'un et l'autre à l'époque des élections présidentielles de 2008 aux États-Unis : le discours de Barack Obama intitulé « A More Perfect Union » (18 mars), et le discours de Jeremiah A. Wright, mentor spirituel d'Obama, donné quelques semaines après (27 avril). Les deux discours abordent la notion de race, et l'enjeu de l'analyse est d'observer les points de vue respectifs qu'ils apportent sur la question, et, plus particulièrement, la façon dont Obama cherche à se distinguer de son père spirituel. Les auteurs notent que dans chacun de ces discours, un terme ou une expression est récurrent et répété comme un mantra : « *time* » chez Obama (le temps, comme par exemple dans l'expression « *this is the time* ») et le

Publication de l'article en ligne : 2017/03
http://www.sens-public.org/spip.php?article1237

couple « *different* » et « *deficient* » chez Wright. Le premier traduirait l'insistance d'Obama sur le fait que nous sommes dans un moment historique exceptionnel et unique, un « temps » de basculement pour la notion de race, et le second reflèterait une autre façon de se positionner sur la question, Wright valorisant la différence que trop d'Américains encore assimilent à une déficience. De ces concordances de mots, Geoffrey Rockwell et Stéfan Sinclair tirent la conclusion suivante : « En tant que candidat à la présidentielle, Obama nous fait regarder les enjeux que nous avons en commun, les besoins unitaires du pays tout entier. En tant que ministre, Wright veut que l'on s'engage à regarder la différence, et non la déficience ».

Ces inférences interprétatives ont aussi leur lot de ratés, comme les auteurs eux-mêmes le relèvent lors du dernier interlude, au chapitre 10. Dans ce dernier, une œuvre de Hume est analysée sous Voyant : les *Dialogues sur la religion naturelle*. L'objectif des deux chercheurs, face à ce texte qui se présente comme une conversation entre trois interlocuteurs – le sceptique Philo, le théiste Cléanthe et le fondamentaliste Demea –, où aucune première per-sonne n'incarne la voix de l'auteur, est de mettre au jour le héros du dialogue, qui exprimerait les vues personnelles de Hume sur la question religieuse. Pour ce faire, ils comptabilisent le nombre de mots de chaque personnage et en déduisent, par la longueur de ses répliques, que le sceptique Philo est la voix principale de l'échange. Et, ayant identifié que les deux derniers chapitres de l'œuvre sont ceux où Philo domine le temps de parole, ils en infèrent que ces chapitres sont les mieux à même de refléter l'opinion de Hume. Mais il se trouve que dans ces chapitres 11 et 12, Philo retourne sa veste. Modèle de scepticisme durant tout le *Dialogue*, le voilà soudain reconnaissant l'évidence de l'existence de Dieu. Et les coauteurs de reconnaître, à leur tour, la distance critique qui reste de mise pour les chercheurs qui appuient leur travail interprétatif sur ces prothèses techniques :

> Le retournement de Philo illustre les dangers de l'analyse et de la visualisation textuelles. Les outils montrent *où* parlent les interlocuteurs et *combien* ils parlent, mais ils ne montrent pas ce qu'ils *disent*.

C'est que le fait littéraire, précisément, consiste moins dans le fait de tenir les propos de l'auteur pour paroles d'évangile que de repérer les espaces figurés où il joue avec son lecteur, les lieux d'ironie ou d'antiphrase où, tel Alphonse Allais avec *Un Drame bien parisien* ou Jonathan Swift avec *Une Modeste proposition*, leurs mots affirment, justement, l'inverse de ce qu'ils disent. Et ces espaces, tout ce que Gérard Genette appellerait des *figures,* la machine peut les comparer et les compter, mais aucunement les comprendre.

Publication de l'article en ligne : 2017/03
http://www.sens-public.org/spip.php?article1237

Discussion

Discussion, conversation, dialogue : tels seraient sans doute les principaux mots-clés d'*Hermeneutica*, les principes premiers à partir desquels refonder l'interprétation à l'heure électronique. C'est bien le dialogue, avec des pairs mais aussi avec des outils qui insèrent leur propre altérité entre le regard et le texte, qui forme la trame de cet essai hybride – dont le processus d'écriture est lui-même le résultat d'un échange collaboratif. Depuis l'ouverture sur une critique du solipsisme cartésien jusqu'à l'exemple de la liste de diffusion *Humanist,* paradigme d'une recherche en humanités pensée à l'image de la discussion ouverte ou du banquet, le dialogue est la pierre angulaire de l'herméneutique contemporaine qu'ils proposent. S'il s'agit de réécrire le *Discours de la méthode* en 2016, le dialogisme en sera donc le correctif.

Nous aurions très envie, à notre tour et sur leur invitation, de prolonger la discussion qu'amorcent les auteurs sur l'interprétation textuelle dans les humanités numériques. Notamment, malgré les précautions que prennent Geoffrey Rockwell et Stéfan Sinclair, et leur volonté de subordonner la lecture machinique à la lecture humaine, nous pourrions interroger leur présupposé fondamental : à savoir qu'il n'y a à *interpréter* que là où il y a quelque chose à *lire*. Mais est-il toujours vrai que l'on puisse déduire le thème d'une œuvre de ses termes les plus fréquents ? Que, retournant à l'exemple de *Frankenstein*, le nombre d'occurrences du mot « *man* » soit un indice, voire une preuve, du fait que la problématique du roman se concentre avant tout sur l'homme ? N'a-t-on pas affaire ici à un texte trop plein, qui dit tout ce qu'il y a à comprendre, et minimise ainsi le rôle des lecteurs ? Face à une conception si saturée de l'œuvre, où sont les blancs, les espaces vierges où le lecteur peut insérer son imaginaire ?

Pour continuer le dialogue, on gagnerait à faire converser *Hermeneutica* avec des théories de la lecture comme celle d'Umberto Eco ou avec l'esthétique de la réception, représentée par Hans Robert Jauss et Wolfgang Iser. Aux yeux d'Umberto Eco (*Lector in fabula*), il n'y a à interpréter que là où le texte se tait. Ce sont tous les lieux d'ambivalence, les propositions implicites et les vides de l'œuvre, suscitant la coopération d'un lecteur qui met du sien dans le texte pour combler les blancs, qui font le propre du fonctionnement littéraire. Wolfgang Iser (*L'Appel du texte*) affirme de son côté que, loin de déduire le sens d'une œuvre de ses mots les plus utilisés, « l'essentiel d'un texte est ce qu'il passe sous silence ».

Publication de l'article en ligne : 2017/03
http://www.sens-public.org/spip.php?article1237