

# Des textes au sens. Ce que les innovations technologiques ne prouvent pas.

*From text to sense. What technological innovations don't prove.*

< Sophie Anquetil <sup>1</sup> > < Nicole Pignier <sup>2</sup> >

1. Université de Limoges

2. Université de Limoges

Reproduisant les pratiques scientifiques des sciences dites exactes, les Sciences du Langage, de l'Information et de la Communication semblent avoir délégué, depuis une trentaine d'années, l'interprétation des faits du monde, de leur signification aux innovations technologiques et à leur mode de catégorisation des objets observés. Le numéro 3/2015 de la revue *Interfaces Numériques*, intitulé *Cultiver « le numérique » ?* précisait déjà la nécessité de mettre les technologies digitales en culture ; il interrogeait les capacités des Sciences Humaines et Sociales à appréhender « le numérique » à l'aide de leurs propres catégories (Pignier et Robert, 2015 : 339). Nous expliquons alors que l'épistémologie des données, devenue la *doxa*, tendait à aspirer les sciences, laissant croire que seule l'exploitation algorithmique des *big data* « permet l'avènement du sens » (*Ibid.* : 340).

Dans le même dossier, Bruno Bachimont pointait la perte d'intelligibilité sémiotique qui accompagne la formalisation des faits et des discours extraits de leur contexte pour se soumettre à la logique calculatoire :

*« Des séries statistiques à l'indice permettant de remonter une série de causes ou d'analyse, le scientifique de la culture pratique lui aussi le recours aux sciences formelles pour interroger la réalité qu'il*

*étudie. Mais toutes les statistiques sur le cours du pain à la veille de la révolution ne permettront en rien d'expliquer à quiconque n'a jamais eu faim ni ne sait ce que c'est que manger du pain frais ou cuit au four en quoi le cours du pain à la révolution peut avoir une quelconque relation avec la révolution, les tensions sociales, les émeutes populaires, bref la misère qui fait descendre dans la rue » (Bachimont, 2015 : 398).*

Parmi les “innovations” pointées dans le présent numéro, on recense la textométrie, les outils de fouille textuelle au sens large ainsi que le web sémantique. Le point de vue épistémologique qui sous-tend cette approche nie l'existence d'un processus de signification en deçà d'un matériau palpable, observable. Pourtant, le texte ne fournit que des « indications » et montre une « réticence » à tout dire selon l'expression d'Umberto Eco (1985) incitant ainsi le lecteur « à tirer du texte ce que le texte ne dit pas mais qu'il présuppose, promet, implique ou implicite, à remplir d'espaces vides, à relier ce qu'il y a dans ce texte au reste de l'intertextualité, d'où il naît et où il ira se fondre » (*Ibid.* : 7). Ainsi les analyses qui empruntent ces « innovations » technologiques, parce qu'elles prennent pour point de départ le matériau linguistique pour saisir le sens, occultent le fait que le discours est effet de sens entre locuteurs ou coénonciateurs. Suivant Michel Pécheux (1975), Eni Puccinelli Orlandi (1996) rappelle pourtant que « comprendre ce qu'est un effet de sens, c'est comprendre que le sens n'est (placé) nulle part, mais produit dans les rapports : des sujets, des sens » (*Ibid.*, p. 20). Ainsi pour faire émerger le sens d'un texte, toute analyse doit considérer son historicité, sa relation avec les textes qui l'ont précédé, son incomplétude, son extériorité, et la dimension plurielle des sujets qui l'instituent.

Qui plus est, le postulat de l'observabilité du sens suit une tradition positiviste qui appréhende les données comme les garants de l'objectivité, ignorant là la capacité de symbolisation des êtres humains. Pourtant, les données récoltées à partir de *cesdites* « innovations » ne constituent pas des *faits bruts indépendants des croyances des observateurs* (Searle 1995, 1998) : elles émanent au contraire de catégorisations ontologiques opérées par ceux qui les ont constituées. En cela, on ne peut affirmer leur objectivité épistémologique. Tout au plus, l'analyste tient compte de leur dimension idéologique.

En outre, a-t-on bien affaire à des innovations ? Le terme, dont la première occurrence date de 1297, est emprunté au bas latin *innovatio* qui signifie alors « changement, renouvellement ». Au XVIII<sup>e</sup> siècle, le sens s'affaiblit et se spécialise ; le terme « innovation » s'applique aux domaines de l'industrie et des affaires pour nommer n'importe quel produit nouveau (Rey, 2016 : 1156), qu'il marque un changement profond de manière de penser et de vivre ou qu'il vienne s'inscrire dans une continuité. C'est ce sens qui prévaut encore aujourd'hui dans les discours d'escorte auxquels se plie la recherche scientifique, y compris en Sciences Humaines et Sociales. Les logiciels de textométrie, de fouille ainsi que le web sémantique ne s'inscrivent en rien dans une mise à l'épreuve de la *doxa* que constitue « la silicolonisation du monde » selon l'expression d'Éric Sadin (Sadin, 2016). Leur design ne fait qu'exploiter le « cadre technologique » présent, laissant les acteurs des infrastructures numériques imposer leur « loi aux prétendus innovateurs » (Sadin, 2016 : 144). Il ne s'agit donc pas tant d'innovations que de disruptions. Le terme de « disruption », provenant des Sciences Physiques, désigne l'ouverture soudaine d'un circuit électrique avec pour effet la production d'étincelles. Éric Sadin précise que, dans les années 2000, il nomme l'accélération des développements techniques majoritairement dus aux sciences informatiques (*Ibid.* : 140). Ainsi, les objets logiciels auxquels recourent entre autres les recherches en Sciences du Langage, intégrant les spécificités des technologies numériques dont le traçage, le comptage, la réticularité (Pignier, 2020a : 166-167), témoignent « d'une forme passive de l'innovation, se contentant de profiter de ce qui est à portée de main et de l'appliquer en théorie à n'importe quel domaine » (Sadin, 2016 : 142).

L'objectif du présent numéro est précisément d'interroger l'(in)capacité des « innovations » technologiques à appréhender l'en deçà de la matérialité discursive – interdiscours, intertexte, silence fondateur – dans l'analyse des processus signifiants et ainsi à participer d'une science du sens. Ainsi les « innovations » technologiques sont ici sollicitées, non pour prouver qu'elles prouvent, mais pour exhiber ce qu'elles ne prouvent pas. En prenant appui sur des faits éprouvés, il s'agira de montrer que ces outils contraignent l'analyste à « travailler à partir d'indices lacunaires » « peu déterminés », « de données filtrées par les catégorisations du logiciel », « avec un cotexte peu élargi » (Anquetil, Duteil & Lloveria, dir., 2019 : 15). Dans le cadre de ce dossier, des

contributeurs de Sciences de l'Information et de la Communication, de Sciences du Langage, de Philosophie de la technique et des épistémologies des sciences, caractérisent la manière dont l'énonciation comme expérience concrète, située, sociale et existentielle est évacuée dans l'analyse lexicale des données, les conséquences sémiotiques de ces Faïres scientifiques ?

### **Entretien avec Sylviane Sambor : Pratiques éditoriales et usages du numérique**

L'entretien d'une éditrice, Sylviane Sambor (L'Escampette), ouvre ce numéro en déterminant les manières dont les mutations du numérique ont reconfiguré, non seulement l'objet « livre », mais aussi et surtout l'expérience sensorielle du lecteur et son rapport à la lecture. L'échange questionne également l'influence du numérique sur la « bibliodiversité », la médiation culturelle et l'émergence du sens au sein de dispositifs numériques. L'éditrice y propose enfin de redéfinir le concept même d'innovation en y associant les concepts de *lecture*, d'*édition*, de *saveurs*, de *solidarité*, d'*alimentation*.

### **Emmanuel Souchier et Joanna Pomian : Qui cause ? Qui dose ? Qui ose ? À l'infini poème apporter une fin.**

L'analyse lexicale instrumentée serait un gage d'objectivité. Mais une telle croyance ignore que les data relèvent toujours d'un geste d'interprétation rappellent Bruno Bachimont Emmanuel Souchier et Joanna Pomian. Ils invitent à questionner la multiplicité des acteurs confortant les dispositifs qui rendent les données accessibles à l'utilisateur, par exemple les résultats d'une requête sur le moteur de recherche de Google. Le pouvoir donné à l'informaticien, au statisticien configure une relation statistique au monde influant sur la société et ses pratiques intellectuelles. Mais l'analyse lexicale instrumentée coupe-t-elle avec la pensée linguistique ? Assurément non ; celle-ci fait abstraction de la corporéité, de la matérialité des textes. Elle reste indifférente à tout ce qui – encre, papier, format, typographie, rythmes des lignes et des blancs, couverture, volume, expérience tactilo-kinesthésique – en relation avec

le corps et la perception, entre en compte dans le processus d'énonciation ou d'écriture, de co-énonciation ou de lecture, de mémorisation. Or, c'est précisément de cet évacué que « l'énonciation éditoriale » prend soin ; en quoi et comment le travail éditorial, en coénonciation avec l'énonciation auctoriale, contribue-t-il à offrir des expériences de lecture plus ou moins spécifiques, en tous les cas toujours situées ?

Les auteurs de « Qui cause ? Qui dose ? Qui ose ? À l'infini poème apporter une fin » donnent en partage, à partir d'une expérience scientifique située, ce qui contre leur attente initiale, fait non-sens ou tout au moins contre-sens dans l'automatisation des établissements des variantes, repentirs (signes, mots, phrases, paragraphes) existants entre les états multiples d'un même texte. En effet, c'est en mettant en place un logiciel d'aide à l'établissement des variantes que Joanna Pomian et Emmanuel Souchier, ont pris conscience de ce qui se joue dans le geste artisanal lent et complexe de celle ou celui qui « saisit » les manuscrits pour les rendre décodables, calculables par la machine. Rien moins que la rencontre de l'autre, l'expérience de sa présence, le fait de « rejouer » sa situation énonciative, créative, « d'entrer en empathie » avec l'écrivain.e. Le texte « factuel », fait de données, rendu possible par un logiciel d'I.A. peut avoir son intérêt mais en aucun cas ne peut remplacer le travail anthropologiquement fondateur consistant à se couler dans la peau de l'écrivain.e. pour une « compréhension située des processus de rédaction ayant entraîné les multiples repentirs de l'écriture », concluent les auteur.e.s. Ainsi, le traitement algorithmique des données ne peut remplacer les gestes manuel, perceptif, sensible, intellectuel par lesquels les chercheur.se.s comprennent les textes, les savourent, les analysent.

### **Bruno Bachimont : La complexité herméneutique à l'épreuve du calcul**

En pleine résonance avec la contribution précédente, Bruno Bachimont précise en quoi et pourquoi le traitement algorithmique (n') est (qu') un intermédiaire, une étape qui ne doit jamais se suffire à elle-même ni se solder par une interprétation superficielle et rapide des résultats. Les textes renvoient à des situations d'énonciation, d'adresse aux autres qui y répondent dans une co-énonciation située, symbolique

et concrète. Cependant, pour la machine, il n'y a pas de texte, il n'y a pas de sens. La matérialité, le texte en situation sont hors-jeu. Le traitement numérique des textes nécessite qu'ils soient transformés en « données », c'est-à-dire en « calculi » qui permettent de réaliser des manipulations opératoires. Les calculi, explique Bruno Bachimont, sont donc a-sémiotiques, ils ne font signe à personne, à rien, ils se caractérisent par une neutralité fondamentale. Ce sont des « unités techniques de manipulation » (UTM) ou « alphabets numériques » ; ASCII ou UNiCode pour un texte linguistique, pixel pour une image ; point pour un graphique. Il est alors illusoire de croire que les calculi manipulés sont remis en situation si on prend des données plus massives ; le big data. Ils relèvent en effet d'une complexité technique (in)différente à la complexité sémiotique. Celle-là se veut « analytique » - elle consiste à « couper le réel au moyen de lois qui vont rendre compte de la diversité et de la variabilité des phénomènes » - et « synthétique » - elle consiste à « convoquer dans un même système des éléments hétérogènes pour lesquels on ne dispose pas de théorie globale, mais seulement locale ou propre à quelques-unes de ses parties fonctionnelles ou matérielles » précise Bruno Bachimont. Mais si des assemblages formels inédits émergent grâce au traitement algorithmique a-sémiotique des données, quel sens leur donner ? Comment les appréhender ?

### **Rossana De Angelis : Qu'est-ce que veut dire interpréter dans le cadre d'une herméneutique (du numérique) ?**

Plaidant en faveur d'une « herméneutique du numérique », Rossana De Angelis interroge le sens même d'« interpréter » afin de mieux en comprendre les processus d'*interprétation* et d'*objectivation* des données textuelles. Elle montre que deux médiations sont nécessaires pour passer des textes au sens : 1-une première médiation concernant « l'inscription » des traces numériques calculables, ce qui suppose une manipulation asémantique et une rupture de la relation référentielle avec le monde ; 2-une deuxième médiation concernant la « transcription » des traces numériques interprétables, ce qui suppose une manipulation sémantique. De l'analyse de ces deux processus de médiation découle la conviction, pour la chercheuse, d'ajouter une attitude synthétique afin de rétablir la relation entre le langage et le monde.

### **Olga Galatanu : Le traitement du potentiel argumentatif des mots dans l'analyse de discours**

Dans la perspective de la Sémantique des possibles argumentatifs, Olga Galatanu propose d'interroger la nature sémantique et le statut épistémologique des observables dans la matérialité du texte, tels qu'ils sont proposés par son traitement informatique : mots, associations de mots, environnements sémantique et syntaxique. La thèse défendue est que l'analyse de ces observables n'épuise pas le sens discursif et qu'elle doit s'appuyer sur une démarche de prise en compte dans l'interprétation du sens non seulement du contexte, mais aussi et surtout de la complexité du potentiel argumentatif des mots. Le corollaire de ce point de vue est une triangulation des méthodologies dans laquelle l'analyse qualitative sémantico-discursive des conceptualisations sémantiques offre un modèle conceptuel au traitement textométrique qu'elle précède.

### **Sophie Anquetil : Pourquoi les *actes de langage politiques* échappent-ils au traitement linguistique des données ?**

S'inscrivant dans une perspective pragmatique, l'article de Sophie Anquetil se propose de mettre les outils de fouille textuelle à l'épreuve du traitement des actes de langage *politiques*. L'autrice y défend l'idée selon laquelle un macro-acte de langage *politique*, parce qu'il s'inscrit dans un processus d'institutionnalisation, impose un mode de réalisation « complexe » qui construit un implicite institutionnel et dilue sa valeur illocutoire de base. En prenant appui sur l'exemple d'ALERTER, elle détermine comment le processus d'institutionnalisation en trois phases – 1. l'assignation d'une fonction agentive aux faits institutionnels ; 2 la construction d'une croyance et intentionnalité collectives ; 3. la construction d'une éthique politique – influe sur le marquage de l'acte de langage et fragilise la possibilité d'en saisir les données à l'aide d'outils de fouille textuelle.

## Bibliographie

- ANQUETIL Sophie, DUTEIL Carine et LLOVERIA Vivien (dir.), 2019, *Le sens des données. Statut du corpus et herméneutique à l'aune des Humanités numériques*. Collection Humanités numériques, Paris, L'Harmattan.
- BACHIMONT Bruno, 2015, « Le numérique comme milieu : enjeux épistémologiques et phénoménologiques », in Nicole PIGNIER et Pascal ROBERT (dir.), *Cultiver le « numérique » ? Interfaces Numériques*, vol. 4, 3, p. 385-402 : <https://www.unilim.fr/interfaces-numeriques/383>
- ÉCO Umberto, 1985, *Lector in fabula*, Paris, Grasset.
- PÉCHEUX Michel, 1975, *Les vérités de la Palice*, Paris, Maspero.
- PIGNIER Nicole, 2020a « L'éducation critique aux médiations informationnelles et communicationnelles en milieu numérique », in Sophie JEHEL et Alexandra SAEMMER (dir.), *Education critique aux médias et à l'information en contexte numérique*, Presses de l'ENSSIB, Lyon, p. 163-173.
- PIGNIER Nicole, 2020b, « Le sens, le vivant ou ce qui nous relie à la Terre », in Nicole PIGNIER et Lina Marcela LIÑÁN DURÁN (dir.), *Le design de « l'Intelligence artificielle » à l'épreuve du vivant*, *Interfaces Numériques*, vol. 9, n° 1 : <https://www.unilim.fr/interfaces-numeriques/4144>
- PIGNIER Nicole et ROBERT Pascal, 2015, Introduction au dossier *Cultiver le « numérique » ?*, *Interfaces Numériques*, vol. 4, n° 3, p. 339-349 : <https://www.unilim.fr/interfaces-numeriques/383>
- PUCCINELLI ORLANDI Éni, 1996, *Les formes du silence. Dans le mouvement du sens*. Éditions des Cendres.
- REY Alain (Dir.), 2016, *Dictionnaire historique de la langue française*, Éditions Le Robert, Paris.
- SADIN Éric, 2016, *La silicolonisation du monde*, Paris, Éditions L'Echappée.
- SEARLE John, 1995, *The Construction of Social Reality*, New York, Free Press (*La construction de la réalité sociale*, trad. C. Tiercelin, Paris, Gallimard, 1998).