Émilie NÉE, dir., *Méthodes et outils informatiques pour l'analyse des discours*

Rennes, Presses universitaires de Rennes, coll. Didact méthodes, 2017,
250 pages

Abdelkader Sayad

Édition électronique

URL : <http://journals.openedition.org/questionsdecommunication/13173>

DOI : 10.4000/questionsdecommunication.13173

ISSN : 2259-8901

Éditeur

Presses universitaires de Lorraine

Édition imprimée

Date de publication : 1 septembre 2018

Pagination : 410-411

ISBN : 978-2-8143-0519-9

ISSN : 1633-5961

Référence électronique

Abdelkader Sayad, « Émilie NÉE, dir., *Méthodes et outils informatiques pour l'analyse des discours* », *Questions de communication* [En ligne], 33 | 2018, mis en ligne le 01 septembre 2018, consulté le 05 janvier 2021. URL : <http://journals.openedition.org/questionsdecommunication/13173> ; DOI : <https://doi.org/10.4000/questionsdecommunication.13173>

emblématiques de cette tradition sont reproduits : « Le Voleur de prunes », « La Feuille à l'envers » et « Le Nez de Martin ». Bertrand Masquelier remplace la Vendée par la Caraïbe. Dans son article intitulé « Acte locutoire et *double-entendre*. Le *calypso* grivois à Trinidad (Caraïbes) » (pp. 97-138), il analyse une forme liée au carnaval : le *calypso*. Il s'intéresse plus précisément à sa variante érotique, appelée *saucy calypso*. Cette forme, liée à un instrument de percussion appelé *steel drum*, fonctionne sur le double-entendre, à la manière du texte de Black Stalin « *Wait Dorothy wait* ». Ce morceau de *calypso* superpose la musique et le sexe : « *I jam she/She jam me* ». Bertrand Masquelier indique que le but de cette forme est de « mettre en fiction la rencontre des corps, ce qui nous en est familier, et éviter le réalisme de la description pornographique » (p. 108).

Une troisième série de contributions met en perspective le non-dit, entre ordre et désordre. Dans « "Il m'a traité !" "Ils me traitent !" : de quoi ? L'injure ou le leurre des mots » (pp. 221-248), Éveline Larguèche théorise une situation de communication propre à l'injure dans laquelle elle identifie l'« injurier » d'une part, c'est-à-dire celui qui insulte, et, de l'autre, l'« injurié », la cible de l'insulte, ou l'« injuriaire », celui à qui on dit une insulte qui concerne un tiers. S'intéressant également à l'Afrique et en particulier à la culture arabe, Éveline Larguèche indique la manière dont l'insulte est liée, dans ce contexte culturel, à l'accusation de fornication. Ainsi l'insulte possède-t-elle une dimension à la fois morale et sociale qui permet également de comprendre la réception de certaines formes artistiques comme le nu pictural. Élisabeth Motte-Florac, dans « L'Implicite dans une histoire *purhépecha* de pacte avec le diable (Mexique) » (pp. 249-286), interprète le non-dit dans le cas du pacte avec le diable. Dans le contexte du Mexique, les difficultés liées à l'adaptation aux changements sont envisagées selon un prisme chrétien. Le diable, représenté comme un étranger élégant, invite à gifler le Christ pour devenir riche. Le non-dit de ce canevas est celui de la désorientation, traduite par la perte de l'âme. Un canevas sacré sert à traduire une situation profane. L'Amazonie colombienne succède au Mexique dans l'article de Laurent Fontaine intitulé « Ce qu'on ne dit pas chez les *Yucuna* (Amazonie colombienne) » (pp. 139-173). L'ethnologue s'intéresse à ce qu'on ne dit pas parce que cela va de soi dans une société donnée pour les membres qui en font partie. Chez les *Yucuna*, qui font partie de la famille des Indiens Arawak, de filiation patrilinéaire exogamique, vivant de la chasse, de la pêche, de la cueillette et de l'horticulture, dans la moitié basse du fleuve Miriti-Parana et le Bas Caqueta, en particulier le village

La Pedrera, les groupes se distinguent selon leur rapport à la sexualité et à la religion. Il y a d'un côté ceux qui sont vierges et, de l'autre, ceux qui ne le sont plus ; d'un côté, ceux qui sont initiés, et de l'autre, ceux qui ne le sont pas. En contexte domestique, le non-dit concerne la répartition des tâches domestiques entre homme et femme au sein de la famille. Dans le domaine public, le non-dit relève du bon déroulement de la vie sociale et le dit permet d'accuser ou d'excuser. Laurent Fontaine indique aussi la manière dont les *Yucuna*, au lieu de ne pas dire les mots du sexe, les étendent métaphoriquement aux domaines qui opposent les hommes et les femmes, ceux qui ont l'expérience de la sexualité et ceux qui ne l'ont pas. Dans « Entre oral, anal et carnaval. Maupassant ou les ruses du dire » (pp. 287-312), Jean-Marie Privat analyse le non-dit dans le conte normand intitulé *Toine* (1885). Cette lecture se focalise davantage sur le malentendu et les possibilités de l'interprétation que sur le non-dit et se situe entre la psychanalyse freudienne et la culture populaire de François Rabelais telle que l'envisage Mikhaïl Bakhtine.

En conclusion, ce livre propose une analyse linguistique des formes culturelles et artistiques du non-dit dans diverses aires culturelles. Ces contextes culturels et ces formes artistiques déterminent la fonction et le sens du non-dit. Ainsi le non-dit du proverbe africain est-il un dire africain qui diffère du dire occidental. Le non-dit dépend également des situations de communication, du discours généalogique du *jasare* à celle de l'« injurier » face à un « injurié » ou à un « injuriaire ». D'un point de vue axiologique, le non-dit apparaît normatif ou transgressif. Autrement dit, il rappelle à l'ordre, ou transgresse ce dernier. Dans l'approche empirique de cet ouvrage collectif, la méthodologie linguistique indique le non-dit comme un phénomène transculturel lié à la responsabilité. Le non-dit, conformément à la figure de style qu'est la prétérition, dit paradoxalement ce qu'il ne dit pas et réciproquement.

Christophe Cosker

HCTI, université de Bretagne occidentale, F-29000

Christophe.cosker@free.fr

Émilie Née, dir., *Méthodes et outils informatiques pour l'analyse des discours*

Rennes, Presses universitaires de Rennes, coll. Didact méthodes, 2017, 250 pages

Les chercheurs en sciences humaines et sociales sont confrontés aujourd'hui à une masse importante de données discursives et textuelles. Ces données, qui sont en continue expansion grâce la généralisation

de l'usage de l'internet, et même si elles constituent de prime abord un champ des plus intéressants pour tout chercheur, posent un certain nombre de problèmes : quels outils (statistiques et informatiques) utiliser pour rendre compte de ces données ? Quelles méthodes adopter pour les organiser et en constituer des corpus ? Et dans quels cadres théoriques ?

Cet ouvrage collectif, dirigé par Émilie Née se propose de répondre à ces questions récurrentes dans de nombreuses disciplines comme les sciences de l'information et de la communication, la sociologie, l'histoire, ou encore l'analyse du discours (AD). Cette dernière constituera d'ailleurs un point d'appui pour cet ouvrage d'initiation à l'analyse textuelle assistée par l'outil informatique, où l'on montre « l'intérêt des outils de traitement automatique pour analyser des discours » (p. 11).

L'intérêt de ce manuel d'initiation réside dans le fait qu'il se propose d'articuler les outils et méthodes fondés sur les statistiques lexicales « à des problématiques d'analyse du discours » (p. 15). En effet, comme l'expliquent ses auteurs, l'outillage informatique ne dispense pas le chercheur de s'interroger sur la validité de sa démarche en la confrontant à son objet d'étude. Cette contribution entend donc répondre à des besoins concrets chez les étudiants et chercheurs qui s'intéressent à ce domaine, en abordant des problématiques et des questions de recherches concrètes qui relèvent de l'analyse du discours, dans un contexte où « mesurer les discours semble aujourd'hui devenu une étape incontournable pour de nombreuses recherches en sciences humaines et sociales » (p. 16).

Dans l'introduction (pp. 9-16), Frédérique Sitri et Christine Barats passent en revue quelques postulats relatifs à l'analyse du discours. En effet, après un bref panorama des principales disciplines qui analysent les textes avec des outils informatiques, comme la lexicométrie ou la textométrie, les deux chercheuses présentent efficacement les principes de base de l'analyse du discours, en insistant surtout sur l'importance du contexte et de la situation. De fait, « Pour l'AD, la situation, le contexte, l'environnement du discours ne sont pas ou pas seulement conçus comme matériels : ces extérieurs au discours, ce sont aussi des discours, qui le conditionnent en partie » (p. 12). Face à ces principes, notamment ceux en rapport avec les procédures de constitution et d'analyse des corpus, l'outillage informatique s'avère être « une sorte de garant méthodologique » (p. 14) qui permet une « lecture objective du texte », donnant une valeur de science exacte à cette discipline qui relève des sciences humaines et sociales (SHS).

L'ouvrage se compose de 6 chapitres complémentaires, allant de la présentation de quelques notions, à l'application de ces outils sur des corpus concrets. Le premier, intitulé « Compter les mots ? Pas si simple » de Serge Fleury, Émilie Née et Christine Barats (pp. 17- 39), revient à juste titre sur les difficultés inhérentes aux procédures de comptage des mots en fonction des outils informatiques utilisés, ces derniers ne fournissant pas forcément les mêmes résultats. Ce chapitre se veut aussi une sorte de présentation des prérequis indispensables pour une analyse informatisée de données textuelles. Les auteurs passent en revue plusieurs idées reçues sur l'outil informatique, en menant « une expérience ludique » sur le poème de Charles Baudelaire « La mort des amants ». Des présentations très efficaces de quelques outils d'analyse grands publics, comme « les nuages de mots » ou « Google NGram viewer » sont également proposées dans ce chapitre. Dans « Constituer un corpus en analyse du discours, un moment crucial » (pp. 41-62), Frédérique Sitri et Christine Barats s'intéressent à la notion de corpus en AD, tout en reprenant quelques principes théoriques incontournables qui sous-tendent sa constitution et son organisation. Cette étape, incontournable pour tout chercheur dans ce domaine, nécessite des précautions particulières. La délimitation d'un corpus en AD est le résultat de décisions et de choix, qui doivent être liés aux questions et objectifs de la recherche. Aussi les auteurs proposent-ils « des exemples concrets de structuration de corpus avant d'aborder le cas particulier des corpus recueillis sur le web » (p. 41). Le troisième chapitre, « Constituer un corpus en trois scénarios » (pp. 63-101), d'Émilie Née et Serge Fleury, se situe dans la continuité du deuxième, en traitant des considérations plus concrètes, liées à la mise en œuvre pratique du corpus. L'objectif est de « traiter des implications pratiques liées à l'informatisation des données et à l'automatisation des analyses » (p. 63) en présentant trois scénarios de constitution de corpus : le premier relevant d'un corpus médiatique, le deuxième sociopolitique et le dernier concernant un discours de campagne électorale (un genre particulier de discours politique). Il est à noter que ces scénarios, qui sont « inspirés de travaux réalisés avec des étudiants de master (scénario 1), de travaux existants ou de recherches en cours (scénarios 2 et 3) » (p. 64), illustrent très bien les protocoles présentés dans les deux premiers chapitres de l'ouvrage. « Compter dans les textes, quelles unités ? » d'Émilie Née, Jean-Marc Leblanc et Serge Fleury (pp. 103-121) revient sur l'épineuse question des unités de décompte en analyse de données textuelles. Savoir ce que l'on compte est une étape importante dans tout travail

sur des données textuelles : « Pour permettre des mesures sur les textes, il faut définir les unités qui composent ces textes et dont on va tenir compte pour les mesures » (p. 103). Ces mesures varient le plus souvent et dépendent de plusieurs paramètres comme le « point de vue adopté » ou « les postulats théoriques » de chaque chercheur. En tout état de cause, ces « unités de décompte » poussent les chercheurs « à dépasser les catégories classiques de description des langues » (p. 103). « Quels outils logiciels et pour quoi faire ? » de Jean-Marc Leblanc, Serge Fleury et Émilie Née (pp. 123-161) présente, comme son titre l'indique, « les principaux outils utilisés en ADT » (p. 123). Il se propose également de montrer comment utiliser ces logiciels pour manipuler les types d'unités préalablement définies ainsi que les procédures d'interprétation des résultats, et ce, en centrant la présentation sur les principes de l'analyse factorielle des correspondances (AFC).

« Problématiques d'analyse du discours et méthodes » de Marie Veniard et Frédérique Sitri (pp. 163-202) se situe dans la continuité du chapitre précédent, en proposant une illustration pratique fondée sur plusieurs problématiques abordées en analyse du discours. Il s'agit de montrer comment ces questions peuvent être analysées par des logiciels de traitement automatique des données textuelles. Les outils informatiques ne donnent pas un accès direct à l'interprétation discursive et nécessitent en plus une réflexion méthodologique sur les catégories d'analyse à privilégier. Accéder au sens n'est pas une entreprise aisée. Cela nécessite la prise de nombreuses précautions afin de ne pas tomber dans les pièges de la « transparence des discours ». Ce chapitre central, qui propose quelques clés de l'interprétation discursive, revient sur plusieurs exemples de problématiques relatives à l'identification des thèmes dans un corpus, à l'évolution du sens des vocables résultats de changements socio-historiques, aux genres discursifs, etc.

D'un point de vue pédagogique, des « fiches pratiques et approfondissements » (pp. 203-228) sont proposées à la fin de cet ouvrage d'initiation afin de compléter les orientations proposées dans les différents chapitres. Ces fiches reviennent à titre d'exemple sur les procédures de constitution des corpus et abordent « le langage des expressions régulières » (p. 203) ou encore les « commandes Unix pour manipuler des données » (p. 206). De la même manière, il convient de noter que chaque chapitre est savamment ponctué par différents types d'encadrés, proposant des éclairages bibliographiques, documentaires ou notionnels, ou invitant à des recherches plus développées. Par ailleurs,

il est proposé une synthèse de « ce qu'il faut retenir au terme de ce chapitre... », ce qui facilite grandement la lecture. Même si les références en la matière ne manquent pas, cet ouvrage pourra constituer une ressource pédagogique très utile pour tout étudiant souhaitant s'initier au traitement informatique des données discursives et textuelles ou approfondir ses connaissances dans le domaine.

Abdelkader Sayad

Université de Mostaganem, Algérie, DZ-27000/
MoDyCo, université Paris Ouest Nanterre La Défense,
F-92000
sayadaek@yahoo.fr

Médias, technologies, information

Camille ALLOING, Julien PIERRE, *Le Web affectif. Une économie numérique des émotions*

Bry-sur-Marne, Ina Éd., coll. Études et controverses, 2017, 124 pages

L'ouvrage de Camille Alloing et Julien Pierre présente les premiers résultats d'un programme de recherche qui questionne l'émergence d'une économie numérique des émotions. Cependant, dès l'introduction (pp. 9-16), les auteurs mobilisent un concept plus large, l'affect, qui est selon leur propos « un élément qui circule entre les corps, et certains dispositifs numériques tendent à faciliter cette circulation » (p. 11). À l'aube d'une période où « un mélange s'opère entre le vivant et l'artificiel » (Franck Renucci, Benoît Le Blanc, Samuel Lepastier, « Introduction générale », *Hermès, La Revue*, 68, 2014, pp. 11-14, ici p. 11), imaginer un espace de médiation où s'entremêleraient affects, corps, messages et technique semble s'inscrire dans un débat pertinent en sciences de l'information et de la communication (SIC). De plus, les auteurs – étant eux-mêmes chercheurs en SIC – proposent une approche identifiée de la discipline en s'inscrivant dans la « circulation (des signes, des informations, des idées, etc.) » (p. 102) d'une part, et dans les médiations humaines et techniques, d'autre part.

Les affects numériques « supposent des médiations qui ne sont pas uniquement humaines » (*ibid.*). Ils circuleraient en surface, entre les interfaces, mais aussi en profondeur, dans les bases de données. L'introduction se charge ainsi de nous plonger dans un environnement théorique marqué par les dispositifs et la circulation d'éléments économiques autour de références comme Michel Foucault et Adam Arvidsson. La valeur affective des données personnelles des utilisateurs tendrait de la sorte à faire émerger une « forme de capital » (p. 31).