

Émilie NÉE (dir.) (2017), *Méthodes et outils informatiques pour l'analyse des discours*

Rennes, Presses universitaires de Rennes

Maude Vadot



Édition électronique

URL : <http://journals.openedition.org/communication/10428>

ISBN : 978-2-921383-90-5

ISSN : 1920-7344

Éditeur

Université Laval

Référence électronique

Maude Vadot, « Émilie NÉE (dir.) (2017), *Méthodes et outils informatiques pour l'analyse des discours* », *Communication* [En ligne], Vol. 36/2 | 2019, mis en ligne le 15 juillet 2019, consulté le 22 juillet 2019.

URL : <http://journals.openedition.org/communication/10428>

Ce document a été généré automatiquement le 22 juillet 2019.



Les contenus de la revue *Communication* sont mis à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Émilie NÉE (dir.) (2017), *Méthodes et outils informatiques pour l'analyse des discours*

Rennes, Presses universitaires de Rennes

Maude Vadot

RÉFÉRENCE

Émilie NÉE (dir.) (2017), *Méthodes et outils informatiques pour l'analyse des discours*, Rennes, Presses universitaires de Rennes

- 1 Depuis les premiers essais et propositions, les outils informatiques pour l'analyse de données textuelles se sont diversifiés. L'ouvrage dirigé par Émilie Née, fruit de deux années de recherche et de rédaction collectives d'une équipe constituée de Christine Barats, Jean-Marc Leblanc, Serge Fleury, Émilie Née, Frédérique Sitri et Marie Veniard, vient utilement faire le point sur la gamme de possibilités ouvertes par ces logiciels, dans une démarche ancrée en analyse du discours (désormais AD). Si l'ouvrage n'hésite pas à aborder les aspects techniques de la recherche pour permettre à l'utilisateur débutant de déjouer les obstacles les plus courants, les auteurs insistent sur la nécessité d'un questionnement épistémologique et méthodologique préalable, étape incontournable pour le choix de l'outil d'analyse. Conçu comme un manuel à l'usage des étudiants et des chercheurs en sciences humaines et sociales qui s'intéressent au discours, l'ouvrage s'enracine dans l'expérience de recherche de ses auteurs, issus de différentes disciplines : de nombreuses études de cas sont exploitées afin de montrer l'articulation entre questions de recherche, données et outils. Le format et la structuration de l'ouvrage correspondent bien au genre « manuel » : des figures, notamment des captures d'écran, illustrent pas à pas le propos — on peut cependant regretter le manque de lisibilité de certaines de ces illustrations, dont le format est un peu petit. Des zooms et des encadrés permettent d'approfondir certains points spécifiques ; chaque chapitre se clôt sur une

conclusion synthétique et pratique. Enfin, cinq fiches pratiques sont proposées en fin d'ouvrage.

- 2 L'introduction rappelle les postulats théoriques de l'AD pratiquée par ses auteurs, qui prolonge les travaux de l'AD dite « à la française », prenant appui sur les travaux de Jean Dubois, Michel Pêcheux et, dans une certaine mesure, Michel Foucault. L'explicitation de ces positionnements constitue d'ailleurs une très bonne introduction succincte à la circonscription de l'AD pour les lecteurs qui n'en seraient pas familiers. Les auteurs soulignent les apports, dès l'origine de l'AD, des méthodes informatiques à cette discipline interprétative : elles obligent le chercheur à expliciter ses procédures, fournissent une aide à la délinéarisation du texte et rendent possible le traitement de très grands corpus. Les méthodes et outils informatiques contribuent ainsi à un enrichissement des questions de recherche. Des précautions sont cependant à prendre avant, pendant et après le traitement, car ces méthodes et outils ne libèrent pas le chercheur de la nécessité de comprendre comment les résultats sont obtenus et de savoir les interpréter.
- 3 Le chapitre 1 développe ce propos en repérant les étapes préalables au choix de l'outil informatique. À partir d'un exemple, les auteurs montrent que les différents logiciels courants qui proposent des statistiques textuelles, à commencer par les traitements ou éditeurs de texte, ne comptent pas les unités « mots » de la même manière. Pour l'utilisateur non averti, le résultat paraît infaillible et incontestable, alors qu'il repose sur des choix méthodologiques discutables. L'apparente simplicité de l'outil informatique ne dispense donc pas le chercheur de se poser les questions nécessaires à l'interprétation des résultats, ce qui passe par la compréhension du fonctionnement de l'outil mobilisé. Les auteurs rappellent par ailleurs que la démarche de l'AD, qu'elle soit ou non outillée informatiquement, commence par la formulation d'une problématique, suivie de la construction d'un corpus destiné à valider ou à infirmer des hypothèses : l'utilisation de l'outil informatique nécessite d'avoir, en amont, identifié les données, les méthodes et les outils qui permettent de répondre à la problématique.
- 4 Les chapitres 2 et 3 traitent de la constitution du corpus, dans ses principes généraux en AD d'abord, puis dans sa mise en œuvre pratique à partir de trois exemples d'analyses outillées informatiquement. Les auteurs rappellent qu'un corpus est un objet construit — en fonction de questions, d'hypothèses, d'une conception de l'AD, des outils du chercheur ; qu'il est contextualisé, d'où la nécessité de récolter des métadonnées ; enfin, que le corpus est souvent contrastif / comparatif. Les trois scénarios de constitution de corpus présentés permettent de faire ressortir « les implications pratiques liées à l'informatisation des données et à l'automatisation des analyses » (p. 63) dans les différentes étapes de constitution du corpus : sa délimitation, le recueil ou la collecte de données brutes, le formatage de celles-ci pour les rendre exploitables, la structuration, le partitionnement puis le balisage des données et, enfin, l'annotation du corpus. À travers ces exemples, les auteurs mettent en évidence la nécessité de préparer les données à l'aide d'outils utilisés en amont du logiciel d'analyse, et indiquent la possibilité d'automatiser un certain nombre d'opérations. Ce processus de modelage permet en outre au chercheur de connaître son corpus et d'affiner ainsi ses questionnements.
- 5 Le chapitre 4 revient sur la question des unités textuelles de comptage à retenir pour l'analyse et passe en revue les choix possibles — des N-grammes et des racines lexicales aux segments répétés, cooccurrences et cooccurrences généralisées, en passant par les formes graphiques, lemmes et parties du discours. Les auteurs signalent l'émergence de

recherches sur les motifs, les patrons et les routines, dont le repérage n'est pas encore pris en charge par les logiciels, mais qui constitue une piste prometteuse.

- 6 Le chapitre 5 aborde les principaux outils utilisés en analyse des données textuelles, les approches qui leur sont associées ainsi que leurs fonctionnalités. Retenant les outils qu'ils qualifient d'« essentiellement contrastifs et longitudinaux » (p. 125), les auteurs en présentent les principales méthodes et fonctionnalités : construction d'un index lexical hiérarchique, étude des ventilations de fréquences entre différentes partitions du corpus, calcul de degrés de spécificités d'une unité donnée, étude des concordances, exploitation d'outils de topographie textuelle. Une introduction aux principes généraux de l'analyse factorielle des correspondances (AFC) est présente, mais des recherches plus approfondies seront nécessaires au lecteur qui veut s'en saisir. Une synthèse finale présente un protocole d'exploration méthodique d'un corpus, en l'occurrence le corpus *Vœux* (Leblanc, 2016). La conclusion du chapitre présente un utile tableau de correspondances entre questions méthodologiques, exemples de fonctionnalités textométriques et outils informatiques.
- 7 Enfin, le chapitre 6 met en relation des problématiques en AD et des méthodes en analyse outillée par informatique. Deux problématiques sont abordées : faire émerger ce dont parle le corpus (thèmes, événements, faits sociaux) et prendre le mot comme témoin (AD à entrée lexicale) ; chaque problématique est déclinée à partir de plusieurs approches. Là encore, les études de cas viennent appuyer des développements plus généralisants, ce qui permet au lecteur de saisir les enjeux de ce qui est présenté.
- 8 Comme les auteurs l'expliquent en début d'ouvrage, la méthode d'élaboration collective du volume ne correspond pas à la traditionnelle répartition « un chapitre — un auteur » ; elle repose plutôt sur un tissage plus serré des contributions, dont les équipes de rédaction sont partiellement communes. Si cette méthode est très efficace pour générer des croisements de perspectives et des rapprochements, elle suscite parfois quelques difficultés de lecture (terme mobilisé avant d'être défini, ou inversement défini à plusieurs reprises). Il n'en reste pas moins que la lecture de cet ouvrage peut être recommandée à tous les étudiants et les chercheurs désireux de s'outiller. Si un seul volume ne peut suffire pour devenir compétent, ce manuel est à concevoir comme une très bonne boussole pour cartographier le domaine, prendre ses repères et pouvoir poursuivre l'exploration.

BIBLIOGRAPHIE

LEBLANC, Jean-Marc (2016), *Analyses lexicométriques des vœux présidentiels*, Londres, ISTE Éditions.

AUTEURS

MAUDE VADOT

Maude Vadot est maîtresse de conférences en sciences du langage et didactique des langues,
Université Savoie Mont Blanc. Courriel : maude.vadot@univ-smb.fr