Lidil

Revue de linguistique et de didactique des langues

57 | 2018

**Démarches créatives, détours artistiques et
appropriation des langues**

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

Presses universitaires de Rennes, 2017, 250 p.

Olivier Kraif

Édition électronique

URL : <http://journals.openedition.org/lidil/4852>

DOI : 10.4000/lidil.4852

ISSN : 1960-6052

Éditeur

UGA Éditions/Université Grenoble Alpes

Édition imprimée

ISBN : 978-2-37747-048-8

ISSN : 1146-6480

Référence électronique

Olivier Kraif, « Émilie Née (dir.), *Méthodes et outils informatiques pour l'analyse des discours* », *Lidil* [En ligne], 57 | 2018, mis en ligne le 01 mai 2018, consulté le 24 septembre 2020. URL : <http://journals.openedition.org/lidil/4852> ; DOI : <https://doi.org/10.4000/lidil.4852>

Ce document a été généré automatiquement le 24 septembre 2020.

© Lidil

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

Presses universitaires de Rennes, 2017, 250 p.

Olivier Kraif

RÉFÉRENCE

Émilie Née (dir.), *Méthodes et outils informatiques pour l'analyse des discours*, Presses universitaires de Rennes, 2017, 250 p.

- 1 Cet ouvrage collectif est conçu comme un manuel à vocation didactique, qui s'adresse aux étudiants et aux chercheurs intéressés par les méthodes et outils informatiques mis en œuvre dans le cadre de l'analyse des discours. Comme le notent les auteurs « la spécificité du présent ouvrage est de questionner ces outils et leurs méthodes en les articulant directement à des problématiques d'analyse du discours » (p. 15). Il ne s'agit donc pas simplement de présenter une série de ressources et d'outils, en détaillant sous un angle technique une gamme d'instruments et de fonctionnalités utiles à la recherche, pour aider le lecteur à se les approprier et à comprendre comment s'en servir. Dans sa construction même, à partir de nombreuses études de cas, l'ouvrage se propose d'articuler les données et les outils manipulant ces données aux questions de recherche posées en amont de la constitution des données.
- 2 La démarche interprétative inhérente à l'analyse des discours, qui cherche à en reconstruire le sens à travers leur contextualisation, en les situant non seulement par rapport à des situations spécifiques d'énonciation mais aussi par rapport aux autres discours auxquels ils répondent dans le dialogisme de la circulation des formes et des expressions, impose également de questionner la manière dont on a constitué ses données (*i.e.* son corpus d'étude) et les traitements informatiques qu'on leur a appliqués. De même que les discours ne sont pas « transparents », les données ne

parlent pas d'elles-mêmes, et les outils ne peuvent les faire parler que si l'on est à même d'avoir le recul épistémologique pour comprendre comment on les a construites, afin d'être à même de les interpréter. L'immédiateté intuitive de certaines représentations informatiques, telles que les nuages de points d'une analyse factorielle des correspondances (AFC) qui peuvent montrer visuellement des proximités entre différents textes et/ou différentes unités lexicales, peut s'avérer par exemple trompeuse. Interpréter ces observables n'est possible que si l'on est capable de reparcourir, à rebours, tout le cheminement qui a permis d'y mener, et de répondre à toute une série de questions, telles que : comment lire une AFC ? quelles unités y sont mises en corrélation ? comment a-t-on identifié ces unités ? comment a-t-on segmenté le corpus en unités ? a-t-on utilisé des lemmes ou des formes ? a-t-on enrichi le corpus par une annotation automatique telle que l'étiquetage ? quels choix a-t-on faits dans la sélection des textes ? comment et pourquoi a-t-on constitué le corpus ?, etc.

- 3 Si les outils informatiques aboutissent à une délinéarisation du texte, une fragmentation en facettes plus ou moins disjointes ou éclatées (tables de cooccurents, index hiérarchique, segments répétés, concordances, etc.), il convient de ne jamais rompre le fil d'Ariane interprétatif, qui seul permet de remonter au sens à travers le labyrinthe de l'analyse, tout en reliant les discours entre eux : c'est là le projet ambitieux de cet ouvrage, dont le grand mérite est d'y parvenir de manière didactique et simple, en illustrant toujours les questions méthodologiques par un retour aux données, un grand nombre d'illustrations étant fournies dans des encadrés consacrés à la description de cas d'études concrets.
- 4 Ainsi, après une entrée en matière illustrant les difficultés liées à une opération faussement simple, le comptage des mots, les questions abordées dans l'ouvrage abordent de manière progressive les différentes étapes d'une démarche d'analyse instrumentée, en soulevant les questions méthodologiques et pratiques posées par chacune d'elles : la constitution du corpus, le dépouillement des unités (formes, lemmes, parties du discours, patrons morphosyntaxiques, segments répétés, cooccurrences), le choix des logiciels (outils contrastifs vs outils structurants) et de leurs fonctionnalités (index hiérarchique, tableau lexical entier, ventilation des fréquences, calcul de spécificités, concordances, représentation topographique, AFC). Un dernier chapitre synthétique montre comment, de manière plus globale, ces outils peuvent s'intégrer dans la méthodologie de l'analyse du discours. Avec un souci constant de clarté et d'efficacité pédagogique, chaque chapitre se termine sur un encadré résumant les notions-clés par une synthèse de ce qu'il faut retenir.
- 5 Enfin, l'ouvrage se ferme sur des fiches pratiques synthétiques permettant d'approfondir, sur un plan technique, différents outils utiles au chercheur qui voudra aller plus loin : le langage des expressions régulières, les commandes Unix pour la mise en œuvre de chaînes de traitement simples, l'aspiration de sites web pour la constitution de corpus et l'analyse factorielle des correspondances.
- 6 Bien que toujours explicitement enraciné dans la tradition de l'analyse du discours, notons enfin que cet ouvrage aborde la question des méthodes et des outils d'exploration de corpus d'un point de vue élargi : il pourra constituer également une très bonne introduction pour la mise en œuvre des mêmes outils en linguistique de corpus, en textométrie, voire en sémantique interprétative.

AUTEURS

OLIVIER KRAIF

LIDILEM, Université Grenoble Alpes