

Exploration fonctionnelle de dictionnaires de synonymes pour un projet d'humanités numériques. Étude comparative de ressources numériques et imprimées

Nathalie Gasiglia
Université de Lille

Karolina Suchecka
Université de Lille

La lexicographie numérique élargit les raisons qui poussent à consulter des dictionnaires¹ en facilitant de nouvelles consultations, notamment par des outils informatiques d'extractions ciblées d'informations. C'est ce que cette contribution propose d'envisager comme perspective.

Le contexte est un projet de littérature comparée : l'édition savante numérique d'un corpus de 84 textes en français relatant le mythe d'Orphée et Eurydice, dans lequel nous voulons détecter puis présenter les réemplois intertextuels. Pour ce faire, parmi d'autres procédés de réécriture, nous travaillons sur les relations synonymiques. Les ressources lexicales utilisées pour des traitements comparables (comme *WOLF*²) et les méthodes statistiques³ s'avèrent peu opératoires pour ce petit corpus. Afin d'atteindre la plus fine annotation synonymique possible des mots-occurrences (parfois très polysémiques) de notre corpus, nous nous tournons vers des dictionnaires de synonymes.⁴

Notre objectif serait d'automatiser l'appariement d'une sélection pertinente de synonymes à chaque mot-occurrence lors du prétraitement du corpus, ce qui impliquerait, d'une part, de disposer d'un dictionnaire interrogeable par un programme et proposant des articles subdivisés en sous-listes de synonymes, comme dans les dictionnaires de synonymie distinctive⁵, et, d'autre part, de savoir exploiter informatiquement les sélecteurs sémantiques ou contextuels qui introduisent souvent les sous-listes pour les confronter aux indices repérables dans les contextes des mots-occurrences. À défaut, l'association des synonymes pertinents est manuelle, mais appuyée sur un dictionnaire « bien choisi ».

Notre objectif, dans le cadre de cette contribution, est d'analyser une sélection de dictionnaires de synonymes numériques (*Dictionnaire Électronique des Synonymes*⁶, *Atlas sémantiques*⁷, dictionnaires des correcteurs ortho-grammaticaux *Antidote*⁸ et *Le Robert correcteur*⁹, onglets « synonymes » de sites de dictionnaires généraux *Larousse.fr*¹⁰, *Dico en Ligne du Robert*¹¹, *Wiktionnaire*¹²) mais aussi imprimés¹³ (*Dictionnaire des synonymes de Hachette* [2003]¹⁴, *Grand Dictionnaire Synonymes & Contraires* de Larousse [2004]¹⁵,

Dictionnaire de synonymes et nuances du Robert [2005]) afin de définir quel modèle de traitement dictionnaire est le mieux exploitable pour l'annotation de corpus que nous voulons réaliser.

- 1 Schématiquement : aides à la compréhension, à l'expression ou à l'amélioration de la compétence métalinguistique.
- 2 *WordNet Libre du Français*, Inria, ANR EDyLex, cf. <http://alpage.inria.fr/~sagot/wolf.html>.
- 3 Comme le « plongement lexical », par exemple avec l'algorithme *Word2vec* développé par le groupe de recherche de Google, sous la direction de Thomas Mikolov.
- 4 Les synonymes introduits comme annotations sont exploités par les outils de détection automatique de correspondances intertextuelles (cf. Suchecka, Gasiglia & Ziger 2019 et Suchecka & Gasiglia à paraître). Suchecka, Karolina et Nathalie Gasiglia [à paraître]. « Digital Comparative Edition and Text Similarity Detection Tools: Hyper-textual Cartography of Mythological Rewriting ». *Plotting Poetry 4. Tackling the Toolkit*.
Suchecka, Karolina, Nathalie Gasiglia et Karl Zieger, 2019. « Édition comparative intermédiaire de séries traductives : exploiter les homologies pour créer des visualisations modulables ». *Traitement Automatique des Langues*, TAL et humanités numériques, 60 (3) : 37-62.
- 5 C'est souvent une liste unique de synonymes qui est fournie en ligne, comme dans les dictionnaires de synonymie cumulative.
- 6 Cf. <http://crisco.unicaen.fr/dictionnaire-electronique-des-synonymes/actualites-des/>.
- 7 Cf. <https://www.atlas-semantic.eu/>.
- 8 Cf. <https://www.antidote.info/fr> et <https://www.antidote.info/fr/antidote-10/documentation/guide-utilisation/introduction/les-dictionnaires>.
- 9 Cf. <https://robert-correcteur.lerobert.com/aide-redaction.php>.
- 10 Cf. <https://www.larousse.fr/dictionnaires/francais/>.
- 11 Cf. <https://dictionnaire.lerobert.com/>. Il reprend le contenu du *Dictionnaire de synonymes et nuances* imprimé du Robert (2005).
- 12 Cf. <https://fr.wiktionary.org/>.
- 13 Leur exploitation impliquerait de numériser les articles utiles.
- 14 Également sur disque : *Dictionnaire Hachette. Langue française & synonymes*, CD-ROM PC / Mac, Paris, Hachette Multimédia, 2001; nouv. pressage, 2002.
- 15 Également consultable en ligne, cf. <https://fr.thefreedictionary.com/>.