KURDI, Mohamed Zakaria (2018) : *Traitement automatique des langues et linguistique informatique. Vol. 2. Sémantique, discours et applications. Sciences cognitives.* Londres : ISTE Éditions, 323 p.

Andreea Ghiță

Volume 64, numéro 3, décembre 2019

URI : <https://id.erudit.org/iderudit/1070547ar>

DOI : <https://doi.org/10.7202/1070547ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer ce compte rendu

Ghiță, A. (2019). Compte rendu de [KURDI, Mohamed Zakaria (2018) : *Traitement automatique des langues et linguistique informatique. Vol. 2. Sémantique, discours et applications. Sciences cognitives.* Londres : ISTE Éditions, 323 p.] *Meta*, 64(3), 836–839. <https://doi.org/10.7202/1070547ar>

NOTES

- Comme le note l'auteur, la néologie est peu conceptualisée dans les pays où l'héritage wüstérien est plus direct; la faible considération pour la dimension diachronique semble en être la raison principale: «[...] on peut considérer qu'il existe un hiatus entre la théorie de Wüster, pour qui la terminologie est complètement synchronique, et le postulat de la néonymie, qui est par définition diachronique, voire panchronique» (Guilbert 1973: 26).
- «Les familles lexicales semblant avoir horreur du vide, quelque linguiste avisé créera, si ce n'est déjà fait, le *néographe* nécessaire à *néologie*. [...] 500 MN [500 mots nouveaux définis et expliqués, par Jacques Cellard et Micheline Sommant, 1979, dont J.-Cl. Boulanger faisait alors le compte rendu] propose le terme *néonyme*, qui tendrait à remplacer *néologisme!* [sic] [...] [R]appelons que G. Rondeau a proposé dans un récent colloque sur la terminologie (Moscou, novembre 1979) le terme *néonymie* pour désigner les néologismes techniques et scientifiques» (Boulanger 1981: 382). John Humbley fait état (p. 35) de divers termes forgés pour parler de néonymes (dont *néoterminisme* – proposé par Alain Rey [Rey 1979 – référence non citée dans la bibliographie], repris ponctuellement par l'école fonctionnelle de Martinet –, et *néonyme*). Il fait l'hypothèse que le terme *néoterminologie*, employé par Julie Pelletier et Andy Van Drom – alors étudiants-chercheurs à l'Université Laval –, aurait pu être inspiré par la norme ISO 1087-1: 2000. Il est plus probable que le néonyme *néoterme* soit repris de Jean-Claude Boulanger, qui en faisait usage dès le début des années 1980, en lui associant une nuance qui le distingue de son proche voisin *néonyme*: «On nous permettra une remarque supplémentaire à propos de *néonymie* et du dérivé *néonyme*. Le raffinement terminologique et sémantique peut tendre vers une telle perfection que certains objecteront que *néonymie* et *néonyme* ne sont pas satisfaisants. Étymologiquement *néonymie* et *néonyme* incluent *nom*, *mot* et *terme* (ou *terminologisme*). Il serait peut-être plus juste d'utiliser *néoterme* ou *néoterminologisme* plutôt que *néonyme* lorsqu'il est question de terminologie?» (Boulanger 1983: 319 – un peu plus tôt dans ce texte, Boulanger propose aussi *néonymiste* pour parler des acteurs du domaine [Boulanger 1983: 315]; voir aussi Boulanger 1989). *Néoterme* était, chez Jean-Claude Boulanger, le pendant, pour les usages spécialisés, de *néomot*.
- Ce dont l'auteur est bien conscient: «Les historiens des sciences ne se trompent pas lorsqu'ils se penchent sur les questions de terminologie dans le but de comprendre les évolutions scientifiques» (p. 27).

RÉFÉRENCES

- BOULANGER, Jean-Claude (1981): Compte rendu de *500 mots nouveaux définis et expliqués* par Jacques CELLARD et Micheline SOMMANT [1979, Paris/Gembloux: Duculot]. *Le français moderne*. 49(1):379-382.
- BOULANGER, Jean-Claude (1983): Synonymie, néonymie et normalisation en terminologie. Commentaire d'un exposé d'Alain Rey. In: Diane DUQUET-PICARD et Marian BUGARA-ADSHEAD, dir. *Problèmes de la définition et de la synonymie en terminologie*. (Colloque international de terminologie, Québec, 23-27 mai 1982). Québec: Girsterm, 311-327.
- BOULANGER, Jean-Claude (1989): L'évolution du concept de néologie de la linguistique aux industries de la langue. In: Caroline DE SCHAEZTEN, dir. *Terminologie diachronique*. (Colloque «Terminologie diachronique, Bruxelles, 25-26 mars 1988). Paris: Conseil international de la langue française/Ministère de la communauté française de Belgique, 193-211.
- GARDIN, Bernard (1974): La néologie, aspects sociolinguistiques. *Langages*. 36:67-73.
- GUILBERT, Louis (1975): *La créativité lexicale*. Paris: Larousse.
- HUMBLEY, John (1993): L'observation de la néologie terminologique. *La banque des mots*. 5:65-74.
- HUMBLEY, John (2006): La néologie: interface entre ancien et nouveau. In: Rosalind GREENSTEIN, dir. *Langues et cultures: une histoire d'interface*. Paris: Publications de la Sorbonne, 93-103.
- HUMBLEY, John (2012): Retour aux origines de la terminologie: l'acte de dénomination. *Langue française*. 174:111-129.
- REY, Alain (1979): *La terminologie*. Paris: Presses universitaires de France.
- QUEMADA, Gabrielle, dir. (1983): *Dictionnaire de termes nouveaux des sciences et des techniques*. Paris: Conseil international de la langue française.
- SABLAYROLLES, Jean-François (2000): *La néologie en français contemporain*. Paris: Honoré Champion.
- KURDI, Mohamed Zakaria (2018): *Traitement automatique des langues et linguistique informatique*. Vol. 2. *Sémantique, discours et applications*. Sciences cognitives. Londres: ISTE Éditions, 323 p.
- Dans son ouvrage, Mohamed Zakaria Kurdi, professeur au Département d'informatique du Lynchburg College (Virginie, États-Unis), s'applique à dresser la deuxième partie de son «bilan panoramique» (p. 12) du traitement automatique des langues (TAL), comme lui-même l'avance, un effort qui est couronné de succès en février

2018, grâce à la parution du présent livre chez les éditions ISTE. Publié sous la direction de Patrick Paroubek, ingénieur de recherche au CNRS, ce tome, tout comme le premier, passe en revue, de manière scientifique et avertie, des théories issues des disciplines comme l'intelligence artificielle (IA), la linguistique, l'informatique, la psychologie cognitive et les neurosciences, qui toutes influent sur le TAL. De plus, les thématiques abordées dans ce deuxième volume ainsi que celles incluses dans le premier (Kurdi 2017) s'avèrent un choix évident pour l'auteur. Ce dernier s'est, à présent, investi dans les recherches portant sur le dépouillement de textes ainsi que sur les systèmes de tutorat intelligent pour l'enseignement des langues étrangères qui intègrent, parmi d'autres applications, le traitement de la langue naturelle et les réseaux neuronaux.

Cet ouvrage, disponible en format papier et électronique, est interdisciplinaire, suivant un modèle de succès qui marie des approches tantôt traditionnelles tantôt modernes, cherchant, peut-être, à toucher un nombre plus grand de lecteurs avec des intérêts disciplinaires divers. Il se veut aussi équitable, dans le sens où l'on a voulu y incorporer et décrire des approches tant linguistiques que computationnelles. À l'instar d'ouvrages d'autres auteurs qui l'ont précédé et qui ont cherché à aboutir au même type de structure, comme par exemple, Fuchs, Lacheret-Dujour *et al.* (1993), Enjalbert (2005) ou encore Biskri et Jebali (2011), pour n'en citer que quelques-uns, les deux tomes de Kurdi ont l'avantage de décrire des applications élaborées dans les dix dernières années.

Le lecteur aurait cependant apprécié un aperçu de la technique de traduction neuronale; cette lacune est quelque peu frustrante. Il faut néanmoins souligner le fait que l'auteur a choisi de diviser son travail de synthèse en deux tomes, ce qui lui a permis d'intégrer plusieurs notions et théories se rapportant aux champs d'investigation qui s'entrecroisent avec le TAL.

Quant à l'architecture du deuxième tome qui fait l'objet du présent compte-rendu, elle renferme quatre chapitres et un nombre considérable de sous-chapitres, sections et sous-sections. Une telle structure est cependant facile à suivre pour le lecteur. L'ouvrage va à l'essentiel, ne comportant ni dédicace, ni préface, ni avant-propos, trait qui se manifeste davantage à travers les sections qui ne dépassent que sporadiquement trois pages. Par contre, après la bibliographie, l'auteur ajoute un index qui n'est pas du tout redondant et un sommaire particulièrement utile.

Le livre s'ouvre sur une introduction qui parcourt tant schématiquement que diachroniquement, d'abord, la connexion entre la langue et le TAL. Il s'agit de retracer le développement de la

traduction automatique (TA), une des applications centrales du TAL, qui à la suite du fameux rapport ALPAC¹ de 1966 va suivre deux voies différentes: la TA traditionnelle qui joue sur la syntaxe et la TA rattachée à l'IA qui repose sur une analyse sémantique (Villard 1989: 56-57). Ensuite, le deuxième paragraphe reconnaît l'importance du domaine de la recherche d'information et de son pendant, la science des données, en ce qui concerne la conception des applications de TAL de celles-ci. L'objectif du livre, quant à lui, est exposé dans le paragraphe suivant, tandis que le dernier paragraphe de l'introduction fait l'inventaire du contenu de chacun des quatre chapitres de l'ouvrage.

Le premier chapitre nous plonge dans la sphère du lexique et des connaissances. La partie qui lui est consacrée est divisée entre un recensement des notions ayant trait à la sémantique lexicale et une introduction aux bases de données lexicales comme Wordnet, Prolex et Brulex. La fin du chapitre offre au lecteur une vue d'ensemble sur les ontologies ainsi que sur les formalismes de représentation des connaissances comme les *réseaux sémantiques*, les *graphes conceptuels* et les *schémas*.

La section 1.1, intitulée «Sémantique lexicale», aurait bénéficié, à notre avis, d'une perspective sur les relations syntagmatiques, par exemple en couvrant le concept de *degré de figement* et les *expressions figées*, qui depuis longtemps font l'objet de travaux dans le domaine du TAL. Notons toutefois que l'auteur avait consacré une page aux *termes*, *collocations* et *colligations* dans son premier tome.

Tout au long des pages consacrées au chapitre suivant, Kurdi mène une étude approfondie de la sémantique en recensant quelques théories fondamentales de sémantique combinatoire et, ensuite, des notions propres à la sémantique formelle. La section 2.1 offre des informations pertinentes sur la sémantique interprétative, la sémantique générative, la grammaire des cas, la sémantique interprétative de Rastier et la théorie sens-texte.

Bien que les deux premiers chapitres soient convaincants, il aurait été préférable de créer une section distincte portant sur les approches cognitives en sémantique lexicale, qui sont la nouvelle tendance en linguistique. Cette suggestion résulte du fait que le livre touche à des notions comme la métaphore conceptuelle et la théorie du prototype, historiquement importantes pour le développement de la sémantique cognitive. D'autres théories issues de la sémantique lexicale cognitive auraient mérité d'y être couvertes davantage. On peut citer, notamment, la métonymie, la polysémie ou la sémantique des cadres de Fillmore (1982) qui a conduit à la création du modèle computationnel *FrameNet*².

Néanmoins, il faut absolument noter que l'auteur apporte des précisions essentielles quant aux théories et notions de la linguistique cognitive qu'il passe en revue. C'est précisément le cas de la métaphore conceptuelle qu'il situe dans le courant cognitif bien qu'il le fasse de manière assez laconique, lui consacrant environ trois paragraphes et deux illustrations (p. 14-16).

Le troisième chapitre, quant à lui, fait l'objet d'un inventaire non seulement des notions principales qui tiennent du discours mais aussi des approches logiques du discours, dont la segmentation linéaire, l'interprétation du discours et le traitement de l'anaphore.

Dans le dernier chapitre qui fait le point sur les applications du TAL, l'auteur entreprend de décrire la traduction automatique et dresse aussi le portrait de la recherche d'information et de l'extraction de celle-ci. Nous regrettons toutefois la non-prise en compte de la *traduction automatique neuronale*, étant donné que ce type de technologie ne cesse de s'implanter et d'évoluer. De fait, ce type de traduction, qui est né de l'intelligence artificielle en 2003 (Bengio, Ducharme *et al.* 2003: 1137-1155) et qui utilise des réseaux neuronaux similaires à ceux du cerveau humain pour encoder et décoder le langage, est très prometteur. Il est dommage que l'auteur n'ait pas brossé un tableau raisonné de la traduction automatique neuronale.

Le succès de ce type de technologie s'est concrétisé quand il est devenu accessible gratuitement à l'échelle mondiale, notamment par l'entremise de Google Translate et de DeepL³. Google Brain Team⁴ a créé GTNM⁵, la machine de traduction neuronale de Google, dont la qualité des résultats surpasse nettement celle des systèmes fondés sur la TA statistique⁶. En effet, son module d'intelligence artificielle est à même d'apprendre et de se perfectionner seul.

Outre une nouvelle présentation du contenu des chapitres de l'ouvrage, la conclusion du volume contient des propos qui viennent renforcer l'idée que le TAL doit répondre à de nombreux défis dans des domaines aussi divers que la politique, le marketing et la traduction. L'auteur met à nouveau l'accent sur les applications du TAL comme instrument de travail de nombreux acteurs d'horizons différents, d'où leur importance. En guise de conclusion, Kurdi affirme (p. 283) qu'il reste quand même confiant que la linguistique développera des modèles mieux adaptés à la réalité, en dépit du fait qu'à présent, elle semble plutôt éclipsée par les avancées technologiques.

Des questions de forme viennent cependant perturber la lecture de cet ouvrage, notamment le format des citations. Par exemple, la linguiste Eleanor Rosch est citée comme [ROS 73] ou [ROS 76]. Cela est désagréable, mais pas insurmontable,

car l'auteur utilise son nom et prénom dans la même section. Par contre, le symbole [RUL 00] cache en fait le nom du linguiste Ruland Tobias, dont on ne cite pas le nom dans la section où il est mentionné. Cela pose problème, surtout dans la version numérique du livre, dans laquelle on doit abuser de la fonction *rechercher*.

D'une façon générale, ce livre est plaisant à parcourir; l'information y est bien structurée et présentée dans une langue facile à comprendre, ce qui, pour un novice en matière de linguistique informatique, représente un attrait particulier. Un autre avantage de ce livre est qu'il renferme de nombreux tableaux et graphiques qui décrivent de manière visuelle les notions linguistiques ou les principes de fonctionnement computationnel.

Bref, cet ouvrage serait parfait si ce n'était des lacunes informationnelles identifiées plus haut et le format de citation employé. Je n'hésite pas à le recommander.

ANDREEA GHIȚĂ

Universitatea Transilvania, Braşov, Roumanie

NOTES

1. Le Automatic Language Processing Advisory Committee a été fondé en 1964 à la demande du gouvernement des États-Unis afin d'évaluer les progrès faits au sein de la linguistique computationnelle et, en particulier, de la traduction automatique.
2. Un projet dont l'objectif est de créer une base de données lexicales électronique à partir de la théorie des cadres, qui décrit les cadres conceptuels des phrases.
3. Un service de traduction performant dont le fonctionnement repose sur des algorithmes d'apprentissage par réseau neuronal profond qui ont conduit à des progrès importants et rapides dans la traduction automatique.
4. L'équipe de recherche de Google qui utilise l'intelligence artificielle pour développer des logiciels de pointe dans des domaines divers.
5. *Google Translation Neural Machine*, logiciel de traduction automatique neuronale qui a été rendu accessible au public à la fin de 2016.
6. En anglais, *statistical machine translation* (STM), traduction automatique utilisant des modèles statistiques pour apprendre les correspondances bilingues dans des corpus parallèles.

RÉFÉRENCES

- BENGIO, Yoshua, DUCHARME, Réjean, VINCENT, Pascal *et al.* (2003): A neural probabilistic language model. *Journal of Machine Learning Research*. 3:1137-1155.
- BISKRI, Ismaïl et JEBALI, Adel (2011): *Traitement automatique des langues naturelles: de l'analyse à l'application*. Paris: Hermès.

- ENJALBERT, Patrice, dir. (2005): *Sémantique et traitement automatique du langage naturel*. Paris: Hermès/Lavoisier.
- FILLMORE, Charles John (1982): *Frame Semantics*. In: LINGUISTIC SOCIETY OF KOREA, dir. *Linguistics in the Morning Calm: Selected Papers from SICOL-1981*. (First Seoul International Conference on Linguistics, Séoul, 29 juillet au 7 août 1981). Séoul: Hanshin, 111-137.
- FUCHS, Catherine, LACHERET-DUJOUR, Anne et VICTORRI, Bernard (1993): *Linguistique et traitement automatique des langues*. Paris: Hachette.
- KURDI, Mohamed-Zakaria (2017): *Traitement automatique des langues et linguistique informatique*. Vol. 1. *Parole, morphologie et syntaxe*. Sciences cognitives. Londres: ISTE Éditions.
- VILLARD, Masako (1989): Traduction automatique et recherche cognitive. *Histoire Épistémologie Langage*. 11(1):55-84.

VIDAL CLARAMONTE, María Carmen África (2018): *La traducción y la(s) historia(s). Nuevas vías para la investigación*. Granada: Comares, 143 p.

Résultat du projet de recherche intitulé «Violencia simbólica y traducción: retos en la representación de identidades fragmentadas en la sociedad global», *La traducción y la(s) historia(s). Nuevas vías para la investigación* réunit les plus récentes réflexions d'África Vidal. Comme l'affirme Edwin Gentzler dans la présentation du livre, il s'agit d'une étude novatrice faite par l'une des responsables du «power turn» en traductologie qui, cette fois, amplifie considérablement ce paradigme. L'étude a sa place parmi celles qui depuis les années 1980 portent sur les rapports entre traduction et pouvoir, comme les travaux de Gayatri Spivak (1988), Paul Bandia (2009) et Georges Bastin (2010, 2011). L'auteur fait entendre les voix des femmes, des immigrants et des autochtones par le biais d'un ensemble de nouvelles méthodologies développées au long de l'ouvrage. Pour ce faire, África Vidal s'appuie sur des historiens tels que Hayden White (1987), Dominick LaCapra (1985) et Alum Munslow (2013). Le livre comporte, à part la préface de Gentzler, six chapitres et une conclusion.

Dans le premier chapitre, «La traducción y el mar de historias», l'auteur part de l'œuvre de Salman Rushdie, *Haroun and the Sea of Stories*, où il compare le monde des récits à l'océan. Comme celui-ci, les narrations sont fluides et se transforment toujours. On se rapproche ainsi d'un contexte épistémologique post-positiviste et d'une conception ouverte de l'histoire. L'idée de simulacre de Baudrillard, pour qui «la réalité est un reflet du réel», c'est-à-dire la réalité

est une construction, permet à África Vidal de penser la traduction comme un médium pour le changement politique, social ou culturel. D'où son importance pour d'autres disciplines, comme l'histoire, qui, grâce à la traduction, développent d'autres outils d'analyse, menant, selon Backmann-Medick (2009), à une sorte de «translation turn» dans les sciences humaines. Ce rapprochement à l'histoire nous dirige vers une historiographie critique pour laquelle l'histoire est narration. Des auteurs comme Hayden White, Michel Certeau et Dominick LaCapra permettent d'envisager une histoire anti-téléologique, faite de fragments et qui n'accepte pas de hiérarchisation. Ainsi, écrire l'histoire, c'est réécrire des réalités, ce qui ferait de l'historien un traducteur.

Cette conception élargie de la traduction offre la possibilité de voir l'Histoire comme un concept problématique. Dans «La Historia, un concepto problemático», deuxième chapitre du livre, África Vidal nous rappelle que ce n'est qu'au 18^e siècle que l'histoire devient une discipline autonome. En quelques pages, on va de Quintilien et Cicéron, pour qui traduction et narration se confondent, en passant par Durkheim et Benjamin qui mettent en cause l'objectivité scientifique de l'observation du passé, pour arriver à l'histoire critique de Lawrence Stone, Michel de Certeau et Hayden White qui rapprochent à nouveau et de façon décisive histoire et narration. Le rôle central du langage comme espace de construction de l'histoire amène l'auteur au post-structuralisme de Derrida, Barthes et Foucault qui corroborent les principes énoncés par les historiens critiques. On touche ainsi à l'une des questions clés du livre: qui sommes-nous quand nous nous racontons l'histoire? La réponse varie en fonction des contextes, une fois que les représentations qu'on se fait sont le résultat des interprétations que chaque groupe élabore selon ses intérêts.

Le troisième chapitre, «El peligro de una sola historia», est un complément nécessaire au chapitre précédent. On ne peut accepter une conception ouverte de l'histoire sans dénoncer les dangers de l'imposition d'une seule version de l'histoire. Contre l'hégémonie d'une seule interprétation, souvent eurocentrique, África Vidal évoque «la radicalisation des politiques culturelles de la différence», partagée par des courants théoriques aussi diversifiés que l'École de Francfort (Marcuse, Adorno, Horkheimer), les marxismes italien et français (Sartre, Althusser, Gramsci), le structuralisme (Lévi-Strauss, Todorov), le post-structuralisme (Foucault, Derrida, Deleuze), le néo-pragmatisme de Richard Rorty en histoire, ainsi que l'orientalisme de Said ou le féminisme de Cixous. Mais c'est le nouvel historicisme qui va l'intéresser davantage, surtout la microhistoire, spécialement par le fait