

Introduction. Traductologie de corpus : 20 ans après

Natalie Kübler, Rudy Loock et Mojca Pecman

Volume 63, numéro 3, décembre 2018

Traductologie de corpus : 20 ans après

URI : <https://id.erudit.org/iderudit/1060163ar>

DOI : <https://doi.org/10.7202/1060163ar>

[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 document

Kübler, N., Loock, R. & Pecman, M. (2018). Introduction. Traductologie de corpus : 20 ans après. *Meta*, 63(3), 577–582. <https://doi.org/10.7202/1060163ar>

INTRODUCTION

Traductologie de corpus : 20 ans après

NATALIE KÜBLER*

Université Paris Diderot, Paris, France
nkubler@eila.univ-paris-diderot.fr

RUDY LOOCK

Université de Lille, Lille, France
Laboratoire « Savoirs, Textes, Langage » du CNRS, Lille, France
rudy.loock@univ-lille3.fr

MOJCA PECMAN*

Université Paris Diderot, Paris, France
mpecman@eila.univ-paris-diderot.fr

C'est il y a exactement vingt ans que Sara Laviosa dirigeait un numéro de *Meta* intitulé *L'approche basée sur le corpus* dans lequel cette approche était présentée comme un nouveau paradigme (Laviosa 1998). Depuis, *Meta* a publié de nombreux articles dans lesquels l'approche basée sur le corpus est devenue familière à ses lecteurs. En 2011, un deuxième numéro, intitulé *Les corpus et la recherche en terminologie et traductologie*, et dirigé par Marc Van Campenhoudt et Rita Temmerman, confirmait le changement de paradigme dans les études en traductologie et dans des disciplines liées, comme la terminologie et la lexicologie. C'est donc le troisième numéro en vingt ans que *Meta* consacre intégralement à la traductologie de corpus. Si Laviosa (1998 : 1) présentait l'approche basée sur le corpus (*corpus-based translation studies*) comme un domaine prometteur dans les études en traduction, en analyse contrastive et dans la formation des traducteurs, cette approche est aujourd'hui bien installée et a permis de développer des recherches qui se sont grandement élargies à d'autres territoires, tout en tirant parti des progrès rencontrés dans les outils statistiques de traitement de corpus. L'objectif de ce numéro¹ consiste à illustrer la richesse des différents territoires explorés par cette approche, son dynamisme et la solidité du tournant pris il y a une vingtaine d'années avec ce que Van Campenhoudt et Temmerman (2011 : 224) nomment le « bras informatisé de la linguistique descriptive », et donc, aujourd'hui, de la traductologie. À cela, il convient d'ajouter le « bras statistique », en raison du tournant statistique pris en sciences du langage et, simultanément, en traductologie (Oakes et Ji 2012).

Le numéro est divisé en trois parties dans lesquelles les différents auteurs abordent à la fois des questions épistémologiques et théoriques et présentent des analyses empiriques et/ou à visée pédagogique. La première partie rassemble des recherches portant sur des textes traduits, la deuxième présente des travaux dans un domaine qui s'est grandement développé ces dernières années, celui de l'interprétation ; elle se clôt par un article qui cherche à montrer que les barrières entre l'interprétation et la traduction tendent à disparaître. Enfin, la troisième partie aborde les corpus en traduction et traductologie comme des outils d'aide à la traduction ou à

l'enseignement. Nous cherchons ici à donner un aperçu, non exhaustif, non seulement de la diversité des objets d'étude du domaine, mais aussi à soulever la question de la distinction entre traductologie, traduction, interprétation, enseignement et outils. On peut en effet faire l'hypothèse que l'approche basée sur le corpus a permis de rapprocher de plus en plus ces différentes branches de la traductologie, en fondant les analyses et les résultats sur des attestations authentiques et concrètes, issues d'analyses quantitatives et qualitatives structurées sur le corpus.

1. Les textes traduits : idéologie, stratégies et influences

Le premier article, proposé par Kaibao Hu et Xiaoqian Li, a pour objectif de synthétiser et poser les bases théoriques et méthodologiques de la traductologie critique reposant sur le corpus (*corpus-based critical translation studies*, ou CCTS). Les auteurs expliquent comment la traductologie descriptive et l'analyse de discours critique (*critical discourse analysis* ou CDA) ont pu fusionner pour en arriver à une approche critique de l'idéologie dans la traduction. Ils argumentent en faveur de l'approche basée sur le corpus qui permet des analyses approfondies et précises, tirant parti des études sur l'idéologie en corpus. Leur objectif consiste à clarifier l'approche CCTS par rapport à la traductologie descriptive (*descriptive translation studies*: DTS), à expliciter les domaines de recherche de l'approche CCTS, à évoquer les méthodologies possibles et à délimiter les contributions théoriques et pratiques de cette approche pour la traductologie critique (CTS). Des propositions concrètes d'analyses s'inspirant de la linguistique systémique fonctionnelle sont abordées et illustrées. Enfin, pour les auteurs, la traductologie de corpus représente non seulement un changement important dans la méthodologie de l'étude de l'idéologie en traduction, mais elle permet aussi d'explorer de nouvelles questions et de réduire la subjectivité du chercheur en s'appuyant sur une approche statistique et sur des faits authentiques.

Dorothy Kenny et Mali Saththachai s'attaquent ensuite à relever des caractéristiques appartenant aux universaux relatifs de la traduction en recherchant des attestations d'explicitation et d'éléments dits « uniques » dans la traduction juridique en thaï. Comme les études de traductologie sont encore rares en thaï, malgré un intérêt croissant, cet article vise à développer une recherche encore embryonnaire. La réforme juridique entreprise en Thaïlande entre 1890 et 1930 présente l'étonnante caractéristique d'avoir été rédigée en anglais et ensuite traduite en thaï. Bien qu'aujourd'hui, la législation soit rédigée en thaï directement, l'emploi du passif, influencé par la présence prégnante du passif en anglais juridique, peut poser des problèmes en thaï puisque certaines formes présentent une coloration négative. Cette question est étudiée de deux manières: d'une part, par l'analyse des traductions du passif anglais en thaï dans un corpus parallèle, et d'autre part, par l'analyse des structures passives dans un corpus comparable de thaï traduit et de thaï original. Les deux analyses se font au travers du prisme de l'explicitation et de la recherche d'éléments « uniques », à savoir une forme de la langue cible qui n'existe pas dans la langue source, ici des formes du passif thaï qui peuvent avoir une valeur négative. Les résultats montrent, d'une part, que, contrairement à ce qui était attendu, la stratégie d'explicitation n'est pas caractéristique pour ce type de texte dans cette paire de langues. D'autre part, il semble que la langue source, l'anglais, transparaisse dans la

traduction thaïe par une surreprésentation des passifs. Cette étude aborde donc les difficultés de travailler avec une langue peu traitée, propose une méthode d'extraction des données pertinente et contribue à approfondir les connaissances sur la stratégie de l'explicitation et des éléments uniques dans les textes traduits.

Poursuivant sur l'influence de la langue source dans les textes traduits, Lore Vandevoorde cherche à savoir dans quelle mesure la langue source, le type de texte et les cognats (ou mots apparentés) influent sur les choix lexicaux dans les textes traduits. Se détachant de la notion d'universaux absolus de la traduction, elle se tourne vers la recherche de tendances générales, ce qui nécessite des techniques statistiques sophistiquées, dont elle démontre l'intérêt. Ce travail est donc pertinent dans ce volume pour deux raisons. D'une part, il illustre le développement et l'utilité de l'outil statistique complexe dans les études sur les textes traduits en permettant de construire deux visualisations des choix lexicaux, et d'autre part, il relativise la notion d'universaux de la traduction en s'appuyant sur trois facteurs: la plus ou moins grande proximité de la langue source (pour des traductions du français et de l'anglais vers le néerlandais, la première langue source étant plus éloignée que la seconde de la langue cible), le type de texte (les registres selon Biber) et la proximité des cognats (une variable calculée statistiquement). Vandevoorde démontre ainsi l'importance de ces trois critères et poursuit des recherches approfondies et systématiques sur les choix lexicaux des traducteurs d'un point de vue onomasiologique.

2. De l'interprétation à la médiation linguistique

Si Shlesinger (1998) proposait d'explorer l'interprétation de manière efficace en se basant sur des corpus d'interprétation, la collecte de corpus d'interprétation diffère de celle des corpus de textes traduits et reste délicate. Caterina Falbo identifie deux interrogations concernant la collecte de corpus d'interprétation: celle de la représentativité et celle de la transcription qui doit tenir compte de l'objectif pour lequel le corpus est collecté. Par la description de deux corpus qu'elle a constitués, elle aborde les difficultés à la fois techniques et humaines qui président à la constitution de tels corpus et propose une réflexion sur la représentativité qui s'appuie sur les réflexions développées en linguistique de corpus.

Dans une perspective proche, Natacha Niemants problématise la constitution d'un corpus d'interprétation de dialogues qui pose des problèmes encore différents, notamment concernant la temporalité et l'interaction. Elle décrit les choix effectués pour constituer son corpus et teste deux logiciels (*EXMARaLDA*² et *ELAN*³) sur un sous-ensemble de 65 rencontres. Enfin, par l'analyse d'éléments lexicaux potentiellement intéressants pour décrire les structures de l'interaction, ainsi que par l'extraction de segments alignés à l'audio dont elle démontre qu'ils peuvent être appliqués utilement à l'enseignement, elle se livre à un véritable plaidoyer en faveur de l'approche corpus pour l'interprétation de dialogue en milieu médical. Son travail permet d'ouvrir de nombreuses perspectives de recherche dans un domaine qui met encore peu en œuvre l'approche basée sur le corpus.

C'est en s'attaquant à une question encore peu traitée dans le domaine de l'interprétation de conférence simultanée que Camille Collard, Bart Defrancq et Heike Przybyl démontrent la variété des questions que l'approche basée sur le corpus permet non seulement d'aborder, mais aussi auxquelles elle peut apporter des réponses. Si

un certain nombre de chercheurs ont abordé la question de l'effort cognitif exigé d'un interprète pour passer d'une langue SOV, comme l'allemand ou le néerlandais, à une langue SVO, comme l'anglais ou le français par exemple, peu d'études se sont penchées sur l'interprétation vers une langue SOV à partir d'une langue SVO. En analysant un corpus de discours et de leurs interprétations au Parlement européen, les auteurs cherchent à voir si les interprètes développent des stratégies de raccourcissement de la parenthèse verbale et d'extraposition dans deux langues SOV, l'allemand et le néerlandais. Afin de répondre à cette question, les auteurs détaillent avec soin la méthodologie de constitution des données et de l'annotation du corpus. Leur étude tire parti des développements statistiques que connaît l'approche basée sur le corpus, ce qui leur permet de démontrer quelles stratégies sont mises en œuvre pour minimiser l'effort cognitif des interprètes. Bien que celui-ci semble être en effet similaire, leur étude laisse voir que ces stratégies peuvent différer entre ces deux langues SOV.

Alors que longtemps on a travaillé de manière séparée sur les textes traduits et sur l'interprétation, Adriano Ferraresi, Silvia Bernardini, Maja Milicevic Petrovic et Marie-Aude Lefer interrogent la problématique de la simplification dans ce qu'ils appellent l'anglais « médié » (*mediated English*) dans le corpus intermodal du Parlement européen EPTIC. En étudiant l'anglais traduit et interprété à partir du français et de l'italien, ils cherchent à détecter si les stratégies de simplification diffèrent selon la langue source ou selon la modalité de médiation linguistique. Outre les résultats qui sont présentés et montrent que la langue source est un facteur relativement prédominant sur la modalité pour certaines stratégies de simplification, l'intérêt de ce travail réside dans la démonstration d'une tendance de plus en plus marquée en traductologie de corpus, à savoir le gommage des frontières entre traduction et interprétation entre autres modes de médiation, et l'application de méthodologies utilisées en traduction à l'interprétation et vice-versa (Vandevoorde, Daems et Defrancq, à paraître).

3. Corpus, aide à la traduction et à la formation

L'autre face de la recherche dans l'approche basée sur le corpus est représentée non plus par l'analyse de textes traduits ou d'interprétation, mais par l'utilisation des corpus comme outils d'aide à la traduction et d'aide à la formation des traducteurs.

Une première approche, proposée par Mavina Pantazara et Eleni Tziafa, consiste à vérifier et valider les glossaires multilingues à l'aide des corpus. Se penchant sur un glossaire de 800 entrées sur la crise économique grecque et qui propose des équivalents dans plusieurs langues, les auteurs adoptent une démarche systématique dans laquelle l'élaboration de corpus multilingues spécialisés leur permet de valider et mettre à jour les termes existants, ainsi que de relever les variantes terminologiques en français, anglais et grec. Elles illustrent ainsi l'intérêt de l'approche corpus dans la confection de bases de données multilingues destinées aux traducteurs professionnels.

Kyo Kageura adopte une approche de traitement quantitatif de terminologie en corpus dans laquelle, en s'appuyant sur un traitement statistique avancé, il illustre comment évaluer deux aspects importants permettant de choisir des textes spécialisés pour l'enseignement de la traduction spécialisée. L'apprenti traducteur doit en effet maîtriser un certain nombre de facteurs pour pouvoir traduire un texte, dont les deux facteurs étudiés ici, à savoir, d'une part, l'ensemble des concepts à maîtriser

dans un texte pour pouvoir le traduire, et d'autre part, les unités de discours basé sur des termes techniques dans un texte spécialisé. À l'aide de techniques statistiques avancées, Kageura propose ainsi un outil élégant de caractérisation des textes spécialisés.

Avec l'apparition de la traduction automatique neuronale, on constate que le métier de traducteur est en train de subir un certain nombre de bouleversements, ce qui amène Rudy Loock à comparer les résultats de la traduction automatique (TA) neuronale par rapport à la TA statistique et à la langue originale, dans le but de former les futurs traducteurs à une utilisation informée de la TA neuronale en sachant quels sont les éléments à vérifier absolument. Dans son étude, il démontre en effet que certaines structures dans une traduction automatique de l'anglais vers le français sont surreprésentées, en faisant l'hypothèse que la langue source est peut-être à l'origine de ce phénomène. Cependant, une étude qualitative reste à faire, afin de vérifier l'hypothèse.

L'approche basée sur le corpus a aussi permis de développer les corpus d'apprenants de la traduction. Natalie Kübler, Alexandra Mestivier et Mojca Pecman présentent une méthodologie de l'enseignement de la traduction basée sur le corpus et qui leur permet de collecter des corpus d'apprenants de la traduction. Elles cherchent ainsi à vérifier si l'utilisation des corpus dans le processus de traduction chez des apprentis traducteurs donne de meilleurs résultats que la traduction sans corpus, en analysant le corpus d'apprenants de la traduction qu'elles ont constitué et annoté à l'aide d'une typologie d'erreurs. Les résultats montrent une amélioration des traductions avec corpus sur certains éléments linguistiques, confirmant en cela l'hypothèse que l'utilisation des corpus dans la formation des traducteurs joue un rôle important. L'ensemble de ces contributions permet donc d'avoir un panorama, certes non exhaustif, mais illustrant la variété et le développement des thématiques que l'approche sur corpus permet d'aborder.

Ce numéro devait être dirigé avec l'aide du professeur Guy Aston, de la Università di Bologna en Italie. Guy Aston nous a quittés le 31 octobre 2018. C'est donc en son honneur, et avec le cœur lourd, que nous lui dédions ce numéro de *Meta*. Après avoir complété une maîtrise ès arts (*Master of Arts*) à l'Université d'Oxford et une maîtrise ès sciences (*Master of Science*) à l'Université d'Édimbourg, il a soutenu sa thèse au London Institute of Education. Le professeur Aston a travaillé tout d'abord à la Università degli studi di Bari, puis à la Università politecnica delle Marche (Ancône), pour finalement prendre une chaire à Bologne, où, depuis 1991, il a enseigné à la Scuola Superiore di Lingue Moderne per Interpreti e Traduttori (SSLMIT) [École supérieure de langues et littératures modernes pour interprètes et traducteurs] à Forlì. Il a été doyen de la SSLMIT de 2002 à 2005. Expert en linguistique de corpus, il est l'un des fondateurs du congrès bisannuel *Teaching and Language Corpora* (TaLC) et a participé à la création du congrès *Corpus Use and Learning to Translate* (CULT). Guy Aston est internationalement reconnu et a été un pionnier dans l'utilisation des corpus pour l'enseignement et l'apprentissage des langues étrangères, ainsi qu'en traduction et interprétation. Sa recherche se poursuivra grâce aux nombreux étudiants qu'il a formés, mais aussi aux nombreux collègues auxquels il a fait découvrir les corpus, avec sa grande générosité et finesse d'esprit. Guy Aston restera toujours dans nos mémoires.

NOTES

- * Membres du Centre de Linguistique Inter-langues, de Lexicologie, de Linguistique Anglaise et de Corpus-Atelier de Recherche sur la Parole (CLILLAC-ARP), Université Paris Diderot.
1. Une partie des articles est issue du 1^{er} Congrès Mondial de Traductologie qui s'est tenu à l'Université de Paris X à Nanterre du 10 au 14 avril 2017. Voir le site Web du Congrès, <<https://cmt.u-paris10.fr/cmt-2016/>>.
 2. HAMBURGER ZENTRUM FÜR SPRACHKORPORA (9 mai 2017) : *EXMARALDA*. Hambourg : Universität Hamburg. Consulté le 1 décembre 2018, <<http://exmaralda.org>>.
 3. THE LANGUAGE ARCHIVE (4 avril 2018) : *ELAN*. Version 5.2. Nimègue : Max Planck Institute for Psycholinguistics. Consulté le 1^{er} décembre 2018, <<https://tla.mpi.nl/tools/tla-tools/elan/>>.

RÉFÉRENCES

- LAVIOSA, Sara, dir. (1998) : L'approche basée sur le corpus. *Meta*. 43(4).
- LAVIOSA, Sara (1998) : The Corpus-based Approach: A New Paradigm in Translation Studies. In: Sara LAVIOSA, dir. *L'approche basée sur le corpus*. *Meta*. 43(4):474-479.
- OAKES Michael P. et Ji, Meng, dir. (2012) : *Quantitative Methods in Corpus-Based Translation Studies*. Amsterdam/Philadelphie : John Benjamins.
- SHLESINGER, Miriam (1998) : Corpus-based Interpreting Studies as an Offshoot of Corpus-based Translation Studies. *Meta*. 43(4):486-493.
- VAN CAMPENHOUDT, Marc et TEMMERMAN, Rita (2011) : Les corpus et la recherche en terminologie et en traductologie. In: Marc VAN CAMPENHOUDT et Rita TEMMERMAN, dir. *Les corpus et la recherche en terminologie et en traductologie*. *Meta*. 56(2):223-225.
- VAN CAMPENHOUDT, Marc et TEMMERMAN, Rita, dir. (2011) : Les corpus et la recherche en terminologie et en traductologie. *Meta*. 56(2).
- VANDEVOORDE, Lore, DAEMS, Joke and DEFRANCO, Bart (à paraître) : *New Empirical Perspectives on Translation and Interpreting*. Routledge Advances in Translation and Interpreting Studies. Londres/New York : Routledge.