Traduire
Revue française de la traduction

237 | 2017
La tête dans la toile

Web, corpus, traduction : l'exploitation par les traducteurs des données du web

Rudy Loock

Édition électronique
URL : http://journals.openedition.org/traduire/939
DOI : 10.4000/traduire.939
ISSN : 2272-9992

Éditeur
Société française des traducteurs

Édition imprimée
Date de publication : 1 décembre 2017
Pagination : 23-32
ISSN : 0395-773X

Référence électronique
Rudy Loock, « Web, corpus, traduction : l'exploitation par les traducteurs des données du web », *Traduire* [En ligne], 237 | 2017, mis en ligne le 01 décembre 2017, consulté le 29 juin 2019. URL : http://journals.openedition.org/traduire/939 ; DOI : 10.4000/traduire.939

Web, corpus, traduction : l'exploitation par les traducteurs des données du web

Rudy Loock

Au moment où nous rédigeons ces lignes, on dénombre plus de 1,3 milliard de sites internet(1). Le web, depuis son apparition, mais surtout depuis sa démocratisation avec l'arrivée de l'ADSL, contient des données en un nombre que l'on peut considérer aujourd'hui comme infini. Leur consultation par tout un chacun est désormais des plus aisées grâce aux moteurs de recherche bien connus tels que Google, Bing, ou encore le français Qwant, faisant du web une base de données gigantesque au sein de laquelle il est possible d'effectuer des recherches très pointues. Alors que nous sommes au cœur d'une révolution numérique, véritable révolution industrielle qui est loin d'être terminée et qui bouleverse bon nombre de nos habitudes (communication, consommation, mais aussi marché du travail…) comme l'ont fait d'autres révolutions industrielles avant elle, l'exploitation de ces données représente un véritable enjeu. Nous sommes entrés dans l'ère des *big data* (mégadonnées), et pour de nombreuses professions, il s'agit d'un réel défi auquel le secteur de la traduction n'échappe pas. Dire que le problème aujourd'hui n'est pas le manque mais la pléthore d'informations et la façon de les sélectionner, de les trier, de les organiser, afin d'opérer des choix pertinents correspondant à des besoins spécifiques est devenu un lieu commun. Les traducteurs, pour qui la question des outils informatiques s'est posée dès les années 1990 avec l'arrivée des logiciels d'aide à la traduction et des mémoires de traduction, peuvent effectuer sur le web leurs recherches documentaires. Ils peuvent également y trouver toute une série d'outils à l'aide desquels ils parviennent à accroître à la fois leur productivité et la qualité de leurs traductions. Ainsi, il est possible d'accéder facilement et gratuitement à d'excellents dictionnaires et glossaires terminologiques en ligne (IATE, Termium Plus…), à des lexiques spécialisés (tels que ceux publiés par la Délégation générale à la langue française et aux langues de France), ou encore à des mémoires de traduction en accès libre (par exemple celles de la Direction générale de la traduction de la Commission européenne). Mais outre le fait de favoriser la consultation, le web lui-même peut également être exploité en tant que base de données linguistiques afin d'effectuer des

(1) Source : http://www.internetlivestats.com/

recherches terminologiques, phraséologiques, ou encore liées à l'organisation du discours, notamment spécialisé. Bien loin de la simple utilisation des moteurs de recherche, qui n'est pas sans danger, il s'agit alors de compiler et d'exploiter ce que l'on appelle des corpus électroniques, dont l'utilisation s'est développée en linguistique dès les années 1960 avec la linguistique de corpus, et qui sont soit disponibles et prêts à l'emploi grâce à une interface en ligne, soit à compiler soi-même afin de répondre à un besoin donné, comme un projet de traduction sur un sujet très spécialisé. Contrairement aux idées reçues, l'exploitation et la compilation de corpus ne sont ni techniques ni chronophages ; elles ne sont pas non plus réservées aux chercheurs en linguistique et ne génèrent pas de dépenses supplémentaires. Les traducteurs semblent ne pas s'y tromper puisqu'après une longue période de méfiance (on a longtemps parlé d'un rendez-vous manqué entre les traducteurs et les outils de corpus), ils s'approprient désormais ces outils qui leur donnent au final davantage de liberté dans leurs prises de décision, comme en témoignent le nombre de formations sur le sujet, dans les universités ou ailleurs (écoles d'été, ateliers lors de conférences, formations continues, etc.).

L'objectif de cet article est d'expliquer dans quelle mesure il est possible d'exploiter les données du web afin de compiler des corpus avec pour objectif la résolution de problèmes de traduction. Une telle démarche exige en effet un certain nombre de précautions afin de s'assurer de la pertinence et de la fiabilité des données exploitées. Toutes les données ne se valent pas. Les traducteurs savent mieux que quiconque l'importance de l'adéquation optimale entre la recherche documentaire et le projet de traduction. Tout aussi cruciale est la rapidité avec laquelle cette recherche documentaire peut être effectuée : le rapport coût/bénéfice revêt une grande importance et l'utilisation de tout outil doit se révéler rentable et déboucher sur une plus grande productivité et/ou sur une meilleure qualité du texte traduit. L'exploitation de données issues du web sous forme de corpus ne fait pas exception à cette règle et doit répondre à ces exigences du métier.

Dans un premier temps, nous allons donc définir ce que l'on entend par corpus et préciser en quoi ces derniers peuvent être utiles aux traducteurs, en opérant une distinction entre les corpus prêts à l'emploi et ceux nécessitant une compilation préalable. Nous nous interrogerons également sur la pertinence de considérer le web comme un corpus, et mettrons en garde contre un certain nombre d'outils séduisants mais, au final, peu pertinents pour les traducteurs.

1. Corpus et traduction

1.1. Qu'est-ce qu'un corpus ?

Un corpus électronique est un ensemble de données linguistiques (textes écrits ou retranscriptions de discours oraux) rassemblées en un seul et même endroit, qui peut être interrogé

automatiquement via une interface en ligne ou via un programme informatique hors ligne (concordancier). En effet, un corpus ne se lit pas, il s'interroge. Il contient un échantillon représentatif d'une langue ou d'une variante de langue (variante géographique ; registre, domaine de spécialité particulier) et peut être associé à d'autres corpus à des fins de comparaison. Les données peuvent être brutes (le corpus ne contient que le texte) ou annotées (le corpus peut par exemple être étiqueté, chaque mot étant associé à sa catégorie grammaticale). Les données sont par ailleurs authentiques ; on ne les « crée » pas pour la circonstance.

1.2. Quels outils pour quelles tâches ?

Les traducteurs peuvent exploiter des corpus électroniques pour différentes raisons, déterminant l'utilisation de différents types de corpus.

Ainsi, ils exploiteront un corpus de langue étrangère afin d'affiner leur compréhension du texte source. En cela, les corpus sont d'excellents compléments aux outils plus traditionnels (dictionnaires, grammaires, etc.) car ils peuvent fournir de nombreux exemples de l'utilisation, en contexte, d'un mot, d'une expression, ou d'une structure syntaxique. Au-delà de ces exemples, les corpus pourront donner des informations précieuses en lien avec l'usage réel de la langue, que les dictionnaires ne peuvent pas toujours fournir : le terme recherché est-il limité à certains registres ? est-il technique ? correspond-il à une variante géographique de la langue précise ? a-t-il une connotation ou une prosodie sémantique particulières ? le trouve-t-on systématiquement au sein de schémas syntaxiques bien précis ?

Les traducteurs pourront également exploiter des corpus en langue cible, bien souvent leur langue maternelle, possibilité particulièrement intéressante lorsqu'il s'agit de rédiger une traduction dans un domaine très spécialisé où une vérification de la terminologie, de la phraséologie, mais aussi de l'organisation discursive s'impose. Ainsi, à l'aide d'un corpus électronique constitué de rapports économiques du Fonds monétaire international, par exemple, le traducteur pourra vérifier la manière dont un terme est utilisé en contexte par les spécialistes du domaine. En particulier, ce type d'observation permet de mettre au jour les phénomènes de collocation : avec quels adjectifs, verbes, ou adverbes le terme s'associe-t-il de façon naturelle ? L'objectif du traducteur est alors d'atteindre l'invisibilité en rédigeant son texte comme un expert du domaine le ferait.

Par extension, en exploitant un corpus en langue source et un corpus en langue cible pour un même domaine (on parle alors de **corpus comparable**), le traducteur pourra établir des glossaires bilingues. En compilant deux corpus contenant des articles scientifiques d'un même domaine, il est possible d'établir une liste des termes les plus fréquents dans les deux sous-corpus et de faire ainsi émerger des équivalences entre les deux langues considérées. L'accès aux termes en contexte permet également d'observer leur utilisation linguistique naturelle par des spécialistes du domaine. Ce type d'observation peut concerner la terminologie, la

phraséologie, mais aussi l'organisation du discours (préférence pour des structures différentes selon la langue afin d'exprimer le même contenu logico-sémantique).

Enfin, et il s'agit probablement là du type de corpus le plus connu et le plus utilisé des traducteurs (même si l'appellation « corpus » n'est pas employée), il est possible de consulter des textes traduits associés aux textes originaux, afin de prendre connaissance de traductions déjà effectuées par d'autres professionnels. On parle alors de **corpus parallèles**, qui sont souvent segmentés et alignés au niveau de la phrase, ce qui n'est pas sans rappeler les mémoires de traduction que l'on trouve au sein des logiciels de traduction assistée par ordinateur (TAO). En ce sens, il s'agit pour les traducteurs d'exploiter les corpus comme source d'inspiration.

Par manque de place, il ne nous est pas possible de donner des exemples concrets, mais nous pouvons renvoyer à Loock (2016a : 103-179) ou encore (2016b), disponible en ligne, où sont détaillées les différentes utilisations des corpus électroniques dans le cadre de projets de traduction. De nombreuses expériences ont été menées dans les universités françaises et étrangères afin de former les futurs traducteurs à l'exploitation de données issues de corpus électroniques en vue d'améliorer leurs traductions, s'agissant notamment de la fluidité de la langue cible dans des domaines spécialisés (cf. travaux de Lynne Bowker, Ana Frankenberg-Garcia, Cécile Frérot, Natalie Kübler, pionnière en la matière en France, Federico Zanettin, ou encore nos propres travaux). Ces expériences ont souvent été concluantes et montrent que les corpus électroniques sont de véritables outils d'aide à la traduction, qui viennent s'ajouter à la boîte à outils des traducteurs déjà très fournie en ce début de XXIe siècle. On constate d'ailleurs à l'heure actuelle une intégration progressive des outils de corpus au sein des logiciels de TAO, ce qui témoigne de la popularité de ces outils auprès des traducteurs.

1.3. Corpus prêts à l'emploi et corpus à compiler soi-même

Le traducteur peut être amené à utiliser deux grands types de corpus : les « **corpus officiels** » prêts à l'emploi et les corpus à compiler soi-même. Les premiers sont disponibles directement sur internet et leurs données accessibles, gratuitement ou moyennant un abonnement, via une interface prête à l'emploi, plus ou moins sophistiquée en fonction de la présence, ou non, d'annotations. Il en existe pour de nombreuses langues, même si l'on note une prépondérance pour la langue anglaise. Citons à cet égard les corpus disponibles via le site de la Brigham Young University(2), qui propose toute une série de corpus de langue anglaise interrogeables depuis une interface ergonomique et conviviale : le *British National Corpus*, le *Corpus of Contemporary American English*, le *Wikipedia Corpus*, le *News On the Web Corpus*, le *Global Web-Based English Corpus*, etc. Ces corpus contiennent des centaines de millions de mots, parfois au-delà du milliard, englobant différentes variantes géographiques mais aussi différents registres.

(2) (https://corpus.byu.edu/)

S'agissant de la langue française, le corpus le plus connu est sans doute Frantext(3), qui contient près de 300 millions de mots de français du XIIe au XXIe siècles, mais qui fait la part belle aux textes littéraires et ne sera que peu utile au traducteur spécialisé. Bien qu'il existe d'autres corpus en ligne, nous ne disposons encore pour la langue française d'aucun corpus de référence comparable au *British National Corpus* et au *Corpus of Contemporary American English* mentionnés ci-dessus.(4) Bon nombre de corpus électroniques pour de nombreuses autres langues sont accessibles en ligne et il nous est impossible d'en proposer une énumération exhaustive ici. Mentionnons néanmoins l'existence du très intéressant site internet *Sketch Engine*(5), portail (payant) à partir duquel plus de 400 corpus pour près d'une centaine de langues sont interrogeables.

Lorsque ces corpus ne suffisent pas, ce qui est souvent le cas pour le traducteur spécialisé car les corpus en ligne ne sont pas toujours suffisamment exhaustifs ou récents, il est possible de compiler ses propres corpus. On parle alors de « **corpus maison** » ou encore de « **corpus DIY** » (pour *do-it-yourself*). À partir de données linguistiques personnelles ou collectées sur internet (rapports ou articles au format *.pdf, *.doc, *.docx, *.html…), le traducteur peut créer des corpus en rassemblant ces données sous forme de fichiers standardisés et convertis dans un format spécifique(6), exploitables hors ligne à l'aide d'un programme informatique appelé « concordancier », qui permet d'effectuer des recherches simultanées au sein de ces différents documents. Cela évite au traducteur d'avoir à lire l'ensemble des documents ou même d'y effectuer de simples recherches grâce à l'outil « rechercher au sein du document ». On sait en effet, notamment grâce à l'enquête menée dans le cadre du projet MeLLANGE *(Multilingual eLearning in LANGuage Engineering)*, que les traducteurs consultent des pages web pour leurs recherches documentaires, qu'ils lisent en intégralité ou qu'ils parcourent à l'aide de la fonction « rechercher » de leur navigateur. Il s'agit pourtant là d'une perte de temps, car si la collecte des documents est incontournable (phase de compilation), ces derniers peuvent ensuite être facilement consultés de façon automatique (phase d'exploitation), par exemple en visualisant simultanément toutes les occurrences d'un même terme au sein de *n* documents différents dans un contexte restreint (lignes de concordance, cf. Figure 1) ou plus large (cf. Figure 2). Rappelons en effet qu'un corpus ne se lit pas, mais s'interroge, ce qui au passage permet d'augmenter le nombre de documents consultés lors d'une recherche documentaire,

(3) (http://www.frantext.fr/)

(4) Cela ne saurait tarder. Divers projets sont en cours, notamment le Corpus de Référence du Français Contemporain (CRFC), constitué de plus de 300 millions de mots de français écrit et parlé (Siepman et al. 2017).

(5) (https://www.sketchengine.co.uk/)

(6) Le format de fichier à utiliser est le format *.txt (texte brut) avec un encodage UTF-8 (*Universal Character Set Transformation Format* - 8 bits) permettant de prendre en compte l'ensemble des caractères possibles, comme les caractères diacritiques ou les alphabets autres que l'alphabet latin. Cet encodage est disponible au sein des traitements de texte et doit être sélectionné lors de l'enregistrement au format *.txt. Il s'agit là de la seule manipulation un peu technique, mais qui n'est guère plus complexe qu'un enregistrement au format *.pdf.

et donc de diminuer la prise de risques au moment de la décision. Des fichiers de tous types peuvent être très aisément convertis en corpus : le corpus exploité au sein des Figures 1 et 2 a été compilé à partir de 5 rapports économiques du FMI (rapports *World Economic Outlook*) au format *.pdf téléchargés sur le site de l'institution. Cette compilation d'un corpus de plus de 225 000 mots s'est faite en à peine quelques minutes, l'étape la plus longue étant la conversion au format requis pour l'exploitation au sein du concordancier(7), mais celle-ci s'opère automatiquement grâce à un convertisseur de fichiers(8). Les concordanciers offrent naturellement davantage de fonctionnalités, telles que les listes de mots les plus fréquents, la recherche de collocations ou encore les expressions régulières(9).

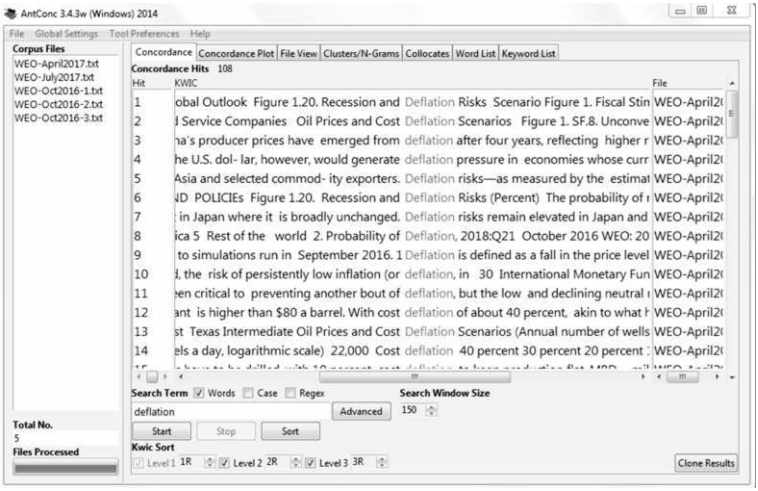

Figure 1. Lignes de concordance pour la requête *deflation*

(7) Le concordancier utilisé ici est AntConc version 3.4.4, très facile à prendre en main et téléchargeable gratuitement (http://www.laurenceanthony.net/software/antconc/). Il en existe d'autres (ex. : Kwic Concordance, #LancsBox, WordSmith Tools, TextStat). Nous souhaitons remercier chaleureusement Laurence Anthony de nous avoir permis d'utiliser deux captures d'écran afin d'illustrer cet article.

(8) On peut citer le logiciel AntFileConverter (http://www.laurenceanthony.net/software/antfileconverter/).

(9) Pour de plus amples informations, nous invitons le lecteur à consulter les tutoriels disponibles sur la page d'accueil d'AntConc, le didacticiel de Dominique Legallois disponible via le site de l'Université Ouverte des Humanités (www.uoh.fr), l'ouvrage Zanettin (2012), ou encore notre propre ouvrage (Loock 2016a : 131-138).

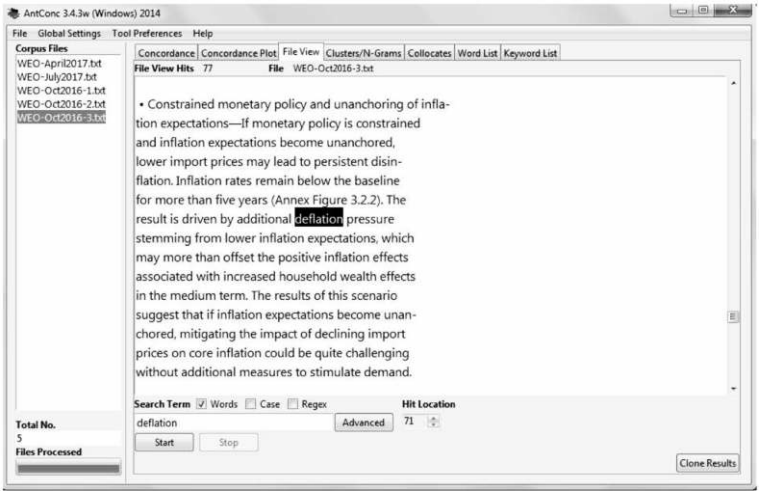

Figure 2. Contexte linguistique élargi pour l'une des occurrences du terme *deflation*

2. Le web est-il un corpus ?

Puisqu'il est possible de compiler des corpus à partir de données issues du web, on peut légitimement se demander si le web lui-même ne peut pas être considéré comme un « méga » corpus électronique interrogeable directement depuis un moteur de recherche. Quoi de plus simple, en effet, que de vérifier le caractère attesté d'une expression en la saisissant dans une barre de recherche afin d'en vérifier d'éventuelles occurrences ! Il est par ailleurs indéniable que le web représente une masse de données impressionnante, en croissance permanente, constamment mise à jour, accessible de façon simple et gratuite. Nous mettons néanmoins en garde contre l'utilisation du web comme corpus. Les données issues du web ne font l'objet d'aucun contrôle et ne peuvent pas être considérées comme représentatives d'une langue ou d'une variante de langue : les textes ont-ils été rédigés par des experts ? par des natifs de la langue ? Ont-ils été traduits, par un traducteur humain ou par une machine ? On sait également que les résultats obtenus par un moteur de recherche ne sont pas objectifs : seules les pages nouvelles ou les plus populaires sont indexées, les résultats de fréquence fournis ne sont pas fiables et peuvent varier d'un utilisateur à l'autre, voire d'un moteur de recherche à l'autre (Thelwall 2008, Kilgarriff 2007). En ce sens, comme l'affirme le titre de l'article de Kilgarriff (2007), « la Googleologie, c'est de la mauvaise science ». Il est donc dangereux de « pêcher » dans l'océan que représente le web ; comme dans les vrais océans, on y trouve d'excellents poissons mais aussi tout un tas de choses indésirables. Nous opérons donc la distinction proposée par De Schryver (2002) entre « web comme corpus » et « web pour corpus », qui sont deux approches différentes vis-à-vis des données du web, la seconde étant à privilégier. Car

ce qui importe avant tout pour l'expert linguistique qu'est le traducteur, c'est la fiabilité des données : exploiter des données issues du web pour compiler un corpus est acceptable à partir du moment où il y a eu vérification de la source des données (sites internet reconnus, documents officiels rédigés par des experts, articles scientifiques en accès libre, etc.) et de leur adéquation avec le projet de traduction en cours.

Une solution intermédiaire, peu connue mais pourtant pertinente pour le traducteur, est la création d'un moteur de recherche Google personnalisé *(Google Custom Search Engine* en anglais). Il est en effet possible de paramétrer des moteurs qui n'effectueront des recherches que parmi une série de sites choisis avec soin par l'utilisateur. Cette fonctionnalité proposée par Google(10) et normalement destinée à permettre l'intégration d'un moteur de recherche au sein de son propre site web, peut tout à fait être « détournée » à des fins linguistiques. Il est ainsi possible de se créer un moteur de recherche personnalisé limité à des sites internet spécialisés dans les questions relatives à l'informatique, par exemple, dont les requêtes ne cibleront que les sites retenus par l'utilisateur selon ses propres critères (pour le traducteur, la qualité linguistique et le degré de spécialisation). Ces moteurs de recherche personnalisés ont par ailleurs l'avantage de pouvoir être partagés et conservés pour des usages ultérieurs, contrairement au raccourci « site : » de la boîte de recherche du moteur générique.

Il existe également des logiciels permettant de récupérer automatiquement des données issues du web à partir de mots clé, données dont les sources peuvent néanmoins être filtrées et contrôlées avant validation. C'est le cas de BootCaT (http://bootcat.dipintra.it/), téléchargeable gratuitement. Plus précis, le logiciel AntCorGen (http://www.laurenceanthony.net/software/antcorgen/) permet de récupérer des données (en anglais uniquement) issues d'articles scientifiques de la revue *PLOS ONE* de la *Public Library of Science* aux États-Unis, soit des données spécialisées rédigées par des experts dans les domaines en question. Dans les deux cas, les données peuvent ensuite être exploitées hors ligne à l'aide d'un concordancier.

À la question « le web est-il un corpus ? », nous répondons donc par la négative. Il est, en revanche, tout à fait possible de collecter des données sur le web afin de compiler des corpus personnels en fonction de ses besoins, l'important étant la fiabilité des sources. Au-delà de la fiabilité, il convient d'insister également sur l'adéquation entre les données collectées et le projet de traduction : des données juridiques ne peuvent pas être exploitées pour traduire un texte médical et vice-versa, des données datant de 2006-2010 ne peuvent pas être exploitées pour traduire un texte scientifique récent, la terminologie et la phraséologie évoluant aussi avec le temps. Pour ces différentes raisons, nous mettons en garde contre l'utilisation d'outils pourtant populaires comme le *Google Books Corpus*, dont le contenu est une sorte de boîte noire inaccessible et dont la provenance des données statistiques affichées au sein du

(10) (https://cse.google.fr/cse/)

Google n-gram viewer (https://books.google.com/ngrams) reste incertaine, ou encore *Google Fight* (https://www.googlefight.fr/), qui permettrait de comparer la fréquence de l'occurrence des termes.

Conclusion

Cet article constitue une mise au point sur les éventuelles exploitations du web en tant que base de données linguistiques. Nous y défendons le principe d'une utilisation précise et réfléchie, non pas en recourant aux moteurs de recherche pourtant souvent utilisés à des fins de vérification linguistique, mais par le biais des corpus électroniques, accessibles via une interface en ligne ou exploitables hors ligne à l'aide d'un concordancier. Une telle approche ne doit pas être réservée aux chercheurs en linguistique, et les traducteurs ont tout intérêt à s'emparer des outils de corpus comme outils d'aide à la traduction pour les aider dans l'exercice de leur métier. Et nous répétons ici que les opérations de compilation et d'exploitation de corpus ne sont ni techniques ni chronophages ; elles ne génèrent pas non plus de coûts supplémentaires, mais donnent au traducteur une véritable marge de manœuvre dans ses choix décisionnels.

rudy.loock@univ-lille3.fr

Rudy Loock *est Professeur des universités en linguistique anglaise et en traductologie au sein de l'UFR « Langues Étrangères Appliquées » de l'Université de Lille. Responsable du master « Traduction Spécialisée Multilingue », il enseigne la traduction spécialisée, la traductologie, la grammaire comparée, ainsi que l'utilisation des corpus électroniques comme outils d'aide à la traduction et comme outils de recherche en traductologie. Membre de l'UMR « Savoirs, Textes, Langage » du CNRS, il consacre actuellement ses travaux de recherche à la traductologie de corpus, la qualité des traductions, et la formation des futurs traducteurs. Il a publié en 2016* La Traductologie de corpus *aux Presses Universitaires du Septentrion.*

Bibliographie

ANTHONY Laurence, 2014, AntConc (Version 3.4.3), Tokyo, Waseda University.

ANTHONY Laurence, 2015, AntFileConverter (Version 1.2.0), Tokyo, Waseda University.

ANTHONY Laurence, 2017, AntCorGen (Version 1.1.0), Tokyo, Waseda University.

DE SCHRYVER Gilles-Maurice, 2002, « Web for/as Corpus: a Perspective for the African Languages », in *Nordic Journal of African Studies* 11, p. 266-282.

KILGARRIFF Adam, 2007, « Googleology is Bad Science », in *Computational Linguistics* 33(1), p. 147-151.

LOOCK Rudy, 2016a, *La Traductologie de corpus*, Villeneuve d'Ascq, Presses Universitaires du Septentrion.

LOOCK Rudy, 2016b, « L'utilisation des corpus électroniques chez le traducteur professionnel : quand ? comment ? pour quoi faire ? », in *ILCEA 27* (Numéro spécial « Approches ergonomiques des pratiques professionnelles et des formations des traducteurs »), https://ilcea.revues.org/3835.

SIEPMANN Dirk, BÜRGEL Christoph, et DIWERSY Sascha, 2017, « The 'Corpus de référence du français contemporain (CRFC)' as the first genre-diverse mega-corpus of French », in *International Journal of Lexicography* 30(1), p. 63-84.

THELWALL Michael, 2008, « Extracting Accurate and Complete Results from Search Engines: Case study Windows Live », in *Journal of the American Society for Information Science and Technology* 59(1), p. 38-50.

ZANETTIN Federico, 2012, *Translation-driven Corpora*, Manchester, St Jerome Publishing.