

---

# Outils d'aide à la traduction : pour une intégration des corpus et des outils d'analyse de corpus dans l'enseignement de la traduction et la formation des traducteurs

**Cécile Frerot**

---

 <https://www.ouvroir.fr/cpe/index.php?id=232>

**DOI** : 10.57086/cpe.232

## **Electronic reference**

Cécile Frerot, « Outils d'aide à la traduction : pour une intégration des corpus et des outils d'analyse de corpus dans l'enseignement de la traduction et la formation des traducteurs », *Cahiers du plurilinguisme européen* [Online], 2 | 2010, Online since 01 janvier 2010, connection on 08 novembre 2023. URL : <https://www.ouvroir.fr/cpe/index.php?id=232>

## **Copyright**

Licence Creative Commons – Attribution – Partage dans les mêmes conditions 4.0 International (CC BY-SA 4.0)

# Outils d'aide à la traduction : pour une intégration des corpus et des outils d'analyse de corpus dans l'enseignement de la traduction et la formation des traducteurs

Cécile Frerot

## OUTLINE

---

1. Introduction
2. Les mémoires de traduction
  - 2.1. L'enseignement des mémoires de traduction
  - 2.2. Les différents systèmes présents sur le marché
3. Corpus et traduction
  - 3.1. L'utilisation des corpus dans l'enseignement de la traduction
  - 3.2. L'exploitation des corpus par les traducteurs professionnels
4. Un cours de traduction fondé sur les corpus
  - 4.1. Objectifs
  - 4.2. Introduction aux corpus et aux outils d'analyse de corpus
    - 4.2.1. Concepts-clés liés aux corpus
    - 4.2.2. Outils d'analyse de corpus : concordanciers
  - 4.3. Exploitation des corpus pour la traduction
    - 4.3.1. Utilisation d'un corpus comparable
    - 4.3.2. Utilisation d'un corpus parallèle
5. Conclusion

## TEXT

---

### 1. Introduction

- 1 Les outils informatiques d'aide à la traduction font de plus en plus partie de l'environnement professionnel du traducteur. Une initiation à ces outils est donc devenue incontournable pour toute formation qui se donne comme objectif de former de futurs traducteurs professionnels. Les outils d'aide à la traduction désignent généralement les logiciels de Traduction Assistée par Ordinateur (TAO), plus communément appelés mémoires de traduction – bien qu'une mémoire de traduction (MT) ne soit qu'un composant du logiciel où sont stockées

les phrases en langue source assorties de leur traduction en langue cible.

- 2 Aujourd'hui, l'utilisation des MT tend à se banaliser dans les formations de traducteurs, même si la forme que peut prendre ce type d'enseignement, loin d'être clairement définie, soulève un certain nombre de questions et de difficultés liées au mode d'apprentissage (Arrouart, 2003 ; Sauron, 2007). À côté des MT, d'autres outils issus de la linguistique de corpus sont aujourd'hui utilisés pour l'enseignement de la traduction, et l'exploitation des corpus en traduction a fait l'objet de publications au cours des dix dernières années (Bowker, Pearson, 2002 ; Zanettin, Bernardini, Stewart, 2003). L'utilisation des corpus en traduction n'est cependant pas très courante dans les formations universitaires françaises et les travaux relatant d'études sur corpus menées à des fins didactiques n'émanent que très rarement du domaine français. Hors de l'hexagone, l'approche par corpus est davantage répandue même si des efforts notables doivent être faits pour convaincre une grande partie des formations universitaires du bien-fondé des corpus comme le souligne Silvia Bernardini (2006).
- 3 C'est dans ce contexte que nous nous sommes intéressées à l'élaboration d'un cours de traduction anglais-français articulé autour de la notion de corpus, destiné à un public d'étudiants de L3 (dernière année de licence, parcours « Traduction Spécialisée ») au département de Langues Etrangères Appliquées (LEA) de l'université Stendhal de Grenoble et conçu comme un pré-requis pour l'entrée en Master de Traduction Spécialisée Multilingue.
- 4 Dans cet article, nous commencerons par rappeler l'importance d'une « culture de l'informatique » (Arrouart, 2003), qui implique l'enseignement des MT, et soulèverons quelques questions liées à l'enseignement de ces outils avant de présenter brièvement les différents systèmes disponibles sur le marché. Nous soulignerons dans une deuxième partie l'intérêt grandissant porté aux corpus pour l'enseignement de la traduction, et la nécessité d'intégrer les besoins des traducteurs professionnels dans leur formation universitaire. La troisième et dernière partie, qui constitue le cœur de notre article, sera consacrée à l'exploitation des corpus dans un cours de traduction. Après avoir défini les objectifs du cours, nous proposerons une introduction aux corpus qui permet de familiariser les apprentis traduc-

teurs avec les principaux concepts-clés du domaine et les outils d'analyse de corpus. Nous montrerons ensuite, au moyen d'activités pédagogiques ciblées, comment l'utilisation d'un corpus comparable anglais-français permet de sensibiliser les apprenants à l'apport que représentent les corpus pour l'acte même de traduction, contribuant à leur faire mener une véritable réflexion sur la langue. Avant de conclure, nous illustrerons l'utilisation d'un corpus parallèle anglais-français et montrerons comment nous confrontons les apprentis traducteurs aux outils de toute dernière génération présents sur le marché qui associent aux fonctionnalités classiques des MT des possibilités puissantes de recherche en corpus.

## 2. Les mémoires de traduction

### 2.1. L'enseignement des mémoires de traduction

- 5 Parmi les qualités et les compétences fondant l'employabilité d'un (apprenti)-traducteur, Daniel Gouadec (2006) définit comme *cruciale* la maîtrise d'outils spécifiques, au sein desquels figurent les logiciels de MT. Leur intégration dans la formation des traducteurs s'est donc imposée en réponse aux besoins du marché de la traduction et aux nouvelles compétences et pratiques qui déterminent l'environnement du traducteur professionnel. On s'inscrit plus largement dans une « culture de l'informatique » qu'évoque Catherine Arrouart (2003), présente dans tous les aspects des activités professionnelles du traducteur, et plus particulièrement dans le domaine de la localisation. Aujourd'hui, l'enseignement des MT se trouve ainsi généralisé et « toute formation en traduction digne de ce nom comprend une formation aux outils informatiques et aux programmes d'aide à la traduction » (Sauron, 2007 : 208).
- 6 Même si l'enseignement des MT tend aujourd'hui à se banaliser, il soulève un certain nombre de questions et pose des contraintes particulières dont il faut tenir compte dans le cadre de l'enseignement. Nous voudrions tout d'abord souligner l'intérêt d'ancrer l'utilisation des MT, avant tout perçues par les apprenants comme des outils *informatiques*, au sein de l'enseignement de traduction afin de sensibili-

ser les futurs traducteurs à l'acte même de traduction à l'aide d'un outil informatique. L'utilisation de systèmes dont l'approche repose sur des corpus de texte entiers nous semble à ce titre fondamentale, la recherche contextuelle étant, comme le rappelle Véronique Sauron (2007 : 214) :

l'une des fonctions les plus utilisées par les traducteurs. En effet, nombre de textes auxquels ils [les apprenants] sont confrontés dans la pratique ne sont pas vraiment répétitifs et il est important pour eux de pouvoir trouver des solutions de traduction aux problèmes terminologiques et sous-phrastiques qui se présentent à eux.

- 7 Le choix des outils à intégrer dans l'enseignement s'avère par ailleurs fondamental ; il doit refléter la réalité du marché, c'est-à-dire la diversité des MT, confronter les futurs traducteurs à des outils variés (en termes d'interface et de fonctionnalité notamment) et les doter ainsi des capacités d'adaptation qu'exige l'utilisation des outils présents dans l'environnement professionnel. Le choix doit également tenir compte de l'évolution des outils disponibles sur le marché et des différentes approches des concepteurs, notamment celle fondée sur les corpus de textes entiers.
- 8 Enfin, l'apprentissage des MT doit être progressif et dépasser le cadre d'une initiation pour garantir aux apprenants une maîtrise suffisante de ces outils dont la complexité ne cesse de croître. L'enseignement doit être mis en place suffisamment tôt dans les cursus de traduction pour (i) familiariser les apprenants avec des concepts liés aux MT qu'ils manipuleront dans d'autres enseignements (c'est le cas notamment au sein du Master de Traduction Spécialisée Multilingue du cours de *Constitution et Gestion de Documents Electroniques* et du cours de *Terminologie*), et (ii) établir ainsi des passerelles entre les différents enseignements de manière à garantir une cohésion dans la construction pédagogique.

## 2.2. Les différents systèmes présents sur le marché

- 9 De manière schématique, les systèmes à MT reposent sur deux techniques distinctes, fruits d'une évolution au cours de la dernière dé-

cennie. Dans son travail sur l'évaluation des différents systèmes, Francie Gow (2003) distingue en effet les MT conventionnelles, dont l'approche repose sur la segmentation des textes alignés en phrases, des MT plus récentes fondées sur l'utilisation de corpus de textes entiers. Les outils représentatifs des MT conventionnelles sont par exemple *Trados*, *Déjà Vu* ou *SDL* alors que les logiciels *Logiterm* ou *Multitrans* représentent les outils de dernière génération.

- 10 Ainsi, à l'approche phrastique des MT conventionnelles a succédé une approche plus textuelle, permettant notamment aux traducteurs d'accéder aux textes entiers (et au contexte) dans deux langues tout en traduisant. En effet, les principales critiques émises par les utilisateurs à l'encontre des MT conventionnelles concernent la perte de contexte et le recyclage des phrases, qui appauvrissent la qualité de la traduction ; Claude Bédard (2000) évoque à ce titre le manque de cohésion dans le style, le manque de cohésion textuelle ou bien encore la discontinuité terminologique, et parle de « salade de phrases » à laquelle se trouve confronté le traducteur.

## 3. Corpus et traduction

### 3.1. L'utilisation des corpus dans l'enseignement de la traduction

- 11 L'intérêt grandissant porté aux corpus en traduction a fait l'objet de nombreuses descriptions linguistiques contrastives mais aussi d'études sur la théorie de la traduction, le développement d'une méthodologie à partir de corpus permettant en effet de « comprendre les contraintes, les pressions et les motivations qui influencent spécifiquement l'acte traductionnel et sous-tendent cette forme unique de communication » comme l'avait souligné dans les années 1990 l'une des fondatrices des travaux à partir de corpus, Mona Baker (1998 : 1).
- 12 Dans le cadre de l'enseignement de la traduction, cet intérêt s'est manifesté par des publications au cours des dix dernières années (Laviosa, 2002 ; Zanettin, Bernardini et Stewart, 2003), l'ouvrage à visée didactique co-écrit par Jennifer Pearson et Lynne Bowker (2002) illustrant de manière très concrète l'exploitation qui peut être faite des corpus dans un cours de traduction<sup>1</sup>.

- 13 L'utilité des corpus a par ailleurs fait l'objet de quelques expériences impliquant des apprentis-traducteurs et visant à évaluer l'apport que représentent les corpus pour la traduction. L'étude expérimentale menée par Lynne Bowker (1998) a contribué à préciser dans quelle mesure le recours à un corpus comparable (corpus dans deux langues portant sur le même domaine de spécialité) améliore la qualité globale des traductions par rapport à des ressources plus traditionnelles telles que les dictionnaires. Dans cette étude, des apprentis-traducteurs ont réalisé d'une part une traduction en utilisant des dictionnaires généraux et spécialisés, et d'autre part une traduction en exploitant des corpus électroniques constitués pour l'expérience. Les résultats montrent que le recours au corpus améliore la qualité des traductions, notamment au niveau de la compréhension du domaine, du choix des termes et de l'utilisation d'expressions idiomatiques.
- 14 Dans la lignée de cette étude, Federico Zanettin (1998) a montré comment un corpus comparable pouvait être utilisé pour développer des activités pédagogiques qui améliorent la compréhension de la langue source des apprenants ainsi que leur production dans la langue cible. Dans une étude plus récente (Zanettin, 2001), l'auteur évalue l'apport des corpus pour la traduction en évaluant des traductions produites par des apprentis traducteurs (niveau licence, italien-anglais) qui ont exploité un corpus comparable informatisé à l'aide d'un concordancier (outil permettant de visualiser les mots d'un corpus dans leur contexte d'apparition). Federico Zanettin souligne l'apport des corpus comparables notamment pour respecter le genre textuel lors du transfert en langue cible (journalistique en l'occurrence) mais aussi pour mettre au jour des stratégies d'apprentissage chez les apprenants. Williams (1996) met quant à lui en évidence le fait que l'utilisation de corpus parallèles (textes en langue source assortis de leur traduction en langue cible) améliore considérablement la précision des équivalents de traduction par rapport à l'utilisation de dictionnaires bilingues. Les études visant à mesurer l'impact des corpus parallèles sur la qualité de la traduction restent cependant rares, comme le souligne Jennifer Pearson (2003) ou bien encore Sylviane Granger (2003).
- 15 Ainsi, l'ensemble de ces travaux concourt à motiver l'intégration des corpus dans un enseignement de traduction. Par ailleurs, la vocation professionnelle du Master de Traduction Spécialisée Multilingue au

sein duquel ce cours est mis en place nous incite à nous tourner du côté des professionnels pour observer dans quelle mesure ils utilisent des corpus et des outils d'analyse de corpus.

## 3.2. L'exploitation des corpus par les traducteurs professionnels

- 16 Comme nous l'avons mentionné en introduction, la mise en œuvre de cours de traduction fondés sur les corpus n'est pas courante dans le contexte français. Hors de l'hexagone, l'approche par corpus est davantage répandue même si elle n'a rien de systématique (Bernardini, 2006). Lynne Bowker et Mickael Barlow (2008) constatent à ce titre que si les MT sont bien connues des professionnels, les concordanciers semblent quant à eux être davantage utilisés dans le milieu universitaire. Étudiant le marché canadien, Lynne Bowker (2002) montre en effet que si les organismes professionnels connaissent généralement l'existence des corpus, leur intérêt se porte davantage sur les MT. D'après l'auteur, l'une des principales raisons liées à l'impopularité des corpus chez les professionnels de la traduction tient au temps requis pour leur constitution et leur exploitation, et au fait que le gain de productivité n'est pas immédiat.
- 17 Par ailleurs, une enquête menée en Europe à plus grande échelle (Mellange, 2006) contribue à apporter un éclairage intéressant sur l'utilisation des corpus comme outil de traduction. Nous retiendrons quelques éléments essentiels de cette enquête ayant impliqué 740 participants qui ont répondu à un questionnaire (90 % sont des traducteurs professionnels exerçant au Royaume-Uni, en France, en Allemagne et en Italie et 10 % sont des apprentis traducteurs).
- 18 Sur l'ensemble des participants au questionnaire : (i) plus de 40 % collectent des textes spécifiques à un domaine ; (ii) parmi ces derniers, plus de 69 % le font sous format électronique ; (iii) plus de la moitié ( $\approx 54$  %) « lisent » les corpus plutôt que d'utiliser un outil informatisé ; (iv) sur les utilisateurs d'outils informatisés ( $\approx 46$  %), environ 20 % utilisent un concordancier (rappelons ici que ces données concernent des traducteurs et des étudiants), et plus de 65 % utilisent les fonctions de recherche d'un traitement de texte.



- 19 Ces données révèlent dans l'ensemble une utilisation – relativement – faible des corpus et des outils d'analyse de corpus. Cependant, l'enquête montre le vif intérêt que les participants portent à la possibilité d'utiliser un logiciel de constitution de corpus portant sur des domaines spécifiques (84 %) et d'extraction terminologique à partir de domaines spécifiques (83,4 %). Enfin, et c'est là aussi une donnée fondamentale, environ 86 % des participants souhaitent être informés sur les possibilités liées à l'utilisation de corpus.
- 20 Ainsi, ces résultats plaident en faveur d'une meilleure diffusion des possibilités offertes par les corpus qui passe par leur intégration dans la formation des traducteurs, avec notamment l'utilisation de logiciels facilitant la constitution et l'exploitation de corpus.

## **4. Un cours de traduction fondé sur les corpus**

### **4.1. Objectifs**

- 21 Comme nous l'avons précisé dans l'introduction, le cours de traduction (anglais-français) mis en place au sein du parcours de Traduction Spécialisé et articulé autour de la notion de corpus s'adresse à des étudiants de L3 (parcours « Traduction Spécialisée ») du département de LEA de l'université Stendhal de Grenoble.
- 22 A travers cet enseignement, nous souhaitons sensibiliser les apprenants à l'apport des corpus pour la traduction et ce, à travers des activités pédagogiques ciblées impliquant l'utilisation d'un corpus comparable et d'un concordancier. Ces activités s'articulent en partie autour de l'utilisation comparée des données présentes dans le corpus avec celles fournies par les dictionnaires bilingues, ce type de ressource étant largement utilisé par les traducteurs professionnels comme le montre les résultats d'un sondage réalisé par Amélie Josselin-Leray (2005). Par ailleurs, nous souhaitons confronter les apprentis-traducteurs aux outils d'aide à la traduction présents sur le marché, qui associent aux fonctionnalités classiques des MT des possibilités puissantes de construction et de recherche dans des corpus parallèles, et ce, afin notamment de répondre aux besoins exprimés par les traducteurs professionnels.

- 23 Nous commençons dans un premier temps par proposer une introduction à l'utilisation des corpus et des concordanciers afin de familiariser les apprentis traducteurs à ce type de ressources.

## **4.2. Introduction aux corpus et aux outils d'analyse de corpus**

### **4.2.1. Concepts-clés liés aux corpus**

- 24 Nous définissons tout d'abord un certain nombre de concepts-clés liés aux corpus. C'est ainsi que nous abordons des questions relatives à la définition, la typologie et la constitution des corpus en insistant notamment sur les deux types de corpus utilisés en amont dans le cours (corpus comparable et corpus parallèle) et sur les critères de constitution des corpus (objectif visé, représentativité).

### **4.2.2. Outils d'analyse de corpus : concordanciers**

- 25 Nous avons mentionné dans la partie sur l'enseignement des MT (2.1.) l'importance de la recherche contextuelle pour les traducteurs. Une des techniques les plus simples utilisées en linguistique de corpus pour explorer l'usage des mots consiste à produire des *concordances*, c'est-à-dire l'ensemble des occurrences d'un mot donné dans son contexte d'apparition dans le corpus. Le format le plus répandu pour les concordances est le format KWIC (Key-Word in Context) et les concordances sont produites à l'aide d'un outil informatisé appelé *concordancier*. Un concordancier monolingue exploite un corpus dans une langue donnée alors qu'un concordancier bilingue repose sur l'utilisation d'un corpus parallèle. Dans cette première étape du cours, et dans un souci de progression graduelle, nous nous concentrons sur les concordanciers monolingues afin de familiariser les apprenants avec les techniques les plus simples utilisées pour exploiter des corpus.
- 26 Nous montrons aux étudiants deux types de concordanciers :
- des concordanciers interrogeables en ligne qui fonctionnent avec des corpus déterminés (c'est le cas du *British National Corpus*, consultable à l'adresse [ht](#)

[p://www.natcorp.ox.ac.uk/](http://www.natcorp.ox.ac.uk/) et du *Corpus of Contemporary American English* – <http://www.americancorpus.org/>), et

- des concordanciers téléchargeables en ligne qui fonctionnent avec des corpus que l'utilisateur choisit lui-même (concordancier *AntConc* – <http://www.antlab.sci.waseda.ac.jp/> – ou *MonoConc* – <http://www.athel.com/mono.html> entre autres).

27 Avant d'exploiter les corpus pour la traduction, les apprentis traducteurs s'initient à la manipulation d'un concordancier interrogeable en ligne avec le *British National Corpus*. Ce corpus représentatif de l'anglais britannique a été constitué entre 1991 et 1996 et comprend 100 millions de mots. Le but est d'amener les apprentis traducteurs à faire appel à leur intuition pour définir les différents sens et structures syntaxiques associés à un mot donné puis de confronter leurs résultats aux données présentes dans le corpus en observant les lignes de concordances. Ce tout premier exercice vise à sensibiliser les apprenants au rôle que joue les corpus dans la vérification d'une hypothèse de langue.

## 4.3. Exploitation des corpus pour la traduction

### 4.3.1. Utilisation d'un corpus comparable

28 À l'aide du concordancier *AntConc*, les apprenants exploitent un corpus comparable, ce type de corpus étant particulièrement utile pour identifier la terminologie et la phraséologie spécifiques à un domaine. Le corpus de travail comporte des textes en langue anglaise (variété britannique, américaine et canadienne) et en langue française (variété française et canadienne) portant sur le domaine de la volcanologie. Ce corpus comparable<sup>2</sup> comporte environ 400 000 mots par langue, soit un total d'environ 800 000 mots. La période couvre une vingtaine d'années (de la fin des années 1970 jusqu'aux années 2000). Les textes inclus sont sous forme écrite, entiers, émanant de sources fiables et que l'on peut qualifier de manière globale comme étant des « textes de vulgarisation », même si l'appellation recouvre en réalité plusieurs types de discours (du discours du quotidien généraliste au discours de semi-vulgarisation, en passant par le discours pédagogique).

- 29 À partir de l'utilisation de ce corpus, et à travers des activités pédagogiques ciblées, les étudiants vont découvrir en quoi les corpus sont utiles pour la traduction. Ces activités sont en lien avec la traduction d'un texte soumis aux étudiants (en l'occurrence un extrait de corpus) qui comporte un certain nombre de difficultés liées à la recherche terminologique et phraséologique.
- 30 La première activité pédagogique proposée aux apprenants consiste à vérifier une intuition, en d'autres termes à tester une hypothèse de traduction comme le font les traducteurs en utilisant un dictionnaire pour confirmer (ou infirmer) leur intuition. Par exemple, les apprenants font une hypothèse de traduction pour les termes *erupting volcano* et *active volcano*. Dans ce premier exercice, les connaissances que les étudiants possèdent de la langue anglaise et de la langue française (non-correspondance des structures en français et en anglais) les aident généralement à formuler une hypothèse de traduction. C'est le cas notamment pour *erupting volcano* étant donné que la structure syntaxique de cette collocation (adjectif + nom) n'a pas la même équivalence en français (*volcan en éruption*, nom + préposition + nom). Concernant le terme *active volcano*, les étudiants privilégient généralement comme première hypothèse *volcan actif* (120 occurrences dans le corpus), mais, encouragés à faire appel à leur propres connaissances linguistiques (une structure anglaise composé d'un adjectif et d'un nom peut se traduire en français par une structure nom + préposition + nom), ils approfondissent la recherche dans le corpus et testent alors comme hypothèse *volcan en activité* ; l'usage vient confirmer leur hypothèse de traduction (21 occurrences).
- 31 Comme nous l'avons déjà mentionné en citant l'étude de Lynne Bowker (1998), le recours au corpus améliore la qualité des traductions par rapport à la seule utilisation de dictionnaires, notamment au niveau du choix des termes et de l'utilisation d'expressions idiomatiques. C'est ce que les étudiants vont tester en confrontant les données du corpus à celles présentes dans des dictionnaires.
- 32 Une première recherche porte sur le terme *shield volcano* et consiste à comparer les données issues du (<http://www.granddictionnaire.com>) *Grand Dictionnaire Terminologique* à celles présentes dans le corpus. Après avoir relevé les équivalents de traduction proposés par le *Grand Dictionnaire Terminologique* (GDT), les apprenants pro-

duisent des concordances pour chaque équivalent afin de vérifier si l'ensemble des traductions proposées par le GDT est attesté dans la partie française du corpus de volcanologie. A l'issue des recherches, un constat s'impose : les termes *volcan en bouclier* et *volcan surbaissé* proposés par le GDT ne font pas partie du corpus et ne reflètent pas l'usage.

- 33 Une deuxième recherche implique la comparaison d'informations présentes dans le corpus avec celles recensées dans des dictionnaires bilingues généraux, qui, comme nous l'avons indiqué en mentionnant le sondage effectué par Amélie-Josselin Leray (2005), constituent une ressource largement utilisée par les traducteurs.
- 34 La recherche porte sur le terme *dormant volcano*. Les étudiants cherchent à vérifier les équivalents proposés par les deux dictionnaires bilingues suivants : *Dictionnaire Oxford-Hachette*<sup>3</sup> (OXHA) et *Harrap's Shorter*<sup>4</sup> (HAR). Ces équivalents de traduction sont les suivants : *volcan dormant*, *volcan en repos* et *volcan au repos*.
- 35 Les apprenants, amenés à commenter ces équivalents, s'interrogent sur l'équivalent de traduction *volcan dormant* et font l'hypothèse qu'il s'agit de l'équivalent le moins attesté par l'usage (calque de l'anglais). Le corpus montre que les deux premières traductions proposées par les dictionnaires analysés ne sont pas attestées par l'usage. L'intuition des étudiants a notamment été confirmée concernant le terme *volcan dormant*, absent du corpus. Ainsi, le croisement des données du corpus avec celles de dictionnaires bilingues permet de découvrir que certains équivalents fournis par ces dictionnaires ne font pas partie de l'usage. En ce qui concerne le terme *volcan au repos*, une seule occurrence apparaît dans le corpus ce qui ne permet pas de le choisir comme équivalent de traduction. Les apprenants approfondissent l'analyse et utilisent les fonctions avancées du concordancier qui permettent d'effectuer une recherche contextuelle sur l'association de deux mots afin d'observer comment ils se combinent dans l'usage. Les apprenants sont donc amenés à s'interroger sur l'association possible des deux mots *volcan + repos* au sein d'expressions. Leur recherche fournit les données suivantes :

alternant avec des **phases de repos** qui peuvent être très longues. Le **volcan** est alors dit éteint.

le **volcan**, qui ne connaît que des **phases de repos** rarissimes, entre en éruption.

tenu pour toujours actif un **volcan** dont l'actuelle **période de repos** n'est pas considérablement plus étendue

le **volcan** napolitain se trouve dans une **phase de repos** qui n'augure rien de bon pour le futur.

La période de repos actuelle des volcans de la chaîne des Puys n'est pas singulièrement...

- 36 Ces contextes semblent indiquer comme équivalents de *dormant volcan*, *volcan en phase de repos* ou *volcan en période de repos*. L'utilisation combinée d'un corpus comparable et d'un concordancier a permis de mettre au jour ces équivalents de traduction non recensés dans les dictionnaires analysés, et cet exercice montre aux apprentis traducteurs l'apport des corpus dans l'enrichissement des dictionnaires bilingues.
- 37 Dans l'exercice suivant, les apprenants vont explorer le corpus pour identifier des contextes informatifs susceptibles de faire émerger d'autres équivalents de traduction pour *dormant volcano*. En produisant des concordances pour l'adjectif *dormant*, ils découvrent qu'un volcan peut être *dormant* ou *active* comme le montre le contexte suivant : *Volcanoes are classified as extinct, dormant or active*. Ils utilisent ensuite cette information dans le corpus français et en produisant des concordances pour *actif*, ils obtiennent des contextes tels que : (...) *distinguer un volcan actif d'un volcan éteint car un volcan actif peut être soit **en activité** soit **en sommeil**. La difficulté réside à distinguer un volcan **en sommeil** (...) et susceptible de se réveiller d'un volcan éteint*. Les apprenants mettent donc en évidence qu'un volcan est dit *en activité* ou *en sommeil*, et non pas *dormant* comme lexicalisé dans les dictionnaires OXHA et HAR. Cette information permet aux étudiants de faire de nouvelles hypothèses de traduction en travaillant sur le champ sémantique lié au *sommeil*, comme par exemple *volcan assoupi* ou *volcan endormi*. Dans le tableau 1 figurent les équivalents de traduction mis au jour par les apprenants pour *dormant volcano*.

**Tableau 1. Équivalents de traduction fournis par le corpus pour *dormant volcano***

Terme anglais	Fréquence dans le corpus	Structure	Terme français	Fréquence dans le corpus	Structure
Dormant <i>volcano</i>	16	Adj. + Nom	<i>Volcan endormi</i>	8	Nom + Adj.
			<i>Volcan en sommeil</i>	4	Nom + Prep. + Nom
			<i>Volcan assoupi</i>	3	Nom + Adj.

38 Enfin, dans ce dernier exercice sur l'utilisation d'un corpus comparable, nous sensibilisons les étudiants à l'apport des corpus pour la traduction à travers l'analyse d'un verbe central dans le domaine de la volcanologie, le verbe *erupt* qui a fait l'objet d'une analyse détaillée chez Frérot et Josselin-Leray (2008). Les apprenants analysent le fonctionnement lexical et syntaxique de ce verbe en corpus et le comparent aux descriptions des deux dictionnaires OXHA et HAR. Cette étude implique notamment de recenser les équivalents de traduction proposés par les dictionnaires ainsi que les structures lexico-syntaxiques associées au verbe. L'observation des occurrences de ce verbe dans la partie anglaise du corpus révèle un fonctionnement bien plus complexe que celui décrit dans les dictionnaires analysés, comme l'illustrent les données du tableau 2.

**Tableau 2. Entrées du HAR et du OXHA pour le verbe *erupt***

Entrée <i>erupt</i> du HAR	Entrée <i>erupt</i> du OXHA
(a) (of volcano) entrer en éruption (b) (of violence, anger, sound) éclater ; (of person) exploser <i>the stadium erupted in a huge roar</i> le stade a éclaté d'un énorme rugissement (c) (of teeth) percer ; (of spot) <i>sortir his face erupted in spots</i> son visage s'est couvert de boutons	1. [volcano] entrer en éruption ; 2. figurative [war, violence, gunfire] éclater ; [laughter, cry] jaillir, éclater ; [person] (with anger) éclater (with de) ; 3. medicine [rash] apparaître ; dentistry [tooth] percer.

39 Il ressort notamment de l'analyse du corpus que le verbe *erupt* est utilisé de manière transitive, emploi qui n'est mentionné dans aucun des deux dictionnaires bilingues analysés (exemple issu du corpus : *Volcanoes on other planets, and their moons, do not necessarily erupt the same materials as Earth's volcanoes*). Par ailleurs, les apprentis

traducteurs identifient dans le corpus d'autres noms en position sujet que ceux recensés par les dictionnaires (*volcan*) comme par exemple *molten rock*, *ash*, *magma* ou bien encore *lava*. Ces noms rentrent dans une catégorie que nous pourrions appeler « divers » représentant, de manière générale, tout ce qui peut être éjecté d'un volcan. Nous fournissons ici un exemple extrait du corpus : « In July 1986, just as it appeared that Puu Oo would start the 48th episode, **magma** began to **erupt** from **new fissures** at the base of the cone ».

- 40 L'équivalent français du verbe *erupt* ne peut pas être *entrer en éruption* dans cet exemple ; or *entrer en éruption* est le seul équivalent de traduction proposé par les dictionnaires. Les apprenants utilisent alors les fonctionnalités de recherche contextuelle dans *AntConc* afin d'observer des associations de mots en corpus pour *magma* + *fissure* et *lave* + *fissure* et tenter ainsi de mettre en évidence de nouveaux équivalents de traduction. C'est ainsi qu'ils identifient les contextes suivants : *du magma s'épanche de cette fissure* ; *de la lave rouge-orange sort d'une fissure* ; *les coulées proviennent d'une fissure*. Leur recherche en corpus permet donc de mettre au jour de nouveaux équivalents de traduction pour *erupt* : *provenir de, sortir, s'épancher*.

### 4.3.2. Utilisation d'un corpus parallèle

- 41 Nous venons de voir, à travers l'élaboration d'activités pédagogiques ciblées, comment un corpus comparable peut être utilisé dans un cours de traduction pour sensibiliser les étudiants à l'apport des corpus pour améliorer la qualité des traductions (respect des usages, précision terminologique) et leur faire mener une réflexion sur la langue tout en les initiant à une méthodologie de recherche en corpus.
- 42 À côté des corpus comparables, les corpus parallèles présentent un grand intérêt pour la traduction comme l'ont démontré notamment Bowker et Pearson (2002). Les outils utilisés pour explorer ce type de corpus sont des concordanciers bilingues (un concordancier bilingue fonctionne avec un corpus parallèle qui est aligné afin de permettre une recherche « parallèle »). Les MT (cf. *supra* 2.) fonctionnent également avec des corpus parallèles. Il existe par ailleurs des outils d'aide à la traduction qui associent aux fonctionnalités classiques des MT des possibilités puissantes de recherche en corpus. C'est à ce



type d'outils que nous allons nous intéresser à travers le logiciel *Multitrans*.

- 43 *Multitrans* est un outil commercialisé (<http://www.multicorpora.ca/>) utilisé dans un nombre croissant de services de traduction (Bureau de la Traduction au Canada, ministères du gouvernement du Canada mais aussi organisations internationales telles que l'UNESCO, Interpol et des sociétés comme Ford, Sobeys, Kraft ou bien encore Pfizer).
- 44 Ce logiciel fait partie des outils de dernière génération disponibles sur le marché dont un des principes forts de conception repose sur l'accès aux textes entiers du corpus aligné et la possibilité d'y mener des recherches tout en traduisant. Etant donné l'apport des corpus pour la traduction, ce point constitue une motivation forte pour intégrer cet outil dans notre enseignement. Par ailleurs, les apprentis traducteurs, à cette étape de l'enseignement, sont suffisamment familiarisés avec les corpus pour se confronter à un outil tel que *Multitrans* et en appréhender l'utilité. Enfin, comme nous l'avons déjà mentionné (2.2.), des enquêtes réalisées auprès de traducteurs professionnels montrent leur vif intérêt pour l'utilisation de logiciels de constitution et d'exploration de corpus ainsi que d'extraction terminologique à partir de corpus portant sur des domaines spécifiques, ce que permet précisément *Multitrans*. Ainsi, le choix de cet outil est motivé par :
- l'apport des corpus pour la traduction,
  - l'adéquation aux besoins et intérêts des traducteurs,
  - la création rapide de corpus alignés,
  - l'accès aux textes entiers dans les deux langues,
  - la navigation dans le corpus (accès à l'intégralité du corpus), et,
  - l'extraction automatique de terminologie et la gestion de terminologie.
- 45 Les étudiants apprennent à utiliser *Multitrans* pour constituer un corpus aligné, l'enrichir, et faire des recherches terminologiques et phraséologiques à l'aide du concordancier bilingue qui permet de naviguer dans le corpus parallèle. Ils s'initient également à la création de fiches terminologiques avec le gestionnaire de terminologie intégré au logiciel.

## 5. Conclusion

- 46 L'élaboration d'activités pédagogiques ciblées à partir de corpus, en particulier lorsqu'elle s'appuie sur une comparaison avec des ressources plus traditionnelles telles que les dictionnaires bilingues, peut être très utile pour sensibiliser les apprentis traducteurs à l'apport que représentent les corpus pour améliorer la qualité des traductions (respect des usages, précision terminologique).
- 47 Par ailleurs, la prise en compte dans la conception de cours de traduction des besoins exprimés par les traducteurs professionnels vis-à-vis des outils informatisés d'aide à la traduction est déterminante pour former des traducteurs dont la compétence sera reconnue sur le marché du travail. Ainsi, l'intégration dans la formation universitaire des traducteurs d'outils d'aide à la traduction telles que les mémoires de traduction, mais, comme nous l'avons plus particulièrement montré, d'outils permettant de construire, d'enrichir, et de naviguer dans un corpus parallèle s'avère essentielle pour répondre aux exigences du métier.
- 48 Du point de vue de la performance des outils d'aide à la traduction, il semblerait pertinent de pouvoir évaluer avec précision quel est l'apport des outils de recherche sur corpus tels que les concordanciers par rapport aux mémoires de traduction qui restent, tout au moins dans le contexte universitaire français, bien plus souvent intégrées aux formations de traducteurs.

## BIBLIOGRAPHY

---

Arrouart Catherine, 2003, « Les mémoires de traduction et la formation universitaire : quelques pistes de réflexion », dans *Meta*, vol. 48, n° 3, Les Presses de l'Université de Montréal.

Bédard Claude, 2000, « Mémoire de traduction cherche traducteur de phrases... », dans *Traduire*, n° 186.

Bernardini Silvia, 2006, « Corpora for translators and translation practices. Achievements and challenges », dans *Proceedings of LREC (Language Resources and Evaluation Conference)*, p. 17-22.

Bowker Lynne, 1998, « Using Specialized Monolingual Native-Language Corpora as a Translation Resource: a Pilot

Study », dans *Meta*, vol. 43, n° 4, p. 631-651.

Bowker Lynne, 2002, *Computer-Aided Translation Technology : A Practical Introduction*, University of Ottawa, Ottawa Press, Didactic of Translation Series.

Bowker Lynne, Barlow Mickael, 2008, « A comparative evaluation of bilingual concordancers and translation memory systems », dans *Topics for Language Resources in Translation and Localisation*, Yuste Rodrigo, Elia (ed.), p. 1-22.

FRÉROT Cécile, JOSSELIN-LERAY Amélie, 2008, « Contribution des corpus à l'enrichissement des dictionnaires bilingues généraux. Application au domaine de la volcanologie », dans *Autour des langues et du langage. Perspective interdisciplinaire*, Presses Universitaires de Grenoble, p. 415-422.

Gouadec Daniel, 2006, « Profils de compétences et formation universitaire », Colloque de l'université Rennes 2 : *Quelle qualification universitaire pour les traducteurs ?*

Gow Francie, 2003, *Metrics for Evaluating Translation Memory Software*, Thesis under the supervision of L. Bowker, School of Translation and Interpretation, University of Ottawa.

Granger Sylviane, 2003, « The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies », dans Granger S., Lerot J. et Petch-Tyson S. (ed.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*, Amsterdam & Atlanta, Rodopi, p. 17-29.

Josselin-leray Amélie, 2005, *Place et rôle des terminologies dans les dictionnaires généraux unilingues et bilingues*.

*Étude d'un domaine de spécialité : volcanologie*, thèse de doctorat, Université Lyon II.

Laviosa Sara, 2002, *Corpus-Based Translation Studies: Theory, Findings, Applications*, Amsterdam, New York, Rodopi.

MeLLANGE, 2006, *Corpora & E-learning Survey Results*, en ligne [http://www.iti.org.uk/uploadedFiles/survey\\_s/Mellange\\_Survey\\_Updated\\_\(05-06\).pdf](http://www.iti.org.uk/uploadedFiles/survey_s/Mellange_Survey_Updated_(05-06).pdf) consulté le 20 novembre 2008.

Pearson Jennifer, 2003, « Using Parallel Texts in Translation Training Environment », dans Zanettin Federico, Bernardini Silvia et Stewart Dominic (eds), *Corpora in translator education*, Manchester, St Jerome, p. 15-24.

Pearson Jennifer, Bowker Lynne, 2002, *Working with Specialized Language. A practical guide to using corpora*, Routledge.

Sauron Véronique, 2007, « Les nouvelles technologies dans l'enseignement de la traduction : l'exemple de la traduction juridique », dans Lavault Elisabeth. (éd.), *Traduction spécialisée : pratiques, théories, formations*, Peter Lang, Berne, p. 207-224.

Williams Ian A, 1996, « A translator's reference needs: dictionaries or parallel texts », dans *Target*, n° 8, p. 277-299.

Zanettin Federico, 1998, « Bilingual Comparable Corpora and the Training of Translators », in *Meta*, vol. 43, n° 4, p. 613-630.

Zanettin Federico, 2001, « Swimming in Words : Corpora, Translation and Language Training », in Aston Guy, (éd.), *Learning with Corpora*, Athelstan, p. 177-197.

Zanettin Federico, Bernardini Silvia, *translator education*, Manchester, St Jerome, 2003, *Corpora in* rome.

## NOTES

---

- 1 Voir plus particulièrement le chapitre 11, p. 193-210 « Using LSP as a translation resource ».
- 2 Nous remercions vivement Amélie Josselin-Leray de nous permettre d'utiliser ce corpus qui a été constitué dans le cadre de sa thèse. Josselin-Leray, Amélie (2005), *Place et rôle des terminologies dans les dictionnaires généraux unilingues et bilingues. Etude d'un domaine de spécialité : volcanologie*, thèse de doctorat, Université Lyon II.
- 3 *The Oxford-Hachette French Dictionary/Le Dictionnaire Hachette-Oxford* (1996). Hachette Livre and Oxford University Press. Version Cd-Rom.
- 4 *Harrap's Shorter. Dictionnaire anglais-français / français-anglais* (2000). Chambers Harrap Publishers. Version Cd-Rom.

## ABSTRACTS

---

### Français

Les outils informatiques d'aide à la traduction font de plus en plus partie de l'environnement professionnel du traducteur. Une initiation à ces outils est donc devenue incontournable pour toute formation qui se donne comme objectif de former de futurs traducteurs.

À côté des mémoires de traduction, dont l'enseignement tend aujourd'hui à se banaliser dans les formations de traducteurs, des outils issus de la linguistique de corpus méritent qu'on leur accorde une place, d'autant que les outils d'aide à la traduction visant à constituer, enrichir et exploiter facilement des corpus sont largement plébiscités par les traducteurs. Ainsi, à travers des activités pédagogiques ciblées, nous montrons comment nous pouvons sensibiliser les apprentis traducteurs à l'apport que représentent les corpus pour la traduction ; nous insistons sur la nécessité d'intégrer dans les cursus des outils de toute dernière génération qui offrent des possibilités puissantes de recherche en corpus.

### English

Over the past decade, the use of computer-aided translation tools has grown increasingly attractive among professional translators and training students with those tools has become an essential part in translator training.

While it is now well-established that translation memories should be part of any translator training, other tools such as corpora and concordancers have proved to be of considerable value in a great number of translation studies. Besides, a European survey shows that professional translators are very interested in corpus-based tools that make it possible to easily create and enrich corpora as well as search through them. In this paper, we show how corpus-based translation-related tasks can help teachers raise trainee awareness of corpora as a translation resource. We focus also on leading-edge computer-aided translation tools and show how they can be integrated in the classroom in order to provide trainees with tools that best meet the professional translators' needs.

## AUTHOR

---

### **Cécile Frerot**

Maître de conférences au département d'anglais (LEA) de l'université Stendhal où elle assure notamment des enseignements en traduction, terminologie et TAO.

Traductrice et terminologue de formation, spécialisée dans le traitement automatique des langues, elle s'intéresse principalement à l'utilisation d'outils informatisés dans la formation des traducteurs et à l'exploitation des corpus pour l'enseignement de la traduction.

IDREF : <https://www.idref.fr/095113665>

ORCID : <http://orcid.org/0000-0001-5106-3536>

HAL : <https://cv.archives-ouvertes.fr/cecile-frerot>

ISNI : <http://www.isni.org/0000000359149195>