



**HAL**  
open science

# Apport de la traduction dans l'analyse des marqueurs de relations conceptuelles. Une étude en corpus aligné français-italien

Anne Condamines, Marie-Pierre Escoubas-Benveniste, Silvia Federzoni

## ► To cite this version:

Anne Condamines, Marie-Pierre Escoubas-Benveniste, Silvia Federzoni. Apport de la traduction dans l'analyse des marqueurs de relations conceptuelles. Une étude en corpus aligné français-italien. Cécile Frérot; Mojca Pecman. Des corpus numériques à l'analyse linguistique en langues de spécialité, UGA Editions, pp.313-335, 2021, Langues gestes paroles, 978-2-37747-261-1. halshs-03721260

**HAL Id: halshs-03721260**

**<https://shs.hal.science/halshs-03721260>**

Submitted on 12 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Chapitre 12 : Apport de la traduction dans l'analyse des marqueurs de relations conceptuelles. Une étude en corpus aligné français-italien

Anne Condamines<sup>1</sup>, Marie-Pierre Escoubas-Benveniste<sup>2</sup> et Silvia Federzoni<sup>1</sup>

<sup>1</sup>Cognition, Langues, Langage, Ergonomie (CLLE) CNRS et Université Toulouse Jean Jaurès,

<sup>2</sup>Metodi e Modelli per l'Economia il Territorio e la Finanza (MEMOTEF), Université de Rome, La Sapienza

**Résumé :** Ce chapitre s'intéresse aux possibilités d'utiliser un corpus spécialisé aligné dans la perspective de construire un réseau conceptuel unique pour les deux langues concernées, en l'occurrence le français et l'italien. Il s'inspire des travaux en terminologie textuelle et en traductologie qui ont montré que le processus de traduction conduisait parfois à des modifications du point de vue informationnel. Il se focalise sur les marqueurs de relation conceptuelle et leur traduction, plus précisément, dans cette étude, sur les marqueurs de la relation d'hyponymie. L'étude est réalisée sur un corpus journalistique spécialisé, *Le Monde diplomatique*, et sa traduction en italien, les deux corpus comportant environ 500 000 mots chacun. Nous justifions le choix d'un corpus de presse en lien avec l'objectif de l'étude. Une liste de 39 marqueurs d'hyponymie a été constituée pour chaque langue. Grâce au logiciel TXM, chacun de ces marqueurs est recherché dans le corpus correspondant à la langue concernée et, grâce à l'alignement, son équivalent dans l'autre langue est récupéré. L'analyse des résultats obtenus permet de repérer des phénomènes intéressants (apparition de nouveaux marqueurs, ajout d'hyponymes, entre autres) qui pourraient contribuer à la construction d'un réseau terminologique. Il s'avère ainsi que le processus de traduction, qui conduit à des modifications dans l'information véhiculée, pourrait aider à enrichir le réseau conceptuel qui sous-tend les deux langues.

**Mots-clés :** corpus aligné, corpus journalistique, marqueurs de relation, relations conceptuelles, terminologie textuelle, traductologie

## 1 Introduction

Dans ce nouveau chapitre qui prolonge les contributions précédentes dédiées aux analyses syntaxiques et sémantiques des textes et des discours spécialisés, nous cherchons à examiner l'apport de l'utilisation de corpus alignés en terminologie, et plus particulièrement dans la construction de réseaux terminologiques. En traductologie, des phénomènes de modification impliquant les éléments traduits (ajout ou suppression d'éléments) ont été repérés, par exemple par Baker ([BAKER, 1996](#)). Néanmoins, seule Pearson ([PEARSON, 2000](#)) a évoqué la possibilité d'utiliser ces modifications au moment de la traduction pour construire un réseau relationnel en se focalisant sur les marqueurs de relations conceptuelles. Rappelons que les marqueurs de relations conceptuelles concernent l'ensemble des éléments langagiers voire typodispositionnels qui peuvent être utilisés pour construire une relation conceptuelle.

Nous avons travaillé sur un corpus aligné franco-italien (*Le Monde diplomatique* traduit en italien) et nous nous sommes focalisées sur la relation d'hyponymie. À l'aide des logiciels TXM ([HEIDEN et coll., 2010](#)) et Alinea ([KRAIF, 2001](#)), nous avons étudié, dans les deux corpus, l'ensemble des contextes contenant un marqueur d'hyponymie et leur traduction

dans l'autre langue. Une fois tous les contextes extraits, après avoir réduit le plus possible les occurrences non pertinentes obtenues, nous avons examiné chacun des couples marqueur/traduction afin de caractériser l'ensemble des phénomènes apparus au moment de la traduction. Cette étude a donc nécessité que nous disposions de listes de marqueurs d'hyponymie en français et en italien.

Dans la deuxième section, nous faisons état des travaux existant en terminologie textuelle et en traductologie afin de situer notre apport au carrefour de ces deux disciplines. Dans la troisième section, nous présentons la mise en place de l'étude : choix d'un corpus de presse dans le cadre d'une étude terminologique, constitution des listes de marqueurs, mise en œuvre de ces marqueurs grâce à TXM. Nous rendons compte ensuite des résultats que nous avons obtenus (section 4).

## 2 Contexte de l'étude, travaux existants

### 2.1 Travaux en terminologie textuelle

La terminologie textuelle s'est développée depuis le début des années 1990. Pour le dire brièvement, il s'agissait de construire des réseaux terminologiques à partir de corpus considérés comme réservoirs de connaissances (CONDAMINES & REBEYROLLE, 2001 ; PEARSON, 1998) pour constituer des bases de connaissances terminologiques (CONDAMINES, 2018). Divers paramètres ont rendu possible l'émergence de ce courant :

- le développement de la linguistique de corpus et la volonté de rapprocher les études en terminologie des études en lexicologie (SLODZIAN, 2000) ;
- le développement d'outils de Traitement Automatique de la Langue et en particulier ceux dédiés à l'analyse terminologique (BOURIGAULT et coll., 2001) ;
- le rapprochement avec l'ingénierie des connaissances qui utilise elle aussi les corpus spécialisés comme réservoirs de connaissance (SKUCE & MEYER, 1991).

Dans cette approche, les corpus sont donc utilisés pour repérer des termes et des relations entre les termes (générique/spécifique, méronymie, cause, pour ne citer que les plus fréquentes). Depuis lors, des travaux ont été menés pour essayer de trouver des moyens aussi systématiques (et automatisables) que possible pour repérer les termes et les relations conceptuelles. Pour les relations conceptuelles, des travaux en lexicologie sur corpus avaient déjà été développés et avaient montré que certains éléments linguistiques (lexicaux, syntaxiques, morphologiques voire typographiques) pouvaient être considérés comme des marqueurs de relation conceptuelle (FLOWERDEW, 1992 ; SINCLAIR, 1991). Par exemple [de tous les X, Y...] peut être considéré comme un marqueur d'hyponymie (lien entre X et Y) :

[de tous les X, dét Y...]

*de tous les oiseaux, le colibri est le plus petit*

Si l'ingénierie des connaissances a évolué vers l'utilisation de méthodes tout à fait différentes pour repérer les relations – méthodes dites de *deep learning*, basées sur des approches de sémantique distributionnelle (LENCL, 2008) – les linguistes-terminologues ont continué à travailler sur la notion de marqueurs de relation (AUGER & BARRIÈRE, 2008). Un aspect a été en particulier examiné, celui de la variation, qui concerne la présence *vs* absence de certains marqueurs dans certains corpus mais aussi l'interprétation de ces marqueurs dont on sait que beaucoup sont polysémiques. C'est alors le rôle du domaine et celui du genre textuel qui ont été examinés (CONDAMINES, 2002 ; MARSHMAN & L'HOMME, 2008).

Il existe un autre type de variation qui, à notre connaissance, n'est évoqué que dans l'étude de Pearson (2000) : il s'agit de la variation qui se produit au moment de la traduction. Le

processus de traduction entraîne parfois l'ajout, la suppression ou la modification d'éléments (voir section 4), constituant des phénomènes qui pourraient être précieux pour le linguiste-terminologue. C'est précisément l'intérêt de travailler sur la traduction des marqueurs de relation, en utilisant un corpus aligné, que nous examinons dans ce chapitre.

## 2.2 Travaux en traductologie

Cette section aborde les travaux en traductologie, tout particulièrement ceux qui se sont intéressés aux phénomènes de variation liés au processus de la traduction.

### 2.2.1 Le texte traduit : un genre communicatif à part entière

L'idée que la traduction est une « forme de communication unique » (BAKER, 1998) fait du texte traduit, en tant que résultat d'un processus langagier complexe, le fertile terrain d'études scientifiques qui vont pouvoir s'appuyer sur un très grand nombre de données attestées dans différentes langues utilisées pour traduire. Dans le domaine particulier de la traduction « pragmatique » ou *traduzione specializzata* (DELISLE, 1980 ; SCARPA, 2001) auquel nous nous intéressons ici (voir section 2.1), la traduction est une communication particulièrement contrainte : le texte traduit doit assurer la même fonction communicative, c'est-à-dire les mêmes effets que le texte source et, hormis pour les textes scientifiques et techniques, souvent au sein d'un environnement socioculturel distinct de celui du texte original. Le texte traduit résulte donc d'une série de négociations entre le traducteur et les multiples contraintes auxquelles il est soumis.

Tout d'abord dans sa pratique, le traducteur est très fortement tributaire du texte original et de la langue source pour les connaissances et les contenus encyclopédiques en général. Il doit également tenir compte de contraintes concernant la forme et la fonction des textes dans la langue et la culture sources : le genre textuel tout comme le ou les domaines de spécialisation impliqués et les pratiques sociales qui y sont liées, vont imposer leurs propres nécessités plus ou moins normatives en ce qui concerne le traitement des connaissances spécialisées dans la langue-cible. Ainsi, dans la traduction pragmatique, non seulement deux langues se confrontent, mais aussi et surtout deux pratiques terminologiques qui sont socialement ancrées et caractérisées localement dans les langues et les usages sociaux en présence.

Enfin, le traducteur doit tenir compte du rapport que les lecteurs présumés du texte traduit entretiennent avec cette connaissance spécialisée. Il doit donc anticiper les besoins informatifs des destinataires et adapter son action langagière en conséquence. C'est ainsi que, dans la traduction en contexte de communication asymétrique – quand les récepteurs du texte-cible n'ont pas le même degré d'expertise dans le domaine de connaissance que l'émetteur du texte-source – le traducteur peut faire preuve d'initiative voire de créativité (BECIRI, 2008 ; voir section 4.5) dans le but d'optimiser la réception de son texte par ses lecteurs.

### 2.2.2 Corpus, traductologie et analyse contrastive

À partir des années 1990, les méthodes de la linguistique de corpus sont appliquées à de grands ensembles de textes traduits et la traductologie connaît une avancée théorique majeure (BAKER, 1996). Par traductologie, il faut entendre ici les études à orientation descriptive plutôt que prescriptive de la traduction (HOLMES, 1988 ; TOURY, 1995), celle-ci signifiant aussi bien la pratique langagière spécifique qui correspond à un « comportement d'humain traduisant » (*translational behavior*), que son résultat, le texte traduit. Dans ses travaux pionniers, Baker montre qu'une analyse basée sur corpus de la « langue de la traduction », c'est-à-dire de la langue quand elle est utilisée pour traduire, peut permettre de confirmer statistiquement et de théoriser l'existence de régularités propres aux textes traduits, observées sur différentes

langues à partir des années 1970. L'utilisation de corpus, qu'il s'agisse de corpus bilingues formés de textes originaux et de leur traduction, ou encore de corpus monolingues comparables faits d'un ensemble de textes en langue originale et d'un ensemble de textes de même genre communicationnel, où cette même langue est utilisée pour traduire, a permis de vérifier certains postulats concernant les « lois » de la traduction. Baker (1996) montre comment une analyse basée sur corpus permet de théoriser ces phénomènes réguliers et transversaux, et de poser l'existence de traits typologiques communs à toute langue utilisée pour traduire.

Les objectifs de la traductologie descriptive ou « science de la traduction » se distinguent de ceux de la linguistique contrastive. Toutes deux consistent à comparer des ensembles ou échantillons de langues pour en faire émerger des écarts et des points communs, mais leurs objectifs divergent fondamentalement. L'analyse contrastive vise à décrire les similarités et les différences entre deux langues, ou plus, par le biais d'une comparaison systématique entre les différents systèmes linguistiques, qu'il s'agisse de la langue générale ou de la langue de spécialité (JOHANSSON, 2003, p. 15), tandis que les études traductologiques ont pour objectif d'étudier les traits typologiques de toute langue dès lors qu'elle est utilisée pour traduire. Pourtant, ces deux disciplines sont liées par un objectif commun : « Elles s'intéressent à la façon dont 'la même chose' peut être dite de plusieurs autres manières, même si dans chacun des domaines cette information est utilisée à des fins différentes »<sup>1</sup> (CHESTERMAN, 1998, p. 39, cité par GRANGER, LEROT & PETCH-TYSON, 2003, p. 9). Les corpus bilingues, multilingues, qu'ils soient parallèles ou comparables sont utilisés aussi bien pour des études contrastives que traductologiques. Toutefois, ce n'est qu'assez récemment que ces deux champs de recherche ont commencé à converger grâce en particulier à la diffusion des corpus de textes électroniques qui multiplie les potentialités de l'analyse. Tout comme les traductologues, les linguistes contrastivistes trouvent la possibilité de confronter certaines généralisations, fondées initialement sur l'intuition, à de plus grands nombres de textes<sup>2</sup>. Les méthodes basées sur corpus sont partiellement automatisées et se fondent sur le repérage de formes. Il est donc méthodologiquement impossible d'abstraire la recherche des traits typiques universels des genres textuels traduits de leurs formes proprement linguistiques, et l'approche contrastive reste à l'ordre du jour, qu'il s'agisse de comparaisons de systèmes linguistiques actualisés dans les textes ou de comparaison de types d'usage langagier (langage original vs langage traduisant). La description des spécificités de la langue de la traduction tire aussi bien parti de corpus monolingues comparables pour le genre textuel et le domaine et se distinguant par la fonction de la langue (cf. anglais du tourisme langue originale / langue de la traduction, (GANDIN, 2015)), que des résultats obtenus par des études contrastives sur deux ou plusieurs langues basées sur des corpus spécialisés parallèles. Citons par exemple Patin (PATIN, 2014) qui rend compte d'interférences morphosyntaxiques entre le français langue originale des textes parlementaires européens et l'espagnol de la traduction dans l'expression de certains termes, et Cuenca (CUENCA, 2003) qui étudie la variation des structures linguistiques marquant la reformulation dans trois langues. Pour le couple français-italien qui fait l'objet de notre recherche, si les études de linguistique contrastive jouissent d'une grande vitalité à partir des années 1980 (DRUETTA, 1996), il faut attendre les années 2000 pour qu'elles s'orientent vers l'utilisation des méthodes de la linguistique de corpus (HEDIARD, 2002, 2007, 2008) en recourant à une exploration du corpus assistée en particulier par des concordanciers mono- et bilingues. Pour le domaine économique, Janot (JANOT, 2014)

---

<sup>1</sup> Notre traduction du texte original : « *They are interested in seeing how 'the same thing' can be said in other ways, although each field uses this information for different ends* ».

<sup>2</sup> « *Contrastive linguists now have a way of testing and quantifying intuition-based contrastive statements in a body of empirical data that is vastly superior – both qualitatively and quantitatively – to type of contrastive data that had hitherto been available to them.* » (GRANGER, LEROT & PETCH-TYSON, 2003, p. 9).

analyse la présence/absence d'un discours métalinguistique d'accompagnement d'un terme à partir d'un corpus parallèle de textes italiens traduits en français. La plupart de ces études linguistiques ont pour finalité l'application à l'apprentissage des langues ou à la formation des traducteurs. Il n'existe pas, à notre connaissance, de travaux d'analyse contrastive sur l'italien et le français concernant l'apport que peut représenter la traduction à l'enrichissement des marqueurs de la relation d'hyperonymie à partir d'un corpus aligné.

### 2.2.3 Les dimensions de la variation du texte traduit

L'analyse des textes traduits basée sur corpus offre la possibilité de repérer les traces formelles de l'intervention traduisante et de documenter des phénomènes récurrents qui sont imputables à la nature communicative de la traduction même. Déjà intuitivement qualifiée de « troisième code » résultant d'une confrontation et d'un compromis entre deux « codes » linguistiques en présence (FRAWLEY, 1984), la langue, quand elle est utilisée pour traduire et non pas dans son usage ordinaire, est donc théorisée comme un *langage du troisième type*, c'est-à-dire comme un *usage particulier de la langue* qui, indépendamment des « codes » linguistiques en présence, possède des traits typologiques spécifiques qui dépendent du type communicatif que constitue la traduction. Sur la base des faits vérifiés quantitativement et qualitativement dans les analyses basées sur corpus, Baker (1996) définit quatre universaux de la traduction, c'est-à-dire quatre traits linguistiques typologiques prédictibles dans les textes traduits :

- l'« explicitation », qui désigne la tendance des traducteurs à exprimer (sous différentes formes) des informations qui sont implicites dans le texte-source<sup>3</sup> ;
- la « simplification », qui indique le fait que les textes traduits tendent à être plus simples que les textes-sources d'un point de vue syntaxique, lexical et stylistique. L'auteur souligne aussi que la modification de la ponctuation par rapport au texte original fait partie des moyens utilisés pour simplifier et rendre plus clair le texte d'arrivée<sup>4</sup> ;
- la « normalisation », qui désigne la tendance des traducteurs à se conformer aux caractéristiques linguistiques du texte de départ. Ce trait spécifique de la traduction fait que l'on retrouve dans les textes-cibles des interférences dues au texte-source. Ces interférences s'expliquent par le fait que, à différents niveaux (lexical, syntaxique, grammatical, typographique), le texte d'arrivée tend à être calqué par le traducteur sur des patrons propres au texte original. Ce trait expliquerait pourquoi les textes traduits sont souvent critiqués pour leurs écarts par rapport à l'usage réel du langage, tel qu'on l'utiliserait si le texte était rédigé directement en langue originale<sup>5</sup> ;
- le « nivellement » ou « *levelling out* » opéré par le processus traduisant désigne la tendance des textes traduits à être tous homogènes, c'est-à-dire par exemple, à présenter une variation lexicale réduite par rapport à celle du texte source<sup>6</sup>.

Il existe donc des caractéristiques communes à tout processus de traduction, et tout texte traduit serait susceptible de varier par rapport à sa source selon l'une des quatre dimensions

---

<sup>3</sup> Notre traduction du texte original : « *a tendency to spell things out rather than leave them implicit in translation.* » (BAKER, 1996, p. 180).

<sup>4</sup> Notre traduction du texte original : « *the tendency to simplify the language used in translation* » (*ibid.*, p. 181).

<sup>5</sup> Notre traduction du texte original : « *the tendency to exaggerate features of the target language and to conform to its typical patterns* » (*ibid.*, p. 183).

<sup>6</sup> Notre traduction du texte original : « *the tendency of translated text to gravitate towards the centre of a continuum* » (*ibid.*, p. 184).

considérées. On peut ainsi reconnaître parmi ces axes de variation que représentent les universaux de la traduction certains des phénomènes observés par Pearson (PEARSON, 2000) lors d'une étude sur les marqueurs de relation en anglais et en français. À partir d'un corpus parallèle bilingue, l'auteur montre qu'au moment de la traduction, les marqueurs de relation sont restitués différemment, parfois simplifiés, parfois adaptés à la langue cible. Ces variations du marquage sont l'une des manifestations linguistiques des universaux de la traduction, il est donc possible d'établir un lien entre le processus de traduction et la variation du marquage de relation. C'est cette possibilité que nous exploitons ici, en essayant de montrer qu'un corpus bilingue de presse spécialisée et ses traductions constitue une ressource exploitable selon une exploration bidirectionnelle pour aider à enrichir le réseau conceptuel qui sous-tend les deux langues.

### 3 Mise en place de l'étude pour l'hyponymie

L'étude que nous présentons concerne l'hyponymie, qui est sans doute la relation la plus présente dans les corpus. D'un point de vue logique, l'hyponymie a pour principale caractéristique de permettre l'héritage des propriétés, ce qui, dans la perspective de la construction d'un réseau est particulièrement économique. Dans cette section, nous présentons le corpus, les marqueurs de relation et la façon dont nous avons mis en œuvre ces marqueurs de façon systématique.

#### 3.1 Le Corpus

Notre corpus d'étude est un corpus parallèle de type journalistique *Le corpus français-italien du Monde diplomatique - Il Manifesto* (DI FIORE & DI VITO, 2008 ; désormais MDIL). Il constitue un échantillon de textes de presse spécialisée, attendu que chacun des articles publiés fait l'objet d'une catégorisation par domaine. Il se compose de six sous-corpus, pour un total de 522 180 mots pour le français et de 496 966 mots pour l'italien (voir Tableau 1). Les six sous-corpus correspondent à une sous-catégorisation des textes en fonction des domaines de l'économie, de l'environnement, de la justice, de la médecine, de la politique et de la société (colonne 1 du Tableau 1 : md-eco, md-env, md-jus, md-med, md-pol, md-soc). Les articles français ont été extraits de la page Web du *Monde Diplomatique* et les articles traduits en italien ont été tirés du *Manifesto*. Les traducteurs qui ont participé à la traduction des articles pris en considération sont au nombre de 12, ce qui nous autorise à supposer que les biais liés à des pratiques individuelles spécifiques de traduction sont réduits, même si le nom du traducteur n'est pas toujours précisé pour chaque article.

Le corpus n'est pas homogène : le sous-corpus md-eco est prédominant. Le corpus md-eco-fr se compose de 250 222 mots et md-eco-it de 241 191 (il représente donc 47,91 % de MDIL-fr et 48,53 % de MDIL-it) ; les autres sous-corpus se répartissent de manière plus homogène, comportant entre 50 000 et 55 000 mots.

Cette différence s'explique par les intervalles temporels pris en considération pour les années de publication des articles. Les textes dont se compose md-eco sont des articles publiés entre 1998 et 2008, alors que les autres sous-corpus comprennent des articles publiés entre 2006 et 2008. On peut supposer que des contraintes externes ont réglé ce choix, mais les auteurs du corpus ne se prononcent pas sur ce point.

TABLEAU 1 – COMPOSITION DU CORPUS DE PRESSE SPECIALISEE

Sous-corpus	Nb mots (MDIL-fr)	Nb mots (MDIL-it)
md-eco	250 222	241 191
md-env	53 517	50 551
md-jus	55 149	52 313
md-med	52 575	49 293
md-pol	55 361	51 282
md-soc	55 356	52 336
Total	522 180	496 966

Le corpus est disponible aux formats texte, XML et HTML. Les différentes versions proposent les textes parallèles subdivisés en paragraphes par les balises <p> et </p>. Cette division vise à faciliter la phase d'alignement des textes. Les fichiers du corpus sont en texte brut, aucune analyse morpho-syntaxique n'a été effectuée.

Pour mener notre recherche, nous avons utilisé le corpus au format XML. Nous avons procédé à l'alignement et à l'étiquetage morpho-syntaxique, afin d'exploiter ces informations pour la construction de nos requêtes en vue du repérage des marqueurs. Pour l'étiquetage morpho-syntaxique, comme nous le montrons dans la section 3.3, nous avons exploité TreeTagger, intégré dans la chaîne de traitement du logiciel TXM<sup>7</sup>.

Nous avons aligné les textes au niveau des paragraphes, à l'aide du logiciel Alinea<sup>8</sup>. Nous avons effectué l'alignement pour chaque sous-corpus de notre corpus. Nous n'avons pas vérifié la qualité de l'alignement manuellement faute de temps. Seul le md-env a été supprimé pour effectuer notre recherche parce que dans 90 % des cas, les segments alignés ne correspondaient pas, ce qui aurait compromis nos analyses quantitatives et rendu trop longue la phase de correction. Pour les autres sous-corpus, le problème était beaucoup plus limité : sur un total de 1 071 segments extraits (total des segments français et italiens), 55 segments ne correspondaient pas, et donc 5,14 % des cas ont été corrigés.

### 3.2 Le choix d'un corpus de presse

Avant de poursuivre la présentation de notre méthodologie, il est nécessaire de justifier le choix d'un corpus de presse – même si les articles exploités portent sur des domaines spécialisés – dans la perspective d'une analyse terminologique.

Nous avons fait le choix d'utiliser un corpus de presse pour différentes raisons. Tout d'abord, obtenir un corpus aligné en domaine spécialisé, avec un volume assez important de données dans les deux langues – surtout si l'une n'est pas l'anglais – n'est pas aisé. La disponibilité du quotidien *Le Monde diplomatique* en français avec sa traduction en italien nous a orientées vers ce corpus.

Toutefois, une raison plus scientifique a aussi étayé notre choix. Les auteurs s'accordent pour considérer que la presse, surtout lorsqu'elle est spécialisée, s'intègre dans une perspective de

<sup>7</sup> Disponible en ligne sur <http://textometrie.ens-lyon.fr> [consulté en décembre 2020].

<sup>8</sup> Disponible en ligne sur [http://turing3.univ-grenoble-alpes.fr/olivier.kraif/index.php?option=com\\_content&task=view&id=27&Itemid=43](http://turing3.univ-grenoble-alpes.fr/olivier.kraif/index.php?option=com_content&task=view&id=27&Itemid=43) [consulté en décembre 2020].



vulgarisation (JACOBI, 2007 ; BEACCO & MOIRAND, 1995). Pour Pearson (1998), la situation dans la presse spécialisée correspond à une communication émanant d'« un expert relatif [du domaine s'adressant] au non-initié<sup>9</sup> ». De la même façon, il y a unanimité pour considérer que la presse spécialisée, relevant de la vulgarisation, est riche en éléments métalinguistiques permettant la reformulation et la définition (JACOBI, 1987 ; BEACCO & MOIRAND, 1995 ; MORTUREUX, 1982). Certains travaux en terminologie ont montré que les marqueurs de relation conceptuelle étaient plus nombreux dans les corpus de presse spécialisée que dans les corpus comportant des textes rédigés par des experts pour des experts (MARSHMAN & L'HOMME, 2008 ; REBEYROLLE, 2000). Pour une étude portant sur la traduction des marqueurs de relation, l'utilisation d'un corpus de presse est donc pertinente.

Cependant, pour la construction d'un réseau conceptuel concernant un domaine spécialisé, un corpus de presse présente une difficulté. On a parfois caractérisé la vulgarisation comme étant une forme de traduction (JACOBI, 1987) même si cette transmission à trois protagonistes – l'expert, le médiateur, le néophyte – peut paraître trop stéréotypée (MOIRAND & TREGUER-FELTEN, 2007). Il s'agit d'une médiation « cognitive » permettant de mettre à la portée d'un lecteur plus ou moins néophyte des concepts qu'il ne connaît pas. Cela n'est pas sans conséquences sur les formes utilisées pour adapter un savoir technique ou scientifique. Ainsi, si l'on considère que la première opération définitoire consiste à rattacher un terme à un autre, superordonné (un hyperonyme), il est prévisible que le terme générique choisi pour définir se situera plus « haut » dans la hiérarchie des concepts qu'il ne serait dans un contexte hautement spécialisé. Ainsi *Le pangolin est un mammifère originaire d'Amérique et d'Asie* introduit *mammifère* comme hyperonyme direct de *pangolin* alors que dans *Le pangolin appartient aux manidés, une famille de mammifères pholidotes*, il y a deux hyperonymes intermédiaires (manidés et mammifères pholidotes) entre *pangolin* et *mammifère*. En réalité, tout dépend de la représentation que se fait le journaliste de la connaissance de ses lecteurs. S'il estime que les lecteurs sont semi-experts, il utilisera moins de contextes définitoires et les hyperonymes seront déjà très spécialisés. Quoi qu'il en soit, on peut penser que si l'on construit une base de connaissance terminologique à partir d'un corpus de presse, elle visera un public peu expert et aura pour vocation la vulgarisation.

En revanche, pour ce qui concerne les marqueurs proprement dits, on peut s'attendre à ce que les marqueurs spécifiques au domaine, découverts grâce à l'utilisation d'un corpus de presse aligné, soient également pertinents pour découvrir des relations dans un corpus d'experts du même domaine, puisqu'en effet, les marqueurs sont moins sujets à variation que dénomination de concepts.

### 3.3 Les marqueurs de relation conceptuelle

Bien qu'ils aient fait l'objet d'un nombre d'études assez important, les marqueurs de relations ont été très peu recensés systématiquement. Or, pour cette étude, nous avons besoin de disposer de listes de marqueurs français et italiens pour pouvoir les projeter systématiquement sur le corpus. Pour le français, nous avons utilisé la liste construite par Lefeuvre dans sa thèse, dans le cadre du projet ANR Cristal (LEFEUVRE, 2017)<sup>10</sup>. La liste pour l'italien a été construite à partir de la liste de Lefeuvre, par traduction en italien et ajout de nouveaux marqueurs (FEDERZONI, 2017). Au total, pour l'hyperonymie, nous disposons de 39 marqueurs pour le français et 39 pour l'italien. Une fois la liste des marqueurs arrêtée, nous nous sommes employées à leur donner une forme qui soit utilisable de manière

---

<sup>9</sup> Notre traduction du texte original : « [...] *relative expert to the uninitiated* [...] », (PEARSON, 1998, p. 38).

<sup>10</sup> Liste disponible sur le site : [http://redac.univ-tlse2.fr/misc/mar-rel\\_fr.html](http://redac.univ-tlse2.fr/misc/mar-rel_fr.html) [consulté en décembre 2020].

automatique. Il a donc fallu constituer des grammaires locales, proches de celles que l'on trouve dans l'extraction d'information (BARNBROOK & SINCLAIR, 2001).

### 3.4 Mise en œuvre des marqueurs avec TXM

Nous avons importé notre corpus aligné dans TXM. Cette plate-forme est conçue pour aider les utilisateurs dans la construction et l'analyse de différents types de corpus numériques, en texte brut ou déjà étiqueté, aux différents formats. TXM permet d'intégrer facilement dans sa chaîne de traitement l'étiqueteur TreeTagger pour différentes langues. La plate-forme rend accessibles les résultats de l'étiquetage sous forme de propriétés de mots que l'utilisateur peut spécifier dans ses requêtes (Tableau 2). Le choix de la plate-forme TXM est motivé par les facilités qu'elle offre en termes de fonctionnalités de recherche (intégration du langage CQL et des expressions régulières). Nous avons donc opté pour ce logiciel afin de projeter les marqueurs de la relation d'hyponymie sur le corpus.

TABLEAU 2 – PROPRIETES DE MOTS ACCESSIBLES DANS TXM

	Propriété de mots	Exemple de requête
<b>Mot</b>	word	[word='mot à chercher']
<b>Lemme</b>	lemma	[(langue)lemma='lemme à chercher']
<b>Catégorie grammaticale</b>	pos	[(langue)pos='catégorie à chercher']

Dans un premier temps, en partant des deux listes à notre disposition, nous avons construit une requête ciblée pour chaque marqueur. Pour ce faire, nous avons exploité les expressions régulières et les informations de l'étiquetage morpho-syntaxique pour formuler des contraintes. Nous ne nous sommes pas limitées à rechercher les lemmes (comme par exemple *être*, *sorte*, *comme*), mais, dans le but de réduire le bruit produit par nos requêtes, nous avons exploité la possibilité de compléter les requêtes en explicitant et combinant différentes informations, comme le lemme, la catégorie grammaticale ; nous avons aussi exprimé des contraintes sur le temps verbal.

Ensuite, nous avons projeté nos requêtes sur le corpus étiqueté par TXM en exploitant ce même logiciel. La projection des requêtes sur le corpus nous a permis de les améliorer ; nous avons procédé, de manière itérative, à la modification des requêtes et à la projection en corpus. Une fois les requêtes finalisées, nous avons exporté les résultats de chacune d'elles afin de calculer les performances du marqueur correspondant.

Pour évaluer la performance effective de nos requêtes, il était indispensable de passer par une phase de validation manuelle. À la fin de notre annotation, nous avons été en mesure de calculer la précision pour chaque requête, c'est-à-dire la part des occurrences extraites qui contiennent effectivement la relation d'hyponymie. En revanche, nous n'avons pas pu calculer le silence, par manque d'un corpus de référence annoté manuellement, avec lequel confronter nos résultats.

## 4 Étude des résultats obtenus

Faute de place, nous ne détaillerons pas ici l'ensemble des résultats, ceux-ci sont développés dans une autre étude (FEDERZONI, 2018). Notons que sur les 1 071 contextes bilingues extraits, 487 correspondaient vraiment à une hyponymie (comme l'a montré l'annotation manuelle), soit près de 45,5 %. Parmi ces 487, 74 présentaient une variation en

lien avec la traduction, soit 15,2 %. La Figure 1 présente les types de variation rencontrés (manuellement validés), exprimés en pourcentage, par rapport au total des occurrences traitées (fr-it et it-fr).

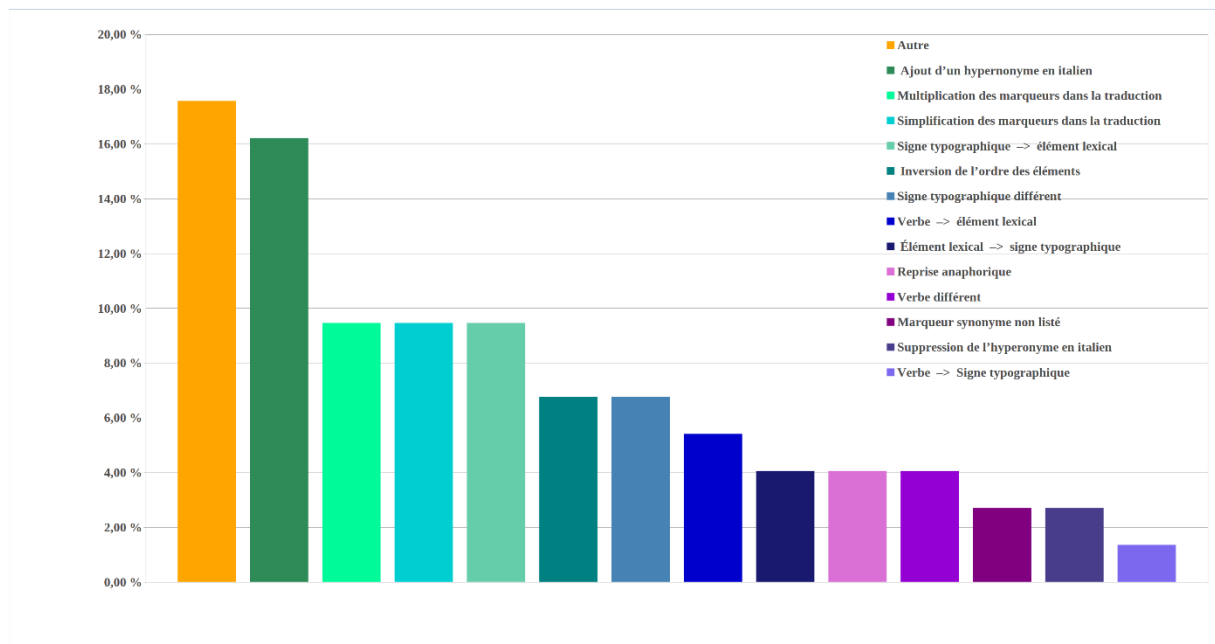


FIGURE 1 - TYPES DE VARIATION RENCONTRES

Voici le nombre de contextes concernés par chaque type de variation :

- Autre : 17,57 % ; 13 occurrences
- Ajout d'un hyperonyme en italien : 16,22 % ; 12 occurrences
- Multiplication des marqueurs dans la traduction : 9,46 % ; 7 occurrences
- Simplification des marqueurs dans la traduction : 9,46 % ; 7 occurrences
- Signe typographique → élément lexical : 9,46 % ; 7 occurrences
- Inversion de l'ordre des éléments : 6,76 % ; 5 occurrences
- Signe typographique différent : 6,76 % ; 5 occurrences
- Verbe → élément lexical : 5,41 % ; 4 occurrences
- Élément lexical → signe typographique : 4,05 % ; 3 occurrences
- Reprise anaphorique : 4,05 % ; 3 occurrences
- Verbe différent : 4,05 % ; 3 occurrences
- Marqueur synonymique non listé : 2,70 % ; 2 occurrences
- Suppression d'un hyperonyme en italien : 2,70 % ; 2 occurrences
- Verbe → signe typographique : 1,35 % ; 1 occurrence

En plus de l'ajout d'hyperonyme, dont on voit immédiatement l'intérêt, le dépouillement manuel des résultats nous a permis de recenser plusieurs phénomènes qui pourraient être utiles pour la constitution du réseau conceptuel sous-jacent au corpus par la relation d'hyperonymie. Ces phénomènes de variation sont étroitement liés aux traits typiques de la traduction, en particulier à l'« explicitation », à la « simplification » et au « nivellement » (BAKER, 1996).

## 4.1 Ajout d'un hyperonyme au moment de la traduction

Ce type de fonctionnement peut être rattaché à ce que Baker appelle « l'explicitation ». Dans ce cas, le traducteur ajoute un hyperonyme au moment de la traduction, sans doute pour donner une information supplémentaire au lecteur. Ainsi dans l'exemple 1 ci-dessous, le texte italien ajoute et précise donc que *La Repubblica* et *il Corriere della Serra* sont les deux plus grands journaux de la péninsule (*I due più grandi giornali della penisola*).

Exemple 1)

FR : « [...] Classés au centre gauche, La Repubblica et le Corriere della Serra [...] »

IT : [...] *I due più grandi giornali della penisola, la Repubblica e il Corriere della Serra.*

Ce type de traduction correspond précisément aux passages qui pourraient être utilisés dans la perspective d'enrichir un réseau conceptuel. En effet, un réseau conceptuel est censé représenter la connaissance d'un corpus constitué pour représenter un domaine (au sens large)<sup>11</sup>. Il est parfois difficile de construire des corpus assez volumineux pour que la constitution du réseau conceptuel soit productive. Si un corpus aligné existe, ce corpus pourrait contribuer à enrichir le réseau, grâce à l'apport de connaissances au moment de la traduction (comme l'ajout d'un hyperonyme). En effet, spécifiquement dans les corpus spécialisés et par hypothèse, les concepts et le réseau que l'on peut élaborer sont les mêmes dans un corpus et sa traduction. Dans le corpus, nous avons trouvé 12 cas de ce type (16,22 %). D'autres cas de variations au moment de la traduction permettraient, quant à eux, d'enrichir la liste des marqueurs de relations conceptuelles.

## 4.2 Suppression du marqueur dans la traduction

Ce fonctionnement peut être rattaché à la « simplification » de Baker. En simplifiant la syntaxe, le traducteur fait disparaître un marqueur.

Exemple 2)

FR : « [...] mais aussi par des taxes du type Tobin. »

IT : [...] *ma anche attraverso delle Tobintax.*

La confrontation des deux segments suggère donc pour l'italien un nouveau candidat-marqueur (*del tipo*). La traduction est aussi utile pour un examen des termes économiques. L'élimination du marqueur d'hyperonymie [du type X] nous permet de supposer la présence d'un terme intégré à l'usage général en italien, qui ferait donc partie des informations que le traducteur juge interprétables par le lecteur du *Manifesto*. Dans ce cas, l'apport de la traduction aussi réside dans la possibilité qu'elle offre de repérer un terme potentiel dans le domaine de l'économie en français : *Tobintax*, qui, d'après la traduction, semble être un calque terminologique entré dans l'usage de la presse en italien, pourrait être un candidat-terme en français, hypothèse qui demande à être vérifiée.

## 4.3 Ajout d'un marqueur dans la traduction

Dans ce type d'exemple, comme dans le précédent, l'apport se situe au niveau de l'enrichissement de la base des marqueurs. Il se peut en effet qu'apparaisse un marqueur non listé dans la base initiale.

---

<sup>11</sup>Notons que ce même exemple illustre un cas de suppression d'une information explicitée dans le texte de départ (« Classés au centre gauche ») qui pourrait participer elle aussi à l'enrichissement dans les deux langues du réseau conceptuel unique commun.

Exemple 3)

FR : « Bien qu'il s'agisse d'une association privée, la Ned figure au budget du département d'État. »

IT : *Sul piano giuridico, la Ned è un'associazione privata, eppure figura nel bilancio del dipartimento di stato.*

En simplifiant la syntaxe de la concession, le traducteur élimine la subordonnée concessive et l'anaphore entre les deux éléments de la relation (*association privée, Ned*) non repérables avec les marqueurs disponibles. Il opte pour des propositions indépendantes : l'usage du marqueur canonique de la relation d'hyponymie (*la Ned è un'associazione privata*), qu'il fait précéder d'une information ajoutée (*Sul piano giuridico*) et exprime la concession par un connecteur (*eppure / et pourtant*). Ces adaptations, qui ont sans doute pour but de rétablir les informations dans un ordre et une quantité jugés plus facilement interprétables, montrent comment trois des dimensions de la traduction identifiées par Baker – l'explicitation, la simplification et le nivellement – peuvent opérer dans une même phrase.

C'est grâce au marqueur canonique [X è det Y], ajouté dans la traduction, qu'il est possible de repérer, dans le corpus français, le verbe impersonnel (*s'agir de*) dans son usage de marqueur d'hyponymie en relation avec une anaphore. C'est aussi grâce à l'information ajoutée concernant le domaine juridique (*Sul piano giuridico*) que le statut terminologique respectif de *association privée* et *associazione privata* peut être questionné.

#### 4.4 Utilisation d'un autre type de marqueur dans la traduction

Ici, la variation ne concerne pas l'absence vs présence du marqueur, mais le fait que la traduction n'utilise pas le marqueur « isomorphe ». Cette pratique du traducteur présente le même intérêt que dans le cas précédent : un nouveau marqueur peut être identifié. L'exemple 4 illustre un cas de ce type. La mise en œuvre du marqueur italien [X è ... dét Y] et l'alignement de la version en français a permis d'identifier un marqueur non listé en français : [X arriver en tête de Y] qui permet de repérer une relation d'hyponymie entre Y (*pays industrialisés avancés*) et X (*Japon*).

Exemple 4)

IT : [...] *Secondo l'Organizzazione per la cooperazione e lo sviluppo economico (Ocse), nel 2005, il Giappone era il primo dei paesi più industrializzati per tasso di ricerca e sviluppo.* [...]

FR : « [...] Selon l'Organisation de coopération et de développements économiques (OCDE), en 2005, le Japon arrivait en tête des pays industriels avancés pour son taux de recherche et développement [...] »

En exploitant l'alignement du corpus et sa bidirectionnalité, il est possible de tirer parti des traits spécifiques de la traduction pour repérer des segments de texte original qui ne l'auraient pas été autrement.

Ces différents exemples font apparaître la double utilité du corpus aligné, d'une part pour l'enrichissement des réseaux de concepts, d'autre part pour l'établissement de listes plus complètes de marqueurs.

#### 4.5 Métalangage et métadiscours dans la traduction en italien

Un autre type d'indice pourrait être utile dans la perspective d'enrichir le réseau terminologique dans les deux langues de certaines informations, même s'il est d'un maniement difficile : il s'agit de la traduction du métadiscours. Il n'est pas rare que le

discours métalinguistique fasse apparaître les noms de concepts « en mention », c'est-à-dire marqués comme autonymiques, par le biais d'énoncés de désignation et de dénomination (REY-DEBOVE, 1978). L'apport de ce type de contexte réside dans le fait que le nom marqué par ce métadiscours est posé comme une dénomination « autorisée » du concept (ESCOUBAS-BENVENISTE, 2010). Par l'adaptation de son métadiscours, le traducteur ajoute sa propre connaissance terminologique au texte-cible en la donnant à voir. Ainsi, dans l'exemple suivant, la variation que subit le métadiscours quand il est traduit permet de « mentionner » le sigle *Lbo* comme autoréférentiel, c'est-à-dire comme une autre dénomination autorisée du concept de « *fondi di private equity* ». Il devient alors un terme synonyme, en vertu de l'autorité que le traducteur assume en matière économique.

Exemple 5)

FR : « [...] les fonds de private equity (ce qu'en France on appelle les LBO) [...] »

IT : [...] *i fondi di private equity (chiamati Lbo, in Francia o in Italia)* [...]

Pour être maintenue dans le passage à l'autre langue-culture, la visée didactique du genre textuel de presse dans des domaines spécialisés doit être aménagée et adaptée au contexte local du lecteur italoophone. Dans cet objectif, la traduction du métadiscours peut varier elle aussi selon la dimension de l'explicitation.

Dans l'exemple 5, l'information métadiscursive typographiquement délimitée par les parenthèses et repérable par le métalangage (*ce qu'... on appelle*) concerne le terme économique LBO (*Leveraged Buy-Out*), posé comme synonyme de *fonds de private equity*. La traduction du métadiscours a simplifié la syntaxe de l'opération de dénomination (*chiamati* / « appelés ») et nous apprend que ce sigle est aussi un terme en italien (*o in Italia*), qui se présente sous une forme graphique différente (*Lbo*).

Cette créativité discursive du traducteur par rapport au texte source est indispensable, si l'on considère le public visé par la traduction et la situation de communication : nous l'avons dit, les textes sont des articles de presse concernant des domaines de connaissance spécialisée que le traducteur expert doit transmettre à des lecteurs potentiellement novices. Le discours qui résulte de la traduction, pour être interprétable, se doit d'être en quelque sorte « didactisé » : il doit combler certaines lacunes informationnelles en explicitant des connaissances, spécifiques à la langue spécialisée du texte cible. Ce faisant, le traducteur se conforme aux contraintes du genre textuel – la presse spécialisée (*Le Monde diplomatique - Il Manifesto*) adressée aux non-spécialistes. Il répond, par des interventions dans le texte traduit, aux besoins de son lecteur en fonction de la représentation qu'il se fait – ou que la rédaction en chef du journal lui impose – du profil cognitif de son lectorat. Il devient « traducteur-rédacteur » et « doit anticiper le processus d'interprétation de son texte par ses lecteurs » en fonction du « genre interprétatif » attendu (BECIRI, 2008, p. 245-246).

Finalement, on voit que le corpus comparable, mis en œuvre dans les deux sens, sans que l'un prévale sur l'autre, permet d'identifier un ensemble de phénomènes qui peuvent être utiles pour le travail de constitution d'un réseau terminologique : repérage des concepts via leur réalisation en français ou en italien, repérage des relations, enrichissement de la base des marqueurs, repérage de dénominations alternatives, etc.

## 6 Conclusion

Ce chapitre s'est attaché à étudier en profondeur ce que l'utilisation d'un corpus de vulgarisation, comportant des textes et leur traduction en italien, pouvait apporter à l'analyse terminologique grâce au processus de traduction. Au carrefour de la terminologie textuelle et de la traductologie, l'étude s'est focalisée sur la traduction des marqueurs d'hyponymie

dans un corpus journalistique : *Le Monde diplomatique* et sa traduction en italien. L'ensemble des occurrences des 39 marqueurs d'hyponymie recensés pour le français et 39 recensés pour l'italien ainsi que leur traduction a été recherché grâce à TXM sur le corpus aligné français/italien. Au total, nous avons examiné 1 071 couples (segments) français/italien ou italien/français et, dans les 487 qui présentaient des variations liées à la traduction, nous avons repéré des phénomènes qui, en effet, présentent un intérêt pour la terminologie : ajout d'un hyperonyme, modification du marqueur (et donc possibilité d'enrichir les bases de marqueurs), ajout d'informations métalinguistiques. Ces résultats ouvrent de nombreuses perspectives et questionnements. Tout d'abord, il faudrait évaluer d'autres marqueurs de relation (l'étude pour la méronymie est en cours). Il se peut en effet que l'hyponymie soit particulièrement concernée par ces phénomènes de variation liée à la traduction. Il faudrait aussi construire réellement un (ou plutôt sans doute des) réseau(x) terminologique(s) afin d'évaluer l'apport réel de ces phénomènes de traduction. Rappelons que, en tout cas pour l'hyponymie, nous espérons augmenter la taille du réseau par l'ajout d'hyperonymes qui pourraient « raccrocher » différentes parties du réseau obtenu par l'utilisation de marqueurs de relation sur chaque corpus.

On pourrait penser que les cas intéressants recensés ne sont, au fond, pas très nombreux. Mais s'ils permettent de rapprocher des parties de réseau éparpillées, c'est un gain important.

Une question se pose toutefois à propos de la nature du corpus. En effet, nous avons utilisé un corpus journalistique. Il se peut que, dans ce genre textuel, les phénomènes d'ajout soient particulièrement présents en lien avec la volonté du traducteur de donner le plus d'informations possible. En d'autres termes, il se peut ainsi que le recours aux phénomènes d'ajout d'un hyperonyme lors de la traduction constitue une caractéristique du processus de traduction dans les articles de presse spécialisée.

Il n'est pas certain que ce même type de fonctionnement se retrouve dans des corpus très spécialisés ; en effet, le traducteur ne s'autorisera sans doute alors pas à rajouter des éléments. Il serait intéressant de vérifier cette hypothèse. Par ailleurs, comme nous l'avons signalé, les corpus les plus riches en marqueurs de relation ne sont pas les corpus les plus spécialisés mais plutôt les corpus qui ont une dimension didactique ou vulgarisatrice. Or les corpus journalistiques rentrent dans cette catégorie. Il serait nécessaire de construire véritablement un réseau conceptuel pour étudier jusqu'où cet apport est vérifié pour la pratique terminologique.

Quoi qu'il en soit, l'étude systématique d'un corpus dont la dimension de vulgarisation est avérée a permis de montrer la complémentarité de la traduction et de la terminologie textuelle pour construire des réseaux de concepts.

## Références bibliographiques

- AUGER Alain & BARRIERE Caroline, 2008, "Pattern Based Approaches to Semantic Relation Extraction: a state-of-the-art" , *Terminology*, vol. 14, n° 1, p. 1-19.
- BAKER Mona, 1996, "Corpus-based Translational Studies: the Challenges that Lie Ahead", dans H. Somers (dir.), *Terminology, LSP and Translation: Studies in Language Engineering, in Honour of Juan C. Sager*, Amsterdam/Philadelphia, John Benjamins Publishing Company, p. 175-186.
- BAKER Mona, 1998, « Réexplorer la langue de la traduction. L'approche basée sur le corpus », *Meta*, vol. 43, n° 4, p. 480-485.

- BARNBROOK Geoff & SINCLAIR John, 2001, "Specialised Corpus, Local and Functionnal Grammars", dans M. Ghadessy, A. Henry and R.L. Roseberry (dir.), *Studies in Corpus Linguistics*, Amsterdam/Philadelphie, John Benjamins Publishing Company, p. 237-75.
- BEACCO Jean-Claude & MOIRAND Sophie, 1995, « Autour des discours de transmission des connaissances », *Langages*, n° 117 (Les analyses du discours en France), p. 32-53.
- BECIRI Hélène, 2008, « Traduction spécialisée : quelques spécificités de la communication technique asymétrique », *Cahiers du CIEL*, (Aspects de la recherche en langue de spécialité), p. 243-268.
- BOURIGAUT Didier, JACQUEMIN Christian & L'HOMME Marie-Claude, 2001, *Recent Advances in Computational Terminology*, Amsterdam/Philadelphie, John Benjamins Publishing Company, coll.« Natural Language Processing, vol. 2 ».
- CHESTERMAN Andrew, 1998, *Contrastive Functional Analysis*, Amsterdam/Philadelphie, John Benjamins Publishing Company.
- CONDAMINES Anne, 2002, "Corpus Analysis and Conceptual Relation Patterns", *Terminology*, vol. 8, n° 1, p. 141-162.
- CONDAMINES Anne, 2018, "Terminological Knowledge Bases", dans P. Fuentes (dir.), *The Routledge Handbook of Lexicography*, London, Routledge, p. 335-349.
- CONDAMINES Anne & REBEYROLLE Josette, 2001, "Searching for and Identifying Conceptual Relationships via a corpus-based approach to a Terminological Knowledge Base (CTKB): method and results", dans D. Bourigault, M.C. L'Homme et C. Jacquemin (dir.), *Recent Advances in Computational Terminology*, Amsterdam/Philadelphie, John Benjamins Publishing Company, coll.« Natural Language Processing, vol. 2 », p. 127-148.
- CUENCA Maria-Josep, 2003, "Two Ways to Reformulate: a Contrastive Analysis of Reformulation Markers", *Journal of Pragmatics*, vol. 35, n° 7, p. 1069-1093.
- DELISLE Jean, 1980, *L'analyse du discours comme méthode de traduction. Initiation à la traduction française de textes pragmatiques anglais. Théorie et pratique*, Ottawa, Éditions de l'Université d'Ottawa, 282 p.
- DI FIORE Elena & DI VITO Sonia, 2008, « Le corpus français-italien du Monde diplomatique-II Manifesto », dans M. Hédiard (dir.), *Corpus et logiciels pour l'analyse linguistique*, Centro editoriale dell'Università di Cassino, [CD-Rom], ISBN : 978-88-902324-7-3.
- DRUETTA Ruggero, 1996, « Dix années de recherches contrastives (1984-1994) », *Franco-Italica*, n° 9, p. 11-66.
- ESCOUBAS BENVENISTE Marie-Pierre, 2010, « La définition dans le texte économique écrit de vulgarisation savante – première partie », dans M. Rossi (dir.) *Autour de la définition*, vol. 6, Publif@rum, n° 11, Università degli Studi di Genova, disponible en ligne sur <http://publiforum.farum.it/index.php/publiforum/article/view/201/1064> [consulté en décembre 2020].
- FRAWLEY William, 1984, "Prolegomenon to a Theory of Translation", dans W. Frawley (dir.), *Translation: Literary, Linguistic, and Philosophical Perspectives*, London & Toronto, Associated University Presses, p. 159-175.
- FEDERZONI Silvia, 2017, *Les marqueurs de la relation d'hyponymie en français et en italien*, Tesi di Magistrale, Corso di Scienze linguistiche letterarie e della traduzione, Università La Sapienza, Rome, non disponible en ligne.
- FEDERZONI Silvia, 2018, *Étude de la présence et du fonctionnement des marqueurs de relations conceptuelles dans un corpus aligné franco-italien*, mémoire de master, parcours Linguistique, informatique et technologie du langage, université Toulouse



- Jean Jaurès, Toulouse, disponible en ligne sur <http://dante.univ-tlse2.fr/9090> [consulté en décembre 2020].
- FLOWERDEW John, 1992, "Definitions in Science Lectures", *Applied Linguistics*, vol. 13, n° 2, p. 203-221.
- GANDIN Stefania, 2015, "Translating the Language of Tourism. A Corpus-based Study on the Translational Tourism English Corpus (T-TourEC)", *Social and Behavioral Sciences*, vol. 95, p. 325-335.
- GRANGER Sylviane, LEROT Jacques & PETCH-TYSON Stephanie, 2003, *Corpus-based Approaches to Contrastive Linguistics and Translation*, Amsterdam, Rodopi.
- HEDIARD Marie, 2002, « Le nuove tecnologie nella ricerca in linguistica e glottodidattica », Cassino, Centro editoriale di Ateneo dell'Università degli Studi di Cassino.
- HEDIARD Marie, 2007, *Linguistica dei corpora. Strumenti e applicazioni*, Cassino, Centro editoriale di Ateneo dell'Università degli Studi di Cassino, coll. « Collana Scientifica, 20 ».
- HEDIARD Marie, 2008, *Corpus et logiciels pour l'analyse linguistique*, Cassino, Centro editoriale di Ateneo dell'Università degli Studi di Cassino, [CD-Rom], ISBN :978-88-902324-7-3.
- HEIDEN Serge, MAGUE, Jean-Philippe & PINCEMIN Bénédicte, 2010, « TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement » (proceedings of 10th International Conference on the Statistical Analysis of Textual Data (JADT 2010), vol. 2, S. Bolasco (dir.)) Roma, Edizioni Universitarie di Lettere Economia Diritto, p. 1021-1032, disponible en ligne sur [http://lexicometrica.univ-paris3.fr/jadt/jadt2010/allegati/JADT-2010-1021-1032\\_025-Heiden.pdf](http://lexicometrica.univ-paris3.fr/jadt/jadt2010/allegati/JADT-2010-1021-1032_025-Heiden.pdf) [consulté en avril 2017].
- HOLMES James S, 1988, "The Name and Nature of Translation Studies", dans J. S. Holmes (dir.), *Translated Papers on Literary Translation and Translation Studies*, Amsterdam, Rodopi, p. 66-80.
- JACOBI Daniel, 1985, « Sémiotique du discours de vulgarisation scientifique », *Semen*, n° 2 (De Saussure aux médias), disponible en ligne sur <http://journals.openedition.org/semen/4291> [consulté en novembre 2020].
- JANOT Pascale, 2014, « L'escorte métalinguistique de "spread" dans les discours de vulgarisation économique traduits de l'italien vers le français : quels enjeux discursifs pour le traducteur ? », dans R. Druetta et C. Falbo (dir.), *Cahiers de recherche de l'École doctorale en linguistique française*, (Docteurs et recherche... une aventure qui continue), Università degli Studi Di Trieste, p. 111-127.
- JOHANSSON Stig, 2003, "Contrastive Linguistics and Corpora", dans S. Granger, J. Lerot et S. Petch-Tyson (dir.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*, Amsterdam, Rodopi, p. 31-45.
- KRAIF Olivier, 2001, « Exploitation des cognats dans les systèmes d'alignement bi-textuel : architecture et évaluation », *TAL*, vol. 42, n° 3 (Lexiques sémantiques), p. 833-867.
- LEFEUVRE Luce, 2017, *Analyse des marqueurs de relations conceptuelles en corpus spécialisé : Recensement, Évaluation en fonction du domaine et du genre textuel*, thèse de doctorat, université Toulouse Jean Jaurès, Toulouse.
- LENCI Alessandro, 2008, "From Context to Meaning: Distributional Models of the Lexicon in Linguistics and Cognitive Science", *Italian Journal of Linguistics*, vol. 20, n° 1, p. 1-31.
- MARSHMAN Elizabeth & L'HOMME Marie-Claude, 2008, « Portabilité des marqueurs de la relation causale : étude sur deux corpus spécialisés », dans F. Maniez, P. Dury, N. Arlin

- et C. Rougemont (dir.), *Corpus et dictionnaires de langues de spécialité*, Grenoble, Presses universitaires de Grenoble, p. 87-110.
- MARSHMAN Elizabeth, L'HOMME Marie-Claude & SURTEES Victoria, 2008, "Portability of Cause-effect Relation Markers across Specialised Domains and Text Genres: A Comparative Evaluation", *Corpora*, vol. 3, n° 2, p. 141-172.
- MOIRAND Sophie & TREGUER-FELTEN Geneviève, 2007, « Des mots de la langue aux discours spécialisés, des acteurs sociaux à la part culturelle du langage : raisons et conséquences de ces déplacements », *ASp*, n° 51-52 (Varia), p. 7-33.
- MORTUREUX Marie-Françoise (dir.), 1982, *Langue française*, n° 53 (La vulgarisation), 1982.
- PATIN Stéphane, 2014, « Les discours parlementaires européens : regard croisé français-espagnol », *Revue française de linguistique appliquée*, vol. XIX, n° 1, p. 71-86.
- PEARSON Jennifer, 1998, *Terms in Context*, Amsterdam/Philadelphia, John Benjamins Publishing Company.
- PEARSON Jennifer, 2000, « Une tentative d'exploitation bi-directionnelle d'un corpus bilingue », *Cahiers de Grammaire*, vol. 25 (Sémantique et Corpus), p. 53-69.
- REBEYROLLE Josette, 2000, *Forme et fonction de la définition en discours*, thèse de doctorat, université Toulouse Le Mirail, Toulouse.
- REY-DEBOVE Josette, 1978, *Le Métalangage*, Paris, Le Robert, coll. « L'Ordre des mots ».
- SCARPA Federica, 2001, *La traduzione specializzata – Lingue speciali e mediazione linguistica*, Milano, Hoepli Editore.
- SINCLAIR John, 1991, *Corpus, Concordance, Collocation: Describing English language*, Oxford, Oxford University Press.
- SKUCE Douglas & MEYER Ingrid, 1991, "Terminology and Knowledge Engineering: Exploring a Symbiotic Relationship", (proceedings of Banff Knowledge Acquisition For Knowledge-Based Systems Workshop (KAW'91)), p. 1-29.
- SLODZIAN Monique, 2000, « L'émergence d'une terminologie textuelle et le retour du sens », dans H. Béjoint et P. Thoiron, (dir.), *Le sens en terminologie*, Lyon, PUL, p. 61-85.
- TOURY Gideon, 1995, *Descriptive Translation Studies and Beyond*, Amsterdam/Philadelphia, John Benjamins Publishing Company.