

# Les corpus dans le cadre de la maîtrise en traduction

## [The use of corpora in the context of MA seminars in translation studies]

TERESA TOMASZKIEWICZ

*Université Adam Mickiewicz à Poznań, Pologne*

tomaszki@amu.edu.pl

### Resumen

Dans les programmes de formation des traducteurs nous retrouvons trois blocs de matières : une formation linguistico-traductologique, une formation professionnelle et une formation académique. Dans le cadre de cette dernière se trouve habituellement le séminaire de maîtrise qui aboutit à la soutenance des mémoires. Un professeur qui est censé diriger un tel séminaire affronte avec ses étudiants les différentes dimensions de l'analyse des corpus pour mener à bien leurs recherches.

Dans cet article, reposant sur une expérience de direction de plus de 150 maîtrises à l'Université Adam Mickiewicz en Pologne, on présente les problèmes suivants : la place de la recherche en maîtrise intégrée dans les programmes de formation des traducteurs au sein des universités, les corpus avant les années 90 vs les corpus contemporains, la collecte des corpus, la variété de sujets vs la variété de corpus, la compatibilité des corpus par rapport aux sujets de maîtrise et la méthodologie d'analyse des corpus.

### Abstract

Programs for the training of translators, especially in relation to Master's Translation Studies contain three subject blocks : linguistic and translation training, professional training and academic training. The academic training consists, as a rule, of the MA seminar which ends in the defense of the Senior thesis (Master's thesis). To properly carry out scientific research the Professor who conducts the students must tackle a variety of research corpora. In this article – which is based on our experience of 15 years as a supervisor of over 150 MA theses – we present various statistics which show how students choose the corpora. The following issues will be discussed: the role of the MA seminar in the overall training of translators at the university level, corpora dating back to the nineteen eighties as opposed to present-day corpora, literary and non-literary corpora and the diversity of the issues studied, suitability of corpora in connection with the issue under analysis, methodology of analysis.

### Mots clés

Maîtrise en traductologie , corpus littéraires , corpus non-littéraires , collecte des corpus, méthodologie d'analyse des corpus

### Keywords

Master's in translation , literary corpora , non-literary corpora , obtaining corpora , methodology of corpora analysis



Esta obra se publica con una licencia Creative Commons **Reconocimiento – NoComercial – SinObrasDerivada (by-nc-nd)**: No se permite un uso comercial de la obra original ni la generación de obras derivadas.

## 1. Remarques préliminaires

Le séminaire de maîtrise qui prépare les étudiants à l'élaboration d'un mémoire de fin d'études constitue un point important dans la formation universitaire des traducteurs et des interprètes, au moins dans les pays de l'Europe centrale et orientale. Pendant un an ou plus souvent deux ans les étudiants de deuxième cycle, sous la surveillance d'un professeur, sont censés effectuer des recherches ponctuelles dans le domaine de la traductologie, très souvent reposant sur l'analyse de différents corpus. La nature, la forme, l'étendue de ces corpus sont très variés, mais leur choix et leur adéquation par rapport à la recherche menée font, entre autres, l'objet d'une évaluation du jury pendant l'examen final. Dans cet article nous pensons présenter nos expériences en tant que directrice de ces séminaires de maîtrise depuis 20 ans pour démontrer la façon de constituer les corpus par les étudiants ainsi que les possibilités de les exploiter.

Pour ouvrir le débat sur le rôle des corpus dans l'élaboration d'une maîtrise en traduction nous allons tout d'abord rappeler quelques compréhensions de cette notion, pour ensuite introduire des informations concernant le fonctionnement des études traductologiques en Pologne avant 1990 et après. Nous allons aussi nous pencher sur le rôle du mémoire de maîtrise dans les programmes contemporains de formation des traducteurs et d'interprètes pour donner à la fin, à titre d'exemple, un aperçu de différentes recherches réalisées par les étudiants à l'Université Adam Mickiewicz à Poznań pendant ces vingt dernières années. Cette présentation n'a aucune prétention d'être exhaustive, mais doit servir comme point de départ pour les discussions ultimes concernant le rôle de ce mémoire de fin d'études pour la totalité de la formation traductologique.

## 2. Notion de corpus

Les études sur corpus semblent promises à un avenir fructueux, car elles ont déjà fait la preuve de leur efficacité et rencontrent de ce fait un intérêt croissant auprès des linguistes et des traductologues. Il n'empêche qu'il existe différentes façons de constituer des corpus et il n'y a pas, semble-t-il, une seule conception scientifique reconnue mais des méthodes et des acceptions aussi nombreuses que les objectifs des chercheurs. Passons en revue quelques opinions concernant la compréhension de la notion de *corpus* :

- La notion de corpus est définie dans la tradition comme un regroupement de textes fondateurs, en particulier dans la religion ou dans le droit. Cette conception canonique se réfère par exemple au corpus biblique dont la logique propre est de rassembler des textes qui par nature sont des éléments d'un ensemble. [...]
- Par opposition à cette notion canonique de corpus, on trouve des regroupements d'écrits ou de transcriptions de l'oral qui se justifient pour

des raisons de démonstration littéraires, linguistiques ou sociologiques [...].

- Depuis quelques années s'est renforcée une réflexion sur la 'linguistique de corpus' [...] Cette conception reconnaît le corpus comme un ensemble raisonné de textes, qui peut être éventuellement de grand volume (Trésor de la Langue Française, par exemple) et qui est traversé par une cohérence interne, nécessitant parfois une représentativité externe. (Ballard et Pineira-Tresmontant 2007 :7).

Il résulte de cette première ébauche de définition que les corpus aussi bien en linguistique qu'en traductologie sont des regroupements de textes authentiques, écrits ou oraux. Ces textes de plus en plus prennent une forme électronique qui facilite leur exploitation. Evidemment ils peuvent être analysés de différents points de vue qui concernent les différents aspects du fonctionnement de l'homme dans la société. N'oublions pas que la notion de corpus entre dans un champ beaucoup plus large qui concerne en général une réflexion commune et croisée, d'ordre épistémologique et méthodologique sur la notion de corpus telle qu'elle est exploitée aujourd'hui dans les sciences du langage, et plus largement, dans les sciences humaines et sociales. « On y traite des questions relatives aux notions de donnée, d'observables, d'empiricité, de théorie (lien type/ occurrence), de variable contextuelle, d'annotation, de codage, de catégorisation etc. On s'interroge sur la fonction des corpus dans l'activité de recherche. » (Laboratoire STIH -Sens, Texte, Informatique, Histoire, dirigé par Franck Neveu) ([www.stih.paris-sorbonne.fr](http://www.stih.paris-sorbonne.fr)).

Il est clair, que nous ne prétendons pas à résoudre dans cette courte présentation tous les dilemmes concernant la définition des corpus dans les sciences humaines. Dans le cadre de cet article nous allons nous limiter au contexte linguistique et traductologique.

G.Vetulani (2013 : 14) constate :

Pour un linguiste la réalité qu'il observe est constituée par les données provenant des corpus, c'est à dire des réalisations linguistiques pratiques au niveau de la *parole*, tels que: mots, fragments plus ou moins longs de la parole, provenant des énonciations concrètes, phrases, textes écrits ainsi que les enregistrements appartenant aux différents registres ou types de discours. (trad.T.T.)<sup>1</sup>

La collecte des unités soumises à une investigation linguistique peut se faire par une méthode traditionnelle, à travers une analyse, nous dirions manuelle, des textes ou des définitions dans les dictionnaires ou à l'aide de programmes électroniques spécialisés qui soutiennent un linguiste dans ses recherches grâce à une grande capacité et une vitesse beaucoup plus puissante de dépouillement des données que celle humaine. Or, dans les deux cas nous avons affaire aux recherches empiriques ce qui réserve la place à la linguistique parmi les scien-



ces se référant tout d'abord à l'observation des données réelles. La « matière » de cette « observation » est justement fournie par des corpus.

Il faudrait alors réfléchir jusqu'à quel point l'observation du fonctionnement de la langue et de la parole par un linguiste s'appuie sur les mêmes données que l'analyse du transfert linguistique et culturel pendant la traduction. Autrement dit, il faut se poser la question sur la façon d'exploiter des corpus par un linguiste et un traductologue.

Or, déjà dans cette courte présentation de la notion, nous avons pu voir que la signification et la pragmatique des corpus a évolué pendant les différentes périodes. En plus, les chercheurs signalent qu'il n'y a pas de corpus bon à tout et qu'il faut les personnaliser en fonction de la recherche.

Il ne peut y avoir de théorie générale des corpus. Réunir dans un ensemble des éléments langagiers apparentés dépend des items et du type de textes, de la méthode et des moyens d'analyse qui renvoient eux-mêmes aux projets de recherche. (Tournier 2007 : 49).

Il en résulte que la constitution du corpus, son étendue, son caractère sont dictés par le champ d'investigation, les hypothèses posées, les résultats recherchés. Aussi bien les linguistes que les traductologues observent les phénomènes ayant trait à l'utilisation de la langue, or dans le cas de la recherche en traduction les questions et les hypothèses surgissent dans le cadre de la rencontre d'au moins deux zones linguistico-culturelles, ce qui n'est pas forcément le cas de la recherche en linguistique. C'est probablement pourquoi C. Wecksteen (2007 : 261) signale l'existence « d'un fossé [...] entre la façon dont linguistes et traductologues conçoivent et utilisent un corpus », ce qui d'après cette auteure n'exclue pas « une discussion qui puisse être fructueuse pour les deux parties. » Cette auteure, par exemple, a entrepris de montrer comment une analyse fondée sur l'observation de corpus parallèles permettait à la traductologie de démontrer des phénomènes liés à l'opération même de la traduction qui met en relation non seulement des langues, mais aussi des textes. Elle souligne aussi que « même si linguistique contrastive et traductologie utilisent toutes deux les corpus parallèles, la méthode d'exploration, les objectifs visés, et par conséquent les enseignements qui en sont tirés, sont différents. » (Wecksteen 2007 : 262). Nous allons revenir à cette constatation dans notre analyse concernant l'exploitation des corpus par les étudiants. Toutefois, avant de passer à cette analyse il faudrait s'arrêter encore sur une autre observation.

L'intérêt énorme qu'attachent aujourd'hui les chercheurs en sciences humaines aux corpus en général, vu le nombre de publications, rencontres, débats autour de cette problématique pourrait faire penser que la notion de corpus est une invention nouvelle, caractéristique pour les recherches contemporaines. Cette confusion résulte du fait qu'on s'émerveille devant les nouvelles technologies

qui permettent de collecter des corpus et des bases de données énormes et de les traiter d'une manière rapide, beaucoup plus rapide qu'il y a encore une vingtaine d'années. Or, ceci ne veut pas dire que la notion de corpus est née à présent. Cl. Bocquet constate :

En lisant la littérature ancienne sur la traduction, celle qui a précédé la littérature traductologique au sens strict [...] on s'aperçoit qu'elle est largement fondée sur l'analyse de ce qu'on pourrait appeler aujourd'hui un corpus, dans la définition très classique qu'en donne, par exemple, Jean Dubois, dans la perspective de l'analyse de discours. 'Collection de données langagières sélectionnés et organisées selon des critères linguistiques explicites pour servir d'échantillon du langage' (Dubois 1969 : 116).<sup>2</sup>

Cl. Bocquet continue :

Étienne Dolet déjà fonde l'essentiel de ses considérations relatives à la manière de bien traduire sur un ensemble, certes très modeste, de traduction de Ciceron [...] Gaspar de Tende, le premier de son genre il est vrai, utilise, quant à lui, un véritable corpus, au demeurant gigantesque [...] (Bocquet, 2007 : 158-159).

Il faut donc admettre que la majorité des recherches en linguistique ou en traductologie doit s'appuyer sur l'observation des données langagières ou textuelles, mais cette observation, cette analyse sont différentes aujourd'hui par rapport à celles d'antan.

On peut remarquer sans problèmes que l'accessibilité de données linguistiques, surtout en ce qui concerne les langues pour lesquelles on a déjà élaboré des corpus spécialisés, est nettement plus grande que dans les époques précédentes. [...] Les recherches sur les corpus se développent de plus en plus. Grâce à elles on enrichie le savoir à la base des preuves enregistrées (usages) et non pas à la base des exemples préfabriqués, comme c'était parfois le cas avant (Vetulani, 2013 : 15) (trad.T.T.).<sup>3</sup>

Cette courte revue de quelques points de vue sur la compréhension et l'utilisation des corpus nous montre qu'il est relativement difficile d'en donner une définition pouvant embrasser tous les aspects. Nous pouvons seulement constater que c'est un phénomène en plein épanouissement. Les nouvelles technologies qui permettent de créer des corpus de plus en plus gigantesques imposent aussi un nouveau traitement de ces données et de nouvelles exploitations. Il est clair que cet outil technique est exploité par les domaines de plus en plus vastes et se trouve au croisement de plusieurs disciplines.

Dans ce contexte, il faut de nouveau réfléchir à la façon dont toutes ces possibilités d'exploitation des bases de données influencent le choix des corpus par les étudiants dans leurs recherches. Pour répondre à cette question nous allons tout d'abord nous arrêter sur quelques exemples concrets concernant l'organisation et le fonctionnement des séminaires de maîtrise en Pologne.



### 3. Un peu d'histoire

Toutes ces réflexions qui viennent d'être présentées concernent sans aucun doute tous les professeurs qui sont censés diriger les travaux de maîtrise aujourd'hui dans le cadre des recherches traductologiques largement comprises. Or, nous aimerions rappeler une époque un peu plus éloignée de la réalité contemporaine, à savoir le fonctionnement des études traductologiques en Pologne avant 1990. Ces expériences sont comparables avec celles d'autres pays de l'Europe centrale et orientale et peuvent servir comme modèle « d'antan ».

Dans ces pays les formations en langues après la Deuxième Guerre mondiale se faisaient dans le cadre de ce qu'on appelait - la philologie. D'ailleurs cette terminologie a été conservée de nos jours car à présent où plusieurs programmes de formation des traducteurs sont enregistrés, par exemple, sur la plateforme OPTIMALE (voir la note 5), la majorité sont intégrés dans la formation dite « philologique ». Or, il nous semble que l'aspect terminologique n'est pas crucial pour ce débat. Le centre du problème se trouve dans la reconnaissance du métier du traducteur par le système administratif en vigueur. Ce métier, au moins en Pologne, n'était pas reconnu par les autorités communistes. Il en résulte qu'on ne pouvait pas proposer aux autorités ministérielles de l'époque un programme universitaire de formation de traducteurs qui s'éloignait d'un programme général prévu pour « la philologie » ou à partir des années soixante-dix aussi pour « la linguistique appliquée ».

Ceci dit, la plus ancienne formation traductologique en Pologne a été réalisée à l'Université de Varsovie sous l'étiquette : « linguistique appliquée ». Or, ce programme devait répondre à plusieurs exigences : montrer qu'on formait des gens qui étaient préparés aussi bien pour assurer la didactique des langues que la traduction (linguistique appliquée) et qui en plus réalisaient le minimum obligatoire du programme prévu par le Ministère pour la formation universitaire appelée « la philologie ». Ce contexte paraissait relativement hostile à toute tentative de donner à la formation des traducteurs et des interprètes un statut de formations universitaire. En plus, il faut le dire pour une fois ouvertement, dans un système totalitaire toute personne maîtrisant des langues étrangères et par définition ayant des contacts avec des étrangers était nuisible pour le système en tant que tel. Par conséquent, toutes les formations philologiques et encore plus la formation des traducteurs étaient constamment surveillées par les autorités du pays.

Néanmoins, malgré ce contexte un peu hostile aux recherches et à la formation en traductologie, dans un contexte d'un pays au-delà du « rideau de fer », beaucoup de travaux de maîtrise ont été réalisés, contre vents et marées. Pour faire face à ces inconvénients il a fallu créer une façon spécifique de constituer des corpus estudiantins afin qu'ils réalisent leurs projets. N'oublions pas cette bana-

lité qu'à l'époque il n'y avait pas d'Internet, de tablettes, de communication électronique, de corpus électroniques et en général l'accès aux documents écrits en deux versions linguistiques était très restreint. Il a fallu créer des corpus un peu différemment et profiter de ce qu'on pouvait trouver sur le marché, p.ex. les documents parallèles des conventions et contrats internationaux publiés par une institution importante de cette époque et accessible au public, à savoir PISM (L'Institut Polonais des Affaires Internationales). En ce qui concerne la presse, par exemple française, un seul titre était disponible l'Humanité (l'organe officiel du Parti Communiste Français mais sans traduction en polonais). Cette situation imposait aux jeunes chercheurs le recours aux corpus littéraires.

Pourtant, Skibińska (2008) a bien démontré que la politique éditoriale en Pologne après la guerre a été largement marquée par une idéologie sous-jacente. Il en résulte que les étudiants en maîtrise n'avaient accès qu'aux corpus constitués par les traductions des œuvres littéraires françaises acceptées par les autorités en vigueur, à savoir à tout ce qui correspondait à l'idéologie communiste-marxiste. Dans ce type de conditions le type de corpus auquel les étudiant pouvaient avoir accès dictait le type et le but de la recherche. Évidemment nous parlons de petites recherches dans le cadre de maîtrise, mais même à ce niveau-là il faut se poser la question de la méthodologie. « On est donc conduit à discuter les méthodologies de constitution des corpus en termes de représentativité quantitative et qualitative par rapport aux phénomènes à décrire et à analyser, [...] » (Beacco 2002 : 149).

À l'époque que nous décrivons la réalité de la rédaction d'une maîtrise dans un contexte dépourvu d'accès à une vraie base de données limitait les recherches à une analyse du corpus qu'on pouvait difficilement constituer. Nous pouvons donc avancer que l'accès limité aux bases de données très souvent dictait le type de recherches effectuées. Nous pouvons paraphraser cette situation de façons suivante :

« Analysons de quoi nous disposons et en fonction de ceci nous pouvons préciser l'objet de la recherche ». Du point de vue scientifique cette attitude paraît minimaliste, mais elle a été acceptée par plusieurs directeurs de mémoires.

Les époques et les conditions du travail à l'université ont largement changé en 1990, c'est-à-dire à partir du changement du système politique et économique dans les pays de l'Europe de l'Est. Il n'empêche que la façon de travailler des étudiants en maîtrise dans notre nouvelle époque n'a pas beaucoup changé. Les observations de la façon de faire de plusieurs étudiants en maîtrise depuis une vingtaine d'années nous permet de constater majoritairement deux tendances :

- ils ont une idée de réflexion et essaient de choisir les corpus en fonction du but de la recherche, mais aussi souvent



- ce sont les corpus auxquels ils ont accès qui dictent le but de leur investigation.

Devant une telle réalité des choses on devrait peut-être revenir à la discussion sur le rôle de cette recherche ponctuelle dans le domaine de traductologie qui doit se terminer par un mémoire de fin d'études de maîtrise. Dans le monde contemporain où tout le monde essaie d'ajuster les programmes de formation universitaire aux besoins du marché il faut se poser la question dans quelle mesure la rédaction d'un mémoire de maîtrise peut contribuer à la réalisation de cet objectif.

#### **4. Rôle du séminaire de maîtrise en formation traductologique à l'époque contemporaine**

Passons donc à la réflexion concernant le rôle des séminaires de maîtrise dans le cadre d'une formation traductologique à l'époque contemporaine. La formation des traducteurs et d'interprètes fait l'objet depuis une dizaine d'années d'un nombre impressionnant de débats, de conférences, de projets internationaux et nationaux. Les universitaires et les professionnels se penchent sur le contenu des programmes d'enseignement, sur la qualité des formateurs, sur des outils aidant cette formation. Plusieurs chaînes et plateformes de coopération ont été créées, au moins au niveau européen, comme par exemple, EMT (European Master's in Translation)<sup>4</sup> ou la plateforme OPTIMALE<sup>5</sup> qui résulte de la coopération pendant 3 ans des directeurs de plus de 70 programmes de formation dans ce domaine dans toute l'Europe. Dans ces débats revient toujours la question concernant la place du séminaire de maîtrise dans ces formations.

Si tous, nous sommes d'accord que les études dans le domaine de traduction/interprétation sont des études professionnels, il n'empêche que dans les programmes de formation des traducteurs /interprètes<sup>6</sup> nous retrouvons trois blocs de matières enseignées: une formation linguistico-traductologique, une formation professionnelle et une formation académique. Dans le cadre de cette dernière se trouve habituellement le séminaire de maîtrise qui aboutit à la soutenance des mémoires. Ceci ne veut pas dire que tous les pays européens ont adopté cette façon de délivrer les diplômes (voir p.ex. la Grande Bretagne ou l'Espagne ou l'obligation de ce travail de fin d'études n'est pas nécessaire pour obtenir le diplôme). Mais, dans beaucoup d'entre eux, la Pologne y comprise, le mémoire de maîtrise à la fin des études est obligatoire. Evidemment la préparation de ce mémoire peut durer un ou deux ans, mais le plus souvent deux. Or, en général il n'existe aucun document qui préciserait les critères d'une maîtrise dans ce domaine. Alors, nous pouvons nous référer uniquement aux exemples concrets des directeurs de ces travaux. Ainsi, par la suite nous pensons parler de nos expériences en tant que directrice de ce séminaire de maîtrise depuis 20 ans.



## 5. Les séminaires de maîtrise en traduction à l'Université Adam Mickiewicz à Poznań

L'année 1990 marque une rupture entre l'époque du communisme et l'ouverture à un système démocratique. Toute sorte de bouleversements dans tous les domaines de la vie sociale et politique ont traversé le pays. Les premières institutions ouvertes aux changements étaient les universités. On a rejeté les programmes de formation imposés par le Ministère de l'Éducation et chaque université cherchait son chemin éducatif. La situation était toutefois difficile car pour réaliser de nouvelles visions de formation on avait besoin de spécialistes en formation qui manquaient dans beaucoup de domaines, la traductologie y comprise. Il a fallu plusieurs années pour constituer des équipes d'enseignants pouvant assurer la réalisations de nouveaux programmes, adaptés aux besoins de la société démocratique. A l'Université Adam Mickiewicz nous avons aussi travaillé pour former des formateurs dans le cadre des études traductologiques. Cette période a duré un certain temps car on ne peut pas former des cadres spécialisés en une ou deux années. Pourtant le séminaire de maîtrise couronné par une recherche a toujours existé. Finalement, notre programme de formation de traducteurs a été accepté par la Direction Générale de la Traduction de la Commission Européenne et nous avons pu adhérer au réseau de l'EMT ayant obtenu le certificat de qualité en 2011. Dans ce programme le séminaire de maîtrise est un élément important. Nous dirions plus, sans ce mémoire final un étudiant en Pologne ne peut recevoir aucune certification de fin d'études.

Au cours des quinze dernières années 153 maîtrises ont été rédigées sous notre direction. En principe, les étudiants ont la liberté de préciser leurs sujets ou leurs domaines d'intérêt. Presque tous ces travaux (151) reposaient sur l'analyse de différents corpus, dans la majorité des cas choisis librement par ces jeunes chercheurs. Si nous voulions caractériser ces corpus il faudrait les regrouper selon certaines catégories. La première grande ligne de partage (tout d'abord très générale) concerne la distinction entre les *corpus littéraires* et *non-littéraires*. Alors parmi ces 151 corpus 68 concernaient les textes littéraires, donc 45 % et le reste 55 % (83) toute sorte de textes non-littéraires.

### 5.1. Corpus littéraires

Parmi les textes littéraires une grande partie sont des romans ou des nouvelles, aussi bien contemporains que plus anciens : moyenâgeux, romantiques, du XIX s. ou tout à fait contemporains.

- 4 de ces corpus embrassaient des textes d'opéras ou de comédies musicales (6 %),
- et 4 des œuvres poétiques, chansons y comprises (6 %),



- 2 des textes de littérature enfantine (3 %),
- 6 (9 %) étaient constitués de pièces théâtrales.
- Nous avons classé dans cette catégorie aussi les textes bibliques (3) (4 %) non seulement parce qu'ils possèdent beaucoup des traits d'un texte littéraire mais aussi à cause de la problématique analysée.

Parmi les phénomènes analysés à la base de ces corpus littéraires nous pouvons évidemment citer :

- le transfert des différentes figures de style se plaçant au niveau paradigmatique ( métaphores, personnifications), syntagmatique ( apostrophes, répétitions, reformulations) et phonétique (assonances , allitérations) etc. mais aussi
- le transfert du comique ou des néologismes.

Toutefois, *le comique* comme problème de recherche peut être étudié à la base des corpus littéraires, mais aussi audiovisuels. Par contre, nous pouvons parler *des formations néologiques* dans les productions littéraires mais aussi dans le langage de presse ou en tant que nouvelle terminologie. En plus, les corpus littéraires permettent de réfléchir sur les prototypes textuels (descriptif, narratif, mais aussi dialogal).

Ceci dit, nous pouvons constater aussi que parmi ces 68 corpus littéraires 10 ont été utilisés pour étudier les problèmes lexicaux et morphosyntaxiques apparaissant pendant les transferts d'une langue à une autre.

Une autre problématique traitée à l'aide des corpus littéraires, mais aussi audiovisuels concerne les *différentes stylisations* : archaïque, dialectale, langage des subcultures des jeunes, langage érotique etc. (10 mémoires).

Dans la majorité des cas tous ces corpus sont principalement bilingues polonais-français ou français polonais mais, dans 7 cas trois systèmes linguistiques entraient en jeu : polonais + français + espagnol ou + anglais, ou + allemand. Il y avait encore 3 travaux où on a pris en compte deux ou plus de traduction d'un même texte.

Une autre forme de constituer les corpus amenait à pratiquer la traduction commentée des œuvres n'existant pas en version polonaise. Il s'agit d'un projet scientifique réalisé entre 2005 et 2010 avec les étudiants en maîtrise. L'objectif principal de ce projet au début était la diffusion et la promotion en Pologne de la littérature, langue et culture des communautés minoritaires francophones canadiennes, notamment: acadienne, franco-ontarienne, franco-saskatchewanaise et franco-manitobaine. Nous avons prévu trois éditions du séminaire pendant lesquelles les étudiants ont préparé leurs maîtrises sous forme de traduction commentée de leurs traductions des textes des auteurs choisis<sup>7</sup>. Dans la deuxième étape nous avons pris en compte la littérature belge contemporaine francophone. La majorité des réflexions concernaient les problèmes traductologiques liés aux

transferts culturels, stylisation dialectale, langage des jeunes, universalité des valeurs des cultures éloignées etc. Nous pouvons dire que dans ce projet les corpus étaient en quelque sorte imposés par des enseignants.

## 5.2. Corpus non-littéraires

En ce qui concerne *les corpus non-littéraires*, comme nous avons dit ils constituent 55 % du total. Or, leur caractère est beaucoup plus varié. Tout d'abord il faut parler des corpus écrits, des corpus oraux et des corpus mixtes qui intègrent des éléments oraux, écrits et sémiotiques (gestes, mimiques, images, icônes) par exemple dans la traduction audiovisuelle. Dans cette catégorie s'intègrent aussi des corpus parallèles. Cette énumération ne présente peut-être pas quelque chose de nouveau, mais il est intéressant d'évoquer certaines statistiques.

### 5.2.1. Langues de spécialité

Les corpus écrit dans le domaine *des langues de spécialité* constituent 30 % de tous les corpus. Évidemment la notion de langues de spécialité est très vaste et elle peut être exploitée de manières différentes c'est pourquoi nous mentionnons seulement la répartition proportionnelle de ce différents domaines qui faisaient l'objet de recherches dans le cadre de notre séminaire. Le domaine qui est le plus représenté c'est *le droit* (11), suivi de *médecine* (4), et ensuite un travail dans les domaines suivants : *économie, technique, musique, tourisme, église, torture, marin, pêche, chasse, militaire*. Treize de ces travaux peuvent être classifiés en tant que recherches typiquement terminologiques et sept ont envisagé aussi des aspects morfo-syntaxiques.

### 5.2.2. Textes de presse

Les deuxième groupe de textes écrits ce sont *les textes de presse* (17 - 20 %), qui par leur nature sont eux aussi très variés : presse politique, presse féminine, vulgarisation scientifique, mais aussi publicité. A un certain moment quelques travaux ont été consacrés aux *modèles de textes* (8) fonctionnant dans deux ou trois sociétés et ayant des matrices un peu différentes, comme des recettes de cuisine, des faire-part de décès, des modes d'emploi, des petites annonces, mais aussi les contrats de droit civil (de bail ou de location) etc. Dans ces cas souvent les corpus avaient la structure des textes parallèles. Par conséquent, nous avons aussi proposé d'étudier *des textes parallèles* télévisés comme des journaux télévisés, des bulletins météo ou des talk-show, mais cette fois-ci nous avons eu affaire aux corpus audiovisuels.



### 5.2.3. *Textes audiovisuels*

D'une manière plus sommaire nous pouvons dire que parmi 83 corpus non-littéraires 20 (24 %) étaient constitués par les corpus audio-visuels (films compris). Nous avons classé parmi cette catégorie aussi les travaux consacrés aux traductions des films pour les malentendants (2) et l'audio description (1). Il est clair que l'analyse de ces corpus doit aussi intégrer la dimension sémiologique largement comprise, c'est qui est encore plus important dans le groupe suivant.

### 5.2.4. *Corpus médiaux à dominante sémiotique*

Dans cette catégorie nous intégrons non seulement des textes audiovisuels mais aussi les corpus de presse, étudiés du point de vue non seulement linguistique mais aussi par rapport à la *dimension sémiotique du document* (images, graphismes, icônes, dimensions mimique et gestuelle). Pendant la période analysée dix travaux ont concerné ce genre de réflexion. Quelque uns ont débattu, par exemple, du problème de l'image stéréotypique de la femme ou de l'homme dans la publicité. D'autres ont comparé les stéréotypes en langue et en images. Nous avons donc pris en compte les différentes dimensions de la traduction intersémiotique.

### 5.2.5 *Corpus oraux*

Et finalement il faut parler des corpus oraux se composant d'enregistrements audio ou audiovisuels. Il s'agit donc des enregistrements des conversations authentiques, des interprétations de conférence ou du matériel audiovisuel. Dans le premier cas l'objet de l'investigation concernait le déroulement de la conversation. Les étudiants ont pris en compte des phénomènes tels que les échanges comportant des « excuses », des « offres », des « termes d'adresse », des « titres honorifiques », les phénomènes de textualisation conversationnelle (8 travaux = 10 %). Pendant l'analyse de l'interprétation de conférence (2 travaux) et la traduction à vue (1) on a analysé le problème de reformulation, la qualité de l'interprétation ou le passage du code écrit au code oral.

### 5.2.6. *Traduction automatique*

La dernière catégorie qui malheureusement est peu représentée dans ces travaux que nous venons de présenter concerne la traduction automatique et l'utilisation des outils informatiques dans la traduction (3 = 4 %). Malgré la formation de ces étudiants qui intègre l'utilisation des outils informatiques pendant la traduction ou le recours aux bases de données les étudiants ne paraissent pas tellement intéressés par une recherche dans ce domaine. Nous supposons qu'avec le développement de plus en plus poussé des techniques informatiques en traductologie cette tendance peut changer.

Pour terminer cette revue des corpus et de leur utilisation dans les maîtrises il faut encore mentionner un aspect important. Un corpus peut être aussi utilisé pour analyser des techniques traductologiques menant à la résolution des problèmes, mais aussi pour détecter des erreurs de traduction, erreurs commises par l'homme mais aussi par la machine (11 travaux).

## 6. Conclusion

Aujourd'hui, les jeunes chercheurs qui préparent leur maîtrise travaillent sur des corpus de plus en plus diversifiés et multimodaux. Comme nous avons démontré, les corpus peuvent être livresques, numériques, multimédia (sites internet de presse en ligne, forum...), plurilingues ou bien encore plurisémiotiques, comme une collection d'objets ou d'œuvres d'art contextualisant les discours d'exposition artistiques, des enregistrements audio-visuels dans des espaces publics (avec la prise en compte de la gestuelle).

Très souvent les disciplines concernées dépassent le cadre strict de la linguistique ou de la littérature. Il n'empêche que les étudiants en maîtrise doivent faire face à ces corpus hétérogènes, hybrides et la traductologie n'apporte pas toujours un cadre d'analyse adéquat. Nous affrontons alors la notion d'hybridité des corpus qui doit être entendue comme un ensemble de données en contexte, c'est-à-dire situées dans un environnement signifiant et constituant de nouveaux points de questionnement pour toutes les disciplines. Il en résulte, qu'il faut que le directeur de thèse soit capable de suggérer tout d'abord le champ d'exploitation, de donner des outils d'analyse pour aider à intégrer la recherche en question à certains courants en linguistique ou en étude littéraire.

## 7. Références

- Ballard, Michel et Carmen Pineira-Tresmontant, eds, 2007. *Les corpus en linguistique et en traductologie*. Arras : Artois Presses Université Traductologie. ISBN 978-2-84832-063-2.
- Beacco, Jean-Claude. 2002. Corpus. @ P. Charadeau et D.Mainguenau, eds. *Dictionnaire d'Analyse du Discours*. Paris : Éditions du Seuil. ISBN 2-02-037845-0, pp.148-151.
- Bocquet Claude. 2007. D'un corpus pour la traduction à un corpus pour une véritable histoire de la traduction. @M.Ballard et C. Pineira-Tresmontant, eds. *Les corpus en linguistique et en traductologie*. Arras : Artois Presses Université Traductologie. ISBN : 978-2-84832-063-2. pp. 159-174.
- Bocquet, Claude. 2008. *La traduction juridique. Fondement et méthode*. Bruxelles : De Boeck. ISBN 978-2-8041-5928-3.



- Dubois, Jean. 1973. Corpus. @ J.Dubois et al, eds. *Dictionnaire de linguistique*. Paris : Larousse. ISBN 20299-10-72, pp. 128-129.
- Skibińska, Elżbieta. 2008. *Kuchnia tłumacza. Studia o polsko-francuskich relacjach przekładowych*. Kraków : Universitas. ISBN 97883-242-0953-8.
- Tomaszkiewicz, Teresa. 2008. Transfert de la littérature acadienne en Pologne : bilan d'une expérience traductologique et didactique. @ *Alternative francophone 1 /1*: 98-118. <<http://ejournals.library.ualberta.ca/index.php/af/issue/current>>.
- Tomaszkiewicz, Teresa. 2009. Granice przekładalności współczesnej literatury akademickiej w świetle myśli Olgierda Wojtasiewicza. @ Hejwowski K., Szczęsny A. et Topczewska U. eds. *50 lat polskiej translatoryki*. Warszawa: ILS, Uniwersytet Warszawski. ISBN 978-83-929439-0-7, pp.163-174.
- Tomaszkiewicz, Teresa. 2012. Croisement des langues et des cultures dans la traduction. @ Kacprzak A., A. Konowska et M. Gajos, eds. *Pluralité des cultures : chances ou menaces ? Analyses linguistiques et didactiques*. Łódź – Łask : Oficyna Wydawnicza LEKSEM. ISSN 2107-7029, pp. 275-285.
- Tournier, Maurice. 2007. Corpus de textes en lexicométrie sociopolitique. @ M.Ballard et C. Pineira-Tresmontant, eds. *Les corpus en linguistique et en traductologie*. Arras : Artois Presses Université Traductologie. ISBN 978-2-84832-063-2, pp. 49-60.
- Vetulani, Grażyna. 2013. Problemy i korzyści wynikające z automatycznego przetwarzania korpusów – na przykładzie badań z zakresu predykcji rzeczownikowej w języku polskim. @ D. Śliwa (ed.) *Roczniki Humanistyczne, z. 8, Lingwistyka korpusowa i translatoryka*. pp. 13-24. ISSN: 0035-7707.
- Wecksteen, Corinne. 2007. Le corpus en traductologie : un moyen d'observation pour une approche réaliste de la traduction : application à quelques phénomènes connotatifs. @ M.Ballard et C. Pineira-Tresmontant, eds. *Les corpus en linguistique et en traductologie*. Arras : Artois Presses Université Traductologie. ISBN 978-2-84832-063-2, pp. 261-282.

---

<sup>1</sup> Vetulani 2013, citation en original: « Dla lingwisty rzeczywistością obserwowaną jest zgromadzona w korpusach "materia", czyli konkretne realizacje językowe (wykonania na poziomie parole), takie jak: słowa, krótsze i dłuższe fragmenty wypowiedzi, zdania, teksty pisane, ale także nagrania mowy pochodzące z różnych rejestrów i typów dyskursu.»

<sup>2</sup> Pour une définition plus ample on peut consulter Dubois et al., 1973.

<sup>3</sup> Vetulani 2013, citation en original: „Rzeczą oczywistą jest, iż obecnie obserwacja zjawisk językowych różni się w swoim charakterze od tej, która miała miejsce dawniej. Bez trudu da się zauważyć - przynajmniej jeśli chodzi o języki, dla których stworzono już nowoczesne i reprezentatywne korpusy - że dostępność do faktów językowych jest dzisiaj dużo większa niż dawniej. [...] Badania korpusowe są dziś niezwykle zaawansowane. Dzięki nim wnosi się do nauki wiedzę na podstawie zarejestrowanych zaświadczeń (uzusu), a nie - jak to bywało niekiedy w przeszłości - poprzez preparowane przykłady językowe.»

---

<sup>4</sup> European Master's in Translation : [http://ec.europa.eu/dgs/translation/programmes/emt/universities/index\\_en.htm](http://ec.europa.eu/dgs/translation/programmes/emt/universities/index_en.htm).

<sup>5</sup> OPTIMALE : <http://www.translator-training.eu/>.

<sup>6</sup> Nous pensons qu'à ce point il faut faire une remarque importante qui concerne la distinction entre les notions de traducteur et d'interprète. Toutes les nouvelles formations universitaires font clairement la différence entre les traducteurs des textes écrits et les interprètes des textes oraux. Or, Cl. Bocquet (2008 : 8) constate : « [...] il aurait été également possible, et peut-être bienvenu, d'accueillir en français, comme cela a été fait en allemand, le terme de 'translatologie' à côté du terme 'traductologie', pour englober l'interprétologie'. Cela aurait été d'autant plus justifié qu'avant Robert Estienne en 1539 et Étienne Dolet en 1540, on employait en français le verbe 'translater' dans un sens correspondant aux deux verbes modernes de traduire et d'interpréter [...] ». Il en résulte que quand nous allons parler dans la suite des séminaires de maîtrise, ces séminaires embrassent aussi bien des étudiants en traduction qu'en interprétation.

<sup>7</sup> A propos de ce sujet on peut lire plus dans Tomaszkiwicz (2008, 2009, 2012).

