

Utilisations de *mkAlign* pour la traduction philologique

Serge Fleury, Maria Zimina

EA2290 SYLED/CLA²T, Université de la Sorbonne nouvelle – Paris 3 – France

Abstract

We propose an approach of computer-aided translation which enables the comparison between the original text and its translation variants, emphasizing the similarities/divergences of different versions. This approach presents genuine advantages when working on philological translations.

Résumé

Nous proposons une approche de l'analyse de la traduction assistée par ordinateur qui permet de comparer le texte original avec ses variantes de traduction en soulignant les convergences/divergences des versions. Elle peut présenter de réels avantages pour le travail de traduction philologique

Mots-clés : alignement, corpus multilingues, textométrie, traduction philologique.

1. La traduction philologique

Le concept de traduction philologique a été introduit par la célèbre philologue russe Olga Akhmanova et développé dans des travaux ultérieurs (Akhmanova et Zadornova, 1972 ; 1977a ; 1977b). La traduction philologique d'une œuvre littéraire est élaborée par le philologue sur la base des variantes de traduction existantes. Après une étude attentive du texte original, les traductions disponibles sont analysées et comparées entre elles. Les imperfections de chacune des variantes sont identifiées et écartées. Le philologue tente ensuite de réunir les avantages de chacune des traductions dans une nouvelle version, tout en épurant l'expression et le style. Au final, la nouvelle version améliorée est considérée comme une traduction « sans auteur » qui a pour objectif de refléter le plus fidèlement possible le texte original. Cette approche est utilisée à des fins de recherche.

La méthodologie de la traduction philologique s'appuie sur une analyse stylistique comportant trois niveaux : *sémantique*, *metasémantique* et *linguo-poétique*. L'analyse sémantique constitue le niveau de base : on n'accède à la compréhension de l'ensemble du texte qu'après avoir établi le sens des mots dont il se compose. Cette première phase de l'analyse stylistique est souvent laborieuse et nécessite une exploration textuelle approfondie (Zimina, 1996).

La traduction philologique peut être assistée par ordinateur et bénéficier de l'expérience méthodologique acquise dans le domaine de l'analyse quantitative des données textuelles. La méthodologie de l'étude quantitative des textes alignés, ou *textométrie multilingue*¹, que nous mettons en œuvre en ce moment autour du logiciel *mkAlign* (Fleury et Zimina, 2007a ; 2007b) a été spécialement conçue pour prendre en compte ce besoin particulier de comparaison synchronisée des fragments de textes, écrits dans une même langue, ou dans des langues différentes (Zimina, 2002 ; 2004). Pour présenter ces méthodes de travail mises en œuvre à

l'aide de *mkAlign*, nous utiliserons un corpus rassemblant deux traductions anglaises du roman « Le Maître et Marguerite » de Mikhaïl Boulgakov.

2. Corpus des variantes de traduction

2.1. Le roman « *Le Maître et Marguerite* »

Le roman « Le Maître et Marguerite » (en russe : « Мастер и Маргарита ») a été commencé par M. Boulgakov au début des années 1928-1929. Le manuscrit a ensuite été abandonné, repris, détruit, restitué, puis réorganisé à plusieurs reprises. La première partie du roman est parue en 1966, la deuxième en 1967. Cette œuvre classique de la littérature russe a marqué les esprits des lecteurs au point que certaines expressions sont devenues proverbiales : « *Les manuscrits ne brûlent pas.* » « *La lâcheté est le pire de tous les défauts !* »².

Le roman est composé de trois histoires entrelacées de manière complexe. La première offre une représentation satirique de la vie de Moscovite des années trente, avec ses normes et ses dogmes, fortement remis en question par l'apparition du personnage de Satan (Woland). La deuxième histoire se situe à Jérusalem, sous le gouvernement du procureur Ponce Pilate. Cette histoire, la plus controversée de toutes, décrit un acte de lâcheté, regretté par la suite, et un dialogue interrompu entre Ponce Pilate et Yeshua Ha-Notsri (Jésus de Nazareth). La troisième histoire est un roman au sens plus traditionnel : la célébration de l'amour entre une jolie femme (Marguerite) et son amant (le Maître), homme de talent (cf. figure 1).

Le Maître et Marguerite se prête facilement à plusieurs interprétations. Dans la littérature critique, Mikhaïl Boulgakov est considéré proche de Mikhaïl Bakhtine, dont la vision carnavalesque de la culture et de la littérature est reflétée dans les termes *polyphonie*, *plurivalence*, *hétérogénéité*, *mélange des styles* (Belleau, 1970).

2.2. Corpus de travail : volet *SOURCE* – deux volets *CIBLES*

Nous avons réunis dans un même corpus de travail le texte original du roman (en russe) et deux variantes de traductions en anglais espacées dans le temps. La première traduction de Michael Glenly date de 1969. La deuxième, réalisée par Richard Pevear et Larissa Volokhonsky, est parue en 1997 (cf. Corpus *infra*). De façon générale, les deux traductions respectent l'agencement des chapitres du texte original.

Dans les explorations qui suivent, nous appellerons le corpus original *M&M_ru* et les deux traductions *M&M_en69* et *M&M_en97*. Chacun de ces trois textes contient environ 150 000 occurrences.

2.3. Variations dans la traduction

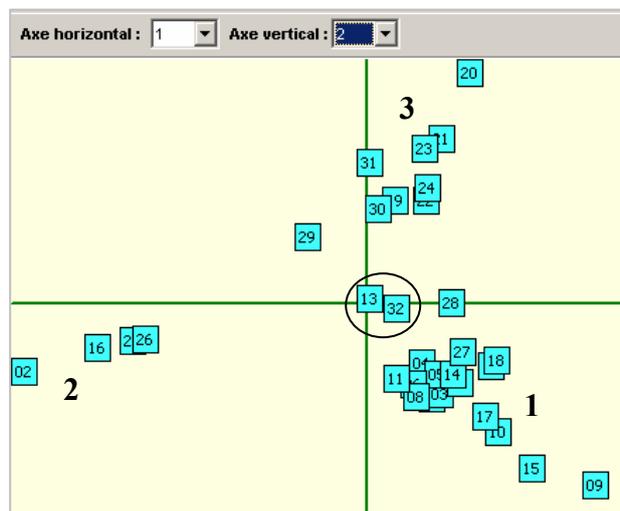
Les résultats de l'analyse factorielle des correspondances (AFC) du corpus de travail partitionné en 32 chapitres (parties) montrent que les trois volets se construisent autour des trois pôles thématiques (trois actions du roman interconnectées entre elles, cf. figure 1).

En reflétant l'ensemble de l'œuvre originale de façon similaire, les deux traductions ont chacune leur identité stylistique. Lorsque l'on compare les deux textes (en commençant par le niveau sémantique, comme le suggèrent les principes de l'analyse stylistique évoquée plus haut), on constate que les différences sont extrêmement nombreuses.

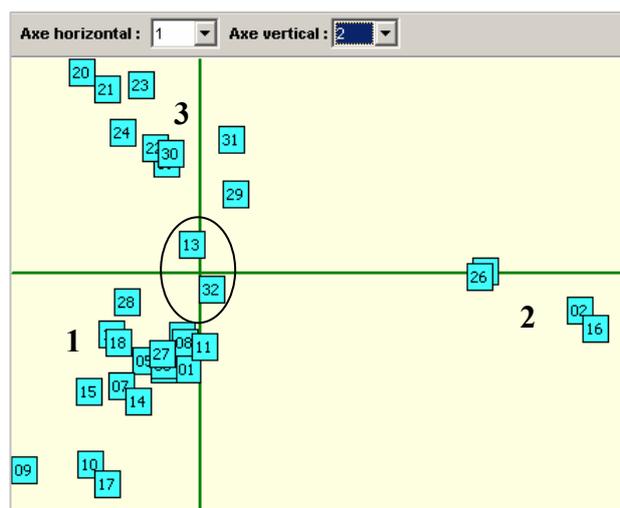
Sur les figures 2 et 3, la comparaison du texte original en russe avec chacune de ces deux traductions en anglais est facilitée grâce à l'éditeur à double entrée : les textes sont présentés

côté à côté, par paquets de fragments alignés. Cette représentation est obtenue à l'aide de *mkAlign*.³

M&M_ru (texte SOURCE)



M&M_en69 (texte CIBLE 1)



M&M_en97 (texte CIBLE 2)

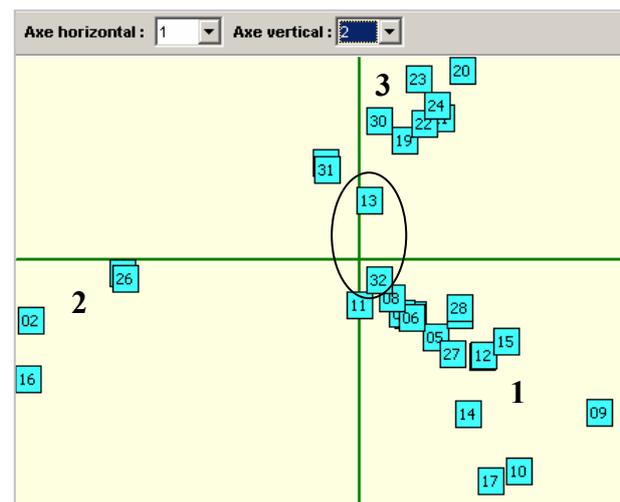


Figure 1 : Résultats de l'analyse factorielle des correspondances du corpus *M&M*.

Guide de lecture de la figure 1:

Sur les graphiques générés à l'aide de *Lexico3* (Lamalle *et al.* 2003), chacun des volets du corpus est partitionné en 32 chapitres.

Le premier graphique correspond au texte original du roman « Le Maître et Marguerite » (en russe). Les autres graphiques représentent ces traductions en anglais.

Sur les trois graphiques, les regroupements des parties observés sur le premier plan de l'AFC (premier et deuxième facteurs) montrent avec beaucoup de netteté la présence de trois pôles distincts (1, 2 et 3). Ces trois pôles correspondent aux trois actions qui se mélangent au fil du roman.

Au centre de chacun des graphiques, nous retrouvons les chapitres 32 et 13. Le chapitre 32 est le dernier chapitre du roman où tous les personnages se retrouvent. Le chapitre 13 « Apparition du héros » décrit la première rencontre des deux personnages principaux : c'est le début de l'action dans l'histoire du Maître et Marguerite.

Les résultats de l'analyse factorielle de ces trois ensembles textuels montrent que les deux traductions espacées dans le temps reflètent l'ensemble de l'œuvre originale de façon similaire.

mkAlign 2.00	<i>M&M_ru</i>	<i>M&M_en69</i>
	<p>-- она несла в руках отвратительные, тревожные желтые цветы. черт их знает, как их зовут, но они первые почему-то появляются в москве. и эти цветы очень отчетливо выделялись на черном ее весеннем пальто. она несла желтые цветы! нехороший цвет.</p>	<p>'she was carrying some of those repulsive yellow flowers. god knows what they're called, but they are somehow always the first to come out in spring. they stood out very sharply against her black dress. she was carrying yellow flowers! it's an ugly colour.</p>
	<p>она повернула с тверской в переулочек и тут обернулась. ну, тверскую вы знаете? по тверской шли тысячи людей, но я вам ручаюсь, что увидела она меня одного и поглядела не то что тревожно, а даже как будто болезненно.</p>	<p>she turned off tverskaya into a side-street and turned round. you know the tverskaya, don't you? there must have been a thousand people on it but i swear to you that she saw no one but me.</p>
	<p>и меня поразила не столько ее красота, сколько необыкновенное, никем не виданное одиночество в глазах! повинувшись этому желтому знаку, я тоже свернул в переулочек и пошел по ее следам. мы шли по кривому, скучному переулочку безмолвно, я по одной стороне, а она по другой.</p>	<p>she had a look of suffering and i was struck less by her beauty than by the extraordinary loneliness in her eyes. obeying that yellow signal i too turned into the side-street and followed her. we walked in silence down that dreary, winding little street without saying a word, she on one side, i on the other.</p>
	<p>и не было, вообразите, в переулочке ни души. я мучился, потому что мне показалось, что с нею необходимо говорить, и тревожился, что я не вымолвлю ни одного слова, а она уйдет, и я никогда ее более не увижу... и, вообразите, внезапно заговорила она:</p>	<p>there was not another soul in the street. i was in agony because i felt i had to speak to her and was worried that i might not be able to utter a word, she would disappear and i should never see her again. then, if you can believe it, she said :</p>

Figure 2 : « Le Maître et Marguerite » : texte original et sa traduction en anglais de 1969

<i>M&M_ru</i>	<i>M&M_en97</i>
<p>-- она несла в руках отвратительные, тревожные желтые цветы. черт их знает, как их зовут, но они первые почему-то появляются в москве. и эти цветы очень отчетливо выделялись на черном ее весеннем пальто. она несла желтые цветы! нехороший цвет.</p>	<p>'she was carrying repulsive, alarming yellow flowers in her hand. devil knows what they're called, but for some reason they're the first to appear in moscow. and these flowers stood out clearly against her black spring coat. she was carrying yellow flowers! not a nice colour.</p>
<p>она повернула с тверской в переулочек и тут обернулась. ну, тверскую вы знаете? по тверской шли тысячи людей, но я вам ручаюсь, что увидела она меня одного и поглядела не то что тревожно, а даже как будто болезненно.</p>	<p>she turned down a lane from tverskaya and then looked back. well, you know tverskaya! thousands of people were walking along tverskaya, but i can assure you that she saw me alone, and looked not really alarmed, but even as if in pain.</p>
<p>и меня поразила не столько ее красота, сколько необыкновенное, никем не виданное одиночество в глазах! повинувшись этому желтому знаку, я тоже свернул в переулочек и пошел по ее следам. мы шли по кривому, скучному переулочку безмолвно, я по одной стороне, а она по другой.</p>	<p>and i was struck not so much by her beauty as by an extraordinary loneliness in her eyes, such as no one had ever seen before! obeying this yellow sign, i also turned down the lane and followed her. we walked along the crooked, boring lane silently, i on one side, she on the other.</p>
<p>и не было, вообразите, в переулочке ни души. я мучился, потому что мне показалось, что с нею необходимо говорить, и тревожился, что я не вымолвлю ни одного слова, а она уйдет, и я никогда ее более не увижу... и, вообразите, внезапно заговорила она:</p>	<p>and, imagine, there was not a soul in the lane. i was suffering, because it seemed to me that it was necessary to speak to her, and i worried that i wouldn't utter a single word, and she would leave, and i'd never see her again. and, imagine, suddenly she began to speak:</p>

Figure 3 : « Le Maître et Marguerite » : texte original et sa traduction en anglais de 1997

Comme le montrent les résultats de la comparaison de ces trois extraits du corpus (cf. figures 2 et 3), les différences concernent le découpage en phrases, le choix du lexique, l'importance attribuée à certains détails du texte original, par exemple :

<i>M&M_ru</i> (SOURCE)	<i>M&M_en69</i> (CIBLE, 1969)	<i>M&M_en97</i> (CIBLE, 1997)
отвратительные, тревожные желтые цветы	repulsive yellow flowers	repulsive, alarming yellow flowers
черт их знает, как их зовут	god knows what they're called	devil knows what they're called
нехороший цвет	it's an ugly colour	not a nice colour
по кривому, скучному переулку	down that dreary, winding little street	along the crooked, boring lane
повинуясь этому желтому знаку	obeying that yellow signal	obeying this yellow sign
я мучился	I was in agony	I was suffering
она уйдет	she would disappear	she would leave
и, вообразите, внезапно заговорила она :	then, if you can believe it, she said:	and, imagine, suddenly she began to speak:

Pour explorer l'univers de l'œuvre originale et analyser les choix des traducteurs, il est utile d'avoir recours à des méthodes textométriques de navigation textuelle. Ces méthodes ont déjà été utilisées pour l'extraction de ressources traductionnelles à partir de corpus de textes rédigés dans des langues différentes (Zimina, 2005a ; 2005b). Nous tentons, dans ce qui suit d'élargir leur champ d'application à des corpus plus complexes, à la fois parallèles et comparables, dont les différents volets autorisent des comparaisons type *source* – *cible* (correspondances bilingues) et *version 1* – *version n* (correspondances unilingues).

3. Exploration cartographique des variantes de traduction sous *mkAlign*

3.1. Rappel sur l'application *mkAlign*

L'application *mkAlign* permet de construire, de corriger et de visualiser un alignement de deux textes via un éditeur à double entrée (Fleury et Zimina, 2007a ; 2007b). Il permet la prise en compte de plusieurs encodages des fichiers textes, comme en témoignent des explorations textométriques récentes menées au sein de l'équipe non seulement sur des langues indo-européennes mais aussi sur des langues sino-tibétaines dotées d'un système d'écriture idéographique.³

L'outil intègre également plusieurs fonctionnalités qui facilitent la mise en correspondance des corpus : segmentation automatique en « caractères » (type kanji, par exemple), en formes graphiques, en phrases ou en paragraphes, module d'alignement par mots apparentés, module d'édition manuelle du bi-texte généré au cours de l'alignement automatique, etc. C'est l'utilisateur qui supervise l'ensemble des processus liés à la gestion des corpus, depuis la mise en correspondance initiale des segments équivalents jusqu'à l'export final du bi-texte produit. Il appartient à l'utilisateur de construire un alignement et de définir son degré de précision (résolution). Cette résolution peut être variée pour mettre en évidence les correspondances entre les segments textuels qui se situent à des niveaux différents : phrases, paragraphes, sections, parties, etc. (Véronis, 2000).

Sous *mkAlign*, l'analyse du corpus peut être cyclique : importation des données, exploration quantitative, export des résultats (ou création des sous-corpus autour des phénomènes ciblés) éventuellement suivies de la réimportation des données textuelles et d'une nouvelle

exploration. La notion de sauvegarde des sessions de travail (création des fichiers d'export/import de bi-textes aux formats TXT, XML et HTML) permet de commencer le travail d'exploration sur un corpus à deux volets textuels, de l'exporter au format désiré, puis de le réimporter ensuite pour y apporter des modifications.

3.2. Cartographie de la traduction

En plus des outils de gestion de corpus à deux volets, *mkAlign* permet de cartographier leur alignement.

Sur la figure 4, nous avons montré une représentation cartographique d'un extrait du corpus *M&M* aligné au niveau des paragraphes à l'aide de *mkAlign*. Lorsque l'on sélectionne l'un des carrés de la carte, le texte du paragraphe correspondant s'affiche en bas de l'écran ainsi que la section équivalente affichée, elle, dans l'autre volet du corpus (sélection parallèle). L'extrait que nous avons représenté correspond au chapitre 13 du roman. Ce chapitre a retenu notre attention en raison de sa position charnière au sein de l'œuvre, comme le montrent les résultats de l'analyse factorielle des correspondances (cf. figure 1). Les traits horizontaux découpent les cartes en deux parties : la première montre le texte-source en russe et sa première traduction en anglais (source *M&M_ru* ~ cible *M&M_en69*) ; la deuxième montre ce même texte-source et sa traduction plus récente (source *M&M_ru* ~ cible *M&M_en97*). La carte est symétrique : chaque couple des carrés sélectionnés correspond à deux fragments textuels alignés (paragraphe source / paragraphe cible).

La carte des sections alignées constitue un outil d'analyse puissant pour l'exploration des variantes de traduction grâce à ses trois fonctions principales :

- Affichage de la distribution de plusieurs types d'unités textuelles (formes, segments, unités définis à l'aide des expressions régulières, patrons morphosyntaxiques, etc.) au sein des sections alignées.
- Sélection automatique des fragments textuels correspondant à l'intersection ou à la différence constatée dans les distributions lexicales affichées sur les deux volets de la carte.
- Application des méthodes statistiques de textométrie (type analyse des spécificités lexicales) aux fragments textuels sélectionnés dans les deux volets du corpus [Zimina, 2004 ; 2005b].

Pour illustrer ces principes, nous avons affiché la distribution de l'unité visée sur le volet russe de la carte (cf. figures 4-7). Dans la section suivante nous montrerons la mise en pratique de l'exploration en utilisant des exemples concrets.

3.3. Visualisation des correspondances lexicales : intersection et différence

Les méthodes de navigation textuelle dans les corpus parallèles ont déjà fait l'objet d'une série de travaux consacrés à l'extraction de ressources traductionnelles à base de corpus [Zimina 2004, 2005, 2006] ; [Fleury et Zimina, 2007]. Cette approche peut être généralisée à des corpus des variantes de traduction. Nous résumons ici les principales étapes d'une telle exploration.

Première étape :

Sur la figure 4, nous avons affiché la distribution de l'unité *zocm+* (fr. hôte) qui couvre toutes les formes fléchies du substantif russe *zocmь* dans le volet source du corpus. Les paragraphes de la version originale où cette forme est attestée au moins une fois sont marqués par une coche. Les sections correspondant aux traductions dans la langue cible sont mises au jour de l'autre côté de la carte dans les deux textes cibles.

Le premier volet correspond au texte en russe et le deuxième au texte en anglais :

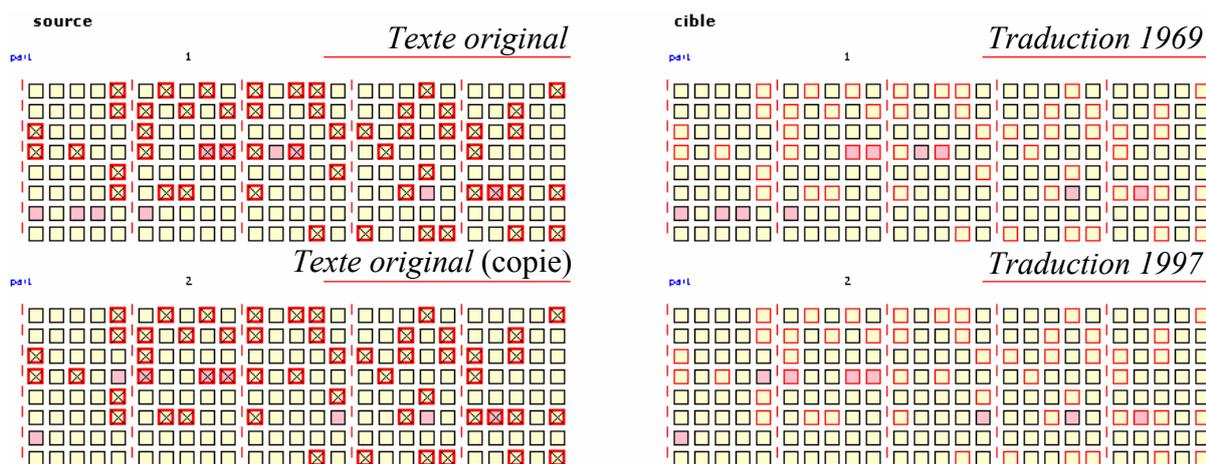


Figure 4 : Extrait de la carte du corpus *M&M* qui affiche la ventilation de l'unité *zocm+* 

Guide de lecture de la figure 4 :

Source  = Cible  : présence de l'unité recherchée / section parallèle sélectionnée ;

Source  = Cible  : appariement avec une section vide.

A cette première étape de l'exploration, le calcul des spécificités appliqué aux sections de la cible, met en évidence deux principales variantes de traductions de l'unité russe *zocm+* en anglais : la forme *guest* est majoritairement utilisée dans la traduction de 1997 alors que c'est la forme *visitor* qui est typique de la traduction de 1969.

Deuxième étape :

La ventilation de *zocm+* est toujours affichée dans le texte source. Les deux traductions mises en évidence par le calcul des spécificités au cours de l'étape précédente sont regroupées au sein d'une seule unité textuelle (à l'aide des expressions régulières disponibles sous *mkAlign*).³

La ventilation de cette nouvelle unité est représentée sur le volet cible de la carte. Comme c'est indiqué sur la figure 5, en écartant les sections sélectionnées parallèlement sur les deux volets de la carte (intersection Source  = Cible ), on parvient à matérialiser la différence qui existe entre les distributions source/cibles .

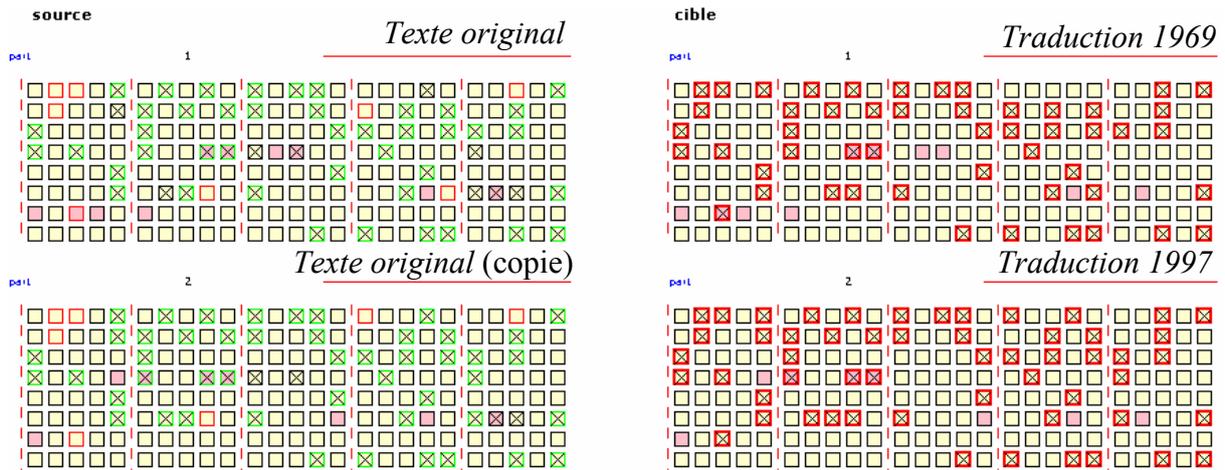


Figure 5 : Extrait de la carte du corpus M&M qui affiche la différence □ dans les ventilations des unités *zocm+* ⊗ et *visitor|guest* ⊗

Ces contextes singuliers qui ne comportent pas *zocm+* mais qui ont quand même été traduits par *guest* ou *visitor* peuvent être analysés en termes des spécificités lexicales. Cette opération permet de découvrir des synonymes de *zocm+* dans le texte russe : *посетитель* (fr. visiteur), *пришедший* (fr. celui qui est venu), etc.

La différence peut aussi être exportée dans un tableau au format XML ou HTML (cf. figure 6). Ce type d'export permet une sélection ciblée des contextes (ou vecteurs contextuels) décrits en termes d'intersection/différence des distributions lexicales. Sur la figure 6, nous avons indiqué par la couleur jaune les équivalences lexicales mises en évidence grâce à ce type d'export :

MKA-ed	SOURCE	CIBLE
CIBLE Seulement		
2	убедившись в том, что иван один, и прислушавшись, таинственный посетитель осмелел и вошел в комнату. тут увидел иван, что пришедший одет в больничное. на нем было белье, туфли на босу ногу, на плечи наброшен бурый халат. §	the mysterious visitor listened awhile then, satisfied that ivan was alone, entered the room. as he came in ivan noticed that the man was wearing hospital clothes--pyjamas, slippers and a reddish-brown dressing gown thrown over his shoulders. §
3	пришедший подмигнул ивану, спрятал в карман связку ключей, шепотом осведомился: "можно присесть?" -- и, получив утвердительный кивок, поместился в кресле. §	the visitor winked at ivan, put a bunch of keys into his pocket and asked in a whisper : ' may i sit down?' receiving an affirmative reply he settled in the armchair. §
23	-- никаких я ваших стихов не читал! -- нервно воскликнул посетитель. §	'i've never read any of your poetry!' said the visitor tetchily. §
27	-- не пишите больше! -- попросил пришедший умоляюще. §	'don't write any more!' said the visitor imploringly. §
41	описание ужасной смерти берлиоза слушающий сопроводил загадочным замечанием, причем глаза его вспыхнули злобой: §	listening to the terrible description of berlioz's death, the visitor made an enigmatic comment, his eyes flashing with malice : §

Figure 6 : mkAlign : résultats de l'export de la différence source/cible (extrait de quelques contextes de M&M_en69)

Troisième étape :

La visualisation simultanée de la distribution de l'unité *guest|visitor* (*guest* ou *visitor*) au sein des deux textes cibles permet de localiser les discordances qui existent entre les traductions.

Pour le faire, nous nous intéresserons aux sections des textes cibles où l'une des formes *guest* ou *visitor* est utilisée dans une des traductions seulement (la deuxième traduction ayant une autre variante). Nous mettons ainsi en évidence la variation suivante : « his guest and his secret wife » (*M&M_en69*) / « the man and his mistress » (*M&M_en97*) :

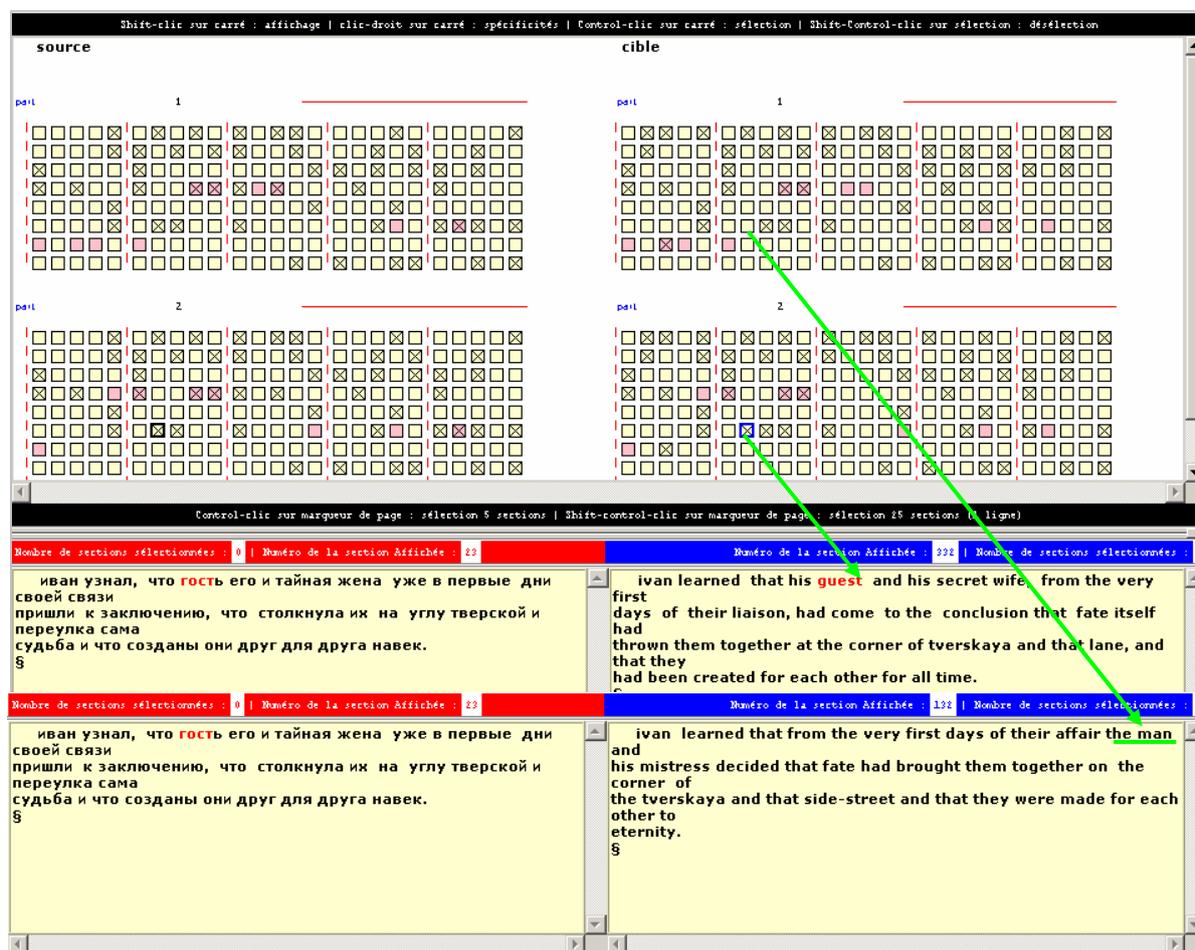


Figure 7 : extrait de la carte du corpus *M&M* sous *mkAlign* : variantes de traduction

Comme nous l'avons montré au cours de cette expérience, la cartographie de la traduction simplifie les accès contextuels tout en respectant l'objet d'étude qui se construit au fur à mesure de l'exploration en fonction des critères définis par le chercheur. L'analyse de l'espace intertextuel est itérative et permet une véritable navigation dans le bi-texte (Zimina-Poirot, 2005 – sous presse).

4. Conclusion et perspectives

Nous avons proposé une approche de l'analyse de la traduction assistée par ordinateur qui permet de comparer le texte original avec ses variantes de traduction en soulignant les convergences/divergences des versions. Elle peut présenter de réels avantages pour le travail de traduction philologique.

Notre approche s'appuie sur les principes de navigation textométrique en corpus, renforcée par des aides visuelles (Zimina, 2002 ; 2004 ; 2005c). L'application *mkAlign* utilisée au cours de nos expériences permet la comparaison simultanée des fragments textuels, écrits dans une même langue, ou dans des langues différentes (Fleury et Zimina, 2007ab).

Les expériences entreprises sur le corpus du roman « Le Maître et Marguerite » et ses deux traductions anglaises seront développées et enrichies en prenant en compte la traduction française du roman. En nous intéressant à la problématique de la traduction littéraire, nous nous fixons deux objectifs connexes : 1) appréhender l'ensemble des phénomènes complexes de la traduction littéraire ; 2) mettre au point de nouveaux modules d'exploration susceptibles de compléter la boîte à outils du chercheur s'intéressant à l'exploration et/ou à la comparaison des textes traduits.

Ces recherches ouvrent la voie à de nombreuses pistes d'exploration des corpus de traductions prenant en compte, notamment, les résultats des expériences sur la cartographie des proximités segmentales développés dans les travaux des membres de l'équipe SYLED-CLA2T (Salem, 2006) ; (Fleury, 2007).⁴

Références

- Akhmanova O., Zadornova V. (1972). On Linguopoetic Stratification of Literary Texts. *Poetica: An International Journal of Linguistic-Literary Studies*, Vol.7: 50-60.
- Akhmanova O., Zadornova V. (1977a). Où en est la linguo-poétique ? *Estudios ofrecidos a Emilio Alarcos Llorach*, Vol.1: 249-260.
- Akhmanova O., Zadornova V. (1977b). The Philology of Translation. *Shakespeare Translation*, Vol.4.
- Belleau A. (1970). Bakhtine et le multiple. *Études françaises*, Vol.6(4).
- Fleury S., Zimina M. (2007a). MkAlign, Manuel d'utilisation. EA2290 SYLED/CLA2T, Paris 3. Disponible en ligne : <http://tal.univ-paris3.fr/mkAlign/mkAlignDOC.htm>.
- Fleury S., Zimina M. (2007b). Exploring Translation Corpora with *mkAlign*. *Translation Journal*, vol.11(1). Disponible en ligne : <http://accurapid.com/journal/39mk.htm>.
- Lamalle C. et al. (2003). Lexico3 – Outils de statistique textuelle. EA2290 SYLED/CLA2T, Paris 3. Disponible en ligne : <http://www.cavi.univ-paris3.fr/Ilpga/ilpga/tal/lexicoWWW/manuel3.htm>.
- Salem A. (2006). Proximités segmentales. *Actes des JADT'06 (Journées internationales d'Analyse statistique des Données Textuelles)*, pages 843-854.
- Véronis J. (Ed.). (2000). *Parallel Text Processing: Alignment and use of translation corpora*. Dordrecht, Kluwer Academic Publishers.
- Zimina M. (1996). Philological Translation of William Blake's Poem "The Tyger". In *Proceedings of Lomonosov-96 (Undergraduate and Postgraduate Student International Conference on Fundamental Sciences)*, page 137.
- Zimina M. (2002). Repérages lexicométriques des équivalences à basse fréquence dans les corpus bilingues. *Lexicometrica: Corpus alignés*. Disponible en ligne : <http://www.cavi.univ-paris3.fr/lexicometrica/thema/thema6/Zimina.pdf>.
- Zimina M. (2004). L'alignement textométrique des unités lexicales à correspondances multiples dans les corpus parallèles. *Actes des JADT'04 (Journées internationales d'Analyse statistique des Données Textuelles)*, pages 1195-1202.
- Zimina M. (2005a). Exploration textométrique de corpus de traduction. *META*, Vol.50 (cd-rom). Disponible en ligne : <http://www.erudit.org/livre/meta/2005/000264co.pdf>.
- Zimina M. (2005b) Bi-text Topography and Quantitative Approaches of Parallel Text Processing. In *Proceedings of CL'05 (Corpus Linguistics Conference)*. Disponible en ligne : <http://www.corpus.bham.ac.uk/PCLC/>.
- Zimina M. (2005c). Topographie bi-textuelle et approches quantitatives de l'extraction de ressources traductionnelles à partir de corpus parallèles. *Actes des LTT'05 (Journées scientifiques du Réseau*

de chercheurs « *Lexicologie, Terminologie, Traduction* »), pages 175-186. Disponible en ligne : <http://perso.univ-lyon2.fr/~thoiron/JS%20LTT%202005/pdf/Zimina.pdf>.

Zimina-Poirot M. (2005 – sous presse). Corpus multilingues: exploration textométrique de l'espace intertextuel. *Actes des journées d'études sur les corpus, C.E.R.T.A. (Centre d'Etudes et de Recherches en Traductologie de l'Artois)*, pages 107-122.

Corpus

Nous remercions Olga Semenova (M2 Professionnel, CRIM/INaLCO, 2006-07) qui a localisé sur le Web les trois volets du corpus **M&M** :

Булгаков М. (1967-1984). Мастер и Маргарита [texte du manuscrit original]: <http://www.lib.ru/BULGAKOW/master.txt>

Glenny M. (1967). *The Master and Margarita*. Collins and Harvill Press, London: http://www.lib.ru/BULGAKOW/master_engl.txt

Pevear R., Volokhonsky L. (1997). *The Master and Margarita*. Penguin books. http://www.lib.ru/BULGAKOW/master97_engl.txt

Notes :

¹ Textométrie Multilingue : <http://www.cavi.univ-paris3.fr/lexicometrica/jadt/textometrie-multilingue/>

² La traduction française du roman « Le Maître et Marguerite » est disponible sur le web : http://www.ebooksgratuits.org/html/boulgakov_maitre_et_marguerite.html

³ *MkAlign* : <http://tal.univ-paris3.fr/mkAlign/>.

Documentation : <http://tal.univ-paris3.fr/mkAlign/mkAlignDOC.pdf>

⁴ Le projet « Projection lexicométrique » [Fleury, 2007] est également consacré à l'étude des proximités segmentales : <http://www.cavi.univ-paris3.fr/ilpga/ilpga/sfleury/WIP/variation-sroyal-11022007/>.