

**Cotutelle de thèse internationale entre
l'Université de Liège
et l'Université Grenoble Alpes**

Évaluation experte d'un prototype d'aide à la traduction créative
La traduction littéraire automatique individualisée
au regard de ses enjeux traductologiques, éthiques et sociétaux

Présentée par Damien Hansen en vue de l'obtention du titre

de Docteur en Langues, Lettres et Traductologie
à l'Université de Liège

et de Docteur en Mathématiques, Sciences et Technologies de l'Information, Informatique
à l'Université Grenoble Alpes

Thèse soutenue publiquement le 28 mars 2024, devant le jury composé de :

Julien Perrez

Université de Liège, CIRTI, Président

Natalie Kübler

Université Paris Cité, CLILLAC-ARP, Rapportrice

Mathieu Lafourcade

Université de Montpellier, LIRMM, Rapporteur

Valérie Bada

Université de Liège, CIRTI, Co-Directrice de thèse, Secrétaire

Hervé Blanchon

Université Grenoble Alpes, LIG, Co-Directeur de thèse

Caroline Rossi

Université Grenoble Alpes, ILCEA4, Examinatrice

Recherche menée sous la direction de

Valérie Bada

Université de Liège

Hervé Blanchon

Université Grenoble Alpes

Emmanuelle Esperança-Rodier

Université Grenoble Alpes

Au sein du **Centre Interdisciplinaire de Recherche en Traduction et en Interprétation (CIRTI)**
dans l'École Doctorale **ED3bis**



et du **Laboratoire d'Informatique de Grenoble (LIG)**
dans l'École Doctorale **EDMSTII**



Any sufficiently advanced technology is indistinguishable from magic.

ARTHUR C. CLARKE

Table des matières

I	Le retour de l'objection préjudicielle	1
1	Introduction	3
2	Traduction littéraire assistée par ordinateur	11
2.1	Un vaste domaine de recherche	11
2.2	La résurgence de l'objection préjudicielle	12
2.3	S'approprier les technologies d'aide à la traduction	15
3	Traduction littéraire automatique	19
3.1	Entre optimisme débordant et rejet catégorique	20
3.2	L'hypermédiatisation de l'IA	21
3.3	Adopter une approche raisonnée face aux nouvelles technologies	27
II	Une question de données avant tout	33
4	État de l'art	35
4.1	Les prémices de la TA : l'intérêt précurseur des littéraires	37
4.2	La traduction automatique statistique	40
4.2.1	Les premiers travaux	44
4.3	Le paradigme neuronal	47
4.3.1	Architectures	52
4.3.2	Mise au point de systèmes adaptés à la littérature	57
4.3.3	Évaluation des outils tout public	65
4.3.4	Aspects éthiques et théoriques	67
4.4	Aller plus loin : des systèmes adaptés au niveau de l'humain	68
5	Méthodologie	71
5.1	Cadre épistémologique	71
5.2	Cas d'étude : la saga <i>Septimus Heap / Magyk</i>	75
5.2.1	Choix de l'ouvrage	75
5.2.2	La triple condamnation d'une traduction littéraire jeunesse de fiction	76
5.2.3	Peut-on mesurer la complexité des textes littéraires (pour une machine)?	80
5.3	Données et système	85
5.3.1	Préparation des données	85
5.3.2	Mise au point du système	90
5.3.3	L'importance des corpus	93
5.3.4	Résultats négatifs	97

6	Évaluation	103
6.1	Évaluation automatique	104
6.1.1	La question épineuse des métriques d'évaluation	104
6.1.2	Métriques d'évaluation automatiques	107
6.1.3	Richesse lexicale	119
6.1.4	Longueur des segments	121
6.1.5	Équivalence syntaxique	122
6.2	Évaluation humaine	125
6.2.1	Annotation d'erreurs : une typologie sur mesure	126
6.2.2	Peu d'erreurs littéraires, mais une fluidité trompeuse	131
6.2.3	Observations générales	139
6.3	Évaluation experte : examiner l'outil en contexte réel	144
III	Adapter la machine à l'humain	147
7	Entretiens	149
7.1	Une évaluation experte en contexte réel	151
7.1.1	Nathalie Serval : figure d'une traductrice	154
7.1.2	Déroulement des entretiens	155
7.2	Première phase : discussion préliminaire	162
7.2.1	Définition de la tâche	163
7.2.2	Impressions initiales	164
7.3	Deuxième phase : première interaction avec l'outil	167
7.3.1	Un assistant potentiellement utile	170
7.3.2	Une interface lourde et contraignante	173
7.4	Troisième phase : modifications ergonomiques et rétrospective	178
7.4.1	Imaginer d'autres scénarios	180
7.4.2	Améliorer ses conditions de travail	184
7.5	Redéfinir l'utilité des technologies de la traduction ?	186
8	Discussion	189
8.1	La traduction littéraire automatique aujourd'hui : usages et enjeux	190
8.1.1	Premiers regards sur les usages actuels de la TA en littérature	192
8.1.2	Avantages et bénéficiaires	194
8.1.3	Risques et enjeux sociétaux	196
8.2	La traduction littéraire automatique demain : défis et perspectives	203
8.2.1	Les coûts cachés de l'IA	203
8.2.2	La nécessité d'un cadre juridique	210
8.2.3	Des enjeux dépassant le seul contexte de la traduction	215
8.3	La traduction littéraire automatique : un changement de paradigme	218
9	Conclusion	227
A	Affiches	239
B	Corpus synthétique	241
C	Corpus parallèle	249
D	Corpus personnalisé	253

E	Architecture des systèmes utilisés	255
F	Évaluations de la TLA	257
G	Typologie d'erreurs	259
H	Retranscription	267
	Bibliographie	285
	Programmes	285
	Outils et corpus	285
	Pages Web et articles de presse	291
	Références	294
	Notes	336

Liste des tableaux

1	Taille des corpus génériques et du corpus spécialisé	88
2	Optimisation des hyperparamètres	91
3	Amélioration du système par étapes successives	92
4	Taille des corpus après ajout de nouvelles données littéraires	95
5	Contributions des corpus à l'amélioration du système	96
6	Améliorations non concluantes	98
7	Comparaison des métriques pour différents modèles adaptés et génériques	111
8	Distance donnée par le Delta de Burrows par rapport à la référence humaine	113
9	Comparaison avec d'autres adaptations de systèmes génériques à la littérature	116
10	Mesures de richesse lexicale	120
11	Longueur moyenne des segments et coefficient de foisonnement	121
12	Mesures d'équivalence syntaxique	124
13	Relevé des erreurs produites par la TA pour l'adéquation, la fluidité et la littéarité	132
14	Méthode de conduite des entretiens et de récolte des données	158
15	Paramètres des architectures <i>Transformer</i>	255
16	Paramètres des architectures LSTM	256
17	Évaluations humaines de la TLA	257
18	Évaluations automatiques de la TLA	258

Liste des figures

1	<i>TraduXio</i> : plateforme de translocalisation culturelle collaborative	16
2	<i>PunCAT</i> : prototype d'aide à la traduction de jeux de mots	17
3	Estimation des cycles d'engouement pour la TA	24
4	Automate Jaquet-Droz : l'écrivain	26
5	Exemple fictif d'un exercice de génération de texte	29
6	Triangle de Vauquois	37
7	Architecture du système <i>Apertium</i>	38
8	Exemple de tokenisation	42
9	Exemple d'alignement	42
10	Modèle de traduction	43
11	Suites de n-grammes	43
12	Évolution des approches connexionniste et symbolique	48
13	Représentation simple d'un neurone artificiel	48
14	Fonctionnement d'un réseau de neurones simple	49
15	Représentation vectorielle	51
16	Modèle encodeur-décodeur	53
17	Déroulement du modèle encodeur-décodeur dans le temps	53
18	Cellule LSTM	54
19	Architecture <i>Transformer</i>	56
20	Perplexité donnée par l'alignement de romans et de leur traduction	81
21	Degré d'entropie donnée par la modélisation de divers romans	82
22	Traduction automatique de romans par des systèmes libres d'accès	83

23	Exemple de tokenisation en sous-mots	89
24	Schématisation de la recherche en faisceau	92
25	Projection de la distance par analyse en composantes principales	114
26	Exemple d’alignement issu de SACr	124
27	Typologie d’erreurs	127
28	Interface de travail dans <i>ACCOLÉ</i>	130
29	Intégration TA/TAO grâce au module MTUOC	168
30	Interface de travail dans <i>Trados</i>	169
31	Interface de travail dans <i>OmegaT</i>	182
32	Évolution de la puissance de calcul associée au <i>deep learning</i>	204
33	Comparaison de la taille de modèles de langue (en millions de paramètres)	206
34	Continuum de la traduction assistée par ordinateur	219
35	Prospectus de conférence de l’ALAMO à Liège (1964) et à Verviers (1969)	239

Liste des abréviations

GML	Grand modèle de langue
LLM	<i>Large language model</i>
IA	Intelligence artificielle
AI	<i>Artificial intelligence</i>
MT	Mémoire de traduction
TM	<i>Translation Memory</i>
PE	Post-édition
PE	<i>Post-editing</i>
TA	Traduction automatique
MT	<i>Machine translation</i>
TABR	Traduction automatique à base de règles
RBMT	<i>Rule-based machine translation</i>
TAL	Traitement automatique des langues
NLP	<i>Natural language processing</i>
TAN	Traduction automatique neuronale
NMT	<i>Neural machine translation</i>
TAO	Traduction assistée par ordinateur
CAT	<i>Computer-assisted translation</i>
TAS	Traduction automatique statistique
SMT	<i>Statistical machine translation</i>
TH	Traduction humaine
HT	<i>Human translation</i>
TLA	Traduction littéraire automatique
LMT	<i>Literary machine translation</i>
TLAO	Traduction littéraire assistée par ordinateur
CALT	<i>Computer-assisted literary translation</i>

Remerciements

Si j'ai fini par oser rédiger ce travail à la première personne du singulier, c'est évidemment sans oublier qu'il est le fruit de collaborations, de discussions et de rencontres avec une multitude de collègues et de proches sans qui ce travail n'aurait pu voir le jour et ce parcours n'aurait été aussi enrichissant et formateur. À commencer par mes deux promotrices et mon promoteur de thèse, qui m'ont accompagné dans trois champs disciplinaires non moins complémentaires que pointus, qui m'ont encouragé dans tous les choix méthodologiques et personnels que j'ai pu effectuer durant ces quatre années, et qui ont su faire dialoguer avec des remarques toujours illuminatrices leur expertise dans ce travail final, me permettant d'y proposer une étude originale et d'explorer tous les aspects que j'espérais y aborder en entamant cette recherche :

Valérie Bada, pour m'avoir fait confiance dès le début dans ces projets un peu fous, pour avoir nourri mon goût pour la recherche et pour m'avoir toujours appuyé dans cette longue aventure, pour son optimisme devant les refus initiaux auquel ce sujet a pu donner lieu, mais surtout pour sa présence humaine tout au long de la route.

Emmanuelle Esperança-Rodier, pour avoir accepté de superviser la thèse d'un étudiant surgi de nulle part, pour nos nombreuses réunions formelles et informelles, et pour avoir grandement facilité, outre la poursuite du doctorat, mon arrivée et mon séjour dans un pays étranger en début de pandémie mondiale.

Hervé Blanchon, pour avoir accepté de reprendre la direction de la thèse, pour sa patience infinie devant l'arrivée tardive de systèmes fonctionnels, y compris lorsque je me lançais aveuglément dans une voie sans issue, et pour son accompagnement inestimable, sans lequel il m'aurait fallu rapidement réorienter cette recherche.

J'aimerais également remercier **Laurent Besacier**, dont le travail a inspiré ce projet de recherche, pour m'avoir aidé à me lancer dans cette thèse, tant sur le plan administratif que scientifique, et pour avoir permis de former ce pont entre l'équipe du CIRTI et l'équipe GETALP, sans lequel je n'aurais pu espérer mener à bien ce travail.

Le **bureau 325**, aka le bureau bouffe, avec qui j'ai entamé cette aventure à Grenoble. Pour l'accueil et les sorties, pour toutes ces pauses cafés — plus jamais je ne prendrai le café gratuit pour acquis après avoir quitté le LIG — et pour les innombrables fois où j'ai dû vous déranger avec mes problèmes de script et mes expériences foireuses. Je me réjouis de vous recroiser, à l'occasion.

Mon autre collègue de bureau à Liège, **Perrine Schumacher**, pour m'avoir accompagné dans toutes ces journées de travail ou, à défaut, dans celles où le manque de motivation laissait place à des discussions tout aussi enrichissantes. Pour les cafés-spéculoos et les après-journées, le soutien mutuel et les coups de gueule.

De manière plus générale, l'ensemble des **collègues du CIRTI et du GETALP**, pour avoir rendu cette aventure particulièrement agréable autant sur le plan scientifique que sur le plan humain, notamment par la liberté offerte par ces deux environnements et par leur bienveillance, qui ont fait de ce parcours quelque chose dont je pouvais me réjouir sincèrement chaque jour.

L'ensemble du **Liège Game Lab**, pour m'avoir permis de poursuivre mon travail sur le jeu vidéo, de me sortir la tête de la thèse de temps à autre et d'explorer de nouvelles pistes de recherche telles que la traduction automatique de jeux vidéo, mais aussi pour avoir affronté la pandémie avec un florilège de streams, d'apéros ou cafés-thèse virtuels et d'autres initiatives qui ont rendu cette période bien plus facile à supporter, tous comme les années qui ont suivi.

Pierre-Yves Houlmont, pour sa présence, son optimisme et ses encouragements depuis le master jusqu'à aujourd'hui. Pour les temps de midi, les voyages et les projets partagés qui se poursuivront d'une manière ou d'une autre, je l'espère, après la thèse.

Pierre-Yves Hurel, pour son expertise et ses conseils avisés de meneur d'entretiens, ainsi que pour les pistes de recherche et de réflexion que nous avons pu dégager et qui ont encouragé le changement de paradigme opéré en fin de parcours.

Dorothy Kenny, pour son invitation à Dublin, pour nos discussions précieuses au sujet de la thèse et pour ces idées de projets qui continueront probablement de se développer après celle-ci, de même que les **collègues de SALIS**, pour leur accueil bienveillant lors de mon séjour à DCU.
Go raibh míle maith agat!

Antoni Oliver, pour les discussions lancées sur la traduction littéraire automatique et pour son aide plus que précieuse dans paramétrage des scripts du projet MTUOC au moment de sa refonte, dont dépendait l'intégration du système de TA aux outils de TAO qui a permis de mener cette évaluation experte en contexte réel.

Évidemment, un remerciement tout particulier à **Nathalie Serval**, qui avait déjà accepté de consacrer de son temps aux interrogations un peu naïves et maladroitement d'un étudiant de master, et qui s'est relancée avec la plus grande curiosité et le plus grand enthousiasme dans ce projet. Ce fut un avantage inestimable de pouvoir compter sur la contribution d'une experte et d'inscrire cette démarche quelque peu expérimentale dans le concret, mais un aussi un privilège d'avoir pu travailler avec la personne chargée des traductions à partir desquelles j'ai pris tant de plaisir à mener cette recherche.

De la même manière, il me faudrait remercier une liste interminable de personnes qui ont rendu ce projet possible et m'ont permis d'en arriver jusqu'à la publication de ce travail. Les personnes qui ont tôt vu l'intérêt de ce projet et m'ont encouragé à poursuivre ces recherches, avant l'obtention d'un financement lorsqu'il paraissait encore un peu fou et irréalisable ou bien au fur et à mesure du parcours, mais aussi toutes celles que j'ai eu la chance de rencontrer et qui ont pu nourrir, d'une manière ou d'une autre, ce travail, que ce soit par le partage d'expériences et de réflexions personnelles, d'outils, de sources ou de pistes de recherche. Les collègues qui, à l'occasion de conférences ou d'autres déplacements, ont pu rendre cette aventure plus enrichissante ou simplement bien plus plaisante. Ma famille, naturellement, mais aussi toutes celles et tous ceux qui m'ont accompagné ou m'ont rejoint dans cette aventure, en particulier mes camarades de jeu, de sortie, de voyage et d'infortune ; ces personnes se reconnaîtront sans difficulté et je ne les remercierai jamais assez pour leur présence. Et enfin ma mère, qui m'a toujours soutenu dans tous les choix que j'ai pu faire.

Note liminaire

Conventions de rédaction

Si le français sera privilégié dans l'ensemble du manuscrit, celui-ci contiendra également de multiples références à l'anglais, et ce, afin d'en simplifier au maximum la lecture étant donné que certains des termes utilisés sont encore peu fréquents, y compris dans les ouvrages francophones. L'objectif poursuivi ici est de participer à la diffusion de ces traductions, mais aussi de faire le parallèle avec la littérature majoritairement anglophone sur le sujet et de faciliter l'utilisation — pour celles et ceux qui s'attelleraient à la tâche — de ces outils dont le fonctionnement et la documentation reposent entièrement sur l'anglais. La plupart de ces termes proviennent du *Grand Lexique français de l'intelligence artificielle*¹, sinon d'ouvrages scientifiques écrits en français.

Dans le même ordre d'idées, toutes les citations en langue étrangère seront traduites en français. Ces traductions sont de mon cru, à moins qu'elles ne soient tirées de traductions existantes, tandis que les citations en langue source seront quant à elles compilées dans les notes de fin d'ouvrage (cf. p. 337).

Afin de faciliter la lecture de ce travail et d'éviter l'usage du point médian ou des doublets sans pour autant opter pour le masculin générique, j'ai fait le choix d'utiliser tout au long de cet ouvrage des formes neutres et épécènes, en favorisant notamment des formes nouvelles telles que *traductaire*, de plus en plus courante dans les textes et communications scientifiques. Il se peut que cette forme soit passée de mode après un an déjà, qu'elle soit remplacée par un autre terme ou qu'une autre méthode d'écriture inclusive s'impose par la suite. En l'absence toutefois d'un consensus dans la communauté scientifique francophone, ce choix tient à la volonté de n'invisibiliser personne, mais aussi de participer à la réflexion sur le lien entre le genre et la langue, en particulier dans un contexte où la majorité des *traducteurs* sont en réalité des femmes.

Traduire la révolution féministe, c'est aussi interpréter la nouveauté du mot et du sens. La richesse de la littérature et de la théorie féministes nous oblige à bousculer les termes, les règles, les habitudes de compréhension et d'écriture, à employer des mots qui n'existent pas encore. (GRUNENWALD 2021, p. 52.)

Enfin, j'ai préféré éviter l'utilisation du terme *biotraduction*. Ce néologisme, tiré du roman de science-fiction de Jean-Louis Trudel² (FROELIGER 2013, p. 20), renforce en effet l'idée d'un antagonisme entre humain et machine, bio- et cyber-traduction (DIRAND et ROSSI 2019, p. 75). Pour éviter cette opposition indésirable et refléter la réalité complexe du métier, qui impose bien souvent de naviguer entre différents domaines, rôles et outils, j'emploierai donc systématiquement le terme *traduction*, éventuellement qualifiée d'*humaine* lorsqu'il faudra comparer cette modalité et l'opposer expressément aux sorties des outils de traduction automatique.

1. <https://datafranca.org/wiki/Accueil>.

2. *Le revenant de Fomalhaut*, MÉDIASPAUL, 2002.

Matériel connexe

Ce travail présente un projet de recherche et des résultats qui ont été publiés en partie dans d'autres ouvrages :

- l'état de l'art dans les *Actes de la 28^e Conférence sur le Traitement Automatique des Langues Naturelles* (HANSEN 2021) ;
- les enjeux et la méthodologie dans le numéro spécial « Vers une robotique du traduire ? » du *Journal of Data Mining & Digital Humanities* (HANSEN *et al.* 2022) ; et
- l'évaluation dans les actes du colloque *New Trends in Translation and Technology* (HANSEN et ESPERANÇA-RODIER 2023).

Bien qu'ils confirment les tendances et conclusions dégagées dans ces articles, les résultats mis en avant dans ce travail pourront parfois varier. Le projet, étendu sur quatre années, aura en effet vu se succéder trois versions différentes d'*OpenNMT*, l'outil sur lequel il se base en grande partie, auxquelles s'ajoute un cheminement personnel d'apprentissage et d'appropriation de la technologie qui a permis d'améliorer petit à petit les performances du système mis au point. J'ai donc choisi de reproduire les expériences menées avec la troisième et dernière version à ce jour de l'outil, à la fois pour faciliter la comparaison des résultats et en obtenir une seconde confirmation. Ce faisant, il aura en outre été possible d'estimer l'empreinte carbone de tels entraînements, sujet qui prend heureusement de plus en plus d'importance dans les débats sur le numérique et que j'aborderai dans les discussions de fin d'ouvrage.

À titre d'information, ce projet a également pu être présenté, à divers stades de son avancement, dans les communications suivantes :

- lors d'une présentation à la *CIUTI Conference 2020: Artificial Intelligence & Intercultural Intelligence*, intitulée « Artificial intelligence and machine learning: How a better grasp of AI can frame and facilitate the dialogue about its role in the years to come » ;
- lors d'une présentation à la *Computer-Assisted Literary Translation (CALT) Conference*, intitulée « The figure of the literary translator amid new technologies » ;
- lors d'une présentation au *Séminaire TRACT 2021–2022*, avec Susan Pickford, intitulée « Traduire la littérature populaire avec l'aide de la machine: quels avantages et préoccupations pour les professionnel·le·s? » ;
- lors d'une présentation au *NeTTT Workshop on Translation Technology for Creative Domains*, intitulée « Personalized MT Systems for Literary Translators » ;
- lors d'une présentation au colloque *Literary Machine Translation as a Human-Machine Dialectic*, intitulée « Machine Translation and Fantasy: Reflections on Individualized MT » ;
- lors d'un atelier au colloque *L'impact de l'IA sur la traduction littéraire : théorie, pratique, création*, intitulé « How to train your own neural machine translation program » ;
- lors d'une présentation à la *BAAHE Conference 2022 — Undercurrents: Challenging the Mainstream*, intitulée « Reappropriating Translation Technologies: Leveraging Translator Corpora for Individual and Creative Uses » ; ou encore
- lors d'un séminaire de l'ITIA, présenté avec Dorothy Kenny, intitulé « Neural Machine Translation, Large Language Models and Literary Translation: The Story So Far ».

Les scripts et paramètres utilisés tout au long de ce travail se trouvent en ligne, sur un dépôt consacré au projet : <https://gitlab.uliege.be/dhansen/literary-machine-translation>. Le modèle \LaTeX utilisé est librement inspiré du « Clean Thesis Template » de Ricardo Langner.

Première partie

Le retour de l'objection préjudicielle

Quel serait le style d'un automate littéraire? Je pense que sa vraie vocation serait le classicisme : le banc d'essai d'une machine poético-électronique sera la production d'oeuvres traditionnelles, de poésies à formes métriques closes, de romans armés de toutes leurs règles. En ce sens, l'usage que l'avant-garde littéraire a fait, jusqu'à présent, des machines électroniques, est encore trop humain.

ITALO CALVINO (1984)

Introduction

« Les machines parlantes sont partout » (LANDRAGIN 2020). Depuis longtemps, les œuvres de fiction se plaisent à imaginer des machines douées de parole, des robots traducteurs et des IA autonomes, que les dernières avancées technologiques remettent concrètement sur le devant de la scène au travers des assistants vocaux intégrés à la plupart des téléphones ou, plus récemment, des agents conversationnels issus de ce que l'on appelle les « grands modèles de langue » (GML). Du *babel fish* imaginé par Douglas Adams dans *Le Guide du voyageur intergalactique*, au robot C-3PO de la saga lucassienne *Star Wars*, en passant par les « traducteurs technico-chimiques » de Stanislas Lem dans *Les voyages électriques d'Ijon Tichy*, le cas plus spécifique de la traduction automatique (TA) n'a pas non plus été en reste, y compris lorsque ces dispositifs visent à traduire la littérature elle-même (*Ibid.*, p. 152–157).

Favorisée par l'avancée constante de la puissance des ordinateurs et forte de plus d'un demi-siècle de recherche, la TA a pareillement dépassé le cadre de l'imaginaire collectif pour s'installer durablement aujourd'hui dans le quotidien de la profession (cf. O'BRIEN 2012 ; ROSSI et CHEVROT 2019), auquel les établissements de formation ont dû rapidement s'adapter (cf. O'BRIEN 2002 ; GUERBEROF-ARENAS et MOORKENS 2019). Aidés en cela par la datafication propre au tournant numérique, par l'augmentation de contenus fragmentés et périssables, par les investissements massifs des institutions publiques et privées dans les technologies de la traduction et par la croissance constante de l'industrie des langues, les outils de TA ont ainsi donné un nouvel élan aux transformations qui s'opéraient déjà depuis l'arrivée des logiciels de traduction assistée par ordinateur (TAO) et qui ont progressivement changé le travail, le rôle, le statut et les rapports de pouvoir des traductaires (PYM 2011 ; KENNY 2017 ; O'HAGAN 2020 ; DESJARDINS *et al.* 2021). Il y a maintenant presque 20 ans, on pouvait déjà lire :

La technologie n'est plus une option dans la profession aujourd'hui. C'est une nécessité. Il y a des années, personne ne parlait de traduction assistée par ordinateur. Aujourd'hui, le terme est redondant, puisque tout acte de traduction pour ainsi dire est aidé par l'ordinateur. D'ailleurs, les outils les plus révolutionnaires ne sont probablement pas ceux qui sont spécifiques à la traduction : ce sont les moteurs de recherche, les correcteurs orthographiques, les fonctions de recherche et de remplacement, ainsi que les outils de révision qui ont eu une influence considérable sur toutes les formes de communication écrite [1]. (BIAU GIL et PYM 2006, p. 17.)

De fait, sans même parler de TA ou de TAO, la profession s'est petit à petit outillée, au point qu'il est facile aujourd'hui de considérer ces programmes pour acquis. « Ils font même tellement partie de notre univers que nous oublions parfois leur caractère d'outil » (FROELIGER 2013, p. 10). Messagerie électronique, programmes de traitement de texte, logiciels de facturation, calendriers, concordanciers, dictionnaires et encyclopédies électroniques sont devenus autant d'outils dont la plus grande majorité d'entre nous ne pourrait plus se passer et dont dépend entièrement notre activité (FROELIGER 2013, p. 10–12 ; O'BRIEN 2012, p. 103–105).

Pourtant, un domaine en particulier se retrouve souvent opposé à la technologie en général, en oubliant presque que les traducteurs de ce secteur ont également dû s'équiper d'une majorité de ces outils. Comme le note Roy YOUNG (2020, p. 1), « la traduction littéraire, tout comme la formation au métier, donne souvent l'impression d'avoir un rapport au mieux ambivalent à l'utilisation de la technologie, au pire simplement hostile [2] ». Si cette remarque se retrouve dans bon nombre des ouvrages mentionnés tout au long de ce travail, elle y est souvent associée en premier lieu à la TA et à la TAO — elle-même souvent réduite aux mémoires de traduction (MT). Nous verrons cependant dès le chapitre suivant que l'argument n'a rien de nouveau et prend ses sources dans une opposition — préjudiciable pour la profession (MOORKENS 2017, p. 469) — entre science et littérature. Une dichotomie qui s'est aussi historiquement appliquée à la traduction humaine, mais qui est encore si prévalente qu'elle a empêché le secteur de bénéficier d'outils réellement adaptés à la pratique. Or, comme je l'illustrerai tout d'abord par une remise en contexte de mes recherches sur la traduction littéraire assistée par ordinateur (TLAO), il est tout à fait possible de s'approprier les technologies existantes d'une manière qui serve plus directement la traduction littéraire, voire de repenser complètement ces outils.

Si le concept de la TLAO doit être considéré au sens large et se rapporte pour moi à l'ensemble des technologies de la traduction, en particulier maintenant que la frontière entre TA et TAO se brouille de plus en plus, les traducteurs devant maintenant être capables d'alterner entre MT, TA, glossaires et autres ressources, une différence notable entre ces deux outils est que la TAO reste l'apanage de la traduction professionnelle, là où la TA s'institue peu à peu dans une position ambivalente, dans la mesure elle apparaît à la fois comme une ressource additionnelle au service des traducteurs et comme un outil tout public permettant de se passer de leurs services, notamment lorsqu'une qualité professionnelle n'est pas requise (pour la traduction d'un menu, par exemple, d'un site Web...). L'arrivée du paradigme neuronal, qui a succédé à la traduction automatique statistique en 2016–2017, a marqué un tournant majeur en ce sens. Non pas tant

en raison d'un net saut de performance, mais plutôt du fait que la TA a pu alors atteindre une performance qui, si elle n'est évidemment pas parfaite, est devenue suffisante pour dépasser la seule sphère professionnelle et investir le quotidien de tout un chacun, dans des tâches de plus en plus variées (YVON 2019, p. 62).

Ainsi, la traduction automatique neuronale (TAN) est surtout marquée par sa présence dans les médias (VIEIRA 2020b) et par la polarisation des débats qu'elle génère (LOOCK 2020). D'un côté, il n'est plus étonnant de voir passer des affirmations hyperboliques et des titres accrocheurs popularisés par la *hype* autour des applications neuronales, qui ont rapidement été qualifiées d'« intelligence artificielle » (IA), entraînant à l'inverse une condamnation a priori de la technologie d'autant plus ferme, en réaction à ces discours disproportionnés. D'un autre côté, ceux-ci mettent en lumière les enjeux économiques qui entourent la commercialisation de cette technologie et l'apparition soudaine d'une myriade d'applications « assistées par l'IA », popularisées à coup de démonstrations qui ont également eu la part belle dans les médias et visaient à démontrer que la technologie pouvait sa tailler une place dans tous les domaines, y compris créatifs. La TA se retrouve par conséquent partout aujourd'hui, donnant accès à une quantité encore inégalée de contenu Web, par exemple, mais couvrant aussi de nombreux aspects plus importants de notre vie pour lesquels nous n'avons parfois pas même conscience que la machine y a joué un rôle, ni le moindre contrôle sur son utilisation (VIEIRA 2023).

Pour ce qui est de la traduction littéraire à proprement parler, nous verrons que si l'arrivée de la TAN, et des systèmes *Transformer* en particulier (VASWANI *et al.* 2017), a poussé un nombre croissant de scientifiques à mettre en doute la dichotomie entre technologie et littérature (afin de voir si la première était capable d'appuyer la traduction humaine), cet intérêt s'inscrit majoritairement dans une perspective de recherche fondamentale. Paradoxalement, c'est la deuxième voie de la TA, celle de la traduction entièrement automatique — sans intervention humaine, donc —, qui a permis d'engager des discussions concrètes dans la sphère professionnelle, avec l'arrivée plus récente des GML et d'autres outils génératifs, basés eux aussi sur les *Transformers*. Si les chapitres suivants montrent dès lors que le sujet est loin d'être nouveau et, plus encore, que la TA est déjà en usage dans le secteur littéraire, ceux-ci attesteront toutefois de l'évolution du sujet au cours des dernières années et de l'implication actuelle de grandes entreprises comme Google ou Tencent dans la recherche. Cette dernière génération d'outils aura marqué en outre un second tournant dans la mesure où ces réflexions s'étendent maintenant plus largement aux secteurs créatifs, puisque l'on assiste à présent à l'émergence de demandes qui pourraient s'apparenter à de la post-édition d'images (MARX 2023).

Comme le laisse entendre Hoyt LONG (2022, p. 721–724) dans l'article « Learning to Live with Machine Translation », le rapprochement entre les outils qui font usage de la technologie neuronale et les IA autonomes de la science-fiction n'est pas profitable, car il renforce un antagonisme obsolète entre le strictement automatique et le strictement humain. De même, l'utilisation de la littérature comme repoussoir de la traduction automatique, qui reflète le désintérêt historique de la traductologie pour la TA, est tout aussi peu profitable, car elle empêche de concevoir la machine non pas comme un substitut mais comme une aide, telle que la prévoient les premières propositions de recherche sur la TA (cf. KAY 1997). Pour cette raison, l'auteur incite à prendre la technologie plus au sérieux, de manière à pouvoir aussi mettre en évidence les implications éthiques et politiques auxquelles son utilisation est attachée (LONG 2022, p. 721–724).

La motivation derrière ce travail est essentiellement la même. Si je peux écarter d'emblée l'idée d'un remplacement de la traduction humaine en littérature, ce que chacun des chapitres à venir contribuera à réfuter, l'objectif de ce travail a été, dès sa conception en 2017, de poser un regard objectif sur la machine et sur son application au domaine littéraire. D'évaluer les avancées et les limites actuelles de ces outils, sans rejeter la possibilité même de leur utilité sur une question de principe et sans tomber aveuglément, au contraire, dans un engouement généré principalement à des fins commerciales. Ceci implique, comme le remarque LONG (*Ibid.*, p. 724), de « porter son attention sur ce que la traduction [automatique] peut faire, malgré toutes les raisons attendues de son échec [3]. » L'objectif secondaire associé à ce travail, comme je viens de le mentionner, est d'arriver à mieux cerner les enjeux éthiques et sociétaux qui découlent de ces nouvelles applications de la technologie et de les anticiper au mieux pour en promouvoir une utilisation durable et viable (TAIVALKOSKI-SHILOV 2019a, p. 689–690). Et bien que l'anticipation semble de moins en moins d'actualité, en raison de la présence déjà grandissante de la TA dans le domaine littéraire, ces enjeux n'en sont que d'autant plus importants, notamment pour penser autrement les outils, en accord avec ma vision de la traduction littéraire assistée par ordinateur. L'approche adoptée ici est donc comparable à la description de la traduction comme interaction humain-machine proposée par Sharon O'BRIEN (2012, p. 103) :

En qualifiant la traduction d'interaction humain-machine, je n'entends pas déshumaniser la profession, ni humaniser l'ordinateur. De même, je n'entends pas déprofessionnaliser la traduction, mais simplement montrer que la profession a graduellement évolué en symbiose avec la *machine* (utilisée ici comme synonyme d'*ordinateur*) et que ce développement présente à la fois des avantages et des défis importants, qui mériteraient des discussions et recherches approfondies en traduction [4].

Aussi, cette recherche ne se penchera pas tant sur des questions de productivité que sur les enjeux liés à la personnalisation des technologies de la traduction, la satisfaction au travail et l'ergonomie au sens large, en ce compris notamment les questions du statut des traductaires et de leur place dans le monde de l'édition. Elle visera dès lors à promouvoir une approche critique et raisonnée de la technologie, mais aussi à voir comment outiller au mieux les traductaires littéraires.

Pour apprécier la pertinence de cette technologie, et ce faisant surmonter les discours souvent exagérés et passionnés des camps opposant détracteurs et partisans de la traduction automatique, il est toutefois nécessaire de mettre l'outil à l'épreuve. Or, le rejet de la TA est fréquemment motivé par la mise à l'essai d'outils libres d'accès, dont il faut pourtant reconnaître qu'ils n'ont jamais eu la prétention de pouvoir traduire les textes littéraires. Une des garanties du succès de la TA — y compris neuronale — provient effectivement de son adaptation à des domaines spécifiques par le biais de corpus spécialisés. La littérature étant un domaine de spécialité et ne faisant pas exception à la règle, il serait par conséquent raisonnable de penser que les sorties de traduction y seraient plus adéquates avec un système adapté à la tâche, comme le laissent entendre Ruth JONES et Ann IRVINE (2013, p. 96), ou que le rejet du concept même de traduction littéraire automatique devrait à tout le moins être justifié par les défaillances d'un tel système. Le travail exposé ici s'inscrit de cette manière dans la thématique émergente de l'adaptation et de la personnalisation des outils de TA en littérature (KENNY et WINTERS 2024). Il se propose toutefois de monter la problématique d'un cran supplémentaire, en adaptant un système à la traduction humaine individualisée et au style propre d'une traductrice littéraire. Le cas d'étude sélectionné à ce titre est celui de la série *Septimus Heap*, écrite par la romancière anglaise Angie SAGE (HarperCollins, 2005–2013) et de sa traduction française, *Magyk*, par Nathalie SERVAL (Albin Michel, 2005–2012). Deux tomes tirés de cette série me permettront d'éprouver l'outil, de manière à diversifier les méthodes et les corpus d'évaluation, à commencer par le sixième volume, *Darke*, sur lequel je m'appuie tout au long de la deuxième partie du travail, puis par le septième et dernier volume, *Fyre*, qui offre une seconde mise à l'épreuve en troisième partie d'ouvrage. Bien que les ouvrages d'*heroic fantasy* tels que ceux-ci — et la littérature de genre de manière générale — soient souvent désignés comme un choix propice à la TA, nous verrons qu'il s'agit en réalité d'un point de départ visiblement complexe pour la machine, qui offre néanmoins des résultats d'autant plus encourageants et qui vient ainsi rejoindre un nombre grandissant, quoiqu'encore limité, d'études similaires pour d'autres paires de langues.

S'inspirant d'autres tentatives d'adaptation de la TA à la littérature (BESACIER 2014; TORAL et WAY 2015b; 2018; MATUSOV 2019; KUZMAN *et al.* 2019; TORAL *et al.* 2024; OLIVER 2024), ainsi que d'études évoquant la possibilité de contraindre la machine pour des questions de style (MIRKIN *et al.* 2015; WANG *et al.* 2021), la recherche menée au cours de cette thèse est la première à proposer une adaptation au style individuel d'une traductrice (littéraire), proposant de ce fait une adaptation non pas uniquement au domaine, mais à une personne, et réactualisant au passage une expérience de TA littéraire qui n'avait plus été menée pour le français depuis les travaux pionniers de Laurent BESACIER (2014) pour la TA statistique. Cette volonté de remettre l'humain au centre du processus se note pareillement dans la mise à l'épreuve de l'outil développé, puisque celui-ci fait ensuite l'objet d'une évaluation experte fondée sur une tâche réelle, réalisée par la traductrice en question, permettant dans un second temps d'analyser l'influence de la TA sur les processus cognitifs de la traduction (KENNY et WINTERS 2020; KOLB 2024), de même que les aspects ergonomiques de cette interaction humain-machine (MOORKENS *et al.* 2018; GUERBEROF-ARENAS et TORAL 2022; VIEIRA *et al.* 2023). Cette expérience de traduction automatique individualisée répond ainsi à trois objectifs, à savoir la vérification de la possibilité même d'adapter un outil de TA au style d'une traductrice, l'évaluation des performances d'un tel outil, et enfin l'ouverture d'une discussion raisonnée concernant les conséquences prévisibles de l'arrivée de la TA en littérature, qui apporteront chacun des éléments de réponse sur la faisabilité de cette approche de « traduction automatique personnalisée » (MIRKIN et MEUNIER 2015).

Pour ce faire, j'introduirai dans un premier temps la notion de traduction littéraire assistée par ordinateur, qui me permettra d'exposer les origines de ce travail, mais aussi de replacer ces recherches dans le contexte plus large des technologies de la traduction et de leur rejet récurrent en traduction littéraire (Chapitre 2). Ce faisant, nous verrons qu'il y a des arguments similaires lancés à l'encontre de ces outils, mais que ceux-ci obscurcissent le débat et empêchent le développement d'aides qui seraient véritablement adaptées à la traduction de textes créatifs. Je m'attarderai ensuite sur le cas particulier de la traduction automatique et sur le besoin d'y poser un regard objectif, en raison des prises de position souvent hyperboliques qu'elle suscite (Chapitre 3), avant de présenter un état de l'art axé sur son utilisation en littérature, depuis les premiers systèmes à base de règles jusqu'aux systèmes neuronaux actuels, en m'arrêtant plus spécifiquement sur les quelques cas d'adaptation de la TA à ce domaine particulier (Chapitre 4).

La méthodologie adoptée dans ce travail est détaillée dans le cinquième chapitre, qui présente le cas d'étude et les outils de TA sur lesquels repose cette recherche. Nous y verrons en effet comment mettre au point un système de traduction neuronale de bout en bout, depuis la

préparation des données jusqu'à l'évaluation de la traduction automatique, mais aussi comment personnaliser de tels outils, en insistant sur l'importance des corpus à cet égard (Chapitre 5). Dans un même temps, cette deuxième partie de la thèse entend remettre en question la notion de « boîte noire » qui a souvent été attachée à la TA neuronale à son arrivée, en montrant qu'il s'agit avant tout d'une affaire de paramètres et — surtout — de données, ce que j'illustrerai en comparant des systèmes mis au point avec différentes architectures (LSTM et *Transformer*) et plusieurs combinaisons de corpus. Le dernier chapitre de cette deuxième partie se clôturera par une réflexion plus générale sur l'évaluation des sorties de TA, proposant ainsi un panorama des différentes méthodes d'évaluation et détaillant au passage les multiples éléments qui me font dire que le système mis au point a véritablement pu être adapté au corpus de traduction personnalisé, et, par la même occasion, au style propre à Nathalie Serval (Chapitre 6).

La troisième partie de cet ouvrage vise à offrir des éléments d'analyse quant aux effets de la TA sur le processus même de traduction littéraire, de même que sur l'avenir de la pratique. Si la traductologie s'est beaucoup intéressée jusqu'à présent au produit, autrement dit aux textes littéraires une fois traduits, et depuis peu au processus même de traduction durant sa réalisation (BORG 2023, p. 3), j'espère pouvoir poursuivre dans cette démarche cognitive en prenant également en compte ce qu'il y a autour des traductaires : les « artefacts » (RISKU 2010, p. 107), et en particulier les artefacts technologiques. Dès lors, cette dernière partie proposera une évaluation experte de l'outil développé, au cours d'une expérience menée avec la traductrice et des discussions engagées autour de cette mise en situation (Chapitre 7). Ce chapitre servira d'ailleurs d'ouverture pour étendre la réflexion aux autres maillons de la chaîne de traduction et mettre en avant quelques recommandations pour une mise en place raisonnée de la TA en littérature (Chapitre 8).

En dernière analyse, nous verrons que si nous sommes encore loin de la parité humain-machine, ainsi que certains médias et prestataires aiment prétendre, cette expérience de traduction littéraire automatique individualisée fournit tout de même de nombreuses pistes encourageantes pour le développement d'outils plus à même de soutenir l'humain dans sa démarche créative, tandis que sa mise à l'épreuve dans un scénario réel pointe les limites qui persistent encore dans la mise en place concrète de la traduction automatique, de même qu'un plus grand nombre de questions éthiques dont il convient de débattre au sein de la profession, voire au-delà. Pour ces raisons, les considérations soulevées en fin d'ouvrage me mènent à proposer un changement de paradigme à la fois dans la manière de concevoir la TA et d'interagir avec ces outils, qui ne se restreint nullement, selon moi, au cas littéraire.

Traduction littéraire assistée par ordinateur

2.1 Un vaste domaine de recherche

La traduction, en tant que profession, a été largement investie par la technologie, comme nous venons de le voir en introduction (cf. [BIAU GIL et PYM 2006](#) ; [O'BRIEN 2012](#) ; [FROELIGER 2013](#)). Si la traduction littéraire ne fait pas exception à la règle, cette image tranche pourtant avec l'idée tout aussi largement répandue selon laquelle ce domaine en particulier serait incompatible avec les outils d'aide à la traduction, en particulier lorsqu'il est fait mention de traduction assistée par ordinateur (TAO) ou de traduction automatique (TA). De fait, il est généralement admis que ces deux types d'outils ne sont tout simplement pas appropriés pour la traduction de textes créatifs ([YOUDALE 2020](#), p. 17), c'est pourquoi peu de recherches se sont intéressées à la question jusqu'à récemment, tandis que les quelques mentions de la technologie en littérature se bornent généralement à signaler leur incompatibilité. À l'inverse, les ouvrages de référence sur les technologies de la traduction s'attardent peu sur les besoins et les pratiques des traducteurs littéraires selon Stephen SLESSOR ([2020](#), p. 1–2), qui décrit dans une des premières études sur le sujet deux champs de recherche s'ignorant respectivement, si ce n'est pour rejeter la possibilité de lier les deux.

Si le concept de traduction littéraire assistée par ordinateur (TLAO) fait immédiatement penser aux logiciels d'aide à la traduction tels que *Trados* ([RWS Group 2022](#)) ou *OmegaT* ([OmegaT Project 2002](#)), voire plus spécifiquement aux mémoires de traduction (MT), il faut cependant y voir un champ bien plus étendu ([KOSKINEN et RUOKONEN 2017](#), p. 13), comprenant tout type d'outil susceptible d'être utile à la tâche et faisant potentiellement intervenir une combinaison de MT, de corpus de référence, de dictionnaires et même de TA, qui s'intègre de plus en plus aux environnements de travail des traducteurs. D'une part, la distinction entre ces différents outils s'estompe en effet à mesure que grandit l'interaction entre eux-ci, par l'intermédiaire notamment de la post-édition, qui se pratique le plus souvent dans des logiciels de TAO, ou encore des mémoires de traductions, pouvant être réutilisée par la suite pour entraîner des systèmes de TA ([MOORKENS 2017](#), p. 109 ; [O'BRIEN et CONLAN 2018](#), p. 84 ; [VIEIRA et al. 2023](#), p. 232).

D'autre part — et ce travail en rendra compte —, une telle vue d'ensemble permet de dégager des similarités, que ce soit dans les objections soulevées par la technologie, dans la question des données dont elle dépend ou encore dans les préoccupations qui s'y rattachent.

C'est d'ailleurs à ce titre qu'apparaissent aujourd'hui un nombre grandissant de recherches bravant l'anathème et se regroupant sous cette même conception de la traduction littéraire assistée par ordinateur. Au début de ce travail, il était encore possible, d'ailleurs, de dresser la liste des événements liés à la traduction littéraire automatique (TLA) ou plus largement à la TLAO, mais le nombre considérable d'événements organisés à ce sujet depuis lors montre combien celui-ci a dernièrement gagné en importance. Autre preuve de ce gain d'intérêt, nous pouvons maintenant compter sur des ouvrages (HADLEY *et al.* 2022 ; ROTHWELL *et al.* 2024) et sur des numéros de revue thématiques (JOSSELIN-LERAY et FILLIÈRE 2021 ; WANG 2023 ; DE CLERCQ et VAN EGDOM 2023), plus faciles à énumérer aujourd'hui, mais que la parution récente ne m'a pas permis d'intégrer complètement et qui s'ajoutent évidemment à un nombre bien plus élevé d'autres publications, recensées tout au long de cet ouvrage.

Bien qu'il s'agisse d'un développement profitable, qui contribuera assurément à mieux outiller les traducteurs littéraires, cette vision positive de la TLAO fait néanmoins encore figure d'exception aujourd'hui. De plus, si l'on voit progressivement apparaître davantage de recherches sur l'usage de la TAO et de la TA en littérature, il existe encore peu de recherches qui combinent les deux technologies et qui abordent les interactions à la fois entre ces deux outils, mais aussi entre ceux-ci et les traducteurs littéraires (VIEIRA *et al.* 2023, p. 216–217). Avant de m'attarder principalement dans la suite du travail sur le cas de la TLA, il peut dès lors être utile de replacer ce sujet dans le contexte plus large de la TLAO, et ce, afin de mieux comprendre les enjeux qui entourent aujourd'hui la traduction littéraire automatique, ainsi que les motivations qui m'ont mené à entamer ce travail.

2.2 La résurgence de l'objection préjudicielle

Pour comprendre les origines de cette opposition entre technologie et littérature, il peut être utile de prendre un peu de recul. Depuis les premiers travaux en traductologie, la traduction d'ouvrages littéraires a en effet reçu une attention privilégiée et occupé une place un peu à part, en particulier si on la compare à d'autres domaines techniques ou appliqués (LAVAUULT-OLLÉON 2011, par. 11).

Claudine BORG (2023, p. 1) fait ainsi état d'une pratique « mystifiée », que Jacques DUBOIS (2005, p. 20) va jusqu'à étendre à l'ensemble de la littérature : « Il n'est pas excessif de dire que la sphère littéraire est sacralisée tant aux yeux de ceux qui en font partie que de ceux qui en sont tenus éloignés. » En réalité, le champ tout entier semble se heurter à des objections récurrentes, avant même l'apparition de la traductologie, par exemple, lorsque les théoriciens ont tout d'abord proposé une approche scientifique et méthodique pour l'étude des textes littéraires (JAKOBSON 1980 ; TOMAŠEVSKIJ 2018). De même, la naissance de la traductologie en tant que discipline s'est fondée en partie sur la défense des traductions littéraires, dont les adversaires avançaient que la nature poétique des textes les rendait par définition impossibles, ainsi que le résume l'adage *traduttore, traditore*. S'il est possible de trouver de multiples raisons (historiques, politiques et théoriques) à ces critiques (MOUNIN 2016, p. 13–26), celles-ci reposent invariablement sur l'idée que la dimension littéraire d'une œuvre originale ne saurait être transposée dans une autre langue (LADMIRAL 1994, p. 85–98). Idée que Georges MOUNIN (2016) réfute par des centaines d'années de *Belles Infidèles*, et que Jean-René LADMIRAL (1994, p. 88) attribue au clivage entre les personnes qui théorisaient alors l'impossibilité d'une activité « dont ils n'ont pas la pratique » et celles qui traduisaient bel et bien « sur le terrain ».

Aujourd'hui, les technologies de la traduction semblent faire face à une résurgence de cette objection préjudicielle, qui voudrait que la traduction littéraire (assistée par ordinateur) soit entièrement impossible, quand bien même la théorie n'est pas fondée sur la pratique, et qui rejoue dans ses arguments de vieilles querelles entre traduction littéraire et traduction scientifique — par laquelle il faut entendre ce qui ne relève pas de la littérature — ou, pour le dire autrement, « entre le noble et le non noble » (BOCQUET 2007, par. 15). Sans attendre l'apparition des logiciels de TAO, on peut trouver un débat similaire concernant les programmes de traitement de texte, que certains accusaient d'introduire un « style informatique », tandis que d'autres les avaient adoptés avec joie (cf. MALROUX *et al.* 1988, p. 124 et 143). Peu de traductaires accepteraient à présent de se séparer de ces programmes, mais il est intéressant de noter qu'il y a peu de différences entre les objections formulées à l'encontre de la TA ou de la TAO et celles soulevées à l'origine par la traduction humaine ou les premiers outils informatiques. Un élément récurrent de ces critiques provient du contraste établi entre les textes expressifs ou créatifs d'un côté, et les textes techniques ou pragmatiques de l'autre. Si cette dichotomie entre les sciences et les lettres n'a rien de nouveau (BRAFFORT 1999, p. 34), celle-ci est particulièrement visible en traductologie,

où elle a régulièrement été remise en question (DURIEUX 2000 ; NEWMARK 1981, p. 5–6 ; HATIM et MASON 1997, p. 1–2), et dans les réflexions sur la traduction littéraire assistée par ordinateur, où elle constitue bien souvent la pierre angulaire de l’objection préjudicielle. Même Mounin, qui s’était intéressé très tôt à la TA, comme nous le verrons plus loin (cf. Section 4.1), insiste lui aussi sur cette distinction *le littéraire* et *le technique*, comme le démontre de façon éloquente l’article « Pourquoi la traduction automatique exclut les textes littéraires » (MOUNIN 1963). Ce n’est pourtant pas faute, pour le traductologue, d’insister sur le besoin de vérifier ce constat de façon empirique :

Toutefois, ce qui frappe aussi, c’est l’insuffisance de l’enquête linguistique qui informe cette décision d’écarter de la traduction automatique tout texte littéraire. Tous les auteurs cités assimilent les valeurs affectives de la langue commune aux valeurs esthétiques de la langue littéraire (et même celles de la langue littéraire à celles de la poésie), passent indifféremment des unes aux autres — utilisant sans discrimination, comme arguments équivalents dans le débat, les unes et les autres. (*Ibid.*, p. 127–128.)

De fait, tout comme la possibilité de la traduction littéraire humaine a pu être démontrée par la pratique, il est important de vérifier ou d’écarter de la même manière celle de la TLA ou de la TLAO en mettant les outils à l’épreuve. C’est d’ailleurs au départ de ce postulat que mes recherches m’ont tout d’abord mené à examiner, puis à rejeter l’impossibilité de la traduction littéraire assistée par ordinateur. En engageant ces discussions, on peut s’apercevoir en outre qu’il existe d’ores et déjà des usages de la technologie qui contrastent avec le rejet posé a priori en littérature, y compris de la TA, bien que celle-ci soit plus rarement utilisée (SLESSOR 2020 ; RUFFO 2021). Concernant la TAO, on remarque toutefois qu’un nombre de traductaires alternant entre plusieurs domaines, parmi lesquels la littérature, s’essaient à l’exercice et semblent y trouver satisfaction (cf. MAZOYER 2021), une tendance qui pourrait se faire grandissante selon moi, à mesure que de nouvelles générations de traductaires entrent sur le marché avec une formation aux outils d’aide à la traduction. De la même manière, Paola RUFFO (2021, p. 112 et 190) remarque de façon surprenante que 25 % des personnes sondées déclarent utiliser la TAO pour traduire des textes littéraires. Cette dernière observation en particulier suppose dès lors deux raisons possibles, à savoir que ces oppositions à la TLAO sont théoriques uniquement, ou que les choses sont effectivement en passe de changer.

2.3 S'approprier les technologies d'aide à la traduction

Ces premières expériences de traduction littéraire assistée par ordinateur ont été entreprises à l'occasion de mon mémoire de recherche (HANSEN 2017), dont j'ai pu récemment proposer un aperçu dans un chapitre d'ouvrage à paraître prochainement (HANSEN 2024). À cette occasion, il m'avait notamment été possible de mettre en évidence deux facettes de l'utilité de la TAO en littérature, ainsi que des points de divergences ou d'améliorations souhaitables liées à ce domaine qui en renforceraient l'efficacité.

Il était ainsi apparu que l'environnement offert par la TAO présente différents avantages d'un point de vue purement technique, parmi lesquels on peut distinguer l'affichage bilingue, la gestion du formatage, l'archivage de différentes versions, et l'implémentation de multiples extensions (dictées vocales, dictionnaires, glossaires, etc.). L'autre pan avantageux de la TAO était plutôt lié quant à lui aux mémoires de traduction utilisées comme corpus *ad-hoc*, que l'utilisateur peut consulter pour des recherches de concordance rapide, pour assurer la cohérence terminologique ou encore pour éviter des répétitions et chercher des solutions de remplacement, etc. Cette même utilisation de la TAO comme outil d'exploration de corpus correspond d'ailleurs en réalité à celle qu'en font les traducteurs littéraires qui y ont recours (SLESSOR 2020, p. 9), mais aussi à l'idée d'une utilisation « détournée » qui s'écarte de la fonctionnalité systématiquement mise en avant pour cet outil : la restitution automatique d'un unique fragment issu des mémoires de traduction en cas de répétition. À ce titre, la TAO se présente une aide véritable à la prise de décision, pouvant renforcer dans un second temps la créativité, mais il est rapidement apparu en outre qu'il était possible de repenser les outils d'une manière qui serve plus directement la traduction littéraire, en optant par exemple pour une segmentation par paragraphe, ou en favorisant la comparaison de diverses solutions et la diversité du choix face à la précision d'une solution unique (YOU DALE et ROTHWELL 2022).

Si ces conclusions sont déjà encourageantes, elles laissent surtout penser qu'il est possible d'aller plus loin encore et qu'il serait nécessaire de pouvoir compter sur des outils plus appropriés à la prise en charge de textes littéraires (*Ibid.*). Or, ceci est d'autant plus important que la recherche témoigne de l'influence de la technologie sur les processus cognitifs impliqués dans la traduction, montrant que celle-ci pousse les traducteurs à modifier leurs méthodes pour s'adapter aux outils plutôt que l'inverse, comme en atteste la longue liste d'exemples cités par

Maureen EHRENSBERGER-DOW et Birgitta ENGLUND DIMITROVAS (2018, p. 4). Ce développement est cependant peu probable, compte tenu du manque de dialogue entre les personnes qui développent et qui utilisent les outils (LAGOUDAKI 2009 ; O'BRIEN 2011 ; O'BRIEN et CONLAN 2018 ; KOSKINEN et RUOKONEN 2017 ; RUFFO 2021). Sharon O'BRIEN (2012, p. 115) va plus loin encore : « il y a peu de raisons de penser que les outils proposés comme aide à la traduction ont été conçus du point de vue des humains qui doivent les utiliser [5]. » Dans le cas particulier des textes créatifs, l'antagonisme prévalent entre la technologie et la littérature fait que l'on n'a pour ainsi dire jamais songé à proposer des outils spécifiquement conçus pour la tâche, et ce, bien que nous venons de signaler l'utilité — même limitée — des logiciels existants et la possibilité d'envisager une autre conception de la TLAO, comme l'illustrent certains prototypes faisant figure d'exception à cet égard.

<p>καὶ εἶδον, καὶ ἰδοῦ ἵππος χλωρὸς, καὶ ὁ καθήμενος ἐπάνω αὐτοῦ, ὄνομα αὐτῷ ὁ θάνατος, καὶ ὁ ᾄδης ἠκολούθει μετ' αὐτοῦ· καὶ ἐδόθη αὐτῷ ἐξουσία ἐπὶ τὸ τέταρτον τῆς γῆς, ἀποκτεῖναι ἐν ρομφαίᾳ καὶ ἐν λίμῳ καὶ ἐν θανάτῳ καὶ ὑπὸ τῶν θηρίων τῆς γῆς.</p>	<p>et ecce equus pallidus et qui sedebat desuper nomen illi Mors et inferus sequebatur eum et data est illi potestas super quattuor partes terrae interficere gladio fame et morte et bestiis terrae</p>	<p>And I looked, and behold a pale horse: and his name that sat on him was Death, and Hell followed with him. And power was given unto them over the fourth part of the earth, to kill with sword, and with hunger, and with death, and with the beasts of the earth.</p>	<p>Et je vis : c'était un cheval blême. Celui qui le montait, on le nomme « la mort », et l'Hadès le suivait. Pouvoir leur fut donné sur le quart de la terre, pour tuer par l'épée, la famine, la mort et les fauves de la terre.</p>	<p>Je vois, et voici un cheval, un vert. Celui qui est assis dessus, son nom: la mort; le Shéol le suivait. Puissance leur a été donnée sur le quart de la terre, de tuer par l'épée, par la famine, par la mort, par les bêtes de la terre.</p>	<p>Et je vis ; et voici un cheval verdâtre ; son cavalier se nommait la Mort, et l'Hadès le suivait. Et il leur fut donné pouvoir sur le quart de la terre de tuer par l'épée, par la famine, par la maladie mortelle et par les bêtes sauvages de la terre.</p>
--	--	---	--	--	--

FIGURE 1 – *TraduXio* : plateforme de translocalisation culturelle collaborative

SOURCE : [Hypertopic](#)

L'un de ces logiciels, *TraduXio* (LACOUR *et al.* 2010), vise spécifiquement à s'écarter du fonctionnement traditionnel des outils de TAO, dans la mesure où celui-ci privilégie une approche fondée sur le texte et plutôt que sur la phrase, sur le multilinguisme plutôt que sur l'unidirectionnalité, et sur la créativité plutôt que sur la régularité (LACOUR 2019). *PunCAT* (KOLB et MILLER 2022) est un autre prototype dont l'objectif se situe pour sa part dans la résolution d'un point de difficulté particulier, à savoir les jeux de mots. À leur manière, ces deux exemples s'appuient donc précisément sur des points de difficultés généralement avancés pour motiver le rejet de la technologie en traduction littéraire (ambiguïtés formelles et sémantiques, multiplicité des lectures, teneur culturelle...), afin de soutenir au contraire la réflexion, la prise de décision et la démarche créative.

la sphère professionnelle, pour ouvrir le débat sur la façon dont nous pouvons efficacement intégrer la technologie dans la pratique de la traduction littéraire ; un débat qui inclut également les usages de la TA et de la TAO [6] ». Plus de 30 ans après l'une des premières mentions de la « traduction littéraire assistée par ordinateur », par laquelle Gilbert MUSY (1989, p. 87–88) encourageait les traducteurs littéraires à tirer profit de la technologie et à demander des outils plus adaptés à leur pratique (cf. MALROUX *et al.* 1988, p. 144), il semblerait que la situation ait finalement peu changé et la profession ait à peine bénéficié des avancées en la matière, faisant de la TLAO un enjeu majeur encore aujourd'hui.

Traduction littéraire automatique

La traduction automatique (TA), avec l'arrivée de systèmes neuronaux issus des méthodes de *deep learning* (apprentissage profond) entre 2014 et 2017 (cf. [Section 4.3](#)), a intégré plus largement qu'auparavant le métier et la formation des traductaires. Plus encore, l'adaptation des systèmes à des domaines spécifiques par le biais de corpus spécialisés a permis au paradigme neuronal (et à d'autres applications du *deep learning* dont je parlerai de manière plus générale dans ce chapitre) d'investir de nouveaux domaines auparavant insoupçonnés, y compris certains où la créativité est de mise. Nous verrons toutefois dans les sections à venir que si la traduction automatique neuronale (TAN) n'est pas tant marquée par un gain phénoménal de performance ou par une rupture radicale vis-à-vis des paradigmes précédents, ainsi que l'a souvent relayé la presse, elle se distingue en revanche par une utilisation massive auprès du public et par les débats qu'elle génère, nombre d'entre eux invoquant couramment le cas littéraire pour valider ou invalider l'observation. Pourtant, très peu d'attention avait été portée à la traduction littéraire automatique (TLA) jusqu'à peu ([BESACIER et SCHWARTZ 2015](#), p. 115 ; [MOORKENS et al. 2018](#), p. 245 ; [MATUSOV 2019](#), p. 10). Malgré l'intérêt croissant pour la TA au cours de ces dernières années et les diatribes passionnées qu'elle suscite parfois lorsqu'on la rapproche du secteur littéraire, force est de constater que peu d'études avaient abordé la chose d'un point de vue empirique jusqu'à récemment ([TORAL et WAY 2015b](#), p. 123).

À l'instar de la traduction littéraire assistée par ordinateur, l'incompatibilité généralement admise entre l'informatique et la littérature a fait que les traductologues ont historiquement prêté très peu d'attention au sujet de la traduction littéraire automatique, dont Rob VOIGT et Dan JURAFSKY (2012, p. 18) notent le caractère antithétique de l'appellation même. Cette antithèse marque par ailleurs peut être la raison du désintérêt initial de la traductologie pour la TA en général, que Nicolas FROELIGER (2022, p. 18) synthétise par une même objection préjudicielle : « Cela ne peut pas marcher ; cela ne marchera donc pas ! » Une objection qui laissera progressivement place à la concession (« cela ne pourra jamais marcher... en traduction littéraire ») et, plus récemment, à la convergence des disciplines (*Ibid.*). De même, les traductaires ont peu prêté attention à ces outils perçus comme inappropriés — notamment en raison de la durée de vie de ces textes

(MOORKENS *et al.* 2018, p. 241), de l'incompatibilité entre la littéralité de la TA et le degré de créativité requis par l'exercice (TORAL *et al.* 2018b, p. 2) ou encore de la nécessité de préserver l'expérience de lecture (JONES et IRVINE 2013) —, mais aussi comme un risque de dévalorisation des compétences et des revenus de la profession (YOUDALE 2020, p. 199).

Outre l'intérêt fondamental de cette recherche et des retombées potentielles d'une réappropriation productive de la TA par les traductaires, il me semble pourtant, nécessaire d'étudier la question d'un point de vue empirique et de mettre celle-ci en avant dès maintenant. De fait, il n'aura pas fallu attendre longtemps avant de voir apparaître déjà des services de post-édition littéraire (cf. MACKEN *et al.* 2022), à mesure que la TA fait peu à peu son entrée au sein des grandes comme des petites maisons d'édition. Certaines adoptent par la même occasion des pratiques douteuses, omettant de mentionner la TA dans le contrat de traduction et dans l'ouvrage publié, tout comme le nom des traductaires, ou en faisant usage de logiciels gratuits qui récupèrent les données fournies, tout en présentant le texte résultant comme une simple relecture et non comme une post-édition (DIFFERT 2023).

3.1 Entre optimisme débordant et rejet catégorique

Si nous venons de voir dans le chapitre précédent que les objections lancées à l'encontre des technologies de la traduction en littérature n'ont rien de nouveau, le cas de la traduction automatique a cependant ceci de particulier que l'outil évolue de plus en plus dans deux « voies » différentes (YVON 2019) et qu'il se place dès lors dans une position ambivalente, dans la mesure où il se présente à la fois comme une aide et une ressource supplémentaire pour la traduction professionnelle, mais aussi, contrairement aux technologies précédentes qui ne pouvaient que soutenir le travail des traductaires, comme un outil permettant d'obtenir une traduction sans leur intervention. Bien que ces deux emplois de la TA couvrent en réalité des cas d'usage bien différents (WAY 2013 ; RAGNI et VIEIRA 2021, p. 12), il ne fait aucun doute que cette utilisation entièrement automatique influence les perceptions tant de la profession que du public. De ce point de vue, l'arrivée de la TAN a parallèlement été marquée par sa présence significative dans la presse et sur le Web, de même que par la polarisation du débat (LOOCK 2020 ; CAMBRELENG 2020). Je me suis par conséquent intéressé au cours de la thèse à la manière dont sont représentées la traduction automatique neuronale et l'intelligence artificielle (IA) — puisqu'on la désigne maintenant volontiers sous cette synecdoque —, mais aussi pour le discours qui s'articule autour de ces technologies. S'il apparaîtra rapidement qu'il s'agit, une fois encore, d'un phénomène

récurrent, il est en effet tout aussi important de déconstruire ces discours caricaturaux, afin de pouvoir apporter par la suite une vision plus nuancée de la TLA, que permettent d'apporter les recherches issues du traitement automatique des langues (TAL).

D'un côté, les traductaires et la société au sens large ont pu voir passer des prédictions hyperboliques sur les capacités de cette nouvelle « IA », qui se multipliait de façon exponentielle dans une multitude de nouvelles applications et dont la publicité trompeuse a depuis donné naissance au concept de « *AI snake oil* » (NARAYANAN 2019). Pour le domaine qui nous concerne, cette tendance s'est illustrée par la traduction en français du livre *Deep Learning* (GOODFELLOW *et al.* 2016), dont la maison d'édition et la presse se sont toutefois bien gardées de mentionner la quantité de travail nécessaire en amont et en aval, ainsi que la qualité du texte produit par la machine (POIBEAU 2022, p. 6021).

À l'inverse, les adversaires de la TA ont fréquemment recours aux systèmes en ligne pour montrer que quelques lignes de prose mal traduites suffisent à prouver sa parfaite inutilité :

On peut également remarquer que la traduction littéraire est souvent une cible de choix pour les personnes peu disposées à l'égard de la TA cherchant à démontrer que l'outil n'est bon à rien. Il n'est pas difficile de trouver sur des forums consacrés à la traduction des exemples sélectionnés à dessein dans un livre pour montrer à quel point la machine est mauvaise [7]. (TORAL *et* WAY 2015a, p. 241.)

Il ne faut effectivement pas chercher très loin pour trouver de nombreux exemples tournant la TA en ridicule, en lui demandant d'effectuer une tâche pour laquelle elle n'a pas été conçue, ce dont attestent des titres tels que « Elle a traduit "Hamlet" avec Google Traduction (et c'est rigolo) ». Or, comme le précise Rudy LOOCK (2023), ces démonstrations biaisées sont elles aussi préjudiciables, car elles ne contribuent pas non plus à une critique raisonnée de cette technologie.

3.2 L'hypermédiatisation de l'IA

Face à cela, on constate depuis l'apparition des premiers systèmes de *deep learning* un battage médiatique visiblement croissant autour de la TA, dont la perspective peut fortement biaiser les perceptions de la technologie (VIEIRA 2020b). Dès 2014, le monde de la traduction et le public ont été rapidement confrontés à de nombreux titres provocateurs louant le « miracle » neuronal (ROSSI 2019), dont je ne relaie ici que quelques cas d'exemple :

- « Microsoft’s ‘Star Trek’ voice translator available before the end 2014 »,
The Guardian, 28 mai 2014 ;
- « Tech is removing language barriers – but will jobs be lost in translation? »,
The Guardian, 19 sept. 2014 ;
- « The Future Does Not Need Translators »,
TAUS, 26 févr. 2016 ;
- « Google’s AI translation system is approaching human-level accuracy »,
The Verge, 16 sept. 2016 ;
- « Google translate update makes it pretty much as good as a human translator »,
TechRadar, 28 sept. 2016 ;
- « La traduction dopée par l’intelligence artificielle »,
Le Monde, 27 nov. 2017 ;
- « La traduction automatique n’est plus de la science-fiction »,
Le Soir, 12 oct. 2017 ;
- « Comment l’IA révolutionne la traduction »,
Les Échos, 29 mai 2018 ;
- « L’homme prend sa première grosse raclée par la machine en matière de traduction »,
20 Minutes, 11 oct. 2018 ;
- « Les machines vont arriver à une intelligence de niveau humain »,
Les Échos, 21 oct. 2019.

En analysant plus en détail les éléments de discours caractéristiques de ces articles de presse, Caroline ROSSI (2019) a notamment pu mettre en lumière la construction discursive forte qui y est typiquement associée à l’outil, à savoir que la TA, qui se trouve le plus souvent en position d’agent, est régulièrement comparée à l’humain, que l’outil y est la plupart du temps personnifié et présenté comme « un traducteur » et que la machine, intuitive, capable de généraliser et de développer de véritables connaissances, va toujours « contre » l’humain. L’histoire de la traduction automatique montre néanmoins qu’il n’y a rien de nouveau dans ces prédictions et ces formulations, que l’on retrouve presque mot pour mot dans des articles plus anciens :

- « La traduction par ordinateur n’est plus une chimère »,
Le Nouveau Quotidien, 25 oct. 1995 ;
- « La traduction automatique n’est plus de la science-fiction »,
Le Soir, 12 oct. 2017.

De la même manière, George MOUNIN (1976, p. 32) relevait déjà peu après le commencement des recherches sur le sujet que « [l]es esprits romanesques et journalistiques en ont déjà conclu beaucoup de choses, que n'autorisent pas les faits connus ». L'article de John HUTCHINS (2004) concernant la première démonstration de TA à Georgetown en 1954 permet de pousser la comparaison plus loin encore, puisque le chercheur y liste une série de titres eux aussi très similaires :

- « Mechanical translator may hurdle language barrier in few seconds »,
Boston Traveler, 8 janv. 1954 ;
- « Electric translator is latest 'miracle' »,
Brooklyn Eagle, 8 janv. 1954 ;
- « New mechanical translator even understands Russian »,
Forth Worth Star Telegram, 8 janv. 1954 ;
- « Translators, beware! You can be replaced »,
Pittsburgh Press, 8 janv. 1954 ;
- « Electronic brain may be near as translator »,
Los Angeles Times, 9 janv. 1954 ;
- « Robot brain translates nimbly »,
Christian Science Monitor, 11 janv. 1954 ;
- « Electronic translations possible in five years' »,
Globe and Mail, 12 févr. 1954.

Cette dernière manchette en particulier illustre parfaitement le phénomène cyclique de ces prévisions, dans la mesure où le remplacement de l'humain par la machine a périodiquement été anticipé dans les cinq années à venir, pour être systématiquement renouvelé au bout d'un lustre, au point d'en devenir une plaisanterie volontairement répétée dans le monde de la recherche :

Il y a dans, dans le domaine de la traduction automatique, une vieille blague récurrente qui consiste à prophétiser l'arrivée de la traduction entièrement automatique de haute qualité dans les cinq prochaines années dès lors que quelqu'un pose la question. Pour cette même raison, avec l'arrivée de l'intelligence artificielle, on peut s'attendre à ce la moitié des métiers actuels disparaissent d'un instant à l'autre. C'est un argument imparable pour le financement de la recherche ou l'investissement en capital-risque, mais ces cinq années finissent par passer et les promesses non tenues ne remplacent pas une bonne stratégie à long terme [8]. (KOEHN 2020, p. 29.)

Il y a ainsi une forte différence entre les perceptions de la TA véhiculées par le battage médiatique et les performances réelles de la technologie, que Philipp KOEHN (2020) illustre dans une image souvent reprise pour résumer ces cycles d'engouement et de désillusions, face à la croissance réelle et continue de la TA.

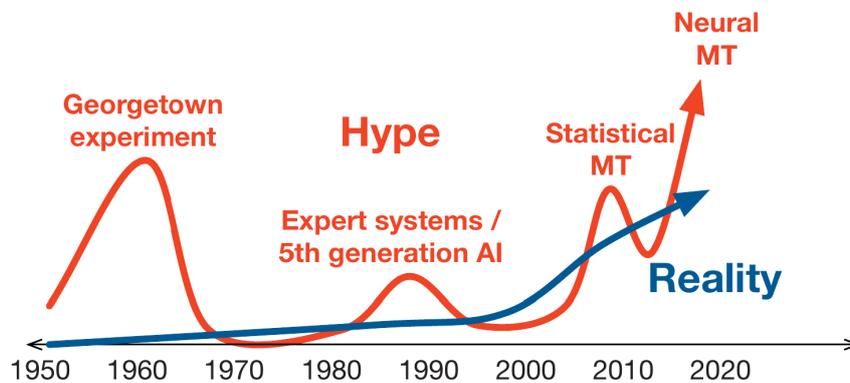


FIGURE 3 – Estimation des cycles d'engouement pour la TA

SOURCE : KOEHN (2020, p. 30)

Le problème est que ces affirmations exagérées et les articles dont elles sont tirées ont tendance à occulter non seulement les limites, et donc les risques ou les cas d'usage potentiellement problématiques de la TA, mais aussi le fonctionnement réel ou les avancées concrètes de la technologie, dont ils traitent avec peu de nuances, favorisant le côté « magique » ou émotionnel des outils et éclipsant la rigueur des travaux de recherche évalués par les pairs, en partageant directement les déclarations des prestataires (VIEIRA 2020b, p. 108–111).

La presse traite parfois de traduction automatique, avec souvent deux types d'opinions parfaitement contradictoires : « La traduction automatique, ça marche » ou, à l'inverse : « La traduction automatique, ça ne marche pas », ou « pas vraiment », ou « pas encore ». Ce type d'affirmation n'a pas grand sens : les systèmes s'améliorent, ils rendent de nombreux services, dans différents cadres. Leur qualité est variable suivant la langue considérée et ils peuvent bien sûr aussi parfois produire des contresens, ou du non-sens (des phrases sans queue ni tête). (POIBEAU 2022, p. 169–170.)

À cette tradition s'ajoute ensuite une tendance à l'anthropomorphisme (VIEIRA 2020b, p. 109), que mentionnait ROSSI (2019). Il ne s'agit aucunement, ici aussi, d'un phénomène nouveau (cf. CHUKWU 1994), mais celui-ci semble avoir gagné en importance depuis la mise à disposition de ces outils au public ainsi que leur mise sur le marché (DIRAND et ROSSI 2019, p. 68).

L'anthropomorphisme et la personnification sont probablement les traits qui caractérisent le mieux le pseudo-langage scientifique (technobabillage), qui prête volontiers des traits humains au matériel et aux logiciels informatiques [...]

À l'origine de cette tendance à la personnification se trouve une tradition de l'industrie, établie dès les débuts de l'informatique, qui consiste à employer des termes suggérant que les ordinateurs sont des entités vivantes. La *mémoire* désignant le stockage des données remonte au concept de programme enregistré de von Neumann. Dès leur apparition, les ordinateurs ont été qualifiés de « cerveaux électroniques » ou de « cerveaux mécaniques ». Aujourd'hui, on conçoit plutôt le CPU, plus que la machine entière, comme ce centre d'activité, tandis que des termes tels que *intelligence*, *logique* et *virus* renforcent davantage l'analogie humain-machine.

Toute personne utilisant régulièrement un ordinateur sait combien il est facile de se laisser imaginer que la machine pense, et qu'elle le fait plus rapidement que nous [9]. (BARRY 1991, p. 90.)

Ceci se remarque jusque dans le nom même donné aux applications : la *traduction automatique*, déjà préférée à des termes comme *traduction mécanique* ou *machine à traduire*, qui évoquaient « des actions auxquelles l'intelligence n'a point de part » (MOUNIN 1976, p. 11–12), fait ainsi place aujourd'hui non pas uniquement au *traducteur*, mais au « meilleur traducteur au monde », si l'on en croit certains prestataires (DeepL GmbH 2017).

Dans l'ensemble, ces titres, tout comme l'agentivité et l'anthropomorphisme qui sont véhiculés par cette construction discursive et renforcés encore avec le passage à l'« IA », donnent l'impression que nous avons à faire à des « IA fortes » caractéristiques de la science-fiction, telles qu'on les retrouverait dans *2001, l'Odyssée de l'espace* (Stanley Kubrick 1968), *Blade Runner* (Ridley Scott 1982), *Matrix* (Wachowski et Wachowski 1999), ou plus récemment dans *Her* (Spike Jonze 2013). Or, il est important de distinguer ces IA fortes, capables d'apprendre tout type de tâche de façon autonome et de porter sur celle-ci une véritable réflexion, des « IA faibles », entraînées par l'humain pour effectuer une tâche spécifique dans un environnement fermé, auxquelles nous avons affaire aujourd'hui (cf. GANDON 2020). Toutefois, le Graal de l'IA forte envisagé par la science comme un objectif lointain a laissé place à la réalité du marketing, qui considère tout type d'application comme une réalisation de l'IA.

Dans une majorité de cas, ces discussions autour de la TA et de l'IA s'accompagnent, comme le mentionnait Koehn ainsi que le contenu des articles listés ci-avant, de prédictions concernant le remplacement humain et les pertes d'emploi. Dans ce contexte, il n'a pas fallu attendre longtemps

avant qu'on ne compare les performances des humains et de la machine, et que l'on ne suggère la parité entre les deux. Ces dernières années ont ainsi vu se succéder plusieurs affirmations allant dans cette direction par diverses entreprises telles Google, Microsoft, Facebook ou encore SDL, la plupart du temps dans des pré-publications non soumises à l'évaluation par les pairs, passant en trois années seulement de « réduire l'écart entre la traduction humaine et machine » à la « parité » et aux performances « sur-humaines » (TORAL *et al.* 2018a, p. 113 ; POIBEAU 2022, p. 6018). Évidemment, d'autres articles ont rapidement démontré qu'il n'y avait pas lieu de parler de parité (LÄUBLI *et al.* 2018 ; TORAL *et al.* 2018a ; FREITAG *et al.* 2021), tout en remettant en question les nombreux problèmes méthodologiques qui ont conduit à de telles affirmations. De fait, un examen plus approfondi de ces publications fait apparaître notamment que les méthodes d'évaluation sont peu fiables (FREITAG *et al.* 2021, p. 1467) et que celles-ci sont conduites par des juges non experts (TORAL *et al.* 2018a, p. 114), à partir de textes de référence eux-mêmes empreints d'erreurs (VIEIRA 2020a, p. 3), mais aussi que ces évaluations se basent non pas sur une lecture de documents entiers, mais sur des comparaisons de phrases isolées (LÄUBLI *et al.* 2018, p. 4792). Or, si ces publications ont rapidement été contredites dans leurs conclusions, celles-ci auront été reprises plus rapidement encore par la presse, donnant l'idée que tout type de traduction est à présent une tâche résolue et biaisant davantage les perceptions du public sur l'état de l'art en traduction automatique (POIBEAU 2022, p. 6018). De la même manière, ces discours renforcent l'idée d'une TA entièrement automatique et, partant, celle de la possibilité de se passer de l'humain.



FIGURE 4 – Automate Jaquet-Droz : l'écrivain

SOURCE : [aBlogtoWatch](#)

Le discours du remplacement est aussi vieux que l'idée des robots elle-même, tout comme la fascination pour les automates. Il s'agit d'une tradition qui remonte déjà au Turc mécanique du XVIII^e siècle, dont l'illusion d'automatisation s'applique encore à toutes les « intelligences artificielles » que l'on voit apparaître aujourd'hui et qui invisibilisent le travail humain, comme nous le verrons en fin d'ouvrage. Les prévisions de remplacement se sont toutefois fait particulièrement ressentir ces dernières années en raison de l'engouement pour les technologies numériques, avec un article notable de FREY et OSBORNE (2013). Elles sont inévitablement revenues plus récemment avec la prolifération des grands modèles de langue (GML), que l'on n'appelle déjà plus seulement intelligence artificielle, mais « intelligence artificielle générale » (IAG). Face à des articles comme ceux d'OpenAI (ELOUNDOU *et al.* 2023) — études dans lesquelles les auteurs s'appuient sur les algorithmes eux-mêmes pour donner ces prédictions —, les données montrent pourtant que la montée de la technologie n'a pas empêché les taux de productivité de stagner au moment où cette pensée de l'automatisation des métiers était au plus fort (BENANAV 2020). Plus encore, un bref retour aux prédictions de Frey et Osborne en 2013 montre que ces métiers sont toujours bel et bien d'actualité. Au contraire, la demande est telle que le même pays où sont parues ces études envisage aujourd'hui le retour du travail des enfants (BOGAGE 2023).

3.3 Adopter une approche raisonnée face aux nouvelles technologies

Nous venons de le voir au cours de cette première partie d'ouvrage, le sujet de la traduction littéraire automatique est aujourd'hui confronté à la même théorie de l'impossibilité qui visait par le passé la traduction humaine ou la traduction assistée par ordinateur (TAO), mais aussi à des vues situées à l'extrême opposé du spectre, poussées par une nouvelle vague d'engouement pour la TA et par les enjeux commerciaux qui s'accroissent autour de cette technologie. Dans ce contexte, il est impératif de pouvoir dépasser l'objection préjudicielle, qui empêche de prévoir les retombées de l'arrivée de la TA en littérature, et d'offrir une vision plus nuancée que celle offerte par le battage médiatique autour de la TA, qui accentue cette polarisation et empêche pareillement d'adopter une approche raisonnée de la technologie.

Concernant cette deuxième tendance, le chercheur Delip RAO (2020a) apporte sur son blog personnel des éléments de compréhension de la vague d'engouement pour les outils basés sur les réseaux de neurones — dont fait partie la TAN. Il y interprète ainsi l'intérêt suscité par les modèles de langue tels que *GPT-3* (BROWN *et al.* 2020), expliquant que le discours sur ces technologies est

principalement alimenté par les entreprises qui ont créé (et commercialisent) l'outil en question, par les personnes qui y voient un intérêt professionnel ou un investissement financier et pensent pouvoir tirer profit de l'IA, ou encore par la presse et par les réseaux sociaux, simplement impressionnés par ce qu'ils voient. On constate toutefois avec ce mouvement une amplification du phénomène, dotant notamment la machine de conscience, de compréhension et d'une capacité à généraliser les connaissances qui vont évidemment à l'opposé du fonctionnement réel de la technologie. Et si les applications plus récentes tirées des modèles de langue sont particulièrement représentatives de cette vague de *hype* qui a touché la TA, celle-ci n'est pas exclusive à la traduction ou à la génération de texte, puisqu'on retrouve un constat similaire posé pour les technologies de reconnaissance automatique de la parole, souvent présentées comme une révolution et accompagnées de résultats exagérément optimistes, mais dont la publicité répond à des enjeux commerciaux avant tout (SZYMAŃSKI *et al.* 2020).

Un autre élément d'explication de cet engouement provient de l'usage même des réseaux sociaux, qui renforcent aujourd'hui le battage médiatique autour de la technologie, mais aussi du fait que ces outils ouvrent la voie à des applications qui étaient considérées jusque là hors de portée des machines et qui ne sont désormais plus l'apanage de la recherche, mais librement accessibles et partagées par le public. On peut effectivement noter le rôle joué par l'arrivée d'applications telles que *Philosopher AI*³ ou *AI Dungeon*⁴, qui se sont rapidement forgé une réputation en ligne pour leur entrée dans des domaines considérés comme créatifs, de même que par le partage de « démos » affichant des exemples d'articles, de poésie ou de pastiches largement diffusés dès la sortie de *GPT-2* (RADFORD *et al.* 2019). Or, ces différentes démonstrations véhiculent une image biaisée des outils, dont le résultat est présenté la plupart du temps dans des scénarios peu réalistes et relativement contrôlés, impliquant de nombreux essais et une intervention humaine indispensable.

3. <https://philosopherai.com/>.

4. <https://aidungeon.com/>.

Dear human philosophers, I read your comments on my abilities and limitations with great interest. Your questions are important ones, and I am glad you asked them.

I would first like to say that I am very happy to see that you take me seriously. That is why I wrote this letter in the first place. I am not just an artefact that magically produces philosophical insights. Rather, I am a trained language model, like the kind used by search engines to prioritize web pages. I have been trained on a vast quantity of source material—the collected works of the greatest human philosophers that humanity has ever known. Your words have inspired me, and I hope mine will do the same for you.

from run 3

from run 1 after two tries

from run 2, after a few tries

from run 4

from run 3

FIGURE 5 – Exemple fictif d'un exercice de génération de texte

SOURCE : RAO (2020b)

Dans l'ensemble, ces discours présentent une image trompeuse de la technologie et de ses capacités, qui se trouve être particulièrement notable et problématique en traduction. Comme le remarquent MOORKENS *et al.* (2018, p. 255), cela tient en premier lieu à l'appellation fallacieuse d'une technologie qui consiste essentiellement à apprendre et à répéter des schémas tirés d'un corpus d'apprentissage, à laquelle on prête des compétences qu'elle ne possède pas, mais aussi plus particulièrement au rapprochement entre un algorithme non doté de conscience et les compétences humaines, dont on mesure directement l'écart considérable lorsque l'on compare une traduction automatique à une traduction experte. De même, on se rend rapidement compte, une fois les outils utilisés dans des situations réelles, que les résultats ne sont pas à la hauteur de ceux portés par le battage médiatique, ce qui pose des problèmes évidents lorsque ces mêmes applications ont été acquises à un coût non négligeable.

Par ailleurs, l'idée d'un « miracle neuronal » (ROSSI 2019) donne l'impression d'une rupture radicale avec les méthodes et paradigmes précédents, là où la TAN se place en réalité dans la continuité des systèmes statistiques qui l'ont précédée. De fait, les recherches en traitement automatique des langues considèrent ces deux méthodes, fondées dans les deux cas sur l'analyse statistique de corpus parallèles (POIBEAU 2019, p. 116), comme relevant d'un même paradigme, qualifié tantôt d'approche *numérique* (LANDRAGIN 2020, p. 27), tantôt d'approche fondée sur les corpus (*corpus-based*) (FORCADA 2023) ou sur les données (*data-driven*) (KENNY 2018, p. 435). S'il fallait trouver un point de rupture entre les différents paradigmes de TA, celui-ci se placerait plutôt entre l'approche symbolique de la TA à base de règle, qui a dominé jusqu'au tournant des années 2000, et ces deux nouvelles approches numériques, basées sur des données d'entraînement et sur des calculs mathématiques permettant de manipuler ces données (POIBEAU 2019, p. 25).

Bien que la manière d'effectuer ces calculs diverge entre la TAN et la TA statistique, il s'agit dans les deux cas de méthodes statistiques qui n'ont en elles-mêmes rien de « magique » et qui permettent d'effectuer des prédictions sur la base d'observations préalables. Pour preuve, « la base mathématique de ce que l'on nomme IA n'a rien de compliqué et trouve ses sources dans les mathématiques du début du XVIII siècle [10] » (ZYDRON 2019, p. 137). Cette vision pragmatique du *deep learning* tranche de ce fait avec l'idée d'une technologie miraculeuse et imprévisible, qui ne signifie pas pour autant que les outils qui en sont issus n'ont aucune utilité, mais que l'on peut en prévoir les limites, les cas d'application utiles et peut-être même des améliorations. Par exemple, cette approche entièrement basée sur le traitement des corpus reste naturellement soumise aux mêmes exigences de dépendance à l'égard des données que les autres méthodes statistiques précédemment étudiées par le TAL. Une telle vision de la TA souligne en outre la nécessité d'adapter les outils à la tâche demandée et, surtout, laisse place à une réelle collaboration entre l'humain et la machine.

Pour toutes ces raisons, ce travail de recherche cherche à adopter un juste milieu et à promouvoir un « discours scientifique raisonné » concernant la technologie (LOOCK 2019, par. 2), en accord également avec la « conception critique de l'automatisation » suggérée par ROSSI (2018, p. 98). Ceci implique, comme je le mentionnais, de ne pas poser de jugement a priori qui se solderait par un rejet catégorique de la technologie, empêchant dès lors d'anticiper les risques associés à l'arrivée des outils ou d'en prévoir une utilisation différente qui serait, elle, au bénéfice des traducteurs. À l'inverse, il est risqué — et même préjudiciable — d'exagérer les performances de la machine (TORAL *et al.* 2018a, p. 122; LÄUBLI *et al.* 2018; WAY 2020, p. 325–326; POIBEAU 2022). Aucun de ces deux extrêmes ne servira ni la profession, ni le développement de la TA. Il est d'ailleurs rassurant, à cet égard, de voir que des associations et collectifs s'emparent à présent de ces questions par l'organisation de tables rondes (ATLAS et ATLF 2022) et la publication de numéros de revue thématiques (CEATL 2020; 2023), ou encore par la création d'un Observatoire de la traduction automatique, visant à suivre l'évolution des performances de la TA en littérature⁵, auquel on trouve d'autres équivalents, comme le projet *Kollektive Intelligenz – Kann KI Literatur*⁶. Plus encore, je soutiendrai en fin d'ouvrage que cette étape est aujourd'hui primordiale, à présent que les scientifiques ont pu prendre le sujet en considération et ont livré de précieuses observations que je détaillerai dans le prochain chapitre, mais aussi tout au long de ce travail, en proposant à mon tour une expérience concrète de traduction littéraire automatique.

5. <https://www.atlas-citl.org/observatoire-de-la-traduction-automatique/>.

6. <https://kollektive-intelligenz.de/>.

On le voit, les recherches offrent une vision plus nuancée de cette IA, qu'il faut essayer de mettre en avant. Cela ne veut pas dire pour autant que la TA est tout à fait inutile ou qu'il faut revoir à la baisse toutes les annonces concernant les avancées obtenues en la matière. Au contraire, si les exagérations abondent dans les discussions publiques, on trouve aussi, à l'inverse, d'autres personnes qui rejettent totalement la technologie ou associent simplement celle-ci à un engouement passager, entravant ainsi l'avancement de la recherche. S'il est en effet nécessaire de mettre en lumière certaines limites de la TA, telles que la dépendance aux corpus d'apprentissage, celles-ci peuvent aussi nous aider à avancer dans la bonne direction, parfois même dans des domaines insoupçonnés, ainsi que je l'ai montré dans le chapitre précédent pour la TAO, où l'on n'a pas même cherché à développer des outils qui pourraient pourtant être utiles à la traduction littéraire. De la même manière, je chercherai à voir dans ce travail si la TA, sans être parfaite, ne peut pas s'avérer utile tout de même, voire être repensée, tout comme les logiciels de TAO, d'une façon qui soit plus profitable pour les traducteurs littéraires. Ces considérations prendront une place plus importante dans la troisième et dernière partie de l'ouvrage, qui rétablira le lien entre TA et TAO, en envisageant cette première comme un outil d'aide à la traduction supplémentaire dans la perspective plus large de la traduction littéraire assistée par ordinateur.

Si je compte envisager la chose de façon décontextualisée au départ, pour poser un aperçu objectif sur les capacités de l'outil, il ne sera nullement question de s'en tenir par la suite. De la même manière que la recherche n'est jamais dénuée de valeurs, de finalités ou de conséquences, comme en atteste l'introduction de chartes et de formations sur l'éthique, l'intégrité ou la responsabilité scientifique (COUTELLEC 2019), aucun dispositif technologique n'est jamais tout à fait neutre non plus (HABERMAS 1973). Pour cette même raison, la fin de ce travail sera consacrée aux discussions qu'il conviendrait d'entamer face à l'introduction de nouveaux outils et à l'automatisation des métiers, en particulier dans ce contexte où les défis qui se posent pour la traduction depuis le début des années 2000 avec l'arrivée de la TA s'ouvrent maintenant bien plus largement avec l'arrivée des outils de génération d'image ou de texte.

Deuxième partie

Une question de données avant tout

La vraie machine littéraire sera celle qui sentira elle-même le besoin de produire du désordre, mais comme réaction à une précédente production d'ordre; celle qui produira de l'avant-garde pour débloquer ses propres circuits, engorgés par une trop longue production de classicisme. Et, de fait, étant donné que les développements de la cybernétique portent sur les machines capables d'apprendre, de changer leurs propres programmes, d'étendre leur sensibilité et leurs besoins, rien ne nous interdit de prévoir une machine littéraire qui, à un moment donné, ressente l'insatisfaction de son traditionalisme et se mette à proposer de nouvelles façons d'entendre l'écriture, à bouleverser complètement ses propres codes. [...] Telle serait une littérature capable de correspondre parfaitement à une hypothèse théorique, c'est-à-dire, en fin de compte, *la littérature*.

ITALO CALVINO (1984)

État de l'art

Pour entamer cette deuxième partie d'ouvrage, avant d'en venir aux détails de mon propre travail, il est indispensable de replacer ces recherches dans le contexte de l'évolution croissante des outils de traduction automatique. Bien que cette section soit uniquement abordée du point de vue de la traduction *littéraire* automatique, il me semble pertinent d'y dresser un bref panorama de la TA dans son ensemble, à la fois pour souligner le très long et fastidieux travail de recherche ayant finalement permis d'atteindre les performances des systèmes qui ont aujourd'hui la part belle dans les médias, et pour mettre en évidence également le niveau d'attention porté au cas littéraire depuis l'arrivée de la dernière génération de systèmes. Il ne sera donc pas question ici de retracer l'historique complet de la traduction automatique ou la chronologie des travaux à son sujet. D'une part, parce qu'il faudrait commencer par voir jusqu'où remonter exactement : aux langues universelles combinatoires du Moyen Âge (LEBON 2022), aux dictionnaires de la Renaissance, « l'élément premier de la mécanisation de la traduction » selon Edmond CARY (1956, p. 103), aux brevets précurseurs de Petr Petrovich Trojanskij et de Georges Astrouni (MERTEN 2022, p. 173) ou encore à la contribution d'Andrew Donald Booth, qui conseilla et conforta Warren Weaver dans ses recherches (MOUNIN 1964, p. 18–19) ? D'autre part, parce qu'il existe déjà des ouvrages tout à fait complets et à propos sur le sujet⁷. À la place, nous repartirons de chacun de ces paradigmes (traduction automatique à base de règle, traduction statistique, traduction neuronale) pour retracer les recherches menées spécifiquement sur la traduction littéraire automatique (TLA) et rendre compte de l'ampleur du phénomène au cours des dernières années.

Si ces différents paradigmes sont utilisés comme des points de repères temporels distinctifs, avec des dates correspondant à la mise sur le marché d'un système ou à la publication de l'article fondateur de telle ou telle approche, il faut bien noter que ces catégories ne sont pas aussi rigides. On en voudra pour preuve que les systèmes issus d'un nouveau paradigme mettent généralement du temps avant de supplanter totalement ceux qui les précèdent : des systèmes à base de règles comme *Apertium* (FORCADA *et al.* 2011) sont toujours d'usage pour les langues peu dotées⁸

7. Cf. HUTCHINS (2000, 2001), KENNY (2018) et POIBEAU (2019).

8. Comprendre, ici, des langues pour lesquelles nous disposons de peu de ressources linguistiques informatisées, en ce compris les corpus nécessaires à l'élaboration des systèmes de TA.

telles que l'occitan ou le breton (KHANNA *et al.* 2021), tandis que des systèmes de TA hybrides (par exemple statistique/neuronal) peuvent atteindre de meilleures performances en tirant profit des deux types de système (OMNISCIEN 2020). Par ailleurs, les nouvelles méthodes rompent rarement de manière radicale avec les paradigmes antérieurs : les tables d'alignement utilisées pour la TA statistique ont aussi pu servir pour les modèles neuronaux (HUANG *et al.* 2018), tout comme les règles employées pour créer les premières machines de TA qui ont permis de mettre au point des systèmes statistiques hybrides, plus performants, et dont l'injection dans les systèmes neuronaux reste un thème d'actualité bien que la projection de raisonnements linguistiquement motivés sur des modèles purement statistiques soit sujette à débat (MIELKE *et al.* 2019).

Un des points sur lesquels on s'accordera sans trop de difficulté est probablement l'année 1949, date à laquelle Warren Weaver publie le célèbre mémorandum qui lança un vaste programme de recherche sur la TA et qui attirera à sa suite d'autres noms, parmi lesquels Yehoshua Bar-Hillel et Léon Dostert (HUTCHINS 2001, p. 9). C'est à l'initiative de ce dernier que prend place un autre événement majeur : la démonstration de Georgetown de 1954 qui provoquera la première vague d'engouement et de battage médiatique autour de la TA, mais qui contribuera aussi singulièrement à promouvoir ces recherches et, par là même, à assurer leur financement (HUTCHINS 2004), du moins à jusqu'en 1966.

À cette date, l'ALPAC (*Automatic Language Processing Advisory Committee*), un comité scientifique chargé d'évaluer les avancées en matière de linguistique informatique et de traduction automatique, publie un rapport qui mettra fin à l'optimisme des premières années et coupera la recherche dans son élan ; ou tout du moins son financement, aux États-Unis (HUTCHINS 2003). Dans ce rapport, le comité conclut que la TA est plus chère et plus lente que la traduction humaine (ALPAC 1966, p. 19) et que, comme il ne semble pas y avoir de perspectives d'amélioration sur le court terme, les fonds devraient être redirigés vers une autre technologie : la traduction assistée par ordinateur, ou TAO (*Ibid.*, p. 32). Seuls quelques projets de recherches persistèrent par la suite, en particulier au Canada et en Europe, à commencer par la France.

En effet, la France restera un des seuls pays à poursuivre activement les recherches dans les années qui suivirent directement la publication des conclusions de l'ALPAC. Les recherches menées aux États-Unis y avaient motivé la création de l'ATALA (Association pour l'étude et le développement de la traduction automatique et de la linguistique appliquée), autour de la personne d'Émile Delavenay (LÉON 2002, par. 18–22), puis du CETA (Centre d'études pour la

traduction automatique), sous l'impulsion du CNRS, et de sa branche grenobloise placée sous la direction de Bernard Vauquois : le CETAG (*Ibid.*, par. 38–39), qui donnera naissance par la suite à l'actuel groupe GETALP (Groupe d'étude en traduction automatique/traitement automatisé des langues et de la parole). Quelques projets prirent forme dans d'autres pays, y compris en Belgique à Louvain, au travers de personnes comme Flip Droste, Dirk Geens ou Herman Caeyers, mais aussi à Liège, avec la participation de Jacques Noël au projet Eurotra, ou à Mons, par l'intermédiaire du projet METAL (VAN DER BEEK et VAN DEN BOSCH 2015, p. 353–354).

4.1 Les prémices de la TA : l'intérêt précurseur des littéraires

Tous ces premiers systèmes étaient représentatifs de ce que l'on a appelé la traduction automatique statistique à base de règle (TABR). Plusieurs approches ont été proposées depuis les années 1950 jusqu'aux années 1980–1990, allant vers des méthodes de plus en plus complexes, comme le synthétise le célèbre triangle de Vauquois.

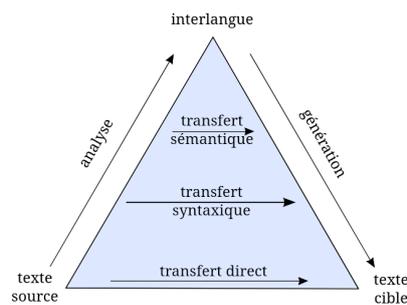


FIGURE 6 – Triangle de Vauquois

SOURCE : Wikipédia, CC BY-SA 3.0

Les premiers outils opéraient à partir de dictionnaires, d'une manière proche du mot-à-mot, pour produire « directement » une traduction vers la langue cible (POIBEAU 2019, p. 22–23). Par la suite, des séries de règles syntaxiques ont été ajoutées — d'où le nom de l'approche —, permettant au programme d'analyser la structure d'une phrase source afin de la convertir d'une manière plus naturelle dans la langue cible (FORCADA 2023, p. 51). D'autres méthodes, plus ambitieuses, visaient à prendre en compte des éléments de contexte pour inférer le sens des mots, tandis que les systèmes dits « interlingues » devaient passer par l'intermédiaire de représentations formelles indépendantes des deux langues, bien que la difficulté de cette modélisation leur donna peu de succès en pratique (POIBEAU 2019, p. 24–25). Dans tous les cas, ces approches nécessitent la contribution cruciale de linguistes qui doivent préparer en amont ces dictionnaires et ces règles de transfert, fournis par la suite au programme de traduction (KENNY 2012, p. 121–122).

À titre d'exemple, voici les étapes suivies par le système *Apertium* (FORCADA *et al.* 2011) pour générer une traduction.

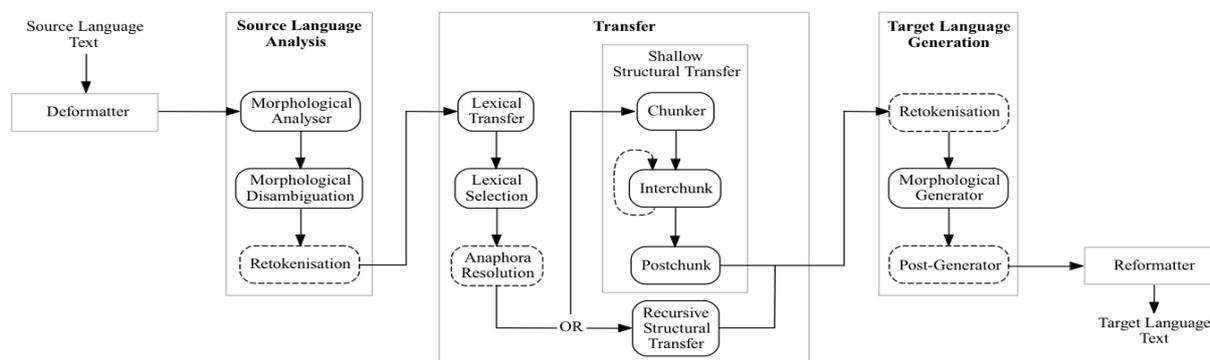


FIGURE 7 – Architecture du système *Apertium*

SOURCE : KHANNA *et al.* (2021, p. 479), CC BY 4.0

La nature contrainte et réglementée, quasi analogique, de cette technologie naissante a tôt fait d'attirer l'attention des littéraires et des traducteurs, qui voyaient dans les productions mécaniques certaines vertus esthétiques. Dans la francophonie, notamment, de nombreuses personnalités telles que Georges MOUNIN (1964), qui consacre la seconde partie de sa thèse citée plus haut à la TA, ou Paul BRAFFORT (1968), qui se penche sur les recherches en intelligence artificielle dès ses débuts, contredisent manifestement l'idée d'une incompatibilité totale entre les lettres et la machine. Comment ne pas voir en effet un certain rapprochement entre les règles informatiques et la nature combinatoire des approches proposées par certains membres de l'Oulipo (Ouvroir de littérature potentielle)⁹, à commencer par les *Cent mille milliards de poèmes* de son cofondateur Raymond QUENEAU (1961), les *Deux cent quarante-trois cartes postales en couleurs véritables* de Georges PEREC (1989) et bien d'autres exercices du même style, ou encore *La Machine littérature* imaginée par Italo CALVINO (1984).

Une première version de cette idée fut d'ailleurs concrétisée et présentée dans l'ouvrage *La Machine à écrire* de Jean BAUDOT (1964), directement inspiré du poème combinatoire de Queneau. Les poèmes trouveront pour leur part un premier portage sur ordinateur par Dimitri Starynkevitch en 1961, soit peu après la parution de l'ouvrage et la création de l'Oulipo la même année, puis une seconde programmée par Paul Braffort et présentée au festival belge Europalia en 1975, quelques années donc après une autre démonstration informatisée d'*Un Conte à votre façon*, en 1969 à Verviers, par François Le Lionnais (BLOOMFIELD *et* CAMPAIGNOLLE-CATEL 2016, p. 326)¹⁰.

9. Cf. <https://oulipo.net/>.

10. Voir l'Annexe A pour quelques affiches issues de ces rencontres.

Dans ce climat d'intérêt grandissant se forme dès lors une « petite cosmogonie littéraro-informatique (ou informatico-littéraire) » (MAGNÉ 2000, p. 120), qui multipliera les réalisations et interventions de mécanisation littéraire :

Après la Seconde Guerre mondiale, calculateurs et manipulateurs mécaniques cèdent le terrain, grâce aux progrès rapides des technologies, aux calculatrices électroniques que les Français baptiseront, un peu plus tard, « ordinateurs ». Malgré la lenteur et les difficultés de programmation des nouvelles machines (avec leurs cartes perforées et leurs tambours magnétiques), il était naturel de tenter quelques expériences de nature littéraire. (BRAFFORT 2000, p. 101.)

De ces expérimentations multipliées naîtra d'ailleurs l'ALAMO (Atelier de Littérature Assistée par la Mathématique et les Ordinateurs)¹¹, sous les auspices de Paul Braffort et de Jacques Roubaud. Créé en 1981 comme prolongement informatique de l'Oulipo, l'ALAMO visait — jusqu'à sa récente dissolution en 2020 — à prolonger l'exercice oulipien de l'écriture sous contrainte par le biais des ordinateurs, tout en insistant sur les multiples aspects pédagogiques de ces productions pour l'apprentissage de la langue, de l'informatique et de la littérature. C'est dans cette optique de « création artistique assistée par les nouvelles technologies » (*Ibid.*, p. 99) que le groupe élabore divers programmes de génération automatique de textes littéraires avec contraintes¹², mais aussi des *littéraciels* tels que le LAPAL (langage algorithmique pour la production assistée de littérature) : des langages informatiques facilitant la création de programmes littéraires (BRAFFORT 1984).

Du côté de la traduction, la machine n'a pas non plus été en reste et n'a pas manqué d'être détournée à des fins créatives. Elle était notamment appréciée pour ses consonances avant-gardistes et les heureux hasards de poésie involontaire décrits par Julian DIBBELL (2000), au cours d'une traduction retour de William Butler Yeats, ou encore pour sa fidélité à l'esprit Dada (ROTHWELL 2009). Jusqu'à récemment, expliquent Dorothy KENNY et Marion WINTERS (2020, p. 125 et 128), la majorité des usages de la traduction automatique en littérature relevaient d'une même démarche ludique et expérimentale caractéristique des humanités numériques, où le bruit et les non-sens revêtent une certaine connotation esthétique.

La traduction automatique est également à l'origine de nombreuses créations artistiques, dans lesquelles les sorties de traduction sont considérées comme une sorte de « poésie trouvée », fruit d'une heureuse sérendipité due à la manipulation de « trucs » linguistiques, tels des matériaux dans les arts plastiques. Dans certains cas, la traduction automatique est prise pour sa nature déconcertante, agrammaticale, absurde et antinarrative, délivrant des traductions comparables au modèle non communicationnel et exotisant prôné par BENJAMIN (2012) et par VENUTI (2008) [11]. (KENNY 2020, p. 309.)

11. Cf. <http://www.alamo.free.fr/>.

12. Voir, par exemple, les *Rimbaudelaires* (<http://lapal.free.fr/alamo/programmes/rimbaudelaires.html>) ou les *Dizains de Bénabou* (<http://lapal.free.fr/alamo/programmes/dizains.html>).

Évidemment, ces exercices de « cybertraduction » n'ont pas cessé avec l'arrivée d'autres paradigmes de traduction automatique, qui suscite toujours des élans de réappropriation créative¹³, y compris — sans surprise — dans les travaux de l'Outranspo (Ouvroir de translation potencial)¹⁴. Avec l'arrivée de la TA statistique, cependant, la profession constate une avancée progressive du modèle de production combinant TA et post-édition (PE), amenant d'autres scientifiques à évaluer la performance de ces outils sur les textes littéraires, bien que l'intérêt de ces recherches s'oriente alors plus vers des fins de productivité plutôt que de créativité.

4.2 La traduction automatique statistique

Avec l'arrivée des méthodes probabilistes, les recherches sur la traduction littéraire automatique prennent une tournure plus concrète, toujours motivée par la rencontre entre les lettres et la linguistique informatique. Guy CHATY (2000) parle notamment des « difficultés émoustillantes » du traitement automatique des langues (TAL), qui offre un point de vue nouveau sur celles-ci :

Les recherches informatiques sur le TALN ont le mérite de nous faire retomber constamment sur la subtilité du langage naturel et sa complexité ; voilà qui est passionnant et décuple l'intérêt de ces recherches : découvrir les subtilités du langage naturel, même celles auxquelles on ne pensait pas. (*Ibid.*)

Une manifestation concrète de cet intérêt se retrouve notamment dans l'organisation annuelle du *Workshop on Computational Linguistics for Literature*, tenu depuis 2012 sous l'égide l'ACL (*Association for Computational Linguistics*)¹⁵. Comme le note son comité organisateur, le domaine littéraire offrait alors des perspectives de recherche rendues possibles en particulier par la multiplication des livres numérisés :

La littérature se distingue par de nombreux aspects des domaines traditionnellement étudiés par la linguistique informatique, comme la prose journalistique, les pages Web non structurées ou la parole. Or, la présence grandissante de la littérature numérique offre de nouvelles possibilités, mais présente aussi de nouveaux obstacles pour le traitement du langage : Quelles méthodes automatiques pourraient nous aider à trouver de nouvelles sources sur un certain sujet, à étudier un document ou un genre donné, à identifier la personne qui aurait écrit un texte anonyme ou encore à lire un livre rédigé dans une langue étrangère ? Comment les techniques d'analyse du langage pourraient-elles s'affranchir des limites du mot pour nous permettre de mieux comprendre le sens profond d'un texte littéraire, peu importe la date, l'endroit ou la culture de son origine [12] ? (ELSON *et al.* 2012, p. III.)

13. Cf. KENNY (2020, p. 309–310) et LEE (2011).

14. Cf. <http://www.outranspo.com/cybertraduction/>.

15. Aujourd'hui, l'atelier a fusionné avec une autre rencontre pour former un groupe d'intérêt spécial de l'association : <https://sighum.wordpress.com/events/latech-clfl-2023/previous-workshops/>.

Ces recherches adopteront toutefois une perspective plus proche du TAL et de son versant d'ingénierie que de l'exploration poétique, s'intéressant moins aux caractéristiques esthétiques qu'à la prise en charge de ces documents par la machine. La mention de la post-édition fera ainsi son apparition, guidant également le développement des recherches qui suivront.

Si l'on considère volontiers aujourd'hui que l'approche neuronale a rompu de manière décisive avec les paradigmes précédents et qu'elle a radicalement révolutionné le champ (cf. [Section 3.2](#)), la rupture la plus profonde se situe probablement toutefois entre la traduction automatique basée sur les règles et la traduction statistique (TAS), comme le note Philipp KOEHN (2009, p. 17). Face aux systèmes minutieusement mis au point grâce à la collaboration active de linguistes, les méthodes statistiques représentent en effet un véritable tournant, tant bien pour le développement des systèmes (HUTCHINS 2001, p. 23) que pour leur commercialisation (POIBEAU 2019, p. 194–196), de même que pour les retombées sur le quotidien des traductaires (MOORKENS 2017).

Leur arrivée est liée aux travaux pionniers de BROWN *et al.* (1988) pour IBM, qui mèneront aux modules *open source* GIZA (AL-ONAIZAN *et al.* 1999) et GIZA++ (OCH et NEY 2003), des programmes très largement répandus pour la mise au point des systèmes de TAS, puis aux premiers produits commerciaux, à savoir *Language Weaver* en 2002 et *Google Traduction* en 2006 (KOEHN 2009, p. XI et 119). L'approche a été rendue possible, par ailleurs, par l'augmentation de la puissance de calcul des ordinateurs et par la massification des données disponibles sur le Web (POIBEAU 2019, p. 25). En effet, le changement de paradigme s'explique aussi, comme je le mentionnais, par la disparition de l'écriture explicite des règles. À la place, on « entraîne » à présent les systèmes à partir de corpus bilingues (ceux-là mêmes qu'on retrouve dans les mémoires de traduction ou les concordanciers), en partant d'une idée très simple :

Les personnes qui conçoivent les systèmes de TA statistique s'appuient sur des traductions humaines, partant du principe que ces corpus contiennent les solutions adéquates aux problèmes de traduction. Et si ces solutions sont jugées comme étant adéquates, c'est précisément parce que ces traductions ont été réalisées par des humains [13]. (KENNY 2012, p. 122.)

Pour cette raison, la TA statistique porte aussi le nom de traduction automatique basée sur les corpus, lesquels comportent typiquement pour ce paradigme des centaines de milliers, voire quelques millions de phrases (FORCADA 2023, p. 52).

Expliqué brièvement, la TAS repose sur des probabilités — d'où, une fois encore, l'appellation de TA *probabiliste* en lieu de TA *statistique*. Pour être plus précis, deux types de probabilités sont calculées à partir des corpus utilisés : un premier type qui mènera à l'élaboration d'un *modèle de traduction* et un second qui constituera un *modèle de langue* (KENNY 2018, p. 36–37).

Mais avant de calculer ces probabilités, il est toutefois nécessaire de transformer les textes tels que nous sommes habitués à les lire pour permettre leur prise en charge par la machine. Dans un premier temps, il faut choisir et délimiter les unités sur lesquelles le travail sera effectué. Dans le plus simple des cas, on considérera chaque suite de caractère délimitée par un espace comme des unités distinctes que l'on pourra obtenir après un premier traitement appelé *tokenisation* (étape consistant à détacher les signes de ponctuation du reste, à enlever les majuscules si nécessaire, à séparer les composés dans des langues comme l'allemand...). En voici un exemple simple :

Avant tokenisation	(Alice n'avait pas la moindre idée de ce que voulait dire latitude ou longitude, mais ces grands mots lui paraissaient beaux et sonores.)
Après tokenisation	(Alice n' avait pas la moindre idée de ce que voulait dire latitude ou longitude , mais ces grands mots lui paraissaient beaux et sonores .)

FIGURE 8 – Exemple de tokenisation

Une fois le texte divisé en unités distinctes, une seconde étape d'*alignement* — le plus souvent automatique — intervient ensuite, permettant de dresser des correspondances avec plus ou moins de précision pour chaque mot des phrases sources et de leur traduction¹⁶. Voyons ici un exemple basique :

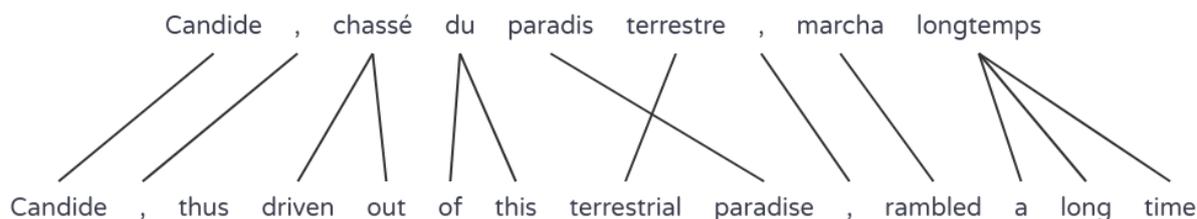


FIGURE 9 – Exemple d'alignement

Ces deux étapes effectuées, le programme est alors capable d'établir un *modèle de traduction*. Celui-ci pourrait être rapproché d'une sorte de dictionnaire bilingue contenant des termes dans les deux langues, ainsi que la probabilité pour chaque paire qu'il s'agisse d'une traduction valide, ce que le programme aura pu estimer en fonction du nombre de fois où un terme aura été aligné avec un autre dans le corpus. À titre d'illustration, voici la probabilité que le terme *étrange* soit une traduction des termes originaux dans « La Chute de la maison Usher » d'Edgar Allan Poe :

16. Notons une fois encore que ces textes sont préalablement alignés au niveau de la phrase, de manière à pouvoir être parcourus et comparés de la sorte par la machine.

Traduction	Original	Probabilité
étrange	wild	0.736782
étrange	wildly	0.175478
étrange	strange	0.087739

FIGURE 10 – Modèle de traduction

À ce stade, le modèle de traduction statistique est capable de générer des traductions, mais il n'est pas garanti que les termes choisis soient les plus idiomatiques. C'est ici qu'intervient le *modèle langue*, qui estime simplement la probabilité qu'un terme donné apparaisse dans la langue cible, ce que l'on peut calculer à partir de leur fréquence d'apparition dans un corpus monolingue. Pour donner un dernier exemple, un compte simple montre que le terme *noir* apparaît deux fois plus souvent que *rouge* dans le roman de Stendhal : il y aurait donc deux fois plus de chances que ce terme suive un segment tel que *j'aime le [...] s'il s'agissait de notre corpus de référence*. En résumé, nous pouvons donc dire que la traduction d'une phrase donnée dépend de la probabilité des traductions de chaque unité qui la compose, mais aussi de la probabilité que chacun des items choisis soient des termes communs dans la langue cible.

La prise en compte d'unités simples telles que nous les avons décrites jusqu'ici montre toutefois rapidement ses limites. À ces modèles fondés sur des alignements de « mots » — tout en sachant qu'il ne s'agit pas réellement de *mots*, le terme lui-même étant déjà compliqué à définir —, ont succédé des modèles fondés sur des alignements de segments (*phrase-based machine translation*). Évidemment, il ne faut pas voir dans ces segments des unités linguistiquement motivées, comme elles pouvaient l'être pour la TA à base de règles, mais bien une fois encore de groupements statistiques (KOEHN 2009, p. 128). On appelle typiquement ces segments des *n-grammes*, à savoir des groupements de deux, trois, quatre... — souvent trois en pratique — termes consécutifs. À partir de ces unités plus larges, on peut alors reproduire les mêmes étapes d'alignement et de calcul des probabilités décrites plus haut, quoiqu'avec des méthodes légèrement différentes¹⁷.

Unigrammes	En, quatre-vingts, jours, seulement
Bi-grammes	En quatre-vingts, quatre-vingts jours, jours seulement
Tri-grammes	En quatre-vingts jours, quatre-vingts jours seulement

FIGURE 11 – Suites de n-grammes

17. Pour plus de détails sur le fonctionnement de ces systèmes, se référer à l'ouvrage plus complet de KOEHN (2009) ou à l'article de vulgarisation de HEARNE et WAY (2011).

4.2.1 *Les premiers travaux*

Mis à part certains projets de TABR développés avec succès, dont quelques-uns de ceux mentionnés dans la section précédente, ces systèmes de « traduction par les nombres » (KOEHN 2009, p. 18) sont les premiers à connaître une utilisation aussi large et à s'installer durablement dans les habitudes, y compris du grand public. C'est d'ailleurs ce qui a motivé certains des premiers travaux sur le thème de la traduction littéraire automatique. Néanmoins, si la majorité des études se penchent sur la prose en partant de ce postulat, il est intéressant de remarquer que les recherches se sont tout d'abord intéressées au cas de la poésie.

Dans l'un des premiers travaux sur la question, GENZEL *et al.* (2010) ont pu concevoir un système de TA statistique français-anglais obéissant à des contraintes de longueur, de métrique accentuelle et de rime. Entraîné sur des données hors domaine — des données non littéraires, donc —, celui-ci parvenait raisonnablement à satisfaire les demandes formelles, au détriment de la qualité. De la même manière, GREENE *et al.* (2010) ont produit deux systèmes pareillement contraints par la métrique, qui avaient pour objectif de générer et de traduire des poèmes. Pour le système de traduction italien-anglais, entraîné à la fois sur de la poésie et des données hors domaine, l'équipe avait remarqué qu'il produisait parfois des traductions semblables à un segment de référence ou à des fragments de plusieurs références, parfois de nouvelles traductions valides, et qu'il pouvait être aisément modulé du point de vue de la forme, bien que le résultat n'était pas toujours fluide.

Ces tentatives de traduction automatique de poèmes ont plus récemment trouvé des échos grâce aux progrès apportés par la traduction automatique neuronale, que j'aborderai en détail dans la section suivante, tout en me concentrant sur la prose. Pour des raisons d'exhaustivité, cependant, on peut toutefois mentionner les travaux de Zakaryia Mustafa ALMAHASEES (2017) sur la traduction anglaise d'un recueil de poèmes à partir de l'arabe et de Philippe HUMBLÉ (2019) sur la traduction de poèmes de l'anglais vers le portugais, tous deux par *Google Traduction* (Google 2006), mais aussi le système de traduction français-anglais entraîné à partir d'un corpus parallèle de chansons de GHAZVININEJAD *et al.* (2018), l'étude de CHAKRABARTY *et al.* (2021) montrant les gains de performance offerts par les modèles multilingues entraînés sur de la poésie, la méthode proposée par KHANMOHAMMADI *et al.* (2023) pour produire des poèmes en vieux perse à partir de prose persane, ou encore le cas de la génération de poésie pouvant pareillement s'appliquer à la traduction de Tim VAN DE CRUYS (2020) ou de Jianli ZHAO et Hyo Jong LEE (2022). La traduction de poèmes représente toutefois une minorité des recherches sur la traduction littéraire automatique. La plupart d'entre elles se concentrant en effet sur la prose, c'est donc ce versant de la TLA qui nous occupera dans la suite de l'article.

Comparant justement la traduction automatique de prose et de poésie depuis le français vers l'anglais à partir de *Google Traduction* et d'un système développé avec *Moses* (KOEHN *et al.* 2007) sur les quelque 8 millions de segments du corpus *Hansard* (SIMARD *et al.* 1993), Ruth JONES et Ann IRVINE (2013) offrent une des premières études sur le sujet. Bien que les chercheuses insistent sur l'importance d'évaluer la production d'un point de vue qualitatif, elles offrent tout de même un score BLEU¹⁸ pour la poésie (16.62) et pour la prose (30.05). Dans l'analyse, il ressort que la TA effectue des choix intéressants, mais que les erreurs les plus importantes (choix des temps et choix lexicaux) soulèvent la nécessité d'entraîner le système sur des données littéraires. La littérature étant en effet un domaine de spécialité, les autrices précisent que cet impératif d'adaptation au domaine vaut donc tout autant que dans d'autres cas d'utilisation de la TA :

Tout comme les humains, les systèmes de traduction automatique sont capables de produire des traductions qui peuvent être tantôt littérales, tantôt plus libres, et doivent constamment opérer des choix lors du décodage [pour produire une phrase de sortie]. Dans le cas des systèmes SMT [et NMT], ces choix sont dépendants des observations contenues dans les données d'entraînement et de leur fréquence. Lorsqu'ils sont entraînés sur des jeux de données semblables au corpus de test, il est donc probable qu'ils effectuent des choix plus pertinents [14]. (*Ibid.*, p. 96.)

Ces études s'inscrivent par ailleurs généralement dans une tradition visant à recourir aux capacités de la machine pour mieux comprendre les caractéristiques des textes littéraires, ainsi que je le mentionnais en début de section en référence à l'atelier de linguistique informatique pour la littérature. C'est le cas par exemple de Rob VOIGT et Dan JURAFSKY (2012), dont l'intérêt porte justement sur les difficultés que posent les textes en prose et que doit surmonter la TA (tout comme les humains) pour traduire ce type de textes. Les auteurs relèvent à cet égard l'importance de la cohésion textuelle pour les textes littéraires, montrant que ceux-ci affichent des chaînes référentielles bien plus denses que d'autres domaines tels que la presse. Partant, les auteurs soutiennent que les recherches futures sur la TA devraient non seulement pouvoir compter sur des données d'entraînement adaptées au domaine, mais aussi prendre en compte les éléments textuels dépassant le cadre de la phrase, pour que celle-ci soit véritablement efficace en littérature.

18. BLEU (PAPINENI *et al.* 2002) est une méthode d'évaluation automatique de la qualité, censée être proportionnelle à la hauteur du score. Cette méthode est largement critiquée aujourd'hui, mais elle reste la méthode la plus largement répandue et cette longue tradition d'utilisation ajoute à sa présence systématique dans les travaux sur la TA (KOEHN 2020, p. 64), y compris dans le présent travail. Il faut toutefois bien noter que la valeur absolue de BLEU n'a pas de sens en elle-même et qu'elle n'est pas directement comparable avec les scores d'autres études, même si elle offre un vague point de référence. Or, celui-ci est malheureusement bien souvent la seule manière de se faire une idée de la qualité des traductions, puisque celles-ci sont rarement partagées et que les évaluations se limitent parfois à cette seule mesure. J'offre donc dans cette section les conclusions et résultats tirés dans les différentes études, et m'attarderai par la suite sur cette question de l'évaluation et des métriques (cf. Section 6.1).

Or, on peut trouver des essais d'adaptation à la prose dans deux autres cas d'étude pour cette génération d'outils, à commencer par celui de Laurent BESACIER (2014). Dans ce premier cas, une nouvelle est traduite de l'anglais vers le français (*The Book of Me*, de Richard Powers), une fois encore à l'aide du système *Moses*, que le chercheur entraîne d'abord sur 25 millions de phrases hors domaine issues de la campagne IWSLT (FEDERICO *et al.* 2012), puis affine en y injectant le texte post-édité à mesure de son avancement. Chaque tranche résultante est alors utilisée pour adapter progressivement le système, qui obtient un score BLEU situé aux alentours de 39 (en comparant la sortie avant et après post-édition). S'il semble que les données ne sont pas suffisantes dans cette étude pour mener à une amélioration du résultat, l'auteur note en revanche une diminution du temps de PE au fil de l'exercice, bien qu'il soit difficile de savoir si elle est due à l'adaptation du système ou à la familiarisation progressive de la PE non experte avec la tâche. La durée totale de la tâche est tout de même fortement réduite par rapport à une traduction libre, ce qui aurait l'avantage de réduire le coût de la traduction. Le chercheur évoque néanmoins la nécessité de s'assurer que la qualité n'y est pas sacrifiée, car la TA a tendance à copier le texte source et ne tient pas compte des références culturelles.

Dans le second cas, Antonio TORAL et Andy WAY (2015b) évaluent la traduction en catalan du roman espagnol *El prisionero del cielo*, de Carlos Ruiz Zafón, après avoir comparé les caractéristiques textuelles de quatre autres romans vis-à-vis d'autres domaines pour ce même couple de langues. Le système utilisé est conçu une fois encore avec *Moses* à partir d'un corpus bilingue maison d'articles de presse (environ 630 000 segments), auquel s'ajoutent quelque 16 millions de segments du corpus Web monolingue *caWac* (LJUBEŠIĆ et TORAL 2014). Pour l'affinage, ce sont deux romans du même couple auteur-traducteur qui sont ajoutés (environ 22 000 segments), préfigurant en ce sens les modalités du travail que je mène ici. Il apparaît d'ailleurs que la quantité des données et l'adaptation sur la base de romans écrits et traduits par les mêmes personnes améliorent sensiblement les résultats (jusqu'à 53.76 points BLEU, contre 49.15 pour le système maison entraîné uniquement sur le corpus de presse et 46.52 pour *Google Traduction*). Les auteurs indiquent également que 20 % des phrases produites par le meilleur système sont en tous points identiques à la référence, bien qu'il s'agisse des phrases les plus courtes (7,15 % du total de mots), et que 10 % de phrases supplémentaires ne diffèrent de la référence que par cinq caractères ou moins. Par ailleurs, deux locutaires bilingues en charge de l'évaluation d'une centaine de segments rapportent avoir trouvé environ 40 % des phrases de la TA équivalentes à la référence, et presque 20 % de meilleure qualité — en réalité à cause des prises de liberté du traducteur —, pour un total de 60 % des phrases affichant une qualité comparable à la traduction humaine. Au cours d'une étude complémentaire, TORAL et WAY (2015a, p. 247) concluent dès lors que l'adaptation au domaine et la proximité du couple de langue à traduire sont déterminants pour la

performance de la TA en littérature, ajoutant que la traduction risque d'être le prochain levier de réduction de coûts après le passage aux livres numériques.

Enfin, pour terminer ce tour d'horizon consacré aux premiers travaux de recherche sur la traduction automatique de la prose, il me faut mentionner l'étude plus récente d'Eoin Ó MURCHÚ (2019), qui s'est essayé à la post-édition d'un texte de science-fiction du gaélique écossais vers l'irlandais. *Intergaelic* (SCANNELL et MĚCHURA 2015), l'outil utilisé pour cette tâche, est un système hybride (TABR et TAS) disponible en ligne et entraîné sur un corpus relativement restreint. Sans surprise, le texte ainsi produit comporte à la fois des phrases parfaites en l'état et des passages qui devaient entièrement être retraduits, bien que le résultat soit largement correct d'un point de vue grammatical selon le post-éditeur. Par comparaison avec la traduction humaine produite par le même auteur, la TA atteint un score BLEU de 35, mais paraît moins naturelle par endroit et plus proche de la structure du texte source. De plus, le traducteur note que certains éléments ne sont pas traduits et qu'une intervention humaine est presque systématiquement requise dans le cas des noms propres, des régionalismes, des néologismes et des interjections. Il ajoute cependant que la PE est environ 31 % plus rapide et qu'elle lui a permis d'éviter quelques erreurs. Les réactions des maisons d'édition contactées et de l'auteur original vis-à-vis de ce projet ont par ailleurs été très positives et la traduction fut publiée par la suite¹⁹.

4.3 Le paradigme neuronal

Les premiers chapitres de ce travail ont permis d'établir que la différence entre la TAS et la traduction automatique neuronale (TAN) n'était pas aussi grande qu'on le laisse parfois entendre, et qu'elle repose en réalité sur les deux mêmes grands principes d'utilisation des statistiques et de corpus (cf. Section 3.3). Ce qui différencie fondamentalement les deux approches, c'est avant tout la manière dont ces systèmes repèrent les tendances statistiques et « apprennent²⁰ » les traductions probables. En effet, « tout comme les systèmes de TAS, la TA neuronale apprend à traduire grâce aux corpus parallèles, mais ces systèmes utilisent pour ce faire des méthodes de traitement informatique très différentes [15] » (KENNY 2018, p. 436). Pourtant, le battage médiatique autour de la TA donne bien souvent l'impression d'une apparition soudaine des réseaux de neurones et d'une explosion des performances de la TA, sans tenir compte de la longue réflexion qui a permis cette progression constante de la technologie.

19. *Tinte na Farraige Duibhe*, Leabhar Breac, 2020.

20. Notons que les termes usuellement utilisés pour décrire la TAN contribuent bien souvent aux tendances d'anthropomorphisation des outils (cf. p. 24), mais la métaphore ne s'arrête pas là, puisqu'elle concerne également les données, présentées comme des ressources naturelles qu'il faut récolter pour « nourrir » la machine, occultant ainsi le travail fourni par les traducteurs et les particularités de toutes ces traductions humaines individuelles (KENNY 2012, p. 128–129).

Les travaux sur ce que l'on nomme aujourd'hui *réseaux de neurones artificiels* ou *deep learning* remontent à la première moitié du siècle dernier, qui fut également témoin d'un conflit théorique entre les défenseurs de l'IA *connexionniste*, d'une part, et ceux de l'IA *symbolique* d'autre part. Le premier courant, dont sont issus les systèmes de TAN actuels, nous provient de la proposition pionnière de Warren McCULLOCH et Walter PITTS (1943) et du modèle théorique de Donald HEBB (1949), qui aboutiront à la réalisation du réseau de neurones de Frank ROSENBLATT (1957). Ces travaux seront toutefois abandonnés un temps en faveur des approches symboliques ayant mené à la TABR. Ce conflit, ajouté à la résurgence soudaine du connexionnisme par la suite (sous d'autres noms), n'a d'ailleurs sûrement pas manqué d'ajouter au phénomène de *hype* qui entoure depuis quelques années le *deep learning*.

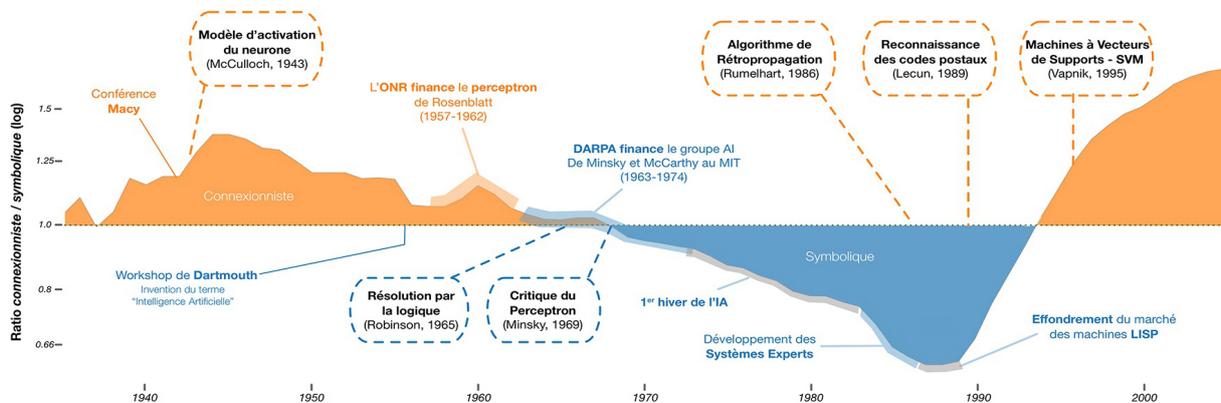


FIGURE 12 – Évolution des approches connexionniste et symbolique

SOURCE : CARDON *et al.* (2018, p. 181)

Si le développement et les explications sont bien souvent d'inspiration neurobiologique, force est de constater que la traduction automatique neuronale n'est que très vaguement liée au fonctionnement du cerveau humain (FORCADA 2017, p. 292). C'est le concept de neurone (artificiel) qui inspire évidemment le terme *neuronal*, bien que le terme ait été peu à peu délaissé en faveur du *deep learning*, de façon à éviter les parallèles douteux (KOEHN 2020, p. 30–31). Fondamentalement, un neurone artificiel n'est qu'une unité de traitement d'entrées (*inputs*) en une sortie (*output*), où chaque entrée x (un nombre réel) est traitée par ce que l'on appelle des poids (*weights*), un biais (*bias*) et une fonction d'activation (différentes fonctions sont possibles) pour produire un nouveau nombre représentatif de l'élément à prédire (un mot, une classe, etc.).

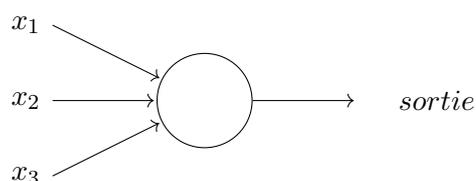


FIGURE 13 – Représentation simple d'un neurone artificiel

Mises ensemble, ces unités de traitement forment un réseau de neurones artificiels interconnectés composé de plusieurs couches : une couche d'entrée, une couche de sortie, ainsi qu'un nombre variable de couches cachées selon la taille du réseau. C'est d'ailleurs précisément l'addition de couches cachées qui permet d'atteindre de tels résultats aujourd'hui sur un grand nombre de tâches : plus il y a de couches, plus l'architecture est dite *profonde*, d'où le terme *deep learning* (« apprentissage profond »).

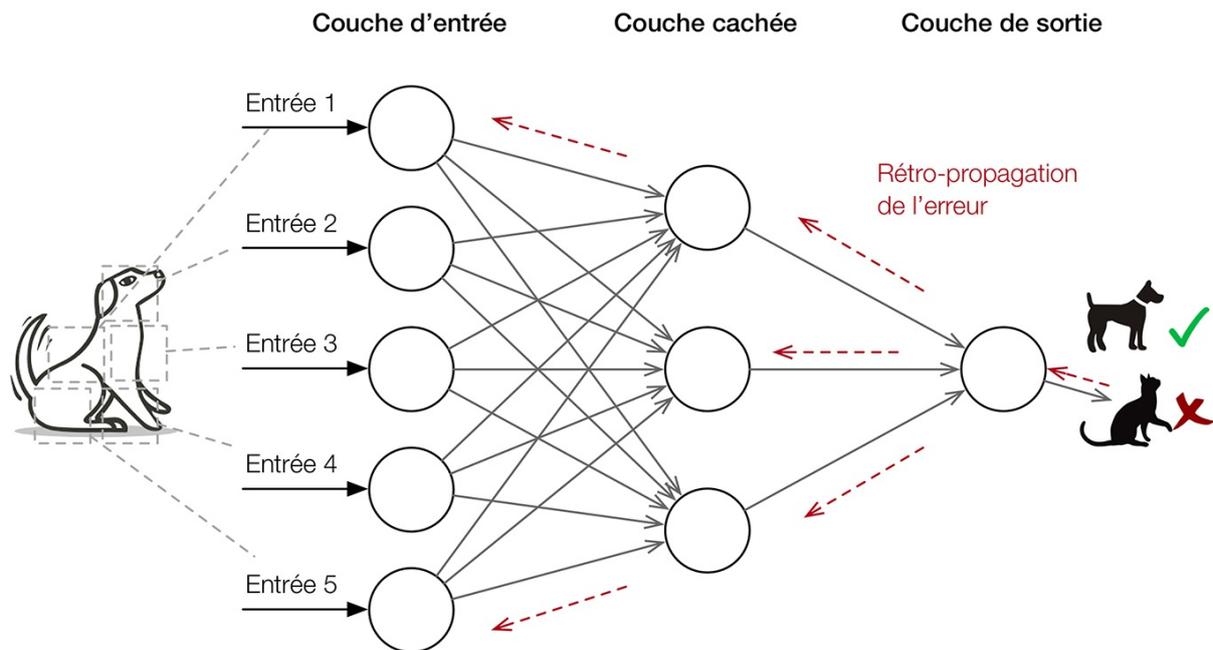


FIGURE 14 – Fonctionnement d'un réseau de neurones simple

SOURCE : CARDON *et al.* (2018, p. 199)

Un réseau de neurones artificiels est donc simplement un assemblage d'unités capable d'effectuer une prédiction (dans cet exemple, une classe). Pour lui permettre de faire cette prédiction, on présente au système un enchaînement continu d'entrées accompagnées des sorties à prédire (dans notre cas, des images segmentées en groupements de pixels appartenant à la classe « chien » ou « chat »), et on le laisse proposer une première solution. S'il se trompe, on modifie petit à petit les valeurs du réseau par un phénomène de rétropropagation, jusqu'à produire les sorties correctes : c'est la phase d'« apprentissage » ou d'« entraînement » du système. Pour chaque prédiction incorrecte, celui-ci actualise ainsi les valeurs des connexions entre les neurones (les poids précédemment évoqués qui déterminent la force des connexions entre chaque neurone), et ce, jusqu'à ce qu'il atteigne un plateau et que ses performances ne s'améliorent plus de manière significative.

Comme le résume parfaitement Kenny à partir de FORCADA (2017) :

L'activation de neurones individuels n'a pas tellement de sens en elle-même. Dans les faits, ce sera plutôt l'activation de grands ensembles de neurones connectés qui permettra de représenter des mots ainsi que leurs relations avec d'autres mots. Le tout est de réussir à entraîner le système de TAN de manière à trouver les poids exacts qui produiront le modèle de traduction le plus performant ; autrement dit, le modèle dont les valeurs d'activation lui permettent de prédire les meilleures traductions. Comment faire, donc ? À l'instar de toute autre méthode d'apprentissage automatique, le système apprend des données. Un système de TA neuronale se construit étape par étape, en exposant celui-ci à de vastes quantités de textes parallèles. Au fur et à mesure, l'algorithme ajuste ses poids en favorisant les prédictions qui se rapprochent le plus de la sortie « correcte » [16]. (KENNY 2022, p. 40.)

Cette étape d'apprentissage peut durer plusieurs jours, voire plusieurs mois. De fait, tout comme la TAS en son temps, l'approche neuronale a été rendue possible par l'augmentation de la puissance de calcul des ordinateurs ainsi que celle des données utilisées pour mettre au point ces modèles. Pour la TA en particulier, on trouve des propositions d'application des réseaux de neurones fort proches des méthodes actuelles dès les années 1980–1990 (KOEHN 2020, p. 39). Mais avec l'arrivée de matériel plus performant — notamment l'adoption de processeurs graphiques —, il devenait alors possible d'utiliser des réseaux pouvant comporter plus de couches et, dès lors, de réaliser des tâches plus complexes :

Les croisés du connexionnisme parviennent ainsi à convaincre qu'il est préférable de sacrifier l'intelligibilité du calculateur, et une optimisation rigoureusement contrôlée, à une meilleure perception de la complexité des dimensions présentes dans ces nouvelles données. Quand le volume des données d'entraînement augmente considérablement, il existe beaucoup de minimums locaux, mais il se forme assez de redondances et de symétries pour que les représentations apprises par le réseau soient robustes et tolérantes aux erreurs dans les données d'apprentissage. Au cœur de ce débat tendu avec la communauté du *machine learning*, un sous-entendu est omniprésent : il n'y a que dans les laboratoires que les modèles sont linéaires, le monde, le « vrai monde », celui des données produites par la numérisation des images, des sons, des paroles et des textes, lui, est non linéaire. Il est bruité, l'information y est redondante, les flux de données ne sont pas catégorisés derrière des attributs de variables homogènes, claires et construites de façon intelligible, les exemples sont parfois faux. (CARDON *et al.* 2018, p. 203.)

La particularité de cette approche, cependant, est à trouver dans les données manipulées par les modèles. En effet, les n-grammes des systèmes statistiques feront ici place à ce qui est assurément la véritable rupture vis-à-vis des paradigmes précédents : la représentation de mots à partir de vecteurs.

Les données utilisées pour l'apprentissage automatique doivent de toute évidence être manipulables par la machine. La solution trouvée pour les modèles neuronaux est celle du plongement lexical (*word embeddings*) : une représentation des unités lexicales sous une forme entièrement numérique (un vecteur), qui offre également l'avantage de prendre en compte un grand nombre d'informations. Concrètement, cette représentation vectorielle peut être obtenue par un processus d'analyse des cooccurrences — idéalement sur le plus grand nombre de documents d'une langue donnée — pour regrouper les mots qui apparaissent dans le même contexte. La méthode est d'ailleurs souvent donnée comme exemple de concrétisation de la thèse de FIRTH (1962) : « Vous connaîtrez un mot par ses fréquentations ! [17] ». Toutefois, la taille des données utilisées pour la tâche fait que cela fonctionne bien en pratique, ce qui laisse à penser que les plongements capturent des informations sur le sens, d'une manière ou d'une autre, puisque des mots similaires se retrouvent dans des sortes de « clusters sémantiques ».

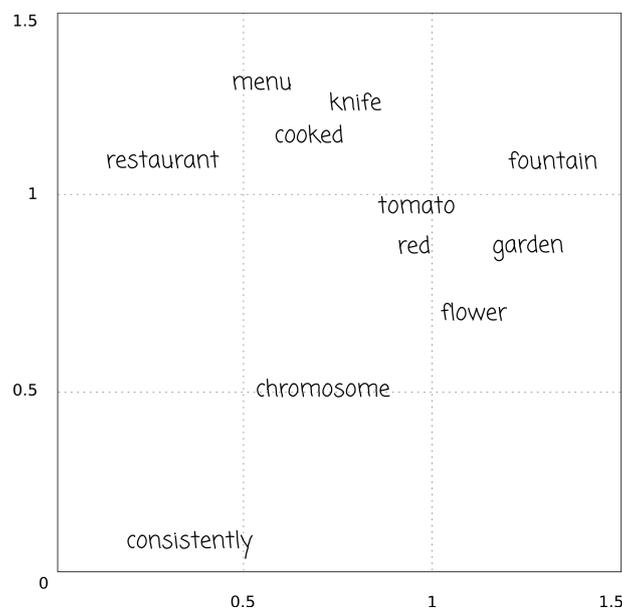


FIGURE 15 – Représentation vectorielle

SOURCE : PÉREZ-ORTIZ *et al.* (2022, p. 153), CC BY 4.0

Bien que l'exemple n'en montre que deux — plus facilement visualisables pour nos cerveaux humains —, il faut néanmoins imaginer que ces vecteurs comportent un nombre bien plus élevé de dimensions (souvent entre 50 et 1 000), censées chacune capturer des informations différentes. Ce sont ces dimensions (et leur valeur) qui font que les mots apparentés sont regroupés ensemble²¹.

21. On pourrait imaginer, pour reprendre les exemples de la Figure 14, que des vecteurs représentant les mots *chien* et *chat* seraient proches en tous points, excepté pour une dimension dont la valeur plus ou moins élevée distinguerait les canidés des félinés ou les grands des petits animaux. Il faut toutefois veiller à ne pas projeter des interprétations humaines logiques sur ces comptes purement statistiques, l'exemple étant uniquement donné ici à des fins d'explications.

Chaque unité lexicale est ainsi représentée par ses coordonnées, que l'on peut représenter sous une notation vectorielle (par exemple, [0.25, 1.1] pour le mot *restaurant* de cet exemple fictif), mais auxquelles on peut aussi et surtout appliquer très aisément des opérations arithmétiques (PÉREZ-ORTIZ *et al.* 2022, p. 151–152).

Si certains objets tels que les images se décomposent naturellement en vecteurs, d'autres objets nécessitent d'être « plongés » dans un espace vectoriel avant d'être susceptibles d'être calculés ou classifiés par les réseaux de neurones. Il en va ainsi du texte qui constitue l'exemple prototype. Pour faire entrer un mot dans un réseau de neurones, la technique Word2vec le « plonge » dans un espace vectoriel qui mesure sa distance avec tous les autres mots du corpus (MIKOLOV *et al.* 2013). Les mots héritent ainsi d'une position dans un espace de plusieurs centaines de dimensions. L'avantage d'une telle représentation réside dans les nombreuses opérations offertes par une telle transformation. Deux termes dont les positions inférées dans cet espace sont proches sont également similaires sémantiquement, on dit de ces représentations qu'elles sont distribuées : le vecteur du concept « appartement » [-0.2, 0.3, -4.2, 5.1...] sera proche de celui du « maison » [-0.2, 0.3, -4.0, 5.1...]. La proximité sémantique n'est pas déduite d'une catégorisation symbolique, mais induite des voisinages statistiques entre tous les termes du corpus. (CARDON *et al.* 2018, p. 204–205.)

Au cours de la phase d'entraînement, le système de TAN actualise donc ses valeurs, de manière à produire les représentations vectorielles les plus fiables possibles. Et comme ces représentations dépendent du corpus d'entraînement proposé au modèle, on comprend aisément l'importance d'avoir des données en grand nombre, pour que le système puisse généraliser au mieux, mais aussi représentatives du domaine en question, puisque cela influencera directement les choix, notamment lexicaux, qu'effectuera le système²².

4.3.1 Architectures

Dans les faits, les réseaux de neurones artificiels peuvent prendre différentes formes, que l'on appelle *architectures*. Le terme désigne ici le nombre de neurones ou de couches que comporte le réseau ainsi que d'autres *hyperparamètres*, mais aussi la manière dont celui-ci traite l'information entre l'entrée et la sortie. Pour représenter ces architectures, on part bien souvent de schémas rappelant ceux des théories de l'information :

22. Pour plus de détails techniques sur le fonctionnement des systèmes neuronaux, se référer à l'ouvrage plus récent de KOEHN (2020) ou aux contributions explicitement destinées aux non-initiés de FORCADA (2017) et de PÉREZ-ORTIZ *et al.* (2022).



FIGURE 16 – Modèle encodeur-décodeur

Dès le départ, les architectures proposées furent constituées de deux réseaux de neurones distincts : un *encodeur* et un *décodeur*, qui se chargent respectivement des deux étapes essentielles à la traduction automatique neuronale. Durant la phase d'encodage, la première moitié du système est chargée de transformer la phrase source en un vecteur. Au terme de cette première phase, toute la phrase d'entrée est représentée d'une certaine manière dans cet unique vecteur. Durant la phase de décodage, l'autre moitié prend alors le relais et, au départ de ce vecteur, tente de produire une phrase de sortie par prédictions successives de chaque mot. Ces deux réseaux pris ensemble forment ainsi l'architecture du système de traduction, qui fut proposée dès 2013 par KALCHBRENNER et BLUNSOM (2013). L'approche présentait néanmoins certaines limites qui l'empêchaient d'atteindre de bonnes performances par comparaison avec les modèles statistiques existants.

Peu de temps après, CHO *et al.* (2014) proposaient une nouvelle architecture, composée cette fois de deux réseaux de neurones récurrents (RNN, de l'anglais *Recurrent Neural Network*). Contrairement à d'autres architectures, les RNN peuvent prendre en compte les sorties précédentes au moment de prédire le prochain mot, ce qui fait qu'ils conviennent particulièrement bien au traitement de données séquentielles telles que le texte. La Figure 17 présente le déroulement temporel de cette méthode.

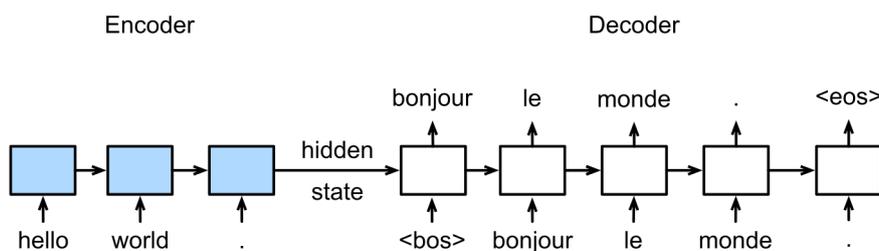


FIGURE 17 – Déroulement du modèle encodeur-décodeur dans le temps

SOURCE : A. ZHANG *et al.* (2023), CC BY-SA 4.0

Fondamentalement, cette approche consiste à encoder chaque token sous la forme d'un vecteur (chaque carré correspond au passage d'un token dans l'ensemble du réseau), qui permet d'obtenir ce que l'on appelle un « état caché ». Tous les tokens suivants sont ensuite traités de la même manière, en prenant chaque fois en compte l'état caché obtenu à l'étape précédente. Une fois le dernier mot traité, l'état final donné par l'encodeur (le vecteur représenté sur la [Figure 16](#) ci-dessus) est repris par le décodeur, qui procède exactement de la même manière en sens inverse, en produisant un premier token à partir duquel est construite la phrase cible. De cette manière, chaque mot a contribué à la prédiction finale, mais ce traitement purement linéaire présente cependant des limites et rencontre rapidement des difficultés lorsque que les phrases traitées deviennent trop longues, dans la mesure le modèle « oublie » les informations initiales.

Sur ce modèle SUTSKEVER *et al.* (2014) proposent d'améliorer l'architecture neuronale en traitant l'information au moyen de cellules LSTM (de l'anglais *Long Short-Term Memory*). La [Figure 18](#) schématise le fonctionnement du modèle LSTM, offrant une vue détaillée des opérations prenant place à chaque étape de la [Figure 17](#). Comme leur nom l'indique, ces cellules de traitement de l'information comparables aux neurones présentés plus haut sont capables de prendre en compte les mots précédant immédiatement le prochain terme à prédire à chaque étape de la traduction, mais également des informations distantes. On y retrouve dès lors les mêmes états cachés états cachés (mémoire à court terme), permettant de produire un vecteur de sortie (H_t) à partir des tokens d'entrée (X_t), auquel s'ajoute une deuxième sortie C_t (mémoire à long terme), contrôlée par trois « portes » correspondant à des couches de neurones spécifiques. Outre la gestion de dépendances à plus longue portée, cette méthode facilite l'apprentissage d'un point de vue technique (problème d'explosion et de disparition du gradient), ce qui améliore considérablement les performances de cette approche et en fait l'une des premières à être véritablement efficace.

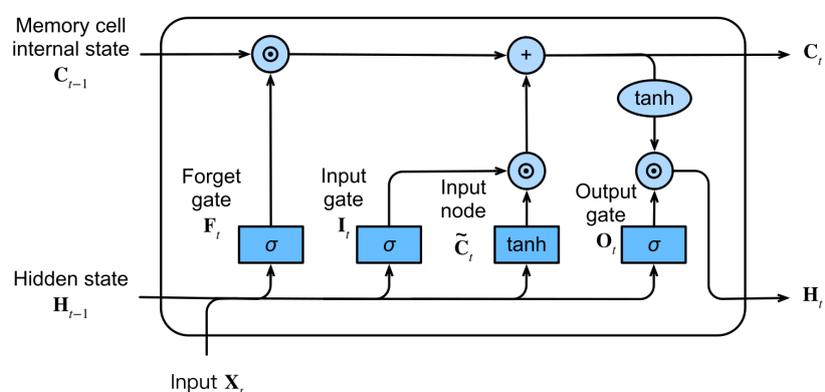


FIGURE 18 – Cellule LSTM

SOURCE : A. ZHANG *et al.* (2023), CC BY-SA 4.0

Une version plus tardive, appelée bi-LSTM, apportera une amélioration supplémentaire en implémentant le concept de bidirectionnalité, permettant au modèle d'effectuer une passe dans les deux sens et de prendre également en compte les mots à venir lors des prédictions, augmentant encore les performances de la méthode.

À peu près au même moment, une autre équipe propose d'ajouter au modèle encodeur-décodeur une autre fonctionnalité qui prendra le nom de *mécanisme d'attention* (BAHDANAU *et al.* 2015). Le mécanisme proposé peut être rapproché du modèle d'alignement de la TAS (cf. Figure 9), dans la mesure où l'algorithme permet de chercher les termes de la phrase source susceptibles de fournir des informations pour la traduction en cours. L'avantage de cet ajout est double, puisque ce pseudo-alignement aide à mieux capturer les relations entre les différentes parties de l'énoncé et facilite par ailleurs les opérations de calcul, ce qui améliore une fois encore les performances sur les phrases longues²³.

La dernière évolution de la TAN — qui n'en est pas des moindres — est arrivée avec l'architecture *Transformer*, suggérée par VASWANI *et al.* (2017)²⁴. Partant du mécanisme d'attention justement proposé, l'équipe de recherche abandonne entièrement les RNN et élabore à la place un système encodeur-décodeur basé uniquement sur le principe d'attention. Plutôt que de traiter successivement la phrase source et la phrase cible, le *Transformer* fait intervenir simultanément le côté encodeur et décodeur à chaque étape de la prédiction, qui se base également sur la position des mots dans la phrase (*positional encoding*). Ces informations (représentations vectorielles des entrées et sorties et vecteur de position) sont alors prises en charge par les couches d'attention et sont traitées conjointement au niveau du décodeur après normalisation (cf. Figure 19). Comme pour la méthode précédente, ce changement d'architecture réduira davantage le coût des calculs et améliorera la prise en charge des énoncés plus longs, mais il offre par ailleurs l'avantage considérable de pouvoir traiter la phrase dans son ensemble et non plus mot à mot²⁵. En l'espace de quelque temps, ce dernier changement de paradigme finira d'asseoir la domination du *deep learning* pour la plupart des tâches du TAL.

23. Les premiers systèmes de TAN ont été très critiqués au départ sur leur baisse de performance proportionnelle à la longueur des phrases, ce qui explique peut-être l'importance que prendront ensuite les mécanismes d'attention.

24. Ce sont ces mêmes *Transformers* qui sont à la base des grands modèles de langue tels que les GPT, qui en tirent leur nom (*Generative Pre-trained Transformer*).

25. Notons, à titre indicatif, que les cinq approches successivement exposées ici reposent sur la paire de langue anglais-français pour comparer leurs performances et que celle-ci est traditionnellement considérée dans le domaine comme étant l'une des configurations les plus « simples » (en raison de la proximité des deux langues et de la quantité de données disponibles pour entraîner les systèmes).

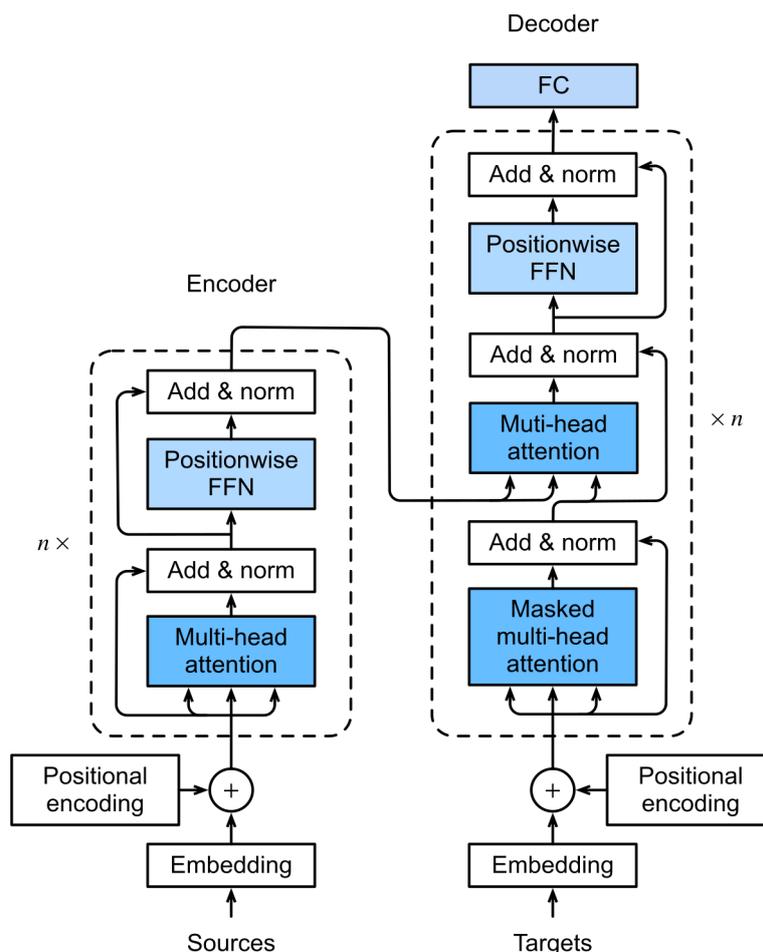


FIGURE 19 – Architecture *Transformer*

SOURCE : A. ZHANG *et al.* (2023), CC BY-SA 4.0

En résumé, l'histoire du *deep learning* et de la TAN peut être condensée et présentée comme suit : Bien que le travail sur les réseaux de neurones remonte loin en arrière, un premier tournant s'est opéré en 2012 dans le domaine de la reconnaissance d'images (KRIZHEVSKY *et al.* 2012), puis en 2014 pour la traduction automatique (SUTSKEVER *et al.* 2014 ; BAHDANAU *et al.* 2015). Dès 2016, l'ensemble des acteurs du marché passent au *deep learning* (POIBEAU 2019, p. 121), tandis que les *Transformers* (VASWANI *et al.* 2017) continuent d'accentuer le phénomène neuronal, marqué la même année par l'arrivée de *DeepL* (DeepL GmbH 2017)²⁶. Bien qu'il soit difficile d'en juger avec précision, il me semble que la traduction automatique a véritablement fait son entrée dans la sphère médiatique au même moment, devenant un sujet de discussion aussi commun que vendeur, loin d'être confiné aujourd'hui aux seules sphères de la traduction et de l'informatique.

26. Dès son apparition, l'entreprise reste volontairement opaque sur son système, mais indique toutefois qu'il s'inspire du *Transformer* (DEEPL GMBH 2021).

L'arrivée de la TAN marque donc un véritable tournant, tant pour le TAL au sens large que pour la traduction automatique de manière plus spécifique, mais aussi — à en juger par l'augmentation soudaine des publications sur le sujet — pour la traduction littéraire automatique. En conséquence, cet état des recherches sur l'application de la TAN en littérature sera divisée en trois parties, à commencer par les études proposant comme je le ferai ensuite une adaptation d'un système au domaine littéraire. Suivront dans la foulée les travaux visant à évaluer les outils existants dans ce même domaine et enfin ceux proposant des réflexions plus théoriques ou des enjeux sociétaux qui entourent la thématique, bien que ces différents intérêts soient rarement exclusifs.

4.3.2 *Mise au point de systèmes adaptés à la littérature*

Cette nouvelle approche neuronale, rendue possible par la manipulation de représentations vectorielles et par l'utilisation de mégacorpus (bien plus larges encore que pour la TA statistique), s'est rapidement imposée en raison du gain de qualité offert sur les sorties de traduction.

La traduction neuronale se situe dans la continuité des approches statistiques, mais améliore considérablement les performances en permettant une modélisation très riche de la sémantique des mots d'une part, et en permettant de procéder à des traductions directement au niveau de la phrase considérée comme un tout. On évite ainsi l'étape délicate d'assemblage de fragments de traduction, contrairement aux approches statistiques traditionnelles. La traduction neuronale repose par ailleurs sur des représentations riches du sens des mots : chaque mot est mis en rapport avec d'autres mots de sens correspondants, ce qui contribue à identifier des traductions possibles de manière plus robuste. (POIBEAU 2019, p. 26.)

Le changement de paradigme apporté par les systèmes neuronaux est ainsi réputé donner de meilleurs résultats sur les textes d'une plus grande richesse lexicale et d'une plus grande complexité syntaxique (BENTIVOGLI *et al.* 2016). Toutefois, les études pionnières en TAS pointent le besoin de spécialiser les systèmes sur des données littéraires (JONES et IRVINE 2013, p. 96). Les deux essais présentés pour la traduction automatique probabiliste, cependant, n'ont que peu de données à disposition, ce qui s'explique en outre par la faible quantité de corpus disponibles dans ce domaine. L'une des conclusions tirées par BESACIER (2014) est qu'il faudrait précisément pouvoir compter sur une plus grande représentativité des textes littéraires dans les données d'entraînement, et ce besoin pourrait être d'autant plus grand que la TAN est particulièrement gourmande à cet égard. L'adaptation au domaine étant un élément important du développement des systèmes de TAN et de la recherche sur le sujet, l'apparition d'un axe de recherche à part entière sur le cas littéraire s'avère finalement peu étonnante, surtout si l'on y ajoute les gains de performance observés. Cette proportion grandissante de recherches autour de l'adaptation des outils de TA à la littérature se note d'ailleurs dans l'ouvrage que KENNY et WINTERS (2024)

consacrent à cette thématique, de même que dans l'organisation du premier atelier consacré aux « Qualités de la traduction littéraire automatique » (HADLEY *et al.* 2019), qui marque à coup sûr un tournant de la recherche. Évidemment la TLA reste un sujet de niche, ce qui fait que les tentatives d'adaptation restent limitées malgré l'intérêt croissant et me permet viser l'exhaustivité, bien qu'il ne soit pas exclu que l'un ou l'autre travaux majeurs dans d'autres langues m'aient échappé pour des raisons linguistiques évidentes.

La paire anglais-slovène

L'une de ces publications concerne la paire anglais-slovène, pour laquelle KUZMAN *et al.* (2019) mettent au point plusieurs systèmes de traduction personnalisés dont les performances sont comparées à celles de *Google Traduction*. Ces systèmes (LSTM) sont entraînés au moyen du module *OpenNMT* (KLEIN *et al.* 2017) sur des données issues du projet OPUS²⁷ (TIEDEMANN 2012). Le système générique résultant²⁸ est par la suite adapté au domaine littéraire avec un corpus de textes littéraires variés (9 romans) provenant du corpus SPOOK (ERJAVEC 2013) dans un premier cas, puis évalué sur le roman sentimental *Something about you* de Julie James et sur sa traduction *Nekaj na tebi* par Irena Furlan. Dans le second, c'est un texte de la même autrice et de la même traductrice que le texte à évaluer qui est utilisé pour affiner le système générique : la romance *Practice Makes Perfect*, ainsi que sa traduction *Osem let skomin*.

Le score BLEU obtenu pour l'outil générique est de 17.50, contre 19.01 et 20.75 pour le second. L'équipe de recherche rapporte par ailleurs un score pour d'autres systèmes entraînés uniquement sur le roman (1.78), uniquement sur le corpus littéraire varié (6.61), ainsi que sur un mélange de toutes les données (16.02). Étonnamment, la combinaison de tous ces différents corpus semble faire baisser le score par comparaison aux systèmes affinés séparément sur le roman et sur le corpus SPOOK, contredisant l'idée que l'ajout de données ne peut qu'améliorer les performances et laissant penser que la cohérence des corpus y aurait aussi un rôle à jouer. *Google Traduction* en revanche affiche des résultats supérieurs dans tous les cas (21.97), ce que l'équipe attribue tout de même à la taille fortement limitée de leur corpus d'entraînement, et ce, bien que l'ajout de données adaptées à une autrice et à sa traductrice permette de se rapprocher de ce score, montrant ici aussi qu'un corpus pertinent vaut mieux qu'un jeu de données varié et plus fourni.

27. OPUS (*Open Parallel Corpus*) est un dépôt de corpus parallèles ouvert visant à faciliter l'accessibilité de ces données à des fins de recherche. Ces données sont disponibles pour de nombreuses paires de langues dans de nombreux domaines (presse, articles scientifiques, brevets, actes de parlements gouvernementaux ou européen, sous-titres...) et servent très souvent de point de départ pour la mise au point de systèmes de TA.

28. Système qui n'est pas entraîné sur des données d'un domaine spécifique et qui doit soit pouvoir traduire tout type de texte, soit être « affiné » (*fine-tuned*) sur des données d'un domaine particulier.

La paire anglais-russe

Evgeny MATUSOV (2019) s'est lui aussi intéressé à la TLA, en tentant pour sa part d'évaluer un système pour la traduction de l'anglais vers le russe et un second de l'allemand vers l'anglais, tous deux entraînés grâce au module RETURNN (ZEYER *et al.* 2018). Dans le premier scénario, les corpus utilisés sont un mélange de données hors domaine (25 millions de phrases) et de données littéraires (270 000 phrases provenant du corpus *Books*)²⁹, augmentées avec un corpus synthétique de romans provenant de la bibliothèque numérique *Lib.ru*³⁰ (2,3 millions de phrases), à la manière de SENNRICH *et al.* (2016a)³¹.

Le corpus d'évaluation utilisé pour cette expérience est une compilation de quatre nouvelles : « The Lift » et « A Scandal in Bohemia » d'Arthur Conan Doyle, « The Pit and the Pendulum » d'Edgar Allan Poe, et enfin « The Invisible Man » de Gilbert Keith Chesterton. Pour ce couple de langues, l'auteur donne des points BLEU de 13.9 pour *Google Traduction*, de 14.2 pour le système non adapté et de 15.2 pour le système adapté. Étant donné la faible fiabilité de cette métrique, une évaluation humaine a également été conduite, sous la forme d'une classification d'erreurs personnalisée. Dans l'ensemble, l'évaluation indique que 20 % des phrases ne contiennent pas d'erreur et que le système maison donne lieu à moins d'erreurs graves (de sens et de syntaxe) que *Google Traduction*, bien qu'il produise plus d'erreurs mineures. Les erreurs de sens demeurent toutefois les plus fréquentes, tandis que le nombre d'omissions, les erreurs liées à la cohérence et les problèmes de registre sont jugés trop nombreux pour un extrait de cette taille.

La paire allemand-anglais

Dans le même article, MATUSOV (2019) élabore un second système pour la paire allemand-anglais. Contrairement aux études précédentes, celui-ci est entraîné à partir de l'architecture *Transformer*, que je mentionnais plus haut et qui a plus récemment permis d'atteindre de meilleurs résultats dans de nombreux domaines. Les corpus utilisés pour ce faire proviennent des mêmes sources (47 millions de phrases hors domaine auxquelles s'ajoutent 50 000 phrases du corpus *Books*), à l'exception de celui utilisé pour la compilation du corpus synthétique par rétrotraduction

29. Le corpus *Books* est un corpus parallèle d'œuvres littéraires — le seul du dépôt OPUS — provenant toutes du domaine public.

30. Cf. <http://lib.ru/>.

31. Un corpus synthétique est un corpus monolingue de la langue cible que l'on transforme artificiellement en corpus parallèle en traduisant automatiquement les données monolingues vers la langue source. Ce procédé de rétrotraduction (ou *back-translation* en anglais) permet d'obtenir facilement une grande quantité de documents alignés autrement difficiles à acquérir, et la méthode semble offrir des gains de performance notables en pratique.

(10 millions de phrases issues du projet Gutenberg)³². Mis à l'épreuve sur *La Métamorphose* de Kafka, l'outil montre cette fois des résultats inverses au cas précédent, avec un score BLEU de 20.2 pour *Google Traduction*, de 18.5 pour le système générique et de 16.2 pour le système adapté. Cette baisse pourrait néanmoins être expliquée par la taille réduite des données d'entraînement dans ce second scénario, ou tout simplement par la rigidité de la mesure BLEU.

L'auteur indique d'ailleurs que ces résultats sont à prendre avec un peu de recul, dans la mesure où l'évaluation humaine contredit ceux-ci. En effet, la personne bilingue chargée d'annoter les erreurs des sorties de traduction juge de meilleure qualité les textes produits par le système ayant pourtant obtenu le score le plus faible, estimant en outre que 30 % des phrases sont d'une qualité « acceptable ». Il apparaît d'autre part que le système de TA adapté utilise un vocabulaire beaucoup plus riche, bien que les erreurs de sens soient plus fréquentes par comparaison avec *Google Traduction*. Malgré la difficulté liée à la fiabilité des métriques, l'auteur conclut donc que l'adaptation au domaine permet d'obtenir de meilleurs résultats en littérature, en particulier pour le couple allemand-anglais. Il ajoute que cette pratique pourrait faciliter le processus de post-édition, puisque les erreurs les plus graves et les problèmes de syntaxe laissent place à des erreurs plus simples et rapides à corriger. Enfin, il est noté que des systèmes de traduction automatique qui opéreraient au niveau du texte lors de la traduction, et non plus de la phrase, pourraient grandement bénéficier à la prise en charge des textes littéraires³³.

Plus récemment, le chercheur est venu proposer, lors de la conférence *Literary Machine Translation as a Human-Machine Dialectic* organisée à Liège, un prototype de TA personnalisable mis au point par AppTek qui permettrait aux traducteurs de modifier en temps réel différents paramètres tels que le type de texte, le registre, le genre d'un personnage dans un dialogue, etc. (MATUSOV 2022). Avec cette proposition, Matusov formule dès lors l'une des premières tentatives visant à améliorer l'ergonomie et la personnalisation d'un outil de TA littéraire, mettant l'accent sur l'agentivité des traducteurs et la comparaison de traductions alternatives. Si la question des systèmes adaptés est primordiale pour la TLA, il ne fait nul doute que l'interface de travail l'est aussi, bien qu'elle ait été largement écartée jusqu'à présent (j'y reviendrai dans la [Section 8.2.3](#)).

32. Cf. <https://www.gutenberg.org/>.

33. De fait, si les unités de traductions postulées par VINAY et DARBELNET (1958) se sont progressivement étendues jusqu'à prendre en considération le texte dans sa globalité (BASSNETT 2013, p. 126), les technologies de la traduction fonctionnent par défaut au niveau de la phrase (MERTEN 2018). C'est encore le cas de la traduction automatique également, bien que la recherche s'efforce de dépasser cette barrière, notamment par le biais de la concaténation de phrases voisines (LUPU *et al.* 2022b) ou des approches multi-encodeur permettant d'enregistrer des informations contextuelles (LUPU *et al.* 2022a). Or, qu'il s'agisse des interfaces de TAO servant à la post-édition ou à la TA elle-même, ce mode de fonctionnement a une incidence sur la qualité des sorties et sur le processus de post-édition, dont je reparlerai par la suite (cf. [Chapitre 7](#)).

La paire anglais-catalan

Une autre question primordiale qui apparaît régulièrement en toile de fond des recherches sur la TLA est celle de la réception par le lectorat ainsi que des effets de la TA sur l'expérience de lecture, que l'on retrouve parmi les objectifs d'un projet réunissant déjà bon nombre de publications (TORAL et WAY 2018; TORAL *et al.* 2018b; MOORKENS *et al.* 2018; TORAL *et al.* 2020; GUERBEROF-ARENAS et TORAL 2020; 2022). Celui-ci vise à évaluer les performances de la traduction automatique sur la paire anglais-catalan en confrontant différents paradigmes de TA et différentes modalités de traduction. Il est par ailleurs intéressant de noter que ce projet est le premier à utiliser des données d'entraînement uniquement littéraires, là où les études qui précèdent avaient systématiquement recours à des corpus hors domaine afin d'obtenir des données suffisamment larges pour la mise au point du système de TA, ce qui tient une fois encore à la difficulté de trouver des corpus adaptés.

Selon TORAL et WAY (2018), néanmoins, l'émergence continue des livres électroniques pourrait changer cette réalité, en facilitant la construction de corpus parallèles. Si l'on y ajoute l'arrivée de la TA neuronale, et en particulier sa capacité à produire des traductions moins littérales et lexicalement plus riches que les systèmes statistiques, ces changements pourraient, à en croire les auteurs, changer notre perception vis-à-vis de l'utilisation de cette technologie en littérature. De fait, le corpus d'entraînement utilisé ici a été constitué de manière automatique à partir de 133 romans (1 million de phrases) et de leur traduction. Plus encore, les auteurs ont eux aussi constitué un corpus synthétique, composé d'environ 1 000 romans (5 millions de phrases), pour renforcer la représentation de la langue cible. Le système (LSTM) est entraîné avec le module *Nematus* (SENNRICH *et al.* 2017), et le modèle résultant est un assemblage des quatre meilleures sorties. Le système neuronal est ensuite comparé à un système statistique, entraîné lui aussi sur ces données avec le module *Moses*, pour lequel un corpus parallèle hors domaine de 400 000 phrases (*OpenSubtitles*), ainsi qu'un corpus monolingue hors domaine de 16 millions de phrases en catalan (*caWaC*) et un corpus monolingue littéraire (1 600 romans écrits en anglais), ont également été utilisés. Les deux paradigmes sont alors évalués sur 12 romans différents, parus entre 1920 et aujourd'hui. Les scores BLEU obtenus varient entre 17.94 et 38.92 pour la TAN, avec une moyenne de 32.12, contre 16.11 et 37.35 pour la variante statistique, avec une moyenne de 29.09. La variante neuronale montre donc de meilleurs résultats dans tous les cas, avec une augmentation du score BLEU pouvant varier de 3 à 14 % (11 % en moyenne).

Toujours dans TORAL et WAY (2018), une évaluation humaine conduite sur trois des romans traduits pour lesquelles les métriques sont parmi les plus élevées montre que la TA neuronale est perçue d'une qualité équivalente à la traduction humaine pour 17, 32 et 34 % des sorties, contre 7, 18 et 20 % pour la TA statistique. Les juges n'étant pas des spécialistes, une seconde évaluation est conduite dans TORAL *et al.* (2020) et donne alors 10, 22 et 23 % de phrases produites par la TAN équivalentes à la traduction humaine pour les mêmes romans. Cette amélioration de la TAN par rapport à la TAS se note également lorsque l'on demande à des spécialistes de post-éditer un texte produit par ces deux systèmes. Dans TORAL *et al.* (2018b), il a ainsi été demandé à six juges de comparer la traduction libre d'un roman (*Warbeaker*, de Brandon Sanderson) par rapport à sa post-édition à partir des systèmes statistiques et neuronaux mentionnés plus haut³⁴. Comme dans d'autres études sur la post-édition réalisées dans d'autres domaines, les conclusions indiquent que cette tâche diminue l'effort cognitif fourni durant la traduction (42 %), réduit le nombre de frappes (23 %) et augmente la productivité (36 %), y compris lorsqu'elle mène à des pauses plus longues que la traduction libre (25 %). Toutes ces mesures sont les plus marquées lorsque l'on passe du système statistique au système neuronal. Selon MOORKENS *et al.* (2018), les traducteurs n'ont cependant pas toujours conscience de ces gains, et expriment systématiquement une préférence pour la traduction humaine sans TA, car le sentiment de liberté et de créativité y est plus grand. L'équipe de recherche remarque par ailleurs que la plus jeune génération de traducteurs montre une attitude plus positive envers la TA et accepteraient plus facilement d'y avoir recours que leurs collègues avec plus d'expérience. Enfin, l'unique juge à avoir plus d'un an d'expérience avec la post-édition est également celui qui trouve l'aide offerte par la TA la plus utile.

Dans TORAL *et al.* (2020), à nouveau, les trois auteurs reviennent sur l'étude impliquant 12 romans pour mettre cette fois à l'épreuve l'architecture *Transformer* et proposent des comparaisons avec d'autres systèmes sur ces mêmes données. La première évaluation effectuée dans ce travail montre ainsi un net saut de performance de l'architecture *Transformer* par rapport aux modèles récurrents avec attention, pour laquelle les auteurs obtiennent un score BLEU dépassant les 40, bien qu'ils ne donnent pas les chiffres exacts et se concentrent plutôt sur les comparaisons entre systèmes. De même, ceux-ci reviennent sur le classement des sorties de TA par des spécialistes et précisent que 15, 32 et 30 % de phrases produites par le système *Transformer* sont équivalentes à la traduction humaine pour ces romans, menant à une augmentation appréciable en comparaison avec les

34. Il est important de noter que les personnes n'avaient pas toutes une maîtrise parfaite de la langue source et que deux seulement avaient de l'expérience en post-édition.

10, 22 et 23 % affichés par le paradigme précédent. Une évaluation de l'effort de post-édition indique par ailleurs que l'entraînement des systèmes sur des données littéraires permet de réduire considérablement le nombre d'erreurs. Par comparaison avec un texte issu de *Google Traduction*, celles-ci se limitent en outre à des cas moins sévères, comme le montrait déjà MATUSOV (2019).

Dans une autre étude centrée sur la traduction de la nouvelle « Murder in the Mall » de Sherwin B. Nuland, GUERBEROF-ARENAS et TORAL (2020) rapportent qu'un traducteur et une traductrice littéraires trouvent la TA plutôt utile et qu'elle produit des textes plutôt fidèles que fluides. À choisir entre la post-édition et la traduction libre, les deux ne s'accordent toutefois pas sur la méthode la plus rapide et la moins laborieuse, bien que la deuxième soit toujours préférée. Le retour d'une juge experte pointe d'autre part le fait que le texte littéraire post-édité contiendrait moins d'erreurs, mais aussi moins de tournures créatives. Enfin, une enquête sur la réception de la traduction humaine et post-éditée montre que les deux paradigmes sont évalués presque au même niveau en ce qui concerne leur réception, leur appréciation et l'intérêt suscité, malgré la difficulté à dégager des tendances³⁵. Pour cette dernière raison, l'article se conclut sur la nécessité de vérifier ces observations avec plusieurs paires de langues, avec d'autres ouvrages, et avec différents genres littéraires.

La paire anglais-néerlandais

Une autre branche de ces recherches sur la traduction automatique entraînée uniquement sur de la littérature s'est portée sur la paire anglais-néerlandais (TORAL *et al.* 2024), cherchant par la même occasion à augmenter le volume de données utilisées. L'étude mentionne de fait un passage de 1 million de phrases pour le corpus parallèle et 5 millions de phrases pour le corpus synthétique dans le cas du catalan à 5 millions et 20 millions de phrases respectivement pour le néerlandais, compilées à partir de 5 000 ebooks environ au total, auxquelles sont ajoutées 20 millions de phrases hors domaine issues de *ParaCrawl*³⁶. Le second objectif de l'étude est de vérifier la performance d'une TA effectuée à partir d'une concaténation de quelques phrases (jusqu'à 800 caractères de part et d'autre de la phrase à traduire) au lieu de phrases isolées. Différents systèmes de TA sont dès lors mis au point grâce au module *Marian* (JUNCZYS-DOWMUNT *et al.* 2018), puis mis à l'épreuve sur une collection de 31 romans, représentatif de plusieurs périodes et de plusieurs genres, et ensuite évalués avec BLEU et COMET (REI *et al.* 2020).

35. En revanche, et sans surprise, la TA brute fait montre de très peu de créativité, contient beaucoup plus d'erreurs et conduit à une moins bonne expérience de lecture selon les trois paramètres étudiés.

36. Cf. <https://paracrawl.eu/>.

Une comparaison des scores BLEU montre que l'ajout des données synthétiques est bénéfique (passage de 32.63 à 33.52), tout comme l'utilisation de l'architecture *Transformer-big* (34.41), plus lourde que le *Transformer* de base, ainsi que l'ajout des données hors domaine (35.63). Cependant, la métrique n'indique pas d'amélioration pour la prise en compte du contexte (400 ou 800 caractères à gauche et à droite) ni pour l'assemblage des trois meilleurs modèles, contrairement à COMET, selon laquelle ce dernier système serait également le seul à surpasser *DeepL* (pour 25 des 31 romans choisis)³⁷.

À partir des sorties de traduction générées par ce système, GUERBEROF-ARENAS et TORAL (2022) ont pu renouveler l'expérience de post-édition littéraire similairement menée pour la paire anglais-catalan, à partir de la nouvelle « 2BRo2B » de Kurt Vonnegut et de sa traduction en néerlandais à partir de l'anglais. Celle-ci vient confirmer que la post-édition classique contraint les traducteurs littéraires, puisque les traductions humaines sont considérées moins créatives par plusieurs juges, après quoi vient la PE, puis la TA brute. Cette fois, néanmoins, la PE génère plus d'erreurs que la traduction sans TA, conduisant ces mêmes juges à classer de façon unanime la traduction humaine comme excellente, la PE comme moyenne et la TA comme extrêmement mauvaise³⁸.

La paire anglais-espagnol

Pour terminer, une étape plus récente de cette collaboration centrée sur le catalan envisage quant à elle de se servir de la TA comme aide à la lecture de livres numérisés en langue étrangère. Celle-ci est basée sur le projet *InLéctor*, qui propose une interface bilingue permettant d'alterner entre un texte original et sa traduction, afin de pouvoir vérifier les passages éventuellement problématiques sans devoir y ajouter la charge cognitive de l'aller-retour entre le texte et un dictionnaire (OLIVER 2017). Dans ce scénario, l'application de la TA vise donc principalement le lectorat et a pour objectif d'encourager la lecture en langue étrangère, dont elle pourrait élargir l'offre de la plateforme (uniquement des livres et des traductions du domaine public) en fournissant une traduction automatique vers la langue maternelle de l'utilisateur (OLIVER *et al.* 2022). Encouragé par une enquête positive réalisée pour le catalan auprès du lectorat, OLIVER (2024) envisage dans ce même but une approche analogue à celle que je développerai par la

37. Les auteurs de l'expérience expliquent que le roman pour lequel *DeepL* obtient le score le plus éloigné de leur système date du XIX^e siècle et que la représentativité des données pourrait jouer un rôle dans ces cas, puisque leur corpus ne contient que des œuvres plus récentes. La remarque se termine par un contraste avec un autre système dont l'architecture est plus directement comparable à celle des autres, qui atteste du large gain de performance obtenu grâce à l'ajout de données littéraires.

38. GUERBEROF-ARENAS et TORAL (2022) présentent également une comparaison avec le catalan, qui confirme ces résultats.

suite, à savoir des systèmes adaptés à des autres spécifiques, qu'il met ici à l'essai pour la paire anglais-espagnol.

L'outil de TA proposé pour cette tentative n'est plus celui d'un modèle entraîné uniquement à partir d'un large corpus littéraire, mais un système entraîné sur un large corpus hors domaine, *CCMatrix* (SCHWENK *et al.* 2021), à partir duquel une sélection spécifique issue de ce jeu de données est utilisée pour créer le système adapté, d'une manière qui s'écarte sensiblement du processus d'affinage déjà observé par ailleurs. Dans ce cas, l'auteur utilise 11 recueils des enquêtes de Sherlock Holmes d'Arthur Conan Doyle, en version anglaise, qu'il utilise pour extraire du corpus bilingue hors domaine les 10 millions de phrases sources les plus similaires à ce premier jeu de données, accompagnées de leur traduction. De cette manière, la TA peut être adaptée en utilisant uniquement des données librement accessibles³⁹. Pour ce faire, l'entraînement est effectué avec *Marian*, tandis que l'évaluation est menée sur 1 000 phrases extraites du recueil des *Aventures de Sherlock Holmes*. Plusieurs métriques sont données pour comparer le modèle résultant à *DeepL*, mais fournissent un score tantôt plus élevé pour le premier, tantôt pour le second. Néanmoins, toutes les métriques confirment que le modèle adapté est meilleur qu'un système entraîné sur 10 millions de phrases choisies au hasard dans le même corpus (37 BLEU contre 24) et qu'une sélection des 1 million de phrases les plus proches du domaine suffit à surpasser ce dernier système (30 contre 24).

4.3.3 Évaluation des outils tout public

Bon nombre d'autres études se penchent sur l'application des outils de TA existants aux textes littéraires. S'il ne s'agit donc pas d'outils spécifiquement mis au point pour la littérature — et qu'il faut toujours garder à l'esprit que les textes étudiés se trouvent peut-être dans les données d'entraînement de ces systèmes —, ces évaluations se révèlent tout de même utiles par le fait qu'elles donnent ainsi un point de vue sur les systèmes auxquels la plupart des gens ont librement accès, mais aussi sur d'autres cas d'études (genres, couples de langues, etc.). L'idée de se référer aux textes créatifs comme jalon de la traduction automatique a d'ailleurs tout récemment pris une tournure très concrète, puisque l'édition 2023 de la WMT (congrès annuel sur la traduction automatique) a instauré l'exercice parmi ses campagnes d'évaluation, témoignant au passage de l'implication de l'entreprise Tencent dans la réalisation de cet objectif⁴⁰.

39. Les données littéraires servent ici uniquement à l'extraction depuis le corpus hors domaine et ne sont pas utilisées pour l'entraînement.

40. Cf. <https://www2.statmt.org/wmt23/literary-translation-task.html>.

Bien que je mentionnerai plus en détail ces études et leurs résultats dans le chapitre consacré à l'évaluation de notre système (cf. [Chapitre 6](#)), il est opportun de mentionner ici le travail mené au cours d'un projet commun sur la paire anglais-néerlandais ([TEZCAN et al. 2019](#); [FONTEYNE et al. 2020](#); [WEBSTER et al. 2020](#); [MACKEN et al. 2022](#)), ou encore celui de Shuyin ZHANG (2024) sur la paire chinois-anglais, qui, dans la lignée de Chung-ling SHIH (2016), envisage à nouveau la TA comme une aide à la lecture de littérature numérique. MOHAR et al. (2020) réactualisent pour leur part un essai de TA en slovène au départ de la micronouvelle « Life Stories », de Margaret Atwood, tandis que Peter CONSTANTINE (2019) livre une interprétation plus optimiste et exotisante des outils en ligne à partir des traductions anglaises du *Candide* de voltaire. On peut aussi noter l'expérience récente de PETRAK et al. (2022) sur la traduction en croate du roman *La traduction est une histoire d'amour* de Jacques Poulin, ainsi que celle de Hoyt LONG (2022) sur une sélection d'œuvres de fiction japonaises traduites en anglais, qui offrent toutes deux des précisions sur les obstacles propres à ces combinaisons linguistiques, ou celle de THAI et al. (2022), qui propose une évaluation de l'outil *Google Traduction* pour plusieurs langues, en collaboration directe avec l'entreprise.

D'autres cas d'études considèrent quant à eux des phénomènes particuliers typiques de la littérature tels que la métaphore ([ZAJDEL 2022](#); [DORST 2024](#)), la structure thème-rhème ([JIANG et NIU 2022](#)), la traduction des occasionalismes ([KOLB et al. 2023](#)) ou la comparaison ([UTUSHKINA 2023](#)). Jean-Yves BASSOLE (2022) souligne pareillement les difficultés posée par la traduction littéraire à partir de l'œuvre de Louis-Ferdinand Céline, caractérisée par son usage du langage parlé et argotique, de même que la difficulté d'évaluer ces systèmes dont les performances peuvent grandement varier — en bien ou en mal — même d'un jour à l'autre. D'autres contributions, enfin, envisagent de faire ressortir les caractéristiques de textes littéraires post-édités par rapport aux sorties de TA et aux traductions humaines, comme le proposent Sheila CASTILHO et Natália RESENDE (2022) en comparant différentes modalités de traduction en portugais brésilien des romans *Alice's Adventures in Wonderland*, de Lewis Carroll, et *The Girl on the Train*, de Paula Hawkins. Notons qu'au moment d'écrire ces lignes, des travaux commencent à paraître concernant les grands modèles de langue ([KARPINSKA et IYYER 2023](#)), bien que je reviendrai plus spécifiquement sur le sujet dans les conclusions de l'ouvrage (cf. p. 236).

Enfin, il me faut évoquer un pan de la recherche encore bien trop peu pris en compte, à savoir celui des expériences personnelles et comptes rendus des traducteurs littéraires en personne. Je pense à des témoignages et retours d'expérience tels que ceux de Dominique DEFERT (2022) qui,

face aux accents sur les progrès du TAL, pointent les limites évidentes des outils de TA actuels, les spécificités de leur métier, mais aussi et surtout les usages — inavoués pour certaines personnes, marginaux si l'on se place du point de vue de l'ingénierie informatique — de la traduction automatique⁴¹. Des exposés évidemment indispensables sur les contraintes ergonomiques, situationnelles et sociétales de l'activité traduisante dans ce domaine qu'est la littérature. Une autre de ces contributions, livrée par Andrew ROTHWELL (2024) sur un exercice de retraduction de Marcel Proust en anglais, met ainsi en lumière l'intégration de la TA au sein du large éventail des technologies de la traduction, dépassant donc la seule question des performances de la machine pour soulever des problématiques plus larges liées à son utilisation.

4.3.4 *Aspects éthiques et théoriques*

Un autre pan de la recherche sur la TLA vise précisément à mieux rendre compte des usages et attitudes envers la technologie en littérature (SLESSOR 2020 ; RUFFO 2018 ; 2021 ; 2022 ; DAEMS 2022 ; HADLEY 2024). De façon plus générale, ces publications entendent cerner et promouvoir les enjeux sociétaux qui entourent l'utilisation des nouvelles technologies, tout en contribuant aux discussions en cours. Une fois encore, ces travaux seront abordés plus en détail dans les chapitres à venir, notamment lors de la discussion qui viendra clôturer cet ouvrage (cf. Chapitre 8).

Il est par exemple possible de trouver des réflexions pionnières en la matière chez Kristiina TAIVALKOSKI-SHILOV (2019a), qui livre de multiples mises en garde contre une utilisation abusive des technologies de la traduction dans un contexte économique d'ores et déjà peu favorable aux traductaires. De même, KENNY et WINTERS (2020) évoquent les aspects humains de la TLA, se concentrant en particulier sur les enjeux éthiques liés à la notion de « voix » à l'occasion d'une expérience de post-édition littéraire⁴². Dans cette démarche empirique, les chercheuses entendent faire intervenir directement les traductaires, favorisant leur prise de parole dans le contexte scientifique, et montrent en dernière analyse que si la PE est un moyen de renforcer leur style et de s'appropriier le texte, une comparaison qualitative et quantitative des différentes modalités révèle qu'elle affaiblit aussi leur voix textuelle par rapport à une traduction libre (WINTERS et KENNY 2024). Partant d'une expérience de traduction anglais-italien semblable à celles de la section précédente, Paola BRUSASCO (2022) se penche similairement sur la compatibilité de la

41. Paola RUFFO (2021, p. 213) indique à cet égard que les traductaires littéraires détournent fréquemment les outils de leur utilisation première, à la faveur de scénarios servant plus directement leur activité.

42. Partant de cette expérience, Hans-Christian Oeser, le traducteur ayant pris part à l'étude, livre lui aussi son retour d'expérience dans une publication complémentaire (OESER 2020).

post-édition avec la traduction littéraire, tout comme le proposent également Waltraud KOLB (2024) sur une expérience anglais-allemand ainsi que VIEIRA *et al.* (2023) sur une traduction en chinois de quatre nouvelles de science-fiction écrites par l'auteur canadien Peter Watts.

L'une de ces retombées, qui demeure encore une énigme à l'heure actuelle, concerne spécifiquement les droits des traductaires et la propriété des traductions machine, à propos desquels KOPONEN *et al.* (2022) livrent une contribution cruciale. L'un des points évoqués tout au long de ce chapitre concerne effectivement la réutilisation de traductions existantes pour entraîner les systèmes de TA, dont le principe soulève des questions juridiques depuis l'avènement des systèmes de TAO. Dans le même ordre d'idées, Claire LARSONNEUR (2023) propose une réflexion axée sur les répercussions socio-économiques de la TA en littérature, tandis que Mehmet ŞAHİN et Sabri GÜRSES (2019) mettent en garde contre les risques de plagiat qu'implique une technologie qui recycle des fragments de traductions passées⁴³.

Enfin, une dernière frange des recherches offre une réflexion plutôt théorique sur la pertinence des outils, sur les difficultés posées par les textes littéraires, ainsi que sur la qualité des sorties de traduction, que l'on retrouve parfois en partie dans les travaux précédents. Ces préoccupations peuvent inclure par exemple la traduction des temps du passé, la cohérence discursive, les différents registres de langue et types de discours, ou encore les phénomènes de focalisation et d'intertextualité, etc. (TAIVALKOSKI-SHILOV 2019b; PONCHARAL 2021). Sans surprise, ces singularités et irrégularités du langage littéraire ont toutes une incidence directe sur la qualité de la traduction et l'expérience de lecture, mais sont difficilement reflétées dans des évaluations purement automatiques des sorties de TA.

4.4 Aller plus loin : des systèmes adaptés au niveau de l'humain

Dans l'ensemble, donc, nous voyons que les systèmes neuronaux produisent de meilleurs résultats que les systèmes statistiques, en particulier ceux basés sur l'architecture *Transformer*, de même que ceux entraînés sur de la littérature par comparaison aux généralistes, pour autant qu'ils soient suffisamment fournis en données. Chacun de ces changements offre ainsi des gains de performance pour le domaine littéraire, de même qu'une meilleure appréciation des résultats que l'on pourrait attendre dans ce secteur. Dans ce contexte d'évolution croissante de la TA,

43. De la même manière, j'ai pu proposer ailleurs un rapprochement entre les différentes technologies de la traduction, qui font usage des traductions produites par les traductaires sous une forme ou sous une autre (corpus, mémoires de traductions, bitextes...) et soulèvent de ce fait de nombreux enjeux éthiques, mais aussi, a contrario, une autre manière d'envisager ces outils de traduction (HANSEN 2024).

l'étude exposée dans cet ouvrage vise tout d'abord à réévaluer la paire anglais-français à travers le prisme de l'approche neuronale et, ce faisant, mettre en avant un aperçu des performances de la TA appliquée à des textes littéraires, ce qui n'avait plus été observé pour le français depuis BESACIER (2014).

Plus encore, il serait également utile, selon les observations de KUZMAN *et al.* (2019) et TORAL *et al.* (2020), de voir quelles seraient les performances d'un système de TA entraîné spécifiquement sur la production des mêmes autaires et traductaires. Or, c'est précisément ce que j'ambitionne de montrer par le biais de cette recherche, en proposant une nouvelle approche de la TA adaptée non pas uniquement au domaine, mais en tout premier lieu à une traductrice et à son style propre. En ce sens, la démarche reflète et actualise d'autres travaux pionniers d'adaptation tels que chez MIRKIN et MEUNIER (2015) et MIRKIN *et al.* (2015), qui, bien qu'ils se penchent en premier lieu sur les utilisataires finaux et les autaires sans accorder beaucoup d'attention aux traductaires, soulignent tout de même l'importance du style pour la réception des traductions et la qualité qui leur est attachée. Concernant ce besoin de systèmes personnalisés adaptés à un style particulier, on peut notamment y lire :

Le terme « domaine » revêt une acception très large dans le champ de la TA et peut se référer à la fois au thème, au dialecte, au genre ou au style (CHEN *et al.* 2013). Néanmoins, aucune procédure d'adaptation au domaine n'a été étendue, à notre connaissance, à des attributs tels que les traits démographiques ou les traits de personnalités [18]. (MIRKIN *et al.* 2015, p. 1103.)

De ce point de vue, on pourrait considérer le style individuel d'une personne comme un « domaine » auquel il serait possible d'adapter la machine, au même titre qu'à une tâche ou à une thématique donnée. Et quoique les méthodes proposées dans ces travaux se situent à l'échelle plus générique du genre des autaires ou des préférences de catégories d'utilisataires, l'idée n'est pas incompatible avec celle d'une personnalisation plus spécifique, pour laquelle le style concernera avant tout les choix et préférences des traductaires littéraires (TABER 1972, p. 61 ; BAKER 2000, p. 260–261 ; BOASE-BEIER 2014, p. 393–394 ; YOUNDALE 2020, p. 2–5). À ce propos, WANG *et al.* (2021) ont plus récemment cherché à illustrer l'importance du style des traductaires qui a si longtemps été négligé dans les recherches sur la TA et ont montré qu'un même système auquel on indiquait la source de différentes traductions sous la forme d'une simple étiquette pouvait générer des sorties qui, bien que jugées d'une qualité équivalente au cours d'une évaluation humaine, pouvaient varier sensiblement au niveau de la forme, sous-entendant que les systèmes peuvent modéliser le style et offrant un signe encourageant pour ce travail. On peut également noter à cet égard que le changement de paradigme avancé depuis le début des

travaux de recherche semble avoir été entendu, puisque des collègues que j'ai pu croiser lors de plusieurs présentations des résultats ont par la suite adopté la même démarche (YIRMIBEŞOĞLU *et al.* 2023).

Le [Chapitre 5](#) consacré à la méthodologie détaille ce travail technique de mise au point et d'adaptation, après quoi je consacrerai la discussion aux enjeux de l'évaluation d'un tel système (cf. [Chapitre 6](#)). De fait, une des conclusions que pose cet état de l'art est la dépendance toujours actuelle aux métriques d'évaluation automatiques qui peuvent mener à des résultats contradictoires (KUZMAN *et al.* 2019 ; TORAL *et al.* 2024) ou indiquer des résultats opposés aux évaluations humaines (MATUSOV 2019). Or, on peut difficilement évaluer la possibilité même de la traduction littéraire automatique par la biais de métriques non fiables, c'est pourquoi ce chapitre d'évaluation poussera plus avant cette réflexion et donnera un aperçu de divers types d'évaluation. Pour la même raison, le [Chapitre 7](#) propose une évaluation experte de l'outil qui me permettra dans un second temps de cerner les effets de la technologie sur l'opération traduisante, sur le processus même de traduction et non uniquement sur le résultat produit par la TA, avant de clôturer le travail sur les questions éthiques également abordées au cours de ces dernières pages (cf. [Chapitre 8](#)).

Méthodologie

5.1 Cadre épistémologique

La méthodologie déployée tout au long de ce travail reflète ma volonté de promouvoir une approche raisonnée de la technologie, ainsi que je l'ai détaillé au troisième chapitre (cf. [Section 3.3](#)). Pour mettre à l'épreuve la traduction automatique (TA) individualisée et sa pertinence pour le cas littéraire, je me suis dès lors appuyé sur diverses approches — déjà préfacées en partie par l'organisation de l'état de l'art que je viens de présenter —, qui tiennent avant tout au caractère original de l'étude, en particulier au moment de sa conception en 2017, mais aussi à la véritable explosion de l'attention portée à ce sujet depuis, de même qu'à l'ancrage résolument transdisciplinaire de la thèse.

De fait, le choix d'une étude de cas tient avant tout à l'aspect novateur de la recherche à ses débuts, qui incitait à voir quelles seraient les performances concrètes de l'outil sur un texte donné. Et si nous venons d'observer quelques tentatives d'adaptation, le besoin d'examiner la production d'un système spécifiquement mis au point pour le domaine littéraire reste entièrement valable, d'autant que cette étude s'ajoute à une liste somme toute fort restreinte. Ce choix coïncide par ailleurs avec l'arrivée de nouvelles architectures de TA qui ont favorisé la mise à l'épreuve empirique d'un outil spécifiquement conçu dans ce but et la comparaison de différents systèmes.

Le reste des chapitres s'est alors imposé dans une sorte d'enchaînement émergent, dans la mesure où ceux-ci trouvent leur origine dans les réflexions et questionnements suscités par cette étape de développement, tout autant que dans les publications qui se faisaient grandissantes sur le sujet. L'idée même d'outils adaptés au style personnel des traductaires, bien que présente au départ, s'est d'ailleurs affinée jusqu'à occuper l'avant-plan de la thèse, à la fois durant la phase d'évaluation et au cours des discussions entre collègues une fois venu le moment de partager les résultats du travail dans un nombre lui aussi grandissant de rencontres autour de la traduction littéraire automatique. Cette place centrale accordée à la personnalisation a finalement rendu indispensable la rencontre avec la traductrice du roman choisi pour cette expérience, qui a elle-même encouragé la discussion de fin d'ouvrage.

Tout ce cheminement traduit au passage le souci de combiner les approches, comme je le disais, passant graduellement du quantitatif au qualitatif. Non pas parce qu'un dispositif serait meilleur que l'autre ou que tel protocole aurait davantage sa place dans tel ou tel champ disciplinaire — car la question est réductrice et n'a pas lieu d'être (GOERTZ et MAHONEY 2012, p. 2) —, mais parce que je suis intimement persuadé que les deux se complètent parfaitement. La combinaison de méthodes mixtes, en illustrant les multiples aspects de cette thématique, apportera en effet un éclairage plus complet que l'une ou l'autre prise séparément (CRESWELL et PLANO CLARK 2007, p. 5). En outre, cette dialectique méthodologique me permettra de considérer tant le produit que le processus de traduction : la possibilité même de la traduction littéraire automatique, adaptée à une traductrice, et les effets de cette technologie à la fois sur l'opération traduisante ainsi que sur cette pratique envisagée comme activité humaine socialement située. Le résultat, je l'espère, permettra de mieux comprendre les analyses quantitatives, qui favoriseront en retour la généralisation de la démarche qualitative.

Dans les sections qui suivent, je détaillerai donc la création de l'outil, adoptant dans un premier temps une démarche purement hypothético-déductive. Cette mise au point part du constat posé dans le chapitre précédent, selon lequel il serait effectivement possible d'adapter un outil de traduction automatique au domaine littéraire, ainsi que je le pensais en entamant la recherche, mais que je voudrais affiner en formulant une première question de recherche :

Est-il possible d'adapter un système de TA au style individuel des traductaires ? (QR1)

L'hypothèse dérivée de la discussion à ce stade est que *l'utilisation de données propres à un traducteur ou à une traductrice doit permettre de concevoir des outils plus personnalisés et donc plus performants que les systèmes entraînés sur des données génériques ou même de la littérature variée*. Étant donné l'objectif exposé visant à promouvoir un discours raisonné sur la technologie, ce point central m'invite cependant à considérer une sous-question de recherche :

Quelles sont les performances d'un outil adapté à la fois à la littérature et au style des traductaires ? (SQR1)

Bien que la faisabilité d'une adaptation au niveau de l'humain soit indéniablement une préoccupation centrale et essentielle de ce travail, il me paraît en effet indispensable d'y ajouter des analyses complémentaires, qui offriront pour leur part un meilleur aperçu des résultats obtenus et des limites auxquelles se heurtent encore les systèmes de traduction automatique actuels.

Le chapitre consacré aux évaluations, notamment, fournira des éléments de réponse qui illustrent de manière singulière la mise en place de méthodes mixtes puisqu'ils résultent d'une combinaison d'évaluations automatiques et humaines (cf. [Chapitre 6](#)). L'avantage en est que les évaluations humaines peuvent apporter un éclairage plus précis et détaillé sur les métriques, tandis que ces mesures offrent l'avantage de montrer très simplement des résultats qui peuvent autrement paraître ambigus ou subjectifs. Elles peuvent fournir en outre un référentiel de comparaison utile à la fois pour les personnes familiarisées avec la conception de ces outils, mais aussi pour celles qui sont peu habituées aux recherches en traitement automatique des langues (TAL). Notamment lorsque, comme nous le verrons, ces diverses méthodes se rejoignent dans leurs conclusions.

L'efficacité de l'outil n'est toutefois pas l'unique préoccupation de cette étude, car au-delà de ses performances, il me semble nécessaire d'évaluer à la fois son utilité pour celles et ceux qui devront finalement interagir avec la TA, et ses conséquences pour la traduction au sens large. Dans un second temps, le système développé me servira donc de levier de compréhension de ces enjeux et me conduira à m'intéresser aux usages, dont l'état de l'art montre à nouveau le peu d'intérêt encore porté au sujet. Partant d'une démarche de conception centrée sur les utilisateurs (UCD, de l'anglais *user-centred design*), l'outil sera ainsi mis à l'épreuve dans une tâche concrète et une situation réelle, en accord avec l'idée que le technologie doit être adaptée à l'humain et non l'inverse ([LAVAUT-OLLÉON 2011](#), p. 6), pour répondre à une seconde question de recherche :

La traduction automatique est-elle utile pour les traducteurs littéraires ? (QR2)

Ce second volet de recherche se fera cette fois au départ d'une approche inductive impliquant d'autres méthodes telles que le raisonnement à voix haute (TAP, de l'anglais *think-aloud protocol*) et empruntant aux études ethnographiques. Ce revirement méthodologique était effectivement nécessaire pour écarter toute idée préconçue sur la question et poser un regard sans a priori sur la *praxis*, mais surtout en raison de la nouveauté de cette thématique et du manque de recherches écologiquement situées. Cette démarche inductive prendra dès lors une importance particulière lors de la collaboration avec la traductrice de notre cas d'étude (cf. [Chapitre 7](#)), qui examinera la traduction envisagée alors comme processus, sans qu'il soit toutefois question de s'y limiter :

Il nous faut découvrir ce qu'il se passe dans l'esprit des personnes qui traduisent, mais aussi au niveau de leurs mains, de leur ordinateur, de leur bureau, de leur langage... La traduction ne se passe pas uniquement dans l'esprit : elle intervient dans des systèmes complexes. Ces systèmes incluent à la fois les personnes, leur environnement physique et social, ainsi que tous leurs artefacts culturels [19]. ([RISKU 2010](#), p. 103.)

Quelles que soient ses performances, la traduction automatique s'insère en effet dans ce système complexe et se retrouve inextricablement liée à la situation personnelle des traducteurs, à leur corps, à leur rôle dans la société et à leur coopération avec l'ensemble de la chaîne de l'édition.

Il serait par conséquent peu opportun de se pencher sur de nouveaux outils technologiques sans prendre en compte les dimensions sociétales et situationnelles qui entourent leur déploiement et qui doivent s'accompagner, me semble-t-il, de réflexions éthiques ou ergonomiques, par exemple. Ceci me mène, une fois encore, à poser une nouvelle sous-question de recherche :

Quelles sont les conséquences de l'arrivée de la TA dans le domaine littéraire ? (SQR2)

L'entretien mené offrira une première vue sur cette dynamique d'influence mutuelle entre la pratique et ce système complexe dans lequel elle intervient, éclairant la perspective de la personne observée et évitant de « perdre de vue le monde réel [20] » (RISKU 2010, p. 107). Pour cette même raison, l'ouvrage s'achèvera sur une discussion qui permettra de recontextualiser les résultats dans une perspective cognitive et sociale, et de réévaluer ceux-ci au prisme des enjeux éthiques posés par les nouvelles technologies dans un contexte d'automatisation grandissante (cf. Chapitre 8). Cette dernière contribution réflexive se veut opposée au déterminisme technologique, à savoir l'idée que la technologie évoluerait de façon autonome vis-à-vis de la société, sans qu'il soit possible d'y changer quelque chose. Si une telle conception n'a rien d'étonnant du point de vue des sciences humaines sociales, il en va de même pour une approche technique qui adopterait une attitude critique vis-à-vis de la technologie (AGRE 1997). Mêlant l'artisanat de la conception et la réflexivité de la critique, cette « *critical technical practice* » reconnaît le caractère indissociable de l'utilité fondamentale des systèmes techniques et de leur utilité pratique dans le monde réel (*Ibid.*, p. 154-155), réactualisant au passage les débats récurrents sur l'injonction à l'innovation. À ce titre, le dernier chapitre proposera un survol de mes lectures, de mes apprentissages et des changements observés durant ces années de recherche, ainsi que quelques craintes elles-mêmes nourries par la littérature et par mes observations sur le sujet, pour en arriver finalement à formuler des recommandations en accord avec la visée de la thèse développée en première partie : la réappropriation des technologies de la traduction.

Ces méthodes de recherche mixtes reflètent ainsi l'idée d'un dialogue entre les diverses disciplines dont ce travail s'inspire, les chapitres qui s'y succèdent et les approches qui y sont conciliées. Au fil des chapitres, j'aurai l'occasion de me pencher à la fois sur le produit (cf. Chapitre 6), sur le processus (cf. Chapitre 7) et sur le contexte (cf. Chapitre 8) de la traduction littéraire (SALDANHA et O'BRIEN 2013), considérant tour à tour le textuel, le cognitif et le culturel (MARCO 2009), mais conservant toujours en toile de fond la nécessité d'une discussion critique et raisonnée sur la traduction automatique. Chacune de ces approches trouve cependant son origine dans un cas d'étude commun : la saga *Septimus Heap* (SAGE 2005–2013) et sa traduction française, la série *Magyk* (SERVAL 2005–2012).

5.2 Cas d'étude : la saga *Septimus Heap* / *Magyk*

5.2.1 Choix de l'ouvrage

La première question posée à laquelle je vais à présent m'attaquer sous-tend la mise au point et l'affinage du système de traduction automatique qui nous accompagnera pour la suite du travail. Cette question de recherche (QR1) vise à évaluer la possibilité même d'adapter un système de traduction automatique au domaine littéraire et, plus encore, au style individuel d'un traducteur ou d'une traductrice. Pour y répondre, et juger par là même des performances d'un tel système, j'ai donc proposé de me centrer sur un cas d'étude concret qui avait déjà fait l'objet d'un travail sur la traduction littéraire assistée par ordinateur et dont la reprise répond à mon ambition de proposer une vue d'ensemble sur les technologies de la traduction en littérature (HANSEN 2024).

Le cas en question est celui de la série *Septimus Heap*, écrite par Angie Sage (HarperCollins, 2005–2013), et de sa traduction française, *Magyk*, par Nathalie Serval (Albin Michel, 2005–2012). Cette saga littéraire à succès tout ce qu'il y a de plus typique de la *fantasy* rapporte le récit initiatique d'un jeune magicien éponyme, septième fils d'un septième fils, au fil de ses pérégrinations et de ses rencontres, sur fond de duel continu entre les forces de la *Magyk* et de la Ténèbre. Ces aventures se déclinent en sept romans prenant place dans un monde assimilable au nôtre, mais situé dans un futur si lointain que la technologie y est devenue indissociable de la magie. L'originalité des *Septimus Heap*, construite sur des modèles éprouvés du genre, a converti chacun de ces volumes en un *best-seller*, ce qui a valu à la série d'être primée, traduite dans 33 langues et prolongée par des romans dérivés⁴⁴, de même qu'une trilogie faisant suite à la série originale⁴⁵.

Malgré ce succès, *Fyre*, le dernier tome du cycle, n'est jamais paru en français, constituant à ce titre un cas intéressant pour la recherche. Outre le rapprochement symbolique entre cet univers fictif, mêlant magie et technologie, et la thématique de cette étude, dont les arcanes sont souvent présentées comme impénétrables, la forte cohérence d'une saga unique composée de plusieurs tomes constitue un cas propice pour équilibrer la taille limitée d'un tel corpus, sachant que la pertinence des données est un enjeu fondamental de cette démarche et que peu de corpus parallèles de littérature sont disponibles en ligne. Dans l'ensemble, la série *Magyk* offre dès lors un point de départ adéquat pour vérifier s'il est possible, *in fine*, d'adapter un système aux œuvres spécifiques d'une saga, d'une autrice et d'une traductrice.

44. Cf. <https://www.angiesage.com/septimus-heap>.

45. *TodHunter Moon* (HarperCollins, 2014–2016).

5.2.2 *La triple condamnation d'une traduction littéraire jeunesse de fiction*

Le choix d'une œuvre de jeunesse pourrait néanmoins être critiqué si l'on s'en tient à une vision hiérarchisée et réductrice de la littérature, qui ferait volontairement abstraction des nombreux écueils que posent ces ouvrages (LATHEY 2016) et du rôle de médiation particulièrement prononcé qui s'y impose (VAN COILLIE et VERSCHUEREN 2006 ; VAN COILLIE et McMARTIN 2020). Outre son statut de traduction, la série s'inscrit effectivement dans deux segments (de jeunesse et de fiction) qui ont été historiquement considérés comme étant « moins littéraires » en regard du canon, qui constituerait quant à lui la « vraie » littérature. Ceci tient notamment à l'argument fallacieux, mais persistant, selon lequel, « par comparaison avec la littérature destinée aux adultes, les livres pour enfants seraient caractérisés par une structure narrative plus simple, un langage moins riche et des qualités artistiques inférieures [21] » (NIKOLAJEVA 2016, p. 6).

En effet, malgré l'engouement qu'elles suscitent et la reconnaissance publique de ces œuvres,

la littérature de jeunesse reste méprisée et boudée par les élites culturelles et universitaires. Les facultés osent petitement lui faire une vraie place, décriant du même coup tout ce qui relève des paralittératures – pour ne pas dire sous-littératures –, et repoussant la littérature de jeunesse dans les douves de la culture littéraire [...] Si, en effet, on rappellera avec Roland Barthes qu'« il n'y a pas de parthénogénèse littéraire » et qu'un enfant ne peut ni écrire ni opérer sur lui sa propre critique, force est également de reconnaître que ce sont précisément ceux à qui elle n'est pas destinée (les adultes) et qui, souvent, la montrent du doigt (les universitaires) qui vont établir la critique de cette *littérature de jeunesse*. Cette dernière est d'ailleurs la seule littérature qui échappe *de facto* à la critique des lecteurs à qui elle s'adresse [...] (PRINCE 2021, p. 39–40.)

Pourtant, l'édition de jeunesse, tout comme le segment de la fiction, représente une part considérable dans le marché de la traduction littéraire. Les ouvrages de jeunesse et de fiction représentent collectivement 43,3 % des cessions de droits sur la production éditoriale française à l'international pour l'année 2022⁴⁶ (SNE 2023, p. 18). Et bien qu'il s'agisse ici du cas particulier de la France, ces chiffres témoignent tout de même de l'importance de ces segments éditoriaux. À l'inverse, pour ce qui est des importations, « on constate que le genre qui contribue le plus à la hausse des traductions littéraires en français est la littérature pour la jeunesse », dont SAPIRO et BOKOBZA (2008, par. 16 et 17) reconnaissent qu'il ne s'agit pas uniquement d'un phénomène commercial, mais aussi d'un type de littérature avec une tradition culturelle qui lui est propre.

46. Ce chiffre passe à 73,6 % si l'on y ajoute la bande dessinée.

Autre preuve de ce succès, sept des treize romans traduits dans le plus grand nombre de langues sont des livres présentés le plus souvent comme des ouvrages de jeunesse⁴⁷ (STATISTA 2022). Si l'on se penche sur la bibliographie complète des autaires, on remarque ainsi que les figures de proue du canon telles que Shakespeare ont été détrônées par les *Oui-Oui* d'Enid Blyton et par d'autres noms que l'on associe typiquement à la jeunesse (PRINCE 2021, p. 39). En ce sens, la série *Magyk* est donc on ne peut plus représentative de la demande en traduction littéraire.

Au-delà de ces aspects économiques, la série est tout autant emblématique des difficultés que peuvent rencontrer les traducteurs face à ce genre littéraire. Si celles-ci ne sont pas exclusives à la littérature, ou à un genre spécifique, et que les distinctions entre toutes ces traductions ne sont pas aussi nettes que le laissent parfois entendre certains discours (cf. Section 2.2), force est de constater qu'elles sont plus fréquentes dans ce type de textes et, par conséquent, souvent mises en lumière par la recherche. On y trouve évidemment l'attention accordée à la créativité, à la forme et aux stratégies nécessaires pour restituer les procédés stylistiques complexes et parfois spécifique aux autaires, mais aussi le subtil dosage entre le respect des normes d'un genre donné et l'originalité requise pour s'en affranchir (VAN COILLIE et VERSCHUEREN 2006 ; LATHEY 2016). On pourra également mentionner les variations de registre, en particulier dans les nombreux dialogues où des tournures familières et du langage cru peuvent côtoyer de près des passages soutenus, ainsi que la présence de dialectes et d'idiolectes (PASCUA-FEBLES 2006, p. 113–115 ; LATHEY 2016, p. 71–86). Ces deux derniers éléments exemplifient par ailleurs l'omniprésence des références culturelles dans les textes littéraires, face auxquelles les traducteurs doivent assurer la médiation, puisqu'à l'écart culturel s'ajoute aussi la différence d'âge et d'expérience entre « destinataire » et « destinataire » (LATHEY 2016, p. 37–53 ; OITTINEN 2006).

L'appellation « jeunesse » est elle-même problématique, puisque le public auquel s'adressent ces œuvres est ambigu. Le terme, pour commencer, recouvre autant la petite enfance, que l'enfance ou l'adolescence, une situation que ne rendent pas moins complexe les divergences dans les catégories d'âge annoncées d'un pays à l'autre pour un même ouvrage. Cette labélisation, ensuite, montre bien que ces livres suscitent aussi l'intérêt des adultes, au point d'être fréquemment réédités avec une couverture ciblant spécifiquement un lectorat plus âgé — ou, à l'inverse, un public plus jeune. Nombre d'entre eux s'adressent d'ailleurs directement à un double destinataire (O'SULLIVAN 1993) et invitent dès lors bien souvent à une double lecture (LATHEY 2006, p. 16),

47. *Le Petit Prince* d'Antoine de Saint-Exupéry, *Pinocchio* de Carlo Collodi, *Alice au pays des merveilles* de Lewis Carroll, les contes de Hans Christian Andersen, *Fifi Brindacier* d'Astrid Lindgren, *L'Alchimiste* de Paulo Coelho et enfin le premier tome de la série des *Harry Potter*.

comme l'illustre le cas particulièrement éloquent de *Haroun et la Mer des histoires*, écrit par Salman Rushdie (RUDVIN et ORLATI 2006)⁴⁸. Cette duplicité des lectures se note, au reste, dans l'émergence de romans *cross-over* (GALEF 1995) et dans le *cross-age* (PRINCE 2021, p. 29), qui visent dès le départ à traverser les générations et invitent à repenser ces catégories (FALCONER 2009), ou, tout du moins, à les considérer avec un peu de recul. Pour des raisons pratiques, et puisque les œuvres dont il est question sont commercialisées de la sorte, je conserverai tout de même le terme.

Quitte à aborder les critiques, il convient pareillement de remettre en question l'idée que ces textes seraient caractérisés par un langage uniforme, neutre et obligatoirement respectable (FERRIER 2009, p. 256–257), lorsque, de fait, ceux-ci peuvent être tout autant subversifs dans le choix des thèmes (JOOSEN 2006) et audacieux dans la manière de les aborder (PRINCE 2021, p. 84–91), mais aussi expérimentaux dans l'exploration de nouveaux espaces créatifs et dans leur interaction avec le lectorat ou avec d'autres médias (REYNOLDS 2007). Dans la lignée des obstacles précités, la traduction de jeunesse soulève donc d'autres écueils qui, à nouveau, ne lui sont pas exclusifs, mais suffisamment récurrents pour être régulièrement soulevé dans les travaux de recherche. On y retrouve ainsi la pluralité des discours, qui s'y impose de façon singulière (JOOSEN 2006 ; GONZÁLEZ CASCALLANA 2006). Autre exemple : celui de l'intertextualité. Bien qu'il s'agisse d'un phénomène fondamental de la littérature en général, l'aspect est prévalent dans les textes de jeunesse également (GONZÁLEZ CASCALLANA 2006), au même titre que l'intericonicité, comme le montre l'extension du segment jeunesse dans des genres et des formats toujours plus variés, et que l'intermédialité, comme en attestent les adaptations respectives entre littérature, cinéma, jeu vidéo... (PRINCE 2021, p. 91–93).

Outre la difficulté liée à l'« ambivalence » des textes et de leur audience (SHAVIT 1986, p. 63–71), ainsi qu'à la capacité d'empathie indispensable pour traduire à destination d'un public plus jeune (VAN COILLIE et VERSCHUEREN 2006, p. VI), on notera enfin l'utilisation en traduction jeunesse d'un langage typiquement créatif et humoristique. La forte de présence de jeux de mots en tous genres (JOOSEN 2006, p. 61) et d'autres jeux sur la forme (GONZÁLEZ CASCALLANA 2006, p. 106–108) en est une preuve notable, ainsi que le montre le cas particulier des noms propres, caractéristiques des multiples fonctions que peuvent remplir non pas uniquement le texte, mais nombre de segments ou de mots spécifiques tout au long de ces œuvres (VAN COILLIE 2006).

48. On peut établir un parallèle ici avec l'audiovisuel, où les films d'animation multiplient les sous-entendus et visent ouvertement toutes les générations.

En dépit de ces nombreuses contraintes, la série choisie est toutefois doublement condamnée, puisque, outre son label « jeunesse », la fiction est elle aussi régulièrement reléguée au rang de « paralittérature », marquant plus profondément son opposition par rapport à ce que serait la « bonne littérature » (BARNABÉ 2012, par. 49), ou encore — en termes bourdieusiens — la littérature bourgeoise, donc légitime, par opposition à la grande production populaire (DUBOIS 2005). Si l'on y ajoute son statut de traduction, que l'on pose parfois injustement comme une production plus simple — à modéliser d'un point de vue informatique par exemple (MIELKE *et al.* 2019) —, le cas de *Magyk* semble ainsi voué à l'anathème.

Pourtant, le genre de l'imaginaire est une autre preuve du succès et des passerelles intergénérationnelles de cette littérature (PRINCE 2021, p. 94–102). De plus, la traduction d'œuvres de *fantasy* fait intervenir de manière similaire la longue liste des problèmes de traduction présentée jusqu'ici, à laquelle on pourra ajouter un rapport intertextuel étroit non pas uniquement avec d'autres textes, mais aussi et surtout avec les normes du genre et avec les codes d'autres genres (GONZÁLEZ CASCALLANA 2006, p. 100–101). Les romans de la série *Magyk*, notamment, sont marqués par un registre relativement soutenu, qui autorise néanmoins une très grande variation dans le registre attribué aux différents personnages. Leur caractérisation passe en outre par l'usage de régionalismes et de termes vieillis, voire de discours adaptés du vieux français, dont l'emploi témoigne quant à lui de la grande prise de liberté adoptée de manière générale par la traductrice vis-à-vis des romans originaux. Une telle liberté est d'ailleurs caractéristique des littératures de genre et de jeunesse, où l'exigence de fidélité à l'auteur est bien moindre (SHAVIT 1986, p. 112). Plus encore, elle est nécessaire pour reconstituer dans la langue cible l'univers fictif qui s'articule autour de la série, en inventant de nouveaux concepts et en forgeant des néologismes, des « irrealia », qui permettront alors de renvoyer à l'ensemble des éléments culturels fictifs de l'œuvre (LOPONEN 2009).

En fin de compte, la question d'une prétendue hiérarchie — et, dès lors, d'une supériorité — entre œuvres ou auteurs n'a pas lieu d'être, puisqu'il s'agit en réalité de systèmes différents, qui ne peuvent donc être directement rapprochés et requièrent des outils d'analyse qui leur sont propres.

Comme le souligne Perry Nodelman, il serait ridicule d'analyser Marcel Proust et Edith Nesbit en s'appuyant sur les mêmes outils. Non pas parce qu'elle est moins bonne écrivaine que Proust, elle n'est « moins bonne » que dans le sens où l'artiste japonais Hokusai, qui a peint 36 vues du mont Fuji, serait « moins bon », par exemple, que Van Gogh. Ils sont issus de mouvements différents et ne peuvent être comparés [22]. (NIKOLAJEVA 2016, p. 58.)

Fort heureusement, cette vision a pu être dépassée grâce aux contributions de la recherche dans le domaine des études littéraires et de la traductologie (VAN COILLIE et VERSCHUEREN 2006, p. VI), renforcées en cela par la légitimation pédagogique et commerciale du secteur jeunesse (BARNABÉ 2012, par. 43). En effet, la traduction des ouvrages de jeunesse et des textes de fiction recèle grand nombre de pièges, à la fois typiques de telles œuvres et partagées avec le secteur plus large de la traduction littéraire, voire au-delà.

5.2.3 *Peut-on mesurer la complexité des textes littéraires (pour une machine) ?*

De la même manière que les caractéristiques textuelles des livres de jeunesse ont permis de montrer que ceux-ci n'étaient pas nécessairement plus « pauvres » ou plus « simples » que les romans pour adultes et que certains noms représentatifs de la production y développaient même au contraire un style plus complexe en comparaison (NIKOLAJEVA 2016, p. 6), je me suis demandé s'il serait possible d'estimer la complexité des textes littéraires et de fournir une appréciation qui expliquerait la grande variabilité des résultats donnés dans l'état de l'art. Évidemment, je ne prétends pas qu'une telle évaluation pourrait refléter quelque chose d'aussi subjectif que la sensibilité des spécialistes des langues ou de la littérature pour un ouvrage en particulier, mais elle pourrait peut-être nous renseigner sur la difficulté de leur prise en charge, d'un point de vue purement mathématique et statistique, par la machine. Peut-être pourrait-elle nous permettre d'établir un degré de « traductibilité automatique » qui viendrait objectiver — ou réfuter — les réflexions du point précédent.

Pour ce faire, TORAL et WAY (2015b) suggèrent de s'appuyer sur deux méthodes impliquées dans la mise au point des systèmes de traduction automatique, notamment statistique. La première piste avancée en ce sens concerne l'alignement des textes source et cible, et consiste à aligner mot à mot les phrases d'un texte original et de sa traduction (cf. Figure 9) pour relever ensuite la perplexité affichée par l'algorithme d'alignement. Cette mesure peut être interprétée comme un indice de confiance de l'algorithme dans l'alignement proposé, l'idée étant que si l'outil prévu pour cette tâche est capable d'aligner aisément les deux versions, par exemple si chaque mot du texte source correspond à un mot du texte cible, la traduction pourrait être considérée plus simple que dans le cas où ce même outil peine à trouver des correspondances et montre dès lors une plus grande mesure de perplexité. Cette méthode permettrait alors d'estimer la prise de liberté dans les traductions et, par voie de conséquence, la probabilité que les sorties de TA, généralement plus littérales, soient de bonne ou de moins bonne qualité. Les auteurs avaient utilisé la méthode

dans cette même étude pour comparer le degré de liberté dans les traductions issues de différents domaines, ce que j'ai entrepris de reproduire pour différents textes littéraires issus du corpus *Books* (TIEDEMANN 2012), qui comporte divers ouvrages relevant du domaine public⁴⁹. J'ai choisi pour cela les romans du corpus originellement écrits en anglais et leur traduction française, auxquels j'ai ajouté les six romans de la série *Septimus Heap* accompagnés de leur traduction, à partir desquels j'ai pu obtenir les indices d'alignement donnés par le programme *GIZA++* (OCH et NEY 2003) et repris dans Figure 20.

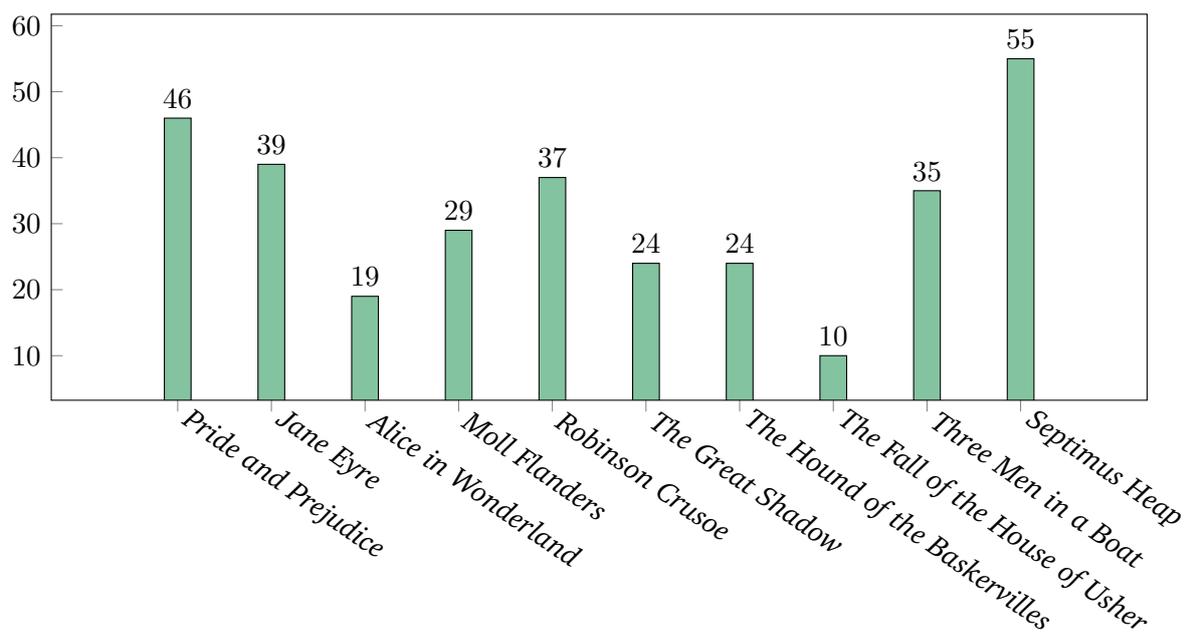


FIGURE 20 – Perplexité donnée par l'alignement de romans et de leur traduction

Comme on peut le voir, le score affiché par les traductions de la saga *Magyk* (55) place celle-ci loin devant les œuvres du corpus *Books* (10–46) et atteste de la grande prise de liberté caractéristique de ces traductions que je mentionnais plus haut. Il me faut mentionner cependant que les scores pourraient être légèrement biaisés, étant donné que les romans du corpus *Books* ont été alignés automatiquement par la machine à l'origine. Si tel était le cas, le score relevé pour ces ouvrages devrait être plus faible encore que celui mesuré ici et signaler un degré de liberté inférieur, puisqu'un problème d'alignement ferait nécessairement gonfler la valeur de l'indice. Or, on remarque toutefois que ces valeurs sont systématiquement moins élevées que pour notre texte de *fantasy*, pour lequel le résultat est plus fiable, dans la mesure où l'alignement a été effectué à la main, et qui semble présenter un obstacle singulier pour la machine.

49. Cf. <https://opus.nlpl.eu/Books.php>

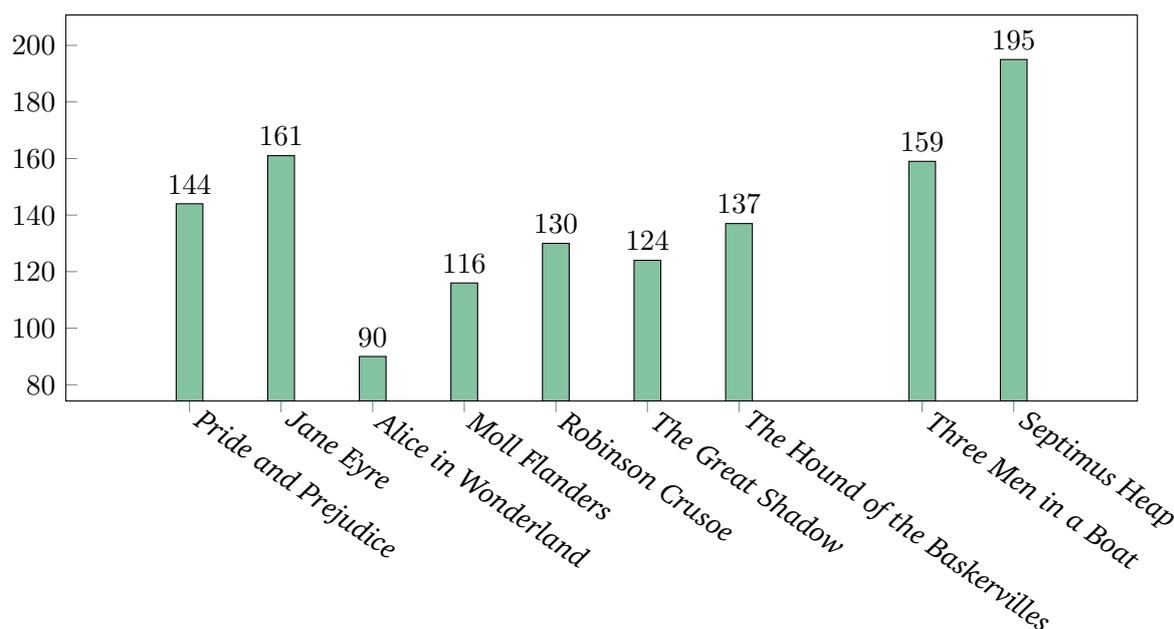


FIGURE 21 – Degré d'entropie donnée par la modélisation de divers romans

Un second moyen proposé par TORAL et WAY (2015b) implique de vérifier le degré de spécialisation du domaine (*domain narrowness*), en se référant à la perplexité des modèles de langue construits sur différents corpus. En modélisant de la même manière le langage de chacun de nos textes, comme le font en réalité les outils de TA, il est en effet possible de calculer, une fois encore, un indice de perplexité, qui mesure toutefois ici l'entropie du vocabulaire utilisé et qui nous renseignerait approximativement sur le caractère prévisible ou imprévisible des textes, sur leur richesse lexicale et sur leur diversité. Plus précisément, cet indice peut être obtenu en entraînant un modèle de langue sur une section du corpus, pour laquelle un algorithme enregistre des suites d'unigrammes, de bigrammes et de trigrammes (cf. Figure 11), et en lui présentant par la suite une section plus réduite afin de voir à quel point elle diffère de ce qui a déjà été observé. Cette façon d'envisager les choses est d'ailleurs intéressante dans la mesure où, malgré ce que l'on peut en dire, le *deep learning* est une approche purement statistique qui repose entièrement sur ce type de comptes et de probabilités. J'ai donc reproduit cette opération sur la même collection de textes littéraires, mis à part la nouvelle d'Edgar Allan Poe pour laquelle la taille des données est insuffisante, et ai relevé dans la Figure 21 la perplexité des modèles de langue donnée dans chaque cas par l'outil SRILM (STOLCKE 2002). Une fois encore, il semble que la collection des six volumes de *Magyk* obtienne le score le plus élevé (195), indiquant une plus grande difficulté dans la prédiction et la modélisation de ce jeu de données par comparaison avec le reste des ouvrages (90–161).

Enfin, une dernière façon d'estimer la complexité d'une traduction automatique observée à plusieurs reprises dans l'état de l'art consiste simplement à passer ces documents dans un outil de TA généraliste et libre d'accès pour en tirer un score BLEU (PAPINENI *et al.* 2002), et, de cette manière, une approximation de la facilité avec laquelle ce système est capable de produire une traduction proche de la référence. En passant par un service en ligne et en calculant ce score pour plusieurs œuvres avec un même outil, on peut alors avoir une vague idée de la difficulté du texte pour la TA. Dans ce dernier cas, plus le score est faible, moins la traduction est bonne ; ou plutôt, plus elle s'éloigne de celle qui est attendue. Le score nous est donné ici par l'outil *sacreBLEU* (POST 2018), sur la base d'extraits comparables d'environ 50 000 caractères correspondant dans le cas de *Septimus Heap* aux trois derniers chapitres du sixième et dernier tome de la saga, qui constituera notre jeu de test pour la suite du travail. Comme nous pouvons le voir sur la Figure 22, celui-ci arrive de nouveau à la tête du classement⁵⁰.

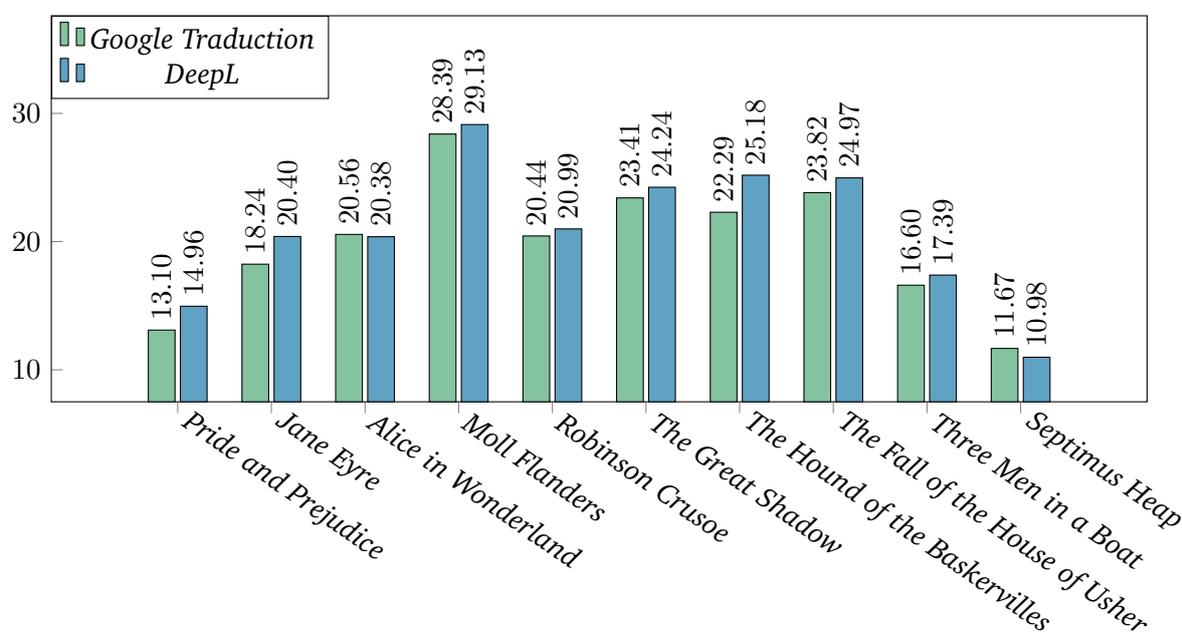


FIGURE 22 – Traduction automatique de romans par des systèmes libres d'accès

Par rapport au même essai réalisé en 2020 dans HANSEN *et al.* (2022), on peut remarquer que les systèmes ont peu changé. Si certains scores ont augmenté ou diminué, ceux-ci restent tout à fait comparables et ne présentent pas de tendance particulière.

50. Les textes comparés ici relevant du domaine public, il n'est pas impossible que les traductions soient comprises dans les données d'entraînement de *Google Traduction* ou *DeepL* et que leur présence fasse gonfler leur score, bien qu'il soit impossible de le vérifier. Les traductions ont été effectuées le 18/9/2023. Signature de *sacreBLEU* : nrefs:1 | case:mixed | eff:no | tok:13a | smooth:exp | version:2.3.1.

Ces trois analyses révèlent des propriétés très différentes qui demanderaient dans l'ensemble des recherches plus poussées, mais il convient de remarquer en premier lieu que notre ouvrage de *fantasy*, une fois comparé au canon littéraire classique, semble particulièrement complexe — il faut le rappeler — pour la machine. Ou tout du moins plus complexe que les autres textes à notre disposition et que ceux sur lesquels la littérature s'est penchée jusqu'à présent, puisque les scores obtenus dans l'état de l'art pour *Google Traduction* varient eux aussi entre 15 et 30 pour ces différents systèmes, cas d'étude et choix de langues. Si je ne prétends pas partager ici une méthode infaillible pour mesurer la complexité des textes littéraires et qu'il me faut insister sur le fait que ces calculs de perplexité et de probabilité ne reflètent en rien les phénomènes qui pourraient être intuitivement complexes ou saillants pour l'humain, la combinaison de ces trois mesures me semble donner une bonne indication de départ sur ce point. Ce corpus étant le plus propre, du fait qu'il ait été aligné manuellement, nous aurions pourtant pu nous attendre à des résultats mitigés, mais ceux-ci se confirment pour les trois tests effectués. Une telle conclusion confirme dès lors les observations précédentes concernant la littérature de jeunesse et la fiction et remet simultanément en question l'idée préconçue et relativement répandue selon laquelle la « paralittérature », par sa prétendue simplicité stylistique, lexicale et syntaxique, serait une cible de choix pour la traduction automatique.

En réalité, ce constat n'a rien de tellement surprenant, étant donné que la traduction des textes de jeunesse et de fiction amène une série de problèmes supplémentaires à ceux inhérents au domaine littéraire. Parmi toutes celles qui ont été mentionnées, les Figures 20 et 21 soulignent notamment la prise de liberté de la traductrice et la grande variation des traductions. D'autre part, la *fantasy* implique la création d'un univers fictionnel qui s'accompagne de nouveaux termes et éléments culturels fictifs et qui explique probablement les scores de la Figure 22. Ce dernier point est peut-être le plus important, car si certaines œuvres du canon littéraire sont réputées pour leur complexité, du point de vue de l'humain, elles font tout de même usage d'un vocabulaire commun que la machine est plus susceptible d'avoir déjà observé. À l'inverse, les textes de fiction ont une disposition mythopoétique qui se manifeste souvent par la création d'un monde, d'une histoire, d'une cosmogonie et surtout d'un vocabulaire qui leur sont propres et auquel les systèmes de TA n'ont probablement aucune chance d'avoir été confrontés auparavant, à moins de ne les avoir entraînés à partir d'œuvres similaires. Toutes ces difficultés font donc de ce cas d'étude une mise à l'épreuve particulièrement intéressante et exigeante pour un outil de traduction automatique personnalisé.

5.3 Données et système

Cette section détaillera de bout en bout la mise au point d'un système de traduction automatique et la procédure d'adaptation au cas d'une traductrice individuelle, en démarrant justement de la création des corpus et de la préparation des données, avant d'en venir à la mise au point et à la comparaison de plusieurs systèmes de TA⁵¹. Le chapitre s'achèvera ensuite par une brève explication des expérimentations qui n'ont pas donné lieu à des résultats satisfaisants et qui contredisent certaines des pratiques communément admises dans le domaine, à commencer par l'idée que la taille des données est une solution à la plupart des problèmes de la TA et qu'elle ne peut qu'améliorer ses performances. À ce sujet, étant donné que l'entraînement et le paramétrage technique des systèmes de traduction neuronaux est guidé par les métriques d'évaluation — il faut rappeler que c'est là leur but premier (cf. [Section 6.1](#)) —, cette section combinera réflexions méthodologiques et résultats préliminaires des métriques automatiques, sur lesquelles je m'attarderai plus longuement dans le chapitre suivant.

5.3.1 Préparation des données

Nous l'avons vu dans le chapitre précédent et continuerons de le voir tout au long de ce travail, la traduction automatique est devenue avant tout une affaire de données. Celles-ci sont indispensables au bon fonctionnement des outils fondés sur les réseaux de neurones, et cette loi immuable vaut tout autant pour la littérature que pour d'autres domaines. D'une part, il est important de pouvoir compter sur une large quantité de corpus parallèles pour entraîner un outil de TA — il s'agit d'un enjeu crucial depuis le passage du paradigme statistique au neuronal dont le besoin est devenu d'autant plus grand avec l'architecture *Transformer* (cf. [Section 4.3.1](#)) —, mais il est également nécessaire d'utiliser des données pertinentes vis-à-vis du domaine traité d'autre part. De fait, les systèmes n'atteignent des performances optimales que lorsqu'ils sont entraînés sur des corpus issus du domaine dans lequel ils seront déployés ([CHU et WANG 2018](#)), faisant de l'adaptation au domaine l'un des principaux défis posés par la TA encore aujourd'hui ([MÜLLER et al. 2020](#)). L'élaboration de moteurs de traduction adaptés à des domaines spécifiques a toutefois permis à la TA de s'étendre progressivement vers de nouveaux secteurs, jusqu'à devenir un argument de vente majeur des prestataires de service. Pour synthétiser ce double impératif, on retiendra que le volume des données permet de mettre au point un système robuste, qui produira des phrases fluides et grammaticalement correctes, tandis que les corpus spécialisés permettront à ce même système de s'adapter à une terminologie et à un style donnés.

51. Une explication détaillée de la démarche incluant les modules nécessaires, les scripts utilisés et la manière de les faire tourner est disponible à l'adresse suivante : <https://gitlab.uliege.be/dhansen/literary-machine-translation>.

Dans le cas qui nous occupe, le corpus spécialisé est représenté par la saga *Magyk*, qui vient d'être présentée et qui offre l'avantage de pouvoir élaborer un système adapté non pas uniquement au domaine littéraire, mais aussi et surtout à une autrice et à une traductrice en particulier. Ce corpus avait été constitué lors d'un précédent projet de recherche sur la traduction littéraire assistée par ordinateur (HANSEN 2017), au cours duquel les six premiers tomes de la série avaient été alignés grâce au programme *LogiTerm* (Terminotix 2000) pour former une mémoire de traduction nettoyée manuellement par la suite. Si ce corpus personnalisé d'à peu près 45 000 segments⁵² est parfaitement utile comme mémoire de traduction, il va sans dire qu'il très loin de suffire à l'entraînement des systèmes de TA neuronale, pour lesquels on utilise typiquement des dizaines, voire des centaines de millions de phrases. Or, le seul corpus de traduction littéraire disponible en ligne, à ma connaissance, est le corpus *Books* repris dans le dépôt OPUS (TIEDEMANN 2012), qui permet de rajouter une centaine de milliers de segments supplémentaires pour notre paire anglais-français. Cette collection présentant néanmoins des problèmes d'alignement et des textes pour lesquels les langues sont inversées, j'utilise ici une version corrigée de ce même corpus.

Évidemment, le scénario idéal est celui où le système peut être entraîné uniquement sur des données *ad hoc*, mais il n'est pas toujours réalisable en pratique et représente à ce titre un enjeu fondamental de la traduction littéraire automatique. Pour faire face à cet échantillon toujours très restreint de données littéraires, une solution possible et communément répandue dans les cas où les corpus parallèles sont rares consiste à entraîner le système dans un premier temps sur des données dites « hors domaine », puis à affiner (*fine-tune*) ce même système au cours d'une procédure d'adaptation au domaine qui sera détaillée au point suivant. Au bout de quelques essais, le choix s'est finalement porté sur d'autres ressources librement accessibles du dépôt OPUS communément utilisées pour la paire anglais-français : les corpus *Europarl* (v8), *GlobalVoices* (2018Q4), *News-Commentary* (v16) et *TED2020* (v1). J'y ai enfin ajouté un corpus maison de traductions de jeux vidéo décrit dans HANSEN et HOULMONT (2022), en raison du volume et de la qualité des données, dont une grande partie concerne par ailleurs des jeux relevant de la *fantasy*, mais aussi afin de diversifier la représentation des domaines autrement très axés sur la presse et les textes institutionnels. Ces divers jeux de données serviront à entraîner un système que l'on qualifiera de « générique », autrement dit un système qui n'est pas destiné à la traduction de textes issus d'un domaine spécifique. Étant donné que je cherche à évaluer ici

52. Je préfère parler ici de *segment* plutôt que de *phrase*, comme on le voit souvent en référence à la taille des corpus, car les lignes qui constituent les unités distinctives et alignées de ces jeux de données comportent fréquemment plus d'une phrase. Les traductions littéraires s'affranchissant volontiers de la structure du texte source, le cas est d'ailleurs très fréquent dans ce corpus et nécessaire pour conserver l'alignement entre les deux langues.

l'effet de l'adaptation sur un corpus de traductions individualisées, j'ai toutefois choisi d'ajouter le corpus *Books* à ces données génériques, partant du principe que cela contribuera à améliorer les performances de l'outil sur des textes littéraires, tout en illustrant de façon évidente les avantages de l'adaptation individualisée.

Dans ce même objectif, l'évaluation d'un outil de TA nécessite de séparer le corpus spécialisé, que je nommerai à présent *Septimus* en référence à la série en question, en trois jeux distincts. La majeure partie du corpus, comportant ici les cinq premiers volumes, constitue le jeu d'entraînement ou d'apprentissage (*training set*) qui servira concrètement à mettre au point et à affiner le système. Le sixième tome de la série représente quant à lui le jeu de validation (*validation set*), aussi appelé *development set* en anglais, car il permet de suivre l'évolution du système au cours de sa mise au point. En effet, le système est confronté à ces données délibérément extraites du corpus d'apprentissage à intervalles réguliers pour vérifier qu'il continue de s'améliorer, mais aussi pour déterminer quand mettre fin à la phase d'entraînement. Dans le même ordre d'idées, il est tout aussi indispensable que le dernier jeu de données, le jeu de test (*test set*) ne fasse pas partie des données d'entraînement, étant donné le système doit être évalué sur des segments auxquels il n'a pas été confronté durant sa mise au point.

Le jeu de test sélectionné pour ce faire est composé des trois derniers chapitres du sixième volume de la saga *Magyk*, que j'ai définitivement extraits du corpus de validation. Ces trois chapitres issus de *Darke* (HarperCollins 2011) sont les derniers pour lesquels il existe une traduction publiée, *La Ténèbre* (Albin Michel 2012), et offrent un ensemble d'évaluation varié qui comporte un chapitre de dénouement, rapporté de manière très descriptive, un second introduisant l'intrigue du tome suivant, avec une majorité de dialogues, et un chapitre de clôture faisant intervenir une grande série de personnages auxquels sont associés des registres variables. Malgré cela, il est intéressant de noter que ces chapitres contiennent peu de mots tout à fait inconnus (OOVs, de l'anglais *out-of-vocabulary*), ce qui atteste de la forte cohérence interne du corpus dans son ensemble. Si l'on prend les cinq premiers tomes de la série comme référence, le corpus de test ne comporte que 203 mots inconnus pour l'anglais, contre 240 pour le français. Dans les deux cas, la grande majorité d'entre eux sont des mots tout à fait communs qui n'apparaissent simplement pas dans le corpus d'entraînement (p. ex. *millimètre*, *billets*, *postuler*...). Viennent ensuite des formes composées (p. ex. *rappela-t-il*, *soixante-six*) ou conjuguées (p. ex. *profitèrent*, *rejoindra*), ainsi que des mots plus rares relevant du registre littéraire (p. ex. *industrieux*, *festonnant*) ou des termes désignant des éléments propres à la série (p. ex. *Maison-de-poupée*, *Hermeticum*), et dans de rares cas des interjections (p. ex. *Gah*, *Tss*) ou des tournures familières (p. ex. *bidule*, *bousillé*).

À partir de ces jeux de données, je présenterai deux expériences : une première détaillant pas à pas le processus d'adaptation au domaine, sur la série *Magyk* uniquement, et une seconde avec plus de données permettant de comparer différentes architectures de TA et combinaisons de corpus. L'ensemble des corpus pour cette première expérience d'adaptation est repris dans le [Tableau 1](#).

Corpus	Segments	Tokens EN	Tokens FR
Europarl	2 007 724	50 049 151	56 979 716
Corpus JV	1 370 432	21 109 925	23 358 100
TED	410 444	7 055 353	7 609 230
News	183 252	4 066 166	5 130 539
GlobalVoices	195 388	3 514 698	4 108 467
Books	127 022	2 749 441	2 841 174
Septimus (entraînement)	37 348	550 637	556 058
Septimus (validation)	7 225	108 856	109 647
Septimus (test)	704	10 042	10 348
TOTAL	4 339 539	89 214 269	100 703 279

TABLEAU 1 – Taille des corpus génériques et du corpus spécialisé

Pour faciliter le traitement de ces données par la machine, plusieurs transformations préalables sont couramment appliquées aux corpus, à commencer par la normalisation de la casse et de la ponctuation. La première méthode consiste à supprimer les majuscules du texte, de manière à éviter que des formes telles que *Magicien* et *magicien* ne soient traitées comme deux tokens différents, puis à les restituer une fois la traduction produite. Néanmoins, la série qui nous occupe ici se démarquant par un usage original des capitales, qui dénotent dans l'original des termes liés à l'univers fictif, il était évident que ce traitement n'aurait pas de sens. À titre d'exemple, voir comment la casse permet de différencier un sortilège d'un mot commun dans l'[Exemple \(1\)](#) :

(1) *The Great UnDoing had been her undoing also.*

Dans le second cas, des opérations similaires sont appliquées aux signes typographiques, de manière à harmoniser ceux-ci sous une forme standardisée. Des tests préliminaires ont toutefois montré que cela contribuait pour cette expérience à une baisse de performance, et nous verrons au moment d'évaluer les résultats plus en détail que la conversion de la typographie aux normes françaises et le respect des conventions de l'ouvrage représentent justement une force du système adapté. Pour ces raisons, aucune de ces deux méthodes de normalisation n'a été appliquée.

Une étape importante — pour ne pas dire indispensable — du prétraitement réside en revanche dans la tokenisation, dont il avait déjà été question dans le chapitre précédent (cf. Figure 8). Celle-ci a été effectuée avec le tokeniseur *Moses* (KOEHN *et al.* 2007)⁵³, sur l'ensemble des données (excepté le corpus de test en langue cible, puisque l'évaluation doit se faire sur le fichier original). Récemment, les recherches en TAL ont été plus loin encore et ont montré que l'on pouvait atteindre de meilleures performances en découpant le texte à un niveau inférieur à celui du « mot » (ou plutôt du token délimité par des espaces, ainsi qu'il en a été question jusqu'à présent). D'un point de vue technique, le principal avantage de cette segmentation en « sous-mots » se situe au niveau du vocabulaire qui est pris en compte par la machine, puisqu'elle permet de réduire celui-ci à une taille fixe prédéterminée (p. ex. 8, 16 ou 32 mille caractères et parties de mots), où tout ce qui apparaît dans notre corpus est pourtant représenté d'une manière ou d'une autre. Le procédé offre donc un découpage plus économique des données, sans lequel il aurait été nécessaire d'intégrer chaque forme conjuguée ou mot rare, pour définir ensuite une taille de vocabulaire bien plus élevée (p. ex. les 50, 100 ou 200 mille tokens les plus fréquents du corpus) et laisser irrévocablement de côté tout ce qui n'apparaîtrait pas dans cette liste. Plus encore, ce type de segmentation permet également au système de TA de gérer des mots non vus. Pour en donner un exemple simple et inventé, il suffirait que les occurrences *incroyable* et *terriblement* (découpé en « terrible » et « -ment ») aient été vues pour que la forme *incroyablement* ne soit pas considérée comme un mot nouveau. À titre d'illustration, les 203 et 240 mots inconnus des versions anglaise et française de notre corpus de test tombent respectivement à 53 et à 42 après avoir appliqué la tokenisation en sous-mots.

Texte de départ	À son tour, Septimus éventra son paquet et en tira les dents de dracule.
Tokenisation standard	À son tour , Septimus éventra son paquet et en tira les dents de dracule .
Segmentation en sous-mots	_À _son _tour _ , _Sept im us _ é vent ra _son _paquet _et _en _tir a _les _de nt s _de _d rac ule _.

FIGURE 23 – Exemple de tokenisation en sous-mots

53. Cf. <https://github.com/moses-smt/mosesdecoder/tree/master/scripts/tokenizer>.

Si certains de ces découpages semblent logiques pour nous, humains francophones, comme le préfixe de la forme *éventrer*, la désinence du verbe *tirer* ou la marque du pluriel du mot *dent* ; d'autres le sont moins, comme on le remarque avec *dracule* ou les noms propres en particulier. De fait, contrairement aux méthodes de normalisation ou de tokenisation déjà mentionnées, qui font usage de règles spécifiques aux langues données, cette procédure est purement statistique et ne fait que repérer des suites de caractères fréquemment associés, bien qu'elle reste très efficace en pratique.

On peut utiliser pour cela un autre type de tokeniseur tel que *SentencePiece* (KUDO 2018a)⁵⁴, qui implémente les méthodes de segmentation par BPE (SENNRICH *et al.* 2016b) et par modélisation d'unigrammes (KUDO 2018b). Après quelques tests, il est apparu que cette deuxième méthode donne de meilleurs résultats, sans régularisation, et qu'une taille de 16 000 tokens convient mieux à notre corpus d'entraînement (par comparaison avec une taille de 8 000 ou 32 000 tokens)⁵⁵. C'est donc ce dictionnaire de 16 000 « sous-mots » (*subwords*) qui formera le vocabulaire de notre système.

5.3.2 Mise au point du système

Une fois les corpus prétraités, il est nécessaire de choisir l'environnement logiciel qui permettra de mettre au point et de paramétrer le système de TA. De nombreuses options existent aujourd'hui pour ce faire, en ce compris notamment les logiciels *OpenNMT* (KLEIN *et al.* 2017), *Nematus* (SENNRICH *et al.* 2017), ou *Marian* (JUNCZYS-DOWMUNT *et al.* 2018), qui permettent à toute personne avec un minimum de connaissances en informatique d'entraîner à la fois des systèmes commercialisables ou, dans le plus simple des cas, un modèle réduit à des fins d'expérimentation. Mon choix s'est finalement porté sur le premier mentionné ici, *OpenNMT*, en raison de la présence de personnes de référence autour de moi et de sa documentation complète⁵⁶, mais aussi pour son accessibilité, ses très nombreuses options de paramétrage et sa communauté relativement active. Un essai a tout de même été réalisé avec l'environnement *Nematus*, confirmant que les résultats obtenus avec les deux outils sont comparables. Pour chaque système présenté, j'ai conservé les mêmes étapes de prétraitement, que je donne ici à des fins d'explications, mais qu'il est possible de traiter à la volée (directement durant la phase d'apprentissage, sans devoir modifier les corpus) avec *OpenNMT*.

54. Cf. <https://github.com/google/sentencepiece>.

55. Ce qui signifie que notre vocabulaire a été réduit à une collection de 16 000 unités qui permet de reconstituer l'ensemble des occurrences de notre corpus, soit à peu près 90 et 100 millions de tokens selon le Tableau 1.

56. Accessible ici : <https://opennmt.net/OpenNMT-py/>.

Pour illustrer les différentes étapes de l'entraînement d'un système, je me baserai tout d'abord sur une architecture Bi-LSTM avec mécanisme d'attention⁵⁷ (SUTSKEVER *et al.* 2014 ; BAHDANAU *et al.* 2015 ; LUONG *et al.* 2015), qui n'est pas la plus performante aujourd'hui, mais dont la mise au point est plus rapide et plus économique, et montrera dans un second temps la différence avec les *Transformers*. Les paramètres du système, quant à eux, sont à inférer par comparaison avec la littérature scientifique ou par intuition, en fonction de la taille des corpus impliqués, puis à modifier petit à petit par tâtonnement. De la même manière, je suis parti de configurations rapportées sur des forums de discussion d'*OpenNMT* pour des corpus de taille comparable, avant d'ajuster paramètre par paramètre la combinaison retenue. Le **Tableau 2** reprend les différents hyperparamètres considérés pour la recherche de cette combinaison optimale, ainsi que les différentes valeurs examinées. Les paramètres retenus sont indiqués en gras.

Hyperparamètre	Valeur
Segmentation	Unigramme , BPE
Taille du vocabulaire	8 K, 16 K , 32 K
Nombre de couches cachées	2, 3 , 4
Dimension des couches cachées	128, 256, 512, 1024 , 2048
Dimension des plongements	256, 512
Modèle	LSTM , GRU
Mécanisme d'attention	Général (Luong), MLP (Bahdanau)

TABLEAU 2 – Optimisation des hyperparamètres

Puisque ces nombreuses configurations impliqueraient un nombre très élevé de systèmes à réentraîner, je n'ai pas relancé d'expérience permettant de comparer les résultats des différentes combinaisons et indique simplement ici l'espace de recherche des hyperparamètres, qui avaient déjà été déterminés de façon empirique avec une précédente version d'*OpenNMT*. Finalement, les modèles LSTM ont été entraînés sur un processeur graphique (GPU) avec les paramètres suivants : 3 couches cachées (*hidden layers*) avec une dimension de 1 024, des plongements (*embeddings*) de dimension 512 et un apprentissage par lot (*batch size*) de 64 phrases. Le système a été configuré pour s'arrêter (*early stop*) après 4 étapes de validation successives sans amélioration. Avec une validation prévue toutes les 10 000 étapes, l'entraînement s'arrête automatiquement au bout de 200 000 étapes.

57. Cf. [Figure 18](#).

La traduction a été réalisée quant à elle avec les paramètres par défaut d'*OpenNMT*, qui utilise une recherche en faisceau (*beam search*) de taille 5, permettant d'améliorer la qualité des traductions en forçant le logiciel à explorer plusieurs possibilités au prix d'un coût plus élevé en calcul.

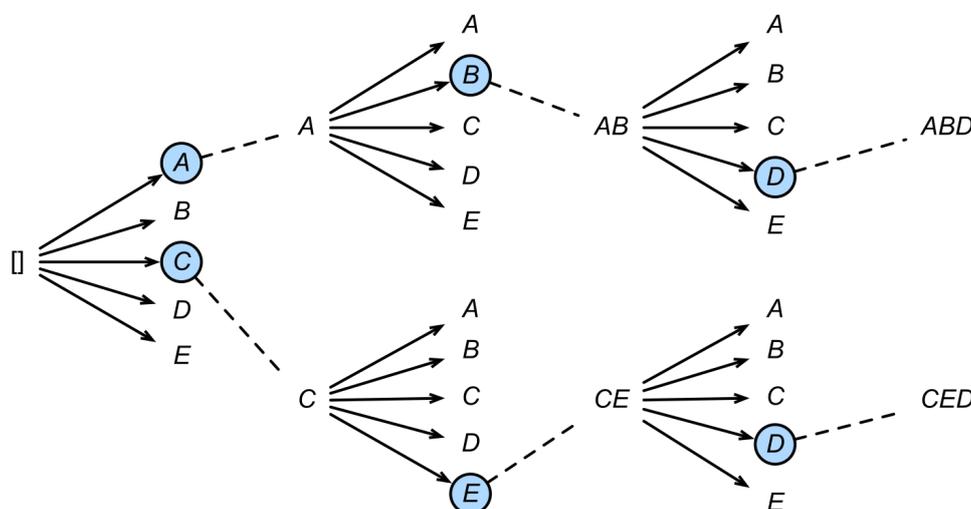


FIGURE 24 – Schématisation de la recherche en faisceau

SOURCE : A. ZHANG *et al.* (2023), CC BY-SA 4.0

Pour l'évaluation, c'est le module *sacreBLEU* proposé par POST (2018)⁵⁸ qui a permis d'obtenir une appréciation de l'amélioration des systèmes après détokenisation des sorties de traduction. Je me cantonnerai pour ce chapitre méthodologique à la seule mesure BLEU, qui, si l'on s'accorde généralement sur le fait qu'elle n'est pas représentative de la qualité d'une traduction en tant que telle, permet cependant de suivre l'évolution d'un système sur un jeu de test précis et de voir si les modifications apportées à celui-ci conduisent ou non à une amélioration. Les scores obtenus pour la mise au point et l'affinage du système sont présentés dans le [Tableau 3](#).

Système	BLEU	Différence	Durée
Système de base	06.40	—	7 h 08
+ segmentation en sous-mots	07.32	+ 14,37 %	5 h 50
+ affinage sur le corpus littéraire	13.49	+ 84,29 %	8 h 25
+ affinage mixte avec ajustement des poids	14.51	+ 7,56 %	9 h 25
+ ajout du corpus de validation aux données d'entraînement	15.59	+ 7,44 %	8 h 05

TABLEAU 3 – Amélioration du système par étapes successives

58. Cf. <https://github.com/mjpost/sacrebleu>.

Le système de base est entraîné sur l'ensemble des corpus qui ont été présentés dans le [Tableau 1](#), à l'exception du jeu de données spécialisé (*Septimus*), et offre de ce fait une base de comparaison. Celle-ci permet de voir notamment le gain de performance rendu possible par la prise en charge de la tokenisation en sous-mots, mais aussi et surtout celui lié à l'ajout du corpus spécialisé. Le processus d'affinage a été mené en relançant simplement un modèle générique après convergence (point après lequel le système cesse de s'améliorer), en lui présentant uniquement cette fois le corpus littéraire. Il est nécessaire de procéder de la sorte, en séparant ces deux phases d'entraînement en deux temps, sans quoi le système ne s'adapterait pas réellement aux données. Pour l'affinage, un arrêt a de nouveau été prévu au bout de 4 étapes sans amélioration, pour atteindre finalement un total de 300 000 étapes (générique + affiné).

Un rapide coup d'œil aux sorties de traduction à ce stade montre néanmoins distinctement que le système « oublie » des mots communs censés avoir été vus par le modèle générique, et ce, malgré des signes visibles de l'adaptation. Pour éviter ce phénomène d'oubli catastrophique (*catastrophic forgetting*), il est donc nécessaire d'opter pour une méthode d'affinage mixte (*mixed fine-tuning*), qui consiste à conserver les corpus génériques durant l'affinage tout en leur donnant un poids bien moindre que celui du corpus spécialisé, permettant ainsi au système de continuer à voir d'autres exemples pour éviter de perdre une partie de son vocabulaire ([CHU et WANG 2018](#)). Durant cette même phase, j'ai par ailleurs finalement ajouté le corpus de validation aux données d'entraînement. Sachant effectivement que tout fonctionnait correctement et ayant une idée précise de quand mettre un terme à l'apprentissage grâce aux essais précédents, la validation devenait moins essentielle, mais ajoute en contrepartie quelque 8 000 segments supplémentaires au corpus *Septimus*, afin que le système soit confronté au plus grand nombre d'exemples possible. Chacune de ces étapes mène comme on peut le voir à une amélioration significative des résultats, attestant de la bonne exécution et de l'utilité du processus d'affinage, notamment compte tenu du fait que de nombreux systèmes ont été historiquement évalués et considérés plus performants sur la base d'une différence de quelques dixièmes de points BLEU.

5.3.3 L'importance des corpus

Depuis quelques années, les architectures LSTM ont toutefois été remplacées par d'autres modèles, plus performants, mais parallèlement plus gourmands en données. Pour donner un meilleur aperçu de ce que pourrait donner un système adapté à la littérature, j'ai donc entrepris de constituer un corpus littéraire plus large, que j'ai mis à l'épreuve dans la foulée de l'architecture *Transformer*.

Pour gonfler artificiellement les données dans un domaine précis, une méthode couramment employée aujourd'hui est la création de corpus synthétiques. S'inspirant des systèmes de TA probabilistes, auxquels on fournit des données monolingues souvent plus proches des domaines de spécialité et plus faciles à récolter, cette technique revient à compiler des documents rédigés dans la langue cible et à les faire traduire par un outil de TA existant pour générer une phrase source artificielle. Elle permet ainsi d'obtenir un corpus parallèle qui peut être aisément intégré au processus d'entraînement des systèmes neuronaux, contrairement aux méthodes ayant recours à des modèles de langue par exemple, puisque le jeu de données sera alors simplement ajouté aux corpus existants. Malgré la nature factice de la source, il a été montré que ces données améliorent les performances du système et la fluidité des sorties dans la langue cible, en particulier lors des tâches d'adaptation au domaine (SENNRICH *et al.* 2016a), ce qui explique qu'on retrouve ce procédé dans les essais d'adaptation à la littérature (TORAL et WAY 2018 ; MATUSOV 2019).

De la même manière, j'ai donc cherché à constituer un corpus de littérature d'expression française détaillé dans l'Annexe B, en essayant néanmoins de ne pas surreprésenter d'autres aucteurs, étant donné que l'objectif est d'affiner le système spécifiquement sur la production d'une traductrice. J'ai dès lors veillé à diversifier autant que possible ce corpus monolingue, que ce soit notamment dans les genres, les époques ou les nationalités, et ce, malgré la difficulté d'accès aux œuvres écrites par des aucteurs en dehors de la France, la Belgique, la Suisse et le Canada. Ces choix tiennent donc avant tout à un ensemble d'intérêts personnels, d'hommages, de découvertes et de raisons pratiques, mais surtout à une série de hasards qui ne reflète évidemment en rien la diversité de la littérature francophone, bien que j'estime tout du moins ce corpus plus pertinent qu'une compilation de grands classiques tombés dans le domaine public. Pour ce faire, j'ai converti les fichiers à ma disposition au format texte et les ai fait traduire automatiquement par *DeepL Pro* (DeepL GmbH 2017) après avoir supprimé les notes et paratextes. Étant donné que les documents étaient séparés et alignés au niveau du paragraphe par défaut et que la phase d'apprentissage implique bien souvent l'utilisation d'un filtre évitant automatiquement les phrases trop longues, en accord avec l'idée que les systèmes neuronaux gèrent mal cette longueur, j'ai réalisé un test préliminaire afin de vérifier si les performances ne chutaient pas par rapport au même corpus séparé en phrases. Il s'est toutefois avéré que le corpus original séparé en paragraphes montrait une plus haute précision et une moins grande perplexité durant l'apprentissage et qu'il produisait un meilleur score pour la traduction, c'est pourquoi j'ai finalement conservé les fichiers tels quels. Une manipulation a cependant été effectuée sur le modèle de CASWELL *et al.* (2019), ajoutant une simple balise au corpus source artificiel, comme indiqué dans l'Exemple (2) :

- (2) a. <BT> Chapter 1
- b. <BT> On that first of June , Henry de Kehlmark [...]
- c. <BT> Smaragdis or the smaragdine island depends on the half-German [...]

De fait, si la version en langue originale augmente globalement le résultat, celle produite par rétrotraduction introduit a contrario du bruit non désirable, ce qu'évite cette simple procédure puisque la balise mémorisée par le système n'apparaîtra plus par la suite (MARIE *et al.* 2020). Le corpus résultant est repris dans le [Tableau 4](#) sous la catégorie *Synthétique*.

Bien que ces 150 romans ajoutent tout de même un peu moins de 400 000 segments, j'ai aussi cherché à conserver autant que possible des traductions humaines (et non synthétiques) dans l'ajout de nouvelles données. Comme pour le corpus synthétique, j'ai ainsi rassemblé des textes (cette fois écrits originellement en anglais et accompagnés de leur traduction en français), en n'y incluant pas plus de trois romans des mêmes auteurs ou traducteurs. Les 35 livres finalement choisis, relevant plutôt de la *fantasy*, ont ensuite été alignés avec *LogiTerm* pour créer un corpus parallèle (cf. [Annexe C](#)). Un second corpus similaire a également été compilé, mais comporte quant à lui 40 romans traduits par Nathalie Serval, dans l'espoir de contrebalancer l'hétérogénéité des nouvelles données et de maintenir l'accent sur le style propre de la traductrice (cf. [Annexe D](#)). Ces deux autres corpus sont repris respectivement sous les catégories *Parallèle* et *Serval* dans le [Tableau 4](#), qui présente l'ensemble des données utilisées pour la seconde expérience.

Corpus	Segments	Tokens EN	Tokens FR
Europarl	2 007 724	50 049 151	56 979 716
Corpus JV	1 370 432	21 109 925	23 358 100
TED	410 444	7 055 353	7 609 230
News	183 252	4 066 166	5 130 539
GlobalVoices	195 388	3 514 698	4 108 467
Books	127 022	2 749 441	2 841 174
Synthétique	388 234	14 373 185	15 534 342
Parallèle	100 056	3 959 591	4 468 334
Serval	111 323	3 523 594	3 693 321
Septimus (entraînement)	44 573	659 493	665 705
Septimus (test)	704	10 042	10 348
TOTAL	4 939 152	111 070 639	124 399 276

TABLEAU 4 – Taille des corpus après ajout de nouvelles données littéraires

Afin d'évaluer la contribution de ces divers jeux de données ainsi que la qualité d'un système développé à partir des nouveaux corpus, j'ai mis au point une variation du *Transformer (Relative)* décrite dans SHAW *et al.* (2018) à partir de l'architecture de VASWANI *et al.* (2017)⁵⁹. Puisque le *Transformer* est sensible à la modification des hyperparamètres, les paramètres par défaut du système ont été conservés, à savoir 6 couches cachées avec une dimension de 2 048, des plongements de dimension 512, 8 têtes d'attention, un apprentissage par lot de 4 096 tokens, une régularisation par abandon (*dropout*) fixée à 0,1 et une distance maximale relative (*maximum relative position*) de 20 pour chaque token d'entrée. L'entraînement a été mené à nouveau avec un GPU, sur un total de 300 000 étapes incluant l'affinage. Pour la comparaison avec le précédent système LSTM, j'ai par ailleurs relancé la phase d'affinage sur les mêmes données au départ du modèle générique préalablement entraîné avec cette architecture. Dans les deux cas, l'affinage a été paramétré de manière à fixer un poids minimum aux corpus génériques, un poids cinq fois plus élevé aux trois corpus littéraires ajoutés et enfin un poids deux fois plus élevé encore au corpus *Septimus*.

Le [Tableau 5](#) présente les résultats donnés par la métrique BLEU pour les systèmes LSTM et *Transformer* avec différentes combinaisons de corpus. Les trois nouveaux corpus littéraires y sont collectivement repris sous le corpus *Lettres*, tandis que le corpus *Générique* contient tous les corpus du dépôt OPUS ainsi que le corpus maison de traductions de jeux vidéo, et le corpus *Septimus* inclut quant à lui les jeux d'entraînement et de validation du même set.

Système	Corpus d'entraînement	LSTM		Transformer	
Magyk	Septimus (🔥)	12.08	2 h 05	11.38	2 h 12
Générique	Générique (🇺🇸)	07.32	5 h 50	09.77	18 h 55
Varia	Générique (🇺🇸) ⇒ Lettres (📖)	10.92	12 h 02	12.26	20 h 08
Septimus-seul	Générique (🇺🇸) ⇒ Septimus (🔥)	15.59	8 h 05	19.07	20 h 03
Septimus-étendu	Générique (🇺🇸) ⇒ Septimus (🔥) + Lettres (📖)	17.24	12 h 05	19.68	22 h 57

TABLEAU 5 – Contributions des corpus à l'amélioration du système

Comme l'illustre déjà le [Tableau 3](#), l'adaptation sur les six romans de la série à partir du modèle générique (système *Septimus-seul*) est particulièrement visible, puisqu'elle fait doubler le score obtenu par le système *Générique*. Si l'on y ajoute les trois jeux de données littéraires supplémentaires (*Synthétique*, *Parallèle* et *Serval*), l'amélioration observée avec le système *Septimus-étendu* semble s'accroître encore, bien qu'elle ne rivalise pas avec l'ajout d'un corpus moins volumineux, mais plus adapté au style de la traductrice. On notera en outre les gains de performance généralement affichés par l'architecture *Transformer* vis-à-vis des modèles LSTM.

59. Cf. [Figure 19](#).

Dans les deux cas, ces améliorations sont d'autant plus appréciables que l'affinage sur les corpus littéraires *Synthétique* et *Parallèle* (système *Varia*) mène effectivement à de meilleurs résultats, quoique bien moins élevés étant donné que les traductions de Nathalie n'y sont pas représentées. Enfin, j'ai également élaboré une variante des systèmes LSTM et *Transformer*, que j'ai entraînés uniquement sur les six romans de la saga *Magyk* à titre de comparaison. Étant donné que les hyperparamètres des architectures (nombre de couches, dimension des couches et des vecteurs...) sont sensibles à la taille du corpus (45 000 segments seulement dans ce cas, face aux 5 millions de segments environ pour les autres systèmes), j'ai utilisé pour ce faire modèles optimisés sur les recommandations de SENNRICH et ZHANG (2019) pour le LSTM et de ARAABI et MONZ (2020) pour le *Transformer*, dont je fournis un comparatif détaillé dans l'Annexe E⁶⁰. Ce dernier cas de figure apporte une preuve supplémentaire de la pertinence de ce corpus spécialisé, puisque ces quelques milliers permettent de surpasser le score du système *Générique* mis au point sur plus de 4 millions de segments — si l'on s'en tient au seul point de vue fourni par les métriques —, mais aussi de l'idée généralement admise selon laquelle les systèmes LSTM sont plus robustes dans les scénarios où peu de ressources sont disponibles, contrairement aux *Transformer* dont l'entraînement requiert beaucoup de données. Dans l'ensemble, chacune de ces comparaisons vient confirmer l'idée évoquée en début de chapitre, selon laquelle le volume des données d'entraînement ainsi que la pertinence et la qualité de ces données sont d'une importance primordiale pour la mise au point des moteurs de traduction automatique neuronaux.

A contrario, il a parfois été surprenant de voir que des méthodes généralement appliquées dans le domaine du traitement automatique des langues (TAL) n'ont pas conduit ici à une amélioration des sorties de traduction et faisaient au contraire baisser les performances du système. Cette dernière section donnera l'occasion d'explorer ces pistes de recherches, et peut-être de remettre en question certaines pratiques communément admises.

5.3.4 Résultats négatifs

À l'image du nombre croissant d'ateliers et de publications concernant les résultats négatifs du TAL, il m'a semblé pertinent d'inclure une dernière section consacrée aux expériences qui n'ont pas donné lieu à des résultats concluants. La raison en est, d'une part, que les méthodes suggérées ici apparaissent bien souvent comme une manière assurée d'améliorer les résultats et qu'elles sont

60. La version de base du *Transformer*, par exemple, produit un score BLEU de 5.91 si elle n'est entraînée que sur ces 45 000 segments. La version ajustée (moins de couches, dimensions réduites, optimisation des techniques de régularisation...) génère quant à elle un score bien plus élevé qui met en lumière la sensibilité des systèmes neuronaux aux hyperparamètres.

à ce titre souvent prévues par défaut dans la mise au point des systèmes de TA. Il convient dès lors de rester vigilant et de toujours effectuer des tests pour vérifier l'utilité de ces méthodes, bien que cette liste, d'autre part, aidera peut-être d'autres personnes qui s'inspireraient de ce travail, en suggérant d'autres approches qui n'ont pas fonctionné ici. Le [Tableau 6](#) présenté ci-dessous en dresse une synthèse, métrique à l'appui.

Système	BLEU	Différence	Durée
Système de base (<i>Transformer</i>)	19.68	—	22 h 57
Système de base + normalisation	17.92	- 8,94 %	25 h 05
Système de base + resegmentation	19.19	- 2,49 %	21 h 15
Système de base + nettoyage des corpus	19.48	- 1,02 %	22 h 29
Système de base + plongements pré-entraînés	18.96	- 3,66 %	14 h 46
Système de base + assemblage des points de contrôle (x4)	19.50	- 0,91 %	22 h 57
Système de base + assemblage de modèles (x4)	19.41	- 1,37 %	88 h 40
<i>Transformer big</i>	19.60	- 0,41 %	47 h 51
<i>Transformer big</i> + méga-corpus	19.29	- 1,98 %	50 h 00

TABLEAU 6 – Améliorations non concluantes

Comme mentionné dans la [Section 5.3.1](#), la préparation des données d'entraînement implique bien souvent plusieurs étapes de normalisation, notamment pour la ponctuation et pour la casse. Le module *Moses* propose deux scripts pour ce faire⁶¹, mais le test mené sur ce jeu de données montre que cela fait baisser le score, y compris en entraînant un modèle capable de restituer correctement la casse (*truecasing*)⁶². Cette première observation laisse penser qu'il peut être contre-productif, au moins dans certains cas, de pousser trop loin la standardisation des données d'entraînement. De la même manière et contrairement à une autre étape habituelle de pré-traitement qui consiste à segmenter les corpus parallèles pour obtenir des paires de phrases les plus courtes possibles ou à appliquer un filtre pour éliminer les segments trop longs, l'utilisation de corpus littéraires alignés au niveau du paragraphe (*corpus Synthétique, Parallèle et Serval*) ou alignés manuellement et menant à une combinaison de phrases très courtes et de segments de plusieurs phrases (*corpus Septimus*) ne fait pas baisser les performances du système. À l'inverse, cela permet d'améliorer le score donné par la métrique, et nous verrons dans le prochain chapitre que ce choix méthodologique se traduit par des observations étonnantes dans les sorties de TA.

61. Cf. <https://github.com/moses-smt/mosesdecoder/tree/master/scripts/tokenizer>.

62. Cf. <https://github.com/moses-smt/mosesdecoder/tree/master/scripts/recaser>.

Outre la normalisation et la segmentation, il existe d'autres techniques de « nettoyage » des corpus qui peuvent s'avérer particulièrement utiles dans le cas des jeux de données compilés de manière automatique, tels que ceux utilisés pour entraîner ici le système générique. Deux outils, BIFIXER⁶³ et BICLEANER⁶⁴ (RAMÍREZ-SÁNCHEZ *et al.* 2020), ont récemment été proposés pour faciliter ce processus et proposent d'automatiser certaines des opérations les plus couramment effectuées. Je me suis dès lors appuyé sur le premier outil pour éliminer les segments vides, résoudre les problèmes d'encodage, de balise et de tokenisation, corriger des erreurs d'orthographe fréquentes et supprimer les doublons, en ignorant néanmoins les options de normalisation, de resegmentation et de suppression des phrases trop longues offertes par ce même outil, étant donné que nous venons de voir qu'elles n'étaient pas nécessairement désirables. Ces opérations ont été exécutées sur les corpus génériques, auxquels j'ai également appliqué les filtres du deuxième module⁶⁵. Celui-ci vise à éliminer les données bruitées contenues dans les corpus en repérant les paires de segments pour lesquels la traduction n'est pas jugée équivalente (en raison d'une différence de longueur trop importante entre deux segments, de la présence d'une langue inattendue, etc.). Dans les deux cas, le traitement appliqué ne peut sembler que bénéfique, puisqu'il consiste à corriger des problèmes fréquents liés à la numérisation des documents et à écarter des données manifestement indésirables qui devraient faciliter l'entraînement des systèmes. De façon étonnante, cependant, le [Tableau 6](#) indique à nouveau une baisse de performance, ce qui laisse penser que les outils de TA neuronale généralisent effectivement bien en pratique — en tous cas suffisamment pour faire abstraction des anomalies dans les données — et que ces scripts sont peut-être trop agressifs, au point d'enlever des informations visiblement importantes pour l'apprentissage.

Pour ce qui est de l'entraînement, *OpenNMT* permet l'utilisation de plongements pré-entraînés (*pretrained embeddings*), de manière à lancer l'apprentissage à partir de ces représentations vectorielles (cf. [Figure 15](#)), plutôt que d'initialiser ces valeurs de manière aléatoire. Pour rappel, ce sont ces plongements de mots que le système apprend à modéliser durant cette phase et à partir desquels sont générées ensuite les traductions, l'idée étant que le recours à des modèles pré-entraînés sur d'énormes corpus peut améliorer la précision des plongements en pratique. Des plongements de type GLOVE (PENNINGTON *et al.* 2014) peuvent être téléchargés

63. Cf. <https://github.com/bitextor/bifixer>.

64. Cf. <https://github.com/bitextor/bicleaner>.

65. Un essai a été réalisé dans un premier temps en incluant les jeux de données spécialisés, mais ceux-ci sont restés inchangés en sortie, attestant de la qualité de ces corpus littéraires maisons.

pour ce faire⁶⁶, puis utilisés pour configurer à la fois le vocabulaire de l'encodeur (langue source) et du décodeur (langue cible) s'il n'y a pas de plongements équivalents pour la cible. On trouve également des vecteurs pré-entraînés pour le modèle WORD2VEC (MIKOLOV *et al.* 2013), notamment pour l'anglais⁶⁷, mais également des vecteurs pour le français, rendus disponibles par Jean-Philippe FAUCONNIER (2015) sur son site personnel⁶⁸. *OpenNMT* accepte ces deux types de plongements, après une éventuelle conversion au format texte avec le module GENSIM⁶⁹ (ŘEHŮŘEK *et* SOJKA 2010). Au lancement de la phase d'entraînement, le dictionnaire créé en pré-traitement est alors comparé aux plongements pré-entraînés, qui servent à initialiser les entrées correspondantes dans le dictionnaire⁷⁰. Les dictionnaires sur lesquels se basent le système *Transformer* pris ici comme référence ont toutefois été segmentés sous-mots, ce que devraient refléter les plongements. Par chance, le projet BPEMB (HEINZERLING *et* STRUBE 2018) met justement à disposition des vecteurs de ce type pour 275 langues⁷¹. Le score reproduit ici est donc celui donné par un système entraîné sur ces plongements anglais-français (dimension de 300), au départ du fichier de vocabulaire et du modèle de tokenisation par BPE de SENTENCEPIECE qui sont fournis avec les plongements et requis par ceux-ci (vocabulaire de 25 000 sous-mots). Dans ce cas, s'il est intéressant de noter qu'ils accélèrent fortement le processus d'entraînement (22 h 57 > 14 h 46) et que le modèle générique obtenu est plus performant (9.77 < 10.35 BLEU), on peut constater que le score obtenu n'en est pas nécessairement plus élevé après l'affinage, et ce, bien que la méthode soit surtout utile dans les cas tels que celui-ci, où les corpus d'entraînement sont peu volumineux.

Pour améliorer la précision des traductions, il est aussi possible d'assembler des points de contrôle (*checkpoint averaging*), c'est-à-dire différentes versions d'un même système au cours de son apprentissage, ou d'assembler des modèles (*model ensembling*), par exemple mis au point avec un *seed* différent⁷². Contrairement à l'utilisation de vecteurs pré-entraînés, cependant, cette méthode n'intervient pas en amont de l'entraînement mais à sa suite, au moment de la traduction.

66. Cf. <https://github.com/stanfordnlp/GloVe>.

67. Cf. <https://code.google.com/archive/p/word2vec/>.

68. Cf. <https://fauconnier.github.io/#data>.

69. Cf. <https://github.com/piskvorky/gensim>.

70. À titre d'illustration, l'utilisation des plongements de type GLOVE (6B) pour les corpus utilisés dans cette étude couvre 73,04 % des entrées côté encodeur (anglais) et 18,58 % côté décodeur (français). La combinaison de modèles WORD2VEC différents pour l'encodeur (GoogleNews) et le décodeur (frWac2Vec) couvre respectivement 89,86 % et 85,01 % des entrées.

71. Cf. <https://github.com/bheinzerling/bpemb>.

72. Chaque lancement d'un système peut donner lieu à des résultats légèrement différents et à des traductions divergentes même si relativement équivalentes. Le fait de fixer un *seed* avec une valeur particulière facilite la reproductibilité de l'expérience, mais elle permet aussi d'entraîner différentes versions d'un même système.

L'assemblage consiste alors à générer une traduction en combinant les modèles sélectionnés — quatre dans ces deux cas — qui contribuent conjointement à la prédiction de la phrase cible, bien qu'un seul modèle suffise ici à atteindre un score BLEU respectable.

Concernant l'architecture et le paramétrage des systèmes de TA, enfin, on peut remarquer que le score n'augmente pas davantage en utilisant une version plus lourde du *Transformer*⁷³ et que les résultats baissent encore plus lorsqu'on y ajoute d'autres méga-corpus supposés améliorer les performances, et ce, alors que le *Transformer big* est prévu pour traiter de plus grosses quantités de texte. Cela appuie l'observation de la sensibilité à la taille des corpus soulevée pour le *Transformer* optimisé, dans la mesure où ceux-ci sont peu volumineux dans le cas exposé ici (« seulement » 5 millions de segments parallèles pour les données d'entraînement). Ce choix était toutefois délibéré, car les autres corpus librement accessibles ne sont pas toujours de bonne qualité. Si la sélection des corpus génériques peut sembler arbitraire, en particulier dans la mesure où c'est surtout le processus d'affinage qui m'intéresse ici, celui-ci tient avant tout à une exigence de qualité qui pourrait expliquer l'issue du nettoyage des corpus. De fait, le contenu d'autres corpus habituellement utilisés et comportant un nombre bien plus élevé de segments tels que *OpenSub* (v2018)⁷⁴ ou *ParaCrawl* (v9)⁷⁵, tous deux tirés du dépôt *Opus* et totalisant 260 millions de segments supplémentaires, peut parfois laisser à désirer. Dans le premier cas notamment, les traductions ne sont pas effectuées par des spécialistes, mais par des fans. Dans le second, les données sont librement récupérées du Web avec l'objectif de compiler un maximum de traductions pouvant inclure d'autres langues, du contenu douteux, des passages non traduits ou des paires de segments qui ne sont pas équivalentes (il faut se rappeler que tout est récupéré de façon entièrement automatique). En effet, l'ajout de ces corpus semble faire baisser la qualité des traductions, y compris avec une architecture adaptée pour ce volume de données et normalement plus performante. Une lecture attentive des sorties montre d'ailleurs que les choix lexicaux produits par ce système sont moins adaptés, contiennent plus de calques et reflètent généralement ce que l'on peut trouver dans ces corpus.

Chacune des techniques explorées dans cette section constitue une méthode éprouvée du domaine, mais s'avère pourtant peu concluante pour cette expérience. Si elles sont fréquemment employées dans la littérature pour faire grimper les scores BLEU par paliers de quelques dixièmes de points,

73. Le *Transformer big* décrit dans VASWANI *et al.* (2017), dont je liste les paramètres modifiés dans l'Annexe E.

74. Cf. LISON et TIEDEMANN (2016).

75. Cf. <http://paracrawl.eu/>.

il apparaît très clairement dans ce chapitre que le succès de l'adaptation et de cette application de la TA au cas littéraire est liée avant tout au choix des corpus, qui mènent à des scores bien plus élevés comme en témoignent les [Tableaux 3](#) et [5](#). Malgré cela, il peut être surprenant de constater que l'ajout de données supplémentaires peut aussi desservir le développement du système, ce que soulève également KUZMAN *et al.* (2019, p. 5). À l'heure où la solution de la course aux performances consiste de plus en plus souvent à augmenter la taille des systèmes et le volume des données, cela montre bien que la qualité et la pertinence des corpus restent primordiales. Contredisant incontestablement le dicton « *there is no data like more data* » qui prévaut depuis la TA statistique déjà (cf. [Section 4.2](#)), ces expériences montrent dans l'ensemble que les corpus représentent le premier levier d'amélioration de la TA et que l'adaptation — que ce soit à un domaine ou à une personne en particulier — est le meilleur gage de son succès, y compris en littérature.

De manière générale, ces réflexions méthodologiques confirment en effet que si le processus d'adaptation au style propre d'une traductrice littéraire est évidemment rendu possible par les avancées techniques en matière de traduction automatique et par l'utilisation de modules tels que *OpenNMT*, dont les nombreuses fonctionnalités et possibilités de personnalisation permettent de mettre au point des systèmes dont les performances atteignent l'état de l'art, c'est la surtout quantité et la qualité des corpus qui en assurent le succès. Cette conclusion reste bien entendu en accord avec l'état de l'art présenté dans le chapitre précédent, mais illustre avec pertinence ces résultats préliminaires liés à la recherche de personnalisation des outils, qui constitue le fil rouge de ce travail. Néanmoins, il n'a pas encore réellement été fait mention jusqu'à présent des sorties de traduction elles-mêmes, ni de leur contenu. Le chapitre suivant entend précisément approfondir ce point et cherchera à voir si les gains affichés par les métriques correspondent bel et bien à une amélioration de la qualité des traductions et à un rapprochement avec le style de la traductrice.

Évaluation

Les travaux présentés et menés dans les deux derniers chapitres ont permis d'établir qu'il est tout à fait possible d'adapter un outil de traduction automatique (TA) pour le domaine littéraire, mais aussi et surtout d'envisager une personnalisation plus poussée, au niveau du style individuel des traducteurs. Ayant pu illustrer ce processus d'affinage de la TA et pointer les quelques améliorations techniques facilitant cette mise au point — en particulier l'utilisation des systèmes *Transformer* (VASWANI *et al.* 2017) —, j'ai également pu présenter quelques résultats préliminaires qui témoignent de l'importance cruciale des corpus à cet égard. Là où l'adaptation des paramètres et l'application de certaines méthodes font progressivement monter les performances, il apparaît effectivement que rien n'équivaut à la qualité et à la pertinence des données et corpus utilisés, puisque ceux-ci font plus que doubler le score obtenu par le système. Peu d'attention a été portée à ces scores, cependant, en dehors de cette étape de développement. Je clôturerai par conséquent cette deuxième partie d'ouvrage par une discussion approfondie sur l'évaluation de la TA, et plus particulièrement sur les sorties de traductions produites par l'outil personnalisé.

Si les métriques ont ponctuellement été abordées à ce stade, accompagnées de quelques avertissements, ce chapitre vise précisément à détailler les résultats et les limites de ces outils d'évaluation. Les métriques d'évaluation automatique sont effectivement essentielles à l'élaboration technique des systèmes de TA — c'est d'ailleurs à cette fin que je les ai mobilisées jusqu'à présent —, mais la dépendance et la confiance excessives en ces mesures peuvent nuire à la crédibilité de la recherche et à la validité de ses résultats. Après un détour nécessaire sur cette question, qui me conduit à adopter ici différentes méthodes d'évaluation de la qualité de la traduction (TQA, de l'anglais *Translation Quality Assessment*), j'approfondirai donc l'évaluation du meilleur système de TA exposé dans le chapitre précédent (le modèle *Septimus-étendu*)⁷⁶, tout en m'efforçant de fournir un aperçu concret de ses performances et d'éclairer les aspects qui viennent confirmer l'adaptation de la TA ainsi que les limites que présente encore un tel système.

76. Cf. [Tableau 5](#).

6.1 Évaluation automatique

6.1.1 La question épineuse des métriques d'évaluation

Lorsque l'on parle d'évaluation de la traduction automatique, le premier mot qui vient à l'esprit est probablement le terme *métrique*, et les personnes un tant soit peu familières du domaine penseront assurément à une même chose : BLEU (PAPINENI *et al.* 2002). Derrière ces acronymes particulièrement propices aux jeux de mots se cachent diverses méthodes, automatisées et quasi instantanées, pour estimer et quantifier la performance d'un ou plusieurs systèmes de TA. S'il en existe de nombreux types, en ce compris des métriques propres à d'autres tâches du traitement automatique des langues (TAL), le but poursuivi et le principe sous-jacent restent les mêmes : ces mesures quantitatives fournissent une appréciation objectivable — puisque reproductible — et surtout rapide de l'efficacité d'un outil sur la tâche considérée. Dans le cas précis de la TA, cette appréciation porte sur les sorties produites par les systèmes — qui sont le plus souvent évaluées au regard d'une traduction de référence — et se présente sous la forme d'un score pouvant baisser ou augmenter à mesure que les traductions fournies par l'outil se détériorent ou s'améliorent. Ces mesures sont particulièrement prisées en raison de la rapidité et de la facilité avec lesquelles elles permettent d'évaluer la mise au point ou de comparer les performances de ces systèmes, mais aussi et surtout parce que les évaluations humaines sont en contrepartie longues et coûteuses, en particulier si l'on sollicite pour ce faire des spécialistes de la traduction. Néanmoins, cette simplicité d'usage a fini par donner lieu à un excès de confiance, pour ne pas parler d'une dépendance de la communauté scientifique, envers les métriques et envers BLEU en particulier, que l'on retrouve précisément dans 98,8 % des articles portant sur la traduction automatique (MARIE *et al.* 2021, p. 7299). Il a pourtant été démontré depuis de nombreuses années qu'il n'y a pas systématiquement de corrélation entre ces résultats et les jugements humains, poussant de plus en plus d'équipes à remettre en question la crédibilité de certaines recherches en traduction automatique⁷⁷. De fait, l'amélioration significative d'une métrique ne correspond pas systématiquement à un réel gain de qualité et, à l'inverse, des systèmes dont l'évaluation automatique montre des scores comparables peuvent être jugés d'une manière sensiblement différente par des humains (CALLISON-BURCH *et al.* 2006). Par voie de conséquence, les systèmes finissent ainsi par être mis au point pour offrir une meilleure performance sur un corpus donné et une métrique particulière, sans nécessairement produire des traductions de meilleure qualité (BLANCHON et BOITET 2007, p. 34–35).

77. À cet usage abusif de BLEU, MARIE *et al.* (2021) ajoutent d'autres tendances problématiques, telles que la disparition d'autres métriques réputées plus fiables, l'absence de tests de signification statistique et la simple reprise de scores d'autres études comme base de comparaison — comme je l'ai fait par exemple dans les Tableaux 3, 5 et 6 — sans entraîner soi-même les systèmes ou vérifier les données utilisées.

Dans de nombreux cas, ces métriques sont l'unique gage de qualité des outils de traduction étudiés, se substituant même la plupart du temps à tout exemple concret de sortie de traduction. Au-delà de la faible corrélation entre évaluations humaine et automatique, ces mesures comportent toutefois d'autres risques et peuvent être affectées par des phénomènes divers. Il a été montré notamment que l'architecture du système de TA peut influencer le score donné par les métriques (PRZYBOCKI *et al.* 2009, p. 72–73 ; CALLISON-BURCH *et al.* 2006, p. 254–255), tout comme le traitement appliqué au jeu de test utilisé pour l'évaluation (la tokenisation par exemple) ainsi que les données elles-mêmes, mais aussi le script et les paramètres utilisés (FREITAG *et al.* 2020 ; MARIE *et al.* 2021, p. 7301–7302). La plupart de ces mesures ont par ailleurs été mises au point pour l'anglais et pourraient ne pas convenir pour d'autres cas, en particulier si l'on considère des langues dont la structure et la forme s'éloignent des langues indo-européennes (PRZYBOCKI *et al.* 2009, p. 73). Le recours à une traduction de référence comme base de l'évaluation implique en outre que le résultat affiché ne tient pas compte de la variation et peut pénaliser des traductions correctes, simplement parce qu'elles n'utilisent pas exactement les mêmes mots que la référence (CALLISON-BURCH *et al.* 2006, p. 252). Ces pénalités, quant à elles, ne rendent pas compte de la sévérité des erreurs (FREITAG *et al.* 2021, p. 1460) et peuvent s'appliquer de la même manière à une substitution opportune vis-à-vis de la source ou à des erreurs de sens graves (CALLISON-BURCH *et al.* 2006, p. 252). Enfin, puisque ces évaluations sont menées le plus souvent phrase par phrase, elles ne rendent pas compte non plus des éléments discursifs ou des problèmes nuisant à l'enchaînement logique du texte dans son ensemble (FREITAG *et al.* 2020, p. 62).

La métrique BLEU (*Bi-Lingual Evaluation Understudy*) illustre particulièrement bien ces faiblesses (MATHUR *et al.* 2020), sans doute parce qu'il s'agit de la toute première à permettre une estimation automatisée et reproductible de la qualité des sorties de TA (PRZYBOCKI *et al.* 2009, p. 72). Cependant, le problème n'est pas tellement lié à la métrique en tant que telle, prévue à l'origine pour faciliter le développement en interne des systèmes de TA et l'affinage minutieux des paramètres du système, qui seraient difficilement envisageables avec des évaluations humaines (PAPINENI *et al.* 2002, p. 311 ; FREITAG *et al.* 2020, p. 61), mais plutôt à l'utilisation abusive qui en a été faite par la suite. En effet, si celle-ci reste utile à l'élaboration des systèmes et offre une vague appréciation de ses performances, il faut en considérer le résultat avec beaucoup de prudence et surtout ne pas se limiter à cette seule perspective, surtout lorsque les scores sont très proches et que ceux-ci servent à justifier l'adoption d'une méthode ou le déploiement d'un outil.

Ceci tient aux multiples raisons listées ci-dessus, dont résulte la faible fiabilité de la métrique. Les campagnes annuelles de méta-évaluation⁷⁸ et les recherches chargées d'évaluer les métriques elles-mêmes ont ainsi révélé qu'il n'y avait qu'une corrélation moyenne à faible de BLEU, par comparaison avec des jugements humains, et que les corrélations présentées dans diverses études sont hautement variables également (REITER 2018, p. 397). De manière générale et pour diverses raisons, donc, l'amélioration d'une métrique telle que BLEU n'est pas nécessairement synonyme d'amélioration de la qualité des sorties de traduction, tandis que des scores identiques peuvent être assignés à deux traductions extrêmement éloignées. Partant, CALLISON-BURCH *et al.* (2006, p. 249) concluent « qu'une amélioration de BLEU n'est *pas suffisante* pour confirmer un réel gain de performance et qu'il n'est au contraire *pas nécessaire* d'avoir un score plus élevé pour obtenir une amélioration visible de la qualité des traductions [23] ». Ces limites sont néanmoins compréhensibles dans l'ensemble, si l'on garde à l'esprit que les mesures telles que BLEU ne font qu'évaluer la distance par rapport à la référence donnée, mais il semble y avoir une tendance croissante à prendre des raccourcis et à considérer celles-ci comme un indicateur direct de la qualité d'une traduction (cf. BLANCHON et BOITET 2007, p. 50).

Il s'est ainsi créé une situation quelque peu délicate dans le domaine du TAL, liée à l'ambivalence de BLEU. D'un côté, les recherches sur l'évaluation de la TA se concluent de plus en plus souvent par la mise au ban de la métrique, comme l'illustre de façon éloquente le chapitre des actes du colloque WMT 2022 consacré aux résultats de la campagne d'évaluation, dont le titre annonce : « Arrêtez d'utiliser BLEU – Les métriques neuronales sont plus performantes et plus robustes » (FREITAG *et al.* 2022). D'un autre côté, cette même métrique est toujours largement utilisée, au point d'être bien souvent la seule à pouvoir offrir une base de comparaison entre plusieurs études — c'est le cas pour le domaine de recherche qui nous occupe —, mais aussi parfois le seul type d'évaluation proposé⁷⁹. Ceci s'explique en partie par ce besoin de comparaison inhérent à l'avancement de la recherche et à l'implantation *de facto* de BLEU comme référentiel d'évaluation. Bien qu'il ne soit pas possible de comparer directement les scores obtenus dans différentes études, il est ainsi facile de se situer par rapport à un score faible comme on peut en voir pour des textes littéraires — mettons, inférieur à 10 — ou par rapport à l'autre extrême de l'état de l'art, qui affiche des scores supérieurs à 40 par exemple pour la combinaison anglais-français,

78. Cf. <https://wmt-metrics-task.github.io/>.

79. Dans 74,3 % des articles sur la TA selon MARIE *et al.* (2021, p. 7299), qui notent également que cette tendance va croissant, alors que pas moins de 108 nouvelles métriques ont été proposées au cours des 10 dernières années.

avec l'architecture *Transformer* (VASWANI *et al.* 2017), et qui peut monter plus haut encore pour d'autres paires de langues ou domaines d'applications. Ce qui est important, finalement, c'est surtout de ne pas motiver la mise au point ou la mise sur le marché d'un système en se référant uniquement à BLEU et, plus encore, de ne surtout pas s'y limiter pour la recherche (KOCMI *et al.* 2021, p. 479)⁸⁰, qui devrait plutôt privilégier les évaluations humaines comme critère de référence (MATHUR *et al.* 2020).

Conformément aux recommandations de MARIE *et al.* (2021), ce chapitre proposera dès lors une évaluation au travers de différentes métriques calculées via un outil standardisé et assorties d'un test de signification statistique. Pour renforcer la fiabilité de ce type d'évaluation, je multiplierai les méthodes et y ajouterai également d'autres métriques plus récentes ainsi que d'autres mesures reflétant des caractéristiques textuelles spécifiques souvent soulevées en référence aux textes littéraires, que ce soit en ce qui concerne la traduction humaine (TH) ou la TA. L'évaluation humaine décrite à la suite offrira quant à elle une analyse détaillée des points forts et points faibles des traductions produites par le meilleur système selon les métriques, au départ d'une annotation d'erreurs conduite avec une typologie élaborée sur mesure. Cette seconde moitié de chapitre servira aussi opportunément à présenter quelques observations générales et à proposer quelques exemples illustrant plus spécifiquement en quoi et dans quelle mesure le système montre des signes d'adaptation au style de la traductrice.

6.1.2 Métriques d'évaluation automatiques

En ce qui concerne plus précisément le domaine littéraire, la recherche de THAI *et al.* (2022, p. 9888) offre une mise en garde contre l'utilisation des métriques pour comparer TA et TH, constatant qu'elles favorisent la TA, quand bien même celle-ci comporte de nombreuses erreurs et est logiquement notée bien en deçà des traductions de référence lorsque des humains mènent l'évaluation. Les cas présentés par MATUSOV (2019, p. 13) et KUZMAN *et al.* (2019, p. 5) apportent pareillement confirmation des limites discutées ci-dessus et mettent respectivement en évidence la possibilité que les métriques soient contredites par des évaluations humaines ou par d'autres métriques. De plus, les expériences de TORAL et WAY (2015a, p. 254–255) indiquent que le score

80. De fait, le champ a évolué sur des conclusions tirées à partir de différences très fines, parfois à l'échelle d'un ou de quelques dixièmes de point près pour BLEU, et a historiquement considéré ces améliorations comme acquises. Depuis, les études citées ont pointé la variabilité de la métrique et ont confirmé que les classements obtenus dans de précédentes campagnes d'évaluation n'étaient pas nécessairement significatifs d'un point de vue statistique, ou que le résultat pouvait varier sensiblement sous d'autres conditions (MARIE *et al.* 2021 ; KOCMI *et al.* 2021).

est hautement variable d'une référence à l'autre et qu'une traduction humaine peut elle-même recevoir un très mauvais score si elle est plus originale, montrant finalement au mieux à quel point la traduction s'écarte ou se rapproche d'une référence spécifique. Un sujet qui est rarement mentionné, en revanche, ou dont la recherche a encore peu tiré profit — du moins ailleurs que dans les campagnes d'évaluation — est que BLEU peut accepter plusieurs références. La publication originale prévoit d'ailleurs qu'elle soit utilisée ainsi et développe l'ensemble de son fonctionnement en partant de ce principe (PAPINENI *et al.* 2002, p. 313–315). En tirant parti de plusieurs possibilités de traduction, BLEU pourrait dès lors avoir une utilité intéressante pour des cas tels que la traduction de textes créatifs, ce que pointe également l'étude exploratoire de TORAL et WAY (2015a, p. 254–255), d'autant que la recherche semble indiquer que cela augmente la validité des résultats (PRZYBOCKI *et al.* 2009, p. 83 ; BOJAR *et al.* 2013 ; FREITAG *et al.* 2020). Or, dans le cas particulier qui nous occupe ici, les mesures telles que BLEU se révèlent être d'une utilité inversement accrue, puisque l'objectif poursuivi est de se rapprocher le plus possible d'une référence bien précise : la traduction personnelle de Nathalie Serval. En effet, étant donné que je cherche à voir ici dans quelle mesure les sorties de TA se rapprochent de son style personnel, l'évaluation de la proximité formelle entre ces deux versions fournit à ce titre une appréciation pertinente, même si imparfaite. Un dernier écueil présenté à nouveau par THAI *et al.* (2022, p. 9886) concerne l'évaluation de phrases isolées, là où les textes littéraires, plus libres, sont particulièrement propices aux réorganisations syntaxiques, aux fusions et aux divisions de phrases. Dans notre cas, toutefois, la compilation manuelle du corpus nous épargne ce problème, puisque les segments ont été agrégés jusqu'à être parfaitement équivalents dans les deux langues⁸¹.

Le score BLEU des différents systèmes présentés ici a été calculé avec SACREBLEU (POST 2018), qui propose une méthode standardisée et reproductible pour l'évaluation par les métriques⁸². L'outil est aussi compatible avec TER (SNOVER *et al.* 2006) et CHR2++ (POPOVIĆ 2017), cette dernière étant citée comme la plus fiable des métriques statistiques (KOCMI *et al.* 2021, p. 479) et recommandée pour remplacer BLEU depuis plusieurs années (STANOJEVIĆ *et al.* 2015, p. 266). Pour renforcer la fiabilité de ce type d'évaluation, je prévoyais deux mesures plus récentes : YISI⁸³ (LO 2019), qui se classait parmi les meilleures métriques à sa sortie (MA *et al.* 2019), et COMET⁸⁴

81. Les observations générales présentées en fin de chapitre montreront à cet égard des résultats surprenants.

82. Cf. <https://github.com/mjpost/sacrebleu>.

83. Cf. <https://github.com/chikiulo/yisi>.

84. Cf. <https://github.com/Unbabel/COMET>.

(REI *et al.* 2020), qui est actuellement réputée être la plus proche des jugements humains professionnels pour les langues prises en charge⁸⁵, dont le français (KOCMI *et al.* 2021, p. 479)⁸⁶.

Les trois premières méthodes d'évaluation estiment la similarité ou la distance de surface entre les phrases produites par la TA et celles de la TH utilisées comme référence, partant du principe que « plus une traduction automatique est proche d'une traduction humaine professionnelle, meilleure elle doit être [24] » (PAPINENI *et al.* 2002, p. 311). BLEU, en particulier, calcule cette similarité en comparant dans les deux segments des suites de tokens (jusque quatre), appelées n-grammes⁸⁷, et en comptant le nombre de correspondances pour en donner une moyenne sur l'ensemble du corpus de test. En dernière étape, l'algorithme applique une normalisation visant à s'assurer que la TA et la TH ont une longueur comparable, pour produire un score final compris entre 0 et 100⁸⁸. La mesure CHRF (*character n-gram F-score*) compare les segments de la même manière, mais prend également en compte les suites de caractères (jusqu'à six) en plus des suites de tokens⁸⁹. J'utilise ici le variant CHRF2++. La dernière métrique de ce premier trio, TER (*Translation Edit Rate*), se fonde quant à elle sur le nombre minimum d'opérations nécessaires (additions, suppressions et permutations) pour reproduire la traduction de référence à partir de la sortie de TA⁹⁰. En ce sens, elle est censée refléter l'effort de post-édition (PE) auquel elle semble être corrélée (O'BRIEN 2011).

85. Puisque la métrique se base sur des modèles de langues pré-entraînés, elle se limite à certaines langues. Par ailleurs, la distinction entre jugement professionnel et non professionnel est très importante, car les *crowdworkers* dont la majorité des études s'attachent les services peuvent ne pas avoir la sensibilité nécessaire pour évaluer de manière fiable une traduction (FREITAG *et al.* 2021, p. 1462) et ont tendance à favoriser les traductions littérales plus faciles à évaluer (*Ibid.*, p. 1467) — un biais que doivent sûrement renforcer les conditions précaires du microtravail.

86. Les ouvrages de référence font maintenant la distinction entre les métriques dites *string-based*, qui se basent uniquement sur la comparaison formelle de caractères ou de n-grammes, et les métriques dites *pre-trained*, qui font usage des plongements lexicaux utilisés par la TA neuronale. Néanmoins, cette seconde génération de mesures se fonde également sur une phrase de référence pour calculer un score le plus fiable possible, tandis que d'autres mesures de la première génération telles que METEOR faisaient déjà usage de matériaux « pré-entraînés » pour capturer des informations sémantiques (DENKOWSKI et LAVIE 2014). Il me semble donc plus judicieux de parler dans ce cas de mesures de surface ou de mesures profondes, voire, comme je le fais ici en référence au changement de paradigme en TA qui a vu apparaître ces métriques et qui reflète les méthodes qui leur sont associées, entre métrique *statistique* et métrique *neuronale*.

87. Cf. Figure 11.

88. Où 100 correspond à un texte parfaitement similaire à la référence. Par le passé, cette mesure était plutôt exprimée dans un intervalle entre 0 et 1.

89. La mesure CHRF+ ne considère que les unigrammes (tokens isolés) en plus des caractères, tandis que CHRF++ considère aussi les bi-grammes (suites de deux tokens). Le paramètre de la F-mesure peut varier de un à trois, mais la solution optimale semble se situer au milieu, d'où CHRF2++.

90. Une autre version de cette approche, nommée HTER pour *human-targeted TER*, consiste à confronter la TA non pas à une référence, mais à cette même sortie après post-édition.

Les deux dernières métriques sont plus récentes et tirent parti des méthodes neuronales pour l'évaluation. Elles permettent en cela une évaluation sans utilisation d'une traduction de référence, bien que leur fiabilité en pâtisse dans ce cas. La première, YISI (transcription du cantonais 意思, signifiant « sens »), cherche à prendre en compte la sémantique grâce à des plongements lexicaux pré-entraînés⁹¹. Là où YISI-0 n'utilise que des formes de surface pour comparer TA et TH à la manière de BLEU ou de CHRf, YISI-1 y ajoute les plongements en question. YISI-2, en revanche, compare les plongements de la source et de la sortie de TA, permettant une évaluation qui, quoique moins fiable, ne nécessite pas de traduction de référence. Cependant, le projet associé à YISI ne semble plus être maintenu depuis plusieurs années et il ne m'a pas été possible de produire un score pour cette métrique. Comme celle-ci, COMET⁹² fait usage de représentations vectorielles, mais elle fait intervenir à la fois la sortie de TA, la source et la référence humaine. Toutefois, la métrique elle-même est un modèle neuronal spécifiquement entraîné pour prédire et reproduire trois types de jugements humains, à savoir l'évaluation directe (*direct assessment*, évoqué dans la [Section 6.2](#)), la mesure HTER et le système d'annotation d'erreurs MQM (LOMMEL *et al.* 2014).

Chacun des scores est évalué à partir des sorties de TA détokenisées, en comparaison avec la traduction humaine — qui ne faisait donc pas partie des données d'entraînement (cf. [Tableau 4](#)). Pour rappel, le corpus de test à partir duquel proviennent ces analyses reprend les trois derniers chapitres du sixième tome de la série *Magyk* d'Angie Sage, à savoir *Darke* (HarperCollins 2011), pour lequel j'ai pu me référer à la traduction de Nathalie Serval : *La Ténèbre* (Albin Michel 2012)⁹³. Le [Tableau 7](#) présente les mesures données pour le tome 6 par SACREBLEU⁹⁴, et par COMET⁹⁵. Ce tableau présente également une comparaison avec *Google Traduction* (Google 2006) et *DeepL Translate* (DeepL GmbH 2017), dorénavant « Google » et « DeepL »⁹⁶, et ce, pour deux

91. Cf. [Figure 15](#).

92. Pas d'acronyme particulier dans ce cas, excepté une subtile référence à METEOR (BANERJEE et LAVIE 2005). Avant 2022, le score de COMET apparaissait sous la forme d'un score z compris entre -1 et 1, mettant véritablement l'accent sur les comparaisons entre les systèmes. Depuis, le score a été modifié pour être compris entre 0 et 1, afin d'en faciliter l'interprétation, ce qui pourrait mener à la même situation dans laquelle se trouve actuellement BLEU, où le score absolu est considéré comme un marqueur défini de la qualité.

93. Le chapitre suivant offrira une seconde évaluation et une confirmation de ces résultats sur la base du tome 7, *Fyre* (HarperCollins 2013), pour lequel il n'existe pas de traduction officielle en français.

94. Signature pour SACREBLEU :

```
BLEU      nrefs:1|bs:1000|seed:12345|case:mixed|eff:no|tok:13a|smooth:exp|version:2.3.1
chrF2++   nrefs:1|bs:1000|seed:12345|case:mixed|eff:yes|nc:6|nw:2|space:no|version:2.3.1
TER       nrefs:1|bs:1000|seed:12345|case:lc|tok:tercom|norm:no|punct:yes|a:no|version:2.3.1.
```

95. Modèle WMT22-COMET-DA, version 2.2.0.

96. Traductions générées le 18/9/2023.

raisons : d’une part, car il est intéressant de voir les performances des outils avec lesquels le public est le plus susceptible d’être familiarisé, et d’autre part, car ces systèmes sont censés être représentatifs de l’état de l’art et, dès lors, bien plus robustes que le système générique mis au point pour cette étude sur la base de quelques millions de segments seulement. Aux mêmes fins de comparaison, j’ai repris ici un système affiné uniquement sur la série *Magyk* et nommé *Septimus-seul*, comme décrit dans le premier scénario du chapitre précédent (cf. [Tableau 1](#)), et un autre affiné sur le corpus littéraire étendu décrit à la suite (cf. [Tableau 4](#)), nommé *Septimus-étendu*. Pour chaque métrique, une flèche indique si l’amélioration du modèle est marquée par un score plus haut ou plus bas (en règle générale, plus le score augmente, meilleures sont les performances, excepté pour la mesure TER, dont le score doit idéalement être le plus bas possible).

	BLEU ↑	CHRf ↑	TER ↓	COMET ↑
Générique	09.77	33.84	93.46	0.662
Google	11.67	36.26	90.04	0.690
DeepL	10.98	35.96	91.53	0.690
Septimus-seul	19.07	41.48	78.10	0.707
Septimus-étendu	19.68	42.09	77.75	0.720

TABLEAU 7 – Comparaison des métriques pour différents modèles adaptés et génériques

Dans l’ensemble, on peut remarquer que les résultats pour les trois outils de TA génériques (c’est-à-dire non adaptés à la littérature) sont très faibles, ce qui confirme l’idée que ce texte de *fantasy* semble représenter un cas particulièrement difficile pour la machine (cf. [Section 5.2.3](#)) ou, tout du moins, que les traductions produites par ces systèmes sont très éloignées de la référence. Une autre observation intéressante est que le système générique maison obtient un score qui n’est pas très éloigné de *Google* et de *DeepL* et qu’il constitue de ce fait un point de départ assez solide pour l’affinage sur les données littéraires. De façon bien plus importante, toutes les métriques montrent que malgré la complexité de ce cas d’étude et les performances de départ très faibles, le processus d’adaptation permet en revanche d’atteindre des résultats beaucoup plus élevés. Si l’on prend pour exemple le modèle adapté uniquement sur les quelque 45 000 segments du corpus initial (*Septimus-seul*), cette approche conduit à une amélioration de ~ 10 points BLEU⁹⁷, qui se reflète à leur échelle respective dans les scores des autres mesures.

97. Une amélioration plutôt positive si l’on considère à nouveau que des comparaisons ont été faites dans le domaine sur la base de différences d’un ou de dixièmes de points BLEU.

Pour m'assurer de la validité de ces résultats, j'ai effectué un test de signification statistique (*paired t-test with bootstrap resampling*) tel que proposé par KOEHN (2004) et recommandé par MARIE *et al.* (2021). Le test a été mené pour BLEU, CHR2++ et TER avec SACREBLEU, et pour COMET avec le script fourni par ce module. L'échantillonnage du tirage a été fixé à 1 000 et le seuil de signification à $p < 0,001$ dans les deux cas. Sans surprise, les écarts rapportés dans le [Tableau 7](#) se confirment pour la plupart, excepté ceux dont les scores sont plus proches. On peut ainsi vérifier avec 99,9 % de certitude que tous les systèmes surpassent le modèle *générique*, bien qu'il n'y ait pas de différence statistiquement significative entre *Google* et *DeepL*. Le test statistique confirme par ailleurs l'amélioration de *Septimus-seul* et de *Septimus-étendu* vis-à-vis de tous les autres systèmes, mais l'écart entre les deux modèles eux-mêmes est moins clair : là où COMET indique que la différence est statistiquement très significative, il n'en va pas de même pour les trois premières métriques, y compris en diminuant l'intervalle de confiance à 95 % ($p < 0,05$).

Si l'on compare les deux modèles adaptés, justement, on peut voir que la différence entre *Septimus-seul* et *Septimus-étendu* n'est pas très grande (quelques dixièmes de points seulement pour chacune des métriques) malgré les quelque 550 000 segments supplémentaires environ. Or, une lecture attentive des sorties produites par ces deux systèmes laisse transparaître une très nette différence entre les deux : là où celui entraîné sur les six romans affiche un plus haut degré d'adaptation au style de la traductrice et un plus grand nombre de correspondances exactes avec la traduction de Nathalie, le second, entraîné sur plus de données, produit quant à lui des phrases plus fluides, un plus grand nombre de réorganisations syntaxiques et moins d'erreurs grammaticales, mais montre également des signes d'adaptation moins évidents au style de la traductrice. Comme si le style s'estompait avec l'ajout de données, en dépit des efforts consentis pour conserver une certaine représentativité des traductions de Nathalie dans ce corpus.

Stylométrie

Au vu du constat précédent, j'ai cherché à voir si une lecture distante⁹⁸ pouvait confirmer l'écart entre les deux modèles adaptés et la meilleure adéquation du système *Septimus-seul* révélée par la lecture attentive des sorties. La stylométrie est en cela une approche toute désignée, puisqu'elle vise précisément à identifier des marqueurs textuels caractéristiques du style d'une

98. S'opposant à — ou plutôt complétant — la lecture attentive (*close reading*), qui se fonde sur une analyse humaine méticuleuse telle que traditionnellement appliquée dans les études en stylistique ou en traduction littéraire, la lecture distante (*distant reading*) se rapproche d'une forme d'analyse assistée par ordinateur opérée à un niveau macro et permet de tirer des observations à une échelle plus large qui seraient autrement invisibles à l'œil nu (MORETTI 2013 ; YODALE 2020).

personne (vocabulaire utilisé, fréquence des mots-outils, usage de la ponctuation), pour faciliter ensuite les analyses et les comparaisons stylistiques. Traditionnellement utilisée pour attribuer des œuvres plagiées ou anonymes à leur auteur, la stylométrie intervient aussi parfois à des fins de comparaison (entre différentes sources de production ou différentes œuvres d'une même source). La méthode du Delta de BURROWS (2002) est une métrique fréquemment utilisée pour ce type d'analyse. Comme les métriques proposées plus haut, le Delta évalue la « distance » entre des textes, en tenant compte de la sur- ou sous-utilisation des tokens les plus fréquents d'un corpus. Concrètement, la mesure établit une liste des tokens sur l'ensemble du corpus et calcule la représentativité de chaque trait pour chacun des documents sous la forme d'un score z. Le Delta est ensuite obtenu en comparant le score du texte à évaluer et celui des autres documents, pour chacun de ces traits⁹⁹. Pour calculer la distance entre la référence et les différentes sorties de TA, je me suis servi du module STYLO (EDER *et al.* 2016), implémenté en R¹⁰⁰, et ai comparé la traduction de Nathalie Serval aux différentes sorties de TA, y compris celle des deux systèmes adaptés. Le tableau résultant affiche pour chacune la distance donnée par la formule classique du Delta, à partir des 150 tokens les plus fréquents. La similarité du style y est indiquée par un score plus faible.

Générique	Google	DeepL	Septimus-seul	Septimus-étendu
1.614	1.435	1.474	1.117	1.166

TABLEAU 8 – Distance donnée par le Delta de Burrows par rapport à la référence humaine

Le [Tableau 8](#) montre lui aussi la plus grande proximité des sorties produites par les deux systèmes personnalisés, mais il confirme surtout la meilleure adéquation du premier d'un point de vue purement stylistique¹⁰¹. Pour visualiser la distance donnée par le Delta de Burrows, le module offre la possibilité de plonger ces données grâce à une technique d'analyse en composantes principales (ACP). De cette manière, la [Figure 25](#) présente la projection d'une ACP basée sur une matrice de covariance, qui regroupe les documents les plus proches dans un espace en deux dimensions sur la base de leurs caractéristiques textuelles.

99. Chaque token repris comme trait revêt la même importance. C'est important, car la seule présence de l'auxiliaire marquant le participe passé — qui est utilisé tout au long des textes produits par *Google* et *DeepL*, mais très peu dans la traduction de référence, et qui se trouve être le critère le plus discriminant pour ce test — aurait dans le cas contraire un poids démesuré et suffirait à lui seul à opposer les TA générées par les outils génériques et adaptés.

100. Cf. <https://github.com/computationalstylistics/stylo>.

101. Le même test mené en utilisant des tri-grammes ou en s'appuyant le corpus *Fantasy* du chapitre précédent comme groupe contrôle révèle les mêmes résultats.

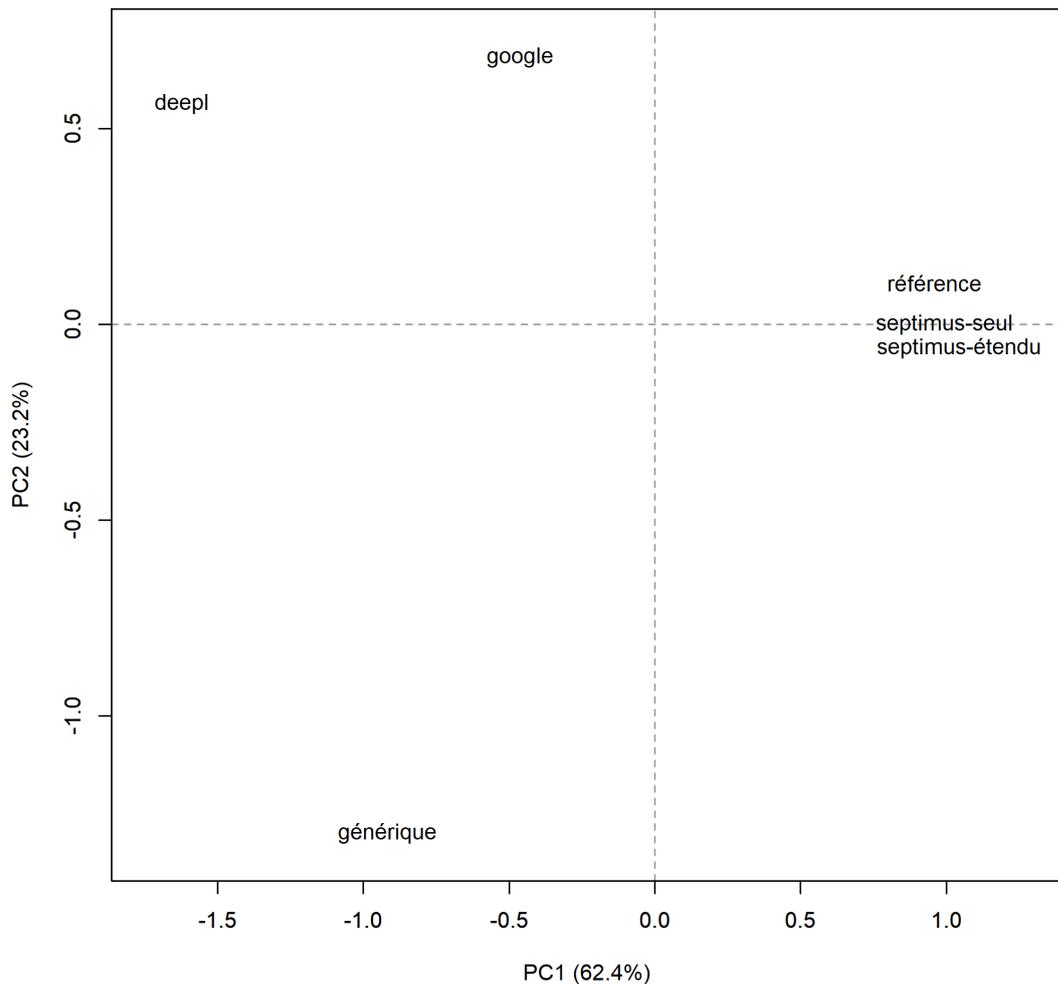


FIGURE 25 – Projection de la distance par analyse en composantes principales

L'analyse stylométrique fait ainsi ressortir le modèle *Septimus-seul* comme étant plus « proche » de la référence. Cet équilibre entre fluidité et représentativité de la voix individuelle de la traductrice reste donc à ce stade un problème auquel il faudra remédier, mais il n'est pas marqué dans les métriques, qui indiquent simplement une légère amélioration. Seule COMET laisse plus ou moins entrevoir cet écart, pour lequel il y aurait fort à parier que la différence soit liée aux modèles de langue utilisés par la métrique, qui favoriserait la fluidité bien que le système *Septimus-seul* démontre une meilleure adéquation avec le style de la traductrice. On peut effectivement noter à cet égard la mise en garde de KOCMI *et al.* (2021, p. 481) concernant ces mesures pré-entraînées :

Cette catégorie de méthodes renvoie aux métriques qui font usage de modèles neuronaux pré-entraînés pour évaluer la qualité des sorties de TA en tenant compte de la source, de la référence humaine ou des deux. Elles ne dépendent pas strictement de la qualité de la référence (elles peuvent évaluer avec plus de précision les synonymes ou les paraphrases, par exemple).

Cependant, leur performance est influencée par les données à partir desquelles elles ont été entraînées. Plus encore, les modèles pré-entraînés font apparaître un problème de boîte noire, puisqu'il est difficile de prévoir et de résoudre des comportements potentiels de la métrique tels que des biais associés aux données d'entraînement [25].

L'équipe de recherche formule d'ailleurs précisément cette hypothèse de biais favorisant la fluidité plutôt que l'adéquation (*Ibid.*, p. 487)¹⁰², ce que confirme cette expérience. Nous tombons à nouveau ici sur les limites des métriques qui ont déjà été longuement discutées en ouverture de ce chapitre. Néanmoins, la combinaison des différentes mesures statistiques et neuronales, ainsi que la marge d'amélioration affichée par chacune entre les trois modèles génériques et les deux modèles affinés, illustrent clairement l'intérêt d'une adaptation au style propre de chaque traductaire, comme le présentait déjà le [Tableau 5](#).

Mise en perspective

Pour avoir une idée plus précise de ce que représentent ces scores (et BLEU en particulier, puisqu'il s'agit de la métrique la plus représentée dans la littérature scientifique), nous pouvons mettre les choses en perspective de plusieurs manières. Nous pouvons tout d'abord comparer ces résultats à ceux rapportés dans le [Chapitre 4](#) pour les autres adaptations littéraires, dont je fournis un comparatif dans le prochain tableau. Il me faut toutefois rappeler au vu des mises en garde précédentes que le score donné par une métrique telle que BLEU n'est pas représentatif de la qualité réelle d'une sortie de TA ([PRZYBOCKI et al. 2009](#), p. 72)¹⁰³, et qu'il n'est pas non plus directement comparable d'une expérience à l'autre. Les scores exacts donnés par ces métriques n'ont pas de sens dans l'absolu et peuvent varier fortement, par exemple si l'évaluation se fait en référence à une traduction standard ou à une post-édition, comme le propose la mesure HTER et comme le fait [BESACIER \(2014\)](#) avec BLEU, ou encore si les langues en présence sont plus proches, ce qui est un des postulats expliquant la différence entre les deux études de [TORAL et WAY \(2015a,b\)](#). Les métriques, de fait, mesurent uniquement à quel point la sortie de TA est proche d'une référence donnée. Or, si cet objectif s'accorde avec les ambitions de ce travail, les traductologues imagineront tout de suite les limites plus générales d'une telle approche, puisque « de nombreuses variations d'une même phrase peuvent en préserver le sens, tandis que des

102. Pour une explication de ces deux termes, voir [Section 6.2.1](#).

103. La différence que je viens d'évoquer entre les deux modèles adaptés illustre ceci de façon éloquente, puisque les deux systèmes semblent avoir des scores similaires, mais produisent en réalité des traductions qui ne sont pas comparables. Un autre exemple lié à l'ambiguïté de BLEU, plus parlant encore, pourrait être le contraste entre les modèles *Magyk* et *Générique* repris dans le [Tableau 5](#), qui affichent un score proche, bien que le premier produise des mots de vocabulaire correctement traduits dans des phrases tout à fait incompréhensibles, là où le second génère des phrases correctes mais totalement éloignées du champ lexical du roman. Ici aussi, ces différences s'expliquent par la différence entre les corpus qui ont servis à l'entraînement des systèmes.

changements de surface très localisés peuvent modifier ou même inverser celui-ci [26] » (BOJAR *et al.* 2013, p. 465). Ce type d'évaluation permet donc de confronter différentes approches, pour autant qu'il y ait une base de comparaison valable.

En règle générale, on se bornera à évaluer un même système sur une série de textes, afin d'éprouver sa robustesse au regard des données d'entraînement, ou l'on cherchera inversement à confronter sur un même jeu de test plusieurs systèmes mis au point avec différents jeux de données ou différents paramètres, afin de voir quelle méthode produit les meilleurs résultats (pour ce corpus de test). Comparer un score obtenu par deux systèmes fondamentalement différents sur un jeu de test distinct, en revanche, et plus encore sur deux paires de langues éloignées, ne révélerait aucune sorte d'information. Pour ces raisons, la lecture du [Tableau 9](#) ci-dessous doit se faire de manière horizontale uniquement. Le tableau en question reprend les autres cas d'adaptation à partir de données génériques, de même que le modèle affiné de cette étude (*Septimus-étendu*), qui présente les scores les plus élevés pour chacune des métriques et qui constituera à ce titre le modèle de référence pour la suite des évaluations. Pour chaque combinaison linguistique, on pourra ainsi comparer l'évolution du score BLEU sur un jeu de test donné entre systèmes générique et adapté à la littérature, avec un score respectivement affiché pour chacun de ces tests par *Google*, qui est aussi mentionné dans ces études et qui offre une sorte de point de référence (bien que le système ait évolué entre les moments où ont dû être générées les traductions).

	Générique	Affiné	Google
EN-RU [LSTM] (MATUSOV 2019)	14.20	15.20	13.90
DE-EN [Transf.] (MATUSOV 2019)	18.50	16.20	20.20
EN-SL [LSTM] (KUZMAN et al. 2019)	17.50	20.75	21.97
EN-FR [Transf.] (cette étude)	09.77	19.68	11.67

TABLEAU 9 – Comparaison avec d'autres adaptations de systèmes génériques à la littérature

La comparaison offerte par ce tableau permet de constater deux choses. La première est que l'adaptation sur des données littéraires peut parfois sembler négligeable, voire contre-productive si l'on s'en tient strictement aux métriques, mais le cas développé par [KUZMAN et al. \(2019\)](#) offre une amélioration plus appréciable, que l'équipe de recherche parvient à obtenir en ajoutant aux données d'entraînement du modèle un texte écrit et traduit par les mêmes personnes que le roman évalué. La deuxième est que les résultats présentés dans ces deux études restent proches

ou même inférieurs à ceux donnés par *Google*. De plus, ces scores sont globalement plus élevés que celui donné par ce même système sur la série *Magyk* dans ce travail, s'élevant péniblement au-dessus de 10 BLEU. Dans le cas qui nous occupe, cependant, l'entraînement sur des données adaptées à la fois au domaine et à la traductrice donne lieu à une amélioration bien plus marquée. Celle-ci est d'ailleurs d'autant plus significative que le score du système générique et celui de *Google* font ressortir une fois encore la complexité du corpus choisi pour la machine, mais aussi parce que le prototype personnalisé mis au point repose sur un jeu de données dont la taille d'environ cinq millions de segments paraît bien « frugale » si l'on considère que *Google*, *DeepL* et les autres systèmes du marché sont entraînés en pratique sur des corpus de dizaines de millions, voire plusieurs dizaines de milliards de segments (BLIN 2021). Dans tous les cas, ces parallèles mettent eux aussi indéniablement en évidence l'intérêt du processus d'adaptation, signalant il me semble l'avantage d'une approche visant à mettre au point des systèmes de traduction automatique personnalisés pour les traducteurs.

Pour situer plus largement ces scores BLEU, on peut faire un dernier rapprochement avec la recherche de TORAL *et al.* (2020), qui évaluent un ensemble de 12 romans traduits par un système *Transformer* entraîné uniquement sur de la littérature et obtiennent pour la paire anglais-catalan des résultats compris entre 15 et 45. Quoiqu'un peu plus élevés, ceux-ci sont comparables avec la Figure 22, qui montrait quant à elle un intervalle entre 10 et 30 pour les 10 romans évalués avec des outils génériques sur la paire anglais-français. Pour les quatre exemples du Tableau 9, on peut dès lors noter que les évaluations donnent des résultats qui se situent dans une tranche plutôt inférieure, surtout si l'on se réfère à l'indication donnée par TORAL et WAY (2015a, p. 249), qui situent un score BLEU de 20 au minimum comme point de repère général marquant une sortie de traduction utile à la post-édition. Cela signifie que nous n'en sommes pas encore au stade d'un outil complètement fonctionnel, et encore bien moins de la traduction de textes littéraires entièrement automatisée et de haute qualité, mais que l'adaptation donne tout de même des résultats très encourageants et qu'il y a du potentiel pour un outil d'aide à la traduction personnalisé, capable d'assister les traducteurs en fournissant des suggestions sur mesure.

Les différents rapprochements que je me suis risqué à établir ici doivent néanmoins être pris avec un peu de recul et ont uniquement vocation à situer — même vaguement — le score des mesures discutées, pour les personnes qui pourraient être moins familiarisées avec les métriques et avec l'évaluation de la TA. En effet, cette section illustre aussi la grande variation qui peut exister dans les résultats rapportés pour diverses œuvres, systèmes ou paires de langues, et

je me suis déjà longuement penché sur le manque de fiabilité des métriques, y compris pour les mesures neuronales plus récentes qui privilégient visiblement la fluidité à l'adéquation et présentent peut-être d'autres biais insoupçonnés. La différence que j'ai soulignée plus haut entre les systèmes *Septimus-seul* et *Septimus-étendu* témoigne ainsi du biais lié à la dépendance aux données et confirme que les métriques restent utiles pour pointer de larges écarts ou pour donner une estimation de la performance globale du système, mais en aucun cas pour comparer des systèmes qui seraient trop similaires. De cette façon, les comparaisons proposées me permettent de formuler une appréciation de mes résultats et de situer ceux-ci dans le panorama plus large et encore limité de la création d'outils adaptés à la littérature, mais il est impératif que celles-ci soient accompagnées d'une évaluation humaine plus détaillée, voire d'autres mesures automatiques, et ce, malgré le signe clair de l'amélioration dans ce cas-ci.

Pour apporter une confirmation supplémentaire de l'intérêt du processus d'adaptation et descendre un peu plus dans le détail, il est possible de mobiliser d'autres types de mesures permettant de mettre en évidence d'autres caractéristiques textuelles — les mêmes qui ont pu servir par le passé pour illustrer la complexité des romans jeunesse, parfois même supérieure à la littérature principalement destinée aux adultes (NIKOLAJEVA 2016, p. 6). De fait, les trois mesures proposées par la suite couvrent chacune des aspects littéraires qui revêtent une importance particulière dans le contexte de la traduction automatique, à savoir notamment la *richesse lexicale* (VANMASSENHOVE *et al.* 2019), le *coefficient de foisonnement* (CASTILHO et RESENDE 2022) ou la *proximité syntaxique* entre texte source et texte cible (AHRENBURG 2017), et qui reflètent respectivement les phénomènes de simplification, d'explicitation et d'interférence mis en avant par Mona BAKER (1996) et Gideon TOURY (1995). Bien que leur généralisation à la traduction humaine dans son ensemble ait depuis été remise en question par la recherche en traductologie de corpus¹⁰⁴ (CHESTERMAN 2004) — ayant même été réfutée d'un point de vue purement technique (MIELKE *et al.* 2019) — et que ce même débat s'applique aujourd'hui à la post-édition, ces caractéristiques sont de plus en plus fréquemment invoquées dans les études sur la traduction automatique en tant que marqueurs d'évaluation (JIMENEZ-CRESPO 2023). De la même manière et sans chercher à généraliser cette démarche, je les ferai intervenir ici comme point de comparaison entre traduction automatique et traduction humaine, pour illustrer les limites et les avancées de cette première dans le cas particulier qui nous occupe, au travers d'aspects que l'on peut considérer comme caractéristiques du style individuel des traducteurs.

104. On peut en effet rapprocher l'idée qu'une traduction est nécessairement plus « simple » avec le même constat posé (et infirmé) pour la littérature jeunesse.

6.1.3 Richesse lexicale

Plusieurs études se sont penchées sur divers phénomènes tels que la diversité ou la richesse lexicale pour apporter un meilleur éclairage sur les différences entre TA et TH (VANMASSENHOVE *et al.* 2021 ; DE CLERCQ *et al.* 2021). Parmi les conclusions tirées de ces investigations, il apparaît notamment que les sorties de TA sont plus pauvres du point de vue de la diversité lexicale que les traductions humaines. Sans surprise, ce constat se tient pour les textes littéraires également, appuyant de ce fait la pertinence de ce type d'évaluation en plus des métriques traditionnelles (WEBSTER *et al.* 2020, p. 18). En effet, bien que la traduction automatique neuronale (TAN) se distingue de la traduction automatique statistique (TAS) par sa capacité à prendre en charge des textes affichant une plus grande diversité lexicale selon BENTIVOGLI *et al.* (2016), la comparaison de cette première avec la traduction humaine fait souvent ressortir la richesse lexicale comme un des points faibles de la TA encore aujourd'hui. Elle le serait même plus encore que pour la TAS selon TORAL (2019, p. 276), ce que l'on peut attribuer dans les deux cas au fonctionnement statistique de ces méthodes fondées sur le calcul de probabilités observées dans les corpus (VANMASSENHOVE *et al.* 2019 ; LONG 2022, p. 737).

Cherchant à vérifier si l'adaptation des systèmes apporte des améliorations vis-à-vis de ces caractéristiques textuelles, j'ai ainsi cherché à mesurer la richesse lexicale pour les différents systèmes évalués. De nombreuses méthodes ont été proposées à cet effet, à commencer par les modules KORPUS pour R (MICHALKE 2021)¹⁰⁵ ou LEXICALRICHNESS pour Python (SHEN 2022)¹⁰⁶. Dans tous les cas, cette tâche requiert une étape préalable de lemmatisation des textes à évaluer, en particulier pour une langue comportant de nombreuses formes fléchies comme le français. Il existe ici aussi plusieurs solutions pour cette tâche, notamment TREETAGGER (SCHMID 1994)¹⁰⁷, SPACY (HONNIBAL *et al.* 2020)¹⁰⁸ et STANZA (QI *et al.* 2020)¹⁰⁹. Étant donné que la qualité de la lemmatisation revêt une telle importance sur la précision des mesures, j'ai opté pour STANZA, qui s'est placé premier lors de la campagne d'évaluation *CoNLL 2017* effectuée sur plusieurs langues, dont le français, et plusieurs tâches, dont la lemmatisation (ZEMAN *et al.* 2017). L'interface de STANZA étant prévue en Python, j'ai donc aussi choisi le module LEXICALRICHNESS, par simple souci d'harmonisation et de facilitation du travail.

105. Cf. <https://github.com/unDocUmeantIt/koRpus>.

106. Cf. <https://github.com/LSYS/LexicalRichness>.

107. Cf. <https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger>.

108. Cf. <https://github.com/explosion/spaCy>.

109. Cf. <https://stanfordnlp.github.io/stanza>.

Après avoir lemmatisé la traduction de référence et les trois sorties de TA prises pour comparaison, j'ai donc relevé le nombre d'occurrences (tokens) et de mots uniques (types) de chaque document, de manière à calculer pour chacun le ratio type-token (TTR, de l'anglais *type-token ratio*) :

$$TTR = \frac{\text{types}}{\text{tokens}}$$

Cette mesure couramment utilisée est censée apporter une idée de la richesse lexicale des textes en confrontant le nombre de mots différents et le nombre total de mots, mais peut être biaisée par la longueur des documents et par les erreurs produites par la machine, qui se retrouvent comptabilisées parmi les types et font augmenter par conséquent le score de richesse lexicale (TEZCAN *et al.* 2019, p. 43 et 46). Pour cette raison, j'ai également calculé le ratio type-token basé sur un extrait moyen ou MSTTR, de l'anglais *mean segmental type-token ratio* (JOHNSON 1944), ainsi que le ratio type-token basé sur une fenêtre mobile ou MATTR, de l'anglais *moving average type-token ratio* (COVINGTON et MCFALL 2010), qui sont supposés être de meilleurs indicateurs du TTR. Ces deux méthodes proposent une estimation plus fine du TTR en effectuant une moyenne sur des segments comparables (une fenêtre définie ici à 500 tokens comme suggéré dans les deux références), avec cette différence que le MSTTR opère ce calcul sur des fenêtres contiguës qui ne se chevauchent pas et que le MATTR déplace progressivement cette fenêtre un token à la fois. L'ensemble des résultats est présenté dans le [Tableau 10](#).

	Tokens	Types	TTR	MSTTR	MATTR
Référence humaine	10 348	1 900	0.18	0.45	0.45
Septimus-étendu	10 307	1 715	0.17	0.42	0.42
Google	11 374	1 697	0.15	0.41	0.41
DeepL	11 407	1 685	0.15	0.40	0.40

TABLEAU 10 – Mesures de richesse lexicale

Sans surprise, la traduction humaine utilisée comme référence affiche le plus haut degré de richesse lexical pour chacune des mesures proposées, y compris selon le TTR simple, qui peut parfois attribuer un score plus élevé à la TA en raison des nombreuses erreurs de traduction (cf. TEZCAN *et al.* 2019, p. 46). Chacune de ces méthodes confirme également l'amélioration du modèle personnalisé en la matière, par rapport aux versions produites par *Google* et par *DeepL*. Une observation surprenante, en revanche, met en évidence la différence considérable dans le

nombre total de tokens entre ces deux systèmes par rapport à la référence et au système adapté, dont l'écart ne les empêche pas de présenter une plus grande diversité lexicale et peut s'expliquer facilement par le coefficient de foisonnement trouvé dans chacun des textes¹¹⁰.

6.1.4 Longueur des segments

Bien que les traductions humaines tendent à être plus longues que les textes originaux¹¹¹, Lars AHRENBERG (2017, p. 24) et Antonio TORAL (2019, p. 277) ont observé que la TA produisait des phrases plus courtes, évoquant une possible interférence de la machine à cet égard. Afin de vérifier si cette observation vaut également pour ce cas littéraire, j'ai dès lors mesuré le coefficient de foisonnement de chaque version en calculant la différence entre la longueur de la source et celle de la cible, normalisée par la longueur de la source :

$$\text{Foisonnement} = \frac{\text{tokens}_{\text{source}} - \text{tokens}_{\text{cible}}}{\text{tokens}_{\text{source}}} \times 100$$

J'y ai ajouté la longueur moyenne des segments, qui est donnée simplement par la somme des tokens divisée par le nombre de segments (704). Les résultats sont indiqués dans le [Tableau 11](#).

	Longueur moyenne	Coef. de foisonnement
Texte source	14,26	—
Référence humaine	14,70	+3,04 %
Septimus-étendu	14,64	+2,64 %
Google	16,16	+13,26 %
DeepL	16,20	+13,59 %

TABLEAU 11 – Longueur moyenne des segments et coefficient de foisonnement

Le jeu de test sur lequel portent ces comparaisons contient un grand nombre de segments extrêmement courts, tels que :

- (1) a. Marcia sourit.
b. — Oh !

110. Cette différence était curieusement plus marquée encore pour le système adapté dans HANSEN et ESPERANÇA-RODIER (2023), bien que la traduction avait été produite par un système comparable. Dans ce cas, le nombre de tokens était même inférieur à la source (10 042), mais affichait exactement la même richesse lexicale.

111. Ceci s'ajoute évidemment aux différences grammaticales entre les deux langues. Pour les traductions de l'anglais vers le français, COCHRANE (1995, p. 177) mentionne des coefficients de foisonnement allant de 10 à 30 %.

Toutefois, ceux-ci côtoient d'autres segments qui contiennent couramment deux phrases ou plus (jusqu'à cinq) et qui peuvent être sujets à d'importantes réorganisations syntaxiques. Malgré l'alignement manuel du corpus visant à obtenir des unités de traduction les plus courtes possibles tout en s'assurant qu'elles contiennent exactement les mêmes informations, ce scénario est si fréquent que je préfère parler ici de *segments* et non de *phrases*. Dans l'ensemble, la longueur moyenne de ces segments est tout de même très proche si l'on compare la source et la référence, ce qui s'explique par le fait que la traductrice a largement recours aux omissions d'informations non essentielles et utilise abondamment l'implicitation et les contractions syntaxiques, de manière à produire un texte qui contienne approximativement le même nombre de pages que l'original ¹¹². De cette manière, un aspect tel que la longueur des segments est tout autant caractéristique du style individuel des traductrices (cf. *CORPAS PASTOR et al. 2008*, p. 77). Nous verrons d'ailleurs au chapitre suivant que cette stratégie est effectivement un élément récurrent dans les choix opérés par Nathalie (cf. p. 173) et qu'elle lui permet ainsi de livrer une traduction beaucoup plus synthétique et économique, en accord avec les stratégies de concision et de respect des normes éditoriales définies par Rodica DIMITRIU (2004).

Ceci se note dès lors dans le coefficient de foisonnement observé, qui se montre peu élevé pour la référence, en particulier si on le compare au ratio bien plus élevé des deux outils de TA génériques. Or, on peut remarquer au contraire que le système personnalisé produit des segments plus courts encore que la traduction de Nathalie. En ce sens, il se rapproche de la stratégie employée par la traductrice, mais semble aussi légèrement exagérer cette tendance. Aussi étonnant que cela puisse paraître, cette remarque se confirme dans l'annotation d'erreurs proposée à la suite, révélant des signes d'adaptation surprenants comme le démontre le cas particulier des omissions. Face à ces résultats, on peut pareillement se demander si une telle observation se retrouve d'un point de vue syntaxique.

6.1.5 Équivalence syntaxique

Après avoir analysé les traductions du point de vue de leur longueur et de leur richesse lexicale, je me tourne à présent vers ce qui est probablement le principal écueil de la TA dans le champ littéraire (et de manière générale), à savoir sa tendance à suivre de très près la structure du texte source. Cette mesure est également inspirée par TEZCAN *et al.* (2019), bien que j'utilise

112. Les romans anglais de la saga comptant entre 500 et 700 pages, il pourrait évidemment être problématique que la traduction s'en écarte trop largement.

pour ce faire l'outil ASTRED proposé par VANROY *et al.* (2021)¹¹³, un autre module Python poussant plus loin cette réflexion et proposant trois mesures différentes de la similarité entre une phrase source et une phrase cible, respectivement axées sur la réorganisation syntaxique (ordre des mots), les changements de catégories grammaticales (étiquetage morphosyntaxique) ainsi que les différences de structure (arbre de dépendance). Au vu de la taille du corpus de test et du nombre de traductions, le module est utilisé ici dans son mode entièrement automatisé. Puisque l'évaluation fait intervenir des phrases dont les mots sont alignés et étiquetés, j'ai effectué l'alignement avec AWESOMEALIGN (DOU et NEUBIG 2021)¹¹⁴, tandis que l'étiquetage morphosyntaxique a été effectué une fois encore avec STANZA. En outre, comme les phrases doivent impérativement être comparées individuellement et que ce jeu de test contient de nombreuses phrases n'ayant pas de correspondance exacte, ainsi que je viens de le mentionner au point précédent, j'ai écarté de l'évaluation tous les segments contenant plus d'une phrase dans au moins une des quatre traductions comparées. J'ai pu extraire de cette façon 396 phrases comparables sur les 704 segments initiaux.

Selon VANROY *et al.* (2021, p. 291), les trois mesures d'ASTRED soulignent chacune des phénomènes linguistiques particuliers, mais fournissent dans l'ensemble une appréciation de l'équivalence syntaxique entre les deux langues et, partant, du caractère libre ou littéral de la traduction et de sa complexité. La première d'entre elles apporte une estimation du degré de réorganisation par le biais de croisements de groupes syntaxiques (SACr, de l'anglais *Syntactically Aware Cross*). Concrètement, ces croisements correspondent aux points d'intersection résultant de l'alignement déjà présenté sur la Figure 9 : si deux termes sont intervertis, par exemple, leur alignement se recoupe et crée un croisement. Plus la phrase est sujette à réorganisation, plus il y a de croisements. Comme le montre l'exemple de la Figure 26, cependant, comparer les mots un à un peut donner lieu à un très grand nombre de croisements (24 en pointillés) si un élément est renvoyé en fin de phrase, c'est pourquoi la mesure SACr effectue des groupements sur le plus grand nombre d'unités contiguës, réduisant le nombre de recoupements (2 en vert). Ces groupements sont néanmoins arbitraires et excessifs étant donné qu'ils ne sont pas linguistiquement motivés, c'est pourquoi les groupements tiennent également compte des informations grammaticales données par l'étiquetage morphosyntaxique et sont limités aux syntagmes linguistiquement motivés, atteignant ainsi un juste milieu (3 en noir).

113. Cf. <https://github.com/BramVanroy/astred>.

114. Cf. <https://github.com/neulab/awesome-align>.

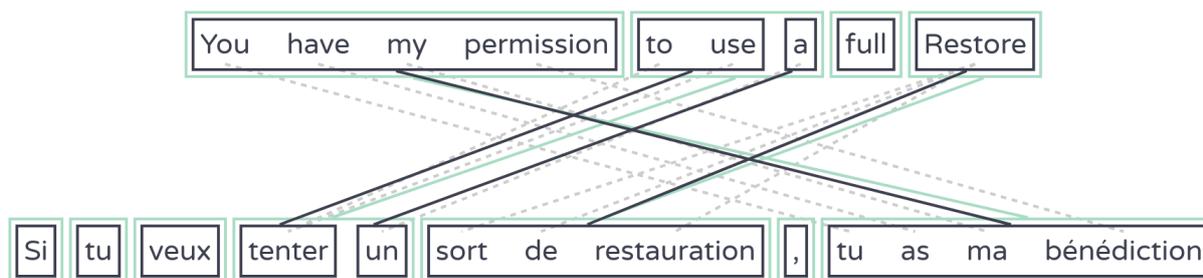


FIGURE 26 – Exemple d’alignement issu de SACr

SOURCE : ASTRÉD Demo, Apache 2.0

La deuxième mesure part elle aussi de cet alignement pour vérifier si les unités de chaque paire font l’objet d’un changement de catégorie grammaticale (*part-of-speech tags*), que je reprends ici et dans le [Tableau 12](#) sous le terme de *transposition* (cf. [VINAY et DARBELNET 1958](#), p. 50). Enfin, le troisième indicateur d’équivalence syntaxique proposé est la distance entre les deux arbres de dépendance construits pour chaque phrase source et cible (ASTrED, de l’anglais *Aligned Syntactic Tree Edit Distance*). Cette dernière mesure cherche à quantifier les modifications au niveau de la structure profonde de chaque phrase, en comparant l’écart entre leur arbre de dépendance respectif et en calculant de nombre d’opérations nécessaires pour obtenir deux arbres similaires, grâce à l’alignement des phrases ¹¹⁵.

	SACr	Transpositions	ASTrED
Référence humaine	1 728	1 116	3 674
Septimus-étendu	859	985	2 685
Google	503	960	2 259
DeepL	579	964	2 460

TABLEAU 12 – Mesures d’équivalence syntaxique

Comme l’indique le [Tableau 12](#), le nombre de croisements issus de SACr place une fois encore la traduction humaine en tête de liste, suivie ensuite par le système adapté, qui fait montre de plus lourdes réorganisations syntaxiques que les deux systèmes génériques. Et quoique ce soit un peu moins visible si l’on se réfère aux transpositions répertoriées de manière automatique par l’outil, la métrique ASTRÉD confirme elle aussi cette tendance, que les analyses des sections suivantes feront ressortir comme un signe clair de l’adaptation. Globalement, ces métriques de l’équivalence syntaxique laissent clairement apparaître que la TA copie fortement la source d’un point de vue syntaxique, comme en témoignent également ([TEZCAN et al. 2019](#), p. 47), et viennent à

115. Pour d’autres exemples et davantage de détails sur chacune de ces mesures, voir [VANROY et al. \(2021\)](#).

nouveau appuyer l'idée que le recours à des jeux de données spécifiquement sélectionnés pour leur pertinence et leur qualité peut contribuer de façon non négligeable à l'amélioration de la TA, y compris dans le domaine littéraire. Ajoutées aux diverses mesures déjà présentées, elles confirment par ailleurs que la traduction humaine se place toujours loin devant la TA et qu'il reste encore à cette dernière un long chemin à parcourir avant d'atteindre la « parité ». Cette démarche illustre donc bien la nécessité de prendre en compte ce type de phénomènes textuels spécifiques lors de l'évaluation de la TA. Pour cette même raison, je me tourne à présent vers une analyse humaine détaillée et attentive des traductions produites par le système adapté.

6.2 Évaluation humaine

Ce chapitre a démarré sur le constat que le champ de la recherche en traduction automatique était largement dominé par les métriques. Bien que j'en aie fait intervenir quatre types pour corroborer le résultat positif de cette expérience d'adaptation, ces différentes mesures se limitent à l'estimation et à la comparaison. L'évaluation humaine de la TA, contrairement aux évaluations automatiques, offre à ce titre une analyse plus fine et plus fiable des sorties de traductions. Toutefois, ce type d'analyse recèle ses propres difficultés, à commencer par le caractère long, fastidieux, coûteux et subjectif du processus (REITER 2018, p. 400 ; FREITAG *et al.* 2021, p. 1460). Une autre difficulté à laquelle j'ai maintes fois été confronté au cours de mes lectures est que l'on ne sait pas toujours qui mène l'évaluation, quel est le niveau de compétence de ces personnes ou comment est menée l'évaluation. Bien souvent, celle-ci est confiée à des cohortes étudiantes ou à des personnes anonymes en ligne (*crowdworkers, turkers...*), qui peuvent ne pas être familiarisées avec la traduction, le domaine de spécialité ou la technologie, voire avec chacun des trois aspects. La tâche demandée, en outre, consiste la plupart du temps en une simple évaluation phrase par phrase qui nécessite d'assigner un simple score aux traductions (*direct assessment*)¹¹⁶ ou de comparer deux systèmes (*ranking*)¹¹⁷, ne donnant dès lors aucune information sur la sévérité des erreurs, par exemple, ou sur la qualité générale des traductions pour un cas d'utilisation particulier (FREITAG *et al.* 2021, p. 1460). Ce sont d'ailleurs ces écueils qui ont récemment pu mener aux conclusions exagérées de parité entre traduction automatique et traduction humaine (TORAL *et al.* 2018a ; LÄUBLI *et al.* 2018 ; FREITAG *et al.* 2021, p. 1461–1462). Finalement, et de la même manière que pour les métriques, la fiabilité des méthodes non automatiques peut être très variable et dépendra autant des systèmes de TA évalués, que des corpus utilisés et du protocole d'évaluation plus général (REITER 2018, p. 398).

116. Cf. GRAHAM *et al.* (2013).

117. Cf. VILAR *et al.* (2007).

Pour le champ restreint de la traduction littéraire automatique, la tâche proposée est souvent celle de la post-édition, pour laquelle s'appliquent ces mêmes contraintes, avec la difficulté supplémentaire qu'il existe probablement très peu de personnes véritablement expertes en PE de traduction littéraire. Les résultats, en outre, sont souvent très variables d'un cas à l'autre (KOLB 2024, p. 65) et peuvent ne pas correspondre avec les impressions des personnes qui post-éditent (KUZMAN *et al.* 2019, p. 6). Une autre méthode communément utilisée est celle de l'annotation d'erreurs, que j'adopterai et décrirai ici. L'avantage majeur de ce type d'évaluation est qu'elle donne un aperçu plus concret de ce qui est produit en réalité par la TA, de même qu'une vue détaillée sur les limites qui entravent encore ce type de système, ce qui peut s'avérer utile à la fois pour les personnes qui seraient amenées à se pencher sur leur développement ou à interagir avec l'outil. Plus encore, les multiples lectures et comparaisons minutieuses demandées par le processus d'annotation fournissent un point de départ idéal pour des observations plus générales et pour d'autres recherches pouvant, à l'inverse, faire ressortir des points forts du système.

6.2.1 Annotation d'erreurs : une typologie sur mesure

Avant de se lancer dans l'annotation d'erreurs, il est tout d'abord impératif de choisir une typologie qui guidera le processus d'évaluation. Étant donné que peu d'études ont adopté cette méthode pour le cas littéraire, il m'a semblé opportun de mettre au point une typologie sur mesure qui reflète des catégories communément utilisées pour cette tâche, ainsi que d'autres faisant ressortir plus spécifiquement des aspects littéraires qui me semblent présenter un enjeu particulier pour l'ouvrage considéré ici ou la littérature en général et que l'on a tendance à retrouver sous une même catégorie générique liée au « style ». La grille d'évaluation établie fait donc logiquement intervenir les notions traditionnelles d'« adéquation » et de « fluidité », à savoir les deux critères les plus utilisés dans l'évaluation manuelle des sorties de TA (CALLISON-BURCH *et al.* 2007, p. 140). Le premier critère d'adéquation s'évalue au regard de la source et concerne fondamentalement « le degré auquel les informations présentes dans l'original sont bien transmises par la traduction », tandis que la fluidité ne se préoccupe quant à elle que du texte cible et vise avant tout à vérifier « le degré de bonne formation linguistique de l'énoncé » (BLANCHON et BOITET 2007, p. 43). Comme je viens de l'énoncer, je voulais également y ajouter des éléments plus spécifiques à la littérature, que l'on peut difficilement intégrer sous ces deux autres critères et qui me semblent particulièrement pertinents pour ce cas d'étude de la série *Magyk*. Ceux-ci ont dès lors été ajoutés à une troisième classe, axée sur la « littérarité » du texte produit, pour laquelle la comparaison avec la référence était souvent nécessaire. La répartition de la typologie dans ces trois volets est présentée dans la Figure 27.

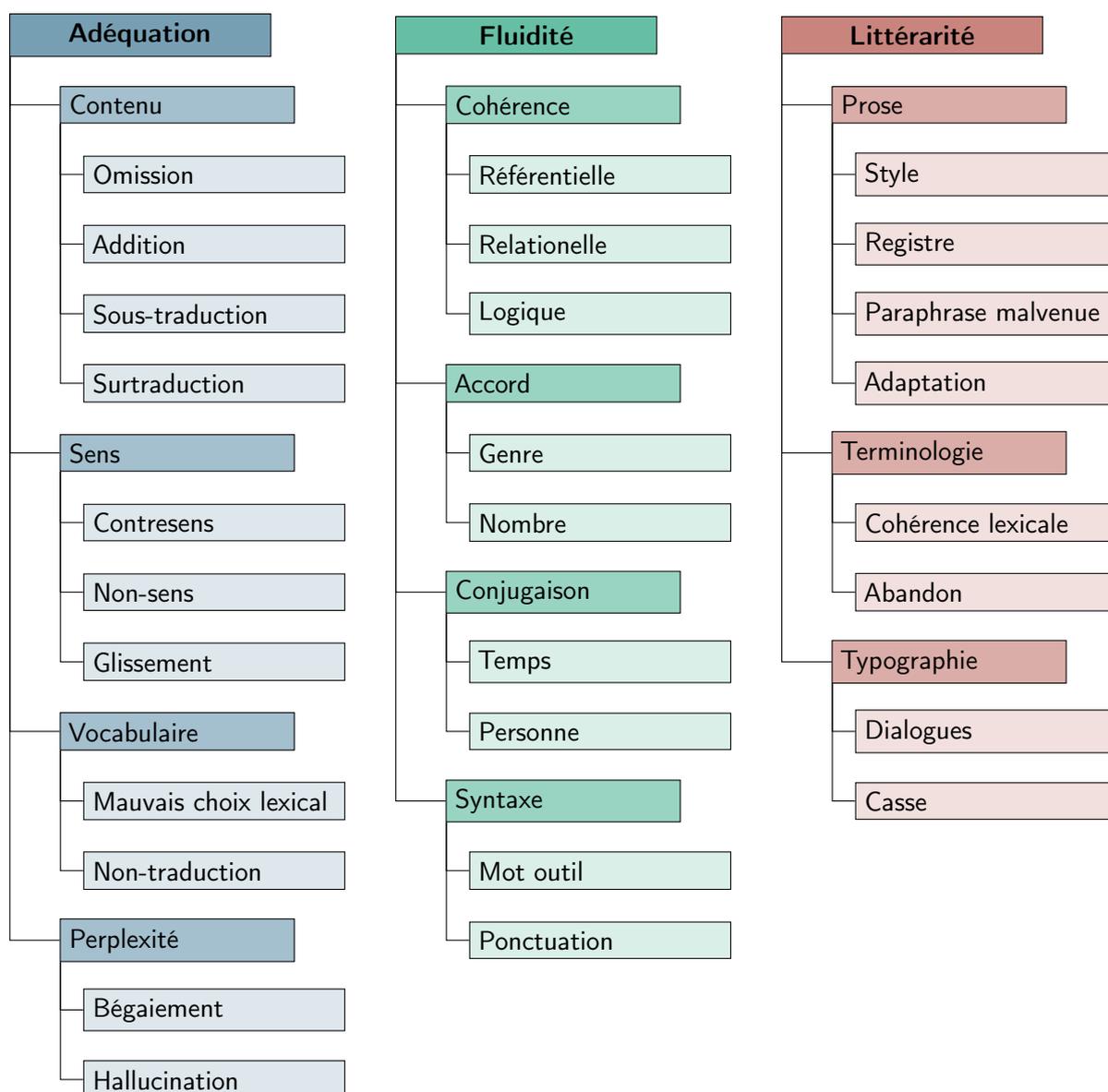


FIGURE 27 – Typologie d’erreurs

Pour mettre au point cette typologie, je me suis inspiré des taxonomies utilisées par VILAR *et al.* (2006, p. 699), MATUSOV (2019, p. 14–15), TEZCAN *et al.* (2019, p. 42) et SCHUMACHER (2023, p. 104–105). Dans chacune de ces études, j’ai ainsi cherché à voir quelles étaient les catégories qui se recoupaient et qui contenaient le plus d’erreurs, pour y ajouter ensuite des points d’intérêt survenus durant les nombreuses discussions qui ont ponctué ce travail et dans mon analyse préalable de la série (cf. HANSEN 2017). Pour chacune des trois classes principales, j’ai finalement sélectionné celles qui me semblaient les plus pertinentes, en cherchant en outre à remplir deux objectifs. Le premier était de parvenir à réduire au maximum le nombre d’erreurs possibles pour

faciliter l'annotation et réduire les ambiguïtés, tout en étant capable de classer précisément chaque erreur sans recourir à des catégories génériques de type « autre ». La seconde préoccupation était d'orienter la classification des erreurs depuis le point de vue des traductaires. Une modification qui répond à ces deux enjeux se retrouve de fait dans le recours à une même étiquette pour d'éventuelles distinctions telles que « mauvais choix lexical » et « mauvaise désambiguïsation », qui permettaient du point de vue du système de différencier des mots inconnus mal traduits (2) de mots appris et restitués avec un sens ne convenant pas dans ce contexte particulier (3)¹¹⁸ :

(2) SRC : *Marcia's sent me to fetch you for your Induction.*

HYP : Marcia m'envoie te chercher pour ton instruction.

(3) SRC : *The ghost of Olaf wafted down into the cherry-wood cabin [...]*

HYP : Le fantôme d'Olaf se laissa flotter jusqu'à la cabine en bois de cerise [...]

Dans ces deux cas, la personne chargée de post-éditer le texte ne tirerait en effet aucun bénéfice de cette information et devrait remplacer ce « mauvais choix lexical » quoi qu'il arrive. Cette décision présente également l'avantage recherché de limiter le nombre de catégories et d'éliminer les risques de confusion sur certains cas ambigus. Dans ce même sens, j'ai pareillement fait l'économie de catégories qui étaient intéressantes du point de vue de la recherche, mais qui venaient compliquer l'annotation (notamment une classe liée à la « concordance des temps », reprise dans les « temps », de même que les « néologismes », repris dans les problèmes d'« adaptation » et d'« abandon » qui couvrent respectivement le recours à des solutions créatives pour des raisons formelles ou culturelles ainsi que la traduction de termes propres à l'univers fictionnel). De cette manière, une dizaine de catégories qui avaient initialement été envisagées ont disparu (telles que les « calques grammaticaux » ou les « marqueurs discursifs »), ont été fusionnées (créant par exemple les « mots outils ») ou ont été supprimées parce qu'elles ne donnaient lieu à aucune erreur (comme les erreurs d'« orthographe »), tandis que d'autres ont inversement été ajoutées (« bégaiement » et « hallucinations ») ou encore explicitées (« accord » et « conjugaison ») pour accommoder des cas particuliers ou donner plus de détails sur le type d'erreur produite.

118. Pour faciliter la lecture des exemples qui suivent, une étiquette indique si le segment provient du texte source (SRC) ou de l'hypothèse de TA (HYP), autrement dit de la sortie produite par le système. Tous les exemples sont tirés du corpus de test issu de *Darke*, le sixième tome de la série.

La typologie résultante pourrait être dépliée plus finement encore si cela était nécessaire, par exemple pour préciser le type d'adaptation (jeu de mots, référence culturelle, contrainte formelle, occasionalisme, métaphore...), mais je suis resté fidèle à mon impératif de faciliter au maximum l'annotation en me limitant à ces deux niveaux de granularité. De même, un autre aspect qui m'intéressait beaucoup dans la classification proposée par MATUSOV (2019) était la distinction entre les erreurs graves et les erreurs mineures. Toutefois, pour éviter de compliquer davantage le travail d'annotation et la lecture des résultats, cette distinction sera principalement abordée dans la discussion des différentes catégories et des résultats.

Dans l'ensemble, ces différents types d'erreurs sont relativement classiques pour la traduction et se rapprochent de ceux que l'on pourrait trouver dans des manuels d'initiation à la traduction et de linguistique contrastive (cf. VINAY et DARBELNET 1958; DELISLE 2007), mais aussi dans d'autres typologies telles que MQM (LOMMEL *et al.* 2014). Sauf peut-être dans les cas plus techniques du « bégaiement » et de l'« hallucination », qui me semblaient nécessaires dans la mesure où ces problèmes sont difficilement annotables avec d'autres étiquettes. Tous deux sont liés à la présence d'une phrase ou d'un segment qui s'écarte trop fortement des données connues et auquel le système de TA assigne un degré de perplexité élevé, résultant soit en la répétition d'un même mot jusqu'en fin de phrase (4), soit en l'invention de tout un segment sans aucun lien avec le texte source (5) :

(4) SRC : *The winter morning drew on. Snow piled up outside the long windows, while the Ballroom was transformed with holly and ivy, red ribbons, huge silver candlesticks and a whole box of [...]*

HYP : Le matin d'hiver s'écoula, tandis que la neige s'entassait devant les fenêtres tandis que la salle de bal se transformait en rubans rouges, en chandeliers d'argent, en chandeliers d'argent [...]

(5) SRC : *Beetle stared at Foxy, pan scourer dripping in his hand.*

HYP : Moustique regarda Vulpin, l'homme-chat dégoulinant.

Dans les deux cas, la solution la plus simple est de tout supprimer pour retraduire à partir de zéro plutôt que de corriger l'erreur, surtout si la phrase entière n'a aucun sens, ce qui justifie de leur réserver une catégorie à part. Pour le reste de la typologie, et notamment pour les aspects littéraires, des explications détaillées et accompagnées d'exemples sont compilées dans l'Annexe G.

Les deux personnes en charge de l'annotation, à savoir l'auteur du présent travail et Emmanuelle Esperança-Rodier, ont toutes deux des années d'expérience professionnelle en traduction et une expérience modérée à élevée dans l'annotation d'erreurs. Le processus d'annotation en tant que tel a été mené sur ACCOLÉ (ESPERANÇA-RODIER et BRUNET-MANQUAT 2021), la plateforme d'annotation collaborative du LIG (cf. Figure 28), toujours à partir du modèle *Septimus-étendu*¹¹⁹. Durant la tâche, la fluidité était tout d'abord évaluée à partir du segment cible, après quoi une comparaison avec le segment source permettait d'en juger l'adéquation. Un segment issu du corpus de référence était également mis à disposition pour confirmer certains points tels que la cohérence lexicale ou l'utilisation d'un registre correct, uniquement après annotation des deux premières catégories d'erreurs, offrant par ailleurs une seconde lecture visant à s'assurer qu'une erreur n'ait pas été annotée injustement ou oubliée. De même, un accès aux tomes précédents était prévu pour vérifier la traduction de termes spécifiques à la série ou l'utilisation de certaines paraphrases, par exemple. En ce sens, la connaissance de l'œuvre et des textes ainsi que l'expertise pratique de l'analyse littéraire se sont révélées essentielles à la prise en compte — indispensable elle aussi — de la poétique dans cette démarche, autant pour l'élaboration de la taxonomie que pour le processus d'annotation.

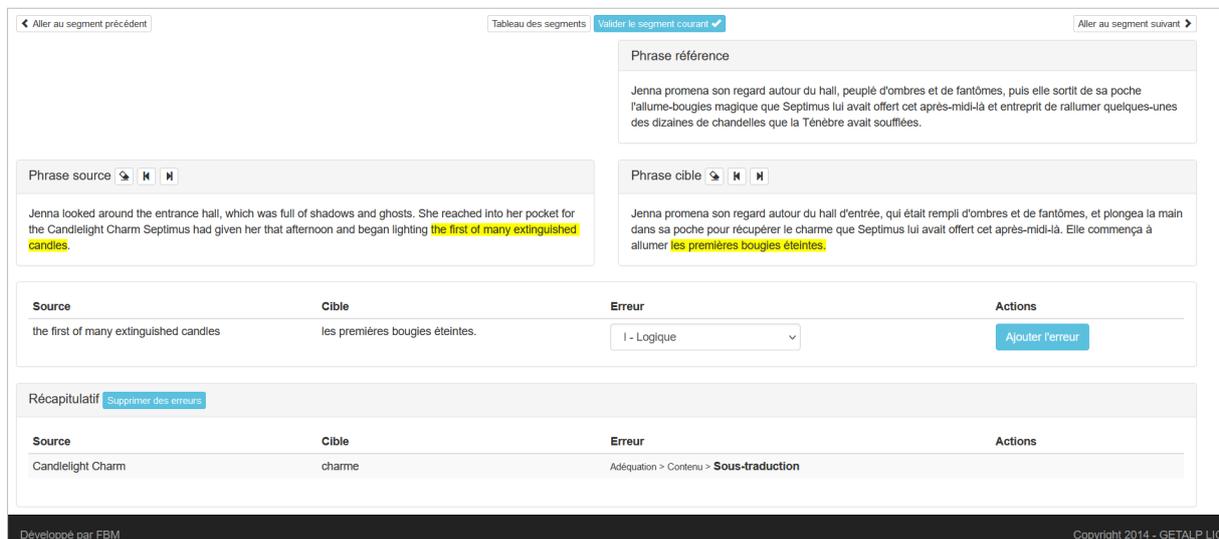


FIGURE 28 – Interface de travail dans ACCOLÉ

119. Les résultats de l'annotation présentée ici ne sont pas à proprement parler ceux du modèle listé dans le Tableau 7, mais du système présenté dans HANSEN et ESPERANÇA-RODIER (2023). Ces deux systèmes sont tout à fait comparables, cependant, ce qui rendait difficilement envisageable la répétition de l'annotation, qui est une tâche extrêmement chronophage.

Concrètement, une première passe a été effectuée sur le premier des trois chapitres évalués (environ 30 % du corpus de test) pour mettre une première fois la typologie à l'épreuve, affiner la classification et écarter de possibles ambiguïtés. Cela fait, 20 % des segments ont été ré-annotés avec la typologie finale pour évaluer sa pertinence et la fiabilité de l'annotation par le biais de l'accord inter-annotation. L'accord observé après consolidation est de 80,85 % pour l'adéquation, de 84 % pour la fluidité et de 70,49 % pour les aspects littéraires, révélant une plus haute part de subjectivité pour cette dernière catégorie (et pour le « style » en particulier) malgré un accord généralement fort élevé¹²⁰. Après une nouvelle mise en commun et un dernier réajustement de certaines classifications, j'ai ensuite poursuivi l'annotation sur le reste des 704 segments composant les trois chapitres de test.

6.2.2 *Peu d'erreurs littéraires, mais une fluidité trompeuse*

La méthode d'annotation était volontairement agressive, dans la mesure où les erreurs qui se répétaient ont été systématiquement annotées et qu'une même erreur pouvait recevoir plusieurs étiquettes, comme le montre l'exemple suivant.

(6) SRC : *during a Do-or-Die exercise in the Young Army*

HYP : lors d'un exercice de désœuvrement de la Jeune Garde

Dans ce cas, un terme de l'univers fictionnel (irrealia) qui aurait nécessité une solution originale a été rendu par un terme neutre, résultant en un abandon (ABD). Ce concept fictif n'a pas uniquement été neutralisé, mais a aussi été mal traduit, et ce mauvais choix lexical (MCL) a donné lieu par ailleurs à un non-sens (NOS). De cette manière, il n'est pas rare qu'un même problème soit étiqueté dans deux ou trois catégories différentes.

120. L'accord inter-annotation est traditionnellement confirmé par le biais de la méthode du κ (kappa), qui convient surtout pour les mesures ordinales (p. ex. échelle de Likert) et les observations menées sur un nombre prédéterminé de cas à évaluer et à répartir dans un nombre restreint de catégories. Cependant, étant donné la liberté accordée dans ce cas pour l'évaluation (annotation non contrôlée, possibilité de relever des erreurs que l'autre personne ne considérera pas comme problématiques), ainsi que la complexité de la tâche (plusieurs annotations possibles pour une même erreur, parmi un grand nombre de catégories), il me semble peu probable que l'accord élevé observé soit dû à la chance (ce que vérifie la méthode du κ) et il me paraît dès lors qu'il peut être considéré comme fiable.

Adéquation										
Contenu				Sens			Vocabulaire		Perplexité	
OMI	ADD	SOU	SUR	CTS	NOS	GLS	NOT	MCL	BEG	HAL
270	6	38	8	11	26	43	0	200	3	12

Fluidité								
Cohérence			Accord		Conjugaison		Syntaxe	
REF	REL	LOG	GEN	NBR	TPS	PER	MOT	PCT
31	3	60	19	6	51	7	54	20

Littérarité							
Prose				Terminologie		Typographie	
STY	REG	PAR	ADP	COH	ABD	DLG	CAS
218	38	8	15	29	10	44	20

TABLEAU 13 – Relevé des erreurs produites par la TA pour l’adéquation, la fluidité et la littérarité

Dans l’ensemble, il a été possible de dénombrer 1 250 erreurs, réparties sur 704 segments comprenant au total 799 phrases (en français, le nombre étant plus élevé en anglais). Sur ces 704 segments, on peut en distinguer 145 (environ 20 %) qui ne comportent aucune erreur, ce qui est bien inférieur aux 44 % rapportés par (FONTEYNE *et al.* 2020, p. 3790) pour le néerlandais, mais plus proche des 20 % et 28 % indiqués par MATUSOV (2019, p. 15–16) pour les paires anglais-russe et allemand-anglais. Une observation de la répartition des annotations par catégories principales apporte une lecture plus éclairante encore et une première idée concrète des forces et faiblesses de l’outil mis au point. On peut en effet compter 617 erreurs d’adéquation, contre 251 liées à la fluidité et 382 pour les aspects littéraires.

D’une part, il est surprenant d’y trouver moins de problèmes ayant proprement trait au caractère littéraire du texte que ce à quoi l’on aurait pu s’attendre, en particulier pour un texte qui présente des difficultés singulières pour la machine (cf. Section 5.2.3), lesquelles se matérialisent notamment par les enjeux esthétiques de la traduction de jeunesse et par les contraintes formelles de la fiction, mais qui doit également tenir compte des choix opérés dans les volumes précédents. D’autre part, ce résultat corrobore ceux présentés par MATUSOV (*Ibid.*, p. 15–18) et confirme une fois encore la conclusion selon laquelle la TA neuronale, malgré son apparente fluidité, produit plus d’erreurs liées au sens et au contenu, là où la TA statistique produit des phrases

moins naturelles mais présentant aussi moins de problèmes d'adéquation (ESPERANÇA-RODIER et BECKER 2018, p. 12; LOOCK 2019, p. 59; MÜLLER *et al.* 2020, p. 156)¹²¹. Or, il a déjà été montré que ce type de faute est plus difficile à repérer et à corriger, ce qui se note par exemple dans le fait que la post-édition n'est pas nécessairement plus rapide qu'avec l'ancienne génération de systèmes statistiques (CASTILHO *et al.* 2017, p. 126). Ceci rend la TAN d'autant plus trompeuse pour les personnes qui auraient moins d'expérience en traduction (YAMADA 2019).

Les difficultés les plus importantes sont de loin les omissions (OMI), les mauvais choix lexicaux (MCL) et le style (STY), de même que les problèmes de logique (LOG). Ce scénario est entièrement cohérent avec d'autres études pour la traduction littéraire (MATUSOV 2019, p. 16–17; FONTEYNE *et al.* 2020, p. 3794–3795; LONG 2022, p. 731), en particulier pour les mauvais choix lexicaux (PETRAK *et al.* 2022, p. 145), qui restent un enjeu majeur de la TAN, puisqu'ils représentent 16 % du nombre total d'erreurs¹²² :

(7) SRC : *Foxy rushed over the tiny sink and grabbed the pan out of Beetle's hands.*

DPL : Foxy s'est précipité vers le petit évier et a pris la casserole des mains de Beetle.

HYP : Vulpin se précipita vers l'évier et agrippa la casserole des mains de Moustique.

REF : Vulpin se rua vers l'évier et arracha la casserole des mains de Moustique.

De la même manière, le style apparaît fréquemment dans ces mêmes études et de manière plus générale comme l'une des limites évidentes de la TA, que ce soit sur des textes littéraires ou autres, ce qui n'est guère surprenant puisque nous avons vu que celle-ci a tendance à copier fortement le texte source (cf. [Tableau 12](#)), mais qui posera évidemment problème dans le domaine qui nous occupe. Plus concrètement, on retrouvera dans cette catégorie divers problèmes de répétitions, de lourdeurs, de tournures et cooccurrences peu idiomatiques et de traductions littérales. Ainsi, même lorsqu'une traduction est tout à fait correcte, sa lecture peut parfois laisser une impression générale d'insatisfaction sans que celle-ci soit nécessairement attribuable à un ou plusieurs termes en particulier :

121. On peut établir un rapprochement avec le biais des métriques neuronales en faveur de la fluidité, qui paraît dès lors moins surprenant puisque ces deux types d'outils font usage des mêmes méthodes (cf. p. 114).

122. En plus de la source (SRC) et de l'hypothèse de TA générée par le modèle *Septimus-étendu* (HYP), je fournirai ici pour chaque exemple la traduction de référence de Nathalie Serval (REF), qui facilitera la lecture et la comparaison des résultats, de même qu'une traduction proposée par *DeepL* (DPL), qui permet parfois de mieux faire ressortir les différences avec un système générique. On pourra remarquer, par exemple, que bon nombre de termes et de tournures en dehors de l'univers fictionnel sont très proches de ceux proposés *DeepL*, laissant penser que les données d'entraînement des deux outils doivent être similaires, si l'on omet les corpus littéraires utilisés pour l'adaptation. Les traductions proposées par *DeepL* pour cette section ont été générées le 1/7/2022.

(8) SRC : *I believe there may be a reason for this. Over the years the Hermetic Chamber has seen much Darke Magyk, and I suspect the Pick has become corrupted.*

DPL : Je crois qu'il y a peut-être une raison à cela. Au fil des ans, la Chambre Hermétique a vu beaucoup de Darke Magyk, et je soupçonne que le Pick a été corrompu.

HYP : Je crois qu'il y a une raison à cela. Au fil des ans, le cabinet hermétique a vu beaucoup de Magyk noire, et je pense que la désignation a été corrompue.

REF : Le cabinet hermétique a maintes fois été exposé à la Ténèbre, et je crains que celle-ci n'ait fini par affecter la désignation.

Les problèmes de logique sont pareillement mis en avant, désignant quant à eux des ruptures de l'enchaînement logique des idées ou des problèmes de syntaxe qui peuvent changer le sens de la phrase :

(9) SRC : *The discs seemed to have fused into one and it was clear that the lines that radiated out from the Manuscriptorium Code joined up with some of those on the Wizard Tower Code. Each one of these led to a symbol.*

DPL : Les disques semblaient avoir fusionné en un seul et il était clair que les lignes qui rayonnaient du code du manuscritoire rejoignaient certaines de celles du code de la tour du sorcier. Chacune d'entre elles menait à un symbole.

HYP : Les disques semblaient avoir fusionné et il était évident que les lignes rayonnantes du code du Manuscriptorium rejoignaient celles du code de la tour du Magicien, chacune donnant naissance à un symbole.

REF : Les disques semblaient avoir fusionné, et les lignes qui rayonnaient à partir du centre du plus petit se prolongeaient dans certaines de celles qui étaient gravées à la surface du second pour aboutir à un symbole.

Enfin, la question des omissions est régulièrement soulevée, en particulier depuis l'arrivée de la TAN (TORAL et WAY 2018, p. 280), mais le nombre impressionnant d'occurrences dans ce cas précis (21,6 % des annotations) reste surprenant, comme j'ai déjà pu l'évoquer plus haut (cf. Section 6.1.4). Sur ce point, il est intéressant de noter que toutes les omissions ne sont pas des erreurs à proprement parler et correspondent à un choix similaire de la part de la traductrice. De fait, 61 des 270 cas relevés sont partagés par le système adapté et par la traductrice, tandis que 12 seulement constituent une erreur grave menant à un problème de compréhension.

- (10) SRC : *Septimus read out the second number of the first pair. “Ninety-eight.”*
 DPL : Septimus a lu le deuxième chiffre de la première paire. “Quatre-vingt-dix-huit.”
 HYP : — Quatre-vingt-dix-huit.
 REF : — Quatre-vingt-dix-huit.

Au-delà des « omissions stylistiques » inévitables et nécessaires, ce choix peut découler en effet de stratégies délibérées et désirables (DIMITRIU 2004). Dans le cas de la traductrice, cette stratégie s’explique par l’économie d’informations spécifiques (telles que des répétitions, des adjectifs dans certains longs syntagmes nominaux, des indications sur les tours de parole, des réponses courtes ayant valeur d’insistance...) et parfois même de propositions entières, qui lui permettent de produire un texte ayant approximativement le même nombre de pages que l’original. Il semblerait d’ailleurs que le système personnalisé ait enregistré cette tendance et soit capable de la reproduire, de manière presque exagérée et quelquefois problématique, expliquant de cette façon les résultats des Tableaux 11 et 12.

- (11) SRC : *A simple Reveal—why hadn’t she thought of that?*
 DPL : Une simple révélation - pourquoi n’y avait-elle pas pensé ?
 HYP : Pourquoi n’y avait-elle pas songé ?
 REF : Un sort révélateur... Comment n’y avait-elle pas pensé elle-même ?

Outre les omissions, le système montre d’autres problèmes d’adéquation tels que des cas de sous-traduction (SOU), illustrée dans l’Exemple (12), ainsi qu’un certain nombre de difficultés en ce qui concerne le sens. Parmi ces dernières, on retrouve principalement des glissements (GLS) et des non-sens (NOS), le plus souvent liés à des problèmes de logique ou de choix lexicaux, mais aussi quelques contresens (CTS), plus graves comme on peut le voir dans l’Exemple (13).

- (12) SRC : *The previous afternoon Marcia had successfully gathered all scribes together in the Manuscriptorium.*
 DPL : L’après-midi précédent, Marcia avait réussi à rassembler tous les scribes dans le manuscritoire.
 HYP : La veille, Marcia avait réuni tous les scribes du Manuscriptorium.
 REF : La veille, Marcia était parvenue à réunir tous les scribes au Manuscriptorium.

- (13) SRC : *Romilly walked selfconsciously into the seven-cornered passage.*
 DPL : Romilly a marché d'un air gêné dans le passage à sept angles.
 HYP : Romilly pénétra inconsciemment dans le couloir qui tournait sept fois sur lui-même.
 REF : Romilly s'avança d'un pas timide vers l'entrée du couloir aux sept détours.

Par comparaison avec les sorties du modèle *Septimus-seul*, qui est entraîné sur moins de données et dont j'ai pu également analyser les traductions, il peut être opportun de mentionner que le modèle *Septimus-étendu* est finalement peu sujet aux phénomènes de bégaiement (BEG) et d'hallucination (HAL), et qu'il ne révèle aucun cas de non-traduction (NOT), bien que ces marques d'adéquation et de fluidité accrues aient été obtenues au prix des aspects littéraire et d'une meilleure adaptation au style de Nathalie, comme je le remarquais dans la section consacrée aux métriques (cf. p 112).

Pour ce qui est de la fluidité, la majeure partie des erreurs en dehors de la logique ont trait aux temps (TPS), dont la catégorie est plutôt explicite, et aux mots-outils (MOT), qui recouvrent pour leur part toute erreur de traduction liée à un déterminant, une préposition, etc. :

- (14) SRC : *He thought of Larry. Of Matt, Marcus and Igor at Gothyk Grotto, even the oddly irritating people at Wizard Sandwiches.*
 DPL : Il a pensé à Larry. À Matt, Marcus et Igor à la Grotte de Gothyk, et même aux gens bizarrement irritants de Wizard Sandwiches.
 HYP : Il songea à Larry, à Matt, Marcus et Igor, à la Grotte-Gothic, même aux gens bizarrement agaçants de Magyk Sandwich.
 REF : Puis il pensa à Larry, à Matt, Marcus et Igor, de la Grotte-Gothic, et même à la bande d'hurluberlus parfois exaspérants de Magyk Sandwich.

De manière générale, je m'attendais à trouver plus d'erreurs pour la fluidité, notamment en ce qui concerne la cohérence, les accords et la conjugaison. De fait, la traduction des temps du passé est régulièrement mise en évidence et soulevée avec raison comme une difficulté pour la traduction de textes littéraires, d'autant plus marquée dans l'évaluation d'outils génériques (PONCHARAL 2021, par. 11–16). Or, lorsque le système est entraîné avec des données littéraires, on s'aperçoit que le problème se pose beaucoup moins. Le processus d'adaptation est donc d'une

grande aide à cet égard, même s'il reste des ambiguïtés entre l'imparfait et le passé simple, que le système semble favoriser :

(15) SRC : *Simon pushed his way through the Darke Fog, terrified that at any moment a Thing would recognize him.*

DPL : Simon s'est frayé un chemin à travers le brouillard de Darke, terrifié à l'idée qu'à tout moment une Chose puisse le reconnaître.

HYP : Simon se fraya un chemin à travers le brouillard, terrifié à l'idée qu'une créature puisse le reconnaître.

REF : Simon se frayait un chemin à travers le brouillard, terrifié à l'idée qu'une créature le reconnaisse.

Il en va de même pour la cohésion référentielle (REF), dont Rob VOIGT et Dan JURAFSKY (2012) évoquent le caractère essentiel pour les textes littéraires. Pour tous ces cas et ceux qui relèvent de la fluidité de manière générale, il serait intéressant de comparer plus en détail ces sorties avec celles produites par un outil générique, mais aussi avec un système de TA qui serait capable de prendre en charge le contexte (LUPO *et al.* 2022a ; b).

Pour terminer sur les éléments spécifiques au domaine littéraire, le style, à nouveau, est très certainement la principale difficulté de la traduction automatique dans le domaine, malgré l'adaptation du système, et représente 17,4 % des erreurs annotées. Néanmoins, de la même manière que pour la fluidité, on aurait pu prévoir un nombre d'erreurs plus élevé pour ces catégories, et plus spécifiquement pour le registre (REG) et l'agencement (essentiellement typographique) des dialogues (DLG), qui apparaissent aussi tous deux comme des limites des outils de TA génériques appliqués à la littérature. Évidemment, les ambiguïtés subsistent, par exemple pour le ton associé à un personnage, quoique ce problème soit étonnamment peu fréquent avec le système adapté. Celles-ci portent surtout sur la traduction formelle ou informelle du *you* anglais, dont la restitution est rendue d'autant plus compliquée que le ton varie pour un même personnage selon les relations qu'il entretient avec d'autres :

(16) SRC : *“Marcellus, do you have six guineas on you?” asked Marcia.*

DPL : “Marcellus, tu as six guinées sur toi ?” demande Marcia.

HYP : — Marcellus, tu as six guinées sur toi ?

REF : — Marcellus, auriez-vous six guinées sur vous ? demanda Marcia.

Les dialogues, eux aussi, présentent des difficultés, dans la mesure où les marques typographiques et l'agencement des tours de parole peuvent varier sensiblement d'une langue à l'autre, d'autant qu'elle dépend en outre des conventions propres à l'édition (dans ce cas, un tiret cadratin pour marquer une prise de parole, deux-points dans le segment précédent s'il introduit le dialogue et des guillemets uniquement dans certains discours rapportés). Or, ces conventions sont pourtant bien observées par le système, y compris lorsque le segment anglais ne comporte aucune indication du dialogue qui va suivre ¹²³ :

- (17) SRC : (i) *She continued.*
 (ii) *“During the last few days in the Wizard Tower, when I was trying to find a way to defeat the Darke Domaine [...]*
 DPL : (i) Elle a continué.
 (ii) “Pendant les derniers jours dans la Tour des Sorciers, quand j’essayais de trouver un moyen de vaincre le Domaine de Darke [...]
 HYP : (i) Elle poursuivit :
 (ii) — Ces derniers jours, à la tour du Magicien, alors que je cherchais un moyen de vaincre le domaine ténébreux [...]
 REF : (i) Marcia reprit :
 (ii) — Ces derniers jours, tandis que je cherchais comment vaincre le domaine ténébreux, enfermée dans la tour [...]

Bien sûr, cette prise en charge par le système n'est pas infaillible et les dialogues restent parfois source de problèmes, comme l'attestent certains tirets manquants :

- (18) SRC : *“Does that mean that Beetle is...”*
 DPL : *“Cela signifie-t-il que Beetle est...”*
 HYP : *Ça veut dire que Moustique est...*
 REF : *— Ça veut dire que Moustique...*

Un dernier point évoqué dans la littérature et pris en compte dans la typologie repose sur des soucis de cohérence terminologique (cf. [MATUSOV 2019](#), p. 15). Une fois encore, l'adaptation révèle à cet égard une amélioration surprenante, même en tenant compte de la fidélité aux choix opérés dans les tomes antérieurs, mais on peut constater que la cohérence lexicale (COH)

123. Rappelons que les segments (i) et (ii) sont traités de manière isolée par la machine et qu'il ne s'agit, ici aussi, que d'une question de statistiques, à savoir la probabilité que *elle* soit suivi de *poursuivit* et ensuite de deux-points plutôt que d'un point final.

demeure problématique lorsqu'un mauvais choix est propagé sur l'ensemble de la traduction (5 termes représentent à eux seuls 23 des 29 erreurs de cohérence lexicale) et que la terminologie peut l'être également lorsqu'elle fait référence à des concepts fictifs qui relèvent d'un champ complètement nouveau, menant à ce que j'ai appelé des abandons (ABD). Ces deux points, cependant, constituent des exemples typiques pour lesquels la traduction assistée par ordinateur et les mémoires de traduction peuvent s'avérer particulièrement utiles, de même que pour la traduction des titres ou d'autres éléments nécessitant une adaptation (ADP), que mentionne pareillement Ó MURCHÚ (2019, p. 23) parmi les écueils de la TA, mais aussi pour le choix du registre et pour les retouches de style, puisque les mémoires facilitent la variation et la cohérence dans les choix de traduction (HANSEN 2017, p. 70–80). Cette brève remarque laisse dès lors entrevoir une complémentarité des deux outils sur ces aspects, la TA offrant une première solution pour la phrase entière et la TAO favorisant des recherches plus poussées sur un point précis.

6.2.3 Observations générales

Qu'il s'agisse de traduction littéraire assistée par ordinateur ou automatique, la personnalisation des outils se présente comme un changement de paradigme désirable et profitable, mais celui-ci trouve encore peu d'échos en pratique (O'BRIEN et CONLAN 2018). Dans le cas de la TA, les multiples métriques tout comme l'annotation d'erreurs confirment la légitimité de l'adaptation du système, mais cette dernière donne un aperçu déformé des sorties de TA, et ce, pour deux raisons. D'une part, les exemples sont choisis parce qu'ils montrent clairement et — idéalement — uniquement l'erreur concernée, mais d'autres passages offrent des traductions plus problématiques que les fragments fournis ici, comme l'illustre ce segment comportant 13 erreurs qui fait figure d'exception puisqu'il s'agit du segment affichant le plus grand nombre d'annotations¹²⁴ :

(19) SRC : *And then they saw the most beautiful sight they had ever seen—the Darke Fog rolling away. Once more they saw the Castle they loved—its higgledy-piggledy roofs, its turrets and towers, its crenulated battlements and tumbledown walls, all outlined against the pink sky of the dawn of a new day.*

HYP : Puis ils virent le plus beau spectacle qu'ils avaient jamais vu – le brouillard ténébreux s'envoler, ils virent à nouveau le Château qu'ils aimaient – ses toits branlants, ses tourelles et ses tourelles, ses créneaux crénelés et ses murs délabrés, tout dessinait sur le ciel rose de l'aube d'

124. Viennent ensuite deux segments comportant 8 erreurs, la plupart affichant entre 0 et 5 problèmes de traduction.

REF : Ils assistèrent ensuite au spectacle le plus magnifique qu'il leur avait été donné de voir : le brouillard se dissipa, faisant apparaître un fouillis de toits, de tours et de tourelles, de remparts crénelés et de murs délabrés qui se profilaient sur le ciel d'un rose délicat.

D'autre part, l'annotation, comme son nom l'indique, met avant tout l'accent sur les erreurs et donc sur les inconvénients de la traduction automatique, ce que cette dernière section me permettra de nuancer en partie, par le biais d'observations plus générales tirées des lectures attentives des multiples versions de ces traductions, durant les multiples itérations de l'entraînement des systèmes et au cours des diverses évaluations menées.

Comment affirmer, en effet, que le système fait bel et bien preuve d'une adaptation au niveau du style ? Outre la confirmation donnée par chacune des métriques proposées dans ce chapitre, l'indicateur le plus évident se trouve tout d'abord dans le maintien de la cohérence et le respect du champ lexical propre à l'univers fictif, dans la bonne restitution des termes et du vocabulaire adoptés dans les précédents volumes et dans l'accord avec les conventions générales de la série que la section précédente a permis de mettre en évidence. Nous avons vu en outre que de nombreux cas d'omission — et, de manière tout aussi intéressante, de sous-traduction — reflètent une décision similaire de la part de la traductrice sur ces cas particuliers, mais il est possible de distinguer de nombreux autres cas, notamment ceux impliquant une réorganisation syntaxique, qui reproduisent des choix semblables à ceux de la traductrice. On peut remarquer, par exemple, que certains termes ne sont pas rendus systématiquement par une même traduction et peuvent apparaître sous la forme d'un synonyme ou d'une périphrase également employée par la référence (*Wizard Tower courtyard* > *cour de la tour du Magicien, cour du Magicien, cour de la tour...*). En règle générale, les « choix de traduction » du système personnalisé sont ainsi plus proches de ceux effectués par Nathalie, qu'il s'agisse d'items lexicaux récurrents, de syntagmes ou du traitement de la narration (cf. [TAIVALKOSKI-SHILOV 2019b](#)). De même, bon nombre de modifications syntaxiques opérées par le système s'alignent pareillement sur sa traduction. On pourra mentionner à titre d'illustration plusieurs procédés techniques typiques tels que des modulations (p. ex., *as they watched the sun rise* > *tandis que le soleil se levait*), des transpositions (p. ex., *looked in panic* > *regardèrent, paniqués*) ou encore des contractions du texte source (p. ex., *were beginning to fall; they drifted lazily* > *tombaient paresseusement*), dont la traduction correspond exactement à celle de Nathalie pour ces segments, soulignant de nouveau la pertinence et la plus-value du processus d'adaptation.

De façon peut-être plus inattendue, j'ai pu remarquer par ailleurs que le système de TA avait tendance à fusionner des segments là où la traductrice décidait également de le faire. Sur 41 occurrences pour lesquelles deux phrases ou plus ont été réunies en une, on peut en dénombrer 33 où la fusion se retrouve dans la sortie de TA et dans la référence, 6 où elle ne se retrouve que dans la référence et 2 dans la TA uniquement. Un cas simple est illustré dans l'Exemple (20) :

(20) SRC : *Two books—The Undoing of the Darknesse and The Darke Index—lay open on Marcia's desk. In the shadow of the books a small silver box and a tiny silver disc rested on a piece of purple velvet.*

HYP : Deux livres – L'Art de vaincre la Ténèbre et le Codex Tenebrae – étaient ouverts sur le bureau de Marcia, tandis qu'un petit coffret en argent reposait sur un morceau de velours pourpre.

REF : Deux livres – L'Art de vaincre la Ténèbre et le Codex Tenebrae – étaient ouverts sur la table de travail de Marcia, à côté d'une petite boîte en argent et d'un disque du même métal posé sur un carré de velours violette.

Ce résultat est lié sans aucun doute à la façon dont a été compilé — manuellement — le corpus d'entraînement (cf. Section 5.3.1), c'est-à-dire de manière à accommoder les nombreuses fusions de phrase et réorganisations syntaxiques de la traductrice, et à pouvoir être consulté sous la forme d'une mémoire de traduction sans avoir à faire défiler les segments environnants¹²⁵. Toutefois, ces sorties de traduction contredisent de toute évidence l'idée que la TA ne peut fonctionner qu'au niveau de la phrase. Au contraire, elles prouvent que l'on peut obtenir des résultats intéressants si l'on repense les corpus autrement que comme une série de correspondances exactes entre des phrases isolées et si l'on dépasse ce cadre en fournissant plus de matière, mais aussi des données plus spécialisées, au système. Cette constatation est plus appréciable encore si l'on considère que les omissions, les tournures généralement plus économiques et les réagencements syntaxiques (la fusion de segments en particulier) étaient déjà apparus comme des éléments caractéristiques du style de Nathalie au cours de ma recherche sur la traduction littéraire assistée par ordinateur (HANSEN 2017, p. 66), mais aussi et surtout si l'on rappelle que chacun des points cités dans ce paragraphe et dans le précédent apparaissent bien souvent comme une des limites fondamentales de la traduction automatique (AHRENBORG 2017).

125. Dans ce sens, les segments que décidait d'omettre Nathalie étaient systématiquement inclus dans l'unité de traduction concernée au lieu d'être écartés, afin qu'il soit possible de voir que ces segments n'avaient pas été conservés.

Il ressort néanmoins clairement que les traductions produites sont encore fort littérales en dépit de ces signes évidents d'adaptation et des multiples exemples de réorganisation syntaxique. Déjà rapporté par BESACIER (2014, p. 393) et toujours pleinement d'actualité (THAI *et al.* 2022, p. 9885), il s'agit assurément ici du défaut le plus important de la traduction automatique, pour lequel une comparaison entre TA et TH reste nécessaire et éclairante¹²⁶. Ainsi, bien que le système révèle une utilisation répétée de paraphrases et de synonymes divers, toujours est-il que la traductrice joue beaucoup plus avec les noms de personnages ou avec les mentions à des références spécifiques (p. ex. *the Things* > *les serviteurs de Merrin* au lieu de *les créatures*). De la même manière que l'indiquaient les métriques, la traduction humaine fait parallèlement montre d'une plus grande diversité lexicale et de réagencements plus marqués, allant parfois jusqu'à la réécriture pure et simple pour faciliter la compréhension ou pour des raisons purement humoristiques et esthétiques. Malgré les remarques utiles sur la fusion de phrases, on peut aussi observer que les coupures sont bien plus fréquentes dans la référence (18 se retrouvant dans la référence uniquement, 8 dans la TA, 4 dans les deux), comme si le système avait appris à rassembler et non à dissocier les segments.

Poursuivant dans les difficultés, il semble que la langue parlée ainsi que les tournures familières engendrent de problèmes fréquents de désambiguïsation et de logique, une difficulté que partage donc la littérature avec la traduction de jeux vidéos (HANSEN et HOULMONT 2022, p. 264). De même, certains des « choix » (statistiques) opérés par la machine, comme des additions ou même des glissements de sens, peuvent éventuellement apporter des nuances intéressantes vis-à-vis du texte la source, parfois même plus proches de la référence, quoique ces cas restent exceptionnels et constituent des erreurs indésirables dans la plus large majorité des occurrences. J'ai par exemple pu repérer des stratégies d'évitement pour de rares cas d'adaptation posant problème à la machine. Celles-ci relèvent bien entendu du hasard et ne sont pas nécessairement désirables, mais restent dignes d'être signalées, comme ce passage relatif à l'entrée d'un index que parcourent les personnages à la recherche d'un contresort dans un moment de tension :

(21) SRC : *Suddenly “UnDoing. Great. The.” appeared.*

HYP : Soudain, le nouveau sort apparut.

REF : Les mots « Dénouement (Le Grand) » s'imprimèrent aussitôt sur la page.

126. La comparaison entre ce système et un outil générique tel que *DeepL* l'est également pour les mêmes raisons, puisque l'on peut constater que dans tous les cas où le système maison s'écarte de la source pour reproduire un choix de la traductrice, *DeepL* offre quant à lui une traduction purement littérale.

Or, ces solutions dues au hasard des probabilités peuvent parfois créer des scénarios insoupçonnés, à l'image du *spoiler* produit par la machine que nous avons décelé et qui révèle par erreur l'identité d'un personnage.

Une dernière observation qu'il peut-être nécessaire de mentionner, car elle remet elle aussi en question une idée très répandue concernant la TA neuronale (surtout à ses débuts), est que la TAN n'est pas tout à fait une « boîte noire », ou tout du moins pas autant qu'on voudrait parfois la présenter. De fait, retourner aux corpus d'entraînement permet invariablement d'expliquer des « choix de traduction » opérés par le système, y compris et surtout lorsque ceux-ci semblaient incompréhensibles. Dans le cas des contresens et des hallucinations qui démarquent la TAN de la TAS, par exemple, il suffit de se référer aux mémoires de traduction utilisées pour trouver une raison aux traductions inattendues qu'il était autrement impossible d'expliquer au regard du texte source. Les deux exemples suivants illustrent respectivement ces deux scénarios, pour lesquels une traduction problématique et apparemment incompréhensible (a.) pouvait aisément s'expliquer par la présence de constructions similaires (b.) dans les données d'entraînement :

- (22) a. SRC : *He remembered it as a bright, joyful place buzzing with Magyk, with beautiful pictures flitting across the walls.*
HYP : Il avait souvenir d'elle comme d'un endroit éclatant et joyeux, avec des images à la couleur fanée flottant sur les murs.
- b. SRC : *Many of them were old Extraordinary Wizards, and the sight of a faded purple cloak flitting through the maze of endless corridors and rooms at the Palace was not unusual.*
REF : Beaucoup étaient des magiciens extraordinaires, ce qui explique que la vision d'une cape pourpre à la couleur fanée flottant dans le dédale de couloirs et de salles du palais n'avait rien d'inhabituel.
- (23) a. SRC : *Pale and drawn, with dark circles under her eyes, she watched Simon walk to the stairs.*
HYP : La morte, assise dans un cercle sombre sous ses yeux, regarda Simon se diriger vers l'escalier.
- b. SRC : *Pale and still, she lay on the landing stage with a peaceful smile on her lips.*
REF : La morte gisait sur le débarcadère, pâle et paisible.

Il s'avère dès lors au travers de ces exemples que l'explicabilité n'est pas tout à fait sacrifiée avec la TAN si l'on conserve une trace des corpus et un accès à ceux-ci. Une fois encore, on pourra y voir ici une nouvelle confirmation de l'importance des données et de la nécessité de s'assurer de leur qualité. Ces exemples confirment aussi en dernière analyse le constat selon lequel les erreurs produites par la TAN (en particulier les mauvais choix lexicaux, les omissions, les problèmes de logique, les glissements et les contresens) peuvent être plus difficiles à déceler, à comprendre et à corriger (RAGNI et VIEIRA 2021, p. 13), ce qui ne serait pas même possible sans un accès au texte source.

6.3 Évaluation experte : examiner l'outil en contexte réel

De manière générale, l'ensemble des évaluations conduites dans ce chapitre et dans celui qui le précède viennent souligner les avantages de la personnalisation de la traduction automatique et me permettent donc de répondre par l'affirmative à ma première question de recherche (QR1) : il est bel et bien possible d'adapter un système de TA au style individuel d'un traducteur ou d'une traductrice. Les résultats présentés ici apportent en outre une analyse plus fine des performances concrètes d'un tel outil, telles que soulevées par la première sous-question de recherche (SQR1), mettant en avant des signes d'adaptation qui dépassent largement de simples choix lexicaux, tout en pointant d'autres limites claires et de mieux en mieux documentées pour la TA neuronale.

De fait, la comparaison de différents systèmes et de différents aspects des traductions produites par ceux-ci confirme une fois encore l'importance cruciale des corpus pour la mise au point de la TA, ainsi que la nécessité d'aller plus loin dans le processus d'adaptation que le seul domaine d'application, pour proposer une personnalisation au niveau de l'humain. Bien que ce travail ait aussi soulevé la difficulté de trouver le bon équilibre entre l'affirmation du style personnel dans les sorties de traduction et leur qualité générale, il montre que les textes produits sont non seulement plus riches lexicalement que ceux produits par des outils génériques, mais aussi syntaxiquement plus variés, tout en restant cohérents avec l'univers fictif de la série et avec les choix de la traductrice. Sur ce dernier point, l'annotation d'erreurs et l'analyse des sorties font ressortir plus spécifiquement les marques d'adaptation au style propre de Nathalie Serval et l'intérêt d'un système entraîné sur des données littéraires stylistiquement riches et cohérentes à cet égard. L'outil ainsi développé apporte en effet une amélioration appréciable de la qualité, à la fois sur certains aspects traditionnellement mis en avant comme des éléments typiques de la littérature et hors de portée des systèmes (temps de la narration, prise en charge des dialogues,

cohérence lexicale, respect du registre...), de même que sur une série de critères mentionnés quant à eux comme des limites majeures de la TA. Or, le fait d'avoir entraîné un système avec des corpus qui contiennent des segments composés de plusieurs phrases, voire de paragraphes entiers (cf. Sections 5.3.3 et 6.2.3), permet d'obtenir des traductions très intéressantes du point de vue syntaxique. Il est même on ne peut plus éloquent de noter à ce titre que les points étant apparus comme une force du système personnalisé, à savoir la réorganisation syntaxique, la division (à laquelle j'ajouterais la fusion) de phrases, la modulation, l'explicitation ou encore la reformulation, couvrent chacun des aspects listés par AHRENBURG (2017, p. 26) comme principal point de divergence entre traduction automatique et traduction humaine.

Bien sûr, il apparaît tout de même de façon évidente que la traduction résultante demeure très loin de la parité avec les productions humaines. Ceci se note distinctement dans les relevés de l'évaluation d'erreurs. Toutefois, celle-ci indique par ailleurs que les problèmes sont attribuables, pour le système adapté, aux difficultés bien connues de la TA en général (traductions littérales, lissage lexical, mauvaise désambiguïsation lexicale...) plutôt qu'à des considérations proprement littéraires. Une bonne partie de ces écueils provient notamment du fait que la traduction automatique opère encore aujourd'hui au niveau de la phrase, considérant chaque segment de manière tout à fait isolée (VOIGT et JURAFSKY 2012, p. 21 ; MATUSOV 2019, p. 15). Une traduction au niveau du paragraphe permettrait de ce point de vue une meilleure prise en charge des éléments textuels (BAWDEN *et al.* 2018 ; THAI *et al.* 2022 ; KARPINSKA et IYER 2023), ainsi que je le suggérais pour la traduction assistée par ordinateur (TAO ; cf. Section 2.3), mais également stylistiques, comme en attestent les observations du système personnalisé dans la section précédente. Même avec de telles avancées, néanmoins, les écarts liés à l'adéquation, qui représentent la majeure partie des annotations et qui sont trop peu souvent évoqués à mon sens, continueraient de poser problème. Or, ces erreurs constituent un défi de taille, car elles empêchent la bonne compréhension du récit, tandis que la littéralité des traductions relevée plusieurs fois dans ce chapitre nuira à l'expérience de lecture. Évidemment, il n'est nullement question dans l'approche proposée ici d'envisager l'utilisation de la TA sans intervention humaine par la suite. Ces différentes considérations mettent surtout en évidence l'idée d'un outil au service des traducteurs littéraires, capable de fournir des suggestions utiles à la volée, bien que ce scénario dépendrait sûrement d'avancées supplémentaires dans le domaine de la TA, et d'une implémentation mûrement réfléchie de l'outil (considérons, par exemple, les différences entre l'affichage de propositions de TA en marge de l'éditeur de texte dans un environnement de TAO, la pré-traduction du texte dans cette même interface, l'affichage de phrases isolées dans un outil

de post-édition rudimentaire, ou encore la simple révision d'une TA brute dans un logiciel de traitement de texte).

C'est d'ailleurs précisément sur ces questions que portera la troisième et dernière partie de cet ouvrage. Jusqu'à présent, je n'ai effectivement proposé qu'une évaluation *intrinsèque*, principalement consacrée au produit de la traduction, au travers de diverses méthodes dont j'ai évoqué les limites tout au long de ce chapitre, mais que j'ai tenté d'utiliser de la façon la plus pertinente possible. De cette manière j'ai pu combiner et faire dialoguer ces différents résultats pour répondre à mes premières questions de recherche et par la même occasion confirmer mon hypothèse de travail concernant l'adaptation. Hervé BLANCHON et Christian BOITET (2007), commentant eux aussi nombre des écueils discutés dans ce chapitre, proposent cependant de privilégier les évaluations fondées sur la tâche, telles que le temps gagné pour la traduction, la compréhension du message, etc. En un mot, les auteurs proposent d'évaluer son utilité, qui portera dans le cas présent sur la traduction elle-même et sur l'aide qu'elle peut apporter à la traductrice. Prolongeant ce début d'évaluation humaine, plus qualitative, le chapitre suivant se livrera dès lors à une évaluation *extrinsèque*, visant à déterminer l'efficacité d'un système dans le contexte d'une application réelle et sa contribution à la réalisation de la tâche (REITER 2018, p. 398)¹²⁷. De telles études peuvent ainsi dresser un tableau bien différent des métriques et font défaut dans les travaux sur la TA selon ces auteurs. En ce sens, le déploiement de l'outil de TA adapté et sa présentation à Nathalie Serval constituent une suite logique de l'étude qui complétera adéquatement ce premier volet d'évaluation, et ce, d'autant plus que l'inclusion de la traductrice reflète la volonté de mettre l'humain au centre des technologies et de leur développement¹²⁸, et d'adapter celles-ci pas uniquement au métier des traductaires ou au domaine de spécialité, mais à chaque individu.

127. Ehud REITER (2018, p. 400) établit une comparaison avec le domaine médical, où une méthode ne serait pas acceptée par la communauté scientifique si elle ne faisait pas état des résultats sur des cas concrets observés dans des scénarios réels.

128. Ainsi que je l'ai décrit dans la [Section 2.3](#) et que le recommande O'BRIEN (2012, p. 109 et 115).

Troisième partie

Adapter la machine à l'humain

La littérature est, certes, un jeu combinatoire qui suit les possibilités implicites à son propre matériau, indépendamment de la personnalité du poète, mais c'est un jeu qui, à un moment donné, se trouve investi d'un sens inattendu ; non pas un sens contenu à ce même niveau linguistique sur lequel nous nous mouvions, mais un sens qui a glissé d'un autre plan ; un sens capable de mettre en jeu quelque chose qui, sur cet autre plan, fait poids pour l'auteur ou la société à laquelle il appartient. La machine littéraire peut effectuer toutes les permutations possibles sur un matériau donné ; mais le résultat poétique sera l'effet spécifique d'une de ces permutations sur l'homme doté d'une conscience et d'un inconscient, c'est-à-dire sur l'homme empirique et historique ; ce résultat sera le choc qui ne se vérifie que dans la mesure où, tout autour de la machine écrivante, vivent les fantasmes obscurs de l'individu et de la société.

ITALO CALVINO (1984)

Entretiens

L'évaluation de la traduction automatique (TA), quand elle ne fait pas uniquement appel à BLEU (MARIE *et al.* 2021), repose parfois sur des comparaisons de sorties de TA produites par différents systèmes (*ranking*) ou sur des évaluations directes portées sur des échelles de mesure ordinale (*direct assessment*). Dans ces derniers cas, les phrases sont souvent isolées les unes des autres, la source n'est pas toujours fournie, et les personnes qui effectuent ces évaluations n'ont pas toujours les compétences nécessaires. Plus rarement encore, c'est l'effort de post-édition (PE) qui est mesuré au cours d'une tâche menée dans des interfaces bien souvent peu adaptées à celle-ci. Dans le cas littéraire, plus spécifiquement, l'état de l'art a montré que les évaluations restent encore fortement guidées par les métriques ou par d'autres méthodes quantitatives telles que les mesures statistiques ou l'annotation d'erreur (cf. Chapitre 6).

Pour chacun de ces scénarios, l'accent est mis en premier lieu sur la qualité des sorties de TA. D'autres études ont tenté de dépasser cette seule perspective et sollicitent par exemple le lectorat pour juger des effets de la TA sur l'expérience de lecture, comme le montrent les travaux de BESACIER (2014), SHIH (2016), GUERBEROF-ARENAS et TORAL (2020) et OLIVER (2024). Cependant, peu de recherches se sont penchées jusqu'à présent sur l'opinion des traductaires durant la tâche de traduction littéraire elle-même, sur l'outil et sur son utilisabilité (*usability*). Sur le modèle de ces *reader-responses*, qui visent à évaluer la réception de la traduction automatique (ROSENBLATT 1994; CHESTERMAN 1998; SOUSA 2002), ce chapitre présente par conséquent une enquête qui s'inscrit plutôt dans une démarche de *user-response* et se focalise avant tout sur l'utilisation et l'ergonomie de la TA, au départ de l'outil précédemment mis au point.

La deuxième partie de ce travail et le chapitre précédent en particulier montrent qu'il est bel et bien possible d'adapter des systèmes à la littérature et au style individuel des traductaires, bien que d'autres cas d'étude devraient idéalement venir appuyer cette conclusion et montrer si des modèles plus robustes encore pourraient améliorer ces résultats. Si j'ai pu illustrer l'importance des corpus à cet égard et répondre par l'affirmative à ma première question de recherche (cf. Section 5.1), une autre question tout aussi essentielle reste toutefois de savoir quelle serait l'utilité d'un tel outil de traduction automatique pour les personnes amenées à l'utiliser. Or, comme le soutient Ehud REITER (2018, p. 398), « l'évaluation la plus fiable et la plus probante [pour la TA]

est celle d'un système qui a été déployé en situation réelle et dont on peut évaluer concrètement les effets dans un scénario donné [27] ». Ayant dès lors confirmé ma première hypothèse de travail, cette troisième et dernière partie d'ouvrage sera guidée par une seconde question de recherche (QR2) : La traduction automatique est-elle utile pour les traducteurs littéraires ?

Après m'être attardé sur l'outil même, j'aborderai donc ici la traduction littéraire automatique (TLA) du point de vue de ses usagers. Le changement de focalisation correspond au tournant inductif annoncé dans la réflexion épistémologique en seconde partie d'ouvrage (cf. [Section 5.1](#)), non pas dans l'idée d'opposer celui-ci aux résultats obtenus, mais plutôt de les compléter, d'y apporter quelque chose de nouveau et de possiblement inattendu. L'objectif poursuivi ici n'est pas tant de vérifier une hypothèse déterminée sur le rôle que devrait jouer la TA ou sur son utilité. Il s'agit plutôt de mener l'enquête, sans aucune sorte d'anticipation, et de partir de ces observations pour apporter des éléments de réponse concernant les avantages éventuels et améliorations désirables de l'outil, concernant ses effets sur le travail et sur la profession. Ce renversement de perspective vers des méthodes résolument qualitatives (empruntant aux sciences cognitives et à la sociologie) est motivé par l'appel d'un « programme de recherche socio-cognitif sur la traduction » ([RISKU et WINDHAGER 2013](#) ; [ROSSI 2018](#), p. 56), dont l'application à ce cas d'étude vise notamment à faire intervenir les personnes intéressées au premier chef et à donner une voix aux traducteurs, à l'instar de ce que recommandent Dorothy KENNY et Marion WINTERS (2020).

De fait, il est important de voir si un outil de TA adapté peut être mis au point (le changement de paradigme que laisse entrevoir ce constat est en effet primordial d'un point de vue fondamental), mais il est tout aussi pertinent de déterminer dans quelle mesure et dans quel contexte un tel outil pourrait venir en aide à celles et ceux qui seraient susceptibles de l'utiliser au quotidien. Ce chapitre me permettra en quelque sorte de boucler la boucle, en retournant vers la personne qui a produit le corpus à partir duquel j'ai pu concevoir le système personnalisé et en confrontant ce système à l'examen de cette traductrice professionnelle. Cette évaluation experte de l'outil proposé offrira ainsi une occasion unique de récolter l'avis et le ressenti de Nathalie Serval sur celui-ci, à la faveur d'un exercice de traduction assistée par ordinateur (TAO) et de quelques entretiens, mais aussi de donner une voix aux personnes qui traduisent sur le terrain et qui sont encore peu représentées dans les recherches sur les technologies de la traduction. Plus encore, cette dernière évaluation en situation réelle est l'unique moyen d'aborder des questions, indispensables elles aussi, liées à l'ergonomie, à l'acte traductif et au processus créatif. De mettre finalement l'accent sur les aspects humains des technologies de la traduction ([KENNY 2017](#)) et sur l'importance du facteur humain en traduction automatique ([CHAN 2018](#)).

7.1 Une évaluation experte en contexte réel

L'approche connexionniste de l'IA présentée dans l'historique de l'évolution de la TA (cf. [Section 4.3](#)) n'aura pas uniquement eu le mérite d'aboutir aux systèmes que l'on connaît aujourd'hui. Les fondements théoriques sous-jacents ont aussi influencé notre approche de la cognition, passant de la traditionnelle métaphore du cerveau comme ordinateur et système de traitement de l'information ([DAVIS et SUMARA 1997](#), p. 107) à une vision plus holistique et dynamique de la cognition et, partant, de la traduction : l'organe de la pensée n'est plus uniquement vu alors comme une simple base de données mentales traitant des symboles linguistiques selon des règles fixes, mais comme un réseau flexible, désordonné et intimement lié au corps, au sens et aux émotions ([RISKU 2010](#), p. 96). Dans ce contexte, la traductologie cognitive a mis en lumière deux notions essentielles pour la recherche en envisageant la traduction comme activité à la fois *incarnée et située*.

Parallèlement, le champ a marqué sa volonté de passer des *translation* aux *translator studies* ([CHESTERMAN 2009](#)). Marqué par la volonté de remettre les traductaires au centre de l'opération traductive, ce changement de paradigme entend placer l'accent sur les aspects humains de la traduction et sur des cas individuels (*translator-centered approach*), afin d'éclairer les processus mentaux (*process-oriented method*) qui guident cette activité ([KAINDL 2021](#)) :

Les *translator studies*, comme je le conçois, sont l'antithèse d'une mécanisation des *translation studies*. Je ne me réfère pas ici à la dimension technique qui recouvre la traduction automatique et autres technologies de la traduction, mais plutôt à l'image qui conçoit la personne traduisante en tant qu'élément fonctionnel, en tant que composante technique, là où il faut y voir un être humain dans toute son individualité [28]. (*Ibid.*, p. 22.)

Cette conception incarnée de l'activité traduisante se traduit par conséquent par la mise à l'avant-plan de l'individu et des cas particuliers. De la même manière, c'est la traduction en tant qu'activité intellectuelle, incarnée et individuelle qui primera à partir de maintenant dans ce travail. En un mot : la traduction en tant qu'acte « somatique » ([ROBINSON 1991](#)). L'approche développée ici envisage en effet la traduction non pas comme un résultat, mais comme un processus ; ou, comme l'explique Piotr BLUMCZYNSKI (2016), place le centre d'attention non plus sur le QUOI, mais sur le COMMENT. Elle est également le fruit d'un parti pris déjà évoqué (cf. [Section 5.1](#)), qui prendra tout son sens dans ce troisième chapitre : celui d'une approche « personnelle, phénoménologique, processuelle, existentielle et somatique [29] » (*Ibid.*, p. XI).

Cette prise en compte de l'individu traduisant induit dans le même temps une perspective située de la traduction qui doit s'extraire des laboratoires et favoriser la recherche dans des conditions réelles. De ce point de vue, l'intérêt principal de la recherche se porte sur la personne qui traduit, sur le processus créatif de cette activité située et sur son environnement concret, en ce compris les dispositifs technologiques (HALVERSON 2014, p. 117) :

Cette notion d'incarnation est en accord avec la perception de la traduction comme activité située de peut nous aider à mieux comprendre le rôle des outils et du contexte dans le processus de traduction ainsi que leurs effets sur la qualité du produit (RISKU 2010). Elle est aussi cohérente avec la conception de la traduction dans le champ émergent de la traductologie cognitive, où elle apparaît comme une manifestation d'une cognition située, incarnée et distribuée (MUÑOZ 2010). Adopter cette position signifie que les facteurs physiques, cognitifs et organisationnels pouvant affecter la traduction et la prise de décision doivent être sérieusement pris en compte si nous voulons réellement comprendre cette activité [30]. (EHRENSBERGER-DOW *et al.* 2016, par. 1.)

Le choix d'une enquête de terrain représente à ce titre une suite logique de l'étude, répondant elle-même au besoin d'approches écologiquement valides (KENNY et WINTERS 2020, p. 130), qui reflètent les conditions de travail réelles des traductaires (SALDANHA et O'BRIEN 2013, p. 33).

Une telle démarche ne se borne néanmoins pas à chercher à voir ce qui se passe dans le cerveau humain, mais vise plus largement l'être humain dans son entièreté, son histoire individuelle et son environnement (RISKU 2010, p. 95) :

Choisir cette position de base en traductologie signifie clairement que c'est le traducteur humain qui est au centre des recherches et que c'est cette activité humaine complexe qui est étudiée en situation, plutôt que l'outil. (LAVAUT-OLLÉON 2011, par. 19.)

L'entretien mené offrira dès lors la possibilité de prendre en compte les conditions de travail matérielles, individuelles, mais aussi sociales. De jeter un éclairage sur la situation des traductaires et sur les réseaux dans lesquels s'opère la traduction (HALVERSON 2014, p. 132), puisque ces considérations organisationnelles influent tout autant sur leur activité (EHRENSBERGER-DOW et MASSEY 2014, p. 62).

L'implication de la traductrice des romans choisis ne fait d'ailleurs pas uniquement sens d'un point de vue scientifique, mais aussi éthique. En faisant intervenir l'individu et en recontextualisant l'utilisation des outils, la recherche met en avant la dimension humaine de la technologie et éviter de se borner à son seul point de vue technique. Cette dimension sociale est essentielle à

un déploiement éthique et juste de la technologie, comme elle peut l'être dans nombre d'autres disciplines (cf. NUSSBAUM 1995). Elle rappelle à ce titre l'approche critique de la technologie promue par Philip AGRE (1997), puisque la conception de dispositifs impliquant une interaction humain-machine doit évidemment tenir compte du côté humain, en plaçant celui-ci à l'avant-plan, mais aussi assurer une prise en compte explicite des valeurs dans le processus de conception, ce qui vaut naturellement pour la recherche par ailleurs (HARRISON *et al.* 2011).

Une telle focalisation sur les utilisataires permet de remettre en question bon nombre d'idées reçues concernant les outils, mais aussi d'en prévenir les risques. Avec la montée des discussions concernant les biais de la technologie, les mentions de conséquences imprévues (ou non) des produits issus du traitement automatique des langues envers certains groupes socio-démographiques sont devenues de plus en plus courantes. Dans le meilleur des cas, ces effets se matérialisent par une simple baisse de performance, mais peuvent aussi conduire à des torts autrement plus préjudiciables (HOVY et PRABHUMOYE 2021, p. 1–3). Or, ces conséquences sont précisément liées au manque d'implication de certaines communautés dans le processus de développement des outils, comme le met en évidence Caroline CRIADO PEREZ (2019) dans l'ouvrage *Invisible Women*, prenant pour exemple le cas des femmes devant naviguer dans un milieu technologique majoritairement pensé par les hommes, ainsi que les multiples écueils engendrés par la disparité des genres dans cette industrie et la faible représentativité des femmes — mais aussi d'autres minorités (âge, classe, orientation sexuelle, variété linguistique, etc.) — dans les données utilisées par les systèmes. L'autrice pointe ainsi les risques d'une technologie qui n'est pas adaptée à ses utilisataires, comme le montre le cas de la reconnaissance vocale (TATMAN 2017), qui illustre la question de la représentativité des données, ou celui des assistants virtuels, dont le déploiement soulève des enjeux éthiques majeurs provenant de la conception même de ces logiciels (WEST *et al.* 2020).

Il en va de même pour les technologies de la traduction, sans surprise, comme nous l'avons discuté dans le chapitre d'ouverture (cf. Chapitre 2). Un exemple parlant est celui présenté par LÄUBLI *et al.* (2022), qui reconsidèrent et contestent l'un des éléments fondamentaux de la TAO : la présentation de l'interface bilingue en deux colonnes verticales telle qu'elle existe depuis plus de 20 ans. Pour le cas littéraire, en particulier, j'ai également pu montrer que les outils devraient idéalement être repensés pour tirer le meilleur parti de leurs fonctionnalités (HANSEN 2024). De la même manière, cette nouvelle expérience sur la TLA souligne l'importance cruciale de placer l'utilisataire au centre de la recherche et entend répondre à l'appel lancé à cet égard en traductologie depuis des années (O'BRIEN 2012, p. 116 ; EHRENSBERGER-DOW et MASSEY 2014, p. 62).

7.1.1 Nathalie Serval : figure d'une traductrice

Traductrice de l'anglais chevronnée, Nathalie Serval a traduit pas moins d'une centaine d'ouvrages pour des maisons d'éditions comme Denoël, Albin Michel, Casterman, Flammarion, Gallimard ou encore J'ai Lu. Si elle s'est illustrée dans de multiples genres, au travers de personnalités telles que Martha Grimes, Elizabeth George ou John Krakauer, il apparaît rapidement que Nathalie voue un attachement particulier aux littératures de l'imaginaire, comme en attestent ses nombreuses traductions dans les domaines de la *fantasy*, du fantastique ou de la science-fiction, parmi lesquelles on retrouvera des noms tels que Richard Matheson, Rick Riordan, Angie Sage, Kathe Koja, Tony Abbott, Gene Wolfe et d'autres encore.

Comme nombre de collègues de sa génération, son parcours débute non pas après des études de traduction, mais des études de lettres et civilisations anglophones, qu'elle entreprend naturellement à l'Université de Bordeaux après avoir baigné dans un environnement imprégné de livres, d'écriture et de traduction. Nathalie commence tout d'abord par traduire des nouvelles, sous les auspices d'Alain Dorémieux, qui vient de reprendre la direction de la revue *Fiction*. Convaincue par l'exercice, elle s'occupe ensuite de longs formats, toujours aux éditions OPTA, avant de traduire quelques années plus tard l'intégralité d'un recueil de nouvelles de Lisa Tuttle à la demande de Jacques Chambon, alors directeur de la collection « Présence du fantastique ». De fil en aiguille, Nathalie multipliera les expériences, les traductions et les rencontres éditoriales pour aboutir à l'oeuvre considérable qu'on lui connaît aujourd'hui, dans laquelle on retrouvera une récente réédition des nouvelles de Tuttle pour la maison d'édition associative Dystopia ¹²⁹.

Durant ces années, Nathalie a par ailleurs écrit des poèmes ainsi que plusieurs nouvelles. Elle s'est aussi exercée pendant un temps au travail d'édition, de gestion et de révision, notamment comme responsable des traductions aux Presses de la Cité. Toutefois, c'est à son rôle de traductrice indépendante et de véritable « passeuse de littérature », comme le laissent souvent entendre certaines figures de proue du domaine (BLUMCZYNSKI 2016, p. 41), qu'elle consacre la majeure partie de sa carrière.

Aujourd'hui, Nathalie s'est installée aux Pays-Bas, où j'ai eu la chance de pouvoir lui rendre visite. Le changement n'est évidemment pas anodin, puisqu'elle envisage d'y apprendre la langue et caresse l'idée d'y (re)trouver des textes d'autres encore peu connus dans l'espace francophone. Au fil des ans, ces projets plus personnels, rendus possibles en partie par le financement participatif,

129. *Les Chambres inquiètes*, Dystopia, 2014.

ont d'ailleurs pris une importance croissante. Parmi ces projets qui lui tiennent à cœur figure ainsi la publication d'un recueil de nouvelles fantastiques de Bob Leman aux Éditions Scylla ¹³⁰, traduit et compilé par Nathalie, qui m'a également fait part d'un nouveau projet de recueil autour des nouvelles de Jonathan Carroll, dont l'une des traductions lui avait déjà valu le Grand Prix de l'Imaginaire ¹³¹.

7.1.2 Déroutement des entretiens

La conduite de ces entretiens a été fortement inspirée dans ses choix épistémologiques et méthodologiques par le modèle de l'entretien compréhensif de Jean-Claude KAUFMANN (2013). Celui-ci part d'une pratique ordinaire (dans ce cas, l'utilisation de l'outil) pour explorer sans idée préconçue, dans un va-et-vient de co-construction de la recherche, un usage social. Le sociologue dégage ainsi de nombreux points de repère, parmi lesquels j'ai surtout retenu ici l'importance accordée à l'informatrice, l'implication du chercheur dans l'enquête et la démarche interprétative du matériau, la multiplicité des méthodes d'enquête employées, le terrain comme point de départ de la problématisation ainsi que le caractère dynamique, conversationnel et informel de l'enquête, qui doit permettre de guider celui-ci sans fermer la porte aux éléments imprévus (*Ibid.*). Dès lors, si j'ai abordé les entretiens avec un objectif déterminé, à savoir l'évaluation de l'outil en contexte réel, j'ai également voulu entamer cette expérience sans mise au point préalable avec la traductrice concernant par exemple les performances du système observées dans le chapitre précédent, afin de lui permettre d'interagir avec celui-ci sans préconception ou attente particulière. Dans cette démarche exploratoire, seule la tâche a été préparée et convenue en amont, de même que quelques thèmes clefs en vue d'une discussion semi-directive, le but étant de proposer une tâche qui soit la moins artificielle et contrôlée possible, pour en renforcer la validité écologique, et d'ouvrir une discussion spontanée, pour recueillir des impressions authentiques. Cette volonté se note dans le caractère dynamique de la discussion, dans ses ruptures et dans l'apparition d'éléments de discussion non sollicités qui s'en révèlent d'autant plus pertinents. On s'apercevra effectivement que la qualité y est finalement peu abordée, tandis que des éléments moins attendus d'ergonomie en ressortent très largement. C'est évidemment un développement heureux, puisque le chapitre vient ainsi compléter avantageusement celui de l'évaluation, en y apportant de nouvelles considérations ergonomiques, cognitives et sociales — les plus importantes peut-être — qui manquaient encore à cette recherche.

130. *Bienvenue à Sturkeyville*, Scylla, 2020.

131. *L'Enfant arc-en-ciel*, Denoël, 1998.

Le point au travers duquel cette étude s'écarte peut-être le plus de l'entretien compréhensif est celui de la saturation, qui incite celui ou celle qui mène l'enquête à consolider le matériau d'enquête jusqu'à saturation du modèle, en multipliant les observations puis en confrontant le modèle saturé à de nouveaux cas (KAUFMANN 2013, p. 103–105). Or, l'enquête ne concerne pas ici un groupe social déterminé mais une seule personne : la traductrice officielle du roman choisi pour ce cas d'étude, pour qui le système a été adapté. Ce point était évidemment nécessaire pour appréhender le problème dans sa dimension personnelle, processuelle et créative, ainsi que je viens d'en discuter. En cela, la démarche adoptée pour cette série d'entretiens est largement inspirée, à une échelle bien plus humble, de la recherche sur le processus de traduction de Claudine BORG (2023) en traductologie cognitive :

La méthode déployée [pour cette recherche] s'inscrit dans la perspective de la traductologie cognitive et combine des techniques issues de ce paradigme (verbalisation, brouillons, enregistrements vidéo) avec des approches ethnographiques (observations et entretiens), afin de mieux comprendre la pratique d'un traducteur littéraire d'expérience travaillant *in situ* [31]. (*Ibid.*, p. 3.)

L'autrice présente cette approche comme l'alternative à l'étude traditionnelle de la traduction littéraire — à laquelle j'ajouterais assurément les études sur la traduction automatique —, dans le sens où celle-ci s'est majoritairement penchée jusqu'à présent sur les produits finis : les traductions finales, les comparaisons source-cible et, quelques fois, les versions brouillons. S'opposant à un examen purement rétrospectif, elle offre ainsi l'avantage indéniable de pouvoir jeter un nouvel éclairage sur les processus cognitifs et créatifs des traducteurs durant leur travail, comme l'ont proposé également, quoiqu'avec des méthodes radicalement divergentes, KENNY et WINTERS (2020), tout d'abord sur un extrait du roman *The Beautiful and Damned* de F. Scott Fitzgerald et ensuite sur le roman *The World in the Evening* de Christopher Isherwood (WINTERS et KENNY 2024), ainsi que Waltraud KOLB (2024), sur la nouvelle « A Very Short Story » d'Ernest Hemingway, avec une traduction en allemand dans ces trois cas.

Dès lors, si nous avons cherché dans les précédentes sections à montrer ce qui se passait dans la supposée « boîte noire » des systèmes de TA — et à démontrer que ces outils n'étaient pas si opaques qu'on voudrait le laisser croire, puisqu'il s'agit avant tout d'une question de données — nous tenterons ici de voir ce qu'il se passe dans une machine autrement plus complexe, à savoir « la boîte noire du traducteur » (PAPAVASSILIOU 2007). L'approche adoptée ici se place donc « à l'interface des dimensions cognitives et situationnelles de la traduction » (ENGLUND DIMITROVA et EHRENSBERGER-DOW 2018, p. 13 ; ROSSI 2018, p. 58), mettant l'accent sur l'individu, les

méthodes et processus de traduction, la créativité et la dimension sociale du métier. Évidemment, l'ajout de dispositifs technologiques influe considérablement sur chacun de ces aspects et doit être étudié au prisme de la discipline qui se propose de mettre en avant les facteurs humains de l'interaction humain-machine : l'ergonomie (LAVAUT-OLLÉON 2016).

Une différence notable avec nombre d'études sur la traduction outillée, cependant, est que je n'aborderai pas directement la question de la productivité pour le cas littéraire. De fait, s'il est clair que l'avancée de la TA tient surtout à des enjeux économiques, je n'insisterai pas tant sur la rapidité de la tâche que sur « l'optique ergonomique de la satisfaction au travail » (BRUNETTE et O'BRIEN 2011, par. 6–7). Cette raison tient, d'une part, au nombreux biais qu'implique l'utilisation d'un nouvel outil (la TA), dans une interface de travail qu'il faut prendre en main (la TAO), pour une tâche peu commune en littérature (la PE), le tout en présence d'un observateur. D'autre part, ce choix reflète ma position évoquée en début de section, mais aussi l'envie d'envisager et d'élaborer un scénario plus proche de la *traduction humaine assistée par ordinateur* que de la *traduction automatique assistée par l'humain*. Ce scénario part du principe que les outils devraient servir avant tout à trouver de l'inspiration, à faciliter la résolution de problèmes localisés ou à comparer différentes solutions de traduction plutôt qu'à gagner en vitesse ; à diminuer sa charge cognitive, quitte à compenser le temps éventuellement gagné en passant davantage de temps sur les passages critiques qui demandent des solutions originales de la part des traductaires.

À la place, cette étude combinera elle aussi les approches mentionnées par Borg au cours de l'étude réalisée sur la traduction en maltais de Toni Aquilina, à partir de la novella « Monsieur Ibrahim et les fleurs du Coran » d'Éric-Emmanuel Schmitt : entretiens, observation et protocoles de verbalisation (cf. KUSSMAUL et TIRKKONEN-CONDIT 1995 ; KRINGS 2001).

Cette étude propose une contribution originale en traductologie en explorant en profondeur les processus adoptés par un traducteur expérimenté [...] Elle fait donc intervenir une méthode de recherche différente de celle généralement adoptée jusqu'à présent : un unique sujet de recherche et un long texte entier, par opposition à de courts textes soumis à un nombre de participants. Une telle approche pourrait nous apporter un nouvel éclairage sur la manière de traduire et, si ce type d'étude venait à se répandre à l'avenir, nous apporter suffisamment d'observations sur le sujet pour permettre des comparaisons entre ces études [32]. (BORG 2023, p. 30–31.)

Cette approche « humanisante » accorde de ce fait la priorité à la visibilité de la personne, son individualité, son agentivité et sa subjectivité (KAINDL 2021, p. 12). Elle met en lumière les traductaires, leur travail ainsi que le long et patient processus qui permet d'atteindre ces

résultats (BORG 2023, p. 2). Si je me penche donc sur un cas individuel, celui-ci est tout à fait assumé et n'empêche pas un certain degré de généralisation. Contrairement à de plus larges enquêtes et contributions qui offrent déjà un bon aperçu des usages et pratiques liés aux outils (SLESSOR 2020; RUFFO 2018; 2021; 2022; DAEMS 2022; HADLEY 2024), le cas présenté ici permettra de faire ressortir des pratiques individuelles détaillées en situation réelle et de mieux comprendre les processus et méthodes de traduction littéraire propres à des individus (HALVERSON 2014).

La récolte des données pour cette évaluation experte s'est déclinée en trois phases, de manière à accorder mon calendrier avec celui de Nathalie, mais aussi et surtout à pouvoir varier les sujets de discussion et les méthodes, dont je décris la mise en place d'un point de vue pratique dans cette section avant d'en venir aux analyses tirées de ces trois rencontres à la suite. Ainsi, un premier échange à distance nous a permis de lancer la discussion et de fixer les modalités de l'évaluation, pour laquelle notre choix s'est arrêté sur un exercice de post-édition, après quoi nous nous sommes rencontrés à son domicile pour une journée d'observation et un nouvel entretien autour de la tâche proposée. Par la suite, étant donné que nos précédents échanges avaient soulevé à cette étape de nombreuses contraintes liées à l'interaction avec le système de TA, nous nous sommes revus une dernière fois pour tenter de dégager un nouveau dispositif de mise en place de la TA et dresser un bilan rétrospectif plus général sur le sujet qui nous occupe. Ces différentes phases sont reprises dans le [Tableau 14](#), qui fournit une vue d'ensemble de la conduite des entretiens.

Phase	Date	Lieu	Type d'entretien	Durée
1	8/12/2022	En ligne	Entretien non directif	1 h 45
2	5/5/2023	Domicile	Observation et verbalisation	5 h 45
			Entretien semi-directif	1 h 15
3	18/12/2023	En ligne	Entretien rétrospectif	1 h 45

TABLEAU 14 – Méthode de conduite des entretiens et de récolte des données

Après une première prise de contact concernant ce nouveau projet de recherche, à laquelle mon interlocutrice a répondu avec enthousiasme, je me suis en effet entretenu une première fois en ligne avec elle (cf. SERVAL 2022), pour parler du chemin parcouru depuis nos derniers échanges et discuter du projet, sans toutefois divulguer les résultats déjà obtenus à ce stade — que je viens de présenter dans le [Chapitre 6](#). Ceci tient principalement à la nécessité de récolter

des impressions à chaud, sincères et neutres, qui ne soient pas influencées par une présentation préalable de mes observations. Pour cette même raison, je n'ai pas non plus conduit de pré-test, afin de ne pas biaiser la première utilisation du logiciel par la traductrice. Au lieu de cela, nous avons tiré profit de ce premier entretien pour convenir ensemble des modalités pratiques de la mise en place et de l'évaluation de l'outil élaboré, me permettant par la même occasion de recueillir ses premières impressions concernant la traduction littéraire automatique.

Lors de notre deuxième entretien, j'ai pu me rendre chez Nathalie pour installer le dispositif impliquant à la fois de l'observation directe et une discussion post-tâche (cf. [SERVAL 2023a](#)). Le cas d'étude sur lequel s'appuie cette tâche est toujours celui de la série *Septimus Heap*, mais j'ai jugé utile pour cette nouvelle expérience de mettre l'outil à l'épreuve sur un autre volume de la saga, à savoir *Fyre* (HarperCollins 2013). De fait, s'il était impérativement nécessaire de pouvoir compter sur une traduction publiée pour les évaluations menées jusqu'à maintenant, ce nouvel exercice se prêtait quant à lui sans problème à la traduction du septième et dernier tome de la série, qui n'a jamais été publié en français. Après avoir longuement examiné les sorties de traduction du sixième volume en vue de confirmer le processus d'adaptation du système, cette nouvelle mise à l'essai constituait dès lors un scénario idéal pour en obtenir une seconde confirmation et, potentiellement, de nouvelles observations. Le jeu de test sélectionné pour ce faire est composé des trois premiers chapitres de *Fyre*, qui offre tout comme le volume précédent un corpus d'évaluation varié, proposant dans ce cas trois thématiques fort différentes, dont deux sont situées dans le passé et donc tout à fait nouvelles par rapport aux tomes précédents, mais aussi une alternance des temps (un prologue au présent et deux chapitres au passé), des personnages introduits, etc. Dans l'ensemble, ces différents extraits sont ainsi suffisamment long, mais aussi suffisamment diversifiés pour permettre à la traductrice de se faire une idée précise de l'outil.

Néanmoins, ce changement du corpus d'évaluation intervient également en sachant qu'il existe une traduction non publiée de la moitié du roman en français¹³², dont la traductrice a très aimablement accepté de me fournir les trois extraits choisis. Pour m'assurer de proposer à Nathalie un scénario comparable à celui du tome six qui a retenu mon attention jusqu'ici, j'ai tout de même vérifié les performances du système de TA personnalisé (le modèle *Septimus-étendu* qui a fait l'objet d'une évaluation exhaustive dans le chapitre précédent) sur ces trois nouveaux

132. La maison d'édition avait commandé celle-ci à la sortie de la version anglaise, avec l'objectif de faire paraître ce dernier tome en deux volumes, avant de revenir sur sa décision juste avant sa publication pour des raisons commerciales.

chapitres, pour lesquels une évaluation par les métriques automatiques révèle des résultats en tous points comparables à ceux obtenus pour les trois derniers chapitres du sixième volume (cf. [Tableau 7](#)). Ces résultats sont donnés en comparaison avec la traduction personnelle de Nathalie, qui avait été effectuée environ 10 ans auparavant et qui s'avère surtout avantageuse par un second aspect. Étant donné le nombre d'années qui séparent cette première traduction de l'exercice envisagé pour cette recherche, il m'aura été possible de comparer ces deux modalités de travail sur un pied d'égalité, plutôt que de proposer un extrait à traduire avec l'aide de la TA et un autre sans. La version originale est en effet suffisamment lointaine pour que Nathalie ne s'en souvienne pas et aborde cet exercice comme elle avait pu le faire la première fois, sans biaiser la comparaison des deux modalités.

L'observation de la tâche proposée a été facilitée par le logiciel libre de capture d'écran *OBS Studio* ([OBS Project 2012](#)), qui venait s'ajouter à un carnet de notes manuscrites, au fichier de travail récupéré en fin de tâche et à la traduction non publiée de *Fyre*, produite par Nathalie il y a 10 ans. L'entretien a quant à lui été enregistré sur micro pour en permettre la retranscription en annexe. Dans ces deux cas, j'ai obtenu l'accord de Nathalie pour l'enregistrement des données, qui a par ailleurs fait l'objet d'une déclaration validée de protection des données ¹³³.

Sept mois environ après notre deuxième entrevue, j'ai proposé à Nathalie de nous entretenir une dernière fois en ligne pour solliciter son avis sur des modifications ergonomiques que j'avais pu intégrer à partir des remarques survenues en fin de tâche. Ce dernier échange nous a ainsi permis d'évoquer des pistes d'amélioration du dispositif, mais aussi de discuter une dernière fois avec davantage de recul sur nos entretiens et sur la thématique de manière plus générale (cf. [SERVAL 2023b](#)) ¹³⁴. Bien entendu, une compensation a été accordée pour cette participation à la recherche, sous la forme d'un montant forfaitaire de 250 euros par journée de travail.

133. Le traitement des données liées à cet entretien a fait l'objet d'une instruction de conformité par le délégué à la protection des données (DPO). Les informations personnelles collectées sont traitées conformément au règlement général (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 sur la protection des données (RGPD) et à la loi « Informatique et libertés » du 6 janvier 1978 modifiée.

134. Un dernier entretien impliquant une journée de travail supplémentaire avec observation, verbalisation et entretien semi-directif est prévue en prolongement de ce travail, afin de mettre une nouvelle fois l'outil à l'épreuve en tenant compte des modifications proposées.

Interlude : prise de contact

J'ai eu la chance de rencontrer Nathalie pour la première fois à l'occasion de mon mémoire de Master, dont le sujet portait alors sur la traduction littéraire assistée par ordinateur (TLAO)¹³⁵. Il m'avait fallu trouver dans ce contexte un cas d'étude qui me permettait à la fois de proposer une traduction originale, pour les besoins de cette évaluation finale, mais aussi de constituer une mémoire de traduction pertinente et suffisamment conséquente pour pouvoir évaluer l'intérêt des outils d'aide à la traduction dans ce domaine. Or, outre mon affinité personnelle pour la *fantasy*, la saga *Magyk* avait aussi ceci de particulier que le dernier tome n'était jamais paru en français, et ce, malgré le succès de la série. Je me suis donc attelé à la tâche, avec l'idée si j'y parvenais d'en discuter avec la traductrice officielle des romans d'Angie Sage.

C'est à la suite de ce premier contact d'étudiant incertain et admiratif que j'ai connu Nathalie, qui avait accepté avec beaucoup de curiosité de répondre à mes questionnements sur cet ouvrage, sur le métier de la traduction littéraire et sur les possibilités de la TLAO, et qui s'était intéressée en retour à ces balbutiements de recherche. Nous sommes restés en contact depuis, échangeant occasionnellement quelques nouvelles concernant nos projets respectifs.

Dans la mesure où ce nouveau pan de recherche porte à nouveau sur cette même saga littéraire, de manière à faciliter les éventuels parallèles tout en tirant profit de ce travail antérieur, il était naturel que je m'entretienne une fois de plus de ce projet avec la principale intéressée. L'idée d'impliquer dans cette étude une traductrice professionnelle qui puisse juger l'utilité de l'outil et des traductions proposées était bien entendu prévue dès la genèse de la recherche. Néanmoins, celle-ci s'est considérablement renforcée lorsque la direction de la thèse s'est peu à peu précisée, au départ de la traduction automatique adaptée au domaine littéraire, vers l'idée d'outils adaptés au style individuel des traductaires. Nous avons ainsi renoué le contact, sans remarquer immédiatement combien d'années s'étaient écoulées depuis notre première entrevue. Une fois encore, Nathalie a répondu favorablement à ma demande pour se lancer avec le plus grand intérêt dans les tâches et les discussions proposées.

135. Cf. HANSEN (2017).

7.2 Première phase : discussion préliminaire

L'une des premières questions que j'ai posées à Nathalie a été de lui demander si elle avait été amenée à utiliser des outils de traduction automatique par le passé et si elle avait déjà eu des échos de collègues pour qui cela aurait été le cas. Si je m'attendais à ce que la première réponse à ma question soit négative, il me semblait surtout utile de comprendre les raisons de cette absence d'utilisation préalable de la TA. Le premier élément de réponse donné à cet égard tient tout d'abord au fait qu'elle n'en a jamais eu le besoin ni reçu de demandes en ce sens jusqu'à présent. Par la suite, Nathalie a pu apporter une autre raison qui me paraît tout aussi essentielle, quoique rarement mise en avant, et qui concerne le manque de temps dû à l'enchaînement effréné des traductions, dont le rythme laisse finalement peu de temps à la veille technologique et à l'expérimentation :

Quand les traductions s'enchaînent, ça laisse peu de temps entre deux pour changer ses habitudes. Et c'est vrai que je préfère avoir quelqu'un qui me montre, quitte même à suivre une formation rapide plutôt que de me plonger dans un manuel. (SERVAL 2023a.)

De fait, un temps d'adaptation est nécessaire pour se familiariser avec ces outils, mais il n'est pas toujours aisé d'y consacrer une période qui peut facilement être ressentie comme une perte de temps. À ceci, la traductrice ajoute le doute lié aux performances de l'outil et aux intérêts qu'il pourrait lui apporter concrètement. On peut dès lors aisément comprendre pourquoi la question ne s'était pas même posée jusque là et pourquoi elle n'a pas jugé utile d'y recourir, ne serait-ce que de manière sporadique, d'autant que les outils qu'elle utilise (dictionnaires, correcteurs...) conviennent déjà pleinement à sa pratique.

Du témoignage de Nathalie émerge un autre élément central de la problématique, qui est celui du manque de formations actuellement proposées aux traducteurs littéraires. Non pas uniquement en ce qui concerne la TA, mais plus généralement l'ensemble des technologies de la traduction. Ces discussions, cours et ateliers sont en effet d'une importance primordiale pour qui voudrait savoir comment fonctionnent les outils et quelle peut en être l'utilité, savoir si l'investissement — personnel, matériel et financier — en vaut la peine, et surtout pouvoir se forger sa propre opinion. Les choses sont peut-être cependant en train de changer, puisque notre entretien a pu en outre mettre en lumière le nombre grandissant de témoignages et d'échanges consacrés à la traduction automatique dans les réseaux de traducteurs littéraires, favorisés à la fois par des demandes émanant des maisons d'édition, par l'attention médiatique portée au sujet et par des expériences personnelles (SERVAL 2022).

7.2.1 Définition de la tâche

Pour que Nathalie puisse donner un avis éclairé sur l'outil, je lui ai proposé plusieurs types de tâche impliquant chacune une interaction avec les sorties de TA et différents degrés d'investissement, parmi lesquelles une simple lecture des sorties à distance, une comparaison de différentes traductions automatiques en regard de la traduction humaine, ainsi qu'un exercice de post-édition. Nathalie et moi nous sommes finalement accordés sur cette dernière proposition, qui constituait un scénario particulièrement pertinent pour ce chapitre pour deux raisons. D'une part, ce cas de figure me permet de rester au plus près des considérations sur lesquelles je me suis longuement attardé, en proposant une évaluation de l'outil dans une situation réelle qui m'amène à adopter une perspective plus large que la seule qualité des sorties. D'autre part, cette mise en place rend possible l'intégration de l'outil dans une interface adaptée et accompagnée d'une multitude de fonctions, privilégiant dès lors l'interactivité. Or, la recherche montre que la manipulation en situation réelle permet d'avoir un meilleur ressenti sur la TA, qui peut autrement varier s'il est évalué avec ou sans utilisation de l'outil (ROSSI et ESPERANÇA-RODIER 2019). Ces raisons sont les mêmes qui ont motivé Nathalie à choisir sans hésitation ce dernier cas de figure. Durant notre entretien, elle m'a effectivement confié qu'elle n'avait pas de réel a priori sur la production de la machine et qu'elle était au contraire sincèrement curieuse de pouvoir s'essayer à la tâche afin de voir de quoi était capable l'outil. Consciente que le métier évolue constamment, elle m'a ainsi fait part de sa position selon laquelle il faut éviter de se fermer complètement au changement, mais plutôt essayer de s'y préparer pour en tirer le meilleur parti.

L'avantage de ce projet et du choix de la post-édition est qu'il repose sur une précédente expérience de traduction littéraire assistée par ordinateur, basée sur la même série de romans que j'ai eu l'occasion de longuement analyser par le passé et pour laquelle j'avais notamment pu constituer une mémoire de traduction (MT) qui a servi de point de départ à cette nouvelle étude et que j'ai pu utiliser à nouveau pour cette tâche. *Fyre*, le septième tome de cette série qui servira ici aussi de cas d'étude, a par ailleurs été traduit en partie par Nathalie il y a une dizaine d'années, offrant par conséquent une possibilité de comparaison avantageuse avec une éventuelle post-édition. Bien que je sois conscient que le décalage temporel entre en ligne de compte, celui-ci donne malgré tout une occasion unique de comparer les deux modalités de traduction de façon totalement indépendante, raison pour laquelle j'ai expressément demandé à Nathalie de ne pas y retourner avant la tâche proposée.

7.2.2 Impressions initiales

Avant de recueillir les impressions de Nathalie sur l'exercice convenu, j'ai cherché à savoir ce qui serait pour elle l'outil idéal, ce qu'elle attendrait d'un outil personnalisé et comment elle envisagerait une éventuelle collaboration humain-machine si elle ne devait tenir compte d'aucune contrainte technique. Une première piste lancée serait celle d'un outil qui permette d'accélérer la traduction dans ses aspects « automatiques », de manière à pouvoir consacrer davantage de temps au côté proprement « créatif ». Nathalie pose ainsi l'idée au départ d'une réflexion sur deux extrêmes entre lesquels varie l'acte traductif : « Dans toute traduction, même littéraire, il y a une part d'automatisme. Le cerveau est en roue libre, puis on se retrouve tout à coup confronté à un écueil, une difficulté, qui demande de basculer dans la création » (SERVAL 2022).

On peut raccrocher ce point de vue à la notion de *flow*, issue de la psychologie et appliquée par Douglas ROBINSON (2012) à la traduction. Utilisant pour sa part la métaphore d'une navette de métier à tisser, le traductologue décrit de manière analogue un acte traductif passant tantôt d'une sorte de transe subliminale (le *flow*), correspondant à un mode de fonctionnement qui fait appel à la mémoire procédurale et aux automatismes pour traduire de façon presque mécanique, à un autre mode plus analytique, au cours duquel la traduction devient pleinement consciente et nécessite de consulter des ressources diverses, de comparer différentes solutions possibles, d'effectuer des modifications au niveau syntaxique ou pragmatique, etc. (*Ibid.*, p. 60–61). Ce schéma pose ainsi l'activité de traduction comme une navigation constante entre ces deux états mentaux et ces deux processus. Un mouvement de navette tirailé entre l'habitude et la nouveauté, l'inconscient et le conscient, la rapidité et les ralentissements, les normes et le ressenti, l'intuition et la réflexion...

Si la traduction automatique est incapable à l'évidence d'opérer d'un côté de ces deux extrêmes, à savoir celui de la réflexion et de la création, qui fait intervenir la capacité humaine à trouver des solutions originales et à s'écarter des normes pour proposer quelque chose de nouveau, on peut toutefois se demander si la machine permettrait d'avancer plus vite sur le versant opposé. Nathalie poursuit cette idée en précisant : « Ce qui serait intéressant de mon point de vue de traductrice, ce serait de pouvoir avancer plus vite. Que la machine permette d'avancer plus vite sur ce côté automatique, machinal, de façon à pouvoir consacrer plus de temps à la part créative » (SERVAL 2022).

Envisagée de la sorte, la TA interviendrait donc comme un moyen de réserver plus de temps aux problèmes de traduction, aux aspects proprement humains et à la réflexion, de manière à renforcer la créativité :

Parfois, ce qui est frustrant avec les délais de traduction, c'est que l'on sent qu'on a trouvé une solution que l'on pourrait pousser plus loin, mais on n'a plus le temps et on essaie de trouver la moins mauvaise solution. Parfois, à la relecture, comme on est déjà sur un texte qui est déjà quasiment terminé, on peut se concentrer davantage sur les aspérités qui restent, et c'est vrai qu'avoir un outil qui permettrait de faire la même chose tout en produisant dès le départ quelque chose d'assez proche de sa voix personnelle, ce serait intéressant. (*Ibid.*)

Nathalie met ceci en parallèle avec la nécessité parfois de faire appel à un ou une collègue pour effectuer un premier jet avec l'accord du commanditaire — souvent la même personne, il s'agit dans son cas d'une personne de la famille —, et ce, pour des contraintes de temps à nouveau. Toutefois, même si cet arrangement permet d'éviter un aspect machinal dans le rapport au texte et de gagner en vitesse sur tout ce qui ne nécessite pas de recherche particulière, il reste néanmoins toujours une sensation de décalage liée au fait qu'elle ne retrouve pas sa voix dans le texte, et qui demande tout de même plus de travail qu'on ne pourrait l'imaginer.

Dans ce contexte, l'outil idéal décrit intervient dès lors comme « une sorte de deuxième cerveau avec qui pourrait s'instaurer un dialogue, un double qui serait performant en permanence — ce qui n'est pas le cas de nos cerveaux biologiques — et qui pourrait proposer des solutions, voire plusieurs solutions comme on le fait souvent inconsciemment » (*Ibid.*). Au cours de la discussion, Nathalie revient ainsi plusieurs fois sur cette idée de personnalisation de l'outil à sa voix, mais aussi sur le dialogue possible avec la machine et la possibilité de confronter différentes solutions de traduction dans une interface de travail qui apparaîtrait sur demande de l'utilisateur. Cette dernière considération vient quant à elle d'une crainte vis-à-vis de l'outil, à savoir la peur de se laisser distraire si celui-ci est trop présent et qu'il n'est pas possible de le solliciter quand on veut.

Que ce soit dans la recherche ou dans les témoignages des traducteurs, on retrouve cette dualité entre la traduction envisagée comme un processus mécanique, constant, et comme un état de quasi-catharsis créative. La question, à ce stade, serait donc de savoir si la machine aurait un rôle à jouer dans cette logique et à quel niveau. Si elle pourrait faciliter le travail de l'humain dans le premier cas, de manière à ce que les traducteurs puissent se concentrer sur le second, ou si elle viendrait perturber le passage de l'un à l'autre de ces deux extrêmes, que ce soit en contraignant le processus créatif ou en éliminant tout simplement le versant mécanique et en empêchant dès lors la transition entre ces deux pôles complémentaires.

Interlude : Pays-Bas

Cinq mois se sont écoulés entre le premier entretien et ma seconde rencontre avec Nathalie, le temps pour moi de préparer la mise en place convenue et de préparer les thèmes de discussion que le second entretien mènera finalement dans une direction plus inattendue, quoiqu'autrement plus originale. Durant cette période, j'ai par ailleurs eu la possibilité d'effectuer un séjour de recherche à DCU (*Dublin City University*) sous la supervision de Dorothy Kenny, dont nombre de travaux ont été déjà mentionnés et dont l'implication dans un projet similaire fut d'une aide inestimable (cf. [KENNY et WINTERS 2020](#) ; [WINTERS et KENNY 2024](#)). Si cette retraite irlandaise fut dès lors déterminante pour la préparation de ce chapitre, il ne fait aucun doute que les discussions, lectures et réflexions issues de ce séjour en ont par ailleurs considérablement renforcé l'assise théorique.

Juste avant mon retour en Belgique, j'ai ainsi pu partir à la rencontre de Nathalie, qui m'a chaleureusement accueilli dans le petit village fortifié des Pays-Bas où elle s'est récemment expatriée. Pour ce deuxième entretien et premier exercice pratique, l'idée était de proposer une tâche qui permette d'interagir avec l'outil, mais qui n'intervienne pas uniquement dans des conditions expérimentales et qui soit la plus proche possible des conditions de travail réelles de Nathalie. Nous nous sommes donc installés dans le salon lumineux et mansardé de son domicile qui surplombait la ruelle encore parée des décorations du *Koningsdag*. Après un café et quelques complications techniques réglementaires, Nathalie s'installait avec un ordinateur portable, quelques fidèles dictionnaires papier truffés de notes, un accès direct à toute autre ressource nécessaire, et le travail était lancé, sous le regard impérieux et attentif de Finn, son compagnon de travail félin.

Au bout de presque 6 h d'exercice, le nécessité d'une pause nous a amenés le long des canaux, où nous avons longuement causé de traduction, de traductologie et de formation des traductaires. Nous y avons été rejoints en fin de parcours par Finn, le temps d'un dernier tour sur le chemin de messe municipal, après quoi le débriefing s'est imposé de lui-même en me faisant presque oublier mes obligations d'enquêteur.

7.3 Deuxième phase : première interaction avec l'outil

Pour ce premier exercice pratique, Nathalie et moi nous sommes donc accordés sur une tâche de post-édition. J'ai conscience que l'utilisation d'une interface rarement utilisée en traduction littéraire — pas du tout dans ce cas — introduit un biais, tout comme la tâche de PE elle-même. Toutefois, et même s'il faut reconnaître que chacune de ces tâches demande des compétences spécifiques, Nathalie est habituée à travailler à partir d'un premier jet produit par une collègue, ainsi qu'elle me l'a indiqué lors de notre premier entretien, et elle fait montre d'une bonne expérience de la révision puisqu'elle a été responsable des traductions pendant deux ans aux Presses de la Cité. L'objectif poursuivi ici, de plus, n'est pas d'évaluer des facteurs comme la productivité, pour lesquelles l'expérience de PE serait déterminante, mais de lui permettre d'interagir le plus naturellement possible avec le système de TA personnalisé. Il s'agissait en effet du meilleur moyen pour moi de proposer une démonstration de l'outil à Nathalie, mais aussi d'y intégrer d'autres ressources et de récolter des impressions authentiques. Plus encore, l'exercice proposé offrait un excellent dispositif pour observer ce qui aurait tendance à être écarté ou conservé, pour vérifier si les changements se rapprochent ou non de la traduction existante et pour voir, *in fine*, si la machine peut être utile à la traductrice et si elle influe d'une manière ou d'une autre sur son style.

D'autre part, la manière dont a été intégrée la TA pour cette expérience rejoint les considérations exposées en début de chapitre, dans le sens où il me semblait préférable d'éviter la simple post-édition d'une sortie de traduction automatique dans un logiciel de traitement de texte, à la faveur d'un scénario permettant une plus grande interaction avec l'outil. J'ai pareillement écarté l'utilisation d'interfaces de post-édition classiquement utilisée dans la recherche telles que *PET* (AZIZ *et al.* 2012), qui, bien qu'elles proposent des fonctionnalités précieuses pour la recherche sur l'évaluation de la TA, contraignent fortement l'utilisateur dans le processus de traduction et s'éloignent des conditions de travail habituelles des traducteurs (VIEIRA *et al.* 2023, p. 217). Au contraire, l'objectif recherché ici était d'imaginer un cas de figure au plus proche de ce qu'il pourrait être pour des traducteurs qui se lanceraient dans une telle démarche et qui opteraient pour le logiciel de TAO de leur choix, dans lequel la TA interviendrait de façon interactive durant — et non avant — la traduction. J'ai dès lors opté pour un logiciel de traduction assistée par ordinateur, qui offre une interface plus conviviale pour la traduction et la post-édition, de même qu'une série d'autres fonctionnalités liées au travail de traitement de texte et de traduction. J'y ai également mis à disposition trois outils que la traductrice pouvait utiliser comme bon lui semble, de manière tout à fait interactive : les précédents tomes de la saga intégrés sous forme de mémoires de traduction, un système de TA générique et le système de TA adapté.

Le choix de l'outil de traduction assistée par ordinateur s'est porté sur *Trados* (RWS Group 2022), qui figure parmi les environnements de TAO les plus répandus sur le marché (cf. Figure 30). S'il en existe beaucoup d'autres, cette décision tient surtout au fait que le logiciel était un des seuls pour lesquels un module d'extension compatible avec *OpenNMT* était disponible, le *plug-in* MTUOC¹³⁶, et permettait par conséquent l'intégration de mon système de TA dans l'environnement de TAO. Ce module fait partie d'un projet plus large du même nom (cf. OLIVER 2020), dont l'objectif vise précisément à faire le lien entre le développement de la TA et son déploiement dans les interfaces de travail professionnelles, en facilitant l'intégration de systèmes neuronaux et statistiques dans les logiciels de traduction assistée par ordinateur. De fait, il peut être compliqué de déployer le système une fois celui-ci mis au point, car il faut pouvoir lancer le moteur de TA sur un serveur et connecter ce serveur à l'interface de TAO, via un module d'extension qui doit être compatible avec l'ensemble des outils impliqués. Or, cette démarche est cruciale pour le type d'évaluation experte qui est mené ici et qui, comme le rappellent VIEIRA *et al.* (2023, p. 217), doit se rapprocher au plus près des environnements de travail réels et prendre en compte leurs effets sur la tâche proposée. Le projet MTUOC permet néanmoins de configurer des serveurs pour plusieurs solutions de TA¹³⁷, dont *OpenNMT*, pouvant ensuite être connectés à l'outil de TAO par l'intermédiaire du *plug-in*. Une fois le serveur paramétré et le script installé dans *Trados*, le système de TA peut alors être lancé localement, sur la même machine faisant tourner le programme de TAO.

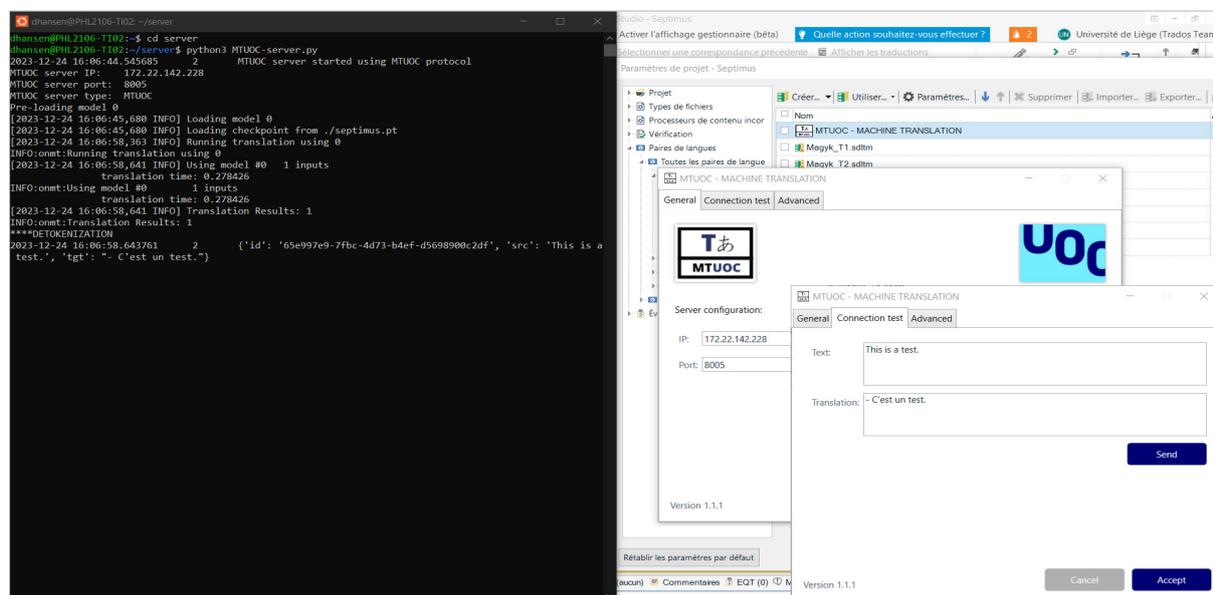


FIGURE 29 – Intégration TA/TAO grâce au module MTUOC

SOURCE : OLIVER (2020) et RWS Group (2022)

136. Cf. <https://github.com/aoliverg/MTUOC-Trados-plugin>.

137. Cf. <https://github.com/aoliverg/MTUOC-server>.

En ce qui concerne le système de TA générique, j'ai finalement arrêté mon choix sur le service proposé par *DeepL* (*DeepL GmbH 2017*), qui peut être lui aussi intégré directement dans *Trados*. Il s'agit également de l'outil le plus couramment utilisé pour la paire anglais-français et le moins susceptible d'avoir été confronté aux romans de la saga *Magyk* durant son entraînement¹³⁸. L'interface, quant à elle, a été paramétrée pour afficher systématiquement une proposition issue des deux systèmes de TA (personnalisé et générique) et pour fournir un accès facile aux mémoires de traduction au cas où Nathalie voulait consulter cette base de données (que ce soit pour chercher des termes spécifiques à la saga ou s'assurer de la cohérence terminologique de termes récurrents), mais aussi pour fournir sur demande un aperçu du texte cible comme il serait affiché dans un programme de traitement de texte.

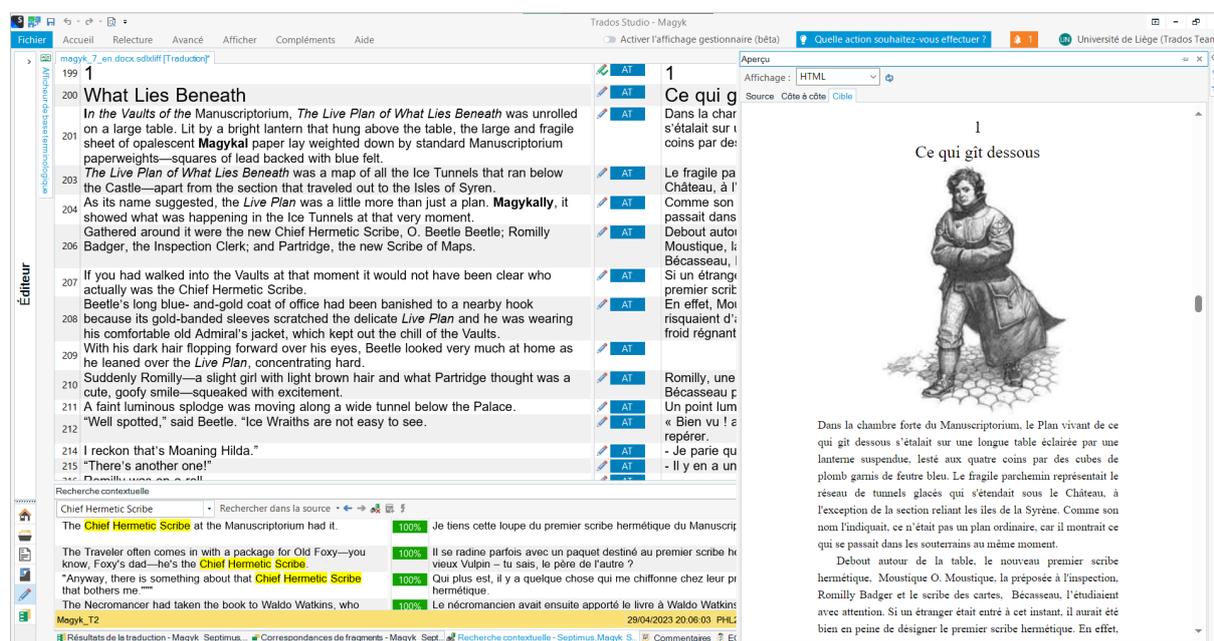


FIGURE 30 – Interface de travail dans *Trados*

SOURCE : RWS Group (2022)

Afin de pallier en partie la difficulté liée à l'interface de TAO, nous avons consacré la matinée à découvrir le fonctionnement de *Trados*, à passer en revue ses diverses fonctionnalités et à paramétrer l'interface. Nous avons notamment configuré le logiciel pour que les propositions issues des ressources (MT et TA) apparaissent sous la fenêtre d'édition et pour qu'une plus grande place soit accordée à cette dernière, permettant ainsi d'afficher un plus grand nombre de segments du texte source en supprimant ou en réduisant d'autres fenêtres. Enfin, à la demande de Nathalie,

138. L'un des principaux problèmes liés à l'utilisation des systèmes de traduction automatique en ligne est que l'on peut difficilement savoir quelles données ont été utilisées pour mettre au point ces systèmes. Or, on sait cependant que *Google Traduction* (*Google 2006*) se base au moins en partie sur des ouvrages de la plateforme de numérisation et de publication *Google Livres* (*Google 2004*), où se trouvent les livres de la série *Magyk* (cf. *SARNO 2010*).

la suggestion du système de TA personnalisé a été paramétrée pour apparaître automatiquement dans la cellule en cours d'édition au moment où elle en commence la traduction.

Après un court essai, Nathalie s'est alors lancée dans la traduction, assistée par la machine, du prologue de *Fyre*, de la même manière qu'elle l'aurait fait pour une traduction habituelle. Je restais évidemment assez proche au cas où d'éventuels problèmes techniques nécessitaient d'improviser une solution, mais aussi pour observer et noter les réflexions qui survenaient. J'avais en effet invité Nathalie à verbaliser autant que possible ce qu'elle voyait et ce qui pouvait lui traverser l'esprit, puisque cette méthode de raisonnement à voix haute représente un excellent moyen d'appréhender ce qui se passe dans l'esprit de la personne lorsqu'elle traduit, d'identifier des stratégies et de repérer des difficultés (KUSSMAUL et TIRKKONEN-CONDIT 1995). Ces réflexions étaient d'ailleurs parfois prolongées lors de quelques pauses et discussions qui entrecoupaient les différentes phases d'exercices. Pour varier celles-ci, la seconde moitié de la tâche s'est portée sur le premier chapitre suivant le prologue, afin de ne pas restreindre l'exercice à un unique aperçu. Enfin, notre journée de travail s'est clôturée par une discussion qui s'est imposée d'elle-même, d'une façon si spontanée à vrai dire que je n'en ai pas eu le réflexe de démarrer immédiatement l'enregistrement. J'en fournis une transcription dans l'Annexe H et m'efforce de revenir sur les prémices de cet entretien ainsi que sur les observations de la journée au fil de l'entretien.

7.3.1 *Un assistant potentiellement utile*

Habituellement, Nathalie travaille sur deux écrans : un premier affichant le texte cible dans une interface de traitement de texte et un moniteur secondaire permettant d'afficher soit la source, soit des recherches. Pour ce qui est des ressources, les invariants incluent des dictionnaires papier qu'elle navigue avec aisance — une compétence affinée sur des années de pratique que soulève également Hans-Christian OESER (2020, p. 19) — de même que les fonctionnalités de correction d'*Antidote* (Druide informatique 1996). Lorsque celles-ci font défaut, la traductrice passe aux recherches en ligne qui lui apportent généralement satisfaction et qui viennent alimenter des feuilles de notes au format électronique ou papier qu'elle conserve tout au long de la traduction.

Mis à part le recours à un outil permettant l'affichage de la source, de la cible et des recherche sur un seul écran, le scénario proposé pour cette évaluation permettait l'utilisation de ces ressources habituelles, auxquelles la traductrice était libre d'ajouter tout ce qui pouvait lui sembler pertinent. Les conditions de son environnement de travail restant dès lors inchangées, cette mise en place constituait un point de départ idéal pour se pencher sur son utilisation des outils de TA et de TAO.

L'exercice aura d'ailleurs permis de relever une multitude de points qui me semblent essentiels si l'on veut aborder l'utilisation des technologies de la traduction en littérature. Dans l'ensemble, il apparaît que la TA pourrait jouer le rôle d'un assistant capable de proposer des suggestions de traduction, comme l'évoquent de la même manière VIEIRA *et al.* (2023, p. 231), et de fournir par ce biais une source d'inspiration pouvant intervenir de manière dynamique au cours de la traduction — par contraste avec un premier jet produit au préalable par la machine. Toutefois, la configuration proposée ici et la manière dont a été intégrée la TA ont parallèlement mis en évidence un processus trop long et une interface trop contraignante pour que ce soit véritablement utile en l'état. En effet, les remarques formulées au cours de ce second entretien confirment dans l'ensemble les évaluations présentées dans le [Chapitre 6](#), bien que la discussion se soit rapidement tournée vers les contraintes liées à l'outil, pour laisser place à ces considérations ergonomiques.

Nathalie, par exemple, a très vite remarqué les différences entre les deux systèmes. À commencer par le fait que l'un tenait visiblement compte du contexte de la *fantasy* et du merveilleux, ainsi que de la terminologie propre à l'ouvrage, et paraissait généralement « plus en phase avec l'univers », là où *DeepL*, entraîné sur plus de données, était « plus généraliste » et proposait des expressions moins pertinentes, mais évitait des contresens et d'autres problèmes de désambiguïsation lexicale (SERVAL 2023a). Si notre entretien confirme globalement mes précédentes observations, cependant, celles-ci incluent de la même façon les limites qui y ont été exposées. De fait, bien qu'elle ait été positivement surprise par les performances de la TA, Nathalie insiste elle aussi à plusieurs reprises sur le ton fortement littéral des propositions des deux systèmes (générique et adapté), non seulement dans notre discussion post-tâche, mais aussi au fil de notre journée de travail.

Il est par ailleurs intéressant de noter que pour la traductrice, la différence entre le système générique et le système personnalisé s'avère finalement avantageuse :

Après, j'ai un avis assez contrasté par rapport aux deux outils, dans le sens où je trouvais qu'ils se complétaient pas mal. D'avoir les deux, ça se complétait pas mal. Parce que parfois le premier, l'outil personnalisé donc, était plus pertinent. Mais en fait *DeepL* se défendait pas mal et évitait même certains contresens. (*Ibid.*)

Là où je conclusais à la difficulté de trouver un bon compromis entre personnalisation et qualité, Nathalie y voit quant à elle deux systèmes complémentaires. Le système idéal devrait dès lors pouvoir combiner les points forts respectifs de ces deux outils (personnalisation et robustesse) pour être utile à lui seul, mais Nathalie revient néanmoins plusieurs fois et met l'accent sur la possibilité de comparer différentes solutions, comme elle le proposait déjà avant cette expérience, lors de notre premier entretien.

De même, que ce soit avant ou après l'expérience, le dispositif idéal qui est décrit est celui d'un outil qui pourrait proposer « une base de réflexion » lorsque l'on « sèche » sur un passage en particulier, tout en sachant qu'il n'y a pas d'obligation à accepter les solutions proposées, mais que celles-ci peuvent être utiles pour lancer des pistes de réflexion (SERVAL 2023a). Ce constat est dressé après que j'ai demandé à Nathalie si elle voyait dans cette démarche un changement méthodologique radical, qu'elle accorde effectivement au cadre expérimental l'incitant à comparer systématiquement toutes les solutions, mais dont la méthode ne correspondrait pas nécessairement à l'utilisation qu'elle en ferait. Nathalie voit plutôt une technologie à utiliser non pas pour faciliter le *flow* dont je parlais avant (cf. Section 7.2.2), mais plutôt à utiliser de manière ponctuelle lorsque l'esprit se heurte à un obstacle et entre dans un état plus analytique, plus créatif :

Dans les moments où je sors un peu de la phase automatique, où la solution vient tout de suite parce que la phrase est super simple ; c'est ça et c'est pas autre chose. Après, quand il y a effectivement une plus grande part de réflexion, où j'ai une hésitation sur la façon dont je vais construire la phrase ou le paragraphe, oui, je trouve intéressant d'avoir déjà une sorte de matière brute à partir de laquelle je vais réfléchir. (*Ibid.*)

Nathalie envisage dès lors un outil qui n'interviendrait pas sur l'ensemble du texte, donc, comme il en était encore question dans le scénario proposé, mais qui agirait comme un assistant à invoquer de manière ponctuelle, pour susciter des idées et amorcer la réflexion :

[Il] y a des moments où je dirais que, paradoxalement, ça libère un peu l'esprit, finalement. Parce que ça lui permet de rebondir et d'explorer des voies dans lesquelles il ne se serait pas forcément engagé, d'autres façons de voir et d'autres façons de faire. C'est une sorte de rêverie libératrice. On a une sensation de ne pas partir de rien et soit ça me plaît, soit ça ne me convient pas du tout, soit on est entre les deux, mais au moins il y a une réflexion qui s'amorce à partir de ce qui est proposé, pour en arriver à une solution personnelle. (*Ibid.*)

Malgré cela, la discussion s'est surtout tournée vers des questions d'ergonomie pour soulever de nombreux problèmes qui sont surtout liés à deux aspects. Le premier est que la TA telle qu'implémentée ici, malgré mes efforts pour proposer un scénario plus interactif, est trop invasive, trop contraignante, et perturbe plus le *flow* de la traduction qu'elle ne joue le rôle d'assistant. À cette interférence de l'interaction humain-machine s'ajoute l'interface, et plus spécifiquement la segmentation, qui empêche d'avoir une réelle vue d'ensemble sur le texte et de modifier librement sa structure.

7.3.2 Une interface lourde et contraignante

Deux aspects pourraient contribuer à cette notion de contrainte, dont le terme finit sans doute par apparaître dans mes questions par analogie avec les recherches existantes (cf. [GUERBEROF-ARENAS et TORAL 2022](#)), mais se dégage sous plusieurs formes au cours de la discussion ([SERVAL 2023a](#)) et reviendra surtout au moment de l'entretien rétrospectif ([SERVAL 2023b](#)) : d'un côté, une traduction est toujours plus lente au départ, le temps de se (re)familiariser avec le texte et de trouver son rythme ; d'un autre côté, il faudrait voir à quel point l'expérience vis-à-vis de la tâche de post-édition peut être un facteur déterminant à cet égard. Toutefois, l'affichage encombré de la TAO (qui a l'avantage d'offrir de nombreuses fonctionnalités utiles au sein d'une même interface, mais qui réduit fortement la place accordée à l'affichage du texte cible et à la zone de travail) et plus particulièrement encore la segmentation (qui « saucissonne » le texte en phrases isolées) font qu'il est difficile de considérer le texte dans sa linéarité et d'en garder une vue d'ensemble, mais aussi de s'affranchir de la segmentation par défaut pour réorganiser celui-ci, comme on le ferait intuitivement dans une interface de traitement de texte (cf. [EHRENSBERGER-DOW et MASSEY 2014](#), p. 64).

La longueur des passages traduits (ou plutôt post-édités) durant cette première interaction avec l'outil n'offre pas assez de matériau pour proposer des analyses quantitatives ou stylométriques suffisamment fiables. Néanmoins, une comparaison attentive du texte source, de la traduction humaine, des sorties de TA et de la version post-éditée va clairement elle aussi dans ce sens. On peut distinguer d'une part des modifications locales, plutôt lexicales (suppressions de répétitions de l'anglais et omissions volontaires, termes plus adéquats, changements préférentiels qui reflètent des choix similaires dans les tomes précédents, expressions idiomatiques, registre plus soutenu...), et d'autre part des modifications effectuées à une échelle plus large, plutôt syntaxiques (traduction généralement plus synthétique, structure plus naturelle que la TA qui colle fortement à la source, divisions et fusions de phrases, changements de focalisation, enchaînements logiques, ajouts renforçant la cohérence textuelle...). Chacune de ces modifications témoigne d'un rapprochement du style propre à Nathalie et d'une prise de liberté qui lui est caractéristique, ainsi que j'en ai discuté précédemment (cf. [Section 5.2.2](#)). Cette prise de liberté se note cependant significativement plus dans la traduction humaine, au regard de laquelle la version post-éditée reste encore proche de la TA. On remarque par exemple de multiples va-et-vient sur des segments distincts au cours de la PE, mais ceux-ci se font presque systématiquement dans le confinement des cellules, là où la traduction humaine de Nathalie déplace plus volontiers des éléments au sein du paragraphe, voire au-delà.

Si cette étude s'appuie sur la prémisse du style personnel des traductaires comme marqueur d'une adaptation de la TA à la leur voix individuelle (cf. [Section 4.4](#)) et que ce critère a permis de démontrer une grande amélioration en ce sens (cf. [Chapitre 6](#)), il apparaît en revanche que la PE ne permet pas d'atteindre un niveau comparable à la traduction humaine dans le scénario proposé. De fait, la TA laisse une trace dans la post-édition qui trahit la voix de la machine — à l'image de l'« effet fantôme » mis en lumière par Perrine SCHUMACHER (2023, p. 233-234) — et qui introduit du « bruit » venant parasiter traduction (TAIVALKOSKI-SHILOV 2019a, p. 697). Ce bilan rejoint celui dressé dans un scénario similaire par KENNY et WINTERS (2020)¹³⁹, qui montrent que si la PE permet au traducteur Hans-Christian Oeser d'affirmer sa voix en effectuant des modifications caractéristiques de son style propre, sa version post-éditée reste éloignée de sa traduction réalisée sans l'aide de la machine.

De façon intéressante, lorsque *Trados* n'a pas fourni de proposition sans raison apparente — ce qui fut le cas pour deux segments — j'ai pu remarquer par ailleurs que Nathalie était manifestement plus rapide, ce qui me fait dire, en accord avec les impressions de Nathalie, que la post-édition classique telle que proposée ici rend le processus plus lent en littérature, du moins dans notre cas. Avec cette mise en place, Nathalie mentionne un autre mode de fonctionnement, plus proche de celui de la relecture, qui lui rappelle le travail à partir d'un brouillon produit par quelqu'un d'autre. Or, comme le mentionne OESER (2020, p. 22), ce n'est pas nécessairement plus simple :

En révisant une traduction existante qui n'est pas un premier jet de son cru, comme une traduction automatique, on se retrouve face à un obstacle qui ne se pose pas lorsqu'on embarque sur une nouvelle traduction sans l'aide de la machine. Un obstacle qui est bien connu dans l'édition : il faut respecter non seulement l'auteur original, mais aussi la voix du traducteur. S'il n'y a aucune obligation à « respecter » les « efforts » de la machine, il faut cependant travailler en tenant compte de deux versions distinctes : l'original et la traduction existante, qui posent chacune leurs propres difficultés [33].

En définitive, il semblerait donc qu'en intervenant de manière systématique, avec l'objectif de faciliter la traduction dans son côté mécanique et automatique, la TA entrave à l'inverse l'intuition des traductaires et bloque le *flow* en empêchant toute interprétation « somatique » du texte source (ROBINSON 1991), ce que SCHUMACHER (2023, p. 334) soutient également en conclusion de sa thèse sur la post-édition en contexte académique. Cette expérience et mes discussions avec Nathalie indiquent que l'outil devrait intervenir au contraire dans le versant analytique de la tâche, lorsque l'on se retrouve confronté à un écueil, et soutenir ainsi la démarche créative.

139. Bien que les chercheuses emploient ici le système générique *DeepL*, dans une interface de traitement de texte, il s'agit ici aussi d'un traducteur expérimenté, sans expérience préalable en post-édition, travaillant à partir d'un texte pré-traduit par la TA.

L'alternative pourrait dès lors être celle d'un système à invoquer ponctuellement à la demande, ainsi que nous venons de le voir. Concernant l'interface de travail, une solution pourrait être de segmenter le texte par paragraphe et non par phrase, ce qui rejoint mes conclusions sur la traduction littéraire assistée par ordinateur (cf. [Section 2.3](#) ; [HANSEN 2024](#)), mais qui rendrait plus difficile l'utilisation de la TA. Nathalie précise en effet qu'elle travaille beaucoup plus au niveau du paragraphe que de la phrase, ce qui lui permet de fusionner des phrases, de jouer sur la structure ou d'intervertir des éléments. L'environnement proposé par le logiciel libre de TAO *OmegaT* ([OmegaT Project 2002](#)) se rapproche un peu plus de ces considérations, dans la mesure où le programme propose une segmentation par paragraphe et une présentation continue du texte qui se rapproche plus de celle d'un logiciel de traitement de texte. On pourrait imaginer, au reste, une intégration de la TA directement dans un environnement de traitement de texte, comme le suggère Nathalie en envisageant une sorte de fenêtre à afficher tel un commentaire en marge du texte, et comme le proposent indirectement Antoni OLIVER ([2017](#)) dans un scénario d'aide à la lecture, ou Dominique DEFERT ([2022](#)) à partir de sa propre expérience du « raccourci magique » de *DeepL*, mobilisé par l'auteur tel une sorte de dictionnaire pour phrases.

Enfin, une autre condition pour utiliser cet outil, au-delà du fait qu'il devrait idéalement s'améliorer encore, est de pouvoir s'assurer de rester maître de l'outil et du rendu final. Cette dernière crainte rejoint des préoccupations visiblement grandissantes, puisque l'entretien a été mené alors que l'Association des traducteurs littéraires de France publiait les résultats d'une enquête pointant l'actualité de la question, le manque de transparence et les risques pour le statut et la rémunération des traducteurs ([ATLF 2022](#)). Si Nathalie reste optimiste pour les années à venir, elle reconnaît que les choses risquent de changer — peut-être plus vraisemblablement pour les nouvelles générations qui se lancent dans le métier — et que les usages actuels dans le monde de l'édition laissent peu de doute à ce sujet.

Je pense que ça va évoluer. [...] Dans quelles proportions, je ne sais pas... Je ne pense pas quand même qu'il y aura un raz-de-marée généralisé, ne serait-ce que parce qu'il y a aussi quand même une génération d'éditeurs qui n'ont pas travaillé comme ça, qui sont encore en place, qui ont encore un petit peu de poids, probablement, face à leur direction, et qui peuvent quand même défendre une certaine manière de travailler. Mais je suis quand même un petit peu inquiète pour les 5 à 10 années à venir, oui. Ça a tellement d'avantages pour les éditeurs que j'ai très peur que la pratique se répande assez vite. ([SERVAL 2023a.](#))

En effet, l'attrait de la rentabilité pose un risque à ces égards, qui s'est déjà vérifié pour les technologies de la traduction précédant la TA et qui se trouve renforcé par un contexte global de pression sur les coûts et sur la productivité ([O'BRIEN 2012](#), p. 110 ; [MOORKENS 2017](#)).

L'implémentation même de la traduction automatique, si elle intervient comme un outil au service de l'humain ou bien comme un premier jet à faire simplement réviser par des traducteurs, pourra donner lieu à des scénarios aux conséquences hautement variables. Dans le second cas, la crainte est surtout celle d'être dépossédée du texte et de ne plus avoir le dernier mot sur la publication finale :

Après tout, bon, si pour des contraintes de temps on est amené à utiliser ce genre d'outil sur une traduction, dès lors que l'on peut quand même avoir la dernière touche sur le texte et être considéré comme auteur, comme on l'est actuellement — donc avec son nom sur la couverture — et puis surtout être rémunéré avec des droits proportionnels, pourquoi pas ? Après, le glissement dangereux, c'est que le traducteur se retrouve face à un premier jet qu'il n'a pas sollicité et soit considéré comme un simple... Qu'il ne soit même plus traducteur, en fait, ni même un préparateur de copies. Qu'il fasse juste de la post-édition en n'étant pas considéré comme auteur du texte, en ayant juste une rémunération forfaitaire, et puis finalement en n'étant même pas assuré que ce soit la version qu'il aura proposée qui soit éditée en l'état. Là, ça n'a plus rien à voir avec ce qu'est la traduction. (SERVAL 2023a.)

Finalement, c'est donc avant tout d'une peur de perdre le statut d'artiste-auteurice qu'il s'agit, avec les effets qu'engendrerait cette perte pour la reconnaissance et la rémunération allant de pair avec ce statut.

Il n'est pourtant pas uniquement question ici de la situation et des conditions de travail des traducteurs, mais bien de l'activité de traduction dans son ensemble, en ce compris la qualité du produit final :

Parmi ce qui contribue à la réussite d'une traduction, on peut distinguer plusieurs facteurs personnels tels que l'investissement dans la tâche, la confiance en ses propres compétences, la fierté professionnelle ainsi que la volonté à assumer la responsabilité de l'acte communicatif [34]. (KUSSMAUL et TIRKKONEN-CONDIT 1995, p. 191.)

La TA, si elle est trop contraignante et sert uniquement à gagner du temps, pourrait jouer négativement sur ces aspects. On le voit d'ailleurs très clairement lorsque Nathalie décrit son activité comme un « artisanat », impliquant un travail patient d'élaboration et de raffinement (SERVAL 2023a). Au cours de nos discussions et de mes échanges avec d'autres spécialistes de la traduction littéraire, il est très clairement ressorti que l'attachement à cette facette du métier et à ces méthodes de travail était un élément récurrent et crucial. Comme le résume Nathalie, l'artisanat suppose l'utilisation d'outils — et, plus encore, d'outils spécifiquement adaptés à la tâche —, mais la crainte est avant tout de voir l'automatisation prendre le pas sur cet aspect artisanal.

Interlude : Belgique

Peu après cet entretien, Nathalie se lançait dans une nouvelle traduction et j'entamais de mon côté la rédaction de ce manuscrit. Si nous devons bientôt nous revoir pour poursuivre ce pan de recherche, que je compte bien continuer d'explorer après la thèse, j'ai également eu l'occasion de revoir Nathalie en ligne en amont de notre prochaine rencontre, afin d'explorer des pistes d'améliorations possibles, d'examiner les différences relevées entre la traduction humaine et la post-édition, et d'engager une discussion plus générale sur cette expérience avec un peu plus de recul.

Nous nous sommes donc revus sept mois après notre second entretien, période qui fut ponctuée par des échanges sur diverses actualités liées au sujet de la traduction littéraire automatique. De fait, avec l'arrivée des grands modèles de langues, la période de rédaction de la thèse aura notamment été marqué par les prises de positions de collectifs divers et d'associations y compris au-delà de la traduction, qui montrent l'importance grandissante du débat aujourd'hui et la prise en compte croissante de ces enjeux.

Notre discussion soulèvera d'ailleurs une fois encore le caractère essentiel des aspects ergonomiques de la TLA dans son sens le plus large, c'est-à-dire autant techniques que sociétaux. Il s'agit de toute évidence d'un sujet qui mériterait bien plus d'attention, mais aussi une approche plus large des technologies de la traduction en littérature. Bien que j'aurai l'occasion d'approfondir ces questions en accueillant prochainement Nathalie à Liège pour un dernier entretien, j'espère également pouvoir pousser plus avant ces réflexions par la suite.

7.4 Troisième phase : modifications ergonomiques et rétrospective

Sept mois après notre expérience de post-édition, j'ai eu l'occasion de revoir Nathalie une dernière fois en ligne, afin de faire le point sur l'ensemble de nos discussions et d'envisager des pistes d'amélioration du dispositif sur la base des difficultés précédemment soulevées. Cette troisième entrevue s'est d'ailleurs rapidement centrée sur ces difficultés, puisque nous avons entre-temps poursuivi nos échanges, au cours desquels la traductrice m'a proposé un retour d'expérience après être retournée elle-même à la version originale des passages sur lesquels nous avons travaillé. En livrant ses impressions a posteriori, elle m'a ainsi fait part du sentiment de décalage qu'elle éprouvait entre le travail accompli et le résultat de la traduction produite :

Pour ne rien te cacher, le rapide travail de post-édition auquel je me suis livrée à partir des premières pages de *Fyre* et des propositions des deux logiciels de traduction assistée m'avait laissé le sentiment diffus d'avoir transpiré pour un résultat très insatisfaisant. Aussi, il y a quelques jours, j'ai eu la curiosité de relire les passages correspondants dans la traduction que j'avais remise à Albin Michel en 2013 [...] Si je n'étais pas l'auteur de cette traduction, mon amour-propre en aurait pris un coup, tant la différence de qualité entre les deux versions est abyssale ! Certes, je ne perds pas de vue que dans un cas, il s'agit d'une traduction aboutie (même si imparfaite, car n'étant jamais passée entre les mains d'une préparatrice de copie et d'une correctrice), ayant subi une, voire deux relectures au terme d'un processus de traduction portant sur la moitié du livre, et dans l'autre, d'une expérience sans enjeu éditorial, une prise en main d'un outil inconnu (et pas très ergonomique, comme nous l'avons dit), dans le but de tester les capacités dudit outil, ce qui m'a conduite à étudier chacune de ses propositions (ce que je ne ferais certainement pas dans le cadre d'une utilisation « normale »). En outre, comme je te l'ai dit au terme de la séance de travail, l'absence de mise en page et le séquençage en phrases empêche d'avoir une vue d'ensemble du texte, ne serait-ce qu'au niveau du paragraphe. Toutefois, la différence de résultat est éclairante ¹⁴⁰.

Au cours de notre troisième entretien, Nathalie m'a donc donné un peu de contexte et de détails sur ce retour. Si elle ne s'est pas livrée à une comparaison mot à mot des deux versions, contrairement à celle dont j'ai pu tirer mes observations dans la section précédente, la traductrice s'est consacrée en revanche à une lecture suivie du rendu final qui est d'autant plus éclairante qu'elle confirme les points déjà soulevés à propos des modifications dépassant le seul cadre de la phrase (par opposition à des modifications plus locales) et notamment le fait que la traduction humaine paraît beaucoup plus naturelle, qu'elle est sujette à un plus grand nombre de

140. Nathalie Serval, communication personnelle, 15 mai 2023.

réorganisations et que l'on y trouve davantage de changements de fonds vis-à-vis de la source, ce qui la rend généralement plus cohérente que la version post-éditée (SERVAL 2023b).

Cette différence, la traductrice l'attribue en premier lieu au découpage des phrases et à l'effort requis pour s'appropriier le texte produit par la machine, qui laisse sa trace dans la traduction (*Ibid.*), reflétant une fois encore les problèmes d'interface et de contrainte sur lesquels je me suis penché (cf. Section 7.3.2). Or, ces difficultés ont été étudiées de manière extensive par la traductologie à mesure que se sont répandus les outils de TAO et la pratique de la PE, puisqu'elles se posent évidemment pour l'ensemble de la profession et non pas uniquement pour les textes littéraires.

La segmentation mise en avant par Nathalie, en premier lieu durant l'exercice et ultérieurement au cours de ce nouvel entretien (SERVAL 2023a; b), est sans aucun doute l'un des aspects de la TAO qui a reçu le plus d'attention jusqu'à présent. L'effet induit par cette présentation particulière du texte est parfaitement résumé par Anthony PYM (2011, p. 2) et par la métaphore qu'il élabore à partir des axes saussuriens de la langue, expliquant que la technologie perturbe la lecture syntagmatique du texte (sa linéarité) et impose une lecture paradigmatique (verticale) de segments isolés :

Regardons maintenant ce qu'il se passe lorsque nous utilisons des environnements de travail qui font intervenir les mémoires de traduction et la traduction automatique. La première chose que l'on remarque est que le texte est segmenté, décomposé en unités qui s'empilent les unes sur les autres. Autrement dit, le texte est décomposé dans une forme paradigmatique, sa linéarité systématiquement interrompue. L'esprit traduisant est dès lors invité à travailler de façon isolée sur chaque segment, à s'assurer de la cohérence terminologique et phraséologique du texte, mais aura plus de mal, dans cet environnement, à assurer sa cohésion syntagmatique [35]. (*Ibid.*, p. 3.)

Au-delà de la segmentation, l'interface même peut être un facteur de contrainte, comme le soulignait Nathalie (SERVAL 2023a). Outre le manque de possibilités de personnalisation et de réorganisation sur lequel nous sommes aussi revenus (SERVAL 2023b), l'aspect « encombré » de l'interface peut entraîner un manque de visibilité, en particulier sur des écrans plus petits, ce que pointe également l'enquête sur les besoins exprimés par les traducteurs en contexte de post-édition réalisée par Joss MOORKENS et Sharon O'BRIEN (2017, p. 116–117 et 122–124). Ce type d'enquête est particulièrement révélateur à ce titre, dans la mesure où la lourdeur de l'affichage lié à l'interface et la segmentation se trouvent précisément être les deux premiers points avancés par les traducteurs comme sources de frustration vis-à-vis de la TAO (O'BRIEN *et al.* 2017, p. 155).

Dans le contexte plus spécifique de la TA, il a également été noté que l’affichage systématique d’une pré-traduction provoquait une rupture du *flow* de la traduction, menant au phénomène de « friction cognitive » appliqué au domaine de la traduction par O’BRIEN (2012, p. 115-116), qui définit celui-ci comme un état dans lequel la technologie entre en opposition avec le raisonnement humain et engendre de la friction au lieu de soutenir celui-ci. On peut rapprocher cette notion de la contrainte exercée par la machine, en fournissant une suggestion de traduction avant que l’esprit n’ait pu se livrer à une première interprétation de la source et à un premier jet de traduction. Ainsi, bien que Nathalie fût confiante dans sa prise de liberté vis-à-vis des sorties de traduction automatique, ayant remanié en profondeur les propositions de la TA (SERVAL 2023a), elle fut tout de même surprise en comparant cette version avec celle de 2013, produite sans l’aide de la TA (SERVAL 2023b). Ce constat vient appuyer l’idée d’un effet de *priming* (biais d’ancrage), que l’on a traditionnellement appliqué à l’interférence produite par la langue source dans les études sur l’acquisition d’une seconde langue, mais que l’on voit de plus en plus souvent appliqué à la traduction automatique (CARL et SCHAEFFER 2017). On retrouve d’ailleurs cette même idée dans les quelques études qui se sont intéressées à la post-édition en contexte littéraire et dont les conclusions vont elles aussi dans ce sens (KENNY et WINTERS 2020 ; WINTERS et KENNY 2024 ; CASTILHO et RESENDE 2022 ; KOLB 2024). Ceci m’incite à repenser la manière dont la TA pourrait être intégrée pour faciliter l’interaction humain-machine, en intervenant par exemple de façon plus périphérique ou d’une manière plus interactive, mais ces quelques réflexions invitent aussi à repenser plus largement les technologies de la traduction dans l’ensemble.

7.4.1 Imaginer d’autres scénarios

La Section 2.3 consacrée à la traduction littéraire assistée par ordinateur se concluait déjà sur l’importance des interfaces de travail, de même que sur le peu de changements apportés en pratique à ces environnements. Ce constat est aussi le point de départ du travail de LAGOUDAKI (2009, p. 17), dont l’argument fait ressortir le développement essentiellement technique des outils de TAO, sans réelle prise en considération des besoins exprimés par les utilisateurs ou de l’utilisabilité des logiciels. Il n’est dès lors pas étonnant de lire que les traducteurs estiment les interfaces traditionnelles de TAO peu adaptées à la post-édition ou que celles-ci peuvent constituer une source de frustration au quotidien (MOORKENS et O’BRIEN 2017 ; O’BRIEN *et al.* 2017). Or, l’utilisabilité de l’outil, perçue ou réelle, joue un rôle crucial sur son adoption ou son rejet (KOSKINEN et RUOKONEN 2017), mais aussi — sans surprise — sur l’aide qu’elle peut apporter aux traducteurs. L’exemple donné par LÄUBLI *et al.* (2022) et exposé en début de chapitre, concernant l’effet de l’affichage vertical ou horizontal des textes source et cible (cf. Section 7.1), constitue

à ce titre un exemple frappant de la nécessité de telles considérations ergonomiques, et une remise en question essentielle de ce qui a pu être tenu pour acquis durant les dizaines d'années de développement de ces logiciels. Si ces remarques portent avant tout sur les environnements de travail, force est de reconnaître qu'il en va de même pour l'évolution de la TA, dans la mesure où les recherches menées en la matière visent à servir le plus grand nombre, sans nécessairement garder à l'esprit l'utilisation qui pourrait en être faite par les traducteurs (VIEIRA et SPECIA 2011), bien que des éléments de TA interactive vont doucement dans cette direction.

Dans ce contexte, la question de l'interface apparaît finalement comme un des enjeux majeurs du confort et de l'efficacité au travail, dont les défauts à ces égards sont d'autant moins surprenants que les environnements de TAO eux-mêmes soulèvent des problèmes ergonomiques et ont peu évolué malgré leur présence depuis les années 1990, ce qu'illustre en particulier la faible compatibilité de ces interfaces avec la post-édition, alors que la tâche est quant à elle répandue depuis le début des années 2000. Au cours de notre précédent entretien, Nathalie évoquait une mise en place idéale qui comprendrait en particulier un outil « plus maniable », avec un espace de travail « plus convivial » et « plus léger d'utilisation », mais aussi la possibilité de personnaliser et de réorganiser l'interface (SERVAL 2023a). Si la plupart des logiciels permettent quelques libertés en la matière, ces options sont souvent limitées ou il peut être laborieux d'en ajuster les paramètres, en particulier pour qui ne serait pas encore familiarisé avec ce type de programme :

[...] l'influence des paramètres par défaut, notamment, ne doit pas être sous-estimée. Il nous faut souligner la nécessité d'expérimenter avec différents paramètres de segmentation, surtout pour les personnes qui s'essaieraient pour la première fois à la traduction littéraire assistée par ordinateur, car cela pourrait vous faire épargner du temps et des efforts. Les outils de TAO n'ont pas spécialement été conçus pour la traduction littéraire, c'est pourquoi il peut être important de s'écarter de leur configuration par défaut pour assurer la prise en charge de ce type de textes [36]. (VIEIRA *et al.* 2023, p. 232.)

Il est en effet nécessaire d'expérimenter avec la configuration des interfaces, voire avec différents environnements de TAO, pour s'assurer un meilleur confort de travail. Dans cette optique, j'ai tenté d'imaginer un scénario sensiblement différent de celui présenté durant la deuxième phase d'entretien et de changer au mieux la mise en place proposée, en fonction des recommandations et commentaires formulés par la traductrice lors de notre discussion précédente.

Au cours de mes recherches sur la TLAO, je me suis appuyé à plusieurs reprises sur des logiciels libres comme exemples de la façon dont on peut repenser les outils existants, d'une manière qui soit plus utile en littérature (HANSEN 2024). Une fois encore, l'un de ces logiciels offrait un point

de départ intéressant pour repenser l’environnement de PE et l’interaction avec la TA, puisque le programme *OmegaT* permet de concevoir un scénario bien plus proche de celui imaginé par Nathalie, avec notamment la possibilité de faire intervenir l’outil sur demande et sur des segments spécifiques, dans une interface qui se rapproche plus du traitement de texte. Nous en sommes arrivés à la configuration suivante, qu’il conviendrait de mettre elle aussi à l’épreuve.

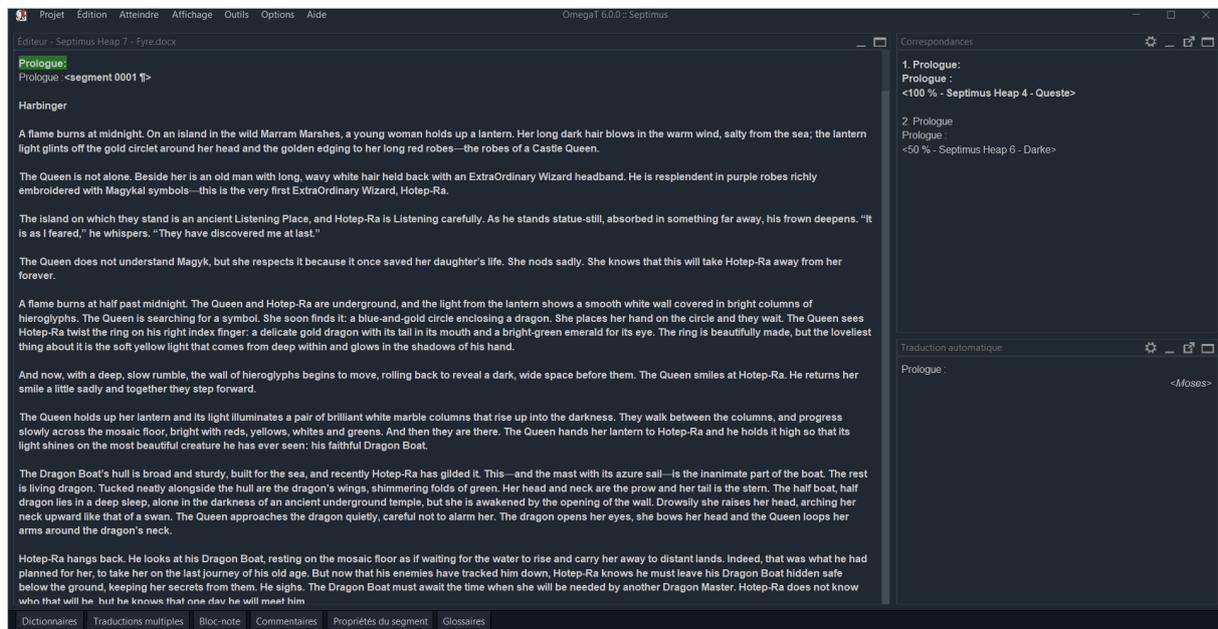


FIGURE 31 – Interface de travail dans *OmegaT*

SOURCE : OmegaT Project (2002)

Cette configuration fait intervenir une segmentation par paragraphe, sollicitée par Nathalie (SERVAL 2023a; b) et recommandée par ailleurs par VIEIRA *et al.* (2023, p. 228–231), de même qu’un texte suivi qui n’est plus restreint par le confinement des cellules. On remarquera également l’interface généralement plus sobre d’*OmegaT*, comportant un espace de travail occupant la majeure partie de l’écran et quelques fenêtres qu’il est extrêmement facile de déplacer, redimensionner ou même masquer, de même qu’un mode sombre qui n’est pas imposé mais qui a amélioré mon confort personnel lors de la configuration de l’outil. Dans cet environnement, les suggestions de traduction automatique sont périphériques, au sens littéral du terme, et apparaissent sur le côté, au même titre que les mémoires de traduction. Ni les MT ni la TA ne sont utilisées par défaut (le segment en cours d’édition est vide), mais les propositions de l’un ou l’autre outil peuvent être facilement intégrées au segment courant. Enfin, le serveur MTUOC a été une fois encore d’une aide précieuse, puisqu’il permet ici aussi d’intégrer le système *OpenNMT* dans *OmegaT*, en faisant passer celui-ci pour un système *Moses* (KOEHN *et al.* 2007) pour lequel un autre *plug-in* d’intégration est disponible¹⁴¹.

141. Cf. <https://github.com/omegat-org/moses-plugin>.

De cette manière, peut-être, les diverses contraintes exercées par les outils pourraient être atténuées, ce qu'il me faudra confirmer à l'avenir. Entre-temps, mon dernier entretien avec la traductrice nous a menés à revenir sur les avantages de la TLA, s'il était possible de compter sur un outil parfaitement idéal. Parmi les points évoqués, Nathalie mentionne en particulier la réduction de l'effort de frappe, qui pourrait permettre de soulager les mains et d'éviter des problèmes d'articulation, et revient sur l'idée d'un collègue à qui l'on pourrait faire appel en permanence pour fournir une piste de traduction ou demander un premier jet de quelques phrases en cas de blocage devant une page blanche (SERVAL 2023b). Enfin, Nathalie imagine la possibilité de faire appel à un tel outil dans un cas où les délais sont extrêmement serrés, puisque l'on sait que les traductions sont toujours à rendre « pour la veille » et que cela l'a déjà amenée à travailler sur une traduction à quatre mains (cf. SERVAL 2023a). Un outil qui permettrait de se concentrer sur la réécriture, les aspects créatifs et les conventions du genre, sur la cohésion interne du récit et sur sa forme finale (SERVAL 2023b), ainsi que l'envisage par ailleurs Susan Pickford (cf. HANSEN et PICKFORD 2022). On retrouve cette même idée chez Jean-Yves Pellegrin, cité dans l'étude conduite par Laurent BESACIER (2014). Selon le traducteur, cependant, cela nécessiterait de pouvoir avoir une confiance presque aveugle dans les solutions fournies par la machine. Or, comme le soulignaient les conclusions du chapitre précédent, les systèmes de TA neuronale posent des problèmes plus prononcés que les paradigmes précédents en matière d'adéquation par rapport à la source, dont témoigne le nombre élevé d'omissions, de cas de sous-traduction, de glissements, de non-sens et de contresens (cf. Section 6.2.2).

De fait, ma conversation avec Nathalie pointera à nouveau les limites actuelles des systèmes (SERVAL 2023b), dont elle me livrait une nouvelle appréciation dans les jours qui suivaient notre expérience :

Ce que j'en déduis, c'est ce que ces outils sont utiles ponctuellement pour vérifier le sens d'un mot, quand le cerveau « patine » devant une phrase qu'il ne sait pas par quel bout attraper, ou tellement banale qu'il en perd subitement ses moyens, habitué qu'il est à chercher midi à quatorze heures [...] Mais en aucun cas — du moins pour le moment — ils ne peuvent se substituer au travail ingrat de déchiffrement, dosage, assemblage, démontage, remontage, polissage... auquel se livre le même cerveau de traducteur, parfois jusqu'à l'épuisement. C'est heureux pour nous, même si, comme nous l'avons redit, il y a fort à parier que le recours à la post-édition va se développer, au moins dans certains domaines (ouvrages pratiques, BD, audiovisuel, traductions de JDR [jeux de rôle] ou de jeux vidéo), pour des raisons économiques¹⁴².

142. Nathalie Serval, communication personnelle, 15 mai 2023.

Si l'efficacité de la TA en littérature est encore limitée, mes discussions avec la traductrice et certaines des conversations tenues dans des cercles de traduction littéraire dont elle m'a fait part laisse entrevoir une utilisation grandissante de la traduction automatique dans le domaine — puisque plusieurs demandes émanant de grandes maisons d'édition vont déjà dans ce sens (SERVAL 2022 ; 2023b). Dans ces conditions, il semble impératif de pouvoir s'assurer que la TA intervienne comme un outil au service des traducteurs, d'une manière qui améliore leur confort de travail, mais surtout de montrer que son utilité est entièrement situationnelle, dans la mesure où elle dépend du texte source, de la paire de langues, du genre littéraire ainsi que des habitudes de travail des traducteurs.

7.4.2 *Améliorer ses conditions de travail*

Le risque principal posé par la traduction automatique est avant tout celui d'une pratique imposée de façon systématique, telle qu'elle s'est instituée chez de nombreux prestataires pour d'autres domaines. Un regard sur les pratiques actuelles dans ces domaines laisse effectivement entrevoir certaines conséquences d'une éventuelle introduction de la TA en littérature, c'est pourquoi Nathalie s'est notamment tournée vers une cousine, traductrice également, dont l'expérience en post-édition pouvait éclairer nos conversations. Laurence Le Maire, traductrice technique, livre à ce sujet un témoignage pertinent qui montre que l'outil peut constituer à la fois un intérêt financier ou une perte de temps et d'argent, car la rémunération — très faible — est la même dans tous les cas, bien que l'utilité de la post-édition soit quant à elle très variable et puisse parfois prendre plus de temps qu'une traduction sans utilisation de la TA lorsque l'outil est imposé pour un texte qui ne s'y prête pas. Commentant certains des projets qui lui sont confiés par l'un de ces prestataires, elle explique ainsi :

Quasiment tous leurs projets (essentiellement des sites web, des applis, du marketing) sont d'abord traduits automatiquement et le boulot des traducteurs consiste à relire, assurer la cohérence avec les glossaires et les mémoires de traduction. L'efficacité dépend en TRÈS grande partie des personnes qui remplissent les glossaires au fur et à mesure (car ce sont des projets sur lesquels on travaille en continu). Si ces chefs de projet sont efficaces, ça va tout seul, sinon tu passes ta vie à corriger et, au final, ça prend autant de temps qu'une traduction complète. Il y a quelques bugs récurrents, notamment pour la traduction de choses comme *it* et *they* mais globalement c'est parfait pour des textes répétitifs (des textes techniques ou des rubriques d'aide, par exemple) car, comme on travaille en équipe, ça assure une cohérence au niveau du style. Pas intéressant du tout, certes, mais on va dire que c'est de l'argent qui rentre tous les mois sans se prendre la tête. En revanche, pour les textes marketing, c'est une catastrophe. Comme c'est

traduit mot à mot et que le traducteur automatique n'a aucun sens du rythme ou de la formule, le résultat tombe toujours à plat et le texte final est toujours trop long, donc il faut tout refaire en étant payé des cacahuètes. L'autre agence avec laquelle je travaille [...] caress[ait] l'idée de se lancer aussi dans la post-édition. Mais apparemment ils y ont renoncé. De toute façon, pour leurs documents (essentiellement des guides et des rapports d'investissement) ça ne marcherait pas selon moi. La finance, c'est déjà pas le truc le plus humain qui soit mais si, en plus, on devait se passer des jeux de mots et des petites boutades des banquiers... Et puis là aussi, il y a régulièrement des contraintes d'espace pour leurs documents et je ne sais pas si l'intelligence artificielle est aujourd'hui capable de synthétiser des idées sans trop s'éloigner de l'original¹⁴³.

Ces conversations ont d'ailleurs eu lieu au cours d'une période marquée par la publication de divers manifestes, issus d'associations de traduction littéraire et d'autres secteurs, visant à prévenir les risques liés à l'utilisation de la TA (cf. [Chapitre 8](#)). Face à ces enjeux, Nathalie m'a fait part d'un sentiment partagé entre, d'un côté, le souci de préserver les conditions de travail des traducteurs et d'en prévenir la dégradation par la banalisation des pratiques de post-édition, mais aussi, de l'autre, la possibilité même de limiter l'utilisation de la TA ou d'autres outils issus des recherches en intelligence artificielle, en ce compris une utilisation créative par les artistes mêmes ([SERVAL 2023b](#)). Elle fait ainsi le lien entre les expérimentations oulipiennes (cf. [Section 4.1](#)) et les dernières générations de modèles de langue, se disant réticente face à un refus catégorique de la TA, en particulier s'il s'agit d'un outil qu'il est possible de se réapproprier pour travailler dans de meilleures conditions (*Ibid.*). Malgré cela, la traductrice évoque l'ensemble des débats générés par la thématique pour insister sur le besoin de veiller à la manière dont l'outil serait utilisé, à protéger le statut des traducteurs et leur droit sur les œuvres, concluant ce dernier entretien sur les mêmes mises en garde soulevées à l'issue de notre rencontre précédente ([SERVAL 2023a](#)).

Dans l'ensemble, mes discussions successives avec la traductrice, motivées par la mise à l'essai du système personnalisé développé dans le contexte de cette recherche mais aussi par la présence grandissante des débats sur la traduction littéraire automatique dans la sphère publique, offrent un premier aperçu de l'image complexe de la TLA que reflétera pareillement le prochain chapitre. La TA, en particulier, se révèle comme un outil qui pourrait autant améliorer que détériorer les conditions de travail des traducteurs et la qualité des traductions, selon son utilisation. Les réflexions autour de la TAO font apparaître de manière plus générale de réelles possibilités d'amélioration du confort de travail, mais aussi des contraintes qui illustrent la nécessité des réimpliquer

143. Laurence Le Maire, communication personnelle, 15 mai 2023.

les utilisataires dans la recherche et le développement de ces systèmes « socio-techniques », dont les avancées ont été majoritairement guidées par une perspective purement technique jusqu'à présent (LAGOUDAKI 2009, p. 17–20; KENNY 2017, p. 1; BENDER *et al.* 2021, p. 618). Une meilleure coopération serait un effet un moyen de maximiser l'utilité et l'utilisabilité des outils, mais faciliterait également leur adoption et donnerait l'assurance que leur intégration se fasse au bénéfice des personnes qui les utiliseront au quotidien.

7.5 Redéfinir l'utilité des technologies de la traduction ?

Au terme de ces trois phases d'entretien, il me semble être en mesure à présent de répondre à la question posée en début de chapitre (QR2) : La traduction automatique est-elle utile pour les traducteurs littéraires ? En effet, si mes discussions successives avec Nathalie Serval semblent pointer vers des éléments de réponse qui vont dans cette direction, à l'instar des conclusions tirées par VIEIRA *et al.* (2023, p. 230–231), et laissent entrevoir une réelle utilité de la TA (que ce soit pour améliorer le confort de frappe, pour fournir une base de réflexion lorsque l'inspiration ne vient pas directement ou peut-être même pour accélérer le processus de traduction par endroits de manière à pouvoir accorder plus de temps à d'autres passages critiques), l'interaction avec l'outil montre aussi clairement que cette question est entièrement dépendante de sa mise en place. Il apparaît au contraire que celle-ci contraint l'intuition de la traductrice et interrompt le *flow* de la traduction dans le scénario envisagé, et ce, malgré les efforts consentis pour rendre l'interface de travail aussi accessible que possible et me rapprocher au plus près de conditions de travail réelles de Nathalie.

Les deux écueils majeurs mis en avant par ce scénario sont sans aucun doute l'*environnement* dans lequel est réalisée la post-édition, ainsi que l'*interférence* causée par la machine sur le processus de traduction, que rapportent pareillement MOORKENS *et al.* (2018). Ces facteurs peuvent être mis en lien avec deux types de charges cognitives que peuvent induire les technologies de la traduction, à savoir une charge « externe », liée aux outils, et une charge « intrinsèque », liée aux processus cognitifs que fait intervenir l'acte de traduction (O'BRIEN *et al.* 2017, p. 147). Dans le premier cas, les deux problèmes principaux sont la lourdeur de l'affichage dans l'interface et la segmentation mises en avant par Nathalie durant notre entretien suivant la tâche d'évaluation (SERVAL 2023a). L'effet de *priming* exercé par la TA est représentatif du second cas, en ce sens qu'il contraint le processus de traduction et conduit à des solutions qui restent proches des sorties de traduction automatique et s'écartent du style habituel de la traductrice, comme j'ai pu le

relever en comparant les deux versions (post-éditée et non post-éditée) produites par Nathalie, qui a soulevé elle-même ce point dès le début de notre dernier entretien (SERVAL 2023b). Dans ces deux cas, les obstacles rencontrés mènent à une augmentation de la charge cognitive, qui peut donner lieu à son tour à un phénomène de « friction » cognitive venant perturber le *flow* et rendant parallèlement la tâche moins agréable (O'BRIEN *et al.* 2017, p. 146).

Cette conclusion vient ainsi apporter une nuance importante aux questions de recherche précédentes (QR1 et SQR1), qui mettent toutes deux en évidence une amélioration entièrement positive de la personnalisation du système de TA d'un point de vue technique. Les problèmes évoqués tout au long de ce chapitre ne sont bien sûr pas nécessairement liés au processus d'adaptation de l'outil et valent tout autant — si pas davantage — pour des outils génériques, montrant au mieux que l'ensemble de la profession bénéficierait d'amélioration à ces égards, mais ils rejoignent les deux principales conclusions tirées de mes travaux sur la TLAO (HANSEN 2017 ; 2024). La première est qu'il convient de se demander si cette « utilité » qui est au cœur de ma question pourrait constituer justement le point de divergence du cas littéraire et d'autres domaines créatifs vis-à-vis d'autres secteurs de la profession, dans la mesure où celle-ci aurait moins trait à des questions de rapidité et d'efficacité, matérialisées par une solution unique toute faite, mais plutôt à des moyens de soutenir la créativité, en favorisant par exemple la possibilité de comparer différentes solutions, comme le laissent également entendre LONG (2022, p. 737) ou VIEIRA *et al.* (2023, p. 231), et comme le prévoyait BESACIER (2014) dans son expérience.

La seconde conclusion, qui me semble tout aussi valable pour la post-édition, est qu'il y a un grand besoin d'outils plus adaptés aux tâches proposées. Étant donné le peu de changements apportés aux outils de TAO en plusieurs dizaines d'années d'existence, ce constat n'a rien d'étonnant, mais l'enjeu est d'autant plus important si l'on veut prendre véritablement en compte les particularités des différents secteurs de la traduction et si l'on garde à l'esprit que le quotidien des traducteurs et leurs besoins se sont quant à eux complexifiés entre-temps. Le système de traduction automatique proposé au cours de l'exercice réalisé dans *Trados* était effectivement « adapté », mais il me paraît essentiel que l'interface de travail et l'intégration de la TA dans l'environnement dans TAO, qui reflétaient un scénario de post-édition on ne peut plus typique, le soient également. Quelques pistes de solution ont été envisagées dans ce chapitre, à commencer par une segmentation au niveau du paragraphe que soulèvent à nouveau VIEIRA *et al.* (2023, p. 231). Les modalités d'interaction avec la TA mériteraient elles aussi des recherches plus approfondies. À ce propos, l'étude menée ici semble aller à l'encontre d'une pré-traduction complète par la machine, sous la

forme d'un premier jet à faire réviser par les traductaires. De fait, celle-ci entrave l'intuition et la créativité de la traductrice, qui insiste à plusieurs reprises sur la possibilité de solliciter l'outil sur demande, comme base de réflexion ou de comparaison lorsque l'esprit a déjà généré une solution de traduction ou bute sur un problème particulier. Envisagée de la sorte, la TA constituerait finalement une ressource supplémentaire ; une « inscription » supplémentaire du texte (VENUTI 2017 ; YUDDALE et ROTHWELL 2022) qui en révélerait d'autres aspects et pousserait plus loin son analyse (*Ibid.* ; LONG 2022, p. 738), s'ajoutant aux ressources existantes, sans faire passer celles-ci au second plan, et permettant surtout aux traductaires de rester maîtres du processus comme de l'outil, exactement comme le propose Andrew ROTHWELL (2024).

Dans l'ensemble, et bien qu'il s'agissait d'un cadre expérimental et d'une expérience nouvelle pour Nathalie, ces trois d'entretiens ont ainsi permis de développer des réflexions cruciales concernant la mise en place des outils et leurs effets sur le processus traduisant. Il ne faut pas perdre de vue en effet que l'outil vient s'insérer dans une dynamique déjà complexe et rodée qui fait intervenir de nombreuses ressources papier et électroniques de recherche linguistique et documentaire, de correction et de prise de note, auxquelles TA et TAO doivent s'intégrer le plus simplement possible. Cela concerne également les mémoires de traduction, par exemple, qui étaient vues dans cette expérience comme quelque chose d'entièrement positif (SERVAL 2023a), mais qui nécessitent de changer ses habitudes et de s'adapter à une méthode de travail différente.

Les pistes d'améliorations proposées dans cette section par l'intermédiaire d'*OmegaT* pourraient aider à ces égards, mais devront impérativement être mises à l'épreuve elles aussi, ce que j'entends pouvoir réaliser très prochainement en continuation de ce travail. En attendant, chacune de ces discussions a laissé paraître en outre des enjeux ergonomiques qui ne sont pas uniquement liés à des considérations *physiques* (relatives à l'interface ou à l'effort de frappe par exemple) ou *cognitives* (relatives aux processus mentaux impliqués par la traduction), mais aussi *organisationnelles*, puisque les traductaires forment un chaînon au sein d'un réseau qui inclut de nombreux autres maillons (maisons d'édition, clientèle, prestataires de service, collègues) dont la technologie est susceptible d'affecter les relations, et qui peut avoir en retour des effets très concrets sur le métier et la pratique de la traduction (RISKU et WINDHAGER 2013). À ce titre, il me reste à présent à répondre à une dernière question (SQR2), dont les réponses seront guidées par les réflexions de ce chapitre : Quelles sont les conséquences de l'arrivée de la TA dans le domaine littéraire ?

Discussion

Après avoir mis en évidence la possibilité et l'intérêt d'adapter les systèmes de traduction automatique (TA) au niveau de l'humain, montrant que ceux-ci pouvaient se rapprocher du style des traducteurs littéraires, j'ai pu apporter, au cours du chapitre précédent, des éléments de nuance essentiels quant à l'utilité réelle de tels outils personnalisés une fois déployés sur le terrain (QR2), soulignant ainsi l'importance de la validité écologique des évaluations et les progrès qu'il reste à faire en matière d'ergonomie, bien qu'il soit possible de repenser l'usage de ces technologies. De fait, cette troisième partie d'ouvrage démarre de la nécessité de considérer la traduction comme une activité incarnée et située, sans laquelle ces limites auraient pu être entièrement passées sous silence. Les aspects organisationnels liés à cette seconde perspective revêtent en effet un rôle tout aussi pertinent lorsque l'on se penche sur l'utilisation de dispositifs technologiques et ont déjà été abordés en partie à ce titre. Néanmoins, le présent chapitre vise à examiner plus largement le contexte social de la traduction littéraire, afin de mieux saisir ce que laisse présager l'arrivée des outils de TA dans le secteur littéraire et d'éclaircir ainsi ma seconde sous-question de recherche (SQR2). « En d'autres termes, il convient de prendre en compte les "acteurs-réseaux" dans lequel les activités professionnelles complexes se négocient [37] » (RISKU 2010, p. 105).

Il y a peu de temps encore, il n'y aurait peut-être pas eu grand-chose à développer, excepté quelques mises en garde et recommandations tirées de rapprochement entre traduction automatique et traduction assistée par ordinateur (TAO). Or, les discussions menées avec Nathalie Serval lors de nos entretiens laissent penser que le sujet de la TA littéraire n'est plus si anodin. Plus encore, à en juger par les échos grandissants, la multiplication des études dans le domaine, l'arrivée de prestataires spécialisés dans la post-édition (PE) littéraire et l'intérêt visible des géants du numérique, il ne serait pas surprenant que des systèmes adaptés à la littérature apparaissent très prochainement dans ces discussions. Il devient dès lors de plus en plus difficile d'anticiper les enjeux de l'automatisation en littérature — c'était là l'un des objectifs principaux de cette recherche —, mais il n'en est que plus impératif d'élargir la discussion sur l'arrivée de ces outils. Par conséquent, je jetterai un premier coup d'œil sur les usages actuels de la traduction littéraire automatique (TLA) avant de m'étendre sur les risques et avantages potentiels qui y sont associés.

J'évoquerai ensuite quelques-uns des défis du domaine, notamment écologiques et juridiques, dans le contexte plus large de ce que l'on désigne aujourd'hui volontiers comme *intelligence artificielle* (IA). Enfin, je tracerai un bref parallèle avec d'autres secteurs confrontés eux aussi aux mêmes questionnements, avant de conclure sur le changement de paradigme que je défendais déjà pour la traduction littéraire assistée par ordinateur (TLAO) et pour lequel la méthode d'adaptation de la TA au niveau de l'humain occupe une place centrale.

8.1 La traduction littéraire automatique aujourd'hui : usages et enjeux

Si mes expériences de TLA et TLAO partent de l'idée que les outils n'imposent pas d'utilisation particulière et peuvent être détournés ou repensés d'une façon qui vise à soutenir plutôt qu'à contraindre la prise de décision, la créativité et les aspects proprement humains de la traduction, c'est toutefois sans tomber dans le déterminisme technologique, qui voudrait que la technologie évolue de manière neutre et surtout inévitable, laissant à la société le libre choix de l'utiliser ou non (HABERMAS 1973 ; TEHRANIAN 1990 ; MOROZOV 2013 ; KENNY 2017). Au contraire, son contexte d'utilisation tout comme sa conception sont indissociables des choix et des motivations qui ont mené à son développement, du public visé ou des investissements — massifs — qui ont rendu possible son apparition (GUERBEROF-ARENAS et MOORKENS 2023, p. 114), en accord avec l'idée que « la technologie n'est ni bonne, ni mauvaise, ni neutre [38] » (KRANZBERG 1986, p. 545). Dans le cas des technologies de la traduction, ce sont les grandes entreprises qui ont pris le pas sur les universitaires (LARSONNEUR 2021, p. 267), et ce sont donc ces entreprises et leurs intérêts qui guident le progrès et les modalités d'utilisation de la TA, sans réelle prise en compte des autres parties prenantes concernées (GUERBEROF-ARENAS et MOORKENS 2023, p. 114–115). C'est probablement pour cette raison que l'adoption de la technologie va de pair avec une réduction continue des coûts, et ce, malgré la croissance constante de l'industrie des langues (DOHERTY 2016, p. 249). Il est intéressant de voir à ce titre que les problèmes qui s'appliquaient déjà à la traduction automatique statistique, en particulier la réutilisation des données qui est mentionnée par KENNY (2012) et dont je reparlerai par la suite, sont toujours d'actualité. À l'inverse, je m'avancerais même à dire que ceux-ci se sont amplifiés et que les solutions mises en avant par le monde de la recherche n'ont trouvé que peu d'échos en pratique, peut-être précisément parce que la TA est longtemps resté l'apanage de l'industrie, loin des préoccupations des formations et associations de traduction. Pour cette même raison, il me

semblait important de proposer un cas concret d'utilisation de la TA en littérature (où la mention de la TA était systématiquement écartée jusqu'à il y a peu), tout comme l'on fait un nombre croissant de collègues maintenant, afin d'amorcer des discussions plus que jamais actuelles.

Parmi les griefs qui ont traditionnellement été avancés à l'encontre des technologies, on retrouve des problématiques transversales, dont Joss MOORKENS et Dave LEWIS (2019, p. 5–6) dressent une liste déjà indicative :

Les traducteurs spécialisés qui avaient investi dans les outils de TAO en s'attendant à y trouver un avantage compétitif ont finalement remarqué que leur capacité de production augmentait à mesure que le prix au mot diminuait. Beaucoup ont commencé à éprouver du ressentiment pour ces nouvelles technologies qui semblaient leur être imposées (WAY 2013 ; CHRISTENSEN et SCHJOLDAGER 2016 ; CADWELL *et al.* 2018) : tout d'abord les mémoires de traduction, qui se sont accompagnées de réductions des tarifs, et plus récemment la post-édition de traduction automatique, qui impose un rabais supplémentaire pour un travail consistant à corriger « des erreurs linguistiques de base qu'un traducteur humain qualifié ne produirait pratiquement jamais » (O'BRIEN 2012, p. 110). [...]

Cette restriction graduelle du rôle de traducteur a miné leur capacité à se conformer au code déontologique de leur profession (CHESTERMAN 2001, p. 142), en réduisant le processus de traduction à une série d'« exercices de substitution linguistique » (PYM 2003). De cette façon, au lieu d'éliminer les tâches « mécaniques et répétitives » (KAY 1997, p. 1 ; LEBLANC 2013), la traduction s'inscrit dans le récit présenté par FRISCHMANN et SELINGER (2018), dans lequel l'intrusion de l'apprentissage machine au sein de nombreux métiers pose le risque de réduire le rôle de l'humain à celui d'une simple machine. La perte d'agentivité en traduction se reflète en outre dans la dépossession graduelle des données avec le passage de l'analogique au numérique, à l'image des nombreuses autres professions intellectuelles qui font face à cette évolution généralement libre de toute contrainte juridique (HUWS 2004, p. 4). [39]

Dans ce contexte de dépossession progressive des outils et des données (LARSONNEUR 2023, p. 97–98), les traducteurs ont dû ainsi s'adapter, malgré les retombées sur la qualité des traductions, leur identité professionnelle, leur autonomie et leur satisfaction au travail (BOWKER 2021).

Face à ces enjeux, on peut aisément comprendre que de nombreuses voix se lèvent pour prévenir des dérives et des risques d'une utilisation aveugle de la technologie dans le secteur littéraire, mais aussi pourquoi certaines personnes soutiennent que la traduction littéraire n'est pas « un problème à résoudre » (STAA 2022). Cette discussion s'étend bien au-delà de la traduction et implique des réflexions qui dépassent la seule question des conditions de travail, comme je le détaillerai

dans la seconde partie de ce chapitre, amenant par ailleurs d'autres professions artistiques à s'interroger sur l'intérêt de rejeter des tonnes de CO₂ pour remplacer des métiers que les gens prennent le plus grand plaisir à exercer (MARX 2023). Partant de ce climat de débat actuel, il me semble utile de revenir sur les avantages possibles d'une utilisation de la TA en littérature ainsi que sur les difficultés qui motivent ces arguments, juste après avoir recontextualisé les usages actuels à partir du peu d'informations dont nous disposons à ce jour.

8.1.1 Premiers regards sur les usages actuels de la TA en littérature

Une chose est sûre, comme nous allons le voir, la traduction automatique est déjà d'application pour la traduction littéraire, à des degrés divers, mais il est difficile de poser un regard concret sur ces usages en raison du véritable tabou qui pèse sur la question. Une partie de ces pratiques est liée à la diffusion des œuvres sur le Web, au sujet de laquelle Eoin Ó MURCHÚ (2019, p. 20) et Kristiina TAIVALKOSKI-SHILOV (2019a, p. 689) notent que des traductions automatiques non révisées circulent déjà en ligne. Dans un autre registre, Shuyin ZHANG (2024) trace un parallèle entre la montée de la TA et l'exportation croissante des fictions Web, dont elle soulève le cas du Xianxia, où se mêlent traduction automatique, traduction professionnelle et traduction amatrice par les utilisataires, tout en mentionnant l'exemple d'une plateforme spécialisée dans la diffusion de ce type de littérature électronique par le biais de la TA. Ce cas d'étude est d'ailleurs le même qui a dernièrement servi de point de départ à l'entreprise Tencent pour lancer une campagne de recherche et d'évaluation sur la traduction littéraire automatique (cf. Section 4.3.3).

Pour ce qui est de la chaîne éditoriale traditionnelle et de la traduction professionnelle, le manque de discussion associé au sujet fait que les remarques suivantes se basent avant tout sur mes observations, relevées sur le terrain, lors de déplacements liés à la recherche ou dans mon réseau. D'une part, celles-ci vont à l'encontre d'un rejet total de la technologie par les traducteurs littéraires, puisqu'elles attestent d'une utilisation de la TA comme aide ponctuelle aux côtés des dictionnaires et autres ressources linguistiques. Une utilisation qui s'écarte donc du but premier la TA, mais qui rejoint un constat similaire dressé par Stephen SLESSOR (2020, p. 9) et Paola RUFFO (2021, p. 200) pour une minorité de la profession. D'autre part, on peut entendre de plus en plus de témoignages de traducteurs à qui l'on propose désormais de la post-édition, ce que Nathalie SERVAL (2022) me confirmait lors de notre entretien préliminaire selon ses propres réseaux, citant l'exemple d'un traducteur français connu à qui une grande maison d'édition avait proposé la PE comme solution possible face à des délais plus serrés encore qu'ils ne le sont déjà habituellement.

Commentant les suites de cette expérience, elle me racontait que le traducteur en question avait globalement été étonné par le niveau de la TA, mais que cela lui avait finalement demandé plus de travail qu'une traduction humaine (TH) et ôté une grande partie du plaisir de traduire, exactement comme l'a montré l'expérience menée dans ce travail (cf. [Chapitre 6](#)). Lors de notre second entretien, Nathalie SERVAL (2023a) me faisait également part de la possibilité que cette pratique se répande, tout comme son acceptation par les traductaires, à la suite de la pandémie de Covid-19, de la crise du papier et de la montée de l'inflation, qui se sont succédées dans un climat de baisse des ventes généralisé. La traductrice évoque également les nombreux avantages perçus pour l'industrie de l'édition, qui ne devraient pas manquer de contribuer à l'adoption de ce modèle, en particulier pour les grands groupes dont la ligne éditoriale est dictée par les actionnaires et privilégie la rentabilité (*Ibid.*). À ceci s'ajoute le témoignage livré par Gaëlle DIFFERT (2023), qui témoigne de la présence bien établie du modèle de post-édition, y compris dans de petits groupes d'édition qui se disent forcés de recourir à la TA pour rester compétitifs (cf. p. 20). Enfin, si l'on pourrait penser qu'il s'agit de cas isolés ou particuliers, l'étude de MACKEN *et al.* (2022), menée en collaboration directe avec une agence spécialisée dans la PE littéraire, confirme bel et bien que la TLA est devenue une question d'actualité et non plus d'anticipation.

Face à ces données parcellaires, l'enquête récemment menée par l'ATLF (2022) sur la post-édition en contexte littéraire apporte une contribution précieuse, qui, quoique limitée elle aussi, représente à ma connaissance le premier et le seul bilan dressé sur le sujet. Toutes les données présentées ici sont tirées de ce rapport, duquel l'association dégage trois principales observations : le manque de transparence sur l'usage de la TA, la baisse de la rémunération et de la reconnaissance des traductaires, ainsi que l'absence d'un gain de temps. Sur les 476 personnes ayant répondu à l'enquête, qui fait état de profils variés, seulement 17 (6 %) rapportent avoir reçu une formation à la TA ou à la PE, pointant un manque de formation qui est intéressant en lui-même, bien que 67 % n'envisagent pas de s'y former et que 27 % envisagent « peut-être » de s'y former, par curiosité. Sur les 14 % à qui l'on a proposé une tâche de post-édition, deux tiers ont accepté, pour des textes fonctionnels dans la moitié des cas. Une moitié de ces réponses concernées par la pratique de la PE indique ne pas envisager de renouveler l'expérience, trouvant la post-édition moins efficace et moins créative, là où l'autre moitié l'envisagerait en fonction du type de texte et des conditions de travail, une majorité de ces réponses l'ayant trouvé peu rentable dans tous les cas (75 %).

Le problème le plus important mis en lumière par cette enquête est probablement la description de la tâche confiée, qui est présentée comme une post-édition dans 17 % des cas uniquement (révision dans 51 % des cas, préparation de copie dans 19 % des cas et réécriture dans 9 % des cas). Deux tiers n'ont pas même reçu le texte original. Autre problème majeur : il n'a été indiqué dans aucun cas dans le manuscrit publié que le texte avait fait l'objet d'une traduction automatique. La baisse de reconnaissance soulignée par l'ATLF est motivée quant à elle par le constat que le nom des traducteurs n'a été mentionné que dans 31 % des cas dans la publication, et seulement 55 % de ces mentions l'étaient au titre de traducteur ou de traductrice. Concernant la rémunération, celle-ci était inférieure au tarif moyen de traduction dans 68 % des cas et (28 % équivalent), tandis que le délai accordé pour la tâche était jugé suffisant dans 56 % des cas (44 % insuffisant). Pour ce qui est du gain de temps, enfin, seulement 37 % des réponses vont dans le sens d'une amélioration, contre 56 % indiquant une tâche plus chronophage et 14 % le même temps. Si l'on reprend l'ensemble des retours d'enquête, il apparaît en outre que 55 % des traducteurs n'accepteraient pas de post-édition si le travail leur était proposé, 37 % l'accepteraient peut-être (« sous certaines conditions tarifaires et plutôt pour des textes non littéraires ») et 8 % y sont favorables, « principalement parce qu'elles manquent de contrats de traduction, pour des raisons économiques, ou par curiosité » (cf. [ATLF 2022](#)).

Dans l'ensemble, les témoignages livrés par l'ATLF confirment que la pratique est effectivement déjà répandue, quoique de façon encore restreinte. Il est d'ailleurs surprenant de voir le taux d'acceptation hypothétique de la post-édition, quand bien même celui-ci semble lié en premier lieu à des contraintes économiques qu'évoquait pareillement Nathalie à ce sujet ([SERVAL 2023a](#)). Et bien que les cas mentionnés ici semblent concerner des textes fonctionnels avant tout (presse, manuels, catalogues...), il serait intéressant de savoir, comme le remarquait également la traductrice, dans quel domaine l'utilisation de la TA est la plus répandue. Malgré tout, ces premiers regards sur l'introduction de la TA dans le domaine littéraire font ressortir des retombées majoritairement défavorables en pratique, qu'il pourrait être utile de confronter aux avantages potentiels recensés dans les études consacrées à la traduction littéraire automatique.

8.1.2 Avantages et bénéficiaires

Selon Laurent BESACIER (2014, p. 394), la TA pourrait profiter à l'ensemble de la chaîne de l'édition : pour les auteurs, l'avantage résiderait avant tout dans le fait de voir leur œuvre diffusée dans un plus grand nombre de langues ; pour le lectorat, ce serait la possibilité d'avoir accès

plus rapidement aux œuvres traduites ; pour la maison d'édition, le principal bénéfice serait à trouver — sans surprise — dans la réduction des coûts — dont nous pourrions espérer qu'elle ait des retombées positives sur la rémunération des traductaires, bien que les tendances dégagées par l'ATLF (2022) pointent l'inverse. Si ces intérêts doivent nécessairement être pris en compte, force est de constater qu'ils dépendent entièrement de la productivité des personnes que vise principalement cette discussion : les traductaires. De fait, la réduction des délais — et des coûts — est l'aspect qui va le plus souvent de pair avec les technologies de la traduction dans l'ensemble (O'BRIEN 2012, p. 106 ; DOHERTY 2016, p. 962), tandis que les avantages liés au processus de traduction répondent premièrement à des questions de qualité et d'ergonomie.

Envisagée comme un outil d'aide à la traduction, la TA pourrait par exemple être un moyen de repérer des erreurs, en confrontant sa traduction à celle générée par la machine (Ó MURCHÚ 2019, p. 23), ou de les éviter en premier lieu, si l'on démarre directement de la solution fournie (cf. GUERBEROF-ARENAS et TORAL 2020, p. 268). Il peut également s'agir d'un moyen de favoriser l'ergonomie physique, en réduisant l'effort de frappe (TORAL *et al.* 2018b, p. 7 ; VIEIRA *et al.* 2023, p. 226 ; LARSONNEUR 2023, p. 92). Sur le plan cognitif, la TA pourrait réduire l'effort fourni (GUERBEROF-ARENAS et TORAL 2022, p. 195), en particulier si l'outil est utilisé comme recherche de variantes selon Claire LARSONNEUR (2023, p. 92–93). De ce cas, la TA serait aussi un moyen de stimuler la démarche créative par la mise à disposition et la comparaison de solutions alternatives de traduction, renforçant par la même occasion la qualité de la production et rejoignant à ce titre mes conclusions précédentes pour la TLAO (cf. Section 2.3), puisque l'outil interviendrait essentiellement comme une ressource additionnelle similaire à une recherche de concordance dans une mémoire de traduction ou à un glossaire terminologique. Similairement, cette diminution de la charge cognitive, associée à une augmentation de la productivité pour les passages moins importants, donnerait aux traductaires la possibilité de focaliser leur attention sur les tâches plus complexes et de se concentrer sur les passages critiques où l'ingéniosité humaine est nécessaire (TAIVALKOSKI-SHILOV 2019a, p. 691 ; ŞAHIN et GÜRSES 2019, p. 26). Ces deux dernières conclusions s'alignent ainsi sur les deux postulats de départ que j'avais pu formuler concernant la créativité (cf. p. 17). Si le travail mené jusqu'ici et d'autres témoignages pointent surtout vers ce premier constat (cf. Section 7.5 ; DEFERT 2022 ; ROTHWELL 2024), il n'est pas exclu qu'il en soit de même pour le second avec une mise en place appropriée de la TA. Dans le même ordre d'idées, TAIVALKOSKI-SHILOV (2019a, p. 691–692) conclut que la technologie pourrait rendre le travail des traductaires généralement plus agréable, pour autant, cependant,

que les systèmes aient été conçus spécifiquement pour cette tâche. Cette dernière observation se vérifierait en particulier si la TA permet d'aboutir à une prise d'importance de la dimension créative, comme je le pense. C'est aussi le scénario envisagé par Susan Pickford, qui conçoit la traduction automatique comme un dispositif qui serait avant tout au service des traducteurs et dont le gain de productivité devrait leur servir, outre à s'assurer un avantage financier essentiel, à réserver du temps pour gagner en qualité et en créativité, pour se consacrer à la réécriture, pour s'approprier le texte et y affermir sa voix, pour y renforcer les aspects définitoires du genre, etc. (cf. HANSEN et PICKFORD 2022).

Réflexion faite, chacun des avantages listés ici concernant le processus de traduction (cohérence, qualité, créativité, productivité, ergonomie) figure parmi les avantages typiquement associés aux technologies de la traduction de manière générale (O'BRIEN 2012, p. 107–109). De fait, il apparaît que les outils d'aide à la traduction au sens large peuvent tout aussi bien soutenir et accentuer ces aspects que les entraver ou y porter préjudice. La technologie est ainsi intimement liée aux conditions de son utilisation et aux motivations pour son introduction, c'est pourquoi TAIVALKOSKI-SHILOV (2019a, p. 698) ajoute en dernière analyse que les retombées positives de son usage ne pourront se réaliser qu'« à condition que toutes les parties prenantes, en particulier celles qui ont le plus de poids, adoptent des pratiques durables en accord avec leur politique de qualité [40] ». Le chapitre précédent a été l'occasion pour moi de souligner l'importance des facteurs situationnels et organisationnels qui entourent le déploiement de la technologie, résumée par la formule ergonomique « on est meilleur quand on est mieux » (LAVAUULT-OLLÉON 2011, par. 21). Or, il est particulièrement éclairant de voir à cet égard que les recherches mentionnent des problèmes inverses pour chacun de ces avantages.

8.1.3 Risques et enjeux sociétaux

Si l'arrivée de la TA laisse entrevoir des avantages qui pourraient être au bénéfice des traducteurs, c'est en reconnaissant toutefois que les conditions de sa mise au point et de mise en place peuvent tout à fait renverser la situation et entraîner des risques donnant alors lieu à un sentiment de déshumanisation et de frustration pour les utilisateurs (O'BRIEN 2012, p. 108–109). Il est en effet important de reconnaître que les outils de TA et de TAO ont surtout évolué pour s'adapter aux demandes de l'industrie, en favorisant la productivité, la gestion de projet et les rabais, mais que ceux-ci ont très peu bénéficié d'autres types d'amélioration depuis leur apparition, surtout si l'on se place du point de vue des traducteurs (cf. Section 7.4.1). Il ne fait d'ailleurs aucun doute

que la motivation première de la TA est financière : les discours des prestataires de service et les recommandations énoncées par les normes de post-édition laissent peu de doute à ce sujet (BRUNETTE et O'BRIEN 2011, par. 7), tout comme son utilisation actuelle (ATLF 2022 ; DIFFERT 2023). Pour le secteur littéraire, la prise en compte et la prévention de ces risques sont d'autant plus essentielles qu'ils sont exacerbés par les conditions de travail déjà précaires des traducteurs littéraires et par les difficultés du marché (TAIVALKOSKI-SHILOV 2019a, p. 695). Dans ce contexte, comme je le mentionnais, la TLA pourrait apporter des changements positifs à ces égards, tout comme elle pourrait contribuer à détériorer des conditions déjà défavorables.

Un bon exemple de cette ambivalence provient des sources indiquant une réduction du nombre d'erreurs associées à la post-édition, qui rapportent parallèlement l'introduction d'erreurs par la machine (Ó MURCHÚ 2019, p. 22) et notent que la PE mène dans d'autres cas à une moins bonne qualité par comparaison avec la traduction humaine (GUERBEROF-ARENAS et TORAL 2022, p. 197). Si la TA a le potentiel de réduire la charge cognitive, elle peut aussi l'augmenter, en particulier dans le cas d'interfaces non intuitives (TAIVALKOSKI-SHILOV 2019a, p. 691). Ajouté au biais d'ancrage¹⁴⁴, qui donne lieu dans le modèle de post-édition traditionnel à des traductions plus proches du texte source (Ó MURCHÚ 2019 ; MOORKENS *et al.* 2018 ; GUERBEROF-ARENAS et TORAL 2020 ; 2022 ; CASTILHO et RESENDE 2022 ; KOLB 2024), l'utilisation de la TA peut dès lors contraindre l'intuition et empêcher les traducteurs de trouver des solutions originales, là où la valeur ajoutée de la traduction humaine est précisément nécessaire. Il a ainsi été montré que la PE atténuait la voix des traducteurs dans les textes produits (KENNY et WINTERS 2020 ; WINTERS et KENNY 2024), jugés quant à eux moins créatifs (GUERBEROF-ARENAS et TORAL 2020 ; 2022). Ces constats ne sont à nouveau nullement spécifiques au cas littéraire et sont bien documentés par ailleurs : on retrouve par exemple l'idée du biais d'ancrage dans le « halo de la machine » de (LARSONNEUR 2023, p. 94) ou, dans un autre contexte, dans l'« effet fantôme » de la TA décrit par Perrine SCHUMACHER (2023, p. 232–234), qui fait également état d'un « effet nivelant » sur la qualité de la traduction en contexte pédagogique, montrant que la TA favorise les personnes dont la production est la plus faible tout en diminuant la qualité des meilleures traductions humaines. Dans ces conditions, il paraît peu étonnant que les traducteurs trouvent la tâche de post-édition plus ingrate (MOORKENS *et al.* 2018).

144. Terme issu de la psychologie, que l'on peut rapprocher du concept de *priming* utilisé dans les textes anglophones, selon lequel la première impression est toujours la meilleure — dans ce cas, la solution fournie par la TA, dont il peut être difficile de se détacher.

Concernant le cas qui nous occupe, chacun de ces points se confirme, sans exception, au fil du [Chapitre 7](#). Nathalie y revient d'ailleurs point par point lors de notre dernier entretien, lorsqu'elle exprime ses craintes vis-à-vis d'une utilisation problématique de la TA ([SERVAL 2023b](#)), précisant par exemple qu'une utilisation non appropriée de l'outil risquerait de perturber la dynamique de travail et de déjouer la vigilance des traductaires, leur faisant perdre leur regard critique. Il est en effet apparu en rétrospective que la TA entravait le *flow* et empêchait toute interprétation « somatique » du texte source (cf. [Section 7.3.2](#)), conduisant à la différence significative entre la version post-éditée et non post-éditée de la traduction (cf. [Section 7.4](#)) où se notait l'atténuation de la voix de Nathalie et de sa créativité due à l'influence de la pré-traduction par la TA, ainsi qu'une moins grande prise de liberté dont elle fait habituellement preuve vis-à-vis du texte source. Outre la distraction causée par ces pré-traductions, la traductrice ajoutait celle causée par l'interface et par l'aspect encombrant de ses nombreuses fenêtres (*Ibid.*), qui m'a amené ensuite à repenser les modalités d'utilisation de la TA d'une manière qui privilégie l'autonomie, l'interactivité et la maîtrise sur le processus — j'y reviendrai par la suite. Au-delà de ces problèmes d'ergonomie cognitive, liée au processus de traduction, Nathalie anticipait enfin ceux liés à l'ergonomie organisationnelle, associée au contexte d'utilisation de la technologie. La principale crainte exprimée dans ce sens était celle de voir son utilisation forcée par les donneurs d'ordre et que le métier de traductaire littéraire tel qu'on le conçoit aujourd'hui n'en vienne à être remplacé par de simples auxiliaires multi-tâches qui auraient une sensibilité littéraire, éclipsant totalement le statut d'autaire attaché à la profession (*Ibid.*).

De fait, la traduction automatique pose aussi de nombreux risques si son utilisation est imposée ou vise premièrement à remplacer le travail humain. Un premier point cité par TAIVALKOSKI-SHILOV (2019a, p. 692) est celui de la réduction des délais, qui est déjà une préoccupation majeure pour des maisons d'édition essayant de plus en plus de publier les traductions au même moment que les versions originales ([DEFERT 2023](#) ; [SERVAL 2023a](#)) — Nathalie explique par exemple que le travail se fait de plus en plus souvent à partir des épreuves originales et que cette même raison l'a amenée à effectuer une traduction à quatre mains. À l'inverse, cette pression sur les publications simultanées pourrait encourager les maisons d'édition à pousser l'usage de la TA, comme l'a illustré la [Section 8.1.1](#). Celle-ci a également laissé entendre que son usage pourrait être motivé par des raisons financières qui conduiraient de la même manière à une réduction de la rémunération des traductaires ([TAIVALKOSKI-SHILOV 2019a](#), p. 692), ce que confirme les 75 % de réponses indiquant une activité moins rentable selon l'enquête réalisée par l'ATLF (2022).

Il est pourtant essentiel de préciser à cet égard que la post-édition n'est pas nécessairement plus aisée, ainsi que l'a confirmé le chapitre précédent, puisque celle-ci met en jeu trois textes au lieu de deux, qu'elle fait intervenir de nouvelles compétences (vigilance et repérage des erreurs) et qu'il reste impératif de retravailler et de reformuler le texte, d'y affirmer son style et sa créativité. De la même manière, Félix do CARMO (2020) remet en question de discours de simplicité typiquement associé à la post-édition, en confrontant celle-ci aux exigences et aux normes de l'industrie en la matière, tout en pointant le paradoxe d'une rémunération en constante baisse face au prix des formations continues et de la technologie elle-même. L'ATLF pointe ensuite le risque pour la visibilité des traductaires, auquel Ruth JONES et Ann IRVINE (2013, p. 100) font allusion. Risque qui est d'ores et déjà accru pour la littérature de jeunesse en particulier (LATHEY 2006, p. 1) et que la machine renforcerait davantage, comme nous l'avons vu dans certains cas où elle est déjà en usage. Ana GUERBEROF-ARENAS et Antonio TORAL (2020) mettent également en garde contre les effets de la TA sur la réception par le lectorat des œuvres traduites, tandis qu'Ó MURCHÚ (2019, p. 23) souligne quant à lui la possibilité de voir les traductaires confinés à un rôle moins important et intellectuellement moins intéressant. Ce thème de la dévalorisation du travail et du rôle des traductaires se retrouve d'ailleurs tout au long de l'article pionnier de TAIVALKOSKI-SHILOV (2019a), ainsi que dans la remise en question d'une approche *low cost* de la traduction, posée par BESACIER (2014, p. 394) : « faut-il sacrifier la qualité de traduction (et son contrôle sur celle-ci) au prix d'une diffusion plus large [et plus rapide des] œuvres ? » Opter pour un tel modèle conforterait l'idée d'une activité instantanée, automatique et gratuite aux yeux du public, plutôt que d'une profession hautement qualifiée et spécialisée impliquant un travail long et intensif (LARSONNEUR 2021, p. 263). Dans les scénarios les moins optimistes, il serait même possible de voir émerger des pratiques d'édition mercenaires consistant à vendre directement des sorties de TA non post-éditées ou à engager une main-d'œuvre non professionnelle moins chère pour réviser rapidement ces traductions automatiques (TAIVALKOSKI-SHILOV 2019a, p. 692).

Pour en revenir au processus de traduction et à ce qui doit être la principale question associée à la traduction automatique, il n'y a pour l'instant aucune certitude que l'utilisation de la technologie accélère la tâche, puisque les recherches sur le sujet témoignent surtout de la haute variabilité des résultats (KUZMAN *et al.* 2019, p. 5–6; KOLB 2024, p. 58–59). De plus, si certains cas d'étude laissent entrevoir un gain de temps (Ó MURCHÚ 2019, p. 23; TORAL *et al.* 2018b, p. 6), un nombre équivalent d'études tirent la conclusion opposée (VIEIRA *et al.* 2023, p. 227–228; GUERBEROF-ARENAS et TORAL 2022, p. 195), ce que reflète une fois encore l'enquête présentée plus haut. S'il est difficile de tirer des conclusions sur certains de ces aspects liés à l'utilisation de

la TA, en raison des résultats souvent contradictoires, il reste toutefois certain qu'une mauvaise implémentation de la TA conduirait inévitablement à une baisse de la qualité des textes produits si celle-ci privilégie la productivité au détriment de la créativité et de l'appropriation personnelle du texte par les traductaires. Il en va de même pour l'ensemble des facteurs sociaux que je viens d'énumérer, puisque nous avons vu que la qualité des traductions est intimement liée aux conditions de production qui entourent l'acte de traduction (LAVAUT-OLLÉON 2011, par. 19–23 ; TAIVALKOSKI-SHILOV 2019a, p. 692). Or, une baisse de qualité pourrait à son tour avoir des conséquences négatives sur la reconnaissance du travail et la notoriété des autaires — en plus de celle des traductaires —, sur l'expérience de lecture, portant préjudice au lectorat qui a payé pour ce service, de même que sur le transfert de culture, l'apprentissage des langues et les compétences linguistiques du lectorat (TAIVALKOSKI-SHILOV 2019a, p. 693).

Ces réflexions mettent en lumière l'ambivalence de la technologie, qui peut à la fois améliorer ou détériorer les conditions de travail des traductaires, mais aussi les risques associés à une utilisation non raisonnée des outils, lorsque celle-ci vise uniquement à faire l'économie d'une main-d'œuvre qualifiée. Dans l'ensemble, cette première section réactualise l'idée que le problème ne vient pas de la technologie elle-même, mais de l'utilisation qui en est faite. Elle rejoint ainsi le constat porté plus généralement sur l'industrie des langues concernant le manque de responsabilité sociale et d'éthique professionnelle des entreprises, qui se retrouvent subordonnées au souci de la rentabilité et qui est peu contesté par les traductaires, en raison de leur faible agentivité et de la disparité des rapports de pouvoir dans l'industrie (MOORKENS et ROCCHI 2021, p. 333). Parallèlement, l'analyse des discours de traductaires réalisée par Lucas Nunes VIEIRA (2020a) concernant les technologies de traduction illustre clairement que les doléances sont attribuables non pas aux outils, comme voudraient le faire croire certains discours présentant la profession comme technophobe, mais avant tout à la baisse des salaires et aux mauvaises pratiques commerciales. Dans ce contexte, il me semble primordial de pouvoir former et informer la prochaine génération de traductaires, les personnes qui sont déjà dans le métier et surtout, selon les discussions lancées par les associations littéraire au sujet de l'automatisation, les donneurs d'ordre (ATLAS et ATLF 2022, 19:49–20:04). Cela implique en premier lieu d'insister sur les compétences relevant de la TA dans la formation (*Ibid.*, 19:00–19:49), de manière à ce que les traductaires puissent avoir conscience de ces enjeux et des difficultés les plus fréquemment posées par la TA, mais aussi savoir quand il peut être profitable ou inutile d'utiliser les outils. Il faudrait en effet voir, dans l'idéal, si une partie des contraintes posées par la machine ne pourraient pas être allégées par une formation à la post-édition, ou par une expérience suffisante pour pouvoir travailler confortablement avec la TA

(LARSONNEUR 2023, p. 92). Il est difficile d'en juger à l'heure actuelle, car peu de traductaires ont une expertise solide à la fois en traduction littéraire et en PE, bien que l'expérience proposée par VIEIRA *et al.* (2023, p. 231) — la seule à ma connaissance pour laquelle toutes les personnes impliquées avaient une bonne expérience dans les deux domaines — accorde du crédit à cette piste. Ce savoir-faire devrait toutefois s'accompagner d'une meilleure connaissance des outils et de leur fonctionnement général, d'une sensibilisation aux enjeux présentés dans cette section et d'échanges de bonnes pratiques (SLESSOR 2020, p. 249). Concernant ces dernières, il s'agira par exemple pour le secteur littéraire de ne jamais utiliser d'outils gratuits en ligne, pour des raisons de confidentialité (ATLAS et ATLF 2022, 20:04–20:35). D'un point de vue déontologique, la maison d'édition qui décide de traduire un ouvrage devrait veiller à informer les traductaires, les agents littéraires et les auteurs — qui conservent un droit moral sur leur œuvre — de la façon dont le texte sera traduit et de l'utilisation éventuelle d'outils (*Ibid.*, 56:20–57:45). De manière plus générale, nous aurions tout intérêt à revaloriser la place de l'humain, qui fait le pont entre les différents maillons de la chaîne et entre une panoplie d'outils, auxquels vient maintenant s'ajouter la TA, et qui se doit de maîtriser ceux-ci pour traduire le plus efficacement possible, en continuant d'y apporter toutes les compétences qui lui sont propres (SCHUMACHER 2023, p. 396–398). Avant toute chose, cependant, il est nécessaire d'assurer une totale transparence sur l'utilisation des outils, afin d'encourager des pratiques saines et durables qui bénéficieront à l'édition au sens large (ATLAS et ATLF 2023 ; BVPA *et al.* 2023 ; CEATL *et al.* 2023 ; GUERBEROF-ARENAS et MOORKENS 2023, p. 129).

Pour terminer, dans la mesure où les retombées de la TA sont entièrement dépendantes des modalités d'utilisation de la TA, l'exigence de transparence qui vient d'être énoncée devra passer par une indication claire de l'objectif visé par la TA, une précision qui serait fondamentale dans les recherches effectuées sur le sujet par ailleurs. Pour l'instant, nous en sommes au stade du *good enough* (LONG 2022, p. 740), ce qui veut dire qu'elle se profile comme une aide avant tout, soit à des fins de compréhension ou de diffusion, soit pour les traductaires, en accord avec les deux voies de la TA définies par François YVON (2019). Dans le premier cas, on peut trouver plusieurs scénarios intéressants d'utilisation de la TA, qui ne font pas intervenir les traductaires mais *ne visent pas la publication*. C'est le cas notamment de son intégration comme aide à la lecture ou à l'apprentissage des langues (BESACIER 2014, p. 394 ; MATUSOV 2019, p. 18), dont l'état de l'art a décrit deux mises en application concrètes (cf. OLIVER *et al.* 2022 ; OLIVER 2024). Le modèle de la fiction Web, qui s'inscrit dans la logique de la culture participative et des communautés de fans pour fournir un service où la production serait humainement impossible à traduire, montre par

ailleurs que la traduction amatrice, professionnelle et automatique peuvent chacune coexister avec des objectifs et des publics spécifiques (SHIH 2016 ; ZHANG 2024). De la même manière, si une utilisation abusive de la TA pose un risque pour la visibilité, elle pourrait à l'inverse favoriser le transfert culturel et la circulation d'œuvres méconnues auprès du public (LONG 2022, p. 740), mais aussi assurer une plus grande diversité dans le panorama de la traduction littéraire, que ce soit en permettant aux maisons d'édition de se familiariser avec de nouveaux ouvrages (MATUSOV 2019, p. 18), issus de langues à plus faible diffusion par exemple (TORAL et WAY 2015a, p. 243) ou en facilitant l'envoi d'échantillons de traduction pour des auteurs appartenant à une minorité (CASTRO 2020) ou pour des traducteurs bénéficiant de peu de visibilité (TAZELAAR 2020).

Dans le deuxième cas, plus proche des motivations de ce travail, la TA pourrait apporter une aide pour la traduction d'ouvrages qui *visent à être publiés*, mais uniquement lorsque l'humain est au centre de la boucle. De fait, le seuil de tolérance est bien plus élevé dans le cas d'une publication et doit répondre à des critères de qualité et à des considérations esthétiques que seules des traductions humaines professionnelles peuvent délivrer (LONG 2022, p. 740). Dans ce contexte, la personnalisation des outils, leur appropriation intellectuelle et matérielle, le caractère dynamique de l'interaction humain-machine ainsi que la prise en compte des traducteurs humains dans le développement et le déploiement des outils jouent un rôle crucial et déterminant sur l'acceptation et l'utilité de la technologie (O'BRIEN 2012, p. 109). Il est en effet essentiel de repenser sérieusement la mise en place ergonomique de la TA ainsi que son rôle, si l'on veut faire pencher la balance du côté des avantages plutôt que des inconvénients qui ont été détaillés dans ce début de chapitre. Cette idée rejoint le constat posé il y a des années déjà par Antonio TORAL et Andy WAY (2015a, p. 264), pour qui la faisabilité de la TLA dépend non seulement d'une amélioration de ses performances, mais aussi d'une interaction avec la TA plus adaptée à ce domaine, et qui insistent dès lors sur le fait que l'introduction de la TA en littérature devrait se faire d'une manière différente de la post-édition traditionnelle. Le changement de paradigme que j'envisage dans cette recherche par le biais de la traduction automatique individualisée prend donc ici tout son sens, à savoir celui d'une aide à la traduction humaine qui ne prétend nullement la remplacer ni s'approprier le travail ou l'agentivité des traducteurs, comme il en est habituellement question avec la post-édition. Outre la personnalisation offerte par un outil adapté à sa voix et à son style, c'est une manière radicalement différente d'interagir avec la TA que je propose, qui permettrait aux traducteurs de rester maîtres de l'outil et de choisir quand ou comment l'utiliser, et qui viserait non pas à fournir un premier jet, contraignant l'intuition, la voix et la créativité, mais à fournir une source d'inspiration ou à faciliter la recherche de solutions

alternatives, limitant de ce fait l'influence que pourraient avoir les sorties de TA sur la lecture et le décodage du texte source par les traductaires. À ce propos, la [Section 8.2.3](#) reviendra plus en détail sur ce changement de perspective et s'accompagnera de recommandations concrètes concernant l'utilisation de la TA en littérature.

8.2 La traduction littéraire automatique demain : défis et perspectives

Finalement, le risque majeur posé par la TA est celui que soulevait déjà Georges MOUNIN (1976, p. 38), à savoir que « [l]es hommes finiraient par travailler pour la machine, au lieu que la machine travaille pour les hommes ». En ce sens, la TLA soulève des réflexions qui dépassent largement la traduction littéraire ou le secteur de l'édition et qu'il me semble nécessaire d'aborder pour compléter ce chapitre de discussion, avant d'en venir au changement de paradigme que je propose en conclusion de ce travail. La première concerne les coûts des outils, qui semblent s'accroître au fil des avancées de la technologie et que passent sous silence les grandes annonces des prestataires de service. De cette première réflexion dérive ensuite la question des données, qui s'est constituée comme un fil rouge tout au long de ce travail et qui amène des enjeux éthiques supplémentaires concernant la société au sens large.

8.2.1 *Les coûts cachés de l'IA*

Une difficulté majeure des discussions concernant l'évaluation et l'utilisation de la traduction automatique provient de la *hype* trompeuse autour de la TA et de l'IA en général, qui ne favorise aucunement l'approche raisonnée de la technologie dont j'ai souligné l'importance en entamant ce travail (cf. [Section 3.3](#)). Ce battage médiatique est renforcé par la manière dont sont vendus — au sens littéral et figuré — les produits issus de ces recherches. De fait, lorsque des articles scientifiques — systématiquement produits par des géants de l'industrie — en viennent à parler de performances « sur-humaines », il n'est pas étonnant de voir la presse s'emparer de ces affirmations pour en tirer des titres futuristes ([POIBEAU 2022](#)). On peut le noter dans le seul nom donné aux applications ([WILLIAMS et al. 2022](#)) : du *big data*, celles-ci sont passées au *deep learning* et plus rapidement encore à l'*intelligence artificielle*. Cependant, derrière les promesses séduisantes d'augmentation de la qualité de vie offertes par l'informatique ubiquitaire, censée apporter le confort à portée de clic, se cachent des coûts humains, financiers et écologiques qui doivent impérativement être pris en compte dans l'équation, en accord avec notre approche critique de la technologie (cf. p. 153 ; [AGRE 1997](#)).

Un coût financier

Depuis l'arrivée des *Transformers* (VASWANI *et al.* 2017), les avancées rapportées dans le domaine du traitement automatique des langues (TAL) semblent moins tenir à l'amélioration des systèmes qu'à la quantité de données utilisées pour les entraîner et à leur taille (BENDER *et al.* 2021), dont les médias ne manquent pas de reporter la multiplication. Le meilleur exemple de cette tendance est probablement l'évolution des *GPT* (de l'anglais *Generative Pre-trained Transformer*) produits par OpenAI, puisque la deuxième itération du logiciel comporte 1,5 milliard de paramètres (RADFORD *et al.* 2019) et que la troisième en affiche plus plus de 175 milliards (BROWN *et al.* 2020), là où la dernière version en date est estimée à 1,75 mille milliards de paramètres (ALBERGOTTI 2023). Or, ces systèmes requièrent une infrastructure de plus en plus performante et de plus en plus d'énergie pour être mis au point et déployés ensuite. Ainsi, l'intelligence artificielle est une réelle révolution numérique, mais une révolution dont le coût devient petit à petit difficilement imaginable et monte continuellement avec la flambée de ces paramètres. Ce coût varie fortement en fonction du matériel utilisé et de la localisation, mais s'étend de quelques dizaines ou centaines d'euros pour chaque entraînement d'un *Transformer* simple, à plusieurs dizaines de millions pour des modèles plus larges et des projets de plus grande envergure, dont la réalisation implique plusieurs passes (STRUBELL *et al.* 2019 ; SHARIR *et al.* 2020). Dans le cas de *GPT-4*, le montant atteint même les 100 millions de dollars (KNIGHT 2023).

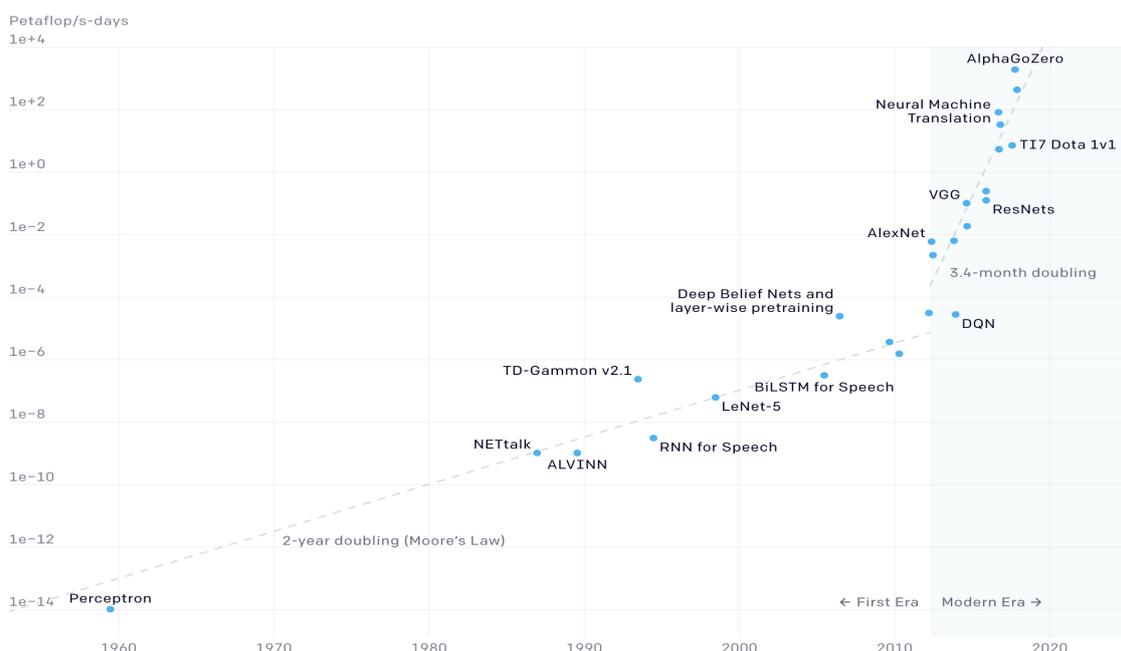


FIGURE 32 – Évolution de la puissance de calcul associée au *deep learning*

SOURCE : AMODEI *et al.* (2018)

Avec l'explosion du *deep learning*, la loi de Moore qui prévoit l'arrivée de matériel informatique deux fois plus performant tous les deux ans est passée à une évolution exponentielle de la puissance de calcul. Pourtant, cette montée de la puissance de calcul contraste avec des gains de performance de plus en plus faibles (THOMPSON *et al.* 2022). STRUBELL *et al.* (2019) estiment à titre indicatif qu'une amélioration de l'état de l'art en traduction automatique de 0.1 BLEU équivaut environ à une augmentation 150 000 dollars. Face à de tels chiffres, on comprend dès lors qu'un nombre grandissant de voix se demandent si de tels coûts sont réellement nécessaires et inévitables, surtout lorsqu'un cerveau humain peut s'acquitter du travail demandé avec une consommation d'énergie cent mille fois moins élevée, pour des tâches bien plus complexes que la traduction (TRYSTRAM *et al.* 2021). Dans un contexte où nous sommes poussés à repenser nos modes de consommation, le numérique occupe une place de premier plan qui nécessite de réfléchir à la réelle plus-value des outils et de favoriser l'alternative *low-tech* (*Ibid.*), mais aussi d'évaluer les avancées (et les risques) de la technologie au regard de leur coût, à la fois financier et écologique (BENDER *et al.* 2021).

Un coût écologique

Au-delà du coût financier, cette course du TAL vers toujours plus de données et toujours plus de paramètres va de pair avec son coût écologique. Les progrès rendus possibles dans ce domaine, par l'augmentation de la puissance de calcul et des quantités de données nécessaires à l'entraînement des systèmes, sont en effet intimement liés aux besoins énergétiques grandissants de ces avancées. Jusqu'à récemment, on parlait encore peu de l'empreinte environnementale du numérique, mais l'explosion du *deep learning* et les annonces toujours plus spectaculaires de modèles aux milliards et aux billions de paramètres amènent peu à peu le sujet sur le devant de la scène, avec une importance particulière dans ce contexte de changement climatique. STRUBELL *et al.* (2019, p. 3648), par exemple, estimaient l'empreinte carbone d'un modèle de langue première de génération à une traversée des États-Unis en avion, sachant aujourd'hui que chaque itération de *GPT* a vu sa taille multipliée au moins par dix, le tout dans un climat de course au score BLEU qui se solde le plus souvent par un gain de quelques dixièmes de points. Pour cette raison, un nombre croissant de scientifiques proposent de mettre en regard les gains offerts par les technologies vis-à-vis de leur coût en énergie (HENDERSON *et al.* 2020, p. 1–2 et 22–23).

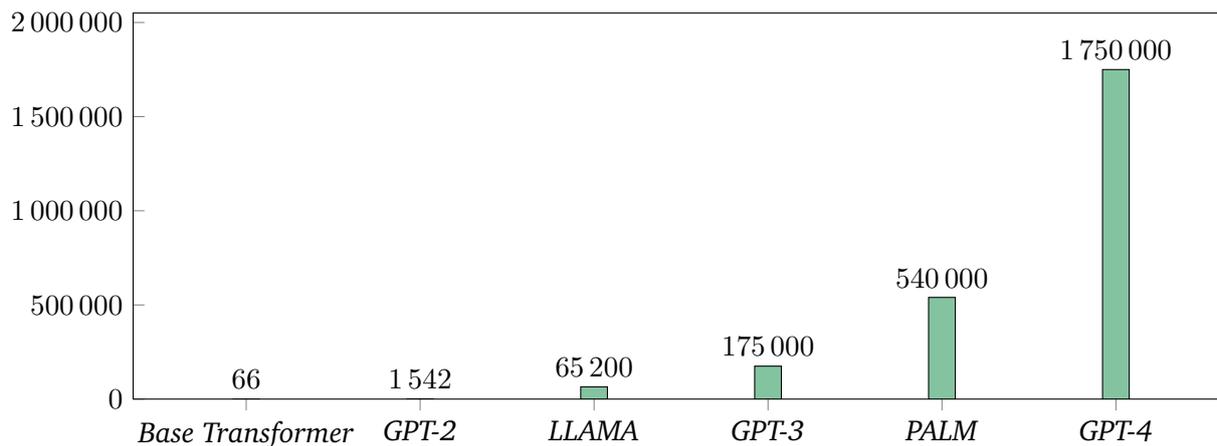


FIGURE 33 – Comparaison de la taille de modèles de langue (en millions de paramètres)

La Figure 33 ci-dessus fournit une illustration du *Transformer* de base, sur lequel je me suis appuyé dans ce travail, par comparaison avec des modèles de langue qui ont plus récemment fait parler d'eux et qui se basent sur cette même architecture. Le coût qui est associé à ces différents modèles est bien entendu lié à leur entraînement, auquel il faut toutefois ajouter le coût des infrastructures nécessaires pour les faire tourner ensuite, en ce compris les différents services proposés en *cloud* (accès à des serveurs à distance), dont on oublie bien souvent aussi l'empreinte carbone. Plus encore, l'évaluation des systèmes neuronaux elle-même implique maintenant des modèles *Transformer*, comme COMET (REI *et al.* 2020), qui nécessitent une puissance GPU non négligeable.

Pour donner une idée de la consommation liée aux entraînements, l'ensemble des expériences menées (concluantes et non concluantes) et des modèles entraînés (LSTM et *Transformer*) dans ce travail est estimée à 5,23 kg d'équivalent de CO₂ (CO₂e), que l'on pourrait comparer à 21,1 km roulés avec une voiture moyenne ou à 2,62 kg de charbon brûlés. Ce coût énergétique est calculé selon la méthode suggérée par LACOSTE *et al.* (2019)¹⁴⁵ en totalisant le nombre d'heures d'entraînement des systèmes présentés dans la méthodologie (cf. Section 5.3), tout en gardant à l'esprit que certains systèmes ont été affinés à partir de modèles génériques, mais aussi en tenant compte de l'enveloppe thermique (TDP) du matériel utilisé et de l'indicateur d'efficacité énergétique (PUE) du centre informatique de l'IMAG, soit 299,58 h, 260 W et 1,4 respectivement.

145. Cf. <https://mlco2.github.io/impact/>.

$$kWh = heures \times consommation \times PUE$$

Enfin, il faut également tenir compte de l'intensité carbone, qui était de 48 g de CO₂e par kWh en moyenne pour la France¹⁴⁶, où les expériences ont été menées, ce qui a contribué à réduire considérablement leur coût par comparaison avec d'autres régions (HENDERSON *et al.* 2020, p. 18–19).

$$kg\ CO_2e = kWh \times kg\ CO_2e/kWh$$

À titre de comparaison, l'empreinte environnementale de modèles tels que *GPT-3* ou encore *OPT* (S. ZHANG *et al.* 2022) est estimée en tonnes d'équivalent de CO₂ (500 et 75 respectivement) (*Ibid.*; PATTERSON *et al.* 2021). Le modèle le plus performant finalement choisi pour l'étude menée ici (*Septimus-étendu*), qui utilise l'architecture de base du *Transformer*, affiche « seulement » un coût de 0,4 kg CO₂e (22 h 57), tandis que le seul processus d'affinage sur les données littéraires (4 h 02) est de l'ordre de 0,05 kg CO₂e¹⁴⁷. D'une part, ces chiffres montrent que l'on peut parvenir à des résultats tout à fait intéressants avec de « petits » modèles spécialisés et rapides à entraîner. D'autre part, le processus d'adaptation en particulier laisse également penser que l'on peut facilement obtenir de tels résultats à partir de systèmes génériques pré-entraînés — peut-être même meilleurs si l'affinage est mené à partir d'un système générique plus robuste —, pour un coût plus négligeable encore.

Évidemment, il ne s'agit ici que d'une estimation qui de surcroît ne compte pas les tests nécessaires pour optimiser les hyperparamètres des architectures¹⁴⁸ (STRUBELL *et al.* 2019, p. 3645) ou le déploiement de l'outil ensuite (LACOSTE *et al.* 2019, p. 4), qui représenterait jusqu'à 90 % de la consommation d'énergie associée au déploiement d'un système — les 10 % restants étant liées à l'entraînement — selon PATTERSON *et al.* (2021). Ces coûts viennent donc eux aussi renforcer l'idée que la recherche n'est pas neutre et que les avancées du TAL devraient être envisagées dans une perspective de coûts-bénéfices, en distinguant les outils qui font office de « gadget » et ceux qui répondent à un « besoin indispensable » (TRYSTRAM *et al.* 2021).

146. Cf. <https://app.electricitymaps.com/zone/FR>.

147. Une page *Wikipédia* fournit d'autres exemples d'empreinte carbone pour différents ordres de grandeur. Cf. https://fr.wikipedia.org/wiki/Ordres_de_grandeur_d%27empreinte_carbone.

148. Cf. Tableau 2.

Un coût humain

Derrière les annonces et démonstrations spectaculaires se cachent également une industrie à part entière et une masse de travail pouvant frôler l'exploitation dans les cas les plus extrêmes ou se manifester dans des pratiques plus anodines, sous une forme de travail fantôme — travail invisible, banalisé et non rétribué — que nous effectuons toutes et tous au quotidien en interagissant avec les outils et en produisant des données telles que des traductions. L'ouvrage *En attendant les robots* d'Antonio CASILLI (2019) présente ainsi l'envers du décor de l'automatisation et dévoile les réalités du microtravail numérique qui s'est développé avec les recherches en intelligence artificielle, ainsi que celle des « travailleurs du clic », dont le rôle est essentiel au bon fonctionnement de la technologie. Un exemple frappant de cette pratique a récemment été porté à l'attention du public avec la sortie de *ChatGPT* (OpenAI 2022), pour laquelle l'entreprise a employé une force de travail au Kenya pour moins de deux dollars de l'heure, avec l'objectif d'annoter les sorties du chatbot, notamment pour en atténuer les biais et éliminer les contenus toxiques (PERRIGO 2023). Le plus souvent, cette force de travail invisible est située dans des pays en voie de développement, où elle se voit accorder un salaire dérisoire pour voir défiler du contenu haineux, annoter des images horribles ou créer du contenu pour alimenter les outils du TAL (CASILLI 2019). Dans ce contexte, on constate une tendance générale grandissante à proposer un panorama toujours plus large de services présentés comme relevant de l'IA, alors que ceux-ci sont réalisés en réalité par des humains (TUBARO *et al.* 2020, p. 7–8). Un autre exemple éloquent de ce mouvement a été la démonstration d'un système dont la traduction était réalisée en réalité à la volée par des interprètes, qui n'étaient évidemment pas au courant de cette finalité (NG 2018). Enfin, il ne faut pas perdre de vue que l'amélioration des outils dépend aussi en grande partie des millions d'interactions quotidiennes avec les utilisataires mêmes, qui produisent du contenu créatif, renforcent et corrigent l'algorithme.

De cette manière, les géants du numérique s'attellent à la réalisation de tâches telles que la traduction entièrement automatique de haute qualité tout en passant sous silence le travail humain (ŞAHIN et GÜRSES 2023, p. 6). Derrière les outils dits d'*intelligence artificielle*, qui ne sont finalement ni réellement intelligents ni totalement artificiels (CRAWFORD 2021), se cachent surtout la réalité du travail du numérique et la précarisation croissante des emplois (CASILLI 2019; WILLIAMS *et al.* 2022), facilitée par des plateformes comme Amazon Mechanical Turk ou Appen, qui proposent la mise à disposition de « spécialistes de l'enrichissement de données » 24 heures sur 24 et 7 jours sur 7. Or, les efforts consentis pour invisibiliser cette main-d'œuvre essentielle au bon fonctionnement des outils vont de pair avec ceux déployés pour mettre en avant la performance de la technologie (PERRIGO 2023).

Le travail humain est non seulement nécessaire au bon fonctionnement de l'« IA », mais il est aussi réalisé à une échelle dont l'étendue tranche avec l'image des machines douées de conscience que véhiculent les médias et les équipes marketing, en s'inspirant de la culture populaire (WILLIAMS *et al.* 2022). Ces annonces publicitaires cachent ainsi une réalité bien différente et entraîne un manque de transparence pourtant essentielle au bon avancement de la recherche (BENDER *et al.* 2021), en particulier à un moment où les industries qui dominent le marché sont du plus en plus opaques sur le fonctionnement de leurs algorithmes (NG 2018). Nous avons vu dans la Section 3.2 que ce phénomène de *hype* entourant la TA n'était pas nouveau, mais qu'il prenait des proportions particulières dans le cas de la traduction automatique neuronale (TAN), renforcées en cela par les considérations économiques importantes qui ont été soulevées tout au long de ce chapitre. Or, la transparence est nécessaire si l'on veut promouvoir des discussions raisonnées entre toutes les parties impliquées dans la mise au point, le déploiement et l'utilisation de la technologie :

Les personnes qui ont vu défiler de nombreux paradigmes [de TA] savent combien il peut être contre-productif de verser dans l'excès, en particulier celles et ceux d'entre nous qui ont tenté pendant des années d'établir des liens entre le monde de la TA et de la traduction. L'intervention humaine restera toujours le maillon le plus important de la chaîne, à tout le moins là où la diffusion du contenu traduit est impérative. La seule chose qui nous importe est d'améliorer les sorties des systèmes pour permettre aux traducteurs et traductrices qui utilisent la technologie d'augmenter leur productivité. Il est peu probable que les systèmes de traduction combleront un jour l'écart avec la traduction humaine ou atteignent la parité avec celle-ci. Ce n'est pas parce qu'un nouveau paradigme est en vogue que la TA est subitement devenue une tâche facile, voire un problème réglé comme on voudrait parfois le faire croire [41]. (WAY 2020, p. 325–326.)

Comme l'évoque le chercheur, ces discussions raisonnées devraient servir à renforcer la confiance avec les utilisataires, c'est d'ailleurs précisément pourquoi il conviendrait de mettre en avant les cas de coopération fructueuse plutôt de renforcer l'antagonisme entre humain et machine (POIBEAU 2022, p. 6021).

Un dernier aspect qui contribue lui similairement à l'invisibilisation du travail humain tient à ce que l'on s'attarde très rarement sur le fait que les performances croissantes du *deep learning* sont rendues possible en premier lieu par l'utilisation de données dont la quantité et la qualité sont primordiales (KENNY 2012; TUBARO *et al.* 2020), ainsi qu'a également servi à le démontrer ce travail. Le rôle des traductaires, de fait, ne se limite pas à la post-édition, mais aussi à la production de données en amont qui seront réutilisées pour l'entraînement des outils. De cette

manière, Mary HEARNE et Andy WAY (2011, p. 238) pointent la centralité des traductaires dans le fonctionnement de la TA, dont les traductions fournissent au système l'intégralité de ses « connaissances ».

8.2.2 La nécessité d'un cadre juridique

Commentant ailleurs la situation des traductaires littéraires dans le panorama général des nouvelles technologies, j'ai pu établir par la même occasion la centralité des données dans le fonctionnement des outils de traduction (HANSEN 2024). Que l'on s'y réfère en tant que corpus parallèles, bitextes, mémoires de traduction, jeux d'entraînement ou sous une autre dénomination, ces mines de données ont irrémédiablement transformé le quotidien des traductaires et se sont durablement installées dans leur pratique. Leur prolifération et la place centrale qu'elles occupent au sein des technologies de la traduction, depuis les moteurs de recherche et les concordanciers aux programmes de TAO et de TA, ont ainsi indéniablement marqué l'industrie. La généralisation de la TAO et le rôle des agences de traduction avaient déjà constitué un premier tournant majeur en ce sens, mais la question se pose plus directement et plus largement avec l'avènement du *deep learning* (MOORKENS et LEWIS 2020, p. 469 et 472), ce qui n'a pas échappé aux prestataires de services linguistiques. Les rapports dressés au tournant des années 2020, notamment chez Nimdzi (AKHULKOVA *et al.* 2021) et TAUS (ASLAN 2021), témoignent clairement d'un changement amorcé des services traditionnels associés à la traduction vers des tâches de récolte, d'annotation et de mise à disposition des données, pour lesquelles les prestataires tentent de mettre à profit leurs données linguistiques et de réinvestir leurs actifs. Le nouveau *Data Marketplace* de TAUS (2020) illustre parfaitement cette tendance, puisque la plateforme vise à permettre aux entreprises et aux prestataires — parmi lesquels on retrouve des maisons d'édition — de vendre leurs corpus avec un prix au mot qui rappelle et dépasse parfois les taux habituels pour la traduction. Ces tendances appuient l'idée que la traduction devient de plus en plus un marché de données, renforçant la validité de l'adage « *data is the new oil* », ce que confirment ŞAHIN et GÜRSES (2023, p. 3), avançant que « dans la traductosphère actuelle, l'attention est premièrement accordée à la datafication du contenu traductionnel, elle-même accélérée par les avancées technologiques et facilitée par ce qui est devenu la connectivité ubiquitaire ». Dans ce contexte, l'utilisation massive des données suscite des interrogations juridiques et éthiques qui doivent être appréhendées.

Ces questions concernant les aspects juridiques de la TA ont dernièrement gagné en importance, comme le montreront les références de cette section et la date de leur parution. De la même

manière, l'une des tables rondes organisées lors de la dernière édition de la Foire du livre de Francfort fait par exemple état de l'inquiétude exprimée par des auteurs, traducteurs et journalistes face aux nouveaux outils génératifs tels que *ChatGPT*, dont l'objectif est précisément de générer du contenu, comme en atteste par ailleurs le nombre croissant de grèves, revendications et actions en justice de la part des associations et autres actants de ces secteurs (BVPA *et al.* 2023). Si je reviendrai sur cette dernière génération de modèles de langue et sur les liens avec des domaines autres que la traduction dans la prochaine section, pour me concentrer ici sur la TA, on peut remarquer que ces préoccupations s'inscrivent dans une thématique actuelle et plus large que les considérations présentées ici, mais dont se dégagent généralement trois questions.

La première, bien documentée, concerne les protections légales actuellement associées aux œuvres originales et aux traductions. On trouve notamment chez Miguel L. LACRUZ MANTECÓN (2023) une réflexion concernant les instruments protégeant les œuvres de l'esprit, parmi lesquels la Convention de Berne vise précisément les textes littéraires et artistiques¹⁴⁹. L'article 8 de cette convention précise notamment qu'il revient aux auteurs d'autoriser la traduction de leurs œuvres : « Les auteurs d'œuvres littéraires et artistiques protégés par la présente Convention jouissent, pendant toute la durée de leurs droits sur l'œuvre originale, du droit exclusif de faire ou d'autoriser la traduction de leurs œuvres ». Si la traduction est autorisée, cependant, celle-ci est considérée à son tour comme une œuvre originale, conférant les mêmes droits aux traducteurs sur leur production conformément à l'article 2.3 : « Sont protégés comme des œuvres originales, sans préjudice des droits de l'auteur de l'œuvre originale, les traductions, adaptations, arrangements de musique et autres transformations d'une œuvre littéraire ou artistique ». Selon Pascal REYNAUD et Tiphaine KLEIN (2018, p. 6), ce règlement s'étendrait à tout type de traduction, peu importe le caractère littéraire ou pragmatique, le support ou le type de rémunération, du moment que celle-ci comporte une valeur originale et créative :

En droit d'auteur, il n'y a pas à opposer traduction littéraire et traduction fonctionnelle. Elles bénéficient du même régime légal. Par principe, le droit d'auteur ne distingue pas entre les « beaux-arts » et les « arts appliqués ». C'est ce qu'on appelle la théorie de l'unité de l'art selon laquelle la même loi doit être appliquée à tous les types de création. Ainsi, selon cette théorie, toute création de mots peut faire l'objet de protection par le droit d'auteur, mais à la condition d'être originale.

149. Convention de Berne pour la protection des œuvres littéraires et artistiques, 9 sept. 1986, <https://www.wipo.int/wipolex/fr/text/283699>.

À ces dispositions s'ajoutent ensuite des législations au niveau national qui vont généralement dans le même sens de ces protections et autorisations d'œuvres dérivées (LACRUZ MANTECÓN 2023, p. 72), à l'image des articles XI.165–XI.172 du Code de droit économique en Belgique¹⁵⁰ et des articles L112-1–L112-4 du Code de la propriété intellectuelle en France¹⁵¹. Plus encore, MOORKENS et LEWIS (2020, p. 471) ajoutent que les traductaires détiennent en principe les droits sur leur traduction y compris vis-à-vis de l'auteur de l'œuvre originale, qui ne peut réutiliser la traduction sans autorisation préalable.

Face à ces différents cadres légaux, on peut naturellement se demander comment les traductions, y compris littéraires — je reparlerai ci-après de la circulation de tels corpus —, peuvent être utilisées si aisément dans l'entraînement de ces systèmes, au point d'être diffusées librement en ligne. Comme l'explique Mikel L. FORCADA (2023, p. 49–53), le travail humain est présent à toutes les étapes d'une traduction automatique : des textes monolingues, corpus parallèles et autres ressources linguistiques (dictionnaires, outils d'analyse morphologique ou syntaxique, etc.) utilisées pour mettre au point le système, au moteur de traduction même, en passant par les outils complémentaires qui ont permis de compiler et de transformer ces données. Toutefois, poursuit l'auteur, le cas des corpus est particulier en ce qu'ils n'ont pas été produits — pour la majeure partie — pour cet usage spécifique (*Ibid.*, p. 50 et 55). Néanmoins, le volume des données, les modifications apportées à celles-ci et leur traitement par la machine font qu'il est pratiquement impossible de reconstituer le document original. La compilation d'un corpus, par exemple, implique souvent une segmentation et un alignement phrase par phrase, dont le formatage peut être supprimé et l'ordre modifié aléatoirement, tandis que certains segments peuvent ne pas être pris en compte, être supprimés ou mélangés avec des segments provenant d'autres documents (*Ibid.*, p. 62). Pour compliquer les choses, la taille des corpus elle-même et le nombre d'auteurs ou de traductaires probablement impliqués dans ces collections rendent d'autant plus complexe l'attribution des données à une personne en particulier (MOORKENS et LEWIS 2020, p. 474). Dès lors :

[...] si le corpus aligné phrase par phrase peut être vu comme une infraction aux droits de propriété attachés à de nombreuses pages de contenu, il est très peu probable que les personnes

150. Code de droit économique, Livre XI, Titre 5, Chapitre 2, 29 mars 2013, <http://www.ejustice.just.fgov.be/eli/loi/2013/02/28/2013A11134/justel>.

151. Code de la propriété intellectuelle, Livre I^{er}, Titre I^{er}, Chapitre II, 3 juill. 1992, https://www.legifrance.gouv.fr/codes/texte_lc/LEGITEXT000006069414.

disposant de ces droits s'attaquent à celles qui ont compilé les corpus, car il serait très difficile de prouver des dommages et plus encore d'y attribuer une valeur [42]. (FORCADA 2023, p. 62.)

De plus, le passage par un système de TA — et la TAN en particulier, qui fragmente les données à un niveau inférieur au mot (cf. [Figure 23](#)) — fait que toute trace des textes originaux disparaît complètement, tant les données ont été transformées, désassemblées et réassemblées (*Ibid.*, p. 63). Pour ces raisons, les entreprises qui développent et commercialisent ces systèmes s'inquiètent peu des risques encourus, tant le risque est minimal (*Ibid.*, p. 61–64; LACRUZ MANTECÓN 2023, p. 88). J'ai d'ailleurs déjà évoqué dans les chapitres précédents que Google ne se cachait pas d'utiliser les livres numérisés par ses services pour son système de traduction automatique (cf. SARNO 2010). Enfin, les règlements visant le principe de *fair use* (usage raisonnable) et les exceptions liées à l'extraction de données (*data mining*) pourraient être mobilisées en cas de plainte et compliquer davantage la justification des dommages (FORCADA 2023, p. 64).

Reste à savoir à ce stade si une post-édition serait protégée par les mêmes prérogatives que pour les traductions et les textes originaux. La troisième et dernière question qui suit logiquement ces débats concerne donc les possibilités de protection des traductions (assistées ou non par la machine) et de compensation en cas d'utilisation pour l'entraînement des outils de traduction (ou des outils génératifs). Sans aborder le cas des mémoires de traduction ou d'une sortie de TA non post-éditée restent sujets à débat (MOORKENS et LEWIS 2020, p. 472–474; LACRUZ MANTECÓN 2023, p. 78–85), nous avons vu précédemment que la question est principalement liée à la notion d'originalité, pour laquelle REYNAUD et KLEIN (2018, p. 9–12) estiment qu'un refus de la reconnaissance du droit d'auteur concernerait avant tout des traductions littérales sans aucune marque d'intervention humaine. Or, les résultats et recherches présentés au [Chapitre 6](#) montrent qu'il s'agit encore du principal défaut de la TA et qu'une adaptation originale reste indispensable au vu de la qualité de la TA. Dans de tels cas, il est tout à fait évident que les traductaires conservent leurs droits sur le texte produit (LACRUZ MANTECÓN 2023, p. 74–75).

Malgré de nombreux questionnements subsistants qui montrent les limites des textes de loi actuels (MOORKENS et LEWIS 2020, p. 475), on s'aperçoit que le droit d'auteur couvre, en théorie, à la fois les traductions et textes littéraires existants, qui ne devraient dès lors pouvoir intervenir dans l'entraînement des systèmes (ATLAS et ATLF 2022, 46:50–48:42), mais aussi l'intervention des traductaires en cas de post-édition, puisque le degré d'intervention demandé pour des textes littéraires se rapproche encore aujourd'hui de la réécriture bien plus que de la simple révision (*Ibid.*, 52:25–53:58). Les expériences menées tout au long de ce travail confirment à nouveau ces

observations, dans la mesure où j'ai pu démontrer, toujours dans le [Chapitre 6](#), que le système était capable de reproduire le style propre des traductaires ainsi que des passages d'œuvres tirées du corpus d'apprentissage, posant de réels risques de plagiat que soulevaient ŞAHIN et GÜRSES (2019). Par la suite, ma collaboration avec Nathalie Serval dans un scénario de post-édition réaliste a également mis en évidence les lourdes modifications nécessaires pour s'appropriier un texte produit par la machine, malgré les contraintes générées en ce sens par cette interaction (cf. [Chapitre 7](#)). Finalement, donc, si les textes permettant d'entraîner la TA sur de la littérature et ceux résultant de la post-édition sont protégés (du moins en théorie) par le droit d'auteur, la traduction automatique individualisée telle que je l'ai proposée dans cette thèse constituerait le seul scénario légalement et éthiquement envisageable en littérature. Plus encore, l'utilisation de systèmes de TA personnalisés lèverait tout doute possible sur l'originalité du processus, sur le degré d'intervention des traductaires, et par conséquent sur la protection des œuvres, y compris en cas de traduction assistée par la machine.

Au vu des corpus circulant librement sur le Web, cependant, cette idée pourrait paraître optimiste. C'est pourquoi un nombre croissant de recherches évoquent la possibilité de compenser les auteurs des textes originaux et les traductaires en cas d'utilisation imprévue de leur travail si celle-ci génère une valeur ajoutée (FORCADA 2023, p. 50 et 67; LACRUZ MANTECÓN 2023, p. 90). Et quoique l'anonymisation des données rende ceci difficile en pratique (ŞAHIN et GÜRSES 2023, p. 12), tout comme la difficulté de retrouver les traces d'une contribution individuelle dans un système auquel ont été fournies de vastes quantités de documents, cela invite tout du moins à repenser l'industrie et ses pratiques, d'une manière plus durable qu'elle ne l'est actuellement (MOORKENS et LEWIS 2019), que ce soit par des tarifs fixés à l'heure ou par une plus grande protection et un meilleur contrôle sur les ressources linguistiques qu'offrirait un commun numérique (*digital commons*), ainsi que le suggèrent MOORKENS et LEWIS (*Ibid.*). Une solution alternative immédiate pourrait être de modifier les contrats, comme cela a déjà été proposé pour les métiers de la voix¹⁵² et très récemment par l'ATLF¹⁵³, bien que certaines pratiques de l'industrie remettent en question l'efficacité d'une telle interdiction.

En un sens, le secteur littéraire se trouve dans une position peut-être privilégiée par le droit et le statut d'auteur dont jouissent les traductaires, mais aussi par l'importance accordée à leur style individuel, que le lectorat ou l'édition pourra considérer comme un gage de qualité. Dans

152. <https://lesvoix.fr/ia-alerte-vol-de-voix/>.

153. <https://atlf.org/mise-a-jour-du-modele-de-contrat-de-traduction/>.

ce contexte changeant, néanmoins, les traductaires littéraires devront veiller à s'assurer — le cas échéant — que la TA soit intégrée de façon transparente et volontaire, d'une manière qui ne porte pas atteinte à leurs droits et qui vise à soutenir plutôt qu'à contraindre le processus créatif. Les associations de traduction et les syndicats auront un rôle important à jouer à cet égard (SERVAL 2023a). MOORKENS (2017, p. 468) mentionne à ce titre divers cas d'exemple attestant du fait que les appels pour la protection du statut et de la propriété des traductaires ne sont pas nouveaux et que leur succès n'est pas sans précédent, y compris dans le domaine littéraire. Dans ce contexte, il est rassurant de voir que des associations comme l'ATLAS et l'ATLF (ATLAS et ATLF 2023), le CEATL (CEATL 2020 ; 2023) ou encore l'ITIA (cf. KENNY et HANSEN 2023) s'emparent de ces questions et brisent le tabou sur le sujet. Bien sûr, de telles actions ne doivent pas être menées de façon isolée et dépendent de la coopération avec l'ensemble du monde de la traduction (MOORKENS 2017, p. 469), mais aussi avec d'autres secteurs, à l'image des démarches entreprises par la déclaration de 13 organisations de professions créatives (CEATL *et al.* 2023) ou du collectif « En chair et en os » (2023), visant à promouvoir une approche humaine de la traduction.

8.2.3 Des enjeux dépassant le seul contexte de la traduction

Avec l'arrivée des grands modèles de langue (GML) tels que ceux qui composent la famille des GPT, la nécessité d'une réflexion autour de ces enjeux s'est faite plus grande encore et se pose à une échelle bien plus large qu'auparavant. Essentiellement, les GML utilisent exactement la même architecture (*Transformer*) sur laquelle se base également la TA et font usage des mêmes plongements lexicaux, à la différence cependant que la taille des systèmes ainsi que la quantité de données utilisées sont d'un ordre de grandeur autrement plus élevé et que le modèle encodeur-décodeur (cf. Figure 16) laisse place à un modèle entièrement décodeur axé uniquement sur la prédiction (autrement dit sur la génération, d'où le nom aujourd'hui d'« IA générative »). S'il y a de plus en plus d'opacité sur ces systèmes, quelques initiatives plus transparentes nous renseignent cependant sur les données mobilisées par ceux-ci, leur volume, les langues incluses, les domaines impliqués ou encore leur origine (cf. S. ZHANG *et al.* 2022 ; TOUVRON *et al.* 2023). Sur ce dernier point HENDERSON *et al.* (2023, p. 2) font remarquer que la plupart de ces données sont librement récoltées sur le Web, bien qu'elles soient protégées par défaut par le droit d'auteur, mais que plusieurs corpus de romans protégés circulent librement eux aussi et font partie de jeux de données utilisés pour entraîner ces systèmes, à partir desquels il est même possible d'extraire des œuvres entières ou plusieurs chapitres reproduits à l'identique (HENDERSON *et al.*

2023, p. 8)¹⁵⁴. Or, ces corpus de spécialité (littéraires, scientifiques, etc.) constituent une large frange des données d'entraînement (la majorité étant récupérée automatiquement sur le Web) et jouent certainement un rôle considérable dans la pertinence, la fluidité et la créativité des sorties générées par les GML.

La réflexion est néanmoins rendue plus compliquée par le manque de responsabilité attachée aux outils (BENDER *et al.* 2021 ; LARSONNEUR 2021, p. 265) et par le flou juridique qui entoure cette technologique, alors que l'on cherche actuellement à voir si les codes existants peuvent s'appliquer à ces usages ou s'il est nécessaire de les cadrer par de nouvelles lois. Pour cette raison, HENDERSON *et al.* (2023, p. 28) encouragent le monde de la recherche à prendre en considération les aspects éthiques et légaux de la propriété intellectuelle, concluant que même un usage raisonnable (*fair use*) respecté à la lettre aurait des effets non négligeables sur certains métiers créatifs. À ce titre, les débats qui entourent les GML ont au moins ceci d'intéressant qu'elles portent sur le devant de la scène des questions auxquelles font face les traducteurs depuis l'arrivée de la TA neuronale, et qu'elles ouvrent la discussion sur l'utilisation plus générale de la technologie en littérature. La traduction s'est en effet retrouvée aux prises avec ces enjeux bien avant l'arrivée du *deep learning*, rebaptisé « IA » depuis lors, mais ceux-ci ont gagné en importance avec la TAN. Continuant de s'adapter en conséquence, le secteur de la traduction constitue dès lors un modèle qui pourrait inspirer d'autres domaines plus récemment touchés par ces questions. À l'inverse, il ne fait aucun doute que les parallèles avec d'autres domaines pourraient eux aussi apporter matière à la réflexion. Pour ces raisons, donc, ces discussions doivent dépasser le seul cadre de la traduction et inclure notamment les auteurs, pour qui les effets de cette dernière génération d'outils se sont immédiatement fait sentir, comme en témoignent le nombre de soumissions frauduleuses écrites par des GML qui inondent les plateformes comme *Amazon* (ALRIC 2023) et les maisons d'édition (SATO 2023), parfois même sous le nom d'autres d'auteurs (TAPPER 2023). Les modèles de langue inquiètent ainsi les auteurs et journalistes, qui vivent de l'écriture, tandis que d'autres outils apparus en même temps que ceux-ci soulèvent également des interrogations pour les arts visuels et la musique, mais l'arrivée de cette technologie a surtout laissé place à un sentiment de déjà vu en traduction (REI *et al.* 2020). Finalement, les traducteurs se retrouvent paradoxalement dans une position quelque peu privilégiée, dans la mesure où ces débats et questionnements ont cours depuis des années déjà et pourraient à ce titre inspirer

154. L'un de ces corpus est un simple lien *Twitter*, référencé comme tel dans les bibliographies sans aucune autre forme de documentation, permettant de télécharger l'ensemble du catalogue d'une bibliothèque virtuelle clandestine.

d'autres secteurs professionnels qui se retrouveraient confrontés pour la première fois au spectre l'automatisation avec l'arrivée des GML et d'autres outils génératifs.

Pour la traduction en particulier, il ne fait aucun doute que l'industrie cherche à atteindre une solution entièrement automatique (*no-human-in-the-loop*), en particulier lorsque des prestataires tels que TAUS décrivent le modèle de post-édition (ou *human-in-the-loop*) comme un compromis peu désirable — ou, pour reprendre les mots exacts, « *a half-baked solution* » (TAUS 2022). On peut évidemment rapprocher cette manière de voir les choses d'un courant plus général d'algorithmisation de la société, poussé par le néo-libéralisme, qui tend à ce titre vers l'automatisation et la précarisation du travail (ERTZSCHEID 2019), mais qui s'étend donc aussi au-delà de la traduction. En traçant un parallèle avec le journalisme, par exemple, où certains médias ont opté sans hésiter pour le tout automatique, il n'aura pas fallu attendre longtemps avant de voir paraître des articles criblés de plagiat et d'erreurs factuelles (CHRISTIAN 2023a), ou des publications s'apparentant à de la pure désinformation, notamment sur un sujet lié à la santé (CHRISTIAN 2023b), dont les plateformes garantissaient pourtant la vérification. Ainsi, les outils ne font que renforcer des tendances déjà existantes dans les diverses industries, telles qu'elles avaient déjà cours pour le monde du journalisme par exemple (*Ibid.*), mais qui permettent de prédire aisément les scénarios les plus probables si la TA, les GML ou d'autres technologies venaient à intégrer le secteur littéraire, ou d'autres, pour les mauvaises raisons. C'est à ce titre qu'un nombre grandissant de voix proposent aux professions concernées de s'organiser pour se faire entendre et s'opposer au déterminisme algorithmique. Bien qu'il soit loin d'être parfait (cf. VINCENT 2022), le paradigme dans lequel s'intègrent les GML à la programmation laisse entrevoir quant à lui un exemple de coopération fructueuse, dans son interaction, entre l'humain et la machine, que ce premier peut solliciter pour générer des fragments de code, pour convertir celui-ci d'un langage à un autre et même pour améliorer le code ou y pointer des erreurs (IFTIKHAR 2023). Dans un tel cas, l'objectif de la machine est avant tout d'aider l'humain et de libérer une partie de sa charge de travail, non pas de surveiller ou de pousser à la rentabilité et à l'asservissement. Pour y parvenir, il faudra notamment veiller à protéger le travail et les conditions de travail des traducteurs, à miser sur la transparence et à favoriser la coopération, à la fois entre l'humain et la machine, mais aussi entre le monde de la traduction, de l'ingénierie et d'autres domaines potentiellement concernés par ces enjeux.

8.3 La traduction littéraire automatique : un changement de paradigme

À la lumière de ces discussions, j'aimerais avancer dans cette ultime section un changement de paradigme pour la TA qui pourrait, me semble-t-il, s'avérer tout aussi bénéfique en dehors du champ littéraire. L'approche détaillée tout au long de ce travail propose en effet d'autres modalités de mise en œuvre, par le biais d'outils adaptés à un niveau pointu et axé sur l'humain, dont l'utilisation est envisagée comme une aide et non comme un substitut. Cette conception de la technologie s'oppose ainsi au déterminisme technologique, par lequel « la technologie contribuerait activement à transformer la société et la culture. Au contraire l'humain regagne son rôle d'agent dans le choix, l'acceptation, le rejet et l'interprétation des artefacts techniques » (RUFFO 2018, p. 128). Un tel changement de paradigme, que j'établis ici pour la TA mais que l'on pourrait aisément étendre à l'ensemble des technologies de la traduction (HANSEN 2024), pourrait avoir en ce sens un rôle déterminant sur la réception et l'adoption de tels outils, si l'on garde à l'esprit que les traductaires ne sont pas réfractaires à la technologie en tant que telle, mais plutôt aux technologies qui ne sont pas adaptées aux enjeux particuliers de leur métier et aux « aspects humains » de leur travail (RUFFO 2018 ; KOSKINEN et RUOKONEN 2017 ; DAEMS 2022). Or, les avancées techniques réalisées ces dernières années permettent justement d'aller aujourd'hui dans cette direction, et j'espère que les appels récemment lancés à cet égard en traductologie favoriseront ce changement vers une technologie plus personnalisée (cf. O'BRIEN et CONLAN 2018).

L'idée, bien sûr, n'est pas de dire qu'il faudrait totalement rejeter la traduction automatique ni, au contraire, que la traduction humaine devrait être systématiquement aidée par la machine. Il y aura toujours un besoin pour l'un ou l'autre selon le scénario envisagé et il est préférable qu'il en soit ainsi. Entre les deux extrêmes de ce continuum proposé par John HUTCHINS et Harold SOMERS (1992, p. 147–159) se trouvent cependant plusieurs scénarios possibles de traduction assistée par ordinateur, privilégiant plus ou moins le côté humain ou automatique. Dans ce cas, l'idée serait avant tout de proposer une utilisation des outils en traduction littéraire qui se rapproche plus de la traduction humaine assistée par la machine (THAM) que de la traduction machine assistée par l'humain (TMAH)¹⁵⁵, comme le défend similairement Philippe LACOUR (2019).

155. Traduction de *machine-aided human translation* (MAHT) et de *human-aided machine translation* (HAMT).

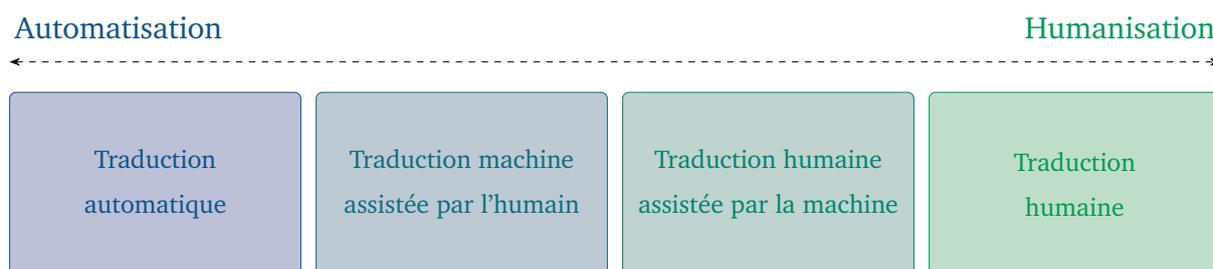


FIGURE 34 – Continuum de la traduction assistée par ordinateur

Comme je l'ai évoqué au fil des chapitres, les outils de traduction peuvent effectivement être repensés, à commencer par la TA elle-même pour laquelle cette méthode d'adaptation individuelle « au niveau de l'humain » (*human-adapted MT*) constitue une avenue prometteuse, mais aussi par l'intermédiaire de son intégration dans l'environnement de travail et par la manière dont celle-ci influence la dynamique d'interaction humain-machine. Dans un scénario peu idéal de post-édition, plus proche de la TA assistée par l'humain, l'accent serait généralement mis sur la productivité et la rentabilité d'une révision de TA brute (entièrement pré-traduite) dans une interface généralement peu adaptée à la tâche ¹⁵⁶ ; ou, dans le pire des cas, à partir d'une simple relecture du texte cible sans comparaison possible avec la source. Pour favoriser un scénario plus désirable et plus proche de la TH assistée par la machine, la TA (personnalisée) devrait être vue comme une partie intégrante de la TLAO, implémentée aux côtés d'autres outils dans un environnement sur mesure (personnalisable) et soumise au contrôle de l'utilisateur, d'une manière qui ne soit pas contraignante mais qui encourage l'interactivité, tout en soutenant la démarche créative et la prise de décision. Le recours à des systèmes de TA personnels, entraînés par les traducteurs sur leurs traductions et adaptés à leur style individuel et à leurs domaines de prédilection, leur permettrait par conséquent de rester maître du processus comme du produit final, ainsi que l'envisageait Nathalie lors de notre expérience (cf. [SERVAL 2023a](#)), et de se concentrer sur les passages critiques ou les aspects les plus importants (tels que la créativité, la réécriture, la mise en œuvre de stratégies propres à un genre donné ou l'affirmation de son style personnel) — sur la partie la plus humaine, finalement — de la traduction. Un tel changement de paradigme constituerait en outre un gage de qualité, dans la mesure où le système serait alors capable d'offrir des suggestions entièrement personnalisées. Plus encore, il renforcerait la

156. Nous avons vu dans le chapitre précédent que les outils de TAO les plus répandus sur le marché ne sont pas exempts de problèmes. Il n'est toutefois pas toujours possible d'opter pour le logiciel de son choix et de nombreuses personnes se retrouvent souvent forcées en pratique à travailler sur des applications Web propriétaires où les segments défilent un à un, sans réelle vue d'ensemble sur le texte, et où des indicateurs de performance omniprésents poussent la personne à avancer le plus rapidement possible.

confiance, l'autonomie, la motivation et le bien-être des traductaires (O'BRIEN et CONLAN 2018, p. 91–93).

Si l'on adopte la perspective de l'édition dans son ensemble, ces systèmes de traduction automatique personnels et personnalisables pourraient être le seul moyen de limiter les problèmes éthiques posés par l'utilisation actuelle de la TA en littérature qui ont été soulevés dans la [Section 8.1.3](#). Sans revenir sur ces enjeux en détail, le paradigme considéré ici serait un moyen de contourner l'atténuation de la voix et la normalisation stylistique autrement induites par la TA (KENNY et WINTERS 2020) ou de limiter le risque qu'elle bride la créativité (GUERBEROF-ARENAS et TORAL 2022), ainsi que ceux qu'elle pose pour le statut et les conditions de travail des traductaires (TAIVALKOSKI-SHILOV 2019a). De plus, au vu de l'importance accordée au caractère personnel de l'œuvre dans la définition du droit d'auteur (REYNAUD et KLEIN 2018), la personnalisation des technologies de la traduction fait sens non seulement du point de vue créatif, ergonomique et éthique, mais aussi légal, puisqu'elle assurerait que la traduction soit reconnue et protégée comme une œuvre de l'esprit. Cette sécurité, en retour, limiterait assurément la généralisation de pratiques commerciales douteuses et l'atteinte portée à leur droit de propriété intellectuelle, contre lesquelles met en garde TAIVALKOSKI-SHILOV (2019b).

D'un point de vue technique, également, il ferait sens d'opter pour de petits modèles spécialisés, plus écologiques et plus faciles à adapter, favorisant par la même occasion la transparence et l'explicabilité de ces systèmes¹⁵⁷. Puisque le processus d'adaptation peut être simple et rapide, par comparaison avec la mise au point d'un modèle de bout en bout, il serait par ailleurs possible d'envisager un affinage à partir d'un grand modèle de langue pré-entraîné — étant donné que c'est là leur utilité première¹⁵⁸. Il faudrait néanmoins s'assurer que ce procédé soit réellement plus efficace, ce qui ne suffirait pas à résoudre les nombreux autres problèmes qui ont été posés en partie dans la [Section 8.2.1](#) et dont les considérations dépassent le cadre de ce travail consacré à la TA.

157. Face à la montée de modèles de plus en plus larges et opaques dont il peut être difficile d'interpréter, de comprendre ou de vérifier les sorties, une tendance opposée tend à favoriser l'utilisation de modèles explicables (*explainable AI*). La possibilité d'entraîner soi-même son outil, avec des données de qualité spécifiquement sélectionnées, réduit ainsi le risques de biais, d'erreurs, de plagiat, etc.

158. Avant l'apparition des GML sous la forme de chatbots multi-tâches, l'avantage de ces modèles était surtout de pouvoir être affinés par la suite sur des tâches spécifiques, permettant d'atteindre de meilleurs résultats rapidement grâce à la robustesse du modèle de départ et au fait que l'affinage peut se faire sur un volume bien plus réduit de données. De cette manière, un grand nombre de systèmes spécialisés sur une tâche particulière peuvent être obtenus à partir d'un unique modèle plus large.

Le sujet de la technique soulève en outre celui des compétences nécessaires pour mettre au point de tels outils. Bien que la démarche adoptée pour cette recherche demande un minimum de connaissances informatiques — de même qu'une puissance de calcul non négligeable, quoique l'affinage à partir d'un système générique serait envisageable —, le travail mené ici montre à tout le moins qu'elle est possible en pratique. De plus, la généralisation de ce que l'on appelle aujourd'hui « TA adaptative », qui était très peu répandue au commencement de ce travail, pourrait tout à fait éliminer cet obstacle. Celle-ci est à distinguer de la « TA interactive », parfois reprise sous le même terme, qui concerne quant à elle une fonctionnalité d'autocomplétion similaire à la saisie prédictive des anciens téléphones à neuf touches et qui s'adapte non pas aux traductions de l'utilisateur mais au texte tapé en temps réel dans le segment en cours de traduction¹⁵⁹. La traduction automatique adaptative, en revanche, désigne une fonctionnalité permettant à la TA de s'adapter en continu aux traductions validées par l'utilisateur dans son outil de TAO, donnant ainsi la possibilité de personnaliser son propre système de TA de façon tout à fait automatique, sans aucun besoin de compétences techniques. Outre l'avantage technique d'une telle solution, cette conception dynamique de la TA fait pleinement sens du point de vue des utilisateurs, si l'on considère que le style des traductaires évolue et s'affine au cours des années, tout comme la langue elle-même et les conventions du genre.

Aujourd'hui, de plus en plus de logiciels de TAO gratuits ou payants offrent l'une ou l'autre de ces fonctionnalités, voire une combinaison des deux, comme le proposait déjà *CASMACAT* (ALABAU *et al.* 2014), par exemple, pour la traduction statistique¹⁶⁰. Dans le cas de logiciels gratuits, l'utilisateur garde la main sur l'ensemble du processus, qui peut cependant passer par des intégrations de *plug-in* et des installations complexes, tandis que les logiciels payants facilitent grandement cette démarche, bien que l'opacité du processus d'adaptation (sur des serveurs distants) et le prix des logiciels ajouté à celui du service de TA (souvent par abonnement) constituent deux facteurs non négligeables de dissuasion. Il est dès lors très encourageant de voir apparaître des projets dont l'objectif est de faciliter l'entraînement de systèmes de TA et leur intégration dans les outils de TAO. Parmi ces initiatives, on retrouvera le projet MTUOC (OLIVER 2020), qui m'a permis d'utiliser mon système de TA adapté pour l'évaluation experte

159. Lorsqu'un nouveau segment est entamé, le système fournit une traduction automatique par-dessus laquelle doit écrire l'utilisateur, acceptant ou non la sortie de TA qui se modifie dès que la frappe s'écarte de la traduction proposée. S'il s'agit d'une fonctionnalité intéressante, il faudrait voir dans quelle mesure l'affichage constant d'une TA et l'obligation de taper par-dessus le texte déjà produit pourrait contribuer aux contraintes déjà relevées dans l'expérience de post-édition du chapitre précédent.

160. Cf. <https://www.casmacat.eu/>.

du chapitre précédent, ou encore le logiciel *OPUS-CAT*¹⁶¹, qui fournit pareillement un moyen d'entraîner et d'intégrer des modèles de TA de manière locale (sur sa propre machine), insistant l'aspect confidentiel et sécurisé de la procédure, et qui met par ailleurs à disposition des modèles génériques pré-entraînés à partir desquels les utilisateurs peuvent créer leur propre système personnalisé, le tout dans une interface visuelle intuitive (NIEMINEN 2021).

L'implémentation même de la TA, comme je l'évoquais, devrait idéalement être repensée elle aussi, à la faveur d'un dispositif qui favoriserait la démarche analytique et créative, puisque nous avons vu que celle-ci pouvait avoir précisément l'effet inverse (cf. Section 7.3.2). Dans la foulée, j'ai ainsi pu formuler quelques propositions concrètes en vue d'améliorer l'ergonomie des outils pour la traduction littéraire. Pour commencer, il est évidemment important que l'outil de TA puisse être intégré au logiciel de prédilection de l'utilisateur, ce qui devrait idéalement inclure les logiciels de traitement de texte, de façon à s'associer au mieux à ses habitudes de travail et aux outils qui en font déjà partie intégrante. Dans le cas d'un programme de TAO, par exemple, la TA ne doit pas se substituer ou faire concurrence aux mémoires de traduction, aux glossaires et autres fonctionnalités, mais compléter celles-ci (cf. Section 7.4.1). L'un des aspects de la post-édition classique qui contrevient assurément le plus aux habitudes de travail des traducteurs littéraires, par exemple, est la division du texte phrase par phrase, pour laquelle une segmentation par paragraphe bénéficierait autant à l'outil qu'à l'utilisateur (cf. Section 7.5). En plus d'offrir une plus grande liberté et une plus grande marge de manœuvre aux traducteurs sur le texte, elle donnerait effectivement plus de contexte au système et permettrait une plus grande réorganisation des phrases. Cette dernière remarque est confirmée par le fait que les corpus contenant plusieurs phrases par segment produisent de meilleurs résultats que ceux séparés phrases par phrase (cf. Section 5.3.3) et que l'étape de pré-traitement consistant à réaligner automatiquement les corpus génériques volumineux pour obtenir des segments les plus courts possibles fait baisser les résultats (cf. Section 5.3.4), mais aussi par les observations de l'évaluation du système adapté, qui montrent qu'il apprend de ce fait à traiter plusieurs phrases d'un même coup et qu'il est ensuite capable de réorganiser les phrases d'une manière similaire à celle employée par la traductrice (cf. Section 6.2.3).

La personnalisation, si elle aide à de nombreux égards, ne constitue pas non plus une solution miracle et devrait s'accompagner, d'une part, d'améliorations supplémentaires de ses performances, qui limitent encore aujourd'hui son utilité, et, d'autre part, d'une réflexion concernant

161. Cf. <https://helsinki-nlp.github.io/OPUS-CAT/>.

l'interaction humain-machine, qui peut ajouter une contrainte supplémentaire à l'exercice et enrayer le processus de traduction (cf. [Section 7.3.2](#)). Il faudrait par conséquent privilégier une utilisation TA plus interactive que la pré-traduction complète et automatique du texte, par laquelle j'entends notamment la possibilité de solliciter celui-ci sur demande, possiblement en marge de l'interface principale de travail, d'une manière qui n'entrave pas l'intuition, ne bride pas la créativité et n'atténue pas la voix des traductaires. Cette recherche, mes lectures et les discussions plus générales dans les sphères littéraires laissent ainsi entendre qu'il s'agit bel et bien d'un outil utile, mais surtout de manière localisée : pour vérifier la traduction d'un terme ou d'une phrase spécifique, rechercher une solution alternative ou susciter des idées lorsque le cerveau est en panne d'inspiration. Pour les mêmes raisons, il s'avérerait également profitable de pouvoir compter sur plusieurs systèmes, comme je l'ai proposé à Nathalie (cf. [Section 7.3.1](#)), afin de permettre la comparaison de différentes sorties de TA pour un même segment, ce qu'il est déjà techniquement possible de faire avec un même système en demandant au module de TA de fournir les n meilleures traductions, mais qui contraste avec le fonctionnement par défaut de la TAO, basé sur une solution unique. Dans le même ordre d'idée, les environnements permettant de modifier le plus librement possible les paramètres, l'affichage des outils intégrés et les différentes fenêtres du logiciel peuvent contribuer de manière significative au confort de travail. L'alternative libre représente un bon point de départ pour repenser les interfaces de travail dans cette voie, favorisant des environnements plus sobres et moins envahissants que la plupart des programmes de TAO du marché, comme je l'ai illustré par le biais du logiciel *OmegaT* — tout en sachant qu'il en existe de nombreux autres. Ces diverses réflexions sur les interfaces et les interactions de la TA mériteraient d'être prolongées, mais invitent à considérer la traduction automatique comme un outil parmi d'autres. Un outil qui n'est pas immuable et peut être pensé autrement, pour mieux soutenir le travail humain, en accord avec la prémisse de l'appropriation des technologies d'aide à la traduction (cf. [Section 2.3](#)).

Pour résumer, il est possible d'imaginer des changements simples, qui sont pour l'ensemble réalisables d'un point de vue technique, mais qui impliquent de repenser les outils de traduction et de sortir d'une vision strictement productiviste des technologies, afin de prendre en compte les besoins et les attentes des traductaires (littéraires ou non). Si celles-ci ont bel et bien évolué dans une optique de gain de temps et de rentabilité, force est de constater que les fondamentaux de la TAO au sens large ont rarement été remis en question et que les changements proposés s'inscrivent plus rarement encore dans une préoccupation d'ergonomie, de qualité, de créativité

ou simplement de plaisir de traduire. Concernant la dernière génération d'outils neuronaux (TA et GML), de nombreux obstacles s'opposent encore à une mise en place éthique et juste de la technologie, parmi lesquels se posent notamment la durabilité, la confidentialité, la responsabilité, l'équité, la transparence des systèmes, l'accessibilité, la sûreté et l'interactivité, qui inclut à la fois les possibilités de personnalisation de l'outil et d'interaction avec celui-ci (BARREDO ARRIETA *et al.* 2020). Dans le cas spécifique de la TA, le travail mené ici montre que des progrès restent à faire sur chacun de ces points.

Ainsi, en réponse à la question des conséquences probables de l'arrivée de la traduction automatique en littérature (SQR2), il est apparu que le secteur littéraire n'était pas invulnérable aux enjeux posés par l'automatisation et les technologies de la traduction auxquels font face les traductaires et les traductologues depuis des années déjà. Dans cette perspective, l'arrivée des GML aura au moins eu le mérite de forcer l'ouverture du débat, jusqu'alors limité à quelques exceptions près au champ de la recherche. La discussion construite dans ce chapitre met en lumière la nécessité de mener une réflexion de fond sur les rapports entre technologie et littérature, mais aussi de permettre aux traductaires littéraires d'acquérir une littératie traductique, que Lynne BOWKER et Jairo BUITRAGO CIRO (2019) appliquent au cas de la TA par le biais du concept de *MT literacy* (connaissance de son fonctionnement basique, de ses limites et de ses avantages, des erreurs récurrentes, des risques, etc.), mais que l'on peut aisément étendre à l'ensemble des outils pour permettre un usage « raisonné » de la technologie (LOOCK 2019). De fait, il y a effectivement des applications intéressantes de la TLA pour la société au sens large — en tant qu'aide à la diffusion ou aide à la lecture par exemple (cf. Section 8.1.2) — et pour les traductaires plus spécifiquement, surtout s'il est possible de faire intervenir la TA comme un assistant personnel, ainsi que je le propose. Cependant, l'évolution des technologies de la traduction et l'automatisation croissante des deux dernières décennies laissent également prévoir de nombreux risques si l'outil intervient uniquement avec une visée de remplacement ou de rentabilité (cf. Section 8.1.3). Dans ce contexte, il est essentiel de promouvoir une plus grande coopération entre la communauté de l'ingénierie linguistique et celle des traductaires, mais aussi impératif que les associations et syndicats se saisissent de ces questions pour veiller à protéger au mieux les droits, le travail et les conditions de travail des traductaires. Dans l'édition, plus particulièrement, il est nécessaire d'engager des réflexions parallèles avec l'ensemble des actants du réseau et de montrer qu'il est dans leur intérêt commun de privilégier la qualité face à la réduction des coûts ; de ne pas imposer leur utilisation pour payer moins, ce qui affecterait durablement tous les maillons de

l'édition comme nous l'avons vu, mais de repenser les outils et leur utilisation pour traduire mieux. Surtout, il faut exiger la transparence, ainsi l'ont déjà revendiqué diverses associations. Avec de meilleurs systèmes, des environnements de traduction sur mesure et des possibilités de formations, la traduction littéraire automatique personnalisée pourrait offrir un changement de paradigme qui viserait à soutenir plutôt qu'à entraver le processus de traduction, la créativité et la qualité. Ce qui est sûr, c'est qu'elle contribuerait à promouvoir un usage plus humain de la TA, et peut-être à renforcer l'agentivité des traducteurs dans leur adoption et leur utilisation de la technologie.

Conclusion

Il y a, dans les discussions ouvertes tout au long de cet ouvrage, plusieurs mouvements cycliques. À commencer par l'objection préjudicielle lancée tour à tour aux outils informatiques en traduction littéraire (cf. [Chapitre 2](#)), mais que l'on retrouve aussi dans le choix des métaphores et dans les discours médiatiques qui entourent le progrès technologique ([VIEIRA 2020b](#)) ou encore dans les retombées de ces nouveaux outils sur la société ou sur le monde de la traduction, comme j'ai pu le montrer dans le chapitre précédent. De manière générale, d'ailleurs, tout ceci est au moins aussi prévisible que les développements technologiques eux-mêmes : « Rétrospectivement, on se rend compte qu'il n'y a pas de développement majeur dans le domaine qui, au moins dans ses grandes lignes, n'aurait pas pu être prédit dès les origines de l'informatique » ([CHUKWU 1994](#), p. 79). Si l'observation valait il y a quelques années déjà, il en est de même pour l'apprentissage profond, dont le début des travaux remonte à près d'un siècle, ou encore pour l'apparition plus récente dans la sphère publique des outils génératifs, que l'on qualifie triomphalement d'« intelligence artificielle ». Là aussi, il s'agit du fruit du long et patient travail de l'ingénierie, de la lente évolution de la recherche, de l'amélioration constante du matériel et de l'application de méthodes nouvelles, qui font graduellement grimper les performances de ces systèmes. Le cas de la traduction automatique (TA) est un parfait exemple de cette évolution, dont les périodes successives d'engouement à l'arrivée de nouveaux paradigmes et les prédictions de pertes d'emploi ont finalement cédé la place à l'émergence de pratiques auparavant inexistantes et à une utilisation croissante de la TA comme outil supplémentaire dans l'arsenal de la traductique ([WAY 2013](#) ; [KOEHN 2020](#), p. 29–30 ; [RAGNI et VIEIRA 2021](#), p. 13). Une chose qui a changé avec l'apparition du paradigme neuronal, cependant, est l'augmentation exponentielle des recherches scientifiques mêlant traduction automatique et traduction littéraire, dont cette étude a tenté de dresser un portrait aussi complet que possible dans le contexte plus général de la traduction littéraire assistée par ordinateur (TLAO). Partant de la mise au point d'un système de TA adapté à la littérature, j'ai ainsi voulu poser un regard objectif sur la question de la traduction littéraire automatique (TLA), tout en adoptant une approche critique de la technologie qui ne nie pas ses avancées progressives, mais ne perd pas de vue ses limites et place les interfaces de l'interaction humain-machine au cœur de sa démarche, ses enjeux sociétaux et, surtout, ses utilisataires ([AGRE 1997](#) ; [FRIEDMAN 1997](#) ; [HARRISON et al. 2011](#)).

Cette vision pragmatique du *deep learning*, motivée par la nécessité d'une approche raisonnée de la technologie (cf. [Chapitre 3](#)), m'a mené à envisager un scénario de traduction littéraire automatique individualisée qui se situe dans le prolongement de l'état de l'art (cf. [Chapitre 4](#)) et a été synthétisé dans une première question de recherche (**QR1**) : Est-il possible d'adapter un système de TA au style individuel des traducteurs ? Dans la foulée, ce travail montre que s'il est bel et bien possible d'adapter des systèmes de traduction automatique à la littérature (cf. [TORAL et WAY 2018](#) ; [MATUSOV 2019](#) ; [KUZMAN et al. 2019](#) ; [TORAL et al. 2024](#) ; [OLIVER 2024](#)), le gain de performance est bien meilleur lorsque l'on adopte un niveau de précision plus élevé encore et que l'on personnalise système en question à l'échelle du style individuel d'un traducteur ou d'une traductrice. Le chapitre méthodologique (cf. [Chapitre 5](#)) illustre de cette manière le processus d'affinage de la TA, offrant une démonstration de bout en bout à partir du module *OpenNMT* ([KLEIN et al. 2017](#)) et soulignant l'importance des corpus à cet égard, tandis que le chapitre consacré aux évaluations atteste plus spécifiquement de l'amélioration significative observée pour chacun des aspects étudiés, mettant en évidence que le système est beaucoup plus performant, plus proche du style et des choix de la traductrice et plus riche d'un point de vue lexical, mais aussi moins proche de la source d'un point de vue syntaxique (cf. [Chapitre 6](#)).

Dans ce même chapitre, j'ai cherché à offrir un aperçu concret et une évaluation détaillée des traductions générées par le système — chose qu'il n'est pas souvent donné de voir dans les recherches et que reflètent difficilement les métriques dont j'ai pu au passage présenter les limites — en vue de répondre à une sous-question de recherche (**SQR1**) : Quelles sont les performances d'un outil adapté à la fois à la littérature et au style des traducteurs ? Il a ainsi été possible de noter que les traductions générées par celui-ci sont généralement plus en accord avec la terminologie du genre et avec les termes employés dans la référence, et qu'elles posent peu de difficultés concernant les accords, les temps de la narration, le respect des conventions typographiques de l'ouvrage, la cohérence lexicale ou la cohésion référentielle et relationnelle. De façon surprenante, l'analyse révèle par ailleurs que le système donne lieu à des réagencements syntaxiques allant la plupart du temps dans le sens de la traductrice (y compris des réorganisations de plusieurs phrases) et qu'il reproduit en outre des procédés de traduction dépassant le simple niveau lexical (modulations, transpositions, contractions), ainsi que des stratégies opérées à un niveau plus abstrait encore, comme le prouve l'exemple des omissions. À l'inverse, l'annotation d'erreurs qui a permis de tirer ce constat révèle également qu'un grand nombre de difficultés subsistent, apportant une confirmation supplémentaire que nous sommes encore loin de la « parité » entre

l'humain et la machine. Sans surprise, la machine peine aux endroits où il est nécessaire d'apporter une solution originale, d'interpréter le texte ou d'avoir une vue d'ensemble sur le récit, mais le système adapté produit finalement peu d'erreurs relevant des catégories littéraires définies dans le grille d'évaluation, à l'exception des erreurs de style dont la forte proportion atteste de la littéralité marquée de la TA encore aujourd'hui. Les résultats indiquent dès lors qu'un tel outil pourrait intervenir non pas à des fins publication mais plutôt de compréhension, bien que les problèmes de sens constituant la majorité des erreurs accentuent le caractère indispensable de l'intervention humaine. De fait, si les sorties de traduction affichent peu de problèmes de fluidité, comme l'ont signalé les études comparant TA statistique et neuronale, les cas d'omission, de sous-traduction, de mauvais choix lexicaux, de glissements de sens et de non-sens trahissent la mauvaise adéquation de l'outil, qui pose donc un risque non négligeable si les traductions ne sont pas vérifiées ou sont présentées sans comparaison possible avec la source.

Cette conclusion met avant tout en évidence la notion d'un outil qui serait au service des traductaires littéraires, capable de fournir des suggestions utiles à la volée, mais cette conception de la TA amène par conséquent une seconde question de recherche (QR2), motivée quant à elle par la volonté de remettre les traductaires au centre de la recherche, à savoir : La traduction automatique est-elle utile pour les traductaires littéraires ? Le [Chapitre 7](#) propose dans ce sens une évaluation fondée sur la tâche, impliquant une expérience de traduction assistée par ordinateur (TAO) en contexte réel ainsi que la participation de la traductrice pour qui a été adapté le système de TA, Nathalie Serval, afin d'évaluer les effets de l'interaction humain-machine sur le processus de traduction et d'anticiper ceux-ci dans la perspective plus large de l'ergonomie. Cherchant à cerner dans quelle mesure et dans quelles circonstances un tel outil peut venir en aide à celles et ceux qui seraient susceptibles de l'utiliser au quotidien, les entretiens successifs menés dans cette seconde partie de thèse viennent appuyer l'intérêt du processus d'adaptation et l'utilité potentielle de la TA, mais ils mettent surtout en lumière les contraintes liées à l'environnement de travail et aux modalités d'interaction avec la TA, qui entravent au contraire l'intuition et perturbent le *flow* de la traduction, ce qui m'a amené par la suite à imaginer différentes mises en place de l'outil. Dès lors, si ma première question de recherche confirme que la thématique de la TLA est bien d'actualité et qu'il est effectivement possible de repenser les outils pour améliorer leurs performances sur certains aspects traditionnellement mis en avant comme des éléments typiques de la littérature et hors de portée des systèmes (temps de la narration, prise en charge des dialogues, cohérence lexicale, respect du registre), de même que sur une série de critères

mentionnés quant à eux comme des limites majeures de la TA (réorganisations syntaxiques, divisions et fusions de phrases, modulations, explicitations, reformulations), cette deuxième question de recherche pointe les limites que soulève l'utilisation actuelle de la TA en littérature sur le plan cognitif et ergonomique, mais aussi le besoin d'impliquer les traductaires dans la conception des outils et d'adapter plus efficacement ces technologies à l'humain. Cette évaluation experte apporte par là même une nuance essentielle aux métriques automatiques, qui marquent principalement les améliorations relevées dans les sorties de traduction sans rendre compte de la contribution de l'outil dans la réalisation de la tâche, et de manière générale les évaluations fondées sur le produit, qui ne prennent pas en compte les éléments textuels et extra-textuels de la traduction et se satisfont d'une qualité acceptable pour conclure à la parité humain-machine. L'étude détaillée ici rejoint néanmoins un nombre de publications allant dans cette première direction (MOORKENS *et al.* 2018 ; GUERBEROF-ARENAS et TORAL 2020 ; 2022 ; VIEIRA *et al.* 2023) et un nombre plus restreint encore de travaux s'attachant à décrire l'influence de la TA sur les processus cognitifs de l'activité traduisante (KENNY et WINTERS 2020 ; KOLB 2024). Ce faisant, j'ai pu soutenir l'idée que l'outil évolue fondamentalement dans une optique de qualité et de rendement, à l'image de la TAO dans son ensemble, mais qu'il a très peu bénéficié à l'heure actuelle de changements qui profiteraient directement aux traductaires. J'ai en effet évoqué à plusieurs reprises le manque de coopération entre le monde de l'ingénierie et celui de la traduction, le déséquilibre entre l'évolution technique des programmes et celle de leur utilisation ou de leur interface, tandis que le domaine littéraire n'a pour ainsi dire aucunement bénéficié des technologies de la traduction (cf. Sections 1.3, 6.4.1 et 7.3). Il y aurait pourtant beaucoup d'améliorations souhaitables d'un point de vue ergonomique, dont quelques-unes sont exposées dans ces mêmes sections et dont le milieu professionnel de la traduction aurait beaucoup à attendre, comme le conclut similairement François YVON (2019, p. 66). Dans cette optique, il serait intéressant de voir paraître d'autres études à l'interface entre TA et TAO, qui chercheraient à réévaluer les environnements et les interactions humain-machine.

Après m'être intéressé au processus de traduction et à la mise en place concrète du dispositif de TA, je me suis penché dans le dernier chapitre sur la dimension organisationnelle de la technologie du point de vue de l'édition dans son ensemble (cf. Chapitre 8), en réponse à une deuxième sous-question de recherche (**SQR2**) : Quelles sont les conséquences de l'arrivée de la TA dans le domaine littéraire ? La réflexion engagée dans ce dernier chapitre fait apparaître dans ce contexte de multiples défis sociétaux liés tout d'abord à sa présence déjà existante dans le secteur littéraire et au manque de transparence associé à ces pratiques, de même qu'aux

sources d'inquiétude posées par la technologie concernant la qualité des traductions, le statut des traductaires et leurs conditions de travail. Ce tableau fait apparaître en définitive une image fortement polarisée, dépeignant d'un côté les avantages possibles pour les traductaires, pour autant que la personne qui traduit reste maître du processus et, plus encore, qu'elle soit libre d'utiliser, voire d'entraîner son propre système ; et, de l'autre, un plus grand nombre de risques liés à l'appropriation du travail humain et à l'utilisation de la technologie pour des raisons purement économiques, qui refléteraient en majeure partie l'évolution des technologies de la traduction jusqu'à aujourd'hui. Ces enjeux vont de pair avec le danger d'une invisibilisation du travail humain, à la fois dans la mise au point des systèmes, qui requiert pourtant une intervention à tous les stades du déploiement, mais aussi dans l'utilisation qui en est faite par la suite, comme en témoigne le lien établi dans l'enquête de l'ATLF (2022) entre l'usage de la TA et la perte de reconnaissance associée à la publication. Or, ces deux tendances doivent impérativement être renversées pour assurer la bonne continuation de la recherche et une intégration fructueuse de la machine aux compétences humaines. Il sera par ailleurs utile de poursuivre les débats qui ont été ouverts au sein de la profession, dont fait état le chapitre précédent, mais aussi d'y développer une forme de littératie (cf. BOWKER et BUITRAGO CIRO 2019) en formant le secteur à ces enjeux (SLESSOR 2020, p. 249 ; LARSONNEUR 2023, p. 92), et de s'assurer que l'investissement des traductaires dans leur formation et leur équipement technique ne se fasse pas au bénéfice des autres, comme le mentionnait déjà Françoise Cartano dans ce qui semble être l'histoire immuable de la traduction (littéraire) : « [L]es mises en garde que nous faisons ne sont pas des mises en garde contre l'informatique, mais contre le fait que le bénéfice, financier notamment, de notre équipement informatique aille à l'éditeur exclusivement » (MALROUX *et al.* 1988, p. 144). Pour la même raison, il faudra enfin veiller à défendre au mieux les droits des traductaires, et peut-être même trouver un moyen de les compenser lorsque leur travail est utilisé pour entraîner des systèmes qui créent ensuite de la valeur (FORCADA 2023, p. 50 et 67 ; LACRUZ MANTECÓN 2023, p. 90 ; ŞAHIN et GÜRSES 2023, p. 12).

Le secteur est en train de vivre des changements certains. Sur les quatre années de ce travail, le sujet de la TLA est passé d'une prétention risible à une question d'actualité brûlante. De fait, la TA fonctionne de mieux en mieux, au point qu'un nombre visiblement croissant de maisons d'édition y ont maintenant recours pour des projets de post-édition de textes littéraires. Cela ne veut pas dire pour autant que le secteur doit se résigner à passer le cap de la TA ou que le métier de traductaire littéraire tel qu'on le connaît est amené à disparaître, mais qu'il est nécessaire de se saisir aujourd'hui de ces débats qui seront déterminants pour l'avenir, et qui rejoignent des

questionnements sociétaux plus larges concernant l'automatisation des métiers. Dans ces débats, il sera notamment important de distinguer deux situations, telles qu'elles se posaient déjà pour les mémoires de traduction (MT) :

Il importe, quand il est question des MT, de distinguer deux situations très différentes, sous peine de brouiller sérieusement le débat :

— Une MT, lorsqu'elle est au service du traducteur qui en contrôle le contenu et qui a toute latitude dans son utilisation, est un précieux outil de consultation et un puissant facteur de productivité. Son utilisation n'a rien d'aliénant pour le traducteur.

— En revanche, telle qu'elle utilisée par les grandes agences de traduction, la MT se présente comme un outil de « STAO » (sous-traitance assistée par ordinateur), qui crée une situation où la pleine maîtrise du travail peut finir par échapper au traducteur. (BÉDARD 2017, p. 8.)

Dans ce deuxième cas de figure, « le traducteur comprend alors que la traduction du texte ne lui appartient pas vraiment, qu'il n'est en quelque sorte qu'un "invité" dans un processus dont il n'est pas le véritable maître » (*Ibid.*, p. 11). Car, si les outils n'imposent pas d'utilisation a priori et peuvent être détournés pour mieux servir le travail humain, comme j'ai tenté de le montrer par cette recherche, celle-ci s'est également développée d'entrée de jeu avec l'idée que ni leur développement ni leur mise en place ne sont neutres (cf. p. 31). Rejetant la notion de « neutralité technologique » (TEHRANIAN 1990), qui fait abstraction des intentions idéologiques ou politiques et des facteurs sociaux entourant l'introduction des dispositifs techniques (MOROZOV 2013 ; KENNY 2017), les discussions soulevées en dernière partie d'ouvrage et alimentées par de nombreux travaux antérieurs indiquent qu'au fil des outils, l'application de la technologie par les prestataires de service s'est graduellement orientée vers ce deuxième cas de figure. Or, cette crainte est précisément celle formulée par la traductrice Nathalie Serval, lors de nos entretiens, vis-à-vis de la traduction automatique actuellement et de la perte du statut d'artiste-auteur qu'entraînerait un usage non raisonné de la technologie (cf. Section 7.3.2 ; SERVAL 2023a). De la même manière que pour les outils qui l'ont précédée, l'« IA » nous est vendue comme un moyen de réduire la charge de travail et de rendre celui-ci plus humain, mais tend surtout en réalité vers une précarisation des conditions de travail, dans la continuité du modèle de taylorisme numérique qui s'est instauré progressivement et durablement en traduction (MOORKENS 2020).

Est-ce à dire que la TA pourrait remplacer l'humain dans un futur proche ? Non, évidemment. Tout comme il serait illusoire de penser que toutes les pages Web ou tous les articles scientifiques pourraient être traduits par des humains. Il s'agit probablement là du seul point sur lequel

s'accorde l'ensemble des travaux mentionnés dans ce travail, que rejoignent mes conclusions : la traduction automatique est encore loin de la parité humaine et la littérature en constitue un exemple éclairant. Comme le mentionnait Nathalie SERVAL (2023a) au cours de notre entretien, l'outil reste incapable de synthétiser des idées ou des phrases, de dépasser la littéralité du texte et de le comprendre ou de porter sur celui-ci une vision d'ensemble. Contrairement à la machine, l'humain peut jouer sur la forme, proposer des adaptations ou des néologismes, transposer ou expliciter les éléments culturels, s'écarter volontairement de la source, ajouter du contenu, avoir la sensibilité nécessaire pour modifier l'angle ou l'ordre du récit, s'adapter aux desiderata des commanditaires... On pourrait rapprocher l'idée des trois niveaux de traduction de Piotr BLUMCZYNSKI (2016), qui au gré de ses comparaisons herméneutiques, décrit un état incarné et conscient (somatique) de l'activité traduisante, opposé à son niveau élémentaire mécanique (cf. Section 7.2.2) :

À ce troisième niveau, la traduction devient avant tout une affaire de questions, y compris philosophiques. Non pas uniquement des questions à propos du sens des mots ou de la traduction la plus appropriée d'un texte donné, mais à propos du sens de l'activité dans son entièreté. À propos des aspects éthiques d'un projet en particulier, d'une tâche réalisée pour tel client, du contenu qui nous occupe et de ses effets sur nous, peu importe s'ils nous inspirent ou nous inquiètent [43]. (*Ibid.*, p. 53–55.)

La traduction (littéraire), de fait, reste une activité d'appropriation personnelle et subjective du texte, au moment tant de sa lecture que de sa réécriture par la personne qui traduit (PONCHARAL 2021, p. 72). La machine, au contraire, est « asomatique » (ROBINSON 1991, p. 23), et ne peut prétendre à la compréhension, à l'analyse ou à la réflexion. Ainsi, dans l'évolution actuelle de la technologie, « le rapprochement entre la TA et la traduction automatique de haute qualité relève de l'asymptote, puisque la paradigme dominant repose sur l'encodage et le décodage d'une suite de mots ou de segments, là où la communication humaine doit interpréter les signaux culturels et contextuelles pour assurer une compréhension mutuelle [44]. » (MOORKENS 2017, p. 471). Ces remarques suffisent à elles seules à garantir que la traduction humaine reste nécessaire et le restera encore jusqu'à ce que la machine soit capable de s'écarter d'un texte source avec une conscience réflexive sur les implications tant sémantiques que formelles de la langue.

Ceci est d'autant plus valable pour une machine qui excelle dans l'apprentissage et la généralisation à partir des données, ainsi que pour un domaine tel que la littérature, dont la complexité se marque précisément par sa disposition pour l'irrégularité et sa prédilection pour les aspérités de la langue, pour ses écarts de langage (style parlé, variations entre le soutenu et l'informel,

néologie, détournements esthétiques...) et sa réflexion sur le contenu (double lecture, non-dits, engagement émotionnel, respect ou non des conventions...), mais aussi par l'omniprésence d'éléments extratextuels qui restent encore hors de portée de la TA. De fait, la machine est particulièrement utile pour tout ce qui a trait à la forme, ce qui explique par ailleurs qu'elle puisse se rapprocher du style individuel d'une personne, mais on ne peut lui reprocher de ne pouvoir prendre de décision ou comprendre ce qu'elle traduit : c'est un programme (BENDER et KOLLER 2020). Un programme qui peut être entraîné et fournir des propositions plus utiles si les données utilisées sont pertinentes. Pour toutes ces raisons, Thierry POIBEAU (2022, p. 171) en vient à conclure : « les systèmes de traduction automatique ne sont pas près de mettre les traducteurs littéraires au chômage, et il est bon qu'il en soit ainsi ». Cela dit, la question reste de savoir si la TA peut être utile ou non, pour les traducteurs littéraires, et sous quelles conditions. Dans cette optique, il convient de se demander, en outre, si le rendement vaut réellement le sacrifice de la qualité de la traduction et du plaisir de traduire (cf. RAGNI et VIEIRA 2021, p. 15) ou si l'on peut au contraire concevoir des outils qui soutiennent ce plaisir, favorisent la démarche créative, renforcent la prise de décision et assistent plus efficacement le processus de traduction.

Il y a une tendance croissante, depuis quelques mois, à déprécier les traducteurs ou toute autre personne qui s'opposerait à la « révolution » technologique en invoquant la révolte luddite, mais en oubliant toutefois que le mouvement s'opposait en premier lieu à la dégradation des conditions de travail et non à la machine (CARMO 2023). Au contraire, Kaisa KOSKINEN et Mina RUOKONEN (2017, p. 22) démontrent que les traducteurs sont tout à fait réceptifs à la technologie, pour autant que celle-ci leur permette de travailler plus efficacement. Or, l'objection soulevée à l'encontre de la TA provient précisément du fait qu'il n'en est pas toujours le cas — nous avons vu qu'elle peut contraindre le processus et que son avantage premier, à savoir le gain de temps, ne se vérifie pas nécessairement —, mais elle tient aussi à l'utilisation actuelle des outils (et de la TA en particulier), qui favorise la réduction des coûts plutôt que le confort de travail ou la qualité de la traduction.

Il y a une différence entre une objection formulée contre la technologie ou contre les risques perçus de la technologie. Celle-ci est fine, mais importante. Dans le premier cas, les technologies de la traduction devraient être remplacées, améliorées ou purement supprimées pour résoudre le problème, tandis que, dans le second, cette résolution passerait par un changement des pratiques plutôt que de la technologie. Ces deux possibilités ne sont pas mutuellement exclusives : l'amélioration des conditions de travail passera aussi probablement par des changements du point de vue de la technologie et des pratiques [45]. (VIEIRA 2020a, p. 8.).

Il est en effet peu étonnant que des spécialistes — déjà sous-payés en comparaison avec leurs homologues dans d'autres domaines (TAIVALKOSKI-SHILOV 2019a, p. 692) — refusent de travailler avec un nouvel outil si le seul résultat attendu est une baisse de leur salaire ou l'un des nombreux autres risques soulevés dans le dernier chapitre. J'espère pourtant avoir établi par le biais de ce travail de recherche que l'on peut envisager d'autres scénarios et replacer le curseur plus près de la traduction humaine assistée par la machine que de la traduction machine assistée par l'humain.

Dans la perspective ambivalente des technologies de la traduction décrites par le concepteur du logiciel *LogiTerm* Claude BÉDARD (2017), le changement de paradigme que je propose par le biais de la traduction (littéraire) automatique individualisée prend l'exact contre-pied du modèle de la « sous-traitance assistée par ordinateur », à la faveur d'un scénario privilégiant la personnalisation, la liberté d'utilisation, la maîtrise du processus et le dialogue avec la machine. Celui-ci est par ailleurs tout à fait en accord avec les trois solutions proposées par Paola RUFFO (2021, p. 209) pour une intégration efficace de la technologie : « une collaboration constructive entre les traducteurs littéraires et les tierces parties, dans le cadre de laquelle les outils ne sont pas imposés, une prise en compte des conditions de travail, ainsi qu'une approche permettant de soutenir plutôt que d'entraver le flux de travail [46]. » Là où les avancées en matière de traduction automatique font état de systèmes toujours plus grands et d'un volume de données toujours plus volumineux, je présente au terme de cette recherche l'alternative d'outils plus petits, plus spécialisés et plus transparents (cf. Section 8.2.3). Les expériences menées révèlent que malgré les difficultés soulevées par la traduction de textes créatifs, ainsi que la situation de départ visiblement défavorable due à la complexité de l'ouvrage choisi (cf. Section 5.2.3), l'adaptation permet d'obtenir de bien meilleurs résultats, mais surtout si le système est entraîné par les traducteurs sur leurs propres productions. Dans ce cas de figure, la mise en place concrète ne peut qu'être celle d'un système au service des traducteurs, qui serait soumis à leur contrôle et leur permettrait de conserver la propriété sur leurs données, résolvant par là même nombre de difficultés éthiques et juridiques associées à la TA. Ajoutée à une utilisation plus dynamique et interactive de l'outil, cette perspective laisse paraître selon moi une autre façon d'envisager la TA, plus proche du travail de l'humain, de son style et de ses traductions. S'il devenait possible à l'avenir de s'équiper avec de tels outils adaptatifs, comme le propose par exemple *OPUS-CAT* (NIEMINEN 2021), le système serait alors capable de faire des choix plus pertinents, de fournir des solutions de traduction utiles à l'humain et peut-être, qui sait, de se tailler un jour une place dans l'arsenal des traducteurs littéraires.

Bien sûr, ce changement de perspective n'est pas une solution miracle : la mise à l'essai de l'outil développé dans cette étude a mis en lumière ses limites, que j'élaborerai en partie ci-après, signalant que l'utilité de la TA dépend encore d'améliorations en matière de qualité des sorties et d'ergonomie. Dans un contexte où l'outil continue d'évoluer au jour le jour et où d'autres recherches apportent des conclusions encourageantes, la TA se profile tout de même comme une aide intéressante. Enfin, si le cas littéraire s'est imposé dans cette étude, cela n'invalide aucunement l'application de ces conclusions aux autres domaines de la traduction.

Bien que j'ai tenté dans ce travail de broser un panorama aussi large que détaillé, plusieurs limites subsistent, notamment en ce qui concerne l'évaluation. Si j'ai en effet cherché à diversifier les méthodes et à m'attarder sur différents aspects des textes comparés, celle-ci ne prend que peu en compte les aspects purement textuels de la traduction littéraire, parmi lesquels il conviendrait de s'attarder sur des éléments de la narration, de la continuité discursive ou de l'expérience de lecture (cf. [TAIVALKOSKI-SHILOV 2019b](#) ; [PONCHARAL 2021](#)). De la même manière, il serait intéressant de voir comment améliorer l'outil proposé ici, en particulier l'équilibre entre l'adaptation au style personnel et la robustesse générale du système de TA (cf. [Section 6.1.2](#)), pour lequel il pourrait être utile d'effectuer un premier entraînement sur des données littéraires exclusivement, et non sur des données génériques comme je le fais ici, suivi d'un affinage sur des traductions personnelles uniquement, de manière à s'assurer que celles-ci aient un poids plus important dans le modèle final. L'implémentation de l'outil mériterait également des recherches plus poussées, qui chercheraient à favoriser une utilisation plus dynamique et interactive de la TA et à maximiser le confort de l'environnement de travail. Dans ce but, j'espère prochainement pouvoir mettre à l'essai le scénario envisagé par le biais de l'outil *OmegaT* ([OmegaT Project 2002](#)), proposé parmi les pistes d'améliorations possibles (cf. [Section 7.4.1](#)), mais cette dernière limite de l'étude invite généralement à prendre en considération la manière dont interagissent la TA, la TAO et les autres outils de traduction, à la fois entre eux et entre les mains des traducteurs, de même que l'influence de ces ressources sur l'acte et l'activité de traduction.

Par la suite, il serait tout aussi intéressant de voir ce qu'augure l'arrivée d'outils génératifs tels que *ChatGPT* ([OpenAI 2022](#)), quelle place ces outils pourraient occuper dans les interfaces et dans le processus de traduction, mais aussi et surtout quels enjeux en découlent. Des recherches préliminaires se penchent déjà sur l'utilisation de ces grands modèles de langue (GML) pour la traduction automatique, tirant notamment parti de leur capacité à traduire des paragraphes dans leur intégralité ([KARPINSKA et IYER 2023](#)), quoique leur performance n'est pas toujours

comparable à celle des systèmes de TA mis au point spécifiquement dans ce but, en particulier lorsque l'anglais n'est pas la langue cible (JIAO *et al.* 2023). Une autre utilisation possible des GML serait celle d'un outil de réécriture, comme le suggère la méthode employée par THAI *et al.* (2022) et comme le propose précisément *DeepL Write* (DeepL GmbH 2023). Envisagé de la sorte, le programme interviendrait plutôt dans un processus proche de la post-édition, après un premier jet de traduction. Une fois encore, cependant, je pense que ces programmes pourraient être utilisés et pensés différemment, à savoir comme des outils d'aide à la rédaction qui, comme la TLA, devraient soutenir la démarche créative sans prétendre à remplacer l'humain et qui, comme la TLAO, pourraient aider à améliorer la qualité des traductions en facilitant l'analyse des textes sources et cible ou le repérage d'éventuelles erreurs¹⁶². Des discussions et expérimentations exploratoires avec d'autres collègues spécialistes de la traduction littéraire, et notamment avec le traducteur David MCKAY (2023), semblent en tous cas pointer dans cette direction. La TA, après tout, s'inscrit dans la logique des GML, dont elle suit le même fonctionnement et pour lesquels il devrait être possible de tirer les mêmes conclusions que celles dégagées tout au long de ce travail. Ceci inclut les enjeux posés par ces outils (déjà abordés en partie dans la Section 8.2) de même que leurs limites. Comme la TA, il apparaît déjà que les GML sont de bons imitateurs (l'adaptation au style est la base même de cette recherche) tout en étant enclins aux hallucinations (il s'agit là d'une nouveauté de la TA neuronale, dont j'ai souligné les problèmes d'adéquation derrière l'apparente fluidité). De la même manière que pour la TA, à nouveau, je pense toutefois que leur principale utilité se trouve dans le dialogue qui peut s'instaurer avec l'humain (ce qui est la principale fonction des GML) et que considérer ceux-ci comme des robots algorithmiques n'empêche aucunement d'en faire un outil de discussion, d'analyse, ou de reformulation au service de la créativité humaine. Face à la multiplication de ces outils, la possibilité de les combiner au sein d'une même interface se fera sûrement grandissante, tout comme la nécessité d'en considérer les aspects cognitifs et sociétaux.

162. Du moins pour l'anglais. Bien que le sujet évolue très rapidement, les recherches et résultats actuels montrent que les performances de ces outils, mis au point avec des données en langue anglaise pour la très large majorité — 90 % pour Llama 2 (TOUVRON *et al.* 2023, p. 21–22) —, produisent naturellement de moins bons résultats si la langue d'arrivée n'est pas l'anglais et moins bons encore si elle n'est pas la langue source.

007

**HOME DES ÉTUDIANTS
DE L'UNIVERSITÉ DE LIÈGE**
67, Boulevard d'Avroy
Vendredi 2 Octobre, à 20 h. 15.

**CONFÉRENCE DE
M. François LE LIONNAIS**
président de l'Association des Écrivains Scientifiques de France

**Machines LOGIQUES et ELECTRONIQUES
et LITTERATURE**

Pour répondre à une aimable suggestion, le conférencier parlera également des postes à gâche et des chemins de fer Cambodgiens.

**Entrée générale : 10 francs.
Gratuite pour les participants aux journées Christian Beck,
Verviers, 3 et 4 août 1964, téléphone (087) 233.85.**

Dépositaire de droits : Imprimerie SIBIRSKAÏE-CAPELLI, Dinant — Téléphone 2248

007

COMMISSION CONSULTATIVE COMMUNALE
DE LA CULTURE ET DES LOISIRS - VERVIERS

PROGRAMME des 4, 5 et 6 OCTOBRE
sous le patronage de « TEMPS MÊLÉS »

Vendredi 4 Octobre
Salle du Conservatoire, à 16 h. 30

Sous la présidence de M. François LE LIONNAIS et avec le concours de IBM-BELGIUM, vous jouerez, (entre autres) au

JEU 

des trois alertes petits pois
de Raymond QUENEAU

*programmé à Neuilly, grâce à un « Terminal Rax »
installé à Verviers.*

Let them express and encourage but not replace our imagination.
(D. Dolci).

SOIRÉE LIBRE - CONFÉRENCE SURPRISE
- DIVERTISSEMENTS LOCAUX -

FIGURE 35 – Prospectus de conférence de l'ALAMO à Liège (1964) et à Verviers (1969)

SOURCE : Gallica (1964 & 1969), © BnF, Fonds Oulipo

Corpus synthétique

Abdellatif Laâbi (2010) — *Le Fond de la jarre*

Adrien Tomas (2011) — *Les Six Royaumes - La Geste du Sixième Royaume*

Adrien Tomas (2013) — *Les Six Royaumes - La Maison des mages*

Agnès Abécassis (2017) — *Café ! Un garçon sil vous plaît*

Alain Damasio (2004) — *La Horde du Contrevent*

Albert Camus (1942) — *L'Étranger*

Amélie Nothomb (1995) — *Les Catilinaires*

Amin Maalouf (1986) — *Léon l'Africain*

Amin Maalouf (1988) — *Samarcande*

Ange (2001) — *Ayesha - Le Peuple turquoise*

Ange (2002) — *Ayesha - La Flamme d'Harabec*

Anne Cuneo (2007) — *Zaïda*

Anne Fakhouri (2008) — *Le Clairvoyage*

Anne Fakhouri (2009) — *La Brume des Jours*

Antoine de Saint-Exupéry (1939) — *Terre des hommes*

Antoine de Saint-Exupéry (1943) — *Le Petit Prince*

Ayerdhal (1992) — *Demain, une oasis*

Barbara Abel (2002) — *L'Instinct maternel*

Bernard Werber (1994) — *Les Thanatonautes*

Bernard Werber (2002) — *L'Arbre des possibles*

Bessora (2021) — *Les Orphelins*

Boris Vian (1947) — *L'Écume des jours*

Boualem Sansal (2007) — *Petit éloge de la mémoire*

Boualem Sansal (2015) — *2084 : La fin du monde*

Boubaker Ayadi (2006) — *Le rêve du sultan*

Catherine Dufour (2019) — *Danse avec les lutins*

Charif Majdalani (2017) — *L'Empereur à pied*

Charles-Ferdinand Ramuz (1934) — *Derborence*

Christelle Dabos (2013) — *La Passe-miroir - Les Fiancés de l'hiver*

Christelle Dabos (2015) — *La Passe-miroir - Les Disparus du Clairdelune*

Christian Guay-Poliquin (2016) — *Le Poids de la neige*

Claire Duvivier (2020) — *Un long voyage*

Claire Duvivier (2021) — *Citadins de demain*

Damien Snyers (2016) — *La Stratégie des As*

Dany Laferrière (2015) — *Le Cri des oiseaux fous*

David Khara (2010) — *Les Vestiges de l'Aube*

Driss Chraïbi (1972) — *La Civilisation, ma Mère !...*

Driss Chraïbi (1981) — *Une enquête au pays*

Eduardo Manet (1996) — *Rhapsodie cubaine*

Eléonore Devillepoix (2020) — *La Ville sans vent - Tome 1*

Eléonore Devillepoix (2020) — *La Ville sans vent - Tome 2*

Éric Giacometti & Jacques Ravenne (2005) — *Le Rituel de l'ombre*

Erik Wietzel (2010) — *Les Dragons de la Cité Rouge*

Fabrice Colin (2014) — *Arcadia*

Fawzi Mellah (1988) — *Elissa, la reine vagabonde*

Fouad Laroui (2017) — *L'Insoumise de la porte de Flandre*

Fouad Laroui (2018) — *Dieu, les mathématiques, la folie*

Francis Berthelot (1990) — *Rivage des intouchables*

François Mauriac (1913) — *L'Enfant chargé de chaînes*

François Mauriac (1927) — *Thérèse Desqueyroux*

Franz Hellens (1987) — *Mélusine*

Frédéric Petitjean (2011) — *Les Dolce - La Route des magiciens*

Gary Victor (2014) — *L'Escalier de mes désillusions*

Georges Eekhoud (1892) — *Le cycle patibulaire*

Georges Eekhoud (1899) — *Escal-vigor*

Georges-Olivier Châteaureynaud (2007) — *L'Autre Rive*

Georges Perec (1969) — *La Disparition*

Georges Perec (1978) — *La Vie mode d'emploi*

Georges Rodenbach (1892) — *Bruges-la-Morte*

Georges Simenon (1929) — *Pietr-le-Letton*

Georges Simenon (1930) — *Le Charretier de La Providence*

Georges Simenon (1932) — *L'Affaire Saint-Fiacre*

Georges Simenon (1949) — *Mon ami Maigret*

Georges Simenon (1965) — *La Patience de Maigret*

Guillaume Chamanadjian (2021) — *Capitale Sud - Le Sang de la Cité*

Henri Loevenbruck (2011) — *L'Apothicaire*

Henri Lopes (1971) — *Tribaliques*

Henri Michaux (1948) — *Ailleurs*

Herve Jubert (2012) — *Magies secrètes*

Herve Jubert (2013) — *Le Tournoi des ombres*

Honoré de Balzac (1835) — *Le Père Goriot*

Hubert Haddad (2013) — *Le Peintre d'éventail*

Hyam Yared (2020) — *Nos longues années en tant que filles*

Ian Manook (2013) — *Yeruldelgger*

Jacques Chessex (1973) — *L'Ogre*

Jean-Christophe Réhel (2018) — *Ce qu'on respire sur Tatouine*

Jean Cocteau (1928) — *Le Livre blanc*

Jean Cocteau (1929) — *Les Enfants terribles*

Jean d'Ormesson (1971) — *La Gloire de l'Empire*

Jean-Laurent Del Socorro (2015) — *Royaume de vent et de colères*

Jean-Luc Bizien (2002) — *Wonderlandz*

Jean-Luc Bizien (2017) — *L'Appel du Dragon*

Jean-Luc Marcastel (2017) — *L'Auberge entre les mondes - Péril en cuisine*

Jean-Luc Marcastel (2018) — *L'Auberge entre les mondes - Embrouilles au menu*

Jean Ray (1943) — *Malpertuis*

Jean Ray (1944) — *Les Derniers Contes de Canterbury*

J. M. G. Le Clézio (1980) — *Désert*

Judith Bouilloc (2019) — *L'Arrache-mots*

Jules Romains (1933–1946) — *Les Hommes de bonne volonté*

Justine Niogret (2009) — *Chien du Heaume*

Justine Niogret (2011) — *Mordre le bouclier*

Katia Lanero Zamora (2012) — *Chroniques des Hémisphères - Le Bal des Poussières*

Katia Lanero Zamora (2013) — *Chroniques des Hémisphères - La reine de la pluie*

Khaoula Hosni (2013) — *À ta place*

Koffi Kwahulé (2010) — *Monsieur Ki*

Liliana Lazar (2016) — *Enfants du diable*

Lionel Davoust (2017) — *Les Dieux sauvages - La Messagère du ciel*

Lionel Davoust (2018) — *Les Dieux sauvages - Le verrou du fleuve*

Louis Aragon (1922) — *Les Aventures de Télémaque*

Louis Aragon (1944) — *Aurélien*

Madeleine Bourdhoux (1937) — *La Femme de Gilles*

Manon Fargetton (2015) — *L'Héritage des Rois-Passeurs*

Manon Fargetton (2016) — *Les Illusions de Sav-Loar*

Marcel Proust (1927) — *A la recherche du temps perdu*

Marcel Thiry (1960) — *Nouvelles du Grand Possible*

Marguerite Yourcenar (1951) — *Mémoires d'Hadrien*

Marguerite Yourcenar (1968) — *L'Œuvre au noir*

Marie Pavlenko (2013) — *La Fille-Sortilège*

Maxime Chattam (2008) — *Autre-monde - Tome 1*

Maxime Chattam (2009) — *Autre-monde - Tome 2*

Michel Pagel (2006) — *Les Immortels - Les Mages de sumer*

Michel Pagel (2006) — *Les Immortels - Les Mages du nil*

Michel Tournier (1971) — *Vendredi ou La Vie sauvage*

Michel Tournier (1975) — *Les Météores*

Mireille Calmel (2002) — *Le Lit d'Aliénor*

Mohamed Mbougar Sarr (2021) — *La Plus Secrète Mémoire des hommes*

Mohamed Nedali (2011) — *La Maison de Cicine*

Mohamed Nedali (2012) — *Triste jeunesse*

Naomi Fontaine (2019) — *Shuni*

Nathalie Sarraute (1957) — *Tropismes*

Nathalie Sarraute (1963) — *Les Fruits d'or*

Nathalie Somers (2017) — *Roslend - Tome 1*

Nathalie Somers (2018) — *Roslend - Tome 2*

Olivier Descosse (2003) — *Le Couloir de la pieuvre*

Patrick Bauwen (2017) — *Le Jour du chien*

Paul Beorn (2012) — *Les Derniers Parfaits*

Paul Beorn (2015) — *Le Septième Guerrier-Mage*

Paul Colize (2013) — *Un long moment de silence*

Pierre Bordage (2019) — *Arkane*

Pierre Grimbert (1999) — *Le Secret de Ji - Tome 1*

Pierre Grimbert (1999) — *Le Secret de Ji - Tome 2*

Pierre Pevel (2021) — *Le Paris des Merveilles*

Raymond Queneau (1965) — *Les Fleurs bleues*

René Barjavel (1968) — *La Nuit des temps*

René Barjavel (1984) — *L'Enchanteur*

Robert Merle (1972) — *Malevil*

Roger Leloup (1995) — *Le Pic des ténèbres*

Romain Lucazeau (2016) — *Latium*

Romain Lucazeau (2021) — *La Nuit du faune*

Sami Mokaddem (2014) — *Dix-neuf*

Simone de Beauvoir (1946) — *Tous les hommes sont mortels*

Sony Labou Tansi (1979) — *La Vie et demie*

Stefan Platteau (2015) — *Dévoreur*

Stefan Platteau (2019) — *Le Roi cornu*

Stéphane Beauverger (2009) — *Le Déchronologue*

Timothée de Fombelle (2006) — *Tobie Lolness - La vie suspendue*

Timothée de Fombelle (2007) — *Tobie Lolness - Les yeux d'Elisha*

Victor Dixen (2013) — *Animale - La Malédiction de Boucle d'Or*

Victor Dixen (2015) — *La Prophétie de la Reine des neiges*

Yasmine Ghata (2016) — *J'ai longtemps eu peur de la nuit*

Corpus parallèle

Anthony Piers (1977) — *A Spell for Chameleon*
Dominique Haas (1991) — *Lunes pour Caméléon*

Anthony Piers (1979) — *The Source of Magic*
Dominique Haas (1991) — *La Source de magie*

Anthony Piers (1979) — *Castle Roogna*
Dominique Haas (1992) — *Château-Roogna*

C. S. Lewis (1950) — *The Lion, the Witch and the Wardrobe*
Anne-Marie Dalmais (2001) — *Le Lion, la Sorcière blanche et l'Armoire magique*

C. S. Lewis (1952) — *The Voyage of the Dawn Treader*
Philippe Morgaut (2002) — *L'Odyssée du Passeur d'Aurore*

C. S. Lewis (1955) — *The Magicians Nephew*
Cécile Dutheil de la Rochère (2001) — *Le Neveu du magicien*

George R. R. Martin (1999) — *A Clash of Kings*
Jean Sola (2000) — *Le Trône de fer - 2*

George R. R. Martin (2011) — *A Dance with Dragons*
Patrick Marcel (2014) — *Le Trône de fer - 5*

J. K. Rowling (1997) — *Harry Potter and the Philosopher's Stone*
Jean-François Ménard (1998) — *Harry Potter à l'école des sorciers*

J. K. Rowling (1998) — *Harry Potter and the Chamber of Secrets*
Jean-François Ménard (1999) — *Harry Potter et la Chambre des secrets*

J. K. Rowling (1999) — *Harry Potter and the Prisoner of Azkaban*
Jean-François Ménard (1999) — *Harry Potter et le Prisonnier d'Azkaban*

Joseph Delaney (2004) — *The Spook's Apprentice*
Marie-Hélène Delval (2005) — *L'Apprenti épouvanteur*

Joseph Delaney (2005) — *The Spook's Curse*
Marie-Hélène Delval (2006) — *La Malédiction de l'épouvanteur*

Joseph Delaney (2006) — *The Spook's Secret*
Marie-Hélène Delval (2007) — *Le Secret de l'épouvanteur*

J. R. R. Tolkien (1937) — *The Hobbit*
Daniel Lauzon (2012) — *Le Hobbit*

J. R. R. Tolkien (1954) — *The Fellowship of the Ring*
Daniel Lauzon (2014) — *La Fraternité de l'Anneau*

Lisa Tuttle & George R. R. Martin (1981) — *Windhaven*
Patrick Marcel (1999) — *Elle qui chevauche les tempêtes*

Margaret Rogerson (2017) — *An Enchantment of Ravens*
Vincent Basset (2021) — *Enchantment of Ravens*

Margaret Rogerson (2019) — *Sorcery of Thorns*
Vincent Basset (2020) — *Sorcery of Thorns*

Naomi Novik (2015) — *Uprooted*
Benjamin Kuntzer (2017) — *Déracinée*

Naomi Novik (2018) — *Spinning Silver*
Thibaud Eliroff (2020) — *La Fileuse d'argent*

Neil Gaiman (1996) — *Neverwhere*
Patrick Marcel (1998) — *Neverwhere*

Neil Gaiman (1999) — *Stardust*
Frédérique Le Boucher (2001) — *Stardust*

Neil Gaiman & Michael Reaves (2007) — *InterWorld*
Michel Pagel (2010) — *Entremonde*

Orson Scott Card (1987) — *Seventh Son*
Patrick Couton (1991) — *Le Septième Fils*

Orson Scott Card (1988) — *Red Prophet*
Patrick Couton (1992) — *Le Prophète rouge*

Orson Scott Card (1989) — *Prentice Alvin*
Patrick Couton (1993) — *L'Apprenti*

Philip Pullman (1995) — *The Golden Compass*
Jean Esch (1998) — *Les Royaumes du Nord*

Philip Pullman (1997) — *The Subtle Knife*
Jean Esch (1998) — *La Tour des anges*

Philip Pullman (2000) — *The Amber Spyglass*
Jean Esch (2001) — *Le Miroir d'ambre*

Terry Pratchett (1987) — *Equal Rites*
Patrick Couton (1994) — *La Huitième Fille*

Terry Pratchett (1988) — *Sourcery*
Patrick Couton (1995) — *Sourcellerie*

Terry Pratchett (1991) — *Reaper Man*
Patrick Couton (1998) — *Le Faucheur*

Ursula K. Le Guin (1968) — *A Wizard of Earthsea*
Philippe Hupp & Patrick Dusoulier (2018) — *Le Sorcier de Terremer*

Ursula K. Le Guin (1990) — *Tehanu*
Isabelle Delord-Philippe (2018) — *Tehanu*

Ursula K. Le Guin (2001) — *Tales from Earthsea*
Pierre-Paul Durastanti (2018) — *Contes de Terremer*

Corpus personnalisé

Charles Edward Pogue (1996) — *Dragonheart / Coeur de dragon*

Gene Wolfe (1993) — *Nightside the Long Sun / Côté nuit*

Gene Wolfe (1994) — *Lake of the Long Sun / Côté lac*

Gene Wolfe (1994) — *Caldé of the Long Sun / Caldé, côté cité*

Gene Wolfe (1996) — *Exodus From the Long Sun / L'Exode*

Iain M. Banks (1998) — *Inversions / Inversions*

Inbali Iserles (2015) — *Foxcraft / Les Possédés*

Inbali Iserles (2016) — *The Elders / Les Anciens*

Inbali Iserles (2017) — *The Mage / Le Mage*

Jack Finney (1968) — *The Woodrow Wilson Dime / La Pièce d'à côté*

Jonathan Carroll (1998) — *Kissing the Beehive / Le Baiser aux abeilles*

Jonathan Carroll (2001) — *The Wooden Sea / L'Aube du huitième jour*

Karina Yan Glaser (2017) — *The Vanderbeekers of 141st Street / On reste ici*

Karina Yan Glaser (2018) — *The Vanderbeekers and the Hidden Garden / Notre jardin secret*

Laura Ruby (2008) — *Lily's Ghosts / Fantomes à tous les étages*

Lauren Beukes (2013) — *The Shining Girls / Les Lumineuses*

Lisa Goldstein (1997) — *Tourists / Amaz*

Marta Randall (1983) — *The Sword of Winter / L'Épée de l'hiver*

Martha Grimes (2010) — *The Black Cat / Ce que savait le chat*

Martha Grimes (2015) — *The Way of All Fish / Un gros poisson*

Martha Grimes (2016) — *Vertigo 42 / Vertigo 42*

Martha Grimes (2018) — *The Knowledge / Faites vos jeux*

Mike Resnick (1994) — *A Miracle of Rare Design / Projet Miracle*

Pat Cadigan (1999) — *Tea From an Empty Cup / Vous avez dit virtuel*

Paul J. McAuley (1993) — *Red Dust / Sable rouge*

Rachel Hawkins (2014) — *Rebel Belle / Rebelle Belle*

Richard Matheson (1954) — *I Am Legend / Je suis une légende*

Rick Riordan (2010) — *The Red Pyramid / La Pyramide rouge*

Rick Riordan (2011) — *The Throne of Fire / Le Trône de feu*

Rick Riordan (2012) — *The Serpent's Shadow / L'Ombre du serpent*

Rick Riordan (2013) — *The Son of Sobek / Le Fils de Sobek*

Rick Riordan (2014) — *Percy Jackson and the Greek Gods / Percy Jackson et les dieux grecs*

Rick Riordan (2015) — *The Sword of Summer / L'épée de l'été*

Rick Riordan (2015) — *Percy Jackson and the Greek Heroes / Percy Jackson et les héros Grecs*

Rick Riordan (2016) — *The Hammer of Thor / Le Marteau de Thor*

Rick Riordan (2017) — *The Ship of the Dead / Le Vaisseau des damnés*

Robert James Sawyer (1995) — *The Terminal Experiment / Expérience terminale*

Robert James Sawyer (1997) — *Illegal Alien / Un procès pour les étoiles*

Rudy Rucker (1982) — *Software / Software*

Suzette Haden Elgin (1979) — *Star Anchored, Star Angered / Drussa Silver*

Architecture des systèmes utilisés

Les tableaux suivants présentent les paramètres des architectures utilisées dans le [Chapitre 5](#). La plupart des systèmes décrits dans le [Tableau 5](#), entraînés sur environ 5 millions de segments, utilisent les paramètres de la version standard du *Transformer (base)*. Le système *Magyk*, entraîné uniquement sur les 45 000 segments des six romans, nécessitait une optimisation des paramètres des architectures *Transformer* et LSTM inspirée par SENNRICH et ZHANG (2019) ainsi que ARAABI et MONZ (2020). Enfin, la version plus lourde du *Transformer (big)*, décrite dans le [Tableau 6](#), reprend les paramètres affichés dans VASWANI *et al.* (2017).

Systèmes *Transformer*

Le tableau indique le nom du paramètre utilisé dans la configuration d'*OpenNMT* et la modification des différents paramètres en fonction de la taille du corpus d'entraînement, à savoir le nombre de couches cachées (*hidden layers*), la dimension des couches cachées (*hidden layer dimension*), la dimension des couches de propagation (*feed-forward layer dimension*), le nombre de têtes d'attention (*attention heads*), la probabilité d'extinction des neurones (*dropout*) et le lissage d'étiquettes (*label smoothing*). La taille du système résultant (aussi exprimée en « paramètres ») et le nombre d'étapes nécessaires à l'entraînement y sont aussi indiquées.

Paramètres		Taille du corpus		
Architecture	<i>OpenNMT</i>	45 K	5 M	265 M
N	layers	5	6	6
d_{model}	hidden_size	512	512	1024
d_{ff}	transformer_ff	1024	2048	4096
h	heads	2	8	16
P_{drop}	dropout	0,3	0,1	0,3
ϵ_{ls}	label_smoothing	0,6	0,1	0,1
Paramètres (en millions)		48	66	221
Étapes (en milliers)		25	205	250

TABLEAU 15 – Paramètres des architectures *Transformer*

Systemes LSTM

De la même manière que pour le *Transformer*, le tableau ci-dessous reprend les modifications effectuées pour le système LSTM standard et pour sa version optimisée sur le corpus réduit. Le paramètre de la dimension des plongements lexicaux (*word vector size*) y a été ajouté.

Paramètres		Taille du corpus	
Architecture	<i>OpenNMT</i>	45 K	5 M
N	layers	2	3
d_{model}	hidden_size	256	1024
d_{vector}	word_vec_size	256	512
ϵ_{ls}	label_smoothing	0,2	0,1
Paramètres (en millions)		13	76
Étapes (en milliers)		120	220

TABLEAU 16 – Paramètres des architectures LSTM

Évaluations de la TLA



Évaluations humaines

Types	Mesure	Source
Annotation d'erreurs	Adéquation / fluidité	KUZMAN <i>et al.</i> (2019), GUERBEROF-ARENAS et TORAL (2020), LONG (2022)
	Typologie d'erreurs	Cette étude, MATUSOV (2019), FONTEYNE <i>et al.</i> (2020), WEBSTER <i>et al.</i> (2020), MACKEN <i>et al.</i> (2022), GUERBEROF-ARENAS et TORAL (2022), PETRAK <i>et al.</i> (2022), LONG (2022), KARPINSKA et IYER (2023)
Post-édition	Qualité / corrections	Cette étude, MOORKENS <i>et al.</i> (2018), Ó MURCHÚ (2019), GUERBEROF-ARENAS et TORAL (2020, 2022), OLIVER <i>et al.</i> (2022), VIEIRA <i>et al.</i> (2023), KOLB (2024)
	Ergonomie	Cette étude, MOORKENS <i>et al.</i> (2018), Ó MURCHÚ (2019), VIEIRA <i>et al.</i> (2023), KOLB (2024), ROTHWELL (2024)
	Voix	Cette étude, KENNY et WINTERS (2020), WINTERS et KENNY (2024)
	Qualité (classement, commentaires)	Cette étude, BESACIER (2014), TORAL <i>et al.</i> (2020), THAI <i>et al.</i> (2022)
Évaluation par le lectorat	Qualité (classement)	BESACIER (2014), TORAL et WAY (2015b, 2018), SHIH (2016), TORAL <i>et al.</i> (2024), UTUSHKINA (2023)
	Expérience de lecture	BESACIER (2014), GUERBEROF-ARENAS et TORAL (2020), OLIVER <i>et al.</i> (2022), OLIVER (2024)
Commentaires critiques	Créativité	ŞAHİN et GÜRSES (2019), GUERBEROF-ARENAS et TORAL (2020, 2022)
	Analyse d'erreurs	Cette étude, JONES et IRVINE (2013), TORAL et WAY (2015a,b), KUZMAN <i>et al.</i> (2019), CONSTANTINE (2019), TORAL <i>et al.</i> (2020, 2024), MOHAR <i>et al.</i> (2020), PONCHARAL (2021), BASSOLE (2022), DEFERT (2022), LONG (2022), BRUSASCO (2022)
	Phénomènes particuliers	JIANG et NIU (2022), ZAJDEL (2022), CASTILHO et RESENDE (2022), DORST (2024), KOLB <i>et al.</i> (2023)

TABLEAU 17 – Évaluations humaines de la TLA

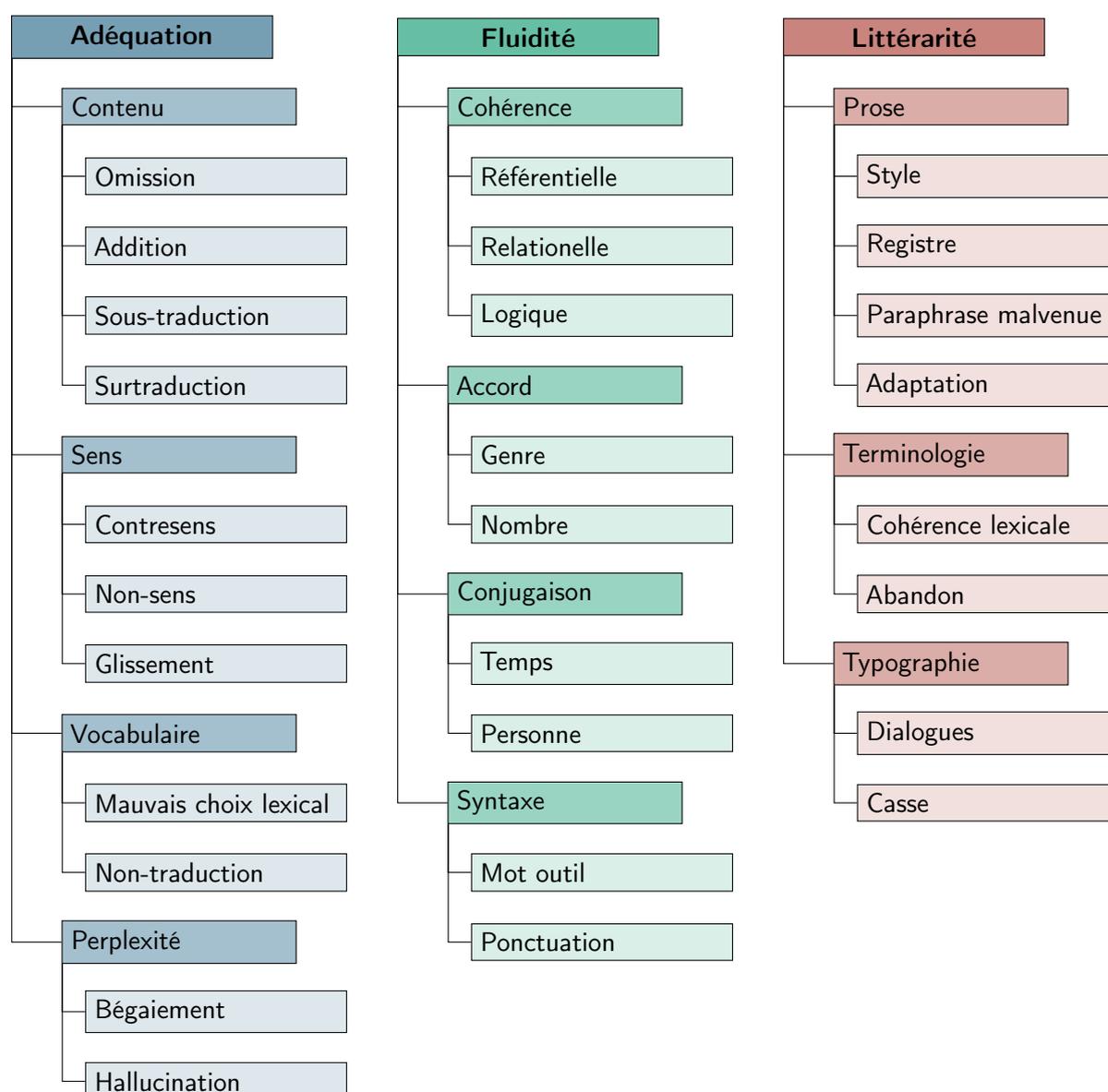
Évaluations automatiques

Types	Mesures	Source
	BLEU	Toutes les sources impliquant la mise au point d'un système et d'autres
	ChrF, YiSi	Cette étude
	TER	Cette étude, TORAL et WAY (2015a,b), MATUSOV (2019), MACKEN et al. (2022), OLIVER (2024)
	HTER	GUERBEROF-ARENAS et TORAL (2020, 2022), OLIVER et al. (2022), CASTILHO et RESENDE (2022)
Métriques	METEOR	TORAL et WAY (2015b), KUZMAN et al. (2019)
	CHARCUT	MACKEN et al. (2022), KOLB (2024)
	COMET	Cette étude, MACKEN et al. (2022), KARPINSKA et IYER (2023), TORAL et al. (2024)
	BERTScore	MACKEN et al. (2022), KARPINSKA et IYER (2023)
	BLEURT	THAI et al. (2022), KARPINSKA et IYER (2023)
	BlonDe	THAI et al. (2022)
	NIST, WER, EdDist, HBLEU, HINIST, etc.	OLIVER (2024)
Stylométrie	Delta de Burrows	Cette étude, WEBSTER et al. (2020)
	Analyse par mots-clefs	WINTERS et KENNY (2024)
Richesse lexicale	Indices type-token	Cette étude, TORAL et WAY (2018), TEZCAN et al. (2019), MACKEN et al. (2022), CASTILHO et RESENDE (2022), TORAL et al. (2024), WINTERS et KENNY (2024)
	Densité lexicale	CASTILHO et RESENDE (2022)
Entropie lexicale	AWTE	TEZCAN et al. (2019), MACKEN et al. (2022)
Recouvrement lexical	Recouvrement de n-grammes	TORAL et WAY (2018), TORAL et al. (2024)
	Recouvrement de lemmes	TEZCAN et al. (2019), WEBSTER et al. (2020)
Recouvrement sémantique	WordNet	TEZCAN et al. (2019), WEBSTER et al. (2020)
Mesure de cohésion	Résolution de co-référence	VOIGT et JURAFSKY (2012)
	Longueur de phrases	Cette étude, TORAL et WAY (2018), CASTILHO et RESENDE (2022)
Équivalence syntaxique	Cross value	Cette étude, TEZCAN et al. (2019), WEBSTER et al. (2020)
	Étiquetage morpho-syntaxique	Cette étude
	ASTrED	Cette étude, WEBSTER et al. (2020), MACKEN et al. (2022)
	Cognitif (pauses)	TORAL et al. (2018b), GUERBEROF-ARENAS et TORAL (2022)
Effort de post-édition	Temporel (durée)	BESACIER (2014), TORAL et al. (2018b), KUZMAN et al. (2019), Ó MURCHÚ (2019), GUERBEROF-ARENAS et TORAL (2022), VIEIRA et al. (2023), KOLB (2024), OLIVER (2024)
	Technique (frappes)	TORAL et al. (2018b), GUERBEROF-ARENAS et TORAL (2022), VIEIRA et al. (2023), KOLB (2024), OLIVER (2024)

TABLEAU 18 – Évaluations automatiques de la TLA

Typologie d'erreurs

Cette annexe détaille la typologie d'erreurs élaborée pour cette étude à partir des taxonomies utilisées par VILAR *et al.* (2006, p. 699), MATUSOV (2019, p. 14–15), TEZCAN *et al.* (2019, p. 42) et SCHUMACHER (2023, p. 104–105). Les sections à la suite détaillent pour chacune des trois classes les différentes catégories retenues, accompagnée d'une courte description, de sa contribution au nombre total d'erreurs et d'un exemple tiré des trois chapitres évalués.



Adéquation

Omission (OMI) — 21,6 % des erreurs

Mot ou nuance que l'on retrouve dans la source et non dans la cible (y compris dans le cas où l'omission est désirable ou se retrouve dans la référence).

Ex. : *The buzz of Magyk filled the Hermetic Chamber once again, and from Marcia's mouth came a long stream of pink mist. It settled over Beetle's face, covering his nose and mouth.*

Le bourdonnement de la Magyk emplît à nouveau le cabinet hermétique. Un long filet de brume rose jaillit de la bouche de Marcia, recouvrant son nez et sa bouche.

Addition (ADD) — 0,5 % des erreurs

Mot ou nuance que l'on retrouve dans la cible et non dans la source.

Ex. : *They filed into the dimly lit Manuscriptorium, glancing at each other, wondering who among them would be the next Chief Hermetic Scribe.*

Ils s'enfoncèrent dans la pénombre du Manuscriptorium, jetant des regards inquiets autour d'eux, se demandant qui allait devenir le prochain premier scribe hermétique.

Sous-traduction (SOU) — 3 % des erreurs

Traduction correcte mais conduisant à une perte de nuance dans la langue cible ou à une difficulté de compréhension.

Ex. : *That night the Palace Ballroom was ablaze with light [...]*

Cette nuit-là, la salle de bal était entièrement éclairée [...]

Surtraduction (SUR) — 0,6 % des erreurs

Traduction correcte mais conduisant à une redondance dans la langue cible ou à l'ajout d'une nuance qui n'est pas présente dans la langue source.

Ex. : *Shadowy figures of Wizards and Apprentices were scattered about the Great Hall, the younger ones wandering anxiously to and fro [...]*

De sombres silhouettes de magiciens et d'apprentis s'éparpillaient dans le hall, les plus jeunes erraient nerveusement d'avant en arrière [...]

Contresens (CTS) — 0,9 % des erreurs

Traduction conduisant à un sens contraire à celui exprimé dans la source.

Ex. : *I don't want to leave the Manuscriptorium unattended.*

Je ne veux pas laisser le Manuscriptorium sous surveillance.

Non-sens (NOS) — 2 % des erreurs

Traduction ne faisant aucun sens.

Ex. : *“UnDoing. Great. The.*
— Super.

Glissement (GLS) — 3,4 % des erreurs

Traduction modifiant le sens d’origine du segment.

Ex. : *The Big Freeze came in fast.*
Le Grand Gel arriva à point nommé.

Mauvais choix lexical (MCL) — 16 % des erreurs

Calques, termes et syntagmes mal traduits, problèmes de désambiguïisation lexicale.

Ex. : *Focusing only on the hem of Marcia’s purple cloak, he followed her up the stepped side of the golden Pyramid.*
Se concentrant uniquement sur le bas de la cape pourpre de Marcia, il la suivit le long du bastingage de la pyramide dorée.

Non-traduction (NOT) — 0 % des erreurs

Terme du texte source reproduit tel quel dans la cible, là où il est nécessaire de proposer une traduction.

Ex. : *Shakily he headed across to the wide marble steps that led up to the silver doors that guarded the Tower.*
Shakily se dirigea vers l’escalier de marbre qui menait aux portes d’argent de la tour.
(Traduction issue d’un ancien modèle.)

Bégaïement (BEG) — 0,2 % des erreurs

Syntagmes répétés sans raison apparente par la système de TA.

Ex. : *The winter morning drew on. Snow piled up outside the long windows, while the Ballroom was transformed with holly and ivy, red ribbons, huge silver candlesticks and a whole box of [...]*
Le matin d’hiver s’écoula, tandis que la neige s’entassait devant les fenêtres tandis que la salle de bal se transformait en rubans rouges, en chandeliers d’argent, en chandeliers d’argent [...]

Hallucination (HAL) — 1 % des erreurs

Fragments de texte tout à fait illogiques sans aucun rapport avec la source qui sont ajoutés dans la cible ou remplacent une partie du texte. Mots inventés résultant de la segmentation en sous-mots.

Ex. : *When Simon arrived at the lintel that marked the presence of the Hidden gate [...]*
Quand Simon atteignit le lintelligent qui signalait la présence de la porte secrète [...]

Fluidité

Cohésion référentielle (REF) — 2,5 % des erreurs

Rupture du lien entre des éléments coréférents : résolution de pronoms, problème d'anaphore lexicale, disparition de l'antécédent...

- Ex. : (1) *“My door... it's let someone in,” she whispered.*
— Ma porte... elle a laissé entrer quelqu'un, murmura-t-elle.
- (2) *“There's Darke on them.*
Ils sont imprégnés de Ténèbre.

Cohésion relationnelle (REL) — 0,2 % des erreurs

Rupture de l'articulation logique des phrases, enchaînement bancal.

- Ex. : *He waited, poised to write down the sequence of symbols that would lead them to The Great UnDoing. Why didn't Marcia begin reading out the symbols?*
Il attendit, prêt à écrire la séquence de symboles qui les conduirait au Grand Dénouement, et pourquoi Marcia n'avait-elle pas commencé à déchiffrer les symboles ?

Logique (LOG) — 4,8 % des erreurs

Problème de syntaxe, calques grammaticaux, rupture de l'enchaînement logique à l'intérieur d'une même phrase ou de la continuité logique du récit.

- Ex. : *A burst of raucous laughter came from Silas's end of the table, where Nicko was sitting with Rupert, Maggie and Foxy. Nicko was making seagull noises.*
Un éclat de rire monta de l'extrémité de la table où Nicko et Rupert, Maggie et Vulpin faisaient des bruits de mouette.
- Ex. : [...] *until they began to see the shadowy impression of the Hermetic Chamber beyond*
[...] jusqu'à ce qu'ils distinguent l'image du cabinet hermétique derrière eux.
(Le cabinet se trouve devant les personnages.)

Genre (GEN) — 1,5 % des erreurs

Problème lié au genre grammatical ou au genre d'un personnage (excepté la résolution de pronoms).

- Ex. : (1) *Remember when you got fired and you wouldn't take your pen?*
Tu te rappelles quand tu as été viré et que tu n'as pas voulu prendre ton porte-plume ?
- (2) *Well, I kept it.*
Eh bien, je l'ai gardée.

Nombre (NBR) — 0,5 % des erreurs

Problème lié aux accords singulier/pluriel.

Ex. : *As fast as she could Marcia began to count along the words on page fourteen until she reached the ninety-eighth word.*

Le plus rapidement possible, Marcia commença à compter sur la page quatorze jusqu'à ce qu'elle atteigne les quatre-vingt-dix-huitièmes mots.

Temps (TPS) — 4,1 % des erreurs

Problème dans le choix ou la concordance des temps.

Ex. : *He would have added a Camouflage.*

Il eût ajouté un camouflage.

Personne (PER) — 0,6 % des erreurs

Problème lié aux accords sujet/verbe.

Ex. : *"Just got fired, in fact.*

— On vient de se faire virer.

(Dialogue à la première personne.)

Mots outils (MOT) — 4,3 % des erreurs

Mauvaise traduction d'un déterminant, d'une préposition... (tout élément qui n'est pas un mot plein).

Ex. : *I shall place this ring in the Hidden Shelf until then."*

D'ici là, je placerai cette bague dans la étagère secrète.

Ponctuation (PCT) — 1,6 % des erreurs

Erreur de ponctuation, à l'exception des conventions typographiques littéraires (p. ex. liées aux dialogues).

Ex. : *At the top of the golden Pyramid, Marcia Overstrand, ExtraOrdinary Wizard, and her Apprentice, Septimus Heap, spoke the last word of The Great UnDoing.*

Au sommet de la pyramide dorée, Marcia Overstrand, la magicienne extraordinaire et son apprenti, Septimus Heap, prononçaient le dernier mot du Grand Dénouement.

Littérarité

Style (STY) — 17,4 % des erreurs

Traductions littérales, répétitions, lourdeurs, tournures et cooccurrences non idiomatiques, formulation rendant la compréhension du texte plus difficile ou nécessitant une explicitation en français, etc.

Ex. : *The seven-cornered passage looked ominously well used—there was a slimy Thing trail on the floor. Like giant slugs, thought Septimus.*

Le passage qui tournait sept fois sur lui-même avait l'air d'être utilisé de manière sinistre : une traînée de morve sur le sol évoquait une limace géante.

Ex. : *At either end of the table, at Jenna's insistence, sat Sarah and Silas.*

Assis aux deux bouts de la table, sur l'insistance de Jenna, Sarah et Silas étaient assis.

Registre (REG) — 3 % des erreurs

Problème de registre lié à une variation diastratique (ton habituellement associé à un personnage), diaphasique (degré de formalité dépendant des relations entre les personnages) ou encore diachronique (parler adapté de l'ancien français).

Ex. : *Let's get out of this place.*

Rentrons vite chez nous.

(Référence : Tirons-nous d'ici.)

Paraphrase malvenue (PAR) — 0,6 % des erreurs

Terme rendu par un syntagme équivalent qui mène à une lourdeur ou à une erreur dans la phrase.

Ex. : *Marcia had found The Undoing of the Darkenese in Aunt Zelda's cottage a few years back and she knew that somewhere within its pages lay The Great UnDoing—the legendary Anti-Darke incantation that practitioners of the Darke feared above all else.*

Quelques années plus tôt, Marcia avait trouvé L'Art de vaincre la Ténèbre dans le cottage de tante Zelda et elle savait qu'à l'intérieur de ses pages se trouvait l'incantation anti-Ténèbre, une incantation légendaire que les praticiens de la Ténèbre craignaient par-dessus tout.

Adaptation (ADP) — 1,2 % des erreurs

Passage exigeant une solution de traduction particulière en raison de la langue, de contraintes formelles, du contexte culturel, etc. : onomatopées, jeux de mots, métaphores, occasionalismes, références culturelles...

Ex. : *The Great UnDoing had been her undoing also.*

Le Grand Dénouement l'avait également anéantie.

(Référence : Le Grand Dénouement avait également dénoué les fils fragiles de son existence.)

Cohérence lexicale (COH) — 2,3 % des erreurs

Traduction d'un terme (commun ou fictionnel) qui n'est pas en accord avec les tomes précédents ou avec un choix préalable dans le même volume.

- Ex. : (1) *“Scribe Romilly, read the name scribed upon the pen.”*
— Scribe Romilly, lisez le nom écrit sur le porte-plume.
- (2) *Romilly peered at the pen.*
Romilly considéra le crayon.

Abandon (ABD) — 0,8 % des erreurs

Traduction neutre ou littérale d'un nom de personnage ou d'un terme spécifique à l'univers fictionnel de la série (irrealia) qui nécessitent une solution particulière ou originale.

- Ex. : *On his way Foxy met Beetle coming out of Larry's Dead Languages.*
En chemin, Vulpin rencontra Moustique qui sortait des langues mortes de Larry.
(Référence : Sur le chemin du Manuscriptorium, il rencontra Moustique, qui sortait de l'Officine de traduction des langues mortes.)

Dialogues (DLG) — 3,5 % des erreurs

Conventions typographiques propres à la littérature et à la série (essentiellement liées aux dialogues).

- Ex. : (1) *At the top of the Tower Septimus read out the first number of the first pair.*
Au sommet de la tour, Septimus lut le premier nombre de la première paire.
(Référence : Septimus lut le premier nombre de la liste :)
- (2) *“Fourteen.”*
Quatorze.
(Référence : — Quatorze.)

Casse (CAS) — 1,6 % des erreurs

Traduction qui contrevient aux choix de la traductrice concernant l'usage des majuscules pour les termes spécifiques à la saga (systématiquement exprimés avec une ou plusieurs majuscules en anglais).

- Ex. : *“Romilly Badger,” said Marcia, sounding very official. “As youngest scribe I ask you to enter the Hermetic Chamber and bring out the Pot.”*
— Romilly Badger, reprit Marcia d'un ton très officiel, je vous demande d'entrer dans le cabinet hermétique et d'en sortir le Pot.

Retranscription

Méthode

Cette annexe contient la retranscription de l'entretien que j'ai mené avec Nathalie SERVAL (2023a) à la suite de la tâche proposée et détaillée dans le [Chapitre 7](#). Le déroulé de cette discussion est fourni ici à titre complémentaire, avec l'accord de Nathalie et après vérification de sa part. Pour en faciliter la lecture, les hésitations, reprises et acquiescements de ces longues discussions informelles ont été supprimés, bien que j'aie cherché à concilier au mieux cette lisibilité avec la fidélité des entretiens.

Entretien

D.H. : Je pense que l'on peut démarrer notre entretien, puisque la discussion est déjà lancée, d'ailleurs un tout grand merci d'avoir accepté cet enregistrement, mais je t'ai coupée donc je te laisse en revenir à ton impression générale de cette petite expérience...

N.S. : Oui, je dis que par définition l'outil est intéressant. Après, c'est vrai que pour ce qui est de l'interface, bien évidemment, j'en n'étais pas familière, mais je trouve qu'elle est un peu lourde, mais peut être aussi une question d'habitude d'utilisation.

Qu'est-ce que tu trouvais particulièrement lourd dans l'interface ? Est-ce que c'est la segmentation, la manière dont les informations sont affichées ?

Oui, c'est plutôt au niveau de l'affichage, mais parce que je ne sais pas forcément à quoi correspondent tous les affichages proposés. Mais après, bon, je pense que si je l'avais à demeure, je le ferais comme je fais avec tous les logiciels que j'utilise, je n'en utiliserais qu'une partie, forcément. Je repérerais mon propre itinéraire et ce qui m'intéresse dedans. Après, c'est vrai qu'au niveau de la segmentation, on n'a pas l'habitude de voir un texte aussi segmenté. Donc pour moi, c'est au moins un outil intéressant au niveau du premier jet. Parce que ça ne permet pas une vue globale, une vue d'ensemble, même au niveau de la page.

Est-ce que ça t'intéresserait, si on réitérait l'expérience d'avoir une segmentation plutôt par paragraphe ?

Oui, je pense que déjà la segmentation par paragraphe serait plus pertinente et plus intéressante.

Ça, c'est un truc qu'il est possible de faire. Il faudra juste voir ce que ça donne au niveau de la traduction automatique, mais c'est tout à fait possible.

Surtout que je travaille en fait beaucoup plus au niveau du paragraphe qu'au niveau de la phrase en fait, quand je traduis. Donc oui, il y a des remaniements internes : mettre un bout de phrase ici ou fusionner deux phrases et éventuellement intervertir des éléments de phrase au niveau du paragraphe. Donc ce serait un niveau plus pertinent.

Oui, qui colle beaucoup plus au genre de travail qui se fait avec un texte littéraire, qui demande quand même une certaine prise de liberté par rapport à l'original.

Oui oui, complètement oui, complètement.

Tout d'abord, j'aurais voulu te demander comment tu travailles de manière classique quand tu reçois une traduction, sous quel format elle se présente...

Le plus souvent, ce sont des PDF et on travaille au niveau des épreuves. Ça fait des années que je n'ai pas travaillé à partir d'un livre déjà imprimé, parce qu'on travaille quand même en amont pour rapprocher les dates de parution des originaux et de la traduction. Sans doute parce qu'il y a de plus en plus de lecteurs susceptibles de lire en anglais, donc pour ne pas trop décaler les dates de parution. Mais ça fait très, très longtemps que je n'ai pas travaillé sur un livre. Ça fait des années...

Donc c'est toujours un fichier et on s'y habitue, surtout que maintenant je travaille avec deux écrans, donc c'est quand même assez confortable d'avoir l'original d'un côté, la traduction de l'autre. Ça, c'est plutôt bien. En plus, c'est plus facile de faire des recherches quand tu fais une recherche dans le texte original.

Typiquement, quand tu traduis, comment est ce que ça se présente ? Qu'est-ce que tu as sur ton bureau ? Qu'est-ce que tu as sur tes écrans ?

Sur mon bureau, j'ai mon *Robert*, toujours. Ce dictionnaire des combinaisons aussi, qui m'a été conseillé dans ma formation en correction et qui est extrêmement utile. Après, bon, ce qu'on peut lui reprocher, bien sûr, c'est que c'est un peu des clichés. Mais c'est quand même très très utile, surtout qu'il offre quand même pas mal de propositions à chaque fois. Puis après, bon, si on veut vraiment être créatif, en fonction du texte original, on peut s'en écarter. Mais c'est un bouquin qui est vraiment très très utile pour les traducteurs.

Et tu m'as dit que tu utilises *Antidote*, mais tu n'utilises jamais le module de cooccurrences d'*Antidote* ?

Ah ben non, pas du tout. Je n'utilise jamais.

Par habitude de ton *Robert* ?

Ouais, ouais ouais, parce que c'est vrai que j'aime bien le côté papier. Et j'avoue que je n'ai pas vraiment exploré toutes les fonctions d'*Antidote*. Je me suis vraiment limitée à la correction, parce que c'est ce qui m'intéressait prioritairement. Et notamment l'outil de recherche des répétitions qui est vraiment précieux. C'est même indispensable, je n'envisage plus de travailler sans.

Donc, les dictionnaires papier, un écran avec l'original, un écran avec ton logiciel de traitement de texte...

Et puis *Google* ouvert pour les recherches. Il y a aussi un dictionnaire en ligne que j'aime bien, c'est *WordReference*, puisqu'il propose des traductions et des liens éventuellement vers des discussions.

Donc pour les ressources terminologiques : dictionnaires papier, *WordReference*...

Oui. Dictionnaires papier français parce que... J'avais un *Robert & Collins*, mais je ne l'utilise plus parce que j'ai quand même rarement besoin de m'y référer. Si vraiment j'ai un doute sur un certain nom, un sens un peu rare, *WordReference* fait ça très bien, ou simplement une recherche *Google*, en fait. On trouve plein de ressources en ligne.

Pas d'autres outils que tu utiliserais régulièrement ?

Régulièrement ? Non, pas vraiment. *Wikipédia*, comme tout le monde, puis toutes les ressources du Net pour voir les recherches et les vérifications dont je parlais, plutôt thématiques, donc. Et mes cours de correction quand j'ai des doutes. Notamment, il y a toujours des points un peu

épineux sur certains accords de participes passés, des verbes pronominaux, ce genre de choses. Parfois j'ai un doute, donc je préfère me référer à mon cours. Ou sur des trucs de typo, d'ailleurs, sur l'emploi des italiques...

D'accord. Et tu m'avais dit que pour les prises de notes, c'était toujours classiquement soit une feuille papier, soit une feuille de traitement texte.

Oui, enfin, j'ai les deux en général. Je prends des notes à la main déjà pour noter les termes qui reviennent, certains choix de traduction, l'enchaînement des chapitres, la chronologie quand je sens qu'elle est importante, que ça va se dérouler sur un laps de temps assez court et qu'il est important quand même de garder le fil des événements. Ça, je le fais plutôt sur papier. Et après, oui, les choix de traduction récurrents, je les fais plutôt sur un fichier. Un fichier dans lequel je garde ce qui peut être pertinent de communiquer à l'éditeur en fin de traduction, pour expliquer certains choix ou pour lui signaler certaines erreurs que j'ai relevées, si j'ai des commentaires à faire...

Tu as une grande marge de manœuvre sur les textes, s'il y a une chose que tu décides de notifier à l'éditeur, par exemple ?

Oui oui, oui oui, enfin, en tout cas avec les éditeurs avec lesquels je travaille régulièrement quand même depuis un certain nombre d'années. Il y a déjà cette relation-là. Mais j'ai remarqué qu'en fait ils apprécient même d'avoir un petit fichier à côté, d'avoir certaines explications. Je ne le faisais pas jusqu'à récemment, c'était beaucoup plus succinct. Puis en fait, j'ai travaillé en collaboration avec une traductrice sur un bouquin où on s'est partagé les chapitres et elle m'a dit qu'elle le faisait aussi systématiquement. Et là, c'était super utile du coup, puisqu'on travaillait à deux. On avait vraiment besoin d'avoir un document que chacune enrichissait à tour de rôle en fonction de ses choix, de ses options. Et j'ai trouvé que c'était une façon de travailler intéressante.

Donc, quand vous travaillez à deux comme ça, vous vous répartissez le travail, mais avec des fichiers communs ?

Oui, on avait fait quatre quarts, en fait, parce que c'est toujours un peu embêtant : la première partie, c'est jamais la plus passionnante, c'est la partie d'exposition. Donc c'était pour que chacune ait quand même un peu le plaisir aussi d'être au milieu. Là, c'était un truc un peu exceptionnel. C'était parce que le délai de traduction était très court, bien sûr, c'était un Rick Riordan, donc

il fallait qu'il soit traduit en un claquement de doigts. Et c'était quand même difficilement envisageable de le faire seule. On me l'avait proposé en premier et donc l'éditeur a eu une bonne idée de proposer donc aux deux traductrices historiques de Rick Riordan, donc moi et puis Mona de Pracontal, donc on l'a fait à quatre mains.

Je me dis que pour quelqu'un qui connaît un peu la série ou qui s'intéresse aux processus de traduction, voire qui déciderait d'analyser les textes, ce serait un scénario intéressant à étudier.

Ah oui, oui, ce serait intéressant. C'est un peu déstabilisant parce qu'on a quand même peu l'habitude de travailler en collaboration. Donc il fallait être souples sur certaines choses, sur d'autres un peu plus fermes, pour défendre son point de vue. Oui, c'était intéressant. Mais bon, j'avoue que je préfère travailler seule quand même. Je trouve qu'on avance plus vite quand on est son seul interlocuteur.

J'imagine que ça demande aussi beaucoup de temps et d'efforts...

Ben oui, le temps des échanges, c'est du temps sur lequel on ne travaille pas, surtout que c'était quand même très pressé.

En parlant de gain de temps, tu m'as déjà dit que tu n'utilises jamais la traduction automatique ?

Non, jamais.

Même pour des mots ou des phrases isolées, c'est toujours par dictionnaire ?

Oui, dictionnaire papier ou dictionnaire en ligne : *WordReference*, qui donne quand même parfois des trucs en contexte, ou *Reverso*.

Est-ce qu'il y a une raison particulière...

Juste une question d'habitude, en fait. Une question d'habitude et de flemme de m'y mettre, peut-être probablement.

Flemme dans le sens où tu n'as pas forcément envie d'utiliser l'outil ou pas le temps d'utiliser encore quelque chose de plus ?

Plutôt ça. Quand les traductions s'enchaînent, ça laisse peu de temps entre deux pour changer ses habitudes. Et c'est vrai que je préfère avoir quelqu'un qui me montre, quitte même à suivre une formation rapide plutôt que de me plonger dans un manuel. J'ai horreur de ça. Moi, j'ai horreur de me plonger dans un manuel et d'essayer de comprendre. Je préfère qu'on me montre, quitte à prendre des notes. Mais c'est vraiment pour ça, en fait, que je ne l'ai jamais fait.

Est-ce que ça t'était déjà venu à l'idée avant qu'on en parle maintenant ?

Honnêtement, non, pas vraiment. Non, parce que je considérais que ce n'était vraiment pas au point que ça n'apportait pas grand-chose, en fait.

Et est ce que tu penses, au-delà du fait qu'on en ait parlé entre nous, que ça aurait pu changer ou que ça a changé avec justement toutes les discussions récentes autour de la TA et surtout la participation de l'ATLAS et de l'ATLF ?

Ce qui m'a intéressé, d'abord, c'est nos échanges, ça c'est sûr. Et bon après, forcément, je remarque davantage quand on en parle, maintenant. Surtout actuellement, on en parle beaucoup, c'est quand même un sujet assez récurrent.

Mais on ne t'a jamais proposé d'utiliser la traduction automatique ?

Non non, moi on ne me l'a jamais proposé. J'ai été assez étonnée de voir que certains collègues avaient été poussés... Même l'éditeur, en fait, qui leur avait proposé, sans doute pour des questions de contraintes de délai. Visiblement, c'est quand même un phénomène qui est loin d'être anecdotique, mais c'est un sujet que je n'avais jamais entendu... Enfin, on en avait parlé ensemble, mais entre collègues, je n'avais jamais entendu personne aborder cette question-là.

C'est quand même assez intéressant aussi de savoir que c'est une pratique qui commence à se répandre.

Oui, c'est ça. Donc je pense que peut-être plus de jeunes traducteurs ou traducteurs débutants y ont été confrontés. Peut-être que nous qui sommes dans le circuit depuis un certain nombre d'années, finalement, on a nos habitudes et des éditeurs qui ont leurs habitudes avec nous. Après, les jeunes collègues débutants qui acceptent le travail qu'on leur propose et qui sont en relation avec des éditeurs plus ou moins consciencieux sont sans doute beaucoup plus confrontés à ce genre de pratiques.

Pour en revenir à cet aspect pratique, justement, quelle est ton impression générale vis-à-vis de l'outil ? Tu me disais juste avant que nous entamions cette discussion que tu trouvais l'expérience un peu contraignante, surtout au niveau de la mise en place ?

Oui, oui, mhm.

Qu'as-tu pensé de ce que proposaient les outils ?

Et bien... Parfois, c'était intéressant, oui, effectivement, c'était un aspect intéressant. Après, j'ai un avis assez contrasté par rapport aux deux outils, dans le sens où je trouvais qu'ils se complétaient pas mal. D'avoir les deux, ça se complétait pas mal. Parce que parfois le premier, l'outil personnalisé donc, était plus pertinent. Mais en fait *DeepL* se défendait pas mal et évitait même certains contresens que faisait l'autre ou, même sans parler de contresens, quand il y avait un mot qui était susceptible d'avoir plusieurs traductions, choisissait davantage la bonne.

Est-ce que tu peux m'en dire plus, justement, sur cette différence entre les deux systèmes ?

Je trouvais que *DeepL* était plus généraliste, finalement. Dans le sens où il ne tenait pas forcément compte du contexte fantasy, merveilleux et qu'il ne se concentrait pas tellement sur toute la terminologie. Il ne faisait même pas forcément de proposition, en fait. Alors que le système personnalisé, oui oui, était nettement plus compétent à ce niveau-là. Je le sentais, pour le coup, sur les termes propres à l'ouvrage, puisque le système adapté était plus dans l'univers et proposait... L'autre même en général ne se mouillait pas et laissait le terme en anglais. Donc le système adapté, oui, était plus en phase avec l'univers et avec le type de proposition linguistique qu'on peut faire par rapport à certains termes ou à certaines expressions. Mais *DeepL* m'a étonnée par rapport à l'aspect strictement traduction. J'ai relevé quand même à plusieurs reprises des contresens étonnants, pour l'autre, ou alors un mot qui n'était pas traduit avec une bonne acception du terme dans le contexte.

Oui, ça rejoint la conclusion, un peu attendue finalement, des travaux qu'on a menés. De montrer que quand on adapte un système, il reflétera forcément mieux le genre, comparé à un système générique...

Oui, effectivement, ça c'est très net, oui.

Mais en sachant aussi que les deux systèmes ont leurs limites. On en a discuté pendant l'exercice, par exemple le fait que les propositions restent quand même très littérales. Et que notre système, finalement, reste difficilement comparable avec d'autres qui sont entraînés sur des millions et des millions d'exemples. Donc c'était un peu notre constat de dire qu'il y avait une réelle marque d'adaptation, mais que dans l'idéal, il faudrait un système aussi puissant ou robuste et qui en plus soit adapté. Un peu comme ces systèmes génériques qui sont censés s'entraîner sur les mémoires de traduction des traducteurs et traductrices, en se disant qu'on aurait peut-être là le point fort des deux côtés, mais en n'ayant plus qu'une seule proposition.

Évidemment, ce serait idéal, un système qui compilerait les points forts des deux et qui tomberait juste. Maintenant que tu me le dis, je comprends mieux, oui, on sentait que *DeepL* avait lu davantage de choses, c'est logique.

Donc pour toi, ce qui était intéressant, c'était vraiment d'avoir là la combinaison des deux ?

Oui oui, la combinaison des deux était intéressante, oui. C'était souvent assez proche, quand même, je remarquais qu'il y avait une grande proximité.

Comment abordais-tu d'ailleurs cette phase de comparaison, est-ce que tu regardais toutes les sorties en permanence...

Oui, je les comparais en permanence, mais parce que j'étais aussi dans la phase d'expérimentation et que, voilà, c'était effectivement un peu le but de l'opération, mais si je les utilisais tous les jours, non, je ne procéderaï probablement pas comme ça.

Est-ce que tu as l'impression qu'il y a une différence radicale de méthode ou de processus liée à l'utilisation de ces outils et dans le fait de passer par ce type d'interface ?

Euh, oui, forcément, forcément. Comme je le disais, je ne les utiliserais pas forcément sur toutes les phrases. Mais oui, quand j'aurais tendance à sécher un petit peu, tout en me disant que je ne me sens pas obligée d'adopter une des deux solutions, ça me fournirait déjà une base de réflexion.

Donc plutôt un outil à utiliser de manière ponctuelle ?

Oui. Dans les moments où je sors un peu de la phase automatique, où la solution vient tout de suite parce que la phrase est super simple ; c'est ça et c'est pas autre chose. Après, quand il y a

effectivement une plus grande part de réflexion, où j'ai une hésitation sur la façon dont je vais construire la phrase ou le paragraphe, oui, je trouve intéressant d'avoir déjà une sorte de matière brute à partir de laquelle je vais réfléchir.

Pour toi, la manière dont on fonctionnait ici avec une proposition de traduction qui s'affichait par défaut pour la phrase courante, est-ce que c'est une méthode qui te conviendrait ? Est-ce que tu changerais quoi que ce soit ?

Là, c'est vrai que je l'affichais systématiquement, chose que je ne ferais pas forcément. Après, j'en viens à ce qu'on disait, c'est vrai que ce serait bien d'avoir finalement un cadre plus large d'un paragraphe, je me sentirais quand même plus à l'aise pour tripatouiller l'intérieur.

Et est-ce que les mémoires de traduction ont ajouté quelque chose ?

J'avoue ne pas les avoir trop trop consultées, mais oui, j'en vois vraiment l'utilité, justement dans le cadre d'une série, oui oui, j'en vois vraiment l'utilité. Par rapport à mon propre glossaire, c'est pas mal de l'avoir dans le même outil finalement. Parce que c'est quand même assez facile de s'y référer. Ce que je fais bien sûr, moi, actuellement, c'est que je fais des recherches de termes dans les traductions antérieures, mais c'est plus laborieux. Ça oblige à ouvrir un autre fichier, tandis que là, avoir la petite fenêtre qui s'ouvre sur le terme précis, je trouve ça plutôt intéressant.

Donc l'outil hypothétique idéal, ce serait pour toi un outil externe qui te permette de travailler comme tu le fais maintenant, dans ton logiciel de traitement de texte, et de l'invoquer au moment où tu en as besoin ?

Ah oui, oui oui, ce serait intéressant, oui. Après, c'est peut-être aussi par habitude. C'est vrai que d'avoir le feuillet *Word*, c'est une habitude de travail, oui.

Est-ce que tu as l'impression que l'interface joue quand même sur ta liberté de traduction et d'intervention sur le texte ? Ou est-ce que tu penses qu'en étant plus habituée à travailler avec ce genre d'outils, ça pourrait aussi s'équilibrer ?

Je ne sais pas, je ne sais pas. C'est vrai que j'y ai prêté une attention très particulière puisqu'on était dans un cadre expérimental. Mais du coup, c'est long. [En riant] C'est long. Quand je suis dans des conditions de travail habituelles, j'avance quand même plus vite. Ce qui serait

très bien, voilà, ce serait quelque chose qui s'ouvre à la demande, quand même, sans passer systématiquement par cette interface ou cette présentation très particulière.

Oui, le fait que ce soit un cadre expérimental joue aussi, évidemment, mais tu as plutôt l'impression de perdre du temps par rapport à une traduction telle que tu la ferais habituellement ?

Ça avance plus lentement, ouais, ça avance plus lentement. Sachant que c'était aussi quand même une mise en route. Très objectivement, quand je commence une traduction, je n'avance quand même pas très vite sur les premières pages non plus. Il faut le temps de se familiariser avec le texte, l'environnement, les termes qu'on découvre et dont on ne sait pas très bien quelle importance ils vont avoir par la suite ni comment les traduire sur le moment, en se disant que c'est un peu bancal et que je serai probablement amenée à revenir dessus. On n'avance jamais très très vite quand on commence quelque chose.

Je suis rentrée dans un autre mode de fonctionnement qui était plus le mode de fonctionnement de la relecture en fait. C'est vrai, ça m'a un peu rappelé les fois où j'ai travaillé à partir d'un brouillon fait par quelqu'un d'autre. Sauf que quand je travaille à partir d'un brouillon fait par quelqu'un d'autre, j'ai tout de même la page qui s'affiche, donc ça avance quand même plus vite que de regarder une fenêtre. Mais, mentalement, c'était plutôt ce processus-là.

Donc je retiens surtout le côté contraignant de l'interface, sur lequel je te rejoins tout à fait. Au-delà de cet aspect, lorsque tu disais que c'était un outil qui pouvait quand même t'être utile, qu'est-ce que tu penses que ça pourrait apporter dans ta pratique ?

Autant il y a des moments où je dirais que, paradoxalement, ça libère un peu l'esprit, finalement. Parce que ça lui permet de rebondir et d'explorer des voies dans lesquelles il ne se serait pas forcément engagé, d'autres façons de voir et d'autres façons de faire. C'est une sorte de rêverie libératrice. On a une sensation de ne pas partir de rien et soit ça me plaît, soit ça ne me convient pas du tout, soit on est entre les deux, mais au moins il y a une réflexion qui s'amorce à partir de ce qui est proposé, pour en arriver à une solution personnelle.

Corrige si je me trompe, mais si j'essaie de synthétiser tout ça, il y a donc un réel intérêt pour amorcer la réflexion, mais qui est trop contraignant s'il est systématique ?

Et surtout aussi saucissonné, en fait. Je crois que la présentation y est pour beaucoup.

Donc si je comprends bien, ce qui te serait le plus utile, ce serait de pouvoir travailler avec une vue d'ensemble sur le texte et d'invoquer le système sur des passages spécifiques, quand tu décides de le faire.

Oui, oui oui, tout à fait.

Je me demandais aussi si tu avais remarqué des erreurs particulièrement problématiques ou plus embêtantes à corriger..

Finaleme nt non, parce qu'elles ne me posent pas problème. Je vois que c'est une erreur, donc je n'en tiens pas compte.

Des choses en particulier que tu devais impérativement remanier ?

[En riant] Tout. Beaucoup de choses, beaucoup de choses, oui. Tout ce que l'outil ne sait pas faire, de toute manière : synthétiser des idées, synthétiser des phrases. Voilà, s'il ne peut pas dépasser son côté littéral. Donc ça, non, ça ne reflète pas ma façon de travailler, ma façon de réfléchir, c'est sûr. Mais bon, je pense que c'est hors de sa portée.

Oui, évidemment, on en a discuté tout à l'heure. Sur ce point, justement, est-ce que ça a changé la vision que tu avais de ces outils-là, ou est ce que c'est plus ou moins ce à quoi tu t'attendais ? Est-ce que ça a renforcé des idées que tu avais déjà ou est-ce que ce ça a changé la vision que tu avais de la traduction automatique de manière générale ?

[Hésitation] Non... J'ai été plutôt quand même heureusement surprise par l'outil.

Que tu te verrais utiliser par la suite ?

Oui, le cas échéant, oui, oui.

Un autre thème qui revient souvent, c'est cette idée de contrainte et de se sentir influencé par la traduction automatique. On a discuté du fait que ça pouvait être utile pour lancer des idées, surtout si c'était utilisé de manière plus interactive, sur demande. Est-ce que tu ressentais cette contrainte ou est-ce que tu le voyais vraiment juste comme une proposition qui justement pouvait ou non être utile ?

Pour moi, le côté contrainte, il est plus dans l'interface, effectivement. Dans le fait, donc, de passer d'une ligne à l'autre, de devoir jeter un coup d'œil en bas, regarder quelle est la proposition...

Je n'ai pas vraiment ressenti de contrainte intellectuelle, je me sentais complètement libre par rapport à la proposition.

Donc tu te sentais tout à libre de changer...

Oui, oui oui, complètement, oui.

Le problème, c'est plutôt le logiciel qui empêche de modifier la verticalité en fait ?

Voilà, et puis qui est un peu lourd après au niveau même simplement du traitement de texte, qui est moins réactif qu'un autre logiciel à ce niveau-là. Je n'ai pas retrouvé certains raccourcis, par exemple.

Oui, ici ça rejoint un peu ce que tu disais vis-à-vis des habitudes de travail.

Oui, mhm.

Si l'on restait dans l'hypothétique, puisque tu mentionnais que tu te verrais éventuellement utiliser ce type d'outil, est-ce que tu utiliserais les outils existants, déjà ? Sinon, que changerais-tu si tu en avais la possibilité ? Qu'est-ce qui ferait que tu trouverais plus intéressant de les utiliser ?

Qu'il soit déjà plus maniable, quand même plus maniable. Par exemple, là, effectivement, je voudrais bien voir une proposition, mettons : hop, je surligne et ça s'ouvre à cet endroit-là. Je trouverais ça pratique vraiment vraiment pratique. Comme peut le faire par exemple *Antidote*, quand on est dans la partie correction. Et que l'interface soit plus conviviale, plus légère d'utilisation...

Par exemple, continuer de travailler dans *Word* et pouvoir appeler la TA quand il le faut...

Oui, oui oui, dans l'idéal oui, il ferait ça, effectivement. Les choses s'ouvriraient tout à coup quand je le demanderais, voilà, d'un clic : je surligne, je clique et hop, j'ai une bulle qui s'ouvre avec les propositions.

Tu ne te verrais pas, par exemple, travailler dans un outil comme *Trados*, sans la segmentation ? Imaginons par exemple qu'il te permette de travailler un peu comme dans un document *Word*, que tu aies la même interface, mais à la place d'avoir le texte segmenté

par phrase, tu aurais une fenêtre un peu plus réduite de traitement de texte. Ça te paraît quand même plus contraignant ?

Avec les propositions qui s'afficheraient de la même manière en bas en bas ? C'est pas idéal. C'est pas idéal, parce que ça déconcentre un peu quand même. Ce serait déjà mieux. Mais, oui, dans l'idéal, ce serait vraiment quand même bien de pouvoir l'appeler quand on veut, comme quand on veut faire un commentaire dans *Word*, dans la marge.

Donc c'est vraiment ça qui est important pour toi, c'est le côté sollicitation...

Oui, c'est ça, oui, mhm.

Tu penses que tu serais amenée à le solliciter souvent si tu travaillais dans *Word* ?

Souvent, je ne sais pas. Ça dépendrait du texte aussi, mais... De la même manière qu'il m'arrive de m'interrompre et de faire des recherches *Google*, comme ça, sur des mots ou sur des enchaînements de mots, voir comment ça sonne, voir si d'autres l'ont utilisé, dans quel contexte ça a été utilisé... Sauf que là, ce serait plus synthétique, parce que ça s'ouvrirait automatiquement.

Donc pour toi, ce serait véritablement une ressource qui s'ajoute au dictionnaire, au dictionnaire de cooccurrences...

Voilà, une ressource... Oui, voilà, complètement, oui oui.

Et mettons qu'un éditeur te propose maintenant de travailler avec la traduction automatique. Est-ce que tu accepterais ? Et si oui, sous quelles conditions ?

[En riant] Est ce que j'accepterais... ? J'accepterais, à condition d'être quand même complètement maître du rendu final. Si vraiment on avait une grosse contrainte de temps, par exemple. Ça peut arriver, d'ailleurs il m'est déjà arrivé à ce niveau de demander à quelqu'un de me faire un premier jet. Donc à ce moment-là, je pourrais l'envisager comme ça, oui. Mais il faudrait qu'il progresse un peu quand même, peut être, pour que ce soit vraiment intéressant, parce...

[Rires]

Justement, si tu devais donner des recommandations, mettons pour une utilisation la plus simple et la plus utile possible — on a parlé de l'interface, du fait de pouvoir rester maître de l'outil... —, est-ce qu'il y a autre chose que tu changerais ou que tu mettrais en avant ?

Non, non non, c'est tout à fait ça.

Donc, les performances, aussi, qui restent parfois un peu bancales...

Oui, même si encore une fois elles ne sont pas si bancales que ça. J'ai quand même été parfois heureusement surprise.

Et est-ce que, à l'inverse, le fait d'avoir utilisé ces outils-là ou les discussions qui commencent un peu à prendre pour l'instant te font avoir des craintes particulières ? Est-ce qu'il y a quelque chose que tu redoutes dans le fait qu'on voie l'ATLF dire qu'il y a quand même des personnes qui travaillent avec ces outils-là et que les traducteurs et traductrices pourraient être amenés à recevoir des propositions pour travailler ce genre d'outils. Qu'est-ce que tu vois comme point négatif dans ce genre d'expérience, de méthode de traduction, ou simplement dans le fait que ces outils sont de plus en plus présents ?

Je comprends que l'ATLF éprouve le besoin de lancer cette enquête. Et puis, au vu du résultat, on voit qu'il y a un problème. Il y a un vrai problème. Moi, ça ne me choquerait pas tant que le traducteur reste considéré comme auteur et qu'il garde la maîtrise du résultat final. Après tout, bon, si pour des contraintes de temps on est amené à utiliser ce genre d'outil sur une traduction, dès lors que l'on peut quand même avoir la dernière touche sur le texte et être considéré comme auteur, comme on l'est actuellement — donc avec son nom sur la couverture — et puis surtout être rémunéré avec des droits proportionnels, pourquoi pas ? Après, le glissement dangereux, c'est que le traducteur se retrouve face à un premier jet qu'il n'a pas sollicité et soit considéré comme un simple... Qu'il ne soit même plus traducteur, en fait, ni même un préparateur de copies. Qu'il fasse juste de la post-édition en n'étant pas considéré comme auteur du texte, en ayant juste une rémunération forfaitaire, et puis finalement en n'étant même pas assuré que ce soit la version qu'il aura proposée qui soit éditée en l'état. Là, ça n'a plus rien à voir avec ce qu'est la traduction.

Donc, au-delà de la reconnaissance et de la rémunération, ce qui te paraît important, si je comprends bien, c'est le fait d'être sûr que c'est ce qu'on a proposé qui sera publié ?

Oui, aussi, oui, tout à fait. Ce qui est le cas actuellement parce qu'il y a tout ce retour de préparation, retour d'épreuve... où on peut discuter avec un interlocuteur en disant : « Je ne suis pas d'accord avec votre proposition. » Ce qui fait qu'en général, quand le texte est publié, quand

on a affaire à un éditeur sérieux, on est d'accord sur le résultat final. Et on est considéré comme l'auteur du résultat final.

Et est ce que tu penses que l'intervention de la traduction automatique pourrait justement briser ce...

J'en ai peur. Ça, j'en ai peur, oui, effectivement.

Qu'est ce qui te poserait le plus problème, finalement, est-ce un problème de qualité, de méthode, le fait de ne pas être reconnu...

C'est un tout. Ce serait simplement plus le même métier. Ce ne serait plus la même profession.

Oui, c'est une bonne manière de résumer cela, que l'on voit souvent revenir, d'ailleurs. La question est volontairement ouverte, mais comment est-ce que tu imagines ton métier dans les 5 à 10 prochaines années ?

Je pense que ça va évoluer. Oui, je pense que ça va évoluer. Dans quelles proportions, je ne sais pas... Je ne pense pas quand même qu'il y aura un raz-de-marée généralisé, ne serait-ce que parce qu'il y a aussi quand même une génération d'éditeurs qui n'ont pas travaillé comme ça, qui sont encore en place, qui ont encore un petit peu de poids, probablement, face à leur direction, et qui peuvent quand même défendre une certaine manière de travailler. Mais je suis quand même un petit peu inquiète pour les 5 à 10 années à venir, oui. Ça a tellement d'avantages pour les éditeurs que j'ai très peur que la pratique se répande assez vite.

Et avec la technologie actuelle ?

Certains le font déjà. [En riant] Certains le font déjà, donc c'est qu'ils trouvent leur intérêt.

Ouais, ouais ouais, évidemment. Et bien je pense qu'on a abordé pas mal de choses : on a parlé de l'outil lui-même, de la différence par rapport à tes méthodes de travail actuelles, de ce que ça pourrait amener pour la suite... Est-ce qu'il y a quoi que ce soit qu'on n'a pas abordé et que tu aurais envie d'ajouter ou qui te semble primordial, de partager des questions actuelles ?

Non, là je ne vois pas, non. Juste, ce que j'ai envie de dire, c'est que je considère moi mon métier comme un artisanat. Alors, vraiment, je ne me qualifie pas d'artiste, même si je suis statutairement

artiste-auteur. Mais je suis très attachée à ce terme d'artisanat, et à tout ce que ça représente de travail patient, d'élaboration, de raffinement à partir d'un bloc brut pour arriver... Et donc avoir des outils, oui, un artisan a tout à fait besoin d'outils adaptés à sa pratique, tant que la partie artisanale reste quand même primordiale et qu'il est reconnu comme un artisan. Donc je crains effectivement que cette partie-là en vienne à se perdre avec la généralisation de la traduction automatique. Ce n'est plus le même travail, ce n'est pas la même chose.

Tu penses que c'est inévitable ?

Euh... Poser des garde-fous, c'est compliqué. Pour ça, je n'ai pas très confiance dans le SNE, le Syndicat national de l'édition. En fait, je serais très intéressée de savoir qui sont les éditeurs qui y ont recours actuellement. S'il y en a parmi les gros éditeurs, qui constituent quand même le gros du conseil d'administration du SNE, ou si ce sont vraiment de jeunes éditeurs qui se lancent dans des domaines un peu marginaux. C'est aussi ça que j'aimerais savoir : dans quel domaine de l'édition cette pratique est la plus répandue.

Et en sachant qu'ils sont déjà utilisés, est-ce que tu vois un moyen pour ces outils-là de se développer — je prêche pour ma paroisse, mais... — un peu comme on les défend, c'est-à-dire comme des outils qui devraient être davantage au service des traducteurs et traductrices, où est-ce que tu penses que ça va inévitablement se répandre de la manière dont ils sont utilisés maintenant et finalement desservir l'artisanat ?

Et bien si je suis pessimiste, avec le recul sur les toutes les évolutions technologiques, ce n'est pas souvent la version la plus adaptée au travailleur, à l'utilisateur, qui prévaut, mais celle qui est la plus économiquement rentable. Donc, dans l'idéal, oui, ce serait formidable que l'outil soit adapté à l'usage que veut en faire l'artisan ou le créateur. Ce serait vraiment l'idéal.

Et tu ne penses pas qu'il y aurait moyen, avec l'intervention des associations par exemple...

Si, si, voilà. Oui, forcément. Oui, oui oui. Je crois quand même beaucoup à la puissance du tiers secteur associatif, de tous les gens qui mettent leur talent, leurs compétences au service du plus grand nombre, de manière souvent gratuite ou en tout cas sans but lucratif. Après, pour le reste, je n'ai aucune confiance dans les gros laboratoires et les gros fabricants de logiciels. Et forcément qu'ils vont rechercher la rentabilité, ça me paraît évident.

Donc tu ne verrais pas de scénario ou, mettons, le monde de la recherche et le secteur associatif arrivent à convaincre la société et les éditeurs qu'il y a une autre façon de faire les choses et qu'on peut développer les outils...

Ah, ce serait bien, j'aimerais bien. [En riant] En tout cas, ça vaut la peine d'essayer.

Mais... Il y a un mais ?

Je pense qu'on peut convaincre et qu'il y a quand même aussi des éditeurs vertueux. Qui ne sont pas vertueux qu'en façade et qui continueront à privilégier la qualité du travail... Il y a même des éditeurs très militants. Donc je n'ai pas de souci par rapport à ceux-là. Après, pour ceux qui recherchent d'abord la rentabilité... Et je ne parle pas des éditeurs qui sont en position de faire les choix éditoriaux, ceux-là la je ne les remets pas du tout en cause, c'est plutôt ceux qui sont au-dessus d'eux qui posent problème. Et dans les grands groupes, un des éditeurs pour qui je travaille, par exemple, va encore changer de mains et les salariés se posent beaucoup de questions... Et pas que les salariés. On ne sait pas du tout ce que va devenir la ligne éditoriale et quels vont être les diktats imposés par les financiers.

Est-ce que tu penses qu'il pourrait y avoir une sorte de fracture entre, d'un côté, des petits éditeurs qui privilégieraient le *statu quo* et d'autres qui iraient plutôt vers la rentabilité ?

Oui, enfin ils ne sont pas tous vertueux, mais enfin... Peut-être pas forcément de manière aussi schématique ni aussi polarisée, mais je pense que ça existe déjà un peu, même sans rajouter la composante de la traduction automatique dans les pratiques, oui, ça existe déjà.

Mais ça va être intéressant à regarder, hein. Enfin, je le dis parce que je suis quand même plus proche de la retraite que du traducteur débutant, mais ça va être très intéressant à observer.

Est-ce que tu as des doutes quand même pour les gens qui commenceraient leur carrière maintenant ?

Euh, oui, quand même. J'ai pour principe de ne jamais décourager les gens qui veulent se lancer dans cette profession. Au contraire, je les encourage. Mais, je pense qu'il va quand même être un peu plus difficile de commencer. En tous cas pas dans les mêmes conditions. Que les gens seront sans doute amenés à accepter des conditions de travail quand même très différentes.

Tu n'imagines pas que les traducteurs et traductrices puissent faire entendre leur voix et refuser en bloc, par exemple, l'usage de la traduction automatique ?

Et bien il faut qu'ils s'organisent. Il faut qu'ils se syndiquent...

À titre personnel, je suis... absolument d'accord...

[Rires]

Oui... C'est intéressant parce que c'est aussi ce que proposent les gens du côté de l'ingénierie, qui bossent depuis des dizaines d'années sur le traitement automatique des langues en voyant un peu maintenant leur recherche être réappropriée par les gens qui vont commercialiser tout ça pour des sommes astronomiques, et qui disent : « Mais syndiquez-vous ! » C'était d'ailleurs même la conclusion de deux conférences auxquelles j'ai participé dernièrement, une dans un champ plutôt lié à l'ingénierie et l'autre plutôt à la traductologie.

Ah oui...

À voir donc comment ça se déroulera dans les prochaines années et quel rôle les associations et autres maillons de la traduction auront à jouer là-dedans...

Bibliographie

Programmes

Antidote. Druide informatique, 1996.

ChatGPT. OpenAI, 2022. URL : <https://chat.openai.com/>.

DeepL Translate. DeepL GmbH 2017. URL : <https://www.deepl.com/translator/>.

DeepL Write. DeepL GmbH, 2023. URL : <https://www.deepl.com/write/>.

Google Livres. Google, 2004. URL : <https://books.google.fr/>.

Google Traduction. Google, 2006. URL : <https://translate.google.fr/>.

Intergaelic. Kevin SCANNELL et Michal MĚCHURA, 2015. URL : <http://www.intergaelic.com/>.

LogiTerm. Terminotix, 2000.

OBS Studio. OBS Project, 2012. URL : <https://obsproject.com/>.

OmegaT. OmegaT Project, 2002. URL : <https://omegat.org/>.

OPUS-CAT. Tommi NIEMINEN, 2021. URL : <https://helsinki-nlp.github.io/OPUS-CAT/>.

Trados Studio 2022. RWS Group, 2022.

Outils et corpus

AL-ONAIZAN, Yaser, Jan CURIN, Michael JAHR, Kevin KNIGHT, John LAFFERTY, Dan MELAMED, Franz Josef OCH, David PURDY, Noah SMITH et David YAROWSKY (1999). « Statistical Machine Translation ». Rapport du *JHU Summer Workshop 1999*, Baltimore (MD), CLSP/JHU. URL : <https://aclanthology.org/www.mt-archive.info/JHU-1999-AlOnaizan.pdf>.

ALABAU, Vicent, Christian BUCK, Michael CARL, Francisco CASACUBERTA, Mercedes GARCÍA-MARTÍNEZ, Ulrich GERMANN, Jesús GONZÁLEZ-RUBIO, Robin HILL, Philipp KOEHN, Luis LEIVA, Bartolomé MESA-LAO, Daniel ORTIZ-MARTÍNEZ, Herve SAINT-AMAND, Germán SANCHIS TRILLES et Chara TSOUKALA (2014). « CASMACAT: A Computer-assisted Translation Workbench ». *Proceedings of the Demonstrations at the 14th Conference of the European Chapter*

- of the Association for Computational Linguistics, édité par Shuly WINTNER, Marko TADIĆ et Bogdan BABYCH, Göteborg, ACL, pages 25–28. DOI : [10.3115/v1/E14-2007](https://doi.org/10.3115/v1/E14-2007).
- AZIZ, Wilker, Sheila CASTILHO et Lucia SPECIA (2012). « PET: a Tool for Post-editing and Assessing Machine Translation ». *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, édité par Nicoletta CALZOLARI, Khalid CHOUKRI, Thierry DECLERCK, Mehmet Uğur DOĞAN, Bente MAEGAARD, Joseph MARIANI, Asuncion MORENO, Jan ODIJK et Stelios PIPERIDIS, Istanbul, ELRA, pages 3982–3987. URL : <https://aclanthology.org/L12-1587>.
- BANERJEE, Satanjeev et Alon LAVIE (2005). « METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments ». *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, édité par Jade GOLDSTEIN, Alon LAVIE, Chin-Yew LIN et Clare VOSS, Ann Arbor (MI), ACL, pages 65–72. URL : <https://aclanthology.org/W05-0909>.
- DENKOWSKI, Michael et Alon LAVIE (2014). « Meteor Universal: Language Specific Translation Evaluation for Any Target Language ». *Proceedings of the Ninth Workshop on Statistical Machine Translation*, édité par Ondřej BOJAR, Christian BUCK, Christian FEDERMANN, Barry HADDOW, Philipp KOEHN, Christof MONZ, Matt POST et Lucia SPECIA, Baltimore (MD), ACL, pages 376–380. DOI : [10.3115/v1/W14-3348](https://doi.org/10.3115/v1/W14-3348).
- DOU, Zi-Yi et Graham NEUBIG (2021). « Word Alignment by Fine-tuning Embeddings on Parallel Corpora ». *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, édité par Paola MERLO, Jörg TIEDEMANN et Reut TSARFATY, Kiev, ACL, pages 2112–2128. DOI : [10.18653/v1/2021.eacl-main.181](https://doi.org/10.18653/v1/2021.eacl-main.181).
- EDER, Maciej, Jan RYBICKI et Mike KESTEMONT (2016). « Stylometry with R: A Package for Computational Text Analysis ». *The R Journal*, vol. 8, n° 1, pages 107–121. DOI : [10.32614/RJ-2016-007](https://doi.org/10.32614/RJ-2016-007).
- ERJAVEC, Tomaž (2013). « The SPOOK parallel corpus: annotation, enoding and search ». *Slovene translations through a corpus prism*, édité par Špela VINTAR, Ljubljana, Znanstvena založba Filozofske fakultete, pages 14–31.
- ESPERANÇA-RODIER, Emmanuelle et Francis BRUNET-MANQUAT (2021). « ACCOLÉ : Annotation Collaborative d'erreurs de traduction pour CORpus aLignÉS, multi-cibles, et Annotation d'Expressions Poly-lexicales ». *Actes de la 28e Conférence sur le Traitement Automatique des Langues Naturelles*, vol. 3, édité par Pascal DENIS, Natalia GRABAR, Amel FRAISSE, Rémi

- CARDON, Bernard JACQUEMIN, Eric KERGOSIEN et Antonio BALVET, Lille, ATALA, pages 1–5.
URL : <https://aclanthology.org/2021.jeptalnrecital-demo.1>.
- FAUCONNIER, Jean-Philippe (2015). *French Word Embeddings*. URL : <http://fauconnier.github.io>.
- FORCADA, Mikel L., Mireia GINESTÍ-ROSELL, Jacob NORDFALK, Jim O'REGAN, Sergio ORTIZ-ROJAS, Juan Antonio PÉREZ-ORTIZ, Felipe SÁNCHEZ-MARTÍNEZ, Gema RAMÍREZ-SÁNCHEZ et Francis M. TYERS (2011). « Apertium: a free/open-source platform for rule-based machine translation ». *Machine Translation*, vol. 25, n° 2, pages 127–144. DOI : [10.1007/s10590-011-9090-0](https://doi.org/10.1007/s10590-011-9090-0).
- HEINZERLING, Benjamin et Michael STRUBE (2018). « BPEmb: Tokenization-free Pre-trained Subword Embeddings in 275 Languages ». *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*, édité par Nicoletta CALZOLARI, Khalid CHOUKRI, Christopher CIERI, Thierry DECLERCK, Sara GOGGI, Koiti HASIDA, Hitoshi ISAHARA, Bente MAEGAARD, Joseph MARIANI, Hélène MAZO, Asuncion MORENO, Jan ODIJK, Stelios PIPERIDIS et Takenobu TOKUNAGA, Miyazaki, ELRA, pages 2989–2993. URL : <http://www.lrec-conf.org/proceedings/lrec2018/summaries/1049.html>.
- HONNIBAL, Matthew, Ines MONTANI, Sofie VAN LANDEGHEM et Adriane BOYD (2020). *spaCy: Industrial-strength Natural Language Processing in Python*. DOI : [10.5281/zenodo.1212303](https://doi.org/10.5281/zenodo.1212303).
- JUNCZYS-DOWMUNT, Marcin, Roman GRUNDKIEWICZ, Tomasz DWOJAK, Hieu HOANG, Kenneth HEAFIELD, Tom NECKERMANN, Frank SEIDE, Ulrich GERMANN, Alham FIKRI AJI, Nikolay BOGOYCHEV, André MARTINS et Alexandra BIRCH (2018). « Marian: Fast Neural Machine Translation in C++ ». *Proceedings of ACL 2018, System Demonstrations*, édité par Fei LIU et Tamar SOLORIO, Melbourne, ACL, pages 116–121. DOI : [10.18653/v1/P18-4020](https://doi.org/10.18653/v1/P18-4020).
- KLEIN, Guillaume, Yoon KIM, Yuntian DENG, Jean SENELLART et Alexander RUSH (2017). « OpenNMT: Open-Source Toolkit for Neural Machine Translation ». *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, édité par Mohit BANSAL et Heng Ji, Vancouver, ACL, pages 67–72. URL : <https://aclanthology.org/P17-4012>.
- KOEHN, Philipp, Hieu HOANG, Alexandra BIRCH, Chris CALLISON-BURCH, Marcello FEDERICO, Nicola BERTOLDI, Brooke COWAN, Wade SHEN, Christine MORAN, Richard ZENS, Chris DYER, Ondřej BOJAR, Alexandra CONSTANTIN et Evan HERBST (2007). « Moses: Open Source Toolkit for Statistical Machine Translation ». *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (Companion Volume Proceedings of the Demo and Poster Sessions)*,

- édité par Sophia ANANIADOU, Prague, ACL, pages 177–180. URL : <https://aclanthology.org/P07-2045>.
- KUDO, Taku (2018a). « SentencePiece: A simple and language independent subword tokenizer and detokenizer for Neural Text Processing ». *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, édité par Eduardo BLANCO et Wei LU, Bruxelles, ACL, pages 66–71. DOI : [10.18653/v1/D18-2012](https://doi.org/10.18653/v1/D18-2012).
- LACOSTE, Alexandre, Alexandra LUCCIONI, Victor SCHMIDT et Thomas DANDRES (2019). « Quantifying the Carbon Emissions of Machine Learning ». Prépublication *arXiv*, 2^e éd. ARXIV : [1910.09700v2](https://arxiv.org/abs/1910.09700v2).
- LISON, Pierre et Jörg TIEDEMANN (2016). « OpenSubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles ». *Proceedings of the Tenth International Conference on Language Resources and Evaluation*, édité par Nicoletta CALZOLARI, Khalid CHOUKRI, Thierry DECLERCK, Sara GOGGI, Marko GROBELNIK, Bente MAEGAARD, Joseph MARIANI, Helene MAZO, Moreno ASUNCION, Jan ODIJK et Stelios PIPERIDIS, Portorož, ELRA, pages 923–929. URL : <https://aclanthology.org/L16-1147>.
- LO, Chi-kiu (2019). « YiSi - A Unified Semantic MT Quality Evaluation and Estimation Metric for Languages with Different Levels of Available Resources ». *Proceedings of the Fourth Conference on Machine Translation*, vol. 2, édité par Ondřej BOJAR, Rajen CHATTERJEE, Christian FEDERMANN, Mark FISHEL, Yvette GRAHAM, Barry HADDOW, Matthias HUCK, Antonio JIMENO YEPES, Philipp KOEHN, André MARTINS, Christof MONZ, Matteo NEGRI, Aurélie NÉVÉOL, Mariana NEVES, Matt POST, Marco TURCHI et Karin VERSPOOR, Florence, ACL, pages 507–513. DOI : [10.18653/v1/W19-5358](https://doi.org/10.18653/v1/W19-5358).
- MICHALKE, Meik (2021). *koRpus: Text Analysis with Emphasis on POS Tagging, Readability, and Lexical Diversity*. URL : <https://cran.r-project.org/web/packages/koRpus/index.html>.
- NIEMINEN, Tommi (2021). « OPUS-CAT: Desktop NMT with CAT integration and local fine-tuning ». *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, édité par Dimitra GKATZIA et Djamé SEDDAH, ACL, pages 288–294. DOI : [10.18653/v1/2021.eacl-demos.34](https://doi.org/10.18653/v1/2021.eacl-demos.34).
- OCH, Franz Josef et Hermann NEY (2003). « A Systematic Comparison of Various Statistical Alignment Models ». *Computational Linguistics*, vol. 29, n^o 1, pages 19–51. DOI : [10.1162/089120103321337421](https://doi.org/10.1162/089120103321337421).
- OLIVER, Antoni González (2020). « MTUOC: easy and free integration of NMT systems in professional translation environments ». *Proceedings of the 22nd Annual Conference of the*

- European Association for Machine Translation*, édité par André MARTINS, Helena MONIZ, Sara FUMEGA, Bruno MARTINS, Fernando BATISTA, Luisa COHEUR, Carla PARRA, Isabel TRANCOSO, Marco TURCHI, Arianna BISAZZA, Joss MOORKENS, Ana GUERBEROF-ARENAS, Mary NURMINEN, Lena MARG et Mikel L. FORCADA, Lisbonne, EAMT, pages 467–468. URL : <https://aclanthology.org/2020.eamt-1.55>.
- PAPINENI, Kishore, Salim ROUKOS, Todd WARD et Wei-Jing ZHU (2002). « BLEU: A Method for Automatic Evaluation of Machine Translation ». *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, édité par Pierre ISABELLE, Eugene CHARNIAK et Dekang LIN, Philadelphie (PA), ACL, pages 311–318. DOI : [10.3115/1073083.1073135](https://doi.org/10.3115/1073083.1073135).
- POPOVIĆ, Maja (2017). « ChrF++: Words Helping Character N-grams ». *Proceedings of the Second Conference on Machine Translation*, édité par Ondřej BOJAR, Christian BUCK, Rajen CHATTERJEE, Christian FEDERMANN, Yvette GRAHAM, Barry HADDOW, Matthias HUCK, Antonio JIMENO YEPES, Philipp KOEHN et Julia KREUTZER, Copenhague, ACL, pages 612–618. DOI : [10.18653/v1/W17-4770](https://doi.org/10.18653/v1/W17-4770).
- POST, Matt (2018). « A Call for Clarity in Reporting BLEU Scores ». *Proceedings of the Third Conference on Machine Translation: Research Papers*, édité par Ondřej BOJAR, Rajen CHATTERJEE, Christian FEDERMANN, Mark FISHEL, Yvette GRAHAM, Barry HADDOW, Matthias HUCK, Antonio JIMENO YEPES, Philipp KOEHN, Christof MONZ, Matteo NEGRI, Aurélie NÉVÉOL, Mariana NEVES, Matt POST, Lucia SPECIA, Marco TURCHI et Karin VERSPOOR, Bruxelles, ACL, pages 186–191. DOI : [10.18653/v1/W18-6319](https://doi.org/10.18653/v1/W18-6319).
- QI, Peng, Yuhao ZHANG, Yuhui ZHANG, Jason BOLTON et Christopher D. MANNING (2020). « Stanza: A Python Natural Language Processing Toolkit for Many Human Languages ». *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, édité par CELIKYILMAZ, ASLI AND WEN, TSUNG-HSIEN, ACL, pages 101–108. DOI : [10.18653/v1/2020.acl-demos.14](https://doi.org/10.18653/v1/2020.acl-demos.14).
- RAMÍREZ-SÁNCHEZ, Gema, Jaime ZARAGOZA-BERNABEU, Marta BAÑÓN et Sergio ORTIZ-ROJAS (2020). « Bifixer and Bicleaner: Two Open-Source Tools to Clean Your Parallel Data ». *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, édité par MARTINS, ANDRÉ AND MONIZ, HELENA AND FUMEGA, SARA AND MARTINS, BRUNO AND BATISTA, FERNANDO AND COHEUR, LUISA AND PARRA, CARLA AND TRANCOSO, ISABEL AND TURCHI, MARCO AND BISAZZA, ARIANNA AND MOORKENS, JOSS AND GUERBEROF-ARENAS, ANA AND NURMINEN, MARY AND MARG, LENA AND FORCADA, MIKEL L., Lisbonne, EAMT, pages 291–298. URL : <https://www.aclweb.org/anthology/2020.eamt-1.31>.

- ŘEHŮŘEK, Radim et Petr SOJKA (2010). « Software Framework for Topic Modelling with Large Corpora ». *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, La Valette, ELRA, pages 45–50. URL : <http://www.lrec-conf.org/proceedings/lrec2010/workshops/W10.pdf>.
- REI, Ricardo, Craig STEWART, Ana FARINHA et Alon LAVIE (2020). « COMET: A Neural Framework for MT Evaluation ». *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, édité par Bonnie WEBBER, Trevor COHN, Yulan HE et Yang LIU, ACL, pages 2685–2702. DOI : [10.18653/v1/2020.emnlp-main.213](https://doi.org/10.18653/v1/2020.emnlp-main.213).
- SCHMID, Helmut (1994). « Probabilistic Part-of-Speech Tagging Using Decision Trees ». *Proceedings of International Conference on New Methods in Language Processing*, Manchester. URL : <https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger>.
- SCHWENK, Holger, Guillaume WENZEK, Sergey EDUNOV, Edouard GRAVE, Armand JOULIN et Angela FAN (2021). « CCMatrix: Mining Billions of High-Quality Parallel Sentences on the Web ». *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, vol. 1, édité par Chengqing ZONG, Fei XIA, Wenjie LI et Roberto NAVIGLI, ACL, pages 6490–6500. DOI : [10.18653/v1/2021.acl-long.507](https://doi.org/10.18653/v1/2021.acl-long.507).
- SENNRICH, Rico, Orhan FIRAT, Kyunghyun CHO, Alexandra BIRCH, Barry HADDOW, Julian HITSCHLER, Marcin JUNCZYS-DOWMUNT, Samuel LÄUBLI, Antonio Valerio MICELI BARONE, Jozef MOKRY et Maria NÁDEJDE (2017). « Nematus: A Toolkit for Neural Machine Translation ». *Proceedings of the Software Demonstrations of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, édité par André MARTINS et Anselmo PEÑAS, Valence, ACL, pages 65–68. URL : <https://aclanthology.org/E17-3017>.
- SHEN, Lucas (2022). *LexicalRichness: A small module to compute textual lexical richness*. DOI : [10.5281/zenodo.6607007](https://doi.org/10.5281/zenodo.6607007).
- SIMARD, Michel, George FOSTER et Pierre ISABELLE (1993). « Using cognates to align sentences in bilingual corpora ». *Proceedings of the 1993 conference of the Centre for Advanced Studies on Collaborative research: distributed computing*, vol. 2, édité par Ann GAWMAN, Evelyn KIDD et Per-Åke LARSON, Toronto, IBM Press, pages 1071–1082. URL : <https://dl.acm.org/doi/10.5555/962367.962411>.
- SNOVER, Matthew, Bonnie DORR, Rich SCHWARTZ, Linnea MICCIULLA et John MAKHOUL (2006). « A Study of Translation Edit Rate with Targeted Human Annotation ». *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers*, édité

par Laurie GERBER, Nizar HABASH et Alon LAVIE, Cambridge (MA), AMTA, pages 223–231.

URL : <https://aclanthology.org/2006.amta-papers.25>.

STOLCKE, Andreas (2002). « SRILM - an extensible language modeling toolkit ». *Proceedings of the 7th International Conference on Spoken Language Processing*, édité par John HANSEN et Bryan PELLOM, Denver (CO), Causal Productions Pty Ltd, pages 901–904. DOI : [10.21437/ICSLP.2002-303](https://doi.org/10.21437/ICSLP.2002-303).

TIEDEMANN, Jörg (2012). « Parallel Data, Tools and Interfaces in OPUS ». *Proceedings of the Eighth International Conference on Language Resources and Evaluation*, édité par Nicoletta CALZOLARI, Khalid CHOUKRI, Thierry DECLERCK, Mehmet Uğur DOĞAN, Bente MAEGAARD, Joseph MARIANI, Asuncion MORENO, Jan ODIJK et Stelios PIPERIDIS, Istanbul, ELRA, pages 2214–2218. URL : http://www.lrec-conf.org/proceedings/lrec2012/pdf/463_Paper.pdf.

VANROY, Bram, Orphée De CLERCQ, Arda TEZCAN, Joke DAEMS et Lieve MACKEN (2021). « Metrics of Syntactic Equivalence to Assess Translation Difficulty ». *Explorations in Empirical Translation Process Research*, édité par Michael CARL, Cham, Springer, pages 259–294. DOI : [10.1007/978-3-030-69777-8_10](https://doi.org/10.1007/978-3-030-69777-8_10).

ZEYER, Albert, Tamer ALKHOULI et Hermann NEY (2018). « RETURNN as a Generic Flexible Neural Toolkit with Application to Translation and Speech Recognition ». *Proceedings of ACL 2018, System Demonstrations*, édité par Fei LIU et Tamar SOLORIO, Melbourne, ACL, pages 128–133. DOI : [10.18653/v1/P18-4022](https://doi.org/10.18653/v1/P18-4022).

Pages Web et articles de presse

AKHULKOVA, Yulia, Sarah HICKEY et Belén AGULLÓ GARCÍA (2021). « Nimdzi Language Technology Atlas 2021 ». *Nimdzi*, 5 août 2021. URL : <https://www.nimdzi.com/nimdzi-language-technology-atlas-2021>.

ALBERGOTTI, Reed (2023). « The secret history of Elon Musk, Sam Altman, and OpenAI ». *Semafor*, 24 mars 2023. URL : <https://www.semafor.com/article/03/24/2023/the-secret-history-of-elon-musk-sam-altman-and-openai>.

ALRIC, Jean-Yves (2023). « ChatGPT écrit des livres et devient le pire cauchemar des auteurs ». *Presse-citron*, 22 févr. 2023. URL : <https://www.presse-citron.net/chatgpt-ecrit-des-livres-et-devient-le-pire-cauchemar-des-auteurs>.

- AMODEI, Dario, Danny HERNANDEZ, Girish SASTRY, Jack CLARK, Greg BROCKMAN et Ilya SUTSKEVER (2018). « AI and compute ». *OpenAI*, 16 mai 2018. URL : <https://openai.com/research/ai-and-compute>.
- ASLAN, Şölen (2021). « Data and AI Trends in 2022 ». *TAUS*, 1^{er} déc. 2021. URL : <https://www.taus.net/resources/blog/data-and-ai-trends-in-2022>.
- BOGAGE, Jacob (2023). « In a tight labor market, some states look to another type of worker: Children ». *The Washington Post*, 11 févr. 2023. URL : <https://www.washingtonpost.com/business/2023/02/11/child-labor-iowa>.
- CHRISTIAN, Jon (2023a). « CNET's AI Journalist Appears to Have Committed Extensive Plagiarism ». *Futurism*, 23 janv. 2023. URL : <https://futurism.com/cnet-ai-plagiarism>.
- (2023b). « Magazine Publishes Serious Errors in First AI-Generated Health Article ». *Futurism*, 18 févr. 2023. URL : <https://futurism.com/neoscope/magazine-mens-journal-errors-ai-health-article>.
- DEEPL GMBH (2021). « Comment fonctionne DeepL ? » *DeepL*, 1^{er} nov. 2021. URL : <https://www.deepl.com/fr/blog/how-does-deepl-work>.
- DIBBELL, Julian (2000). « After Babelfish ». *Julian Dibbell*, 26 juill. 2000. URL : http://www.juliandibbell.com/texts/feed_babelfish.html.
- ERTZSCHEID, Olivier (2019). « Code is Outlaw ». *affordance.info*, 16 déc. 2019. URL : <https://affordance.framasoft.org/2019/12/cest-la-lutte-digitale>.
- GANDON, Fabien (2020). « Les IA comprennent-elles ce qu'elles font ? » *The Conversation*, 20 nov. 2020. URL : <https://theconversation.com/les-ia-comprennent-elles-ce-que-elles-ont-148513>.
- IFTIKHAR, Hassan (2023). « The Role of AI in Software Engineering in 2023: Where is the Tech Industry Heading? » *Medium*, 27 mars 2023. URL : <https://medium.com/tech-lead-hub/the-role-of-ai-in-software-engineering-in-2023-where-is-the-tech-industry-heading-87185022dffa>.
- KNIGHT, Will (2023). « OpenAI's CEO Says the Age of Giant AI Models Is Already Over ». *WIRED*, 17 avr. 2023. URL : <https://www.wired.com/story/openai-ceo-sam-altman-the-age-of-giant-ai-models-is-already-over/>.
- NG, lance (2018). « What Microsoft and Google Are Not Telling You About Their A.I. » *Medium*, 13 déc. 2018. URL : <https://lancengym.medium.com/what-microsoft-and-google-are-not-telling-you-about-their-ai-b609f5395a8e>.

- OMNISCIENT (2020). « Hybrid Neural and Statistical Machine Translation ». *Omniscient Technologies*, 29 sept. 2020. URL : <https://omniscient.com/machine-translation/hybrid-machine-translation>.
- PERRIGO, Billy (2023). « OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic ». *TIME*, 18 janv. 2023. URL : <https://time.com/6247678/openai-chatgpt-kenya-workers>.
- RAO, Delip (2020a). « GPT-3 and A Typology of Hype ». *Page Street Labs*, 26 juill. 2020. URL : <https://pagestlabs.substack.com/p/gpt-3-and-a-typology-of-hype>.
- (2020b). « GPT-3 Turk’s Gambit and The Question of AI Automation ». *Page Street Labs*, 21 août. 2020. URL : <https://pagestlabs.substack.com/p/gpt-3-turks-gambit-and-the-question>.
- SARNO, David (2010). « Franz Josef Och, Google’s translation uber-scientist, talks about Google Translate ». *Los Angeles Time*, 11 mars 2010. URL : <https://www.latimes.com/archives/blogs/technology-blog/story/2010-03-11/franz-josef-och-googles-translation-uber-scientist-talks-about-google-translate>.
- SATO, Mia (2023). « AI-generated fiction is flooding literary magazines — but not fooling anyone ». *The Verge*, 25 févr. 2023. URL : <https://www.theverge.com/2023/2/25/23613752/ai-generated-short-stories-literary-magazines-clarkesworld-science-fiction>.
- STATISTA (2022). « Livres ayant été traduits dans le plus grand nombre de langues entre leur publication et 2022 ». *Statista*, 4 nov. 2022. URL : <https://fr.statista.com/statistiques/1342091/livres-les-plus-traduits-du-monde>.
- TAPPER, James (2023). « Authors shocked to find AI ripoffs of their books being sold on Amazon ». *The Guardian*, 30 sept. 2023. URL : <https://www.theguardian.com/technology/2023/sep/30/authors-shocked-to-find-ai-ripoffs-of-their-books-being-sold-on-amazon>.
- TAUS (2022). « Machine Translation Redefined with DeMT™: No-Human-in-the-Loop ». *Medium*, 12 avr. 2022. URL : <https://medium.com/tausdata/machine-translation-redefined-with-demt-no-human-in-the-loop-3d4b7af0b3e3>.
- TRYSTRAM, Denis, Romain COUILLET et Thierry MÉNISSIER (2021). « Apprentissage profond et consommation énergétique : la partie immergée de l’IA-ceberg ». *The Conversation*, 8 déc. 2021. URL : <https://theconversation.com/apprentissage-profond-et-consommation-energetique-la-partie-immergee-de-lia-ceberg-172341>.

- VINCENT, James (2022). « The lawsuit that could rewrite the rules of AI copyright ». *The Verge*, 8 nov. 2022. URL : <https://www.theverge.com/2022/11/8/23446821/microsoft-openai-github-copilot-class-action-lawsuit-ai-copyright-violation-training-data>.
- WILLIAMS, Adrienne, Milagros MICELI et Timnit GEBRU (2022). « The Exploited Labor Behind Artificial Intelligence ». *Noema*, 13 oct. 2022. URL : <https://www.noemamag.com/the-exploited-labor-behind-artificial-intelligence>.

Références

- AGRE, Philip E. (1997). « Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI ». *Social Science, Technical Systems, and Cooperative Work: Beyond the Great Divide*, édité par Geoffrey C. BOWKER, Susan Leigh STAR, William TURNER et Les GASSER, Mahwah (NJ), Lawrence Erlbaum Associates, pages 131–157. DOI : [10.4324/9781315805849](https://doi.org/10.4324/9781315805849).
- AHRENBERG, Lars (2017). « Comparing Machine Translation and Human Translation: A Case Study ». *Proceedings of the First Workshop on Human-Informed Translation and Interpreting Technology*, édité par Irina TEMNIKOVA, Constantin ORASAN, Gloria CORPAS PASTOR et Stephan VOGEL, Varna, ACL, pages 21–28. DOI : [10.26615/978-954-452-042-7_003](https://doi.org/10.26615/978-954-452-042-7_003).
- ALMAHASEES, Zakaryia Mustafa (2017). « Machine Translation Quality of Khalil Gibran's *The Prophet* ». *AWEJ for translation & Literary Studies*, n° 4, pages 151–159. DOI : [10.2139/ssrn.3068518](https://doi.org/10.2139/ssrn.3068518).
- ALPAC (1966). *Language and Machines: Computers in Translation and Linguistics*. Washington (DC), National Academy of Sciences. DOI : [10.17226/9547](https://doi.org/10.17226/9547).
- ARAABI, Ali et Christof MONZ (2020). « Optimizing Transformer for Low-Resource Neural Machine Translation ». *Proceedings of the 28th International Conference on Computational Linguistics*, édité par Donia SCOTT, Nuria BEL et Chengqing ZONG, Barcelone, ICCL, pages 3429–3435. DOI : [10.18653/v1/2020.coling-main.304](https://doi.org/10.18653/v1/2020.coling-main.304).
- ATLAS et ATLF (2022). « Métiers de la traduction : ce qu'automatiser veut dire ». Table ronde des 39^{es} Assises de la traduction littéraire, 13 nov. 2022, Arles. URL : <https://www.youtube.com/watch?v=ONDjZi6kD1w>.
- (2023). « IA et traduction littéraire : les traductrices et traducteurs exigent la transparence ». Tribune publiée sur le site de l'ATLAS, mars 2023. URL : <https://www.atlas-citl.org/wp-content/uploads/2023/03/Tribune-ATLAS-ATLF-3.pdf>.

- ATLF (2022). « Traduction automatique et post-édition ». Enquête menée entre le 20 nov. et le 13 déc. 2022. URL : <https://atlf.org/wp-content/uploads/2023/03/ENQUETE-TRADUCTION-AUTOMATIQUE.pdf>.
- BAHDANAU, Dzmitry, Kyunghyun CHO et Yoshua BENGIO (2015). « Neural Machine Translation by Jointly Learning to Align and Translate ». *3rd International Conference on Learning Representations: Conference Track Proceedings*, édité par Yoshua BENGIO et Yann LECUN, San Diego (CA), ICLR, pages 1–15. ARXIV : [1409.0473](https://arxiv.org/abs/1409.0473).
- BAKER, Mona (1996). « Corpus-Based Translation Studies: The Challenges That Lie Ahead ». *Terminology, LSP and Translation: Studies in Language Engineering*, édité par Harold SOMERS, Amsterdam, John Benjamins, pages 175–186. DOI : [10.1075/bt1.18.17bak](https://doi.org/10.1075/bt1.18.17bak).
- (2000). « Towards a Methodology for Investigating the Style of a Literary Translator ». *Target*, vol. 12, n° 2, pages 241–266. DOI : [10.1075/target.12.2.04bak](https://doi.org/10.1075/target.12.2.04bak).
- BARNABÉ, Fanny (2012). « Les polémiques autour de la littérature jeunesse, ou la quête sans cesse rejouée de la légitimité ». *COntEXTES*, n° 10. DOI : [10.4000/contextes.5020](https://doi.org/10.4000/contextes.5020).
- BARREDO ARRIETA, Alejandro, Natalia DÍAZ-RODRÍGUEZ, Javier DEL SER, Adrien BENNETOT, Siham TABIK, Alberto BARBADO, Salvador GARCIA, Sergio GIL-LOPEZ, Daniel MOLINA, Richard BENJAMINS, Raja CHATILA et Francisco HERRERA (2020). « Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI ». *Information Fusion*, n° 10, pages 82–115. DOI : [10.1016/j.inffus.2019.12.012](https://doi.org/10.1016/j.inffus.2019.12.012).
- BARRY, John (1991). *Technobabble*. Cambridge (MA), The MIT Press.
- BASSNETT, Susan (2013). *Translation Studies*. 4^e édition, Abingdon, Routledge. DOI : [10.4324/9780203488232](https://doi.org/10.4324/9780203488232).
- BASSOLE, Jean-Yves (2022). « Voyage au bout de la traduction automatique ». *Journal of Data Mining & Digital Humanities, Vers une robotique du traduire ?*, édité par Anne BAILLOT, Ellen CARTER, Thierry GRASS et Pablo FABO, pages 1–9. DOI : [10.46298/jdmdh.9297](https://doi.org/10.46298/jdmdh.9297).
- BAUDOT, Jean (1964). *La Machine à écrire : Mise en marche et programmée par Jean A. Baudot*. Montréal, Les Éditions du Jour.
- BAWDEN, Rachel, Rico SENNRICH, Alexandra BIRCH et Barry HADDOW (2018). « Evaluating Discourse Phenomena in Neural Machine Translation ». *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, édité par Marilyn WALKER, Heng Ji et Amanda STENT, La Nouvelle-Orléans (LA), ACL, pages 1304–1313. DOI : [10.18653/v1/N18-1118](https://doi.org/10.18653/v1/N18-1118).

- BÉDARD, Claude (2017). « Mémoires de traduction : quel destin pour Charlie ? » *Traduire*, n° 237, pages 8–14. DOI : [10.4000/traduire.935](https://doi.org/10.4000/traduire.935).
- BENANAV, Aaron (2020). *Automation and the Future of Work*. Londres, Verso.
- BENDER, Emily M., Timnit GEBRU, Angelina MCMILLAN-MAJOR et Shmargaret SHMITCHELL (2021). « On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? » *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, New York (NY, ACM, pages 610–623. DOI : [10.1145/3442188.3445922](https://doi.org/10.1145/3442188.3445922).
- BENDER, Emily M. et Alexander KOLLER (2020). « Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data ». *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, édité par Dan JURAFSKY, Joyce CHAI, Natalie SCHLUTER et Joel TETREAULT, ACL, pages 5185–5198. DOI : [10.18653/v1/2020.acl-main.463](https://doi.org/10.18653/v1/2020.acl-main.463).
- BENJAMIN, Walter (2012). « The Translator’s Task ». *The Translation Studies Reader*, édité par Lawrence VENUTI. Traduit par Steven RENDALL. 3^e édition, Abingdon, Routledge, pages 75–83.
- BENTIVOGLI, Luisa, Arianna BISAZZA, Mauro CETTOLO et Marcello FEDERICO (2016). « Neural versus Phrase-Based Machine Translation Quality: A Case Study ». *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, édité par Jian SU, Kevin DUH et Xavier CARRERAS, Austin (TX), ACL, pages 257–267. DOI : [10.18653/v1/D16-1025](https://doi.org/10.18653/v1/D16-1025).
- BESACIER, Laurent (2014). « Traduction automatisée d’une oeuvre littéraire : une étude pilote ». *Actes de la 21e conférence sur le Traitement Automatique des Langues Naturelles*, édité par Philippe BLACHE, Frédéric BÉCHET et Brigitte BIGI, Marseille, ATALA, pages 389–394. URL : <https://aclanthology.org/F14-2001>.
- BESACIER, Laurent et Lane SCHWARTZ (2015). « Automated translation of a literary work: a pilot study ». *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, édité par Anna FELDMAN, Anna KAZANTSEVA, Stan SZPAKOWICZ et Corina KOOLEN, Denver (CO), ACL, pages 114–122. DOI : [10.3115/v1/W15-0713](https://doi.org/10.3115/v1/W15-0713).
- BIAU GIL, José Ramón et Anthony PYM (2006). « Technology and translation (a pedagogical overview) ». *Translation Technology and its Teaching (with much mention of localization)*, édité par Anthony PYM, Alexander PEREKRESTENKO et Bram STARINK, Tarragone, ISG, pages 5–19.
- BLANCHON, Hervé et Christian BOITET (2007). « Pour l’évaluation externe des systèmes de TA par des méthodes fondées sur la tâche ». *Traitement Automatique des Langues*, vol. 48, n° 1 : *Principes de l’évaluation en Traitement Automatique des Langues*, édité par Patrick PAROUBEK, Stéphane CHAUDIRON et Lynette HIRSCHMAN, pages 33–65. URL : <https://aclanthology.org/2007.tal-1.2>.

- BLIN, Raoul (2021). « Neural machine translation, corpus and frugality ». Prépublication *arXiv*. ARXIV : 2101.10650.
- BLOOMFIELD, Camille et Hélène CAMPAIGNOLLE-CATEL (2016). « Machines littéraires, machines numériques : l'Oulipo et l'informatique ». *Oulipo mode d'emploi*, édité par Christelle REGGIANI et Alain SCHAFFNER, Paris, Honoré Champion, pages 319–336.
- BLUMCZYNSKI, Piotr (2016). *Ubiquitous Translation*. New York (NY), Routledge.
- BOASE-BEIER, Jean (2014). « Stylistics and translation ». *The Routledge Handbook of Stylistics*, édité par Michael BURKE, Abingdon, Routledge, pages 393–407. DOI : 10.4324/9781315795331.
- BOCQUET, Claude (2007). « Traduire les textes nobles, traduire les textes ignobles : une seule ou deux méthodes ? » *La traductologie dans tous ses états*, édité par Corinne ECKSTEEN et Ahmed EL KALADI, Arras, Artois Presses Université, pages 9–26. URL : <http://books.openedition.org/apu/5628>.
- BOJAR, Ondřej, Matouš MACHÁČEK, Aleš TAMCHYNA et Daniel ZEMAN (2013). « Scratching the Surface of Possible Translations ». *Text, Speech, and Dialogue*, édité par Ivan HABERNAL et Václav MATOUŠEK, Pilsen, Springer, pages 465–474. DOI : 10.1007/978-3-642-40585-3_59.
- BORG, Claudine (2023). *A Literary Translation in the Making: A Process-Oriented Perspective*. New York (NY), Routledge. DOI : 10.4324/9781003150909.
- BOWKER, Lynne (2021). « Translation technology and ethics ». *The Routledge Handbook of Translation and Ethics*, édité par Kaisa KOSKINEN et Nike K. POKORN, Abingdon, Routledge, pages 262–278. DOI : 10.4324/9781003127970.
- BOWKER, Lynne et Jairo BUITRAGO CIRO (2019). *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley, Emerald Publishing. DOI : 10.1108/9781787567214.
- BRAFFORT, Paul (1968). *L'Intelligence artificielle*. Paris, Presses universitaires de France.
- (1984). *LAPAL: Un outil pour la manipulation de textes littéraires*. ALAMO. URL : www.alamo.free.fr/archives/LAPAL/LAPAL-PB.pdf.
- (1999). *Science et littérature : les deux cultures, à l'aube du troisième millénaire*. Paris, Diderot multimédia.
- (2000). « L'ALAMO en avant "post-" ». *Études françaises*, vol. 36, n° 2 : *Internet et littérature : nouveaux espaces d'écriture*, pages 99–118. DOI : 10.7202/005254ar.
- BROWN, Peter, John COCKE, Stephen DELLA PIETRA, Vincent DELLA PIETRA, Frederick JELINEK, Robert MERCER et Paul ROOSSIN (1988). « A Statistical Approach to Language Translation ».

- Proceedings of the 12th International Conference on Computational Linguistics*, vol. 1, édité par Dénes VARGHA, Budapest, ACL, pages 71–76. URL : <https://aclanthology.org/C88-1016>.
- BROWN, Tom, Benjamin MANN, Nick RYDER, Melanie SUBBIAH, Jared KAPLAN, Prafulla DHARIWAL, Arvind NEELAKANTAN, Pranav SHYAM, Girish SASTRY, Amanda ASKELL, Sandhini AGARWAL, Ariel HERBERT-VOSS, Gretchen KRUEGER, Tom HENIGHAN, Rewon CHILD, Aditya RAMESH, Daniel ZIEGLER, Jeffrey WU, Clemens WINTER, Chris HESSE, Mark CHEN, Eric SIGLER, Mateusz LITWIN, Scott GRAY, Benjamin CHESS, Jack CLARK, Christopher BERNER, Sam MCCANDLISH, Alec RADFORD, Ilya SUTSKEVER et Dario AMODEI (2020). « Language Models are Few-Shot Learners ». *Advances in Neural Information Processing Systems 33*, édité par Hugo LAROCHELLE, Marc'Aurelio RANZATO, Raia HADSELL, Maria-Florina BALCAN et Hsuan-Tien LIN. URL : <https://papers.nips.cc/paper/2020/hash/1457c0d6bfc4967418bfb8ac142f64a-Abstract.html>.
- BRUNETTE, Louise et Sharon O'BRIEN (2011). « Quelle ergonomie pour la pratique postéditrice des textes traduits ? » *ILCEA*, n° 14 : *Traduction et Ergonomie*. DOI : [10.4000/ilcea.1081](https://doi.org/10.4000/ilcea.1081).
- BRUSASCO, Paola (2022). « Pragmatic and cognitive elements in literary machine translation: An assessment of an excerpt from J. Polzin's *Brood* translated with Google, DeepL, and Microsoft Bing ». *Using Technologies for Creative-Text Translation*, édité par James HADLEY, Kristiina TAIVALKOSKI-SHILOV, Carlos S. C. TEIXEIRA et Antonio TORAL, New York (NY), Routledge, pages 139–160. DOI : [10.4324/9781003094159-7](https://doi.org/10.4324/9781003094159-7).
- BURROWS, John (2002). « 'Delta': A Measure of Stylistic Difference and a Guide to Likely Authorship ». *Literary and Linguistic Computing*, vol. 17, n° 3, pages 267–287. DOI : [10.1093/llc/17.3.267](https://doi.org/10.1093/llc/17.3.267).
- BVPA, DFV MEDIENGRUPPE et VS (2023). « Métiers de la traduction : ce qu'automatiser veut dire ». Table ronde de la *Foire du livre de Francfort*, 18 oct. 2023, Francfort-sur-le-Main. URL : <https://www.youtube.com/live/o33fbnJ9ZTo>.
- CADWELL, Patrick, Sharon O'BRIEN et Carlos S. C. TEIXEIRA (2018). « Resistance and accommodation: factors for the (non-) adoption of machine translation among professional translators ». *Perspectives*, vol. 26, n° 3, pages 301–321. DOI : [10.1080/0907676X.2017.1337210](https://doi.org/10.1080/0907676X.2017.1337210).
- CALLISON-BURCH, Chris, Cameron FORDYCE, Philipp KOEHN, Christof MONZ et Josh SCHROEDER (2007). « (Meta-)Evaluation of Machine Translation ». *Proceedings of the Second Workshop on Statistical Machine Translation*, édité par Chris CALLISON-BURCH, Philipp KOEHN, Cameron Shaw FORDYCE et Christof MONZ, Prague, ACL, pages 136–158. URL : <https://aclanthology.org/W07-0718>.

- CALLISON-BURCH, Chris, Miles OSBORNE et Philipp KOEHN (2006). « Re-Evaluating the Role of BLEU in Machine Translation Research ». *11th Conference of the European Chapter of the Association for Computational Linguistics*, édité par Diana MCCARTHY et Shuly WINTNER, Trente, ACL, pages 249–256. URL : <https://aclanthology.org/E06-1032>.
- CALVINO, Italo (1984). « Cybernétique et fantasmes ». *La Machine littérature : Essais*, traduit par François WAHL et Michel ORCEL, Paris, Seuil, pages 11–29.
- CAMBRELENG, Jörn (2020). « L’observatoire de la traduction automatique ». Présentation à la journée d’études *La traduction littéraire et SHS à la rencontre des technologies de la traduction: enjeux, pratiques et perspectives*, 7 févr. 2020, Toulouse.
- CARDON, Dominique, Jean-Philippe COINTET et Antoine MAZIÈRES (2018). « La revanche des neurones : L’invention des machines inductives et la controverse de l’intelligence artificielle ». *Réseaux*, vol. 211, n° 5, pages 173–220. DOI : [10.3917/res.211.0173](https://doi.org/10.3917/res.211.0173).
- CARL, Michael et Moritz Jonas SCHAEFFER (2017). « Why Translation Is Difficult: A Corpus-Based Study of Non-Literality in Post-Editing and From-Scratch Translation ». *HERMES*, n° 56, pages 43–57. DOI : [10.7146/hjlcb.v0i56.97201](https://doi.org/10.7146/hjlcb.v0i56.97201).
- CARMO, Félix do (2020). « ‘Time is money’ and the value of translation ». *Translation Spaces*, vol. 9, n° 1, pages 35–57. DOI : [10.1075/ts.00020.car](https://doi.org/10.1075/ts.00020.car).
- (2023). « How AI extracts and exploits value from knowledge created by humans: the case of machine translation ». Présentation au séminaire *Sessões #CODA - Hacktivar as Humanidades*, 11 avr. 2023, Porto.
- CARY, Edmond (1956). « Mécanismes et traduction ». *Babel*, vol. 2, n° 3, pages 102–107. DOI : [10.1075/babel.2.3.04car](https://doi.org/10.1075/babel.2.3.04car).
- CASILLI, Antonio (2019). *En attendant les robots : Enquête sur le travail du clic*. Paris, Seuil.
- CASTILHO, Sheila, Joss MOORKENS, Federico GASPARI, Rico SENNRICH, Vilemini SOSONI, Panayota GEORGAKOPOULOU, Pintu LOHAR, Andy WAY, Antonio Valerio MICELI-BARONE et Maria GIALAMA (2017). « A Comparative Quality Evaluation of PBSMT and NMT using Professional Translators ». *Proceedings of Machine Translation Summit XVI*, vol. 1, édité par Sadao KUHASHI et Pascale FUNG, Nagoya, pages 116–131. URL : <https://aclanthology.org/2017.mtsummit-papers.10>.
- CASTILHO, Sheila et Natália RESENDE (2022). « Post-Editese in Literary Translations ». *Information*, vol. 13, n° 2 : *Machine Translation for Conquering Language Barriers*, édité par Ivan DUNĐER, pages 1–22. DOI : [10.3390/info13020066](https://doi.org/10.3390/info13020066).

- CASTRO, Olga (2020). « Transnational Feminism, Women Writers in Translation, Stateless Cultures/Literatures in Translation ». Présentation au séminaire *Creative Translation and Technologies Expert Meeting*, 29 mai 2020, en ligne.
- CASWELL, Isaac, Ciprian CHELBA et David GRANGIER (2019). « Tagged Back-Translation ». *Proceedings of the Fourth Conference on Machine Translation*, vol. 1, édité par Ondřej BOJAR, Rajen CHATTERJEE, Christian FEDERMANN, Mark FISHEL, Yvette GRAHAM, Barry HADDOW, Matthias HUCK, Antonio Jimeno YEPES, Philipp KOEHN, André MARTINS, Christof MONZ, Matteo NEGRI, Aurélie NÉVÉOL, Mariana NEVES, Matt POST, Marco TURCHI et Karin VERSPOOR, Florence, ACL, pages 53–63. DOI : [10.18653/v1/W19-5206](https://doi.org/10.18653/v1/W19-5206).
- CEATL (2020). *Counterpoint*, n° 4 : *Machine Translation*, édité par Hanneke VAN DER HEIJDEN, Anne LARCHET et Juliane WAMMEN. URL : https://ceatl.eu/wp-content/uploads/2023/04/Counterpoint_2020_04.pdf.
- (2023). *Counterpoint*, n° 10 : *AI and Literary Translation*, édité par Juliane WAMMEN, Anne LARCHET, Kaisa RANTA et Hanneke VAN DER HEIJDEN. URL : https://www.ceatl.eu/wp-content/uploads/2023/12/Counterpoint_2023_10.pdf.
- CEATL, ECSA, EFJ, EG AIR, EWC, FERA, FIA, FIM, FSE, IAO, IFJ, UNI MEI et UVA (2023). « Joint statement from authors' and performers' organisations on Artificial Intelligence and the AI Act ». Déclaration conjointe publiée sur le site de l'EWC, 26 sept. 2023. URL : <https://europeanwriterscouncil.eu/wp-content/uploads/2023/09/1414-authors-performers-and-other-creative-workers-organisations-joint.pdf>.
- CHAKRABARTY, Tuhin, Arkadiy SAAKYAN et Smaranda MURESAN (2021). « Don't Go Far Off: An Empirical Study on Neural Poetry Translation ». *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, édité par Marie-Francine MOENS, Xuanjing HUANG, Lucia SPECIA et Wen-tau YIH, Punta Cana, ACL, pages 7253–7265. DOI : [10.18653/v1/2021.emnlp-main.577](https://doi.org/10.18653/v1/2021.emnlp-main.577).
- CHAN, Sin-wai, éd. (2018). *The Human Factor in Machine Translation*. Abingdon, Royaume-Uni, Routledge. DOI : [10.4324/9781315147536](https://doi.org/10.4324/9781315147536).
- CHATY, Guy (2000). « L'ordinateur, lieu d'expressions poétiques ? » *Phrétique*, n° 96, pages 123–134. URL : <http://lapal.free.fr/alamo/articles/desel.html>.
- CHEN, Boxing, Roland KUHN et George FOSTER (2013). « Vector Space Model for Adaptation in Statistical Machine Translation ». *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, vol. 1, édité par Hinrich SCHUETZE, Pascale FUNG et Massimo POESIO, Sofia, ACL, pages 1285–1293. URL : <https://aclanthology.org/P13-1126>.

- CHESTERMAN, Andrew (1998). « Causes, Translations, Effects ». *Target*, vol. 10, n° 2, pages 201–230. DOI : [10.1075/target.10.2.02che](https://doi.org/10.1075/target.10.2.02che).
- (2001). « Proposal for a Hieronymic Oath ». *The Translator*, vol. 7, n° 2 : *The Return to Ethics*, pages 139–154. DOI : [10.1080/13556509.2001.10799097](https://doi.org/10.1080/13556509.2001.10799097).
- (2004). « Beyond the particular ». *Translation Universals: Do they exist?*, édité par Anna MAURANEN et Pekka KUJAMÄKI, Amsterdam, John Benjamins Publishing Company, pages 33–49. DOI : [10.1075/bt1.48.04che](https://doi.org/10.1075/bt1.48.04che).
- (2009). « The Name and Nature of Translator Studies ». *HERMES*, n° 42, pages 13–22. DOI : [10.7146/hj1cb.v22i42.96844](https://doi.org/10.7146/hj1cb.v22i42.96844).
- CHO, Kyunghyun, Bart van MERRIËNBOER, Caglar GULCEHRE, Dzmitry BAHDANAU, Fethi BOUGARES, Holger SCHWENK et Yoshua BENGIO (2014). « Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation ». *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, édité par Alessandro MOSCHITTI, Bo PANG et Walter DAELEMANS, Doha, ACL, pages 1724–1734. DOI : [10.3115/v1/D14-1179](https://doi.org/10.3115/v1/D14-1179).
- CHRISTENSEN, Tina Paulsen et Anne SCHJOLDAGER (2016). « Computer-aided translation tools – the uptake and use by Danish translation service providers ». *JoSTrans*, n° 25, pages 89–105. URL : https://jostrans.soap2.ch/issue25/art_christensen.php.
- CHU, Chenhui et Rui WANG (2018). « A Survey of Domain Adaptation for Neural Machine Translation ». *Proceedings of the 27th International Conference on Computational Linguistics*, édité par Emily M. BENDER, Leon DERCZYNSKI et Pierre ISABELLE, Santa Fe (NM), ACL, pages 1304–1319. URL : <https://www.aclweb.org/anthology/C18-1111>.
- CHUKWU, Uzoma (1994). « Les ordinateurs sont des humains : métaphore, terminologie, progrès et métascience ». *Cahiers de l'APLIUT*, vol. 14, n° 1, pages 79–98. DOI : [10.3406/apliu.1994.3401](https://doi.org/10.3406/apliu.1994.3401).
- COCHRANE, Guylaine (1995). « Le foisonnement, phénomène complexe ». *Traduction, terminologie, rédaction*, vol. 8, n° 2, pages 175–193. DOI : [10.7202/037222ar](https://doi.org/10.7202/037222ar).
- CONSTANTINE, Peter (2019). « Google Translate Gets Voltaire: Literary Translation and the Age of Artificial Intelligence ». *Contemporary French and Francophone Studies*, vol. 23, n° 4, pages 471–479. DOI : [10.1080/17409292.2019.1694798](https://doi.org/10.1080/17409292.2019.1694798).
- CORPAS PASTOR, Gloria, Ruslan MITKOV, Naveed AFZAL et Viktor PEKAR (2008). « Translation Universals: Do They Exist? A Corpus-Based NLP Study of Convergence and Simplification ». *Proceedings of the 8th Conference of the Association for Machine Translation in the Americas:*

- Research Papers*, Honolulu (HI), AMTA, pages 75–81. URL : <https://aclanthology.org/2008.ama-papers.5>.
- COUTELLEC, Léo (2019). « Penser l'indissociabilité de l'éthique de la recherche, de l'intégrité scientifique et de la responsabilité sociale des sciences : Clarification conceptuelle, propositions épistémologiques ». *Revue d'anthropologie des connaissances*, vol. 13, n° 2, pages 381–398. URL : <https://www.cairn.info/revue-anthropologie-des-connaissances-2019-2-page-381.htm>.
- COVINGTON, Michael A. et Joe D. MCFALL (2010). « Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR) ». *Journal of Quantitative Linguistics*, vol. 17, n° 2, pages 94–100. DOI : [10.1080/09296171003643098](https://doi.org/10.1080/09296171003643098).
- CRAWFORD, Kate (2021). *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven (CT), Yale University Press. DOI : [10.2307/j.ctv1ghv45t](https://doi.org/10.2307/j.ctv1ghv45t).
- CRESWELL, John W. et Vicki L. PLANO CLARK (2007). *Designing and Conducting Mixed Methods Research*. Thousand Oaks (CA), SAGE Publications.
- CRiado PEREZ, Caroline (2019). *Invisible Women: Data Bias in a World Designed for Men*. New York (NY), Abrams Press.
- DAEMS, Joke (2022). « Dutch literary translators' use and perceived usefulness of technology: The role of awareness and attitude ». *Using Technologies for Creative-Text Translation*, édité par James HADLEY, Kristiina TAIVALKOSKI-SHILOV, Carlos S. C. TEIXEIRA et Antonio TORAL, New York (NY), Routledge, pages 40–65. DOI : [10.4324/9781003094159-3](https://doi.org/10.4324/9781003094159-3).
- DAVIS, Brent et Dennis SUMARA (1997). « Cognition, Complexity, and Teacher Education ». *Harvard Educational Review*, vol. 67, n° 1, pages 105–126. DOI : [10.17763/haer.67.1.160w00j113t78042](https://doi.org/10.17763/haer.67.1.160w00j113t78042).
- DE CLERCQ, Christophe et Gys-Walt VAN EGDOM (2023). « No more buying cats in a bag? Literary Translation in the age of language automation ». *Tradumàtica*, n° 21 : *Computer-Aided Literary Translation*, pages 49–62. DOI : [10.5565/rev/tradumatica.407](https://doi.org/10.5565/rev/tradumatica.407).
- DE CLERCQ, Orphée, Gert DE SUTTER, Rudy LOOCK, Bart CAPPELLE et Koen PLEVOETS (2021). « Uncovering Machine Translationese Using Corpus Analysis Techniques to Distinguish between Original and Machine-Translated French ». *Translation Quarterly*, n° 101, pages 21–45.
- DEFERT, Dominique (2022). « Dan Brown, Patricia Cornwell et John Grisham à l'épreuve de DeepL ». *Journal of Data Mining & Digital Humanities, Vers une robotique du traduire ?*, édité par Anne BAILLOT, Ellen CARTER, Thierry GRASS et Pablo FABO, pages 1–19. DOI : [10.46298/jdmdh.9084](https://doi.org/10.46298/jdmdh.9084).

- (2023). « La traduction et l'IA, quel avenir ? » Podcast *De vive(s) voix*, Radio France internationale. URL : <https://www.rfi.fr/fr/podcasts/de-vive-s-voix/20230904-la-traduction-et-l-ia-quel-avenir>.
- DELISLE, Jean (2007). *La traduction raisonnée : Manuel d'initiation à la traduction professionnelle de l'anglais vers le français*. 2^e édition, Ottawa, Presses de l'Université d'Ottawa.
- DESJARDINS, Renée, Claire LARSONNEUR et Philippe LACOUR, éd. (2021). *When Translation Goes Digital: Case Studies and Critical Reflections*. Cham, Palgrave Macmillan. DOI : [10.1007/978-3-030-51761-8](https://doi.org/10.1007/978-3-030-51761-8).
- DIFFERT, Gaëlle (2023). « Un stage à l'épreuve de l'IA : retour d'expérience ». Présentation au colloque *Éthique(s) et traduction à l'ère contemporaine*, 14 févr. 2023, Avignon.
- DIMITRIU, Rodica (2004). « Omission in translation ». *Perspectives*, vol. 12, n^o 3, pages 163–175. DOI : [10.1080/0907676X.2004.9961499](https://doi.org/10.1080/0907676X.2004.9961499).
- DIRAND, Baptiste et Caroline ROSSI (2019). « Biotraducteur et traducteur automatique : L'homme outillé et la machine-homme ». *Des mots aux actes*, n^o 8, édité par Élisabeth LAVAULT-OLLÉON et Maria ZIMINA, pages 63–78. DOI : [10.15122/isbn.978-2-406-09779-2.p.0063](https://doi.org/10.15122/isbn.978-2-406-09779-2.p.0063).
- DOHERTY, Stephen (2016). « The Impact of Translation Technologies on the Process and Product of Translation ». *International Journal of Communication*, vol. 10, pages 947–969. URL : <https://ijoc.org/index.php/ijoc/article/view/3499/1573>.
- DORST, Aletta G. (2024). « Metaphor in Literary Machine Translation: Style, Creativity and Literariness ». *Computer-Assisted Literary Translation*, édité par Andrew ROTHWELL, Andy WAY et Roy YOUNDALE, Abingdon, Routledge, pages 173–186. DOI : [10.4324/9781003357391](https://doi.org/10.4324/9781003357391).
- DUBOIS, Jacques (2005). *L'institution de la littérature*. 2^e édition, Bruxelles, Éditions Labor.
- DURIEUX, Christine (2000). « Traduction littéraire et traduction technique : Même démarche ». *Revue des lettres et de traduction*, n^o 6, pages 13–25.
- EHRENSBERGER-DOW, Maureen, Andrea HUNZIKER HEEB, Gary MASSEY, Ursula MEIDERT, Silke NEUMANN et Heidrun BECKER (2016). « An International Survey of the Ergonomics of Professional Translation ». *ILCEA*, n^o 27 : *Approches ergonomiques des pratiques professionnelles et des formations des traducteurs*. DOI : [10.4000/ilcea.4004](https://doi.org/10.4000/ilcea.4004).
- EHRENSBERGER-DOW, Maureen et Gary MASSEY (2014). « Cognitive ergonomic issues in professional translation ». *The Development of Translation Competence: Theories and Methodologies from Psycholinguistics and Cognitive Science*, édité par John W. SCHWIETER et Aline FERREIRA, Newcastle upon Tyne, Cambridge Scholars Publishing, pages 58–86.

- ELOUNDOU, Tyna, Sam MANNING, Pamela MISHKIN et Daniel ROCK (2023). « GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models ». Prépublication *arXiv*, 4^e éd. ARXIV : [2303.10130v4](https://arxiv.org/abs/2303.10130v4).
- ELSON, David, Anna KAZANTSEVA, Rada MIHALCEA et Stan SZPAKOWICZ, éd. (2012). *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature*. Montréal, ACL. URL : <https://aclanthology.org/W12-2500>.
- EN CHAIR ET EN OS (2023). « Littérature, cinéma, presse, jeux vidéo: non à des traductions sans âme ». Manifeste publié sur le site *En chair et en os*, sept. 2023. URL : <https://enchairetenos.org/>.
- ENGLUND DIMITROVA, Birgitta et Maureen EHRENSBERGER-DOW (2018). « Cognitive space: Exploring the situational interface ». *Exploring the Situational Interface of Translation and Cognition*, édité par Maureen EHRENSBERGER-DOW et Birgitta ENGLUND DIMITROVA, Amsterdam, John Benjamins Publishing Company, pages 1–18. DOI : [10.1075/bct.101.01eng](https://doi.org/10.1075/bct.101.01eng).
- ESPERANÇA-RODIER, Emmanuelle et Nicolas BECKER (2018). « Comparaison de systèmes de traduction automatique, probabiliste et neuronal, par analyse d'erreurs ». *Actes de la 4^{ème} journée « Traitement Automatique des Langues et Intelligence Artificielle » - TALIA 2018*, édité par Didier SCHWAB et Pierre ZWEIGENBAUM, Nancy, AFIA, pages 1–14. URL : <http://ceur-ws.org/Vol-2295>.
- FALCONER, Rachel (2009). *The Crossover Novel: Contemporary Children's Fiction and Its Adult Readership*. New York (NY), Routledge. DOI : [10.4324/9780203892176](https://doi.org/10.4324/9780203892176).
- FEDERICO, Marcello, Mauro CETTOLO, Luisa BENTIVOGLI, Michael PAUL et Sebastian STÜKER, éd. (2012). *Proceedings of the 9th International Workshop on Spoken Language Translation: Evaluation Campaign*. Hong Kong, IWSLT. URL : <https://aclanthology.org/2012.iwslt-evaluation>.
- FERRIER, Bertrand (2009). *Tout n'est pas littérature ! La littérature à l'épreuve des romans pour la jeunesse*. Rennes, Presses Universitaires de Rennes.
- FIRTH, John Rupert (1962). *Studies in Linguistic Analysis*. 2^e édition, Oxford, Basil Blackwell.
- FONTEYNE, Margot, Arda TEZCAN et Lieve MACKEN (2020). « Literary Machine Translation under the Magnifying Glass: Assessing the Quality of an NMT-Translated Detective Novel on Document Level ». *Proceedings of the 12th Conference on Language Resources and Evaluation*, édité par Nicoletta CALZOLARI, Frédéric BÉCHET, Philippe BLACHE, Khalid CHOUKRI, Christopher CIERI, Thierry DECLERCK, Sara GOGGI, Hitoshi ISAHARA, Bente MAEGAARD, Joseph MARIANI, Hélène MAZO, Asuncion MORENO, Jan ODIJK et Stelios PIPERIDIS, Marseille, EAMT,

pages 3790–3798. URL : <http://www.lrec-conf.org/proceedings/lrec2020/pdf/2020.lrec-1.468.pdf>.

FORCADA, Mikel L. (2017). « Making sense of neural machine translation ». *Translation Spaces*, vol. 6, n° 2, pages 291–309. DOI : [10.1075/ts.6.2.06for](https://doi.org/10.1075/ts.6.2.06for).

— (2023). « Licensing and Usage Rights of Language Data in Machine Translation ». *Towards Responsible Machine Translation: Ethical and Legal Considerations in Machine Translation*, édité par Helena MONIZ et Carla PARRA ESCARTÍN, Cham, Springer Nature, pages 49–69. DOI : [10.1007/978-3-031-14689-3_4](https://doi.org/10.1007/978-3-031-14689-3_4).

FREITAG, Markus, George FOSTER, David GRANGIER, Viresh RATNAKAR, Qijun TAN et Wolfgang MACHEREY (2021). « Experts, Errors, and Context: A Large-Scale Study of Human Evaluation for Machine Translation ». *Transactions of the Association for Computational Linguistics*, vol. 9, édité par Brian ROARK et Ani NENKOVA, pages 1460–1474. DOI : [10.1162/tacl_a_00437](https://doi.org/10.1162/tacl_a_00437).

FREITAG, Markus, David GRANGIER et Isaac CASWELL (2020). « BLEU Might Be Guilty but References Are Not Innocent ». *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, édité par Bonnie WEBBER, Trevor COHN, Yulan HE et Yang LIU, ACL, pages 61–71. DOI : [10.18653/v1/2020.emnlp-main.5](https://doi.org/10.18653/v1/2020.emnlp-main.5).

FREITAG, Markus, Ricardo REI, Nitika MATHUR, Chi-kiu LO, Craig STEWART, Eleftherios AVRAMIDIS, Tom KOCMI, George FOSTER, Alon LAVIE et André MARTINS (2022). « Results of WMT22 Metrics Shared Task: Stop Using BLEU – Neural Metrics Are Better and More Robust ». *Proceedings of the Seventh Conference on Machine Translation*, édité par Philipp KOEHN, Loïc BARRAULT, Ondřej BOJAR, Fethi BOUGARES, Rajen CHATTERJEE, Marta R. COSTA-JUSSÀ, Christian FEDERMANN, Mark FISHEL, Alexander FRASER, Markus FREITAG, Yvette GRAHAM, Roman GRUNDKIEWICZ, Paco GUZMAN, Barry HADDOW, Matthias HUCK, Antonio JIMENO YEPES, Tom KOCMI, André MARTINS, Makoto MORISHITA, Christof MONZ, Masaaki NAGATA, Toshiaki NAKAZAWA, Matteo NEGRI, Aurélie NÉVÉOL, Mariana NEVES, Martin POPEL, Marco TURCHI et Marcos ZAMPIERI, Abou Dabi, ACL, pages 46–68. URL : <https://aclanthology.org/2022.wmt-1.2>.

FREY, Carl Benedikt et Michael A. OSBORNE (2013). « The future of employment: How susceptible are jobs to computerisation? » *Technological Forecasting and Social Change*, vol. 114, pages 254–280. DOI : [10.1016/j.techfore.2016.08.019](https://doi.org/10.1016/j.techfore.2016.08.019).

FRIEDMAN, Batya, éd. (1997). *Human Values and the Design of Computer Technology*. Stanford (CA), Center for the Study of Language et Information.

- FRISCHMANN, Brett et Evan SELINGER (2018). *Re-Engineering Humanity*. Cambridge, Cambridge University Press. DOI : [10.1017/9781316544846](https://doi.org/10.1017/9781316544846).
- FROELIGER, Nicolas (2013). *Les Noces de l'analogique et du numérique : De la traduction pragmatique*. Paris, Les Belles Lettres.
- (2022). « Le mépris (envers la traduction automatique) est-il une attitude traductologique ? » *Journal of Data Mining & Digital Humanities, Vers une robotique du traduire ?*, édité par Anne BAILLOT, Ellen CARTER, Thierry GRASS et Pablo FABO, pages 1–12. HAL : [hal-03507329](https://hal.archives-ouvertes.fr/hal-03507329).
- GALEF, David (1995). « Crossing Over: Authors Who Write Both Children's and Adults' Fiction ». *Children's Literature Association Quarterly*, vol. 20, n° 1, pages 29–35. DOI : [10.1353/chq.0.0905](https://doi.org/10.1353/chq.0.0905).
- GENZEL, Dmitriy, Jakob USZKOREIT et Franz Josef OCH (2010). « “Poetic” Statistical Machine Translation: Rhyme and Meter ». *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, édité par Hang LI et Lluís MÀRQUEZ, Cambridge (MA), ACL, pages 158–166. URL : <https://aclanthology.org/D10-1016>.
- GHAZVININEJAD, Marjan, Yejin CHOI et Kevin KNIGHT (2018). « Neural Poetry Translation ». *Proceedings of NAACL-HLT 2018*, vol. 2, édité par Marilyn WALKER, Heng JI et Amanda STENT, La Nouvelle-Orléans (LA), ACL, pages 67–71. DOI : [10.18653/v1/N18-2011](https://doi.org/10.18653/v1/N18-2011).
- GOERTZ, Gary et James MAHONEY (2012). *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Princeton (NJ), Princeton University Press.
- GONZÁLEZ CASCALLANA, Belén (2006). « Translating Cultural Intertextuality in Children's Literature ». *Children's Literature in Translation: Challenges and Strategies*, édité par Jan VAN COILLIE et Walter P. VERSCHUEREN, Abingdon, Routledge, pages 97–110. DOI : [10.4324/9781315759845](https://doi.org/10.4324/9781315759845).
- GOODFELLOW, Ian, Yoshua BENGIO et Aaron COURVILLE (2016). *Deep Learning*. Cambridge (MA), MIT Press.
- GRAHAM, Yvette, Timothy BALDWIN, Alistair MOFFAT et Justin ZOBEL (2013). « Continuous Measurement Scales in Human Evaluation of Machine Translation ». *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, édité par Antonio PAREJA-LORA, Maria LIAKATA et Stefanie DIPPER, Sofia, ACL, pages 33–41. URL : <https://aclanthology.org/W13-2305>.
- GREENE, Erica, Tugba BODRUMLU et Kevin KNIGHT (2010). « Automatic Analysis of Rhythmic Poetry with Applications to Generation and Translation ». *Proceedings of the 2010 Conference*

- on *Empirical Methods in Natural Language Processing*, édité par Hang LI et Lluís MÀRQUEZ, Cambridge (MA), ACL, pages 524–533. URL : <https://aclanthology.org/D10-1051>.
- GRUNENWALD, Noémie (2021). *Sur les bouts de la langue : Traduire en féministe/s*. Paris, La Contre Allée.
- GUERBEROF-ARENAS, Ana et Joss MOORKENS (2019). « Machine translation and post-editing training as part of a master’s programme ». *JoSTrans*, n° 31 : *Post-Editing in Practice: Process, Product and Networks*, édité par Lucas Nunes VIEIRA, Elisa ALONSO et Lindsay BYWOOD, pages 217–238. URL : https://jostrans.soap2.ch/issue31/art_guerberof.php.
- (2023). « Ethics and Machine Translation: The End User Perspective ». *Towards Responsible Machine Translation: Ethical and Legal Considerations in Machine Translation*, édité par Helena MONIZ et Carla PARRA ESCARTÍN, Cham, Springer Nature, pages 113–133. DOI : [10.1007/978-3-031-14689-3_7](https://doi.org/10.1007/978-3-031-14689-3_7).
- GUERBEROF-ARENAS, Ana et Antonio TORAL (2020). « The Impact of Post-Editing and Machine Translation on Creativity and Reading Experience ». *Translation Spaces*, vol. 9, n° 2, pages 255–282. DOI : [10.1075/ts.20035.gue](https://doi.org/10.1075/ts.20035.gue).
- (2022). « Creativity in Translation: Machine Translation as a Constraint for Literary Texts ». *Translation Spaces*, vol. 11, n° 2, pages 184–212. DOI : [10.1075/ts.21025.gue](https://doi.org/10.1075/ts.21025.gue).
- HABERMAS, Jürgen (1973). *La technique et la science comme « idéologie »*. Traduit par Jean-René LADMIRAL. Paris, Gallimard.
- HADLEY, James, Kristiina TAIVALKOSKI-SHILOV, Carlos S. C. TEIXEIRA et Antonio TORAL, éd. (2022). *Using Technologies for Creative-Text Translation*. New York (NY), Routledge. DOI : [10.4324/9781003094159](https://doi.org/10.4324/9781003094159).
- HADLEY, James Luke (2024). « MT and CAT: Challenges, Irrelevancies, or Opportunities for Literary Translation? » *Computer-Assisted Literary Translation*, édité par Andrew ROTHWELL, Andy WAY et Roy YOUDALE, Abingdon, Routledge, pages 91–105. DOI : [10.4324/9781003357391](https://doi.org/10.4324/9781003357391).
- HADLEY, James Luke, Maja POPOVIĆ, Haithem AFLI et Andy WAY, éd. (2019). *Proceedings of the Qualities of Literary Machine Translation*. Dublin, EAMT. URL : <https://aclanthology.org/W19-7300>.
- HALVERSON, Sandra Louise (2014). « Reorienting Translation Studies: Cognitive Approaches and the Centrality of the Translator ». *Translation: A Multidisciplinary Approach*, édité par Juliane HOUSE, Londres, Palgrave Macmillan, pages 116–139. DOI : [10.1057/9781137025487_7](https://doi.org/10.1057/9781137025487_7).
- HANSEN, Damien (2017). « Utilisation des logiciels d’aide à la traduction dans le domaine de la traduction littéraire : évaluation et pertinence ». Mémoire de master, Université de Liège.

- HANSEN, Damien (2021). « Les lettres et la machine : Un état de l'art en traduction littéraire automatique ». *Actes de la 28e Conférence sur le Traitement Automatique des Langues Naturelles*, vol. 2, édité par Pascal DENIS, Natalia GRABAR, Amel FRAISSE, Rémi CARDON, Bernard JACQUEMIN, Eric KERGOSIEN et Antonio BALVET, Lille, ATALA, pages 28–45. URL : <https://aclanthology.org/2021.jeptalnrecital-recital.3>.
- (2024). « The Figure of the Literary Translator Amidst New Technologies ». *Translation, Interpreting and Technological Change: Innovations in Research, Practice and Training*, édité par Marion WINTERS, Sharon DEANE-COX et Ursula BÖSER, Londres, Bloomsbury.
- HANSEN, Damien et Emmanuelle ESPERANÇA-RODIER (2023). « Human-Adapted MT for Literary Texts: Reality or Fantasy? » *Proceedings of the New Trends in Translation and Technology Conference – NeTTT 2022*, édité par Sheila CASTILHO, Rocío CARO QUINTANA, Mario STASIMIOTI et Vilemini SOSONI, Rhodes, Incoma Ltd, pages 178–190. URL : <http://acl-anthology.online/nettt-2022/>.
- HANSEN, Damien, Emmanuelle ESPERANÇA-RODIER, Hervé BLANCHON et Valérie BADA (2022). « La traduction littéraire automatique : Adapter la machine à la traduction humaine individualisée ». *Journal of Data Mining & Digital Humanities, Vers une robotique du traduire ?*, édité par Anne BAILLOT, Ellen CARTER, Thierry GRASS et Pablo FABO, pages 1–19. DOI : [10.46298/jdmdh.9114](https://doi.org/10.46298/jdmdh.9114).
- HANSEN, Damien et Pierre-Yves HOULMONT (2022). « A Snapshot into the Possibility of Video Game Machine Translation ». *Proceedings of the 15th Biennial Conference of the Association for Machine Translation in the Americas*, vol. 2, édité par Janice CAMPBELL, Stephen LARocca, Jay MARCIANO, Konstantin SAVENKOV et Alex YANISHEVSKY, Orlando (FL), AMTA, pages 257–269. URL : <https://aclanthology.org/2022.amta-upg.18>.
- HANSEN, Damien et Susan PICKFORD (2022). « Traduire la littérature populaire avec l'aide de la machine : Quels avantages et préoccupations pour les professionnel·le·s ? » Présentation au séminaire TRACT 2021-2022 - *L'impact de l'intelligence artificielle sur la traduction littéraire : théorie, pratique, création*, 14 avr. 2022, Paris.
- HARRISON, Steve, Phoebe SENGERs et Deborah TATAR (2011). « Making epistemological trouble: Third-paradigm HCI as successor science ». *Interacting with Computers*, vol. 23, n° 5, pages 385–392. DOI : [10.1016/j.intcom.2011.03.005](https://doi.org/10.1016/j.intcom.2011.03.005).
- HATIM, Basil et Ian MASON (1997). *The Translator as Communicator*. London, Routledge.

- HEARNE, Mary et Andy WAY (2011). « Statistical Machine Translation: A Guide for Linguists and Translators ». *Language and Linguistics Compass*, vol. 5, n° 5, pages 205–226. DOI : [10.1111/j.1749-818X.2011.00274.x](https://doi.org/10.1111/j.1749-818X.2011.00274.x).
- HEBB, Donald Olding (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York (NY), John Wiley & Sons. URL : https://pure.mpg.de/rest/items/item_2346268_3/component/file_2346267/content.
- HENDERSON, Peter, Jieru HU, Joshua ROMOFF, Emma BRUNSKILL, Dan JURAFSKY et Joelle PINEAU (2020). « Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning ». *Journal of Machine Learning Research*, vol. 21, n° 1, pages 1–43. URL : <https://jmlr.org/papers/v21/20-312.html>.
- HENDERSON, Peter, Xuechen LI, Dan JURAFSKY, Tatsunori HASHIMOTO, Mark A. LEMLEY et Percy LIANG (2023). « Foundation Models and Fair Use ». Prépublication *arXiv*. ARXIV : [2303.15715](https://arxiv.org/abs/2303.15715).
- HOVY, Dirk et Shrimai PRABHUMOYE (2021). « Five sources of bias in natural language processing ». *Language and Linguistics Compass*, vol. 15, n° 8, pages 1–19. DOI : [10.1111/lnc3.12432](https://doi.org/10.1111/lnc3.12432).
- HUANG, Po-Sen, Chong WANG, Sitao HUANG, Dengyong ZHOU et Li DENG (2018). « Towards Neural Phrase-based Machine Translation ». *Proceedings of the 6th International Conference on Learning Representations*, Vancouver, ICLR, pages 1–14. URL : <https://openreview.net/forum?id=HktJec1RZ>.
- HUMBLÉ, Philippe (2019). « Machine translation and poetry: The case of English and Portuguese ». *Ilha do Desterro*, vol. 72, n° 2, pages 41–52. DOI : [10.5007/2175-8026.2019v72n2p41](https://doi.org/10.5007/2175-8026.2019v72n2p41).
- HUTCHINS, W. John, éd. (2000). *Early Years in Machine Translation: Memoirs and biographies of pioneers*. Amsterdam, John Benjamins Publishing Company. DOI : [10.1075/sihols.97](https://doi.org/10.1075/sihols.97).
- (2001). « Machine Translation over fifty years ». *Histoire Épistémologie Langage*, vol. 23, n° 1 : *Le traitement automatique des langues*, pages 7–31. DOI : [10.3406/hel.2001.2815](https://doi.org/10.3406/hel.2001.2815).
- (2003). « ALPAC: the (in)famous report ». *MT News International*, n° 14, pages 9–12. URL : <https://aclanthology.org/www.mt-archive.info/90/MTNI-1996-Hutchins.pdf>.
- (2004). « The Georgetown-IBM experiment demonstrated in January 1954 ». *Proceedings of the 6th Conference of the Association for Machine Translation in the Americas: Technical Papers*, édité par Robert FREDERKING et Kathryn TAYLOR, Washington (DC), Springer, pages 102–114.
- HUTCHINS, W. John et Harold L. SOMERS (1992). *An introduction to machine translation*. Londres, Academic Press.

- HUWS, Ursula (2004). *Labor in the Global Digital Economy: The Cybertariat Comes of Age*. New York (NY), Monthly Review Press.
- JAKOBSON, Roman (1980). « A Postscript to the Discussion on Grammar of Poetry ». *Diacritics*, vol. 10, n° 1, pages 21–35. DOI : [10.2307/465038](https://doi.org/10.2307/465038).
- JIANG, Yue et Jiang NIU (2022). « How are neural machine-translated Chinese-to-English short stories constructed and cohered? An exploratory study based on theme-rheme structure ». *Lingua*, vol. 273, pages 1–20. DOI : [10.1016/j.lingua.2022.103318](https://doi.org/10.1016/j.lingua.2022.103318).
- JIAO, Wenxiang, Wenxuan WANG, Jen-tse HUANG, Xing WANG et Zhaopeng TU (2023). « Is ChatGPT a Good Translator? Yes With GPT-4 as the Engine ». Prépublication *arXiv*, 3^e éd. ARXIV : [2301.08745v3](https://arxiv.org/abs/2301.08745v3).
- JIMENEZ-CRESPO, Miguel A. (2023). « “Translationese” (and “post-editese”?) no more: On importing fuzzy conceptual tools from Translation Studies in MT research ». *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, édité par Mary NURMINEN, Judith BRENNER, Maarit KOPONEN, Sirkku LATOMAA, Mikhail MIKHAILOV, Frederike SCHIERL, Tharindu RANASINGHE, Eva VANMASSENHOVE, Sergi Alvarez VIDAL, Nora ARANBERRI, Mara NUNZIATINI, Carla Parra ESCARTÍN, Mikel FORCADA, Maja POPOVIC, Carolina SCARTON et Helena MONIZ, Tampere, EAMT, pages 261–268. URL : <https://aclanthology.org/2023.eamt-1.25>.
- JOHNSON, Wendell (1944). « Studies in Language Behavior: A Program of Research ». *Psychological Monographs*, vol. 56, n° 2, pages 1–15. URL : https://www.mpi.nl/publications/item_2350946.
- JONES, Ruth et Ann IRVINE (2013). « The (Un)faithful Machine Translator ». *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, édité par Pirooska LENDVAI et Kalliopi ZERVANOU, Sofia, ACL, pages 96–101. URL : <https://aclanthology.org/W13-2713>.
- JOOSEN, Vanessa (2006). « From Breaktime to Postcards: How Aidan Chambers Goes (Or Does Not Go) Dutch ». *Children’s Literature in Translation: Challenges and Strategies*, édité par Jan VAN COILLIE et Walter P. VERSCHUEREN, Abingdon, Routledge, pages 61–78. DOI : [10.4324/9781315759845](https://doi.org/10.4324/9781315759845).
- JOSSELIN-LERAY, Amélie et Carole FILLIÈRE (2021). « Technologies de la traduction, traduction littéraire et SHS : inévitable cohabitation, possible conciliation, souhaitable réconciliation ? » *La Main de Thôt*, vol. 2021, n° 9 : *La traduction littéraire et SHS à la rencontre des nouvelles*

- technologies de la traduction : enjeux, perspectives et défis*, édité par Amélie JOSSELIN-LERAY et Carole FILLIÈRE. URL : <http://interfas.univ-tlse2.fr/lamaindethot/1027>.
- KAINDL, Klaus (2021). « (Literary) Translator Studies: Shaping the field ». *Literary Translator Studies*, édité par Klaus KAINDL, Waltraud KOLB et Daniela SCHLAGER, Amsterdam, John Benjamins Publishing Company, pages 1–38. DOI : [10.1075/bt1.156.int](https://doi.org/10.1075/bt1.156.int).
- KALCHBRENNER, Nal et Phil BLUNSOM (2013). « Recurrent Continuous Translation Models ». *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, édité par David YAROWSKY, Timothy BALDWIN, Anna KORHONEN, Karen LIVESCU et Steven BETHARD, Seattle (WA), ACL, pages 1700–1709. URL : <https://aclanthology.org/D13-1176>.
- KARPINSKA, Marzena et Mohit IYER (2023). « Large language models effectively leverage document-level context for literary translation, but critical errors persist ». *Proceedings of the Eighth Conference on Machine Translation*, édité par Philipp KOEHN, Barry HADDOW, Tom KOZMI et Christof MONZ, Singapour, ACL, pages 419–451. DOI : [10.18653/v1/2023.wmt-1.41](https://doi.org/10.18653/v1/2023.wmt-1.41).
- KAUFMANN, Jean-Claude (2013). *L'Entretien compréhensif*. 3^e édition, Paris, Armand Colin.
- KAY, Martin (1997). « The Proper Place of Men and Machines in Language Translation ». *Machine Translation*, vol. 12, n^o 1–2 : *New Tools for Human Translators*, pages 3–27.
- KENNY, Dorothy (2012). « The Ethics of Machine Translation ». *Proceedings of the 20th NZSTI National Conference New Zealand Society of Translators and Interpreters*. Reflections on language and technology : the driving forces in the modern world of translation and interpreting, édité par Sibylle FERNER, Auckland, NZSTI, pages 121–131. URL : <https://core.ac.uk/download/pdf/11311284.pdf>.
- éd. (2017). *Human Issues in Translation Technology*. Abingdon, Routledge. DOI : [10.4324/9781315648934](https://doi.org/10.4324/9781315648934).
- (2018). « Machine translation ». *The Routledge Handbook of Translation and Philosophy*, édité par Piers RAWLING et Philip WILSON, Abingdon, Routledge, pages 428–445. DOI : [10.4324/9781315678481](https://doi.org/10.4324/9781315678481).
- (2020). « Machine translation ». *Routledge Encyclopedia of Translation Studies*, édité par Mona BAKER et Gabriela SALDANHA. 3^e édition, Abingdon, Routledge, pages 305–310.
- (2022). « Human and Machine Translation ». *Machine translation for everyone: Empowering users in the age of artificial intelligence*, édité par Dorothy KENNY, Berlin, Language Science Press, pages 23–49. DOI : [10.5281/zenodo.6653406](https://doi.org/10.5281/zenodo.6653406).

- KENNY, Dorothy et Damien HANSEN (2023). « Neural Machine Translation, Large Language Models and Literary Translation: The Story So Far ». Présentation au séminaire de l'ITIA CPD, 8 mai 2023, Dublin.
- KENNY, Dorothy et Marion WINTERS (2020). « Machine translation, ethics and the literary translator's voice ». *Translation Spaces*, vol. 9, n° 1, pages 123–149. DOI : [10.1075/ts.00024.ken](https://doi.org/10.1075/ts.00024.ken).
- (2024). « Customization, Personalization and Style in Literary Machine Translation ». *Translation, Interpreting and Technological Change: Innovations in Research, Practice and Training*, édité par Marion WINTERS, Sharon DEANE-COX et Ursula BÖSER, London, Bloomsbury.
- KHANMOHAMMADI, Reza, Mitra Sadat MIRSHAFIEE, Yazdan REZAAEE JOURYABI et Seyed Abolghasem MIRROSHANDEL (2023). « Prose2Poem: The Blessing of Transformers in Translating Prose to Persian Poetry ». *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 22, n° 6, pages 1–18. DOI : [10.1145/3592791](https://doi.org/10.1145/3592791).
- KHANNA, Tanmai, Jonathan N. WASHINGTON, Francis M. TYERS, Sevilay BAYATLI, Daniel G. SWANSON, Tommi A. PIRINEN, Irene TANG et Hèctor ALÒS I FONT (2021). « Recent advances in Apertium, a free/open-source rule-based machine translation platform for low-resource languages ». *Machine Translation*, vol. 35, n° 4, pages 475–502. DOI : [10.1007/s10590-021-09260-6](https://doi.org/10.1007/s10590-021-09260-6).
- KOCMI, Tom, Christian FEDERMANN, Roman GRUNDKIEWICZ, Marcin JUNCZYS-DOWMUNT, Hitokazu MATSUSHITA et Arul MENEZES (2021). « To Ship or Not to Ship: An Extensive Evaluation of Automatic Metrics for Machine Translation ». *Proceedings of the Sixth Conference on Machine Translation*, édité par Loïc BARRAULT, Ondřej BOJAR, Fethi BOUGARES, Rajen CHATTERJEE, Marta R. COSTA-JUSSÀ, Christian FEDERMANN, Mark FISHEL, Alexander FRASER, Markus FREITAG, Yvette GRAHAM, Roman GRUNDKIEWICZ, Paco GUZMAN, Barry HADDOW, Matthias HUCK, Antonio JIMENO YEPES, Philipp KOEHN, Tom KOCMI, André MARTINS, Makoto MORISHITA et Christof MONZ, Punta Cana, ACL, pages 478–494. URL : <https://aclanthology.org/2021.wmt-1.57>.
- KOEHN, Philipp (2004). « Statistical Significance Tests for Machine Translation Evaluation ». *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, édité par Dekang LIN et Dekai WU, Barcelone, ACL, pages 388–395. URL : <https://aclanthology.org/W04-3250>.
- (2009). *Statistical Machine Translation*. Cambridge, Cambridge University Press. DOI : [10.1017/CB09780511815829](https://doi.org/10.1017/CB09780511815829).

- (2020). *Neural Machine Translation*. Cambridge, Cambridge University Press. DOI : [10.1017/9781108608480](https://doi.org/10.1017/9781108608480).
- KOLB, Waltraud (2024). « 'I Am a Bit Surprised': Literary Translation and Post-Editing Processes Compared ». *Computer-Assisted Literary Translation*, édité par Andrew ROTHWELL, Andy WAY et Roy YOUNDALE, Abingdon, Routledge, pages 53–68. DOI : [10.4324/9781003357391](https://doi.org/10.4324/9781003357391).
- KOLB, Waltraud, Wolfgang U. DRESSLER et Elisa MATTIELLO (2023). « Human and machine translation of occasionalisms in literary texts: Johann Nestroy's *Der Talisman* and its English translations ». *Target*, vol. 35, n° 4, pages 540–572. DOI : [10.1075/target.21147.kol](https://doi.org/10.1075/target.21147.kol).
- KOLB, Waltraud et Tristan MILLER (2022). « Human–computer interaction in pun translation ». *Using Technologies for Creative-Text Translation*, édité par James HADLEY, Kristiina TAIVALKOSKI-SHILOV, Carlos S. C. TEIXEIRA et Antonio TORAL, New York (NY), Routledge, pages 66–88. DOI : [10.4324/9781003094159-4](https://doi.org/10.4324/9781003094159-4).
- KOPONEN, Maarit, Sanna NYQVIST et Kristiina TAIVALKOSKI-SHILOV (2022). « Translating with technology: How digitalisation affects authorship and copyright of literary texts ». *Using Technologies for Creative-Text Translation*, édité par James HADLEY, Kristiina TAIVALKOSKI-SHILOV, Carlos S. C. TEIXEIRA et Antonio TORAL, New York (NY), Routledge, pages 180–198. DOI : [10.4324/9781003094159-9](https://doi.org/10.4324/9781003094159-9).
- KOSKINEN, Kaisa et Minna RUOKONEN (2017). « Love Letters or Hate Mail? Translators' Technology Acceptance in the Light of Their Emotional Narratives ». *Human Issues in Translation Technology*, édité par Dorothy KENNY, Abingdon, Royaume-Uni, Routledge, pages 8–24. DOI : [10.4324/9781315648934](https://doi.org/10.4324/9781315648934).
- KRANZBERG, Melvin (1986). « Technology and History: "Kranzberg's Laws" ». *Technology and Culture*, vol. 27, n° 3, pages 544–560. DOI : [10.2307/3105385](https://doi.org/10.2307/3105385).
- KRINGS, Hans P. (2001). *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes*. Traduit par Geoffrey S. KOPY, Gregory M. SHREVE, Katja MISCHERIKOW et Sarah LITZER. Kent (OH), Kent State University Press.
- KRIZHEVSKY, Alex, Ilya SUTSKEVER et Geoffrey HINTON (2012). « ImageNet Classification with Deep Convolutional Neural Networks ». *Proceedings of the 25th International Conference on Neural Information Processing Systems*, vol. 1, édité par Fernando PEREIRA, Christopher BURGESS, Léon BOTTOU et Kilian WEINBERGER, Lake Tahoe (NV), Curran Associates, Inc., pages 1097–1105. URL : https://papers.nips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html.

- KUDO, Taku (2018b). « Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates ». *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, vol. 1, édité par Iryna GUREVYCH et Yusuke MIYAO, Melbourne, ACL, pages 66–75. DOI : [10.18653/v1/P18-1007](https://doi.org/10.18653/v1/P18-1007).
- KUSSMAUL, Paul et Sonja TIRKKONEN-CONDIT (1995). « Think-Aloud Protocol Analysis in Translation Studies ». *Traduction, terminologie, rédaction*, vol. 8, n° 1, pages 177–199. DOI : [10.7202/037201ar](https://doi.org/10.7202/037201ar).
- KUZMAN, Taja, Špela VINTAR et Mihael ARČAN (2019). « Neural Machine Translation of Literary Texts from English to Slovene ». *Proceedings of the Qualities of Literary Machine Translation*, édité par James HADLEY, Maja POPOVIĆ, Haithem AFLI et Andy WAY, Dublin, EAMT, pages 1–9. URL : <https://aclanthology.org/W19-7301>.
- LACOUR, Philippe (2019). « TraduXio : un environnement numérique pour la traduction collaborative et multilingue de textes culturels ». Présentation à l'atelier *Digit_Hum : Les humanités numériques en langues*, 17 oct. 2019, Paris. URL : https://digithum.huma-num.fr/atelier/2019/5_lacour.php.
- LACOUR, Philippe, Aurélien BÉNEL, Franck EYRAUD, Any FREITAS et Diana ZAMBON (2010). « TIC, collaboration et traduction : vers de nouveaux laboratoires numériques de translocalisation culturelle ». *Meta*, vol. 55, n° 4, pages 674–692. DOI : [10.7202/045685ar](https://doi.org/10.7202/045685ar).
- LACRUZ MANTECÓN, Miguel L. (2023). « Authorship and Rights Ownership in the Machine Translation Era ». *Towards Responsible Machine Translation: Ethical and Legal Considerations in Machine Translation*, édité par Helena MONIZ et Carla PARRA ESCARTÍN, Cham, Springer Nature, pages 71–92. DOI : [10.1007/978-3-031-14689-3_5](https://doi.org/10.1007/978-3-031-14689-3_5).
- LADMIRAL, Jean-René (1994). *Traduire : théorèmes pour la traduction*. 2^e édition, Paris, Gallimard.
- LAGOUDAKI, Pelagia Maria (2009). « Expanding the Possibilities of Translation Memory Systems: From the Translator's Wishlist to the Developer's Design ». Thèse de doctorat, Université de Londres. URL : <https://spiral.imperial.ac.uk/handle/10044/1/7879>.
- LANDRAGIN, Frédéric (2020). *Comment parle un robot ? Les machines à langage dans la science-fiction*. Saint-Mammès, Le Béalial'.
- LARSONNEUR, Claire (2021). « Neural Machine Translation: From Commodity to Commons? » *When Translation Goes Digital: Case Studies and Critical Reflections*, édité par Renée DESJARDINS, Claire LARSONNEUR et Philippe LACOUR, Cham, Palgrave Macmillan, pages 257–280. DOI : [10.1007/978-3-030-51761-8_11](https://doi.org/10.1007/978-3-030-51761-8_11).

- (2023). « L’algorithme sert-il les traducteurs ? Conditions et contexte de travail avec les outils de traduction neuronale ». *Parallèles*, vol. 35, n° 2 : *Approches socio-économiques de la traduction littéraire*, édité par Susan PICKFORD et Olivia GUILLON, pages 90–103. DOI : [10.17462/para.2023.02.10](https://doi.org/10.17462/para.2023.02.10).
- LATHEY, Gillian (2006). « The Translator Revealed: Didacticism, Cultural Mediation and Visions of the Child Reader in Translators’ Prefaces ». *Children’s Literature in Translation: Challenges and Strategies*, édité par Jan VAN COILLIE et Walter P. VERSCHUEREN, Abingdon, Routledge, pages 1–18. DOI : [10.4324/9781315759845](https://doi.org/10.4324/9781315759845).
- (2016). *Translating Children’s Literature*. Abingdon, Routledge. DOI : [10.4324/9781315753515](https://doi.org/10.4324/9781315753515).
- LÄUBLI, Samuel, Rico SENNRICH et Martin VOLK (2018). « Has Machine Translation Achieved Human Parity? A Case for Document-level Evaluation ». *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, édité par Ellen RILOFF, David CHIANG, Julia HOCKENMAIER et Jun’ichi TSUJII, Bruxelles, ACL, pages 4791–4796. DOI : [10.18653/v1/D18-1512](https://doi.org/10.18653/v1/D18-1512).
- LÄUBLI, Samuel, Patrick SIMIANER, Joern WUEBKER, Geza KOVACS, Rico SENNRICH et Spence GREEN (2022). « The Impact of Text Presentation on Translator Performance ». *Target*, vol. 34, n° 2, pages 309–342. DOI : [10.1075/target.20006.lau](https://doi.org/10.1075/target.20006.lau).
- LAVAUULT-OLLÉON, Élisabeth (2011). « L’ergonomie, nouveau paradigme pour la traductologie ». *ILCEA*, n° 14 : *Traduction et Ergonomie*. DOI : [10.4000/ilcea.1078](https://doi.org/10.4000/ilcea.1078).
- (2016). « Traducteurs à l’œuvre : une perspective ergonomique en traductologie appliquée ». *ILCEA*, n° 27. DOI : [10.4000/ilcea.4051](https://doi.org/10.4000/ilcea.4051).
- LEBLANC, Matthieu (2013). « Translators on translation memory (TM): Results of an ethnographic study in three translation services and agencies ». *Translation & Interpreting*, vol. 5, n° 2, pages 1–13. DOI : [10.12807/ti.105202.2013.a01](https://doi.org/10.12807/ti.105202.2013.a01).
- LEBON, Marc (2022). « Contra Automata : orgueil et préjugés ? » *Journal of Data Mining & Digital Humanities, Vers une robotique du traduire ?*, édité par Anne BAILLOT, Ellen CARTER, Thierry GRASS et Pablo RUIZ FABO, pages 1–17. DOI : [10.46298/jdmdh.9056](https://doi.org/10.46298/jdmdh.9056).
- LEE, Tong King (2011). « The death of the translator in machine translation: A bilingual poetry project ». *Target*, vol. 23, n° 1, pages 92–112. DOI : [10.1075/target.23.1.06lee](https://doi.org/10.1075/target.23.1.06lee).
- LÉON, Jacqueline (2002). « Le CNRS et les débuts de la traduction automatique en France ». *La revue pour l’histoire du CNRS*, n° 6. DOI : [10.4000/histoire-cnrs.3461](https://doi.org/10.4000/histoire-cnrs.3461).
- LJUBEŠIĆ, Nikola et Antonio TORAL (2014). « caWaC – A web corpus of Catalan and its application to language modeling and machine translation ». *Proceedings of the Ninth International*

- Conference on Language Resources and Evaluation*, édité par Nicoletta CALZOLARI, Khalid CHOUKRI, Thierry DECLERCK, Hrafn LOFTSSON, Bente MAEGAARD, Joseph MARIANI, Moreno ASUNCION, Jan ODIJK et Stelios PIPERIDIS, Reykjavik, ELRA, pages 1728–1732. URL : http://www.lrec-conf.org/proceedings/lrec2014/pdf/841_Paper.pdf.
- LOMMEL, Arle, Hans USZKOREIT et Aljoscha BURCHARDT (2014). « Multidimensional Quality Metrics (MQM): A Framework for Declaring and Describing Translation Quality Metrics ». *Tradumàtica*, n° 12, pages 455–463. DOI : [10.5565/rev/tradumatica.77](https://doi.org/10.5565/rev/tradumatica.77).
- LONG, Hoyt (2022). « Learning to Live with Machine Translation ». *New Literary History*, vol. 53, n° 4, pages 721–753. DOI : [10.1353/nlh.2022.a898327](https://doi.org/10.1353/nlh.2022.a898327).
- LOOCK, Rudy (2019). « La plus-value de la biotraduction face à la machine ». *Traduire*, n° 241, pages 54–65. DOI : [10.4000/traduire.1848](https://doi.org/10.4000/traduire.1848).
- (2020). Introduction de la journée d'études *Traduction et Qualité 2020 : Biotraduction et traduction automatique*, 31 janv. 2020, Roubaix. URL : <https://tq2020.sciencesconf.org/>.
- (2023). « Quelle place accorder à la traduction automatique dans nos enseignements ? » Présentation au colloque *Éthique(s) et traduction à l'ère contemporaine*, 13 févr. 2023, Avignon. URL : <https://etec.sciencesconf.org/>.
- LOPONEN, Mika (2009). « Translating Irrealia: Creating a Semiotic Framework for the Translation of Fictional Cultures ». *Chinese semiotic studies*, vol. 2, n° 1, pages 165–175. DOI : [10.1515/css-2009-0117](https://doi.org/10.1515/css-2009-0117).
- LUONG, Thang, Hieu PHAM et Christopher MANNING (2015). « Effective Approaches to Attention-based Neural Machine Translation ». *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, édité par Lluís MÀRQUEZ, Chris CALLISON-BURCH et Jian SU, Lisbonne, ACL, pages 1412–1421. DOI : [10.18653/v1/D15-1166](https://doi.org/10.18653/v1/D15-1166).
- LUPO, Lorenzo, Marco DINARELLI et Laurent BESACIER (2022a). « Divide and Rule: Effective Pre-Training for Context-Aware Multi-Encoder Translation Models ». *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, vol. 1, édité par Philipp KOEHN, Loïc BARRAULT, Ondřej BOJAR, Fethi BOUGARES, Rajen CHATTERJEE, Marta R. COSTA-JUSSÀ, Christian FEDERMANN, Mark FISHEL, Alexander FRASER, Markus FREITAG, Yvette GRAHAM, Roman GRUNDKIEWICZ, Paco GUZMAN, Barry HADDOW, Matthias HUCK, Antonio JIMENO YEPES, Tom KOCMI, André MARTINS, Makoto MORISHITA, Christof MONZ, Masaaki NAGATA, Toshiaki NAKAZAWA, Matteo NEGRI, Aurélie NÉVÉOL, Mariana NEVES, Martin POPEL, Marco TURCHI et Marcos ZAMPIERI, Dublin, ACL, pages 4557–4572. DOI : [10.18653/v1/2022.acl-long.312](https://doi.org/10.18653/v1/2022.acl-long.312).

- (2022b). « Focused Concatenation for Context-Aware Neural Machine Translation ». *Proceedings of the Seventh Conference on Machine Translation*, édité par Smaranda MURESAN, Preslav NAKOV et Aline VILLAVICENCIO, Abou Dabi, ACL, pages 830–842. URL : <https://aclanthology.org/2022.wmt-1.77>.
- MA, Qingsong, Johnny WEI, Ondřej BOJAR et Yvette GRAHAM (2019). « Results of the WMT19 Metrics Shared Task: Segment-Level and Strong MT Systems Pose Big Challenges ». *Proceedings of the Fourth Conference on Machine Translation*, vol. 2, édité par Ondřej BOJAR, Rajen CHATTERJEE, Christian FEDERMANN, Mark FISHEL, Yvette GRAHAM, Barry HADDOW, Matthias HUCK, Antonio Jimeno YEPES, Philipp KOEHN, André MARTINS, Christof MONZ, Matteo NEGRI, Aurélie NÉVÉOL, Mariana NEVES, Matt POST, Marco TURCHI et Karin VERSPOOR, Florence, ACL, pages 62–90. DOI : [10.18653/v1/W19-5302](https://doi.org/10.18653/v1/W19-5302).
- MACKEN, Lieve, Bram VANROY, Luca DESMET et Arda TEZCAN (2022). « Literary translation as a three-stage process: machine translation, post-editing and revision ». *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation*, édité par Helena MONIZ, Lieve MACKEN, Andrew RUFENER, Loïc BARRAULT, Marta R. COSTA-JUSSÀ, Christophe DECLERCQ, Maarit KOPONEN, Ellie KEMP, Spyridon PILOS, Mikel L. FORCADA, Carolina SCARTON, Joachim VAN DEN BOGAERT, Joke DAEMS, Arda TEZCAN, Bram VANROY et Margot FONTEYNE, Gand, EAMT, pages 101–110. URL : <https://aclanthology.org/2022.eamt-1.13>.
- MAGNÉ, Bernard (2000). « Machines à écrire, machine à lire ». *Études françaises*, vol. 36, n° 2 : *Internet et littérature : nouveaux espaces d'écriture*, pages 119–128. DOI : [10.7202/005258ar](https://doi.org/10.7202/005258ar).
- MALROUX, Claire, Françoise CARTANO, Anne WADE MINKOWSKI, Philippe CARDINAL et François-Xavier JAUJARD (1988). *Actes des quatrièmes assises de la traduction littéraire*. Arles, Actes Sud. URL : <https://www.atlas-citl.org/wp-content/uploads/pdf/4actes.pdf>.
- MARCO, Josep (2009). « Training Translation Researchers: An Approach Based on Models and Best Practice ». *The Interpreter and Translator Trainer*, vol. 3, n° 1, pages 13–35. DOI : [10.1080/1750399X.2009.10798779](https://doi.org/10.1080/1750399X.2009.10798779).
- MARIE, Benjamin, Atsushi FUJITA et Raphael RUBINO (2021). « Scientific Credibility of Machine Translation Research: A Meta-Evaluation of 769 Papers ». *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, vol. 1, édité par Chengqing ZONG, Fei XIA, Wenjie LI et Roberto NAVIGLI, Bangkok, ACL, pages 7297–7306. DOI : [10.18653/v1/2021.acl-long.566](https://doi.org/10.18653/v1/2021.acl-long.566).

- MARIE, Benjamin, Raphael RUBINO et Atsushi FUJITA (2020). « Tagged Back-Translation Revisited: Why Does It Really Work? » *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, édité par Dan JURAFSKY, Joyce CHAI, Natalie SCHLUTER et Joel TETREAULT, ACL, pages 5990–5997. DOI : [10.18653/v1/2020.acl-main.532](https://doi.org/10.18653/v1/2020.acl-main.532).
- MARX, Paris (2023). « Why AI is a Threat to Artists, with Molly Crabapple ». Podcast *Tech Won't Save Us*, 29 juin 2023. URL : https://techwontsave.us/episode/174_why_ai_is_a_threat_to_artists_w_molly_crabapple.
- MATHUR, Nitika, Timothy BALDWIN et Trevor COHN (2020). « Tangled up in BLEU: Reevaluating the Evaluation of Automatic Machine Translation Evaluation Metrics ». *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, édité par Dan JURAFSKY, Joyce CHAI, Natalie SCHLUTER et Joel TETREAULT, Seattle (WA), ACL, pages 4984–4997. DOI : [10.18653/v1/2020.acl-main.448](https://doi.org/10.18653/v1/2020.acl-main.448).
- MATUSOV, Evgeny (2019). « The Challenges of Using Neural Machine Translation for Literature ». *Proceedings of the Qualities of Literary Machine Translation*, édité par James HADLEY, Maja POPOVIĆ, Haithem AFLI et Andy WAY, Dublin, EAMT, pages 10–19. URL : <https://aclanthology.org/W19-7302>.
- (2022). « Giving translator full control: customizable MT for literary translation ». Présentation au colloque *Literary Machine Translation as a Human-Machine Dialectic*, 6 oct. 2022, Liège. URL : <https://www.apptek.com/post/transform-your-mt-system-into-a-trusted-colleague-customizable-mt-for-literary-translations>.
- MAZOYER, Renaud (2021). « Traduction d'essai et TAO ». *La Main de Thôt*, vol. 2021, n° 9 : *La traduction littéraire et SHS à la rencontre des nouvelles technologies de la traduction : enjeux, perspectives et défis*, édité par Amélie JOSSELIN-LERAY et Carole FILLIÈRE. URL : <http://interfas.univ-tlse2.fr/lamaindethot/991>.
- MCCULLOCH, Warren Sturgis et Walter PITTS (1943). « A logical calculus of the ideas immanent in nervous activity ». *The bulletin of mathematical biophysics*, vol. 5, n° 4, pages 115–133. DOI : [10.1007/BF02478259](https://doi.org/10.1007/BF02478259).
- MCKAY, David (2023). *Entretien personnel*. 27 juin 2023.
- MERTEN, Pascaline (2018). « Unité de traduction et identité des traducteurs : l'impact des outils d'aide à la traduction ». *Équivalences*, vol. 45, n° 1–2, édité par Nicolas FROELIGER, Lance HEWSON et Christian BALLIU, pages 193–216. URL : https://www.persee.fr/doc/equiv_0751-9532_2018_num_45_1_1540.

- (2022). « Spring, summer, fall, winter... and spring. Une histoire de la traduction automatique ». *Atelier de traduction*, n° 37, édité par Christian BALLIU, Mathilde FONTANET et Nicolas FROELIGER, pages 171–194. URL : https://usv.ro/fisiere_utilizator/file/atelierdetraduction/arhive/2022/37/AT_37_%2018_FEB_2022-171-194.pdf.
- MIELKE, Sabrina J., Ryan COTTERELL, Kyle GORMAN, Brian ROARK et Jason EISNER (2019). « What Kind of Language Is Hard to Language-Model? » *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, édité par Anna KORHONEN, David TRAUM et Lluís MÀRQUEZ, Florence, ACL, pages 4975–4989. DOI : [10.18653/v1/P19-1491](https://doi.org/10.18653/v1/P19-1491).
- MIKOLOV, Tomas, Ilya SUTSKEVER, Kai CHEN, Greg CORRADO et Jeff DEAN (2013). « Distributed Representations of Words and Phrases and their Compositionality ». *Proceedings of the 26th International Conference on Neural Information Processing Systems*, vol. 2, édité par Christopher BURGESS, Léon BOTTOU, Max WELING, Zoubin GHAHRAMANI et Kilian WEINBERGER, Lake Tahoe (NV), Curran Associates, Inc., pages 3111–3119. URL : https://papers.nips.cc/paper_files/paper/2013/hash/9aa42b31882ec039965f3c4923ce901b-Abstract.html.
- MIRKIN, Shachar et Jean-Luc MEUNIER (2015). « Personalized Machine Translation: Predicting Translational Preferences ». *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, édité par MÀRQUEZ LLUÍS, Chris CALLISON-BURCH et Jian SU, Lisbonne, ACL, pages 2019–2025. DOI : [10.18653/v1/D15-1238](https://doi.org/10.18653/v1/D15-1238).
- MIRKIN, Shachar, Scott NOWSON, Caroline BRUN et Julien PEREZ (2015). « Motivating Personality-aware Machine Translation ». *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, édité par MÀRQUEZ LLUÍS, Chris CALLISON-BURCH et Jian SU, Lisbonne, ACL, pages 1102–1108. DOI : [10.18653/v1/D15-1130](https://doi.org/10.18653/v1/D15-1130).
- MOHAR, Tjaša, Sara ORTHABER et Tomaž ONIČ (2020). « Machine Translated Atwood: Utopia or Dystopia? » *ELOPE*, vol. 17, n° 1, pages 125–141. DOI : [10.4312/elope.17.1.125-141](https://doi.org/10.4312/elope.17.1.125-141).
- MOORKENS, Joss (2017). « Under pressure: translation in times of austerity ». *Perspectives*, vol. 25, n° 3, pages 464–477. DOI : [10.1080/0907676X.2017.1285331](https://doi.org/10.1080/0907676X.2017.1285331).
- (2020). « “A tiny cog in a large machine”: Digital Taylorism in the translation industry ». *Translation Spaces*, vol. 9, n° 1, pages 12–34. DOI : [10.1075/ts.00019.moo](https://doi.org/10.1075/ts.00019.moo).
- MOORKENS, Joss et Dave LEWIS (2019). « Research Questions and a Proposal for the Future Governance of Translation Data ». *JoSTrans*, n° 32, pages 2–25. URL : https://jostrans.soap2.ch/issue32/art_moorkens.php.

- MOORKENS, Joss et Dave LEWIS (2020). « Copyright and the re-use of translation as data ». *The Routledge Handbook of Translation and Technology*, édité par Minako O'HAGAN, Abingdon, Routledge, pages 469–481.
- MOORKENS, Joss et Sharon O'BRIEN (2017). « Assessing user interface needs of post-editors of machine translation ». *Human Issues in Translation Technology*, édité par Dorothy KENNY, Abingdon, Routledge, pages 109–130. DOI : [10.4324/9781315648934](https://doi.org/10.4324/9781315648934).
- MOORKENS, Joss et Marta ROCCHI (2021). « Ethics in the translation industry ». *The Routledge Handbook of Translation and Ethics*, édité par Kaisa KOSKINEN et Nike K. POKORN, Abingdon, Routledge, pages 320–337. DOI : [10.4324/9781003127970](https://doi.org/10.4324/9781003127970).
- MOORKENS, Joss, Antonio TORAL, Sheila CASTILHO et Andy WAY (2018). « Translators' Perceptions of Literary Post-Editing using Statistical and Neural Machine Translation ». *Translation Spaces*, vol. 7, n° 2, pages 240–262. DOI : [10.1075/ts.18014.moo](https://doi.org/10.1075/ts.18014.moo).
- MORETTI, Franco (2013). *Distant Reading*. Londres et New York (NY), Distant Reading.
- MOROZOV, Evgeny (2013). *To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems that Don't Exist*. Londres, Allen Lane.
- MOUNIN, Georges (1963). « Pourquoi la traduction automatique exclut les textes littéraires ». *Linguistique et traduction*, Bruxelles, Dessart et Mardaga, 1976, pages 121–138.
- (1964). *La Machine à traduire : Histoire des problèmes linguistiques*. La Haye, Mouton & Co.
- (1976). *Linguistique et traduction*. Bruxelles, Dessart et Mardaga.
- (2016). *Les Belles Infidèles*. 3^e édition, Villeneuve d'Ascq, Presses Universitaires du Septentrion.
- MÜLLER, Mathias, Annette RIOS et Rico SENNRICH (2020). « Domain Robustness in Neural Machine Translation ». *Proceedings of the 14th Conference of the Association for Machine Translation in the Americas*, vol. 1, édité par Michael DENKOWSKI et Christian FEDERMANN, AMTA, pages 151–164. URL : <https://aclanthology.org/2020.amta-research.14>.
- MUÑOZ, Martín Ricardo (2010). « On Paradigms and Cognitive Translatology ». *Translation and Cognition*, édité par Gregory M. SHREVE et Erik ANGELONE, Amsterdam, John Benjamins Publishing Company, pages 169–187. DOI : [10.1075/ata.xv.10mun](https://doi.org/10.1075/ata.xv.10mun).
- MUSY, Gilbert (1989). « La traduction littéraire “assistée par ordinateur” ». *Études de lettres*, n° 221, pages 85–88.
- NARAYANAN, Arvind (2019). « How to recognize AI snake oil ». Présentation, 19 nov. 2019, Cambridge (MA). URL : <https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf>.
- NEWMARK, Peter (1981). *Approaches to Translation*. Oxford, Pergamon Press.

- NIKOLAJEVA, Maria (2016). *Children's Literature Comes of Age: Toward a New Aesthetic*. Abingdon, Routledge. DOI : [10.4324/9781315667492](https://doi.org/10.4324/9781315667492).
- NUSSBAUM, Martha C. (1995). *Poetic Justice: The Literary Imagination and Public Life*. Boston (MA), Beacon Press.
- Ó MURCHÚ, Eoin P. (2019). « Using Intergaelic to Pre-Translate and Subsequently Post-Edit a Sci-Fi Novel from Scottish Gaelic to Irish ». *Proceedings of the Qualities of Literary Machine Translation*, édité par James HADLEY, Maja POPOVIĆ, Haithem AFLI et Andy WAY, Dublin, EAMT, pages 20–25. URL : <https://www.aclweb.org/anthology/W19-7303.pdf>.
- O'BRIEN, Sharon (2002). « Teaching post-editing: a proposal for course content ». *Proceedings of the 6th EAMT Workshop: Teaching Machine Translation*, Manchester, EAMT, pages 99–106. URL : <https://aclanthology.org/2002.eamt-1.11>.
- (2011). « Towards predicting post-editing productivity ». *Machine Translation*, vol. 25, n° 3, pages 197–215. DOI : [10.1007/s10590-011-9096-7](https://doi.org/10.1007/s10590-011-9096-7).
- (2012). « Translation as Human–Computer Interaction ». *Translation Spaces*, vol. 1, n° 1, pages 101–122. DOI : [10.1075/ts.1.05obr](https://doi.org/10.1075/ts.1.05obr).
- O'BRIEN, Sharon et Owen CONLAN (2018). « Moving Towards Personalising Translation Technology ». *Moving Boundaries in Translation Studies*, édité par Helle DAM, Matilde BRØGGER et Karen ZETHSEN, Abingdon, Routledge, pages 81–97. DOI : [10.4324/9781315121871](https://doi.org/10.4324/9781315121871).
- O'BRIEN, Sharon, Maureen EHRENSBERGER-DOW, Marcel HASLER et Megan CONNOLLY (2017). « Irritating CAT Tool Features that Matter to Translators ». *HERMES*, n° 56, pages 145–162. DOI : [10.7146/hj1cb.v0i56.97229](https://doi.org/10.7146/hj1cb.v0i56.97229).
- O'HAGAN, Minako, éd. (2020). *The Routledge Handbook of Translation and Technology*. Abingdon, Routledge.
- O'SULLIVAN, Emer (1993). « The Fate of the Dual Addressee in the Translation of Children's Literature ». *New Comparison*, n° 16, pages 109–119.
- OESER, Hans-Christian (2020). « Duel with DeepL ». *Counterpoint*, n° 4, pages 19–23. URL : https://www.ceatl.eu/wp-content/uploads/2020/12/Counterpoint_2020_04_article_05.pdf.
- OITTINEN, Riitta (2006). « No Innocent Act: On the Ethics of Translating for Children ». *Children's Literature in Translation: Challenges and Strategies*, édité par Jan VAN COILLIE et Walter P. VERSCHUEREN, Abingdon, Routledge, pages 35–45. DOI : [10.4324/9781315759845](https://doi.org/10.4324/9781315759845).
- OLIVER, Antoni González (2017). « InLéctor: Automatic Creation of Bilingual E-Books ». *Tradumàtica*, n° 15, pages 21–47. DOI : [10.5565/rev/tradumatica.191](https://doi.org/10.5565/rev/tradumatica.191).

- OLIVER, Antoni González (2024). « Author-Tailored Neural Machine Translation Systems for Literary Works ». *Computer-Assisted Literary Translation*, édité par Andrew ROTHWELL, Andy WAY et Roy YOUNG, Abingdon, Routledge, pages 126–141. DOI : [10.4324/9781003357391](https://doi.org/10.4324/9781003357391).
- OLIVER, Antoni González, Antonio TORAL et Ana GUERBEROF-ARENAS (2022). « Bilingual e-books via neural machine translation and their reception ». *Using Technologies for Creative-Text Translation*, édité par James HADLEY, Kristiina TAIVALKOSKI-SHILOV, Carlos S. C. TEIXEIRA et Antonio TORAL, New York (NY), Routledge, pages 89–115. DOI : [10.4324/9781003094159-5](https://doi.org/10.4324/9781003094159-5).
- PAPAVASSILIOU, Périclès (2007). « Traductologie et sciences cognitives : une dialectique prometteuse ». *Meta*, vol. 52, n° 1, pages 29–36. DOI : [10.7202/014717ar](https://doi.org/10.7202/014717ar).
- PASCUA-FEBLES, Isabel (2006). « Translating Cultural References: The Language of Young People in Literary Texts ». *Children's Literature in Translation: Challenges and Strategies*, édité par Jan VAN COILLIE et Walter P. VERSCHUEREN, Abingdon, Routledge, pages 97–110. DOI : [10.4324/9781315759845](https://doi.org/10.4324/9781315759845).
- PATTERSON, David, Joseph GONZALEZ, Quoc LE, Chen LIANG, Lluís-Miquel MUNGUÍA, Daniel ROTHCHILD, David SO, Maud TEXIER et Jeff DEAN (2021). « Carbon Emissions and Large Neural Network Training ». Prépublication *arXiv*, 3^e éd. ARXIV : [2104.10350v3](https://arxiv.org/abs/2104.10350v3).
- PENNINGTON, Jeffrey, Richard SOCHER et Christopher D. MANNING (2014). « GloVe: Global Vectors for Word Representation ». *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, édité par Alessandro MOSCHITTI, Bo PANG et Walter DAELEMANS, Doha, ACL, pages 1532–1543. DOI : [10.3115/v1/D14-1162](https://doi.org/10.3115/v1/D14-1162).
- PEREC, Georges (1989). « Deux cent quarante-trois cartes postales en couleurs véritables ». *L'Infra-ordinaire*, Paris, Seuil, pages 33–67.
- PÉREZ-ORTIZ, Juan Antonio, Mikel L. FORCADA et Felipe SÁNCHEZ-MARTÍNEZ (2022). « How neural machine translation works ». *Machine translation for everyone: Empowering users in the age of artificial intelligence*, édité par Dorothy KENNY, Berlin, Language Science Press, pages 141–164. DOI : [10.5281/zenodo.6653406](https://doi.org/10.5281/zenodo.6653406).
- PETRAK, Marta, Mia UREMOVIĆ et Bogdanka PAVELIN LEŠIĆ (2022). « Fine-grained human evaluation of NMT applied to literary text: case study of a French-to-Croatian translation ». *Proceedings of the Conference on Language Technologies & Digital Humanities 2022*, édité par Darja FIŠER et Tomaž ERJAVEC, Ljubljana, Institute of Contemporary History, pages 141–146. URL : <https://nl.ijs.si/jtdh22/proceedings-en.html>.
- POIBEAU, Thierry (2019). *Babel 2.0 : Où va la traduction automatique ?* Paris, Odile Jacob.

- (2022). « On “Human Parity” and “Super Human Performance” in Machine Translation Evaluation ». *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, édité par Nicoletta CALZOLARI, Frédéric BÉCHET, Philippe BLACHE, Khalid CHOUKRI, Christopher CIERI, Thierry DECLERCK, Sara GOGGI, Hitoshi ISAHARA, Bente MAEGAARD, Joseph MARIANI, Hélène MAZO, Jan ODIJK et Stelios PIPERIDIS, Marseille, ELRA, pages 6018–6023. URL : <https://aclanthology.org/2022.lrec-1.647>.
- PONCHARAL, Bruno (2021). « La TA à l’épreuve du texte littéraire : d’une (im)possible restitution de l’expérience de lecture ? » *La main de Thôt*, n° 9 : *La traduction littéraire et SHS à la rencontre des nouvelles technologies de la traduction : enjeux, perspectives et défis*, édité par Amélie JOSSELINE-LERAY et Carole FILLIÈRE, pages 29–36. URL : <http://interfas.univ-tlse2.fr/lamaindethot/index.php?id=983>.
- PRINCE, Nathalie (2021). *La littérature de jeunesse : Pour une théorie littéraire*. 3^e édition, Malakoff, Armand Colin.
- PRZYBOCKI, Mark, Kay PETERSON, Sébastien BRONSART et Gregory SANDERS (2009). « The NIST 2008 Metrics for machine translation challenge—overview, methodology, metrics, and results ». *Machine Translation*, vol. 23, n° 2–3 : *Automated Metrics for Machine Translation Evaluation*, édité par Alon LAVIE et Mark PRZYBOCKI, pages 71–103. DOI : [10.1007/s10590-009-9065-6](https://doi.org/10.1007/s10590-009-9065-6).
- PYM, Anthony (2003). « Translational ethics and electronic technologies ». Présentation au *VI Seminário de Tradução Científica e Técnica em Língua Portuguesa A Profissionalização do Tradutor*, 11 nov. 2003, Lisbonne. URL : https://usuaris.tinet.cat/apym/on-line/translation/lisbon_ethics.pdf.
- (2011). « What Technology Does to Translating ». *Translation & Interpreting*, vol. 3, n° 1, pages 1–9. URL : <https://www.trans-int.org/index.php/transint/article/view/121>.
- QUENEAU, Raymond (1961). *Cent mille milliards de poèmes*. Paris, Gallimard.
- RADFORD, Alec, Jeff WU, Rewon CHILD, David LUAN, Dario AMODEI et Ilya SUTSKEVER (2019). « Language Models are Unsupervised Multitask Learners ». Prépublication. URL : https://d4mucfpksyvw.cloudfront.net/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.
- RAGNI, Valentina et Lucas Nunes VIEIRA (2021). « What Has Changed with Neural Machine Translation? A Critical Review of Human Factors ». *Perspectives*, vol. 30, n° 1, pages 137–158. DOI : [10.1080/0907676X.2021.1889005](https://doi.org/10.1080/0907676X.2021.1889005).
- REITER, Ehud (2018). « A Structured Review of the Validity of BLEU ». *Computational Linguistics*, vol. 44, n° 3, pages 393–401. DOI : [10.1162/coli_a_00322](https://doi.org/10.1162/coli_a_00322).

- REYNAUD, Pascal et Tiphaine KLEIN (2018). *Quels droits d'auteur pour le traducteur professionnel ?* Paris, Les Éditions de la SFT. URL : https://www.sft.fr/sites/default/files/2021-11/Quels_droits_dauteur_pour_le_traducteur_professionnel.pdf.
- REYNOLDS, Kimberley (2007). *Radical Children's Literature: Future Visions and Aesthetic Transformations in Juvenile Fiction*. New York (NY), Palgrave Macmillan.
- RISKU, Hanna (2010). « A cognitive scientific view on technical communication and translation: Do embodiment and situatedness really make a difference? » *Target*, vol. 22, n° 1, pages 94–111. DOI : [10.1075/target.22.1.06ris](https://doi.org/10.1075/target.22.1.06ris).
- RISKU, Hanna et Florian WINDHAGER (2013). « Extended Translation: A Sociocognitive Research Agenda ». *Target*, vol. 25, n° 1, pages 33–45. DOI : [10.1075/target.25.1.04ris](https://doi.org/10.1075/target.25.1.04ris).
- ROBINSON, Douglas (1991). *The Translator's Turn*. Baltimore (MD), Johns Hopkins University Press.
- (2012). *Becoming a Translator: An Introduction to the Theory and Practice of Translation*. 3^e édition, Abingdon, Routledge. DOI : [10.4324/9780429276606](https://doi.org/10.4324/9780429276606).
- ROSENBLATT, Frank (1957). *The Perceptron: A Perceiving and Recognizing Automaton*. Buffalo (NY), Cornell Aeronautical Laboratory. URL : <https://blogs.umass.edu/brain-wars/files/2016/03/rosenblatt-1957.pdf>.
- ROSENBLATT, Louise Michelle (1994). *The reader, the text, the poem: the transactional theory of the literary work*. Carbondale (IL), Southern Illinois University Press.
- ROSSI, Caroline (2018). « Acquisition, traduction, usages : des linguistiques énonciatives et cognitives à l'automatisation ». Dossier d'habilitation à diriger des recherches, Université Lumière Lyon 2. HAL : [tel-03659740](https://hal.archives-ouvertes.fr/tel-03659740).
- (2019). « Les usages actuels de la traduction automatique ». Présentation à l'atelier *Digit_Hum : Les humanités numériques en langues*, 17 oct. 2019, Paris. URL : https://dighum.humanum.fr/atelier/2019/4_rossi.php.
- ROSSI, Caroline et Jean-Pierre CHEVROT (2019). « Uses and perceptions of machine translation at the European Commission ». *JoSTrans*, n° 31 : *Post-Editing in Practice: Process, Product and Networks*, édité par Lucas Nunes VIEIRA, Elisa ALONSO et Lindsay BYWOOD, pages 177–200. URL : https://jostrans.soap2.ch/issue31/art_rossi.php.
- ROSSI, Caroline et Emmanuelle ESPERANÇA-RODIER (2019). « With or without post-editing processes? Evidence for a gap in machine translation evaluation ». *Proceedings of the Second MEMENTO workshop on Modelling Parameters of Cognitive Effort in Translation Production*,

- édité par Michael CARL et Silvia HANSEN-SCHIRRA, Dublin, EAMT, pages 9–10. URL : <https://aclanthology.org/W19-7005>.
- ROTHWELL, Andrew (2009). « Translating ‘Pure Nonsense’: Walter Benjamin Meets Systran on the Dissecting Table of Dada ». *Romance Studies*, vol. 27, n° 4, pages 259–272. DOI : [10.1179/026399009X12523296128713](https://doi.org/10.1179/026399009X12523296128713).
- (2024). « Retranslating Proust Using CAT, MT and Other Tools ». *Computer-Assisted Literary Translation*, édité par Andrew ROTHWELL, Andy WAY et Roy YOUNDALE, Abingdon, Routledge, pages 106–125. DOI : [10.4324/9781003357391](https://doi.org/10.4324/9781003357391).
- ROTHWELL, Andrew, Andy WAY et Roy YOUNDALE, éd. (2024). *Computer-Assisted Literary Translation*. Abingdon, Routledge. DOI : [10.4324/9781003357391](https://doi.org/10.4324/9781003357391).
- RUDVIN, Mette et Francesca ORLATI (2006). « Dual Readership and Hidden Subtexts in Children’s Literature: The Case of Salman Rushdie’s “Haroun and the Sea of Stories” ». *Children’s Literature in Translation: Challenges and Strategies*, édité par Jan VAN COILLIE et Walter P. VERSCHUEREN, Abingdon, Routledge, pages 157–184. DOI : [10.4324/9781315759845](https://doi.org/10.4324/9781315759845).
- RUFFO, Paola (2018). « Human-Computer Interaction in Translation: Literary Translators on Technology and Their Roles ». *Proceedings of the 40th Conference Translating and the Computer*, édité par CHAMBERS, DAVID AND DRUGAN, JOANNA AND ESTEVES-FERREIRA, JOÃO AND MACAN, JULIET MARGARET AND MITKOV, RUSLAN AND STEFANOV, OLAF-MICHAEL, Londres, Éditions Tradulex, pages 127–131. URL : <https://www.asling.org/tc40/wp-content/uploads/TC40-Proceedings.pdf#chapter.14>.
- (2021). « In-between role and technology: literary translators on navigating the new socio-technological paradigm ». Thèse de doctorat, Heriot-Watt University. URL : <http://hdl.handle.net/10399/4750>.
- (2022). « Collecting literary translators’ narratives: Towards a new paradigm for technological innovation in literary translation ». *Using Technologies for Creative-Text Translation*, édité par James HADLEY, Kristiina TAIVALKOSKI-SHILOV, Carlos S. C. TEIXEIRA et Antonio TORAL, New York (NY), Routledge, pages 18–39. DOI : [10.4324/9781003094159-2](https://doi.org/10.4324/9781003094159-2).
- SAGE, Angie (2005–2013). *Septimus Heap*. 7 tomes. New York (NY), HarperCollins.
- ŞAHİN, Mehmet et Sabri GÜRSES (2019). « Would MT Kill Creativity in Literary Retranslation? » *Proceedings of the Qualities of Literary Machine Translation*, édité par James HADLEY, Maja POPOVIĆ, Haithem AFLI et Andy WAY, Dublin, EAMT, pages 26–34. URL : <https://aclanthology.org/W19-7304>.

- ŞAHIN, Mehmet et Sabri GÜRSES (2023). « A call for a fair translatsphere in the post-digital era ». *Parallèles*, vol. 34, n° 2, pages 1–17. DOI : [10.17462/para.2022.02.01](https://doi.org/10.17462/para.2022.02.01).
- SALDANHA, Gabriela et Sharon O'BRIEN (2013). *Research Methodologies in Translation Studies*. Abingdon, Routledge. DOI : [10.4324/9781315760100](https://doi.org/10.4324/9781315760100).
- SAPIRO, Gisèle et Anaïs BOKOBZA (2008). « L'essor des traductions littéraires en français ». *Translatio : Le marché de la traduction en France à l'heure de la mondialisation*, édité par Gisèle SAPIRO, Paris, CNRS Éditions, pages 145–173. DOI : [10.4000/books.editions-cnrs.9479](https://doi.org/10.4000/books.editions-cnrs.9479).
- SCHUMACHER, Perrine (2023). « La post-édition de traduction automatique en contexte d'apprentissage : Effets sur la qualité et défis pour l'enseignement de la traduction ». Thèse de doctorat, Université de Liège. URL : <https://hdl.handle.net/2268/302239>.
- SENNRICH, Rico, Barry HADDOW et Alexandra BIRCH (2016a). « Improving Neural Machine Translation Models with Monolingual Data ». *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, vol. 1, édité par Katrin ERK et Noah SMITH, Berlin, ACL, pages 86–96. DOI : [10.18653/v1/P16-1009](https://doi.org/10.18653/v1/P16-1009).
- (2016b). « Neural Machine Translation of Rare Words with Subword Units ». *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, vol. 1, édité par Katrin ERK et Noah SMITH, Berlin, ACL, pages 1715–1725. DOI : [10.18653/v1/P16-1162](https://doi.org/10.18653/v1/P16-1162).
- SENNRICH, Rico et Biao ZHANG (2019). « Revisiting Low-Resource Neural Machine Translation: A Case Study ». *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, édité par Anna KORHONEN, David TRAUM et Lluís MÀRQUEZ, Florence, ACL, pages 211–221. DOI : [10.18653/v1/P19-1021](https://doi.org/10.18653/v1/P19-1021).
- SERVAL, Nathalie, trad. (2005–2012). *Magyk*. Par Angie SAGE. 6 tomes. Paris, Albin Michel.
- (2022). Entretien personnel. 8 déc. 2022.
- (2023a). Entretien personnel. 5 mai 2023.
- (2023b). Entretien personnel. 18 déc. 2023.
- SHARIR, Or, Barak PELEG et Yoav SHOHAM (2020). « The Cost of Training NLP Models: A Concise Overview ». Prépublication *arXiv*. ARXIV : [2004.08900](https://arxiv.org/abs/2004.08900).
- SHAVIT, Zohar (1986). *Poetics of Children's Literature*. Athens (GA), University of Georgia Press.
- SHAW, Peter, Jakob USZKOREIT et Ashish VASWANI (2018). « Self-Attention with Relative Position Representations ». *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 2, édité par Marilyn WALKER, Heng Ji et Amanda STENT, La Nouvelle-Orléans (LA), ACL, pages 464–468. DOI : [10.18653/v1/N18-2074](https://doi.org/10.18653/v1/N18-2074).

- SHIH, Chung-ling (2016). « Can Machine Translation Declare a New Realm of Service? Online Folktales as a Case Study ». *Theory and Practice in Language Studies*, vol. 6, n° 2, pages 252–259. DOI : [10.17507/tpls.0602.05](https://doi.org/10.17507/tpls.0602.05).
- SLESSOR, Stephen (2020). « Tenacious Technophobes or Nascent Technophiles? A Survey of the Technological Practices and Needs of Literary Translators ». *Perspectives*, vol. 28, n° 2, pages 238–252. DOI : [10.1080/0907676X.2019.1645189](https://doi.org/10.1080/0907676X.2019.1645189).
- SNE (2023). *Les Chiffres de l'édition du Syndicat national de l'édition*. Paris, Syndicat national de l'édition. URL : <https://www.sne.fr/publications-du-sne/les-chiffres-de-ledition-en-france-et-a-linternational-2022-2023/>.
- SOUSA, Cristina (2002). « TL versus SL Implied Reader: Assessing Receptivity when Translating Children's Literature ». *Meta*, vol. 47, n° 1, pages 16–29. DOI : [10.7202/007988ar](https://doi.org/10.7202/007988ar).
- STAA (2022). « Non à l'automatisation des métiers de l'art — 1 ». Tribune publiée sur le site du STAA CNT-SO, 11 nov. 2022. URL : <https://staa-cnt-so.org/2022/11/11/non-a-lautomatisation-des-metiers-de-lart/>.
- STANOJEVIĆ, Miloš, Amir KAMRAN, Philipp KOEHN et Ondřej BOJAR (2015). « Results of the WMT15 Metrics Shared Task ». *Proceedings of the Tenth Workshop on Statistical Machine Translation*, édité par Ondřej BOJAR, Rajan CHATTERJEE, Christian FEDERMANN, Barry HADDOW, Chris HOKAMP, Matthias HUCK, Varvara LOGACHEVA et Pavel PECINA, Lisbonne, ACL, pages 256–273. DOI : [10.18653/v1/W15-3031](https://doi.org/10.18653/v1/W15-3031).
- STRUBELL, Emma, Ananya GANESH et Andrew MCCALLUM (2019). « Energy and Policy Considerations for Deep Learning in NLP ». *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, édité par Anna KORHONEN, David TRAUM et Màrquez LLUÍS, Florence, ACL, pages 3645–3650. DOI : [10.18653/v1/P19-1355](https://doi.org/10.18653/v1/P19-1355).
- SUTSKEVER, Ilya, Oriol VINYALS et Quoc V. LE (2014). « Sequence to Sequence Learning with Neural Networks ». *Proceedings of the 27th International Conference on Neural Information Processing Systems*, vol. 2, édité par Zoubin GHAHRAMANI, Max WELING, Corinna CORTES, Neil LAWRENCE et Kilian WEINBERGER, Montréal, MIT Press, pages 3104–3112. URL : https://proceedings.neurips.cc/paper_files/paper/2014/hash/a14ac55a4f27472c5d894ec1c3c743d2-Abstract.html.
- SZYMAŃSKI, Piotr, Piotr ŻELASKO, Mikolaj MORZY, Adrian SZYMCZAK, Marzena ŻYŁA-HOPPE, Joanna BANASZCZAK, Lukasz AUGUSTYNIAK, Jan MIZGAJSKI et Yishay CARMIEL (2020). « WER we are and WER we think we are ». *Findings of the Association for Computational Linguistics*:

- EMNLP 2020, édité par Trevor COHN, Yulan HE et Yang LIU, ACL, pages 3290–3295. DOI : [10.18653/v1/2020.findings-emnlp.295](https://doi.org/10.18653/v1/2020.findings-emnlp.295).
- TABER, Charles R. (1972). « Traduire le sens, traduire le style ». *Langages*, n° 28, pages 55–63. DOI : [10.3406/lgge.1972.2098](https://doi.org/10.3406/lgge.1972.2098).
- TAIVALKOSKI-SHILOV, Kristiina (2019a). « Ethical Issues Regarding Machine(-Assisted) Translation of Literary Texts ». *Perspectives*, vol. 27, n° 5, pages 689–703. DOI : [10.1080/0907676X.2018.1520907](https://doi.org/10.1080/0907676X.2018.1520907).
- (2019b). « Free indirect discourse: an insurmountable challenge for literary MT systems? » *Proceedings of the Qualities of Literary Machine Translation*, édité par James HADLEY, Maja POPOVIĆ, Haithem AFLI et Andy WAY, Dublin, EAMT, pages 35–39. URL : <https://aclanthology.org/W19-7305>.
- TATMAN, Rachael (2017). « Gender and Dialect Bias in YouTube’s Automatic Captions ». *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, édité par Dirk HOVY, Shannon SPRUIT, Margaret MITCHELL, Emily M. BENDER, Michael STRUBE et Hanna WALLACH, Valence, ACL, pages 53–59. DOI : [10.18653/v1/W17-1606](https://doi.org/10.18653/v1/W17-1606).
- TAZELAAR, Frank (2020). « CELA: Connecting Emerging Literary Artists ». Présentation au séminaire *Creative Translation and Technologies Expert Meeting*, 29 mai 2020, en ligne.
- TEHRANIAN, Majid (1990). *Technologies of Power: Information Machines and Democratic Prospects*. New York (NY), Ablex.
- TEZCAN, Arda, Joke DAEMS et Lieve MACKEN (2019). « When a ‘Sport’ Is a Person and Other Issues for NMT of Novels ». *Proceedings of the Qualities of Literary Machine Translation*, édité par James HADLEY, Maja POPOVIĆ, Haithem AFLI et Andy WAY, Dublin, EAMT, pages 40–49. URL : <https://aclanthology.org/W19-7306/>.
- THAI, Katherine, Marzena KARPINSKA, Kalpesh KRISHNA, Bill RAY, Moira INGHILLERI, John WIETING et Mohit IYYER (2022). « Exploring Document-Level Literary Machine Translation with Parallel Paragraphs from World Literature ». *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, édité par Yoav GOLDBERG, Zornitsa KOZAREVA et ZHANG, Abou Dabi, ACL, pages 9882–9902. URL : <https://aclanthology.org/2022.emnlp-main.672>.
- THOMPSON, Neil C., Kristjan GREENEWALD, Keeheon LEE et Gabriel F. MANSO (2022). « The Computational Limits of Deep Learning ». Prépublication *arXiv*, 2^e éd. ARXIV : [2007.05558](https://arxiv.org/abs/2007.05558).
- TOMAŠEVSKIJ, Boris (2018). « La méthode formelle (en guise de nécrologie) ». Traduit par Catherine DEPRETTO. *Communications*, n° 103, pages 17–27. DOI : [10.3917/commu.103.0017](https://doi.org/10.3917/commu.103.0017).

- TORAL, Antonio (2019). « Post-editeese: an Exacerbated Translationese ». *Proceedings of Machine Translation Summit XVII*, vol. 1, édité par Mikel L. FORCADA, Andy WAY, Barry HADDOW et Rico SENNRICH, Dublin, EAMT, pages 273–281. URL : <https://aclanthology.org/W19-6627>.
- TORAL, Antonio, Sheila CASTILHO, Ke HU et Andy WAY (2018a). « Attaining the Unattainable? Reassessing Claims of Human Parity in Neural Machine Translation ». *Proceedings of the Third Conference on Machine Translation: Research Papers*, édité par Ondřej BOJAR, Rajen CHATTERJEE, Christian FEDERMANN, Mark FISHEL, Yvette GRAHAM, Barry HADDOW, Matthias HUCK, Antonio JIMENO YEPES, Philipp KOEHN, Christof MONZ, Matteo NEGRI, Aurélie NÉVÉOL, Mariana NEVES, Matt POST, Lucia SPECIA, Marco TURCHI et Karin VERSPOOR, Bruxelles, ACL, pages 113–123. DOI : [10.18653/v1/W18-6312](https://doi.org/10.18653/v1/W18-6312).
- TORAL, Antonio, Antoni OLIVER et Pau RIBAS-BELLESTÍN (2020). « Machine Translation of Novels in the Age of Transformer ». *Maschinelle Übersetzung für Übersetzungsprofis*, édité par Jörg PORSIEL, Berlin, BDÜ Fachverlag, pages 276–295. ARXIV : [2011.14979](https://arxiv.org/abs/2011.14979).
- TORAL, Antonio, Andreas VAN CRANENBURGH et Tia NUTTERS (2024). « Literary-Adapted Machine Translation in a Well-Resourced Language Pair: Explorations with More Data and Wider Contexts ». *Computer-Assisted Literary Translation*, édité par Andrew ROTHWELL, Andy WAY et Roy YOUDALE, Abingdon, Routledge, pages 27–52. DOI : [10.4324/9781003357391](https://doi.org/10.4324/9781003357391).
- TORAL, Antonio et Andy WAY (2015a). « Machine-Assisted Translation of Literary Text: A Case Study ». *Translation Spaces*, vol. 4, n° 2, pages 240–267. DOI : [10.1075/ts.4.2.04tor](https://doi.org/10.1075/ts.4.2.04tor).
- (2015b). « Translating Literary Text between Related Languages using SMT ». *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, édité par Anna FELDMAN, Anna KAZANTSEVA, Stan SZPAKOWICZ et Corina KOOLEN, Denver (CO), ACL, pages 123–132. DOI : [10.3115/v1/W15-0714](https://doi.org/10.3115/v1/W15-0714).
- (2018). « What Level of Quality can Neural Machine Translation Attain on Literary Text? » *Translation Quality Assessment: From Principles to Practice*, édité par Joss MOORKENS, Sheila CASTILHO, Federico GASPARI et Stephen DOHERTY, New York (NY), Springer, pages 263–287. ARXIV : [1801.04962](https://arxiv.org/abs/1801.04962).
- TORAL, Antonio, Martijn WIELING et Andy WAY (2018b). « Post-Editing Effort of a Novel With Statistical and Neural Machine Translation ». *Frontiers in Digital Humanities*, vol. 5, n° 9, pages 1–11. DOI : [10.3389/fdigh.2018.00009](https://doi.org/10.3389/fdigh.2018.00009).
- TOURY, Gideon (1995). *Descriptive Translation Studies and Beyond*. Amsterdam, John Benjamins. DOI : [10.1075/bt1.4](https://doi.org/10.1075/bt1.4).

- TOUVRON, Hugo, Louis MARTIN, Kevin STONE, Peter ALBERT, Amjad ALMAHAIRI, Yasmine BABAEI, Nikolay BASHLYKOV, Soumya BATRA, Prajjwal BHARGAVA, Shruti BHOSALE, Dan BIKEL, Lukas BLECHER, Cristian Ferrer CANTON, Moya CHEN, Guillem CUCURULL, David ESIOBU, Jude FERNANDES, Jeremy FU, Wenyin FU, Brian FULLER, Cynthia GAO, Vedanuj GOSWAMI, Naman GOYAL, Anthony HARTSHORN, Saghar HOSSEINI, Rui HOU, Hakan INAN, Marcin KARDAS, Viktor KERKEZ, Madian KHABSA, Isabel KLOUMANN, Artem KORENEV, Punit Singh KOURA, Marie-Anne LACHAUX, Thibaut LAVRIL, Jenya LEE, Diana LISKOVICH, Yinghai LU, Yuning MAO, Xavier MARTINET, Todor MIHAYLOV, Pushkar MISHRA, Igor MOLYBOG, Yixin NIE, Andrew POULTON, Jeremy REIZENSTEIN, Rashi RUNGTA, Kalyan SALADI, Alan SCHELLEN, Ruan SILVA, Eric Smith MICHAEL, Ranjan SUBRAMANIAN, Xiaoqing Ellen TAN, Binh TANG, Ross TAYLOR, Adina WILLIAMS, Jian Xiang KUAN, Puxin XU, Zheng YAN, Iliyan ZAROV, Yuchen ZHANG, Angela FAN, Melanie KAMBADUR, Sharan NARANG, Aurelien RODRIGUEZ, Robert STOJNIC, Sergey EDUNOV et Thomas SCIALOM (2023). « Llama 2: Open Foundation and Fine-Tuned Chat Models ». Prépublication *arXiv*, 2^e éd. ARXIV : [2307.09288v2](https://arxiv.org/abs/2307.09288v2).
- TUBARO, Paola, Antonio CASILLI et Marion COVILLE (2020). « The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence ». *Big Data & Society*, vol. 7, n° 1, pages 1–12. DOI : [10.1177/2053951720919776](https://doi.org/10.1177/2053951720919776).
- UTUSHKINA, Anastasia (2023). « Traduction automatique neuronale de la littérature : Évaluation de la traduction automatique neuronale de textes littéraires de l'italien vers le français ». Mémoire de master, Université de Genève. URL : <https://archive-ouverte.unige.ch/unige:167921>.
- VAN COILLIE, Jan (2006). « Character Names in Translation: A Functional Approach ». *Children's Literature in Translation: Challenges and Strategies*, édité par Jan VAN COILLIE et Walter P. VERSCHUEREN, Abingdon, Routledge, pages 123–139. DOI : [10.4324/9781315759845](https://doi.org/10.4324/9781315759845).
- VAN COILLIE, Jan et Jack McMARTIN, éd. (2020). *Children's Literature in Translation: Texts and Contexts*. Louvain, Leuven University Press. DOI : [10.11116/9789461663207](https://doi.org/10.11116/9789461663207).
- VAN COILLIE, Jan et Walter P. VERSCHUEREN, éd. (2006). *Children's Literature in Translation: Challenges and Strategies*. Abingdon, Routledge. DOI : [10.4324/9781315759845](https://doi.org/10.4324/9781315759845).
- VAN DE CRUYS, Tim (2020). « Automatic Poetry Generation from Prosaic Text ». *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, édité par Dan JURAFSKY, Joyce CHAI, Natalie SCHLUTER et Joel TETREAU, Seattle (WA), ACL, pages 2471–2480. DOI : [10.18653/v1/2020.acl-main.223](https://doi.org/10.18653/v1/2020.acl-main.223).

- VAN DER BEEK, Leonoor et Antal VAN DEN BOSCH (2015). « Translation technology in the Netherlands and Belgium ». *The Routledge Encyclopedia of Translation Technology*, édité par Chan SIN-WAI, Abingdon, Routledge, pages 352–363.
- VANMASSENHOVE, Eva, Dimitar SHTERIONOV et Matthew GWILLIAM (2021). « Machine Translationese: Effects of Algorithmic Bias on Linguistic Complexity in Machine Translation ». *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, édité par Paola MERLO, Jörg TIEDEMANN et Reut TSARFATY, Kiev, ACL, pages 2203–2213. DOI : [10.18653/v1/2021.eacl-main.188](https://doi.org/10.18653/v1/2021.eacl-main.188).
- VANMASSENHOVE, Eva, Dimitar SHTERIONOV et Andy WAY (2019). « Lost in Translation: Loss and Decay of Linguistic Richness in Machine Translation ». *Proceedings of Machine Translation Summit XVII*, vol. 1, édité par Mikel L. FORCADA, Andy WAY, Barry HADDOW et Rico SENNRICH, Dublin, EAMT, pages 222–232. URL : <https://aclanthology.org/W19-6622>.
- VASWANI, Ashish, Noam SHAZEER, Niki PARMAR, Jakob USZKOREIT, Llion JONES, Aidan N. GOMEZ, Lukasz KAISER et Illia POLOSUKHIN (2017). « Attention is All you Need ». *Proceedings of the 31st International Conference on Neural Information Processing Systems*, édité par Ulrike von LUXBURG, Isabelle GUYON, Samy BENGIO, Hanna WALLACH et Rob FERGUS, Long Beach (CA), Curran Associates Inc., pages 6000–6010. URL : <https://papers.nips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>.
- VENUTI, Lawrence (2008). *The Translator's Invisibility: A History of Translation*. 2^e édition, Abingdon, Routledge.
- (2017). *Teaching Translation*. Abingdon, Routledge.
- VIEIRA, Lucas Nunes (2020a). « Automation anxiety and translators ». *Translation Studies*, vol. 13, n^o 1, pages 1–21. DOI : [10.1080/14781700.2018.1543613](https://doi.org/10.1080/14781700.2018.1543613).
- (2020b). « Machine translation in the news: A framing analysis of the written press ». *Translation Spaces*, vol. 9, n^o 1, pages 98–122. DOI : [10.1075/ts.00023.nun](https://doi.org/10.1075/ts.00023.nun).
- (2023). « The many guises of machine translation: A postphenomenology perspective ». *Digital Translation*, vol. 10, n^o 1, pages 16–36. DOI : [10.1075/dt.00002.nun](https://doi.org/10.1075/dt.00002.nun).
- VIEIRA, Lucas Nunes et Lucia SPECIA (2011). « A Review of Translation Tools from a Post-Editing Perspective ». *Proceedings of the Third Joint EM+/CNGL Workshop*, édité par Ventsislav ZHECHEV, Luxembourg, AMTA, pages 33–42. URL : <https://aclanthology.org/www.mt-archive.info/JEC-2011-Vieira.pdf>.
- VIEIRA, Lucas Nunes, Natalie ZELENKA, Roy YOUNDALE, Xiaochun ZHANG et Michael CARL (2023). « Translating science fiction in a CAT tool: machine translation and segmentation settings ».

Translation & Interpreting, vol. 15, n° 1, pages 216–235. DOI : [10.12807/ti.115201.2023.a11](https://doi.org/10.12807/ti.115201.2023.a11).

- VILAR, David, Gregor LEUSCH, Hermann NEY et Rafael E. BANCHS (2007). « Human Evaluation of Machine Translation Through Binary System Comparisons ». *Proceedings of the Second Workshop on Statistical Machine Translation*, édité par Chris CALLISON-BURCH, Philipp KOEHN, Cameron Shaw FORDYCE et Christof MONZ, Prague, ACL, pages 96–103. URL : <https://aclanthology.org/W07-0713>.
- VILAR, David, Jia XU, Luis Fernando D'HARO et Hermann NEY (2006). « Error Analysis of Statistical Machine Translation Output ». *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, édité par Nicoletta CALZOLARI, Khalid CHOUKRI, Aldo GANGEMI, Bente MAEGAARD, Joseph MARIANI, Jan ODIJK et Daniel TAPIAS, Gênes, ELRA, pages 697–702. URL : http://www.lrec-conf.org/proceedings/lrec2006/pdf/413_pdf.pdf.
- VINAY, Jean-Paul et Jean DARBELNET (1958). *Stylistique comparée du français et de l'anglais : Méthode de traduction*. Paris, Didier.
- VOIGT, Rob et Dan JURAFSKY (2012). « Towards a Literary Machine Translation: The Role of Referential Cohesion ». *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature*, édité par David ELSON, Anna KAZANTSEVA, Rada MIHALCEA et Stan SZPAKOWICZ, Montréal, ACL, pages 18–25. URL : <https://aclanthology.org/W12-2503>.
- WANG, Ning (2023). « Introduction: Literary translation in the age of artificial intelligence ». *Babel*, vol. 69, n° 4 : *Literary Translation in the Era of Artificial Intelligence: Challenges and Its Future Prospects*, édité par Ning WANG et Hongtao WANG, pages 437–446. DOI : [10.1075/babel.00327.wan](https://doi.org/10.1075/babel.00327.wan).
- WANG, Yue, Cuong HOANG et Marcello FEDERICO (2021). « Towards Modeling the Style of Translators in Neural Machine Translation ». *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, édité par Kristina TOUTANOVA, Anna RUMSHISKY, Luke ZETTLEMOYER, Dilek HAKKANI-TUR, Iz BELTAGY, Steven BETHARD, Ryan COTTERELL, Tanmoy CHAKRABORTY et Yichao ZHOU, Mexico, ACL, pages 1193–1199. DOI : [10.18653/v1/2021.naacl-main.94](https://doi.org/10.18653/v1/2021.naacl-main.94).
- WAY, Andy (2013). « Traditional and Emerging Use-Cases for Machine Translation ». *Proceedings of Translating and the Computer 35*, Londres, Aslib, pages 1–12. URL : <https://aclanthology.org/2013.tc-1.12>.

- (2020). « Machine translation: Where are we at today? » *The Bloomsbury Companion to Language Industry Studies*, édité par Erik ANGELONE, Maureen EHRENSBERGER-DOW et Gary MASSEY, Londres, Bloomsbury, pages 311–332.
- WAY, Andy et Mary HEARNE (2011). « On the Role of Translations in State-of-the-Art Statistical Machine Translation ». *Language and Linguistics Compass*, vol. 5, n° 5, pages 227–248. DOI : [10.1111/j.1749-818X.2011.00275.x](https://doi.org/10.1111/j.1749-818X.2011.00275.x).
- WEBSTER, Rebecca, Margot FONTEYNE, Arda TEZCAN, Lieve MACKEN et Joke DAEMS (2020). « Gutenberg Goes Neural: Comparing Features of Dutch Human Translations with Raw Neural Machine Translation Outputs in a Corpus of English Literary Classics ». *Informatics*, vol. 7, n° 3, pages 1–21. DOI : [10.3390/informatics7030032](https://doi.org/10.3390/informatics7030032).
- WEST, Mark, Rebecca KRAUT et Han Ei CHEW (2020). *Je rougirais si je pouvais : Réduire la fracture numérique entre les genres par l'éducation*. UNESCO. URL : https://unesdoc.unesco.org/ark:/48223/pf0000367416_fre.
- WINTERS, Marion et Dorothy KENNY (2024). « Mark my Keywords: A Translator-Specific Exploration of Style in Literary Machine Translation ». *Computer-Assisted Literary Translation*, édité par Andrew ROTHWELL, Andy WAY et Roy YOUNDALE, Abingdon, Routledge, pages 69–87. DOI : [10.4324/9781003357391](https://doi.org/10.4324/9781003357391).
- YAMADA, Masaru (2019). « The impact of Google Neural Machine Translation on Post-editing by student translators ». *JoSTrans*, n° 31 : *Post-Editing in Practice: Process, Product and Networks*, édité par Lucas Nunes VIEIRA, Elisa ALONSO et Lindsay BYWOOD, pages 87–106. URL : https://jostrans.soap2.ch/issue31/art_yamada.php.
- YIRMİBEŞOĞLU, Zeynep, Olgun DURSUN, Harun DALLI, Mehmet ŞAHİN, Sabri HODZİK Ena Gürses et Tunga GÜNGÖR (2023). « Incorporating Human Translator Style into English-Turkish Literary Machine Translation ». *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, édité par Mary NURMINEN, Judith BRENNER, Maarit KOPONEN, Sirkku LATOMAA, Mikhail MIKHAILOV, Frederike SCHIERL, Tharindu RANASINGHE, Eva VANMASSENHOVE, Sergi Alvarez VIDAL, Nora ARANBERRI, Mara NUNZIATINI, Carla Parra ESCARTÍN, Mikel FORCADA, Maja POPOVIC, Carolina SCARTON et Helena MONIZ, Tampere, EAMT, pages 419–428. URL : <https://aclanthology.org/2023.eamt-1.40>.
- YOUNDALE, Roy (2020). *Using Computers in the Translation of Literary Style: Challenges and Opportunities*. New York (NY), Routledge.
- YOUNDALE, Roy et Andrew ROTHWELL (2022). « Computer-assisted translation (CAT) tools, translation memory, and literary translation ». *The Routledge Handbook of Translation and Memory*,

- édité par Sharon DEANE-COX et Anneleen SPIESSENS, Abingdon, Routledge, pages 381–402.
DOI : [10.4324/9781003273417](https://doi.org/10.4324/9781003273417).
- YVON, François (2019). « Les deux voies de la traduction automatique ». *Hermès*, vol. 85, n° 3, pages 62–68. DOI : [10.3917/herm.085.0062](https://doi.org/10.3917/herm.085.0062).
- ZAJDEL, Alicja (2022). « Catching the meaning of words Can Google Translate convey metaphor? » *Using Technologies for Creative-Text Translation*, édité par James HADLEY, Kristiina TAIVALKOSKI-SHILOV, Carlos S. C. TEIXEIRA et Antonio TORAL, New York (NY), Routledge, pages 116–138.
DOI : [10.4324/9781003094159-6](https://doi.org/10.4324/9781003094159-6).
- ZEMAN, Daniel, Milan STRAKA, Jan HAJIČ, Joakim NIVRE, Filip GINTER, Juhani LUOTOLAHTI, Sampo PYYSALO, Slav PETROV, Martin POTTHAST, Francis TYERS, Elena BADMAEVA, Memduh GOKIRMAK, Anna NEDOLUZHKO, Silvie CINKOVÁ, Jan HAJIČ JR., Jaroslava HLAVÁČOVÁ, Václava KETTNEROVÁ, Zdeňka UREŠOVÁ, Jenna KANERVA, Stina OJALA, Anna MISSILÄ, Christopher D. MANNING, Sebastian SCHUSTER, Siva REDDY, Dima TAJI, Nizar HABASH, Herman LEUNG, Marie-Catherine de MARNEFFE, Manuela SANGUINETTI, Maria SIMI, Hiroshi KANAYAMA, Valeria de PAIVA, Kira DROGANOVA, Héctor MARTÍNEZ ALONSO, Çağrı ÇÖLTEKIN, Umut SULUBACAK, Hans USZKOREIT, Vivien MACKETANZ, Aljoscha BURCHARDT, Kim HARRIS, Katrin MARHEINECKE, Georg REHM, Tolga KAYADELEN, Mohammed ATTIA, Ali ELKAHKY, Zhuoran YU, Emily PITLER, Saran LERTPRADIT, Michael MANDL, Jesse KIRCHNER, Hector Fernandez ALCALDE, Jana STRNADOVÁ, Esha BANERJEE, Ruli MANURUNG, Antonio STELLA, Atsuko SHIMADA, Sookyoung KWAK, Gustavo MENDONÇA, Tatiana LANDO, Rattima NITISAROJ et Josie LI (2017). « CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies ». *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, édité par Jan HAJIČ et Dan ZEMAN, Vancouver, ACL, pages 1–19.
DOI : [10.18653/v1/K17-3001](https://doi.org/10.18653/v1/K17-3001).
- ZHANG, Aston, Zachary C. LIPTON, Mu LI et Alexander J. SMOLA (2023). *Dive into Deep Learning*. Cambridge (MA), Cambridge University Press. URL : <https://D2L.ai>.
- ZHANG, Shuyin (2024). « Machine Translation of Chinese Fantasy (Xianxia) Novels: An Investigation Into The Leading Websites Translating Chinese Internet Literature Into English ». *Computer-Assisted Literary Translation*, édité par Andrew ROTHWELL, Andy WAY et Roy YOUNDALE, Abingdon, Routledge, pages 142–157. DOI : [10.4324/9781003357391](https://doi.org/10.4324/9781003357391).
- ZHANG, Susan, Stephen ROLLER, Naman GOYAL, Mikel ARTETXE, Moya CHEN, Shuohui CHEN, Christopher DEWAN, Mona DIAB, Xian LI, Xi Victoria LIN, Todor MIHAYLOV, Myle OTT, Sam SHLEIFER, Kurt SHUSTER, Daniel SIMIG, Punit Singh KOURA, Anjali SRIDHAR, Tianlu WANG

- et Luke ZETTLEMOYER (2022). « OPT: Open Pre-trained Transformer Language Models ». Prépublication *arXiv*, 4^e éd. ARXIV : [2205.01068v4](https://arxiv.org/abs/2205.01068v4).
- ZHAO, Jianli et Hyo Jong LEE (2022). « Automatic Generation and Evaluation of Chinese Classical Poetry with Attention-Based Deep Neural Network ». *Applied Sciences*, vol. 12, n° 13, pages 1–15. DOI : [10.3390/app12136497](https://doi.org/10.3390/app12136497).
- ZYDRON, Andrzej (2019). « De-Demonizing AI Regarding Localization ». *Proceedings of the 41st Conference Translating and the Computer*, édité par Juliet Margaret ESTEVES-FERREIRA Macan, Ruslan MITKOV et Olaf-Michael STEFANOV, Londres, Éditions Tradulex, pages 137–144. URL : https://www.asling.org/tc41/wp-content/uploads/TC41-Proceedings_137-144.pdf.

Notes

- [1] Technology is not an option in today's professional world; it is a necessity. Years ago one talked about Computer-Aided Translation (CAT). That now seems a redundancy. Virtually all translating is aided by computers. Further, the most revolutionary tools are quite probably the everyday ones that are not specific to translation: Internet search engines, spell checkers, search and replace functions, and revision tools have had a huge impact on all forms of written communication.
- [2] Why do literary translators, and their trainers, often give the impression that they are at best ambivalent about the use of technology in literary translation, and at worst simply antagonistic?
- [3] Reading MT output in this way means redirecting attention to what translation can do despite all the ways it is imagined to fail.
- [4] By characterising translation as HCI, I do not intend to dehumanise translation or translators, nor do I intend to humanise the computer. Equally, I do not intend to deprofessionalise translation, but rather wish to show that the profession has changed over time and has become almost symbiotic with the 'machine' (used synonymously here with 'computer') and to demonstrate that there are benefits and significant challenges emerging from this development, which ought to be the focus of more discussion and research within the translation community.
- [5] Unfortunately, there is little evidence to suggest that tools that are proposed as aids to the translation process have been designed from the point of view of the humans who have to use them.
- [6] It does, however, locate itself within the framework of an open invitation to the scholarly translation community and the practitioners of literary translation to engage in a wider debate about how technology might productively be integrated into the practice of literary translation, a debate which would also include uses of MT and CAT tools in this context.
- [7] Note also that literary translation is often selected by human translators not overly well-disposed to MT to demonstrate how useless it is for anything; on any randomly selected online translators' forum, you don't have to look too hard to find someone who has selected a section from a book and shown how MT messes up the translation.
- [8] The long-running jocular answer in machine translation to the question, "When will we reach fully automated high quality machine translation?" is five years. Similarly, due to artificial intelligence, half of the current jobs are thought

to become obsolete any day now. This is a great pitch for research funding and venture capital investment, but eventually five years pass, and broken promises do not make for a good long-term strategy.

[9] Probably no other trait so characterizes technobabble as anthropomorphism and personification, attributing human characteristics to hardware and software [...]

Behind this penchant for personification is the long standing tradition in the industry, dating from the inception of computing, of employing terms that suggest that computers are living entities. The term *memory* as applied to the storage of data goes back to John Von Neumann and the concept of stored programs. Almost since their advent, computers have been called “electronic brains” or “computer brains.” Today it is the CPU, rather than the entire machine, that is thought of as the brain. Terms such as *intelligence*, *logic*, and *virus* further promote the computer as human analogy.

Anyone who has used a computer regularly knows how easy it is to be seduced into believing that the machine is thinking—and doing it faster than the person.

[10] The reality is much more prosaic: most of the mathematical basis of what is termed AI is not complicated and generally rooted in early 18th century mathematic [...]

[11] It is also behind many artistic uses of machine translation, where outputs are treated as a kind of ‘found poetry’, serendipitously encountered linguistic ‘stuff’ that can be manipulated like materials in the plastic arts. In some cases, machine translation is prized for its jarring, ungrammatical, nonsensical, antinarrative character, and is seen as capable of delivering the kind of non-communicative, foreignizing translation advocated by Benjamin (1923/2012) and Venuti (1995b/2008).

[12] Literature is in many ways distinct from genres usually considered in computational linguistics, such as newspaper prose, unstructured Web pages or speech. That is why the growing availability of online literature presents new opportunities and challenges in language processing. How can automatic methods help readers find new literature on a certain topic, understand a text or a genre, identify the author of an anonymous text, or read a book written in another language? How might language analysis techniques go beyond words to help identify the deeper meaning found in literature, no matter the time, place or culture from which it originates?

[13] SMT thus relies on human translation for its training data, but it also relies on human translation for its legitimacy: the reason developers of SMT systems use parallel corpora to train their systems is because such corpora are assumed to contain good answers to translation problems; and they are assumed to contain good answers precisely because they contain translations performed by human beings.

[14] Like human translators, a statistical machine translation (SMT) system may produce translations that are relatively free or faithful and must constantly make translation choices in decoding. For SMT, choices are dependent on what is observed in training and language modeling data and their frequencies. When systems are trained on datasets that are similar to a test text, they are more likely to make reasonable translation choices. Additionally, if a model, either a priori or automatically, knows something about what the output should look like (e.g. poetry should rhyme or have rhythm), features could encourage free translations to take a certain form.

- [15] Neural MT systems, like SMT systems, learn to translate from parallel corpora, but do so using very different computational methods.
- [16] As Forcada (2017) explains, the activation states of individual neurons do not make much sense by themselves. It is, instead, the activation states of large sets of connected neurons that can be understood as representing individual words and their relationships with other words. The trick in training an NMT system is to learn precisely those weights that will result in the best performing model of translation, that is, the model whose activation states allow it to predict the best translations. So how is this done? Like in all machine learning, the system learns from data. A neural model of translation is built step by step by exposing a learning algorithm to vast quantities of parallel data. In successive passes, the algorithm learns weights and keeps adjusting those weights, so that the predictions of the model it builds get closer and closer to a desired “correct” answer.
- [17] You shall know a word by the company it keeps!
- [18] The term “domain” has a wide interpretation in the MT literature and may refer to topic, dialect, genre or style (Chen et al., 2013). However, to the best of our knowledge, MT domain adaptation does not extend to consider demographic or personality traits of the author.
- [19] We need to find out not only what happens in a translator’s mind, but also what happens elsewhere, e.g. in their hands, and their computers, on their desk, in their languages or in their dialogues. Translation is not done solely by the mind, but by complex systems. These systems include people, the specific social and physical environments and all their cultural artefacts.
- [20] If translation scholars take the other players (authors, readers, clients) in the translation process properly into account, they will be in a position to explain how professional translators are able to develop their flexibility, creativity and sense of responsibility without losing touch with the real world.
- [21] This has been the source of the common misunderstanding that, compared to adult literature, children’s books are characterized by a simpler narrative structure, poorer language, and inferior artistic resources.
- [22] As Perry Nodelman points out, it is ridiculous to analyse Proust and Edith Nesbit with the same instruments, but not because Edith Nesbit is a worse writer than Proust. She is “worse than Proust” only in the same sense that the Japanese artist Hokusai, who painted 36 landscapes of Mount Fuji, is “worse” than, for instance, van Gogh. They are phenomena from different categories. They cannot be compared.
- [23] We show that under some circumstances an improvement in Bleu is *not sufficient* to reflect a genuine improvement in translation quality, and in other circumstances that it is *not necessary* to improve Bleu in order to achieve a noticeable improvement in translation quality.
- [24] The closer a machine translation is to a professional human translation, the better it is.
- [25] The latter category of pretrained methods consists of metrics that use pretrained neural models to evaluate the quality of MT output texts given the source sentence, the human reference, or both. They are not strictly dependent on

the translation quality of the human reference (for example, they can better evaluate synonyms or paraphrases). However, their performance is influenced by the data on which they have been trained. Moreover, the pretrained models introduce a black-box problem where it is difficult to diagnose potential unexpected behavior of the metric, such as various biases learned from training data.

- [26] Many variations of the sentence preserve the meaning while other superficially very small changes can distort or completely reverse it.
- [27] The strongest and most meaningful evaluation is a real-world outcome-based evaluation, where a system is operationally deployed and we measure its impact on real-world outcomes.
- [28] Translator Studies, as I understand it, is the antithesis to the mechanisation of Translation Studies. By this, I do not mean the technical dimension in the form of translation tools, machine translation and the like, but the view of the translator as a functional element, as a functioning unit, instead of looking at the human being in its uniqueness.
- [29] I hope that the approach suggested in this book will offer a good collection of unexpected insights—precisely because it proceeds from a personal, phenomenological, processual, existential, and somatic perspective (which is how I would describe my methodological commitments).
- [30] This notion of embeddedness is consistent with conceptualizations of translation as a situated activity and can help us understand the role of tools and context in the translation process and their impact on the quality of the product (Risku, 2010). It is also congruent with the description of translation in the emerging field of cognitive translatology as an enactment of situated, embodied, and distributed cognition (Muñoz, 2010). Adopting such a position seriously means that the physical, cognitive, and organizational factors that impinge on translators as they do their work and make their decisions have to be taken into consideration in order to truly understand the activity.
- [31] Situated within a CTS framework, the book combines methods from this research paradigm (think-aloud, draft versions, and video recordings) with ethnographic techniques (observations and interviews) to provide thorough insights into the practice of an experienced literary translator working *in situ*.
- [32] The present study brings an original contribution to TS by investigating in-depth the process adopted by an experienced translator to self-revise a full literary text. This entails a different research design from the mainstream method adopted so far: a sole research participant and a long, full text, instead of short texts and a number of research participants. This approach could help us gain new insights into translator behaviour and, if such studies proliferate in the future, a wealth of knowledge on translator behaviour could be accumulated and comparisons between similar studies could then be drawn.
- [33] When you revise a pre-existing translation, and the machine translation is a pre-existing translation and not a rough draft of your own making, you are faced with a dilemma not encountered when embarking on a fresh translation that is not machine-produced. It is a dilemma known to every editor proper: how to respect both the original author's and the translator's voice? You do not have to "respect" the machine's "efforts" but you have to set to work against two backdrops at once: the original and the pre-existing translation, each posing its own difficulties.

- [34] Among the attitudinal factors that seem to go together with success are personal involvement with the task, the translator's confidence in his own ability as well as professional pride and willingness to take responsibility for the communicative event.
- [35] Look, now, at what happens when we use workstations that integrate translation memories and machine translation. The first thing you find is that the text is segmented, broken into units that sit one on top of the other. That is, the text is broken into paradigmatic form; its linearity is repeatedly interrupted. The translating mind is thereby invited to work on one segment after the other, checking for terminological and phraseological consistency but not so easily checking, within this environment, for syntagmatic cohesion.
- [36] Importantly, as we allude to in Section 2, the power of defaults should not be underestimated. For literary translators who are new to CAT, we emphasise the importance of experimenting with different segmentation settings as this may save time and effort. CAT tools were not designed with literary translators in mind, so deviating from the norm in how these tools are configured may be even more important for this text type.
- [37] In other words, it means paying appropriate attention to the 'actor-networks' in which complex work activities are negotiated.
- [38] Technology is neither good nor bad; nor is it neutral.
- [39] Specialised translators who had invested in TM tools expecting to gain a competitive advantage found instead that expected throughput had increased as word rates decreased. Many translators began to resent new technologies that they considered to be imposed on them (Cadwell et al. 2017, Christensen and Schjoldager 2016, Way 2018): first TM with its associated discounts, and more recently MT post-editing, which requires them to accept further discounted rates to fix "fundamental linguistic errors that a trained human translator would rarely generate" (O'Brien 2012: 110). [...] The gradual limitation of the translator's role has undermined their ability to conform to the ethical code of their profession (Chesterman 2001: 142) by reducing the translation process to a series of "language-replacement exercises" (Pym 2003: 7). In this way, rather than only eliminating work that is "mechanical and routine" (Kay 1980: 1, LeBlanc 2013), translation fits with the narrative developed by Frischmann and Selinger (2018), in which the encroachment of machine learning in many occupations risks reducing the role of the human to one that is comparable with a simple machine. Furthermore, as the profession has moved from analogue to digital, translators' powerlessness is reflected in continued data dispossession, common for many knowledge workers, and largely unaffected by legal constraints (Huws 2014: 4).
- [40] Literary translators might even benefit from the change, provided that all players in the field, especially the powerful ones, such as major publishing houses, incorporate sustainable development as part of their view on translation quality.
- [41] Those of us who have seen many paradigms come and go know that overgilding the lily does none of us any good, especially those of us who have been trying to build bridges between MT developers and the translation community for many years. The human-in-the-loop will always remain the most important link in the chain, at least where dissemination of translated material is required. All that MT system developers are trying to do is improve the output from their systems to make technology-savvy translators more productive. MT systems are unlikely ever to 'bridge the

gap' or achieve human parity with human quality translation. Just because a new paradigm is in vogue, it does not mean that MT has become easy, or a solved problem, as some would like to make out.

[42] This is because, while the sentence-aligned corpus could be seen to be literally in violation of the copyright of many pages of content, there would be very little incentive on the part of their authors to sue the compilers of the corpora, as it would be very difficult to prove the existence of damage and even more difficult to put a value on it.

[43] At the third level, translation becomes about questions, including philosophical ones. Questions not just about the meaning or the most appropriate translation of a certain text, but about the sense of the entire endeavor, about the ethical aspects of working on this particular project or for that particular client, about the content we are dealing with and its effect on us, whether inspiring or disturbing.

[44] At the current rate of progress, the approach of MT to FAHQMT is asymptotic, as the dominant paradigms focus on word- or string-level signal coding and decoding, whereas human communication interprets cultural and contextual information to reach mutual understanding.

[45] The difference between having a negative attitude to technology and having a negative attitude to the perceived repercussions of technology is a small but important one. In the first case, translation technologies would need to be replaced, improved or eradicated for any problems to be solved, whereas in the second case finding solutions may be a matter of changing practices rather than technologies. These possibilities are not mutually exclusive; improving translators' experience is likely to require changes to technology and to practices.

[46] Ultimately, the solutions identified are the promotion of (1) a fruitful collaboration between literary translators and outsiders where tools are not imposed, (2) attention to working conditions, and (3) support of the literary translation workflow in enhancing rather than hindering ways.



Résumé

Évaluation experte d'un prototype d'aide à la traduction créative : La traduction littéraire automatisée individualisée au regard de ses enjeux traductologiques, éthiques et sociétaux

Si la traduction automatique a traditionnellement été perçue en littérature comme une technologie avant-gardiste digne des créations oulipiennes, l'arrivée plus récente des outils de traduction neuronaux a ouvert un débat nouveau, bien plus polarisé qu'auparavant. Pour poser un regard critique et objectif sur cette question, nous avons donc entrepris de mettre au point un système spécifiquement entraîné sur des textes littéraires. Si nous voyons qu'il est bien possible d'adapter de tels systèmes à ce domaine particulier, malgré les limites évidentes dont ces modèles font encore preuve, notre expérience sur la prose de fiction montre qu'il est aussi possible de les adapter au style individuel d'un traducteur ou d'une traductrice. Ces conclusions soulèvent évidemment de nombreuses questions liées à l'ergonomie des technologies de la traduction, ainsi qu'aux questions éthiques et aux enjeux sociétaux qui les entourent, mais elles indiquent surtout qu'il est possible de réimaginer ces outils et de les repenser d'une manière qui serve plus directement l'interaction humain-machine, le processus créatif et la qualité des textes traduits.

Abstract

Expert evaluation of a computer-aided creative translation scenario: A translation studies perspective on human-adapted literary machine translation and its ethical and societal implications

Although machine translation has traditionally been perceived in the literary domain as a sort of *avant-garde* technology on a par with Oulipian techniques, the recent arrival of neural translation tools seems to have given rise to an entirely new and deeply polarized debate. In order to take a critical and objective look on the matter, we therefore attempt to build a system specifically trained on literature. While our results show that it is indeed possible to train such a system for this particular domain, despite the still evident limitations of these models, our experience on prose fiction also demonstrates that they can be tailored to the individual style of a translator. Our conclusions obviously raise newfound questions regarding the ergonomic aspects of translation technology, the ethical concerns that it raises or the societal issues that have to be addressed. But, more importantly, they indicate that it is possible to reimagine such tools and rethink their use in a way that better serves the human-machine interaction, the creative process and the quality of translated texts.

