



HAL
open science

Le traducteur automatique comme outil du traducteur indépendant spécialisé en médecine

Magali Vidrequin

► **To cite this version:**

Magali Vidrequin. Le traducteur automatique comme outil du traducteur indépendant spécialisé en médecine. 2022. hal-03583630v1

HAL Id: hal-03583630

<https://hal.science/hal-03583630v1>

Preprint submitted on 22 Feb 2022 (v1), last revised 6 Jan 2023 (v3)

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

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

Le traducteur automatique comme outil du traducteur indépendant spécialisé en médecine

Magali Vidrequin¹

1 Université Rennes 2, France

*Corresponding author: magali.vidrequin@univ-rennes2.fr

Abstract

L'évolution et le progrès de la traduction automatique ont entraîné une nouvelle pratique sur le marché de la traduction : la post-édition est aujourd'hui une nouvelle tâche. Cet article est consacré à l'analyse d'une enquête réalisée auprès de traducteurs indépendants intervenant dans le domaine médical, afin de déterminer les différents profils et pratiques des professionnels en matière de post-édition, et d'observer l'intégration de la traduction automatique dans leur environnement de traduction.

keywords

traduction automatique, post-édition, indépendants, traduction médicale

INTRODUCTION

Depuis 2014, la traduction automatique a fait un important bond en avant en matière de qualité grâce aux réseaux neuronaux [Loock, 2018]. Cette qualité peut cependant varier d'un domaine à l'autre et d'un type de document à un autre. Par ailleurs, si les recherches en la matière sont très nombreuses, l'intérêt accordé plus spécifiquement à l'utilisation de la traduction automatique par les traducteurs indépendants n'est que très récent. Il est en effet plus difficile pour un professionnel indépendant d'intégrer la traduction automatique dans sa pratique, premièrement pour des raisons technologiques : la plupart des moteurs de traduction automatique employés en traduction professionnelle exigent des compétences approfondies en informatique : les résultats favorables obtenus dans certains domaines par ces moteurs [Castilho et Al, 2017] nécessitent un entraînement du moteur, réalisé le plus souvent par des ingénieurs ou des informaticiens. En deuxième lieu, viennent les contraintes financières : si désormais plusieurs moteurs – généraux - sont techniquement accessibles au grand public, sur navigateur internet, les fonctionnalités proposées gratuitement sont souvent limitées pour des professionnels (travaillant avec la TAO et sur des volumes de traduction importants, par exemple), et une licence payante représente un coût supplémentaire pour les indépendants. L'exercice de la post-édition doit donc être différencié selon qu'il est réalisé par des

traducteurs salariés ou proposé par des agences de traduction, avec des outils imposés, ou qu'il est effectué de manière autonome par un traducteur indépendant, qui doit mettre en place sa propre configuration. Comme l'affirmait [O'Brien, 2012], la traduction moderne est une interaction entre le traducteur et les technologies. Les professionnels doivent donc intégrer ce nouvel outil dans un environnement aujourd'hui déjà composé de mémoires de traduction, de glossaires terminologiques ou de corpus de toutes sortes.

C'est dans ce contexte que nous avons souhaité évaluer la qualité obtenue par des traducteurs professionnels dans une configuration indépendante, lors de la traduction de documents spécialisés en sciences de la santé, domaine de spécialisation complexe [Garcia-Castillo and Fetters, 2007]. Pour cela, nous avons cherché en premier lieu à définir la situation actuelle de ces traducteurs, ainsi que leurs pratiques en matière d'exploitation de la traduction automatique.

Notre article est plus généralement organisé de la manière suivante : en premier lieu, nous passerons en revue un certain nombre d'enquêtes réalisées ces dernières années auprès de traducteurs sur la traduction automatique. Puis, nous présenterons les résultats de notre propre enquête. Enfin, nous discuterons des orientations suscitées par ces résultats et des développements poursuivis dans notre thèse suite à ce questionnaire. Nous résumerons notamment les observations liées à l'utilisation de la traduction automatique dans l'échantillon que nous avons soumis à évaluation et quelques marqueurs significatifs d'acceptabilité de la post-édition.

I ENQUÊTES EXISTANTES ET MÉTHODOLOGIE DE L'ENQUÊTE

1.1. Enquêtes existantes

La SFT (Société française des traducteurs) a déjà réalisé plusieurs enquêtes visant à étudier les pratiques professionnelles des traducteurs. Aucune des questions posées lors de leur dernière enquête, réalisée en 2015, ne mentionne la traduction automatique, et la liste des prestations pouvant être proposées par les traducteurs ne contient pas la post-édition.

C'est en 2018 que le marché de la traduction opère un véritable basculement dans l'utilisation de la traduction automatique : l'enquête European Language Industry Survey indique pour la première fois que plus de la moitié des prestataires de services linguistiques et des traducteurs indépendants utilisent la traduction automatique, respectivement 64 % et 52 %, contre 20 % et 31 % en 2016. Une enquête réalisée sur la même période par l'association des métiers de la traduction et de l'interprétation en Belgique observait des chiffres similaires, avec 41% traducteurs ayant répondu ne jamais pratiquer la post-édition et ne pas le souhaiter. Cette enquête révèle que le marché n'est pas divisé seulement entre les traducteurs qui réalisent ou non de la post-édition : alors que 30 % des répondants ont affirmé effectuer de la post-édition, à différentes fréquences, 29 % ont, quant à eux, répondu qu'ils n'en réalisaient pas encore, mais qu'ils n'étaient pas opposés à la possibilité d'ajouter cette tâche parmi leurs services proposés. Cette dualité apparente dans l'acceptation ou non de la traduction automatique offre en réalité des nuances dans la situation actuelle de la traduction automatique.

1.2. Méthodologie de l'enquête

Entre janvier et mai 2020, nous avons réalisé une enquête par questionnaire, diffusé à des associations de traducteurs ou par des réseaux professionnels et destiné aux traducteurs indépendants intervenant sur des contenus médicaux. Disponible en ligne et composé de 35 questions, le questionnaire a reçu un ensemble de 139 réponses exploitables.

La catégorie d'âge la plus représentée est celle des traducteurs de plus de 50 ans, et près de 60 % des répondants ont plus de 40 ans. Concernant l'expérience professionnelle, près de 70 % ont répondu avoir au moins 10 ans d'expérience en traduction. 98 % des répondants disposent au moins d'une formation de niveau Bac+3, et 65 % disposent d'une formation niveau Bac+5.

Cette enquête s'adressait aux traducteurs intervenant dans le domaine médical, mais les pratiques de nos répondants se sont montrées variées : une sélection de 8 sous-domaines de traduction médicale a été proposée afin de définir les sous-domaines les plus présents. Le sous-domaine le plus mentionné est la pharmacologie, pour 75 % des répondants, suivi par les documents d'essais cliniques (70 % des répondants), puis la documentation technique (69 % des répondants). Les autres sous-domaines proposés sont les contenus de recherche, les écrits promotionnels, les rapports médicaux, les contenus de vulgarisation et les documents règlementaires.

La présence du domaine médical dans la pratique de nos répondants est également variée, la traduction médicale représentant en moyenne 56 % du chiffre d'affaires des traducteurs, tandis que 15,6 % des répondants travaillent uniquement dans le domaine médical. 65 % des répondants ont indiqué avoir suivi une formation spécifique en traduction médicale. La plupart d'entre eux ont suivi une formation payante (61 %). Seuls 18,9 % des répondants ayant suivi une formation spécifique à la traduction médicale ont réalisé une formation dans le cadre d'un cursus universitaire spécialisé. Parmi les réponses « Autre », des traducteurs ont précisé s'être formés sur le terrain, tandis que d'autres disposent de formations universitaires scientifiques ou médicales (diplôme en biologie, par exemple).

II RÉSULTATS DE L'ENQUÊTE

2.1 Utilisation de la traduction automatique

Dans cette section, nous présentons les résultats concernant l'utilisation de la traduction automatique et la pratique de la post-édition. Au total, 48 % des répondants utilisent la traduction automatique, mais ce chiffre doit être nuancé, puisque seuls 14 % des traducteurs ont un chiffre d'affaires qui correspond pour plus de 50 % à des projets de post-édition. En moyenne, la post-édition représente 21 % du chiffre d'affaires des traducteurs qui acceptent de mener cette activité/proposer ces prestations, la médiane étant de 10 % du chiffre d'affaires.

Parmi les répondants qui acceptent de réaliser de la post-édition, 62 % l'ont déjà effectuée dans le domaine médical. Pour ceux qui indiquent ne pas effectuer de post-édition en traduction médicale, certains évoquent parmi les raisons exprimées les spécificités de la

traduction médicale et ses exigences de qualité, mais d'autres indiquent également que la post-édition ne leur a jamais été proposée dans ce domaine. Parmi les traducteurs ayant déjà réalisé de la post-édition dans le domaine médical, 13 % ont indiqué que ces traductions étaient réalisées pour des clients directs.

Nous nous sommes ensuite intéressés à la qualité de post-édition exigée par les projets. La majorité (71 %) des projets de post-édition exige un niveau de *Full post-editing* (ou « post-édition complète »), mais 24 % des répondants ont répondu ne pas savoir précisément quelle qualité leur était demandée. Cette notion a pourtant des implications concrètes en matière de productivité, et donc de tarifs. Les conséquences sont également importantes sur la perception du traducteur par rapport à l'activité de post-édition.

Quelques questions de notre enquête concernaient plus précisément les outils utilisés par les traducteurs. Parmi nos répondants, 12 % d'entre eux ont affirmé ne pas utiliser de logiciel de TAO. Les traducteurs utilisant la TAO dans leur environnement de traduction ont mentionné SDL Trados (67 %), suivi de MemoQ (34 %) et Wordfast (16 %). Seuls 23 % des traducteurs acceptant de réaliser de la post-édition ont investi dans un outil de traduction automatique, tandis que les autres répondants utilisent un moteur fourni par leur client.

Nous avons cherché à déterminer si les traducteurs avaient recours à la post-édition uniquement lorsque cela leur était imposé par un projet, ou s'ils avaient recours à une utilisation « autonome » de la traduction automatique. 62 % des traducteurs acceptant de réaliser de la post-édition ont précisé ne jamais utiliser la traduction automatique si elle ne leur était pas imposée par le projet. Parmi les 38 % restants, nous avons identifié deux schémas d'utilisation autonome : la traduction automatique est appliquée sur certains segments d'un projet pour 27 % des répondants, alors que 13 % l'appliquent sur l'ensemble d'un projet. Lorsqu'ils utilisent la traduction automatique alors qu'elle n'est pas imposée par leur client, la majorité des traducteurs ont indiqué ne pas en informer le client (65 %). La principale raison évoquée pour une utilisation « autonome » de la traduction automatique est l'augmentation de la productivité (73 %), puis l'utilisation de la traduction automatique comme un outil d'aide en cas de traduction difficile (54 %). Concernant les outils utilisés en cas de post-édition autonome, 54 % des traducteurs ont indiqué utiliser un outil disponible sur une interface web, tandis que 38 % utilisent un outil intégré à leur système de TAO. Les critères de choix de l'outil se portent principalement sur la qualité de la traduction automatique, puis sur la confidentialité des données, la facilité d'utilisation ou d'installation, la possibilité d'intégration de la traduction automatique dans un outil de TAO et le prix. Le moteur de traduction automatique le plus mentionné est DeepL (pour 77 % des répondants), mais les moteurs de SDL, Google Traduction et Systran sont également utilisés.

Les traducteurs ont eu la possibilité de répondre à une question ouverte sur les différences entre révision et post-édition dans le domaine médical. Des traducteurs ont précisé que, selon eux, la différence entre la révision et la post-édition se situe dans le rapport de confiance par rapport au document final. En effet, ils indiquent que dans le cadre d'une révision, ils font confiance au traducteur humain, qui a déjà effectué des recherches documentaires afin de définir la terminologie ou la phraséologie de la traduction, par exemple. Dans le cas de la

post-édition, ce travail de recherche et d'harmonisation n'a pas été réalisé en amont par une personne. Quelques répondants ont également précisé qu'ils vérifient plus attentivement le sens lors d'une post-édition, car c'est un point sur lequel ils font moins confiance à un outil automatique qu'à un traducteur. L'erreur la plus souvent mentionnée porte sur la terminologie, parfois citée comme étant incohérente, ce qui nécessite une attention particulière. Un seul répondant a cependant indiqué estimer que la terminologie est de meilleure qualité dans le cadre d'une post-édition. Globalement, les traducteurs ont indiqué qu'ils ne recherchent pas le même type d'erreurs dans une révision que dans une post-édition. Certains ont également ajouté que la post-édition leur demande plus de temps, du fait des recherches qu'ils approfondissent et du nombre d'erreurs plus important.

Nous leur avons également demandé de nous préciser les erreurs les plus souvent rencontrées lors d'une tâche de post-édition : les erreurs les plus citées sont les erreurs de sens, de terminologie et de style. Lorsque nous les avons interrogés sur les erreurs spécifiques au domaine médical, ils indiquent également avoir observé des erreurs de terminologie et de sens, et certains ont précisé que ces erreurs sont causées par le manque de prise en compte du contexte. Un traducteur cite des exemples d'erreurs de niveau de langue, entre la vulgarisation et la langue des experts. D'autres erreurs sont également citées, par exemple des erreurs liées à la mauvaise gestion des abréviations ou des formules chimiques.

2.2 Refus de post-édition

52% des répondants de notre enquête ont indiqué ne pas réaliser de post-édition ou ne pas accepter ce type de prestation. Ces traducteurs n'ont donc pas répondu aux questions concernant l'environnement de post-édition, mais ont été invités à développer leurs raisons, dans une question ouverte. Dans cette section, nous allons explorer les différents motifs de refus exprimés par les traducteurs.

Nous notons que l'acceptabilité ou non de la traduction automatique est associée dans notre échantillon à l'acceptabilité générale des technologies de traduction, notamment à celle des outils de TAO. En effet, 16 % des traducteurs qui ne réalisent pas de post-édition ont également répondu ne pas utiliser les outils de TAO dans leur pratique, tandis que ce chiffre est de 8 % chez les traducteurs qui acceptent de réaliser de la post-édition.

La première raison motivant le refus de réaliser de la post-édition est la qualité insuffisante du document à post-éditer. Les explications des répondants semblent également indiquer que ce n'est parfois pas seulement le niveau de qualité de la traduction brute qui leur pose souci, mais plutôt les spécificités mêmes du processus de post-édition qui, lorsqu'elles sont comparées à celles du processus de traduction, ne semblent pas permettre de corriger correctement ces erreurs. Deux répondantes ont, par exemple, expliqué : « [...] j'ai trouvé qu'il y avait beaucoup trop d'erreurs graves ainsi que des propositions de traduction qui pouvaient m'induire en erreur » ou « J'estime que cela est trop risqué. C'est plus rapide mais source d'oublis et de confusions ». La comparaison entre biotraduction et post-édition est très présente dans les réponses, principalement car le processus de post-édition semble être cognitivement plus complexe que celui de biotraduction pour ces répondants : « aucun gain

de temps, bien au contraire (la correction d'une traduction automatique erronée sur le plan du sens et de la grammaire nécessite plus d'attention que la traduction directe) ». Les répondants se sentent donc également lésés sur le plan de la productivité, et précisent également que cela est lié aux tarifs généralement proposés pour la post-édition : il n'est donc pas rentable pour eux d'abandonner la biotraduction au profit de la post-édition.

Une des raisons citées est également le manque d'attrance personnelle pour cette tâche, parfois comparée à la révision qu'ils n'apprécient pas non plus. Deux traducteurs disent même qu'ils trouvent/indiquent même trouver la post-édition « ennuyante ».

Concernant les aspects plus pratiques, la question de la confidentialité des documents traités par les moteurs automatiques revient également dans les réponses : le manque de certitude sur la confidentialité lors du traitement par les moteurs de traduction automatique est également un facteur de refus.

Alors qu'ils ont répondu qu'ils n'effectuaient jamais de post-édition, quelques traducteurs ont cependant répondu qu'ils pourraient revenir sur ce choix dans le futur. En effet, certains ont indiqué qu'ils ne réalisaient pas de post-édition, car cela ne leur avait jamais été proposé par leurs clients. D'autres traducteurs ont répondu qu'ils n'étaient pas contre la post-édition, mais qu'ils estimaient ne pas être suffisamment formés pour accepter ce type de tâche.

2.3. Facteurs d'acceptabilité de la post-édition

À partir de nos résultats, nous avons recherché quels sont les profils de traducteurs susceptibles d'accepter plus facilement de réaliser de la post-édition.

Notre hypothèse est que les traducteurs les plus jeunes et avec moins d'expérience en traduction acceptent mieux la post-édition.

Nous avons donc tout d'abord testé l'interdépendance entre l'âge des traducteurs et l'acceptabilité de la post-édition. Pour le traitement de cette analyse, les âges sont regroupés par tranche de 10 ans.

	Acceptation de la PE	Refus de la PE
De 20 à 30 ans	11	6
De 31 à 40 ans	25	15
De 41 à 50 ans	17	29
Plus de 50 ans	14	16

Table 1. Acceptation de la post-édition en fonction de l'âge.

C'est à partir de ces données que nous avons appliqué le test statistique du χ^2 afin de mesurer l'indépendance entre les deux variables.

Pour ce test, le χ^2 global obtenu est de 10,175 et la valeur p est de 0,01714. Avec un seuil de significativité de 0,05, on peut donc en conclure une forte présomption contre l'hypothèse nulle, et donc que l'association entre les variables est statistiquement significative dans notre échantillon. Ceci signifie, pour notre étude, qu'un traducteur plus jeune a tendance à mieux accepter la post-édition.

L'enseignement de la post-édition et du fonctionnement de la traduction automatique ayant été intégré seulement récemment dans les formations universitaires, le facteur de l'âge peut être corrélé à un manque de formation en post-édition lors de la formation initiale de certains traducteurs. Dans nos résultats, nous observons une réticence majoritaire à partir de la tranche d'âge de 41 à 50 ans.

Ce facteur de l'âge pourrait également être lié à l'expérience professionnelle des traducteurs, et donc à une certaine routine professionnelle qui serait installée et plus difficile à bousculer pour y intégrer une nouvelle technologie. Nous avons donc ensuite réalisé un test statistique d'interdépendance entre le nombre d'années d'expérience des traducteurs et l'acceptabilité de la post-édition.

	Acceptation de la PE	Refus de la PE
De 0 à 5 ans	17	3
De 5 à 10 ans	14	14
De 10 à 15 ans	14	13
De 15 à 20 ans	17	19
De 20 à 30 ans	8	17
Plus de 30 ans	4	6

Table 2. Acceptation de la post-édition en fonction de l'expérience.

Pour ce test, le χ^2 global obtenu est de 13,942 et la valeur p est de 0,01598. Avec un seuil de significativité de 0,05, on peut donc en conclure une forte présomption contre l'hypothèse nulle, et donc que l'association entre les variables est statistiquement significative dans notre échantillon.

Ces résultats confortent notre hypothèse selon laquelle l'âge n'est pas le seul facteur déterminant l'acceptabilité de la post-édition, mais que l'expérience est également une variable à prendre en compte. La différence est d'autant plus marquée pour les traducteurs ayant moins de dix ans d'expérience, ce qui correspond plus ou moins à la période ayant vu l'apparition de la traduction neuronale et l'évolution de la traduction automatique sur le marché de la traduction.

Ces facteurs, liés aux profils des traducteurs, s'intègrent bien dans le modèle d'acceptabilité des technologies (TAM¹) développé par [Davis, 1993]. Selon le TAM, l'attitude des professionnels repose sur deux éléments :

- l'utilité perçue, définie comme « le degré avec lequel une personne pense que l'utilisation d'un système améliore sa performance au travail », peut être affectée par de nombreuses variables selon le profil de l'utilisateur (âge, expérience...)
- et la facilité d'utilisation, qui correspond « au degré auquel une personne pense que l'utilisation d'un système ne nécessite pas d'efforts ».

¹ Technology Acceptance Model

Ces deux éléments sont annoncés par le TAM comme des marqueurs prédictifs de l'utilisation d'une technologie. Les facteurs que nous avons identifiés s'inscrivent donc dans la notion d'utilité perçue, et sont confirmés par les ressentis exprimés par les traducteurs qui refusent la post-édition : le sentiment que la post-édition leur prend plus de temps et que la qualité finale n'en sera pas améliorée. Le manque de formation ou d'information sur la traduction automatique et la post-édition intervient directement sur la facilité d'utilisation.

Conclusion

Cette enquête nous a permis d'observer l'intégration de la traduction automatique auprès d'un échantillon de traducteurs indépendants intervenant dans le domaine médical, et d'en déduire des profils et des pratiques très variées, qui dépassent la dualité entre l'acceptation ou le refus de la traduction automatique par les professionnels. L'acceptation de la post-édition est en effet nuancée par son degré d'adhésion, mesuré par la part de la post-édition dans l'activité des traducteurs, mais également par une utilisation parfois volontaire et autonome, c'est-à-dire non imposée par les clients. Les motifs exprimés par les traducteurs en cas de refus de la post-édition montrent également plusieurs tendances : le manque de confiance en matière de qualité est souvent évoqué dans notre enquête, tout comme l'absence de formation en post-édition, qui est parfois citée comme un frein pour accepter des projets de post-édition. Les réponses marquent également les différences de pratiques et de processus cognitifs entre les tâches de biotraduction et de post-édition.

Ces retours seront complétés par des entretiens directs avec des indépendants représentant les différents profils observés lors de l'analyse de cette enquête. L'objectif est de recueillir les retours des professionnels au cours de l'expérimentation d'un environnement de traduction automatique mis en place dans un contexte d'utilisation autonome. Cet environnement fera également l'objet, en parallèle, d'une évaluation de la qualité obtenue sur trois types de discours médicaux, en s'appuyant sur une analyse objective réalisée à l'aide d'une grille d'évaluation quantitative. Ces recherches sont réalisées dans le cadre d'une thèse cherchant à explorer un cadre méthodologique et un environnement optimisant l'utilisation de la post-édition en traduction médicale indépendante, lorsque celle-ci est adaptée au contexte.

References

- Castilho S., Moorkens J., Gaspari F., Calixto I., Tinsley J. and Way A. Is neural machine translation the new state of the art?. *The Prague Bulletin of Mathematical Linguistics*. 2017;108(1):109-120.
- Davis F. User acceptance of information technology. *International Journal of Man-Machine studies*. 1993;38.
- Loock R. La plus-value de la biotraduction face à la machine. *Traduire*. 2019;241:54-65.
- Garcia-Castillo D., Fetters M.D. Quality in medical translations: a review. *J Health Care Poor Underserved*. 2007;18(1):74-84.
- O'Brien S. Translation as human-computer interaction. *Translation spaces*. 2012;1(1):101-122.
- http1 <https://www.sft.fr/fr/chiffres-cles>
- http2 <https://euatc.org/industry-surveys/2018-european-language-survey-results-announced/>
- http3 <https://www.cbti-bkvt.org/fr/publications/market-survey-report>