

# Analyse des erreurs de traduction automatique pour la combinaison de langues slovène-français et perspectives pour une formation en post-édition\*

*Análisis de los errores de traducción automática para la combinación de idiomas esloveno-francés y perspectivas para una formación en posesición*

*Error Analysis in Machine Translation for Language Pair Slovenian-French and Perspectives for Post-Editing Training*

Sonia Vaupot<sup>a</sup>

*Université de Ljubljana, Ljubljana, Slovénie*

---

\* Article de réflexion

Reçu: 1er novembre 2020    Accepté: 16 juillet 2021

<sup>a</sup> Ph.D. de Traductologie.  
Maître de conférences.

sonia.vaupot@ff.uni-lj.si – ORCID: <https://orcid.org/0000-0003-1410-0937>

**Cómo citar:** Vaupot, S. (2020). Analyse des erreurs de traduction automatique pour la combinaison de langues slovène-français et perspectives pour une formation en post-édition. *Matices en Lenguas Extranjeras*, 14(2), 83-110. <https://doi.org/10.15446/male.v14n2.91299>

Este trabajo se encuentra bajo la licencia Creative Commons Attribution 4.0.

## RÉSUMÉ

Ces dernières années, la traduction automatique (TA) a considérablement progressé de sorte que des connaissances en post-édition (PE) deviennent nécessaires dans l'industrie de la traduction. Cela nécessite de nouvelles connaissances et la mise en place d'un enseignement dans le domaine de la PE. Dans le présent article, nous vérifions l'usage de la TA dans une formation en traduction de niveau Master (Université de Ljubljana), à partir d'une langue maternelle à moindre diffusion (le slovène) vers une langue étrangère à large diffusion (le français), et nous en dégageons les erreurs dues à la TA. L'analyse des erreurs indique que la qualité de la TA diffère notamment en fonction du type de texte et du niveau de difficulté du texte à traduire. Enfin, nous proposons quelques perspectives pour une formation en PE axée sur la traduction et l'apprentissage d'une langue étrangère.

**Mots clés:** *analyse des erreurs, formation, post-édition, traduction automatique.*

## RESUMEN

En los últimos años, la traducción automática (TA) ha avanzado considerablemente, por lo que tener conocimientos de posesición (PE) es cada vez más necesario en la industria de la traducción. Esto requiere adquirir nuevos conocimientos y que se eduque en el campo de la PE. En este artículo, verificamos el uso de la TA en la formación en traducción a nivel de Maestría (Universidad de Ljubljana); basándonos en la TA desde una lengua materna no muy difundida (esloveno) hacia una lengua extranjera de mayor difusión (francés), e identificamos errores debidos a la TA. El análisis de errores indica que la calidad de la TA difiere notablemente según el tipo de texto y el nivel de dificultad del texto a traducir. Por último, ofrecemos algunas perspectivas para una formación en PE centrada en la traducción y el aprendizaje de una lengua extranjera.

**Palabras clave:** *análisis de errores, formación, posesición, traducción automática.*

## ABSTRACT

Machine translation (MT) has advanced significantly in recent years, making knowledge of post-editing (PE) essential in the translation industry. This requires new knowledge and training in PE. In this paper, we investigate the use of MT in a Master's level translation course at the University of Ljubljana, from a less common native language (Slovenian) to a more common foreign language (French) to identify MT errors. The failure analysis shows that the quality of MT varies depending particularly on the type of text, and the level of difficulty of the text that is to be translated. Finally, we propose some perspectives for PE training focused on translation and foreign language learning.

**Keywords:** *failure analysis, machine translation, post-editing, training.*

**CES DERNIÈRES DÉCENNIES, LE DOMAINE** de la traduction s'est transformé grâce au développement des outils d'aide à la traduction, des mémoires de traduction, etc. En effet, ces progrès imposent au traducteur « une modulation sensible de ses méthodes de travail [et parmi] tous les outils d'aide à la traduction, le logiciel de TA est celui qui impose le plus grand nombre de contraintes » (L'Homme, 2008, p. 245). Ces contraintes sont de nature diverse et concernent notamment l'apprentissage d'un logiciel, la relecture de textes et l'acquisition de certaines habiletés de vérification liées à la traduction automatique (TA). Ces exigences relèvent, entre autres, du domaine de la post-édition (PE).

La PE est une activité étroitement liée aux systèmes de TA. Elle se différencie de l'activité traditionnelle de révision ou de relecture d'un texte en ce sens qu'il s'agit de vérifier et de corriger une traduction issue d'un système de TA. Cela consiste à améliorer et « à repasser derrière un texte pré-traduit automatiquement pour le rendre humainement intelligible » (Robert, 2010, p. 139). Brunette et O'Brien (2011) notent également que « les textes à revoir ont une nature hybride puisqu'ils intègrent remontées de mémoires de traduction et sorties de traduction automatique » (p. 2). L'enseignement de la traduction humaine peut ainsi s'appuyer sur la TA par le biais de la PE. Les auteurs (Brunette & O'Brien, 2011) conseillent, par ailleurs, une formation préalable en PE pour les traducteurs. Lavault-Olléon (2011) estime, pour sa part, qu'une

meilleure connaissance des compétences et stratégies mobilisées en post-édition, de son organisation matérielle, des délais qu'elle nécessite et des degrés de qualité attendus permettraient aux traducteurs de l'accepter comme une tâche faisant partie intégrante de leur métier et aux formateurs de l'enseigner de façon plus organisée. (p. 9)

Toutefois, même si la TA est utilisée comme base pour des traductions professionnelles, elle possède encore de nombreuses lacunes et ne peut actuellement remplacer la traduction humaine. En effet, les traductions générées automatiquement sont loin d'être parfaites. Nous remarquons également que la TA entraîne des erreurs différentes de celles de la traduction humaine. Par conséquent, dans le processus d'enseignement/apprentissage de la langue étrangère et de la traduction à visée professionnelle, les apprenants semblent confrontés à des difficultés qui ont évolué avec la progression de la TA.

Dans le cadre de leur cursus universitaire, les étudiants inscrits en Master au département de traduction (Université de Ljubljana) doivent aussi traduire à partir du slovène, qui est leur langue maternelle (A), vers deux langues étrangères (B), à savoir l'anglais, qui est obligatoire, et le français (les autres langues B sont, au choix, l'italien ou l'allemand). Or, nous constatons que les étudiants ont de plus en plus recours à la TA pour traduire vers la langue maternelle ou vers une langue étrangère. La TA leur permet, outre la rapidité d'exécution, d'éviter certaines fautes de langue, en l'occurrence des fautes de français. En outre, il est difficile, à l'heure actuelle, de s'opposer complètement à l'usage de la TA dans un cursus destiné à former des traducteurs, d'où la nécessité de perfectionner la formation initiale en traduction.

Notre recherche se rapporte à la traduction du slovène, une langue de petite diffusion ou périphérique, vers le français considéré comme une langue supercentrale, de niveau 2 (pour reprendre la classification proposée par Calvet, 1999, pp. 75-83). Ainsi, en Slovénie, un pays d'environ deux millions d'habitants, le marché de la traduction est petit et un traducteur peut être amené à traduire vers sa langue maternelle et/ou vers la langue étrangère. Dans leur cursus, les apprenants sont aussi amenés à traduire d'une langue minorisée vers une langue étrangère dite

« supercentrale » ou une « grande langue. » On peut donc supposer que la traduction par un système de TA comporte des erreurs de langue ou de traduction plus importante que la TA entre deux langues majeures.

Dans cet article, nous présentons une expérience pédagogique menée dans le cadre d'un cours de traduction slovène-français suivi par des étudiants de master en traduction. Nous présentons les résultats de l'analyse qualitative d'un corpus composé de quatre textes sources en slovène, puis de leur TA et PE en français. L'outil utilisé est Google translate. L'objectif principal de cette expérience est de mettre en évidence les limites de la TA, et ce, malgré les progrès réalisés dans ce domaine, mais aussi de vérifier la possibilité d'utiliser la TA et la PE à des fins pédagogiques.

Après quelques généralités sur les liens entre la TA et la PE, nous abordons les typologies d'erreur dans ce domaine à savoir la traduction de la langue maternelle, le slovène, vers une langue étrangère, le français. La TA de quatre textes, de type différent, nous a permis de mettre en lumière une typologie des principales erreurs et d'évaluer la qualité de la TA. Enfin, la PE étant étroitement liée aux systèmes de TA, nous proposons quelques pistes didactiques afin d'intégrer la PE dans une formation destinée aux futurs traducteurs, l'accent étant mis sur la PE en langue étrangère.

## **TRADUCTION AUTOMATIQUE ET POST-ÉDITION**

Si les débuts de la TA remontent aux années 1930, c'est dans les années 60 que l'on commence à s'intéresser au modèle syntaxique des langues. Puis, dans les années 70, face aux besoins croissants dans le domaine de la traduction communautaire, la communauté européenne met en place le système Systran (Kübler, 2007). Dans les années 80, les activités de recherche en matière de TA évoluent et connaissent un véritable essor grâce à

la révolution en informatique, aux nouvelles théories syntaxiques et sémantiques, au développement de l'intelligence artificielle et de la traductologie (Kliffer, 2008). Des projets européens se mettent en place (GETA, Eurotra) et les prémices de l'intelligence artificielle se font sentir (Maxwell, 1990).

Puis, les logiciels d'aide à la traduction apparaissent et les outils en ligne progressent. Pour effectuer une traduction, les systèmes n'utilisent plus de règles définies, mais ils se servent de modèles statistiques (traduction automatique statistique ou SMT) construits à partir de corpus et de banques de données pour chaque langue. À partir des années 90, les mémoires de traduction sont de plus en plus présentes sur le marché et deviennent l'outil indispensable du traducteur professionnel. On note, par ailleurs, un intérêt et un besoin croissants pour les outils d'aide à la traduction. Peu à peu, la TA touche de nombreux domaines comme la syntaxe, le lexique, la linguistique théorique et formelle, la linguistique de corpus, etc. De plus, les systèmes de TA s'associent à de nombreux outils d'aide à la traduction (concordanciers, mémoires de traduction, dictionnaires électroniques, vérificateurs d'orthographe et de grammaire, etc.). C'est aussi à la fin des années 90 que les agences de traduction commencent à utiliser la TA en lien avec la PE. À ce propos, Guerberof Arenas et Moorkens (2019) notent que « In response, many universities have seen the necessity, as with CAT tools in the 1990s, of introducing MT and MTPE as part of their curricula for new generations of translation graduates. » (p. 219)

Le domaine de la TA progresse rapidement dans les années 2000 combinant l'usage de l'intelligence artificielle et des réseaux neuronaux. Depuis 2015, la traduction automatique neuronale (TAN) remplace la traduction automatique statistique. Aujourd'hui, la TAN évolue rapidement grâce notamment à l'auto-apprentissage basé sur l'Intelligence artificielle, le Big Data et le Deep Learning (Bentivogli et al., 2016 ; Toral &



Sánchez-Cartagena, 2017). Face à la traduction humaine, la TAN est donc considérée comme une technologie nouvelle, voire révolutionnaire, qui pousse le traducteur à acquérir de nouvelles compétences. De ce fait, les services de traduction utilisent de plus en plus la PE afin d'être plus rapides et compétitifs, mais aussi pour améliorer leur productivité. Certaines études ont en effet montré des gains de productivité significatifs lorsque les traducteurs optent d'abord pour la TAN (par la suite TA) et ont ensuite recours à la PE (Federico et al., 2012 ; Green et al., 2013). La PE est ainsi étroitement liée à la TA.

De l'anglais *to postedit*, la PE consiste à réviser ou corriger un texte traduit au préalable par un logiciel de TA, l'objectif étant de traduire plus rapidement tout en produisant une traduction de qualité. On distingue quatre types de PE : la PE est dite *complète* lorsque le texte final est révisé dans un souci de qualité (Wagner, 1985) ; elle est *professionnelle* lorsque le texte est produit par la TA et la traduction assistée par ordinateur (Robert, 2013). En revanche, la PE est dite *rapide* quand la révision du texte est superficielle (Wagner, 1985), et on la considère comme *brute* si la révision du texte traduit passe par un moteur de TA (Robert, 2013). Ces différentes manières de post-éditer se concentrent essentiellement sur le transfert correct du sens. Allen (2003) reprend la subdivision de Wagner (1985), mais ajoute à cette typologie la PE *minimale* qui vise à produire une qualité finale acceptable avec le minimum de modifications. Selon Allen (2003), le niveau de PE dépend de plusieurs facteurs : l'utilisateur final ou le client, le volume de documentation à traiter, le niveau de qualité attendu, le temps à disposition du post-éditeur, la durée de vie du document et du matériel informatique ou destiné à la publication.

Par conséquent, les progrès dans le domaine de la TA et le recours à la PE mènent à repenser les formations en traduction. En

effet, pour que les étudiants puissent « post-éditer correctement un texte, il importe de bien analyser ce que l'on voit, de bien déterminer ce qui doit être vérifié et éventuellement modifié. » (Loock, 2019, p. 60). Ne pouvant empêcher, à moyen ou long terme, les étudiants d'avoir recours à la TA, il est donc indispensable de les sensibiliser à la PE vers la langue maternelle, mais aussi vers la langue étrangère.

### **PARTICIPANTS ET DÉROULEMENT**

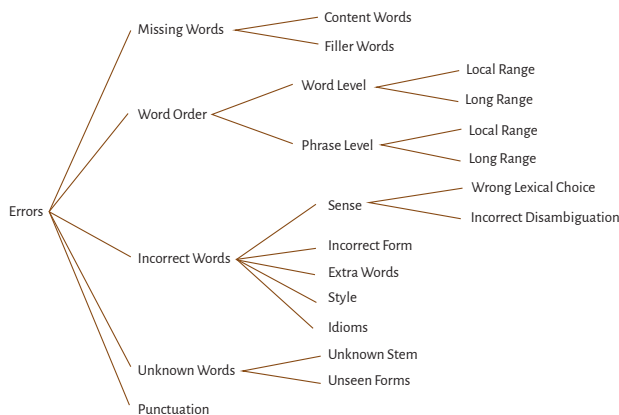
Dans le cadre d'un cours de Master, visant la traduction du slovène (la langue maternelle) vers le français (la langue étrangère), qui s'est déroulé au sein du département de traduction (Faculté des Lettres), nous avons effectué l'annotation des erreurs de traduction au cours de l'année académique 2019/2020. Notre objectif consistait, d'une part, à évaluer la qualité de la TA d'une langue minorisée, le slovène, vers une langue majeure, le français, et les éventuelles variations ou niveaux de difficulté en fonction du type de texte. Nous supposons en effet que les langues minorisées sont moins bien représentées que les langues centrales dans le domaine de la TA et que ce fait donne lieu à davantage d'erreurs. Nous souhaitions également vérifier les capacités des apprenants à améliorer une TA vers la langue étrangère. À cette fin, nous avons d'abord identifié les erreurs produites par le système de TA, puis vérifié les corrections effectuées par les étudiants. Les résultats obtenus visaient à orienter plus facilement l'étudiant dans l'activité de PE d'un texte vers le français langue étrangère.

Afin de mener à bien cette tâche, nous avons sélectionné quatre textes de genres différents et de difficultés croissantes. Les quatre textes sont les suivants : l'article de presse comprend 302 mots ; la présentation de la collection sur le site du Musée national slovène de Ljubljana est de 271 mots ; le texte touristique contient 305 mots ; enfin, l'article sur le patrimoine slovène, plus spécialisé, est

constitué de 375 mots. L'annotation s'est faite en deux temps. Nous avons d'abord vérifié la qualité de la TA des quatre textes. Parmi les systèmes et outils de traduction en ligne les plus connus et disponibles gratuitement pour le couple de langues slovène-français, on note (depuis peu) DeepL Translator et Google Translate. Notre choix s'est porté sur le système Google translate qui propose, depuis plusieurs années, des traductions du slovène vers le français et inversement. Puis, nous avons soumis ces traductions à un groupe de 25 étudiants. Les étudiants devaient post-éditer les textes traduits automatiquement. L'expérience s'est déroulée dans un contexte pédagogique, les étudiants disposant de toutes les ressources qu'il jugeait nécessaires (dictionnaires, glossaires, grammaire, etc.). Il leur a ainsi été demandé de relire les traductions, de corriger les erreurs et d'améliorer la traduction brute générée par le moteur de TA. Les productions ont été traitées de manière anonyme.

### **TPOLOGIE DES ERREURS**

Diverses classifications typologiques des erreurs sont possibles. Les recherches dans le domaine de la TA peuvent être classées notamment en trois grandes catégories : (1) les typologies d'erreurs au niveau linguistique qui permettent de classer les erreurs dans les sorties des systèmes (Farrus et al., 2010 ; Lommel et al., 2014 ; Vilar et al., 2006) ; (2) les typologies d'erreurs au niveau de la traduction (Koponen, 2010) ; (3) les typologies visant la métrique et la qualité du texte (Lommel & Melby, 2018). Certains travaux (Lommel et al., 2014 ; Vilar et al., 2006) portent notamment sur des typologies d'erreurs au niveau linguistique. Ainsi, Vilar et al. (2006) identifient cinq catégories principales d'erreurs de traduction, à savoir les mots absents ou non traduits, l'ordre des mots, les mots incorrects (mauvais choix lexical ou mauvaise désambiguïsation), les mots inconnus et le non-respect de la ponctuation de la langue cible ainsi que des sous-catégories. Vilar et al. (2006) dressent également la Figure 1 :

*Figure 1. Typologie des erreurs de traduction.*

*Source* : récupérée de Vilar et al. (2006, p. 699).

Pour leur part, Farrus et al. (2010) dressent une typologie des erreurs linguistiques les plus fréquentes dans le domaine de la TA. Cette classification est essentiellement d'ordre orthographique, morphologique, lexical, sémantique et syntaxique, chaque catégorie comprenant aussi des sous-catégories. Dans sa recherche, Koponen (2010) met l'accent sur les aspects sémantiques et propose une première typologie qui se base sur six catégories d'erreur : le concept est omis ; le concept est ajouté ; le concept n'est pas traduit ; le concept est mal traduit ; le concept est substitué et le concept est expliqué. À cette classification de base viennent s'ajouter les concepts d'omission ou d'ajout : le participant est omis ou ajouté ; la relation est omise ou ajoutée en raison d'erreurs morphosyntaxiques ; le participant ou la relation n'est pas identique ; le participant ou la relation a été substitué(e). Ces normes sont censées veiller à la cohérence et la clarté d'une traduction.

Pour mener à bien notre recherche, nous nous sommes basées sur ces études antérieures et avons distingué dans la grille ci-dessous (Tableau 1), six types d'erreurs qui reposent sur des catégories linguistiques traditionnelles : grammaire/(morpho) syntaxe/style ; sens/sémantique ; orthographe/ponctuation ; lexicale/terminologie ; phraséologie/collocation ; omission/ajout/calque.

**TABLEAU 1.** Grille générale des erreurs de langue et de traduction conçue à partir des études précitées.

grammaire/ (morpho)- syntaxe/ style	sens/ sémantique	orthographe/ ponctuation	lexique/ terminologie	phraséologie/ collocation	omission/ ajout/ calque
ordre des mots	ambiguïté	orthographe non conforme	variante inappropriée	phraséologie inappropriée	mots/noms propres non connus
structure et longueur de la phrase	perte de sens	ponctuation non conforme	terme appartenant à un autre domaine	phraséologie appartenant à un autre domaine	formes non connues
temps verbal, concordance des modes et des temps	sens incorrect, faux-sens, contresens		incohérence terminologique	incohérence /mauvaise collocation ou expression idiomatique	mot inséré
déterminant			néologisme		
préposition			abréviation		
accord en genre et en nombre					
registre de langue					

*Source* : élaboré par l'auteur.

Nous reproduisons dans le Tableau 2, sur la base du Tableau 1, la typologie des erreurs obtenues à partir de la TA des quatre textes slovènes traduits vers le français.

**TABLEAU 2.** *Typologie des erreurs de TA pour les textes slovènes.*

type d'erreurs	texte 1	texte 2	texte 3	texte 4
grammaire/(morpho syntaxe/style	4	7	13	17
sens/sémantique	2	9	6	9
orthographe/ ponctuation	3	2	3	7
lexique/terminologie	4	4	5	2
phraséologie/collocation	-	-	3	3
omission/ ajout/calque	-	-	-	4

*Source :* élaboré par l'auteur.

### Diagnostics des erreurs

Le Tableau 3, ci-dessous, indique le pourcentage des erreurs de traduction, tous textes compris.

**TABLEAU 3.** *Pourcentage des erreurs de langue et de traduction pour les textes slovènes.*

grammaire/ (morpho) syntaxe/ style	sémantique	orthographe/ ponctuation	lexique/ terminologie	phraséologie/ collocation	omission/ calque
38%	24%	14%	14%	6%	4%

*Source :* élaboré par l'auteur.

Il apparaît que les erreurs les plus fréquentes dans le texte 1 (article de presse sur les tissus coptes) concernent l'erreur syntaxique, à savoir l'ordre des mots, et l'erreur lexicale pour les répétitions de termes et le manque de nuance de sens. Dans le texte 2 (présentation internet de l'exposition du Musée national – la collection des tissus coptes), les erreurs les plus fréquentes s'appliquent aux erreurs syntaxiques (l'ordre des mots), stylistiques (lourdeur des phrases) et sémantiques (sens incorrect et faux-sens). Pour le texte 3 (texte touristique – présentation de la région karstique Planinsko polje), l'erreur grammaticale (problème de déterminant et emploi de la préposition) apparaît le plus souvent. Sur le plan sémantique, le système de TA a commis le plus de faux-sens. Enfin, dans le texte 4 (l'article sur l'histoire des costumes traditionnels slovènes), les erreurs relatives à la syntaxe et la stylistique (lourdeur des phrases, omission de l'article, phrase incompréhensible) sont les plus fréquentes. Sur le plan sémantique, le système de TA a commis le plus d'erreurs de faux-sens et sur le plan orthographique, les virgules et les points sont omis ou mal placés. La traduction est, quant à elle, littérale.

L'analyse des erreurs indique que, pour le couple de langue slovène-français, la qualité de la TA diffère bien en fonction du type de texte et que les articles de presse ou textes journalistiques comportent le moins d'erreurs et semblent les plus appropriés pour la TA du slovène vers le français. Les textes proposés étant de niveaux linguistiques différents, ce fait peut expliquer les niveaux variables des traductions automatiques. Quelques exemples d'erreurs générées par le moteur de traduction Google dans tous les textes sont indiqués dans le Tableau 4 :

TABLEAU 4. Nature des erreurs et exemples pour les textes slovènes.

erreurs	nature de l'erreur	exemples
syntaxiques	ordre des mots incorrect	<i>*la timidement cachée Rakov Škocjan dans les bois ;</i>
grammaticales	(1) accord en nombre,	(1) <i>*Il n'est pas non plus indépendant de la mode (le singulier est utilisé à la place du pluriel, cela concerne les costumes folkloriques).</i>
	(2) mauvais emploi de la préposition,	(2) <i>*Quand nous regardons de Tromostovje dans le centre de Ljubljana à la Ljubljana verte (mis pour « quand nous observons le centre de Ljubljana depuis les trois ponts Tromostovje et la verte rivière Ljubljana » ; *des éléments encore plus ou moins clairs et reconnaissables de leurs habitants (au lieu de « par/pour les habitants »).</i>
	(3) omission de l'article défini	(3) <i>*population urbaine ; *cachés dans dépôts depuis 130 ans.</i>
orthographiques	virgule mal placée	<i>*Les Coptes sont les descendants des anciens Égyptiens, qui ne se sont pas convertis à l'islam après la conquête arabe ou turque ; *C'est l'un des sept noms du gouffre karstique, qui a coulé à la surface avant même l'ère glaciaire.</i>
stylistiques	répétitions erronées, lourdeur des phrases, phrases peu compréhensibles	<i>*ils ont également été endommagés par les colles avec lesquelles ils ont été collés aux bases lors de la compilation des collections ; *Avec de nombreux sites détaillés de Planinsko polje, il ne serait pas difficile de remplir au moins un an de suppléments Naša družina, donc ce dossier devrait vous encourager à en trouver au moins une poignée vous-même !</i>
sémantiques et lexicales	sens incorrect ou faux sens, répétitions, manque de nuances de sens	<i>*le défunt était enterré dans le plus beau travail qu'il ait jamais eu de son vivant ; *des changements plus visibles ont commencé à apparaître dans les costumes paysans, à la suite desquels les versions de paysage ont été perdues.</i>
phraséologiques	collocations incorrectes	<i>*Ainsi, pour nous, observateurs de surface et généralement superficiels.</i>

Source : élaboré par l'auteur.



Par conséquent, les erreurs dans les TA sont avant tout des fautes de langue (grammaire ou style), suivies par des erreurs de traduction (erreur de sens, calque, traduction littérale). Les exemples cités dans le Tableau 4 montrent que certaines phrases, notamment dans les textes 3 et 4, sont incohérentes. Les erreurs de ponctuation portent essentiellement sur un mauvais usage de la virgule, et les erreurs lexicales sont des variantes inappropriées. Enfin, les fautes d'ordre phraséologique sont moindres.

La majorité des étudiants ont rectifié les erreurs suivantes syntaxiques (rétablissement de l'ordre des mots correct), grammaticales (accord en nombre), stylistiques (répétitions erronées, phrases peu compréhensibles, lexique erroné). En revanche, peu d'étudiants ont apporté une correction concernant la grammaire (mauvais emploi de la préposition ou omission de l'article), l'orthographe (virgule mal placée), la syntaxe (traduction littérale ou calque), le style (phrases trop lourdes), la sémantique ou le lexique (nuances de sens). Les erreurs de ponctuation s'expliquent sans doute par le fait que la virgule est placée obligatoirement devant un pronom relatif en slovène, ce que les apprenants ont tendance à reproduire en français. Les difficultés liées à l'emploi de la préposition subsistent, même si elles sont moindres. En outre, les étudiants n'ont pas corrigé les lourdeurs de style ni le manque de cohésion. Nous estimons toutefois que, malgré les nombreuses lacunes de la TA pour la combinaison de langues slovène-français, il peut être intéressant d'exploiter le domaine de la TA dans un cours universitaire de traduction, vers ou à partir de la langue maternelle. Il est donc indispensable de former les futurs traducteurs à l'activité de correction et de PE, mais nous confirmons que les activités seront différentes lorsqu'il s'agit de la PE dans la langue maternelle ou une langue étrangère.

## QUELQUES PERSPECTIVES DIDACTIQUES À L'ACTIVITÉ DE POST-ÉDITION

La présente étude marque la nécessité de sensibiliser l'apprenant aux erreurs récurrentes en matière de TA. Il est toutefois utile de départager la formation de la PE en lien avec la langue maternelle, et celle destinée à approfondir les connaissances ou l'apprentissage en langue étrangère.

De plus en plus d'auteurs se penchent sur les besoins d'une formation en PE, mais essentiellement vers la langue maternelle. Vasconcellos (1986, p. 145) préconise de développer progressivement les compétences dans ce domaine. Il ajoute que les compétences essentielles d'un post-éditeur sont une excellente connaissance de la langue source, une parfaite maîtrise de la langue cible, ainsi que des connaissances techniques et en traduction. L'auteure estime également que l'étudiant doit pouvoir pratiquer la PE pendant et en dehors de ses cours. O'Brien (2002) propose de former les futurs traducteurs, dans le cadre d'un module réparti sur deux semestres, à l'activité de PE. Le cours comprendrait un premier volet théorique et un second plus pratique. La partie théorique serait une introduction à la PE, aux technologies de TA, au système auteur d'édition contrôlée, à la gestion de bases terminologiques, à la linguistique textuelle et aux compétences de base en programmation (O'Brien, 2002, p. 103). La partie pratique (O'Brien, 2002, pp. 104-105) consisterait à acquérir une certaine aisance en PE, notamment une bonne maîtrise linguistique, une aptitude à la gestion terminologique, des compétences en informatique et en traduction spécialisée, ainsi qu'aux technologies du langage. Pour ce faire, l'étudiant peut s'entraîner sur plusieurs types de textes et différentes traductions automatiques. Il peut aussi s'exercer sur au moins deux niveaux de PE (par exemple, rapide et complète). Il peut soumettre des textes à traduire au système de TA, puis analyser et comparer les résultats.

Il peut également étudier les avantages et les inconvénients de l'intégration d'un système de TA à un outil de mémoire de traduction, créer une base de données terminologique à l'aide d'un outil de gestion terminologique spécifique. L'étudiant doit donc « apprendre à gérer les transformations de paramètres des systèmes utilisés, la terminologie spécifique et les changements de règles linguistiques. Il est également nécessaire d'aborder l'intégration de la traduction automatique aux outils de traduction assistée par ordinateur » (De Faria Pires, 2018, p. 54).

Pour sa part, Depraetere (2010) distingue, dans un premier temps, les besoins des apprenants qui diffèrent de ceux des traducteurs professionnels. Partant de l'analyse des activités de PE des étudiants, l'auteure suggère de se concentrer sur l'analyse des erreurs en matière de TA. Elle propose, entre autres, d'aborder les différents niveaux de PE (rapide, minimale et complète) en lien avec l'apprenant qui doit pouvoir les distinguer et les appliquer. D'autres auteurs visent plutôt une collaboration avec les agences de traductions. Ainsi, Robert (2013) soutient que la TA doit être associée aux technologies de traduction assistée par ordinateur (TAO). Or cette association doit être préparée auparavant par des spécialistes dans les agences de traduction, d'où la difficulté de mettre en place un cours destiné à la PE. L'auteure (Robert, 2013, pp. 37-38) propose notamment des exercices qui peuvent être effectués à partir de traductions automatiques générées par des moteurs disponibles gratuitement sur internet. Toutefois, elle ajoute que ces moteurs semblent peu fiables et ne conviennent pas à une formation professionnelle. Pym (2013) estime également que la formation doit se baser sur certaines aptitudes et compétences nécessaires pour travailler avec les mémoires de traduction et la traduction automatique statistique dans un environnement professionnel.

Les nombreuses expériences récentes (pour ne citer que celles de Bouillon et al., 2021 ; Casas, 2021 ; Guerberof Arenas & Moorkens, 2019 ; Plaza Lara, 2019 ; Yamada, 2019 ; etc.) prouvent bien les tendances actuelles et les progrès dans la formation en PE. D'une manière générale, il semble préférable de proposer une formation en PE (vers la langue maternelle ou la langue étrangère), incluant des compétences générales, techniques ou spécialisées, plutôt en fin de parcours. Un tronc commun peut inclure des cours théoriques dispensés dans un cursus universitaire de traduction. Il est possible d'évoquer en cours le parcours historique et l'évolution de la TA ou encore les diverses approches (neuronale, statistique, etc.). L'enseignant peut également se baser sur les différents types de PE pour créer des cours ou pratiquer des exercices qui permettront de renforcer la capacité à produire des traductions professionnelles et à sensibiliser les apprenants aux différences entre la révision et la PE. Les cours peuvent ainsi aborder les différences entre la révision, qui se réfère à la traduction humaine dans une perspective de lecture et de vérification, mais aussi la PE qui concerne la lecture d'une traduction à l'état brut et nécessite des opérations de correction ou de remaniement du texte. Il est, par ailleurs, indispensable de développer le sens critique des apprenants afin qu'ils puissent fournir des traductions de qualité. Enfin, si les cours théoriques sur la PE sont indispensables, l'étudiant doit au préalable avoir été formé aux outils d'aide à la traduction ou outils de TAO avant d'aborder la PE de manière pratique. En plus des compétences linguistiques, il doit en effet pouvoir développer des compétences technologiques (De Faria Pires, 2018, pp. 51-52).

En ce qui concerne l'activité de PE vers la langue étrangère, des activités langagières supplémentaires sont nécessaires. Ces activités vont sans doute varier en fonction des langues. Pour

notre part, nous enseignons la traduction vers le français langue étrangère. Or, nous remarquons certains faits depuis l'avènement de la TA. D'une part, la nature des erreurs et les corrections de l'enseignant ont changé depuis que les apprenants utilisent la TA : par exemple, nous ne corrigeons quasiment plus les fautes d'orthographe (accent, cédille, accord de genre ou de nombre) ou certaines fautes de grammaire (mauvais emploi du déterminant, accord entre le verbe et le participe passé, etc.). En revanche, les fautes de grammaire qui subsistent, malgré le recours à la TA, pour le couple de langues slovène-français, sont : les erreurs de concordance des temps ; les différences entre le passé composé/passé simple/imparfait ; quelques prépositions (mais bien moins que lors d'une traduction humaine), etc. Par conséquent, nous confirmons que les compétences linguistiques (grammaire, syntaxe, lexique) et stylistiques sont indispensables, même après une traduction générée par un logiciel, pour améliorer ou parfaire la TA en français. Dans un premier temps, on peut suggérer des exercices de correction afin d'améliorer la TA et accroître les compétences en langue (syntaxe, style, par exemple) ou en traduction (vérification du sens, traduction de sens). Un travail sur les fautes récurrentes peut aussi s'avérer utile afin de sensibiliser les apprenants à anticiper les erreurs de TA.

D'autre part, après une première phase de correction, on peut envisager des exercices de PE même en langue étrangère. Nous suggérons un travail sur le texte dans son ensemble, et non uniquement sur des phrases isolées qui, selon nous, s'avère nécessaire. Il nous paraît en effet utile de développer les compétences stylistiques et textuelles des apprenants. On peut notamment viser la capacité à augmenter et, si possible, à contrôler la lisibilité du texte cible, mais aussi la faculté à s'adapter aux différents types de textes, ainsi que la capacité à postéditer la traduction. Ainsi, les cours pratiques pourraient comprendre un

travail ou des exercices sur les différents types de texte afin que l'étudiant prenne conscience des différents stades de TA. L'étudiant peut, par exemple, comparer deux versions pré-traduites, en examiner les différences, travailler sur les différents niveaux de PE (rapide, minimale, complète) avant de modifier, dans un contexte informatique, la version qu'il considère comme étant la plus simple à améliorer. Il peut notamment choisir quelques phrases du texte pré-traduit, parmi les plus problématiques, et parfaire la traduction avant d'harmoniser, articuler et finaliser l'ensemble. Il est aussi possible de travailler à partir de textes, plus ou moins complexes et prétraduits, en respectant les normes de correction linguistique et de contrôle de la qualité. Afin de reconnaître une TA de qualité, on peut proposer des exercices sur la lisibilité du texte (cohésion, cohérence), le style et les compétences rédactionnelles. Les critères de PE peuvent aussi varier en fonction de la typologie des textes (texte de vulgarisation, texte spécialisé). Pour les textes plus spécialisés, il est possible de consulter des supports de référence spécifiques (les glossaires terminologiques, notamment), ou de travailler la terminologie (Vaupot, 2020) et/ou la phraséologie (Vaupot, 2021) à partir d'un corpus. L'apprentissage par projet qui incite les apprenants à traduire de manière autonome et en contexte, voire à facturer pour le travail effectué, est une autre possibilité.

With this methodology we expect to develop the competences that the translation market requires, to improve second language learning and to raise awareness of what machine translation and post-editing imply in the translation process, but also what it takes to properly post-edit a machine-translated text in professional practice. (Castillo Bernal, 2021, p. 15)

Enfin, comme il est difficile d'atteindre la perfection en langue étrangère, il sera toujours conseillé de faire relire la traduction finale par une tierce personne lorsqu'on ne traduit pas vers sa langue maternelle.

## CONCLUSION

La pratique de la traduction est aujourd'hui indissociable des technologies. Dans cet article, nous avons dégagé les principales erreurs de TA et présenté les typologies des erreurs de TA, pour la paire de langues slovène-français, à partir de quatre textes de genre différent. Nous avons poursuivi en décrivant les recherches et les études qui préconisent une formation préalable dans le domaine de la PE. Enfin, partant du principe qu'il est de plus en plus difficile d'empêcher les apprenants d'avoir recours à la TA, surtout lorsqu'il s'agit de traduire vers une langue étrangère supercentrale, nous avons proposé quelques pistes susceptibles d'être prises en compte afin d'améliorer le texte brut, issu d'une TA, avant de le modifier dans un contexte informatique. Le gain de temps et la maîtrise de certains éléments langagiers sont des avantages certains de la TA. Toutefois, l'expérience a montré également les limites de la TA pour le couple de langues slovène-français. D'une part, la PE nécessite une bonne, sinon une excellente connaissance de la langue cible. Les apprenants ont certes amélioré les textes, mais surtout d'un point de vue langagier, en commettant parfois des fautes. La TA ayant tendance à traduire les textes littéralement, il restait encore à améliorer les traductions. Finalement, le gain de temps était moindre. Afin d'éviter que les étudiants ne se satisfassent d'une traduction trop rapide et trop imparfaite, nous avons avancé quelques aptitudes, compétences et activités linguistiques susceptibles, selon nous, d'être incluses dans une formation pratique en PE. Ces activités méritent encore d'être approfondies.

L'évolution rapide de la TAN implique certains changements dans la profession de traducteur et, de ce fait, une adaptation relativement rapide du cursus universitaire. Comme c'est souvent le cas en traduction professionnelle, la PE requiert notamment une part de travail relativement importante en terminologie et en phraséologie spécialisées, majoritairement vers la langue maternelle, mais aussi en langue étrangère et en gestion des projets. Il faut en effet pouvoir s'organiser pour réviser, corriger ou améliorer rapidement de gros volumes de traductions générées par un logiciel de TA, à rendre souvent dans des délais très courts. Les logiciels de TA sont des outils de plus en plus performants qui permettent un gain de temps et de productivité lorsque la traduction est cohérente ou s'ils sont adaptés à une langue spécialisée. Une formation en PE est utile à l'heure actuelle afin de s'adapter aux changements, même lorsque le logiciel de TA n'est pas encore suffisamment développé, comme c'est le cas pour quelques petites langues ou langues minorisées. Les apprenants doivent toutefois s'assurer que le texte cible correspond bien au texte source et surtout prendre conscience que, malgré l'extrême rapidité d'une TA, le résultat final doit nécessairement tendre vers la qualité.

Le métier de traducteur devient donc de plus en plus polyvalent. La traduction professionnelle semble en effet requérir des compétences technologiques alliées aux compétences langagières, traductionnelles, terminologiques et rédactionnelles. De ce fait, le profil semble, à l'ère actuelle, évoluer de pair avec la technologie et des partenariats avec l'industrie de la traduction.

## RÉFÉRENCES

- Allen, J. (2003). Post-Editing. In H. Somers (Ed.), *Computers and translation: A translator's guide* (pp. 297-317). Benjamins Translation Library. <https://doi.org/10.1075/btl.35.19all>



- Bentivogli, L., Bisazza, A., Cettolo, M., & Federico, M. (2016). Neural versus phrase-based machine translation quality: A case study. In J. Su, K. Duh, & X. Carreras (Eds.), *Proceedings of the 2016 conference on empirical methods in natural language processing* (pp. 257-267). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D16-1025>
- Bouillon, P., Girletti, S., & Volkart, L. (2021). Advanced machine translation post-editing training for in house professional translators. In *Proceedings of the Affumt 2021 Conference: Former aux métiers de la traduction aujourd'hui et demain* (pp. 10-11). Livret Abstracts.
- Brunette, L. & O'Brien, S. (2011). Quelle ergonomie pour la pratique postéditrice des textes traduits ? *ILCEA*, 14, 1-10. <https://doi.org/10.4000/ilcea.1081>
- Calvet, L.-J. (1999). *Pour une écologie des langues du monde*. Plon.
- Casas, S. (2021). Intégration de la post-édition dans un cours de traduction et révision. In *Proceedings of the Affumt 2021 Conference: Former aux métiers de la traduction aujourd'hui et demain* (pp. 13-14). Livret Abstracts.
- Castillo Bernal, P. (2021). Post-editing machine-translated specialised texts and second-language learning. In *Proceedings of the Affumt 2021 Conference: Former aux métiers de la traduction aujourd'hui et demain* (pp. 14-15). Livret Abstracts.
- De Faria Pires, L. (2018). Intégration de la traduction automatique neuronale à la formation universitaire des futurs traducteurs : pistes d'exploration. *Myriades*, 4, 50-61.
- Depraetere, I. (2010). What counts as useful advice in a university post-editing training context? Report on a case study. In V. Hansen & F. Yvon (Eds.), *Proceedings of the 14th annual conference of the European Association for Machine Translation*. European Association for Machine Translation.

- Farrus, M., Costa-Jussà, M. R., Marino, J. B., & Fonollosa, J. A. R. (2010). Linguistic-based evaluation criteria to identify statistical machine translation errors. In V. Hansen & F. Yvon (Eds.), *Proceedings of the 14th annual conference of the European Association for Machine Translation*. European Association for Machine Translation
- Federico, M., Cattelan, A., & Trombetti, M. (2012). Measuring user productivity in machine translation enhanced computer assisted translation. In J. DeCamp (Ed.), *Proceedings of the 10th conference of the Association for Machine Translation in the Americas, AMTA*. Association for Machine Translation in the Americas.
- Guerberof Arenas, A. & Moorkens, J. (2019). Machine translation and post-editing training as part of a master's programme. *Jostrans: The Journal of Specialised Translation*, (31), 217-238.
- Green, S., Heer, J., & Manning, C. D. (2013). The efficacy of human post-editing for language translation. In *Proceedings of the Sigchi Conference on Human Factors in Computing Systems* (pp. 439-448). Association for Computing Machinery. <https://doi.org/10.1145/2470654.2470718>
- Kliffer, M. (2008). Post-Editing machine translation as an FSL Exercise. *Porta Linguarum: Revista Internacional de Didáctica de las Lenguas Extranjeras*, (9), 53-68. <https://doi.org/10.30827/Digibug.31745>
- Koponen, M. (2010). Assessing machine translation quality with error analysis. In *MikaEL: Electronic proceedings of the KäTu symposium on translation and interpreting studie* (Vol. 4). Suomen kääntäjien ja tulkkien liitto.
- Kübler, N. (2007). La traduction automatique : traduction machine. *Tribune internationale des langues vivantes*, 45, 1-14.
- Lavault-Olléon, É. (2011). L'ergonomie, nouveau paradigme pour la traductologie. *ILCEA, 14*. <https://doi.org/10.4000/ilcea.1078>

- L'Homme, M.-C. (2008). *Initiation à la traductique*. Linguattech éditeur.
- Loock, R. (2019). La plus-value de la biotraduction face à la machine. *Traduire*, 241, 54-65. <https://doi.org/10.4000/traduire.1848>
- Lommel, A., Burchardt, A., Popovic, M., Harris, K., Avramidis, E., & Uszkoreit, H. (2014). Using a new analytic measure for the annotation and analysis of MT errors on real data. In M. Cettolo, M. Federico, L. Specia, & A. Way (Eds.), *Proceedings of the 17th conference of the European Association for Machine Translation* (pp. 65-172). European Association for Machine Translation.
- Lommel, A. & Melby, A. K. (2018). Tutorial: MQM-DQF: A good marriage (Translation quality for the 21st Century). In J. Campbell, A. Yanishevsky, J. Doyon, & D. Jones (Eds.), *Proceedings of the 13th conference of the Association for Machine Translation in the Americas* (Vol. 2). Association for Machine Translation in the Americas.
- Maxwell, D. (1990). *Distributed language translation: A multilingual project*. The Indiana University Linguistics Club.
- O'Brien, S. (2002). Teaching post-editing: A proposal for course content. In *Proceedings of the 6th EAMT workshop: Teaching machine translation* (pp. 99-106). European Association for Machine Translation.
- Plaza Lara, C. (2019). Análisis DAFO sobre la inclusión de la traducción automática y la posesición en los másteres de la red EMT. *Jostrans: The Journal of Specialised Translation*, (31), 260-280.
- Pym, A. (2013). Translation skill-sets in a machine-translation age. *Meta: Journal des traducteurs Translators' Journal*, 58(3), 487-503. <https://doi.org/10.7202/1025047ar>
- Robert, A.-M. (2010). La post-édition : l'avenir incontournable du traducteur ? *Traduire*, 222, 137-144. <https://doi.org/10.4000/traduire.460>

- Robert, A.-M. (2013). Vous avez dit post-éditrice ? Quelques éléments d'un parcours personnel. *The Journal of Specialised Translation*, (19), 29-39.
- Toral, A., & Sánchez-Cartagena, V. M. (2017). A multifaceted evaluation of neural versus phrase-based machine translation for 9 language directions. In M. Lapata, P. Blunsom, & A. Koller (Eds.), *Proceedings of the 15th conference of the European Chapter of the Association for Computational Linguistics* (Vol. 1, pp. 1063-1073). Association for Computational Linguistics. <https://doi.org/10.18653/v1/E17-1100>
- Vasconcellos, M. (1986). Post-editing on-screen: Machine translation from Spanish into English. *Translating and the Computer*, 8, 133-146.
- Vaupot, S. (2020). Création de ressources lexicographiques Français-Slovène d'aide à la traduction spécialisée. *Lexikos*, 30, 1-21. <http://doi.org/10.5788/30-1-1579>
- Vaupot, S. (2021). Creating a bilingual dictionary of collocations: a learner-oriented approach. *Indonesian Journal of Applied Linguistics*, 10(3), 762-770. <https://doi.org/10.17509/ijal.v10i3.31888>
- Vilar, D., Xu, J., D'Haro, L. F., & Ney, H. (2006). Error analysis of statistical machine translation output. In N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk, & D. Tapias (Eds.), *Proceedings of the fifth international conference on language resources and evaluation*. European Language Resources Association (ELRA).
- Wagner, E. (1985). Post-Editing systran – A challenge for commission translator. *Terminologie et Traduction*, 3, 1-7.
- Yamada, M. (2019). The impact of Google Neural Machine Translation on post-editing by student translators. *Jostrans: The Journal of Specialised Translation*, (31), 87-106.