

Intégrer des plateformes de traduction automatique neuronale dans l'enseignement de la traduction spécialisée

Nicolas Froeliger, Christopher Gledhill, Maria Zimina-Poirot
Université Paris Cité, France

Abstract This paper covers three different kinds of cyclical interaction implemented on the ILTS Master's programme (Université Paris Cité). We first describe an integrated translation project (inspired by similar programmes) with emphasis on neural translation, post-editing, project management and exchanges with an external third party. We then explain the importance of considering translator training today within a broader societal framework. Finally, we underline the key role played by blending together training, research and professional life in the rapidly developing universe of professional translation.

Keywords Neural machine translation. Post-editing. Project management. Service translation. Integrated translation project.

Sommaire 1 Trois défis discursifs au cœur du projet TSA (Traduction de site web vers l'anglais). – 1.1 Comment reconnaître et reproduire un discours de spécialiste quand on n'est pas soi-même spécialiste du domaine ?. – 1.2 Comment apprendre et appliquer les normes d'un 'guide de style' ?. – 1.3 Comment respecter et restituer la 'voix institutionnelle' du client ?. – 2 Le défi des genres discursifs spécifiques : l'exemple de la 'page d'accueil'. – 2.1 Le défi des normes rédactionnelles : l'exemple du 'Guide de style de la DGT'. – 2.2 Le défi des éléments du langage : l'exemple du *person-first language*. – 3 Intégrer les technologies de la traduction automatique neuronale dans l'enseignement de la pratique outillée de la traduction. – 4 Faire converger biotraduction et compétences technologiques. – 5 Conclusion.

Trop longtemps, la réflexion sur la traduction et même sur les métiers qui constituent cette profession a laissé de côté les aspects et les besoins économiques et sociaux. Ce n'est plus le cas aujourd'hui, et l'on constate que la composante 'prestation de service' est de plus en plus prise en compte, d'abord sous l'influence de certains masters innovants, puis par l'intermédiaire de référentiels de compétences successifs, établis en particulier dans le cadre du Master européen en traduction (EMT 2009 ; 2017 ; 2022). C'est à vrai dire une nécessité, car le développement rapide et la métamorphose constante des métiers en question suppose une formation qui perçoive, qui intègre, et même qui soit en mesure d'anticiper ces nouvelles tendances. Cela passe par deux formes de transversalité. D'une part, il s'agit de nouer des liens étroits entre formation et recherche ; d'autre part, il importe de mettre en cycle, à l'intérieur-même des formations, différents enseignements dans le cadre de projets intégrés. Le pionnier en la matière est le célèbre Tradutec, né à Rennes dans les années 1990 à l'instigation de Daniel Gouadec (2002), et poursuivi ensuite dans toute l'Europe (Hernandez Morin, Barbin, Phuez-Favris 2017). Mais beaucoup d'autres projets ont suivi, plus ou moins inspirés de ce pionnier, dans différents masters, en Europe et en France. C'est d'un tel projet que traite le présent article.

Tous les projets de ce type partagent un certain nombre de traits communs. Tout d'abord, l'idée – qui est aujourd'hui une évidence mais qui ne l'était pas il y a encore 15 à 20 ans –, que traduction humaine et traduction automatique n'existent pas dans des mondes séparés. Ensuite, le désir de penser une formation qui soit réellement professionnalisante, c'est-à-dire qui permette aux diplômés de s'intégrer avec succès dans le monde du travail, en qualité d'indépendants comme de salariés, avec une juste rémunération et un véritable pouvoir de négociation – une agentivité (*agency*), pourrait-on dire. Enfin, le souci de peser dans les évolutions à venir du marché. Trop longtemps, la traduction a été perçue – et s'est elle-même perçue – comme intervenant en bout de chaîne, sans possibilité d'influence sur son environnement et dans une forme d'invisibilité (Venuti 1995-2008). Ce n'est plus le cas aujourd'hui : professionnels, chercheurs, formateurs ont compris qu'ils pouvaient être force de proposition, dès lors qu'ils s'organisaient entre eux : agentivité, là encore.

C'est donc sur ce fond que se dessine notre projet d'intégration de plateformes de traduction automatique neuronale dans l'enseignement de la traduction spécialisée. Dans les paragraphes qui suivent, nous entendons en présenter les grandes lignes, revenir sur certaines des difficultés professionnelles auxquelles un tel projet permet de confronter les apprenants et enfin décrire la plate-forme technique mise en place pour mener à bien cette entreprise.

1 Trois défis discursifs au cœur du projet TSA (Traduction de site web vers l'anglais)

Le projet TSA (Traduction de site web vers l'anglais) est un projet de traduction pédagogique, mené dans le cadre de la deuxième année du Master ILTS (Industrie de la langue et traduction spécialisée, Université Paris Cité). L'objectif est de rendre en un anglais de qualité une sélection de pages d'un site internet francophone. Deux cours du master sont consacrés à ce projet (Traduction vers l'anglais ; Outils de traduction de site). La nature transversale de ce dernier implique plusieurs compétences : maîtriser les nouveaux outils de la traduction automatique neuronale (TAN), organiser une équipe de post-éditeurs/réviseurs, mettre en place plusieurs cycles de révision, gérer les communications entre les différents partenaires, etc.

Nous avons exploré précédemment les particularités du projet TSA (Gledhill, Zimina 2019 ; Zimina, Gledhill 2021a ; 2021b), en soulignant les défis soulevés par la gestion d'un projet de grande envergure, et en proposant quelques solutions pédagogiques (adoption d'un flux de travail innovant, évaluation par formulaires, etc.).

Dans cette contribution, nous nous concentrons d'abord sur trois défis qui restent - pour l'instant au moins - insolubles. En effet, leur nature repose sur l'incapacité de la technologie actuelle à prendre en compte des phénomènes discursifs complexes pour sensibiliser les étudiants aux exigences d'une prestation vraiment professionnelle avec des cycles de révision multiples dans un projet intégré. Nous montrons ensuite comment faire converger traduction humaine et compétences technologiques pour former les traducteurs et traductrices de demain.

1.1 Comment reconnaître et reproduire un discours de spécialité quand on n'est pas soi-même spécialiste du domaine ?

Il s'agit dans ce projet de traduire des genres textuels peu rencontrés dans les formations de langues traditionnelles. Puisque les partenaires choisis pour le projet TSA sont souvent des organismes à but non lucratif, les étudiants rencontrent une grande diversité de genres discursifs politiques (communiqués de presse, feuilles de route, manifestes, etc.), ainsi que des textes institutionnels (biographies, comptes rendus, procédures, etc.). Les sites web sont également composés de plusieurs micro-genres spécifiques avec lesquels les étudiants sont parfois à l'aise dans leur première langue, mais qui peuvent poser des difficultés dans la langue cible (la majorité sont francophones, et ils traduisent ici vers l'anglais). Par ailleurs - les comparatistes le savent très bien - la distribution des spécificités

lexico-grammaticales dans un genre textuel peut varier considérablement entre le français et l'anglais, *a fortiori* pour des genres qui évoluent rapidement sur internet (Askehave, Ellerup Nielsen 2005). Si l'approvisionnement de ces modes discursifs nécessite des compétences rédactionnelles avancées chez les biotraducteurs, la TAN – au moins dans sa version non-entraînée – rencontre donc également des difficultés comparables. Nous présenterons plus loin quelques exemples de ce phénomène, ainsi que des propositions de remèdes.

1.2 Comment apprendre et appliquer les normes d'un 'guide de style' ?

Pour rendre un texte de qualité, les traducteurs professionnels doivent se familiariser avec les recommandations d'un *Style Guide* (Guide de style, Code rédactionnel, etc.). Cela implique non seulement la reconnaissance de plusieurs centaines de règles essentiellement arbitraires (en langues source et cible), mais aussi la capacité d'intervenir afin d'assurer la conformité du texte dans son ensemble : une tâche délicate lorsqu'il s'agit d'un site web avec une arborescence conséquente. Là aussi, nous verrons que la TAN non-entraînée est loin de pouvoir appliquer ces consignes de manière systématique.

Les deux premiers défis concernent des complexités discursives dans une perspective essentiellement 'cibliste'. Le troisième concerne un problème plutôt 'sourcier' : celui de l'hétérogénéité discursive dans la langue de départ.

1.3 Comment respecter et restituer la 'voix institutionnelle' du client ?

Les organismes pour lesquels nous traduisons adoptent souvent des postures politiquement engagées, et donc rhétoriquement 'marquées', pour employer les termes de la grammaire systémique (Halliday 1976). Dans cette perspective, un discours marqué correspond à une reconfiguration planifiée des formes terminologiques et phraséologiques afin d'être en rupture avec la parole instituée habituelle : il s'agit d'une stratégie de communication 'disruptive'.¹ Encore une fois, les apprentis-traducteurs doivent apprendre à gérer ce genre de complexité, puisque la TAN n'est pas en mesure de reconnaître ces discours atypiques. Par ailleurs, la TAN est même susceptible de

¹ Voir par exemple la page d'USC School for Communication and Journalism, University of Southern California : <https://communicationmgmt.usc.edu/blog/co-innovation-disruption-communication-digital-world>.

‘lisser’ certains discours, et, ce faisant, de les détourner de l’objectif initial (Raus 2013). Dans les sections suivantes, nous proposons une solution consistant à composer et à analyser des corpus de textes comparables : un remède qui demande un certain travail de prospection et de négociation avec les auteurs du site.

Nous nous proposons donc d’illustrer ces trois défis en présentant quelques exemples tirés d’un projet réalisé en deux années universitaires (2020-21 et 2021-22). Pendant la période 2020-21, nous avons traduit une sélection de pages du site de l’*Association des diabétiques de l’Île-de-France* (ADF), créant ainsi une mémoire de traduction (nous revenons sur ce procédé vers la fin de cette contribution). Pendant l’année 2021-22, nous avons exploité ces données constituant cette mémoire en traduisant une sélection de pages du site de la Fédération française des diabétiques (FFD), organisme qui chapeaute cette association locale.



Figure 1 Traduire un micro-genre : la page d’accueil de la FFD

2 Le défi des genres discursifs spécifiques : l’exemple de la ‘page d’accueil’

Pour illustrer notre premier défi (Comment traduire des discours spécialisés, des genres textuels spécifiques, etc. ?), nous proposons d’examiner le micro-genre tout à fait familier mais aussi assez particulier de la ‘page d’accueil’. Ainsi se présente la page de la FFD dans la figure 1 ci-dessus.

Comme nous l’avons indiqué, la traduction d’un site web implique une très grande variété de micro-genres présents sur la même page :

- éléments du péritexte, comme les titres, les éléments de menus, etc.
- liens vers des sites web interconnectés
- formulaires de recherche, liens vers des documents administratifs
- annonces, listes, blogs, notices
- informations méta-textuelles (directives, mentions légales, informations de contact, etc.).

Sur la page d'accueil de la FFD, on ne peut pas manquer de remarquer les spécificités lexicales et grammaticales des différents titres et 'éléments de menu'. On note en particulier une préférence nette pour des propositions déclaratives/performatives écrites à la première personne.

Pour la même page, il est intéressant d'examiner les premiers résultats de la traduction automatique en anglais [fig. 2] ici sans aucune opération de post-édition ou de révision :

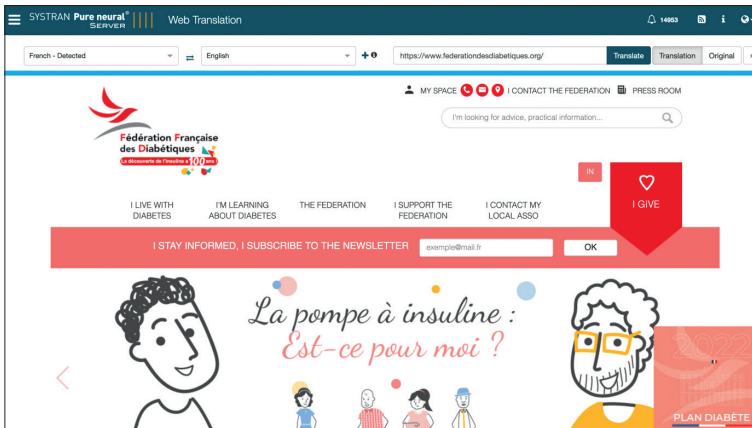


Figure 2 Premiers résultats de la traduction 'machine' réalisée avec SYSTRAN sur le serveur de traduction installé à Université Paris Cité

Comme on peut le voir en haut de la figure 2, nous avons utilisé une plate-forme de traduction automatique professionnelle (SYSTRAN Pure Neural Server, SPNS) pour générer cette version initiale. La plateforme SPNS intègre l'état de l'art des technologies de traduction avec un moteur de traduction automatique neuronale. L'une des fonctionnalités de ce système permet une vue contextuelle spécifique appelée simplement *Web Translation*. Cette interface permet aux biotraducteurs de modifier et/ou commenter n'importe quel segment de texte cible tout en ayant accès au contexte de la page source et à sa traduction.

Il convient de noter que cette page a été traduite à l'aide d'un modèle générique (FR > EN0) qui n'avait pas été entraîné pour un domaine ou un projet spécifique. Or cette version 'non-entraînée' (que nous annotons EN0 'version anglaise sans révision') comporte plusieurs propositions de traduction de mauvaise qualité. Par exemple, l'abréviation informelle pour 'Association' est 'Asso' en français, mais cette forme n'est pas acceptable en anglais. On constate également des erreurs plus subtiles, notamment les éléments de menu exprimés à la première personne (*I live with diabetes, I'm leaning about diabetes, I contact my local Asso*, etc.). De même, l'un des principaux éléments du menu 'JE DONNE' est traduit par 'I GIVE'. Ces formes s'écartent sensiblement des pratiques discursives des sites web comparables en anglais.

Pour des raisons pédagogiques, nous ne proposons pas de modèles entraînés à ce stade du projet. Et au lieu de proposer des corrections immédiates, nous présentons les résultats initiaux de la TAN tels quels, en posant la question suivante aux étudiants : s'agit-il des traductions optimales pour ces segments ? Dans les cas où les étudiants ont des doutes, ils sont invités à reconsidérer les autres choix de traduction sur la même page.

Examinons le cas particulier du discours à la première personne. Comme nous venons de le voir, les auteurs du site web français ont choisi d'adopter cette forme discursive sur l'ensemble de la page. Il s'agit d'une construction tout à fait dynamique, mais elle n'est pas la seule possibilité (comme on peut le voir dans des formes alternatives comme « ma vie avec le diabète », « vivre avec le diabète », etc.). Si ce choix est tout à fait cohérent en français, on peut constater que les sites web comparables du monde anglophone n'adoptent pas les mêmes constructions. Voir par exemple la page d'accueil de *Diabetes UK* [fig. 3] et, à titre de comparaison, la page d'accueil de *Diabetes Ireland* [fig. 4]. On note l'utilisation dans les deux cas de constructions impératives (*Donate, Get involved, Fundraise for us*) ou l'emploi d'un participe (*Preventing Type 2, Living with Diabetes*).

Un bref aperçu de pages comparables en Afrique du Sud, Australie, Nouvelle-Zélande, etc. révèle des tendances similaires. Il est même possible d'observer une alternance grammaticale assez subtile entre les formes impératives en anglais, par exemple *Support* et *Contact* (apparemment utilisées pour les liens vers les pages procédurales), et des formes où l'on préfère un participe, comme dans *LIVING WITH, Preventing type 2* (utilisées pour les liens vers les pages plus descriptives). Ce genre de contraste démontre que nous sommes en présence d'un micro-genre, avec ses conventions rhétoriques et ses préférences lexico-grammaticales particulières dans chaque langue. Nous avançons l'hypothèse que ce type de subtilité échappe aux systèmes de TAN, au moins en ce qui concerne les systèmes actuels non entraînés sur ce mode de discours très particulier.

Nicolas Froeliger, Christopher Gledhill, Maria Zimina-Poirot
Intégrer des plateformes de traduction automatique neuronale



Figure 3 Une page d'accueil comparable en anglais (Royaume-Uni).
Source : <https://www.diabetes.org.uk>

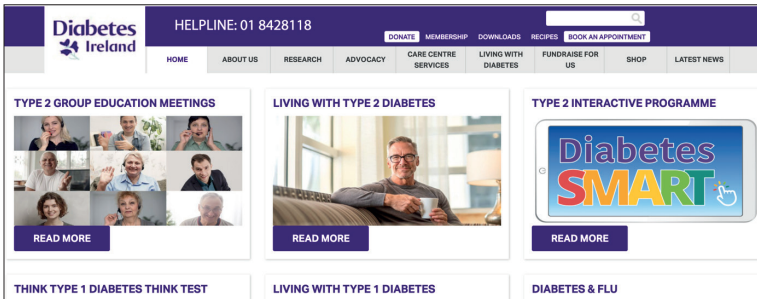


Figure 4 Une page d'accueil comparable en anglais (Irlande). Source : <https://www.diabetes.ie>

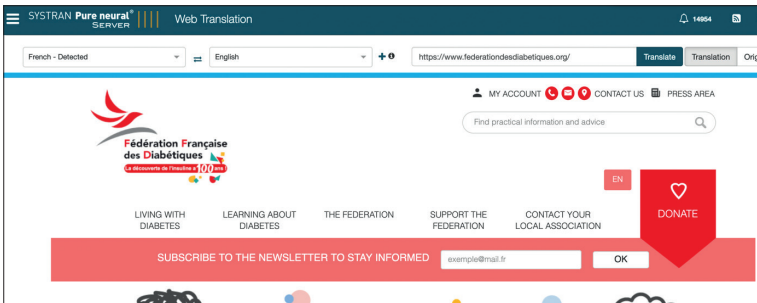


Figure 5 La page d'accueil après un cycle de révision (annotée EN0 > EN1).
Traduction générée avec le serveur de traduction installé à Université Paris Cité

Enfin, on peut voir à la figure 5 la même page en anglais après un cycle de révision (que nous annotons donc 'EN1'). On peut constater ici que les étudiants ont recherché d'autres possibilités phraséologiques, et sont arrivés à des reformulations assez différentes (par exemple *I live with diabetes* > *Living with diabetes, I contact my local Asso* > *Contact your local association, I give* > *Donate, etc.*).

2.1 Le défi des normes rédactionnelles : l'exemple du 'Guide de style de la DGT'

Le deuxième défi que nous avons identifié correspond à l'identification et l'application d'un 'guide de style'. Le guide de style (aussi appelé *Code de rédaction, Feuille de style, Normes éditoriales, etc.*) est un document prescriptif créé par une institution dans le but de définir l'ensemble des normes linguistiques et rédactionnelles auxquelles tout document est censé se conformer avant publication définitive. Ces documents n'ont pas souvent fait l'objet d'études scientifiques, comme le remarque Washbourne (2012, 10) :

The style guide is underutilised as a teaching tool in the translation classroom, in tandem with the fact it is both undertheorised and understudied as an empirical artifact and as a tool of the trade. Much work remains to be done in Translation Studies on this resource.

Cela dit, il existe quelques études en analyse de discours (Cameron 1996 ; Attolino 2007) s'intéressant aux partis pris idéologiques que peuvent receler les guides de style édités par les grands journaux nationaux. Plus proche de notre propos, Svoboda (2013) et Svoboda ; Biel ; Łoboda (2017) ont effectué une série d'études sur la diversité des ressources linguistiques publiées par les instances de l'UE ainsi que sur le degré de conformité atteint par les documents traduits par la Direction générale de la traduction (DGT ; Commission européenne 2022), notamment par rapport aux différentes versions linguistiques du « Code rédactionnel interinstitutionnel ».

Étant donné son intérêt scientifique ainsi que l'importance accordée par les institutions à ce genre de document, nous estimons qu'il est légitime d'utiliser les recommandations de la DGT pour le projet TSA. Nous indiquons ainsi aux étudiants que ces recommandations sont obligatoires pour la réalisation du projet TSA, mais qu'elles sont aussi modifiables (voire remplaçables, dans la mesure où elles sont essentiellement arbitraires et varient d'une institution à une autre). Le besoin de respecter un guide de style est aussi inscrit dans l'accord de principe destiné aux clients. Pour les textes en anglais, les consignes de la DGT existent sous format pdf/papier (avec

des recommandations spécifiques aux documents imprimés) : *English Style Guide* (plusieurs éditions à partir de 2016)² et sous forme d'un wiki (avec des recommandations plus adaptées aux documents en ligne) : *Web Writing Guide*.³

Pour désigner l'intégration du guide de style dans le *workflow* du projet, nous utilisons le terme 'normalisation'. Il s'agit d'un cycle de révision entier consacré à l'harmonisation orthographique, typographique, etc. de l'ensemble du projet en conformité avec le guide de la DGT.

Enfin, le guide de style est également intégré dans l'évaluation des étudiants à la fin de l'année sous la forme d'un document que nous appelons le *Reporting Form*. Ainsi dans une section de ce document nous demandons aux étudiants de classer et de quantifier un échantillon des modifications apportées aux pages dont ils avaient la responsabilité éditoriale. Ces informations permettent non seulement de cartographier la typologie des révisions, mais aussi d'établir une chronologie assez précise de la prise en compte de telle ou telle modification pour chaque cycle du projet.

Dans le Tableau 1 nous présentons une typologie sommaire des différents phénomènes signalés dans le corpus des *Reporting Forms*. Pour des raisons pédagogiques, nous utilisons une version simplifiée du modèle systémique pour classer chaque consigne ou modification.

Quantitativement, il n'est pas étonnant d'observer que la plupart des erreurs ou modifications répertoriées se situent en bas de cette échelle. Les étudiants éprouvent souvent plus de difficulté à trouver des exemples 'pragmatiques' en haut du tableau, mais avec un peu d'entraînement il leur est souvent possible d'identifier une poignée de phénomènes énonciatifs/contextuels sur chaque page du site web.

Plus généralement, cet exercice de classification démontre qu'il existe une très grande diversité de normes présentées dans les guides de style. Et quelle que soit la distribution ou la nature de ces modifications, on note que toutes les consignes concernent des formes d'expression alternatives, c'est-à-dire autant de points de décision éditoriaux qui seraient dans leur ensemble difficiles (voir impossibles) à prendre en charge automatiquement par un moteur générique de TAN.

² *English Style Guide* est disponible sur le site web officiel de l'Union européenne : https://ec.europa.eu/info/sites/default/files/styleguide_english_dgt_en.pdf.

³ *Web Writing Guide* est publié sur le site web officiel de l'Union européenne : <https://wikis.ec.europa.eu/display/WEBGUIDE/02.+Web+writing+guidelines>.

Tableau 1 Typologie des recommandations établies par le guide de style de la DGT

<i>Strate linguistique</i>	<i>Consignes dans le Guide rédactionnel</i>
Situation énonciative / contextuelle	Décisions éditoriales d'ordre culturel (écriture inclusive, néologismes « politiques », noms institutionnels, sigles), modifications d'ordre contextuel (suppression d'informations non pertinentes, actualités éphémères, « localismes », etc.)
Stratégies sémantiques / discursives	Recommandations sur la charge informative du texte (explication du « jargon », longueur des phrases, segmentation en paragraphes et sous-paragraphes etc.)
Structures lexico-grammaticales	Consignes sur l'orientation de la phrase (opposition actif/passif), sur la simplification des groupes nominaux complexes, sur les accords morphologiques (pluriel variable en ENG), etc.
Substance phonologique / graphique	Règles permettant de trancher entre formes alternatives (orthographe, ponctuation, typographie, format de dates), etc.

2.2 Le défi des éléments du langage : l'exemple du *person-first language*

Notre troisième défi concerne les stratégies à adopter afin de respecter les spécificités rhétoriques des organismes avec lesquels nous travaillons. Comme on pourrait s'y attendre, nous encourageons nos étudiants à considérer le contexte culturel et idéologique des pages web que nous traduisons, surtout lorsqu'il s'agit d'associations caritatives. Ainsi, si les sites web de ces organismes comportent souvent des contenus administratifs (formulaires d'adhésion, statuts, mentions légales, etc.), plusieurs pages ont une teneur plus politique (communiqués de presse, manifestes, rubriques sur des questions de société, etc.).

Il est particulièrement pertinent pour le projet TSA d'observer que certaines associations élaborent des politiques de communication explicites. En effectuant la recherche documentaire pour ce projet, nous avons remarqué que l'association australienne pour les diabètes (*Diabetes Australia*) consacre une rubrique entière à la communication dans ses *Position Statements* (« Manifestes », « Déclarations de principe »), comme on peut le voir à la figure 6 ci-dessous.

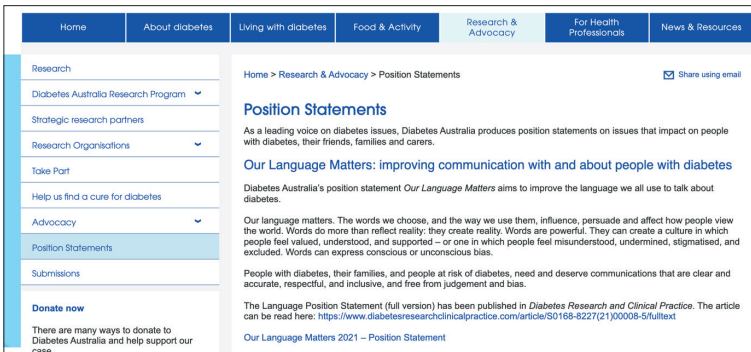


Figure 6 « Activités de sensibilisation » et « Déclaration sur le langage ». Source : <https://www.diabetesaustralia.com.au>

Ce constat nous a mené à consulter d'autres sites d'associations de diabètes anglophones (Afrique du sud, Australie, Etats-Unis, Irlande, Nouvelle-Zélande). Il nous semble que les associations anglophones ont dû se concerter sur ces questions ; on constate ainsi de très fortes correspondances, avec des références croisées à des articles de recherche et à des manifestes rédigés par les mêmes équipes de co-auteurs.

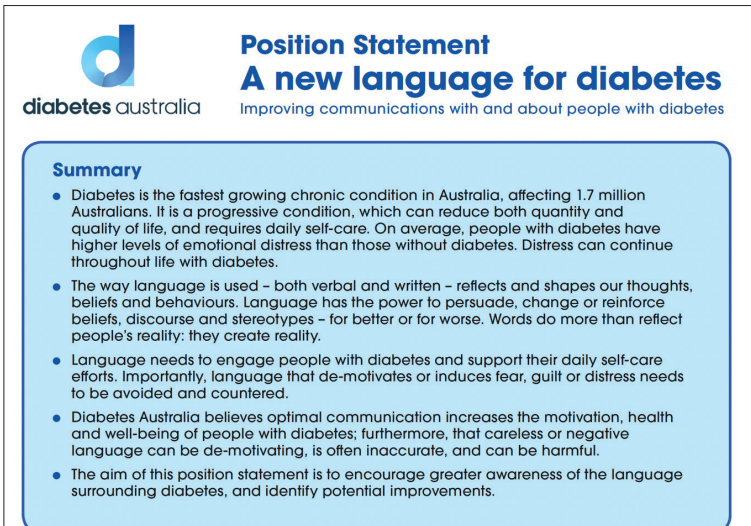


Figure 7 Manifeste : une nouvelle langue pour le diabète. Source : <https://www.diabetesaustralia.com.au>

Pour revenir au site *Diabetes Australia*, nous trouvons plusieurs documents portant plus précisément sur la politique linguistique de l'association. Ainsi, à la figure 7 on trouve une liste de 'déclarations déontologiques'

ainsi qu'un appel à adopter une 'nouvelle langue' pour combattre les difficultés rencontrées par les personnes vivant avec le diabète.

La plupart de ces sites font référence à des articles plus détaillés pour appuyer leurs arguments. Parmi les références les plus citées, on trouve Speight et al. 2012 (*A New Language for Diabetes : Improving Communications with and About People with Diabetes*) Nous en présentons un extrait annoté dans la figure 8 :

Table 1 – We can change the way we communicate about people.		
Instead of	Try ...	Why it matters
Diabetic (noun)	Person Person with diabetes Person living with diabetes	There is so much more to a person than their diabetes. The label 'diabetic' defines a person by their diabetes. It makes it easy to overlook the person's other priorities, e.g. family, work, hobbies. No-one calls a person with cancer a 'cancer', so why do it with diabetes? People with diabetes are free to refer to themselves in whatever way they wish – but others have a responsibility to use inclusive language.
Sufferer (noun)		While some people with diabetes (and its complications) find it challenging and distressing, very few 'suffer'. Labelling people with diabetes as 'sufferers' positions them as helpless victims, rather than empowering them to live well with diabetes.
Normal, Non-diabetic, Healthy (person)		Referring to people who do not have diabetes as 'normal' implies that people with diabetes are 'abnormal'. This is stigmatising.
Patient (noun)		'Patient' comes from the Latin 'patiens', meaning 'to suffer' or 'to bear'. This label places the person in a passive role, tolerating their situation, awaiting an 'expert' or 'health provider' to intervene. It does not reflect the realities of living with and managing diabetes every day. Patients are people. People are individuals, with their own preferences, priorities and lives beyond diabetes.

Figure 8 Recommandations linguistiques tirées du *Manifeste sur le langage* (Speight et al. 2012)

Il est intéressant d'examiner certaines consignes de plus près. Comme on le voit à la figure 8, les auteurs formulent diverses recommandations portant sur le *first-person language*, concept que l'on peut traduire plutôt littéralement par 'langage qui met les personnes devant'). Par exemple les auteurs proposent de remplacer le mot **diabetic* par *> person with diabetes* ainsi que **diabetics > people with diabetes*.

Si cette stratégie de reformulation relève d'un mécanisme syntaxique, d'autres consignes visent plutôt à ré-orienter certains termes-clés sur le plan sémantique. Ainsi on propose la reformulation **diabetic patient > adult diabetic* pour supprimer la transitivité passive de patient ; dans la reformulation **health provider > health professional* on cherche à remplacer l'agentivité de *provider* par un attribut non transitif *professional*, etc. De même, pour certaines constructions verbales, on propose la reformulation **treating diabetes > managing diabetes* : il s'agit ici de remplacer l'aspect accompli du verbe ('guérir/en finir avec le diabète') par une forme non-télique ('gérer/contenir le diabète'). Plus loin dans le texte les auteurs proposent d'éviter des termes exprimant des évaluations explicites jugées trop négatives, telles que **disease*, **normal*, **sufferer*, etc.

Or les préconisations de l'article de Speight et al. (2012) ne constituent pas en elles-mêmes des recommandations formelles ou structurées, contrairement à celles proposées dans le guide de la DGT. Certes, les auteurs font appel à des études en psychologie ou en communication médicale (par ex. Dunn, Andrews 2015.) Dans son ensemble, ce document (Speight et al. 2012) correspond plus à une déclaration de principe qu'à un véritable guide de style. Il nous paraît particulièrement important de souligner ce point, car si nos

étudiant.es sont encouragés à observer des exemples attestés dans des corpus d'usage pour justifier tel ou tel emploi, ici nous sommes face à une logique très différente : l'objectif est de changer les habitudes discursives de la majorité des énonciateurs. Les solutions proposées par les auteurs s'inscrivent ainsi dans une tradition de planification linguistique que l'on peut comparer à la création d'une langue contrôlée (voir à ce propos Gledhill et al. 2019).

Enfin, il faut aussi prendre en compte un autre facteur, qui risque de complexifier encore le travail des biotraducteurs. Comme l'ont signalé plusieurs observateurs, par exemple Halmari (2011), il n'existe pas de consensus parfait en ce qui concerne certaines propositions de reformulation. Ainsi, si la plupart des associations de diabétiques dans le monde anglophone semblent préférer des formulations dites *person-first*, il existe une tendance inverse qui consiste à préférer des formulations *identity-first*. Il est donc possible que dans d'autres domaines les préférences linguistiques se trouvent inversées, ce qui souligne la nécessité d'échanger régulièrement avec les membres de ces associations afin de déterminer leurs positions sur ces questions.

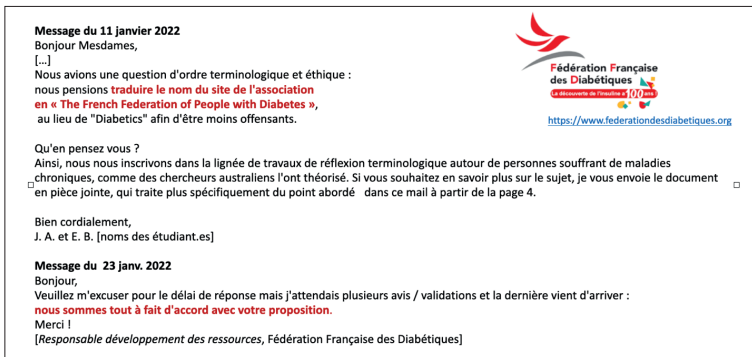


Figure 9 Échange entre gestionnaires du projet et responsables de l'association sur le nom de la FFD en anglais

Pour illustrer ce point, il suffit d'examiner une question soulevée par la dénomination même de la *Fédération française des diabétiques* : comment traduire le nom de l'association en anglais tout en respectant les consignes concernant le *first-person language* ? Si on suit les recommandations de Speight et al. (2012), la solution qui s'impose serait *The French Federation of People with Diabetes* (contrairement à des alternatives comme *The French Federation of Diabetics/Diabetic People*, etc.). Il s'agit là d'un très bel exemple de problème politique nécessitant un dialogue entre les biotraducteurs et les responsables de l'association. Comme on le voit dans la figure 9 ci-dessus, l'équipe des biotraducteurs a proposé cette solution aux responsables de l'association.

Il est intéressant de noter que les responsables de l'association avouent avoir hésité avant d'accepter cette proposition, ce qui indique que le message des traducteurs a dû susciter un certain débat. Le fait que la directrice de la communication a répondu en personne, et avec tant d'enthousiasme, a certainement été très bien reçu par nos étudiants. Cet échange révèle également que le fait de traduire peut mener 'par boucle de rétroaction' à des réflexions plus profondes chez les clients sur leurs propres choix discursifs. Ainsi, si une formulation *first-person* est envisageable pour la version anglaise du site, ne devrait-on aussi modifier le nom de l'association sur les pages en français ? À l'heure où nous rédigeons ces lignes, nous attendons encore les délibérations au sein de la FFD pour leur avis.

En définitive, et au-delà des solutions de traduction que nous avons évoquées ici, il nous semble que la TAN, au moins sous sa forme non entraînée, passerait entièrement à côté de ce type de considérations.

3 Intégrer les technologies de la traduction automatique neuronale dans l'enseignement de la pratique outillée de la traduction

Compte tenu des défis discursifs identifiés, comment intégrer les technologies de la traduction automatique neuronale dans la pratique outillée de la traduction, et comment former les traducteurs spécialisés de demain ? Nous l'avons vu, la réflexion sur ces questions a guidé la conception du projet pédagogique TSA, qui s'articule autour de plusieurs blocs de compétences. Au fil des cours, nos étudiants doivent acquérir des compétences en expertise linguistique et en gestion de projet ; ils doivent apprendre à valider la qualité du résultat TAN, tout en se conformant aux exigences de qualité spécifiques à chaque client ou projet. Les connaissances en ingénierie linguistique, avec notamment la spécialisation de la TAN à l'aide de ressources adaptées au projet (modèles spécialisés, dictionnaires, mémoires de traduction), sont développées parallèlement aux compétences professionnelles au cours du suivi des tâches associées à la révision, le travail collaboratif, la livraison, etc.

Il est parfois difficile pour les étudiants d'apprécier ces différents aspects simultanément. Une des solutions consiste à proposer des rôles spécifiques pour chacun d'entre eux et à établir une liste des responsabilités avec des indications précises concernant les livrables. Parallèlement, nous avons élaboré un système d'optimisation du rendement qualitatif que nous appelons en anglais *Qualitative Translation/Revision Workflow* (Zimina, Gledhill 2021b).

L'architecture 'outils' utilisée dans le cadre du projet TSA repose d'abord sur la spécialisation de la TAN. Le développement de cette approche en cours est possible grâce à l'utilisation intensive des outils

mis à disposition des étudiants et des enseignants d'Université Paris Cité dans le cadre de la plate-forme PAPTAN (Plate-forme pour l'apprentissage profond pour la traduction automatique neuronale).⁴

Tableau 2 Données d'apprentissage : segments révisés dans la mémoire de traduction TSA

Segment en français	Segment en anglais
Je contacte le Diabète LAB	Contact the Diabète LAB
Je contacte mon asso locale	Contact your local association
Je découvre les témoignages	Read testimonies
Je donne	Donate
Je donne pour la recherche	Donate to research
Je m'abonne	Subscribe
Je m'abonne à Équilibre	Subscribe to Équilibre
Je m'informe sur le diabète	Learning about diabetes
Je m'inscris à la newsletter	Subscribe to the newsletter
Je partage mon histoire	Share my story
Je pratique une activité physique	Staying active
Je reste informé, je m'inscris à la newsletter	Subscribe to the newsletter to stay informed
Je soutiens la fédération	Support the Federation
J'évite les complications	Avoiding complications
Je trouve une idée de menu	Meal ideas

Compte tenu de l'état de l'art des technologies TAN, c'est par le biais de la spécialisation que l'on peut agir sur le résultat final de la traduction. Ce type de processus demande des efforts et des compétences, en particulier une compréhension approfondie des données et des ressources linguistiques disponibles, telles que les corpus d'entraînement, les modèles spécialisés disponibles, les mémoires

⁴ La plate-forme PAPTAN (<https://plateformes.u-paris.fr/pure-neural-server-clillac-arp-plateforme-de-traduction-automatique-sur-serveur-universite>) est déployée à Université Paris Cité avec l'aide du partenaire industriel SYSTRAN et le soutien financier d'Université Paris Cité. PAPTAN bénéficie de l'accompagnement logistique de la DSIN (Direction du système d'information et du numérique) de cet établissement. La plateforme prévoit plusieurs niveaux d'accès en fonction du profil de l'utilisateur et sa maîtrise des technologies TAN (allant de l'utilisation de la TAN à l'entraînement de modèles spécialisés avec accès aux GPUs de la plate-forme). Les utilisateurs-étudiants M2 ILTS utilisent deux interfaces : Plat-eforme PNS-UP (SYSTRAN Pure Neural® Server) pour la traduction automatique spécialisée et SYSTRAN Model Studio Lite pour l'entraînement de modèles spécialisés.

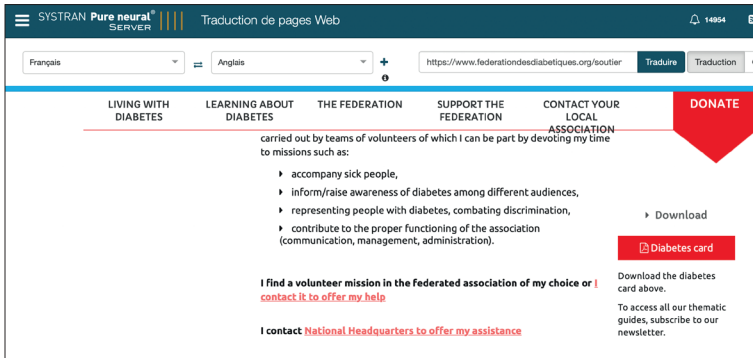


Figure 10 Résultats de la TAN après la spécialisation : vers un nouveau cycle de révision et customisation du modèle (serveur SYSTRAN, Université Paris Cité)

de traduction, la terminologie, etc. Ainsi, nos étudiants doivent être impliqués à chaque étape du processus de mise en œuvre de la TAN spécialisée. Au final, un profil TAN spécialisé et personnalisé permet une amélioration raisonnable de la qualité de la traduction notamment grâce aux données utilisées pour l'apprentissage (les données de la révision présentées dans le Tableau 2).

Pourtant, même après que les modèles TAN personnalisés ont produit des résultats plus précis [fig. 10], il reste plusieurs tâches à gérer, notamment la révision des 'nouveaux' segments traduits, tels que *accompany sick people* et *representing people with diabetes*, etc., ainsi que les tâches liées à la gestion du projet, la livraison, etc.

A partir de ces observations pratiques, nous arrivons à plusieurs questions de recherche. Comment prendre en compte les spécificités discursives aux cours de l'apprentissage automatique ? Comment aller au-delà des correspondances lexicales et tenir compte des spécificités lexico-grammaticales lors de la propagation des segments enregistrés dans les mémoires de traduction aux nouvelles traductions ? Comment renforcer l'apprentissage et l'exploitation des mémoires de traduction qui pourraient devenir des corpus augmentés ? Peut-on envisager une adaptation dynamique de la TAN au cours de la révision ? Quelles sont les fonctionnalités qui manquent encore aux systèmes TAN disponibles sur le marché professionnel de la traduction ? Ce sont des questions qui se posent à nos apprentis traducteurs et à nous, enseignants-chercheurs qui travaillons dans ces domaines.

4 **Faire converger biotraduction et compétences technologiques**

Il s'agit en somme, à travers de tels projets intégrés, de faire converger biotraduction et compétences technologiques, en vue d'une meilleure intégration économique et sociale, ce qui passe par un lien étroit avec la recherche. Bref, savoir faire et faire savoir, c'est-à-dire, depuis 2011, le champ couvert par les colloques Tralogy. Et de toute évidence, vouloir une formation qui soit de plain-pied avec les évolutions de la profession suppose une adaptation constante de l'ingénierie pédagogique. De fait, enseigner la traduction professionnelle aujourd'hui est de plus en plus complexe - et de plus en plus nécessaire. Les statistiques (SFT, 2016) montrent bien qu'il existe, dans les métiers de la traduction, une montée en puissance du diplôme. Néanmoins, la professionnalisation est un impératif qui ne s'impose pas qu'aux étudiants : elle s'adresse tout autant aux formateurs de ces derniers. Il importe donc de surmonter les résistances qui peuvent se faire jour de la part des intervenants comme des apprenants. Et il faut reconnaître que celles-ci sont tout à fait compréhensibles, voire, en fonction des positionnements des uns et des autres, légitimes. Ainsi, les traducteurs peuvent être tiraillés entre les exigences de qualité humaine et la volumétrie des besoins, ce que l'on observe fréquemment dans le monde professionnel. Il est par ailleurs cognitivement difficile pour les apprenants de construire une vue d'ensemble à partir d'enseignements qui peuvent leur sembler parcellisés : telle compétence acquise dans le cadre d'un cours de terminologie ou de gestion de projet ne sera pas toujours naturellement appliquée dans un cours de traduction, où pourtant elle aurait totalement sa place. Et faire en sorte qu'elle le soit, c'est-à-dire mettre en cycle les savoirs, nécessite aussi un effort, matériel et cognitif, de la part des enseignants et des responsables des formations impliqués.

Quels arguments employer pour surmonter ces réticences ? D'abord, apporter la démonstration que biotraduction et traduction automatique, nous l'avons dit en introduction, n'existent pas isolément. Ensuite, communiquer afin de faire percevoir aux uns et aux autres l'intérêt des logiques transversales, en insistant sur les passerelles entre les différents enseignements. Enfin, sensibiliser aux exigences d'une prestation véritablement professionnelle, avec des cycles de révisions multiples. Ce qui suppose d'allier une exigence de qualité forte et la nécessité de s'arrêter un moment donné. Tout est *in fine* question de limites : celles qu'il faut franchir, et celles qu'il faut savoir accepter. L'organisation d'un projet visant à intégrer les plates-formes de traduction automatique neuronale dans l'enseignement de la traduction spécialisée s'inscrit donc bel et bien dans un espace topologique et évolutif.

5 Conclusion

La traduction, ses métiers et les formations qui y préparent opèrent dans un univers dynamique, riche en incertitudes, mais aussi porteur de multiples possibles. Il y a certes dans cet univers de quoi donner le vertige, car rien n'y est plus fixe : système galiléen. Mais le plus important de ces changements n'est pas de nature technologique : il concerne la place elle-même de la traduction dans la société. Là où cette fonction était naguère omniprésente et invisible (voire invisibilisée) :

L'Europe est née de la traduction et dans la traduction. L'Europe ne s'est fondée que sur des traductions. Et elle ne s'est constituée que de l'effacement de cette origine toute de traduction. (Meschonnic 1999, 32)

Elle est maintenant de plus en plus reconnue en tant que facteur essentiel du lien social. La traduction comme agent de changement : c'est autour de ce paradigme que doivent se reconfigurer, ensemble, la formation, la vie professionnelle, les développements technologiques et la recherche. Il s'agit bel et bien de rendre un meilleur service à la société, à l'échelle européenne et mondiale, en étant à la pointe des connaissances, avec l'ambition de le rester, ce qui passe aussi par une action résolue de formation continue et par une mise à jour permanente des formations tout court.

Sources primaires

- Dunn, D.S. ; Andrews, E.E. (2015). « Person-first and Identity-first Language : Developing Psychologists' Cultural Competence Using Disability Language ». *The American Psychologist*, 70(3), 55-64.
- Speight J. et al. (2012). « Diabetes Australia Position Statement. A New Language for Diabetes : Improving Communications with and About People with Diabetes ». *Diabetes Research and Clinical Practice*, 97(3), 425-31. <https://doi.org/10.1016/j.diabres.2012.03.015>.
- Zaks, Z. (2006). *Life and Love : Positive Strategies for Autistic Adults*. Shawnee, KS: AAPC.

Bibliographie

- Askehave, I. ; Ellerup Nielsen, A. (2005). « Digital Genres : A Challenge to Traditional Genre Theory ». *Information Technology & People*, 18(2), 120-41. <https://doi.org/10.1108/09593840510601504>.
- Attolino, P. (2007). « Style Guides and Journalistic Voice ». *Textus*, XX, 205-20.
- Cameron, D. (1996). « Style Policy and Style Politics : A Neglected Aspect of the Language of the News ». *Media, Culture and Society*, 18, 315-33.
- Direction générale de la traduction (DGT) ; Commission européenne (2016-22). *English Style Guide a Handbook for Authors and Translators in the European Commission, Union Européenne / European Union*. https://ec.europa.eu/info/sites/default/files/styleguide_english_dgt_en.pdf.
- European Master's in Translation (EMT) (2009 ; 2017 ; 2023). *European Master's in Translation – Référentiel de compétences*. https://commission.europa.eu/resources-partners/european-masters-translation-emt/european-masters-translation-emt-explained_fr.
- Gledhill, C. ; Zimina, M. (2019). « The Impact of Machine Translation on a Masters Course in Web Translation : From Disrupted Practice to a Qualitative Translation/Revision Workflow ». *Proceedings = Translating and the Computer*, 41 (London, 21-22 November 2019). Geneva : Éditions Tradulex, 60-73. <http://www.tradulex.com/varia/TC41-London2019.pdf>.
- Gledhill, C. et al. (2019). « Towards a Linguistic Definition of 'Simplified Medical English' : Applying Textometric Analysis to Cochrane Medical Abstracts and Their Plain Language Versions ». *LCM*, 91-114. <https://doi.org/10.7359/919-2019-gled>.
- Gouadec, D. (2002). *Profession : traducteur*. Paris : La Maison du Dictionnaire.
- Halliday, M.A.K. (1976). « Anti-languages ». *American Anthropologist*, 78(3), 570-84. <https://doi.org/10.1525/aa.1976.78.3.02a00050>.
- Halmari, H. (2011). « Political Correctness, Euphemism, and Language Change : The Case of 'People First' ». *Journal of Pragmatics*, 43(3), 828-40. <https://doi.org/10.1016/j.pragma.2010.09.016>.
- Hernandez Morin, K. ; Barbin F. ; Phuez-Favris, G. (2017). *Handbook for the Organisation of Intensive Translation Sessions*. Rennes : Presses de l'Université de Rennes.
- Meschonnic, H. (1999). *Poétique du traduire*. Lagrasse : Verdier.
- Raus, R. (2013). *La terminologie multilingue : la traduction des termes de l'égalité : homme/femme dans le discours international*. Louvain la Neuve : Éditions de Boeck.

- Société française des traducteurs (SFT) (2016). *Enquête sur les pratiques professionnelles des interprètes*. https://www.sft.fr/clients/sft/telechargements/file_front/51802_Enquete_interpretes_2016_RP.pdf.pdf.
- Svoboda, T. (2013). « Translation Manuals and Drafting Style Guides at the European Commission ». *Le Bulletin du CRATIL*, 10, 80-106.
- Svoboda, T. ; Biel, L. ; Švoboda, K. (eds) (2017). « Quality Aspects in Institutional Translation ». *Translation and Multilingual Natural Language Processing*, 8. Berlin : Language Science Press.
- Venuti, L. (1995-2008). *The Translator's Invisibility – A History of Translation*. 2nd ed. London ; New York : Routledge.
- Washbourne, K. (2012). « Translation Style Guides in Translator Training : Considerations for Task Design ». *Journal of Specialised Translation*, 17, 2-17.
- Zimina, M. ; Gledhill, C. (2021a). « L'impact de la traduction automatique sur les pratiques langagières et professionnelles des apprentis-traducteurs : entre apports en efficacité et menaces pour la diversité des discours ». Beacco J-C. ; Herreras J.C. ; Tremblay C. (éds), *Traduction automatique et usages sociaux des langues. Quelle conséquences pour la diversité linguistique ?* Observatoire européen du plurilinguisme, 63-81. <https://doi.org/10.3917/oep.beacc.2021.01.0063>.
- Zimina, M. ; Gledhill, C. (2021b). « Comment les étudiants révisent-ils les sorties de la traduction automatique neuronale ? Quelques réflexions sur l'application des directives de la DGT et l'impact de l'anglais lingua franca sur les processus de traduction ». *L'intelligenza artificiale per la traduzione: verso una nuova progettazione didattica? = Translating Europe Workshop* (Torino, 3 dicembre 2021).

