



HAL
open science

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.

Maria Zimina-Poirot, Christopher Gledhill

► To cite this version:

Maria Zimina-Poirot, Christopher Gledhill. 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.. Jean-Claude Beacco; José Carlos Herreras; Christian Tremblay. Traduction automatique et usages sociaux des langues. Quelles conséquences pour la diversité linguistique ?, 2021 (3), Observatoire européen du plurilinguisme, pp.63-81, 2021, Plurilinguisme. hal-03511177

HAL Id: hal-03511177

<https://u-paris.hal.science/hal-03511177>

Submitted on 4 Jan 2022

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

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

Coordonné par Jean-Claude Beacco,
José Carlos Herreras et Christian Tremblay

Traduction automatique et usages
sociaux des langues. Quelles
conséquences pour la diversité
linguistique ?



Collection **Plurilinguisme**
dirigée par l'**Observatoire européen du plurilinguisme**

Conclusions de Jean-Gabriel Ganascia



Autres parutions de l'Observatoire européen du plurilinguisme

Collection *Plurilinguisme*

Plurilinguisme, interculturalité et emploi : Défis pour l'Europe, dir. F.-X. d'Aligny, A. Guillaume, B. Nieder, F. Rastier, C. Tremblay et H. Wismann, 404 p., 2009, L'Harmattan.

Culture et plurilinguisme, coord. C. Tremblay, 189 p., mai 2016, Ed. La Völva.

Plurilinguisme et créativité scientifique, coord. P. Frath et JC Hererras, coll. Plurilinguisme, 148 p., janvier 2017, Thebookedition.com (autoédition), version papier et numérique.

Plurilinguisme et enseignement du français en Afrique subsaharienne, coord. Koffi Ganyo Agbefle, coll. Plurilinguisme, 316 p., février 2017, Bookelis, version papier.

Culture et plurilinguisme, coord. C. Tremblay, février 2017, Bookelis, version numérique.

Plurilinguisme et éducation, volume I, coord. J.-C. Beacco et C. Tremblay, 220 p., Bookelis, décembre 2017.

Plurilinguisme et éducation, volume II, coord. J.-C. Beacco et C. Tremblay, 218 p., Bookelis, décembre 2017.

Plurilinguisme, entreprises, économie et société, préface Claude Truchot, coord. Christian Tremblay, 132 p., Bookelis, février 2018.

Écoles, langues et cultures d'enseignement en contexte plurilingue africain, coord. Koffi Ganyo Agbefle et Robert Yennah, 307 p., Bookelis, mars 2018.

Langues, formations et pédagogies : Le miroir africain, coord. Koffi Ganyo Agbefle et Yelain Constant Aguessy, 527 p., Bookelis, décembre 2018.

Les parlars urbains africains au prisme du plurilinguisme : description sociolinguistique, tome I, coord. Kossonou Kouabena Théodore, Dodo Jean-Claude, Youant Yves-Marcel, Bookelis, 371 p., décembre 2019.

Méthodes et pratiques d'enseignement des langues africaines : Identification, analyses et perspectives, coord. Julia Ndibnu Messina Ethe, Pierre Frath, Bookelis, 354 p., décembre 2019.

Didactique des langues, plurilinguisme et sciences sociales en Afrique francophone : quelles places à l'interdisciplinarité ?, Coord. Koffi Ganyo Agbefle & Pascal Somé, Bookelis, 334 p., janvier 2020.

Plurilinguisme, politique et citoyenneté, coord. José Carlos Herreras, Bookelis, août 2020, 275 p.

L'enseignement-apprentissage en/des langues européennes dans les systèmes éducatifs africains : place, fonctions, défis et perspectives, coord. Koia Jean Martial Kouamé et Djoa Johnson Manda, Bookelis, 337 p., septembre 2020.

Les parlars urbains africains au prisme du plurilinguisme : description sociolinguistique, tome II, coord. Kossonou Kouabena Théodore, Dodo Jean-Claude, Youant Yves-Marcel, Bookelis, 381 p., décembre 2020.

Aux carrefours de la langue, de la littérature, de la didactique et de la société : la recherche francophone en action, coord. Koffi Agbefle, Rahma Barbara, Maman Lawan et Elhadji Yawale, Bookelis, 414 p., avril 2021.

Multilinguisme, multiculturalisme et représentations identitaires, coord. Amina Goron, Bookelis, 335 p., juin 2021

L'impératif plurilingue, 18 ans avec l'Observatoire européen du plurilinguisme, Christian Tremblay, préface de Jean Pruvost, Bookelis, 510 p., à paraître, novembre 2021

Traduction automatique et usages sociaux des langues. Quelles conséquences pour la diversité linguistique ?

Coordonné par :
Jean-Claude Beacco, José Carlos Herreras et
Christian Tremblay

Conclusions de Jean-Gabriel Ganascia

Édité par
L'Observatoire européen du plurilinguisme
(OEP)

Illustration de couverture Danielle Rivier

Commande au numéro en ligne :

<https://www.bookelis.com>

Plurilinguisme
Collection dirigée
par l'Observatoire européen du plurilinguisme

Volume 2021/3

**Traduction automatique et usages sociaux
des langues. Quelles conséquences pour la
diversité linguistique ?**

Coordonné par
Jean-Claude Beacco, José Carlos Herreras
et Christian Tremblay
Conclusions de Jean-Gabriel Ganascia

Liste des auteurs

Nicolas Bacaër (Institut de Recherche pour le Développement)

Janina Di Piero Cárdenas, Renata De Rugeris Juárez (Università degli studi di Urbino Carlo Bo (Italia)

Jacques Coulardeau (Université Paris 1 Panthéon Sorbonne)

Séraphin Personne Feykééré (Université de Bangui)

Jean-Gabriel Ganascia (Sorbonne Université, CNRS)

Christopher Gledhill, Maria Zimina (Université de Paris)

Elena Kokanova, Aleksandra Epimakhova, Maxim Berendyaev, Nicolai Kulikov, Maria Evgrafova, Natalia Pak (Université fédérale Arctique)

Claire Larssonneur (Université Paris 8)

Éric Navé (Université de Lorraine)

Maria Isabel Rivas Ginel (Université de Bourgogne Franche-Comté)

Aurélien Talbot (Université Grenoble Alpes)

Christian Tremblay (Observatoire européen du plurilinguisme)

Jean-Louis Vaxelaire (Université de Namur)

Sommaire

Introduction..... 11

Théorie de la TA..... 17

Jean-Louis Vaxelaire – Les progrès de la traduction automatique par le prisme du turc et du luxembourgeois..... 19

Aurélien Talbot – La « pensée-interprète », les appareils et la diversité linguistique..... 31

Jacques Coulardeau – Humaniser la machine à traduire ?..... 45

TA, enseignement et formation..... 61

Maria Zimina, Christopher Gledhill - 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..... 63

Janina Di Pierro Cárdenas, Renata De Rugeriis Juárez - Inteligencia artificial y SoftPower de la traducción asistida y automática: perspectivas en el proceso de enseñanza-aprendizaje de idiomas..... 83

Elena Kokanova, Aleksandra Epimakhova, Maxim Berendyaev, Nikolay Kulikov, Maria Evgrafova, Nata-

Iya Pak – Traduction automatique, un nouveau défi pour la formation des traducteurs : le cas de l’Université fédérale Arctique (Russie).....	101
Eric Navé – Qu’en pensent-ils ? Réflexion sur les représentations d’apprenants et d’enseignants d’Arabie saoudite vis-à-vis du recours aux outils de traduction automatique en contexte FLE débutant.....	115
TA et pratique de la diversité linguistique.....	131
Nicolas Bacaër - Traduire automatiquement des articles dans les sciences dites dures.....	133
Christian Tremblay - La traduction automatique dans le contexte des institutions européennes - Essai de traduction automatique comparée.....	143
María Isabel Rivas-Ginel – La traducción automática en la localización de videojuegos.....	167
Claire Larssonneur – Alexa, Siri : la diversité linguistique au prisme des agents conversationnels.....	179
Séraphin Personne Feykéré – Dynamique du sängo : Fonction, procédures et méthodologie relatives à l’emploi étendu (production automatique et lecture automatique des textes écrits).....	199
Conclusions : Jean-Gabriel Ganascia.....	211
Documents.....	225

Jean-Gabriel Ganascia - Les langues comme composants des processus de créativité.....227

Maria Zimina, Christopher Gledhill - 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

Résumé

Nous examinons l'impact de la traduction automatique (TA) sur les pratiques langagières et professionnelles et d'un groupe d'apprentis-traducteurs lorsqu'il s'agit de simuler la traduction d'un site web institutionnel. Si les outils d'aide à la traduction permettent à nos étudiant.es de gérer les aspects techniques du projet de manière efficace, la TA ne représente qu'une étape initiale dans la réalisation de ce type de projet.

Abstract

The impact of machine translation on the language and professional practices of trainee translators: between efficiency gains and threats to the diversity of discourse

Here we examine the impact of machine translation (MT) on the linguistic and professional practices of a group of trainee translators when it comes to simulating the translation of an institutional website. While machine-assisted translation allows our students to manage the technical aspects of the project effectively, MT is only an initial step in this type of project.

Resumo

Ni studas la efikojn de aŭtomata tradukado (AT) al la lingvaj kaj profesiaj praktikoj de studentaj tradukistoj dum simulata projekto, kies celo estas traduki institucionan retejon. Kvankam la AT permesas al niaj studentoj efike trakti teknikajn aspektojn, ĝi nur reprezentas komencan etapon de la projekto.

Introduction

Dans cette contribution, nous examinons l'impact de la traduction automatique (TA) sur les pratiques professionnelles et langagières

d'une équipe d'apprentis traducteurs inscrits dans deux cours de niveau Master 2 intitulés « Outils de traduction de site web » et « Traduction de site web vers l'anglais » assurés en co-enseignement.

Dans un premier temps, nous nous intéressons aux différentes stratégies à adopter lorsque ces étudiant.es sont confronté.es à la traduction d'un genre textuel qu'ils pensent connaître (la « page web institutionnelle »), dans une langue qu'ils pensent maîtriser (puisque ce sont majoritairement des diplômé.es issu.es d'une formation en Langues étrangères appliquées, LEA, ayant un bon niveau d'anglais). Ces défis nous amènent à nous interroger sur les diverses compétences techniques et professionnelles que les membres de l'équipe doivent acquérir, ainsi que les décisions qu'ils doivent prendre pendant le projet : connaissent-ils les protocoles à suivre lorsqu'on traduit les parties d'un site qui relèvent du discours politique ou du discours publicitaire ? Sont-ils en mesure d'appliquer les normes dictées par un code rédactionnel ? Dans quelle mesure devraient-ils respecter les premières solutions de traduction proposées par la TA ?

Dans un deuxième temps, nous partons de l'hypothèse qu'il est possible d'étudier l'impact de la TA en s'appuyant sur les réalisations linguistiques des étudiant.es. Nous fondons ces observations sur l'analyse textométrique d'un corpus de textes comparables composé de trois volets, qui correspondent chacun à une étape clé dans tout projet de traduction professionnelle simulé : 1) l'ensemble des textes produits par la traduction automatique, 2) les révisions proposées par les étudiant.es, 3) une réécriture de ces textes proposée par une experte en communication.

Dans la dernière partie de cette contribution, nous proposerons une typologie des défis associés à ce projet pédagogique, ainsi que des pistes de réflexion sur les solutions à mettre en œuvre dans les projets de traduction de ce type.

1. Le contexte pédagogique

Cette étude porte sur la place de la traduction automatique dans deux cours intitulés « Outils de traduction de site web » et « Traduction de site web vers l'anglais », dispensés dans le cadre d'une formation de traduction professionnelle en deuxième année de master (M2 ILTS « Industrie de la langue et traduction spécialisée », Université de Paris). Ces cours dispensés en co-enseignement ont quatre objectifs pédagogiques :

- 1) Simuler un projet de traduction de site web dans des conditions aussi proches que possible de l'exercice du métier.
- 2) Mettre en place un enseignement transversal qui développe plusieurs compétences en traduction professionnelle.
- 3) Initier les étudiant.es à l'utilisation de plateformes de traduction de sites web, en mettant l'accent sur les outils utilisés par les professionnels du secteur.
- 4) Améliorer les compétences traductionnelles et rédactionnelles des étudiant.es (majoritairement francophones) afin qu'ils puissent produire des textes de qualité professionnelle en anglais.

Nous nous concentrons ici sur un projet réalisé pendant l'année universitaire 2019-2020. Il s'agissait de traduire une sélection des pages du site web de l'Université de Paris (UP), une nouvelle université issue de la fusion de plusieurs institutions parisiennes, et dont les pages en français n'existaient pas avant 2019. Afin de déterminer les périmètres du projet, nous avons collaboré avec la Direction de la Communication (DCom) qui gère les contenus des pages de l'UP. La DCom a ensuite communiqué à nos étudiant.es une liste de pages prioritaires pour la diffusion en anglais. Avant la livraison finale du projet, une collègue anglophone de la DCom a été chargée d'effectuer une relecture de l'ensemble des pages livrées par les étudiant.es.

Les différentes étapes de la traduction peuvent être résumées ainsi :

- A. AUDIT : identification des pages prioritaires, rédaction d'un accord de principe.
- B. FR : analyse de la structure du site web publié en ligne et identification des pages rédigées en français et de leur organisation.
- C. EN0 : traduction automatique des pages FR vers l'anglais (EN), gérée par les étudiant.es.
- D. EN1 : post-édition de la traduction (EN0) par les étudiant.es.
- E. EN2 : révision de la traduction (EN1) par les étudiant.es.
- F. EN3 : correction de la traduction (EN2) par les enseignants.
- G. EN4 : réécriture après correction par l'expert(e)
- H. LIVRAISON.

Il est intéressant de remarquer la nature cyclique de ces différentes étapes. Alors que l'évaluation de la traduction met traditionnellement l'accent sur la relation biunivoque entre les segments du texte source et les segments équivalents du texte cible, l'approche cyclique que nous préconisons ici insiste sur la nature itérative des différentes activités de relecture et de révision, ainsi que sur la décomposition de ces étapes en une série de tâches ciblées (par exemple, un cycle entier peut être consacré à la mise en conformité du texte par rapport à un code rédactionnel, DGT 2021). Cette approche présente un intérêt pédagogique et scientifique : pour chacun des cycles identifiés, les versions intermédiaires sont désignées et conservées, permettant ainsi d'étudier l'évolution des choix de traduction effectués par tous les participants à chaque étape du projet. La différenciation des données par étapes (ou par cycles) joue également un rôle important dans l'évaluation des étudiant.es, comme nous le verrons dans la section qui suit.

2. Les modalités d'évaluation adaptées à la démarche pédagogique

Comment évaluer un projet de traduction complexe visant à simuler la gestion de projet et l'utilisation intense des technologies ? Au lieu de fonder notre évaluation uniquement sur la qualité des textes livrés à la fin du projet, nous demandons à chaque étudiant.e de préparer une fiche d'évaluation à la fin du cours. La fiche est divisée en trois parties (Gledhill et Zimina, 2019) et comporte des rubriques correspondant aux compétences de traduction énumérées dans le référentiel de compétences du Master européen en traduction (EMT, 2017).

Dans la première partie de la fiche, il s'agit de répertorier des informations quantitatives de base (liste des pages traduites, nombre de mots traités, nombre de cycles de révision, modifications effectuées etc.). La conservation de ce genre d'informations constitue à la fois une bonne pratique professionnelle et un outil de réflexion : cette partie de la fiche rappelle les protocoles que les traducteurs professionnels appliquent pour la préparation, la gestion et la livraison de leurs projets de traduction.

Dans la deuxième partie de la fiche, les étudiant.es doivent commenter un échantillon de problèmes linguistiques rencontrés au cours des différents cycles de post-édition/révision. Ces problèmes peuvent correspondre à diverses strates du système linguistique (problèmes contextuels, discursifs, lexico-grammaticaux, typographiques, etc.) et à divers niveaux de sévérité (coquilles, erreurs complexes, problèmes sans solution immédiate). Pour chaque problème, les étudiant.es doivent : 1) catégoriser le phénomène selon la grille d'analyse proposée pendant le cours (Gledhill et Zimina 2019) ; 2) commenter les différentes solutions adoptées. L'identification des différentes étapes du projet (EN0, EN1, etc.) s'avère particulièrement utile : l'existence de plusieurs corpus désignés par étapes (et dont les segments sont à chaque fois alignés avec ceux du texte source) permet de tracer systématiquement l'historique des modifications effectuées à chaque étape du processus.

Enfin, dans la troisième partie de la fiche, les étudiant.es doivent mener une réflexion critique sur les tâches effectuées dans le flux opérationnel (*workflow*), sur leurs expériences avec les outils et les plateformes de traduction, ainsi que sur l'utilisation des ressources présentées pendant le cours (corpus de pages web comparables, outils d'analyse de données textuelles, bases de données terminologiques, mémoires de traduction, etc.).

Il convient de souligner que ce système d'évaluation n'existe pas simplement pour générer des notes à des fins d'évaluation : à notre avis (et selon les professionnels du secteur), il est primordial que les apprentis traducteurs puissent non seulement démontrer qu'ils savent exploiter les nouvelles technologies de manière efficace, mais qu'ils peuvent également démontrer un recul critique vis-à-vis du rôle de ces outils par rapport à l'ensemble des compétences qu'ils doivent acquérir dans un projet de ce type.

3. Les particularités des cours liés au projet de traduction de site

L'objectif de nos cours est de fournir un environnement simulé dans lequel des équipes d'étudiant.es travaillent sur un projet de traduction de site web. Il s'agit de traduire du français vers l'anglais des pages web provenant d'une université ou d'un autre établissement français (une moyenne de 1 à 2 pages par étudiant, étant donné qu'il y a environ 30 à 35 participants). Au fil des années, nous avons travaillé sur des sites web tels que celui de l'Institut de Physique du Globe de Paris (IPGP) : <http://www.ipgp.fr>, le site de l'école doctorale STEP'UP : <https://ed560.ed.univ-paris-diderot.fr>, le site du Master Mathématiques et Applications (Sorbonne Université) : <https://www.ljll.math.upmc.fr/MathModel>, etc. Nous ne sélectionnons généralement que les pages des sites qui ont des contenus textuels suffisamment développés : les textes font au moins 200 mots, et les genres concernés vont des rapports scientifiques (descriptions d'instruments, rapports d'activité scientifique, etc.) aux annonces administratives (annonces d'emploi, appels à candidature, instructions

pour l'inscription en doctorat, appels à manifestations scientifiques, etc.)

Nous demandons à nos « clients » (donneurs d'ordre) de reconnaître le travail de traduction/révision effectuée par l'équipe par le biais d'une mention apposée en marge des pages (*Translation carried out by the Masters 2 ILTS team*). Il s'agit notamment de sensibiliser nos partenaires à l'importance de leur collaboration active avec l'équipe ILTS : la disponibilité et la réactivité de toutes les parties sont fondamentales pour le bon déroulement du projet.

Comme indiqué ci-dessous, chaque étudiant a un ou plusieurs rôles au sein du projet :

Responsable de communication : communication entre les enseignants, l'équipe de traduction et les propriétaires du site web.

Traducteur : traduction/post-édition (en binôme avec le réviseur).

Réviseur : post-édition/révision (en binôme avec le traducteur).

Responsable de l'audit : identification des pages à traduire en priorité

Rédacteur de contrats : signature de l'accord de principe avec les propriétaires de site(s) web.

Terminologue(s) : gestion des ressources linguistiques associées au profil de traduction sur la plateforme de traduction collaborative, harmonisation des pages, contrôle qualité.

Gestionnaire de tâches : découpage du travail en tâches, suivi de la réalisation.

Nous utilisons une plateforme de traduction automatique professionnelle *SYSTRAN Pure Neural Server* (SPNS)⁵⁴ intégrant l'état de l'art des technologies de traduction avec le moteur de traduction automatique neuronale. La plateforme est hébergée à l'Université de Paris et

54 <<https://www.systransoft.com/fr/produits-de-traduction/systran-pure-neural-server>>

peut être utilisée pour traduire plusieurs types de documents : fichiers texte, pages web, fichiers audio, etc. Elle permet aux usagers de post-éditer un texte cible, de gérer les résultats de la post-édition dans un processus de correction-révision, de personnaliser le profil de la traduction grâce aux mémoires de traductions et des dictionnaires utilisateurs, d'exporter les résultats pour publication, etc.

Une vue contextuelle spécifique appelée « Traduction de pages web » permet aux traducteurs de modifier n'importe quel segment de texte cible en ayant accès au contexte de la page source et à sa traduction. Tous les segments de texte provenant de différentes pages web (titres, infobulles, éléments de menu, etc.) sont répertoriés automatiquement pour faire des choix pertinents et créer une traduction cohérente du site web. Enfin, les traductions peuvent être améliorées et adaptées à un domaine et à des paramètres de projet spécifiques.

Outre ce flux de travail de traduction, la plateforme présente plusieurs caractéristiques fonctionnelles que nous considérons comme essentielles pour la traduction professionnelle de sites web :

- Traduction collaborative.
- Maintien de la cohérence (gestion des doublons, homogénéisation, etc.).
- Interface graphique et restitution des formats (PDF, HTML, etc.).
- Prévisualisation des pages web.
- Traduction automatique personnalisée (dictionnaire utilisateur, mémoire de traduction, normalisation, etc.).
- Modèle(s) TA enrichi(s) grâce à la prise en compte de la terminologie et phraséologie spécialisée.
- Aide à la localisation (images, audio, etc.).

4. Les particularités du site web en tant que genre textuel

Jusqu'ici nous avons souligné les complexités pédagogiques et technologiques de notre projet pédagogique. Dans cette section, nous examinons une autre dimension de la question : les caractéristiques linguistiques des textes à traduire. Il suffit d'examiner quelques extraits pour constater la grande variété de registres que l'on peut trouver sur le site de l'Université de Paris, avec chacun une configuration spécifique de caractéristiques lexico-grammaticales et revêtant des fonctions rhétoriques contrastées (Bhatia, 2005 ; Maingueneau, 2016) :

Tableau 1 Exemples de modes discursifs distincts sur le site de l'Université de Paris

Discours promotionnel	<i>Forte de ses 24 300 étudiant.e.s, 40 laboratoires de recherche et 2 060 chercheur.e.s, enseignant.e.s-chercheur.e.s, la faculté des sociétés et humanités est une des trois grandes facultés d'Université de Paris. [...]</i> (https://u-paris.fr/faculte-societes-humanites/#)
Discours exhortatif	<i>Ne laissez pas votre mal-être durer. Pour vous aider à faire face, la Fondation FondationMental propose le site Écoute Étudiant.es totalement gratuit et anonyme [...]</i> (https://u-paris.fr/ecoute-etudiants-la-nouvelle-plateforme-de-soutien-psychologique/)
Discours réglementaire	<i>TOUT ENVOI DE PIÈCES SEUL (SANS INSCRIPTION SUR GALAXIE-ALTAIR AU PRÉALABLE), NON CONFORME OU A UNE AUTRE ADRESSE AMÈNERA LA DRH A DÉCLARER VOTRE CANDIDATURE IRRECEVABLE [...]</i> (https://u-paris.fr/ater-lecteurs-maitres-de-langues-enseignants-associes/)

Cette variété de discours constitue à notre avis un défi important pour les apprentis traducteurs. Le plus souvent, leur expérience du texte « expert » se limite à des extraits didactisés de textes semi-spécialisés (modes d'emploi, articles de vulgarisation scientifique, etc.). De même, très peu de ces étudiant.es ont rédigé des textes de niveau professionnel destinés à la publication en ligne dans leur langue natale, encore moins des contenus en anglais.

En outre, les discours institutionnels du projet ont des propriétés très spécifiques. Pour beaucoup d'analystes, les « sites web institutionnels » constituent un genre à part entière (Askehave et Nielsen, 2004 ; Santini *et al.*, 2010, Venuti et Nasti, 2013). D'une part, ces documents comportent des particularités structurelles liées aux interactions de l'utilisateur avec les menus de navigation, les liens hypertexte, les images et le contenu du cadre textuel principal (Ashekawe et Nielsen, 2004). De même, les sites web des universités reprennent les discours promotionnels des brochures universitaires traditionnelles, tout en les adaptant aux attentes d'un nouveau public international (Mauranen, 2010 ; Callahand et Herring, 2012 ; Tréguer-Felten 2019). Mais en parallèle les éditeurs des pages web institutionnelles s'efforcent de bâtir des discours institutionnels reconnaissables et uniques (Ho, 2013 ; Venuti et Nasti, 2013). Ce dernier constat nous interpelle particulièrement : si les auteurs d'un texte institutionnel souhaitent créer une « voix unique » (à l'instar de l'identité visuelle d'une entreprise) en repoussant, par exemple, les limites de la terminologie ou de la phraséologie habituelle, la traduction peut-elle la préserver, notamment si la révision succède à la post-édition de la TA ?

En ce qui concerne notre projet, il faut reconnaître qu'une partie conséquente des textes que nous traduisons est éphémère (annonces, listes d'adresses, organigrammes institutionnels, etc.) et relève des éléments stabilisés sur le plan linguistique. Mais les propriétaires du site nous demandent aussi de traduire en priorité des pages que la plupart des lecteurs vont consulter le plus fréquemment, et qui se

trouvent en général en haut de l'arborescence du site. Pour nous référer à ces textes, nous parlons de « pages vitrines » ou « pages prioritaires » : il s'agit de ce que Halliday (1985) appelle des *high value texts*, des textes dont les qualités rhétoriques se démarquent des autres textes plus prosaïques, et qui constituent des textes de référence reconnus à la fois par les experts et la communauté linguistique plus large.

Pour illustrer ce type de discours « rédactionnel », nous examinons ci-dessous un échantillon de problèmes rencontrés lors de la traduction des premières pages du site de l'Université de Paris. Il ne s'agit pas d'un échantillon d'exemples tirés au hasard : pour sélectionner nos exemples, nous fondons nos observations sur l'analyse textométrique.

5. L'analyse textométrique des différentes étapes de la traduction

Comme nous l'avons expliqué plus haut, le flux opérationnel du projet de traduction de site web nous permet de compiler des corpus de textes comparables correspondant aux différentes étapes du projet : FR (textes originaux), EN0 (traduction automatique), EN1 (post-édition), puis EN2, EN3, etc. Grâce aux méthodes de l'analyse textométrique de ce type de corpus (Gledhill, Patin et Zimina, 2017), nous sommes en mesure d'identifier les différences lexico-grammaticales qui caractérisent les différents volets comparables. Nous utilisons notamment la méthode des *spécificités* appliquée aux résultats de l'étiquetage morphosyntaxique du corpus afin de cibler automatiquement en contexte les éléments caractéristiques *surreprésentés* (spécificités positives) ou *sous-représentés* (spécificités négatives) dans certains volets du corpus par rapport aux autres parties (Patin, Zimina et Fleury 2016).

Tableau 2 Visualisation des résultats du calcul des spécificités

EN0 (TA)

EN1/EN2 (étudiant.es M2)

38	Remember to fill out the address document for your parents or someone who can be used as a contact point for the next 8 years. §	Please remember to fill in the address of your parents or a person who can be contacted over the next 8 years. §
37	After formal authorization from the doctoral school, you will have to go to the registrar of the Doctoral Training Service to pay the amount of the registration (€380) and finalize the registration. §	After formal authorisation from the Doctoral School, you must pay the registration fee (380 EUR) and finalise your enrolment at the designated office in the Doctoral Training Department. §
36	A copy of your Master's degree (or note statement signed by the Administration if you do not - your Master's degree will be due later); §	A copy of your Masters degree (or a transcript signed by the administration, in which case your masters degree will need to be submitted later); §
35	STEPUP Doctors pursue their careers in the fields of Research and Teaching in French or foreign universities and research organizations, in Geophysics or Environment companies, in risk assessment and insurance, in space-related companies, in market statistical analysis companies or using large database processing or image processing, or in the sectors IT and banking. §	What kind of job can you get with a PhD from STEPUP? Our graduates can find work in various sectors including : research and teaching positions at French universities and abroad, research posts in scientific organisations, or in companies involved in geophysics and the environment, or jobs in areas such as risk assessment and insurance, the space industry, statistical and market analysis, big data, image processing, IT and banking sectors, etc. §
34	For courses related to the UFR STEP, you can check here the latest training schedule and the training room. §	For courses relating to the STEP department, here you can check the latest training schedule and the teaching room. §
33	The ED courses are systematically credited with a number of points from 1 to 7. §	The ED courses are systematically credited with a number of points ranging from 1 to 7. §
32	A total of 10 points may be validated during the thesis §	During the thesis, a total of 10 points maximum may be validated §
31	Research and Training Committees are led by the Vice-Presidents. They will participate in the appraisal of subjects and projects before they are submitted to the authorities for decision: the drafting of the establishment contract, the organisation of the evaluation and monitoring of the university's performance in accordance with its strategic plan or the support and definition of new projects or interdisciplinary initiatives. §	Research and Training Committees are led by the Vice-Presidents. They will take part in the study of subjects and projects before they are submitted to the relevant authorities for decision. This will include drafting the establishment contract, organising modalities of assessment as well as monitoring the university's performance in accordance with its strategic plan. Other tasks will include supporting and creating new projects or interdisciplinary initiatives. §
30	A reduced number of faculties §	Reducing the number of faculties §
29	The University of Paris is an important institution in terms of its staff. This size is an asset for its visibility and capacity for action, with central authorities focused on defining the main strategic axes, and faculties with a high degree of operating autonomy; §	Université de Paris is, by the sheer size of its staff, an important institution. This is an asset for its visibility and will enable great capacity for action, with central authorities focused on defining the main strategic axes, and faculties operating with a high degree of autonomy; §

En plus de l'analyse des éléments caractéristiques qui correspondent à des formes ou catégories morphosyntaxiques isolées, nous prenons en compte les récurrences caractéristiques identifiées grâce au calcul des *segments répétés*. Par exemple, les résultats de la comparaison de la traduction automatique (EN0) aux premiers cycles de la révision (EN1/EN2) montrent la présence de plusieurs constructions caractéristiques : 1) <V> to <V-ing>, 2) <Det> <N>, 3) <N> <V-ing>, 4) <V-ed> <Prep> <NP>, 5) <Aux> <Inf> <Det> <N>, 6) <N>

<Aux> <Inf> <PP>, etc. qui témoignent de l'existence des unités phraséologiques étendues (cf. Tableau 2 ci-dessus).

L'une des différences les plus marquées qui caractérisent les révisions par rapport à la TA correspond à un suremploi des « constructions verbales en *-ing* ». Ces emplois reflètent la plupart du temps ce que l'on pourrait appeler la « dé-nominalisation » :

A1) FR	Aménagement d'études
EN0	Student accommodation
EN1-3	Student Development Plan
EN4	Planning your studies

On remarque en premier lieu que la TA a introduit une erreur terminologique : dans le contexte universitaire anglo-saxon, *accommodation* signifie « logement (étudiant) ». Si les étudiant.es ont trouvé que *Student Development Plan* peut traduire le terme « aménagement d'études », la notion visée en français est beaucoup plus spécifique : il s'agit d'un aménagement appliqué aux modalités de suivi de certains étudiant.es handicapé.es (<https://u-paris.fr/en/education-and-disabilities/>). A cet égard, la réorientation interpersonnelle opérée dans la révision de l'experte « *Planning your studies* » nous paraît cohérente.

L'exemple A2 présente une autre configuration du même phénomène : les étudiant.es ont préféré une forme prédicative en anglais (*succeed*), alors que les éditeurs de la version en ligne ont opté pour une construction nominale (*promote + success*) :

A2) FR	Favoriser la réussite
EN0	Favouring success
EN1-3	Helping all students succeed
EN4	Promote the success of all students

Alors que les exemples A1-A2 correspondent à des difficultés essentiellement terminologiques, nous verrons que les exemples B1-3 représentent plutôt des problèmes au niveau de la phraséologie (vue ici

comme la co-sélection de plusieurs lexèmes au sein d'une même construction grammaticale).

Sur ce plan, l'une des régularités que nous avons observée dans notre corpus concerne l'emploi fréquent du verbe *provide* qui apparaît de manière fréquente à la fois dans les résultats de la TA et dans les reformulations proposées par les étudiant.es. Ce verbe apparaît dans des constructions datives (en tant que synonyme plus soutenu que le verbe *to give*), ainsi que dans des constructions à verbe support avec les formulations alternatives telles que *to assist* > *to provide assistance*.

L'exemple B1 présente un exemple d'imbrication phraséologique, où la traduction doit refléter des choix lexicaux et phraséologiques subtiles au sein de la même construction :

B1 FR Université de Paris s'engage à [...] **proposer une offre de formation** toujours en adéquation avec le marché de l'emploi français et international.

EN0 Université de Paris is committed to [...] **offer a training** always in line with the French and international job market.

EN1-4 Université de Paris is committed to [...] **provide a course offering** that is in line with the French and international job market.

Pour traduire *formation*, deux équivalents potentiels existent en anglais : un processus nominalisé (« entraînement » : *teaching, training*) ou un référent nominal (« cursus » : *course, programme*). Dans ce cas, les étudiant.es ont choisi la deuxième option qui est juste. Le choix de *training* dans la traduction automatique pourrait venir de l'influence de l'anglais *lingua franca* dont il faut tenir compte lors de la constitution des corpus d'entraînement. Il s'agit ensuite d'identifier le verbe le plus adapté (*offer* ou *provide*) pour le complément (*course, programme*). Dans ce contexte, *provide* semble mieux adapté, puisqu'il dénote non seulement la mise à disposition d'une offre, mais aussi un degré de responsabilité de la part du sujet (comparable à la notion de « rendre service »).

Le prochain exemple (B2) constitue un problème phraséologique combiné avec une ambiguïté structurale cachée. Ici, une construction à verbe support en français (« prendre » + « en charge ») a été traduite par des structures apparemment comparables *take + responsibility* et *provide + assistance* :

B2 FR L'université **ne prend pas en charge** l'aide aux actes de la vie quotidienne, celle-ci relève de la Prestation de Compensation du Handicap (PCH).

EN0 The university **does not take responsibility** for daily living: this is covered by an allowance known in French as 'Prestation de Compensation du Handicap' (PCH).

EN1-3 The university **does not provide assistance** for everyday life: this falls under the disability compensation benefit (PCH)

EN4 The university **does not provide assistance** with daily living activities, which is covered by the Disability Compensation Service (PCH).

La complexité de cet exemple est révélée par le fait que le moteur de TA (EN0) a omis de traduire le complément principal du verbe français (« l'aide »). Cette erreur est rectifiée dans les versions révisées (où *assistance* correspond à « aide »). Parallèlement, le choix du verbe *provide* correspond à une réorientation interpersonnelle de la phase comparable à celles opérées par l'experte en communication dans l'exemple A2 ci-dessus. Cette reformulation apporte une modification sémantique : sur le plan grammatical il n'y a pas de « bénéficiaire » pour les verbes *prendre/take*, alors que pour les verbes datifs comme *provide* il existe toujours un complément second implicite. Ce changement peut expliquer pourquoi ce verbe est utilisé si fréquemment dans ce type de contexte.

Enfin, le dernier exemple B3 correspond à la co-sélection de plusieurs éléments lexicaux au sein d'une séquence plus étendue :

- B3) FR [...] la valorisation de l'engagement étudiant et l'accès à des programmes d'aide à l'orientation **confèrent** à ses étudiant.es **un cadre d'études propice pour construire** leur projet professionnel.
 EN0 [...] the development [...] of student engagement and access to orientation support programs **provide** students **with a framework for studies to build** their professional project.
 EN1-3 [...] the development [...] of student engagement and access to guidance programmes **provide** students **with a study environment conducive to developing** their professional plan.
 EN4 the development [...] of student engagement and access to guidance programmes **provide** students **with a study environment conducive to developing** their professional plan.

La séquence *conférer + un cadre de N + propice à V* constitue une routine discursive, c'est-à-dire une séquence lexico-grammaticale qui sert de modèle pour de nombreuses séquences qui partagent la même fonction rhétorique (Gledhill, Patin et Zimina, 2017). Pour démontrer non seulement l'existence mais aussi la productivité potentielle de cette construction, il suffit de chercher des occurrences similaires en ligne (exemples B4 et B5) :

- B4 We focus on building teams with complementary talents and **provide** employees with a work **environment conducive to developing** their skills. [...] (<https://bba.cvmanager.com>)
- B5 We **provide** you with a stimulating work **environment, conducive to learning** and professional development [...] (<https://www.wave-stone.com>)

Les exemples B4-5 montrent un schéma comparable à l'exemple B3 avec une forme légèrement différente (*provide + an environment + conducive to V-ing*). Il n'est pas difficile de voir que ces exemples correspondent tous au même type de construction, et que cette routine phraséologique fait partie d'un type de discours plus général que l'on appelle en anglais *mission statement* (« déclaration de mission », « mission d'entreprise »).

Si on revient à la traduction automatique de ce schéma (B3), on remarque que la co-sélection d'éléments lexicaux n'est pas cohérente : le modèle TA permet d'identifier le verbe provide mais n'arrive pas à détecter qu'une structure plus complexe est requise à la fin de la phrase en anglais.

Conclusion

Loock (2019) a récemment souligné la « plus-value » que les « biotraducteurs » (traducteurs humains) peuvent apporter à un projet de traduction (par ex. contrôler la qualité terminologique, faire respecter un code rédactionnel, gérer les liens hypertextes, etc.). Il s'ensuit que dans un projet de traduction professionnel, la traduction automatique occupe une place importante, mais celle-ci est relativement mineure par rapport à l'ensemble des autres activités que les traducteurs doivent accomplir.

Dans cette contribution, nous sommes arrivés à la même conclusion, mais nous avons plutôt souligné les défis auxquels les biotraducteurs sont confrontés, même lorsqu'ils bénéficient d'un moteur de traduction très avancé. En effet, les complexités liées à la gestion d'un projet aussi ambitieux que la traduction d'un site web nous mènent à relativiser l'impact de la TA sur le projet dans son ensemble. De même, si les plateformes de TA actuellement disponibles fournissent des solutions de traduction adéquates, elles ne sont pas en mesure de prendre en charge la traduction de discours « rédactionnels » ou « promotionnels » sans intervention humaine. Comme nous l'avons vu plus haut, nos apprentis-traducteurs apprennent très vite qu'il faut se méfier des premières solutions proposées par la machine. Le fait d'élaborer des discours stratégiquement sophistiqués dans un anglais de haut niveau constitue en effet un défi pour des anglophones non-natifs relativement inexpérimentés. Il nous incombe donc en tant qu'enseignants de leur apprendre les procédures à mettre en place pour garantir que ces textes seront révisés de manière efficace, tout en conservant l'identité discursive des institutions pour lesquelles ils

effectuent leur traduction. Ainsi, si les « pages vitrines » de nos institutions universitaires nécessitent une attention rédactionnelle particulière, cela laisse beaucoup de travail encore pour les équipes d'éditeurs/traducteurs.

BIBLIOGRAPHIE

- Askehave, Inger et Nielsen (dir.). 2004. *Webmediated Genres – a Challenge to Traditional Genre Theory*, Working paper No. 6, Center for Virksomhedskommunikation, Aarhus School of Business.
- Bhatia, Vijay. 2005. Generic patterns in promotional discourse, dans H. Hallamari et T. Virtanen (dir.) *Persuasion Across Genres. A linguistic approach*, Amsterdam/Philadelphia, John Benjamins Publishing Company, p.213-228.
- Callahand, Ewa et Susan Herring. 2012. Language Choice on University Websites: Longitudinal Trends. *International Journal of Communication* 6: 322– 355.
- Direction Générale de la traduction (DGT). 2021. *English Style Guide* (Code rédactionnel pour les rédacteurs / traducteurs de l'Union européenne). Dernier accès (janvier 2021). Accessible ici : <https://ec.europa.eu/info/sites/info/files/styleguide_english_dgt_en.pdf>.
- European Master's in Translation (EMT). 2017. *European Master's in Translation Competence Framework*. Bruxelles : Direction générale de la traduction, Commission de l'Union européenne.
- Gledhill, Christopher, Stéphane Patin et Maria Zimina. 2017. Lexico-grammaire et textométrie : identification et visualisation de schémas lexico-grammaticaux caractéristiques dans deux corpus juridiques comparables en français. *Corpus* 17: 113-144.
- Gledhill, Christopher et Maria Zimina. 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, November 2019. Editions Tradulex, Geneva, p.60-73.
- Halliday, Michael. A.K. 1985. *Introduction to Functional Grammar*. London, Edward Arnold.
- Ho, Victor. 2013. The Need for Identity Construction in Computer-Mediated Professional Communication: A Community of Practice Perspective. In

- Rocci Luppicini (dir.) *Handbook of Research on Technoself: Identity in a Technological Society*. Hershey, USA: IGI Global, p.180-208.
- Loock, Ruddy 2019. Dégager la plus-value de la biotraduction face à la machine. *Traduire* 241, décembre 2019, p.57-68.
- Mangueneau, Dominique. 2016 : L'ethos discursif et le défi du Web. *Itinéraires. Littérature, textes, cultures*, no 2015-3, p.90-96.
- Mauranen, Anna. 2010. Features of English as a Lingua Franca in Academia. *Helsinki English Studies* 6: 6-28.
- Patin Stéphane, Maria Zimina et Serge Fleury. 2016. Lecture Textométrique Différentielle (LTD) de textes législatifs comparables de l'Union européenne. *Actes des 13es Journées internationales d'Analyse statistique des Données Textuelles*, Nice, 7-10 juin 2016.
- Santini, Marina, Alexander Mehler et Serge Sharoff. 2010. Riding the Rough Waves of Genre on the Web. Concepts and Research Questions. Dans Santini, Marina, Mehler, Alexander et Sharoff, Serge (dir.) *Genres on the Web. Computational Models and Empirical Studies*, Dordrecht, Heidelberg, London, New York: Springer, 3–30.
- Tréguer-Felten, Geneviève. 2018. *Langue commune, Cultures distinctes les illusions du 'globish'*. Paris: Editions Hermann.
- Venuti, Marco et Chiara Nasti. 2013. Institutional Academic Discourse in European Universities' websites. *Proceedings of CLAVIER 13 Conference on Discourse in and through the media*, University of Modena, 6-8 November 2013.

OEP

Observatoire Européen du Plurilinguisme

<http://www.observatoireplurilinguisme.eu/>

email : contact@observatoireplurilinguisme.eu

Dépôt légal : novembre 2021

Tous droits de reproduction, d'adaptation et de traduction, intégrale ou partielle réservés pour tous pays. L'auteur est seul propriétaire des droits et responsable du contenu de ce livre.

ISBN : 978-2-492327-14-8

Traduction automatique et usages sociaux des langues. Quelles conséquences pour la diversité linguistique ?

Depuis quelques années les logiciels de traduction automatique ont fait des progrès spectaculaires à la faveur de la constitution sur Internet d'immenses corpus de textes traduits par des traducteurs humains, tels que la base de données de l'Union européenne ou de l'Office européen des brevets, et de l'émergence informatique de la technologie des réseaux neuronaux ou réseaux de neurones artificiels, une des branches de l'intelligence artificielle.

L'objet du colloque en visioconférence organisé le 25 novembre 2020 par l'OEP et par l'Université de Paris sur le thème « Traduction automatique et usages sociaux des langues - Quelles conséquences pour la diversité linguistique », n'était pas prioritairement d'examiner les progrès techniques réalisés ni même de comparer l'efficacité respective des systèmes aujourd'hui présents sur le marché, mais il était évidemment nécessaire d'évoquer le contexte technique pour bien comprendre les enjeux linguistiques et sociaux des applications de traduction automatique ou de traduction assistée.



L'Observatoire européen du plurilinguisme est né des 1ères Assises européennes tenues à Paris en novembre 2005. Structure de mutualisation et de collaboration entre acteurs du plurilinguisme, il a produit la Charte européenne du plurilinguisme, il anime un site Internet, publie une lettre d'information bimestrielle, et organise tous les trois ans les Assises européennes du plurilinguisme.



9 782492 327148

13,00 €
ISBN 978-2-492327-14-8