

Le français, langue du savoir?

Pour une approche structurée de l'usage de la traduction automatique
dans le milieu scientifique

Direction

Benoît Dubreuil, commissaire à la langue française

Recherche, analyse et rédaction

Marc Tremblay-Faulkner

Rodolphe Parent

Mise en page

ookee.ca

Révision

Secrétariat général et direction des affaires administratives et institutionnelles

Date de parution

31 octobre 2023

Comment citer ce document

Commissaire à la langue française. (2023). Le français, langue du savoir?

Pour une approche structurée de l'usage de la traduction automatique dans le milieu scientifique.

<https://commissairelanguefrancaise.quebec/publications/avis/francais-traduction-milieu-scientifique.pdf>

Éditeur

Commissaire à la langue française

875, Grande Allée Est, bureau 1.879

Québec (Québec) G1R 4Y8

Site Web : commissairelanguefrancaise.quebec

Courriel : info@clf.quebec

La reproduction de cette publication est autorisée pour l'enseignement et la recherche, à la condition que l'extrait soit intégral, sans modifications. Tout le document peut aussi être reproduit à ces fins.

La mention de la source est obligatoire.

Toute autre utilisation doit être autorisée par le Commissaire à la langue française, qui détient les droits exclusifs de propriété intellectuelle sur ce document. Pour l'obtenir, communiquez avec le Commissaire à info@clf.quebec.

La présente publication a été produite par le Commissaire à la langue française.

Avis remis à l'Assemblée nationale du Québec.

Dépôt légal

Bibliothèque et Archives nationales du Québec, 2023

ISBN : 978-2-550-96000-3 (version imprimée)

ISBN : 978-2-550-96039-3 (version PDF)

© Commissaire à la langue française, 2023

Dans un souci d'inclusion, nous avons privilégié l'emploi d'un vocabulaire épïcène.

Le Commissaire à la langue française

Le Commissaire à la langue française a été créé en juin 2022, à la suite de l'adoption de la *Loi sur la langue officielle et commune du Québec, le français* (2022, c. 14). Nommé le 8 février 2023 par l'Assemblée nationale, le premier commissaire à la langue française est entré en fonction le 1^{er} mars 2023 pour un mandat de sept ans.

Le commissaire a pour fonction de :

- surveiller le respect des droits fondamentaux et des obligations prévus à la *Charte de la langue française*;
- vérifier la mise en œuvre de ses dispositions par l'Administration, notamment le ministère de la Langue française, l'Office québécois de la langue française et Francisation Québec, ainsi que par les institutions parlementaires;
- surveiller l'évolution de la situation linguistique au Québec;
- recommander des mesures susceptibles de favoriser l'usage du français comme langue commune;
- réaliser des vérifications et des enquêtes sur toute matière relevant de ses fonctions et en faire rapport à l'Assemblée nationale;
- informer le public de toute question relative à la langue française.

En vertu des pouvoirs qui lui sont conférés, le commissaire fournit à l'Assemblée nationale, au gouvernement et au ministre de la Langue française les avis et les recommandations qu'il estime appropriés, s'il le juge nécessaire.

Sommaire

Rédiger sa thèse de doctorat, assister à une conférence internationale, consulter les principaux textes de sa discipline, le tout en français, cela pourrait-il devenir la norme? Le commissaire à la langue française répond oui, et ce, rapidement, à la condition de s'en donner les moyens.

Ces dernières années, les progrès de l'intelligence artificielle ont transformé le domaine de la traduction en rendant les outils de traduction automatique plus accessibles et performants. Ces progrès sont l'occasion de redonner

Les progrès de la traduction automatique sont l'occasion de redonner au français une place de choix en science et en recherche.

au français une place de choix en science et en recherche. Pour ce faire, on doit reconnaître non seulement le potentiel indiscutable de la traduction automatique, mais aussi les problèmes inhérents à sa mauvaise utilisation. La majorité des usages de la traduction automatique ne soulève aucune difficulté, mais un usage mal avisé peut mener

à une violation du droit d'auteur ou de la confidentialité, ou encore à la diffusion de textes contenant des erreurs potentiellement dommageables.

Pour ces raisons, le commissaire à la langue française plaide en faveur d'une approche structurée de l'usage de la traduction automatique dans le milieu scientifique. Pour redonner au français la place qui lui revient en science et en recherche, des actions concertées doivent être menées pour ajuster les pratiques individuelles et faire du multilinguisme une priorité partagée au sein du milieu scientifique, aussi bien à l'échelle du Québec qu'à celle de la francophonie.

Synthèse des propositions

Pour mettre en place une approche structurée de la traduction automatique dans le milieu scientifique, le commissaire à la langue française présente trois propositions.

Le réaménagement linguistique des universités québécoises

Pour tenir compte de la disponibilité et de la performance des outils de traduction automatique, les universités devront analyser la situation linguistique de leur établissement et élaborer un cadre de bonnes pratiques pour y renforcer la place du français dans l'enseignement et la recherche.

Pour atteindre cet objectif, les universités devront :

- 1** rendre accessible à l'ensemble de la communauté universitaire l'information nécessaire pour utiliser de façon raisonnée la traduction automatique;
- 2** assurer la présence du français dans les activités scientifiques qu'elles accueillent;
- 3** cibler les publications de langue anglaise les plus importantes en vue d'en effectuer une traduction française de bonne qualité;
- 4** offrir un accompagnement en traduction aux étudiants et étudiantes, les incitant ainsi à rendre disponibles en français leurs mémoires ou thèses.

La création d'un pôle d'expertise québécois sur la traduction scientifique et le français en science

Pour appuyer le réaménagement linguistique des universités québécoises, le pôle d'expertise aura pour mandat de :

- 1** faire l'inventaire des expertises disponibles au Québec et les réunir;
- 2** de concert avec le réseau universitaire, mettre au point un programme de formation sur une utilisation raisonnée de la traduction automatique dans le contexte scientifique;
- 3** élaborer une stratégie de libre accès et de traduction scientifique de corpus disciplinaires qui sera efficace;
- 4** aider les membres de la communauté universitaire qui le désirent à traduire des articles ou des ouvrages scientifiques selon les meilleures pratiques.

La mise en place d'un centre de coordination sur la traduction scientifique et le français en science à l'échelle de la francophonie

Le centre de coordination devra établir des partenariats et négocier des accords avec les grands acteurs scientifiques mondiaux pour réaliser les mandats suivants :

- 1** piloter la collecte et le traitement de données textuelles scientifiques pour bâtir des corpus multilingues de qualité, perfectionner les algorithmes de traduction automatique et revitaliser la terminologie française;
- 2** élaborer une stratégie de découvrabilité adaptée aux contenus scientifiques qui placera en son cœur la traduction automatique et ses progrès;
- 3** inscrire la question de la traduction scientifique dans les débats, plus larges, du multilinguisme scientifique et de la science ouverte lors de ses représentations auprès d'organisations internationales;
- 4** entreprendre des démarches pour intégrer de façon systématique la traduction automatique dans le processus d'édition scientifique, notamment en libérant les textes de leurs droits d'auteur pour leur traduction.

Table des matières

Introduction	10
1. Les progrès récents de la traduction automatique	12
1.1 Des années 1940 à nos jours	13
1.2 Quelques principes de base	15
2. La traduction automatique et ses usages dans le milieu scientifique	18
2.1 Un continuum d'usages	19
2.2 L'importance d'une approche raisonnée de la traduction automatique	22
3. Plaidoyer pour le multilinguisme en science et en recherche	24
3.1 La prédominance de l'anglais en science et ses effets	25
3.2 Science ouverte, multilinguisme et découvrabilité des contenus scientifiques : prendre le train en marche	29
4. Vers une approche structurée pour résoudre les problèmes liés à la traduction automatique	32
4.1 Les problèmes liés à une mauvaise utilisation de la traduction automatique	33
4.2 Les problèmes liés à une sous-utilisation de la traduction automatique	34
4.3 Pour une approche structurée de l'usage de la traduction automatique	34
5. Trois propositions pour une approche structurée	36
5.1 Le réaménagement linguistique des universités québécoises	38
5.2 Un pôle d'expertise québécois sur la traduction scientifique et le français en science	41
5.3 Un centre de coordination à l'échelle de la francophonie	44
Conclusion	47
Bibliographie	48
Remerciements	51

Introduction

Depuis quelques décennies, l'anglais s'est imposé comme la langue véhiculaire de la science. Non seulement il prédomine dans la publication scientifique, mais aussi il occupe une place de plus en plus grande dans la tenue des rencontres scientifiques, le fonctionnement de la recherche, et l'enseignement aux cycles supérieurs. Même si cette prédominance est avantageuse à certains égards, elle cause des problèmes et des iniquités. Comment des sociétés qui ne sont pas anglophones peuvent-elles tirer pleinement profit d'un savoir scientifique disponible uniquement en anglais?

La traduction est toujours apparue comme une solution limitée à ce problème en raison des coûts et des délais de traitement qui lui sont associés. Cependant, les progrès de l'intelligence artificielle et de la traduction automatique viennent changer la donne. Si traduire un article scientifique prenait auparavant de nombreuses heures de travail, quelques secondes suffisent aujourd'hui. En un simple clic, on peut obtenir une traduction, de qualité généralement acceptable, de n'importe quel article scientifique dans la langue de son choix parmi des dizaines, voire des centaines.

Ainsi, l'enjeu n'est pas de déterminer si la traduction automatique devrait jouer un rôle fondamental dans le milieu scientifique : c'est déjà le cas. L'enjeu est d'en maximiser les bénéfices, tout en évitant certains écueils liés à un usage non averti. En effet, bien que puissants, les systèmes de traduction automatique sont imparfaits et mal compris de leurs utilisateurs. Or, un mauvais usage peut mener à une violation du droit d'auteur ou de la confidentialité ou encore à la diffusion de textes contenant des erreurs de traduction potentiellement dommageables. Par conséquent, l'usage de la traduction automatique en science gagnerait à être encadré par une approche englobante, complémentaire à l'expertise humaine.

Plus exactement, nous recommandons une approche structurée de l'usage de la traduction automatique dans le milieu scientifique. Fondée sur la mutualisation des ressources existantes, elle favoriserait un éventail d'usages en combinant les technologies de la traduction et l'expertise humaine. Elle transformerait aussi la façon dont on utilise la traduction automatique. D'un usage individuel et spontané, on passerait à un usage multiple, institutionnalisé et

Nous recommandons une approche structurée de l'usage de la traduction automatique dans le milieu scientifique.

spécialisé, selon les besoins. Le français deviendrait ainsi une langue usuelle de la science et de la recherche. Le Québec peut — et doit — se positionner rapidement sur cet enjeu. Déjà, des acteurs du Québec participent à un vaste

mouvement international qui milite pour une science multilingue et l'accès libre aux contenus scientifiques. Or, le succès de ce mouvement dépend en partie d'un usage raisonné de la traduction automatique. Le Québec peut jouer un rôle central dans ce mouvement grâce à son réseau d'institutions scientifiques, et à une expertise de pointe dans différents domaines à l'intersection de la traduction automatique.

Dans cet avis, nous présentons les récents progrès de la traduction automatique avant de broser le portrait de ses usages dans le milieu scientifique. Nous plaidons ensuite en faveur du multilinguisme en science et en recherche, puis nous avançons trois solutions qui pourraient être implantées autant au Québec qu'ailleurs dans la francophonie.

1

Les progrès récents de la traduction automatique

Avant d'expliquer les principes de base de la traduction automatique, nous présentons succinctement l'histoire de cette technologie.

1.1 Des années 1940 à nos jours

Publié en 1949, le court document intitulé *Translation* est souvent présenté comme la première intervention écrite sur la traduction automatique par ordinateur. Le mathématicien américain Warren Weaver y proposait l'idée — révolutionnaire à l'époque — d'utiliser les ordinateurs pour traduire des textes. S'inscrivant dans la théorie de l'information, cette idée lui avait été inspirée par les progrès techniques en cryptographie qui avaient été observés durant la Seconde Guerre mondiale.

Au cours des décennies suivant la publication de ce document, des efforts considérables ont été déployés pour développer des systèmes de traduction automatique. Les résultats furent d'abord mitigés, car on utilisait des approches uniquement basées sur des dictionnaires et des règles grammaticales. Puis, dans les années 1990, grâce à l'avancée de la technologie informatique, on avait remplacé l'approche grammaticale pour des approches utilisant le calcul rapide ou la reconnaissance des motifs. Ces dernières années, un immense bond en avant s'est accompli : la performance des moteurs de recherche s'est considérablement améliorée par le recours à des réseaux neuronaux artificiels qui ont été couplés à l'apprentissage automatique¹. On ne s'appuie plus seulement sur la puissance de calcul des ordinateurs pour traduire : on leur donne la capacité d'apprendre à traduire et de s'améliorer en le faisant.

Les technologies de la traduction, et en particulier la traduction automatique neuronale, sont aujourd'hui utilisées, autant par le grand public que par la communauté scientifique ou professionnelle. À titre d'exemple, l'application Google Traduction a atteint en 2021 le milliard de téléchargements². De nombreux autres outils de traduction sont offerts, qu'ils soient libres d'accès ou commerciaux, génériques ou spécialisés, accessibles en ligne ou propres à certains environnements informatiques. On les utilise dans de nombreux domaines, y compris le domaine scientifique³. Par ailleurs, les avancées en recherche se multiplient — par exemple, on peut aujourd'hui traduire automatiquement l'akkadien (une langue utilisée en Mésopotamie il y a 5 000 ans) ainsi qu'un bon nombre de langues autochtones, y compris l'inuktitut⁴. En parallèle, les robots conversationnels de type ChatGPT, dont l'arrivée sur le marché de masse suscite beaucoup d'engouement, permettent aussi à un large public de générer et de traduire des contenus dans un grand nombre de langues.

¹ Pour un exposé complet et pédagogique sur la traduction et les développements de la traduction automatique, voir Bowker, 2023; pour les derniers développements en traduction automatique multilingue, voir notamment Fan et autres, 2021; Costa-jussà et autres, 2022.

² Pitman, J., 2021.

³ Pour un inventaire récent, voir Fiorini et autres, 2020.

⁴ Guthertz et autres, 2023; Microsoft, 2023; CBC News, 2021.

—

Certains domaines de connaissance touchant à la traduction automatique

La traduction automatique est le fruit d'une multitude de domaines. En voici les principaux.

Traduction

Par leurs connaissances fines, les spécialistes de la traduction peuvent comprendre notamment les nuances entre les langues, la syntaxe, la sémantique, la grammaire et le contexte culturel. Ils collaborent souvent avec des spécialistes de l'informatique pour développer des systèmes de traduction automatique. Par ailleurs, si une traduction professionnelle repose sur la traduction automatique, elle peut inclure à la fois la prédiction (préparation du texte de départ avant sa traduction par un algorithme) et la postédition (révision de textes traduits par un logiciel ou un moteur de traduction automatique). Ces tâches sont réalisées par des spécialistes de la traduction ou d'autres domaines langagiers.

Linguistique

La linguistique est l'ensemble des sciences qui étudient le langage. Les linguistes apportent une compréhension fondamentale de la structure des langues, de la sémantique, de la phonétique et d'autres aspects qui peuvent influencer la traduction.

Traitement automatique du langage naturel

Le traitement automatique du langage naturel (TAL) est une branche de l'intelligence artificielle qui traite de l'interaction entre les ordinateurs et le langage humain. La traduction automatique est au cœur du TAL, qui comprend la compréhension du langage, la génération du langage, le résumé automatique, la traduction automatique, et plus encore.

Informatique

Les compétences des spécialistes en génie et programmation informatiques sont essentielles pour développer et maintenir les systèmes de traduction automatique. Ces systèmes reposent notamment sur des algorithmes, des structures de données, de l'apprentissage automatique et des mécanismes d'optimisation.

Apprentissage automatique

L'apprentissage automatique, et en particulier l'apprentissage profond, est largement utilisé dans la traduction automatique moderne. Des modèles sophistiqués, comme les réseaux de neurones récurrents et les transformateurs, sont aujourd'hui couramment utilisés.

Science des données

Les scientifiques des données aident les spécialistes à analyser et à interpréter les ensembles de données complexes utilisés pour former et tester les systèmes de traduction automatique. Ils contribuent également à l'optimisation des modèles d'apprentissage automatique.

Autres domaines d'expertise concernés

Une expertise particulière peut être requise dans plusieurs domaines du langage. Par exemple, la traduction de documents juridiques ou médicaux nécessite une connaissance de la terminologie propre à ces domaines.

1.2

Quelques principes de base

Une des principales caractéristiques des systèmes de traduction automatique⁵ est la facilité avec laquelle on peut les utiliser. Cette facilité explique leur usage généralisé, mais sous-optimal. En un clic, on active la traduction automatique sur son navigateur pour traduire un site Internet et le consulter dans une langue que l'on connaît, ou encore on copie-colle des textes entiers dans un traducteur en ligne pour rédiger un texte dans une langue que l'on maîtrise moins bien. La plupart du temps, on ne comprend pas bien le fonctionnement des outils de traduction, ni les implications des traductions produites, qu'elles soient, par exemple, scientifiques, juridiques ou sociales.

Ainsi, nous vous présentons dans cette section certains principes de base essentiels pour l'évaluation de l'apport de la traduction automatique à la science et la recherche⁶.

1 Chaque outil de traduction automatique se base sur un corpus distinct

Les outils de traduction automatique dépendent étroitement des données qui alimentent leurs modèles. En effet, chaque outil repose sur un corpus qui lui est propre et il est entraîné à partir de corpus comparables (au moins deux sous-corpus du même type, mais indépendants l'un de l'autre) et parallèles (des textes originaux dans une langue qui sont traduits dans au moins une autre langue)⁷. Cette réalité est

l'un des facteurs techniques expliquant pourquoi on n'obtient pas la même traduction d'un moteur de traduction à l'autre, par exemple Google Traduction et DeepL. Les données du corpus qui alimente chaque outil déterminent donc les performances de l'outil selon le contexte et le contenu du texte à traduire. Cette caractéristique est essentielle, et les suivantes en sont tributaires.

⁵ Pour le reste du document, nous nous référons uniquement aux systèmes de traduction automatique neuronale lorsque nous parlons de traduction automatique. Ces systèmes sont de loin les plus performants et les plus courants sur le marché.

⁶ Les enjeux présentés dans cette section sont pour l'essentiel tirés de Bowker, 2023, chap. 6. Il ne s'agit pas d'une liste exhaustive.

⁷ Loock, 2016.

⁸ On commence toutefois à utiliser la traduction automatique brute pour la constitution de corpus parallèles artificiels. Voir notamment Artetxe et autres, 2019; Fan et autres, 2021.

⁹ En effet, les spécialistes de la traduction ont aujourd'hui accès à un environnement technologique très riche, qui inclut des moteurs de traduction automatique comme Google Traduction ou DeepL, ainsi que des bases terminologiques, des mémoires de traduction, des interfaces de postédition ou des outils de contrôle qualité. À ce sujet, voir Fiorini et autres, 2020; Fiorini, 2022.

2 La traduction automatique nécessite de grandes quantités de données

Les outils de traduction automatique nécessitent des bases d'entraînement textuelles extrêmement volumineuses. Par conséquent, la traduction basée sur les données fonctionne mieux pour les langues et les paires de langues dont l'utilisation est fréquente et répandue. Par exemple, le français et la paire anglais-français sont utilisés largement, tandis que l'islandais et la paire islandais-éwé sont peu utilisés. Il est donc probable que la traduction sera de meilleure qualité dans le premier cas que dans le second. De plus, pour les paires de langues peu utilisées, certains moteurs de traduction automatique neuronale utilisent une stratégie de « pivotage », qui nécessite une double traduction par le biais d'une langue pivot. Cette stratégie en deux étapes augmente cependant la probabilité d'introduire des erreurs de traduction. Notons, par ailleurs, que les variantes géographiques d'une même langue ne sont pas toujours prises en compte. Par exemple, le moteur de traduction automatique Bing propose le français canadien, mais pas Google Traduction.

3 La traduction automatique repose sur l'expertise humaine

La traduction automatique ne menace pas les spécialistes en traduction. Bien au contraire, elle s'appuie sur leur expertise de différentes façons, en amont et en aval de son processus. En amont, les outils de traduction automatique neuronale reposent en grande partie sur les traductions effectuées par des spécialistes de la linguistique et de la traduction. L'existence de corpus parallèles — bilingues ou multilingues, issus d'un travail humain — est en effet fondamentale à l'apprentissage automatique⁸. En aval, la production de textes de qualité exige de combiner la traduction automatique et un travail de postédition, qui s'effectue souvent à l'aide d'un des nombreux outils de traduction assistée par ordinateur offerts sur le marché. Ces outils sont compatibles avec des moteurs de traduction automatique et ils intègrent de nombreuses autres fonctionnalités, comme des mémoires de traduction, des lexiques ou un contrôle qualité⁹.

4 La traduction automatique requiert des corpus dans des domaines spécialisés pour être performante

En plus de reposer sur un très grand volume de données, les approches basées sur les données nécessitent également le bon type de données. En traduction, cela peut signifier des textes issus d'un domaine précis ou d'un genre particulier. En raison de leur rareté, les corpus spécialisés comparables ou parallèles constituent une ressource très prisée sur le marché. Les organisations qui les détiennent peuvent s'attendre à des offres alléchantes pour répondre à une demande croissante, d'autant plus qu'un grand nombre de moteurs spécialisés, détenus par les grandes agences de traduction ou par des entreprises, ne sont pas accessibles au grand public¹⁰.

5 La distinction entre langue générale et langue de spécialité est importante

Chaque domaine ou sous-domaine de connaissances comportent des spécificités terminologiques et linguistiques. Ainsi, pour traduire correctement un texte dans un domaine particulier comme l'archéologie ou le droit, un outil de traduction automatique doit être alimenté par un nombre suffisamment élevé de textes de ce domaine. La distinction est donc cruciale entre langue générale et langue de spécialité, soit un sous-système linguistique correspondant à l'emploi de la langue dans un domaine de connaissances unique¹¹. Si un outil est entraîné à partir d'un corpus parallèle de langue générale, il sera probablement moins performant pour traduire des textes spécialisés. S'il est entraîné à partir de textes issus d'un seul domaine, sa performance risque alors d'être médiocre pour traduire un texte dans un autre domaine.

¹⁰ Fiorini et autres, 2020, 32.

¹¹ La langue scientifique peut également être conceptualisée à l'intersection de la langue générale et de la langue de spécialité. À ce sujet, voir Pecman, 2007.

2

La traduction automatique et ses usages dans le milieu scientifique

Dans cette section, nous présentons le continuum d'usages de la traduction automatique, et les questions éthiques qui y sont associées.

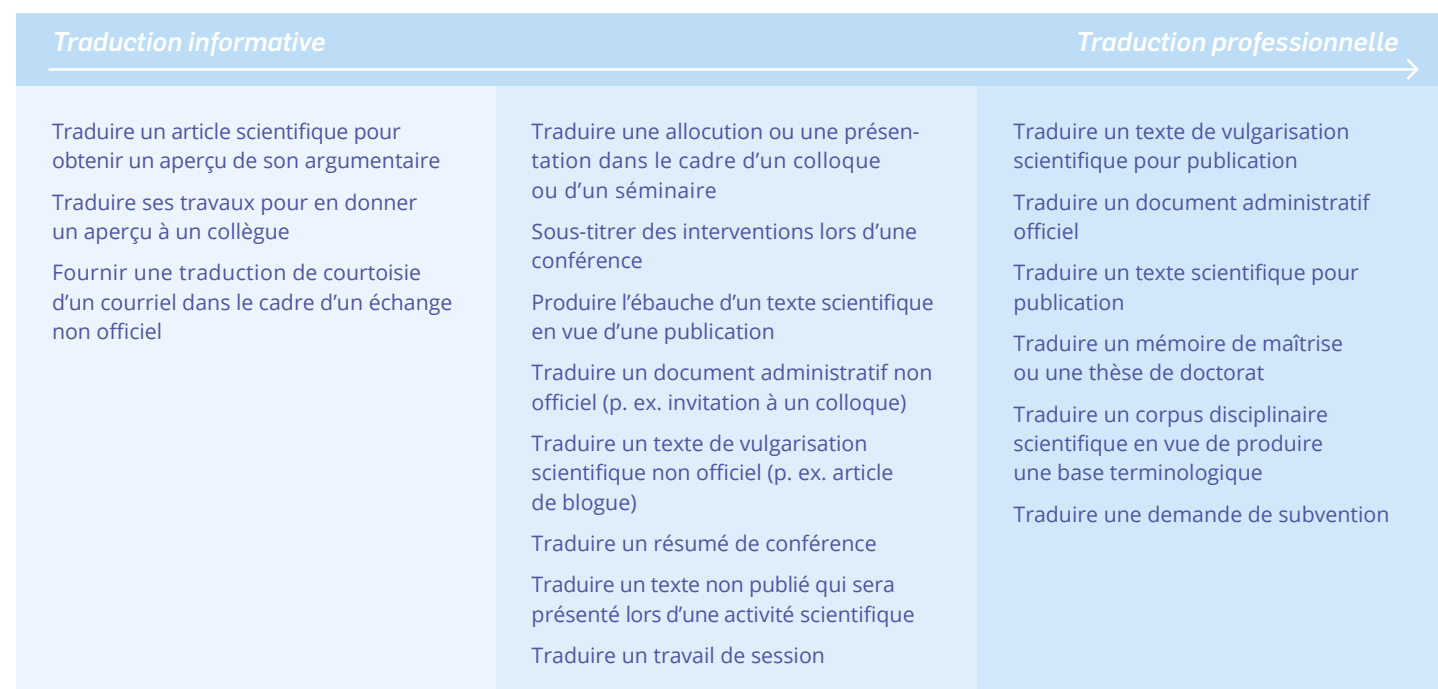
2.1 Un continuum d'usages

Un bon nombre d'études montrent que la communauté étudiante est composée d'avidés utilisateurs de moteurs de traduction automatique¹². Cependant, peu d'études ont analysé l'utilisation de ce type d'outils par les chercheurs et les chercheuses, sauf en didactique des langues étrangères et en traduction. Une enquête menée en 2021 auprès de spécialistes en sciences de l'environnement indique un usage modéré de ce type d'outils¹³. Parmi ceux et celles dont l'anglais n'était pas la langue maternelle, 46,4 % utilisaient régulièrement la traduction automatique pour consulter la recherche scientifique rédigée en anglais. Les auteurs de l'article prévoient une croissance de cette utilisation à la suite de l'introduction des grands modèles de langage comme GPT ou BERT. Étant donné la proportion d'écrits scientifiques publiés en anglais par rapport aux autres langues et le caractère cosmopolite de la recherche, les milieux scientifiques non anglophones sont particulièrement enclins à utiliser ce type de ressources.

Ainsi, à la lumière des principes de base énoncés à la section 1.1, nous pouvons imaginer un continuum sur lequel nous répartirons les usages de la traduction automatique entre traduction informative et traduction professionnelle (voir figure 1).

¹² Voir notamment Mundt et Groves, 2016; Lee, 2020; Holt et autres, 2022; Klimova et autres, 2022; Paterson, 2023.

¹³ Amano et autres, 2023

Figure 1: Continuum d'usages de la traduction automatique dans le milieu scientifique

À gauche de ce continuum apparaît la traduction informative, soit la traduction instantanée de documents pour usage personnel ou pour circulation restreinte. Effectuée strictement à titre informatif (p. ex. traduire un texte pour le comprendre ou améliorer sa compréhension), elle fournit rapidement un aperçu du sujet d'un article ou de l'argument d'un auteur. Les usages qui en relèvent n'exigent pas nécessairement de corrections.

À droite se trouve la traduction professionnelle, par laquelle un ou une spécialiste de la traduction transpose un texte dans une autre langue. Elle est nécessaire quand un document a une visée officielle ou professionnelle (p. ex. traduire un article scientifique en vue de le publier ou un document administratif officiel). Si elle est réalisée au moyen d'outils algorithmiques, elle exige une postédition par une personne qualifiée.

Entre ces deux extrêmes, nous plaçons les usages intermédiaires (p. ex. traduire une présentation pour un cours ou produire un résumé pour un collègue). En ce qui les concerne, une révision non professionnelle de la traduction peut suffire, mais elle doit être réalisée par une personne qui connaît bien le domaine d'expertise abordé et qui a une maîtrise adéquate de la langue source et de la langue cible. Cette catégorie intermédiaire recouvre d'ailleurs de nombreux usages dans le domaine scientifique, où le taux de bilinguisme est élevé et où de nombreux documents non officiels sont produits quotidiennement dans le cadre de l'avancement de la recherche.

Nous pourrions penser que bon nombre des usages répertoriés dans la figure 1 sont déjà répandus dans le milieu scientifique. Comme le grand public, la communauté scientifique utilise régulièrement la traduction informative dans ses tâches quotidiennes, qu'elles soient de nature informelle (p. ex. échanges avec des collègues) ou professionnelle (p. ex. traductions de textes à des fins de consommation ou de production). Cet usage individuel est peu ou pas institutionnalisé.

Évidemment, l'utilisation de la traduction n'est pas nouvelle dans le milieu de la recherche. Jusqu'à tout récemment, par exemple, de nombreux universitaires québécois francophones, quand ils désiraient soumettre un texte à une revue de langue anglaise, écrivaient d'abord leur texte en français, puis ils le faisaient traduire par un ou une spécialiste de la traduction, ou encore ils lui soumettaient une version anglaise de leur cru pour révision. Dans d'autres cas, cette version anglaise était plutôt soumise pour révision à un membre anglophone de l'équipe de recherche.

Aujourd'hui, grâce à la traduction automatique, il devient de plus en plus facile pour les chercheurs et chercheuses francophones de produire une version anglaise de qualité acceptable en s'autotraduisant à l'aide d'un logiciel en ligne. Cette version peut ensuite être revue par un ou une collègue anglophone ou par un ou une spécialiste de la traduction ou de la révision. Pour combler ce besoin, plusieurs revues scientifiques offrent des services de postédition gratuitement aux personnes dont l'anglais n'est pas la langue maternelle¹⁴. De grandes maisons d'édition scientifiques, quant à elles, offrent des services payants de traduction et de postédition professionnelles. À titre indicatif, Elsevier offre sur son site Web des services de traduction vers l'anglais à partir de plusieurs langues. Une traduction du français à l'anglais d'un texte de 5 000 mots coûte ainsi 1 485 \$ US¹⁵.

¹⁴ Pour une liste de ces services, voir les annexes de Steigerwald et autres, 2022.

¹⁵ Elsevier, 2023. Simulation réalisée le 10 juillet 2023.

¹⁶ Dans le cadre de cet avis, nous nous limitons aux enjeux éthiques et à ceux qui concernent les utilisateurs et les utilisatrices de la traduction automatique. Pour un regard exhaustif, qui tient également compte d'enjeux sociaux et environnementaux, voir Moniz et Escartín, 2023.

¹⁷ Consortium Érudit, 2021a.

¹⁸ Steigerwald et autres, 2022, 993.

¹⁹ Bowker, 2023, 103-104.

—

Pour les scientifiques, le fait d'intégrer la traduction automatique dès la phase de rédaction peut être bénéfique en raison des connaissances disciplinaires et terminologiques qu'ils détiennent dans leurs langues maternelles. Pour une personne n'ayant pas une parfaite maîtrise de l'anglais, le fait de rédiger un texte dans sa langue maternelle en s'autotraduisant à l'aide d'un outil en ligne peut représenter un important gain d'efficacité. Il est toutefois essentiel que le chercheur ou la chercheuse comprennent les forces et les faiblesses de l'outil qu'ils utilisent pour déterminer le niveau de révision approprié. Un bon nombre de scientifiques ayant une bonne connaissance de l'anglais s'occupent eux-mêmes de la postédition de leurs textes, mais on doit garder à l'esprit qu'un travail de postédition de niveau professionnel nécessite des compétences particulières et une formation adéquate en traduction.

2.2

L'importance d'une approche raisonnée de la traduction automatique

L'usage de la traduction automatique dans le milieu scientifique soulève certaines questions éthiques¹⁶. La première concerne le respect des droits d'auteur et celui du travail intellectuel des auteurs originaux et des traducteurs humains. En effet, ces droits peuvent être menacés par la réutilisation ou la modification d'un texte non autorisées, ou par la mise en circulation sans encadrement de textes traduits automatiquement. En raison des contraintes liées aux droits d'auteur, les universitaires qui veulent publier leurs travaux dans d'autres langues que l'anglais ne le peuvent pas toujours. La publication croissante d'articles scientifiques dans des revues en libre accès crée néanmoins de nouvelles possibilités. Ces revues optent souvent pour l'une ou l'autre des licences ouvertes pour les œuvres en partage (*Creative Commons*), qui se sont imposées ces dernières années comme le moyen le plus simple de satisfaire aux exigences de libre accès¹⁷. Par rapport à d'autres types de licences, les licences pour œuvres en partage n'obligent pas les chercheurs ni les chercheuses à céder leurs droits d'auteur aux maisons d'édition qui les publient¹⁸. Elles facilitent ainsi la publication de traductions des textes scientifiques originaux, bien qu'elles ne la permettent pas toujours d'emblée.

Une deuxième question porte sur la confidentialité des données utilisées dans les logiciels de traduction en ligne. Plusieurs sociétés, comme Google et Microsoft, précisent dans leurs conditions d'utilisation qu'elles peuvent conserver les données soumises à un moteur de traduction automatique pour ensuite les réutiliser, par exemple pour alimenter des bases de données textuelles d'apprentissage automatique¹⁹. Les utilisateurs et les

utilisatrices doivent par conséquent faire preuve de prudence lorsqu'ils utilisent des moteurs de traduction automatique gratuits sur Internet. Certains documents administratifs, comme des contrats ou des demandes de subvention, peuvent contenir des informations confidentielles, et leur téléversement dans un logiciel en ligne devrait être traité avec prudence.

Une troisième question, celle de la transparence pour le lecteur, doit également être prise en compte par l'utilisateur ou l'utilisatrice de la traduction automatique. Tant qu'un texte traduit par un moteur de traduction automatique n'a pas fait l'objet d'une postédition professionnelle, un risque d'erreur subsiste. La traduction pourrait être globalement bonne, mais la terminologie ou certains passages clés du raisonnement de l'auteur risquent d'être traduits de façon inadéquate, voire trompeuse. Le lecteur devrait donc savoir dans quelle mesure il peut se fier au texte qu'il vient de consulter.

Ces questions éthiques illustrent la nécessité pour la communauté scientifique de faire preuve de discernement et de transparence en adoptant de bonnes pratiques d'utilisation de la traduction automatique. Par exemple, si un ou une spécialiste traduit automatiquement un texte dans le cadre d'un travail de recherche, il ou elle devrait y relever le passage traduit en précisant la langue source, l'outil utilisé, la date de la traduction et les vérifications de la qualité qu'il ou elle a effectuées.

Il est donc nécessaire de promouvoir ce que le spécialiste de la traduction Rudy Loock a appelé une « approche raisonnée » de la traduction automatique²⁰. Cette dernière doit s'appuyer sur la capacité des chercheurs et chercheuses à utiliser de façon efficace la traduction automatique, mais en tenant compte du contexte. Les étudiants et étudiantes, dès le cégep et le premier cycle universitaire, devraient comprendre les principes de fonctionnement de la traduction automatique ainsi que les enjeux pratiques qui découlent de son utilisation.

Il est nécessaire de promouvoir une « approche raisonnée » de la traduction automatique pour l'utiliser efficacement en recherche.

Il existe déjà des modules pédagogiques qui mettent en lumière ces questions et qui expliquent les bonnes pratiques en matière de traduction automatique. Piloté par Lynne Bowker, professeure titulaire à l'École de traduction et d'interprétation de l'Université d'Ottawa, le projet *Machine Translation Literacy* en est un bon exemple²¹. Il regroupe plusieurs outils pédagogiques ainsi que des références utiles pour la communauté scientifique.

²⁰ Loock, 2020.

²¹ Voir Bowker et Ciro, 2019; voir aussi *Machine Translation Literacy Project*, 2023.

—

3

Plaidoyer pour le multilinguisme en science et en recherche

La domination de l'anglais dans la vie scientifique n'est qu'une des nombreuses manifestations de son influence, qui dépend d'une foule de facteurs que nous ne décrivons pas ici²². L'attrait de l'unilinguisme en science n'est pas nouveau. Dans son *Discours préliminaire de l'Encyclopédie* en 1751, d'Alembert parle du latin comme d'un « paradis perdu » en science. Le grec et l'arabe avaient aussi joué autrefois un rôle semblable comme langue véhiculaire savante²³. Au lieu d'être considérée comme un outil d'enrichissement collectif, la diversité linguistique a surtout été conceptualisée comme un problème durant les XVIII^e et XIX^e siècles, une période marquée par l'industrialisation et la mondialisation. Cette conceptualisation a d'ailleurs donné naissance à des langues construites comme le volapük (1880) et l'espéranto (1887)²⁴.

Encore aujourd'hui demeure tenace l'idée selon laquelle la diversité linguistique est un frein à la collaboration et à la communication savante. Par exemple, le multilinguisme serait une « maudite nuisance » selon le politologue hollandais Abram de Swaan et « artificiel, coûteux et juste trop compliqué » d'après le sociologue américain Amitai Etzioni²⁵.

Cette idée relève d'une vision incomplète de l'apport de la diversité et de la cohabitation linguistiques à la science et à la recherche. Les effets indésirables d'un quasi-unilinguisme en science et en recherche sont pourtant bien documentés, et ils suscitent un débat croissant aussi bien au Québec qu'ailleurs²⁶. Le mouvement en faveur du multilinguisme scientifique, longtemps critiqué pour son manque de réalisme, peut aujourd'hui compter sur une alliée de taille : la traduction automatique. Considérant les avancées actuelles et futures des technologies de la traduction, le fatalisme entourant le sort des langues autres que l'anglais en science semble lui-même reposer sur une vision de plus en plus déconnectée de la réalité.

3.1

La prédominance de l'anglais en science et ses effets

Depuis plusieurs décennies, et en particulier depuis le milieu des années 1990, on observe un effondrement des autres langues que l'anglais dans la diffusion des connaissances scientifiques. À l'échelle internationale, l'anglais domine aujourd'hui aussi bien en sciences médicales et naturelles (près de 100 %) qu'en sciences sociales (97 %) ou humaines (90 %)²⁷. Au Québec, ces pourcentages sont plus bas, mais la tendance demeure appréciable. Par exemple, en 2015, 70 % des articles en sciences sociales y étaient publiés en anglais, contre moins de 50 % en 1980. Cette même année en sciences humaines, la proportion d'articles publiés en anglais au Québec se situait plutôt à 30 %, une proportion assez stable dans le temps²⁸.

²² Pour une analyse de la domination de l'anglais en général, voir notamment De Swaan, 2013; pour un état de la situation dans le milieu scientifique québécois, voir Dion, 2012.

²³ Cœuré, 2019, 1; voir aussi Gordin, 2015.

²⁴ Siemund, 2023, 28.

²⁵ Grin, 2017. Nos traductions.

²⁶ Pour un portrait récent de la situation en Europe, voir notamment Wilkinson et Gabriëls, 2021.

²⁷ Larivière et Riddles, 2021.

²⁸ Larivière, 2018.

²⁹ Pour un état de la situation récent au Canada, voir St-Onge et autres, 2021; pour un regard international : Steigerwald et autres, 2022.

³⁰ Voir notamment Rovira et autres, 2021.

³¹ Pour un argumentaire étoffé à ce sujet, voir Montgomery, 2013.

³² Larivière et Gaudreault-DesBiens, 2022.

³³ Longfield, 2022.

—

Non seulement la diffusion des connaissances scientifiques se fait maintenant surtout en anglais, mais on observe également une présence importante de l'anglais dans d'autres sphères de la vie scientifique, comme la rédaction de demandes de subvention, la participation à des conférences internationales ou la recherche documentaire par mot clé dans les bases de données internationales²⁹. Cette évolution est par ailleurs amplifiée par les biais linguistiques des moteurs de recherche³⁰.

La prédominance de l'anglais comme langue de la science présente des avantages indéniables³¹. Elle facilite la communication scientifique et la mobilité internationale en réduisant le nombre de langues que l'on doit apprendre pour mener une carrière en sciences, ou pour diffuser et appliquer les résultats d'une étude scientifique.

Cette prédominance n'est toutefois pas sans conséquence, à la fois pour la communauté scientifique, la population étudiante et la société dans son ensemble. Dans un mémoire³² soumis au Comité permanent de la science et de la recherche de la Chambre des communes du Canada, qui a récemment publié un rapport d'étude sur la recherche et la publication scientifique en français au Québec et au Canada³³, Vincent Larivière et Jean-François Gaudreault-DesBiens offrent une synthèse des conséquences de la prédominance de l'anglais dans la science appliquée au contexte québécois, que nous résumerons en cinq points :

1 L'inégalité dans l'accès aux connaissances pour les francophones

La faible offre de documents savants en français affaiblit la mobilisation du savoir scientifique comme levier de développement social et économique. Ainsi, l'absence de contenus scientifiques francophones rend plus difficile la formation des étudiants et des étudiantes, mais aussi la synergie entre les communautés de recherche et de pratique, dont les actions touchent directement la population.

2 L'inégalité d'accès à la publication savante pour les chercheurs francophones

La rédaction et la communication scientifiques en anglais exigent des efforts des chercheurs et des chercheuses francophones, du Québec ou d'ailleurs, qui doivent doubler leur travail pour atteindre leurs communautés respectives. Menée auprès de chercheurs et de chercheuses en sciences de l'environnement, une enquête récente illustre ce constat. Elle montre que les scientifiques dont l'anglais n'est pas la langue maternelle sont 2,5 fois plus susceptibles de voir une revue scientifique refuser de publier leurs travaux en raison de préoccupations liées à la langue. Ces scientifiques seraient aussi 12,5 fois plus susceptibles de recevoir une demande de révision linguistique³⁴.

3 La baisse de la valeur des objets de recherche associés au français, ou à la diffusion en français

La tendance à l'unilinguisme en science et en recherche a des répercussions sur les objets de recherche des chercheurs et chercheuses francophones, qui sont incités à choisir des objets d'études plus « universels », c'est-à-dire liés à des sujets traités dans le monde anglophone. Cela entraîne du même coup une baisse de la valeur symbolique accordée aux enjeux locaux et appauvrit la recherche qui s'y intéresse. En effet, il existe un risque bien réel³⁵ d'une sous-représentation de la recherche scientifique qui serait produite par les chercheurs francophones dans certaines disciplines, notamment en sciences humaines et sociales. Cette dépréciation peut avoir des effets pervers sur l'appréciation de la recherche francophone par les étudiants et étudiantes, alors qu'ils seront, une fois diplômés, les principaux vecteurs de la diffusion du savoir au sein des entreprises et des autres organisations sociales.

³⁴ Amano et autres, 2023.

³⁵ Voir Daoust et autres, 2022; c'est d'autant plus dommage que la visibilité des écrits scientifiques locaux n'est pas à négliger. Par exemple, l'analyse des téléchargements des revues en sciences sociales et humaines à l'Université de Montréal montre que les articles publiés dans des revues québécoises sont en moyenne plus téléchargés que ceux publiés dans les revues internationales. Voir Larivière, 2014; Cameron-Pesant, 2018.

—

4 L'iniquité subie par les universités francophones dans les classements internationaux

La prédominance de l'anglais influence aussi le positionnement des universités francophones dans les classements internationaux. En effet, ces classements reposent en grande partie sur les publications rédigées en anglais dans des revues internationales. De plus, la production savante francophone du Québec n'y est pas comptabilisée.

5 Le rapport de forces inégal entre les deux langues officielles du Canada

L'inégalité entre les langues française et anglaise dans la science est exacerbée dans le contexte canadien. En effet, au Canada, non seulement l'anglais prédomine comme la langue de la science, mais il est aussi la langue de la majorité. Dans d'autres pays multilingues comme la Belgique ou la Suisse, où l'anglais n'est pas une langue officielle, son usage comme langue scientifique ne bouleverse pas l'équilibre linguistique.

À ces arguments, nous désirons en ajouter un sixième, lié à la valeur instrumentale de la langue³⁶. Nous pouvons, en effet, conceptualiser la langue comme un outil, un instrument qui a de la valeur parce qu'il permet à ses locuteurs d'atteindre les buts et objectifs qui leur tiennent à cœur. Selon cette approche, la valeur instrumentale d'une langue est non seulement tributaire du nombre de personnes qui la parlent et l'écrivent, mais également de la qualité de son corpus³⁷.

Plus une langue est utilisée dans un domaine particulier, par exemple une discipline scientifique, plus on peut s'attendre à ce que son corpus soit façonné par les caractéristiques propres à cette discipline et qu'il en reflète toutes les subtilités et les nuances nécessaires. Ainsi, lorsque les scientifiques cessent d'utiliser une langue pour exprimer des faits liés à un domaine particulier, le statut de cette langue en souffre. Avec le temps, le corpus linguistique propre à cette discipline cessera de s'adapter aux nouvelles connaissances et d'être utile pour exprimer des faits scientifiques complexes et pertinents³⁸.

Contraindre une diversité de points de vue à se conformer à la structure et au vocabulaire d'une seule langue est aussi susceptible d'appauvrir le discours scientifique à la fois en langue anglaise et dans d'autres langues. En effet, la domination de l'anglais a fait émerger une langue à deux vitesses: l'anglais de référence et le « Globish », c'est-à-dire un anglais internationalisé et simplifié, voire appauvri, très présent dans les écrits scientifiques, qui multiplie les néologismes et les synonymes. Toutes les langues ne peuvent pas jouir du même statut ni avoir la même valeur auprès des membres de la communauté scientifique. Par conséquent, on doit s'assurer d'une « symbiose féconde³⁹ » entre, d'un côté, l'anglais comme langue scientifique mondiale et, de l'autre, une diversité de langues scientifiques en santé. Pour cette raison, la traduction nous apparaît comme un élément clé de la solution.

3.2

Science ouverte, multilinguisme et découvrabilité des contenus scientifiques: prendre le train en marche

Depuis quelques années, un mouvement s'organise et prend de l'ampleur en vue de favoriser la science ouverte et d'encourager le multilinguisme dans la communication scientifique. Lancée en 2019, l'initiative d'Helsinki sur le multilinguisme dans la communication savante promeut le multilinguisme pour garantir l'ancrage social des savoirs scientifiques et la pluralité des systèmes de pensée⁴⁰. De nombreux organismes y souscrivent, notamment l'Acfas, le consortium Érudit et le gouvernement du Québec. En 2021, la Conférence générale de l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO) a publié sa *Recommandation pour la science ouverte*. Dans cette recommandation, la définition de la science ouverte inclut la notion de multilinguisme, qui favorise par principe la diffusion de connaissances scientifiques⁴¹.

³⁶ Pour une approche instrumentale appliquée à la langue, voir Robichaud et Schutter, 2012.

³⁷ Robichaud, 2017.

³⁸ Robichaud, 2017, 123-125.

³⁹ Robichaud, 2017, 126-128.

⁴⁰ Federation of Finnish Learned Societies et autres, 2019.

⁴¹ UNESCO, 2021.

⁴² Consortium Érudit, 2021b.

⁴³ St-Onge et autres, 2021, 76.

⁴⁴ Fiorini et autres, 2020, 13.

⁴⁵ Le potentiel de la traduction automatique pour promouvoir le multilinguisme scientifique a d'ailleurs été évoqué plusieurs fois lors du colloque international *Pour des sciences en français et dans d'autres langues*, organisé en 2019 par le ministère français de la Culture, en collaboration avec le réseau OPALE. Pour consulter les actes du colloque : Bonnissent et de Siney, 2021.

Au Québec, les organismes subventionnaires de la recherche des deux ordres de gouvernement ont une politique de libre accès aux publications scientifiques. Depuis 2015, les trois conseils de recherche du Canada imposent un délai d'attente maximal de 12 mois à l'accès gratuit aux résultats de la recherche subventionnée, que ce soit par une revue scientifique ou un dépôt institutionnel. Les Fonds de recherche du Québec (FRQ) ont fait de même en 2019, puis ils se sont joints en 2021 à la cOAlition S. Ce consortium européen promeut la science ouverte, un mouvement qui exige que toute la recherche soit immédiatement publiée en libre accès⁴². Dans son rapport sur la situation du français au Québec et au Canada, l'Acfas recommande d'ailleurs que les organismes subventionnaires de la recherche et les universités appuient ce type d'initiative pour réduire l'importance accordée aux facteurs d'impact des revues lorsque l'on évalue la performance du corps professoral⁴³.

La question de la science ouverte et du libre accès est étroitement liée à celle de la traduction. En effet, la traduction contribue à l'accessibilité du savoir en facilitant la circulation des publications scientifiques auprès d'un plus large public. Ce besoin existe depuis toujours, mais il a récemment été mis en évidence par la crise sanitaire de la COVID-19 : dès les premiers mois de ce qui allait devenir une pandémie, de nombreuses personnes partout sur la planète ont souhaité consulter la documentation scientifique, qui était publiée essentiellement en anglais, du moins au début. Bon nombre d'entre elles ont alors réalisé le potentiel de la traduction automatique pour répondre à ce besoin⁴⁴. En effet, les avancées de la traduction automatique peuvent contribuer à la diffusion de la recherche et au travail des journalistes scientifiques comme intermédiaires entre la science et le public. Elle peut aussi faciliter l'accès à la production scientifique mondiale, en particulier dans un contexte où le multilinguisme scientifique prend de l'ampleur⁴⁵.

La question du multilinguisme scientifique et de la science ouverte est aussi indissociable de celle de la découvrabilité des sciences et des savoirs en français. La découvrabilité, c'est le potentiel pour un contenu, disponible en ligne, d'être aisément repérable par des internautes, notamment par ceux qui ne le cherchaient pas. Plusieurs acteurs de l'espace francophone mondial sont aujourd'hui mobilisés pour cette question⁴⁶. L'Agence universitaire de la Francophonie, par exemple, organise depuis 2021 la Semaine mondiale de la Francophonie scientifique dans le but, entre autres, de promouvoir l'utilisation du français en science dans un contexte multilingue.

Les plateformes numériques d'envergure contrôlent actuellement en grande partie ce qui peut et ce qui ne peut pas être repéré sur le Web, y compris les contenus scientifiques. L'emploi de métadonnées, d'algorithmes de recherche, de mots clés, d'index et de catalogues augmente la découvrabilité d'un contenu au détriment d'un autre. Pour garder vivantes la science et la recherche en français, il est essentiel d'améliorer la visibilité et le repérage des contenus scientifiques de langue française. Un moyen efficace d'y parvenir est le développement de dispositifs de traduction scientifique performants.

Plusieurs acteurs de l'espace francophone mondial sont aujourd'hui mobilisés pour cette question.

⁴⁶ Sur cet enjeu, voir notamment Tchéhoali et Agbobi, 2020.

4

Vers une approche structurée pour résoudre les problèmes liés à la traduction automatique

À la lumière de notre analyse de la situation, nous constatons que le laisser-faire au sujet de la traduction automatique risque d'entraîner deux types de problèmes : certains liés à sa mauvaise utilisation; d'autres, à sa sous-utilisation.

4.1

Les problèmes liés à une mauvaise utilisation de la traduction automatique

Différents problèmes peuvent survenir à cause d'une mauvaise utilisation de la traduction automatique. Le plus évident est le risque de produire des traductions imparfaites, voire erronées. En effet, les textes traduits automatiquement peuvent contenir des erreurs ou des distorsions de sens qui peuvent nuire à la compréhension des travaux scientifiques. Ce problème est important, car une mauvaise traduction peut affecter la crédibilité et la réputation des scientifiques, voire occasionner des préjudices de nature financière, matérielle ou morale. Ce risque varie évidemment selon l'usage qui est fait de la traduction. De plus, des effets négatifs peuvent toucher les utilisateurs et utilisatrices de deuxième ordre, comme les professionnels et professionnelles qui utilisent les résultats de la recherche scientifique dans leur travail. Par exemple, la personne qui utilise sans discernement une traduction de mauvaise qualité en médecine ou génie peut se mettre à risque ou mettre à risque la population.

Un autre problème est la constitution de bases d'entraînement de mauvaise qualité. En effet, si un algorithme est entraîné à partir de textes traduits automatiquement qui n'ont pas fait l'objet d'une postédition, il risque d'être moins performant que s'il avait été entraîné sur un corpus de qualité. Or, l'explosion des traductions instantanées peut mener à l'apparition et à la diffusion de corpus de mauvaise qualité qui réduiront la performance des algorithmes et compliqueront la tâche des spécialistes de la traduction et de l'informatique.

Les lacunes terminologiques sont un autre problème lié à une mauvaise utilisation de la traduction automatique. En effet, les moteurs de traduction automatique performant moins bien s'ils ne sont pas entraînés à partir de corpus spécialisés ni supervisés par des spécialistes qui ont des connaissances fines des disciplines scientifiques pertinentes. Ces lacunes résultent d'un manque de contexte, ou encore de l'évolution rapide du vocabulaire dans certains domaines de recherche.

⁴⁷ Kübler, 2019.

Un dernier problème : la traduction automatique, lorsque mal utilisée, peut occasionner des manquements éthiques. Sans une compréhension claire des enjeux qu'elle soulève, la communauté universitaire peut l'utiliser de façon inappropriée ou non transparente, que ce soit volontaire ou non. Ainsi, la communauté universitaire pourrait utiliser des textes traduits automatiquement sans les indiquer comme tels, ou encore sans respecter leurs droits d'auteur.

4.2

Les problèmes liés à une sous-utilisation de la traduction automatique

Les risques associés à la traduction automatique pourraient décourager son usage dans certaines situations où il est pourtant indiqué. La communauté universitaire pourrait ainsi négliger d'utiliser les outils de traduction disponibles pour produire des traductions de courtoisie, alors que celles-ci peuvent faciliter la communication scientifique dans une foule de situations de la vie courante. Cette sous-utilisation serait ainsi une occasion manquée de rétablir la place du français en science. Pour la communauté universitaire, elle reviendrait à se priver d'un éventail d'usages qui lui permettraient de vivre sa vie scientifique dans la langue qu'elle maîtrise le mieux. De plus, elle la priverait d'outils qui ont le potentiel de « prolonger la validité des collections scientifiques historiques⁴⁷ » tout en conservant la terminologie des registres spécialisés propres aux langues autres que l'anglais.

Sous-utiliser ces nouvelles technologies reviendrait également à ne pas reconnaître les possibilités qu'offre la traduction automatique pour améliorer la qualité du corpus scientifique multilingue et maximiser les retombées de la recherche au bénéfice des sociétés qui les financent. En effet, si la traduction automatique est confinée à un usage individuel, les incitatifs à mener une vie scientifique en anglais pourraient demeurer les mêmes, voire être amplifiés.

4.3

Pour une approche structurée de l'usage de la traduction automatique

Pour surmonter les problèmes exposés dans les sections 4.1 et 4.2, nous croyons nécessaire d'implanter une approche structurée de l'usage de la traduction automatique dans le milieu scientifique. Cette approche poursuivrait un double objectif : miser sur les avancées technologiques récentes en traduction automatique et éviter les écueils liés à sa mauvaise utilisation. Pour réussir, une telle approche doit cependant reconnaître qu'il existe un large éventail d'usages de la traduction automatique.

Il existe un écart considérable entre des situations à risque faible (la traduction courante et instantanée de documents pour usage personnel ou circulation restreinte) et des situations à risque élevé (la traduction professionnelle de documents officiels ou importants). Dans le premier cas, l'usage de la traduction automatique n'entraîne aucun risque sérieux et n'engage aucune ressource. Dans le deuxième cas, l'exercice de traduction est plus risqué et exige des ressources qui doivent être allouées de manière responsable et stratégique. Par exemple, tout effort qui viserait à produire systématiquement une traduction professionnelle de l'ensemble de la littérature scientifique serait à la fois inefficace et coûteux⁴⁸. De la même manière, une utilisation non discriminée de la traduction automatique serait contreproductive.

Par ailleurs, une approche structurée de l'usage de la traduction automatique doit miser sur l'expertise humaine. En effet, les technologies de la traduction peuvent être utilisées à leur plein potentiel si et seulement si elles reposent sur une collaboration entre les spécialistes de la traduction et d'autres disciplines connexes, des technologies, des sciences de l'information et des domaines scientifiques particuliers. Ces spécialistes exercent au sein non seulement des universités, mais également de certaines organisations privées (p. ex. des plateformes de diffusion de contenu scientifique), d'organismes gouvernementaux, d'associations savantes et de regroupements de recherche.

Pour la communauté scientifique, il serait important de mutualiser les ressources pour générer des économies d'échelle et tirer profit des expertises existantes. De plus, pour obtenir des résultats à un coût raisonnable, nous estimons qu'il est crucial de concevoir un modèle d'affaires qui repose sur une mise en commun des ressources à différents niveaux. Cela implique une coordination à l'intérieur des universités, ainsi que l'élaboration d'initiatives interuniversitaires, interdisciplinaires et internationales. Ces initiatives doivent réunir tous les acteurs concernés, notamment les fonds subventionnaires, les associations savantes, les réseaux de recherche, les plateformes de diffusion de contenu scientifique et les organisations gouvernementales.

Il serait important, pour la communauté scientifique, de mutualiser ses ressources pour tirer profit de leurs expertises.

⁴⁸ Pour un exemple de traduction systématique aux résultats décevants en sciences sociales, voir Gingras et Mosbah-Natanson, 2010. Par ailleurs, la plupart des articles scientifiques, même dans leur version anglaise, ne sont lus que par un nombre limité de lecteurs. Seule une très faible proportion de ces articles se démarque et atteint un large public à l'extérieur du lectorat scientifique immédiat. Voir notamment Mohammadi et autres, 2015.

5

Trois propositions pour une approche structurée

Pour rétablir la place du français en science, nous recommandons la mise en place d'une approche structurée de l'usage de la traduction automatique dans le milieu scientifique. Cette approche se décline en trois propositions principales :

- 1 assurer le réaménagement linguistique des universités québécoises;**
- 2 implanter un pôle d'expertise québécois sur la traduction scientifique et le français en science;**
- 3 créer un centre de coordination sur ces domaines à l'échelle de la francophonie.**

À titre indicatif seulement, nous présenterons les mandats et les services que nous envisageons pour chaque instance. Ils seront déterminés en consultation avec les acteurs concernés en fonction de leurs mandats respectifs ainsi que des leviers de financement et des modèles d'affaires qui seront privilégiés.

La mise en place et le déploiement de ces propositions nécessiteront une contribution financière du gouvernement du Québec. Comme le gouvernement du Canada occupe une place importante dans le financement de la recherche, son appui financier serait également souhaitable⁴⁹.

⁴⁹ Le Comité permanent de la science et de la recherche de la Chambre des communes recommandait récemment des actions concrètes pour augmenter la traduction en français des plus importants travaux publiés en anglais et la disponibilité en français des publications scientifiques à des fins d'enseignement, de même qu'un investissement dans des services de soutien à la traduction pour les chercheurs. Voir Longfield, 2022. —

⁵⁰ Dion, 2012; LIRES, 2023; Roy, 2023.

⁵¹ Par « aménagement linguistique », nous entendons l'intervention d'une autorité, ici les directions des établissements universitaires, sur la gestion du français par l'entremise de leur politique linguistique. Depuis 2002, les universités financées par l'État québécois sont tenues d'avoir une politique en matière d'usage et de qualité de la langue française. Ces politiques sont arrimées aux exigences de la *Charte de la langue française*.

⁵² Dans cet avis, nous nous concentrons sur les universités, bien que la situation de l'aménagement linguistique au collégial partage certaines caractéristiques relevées dans cette section. Voir à ce sujet notamment le rapport de LIRES, 2023. Néanmoins, nous considérons que les enjeux ne sont pas les mêmes et qu'ils devraient être traités dans un cadre différent.

5.1

Le réaménagement linguistique des universités québécoises

Dans l'application de leur politique linguistique, les universités québécoises ont depuis longtemps accepté des compromis considérables à l'endroit de l'anglais. En effet, même si l'enseignement en français demeure la norme dans les établissements francophones, des exemples d'utilisation systématique de l'anglais ont été documentés dans divers domaines, dépassant largement la publication scientifique. Par exemple, la rédaction de mémoires de maîtrise et de thèses de doctorat se fait principalement en anglais dans plusieurs disciplines. Il en va de même des communications scientifiques lors des congrès, des conférences et des séminaires. Aux cycles supérieurs, une partie importante du matériel pédagogique est également produite en anglais⁵⁰. Finalement, le recours systématique à l'anglais est favorisé par les collaborations interuniversitaires, car cette langue est souvent la seule connue de tous les participants, qui peuvent toutefois ne pas pleinement la maîtriser.

Le développement d'outils de traduction performants offre aux universités québécoises de nouvelles possibilités en matière d'aménagement linguistique⁵¹. Elle ouvre la voie à des gains rapides dans des situations où le recours systématique à l'anglais est jusqu'à présent vu comme un compromis acceptable.

De ce fait, nous proposons aux universités québécoises d'effectuer un réaménagement linguistique dans lequel elles tiendront compte de la disponibilité et de la performance des outils de traduction automatique.

Ce réaménagement linguistique⁵² s'articulera autour de plusieurs mesures complémentaires.

Les universités devront d'abord **rendre accessible à la communauté universitaire l'information nécessaire pour utiliser de façon raisonnée la traduction automatique** en vue de l'inciter à employer le français. Les universités pourront le faire en concevant elles-mêmes des formations ou en utilisant celles produites par d'autres acteurs. Une telle approche les aidera à maximiser le potentiel de la traduction automatique dans les contextes où le risque est faible, par exemple pour traduire une présentation visuelle, une allocution présentée lors d'un séminaire ou l'ébauche d'une publication scientifique.

Les universités devront aussi déployer des mesures pour **assurer la présence du français dans les activités scientifiques qu'elles accueillent**. Elles devront former les organisateurs d'événements scientifiques internationaux aux bonnes pratiques en matière d'organisation d'événements multilingues, notamment en les incitant à nommer des responsables qui maîtrisent le français pour l'animation, à traduire automatiquement les présentations des conférenciers et à les réviser, à présenter une traduction de courtoisie française ou anglaise des principales allocutions ou à recourir à l'interprétation simultanée. Divers outils existent déjà, par exemple les fonctions de sous-titrages qui peuvent être utilisées lors d'événements hybrides ou virtuels, ou encore qui peuvent projeter la version traduite d'une allocution sur un écran ou la transmettre directement sur les téléphones, les ordinateurs ou les tablettes des participants à un événement.

En partenariat avec leur corps professoral, les universités devront **cibler les publications de langue anglaise les plus importantes qui sont produites par leur communauté ou utilisées en son sein en vue d'en effectuer une traduction française de bonne qualité**. On peut penser aux lectures obligatoires présentées de façon récurrente aux étudiants aux 1^{er} et 2^e cycles, aux listes de références des examens de synthèse au 3^e cycle, ou encore aux publications de leurs chercheurs et chercheuses recueillant le plus de citations. Les universités pourront coordonner les démarches auprès des éditeurs scientifiques en vue d'obtenir l'autorisation de traduction (lorsqu'elle est nécessaire). Pour diffuser les traductions de leurs publications d'importance, les universitaires pourront demander directement aux revues d'annexer une version française aux documents complémentaires de la publication, ou traduire des résumés en langue courante, qui seront par la suite publiés sur un site Web personnel ou institutionnel⁵³.

De plus, **les universités devront offrir un accompagnement en traduction aux étudiants et aux étudiantes, les incitant ainsi à rendre disponibles en français leurs mémoires ou thèses**. Ces traductions permettront aux évaluateurs externes d'effectuer leur travail et ne compromettent pas le rayonnement ni la publication des travaux.

⁵³ Pour une représentation schématique des différentes possibilités de traduction des produits de la recherche, voir la figure 1 dans Steigerwald et autres, 2022.

⁵⁴ Le gouvernement de l'Ontario a procédé à une étude semblable pour cerner les lacunes des ressources éducatives libres en langue française. Voir eCampus Ontario, 2023.

Pour mettre en œuvre ces mesures, les universités devront **analyser la situation linguistique qui prévaut au sein de leur établissement**, notamment en relevant les domaines d'enseignement ou de recherche où l'anglais est utilisé le plus fréquemment, selon les programmes, les départements et les facultés⁵⁴. En plus des publications scientifiques, elles devront **recenser le matériel pédagogique (p. ex., listes de lectures obligatoires), ainsi que les activités scientifiques (p. ex. colloques, séminaires) ou celles liées à la rédaction de mémoires et de thèses**. Ce processus devra mettre à contribution l'ensemble des acteurs concernés, incluant la communauté étudiante. En effet, les étudiants et les étudiantes sont les principaux utilisateurs des services offerts par les universités, et leur point de vue sur l'aménagement linguistique doit être pris en compte. Pour cette analyse, les universités pourront s'appuyer sur des mécanismes de gestion existants, comme l'évaluation du personnel enseignant et l'évaluation de programme.

À partir de cette analyse, elles devront **élaborer un cadre des bonnes pratiques pour renforcer la place du français dans l'enseignement et la recherche dans leur établissement**. Les domaines d'enseignement ou de recherche les plus anglicisés devront y être priorités. Nous sommes conscients que toutes les universités québécoises ne disposent pas d'un service consacré à la valorisation de la langue française. En revanche, toutes ont un service administratif responsable de l'application de leur politique linguistique. Ce service pourrait produire une analyse de la situation linguistique, élaborer le cadre des bonnes pratiques qui en découle et allouer des ressources pour sa mise en œuvre.

5.2

Un pôle d'expertise québécois sur la traduction scientifique et le français en science

Les universités québécoises ne disposent pas toutes des mêmes ressources ni expertises pour mettre en œuvre les mesures que nous leur proposons à la section 5.1. Elles gagneraient donc à collaborer entre elles pour réaliser des économies d'échelle. Cette collaboration pourrait viser, par exemple, l'élaboration d'une formation sur l'usage raisonné de la traduction automatique et la traduction ciblée de publications scientifiques de la communauté universitaire du Québec.

Une telle collaboration serait relativement facile à mettre en place. En effet, le Québec peut déjà compter sur une expertise de pointe dans plusieurs domaines touchant à la traduction automatique, que ce soit dans les départements de linguistique et de traduction, ou les regroupements de recherche qui gravitent autour de l'intelligence artificielle. Au fil des années, le Québec a également établi un important réseau d'institutions scientifiques, qui ont pour mandat la diffusion et le transfert des connaissances scientifiques produites au Québec. À l'heure actuelle, ces organisations n'ont cependant pas le mandat de transformer la traduction en un outil de diffusion et de transfert des connaissances scientifiques en français.

Par conséquent, nous proposons la création d'un pôle d'expertise québécois sur la traduction scientifique et le français en science.

Ce pôle d'expertise devra d'abord **faire l'inventaire des expertises disponibles au Québec et les réunir**. En plus des spécialistes dans les différents domaines touchant à la traduction automatique, d'autres acteurs devront être mis à contribution. Les FRQ et l'Acfas, par exemple, pourront contribuer à l'établissement de partenariats scientifiques, concevoir des formules de financement et accélérer le transfert des connaissances vers les milieux preneurs. Le consortium Érudit est un autre exemple d'acteur à mettre à contribution dans ce pôle d'expertise. Spécialisé dans la valorisation numérique des publications scientifiques et culturelles, il offre déjà des services pertinents, comme le balisage de données ou l'indexation de documents auprès des agrégateurs de contenu internationaux. D'autres ressources devront aussi s'y joindre, notamment celles des bibliothèques universitaires, qui détiennent une expertise importante sur les questions des droits d'auteur et de protection de la propriété intellectuelle.

De plus, le pôle d'expertise doit faciliter la mise en relation entre la communauté scientifique et la communauté professionnelle. Pour cette raison, l'expertise de l'Ordre des traducteurs, terminologues et interprètes agréés du Québec est incontournable. L'expérience de l'Office québécois de la langue française en matière de développement terminologique doit également être mise à profit.

Les mandats confiés à ce pôle d'expertise doivent être complémentaires aux mesures d'aménagement linguistique des universités québécoises que nous proposons à la section 5.1. Parmi ces mesures, nous avons recommandé que les universités québécoises offrent des programmes de formation sur une utilisation raisonnée de la traduction automatique dans le contexte scientifique. Cependant, toutes les universités ne sont pas en mesure de concevoir leur propre programme de formation. Dans tous les cas, elles devraient collaborer avec le pôle d'expertise, **qui mettra au point un tel programme ainsi que des outils complémentaires pour l'ensemble du réseau universitaire.**

Nous avons également proposé que les universités québécoises traduisent les publications de langue anglaise les plus importantes qui sont produites ou utilisées de façon récurrente dans leur établissement (matériel pédagogique, publications d'importance, etc.). Pour ce faire, elles devront collecter des textes libres d'accès, de même qu'obtenir les autorisations ou injecter les sommes nécessaires pour libérer de leurs droits d'auteur les textes qui ne peuvent pas être traduits librement.

Une autre approche possible pour la traduction de corpus disciplinaires serait d'utiliser les réseaux de recherche existants. Au Québec, on retrouve des centaines de regroupements de recherche financés par les universités et les organismes subventionnaires. En partenariat avec ces organismes subventionnaires, le pôle d'expertise pourrait lancer des appels de projets en continu, qui réuniraient à la fois des spécialistes de la traduction, de la terminologie et du traitement automatique du langage naturel et des spécialistes de disciplines thématiques. Ces projets devraient viser la traduction de corpus scientifiques dans les disciplines fortement anglicisées. Les regroupements de recherche devraient appuyer la constitution de corpus de qualité et aider la communauté langagière à réaliser les traductions. Cette stratégie aurait pour avantage de s'intégrer naturellement au mode de fonctionnement de la recherche, en plus d'offrir des objets de recherche intéressants à une génération d'étudiants et d'étudiantes à la maîtrise ou au doctorat en traduction ou dans d'autres disciplines connexes.

Peu importe l'approche retenue, le pôle d'expertise devra élaborer une stratégie de libre accès et de traduction scientifique de corpus disciplinaires qui sera efficace. Une telle stratégie devrait cependant s'assurer d'orienter correctement les efforts à déployer, en tenant compte de l'accélération technologique en matière de traduction automatique. En effet, comme la performance des moteurs de traduction automatique augmente rapidement, il ne serait pas judicieux de traduire dès maintenant des ensembles volumineux de textes. Dans plusieurs cas, il serait préférable de traduire les textes uniquement au besoin, de manière à tirer profit des avancées en matière technologique.

Le pôle d'expertise devra aussi **accompagner les membres de la communauté universitaire qui désirent traduire des articles ou des ouvrages scientifiques selon les meilleures pratiques.** Il servira de point de service pour les utilisateurs et utilisatrices qui se posent des questions pointues sur les enjeux éthiques entourant l'usage de la traduction, la propriété intellectuelle des textes traduits ou les aspects techniques de la traduction automatique. À terme, selon un modèle d'affaires qu'il reste à définir, ce pôle pourra également offrir un service de traduction scientifique à la communauté scientifique québécoise. Pour réduire les chevauchements, on devrait arrimer l'offre de service interne et celle du marché de la traduction professionnelle. Le pôle d'expertise pourrait en effet servir d'intermédiaire entre les membres de la communauté scientifique et celle de la traduction professionnelle spécialisée.

⁵⁵ Fiorini, 2023.

5.3

Un centre de coordination à l'échelle de la francophonie

Les mesures que nous avons proposées aux sections 4.1 et 4.2 établissent la traduction comme un moyen efficace pour rétablir la place du français dans le milieu de la recherche universitaire québécoise. Nous croyons qu'une fois mises en œuvre, elles permettront à la communauté universitaire de mener une plus grande partie de sa vie scientifique dans la langue qu'elle maîtrise le mieux.

Toutefois, nous jugeons que ces mesures ne suffiront pas à contrer la prédominance de l'anglais comme langue véhiculaire de la publication scientifique. En effet, la plupart du temps, les scientifiques doivent rédiger ou traduire leurs travaux en anglais s'ils souhaitent les rendre accessibles à un lectorat mondial ou faire avancer leur carrière. Ces incitatifs sont puissants : ils dépendent d'une structure de production et de diffusion scientifique internationalisée et linguistiquement homogène qui est difficile à réaménager. Pour cette raison, une coordination internationale est nécessaire pour favoriser le multilinguisme en science.

Pour répondre à ces enjeux, nous proposons la mise en place d'un centre de coordination consacré à la traduction scientifique et au français en science à l'échelle de la francophonie.

Les mandats de ce centre de coordination restent à définir, mais **ils seront complémentaires à l'offre de services d'un éventuel pôle d'expertise québécois ou de toute autre organisation similaire dans d'autres États francophones**. Néanmoins, certaines actions doivent être posées pour inscrire la question de la traduction scientifique dans les débats plus larges du multilinguisme scientifique et de la science ouverte.

D'abord, **des corpus scientifiques bilingues ou multilingues de qualité devront être bâtis pour perfectionner les algorithmes de traduction automatique et revitaliser la terminologie française des nombreuses disciplines qui se sont anglicisées au fil du temps**. Or, la collecte et le traitement de tels corpus scientifiques représentent à la fois un défi scientifique et un défi de coordination. Des obstacles doivent être surmontés pour produire les ressources linguistiques nécessaires (bases terminologiques, mémoires de traduction, etc.). En particulier, le cadre juridique contraignant et la qualité inégale des corpus nuisent à l'automatisation des processus de collecte et de traitement des données textuelles⁵⁵.

Nous estimons qu'un effort concerté aiderait à surmonter ces obstacles. En effet, le centre de coordination devra **piloter la collecte et le traitement de données textuelles scientifiques à l'échelle de la francophonie, voire en collaboration avec d'autres régions linguistiques**. Pour y arriver, il pourrait préconiser un mode opératoire similaire à celui du pôle d'expertise québécois, décrit dans la section 5.2, en misant sur la création de corpus scientifiques disciplinaires par les universités ou les regroupements de recherche dans plusieurs pays francophones. Pour faciliter le référencement de ces nouveaux corpus, le centre de coordination devra bâtir des partenariats avec les principales plateformes de diffusion (p. ex. consortium Érudit, Cairn.info).

Le centre de coordination pourrait aussi contribuer à structurer des initiatives déjà en cours. À titre d'exemple, le projet Traductions et science ouverte est une démarche explorant le potentiel de la traduction pour le multilinguisme scientifique. Composé de spécialistes, notamment de la traduction, le groupe de travail scientifique qui pilote ce projet a déjà publié un rapport d'étape dans lequel il propose des recommandations et des pistes d'expérimentation pour évaluer la faisabilité d'un service de traduction scientifique en libre accès⁵⁶. Dans un deuxième rapport, en 2023, il présentera les résultats d'une série d'études préparatoires, incluant une feuille de route et des projections budgétaires⁵⁷.

Des programmes de coopération scientifique internationaux pourraient également inciter les chercheurs et les chercheuses du monde francophone à bâtir de tels corpus. Le volet Recherche du programme Samuel-De-Champlain, financé par la Commission permanente de coopération franco-québécoise, est un bon exemple de ce type de coopération⁵⁸.

Le centre de coordination devra également **élaborer une stratégie de découvrabilité adaptée aux contenus scientifiques qui placera en son cœur la traduction automatique et ses progrès**⁵⁹. Dans ce but, nous sommes d'avis que le centre de coordination devrait d'abord évaluer la visibilité des publications scientifiques en français sur les principales plateformes de diffusion scientifique et les moteurs de recherche. Il devrait ensuite élaborer des lignes directrices sur le balisage des publications scientifiques traduites, la traduction des métadonnées ou le dépôt de traductions dans des référentiels de données, comme GitHub ou Dryad. À partir de ces lignes directrices, il devrait concevoir des formations en rédaction et en publication scientifiques.

⁵⁶ Fiorini et autres, 2020.⁵⁷ Traductions et science ouverte, 2023.⁵⁸ FRQ, 2023⁵⁹ L'enjeu de la découvrabilité fait déjà l'objet d'une concertation importante à l'échelle de la francophonie par le biais de la culture. En effet, la promotion d'un espace numérique francophone fait partie des engagements prioritaires de la Francophonie, confirmés à Djerba en 2022. Une coopération entre le Québec et la France a d'ailleurs permis d'établir les principes de la découvrabilité des contenus culturels francophones en ligne. À ce sujet, voir ministère de la Culture et des Communications, 2020.

Le centre de coordination pourrait également organiser, auprès de la communauté étudiante francophone, des concours d'édition de contenu scientifique sur Wikipédia ou dans d'autres plateformes. L'édition de contenu consiste à sélectionner, éditer et partager les contenus en ligne les plus pertinents pour un sujet donné. Ainsi, de tels concours viseraient le référencement des meilleures publications de langue française sur Wikipédia, la plus grande plateforme de diffusion multilingue de contenu scientifique à l'échelle mondiale. Une telle initiative contribuerait à la diffusion en ligne de contenus scientifiques en français. De plus, elle sensibiliserait la communauté étudiante à la place du français en science et elle enrichirait les outils pédagogiques francophones dans un contexte de changements technologiques.

Pour promouvoir la traduction comme un instrument de valorisation de la science, le centre de coordination devra aussi **établir des partenariats et négocier des accords avec les grandes institutions scientifiques, les plateformes de diffusion, les moteurs de recherche et les éditeurs scientifiques mondiaux**. En parallèle, il devra **inscrire la question de la traduction scientifique dans les débats, plus larges, du multilinguisme scientifique et de la science ouverte** lors de ses représentations auprès d'organisations internationales comme l'Organisation internationale de la Francophonie ou l'Organisation des Nations Unies pour l'éducation, la science et la culture.

Plus particulièrement, le centre devra **entreprendre des démarches pour favoriser le balisage des publications scientifiques traduites, la traduction des métadonnées, le dépôt de traductions dans des référentiels de données et la libération de textes de leurs droits d'auteur pour leur traduction. À plus long terme, lorsque la performance des moteurs de traduction automatique le permettra, des discussions sur la possibilité d'intégrer de façon systématique ce type d'outil dans le processus d'édition scientifique devront avoir lieu**. Il nous est permis de croire à un renversement du fardeau de la traduction, qui reposera alors sur les épaules des institutions scientifiques plutôt que sur celles des producteurs du savoir.

Comme nous l'avons définie, la mission du centre ne se limitera pas à la mise en réseau ou à la veille stratégique. En effet, les retombées de ses actions seront très concrètes pour la communauté scientifique francophone, qui cherche à rayonner non seulement en français, mais aussi à l'international. En ayant l'assurance que leurs travaux sont accessibles, peu importe la langue de départ dans laquelle ils les ont rédigés, les scientifiques seront plus enclins à publier dans leur langue maternelle. Ils seront aussi davantage portés à investir les infrastructures et les plateformes de diffusion de la francophonie scientifique.

Conclusion

Par cet avis, notre objectif est de montrer comment la traduction automatique peut contribuer à valoriser et à renforcer la présence de la langue française dans le milieu scientifique, au Québec comme ailleurs. Comme nous l'avons exposé, le laisser-faire en la matière peut entraîner des problèmes liés à une mauvaise utilisation des technologies de traduction. Nous avons aussi mis en lumière qu'une sous-utilisation de ces technologies serait une occasion manquée pour rétablir le français comme l'une des langues du savoir.

Face à ces enjeux, nous avons proposé une approche structurée de l'usage de la traduction automatique dans le milieu scientifique. Cette approche se décline en trois propositions : la réalisation d'un réaménagement linguistique des universités québécoises, la mise en place d'un pôle d'expertise québécois et la création d'un centre de coordination à l'échelle de la francophonie. Nous avons imaginé ces trois propositions comme faisant partie d'un tout et s'imbriquant les unes dans les autres. Nous croyons que la traduction automatique peut facilement s'intégrer à des démarches déjà en cours et profiter de l'expertise et du savoir existants. Bien sûr, nos propositions misent sur la puissance des nouvelles technologies, ainsi que sur l'engagement actif des milieux scientifique et professionnel.

Nous espérons que les décideurs et les autres acteurs concernés par la place du français en science sauront saisir cette occasion et, peut-être, montrer la voie à d'autres milieux de pratique. Les technologies de la traduction continueront de s'améliorer. Les gens les utiliseront davantage, dans des situations de plus en plus diversifiées. À l'évidence, les mesures proposées dans cet avis pourraient s'appliquer dans d'autres contextes, comme en milieu de travail ou au sein de l'administration publique. Comme société, il est donc temps de maximiser le potentiel de la traduction comme outil d'aménagement linguistique.

Bibliographie

Amano, T., Ramírez-Castañeda, V., Berdejo-Espinola, V., Borokini, I., Chowdhury, S., Golivets, M., González-Trujillo, J. D., Montaña-Centellas, F., Paudel, K., White, R. L. et Verissimo, D. (2023). The manifold costs of being a non-native English speaker in science. *PLOS Biology*, 21(7), article e3002184. <https://doi.org/10.1371/journal.pbio.3002184>

Artetxe, M., Labaka, G. et Agirre, E. (2019). An Effective Approach to Unsupervised Machine Translation. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 194-203. <https://doi.org/10.18653/v1/P19-1019>

Bonnissent, J.-C. et de Sinety, P. (2021). *Pour des sciences en français et dans d'autres langues*. Honoré Champion.

Bowker, L. (2023). *De-mystifying Translation : Introducing Translation to Non-translators*. Routledge. <https://doi.org/10.4324/9781003217718>

Bowker, L. et Ciro, J. B. (2019). *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Emerald Publishing Limited. <https://doi.org/10.1108/9781787567214>

Cameron-Pesant, S. (2018). Usage et diffusion des revues savantes québécoises en sciences sociales et humaines : analyse des téléchargements de la plateforme Érudit. *Recherches sociographiques*, 59(3), 365-384. <https://doi.org/10.7202/1058719ar>

CBC News. (2021). Inuktitut added to language translator software. *CBC*. <https://www.cbc.ca/news/canada/north/inuktitut-added-to-language-translation-software-1.5889771>

Cœuré, S. (2019). La Babel des sciences. *Écrire l'histoire*, 19, 200-203. <https://journals.openedition.org/elh/2153>

Consortium Érudit. (2021a). *Droit d'auteur et licences Creative Commons* — Note de recherche. <https://www.erudit.org/public/documents/Droitdauteur.pdf>

Consortium Érudit. (2021b). *Principes et politiques du libre accès* — Note de recherche. https://www.erudit.org/public/documents/libre_acces.pdf

Costa-jussà, M. R., Cross, J., Çelebi, O., Elbayad, M., Heafield, K., Heffernan, K., Kalbassi, E., Lam, J., Licht, D., Maillard, J., Sun, A., Wang, S., Wenzek, G., Youngblood, A., Akula, B., Barrault, L., Gonzalez, G. M., Hansanti, P., Hoffman... Wang, J. (2022). *No Language Left Behind: Scaling Human-Centered Machine Translation* (document soumis pour publication). Meta AI, UC Berkeley et John Hopkins University. <https://arxiv.org/abs/2207.04672>

Daoust, J.-F., Gagnon, A.-G. et Galipeau, T. (2022). Language Matters: The Study of Canadian Politics through an Exploration of Syllabi and Comprehensive Exams. *Canadian Journal of Political Science/Revue canadienne de science politique*, 55(4), 897-915. <https://doi.org/10.1017/S0008423922000749>

De Swaan, A. (2013). *Words of the world: The global language system*. John Wiley & Sons.

Dion, J. (2012). *Le défi de former une relève scientifique d'expression française : l'usage du français et de l'anglais dans la formation universitaire aux cycles supérieurs au Québec*. Conseil supérieur de la langue française. <https://numerique.banq.qc.ca/patrimoine/details/52327/2249506>

eCampus Ontario. (2023). *Appel au contenu libre* — Soumission rémunérée de Ressources éducatives libres. <https://www.ecampusontario.ca/wp-content/uploads/2023/09/Call-for-Content-FR-Tagging-Update.pdf>

Elsevier. (2023). *Why Elsevier Language Editing services?* <https://webshop.elsevier.com/language-translation-services/>

Fan, A., Bhosale, S., Schwenk, H., Ma, Z., El-Kishky, A., Goyal, S., Baines, M., Celebi, O., Wenzek, G., Chaudhary, V., Goyal, N., Birch, T., Liptchinsky, V., Edunov, S., Grave, E., Auli, M. et Joulin, A. (2021). Beyond english-centric multilingual machine translation. *The Journal of Machine Learning Research*, 22(107), 1-48. <https://jmlr.org/papers/v22/20-1307.html>

Federation Of Finnish Learned Societies, Information, T. C. F. P., Publishing, T. F. A. F. S., Universities Norway, et European Network For Research Evaluation In The Social Sciences And The Humanities. (2019). *Helsinki Initiative on Multilingualism in Scholarly Communication*. <https://doi.org/10.6084/M9.FIGSHARE.7887059>

Fiorini, S. (2022). Traduction automatique et édition scientifique. *Traduire. Revue française de la traduction*, 246, 36-45. <https://doi.org/10.4000/traduire.2805>

Fiorini, S. (2023, 17 mai). Cartographie et collecte de corpus scientifiques bilingues : un premier bilan de l'étude. *Traductions et science ouverte*. <https://tradso.hypotheses.org/285>

Fiorini, S., Barbin, F., Garnier-Rizet, M., Morin, K. H., Humphreys, F., Josselin-Leray, A., Kübler, N., Loock, R., Martikainen, H., Nominé, J.-F., Plag, C., Rossi, C. et Yvon, F. (2020). *Rapport du groupe de travail « Traductions et science ouverte »*. Comité pour la science ouverte. <https://doi.org/10.52949/20>

Fonds de recherche du Québec. (2023). *Programme Samuel-De-Champlain 2023-2024*. <https://frq.gouv.qc.ca/programme/programme-samuel-de-champlain-2023-2024/>

Gingras, Y. et Mosbah-Natanson, S. (2010). Les sciences sociales françaises entre ancrage local et visibilité internationale. *European Journal of Sociology / Archives européennes de sociologie*, 51(2), 305-321. <https://doi.org/10.1017/S0003975610000147>

Gordin, M. D. (2015). *Scientific Babel: How science was done before and after global English*. University of Chicago Press.

Grin, F. (2017). Translation and language policy in the dynamics of multilingualism. *International Journal of the Sociology of Language*, 2017(243), 155-181. <https://doi.org/10.1515/ijsl-2016-0051>

Guthertz, G., Gordin, S., Sáenz, L., Levy, O. et Berant, J. (2023). Translating Akkadian to English with neural machine translation. *PNAS Nexus*, 2(5). <https://doi.org/10.1093/pnasnexus/pgad096>

Holt, B., Loock, R. et Lechauguette, S. (2022). Dealing with the «elephant in the classroom»: Developing language students' machine translation literacy. *Australian Journal of Applied Linguistics*, 5(3), 118-134. <https://doi.org/10.29140/ajal.v5n3.53si2>

Klimova, B., Pikhart, M., Delorme Benites, A., Lehr, C. et Sanchez-Stockhammer, C. (2022). Neural machine translation in foreign language teaching and learning: a systematic review. *Education and Information Technologies*, 28, 663-682. <https://doi.org/10.1007/s10639-022-11194-2>

Kübler, N. (2021). Le français, une langue scientifique internationale : résister au discours « prédateur » de l'anglais scientifique. Dans J.-C. Bonnissent et P. de Sinety (dir.) *Pour des sciences en français et dans d'autres langues*. Honoré Champion.

Laboratoire interdisciplinaire de recherche sur l'enseignement supérieur. (2023). *La place du français en enseignement supérieur au Québec — Rapport de recherche*. https://www.lires.ca/_files/ugd/2474ec_2a9f3f4add6f4e9aa72a3fb953e0eadc.pdf

Larivière, V. (2018). Le français, langue seconde ? De l'évolution des lieux et langues de publication des chercheurs au Québec, en France et en Allemagne. *Recherches sociographiques*, 59(3), 339-363. <https://doi.org/10.7202/1058718ar>

Larivière, V. (2014, 15 septembre). De l'importance des revues de recherche nationales. *Magazine de l'Acfas*. <https://www.acfas.ca/publications/magazine/2014/09/l-importance-revues-recherche-nationales>

Larivière, V. et Gaudreault-DesBiens, J.-F. (2022). *Mémoire sur la recherche et la publication scientifique en français*. <https://www.ourcommons.ca/Content/Committee/441/SRSR/Brief/BR12145326/br-external/ jointly1-f.pdf>

Larivière, V. et Riddles, A. (2021, 19 novembre). Langues de diffusion des connaissances : quelle place reste-t-il pour le français? *Magazine de l'Acfas*. <https://www.acfas.ca/publications/magazine/2021/11/langues-diffusion-connaissances-quelle-place-reste-t-il-francais>

Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157-175. <https://doi.org/10.1080/09588221.2018.1553186>

Longfield, L. (2022). *Un nouvel élan à la recherche et la publication scientifique en français au Canada*. Comité permanent de la science et de la recherche. <https://www.ourcommons.ca/DocumentViewer/fr/44-1/SRSR/rapport-5/>

Loock, R. (2016). *La traductologie de corpus*. Presses universitaires du Septentrion.

Loock, R. (2020, 6 mars). Pour une approche raisonnée de la traduction automatique. *Journée d'études Mots/Machines #2 : Traduction et sentiments : Quel(s) sens pour la machine ?* <https://hal.science/hal-02454708>

Machine Translation Literacy Project. (2023). *Machine Translation Literacy*. <https://sites.google.com/view/machinetranslationliteracy/>

Microsoft. (2023). *Government of Nunavut promotes Indigenous languages with groundbreaking Microsoft machine translation*. Microsoft Customers Stories. <https://customers.microsoft.com/en-au/story/1548201405354496852-government-of-nunavut-azure-cognitive-services>

Ministère de la Culture et des Communications. (2020). *Découvrabilité en ligne des contenus culturels francophones*. <https://www.quebec.ca/culture/aide-financiere/découvrabilité-contenus-culturels-francophones>

Mohammadi, E., Thelwall, M., Haustein, S. et Larivière, V. (2015). Who reads research articles? An altmetrics analysis of Mendeley user categories. *Journal of the Association for Information Science and Technology*, 66(9), 1832-1846. <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/asi.23286>

Moniz, H. et Escartín, C. P. (2023). *Towards Responsible Machine Translation: Ethical and Legal Considerations in Machine Translation* (vol. 4). Springer Nature.

Montgomery, S. L. (2013). *Does Science Need a Global Language? – English and the Future of Research*. University of Chicago Press.

Mundt, K. et Groves, M. (2016).
A double-edged sword: the merits and the policy implications of Google Translate in higher education. *European Journal of Higher Education*, 6(4), 387-401.
<https://doi.org/10.1080/21568235.2016.1172248>

Organisation des Nations unies pour l'éducation, la science et la culture. (2021).
Recommandation de l'UNESCO sur une science ouverte.
https://unesdoc.unesco.org/ark:/48223/pf0000379949_fre

Paterson, K. (2023).
Machine translation in higher education: Perceptions, policy, and pedagogy. *TESOL Journal*, 14(2).
<https://doi.org/10.1002/tesj.690>

Pecman, M. (2007).
Approche onomasiologique de la langue scientifique générale. *Revue française de linguistique appliquée*, XII(2), 79-96.
<https://doi.org/10.3917/rfla.122.0079>

Pitman, J. (2021, 28 avril).
Google Translate : One billion installs, one billion stories. Google.
<https://blog.google/products/translate/one-billion-installs>

Robichaud, D. (2017).
Political philosophy and scientific translation: When individual interest does not translate into collective benefits. Dans G. G. Núñez et R. Meylaerts (Éd.), *Translation and Public Policy: Interdisciplinary Perspectives and Case Studies*. Routledge.

Robichaud, D. et De Schutter, H. (2012).
Language is just a tool! On the instrumentalist approach to language. Dans B. Spolsky (dir.), *The Cambridge Handbook of Language Policy* (p. 124-146). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511979026.009>

Rovira, C., Codina, L. et Lopezosa, C. (2021).
Language Bias in the Google Scholar Ranking Algorithm. *Future Internet*, 13(2), 31.
<https://doi.org/10.3390/fi13020031>

Roy, J.-H. (2023, 9 mai).
En 2023, l'anglais deviendra la principale langue des thèses et mémoires au Québec. *The Conversation*.
<http://theconversation.com/en-2023-langlais-deviendra-la-principale-langue-des-theses-et-memoires-au-quebec-205077>

Siemund, P. (2023).
Multilingual development: English in a global context. Cambridge University Press.

Steigerwald, E., Ramírez-Castañeda, V., Brandt, D. Y. C., Báldi, A., Shapiro, J. T., Bowker, L. et Tarvin, R. D. (2022).
Overcoming Language Barriers in Academia: Machine Translation Tools and a Vision for a Multilingual Future. *BioScience*, 72(10), 988-998.
<https://doi.org/10.1093/biosci/biac062>

St-Onge, S., Forgues, É., Larivière, V., Riddle, A. et Volkanova, V. (2021).
Portrait et défis de la recherche en français en contexte minoritaire au Canada. Acfas.
https://www.acfas.ca/sites/default/files/documents_utiles/rapport_francophonie_final_1.pdf

Tchéhouali, D. et Agbobli, C. (2020).
État des lieux de la découvrabilité et de l'accès aux contenus culturels francophones sur internet.
<https://www.decouvrabilite-francophonie.net/wp-content/uploads/2020/12/Etat-des-lieux.pdf>

Traductions et science ouverte. (2023).
Projet.
<https://tradso.hypotheses.org/le-projet>

Université d'Ottawa. (2023).
Service de soutien linguistique pour le corps professoral.
<https://www.uottawa.ca/notre-universite/institut-langues-officielles-bilinguisme/programmes-cours/cours-personnel-academique>

Weaver, W. (1955).
Translation. Dans W. N. Locke et A. D. Booth (dir.), *Machine Translation of Languages* (p. 15-23). MIT Press.

Warren Weaver (2023, 18 juillet).
Dans *Wikipédia*.
https://en.wikipedia.org/w/index.php?title=Warren_Weaver&oldid=1165939098

Wilkinson, R. et Gabriëls, R. (2021).
The Englishization of Higher Education in Europe. Amsterdam University Press.
<https://doi.org/10.1515/9789048553914>

Remerciements

Nous remercions les personnes et les organisations suivantes d'avoir commenté cet avis :

Christian Agbobli (Université du Québec à Montréal)

Éric Archambault (consortium Érudit)

Donald Barabé (Ordre des traducteurs, terminologues et interprètes agréés du Québec)

Lynne Bowker (Université d'Ottawa)

Guillaume Chicoisne (IVADO)

Monique Cormier (Université de Montréal)

Patrick Drouin (Université de Montréal)

Susanna Fiorini (OPERAS)

Fonds de recherche du Québec

Jean-François Gaudreault-DesBiens (Université de Montréal)

Office québécois de la langue française

Rudy Loock (Université de Lille)

Sophie Montreuil (Acfas)

Tanja Niemann (consortium Érudit)

Jean-Christian Pleau (Université du Québec à Montréal)

Etleva Vocaj (Université de Montréal)

Grégoire Winterstein (Université du Québec à Montréal)

