

LA TRADUCTION AUTOMATIQUE ET ASSISTÉE PAR ORDINATEUR DANS LES PROGRAMMES UNIVERSITAIRES DE FORMATION FRANÇAIS ET ITALIENS

Introduction

L'étude qui fait l'objet du présent article constitue la deuxième phase d'une recherche qui a été lancée en 2017.

À l'époque, la première phase de notre recherche a été basée sur un corpus comparable (FR-IT) constitué pour comparer les descriptifs des offres de formation des masters en traduction¹, collectées à partir des sites Web d'une sélection d'universités françaises et italiennes. L'objectif de ce travail préparatoire était de mettre en lumière les écarts et convergences concernant la formulation des objectifs de formation et des compétences visées par les masters en traduction français et italiens. Les résultats de cette phase de notre étude (Cennamo et Hamon, 2021) ont mis en avant une convergence franco-italienne concernant la place centrale occupée par la professionnalisation au sein des offres de formation de niveau master. Parallèlement, une divergence a été observée au niveau de la mise en discours des compétences visées par les masters : en France, l'accent était mis entre autres sur le développement de compétences technologiques (European Master's in Translation [EMT], 2017 : 9) chez les futurs traducteurs, alors qu'en Italie, au cœur des masters, figuraient les compétences traductives (EMT, 2017 : 7 et 8) de nature davantage méthodologique, axées notamment sur l'acquisition de stratégies de traduction visant les multiples contextes de la profession.

1. C'est-à-dire les descriptifs génériques de présentation des masters sur les sites Internet des formations en question.

Après avoir enquêté sur la présentation des offres de formation, nous avons décidé d'approfondir notre recherche en procédant à une analyse des programmes de formation à la traduction outillée, de niveau master 1 et 2, afin d'accéder aux contenus et aux méthodes prévus par ces enseignements.

Par conséquent, la deuxième phase de notre recherche (cf. Méthodologie de recherche) a été consacrée à l'analyse des mises en discours des termes (FR-IT) liés aux compétences technologiques de Traduction automatique (TA) et Traduction assistée par ordinateur (TAO) visées par ces programmes pour comparer les pratiques de formation à la traduction outillée, dans les deux pays. Nous avons retenu les programmes² des « cours transversaux » axés sur l'acquisition de compétences technologiques appliquées à la traduction professionnelle, c'est-à-dire des cours qui s'adressent à tous les inscrits au master, indépendamment de leurs langues de spécialisation. L'orientation transversale des formations fortement professionnalisantes constitue un aspect saillant pour notre enquête puisqu'elle se veut représentative des compétences technologiques considérées comme fondamentales pour le parcours de master en traduction, et ce conformément à la conception des fiches de formation élaborée par les responsables pédagogiques.

Dans le cadre de notre réflexion traductologique, nous avons pris en compte le modèle de compétence de traduction conçu dans le cadre du référentiel européen des masters en traduction (EMT, 2017), dont nous souhaitons reprendre notamment trois points qui s'avèrent intéressants pour notre recherche.

Le réseau EMT rappelle en premier lieu la nature stratégique de chaque étape de l'opération de traduction (EMT, 2017 : 7), à partir de la définition du projet de traduction jusqu'à la révision du produit final. Par ailleurs, l'interaction avec la machine est définie aujourd'hui comme un versant incontournable de la traduction professionnelle³ (EMT, 2017 : 7), ce qui implique une nouvelle conception des opérations d'évaluation (EMT, 2017 : 9) qui intègrent à la recherche de l'équivalence, celle de la 'pertinence' des choix de traduction.

La conception de la traduction comme un processus contextualisé visant la recherche de mises en équivalences pertinentes (Le Disez, 2013 ; Lederer, 2006), reste aujourd'hui un préalable fondamental pour la forma-

2. Valables pour l'année académique 2019-2020.

3. À cet égard, il s'avère particulièrement intéressant de noter que l'édition 2019 du rapport sur l'industrie des langues publié dans le contexte européen (Language Industry Survey, 2019 : 23) a confirmé la centralité de la TAO dans la pratique de la profession, en indiquant que la TA s'intègre progressivement au flux de travail de la TAO tout en faisant l'objet d'usages hétérogènes. Tout particulièrement, le rapport montre qu'à l'ère de la Traduction automatique neuronale (TAN), si d'une part les sociétés de traduction en 2019 déclaraient prévoir des investissements de plus en plus importants dans la TAN, d'autre part chez les traducteurs indépendants le recours à la TA n'a pas été déclaré comme étant une pratique établie au même degré que la TAO, l'utilité de la TA étant l'objet de perplexités chez les professionnels indépendants de la traduction.

tion à la traduction outillée (Lavault-Olléon et Zimina, 2019; Loock, 2019; Martikainen, 2019; Kenny, 2020; O'Brien, 2020).

Par ailleurs, face à la vague des applications issues de l'Intelligence artificielle (IA) dans l'industrie des langues (O'Brien et Rossetti, 2021), il faut souligner que l'intérêt de cette conception de la traduction ne s'applique pas seulement à la formation de futurs traducteurs spécialisés, mais également à tout parcours de formation à la traduction⁴ lorsqu'il prévoit et/ou admet le recours aux technologies pour la traduction (dictionnaires en ligne, corpus, traducteurs automatiques, mémoires de traduction, logiciels pour l'alignement, etc.). En effet, les résultats obtenus dans le cadre d'expérimentations pédagogiques récentes, centrées sur l'interaction avec la machine, ont signalé la nécessité d'un repositionnement de l'apprenant face à la machine (Ehrensberger-Dow et Massey 2019; Massey et Kiraly, 2019; Rossi, 2019) qui soit axé sur la complémentarité des stratégies de traduction, humaines et automatiques. À la lumière de ces réflexions traductologiques, nous nous sommes interrogés sur les traits caractérisant la conception de la traduction outillée ainsi que sur les modalités d'interaction avec la machine, telles qu'elles sont proposées par les programmes de master en traduction en France et en Italie.

Méthodologie de recherche

Notre enquête se situe dans le cadre la méthodologie de recherche élaborée par les traductologues de corpus (Loock, 2016).

Selon cette méthode, la première étape de notre recherche a concerné la constitution des deux sous-corpus comparables (FR-IT) qui se veulent représentatifs des programmes de cours de traduction outillée (année 2019-2020), de niveau master 1 et 2, en France et en Italie.

Pour la constitution du sous-corpus français (13648 mots), nous avons retenu les programmes de cours accessibles depuis les sites Web des universités françaises membres de l'EMT 2019-2024⁵; en ce qui concerne le sous-corpus italien (10940 mots), nous avons pris en compte les quatre universités membres de l'EMT 2019-2024 ainsi que les sept masters (non EMT) appartenant à la classe ministérielle LM94⁶, et ce en raison de l'écart existant entre le nombre des membres EMT français et le nombre des membres EMT italiens.

4. Cours d'initiation à la traduction, cours de médiation linguistique, cours de traduction pédagogique appliquée à l'acquisition de compétences langagières.

5. La liste des universités membres de l'EMT est accessible au lien suivant : [https://ec.europa.eu/info/resources-partners/european-masters-translation-emt/list-emt-members-2019-2024_fr], consulté le 08-09-2021.

6. La classe ministérielle LM94 comprend les masters visant la formation de traducteurs et interprètes professionnels, comme prévu par le décret ministériel du 16 mars 2007 (Supplément ordinaire de la G.U. n° 155 du 06/07/2007). La liste des masters LM94 est accessible sur [<http://www.cestor.it/atenei/lm094.htm>], consulté le 09-09-2021.

Notre comparaison est fondée notamment sur treize⁷ programmes de cours français et sur seize programmes de cours italiens renvoyant à la même typologie de cours transversaux (cf. Introduction) consacrés à la traduction outillée. Nous précisons que, quant aux programmes proposés par les universités françaises, nous avons identifié trente-quatre cours consacrés aux technologies de la traduction, dont treize programmes seulement étaient détaillés en termes de contenu pédagogique. Du côté italien, parmi les seize programmes retenus, seuls huit programmes sont proposés par les universités membres de l'EMT. Les huit autres programmes de cours sont proposés par les universités de la classe ministérielle LM94 mentionnés plus haut.

Les programmes de cours de traduction outillée qui constituent notre corpus comparable FR-IT ont fait l'objet d'une analyse quantitative⁸ et qualitative : les termes significatifs⁹ liés aux compétences technologiques ont été analysés conformément à une approche socioterminologique où le terme est envisagé comme produit social au sein d'une communauté de pratique (Wenger 1998, 2009) qui constitue en même temps une communauté discursive¹⁰ (Gaudin, 2005 ; Cabré, 2000 ; Cabré, 2016 : 75). Dans le cadre de notre recherche, cette communauté de pratique est représentée par les responsables de cours et les équipes pédagogiques. Notre analyse est centrée sur l'emploi et la mise en discours des unités terminologiques concernant la TA, la TAO, la post-édition, l'IA et la TAN, ainsi que celles qui relèvent, plus généralement, de compétences traductives outillées (cf. Analyse qualitative des résultats : exploration du corpus).

Analyse quantitative : comparaisons de premier niveau

Nous avons recueilli pour chaque programme de cours, les informations concernant quatre points essentiels qui ne relèvent pas de la mise en discours proprement dite, mais qui permettent de dégager :

- 1) le nombre de crédits de chaque cours ;
- 2) les objectifs de la formation ;
- 3) les compétences visées et méthode d'enseignement ;
- 4) les modalités d'examen (entendues en tant que vérification des acquis).

Les résultats de nos comparaisons de premier niveau nous ont montré que les crédits de formation attribués à chaque cours (European Transfer and Accumulation System, ou Système européen de transfert et d'accumulation de crédits [ECTS] français et Crediti Formativi Universitari [CFU] italiens) sont variables aussi bien en France (les ECTS vont de 4 à 9) qu'en Italie (les

7. Pour deux des 15 universités françaises labellisées EMT, nous n'avons pas pu accéder aux programmes de cours détaillés. Les rares informations proposées étaient trop lacunaires pour être intégrées au corpus.

8. [<https://www.laurenceanthony.net/software/antconc/>], consulté le 08-09-2021.

9. Du point de vue quantitatif, en termes d'occurrence dans le corpus.

10. La prise en compte des concordances des termes significatifs a été complétée par l'analyse qualitative de la mise en discours du terme au sein de son programme.

CFU de 5 à 9). Ensuite, pour les deux pays, les objectifs de la formation sont indiqués en fonction des différentes unités d'enseignement constitutives de chaque programme, qui sont généralement prévues afin de distinguer les différentes compétences envisagées¹¹.

En ce qui concerne les compétences visées et la méthode d'enseignement, on relève une différenciation significative dans les deux sous-corpus : presque la moitié des programmes français retenus (5 sur 13) prévoient la mention explicite des logiciels adoptés dans l'enseignement, alors que seul un quart des programmes italiens retenus (4 sur 16) inclut cette indication¹². En guise d'exemple, dans le master Rédacteur/Traducteur proposé par l'université de Brest, pour le cours de M2 TAO et localisation, les logiciels spécifiques sont précisés dans la partie « contenu détaillé » de la page Internet du programme de cours en question :

« Perfectionnement en TAO avec étude des fonctionnalités avancées de Memsources, Studio et MemoQ (traitement des fichiers, formatage, gestion de la terminologie, assurance qualité et facturation). »

Afin de définir des entrées lexicales quantitativement significatives à partir desquelles observer leurs mises en discours, nous avons utilisé la fonction de recherche de mots-clés sur le logiciel Antconc. Nous nous sommes limités aux dix mots les plus récurrents et les plus pertinents pour l'objet de notre recherche, obtenant ainsi un ensemble de données de départ que nous restituons ci-dessous sous forme de tableau :

384 occurrences : traduction	164 occurrences <i>traduzione</i>
89 occurrences : outils	81 occurrences : <i>translation</i> (EN)
82 occurrences : étudiants	75 occurrences : <i>corso</i>
66 occurrences : cours	59 occurrences : <i>strumenti</i>
62 occurrences : gestion	48 occurrences : <i>studenti</i>
58 occurrences : traducteur(s)	43 occurrences : <i>automatica</i>
54 occurrences : technique	43 occurrences : <i>online</i>
53 occurrences : formation	42 occurrences : <i>testi</i>
50 occurrences : compétences	41 occurrences : <i>post-editing</i>
49 occurrences : projet	37 occurrences : <i>assistita</i>

Tab. 1. – Les dix entrées lexicales du corpus les plus fréquentes.

Source : GENNAMO et HAMON.

11. Les unités d'enseignement concernent généralement la distinction entre TA, TAO et gestion de la terminologie. C'est en effet à partir de l'analyse des occurrences des termes en rapport avec la TA, la TAO et la gestion de terminologie que nous avons pu identifier cette articulation des unités d'enseignement.

12. L'analyse des compétences sera abordée plus en détail dans le paragraphe suivant (cf. « Analyse qualitative des résultats : exploration du corpus »).

Ce cadre quantitatif tend à confirmer une orientation professionnalisante plus marquée au sein des programmes français par rapport aux programmes italiens, qui semblent insister davantage sur l'acquisition des principales compétences instrumentales.

En effet, le sous-corpus italien tend à se caractériser, à première vue, par la référence aux trois pratiques technologiques principales : la TAO (37 occurrences), la TA (43 occurrences) et la post-édition (41 occurrences), qui sont également nommées en ayant recours aux termes anglais équivalents : *CAT tools*, *machine translation*, *post-editing*. En ce qui concerne la France, on peut signaler notamment que, si les termes TAO, TA et post-édition ne figurent pas parmi les dix mots les plus fréquents, ce sont les concordances des termes « traducteur » (58 occurrences) « gestion » (62 occurrences) et « projet » (49 occurrences) qui établissent le lien avec les pratiques de la profession. Nous signalons, en particulier, que les occurrences de « traducteur » au singulier renvoient le plus souvent au rôle du traducteur professionnel¹³, alors que les occurrences de « traducteurs » au pluriel concernent, quant à elles, les mises en discours relatives à la projection des apprentis traducteurs dans leur avenir de traducteurs professionnels¹⁴. À titre d'exemple, on trouve dans le corpus des formulations discursives de ce type :

« Les **futurs traducteurs** doivent être capables d'évaluer ces outils en fonction de leurs besoins et des textes qu'ils ont à traiter. Ils sont amenés à développer un regard critique sur l'utilisation de la TAO. »

On peut ici constater que la projection dans l'avenir professionnel des candidats potentiels à la formation suppose la capacité des étudiants à prendre du recul sur le choix et la pertinence des outils de TAO en fonction de la demande du client.

Cet autre exemple témoigne de l'association de la figure professionnelle à sa capacité à intégrer de manière raisonnée la multiplicité des outils d'aide à la traduction :

« L'objectif de cet enseignement est d'initier les **futurs traducteurs** à l'exploitation de corpus électroniques (banques de données linguistiques) comme outils d'aide à la traduction, aux côtés des outils traditionnels comme les dictionnaires, les glossaires ou encore les outils de TAO habituels dotés de mémoires de traduction. »

13. On peut à titre d'exemple trouver des concordances signalant la figure professionnelle générique, associée à la formulation des objectifs de certains cours « Savoir prendre du recul théorique par rapport à l'activité et au rôle du traducteur ».

14. Les étudiants sont à plusieurs reprises désignés comme futurs traducteurs : « Les futurs traducteurs doivent être capables d'évaluer ces outils en fonction de leurs besoins et des textes qu'ils ont à traiter. Ils sont amenés à développer un regard critique sur l'utilisation de la TAO. »

Ces premières observations nous amènent à nous pencher plus en détail sur les mises en discours des termes les plus significatifs et à aborder le volet qualitatif de notre étude.

Analyse qualitative des résultats : exploration du corpus

Dans le cadre de notre analyse qualitative des résultats, nous nous situons conformément à l'approche socioterminologique (Gaudin, 2003 ; 2005) nous permettant de concevoir la rédaction des programmes de cours universitaires sous le prisme de la notion de « pratique langagière » (Boutet, Simonin et Fiala, 1976).

Cette démarche propose « d'adopter une vision dynamique des termes, lesquels ne doivent plus être conçus comme des étiquettes de concepts, mais restitués dans le cadre des échanges langagiers au sein desquels ils apparaissent et se maintiennent » (Gaudin, 2005 : 86 et 87).

Pour ce qui est de la « traduction automatique », le terme a été saisi dans Antconc sous ses deux formes : forme développée et sigle TA. Dans les deux sous-corpus, le cotexte gauche met en avant une mise en discours double de type « théorique et appliquée ». En ce qui concerne la dimension théorique, cet aspect s'exprime dans les programmes de cours des masters français en termes d'évaluation et de compréhension avec des expressions telles que « comprendre », « avoir évalué », « compréhension des limites », « connaissances des principes et du fonctionnement », « questions causées par ». Il s'agit en outre d'aborder, en tant qu'initiation, l'historique de la traduction automatique. Les formes verbales dénotent également le signalement du niveau d'étude de la TA en tant qu'objet d'apprentissage « découvrir les bases », « familiariser ». Le corpus IT fait, quant à lui, émerger des formulations qui mettent en évidence, d'une part, l'accent mis sur la dimension théorique des programmes de cours italiens : *evoluzione della, principi teorici della, ambiti della, storia della*, etc. *metodologie per*, mais aussi, dans une certaine mesure la dimension pédagogique qui d'autre part présente ces contenus comme étant des cours d'initiation à la traduction automatique *conoscenze di base relative alla, rudimenti di, elementi di, introduzione alla, nozioni di*. Ainsi, sur le plan socioterminologique, la mise en discours des termes témoigne de leur inscription dans les champs et domaines de référence (didactique et traductologique). Cette double inscription témoigne de l'influence des contextes extralinguistiques sur l'apparition et la circulation des termes (Gérard, 2018).

Pour ce qui est de la dimension appliquée, la TA est associée à des segments tels que « post-édition », « analyse », « utilisation ». Dans le sous-corpus italien, cette dimension se manifeste par des termes tels que *sistemi/programmi di, motore(i) di*, qui renvoient plus spécifiquement à la dimension technologique, ce qui tendrait à confirmer l'impression d'ensemble véhiculée par la liste des termes les plus fréquents, à savoir pour le sous-corpus italien,

une mise en discours plus axée sur les outils et leur intégration dans les programmes de cours, tandis que le sous-corpus français se caractérise par une mise en discours qui signale davantage la visée professionnelle.

Si les deux sous-corpus sont fortement convergents, malgré tout, il est intéressant d'observer une singularité dans le corpus FR : l'apparition en cotexte droit d'un syntagme verbal plus discursif et réflexif où il est signalé que la TA « en train de bouleverser le secteur ».

Pour ce qui est de la TAO, dans le corpus FR, les cotextes (droit et gauche) sont centrés sur la « pratique » et dénotent un souci de désigner les produits technologiques utilisés (SDL Trados, OmegaT, Memsource). Cela se manifeste par des constructions verbales observables dans le cotexte gauche : « utiliser les logiciels de TAO », « utiliser l'informatique dédiée à la traduction et les outils de TAO », ou dans le cotexte droit « logiciel de TAO en ligne qui sera utilisé dans le projet tuteuré de traduction », « compléter les compétences en TAO en comparant activement plusieurs logiciels à mémoires de traduction concurrents sur le marché ». Des constructions nominales viennent aussi renforcer cet accent mis sur la pratique : « manipulation d'un logiciel de TAO », « Utilisation comparée des outils de TAO ». Ce dont témoigne également les exemples suivants : « pratique de » : « initiation » ou « introduction à », « maîtrise élémentaire ou approfondie de », « manipulation d'un logiciel de », « utilisation des outils de » où l'on remarque la mention du niveau d'étude de l'objet d'apprentissage. Cela rejoint l'observation que nous avons formulée avec les termes pédagogiques utilisés dans le sous-corpus italien, tendant ainsi à confirmer la convergence globale des mises en discours entre les deux sous-corpus.

Enfin, dans l'une des fiches d'offre de formation publiées sur le site de l'une des universités françaises, apparaît aussi un constat exprimé de façon plus narrative, mais qui revêt aussi une valeur prescriptive en mettant en valeur le postulat d'une amélioration de la qualité des produits : « la TAO fait désormais partie intégrante du métier de traducteur. Sa bonne utilisation permet d'automatiser certaines tâches tout en améliorant la cohérence des traductions ».

L'accent mis sur la pratique est également manifeste dans le sous-corpus italien, comme en témoignent les mises en discours relevées dans les cotextes droits et gauches. Cela passe par des constructions verbales : « *gestire e valutare progetti articolati di* », « *usare la* », ou des tournures plus articulées comme « *è auspicabile che gli studenti abbiano già preso confidenza con uno strumento per la* », « *utilizzare i principali applicativi software utilizzati nel campo della...* » ; ou par des syntagmes nominaux ou adjectivaux tels que « *pratica della* », « *una prova pratica di* », « *creazione di progetti e/o pacchetti* », « *l'impiego di memorie di traduzione* », « *l'uso di uno tra i principali strumenti di* », « *utilizzo di software di* ». Ainsi, les termes utilisés pour désigner les cours de TAO démontrent une convergence manifeste entre les deux corpus IT et FR. Les concepteurs des programmes de cours semblent vouloir, des deux côtés des Alpes, mettre

l'accent sur la manipulation des outils de travail professionnels des traductrices et traducteurs, sans pour autant renoncer à introduire une dimension analytique et comparative, comme dans le corpus FR : « Distinguer la TA de la TAO », ou bien comme nous l'avons vu plus haut avec la mention d'une comparaison active, ce qui tend à renforcer l'accent mis sur une utilisation raisonnée des outils de TA et de TAO.

Si les deux sous-corpus semblent fortement converger et opter pour des mises en discours mettant l'accent sur la complémentarité entre théorie et pratique, nous observons en revanche des divergences quant à la présence et à la présentation d'une autre composante de la compétence de traduction liée aux technologies : la post-édition. En effet, dans le corpus IT, le terme « post-editing » est bien représenté (41 occurrences), quasiment toutes sous la forme anglaise. La post-édition est moins représentée dans le corpus FR (19 occurrences). Toutefois, dans les deux cas, le terme « post-édition » est associé à la TA. Dans le sous-corpus français, nous observons des mises en discours intéressantes en ce qu'elles modalisent l'opération comme en témoignent les adjectifs qui lui sont associés (citation ci-après). Nous y constatons également le lien fort avec la dimension professionnelle de la post-édition (utilisation du terme « projets ») :

« Compétence(s) visée(s) : Les étudiants sont en mesure d'évaluer **la qualité d'un outil de TA et d'effectuer une post-édition complète, intelligente et rentable**. Les étudiants acquièrent des connaissances approfondies du logiciel SDL Trados Studio en **y post-éditant** des projets traduits par une machine.

La deuxième partie (8 heures) est consacrée à la pratique de la **post-édition**, c'est-à-dire la traduction humaine faite à partir d'une traduction automatique. Les étudiants testeront et évalueront la pertinence et l'efficacité de plusieurs **systèmes de TA**. »

Pour le corpus IT, la post-édition va également de pair avec la conception de projets de traduction. Cette tendance se manifeste aussi bien par des formes nominales (« *progetto reale* », « *incarichi di progetto* ») que par des constructions verbales (« *integrando traduzione assistita e post-editing* »), voire des développements discursifs plus complets mentionnant la gestion de projets en équipe (« *gestire e valutare progetti articolati di traduzione automatica e post-editing in cui intervengono diverse competenze operative distribuite su più persone*¹⁵ »). Les exemples suivants mettent en évidence, tout particulièrement, l'inscription de la TA dans une démarche professionnelle holistique basée sur son utilisation intégrée avec les outils de TAO :

15. Notre traduction : « gérer et évaluer des projets complexes de traduction automatique et de post-édition impliquant plusieurs compétences opérationnelles réparties sur plusieurs personnes ».

« Alla conoscenza dei **sistemi CAT** verranno anche affiancati dei rudimenti di **traduzione automatica e di relativo post-editing**, al fine di comprenderne le differenze e le possibili **strategie di utilizzo congiunto con i sistemi CAT**.

*Prova pratica: svolgimento di un piccolo **progetto di localizzazione integrato con post-editing** utilizzando **strumenti di traduzione assistita e automatica**¹⁶. »*

À ce stade, nous pouvons donc constater une légère divergence au niveau des mises en discours de la post-édition, étant généralement associée à la TA dans le sous-corpus français, alors qu'elle fait davantage l'objet d'une conception holistique sous forme d'un ensemble articulé TA/TAO/post-édition, dans le sous-corpus italien.

En ce qui concerne l'IA et la traduction neuronale (que nous appellerons TAN) en tant que nouveau paradigme technologique, notre intention n'est pas d'analyser séparément les mises en discours relevant de la TA dans son ensemble et de la TAN dans sa spécificité, mais de voir si l'accent est mis explicitement sur la méthode neuronale.

À cet égard, on observe dans les deux sous-corpus, la très faible représentation de ces deux désignations (IA et TAN) dans les programmes de cours. Les termes n'y apparaissent d'ailleurs ni sous leur forme entière, ni sous forme de sigles. Ils n'apparaissent qu'à deux reprises dans le corpus IT, dont une mise en discours de type définitionnel associant l'IA à la linguistique computationnelle et s'inscrivant ainsi dans un ancrage théorique fort :

« La linguistica computazionale è una disciplina che si occupa di definire e applicare modelli e strumenti per il trattamento del linguaggio naturale. È per sua natura interdisciplinare, combinando metodi e tecniche della linguistica con quelli dell'informatica, dell'intelligenza artificiale e della statistica¹⁷. »

Une autre mise en discours témoigne elle aussi d'une volonté de légitimer théoriquement l'intelligence artificielle en l'associant à des technologies déjà consolidées dans le champ professionnel, mais aussi des questionnements apparus dans la sphère scientifique :

« comprendere il funzionamento degli strumenti di traduzione automatica ed assistita. A tale scopo viene sinteticamente presentata la storia dell'intelligenza

16. Notre traduction : « La connaissance des systèmes de TAO sera également accompagnée d'une initiation à la traduction automatique et à la post-édition, afin de comprendre les différences et les stratégies possibles d'utilisation liées aux systèmes de TAO. Épreuve pratique : réalisation d'un petit projet de localisation en post-édition à l'aide d'outils de traduction assistée par ordinateur et de traduction automatique. »

17. Notre traduction : « La linguistique computationnelle est une discipline qui s'attache à définir et à appliquer des modèles et des outils pour le traitement du langage naturel. Elle est interdisciplinaire par nature, combinant les méthodes et techniques de la linguistique avec celles de l'informatique, de l'intelligence artificielle et des statistiques. »

*artificiale, mettendo in evidenza l'importanza che in essa ha assunto lo sviluppo dei sistemi di traduzione automatica e come il dibattito teorico su di essi ha portato all'idea della traduzione assistita*¹⁸ ».

Enfin, pour ce qui est des modalités d'examen, nous avons observé qu'en Italie, la majorité des programmes (13 sur 16) prévoit l'indication des références bibliographiques nécessaires à la préparation de l'examen, alors que du côté français cette pratique n'est attestée que dans un seul cas. Cela pourrait être interprété comme une volonté de légitimer la présence de technologies de la traduction en tant qu'objet théorique et appliqué d'enseignement/apprentissage, du côté des formations italiennes. Cette forme d'intertextualité (bibliographie) rentre de plein champ dans l'approche que nous avons adoptée et témoigne du lien fort entre le socioterminologique et l'approche discursive.

Si l'on isole la désignation de TAN, on constate que celle-ci est davantage mise en discours dans le sous-corpus FR avec des observations réflexives, dont une très modalisée :

« L'arrivée de la traduction automatique neuronale en 2016 est en train de modifier le modèle économique mais aussi les habitudes de travail. Loin de remplacer les traducteurs, cette nouvelle technologie que l'on peut aisément qualifier de « disruptive » doit être maîtrisée par les professionnels. »

L'offre de formation fait d'ailleurs valoir cette capacité d'adaptation : « Créé en 2004, le Master *** a été conçu en tenant compte de l'évolution rapide de ce marché. » Dans le corpus italien, les quatre occurrences obtenues avec le lemme « neur* » renvoient à des indications bibliographiques, hormis un cas de mise en discours qui, là encore, met l'accent sur la légitimation théorique d'une technologie très récente « *Si prosegue illustrando gli aspetti teorici dei sistemi di traduzione automatica basata su regole* (Rule Based Machine Translation ou RBMT), *statistica* (Statistical Machine Translation ou SMT) e *neurale* (Neural Machine Translation ou NMT) ». Ainsi, dans les deux cas, si la technologie est en soi peu présente, les deux sous-corpus témoignent d'une volonté de la situer sur les plans professionnel (sous-corpus FR) et théorique (sous-corpus IT).

En ce qui concerne le lien entre compétences et maniement des outils de traduction automatique et assistée, quelques divergences minimales apparaissent entre les deux sous-corpus. Si dans les deux corpus FR et IT, les termes couvrent le nuage de compétences technologiques dans son ensemble qu'il s'agisse de l'informatique de base (suite bureautique Office de Microsoft Windows) ou des outils spécialisés appliqués à la traduction (localisation et langages HTML- XML, gestion de bases de données terminologiques,

18. Notre traduction : « comprendre le fonctionnement des outils de traduction automatique et assistée par ordinateur. À cette fin, l'histoire de l'intelligence artificielle est brièvement présentée. L'accent est mis sur l'importance que revêt, dans ce cadre, le développement des systèmes de traduction automatique mais aussi son importance théorique dans la mesure où les travaux sur la TA ont permis l'émergence de la traduction assistée par ordinateur ».

découverte puis maîtrise de la TAO et de la TA, maîtrise des ressources numériques, post-édition dans les domaines spécialisés), le corpus FR se distingue par des renvois explicites au référentiel EMT, voire, dans un cas, à des renvois vers les numéros précis des compétences technologiques EMT correspondantes. Enfin, toujours dans le corpus FR, ces compétences liées aux technologies sont, dans certains cas, associées aux *soft-skills*. Il est en effet question dans le corpus de compétences « comportementales », « personnelles » et « interpersonnelles ». Dans les deux sous-corpus, les concordances du terme soulignent toutefois la pluralité et l'hétérogénéité des compétences technologiques au-delà des seules traductions automatiques et assistées. Au sein même du corpus IT, des divergences apparaissent entre les formations labellisées EMT et les autres universités. En effet, dans les formations EMT, TA et TAO sont envisagées comme deux compétences articulées et complémentaires, tandis que dans les masters italiens non EMT (classe LM94), les deux pratiques technologiques sont envisagées le plus souvent séparément.

Les programmes du corpus FR proposent, quant à eux, des associations entre les compétences liées à la TAO et les compétences liées à d'autres outils (outils de localisation par exemple), mais globalement, les intitulés français « flèchent » la TAO comme un objet de contenus autonome.

Toutefois, au-delà de certains choix distincts dans les formations françaises et italiennes, les programmes concernant les compétences technologiques semblent se caractériser par une certaine homogénéité dans les stratégies discursives. Cela se manifeste notamment par l'accent mis sur la capacité des futurs traducteurs à évaluer la pertinence des instruments d'aide à la traduction, en fonction des différents projets et de la nature des prestations professionnelles. Cela témoigne d'une volonté des concepteurs de programme de se baser sur le référentiel EMT et notamment sur les compétences de « traduction » n° 10 à 14 (EMT, 2017 : 7 et 8) et les compétences « technologiques » n° 15-16, 18-19 (EMT, 2017 : 9) du référentiel.

Dans les deux sous-corpus, il ressort du « fonctionnement discursif des termes » (Gaudin, 2005 : 91) qu'il s'agit avant tout, pour les responsables de cursus, de former des étudiants capables d'apprécier la pertinence des outils en fonction du projet de traduction et de les doter d'un bagage théorique et praxéologique leur permettant d'opérer des choix pertinents en tant que futures traductrices ou futurs traducteurs.

Conclusions et perspectives de recherche

Malgré les quelques divergences observées (la place plus importante accordée à la post-édition dans le sous-corpus IT et les renvois explicites au référentiel EMT dans le sous-corpus FR), l'exploration socioterminologique de notre corpus comparable a mis en lumière un alignement des programmes de cours dans les masters italiens et français qui tendent à converger dans

l'adoption d'une stratégie sociocommunicative de valorisation des programmes orientée sur l'alternance entre théorie et pratique. Tout particulièrement, notre démarche socioterminologique a permis de mieux cerner l'interaction entre deux champs sociaux : le monde de l'université et le monde du travail. Tandis que le champ universitaire est présent dans des mises en discours qui renvoient à l'acquisition de connaissances et de compétences traductives théoriques et appliquées, le champ professionnel transparait par le biais de formulations discursives orientées sur la traduction conçue comme prestation de service. Si, au vu des fréquences des termes utilisés pour définir les objectifs liés à la compétence technologique, la différence la plus notable réside dans l'orientation plus professionnalisante du sous-corpus FR et une orientation plus instrumentale dans le sous-corpus IT, nous constatons dans les deux sous-corpus que l'IA et la TAN sont finalement peu mentionnées explicitement en tant que méthode spécifique de TA. Ce constat soulève des interrogations quant à la prudence adoptée par les personnes chargées de rédiger les programmes de cours. Cela signifie-t-il que la méthode de calcul et les algorithmes propres à l'IA sont encore finalement peu communiqués dans les pratiques professionnelles, ou que les formations nécessitent une recherche poussée d'experts à même d'assurer des cours pour un usage actif des technologies s'appuyant sur l'IA et la TAN, en s'y référant explicitement ? Nous pourrions considérer que la mention de ces nouvelles méthodes de traduction automatique, en tant qu'objet d'apprentissage, constitue un argument promotionnel, car les formations qui mettent en avant ces technologies de dernière génération s'efforcent de souligner leur prise en compte des évolutions du secteur et des capacités des futurs traducteurs à s'y adapter.

Plus globalement, les choix stratégiques manifestes dans l'offre de formation appellent une réflexion prégnante (Bowker et Jairo Buitrago, 2019) dans la formation aux métiers de la traduction : dans quelle mesure et sur quelles bases suivre les évolutions de la profession ? L'université gagne-t-elle dans son offre de formation à suivre pas à pas les évolutions du marché et à réinterroger constamment les contenus et les objectifs des cours visant à former les futurs traducteurs ? Ou gagne-t-elle, au contraire, à se donner le temps et le recul nécessaire pour ajuster et éventuellement renouveler ses programmes ?

Afin d'apporter des éléments de réponse à ces questions et de formuler des hypothèses de recherche plus précises, il nous paraît intéressant de prolonger notre travail en menant des entretiens auprès des responsables des offres de formation universitaire FR et IT, pour mieux comprendre la conception des programmes de formation aux technologies de la traduction, les priorités (pédagogiques, institutionnelles), le choix des compétences EMT, et d'une façon plus générale, pour mieux appréhender les choix stratégiques visant à mettre en valeur les points forts des formations en master ; et enfin, pour mieux cerner le rapport complexe entre les champs professionnels et universitaires de la traduction.

BIBLIOGRAPHIE

- BOUTET Josiane, FIALA Pierre et SIMONIN-GRUMBACH Jenny, 1976, « Sociolinguistique ou sociologie du langage ? », *Critique*, n° 344, p. 68-85.
- BOWKER Lynne et BUITRAGO CIRO Jairo, 2019, *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*, Bingley, Emerald Publishing.
- CABRÉ Maria Teresa, 2016, « La Terminologie », in Werner FORNER et Britta THÖRLE (dir.), *Manuel des langues de spécialité*, Berlin/Boston, De Gruyter, p. 68-81.
- CABRÉ Maria Teresa, 2000, « Terminologie et linguistique : la théorie des portes », in Marcel DIKI-KIDIRI, Sophie DAVID, Danièle DUBOIS et Jacques PORTOU (dir.), *Terminologies nouvelles*, n° 21, « Terminologie et diversité culturelle », p. 10-15.
- CENNAMO Ilaria et HAMON Yannick, 2021, « L'offre de formation en traduction : objectifs et compétences dans les contextes universitaires français et italiens », in Tiffane LEVICK et Susan PICKFORD (dir.), *Enseigner la traduction dans les contextes francophones*, Arras, Artois Presses Université, p. 221-237.
- EHRENSBERGER-DOW Maureen et MASSEY Gary, 2019, « Le traducteur et la machine. Mieux travailler ensemble », in Élisabeth LAVALT-OLLÉON et Maria ZIMINA-POIROT (dir.), *Des mots aux actes*, n° 8, *Traduction et technologie : regards croisés sur des nouvelles pratiques*, Paris, Classiques Garnier, p. 47-62.
- GAUDIN François, 2003, *Socioterminologie, une approche sociolinguistique de la terminologie*, Bruxelles, Duculot.
- GAUDIN François, 2005, « La socioterminologie », *Langages*, n° 157, p. 81-93.
- GERARD Christophe, 2018, *La néologie en contexte : cultures, situations, textes*, Limoges, Éditions Lambert-Lucas.
- GROUPE D'EXPERTS EMT, 2017, *Référentiel de compétences de l'EMT*, [https://ec.europa.eu/info/sites/info/files/emt_competence_fwk_2017_fr_web.pdf], consulté le 07-09-2021.
- HUMBLEY John, 2000, « La terminologie », in Gérald ANTOINE et Bernard CERQUIGLINI (dir.), *Histoire de la langue française*, Paris, CNRS Éditions, p. 315-338.
- KENNY Dorothy, 2020, « Technology in Translator Training », in Minako O'HAGAN (dir.), *The Routledge Handbook of Translation and Technology*, Londres/New York, Routledge, p. 498-515.
- LAVALT-OLLÉON Elisabeth et ZIMINA-POIROT Maria (dir.), 2019, *Des mots aux actes*, n° 8, *Traduction et technologie : regards croisés sur des nouvelles pratiques*, Paris, Classiques Garnier.
- LAVALT-OLLÉON Elisabeth, 2019, « Quelle traductologie pour la traduction outillée d'aujourd'hui ? », in Élisabeth LAVALT-OLLÉON et Maria ZIMINA-POIROT (dir.), *Des mots aux actes*, n° 8, *op. cit.*, p. 25-46.
- LEDERER Marianne, 2006, *La Traduction aujourd'hui. Le modèle interprétatif*, Paris, Lettres Modernes Minard.
- LE DISEZ Jean-Yves, 2013, *Le Bon sens en traduction*, Rennes, Presses universitaires de Rennes.
- LOOCK Rudy, 2016, *La Traductologie de corpus*, Villeneuve-d'Ascq, Presses universitaires du Septentrion.
- LOOCK Rudy, 2019, « La plus-value de la biotraduction face à la machine : Le nouveau défi des formations aux métiers de la traduction », *Traduire*, n° 241, p. 54-65, [<https://journals.openedition.org/traduire/18482019>], consulté le 08-09-2021.

- MARTIKAINEN Hanna, 2019, « Qualité fonctionnelle et distorsion d'interprétation », in Élisabeth LAVALT-OLLÉON et Maria ZIMINA-POIROT (dir.), *Des mots aux actes*, n° 8, *op. cit.*, p. 256-270.
- MASSEY Gary et KIRALY Don, 2019, « The future of translator education: A dialogue », *Cultus: The Journal of Intercultural Mediation and Communication*, n° 12, p. 15-34, [https://www.cultusjournal.com/files/Archives/Cultus_2019_12_002_Kiraly_Massey.pdf], consulté le 07-09-2021.
- O'BRIEN Sharon, 2020, « Translation, human-computer interaction and cognition », in Fabio ALVES et Arnt Lykke JAKOBSEN (dir.), *The Routledge Handbook of Translation and Cognition*, Londres, Routledge, p. 376-388.
- O'BRIEN Sharon et ROSSETTI Alessandra, 2021, « Neural machine translation and evolution of the localisation sector, Implications for training », *Journal of Internationalization and Localization*, vol. 7, n° 1/2, Amsterdam/Philadelphia, John Benjamins, p. 95-121.
- ROSSI Caroline, 2019, « L'apprenti traducteur et la machine. Des connaissances aux perceptions de la traduction automatique » in Élisabeth LAVALT-OLLÉON et Maria ZIMINA-POIROT (dir.), *Des mots aux actes*, n° 8, *op. cit.*, p. 93-105.
- WENGER Étienne, 2009, « Communities of practice », *Communities*, vol. 22, n° 5, p. 57-80.
- WENGER Étienne, 1998, *Communities of Practice: Learning, Meaning, and Identity*, Cambridge, Cambridge University Press.
- 2019 *Language Industry Survey – Expectations and Concerns of the European Language Industry*, [<https://euatc.org/wp-content/uploads/2019/11/2019-Language-Industry-Survey-Report.pdf>], consulté le 08-09-2021.

