

Le poste de travail pour traducteur de l'an 2015 – l'exemple de l'environnement traductionnel de la Direction générale de traduction de la Commission européenne

The translators' workstation for 2015:
the example of the CAT tools of the European Commission's
Directorate General for Translation

Anna Walicka

Uniwersytet im. Adama Mickiewicza, Poznań
walicka@amu.edu.pl

Abstract

The aim of this article is to provide an answer to the question about the current state of advancement of computer-assisted translation tools. We assume that several decades of research in the field carried out by the EU institutions in the context of the European integration process have provided the most advanced computer-assisted translation tools available in the biggest translation service in the world, i.e., the Directorate General for Translation of the European Commission. The present work therefore focuses on the following three main types of CAT tools employed by the EU translators: translation memory tools, terminology management tools and machine translation tools. The same types of tools, offered by the EU providers, i.e. SDL and SYSTRAN, are also used by translators working outside the EU structures. We can therefore presume that the EU translation services set work standards which are then accepted by all professional translators. For that reason, in order to define the most probable directions of future development of these tools, this article also reports the current research conducted by the EU in the CAT tools field.

Keywords: computer-assisted translation tools, translation memory, translation memory tools, terminology management tools, machine translation tools.

INTRODUCTION

La création des Communautés européennes avec une augmentation considérable des besoins en traduction de spécialité marque sur le continent européen le début d'une période des recherches très intenses sur des moyens

d'automatisation des tâches du traducteur. À la différence des travaux menés jusqu'en 1966 dans ce domaine aux Etats-Unis, qui visaient une traduction entièrement automatique de haute qualité, l'objectif des recherches menées au sein de l'UE n'était pas de remplacer l'homme par la machine, mais avant tout d'accélérer le travail des traducteurs humains en leur évitant des tâches répétitives.

Quel est le résultat de cette longue période de recherches ? Voici la question à laquelle nous essaierons de répondre dans le présent article. Afin de définir jusqu'à quel point il est possible d'automatiser les tâches du traducteur (et non pas l'opération de la traduction) à l'heure actuelle, nous avons jugé pertinent d'analyser principalement les outils d'aide à la traduction intégrés actuellement au poste de travail du traducteur interne de l'UE, et plus particulièrement de la Direction générale de traduction de la Commission européenne, qui est le plus grand service de traduction au monde avec plus de 1 800 000 pages traduites annuellement¹.

Notre analyse prendra une telle orientation pour trois raisons principales. Premièrement, vu l'intérêt que porte dès sa création l'UE pour l'automatisation des tâches du traducteur et son engagement dans les recherches en la matière, les traducteurs de l'UE disposent sans doute des outils les plus avancés à ce jour.

Deuxièmement, en parlant des composantes principales du poste de travail du traducteur de l'UE, nous parlerons en effet des tendances globales concernant les fonctionnalités principales du poste de travail de la majorité des traducteurs professionnels exerçant leur métier en dehors des institutions de l'UE. Les enquêtes réalisées auprès de ce deuxième groupe au cours de cinq dernières années (Blancafort & Gornostay, 2010 ; Sikora, 2013) ont confirmé notamment que non seulement les fonctionnalités principales du poste de travail sont les mêmes dans les deux groupes de traducteurs, mais également les solutions informatiques spécifiques privilégiées par les professionnels d'en dehors des institutions de l'UE sont celles fournies par les sociétés SDL et SYSTRAN, fournisseurs principaux des services de traduction de l'UE qui, dans le passé, ont fait l'acquisition des systèmes commerciaux offerts par ces sociétés et les ont ensuite développés pour l'usage interne des institutions européennes.

Ces coïncidences résultent à notre avis justement du fait que les traducteurs en dehors de l'UE choisissent également les outils les plus avancés. L'UE joue ainsi un rôle sans pareil dans le façonnement de la configuration informatique

¹ http://bookshop.europa.eu/fr/outils-d-aide-la-traduction-et-cycle-de-travail-pbHC8108467/downloads/HC-81-08-467-FR-C/HC8108467FRC_002.pdf;pgid=y8dIS7GUWMdSR0EAIIMEUUsWb000b33dk-M2;sid=gD8gRBVi794gQkUG-1a-43dHzVw-NhKxhuw=?FileName=HC8108467FRC_002.pdf&SKU=HC8108467FRCPDF&CatalogueNumber=HC-81-08-467-FR-C

des postes de travail de la majorité des professionnels de la traduction, et par conséquent dans le façonnement des modalités informatiques de leur travail. On peut même oser avancer la thèse selon laquelle le poste de travail du traducteur de l'UE ressemble à un baromètre annonçant les futures tendances de portée générale, d'où l'intérêt particulier d'étudier de près son évolution et les tendances actuelles.

Cet intérêt se voit finalement justifié également par le fait qu'il existe relativement peu d'études consacrées à la description du poste de travail du traducteur interne de l'UE. Les publications existantes sont focalisées sur une description (plus ou moins détaillée) d'outils spécifiques utilisés par les traducteurs de l'UE ou d'outils étant l'objet des recherches menées au sein de l'UE ou en coopération avec elle. Pourtant il est difficile d'en apprendre à quel point chacun de ces outils trouve son usage réel dans le travail quotidien des traducteurs humains de l'UE.

Seulement récemment nous avons pu l'observer grâce à une présentation réalisée en mars 2015 à l'Université Adam Mickiewicz par Andrea Voicu, traductrice auprès de la DGT de la Commission européenne. Cette présentation, focalisée sur des étapes successives du processus de traduction en relation avec les outils intégrés au poste de travail du traducteur de l'UE, nous a permis de « mettre de l'ordre » dans les informations déjà disponibles dans la publication la plus détaillée sur ce sujet (à notre savoir), réalisée en 2009 par la DGT².

1. PRÉCIS TERMINOLOGIQUES ET TAXINOMIQUES

Avant d'entrer au cœur de la problématique qui nous intéresse, un précis terminologique et taxinomique s'impose afin d'éclaircir nos usages terminologiques adoptés tout au long de cet article.

Ainsi, le poste de travail (*workbench*) pour traducteur est un terme générique par lequel on désigne un ensemble formé par différents outils de gestion, de documentation et plus particulièrement d'aide à la traduction. Un tel ensemble est aussi appelé *station traductionnelle* ou *environnement TAO*.

Quant au terme *outils d'aide à la traduction*, il se réfère de nos jours à tout outil ou système qui peut être utile dans le travail du traducteur humain, à la différence des *outils de gestion et de documentation* qui sont des outils destinés à l'usage d'un large public, et pas seulement aux langagiers et traducteurs.

² http://bookshop.europa.eu/fr/outils-d-aide-la-traduction-et-cycle-de-travail-pbHC8108467/downloads/HC-81-08-467-FR-C/HC8108467FRC_002.pdf;pgid=y8dIS7GUWMdSR0EAIMEUUsWb0000b33dk-M2;sid=gD8gRBVi794gQkUG-1a-43dHzVw-NhKxhuw=?FileName=HC8108467FRC_002.pdf&SKU=HC8108467FRC_PDF&CatalogueNumber=HC-81-08-467-FR-C

Par conséquent, le terme *outils d'aide à la traduction* s'applique principalement à trois types d'outils, à savoir : mémoires de traduction et systèmes de mémoire de traduction (1), différents outils terminologiques, tels que dictionnaires, glossaires, base de données terminologiques (2), et finalement systèmes de traduction automatique (3) utilisés principalement comme outil de pré-traduction de documents qui nécessitent ensuite une révision humaine. Les systèmes de mémoire de traduction (SMT) et les outils terminologiques sont souvent désignés aussi par le nom *outils de TAO* (traduction assistée par ordinateur).

La TA et la TAO sont deux approches différentes de l'automatisation de la traduction. Ce qui les différencie, c'est surtout la présence ou l'absence d'intervention de l'homme au cours de l'opération même de traduction (Loffler-Laurian, 1996 : 7³ ; Maegaard & Strørup, 1996⁴). Dans la TA, ou dans la traduction automatique assistée par l'homme, l'homme sert la machine (même si les documents traduits automatiquement demandent souvent une pré-édition ou une post-édition), tandis que dans la TAO la machine sert l'homme, qui contrôle l'opération de traduction au cours de laquelle il *dialogue*⁵ avec l'ordinateur. Par la TAO nous désignerons donc, conformément à une tendance générale⁶, tout système qui aide le traducteur dans son travail sans pour autant générer automatiquement une traduction quelconque.

2. POSTE DE TRAVAIL POUR TRADUCTEUR

L'idée de créer une station de travail pour traducteur date des années 80. Elle semble apparaître pour la première fois dans le fameux article de Martin Kay *The Proper Place of Men and Machines in Language Translation* (Kay, 1980/2003). À l'origine, il s'agissait d'automatiser uniquement les tâches que l'ordinateur savait faire de façon satisfaisante. Le traducteur devrait donc travailler de préférence sur un micro-ordinateur comportant des outils de télé-

³ « La TA consiste à saisir un texte, à le soumettre à un traitement automatique et à récupérer en sortie une traduction 'brute'. Aucune intervention humaine n'est possible pendant le traitement. Les programmes ne permettent pas de choix interactif » (Loffler-Laurian, 1996 : 7).

⁴ „W odróżnieniu od MT, gdzie komputer w miarę swych możliwości wykonuje jak największą część zadania, tłumaczenie wspomagane maszynowo to proces, w którym inicjatywę pozostawia się użytkownikowi” (Maegaard & Strørup, 1996 : 37).

⁵ « Les systèmes de traduction assistée par ordinateur sont davantage des stations de travail où le traducteur « dialogue » avec la machine au sujet des propositions de traduction que celle-ci lui indique » (Loffler-Laurian, 1996 : 7).

⁶ C'est une tendance que nous avons observée entre autres dans les travaux de Maegaard & Strørup (1996) ; Loffler-Laurian (1996) ; Hutchins (1998) ; Simard (2003) ; Bogucki (2009).

communication, de consultation de bases de données lexicographiques et terminologiques, de traitement de texte et de traduction automatique spécialisée qui n'apparaîtrait qu'en fin de traduction, seulement à la demande du traducteur. Le poste de travail du traducteur devrait être équipé également d'un éditeur de texte avec l'option de l'écran partagé.

Une idée pareille a été lancée en 1982 par Alan Melby qui a proposé un poste de travail pour traducteur à trois niveaux (*three-level integrated translator's workstation*). Son premier niveau serait constitué par des outils de traitement de texte et des logiciels de consultation de bases de données (*data communication software*), le deuxième, par un logiciel d'analyse textuelle (*automatic text analysis*) et des outils de comparaison terminologique et d'extraction pour les corpus bilingues (*bilingual text retrieval tool*). Au troisième niveau se trouverait l'interface du système de mémoire de traduction (Melby, 1982 et Melby, 1994 cité par Feder, 2000 : 17).

La réalisation de ces idées est devenue possible à partir des années 90, entre autres grâce à l'apparition et la popularisation des ordinateurs personnels équipés en première étape d'outils de traitement de texte et de logiciels de correction orthographique et grammaticale, et – dans des étapes ultérieures – de dictionnaires informatiques et, finalement, de mémoires de traduction.

2.1. MÉMOIRES DE TRADUCTION ET SYSTÈMES DE MÉMOIRE DE TRADUCTION

Les mémoires de traduction ont pris rapidement une place prépondérante parmi les outils d'aide à la traduction. Ainsi, le noyau du poste de travail pour traducteur de l'an 2015 constitue sans contestation possible un système de mémoire de traduction. Ceci concerne aussi bien les traducteurs internes de l'UE que ceux exerçant leur métier en dehors des institutions de l'UE.

Ainsi, le traducticiel⁷ intégré au poste de travail des traducteurs de la DGT est *Translator's Workbench* (TWB) de la société SDL. C'est un système commercial dont la première version pour Windows date de 1994. Cet outil a été sélectionné à la suite d'un appel d'offres interinstitutionnel et ensuite adapté aux besoins spécifiques des institutions européennes⁸. Le logiciel est régulièrement amélioré.

⁷ C'est le terme provenant du document officiel de la DGT disponible sur http://bookshop.europa.eu/fr/outils-d-aide-la-traduction-et-cycle-de-travail-pbHC8108467/downloads/HC-81-08-467-FR-C/HC8108467FRC_002.pdf;pgid=y8dIS7GUWmDSR0EAlMEUUsWb0000b33dk-M2;sid=gD8gRBVi794gQkUG-1a-43dHzVw-NhKxhuw=?FileName=HC8108467FRC_002.pdf&SKU=HC8108467FRC_PDF&CatalogueNumber=HC-81-08-467-FR-C.

⁸ Ibidem.

La Communauté européenne s'est engagée dans les recherches sur les outils d'aide à la traduction, dont la mémoire de traduction, en 1989 avec le projet TWB⁹ (*Translator's Workbench*) réalisé dans le cadre du Programme stratégique européen pour la recherche en technologie de l'information ESPRIT (*European Strategic Program on Research in Information Technology*). Dans le cadre du projet TRANSLEARN initié en 1993 – *Interactive Corpus-Based Translation Drafting Tool*, la CEE a financé les recherches sur un système de mémoire de traduction (Feder, 2000 : 20).

La solution informatique fournie par SDL est aussi celle privilégiée par les traducteurs professionnels en dehors des institutions de l'UE. L'enquête réalisée par Blancafort & Gornostay en 2010 auprès de 139 langagiers de 31 pays (c'est-à-dire traducteurs, terminologues, professeurs de langue), a démontré notamment que 74% d'entre eux se servaient régulièrement des outils informatiques dédiés à la traduction (*translation software*). À l'intérieur de ce groupe, 67% ont déclaré se servir d'outils de la traduction assistée par ordinateur (*CAT tools*), 9% – de la traduction automatique (TA), 24% – aussi bien de la TA que de la TAO. Parmi les outils de la TAO, les plus populaires étaient : SDL Trados (17%), Similis (12%), Transit (10%), Logoport et *Google translation Kit* (6-7%), Wordfast (5%). SYSTRAN était le système commercial de TA le plus populaire (17%).

En ce qui concerne les traducteurs professionnels polonais, l'enquête réalisée auprès de ce groupe en 2011 a confirmé les tendances européennes générales, à savoir : 65% des personnes enquêtées ont déclaré utiliser dans leur travail un SMT et 16% – un système de TA. Encore une fois, SDL Trados s'est avéré le système le plus populaire (35%) parmi tous les SMT (Sikora, 2013 : 22). L'enquête n'a pas identifié malheureusement le système de TA le plus populaire.

La mémoire de traduction (MT) constitue paradoxalement l'outil le plus simple parmi de nombreux outils issus de longues recherches sur la traduction automatisée. Ce dispositif repose notamment sur un archivage systématique de la production d'un ou de plusieurs traducteurs dans une base de données structurée de façon à permettre la réutilisation de cette production. Les segments du texte source (les plus souvent les phrases) sont alignés avec les segments du texte cible qui leur correspondent et des paires ainsi constituées sont ensuite stockées dans une MT qui peut être alors définie de façon simplifiée comme « une archive de traductions existantes, conçue de façon à faciliter la réutilisation des traductions existantes » (Macklovitch & Russell, 2000 : 138). Des mécanismes de recherche adaptés permettent ensuite le repérage automatique rapide d'un segment dans l'une ou l'autre des deux langues, et ainsi la récupération de sa traduction pour des traductions ultérieures (Simard, 2003).

⁹ http://cordis.europa.eu/project/rcn/8628_fr.html

Un environnement permettant au traducteur d'exploiter effectivement cette base de données forme un système de mémoire de traduction qui peut être alors défini comme :

un type d'outil d'aide à la traduction qui maintient une base de données de paires de phrases en langues source et cible, et qui extrait automatiquement la traduction des phrases d'un nouveau texte qui apparaissent dans cette base de données (Macklovitch et Russell, 2000 : 137).

Un SMT « sert » ainsi à retrouver automatiquement dans la MT les passages d'un texte de départ déjà traduits qu'il soumet alors au traducteur pour acceptation ou correction. Lorsqu'une phrase ou un segment de phrase est présente dans la MT sous une forme identique (on appelle un tel cas un *repérage exact*) ou qu'elle ne coïncide que partiellement avec le contenu de la MT (on parle alors des *repérages flous*), celle-ci sera reprise directement de la MT : le traducteur doit seulement valider (ou non) la version proposée¹⁰.

Le mécanisme privilégié de constitution d'une MT dans les systèmes actuels consiste à aligner des segments du texte de départ avec les segments du texte d'arrivée à mesure que le traducteur les produit. L'autre méthode, plus rapide, quoique basée sur la première technique, consiste à importer une ou plusieurs MT déjà existantes, telles que *DGT-TM* de la Commission ou *VLTM (Very Large Translation Memories)* de Wordfast.

Le premier outil a été élaboré par les services de la DGT de la Commission européenne en coopération avec le Centre commun de recherche (JRC), service scientifique interne de la Commission. La première version de cette MT date de 2007. Elle est régulièrement mise à jour et contient aujourd'hui les documents alignés en phrases, publiés par la Commission dans les années 2004 à 2011 (Steinberger et al., 2012).

VLTM de Wordfast est un ensemble de mémoires de traduction élaboré dans le cadre du projet *Wordfast Global Translation Memory*¹¹. Depuis 2006, les personnes consultant le site web de Wordfast sont invitées à contribuer à créer cette MT. *VLTM*, accessible à ceux qui ont acheté un SMT de Wordfast, peut être utilisée actuellement uniquement pour les traductions des textes rédigés en anglais ou traduits vers l'anglais. *DGT-TM*, en libre accès, permet de peupler les MT de différents SMT également pour les traductions entre le français et le polonais.

¹⁰ <http://www.jway.lu/jway/prod/trados-method.html>

¹¹ <http://67-221-227-91.contegix.com/index.php?whichpage=jobs&lang=pl01> (site consulté en mai 2013).

Une tendance générale présente actuellement chez les fournisseurs de logiciels contribue à doter les SMT d'une capacité, appelée *interopérabilité*, qui permet à divers outils de communiquer. Les MT de différents systèmes sont ainsi interchangeable.

2.2. BASES DE DONNÉES TERMINOLOGIQUES

Les outils terminologiques constituent le deuxième type d'outils d'aide à la traduction intégré au poste de travail pour traducteur. Selon Hutchins, l'avènement de la MT a été technologiquement et conceptuellement préparé justement par la popularisation des banques de données terminologiques et glossaires que cet auteur considère comme le premier outil informatique d'aide à la traduction (Hutchins, 1998 : 291).

Un accès rapide à des dictionnaires et glossaires tenus à jour dans les domaines de la science, des technologies ou de l'économie était un besoin incontestable au sein des CECA, CEE et CEEA. Les traducteurs humains avaient besoin d'un accès rapide à des terminologies spécialisées avec leur traduction officiellement confirmée.

Ces besoins pratiques ont fait apparaître dans les années 60 le système DICAUTOM. Elaboré au sein de la CECA, ce système de consultation automatique de dictionnaires permettait de consulter des bases de données contenant les termes des domaines traités par la Communauté, avec leurs différents sens dans différents contextes. Il permettait également de produire des dictionnaires bilingues spécialisés (Hutchins, 1998 : 291). Dans les années 70, DICAUTOM a été remplacé par EURODICAUTOM.

Parallèlement, d'autres institutions de l'UE ont développé leurs propres banques de données terminologiques contenant, à l'instar de DICAUTOM, définitions, exemples et équivalents traductionnels des mots ou des expressions. Elles servaient aussi à générer glossaires et dictionnaires papier pour le vocabulaire de spécialité. Elles étaient généralement multilingues, presque toutes permettaient un accès direct en ligne et présentaient les termes dans leur contexte aussi large que possible. Ainsi, le Parlement de l'UE a créé EUTERPE, le Conseil de l'UE – TIS et la Cour des comptes – CDCTERM. A l'heure actuelle, EURODICAUTOM, EUTERPE, TIS, CDCTERM et Euroterms (base de données du Centre de traduction) font toutes partie de IATE, base de données terminologique multilingue de l'Union européenne (*Inter-Active Terminology for Europe*).

Les travaux sur IATE ont été initiés en 1999, lorsque les institutions de l'UE ont décidé d'élaborer une base de données entièrement neuve en fusionnant le contenu de toutes leurs bases de données terminologiques. IATE a commencé

à fonctionner en 2004 pour l'usage interne des institutions européennes et a été mise à la disposition du grand public en 2007. Elle contient la terminologie et le *jargon* spécifiques de l'UE ainsi que des termes relevant de toutes sortes de domaines tels que droit, agriculture, technologies de l'information et bien d'autres. Selon l'information officielle, publiée sur le site <http://iate.europa.eu>¹², IATE contient 8,4 millions de termes, dont environ 540 000 abréviations et 130 000 expressions, et couvre les 24 langues officielles de l'UE. De nouveaux termes sont ajoutés quotidiennement et son contenu est constamment actualisé.

Les termes sont introduits dans la base par les terminologues et les traducteurs de l'UE sur la base des informations fournies entre autres par les traducteurs, les administrateurs, les juristes-linguistes et les experts. Le site internet de IATE est administré par le Centre de traduction des organes de l'UE à Luxembourg.

2.3. OUTILS DE GESTION TERMINOLOGIQUE

Différents outils terminologiques sont également utilisés par les traducteurs extérieurs aux institutions de l'UE. Le plus souvent, ces outils sont intégrés aux systèmes de mémoires de traduction dont la majorité est organisée autour d'un logiciel de traitement de texte, soit générique (le plus souvent MS-Word), soit dédié. À l'intérieur de ces logiciels de traitement de texte, le traducteur a généralement accès à des dictionnaires et lexiques spécialisés. La majorité des SMT intègre des fonctionnalités de consultation et de gestion terminologique et de lexiques de projet. À titre d'exemple, une option de contrôle de qualité en temps réel permet de vérifier la conformité de la traduction avec les exigences terminologiques du client¹³. Certains SMT disposent de compléments pour automatiser la gestion de dictionnaires. Ainsi, les glossaires de certains systèmes constituent une ressource documentaire permettant d'accélérer le temps de la recherche et de la traduction par une substitution automatique des termes sources par les termes cibles.

L'usage des outils de gestion terminologique se place néanmoins plutôt loin derrière celui des MT et SMT. Il est de 40% environ selon les enquêtes de Blancafort & Gornostay (40%) et de Sikora (41%). Au sein du groupe enquêté en 2010, 13% des personnes enquêtées se servaient de SDL Termbase (pour la création de bases de données terminologiques), 13% – de MultiTerm (pour la création de glossaires) et 7% d'autres outils intégrés aux systèmes de mémoire de traduction (Blancafort & Gornostay, 2010). Seulement 1,4% des personnes

¹² http://iate.europa.eu/brochure/IATEbrochure_FR.pdf (site consulté en septembre 2015).

¹³ <http://www.wordfast.net/index.php?whichpage=downloadpage&lang=engb>

enquêtées ont déclaré chercher la terminologie nécessaire dans différentes banques de données terminologiques accessibles en ligne. Quelque 33,2% recourent à leurs propres bases de données terminologiques, dictionnaires et glossaires, 27,6% aux ressources terminologiques fournies par le client (Blancafort & Gornostay, 2010).

Les habitudes des traducteurs polonais concernant l'usage de banques de données et outils de gestion terminologiques diffèrent de celles des langagiers européens. Premièrement, pour la gestion terminologique de leurs bases de données, les traducteurs polonais choisissent de préférence les systèmes MultiTerm, Lingo (système de gestion terminologique) et Term Watch (système d'analyse des données textuelles) (Sikora, 2013).

En ce qui concerne la recherche terminologique, 50% des traducteurs polonais, contre 1,4% au sein des langagiers européens, ont déclaré chercher une terminologie nécessaire dans des banques de données terminologiques externes, seulement 9% – dans leurs propres bases de données terminologiques. L'enquête ne fournit aucune information sur les ressources terminologiques fournies par le client aux traducteurs polonais. Notre propre pratique professionnelle confirme néanmoins que le client polonais semble plutôt négliger ce besoin du traducteur qu'il laisse tout simplement « se débrouiller » avec la traduction. Dans le meilleur des cas, le traducteur reçoit des traductions anciennes.

2.4. TRADUCTION AUTOMATIQUE

La traduction automatique est l'outil d'aide à la traduction qui a été intégré au poste de travail du traducteur comme le dernier. La vision de Kay et de Melby s'est ainsi accomplie entièrement, du moins en ce qui concerne le poste de travail du traducteur de l'UE.

La première étape d'un long processus de l'intégration de la TA au poste de travail du traducteur de l'UE a constitué l'acquisition en 1975 par les institutions de la Commission européenne de SYSTRAN (*SYStem TRANslation*), système de TA développé aux Etats-Unis à partir de 1964 par Peter Toma. Son acquisition avait pour objectif d'accroître les possibilités multilingues des services de la CE qui voulait utiliser la TA principalement comme une aide aux centres de documentation (Loffler-Laurian, 1996 : 50).

Depuis 1981, ce système sert à traduire les documents internes de l'UE et dès lors, on assiste au développement et à l'installation de nouvelles versions de ce dispositif dans les institutions européennes (Anis, 1994). Depuis 1996, SYSTRAN est utilisé premièrement comme un outil de pré-traduction, permettant de décider si un document devrait être traduit par l'homme, deuxiè-

mement – comme un outil de traduction *brute* des documents courts ou urgents demandant ensuite une post-édition et troisièmement – comme un outil terminologique qui génère des listes de termes basées sur les dictionnaires SYSTRAN et EURODICAUTOM (Maegaard & Strørup, 1996 : 34).

Les recherches sur le couplage des outils d'aide à la traduction avec la TA ont été lancées dans les années 80. L'UE a soutenu ces recherches dans le cadre du projet Eurolang, lancé en 1991, qui a associé les industriels et laboratoires français (SITE, CENT, GETA, LADL), allemands (Siemens), anglais, espagnols, italiens, danois et suédois (Volvo) qui cherchaient à développer une gamme diversifiée de systèmes de TAO à partir des acquis des systèmes de TA tels que METAL, ARIANE et EUROTRA (Anis, 1994). Le projet visait entre autres l'intégration de la TA comme outil de pré-traduction dans des systèmes de TAO.

L'objectif des recherches les plus récentes en la matière, soutenues par l'UE, est de relier les mémoires de traduction à la TA à base de règles. C'est l'objectif du programme « SYSTRAN – *Acquis Communautaire* » qui est un programme de recherche coopératif entre SYSTRAN et P. Koehn de l'Université d'Edimbourg, leader scientifique du logiciel Moses (système *open source* de TA statistique). Le but du projet est de créer la plus grande technologie de traduction automatique, consacrée à la traduction des documents juridiques de l'UE. Pour atteindre cet objectif, des moteurs de traduction hybrides ont été créés par SYSTRAN pour 56 paires de langues en utilisant *l'acquis communautaire*¹⁴.

La technologie hybride permet d'associer la flexibilité des systèmes de TA statistique et la fiabilité de l'approche à base de règles linguistiques¹⁵. Dans les systèmes de TA statistique, toute connaissance linguistique est acquise de façon entièrement automatique à partir de grandes quantités de données issues de corpus de textes bilingues (tels que justement *l'acquis communautaire*). Cette connaissance est stockée dans un modèle de traduction purement statistique, les règles de traduction sont générées automatiquement grâce à l'analyse de corpus bilingues alignés préalablement au niveau phrastique, appelés corpus d'entraînement. Ces corpus ressemblent en effet aux mémoires de traduction en ce qu'ils contiennent des paires de segments de texte étant des segments de textes originaux alignés avec leur traduction. A partir de corpus bilingues ainsi préparés, il est possible d'apprendre les données nécessaires à la traduction automatique sous forme d'un modèle de langue et d'un modèle de traduction qui permettent ensuite ensemble de trouver, pour une phrase source donnée, la phrase cible qui a la plus grande probabilité d'être sa traduction (Brown et al., 1990 cité par Ozdowska, 2006 : 9).

¹⁴ <http://fr.systranlabs.com/projects/acsystran>

¹⁵ <http://fr.systranlabs.com/projects/acsystran>

Une tendance croissante vers une intégration de la TA au poste de travail pour traducteur est visible également en dehors de l'UE. Il en témoignent entre autres les résultats déjà cités des enquêtes réalisées par Blancafort & Gornostay et par Sikora. Rappelons qu'en 2010, 9% des langagiers européens se servaient de la TA, et 24%, aussi bien de la TA que de la TAO. En ce qui concerne les traducteurs polonais, en 2011, 16% d'entre eux ont déclaré se servir de la TA. Ces pourcentages sont relativement importants vu que la qualité de la traduction entièrement automatique ne jouit généralement pas d'une bonne opinion auprès des traducteurs humains, comme en témoigne par exemple le titre de la communication suivante : *De-mystifying Machine Translation to better equip the next generation of professional translators* de Jeff Allen. Nous avons eu l'occasion de l'entendre pendant la conférence organisée à Gdansk en 2014, consacrée aux outils d'aide à la traduction. Du changement des mentalités témoignent également certaines publications polonaises qui présentent des résultats satisfaisants de certains outils de traduction automatique, tels que *Google Translate* (Solecki, 2013 et 2014).

BILAN

De nombreux chercheurs étaient et restent toujours convaincus que l'ensemble des traductions existantes contient plus de solutions à plus de problèmes de traduction que toute autre ressource...¹⁶ (Langlais & Simard, 2001). Il suffit seulement de savoir en profiter. Etant donné que les corpus bilingues et multilingues sont actuellement exploités dans de nombreux outils de la TAO et aussi comme ressource supplémentaire dans les systèmes de TA traditionnels ou dans la traduction automatique basée sur les corpus, la recherche d'une méthode efficace d'exploitation de ces ressources est au cœur des préoccupations des chercheurs.

Nous avons l'impression que les recherches récentes semblent de plus en plus privilégier les systèmes de TA comme outil permettant une exploitation rapide et efficace de ces ressources. Mais, à la différence des recherches anciennes qui visaient la création des systèmes de traduction entièrement automatique de haute qualité, la TA de 2015 est plutôt un outil d'aide à la traduction humaine. Vu le pourcentage des traducteurs professionnels utilisant déjà la TA et les voix qui s'élèvent au sein de la communauté des traducteurs en faveur de ces outils (Solecki, 2013 et 2014), leur importance est susceptible de croître dans les années à venir.

¹⁶ C'est une paraphrase de la citation de Isabelle et al. (1993) : « L'ensemble des nombres réels contient plus de solutions à plus de problèmes d'arithmétique que toute autre ressource ».

De telles transformations nous semblent avoir des implications entre autres pour la formation des traducteurs qui devrait accentuer davantage le développement auprès des apprentis traducteurs de la compétence de post-édition de textes traduits par des systèmes de TA (et pas seulement par l'homme). C'est une compétence actuellement négligée aussi bien dans la pratique que dans la théorie. Ainsi, aucun modèle de compétence en traduction ne l'énumère de manière implicite.

Elle devient pourtant de plus en plus importante, vu également une tendance croissante au partage de différents types de ressources linguistiques et les stratégies des traducteurs qui cherchent la terminologie nécessaire dans des bases de données en ligne. À cela s'ajoute le fait que des glossaires électroniques privés sont déjà un produit vendu par certains traducteurs¹⁷. Par conséquent, évaluer la pertinence de toutes ces ressources et savoir faire une révision rapide des traductions existantes constituera une composante de plus en plus importante de la compétence globale en traduction.

¹⁷ <http://syntax.home.pl/glosariusze.html>

BIBLIOGRAPHIE

- Allen, J. (2014). De-mystifying Machine Translation to better equip the next generation of professional translators. Communication faite à la conférence : „Wpływ nowych narzędzi IT na jakość tłumaczeń oraz perspektywy zawodowe tłumaczy”, Gdańsk 2014.
- Anis, J. (1994). Ordinateurs et traduction : survol d'un demi-siècle. *Langages*, 116, *Le traducteur et l'ordinateur* (Jean-René Ladmirał éd.), 111-121.
- Blancafort, H. & Gornostay, T. (2010). Calling Professionals: Help Us to Understand Your Needs! http://www.researchgate.net/publication/236626809_Calling_Professionals_-_Help_Us_to_Understand_Your_Needs!
- Bogucki, Ł. (2009). *Tłumaczenie wspomagane komputerowo*. Warszawa : Wydawnictwo Naukowe PWN.
- Feder, M. (2000). *Computer Assisted Translation. A Proposal for Tool Evaluation Methodology*, thèse de doctorat non publiée dirigée par A. Pisarska. Université Adam Mickiewicz, Poznań.
- Hutchins, J. (1998). The Origins of the Translator's Workstation. *Machine Translation*, 13, 287-307.
- Isabelle, P. & Warwick-Armstrong S. (1993). Les corpus bilingues : une nouvelle ressource pour le traducteur. In P. Bouillon & A. Clas (Éds.), *La Traductique. Etudes et recherches de traduction par ordinateur*. Montréal : Presses de l'Université de Montréal, 288-306.
- Kay, M. (1980/2003). The Proper Place of Man and Machines in Language Translation. In S. Nirenburg & H. Somers & Y. Wilks (Éds.), *Readings in Machine Translation*. The MIT Press, 221-232.
- Langlais, P. & Simard, M. (2001). Récupération de segments sous-phrastiques dans une mémoire de traduction. *8^e Conférence sur le Traitement Automatique des Langues Naturelles (TALN)*, Tours, France, 243-252.
- Loffler-Laurian, A.-M. (1996). *La traduction automatique*. Paris : Presses Universitaires du Septentrion.
- Maegaard, B. & Strørup, K. (1996). Tłumaczenie maszynowe : tendencje rozwojowe i perspektywy dla użytkowników. In Z. Vetulani, W. Abramowicz & G. Vetulani (Éds.), *Język i technologia*. Warszawa : Akademicka Oficyna Wydawnicza PLJ, 30-38.
- Macklovitch, E. & Russell, G. (2000). What's been Forgotten in Translation Memory. *Proceedings of AMTA-4*, 137-146.
- Ozdowska, S. (2006). *Alibi, un système d'Alignement Bilingue à base de règles de propagation syntaxique*, thèse de doctorat dirigée par M.-P. Péry-Woodley. Université de Toulouse II-Le Mirail.
- Sikora, I. (2013). Polish translators' workstation: On the usage and adoption of computer-assisted translation tools with some implications for translators' training. In T. Piotrowski & Ł. Grabowski (Éds.), *The Translator and the Computer*. Wrocław : Wydawnictwo Wyższej Szkoły Filologicznej we Wrocławiu, 11-34.
- Simard, M. (2003). *Mémoires de traduction sous-phrastiques*, thèse de doctorat dirigée par Jian-Yun Nie. Université de Montréal, 212 p.
- Solecki, B. (2013). Quantitative and qualitative analysis of sample translations produced by Google Translate and human translators. In T. Piotrowski & Ł. Grabowski (Éds.), *The Translator and the Computer*. Wrocław : Wydawnictwo Wyższej Szkoły Filologicznej we Wrocławiu, 205-219.
- Solecki, B. (2014). Suitability of Google Translate. Quantitative and qualitative analysis of sample translations of fragments of various types and registers produced by Google Translate and human translators. Communication faite pendant le colloque : „Tłumacz a komputer 2”, Wrocław 2014.

Steinberger, R. & Eisele, A. & Klocek, S. & Pilos, S. & Schlüter, P. (2012). DGT-TM: A freely Available Translation Memory in 22 Languages. http://langtech.jrc.ec.europa.eu/Documents/2012_LREC_DGT-TM_Final.pdf

Voicu, A. (2015). Typowy dzień tłumacza Komisji Europejskiej. Communication faite le 20 mars 2015 à l'Université Adam Mickiewicz de Poznań.

Sites internet consultés :

http://bookshop.europa.eu/fr/outils-d-aide-la-traduction-et-cycle-de-travail-pbHC8108467/downloads/HC-81-08-467-FR-C/HC8108467FRC_002.pdf;pgid=y8dIS7GUWMdSR0EAIMUUsWb0000b33dk-M2;sid=gD8gRBVi794gQkUG-1a-43dHzVw-NhKxhuw=?FileName=HC8108467FRC_002.pdf&SKU=HC8108467FRC_PDF&CatalogueNumber=HC-81-08-467-FR-C

http://cordis.europa.eu/project/rcn/8628_fr.html

<http://www.jway.lu/jway/prod/trados-method.html>

<http://67-221-227-91.contegix.com/index.php?whichpage=jobs&lang=pl01>

http://iate.europa.eu/brochure/IATEbrochure_FR.pdf

<http://www.wordfast.net/index.php?whichpage=downloadpage&lang=engb>

<http://fr.systranlabs.com/projects/acsystran>

http://www.sft.fr/statistiques-traduction.html#.VjH_qKIY6Uc

<http://syntax.home.pl/glosariusze.html>

