



Traduire

Revue française de la traduction

224 | 2011

Des outils et des nuages

L'informatique, une aide pour le traducteur

Benjamin Phister



Édition électronique

URL : <http://journals.openedition.org/traduire/251>

DOI : 10.4000/traduire.251

ISSN : 2272-9992

Éditeur

Société française des traducteurs

Édition imprimée

Date de publication : 1 juin 2011

Pagination : 30-37

ISSN : 0395-773X

Référence électronique

Benjamin Phister, « L'informatique, une aide pour le traducteur », *Traduire* [En ligne], 224 | 2011, mis en ligne le 03 février 2014, consulté le 19 avril 2019. URL : <http://journals.openedition.org/traduire/251> ; DOI : 10.4000/traduire.251

L'informatique, une aide pour le traducteur

 Benjamin Phister

Le traducteur est un être curieux. Il passe des heures interminables devant un écran d'ordinateur, et a tendance à s'en plaindre. Mais bien souvent il connaît mal les possibilités de celui-ci, et passe à côté d'outils et de techniques qui lui permettraient... de passer moins d'heures devant l'écran.

Cet article offre un tour d'horizon rapide d'importants outils disponibles pour le traducteur. Après un bref rappel historique, il présente brièvement les types d'outils les plus importants et explique leur intérêt pour le traducteur. Pour terminer il évoque quelques tendances intéressantes pour l'avenir.

Le chemin parcouru

La technologie a toujours été un moteur extraordinaire dans l'informatique. Au fil des années, les machines deviennent plus petites et moins chères, plus puissantes et moins gourmandes en ressources. Les logiciels progressent en même temps, profitant des nouvelles possibilités et d'autres performances accrues du matériel. Les éditeurs inventent de nouveaux outils, puis les améliorent sans cesse en ajoutant de nouvelles fonctions et possibilités. La concurrence entre fabricants d'une part, et entre éditeurs d'autre part, fait que l'informatique évolue sans cesse, à une cadence infernale.

Le travail quotidien du traducteur a été révolutionné par l'arrivée de l'informatique. Il n'y a pas si longtemps, le traducteur tapait son texte à la machine à écrire (s'il n'écrivait pas directement à la main, ou dictait). Après l'avoir révisé, il fallait le saisir à nouveau. Changer deux mots dans une phrase pouvait nécessiter de retaper la page entière... Aujourd'hui il est difficile d'imaginer travailler dans ces conditions, sans accès à un outil de traitement de texte.

Depuis l'arrivée des ordinateurs individuels dans les années 1980, le traducteur dispose de sa propre machine pour saisir sa traduction. Il peut chercher les termes dans les dictionnaires

électroniques, ou dans les traductions précédentes stockées sur son disque dur. Il peut utiliser des outils développés spécifiquement pour lui faciliter la tâche.

Une autre technologie a bouleversé encore d'avantage nos méthodes de travail : l'Internet.

Quand j'ai débuté comme traducteur en 1995, bon nombre des documents à traduire m'étaient envoyés par la poste — ou par coursier pour les urgences. Pour effectuer une recherche concernant le sujet couvert par le texte source, je prenais le bus pour aller à la bibliothèque, en espérant y trouver quelques articles ou des livres sur le sujet.

Aujourd'hui nous recevons et envoyons nos textes *via* courriel. Nous consultons des ressources documentaires en ligne, dans le monde entier, comprenant des dizaines voire des centaines d'articles sur des domaines spécialisés. Nous pouvons faire de la recherche terminologique sur des millions de documents, et tester la validité de nos choix en cherchant le nombre d'occurrences du terme choisi, et en vérifiant les sources. Nous pouvons collaborer sur des projets avec des collègues à l'autre extrémité du pays, ou sur un autre continent — partageant textes, glossaires, et mémoires. Il serait bien difficile de travailler efficacement sans l'Internet.



Les outils importants pour le traducteur

L'éventail d'outils spécifiques pour traducteurs est bien trop important pour tout passer en revue dans cet article. Voici une sélection des outils les plus importants à mes yeux.

Traduction Assistée par Ordinateur (TAO)

La Traduction Assistée par Ordinateur a beaucoup évolué depuis ses origines, et aujourd'hui représente un environnement offrant un ensemble d'outils conçus pour optimiser le travail du traducteur. Les composants principaux de cet environnement sont présentés ci-dessous. La plupart sont disponibles dans des suites de programmes tels que SDL Trados Studio 2009, Wordfast, ou Déjà Vu, pour n'en citer que trois (il y en a bien d'autres). Certains composants listés ci-dessous sont vendus séparément (ex : outils de gestion de la terminologie).

Mémoire de traduction

On peut considérer une mémoire de traduction comme une base de données avec un ensemble de phrases traduites. Quand il faut traduire une nouvelle phrase, l'outil recherche dans la base une phrase similaire ou identique, et affiche la traduction précédente, mettant en évidence le(s) mot(s) qui ont changé et remplaçant les chiffres modifiés.

Pour traduire un document, l'outil divise le texte en « segments ». En général un segment est égal à une phrase mais on peut demander une segmentation en fonction de la marque de fin de paragraphe et, dans ce cas, chaque segment contient toutes les phrases d'un paragraphe. Ceci est utile notamment pour les textes rédactionnels ou littéraires, quand le traducteur veut changer l'ordre des phrases entre les textes source et cible.

La fonction de recherche contextuelle est très pratique pour trouver des phrases, termes ou même mots traduits auparavant. Cette recherche est utile même pour le rédactionnel, quand le traducteur ne veut pas exploiter les *fuzzy matches* mais veut simplement retrouver les termes et tournures de phrases dans la mémoire.

Les systèmes de TAO permettent de gérer des formats de fichier contraignants. L'outil s'occupe des « balises » indiquant le formatage (gras, italique, souligné, changement de polices ou de taille de caractères...); le traducteur doit simplement indiquer à quel endroit dans la phrase il faut ajouter la balise.

Alignement

Pour certains projets, le traducteur dispose de textes de référence en langues source et cible (comme par exemple un rapport annuel de l'année précédente, ou un manuel utilisateur à mettre à jour).



S'il veut simplement rechercher des phrases ou des termes dans le texte de référence, le traducteur peut utiliser des outils pour générer un « bitexte » ; il s'agit d'un document avec 2 colonnes, l'une pour le texte source et l'autre pour le texte cible. Chaque phrase du texte source est affichée avec sa traduction dans la colonne en face.

Par contre, si le traducteur veut inclure les textes de référence directement dans sa mémoire de traduction, il doit « aligner » les textes pour indiquer que telle phrase dans le texte source correspond à telle phrase dans le texte cible. La plupart des systèmes de TAO offrent des outils pour faire un alignement manuel. Mais ce travail d'alignement est fastidieux et chronophage. Certains outils ont été développés pour automatiser ce processus, comme AlignFactory Light de Logiterm. Ils proposent des filtres dotés de multiples critères pour contrôler la qualité de l'alignement (p. ex. : rejeter un segment si un des côtés est vide, ou si l'écart de longueur est excessif, etc.).

Gestion de la terminologie

Pratiquement tous les outils de TAO (Transit, Trados, Déjà Vu, Wordfast, MemoQ...) offrent des moyens de gestion de la terminologie. On peut aussi utiliser des outils indépendants tels que Lingo de Lexicool, ou Anylexic de AIT. Avec ces outils le traducteur peut importer des glossaires – fournis par le client, développés par le traducteur, ou trouvés sur le Web – en divers formats (Word, Excel, HTML...). Il peut ajouter de nouveaux termes, et (avec les systèmes plus avancés) inclure diverses informations sur chaque terme : le nom du projet, du client, et du domaine ; le genre ; une définition ; le sigle ou acronyme associé ; une image...

Quand le traducteur commence à traduire un nouveau segment, l'outil cherche les termes à traduire dans la base terminologique ; quand il en trouve, il indique ces termes dans la fenêtre avec le texte source à traduire. Le traducteur peut insérer la traduction du terme par un simple clic ou un raccourci clavier.

On peut créer de multiples bases terminologiques par domaine et/ou par client. Cette technique permet d'appliquer la terminologie propre à chaque client. On peut ouvrir plusieurs dictionnaires en même temps, afin d'étendre la recherche des termes.

Le traducteur peut faire un contrôle terminologique après avoir traduit le texte. Pour chaque terme dans le texte source qui est connu dans la base terminologique, l'outil vérifie que la traduction du terme défini dans le dictionnaire se trouve bien dans le texte cible.

Contrôle qualité

Beaucoup de systèmes de TAO (comme Trados Studio 2009 et Transit) ainsi que des outils indépendants (comme QA Distiller et ErrorSpy) offrent aussi des fonctions de contrôle qualité.

En plus de la vérification orthographique, ces programmes peuvent vérifier que les chiffres sont identiques entre le texte source et cible et qu'ils sont dans le bon format. Ils peuvent contrôler et corriger automatiquement la ponctuation en fin de phrase, les espaces en trop, la casse, les balises, et la terminologie utilisée. Tous ces contrôles automatiques permettent d'améliorer la qualité de la traduction tout en faisant gagner du temps au stade de la relecture.

Recherche dans des corpora

Depuis quelques années, le traducteur peut utiliser l'Internet pour accéder à des corpora. Il s'agit de grands volumes de textes déjà traduits et alignés. Le traducteur définit la langue source et la langue cible, puis saisit un terme ou quelques mots à traduire. Le système (tel que Linguee ou l'outil du Taus Data Association) affiche une table avec les phrases source et cible, en surlignant le terme recherché. Les nouvelles versions des outils de TAO (notamment MemoQ) permettent de lancer des recherches dans les corpora ou les fichiers de référence, directement depuis l'éditeur.

Traduction automatique

Ces systèmes font traduire le texte entièrement par la machine, sans intervention humaine. Certaines solutions (comme Systran) sont sur le marché depuis des décennies ; d'autres (comme Google Translate) sont plus récentes. La traduction machine est souvent méprisée par le traducteur – nous avons tous vu des exemples peu flatteurs des traductions automatiques. Elle n'est pas appropriée pour tout type de texte, notamment pour le rédactionnel mais ces systèmes s'améliorent petit à petit. Pour certains types de textes pas trop compliqués, ils peuvent rivaliser avec le traducteur humain.

Enfin, on constate l'émergence d'une autre évolution à fort potentiel, rendue possible par la combinaison de différentes technologies. Certaines organisations comme Google ou l'association TAUS ont compilé d'énormes mémoires de traduction de bonne qualité, basées sur des traductions humaines. Ces organisations peuvent aussi trouver d'excellentes bases de données terminologiques. Si elles intègrent une grande mémoire, une base terminologique, et une traduction machine, ces organisations peuvent obtenir – pour certains types de texte – des traductions de qualité entièrement automatisées, pour un prix imbattable (voire gratuitement). En effet, Google génère des bénéfices *via* la publicité sur les sites Internet. Il a donc tout intérêt à augmenter le nombre de sites localisés en les traduisant automatiquement.



Tendances nouvelles

TAO : environnement de travail intégré

Les premiers outils de TAO étaient étroitement liés à MS Word (en fait, ils étaient souvent écrits comme une série de macro-commandes sous Word). Cette conception a permis au traducteur de travailler dans l'environnement Word qu'il connaît bien. Mais il y a un inconvénient important : chaque fois que Microsoft offre une nouvelle version de Word, l'outil TAO n'est plus compatible. Il fallait laisser au développeur le temps d'adapter toutes les macros à la nouvelle version. Autre désavantage : l'éditeur était obligé de développer des outils nouveaux pour traduire des fichiers dans d'autres formats tels que PowerPoint, Excel, ou HTML.

Depuis quelques années, quasiment tous les éditeurs ont changé de modèle. Maintenant leurs outils utilisent un seul environnement de travail, utilisable pour tout type de fichier (Word, PPT, HTML...) et intégrant des dictionnaires et autres outils pour gérer la terminologie. Le traducteur « importe » le(s) fichier(s) à traduire dans cet éditeur, traduit le texte, et « exporte » le texte vers le format cible voulu par le client. Ceci permet d'oublier les questions de format et de mise en page pendant la phase de traduction : le traducteur peut se concentrer sur le texte à traduire. Ce modèle ne plaît pas à tous les traducteurs mais il va falloir s'y habituer...

Concordance au niveau du sous-segment

Lorsque l'on a un segment à traduire, les outils de TAO peuvent rechercher dans la mémoire des phrases similaires. La plupart des outils font cette recherche au niveau du segment entier, utilisant des formules mathématiques pour comparer les mots entre le segment source et le segment cible. Maintenant les outils comme Similis, Déjà Vu, et Transit NXT permettent aussi de trouver juste quelques mots dans un segment déjà traduit. Cette technique permet de beaucoup mieux exploiter le contenu de la mémoire.

Traduction automatique

Les récentes versions des outils de TAO (Trados Studio 2009 et Transit NXT, par exemple) permettent au traducteur de demander la traduction automatique de la phrase courante, voire du document entier. Cette intégration de la traduction machine offre un outil de plus à la disposition du traducteur ; à lui de décider s'il veut s'en servir. Il est à parier que ces fonctions seront utilisées de plus en plus, avec les avancées de la traduction machine.

Ces avancées ont donné naissance à un nouveau métier, la post-édition⁽¹⁾, où le traducteur révise une traduction faite entièrement par la machine. Cette approche est utilisée depuis longtemps dans certains secteurs et elle va certainement se développer avec le temps.

(1) Lire à ce sujet l'article d'Anne-Marie Robert, « La post-édition : l'avenir incontournable du traducteur ? », Traduire n° 222 1/10.



Cloud computing

« L'informatique dématérialisée⁽²⁾ » est un terme couvrant un ensemble de technologies permettant d'utiliser des ressources matérielles et logicielles situées à distance. On externalise la gestion et l'administration des programmes et des ordinateurs : l'utilisateur se sert essentiellement de son navigateur Internet pour accéder à ces ressources lointaines. Ceci peut réduire les coûts et le temps nécessaires pour installer, mettre à jour, et maintenir son système informatique.

La boucle est bouclée

Avec l'arrivée du *cloud computing*, l'informatique referme une grande boucle.

Dans les années 1950 et 1960, les ordinateurs centraux (*mainframes*) étaient des monstres occupant des pièces entières et nécessitant une armée de techniciens pour les alimenter, les dresser, les nourrir, et les manipuler. Très chers à fabriquer et à entretenir, ces mastodontes étaient réservés aux grandes entreprises et instituts de recherche. Les ordinateurs centraux puissants exécutaient les programmes et les utilisateurs tapaient leurs commandes à distance *via* des terminaux simples, sans intelligence.

Dans les années 1970, l'arrivée des micro-processeurs – des calculateurs entiers sur un seul circuit – a permis la création des premiers micro-ordinateurs.

Le concept de « l'ordinateur individuel » était tout nouveau : quelques individus en effet ont rêvé qu'un jour chacun pourrait avoir un ordinateur sur son propre bureau. Je ne sais pas s'ils avaient imaginé que quelques décennies plus tard on en aurait aussi dans nos voitures, nos machines à laver, et jusque dans notre poche.

La percée des micro-ordinateurs dans les années 1980 a donc permis à chacun d'avoir un ordinateur sur son bureau et il est rapidement devenu évident qu'il fallait partager des données entre utilisateurs. On a alors mis en place les réseaux locaux, en tirant un câble entre les machines pour qu'elles puissent communiquer entre elles. Puis on a vu qu'il fallait des machines puissantes dédiées au stockage de données et à l'exécution de certains programmes tels que les bases de données : l'architecture « client-serveur » est née.

Malheureusement le concept même de l'ordinateur individuel pose des difficultés aux utilisateurs. Avec la disparition de l'armée des techniciens qui s'occupait des calculateurs centraux, chaque utilisateur est devenu responsable de l'installation de tous les logiciels sur son propre ordinateur ; il doit les configurer, les faire cohabiter et évoluer, et résoudre les problèmes techniques inévitables – d'où une perte de temps considérable.

(2) Lire à ce sujet l'article d'Anne-Marie Robert, « Informatique dématérialisée et logiciels-services : virtualisation de l'environnement de travail du traducteur », en page 14 de ce numéro.



L'Internet offre la possibilité de se connecter sur des serveurs distants et d'y stocker ses données. Avec l'informatique dématérialisée, les applications tournent aussi sur ces serveurs à distance : nous y accédons *via* un simple navigateur qui affiche des informations calculées et stockées sur les serveurs. Donc plus de maintenance en local, de reconfiguration, etc. Il suffit que l'éditeur du logiciel mette à jour le logiciel sur le serveur central et tous les utilisateurs en bénéficient immédiatement.

Ce concept de terminaux non intelligents reliés à un serveur central puissant ressemble étrangement à ce que l'on a connu il y a 50 ans. Plus ça change...

benjamin@phister.fr

*Après des études d'histoire aux États-Unis, **Benjamin PHISTER** s'est converti dans l'informatique et a passé 16 ans dans cette industrie en France et aux USA. Il travaille depuis 1995 comme traducteur et rédacteur technique spécialisé dans l'informatique et les finances. Ancien membre du Comité directeur de la SFT, il est actuellement Président de SFT Services, la filiale focalisée sur les formations et les autres activités lucratives du syndicat.*

