
À quoi sert encore la traduction automatique ?

Thierry Grass

 <https://www.ouvroir.fr/cpe/index.php?id=308>

DOI : 10.57086/cpe.308

Référence électronique

Thierry Grass, « À quoi sert encore la traduction automatique ? », *Cahiers du plurilinguisme européen* [En ligne], 2 | 2010, mis en ligne le 01 janvier 2010, consulté le 06 novembre 2023. URL : <https://www.ouvroir.fr/cpe/index.php?id=308>

Droits d'auteur

Licence Creative Commons – Attribution – Partage dans les mêmes conditions 4.0 International (CC BY-SA 4.0)

À quoi sert encore la traduction automatique ?

Thierry Grass

PLAN

Introduction

Les fonctions de la traduction automatique (TA)

Les treize péchés capitaux de la TA

Premier problème, polysémie et homonymie

Deuxième problème, l'ambiguïté syntaxique

Troisième problème de la TA, l'ambiguïté référentielle

Quatrième problème, les expressions floues (*fuzzy hedges*)

Cinquième problème, idiotismes et métaphores

Sixième problème, la néologie

Septième problème, les noms propres

Huitième problème, les mots d'origine étrangère et les emprunts

Neuvième problème, les séparateurs

Dixième problème, les sigles et les acronymes

Onzième problème, les synonymes

Douzième problème, la transposition

Treizième problème, l'orthographe

Traduction par écrasement et post-édition

Avantages et inconvénients de la traduction par écrasement

Avantages et inconvénients de la technique de post-édition

Conclusion

TEXTE

Introduction

- 1 Lancée pendant la Guerre froide conjointement aux États-Unis et en Union soviétique, la traduction a connu son heure de gloire jusqu'à la publication du rapport ALPAC en 1966 par la *National Science Foundation*, qui concluait à l'impossibilité d'une traduction automatique de qualité. Elle a connu par la suite un renouveau en Europe jusqu'au constat d'échec du projet Eurotra. Désormais, la traduction assistée par ordinateur et l'essor des mémoires de traduction ainsi que des très grandes mémoires de traduction en ligne, alimentées par les tra-

ducteurs clients, a supplanté la traduction automatique même si cette dernière connaît cependant un regain de popularité sur la toile et en Traitement Automatique des Langues grâce au développement du modèle statistique et à la création de l'outil *Google translate*. Est-ce que celle-ci sert encore à quelque chose ? Est-ce qu'elle présente encore un intérêt pour le traducteur compte tenu des « erreurs » qu'elle génère ? Nous allons tenter de répondre à ces questions en opposant les méthodes de la traduction par écrasement telle qu'elle est pratiquée par la majorité des traducteurs et la post-édition d'une ébauche traduite automatiquement.

Les fonctions de la traduction automatique (TA)

- 2 Selon John Hutchins (Hutchins, 2004 : 13-18), quatre fonctions particulières sont assignées à la traduction automatique :
- 3 La première fonction, ou « fonction de dissémination », consiste à produire un brouillon traduit du texte qui devra par la suite être post-édité manuellement pour aboutir à une traduction correcte.
- 4 La deuxième fonction, ou « fonction d'assimilation », consiste à extraire des informations à partir du texte traduit automatiquement, sans considération pour la qualité du texte cible.
- 5 La troisième fonction, ou « fonction d'échange », consiste à utiliser la traduction automatique comme « interprète » de textes électroniques devant être traduits simultanément, comme par exemple les « chats », les pages web ou les courriers électroniques rédigés dans une langue étrangère.
- 6 La quatrième fonction assignée à la traduction automatique est celle d'outil d'accès à l'information en langue étrangère par l'interrogation d'un système de base de données. Notons que l'accès à une base de données par l'intermédiaire d'un logiciel de traduction automatique permet de recueillir des informations non-textuelles, comme des images, mais ce n'est pas ici notre propos.
- 7 En dehors de la fonction de dissémination, et plus rarement de la fonction d'assimilation, le recours au traducteur humain n'est pas prévu pour le genre de tâches assignées à la traduction automatique.

- 8 Il faut donc se poser la question de savoir si une post-édition d'un document traduit automatiquement peut présenter de l'intérêt par rapport à une traduction par écrasement comme elle est pratiquée par la plupart des traducteurs.
- 9 Pour illustrer notre propos, tous les exemples authentiques ou construits recensés dans cette contribution ont été soumis à une traduction automatique à l'aide de la dernière version (payante) du logiciel de traduction automatique *Systran V6 Premium Translator* fonctionnant sur le modèle du transfert ainsi qu'avec l'outil de traduction (gratuit) de Google, *Google translate*¹, fonctionnant sur le modèle statistique développé par Franz Josef Och. (Och & Ney, 2004). Précisons néanmoins que certains services de traductions gratuits sur la toile comme *Yahoo !* ou *Babelfish* utilisent également *Systran* comme moteur de traduction. Pour simplifier, alors que les systèmes par transfert analysent le texte en langue source, en transfèrent les éléments lexico-syntaxiques dans la langue cible pour générer un texte en langue cible sur la base d'un modèle de langue complexe, les systèmes statistiques puisent à l'aide de modèles mathématiques compliqués dans d'immenses corpus parallèles des portions de textes déjà traduits pour les réassembler dans des phrases en langue cible. D'un point de vue purement linguistique, le modèle de langue est beaucoup plus élégant, il est toutefois nettement plus difficile à mettre en œuvre du fait que la langue a tellement d'irrégularités et d'idiosyncrasies, que les formaliser toutes semble illusoire. Le modèle statistique s'affranchit sinon totalement, du moins en grande partie d'une analyse linguistique. Comme pour les mémoires de traduction, il s'agit de piocher des séquences de textes déjà traduites, l'art résidant dans l'assemblage et la construction d'un texte entier, ce que ne fait pas la mémoire de traduction. Le modèle statistique s'est donc presque affranchi du linguiste et demeure la chasse gardée des informaticiens et des mathématiciens. Bien entendu, comme pour les mémoires de traduction, pour que le modèle fonctionne bien, il faut que les corpus soient à la fois nombreux et de bonne qualité.

Les treize péchés capitaux de la TA

10 Dans un article déjà très ancien, Anne-Marie Loffler-Laurian (Loffler-Laurian, 1983 : 65-78) relève douze catégories d'erreurs générées par le système Systran. Plus d'un quart de siècle plus tard, on relève toujours les mêmes erreurs, même si les systèmes de traduction automatique, notamment avec l'arrivée des systèmes statistiques (Brown *et al.*, 1993) et des modèles factorisés (Koehn, 2007) ont fait de notable progrès quant aux paires de langues désormais accessibles. Pour notre part, nous avons relevé treize types d'erreurs générés par les systèmes de traduction automatique.

- Polysémie et homonymie
- Ambiguïté syntaxique
- Ambiguïté référentielle
- Termes flous ou *Fuzzy hedges*
- Idiotismes et métaphores
- Néologie
- Noms propres
- Mots d'origine étrangère et emprunts
- Sigles et acronymes
- Séparateurs
- Synonymes
- Transposition
- Orthographe

11 Passons en revue ces différents problèmes. Nous illustrerons notre propos d'exemples traduits en français, anglais et allemand.

Premier problème, polysémie et homonymie

12 La polysémie constitue le problème le plus souvent signalé en matière de traduction automatique. Les mots ne fonctionnent pas tant comme des unités discrètes, c'est-à-dire bien délimitées et séparées les unes des autres, mais comme des occurrences sujettes à un certain nombre de variations sémantiques au sein d'un champ.

- 13 La polysémie, pluralité de significations au sein d'un continuum sémantique pour un même mot, pose problème également au traducteur humain pour qui il est parfois difficile de trouver la nuance exacte. Tantôt la polysémie se retrouve d'une langue à l'autre (exemple 1), tantôt pas (exemple 2).

14 **Exemple 1 : la traduction de l'adjectif polysémique « libre »**

la voie est libre (non encombrée)

l'entrée est libre (gratuite)

le prisonnier est libre (n'est plus en captivité)

the way is free (not encumbered)

the entry is free (free)

the prisoner is free (is not any more in captivity) [Systran]

der Weg ist frei (nicht überfüllt)

der Eingang ist frei (kostenlos)

das Gefangene ist frei (ist nicht mehr in Gefangenschaft) [Systran]

15 **Exemple 2 : la traduction de l'adjectif polysémique « léger »**

Ce sac à dos est léger. (a peu de poids)

Le directeur est léger dans son travail. (est négligent)

Marie est une femme légère. (aimant le plaisir sexuel)

This backpack is light. [Google]

The director is light in his work. [Google]

À quoi sert encore la traduction automatique ?

Mary is a woman lightly. [Google]

Dieser Rucksack ist leicht. [Google]

Der Direktor ist leicht in seiner Arbeit. [Google]

Marie ist eine Frau leicht. [Google]

16 L'homonymie qui concerne des mots de même graphie et de sens radicalement différents pose également des problèmes :

17 **Exemple 3 : la traduction de l'homonyme « avocat » (homme de loi + fruit)**

Julia aime son avocat. (le fruit est inconcevable)

Julia aime les avocats. (juristes et fruits sont concevables)

Julia loves its lawyer.

Julia loves lawyers.

« lawyer » est la seule variante connue de Systran et de Google.

Julia mag ihren Rechtsanwalt. [Systran]

Julia mag die Rechtsanwälte. [Systran]

Julia mag sein Anwalt. [Google]

Julia liebt die Anwälte. [Google]

18 Le fruit est ici aussi passé sous silence.

19 L'un des problèmes de la traduction automatique est qu'en règle générale, très peu de variantes lexicales sont présentes. C'est un problème de dictionnaire électronique qui peut être résolu et certains

programmes, comme *Reverso Pro*, proposent des variantes de traduction dans la mesure où une unité comporte des homonymes.

Deuxième problème, l'ambiguïté syntaxique

20 L'ambiguïté syntaxique est le second problème couramment évoqué, il apparaît du fait que certaines structures syntaxiques ne sont pas claires sans connaissance du monde :

21 **Exemple 4 : to fly gliders and *to clean fluids**

*Cleaning fluids can be dangerous (cleaning fluids mais pas *to clean fluids)*

Flying gliders can be dangerous (double interpretation : flying gliders et to fly gliders)

**To clean fluids can be dangerous*

To fly gliders can be dangerous

22 En anglais, les verbes « to fly » et « to clean » sont transitifs. Il existe cependant une restriction quant aux arguments ayant la fonction d'objet direct. Ainsi « to fly » demande comme objet un « objet volant ». Bien que moins sélectif quant à la nature de l'objet, « to clean » est incompatible avec « fluids ». L'ambiguïté syntaxique fait appel au contexte argumental et ne peut être résolue que par la prise en compte du niveau lexico-syntaxique (Gross, 1995 :16-19).

Troisième problème de la TA, l'ambiguïté référentielle

23 La question de la référence (rapport entre le texte et la part non linguistique de la pratique où il est produit et interprété) est une question d'interprétation qui suppose par définition une interprétation cognitive.

- 24 Les pronoms réfèrent ainsi à certains mots ou antécédents qui ne sont pas toujours clairs sans connaissance du monde.
- 25 La traduction automatique est effectuée phrase par phrase et les référents peuvent se situer hors d'une phrase particulière.
- 26 **Exemple 5 : le pronom « le » et son référent**

Paul a heurté le vase du pied et l'a cassé. (le vase ou le pied ?)

Paul ran up against the vase of the foot and broke it. [Systran]

Paul struck the foot of the vase and broke. [Google]

Paul ist sich den Schlamm des Fußes gestoßen und es hat gebrochen. [Systran]

Paul schlug die Vorlage des Fußes und brach. [Google]

- 27 L'ambiguïté référentielle fait aussi appel à la connaissance du monde plus que du contexte et dans l'état actuel de la TA, ne peut être résolue de façon satisfaisante.

Quatrième problème, les expressions floues (*fuzzy hedges*)

- 28 Ce sont des mots ou groupes de mots au caractère idiomatique marqué, donc très dépendants de l'organisation sémantique de la langue source, qui sont difficiles à traduire et dont le rôle est d'exprimer une approximation – "words whose job it is to make things more or less fuzzy" selon Lakoff (Lakoff, 1972 : 183). À titre d'exemple on relèvera « en fait », « d'ailleurs », « en un sens » en français, « somehow », « a sort of », « actually » en anglais, « irgendwie » en allemand, etc.
- 29 **Exemple 6 : traduction de l'expression floue « en un sens » / « in a certain way » / « in einem gewissen Sinne »**

Parler n'est-il pas toujours en un sens donner sa parole ?

To speak isn't always in a direction to give its word? [Systran]

Talking is not always in a sense to give his speech? [Google]

Ist sprechen nicht immer in Sinn, sein Wort zu geben? [Systran]

*Sprechen, ist es nicht immer in eine Richtung geben, sein Wort?
[Google]*

- 30 La question de la traduction des termes flous constitue un problème de lexique qui se situe souvent au niveau polylexical, il n'est pas possible de le résoudre sans prendre en compte le contexte adjacent. Ces termes flous posent problème aux systèmes par transfert et statistiques du fait d'un contexte d'apparition très variable.

Cinquième problème, idiotismes et métaphores

- 31 Les idiotismes ou expressions idiomatiques ou encore phrasèmes chez Mel'čuk (Mel'čuk, 1998), ainsi que les métaphores revêtent une coloration culturelle marquée qu'il est difficile de traduire mot à mot.

- 32 **Exemple 7 : traduction de l'expression idiomatique « à couteaux tirés » / « at daggers drawn » / « in äußerster Feindschaft mit jmdm. Leben »**

Nicolas Sarkozy et Jean-François Copé sont désormais à couteaux tirés. (20minutes.fr)

Nicolas Sarkozy and Jean-François Copé are from now on with drawn knives. [Systran]

Nicolas Sarkozy and Jean-François Copé are now at loggerheads. [Google]

Nicolas Sarkozy und Jean-François Copé sind von nun an an gezogenen Messern. [Systran]

Nicolas Sarkozy und Jean-François Copé sind nun Messer aus. [Google]

- 33 La traduction anglaise de Google « to be at loggerheads (être en désaccord) », bien que moins précise, est acceptable.
- 34 Du fait qu'ils fonctionnent à partir de bases de données de textes traduits humainement, les systèmes statistiques de traduction automatique peuvent se révéler plus performants que les systèmes par transfert pour ce genre de problèmes.

Sixième problème, la néologie

- 35 La langue générale et plus encore la terminologie évoluent et les logiciels de traduction automatique n'incluent pas toujours les dernières évolutions lexicales.

36 **Exemple 8 : traduction des néologismes « internautes » et « Web star »**

Ancienne comédienne, Luna Sentz met son talent au service des internautes en animant des émissions interactives en direct sur le site de Canal+. Une Web star est née. (L'Ordinateur Individuel)

Former actress, Luna Sentz puts her talent at the service of the Net surfers by animating interactive emissions on line on the site of Canal+. A Web star was born. [Systran]

Former actress, Luna Sentz puts his talent to the Internet in facilitating interactive programs live on the site of Canal +. Web is a star born. [Google]

Ehemalige Komödiantin, Luna Sentz stellt ihr Talent im Dienst der Internet-Benutzer, indem sie interaktive Direktmissionen auf dem Standort von Canal+ belebt. Ein Web Star ist geboren geworden. [Systran]

Ehemalige Schauspielerin Luna Sentz stellt ihr Talent in den Dienst der Nutzer in Form von interaktiven Sendungen direkt auf der Website von Canal +. Eine Web-Star ist geboren. [Google]

- 37 La néologie suppose une actualisation régulière des dictionnaires électroniques, avec des équipes de lexicographes qui travaillent en arrière-plan pour les systèmes par transfert. Grâce à ses immenses

corpus de textes traduits relatifs aux nouvelles technologies, Google s'en sort ici particulièrement bien.

Septième problème, les noms propres

- 38 Le problème des noms propres est sans doute l'un des plus difficiles à résoudre en traduction automatique : d'une part, leur nombre est tellement élevé qu'un recensement exhaustif paraît pratiquement impossible. En effet, si l'on considère l'ensemble des noms de personnes, des noms de lieux, des noms de marques, d'associations, d'organismes à l'échelle de la planète, on dépasse de loin pour une langue les dictionnaires de langue générale. A la difficulté du recensement des noms propres vient s'ajouter celle de leur orthographe, souvent fluctuante lorsqu'il s'agit de translittération ou de transcription d'une langue à l'alphabet non latin.

39 **Exemple 9 : un nom propre connu, Vladimir Poutine**

Vladimir Poutine a déclaré vouloir équiper sa chienne, un labrador pré-nommé Koni, du système russe de navigation par satellite Glonass pour pouvoir la retrouver si elle s'enfuit ! On comprend dès lors ce que le Kremlin est capable de mettre en oeuvre pour « pister » ses ennemis politiques ! (lepost.fr)

Vladimir Poutine (Putin) stated to want to equip her bitch, fore-mentioned Labrador Koni, Russian system of Glonass navigation by satellite to be able to find it if she flees! One consequently includes/understands what the Kremlin is able to implement "to track" its political enemies! [Systran]

Vladimir Putin has said it wants to equip his dog, a labrador named Koni, the system of Russian Glonass satellite navigation to find if it escapes! It is therefore understandable that the Kremlin is able to implement to "track" their political enemies! [Google]

Vladimir Poutine (Putin) hat erklärt, seine Hündin, ein oben erwähntes Labrador Koni mit dem russischen Navigationssystem via Glonass-Satellite ausstatten zu wollen, um es wiederfinden zu können, wenn sie sich entflöhe! Man begreift folglich, was der Kremlin (Kreml) fähig ist, zu verwirklichen, um seine politischen Feinde „zu verfolgen“! [Systran]

Wladimir Putin will an seinem Hund, einen Labrador namens Koni, des russischen Satellitennavigationssystem GLONASS, um die wiederzufinden, wenn sie fliehen! Es ist daher einleuchtend, dass der Kreml in der Lage ist, um "Spuren" seiner politischen Gegner! [Google]

40 **Avantage ici à Google pour l'anglais, une chienne n'étant pas toujours « a bitch ». Plus sérieusement, les corpus parallèles traduits sont en mesure de donner des équivalences des noms propres et des pré-noms de façon plus fiable dans les systèmes statistiques que par transfert (transcription de « Vladimir » par « Wladimir » en allemand).**

41 **Exemple 10 : un illustre inconnu, le Dr Michel Maure**

Un mandat d'arrêt a été délivré à l'encontre du Dr Michel Maure, 59 ans, auteur de multiples opérations de chirurgie esthétique ratées. (europe1.fr)

A warrant for arrest was delivered against Dr. Michel Moor, 59 years, author of multiple missed operations of cosmetic surgery. [Systran]

An arrest warrant was issued against Dr Michel Maure, 59, author of multiple cosmetic surgery operations failed. [Google]

Ein Erlaßmandat ist gegen maurischen Dr. Michel 59 Jahre Autor mehrfacher verpaßter Operationen ästhetischer Chirurgie ausgestellt worden. [Systran]

Ein Haftbefehl ausgestellt wurde, gegen die von Dr. Michael Mohr, 59 Jahre, Autor von mehreren Operationen der plastischen Chirurgie Fehldrucke. [Google]

42 **La présence de noms propres vient souvent complètement bouleverser la traduction, le système ne parvenant plus à analyser la phrase de manière correcte, traduisant le nom propre en un adjectif (*maurisch* = *mauresque* !) dans la traduction en allemand de l'exemple précédent par Systran. Net avantage à Google.**

Huitième problème, les mots d'origine étrangère et les emprunts

43 Les mots d'origine étrangère sont extrêmement fréquents dans la langue allemande et proviennent généralement de l'anglais ou du français. Ces mots ne sont généralement pas présents dans les dictionnaires électroniques du système, d'où un net avantage aux systèmes statistiques.

44 **Exemple 11 : traduction des emprunts « Debütantin » et « Favoritin »**

Simone Buess : Von der Debütantin zur Favoritin.

Des Simone payer : Du Debütantin au Favoritin. [Systran]

Simone Buess : de la débutante à la favorite. [Google]

45 De plus en plus, on constate parallèlement à la mondialisation, une tendance dans la langue journalistique à la généralisation des emprunts de mots d'origine étrangère ou à la traduction mot à mot de certaines expressions qui prennent aussi racine dans la langue cible. L'expression anglaise « nothing in the pipeline », est devenue en français « rien dans les tuyaux » et en allemand « nichts in der Pipeline ». Bien entendu, la langue d'emprunt est en général l'anglais pour ces expressions, la langue du « business » international.

46 **Exemple 12 : traduction de l'expression idiomatique empruntée de l'anglais « rien dans les tuyaux » et « nothing in the pipeline »**

Most software companies are one product companies, and have nothing in the pipeline apart from upgrades. (techuser.net)

La plupart des fournisseurs de logiciel sont des compagnies d'un produit, et n'ont rien dans la canalisation indépendamment des mises à niveau. [Systran]

La plupart des entreprises de logiciels sont l'un des produits des entreprises, et n'ont rien en dehors de la canalisation de mise à niveau. [Google]

- 47 L'expression idiomatique est traduite dans la langue cible comme s'il s'agissait d'une proposition libre.

Christine Lagarde, fait valoir de son côté qu'« il n'y a, à ce jour, strictement rien dans les tuyaux ». (tradingsat.com)

Christine Lagarde, puts forward on her side that "there is not, to date, strictly nothing in the pipes". [Systran]

Christine Lagarde, argues in turn that "there has, to date, nothing in the pipes." [Google]

Christine Lagarde macht ihrerseits geltend, dass „es bis zum heutigen Tag strikt nichts in den Rohren gibt“. [Systran]

Christine Lagarde, macht geltend, auf seiner Seite: "Es gibt bis heute nicht unbedingt in den Schläuchen. [Google]

- 48 La traduction inverse montre aussi que l'expression idiomatique n'est pas reconnue.

Doch PR-Chef Franz-J. Weihrauch winkt ab: „Da ist aktuell nichts in der Pipeline." (koelner-brauerei-verband.de)

Mais PR-Chef Franz-J. La fumée consacrer décline : « N'est là actuel rien dans la canalisation. » [Systran]

Mais PR-chef Franz-J. À partir de l'encens du doigt : « Il est actuellement rien dans le pipeline. » [Google]

- 49 Certains emprunts de l'anglais jouissent d'une grande popularité, surtout dans la langue des affaires.

- 50 **Exemple 13 : traduction de l'emprunt « business model »**

Les pirates sont innovants, ils mettent en évidence les problèmes du marché et montrent la voie à de nouveaux *business models*. (ecrans.fr)

The pirates are innovating, they highlight the problems of the market and show the way with new businesses models. [Systran]

The pirates are innovative, they highlight the problems of the market and show the way to new business models. [Google]

Die Piraten sind innovativ, sie heben die Probleme des Marktes hervor und zeigen den Weg an einem neuen Geschäft models. [Systran]

Die Piraten sind innovativ, sie deutlich machen, die Probleme des Marktes und weisen den Weg zu neuen Geschäftsmodellen. [Google]

- 51 La traduction en allemand met en évidence que le passage d'un emprunt de l'anglais d'une langue source vers une langue cible toutes deux différentes de la langue d'emprunt ne s'effectue pas sans heurts. Là encore, il s'agit d'un problème de qualité des dictionnaires électroniques qui peut être résolu dans un système par transfert par création d'un dictionnaire des emprunts. Le système statistique se montre un peu plus performant du fait qu'il se fonde sur des corpus de textes traduits humainement.

Neuvième problème, les séparateurs

- 52 Les signes de ponctuation ainsi que certaines abréviations posent problème aux systèmes de traduction automatique. Le fait que le point n'ait pas toujours une fonction de séparateur de phrases constitue un phénomène bien connu en matière de segmentation en français. Le point peut être rencontré en allemand ou en anglais après une abréviation comme Dr. En allemand, le point sert aussi à identifier les nombres ordinaux que l'on rencontre aussi dans les dates là où le français emploie des cardinaux.

- 53 **Exemple 14 : séparateur et adjectif numéral ordinal**

Montag, den 18. August 2008

Lundi le 18 août 2008 [Systran]

Lundi, le 18 Août 2008 [Google]

Darf man ab dem 16. Lebensjahr ein Girokonto eröffnen? (Yahoo Deutschland)

On peut 16 à partir de cela. Ouvrir une année un compte de virement ? [Systran]

Peut-on, à partir du 16 Ans, d'ouvrir un compte courant ? [Google]

- 54 Le point dans ce dernier exemple sert non de séparateur, mais indique qu'il s'agit d'un nombre ordinal qui devrait être traduit en français par « à partir de la 16^e année » pour « ab dem 16. Lebensjahr ».
- 55 Ce problème des séparateurs peut tout à fait être résolu dans la plupart des cas, ce qui implique des modules supplémentaires dans le moteur de traduction.
- 56 Les sigles ne prennent en principe plus de points entre les différentes lettres qui les composent, ce qui constitue une erreur potentielle de moins pour la question des séparateurs.

Dixième problème, les sigles et les acronymes

- 57 Les sigles (épelés), séparés ou non par des points, ainsi que les acronymes (prononcés comme une unité phonique et donc sans points séparant les différentes lettres) sont couramment employés dans les textes journalistiques. Certains se traduisent, d'autres pas. Certaines langues, comme l'allemand, utilisent aussi les acronymes anglais. Signalons toutefois que les acronymes s'écrivent le plus souvent en capitales et sans points abrégatifs : UNESCO, ONU, OTAN, NASA. Parfois aussi, on les écrit aussi comme des noms propres, avec une majuscule initiale : *Onu*, *Insee*. Les sigles perdent plus difficilement leurs points étant donné qu'il se prononcent lettre après lettre : S.N.C.F.
- 58 **Exemple 15 : traduction du sigle d'une organisation internationale connue**

L'Organisation mondiale du commerce (OMC) est la seule organisation internationale qui s'occupe des règles régissant le commerce entre les pays. (wto.org)

The World Trade Organization (WTO) is the only international organization who deals with the rules governing the trade between the countries. [Systran]

The World Trade Organization (WTO) is the only international organization dealing with the rules governing trade between countries. [Google]

Das Welthandelsorganisation (WHO) ist die einzige internationale Organisation, die sich mit den Regeln befasst, die den Handel zwischen den Ländern leiten. [Systran]

Die Welthandelsorganisation (WTO) ist die einzige internationale Organisation, die sich mit den Regeln für den Handel zwischen den Ländern. [Google]

59 **Système par transfert et système statistique viennent à bout de ce genre de problèmes. Un système comme Google a naturellement à sa disposition les pages traduites desdites organisations, d'où la qualité de la traduction automatique réalisée sur la base de corpus parallèles.**

60 **Exemple 16 : traduction d'une abréviation connue**

L'entreprise a repris la dénomination Total SA le 6 mai 2003. (Wikipédia)

The company took again the denomination Total SA on May 6th, 2003. [Systran]

The company has taken the name Total SA on 6 May 2003. [Google]

Das Unternehmen hat die Bezeichnung Gesamtzahl AG am 6. Mai 2003 wieder aufgenommen. [Systran]

Das Unternehmen hat die Bezeichnung Total SA am 6. Mai 2003.
[Google]

- 61 Une SA (société anonyme) de droit français n'est évidemment pas une AG (Aktiengesellschaft = société par actions) de droit allemand. Le système par transfert pêche ici par une table de correspondances inexactes et un dictionnaire des noms propres d'entreprises insuffisant.

Onzième problème, les synonymes

- 62 La question de la synonymie est l'une des plus cruciales en traduction car elle traduit la richesse lexicale d'une langue et la compétence d'un traducteur. De nombreux mots ne se différencient les uns des autres que par des différences, parfois infimes, mais nécessaires pour reproduire telle ou telle nuance de style ou de sens dans tel ou tel contexte. Un logiciel de traduction ne dispose généralement que d'un nombre limité de variantes pour traduire telle ou telle unité. La traduction peut ainsi apparaître compréhensible, mais peu élégante, voire maladroite.
- 63 **Exemple 16 : traduction de « banner » par « bannière » au lieu de « banderole »**

Two British Free Tibet campaigners are in custody in China after unfurling a Tibetan flag and banner outside the Olympic stadium. (free-tibet.net)

Deux militants libres britanniques du Thibet sont dans la garde en Chine après unfurling un drapeau et une bannière tibétains en dehors du stade olympique. [Systran]

Deux British Free Tibet militants sont en garde à vue après le déploiement en Chine, le drapeau tibétain et la bannière à l'extérieur du stade olympique. [Google]

- 64 Quasi-synonymes, « bannière » = « étendard d'une confrérie, d'une société » ne s'en distingue pas moins de « banderole » = « grande bande de tissu qui porte une inscription (en signe de protestation) ».

Douzième problème, la transposition

65 Au sens classique, la transposition en traduction consiste à traduire une unité lexicale d'une classe (nom, verbe, adjectif, adverbe) par une unité lexicale d'une autre classe. La transposition est assez fréquente lorsqu'on traduit des langues romanes vers les langues germaniques, les premières ayant souvent recours à des nominalisations ou les secondes préféreront des expressions verbales.

66 **Exemple 17 : transposition de « house for sale » en « maison à vendre »**

*Detroit has a bunch of run down houses for sale in the \$30000 range.
(www.productivity501.com/help-me-understand-the-subprime-mess/419/)*

Detroit a un groupe de maisons de course vers le bas à vendre dans la gamme \$30000. [Systran]

Detroit a un tas de courir les maisons en vente dans la gamme \$30,000. [Google]

67 L'expression « house for sale » confine à l'idiotisme et les deux systèmes ont procédé à la dislocation de l'expression qui devient incompréhensible. Mais le problème consistait à ne pas traduire « for sale » par une suite « préposition + nom » mais « préposition + verbe ». C'est une gageure pratiquement impossible à résoudre pour un système par transfert et là encore, même s'il ne brille pas, le système statistique se révèle meilleur. Paradoxalement, les systèmes s'en sortent mieux à la transposition de « maisons à vendre » du français vers l'anglais.

68 **Exemple 18 : transposition de « maison à vendre » en « house for sale »**

Dans tout le pays on organise des foreclosure tours, visites organisées de maisons à vendre.

In all the country one organizes foreclosure turns, visits organized of houses for sale. [Systran]

Across the country on organizes foreclosure tours, tours of homes for sale. [Google]

Treizième problème, l'orthographe

69 L'orthographe, lorsqu'elle est défectueuse, est un ennemi de la traduction automatique, c'est une remarque triviale. Les systèmes de traduction travaillant sur du texte analysent tout mot mal orthographié comme mot inconnu et tout mot inconnu ne peut avoir de traduction et est donc laissé tel quel. Un mot inconnu ne peut être analysé, c'est-à-dire rattaché à une classe de mots comme les noms, les verbes, etc. De ce fait, un mot inconnu provoque systématiquement des erreurs d'analyse dans un système par transfert. L'étape qui doit précéder toute traduction automatique consiste donc en une pré-édition dont le but minimal sera de corriger les fautes d'orthographe.

70 **Exemple 19 : traduction d'une phrase mal orthographiée**

Alors qu'un traducteur humain est capable de traduire une phrase mâle orthographiée, il en va tout ôtrement d'un logiciel.

Whereas a human traducttor is cpable to translate a spelled male sentence, it goes from there all ôtrement a software. [Systran]

While a traductteur human cpable to translate a sentence spelled male, that's a ôtrement of software. [Google]

Während ein menschliches traductteur cpable ist, einen orthographisch richtig geschriebenen männlichen Satz zu übersetzen, geht er davon jedes ôtrement einer Software. [Systran]

Während ein traductteur Menschen cpable übersetzen Mann einen Satz geschrieben, das ganze ôtrement einer Software. [Google]

71 No comment...

72 Si certains problèmes relatifs à la TA (ambiguïté, transposition) sont imputables à la question de la « connaissance du monde », la plupart d'entre eux proviennent d'un codage insuffisant des dictionnaires ne

prenant pas encore en compte toutes les avancées de la linguistique dans les systèmes par transfert. L'avenir semble donc appartenir aux systèmes statistiques disposant de corpus parallèles les plus larges possibles. Quoiqu'il en soit, la traduction automatique connaît, avec de tels systèmes dérivés des mémoires de traduction, un renouveau dont il est impossible de ne pas tenir compte. Ils remettent aussi au goût du jour la technique de la post-édition, c'est-à-dire la correction d'une ébauche traduite automatiquement.

Traduction par écrasement et post-édition

73 Pour pouvoir travailler de façon productive avec un logiciel de traduction automatique, il est nécessaire de connaître ses « erreurs ». La deuxième partie de cet exposé a donc pu paraître comme une critique en règle et quelque peu convenue de la TA. Aussi entreprendrons nous, dans cette troisième partie, de rétablir un peu l'équilibre en faveur de celle-ci en opposant post-édition d'un document traduit automatiquement et traduction par écrasement.

Avantages et inconvénients de la traduction par écrasement

74 La traduction par écrasement d'un fichier source est la méthode employée par une majorité de traducteurs dans leur travail et elle est couramment enseignée comme « la » méthode dans les écoles de traduction. Cette méthode comporte un certain nombre d'avantages mais aussi certains inconvénients.

75 Au crédit de la traduction par écrasement, on relèvera les avantages suivants :

- Elle permet de conserver le formatage du texte source.
- Le texte est traduit dans son intégralité en minimisant le risque d'oublis.
- Associée à une mémoire de traduction, la méthode est rapide pour des textes répétitifs.
- La qualité de la traduction est celle d'une traduction humaine.

76 Mais les inconvénients sont les suivants :

- La traduction doit être saisie au clavier dans son intégralité.
- La qualité de la traduction peut être sujette à des fluctuations en fonction du niveau de compétence du traducteur ou en cas de changement de traducteur sur un même texte.

Avantages et inconvénients de la technique de post-édition

77 La post-édition d'une ébauche traduite automatiquement consiste en une complète correction d'un texte. Elle demande une attention soutenue ainsi qu'une certaine familiarisation avec les logiciels de TA pour connaître les fautes qui sont couramment commises. Peu enseignée dans les écoles de traduction, elle est pourtant assez couramment pratiquée dans l'industrie. Là encore, citons les avantages et les inconvénients les plus marquants.

78 Au crédit de la technique de post-édition, on relèvera :

- Elle permet aussi de conserver le formatage du texte source pour les programmes dédiés comme *Systran* (ce que ne permet pas *Google*).
- Elle permet d'avoir une vue d'ensemble du texte dans la langue cible et de le défricher plus rapidement.
- Certains la considèrent comme plus rapide que la post-édition par écrasement du fait de la présence d'une ébauche traduite préexistante.
- Elle autorise une grande consistance terminologique si les dictionnaires utilisateurs sont bien construits pour les programmes dédiés comme *Systran* (une fonction qui, là encore, n'existe pas dans *Google*).
- On constate une régularité dans la traduction du fait que tout syntagme identique sera toujours traduit de la même façon.
- Elle se prête tout particulièrement à la traduction technique sous réserve de dictionnaires bien élaborés à l'aide de programmes dédiés.
- Elle facilite la circulation du texte d'un traducteur à l'autre du fait de sa tendance à linéariser le style.

79 Les inconvénients sont les suivants :

- Certains traducteurs la considèrent comme plus lente que la méthode par écrasement en arguant qu'il est plus lent, voire plus fastidieux, de corriger que de rédiger.

- La post-édition tend à effacer la « patte » personnelle du traducteur, le style apparaît uniforme, voire monotone.
- La standardisation du vocabulaire tend à se substituer à la richesse lexicale.
- La post-édition n'est pas adaptée à certains types de textes à vocation esthétique, ce n'est pas un outil à utiliser en traduction littéraire ou dans toute tâche où une analyse philologique du texte est nécessaire.
- Il est plus facile d'omettre de traduire certains passages qui resteront sous forme d'ébauche.

Conclusion

- 80 La traduction automatique, n'en déplaise aux professionnels de la traduction, n'est pas morte et connaît même avec les systèmes statistiques de traduction automatique un regain de popularité. Si les mal-adresses d'autrefois prêtaient, et pour certaines, prêtent toujours à sourire, son utilité est indéniable dans certaines paires de langues, et en particulier à partir de et vers l'anglais. En revanche, certaines autres paires de langues posent toujours un plus grand nombre de problèmes, ce constat ne s'appliquant bien entendu pas aux seuls systèmes évoqués ici.
- 81 Comme le souligne Christian Boitet (Boitet, 2007 : 25), « les corpus utilisés en TA de l'écrit et de l'oral ont évolué, depuis les suites de test et les corpus d'essai des débuts, vers des corpus parallèles bilingues ou multilingues, bruts ou enrichis par des métadonnées et une grande variété d'annotations linguistiques. » Les systèmes statistiques reposant sur de tels corpus, comme *Google translate*, du fait qu'ils s'appuient sur des traductions humaines réalisées en amont, peuvent constituer une aide précieuse et tendent par leurs performances à dépasser les systèmes par transfert. Mais il faut toutefois relativiser et quelques exemples ne suffisent pas pour juger d'un système.
- 82 Du fait de l'évolution des systèmes, la réalisation d'une ébauche automatique apparaît de moins en moins absurde et la familiarisation avec la technique de post-édition devrait faire partie de tout cursus de traduction, même à titre accessoire, dans la mesure où elle accentue le recul critique vis-à-vis du texte cible et exerce l'acuité du réviseur.

- 83 Ce texte a été rédigé durant l'été 2008, depuis, une nouvelle version commercialisée en juin 2009, *Systran Enterprise Server 7*, a adopté en partie l'approche statistique grâce à son moteur de traduction hybride entraîné à partir de l'*Acquis Communautaire*, le corpus multilingue de la Commission européenne et réalisé en collaboration avec Philipp Koehn de l'Université d'Edimbourg.

BIBLIOGRAPHIE

- Boitet Christian, 2007, « Corpus pour la TA : types, tailles et problèmes associés, selon leur usage et le type de système », dans *Revue française de linguistique appliquée*, Vol. XII, 2007/1, p. 25-38.
- Brown Peter F., Della Pietra Stephen, Della Pietra Vincent J., Jelinek Frederick, Lafferty John D., Mercer Robert L., Roossin Paul S., 1990, « A Statistical Approach to Machine Translation », dans *Computational Linguistics*, vol. 16/2, p. 79-85.
- Gross Gaston, 1995, « Une sémantique nouvelle pour la traduction automatique : les classes d'objets », dans *La Tribune des Industries de la Langue et de l'Information électronique*, n° 17-18-19, p. 16-19.
- Hutchins John, 2004, « Machine translation and computer-based translation tools. A new spectrum of translation studies », dans Bravo José Maria (ed), *Publicaciones de la universidad de Valladolid*, p. 13-48.
- Koehn Philipp, Hoang Hieu, 2007, « Factored Translation Models », dans *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Prague, p. 868-876.
- Lakoff George, 1972, « Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts », dans Perantean P. M., Levi J. N., and Phares G. C. (ed.), *Papers from the 8th Regional Meeting*, Chicago Linguistics Society, p. 183-228.
- Loffler-Laurian Anne-Marie, 1983, « Pour une typologie des erreurs dans la traduction automatique », dans *Multilingua*, vol. 2, n° 2, p. 65-78.
- Mel'čuk Igor, 1998, « Collocations and Lexical Functions », dans Cowie Anthony P. (ed.), *Phraseology : Theory, Analysis and Applications*, (Oxford Studies in Lexicographie and Lexicology), Oxford, Oxford University Press, p. 79-100.
- Och Franz Josef, Ney Hermann, 2004, « The Alignment Template Approach to Statistical Machine Translation », dans *Computational Linguistics*, vol. 30, n° 4, p. 417-449.

NOTES

1 Accessible sous : <https://translate.google.com/>

RÉSUMÉS

Français

Lancée pendant la Guerre froide conjointement aux États-Unis et en Union soviétique, la traduction automatique a connu son heure de gloire avant d'être supplantée par l'essor des mémoires de traduction. Délaissée par les traducteurs, la traduction automatique sert-elle encore à quelque chose ? Est-ce qu'elle présente encore un intérêt compte tenu des « erreurs » qu'elle génère ? Après avoir recensé les problèmes posés par la traduction automatique, cette contribution tentera de présenter les avantages et inconvénients de la post-édition d'un texte traduit automatiquement par rapport à la traduction par écrasement.

English

The glory days of machine translation, which was developed during the Cold War period in the US and in the Soviet Union, came to an end with the rise of translation memories. Currently machine translation is ignored by translators, and the question may be asked whether it can still be of use. Is it still relevant in light of the "errors" it produces? After identifying the problems of machine translation, this paper will present the advantages and disadvantages of post-editing an automatically translated text, and will compare these findings to a direct translation using translation memories.

INDEX

Mots-clés

erreur de traduction, post-édition, TAL, traduction automatique, franco-allemand

Keywords

machine translation, NLP, post-editing, translation error, French-German

AUTEUR

À quoi sert encore la traduction automatique ?

Thierry Grass

Après s'être spécialisé dans la traduction juridique (allemand-français), Thierry Grass a opté pour une carrière universitaire durant laquelle il a acquis une spécialisation en linguistique informatique. Il a été responsable de cursus de traduction juridique et d'informatique appliquée à la langue et à la traduction aux universités d'Augsbourg et de Fribourg en Allemagne ainsi que de Tours en France. Thierry Grass est désormais professeur à l'UFR de langues et sciences humaines appliquées de l'université de Strasbourg. Responsable du master « Traduction et interprétation », il assure des enseignements de traduction, de terminologie et de traitement automatique des langues.

IDREF : <https://www.idref.fr/053432983>

HAL : <https://cv.archives-ouvertes.fr/grass-thierry>

ISNI : <http://www.isni.org/0000000037768874>

BNF : <https://data.bnf.fr/fr/13594466>