

ANALYSE COMPAREE DE 4 TRADUCTIONS PRODUITES PAR DES LOGICIELS DE TRADUCTION AUTOMATIQUE, SELON LA GRILLE DE BEAUDRANDE ET DRESSLER

Christian Papas

COMPUTER TRANSLATION AND TEXTUAL COMPETENCE

SUMMARY

A survey of 4 computer translations of same excerpt from Newsweek, which appeared in Le Monde (13.1.1999). Evaluation of their communicative value through the 7 standards developed by Dressler and Beaugrande:

- 2 text-centred notions: coherence and cohesion
- 5 user-centred notions: intentionality, acceptability, informativity, situationality and intertextuality.

In case of any of these standards is unfulfilled, the text is not communicative. Differences between these computer translations and a human translation. Remarks on the relative characteristics of the two different modes. Conclusion: an intelligent computer has yet to be devised.

ΜΗΧΑΝΙΚΗ ΜΕΤΑΦΡΑΣΗ ΚΑΙ ΚΕΙΜΕΝΙΚΕΣ ΔΥΝΑΤΟΤΗΤΕΣ

Christian Παπάς

ΠΕΡΙΛΗΨΗ

Στην παρούσα ανακοίνωση προβαίνουμε στην επισκόπηση τεσσάρων μεταφράσεων από αυτόματο μεταφραστή ενός κειμένου από το περιοδικό Newsweek. Οι μεταφράσεις αυτές, προερχόμενες από τέσσερα διαφορετικά συστήματα αυτόματης μετάφρασης (Reverse, Systran, Systran professionnel και Power Translation Pro) δημοσιεύτηκαν στην εφημερίδα Le Monde (13. 1. 1999), κρίνονται και αξιολογούνται ως προς την επικοινωνιακή τους αξία με γνώμονα τους επτά κειμενικούς παράγοντες των Dressler και Beaugrande: συνοχή, συνεκτικότητα, προθετικότητα, αποδεκτότητα, πληροφορητικότητα, καταστασιακότητα και διακειμενικότητα. Στην περίπτωση που κάποιος από τους παράγοντες αυτούς δεν πληρούται το κείμενο θεωρείται μη επικοινωνιακό.

Διαφορές μεταξύ ανθρώπινης και μηχανικής μετάφρασης, ιδιαίτερα χαρακτηριστικά των δύο αυτών προσεγγίσεων; συμπέρασμα: ο έξυπνος υπολογιστής είναι ακόμη μακριά!

INTRODUCTION

Mon attention a ete attirée par un article du Supplement Interactif du journal Le Monde du 13.1.1999 intitulé Quatre logiciels de traduction automatique au banc d'essai. Le jour-

naliste du Monde avait donné à traduire à ces 4 logiciels un même passage tiré d'un numéro spécial de Newsweek consacré aux inventions qui ont marqué ces deux millénaires. Le passage à traduire contenait le titre de l'article *The Power of Big Ideas* et le chapeau c'est-à-dire le résumé de l'article.

En matière de traduction, on trouve aujourd'hui deux écoles:

- ceux qui la considèrent comme un art: les traducteurs et particulièrement les traducteurs littéraires et
- ceux qui la considèrent comme un produit: les créateurs de logiciels de Traduction Automatique et leurs clients (les organismes internationaux: ONU, UNESCO, UE etc., et les grandes entreprises nationales).

Le Monde, qui est nul doute partisan de la première tendance, laisse ses lecteurs juger ces 4 TA et ironise: ou est l'éloquence?

Comme l'écrit Anne-Marie Loffler-Laurian[i], directeur de recherche au CNRS: Si la traduction automatique fonctionnait parfaitement, il n'y aurait rien à en dire.

Je me suis amusé à passer ces 4 TA au crible des 7 critères de textualité que Dressler et Beaugrande ont établi pour juger de la compréhensibilité d'un message - critères réunis dans leur ouvrage *Introduction to Text Linguistics* paru en 1981 et plusieurs fois réédité depuis[ii]. Ils ont défini 7 critères de textualité. Deux critères essentiels, la cohérence et la cohésion qu'exige la logique interne, et cinq critères d'interaction communicative entre l'émetteur et le destinataire du texte, critères d'intentionnalité, d'acceptabilité, d'informativité, de situationalité et d'intertextualité.

Ces sept critères ont été mis à contribution pour analyser à la fois le texte en langue-source et les traductions élaborées par ces quatre logiciels de traduction automatique. Nous essaierons de démontrer - ce qui peut sembler intuitivement une évidence - que les défauts de ces traductions viennent du non respect de certains critères de textualité et nous tenterons d'élargir le débat en proposant une traduction humaine et en en tirant les conclusions qui en découlent.

Les 4 traductions automatiques passées au crible des 7 critères de textualité de Beaugrande et Dressler

A. LES CRITÈRES DE LOGIQUE INTERNE: COHÉRENCE ET COHÉSION

Ce sont les critères les plus importants et les plus utiles au traducteur, au point que Christine Durieux[iii] proposait comme épreuve de sélection à l'entrée de l'ESIT (École Supérieure d'Interprètes et de Traducteurs de l'Université de la Sorbonne nouvelle - Paris III) des tests de raisonnement logique et non plus une épreuve de traduction car celle-ci ne permet de juger qu'une performance et non une aptitude.

Cette aptitude logique garantit la perception de l'homogeneite d'un texte au niveau macro et microstructurel.

a. La coherence

Elle correspond a la macro-organisation d'un texte, a son decoupage en unites de contenu, c'est-a-dire en themes et sous-themes. Elle s'appuie essentiellement sur la continuite semantique. La notion de continuite semantique est indispensable au traducteur, en particulier pour degager le sens des mots polysemiques.

- Dans TA1, la deuxieme proposition de traduction (celle entre parentheses) de "gathering" par "cueillette" n'est pas pertinente. Si "gathering" est polysemique, "cueillette" ne l'est pas et appartient uniquement au champ semantique de l'horticulture (ex. La cueillette des fruits).
- Dans toutes les TA "the light bulb" est traduit par "l'ampoule". Or "l'ampoule" fait partie des champs semantiques de l'electricite et de la pharmacie. Dans une phrase ou apparait egalement le mot "pilule", si on ne renforce pas "ampoule" par l'adjectif "electrique" il plane une certaine ambiguite qui risque de desarconner le lecteur qui ne connatrait pas la langue-source ou n'aurait pas le texte anglais sous les yeux.

B. LA COHESION

Elle concerne la micro-organisation du texte en prenant en compte les relations interphrastiques grace a des connecteurs aussi appeles liens cohesifs. Mais les connecteurs, qu'ils soient des marqueurs spatiaux, temporels, argumentatifs ou explicatifs ne sont pas une contrainte forte. D'autres procedes peuvent assurer l'effet textuel comme l'organisation des contenus, la progression thematique.

L'extrait a traduire est trop bref pour qu'on y distingue des connecteurs, qui seraient d'ailleurs contraires a l'intention du journaliste qui utilise des phrases deconnectees pour engendrer la curiosite. Mais:

-TA1 rompt l'organisation et la progression du texte en traduisant "might" par "pouvaient". La traduction de "might" pose souvent un probleme. Il exprime une probabilite, donc un futur, un peu plus aleatoire que "may". Les TA2 et TA3 du systeme Systran, et la TA3, qui traduisent "might" par "pourraient" et "peuvent" sont toutes aussi pertinentes car la difference de sens est mince entre "pourraient" et "peuvent". Ce qui rend TA1 inacceptable est la rupture de la progression des 3 phrases: la premiere est au passe, la seconde au present, la logique veut que la troisieme dans ce contexte soit au futur et non pas au passe. Les tiroirs verbaux du conditionnel et du present des TA2, TA3 et TA4 ont tous les deux un sens futur et sont donc pertinents.

Ces criteres de logique interne etant evalues, il reste les criteres de communication: intention, acceptabilite, informativite, situation et intertextualite.

C. LES CRITERES DE COMMUNICATION

Ils mettent en jeu la presence implicite de l'emetteur et du destinataire

a. L'intentionalite

L'intentionalite laisse entrevoir la volonte du journaliste. Les regles qui definissent la redaction du chapeau d'un article sont assez bien definies. Il s'agit de ce qu'on appelle un "teaser" en jargon franglais journalistique, c'est-a-dire un court texte qui interpelle le lecteur, qui l'intrigue en lui laissant deviner le corps du sujet et lui donne envie de le lire, de verifier ses hypotheses en l'incitant a pousser plus loin sa lecture. C'est la raison pour laquelle la cohesion n'est pas tres forte et que stylistiquement le texte commence par une interrogation et finit par une interpellation "might surprise you".

Les 4 TA se bornent a traduire des mots et des groupes de mots et les logiciels n'ont pas la capacite de transposer ces elements de fonction phatique (Jakobson) dans le code d'une autre langue.

b. L'acceptabilite

Toutes les TA passent le test d'acceptabilite - c'est-a-dire que le lecteur comprend le sens du message - a condition de ne pas etre puriste en matiere de grammaire.

La plus agrammaticale des quatre TA, la plus "baragouineuse" est de toute evidence TA1.

Le probleme de l'emploi de "de" ou de "des" est particulierement difficile pour des francophones non natifs. Les logiciels de traduction ne sont pas assez avances pour resoudre ce probleme.

Tous les logiciels trebuchent sur le titre: "The power of Big Ideas" traduit par le "Pouvoir de grandes idees". Si en anglais le contexte generique de "Big Ideas" exige un article defini Ø, il est necessaire en francais de reintroduire ce marqueur, c'est-a-dire "les" qui se contracte avec la preposition "de" pour donner "des". C'est donc "le Pouvoir des Grandes Idees". Tous commettent l'erreur; peut-etre confondent-ils "des", contraction de la preposition "de + les", avec "des" article indefini pluriel qui se transforme en "de" quand il est suivi d'un adjectif + nom: exemple elle a de beaux yeux au lieu de elle a des beaux yeux.

A contrario, toutes les TA sauf la TA4 donnent pour "an online gathering of scientists" "une reunion des scientifiques". L'article anglais Ø est ici le pluriel de "a", qui le plus souvent se traduit effectivement par "des". Malheureusement, l'article indefini pluriel "des" s'efface apres la preposition "de" par exemple j'ai achete des livres mais j'ai fait un achat de livres. Seule la TA4 est grammaticale: un rassemblement de scientifiques.

Une autre subtilite que le logiciel de TA1 n'a pas integree est la question de la place de l'adjectif epithete, qui le plus souvent est determinee par le sens (exemple: un homme grand et un grand homme). "grand" dans le sens "d'important" est obligatoirement antepose. C'est donc le Pouvoir des grandes idees et non le Pouvoir d'idees grandes. Par

contre les participes passés sont toujours postposés. "The past 2000 years" rendu dans TA1 par "les 2000 passés ans" est un charabia inacceptable.

D'autre part la TA1 ne respecte pas la différence entre "an" et "année" le "couple infernal" rhématique et thématique selon Henri Adamczewski[iv] (p. 223) "an" est rhématique c'est-à-dire non-presupposant mais "année" est presupposante. "Ces 2000 dernières années" est forcément presupposante.

Concernant "nominates", toutes les TA le rendent par "nommer", ce qui pose le problème des "faux amis" que les logiciels ne savent pas discerner. Si "to nominate" peut être suivi d'une personne ou d'une chose, en français "nommer" ne peut être suivi que d'une personne, d'où le néologisme français les "nommés" pour parler d'une personne ou d'une œuvre sélectionnée pour une récompense.

Il reste à signaler une erreur d'orthographe de "pilule" dans TA2 et TA3 écrite avec 2 l "pillule" et une graphie très anglo-saxonne des chiffres, dans TA1 et TA4 où les unités de mille sont marquées par une virgule (2.000) au lieu d'un espacement.

c. L'informativité

C'est ce qui est la raison d'être d'un texte. On lit rarement un texte pour y trouver tout ce que l'on sait déjà, c'est-à-dire que tout texte doit contenir du nouveau

- faute de quoi il serait dépourvu d'intérêt et du connu
- faute de quoi le lecteur ne pourrait comprendre le nouveau.

Dans ce petit texte, le terme qui pose problème est le mot "on line" dans "an online gathering of scientists" que toutes les TA rendent par "une réunion en ligne de scientifiques".

Le mot "on line" étant d'une fréquence plus grande aux États-Unis que le mot "en ligne" en France, qui manifeste un certain retard dans le domaine de l'Internet. Citons *Le Monde* qui n'offre à ses lecteurs un supplément bimensuel bien documenté que depuis le début de l'année 1999. Il serait plus judicieux de rendre l'hyponyme "on line" par l'hyperonyme "sur Internet" pour une meilleure compréhension (voir ma proposition de traduction).

d. La situationalité

C'est l'adéquation du style du texte au sujet traité. Ici le texte est en quelque sorte une devinette, un problème qui sera résolu en lisant le corps de l'article qui suit. Ce critère n'offre pas de difficultés aux quatre logiciels qui se limitent à traduire littéralement des mots ou des groupes de mots.

Il reste à savoir si un traducteur humain, français natif, poserait la situation de la même façon car le traducteur sait être un médiateur entre deux cultures.

e. L'intertextualité

La prise de conscience de l'intertextualité d'un texte permet de décrypter les allusions et de revenir sur les connaissances présupposées nécessaires à la compréhension d'un texte.

Dans notre cas, il faut savoir que "Big Ideas" est une metonymie pour les "Grandes Inventions". Il est évident que les quatre logiciels ne peuvent prendre en considération cette figure de rhétorique car rien n'est plus subjectif et difficile à transposer dans une autre langue que les figures de style.

Il faut d'autre part savoir que l'ampoule électrique remonte à 1911 et que la pilule anti-conceptionnelle date de 1955, ce qui entraîne obligatoirement en anglais le prétérit "was" que les quatre TA rendent incorrectement par un imparfait. Or l'emploi de l'imparfait est un tiroir verbal difficile à saisir pour un étranger et son emploi ici n'est pas pertinent.

Nous venons de voir que tous les critères de textualité de Beaudrande et Dressler n'avaient pas été respectés: le critère d'acceptabilité a été le plus maltraité, surtout par la TA1. Il était presque certain que nous arriverions à cette conclusion, étant donné que la TA est aujourd'hui encore une "technique balbutiante" comme l'écrit *Le Monde* mais qui progresse très vite grâce à l'incitation que représente la Toile. Il est étonnant de constater que le logiciel le moins cher, numéro 4, de Globalink qui ne vaut que 1450 francs est le meilleur et passe devant le Systran mondialement utilisé. Mais le texte à traduire est trop court pour pouvoir en tirer une conclusion probante.

"Mais où est l'éloquence ?", demande *Le Monde*. Il est évident qu'une traduction faite par une machine même si tous les critères de textualité sont respectés, exige une révision humaine surtout si l'on veut produire un texte publiable. Car comme dit Valéry Larbaud, la traduction est une transfusion de sens et son encodage dans la langue source n'est pas forcément le même que celui de la langue cible. On ne peut demander aux logiciels de T.A. d'avoir une compétence textuelle.

Proposition de traduction humaine

La stylistique du français impose des règles de clarté et de concision. Il a donc fallu procéder à des modifications des TA, notamment à propos de récurrences du même mot ou des mots d'une même famille (importance, important).

- *La portée des grandes inventions.*

J'ai choisi "portée" parce qu'il a un sens métaphorique qui compense la perte de la metonymie de "Big Ideas".

- *L'ampoule électrique a-t-elle eu plus de retentissement que la pilule anticonceptionnelle?*

J'ai procédé au renforcement de "pilule" par l'adjectif "anticonceptionnelle" pour établir un parallèle avec "ampoule électrique" nécessaire à la cohérence.

D'autre part, le passé me semble plus adéquat que l'imparfait qui sous-entend un pré-supposé.

- *Des scientifiques réunis sur Internet désignent les inventions les plus importantes de ces deux millénaires.*

J'ai procede

- a un changement de categorisation de "une reunion en ligne de scientifiques", plutot gauche, pour "des scientifiques reunis sur Internet" (hyperonyme de "en ligne").
- a une compression des "2000 dernieres annees" en un hyperonyme: "de ces deux millenaires".

Certaines reponses peuvent surprendre

J'ai supprime le deictique anaphorique "leurs", ce qui entraine le remplacement de "choix" par "reponses" et qui allège la phrase.

J'ai également supprime le pro-forme "vous" qui n'est pas necessaire a la transmission du message.

CONCLUSION

Le langage et la traduction dependant autant de la rhetorique que de la linguistique, il etait evident que les 4 TA manqueraient "d'eloquence". Les logiciels de T.A. les plus sophistiques se limitent aux problemes de linguistique qu'ils n'arrivent d'ailleurs pas a resoudre completement. Le Monde a tort d'ironiser. Les traductions humaines et les traductions automatiques s'adressent a des destinataires differents. Les T.A., qui sont des traductions brutes, permettent un dechiffrage rapide et peu onereux d'un texte essentiellement technique ou administratif. On ne peut leur demander ce qu'elles ne peuvent pas donner. Comme l'ecrit Anne-Marie Loffler Laurian : Si la traduction automatique fonctionnait parfaitement, l'homme aurait enfin reussi a fabriquer une machine qui travaille a sa place et il pourrait aller se reposer. Elle rejoint la conclusion du journaliste du Monde: On ne verra pas de vraie traduction automatique avant l'avenement complet de l'intelligence artificielle. Ce qui n'est pas pour demain.

Christian Papas

Enseignant a l'Universite Ionienne

Departement de Langues Etrangeres, de Traduction et d'Interpretation

10, rue Rartourou, Corfou

¹ LOFFLER-LAURIAN A.-M., *La Traduction automatique*, Presses Universitaires du Septentrion, 1996.

² DRESSLER-BEAUGRANDE, *Introduction to Text Linguistics*, Longman, 1996.

³ DURIEUX C., *Apprendre a traduire, Prerequis et Tests*, La Maison du Dictionnaire, 1995.

⁴ ADAMCZENSKI H., *Le Francais dechiffre, Cie du langage et des langues*, Armand Colin, 1991.