

La Traduction Automatique : l'état d'art et perspectives

Autor(en): **King, Margaret**

Objektyp: **Article**

Zeitschrift: **Bulletin / Vereinigung Schweizerischer Hochschuldozenten = Association Suisse des Professeurs d'Université**

Band (Jahr): **19 (1993)**

Heft 3

PDF erstellt am: **21.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-894220>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

La Traduction Automatique: l'état d'art et perspectives

Margaret King

1. La traduction.

Même lorsque deux langues proviennent de la même famille, il y a de nombreuses différences entre elles, ce qui fait qu'il est impossible de dire la «même» chose de la même manière dans les deux langues, voire, à la limite, de simplement dire la «même» chose.

Il y a certains exemples classiques de ce phénomène: différentes langues utilisent différemment les mots désignant les couleurs:., p.ex., le russe fait plus de distinctions entre les bleus, et «candidus» en latin caractérise plutôt l'intensité de la lumière que la couleur en soi. Les différences concernant la modalité et l'aspect sont plus subtiles: certaines langues, comme l'italien et le français, permettent au locuteur, lorsqu'il rapporte un événement, d'indiquer, par l'utilisation d'une forme particulière du verbe, qu'il réserve sa position quant à la véracité de ce qu'il exprime, d'autres encore, comme le grec, font la distinction entre une action qui n'est effectuée qu'une seule fois et la même action pouvant être répétée plusieurs fois. Le japonais, comprend, dans la forme de surface de la langue, be-

aucoup de réflexions directes sur l'attitude du locuteur vis-à-vis de ce qu'il dit, ainsi que des réflexions sur ce qu'il pense de son interlocuteur.

Dans ces cas – et ils sont nombreux – le traducteur se trouve dans une situation difficile. Souvent la langue dans laquelle il traduit exige l'utilisation d'informations qui ne se trouvent tout simplement pas dans le texte qu'il traduit. Afin de faire la traduction, il doit calquer le monde de la langue source sur le monde de la langue cible d'aussi près que possible, et, ensuite deviner l'information supplémentaire requise par la langue cible. Et, bien sûr, cela s'applique également dans l'autre sens. Si la langue source décrit des aspects du monde que la langue cible ne peut pas représenter avec des moyens similaires, le traducteur doit soit trouver d'autres moyens pour exprimer ces aspects, soit décider de s'en passer. Ce type de problèmes se pose déjà régulièrement avec des traductions non-littéraires. Comme on peut le supposer, ils atteignent une forme extrême dans le cas de traductions littéraires ou poétiques, où une bonne traduction peut quelquefois exiger, au niveau littéral, un changement du contenu de l'original afin d'en préserver l'essence.

Sans aller jusqu'aux usages figuratifs de la langue, il y a d'autre classes de problèmes de base. Par exemple, les langues permet-

Margaret King est directrice de l'ISSCO et chargée d'enseignement à l'Université de Genève.
Adresse: ISSCO, 54 rte des Acacias, 1227 Genève.

tent différents types de structures syntaxiques: en anglais, on peut dire «I expected John to come», mais en français on dit «Je m'attendais à ce que Jean vienne». Il n'est pas toujours le cas qu'on puisse simplement transformer une structure syntaxique d'une langue en une structure syntaxique différente, mais équivalente, d'une autre langue. Tant l'anglais que le français permettent des constructions passives, mais ils ne les utilisent pas de la même manière: le résumé d'un article scientifique en anglais est souvent écrit au passif («Experiments were carried out ...» «Subjects were examined ...») alors qu'en français l'impersonnel «on» est utilisé («On a déterminé que ...»). Mais on ne pourrait pas, en se basant sur cette observation, transformer tous les passifs anglais en impersonnels français. «The stocks were bought but sold again the same day» est traduit correctement par «Les actions ont été achetées mais revendues le même jour» et non par une phrase utilisant l'impersonnel. Il y a une différence subtile, mais nette, entre «The question was never asked» et «Nobody asked the question» qui n'est pas celle qu'il y a entre «La question n'a pas été posée», «La question ne s'est pas posée» et «On n'a pas posé la question». Donc, même si les deux langues utilisent des structures syntaxiques très similaires, il n'est jamais certain qu'elles soient utilisées de la même manière.

Finalement, il y a tous les cas de lacunes lexicales, lorsqu'un élément lexical existe dans une langue mais n'a pas d'équivalent dans l'autre. On peut se demander, par exemple, comment traduire «méchant» en

anglais; le choix va dépendre du contexte. Il n'y a pas un seul mot anglais qui corresponde dans tous les cas. Le néerlandais a deux mots pour la piste d'atterrissage, un pour quand l'avion atterrit, l'autre pour quand il décolle. Dans ces cas, et, de nouveau, il y en a beaucoup, le traducteur est obligé à recourir à la paraphrase; par définition, donc, il ne peut pas trouver une traduction exacte.

Tout cela revient à dire que, quelle que soit l'ingéniosité du traducteur, quelle que soit l'amplitude de sa connaissance des langues et des cultures impliquées, quelle que soit la similarité entre ces langues et ces cultures, la traduction implique toujours une perte d'information et une déformation du sens original. Le travail du traducteur est de minimiser les dommages autant que possible.

2. La traduction automatique.

Le terme même de «traduction automatique» ou «traduction par ordinateur» nous dit pourquoi la traduction automatique rencontre de plus grandes difficultés que la traduction humaine. Outre les problèmes esquissés brièvement ci-dessus, des complications additionnelles surgissent nécessairement du fait qu'un ordinateur est utilisé.

Le premier problème que nous examinerons ici, celui de l'ambiguïté, est déjà présent lorsqu'un humain traduit mais il est relativement rare qu'il s'en rende compte. Presque chaque phrase, parlée ou écrite, est bourrée d'ambiguïtés d'un type ou d'un autre. Mais du fait que les gens savent de quoi il est question, la plupart du temps ils

résolvent l'ambiguïté sans en être conscients. Prenons la phrase suivante:

«La fermière a vendu la vache parce qu'elle avait besoin d'argent»

La plupart des gens ne se rendent pas compte que «elle» est ambiguë jusqu'ils soient confrontés avec la phrase parallèle:

«La fermière a vendu la vache parce qu'elle ne donnait plus de lait».

Mais un système informatique, à l'exception des rares cas où il traite un texte très limité, doit être nourri d'informations concernant toutes les possibilités d'interprétations possibles. Pour lui, alors, toutes les ambiguïtés sont à résoudre. Un examen un peu plus détaillé du problème servira à nous convaincre de son sérieux.

Commençons par ce qui est le plus évident: les mots peuvent être ambigus. La chaîne de caractères «bank» en anglais peut signifier le bord d'une rivière, l'endroit où l'on change l'argent, un talus au bord d'un champ: chaque possibilité se traduit différemment. Un mot peut également être vague, c'est-à-dire qu'il peut pratiquement n'avoir aucun sens par lui-même. Les verbes comme «avoir» et «faire» sont souvent utilisés de façon vague, par exemple dans des locutions comme «avoir froid», «avoir faim», «faire connaissance», «faire du progrès». Certains mots peuvent être ambigus parce qu'ils peu-

vent jouer des rôles syntaxiques différents, selon la phrase. «Ferme», par exemple, considéré en isolation, peut être un substantif («la ferme de ma tante») un adjectif («le soldat ferme résiste») ou encore un verbe («ferme la porte»).

Si l'on laisse de côté les mots pris de manière isolée et si l'on regarde les mots mis ensemble formant des structures syntaxiques, celles-ci peuvent, elles aussi, être ambiguës. Dans la phrase «Il était assis dans le jardin avec des fleurs», il est clair que «les fleurs» qualifient le jardin, tandis que dans la phrase «Il était assis dans le jardin avec ses amis», «ses amis» n'ont rien à faire avec le jardin; pourtant, rien dans la structure syntaxique de la phrase ne nous indique une différence d'interprétation et donc, à un niveau purement structurel, ces deux phrases sont ambiguës.

De plus, l'ambiguïté lexicale et l'ambiguïté structurelle peuvent agir ensemble. Dans la phrase très simple «Le pilote ferme la porte» presque chaque mot est ambigu. En conséquence, deux structures syntaxiques et deux interprétations sémantiques complètement différentes sont possibles:

(phrase	(le pilote) groupe nominal sujet	(ferme groupe verbal verbe	(la porte)) groupe nominal objet
(phrase	(le pilote ferme) groupe nominal sujet	(groupe verbal verbal	(la) groupe nominal objet
			porte) verbe

Le problème de l'ambiguïté n'est évidemment pas spécifique aux systèmes de traduction automatique. Tout système informatique qui traite des langues naturelles doit aborder ce problème d'une manière ou d'une autre. Cependant, il doit être clair que la résolution correcte d'ambiguïtés dans un texte détermine si la traduction du texte est juste ou fausse. Lorsque les choses vont mal dans un système de traduction automatique, les traductions peuvent devenir carrément farfelues. En conséquence d'une mauvaise résolution d'ambiguïtés, un système a traduit «La cour prévoit l'établissement d'un cinquième poste d'avocat général» par «The playground foresees the establishment of a fifth avocado station»: même les plus généreux auront de la peine à appeler ceci une traduction correcte.

Ensuite, il y a une classe d'ambiguïtés qui est bien spécifique aux systèmes de traduction automatiques. Il y a des cas où un simple mot dans une langue a deux ou plusieurs équivalents dans la langue cible. Le néerlandais, par exemple, a deux mots qui correspondent au mot anglais «put». Leur utilisation dépend de la manière de poser un objet, à plat ou debout. L'italien a deux mots correspondant au mot anglais «wall», faisant la différence entre un mur interne ou externe. Ces ambiguïtés ne font surface que lorsqu'on essaie de trouver la bonne traduction. Ils sont «ambigus» au niveau de la traduction, bien qu'on ne puisse pas vraiment dire qu'ils soient ambigus dans la langue source.

Ensuite se posent encore les problèmes bien connus des prépositions dans les langues européennes. Voici un tout petit

échantillon de traductions françaises de la préposition «by» en anglais:

He came by car	Il est venu <u>en</u> voiture
He came by night	Il est venu <u>pendant</u> la nuit
He came by the town	Il est venu <u>par</u> la ville
He came by three o'clock	Il est venu <u>avant</u> trois heures

L'ambiguïté des prépositions constitue en réalité un cas spécial de l'ambiguïté des sens des mots déjà examinée. Elle est considérée comme un problème spécial surtout à cause de leur utilisation très fréquente et parce qu'il est particulièrement difficile d'établir quelles sont les informations qui conduisent à un bon choix.

Comme nous l'avons déjà dit, l'ambiguïté est toujours présente dans un texte, qu'il soit traduit par un humain ou par une machine. Mais il y a une différence. Si l'on dit à un traducteur humain que «bank» dans la phrase «He sat down on the bank and watched the fish rising» est ambigu, il sera vraisemblablement incrédule et il niera que «bank» dans ce contexte puisse être une banque. Et il aura raison. Le traducteur humain travaille dans le contexte de sa traduction et ne s'attarde donc pas à considérer certaines alternatives. Malheureusement, il est très difficile, sauf dans des circonstances très spéciales, de donner aux systèmes informatiques la même conscience du contexte global: cette conscience lui faisant défaut, un ordinateur prendra en considération toutes les alternatives connues par le système, et tiendra donc compte de possibilités que le traducteur humain ne remarquerait même pas.

Ceci nous amène à une autre raison pour

laquelle la traduction automatique est plus difficile que la traduction humaine. Les machines manquent d'intelligence: elles ont une certaine capacité élémentaire pour effectuer des opérations logiques ou arithmétiques simples lorsqu'on leur a donné des instructions pour le faire, et cette capacité peut être utilisée comme base pour fabriquer des opérations plus complexes. Mais c'est tout: il est inutile de s'attendre à ce qu'un ordinateur sache ce qu'est un substantif, encore moins qu'un certain mot se comporte comme un substantif ou comme un verbe, ou se réfère normalement à un objet animé ou qu'un verbe comme «boire» préfère normalement avoir comme objet quelque chose de liquide.

Ceci signifie que tout ce que nous savons sur les langues entre lesquelles il faut produire une traduction doit être rendu absolument explicite dans le logiciel qui effectuera la traduction. Nous devons y déclarer à quoi ressemble habituellement un groupe nominal ou une phrase, tout comme nous devons lui expliquer où se trouvent les interprétations alternatives d'un mot ou d'une phrase, comment choisir entre les alternatives et sur quelles bases, tout en s'assurant que les informations requises pour pouvoir choisir sont connues par le système au moment où nous voudrions que le choix soit effectué.

D'une certaine manière, un système de traduction automatique ressemble à un système expert où l'expert, dont les connaissances doivent être codifiées à l'intérieur du système, est un traducteur. Mais exprimer les choses ainsi les fait paraître bien plus faciles qu'elles ne le sont en réalité. La

traduction est encore une activité assez mystérieuse: si l'on demande à un traducteur pourquoi il a choisi une traduction plutôt qu'une autre, il est rare qu'il soit capable de répondre en des termes qui puissent être formalisés facilement. Souvent, il n'est même pas conscient lui-même de la façon dont il a fait ses choix. Ce serait même étonnant qu'il le soit. Une partie de la formation d'un traducteur et de l'expérience acquise au cours des années de pratique consiste à rendre instinctif le processus de la traduction: si chaque fois qu'il a un choix à faire, le traducteur devait s'arrêter et chercher par un processus rationnel de réflexion quel choix adopter, il ne pourrait pas faire plus d'une page par jour. Et rappelons-nous de surcroît que certaines alternatives sont considérées par un système informatique comme des choix actifs, alors qu'il ne viendrait jamais à l'esprit du traducteur humain qu'une alternative puisse exister. Dans ces cas, il est inutile de demander au traducteur quel choix il ferait: il n'en ferait pas.

Donc le problème ne consiste pas simplement à parvenir à extraire les connaissances d'un expert dans le domaine et de les codifier sous une forme appropriée pour être utilisées par l'ordinateur. Il peut être nécessaire de concevoir et d'utiliser des types d'informations que le traducteur humain n'utilise jamais ainsi que de rendre explicite ce qu'il utilise réellement. En d'autres termes, on peut considérer la conception et la construction d'un système de traduction automatique comme une entreprise d'ingénierie: il n'y a pas de raison pour laquelle un ordinateur effectuerait le

travail de la traduction de la même manière qu'un traducteur humain le ferait.

3. Les machines traduisent quand même.

Après lecture des deux chapitres précédents, on pourrait fort bien arriver à la conclusion que la traduction automatique est une tâche sans espoir, qu'il vaudrait mieux en oublier l'idée et continuer à faire la traduction à la main et avec le cerveau. Cependant, ce serait une conclusion trop pessimiste; la traduction automatique peut réussir, et réussit déjà.

Mais les conditions de la réussite sont importantes. Il faut d'abord être réaliste en ce qui concerne le type de textes pouvant être traités par ordinateur. Il va sans dire que la traduction littéraire doit être réservée aux traducteurs humains. La même chose vaut pour certains textes non-littéraires, comme des textes juridiques où non seulement tous les éléments du texte et de la traduction doivent correspondre mais également toutes les implications pouvant être inférées du texte. Même lorsqu'il s'agit de traductions humaines, de tels textes sont souvent traduits par des juristes hautement qualifiés.

En outre, toute ambiguïté dans le texte devrait être accidentelle plutôt que délibérée. Normalement, un système de traduction automatique visera à lever l'ambiguïté et non à la conserver. (En effet, il est généralement très difficile de préserver l'ambiguïté lorsqu'on fait une traduction). Ce critère fournit une autre raison pour l'élimination de certains textes littéraires, des poèmes, par exemple, où l'ambiguïté

est voulue et essentielle. D'autres textes seront également éliminés, comme des discours rhétoriques qui essaient de convaincre par tous les moyens.

Il faut aussi être conscient du fait que le texte cible produit par un système de traduction automatique exigera, sauf dans quelques cas bien spécifiques, d'être révisé par un être humain. Souvent, et à tort, ce fait est utilisé comme un argument contre la faisabilité de la traduction automatique par ceux qui oublient que la traduction humaine subit, également, en règle générale, des corrections au moment de la révision. Dans les services de traduction importants, les traducteurs et les réviseurs forment deux groupes distincts. Dans de petits services, le traducteur fait lui-même sa propre révision. Mais il est presque impossible qu'un traducteur finisse sa traduction et arrête là son travail.

Ensuite, il faut se rappeler que les systèmes de traduction automatique seront rarement capables de produire un style très élégant, où de changer de style au cours de la traduction. Souvent, ils peuvent produire un style approprié, par exemple le style qui convient à un résumé scientifique où à un document administratif, quand il est possible de rendre explicite les aspects linguistiques et autres qui caractérisent ce type de document. Mais pour le reste, les systèmes automatiques ont tendance à produire un style plutôt laborieux, respectant, par exemple, la longueur des phrases dans le document source, et à se satisfaire d'une traduction dont la syntaxe soit correcte et qui soit fidèle au contenu de l'original, alors qu'un traducteur humain travaillant

sur le même texte s'efforcerait de tourner les choses de manière à obtenir une traduction plus élégante qui se lise plus facilement.

En résumé, le type de texte le mieux adapté au traitement automatique est celui qui évite l'ambiguïté autant que possible, qui n'est pas trop nuancé, et pour lequel un style un peu lourd dans la traduction ne sera pas trop gênant.

Parfois ceux qui ne sont pas des professionnels de la traduction sont surpris d'apprendre que des textes conformes à cette description forment la majorité du travail des services de traduction. En outre, souvent ces documents n'ont qu'une durée d'existence limitée: ils sont préparés pour une occasion particulière, utilisés intensivement à cette occasion et ensuite classés dans les archives et oubliés: La nature des documents ainsi que leur brève existence rendent ces travaux quelque peu frustrants pour les traducteurs qui préféreraient utiliser leur temps à faire des traductions plus stimulantes et d'une importance plus durable.

Il y a d'autres domaines où la traduction automatique a également fait les preuves de son utilité. Le plus évident est dans le recueil des informations. Prenons un exemple classique, celui d'un spécialiste qui se doit de rester informé sur les travaux faits dans son domaine même si ces travaux sont rapportés dans une langue qu'il ne connaît pas. Pour lui, le système automatique a atteint son but si la traduction qu'il fournit lui permet de juger quels articles méritent d'être traduits correctement – et la traduction automatique peut faire mieux que cela.

Il y a également des domaines où la vitesse de traduction est d'une importance capitale. Ici, l'exemple classique est celui des bulletins de prévision météorologique dont le langage utilisé, bien que souvent complexe, est néanmoins assez stéréotypé, une caractéristique qui favorise un traitement informatisé. Les bulletins météo au Canada sont traduits par ordinateur depuis quinze ans déjà, et à la vitesse d'environ 300 bulletins par jour.

Un dernier exemple des applications possibles de la traduction automatique est la documentation technique. Ici, souvent la contrainte la plus importante est que la terminologie doit être cohérente: les ordinateurs, par manque d'imagination, sont remarquablement bons à cet égard. De plus, la documentation technique doit souvent être produite rapidement, et ne nécessite pas un style élégant.

Jusqu'à présent, notre principale considération a été de déterminer certains domaines où la traduction automatique présentera un résultat valable: mais, sauf en passant, les avantages de la traduction automatique n'ont pas encore été mentionnés. Le premier avantage est assez évident: la traduction automatique est rapide. La vitesse doit être considérée sous deux aspects séparés: la rapidité avec laquelle la traduction brute peut être produite et le temps nécessaire pour la révision de la traduction brute afin d'en faire quelque chose d'utilisable. (Cela ne sert à rien de produire 10 pages en 5 minutes si le réviseur doit mettre deux heures pour corriger chaque page: si la traduction est mauvaise à ce point, tout réviseur normal l'aurait mise depuis

longtemps à la poubelle et aurait commencé à la faire lui-même). Les vendeurs de systèmes automatiques sont un peu réticents à donner des chiffres exacts, simplement parce que ces chiffres pourraient peut-être être améliorés considérablement dans un avenir très proche à cause de l'évolution rapide de la technologie, mais pour donner une indication approximative, on trouve mention dans la littérature scientifique d'un système qui traduit à la vitesse d'environ 150'000 mots par heure, et d'un autre où un réviseur, partant des résultats bruts, arrive en moyenne à produire 29 pages de traduction utilisable par jour. Ces chiffres doivent être comparés aux 5-8 pages propres que produit en moyenne un traducteur humain. Et il ne faut pas oublier non plus qu'un ordinateur n'a pas besoin de sommeil ni de vacances.

Le second avantage s'ajoute au premier. Il est souvent difficile de trouver des traducteurs ayant des paires de langues «rares», danois-grec, par exemple, ou portugais-russe. Si la qualité de la traduction est suffisamment bonne pour que la révision puisse être faite par quelqu'un qui ne connaît que la langue cible, voici un moyen de résoudre la difficulté.

L'avantage final est un peu moins évident. Lors d'une discussion sur la traduction automatique, les traducteurs expriment souvent la crainte d'être remplacés par des machines. Dans la pratique, il est très peu probable que cela se passera ainsi. Les temps où une communauté de langue pouvait imposer son parler au reste du monde sont heureusement presque révolus, tant pour des raisons économiques que politiques.

Les besoins en traduction sont grands et ne cessent de croître. C'est un fait établi que les services existants ne peuvent pas venir à bout du volume de traductions qu'ils devraient traiter. Employer plus de traducteurs constitue rarement une solution faisable ou complète, aussi bien parce qu'il est difficile de trouver suffisamment de traducteurs qualifiés que parce que l'augmentation énorme de la taille des services qui en résulterait causerait autant de problèmes que cela n'en résoudrait. Si une machine peut être utilisée pour soulager le traducteur des travaux les plus fastidieux, le travail serait fait et la vie du traducteur serait plus satisfaisante.

Pour résumer tout ceci en quelques mots; la traduction automatique existe déjà. On voit son utilisation dans des applications très terre-à-terre, comme par exemple, la traduction de la documentation technique ou ailleurs où l'on peut définir une langue restreinte. Pendant les années à venir, des systèmes de traduction automatique seront utilisés de plus en plus pour ce type d'application. On peut imaginer aussi le développement d'autres types d'application ingénieux, inventés par un concepteur qui connaît bien les limites de la technologie actuelle, mais qui voit des possibilités commerciales inattendues. Mais la traduction littéraire n'est pas pour demain; afin d'attaquer les problèmes qu'elle pose, il faut résoudre des problèmes fondamentaux dont, au moins pour le moment, on ne sait pas où chercher la solution.