

SEMAINE D'OUVERTURE SCIENTIFIQUE ET INDUSTRIELLE

Les Traducteurs Automatiques



Auteurs :

Manon ANSART

Etienne BATISE

Grégoire GUTZWILLER

Date de remise du présent rapport :

6 Septembre 2013

Table des matières

Introduction	1
1 Description	2
1.1 Historique	2
1.1.1 Le début de la recherche	2
1.1.2 L'impact de l'ALPAC	2
1.1.3 L'influence du Japon et le renaissance la T.A.	3
1.2 Etat de l'Art	3
1.2.1 La position dominante Systran	4
1.2.2 La force de Google Traduction	4
2 Fonctionnement	5
2.1 Dictionnaire de mots	5
2.2 La traduction mot à mot	5
2.3 La traduction sémantique	6
2.4 La traduction statistique	7
3 L'avenir de la traduction automatique	8
3.1 Les problèmes actuels	8
3.2 Les amélioration possibles	9
3.3 La traduction automatique dans le futur	10
3.3.1 Les nouvelles utilisations	10
3.3.2 Les traducteurs	11
Conclusion	12
Bibliographie	13

Introduction

Avec l'arrivée de nouveaux moyens de transports puis de communications, la terre n'a plus de frontières. Toutes les nations sont amenées à communiquer par le langage. Aussi, il a fallu apprendre à communiquer selon plusieurs langues, et donc de traduire le sens de la parole d'une langue à une autre.

C'est dans ce contexte qu'intervient la traduction automatique. Un traducteur automatique est un programme informatique capable de traduire un texte sans l'intervention humaine. Conçus pour faciliter la tâche des utilisateurs, ceux-ci sont aujourd'hui de plus en plus performants. En effet, aussi bien pour les traducteurs professionnels que pour les traducteurs Web, les prouesses techniques sont de plus en plus impressionnantes. Aujourd'hui, le service de traduction de Google supporte plus de 64 langues différentes, et est utilisé par plus de 250 millions d'utilisateurs chaque mois. Néanmoins, il subsiste toujours des erreurs et des améliorations régulières sont vraiment nécessaire pour proposer un service optimal et sans erreur. En effet, le langage est une caractéristique du cerveau humain très difficile à comprendre et à imiter.

La question de l'évolution passée et à venir de la traduction automatique se pose alors : est-il possible de rendre nul le taux d'erreur ?...

Ce rapport a donc pour but de présenter les améliorations des traducteurs au cours des années et le fonctionnement de ceux-ci, afin de répondre à la problématique précédente.

1 Description

Pour bien comprendre l'importance de cette technologie aujourd'hui, il est important de comprendre d'où elle vient et connaître le chemin parcouru pour en arriver à l'état actuel.

1.1 Historique

1.1.1 Le début de la recherche

D'après Jacqueline Léon chargée de recherche au CNRS au Laboratoire d'histoire des théories linguistiques, la première période dans l'histoire de la traduction automatique s'étend de 1948 à 1960. Elle constitue la phase expérimentale débutée en 1948 par la Grande Bretagne puis par les USA. Par la suite, le développement applicatif et la recherche dans le domaine de la T.A. prennent une envergure mondiale. C'est d'ailleurs dans un contexte de guerre froide que « les traductions concernent essentiellement le russe, pour des raisons militaires et politiques » écrit Jacqueline Léon.

Pour répondre à cette problématique de traduction automatique des langues, de nombreuses études ont débuté dans différentes universités spécialisées à travers le monde. Maurice Gross, nous apprend dans l'article *Notes sur l'Histoire de la traduction automatique* que ces premières études ont permis d'éliminer « quelques idées grossièrement inefficaces : la T.A. comme un problème de cryptographie, de théorie de l'information, de compression de "l'information", de construction d'un nouvel espéranto (maghinois, langue-pivot, etc.). Il semblait alors que la T.A. fût un problème conjoint de linguistique et d'informatique ».

Cette période se conclue avec la création de nouveaux outils. La traduction est littérale, c'est du mot-à-mot, c'est pourquoi les premiers systèmes de T.A. sont des dictionnaires automatiques.

1.1.2 L'impact de l'ALPAC

Durant une courte période, la recherche dans le domaine linguistique se dirige vers une analyse syntaxique. C'est durant cette période, comme l'écrit Jacqueline Léon, que « la production en séries de traductions est reléguée à des temps futurs, ce qui va conduire, à terme, à la suppression des recherches en traduction automatique au profit de la seule linguistique computationnelle. »

Par la suite, en 1966, l'ALPAC¹ va brusquement bouleverser les choses, suite notamment à la publication en 1965 par Noam Chomsky, fondateur de la grammaire générative, du livre *Aspects of Theory and Syntax*. En effet d'après Marie Lebert dans son livre *Livre 010101 : Enquête* « le rapport ALPAC fait une estimation prématurément négative de la valeur des systèmes de traduction automatique, et des perspectives offertes par ceux-ci, mettant fin au financement et à l'expérimentation dans ce domaine pour la décennie suivante. ». L'impact économique conséquent engendré par cette annonce ne permettra qu'à peu de systèmes de surmonter cette crise. C'est ce que confirme Jacqueline Léon,

1. Automatic Language Processing Advisory Committee

Il reste quelques systèmes de recherche (tel Ariane, développé par le GETA à partir de 1971, ou Susy, élaboré à l'université de Sarrebruck) qui ont surmonté la crise parce qu'ils étaient soutenus par de grands organismes de recherche, tel le CNRS, préoccupé seulement de façon secondaire par les aspects économiques de la traduction automatique.

En effet comme l'explique Maurice Gross, « Alors qu'il est très facile de programmer une traduction mot à mot et même correcte à 50%, pour les phrases courtes rencontrées dans un texte usuel, il est clair qu'une traduction est encore inutilisable si elle n'est correcte qu'à 90%. De fait les centres de T.A. ont cessé de fonctionner parce que les essais de traduction en dehors du corpus ayant servi de base à la confection de programmes n'étaient pas généralisables à de nouveaux textes ». C'est pourquoi les groupes à avoir résisté à l'impact de l'ALPAC étaient les groupes dont les traductions n'étaient pas faites dans le but d'être publiées. Elles pouvaient donc se permettre d'être de mauvaise qualité. Ce fut le cas de SYSTRAN, aujourd'hui acteur majeur et spécialisé à l'usage de la traduction automatique dans le domaine entrepreneurial.

1.1.3 L'influence du Japon et le renaissance la T.A.

En 1980, après une forte perte d'attention, le domaine de la T.A. repris sa course grâce aux importants progrès technologiques dans le domaine informatique. D'une part grâce au Japon qui voit dans l'avenir une société basée sur l'utilisation intensive des technologies de l'information. Dans un contexte de pré-mondialisation, on pense déjà à traduire « des modes d'emplois, des descriptifs de produits ou de machine, conditions indispensables pour leur commercialisation à l'échelon mondial » affirme Jacqueline Léon.

Ce n'est qu'en 1990 que les USA revoient leur position dans le domaine de la traduction automatique. En fait, le regain d'attention pour la T.A. se fait pour des raisons économiques : « nécessité de traduire la documentation des concurrents japonais, baisse des coûts de l'informatique et hausse des coûts de la traduction surtout pour les langues à alphabets non latins comme le japonais. ». Cela signifie que par sa politique de promotion des technologies de l'information, le Japon a en partie permis à la traduction automatique de revenir sur le devant de la scène mondiale et notamment sur la scène scientifique car si le besoin de traduction est bien réel, la recherche dans ce domaine n'avait pas avancé.

Du point de vue des méthodes utilisées et du/des « paradigme(s) » linguistique utilisé(s), les systèmes de T.A. sont revenus aux « approches empiriques » et au « traitement des grand corpus » explique Jacqueline Léon. D'après elle, le succès de ces approches est dû au succès des méthodes stochastiques²

1.2 Etat de l'Art

50 ans après le début de la recherche dans le domaine de la traduction automatique, quel est le bilan ?

2. Le calcul stochastique est l'étude des phénomènes aléatoires dépendant du temps. À ce titre, c'est une extension de la théorie des probabilités. www.wikipedia.org/wiki/Calcul_stochastique

Aujourd'hui, seul deux entreprises se démarquent du lot dans cette technologie : d'un côté SYSTRAN, le « leader mondiale des technologies de traduction automatique » et GOOGLE que l'on ne présente plus.

1.2.1 La position dominante Systran

Avec près de 40 ans d'expérience et d'expertise dans ce domaine, la société française SYSTRAN est incontestablement LA référence de la traduction automatique. D'après leur site web : « Les logiciels SYSTRAN traduisent instantanément en 52 paires de langues pour comprendre et publier tous types de contenus. L'utilisation des produits et solutions SYSTRAN améliore l'efficacité et la productivité des entreprises dans de nombreux domaines ». En effet, même son concurrent GOOGLE a longtemps utilisé les services de SYSTRAN à l'exception de quelques langues. C'est effectivement ce que nous apprennent Delphine Graeff, Édouard Lopez dans leur dossier SYSTRAN vs GOOGLE : « Google n'avait développé ses propres processus de traduction uniquement pour certaines langues non latines : arabe, chinois, coréen, japonais, russe ». Yves Drotier, blogueur sur le Journal de Net, nous informe dans l'article *Au cœur de la traduction automatique de SYSTRAN* que l'entreprise compte parmi ses clients de grands groupes internationaux comme Yahoo, Microsoft, des gouvernements et des entreprises tel que EDF ou BNP Paribas et des administrations publique. En effet, la Commission européenne³ qui se trouve être « la plus grande "entreprise de traduction" à l'échelle mondiale » est cliente de SYSTRAN comme le précise le P-DG de SYSTRAN dans l'article *Dimitris Sabatakakis (Systran)* : « Pendant sept ans, la Commission européenne a contesté la propriété de notre logiciel ! ».

La technologie utilisée par SYSTRAN, est notamment l'analyse syntaxique mais depuis quelques années a changé d'approche pour une analyse statistique.

1.2.2 La force de Google Traduction

Le succès de Google Traduction vient de plusieurs points très importants.

D'une part, et cela est sûrement l'un des points les plus importants, le succès de Google Traduction est dû à la popularité du moteur de recherche GOOGLE.

D'autre part c'est une service gratuit disponible sur l'Internet, et cela fait même partie de leur philosophie : « L'objectif de Google Traduction est de rendre l'information utile et accessible pour tous, quelle que soit la langue dans laquelle elle est publiée. » La gratuité du service complet de la part de Google fait forcément de l'ombre à SYSTRAN lorsque l'on voit que l'édition familiale Français-Europe de la firme est vendu à 99,00 €.

D'un point de vue technique, l'approche menée par GOOGLE est une analyse statistique. Néanmoins la spécificité de Google Traduction est, comme nous l'explique Yves Drotier, « qu'il tape sur le Web le mot recherché et détermine la traduction qui semble la plus probable ».

3. La Commission européenne représente l'intérêt général de l'UE. Elle joue un rôle moteur dans l'élaboration de la législation (en proposant des textes législatifs au Parlement et au Conseil), dans la gestion et la mise en œuvre des politiques européennes, dans le contrôle de l'application du droit européen (avec la Cour de justice) et dans les négociations menées par l'UE sur la scène internationale. http://ec.europa.eu/about_fr.htm

2 Fonctionnement

Le fonctionnement d'un traducteur automatique de nos jours est très compliqué, et celui-ci diffère selon les traducteurs. Rappelons que les premiers systèmes de traduction sont apparus peu après l'arrivée des premières machines dans les années 1946. A l'époque, la traduction se fait uniquement via un dictionnaire, et fonctionne au mot par mot.

2.1 Dictionnaire de mots

Un dictionnaire de mots pour un traducteur n'est ni plus ni moins qu'un dictionnaire réel enregistré avec le système. Un mot peut être recherché par le traducteur dans le dictionnaire pour en trouver plusieurs traductions accompagnées d'explications quant à leur contextualisation. Un dictionnaire est complet et généraliste : on y trouve plusieurs traductions d'un mot et le système n'est donc pas limité à un seul équivalent. L'ordinateur ne pourra donc pas faire d'interprétations erronées s'il suit correctement les informations données dans le dictionnaire.

2.2 La traduction mot à mot

Le processus de traduction en lui-même peut être divisé en trois phases successives dont les fondements sont posés par le Triangle de Vauquois.

1. L'analyse, qui correspond à la compréhension du sens du texte ;
2. Le transfert, où on ne conserve que le sens du texte et non les mots ;
3. La génération, c'est-à-dire la retranscription du sens en texte.

La traduction mot à mot agit à un niveau d'analyse particulièrement bas. En effet, on ne se soucie pas vraiment du sens du texte. L'ordinateur va simplement traduire les mots un par un après une étape de lemmatisation.

La lemmatisation en langue source consiste à mettre tous les verbes à l'infinitif, et mettre au singulier les noms dans la phrase. Après cette étape, l'analyse est déjà terminée et le transfert lexical va avoir lieu. Ce dernier va traduire chaque mot de la phrase concernée de la langue source à la langue cible. Il ne reste plus au système qu'à effectuer un réajustement.

Malheureusement les traductions effectuées avec ce genre de système sont hasardeuses, très approximatives, et parfois fausses. La qualité de la traduction est relativement dépendante du dictionnaire utilisé. Il n'est pas utile d'expliquer qu'un dictionnaire plus compliqué ou plus pointu limitera les confusions du système. En outre, il est impossible de traduire certaines expressions idiomatiques, et les mots à plusieurs sens sont en général mal gérés.

Un exemple simple : Essayons de traduire la phrase « Le professeur barbu a mordu le professeur chauve » en anglais. La lemmatisation donnera : « Le professeur barbu mordre le professeur chauve » . L'étape du transfert traduit alors cette phrase en « The

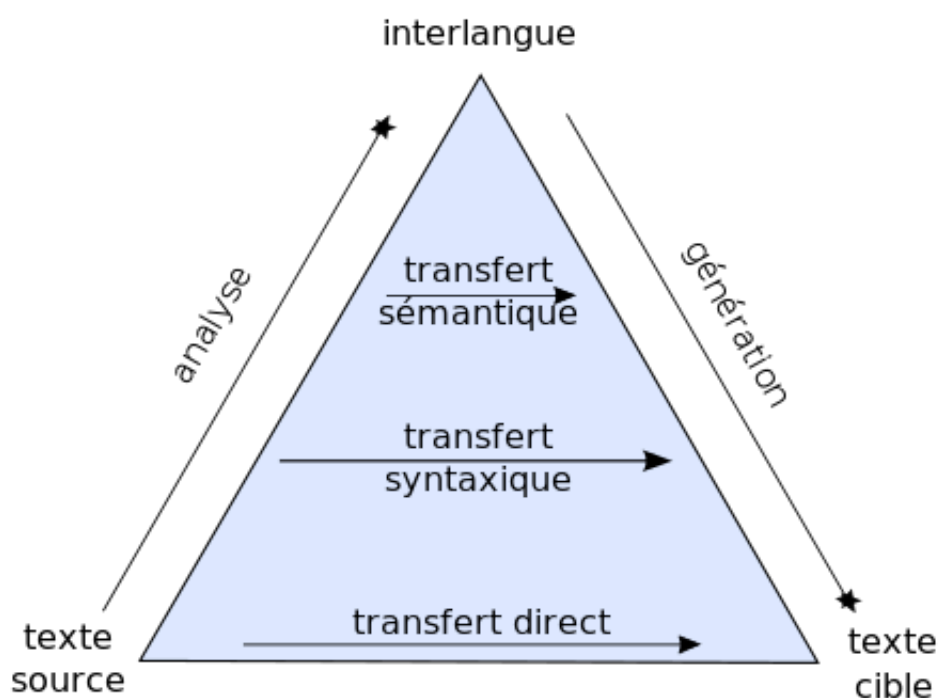


FIGURE 1 – Le triangle de Vauquois

bearded professor bite the bald professor » et le réajustement donnera en conclusion : « The bearded professor bite the bald professor ».

2.3 La traduction sémantique

La traduction sémantique consiste en l'amélioration de la traduction mot-à-mot à la fois grâce à une analyse syntaxique, mais aussi par un système de traduction plus intelligent. Le dictionnaire est enrichi par des informations complémentaires (fonction occupée par le mot, pronoms correspondants), ce qui va permettre au traducteur de trouver la solution la plus pertinente parmi plusieurs traductions possibles.

Pour commencer, le programme découpe la phrase en mots ou groupes de mots, et les compare à son dictionnaire. Les informations qui lui ont été ajoutées sont alors utilisées pour déterminer la fonction de chaque mot dans la phrase et fixer plusieurs sens possibles. Ceci fait, l'ordinateur conserve plusieurs traductions possibles. L'étape la plus importante va alors se dérouler. Dans chacune des traductions retenues, le programme calcule la pertinence de chaque traduction en fonction des mots l'entourant. Enfin le système génère de nouveaux pronoms et remet en place la phrase.

Par exemple, dans le cas de la distinction entre l'avocat en justice, et l'avocat en cuisine, il est courant d'avoir recours au phénomène de traduction sémantique. En effet, « avocado » sera associé au verbe manger tandis que « lawyer » à des cas plus sérieux.

2.4 La traduction statistique

La traduction statistique est plus un principe qu'une vraie méthode de traduction, puisqu'elle est très souvent couplée à une autre forme de traduction. Il s'agit ici d'analyser un énorme corpus de texte, et de relever des données statistiques pour déterminer les sens les plus probables de mots dans les phrases à traduire. Ainsi, chaque tournure est confrontée au corpus de texte, et celle qui obtient le meilleur score a le plus de chances d'être la bonne.

Il s'agit notamment de la solution retenue par Google pour son service de traduction : en confrontant une base de données énorme à la phrase de l'utilisateur, le géant du web comprend quelles sont les tournures les plus utilisées.

3 L'avenir de la traduction automatique

3.1 Les problèmes actuels

Bien que la traduction automatique se soit considérablement développée ces dernières années, devenant de plus en plus précise grâce à des techniques sophistiquées, des problèmes subsistent dans beaucoup de traductions.

Thierry Grass, dans l'article *À quoi sert encore la traduction automatique ?*, relève 13 problèmes persistants dans la traduction automatique. Il explique : « 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

»

Lors de traductions automatiques, on rencontre en effet des erreurs liées aux différents sens que peut avoir un mot en fonction de son contexte. Les logiciels de traduction automatique ne sont pas toujours capables d'analyser le contexte de la phrase pour déterminer quelle traduction convient le mieux pour le mot en question. Par exemple, le mot *kiwi* peut faire référence au fruit ou à l'oiseau, et la traduction de ce mot va différer selon son sens, que l'on détermine grâce au contexte.

Les ambiguïtés syntaxiques et référentielles correspondent à des erreurs liées aux différents sens que peut prendre une phrase si l'on rattache un mot (adjectif, préposition, adverbe, pronom) à un groupe ou à un autre. Par exemple, dans la phrase « J'ai regardé une vidéo de toi jouant du violon avec ta sœur », on ne sait pas si « avec ta sœur » fait référence à « j'ai regardé une vidéo (avec ta sœur) » ou à « tu jouais du violon (avec ta sœur) ». La traduction peut varier selon le sens de la phrase.

Les termes flous sont des mots que l'on rajoute dans une phrase pour « exprimer une approximation » selon Thierry Grass, tels que « en un sens », « d'ailleurs » ou « en fait » en français. Ces mots sont difficiles à traduire car propres à la langue, on ne peut pas les remplacer par une traduction littérale.

Il est également très difficile de traduire des expressions (idiotismes), qui diffèrent selon les langues. Par exemple, la traduction anglaise de « Il pleut comme vache qui pisse » est « It is raining cats and dogs », ce qui est très éloigné de la traduction mot à mot (il pleut des chiens et des chats).

De même, les mots d'origine étrangère et les emprunts, particulièrement présents

dans le domaine du business, sont souvent source d'erreurs lors des traductions, tout comme les noms propres (personnes, lieux...) qui changent selon les langues, ainsi que les sigles et acronymes. Par exemple, l'ONU en français (Organisation des Nations unies) devrait être traduit par UN (United Nations) en anglais. La néologie est la création de nouveaux mots, et pose problème pour les traducteurs automatiques, qui ne connaissent pas toujours les mots récemment créés dans une langue. Les séparateurs ne sont pas utilisés de la même façon dans toutes les langues. Par exemple, comme l'explique Thierry Grass, « 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 ». Ce point sera identifié comme un séparateur de phrases par beaucoup de systèmes de traduction automatique, ce qui induit une mauvaise traduction.

Les synonymes peuvent apporter différentes connotations à une phrase, mais elles ne sont pas toujours saisies par les traducteurs automatiques. L'emploi d'un mauvais synonyme peut engendrer un rendu maladroit et peu naturel à la phrase traduite.

La transposition, c'est-à-dire le fait qu'un mot appartenant à une classe soit traduit par un mot d'une autre classe (adjectif, nom, verbe...), peut également compliquer la tâche des traducteurs automatiques et induire de mauvaises traductions.

Enfin, un mot mal orthographié sera souvent mal traduit, voire pas du tout. Une faute d'orthographe qui remplacerait un mot par un homonyme (comme verre/vers) modifierait même le sens de la phrase.

3.2 Les améliorations possibles

Afin de palier à ces différents problèmes, plusieurs pistes sont à prendre en compte pour améliorer les systèmes de traduction automatique.

Pour empêcher les erreurs liées aux homonymes, aux ambiguïtés syntaxiques et référentielles, aux idiotismes, synonymes et transpositions, il serait préférable que le traducteur soit capable d'analyser le sens de la phrase, de la comprendre pour ensuite traduire cette signification dans la langue voulue. Les phrases seraient alors plus naturelles, et le sens du texte de départ serait conservé. Dans son article *La problématique de la traduction automatique*, Roland Raoul Kouassi écrit : « le traitement du sens et de la signification : ce point est l'un des plus importants de l'activité traduisante. Le traducteur traduit du sens et des significations. (...) Le logiciel doit être capable de calculer les sens – ce qui suppose la prise en compte non seulement des phénomènes de dénotation mais également les faits de connotation. ». Il est important de comprendre le sens de la phrase entière et non des mots séparés, comme le dit Thierry Grauss : « 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 ».

Cette compréhension du sens reposerait sur une base de connaissances très développée, utilisée par le logiciel de traduction. Dans le même article, Roland Raoul Kouassi explique « le logiciel doit être "hypercultivée" (d'où le développement des corpus et autres ressources lexicales générales et spécialisées). Cette culture doit être dynamique en suivant l'évolution des créations morphosyntaxiques, avec, par exemples des liens vers les bases de données et corpus les plus importants du web. ». Dans

sa thèse « Fondements épistémologiques de l’ethnométhodologie » (partie 2, chapitre 5), Mel Vadeker explique que la somme des connaissances nécessaires pour comprendre les deux cultures concernées par la traduction est évidemment énorme, et dépasse les capacités des machines actuelles, en termes de mémoire et de gestion d’une quantité d’informations aussi importante. De plus, « supposons cette difficulté résolue, il faudrait encore trouver le moyen de réunir l’ensemble des connaissances nécessaires », écrit-il.

Les connaissances nécessaires à la compréhension d’un texte sont d’autant plus nombreuses que ce texte est général et semble facile à comprendre à un être humain. Comme l’a écrit Mel Vadeker dans sa thèse, « contrairement à ce qu’on pourrait croire, ce ne sont pas les textes les plus pointus qui demandent le plus de connaissances. Celles qui sont requises sont certes plus précises, mais finalement assez peu nombreuses. (...) [L]a traduction de domaines très pointus demande paradoxalement moins de connaissance que celle de la vie de tous les jours. D’ailleurs, les traducteurs le savent bien le style académique ou technique d’une revue scientifique est incomparablement plus simple à traduire que le style parlé d’un journal comme Libération. » Aussi, les logiciels de traduction automatique spécialisés dans un domaine fonctionnent mieux que les logiciels de traduction générale, puisque les connaissances requises à ces traductions sont moins nombreuses et donc plus faciles à réunir, stocker et utiliser.

Une traduction générale idéale, basée sur le sens des mots, ne semble donc pas réalisable dans un futur immédiat, et nécessiterait une évolution technologique importante.

3.3 La traduction automatique dans le futur

3.3.1 Les nouvelles utilisations

Une des nouvelles utilisations de traduction automatique envisagée par Google est la traduction vocale en temps réel. Selon un article du Monde, *La montée en puissance de la traduction automatique*, rédigé par Damien Leloup en 2010, ce service serait proposé aux utilisateurs d’ici à 2015. Cette évolution serait également basée sur un système de reconnaissance et synthèse vocale. Dans le même article, Marie Candito explique « Le principal problème, c’est que la reconnaissance vocale elle-même est source d’ambiguïtés, même si, dans certains cas, la traduction peut aider à les résoudre : confronté à une ambiguïté lors de l’étape de reconnaissance vocale, le système peut conserver les deux options, et trancher ensuite lors de la traduction. ». Si la difficulté est surmontée, deux personnes parlant deux langues différentes pourraient communiquer par téléphone et se comprendre. Dans un article plus récent *Bientôt un traducteur vocal instantané pour Smartphone signé Google ?*, rédigé par Wassina Zirar en juillet 2013 pour l’Usine digitale, Hugo Barra, le vice président d’Android, confirme cette hypothèse en affirmant que des prototypes ont déjà été réalisés.

On pourrait également envisager d’utiliser les systèmes de traduction automatique pour traduire les fichiers de sous-titres de vidéo et en générer un autre dans la langue souhaitée. Associée à un système de reconnaissance vocale, la traduction automatique

pourrait même permettre d'ajouter directement des sous-titres à une vidéo en version originale.

3.3.2 Les traducteurs

Devant l'avenir apparemment radieux de la traduction automatique, on est en droit de se demander quel sera l'avenir des traducteurs professionnels. Il semble cependant que les logiciels de traduction automatique ne soient pas une menace pour cette profession, au contraire. Ils sont en effet utilisés par beaucoup de traducteurs professionnels car ils proposent des traductions analogues à celles effectuées auparavant. Le traducteur utilise ces propositions puis modifie et homogénéise le texte. Ce système permet aux traducteurs de réduire leurs délais, ce qui diminue également le coût des traductions.

Les traducteurs humains ne sont pour l'instant pas remplacés par des traducteurs automatiques à cause des erreurs que nous avons cité précédemment, qui peuvent nuire à l'image des entreprises utilisant ces textes, et parce qu'ils ne prennent pas en compte les différences culturelles entre les deux pays. Un traducteur humain peut tourner les phrases d'une manière qui semblera correcte à celui qui écoute, ce qu'un traducteur automatique ne sait pas faire et ne fera probablement jamais. Le traducteur humain aura souvent un rôle négociateur et social que la traduction automatique ne pourra jamais endosser. Il semble donc que la traduction automatique, qui a devant elle un bel avenir, ne menace pas pour autant celui des traducteurs humains qui restent irremplaçables de part leur humanité, justement.

Conclusion

La traduction automatique a connu des débuts mouvementés. Elle est née sous la forme de dictionnaires automatiques créés lors des recherches pour la guerre froide. Les recherches qui ont suivi ont été arrêtées suite à un rapport de l'ALPAC, très négatif quant à l'avenir de la traduction automatique, qui paraissait alors bien sombre. C'est le Japon qui relancera la traduction automatique dans les années 1980, suite à d'importants progrès technologiques, suivi par les Etats-Unis dans les années 1990. Les recherches qui sont alors menées vont considérablement améliorer la traduction automatique.

De ces recherches sont nés plusieurs types de traductions, utilisés aujourd'hui, souvent basés sur l'utilisation d'un dictionnaire de mots proposant plusieurs traductions par mot. La méthode la plus basique, la traduction mot à mot, n'est pas la plus efficace puisqu'elle utilise simplement un dictionnaire de mot après avoir effectué une lemmatisation, sans prendre en compte le sens des mots. La traduction sémantique utilise un dictionnaire plus riche, contenant plus d'informations sur chaque mot, et réalise une analyse plus poussée en choisissant la traduction la plus pertinente. La traduction statistique choisit pour chaque mot la traduction la plus utilisée dans le contexte. Différents types de traduction sont souvent associés dans les logiciels de traduction automatique actuels.

Ces traductions, plus pertinentes qu'auparavant, ne sont cependant pas sans failles. On relève différents types d'erreurs plus ou moins fréquents lors des traductions automatiques. Ces problèmes pourraient être résolus dans un avenir plus ou moins proche grâce à une avancée technologique qui permettrait d'utiliser plus de connaissances pour comprendre le sens des phrases. Malgré ces erreurs, de nouvelles applications sont envisagées, comme la traduction vocale en temps réel. On peut donc affirmer que l'ALPAC avait tort quant à l'avenir de la traduction automatique, qui paraissait aujourd'hui radieux. Il semble cependant que les traducteurs professionnels, bel et bien humains, n'aient pas à s'en faire pour le leur, leurs qualités humaines ne pouvant pas être simulées par une machine.

Bibliographie

Références

- [1] THIERRY GRASS « *À quoi sert encore la traduction automatique ?* » <http://www.cahiersdugpepe.fr/index1367.php#tocto1n2> 02/09/2013
- [2] ROLAND RAOUL KOUASSI « *La problématique de la traduction automatique* » http://www.ltml.ci/files/articles4/article_traduction_automatique.pdf 02/09/2013
- [3] DAMIEN LELOUP « *La montée en puissance de la traduction automatique* » *Le Monde* http://www.lemonde.fr/technologies/article/2010/02/23/la-montee-en-puissance-de-la-traduction-automatique_1309888_651865.html 02/09/2013
- [4] WASSINA ZIRAR « *Bientôt un traducteur vocal instantané pour smartphone signé Google ?* » http://www.usine-digitale.fr/article/bientot-un-traducteur-vocal-instantane-pour-smartphone-signe-google_N202319 03/09/2013
- [5] YVES DROTHIER « *Au coeur de la traduction automatique de Systran* » <http://www.journaldunet.com/solutions/intranet-extranet/reportage/au-coeur-de-la-traduction-automatique-de-systran/evolutions-futures-videos-sons-et-synthese-vocale.shtml> 03/09/2013
- [6] DAMIEN LELOUP « *La montée en puissance de la traduction automatique* » http://www.lemonde.fr/technologies/article/2010/02/23/la-montee-en-puissance-de-la-traduction-automatique_1309888_651865.html
- [7] XAVIER BISEUL « *Dimitris Sabatakakis (Systran) : “Pendant sept ans, la Commission européenne a contesté la propriété de notre logiciel !”* »
- [8] MEL VDEKER « *Fondements épistologiques de l’ethnométhodologie* » <http://www.vadeker.net/corpus/loubiere/2-5.html> 02/09/2013
- [9] MAURICE GROSS « *Notes sur l’histoire de la traduction automatique* » http://www.persee.fr/web/revues/home/prescript/article/lgge_0458-726x_1972_num_7_28_2096
- [10] VÉRONIQUE GENDNER « *Analyse de données linguistiques pour le retour d’expérience Partie 3 : Le problème des ambiguïtés pour l’analyse de données textuelles* » http://archives.limsi.fr/Individu/gendner/analyse_texte/2002/03-ambiguite.html
- [11] DELPHINE GRAEFF, ÉDOUARD LOPEZ « *Systran vs. Google* »

- [12] JACQUELINE LÉON « *Le CNRS et les débuts de la traduction automatique en France* »
<http://histoire-cnrs.revues.org/3461>
- [13] Google Traduction http://translate.google.fr/about/intl/fr_ALL/
- [14] Systran <http://www.systran.fr/systran/entreprise>
- [15] Wikipédia - Traduction Automatique https://fr.wikipedia.org/wiki/Traduction_automatique
- [16] Linatec <http://www.linatec.fr/products/tr/information/technology/mtranslation>
- [17] « *Linguistique et informatique : la traduction automatique* » http://www.canal-u.tv/video/universite_de_tous_les_savoirs/linguistique_et_informatique_la_traduction_automatique.872