

I.1. Introduction

La traduction automatique est une traduction effectuée par ordinateur, sans intervention humaine. Les débuts de la traduction automatique, également baptisée traduction par machine ou traduction instantanée, remontent aux années 50, elle est considérée parmi les applications les plus connues de l'intelligence artificielle, le processus consiste en générale de faire passer un texte rédigé dans une langue (« langue source », ou « langue de départ ») à un texte rédigé dans une autre langue (« langue cible », ou « langue d'arrivée »).

Dans ce chapitre nous présentons quelques notions de base sur L'intelligence artificielle IA et traduction automatique ainsi que quelques traitements opérants sur celles-ci

I.2. L'intelligence artificielle IA [int]

L'intelligence Artificielle est le discipline scientifique relative au traitement des connaissances et au raisonnement, dans le but de permettre à une machine d'exécuter des fonctions normalement associées à l'intelligence humaine : compréhension, raisonnement, dialogue, adaptation, apprentissage, etc.

I.2.1. Autre définition [intl]

Le terme « intelligence artificielle », créé par John McCarthy, est souvent abrégé par le sigle « IA » (ou « AI » en anglais, pour Artificial Intelligence). Il est défini par l'un de ses créateurs, Marvin Lee Minsky, comme « la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique »

L'Intelligence Artificielle (IA) est la science dont le but est de faire par une machine des tâches que l'homme accomplit en utilisant son intelligence. La terminologie - malheureuse! -d'Intelligence Artificielle est apparue en 1956. On

peut lui préférer celle d'Informatique Heuristique. On ne parlera pas dans ce cours de machine intelligente, ni de programme intelligent.

Autre définition (J.L.Laurière): Etude des activités intellectuelles de l'homme pour lesquelles aucune méthode n'est a priori connue. (Tout ce qui n'a pas encore été fait en Informatique -quand on sait le faire, ce n'est plus de l'IA ..)

L'Informatique est la science du traitement de l'Information, l'IA s'intéresse à tous les cas où ce traitement ne peut être ramené à une méthode simple, précise, algorithmique. Un algorithme est une suite d'opérations ordonnées, bien définies, exécutables sur un ordinateur actuel, et qui permet d'arriver à la solution en un temps raisonnable (minutes, heures, ou plus,... mais pas des siècles!)

I.3. les domaines d'application de l'IA

Aujourd'hui l'intelligence artificielle regorge de nombreuses capacités et de tâches possibles comme:

- ♦ **les systèmes d'experts:** L'IA peut réaliser la même tâche qu'une personne humaine, notamment dans la robotique (ex: logiciel MYCIN en 1974 qui donne des diagnostics)
- ♦ **le calcul formel:** Cela permet d'exécuter de nombreuses formules symboliques (ex: MATHEMATICA)
- ♦ **la représentation des connaissances:** Pour qu'un robot puisse manipuler des objets, on doit lui les représenter.
- ♦ **le traitement du langage et la résolution de problèmes:** Dans les logiciels de **traduction automatique**, même si ils ne sont pas tous performants car c'est une tâche très délicate. Quant à la résolution de problèmes, on trouve souvent l'IA dans les jeux vidéos.
- ♦ **la reconnaissance:** Qu'elle soit de la parole, de l'écriture ou du visage, l'IA a fait de nombreux progrès dans ces actions.

- ♦ **la robotique:** Les robots se trouvent dans de nombreuses exploitations, comme l'industrie, la médecine, les loisirs, l'exploitation spatiale.
- ♦ **l'apprentissage:** Pour qu'une intelligence artificielle puisse faire son devoir elle doit apprendre les notions données par l'Homme.
- ♦ **les systèmes complexes:** On retrouve cela particulièrement dans les algorithmes, et les modes de vies artificiels.

L'intelligence artificielle déborde de capacités qui se retrouvent dans des domaines du travail comme:

- **la banque:** On retrouve aujourd'hui des systèmes d'évaluation aux risques de vols de crédits
- **les jeux vidéo:** Du simple et ludique morpion, aux jeux vidéo plus perfectionnés l'intelligence artificielle possède une place importante dans les jeux vidéo. Car il y a une interaction entre le joueur humain et "l'ordinateur". Celui-ci doit essayer de jouer aussi bien que le joueur humain et même essayer de le battre, selon la difficulté choisie par le joueur humain.
- **la médecine, l'industrie, l'armée, les loisirs:** Aujourd'hui avec tous les moyens que les inventeurs possèdent, la robotique est en pleine évolution. Ils remplacent les hommes dans des tâches répétitives (ex: travail à la chaîne en usine, particulièrement dans l'automobile. Car ces entreprises souhaitent faire du profit), les professions de la santé, les robots ludiques
- **l'informatique:** Notamment par les moteurs de recherche: Une personne qui recherche une information sur internet va utiliser le moteur de recherche. Ensuite celui-ci donnera l'ensemble des sites qui se rapportent à l'information recherchée par la personne. Les plus célèbres étant Google, Bing...

On retrouve aussi **la traduction automatique:** une personne souhaite traduire un mot, une phrase d'une langue à une autre utiliser la traduction automatique. Les plus populaires sont: Google traduction et Reverso. Ils sont très pratiques pour toutes les personnes, car on a la possibilité de traduire des phrases.

Ce qui est impossible avec un dictionnaire qui ne donne que des traductions de mots.

I.4. La traduction automatique [sys]

L'expression « traduction automatique » désigne au sens strict la traduction d'un texte (ou d'une conversation audio, en direct ou en différé) faite entièrement grâce à un ou plusieurs programmes informatiques, sans qu'un traducteur humain n'ait à intervenir. On englobe parfois également sous ce terme la traduction assistée par ordinateur où la traduction peut se faire en partie manuellement ou de façon interactive avec la machine.

Malgré ses faiblesses, elle peut rendre des services dans des domaines tels que la veille internationale (Veille technologique y compris), où elle permet de prendre connaissance de façon superficielle, mais rapide, de grandes quantités de textes.

Depuis quelques années, elle connaît un essor considérable sur le Web, avec plusieurs systèmes en ligne pouvant traduire automatiquement et en quelques secondes des pages Web ou des textes de plus en plus longs. C'est une aide très appréciée du grand public car elle permet de déchiffrer - de façon grossière - le thème d'une page Web dans une langue totalement inconnue, et les principaux faits ou éléments d'information qu'elle contient. Pour simplifier la navigation, plusieurs moteurs de recherche comme Google, Altavista ou Yahoo! permettent de l'utiliser.

dans le domaine de traduction automatique il existe deux grands pôles (méthodes) :

I.4.1. Traduction Automatique à base de règles TABR

Traduction automatique classique qui comprend des règles faites par des linguistes dans des grammaires. Exemples de systèmes : Systran, Reverso, Prompt et LOGOS. **[sys]**

I.4.2. Traduction automatique statistique TAS

Traduction automatique qui n'utilise que des statistiques, contrairement à la traduction à base de règles. Exemples de systèmes : LanguageWeaver, PORTAGE et MOSES. **[sys]**

I.4.3. Comparaison entre les méthodes

TABR	TAS
Traductions prévisibles et homogènes	Qualité de traduction inégale
Bonne qualité des traductions généralistes	Qualité faible sur les domaines généralistes
Connaissance des règles grammaticales	Pas de connaissance de la grammaire
Performance et robustesse	Besoins en CPU et espace disque importants
Cohérence entre les versions	Pas de cohérence entre les versions
Manque de fluidité	Traductions fluides
Difficulté à gérer les exceptions	Reconnaît bien les exceptions
Coût de développement et de personnalisation élevés	Développement rapide et peu onéreux si des corpus de qualité sont disponibles

Table 1.1 Comparaison entre les méthodes et motivation de choix [sys]

I.5. À quels domaines la traduction automatique peut-elle être appliquée ?

L'utilisation d'un programme de traduction est judicieuse lorsque vous êtes souvent amenés à traduire de grandes quantités de texte ou à travailler avec des textes en langue étrangère que vous souhaitez comprendre rapidement.

Les domaines d'application les plus courants de la traduction automatique sont les suivants :

- ❖ La communication internationale au moyen des e-mails ou de la correspondance commerciale s'effectue bien plus rapidement grâce à la traduction automatique.
- ❖ Les documents déjà existants peuvent être rapidement rendus disponibles aux filiales étrangères ou aux partenaires commerciaux dans leur propre langue.
- ❖ La traduction de textes techniques, par ex. les documentations, manuels, modes d'emploi, etc. Avant tout, la traduction automatique s'avère très efficace pour les textes largement standardisés (grâce notamment à l'utilisation de mémoires de traduction).
- ❖ Traduction de pages Web pour comprendre le contenu (compréhension écrite).
- ❖ Traduction de pages Web pour les mettre à disposition sur Internet en plusieurs langues.
- ❖ Utilisation de la traduction automatique dans les chats en ligne et ventes aux enchères en temps réel.
- ❖ La traduction automatique est moins bien adaptée pour la traduction littéraire de romans et de poésies.

I.6. Description Le processus de traduction [tra]

Le processus de traduction (au sens humain) peut être découpé en trois phases successives :

- ♦ **compréhension** : assimilation du sens véhiculé par un texte, du vouloir dire d'un auteur... ;
- ♦ **déverbalisation** : oubli des mots et conservation du sens ; « Opération par laquelle un sujet prend conscience du sens d'un message en perdant conscience des mots et des phrases qui lui ont donné corps » ;
- ♦ **réexpression** : reformulation du vouloir dire en langue cible.

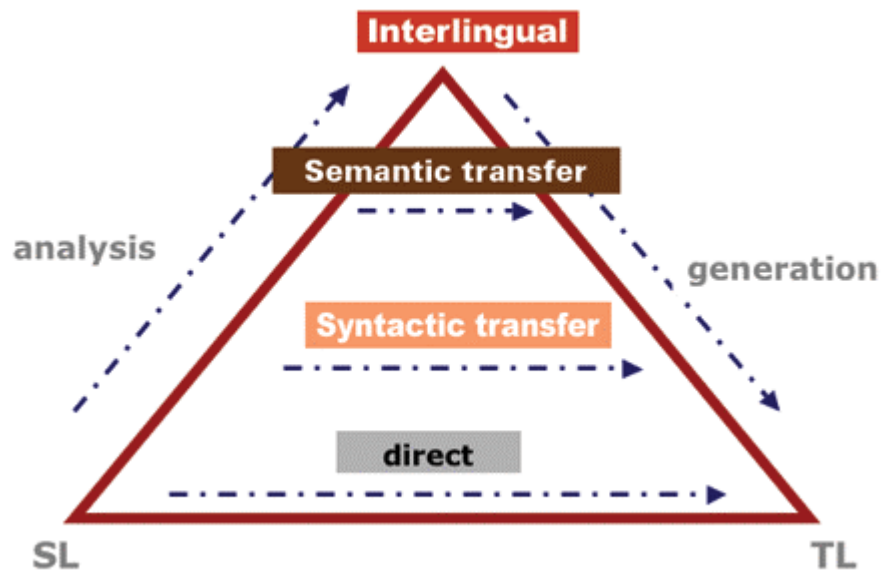


Figure 1.1 Triangle de Vauquois, modèle pour les fondements de la traduction automatique

En termes informatiques, la compréhension devient l'analyse, la déverbalisation devient le transfert et la réexpression devient la génération. Ces étapes du processus sont modélisées dans le triangle de Vauquois. Ce modèle est utile car pour passer de la source à la cible, il existe plusieurs chemins possibles qui constituent les différentes approches envisagées à ce jour. Plus le degré de conceptualisation est élevé, plus court est le chemin du transfert. Nous avons quatre possibilités principales

Le transfert direct: pas de conceptualisation, toute la traduction repose sur le transfert. La traduction par l'exemple et la traduction statistique travaillent à ce niveau. La traduction est vue comme un processus de décodage.

Le transfert syntaxique: le niveau du transfert est syntaxique. Généralement, sa représentation est l'arbre syntaxique. L'analyse produit une représentation syntaxique pour la langue source. Le transfert consiste à produire une représentation syntaxique pour la langue cible à partir de ce dernier. Finalement la génération produit la phrase source. La traduction automatique à

base de règles est représentative de cette catégorie. Les règles permettent les différentes transformations.

Le transfert sémantique: le niveau du transfert est sémantique. Cette voie est celle qu'empruntent les humains. Les modèles de représentation de la sémantique du langage sont décrits par la Pragmatique (linguistique). La sémantique peut être décrite par une Ontologie (informatique). Il n'y a peu d'approches de la traduction automatique représentatives du transfert sémantique.

L'interlangue: Ce niveau supprime la nécessité de transfert. L'Interlangue (traduction automatique) devient universelle. Et seuls restent les processus d'analyse et de génération. l'interlangue est aussi désignée sous le terme de langue pivot. DLT est une tentative inachevée de cette approche. Le langage UNL est aussi un exemple de langage formel informatique permettant de représenter le sens d'un énoncé. L'approche est séduisante car l'effort consiste, pour une langue donnée, à produire un analyseur et un générateur pour l'interlangue. On bénéficie alors de toutes les traductions de ou vers les langues possédant aussi l'analyseur et le générateur. Cette approche reste difficile et n'a pas connu de succès à large échelle.

Actuellement, les moteurs de traduction sont principalement par règles ou statistiques. Une voie dite hybride émerge. Systran, Google Translate et Bing Translator utilisent des approches hybrides.

I.7. System de traducteur en ligne[Atlasia]

Il existe plusieurs system de traduction en ligne permis elle :

Google Traduction, Reverso, Bing, Prompt Translator, Systran, Babylon, Voila Linguattec, Worldlingo, InterTran.



Figure 1.2 System de traducteur en ligne

Apartire de test 2014 « Quel est le meilleur traducteur en ligne gratuit ? » de site ALTISSIA le résultat comme suivent :

ALTISSIA	ADEQUATION DE LA TRADUCTION	CLARITE -SIMPLICITE VISUELLE	VITESSE DE TRADUCTION	POSSIBILITE DE TRADUCTION HUMAINE PAYANTE	POSSIBILITE DE SONORISATION DE LA PHRASE	CORRECTEUR ORTHOGRAPHIQUE DU TEXTE SOURCE	DETECTION DE LANGAGE	POSSIBILITE DE FEEDBACK
Reverso	3.5 /5	⚠	✓	✗	✓	✓	✗	✓
Google translate	4.0 /5	✓	✓✓	✗	✓	✓	✓	✓
SYSTRAN	2.5 /5	✓	✓	✗	✗	✓	✗	✗
babylon	2.5 /5	⚠	⚠	✓	✗	✓	✗	✗
voila	2.5 /5	⚠	✓	✗	✗	✓	✗	✗
WorldLingo	1.5 /5	✓	⚠	✓	✗	✓	✓	✗
linguattec	2.0 /5	⚠	✓	✗	✓	✗	✗	✗
bing	3.0 /5	✓	✓	✗	✓	✓	✓	✓
InterTran	0.5 /5	✓	⚠	✗	✗	✓	✗	✗
PROMT TRANSLATOR	2.5 /5	✓	⚠	✓	✗	✓	✗	✗

SOURCES: ALTISSIA.COM

Table 1.2 comparison entre les system de traduction.

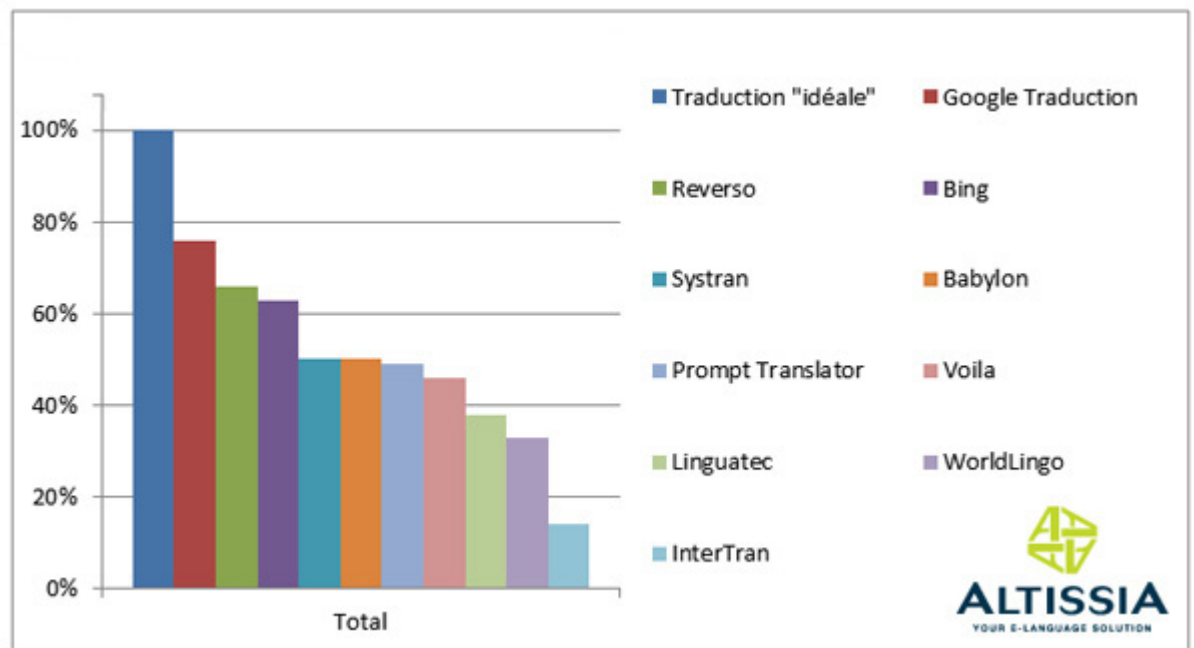


Figure 1.3 comparaison entre les system de traduction

I.8. Avantages de la Traduction Automatique

Les avantages de la traduction automatique varient de la capacité d'offrir au traducteur un bon point de départ pour la traduction à la traduction instantanée d'un document pour une compréhension essentielle. Vous trouverez ci-dessous une liste des principaux avantages de la traduction automatique :

- Procure l'essentiel d'un document ou une traduction « assez bonne »
- Améliorer la communication multilingue
- Publier en plusieurs langues
- Offre la possibilité de passer de la traduction à la postédition – diminuant ainsi vos frais de traduction
- Se libérer des contraintes de capacité et traduire plus
- Propose plus de suggestions de segments traduits aux traducteurs
- Permet à votre organisation de traduire des documents que vous n'auriez pas traduits par le passé

- Élimine le partage de votre contenu dans Google Translate
- Réduire les coûts et les délais de la traduction humaine
- Constituer et valoriser le patrimoine linguistique des entreprises
- Rendre les applications d'entreprise multilingues

I.9. Les difficultés pour la TA

Ils exist plusieurs problems de traduction automatique permit elles

❖ **Ambiguïtés de la langue source**

- **Ambiguïtés lexicales** : polysémie. Dans Le défi des Langues, Claude Piron raconte que le 4 mai 1991 il avait demandé à une machine de traduire: In such a case, you can make a very good case for wooden cases, c'est-à-dire : « En pareil cas, on peut produire une argumentation très convaincante en faveur des caisses en bois », phrase qu'il avait déjà lue et simplement condensée. Au bout d'un instant vint la réponse: « Dans un tel cas vous pouvez faire un très bon cas pour des cas inexpressifs ».
- **Ambiguïtés grammaticales** : un même mot peut relever de deux catégories grammaticales différentes. Toujours dans Le Défi des Langues l'auteur proposait à une machine, mise en mode interactif, He was sorting out food rations and chewing gum, phrase qui peut signifier soit « il triait des rations et du chewing gum », soit « il triait des rations tout en mâchant du chewing-gum ». Au lieu de poser une question sur la valeur grammaticale du -ing dans chewing, au lieu donc d'« interagir », la machine livra la phrase suivante: « Il triait dehors rations de nourriture et mastiquant la gencive.»
- **Ambiguïtés sémantiques** : un même mot peut avoir plusieurs sens dans les contextes différents. Par exemple: «saw» qui peut signifier soit «voir», soit «scie».
- **Ambiguïté de référence des pronoms** : pour la phrase « Elle l'aime » qui peut signifier soit « She likes him/it »,soit « It likes him/it », etc.

qu'il dispose, en plus de connaissances linguistiques, de connaissances factuelles sur l'état du monde (par exemple, il doit savoir que telle personne est de sexe féminin) ;

qu'il soit capable de réaliser une certaine interprétation du texte qui lui est soumis : s'il rencontre le mot « secrétaire », il a besoin de savoir, selon la langue-cible, le sexe de ce/cette secrétaire ou de son employeur. Il est facile de résoudre cette difficulté dans le cas d'une phrase simple comme « John Doe's secretary » : une règle simple permet d'exprimer le fait que l'expression « X's Y » implique que Y "appartient" à X. Mais si le programme rencontre la phrase « je suis allé au service de la comptabilité et j'ai parlé au secrétaire », il est plus difficile pour la machine de faire la relation entre les secrétaires et le service qui vient d'être mentionné ; la tâche est encore plus ardue dans un contexte comme « j'ai vu M. Tartempion, qui m'a donné un document ; je suis allé au service de la comptabilité pour qu'ils en fassent une photocopie et je suis revenu le donner à la secrétaire » : ici, on ne peut même pas s'appuyer sur la proximité entre le mot « secrétaire » et le nom du service, puisqu'il est en fait question de la secrétaire de Tartempion, mentionnée bien avant dans la phrase.

I.10. Conclusion

Dans ce chapitre on a présenté les définitions de quelques conceptions et notions sur l'intelligence Artificielle. IA et Traduction Automatique qui sont nécessaires pour aborder les prochaines étapes de Traduction Automatique