

Introduction générale

Le langage naturel est le mode privilégié par les humains pour communiquer entre eux, de manière parlée, signée ou écrite, ayant évolué depuis des dizaines de milliers d'années. Cette évolution a donné lieu à certains groupes de langues existantes aujourd'hui, ou ayant existé auparavant, gardant une certaine proximité provenant de leur origine commune. Toutefois, d'autres langues ont leurs origines encore incertaines, donnant lieu à de nombreuses questions et débats sur l'origine des langues. Selon Ferdinand de Saussure, "C'est une idée très fausse que de croire que le problème de l'origine du langage soit un autre problème que celui de ses transformations"

De part son besoin de communiquer, l'homme essaye naturellement de comprendre d'autres langues, celles qu'il ne connaît pas. Aujourd'hui, les avancées technologiques et la communication globalisée accentuent ce besoin d'accéder à des données, pouvant être multimédia, dans des langues que nous ne maîtrisons pas. Une solution possible pour prendre en charge la demande grandissante en traduction est l'automatisation de cette tâche au moyen des ordinateurs. Cette idée de traduction "mécanique" (Mechanical Translation) remonte aux prémices de l'informatique. Les recherches scientifiques en traitement de l'information étaient alors majoritairement orientées vers la cryptographie, donnant lieu à l'analogie avec la traduction automatique, pouvant être vu comme un processus de décryptage.

Mais il paraît encore impossible aujourd'hui de traduire parfaitement des textes très longs et génériques, comme des romans par exemple. L'utilisabilité des traductions dépend donc du besoin et de l'usage fait des hypothèses générées automatiquement par les systèmes.

Il existe de nombreux outils de traduction mis à disposition, que ce soit des logiciels (développé par Systran par exemple), ou des interfaces accessibles

en ligne (Google Translate , Babel Fish , Systran , etc.). Des industriels s'intéressent donc à la traduction automatique depuis de nombreuses années et proposent des solutions efficaces selon des contextes particuliers. Systran, IBM, ou encore Xerox, sont des acteurs importants dans le développement et la commercialisation d'outils de traduction automatique. De plus, la communauté scientifique travaille depuis plusieurs années sur des boîtes à outil permettant de construire un système de traduction.

La traduction est une tâche complexe. Traduire ne consiste pas simplement à prendre chaque mot d'une phrase et à le remplacer par sa traduction dans une autre langue. Il faut faire face à de nombreuses difficultés comme les ambiguïtés que peuvent comporter les composants d'une phrase. Un même mot, par exemple, peut appartenir à plusieurs catégories grammaticales, il est alors question d'ambiguïté grammaticale. Le mot français souris en est une parfaite illustration. Il peut à la fois être un nom ou alors un verbe et il est nécessaire de connaître sa catégorie pour le traduire correctement. D'autres mots encore, peuvent avoir plusieurs sens, et l'ambiguïté sera alors dite sémantique

Notre objectif est d'étudier et d'implémenter Un système intelligent pour la traduction automatique du texte.

Pour réaliser ce but, on a suivi les étapes décrites par les chapitres suivants :

Le premier chapitre, Introduction à la traduction automatique.

Le deuxième chapitre, présente La traduction à base de règle et la différente étape de La traduction.

Et Le dernier chapitre, présente notre Architecture et fonctionnement de système.

Finalement, nous clôturons ce mémoire par une conclusion générale.