

Dédicace

Je dédie ce travail qui n'aura jamais pu voir le jour sans les soutiens indéfectibles et sans limite de mes chers parents qui ne cessent de me donner avec amour le nécessaire pour que je puisse arriver à ce que je suis aujourd'hui. Que dieux les protège et que la réussite soit toujours à ma portée pour que je puisse leur combler de bonheur.

Je dédie aussi ce travail à :

Ma sœur, mes frères, ma famille, tous mes amis, mes collègues et tous ceux qui m'estiment.

Remerciement

Je remercie Allah premièrement.

Je dois remercier mon encadreur, [Dr. Bouchana Belkacem](#), qui m'a aidé à finir mon travail.

Je remercie tous ceux qui m'ont aidé, en particulier ma mère, mon père, ma famille.

Enfin et surtout, j'adresse mes salutations sincères aux membres du jury qui évalueront mon travail.

Résumé

Pour la crédibilité de la traduction, il faut strictement ajuster la grammaire des deux langues à traduire ainsi que le contexte.

Dans ce projet, nous créons un interprète qui va traduire le texte (une lettre à un médecin, une analyse médicale) du français vers l'arabe dans le domaine médical, nous choisissons un dictionnaire médical et une méthode de la traduction automatique à base des règles en mode transfert.

Le traducteur effectue quelques opérations sur le texte à traduire en analysant grammaticalement ce dernier, par l'extraction et la sélection des éléments grammaticaux, puis la conversion à l'aide d'un dictionnaire médical et en générant le texte cible à l'aide la grammaire de deux langues.

En premier lieu de ce projet, nous avons commencé par une présentation de l'état de l'art de traduction automatique et nous avons donné quelques travaux existants, ensuite nous avons expliqué brièvement la notion du contexte et de la grammaire .Par la suite, nous avons entamé la conception de notre système .et enfin nous avons réalisé et implémenter notre système.

ملخص

للمصداقية في الترجمة يجب ضبط قواعد اللغة لكلا اللغتين المراد الترجمة بينها و السياق المراد ضبطا دقيقا.

في هذا المشروع سنقوم بصناعة مترجم يقوم بترجمة نص (رسالة الى طبيب مختص و تحاليل طبية) من لغة الفرنسية الى لغة العربية في مجال الطب , عن طريق اختيار قاموس طبي و طريقة الترجمة الالية المعتمدة على القواعد و بالتحديد التحويلية. اذ يقوم المترجم بإجراء بعض العمليات على نص لترجمته حيث يقوم بتحليل النص تحليلا نحويا باستخراج و تحديد العناصر النحوية للنص و من بعدها تحويله دلاليا باستخدام المعجم الطبي و توليده باستخدام قواعد اللغة لكلا اللغتين .

نبدأ هذا المشروع بمقدمة عن الترجمة الالية و بعض الأعمال موجودة , و شرحنا باختصار لكل من قواعد لغتين الفرنسية و العربية ومصطلح السياق ,وبعد ذلك صممناه. و أخيرا قمنا بتحقيقه و تنفيذه.

Table des matières

Introduction généraleErreur ! Signet non défini.

Chapitre 01 Traduction automatique

1.1 Introduction	Erreur ! Signet non défini.
1.2 Définitions.....	Erreur ! Signet non défini.
1.2.1 Le Traitements Automatique des Langues (TAL).....	Erreur ! Signet non défini.
1.2.2 La traduction automatique (TA).....	Erreur ! Signet non défini.
1.2.3 La linguistique	Erreur ! Signet non défini.
1.2.4 Le langage.....	Erreur ! Signet non défini.
1.2.5 Une langue.....	Erreur ! Signet non défini.
1.2.6 La linguistique informatique	Erreur ! Signet non défini.
1.2.7 La grammaire.....	Erreur ! Signet non défini.
1.2.8 Un dictionnaire	Erreur ! Signet non défini.
1.2.9 Vocabulaire.....	Erreur ! Signet non défini.
1.2.10 Intelligence artificielle.....	Erreur ! Signet non défini.
1.3 Historique.....	Erreur ! Signet non défini.
1.3.1 Les années 50-60	Erreur ! Signet non défini.
1.3.2 Le rapport ALPAC et ses conséquences.....	Erreur ! Signet non défini.
1.3.3. Entre les années 1967-1976.....	Erreur ! Signet non défini.
1.3.4 Les années 80.....	Erreur ! Signet non défini.
1.3.5 Le développement des aides à la traduction	Erreur ! Signet non défini.

- 1.3.6 Des années 90 à aujourd'hui.....**Erreur ! Signet non défini.**
- 1.4. Intelligence artificielle et la traduction automatique..... **Erreur ! Signet non défini.**
- 1.5 Les approches de traduction automatique**Erreur ! Signet non défini.**
 - 1.5.1 L'approche à base de règles**Erreur ! Signet non défini.**
 - 1.5.2 L'approche statistique**Erreur ! Signet non défini.**
 - 1.5.3 L'approche à base d'exemples**Erreur ! Signet non défini.**
 - 1.5.4 L'approche hybride**Erreur ! Signet non défini.**
- 1.6 Échantillons Des Logiciels De Traduction Automatique.... **Erreur ! Signet non défini.**
- 1.7 Quels usages et limites ?**Erreur ! Signet non défini.**
- 1.8 La traduction automatique comme outil automatique..... **Erreur ! Signet non défini.**
- 1.9 Personnalisation du système par la création d'un dictionnaire spécialisé**Erreur ! Signet non défini.**
- 1.10 Avantages de la personnalisation**Erreur ! Signet non défini.**
- 1.12 Conclusion.....**Erreur ! Signet non défini.**

Chapitre 02 Grammaire et Contexte

- 2.1 Introduction**Erreur ! Signet non défini.**
- 2.2 La grammaire**Erreur ! Signet non défini.**
 - 2.2.1 Définition.....**Erreur ! Signet non défini.**
 - 2.2.2 La grammaire de la langue française**Erreur ! Signet non défini.**
 - 2.2.2.1 Étapes de l'analyse d'un texte.....**Erreur ! Signet non défini.**
 - 2.2.2.2 Découpages de la phrase.....**Erreur ! Signet non défini.**

2.2.2.3	Division en deux ensembles des mots français.....	Erreur ! Signet non défini.
2.2.2.4	Caractéristiques des éléments	Erreur ! Signet non défini.
2.2.2.5	Autres notions fondamentales.....	Erreur ! Signet non défini.
2.2.2	Grammaire en la langue arabe	Erreur ! Signet non défini.
2.3	Le contexte	Erreur ! Signet non défini.
2.3.1	Définition du contexte	Erreur ! Signet non défini.
2.3.2	Les catégories de contexte.....	Erreur ! Signet non défini.
2.4	Conclusion.....	Erreur ! Signet non défini.

Chapitre 03 Conception

3.1	Introduction	Erreur ! Signet non défini.
3.1	Conception globale.....	Erreur ! Signet non défini.
3.2.1	Architecture générale.....	Erreur ! Signet non défini.
3.2.1.1	Module des opérations sur la base de données	Erreur ! Signet non défini.
3.2.1.2	Module d'analyse du texte.....	Erreur ! Signet non défini.
3.2.1.3	Module de conversion.....	Erreur ! Signet non défini.
3.2.1.4	Module de génération du texte.....	Erreur ! Signet non défini.
3.3	Conception détaillée	Erreur ! Signet non défini.
3.3.1	L'étape préparatoire	Erreur ! Signet non défini.
3.3.2	L'étape d'analyse	Erreur ! Signet non défini.
3.3.2.1	Extraction des éléments grammaticaux (phrases, mots).....	Erreur ! Signet non défini.
3.3.2.2	Sélection des types des éléments grammaticaux (phrases, mots)	Erreur ! Signet non défini.

3.3.3	L'étape de la conversion	Erreur ! Signet non défini.
3.3.3.1	L'étape de la Sélection	Erreur ! Signet non défini.
3.3.3.2	L'étape de la comparaison.....	Erreur ! Signet non défini.
3.3.3.3	L'étape du test	Erreur ! Signet non défini.
3.3.3.4	L'étape de l'extraction	Erreur ! Signet non défini.
3.3.3.5	L'étape de la recherche du synonyme	Erreur ! Signet non défini.
3.3.4	L'étape de la génération	Erreur ! Signet non défini.
3.3.4.1	L'étape de la reformulation	Erreur ! Signet non défini.
3.3.4.1.1	L'étape de l'ordonnancement	Erreur ! Signet non défini.
3.3.4.1.2	L'étape de l'assemblage	Erreur ! Signet non défini.
3.3.5	Module des opérations sur la base de données	Erreur ! Signet non défini.
3.3.5.1	Opération de la suppression	Erreur ! Signet non défini.
3.3.5.2	Opération de l'ajout	Erreur ! Signet non défini.
3.3.5.3	Opération de la mise à jour	Erreur ! Signet non défini.
3.3.5.4	Opération de l'authentification.....	Erreur ! Signet non défini.
3.4	Modélisation de système	Erreur ! Signet non défini.
3.4.1	Diagrammes de cas d'utilisation	Erreur ! Signet non défini.
3.4.1.1	Identification des acteurs	Erreur ! Signet non défini.
3.4.1.2	Identification des cas d'utilisation.....	Erreur ! Signet non défini.
3.5	Conclusion.....	Erreur ! Signet non défini.

Chapitre 04 Implémentation

.....	16
Introduction générale	16
Introduction générale	1

Chapitre 01 La traduction automatique.....	1
1.1 Introduction.....	3
1.2 Définitions	3
1.2.1 Le Traitements Automatique des Langues (TAL).....	3
1.2.2 La traduction automatique (TA)	3
1.2.3 La linguistique.....	4
1.2.4 Le langage.....	4
1.2.5 Une langue.....	4
1.2.6 La linguistique informatique.....	4
1.2.7 La grammaire.....	4
1.2.8 Un dictionnaire.....	5
1.2.9 Vocabulaire.....	5
1.2.10 Intelligence artificielle	5
1.3 Historique	5
1.3.1 Les années 50-60	5
1.3.2 Le rapport ALPAC et ses conséquences.....	6
1.3.3. Entre les années 1967-1976	7
1.3.4 Les années 80	7
1.3.5 Le développement des aides à la traduction.....	7
1.3.6 Des années 90 à aujourd’hui	8
1.4. Intelligence artificielle et la traduction automatique.....	10
1.5 Les approches de traduction automatique	10
1.5.1 L'approche à base de règles	11
1.5.2 L'approche statistique	13
1.5.3 L'approche à base d'exemples	13
1.5.4 L'approche hybride.....	14
1.6 Échantillons Des Logiciels De Traduction Automatique	15
1.7 Quels usages et limites ?	17
1.8 La traduction automatique comme outil automatique.....	17
1.9 Personnalisation du système par la création d’un dictionnaire spécialisé	20
1.10 Avantages de la personnalisation.....	23
1.12 Conclusion	24
Chapitre 02 Grammaire et Contexte.....	85

.....	17
Introduction générale	17
Introduction générale	1
Chapitre 01 La traduction automatique.....	1
1.1 Introduction.....	3
1.2 Définitions	3
1.2.1 Le Traitements Automatique des Langues (TAL).....	3
1.2.2 La traduction automatique (TA)	3
1.2.3 La linguistique.....	4
1.2.4 Le langage	4
1.2.5 Une langue.....	4
1.2.6 La linguistique informatique.....	4
1.2.7 La grammaire.....	4
1.2.8 Un dictionnaire	5
1.2.9 Vocabulaire.....	5
1.2.10 Intelligence artificielle	5
1.3 Historique	5
1.3.1 Les années 50-60.....	5
1.3.2 Le rapport ALPAC et ses conséquences.....	6
1.3.3. Entre les années 1967-1976	7
1.3.4 Les années 80	7
1.3.5 Le développement des aides à la traduction.....	7
1.3.6 Des années 90 à aujourd'hui	8
1.4. Intelligence artificielle et la traduction automatique.....	10
1.5 Les approches de traduction automatique	10
1.5.1 L'approche à base de règles	11
1.5.2 L'approche statistique	13
1.5.3 L'approche à base d'exemples	13
1.5.4 L'approche hybride.....	14
1.6 Échantillons Des Logiciels De Traduction Automatique	15
1.7 Quels usages et limites ?	17

1.8 La traduction automatique comme outil automatique.....	17
1.9 Personnalisation du système par la création d'un dictionnaire spécialisé	20
1.10 Avantages de la personnalisation.....	23
1.12 Conclusion	24
Chapitre 02 Grammaire et Contexte.....	85

Table des figures

Figure (1.1): Traduction automatique à base des règles.....**Erreur ! Signet non défini.**

Figure (1.2): Triangle de Vauquois.**Erreur ! Signet non défini.**

Figure (1.3):Traduction automatique statistique.	Erreur ! Signet non défini.
Figure (1.4):Traduction automatique à base d'exemples.	Erreur ! Signet non défini.
Figure (1.1): Traduction automatique à base des règles.....	11
Figure (1.2):Triangle de Vauquois.	11
Figure (1.3):Traduction automatique statistique.	13
Figure (1.4):Traduction automatique à base d'exemples.	13

Liste des tableaux

Tableau (1.1): Echantillons Des Logiciels De Traduction Automatique.....	17
Tableau (2.1): les types de temps de verbe.	Erreur ! Signet non défini.
Tableau (2.2):Typologie des subordonnées.	30



Introduction générale

Introduction générale

La traduction est un transfert d'un texte d'une langue à une autre, il est important de tenir compte des caractéristiques de toutes les langues traduites entre elles .ces caractéristiques sont mises en pratique dans le contrôle des règles de la langue (grammaire) et la méthode qui consiste à comprendre le contexte dans lequel le texte est traduit.

La traduction automatique est devenue un moyen nécessaire et inévitable, vu la vitesse et la capacité d'échange sur internet

La traduction automatique manque une certaine crédibilité pour être totalement acceptée par les utilisateurs. Quoiqu'on ait dit des progrès de la traduction automatique, celle-ci se trouve à une limite ne dépassant pas les lignes de texte car elle ne permet pas de saisir les caractéristiques des deux langues.

L'objectif de notre projet est de créer un bon traducteur capable de traduire un texte (bilan médical et messages vers un docteur spécialiste) en la langue française à un texte (bilan médical et messages vers un docteur spécialiste) en la langue arabe dans un contexte spécifique (médical) de manière plus efficace.

Pour achever le but, nous choisirons la méthode de traduction automatique à base des règles en mode transfert et un dictionnaire médical pour réaliser la contextualisation.

Ce mémoire structuré de la manière suivante :

Chapitre 01 : Ce chapitre présente l'état de l'art sur traduction automatique. Nous allons présenter des définitions des concepts qui sont utilisés dans notre mémoire, puis nous allons aborder le sujet de traduction automatique selon les titres suivants : historique de la traduction automatique, comment l'IA révolutionne la traduction automatique?, les approches de traduction automatique, échantillons Des

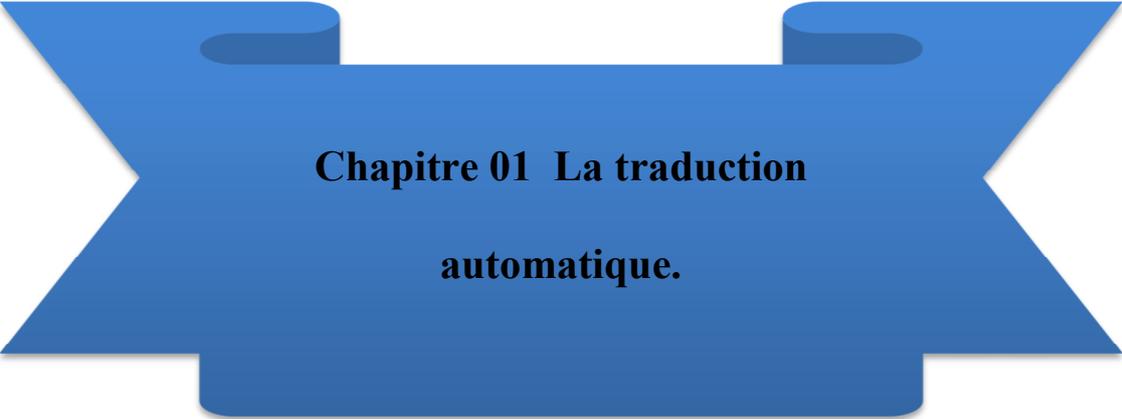
Logiciels, quels usages et limites ?, la traduction automatique comme outil automatique et enfin on parle sur la personnalisation et leur avantages.

Chapitre 02 : Dans ce chapitre, nous parlons du mot contexte dans l'informatique, puis nous présentons en générale tout ce qui concerne la grammaire de la langue arabe et française.

Chapitre 03 : présente l'architecture générale et détaillée de notre système, ainsi que la modélisation de ce dernier.

Chapitre 04 : Parle de la réalisation et l'implémentation de notre projet (nous exposons l'environnement de développement et les outils utilisés dans ce projet, ensuite nous allons présenter les interfaces graphiques en décrivant les différentes fonctionnalités de notre projet et nous présenterons des exemples qui nous permettront d'illustrer les résultats obtenus lors de l'utilisation de notre projet).

Notre mémoire se termine par une conclusion générale résumant notre travail ainsi qu'un ensemble des perceptives.



**Chapitre 01 La traduction
automatique.**

1.1 Introduction

Dans ce chapitre ,nous allons présenter des définitions des concepts qui sont utilisés dans notre mémoire, puis nous allons aborder le sujet de traduction automatique selon les titres suivants : historique de la traduction automatique, comment l'IA révolutionne la traduction automatique?, les approches de traduction automatique, échantillons Des Logiciels, quels usages et limites ?, la traduction automatique comme outil automatique et enfin on parle sur la personnalisation et leur avantages.

1.2 Définitions

1.2.1 Le Traitements Automatique des Langues (TAL)

Le Traitement Automatique des Langues (TAL) 4est un domaine de l'Intelligence Artificielle (IA) qui traite divers aspects de la langue et du discours, à tous les niveaux, lexical, morphologique, syntaxique, sémantique ou pragmatique [1].

Le TAL a pour objectif de développer des logiciels ou des programmes informatiques capables de traiter de façon automatique des données linguistiques [2].

Pour traiter automatiquement ces données, il faut d'abord expliciter les règles de la langue puis les représenter dans des formalismes opératoires et calculables et enfin les implémenter à l'aide de programmes informatiques [2].

Les principaux domaines du TAL sont :

- le traitement de la parole ;
- la traduction automatique ;
- la compréhension automatique des textes ;
- la génération automatique de textes ;
- la gestion électronique de l'information et des documents existants [2].

1.2.2 La traduction automatique (TA)

TA est le domaine de recherche à la base du traitement automatique du langage.

Et aussi c'est la traduction d'un texte réalisée complètement par un ou plusieurs programmes informatiques, sans l'intervention humaine.

1.2.3 La linguistique [3]

Science du langage, c'est-à-dire étude objective, descriptive et explicative de la structure, du fonctionnement (linguistique synchronique) et de l'évolution dans le temps (linguistique diachronique) des langues naturelles humaines. S'oppose ainsi à la grammaire (descriptive et normative) et la philosophie du langage (hypothèses métaphysique, biologiques, psychologiques, esthétiques sur l'origine, le fonctionnement, la signification anthropologiques possibles du langage).

Et aussi bien Étude historique et comparative des langues.

1.2.4 Le langage [4]

Fonction d'expression de la pensée et de communication entre les hommes, mise en œuvre au moyen d'un système de signes vocaux (parole) et éventuellement de signes graphiques (écriture) qui constitue une langue manifestations du langage.

1.2.5 Une langue

Système d'expression et de communication commun à un groupe social (communauté linguistique) [4].

1.2.6 La linguistique informatique

La linguistique informatique est un domaine multidisciplinaire dépend d'une modélisation symbolique (à base de règles) ou statistique du langage naturel établi dans l'informatique.

1.2.7 La grammaire

La grammaire est l'étude méthodologique des Éléments composites de la langue et aussi c'est un manuel ou une collection de documents expliquant des règles grammaticales.

1.2.8 Un dictionnaire

Un dictionnaire est un conteneur d'une collection des mots d'une langue ou d'un domaine d'activité généralement présentés par ordre alphabétique.

1.2.9 Vocabulaire

Collection des mots spécifiques (de domaine, de langue ou de dictionnaire) qui donne la liste des mots.

1.2.10 Intelligence artificielle

Le but de l'**Intelligence artificielle (IA)** est de concevoir des systèmes capables de reproduire le comportement de l'humain dans des activités de raisonnement [13].

1.3 Historique [5]

L'idée moderne en remonte aux années 1930 et on divise généralement son développement en quatre grandes périodes : les années 50-60, le rapport ALPAC (1966), et ses conséquences, la reprise des années 70-80 et la période actuelle qui commence au début des années 90.

1.3.1 Les années 50-60

Durant la deuxième guerre mondiale les premiers ordinateurs permettent aux belligérants de déchiffrer leurs codes respectifs. On tente ensuite d'appliquer ces techniques de déchiffrement à la traduction automatique. L'expert mondial de la traduction automatique, John Hutchins (2001), évoque la première conférence sur la TA qui fut organisée en 1952 par Bar-Hillel. Dès cette époque, les chercheurs sont conscients des difficultés que présente le domaine : ils se sont en effet déjà accordés sur la conclusion qu'une traduction de bonne qualité et totalement automatique relève de la science-fiction, ou du moins, restait du domaine du virtuel. Ils se rendent compte que l'intervention humaine, déjà qualifiée de pré-édition ou de post-édition restera indispensable. De plus, les premiers systèmes de TA partent avec de lourds handicaps : les connaissances en syntaxe et en analyse syntaxique sont encore très insuffisantes, les ordinateurs ont des capacités de stockage très limitées et sont peu puissants. On dispose avant tout de systèmes de mise en correspondance de dictionnaires, ce qui génère des résultats de traduction voués à l'échec ou alors des « systèmes-jouet »,

incapables de traduire plus de quelques phrases. Le premier système de TA est présenté au public en 1954 ; il permet de traduire 49 phrases russes, soigneusement sélectionnées, vers l'anglais se servant d'un dictionnaire de 250 mots et de six règles de grammaire. Ce sont probablement des motivations stratégiques qui poussent alors les gouvernements américain et soviétique à financer ce type de recherches. Comme nous l'avons mentionné plus haut, la recherche en linguistique théorique n'a pas encore atteint toutes les connaissances en syntaxe que l'on a aujourd'hui. Ce n'est qu'avec les découvertes de Noam Chomsky dans *Syntactic Structure* (1957) et *Aspects of the Theory of Syntax* (1962) que l'on commence à avoir des modèles syntaxiques de la langue. C'est aussi dans ces années-là que Lucien Tesnière développe une approche alternative dans ses *Eléments de syntaxe structurale* (1959). Une description syntaxique de la langue donne une structure aux phrases qui apparaissent sinon comme une simple suite de mots sans ordonnancement. La syntaxe est par conséquent indispensable au traitement automatique du langage autre que statistique. On peut donc dire que la recherche en traduction automatique des années 50 et au début des années 60 vient stimuler et appuyer toute la recherche dans le domaine du traitement automatique du langage. En effet, les systèmes de traduction automatique de différents types nécessitent tous, entre autres, un module performant et robuste d'analyse syntaxique.

1.3.2 Le rapport ALPAC et ses conséquences

Mandatée en 1966 par le gouvernement des Etats-Unis, la commission ALPAC (Automatic Language Processing Advisory Committee) publie un rapport qui va avoir une influence déterminante et très négative sur le soutien à la recherche en TA. Ce rapport conclut que la TA est plus lente, moins efficace et deux fois plus chère que la traduction faite par des humains. Il recommande d'arrêter de financer la recherche dans ce domaine et de se tourner plutôt vers des aides à la traduction, comme par exemple les dictionnaires automatiques. Le rapport recommande cependant de continuer à soutenir la recherche fondamentale en traitement automatique du langage. On peut dire que la recherche en TA a été alors virtuellement stoppée pour une décennie aux Etats-Unis

1.3.3. Entre les années 1967-1976

Au cours de cette période, la TA est en sommeil aux Etats-Unis, mais d'autres pays continuent à développer la recherche que stimulent des besoins concrets. Si aux Etats-Unis la seule activité porte sur la traduction du russe en anglais de textes scientifiques et techniques, au Canada, en Europe et au Japon en revanche, le développement de la TA se poursuit. Citons le projet canadien TAUM-Météo (décrit dans Isabelle 1986 par exemple) ; il débute en 1970 et produit deux résultats importants : il mène au développement ultérieur du langage de programmation PROLOG et s'utilise par ailleurs avec succès dans le domaine restreint de la traduction des prévisions météorologiques. En France, on développe à Grenoble un système permettant de traduire du russe des textes mathématiques et physiques (Boitet & Nedobejkine 1980). Dans la communauté européenne, les besoins en traduction deviennent de plus en plus importants et le système Systran est mis en place à la Commission Européenne en 1976. Il est encore utilisé aujourd'hui.

1.3.4 Les années 80

A la fin des années 70 et dans les années 80, on commence à travailler sur de gros systèmes commercialisables, et l'activité de recherche dans le domaine de la TA connaît un nouvel essor. A partir de la première version de Systran produite en 1970 par Peter Toma et qui consiste à traduire du russe vers l'anglais, on assiste au développement et à l'installation de nouvelles versions auprès de la Commission Européenne, et de l'OTAN, chez Xerox notamment. Les systèmes Logos et METAL se développent et sont les principaux concurrents commerciaux de Systran. Du côté de la recherche, on peut citer le premier système Ariane, développé au GETA à Grenoble chez Christian Boitet (Boitet & Nedobejkine 1980), le projet européen EUROTRA (Maegaard 1989 basé sur un système à transfert multilingue, le Distributed Language Translation system à Utrecht (Maxwell 1990), construit sur un système à pivot en passant par l'espéranto, et d'autres encore qui utilisent l'intelligence artificielle

1.3.5 Le développement des aides à la traduction

Dès les années 80, on commence à travailler sur des environnements d'aide à la traduction (Volanschi, ce volume) intégrant des dictionnaires, des concordanciers et des mémoires de traduction. Certaines entreprises qui construisent des systèmes de

TA les abandonnent (ALPSystems aux Etats-Unis) et/ou développent des mémoires de traduction (SITE à Paris). Dans certains systèmes, on commence à travailler sur le couplage de ces outils avec la traduction automatique. Les mémoires de traduction connaissent alors un développement intéressant et prennent une place prépondérante parmi les outils d'aide à la traduction

1.3.6 Des années 90 à aujourd'hui

A partir des années 90, ce sont les mémoires de traduction qui commencent à dominer le marché. Aujourd'hui ces outils font partie de la panoplie indispensable du traducteur professionnel. Avec la montée en puissance des ordinateurs personnels, les systèmes commerciaux de TA commencent à produire des systèmes individuels que l'on peut installer sur PC. Le développement d'Internet permet de mettre en ligne des systèmes de TA. Systran y a recours le premier en 1998 avec Altavista/Babelfish1. Malgré l'intérêt croissant porté aux outils d'aide à la traduction, la recherche en TA ne cesse pourtant pas et certains de ses résultats sont prometteurs. Les systèmes les plus efficaces se basaient jusqu'alors sur des systèmes à transfert, à savoir, un module d'analyse de la langue source, un module de règles de transfert complexes entre la langue source et la langue cible et un module de génération de la langue cible. Ces systèmes fonctionnent avec des règles syntaxiques et de très gros dictionnaires. La quantité croissante de documents disponibles sur support électronique permet d'enrichir toujours davantage les dictionnaires et surtout de créer des dictionnaires spécialisés qui donnent des résultats de traduction appréciables dans les domaines techniques et scientifiques. La recherche est ainsi amenée à tenter de détecter le domaine d'un texte source afin d'indiquer automatiquement au système quels dictionnaires il doit activer (Lang E. & Jin Yang 1999) On constate aussi que les systèmes mettent de plus en plus souvent l'accent sur le lexique et le lexique-grammaire ; on ajoute donc des informations syntaxiques toujours plus nombreuses aux éléments du lexique. C'est l'une des conséquences de l'évolution de la recherche en linguistique théorique et formelle, mais aussi de l'apport de la recherche en linguistique de corpus qui décrit la langue à partir des données concrètes du corpus (Scott & Tribble 2006 par exemple). Par ailleurs, les outils de traitement statistique de corpus développés en TAL ont permis l'émergence d'une nouvelle direction de recherche : la traduction automatique « basée sur les exemples ». Cette approche

permet de dégager les traductions les plus fréquentes en effectuant des analyses statistiques sur des corpus traduits et alignés. Ces systèmes ne sont pas commercialisés pour l'instant, mais ils présentent des résultats prometteurs. Même s'ils sont encore bien loin d'imiter le traducteur humain, les systèmes de traduction automatique de textes écrits sont aujourd'hui reconnus et utilisés sur le marché de la traduction. La traduction de l'oral représente un défi passionnant pour le domaine du traitement automatique du langage ; on parle ici de traduction automatique du langage parlé ou encore d'interprétariat automatique. Le système ATR (Sumita 1999) par exemple, développé au Japon permet d'effectuer des inscriptions téléphoniques à des colloques internationaux ou des réservations d'hôtel par téléphone. Citons encore le projet allemand Verbmobil (1993-2000) dont l'objectif était de créer un traducteur oral transportable pour aider les locuteurs germanophones et japonais à mener des négociations commerciales en anglais. Enfin le projet TC-STAR2 dont le but est d'effectuer la reconnaissance vocale des discours au Parlement Européen, traduit le discours et le restitue oralement dans la langue cible. Dans ce type d'outils, on combine donc la reconnaissance de la parole, la traduction automatique et la synthèse de la parole. Un autre défi consiste à combiner deux approches en traduction automatique. Si les systèmes à transfert ont en effet montré leur efficacité, ils ont aussi révélé leurs limites. Nous l'illustrerons plus bas. Par ailleurs, les systèmes statistiques donnent des résultats prometteurs mais qui révèlent eux aussi un certain nombre de limites. La combinaison des systèmes basés sur exemples avec les systèmes à transfert représente peut-être une possibilité d'améliorer de Notons par ailleurs, que la première interface graphique d'Internet, le World Wide Web n'a été mise en place par Tim Burners-Lee qu'en 1993

manière significative les résultats de la traduction automatique. La recherche en traduction automatique s'y attelle également. Pour conclure, on peut donc relever que le domaine de la TA a représenté et représente un élément moteur en recherche sur le traitement automatique du langage et qu'il conduit au développement de nombreux résultats connexes. Tous les outils d'aide à la traduction que l'on utilise aujourd'hui sont issus des recherches et du développement en TA. Le développement des analyseurs syntaxiques ou de logiciels de génération automatique est lui aussi étroitement lié aux besoins de la recherche en traduction automatique.

1.4. Intelligence artificielle et la traduction automatique [6]

L'histoire de l'intelligence artificielle est intimement liée à celle de la traduction, peut-être parce que savoir donner du sens est ce qu'il y a de plus humain et de plus mystérieux pour la machine.

« La traduction automatique marche tant bien que mal, ça n'est pas parfait, mais c'est suffisamment bon pour être utile ».

D'autant que le domaine a connu des progrès spectaculaires au cours des dernières années, grâce à l'utilisation de systèmes d'apprentissage automatique basés sur des réseaux de neurones - la technologie de l'intelligence artificielle la plus en pointe actuellement. Google, incontournable dans le domaine de la traduction gratuite avec son service Translate, fut l'un des premiers à l'employer pour cette tâche.

La technologie des réseaux de neurones a été introduite chez Google Translate via son laboratoire Google Brain et grâce aux travaux de Jeff Dean, Andrew Ng, Greg Corrado, Geoffrey Hinton et Quoc Le, figures emblématiques de l'intelligence artificielle.

1.5 Les approches de traduction automatique

La traduction automatique s'est développée autour de trois approches essentielles: les approches symboliques, les approches à base de règles, et les approches statistiques [7].

Nous détaillerons dans les approches depuis l'avènement de la traduction automatique (1950) à ce jour.

Les premiers programmes d'ordinateurs relatifs à la traduction étaient destinés à servir d'aide à la traduction. Quelques règles et surtout un dictionnaire bilingue composaient le cœur du système. Les années suivantes voient les dictionnaires grandir ; ce qui engendre une augmentation du nombre de règles régissant le réordonnement des mots. La nécessité d'automatiser l'acquisition des règles et de progresser leur généricité participe au développement de la linguistique informatique [7].

1.5.1 L'approche à base de règles



Figure (1.1): Traduction automatique à base des règles.

Le triangle présenté à la Figure (1.2) est attribué à Vauquois. Il présente de manière synthétique une analyse du processus de traduction encore pleinement pertinente et employée de nos jours [7].

La traduction se fait en plusieurs étapes:

Au niveau le plus bas, on retrouve la traduction directe, qui passe directement des mots de la langue source aux mots de la langue cible [7].

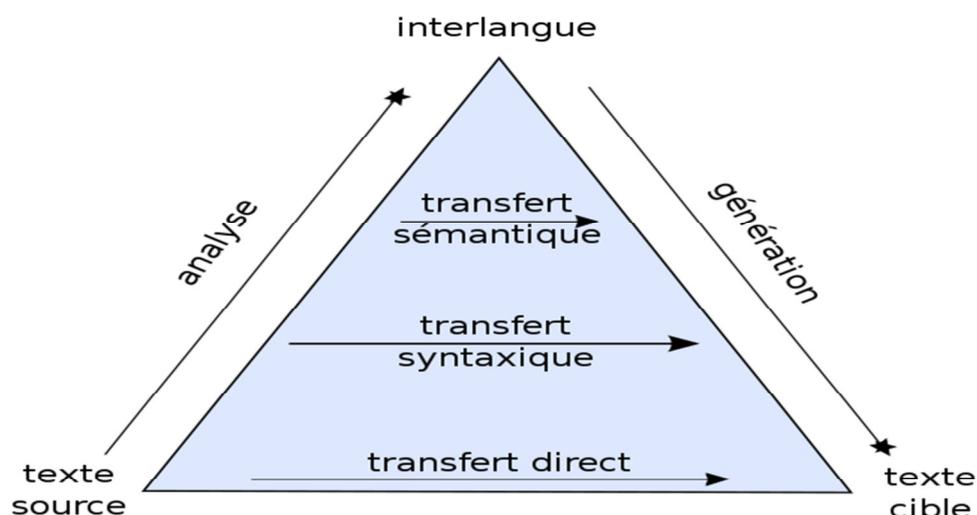


Figure (1.2): Triangle de Vauquois.

Les systèmes semi-directs ont une phase de segmentation ou d'analyse morphologique, voire morphosyntaxique, et une phase de génération morphologique. Si l'on effectue une analyse syntaxique de la phrase source, le transfert à la langue cible devrait être simplifié. À ce niveau, les détails spécifiques à la constitution des groupes nominaux, par exemple, n'ont pas besoin d'être connus des règles régissant le transfert [7].

Avec une analyse plus approfondie de la phrase source, au niveau sémantique, le transfert devient uniquement sémantique. En revanche, la génération des mots après le transfert est plus complexe qu'au niveau inférieur [7].

Enfin, une analyse totale de la phrase source peut aboutir à une représentation de son sens dans une « inter-langue » artificielle, de laquelle on dérive ensuite les mots cibles [7].

Le pivot est un type de représentation utilisant des attributs et relations inter-langue, et des unités lexicales de chacune des langues. Ces systèmes sont à transfert simple, alors qu'on a un double transfert en < pivot > [7].

L'approche reposant sur une inter-langue est attractive car elle remplace le problème de la traduction par deux problèmes unilingues, d'analyse et de synthèse. L'avantage est que les modules unilingues sont a priori réutilisables. Pour couvrir tous les sens de traduction entre n langues, il suffit de n modules d'analyse et de n modules de synthèse, contre $n*(n-1)$ systèmes de transfert [7].

Des systèmes à véritable langue-pivot, on peut citer : ATLASII de Fujitsu ou IVOT/Crossroads de NEC, ou KANT / CATALYST de CMU/Caterpillar, ou UNL, ou MASTOR-1 d'IBM [7]. .

Le début des années 1990 voit le développement d'autres types d'approches. Les ordinateurs se répandent et gagnent en puissance, ce qui permet l'émergence de stratégies qui se fondent sur de grandes quantités de données « corpus based approaches ». On distingue en particulier deux grands types d'approches: la traduction automatique à base d'exemples et la traduction automatique par méthodes statistiques [7].

1.5.2 L'approche statistique



Figure (1.3):Traduction automatique statistique.

Elle doit ses origines aux travaux de Brown et al .en 1990 et en particulier au prototype Candide de Berger et al. en 1994, un système de traduction construit à partir de discours disponibles en français et en anglais de parlementaires canadiens. En effet, comme la traduction à base d'exemples, la traduction par méthodes statistiques repose sur un corpus parallèle [7].

Un modèle statistique de traduction est défini, comprenant une ou plusieurs lois de probabilités. Le corpus est traité afin d'estimer ces lois qui sont souvent constituées de plusieurs milliers, voire millions de paramètres [7].

1.5.3 L'approche à base d'exemples

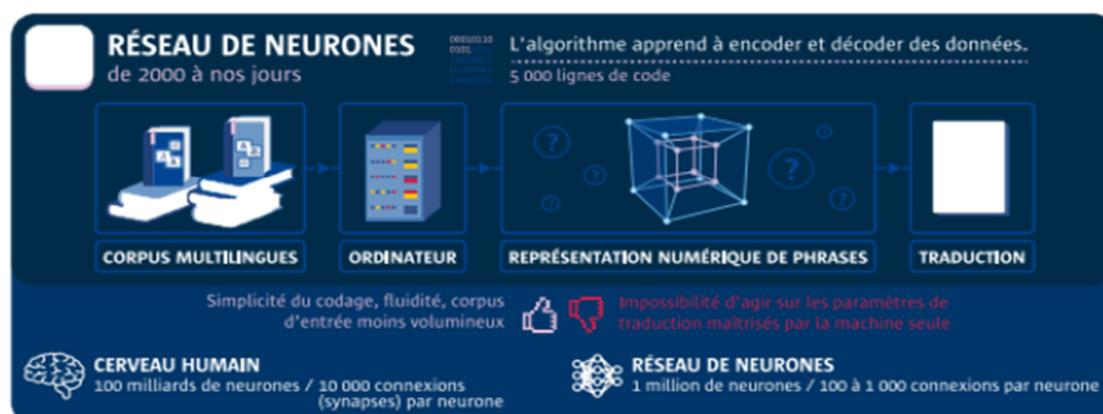


Figure (1.4):Traduction automatique à base d'exemples.

L'approche à base d'exemples (« Example-based machine translation », ou EBMT) repose sur un ensemble « d'exemples » préalablement traduits : un corpus parallèle de phrases traduites l'une de l'autre. Lorsqu'on lui présente une phrase à traduire, le système parcourt sa base d'exemples et produit trivialement une traduction si la phrase s'y trouve. Dans le cas général, la phrase n'apparaît pas dans la base et le système s'emploie alors à rassembler des exemples qui contiennent des fragments communs (des groupes de mots) avec la phrase à traduire. Pour chaque fragment d'exemple dans la langue source, il s'agit ensuite de retrouver sa traduction dans la langue cible : c'est la phase d'alignement. Enfin, la phase de génération assemble les fragments dans la langue cible et produit la traduction. À chacune des trois étapes, il est possible d'utiliser des sources externes de connaissances, telles que des lexiques bilingues, des listes de synonymes, des étiquettes ou des arbres syntaxiques... etc. [7]

1.5.4 L'approche hybride [7]

Leur idée directrice est qu'une approche unique du problème de la traduction, aussi perfectionnée soit-elle, ne parviendra pas à produire une traduction satisfaisante dans tous les cas. Au contraire, une approche par règles peut s'avérer particulièrement adaptée à certaines phrases, tandis que d'autres phénomènes linguistiques sont correctement traités par une approche reposant sur des corpus

Les systèmes hybrides sont actuellement envisagés comme des systèmes combinant les méthodes statistiques ou méthodes à base d'exemples avec des méthodes linguistiques (à base de règles), en particulier pour l'analyse morphologique et syntaxique.

Un système hybride pourrait parvenir à tirer profit des forces de chaque approche. Une première stratégie pour mettre en œuvre un système hybride est d'utiliser les différentes approches en parallèle. Enfin, dans un système statistique, il est courant de traiter par un système de règles spécialisées certains fragments de phrases, typiquement les nombres, les dates, etc. Les morceaux de phrases ainsi identifiés et traduits en isolation par le système à base de règles peuvent être transmis au système statistique.

1.6 Échantillons Des Logiciels De Traduction Automatique [8]

Num	Nom de logiciel	Lien	Remarques
1	Altavasta (de Babelfish)	http://www.babelfish.altavista.com	Ce logiciel permet de traduire en ligne à partir de plusieurs langues (anglais, français, allemand, italien, néerlandais, chinois) en plusieurs autres. (Langue de communication : l'anglais)
2	Bing	http://www.bing.com.translator	Le logiciel de traduction Bing permet également de traduire automatiquement d'une centaine de langues et vice versa. (Langue de communication : l'anglais)
3	Google Traduction	https://googletraduction.fr	Ce site de Google français permet de traduire dans un nombre considérable de combinaisons linguistiques entre l'anglais, le français, l'allemand, l'espagnol, l'italien, le portugais, le coréen, le japonais, le russe et le chinois. (langue de communication : le français)
4	Linguee	http://www.linguee.fr	Linguee est un outil de traduction unique en son genre. Il combine un dictionnaire rédactionnel et un moteur de recherche qui vous permettent de chercher la traduction d'un mot ou d'une expression parmi plusieurs centaines de millions de textes bilingues. Linguee possède mille fois plus de matériel traduit qu'un dictionnaire en ligne traditionnel, et ce sous la forme de phrases complètes et de leur traduction. (langue de communication : le français)
5	Prompt	http://www.translation2.paralink.com	Ce site de Prompt permet de traduire à partir de l'anglais, du français, du russe, de l'allemand, de l'italien et de

			l'espagnol, généralement vers les mêmes langues. (langue de communication : l'anglais)
6	Reverso	http://www.reverso.net	Le logiciel Reverso est un portail d'outils linguistiques y compris la traduction en une variété de combinaisons linguistique, notamment en français, espagnol, allemand, russe, portugais, chinois, arabe, hébreu et japonais (langue de communication : l'anglais)
7	Systran	http://www.systransoft.com	Ce site de Systran permet de traduire en ligne à partir de nombreuses langues, avec une possibilité impressionnante de combinaisons. On est limité à 150 mots à la fois. (langue de communication: l'anglais)
8	Voila	http://www.tr.voila.fr	Ce site de Voila permet de traduire en ligne à partir de plusieurs langues (anglais, français, allemand, portugais, espagnol, russe) en plusieurs autres. Le nombre de combinaisons est relativement limité. (langue de communication : le français)
9	Wanadoo	http://www.outils.wanadoo.fr/traduction	Ce site de Wanadoo permet de traduire dans un nombre considérable de combinaisons linguistiques entre l'anglais, le français, l'allemand, l'espagnol, l'italien, le néerlandais, le portugais et le russe. (langue de communication : le français)
10	WorldLingo	http://www.worldlingo.com/fr	Cet excellent site de WorldLingo permet de traduire à partir des langues sources suivantes : français, chinois, néerlandais, allemand, italien, grec, japonais, portugais, coréen,

			russe et espagnol. Les langues cibles : les mêmes. (langue de communication : le français)
--	--	--	--

Tableau (1.1): Echantillons Des Logiciels De Traduction Automatique.

1.7 Quels usages et limites ? [9]

" La traduction automatique excelle dans certaines tâches, mais échoue dans d'autres. Tout dépend du besoin ".

« Ce qui fait la force de la traduction automatique, c'est moins sa qualité que sa valeur d'usage », constate François Yvon, chercheur au LIMSI/CNRS. Le service offert aux internautes par Facebook ou TripAdvisor suffit à traduire des contenus simples – informations touristiques ou produits grand public, par exemple. Ses usagers en connaissent implicitement les limites et s'amuse d'erreurs parfois grossières. Voire de contresens ; *smoke eater scented candle*, par exemple, devient : « bougie parfumée à la fumée » (au lieu de bougie parfumée anti-tabac). La traduction automatique peut toutefois s'avérer très efficace pour des usages professionnels comme la mise à jour annuelle de manuels techniques ou de documents internes d'entreprise, ou appliqué à un domaine précis comme la météo. À l'instar de Systran, des opérateurs spécialisés proposent d'ailleurs d'adapter les algorithmes aux besoins sectoriels de leurs clients – un secteur au chiffre d'affaires mondial estimé à 150-200 millions d'euros annuels. Dans tous les cas, l'algorithme de traduction automatique décompose le texte en phrases et se révèle donc, selon François Yvon, « incapable de traduire des phénomènes qui dépassent la phrase, par exemple les références pronominales (it en anglais, versus elle ou il en français), les registres de langue ou les choix lexicaux ». Les textes destinés à la publication doivent donc toujours être relus et corrigés par des traducteurs professionnels. Enfin, la traduction automatique échoue – par nature – à restituer les termes inventés ou poétiques, les expressions orales ou les abréviations, les effets humoristiques ou les jeux de mots. Inutile d'y recourir pour traduire des œuvres littéraires !

1.8 La traduction automatique comme outil automatique [5]

Pour aborder la question de la nature de la traduction automatique, on doit poser la question de son utilité et de son utilisation. En effet, si la traduction automatique est « machinale » est-elle vraiment utile ? Les systèmes de TA s'intègrent dans un ensemble plus vaste d'outils d'aide à la traduction. Par conséquent, on doit considérer que la TA est un outil d'aide à la traduction presque au même titre que les dictionnaires électroniques monolingues ou multilingues, les concordanciers, les mémoires de traduction, les vérificateurs orthographiques et grammaticaux, la fonction « recherche » sous Word, etc. (Volanschi, ce volume). On peut considérer l'utilisation de ces divers outils comme étant sur un continuum allant de l'être humain à la machine. Selon que l'on sera plus ou moins proche de l'intervention humaine ou de celle de la machine, la charge de travail sera répartie plus ou moins sur l'humain ou la machine. Se pose alors la question des besoins et des objectifs de la traduction que l'on a à réaliser. Le choix des outils serait donc fonction des besoins ? La faille dans cette logique est que les outils génèrent de nouveaux besoins. Commençons par les besoins et l'évolution du domaine de la traduction au niveau mondial qui justifient l'utilisation de l'ordinateur de manière générale. On trouve chez les experts en TA, mais aussi dans les argumentaires commerciaux des systèmes commercialisés, les explications suivantes:

* Le premier point concerne la quantité de texte à traduire. Aujourd'hui, la quantité de documents qui doivent être traduits est tellement importante et croît si rapidement qu'il n'est pas possible aux traducteurs de tous les traduire. L'ordinateur peut contribuer à augmenter la quantité de documents traduits.

* Nombreux sont les documents qui sont très techniques et assez laborieux à traduire. L'ordinateur peut aider les traducteurs dans cette tâche peut-être ingrate.

* Les grosses entreprises ont pris conscience, depuis un certain nombre d'années maintenant, que la transmission de l'information multilingue demande une terminologie normalisée et cohérente dans toutes les langues utilisées. Les outils d'aide à la traduction proposent systématiquement les termes normalisés par l'entreprise. Cette cohérence est plus longue à réaliser avec des traducteurs humains qui travaillent en équipe et sans outils d'aide à la traduction.

* Les entreprises exigent toujours davantage de texte traduit, dans des délais de plus en plus brefs. L'ordinateur peut aider à intensifier le rythme et à augmenter le volume de texte traduit

* Enfin, dans certaines situations, il n'est pas toujours nécessaire d'avoir une traduction de haute qualité. La traduction automatique permet de traduire très rapidement des pages web rédigées dans une langue que l'on ne connaît pas, avec un résultat approximatif mais qui permet de comprendre globalement le contenu de la page.

Outre ces constatations sur les besoins au niveau industriel, Hutchins (2004) relève aussi des besoins à un niveau plus individuel. Il répartit donc ces besoins en quatre types qui devraient permettre de choisir le bon outil de traduction:

- * La diffusion
- * L'assimilation (ou la veille)
- * Les échanges
- * L'accès de type base de données

Cette classification reflète aussi l'influence d'un outil informatique sur la génération de nouveaux besoins. Nous verrons que le choix de la traduction automatique a des incidences sur la normalisation des documents source et donc probablement, un impact sur la rédaction technique

- La diffusion :

La diffusion représente le niveau le plus élevé en traduction. Elle consiste à produire des traductions - généralement en quantités importantes et de haute qualité - qui seront publiées. Le niveau de qualité doit approcher la perfection: dans le domaine médical, par exemple, les instructions d'utilisation d'un instrument doivent être claires pour éviter les accidents et les litiges ; il est par ailleurs dans l'intérêt d'une entreprise et de son image de marque de publier des documents de qualité. Il est possible de choisir la TA pour la diffusion ; ce choix requiert cependant un certain nombre d'opérations autour de la traduction elle-même : la pré-édition du texte source

avant la TA, la personnalisation du système de TA avant le processus et la post-édition du résultat de traduction.

Pré-édition du texte source avant la TA:

Les systèmes de TA commercialisés sont des systèmes à règles. Bien que le nombre de règles soit très important et que celles-ci soient constamment corrigées en fonction des tests linguistiques effectués, certaines structures - comme les subordonnées - ou certains phénomènes - comme ceux de la dépendance à longue distance ou l'ellipse - posent problème. Par ailleurs, si le texte source contient des erreurs, le système ne pourra pas rétablir le texte correct. C'est donc à un être humain que revient la tâche de corriger le texte source et, plus radicalement, de le simplifier. La simplification est en effet en relation étroite avec le langage dit contrôlé. Le langage contrôlé est une forme de langue naturelle extrêmement prescriptive et simplificatrice au niveau de la syntaxe et du lexique. Il est largement utilisé dans la rédaction de documentation technique - par exemple, dans l'industrie aéronautique ou automobile - de manière à faciliter, entre autres, le processus de TA. Les questions d'effacement sont difficiles à traiter par un système de TA. L'exemple suivant *cable and ADSL connexion* sera traduit par le *câble et la connexion de l'ADSL*, alors qu'il s'agit de la *connexion par câble et par ADSL*. Les recommandations d'un langage contrôlé demandent à ce que soit rétabli le terme effacé : la *connexion par câble et la connexion par ADSL*. C'est aussi ce qui est recommandé pour la pré-édition d'un texte à faire traduire par un système de TA. De même, l'ellipse du *that* en anglais doit être rétablie

Ce type de problèmes est déjà évoqué par Austermühl (2001 :163) qui cite les recommandations faites par un système de TA commercialisé : *Avoid idiomatic expressions Avoid omitting pronouns before a verb (...) Keep to standard, formal English in which grammatical connections are clearly expressed* Ces recommandations sont encore valables aujourd'hui

1.9 Personnalisation du système par la création d'un dictionnaire spécialisé [5]

Un autre problème rencontré par la TA est en effet celui du lexique spécialisé (terminologie) ou tout simplement du figement des termes dans un domaine donné.

L'exemple ci-dessous montre comment un système peut traduire un texte dans un domaine spécialisé, sans l'aide d'un dictionnaire machine spécialisé

Texte source	Traduction par le système	Traduction corrigé
(...)decoded by specific individuals	(...)décodée par les individus spécifiques	(...)décodée par des individus spécifiques

Le système possède une règle de transfert qui spécifie que le déterminant zéro en anglais doit être traduit par un déterminant défini en français, comme c'est le cas dans des affirmations généralisantes (Life is beautiful => La vie est belle). Or, dans ce cas précis, il ne s'agit pas d'une généralisation mais justement d'une situation dans laquelle on ne connaît pas encore les individus dont il est fait mention. En français, il faudra donc utiliser un déterminant indéfini. Mais la règle de transfert citée ci-dessus a été appliquée de manière sur généralisant.

L'exemple suivant montre comment une règle de transfert sur la traduction du passif anglais par un verbe réfléchi en anglais s'applique à tort:

Texte source	Traduction par le système	Traduction corrigé
which has been termed hypermedia	qui s'est nommée hypermedia	qu'on a appelée hypermedia

Comme mentionné plus haut, les systèmes à transfert peuvent être décrits de manière simplifiée comme fonctionnant avec trois modules : un module d'analyse syntaxique de la langue source qui produit une représentation abstraite de la phrase, un module de transfert contenant des règles de traduction et un module de génération qui, à partir d'une représentation abstraite de la langue cible produite par le module de transfert, génère le phrase de la langue cible. Des erreurs peuvent donc se produire à tous les stades du processus. Nous illustrons ci-dessous des erreurs qui interviennent dès la phase d'analyse de la langue source et se propagent dans les deux autres phases pour donner un résultat de traduction complètement erroné.

Dans l'exemple ci-dessous, on peut penser à première vue que le mot *light* a été mal analysé par le système puisqu'il est traduit par un adjectif, alors qu'il s'agit du nom *light* qui se traduit par lumière. Cette erreur provient en fait d'une erreur d'analyse syntaxique de toute une partie de la phrase et ne porte pas seulement sur un mot. C'est tout le groupe nominal complexe *the phase of the light passing through* qui a été mal analysé. Le mot *passing* qui ici fonctionne comme la forme verbale du verbe *to pass* (en fait *to pass through*) a été analysé comme une nominalisation.

Texte source	Traduction par le système	Traduction corrigé
The liquid crystal changes the phase of the light passing through it	Le cristal liquide change la phase du dépassement léger par lui	Le cristaux liquide modifient la phase de la lumière qui les traverse .

De même, dans la phrase suivante, le mot *can* a été analysé comme un nom, rendant l'analyse correcte de la phrase totalement impossible .

Texte source	Traduction par le système	Traduction corrigé
If such a unit is placed between a pair of plane polariser plates then light can pass through it only if the correct voltage is applied.	Si une telle unité est placée entre une paire de passage léger de bidon de plaques plane polarisées puis par elle seulement si la tension correcte est appliquée	Si on place cette unité entre deux plaques planes polarisées, alors la lumière peut les traverser seulement si on applique le voltage correct .

Ces deux exemples démontrent aussi que si la TA ne va pas au-delà de la phrase, elle ne va parfois même pas jusqu'à la phrase. En effet, certains systèmes de TA commercialisés effectuent une analyse de la phrase lorsque c'est possible, mais sinon, ils fonctionnent plutôt avec l'analyse « par chunks », c'est-à-dire, par morceaux

de phrase (la plupart du temps en tentant de reconnaître des syntagmes plus ou moins complexes). Cela expliquerait tout à fait les erreurs d'analyse ci-dessus .

On voit ici que la production d'une traduction de haute qualité à l'aide de la TA ne peut se faire sans intervention humaine. Rappelons cependant que la révision est un processus qui s'applique aussi aux traductions effectuées par les traducteurs. On peut donc envisager d'utiliser la TA pour produire des traductions publiables, mais il est évident que ce choix est aussi dépendant du domaine dans lequel on traduit. Les dictionnaires spécialisés montrent une bien plus grande utilité dans certains domaines spécifiques par exemple, scientifiques et techniques, que dans d'autres – par exemple, dans les domaines économiques et juridiques

Cela tient sans doute au fait que les systèmes de TA ont été surtout développés pour traduire de la documentation technique ; leurs modules d'analyse, de transfert et de génération ne contiennent donc pas les structures de la langue juridique, souvent obsolètes en langue générale ou scientifique et technique. Par ailleurs, ce n'est pas seulement le type de domaine qui doit déterminer l'emploi ou non d'un système de TA, c'est aussi le genre du texte : des documents techniques, comme des manuels d'utilisation ou des spécifications techniques rédigés en langage contrôlé, ainsi que des articles scientifique très spécialisés et de haut niveau sont bien plus adaptés à la TA que des documents de vulgarisation scientifiques, des cours en ligne, des textes de marketing. Le travail de révision est en effet bien moins important dans le premier type de textes.

1.10 Avantages de la personnalisation

- La personnalisation du logiciel est le processus permettant d'améliorer la qualité de traduction selon les besoins spécifiques de chaque client [10].

- La personnalisation est adaptée pour les clients ayant d'importants besoins de traductions dans un domaine spécifique comme l'automobile, l'aéronautique, l'électronique, la pharmacie, la chimie, ou les services financiers. Elles sont mises en œuvre pour entraîner le logiciel qui apprend comment traduire dans le contexte précis de chaque client [10].

- La personnalisation sert le sens (sémantique).

1.12 Conclusion

Dans ce chapitre nous avons illustré les points les plus importants pour bien comprendre la traduction automatique.

Dans le chapitre suivant, nous allons présenter la notion du contexte et de la grammaire dans les deux langues française et arabe.



Chapitre 02 Grammaire et Contexte.

2.1 Introduction

Ce chapitre présente les grammaires de deux langues (française et arabe) et le contexte.

2.2 La grammaire

2.2.1 Définition

La grammaire est l'étude et la description sous forme de règles des structures syntaxiques, morphologiques et phonétiques.

L'enseignement de la grammaire a pour finalité de favoriser la compréhension [14] et l'analyse des textes lus pour la traduction automatique.

2.2.2 La grammaire de la langue française

2.2.2.1 Étapes de l'analyse d'un texte [11]

Lexicale -> syntaxique -> sémantique -> stylistique

Vocabulaire -> ordre, fonction -> sens -> figure de style

Remarque : La grammaire se situe au niveau de la syntaxe mais elle peut déborder un peu sur l'analyse lexicale ou sémantique.

2.2.2.2 Découpages de la phrase [11]

On peut découper la phrase en sujet (ce dont on parle) et prédicat (ce que l'on en dit).

Les compléments peuvent être découpés en compléments essentiels (liés au verbe, non effaçable, non déplaçable ; COD, COI, attribut du sujet, COS) et compléments circonstanciels (effaçable, déplaçable ; CC).

Ou encore, en compléments de verbe (ne modifient que le verbe) et compléments de phrase (apportent une information concernant l'ensemble de la phrase).

Deux décompositions de la phrase

Sujet + prédicat + complément essentiel/circonstanciel

Groupe verbal + groupe nominal sujet + groupes nominaux compléments

2.2.2.3 Division en deux ensembles des mots français

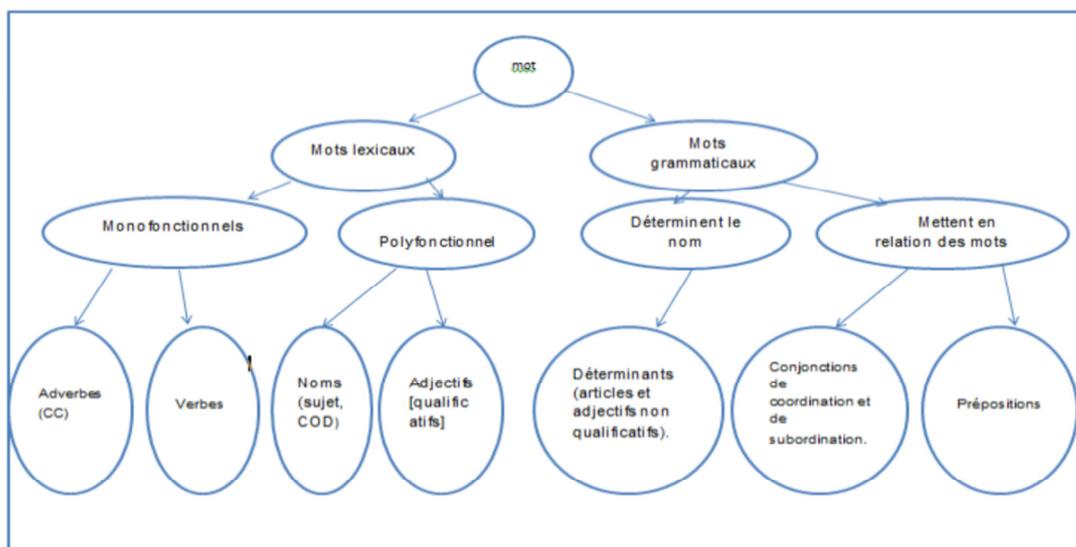


Figure (2.1):Division en deux ensembles des mots français [11].

2.2.2.4 Caractéristiques des éléments [11]

➤ **Déterminant**

Type essentiel (non déplaçable, non supprimable) ; utilité : indique le genre (sauf certains au pluriel), le nombre et le nom (toujours placé avant le nom) ; s'accorde avec le nom-noyau.

- Classement des Déterminants (selon possibilités des combinaisons):

a- Articles, adjectifs possessifs et démonstratifs.

b- Adjectifs indéfinis, numéraux cardinaux, interrogatifs, exclamatifs.

- Les déterminants d'une même série peuvent être combinés entre eux
 - a- Articles définis, adjectifs démonstratifs, possessifs, exclamatifs
 - b- Articles indéfinis, articles partitifs, adjectifs indéfinis, adjectifs numéraux.

➤ **Article**

- défini (le/la/les, contracté : au/aux/du/des).
- indéfini (un/une/des).
- partitif (de + article).

➤ **Adjectif non qualificatif**

- possessifs (ex. mon/ton/son/notre/votre/leur).
- démonstratifs (ex. ce/ces ; nom-ci proximité, nom-là éloignement), numéraux (ordinaux (premier).
- cardinaux (trois cents)).
- indéfinis (nombreux, mot (quelque) ou locution (n'importe quel)).
- Interrogatifs.
- exclamatifs (ex. quel/quels).

➤ **Préposition**

Mise en relation avec le verbe (introduit un CC) ou autre des mots, indique la fonction.

Temps, cause, conséquence, but, condition, comparaison, lieu.

Mot simple ou locution prépositionnelle.

➤ **Conjonction**

- Coordination :

Mise en relation d'éléments de même statut syntaxique. entre élément de même nature et de même fonction.

- Subordination :

Mise en relation d'éléments de statut syntaxique différents ; l'élément subordonné dépend du subordonnant ; préposition uniquement si même nature.

- Juxtaposition :

Mise en relation sans terme de liaison.

➤ **Nom**

C'est le noyau du groupe nominal. Il désigne une personne, un animal ou une chose.

Toujours accompagné d'un déterminant (sauf nom propre).

Ils imposent leur genre et leur nombre.

Classement : animé (humain ou non)/inanimé ; nom propre/commun.

➤ **Pronom**

Le pronom remplace le plus souvent un nom ou un groupe nominal ; mais il peut aussi se substituer à un adjectif ou à une proposition tout entière.

Pronoms personnels (ex. je/tu/il), personnels réfléchis (ex. me/te/se/nous/vous), démonstratifs (ex. ce/ceci), possessifs (ex. le mien/tien/sien), interrogatifs (ex. qui/que/lequel), relatifs (ex. qui/que/quoi/dont)..(

➤ **Adjectif [qualificatif]**

(attribut, épithète, comparatif, superlatif) Modifie le sens des noms.

➤ **Verbe**

Noyau de la phrase, appartient au prédicat.

Exprime une action ou un état.

Les marques des verbes sont la personne, le temps, l'aspect, et le mode, négation. Combinaison des formes = conjugaison.

La fonction des verbes s'oppose à celle des noms, cela est une caractéristique du français. 1er groupe -er, 2e groupe -ir, 3e groupe irrégulier (conjugaison morte).

Intransitifs (pas de COD), transitifs (COD possible).

Voix active, passive, pronominale.

Modes personnels			
Indicatif	Présent	passé composé	
	passé simple	passé antérieur	
	Imparfait	plus-que-parfait	
	futur simple	futur antérieur	
Subjonctif	Present	Passé	
	Imparfait	plus-que-parfait	
Conditionnel	Present	passé 1	passé 2
	Present		
Modes impersonnels			
	Infinitif	Participe	Gérondif

Tableau (2.1): les types de temps de verbe.

➤ **Adverbe**

Modifient le sens des verbes (ce sont les seuls), des adjectifs qualificatifs, d'un autre adverbe.

Ils sont invariables.

Adverbe de temps, de quantité, de manière, de lieu, d'affirmation, de négation, de doute.

2.2.2.5 Autres notions fondamentales

- Mode d'énonciation :

Exclamatif (surprise, indignation, admiration), interrogatif, impératif, déclaratif).

- Types d'appartenance dans la phrase :

Attribut (appartient groupe verbal, insuppressible, complément des verbes d'état) ; Épithète (appartient au groupe nominal, supprimable, directement relié au nom) ; Apposition.

- Il existe trois moyens pour marquer la fonction d'un mot ou groupe
Prépositions, position, sens.

- Typologie des subordonnées:

Subordonnées			
au verbe de la principale		à un nom de la principale	
Conjonctive		Relative	
Complément			
Essentiel	non essentiel	difficilement effaçable	Effaçable
Complétives	Circonstancielle	Déterminative	Explicatives ou Appositives

Tableau (2.2):Typologie des subordonnées.

2.2.2 Grammaire en la langue arabe [12]

➤ **L'adjectif de relation (nisba)**

L'adjectif de relation ou la nisba est un procédé syntaxique très fréquent en arabe. Il se construit en ajoutant **سي** à n'importe quel nom.

لُبْنَانِي لُبْنَان
Libanais Liban

Figure (2.1): Exemple 1 sur L'adjectif de relation

Pour former un adjectif de relation féminin, on rajoute le **ة**.

لُبْنَانِيَّة لُبْنَان
Libanaise Liban

Figure (2.2): Exemple sur L'adjectif de relation en féminin.

Les noms qui se terminent en **ا** comme **فَرَنْسَا** (la France) en terminent par **ِي**:

فَرَنْسَا فَرَنْسَا
Français France

Figure (2.3): Exemple 2 sur L'adjectif de relation.

➤ Annexion (Idhāfa)

L'annexion ou l'Idhāfa est également un procédé syntaxique courant en arabe. Il permet de déterminer un nom en l'annexant à un autre en remplissant la fonction de complément de nom, de la grammaire française, exprimée par : d', de, du, des... . L'annexion se construit par la juxtaposition des deux noms. Ex:

كِتَابُ الْعَرَبِيَّةِ

Le livre d'arabe.

Figure (2.4): Exemple 1 sur Annexion.

Ce groupe nominal se compose ainsi des deux termes dits :

-Premier terme : mudhāf .

-Deuxième terme : mudhāf ilayhi .

Le premier terme de l'annexion ne prend jamais d'article défini, car il sera déterminé par le deuxième. Son cas varie selon sa position dans la phrase. Le deuxième terme prend impérativement le cas indirect. Si l'ensemble de l'annexion est défini, le deuxième terme est défini.

Le livre d'arabe : كِتَابُ الْعَرَبِيَّةِ

Figure (2.5): Exemple 2 sur Annexion.

Si l'ensemble de l'annexion est indéfini, le deuxième terme est indéfini.

Un livre d'arabe : كِتَابٌ عَرَبِيَّةٌ

Figure (2.6): Exemple 3 sur Annexion.

➤ **Adjectif (na't)**

L'adjectif ou le na't se place impérativement après le nom qu'il qualifie. Il s'accorde avec ce nom en quatre points:

a- En genre : féminin ou masculin

طَالِبَةٌ جَمِيلَةٌ OU طَالِبٌ جَمِيلٌ

Figure (2.7): Exemple sur Adjectif en genre.

b- En nombre : singulier, duel ou pluriel

طُلابٌ جَمِيلُونَ ou طالبٌ جَمِيلٌ

Figure (2.8): Exemple sur Adjectif en nombre.

c- En cas : sujet, direct ou indirect

طالبًا جميلًا ou طالبٍ جميلٍ ou طالبٌ جميلٌ

Figure (2.9): Exemple sur Adjectif en cas.

d- En détermination/Indétermination : les deux termes seront déterminés ou indéterminés.

الطالِبُ الجَمِيلُ ou طالبٌ جَمِيلٌ

Figure (2.10): exemple sur Adjectif en détermination/Indétermination.

➤ **Détermination (ta‘rīf)**

La détermination ou le ta‘rīf se réalise en faisant précéder les noms de l’article défini **ال**. Cet article vaut pour tous les noms, qu’ils soient : masculins, féminins, au duel ou au pluriel.

La rose : الوَرْدَةُ
 Le livre : الكتاب
 Les deux roses : الوَرْدَتان
 Les deux livres : الكتابان
 Les roses : الورود
 Les livres : الكُتُب

Figure (2.11): Exemple sur Détermination.

➤ **Indétermination (tankīr)**

L'indétermination ou le tankīr se construit en mettant un tanwīn (doublement de la voyelle finale) sur un nom et en enlevant l'article défini. Le tanwīn se met uniquement sur les noms singuliers (et ceux au pluriel interne), qu'ils soient masculin ou féminin.

Une rose : وَرْدَةٌ (cas indirect), وَرْدَةٌ (cas direct), وَرْدَةٌ (cas sujet)
 Le livre : كِتَابٌ (cas indirect), كِتَابًا (cas direct), كِتَابٌ (cas sujet).
 وَرْدَةٌ : وَرُودٌ/وَرْدٌ/وَرْدًا les roses, est le pluriel interne de

Figure (2.12): Exemple sur Indétermination.

➤ **Auxiliaire « être »**

En arabe, l'auxiliaire être n'existe pas (au présent). Pour former des phrases nominales simples, il suffit de juxtaposer le sujet (mubtada') dont on parle et l'information (habar), le concernant.

L'étudiant est beau.
 الطالب جميل.

Figure (2.13): Exemple sur auxiliaire être.

Ces deux fonctions sont cependant marquées par le cas sujet. L'information prend un tanwīn.

➤ **Auxiliaire «avoir »**

En arabe, l'auxiliaire avoir n'existe pas non plus. Pour exprimer la possession,

l'arabe fait appel à une tournure particulière qui se construit par la particule suivie d'un pronom affixe ; le tout suivi de l'objet possédé.

لِ

J'ai une maison
 لي بيت
 Tu as une maison
 لك بيت

Figure (2.14): Exemple sur auxiliaire avoir.

Notez bien qu'il s'agit ici d'une phrase nominale d'où le cas sujet sur le mot بيت, considéré ici comme un sujet. La traduction littérale de cette tournure est:

À toi une maison = Une maison est à moi.

Figure (2.15) : Exemple sur La traduction littérale.

➤ **Pronoms affixes**

Le pronom affixe ou le Dhamīr muttasil est une « particule » qui s'attache aux trois types de mots:

a- Noms : pour exprimer la possession

كِتَاب + ي = كِتَابِي
 Mon livre = **mon** + livre.
 كِتَاب + ك = كِتَابِكَ
 Ton livre = **Ton** + livre.

Figure (2.16): Exemple sur Pronoms affixes en cas de noms.

b- Verbes : pour exprimer le caractère réfléchi:

عَرَفَ + ك = عَرَفَكَ

Il t'a connu : = **Te** + connaitre.

Figure (2.17): Exemple sur Pronoms affixes en cas de verbes.

c- Particules : pour exprimer le sens de la particule relatif à la personne concernée

فيك = في + ك

Figure (2.18): Exemple sur Pronoms affixes en cas de Particules.

Les pronoms affixes sont au nombre de 15:

Singulier ي-ك-ك-ه-ها
 Duel نا-كُما-كُما-هُما-هُما
 Pluriel نا-كم-كن-هم-هن

Figure (2.19): Exemple sur Pronoms affixes.

➤ **Pronoms personnels**

Les pronoms personnels ou le Dhamīr munfasil sont au nombre de 15. Hormis la distinction entre la première, la deuxième et la troisième personne, on différencie aussi le duel (une paire de personnes ou de choses) et le pluriel (à partir de trois):

Pluriel	Duel	Singulier
نَحْنُ	نَحْنُ	أَنَا
أَنْتُمْ	أَنْتُمَا	أَنْتَ
أَنْتُنَّ	أَنْتُمَا	أَنْتِ
هُمْ	هُمَا	هُوَ
هُنَّ	هُمَا	هِيَ

Figure (2.20): Exemple sur Pronoms personnels.

➤ **Temps et modes (Zaman)**

Les trois temps (passé, présent et futur) ne figurent pas dans la grammaire arabe. Les actions, exprimées par des verbes, ont cependant deux aspects. Si l'action est achevée, le verbe se met à l'accompli : alMādhī, l'équivalent du passé simple ou le passé composé français. Si l'action est en cours, le verbe se met à l'inaccompli : al-Mudhāri', équivalent du présent de l'indicatif. Chaque aspect possède plusieurs modes.

- **Infinitif :**

L'infinitif proprement parler n'existe pas en arabe. Pour l'exprimer, on fait appel à une tournure particulière. On donne le verbe conjugué à l'accompli ensuite à l'inaccompli à la troisième personne du singulier, masculin.

Pour dire : **écrire** : on dira : **كَتَبَ / يَكْتُبُ**

Figure (2.21): Exemple sur Infinitif.

- **Racine :**

La racine ou le Jidhr est la matière première dont se forment les mots (verbes et noms) en arabe. Le plus souvent, à partir d'une racine trilitère, on forme des mots selon un schème (= moule) précis pour désigner un sens. Par exemple, à partir de la

racine: **ك / ت / ب** .l'on forme le verbe écrire, **كَتَبَ** selon le schème des verbes de la première forme, en l'occurrence : **فَعَلَ / يَفْعُلُ** .

- **Accompli:**

L'accompli, correspondant au passé composé français/ passé simple, se construit en mettant des désinences verbales à la fin de la racine trilitère.

جلس الطالب
L'étudiant s'est assis

Figure (2.22): Exemple 1 sur Accompli.

Singulier
أنا جَلَسْتُ
أنتِ جَلَسْتِ
أنتِ جَلَسْتِ
هو جَلَسَ
هي جَلَسَتْ

Figure (2.23): Exemple 2 sur Accompli.

- Inaccompli:

Le deuxième mode est l'inaccompli ou le Mudāri' marfū' **المضارع المرفوع**

Il exprime une action inachevée, une action qui dure ou qui se répète. Le plus souvent, il correspond au présent de l'indicatif français. Ce mode est marqué à la fois par la préfixation d'éléments pronominaux et par des désinences verbales aidant à déterminer, aussi, le genre et le nombre.

أنا أَجْلِسُ
أنتِ تَجْلِسُ
أنتِ تَجْلِسِينَ
هو يَجْلِسُ
هي تَجْلِسُ

Figure (2.24): Exemple sur inaccompli.

- Futur:

Le futur ou le mustaqbal, proprement dit, n'existe pas en arabe. Pour l'exprimer, on fait appel à deux particules, chacune suivie du verbe conjugué à l'inaccompli ; la nuance entre le futur proche et lointain étant plutôt vague.

a- Futur proche

verbe+س سأكتب (J'écrirai).

Figure (2.25): Exemple sur futur proche.

b- Futur lointain

verbe +سوف : ex. سوف أكتب (Je vais écrire).

Figure (2.26): Exemple sur futur lointain.

➤ Négation (nafy)

La négation ou le nafy varie en arabe selon que la phrase est nominale ou verbale. Dans la phrase verbale, la particule de négation varie également selon le mode et l'aspect du verbe. Ici nous nous limitons à la négation des verbes accomplis et inaccomplis :

- **Négation au passé** : Elle se construit en faisant précéder le verbe conjugué à

ما

l'accompli par :

ما كتبتَ الدرسَ

Figure (2.27): Exemple sur Négation au passé.

- **Négation au présent** : Elle se construit en faisant précéder le verbe conjugué à

لا

l'inaccompli par :

لا أكتبُ الدرسَ

Figure (2.28): Exemple sur Négation au présent.

➤ **Formes dérivées**

La langue arabe est une langue **de dérivation** الاشتقاق Les verbes sont dérivés d'une racine trilitère (composée de trois lettres ou radicales). À partir de ces trois radicales, on construit un verbe de la forme I الفعل المجرد.

Il est possible d'augmenter cette forme I de quelques lettres pour obtenir des verbes dérivés/augmentés فعل مزيد Le sens de ces derniers demeure proche de celui de la racine.

➤ **Interjection**

L'interjection ou le nidā' est un procédé syntaxique qui consiste à faire précéder le nom d'une personne par un interpellatif

C'est à dire une particule qui influence le cas du nom. Si ce nom est simple, il prend le cas sujet.

يا أحمدُ Oh ! Ahmad !

Figure (2.29): Exemple 1 sur Interjection.

Si ce nom est annexé, le premier terme prend le cas direct:

يا صاحب البيتِ , Oh toi, maître de la maison !

Figure (2.30): Exemple 2 sur Interjection.

Bien entendu, plusieurs autres particules existent et expriment différentes charges affectives comme **وا، أيا، أيها، آ...**

Trois cas : Les trois cas ou les harkāt sont les voyelles finales qui prennent les mots dans une phrase. Ces cas marquent les fonctions grammaticales. En arabe, il est trois cas :

- Le cas direct, la fatha, notamment pour marquer le complément d’objet direct.
- Le cas indirect, la kasra, notamment pour marquer le complément indirect. Ce

cas est souvent impliqué par les prépositions: **عَنْ، عَلَى، فِي، ب، لِ، ك**

- c- Le cas sujet, la dhamma, notamment pour marquer le sujet.

➤ **Types de phrases**

La grammaire arabe distingue entre deux types de phrase :

- a- **Phrase nominale** : sujet + attribut/prédictat:

L'étudiant est beau. **الطالب جميل**

Figure (2.31): Exemple sur la phrase nominale.

- b **-Phrase verbale** : verbe+ sujet+ (ou sans) complément d’objet (direct ou indirect):

جلس الطالب *L'étudiant s'est assis.*

Figure (2.32): exemple sur la phrase verbale.

➤ **Conjonctions de coordination ('atf)**

La conjonction de coordination ou le ‘atf est un procédé syntaxique qui permet de coordonner deux noms ou deux phrases. Les trois conjonctions de coordination les plus connues sont :

- **ف** exprimant un laps de temps très bref entre les deux actions:

دَخَلَ فَجَلَسَ

Figure (2.33): Exemple1 sur conjonctions de coordination.

; il est entré et [toute de suite] s’est assis.

- **واو** exprimant un laps de temps moins bref

دَخَلَ وَجَلَسَ

Figure (2.34): Exemple 2 sur conjonctions de coordination.

; il est entré et s’est [un peu de temps après] assis.

- **ثم** exprimant un laps de temps plutôt long entre les deux actions.

دَخَلَ ثُمَّ جَلَسَ

Figure (2.35): Exemple 3 sur conjonctions de coordination.

; il est entré [ensuite] il s’est assis.

- **Participe actif (ism al-fā‘il)**

Le participe actif ou ism al-fā'il désigne celui qui fait l'action.

Pour tous les verbes de la première forme, on construit le participe actif sur

le schème **فَاعِلٍ** .

à partir de **كَتَبَ/يَكْتُبُ** (écrire), on obtient **كَاتِبٍ** (écrivain)

Figure (2.36): Exemple sur participe actif.

➤ **Participe passif (ism al-maf'ūl)**

Le participe passif ou ism al-mf'ūl désigne celui qui subit l'action.

Pour tous les verbes de la première forme, on construit le participe passif sur le

schème **مَفْعُولٍ** .

à partir de **كَتَبَ/يَكْتُبُ** (écrire), on obtient **مَكْتُوبٍ** (chose écrite)

Figure (2.37): Exemple sur participe passif.

➤ **Non d'action (masdar)**

Le nom d'action ou le masdar est un substantif dérivé qui désigne l'action. La formation des noms d'action de la forme I ne suit aucun schème.

قِرَاءَةً قَرَأَ lire, lecture.
كِتَابَةً كَتَبَ écrire, écriture

Figure (2.38): Exemple sur non d'action.

Cependant, les noms d'action des 9 formes (verbales) dérivées suivent tous des schèmes constants.

2.3 Le contexte

2.3.1 Définition du contexte

Il existe dans la littérature plusieurs définitions du mot <contexte> suivant les domaines de l'ingénierie. Au regard de la littérature informatique, cette notion a trouvé son essor dans le domaine de l'informatique pervasive ou ambiante où elle a suscité l'intérêt de plusieurs chercheurs [21].

La définition la plus reconnue dans la littérature et proposée par **Dey**, il définit le contexte comme étant : "toutes les informations pouvant être utilisées pour caractériser la situation d'une entité, où une entité est une personne, un lieu, ou un objet qui peut être pertinent pour l'interaction entre l'utilisateur et l'application, y compris l'utilisateur et l'application eux-mêmes[22] "

2.3.2 Les catégories de contexte

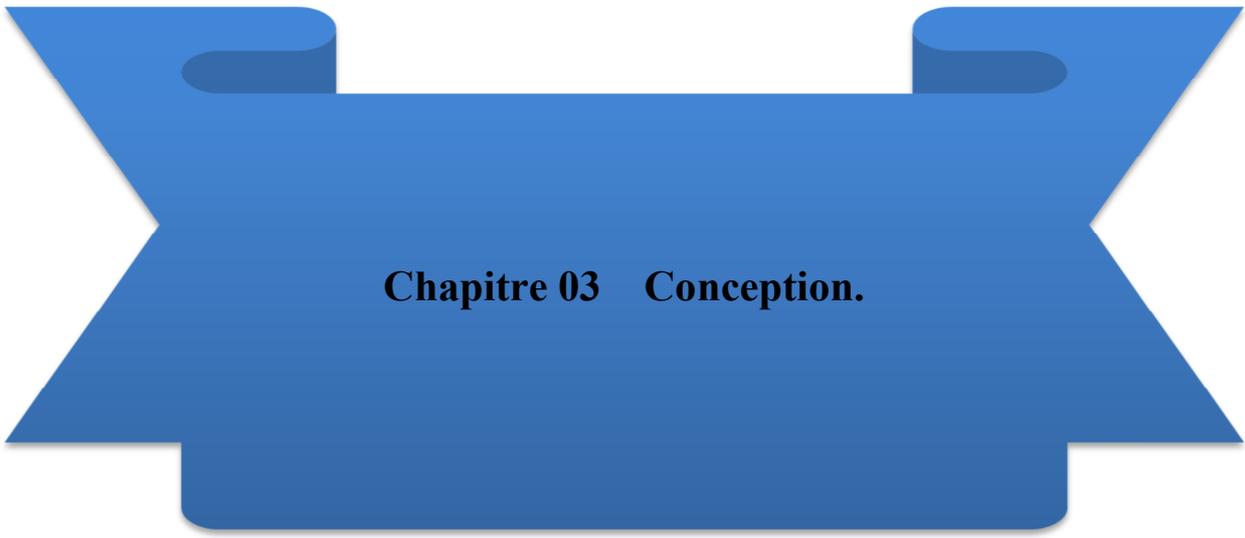
Selon **Tigli** [23] quatre familles de contexte peuvent être identifiées. chacune de ces familles apporte sa propre définition du contexte.

- **Notion de contexte géolocalisé** : la vision de **Schilit** [24] pour le contexte focalise sur la géolocalisation de l'utilisateur. de ceci découlent trois types d'information du contexte.
 - la localisation de l'utilisateur <où est l'utilisateur>.
 - les personnes se trouvant à proximité de l'utilisateur <avec qui ?>.
 - les dispositifs et les ressources que l'utilisateur a à sa disposition <quelles sont les ressources utilisés>
- **Notion de contexte environnementale** : **Brown** [25] relie le contexte aux éléments de l'environnement de l'utilisateur. les autres ajoutent à leur première définition des entités telles que l'heure, la saison, la température, l'identité et la localisation de l'utilisateur. en effet, ce contexte détermine, soit le comportement d'une application, soit les événements issus de l'application et qui intéressent l'utilisateur.

- **Notion de contexte généralisé** : Certains auteurs ont proposé une définition générale du contexte. **Gong [26]** définit le contexte comme étant ce qui entoure et donne du sens à quelque chose d'autre. Tandis que Winograd [27] définit le contexte comme <l'ensemble d'informations structuré et partagé qui évolue et sert l'interprétation>.

2.4 Conclusion

Dans ce chapitre, on parle d'une façon générale de la grammaire dans les deux langues (arabe, français) et du contexte, pour les utiliser dans le chapitre suivant, qui parlera de la conception de notre système.

A blue decorative graphic consisting of a central horizontal bar with a rectangular cutout in the middle. The bar has a 3D effect with a darker blue shadow on the top and bottom surfaces. The text "Chapitre 03 Conception." is centered within the white rectangular cutout.

Chapitre 03 Conception.

3.1 Introduction

La conception est le processus qui consiste à représenter les diverses fonctions du système, c'est certainement la partie la plus importante de notre travail [15].

Dans ce chapitre nous allons présenter la conception de notre système de traduction automatique d'un texte en la langue française vers un texte en la langue arabe (traduction unidirectionnelle) de type fichier TXT avec l'utilisation d'un dictionnaire médical et la méthode de la traduction automatique à base des règles en mode transfert.

Cette conception est présentée par une architecture globale et une architecture détaillée où nous allons expliquer en détail le fonctionnement de notre système.

3.1 Conception globale

Dans cette section, nous allons décrire et présenter le principe et l'architecture générale de notre système.

3.2.1 Architecture générale

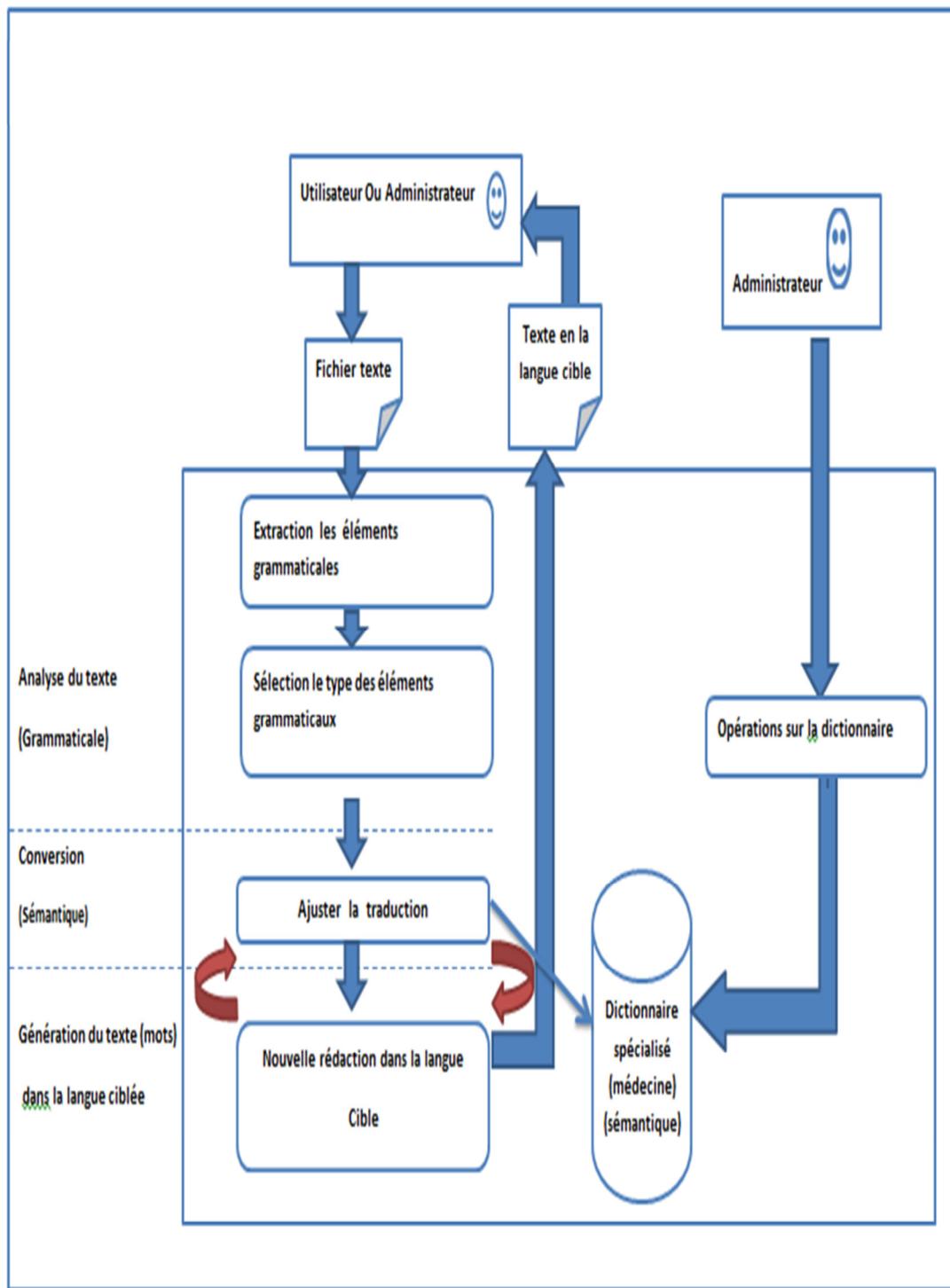


Figure (3.1): Architecture générale de notre système.

Le Figure (3.1) montre la décomposition de système en quatre modules qui sera mentionné ci-dessous où l'administrateur peut traverser le module qui permet des opérations sur base de données mais utilisateur non .

3.2.1.1 Module des opérations sur la base de données

A cette étape, l'administrateur (expert) effectue des opérations sur la base de données telles que suppression, l'ajout et la mise à jour,... etc. il doit avoir de l'expérience dans un domaine (médecine).

3.2.1.2 Module d'analyse du texte

Dans cette étape le contenu de fichier texte (.Txt) est analysé de manière grammaticale à travers l'utilisation des deux sous-étapes suivantes :

- L'étape de l'Extraction des éléments grammaticaux (phrases, mots).
- L'étape de la détermination des types des éléments grammaticaux (phrases nominale, phrase verbale, adjectif, nom, verbe ... etc.).

3.2.1.3 Module de conversion

L'avantage de cette étape est de convertir les mots dans la langue cible afin de préserver le sens sémantique qui leur est attribué à l'aide d'un dictionnaire spécialisé de médecine (synonyme (déduction), le sens de mot dans la langue cible dans le dictionnaire).

3.2.1.4 Module de génération du texte

C'est la dernière étape dans notre système qui permet d'ordonner et d'assembler les phrases et les mots de la langue cible (pronom relatif + nom, pronom relatif +verbe et adjectif + nom ...etc.), à fin d'obtenir le résultat final (texte dans la langue cible).

Remarque : ce module persiste dans la circulation avec le module précédent jusqu'à la fin des mots du texte.

3.3 Conception détaillée

Dans cette section nous allons essayer de détailler les différents modules composants ce système .Les figures suivants présentent l'architecture détaillée de notre système.

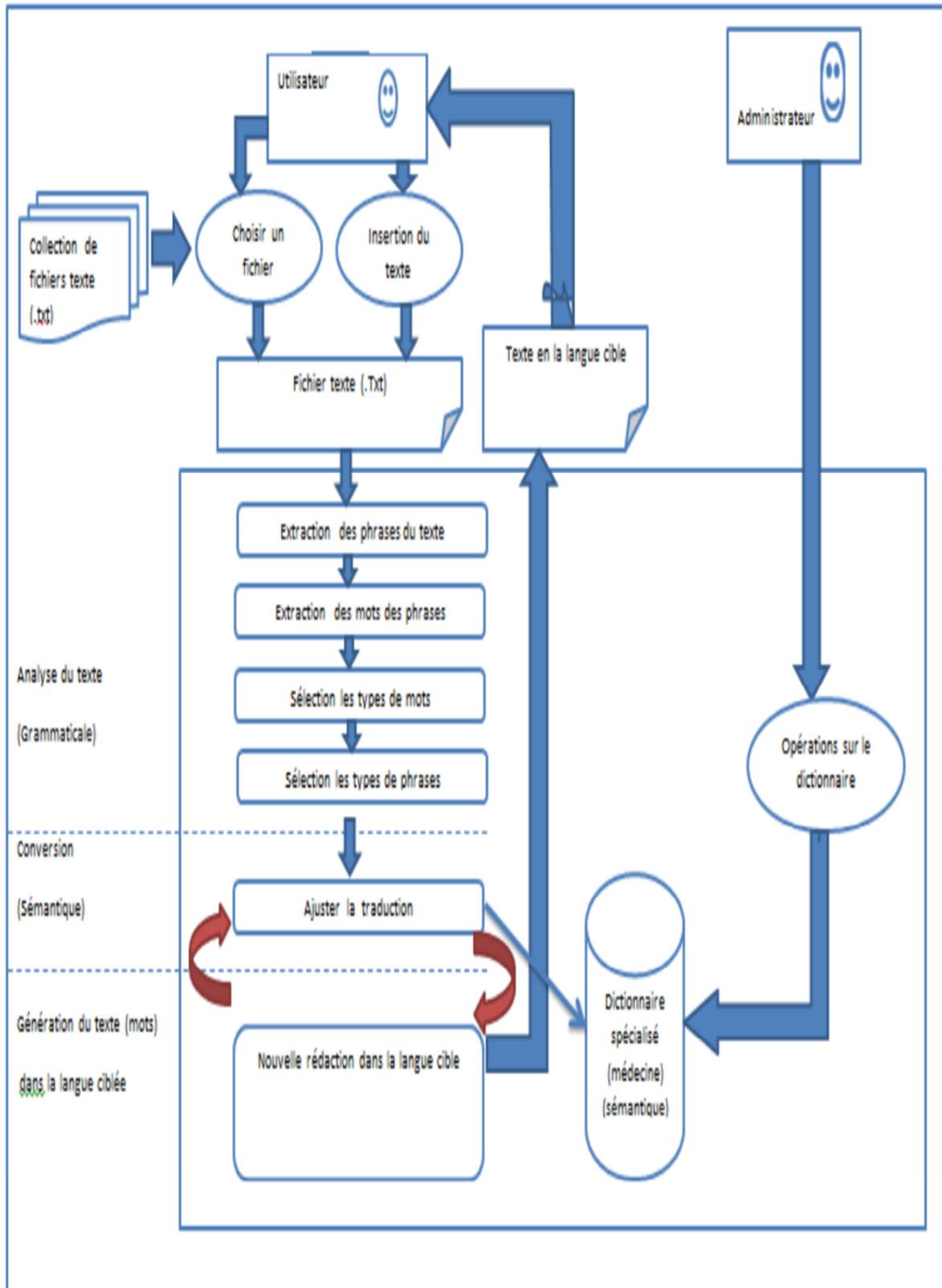


Figure (3.2) : Architecture détaillée de notre système.

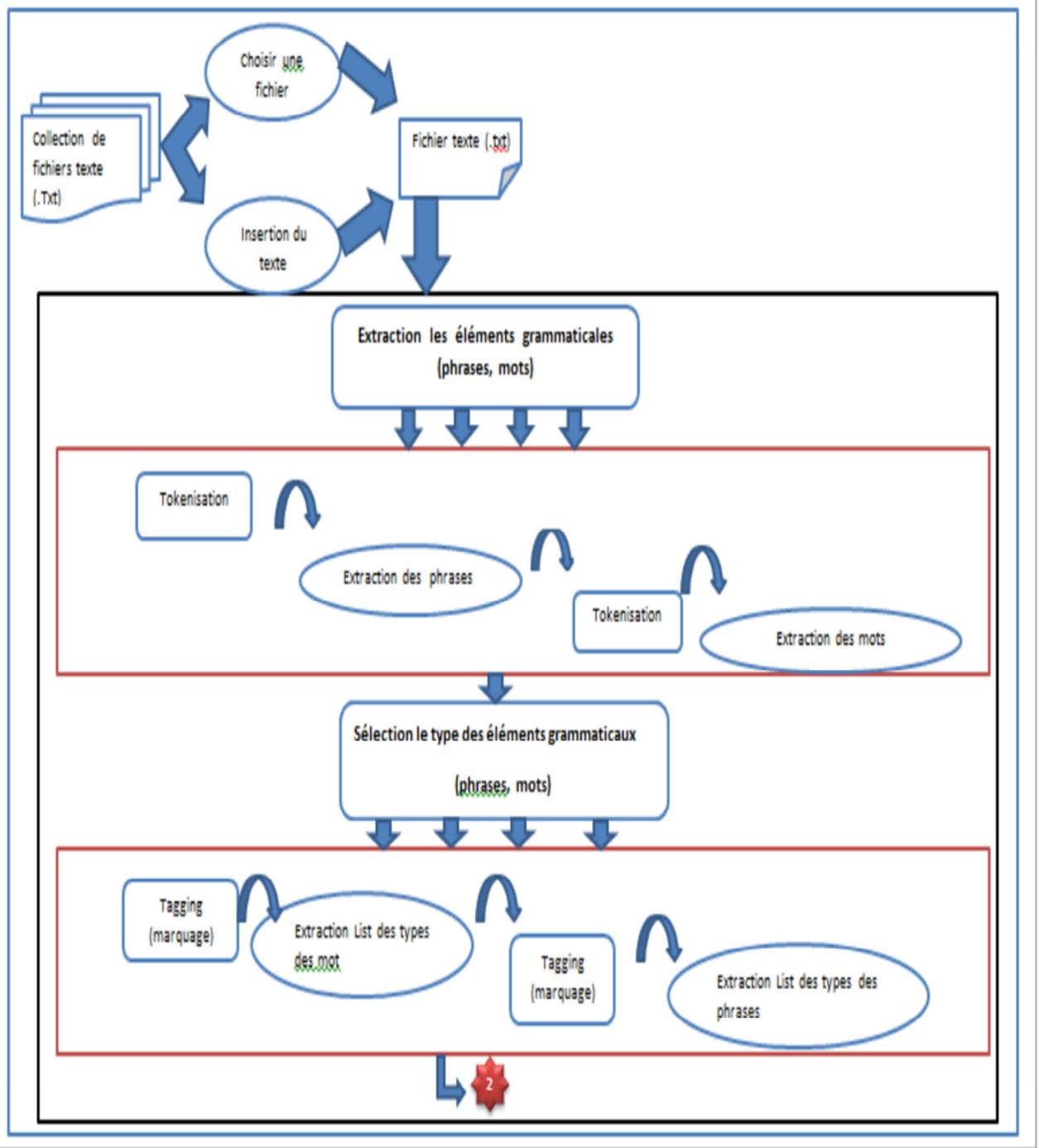


Figure (3.3):Partie 1 d'architecture détaillée de notre système.

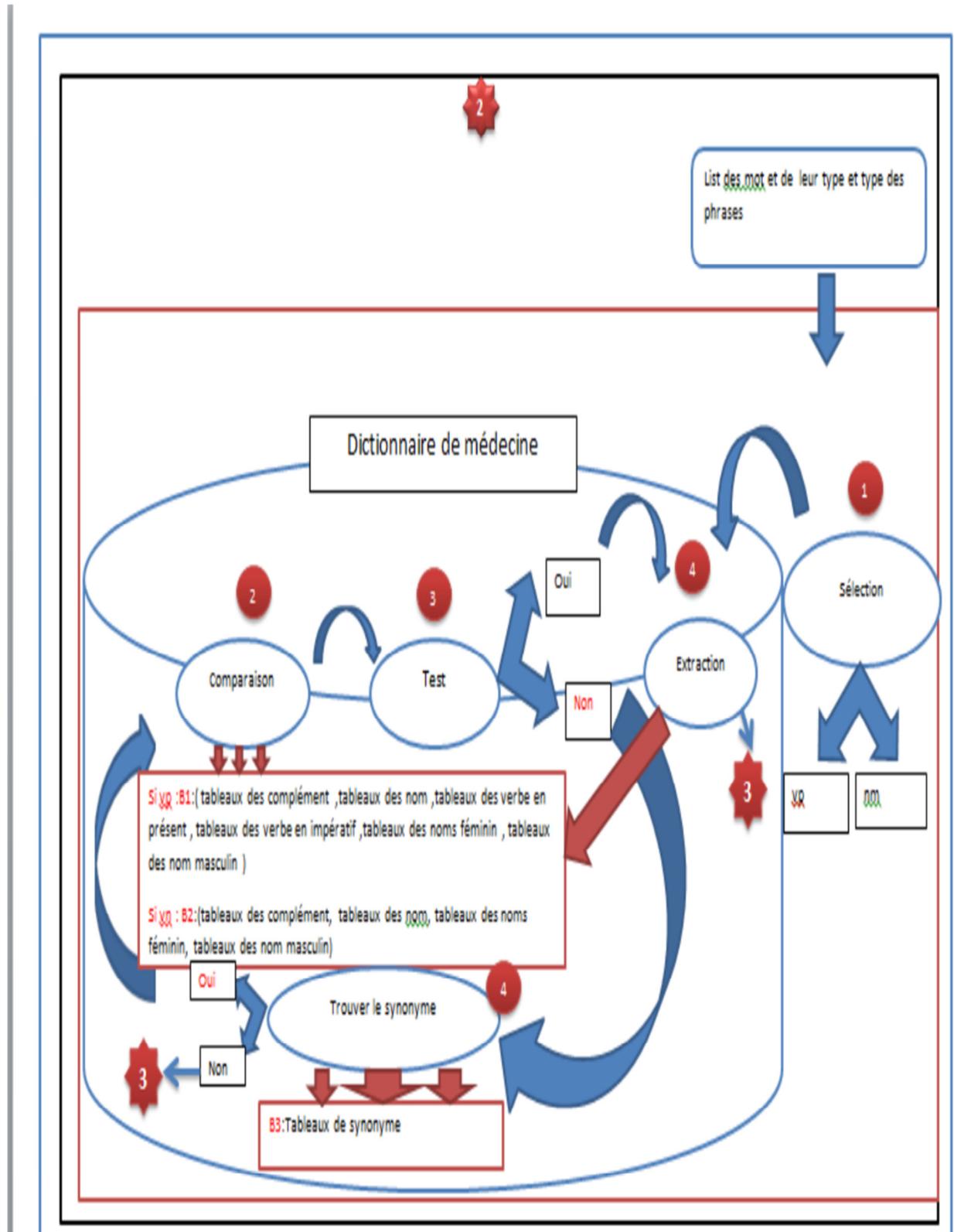


Figure (3.4):Partie 2 d'architecture détaillée de notre système.

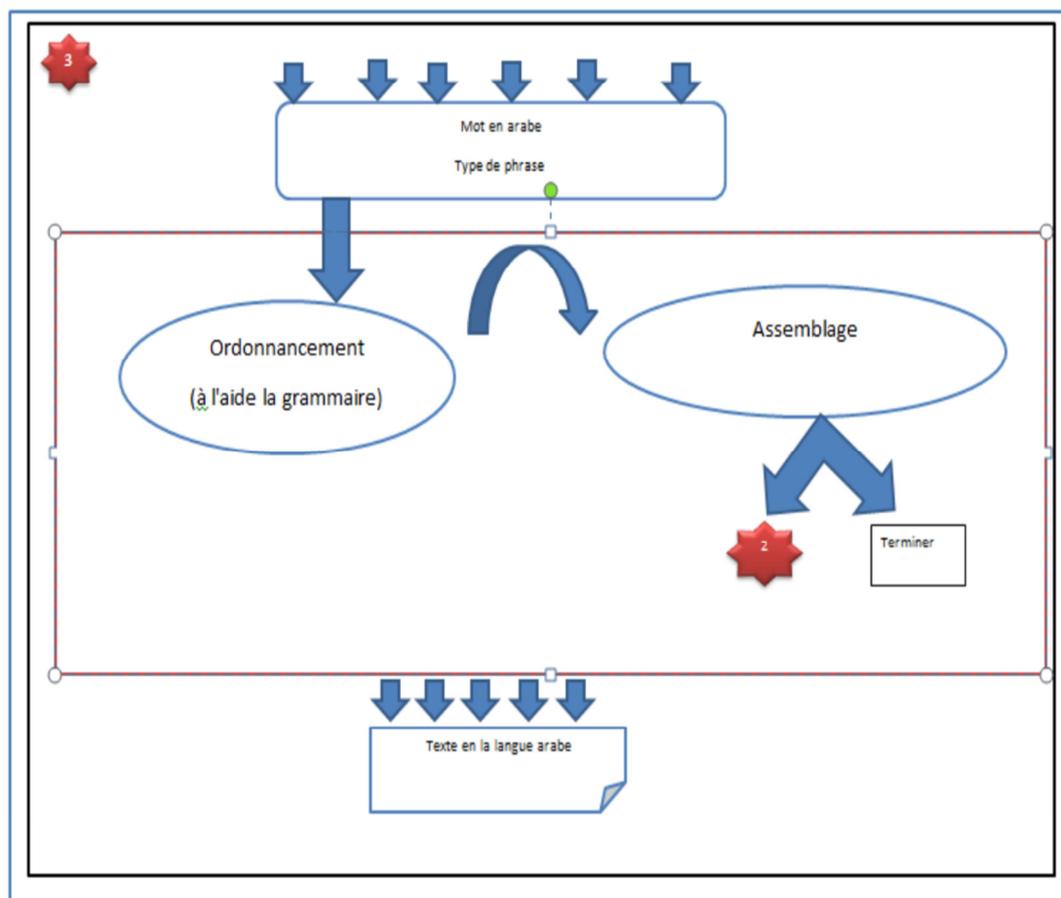


Figure (3.5):Partie 3 d'architecture détaillée de notre système.

3.3.1 L'étape préparatoire

À cette étape, le texte à traduire est importé en choisissant l'une des deux méthodes suivantes :

- Choisir un fichier:
La sélection du texte défini (format .Txt).
- Insertion du texte :
La saisie d'un texte indéfini.

3.3.2 L'étape d'analyse

Elle contient les deux sous-étapes suivantes:

3.3.2.1 Extraction des éléments grammaticaux (phrases, mots)

Dans cette étape, l'exploitation de texte extrait de l'étape précédente : nous extrayons des phrases du texte et des mots de la phrase par Tokenisation pour obtenir une liste des mots et une liste des phrases.

- **Extraction des phrases** : nous savons que dans la grammaire française la phrase commence par une lettre majuscule et se termine par un point et d'après cette règle, nous avons divisé le texte et extraire les phrases, illustré dans la Figure (3.6).

Ma langue arabe est ma belle langue.

Figure (3.6): Exemple illustratif sur l'analyse grammaticale pour trouver les phrases constrictives de texte en la langue française (langue source).

Remarque : nous avons changé la règle de départ (la phrase commence par une lettre majuscule) par une autre règle qui dit (la phrase commence par une lettre majuscule ou minuscule) illustré dans la Figure (3.7).

ma langue arabe est ma belle langue.

Figure (3.7): Exemple illustratif sur le changement de la règle du départ.

- **Extraction des mots** : nous savons aussi que la phrase est un ensemble des mots et que chaque mot est séparé de l'autre par un séparateur (espace), de cette façon les phrases extraites sont divisées en mots, illustré dans la Figure (3.8).

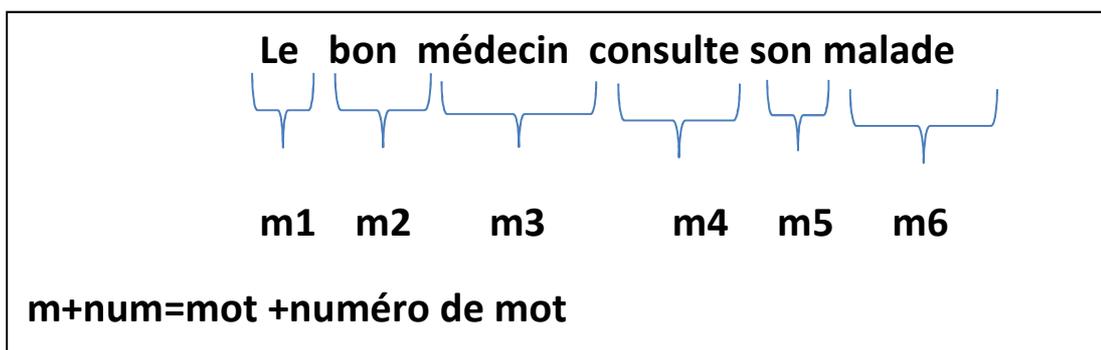


Figure (3.8): Exemple illustratif de l'extraction des mots de la phrase.

3.3.2.2 Sélection des types des éléments grammaticaux (phrases, mots)

Dans cette étape nous utilisons les sorties de l'étape de l'extraction déjà mentionnée.

Au début, nous sélectionnons les types des mots des phrases puis les type des phrases par Tagging (marquage).

- **Extraction de la liste des types des mots :**

On doit préciser chaque type de mot (adjectif, nom, verbe, pronom connecter, ponctuation, complément, déterminent ...etc.) .illustré dans la Figure (3.9).

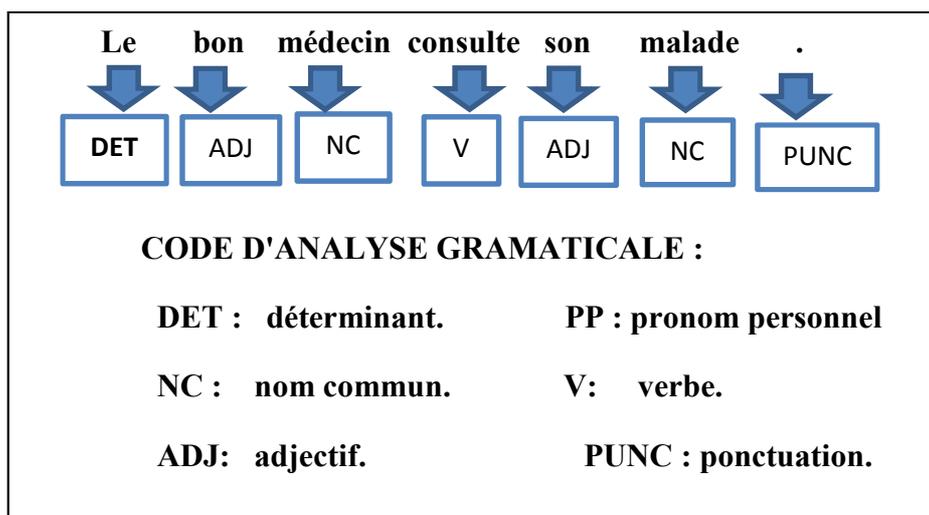


Figure (3.9): Exemple illustratif sur l'analyse grammaticale pour trouver les types des mots.

- **Extraction liste des types des phrases :**

Dans cette dernière étape d'analyse grammaticale, nous exploitons la liste des types des mots pour extraire le type de chaque phrase de l'étape précédente et l'additionner à la liste des mots et leur type.

Si la liste des mots contient au moins un verbe alors la phrase est une phrase verbale si non c'est une phrase nominale, illustré dans Figure (3.10).

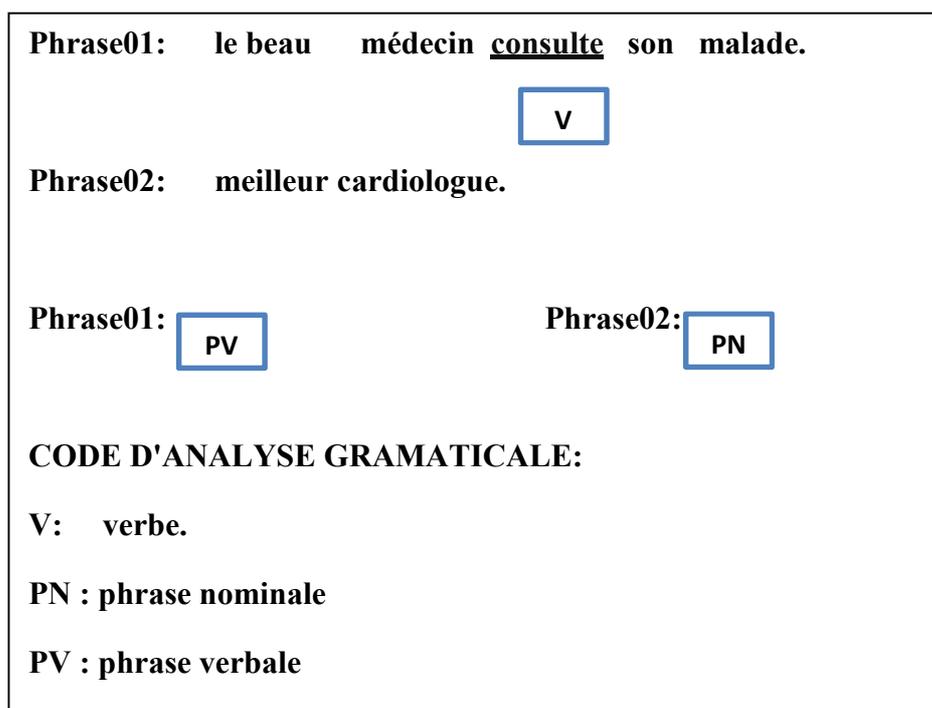


Figure (3.10): Exemple illustratif sur l'analyse grammaticale pour trouver les types des phrases.

3.3.3 L'étape de la conversion

On converti la sortie (List des mots, types des mots et types des phrases) de l'étape d'analyse à la langue cible, en utilisant un dictionnaire médical. Pour plus d'illustration on suit les étapes suivantes :

3.3.3.1 L'étape de la Sélection

Dans cette étape, nous appliquons certaines opérations sur la Liste des mots, types des mots et types des phrases :

- La **Sélection de la phrase** : on sélectionne une phrase (on conservant le reste jusqu'à la fin de la conversion et la génération de la phrase spécifiée), puis on sélectionne le type de la phrase à fin de déterminer les tableaux de dictionnaire à utiliser.
- La **Sélection les type des mots et les mots de la phrase** :
Après l'étape précédente, on sélectionne les mots et leurs types de la phrase choisie. On appelant la sortie de cette étape : <liste temporaire>.
- La **Sélection d'un mot et son type** :
On sélectionne un mot et son type à partir de la liste temporaire (on conserve le reste jusqu' à la fin de la conversion et la génération du mot spécifique).

3.3.3.2 L'étape de la comparaison

On compare le mot obtenu de l'étape précédente avec son équivalent dans le dictionnaire médical (en utilisant les tableaux spécifiques du dictionnaire médical avec le respect des conditions de l'utilisation des tableaux), puis on doit passer par l'étape du test pour choisir la prochaine étape.

➤ les conditions de l'utilisation des tableaux sont les suivantes :

- si le type de phrase est nominal, on n'utilise pas les tableaux spécifiques pour les verbes et synonymes.

- si le type de phrase est verbal, on n'utilise pas le tableau spécifique uniquement pour les synonymes.

3.3.3.3 L'étape du test

Si on trouve l'équivalent, la réponse retournée est oui sinon la réponse retournée est non.

3.3.3.4 L'étape de l'extraction

On extrait l'équivalent du mot en la langue arabe, dans cette étape si la réponse de l'étape de test est retournée <oui> (En utilisant les tableaux spécifiques du dictionnaire médical avec le respect des conditions de l'utilisation des tableaux)

Remarque importante : dans cette étape, on essaye de faire entrer la sémantique avec l'utilisation d'un dictionnaire spécialisé dans la médecine.

ne pour ajuster le sens exact illustré dans la Figure (3.11).



Figure (3.11): Exemple sur la sémantique dans le cas de l'utilisation du dictionnaire spécialisé dans la médecine.

Dans la figure (3.11) le verbe présente un sens يظهر dans un contexte général mais un sens différent dans un contexte médical يشكو.

3.3.3.5 L'étape de la recherche du synonyme

Dans le cas où la réponse de l'étape de test est retournée <non>, nous cherchons le mot en la langue française dans le dictionnaire médical précisément dans le tableau des synonymes pour trouver son synonyme :

- Si on ne trouve pas le mot en la langue française, on passe à l'étape suivante.
- Si on trouve le mot en la langue française, on va chercher son synonyme dans les autres tableaux du dictionnaire médical (avec le respect des conditions de l'utilisation des tableaux (déjà vues)):
 - Si on trouve le synonyme, on retourne à l'étape de la comparaison.
 - Sinon, on passe l'étape suivante.

Remarque importante : Dans cette étape, on essaye de faire entrer la sémantique (l'utilisation des synonymes (déduction) et le processus précédant pour gérer ces synonymes).

Remarque : Dans le cas où le mot est un verbe, on a deux possibilités :

- en cas où le verbe est dans le temps impératif (la phrase commence par un verbe), on va chercher dans le tableau des verbes impératif.

- en cas où le verbe est dans le temps présent, on va faire une analyse grammaticale à fin de trouver des solutions aux problèmes suivants :

Pb1: si le verbe est précédé par un pronom relatif, on va chercher le sujet (qui conjugue le verbe).

Nous avons un médecin qui aime ses malades
Pronom + verbe+ déterminant + nom + pronom
relatif +verbe + adjectif + nom

Figure (3.12): Illustration du problème en cas où le verbe est précédé par un pronom relatif.

Solution1: on conjugue le verbe avec le nom qui précède juste le pronom relatif, on utilisant le tableau des verbes au présent à condition de préciser le genre de nom (féminin ou masculin) (l'utilisation de tableaux des noms féminins et masculins).

Nous avons un médecin qui aime ses malades.
Pronom + verbe+ déterminant + nom + pronom
relatif +verbe + adjectif + nom +ponctuation

Figure (3.13): Illustration de la solution en cas où le verbe est précédé par un pronom relatif.

3.3.4 L'étape de la génération

Après la conversion de mot, on passe à la phase de la reformulation de ce mot.

3.3.4.1 L'étape de la reformulation

3.3.4.1.1 L'étape de l'ordonnancement

On ordonne le mot de la phrase à partir des règles grammaticales de la langue cible.

D'une autre façon quel est le meilleur placement pour mettre le mot dans la liste des mots de la phrase cible (la langue arabe), on utilisant les règles grammaticales de la langue cible ? ,pour détailler ,Nous donnons les exemples suivants :

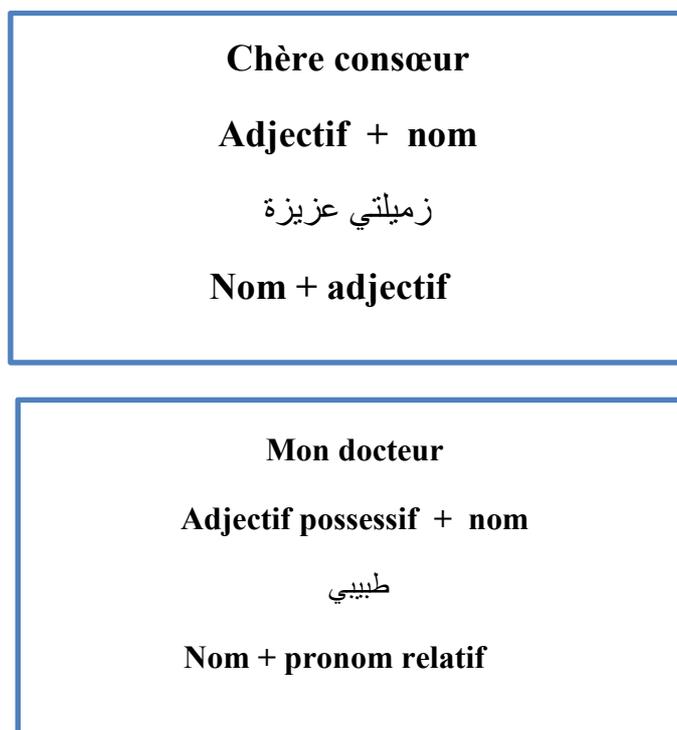


Figure (3.14): Exemples d'un cas d'un adjectif avec un nom.

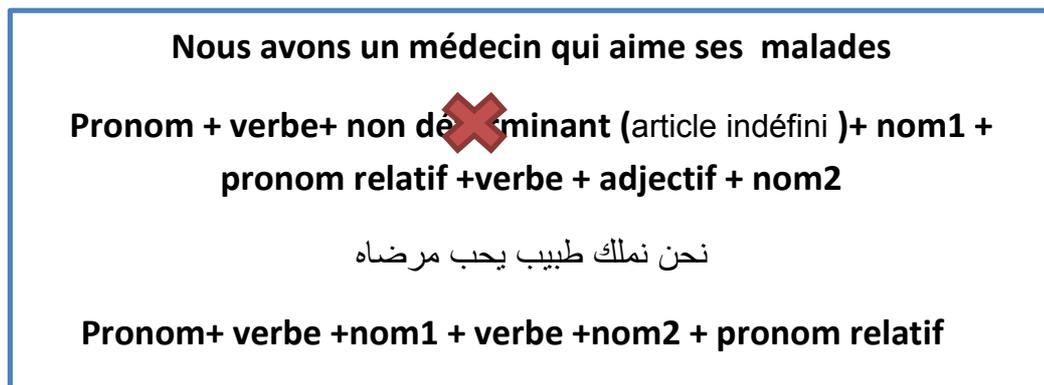


Figure (3.15): Exemple d'un cas d'une élimination d'un article indéfini (un, une).

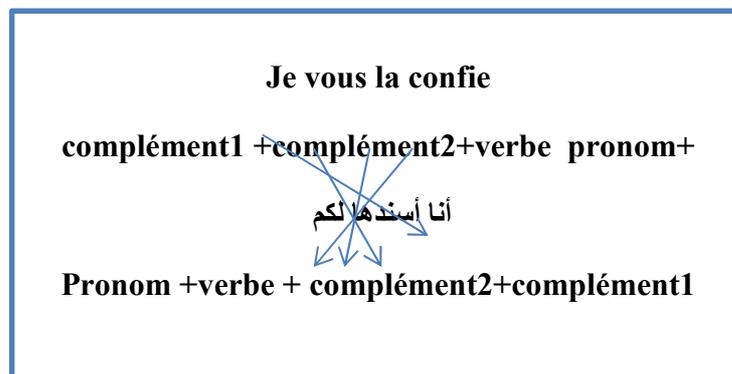


Figure (3.16): Exemple4 en cas de complément.

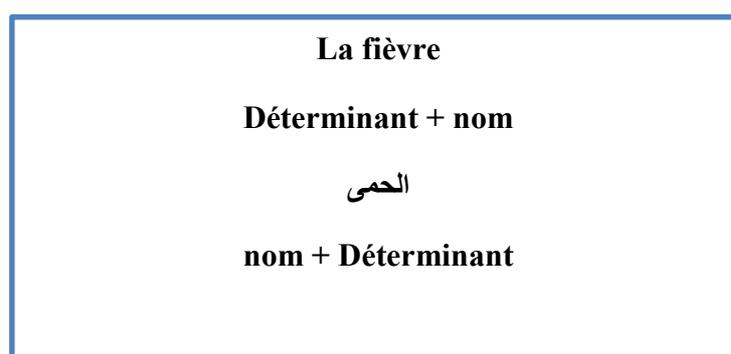


Figure (3.17): Exemple 5 En cas de déterminant et nom (la, le, les, des, de...Etc.).

3.3.4.1.2 L'étape de l'assemblage

- cette dernière étape est caractérisée par la mise du mot dans sa place (trouvée dans l'étape précédente).

Ensuite,

- Si la liste temporaire n'est pas vide, on retourne à la phase de Sélection d'un mot et leur type.
- Si la liste temporaire est vide, on retourne à la phase de Sélection d'une phrase jusqu'à la fin de toutes les phrases de texte, en fin le retourne en la langue cible.

3.3.5 Module des opérations sur la base de données

Dans cette partie, l'utilisateur effectue des opérations sur la base de données comme suivant :

3.3.5.1 Opération de la suppression

Pour faire une suppression dans la base de données, il faut d'abord identifier un id (primary) puis supprimer tous les mots qui portent le même id du même tableau.

3.3.5.2 Opération de l'ajout

Pour faire un ajout d'un mot dans la base de données, il faut identifier les suivants :

- un identificateur (id primary).
- Si on a le tableau des verbes en présent, on identifie le verbe en français, le verbe en arabe et le pronom.

Sinon

Si on a le tableau des synonymes, on identifie le mot en français (synonyme1) et le mot en français (synonyme2).

Si on a un autre tableau, on identifie le mot en français et le mot en arabe.

3.3.5.3 Opération de la mise à jour

Pour faire une mise à jour d'un mot dans la base de données, il faut identifier les suivants :

- un identificateur (id primary).
- Que veux-tu changer ?

Si on veut changer les synonymes, il faut identifier le synonyme 1 ou le synonyme 2 par l'utilisation 1 du tableau des synonymes.

Sinon

-Si on veut changer les verbes en présent, il faut identifier le verbe en français, le Verbe en arabe ou le pronom par l'utilisation du tableau des verbes en présent.

-Sinon il faut identifier le mot en français ou le mot en arabe par l'utilisation du tableau (dictionnaire, féminin ou masculin).

3.3.5.4 Opération de l'authentification

Elle contient trois opérations :

- **L'opération de la connexion** : pour entrer au système et effectuer des opérations sur la base de données, il faut d'abord faire entrer des informations sur l'expert (numéro de certificat (primary) et mot de passe).
- **L'opération de la déconnexion** : pour quitter notre système
- **L'opération de l'inscription** : pour effectuer l'opération de la connexion, il faut d'abord inscrire l'expert (médecin) par quelques informations importantes sur ce dernier, comme :
 - nom
 - email : à fin de se connecter au moment nécessaire.
 - numéro de certificat médical : qui est primary à fin de s'assurer qu'il est un expert (médecin).
 - mot de passe : pour la sécurité.
 -

3.4 Modélisation de système

Cette partie est consacré à l'étape de modélisation, pour cela on va utiliser le formalisme d'UML (Unified Modeling Language) pour la modélisation de notre système .

3.4.1 Diagrammes de cas d'utilisation

Les diagrammes de cas d'utilisation sont des diagrammes UML utilisés pour donner une vision globale du comportement fonctionnel d'un système logiciel. un cas d'utilisation décrit une séquence d'actions réalisées par le système qui produit un résultat observable pour un acteur [25].

3.4.1.1 Identification des acteurs

Les deux acteurs de notre système sont les suivants :

- **L'administrateur:**

L'administrateur c'est l'expert de la base de données qui dispose de toutes les fonctionnalités nécessaires à la gestion de la base de données (ajout, suppression, mise à jour...etc.).

L'administrateur de notre système assure les fonctionnalités suivantes:

- Authentification (connexion, inscription, déconnexion).
- Gérer la base de données (suppression, ajout, mise à jour... etc.).
- Traduction (analyse grammaticale, conversion, génération).
-

- **L'utilisateur :**

- Traduction (analyse grammaticale, conversion, génération).

3.4.1.2 Identification des cas d'utilisation

Chaque un des acteurs précédents a des tâches précises dans le système.

Le diagramme de cas d'utilisation qui décrit les fonctionnalités de l'administrateur, voir la Figure (3.18):

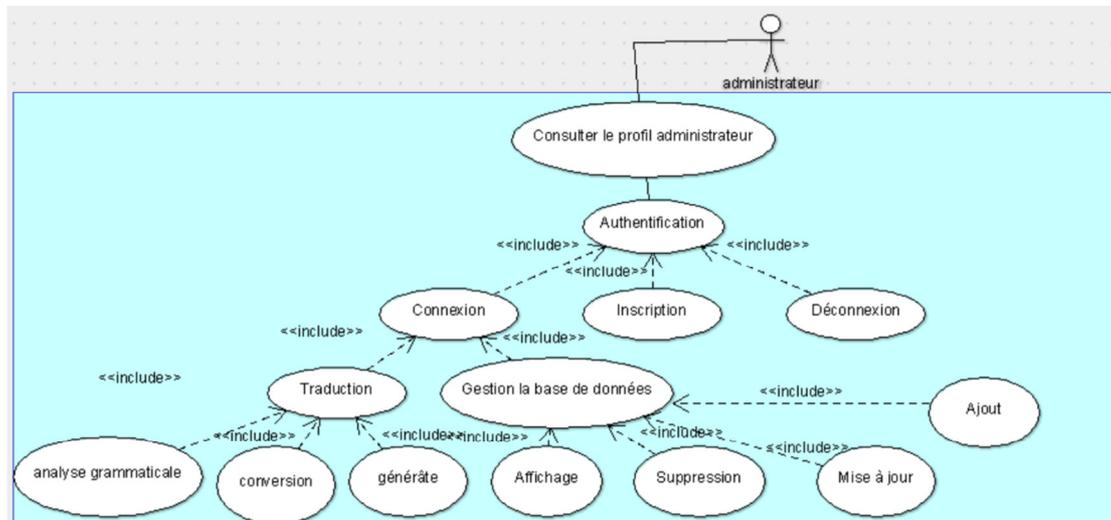


Figure (3.18): Diagramme de cas d'utilisation de l'administrateur du système.

Le diagramme de cas d'utilisation qui décrit les fonctionnalités de l'utilisateur, voir s la Figure (3.19):

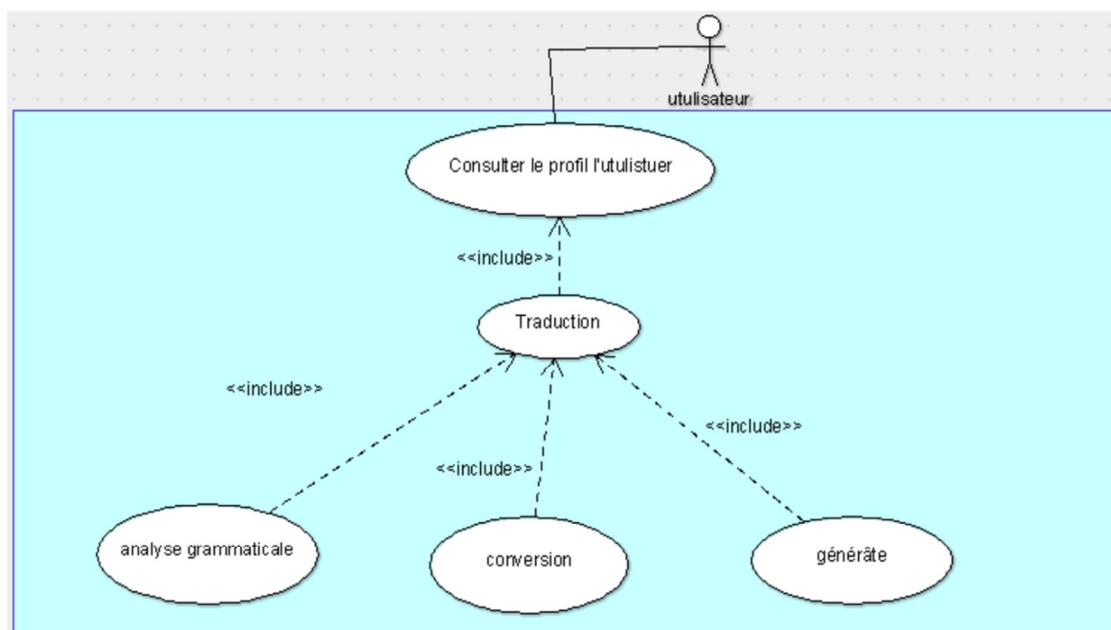
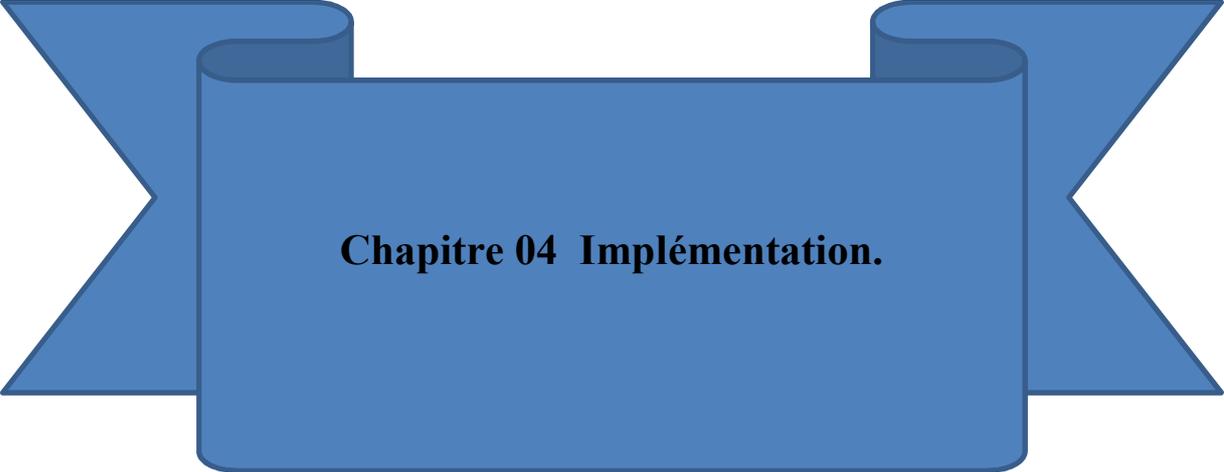


Figure (3.19): Diagramme de cas d'utilisation de l'utilisateur du système.

3.5 Conclusion

Dans ce chapitre, nous avons décrit les besoins fonctionnels et techniques attendus de notre système. Nous avons détaillé le fonctionnement de tous les processus qui constituent ce système.

Dans le chapitre suivant nous allons présenter l'implémentation de tous les composants et les modules de notre système.



Chapitre 04 Implémentation.

4.1 Introduction

Après l'étude conceptuelle, nous allons présenter dans ce chapitre la phase de la réalisation et de l'implémentation de notre système.

Ce chapitre a pour objectif de présenter l'aspect implémentation de notre application, il s'agit donc d'expliquer l'environnement matériel sur lequel notre système a été développé, les langages de programmation et les outils utilisés. Ensuite, nous allons présenter les interfaces graphiques en décrivant les différentes fonctionnalités de notre application et nous présenterons deux exemples qui nous permettront d'illustrer les résultats obtenus lors de l'utilisation de notre application [15].

4.2 Langage de programmation

4.2.1 python [16]

Le langage de programmation Python a été créé en 1989 par Guido van Rossum, aux Pays-Bas. Le nom Python vient d'un hommage à la série télévisée Monty Python's Flying Circus dont G. van Rossum est fan. La première version publique de ce langage a été publiée en 1991. La dernière version de Python est la version 3. Plus précisément, la version 3.7 a été publiée en juin 2018. La version 2 de Python est désormais obsolète et cessera d'être maintenue après le 1er janvier 2020. Dans la mesure du possible évitez de l'utiliser. La Python Software Foundation ¹ est l'association qui organise le développement de Python et anime la communauté de développeurs et d'utilisateurs.

Ce langage de programmation présente de nombreuses caractéristiques intéressantes :

- Il est multiplateforme. C'est-à-dire qu'il fonctionne sur de nombreux systèmes d'exploitation : Windows, Mac OS X, Linux, Android, iOS, depuis les mini-ordinateurs Raspberry Pi jusqu'aux supercalculateurs.

- Il est gratuit. Vous pouvez l'installer sur autant d'ordinateurs que vous voulez (même sur votre téléphone !).

- C'est un langage de haut niveau. Il demande relativement peu de connaissance sur le fonctionnement d'un ordinateur pour être utilisé.

— C'est un langage interprété. Un script Python n'a pas besoin d'être compilé pour être exécuté, contrairement à des langages comme le C ou le C++.

— Il est orienté objet. C'est-à-dire qu'il est possible de concevoir en Python des entités qui miment celles du monde réel (une cellule, une protéine, un atome, etc.) avec un certain nombre de règles de fonctionnement et d'interactions.

— Il est relativement simple à prendre en main.

— Enfin, il est très utilisé en bio-informatique et plus généralement en analyse de données. Toutes ces caractéristiques font que Python est désormais enseigné dans de nombreuses formations, depuis l'enseignement secondaire jusqu'à l'enseignement supérieur.



Figure (4.1): le langage de la programmation python.

4.3 Les environnements de développement

Afin de réaliser notre application, on a utilisé les environnements de développements suivants :

4.3.1 PyCharm [17]

PyCharm est l'IDE le plus populaire pour Python. Il comprend d'excellentes fonctionnalités telles que l'achèvement et l'inspection de code avec un débogueur avancé, la prise en charge de la programmation Web et de divers cadres.

PyCharm est créé par la société tchèque Jet cerveaux qui se concentre sur la création d'un environnement de développement intégré pour divers langages de développement Web tels que JavaScript et PHP .

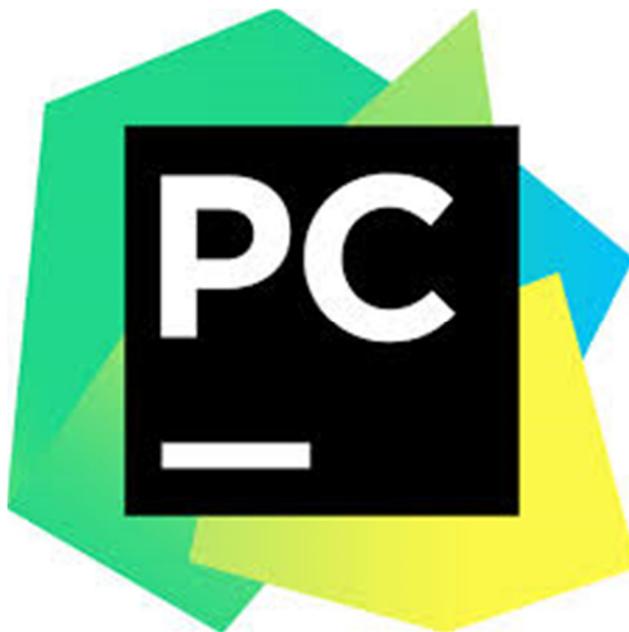


Figure (4.2): la Platform PyCharm .

4.3.2 XAMPP

XAMPP est un ensemble de logiciels permettant de facilement créer une interface web interagissant avec une base de données SQL[18].



Figure (4.3): la Platform XAMPP.

4.3.3 NLPNET

NLPNET est une bibliothèque Python pour les tâches de traitement de langage naturel basées sur des réseaux de neurones. Actuellement, il effectue le marquage d'une partie de la parole, le marquage du rôle sémantique, l'analyse de la dépendance,

Tokenisation , détection de langue, segmentation de phrase ,l'extraction d'entités nommées ,découpage et analyse syntaxique La plupart de l'architecture est indépendante de la langue, mais certaines fonctions ont été spécialement conçues pour travailler avec le portugais. Ce system a été inspiré par SENNA. [19][15].

4.3.4 Tkinter

Le module Tkinter («interface Tk») est l'interface Python standard de la boîte à outils de l'interface graphique Tk et Tkinter sont disponibles sur la plupart des plateformes Unix, ainsi que sur les systèmes Windows.

4.4 Les principales fenêtres de l'application

4.4.1 Les fenêtres de l'accueil

Les figures ci-dessous présentent les fenêtres de l'accueil de notre application.

L'interface d'accueil dans la Figure (4.4) permet l'accès à l'interface principale de la Figure (4.5) par une simple clique sur le Botton (tabib).

Dans la page principale présenté par la Figure (4.5), il y a deux types d'accès:

- l'accès à travers l'administrateur (expert).
- l'accès à travers l'utilisateur.



Figure (4.4):Fenêtre d'accueil de l'application.



Figure (4.5):Fenêtre principale de l'administrateur et de l'utilisateur.

Quand on clique sur le bouton Administrateur de la Figure (4.5), La figure (4.6) apparait.



Figure (4.6):Fenêtre d'accueil de l'administrateur.

Quand on clique sur le bouton Utilisateur de la Figure (4.5), La figure (4.7) apparait.



Figure (4.7):Fenêtre d'accueil de l'utilisateur.

4.4.2 La fenêtre d'authentification de l'administrateur

Quand on clique sur le bouton <bases des données >de La figure (4.6), La figure (4.8) apparait. Dans ce cas l'administrateur doit passer à travers l'authentification par:

- l'opération de la connexion si il a déjà fait l'opération de l'inscription (remplir le champ de numéro de certificat et de mot de passe, puis cliquer sur le Botton Connexion), sinon il doit faire d'abord l'opération de l'inscription (remplir le champ de nom, email, mot de passe et de numéro de certificat, puis cliquer sur le Botton Inscription).

Pour l'opération de déconnexion, on clique sur le bouton Retour de la Figure (4.8).

Pour supprimer les champs, on clique sur le bouton Supprimer Figure (4.8).

traduction automatique

Connexion :

Numéro de certificat :

Mot de pass :

Supprimer Connexion

Inscription :

Name :

Email :

Mot de pass :

Numéro de certificat :

Supprimer Inscription

Retour

Windows تنشيط
انتقل إلى الإعدادات لتنشيط Windows

Figure (4.8):Fenêtre d'authentification.

4.4.3 Les fenêtres des opérations sur les bases de données

Quand on clique sur le bouton Connexion dans de la Figure (4.8), La figure (4.9) apparait.

traduction automatique

 **TABIB** 

AJOUTER 

MISE À JOUR 

SUPPRIMER 

RETOUR 

Windows تنشيط
انتقل إلى الإعدادات لتنشيط Windows

Figure (4.9): Fenêtre de choix de l'opération.

D'après la Figure (4.9), on choisit l'opération à faire.

4.4.3.1 La fenêtre de l'ajout

Quand on clique sur le bouton Ajouter de la Figure (4.8), La figure (4.10) apparait.

Après avoir remplie les champs nécessaires de la Figure (4.10), on clique sur le bouton Ajouter de la Figure (4.10), l'opération de l'ajout démarre.

Quand on clique sur le bouton Affiche de la Figure (4.10), La base de données s'affiche.

Pour supprimer les champs, on clique sur le bouton Supprimer Figure (4.10).

The screenshot shows a web application window titled "traduction automatique". At the top center is the "TABIB" logo. Below the logo, there is a dropdown menu labeled "Sélectionner la base de données:" with the selected option "Base de donnée des noms masculin". To the left of the main content area are four input fields with labels: "mot en francais:", "mot en arabe:", "Pronom:", and "Synonyme de mot en francai:". To the right of these fields is a large, empty rectangular box labeled "BD". At the bottom of the window, there are three buttons: "Afficher la BD", "Ajouter", and "Retour".

Figure (4.10):Fenêtre de l'ajout.

4.4.3.2 Fenêtre de suppression

Quand on clique sur le bouton Supprimer de la Figure (4.9), La figure (4.11) apparait.

Après avoir remplie les champs nécessaires de la Figure (4.11), on clique sur le bouton Supprimer de la Figure (4.11), l'opération de la suppression démarre.

Quand on clique sur le bouton Affiche de la Figure (4.), La base de données s'affiche.

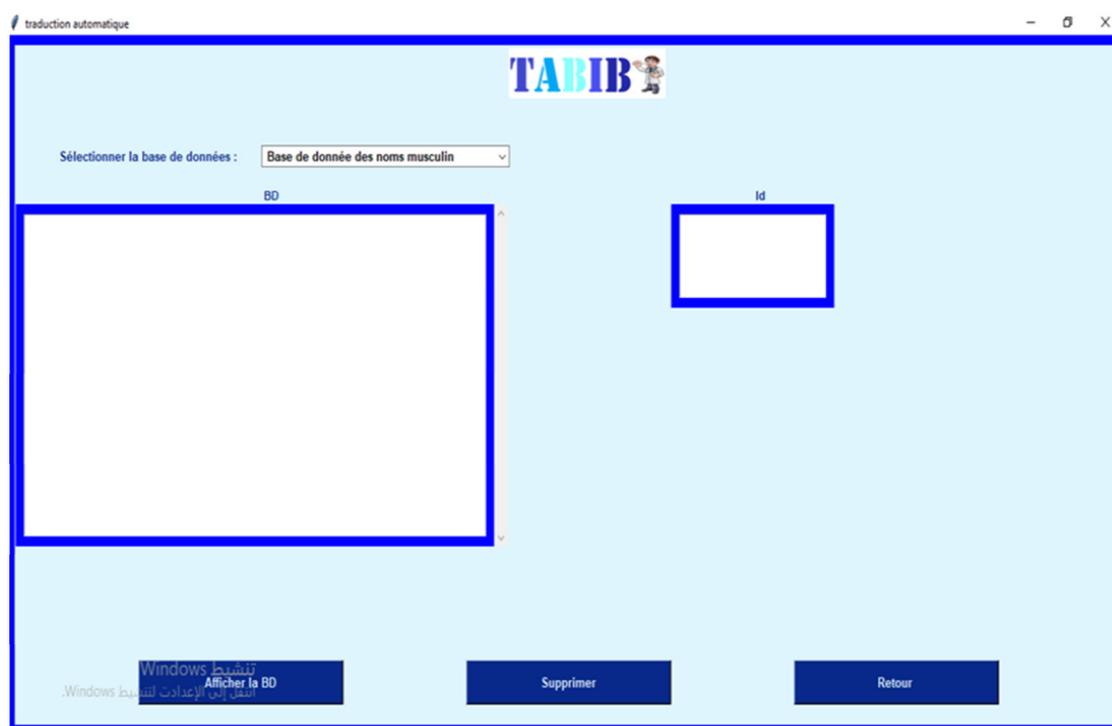


Figure (4.11) : Fenêtre de la suppression.

4.4.3.3 Fenêtre de la mise à jour

Quand on clique sur le bouton Mise à jour de la Figure (4.9), La figure (4.12) apparait.

Après avoir remplie les champs nécessaires de la Figure (4.12), on clique sur le bouton Mise à jour de la Figure (4.12), l'opération de la mise à jour démarre.

Quand on clique sur le bouton Affiche de la Figure (4.12), La base de données s'affiche.

Pour supprimer les champs, on clique sur le bouton Supprimer Figure (4.12).

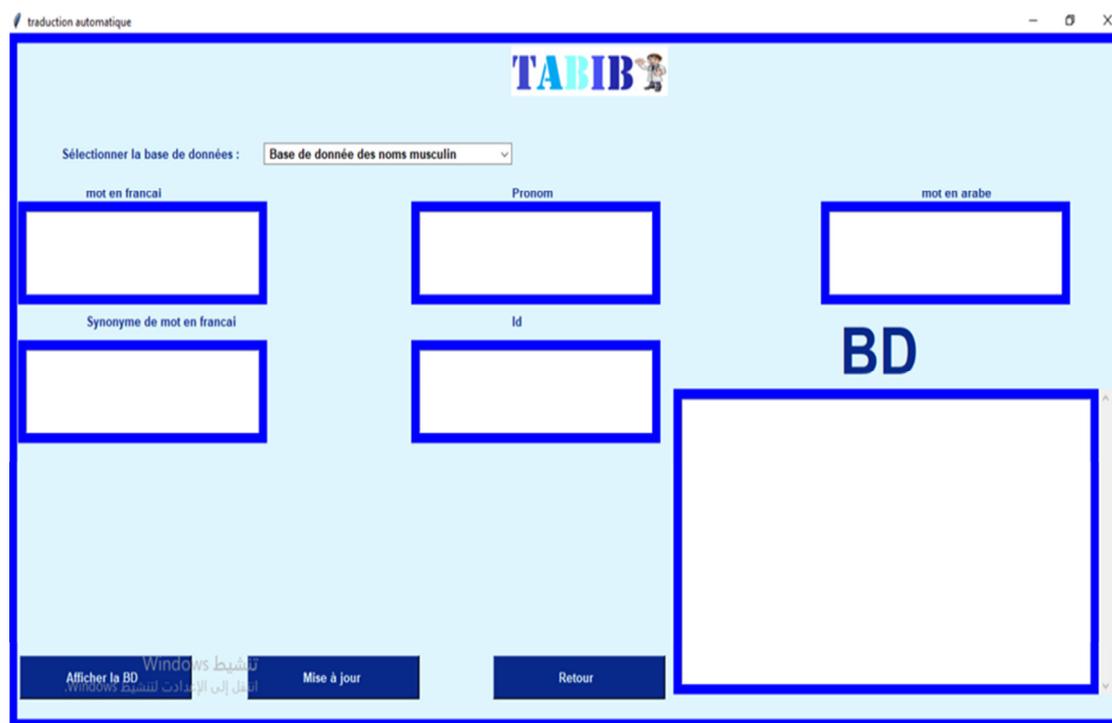


Figure (4.12):Fenêtre de la mise à jour.

4.4.4 Fenêtre de la traduction

Cette fenêtre peut apparaître par l'administrateur ou l'utilisateur de la façon suivante:

- **Par l'administrateur** : quand on clique sur le bouton Traduction de la Figure (4.7), La figure (4.13) apparaît.
- **Par l'utilisateur** : quand on clique sur le bouton Traduction de la Figure (4.8), La figure (4.13) apparaît.



Figure (4.13):Fenêtre de la Traduction d'un texte.

Dans la Figure (2.13) la traction de texte passe par des opérations, comme suit :

- a. **L'opération préparatoire :** On choisit le texte à traduire illustré dans les Figure (4.14).

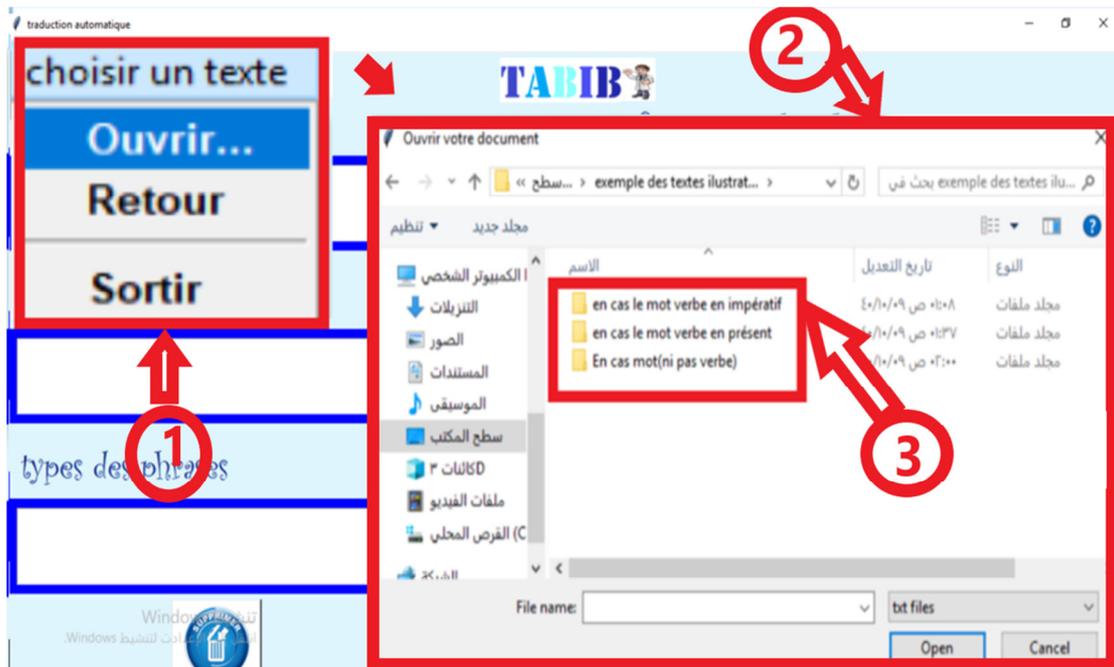


Figure (4.14):Fenêtre de la Traduction d'un texte dans L'opération préparatoire.

Dans La Figure (4.14), on clique sur Ouvrir du cadre (1), va apparaitre le cadre (2) qui contient le cadre (3), ce dernier comprend un nombre important des exemples de textes qui explique les règles de grammaire utilisées dans les opérations de conversion et de génération. Quelques exemples seront détaillés par la suite dans les opérations de conversion et de génération.

b. L'opération d'analyse : On choisit le texte à traduire (un message à un médecin spécialiste), illustré dans les Figure (4.15).



Figure (4.15):Fenêtre de la Traduction d'un texte dans L'opération d'analyse.

Quand on clique sur le bouton (1) de Figure (4.15), le texte va apparaitre sous forme des phrases illustrées dans la Figure (4.16).



Figure (4.16):Fenêtre de la Traduction d'un texte dans L'opération d'analyse (cas de transfert d'un texte vers des phrases).

Quand on clique sur le bouton (2) de Figure (4.15), les phrases vont apparaitre sous forme des mots illustrés dans la Figure (4.17).

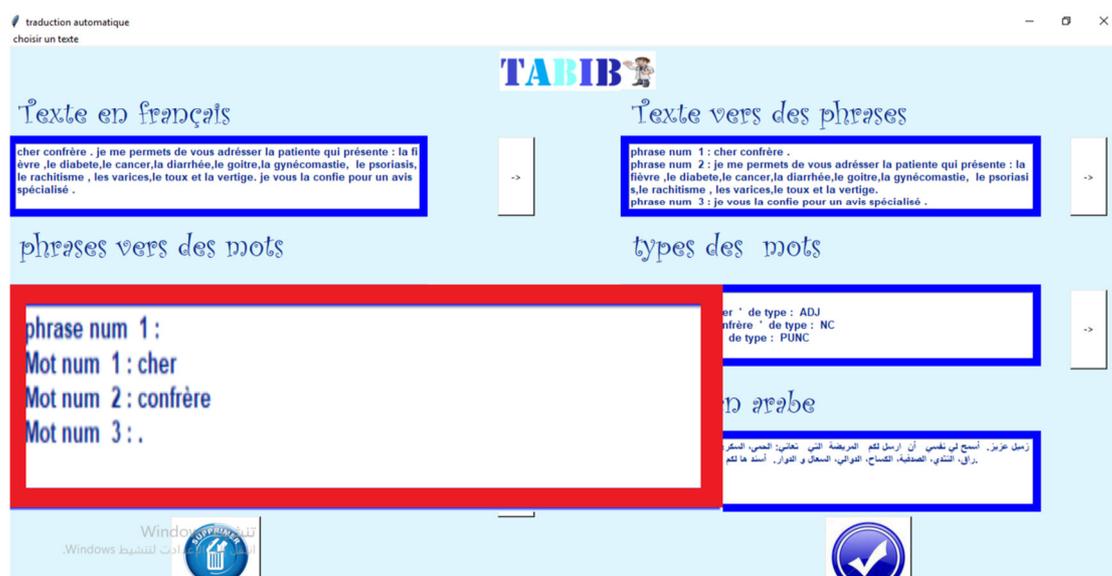


Figure (4.17):Fenêtre de la Traduction d'un texte dans L'opération d'analyse (cas de transfert des phrases vers des mots).



Figure (4.21):Fenêtre de la Traduction d'un texte dans L'opération de conversion (cas du synonyme (déduction)).

Dans la Figure (4.22), le changement de genre de nom de la phrase encadrée dans le champ (A) mentionné dans le champ (B), va induire un changement de la conjugaison du verbe et du pronom relatif de la phrase encadré de champ (A) mentionné dans les champs (C) et (D).

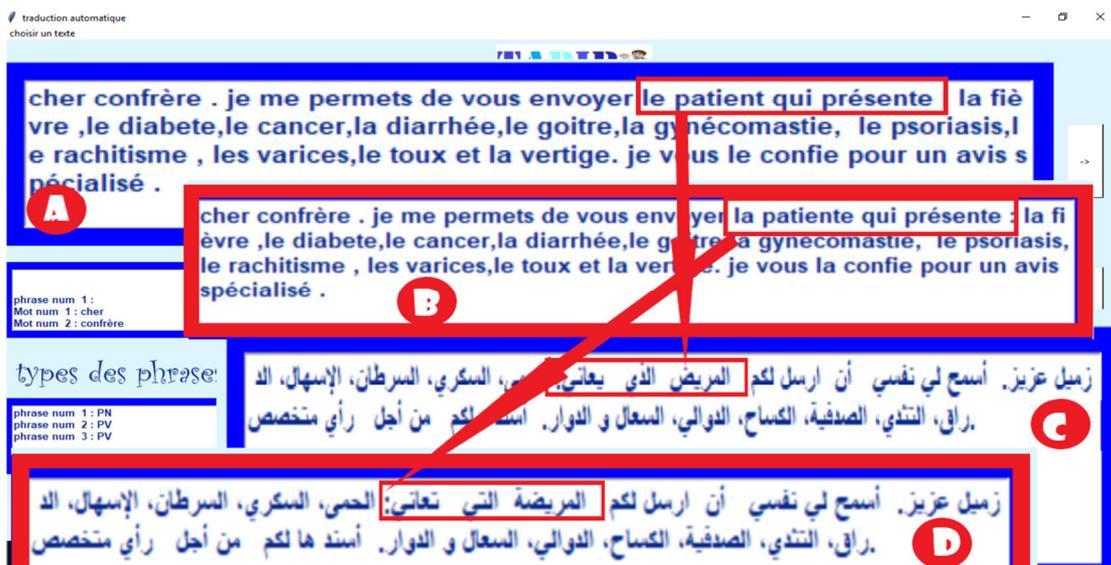


Figure (4.22):Fenêtre de la Traduction d'un texte dans L'opération de conversion (cas changement de genre d'un nom).

d. L'opération de génération :

Quand on clique sur le bouton (5) de Figure (4.15), l'opération de la génération démarre après l'opération de la conversion, illustrée dans la Figure (4.23).



Figure (4.23): Fenêtre de la Traduction d'un texte dans L'opération de génération.

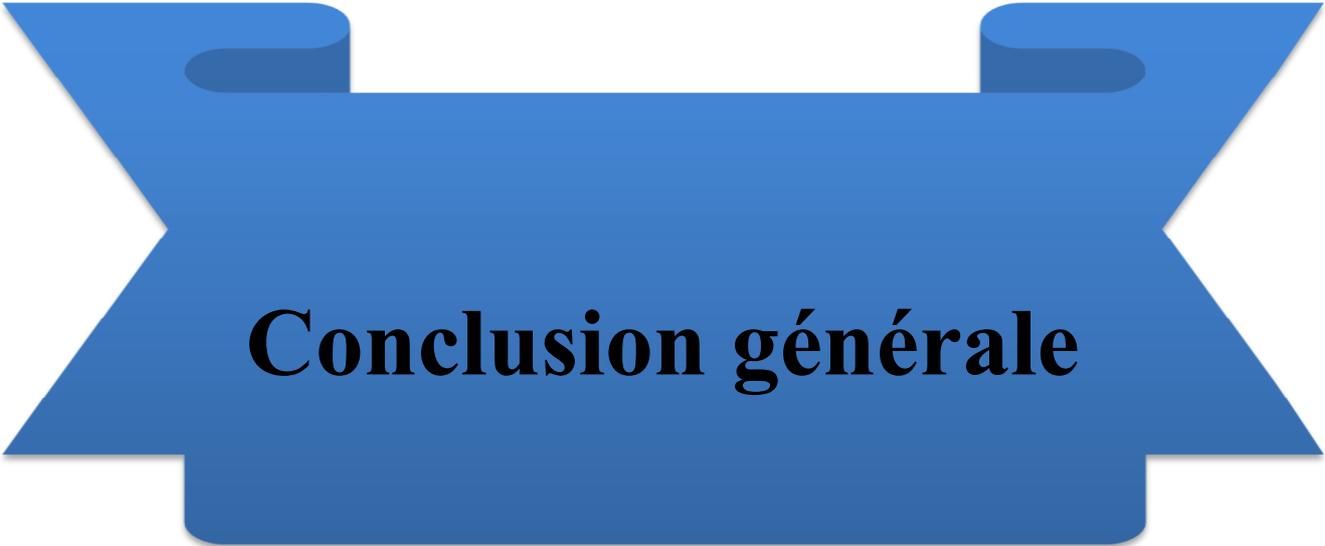
Dans la Figure (4.23), les cadres de champ (B) précisent les résultats de l'opération de génération déroulée dans les cadres de champ (A):

- les cadres (1) de deux champs illustrent le cas où l'adjectif précède le nom.
- le cadre (2) de champ (A) illustre le cas de l'élimination du pronom personnel.
- les cadres (3) et (5) de deux champs illustrent le cas où les pronoms relatifs précèdent le verbe.
- les cadres (4) de deux champs illustrent le cas où l'article défini précède le nom.
- les cadres (6) de deux champs illustrent le cas où l'article indéfini précède le nom.

4.5 Conclusion

Dans ce chapitre, nous avons présenté les langages et les outils de développement utilisés dans la validation et l'implémentation de notre système .nous avons utilisé plusieurs outils talque nlpnet ,tinkinter, xampp et python qui nous a aidé dans l'implémentation de notre système .

Enfin, on a présenté la description du système (les interfaces), et les résultats obtenus.



Conclusion générale

Conclusion générale

Dans notre travail, on a créé un traducteur capable de traduire un texte (bilan médical et messages vers un docteur spécialiste) en la langue française à un texte (bilan médical et messages vers un docteur spécialiste) en la langue arabe dans un contexte spécifique (médical). On a choisi la méthode de traduction automatique à base des règles en ajoutant un dictionnaire médical pour se focaliser sur le contexte médical et ajouter la sémantique.

On a d'abord essayé d'analyser grammaticalement (Sélection et extraction des éléments grammaticale) le texte en langue française, ensuite on a converti et générer le texte en langue arabe en utilisant un dictionnaire médical afin de traiter le sens.

On a essayé de respecter les règles de la grammaire de la langue arabe, à fin d'obtenir un texte bien formaté.

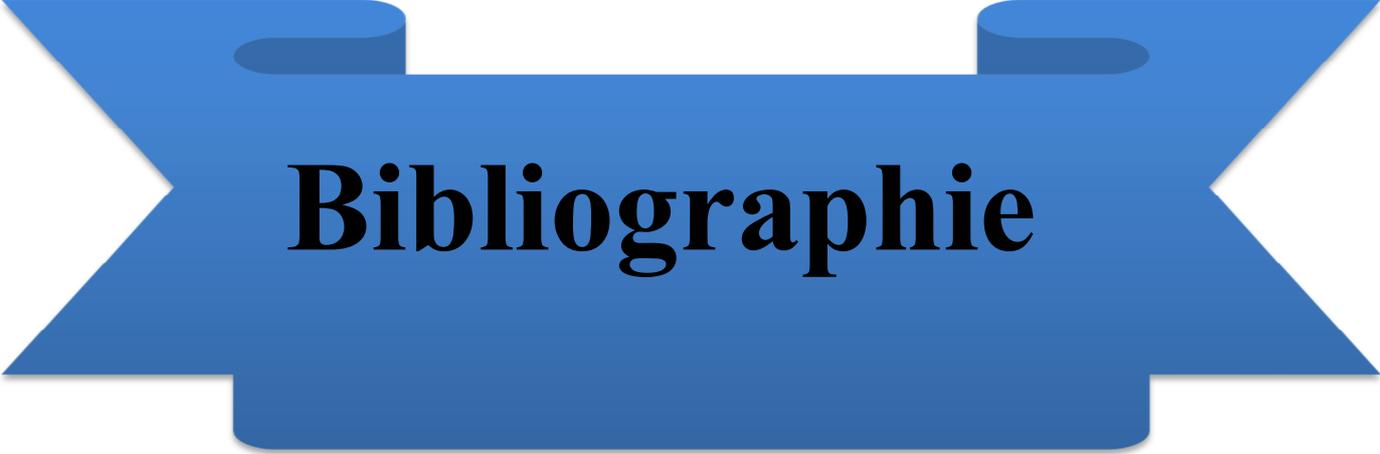
En premier lieu de ce projet, nous avons commencé par une présentation de l'état de l'art de traduction automatique et nous avons donné quelques travaux existants .ensuite nous avons expliqué brièvement la notion du contexte et de la grammaire .par la suite, nous avons entamé la conception de notre système .et enfin nous avons réalisé et implémenter notre système.

Dans notre système, nous pouvons dire que nous avons couvert des parties de la traduction automatique et non totalement, car il y a eu les lacunes suivantes:

- La grammaire est un domaine vaste.
- Certaine ambiguïté dans la recherche de synonyme de mot (cas de plusieurs synonymes).
- Problème d'outil d'analyse grammaticale de texte (nlp) (quelques résultats non fiables, prise de temps considérable).
- Absence d'analyse morphologique.
- le vocabulaire est un domaine vaste.

Nous avons des perceptives pour résoudre le problème de ces lacunes, qui sont :

- Fusionner quelques principes de l'approche statistique et l'approche à base des exemples avec l'approche utilisée.
- Utiliser de nouvelles techniques de traitement de texte tel que le deep learning



Bibliographie

Bibliographie

[1] **Sébastien Haire**. Traitement Automatique des Langues et Apprentissage des Langues .Assisté par Ordinateur : bilan, résultats et perspectives . 2011

[2] Qu'est-ce que le traitement automatique des langues ?
http://www.technolangue.net/imprimer.php3?id_article=274 .15/2/2019.

[3] **Mojca Pecman**. Introduction à la linguistique.

https://www.eila.univ-paris-diderot.fr/_media/enseignement/lea/cours/linguistique/intro-linguistique/introductionlinguistiqueecon1.pdf. 9/3/2019.

[4] **Georges Mounin**. Introduction à la liguistique.2004

[5] **Natalie Kübler**. La traduction automatique : traduction machine?.2007

[6] Comment l'IA révolutionne la traduction ?
<https://www.lesechos.fr/2018/05/comment-lia-revolutionne-la-traduction-991256>. 15/2/2019.

[7] **Sadik Bessou** . Traduction Automatique à Base de Règles de l'Anglais vers l'Arabe.2016

[8] **Segun Afolabi**. La Traduction Automatique par Opposition a La Traduction Humaine Analyse d'un Corpus de Productions Réelles.2014.

[9] Traduction automatique : la révolution de l'intelligence.
<https://leblob.fr/enquetes/traduction-automatique-la-revolution-de-lintelligence-artificielle>. 1/3/2019.

[10] Personnalisation . <http://www.systran.fr/systran/technologie/methodologie-de-personnalisation>.11/4/2019.

[11] Résumé des éléments de grammaire française.
<http://projetconnaissance.free.fr/grammaire-fr.html>.1/3/2019.

- [12] **Nedjmeddine Khalfallah.** Fiches grammaticales.2012
- [13] **Asmai Emeur.** Introduction à Intelligence artificielle.2011.
- [14] **Françoise Picot.** Apprendre la grammaire au CP et au CE1.2012.
- [15] **Khawla Hasni .**Extraction des documents les plus significatifs d'un document textuel .2018.
- [16] **Patrick Fuchs et Pierre Poulain .**Cours de Python.2019.
- [17] Tutorialspoint.
https://www.tutorialspoint.com/pycharm/pycharm_tutorial.pdf.6/4/2019.
- [18] Introduction XAMPP.
http://www.montefiore.ulg.ac.be/services/verif/cours/bd/repet2014/tp7_slides.pdf.
16/5/2019.
- [19] Python software fondation. <https://pypi.org/project/nlpnet>.16/5/2019.
- [20] Tkinter. <https://docs.python.org/2/library/tkinter.html>.16/5/2019.
- [21] **Yamina Kahla .**Découverte des services web sémantique .2014
- [22] Définition du contexte en informatique ubiquitaire.
[http://www-inf-it-sudparis.eu/cours/CSC5004/Archives/EnLigne/Cours/SensibiliteAuContexte/2.1.html](http://www-inf.it-sudparis.eu/cours/CSC5004/Archives/EnLigne/Cours/SensibiliteAuContexte/2.1.html)
. 10/3/2019.
- [23] **Tilgi .**Adaptation dynamique à l'environnement d'exécution.2006.
- [24] **Brown.** Context_aware application.1997.
- [25] **Gong.** Contextual modeling and application.2005.
- [26] **Winograd.** Architectures for context.2001.
- [27] **Khoulod Boukadi .**coopération interentreprises à la demande .2009

Dédicace

Je dédie ce travail qui n'aura jamais pu voir le jour sans les soutiens indéfectibles et sans limite de mes chers parents qui ne cessent de me donner avec amour le nécessaire pour que je puisse arriver à ce que je suis aujourd'hui. Que dieux les protège et que la réussite soit toujours à ma portée pour que je puisse leur combler de bonheur.

Je dédie aussi ce travail à :

Ma sœur, mes frères, ma famille, tous mes amis, mes collègues et tous ceux qui m'estiment.

Remerciement

Je remercie Allah premièrement.

Je dois remercier mon encadreur, [Dr. Bouchana Belkacem](#), qui m'a aidé à finir mon travail.

Je remercie tous ceux qui m'ont aidé, en particulier ma mère, mon père, ma famille.

Enfin et surtout, j'adresse mes salutations sincères aux membres du jury qui évalueront mon travail.

Résumé

Pour la crédibilité de la traduction, il faut strictement ajuster la grammaire des deux langues à traduire ainsi que le contexte.

Dans ce projet, nous créons un interprète qui va traduire le texte (une lettre à un médecin, une analyse médicale) du français vers l'arabe dans le domaine médical, nous choisissons un dictionnaire médical et une méthode de la traduction automatique à base des règles en mode transfert.

Le traducteur effectue quelques opérations sur le texte à traduire en analysant grammaticalement ce dernier, par l'extraction et la sélection des éléments grammaticaux, puis la conversion à l'aide d'un dictionnaire médical et en générant le texte cible à l'aide la grammaire de deux langues.

En premier lieu de ce projet, nous avons commencé par une présentation de l'état de l'art de traduction automatique et nous avons donné quelques travaux existants, ensuite nous avons expliqué brièvement la notion du contexte et de la grammaire. Par la suite, nous avons entamé la conception de notre système. et enfin nous avons réalisé et implémenter notre système.

ملخص

للمصداقية في الترجمة يجب ضبط قواعد اللغة لكلا اللغتين المراد الترجمة بينها و السياق المراد ضبطادقيقا .

في هذا المشروع سنقوم بصناعة مترجم يقوم بترجمة نص(رسالة الى طبيب مختص و تحاليل طبية) من لغة الفرنسية الى لغة العربية في مجال الطب , عن طريق اختيار قاموس طبي و طريقة الترجمة الالية المعتمدة على القواعد و بالتحديد التحويلية. اذ يقوم المترجم بإجراء بعض العمليات على نص لترجمته حيث يقوم بتحليل النص تحليلا نحويا باستخراج و تحديد العناصر النحوية للنص و من بعدها تحويله دلاليا باستخدام المعجم الطبي و توليده باستخدام قواعد اللغة لكلا اللغتين .

نبدأ هذا المشروع بمقدمة عن الترجمة الالية و بعض الأعمال موجودة , وشرحنا باختصار لكل من قواعد لغتين الفرنسية و العربية ومصطلح السياق ,وبعد ذلك صممناه. و أخيرا قمنا بتحقيقه و تنفيذه.

Table des matières

Introduction généraleErreur ! Signet non défini.

Chapitre 01 Traduction automatique

1.1 Introduction	Erreur ! Signet non défini.
1.2 Définitions	Erreur ! Signet non défini.
1.2.1 Le Traitements Automatique des Langues (TAL).....	Erreur ! Signet non défini.
1.2.2 La traduction automatique (TA)	Erreur ! Signet non défini.
1.2.3 La linguistique	Erreur ! Signet non défini.
1.2.4 Le langage.....	Erreur ! Signet non défini.
1.2.5 Une langue	Erreur ! Signet non défini.
1.2.6 La linguistique informatique	Erreur ! Signet non défini.
1.2.7 La grammaire.....	Erreur ! Signet non défini.
1.2.8 Un dictionnaire	Erreur ! Signet non défini.
1.2.9 Vocabulaire.....	Erreur ! Signet non défini.
1.2.10 Intelligence artificielle	Erreur ! Signet non défini.
1.3 Historique	Erreur ! Signet non défini.
1.3.1 Les années 50-60	Erreur ! Signet non défini.
1.3.2 Le rapport ALPAC et ses conséquences.....	Erreur ! Signet non défini.
1.3.3. Entre les années 1967-1976.....	Erreur ! Signet non défini.
1.3.4 Les années 80.....	Erreur ! Signet non défini.
1.3.5 Le développement des aides à la traduction	Erreur ! Signet non défini.

1.3.6 Des années 90 à aujourd'hui.....	Erreur ! Signet non défini.
1.4. Intelligence artificielle et la traduction automatique.....	Erreur ! Signet non défini.
1.5 Les approches de traduction automatique	Erreur ! Signet non défini.
1.5.1 L'approche à base de règles	Erreur ! Signet non défini.
1.5.2 L'approche statistique	Erreur ! Signet non défini.
1.5.3 L'approche à base d'exemples	Erreur ! Signet non défini.
1.5.4 L'approche hybride	Erreur ! Signet non défini.
1.6 Échantillons Des Logiciels De Traduction Automatique....	Erreur ! Signet non défini.
1.7 Quels usages et limites ?	Erreur ! Signet non défini.
1.8 La traduction automatique comme outil automatique.....	Erreur ! Signet non défini.
1.9 Personnalisation du système par la création d'un dictionnaire spécialisé	Erreur ! Signet non défini.
1.10 Avantages de la personnalisation	Erreur ! Signet non défini.
1.12 Conclusion.....	Erreur ! Signet non défini.

Chapitre 02 Grammaire et Contexte

2.1 Introduction	Erreur ! Signet non défini.
2.2 La grammaire	Erreur ! Signet non défini.
2.2.1 Définition.....	Erreur ! Signet non défini.
2.2.2 La grammaire de la langue française	Erreur ! Signet non défini.
2.2.2.1 Étapes de l'analyse d'un texte.....	Erreur ! Signet non défini.
2.2.2.2 Découpages de la phrase.....	Erreur ! Signet non défini.

2.2.2.3	Division en deux ensembles des mots français	Erreur ! Signet non défini.
2.2.2.4	Caractéristiques des éléments	Erreur ! Signet non défini.
2.2.2.5	Autres notions fondamentales	Erreur ! Signet non défini.
2.2.2	Grammaire en la langue arabe	Erreur ! Signet non défini.
2.3	Le contexte	Erreur ! Signet non défini.
2.3.1	Définition du contexte	Erreur ! Signet non défini.
2.3.2	Les catégories de contexte	Erreur ! Signet non défini.
2.4	Conclusion	Erreur ! Signet non défini.

Chapitre 03 Conception

3.1	Introduction	Erreur ! Signet non défini.
3.1	Conception globale	Erreur ! Signet non défini.
3.2.1	Architecture générale	Erreur ! Signet non défini.
3.2.1.1	Module des opérations sur la base de données	Erreur ! Signet non défini.
3.2.1.2	Module d'analyse du texte	Erreur ! Signet non défini.
3.2.1.3	Module de conversion	Erreur ! Signet non défini.
3.2.1.4	Module de génération du texte	Erreur ! Signet non défini.
3.3	Conception détaillée	Erreur ! Signet non défini.
3.3.1	L'étape préparatoire	Erreur ! Signet non défini.
3.3.2	L'étape d'analyse	Erreur ! Signet non défini.
3.3.2.1	Extraction des éléments grammaticaux (phrases, mots)	Erreur ! Signet non défini.
3.3.2.2	Sélection des types des éléments grammaticaux (phrases, mots)	Erreur ! Signet non défini.

3.3.3	L'étape de la conversion	Erreur ! Signet non défini.
3.3.3.1	L'étape de la Sélection	Erreur ! Signet non défini.
3.3.3.2	L'étape de la comparaison.....	Erreur ! Signet non défini.
3.3.3.3	L'étape du test	Erreur ! Signet non défini.
3.3.3.4	L'étape de l'extraction	Erreur ! Signet non défini.
3.3.3.5	L'étape de la recherche du synonyme	Erreur ! Signet non défini.
3.3.4	L'étape de la génération	Erreur ! Signet non défini.
3.3.4.1	L'étape de la reformulation	Erreur ! Signet non défini.
3.3.4.1.1	L'étape de l'ordonnancement	Erreur ! Signet non défini.
3.3.4.1.2	L'étape de l'assemblage	Erreur ! Signet non défini.
3.3.5	Module des opérations sur la base de données	Erreur ! Signet non défini.
3.3.5.1	Opération de la suppression	Erreur ! Signet non défini.
3.3.5.2	Opération de l'ajout.....	Erreur ! Signet non défini.
3.3.5.3	Opération de la mise à jour	Erreur ! Signet non défini.
3.3.5.4	Opération de l'authentification.....	Erreur ! Signet non défini.
3.4	Modélisation de système	Erreur ! Signet non défini.
3.4.1	Diagrammes de cas d'utilisation	Erreur ! Signet non défini.
3.4.1.1	Identification des acteurs	Erreur ! Signet non défini.
3.4.1.2	Identification des cas d'utilisation.....	Erreur ! Signet non défini.
3.5	Conclusion.....	Erreur ! Signet non défini.

Chapitre 04 Implémentation

Introduction générale	16
Introduction générale	1
Chapitre 01 La traduction automatique.....	1

1.1 Introduction.....	3
1.2 Définitions	3
1.2.1 Le Traitements Automatique des Langues (TAL).....	3
1.2.2 La traduction automatique (TA)	3
1.2.3 La linguistique.....	4
1.2.4 Le langage.....	4
1.2.5 Une langue.....	4
1.2.6 La linguistique informatique.....	4
1.2.7 La grammaire.....	4
1.2.8 Un dictionnaire	5
1.2.9 Vocabulaire.....	5
1.2.10 Intelligence artificielle	5
1.3 Historique	5
1.3.1 Les années 50-60.....	5
1.3.2 Le rapport ALPAC et ses conséquences.....	6
1.3.3. Entre les années 1967-1976	7
1.3.4 Les années 80	7
1.3.5 Le développement des aides à la traduction.....	7
1.3.6 Des années 90 à aujourd’hui	8
1.4. Intelligence artificielle et la traduction automatique.....	10
1.5 Les approches de traduction automatique	10
1.5.1 L'approche à base de règles	11
1.5.2 L'approche statistique	13
1.5.3 L'approche à base d'exemples	13
1.5.4 L'approche hybride	14
1.6 Échantillons Des Logiciels De Traduction Automatique	15
1.7 Quels usages et limites ?	17
1.8 La traduction automatique comme outil automatique.....	17
1.9 Personnalisation du système par la création d’un dictionnaire spécialisé	20
1.10 Avantages de la personnalisation.....	23
1.12 Conclusion	24
Chapitre 02 Grammaire et Contexte.....	85

Introduction générale	17
Introduction générale	1
Chapitre 01 La traduction automatique.....	1
1.1 Introduction.....	3
1.2 Définitions	3
1.2.1 Le Traitements Automatique des Langues (TAL).....	3
1.2.2 La traduction automatique (TA)	3
1.2.3 La linguistique.....	4
1.2.4 Le langage.....	4
1.2.5 Une langue.....	4
1.2.6 La linguistique informatique.....	4
1.2.7 La grammaire.....	4
1.2.8 Un dictionnaire.....	5
1.2.9 Vocabulaire.....	5
1.2.10 Intelligence artificielle	5
1.3 Historique	5
1.3.1 Les années 50-60.....	5
1.3.2 Le rapport ALPAC et ses conséquences.....	6
1.3.3. Entre les années 1967-1976	7
1.3.4 Les années 80	7
1.3.5 Le développement des aides à la traduction.....	7
1.3.6 Des années 90 à aujourd’hui	8
1.4. Intelligence artificielle et la traduction automatique.....	10
1.5 Les approches de traduction automatique	10
1.5.1 L'approche à base de règles	11
1.5.2 L'approche statistique	13
1.5.3 L'approche à base d'exemples	13
1.5.4 L'approche hybride.....	14
1.6 Échantillons Des Logiciels De Traduction Automatique	15
1.7 Quels usages et limites ?	17
1.8 La traduction automatique comme outil automatique.....	17
1.9 Personnalisation du système par la création d’un dictionnaire spécialisé	20
1.10 Avantages de la personnalisation.....	23

1.12 Conclusion	24
Chapitre 02 Grammaire et Contexte.....	85

Table des figures

Figure (1.1): Traduction automatique à base des règles.	Erreur ! Signet non défini.
Figure (1.2): Triangle de Vauquois.	Erreur ! Signet non défini.
Figure (1.3): Traduction automatique statistique.	Erreur ! Signet non défini.
Figure (1.4): Traduction automatique à base d'exemples.	Erreur ! Signet non défini.
Figure (1.1): Traduction automatique à base des règles.....	11

Figure (1.2):Triangle de Vauquois.	11
Figure (1.3):Traduction automatique statistique.	13
Figure (1.4):Traduction automatique à base d'exemples.	13

Liste des tableaux

Tableau (1.1): Echantillons Des Logiciels De Traduction Automatique.....	17
Tableau (2.1): les types de temps de verbe.	Erreur ! Signet non défini.
Tableau (2.2):Typologie des subordonnées.	30



Introduction générale

Introduction générale

La traduction est un transfert d'un texte d'une langue à une autre, il est important de tenir compte des caractéristiques de toutes les langues traduites entre elles .ces caractéristiques sont mises en pratique dans le contrôle des règles de la langue (grammaire) et la méthode qui consiste à comprendre le contexte dans lequel le texte est traduit.

La traduction automatique est devenue un moyen nécessaire et inévitable, vu la vitesse et la capacité d'échange sur internet

La traduction automatique manque une certaine crédibilité pour être totalement acceptée par les utilisateurs. Quoiqu'on ait dit des progrès de la traduction automatique, celle-ci se trouve à une limite ne dépassant pas les lignes de texte car elle ne permet pas de saisir les caractéristiques des deux langues.

L'objectif de notre projet est de créer un bon traducteur capable de traduire un texte (bilan médical et messages vers un docteur spécialiste) en la langue française à un texte (bilan médical et messages vers un docteur spécialiste) en la langue arabe dans un contexte spécifique (médical) de manière plus efficace.

Pour achever le but, nous choisirons la méthode de traduction automatique à base des règles en mode transfert et un dictionnaire médical pour réaliser la contextualisation.

Ce mémoire structuré de la manière suivante :

Chapitre 01 : Ce chapitre présente l'état de l'art sur traduction automatique. Nous allons présenter des définitions des concepts qui sont utilisés dans notre mémoire, puis nous allons aborder le sujet de traduction automatique selon les titres suivants : historique de la traduction automatique, comment l'IA révolutionne la traduction automatique?, les approches de traduction automatique, échantillons Des

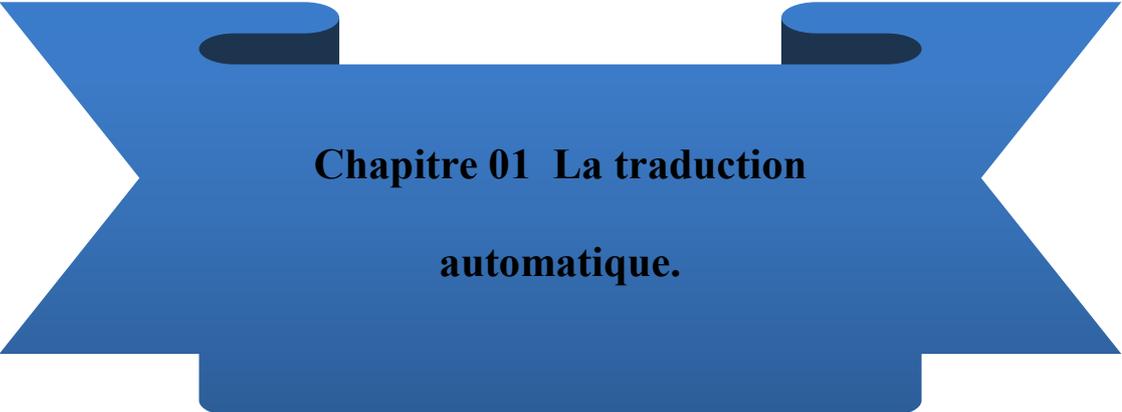
Logiciels, quels usages et limites ?, la traduction automatique comme outil automatique et enfin on parle sur la personnalisation et leur avantages.

Chapitre 02 : Dans ce chapitre, nous parlons du mot contexte dans l'informatique, puis nous présentons en générale tout ce qui concerne la grammaire de la langue arabe et française.

Chapitre 03 : présente l'architecture générale et détaillée de notre système, ainsi que la modélisation de ce dernier.

Chapitre 04 : Parle de la réalisation et l'implémentation de notre projet (nous exposons l'environnement de développement et les outils utilisés dans ce projet, ensuite nous allons présenter les interfaces graphiques en décrivant les différentes fonctionnalités de notre projet et nous présenterons des exemples qui nous permettront d'illustrer les résultats obtenus lors de l'utilisation de notre projet).

Notre mémoire se termine par une conclusion générale résumant notre travail ainsi qu'un ensemble des perspectives.



**Chapitre 01 La traduction
automatique.**

1.1 Introduction

Dans ce chapitre ,nous allons présenter des définitions des concepts qui sont utilisés dans notre mémoire, puis nous allons aborder le sujet de traduction automatique selon les titres suivants : historique de la traduction automatique, comment l'IA révolutionne la traduction automatique?, les approches de traduction automatique, échantillons Des Logiciels, quels usages et limites ?, la traduction automatique comme outil automatique et enfin on parle sur la personnalisation et leur avantages.

1.2 Définitions

1.2.1 Le Traitements Automatique des Langues (TAL)

Le Traitement Automatique des Langues (TAL) est un domaine de l'Intelligence Artificielle (IA) qui traite divers aspects de la langue et du discours, à tous les niveaux, lexical, morphologique, syntaxique, sémantique ou pragmatique [1].

Le TAL a pour objectif de développer des logiciels ou des programmes informatiques capables de traiter de façon automatique des données linguistiques [2].

Pour traiter automatiquement ces données, il faut d'abord expliciter les règles de la langue puis les représenter dans des formalismes opératoires et calculables et enfin les implémenter à l'aide de programmes informatiques [2].

Les principaux domaines du TAL sont :

- le traitement de la parole ;
- la traduction automatique ;
- la compréhension automatique des textes ;
- la génération automatique de textes ;
- la gestion électronique de l'information et des documents existants [2].

1.2.2 La traduction automatique (TA)

TA est le domaine de recherche à la base du traitement automatique du langage.

Et aussic'est la traduction d'un texte réalisée complètement par un ou plusieurs programmes informatiques, sans l'intervention humaine.

1.2.3 La linguistique [3]

Science du langage, c'est-à-dire étude objective, descriptive et explicative de la structure, du fonctionnement (linguistique synchronique) et de l'évolution dans le temps (linguistique diachronique) des langues naturelles humaines. S'oppose ainsi à la grammaire (descriptive et normative) et la philosophie du langage (hypothèses métaphysique, biologiques, psychologiques, esthétiques sur l'origine, le fonctionnement, la signification anthropologiques possibles du langage).

Et aussi bien Étude historique et comparative des langues.

1.2.4 Le langage [4]

Fonction d'expression de la pensée et de communication entre les hommes, mise en œuvre au moyen d'un système de signes vocaux (parole) et éventuellement de signes graphiques (écriture) qui constitue une langue manifestations du langage.

1.2.5 Une langue

Système d'expression et de communication commun à un groupe social (communauté linguistique) [4].

1.2.6 La linguistique informatique

La linguistique informatique est un domaine multidisciplinaire dépend d'une modélisation symbolique (à base de règles) ou statistique du langage naturel établi dans l'informatique.

1.2.7 La grammaire

La grammaire est l'étude méthodologique des Éléments composites de la langue et aussi c'est un manuel ou une collection de documents expliquant des règles grammaticales.

1.2.8 Un dictionnaire

Un dictionnaire est un conteneur d'une collection des mots d'une langue ou d'un domaine d'activité généralement présentés par ordre alphabétique.

1.2.9 Vocabulaire

Collection des mots spécifiques (de domaine, de langue ou de dictionnaire) qui donne la liste des mots.

1.2.10 Intelligence artificielle

Le but de l'**Intelligence artificielle (IA)** est de concevoir des systèmes capables de reproduire le comportement de l'humain dans des activités de raisonnement [13].

1.3 Historique [5]

L'idée moderne en remonte aux années 1930 et on divise généralement son développement en quatre grandes périodes : les années 50-60, le rapport ALPAC (1966), et ses conséquences, la reprise des années 70-80 et la période actuelle qui commence au début des années 90.

1.3.1 Les années 50-60

Durant la deuxième guerre mondiale les premiers ordinateurs permettent aux belligérants de déchiffrer leurs codes respectifs. On tente ensuite d'appliquer ces techniques de déchiffrement à la traduction automatique. L'expert mondial de la traduction automatique, John Hutchins (2001), évoque la première conférence sur la TA qui fut organisée en 1952 par Bar-Hillel. Dès cette époque, les chercheurs sont conscients des difficultés que présente le domaine : ils se sont en effet déjà accordés sur la conclusion qu'une traduction de bonne qualité et totalement automatique relève de la science-fiction, ou du moins, restait du domaine du virtuel. Ils se rendent compte que l'intervention humaine, déjà qualifiée de pré-édition ou de post-édition restera indispensable. De plus, les premiers systèmes de TA partent avec de lourds handicaps : les connaissances en syntaxe et en analyse syntaxique sont encore très insuffisantes, les ordinateurs ont des capacités de stockage très limitées et sont peu puissants. On dispose avant tout de systèmes de mise en correspondance de dictionnaires, ce qui génère des résultats de traduction voués à l'échec ou alors des « systèmes-jouet »,

incapables de traduire plus de quelques phrases. Le premier système de TA est présenté au public en 1954 ; il permet de traduire 49 phrases russes, soigneusement sélectionnées, vers l'anglais se servant d'un dictionnaire de 250 mots et de six règles de grammaire. Ce sont probablement des motivations stratégiques qui poussent alors les gouvernements américain et soviétique à financer ce type de recherches. Comme nous l'avons mentionné plus haut, la recherche en linguistique théorique n'a pas encore atteint toutes les connaissances en syntaxe que l'on a aujourd'hui. Ce n'est qu'avec les découvertes de Noam Chomsky dans *Syntactic Structure* (1957) et *Aspects of the Theory of Syntax* (1962) que l'on commence à avoir des modèles syntaxiques de la langue. C'est aussi dans ces années-là que Lucien Tesnière développe une approche alternative dans ses *Eléments de syntaxe structurale* (1959). Une description syntaxique de la langue donne une structure aux phrases qui apparaissent sinon comme une simple suite de mots sans ordonnancement. La syntaxe est par conséquent indispensable au traitement automatique du langage autre que statistique. On peut donc dire que la recherche en traduction automatique des années 50 et au début des années 60 vient stimuler et appuyer toute la recherche dans le domaine du traitement automatique du langage. En effet, les systèmes de traduction automatique de différents types nécessitent tous, entre autres, un module performant et robuste d'analyse syntaxique.

1.3.2 Le rapport ALPAC et ses conséquences

Mandatée en 1966 par le gouvernement des Etats-Unis, la commission ALPAC (Automatic Language Processing Advisory Committee) publie un rapport qui va avoir une influence déterminante et très négative sur le soutien à la recherche en TA. Ce rapport conclut que la TA est plus lente, moins efficace et deux fois plus chère que la traduction faite par des humains. Il recommande d'arrêter de financer la recherche dans ce domaine et de se tourner plutôt vers des aides à la traduction, comme par exemple les dictionnaires automatiques. Le rapport recommande cependant de continuer à soutenir la recherche fondamentale en traitement automatique du langage. On peut dire que la recherche en TA a été alors virtuellement stoppée pour une décennie aux Etats-Unis

1.3.3. Entre les années 1967-1976

Au cours de cette période, la TA est en sommeil aux Etats-Unis, mais d'autres pays continuent à développer la recherche que stimulent des besoins concrets. Si aux Etats-Unis la seule activité porte sur la traduction du russe en anglais de textes scientifiques et techniques, au Canada, en Europe et au Japon en revanche, le développement de la TA se poursuit. Citons le projet canadien TAUM-Météo (décrit dans Isabelle 1986 par exemple) ; il débute en 1970 et produit deux résultats importants : il mène au développement ultérieur du langage de programmation PROLOG et s'utilise par ailleurs avec succès dans le domaine restreint de la traduction des prévisions météorologiques. En France, on développe à Grenoble un système permettant de traduire du russe des textes mathématiques et physiques (Boitet & Nedobejkine 1980). Dans la communauté européenne, les besoins en traduction deviennent de plus en plus importants et le système Systran est mis en place à la Commission Européenne en 1976. Il est encore utilisé aujourd'hui.

1.3.4 Les années 80

A la fin des années 70 et dans les années 80, on commence à travailler sur de gros systèmes commercialisables, et l'activité de recherche dans le domaine de la TA connaît un nouvel essor. A partir de la première version de Systran produite en 1970 par Peter Toma et qui consiste à traduire du russe vers l'anglais, on assiste au développement et à l'installation de nouvelles versions auprès de la Commission Européenne, et de l'OTAN, chez Xerox notamment. Les systèmes Logos et METAL se développent et sont les principaux concurrents commerciaux de Systran. Du côté de la recherche, on peut citer le premier système Ariane, développé au GETA à Grenoble chez Christian Boitet (Boitet & Nedobejkine 1980), le projet européen EUROTRA (Maegaard 1989 basé sur un système à transfert multilingue, le Distributed Language Translation system à Utrecht (Maxwell 1990), construit sur un système à pivot en passant par l'espéranto, et d'autres encore qui utilisent l'intelligence artificielle

1.3.5 Le développement des aides à la traduction

Dès les années 80, on commence à travailler sur des environnements d'aide à la traduction (Volanschi, ce volume) intégrant des dictionnaires, des concordanciers et des mémoires de traduction. Certaines entreprises qui construisent des systèmes de

TA les abandonnent (ALPSystems aux Etats-Unis) et/ou développent des mémoires de traduction (SITE à Paris). Dans certains systèmes, on commence à travailler sur le couplage de ces outils avec la traduction automatique. Les mémoires de traduction connaissent alors un développement intéressant et prennent une place prépondérante parmi les outils d'aide à la traduction

1.3.6 Des années 90 à aujourd'hui

A partir des années 90, ce sont les mémoires de traduction qui commencent à dominer le marché. Aujourd'hui ces outils font partie de la panoplie indispensable du traducteur professionnel. Avec la montée en puissance des ordinateurs personnels, les systèmes commerciaux de TA commencent à produire des systèmes individuels que l'on peut installer sur PC. Le développement d'Internet permet de mettre en ligne des systèmes de TA. Systran y a recours le premier en 1998 avec Altavista/Babelfish1. Malgré l'intérêt croissant porté aux outils d'aide à la traduction, la recherche en TA ne cesse pourtant pas et certains de ses résultats sont prometteurs. Les systèmes les plus efficaces se basaient jusqu'alors sur des systèmes à transfert, à savoir, un module d'analyse de la langue source, un module de règles de transfert complexes entre la langue source et la langue cible et un module de génération de la langue cible. Ces systèmes fonctionnent avec des règles syntaxiques et de très gros dictionnaires. La quantité croissante de documents disponibles sur support électronique permet d'enrichir toujours davantage les dictionnaires et surtout de créer des dictionnaires spécialisés qui donnent des résultats de traduction appréciables dans les domaines techniques et scientifiques. La recherche est ainsi amenée à tenter de détecter le domaine d'un texte source afin d'indiquer automatiquement au système quels dictionnaires il doit activer (Lang E. & Jin Yang 1999) On constate aussi que les systèmes mettent de plus en plus souvent l'accent sur le lexique et le lexique-grammaire ; on ajoute donc des informations syntaxiques toujours plus nombreuses aux éléments du lexique. C'est l'une des conséquences de l'évolution de la recherche en linguistique théorique et formelle, mais aussi de l'apport de la recherche en linguistique de corpus qui décrit la langue à partir des données concrètes du corpus (Scott & Tribble 2006 par exemple). Par ailleurs, les outils de traitement statistique de corpus développés en TAL ont permis l'émergence d'une nouvelle direction de recherche : la traduction automatique « basée sur les exemples ». Cette approche

permet de dégager les traductions les plus fréquentes en effectuant des analyses statistiques sur des corpus traduits et alignés. Ces systèmes ne sont pas commercialisés pour l'instant, mais ils présentent des résultats prometteurs. Même s'ils sont encore bien loin d'imiter le traducteur humain, les systèmes de traduction automatique de textes écrits sont aujourd'hui reconnus et utilisés sur le marché de la traduction. La traduction de l'oral représente un défi passionnant pour le domaine du traitement automatique du langage ; on parle ici de traduction automatique du langage parlé ou encore d'interprétariat automatique. Le système ATR (Sumita 1999) par exemple, développé au Japon permet d'effectuer des inscriptions téléphoniques à des colloques internationaux ou des réservations d'hôtel par téléphone. Citons encore le projet allemand Verbmobil (1993-2000) dont l'objectif était de créer un traducteur oral transportable pour aider les locuteurs germanophones et japonais à mener des négociations commerciales en anglais. Enfin le projet TC-STAR2 dont le but est d'effectuer la reconnaissance vocale des discours au Parlement Européen, traduit le discours et le restitue oralement dans la langue cible. Dans ce type d'outils, on combine donc la reconnaissance de la parole, la traduction automatique et la synthèse de la parole. Un autre défi consiste à combiner deux approches en traduction automatique. Si les systèmes à transfert ont en effet montré leur efficacité, ils ont aussi révélé leurs limites. Nous l'illustrerons plus bas. Par ailleurs, les systèmes statistiques donnent des résultats prometteurs mais qui révèlent eux aussi un certain nombre de limites. La combinaison des systèmes basés sur exemples avec les systèmes à transfert représente peut-être une possibilité d'améliorer de Notons par ailleurs, que la première interface graphique d'Internet, le World Wide Web n'a été mise en place par Tim Burners-Lee qu'en 1993

manière significative les résultats de la traduction automatique. La recherche en traduction automatique s'y attelle également. Pour conclure, on peut donc relever que le domaine de la TA a représenté et représente un élément moteur en recherche sur le traitement automatique du langage et qu'il conduit au développement de nombreux résultats connexes. Tous les outils d'aide à la traduction que l'on utilise aujourd'hui sont issus des recherches et du développement en TA. Le développement des analyseurs syntaxiques ou de logiciels de génération automatique est lui aussi étroitement lié aux besoins de la recherche en traduction automatique.

1.4. Intelligence artificielle et la traduction automatique [6]

L'histoire de l'intelligence artificielle est intimement liée à celle de la traduction, peut-être parce que savoir donner du sens est ce qu'il y a de plus humain et de plus mystérieux pour la machine.

« La traduction automatique marche tant bien que mal, ça n'est pas parfait, mais c'est suffisamment bon pour être utile ».

D'autant que le domaine a connu des progrès spectaculaires au cours des dernières années, grâce à l'utilisation de systèmes d'apprentissage automatique basés sur des réseaux de neurones - la technologie de l'intelligence artificielle la plus en pointe actuellement. Google, incontournable dans le domaine de la traduction gratuite avec son service Translate, fut l'un des premiers à l'employer pour cette tâche.

La technologie des réseaux de neurones a été introduite chez Google Translate via son laboratoire Google Brain et grâce aux travaux de Jeff Dean, Andrew Ng, Greg Corrado, Geoffrey Hinton et Quoc Le, figures emblématiques de l'intelligence artificielle.

1.5 Les approches de traduction automatique

La traduction automatique s'est développée autour de trois approches essentielles: les approches symboliques, les approches à base de règles, et les approches statistiques [7].

Nous détaillerons dans les approches depuis l'avènement de la traduction automatique (1950) à ce jour.

Les premiers programmes d'ordinateurs relatifs à la traduction étaient destinés à servir d'aide à la traduction. Quelques règles et surtout un dictionnaire bilingue composaient le cœur du système. Les années suivantes voient les dictionnaires grandir ; ce qui engendre une augmentation du nombre de règles régissant le réordonnement des mots. La nécessité d'automatiser l'acquisition des règles et de progresser leur généricité participe au développement de la linguistique informatique [7].

1.5.1 L'approche à base de règles



Figure (1.1): Traduction automatique à base des règles.

Le triangle présenté à la Figure (1.2) est attribué à Vauquois. Il présente de manière synthétique une analyse du processus de traduction encore pleinement pertinente et employée de nos jours [7].

La traduction se fait en plusieurs étapes:

Au niveau le plus bas, on retrouve la traduction directe, qui passe directement des mots de la langue source aux mots de la langue cible [7].

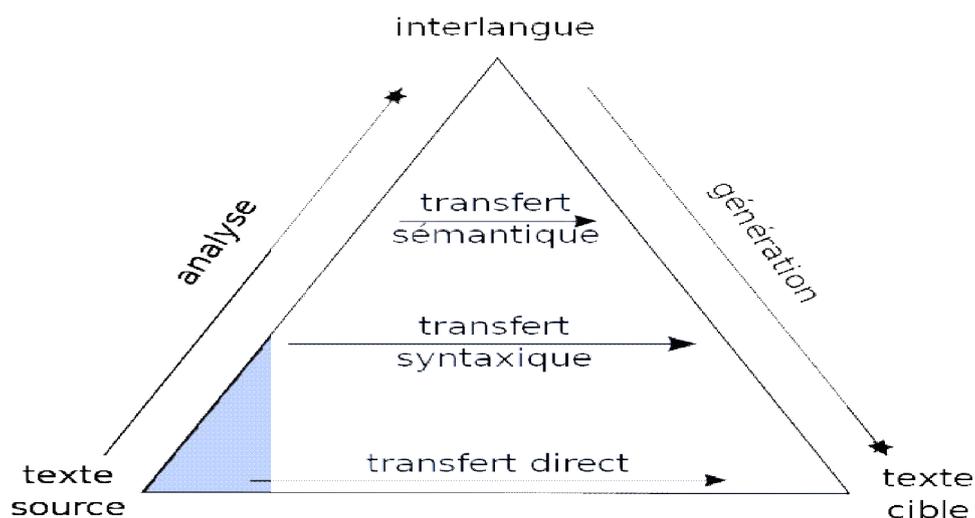


Figure (1.2): Triangle de Vauquois.

Les systèmes semi-directs ont une phase de segmentation ou d'analyse morphologique, voire morphosyntaxique, et une phase de génération morphologique. Si l'on effectue une analyse syntaxique de la phrase source, le transfert à la langue cible devrait être simplifié. À ce niveau, les détails spécifiques à la constitution des groupes nominaux, par exemple, n'ont pas besoin d'être connus des règles régissant le transfert [7].

Avec une analyse plus approfondie de la phrase source, au niveau sémantique, le transfert devient uniquement sémantique. En revanche, la génération des mots après le transfert est plus complexe qu'au niveau inférieur [7].

Enfin, une analyse totale de la phrase source peut aboutir à une représentation de son sens dans une « inter-langue » artificielle, de laquelle on dérive ensuite les mots cibles [7].

Le pivot est un type de représentation utilisant des attributs et relations inter-langue, et des unités lexicales de chacune des langues. Ces systèmes sont à transfert simple, alors qu'on a un double transfert en < pivot > [7].

L'approche reposant sur une inter-langue est attractive car elle remplace le problème de la traduction par deux problèmes unilingues, d'analyse et de synthèse. L'avantage est que les modules unilingues sont a priori réutilisables. Pour couvrir tous les sens de traduction entre n langues, il suffit de n modules d'analyse et de n modules de synthèse, contre $n*(n-1)$ systèmes de transfert [7].

Des systèmes à véritable langue-pivot, on peut citer : ATLASII de Fujitsu ou IVOT/Crossroads de NEC, ou KANT / CATALYST de CMU/Caterpillar, ou UNL, ou MASTOR-1 d'IBM [7]. .

Le début des années 1990 voit le développement d'autres types d'approches. Les ordinateurs se répandent et gagnent en puissance, ce qui permet l'émergence de stratégies qui se fondent sur de grandes quantités de données « corpus based approaches ». On distingue en particulier deux grands types d'approches: la traduction automatique à base d'exemples et la traduction automatique par méthodes statistiques [7].

1.5.2 L'approche statistique



Figure (1.3):Traduction automatique statistique.

Elle doit ses origines aux travaux de Brown et al .en 1990 et en particulier au prototype Candide de Berger et al. en 1994, un système de traduction construit à partir de discours disponibles en français et en anglais de parlementaires canadiens. En effet, comme la traduction à base d'exemples, la traduction par méthodes statistiques repose sur un corpus parallèle [7].

Un modèle statistique de traduction est défini, comprenant une ou plusieurs lois de probabilités. Le corpus est traité afin d'estimer ces lois qui sont souvent constituées de plusieurs milliers, voire millions de paramètres [7].

1.5.3L'approche à base d'exemples

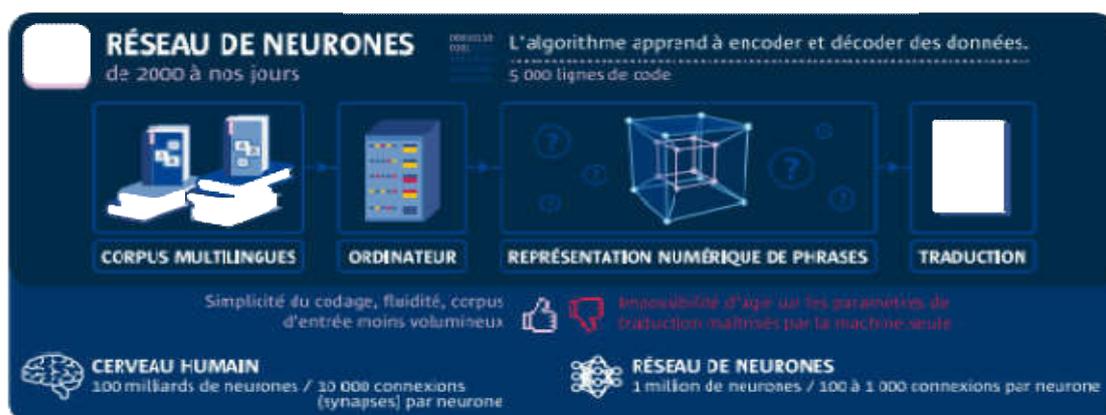


Figure (1.4):Traduction automatique à base d'exemples.

L'approche à base d'exemples (« Example-based machine translation », ou EBMT) repose sur un ensemble « d'exemples » préalablement traduits : un corpus parallèle de phrases traduites l'une de l'autre. Lorsqu'on lui présente une phrase à traduire, le système parcourt sa base d'exemples et produit trivialement une traduction si la phrase s'y trouve. Dans le cas général, la phrase n'apparaît pas dans la base et le système s'emploie alors à rassembler des exemples qui contiennent des fragments communs (des groupes de mots) avec la phrase à traduire. Pour chaque fragment d'exemple dans la langue source, il s'agit ensuite de retrouver sa traduction dans la langue cible : c'est la phase d'alignement. Enfin, la phase de génération assemble les fragments dans la langue cible et produit la traduction. À chacune des trois étapes, il est possible d'utiliser des sources externes de connaissances, telles que des lexiques bilingues, des listes de synonymes, des étiquettes ou des arbres syntaxiques...etc. [7]

1.5.4 L'approche hybride [7]

Leur idée directrice est qu'une approche unique du problème de la traduction, aussi perfectionnée soit-elle, ne parviendra pas à produire une traduction satisfaisante dans tous les cas. Au contraire, une approche par règles peut s'avérer particulièrement adaptée à certaines phrases, tandis que d'autres phénomènes linguistiques sont correctement traités par une approche reposant sur des corpus

Les systèmes hybrides sont actuellement envisagés comme des systèmes combinant les méthodes statistiques ou méthodes à base d'exemples avec des méthodes linguistiques (à base de règles), en particulier pour l'analyse morphologique et syntaxique.

Un système hybride pourrait parvenir à tirer profit des forces de chaque approche. Une première stratégie pour mettre en œuvre un système hybride est d'utiliser les différentes approches en parallèle. Enfin, dans un système statistique, il est courant de traiter par un système de règles spécialisées certains fragments de phrases, typiquement les nombres, les dates, etc. Les morceaux de phrases ainsi identifiés et traduits en isolation par le système à base de règles peuvent être transmis au système statistique.

1.6 Échantillons Des Logiciels De Traduction Automatique [8]

Num	Nom de logiciel	Lien	Remarques
1	Altavasta (de Babelfish)	http://www.babelfish.altavista.com	Ce logiciel permet de traduire en ligne à partir de plusieurs langues (anglais, français, allemand, italien, néerlandais, chinois) en plusieurs autres. (Langue de communication : l'anglais)
2	Bing	http://www.bing.com.translator	Le logiciel de traduction Bing permet également de traduire automatiquement d'une centaine de langues et vice versa. (Langue de communication : l'anglais)
3	Google Traduction	https://googletraduction.fr	Ce site de Google français permet de traduire dans un nombre considérable de combinaisons linguistiques entre l'anglais, le français, l'allemand, l'espagnol, l'italien, le portugais, le coréen, le japonais, le russe et le chinois. (langue de communication : le français)
4	Linguee	http://www.linguee.fr	Linguee est un outil de traduction unique en son genre. Il combine un dictionnaire rédactionnel et un moteur de recherche qui vous permettent de chercher la traduction d'un mot ou d'une expression parmi plusieurs centaines de millions de textes bilingues. Linguee possède mille fois plus de matériel traduit qu'un dictionnaire en ligne traditionnel, et ce sous la forme de phrases complètes et de leur traduction. (langue de communication : le français)
5	Prompt	http://www.translation2.paralink.com	Ce site de Prompt permet de traduire à partir de l'anglais, du français, du russe, de l'allemand, de l'italien et de

			l'espagnol, généralement vers les mêmes langues. (langue de communication : l'anglais)
6	Reverso	http://www.reverso.net	Le logiciel Reverso est un portail d'outils linguistiques y compris la traduction en une variété de combinaisons linguistique, notamment en français, espagnol, allemand, russe, portugais, chinois, arabe, hébreu et japonais (langue de communication : l'anglais)
7	Systran	http://www.systransoft.com	Ce site de Systran permet de traduire en ligne à partir de nombreuses langues, avec une possibilité impressionnante de combinaisons. On est limité à 150 mots à la fois. (langue de communication: l'anglais)
8	Voila	http://www.tr.voila.fr	Ce site de Voila permet de traduire en ligne à partir de plusieurs langues (anglais, français, allemand, portugais, espagnol, russe) en plusieurs autres. Le nombre de combinaisons est relativement limité. (langue de communication : le français)
9	Wanadoo	http://www.outils.wanadoo.fr/traduction	Ce site de Wanadoo permet de traduire dans un nombre considérable de combinaisons linguistiques entre l'anglais, le français, l'allemand, l'espagnol, l'italien, le néerlandais, le portugais et le russe. (langue de communication : le français)
10	WorldLingo	http://www.worldlingo.com/fr	Cet excellent site de WorldLingo permet de traduire à partir des langues sources suivantes : français, chinois, néerlandais, allemand, italien, grec, japonais, portugais, coréen,

			russe et espagnol. Les langues cibles : les mêmes. (langue de communication : le français)
--	--	--	--

Tableau (1.1): Echantillons Des Logiciels De Traduction Automatique.

1.7 Quels usages et limites ? [9]

" La traduction automatique excelle dans certaines tâches, mais échoue dans d'autres. Tout dépend du besoin ".

« Ce qui fait la force de la traduction automatique, c'est moins sa qualité que sa valeur d'usage », constate François Yvon, chercheur au LIMSI/CNRS. Le service offert aux internautes par Facebook ou TripAdvisor suffit à traduire des contenus simples – informations touristiques ou produits grand public, par exemple. Ses usagers en connaissent implicitement les limites et s'amuse d'erreurs parfois grossières. Voire de contresens ; *smoke eater scented candle*, par exemple, devient : « bougie parfumée à la fumée » (au lieu de bougie parfumée anti-tabac). La traduction automatique peut toutefois s'avérer très efficace pour des usages professionnels comme la mise à jour annuelle de manuels techniques ou de documents internes d'entreprise, ou appliqué à un domaine précis comme la météo. À l'instar de Systran, des opérateurs spécialisés proposent d'ailleurs d'adapter les algorithmes aux besoins sectoriels de leurs clients – un secteur au chiffre d'affaires mondial estimé à 150-200 millions d'euros annuels. Dans tous les cas, l'algorithme de traduction automatique décompose le texte en phrases et se révèle donc, selon François Yvon, « incapable de traduire des phénomènes qui dépassent la phrase, par exemple les références pronominales (it en anglais, versus elle ou il en français), les registres de langue ou les choix lexicaux ». Les textes destinés à la publication doivent donc toujours être relus et corrigés par des traducteurs professionnels. Enfin, la traduction automatique échoue – par nature – à restituer les termes inventés ou poétiques, les expressions orales ou les abréviations, les effets humoristiques ou les jeux de mots. Inutile d'y recourir pour traduire des œuvres littéraires !

1.8 La traduction automatique comme outil automatique[5]

Pour aborder la question de la nature de la traduction automatique, on doit poser la question de son utilité et de son utilisation. En effet, si la traduction automatique est « machinale » est-elle vraiment utile ? Les systèmes de TA s'intègrent dans un ensemble plus vaste d'outils d'aide à la traduction. Par conséquent, on doit considérer que la TA est un outil d'aide à la traduction presque au même titre que les dictionnaires électroniques monolingues ou multilingues, les concordanciers, les mémoires de traduction, les vérificateurs orthographiques et grammaticaux, la fonction « recherche » sous Word, etc. (Volanschi, ce volume). On peut considérer l'utilisation de ces divers outils comme étant sur un continuum allant de l'être humain à la machine. Selon que l'on sera plus ou moins proche de l'intervention humaine ou de celle de la machine, la charge de travail sera répartie plus ou moins sur l'humain ou la machine. Se pose alors la question des besoins et des objectifs de la traduction que l'on a à réaliser. Le choix des outils serait donc fonction des besoins ? La faille dans cette logique est que les outils génèrent de nouveaux besoins. Commençons par les besoins et l'évolution du domaine de la traduction au niveau mondial qui justifient l'utilisation de l'ordinateur de manière générale. On trouve chez les experts en TA, mais aussi dans les argumentaires commerciaux des systèmes commercialisés, les explications suivantes:

* Le premier point concerne la quantité de texte à traduire. Aujourd'hui, la quantité de documents qui doivent être traduits est tellement importante et croît si rapidement qu'il n'est pas possible aux traducteurs de tous les traduire. L'ordinateur peut contribuer à augmenter la quantité de documents traduits.

* Nombreux sont les documents qui sont très techniques et assez laborieux à traduire. L'ordinateur peut aider les traducteurs dans cette tâche peut-être ingrate³.

* Les grosses entreprises ont pris conscience, depuis un certain nombre d'années maintenant, que la transmission de l'information multilingue demande une terminologie normalisée et cohérente dans toutes les langues utilisées. Les outils d'aide à la traduction proposent systématiquement les termes normalisés par l'entreprise. Cette cohérence est plus longue à réaliser avec des traducteurs humains qui travaillent en équipe et sans outils d'aide à la traduction.

* Les entreprises exigent toujours davantage de texte traduit, dans des délais de plus en plus brefs. L'ordinateur peut aider à intensifier le rythme et à augmenter le volume de texte traduit

* Enfin, dans certaines situations, il n'est pas toujours nécessaire d'avoir une traduction de haute qualité. La traduction automatique permet de traduire très rapidement des pages web rédigées dans une langue que l'on ne connaît pas, avec un résultat approximatif mais qui permet de comprendre globalement le contenu de la page.

Outre ces constatations sur les besoins au niveau industriel, Hutchins (2004) relève aussi des besoins à un niveau plus individuel. Il répartit donc ces besoins en quatre types qui devraient permettre de choisir le bon outil de traduction:

- * La diffusion
- * L'assimilation (ou la veille)
- * Les échanges
- * L'accès de type base de données

Cette classification reflète aussi l'influence d'un outil informatique sur la génération de nouveaux besoins. Nous verrons que le choix de la traduction automatique a des incidences sur la normalisation des documents source et donc probablement, un impact sur la rédaction technique

- La diffusion :

La diffusion représente le niveau le plus élevé en traduction. Elle consiste à produire des traductions - généralement en quantités importantes et de haute qualité - qui seront publiées. Le niveau de qualité doit approcher la perfection: dans le domaine médical, par exemple, les instructions d'utilisation d'un instrument doivent être claires pour éviter les accidents et les litiges ; il est par ailleurs dans l'intérêt d'une entreprise et de son image de marque de publier des documents de qualité. Il est possible de choisir la TA pour la diffusion ; ce choix requiert cependant un certain nombre d'opérations autour de la traduction elle-même : la pré-édition du texte source

avant la TA, la personnalisation du système de TA avant le processus et la post-édition du résultat de traduction.

Pré-édition du texte source avant la TA:

Les systèmes de TA commercialisés sont des systèmes à règles. Bien que le nombre de règles soit très important et que celles-ci soient constamment corrigées en fonction des tests linguistiques effectués, certaines structures - comme les subordonnées - ou certains phénomènes - comme ceux de la dépendance à longue distance ou l'ellipse - posent problème. Par ailleurs, si le texte source contient des erreurs, le système ne pourra pas rétablir le texte correct. C'est donc à un être humain que revient la tâche de corriger le texte source et, plus radicalement, de le simplifier. La simplification est en effet en relation étroite avec le langage dit contrôlé. Le langage contrôlé est une forme de langue naturelle extrêmement prescriptive et simplificatrice au niveau de la syntaxe et du lexique. Il est largement utilisé dans la rédaction de documentation technique - par exemple, dans l'industrie aéronautique ou automobile - de manière à faciliter, entre autres, le processus de TA. Les questions d'effacement sont difficiles à traiter par un système de TA. L'exemple suivant *cable and ADSL connexion* sera traduit par le *câble et la connexion de l'ADSL*, alors qu'il s'agit de la *connexion par câble et par ADSL*. Les recommandations d'un langage contrôlé demandent à ce que soit rétabli le terme effacé : la *connexion par câble et la connexion par ADSL*. C'est aussi ce qui est recommandé pour la pré-édition d'un texte à faire traduire par un système de TA. De même, l'ellipse du *that* en anglais doit être rétablie

Ce type de problèmes est déjà évoqué par Austermühl (2001 :163) qui cite les recommandations faites par un système de TA commercialisé : *Avoid idiomatic expressions Avoid omitting pronouns before a verb (...) Keep to standard, formal English in which grammatical connections are clearly expressed* Ces recommandations sont encore valables aujourd'hui

1.9 Personnalisation du système par la création d'un dictionnaire spécialisé [5]

Un autre problème rencontré par la TA est en effet celui du lexique spécialisé (terminologie) ou tout simplement du figement des termes dans un domaine donné.

L'exemple ci-dessous montre comment un système peut traduire un texte dans un domaine spécialisé, sans l'aide d'un dictionnaire machine spécialisé

Texte source	Traduction par le système	Traduction corrigé
(...)decoded by specific individuals	(...)décodée par les individus spécifiques	(...)décodée par des individus spécifiques

Le système possède une règle de transfert qui spécifie que le déterminant zéro en anglais doit être traduit par un déterminant défini en français, comme c'est le cas dans des affirmations généralisantes (Life is beautiful => La vie est belle). Or, dans ce cas précis, il ne s'agit pas d'une généralisation mais justement d'une situation dans laquelle on ne connaît pas encore les individus dont il est fait mention. En français, il faudra donc utiliser un déterminant indéfini. Mais la règle de transfert citée ci-dessus a été appliquée de manière sur généralisant.

L'exemple suivant montre comment une règle de transfert sur la traduction du passif anglais par un verbe réfléchi en anglais s'applique à tort:

Texte source	Traduction par le système	Traduction corrigé
which has been termed hypermedia	qui s'est nommée hypermedia	qu'on a appelée hypermedia

Comme mentionné plus haut, les systèmes à transfert peuvent être décrits de manière simplifiée comme fonctionnant avec trois modules : un module d'analyse syntaxique de la langue source qui produit une représentation abstraite de la phrase, un module de transfert contenant des règles de traduction et un module de génération qui, à partir d'une représentation abstraite de la langue cible produite par le module de transfert, génère le phrase de la langue cible. Des erreurs peuvent donc se produire à tous les stades du processus. Nous illustrons ci-dessous des erreurs qui interviennent dès la phase d'analyse de la langue source et se propagent dans les deux autres phases pour donner un résultat de traduction complètement erroné.

Dans l'exemple ci-dessous, on peut penser à première vue que le mot *light* a été mal analysé par le système puisqu'il est traduit par un adjectif, alors qu'il s'agit du nom *light* qui se traduit par lumière. Cette erreur provient en fait d'une erreur d'analyse syntaxique de toute une partie de la phrase et ne porte pas seulement sur un mot. C'est tout le groupe nominal complexe *the phase of the light passing through* qui a été mal analysé. Le mot *passing* qui ici fonctionne comme la forme verbale du verbe *to pass* (en fait *to pass through*) a été analysé comme une nominalisation.

Texte source	Traduction par le système	Traduction corrigé
The liquid crystal changes the phase of the light passing through it	Le cristal liquide change la phase du dépassement léger par lui	Le cristaux liquide modifient la phase de la lumière qui les traverse .

De même, dans la phrase suivante, le mot *can* a été analysé comme un nom, rendant l'analyse correcte de la phrase totalement impossible.

Texte source	Traduction par le système	Traduction corrigé
If such a unit is placed between a pair of plane polariser plates then light can pass through it only if the correct voltage is applied.	Si une telle unité est placée entre une paire de passage léger de bidon de plaques plane polarisées puis par elle seulement si la tension correcte est appliquée	Si on place cette unité entre deux plaques planes polarisées, alors la lumière peut les traverser seulement si on applique le voltage correct .

Ces deux exemples démontrent aussi que si la TA ne va pas au-delà de la phrase, elle ne va parfois même pas jusqu'à la phrase. En effet, certains systèmes de TA commercialisés effectuent une analyse de la phrase lorsque c'est possible, mais sinon, ils fonctionnent plutôt avec l'analyse « par chunks », c'est-à-dire, par morceaux

de phrase (la plupart du temps en tentant de reconnaître des syntagmes plus ou moins complexes). Cela expliquerait tout à fait les erreurs d'analyse ci-dessus .

On voit ici que la production d'une traduction de haute qualité à l'aide de la TA ne peut se faire sans intervention humaine. Rappelons cependant que la révision est un processus qui s'applique aussi aux traductions effectuées par les traducteurs. On peut donc envisager d'utiliser la TA pour produire des traductions publiables, mais il est évident que ce choix est aussi dépendant du domaine dans lequel on traduit. Les dictionnaires spécialisés montrent une bien plus grande utilité dans certains domaines spécifiques par exemple, scientifiques et techniques, que dans d'autres – par exemple, dans les domaines économiques et juridiques

Cela tient sans doute au fait que les systèmes de TA ont été surtout développés pour traduire de la documentation technique ; leurs modules d'analyse, de transfert et de génération ne contiennent donc pas les structures de la langue juridique, souvent obsolètes en langue générale ou scientifique et technique. Par ailleurs, ce n'est pas seulement le type de domaine qui doit déterminer l'emploi ou non d'un système de TA, c'est aussi le genre du texte : des documents techniques, comme des manuels d'utilisation ou des spécifications techniques rédigés en langage contrôlé, ainsi que des articles scientifique très spécialisés et de haut niveau sont bien plus adaptés à la TA que des documents de vulgarisation scientifiques, des cours en ligne, des textes de marketing. Le travail de révision est en effet bien moins important dans le premier type de textes.

1.10 Avantages de la personnalisation

- La personnalisation du logiciel est le processus permettant d'améliorer la qualité de traduction selon les besoins spécifiques de chaque client [10].

- La personnalisation est adaptée pour les clients ayant d'importants besoins de traductions dans un domaine spécifique comme l'automobile, l'aéronautique, l'électronique, la pharmacie, la chimie, ou les services financiers. Elles sont mises en œuvre pour entraîner le logiciel qui apprend comment traduire dans le contexte précis de chaque client [10].

- La personnalisation sert le sens (sémantique).

1.12 Conclusion

Dans ce chapitre nous avons illustré les points les plus importants pour bien comprendre la traduction automatique.

Dans le chapitre suivant, nous allons présenter la notion du contexte et de la grammaire dans les deux langues française et arabe.



Chapitre 02 Grammaire et Contexte.

2.1 Introduction

Ce chapitre présente les grammaires de deux langues (française et arabe) et le contexte.

2.2 La grammaire

2.2.1 Définition

La grammaire est l'étude et la description sous forme de règles des structures syntaxiques, morphologiques et phonétiques.

L'enseignement de la grammaire a pour finalité de favoriser la compréhension [14] et l'analyse des textes lus pour la traduction automatique.

2.2.2 La grammaire de la langue française

2.2.2.1 Étapes de l'analyse d'un texte [11]

Lexicale -> syntaxique -> sémantique -> stylistique

Vocabulaire -> ordre, fonction -> sens -> figure de style

Remarque : La grammaire se situe au niveau de la syntaxe mais elle peut déborder un peu sur l'analyse lexicale ou sémantique.

2.2.2.2 Découpages de la phrase [11]

On peut découper la phrase en sujet (ce dont on parle) et prédicat (ce que l'on en dit).

Les compléments peuvent être découpés en compléments essentiels (liés au verbe, non effaçable, non déplaçable ; COD, COI, attribut du sujet, COS) et compléments circonstanciels (effaçable, déplaçable ; CC).

Ou encore, en compléments de verbe (ne modifient que le verbe) et compléments de phrase (apportent une information concernant l'ensemble de la phrase).

Deux décompositions de la phrase

Sujet + prédicat + complément essentiel/circonstanciel

Groupe verbal + groupe nominal sujet + groupes nominaux compléments

2.2.2.3 Division en deux ensembles des mots français

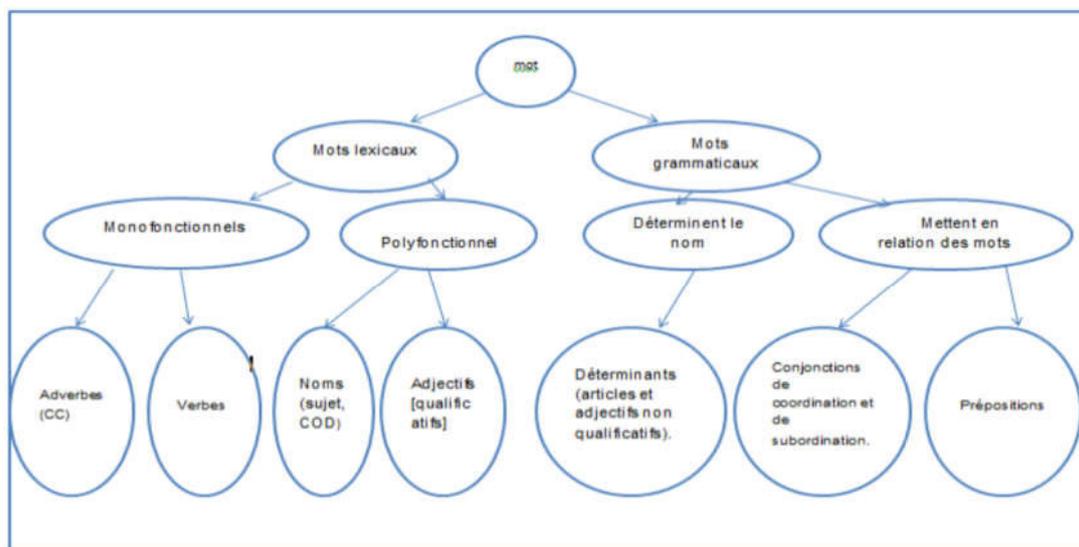


Figure (2.1):Division en deux ensembles des mots français [11].

2.2.2.4 Caractéristiques des éléments [11]

➤ Déterminant

Type essentiel (non déplaçable, non supprimable) ; utilité : indique le genre (sauf certains au pluriel), le nombre et le nom (toujours placé avant le nom) ; s'accorde avec le nom-noyau.

- Classement des Déterminants (selon possibilités des combinaisons):

a- Articles, adjectifs possessifs et démonstratifs.

b- Adjectifs indéfinis, numéraux cardinaux, interrogatifs, exclamatifs.

- Les déterminants d'une même série peuvent être combinés entre eux
 - a- Articles définis, adjectifs démonstratifs, possessifs, exclamatifs
 - b- Articles indéfinis, articles partitifs, adjectifs indéfinis,

adjectifsnuméraux.

➤ **Article**

- défini (le/la/les, contracté : au/aux/du/des).
- indéfini (un/une/des).
- partitif (de + article).

➤ **Adjectif non qualificatif**

- possessifs (ex. mon/ton/son/notre/votre/leur).
- démonstratifs (ex. ce/ces ; nom-ci proximité, nom-là éloignement), numéraux (ordinaux (premier).
- cardinaux (trois cents)).
- indéfinis (nombreux, mot (quelque) ou locution (n'importe quel)).
- Interrogatifs.
- exclamatifs (ex. quel/quels).

➤ **Préposition**

Mise en relation avec le verbe (introduit un CC) ou autre des mots, indique la fonction.

Temps, cause, conséquence, but, condition, comparaison, lieu.

Mot simple ou locution prépositionnelle.

➤ **Conjonction**

- Coordination :

Mise en relation d'éléments de même statut syntaxique. entre élément de même nature et de même fonction.

- Subordination :

Mise en relation d'éléments de statut syntaxique différents ; l'élément subordonné dépend du subordonnant ; préposition uniquement si même nature.

- Juxtaposition :

Mise en relation sans terme de liaison.

➤ **Nom**

C'est le noyau du groupe nominal. Il désigne une personne, un animal ou une chose.

Toujours accompagné d'un déterminant (sauf nom propre).

Ils imposent leur genre et leur nombre.

Classement : animé (humain ou non)/inanimé ; nom propre/commun.

➤ **Pronom**

Le pronom remplace le plus souvent un nom ou un groupe nominal ; mais il peut aussi se substituer à un adjectif ou à une proposition tout entière.

Pronoms personnels (ex. je/tu/il), personnels réfléchis (ex. me/te/se/nous/vous), démonstratifs (ex. ce/ceci), possessifs (ex. le mien/tien/sien), interrogatifs (ex. qui/que/lequel), relatifs (ex. qui/que/quoi/dont)..(

➤ **Adjectif [qualificatif]**

(attribut, épithète, comparatif, superlatif) Modifie le sens des noms.

➤ **Verbe**

Noyau de la phrase, appartient au prédicat.

Exprime une action ou un état.

Les marques des verbes sont la personne, le temps, l'aspect, et le mode, négation. Combinaison des formes = conjugaison.

La fonction des verbes s'oppose à celle des noms, cela est une caractéristique du français. 1er groupe -er, 2e groupe -ir, 3e groupe irrégulier (conjugaison morte). Intransitifs (pas de COD), transitifs (COD possible).

Voix active, passive, pronominale.

Modes personnels			
Indicatif	Présent	passé composé	
	passé simple	passé antérieur	
	Imparfait	plus-que-parfait	
	futur simple	futur antérieur	
Subjonctif	Present	Passé	
	Imparfait	plus-que-parfait	
Conditionnel	Present	passé 1	passé 2
	Present		
Modes impersonnels			
	Infinitif	Participe	Gérondif

Tableau (2.1): les types de temps de verbe.

➤ **Adverbe**

Modifient le sens des verbes (ce sont les seuls), des adjectifs qualificatifs, d'un autre adverbe.

Ils sont invariables.

Adverbe de temps, de quantité, de manière, de lieu, d'affirmation, de négation, de doute.

2.2.2.5 Autres notions fondamentales

- Mode d'énonciation :

Exclamatif (surprise, indignation, admiration), interrogatif, impératif, déclaratif).

- Types d'appartenance dans la phrase :

Attribut (appartient groupe verbal, insuppressible, complément des verbes d'état) ; Épithète (appartient au groupe nominal, supprimable, directement relié au nom) ; Apposition.

- Il existe trois moyens pour marquer la fonction d'un mot ou groupe
Prépositions, position, sens.

- Typologie des subordonnées:

Subordonnées			
au verbe de la principale		à un nom de la principale	
Conjonctive		Relative	
Complément			
Essentiel	non essentiel	difficilement effaçable	Effaçable
Complétives	Circonstancielle	Déterminative	Explicatives ou Appositives

Tableau (2.2):Typologie des subordonnées.

2.2.2 Grammaire en la langue arabe[12]

➤ **L'adjectif de relation (nisba)**

L'adjectif de relation ou la nisba est un procédé syntaxique très fréquent en arabe. Il se construit en ajoutant **سي** . à n'importe quel nom.

لُيْتَانِي لُيْتَان
Libanais Liban

Figure (2.1): Exemple 1 sur L'adjectif de relation

Pour former un adjectif de relation féminin, on rajoute le **ة**.

لُيْتَانِيَّة لُيْتَان
Libanaise Liban

Figure (2.2): Exemple sur L'adjectif de relation en féminin.

Les noms qui se terminent en **ا** comme **فَرَنْسَا** (la France) en termine par **ِي**:

فَرَنْسَا فَرَنْسَا
Français France

Figure (2.3): Exemple 2 sur L'adjectif de relation.

➤ Annexion (Idhāfa)

L'annexion ou l'Idhāfa est également un procédé syntaxique courant en arabe. Il permet de déterminer un nom en l'annexant à un autre en remplissant la fonction de complément de nom, de la grammaire française, exprimée par : d', de, du, des... . L'annexion se construit par la juxtaposition des deux noms. Ex:

كِتَابُ الْعَرَبِيَّةِ

Le livre d'arabe.

Figure (2.4): Exemple 1 sur Annexion.

Ce groupe nominal se compose ainsi des deux termes dits :

-Premier terme : mudhāf .

-Deuxième terme : mudhāfilayhi .

Le premier terme de l'annexion ne prend jamais d'article défini, car il sera déterminé par le deuxième. Son cas varie selon sa position dans la phrase. Le deuxième terme prend impérativement le cas indirect. Si l'ensemble de l'annexion est défini, le deuxième terme est défini.

Le livre d'arabe : كِتَابُ الْعَرَبِيَّةِ

Figure (2.5): Exemple 2 sur Annexion.

Si l'ensemble de l'annexion est indéfini, le deuxième terme est indéfini.

Un livre d'arabe : كِتَابٌ عَرَبِيَّةٌ

Figure (2.6): Exemple 3 sur Annexion.

➤ **Adjectif (na't)**

L'adjectif ou le na't se place impérativement après le nom qu'il qualifie. Il s'accorde avec ce nom en quatre points:

a- En genre : féminin ou masculin

طَالِبَةٌ جَمِيلَةٌ ou طَالِبٌ جَمِيلٌ

Figure (2.7): Exemple sur Adjectif en genre.

b- En nombre : singulier, duel ou pluriel

طَلَابٌ جَمِيلُونَ ou طَالِبٌ جَمِيلٌ

Figure (2.8): Exemple sur Adjectif en nombre.

c- En cas : sujet, direct ou indirect

طَالِبًا جَمِيلًا ou طَالِبٍ جَمِيلٍ ou طَالِبٌ جَمِيلٌ

Figure (2.9): Exemple sur Adjectif en cas.

d- En détermination/Indétermination : les deux termes seront déterminés ou indéterminés.

الطَّالِبُ الْجَمِيلُ ou طَالِبٌ جَمِيلٌ

Figure (2.10): exemple sur Adjectif en détermination/Indétermination.

➤ **Détermination (ta‘rīf)**

La détermination ou le ta‘rīf se réalise en faisant précéder les noms de l’article défini **ال**. Cet article vaut pour tous les noms, qu’ils soient : masculins, féminins, au duel ou au pluriel.

La rose : **الْوَرْدَةُ**
 Le livre : **الْكِتَابُ**
 Les deux roses : **الْوَرْدَتَانِ**
 Les deux livres : **الْكِتَابَانِ**
 Les roses : **الْوُرُودِ**
 Les livres : **الْكِتَابِ**

Figure (2.11): Exemple sur Détermination.

➤ **Indétermination (tankīr)**

L'indétermination ou le tankīr se construit en mettant un tanwīn (doublement de la voyelle finale) sur un nom et en enlevant l'article défini. Le tanwīn se met uniquement sur les noms singuliers (et ceux au pluriel interne), qu'ils soient masculin ou féminin.

Une rose : وَرْدَةٌ (cas indirect), وَرْدَةٌ (cas direct), وَرْدَةٌ (cas sujet)
 Le livre : كِتَابٌ (cas indirect), كِتَابًا (cas direct), كِتَابٌ (cas sujet).
 وَرْدَةٌ : وَرُودٌ/رُودٌ/رُودًا les roses, est le pluriel interne de

Figure (2.12): Exemple sur Indétermination.

➤ **Auxiliaire « être »**

En arabe, l'auxiliaire être n'existe pas (au présent). Pour former des phrases nominales simples, il suffit de juxtaposer le sujet (mubtada') dont on parle et l'information (habar), le concernant.

L'étudiant est beau.
 الطالب جميل.

Figure (2.13): Exemple sur auxiliaire être.

Ces deux fonctions sont cependant marquées par le cas sujet. L'information prend un tanwīn.

➤ **Auxiliaire «avoir »**

En arabe, l'auxiliaire avoir n'existe pas non plus. Pour exprimer la possession,

l'arabe fait appel à une tournure particulière qui se construit par la particule suivie d'un pronom affixe ; le tout suivi de l'objet possédé.

لِ

J'ai une maison
 لي بيت
 Tu as une maison
 لك بيت

Figure (2.14): Exemple sur auxiliaire avoir.

Notez bien qu'il s'agit ici d'une phrase nominale d'où le cas sujet sur le mot بيت, considéré ici comme un sujet. La traduction littérale de cette tournure est:

À toi une maison = Une maison est à moi.

Figure (2.15) : Exemple sur La traduction littérale.

➤ **Pronoms affixes**

Le pronom affixe ou le Dhamīrmuttasil est une « particule » qui s'attache aux trois types de mots:

a- Noms : pour exprimer la possession

كتاب + لي = كتابي
 Mon livre = mon + livre.
 كتاب + لك = كتابك
 Ton livre = Ton + livre.

Figure (2.16): Exemple sur Pronoms affixes en cas de noms.

b- Verbes : pour exprimer le caractère réfléchi:

عَرَفَ + لك = عَرَفَكَ

Il t'a connu : = **Te** + connaitre.

Figure (2.17): Exemple sur Pronoms affixes en cas de verbes.

c- Particules : pour exprimer le sens de la particule relatif à la personne concernée

فيك = في + ك

Figure (2.18): Exemple sur Pronoms affixes en cas de Particules.

Les pronoms affixes sont au nombre de 15:

Singulier ي-ك-ك-ه-ها
 Duel نا-كُما-كُما-هُما-هُما
 Pluriel نا-كم-كن-هم-هن

Figure (2.19): Exemple sur Pronoms affixes.

➤ **Pronoms personnels**

Les pronoms personnels ou le Dhamīrmunfasil sont au nombre de 15. Hormis la distinction entre la première, la deuxième et la troisième personne, on différencie aussi le duel (une paire de personnes ou de choses) et le pluriel (à partir de trois):

Pluriel	Duel	Singulier
تَحْن	تَحْن	أَنَا
أَنْتُمْ	أَنْتُمَا	أَنْتَ
أَنْتُنَّ	أَنْتُمَا	أَنْتِ
هُمْ	هُمَا	هُوَ
هُنَّ	هُمَا	هِيَ

Figure (2.20): Exemple sur Pronoms personnels.

➤ **Temps et modes (Zaman)**

Les trois temps (passé, présent et futur) ne figurent pas dans la grammaire arabe. Les actions, exprimées par des verbes, ont cependant deux aspects. Si l'action est achevée, le verbe se met à l'accompli : alMādhī, l'équivalent du passé simple ou le passé composé français. Si l'action est en cours, le verbe se met à l'inaccompli : al-Mudhāri', équivalent du présent de l'indicatif. Chaque aspect possède plusieurs modes.

- **Infinitif :**

L'infinitif proprement parler n'existe pas en arabe. Pour l'exprimer, on fait appel à une tournure particulière. On donne le verbe conjugué à l'accompli ensuite à l'inaccompli à la troisième personne du singulier, masculin.

Pour dire : **écrire** : on dira : **كَتَبَ / يَكْتُبُ**

Figure (2.21): Exemple sur l'infinitif.

- **Racine :**

La racine ou le Jidhr est la matière première dont se forment les mots (verbes et noms) en arabe. Le plus souvent, à partir d'une racine trilitère, on forme des mots selon un schème (= moule) précis pour désigner un sens. Par exemple, à partir de la

racine: **ك / ت / ب** .l'on forme le verbe **écrire**, **كَتَبَ** selon le schème des verbes de la première forme, en l'occurrence : **فَعَلَ / يَفْعُلُ** .

- **Accompli:**

L'accompli, correspondant au passé composé français/ passé simple, se construit en mettant des désinences verbales à la fin de la racine trilitère.

جلس الطالب
L'étudiant s'est assis

Figure (2.22): Exemple1 sur Accompli.

Singulier
أنا جَلَسْتُ
أَنْتِ جَلَسْتِ
أَنْتِ جَلَسْتِ
هُوَ جَلَسَ
هِيَ جَلَسَتْ

Figure (2.23): Exemple 2 sur Accompli.

- Inaccompli:

Le deuxième mode est l'inaccompli ou le Mudāri' marfū' المضارع المرفوع

Il exprime une action inachevée, une action qui dure ou qui se répète. Le plus souvent, il correspond au présent de l'indicatif français. Ce mode est marqué à la fois par la préfixation d'éléments pronominaux et par des désinences verbales aidant à déterminer, aussi, le genre et le nombre.

أنا أَجْلِسُ
أَنْتِ تَجْلِسِينَ
أَنْتِ تَجْلِسِينَ
هُوَ يَجْلِسُ
هِيَ تَجْلِسُ

Figure (2.24): Exemple sur inaccompli.

- Futur:

Le futur ou le mustaqbal, proprement dit, n'existe pas en arabe. Pour l'exprimer, on fait appel à deux particules, chacune suivie du verbe conjugué à l'inaccompli ; la nuance entre le futur proche et lointain étant plutôt vague.

a- **Futur proche**

verbe+س سأكتب (J'écrirai).

Figure (2.25): Exemple sur futur proche.

b- **Futur lointain**

verbe +سوف : ex. سوف أكتب (Je vais écrire).

Figure (2.26): Exemple sur futur lointain.

➤ **Négation (nafy)**

La négation ou le nafy varie en arabe selon que la phrase est nominale ou verbale. Dans la phrase verbale, la particule de négation varie également selon le mode et l'aspect du verbe. Ici nous nous limitons à la négation des verbes accomplis et inaccomplis :

- **Négation au passé** : Elle se construit en faisant précéder le verbe conjugué à

ما

l'accompli par :

ما كتبتَ الدرسَ

Figure (2.27): Exemple sur Négation au passé.

- **Négation au présent** : Elle se construit en faisant précéder le verbe conjugué à

لا

l'inaccompli par :

لا أكتبُ الدرسَ

Figure (2.28): Exemple sur Négation au présent.

➤ **Formes dérivées**

La langue arabe est une langue **de dérivation** الإشتقاق , verbes sont dérivés d'une racine trilitère (composée de trois lettres ou radicales). À partir de ces trois radicales, on construit un verbe de la forme I الفعل المجرد .

Il est possible d'augmenter cette forme I de quelques lettres pour obtenir des verbes dérivés/augmentés فعل مزيد . Le sens de ces derniers demeure proche de celui de la racine.

➤ **Interjection**

L'interjection ou le nidā' est un procédé syntaxique qui consiste à faire précéder le nom d'une personne par un interpellatif

C'est à dire une particule qui influence le cas du nom. Si ce nom est simple, il prend le cas sujet.

يا أحمدُ *Oh ! Ahmad !*

Figure (2.29): Exemple 1 sur Interjection.

Si ce nom est annexé, le premier terme prend le cas direct:

يا صاحب البيتِ , Oh toi, maître de la maison !

Figure (2.30): Exemple 2 sur Interjection.

Bien entendu, plusieurs autres particules existent et expriment différentes charges affectives comme *وَأَيُّهَا أَيُّهَا، آ...*

Trois cas : Les trois cas ou les harkāt sont les voyelles finales qui prennent les mots dans une phrase. Ces cas marquent les fonctions grammaticales. En arabe, il est trois cas :

- Le cas direct, la fatha, notamment pour marquer le complément d’objet direct.
- Le cas indirect, la kasra, notamment pour marquer le complément indirect. Ce

cas est souvent impliqué par les prépositions: *عَنْ، عَلَى، فِي، ب، لِ، كَ*

- c- Le cas sujet, la dhamma, notamment pour marquer le sujet.

➤ **Types de phrases**

La grammaire arabe distingue entre deux types de phrase :

- a- **Phrase nominale** : sujet + attribut/prédictat:

L'étudiant est beau. الطالب جميل

Figure (2.31): Exemple sur la phrase nominale.

- b -**Phrase verbale** : verbe+ sujet+ (ou sans) complément d’objet (direct ou indirect):

L'étudiant s'est assis. جلس الطالب

Figure (2.32): exemple sur la phrase verbale.

➤ **Conjonctions de coordination ('atf)**

La conjonction de coordination ou le ‘atf est un procédé syntaxique qui permet de coordonner deux noms ou deux phrases. Les trois conjonctions de coordination les plus connues sont :

- **ف** exprimant un laps de temps très bref entre les deux actions:

دَخَلَ فَجَلَسَ

Figure (2.33): Exemple1 sur conjonctions de coordination.

; il est entré et [toute de suite] s’est assis.

- **واو** exprimant un laps de temps moins bref

دَخَلَ وَجَلَسَ

Figure (2.34): Exemple 2 sur conjonctions de coordination.

; il est entré et s’est [un peu de temps après] assis.

- **ثم** exprimant un laps de temps plutôt long entre les deux actions.

دَخَلَ ثُمَّ جَلَسَ

Figure (2.35): Exemple 3 sur conjonctions de coordination.

; il est entré [ensuite] il s’est assis.

- **Participe actif (ism al-fā‘il)**

Le participe actif ou ism al-fā'il désigne celui qui fait l'action.

Pour tous les verbes de la première forme, on construit le participe actif sur

le schème **فَاعِل** .

à partir de **كَتَبَ/يَكْتُبُ** (écrire), on obtient **كَاتِب** (écrivain)

Figure (2.36): Exemple sur participe actif.

➤ **Participe passif (ism al-maf'ūl)**

Le participe passif ou ism al-mf'ūl désigne celui qui subit l'action.

Pour tous les verbes de la première forme, on construit le participe passif sur le

schème **مَفْعُول** .

à partir de **كَتَبَ/يَكْتُبُ** (écrire), on obtient **مَكْتُوب** (chose écrite)

Figure (2.37): Exemple sur participe passif.

➤ **Non d'action (masdar)**

Le nom d'action ou le masdar est un substantif dérivé qui désigne l'action. La formation des noms d'action de la forme I ne suit aucun schème.

قِرَاءَة **قَرَأَ** lire, lecture.
كِتَابَة **كَتَبَ** écrire, écriture

Figure (2.38): Exemple sur non d'action.

Cependant, les noms d'action des 9 formes (verbales) dérivées suivent tous des schèmes constants.

2.3 Le contexte

2.3.1 Définition du contexte

Il existe dans la littérature plusieurs définitions du mot <contexte> suivant les domaines de l'ingénierie. Au regard de la littérature informatique, cette notion a trouvé son essor dans le domaine de l'informatique pervasive ou ambiante où elle a suscité l'intérêt de plusieurs chercheurs [21].

La définition la plus reconnue dans la littérature et proposée par **Dey**, il définit le contexte comme étant : "toutes les informations pouvant être utilisées pour caractériser la situation d'une entité, où une entité est une personne, un lieu, ou un objet qui peut être pertinent pour l'interaction entre l'utilisateur et l'application, y compris l'utilisateur et l'application eux-mêmes [22]"

2.3.2 Les catégories de contexte

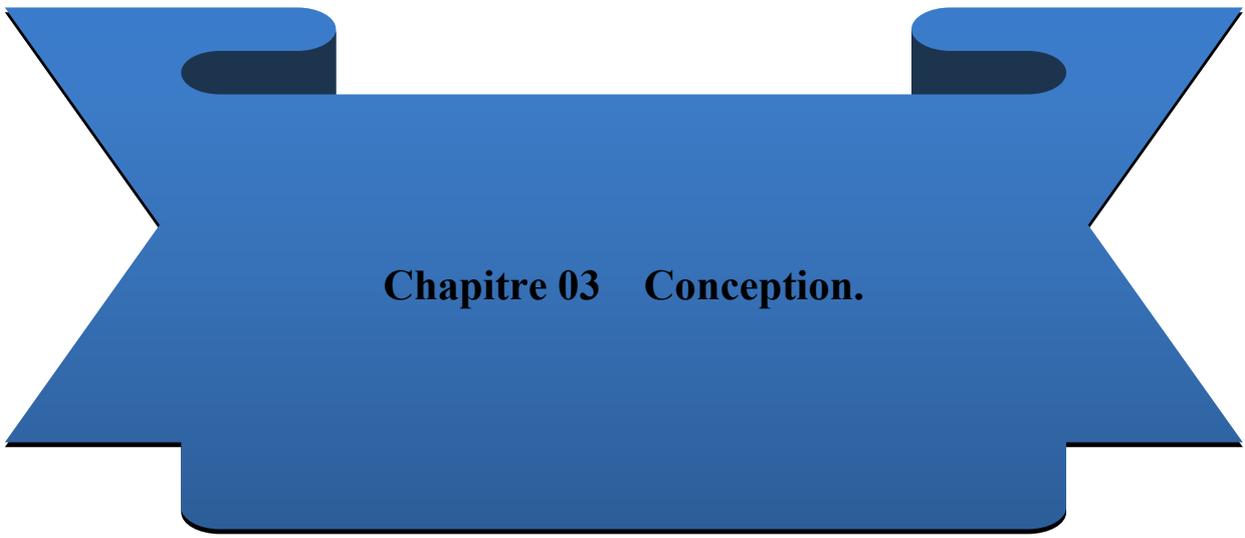
Selon **Tigli** [23] quatre familles de contexte peuvent être identifiées. chacune de ces familles apporte sa propre définition du contexte.

- **Notion de contexte géolocalisé**: la vision de **Schilit** [24] pour le contexte focalise sur la géolocalisation de l'utilisateur. de ceci découlent trois types d'information du contexte.
 - la localisation de l'utilisateur <où est l'utilisateur>.
 - les personnes se trouvant à proximité de l'utilisateur <avec qui ?>.
 - les dispositifs et les ressources que l'utilisateur a à sa disposition <quelles sont les ressources utilisés>
- **Notion de contexte environnementale** : **Brown** [25] relie le contexte aux éléments de l'environnement de l'utilisateur. les autres ajoutent à leur première définition des entités telles que l'heure, la saison, la température, l'identité et la localisation de l'utilisateur. en effet, ce contexte détermine, soit le comportement d'une application, soit les événements issus de l'application et qui intéressent l'utilisateur.

- **Notion de contexte généralisé** :Certaine auteurs ont proposé une définition générale du contexte .**Gong [26]** définit le contexte comme étant ce qui entoure et donne du sens à quelque chose d'autre .Tandis que Winogard [27] définit le contexte comme <l'ensemble d'informations structuré et partagé qui évolue et sert l'interprétation>.

2.4 Conclusion

Dans ce chapitre, on parle d'une façon générale de la grammaire dans les deux langues (arabe, français) et du contexte, pour les utiliser dans le chapitre suivant, qui parle de la conception de notre système.



Chapitre 03 Conception.

3.1 Introduction

La conception est le processus qui consiste à représenter les diverses fonctions du système, c'est certainement la partie la plus importante de notre travail [15].

Dans ce chapitre nous allons présenter la conception de notre système de traduction automatique d'un texte en la langue française vers un texte en la langue arabe (traduction unidirectionnelle) de type fichier TXT avec l'utilisation d'un dictionnaire médical et la méthode de la traduction automatique à base des règles en mode transfert.

Cette conception est présentée par une architecture globale et une architecture détaillée où nous allons expliquer en détail le fonctionnement de notre système.

3.1 Conception globale

Dans cette section, nous allons décrire et présenter le principe et l'architecture générale de notre système.

3.2.1 Architecture générale

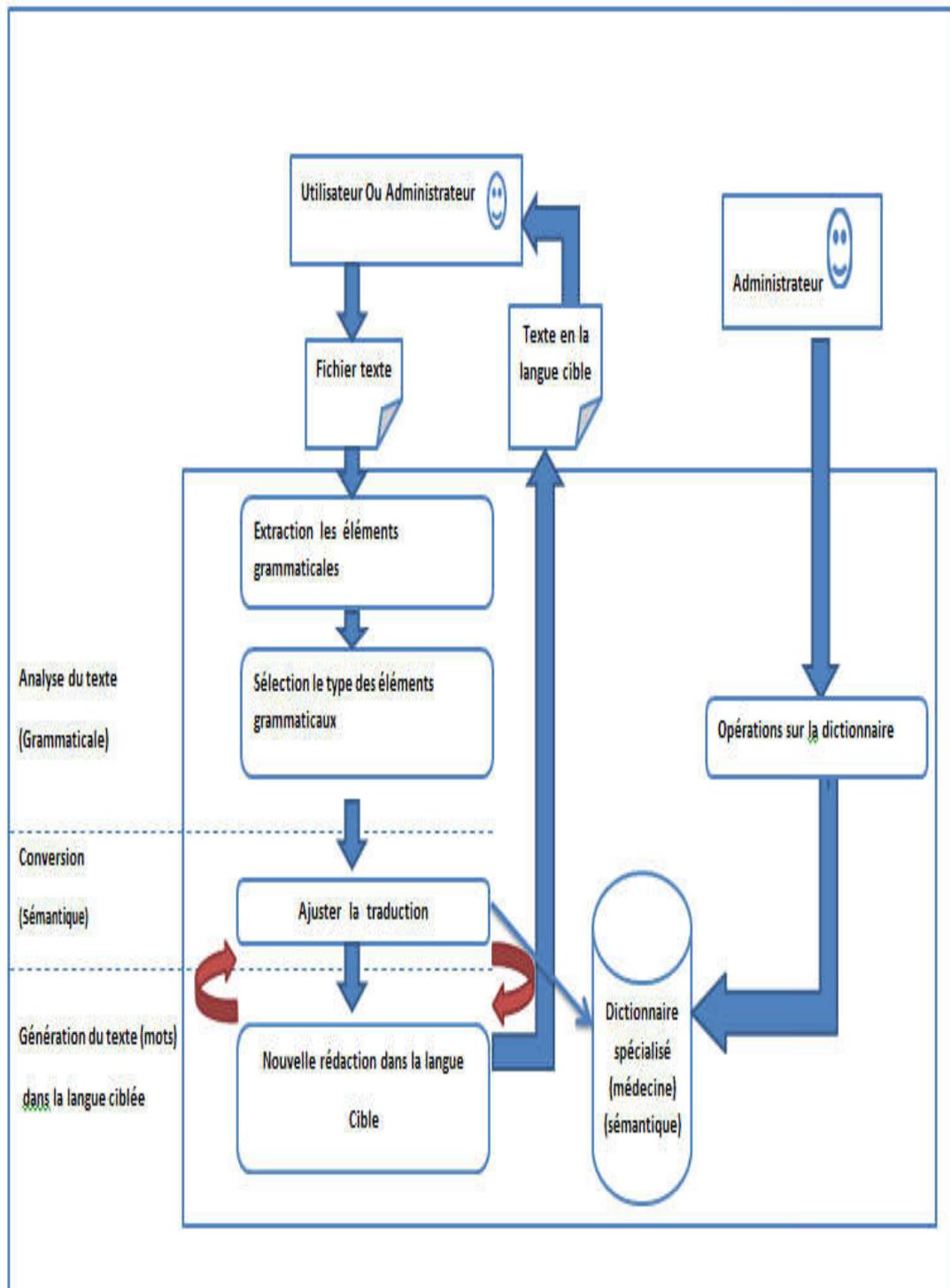


Figure (3.1): Architecture générale de notre système.

Le Figure (3.1) montre la décomposition de système en quatre modules qui sera mentionné ci-dessous où l'administrateur peut traverser le module qui permet des opérations sur base de données mais utilisateur non .

3.2.1.1 Module des opérations sur la base de données

A cette étape, l'administrateur (expert) effectue des opérations sur la base de données telles que suppression, l'ajout et la mise à jour,... etc. il doit avoir de l'expérience dans un domaine (médecine).

3.2.1.2 Module d'analyse du texte

Dans cette étape le contenu de fichier texte (.Txt) est analysé de manière grammaticale à travers l'utilisation des deux sous-étapes suivantes:

- L'étape de l'Extraction des éléments grammaticaux (phrases, mots).
- L'étape de la détermination des types des éléments grammaticaux (phrases nominale, phrase verbale, adjectif, nom, verbe ... etc.).

3.2.1.3 Module de conversion

L'avantage de cette étape est de convertir les mots dans la langue cible afin de préserver le sens sémantique qui leur est attribué à l'aide d'un dictionnaire spécialisé de médecine (synonyme (déduction), le sens de mot dans la langue cible dans le dictionnaire).

3.2.1.4 Module de génération du texte

C'est la dernière étape dans notre système qui permet d'ordonner et d'assembler les phrases et les mots de la langue cible (pronom relatif + nom, pronom relatif + verbe et adjectif + nom ... etc.), à fin d'obtenir le résultat final (texte dans la langue cible).

Remarque : ce module persiste dans la circulation avec le module précédent jusqu'à la fin des mots du texte.

3.3 Conception détaillée

Dans cette section nous allons essayer de détailler les différents modules composants ce système .Les figures suivants présentent l'architecture détaillée de notre système.

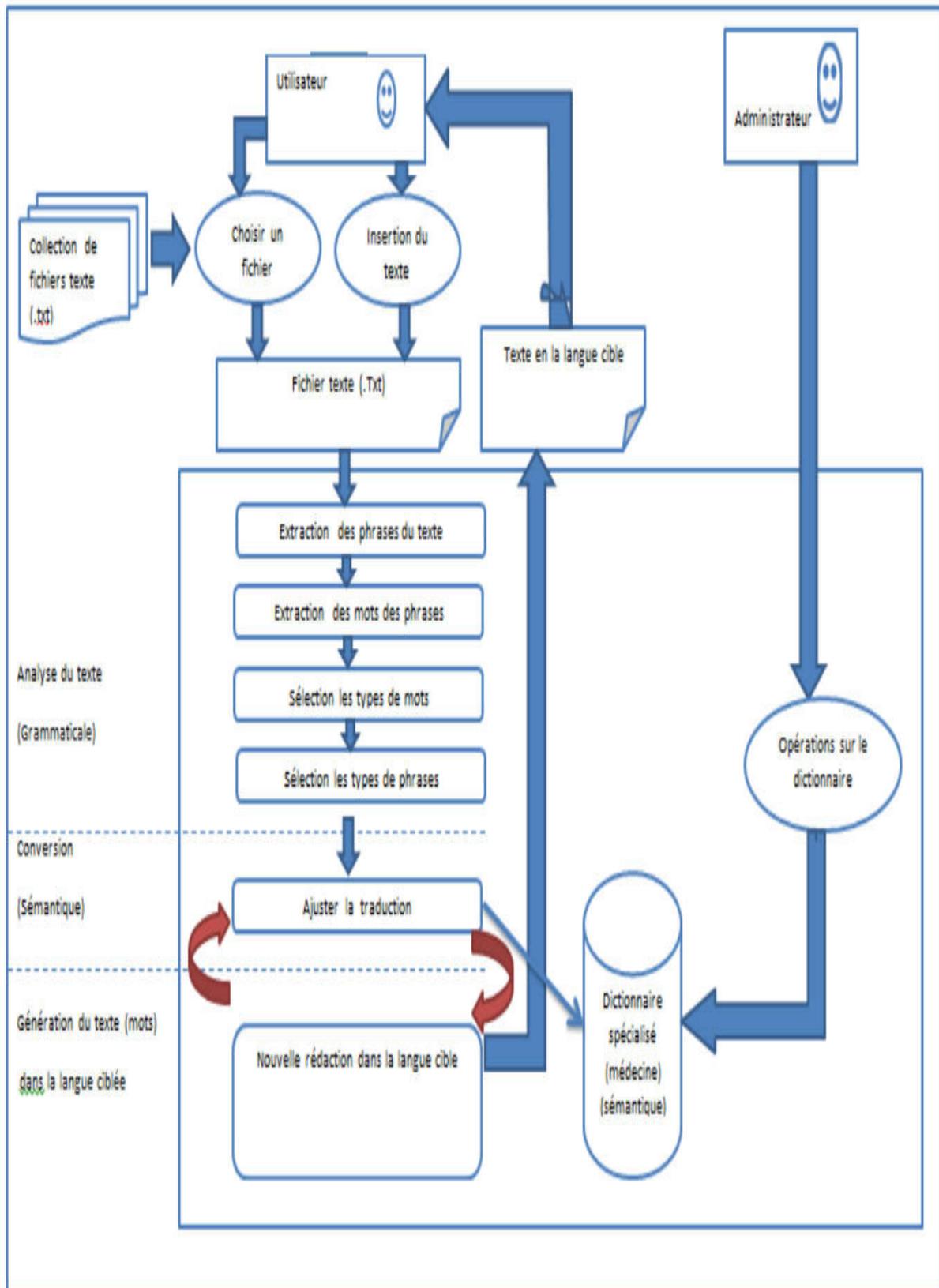


Figure (3.2) : Architecture détaillée de notre système.

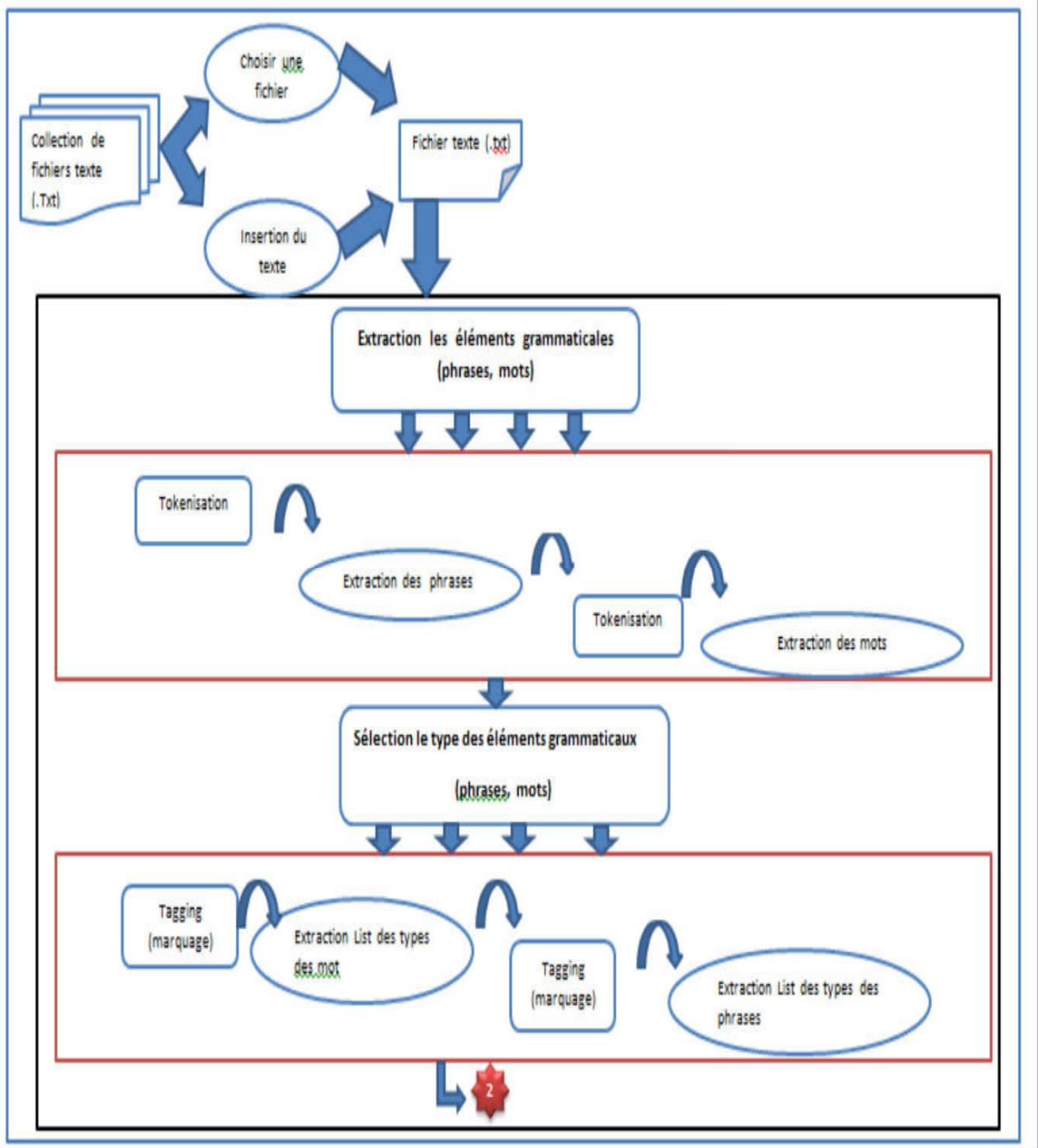


Figure (3.3):Partie 1 d'architecture détaillée de notre système.

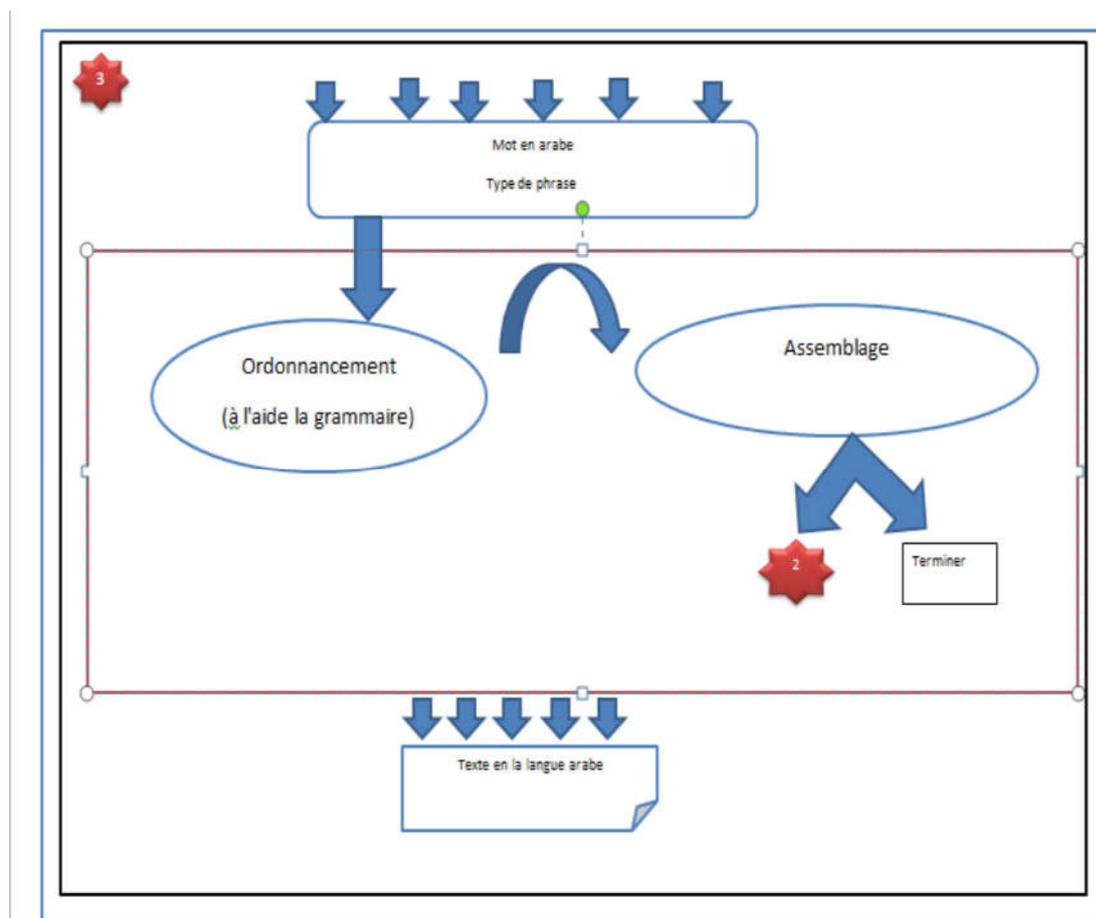


Figure (3.5):Partie 3 d'architecture détaillée de notre système.

3.3.1 L'étape préparatoire

À cette étape, le texte à traduire est importé en choisissant l'une des deux méthodes suivantes :

- Choisir un fichier:
La sélection du texte défini (format .Txt).
- Insertion du texte :
La saisie d'un texte indéfini.

3.3.2 L'étape d'analyse

Elle contient les deux sous-étapes suivantes:

3.3.2.1 Extraction des éléments grammaticaux (phrases, mots)

Dans cette étape, l'exploitation de texte extrait de l'étape précédente : nous extrayons des phrases du texte et des mots de la phrase par Tokenisation pour obtenir une liste des mots et une liste des phrases.

- **Extraction des phrases** : nous savons que dans la grammaire française la phrase commence par une lettre majuscule et se termine par un point et d'après cette règle, nous avons divisé le texte et extrait les phrases, illustré dans la Figure (3.6).

Ma langue arabe est ma belle langue.

Figure (3.6): Exemple illustratif sur l'analyse grammaticale pour trouver les phrases constrictives de texte en la langue française (langue source).

Remarque : nous avons changé la règle de départ (la phrase commence par une lettre majuscule) par une autre règle qui dit (la phrase commence par une lettre majuscule ou minuscule) illustré dans la Figure (3.7).

ma langue arabe est ma belle langue.

Figure (3.7): Exemple illustratif sur le changement de la règle du départ.

- **Extraction des mots** : nous savons aussi que la phrase est un ensemble des mots et que chaque mot est séparé de l'autre par un séparateur (espace), de cette façon les phrases extraites sont divisées en mots, illustré dans la Figure (3.8).

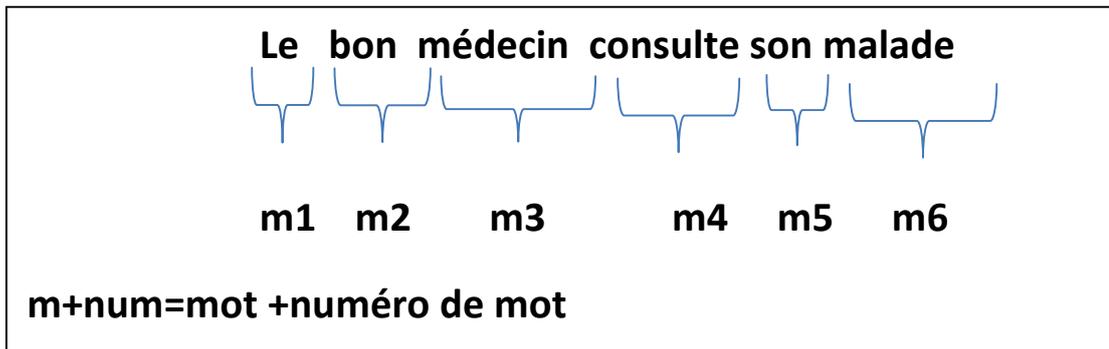


Figure (3.8): Exemple illustratif de l'extraction des mots de la phrase.

3.3.2.2 Sélection des types des éléments grammaticaux (phrases, mots)

Dans cette étape nous utilisons les sorties de l'étape de l'extraction déjà mentionnée.

Au début, nous sélectionnons les types des mots des phrases puis les type des phrases par Tagging (marquage).

- **Extraction de la liste des types des mots :**

On doit préciser chaque type de mot (adjectif, nom, verbe, pronom connecter, ponctuation, complément, déterminent ...etc.) .illustré dans la Figure (3.9).

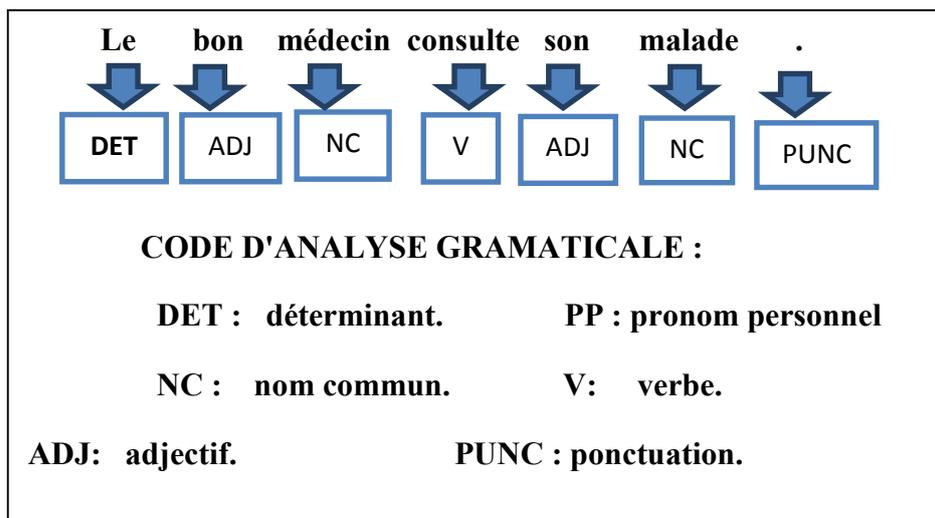


Figure (3.9): Exemple illustratif sur l'analyse grammaticale pour trouver les types des mots.

- **Extraction liste des types des phrases :**

Dans cette dernière étape d'analyse grammaticale, nous exploitons la liste des types des mots pour extraire le type de chaque phrase de l'étape précédente et l'ajouter à la liste des mots et leur type.

Si la liste des mots contient au moins un verbe alors la phrase est une phrase verbale si non c'est une phrase nominale, illustré dans Figure (3.10).

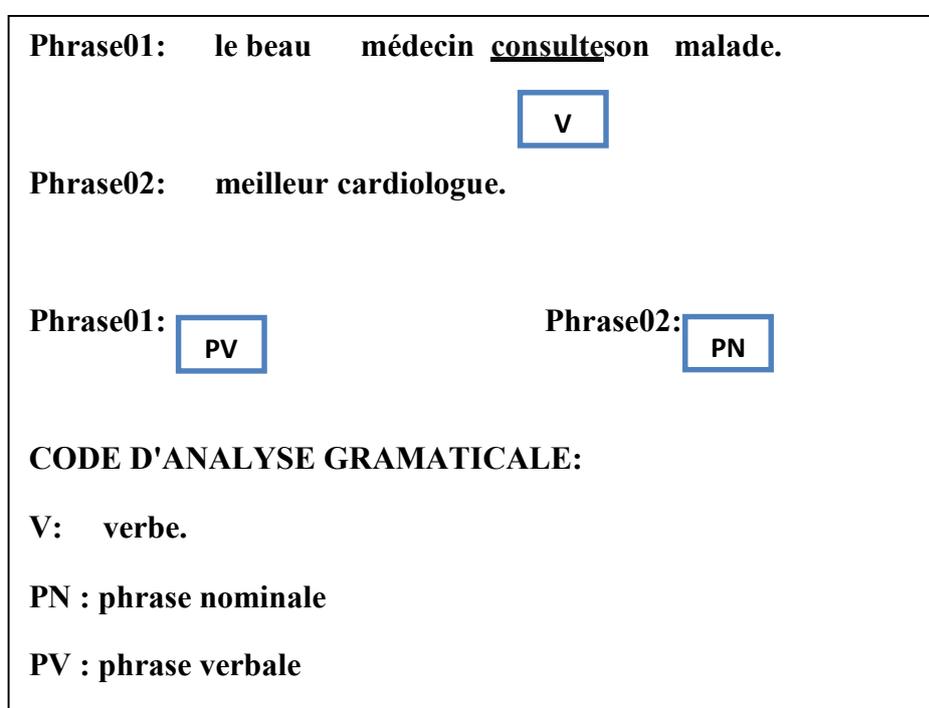


Figure (3.10): Exemple illustratif sur l'analyse grammaticale pour trouver les types des phrases.

3.3.3 L'étape de la conversion

On convertit la sortie (List des mots, types des mots et types des phrases) de l'étape d'analyse à la langue cible, en utilisant un dictionnaire médical. Pour plus d'illustration on suit les étapes suivantes :

3.3.3.1 L'étape de la Sélection

Dans cette étape, nous appliquons certaines opérations sur la Liste des mots, types des mots et types des phrases:

- **La Sélection de la phrase :** on sélectionne une phrase (on conserve le reste jusqu'à la fin de la conversion et la génération de la phrase spécifiée), puis on sélectionne le type de la phrase à fin de déterminer les tableaux de dictionnaire à utiliser.
- **La Sélection les type des mots et les mots de la phrase :**
Après l'étape précédente, on sélectionne les mots et leurs types de la phrase choisie. On appelle la sortie de cette étape: <liste temporaire>.
- **La Sélection d'un mot et son type :**
On sélectionne un mot et son type à partir de la liste temporaire (on conserve le reste jusqu'à la fin de la conversion et la génération du mot spécifique).

3.3.3.2 L'étape de la comparaison

On compare le mot obtenu de l'étape précédente avec son équivalent dans le dictionnaire médical (en utilisant les tableaux spécifiques du dictionnaire médical avec le respect des conditions de l'utilisation des tableaux), puis on doit passer par l'étape du test pour choisir la prochaine étape.

➤ les conditions de l'utilisation des tableaux sont les suivantes :

- si le type de phrase est nominal, on n'utilise pas les tableaux spécifiques pour les verbes et synonymes.

- si le type de phrase est verbal, on n'utilise pas le tableau spécifique uniquement pour les synonymes.

3.3.3.3 L'étape du test

Si on trouve l'équivalent, la réponse retournée est oui sinon la réponse retournée est non.

3.3.3.4 L'étape de l'extraction

On extrait l'équivalent du mot en la langue arabe, dans cette étape si la réponse de l'étape de test est retournée <oui> (En utilisant les tableaux spécifiques du dictionnaire médical avec le respect des conditions de l'utilisation des tableaux)

Remarque importante : dans cette étape, on essaye de faire entrer la sémantique avec l'utilisation d'un dictionnaire spécialisé dans la médecine.

ne pour ajuster le sens exact illustré dans la Figure (3.11).



Figure (3.11): Exemple sur la sémantique dans le cas de l'utilisation du dictionnaire spécialisé dans la médecine.

Dans la figure (3.11) le verbe présente un sens **يظهر** dans un contexte général mais un sens différent dans un contexte médical **يشكو**.

3.3.3.5 L'étape de la recherche du synonyme

Dans le cas où la réponse de l'étape de test est retournée <non>, nous cherchons le mot en la langue française dans le dictionnaire médical précisément dans le tableau des synonymes pour trouver son synonyme:

- Si on ne trouve pas le mot en la langue française, on passe à l'étape suivante.
- Si on trouve le mot en la langue française, on va chercher son synonyme dans les autres tableaux du dictionnaire médical (avec le respect des conditions de l'utilisation des tableaux (déjà vues)):
 - Si on trouve le synonyme, on retourne à l'étape de la comparaison.
 - Sinon, on passe l'étape suivante.

Remarque importante : Dans cette étape, on essaye de faire entrer la sémantique (l'utilisation des synonymes (déduction) et le processus précédent pour gérer ces synonymes).

Remarque : Dans le cas où le mot est un verbe, on a deux possibilités :

- en cas où le verbe est dans le temps impératif (la phrase commence par un verbe), on va chercher dans le tableau des verbes impératif.

- en cas où le verbe est dans le temps présent, on va faire une analyse grammaticale à fin de trouver des solutions aux problèmes suivants :

Pb1: si le verbe est précédé par un pronom relatif, on va chercher le sujet (qui conjugue le verbe).

Nous avons un médecin qui aime ses malades
Pronom + verbe+ déterminant + nom + pronom
relatif +verbe + adjectif + nom

Figure (3.12): Illustration du problème en cas où le verbe est précédé par un pronom relatif.

Solution1: on conjugue le verbe avec le nom qui précède juste le pronom relatif, on utilisant le tableau des verbes au présent à condition de préciser le genre de nom (féminin ou masculin)(l'utilisation de tableaux des noms féminins et masculins).

Nous avons un médecin qui aime ses malades.
Pronom + verbe+ déterminant + nom + pronom
relatif +verbe + adjectif + nom +ponctuation

Figure (3.13): Illustration de la solution en cas où le verbe est précédé par un pronom relatif.

3.3.4 L'étape de la génération

Après la conversion de mot, on passe à la phase de la reformulation de ce mot.

3.3.4.1 L'étape de la reformulation

3.3.4.1.1 L'étape de l'ordonnancement

On ordonne le mot de la phrase à partir des règles grammaticales de la langue cible.

D'une autre façon quel est le meilleur placement pour mettre le mot dans la liste des mots de la phrase cible (la langue arabe), on utilisant les règles grammaticales de la langue cible ? .pour détailler ,Nous donnons les exemples suivants :

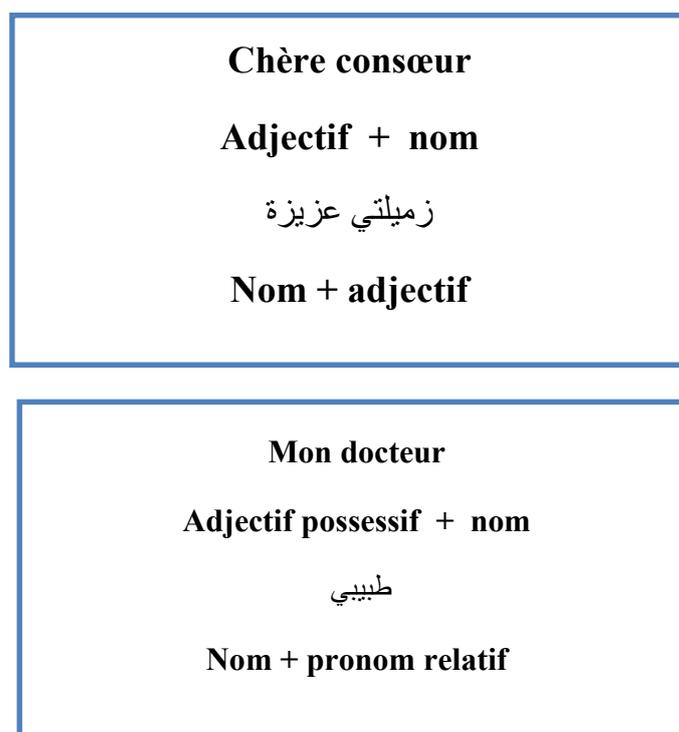


Figure (3.14): Exemples d'uncas d'un adjectif avec un nom.

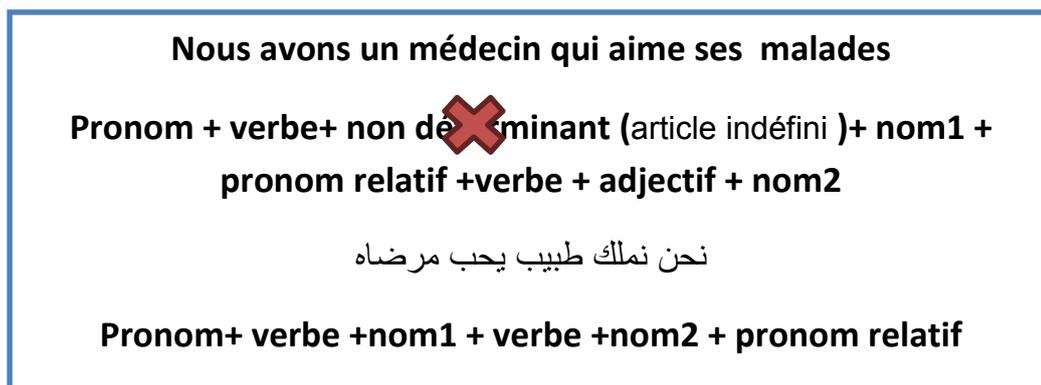


Figure (3.15): Exemple d'un cas d'élimination d'un article indéfini (un, une).

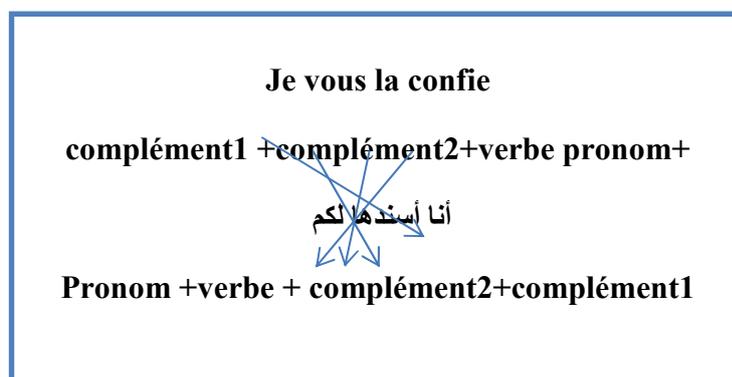


Figure (3.16): Exemple 4 en cas de complément.

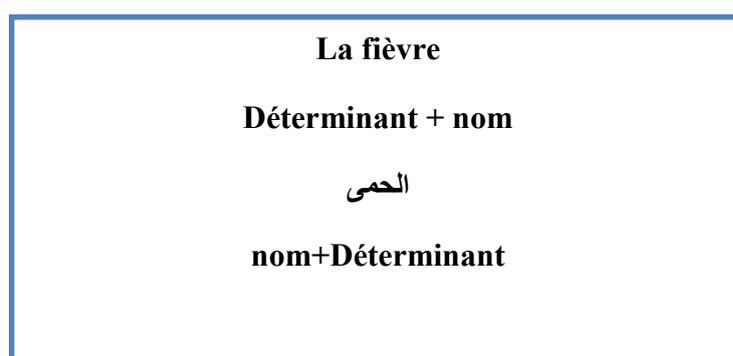


Figure (3.17): Exemple 5 En cas de déterminant et nom (la, le, les, des, de...Etc.).

3.3.4.1.2 L'étape de l'assemblage

- cette dernière étape est caractérisée par la mise du mot dans sa place (trouvée dans l'étape précédente).

Ensuite,

- Si la liste temporaire n'est pas vide, on retourne à la phase de Sélection d'un mot et leur type.
- Si la liste temporaire est vide, on retourne à la phase de Sélection d'une phrase jusqu'à la fin de toutes les phrases de texte, en fin le retourne en la langue cible.

3.3.5 Module des opérations sur la base de données

Dans cette partie, l'utilisateur effectue des opérations sur la base de données comme suivant :

3.3.5.1 Opération de la suppression

Pour faire une suppression dans la base de données, il faut d'abord identifier un id (primary) puis supprimer tous les mots qui portent le même id du même tableau.

3.3.5.2 Opération de l'ajout

Pour faire un ajout d'un mot dans la base de données, il faut identifier les suivants :

- un identificateur (id primary).
- Si on a le tableau des verbes en présent, on identifie le verbe en français, le verbe en arabe et le pronom.

Sinon

Si on a le tableau des synonymes, on identifie le mot en français (synonyme1) et le mot en français (synonyme2).

Si on a un autre tableau, on identifie le mot en français et le mot en arabe.

3.3.5.3 Opération de la mise à jour

Pour faire une mise à jour d'un mot dans la base de données, il faut identifier les suivants :

- un identificateur (id primary).
- Que veux-tu changer ?

Si on veut changer les synonymes, il faut identifier le synonyme 1 ou le synonyme 2 par l'utilisation 1 du tableau des synonymes.

Sinon

-Si on veut changer les verbes en présent, il faut identifier le verbe en français, le Verbe en arabe ou le pronom par l'utilisation du tableau des verbes en présent.

-Sinon il faut identifier le mot en français ou le mot en arabe par l'utilisation du tableau (dictionnaire, féminin ou masculin).

3.3.5.4 Opération de l'authentification

Elle contient trois opérations :

- **L'opération de la connexion** : pour entrer au système et effectuer des opérations sur la base de données, il faut d'abord faire entrer des informations sur l'expert (numéro de certificat (primary) et mot de passe).
- **L'opération de la déconnexion** : pour quitter notre système
- **L'opération de l'inscription** : pour effectuer l'opération de la connexion, il faut d'abord inscrire l'expert (médecin) par quelques informations importantes sur ce dernier, comme :
 - nom
 - email : à fin de se connecter au moment nécessaire.
 - numéro de certificat médical : qui est primary à fin de s'assurer qu'il est un expert (médecin).
 - mot de passe : pour la sécurité.
 -

3.4 Modélisation de système

Cette partie est consacré à l'étape de modélisation, pour cela on va utiliser le formalisme d'UML (Unified Modeling Language) pour la modélisation de notre système .

3.4.1 Diagrammes de cas d'utilisation

Les diagrammes de cas d'utilisation sont des diagrammes UML utilisés pour donner une vision globale du comportement fonctionnel d'un système logiciel .un cas d'utilisation décrit une séquence d'actions réalisées par le système qui produit un résultat observable pour un acteur [25].

3.4.1.1 Identification des acteurs

Les deux acteurs de notre système sont les suivants :

- **L'administrateur:**

L'administrateur c'est l'expert de la base de données qui dispose de toutes les fonctionnalités nécessaires à la gestion de la base de données (ajout, suppression, mise à jour...etc.).

L'administrateur de notre système assure les fonctionnalités suivantes:

- Authentification (connexion, inscription, déconnexion).
- Gérer la base de données (suppression, ajout, mise à jour... etc.).
- Traduction (analyse grammaticale, conversion, génération).
-

- **L'utilisateur :**

- Traduction (analyse grammaticale, conversion, génération).

3.4.1.2 Identification des cas d'utilisation

Chaqueun des acteurs précédents a des tâches précises dans le système.

Le diagramme de cas d'utilisation qui décrit les fonctionnalités de l'administrateur, voir la Figure (3.18):

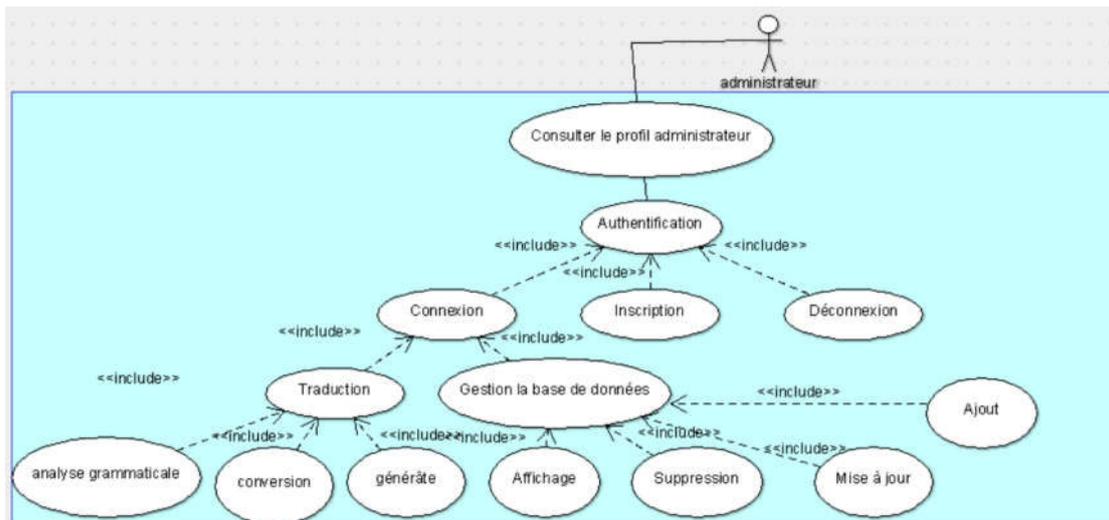


Figure (3.18): Diagramme de cas d'utilisation de l'administrateur du système.

Le diagramme de cas d'utilisation qui décrit les fonctionnalités de l'utilisateur, voir la Figure (3.19):

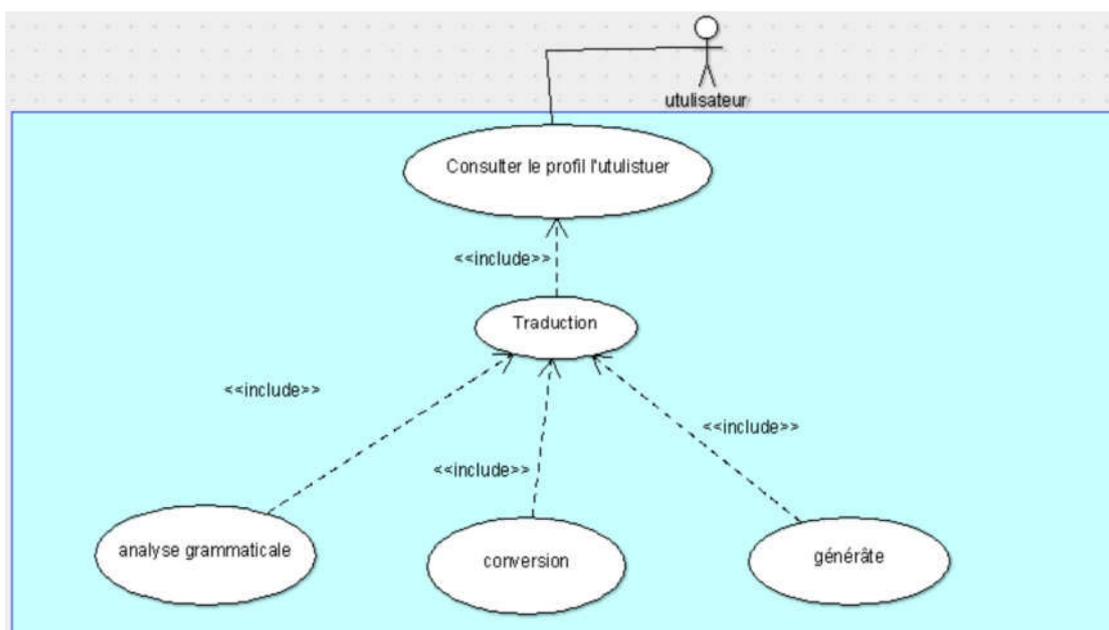
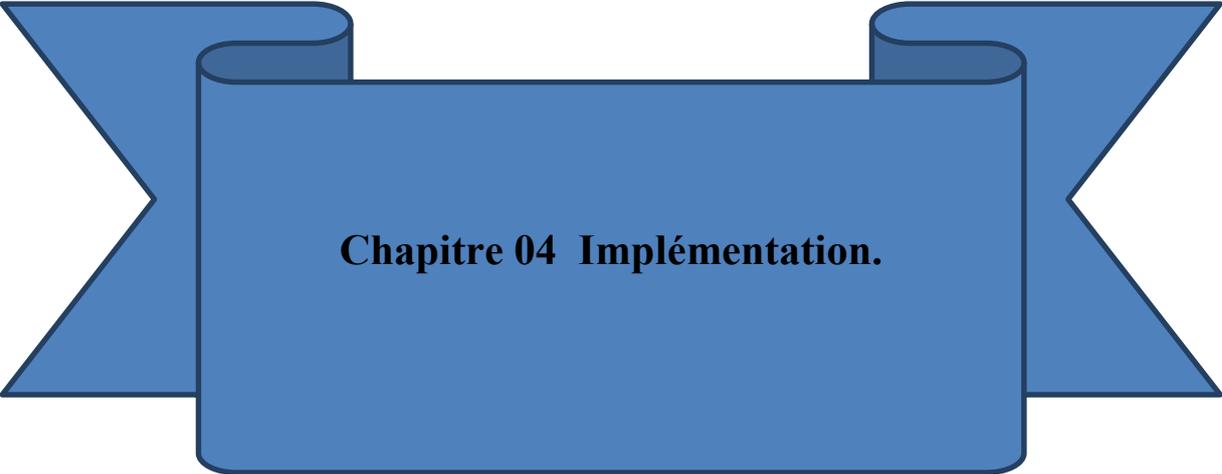


Figure (3.19): Diagramme de cas d'utilisation de l'utilisateur du système.

3.5 Conclusion

Dans ce chapitre, nous avons décrit les besoins fonctionnels et techniques attendus de notre système. Nous avons détaillé le fonctionnement de tous les processus qui constituent ce système.

Dans le chapitre suivant nous allons présenter l'implémentation de tous les composants et les modules de notre système.



Chapitre 04 Implémentation.

4.1 Introduction

Après l'étude conceptuelle, nous allons présenter dans ce chapitre la phase de la réalisation et de l'implémentation de notre système.

Ce chapitre a pour objectif de présenter l'aspect implémentation de notre application, il s'agit donc d'expliquer l'environnement matériel sur lequel notre système a été développé, les langages de programmation et les outils utilisés. Ensuite, nous allons présenter les interfaces graphiques en décrivant les différentes fonctionnalités de notre application et nous présenterons deux exemples qui nous permettront d'illustrer les résultats obtenus lors de l'utilisation de notre application [15].

4.2 Langage de programmation

4.2.1 python [16]

Le langage de programmation Python a été créé en 1989 par Guido van Rossum, aux Pays-Bas. Le nom Python vient d'un hommage à la série télévisée Monty Python's Flying Circus dont G. van Rossum est fan. La première version publique de ce langage a été publiée en 1991. La dernière version de Python est la version 3. Plus précisément, la version 3.7 a été publiée en juin 2018. La version 2 de Python est désormais obsolète et cessera d'être maintenue après le 1er janvier 2020. Dans la mesure du possible évitez de l'utiliser. La Python Software Foundation ¹ est l'association qui organise le développement de Python et anime la communauté de développeurs et d'utilisateurs.

Ce langage de programmation présente de nombreuses caractéristiques intéressantes :

- Il est multiplateforme. C'est-à-dire qu'il fonctionne sur de nombreux systèmes d'exploitation : Windows, Mac OS X, Linux, Android, iOS, depuis les mini-ordinateurs Raspberry Pi jusqu'aux supercalculateurs.

- Il est gratuit. Vous pouvez l'installer sur autant d'ordinateurs que vous voulez (même sur votre téléphone !).

- C'est un langage de haut niveau. Il demande relativement peu de connaissance sur le fonctionnement d'un ordinateur pour être utilisé.

— C'est un langage interprété. Un script Python n'a pas besoin d'être compilé pour être exécuté, contrairement à des langages comme le C ou le C++.

— Il est orienté objet. C'est-à-dire qu'il est possible de concevoir en Python des entités qui miment celles du monde réel (une cellule, une protéine, un atome, etc.) avec un certain nombre de règles de fonctionnement et d'interactions.

— Il est relativement simple à prendre en main.

— Enfin, il est très utilisé en bio-informatique et plus généralement en analyse de données. Toutes ces caractéristiques font que Python est désormais enseigné dans de nombreuses formations, depuis l'enseignement secondaire jusqu'à l'enseignement supérieur.



Figure (4.1): le langage de la programmation python.

4.3 Les environnements de développement

Afin de réaliser notre application, on a utilisé les environnements de développements suivants :

4.3.1 PyCharm[17]

PyCharm est l'IDE le plus populaire pour Python. Il comprend d'excellentes fonctionnalités telles que l'achèvement et l'inspection de code avec un débogueur avancé, la prise en charge de la programmation Web et de divers cadres.

PyCharm est créé par la société tchèque Jet cerveaux qui se concentre sur la création d'un environnement de développement intégré pour divers langages de développement Web tels que JavaScript et PHP .

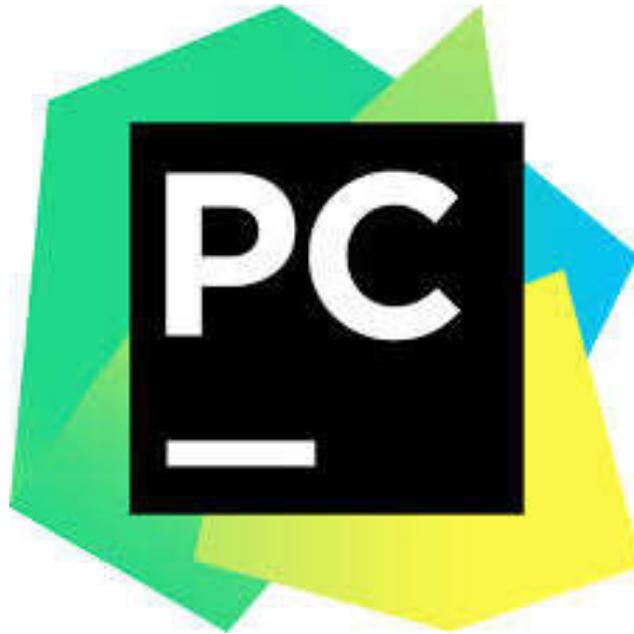


Figure (4.2): la Platform PyCharm .

4.3.2 XAMPP

XAMPP est un ensemble de logiciels permettant de facilement créer une interface web interagissant avec une base de données SQL[18].



Figure (4.3): la Platform XAMPP.

4.3.3 NLPNET

NLPNET est une bibliothèque Python pour les tâches de traitement de langage naturel basées sur des réseaux de neurones. Actuellement, il effectue le marquage d'une partie de la parole, le marquage du rôle sémantique, l'analyse de la dépendance,

Tokenisation , détection de langue, segmentation de phrase ,l'extraction d'entités nommées ,découpage et analyse syntaxique La plupart de l'architecture est indépendante de la langue, mais certaines fonctions ont été spécialement conçues pour travailler avec le portugais. Ce system a été inspiré par SENNA. [19][15].

4.3.4 Tkinter

Le module Tkinter («interface Tk») est l'interface Python standard de la boîte à outils de l'interface graphique Tk et Tkinter sont disponibles sur la plupart des plateformes Unix, ainsi que sur les systèmes Windows.

4.4 Les principales fenêtres de l'application

4.4.1 Les fenêtres de l'accueil

Les figures ci-dessous présentent les fenêtres de l'accueil de notre application.

L'interface d'accueil dans la Figure (4.4) permet l'accès à l'interface principale de la Figure (4.5) par une simple clique sur le Botton (tabib).

Dans la page principale présenté par la Figure (4.5), il y a deux types d'accès:

- l'accès à travers l'administrateur (expert).
- l'accès à travers l'utilisateur.



Figure (4.4):Fenêtre d'accueil de l'application.



Figure (4.5):Fenêtre principale de l'administrateur et de l'utilisateur.

Quand on clique sur le bouton Administrateur de la Figure (4.5), La figure (4.6) apparait.



Figure (4.6):Fenêtre d'accueil de l'administrateur.

Quand on clique sur le bouton Utilisateur de la Figure (4.5), La figure (4.7)apparait.



Figure (4.7):Fenêtre d'accueil de l'utilisateur.

4.4.2 La fenêtre d'authentification de l'administrateur

Quand on clique sur le bouton <bases des données >de La figure (4.6), La figure (4.8) apparait. Dans ce cas l'administrateur doit passer à travers l'authentification par:

- l'opération de la connexion si il a déjà fait l'opération de l'inscription (remplir le champ de numéro de certificat et de mot de passe, puis cliquer sur le Botton Connexion),sinon il doit faired'abord l'opération de l'inscription (remplir le champ de nom, email, mot de passe et de numéro de certificat, puis cliquer sur le Botton Inscription).

Pour l'opération de déconnexion, on clique sur le bouton Retour de la Figure (4.8).

Pour supprimer les champs,on clique sur le bouton Supprimer Figure (4.8).

traduction automatique

Connexion :

Numéro de certificat :

Mot de pass :

Supprimer Connexion

Inscription :

Name :

Email :

Mot de pass :

Numéro de certificat :

Supprimer Inscription

Retour

تنشيط Windows
انتقل إلى الإعدادات لتنشيط Windows

Figure (4.8):Fenêtre d'authentification.

4.4.3 Les fenêtres des opérations sur les bases de données

Quand on clique sur le bouton Connexion dans de la Figure (4.8),La figure (4.9) apparait.

traduction automatique

 **TABIB** 

AJOUTER 

MISE À JOUR 

SUPPRIMER 

RETOUR 

تنشيط Windows
انتقل إلى الإعدادات لتنشيط Windows

Figure (4.9): Fenêtre de choix de l'opération.

D'après la Figure (4.9), on choisit l'opération à faire.

4.4.3.1 La fenêtre de l'ajout

Quand on clique sur le bouton Ajouter de la Figure (4.8), La figure (4.10) apparait.

Après avoir remplie les champs nécessaires de laFigure (4.10), on clique sur le bouton Ajouter de la Figure (4.10), l'opération de l'ajout démarre.

Quand on clique sur le bouton Affiche de la Figure (4.10), La base de données s'affiche.

Pour supprimer les champs,on clique sur le bouton Supprimer Figure(4.10).

The screenshot shows a web application window titled "traduction automatique". At the top center is the "TABIB" logo. Below the logo, there is a dropdown menu labeled "Sélectionner la base de données:" with the selected option "Base de donnée des noms masculin". To the left of the main content area, there are four input fields with labels: "mot en francais:", "mot en arabe:", "Pronom:", and "Synonyme de mot en francais:". To the right of these fields is a large, empty rectangular box labeled "BD". At the bottom of the window, there are three buttons: "Afficher la BD", "Ajouter", and "Retour".

Figure (4.10):Fenêtre de l'ajout.

4.4.3.2 Fenêtre de suppression

Quand on clique sur le bouton Supprimer de la Figure (4.9), La figure (4.11) apparait.

Après avoir remplie les champs nécessaires de laFigure (4.11), on clique sur le bouton Supprimerde la Figure (4.11), l'opération de la suppression démarre.

Quand on clique sur le bouton Affiche de la Figure (4.), La base de données s'affiche.

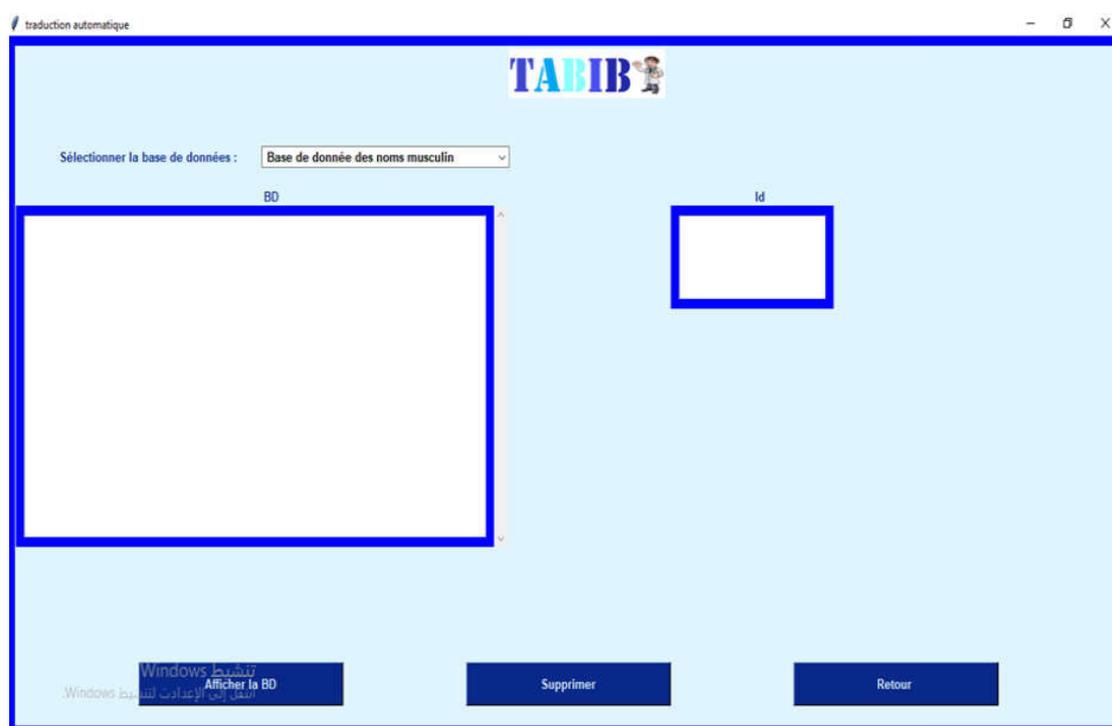


Figure (4.11) : Fenêtre de la suppression.

4.4.3.3 Fenêtre de la mise à jour

Quand on clique sur le bouton Mise à jour de la Figure (4.9), La figure (4.12) apparait.

Après avoir remplie les champs nécessaires de laFigure (4.12), on clique sur le bouton Mise à jour de la Figure (4.12), l'opération de la mise à jour démarre.

Quand on clique sur le bouton Affiche de la Figure (4.12), La base de données s'affiche.

Pour supprimer les champs, on clique sur le bouton Supprimer Figure(4.12).



Figure (4.12):Fenêtre de la mise à jour.

4.4.4 Fenêtre de la traduction

Cette fenêtre peut apparaître par l'administrateur ou l'utilisateur de la façon suivante:

- **Par l'administrateur** :quand on clique sur le bouton Traduction de la Figure (4.7), La figure (4.13) apparaît.
- **Par l'utilisateur** :quand on clique sur le bouton Traduction de la Figure (4.8), La figure (4.13) apparaît.



Figure (4.13):Fenêtre de la Traduction d'un texte.

Dans la Figure (2.13) la traction de texte passe par des opérations, comme suit :

- a. **L'opération préparatoire** :On choisit le texte à traduire illustré dans les Figure (4.14).

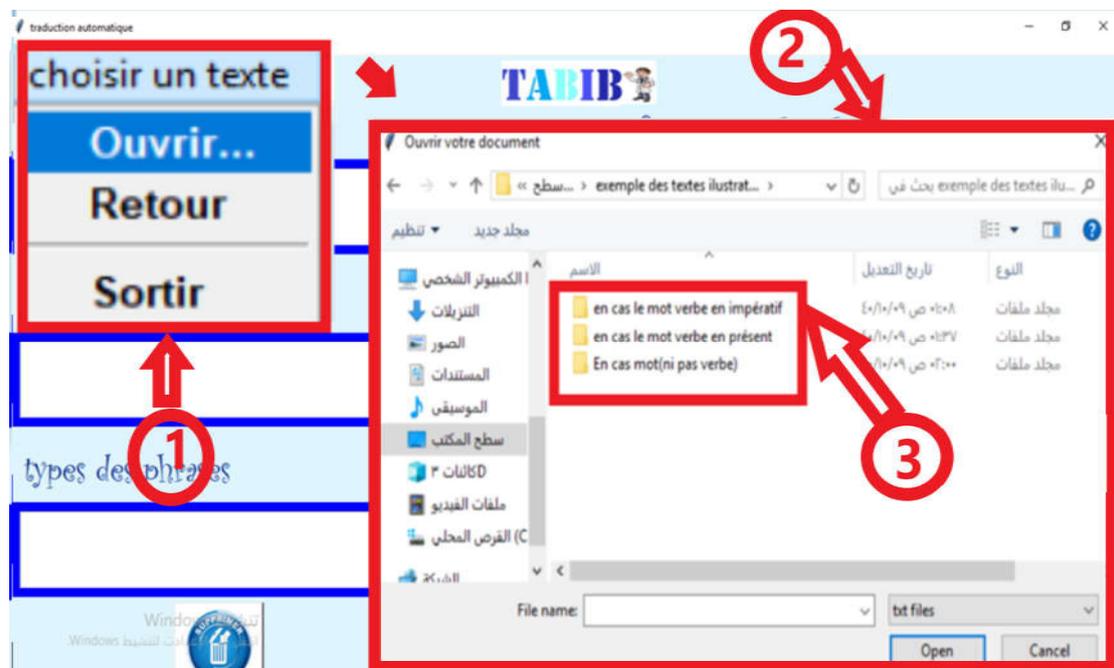


Figure (4.14):Fenêtre de la Traduction d'un texte dans L'opération préparatoire.

Quand on clique sur le bouton (1) de Figure (4.15), le texte va apparaitre sous forme des phrases illustrées dans la Figure (4.16).

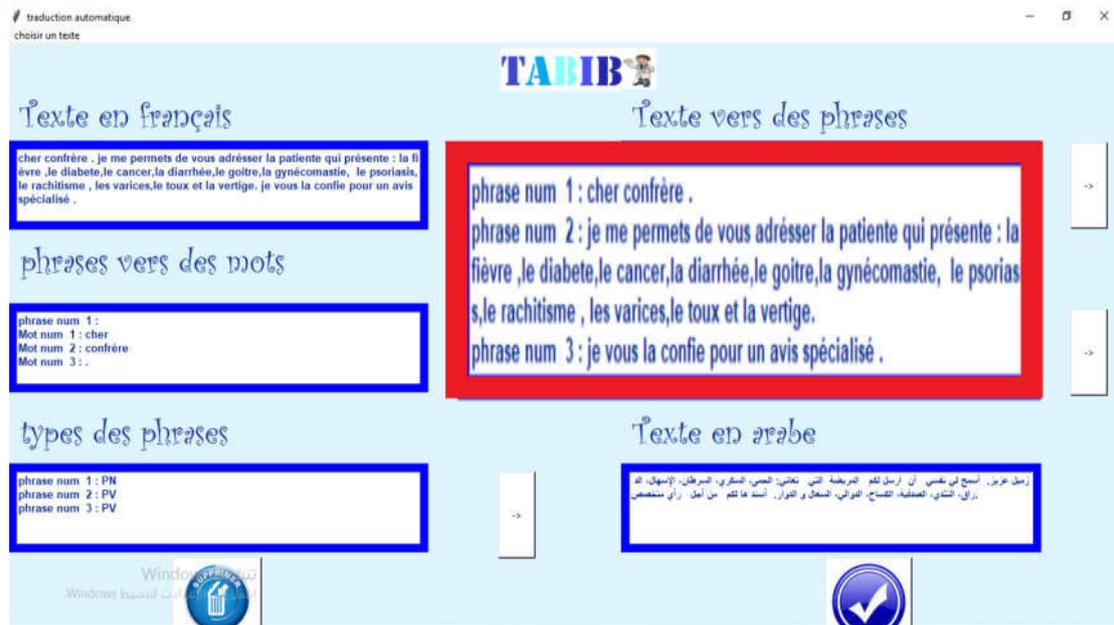


Figure (4.16):Fenêtre de la Traduction d'un texte dans L'opération d'analyse (cas de transfert d'un texte vers des phrases).

Quand on clique sur le bouton (2) de Figure (4.15), les phrases vont apparaitre sous forme des mots illustrés dans la Figure (4.17).

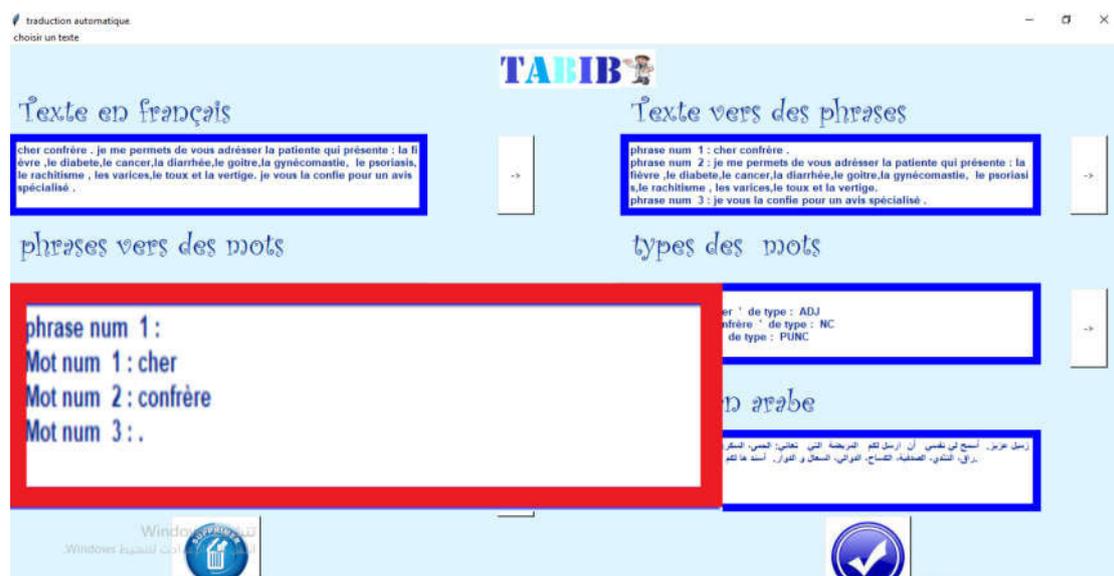


Figure (4.17):Fenêtre de la Traduction d'un texte dans L'opération d'analyse (cas de transfert des phrases vers des mots).

Quand on clique sur le bouton (3) de Figure (4.15), les mots vont apparaitre avec leurs types illustrés dans la Figure (4.18).



Figure (4.18):Fenêtre de la Traduction d'un texte dans L'opération d'analyse (cas de types des mots).

Quand on clique sur le bouton (4) de Figure (4.15), les types des phrases vont apparaitre, illustrés dans la Figure (4.19).

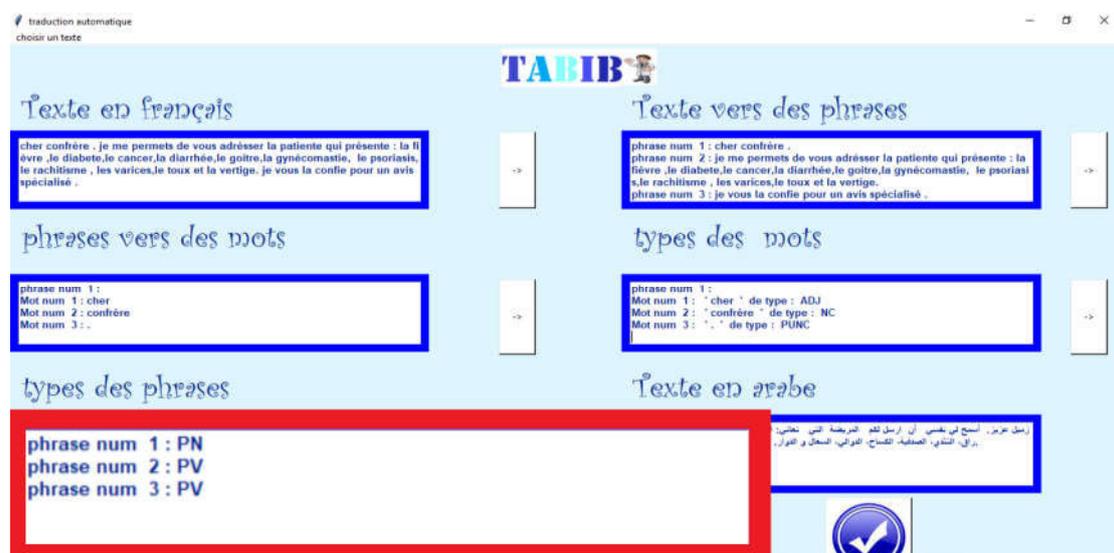


Figure (4.19):Fenêtre de la Traduction d'un texte dans L'opération d'analyse (cas de types des phrases).

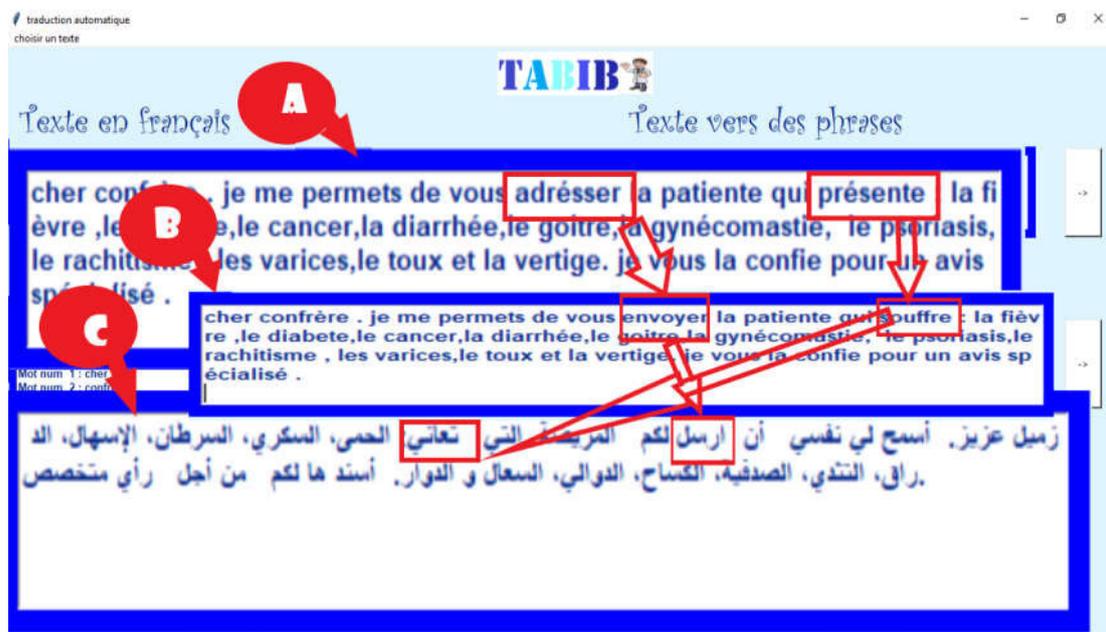


Figure (4.21):Fenêtre de la Traduction d'un texte dans L'opération de conversion (cas du synonyme (déduction)).

Dans la Figure (4.22), le changement de genre de nom de la phrase encadrée dans le champ (A) mentionné dans le champ (B), va induire un changement de la conjugaison du verbe et du pronom relatif de la phrase encadré de champ (A) mentionné dans les champs (C) et (D).

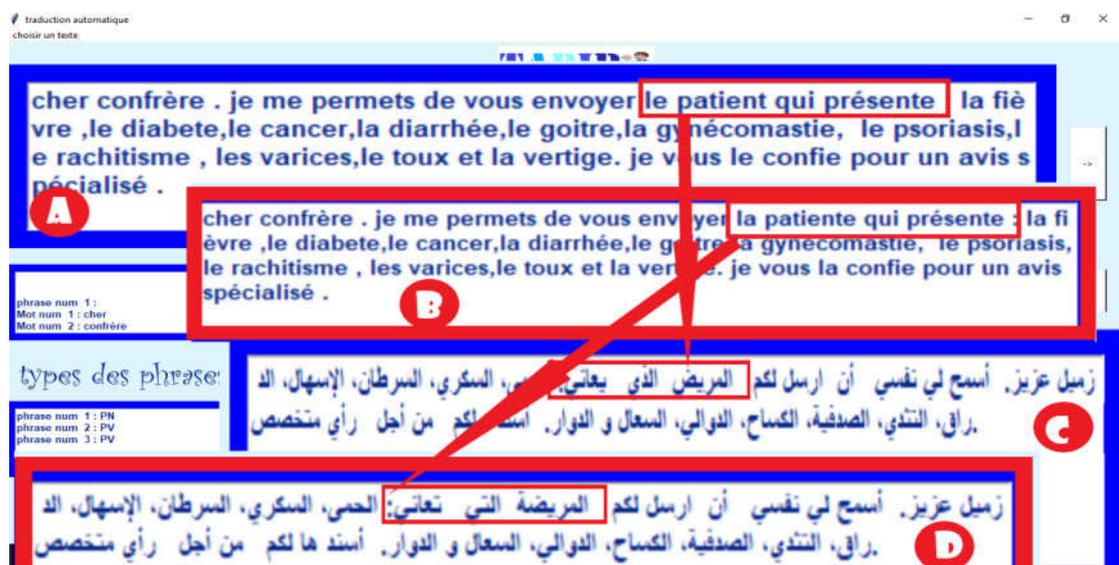


Figure (4.22):Fenêtre de la Traduction d'un texte dans L'opération de conversion (cas changement de genre d'un nom).

d. L'opération de génération :

Quand on clique sur le bouton (5) de Figure (4.15), l'opération de la génération démarre après l'opération de la conversion, illustrée dans la Figure (4.23).

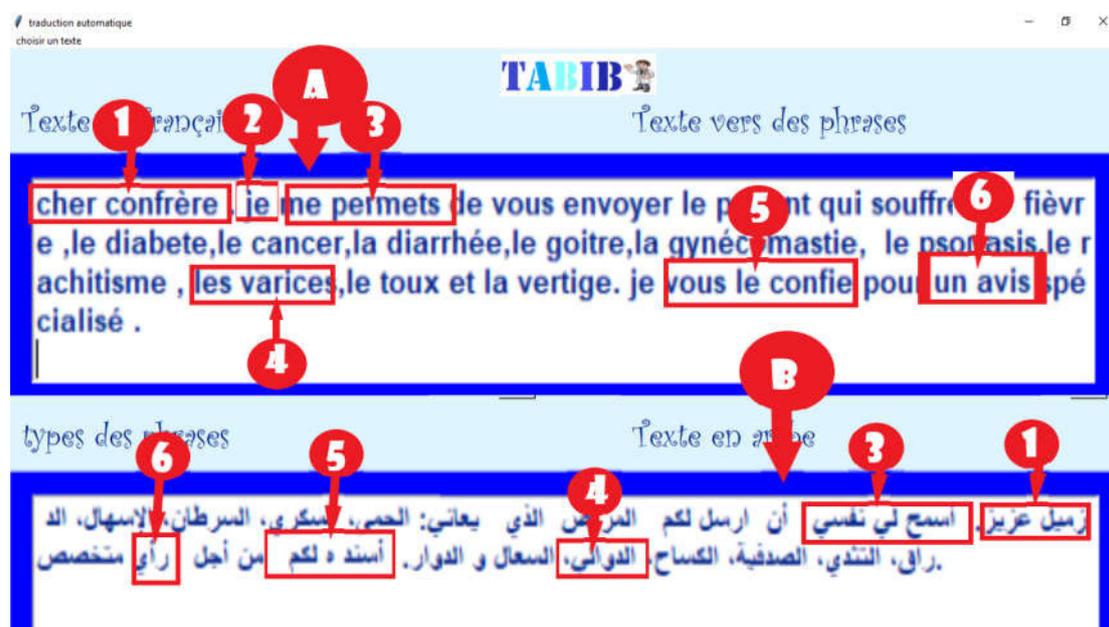


Figure (4.23): Fenêtre de la Traduction d'un texte dans L'opération de génération.

Dans la Figure (4.23), les cadres de champ (B) précisent les résultats de l'opération de génération déroulée dans les cadres de champ (A):

- les cadres (1) de deux champs illustrent le cas où l'adjectif précède le nom.
- le cadre (2) de champ (A) illustre le cas de l'élimination du pronom personnel.
- les cadres (3) et (5) de deux champs illustrent le cas où les pronoms relatifs précèdent le verbe.
- les cadres (4) de deux champs illustrent le cas où l'article défini précède le nom.
- les cadres (6) de deux champs illustrent le cas où l'article indéfini précède le nom.

4.5 Conclusion

Dans ce chapitre, nous avons présenté les langages et les outils de développement utilisés dans la validation et l'implémentation de notre système .nous avons utilisé plusieurs outils talque nlpnet ,tinkinter, xampp et python qui nous a aidédans l'implémentation de notre système .

Enfin, on a présenté la description du système (les interfaces), et les résultats obtenus.



Conclusion générale

Conclusion générale

Dans notre travail, on a créé un traducteur capable de traduire un texte (bilan médical et messages vers un docteur spécialiste) en la langue française à un texte (bilan médical et messages vers un docteur spécialiste) en la langue arabe dans un contexte spécifique (médical). On a choisi la méthode de traduction automatique à base des règles en ajoutant un dictionnaire médical pour se focaliser sur le contexte médical et ajouter la sémantique.

On a d'abord essayé d'analyser grammaticalement (Sélection et extraction des éléments grammaticale) le texte en langue française, ensuite on a converti et généré le texte en langue arabe en utilisant un dictionnaire médical afin de traiter le sens.

On a essayé de respecter les règles de la grammaire de la langue arabe, à fin d'obtenir un texte bien formaté.

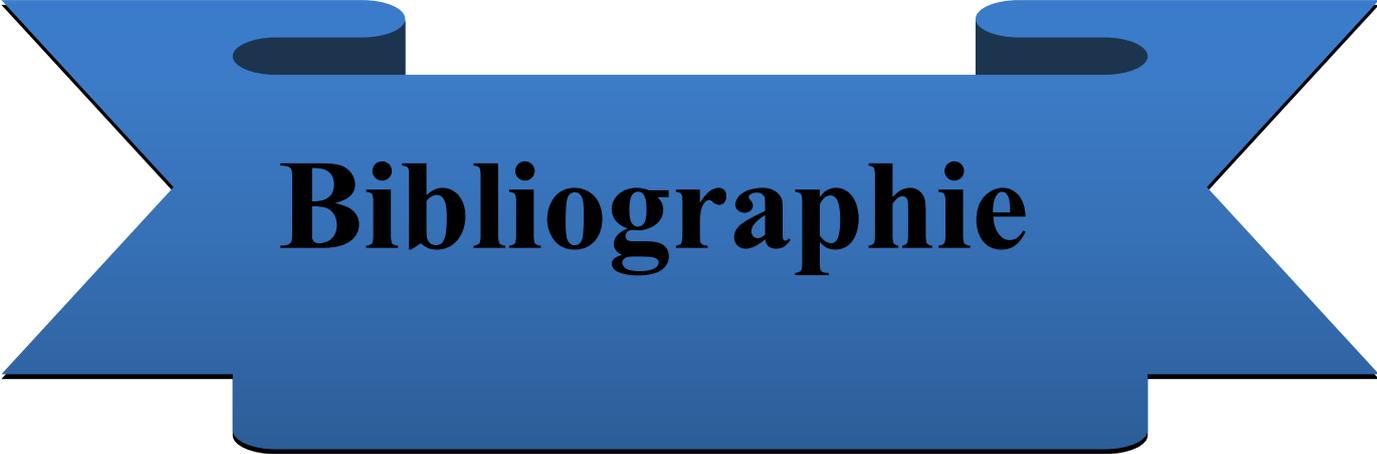
En premier lieu de ce projet, nous avons commencé par une présentation de l'état de l'art de traduction automatique et nous avons donné quelques travaux existants. ensuite nous avons expliqué brièvement la notion du contexte et de la grammaire. par la suite, nous avons entamé la conception de notre système. et enfin nous avons réalisé et implémenter notre système.

Dans notre système, nous pouvons dire que nous avons couvert des parties de la traduction automatique et non totalement, car il y a eu les lacunes suivantes:

- La grammaire est un domaine vaste.
- Certaine ambiguïté dans la recherche de synonyme de mot (cas de plusieurs synonymes).
- Problème d'outil d'analyse grammaticale de texte (nlp) (quelques résultats non fiables, prise de temps considérable).
- Absence d'analyse morphologique.
- le vocabulaire est un domaine vaste.

Nous avons des perspectives pour résoudre le problème de ces lacunes, qui sont :

- Fusionner quelques principes de l'approche statistique et l'approche à base des exemples avec l'approche utilisée.
- Utiliser de nouvelles techniques de traitement de texte tel que le deeplearning



Bibliographie

Bibliographie

[1] **Sébastien Haire**. Traitement Automatique des Langues et Apprentissage des Langues .Assisté par Ordinateur : bilan, résultats et perspectives . 2011

[2] Qu'est-ce que le traitement automatique des langues ?
http://www.technolangue.net/imprimer.php3?id_article=274 .15/2/2019.

[3] **Mojca Pecman**. Introduction à la linguistique.

https://www.eila.univ-paris-diderot.fr/_media/enseignement/lea/cours/linguistique/intro-linguistique/introductionlinguistiqueecon1.pdf. 9/3/2019.

[4] **Georges Mounin**. Introduction à la linguistique.2004

[5] **Natalie Kübler**. La traduction automatique : traduction machine?.2007

[6] Comment l'IA révolutionne la traduction ?
<https://www.lesechos.fr/2018/05/comment-lia-revolutionne-la-traduction-991256>. 15/2/2019.

[7] **Sadik Bessou**. Traduction Automatique à Base de Règles de l'Anglais vers l'Arabe.2016

[8] **Segun Afolabi**. La Traduction Automatique par Opposition à La Traduction Humaine Analyse d'un Corpus de Productions Réelles.2014.

[9] Traduction automatique : la révolution de l'intelligence.
<https://leblob.fr/enquetes/traduction-automatique-la-revolution-de-lintelligence-artificielle>. 1/3/2019.

[10] Personnalisation . <http://www.systran.fr/systran/technologie/methodologie-de-personnalisation>.11/4/2019.

[11] Résumé des éléments de grammaire française.
<http://projetconnaissance.free.fr/grammaire-fr.html>.1/3/2019.

- [12] **NedjmeddineKhalfallah**. Fiches grammaticales.2012
- [13] **AsmaiEmeur**. Introduction à Intelligence artificielle.2011.
- [14] **Françoise Picot**. Apprendre la grammaire au CP et au CE1.2012.
- [15] **KhawlaHasni** .Extraction des documents les plus significatifs d'un document textuel .2018.
- [16] **Patrick Fuchs et Pierre Poulain** .Cours de Python.2019.
- [17]Tutorialspoint.https://www.tutorialspoint.com/pycharm/pycharm_tutorial.pdf.6/4/2019.
- [18] Introduction XAMPP.http://www.monetfiore.ulg.ac.be/services/verif/cours/bd/repet2014/tp7_slides.pdf.
16/5/2019.
- [19] Python software fondation. <https://pypi.org/project/nlpnet>.16/5/2019.
- [20]Tkinter.<https://docs.python.org/2/library/tkinter.html>.16/5/2019.
- [21] **Yamina Kahla** .Découverte des services web sémantique .2014
- [22]Définition du contexte en informatique ubiquitaire.
[http://www-inf-it-sudparis.eu/cours/CSC5004/Archives/EnLigne/Cours/SensibiliteAuContexte/2.1.html](http://www-inf.it-sudparis.eu/cours/CSC5004/Archives/EnLigne/Cours/SensibiliteAuContexte/2.1.html)
. 10/3/2019.
- [23]**Tilgi**.Adaptation dynamique à l'environnement d'exécution.2006.
- [24]**Brown**.Context_aware application.1997.
- [25]**Gong**.Contextual modeling and application.2005.
- [26]**Winograd**.Architectures for context.2001.
- [27]**KhoulodBoukadi** .coopération interentreprises à la demande .2009