

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

**UNIVERSITE DES SCIENCES ET DE LA TECHNOLOGIE
HOUARI BOUMEDIENE**

INSTITUT D'INFORMATION

PROJET DE FIN D'ETUDES

**POUR L'OBTENTION DU DIPLOME D'INGENIEUR
D'ETAT EN INFORMATIQUE
OPTION : SOFTWARE**

**COMPREHENSION AUTOMATIQUE DE
LA LANGUE ARABE NON VOYELLE ,
APPLICATION A L'INTERROGATION
D'UNE BASE DE DONNEE**

Présenté par :

**Mr. HADJ-YAHIA Ouahid
Mr. MAREDJ Azze-Eddine**

Promoteur :

Mlle A. MOKHTARI

Jury:

**Président: Mr. AHMED NACER
Examineurs: Mlle BETOUCHE
Mr. MIHOUBI**

DEDICACES

Je dédie ce mémoire

A ma grand-mère pour laquelle les mots les plus doux ne peuvent exprimer ma gratitude .

A mes parents qui ont tout fait pour que je réussisse, j'exprime ma profonde gratitude .

A mes frères et soeurs .

A tous mes amis Ouahid, Ramadhane, Nacer eddine, Omar, Namick, Hakim, Djamel, Ghalia, Karima, Fatma-zohra pour leur soutien moral .

Azze Eddine

A ma grand-mère, pour les sacrifices qu'elle a fait .

A ma famille, pour ce qu'elle m'a donnée .

A Azze-eddine, Ramdhane. Hakim, Lotfi, Djamel, Mustapha, Abde el kadder, Nacer-eddine, Omar, Namick, Nadia, Karima, Fatma-zohra, Myriam .

A tous ceux qui me sont chers,

je dédie ce mémoire .

Ouahid.

REMERCIEMENTS

Nous tenons à remercier vivement mademoiselle Mokhtari et mademoiselle Madani pour nous avoir encadré, pour la confiance qu'elles nous ont témoignée, **pour** leur constante disponibilité et pour le suivi permanent qu'elles ont accordé à notre travail.

Nous remercions les membres de la sous-commission, madame Benchneb et mademoiselle Mokhtari pour les conseils et critiques judicieuses qu'elles ont apporté à ce travail.

Nos remerciements s'adressent également à monsieur Laribi **pour** ses conseils et pour l'intérêt qu'il a porté à notre travail.

Notre gratitude va aussi à Ramdhan pour le soutien moral et l'aide qu'il nous a apporté et à Saïd du centre de calcul pour ses conseils techniques et sa serviabilité.

Que tous ceux qui nous ont soutenu et encouragé au cours de ce projet trouvent ici l'assurance de nos remerciements sincères et amicaux.

RESUME

Dans le cadre d'une participation aux travaux d'une équipe de recherche de l'U.S.T.H.B dans le domaine de la compréhension automatique du langage naturel, il nous a été demandé de faire une étude sur le langage naturel Arabe non voyellé, puis en faire une application sur l'interrogation d'une base de données.

En premier lieu, notre travail a porté sur la compréhension d'une phrase écrite en langage naturel arabe non voyellé dans un cadre général. Par la suite, nous nous sommes intéressés à la conception et la réalisation d'un interpréteur de requêtes écrites en langage naturel arabe non voyellé, dans le langage LMD/LDD EAQUEL du SGBD EASY réalisé à l'institut d'informatique de l'U.S.T.H.B.

La requête écrite en langage naturel arabe non voyellé est présentée à un processus de traitement, lequel délivre en sortie la représentation de sens en langage EAQUEL.

Le logiciel est implémenté sur IBM PC XT avec le langage de programmation MS PASCAL.

Mots clés:

LANGAGE NATUREL ARABE NON VOYELLE, ANALYSE MORPHOLOGIQUE, ANALYSE LEXICALE, ANALYSE SYNTAXIQUE, ANALYSE SEMANTIQUE, ANALYSE PRAGMATIQUE, REPRESENTATION DU SENS, INTERPRETEUR, A.T.N, DICTIONNAIRE, SGBD EASY, EAQUEL.

SOMMAIRE

| | |
|--|----|
| INTRODUCTION | |
| PREMIERE PARTIE : COMPREHENSION AUTOMATIQUE ET LANGAGE NATUREL | |
| I/INTRODUCTION | 3 |
| II/EVOLUTION | 4 |
| 1/Première époque | 4 |
| 2/Deuxième époque | 4 |
| 3/Troisième époque | 5 |
| DEUXIEME PARTIE : ETUDE LITTERAIRE DE LA LANGUE ARABE | |
| I/INTRODUCTION ET DEFINITIONS | 8 |
| II/STRUCTURE GENERALE DE LA MORPHOLOGIE DE LA LANGUE ARABE . | 9 |
| 1/Morphologie du verbe..... | 9 |
| 1.1/Verbe fort et verbe faible | 9 |
| 1.2/Verbe transitif et verbe intransitif..... | 10 |
| 1.3/Les voix, les temps, le nombre, le genre et les personnes | 10 |
| a - Les personnes..... | 10 |
| b- Les genres | 10 |
| c - Les nombres..... | 10 |
| d- Les voix | 10 |
| e- Les temps et aspects | 10 |
| 1.4/La conjugaison | 11 |
| 2/Morphologie du nom | 11 |
| a- La détermination | 11 |
| b- Le genre | 11 |
| c- Le nombre | 12 |
| d- Le Masdar | 13 |
| e- Le nom d'agent et le nom de patient | 13 |
| f- Le nom d'une fois | 14 |
| g- Le nom de manière | 14 |
| h- Le nom de temps et de lieu | 14 |
| i- Le nom d'instrument | 14 |
| 3/Morphologie du pronom | 15 |
| 3.1/Le pronom personnel | 15 |
| 3.1.1/Le pronom personnel indépendant | 15 |

| | |
|---|----|
| 3.1.2/Le pronom personnel suffixe | 15 |
| a- L e pronom suffixe sujet..... | 15 |
| h- Le pronom suffixe complément du nom | 15 |
| c- L e pronom suffixe complément d u verbe | 16 |
| d- Le pronom suffixe complément de la particule | 16 |
| 3.2/Les démonstratifs | 16 |
| 3.3/Les noms conjoints | 17 |
| 3.4/Les pronoms interrogatifs | 18 |
| 4/Morphologie de l a particule..... | 18 |

| | |
|---|----|
| III/SYNTAXE DE LA LANGUE ARABE | 19 |
| 1/Proposition nominale | 19 |
| 1.1/Moubtada e t Khabar..... | 19 |
| 1.2/Les abrogatifs | 20 |
| a - كان e t ses analogues..... | 20 |
| b - إن e t ses analogues..... | 21 |
| 2/Proposition verbale | 21 |
| 2.1/Le verbe | 22 |
| 2.2/Le sujet actif | 22 |
| 2.3/Le sujet passif | 24 |
| 2.4/i-e complément d'objet | 24 |
| 2.5/ الحال | 24 |
| 2.6/ النعت | 25 |
| 2.7/ الضرف | 25 |
| CONCLUSION | 25 |

TROISIEME PARTIE :
ANALYSE D'UNE PHRASE EN LANGUE ARABE NON VOYELLE

| | |
|--|----|
| I/INTRODUCTION | 26 |
| II/ANALYSE MORPHO-LEXICALE | 26 |
| 1/Découpage d'une entité | 27 |
| 1.1/Recherche et classification des préfixes et des suffixes | 27 |
| a - Les préfixes..... | 27 |
| b.. Les suffixes | 28 |
| c- La table de compatibilité | 32 |
| 1.2/Organisation du dictionnaire | 32 |
| 1.3/Modèle d'analyse | 34 |
| III/ANALYSE SYNTAXIQUE | 36 |
| 1/Différentes techniques d'analyse | 36 |
| 1.1/Syntaxe formelle | 37 |
| a-- Grammaires régulières | 37 |
| b- Grammaires à contexte lib re | 37 |
| 1.2/Les réseaux de transition | 38 |

| | |
|---|-----------|
| 1.3/Réseaux A.T.N (Augmented Transition Network) | 39 |
| 2/Choix de l'outil d'analyse | 39 |
| 2.1/Analyse syntaxique par les A.T.N | 39 |
| a- A.T.N d'une phrase | 39 |
| b- A.T.N de la proposition nominale | 40 |
| c- A.T.N de la proposition verbale | 40 |
| d- Spécification des A.T.N utilisés | 40 |
| IV/ANALYSE SEMANTIQUE | 43 |
| 1/Notions à représenter | 43 |
| 2/Techniques de représentation du sens | 43 |
| a - Les réseaux sémantiques | 43 |
| b- Les grammaires de cas | 44 |
| c- Les dépendances conceptuelles | 44 |
| d- Les frames | 45 |
| 3/Outils d'analyse sémantique | 45 |
| a- Les grammaires sémantiques | 45 |
| b - Les A. T. N sémantiques | 46 |
| 4/Choix de l'outil d'analyse | 46 |
| a- Principe d'analyse | 47 |
| b - Analyse par les A.T.N sémantiques | 47 |
| V/LES CONNAISSANCES PRAGMATIQUES | 48 |
| 1/Analyse pragmatique | 48 |
| 2/Les inférences | 49 |
| 2.1/Définition | 49 |
| 2.2/Pilotage des inférences | 50 |
| VI/EXEMPLES RECAPITULATIFS | 51 |
| 1/Exemple avec ambiguïté levée | 51 |
| 2/Exemple avec ambiguïté | 52 |
| CONCLUSION | 54 |
| QUATRIEME PARTIE : | |
| APPLICATION A L'INTERROGATION D'UNE BASE DE DONNEES | |
| EN LANGAGE NATUREL ARABE NUN VOYELLE | |
| I/PRESENTATION DU PROBLEME | 55 |
| II/PRESENTATION DU MODELE ENTITES/ASSOCIATIONS | 56 |
| 1 /Le modèle entités/associations | 56 |
| 2/Notion de base du modèle | 56 |
| 3/Attributs, valeurs et ensembles de valeurs | 56 |
| III/LE LANGAGE EAQUEL | 57 |
| a - Les commandes du langage EAQUEL | 57 |
| b- Les agrégats | 58 |

| | |
|---|-----|
| IV/PRESENTATION DE LA BASE DE DONNEES CIBLE | 59 |
| 1/Les besoins à satisfaire p a r | |
| 1 a base de données RECTORAT..... | 59 |
| 2/Les ensemh les, les entités et leurs signification | 59 |
| V/REALISATION DE L'ANALYSEUR MORPHO-LEXICAL | 61 |
| 1/Prefixes et suffixes | 61 |
| a - Suffixes..... | 61 |
| b- Prefixes | 64 |
| c- Table de compatibilité | 64 |
| d- Representation des prefixes et suffixes en m&moire . . | 65 |
| 2/Le dirtiannaire | 67 |
| a- Recensement du lexique | 67 |
| b- Organisation du dictionnaire | 67 |
| c- Structure5 de données utilisées | 68 |
| d- Representation des synonymes | 69 |
| e- Representation des homographes | 69 |
| 3/L'analyse morpho-lexicale | 70 |
| a- Pretraitemerct | 70 |
| b- Traitement morphn-lexicale et preparation | |
| de l'analyse syntaxico-sémantique | 71 |
| c - Le caractère 'Y'..... | 7 2 |
| VI/REALISATION DE L'ANALYSEUR SYNTAXICO-SEMANTIQUE | 72 |
| 1/Catégorisation de5 mots du lexique | 72 |
| 2/L'interpréteur d'A.T.N | 74 |
| 2.1/Description des graphes | 74 |
| 2.2/Structure des graphes | 74 |
| 2.3/Réalisation d e l'interpréteur d'A.T.N..... | 76 |
| 3/L'analyse syntaxico-sémantique | 77 |
| VII/L'ANALYSE PRAGMATRIJF FT GENERATION DE LA REMJETE | 77 |
| CONCLUSION | 7 8 |