

Republique Algerienne Democratique Et Populaire

Universite des sciences et de la technologie
HOUARI BOUMEDIENE

Institut D'informatique

PROJET DE FIN D'ETUDE

Pour l'obtention du diplôme d'ingenieur d'etat en informatique

**ESSAI DE DIACRITISATION
D'UN TEXTE ECRIT EN LANGUE ARABE**

Promoteur :

M^r BENHAMA DI

JURY: Présidente :

Melle MADANI

Examineurs :

Mme BOUCHNEB

M^r BOUKRA

Centre de recherche sur l'information scientifique

et technique

C.E.R.I.S.T

N° 9/06,88

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

اِنَّا اَنْزَلْنَاهُ قُرْاٰنًا عَرَبِیًّا

صَدَقَ اللّٰهُ الْعَظِیْمُ

DEDICACES

A mes tres chers parents, a qui je doit tout.

A mes freres et soeurs qui m'ont constament encouragee,
particulierement a MOUNIR.

a ma tres chere SAMY,

avec toute mon affection.

a FATMA, NAIMA, WAHIBA

en signe de ma profonde amitie.

a tous ceux qui me sont chers,

Je dedie ce memoire.

BOUCHERA.

إلى أعظم أم : أمى الحنون،

إلى أعظم أب : أبى،

إلى كل من في قلبه لى ابتسامة

أهدي هذا العمل المتواضع.

حسينة .

-OO REMERCIEMENTS OO-

Nous remercions vivement Monsieur BENHAMADI, directeur du CERIST, pour nous avoir propose ce sujet qui nous a permis de faire une incursion fructueuse dans l'univers de l'intelligence artificielle.

Nous remercions egalement, Monsieur GHEBBACHE pour ses orientations objectives

Nous remercions les membres de la sous-commission, Madame DERDOURI et Madame MAMERI pour les conseils et critiques judicieuses qu'elles ont apporte a ce travail.

Nous tenons a exprimer gratitude et sympathie a Messieurs YOUSFI, BERROUK et SILMI ainsi qu'a Mesdemoiselles SENOUCI et SAIGHI

soit egalement remercie, le personnel de la salle machine, particulierement Monsieur EL-MAOUHAB. Ainsi que tout le personnel du CERIST notamment Mr. TOUDJINE pour son amabilite.

Sans oublier d'exprimer notre amitie aux anges de la bibliotheque Mademoiselle OUDJEHANI et Monsieur MEFTOUH.

SOMHAIRE

AVANT-PROPOS	
INTRODUCTION.....	1
PREMIERE PARTIE: ETUDE BIBLIOGRAPHIQUE	
CHAP I. QU'EST CE QUE LE LANGAGE NATUREL.....	4
CHAP II. LES SYSTEMES DE TRAITEMENT DU LANGAGE NATUREL.....	6
III-1 LA BASE DE CONNAISSANCES.....	8
a/ LES ENTITES.....	8
b/ LES EVENEMENTS.....	9
c/ LES SITUATIONS.....	9
III-2 L'ANALYSEUR.....	10
III-3 LE GENERATEUR.....	10
III-4 L'INTERPRETEUR.....	11
III-5 LA REPRESENTATION DE SENS.....	11
III-6 LE LEXIQUE.....	12
CHAP III SCHEMA CLASSIQUE D'ANALYSE DU LANGAGE NATUREL.....	13
II-1 LES CONNAISSANCES.....	16
a/ LES CONNAISSANCES SYNTAXIQUES.....	17
b/ LES CONNAISSANCES SEMANTIQUES.....	17
c/ LES CONNAISSANCES PRAGMATIQUES.....	18

- III-2 MODELES ET OUTILS.....18
 - a/ MODELES ET OUTILS SYNTAXIQUES.....18
 - a-1/ LES GRAMMAIRES GENERATIVES.....18
 - a-2/ LES RESEAUX DE TRANSITION.....24
 - a-3/ LES RESEAUX DE TRANSITION RECURSIFS...25
 - a-4/ LES RESEAUX DE TRANSITION AUGMENTES...27
 - a-5/ L'APPROCHE STOCHASTIQUE DANS
LA RECONNAISSANCE SYNTAXIQUE.....29
 - b/ LES OUTILS SEMANTIQUES.....31
 - b-1/ LES GRAMMAIRES DE CAS.....31
 - b-2/ LES DEPENDANCES CONCEPTUELLES.....33
 - b-3/ LES FRAMES..... 35

DEUXIEME PARTIE: **LA** LANGUE ARABE

- CHAP I. QUELQUES CARACTERISTIQUES DE LA LANGUE ARABE.....38
 - I-1 GENERALITES.....38
 - ? MORPHOLOGIE ET VOYELLATION 40
 - I-3 FLEXION ET CAS.....46
 - I-4 RAPPORT MORPHOLOGIE-SYNTAXE.....48
- CHAP II. IMPORTANCE DE LA DIACRITISATION DANS UN SYSTEME
LJNGUISTIQUE ARABE.....50

CHAP III. LES DIFFERENTES APPROCHES	53
III-1 VOYELLATION DES MOTS DU LEXIQUE.....	53
III-2 APPROCHE SYNTAXICO-SEMANTIQUE PURE.....	55
III-3 APPROCHE LINGUISTIQUE.....	58
a/ MORPHOLOGIE	58
b/ NIVEAU SYNTAXICO-SEMANTIQUE.....	59
c/ NIVEAU PRE-SYNTAXIQUE.....	60
c-1/ L'ETUDE MORPHOLOGIQUE	60
c-2/ L'ANALYSE CONTEXTUELLE.....	61
TROISIEME PARTIE: REALISATION	
I- PRESENTATION DU PROBLEME.....	63
11- METHODE CHOISIE.....	63
111- LES COMPOSANTS DU SYSTEME.....	65
III-1 LE LEXIQUE.....	65
.. LE LEXIQUE DU VOCABULAIRE DE LA LANGUE.....	66
b/ LE LEXIQUE DES BALANCES MORPHOLOGIQUES.....	70
111-2 LA TABLE INTERMEDIAIRE.....	71
111-3 LE MODULE D'ANALYSE MORPHOLOGIQUE.....	72
a/ PREMIERE METHODE.....	73
. EXEMPLE DE TRAITEMENT.....	74
. CRITIQUES DE LA METHODE.....	75
b/ DEUXIEME METHODE.....	76
111-4 LE MODULE D'ANALYSE CONTEXTUELLE.....	80
CONCLUSIONS	
ANNEXES	
BIBLIOGRAPHIE.	