التوليد الصوري للأسماء البسيطة في اللغة العربية .

مهديوي عمر ثانوية أبن زهر وزان . المغرب .

البريد الإلكتروني :mehdiouiomar @hotmail.com

الهاتف : 68471507 (212) او 37460336(212)

النـــاسوخ : 37460861(212)

ملخص:

يندرج هذا البحث في إطار ما أصبح يصطلح على تسميته اليوم ب " هندسة اللغة العربية " التي تتوخى فتح حوار علمي بين الحاسوب و تطبيقاته و الإنسان العربي بواسطة اللغة العربية بوصفها لغة طبيعية قابلة للمعالجة الآلية ، و هذا لن يتأتى إلا بتعميق النقاش الجاد و المثمر في القضايا الأساسية التي تطرحها المعالجة الحاسوبية للنظام اللغوي العربي في سائر مستوياته : الصرفية ، النحوية ، الدلالية ، و المعجمية ، و البلاغية من قبيل التوليد الآلي ، التحليل الآلي ، التكشيف الآلي للنصوص ، الاسترجاع الآلي للمعارف ، الترجمة الآلية ، الفهم الآلي للنصوص ، التعرف الآلي على النصوص ،التعرف البصري على الخصائص اللغوية العربية، وتوليف الكلام وتركيبه.

الكلمات المفاتيح :

التوليد الصوري –الأسماء البسيطة –هندسة اللغة العربية –المقاربة اللسانية الحاسوبية –النمو التأليفي –قاعدة البيانات.

Abstract

This research paper can be included among what is known nowadays as the Arabic language engineering .the latter aims at opening a scientific dialogue / discussion betwen the computer and it's applications and the arab man through the arabic language whichis a natural language liable to an automatic processing .this can be achieved only through a deep discussion of all the main issues hopics that are due to the computer processing of the arabic language system in all its levels : syntactic , grammatical , symantic , diction and chetoucal such as : automatic generation, automatic analysis , automatic index of the texts , automatic feed-back of knouwledge , automatic translation , automatic

understanding of the texts , automatic recaguition) .Of the literay and linguistic aspects / characteristion of the language and the discourse syntheis and and constuiction .

Key Words :

Generative format grammar – simple words – Arabic language engineering – computer linguistic approch – lexical grammar data-base .

Cette recherche s'inscrit dans le cadre de ce qui est convenu d'appeler actuellement l'architecture de la langue arabe qui se propose d'entamer un rapport dialogique savant entre, d'un côté l'ordinateur et ses application et de l'autre l'être humain arabe, grâce à la langue arabe en sa qualité de langue naturel susceptible d'être manipulée mécaniquement, et cela ne sera pas possible qu'avec l'approfondissement d'un débat sérieux et fructueux à propos des problèmes fondamentaux soumis au traitement attribué par l'ordinateur à l'espace linguistique arabe à tous les niveaux :la conjugaison, la grammaire, la lexicologie et la rhétorique

Mots clés :

la génération automatique, l'analyse automatique, le son linguistique et sa perception, la génération automatique de texte et leur compréhension et élaboration des dictionnaires électroniques

تصميم البحث : سنتطرق في هذه الدراسة اللسانية الحاسوبية إلى العناصر التالية:

1. مقدمة
2. الإطار النظري والمنهجي : مقاربة صورية (خوارزمية)
3. الأسماء البسيطة في اللغة العربية : التوليد من الجذر و الخوار يزم.
4. نتائج وخلا صات وملاحظات.

1- مقدمة تأطيرية عامة:

إن العمل في إطار "هندسة اللغة العربية"يتطلب الإلمام بنوعين من العلوم ،العلوم اللسانية بنوعيها القديمة والحديثة،والعلوم الحاسوبية التي تهتم بمعالجة اللغات الطبيعية. وإذا تأملنا مسار البحث اللساني الحاسوبي العربي، فإننا نجد الباحثين والمهتمين بالبرامج الحاسوبية قد غضوا الطرف عن المعرفة اللسانية أثناء المعالجة اللآلية للغات الطبيعية مكتفين بالمعرفة المعلوماتية الصرفة الشيء الذي دفع بالخبراء العرب، في السنوات الأخيرة إلى التفكير في تطوير وصياغة برمجيات عربية، أقل ما يقال عنها إنها في بداية الطريق ، وهي تبدو متخلفة عن المستوى العلمي الذي وصلت إليه برمجيات اللغات الأجنبية ، ولاشك أن هذا الوضع عطل الحوار بين الإنسان العربي بلسانه العربي والآلة الحاسوبية التي أصبحت تفرض نفسها علينا في جميع الميادين، وتشكل تحدي من تحديات القرن 21 بامتياز.

وهكذا فإن فتح باب التعاون والتنسيق على مصراعيه بين خبراء اللسانيات والحاسوبيات لم يعد مطلبا ثانويا بل ضرورة ملحة أكثر من أي وقت مضى قصد النهوض بمستوى البحث الهندسي اللغوي العربي وجعل لغة الضاد في مصاف اللغات العالمية المنفتحة على علوم العصر الحديثة.

2- الإطار النظري والمنهجي

إن المعالجة الآلية للغة العربية تقتضي بالضرورة الاعتماد على شرطين أساسيين وهما:

- الشرط الأول هو ضرورة الانطلاق من نظرية صورية(خوارزمية) صارمة وواضحة المعالم.
- الشرط الثاني هو ضرورة بناء قاعدة بيانات للمردات اللغوية العربية بنوعيها البسيطة والمركبة.

و فيما يتعلق بالشرط الأول فقد وقع اختيارنا النظري و المنهجي على نظرية النحو التأليفي باعتباره نظرية صورية (خوارزمية) تتمتع بمجموعة من المواصفات و المرتكزات الأساسية و المفهم الإجرائية التي تؤهله لأن يكون نظرية فرعية في اللسانيات العامة و اللسانيات التطبيقية . و نكتفي في هذا الصدد الإشارة إلى الأصول الإبستمية والمنهجية التي يقوم عليها النحو التأليفي من قبيل أسلوب تراكم المعطيات و الملاحظة الوصفية و التفسيرية و مبدأ الوصف

التصنيفي الذي أثبت نجاعته في ميدان البيولوجيا النباتية و الحيوانية و علم الفلك و علم المنطق و الرضيات الجبرية بالإضافة إلى مبدأي التحليل و التوليد الآليين .

أما بالنسبة للشرط الثاني فيمكن القول إنه الهدف الأساس و القاعدة الصلبة التي يسعى إليها مشروعنا الكبير و المتمثل في بناء قواعد بيانات للأسماء البسيطة في اللغة العربية موجهة لخدمة المعاجم الآلية للغة العربية. و غنى عن البيان أن إنجاز عمل من هذا الصنف ليس بالأمر بالهين ،بل هو مغامرة علمية محفوفة بالمخاطر و الصعوبات منها ما يرتبط بالجانب التقني ، و منها ما يرتبط بالجانب النظري المرتبط أساسا بالخصائص الصرفية و التصريفية للأسماء البسيطة في اللغة العربية . و على أية حال ، فقد سبق لنا أن أنجز أنا الجزء الأول من هذا المشروع الضخم في بحث لنيل دوكتوراة السلك الثالث في اللسانيات الحاسوبية العربية في موضوع توليد الأسماء البسيطة من الجذور الثلاثية المعتلة : مقاربة لسانية حاسوبية تحت إشراف الأستاذ الدكتور : **محمد الحناش** بكلية الآداب و العلوم الإنسانية و ظهر المهراز فاس المغرب . و نحن بصدد تهيئ الجزء الثاني لنيل الدكتوراه في موضوع : توليد الأسماء البسيطة من الجذور الثلاثية الصحيحة في اللغة العربية : مقاربة لسانية حاسوبية تحت إشراف الأستاذ **الدكتور عبد الغني أبو العزم** بكلية الآداب و العلوم الإنسانية ، عين الشق الدار البيضاء ، المغرب

3- الأسماء البسيطة في اللغة العربية : التوليد الصوري : التوليد من الجذر و الخوارزم .

3-1 : المفردة في اللسانيات الحاسوبية :

تعرف المفردة من وجهة نظر حاسوبية بأنها متوالية صوتية تتضمنها الأبجدية العربية محاطة ببياض يفصل ما بينها و ما قبلها و ما بعدها . و ينطبق هذا التعريف على جميع المفردات اللغوية العربية بما فيها الأسماء و الأفعال و الحروف و الصفات و المصادر إلخ ، و ويخرج منه المتواليات التي تعرضت لتغيرات تركيبية و إضمارية مثل : أتتذكروننا و بمدرستهم ، كما يخرج من هذا التعريف أيضا الأسماء المركبة المسكوك منها و العادي مثل :

حيص بيص – أخماسا في أسداس – بصيص أمل ..إلخ .

فهذه المتواليات تشكل الأرضية الأساسية لقاعدة بيانات الأسماء البسيطة في اللغة العربية الموجهة لبناء المعاجم الآلية للأسماء البسيطة والذي يجيب عن أسئلة ترتبط بالمورفولوجيا ، و الفونولوجيا .

2-3 : **الأسماء البسيطة في اللغة العربية :**

يقصد بالأسماء المتواليات الصوتية بنوعيها الجامد و المشتق و يندرج ضمن الأسماء البسيطة الجامدة أسماء الأعلام – و الأسماء المشتركة و أسماء الأمكنة و غيرها فيما نعني بالأسماء البسيطة المشتقة المشتقات بما فيها المصدر بأنواعه و اسم الفاعل و الصفة ، و أسماء الزمن و المكان و أسماء الآلة ، و اسم التفضيل ، و اسم المبالغة و اسم المفعول .

3-3 التوليد الصوري : التوليد من الجذر و الخوارزم .

من أبرز الخصائص التي تطوع اللغة العربية للمعالجة الآلية : الطبيعة الصرفية الاشتقاقية ، ذات الإنتاجية العالمية ، و المستمدة أساسا من المكونين الرياضيين : الجذر و الوزن و من شأن هذا أن يدحض بعض الدعوات التي ترى أن اللغة العربية لغة معقدة و صعبة على مستوى المعالجة الآلية ، بناء على ذلك يصح لنا التساؤل : كيف يتأتى للقدرة الإنسانية أن تستنبط الكلمات (الأسماء البسيطة نموذجا) من الجذور اللغوية ثم ما هي العناصر التي تتدخل لتجعل من الجذر المعتل (و.ل.د) اسم المفعول (م 0، و 2 (مولود) ؟ .

للجواب عن هذا الإشكال ننطلق من فرضية أساسية و منهجية مفادها أن الجذر بوصفه علامة صوتية مجردة هو مصدر التوليد و مادة تنظيمية لمداخل المعجم . و لهذا فالعلاقة بين الجذر الصوتي و المدخل المعجمي البسيط (الأسماء البسيطة نموذجا) لا تعد و أن تكون موجهة تتم وفق روزنامة من القواعد الصرفية و التصريفية هي المعبر عنها عادة بخوارزميات الزيادة . يتجلى دور الجذر أو العلامة الصوتية المجردة في صنع الأساس و البنية الأصلية للمفردة في اللغة العربية بينما يقوم الوزن الصرفي بتحديد هيكلها العام و ذلك بتوزيع الحركات (الصوائت أو المصوتات) على مختلف حروف الكلمة ، كما يقوم بتوزيع خوارزميات الزيادة التي تخل على مكونات الجذر و التي تنتج في الأخير الكلمة المولدة ، و هذه المورفيمات الخوارزمية هي في الأساس عبارة عن سوابق أو أ واسط أو لواحق . هذا التشكيل الرياضي و المنطقي للغة العربية جعلها من لغة انصهارية في مقابل اللغات الإلصاقية الأخرى .

أما الحركات فتخضع لتوزيع صارم و مقنن تضبطه خوارزميات التطابق بين الوزن و الكلمة المولدة (المنتجة) . و المراد بالحركات تلك التي يتولى المحلل الصرفي الطبيعي بتوزيعها على الحرفيين الأول و الثاني من الكلمة في حالة الجذر الثلاثي ، و الحروف الثلاثة الأولى في حالة الجذر الرباعي ، أما حركة اللام فيستند توزيعها إلى المحلل النحوي الذي يتعامل مع اللغة بوصفها منظومة من الوظائف الصورية التي لها نظام خاص بها .

و على هذا الأساس ، فاللغة العربية لغة رياضية في الأساس ، و مكونة من أجرومية من الخوارزمات الصورية ، دخلها الجذور الصوتية مرورا بالأوزان الصرفية التي تتمتع بقوة الأصهار المورفيني المبرمج ، و خرجها الكلمات و الجمل و هي كالتالي :

الجذر (ف ع ل) ز ،ك ل،ح

كلمة = ——————————————

الوزن /الميزان

يشكل الجذر دخل البرنامج اللغوي المخزن في الكفاية اللغوية للمتكلم ، في البداية يتم اختيار الجذر المراد تشغيله (جذر ثلثي أو رباعي) ثم تشرع الكفاية اللغوية في تطبيق خوارزميات المطابقة بين الجذر اللغوي والمادة الصورية(ف ع ل) وفي مرحلة لاحقة يتم تفعيل الوزن عن طريق خوارزميات الإقحام والإدماج التي تقوم بعملية إدراج مورفيمات الزيادة (السوابق، الأواسط، اللواحق) والحركات في البنية النظرية للجذر، بهدف توليد الكلمة، أما في مرحلة التحليل فاننا نطبق نوعا أخر من الخوارزميات ، وذلك بتفكيك الكلمة ألي عناصرها الأساسية وما تتألف منه .ولكل صيغة صرفية مقابل معنوي أي دلالي مخزن في الكفاية اللغوية، وهذا المكون الدلالي هو الذي يضمن ربط المستوى الصرفي بالمعجم والدلالة. أما علاقة الصرف بالمستوى الفونولوجي فقائمة بكل تأكيد ولا تستدعي التعليل سيما وأن الأصول المعتلة تشكل مادة خصبة للمتغيرات الصواتية (لمزيد من التوضيح راجع رسالتنا الجامعية لنيل دكتوراه السلك الثلث في موضوع توليد الأسماء البسيطة من الجذور الثلاثية المعتلة : مقاربة لسانية حاسوبية ،1999 تحت إشراف الأستاذ الدكتور محمد الحناش كلية الاداب والعلوم الإنسانية فاس)

وفيما يلي تصور عام لطريقة بناء القواعد التوليدية (التوليد من الجذر والخوار زم):

في الإطار الصوري اللسانيات الحاسوبية، يرتكز التوليد من الجذر والخوار زم على متوالية من القواعد العقلانية(Rationnels) وعلى هذا الأساس ، فتوليد الأسماء من جذورها المعتلة يخضع لقانون خوارزمي، تؤطره المعادلة الصورية التالية :

$$R(X) = [Alg(y)\ El(z)]$$

حيث أن : R (X) = الجذر المعتل ثلاثيا كان أم رباعيا

Alg (y) = : خوارزم الزيادة سواء كانت سابقة أو واسطة أو لاحقة ، وهي مسننة رقميا من 0 إلى 3.

EL (Z) = المدخل المعجمي المولد من جذره المعتل.

أ – خورزميات توليد أسم الفاعل : يشتق إسم الفاعل (سفا) من الجذر المعتل وفق القاعدة الصورية (ق 1) التالية

ج (ن)←——— [خو (سفا)(ن)]

ج (ن) = : الجذر المعتل (3 أو 4) .

خو = : خوارزم الزيادة التي تضاف إلى الجذر

سفا (ن) = : مقولة إسم الفاعل المولدة .

تنطبق (ق1) على جميع الجذور في اللغة العربية المعتل منها و الصحيح ماعدا بعض الحالات الإستثنائية القليلة العدد التي أشرنا إليها في قاعدة البيانات (أنظر الجداول من 1 إلى 8) ، و قد أشرنا إلى هذه الحالات الخاصة بالفراغات أو البياضات) . و للإشارة فإن هذه الحالات الخاصة إما أن ليس لها إسم فاعل البتة، و إما أن إسم الفاعل منها يشتق على صورة مغايرة تماما ل(ق1).

و ع د ← [+ ا1 (واعد)]

ي ن ع ← [+ ا1 (يانع)]

ق و ل ← [+ ا1 (قائل)]

ب ي ع ← [+ ا1 (بائع)]

د ع و ← [+ ا1 (داع)]

ر م ي ← [+ ا1 (رام)]

هـ و ي ← [+ ا1 (هاو)]

و ح ي ← [+ ا1 (واح)]

ج و ر ب ← [+ م0 ، و 1 (مجورب)]

س ل ق ي ← [+ م0 (مسلق)] . إلى اخره

ب- خوارزميات توليد إسم المفعول (سفع) :

يشتق إسم المفعول من الجذر المعتل وفق القاعدة الصورية (ق2) التالية :

ج(ن) ← [خو (سفع(ن)]

ج(ن) = : الجذر المعتل (3أو 4)

خو = :خوارزم الزيادة السابقة أو الواسطة أو اللاحقة التي تدخل على الجذر.

سفع (ن) =:مقولة إسم المفعول المولد من الجذر.

تصلح هذه القاعدة (ق2) للتطبيق على جميع الأسماء المولدة من جذورها المعتلة سواء كانت ثلاثية أو رباعية .ولتوضيح ما قلناه نورد بعض الأمثلة التطبيقية :

و ع د ← [+ م0 ، و2(موعود)]

ي س ر ← [+ م 0 ، و 2 (ميسور)]

ق و ل ← [+ م 0 (مقول)]

ب ي ع ← [+ م 0(مبيع)]

د ع و ← [+ م 0 ،و 2 (مدعو)]

ر م ي ← [+ م 0 ، ي 2 (مرمي)]

ك و ي ← [+ م 0 ،ي 2 (مكوي)]

و ح ي ←——— [+ م 0 (مو حي)]

وللإشارة فالقاعدة (ق2) تطبق على جميع الجذور بحسب المقولة المولدة، وذلك واضح من خلال التغييرات المورفولوجية والفونولوجية التي تطرأ على بنية الجذر.

ج- خوارزمات توليد المصادر (الأصلي، الميمي، العددي، النوعي، والصناعي):

تشتق المصادر من الجذور المعتلة (3 أو 4) وفق القاعدة الصورية التالية

(ق3) :ج(ن) ←——— [خو (مص(ن)]

ج(ن)= : الجذر المعتل (3أو4)،و(ن) متغير .

خو = : خوارزم الزيادة المصدرية

مص = : مقولة المصدر (مص1،مص2، مص3، مص4، مص5)

مثال :

و ع د ←——— [+ ى (وعد)] ،[+ م0(موعد)] ، [+ة3 (عدة)] ..الخ

ي ب س ←——— [+ ى (يبس)] ، [و 2، ة3(يبوسة)] ،[+ت0، ي2(تيبيس)]

ت و ق ←——— [+ ى (توق)] ،[+و2 (تووق)] ،[+ ا2، ة3 (تياقة)]...الخ

ب ي ت ←——— [+ ى (بيت)] ، [+ ا 2 (بيات)]،[+ت0 ، ي 2(تبييت)]....

ج ل و ←——— [+ ى (جلو)] ،[+ ا2 (جلاء)]

س ب ي ←——— [+ى (سبي)]....

د- خوارزميات توليد الصفة (ص) :

تشتق الصفة من الجذر المعتل وفق القاعدة الصورية التالية:

ج(ن) ←———[خو (ص(ن)]

ج(ن)= : الجذر المعتل

خو= : خوار زم الزيادة الوصفية.

ص(ن) = : مقولة الصفة المولدة من جذرها.

مثال :

ي ق ظ ←——— [+ ا3، ن3(يقظان)]

ط و ل ←——— [+ ي 2(طويل)]

ه – خوارزميات توليد المبالغة (سغ).

تشتق المبالغة من الجذر في اللغة العربية وفق القاعدة الصورية التالية (ق5) :

ج(ن) ←———[خو(سغ(ن)]

ج(ن) = : الجذر المعتل

خو = خوار زم الزيادة السابقة أو اللاحقة أو الواسطة التي تدخل على الجذر

سغ (ن) =: مقولة اسم المبالغة المولدة من جذرها

مثال:

و ب ص ⟵ [+ب1، ا2(وباص)]

و ل د ⟵ [+و2 (ولود)]

غ و ر ⟵ [+م 0، ا 2(مغوار)]

غ ي ر ⟵ [+ و 2 (غيور)]

ف ت ي ⟵ [+ ي 2(فتي)]

ن س ي ⟵ [+ س 1، ا2(نساء)]

و- خوارزميات توليد أسماء الزمان والمكان:

تشتق أسماء الزمان والمكان وفق القاعدة الصورية التالية(ق6):

ج(ن) ⟵ [خو(سز/سم(ن)]

ج(ن)=: الجذر الثلاثي أو الرباعي

خو =: خوارزم الزيادة المضافة إلى الجذر؟

سز=: اسم الزمان

سم =: اسم المكان.

أمثلة :

و ع د ⟵ [+م0(موعد)]

و ق ع ⟵ [+م0 (موقع)]

ن ف ي ⟵ [+م0 (منفى)]

غ ر ب ⟵ [+م0 (مغرب)]

ي – خوارزميات توليد أسماء الآلة من الجذور:

يشتق اسم الآلة من الجذر وفق القاعدة الصورية التالية (ق7):

ج (ن) ⟵ [خو (سأ(ن)]

ج(ن) =: الجذر

خو =: خوارزم الزيادة التصريفية

سأ(ن) =: اسم الآلة.

د ر ي ⟵ [+م0 (مدري)]،[+م0، ة3(مدراة)]

ص ف و ⟵ [+م0 ،ة3 (مصفاة)]

ق و د ⟵ [+م0 (مقود)]

ن – خوارزميات توليد اسم التفضيل (سيف):

يشتق إسم التفضيل من الجذر وفق القاعدة الصورية التالية(ق8) :

ج(ن) ⟵ [خو (سف(ن)[

ج(ن) = : الجذر

خو :خوارزم الزيادة التصريفية

سف (ن) =: أسم التفضيل.

ط و ل ←——— [+ أ0 (أطول)]

خ ي ر ←——— [+ ø (خير)]

ه – خوارزميات توليد الأسماء الجامدة (سج):

تشتق الأسماء الجامدة من الجذور المعتلة وفق القاعدة الصورية التالية :

ج(ن) ←——— [خو (سج(ن)]

ج(ن) =: الجذر المعتل

خو =: خوارزميات الزيادة

سج (ن) =: الاسم الجامد

4- نتائج خلا صات وملاحظات :

نسجل من خلال المعطيات الواردة في قاعدة البيانات مايلي :

- تتميز اللغة العربية بقوة إنتاجية صرفية عالية،مما أهلها لأن تكون لغة توليدية بالأساس، لا تباريها في ذلك أية لغة أخرى، ذلك أن الجذر اللغوي الواحد، نولد منه سلسلة من المداخل المعجمية المفردة وفق منظومة من القواعد التصريفية والصرفية(خوارزميات الزيادة).
- يعتبر التوليد من الجذر والخوارزم وسيلة إجرائية لبناء المفردة البسيطة (س، ف، ض، مض،....)باعتبارها مدخلا معجميا للمعجم الآلي، الذي يمكن أن يجيب عن الأسئلة المتعلقة بقضايا الفنولوجيا والمورفولوجيا.
- يمثل الجذر البسيط مدخلا معجميا نسبيا، ومادة نظرية ينطلق منها التوليد الصرفي الالي.
- تنقسم الجذور في اللغة العربية إلى : جذور بسيطة وأخرى مركبة، مطردة وأحرى نظرية، جذور صحيحة، وأخرى معتلة. فالأولى(أي الجذور البسيطة المطردة والمحققة) تتميز بإنتاجية صرفية عالية، بينما الثانية(الجذور النظرية) تتميز بإنتاجية صرفية ضعيفة وعميقة.
- إن الطريقة التي صنفت بها المعاجم العربية جذورها تستدعي المراجعة، قصد إعادة هيكلة بنية المعجم العربي من جديد، وترتيب معماريته وفق المعطيات النظرية المستمدة من اللسانيات الحاسوبية في أفق بناء معاجم آلية للمفردات اللغوية العربية البسيط منها والمركب، على غرار اللغات الأخرى كالفرنسية والإنجليزية.

- تحتل الأسماء البسيطة المشتقة المرتبة الأولى من حيث مجموع النظام اللغوي العربي، بينما تأتي المفردات الأخرى في المرتبة الثانية، مما يدحض الفكرة التقليدية التي تزعم أن الكلام العربي كله مشتق أو جامد، في حين يتضمن النظام اللغوي العربي الصنفين معا.
- ينقسم التوليد (Génération) إلى توليد خارج لغوي (Extra-Linguistique) أي توليد جذر من آخر كتوليد الرباعي من الثلاثي أو الثلاثي من الرباعي وقد أشار القدماء العرب إلى بعض الأدوات الإجرائية التي يتم بواسطتها هذا النوع من التوليد من قبيل القلب، الإبدال، التوسع،.....وتوليد لغوي(Linguistique)

كتوليد المداخل المعجمية من جذورها وفق قواعد مضبوطة، وهذا الصنف هو الذي أسميناه ب"التوليد من الجذر والخوارزم"

- أن ثنائية الجمود والاشتقاق ثنائية غامضة، تحتاج إلى إعادة الصياغة، لأن تطبيقها على المداخل المعجمية من شأنه أن يثير الشكوك في مدى صحة القواعد الموضوعية للفصل بين الجوامد والمشتقات في اللغة العربية. لهذا فأن المعيار الحاسم للتمييز بين الاسم الجامد والاسم المشتق معيار مورفو-تركيبي(Morpho-syntaxique) بالأساس.
- إن قاعدة البيانات تبقى مفتوحة لاستقبال أي مدخل معجمي جديد، مما يعني أنها(قاعدة البيانات) صورة عاكسة لكفاية المستعمل العربي.

عالجت هذه الدراسة اللسانية الحاسوبية مستوى هاما من النظام الصرفي العربي أي توليد الأسماء البسيطة من الجذور المعتلة والخوارزميات(الزوائد التي تدخل على بنية الجذر، سواء أكانت سوابق أو أواسط أو لواحق).

وسعينا انطلاقا من الإطار النظري التأليفي بناء قاعدة توليدية للأسماء في اللغة العربية موجهة بالأساس لخدمة المعجم الآلي للمفردات البسيطة(D.E.A.M.S) وتفاديا للعراقيل والصعوبات التي من المتوقع أن يطرحها الإعراب الصرفي (الجمع، التذكير، التأنيث، النسبة والتصغير) بالنسبة للأسماء لم ندخل هذا المستوى في مجال اشتغالنا لحد الساعة، نظرا للمشاكل التي تطرحها الخصائص التصريفية بالنسبة للبرنامج المعد للمعالجة الآلية، وآثرنا ترك هذا الموضوع جانبا إلى بحوث أخرى إن شاء الله.

وهكذا، فقد واجهتنا صعوبات جمة طوال المدة التي قضيناها في إنجاز هذا البحث المتواضع، منها ما يتعلق بالمراجع، ذلك أن أغلب المراجع الأساسية توجد باللغات الأجنبية ولم نتمكن من الحصول إلا على البعض منها، و منها ما يتعلق بالمعالجة، ذلك أن موضوع الدراسة في جوهره تقني أكثر مما هو لساني مما يستدعي بالضرورة توافر تقنيين معلوماتيين، وهذا ما لم يتحقق، فقد كنا نتردد على مختبر المعلوميات والعلاج الآلي للغة العربية بالمدرسة المحمدية

للمهندسين بالرباط ، بين الفينة والأخرى، لكن هذا التردد لم يكن يحل لنا المشاكل التي استعصى علينا فهمها.

وفيما يلي شرح وتوضيح للطريقة التي اتبعنا في بناء قاعدة البيانات :

أولا : إن قاعدة البيانات عبارة عن أجرومية لنظام السوابق واللواحق والأواسط التي تدخل على بنية الكلمة الاسمية المولدة من جذرها المعتل.

ثانيا : لقد وضعنا ترقيما عدديا للزوائد الأصلية وغير الأصلية (0، 1، 2، 3) بحيث أعطينا رقم (0) للسوابق التي تكون في بداية الاسم البسيط، ورقمي (1،2) للأواسط التي تكون بين حروف الاسم ورقم (3) للواحق التي تكون في نهاية المقولة المولدة.

ثالثا : قسمنا قاعدة البيانات إلى جداول سبعة، بحسب المعلومات التي يتضمنها كل جدول على حدة، وبحسب نظام المجموعات الرياضية. في الجدول الأول وضعنا الجذور المعتلة مرتبة ومصنفة بطريقة تراتبية من المعتل بالفاء، والمعتل بالعين، إلى المعتل اللام). وفي الجدول الثاني وضعنا مجموعة من الأسماء الجامدة، وفي الجدول الثلث وضعنا مجموعة اسم الفاعل المولد من الثلاثي المعتل ومزيده، وفي الجدول الرابع، وضعنا مجموعة اسم المفعول، وفي الجدول الخامس وضعنا مجموعة المصادر(مرتبة في أرقام معينة)، وفي الجدول الخامس وضعنا مجموعة المصادر (مرتبة في أرقام معينة)، وفي الجدول السادس وضعنا مجموعة المختلفات وهي تضم الصفة، المبالغة، الزمان، المكان، الآلة والتفضيل. بالإضافة إلى القواعد الصوتية(الإعلال، الإبدال والقلب) التي تطرأ على بعض الخروف. وفي الجدول السابع والأخير، وضعنا متن المدخل المعجمي مشروحا بشكل موجز ومختصر.

رابعا: إن تصنيفنا للجداول التوليدية لم يكن قائما على الإعتباط بل وضعنا كل مجموعة اسمية في جدولها المناسب انطلاقا من صنف الجذور(المثال الواوي، المثال اليائي، الأجوف الواوي، الناقص الواوي، الناقص اليائي، اللفيف المفروق و اللفيف المقرون).

المراجع المعتمدة بالعربية باللغات الأجنبية

- **الحناش محمد** : اللغة العربية والحاسوب : قراءة سريعة في الهندسة اللسانية، المقال منشور بموقع اللسانيات العربية .

- **غازي عز الدين** : نظام قواعد المعرفة صرافي صوتي للغة العربية، رسالة جامعية لنيل دبلوم الدراسات العليا في اللسانيات الحاسوبية العربية، كلية الاداب والعلوم الإنسانية ظهر المهراز فاس 1999.

- **مهديوي عمر** : توليد الأسماء من الجذور الثلاثية المعتلة : مقاربة لسانية حاسوبية، رسالة لنيل دكتوراة السلك الثالث في اللسانيات الحاسوبية العربية، كلية الآداب والعلوم الإنسانية ظهر المهراز فاس 1999.

: المعالجة الآلية للغة العربية: معجم الجذور نموذجا، بحث لنيل دبلوم الدراسات العليا المعمقة كلية الآداب والعلوم الإنسانية ظهر المهراز 1994.

- استخدام اللغة العربية في المعلوماتية، المنظمة العربية للتربية والثقافة والعلوم تونس 1996.

- Dictionnaires électroniques et analyse automatique de texte, Max silberztein, Masson ,1993.

- Les dictionnaires électroniques, revue de l'association pour le traitement automatique des langues (ATALA), vol 44, N° 2/2003.