الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

جامعة أبي بكر بلقايد
تلمسان- الجزائر

كلية الآداب واللّغات
قسم اللغة العربية و آدابها

Université Abou Bekr Belkaid TLEMCEN
Faculté des lettres et des langues
Département de la Langue et de la Littérature Arabe

أطـروحة مـقدمـة لنيل درجـة الدكتـوراه تخصص:
علـم اللّـغة الحـديث - الموسـومة:

المعجم الإلكتروني للمعالجة الآلية للغة العربية

"الـواقع والآفـاق"

مقاربـة لسـانيـة حـاسـوبيـة

إعـداد الطـالب:
أمـين قـدراوي

إشـراف:
أ. د. سيدي محمد غيثـري

لجنـة المنـاقـشة:

أ. د عبد القـادر سـلامـي	رئـيـسـا	أستاذ التعليم العالي جامعة تلمسان
أ. د سيـدي محمد غيثري	مشرفا ومقررا	أستاذ التعليم العالي جامعة تلمسان
أ. د عـبد المجـيد سـالمي	عـضـوا	أستاذ التعليم العالي جامعة الجزائر2
أ. د طـاهر لـوصـيف	عـضـوا	أستاذ التعليم العالي المدرسة العليا للأساتذة (الأغواط)
د. محمد الأمين عبد الرحيم	عـضـوا	أستاذ محاضر (أ) جامعة تلمسان
د. عـزّ الـدّين حفّـار	عـضـوا	أستاذ محاضر (أ) المركز الجامعي عين تموشنت

السنـة الجـامعية : 1437-1438هـ/2016-2017 م

أطـروحة مـقدمـة لنيل درجـة الدكتـوراه تخصص:
علـم اللّـغة الحـديث - الموسـومة:

المعجم الإلكتروني للمعالجة الآلية للغة العربية
"الـواقع والآفـاق"
مقاربـة لسـانيـة حـاسـوبيـة

إعـداد الطالب
أميـن قـــدراوي
بإشـراف
الأستـاذ الدكتور سيدي محمد غيثـري

1437-1438هـ/2016-2017 م

بسم الله الرحمن الرحيم

أعوذُ بِاللّهِ منَ الشّيطانِ الرَّجيمِ

﴿ قَالُوا سُبْحَانَكَ لَا عِلْمَ لَنَا إِلَّا مَا عَلَّمْتَنَا إِنَّكَ أَنتَ الْعَلِيمُ الْحَكِيمُ [32] ﴾

سورة البقرة، الآية: [32]

شكر و تقدير

وَحُقَّ عَلَيه مِنِّي لَمَّا تَوَضَّحَ جَلالُه قِسْطٌ عَظِيمٌ مِنَ الشُّكْرِ

أُوجهه إلى أستاذي الفاضل الدكتور/ **سيدي محمد غيثري** على خطواته الجبّارة في استنقال الباحث عبر مراحل ارتقائيّة مثمرة, له العون في التأهيل، وله المزيّة في الارتقاء. وأَخصّهُ بوافر شُكري وتقديري وعرفاني على توجيهاته التي أسداها لي أثناء التجربة الفريدة في **المعالجة الآلية للغة العربية** و **المعجم الإلكتروني** بشكل خاص، و**اللسانيات الحاسوبية** بشكل عام، حيث كان معطاءً في تقويمه للزلّات، ونُصحه بالعدول عن الأخطاء التي وقعتُ فيها في أثناء كتابة البحث. والتقدير له موصولٌ على علمه الذي يتمّتع به، وعلى إنسانيّته التي قد انطبع عليها. والشكرُ له أيضاً على صبره وتحمله عبء رعايتي- وفقه الله ورعاهُ- فقد كان لي الأب قبل الشيخ والمعلم، حيثُ ساندني كثيراً وشدّ من عزيمتي، وما فتئ يوجهني، وينيرُ لي السبيل ويذلل الكثير من الصعوبات التي واجهت البحث، ويهدئ من روعي كلّما جنحت سفينتي، حتى يسّر اللّه لهذا البحث أن يخرج على يديه فلهُ الدعاء بدوام العافية...

ومن فضل الله تعالى كذلك أن كان في تقويمه ومناقشته ونقده ونقضه، زمرةٌ ممّن أخلصتُ لهم الحبَّ في الله، أسَاتذَتي الموّقرين الّذين تجشموا عناء قراءته، حفظهم الله تعالى، وشكرَ لهم ذلك، ، وأخصّ هنا بخالص التحية والتقدير والعرفان للأستاذ الدكتور/ **سالمي عبد المجيد** رئيس قسم علوم اللّسان بجامعة الجزائر2، الذي أعطى العربية من جهده فأعطتهُ وَقَارَها، كما أشكر الأستاذين الكريمين : الأستاذ الدكتور/ **لوصيف الطاهر** و**الأستاذ الدكتور/ سلامي عبد القادر** اللّذين أكنّ لهما جلّ التقدير والاحترام .

كما أوجه شكري إلى الدكتور/ **عبد الرحيم محمد الأمين** – على دعمه المستمر الذي به أحسَب نفسي مسترسلاً مع صناعة البحث، وفي ضوئه اطمأن قلبي لما حَسبته. وأخيراً أقدم من الشكر أجزلَه، إلى الدكتور/ **عزّ الدّين حفّار** ؛ إذ مرّ نسيمُ الشكر في عَذَباته، فنشر عُرفَ الطيب في كل مواقف عباراته.

ولا أزعم بهذا الحمد أو الشكر أني قضيتُ ما علي من حقٍّ أو واجب.

أمين قدراوي

إهـداء

إلى أمي
.. أبي
.. إخواني .. أخواتي

بدر

"إنِّي رأيتُ أنَّه لا يكتُبُ إنسانٌ كتابًا في يومِه؛ إلاَّ قالَ في غَدِهِ: لو غُيِّرَ هذا لكان أحسنَ، ولو زِيدَ كذا لكان يُستَحسَنُ، ولو قُدِّمَ هذا لكان أفضلَ، ولو تُرِكَ هذا لكان أجملَ. هذا مِنْ أعظَمِ العِبَرِ، وهو دليلٌ علَى استيلاءِ النَّقصِ علَى جُملةِ البَشَرِ[1]"

القاضي الفاضل عبدالرحيم البيساني

1- كان الأُستاذ أحمد فريد الرِّفاعي(ت 1376 هـ) هو الَّذي شهَّر هذه الكلمةَ؛ حيث وضعها أوَّلَ كلِّ جزء من أجزاء " مُعجم الأدباء " لياقوت الحمويّ، وغيره من الكُتب، وتداولَها النَّاس عنه منسوبةً إلى العماد الأصفهانيّ!! والصَّواب نسبتُها للقاضي الفاضل، بعثَ بها إلى العماد؛ كما في أوَّل " شرح الإحياء " للزَّبيديّ (3/1)، و" الإعلام بأعلام بيت الله الحرام " لقطب الدِّين محمد بن أحمد النهروانيّ الحنفيّ (ت 988) هـ.

مُقَدِّمَة

بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ

الحمد لله ، والصلاة والسّلام على أشرف خلق الله سيّدنا محمّد بن عبد الله ؛ خاتم النبيين وإمامُ المرسلين، وعلى آله وصحبه أجمعــين، إلى يوم الدين.

أما بعد:

إن صناعة المعاجم العربية تتصف في عمومها بالطابع التجريبي الذي يعتمد على منهجية علمية حقيقية وأسس نظرية، زيادة على الخبرة العلمية التي اكتسبها واضعو المعاجم من خلال دراسات سابقة.

و مع تطور اللّسانية الحديثة عامة و علم المفردات خاصة، بدأت تظهر في "المعجم" آثار هذا التطور نحو إرساء أسس علمية موضوعية تستندُ إلى نتائج البحوث التي أُقيمت في مختلف ميادين علم اللسان الحديث، فدراسة المعجم وقضاياه دليل على وعي الأمّة بأهمية لغتها ، ففي هذا العصر أصبح من أعظم ميادين الدراسات اللغوية ، وبقدر تقدّم الأمم يظهر اهتمامها بالمعجم وصناعته .

انطلاقاً من هذه الرؤيا نوّد أن نُسهم بمحاولة للإفادة من هذا التطور العلمي إلى رسم خطوط قد تُفيدُ في صناعة معاجم حاسوبية إستنادا إلى المكتسبات المنهجية العلمية التي أفضت إليها أبحاث اللسانيات الحديثة لتطبيقها على لُغتنا العربية، فلهذا الموضوع أهمية كبرى للاستجابة لطموحات الأمة العربية إزاء تراثها اللّغوي عامةً، والمعجم العربي خاصةً.

إنّ المعاجم في اللّغة العربية صناعة قائمة بذاتها وعلم له أسسه ومنهجياته، وما جهود الخليل بن أحمد، والزبيدي، وابن منظور، والأزهري... وغيرهم إلاّ دلائل تشهد بازدهار هذا العلم وإحكام تلك الصناعة، وإذا كانت الحاجة إلى المعجم حاجة أبدية لا يمكن أن تنتهي إلاّ بانتهاء اللّغة وموتها، فإن الضرورة تقتضي تحديث أساليب التعامل مع المعجم وتطوير سبل الاستفادة منه، ولا سيما أن المعجم من المصادر الأساسية التي لا يمكن لباحث أو دارس أن يستغني عنها مهما كان حقل تخصصه ومجال دراسته؛ ومن هنا بدأ التفكير في حوسبة المعجم العربي للتغلب على صعوبة الوصول إليها من قبل الباحثين والطلاب، والتغلب على الحجم الكبير الّذي يحتله بعضها، دع عنك السرعة الزمنية التي هي من أهم خصائص البحث الحاسوبي؛ وتسهيل تطوير المادة المعجمية وإثرائها.

إنّ ما في أيدينا ليس دراسة في اللّسانيات أو في الحاسوب، بل هي دراسة في حقل اللّسانيات الحاسوبية مطبقة على اللّغة العربية، وعليه- فالمجال المعرفي لهذه الدراسة يمكن

استثماره بين منظومتي اللّغة والحاسوب، فكان لابدّ لها أن تنتقي من اللّغة العربية ما يهم أمور معالجتها آليًا، وتنتقي من الحاسوب تلك الجوانب الفنية ذات الطبيعة اللّغوية متحاشية الدخول في الجوانب التقنية.

فكانت اللّسانيات وسيلة، والحاسوب أداة، والمعجم غاية... يلخصها هذا الثالوث في التقاء اللّغة والحاسوب وهو الالتقاء الّذي أطلق ثورةً حقيقية في معالجة اللّغة العربية آليًا بشكل عام والمعجم العربي بشكل خاص.

إنّ موضوع البحث هو حلقة وصل بين اللّساني والحاسوبي لتحويل قضية معالجة اللّغة العربية آليًا من مجال للصراع العلمي إلى موضوع للحوار المنهجي ولاشك أن اللّساني يمثل الأساس في هذه المعالجة، فعلى عاتقه تقع مسؤولية جمع المادة وتوصيفها وتصنيفها عبر قواعد بيانات مضبوطة تمكن الحاسوبي من معالجتها آليًا بسهولة تامة.

فاستخدام اللّغة العربية في الحاسوب كان دافعًا قويًا للعدد من الباحثين اللّسانيين للتعاون مع الباحثين الحاسوبيين للسعي وراء حلول جذرية تعالج مختلف إشكالات المعالجة الآلية للغة العربية، فأصبح من الضروري إعادة صياغة تراثنا من الدراسات اللّغوية صياغة جديدة وذلك على ضوء الإنجازات التي تحققت في اللّسانيات الحاسوبية في اللّغات الأخرى حتى يصبح بالإمكان التعامل السليم مع اللّغة العربية بخصائصها الذاتية.

ويأتي إذن هذا البحث كمحاولة متواضعة تساهم في طرح أبعاد هذه الدراسة وتحديد متطلبات حلّها حلاً جذريًّا يضمن لنا ألا يكون ثمن استفادتنا من هذه التقنيات الواعدة هو فقداننا لهويتنا الثقافية وتراثنا الحضاري؛ كما يشكل محاولة للإسهام في سدِّ هذا النقص في مثل هذه الدراسات وخاصة ما يتعلق بالمعجم الإلكتروني للمعالجة الآلية للغة العربية كموضوع للدراسة.

ولكي تتجلى أمامنا الآفاق المتاحة للمعالجة الآلية لنظام المعجم العربي، سنحاول التعرف في هذه الدراسة على خاصية من خصائص لغتنا العربية والإمكانيات المتاحة لإقامة علاقات متبادلة مع إمكانيات الحاسوب.

فالمعجم اللّغوي العربي يتصف بحجم هائل من الجذور والصّيغ الصرفية مع تعدد مفرداته الخصبة نتيجة للإنتاجية الصرفية العالية الأمر الذي يجعله موردًا طيبًا للمعالجة الآلية، ولأجل ذلك سيكون موضوعنا- بإذن الله سبحانه- موسومًا: **المعجم الإلكتروني للمعالجة الآلية للغة العربيةالواقع والآفاق- مقاربة لسانية حاسوبية.**

وقد اقتصرت دراستي على مفردات القرآن الكريم كعينة في الدراسة التطبيقية؛ فلا بدّ للمعجم الإلكتروني للمعالجة الآلية للغة العربية أن يحوي مفردات جامعة، وليس هناك ما يحقّق

هاته الخاصية أفضل من مفردات القرآن الكريم هذا بالإضافة إلى أن الكلمات والتعابير الأكثر استعمالاً في اللّغة العربية واردة في القرآن الكريم في غالبيتها العظمى، ذلك ما يؤكده قول الله تعالى: ﴿مَا فَرَّطْنَا فِي الكِتَابِ مِنْ شَيْءٍ﴾*.

والبحث في ترتيب مراحله يركز على عرض الجوانب الأساسية لعلاقة الصرف بالمعالجة الآلية وذلك في محاولة لإبراز أهمية هذه المعالجة في تذليل العقبات أمام المستخدم للوصول إلى غايته المنشودة؛ فقد أظهرت الدراسات أن هذا الأسلوب من المعالجة الآلية يمتلك خصائص إيجابية تتجلى في السرعة وكفاءة المعالجة التي توظفه كأداة ناجحة للبرنامج؛ تمكنه من الوصول إلى جملة من الإيجابيات والارتقاء بهذه التقنية "المعالجة الآلية للغة العربية" إلى مستويات أداء أفضل واعتمادها كإحدى الأدوات الفاعلة لبناء المعجم الإلكتروني المنشود.

ولاختيار موضوع جاد لمثل هذه الدراسات لابدّ من توافر جملة وجيهة من الأسباب أو الدوافع التي تدعو الباحث إلى هذا الاختيار، ويمكن تمييز دوافع البحث بين نوعين لغوية وتقنية.

1- وتتمثل الدوافع اللّغوية فيما يلي:

- قصور الدراسات اللّغوية الحديثة في مجال اللّغة العربية وخصوصًا ما تعلق بالمعالجة الآلية للغة العربية.
- تفنيد الإدعاءات المغرضة ضدّ اللّغة العربية واتهامها بالقصور عن مواكبة عصر المعلومات.
- دخول تطبيقات الحاسوب في مجالات العلوم الإنسانية.

2- وتتجلى الدوافع التقنية فيما يلي:

- تضخم المعلومات وتعقدها يتطلبان سرعة ودقة في معالجتها آليًا، ومن ثم تبرز أهمية اللّغة العربية والحاسوب كمدخل أساسي لاستغلال هذا المورد.
- ضرورة إقامة الحوار بين الآلة والمستعمل باستخدام اللّغة العربية.
- ضرورة الدفع باللّغة العربية إلى أخذ المكانة اللائقة بها كلغة عالمية في مجال البرمجة.

وفي ظل هذه الشروط عالجتُ الموضوع في ضوء تساؤلات مفادها :

- عن أي مستقبل وهدف يقدّمه المعجم العربي في ظلّ مواكبة مستجدات العصر، وعصر المعلومات؟، ما هي الخصائص الآلية التي أصبح يتوفر عليها الحاسوب في سبيل تحويل المعجم

*- سورة الأنعام، الآية: 38.

الورقي إلى إلكتروني؟، وكيف يمكن استثمار وتوظيف منظومة الصرف العربي في بناء معجم إلكتروني للمعالجة الآلية للغة العربية؟، ما هو المعجم الإلكتروني وما مزاياه وما حاجتنا إليه؟، ولماذا المعجم الإلكتروني للمعالجة الآلية للغة العربية؟، وهل يمكننا وضع قواعد بيانات تسهّل إنجاز برنامج حاسوبي لبناء معجم إلكتروني للمعالجة الآلية للغة العربية؟؛ إذن: في ظلّ هذه التحدّيات، **هل يحق لنا اعتماد معجم إلكتروني للمعالجة الآلية للغة العربية؟.**

إن أهمية الموضوع موجهة صوب تحقيق حلول ناجعة لهذه التساؤلات، وحتى نعالج صلب هذه الإشكالية من مستوياتها المختلفة، لا بدّ من توافر منهج يتماشى ومتطلبات المّادة الخاضعة للدراسة والتحليل، فطبيعة الموضوع تقضي الاستعانة بمجموعة من المناهج العلمية والوسائل التحليلية والتفسيرية، بحيث يتسنى لي تغطية مختلف مراحل الدراسة، وهذا ضروري للإحاطة- إن شاء جلَّ وعلاَ- بأهمّ جوانبه، فمنهجية الدراسة تتغير حسب طبيعة مادة البحث في كل طور من أطواره، وأحيانًا تتمازج بمنهجية أخرى فتتكامل معها وتنسجم خدمة للنتائج المرجوة؛ فطبيعة الدراسة اقتضت الاعتماد على منهج علمي أقرب ما يكون إلى المنهج الوصفي التحليلي، فالمناهج تتعدد وتتنوع بتنوع الموضوعات وبما أن اللسانيات التطبيقية تعتمد من باب أولى على المناهج الوصفية التحليلية خدمةً للسانيات، كونها تستفيد من معارف علمية متنوعة لأغراض لسانية تطبيقية، فتعدد المعارف يقتضي تعدد المناهج، أضف إلى ذلك فهي تتكامل عندما تهدف إلى خدمة موضوع واحد ألا وهو معالجة اللغة العربية آلياً، لكن هذا لم يمنعنا من الاستعانة بالمنهج التاريخي على اعتبار أنّه لا يمكن فهم الحاضر إلاّ بدراسة الماضي، فالحاضر وليد الماضي.

وتتأكد الحاجة في الباب الثالث في الفصل الثاني من هذه الدراسة أنّه إذا كانت الغاية من البحث اللّغوي هي التدقيق والتمحيص والوصول إلى الحقيقة وفق أيسر السّبل وأدقّها وأسرعها فلا بدّ إذا بالاستعانة بمنهج هو في الوقت ذاته ضرورة ملحة لهذا الغرض ألا وهو المنهج التجريبي، وهنا تبرز العلاقة بين هذه الغاية والأهداف المرجوة من هذا البحث، ومن هذا المنطلق يسمى المنهج في هذا البحث **منهجًا تكامليًّا.**

لقد ميّز الطابع التداخلي لموضوع البحث أحد الصعوبات التي واجهتها في هذه الدراسة سعيًا مني خلق توازن بين شقيّ الموضوع أي بين اللّغة والحاسوب، ما دفعني إلى الرجوع إلى مراجع متنوعة ومختلفة، تعرض أسس المناهج المعرفية لمثل هذه البحوث، وكذلك الرجوع إلى الكتب المتخصصة بالنظريات والدراسات اللغوية الحديثة، إلى جانب بعض الدراسات التي تتناول أثر التقانات الحاسوبية والمعلوماتية في مجال حوسبة اللغة والمعجم. كما أفاد البحث من

دراسات بعض الباحثين، فيما يتعلق بطبيعة اللغة العربية سواء أكان من ناحية غناها اللغوي أم من ناحية بعض إشكالاتها اللغوية، وقد حاول البحث إثارة أكبر عدد ممكن من المشكلات التطبيقية والعملية التي يمكن أن تواجه مثل هذه الدراسات، من مختلف المستويات الصرفية والنحوية والدلالية والمعجمية. كما اعترضتني صعوبات أخرى أخص بالذكر منها دون الحصر: تعدد المصطلحات للمدلول الواحد في اللّسان العربي الحديث؛ علاوة على قصور الترجمة في اللّغات الأجنبية إلى اللّغة العربية التي وردت من لدن مترجمين غير مختصين في هذا الحقل العلمي ناهيك عن ندرة المصادر والمراجع المتخصصة التي تناولت الموضوع من قريب.

وعلى الرّغم من هذا فقد اعتمدتُ في هذا البحث جملةً من المصادر اللّغوية والمراجع ومن المعاجم وكتب المناهج، تتوزعها جهود العلماء العرب والأجانب ما يدعم البحث ويقوده إلى الحقيقة المنشودة، وقد أفدتُ منها منهجيًا وعلميًا، أذكرُ في مُقَدَّمِها: **القرآن الكريم.**

ومن **القدامى:** المعجم المفهرس لألفاظ القرآن الكريم لمحمد فؤاد عبد الباقي؛ كتاب العين للخليل بن أحمد (ت 175هـ)؛ الخصائص لابن جنّي (ت 392هـ)؛ لسان العرب لابن منظور (ت711هـ)؛ المزهر في علوم اللّغة لجلال الدين السّيوطي (ت 911هـ).

أما ما كان من **المحدثين:** معجم الأبنية العربية (الأسماء والأفعال والمصادر) لأحمد محمد عبد الدّايم؛ مقدمة لدراسة التراث المعجمي لحلمي خليل؛ ، معجم الأوزان الصرفية لإميل بديع يعقوب، والمعجم العربي بين الماضي والحاضر لعدنان الخطيب؛ المعجم العربي، نشأته وتطوره لحسين نصّار؛ صناعة المعجم الحديث لأحمد مختار عمر؛ دراسات في علم اللّغة لــفتح الله سليمان؛ اللّغة العربية والحاسوب لــنبيل علي؛ دليل الباحث إلى اللّسانيات الحاسوبية لــوليد العناتي وخالد جبر؛ بحوث ودراسات في اللّسانيات العربية لــعبد الرحمن الحاج صالح؛ نظم استرجاع المعلومات العربية لــ محمد سالم غنيم؛ المصطلحات العلمية في اللّغة العربية في القديم والحديث للأمير مصطفى الشهابي؛ المعالجة الآلية للغة العربية"المشاكل والحلول" لسلوى حماده.

وعن المناهج: تقنيات ومناهج البحث لعبد الناصر جندلي؛ مناهج البحث العلمي وطرق إعداد البحوث لــ عمار بوحوش والذنيبات محمد محمود.

ومن **المراجع الأجنبية** التي وظفتها في هذا البحث ما يخدم الموضوع من جوانبه المتعددة، قد حصرتها في فهرس خاص في نهاية هذا العمل.

ومهما تعددت الدراسات في مثل هذه الدراسات من كتب إلى مطبوعات سائغة سهلة ، إلاّ أن تلك الجهود – باركها الله- لا تكفي لإزالة الغموض الذي يشوب مختلف مجالات هذا البحث، ثم إن تداخل المصطلحات وتشابه المفاهيم يمثل عائقًا آخر في طريق البحث.

وإنّي وإنْ اخْتَرتُ هذا المسار من البحث، في هذا النّوع من المجالات، وفق المناهج المذكورة، فإنّي لا أدَّعي التجديد أو الابتكار أو النّقد أو التصويب، بقدر ما أرمي إلى وصل حلقات البحث والإسهام في بناء صرح اللّغة العربية، وصونه من التّداعي- إن شاء ربّي جلّ وعلا.

و إنّ موضوع البحث لَيُلْزِمُنِي باتباع خطة علمية ذات نسقٍ ونظام وتماسكٍ وذات منهجٍ علمي قويم ومتكامل، حتى أتمكن – على الأقل- من الإلمام بأهمِّ جوانبه، وكذلك تحقيق الأهداف المرجوة من هذه الدراسة، وتشمل هذه الخطة: مقدمة وتمهيدًا وثلاثة أبوابٍ وخاتمة؛ أمّا **المقدمة:** ففيها بيان تعريفي بالبحث وطرح للإشكال، تم ذكر عناصر البحث ودوافعه ومناهجه وأهم مصادره ومراجعه ثم موّاده، وأمّا التمهيد فهو **مدخل** بعنوان: المعجم العربي بين اللغة والحاسوب –آفاق الالتقاء – أتناولُ فيه عرضًا موجزًا لأهم جوانب هذه الدراسة مبرزًا أهميتها وحدودها وأهدافها من خلال تبيان آفاق حوسبة المعجم العربي في ضوء التحديات التي تواجهها اللّغة العربية ومعالجتها آليًا مستعرضا نموذجا للكفاية المعجمية لبناء معجم إلكتروني للغة العربية ، ثم قمتُ باقتراح مقاربة لسانية حاسوبية يرتكز عليها المشروع في دراسة تطبيقية لإنشاء قاعدة معطيات لبناء المعجم المنشود.

وبعد ذلك أنتقلُ إلى المرحلة الجادّة من البحث وهي من ثلاثة أبوب: واخترتُ **للباب الأول** عنوانًا: " نحو معجم إلكتروني للغة العربية في ضوء اللسانيات الحاسوبية

* قضايا نظرية ومنهجية*، و درستُ في **الفصل الأول:** المعجم العربي: نحوتوصيف جديد في اللسانيات الحاسوبية، تناولتُ فيه جوانب هامة وهي: الإطار العام لمفهوم اللّسانيات الحاسوبية، والتعريفُ بها، من خلال قضاياها وخصائصها، وكذا أهدافها ومجالاتها التطبيقية، مستعرضا في نفس الوقت تجارب علمية ورؤى مستقبلية للسانيات الحاسوبية في خدمة الدراسات اللغة العربية عامة و المعجم العربي خاصة.

وفي **الفصل الثاني** الذي خصصتُهُ للحديث عن حوسبة المعجم العربي: نحو مقاربة لسانية حاسوبية لمعاجم آلية عربية، درستُ فيه المعجم العربي من العين إلى الحاسوب، و علاقة المعجم العربي بالحوسبة العربية، و نظرة تاريخية للمعاجم الورقية نحو المعاجم المحوسبة، و

المعجم الإلكتروني للغة العربية : المشروع الذي لابد منه ، ثم خصصتُ مبحثا مُهّما لقواعد المعطيات النصية لتطوير المعجم الحاسوبي (الإلكتروني)، وختمتُ الفصل بالفرق بين المعجم الورقيّ والمعجم الآلي، ، وأفردتُ **الباب الثاني** للحديث عن المعالجة الآلية للغة العربية لبناء قاعدة البيانات المعجمية آليا.

ويتضمن:

الفصل الأول: – وهو من أهمّ الفصول – خصصناهُ للحديث عن المعالجة الآلية للغة العربية: النظام الصرفي العربي – أنموذجا، أفردتُ الحديث فيه عن الإطار العام لمعالجة اللغة العربية آليًا، وكذا المعالجة الآلية للصرف العربي كأنموذج للدراسة التطبيقية، ثم تطرقتُ إلى أهم التطبيقات لمعالجة اللغة العربية آلياً.

الفصل الثاني: هندسة المعجم الإلكتروني: رؤية جديدة لمعجم آلي للمعالجة الآلية للغة العربية: و جاء فيه التعريفُ بالإطار العام للمعجم العربي الإلكتروني، ثم قدّمتُ دراسة تطبيقية لنماذج عديدة لقاعدة البيانات المعجمية للمعاجم الإلكترونية ، وكيفية بناءها ، وسبل تطوير قاعدة بيانات معجمية لبناء معجم آلي عربي.

أمّا **الباب الثالث** الموسوم: توظيف تقنيات المعالجة الآلية للغة العربية لإعداد وتطوير معجم للغة العربية، فيأخذ الجانب التطبيقي حقّه ، حيث تعتمدُ الدراسة على التحليل الوصفي في شقها الأول لبعض الدراسات العربية والغربية، فجاء **الفصل الأول** بعنوان: المعجم الإلكتروني: نظرة في بعض الدراسات المعجمية– " قراءة في الحصيلة واستشراف للمستقبل" كدراسة تقويمية و خلاصة لما درسناه في فصول البحث كونها تضمّ نماذج متفرقة من دراسات سابقة ، بجمع ما يتّفق من مسائل لبناء قاعدة معطيات معجمية ، ودراستها دراسة تحليلية تقويمية ، فتناولتُ فيه الدراسات الغربية والعربية في بناء قاعدة بيانات لمعاجم إلكترونية.

ثم عقدتُ **الفصل الثاني** بعنوان: نحو بناء معجم إلكتروني للمعالجة الآلية للغة العربية باستخدام نظام آلي لاستخراج جذور الكلمات العربية – مفردات ألفاظ القرآن الكريم "أنموذجا"، وجاء فيه إعداد تصور منهجي للمعجم الإلكتروني اللغة العربية، وكدراسة تطبيقية للموضوع تم عرض خطوات انجاز معجم إلكتروني للمعالجة الآلية للغة العربية باستخدام نظام آلي لاستخراج جذور الكلمات العربية ومفردات ألفاظ القرآن الكريم، نموذجًا كعينة للدراسة التطبيقية.

وفي **خاتمة البحث** ذكرتُ خلاصة النتائج التي كانت عبارة عن حصيلة ما جاء فيه ونتيجتهُ.

وفي الختام أُخلِصُ الشكرَ الجزيلَ مرة أخرى إلى كل من أعانني ، ومدّ لي يد العون في مراحل جمع المادة، والكتابة، والتحليل والمعالجة، وأثرى طرق هذه المراحل بعلمه ، وبنصائحه التي ولّدت لديّ روح المثابرة والاجتهاد. وعلى رأسهم مشرفي **الأستاذ الدكتور سيدي محمد غيثري** لإشرافه على هذا البحث، وما قدّمه من توجيهات سديدة مخلصة وآراء منهجية بنّاءة، وما بذله من جهد ورعاية ومتابعته للعمل حتى وصل إلى نهايته...، والشكر سيكون موصولاً إلى كل من حاول تقويم هذا البحث، أو تصحيح أية معالجة أودت بي إلى نتائج خاطئة، فادّعائي الكمال أمر مرفوض، وادّعاء غيري النقص أو الخلل أمر مقبول، فأَتَمَنَّى أَنْ أَكُونَ قَدْ أَنْجَزْتُ عَمَلِي هذَا عَلَى قَدْرِ حُسْنِ نِيتِي، وَأَنَّنِي حَقَّقْتُ هَدَفِي الَّذِي نَشَدْتُ مَا استَطَعْتُ إِلَيْهِ سَبِيلاً. وَلِي عَزَاءٌ فِي مَقُولَةِ د. جنسن:

"يَتُوقُ كلُّ مَنْ يُؤَلِّفُ كِتَابًا إِلَى الْمَدِيحِ، أَمَّا مَنْ يُصَنِّفُ قَامُوسًا فَحَسْبُهُ أَنْ يَنْجُوَ مِنَ اللَّوْمِ".

ولا أنسى في النهاية تقديم الشكر إلى كلية الآداب واللّغات بجامعة تلمسان التي أقدّم هذا البحث في رحابها .

وأسألُ الله أن يكون هذا العمل حلقة علمية، تليها حلقات أخرى خادمة للغة العربية المجيدة.

والسلام عليكم ورحمة الله وبركاته

باريس، يوم الإثنين 02 شعبان 1437هـ، الموافق: 09 ماي 2016 م

الطالب: أمين قدراوي

تمهيد

المعجم العربي بين اللغة والحاسوب

* آفاق الالتقاء *

المعجم العربي بين اللغة والحاسوب

* آفاق الالتقاء *

أولاً - وضع نموذج الكفاية المعجمية لبناء معجم إلكتروني للغة العربية

ثانيًا - الهندسة المعجمية للغة العربية

ثالثاً- قراءة في موضوع البحث : فرضية الدراسة

رابعاً- خوارزميات توليد و تحليل المفردات العربية

خامساً- مقاربة التحليل الصرفي

سادساً- خوارزميات التوليد الصرفي في بناء معجم إلكتروني للغة العربية

سابعاً-أهداف الدراسة

ثامناً- العوائق والحلول المقترحة للنهوض بمشروع المعجم الإلكتروني للغة العربية

تمهيد

يندرج هذا البحث في مجال يدعى بعلم اللّسانيات الحاسوبية[1] ذلك الفرع من اللّسانيات التّطبيقية الذي يستفيد من البحث اللّساني البحث من جهة[2]، وعلوم الحاسوب من جهة أخرى[3]، وهو العلم الّذي يعدّ العمود الفقري لمثل هذه المشاريع[4] وقد أصبحت معالجة اللّغة آليًا في العقدين المنصرمين أمرًا ملحًا في مختلف مجالات البحث اللّساني البحت وأكثر إلحاحًا في المجالات الحاسوبية لأغراض شتى[5] ولا شك أن اللّساني يمثل قطب الرّحى في هذه المعالجة فعلى عاتقه تقع مسؤولية جمع المادة وتوصيفها وتصنيفًا يعتمد المطرد أمّا الشاذ فلا يتم طرحه وإنّما معالجة آلية تخضع إلى قوانين صارمة، خاصّة وأن اللّغة العربية تتميز بجملة من الخصائص تجعلها أكثر من اللّغات تطويعًا للحاسوب[6] فالدرس اللّساني الحاسوبي والتطبيقات الحاسوبية يمكن أن تفيد من المعالجة الحاسوبية للغة العربية[7] في لغات أخرى وليس العكس[8].

ينحصر الهدف في هذا البحث الذي يعتبر جزءا من مشروع يتمثل في بناء معجم إلكتروني للمعالجة الآلية للغة العربية، في الطريقة التي يجب اتباعها لتقديم وإدخال المعطيات اللسانية مصورنة و**موصوفة**[9] إلى عقل الآلة ، وذلك حتى يتمكن المتخصصون، على ضوء هذه

---

1 –Linguistique informatique.

2– للمزيد من بسط القول ينظر: شارل بوتون، اللسانيات التطبيقية، ترجمة د/ قاسم المقداد، محمد رياض المصري، دمشق، سوريا، د ط، د ت، ص7 ومابعدها.

3– للمزيد من التفصيل في هذا العنصر، أنظر: إشكالية التعريف الواردة في البحث، ص 19 .

4– نبيل علي،اللّغة العربية والحاسوب، دراسة بحثية، دار تعريب، الطبعة الأولى، الكويت، 1988، ص 117.

5– محمد صيني نحو معجم عربي للتطبيقات الحاسوبية، ندوة استخدام اللّغة العربية في تقنية المعلومات، الرياض 1992، ص 511.

6– محمد الزركان، اللّسانيات وبرمجة اللّغة العربية في الحاسب، ندوة استخدام اللّغة العربية في تقنية المعلومات، الرياض، 1992، ص 57.

7– تستند اللغة العربية إلى نظام متسق، تقيده ضوابط دقيقة، وتحكمه قواعد مطردة أو شبه مطردة، إلا أنه قد تغزوها في بعض الأحيان مظاهر شوارد وشذوذ عن بنية النظام الكلي الذي يحاول استيعابها، ويحكم السيطرة على شبكة العلاقات التي تحدد معالمها، وتحكم ظواهرها، وتشكل تطورها، وتكشف عن أعماقها الغنية بالدلالات والمباني.

8– نبيل علي، د. نادية حجازي. الفجوة الرقمية، سلسلة عالم المعرفة، العدد، أوت 2005، ص 231.

9– عبد السلام المسدي، اللسانيات وأسسها المعرفية، الدار التونسية للنشر، تونس ، د ط ، د ت، ص153.

العلاقة من المعرفة اللسانية العميقة "**وصفا وتصنيفا**"[1]، والإلمام بالمعرفة الحاسوبية برمجة وأيللة[2] ، اعتمادا على ما توفره المعلوميات من إمكانية أدواتية و إجرائية ضرورية ، وخاصة تلك المنظومة البرمجية المنطقية التي قوامها الخوارزميات الصارمة، والمرمى طبعا هو معالجة الظواهر اللغوية المتشعبة المسالك، صوتا وصرفا وتركيبا ومعجما ودلالة وتداولا.

فالمشروع –بناءًا على ما ذكرته– يطمح إلى تقديم مساهمة منهجية للتعامل مع البحث اللّساني الحاسوبي، وبالأخص المعالجة الآلية للغة العربية التي تزداد أهمية مع ارتفاع وتيرة تقدّم تكنولوجيا المعلومات، خاصة أن اللّغة أصبحت هي المنهل الطبيعي الّذي تستقي منه التكنولوجيا أسس ذكائها الاصطناعي، والأفكار المحورية بلغات برمجية، وذلك من خلال تقديم مشروع متواضع يتمثل في بناء معجم إلكتروني للمعالجة الآلية للغة العربية.

وهنا تبرز العلاقة بين هذه الغاية ومشروع بحثنا الذي هو التعريف بالإمكانات التي يقدّمها الحاسوب لمعالجة اللّغة العربية آليًا بوصفها لغة طبيعية، وخاصة في المجالين البحثي والتعليمي، فهو يتعرض للطريقة التي يجب إتباعها لإدخال اللّغة العربية إلى عقل الآلة من خلال إعداد قاعدة معطيات بهدف إقامة الحوار بين الإنسان والحاسوب باللّغة العربية، وعندما نقول معالجة اللّغة في هذا السّياق فإنّنا نقصد "الهندسة اللّسانية"[3].

1– تظهر التجربة العملية أن ثمة فارقاً كبيراً بين وصف اللغة وتجريد أمثلتها وضبط أحكامها حين يكون هذا الوصف موجهاً للإنسان، وحين يكون مصمماً ليودعَ الحاسوب.

ولعل هذا ما حمل د. نهاد الموسى على إقامة الفرق بين هذين العملين. فقد سَمَّى ما يُعْمل للإنسان "الوصف"، وما يُعْمَلُ للحاسوب"التوصيف".وبيان الفرق بينهما ماثل في أن وصف العربية ما وقع للعلماء العرب من قواعد مستنبطة من الأداء اللغوي الواقعي، وهو مبني في شطر منه على أن المُستقبِل يسهم إسهاماً فاعلاً في الحدث التواصلي، مضافاً إلى ذلك ما يتحصل للإنسان من معرفة بالحدس والسليقة والخبرة المعرفية والتثقف والعرف اللغوي والمقام.

وأما التوصيف فإنه ينتظم الوصف اللغوي المجرد، مضافاً إليه العناصر التي يتعرفها الإنسان بالحدس والسليقة والقرائن المتعددة اللفظية والمعنوية والموقفية. ولما كان الحدس أظهر ما يتكئ عليه الإنسان في تعرفه اللغة وأدائها، ولما كان الحاسوب يفتقر إلى هذا العنصر البشري الخالص، وجب على المُوَصِّف أن يتدارك هذا النقص، ليبلغ بالحاسوب مَبْلَغَ المعرفة الإنسانية باللغة.

يعد مفهوماً "الوصف والتوصيف" عنصراً رئيساً في الأطروحة التي أسسها الدكتور نهاد الموسى في تهيئة العربية للتمثيل الحاسوبي. ويمكن تحصيل تفاصيل وافية بهذين المفهومين، والأطروحة عموماً في: ←

– نهاد الموسى، العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية.

– نهاد الموسى، من الوصف إلى التوصيف، مقاربة في حوسبة العربية. فصل من كتاب الثنائيات في قضايا اللغة العربية من عصر النهضة إلى عصر العولمة.

– نهاد الموسى، مقدمة في تمثيل الكفاية اللغوية للحاسوب، ورقة مقدمة إلى ندوة "الهوية اللغوية والعولمة" جامعة البترا الأردنية الخاصة، 2003.

2– نستخدم كذلك هذه الترجمة مرادفا إلى جانب مصطلح "الأتمتة" نفلا عن اللغات الأجنبية (Automatisation) (انظر قاموس مصطلحات المعلوماتية فرنسي–عربي لصاحبه إ.حداد مكتبة لبنان حرف A .

3– نبيل علي،الثقافة العربية وعصر المعلومات، عالم المعرفة، المجلس الوطني للثقافة والفنون والآداب، الكويت، العدد 265، يناير 2001، ص261.

فالجمع بين هذين النوعين من المعرفة هو المحك في التقدم والنهوض الذي يحققه مجال البحث في المعالجة الآلية للغة العربية.

ولطالما تجاهل واضعو البرامج الحاسوبية الجانب المعرفي اللّساني مكتفين بالمعرفة الحاسوبية، وهو ما أدى إلى تطوير برامج عربية يمكن القول إنّها ما تزال في بداية الطريق، الأمر الّذي انعكس سلبًا على صناعة البرامج بواسطة قاعدة بيانات أساسها اللغة العربية[1]...

إن الطبيعة الجبرية لنظام اللغة العربية المتمثلة في منظومة الخوارزميات اللسانية التي تجمع بين مختلف مكونات النظام، فإن الربط بين هذه المستويات أمرٌ بالغُ التعقيد لايمكن أن تقوم به الآلة إلا إذا تم تزويدها بالقواعد اللسانية الصورية وبالخوارزميات اللغوية بنوعيها التوليد والتحليل. وهذا ما تم وصفه فعلا في إطار المنهج اللاساني الذي نتبناه هنا والذي يعرف بالمعجم –التركيبي (lexique grammaire)[2]والذي استمد مبادئه النظرية من نحو زيليغ هاريس (1990) (Z.S.Harris) وطوره موريس كروس (M.Gross) في نماذجه منهجاً وصورنةً.

وعلى هذا الأساس:

– ما هي أهم المستويات اللّسانية قابليةً للحوسبة التي يمكن استغلالها لوضع نموذج للكفاية المعجمية ؟.

1– تجدر الإشارة إلى أن علاقة اللّسانيات بهندسة الحاسوب هي علاقة أخذ وعطاء فكما أنّه يمكن استخدام الحاسوب حاليًا لإقامة نماذج لغوية وتحليل الفروع اللّغوية المختلفة فإنّه يمكن تحديد مجالات المعالجة الآلية للغة العربية على المستوى اللّغوي في ما يلي:

– الصرف الحاسوبي La morphologie informatique

– التركيب الحاسوبي La syntaxe informatique

– الدلالة الحاسوبية La sémantique informatique

– المعجمية الحاسوبية La lexicologie informatique

وتكمن أهمية هذه المجالات وتزداد هذه العلاقة أهمية بالنسبة للغة العربية من وجهة نظر الهندسة اللّغوية العربية، لغة قابلة للاستجابة بكل يسر للإجراءات الهندسية اللّسانية في أغلب المستويات ذات الارتباط بالجانب الصوري وخاصة في باب الصرف والتركيب، للمزيد من بسط القول ينظر:عبد الله أبو هيف،مستقبل اللغة العربية " حوسبة المعجم العربي ومشكلاته اللغوية والتقنية أنموذجا"، اللغة العربية، المجلس الأعلى للغة العربية، الجزائر، العدد العاشر، 2004، ص 115،116.

2– نظرية المعجم التركيبي (lexique-grammaire) هي هدف النظرية اللسانية التي انبثق عنها النحو التأليفي (la grmmaire combinatoire) وهو منهج لساني يقوم على الوصف والتصنيف (la taxinomie) السائدين في البيولوجيا والرياضيات والفلك، وهذا التصنيف يعتمد المعطيات النظرية التي تعلل جمع الأصناف المتجانسة أو المتماثلة مع بعضها البعض، كل ذلك في إطار معادلات دقيقة تتميز بالشمولية والاطراد.ففي النظام اللغوي يعتبر الفعل وحدة توزيعية أساسية وعنصر يحدث تركيبا في ذاته (في كتب زيد الدرس) وباعتباره مدخلا معجميا فإن له أكثر من تركيب واحد وهذا ما تقوم به العلمية التوزيعية الحسابية. والنموذج الذي نريده هنا يؤطره الطرح النظري المتبع في ميدان المعلوميات (نحو السياق الحر، والنحو الشكلي مثلا) حيث يتم بناء نماذج وقوالب لسانية = صورية تمثل الأصل النظري للنسق اللساني، وحيث القابلية بالتواصل، هذا المنهج يختلف عن المناهج التقليدية والتوليدية إذ أثبتت قابليتها للتطبيق والمواكبة وتعليل الفرضيات الموضوعة والتأكد منها وذلك ما شرع فيه مختبر الأوتوماتيك والتوثيق اللساني بباريس 7 بفرنسا تحت إشراف العالم الكبير مورسي كروس صاحب هذه النظرية المتصفة بالمرونة وقابلية التكييف في معالجة أنظمة اللغات الطبيعية ومن ضمنها اللغة العربية.

أولا: وضع نموذج الكفاية المعجمية لبناء معجم إلكتروني للغة العربية :

يتطلب العمل في المعالجة الآلية للغة العربية التمكن من نوعين متكاملين من المعرفة: المعرفة اللّسانية العميقة بمختلف جزئيات النظام اللّغوي على ضوء أحدث النظريات اللّسانية المعاصرة وخاصة اللّسانيات الصورية، والإلمام بالمعرفة الحاسوبية ذات الصلة بمعالجة اللّغات الطبيعية وخاصة في جانبها البرمجي، فالمشروع ينتسب إلى درس لساني حاسوبي يتردد بين قطبين: أولهما اللّغة، وثانيهما أداة صماء اسمها الحاسوب[1]،إذن في ظل هذه التحديات، كيف يمكن رصد الكفاية المعجمية؟ وما هي السبل لتقييسها وبناء نماذج لقواعد بياناتها على مستوى الحاسوب؟

يكاد يجمع معظم المعجميين والمعجماتيين على أن المعجم هو"حصيلة تاريخ المجتمع وتجاربه من جوانبه المختلفة الحضارية والثقافية والاقتصادية والفكرية والاجتماعية" [2]، أما القاموس فيجمع بين دفتيه الجزء المستعمل من "المعجم" فهو لا يختلف عن كتب تاريخ الفكر والحضارة والثقافة والمجتمع إلا في طريقة تبويبه، فهو كتاب تاريخ تقرأ عناوينه وموضوعاته من خلال مداخله، أي من خلال ألفاظه وكلماته المرتبة ترتيبا قاموسيا خاصاً[3]. إن اللغة والمعجم يُدرسان لتفسير ظواهر المجتمع بأسره، على اعتبار أن المناسبة بين الدال والمدلول مناسبة اجتماعية محضة، بمعنى آخر، إن العلاقة بين الاسم والمسمى هي علاقة اجتماعية مرتبطة بما يجول في أذهان الأفراد من الأسباب التي تدفعهما إلى وضع اسم ما. والغرض الأساس هنا، هو أن تستفيد **المعجمية**[4] من عطاء ونتائج أذهان الأفراد، كما تستفيد من كل العلوم المجاورة من تاريخ واقتصاد ودلالة، وغير ذلك من العلوم التي لها علاقة وطيدة بالمعجمية.

والمعجم لا بد أن يكون قائماً على نظام معين، لأن الكلمة أو الوحدة اللغوية، "تتم معانيها في بضع لحظة (خمس الثانية أو أقل)... فهناك إذن قدرة خاصة على البحث عن الكلمات وإصدار قرار معجمي في نصف ثانية، إن المعجم الذهني المنظم تنظيماً مُحكماً هو الذي يترجم القدرة على التخزين الكثيف، وعلى الاسترجاع السريع"5 .ذلك أن الإنسان كيفما كان نوعه يتوفر على قدرة هائلة لتذكر آلاف الكلمات التي هي مخزونة في معجمه الذهبي.

1- نبيل علي، اللّغة العربية والحاسوب، دراسة بحثية، دار تعريب، الطبعة الأولى، الكويت، 1988، ص 116- 117.

2- عبد العلي الودغيري، دراسات معجمية، تحت قاموس عربي تاريخي وقضايا أخرى، مطبعة النجاح الجديدة الطبعة الأولى 2001، ص9.

3- المرجع نفسه، ص 10.

4- يعرف الأستاذ الدكتور محمد رشاد الحمزاوي مصطلح المعجمية بأنه مصدر صناعي مضموم الميم " المُعْجَمية" ووضعه مصطلحا عربيا يعني به ما هو معروف في الفرنسية وفي الإنجليزية بـــ Lexicologie و Lexicology . وتعرف المعجمية كذلك بعلم قريب منها، قديم جديد و مشهور، كثيرا ما يختلط بها.للمزيد من التفصيل ينظر : محمد رشاد الحمزاوي، المعجمية، "مقاربة نظرية و مطبقة" "مصطلحاتها ومفاهيمها"، مركز النشر الجامعي ، تونس ، دط، 2004، ص 18 وما بعدها .

5- الفاسي الفهري، المقارنة والتخطيط في البحث اللساني العربي، 1998، ط1، ص 163.

فكلما وجد نفسه في حاجة إلى هذه الكلمات التي هي يسترجعها من خزّانه الذهني. فالمتكلم المؤول من الضروري أن يتوفر على آلة معجمية ذهنية تساعده على التخزين الكثيف والمنظم للمواد المعجمية.

ويؤكد عبد القادر الفاسي الفهري أنه بواسطة هذه الآلة الذهنية يستطيع هذا المؤول أن يُصدر قرارات وأحكاماً معجميةً تمكنه من الحكم على كلمة بأنها تنتمي إلى معجمه الذهني وإلى لغته أو لا تنتمي.أي أنه يستقبل من حين لآخر كلمات ليست كالكلمات التي سمعها من قبل، سواء من الناحية الصوتية أو التركيبية أو الدلالية أو في المقام الذي توضع فيه. ومع ذلك يدرك أنها كلمات. وبالطريقة نفسها يسمع متواليات صوتية ويعرف ليست كلمات وذلك بواسطة ذهنه. فهذه الأحكام وهذا النظام الذي يمكن عملية التخزين لدى الأفراد وكذلك عملية الاسترجاع، هو ما يسميه الفاسي الفهري بالمعجم الذهني (Mental Lexicon)[1]. وهكذا، فإن المعجمية الذهنية أساسها الكلمات أو الوحدات اللغوية باعتبارها الوحدات الدلالية الصغرى التي تؤسس عليها الدراسة المعجمية، لأن هذه الوحدات مرتبطة أشد الارتباط بالفكر والتصور المجتمعي، حيث يتم تحويل الكلمة المفردة إلى شيء جماعي يرمز بها لتدل على شيء معين ومتفق عليه من قبل العشيرة اللغوية الواحدة.

وعلى هذا الأساس، فإن ما وصلت إليه الدراسات اللسانية الحديثة هو ما تحول اهتمام اللسانيين من دراسة السلوك اللغوي العقلي إلى نسق المعرفة الذهنية المسؤول عن هذا السلوك. أي انتقال اللسانيين من دراسة اللغة باعتبارها ظاهرة اجتماعية إلى دراسة النسق اللغوي المعرفي المتمثل في ذهن وعقل المتكلمين. وبذلك فإن هذا التحول في حد ذاته هو: "انتصار للواقعية الذهنية التي تستهدف اكتشاف واقع ذهني يكمن خلف سلوك فعلي. يعرف الواقع الذهني في اللسانيات بأنه القدرة أو الملكة اللغوية" [2].

ولقد أصبح اهتمام المعجمية الحديثة منصباً على المعجم الذهني أو الاجتماعي الذي يكتسبه المتكلم بشكل فطري أيّاً كان نوع اللغة الطبيعية التي تبنى وتؤسس عليها مادة هذا المعجم. ولذلك فإن المعجمي يلزم عليه معرفة أهم النتائج التي توجه تصور الباحث للمعجم الذهني. " تبين أن المتكلم بأية لغة طبيعية كانت، توفر على معجم منظم تنظيماً دقيقاً، ومن مظاهر هذا النظام قدرة المتكلم الفائقة على استذكار ما تم تخزينه من مفردات في معجمه الذهني لتحقيق أغراضه التواصلية المتعددة"[3].

1- المرجع السابق، ص 163.

2- عز الدين البوشيخي، خصائص الصناعة الحديثة، مجلة اللسان العربي، ع 46، ص 22.

3- المرجع نفسه، ص 23.

لما كان المعجم حاوي لمجموع المفردات اللغوية مواد وصوراً تعكس بحسب مجموعة من الدارسين[1] عددا من الخصائص التصورية الإدراكية والدلالية عبر مكترمات التقابل وبين وجهي (interfaces) التمفضلات (articulations) الإدراكية المتصلة بالعالم الخارجي أو الواقع المادي والمجرد والتمثيلات التصورية / المنطقية المتصلة بالتعبيرات اللغوية وعلاقتها بالإحالة أو الواقع الخارجي، فالمدركات ترتبط بالعالم الخارجي مباشرة والتصورات ترتبط بالعالم الذهني المؤطر عموما بما يسمى "الصورة المنطقية"، ولذلك يلعب تشكيل العالم الخارجي وهندسته دوراً هاماً في الوضع والاصطلاح وكذلك في بناء الدلالات والتعابير كما يذهب في ذلك باروايز وبيري 1983 [2].

فعملية معجمة الكتل من أشياء وأحداث في اللغات الطبيعية تلعب دورا أساسياً في إدراك وتصور الأسماء والأفعال وتشكيلهما في الواقع.

إن اكتمال الكفاية المعجمية عند متكلم اللغة العربية ينبني على القدرة على إقامة العلائق الملائمة بين لفظ الكلمة أو رسم المفرد وما يمكن أن تسهم به من معنى داخل بنية تركيبة – دلالية وقد تكون هذه البنية جملة أو نصّاً في السياق الكلامي، ولا يمكن أن تأخذ المفردات معناها الفعلي إلا في بنى ومجال معرفي وسياقي لأن المفردات ما هي إلا عناصر ووحدات في تركيب معين.

وعليه فإن الكفاية المعجمية للغة تتمثل في وجود نظام أو آلة حاسوبية معجمية لا تختلف جوهريا عن نظام القواعد أو النحو، وهو أيضا آلة حاسوبية. ومن هنا تصبح الدخلات المعجمية عناصر يوظفها النظام الحاسوبي[3] للقيام بعمليتين ضروريتين لبلوغ معاني الكلمة: عملية تفكيك أو تحليل، مردها إلى كون الكلمة ليست مفردة، لها معنى ذري لا يقبل التحليل، بل إن الكلمة تحتاج إلى تحليل، لفهم دلالتها الداخلية [4].

والمعجم بهذا الاعتبار يمثل نموذجا لكفاية المتكلمين اللغوية، فهو عبارة عن منظومة الوظائف المثبتة في الكفاية المعجمية على شكل متواليات لغوية تسمى بالمداخل المعجمية بدرجتها

1- من هؤلاء الباحثين بستتجويفسكي (1988 و 1991) وزفات (1992) وتيني (1994) وخيري (1998 أ و 1988ب) وكيرون (2003) وبورر(2003) والفاسي الفهري (2003). انظر، معجمية الإدراكات والتصورات في الأبنية اللغوية، عبد الواحد خيري، وقائع الندوة الدولية حول المعجم العربي واقع وآفاق أيام 27/26/25 سبتمبر 2003 بالدار البيضاء.

2- خصائص الصناعة الحديثة، مجلة اللسان العربي، مرجع سابق، ص 22.

3- محمد رحالي، بعض الخصائص الحاسوبية للغة، أبحاث لسانية، المجلد 13 ، العدد 2/1، منشورات معهد الدراسات والأبحاث للتعريب، 2008، ص 68،69.

4- عبد القادر الفاسي الفهري، أزمة اللغة العربية في المغرب بين اختلالات التعددية وتعثرات "الترجمة". منشورات زاوية الطبعة الأولى 2005، ص 37.

المختلفة وبمستوياتها المعقدة[1]. وإذا كانت التجارب المعجمية العربية قد بينت أن العربية كلغة طبيعية قد مكنت الإنسان العربي من متابعة جميع التطورات الكونية في جميع المجالات وأنها لا زالت في أخذ مع الحضارات العالمية، فقد بات من الضروري بناء معجم موسوعي آلي يستوعب ماضي وحاضر الأمة من تراكيب ومفردات ومصطلحات عملية، كما أنه لابد من تكييف هذا البناء بالمستجدات التكنولوجية السريعة حتى تساير التقدم العلمي. وما يُطرح بحدة هنا هو كيف يمكن تمثيل معرفة هذه الكفاية على مستوى آلة ؟ وما هي الأساليب الكفيلة بتمثيلها تمثيلاً منطقياً وعقلانياً؟ تلك كانت منطلقات الذكاء الاصطناعي وأنظمة قواعد المعرفة والأنظمة الخبيرة في تقييس نماذج من القدرة المعرفية البشرية على الحاسوب، ونريد هنا الاستفادة من نتائج هذه الأبحاث في معالجة اللغات الطبيعية والتطبيقات الهندسية في هذا الإطار.

ثانيا: الهندسة المعجمية للغة العربية:

إن الهدف الحقيقي من معالجة اللغة العربية آليا هو إقامة الحوار بين الإنسان والآلة باللغة العربية، وهذا الحوار لن يتأتى إلا بمحاكاة دماغ الإنسان وتقييسه في سيرورة عمله إدراكاً وتخزيناً وإنجازاً[2]. والقصد بالمعالجة الآلية للغات الطبيعية، إنما يعنى به الهندسة اللسانية أو تكنولوجيا اللغة التي تعنى بتحليل المتواليات اللسانية وتوليدها في جميع المستويات اللغوية عبر خوارزميات مجهزة بتقنات متطورة تشتغل بها آلة الاستدلال بمساراتها المعرفية وعلائقها وتمثيلاتها، وقد استجابت اللغة العربية لهذه التطبيقات الهندسية نظراً لسير ارتباط خصوصيتها الصورية والتركيبية بالإجراءات الهندسية.

وتعتمد هندسة اللغة العربية على وصفين لسانيين عميقين هما: اللسانيات الصورية والصناعة البرمجية المرتبطة بمعالجة اللغات الطبيعية ذات البناء الخوارزمي الصوري الصرف والحاسم في تأويل القضايا دون الاعتماد على الظنية والنسبية[3].

1- جواد حسني سماعنه، المعجم العلمي المختص، (المنهج والمصطلح)، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 48، الرباط، المغرب، ديسمبر 1420هـ/1999م، ص 42، 43.

2- عز الدين غازي، المدخل العام، لرسالة لنيل دبلوم الدراسات العليا في الآداب بعنوان: نظام قواعد صرفي – صواتي للغة العربية، مقاربة الفعل. تحت إشراف د.محمد الحناش، كلية الآداب والعلوم الإنسانية ظهر المهراز ، فاس 1999.

3- انظر محمد الحناش، اللغة العربية والحاسوب (قراءة سريعة في الهندسة اللسانية العربية) أو مقاربة في محاكاة الدماغ العربي لغويا (أكتوبر 2002)، مقال منشور في موقع جامعة عجمان www.ajman.ac.ae. ص 1.

إن البحث في الهندسة اللسانية العربية يجب أن ينصب على وضع نظرية صورية ورمزية متكاملة للغة العربية[1] من قبل العارفين بأمور اللغة وعلمائها حتى يكون المبتغى في الوصف هادفا، وحتى لا تنعكس الصناعة البرمجية سلبا على العربية مما يكون سببا في تأخير الركب التكنولوجي العربي بصفة عامة.

ويجمع علماء اللسانيات عموما على أن نظام اللغات الطبيعية معقدٌ ومتشعبٌ من حيث الكتابة والصوت والصرف والتركيب والدلالة والتداول والمعجم، وعلى الرغم من ذلك استطاع باحثون متخصصون وضعَ برامجَ لسانية حاسوبية طُبقت فيها جميع الخوارزميات الصورية، فبرزت مع صناعة البرمجيات مثل الترجمة الآلية والتطبيقات الصوتية كالتعرف على الكلام وتوليفه آليا والتعرف على الأشكال والخطوط الهندسية والمحلل الصرفي والمشكل الآلي والمدقق الإملائي والتلخيص والتشكيل الآليين والفهرسة والتوثيق ومحركات البحث وبناء قاعدة البيانات المعجمية وقواعد المعـــارف المتعددة اللغـــات وغيـــر ذلك من البرمجيات العربية المتقدمة[2]. ومثل هذه الإنجازات استلهمت جل المستعملين والباحثين والمصنعين على حد سواء وذلك في جميع التطبيقات الصناعية للغات الطبيعية.

في ضوء هذه التحديات، كيف يمكن للغة الطبيعية أن تتوفر على معجم منظم تنظيماً دقيقاً ومحكماً؟ وما هي الإمكانية التي تقدمها الآلة لمعالجة اللغة العربية باعتبارها لغة طبيعية؟ وما هي الطريقة المثلى التي يجب إتباعها لأجل إدماج وصهر هذه اللغة بمعجمها الشاسع في عقل الآلة؟

ثالثا- قراءة في موضوع البحث : فرضية الدراسة:

3 – 1– الطبيعة الرياضية والانصهارية للغة العربية:

إنّ وضع معاجم آلية للمعالجة الآلية للغة العربية يستوجب أخذ بعين الحسبان خصائص الكلمات ومدلولاتها وذلك بغية تجاوز ثغرات المعاجم التقليدية، ومن ثم إيجاد حل ناجع للكثير من المقاربات اللّغوية التي تواجه المعالجة الآلية للغة الطبيعية من قبيل اللّبس الدلالي للمفردات والعبارات، والترجمة الآلية وبنوك المعطيات[3] وغيرها من التطبيقات الآلية؛ إلا أنّ التطبيق على

1- عمر مهديوي، التوليد الصوري للأسماء البسيطة في اللغة العربية، المؤتمر الدولي الثاني حول هندسة اللغة وهندسة اللغة العربية، 27-28، يونيو 2005.

2- وفي ذلك كثير من البرمجيات المتطورة التي استقادت من أساليب التكنولوجيا الرمزية المتعددة على قواعد المعارف، على مستوى اللغة العربية نذكر تمثيلا لا حصراً شركة صخر ميديا لبرامج الحاسوب www.sakhar.com وشركة حرف لتقنية المعلومات والبرمجيات الإسلامية www.harf.com وبرمجيات مايكروسوفت العربية وشركات أخرى كتقنيات اللغات والكمبيوتر ومرآة المعرفة والمركز الهندسي للأبحاث التطبيقية ونولوجي ولاياوت وشركة بروجس سوفت...

3- ليلى المسعودي، علم المصطلحات وبنوك المعطيات، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 28، الرباط، المغرب، 1987، ص 87،86.

اللّغة العربية ما يزال في بداية الطريق، هذا على الرّغم من أنّ هذه اللّغة لا تختلف عن أي لغة في العالم فيما يتعلق ببرنامج الكفاية اللّسانية.

فقد أظهرت نماذج الأنحاء الصورية، من طرف نعوم تشو مسكي[1]، وزيليغ هاريس[2] إمكانية التحديد الرياضي لنحو لغة ما، فأصبحت تنظر إلى اللغة بوصفها مجموعة من الجمل، تتكون كل منها من رمز واحد أو أكثر من مفردات اللغة. وأن نحو اللغة يتكون من تحديد رياضي لهذه المجموعة من الجمل. ويمكن أن يتخذ نحو اللغة أشكالا مختلفة، فإذا كانت جمل اللغة محدودة يمكن وضعها إذن في قائمة، ولكن يستحيل وضع جمل اللغة الطبيعية في قائمة لأن عدد الجمل لا نهائي[3] .

وما تتميز به اللّغة العربية عن غيرها يجب أن يدفع إلى مقدمة اللّغات العالمية في التعامل مع الحاسوب، فهي تقوم على مكونين رياضيين هما **الجذر** و**الوزن** وهما معًا غير موجودين في أغلب لغات العالم.

يتولى الجذر وضع البنية الأساسية للكلمة، ويتولى الوزن هيكلها العام، يقوم الوزن بتوزيع الحركات على مختلف حروف الكلمة كما يقوم بتوزيع المورفيمات التي تضاف إلى مكونات الجذر بغرض توليد الكلمات: (سوابق، لواحق وأواسط) هذا التشكيل الرياضي للغة العربية جعل منها لغة انصهارية[4]، خلافًا للغات الأخرى التي تعد لغات إلصاقية[5] فيما يتعلق بتوزيع المورفيمات داخل بنية الكلمة الأساس؛ والقول بالانصهارية يؤدي حتمًا إلى القول بالطبيعة الرياضية الجبرية للغة العربية.

إنّ اللغة العربية لغة رياضية في أساسها مكونة من منظومة من الخوارزميات الصورية، داخلها الجذور مرورًا بالأوزان التي تتمتع بقوة الانصهار المورفيمي المبرمج وخرجها الكلمات والجمل.

– إن المقاربة في مجالها اللّساني الحاسوبي للمشروع ارتكزت على النقاط التالية:

- فيما يتعلق بالتشكيل بالصوائت في العديد من اللغات، يقوم البعد الرياضي للغة العربية بتطبيق خوارزم التطابق بين الوزن والمفردة، وتقرأ الكلمة في النهاية بدون أي عناء.

---

1 – نعوم تشو مسكي، البنى النحوية، ترجمة يؤيل يوسف عزيز ومراجعة مجيد الشماطة، منشورات عيون، الدار البيضاء، 1987.

وللمزيد من التفصيل ينظر: N.Chomesky , G.A.Miller, l' Analyse Fromelle des naturelles,traduit par Ph. Richard et N.Ruwet Paris 1971

2 –Z.Harris, Structures mathématiques du Langage, Dunod (1968), Paris.

3 –Gérard SABAH, Intelligence Artificielle et langage, Processus de compréhension T 2 hermes. 1990.p.127.

4 –Fusion.

5 –Ensemblist.

- وفي مرحلة لاحقة يتم تفعيل الوزن/الميزان عن طريق خوارزميات الإقحام التي تقوم بإدراج الزوائد في البنية النظرية للجذر بهدف توليد الكلمة[1]. وللمستوى الصرفي دورٌ كبيرٌ في توزيع هذه الوظائف في حالة الجذر الثلاثي أو الرباعي، أما المستوى الصوتي فدوره يبقى في توزيع الوظيفة الصورية على لام الكلمة لأنه ينفذ الأدوار العاملية التي تميز نظام اللغة العربية.

من هذا المنطلق تصبح هذه اللغة، لغةً رياضيةً لكونها تتألف من نظام خوارزمي صوري دخلها الجذور مروراً بالأوزان[2] وخرجها الكلمات البسيطة والمركبة كما في الشكل التالي[3] :

$$\text{الكلمة} = \frac{\text{الجذر} \times (\text{ف ع ل})}{\text{الوزن/ الميزان}}$$

– تقترح المقاربة على طريقة لاستخراج جذور كلمات اللّغة العربية بالاعتماد على التقنيات اللّغوية للمعالجة الآلية للغة العربية نحو بناء معجم إلكتروني للغة العربية.

– ترتكز الطريقة على مفهوم الوزن أو الميزان الصرفي[4].

– تُمثّل الكلمة المدخلة في البرنامج اللّغوي القاعدة الأساسية لاستخراج "الجذر".

– في البداية يتم اختيار الكلمة المدخلة المراد تحليلها؛ ثم يشرع البرنامج في تطبيق خوارزميات المطابقة بين المادة المدخلة "الجانب اللّغوي" والمادة الصورية "ف ع ل"؛ وفي مرحلة لاحقة يتم تفعيل الوزن/ الميزان عن طريق خوارزميات الإقحام التي تقوم بإدراج الزوائد (السوابق واللّواحق) لتفكيك (تقطيع) الكلمة المدخلة بهدف "تحليل" الكلمة إلى بنيتها الأساسية أي "الجذر"[5].

1 –Gérard SABAH, Intelligence Artificielle et langage, Processus de compréhension T 2 hermes. 1990.p.4.

2– يحي هلال، التوليد من الجذور والوزن، مجلة التواصل اللساني، م1 . ع1. سنة 1992. ص. 77.

3– محمد الحناش، اللغة العربية والحاسوب،(قراءة سريعة في الهندسة اللسانية العربية)أو مقربة في محاكاة الدماغ العربي لغويا، منشورات موقع جامعة الإمارات العربية المتحدة، كلية العلوم الإنسانية والإجتماعية، قسم اللغة العربية وآدابها، أكتوبر (2002)، ص2 .

4– الميزان الصرفي هو وسيلة علمية دقيقة تمكن الدارس من تمثيل بنية الكلمة ووصفها من حيث حروفها وحركاتها وزوائدها وصفًا يجمع بين الدقة والإيجاز فلفظ "فعل" الذي يمثل الميزان الصرفي وُضع "ليكون محلاً للهيئة المشتركة" بين الكلمات. للمزيد من التفصيل ينظر الباب الثاني ، الفصل الأول من الأطروحة، ص 199 وما بعدها، وينظر: لطيفة إبراهيم النّجار، "دور البنية الصرفية في وصف الظاهرة النحوية وتقعيدها"، دار البشير، عمّان، الأردن، الطبعة الأولى، 1414– 1994، ص 35. للمزيد من التفصيل في التعريف بالمصطلح ، ينظر الباب الثاني من الأطروحة /الفصل الأول ، ص190 وما بعدها.

5– يمثّل الجذر دخل البرنامج اللغوي في كفاية المتكلم، إذ بعد اختياره تشرع الكفاية اللسانية، أو بالأحرى المنظومة المعرفية المتعددة، في تطبيق خوارزميات التطابق بين مادة الجذر اللغوي والمادة الصورية (ف ع ل) وفي مرحلة لاحقة يتم تطبيق خوارزم الانصهار الذي يقوم بإدراج الزوائد في البنية النظرية للجذر وذلك بهدف توليد الكلمة، أما في مرحلة التحليل فإن العملية تتم عبر نوع آخر من الخوارزميات تتولى تحليل المفردة إلى بنيتها الأساسية في الجذر. للمزيد من التفصيل في التعريف بالمصطلح ، ينظر الباب الثاني من الأطروحة /الفصل الأول ، ص 204، وما بعدها.

– لكل صيغة صرفية مقابل دلالي مخزن في البرنامج وهذا المكون الدلالي الصرفي هو الذي يضمن ربط المستوى الصرفي بالمعجم والدلالة، وتؤمن ظاهرة الانصهار التي تنفرد بها اللّغة العربية حرية الحركة للكلمات داخل الجملة، ويظهر ذلك بجلاء من علاقة المعجم بالدلالة[1]، ذلك أن لكل صيغة صرفية مقابل دلالي مخزن في الكفاية، كما أن علاقة الصرف بالمستوى الصواتي لا تحتاج إلى دليل وخاصة فيما يتعلق بالأصول المعتلة[2].

– اعتمدت المقاربة على الجذور الثلاثية لأنّها تتميز بمرونة واسعة في التحرك داخل البنية اللّغوية خلافًا للجذور الرباعية والخماسية.

– أما في مجال تركيب الجمل يبقى الفعل بمثابة دالة رياضية يتحكم في جميع العناصر الصورية المتغيرة كما في الصوغ التالي[3] :

ج ⇦ ف (ن 1، ن 2،...ن س)

وقد بني النظام الآلي للمعجم الإلكتروني للمعالجة الآلية للغة العربية على هذا الأساس[4]، نظرًا للطبيعة الجبرية لنظام اللّغة العربية المتمثلة في منظومة الخوارزميات اللّسانية التي تجمع بين مختلف مكونات النظام، فالربط بين هذه المستويات أمر بالغ التعقيد لا يمكن أن يقوم به الحاسوب، إلاّ إذا تمّ تزويدها بقاعدة معطيات ترتكز على قواعد لسانية صورية؛ أي بالخوارزميات اللّغوية بنوعيها "التوليد" و"التحليل".

– إذا أجرينا مقارنة بين العمل اللّساني والعمل الحاسوبي نجد هناك تشابه في خطوات العمل وفي طبيعة النتائج التي يتوصل إليها الباحث في كلاّ المجالين وهي تسير وفق ما يلي:

اللّغة الطبيعية ⟸ ترجمة الخوارزميات ⟸ بيانات مشفرة

إن المنهج الذي نطمح إلى تطبيقه على نظام اللغة العربية يستعير من تقنيات البحث الرياضي نظرية الرسوم (théorie des graphes) وجبر المجموعات (algèbre des ensembles)[5] فبفضلها تبنى قواعد مضبوطة للنماذج فكل رسم هو عبارة عن مجموعة من الأزواج (أ ، ب) ليوظفها في دراسة النظام اللغوي بهدف الصورنة والتصنيف وبناء النماذج، وسنرى كيف أن

1– تطرح الدلالة خصوصية كبيرة من داخل المعجم والصناعة المعجمية على وجه الخصوص وقد أثبتت النظرية المعرفية مدى أهمية هذه الخصوصية عند تحديد المعايير الدلالية لمواد المعجم الرئيسية وهذا جزء م الخلفية الكبرى التي نهدف إلى وصفها في إطار هذا البحث (للمزيد حول علاقة المعجم بالدلالة مقال د. ليلي المسعودي، البعد الاجتماعي لإشكال الدلالي في المعجم الثنائي فرنسي– عربي، مجلة الدراسات المعجمية، العدد الثاني، يناير 2003.)

2– عز الدين غازي: نظام قواعد معرفة صرافي – صواتي للغة العربية: مقاربة الفعل، رسالة دبلوم الدراسات العليا، كلية الآداب 1 فاس 1999. ص197 – 183.

3– الحناش محمد: اللغة العربية والحاسوب، ص. 5.

4– يأتي تفصيل ذلك في الجانب التطبيقي من الباب الثالث (الفصل الثاني).

5– محمد الحناش، البحث اللساني بين العمق والعقم "سفر التهافت نموذجا" مجلة دراسات أدبية ولسانية ، ع 4 صيف / خريف 86، ص 126.

العالم موريس كروس (M.Gross)[1] خلص إلى القول بأن كل صنف من المداخل المعجمية التي يبنيها بالطريقة المحددة سلفاً، يشكل جوهر معادلة لوغاريتمية من نوع (Log N2) وعلى الباحث الجاد أن يبحث عن حل لهذه الصيغة الرياضية في ميدان البحث الذي ينجز فيه فرضياته التأليفية، فالخطوة الأولى تمثل قاعدة معطيات قوامها لغة صورية تتبع تركيبًا معيناً في بناء قاعدة بيانات، وفي الخطوة الثانية تتم ترجمة الخوارزميات كلغة برمجة مقدمة لتحليل قاعدة المعطيات باللّغة الطبيعية، ثم تأتي المرحلة الثالثة اللّغة الطبيعية التي تعدّ النتيجة التي يتم من خلالها التأكد من صحة البرنامج.

ومن جهة أخرى، يعتبر التراث المعجم العربي سباقا إلى العمل في وضع المعاجم بعد جمعها وفق رؤية رياضية تنطلق من النظام الجذري وذلك مع الخليل بن أحمد الفراهيدي[2] الذي ألف أول معجم عربي كان له الفضل في تأسيس قاعدة لغوية معجمية تعتمد الإحصاء اللغوي مقدمة لإنجاز معجم عربي مُستخلِصا مفاهيمه من النظرية الرياضية والنظام الصوتي والصرفي للغة العربية، إذ جعل من الرؤيتين نظامين للتعامل في ضوئهما تطبيقا لتأسيس أول معجم عربي على قاعدة رياضية إحصائية[3]، وعلى هذا الأساس تم التفكير في تطوير نظام آلي بهدف تحويله إلى معجم إلكتروني للمعالجة الآلية للغة العربية فقد كان للمستوى الصرفي في هذه المعالجة الآلية واقعية وعلمية هدفه تشجيع وتطوير أنظمة هندسة اللّغة العربية في جميع مستوياتها اللّسانية من خلال[4]:

- تمتين وتقوية نظام قواعد المعرفة الصرفية.
- تمتين وتقوية قواعد المعطيات اللّسانية.
- دراسة قدرة النظام وفق التصورات التجريبية.

فالمعالجة الآلية لا يمكنها إلاّ أن تتعامل مع الدقيق والمضبوط والمكتمل، لهذا تتطلب الكشف عن دخائل البنية الدفينة للغة العربية وتحديد خصائصها لمعالجتها آليًا[5].

1- المرجع السابق، ص 129.

2- كتاب العين هو أول معجم عربي لعبد الرحمان الخليل بن أحمد الفراهيدي (100-175 ه) كانت رؤيته رياضية في الجمع من خلال سعيه إلى تدوين مدونة تضم كلام العرب كله فترى أنه تعرض للمستعمل والمهمل، وفي الوضع حيث اعتمد تصنيف الأصوات لمعرفة العلاقة بين الأصوات في الكلمة المفردة من حيث تقاربه أو تباعدها وهدفه أن يستقر الترتيب في معجمه على منوال الطبقات الصوتية مما جعله عالما بخبايا الموسيقى، وفي الترتيب اهتدى إلى فكرة التقاليب في بناء الثنائي والثلاثي والرباعي والخماسي كما ذكر في مقدمة العين " اعلم أن الكلمة الثنائية تتصرف على وجهين نحو: قد، دق، شد، دش والكلمة الثلاثية تتصرف على ستة أوجه وتسمى مسدوسة...والكلمة الرباعية تتصرف على أربعة وعشرين وجها...والكلمة الخماية تتصرف على مائة وعشرين وجها... " وترتيبه للمنشقات مرتب بحسب مخارج الحروف، وأما في شرحه للمعنى المعجمي فاستخدام الرجل عدة طرق من بينها الشرح أو التعريف بالضد والتفسير بالمغايرة والشرح بكلمة واحدة وشرح المعنى بأكثر من كلمة.لمزيد من التفاصيل أنظر د. حلمي خليل، مقدمة لدراسة التراث المعجمي العربي، دار النهضة العربية، طبعة أولى 1997. ص 124-150.

3- قد ذهب عبد الغني أبو العزم إلى اعتبار أن الخليل كان مهموما بقضايا أصوات الحروف وكلام العرب في غياب أية مرجعية منهجية سابقة عليه ولم يكن أمامه إلا إعمال العقل لربط اللغة بأصواتها وترتيب الكلمات في سياق النسق التدريجي لمخارج الحروف ،وهذا الترابط بين الصوت والكلمة قائم على وجود علاقة بين التصور المنهجي لهذا الرجل وكلام العرب في تطوره التاريخي باعتباره المدخل لإمكانية حصر مفردات اللغة افتراضياً ،للمزيد انظر مقدمات المعاجم العربية القديمة، رؤية معجمية أم مذهبية ؟ عبد الغني أبو العزم، وقائع الندوة الدولية ، المعاجم العربية الواقع والآفاق، أيام 27/26/25 سبتمبر 2003 بالدار البيضاء .

4- نبيل علي، اللغة العربية والحاسوب، سابق، ص 12، 13.

5- المرجع نفسه، ص 03.

لقد اتخذت هذه الدراسة من اللّغة منطلقًا أساسيًا لتأتي المعالجة اللّغوية الآلية كركيزة تقام عليها تطبيقات النّظم المختلفة للمعلومات، وأصبحت وسيلة لتقديم وجهة نظر تتعلق بإعداد قاعدة معطيات موجهة لبناء معجم إلكتروني باستخدام تقنيات معلوماتية حديثة لمعالجة اللّغة العربية آليًا في المستوى الصرفي.

فلا مفر من التعامل مع خصائص اللّغة العربية في علاقتها بالحاسوب وأنّه سيمضي وقت طويل قبل أن تصير النّظم الآلية ذات "شفافية" لغوية عالية بالقدر الّذي تختفي معه الفوارق اللّغوية ظاهريًا، بحيث لا يدري المستخدم شيئا عما يقوم به الحاسوب في استيعاب لغته الخاصة خلال معالجتها آليًا.

فالتقدم الهائل في مجال اللّسانيات الحاسوبية يحاصر تأخرنا في هذا الميدان، فقد برزت الدوافع وتوفرت الأساليب اللازمة لوضع علاقة "العربية" بالحاسوب من أولويات البحوث العلمية.

3-2- **الاطراد الخوارزمي[1] ونظام اللغة العربية:**

يعتبر نظام اللغة العربية أكثر اطرادا من ضمن العديد من اللغات العربية الطبيعية وربما لا تضاهيه في ذلك أي لغة أخرى[2]. ذلك أن التقارب الذي حصل بين اللسانيات والمعلوماتية أقر بشكل كبير الاستفادة المتبادلة بواسطة أساليب النمذجة المطبقة على جميع العلوم. مما أدى إلى معالجة الظواهر اللغوية بأدوات عقلانية متطورة تمثلت في الجانب الخوارزمي الصرف الذي توفره المعلوميات النظرية والتطبيقية، في معالجة الظواهر اللسانية تنظيما وأيللة. لأن ذلك هو السبيل الوحيد لوصف وتجريب مختلف المواقع النظرية والمنهجية للظواهر اللغوية مما يؤدي إلى صياغة قواعد تجريبية عملية هي من صميم علم اللسانيات.

وهذا ما يجعلها أيضا ذات قدرة عالية على وصف اللغات الطبيعية وعلى تطوير تطبيقات معالجاتها.

إن استخدام النماذج الصورية الآتية من نظريات علوم الحاسوب والرياضيات واللسانيات قد منحت القدرة على إبداع قوالب وطرائق تتلاءم مع أنواع الحواسيب وأنظمة القواعد الصورية والمنطق والنمذجة الاحتمالية للغات الطبيعية[3]، كل هذه التطبيقات تؤطرها

1- الاطراد الخوارزمي: ويقصد به التزام منهجية منضبطة وثابتة في التعامل مع المداخل اللغوية المتشابهة، من حيث كم المعلومات، ونوعيتها، وكيفية ورودها في الخوارزمية .

2- محمد الحناش، اللغة العربية والحاسوب،(قراءة سريعة في الهندسة اللسانية العربية)أو مقربة في محاكاة الدماغ العربي لغويا، منشورات موقع جامعة الإمارات العربية المتحدة، كلية العلوم الإنسانية والإجتماعية، قسم اللغة العربية وآدابها، أكتوبر (2002)، ص2 .

3- تلعب النماذج اللغوية دورا أساسيا في تحديد وظيفة أنظمة معالجة اللغات الطبيعية في جميع تطبيقاتها الحاسوبية كالترجمة الآلية والتعرف الآلي على الكلام الطبيعي والبحث عن المعلومات وغيرها وتعمل هذه النمذجة على تكييف أساسيات المعطيات النصية ليس بالاعتبار الكمي بل بالكيفية التي تستغل بها هذه المعطيات وهذا ما يعرف بالتكييف الديناميكي للنماذج اللغوية حيث يمكن تقليص كمية المعطيات الأساسية للتعلم والتي يمكن أن تشمل جميع نماذج (n-grammes) وفق تخمين احتمالي للنماذج المخفية .(نماذج ماركوف مثلا) كما يمكن أيضا اعتماد قواعد المعارف كمنهجية غير رقمية في معالجة المتون النصية بطرق أكثر دقة ونجاعة. للمزيد ينظر: M.Jardino et M.EL-Béze ,Modélisation probabiliste du language naturel,(in T A L ,volume 44-n°1 /2003.p.7-8.

أعداد محدودة من الخوارزميات[1] والتي تعرف الآن أيضا بالبراديغمات الحاسوبية[2] والتي توظف في بناء خوارزميات[3] البحث في أوضاع الفضاءات وخوارزميات البرمجة النشطة التي تُرجمت بواسطة الأدوات الإجرامية والعقلانية المتمثلة في أنواع الأتوماتات والمحولات ذات الأوضاع النهائية ونماذج ماركوف المخفية بالإضافة إلى نحو السياق الحر وأنحاء الحالات ..وكل هذه الخوارزميات تعتمد نظرية الاحتمال في اشتغالها الصوري والتي تستطيع معالجة المستويات اللسانية كالدلالة والتركيب والمعجم والصرف والفونولوجيا.

فهذه الخورزميات[4] المتلائمة مع انظمة القواعد الصورية ومع أوضاع الحواسيب تستلزم طبعًا البحث عن فرضيات لفضاء حالات التمثيل المتمثلة في الدخل الخوارزمي ،فمثلا بالنسبة للمتواليات الفونولوجية يجي اعتماد الكلمة كدخل في نظام التعرف على الخطاب، أو البحث في الشكل الشجري بالنسبة للمعالجة التركيبية، كل ذلك بواسطة خوارزميات الرسوم والأوتوماتات والمحولات المناسبة لكل عملية صورية على حدة، إن نظرية الاحتمال التي تتحكم في هذه الخوارزميات هي العنصر الذري في مجموع هذه التقنيات التي يراد بها معالجة وتحليل المعرفة اللسانية[5] ، فكل النماذج الخاصة بأنظمة القواعد الصورية ومنطق الدرجة الأولى والبحث الأولي المعمق قابلة للتطوير بواسطة نظرية الاحتمال، وما يواجه في الظواهر اللغوية هو كيف يتم معالجة مسألة اللبس[6] في جميع مستويات معالجة اللغات الطبيعية[7]، بمعنى، إعطاء الفرصة

1– خوارزمية: Algorithme: مجموعة من القواعد والعمليات الحسابية والرياضية المتسلسلة على شكل خطوات، يؤدي تنفيذها إلى الوصول إلى النتيجة المطلوبة.

2 - James Martin et al ,Speech and language processing , An introduction to naturel language processing, compititional languistics and Speech recognition ,second edition ,Spring 08.p.5 and 6.see,www.cs.colorado.edu

3– الخوارزمية: لغة جبرية نستطيع بواسطتها برمجة الحسابات الرقمية والخوارزميات بشكل سهل وبسيط مثلاً: لغة Algol.

4– خوارزمية: Algorithmique: لوصف طريقة في الحساب أو في عمل الحاسوب أو أي جهاز آخر ترتكز فيه العمليات على عدد من الخطوات (خوارزم) يسمح بالوصول إلى النتيجة المطلوبة.

5 –James Martin et al ,Speech and language processing , An introduction to naturel language processing , compititional languistics and Speech recognition ,second edition ,Spring 08.p.6

6– إن من خصائص أي لغة الوضوح والإبانة، لكن قد يحدث أن يعتريها اللبس في مستوياتها المختلفة (الصوت، والصرف، والنحو، والدلالة)، وما من لغة في العالم إلا تشتمل على تراكيب صالحة لتعدد المعنى والدلالات.

ومن هنا تأتي أهمية معالجة اللبس أو الغموض في المحتوى العربي على الإنترنت، ودراسة أسبابه وأنماطه وآثاره المختلفة، ومعالجة مشكلاته لاسيما في اللغة العربية المعاصرة، ثم الاستفادة بذلك في برامج معالجة اللغة العربية آلياً، كبرامج الترجمة الآلية ومحركات البحث التي تتعامل مع المحتوى العربي على الإنترنت، وكذا برامج الإحصاء اللغوي.

7– من وسائل معالجة اللبس أو الغموض التركيبي (من خلال تطبيقات معالجة اللغات الطبيعية): القواعد النحوية لفك اللبس ومنها:

أ– التضام: (ويشمل الافتقار والاختصاص والاستغناء والمناسبة المعجمية بين المفردات) ... ب– الرتبة: وقد تكون محفوظة كرتبة الحروف والأدوات من مدخولاتها، وقد تكون غير محفوظة كرتبة الفعل والمفعول به، ورتبة المبتدأ أو الخبر. ج– الربط (إما بالمطابقة وإما بالإحالة): والربط قد يكون بالمطابقة في العدد (الإفراد والتثنية والجمع)، والشخص (المتكلم والمخاطب والغائب)، والنوع(التذكير والتأنيث)، والتعيين (التعريف والتنكير) وفي الإعراب. وقد يكون بالإحالة. والأصل في الإحالة إعادة الذكر. د– البنية: (وتشمل أقسام الكلم والصيغ الصرفية ومعانيها والأدوات والإجراءات التصريفية). وتتناول قرينة البنية دلالات أقسام الكلم والصيغ الصرفية (مبنى ومعنى) وإجراءات تحولاتها، وكذلك الأدوات وحروف المعاني ووظائفها في السياق. للمزيد من التفصيل: ينظر: تمام حسان: "تطوير التأليف في مجالات اللغة العربية"، بحث منشور بمجلة مجمع اللغة العربية.

للخوارزميات أن تقوم "باختبارات متعددة في عملياتها المحتملة حينما تواجه مدخلات غامضة، فتختار المناسب عند الخرج (output)[1] .وما هذا سيكون ايجابيا في بناء نماذج آلات التعلم (machine learning models) بحيث ستتمكن من تمثيل جميع المعارف والمعلومات الموصوفة بواسطة اوتومات وأنظمة القواعد الصورية والبحث الحدسي والمحولات بجميع أصنافها .ان هذه التكنولوجيا قابلة للتجريب وللتطبيق على قواعد بيانات معرفية لغوية باعتبارها تقنيات قوية وصلبة للنمذجة .

انطلاقا من هذه الرؤية المبدئية يتبين ان أدوات الاشتغال متاحة حاليا، في معالجة اللغات الطبيعية، وقد صممت تقنيات أكثر ملائمة للمستوى المعجمي الذي نحن بصدد دراسته[2] :

تقنيات ومناهج وخوارزميات خاصة بعمليات معينة.

نماذج وأشكال صورية تستغل في المعالجة والحساب .

المعطيات اللغوية الخاصة ومنها الأدوات اللسانية كالأنحاء والقواميس الإلكترونية وقواعد البيانات النصية الموسعة .

هذا، وفي السنوات الأخيرة أصبحت هذه المعطيات اللغوية تشكل تحديا كبيرا أمام أنظمة معالجة اللغات الطبيعية .ذلك أنها الكفيلة في تحديد العمليات التطبيقية الملائمة ومنها خوارزميات وبرامج المعالجة المبنية وفق خاصيات مجردة وعقلانية ، مرتكزة على مفاهيم علمية في تصنيف الظواهر اللغوية، هدفها استعاب معطيات جيدة وقادرة على استقطاب الأدوات المعلوماتية والعلمية[3] لبناء أنظمة إلكترونية مؤسسة على القواميس الإلكترونية وعلى أنظمة متون مبنينة وموسعة لتستجيب للكفاية المعجمعية الطبيعية وذلك باحتواء الظواهر اللغوية.

رابعا- خوارزميات توليد و تحليل المفردات العربية :

تمتاز اللغة العربية، شأنها شأن اللغات السامية، بقوة التوليد الصرفي[4]، وفي هذا المبحث نقدم تصورا نعتقد أنه جديد في نظرته إلى الدلالة الصرفية المبنية أساسا على التوليد الخوارزمي الصرفي ويعتمد هذا التصور على مستويين من الدلالة الصرفية، يمكن التمييز بينهما عند تحليل دلالة التوليد، أولهما الدلالة المصاحبة للجذر، والثاني الدلالة المصاحبة

---

1 –James Martin et al ,Speech and language processing , An introduction to naturel language processing compititional languistics and Speech recognition ,second edition ,Spring 08.p.6 .

2 –Eric la porte ,mots et niveau lexicale ,in TAL ,vol n° 2 « Etat de l'art . 1997

3– وهذا ما بات يعرف الآن في الأوساط العلمية بالوعي بوجود (حالة ما) " goulot d'étranglement "انظر:

Eric la porte ,mots et niveau lexicale ,in TAL ,vol n° 2 « Etat de l'art ..1997

4– ابراهيم بن مراد، مقدمة لنظرية المعجم، دار الغرب الإسلامي، بيروت، الطبعة الأولى،1997، ص 143.

للخوارزميات الصرفية. إن تفاعل هذه المستويات يولد مفاهيم متكاملة، يمكن استيعابها على هيئة مجموعات توليدية أو مشتقات عربية.

ولعل الغرض الأساس من هذا المنظور إلى الدلالة الصرفية يتمثل في التعرف على سبيل الربط بين المفاهيم الدلالية والمفردات العربية، فيما يصطلح على تسميته في علم توليد اللغة بالتوليد الخوارزمي الصرفي للمفردات العربية[1].

4-1- الخوارزميات من الدلالة اللغوية إلى الدلالة الاصطلاحية:

قبل أن نتكلم عن الخوارزميات الصرفية في اللغة العربية لا بأس من الوقوف عند مصطلح الخوارزمية نفسه (Algorithme أو Algorithm)[2] والذي لم يعط له ما يستحق من العناية والاهتمام في الأدبيات اللسانية الحاسوبية العربية. إذن ما هي أصوله الإتيمولوجية، وما دلالته، وما معناه في اللسانيات والمعلوماتية ؟

إن الخوارزمية أو الخوارزميات ليست مأخوذة من أية كلمة لاتينية أو يونانية وإنما هي ذات أصول عربية صرفة، وهي تقترن إيتيمولوجيا باسم العالم الرياضي العربي محمد بن موسى الخوارزمي الذي ألف كتابين مهمين: الأول سماه " علم الحساب" والثاني عنوانه "الجبر والمقابلة". وبمرور ثلاثة قرون تقريبا ترجم مصطلح الخوارزمية إلى اللاتينية تحت اسم (Algorismus). وقد نقل المصطلح من مجاله الرياضي الصرف ليستخدم في ميدان المعلوماتية على مستوى بناء البرامج الحاسوبية[3].

4- 2- الخوارزميات الحاسوبية :

هي وسيلة إجرائية تتجلى وظيفتها في القيام بالحساب في الحاسوب، وأول تعريف للخوارزمية في العصر الحديث هو التعريف الذي قدمه العالم الروسي ماركوف: الخوارزميات هي القواعد الدقيقة، والتي توظف باعتبارها وسيلة إجرائية للحساب، وتمكن في نهاية المطاف من الوصول إلى نتيجة دقيقة انطلاقا من معطيات أولية[4].

1- Generation Algorithmique Morphologique des Mots Arabes (GAMA).

2 – عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، الجزء الأول، إشراف الدكتور عبد الغني أبو العزم، كلية الآداب والعلوم الإنسانية، عين الشق، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجمات، جامعة الحسن الثاني، الدار البيضاء،2009/2008، ص 105.

3– المرجع نفسه، ص 105.

4– مازن الوعر، اللسانيات والعلم والتكنولوجيا " نحو تعريب موحد للسانيات التطبيقية العربية وبرمجتها في الحاسبات الإلكترونية"، دورية اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 22، الرباط، المغرب، 1983/1982، ص 20.

وعلى هذا الأساس، إن صياغة خوارزمية معينة هي من قبيل الفعل الإبداعي والابتكاري، أي صياغة الأدوات المعرفية والرياضية والمنطقية التي تقبل البرمجة الحاسوبية ومن ثمّ تترجم إلى خطاطة وظيفية (Organigrame) أي مخطط صندوقي.

إن الخوارزميات المعلوماتية تقوم بعدة وظائف من قبيل جمع المعلومات وجمعها في الذاكرة المركزية في الحاسوب، والبحث السريع عن المعلومة، ومعالجة المعطيات، وترجمة لغات البرمجة، والحساب الصوري وغيرها.

4- 3- الخوارزميات لسانيا:

لقد عرفت الخوارزميات في اللسانيات الحديثة تعريفات متباينة ومختلفة نختار منها:

يعرف موريس غروس الخوارزمية في النحو الصوري على الشكل التالي[1]:

تحدد اللغة الصورية: نص والمعجم: عج بأنها مجموع الكلمات أو الجمل التي تنتمي بشكل حر إلى الزوج:[2] (أو عج*) المولد بواسطة أ (أ بواسطة عج) حيث تمثل مجموع الكلمات (أو الجمل) المقبولة.

إذن، تمثل (ل) على النحو الآتي: خو* / ل.

حيث ترمز الخطاطة إلى مجموعة الجمل الملغاة.

وتعني الخوارزميات عند البعض تعاقبا محددا من التعليمات الموصوفة جيدا، مع توفر الخصائص التالية[3]:

لا يوجد غموض في أية تعليمية.

بعد تنفيذ تعليمية معينة، لا يوجد غموض في تحديد التعليمية التالية.

يجري بلوغ تعليمية انتهت ذاتها بعد تنفيذ عدد من التعليمات.

فيما عرفها آخرون على النحو التالي[4]:

الخوارزميات هي مجموعة من القواعد المنظمة في طريقة معينة، تنطلق من القاعدة البسيطة إلى القاعدة المركبة، ثم إلى القاعدة الأكثر تعقيدا. إنها منظومة من القواعد الصورية المبرمجة في الدماغ البشري، وتقوم بوظيفة مراقبة بالنسبة للغة على مستويي الإدخال والإخراج.

1− Maurice Gross, André Lentin, Notions sur les grammaires formelles, collection « programmation » . Gauthier Villers, Editeurs, Paris 6/1970, p44.

2 − Maurice Gross, Opcit.

3− عبد الإله ديوه جي، مفاهيم أساسية حول تقنية المعلومات، مجلة عالم الفكر، م18. ع.3، 1987.

4− عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، الجزء الأول، إشراف الدكتور عبد الغني أبو العزم، كلية الآداب والعلوم الإنسانية، عين الشق، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجمات، جامعة الحسن الثاني، الدار البيضاء،2009/2008، ص 107.

من خلال التعريفات السابقة الذكر، نستنتج أن الخوارزميات اللغوية هي القوانين والمبادئ العقلانية المتمركزة في منطقة معينة في دماغ الإنسان، والتي بموجبها يستطيع المتكلم اللغوي إنتاج اللغة وفهمها وتوليدها وتحليلها. ونعتقد أن هذه الآليات الصورية قد لعبت دورا كبيرا وأساسيا في الميدانين الرياضي المعلومياتي فيما سبق، ومن المؤكد أن يكون لها مستقبلا دور طلائعيا ورائدا في مجال المعالجة الآلية للغات الطبيعية وهندستها[1].

4- 4- خوارزميات التوليد والتحليل:

تقوم المعالجة الآلية للغات على ركنين أساسيين وهما التحليل والتوليد، اللذان يؤطران عملية الإنتاج اللغوي[2]، وضمن هذا التصور ينظر إلى الدماغ على أنه آلة لها مدخل ومخرج، ففي المدخل (التحليل) توجد قواعد البيانات التي يتمرس عليه الإنسان، فتتولد لديه المعرفة اللسانية التي ستبنى عليها الآلة الخوارزميات اللغوية التي تضبط اللغة في الكفاية وتتحكم فيها تخزينا واسترجاعا. وفي المخرج (التوليد) تتم عملية إنتاج اللغة الإنجازية مفردات وجمل ونصوصا يؤطر ذلك كله الترسيمة التالية[3]:

المدخل	←	الدماغ / الآلة	←	المخرج
Input		**Cerveau / Machine**		**Output**
التحليل		**قاعــدة المعطيات**		**التوليد**
Analysis		**Data Base**		**Génération**

4- 5- خوارزميات التحليل الصرفي:

يعتبر برنامج المحلل الصرفي القاعدة الأساس التي تبنى عليها مختلف التطبيقات الحاسوبية[4] المتعلقة باللغة العربية من قبل أنظمة التصحيح الإملائي والنحوي، والخوارزميات ضغط النصوص وتشفيرها، وبرامج اللغة العربية لغير الناطقين بها ورفع كفاءة محركات البحث، ونظم الاسترجاع، ونظم الترجمة الآلية الثنائية الاتجاه[5].

1- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، سابق، ص 107.

2- أنيس فريحة، نظريات في اللغة، دار الكتاب اللبنانب، بيروت، لبنان، الطبعة الثانية، 1981، ص 70، 71.

3- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، سابق، ص 107.

4- يمكن تصنيف التقيات المتعلقة بالتحليل الصرفي إلى أربع مجموعات هي: أ / خوارزميات التحليل الصرفي القاموسي، ب/ خوارزميات التحليل الصرفي اللغوي، ج/ خوارزميات التحليل الصرفي التبادلي، د/ خوارزميات التحليل الصرفي المصدري.

5- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، سابق، ص 108.

ومن المؤكد أن المعالجة الحاسوبية للصرف العربي ليست جديدة كل الجدة وإنما هنالك دراسات وأبحاث مهمة وأساسية، نذكر منها على سبيل الإشارة: بيزلي 1996، بيزلي 1998، بيزلي 2000، كيراز 2000، الشلبي وأنجليز 1998، بيري زيدوم وعاطف 2001، درويش 2002 وغيرهم كثير، ولكن ما يسجل على بعض هذه الأعمال أنها لم تكن شاملة لجميع المستويات اللسانية، أو أنها همشت مستوى من المستويات لسبب من الأسباب، ولنأخذ على سبيل المثال الشلبي وأنجليز 1998، ركزا اهتمامهما على الأفعال وصيغها، وبيري وزيدوم 2002 الذين قدما نظاما حاسوبيا لا يستخدم المعجم. في حين اعتنى درويش 2002 بجذر الكلمة باعتبارها نتيجة للتحليل الآلي، وتحقيقا لهذا الغرض وظف القواعد اللسانية والإجراءات الإحصائية التي تساعده على معالجة المدونة، ويرتكز المحلل الصرفي في عمله على فحص الخصائص الصرفية والنحوية لقاعدة ضخمة من المفردات العربية المستقاة من تحليل ذخيرة لغوية ضخمة، ولا يعدّ المحلل الصرفي اللغوي تطبيقا نهائيا، ولكن أساس لمجموعة من التطبيقات، فهو يدخل في صلب تطبيقات التشكيل والتصحيح والفهم الآلي ونظم فهرسة المعلومات واسترجاعها، وكذلك تطبيقات التشفير وضغط النصوص. وأهم خاصية تميّز المحلل الصرفي اللغوي المبني على الذخيرة اللغوية بتغطية أكبر وقدرة سريعة على التوسّع وإمكانية تعامله مع النصوص القديمة والحديثة في مختلف التطبيقات الحديثة بالإضافة إلى صغر حجمه في الذاكرة وإمكانية تناقله بين مختلف أنظمة التشغيل.

خامسا- مقاربة التحليل الصرفي:

يقوم المحلل الصرفي بتحليل كلمات الجملة المدخلة إلى عناصرها الأساسية وتحديد بياناتها الصرفية وذلك بعد المرور عبر المراحل التالية[1]:

تجريد الكلمة من السوابق واللواحق ك (الـــ) التعريف والضمائر وغيرها.

استخراج الوزن الصرفي المقابل لكل كلمة.

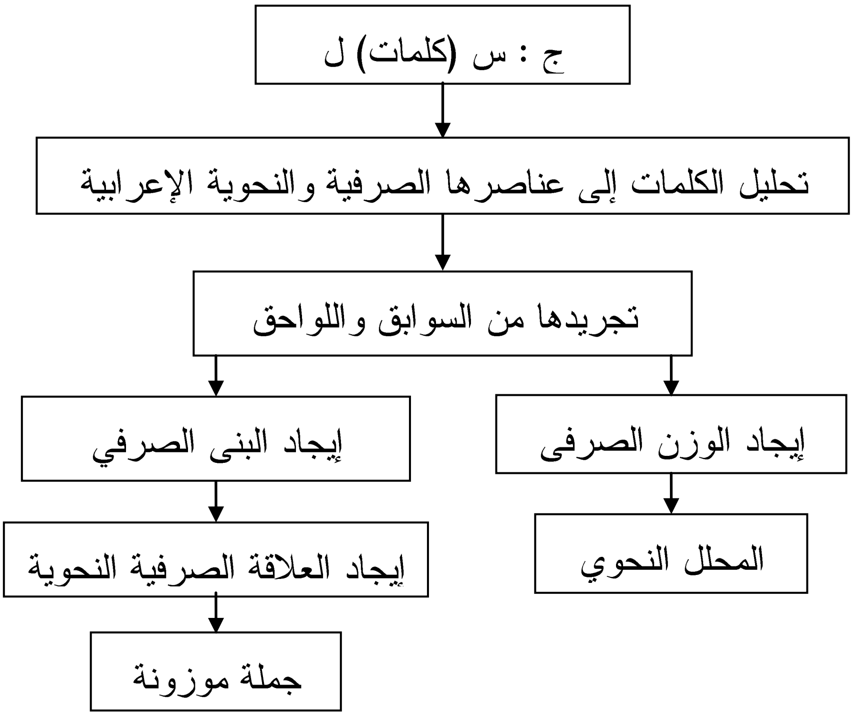
إيجاد العلاقة الصرفية النحوية بين مركبات الجملة وذلك طبقا للقواعد النحوية التي تحدد القالب الصرفي لكل كلمة[2].

وفيما يلي شكل موضح بالخطوات السابقة[3]

1- المرجع السابق، ص 109

2- للمزيد من التفصيل ينظر : عبد الفتاح حمداني، توفيق اليزيدي، التحليل الصرفي للأسماء العربية، أبحاث لسانية، المجلد 12، العدد 2/1، منشورات معهد الدراسات والأبحاث للتعريب، 2007، 37،38.

3- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، سابق، ص 109.



5-1- خوارزميات المحلل الصرفي:

وتنقسم خوارزميات التحليل الصرفي إلى أربعة أقسام رئيسية وهي:

خوارزميات التحليل الصرفي القاموسي (المعجمي)[1]

خوارزميات التحليل الصرفي اللغوي: وتهدف إلى تحديد نوع الكلمة وخصائصها الصرفية[2]، فمثلا كلمة (المحللون) تحلل على الشكل التالي[3]:

ــ الــ: سابقة أداة تعريف

ــ محللون: اسم جمع مذكر سالم.

ــ ون: لاحقة الجمع

ــ حلل: مورفيم الجذر

ــ محلل: ساق الكلمة وقبل تطبيق هذا البرنامج الصرفي، لا بدّ من استحضار العناصر التالية[4]:

1- مع بساطة التحليل الصرفي القاموسي ودقتها إلا أنها تتطلب قائمة ضخمة بجميع الكلمات العربية بأشكالها المختلفة مع جميع التحليلات لها و تتم عملية التحليل الصرفي عن طريق الوصول لسجل الكلمة في القائمة و استخلاص التحليلات الممكنة.

2- تتميز معظم الأعمال البحثيه في مجال التحليل الصرفي بأنها تتبع خوارزميات التحليل الصرفي اللغوي حيث تستفيد من القواعد التي تحكم توليد المفردات العربيةُ مشابه في ذلك عمل اللغوي المتخصص. مع دقتها العالية في التحليل إلا أن هذه العملية تتطلب توفر قوائم عديدة وضخمة.

3- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، سابق، ص 110،109.

4- المرجع نفسه، ص 110.

قائمة بجميع الجدزور اللغوية ثلاثية أم رباعية (قاعدة بيانات الجذور).

قائمة بجميع الوزان الصرفية الأساسية (قاعدة بيانات الخوارزميات الصرفية)

قائمة بجميع السوابق الممكنة (10- 11 سابقة) (مورفيمات الزيادة الصرفية)

قائمة بجميع الكلمات الجامدة، وهي أقل من 500 كلمة (وتضم الأسماء المشتركة واسماء المكنة، وأسماء الأعلام ...).

5-2- خوارزميات التحليل الصرفي التبادلي[1]:

يشتغل هذا البرنامج من خلال إيجاد مجموعة من التباديل من حروف الكلمة المدخلة، ثم يحاول التأكد من أن الجذر صحيح، ومن ثمّ التأكد من أن الحروف الباقية حروف زائدة، فمثلا كلمة (المرتبطون) فسيأخذ الحروف، ويجرب تباديلها، حيث سيجد (رتب) ولكنه (سيجد) من الحروف الباقية حرف الطاء، ولكنه حرف غير زائد، فسيعلم من أن (رتب) ليست الجذر، وسيأخذ (ربط)، وسيجد أن باقي الحروف كلها مزبدة، و(ربط) هي الجذر. لكن ما يعاب على هذه المنهجية أنها مكلفة بالنسبة للبرنامج المستخدم، حيث تثقل كاهله بمهمات متكررة، يستغرق في تنفيذها وقتا أطول، وتطول المدة أكثر، كلما زاد طول الكلمة عن أربعة أحرف.

5- 3- خوارزميات التحليل الصرفي المصدري[2]:

يرتبط المحلل الصرفي بمجموعة من التطبيقات الآلية نذكر منها على سبيل الذكر[3]:

المصححات الإملائية والنحوية ومولدات المفردات والجمل والنصوص العربية المشكلات الآلية للنصوص العربية، التي تعتمد في الأساس على الخواص الصرفية والنحوية التي ترمز، وكذلك على الإحصاءات المتوفرة من الذخيرة، وأنظمة الفهم الآلي وحفظ النصوص واسترجاعها، وتطبيقات التشفير وضغط النصوص، والذخيرة اللغوية وهي تشكل بعد تحليلها وترميزها نواة لكثير من الدراسات الإحصائية اللغوية المختلفة، كما تقوم بدور مساند لبعض التطبيقات التي تستدعي ذلك النوع من الإحصاءات مما يتصل بالتصحيح الإملائي والنحوي وتوليد الجمل، والنظام الخبير والذي سيقوم بالتعليم من الفريق اللغوي الذي يشتغل على تحليل وترميز الذخيرة

1- تعتمد خوارزميات التحليل الصرفي التبادلي على المحاولة وإعادة المحاوله حيث تستخدم في هذه الطريقة نافذة متحركة يتم من خلالها توليد كل تباديل الحروف (3 أو 4 حروف في المرة الواحدة) ومقارنتها بقائمة الجذور المتوفرة. وفي حالة التطابق يتم استخلاص الجذر والوزن الصرفي أما في حالة عدم التطابق فإنه يستخدم تبديل آخر للحروف. تحتاج هذه الطريقة لقوائم للجذور و الأوزان واللواصق ومعالجه طويله نسبياً.

2- أما في خوارزميات التحليل المصدري فتُعتبر محللات مصدرية وليست صرفيه ولا تصل في تحليلها لجذر الكلمة. وتستفيد من تشابه الكلمات المولدة في الشكل لمعرفة تحليلها, وتتميز بعدم حاجتها لقوائم بيانات كبيرة ولا معالجة معقده وبأنها قابلة للتطوير و مناسبه في إمكانية استخدامها كمرحلة أوليه للتحليل الصرفي.

3- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، سابق، ص 110.

اللغوية، واكتساب وتخزين عمله، وسوف يكون كذلك مساندا رئيسا عند الضرورة لتوسع حجم الذخيرة اللغوية، كما سيعمل على تقليل حجم الفريق اللغوي وتسريع عمله.

سادسا- خوارزميات التوليد الصرفي في بناء معجم إلكتروني للغة العربية:

تهدف تقنية التوليد الصرفي[1] إلى توليد المشتقات ومزايداتها انطلاقا من الجذر، مما يعني أن هذا النوع من المعالجة الصرفية الخوارزمية تقوم على عملية التحويل الصرفي، تبدأ بالجذر بوصفه مدخلا input، وتمر بالقناة الصرفية الخوارزمية المؤطرة بالقواعد الصرفية والصوتية، وتنتهي بالمخرج output أي المفردة المعجمية يؤطر ذلك كله الترسيمة التالية[2]:

مدخل ← القناة الصرفية الخوارزمية ← مخرج
قاعدة بيانات الجذور ← القواعد الصرفية والصوتية ← المفردة

إذن، لا بد لكل مولد أي مشتق اسما كان أم فعلا أن يمر عبر القناة الصرفية الخوارزمية، ولا يوجد عنصر لغوي لا يخضع لهذا الشرط التوليدي، والتوليد الصرفي يستطيع توليد جميع الأسماء القياسي منها والسماعي، المجرد والمزيد، الصحيح والمعتل.

وينقسم التوليد الصرفي إلى نمطين[3] :

6-1- التوليد الصرفي المعجمي:

وهو يتوخى توليد المفردات المعجمية الصحيحة والمعتلة، الثلاثية والرباعية، والمفردة في ضوء هذا النمط من التوليد الصرفي، تكون مجردة من اللواصق[4].

6-2- التوليد الصرفي النصي:

وهو يهدف إلى توليد المفردات كما هي في النص اللغوي الطبيعي، والمفردة في هذا النمط تكون ملصقة بالسوابق على اليمين، وباللواحق على جهة اليسار، ويوظف هذا النوع في التطبيقات الترجمة الآلية[5].

1- للمزيد من بسط القول ينظر : يحي هلال، بنوك المصطلحات والتطبيقات الحاسوبية للمعالجة الآلية للغات الطبيعية،اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 49، الرباط، المغرب، 1421هـ/2000م، ص 154.

2- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، سابق، ص 111.

3- يحي هلال، العلاج الآلي للعربية مع تطبيقات، مجلة التواصل اللساني، مج2، ع2س 1990، ص 26- 27.

4- يحي هلال، بنوك المصطلحات والتطبيقات الحاسوبية للمعالجة الآلية للغات الطبيعية،اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 49، الرباط، المغرب، 1421هـ/2000م، ص 154.

5- المرجع نفسه، ص 154.

6- 3 - التحليل المعجمي – التركيبي وبناء المعاجم الآلية للغة العربية:

مما لا شك فيه أن التطبيقات المتطورة للتكنولوجيا المتقدمة قد أبانت عن جدارتها في معالجة النصوص واللغات الطبيعية، باعتبارها أنظمة رمزية. نجد هذا متاحا عبر مجموعة من التقنيات الحاسوبية من نشر إلكتروني وإخراج آلي وتدقيق إملائي وتصحيح آلي وطبع ... إلى غير ذلك من التقنيات المتطورة على شكل أقراص صغيرة الحجم تشكل معطيات أساسية وجمة للمعلومات، وهذا ما سيمكن المتخصصين لا محالة في تطوير أنظمة التوثيق والمعالجة آليين للغات الطبيعية في سائر مستوياتها[1].

إن التحليل الآلي لنصوص منهجية تشبه إلى حد بعيد أساليب مجموعة من المعالجات والمحللات الآلية الأخرى كتقنية الضغط وتسنين النصوص والمعالجة الإحصائية وتجميع المعلومات وغيرها، فضمن العمليات الأولية في معالجة النصوص تلك المتعلقة بتقطيع ملفات المعطيات النصية، أي التعرف على الوحدات الأساسية للنص التي تتكون من متواليات ووحدات مفردة أو مركبة، وهذه المراحل تسمى **بالتحليل المعجمي**[2]. فإن أردنا تحليل النص الآتي: **كثف العدو غاراته،** فإننا نطلب أول الأمر من المحلل المعجمي التعرف على المكونات الثلاثة للجملة وهي " كثف" و " العدو" و " غاراته" غير أن عملية التعرف هنا قد تبدو مُبْهَمَة وغامضة وللوظيفتين التركيبية والدلالية على مستوى هذه البنية دور فعال في فرز الوحدات المعجمية[3].

وفي الحقيقة فإن البرنامج الآلي للتحليل المعجمي لا يمكنه في أول وهلة أن يتعرف ويحدد بدقة تصريف الفعل "كثف" مثلا أهو مضارع أم ماضي؟ فنطلب من المحلل المعجمي أن يقدم لنا هنا مجموعة الإمكانيات المعجمية المتوفرة لديه، ولحل هذه المسألة يطلب المحلل التركيبي أو الدلالي إن أمكن لبسط سياق كل كلمة على حدا[4].

1- عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، أطروحة جامعية لنيل الدكتوراه، إشراف الأستاذ الدكتور: عبد الغني أبو العزم، والأستاذ الدكتور: محمد لهلال، كلية الآداب والعلوم الإنسانية، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجميات، جامعة الحسن الثاني، عين الشق- الدار البيضاء، السنة الجامعية:2007/2008، ص196.

2- المرجع نفسه، ص 196.

3- المرجع نفسه، الصفحة نفسها.

4- المرجع نفسه، الموضع نفسه.

إن المشاكل المطروحة هنا هي من قبيل ما مواجهة المحلل المعجمي عند تطبيقاته المعلوماتية، ولاسيما فيما يتعلق بنظام تجميع المعلومات. فهذه القضايا عُولجت كثيرا وقد أمكن مقاربتها بنظرية أكثر عمقا ألا وهي نظرية الأوتومات[1] بأشكالها المختلفة[2].

إن اعتماد نفس الأدوات المعلوماتية واللسانية عند تطبيق المحلل المعجمي على نصوص لغة ما مسألة في حد ذاتها ناجحة، بيد أن التعرف على المفردات المركبة ، يعتبر من العوائق المثيرة للبحث لأنه مدعى لما قد يظهر خلال عملية التحليل والمعالجة الآلية التي تستدعي أدوات وأساليب عقلانية أكثر دقة[3].

وهكذا تعد المعالجة الصرفية الآلية مقوما أساسيا في تحليل النصوص العربية، واسترجاعها، وكذلك في عمليات الإعراب الآلي للجمل العربية، كما يعد المجال نفسه ميدانا خصبا للبحث، يتقاسمه علماء اللسانيات، وعلماء الحاسوب، وعلماء النفس، والذكاء الاصطناعي، ورغم حداثة هذا المجال، فقد تمكنت لغات أجنبية كالإنجليزية والفرنسية من تطوير برامج مهمة وأساسية نذكر منها على سبيل الإشارة واجهة قواعد البيانات، والتوليد الآلي للحكايات والوثائق والنصوص، والترجمة الآلية، والتعليم بواسطة الحاسوب، ونمذجة العمليات السيكولوجية وغيرها.

أما في مجال التجربة العربية، فقد أنشأت تطبيقات توليدية وتحليلية للعربية في مراكز بحثية على امتداد العالم العربي بالمغرب وسوريا والسعودية والكويت ومصر والإمارات العربية المتحدة، أقل ما يقال عنها أنها لم ترق إلى المستوى المطلوب، ولم تحقق بعد النتائج المرتقبة لسانيا / صوريا وحاسوبيا.

1- إن بناء نظام للتحليل المعجمي يجب أن يتأسس ويرتكز على قواعد معطيات معجمية كبيرة، تحتوي على أكبر عدد ممكن من الأوتومات لتوصيف الظواهر اللغوية، في حين تتطلب مجموعة من الأشكال النحوية إذلال العقبات والقيود الجديدة في المعالجة. لهذا السبب تم توظيف أوتومات الأوضاع النهائية لتحليل الظواهر التيبولوجية (النوعية) والإعرابية والتصريفية والإملائية. وفيما يلي توضيح لمكونات الأوتومات المتمثلة في العقد المساعد كما في الشكل التالي الذي يمثل متغيرين إملائيين لمفردة واحدة نحو كلمة (فعل) "دعى". للمزيد من التفصيل انظر Abdehadi Soudi, A computational Lexem – based Treatment of arabic Etat thesis in Linguistics, faculty of letters and human 'morphology, Doctora d sciences, med v yniversity 2002 p. 10.

2- للمزيد من بسط القول ينظر :الباب الثاني، المعالجة الآلية للغة العربية لبناء قاعدة البيانات المعجمية آليا، الفصل الثاني، هندسة المعجم العربي الإلكتروني، ص 271.

3- محمد الحناش، اللغة العربية والحاسوب،(قراءة سريعة في الهندسة اللسانية العربية)أو مقربة في محاكاة الدماغ العربي لغويا، منشورات موقع جامعة الإمارات العربية المتحدة، كلية العلوم الإنسانية والإجتماعية، قسم اللغة العربية وآدابها، أكتوبر (2002)، ص3 .

سابعا- أهداف الدراسة:

إنّ الأهداف التي نرجو تحقيقها في هذا العمل هي:

– التعرف على النماذج اللّسانية التي تحدد ملامح البنية اللّسانية الحاسوبية للمعجم المحوسب من أجل استخدامها في المعالجة الآلية للغة العربية.

– العمل على إنشاء قاعدة بيانات لغوية يمكن استغلالها في بناء برامج حاسوبية (المعجم الإلكتروني).

– التعرف على بعض المشكلات التي تواجه معالجة اللّغة العربية آليًا ومحاولة إيجاد حلول لها.

– التعرف على بعض البرامج الحاسوبية التي يكمن استخدامها في المعالجة الآلية للغة العربية.

– تطويع تقنية المعلومات لخدمة اللّغة العربية.

– محاولة بناء معجم إلكتروني للمعالجة الآلية للغة العربية يراعي في الوقت ذاته خصائص اللّغة العربية[1] يكون نموذجًا يمكن لتطويره في المستقبل واستثماره في إذكاء وتطوير برامج حاسوبية أخرى.

كما أُلفتُ انتباه القارئ الكريم إلى أنّ هناك جهود و دراسات مبذولة في مجال التحليل الصرفي للغة العربية سبقني إليها غيري[2]، إذ أنّها ترتبطُ بعملي من جانب أو آخر، لذلك يتوجبُ عليّ أن أُشيرَ إلى بعضها، من باب التعريف بها أو استغلالها في عملي هذا، ومنهم على سبيل الذكر لا الحصر:

فتم اقتراح وتنفيذ عدد من خوارزميات التحليل الصرفي، شرح كل من حجازي والشرقاوي شرحاً لنظام صرفي حاسوبي للنص العربي المشكل، وقد تم بناء عملهما على قواعد صرفية و صوتية للغة العربية . قدم كل من غيث والسعدني محللاً صرفياً يمكنه تحديد الجذر والتركيب الصرفي لأي كلمة مشكلة ولها جذر ثلاثي، و قام كل من الفداغي والعنزي ببناء طريقة رياضية بسيطة ، لاستنباط الجذور والصيغ الصرفية للكلمة المعطاة ، ثم أعطى هلال طريقة أكثر تفصيلاً لتحليل النص العربي غير المشكل ، وبعدها قام علي والشامي ببناء نموذج للتحليل الصرفي وتم تنفيذه في النظام المعروف بالمعالج الصرفي المتعدد الأطوار . ويقوم هذا النظام بالتعامل مع حالات التشكيل المختلفة تماماً للكلمة العربية .

ومن الأعمال الرائدة في مجال التحليل الصرفي للغة العربية أذكر منها:

1– محمد محمد داود، العربية وعلم اللغة الحديث، دار غريب للطباعة والنشر والتوزيع، القاهرة، د ط، 2001م، ص 45.

2– ترجع بدايات العمل والبحث في مجال التحليل الصرفي للغة العربية إلى الثمانينات . فقد قام عدد من الباحثين قي البلاد العربية وخارجها بعمل نظم للتحليل والتوليد الصرفي للغة العربية.

– مجموعة من الأعمال و البحوث العلمية التي قام بها البروفيسور يوسف ديشي[1] (JOSEPH DICHY)، في ميدان المعالجة الآلية للغة العربية.

– أطروحة محمد حسون[2]:" تطوير معجم للمعالجة الآلية للغة العربية في مختلف سياقات التطبيق"، جامعة ليون1987، وقد إستفدتُ من مضمونها العام في دراستي من خلال كيفية تعريف وتكوين القاعدة المعرفية للمعجم الإلكتروني، وكيفية بناء قاعدة المعطيات للمداخل المعجمية.

– الأستاذ مروان البواب، من خلال عرضه لتصور منهجي لإعداد معجم حاسوبي للغة العربية، وقد استفدتُ من بعض الجوانب الشكلية المتعلقة بهيكلة الجدولة الخاصة بقاعدة المعطيات للمعجم الإلكتروني.

ـ الباحث: بسو صديق، فقد استفدت من عمله المنجز في الماجستير على تقنيات وآليات إستخراج الجذر والوزن من العناصر اللّغوية[3]

ـ أمّا أُستاذي الكريم، عبد الرّحيم محمد الأمين، فاستفدتُ من عمله المنجز في الدكتوراة، والمتعلق بالتعرف الآلي على العناصر اللّغوية[4]، وفق جدولة صورية مكنتني من السير قُدُماً في عملٍ يعود الفضل في ذلك إلى توجيهات أُستاذي و مُعلمي **الدكتور سيدي محمد غيثري**، الذي ما فتئ يرشدني إلى مجالات لسانية وأخرى حاسوبية مرتبطة بالتدرج مع المستويات لخوض غمار هذا البحث.

ثامنا- العوائق والحلول المقترحة للنهوض بمشروع المعجم الإلكتروني للغة العربية :

- **لماذا تخلفت صناعة البرمجيات العربية؟**

إن صناعة برنامج عربي في صورته اللسانية الحاسوبية وفق معطيات وقاعدة بيانات لم تصل بعد إلى المستوى المطلوب في اللغة العربية، مع أن هذه اللغة من أكثر اللغات العالمية استجابة للقالب الرياضي سواء بأصواتها أو بصرفها أو بتركيبها، تنطلق الكلمة فيها من الجذور لتفرغ في قوالب صرفية وتصريفية قبل أن تنتقل إلى مكانها المناسب في التركيب الذي تحكمه

---

1 ـ جامعة لوميير ليون 2 : مجموعة من الأعمال التى أدت إلى قاعدة المعرفة اللغوية DIINAR.1) (القاموس المحوسب للغة العربية)، صممت ونفذت بالاشتراك مع باحثين من المعهد الوطنى العالى للعلوم و'المعلومات، والمكتبات (ENSSIB، فرنسا، محمد حسون، والمعهد الاقليمي لعلوم الحاسوب والاتصالات اللاسلكية، تونس (أ. براهام، الغزالي) وجامعة lumière، ليون 2 (J. Dichy).

2 ـ Hassoun M, Conception d’un dictionnaire pour le traitement automatique de l’arabe dans différents contextes d’application, Thèse de doctorat d’état, Univ. Lyon-1,1987.

3 ـ Bessou Sadik, Analyse de données Textuelles pour la classification Automatique par les techniques de Text Mining à la Langue Arabe, Mémoire de Magister, option: Sciences et technologies de l'information et de la Communication, Université FERHAT ABBAS- SETIF (UFAS) 2007.

4 ـ Mohamed El Amine ABDERRAHIM, Reconnaissance des unités linguistiques signifiantes. Thèse de doctorat spécialité informatique. Université de Tlemcen 2007- 2008.

قوانين صورية مضبوطة. مما يجعل من هذه اللغة بحق لغة جبرية على حد قول المختصين في اللسانيات الحاسوبية. **فما السبب؟**

- **الإجابة: نقص في الخبرة اللسانية الحاسوبية ونقص في فهم الحاسوبي لمتطلبات اللغوي، وانعدام التعاون بين المتخصصين في كلا الاتجاهين**:

وإن كانت هناك في الماضي صيحات تحاول أن تُهوِّن من شأن مشروع المعجم الإلكتروني للغة العربية، وتُشَكك في نجاحاته، فإنها الآن بدأت تطالب بإلحاح بتطوير وبناء أنظمة حاسوبية عربية للنهوض بهذا المشروع النهضوي، فبعد أن حققت نتائج ملحوظة وخاصة في ميدان الحوسبة والمعالجة الآلية للغة العربية[1] ، فبناء معجم صرفي إلكتروني للغة العربية يعتبر مشروعا علميا كبيرا لابد و أن تتوافر فيه جهود الباحث اللساني المُلِم بالنظريات اللسانية قديمها وحديثها، وجهود المهندس المتمكن من علوم الحاسوب ولغة البرامج والذكاء الإصطناعي.

بل لا نجانب الصواب إذا قلنا إن تنفيذ مثل هذه المشاريع اللغوية والتقنية يتطلب وجود مؤسسة حاضنة وداعمة للمشروع[2]، تنفق بسخاء من أجل تصميمه وانجازه.وكثيرة هي المشاريع اللسانية الحاسوبية العربية التي بدأها أفراد، وبقيت حبيسة الرفوف، ولم تتجاوز أسوار الجامعة أو المخابر العلمية ، حتى أنه لم يستفاد منها.

وبالنسبة إلى بحثي هذا ، فقد اعترضت سبيلي مشاكل وعوائق، منها ما يرتبط بالمصادر والمراجع النادرة، والتي لم أحصل عليها كاملة، باستثناء القليل منها الذي مكنني من النفاذ إليه عبر الشبكة العنكبوتية أو عبر المكتبات الخاصة، ومنها ما يرتبط بالعوائق التقنية التي تتمثل في صعوبة إنجاز برنامج حاسوبي يتمثل في نظام آلي قادر على توليد و استخراج مفردات اللغة العربية وتحليلها، نظرا لضعف تجربتنا في هذا الميدان. ولتجاوز هذا الوضع ، كنت أستشير أهل العلم شرقا وغربا ، كلما أتيحت لي فرصة أمامي، ناهيك عن زيارتي لبعض المركز

1- إنّ هذه التحدّيات العلمية واللّغوية، تضع على كاهل الباحثين في هذا المجال مسؤولية كبيرة ومصيرية فقد بات من الواجب تأسيس منظومة علمية عربية يكون هدفها النّهوض باللّغة العربية وتوجيهها لتنفيذ مشروعات في مجال المعالجة الآلية للغة العربية للدخول في عصر المعلومات واللّحاق بركب الأمم المتقدمة في مجال التقانات اللّغوية، خاصة في مجال المعالجة الآلية للغة العربية على المستوى الصرفي والنحوّي والدّلالي، فقضية معالجة اللّغة العربية آليًا قضية شديدة التعقيد، كونها تتطلب دراسة نظرية معقمة حول كيفية تمثيل المعارف اللّغوية من قواعد، واستنباط طرق ناجعة للمعالجة الآلية. ولا شك أن هذه القضية المحورية تشكل تحدّيًا أساسيًا للغة العربية من أجل بناء برامج حاسوبية تشمرُ في مثل هذه الدراسات والبحوث اللّغوية، من خلال تطوير قدراتنا الذاتية في هذا الصدد، وذلك لعدم جدوى استيراد الحلول الجاهزة أو المختصرة أو قصيرة الأجل.للمزيد من بسط القول ينظر : عبد الكريم خليفة، اللّغة العربية على مدارج القرن الواحد والعشرين، دار الغرب الإسلامي، الطبعة الأولى، 2003، ص 9،8.

2- ونظرا لما لمسناه من غياب في الخبرة اللسانية التي تعتمدها جل التطبيقات الحاسوبية المدونة أعلاه، قمنا بالعمل في مشروع علمي بالتعاون مع عدد من الباحثين في مخبر المعالجة الآلية للغة العربية تحت إشراف الدكتور سيدي محمد غيثري في جانبه اللساني الحاسوبي و الدكتور عبد الرحيم محمد الأمين كخبير حاسوبي فباعداد قاعدة معطيات لسانية نحو اعداد وتطوير معجم إلكتروني للمعالجة الآلية للغة العربية.

البحثية والجامعات خارج الوطن[1] التي استفدت من أصحابها الشيء الكثير، مما ذلل علينا صعوبات البحث.

وهكذا، إن النهوض بالبحث اللساني الحاسوبي العربي يتطلب إعادة الاعتبار لتخصص اللسانيات الحاسوبية والهندسة اللغوية في الجامعات ومراكز البحث العلمي، وتشجيع الطلبة على ولوج هذا الميدان، وحثهم عل تسجيل بحوث فيه، ثم تخصيص منحة لتمويل مثل هذه المشاريع، وتكوين فرق بحثية تضم في ثناياها تخصصات مختلفة، تسهر على تنفيذ تطبيقات المعالجة الآلية للغة العربية بشكل عام وخوض غمار حوسبة المعجم العربي كمشروع الأمة العربية الذي لا مفر منه، فاللغويون والحاسوبيون العرب مطالبون أكثر من أي وقت مضى بالإفادة القصوى من معطيات المشاريع المعاصرة لدى الغرب، وتسخيرها لخدمة اللغة العربية .

1- لقد قمت بزيارت علمية خارج الوطن منها على سبيل الذكر لا الحصر: جامعة ليون، و جامعة باريس 8 أثناء إقامتي العلمية بفرنسا، وفي إطار التربصات العلمية أثمرت زيارتي لمعهد التعريب ومختبر المعلوميات والمعالجة الآلية للغة العربية(م، م، م) بالرباط، المغرب، وتوجت فترة اقامتي بالمغرب بزيارة الجمعية المغربية للدراسات المعجمية بالرباط ولقاء رئيس الجمعية الأستاذ الدكتور عبد الغني أبو العزم.

الباب الأول

نحو معجم إلكتروني للغة العربية
في ضوء اللسانيات الحاسوبية

* قضايا نظرية ومنهجية *

الفصل الأول

المعجم العربي

* نحو توصيف جديد في اللسانيات الحاسوبية *

أولاً- الإطار العام لمفهوم اللسانيات الحاسوبية

ثانياً- اللسانيات الحاسوبية والمعالجة الآلية للغات الطبيعية

ثالثا- مكونات اللسانيات الحاسوبية ووظائفها

رابعا- الإطار العام للمعالجة الآلية للصرف العربي

خامساً- الجوانب التطبيقية للسانيات الحاسوبية

سادساً- اللسانيات الحاسوبية في خدمة الدراسات اللغة العربية: تجارب ورؤى مستقبلية

أولا- الإطار العام لمفهوم اللسانيات الحاسوبية:

1- تحديدات أولية :

إن الحاجة إلى الحوسبة أدّت إلى ظهور اللسانيات الحاسوبية التي تحاول في فهم الظاهرة اللغوية تنظيراو انجازا[1]، فلا يمكن للباحث أن يتخيل الفوائد النظرية والتطبيقية التي يمكن الحصول عليها من هذا العلم، فعندما يدرس اللسانيون المواد اللغوية دون استخدام الحاسوب، فإنه لا بد من استخدام منهج لساني معين. ولكن هذا المنهج اللساني والمطبق على المواد اللغوية، فإنه لا بد من تخزينه في الذاكرة الإنسانية ذات الصفات المحدودة ...[2]

فمنذ ظهور الحاسوب في أواخر الأربعينيات وصلته باللغة تتوثق وتتأصل[3]، فاللّغة تقع في قمة الموضوعات التي تهتم **بها العلوم الإنسانية، فكون اللّغة تجسيدًا لما في الذّهن البشري من نشاط، واتجاه الحاسوب لمحاكاة وظائف الإنسان وقدراته الذهنية وذلك مع تفاعل علمي وتقني بصورة لا مثيل لها[4].**

1- خليفة الميساوي، المصطلح اللساني وتأسيس المفهوم، دار الآمان، الرباط ، الطبعة الأولى، 1434 هـ ،2013 مـ ، ص 30

2- محمد علي الزركان، اللسانيات وبرمجة اللغة العربية في الحاسوب، مجلة المنهل، مايو 1993، ص182 .

3- نبيل علي، اللغة العربية والحاسوب، دراسة بحثية، تقديم: د.أسامة الخولي، دار تعريب، 1988، ص113.

4- تنبثق الدراسات اللغوية التي تتخذ الدراسات الحاسوبية محوراً لها من فرضية مفادها أن وصف النحاة للنظام اللغوي يتجاوز عن كثير من القواعد التي يفترض أن العقل البشري يعيها بالحس؛ فلا يصرح بتلك القواعد الضمنية لأن المقدرة العقلية البشرية الفطرية قادرة على احتوائها. وسواء كان الهدف الذي وضعت لأجله القواعد النحوية صيانة ألسنة العرب أنفسهم من اللحن، أو إعانة الناطق بغير العربية للتمكن منها إدراكاً وفهماً وإنتاجاً؛ فإن اللسانيات الحاسوبية تهدف إلى إعادة توصيف ذلك النظام اللغوي – انطلاقاً من أعمال الأوائل بالطبع – على نحو يستدرك ما لم ينص عليه علماء العربية معولين على الحدس البشري. إن غاية ما تحاول اللسانيات الكشف عنه هو استبطان المعطيات والأدلة الضمنية المدركة بالحدس لدى ابن اللغة، وصياغة تلك المعطيات في خطوات إجرائية وأدلة شكلية يمكن تمثيلها لجهاز الحاسوب المفتقر إلى (الفهم).

إن غاية ما تسعى اللسانيات الحاسوبية إلى كشف النقاب عنه هو وصف العمليات الذهنية التلقائية التي يقوم بها العقل البشري لدى استيعابه التراكيب اللغوية المتباينة أو عند تلقيه خطاباً لغوياً يحتمل عدة خيارات صحيحة. وإن كان العقل البشري يستعين بالقرائن السياقية أجل تحديد المعنى المراد من خلال استعراض الاحتمالات الممكنة ثم تقليلها بإقصاء معظمها حتى يقرر واحداً منها؛ فإن تلك القرائن السياقية التي يستعين بها العقل البشري هي ما تجهد اللسانيات الحاسوبية في ترجمتها إلى ألفاظ إجرائية يمكن توصيفها لمبرمج الحاسوب لغرض التمثيل الحاسوبي.

ونمثل على ذلك بالعملية الذهنية التي يحدد من خلالها العقل البشري تشكيل (ضرب) في ثلاث جمل مختلفة:

(أ) ضرب الأستاذ الطالب. (ب) ضرب الطالب من قبل الأستاذ لإهماله. (ج) ضرب الأستاذ الطالب سلوك غير مقبول.

يتعين في (أ) أن يكون الفعل مبنياً للمعلوم، وفي (ب) تعيننا الأدلة السياقية على أن (ضرب) فعل مبني للمجهول، وأنه مصدر في (ج). وفي واقع الأمر أن العقل البشري في تحديده تشكيل الكلمة (ضرب) في الأمثلة المذكورة يقوم بعمليات تلقائية فائقة السرعة قد لا نكترث لأهميتها، ولا نجهد في أدائها. ولعل تلك العمليات تتلخص في الابتداء باحتمال رئيس يتمثل في المبني للمعلوم، ويوازي هذا الاحتمال الرئيس ما يعرف في الحاسوب بالاحتمال الافتراضي Default؛ فإذا ما تبين بطلان الاحتمال الرئيسي استعرض العقل مجموعة من الاحتمالات أو البدائل Alternatives يقصي منها ما لا يأتلف وسياق الجملة، ثم يقرر في نهاية المطاف أن أحد تلك الاحتمالات هو الأنسب بناء على ما يتوافر لديه من المحددات والأدلة السياقية. للمزيد من التفصيل ينظر : */ سيبويه، الكتاب، تحقيق: عبد السلام هارون، دار الجيل، بيروت، 1991م، ص 39.

*/ عبد ذياب العجيلي، الحاسوب واللغة العربية، جامعة اليرموك، إربد، 1996م، ص 20. / محمد الزركان، اللسانيات وبرمجة اللغة العربية، السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات، مكتبة الملك عبد العزيز، الرياض، 1993م، ص 125.

*/ كارول شابل، تطبيقات الحاسب الآلي في اكتساب اللغة الثانية، أسس للتعليم والقياس والبحث العلمي، ترجمة: سعد القحطاني، جامعة الملك سعود، الرياض، 2007م، ص 60.

ممّا يمكنا من تحليل النظام اللغوي تحليلا آليا متعدد المستويات وبأسرع وقت ممكن[1].

فالحاسوب يحتاج إلى أدلة أقرب ما تكون إلى الإجرائية بحيث يمكن أو توصف بأنها (أدلة شكل)[2]

وفي سياق هذا التعريف بالإطار العام لهذا المجال اللساني أن ننتقل من النظري إلى التطبيق، ومن التجريد إلى التجسيد فأدّى إلى الثالوث: "اللّغة– الحاسوب– التطبيق"، فقد تهيأت اللّغة للحوسبة (المعالجة الآلية) بدخولها مجالات التحليل الرياضي والمنطقي والإحصائي، وتهيأ الحاسوب للقائه مع اللّغة بالسرية الفائقة، وفخامة الذاكرة، وصغر الحجم، وأساليب الذكاء الاصطناعي، ولغات البرمجة الراقية، وأمّا التطبيق، فقد شق السبيل إلى مجالات التعليم والإنسانيات والنّظم الكبيرة[3].

هكذا برزت الأسباب وتوفرت الأساليب، وتحددت الأهداف، مؤكدة حتمية التقاء اللّغة والحاسوب، وهو الالتقاء الّذي أطلق ثورة حقيقية في حقلي اللّغة والحاسوب[4]. ممّا أدى إلى تطوّر فرع معرفي جديد هو اللّسانيات الحاسوبية[5] وهو العامل المشترك لعلمي اللّسانيات والحاسوب، فاللّسانيات هي الأرضية النظرية لهذه التطبيقات بكل ما يتعلق باللّغة وأساليب وطرق صياغتها بشكل علاقات رياضية ومنطقية، أما علم الحاسوب فيتولى بلورة الجوانب اللّغوية وصياغتها بشكل برامج آلية تكون قادرة على جعل الحاسوب يفهم ويتعامل معها.

فاللسانيات الحاسوبية تطمح من خلال توصيف قواعد العربية لأغراض البرمجة الحاسوبية[6] أن يصبح الحاسوب (محللا لغوياً) قادراً على معالجة اللغة العربية تحليلاً وتوليداً. أما أن يكون

1– خليفة الميساوي، المصطلح اللساني وتأسيس المفهوم، مرجع سابق ، ص 31.

2– نهاد الموسى،العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية،المؤسسة العربية للدراسات والنشر، بيروت، ط1، 2000م، ص 20–27.

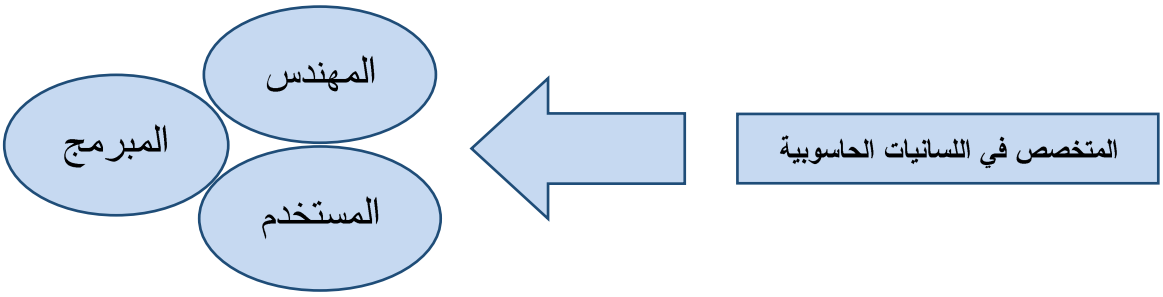
3– المرجع نفسه، ص116.

4– المرجع نفسه، ص116–117.

5– لمزيد من التوضيح، أنظر إشكالية التعريف الواردة في البحث، ص 19 .

6– مع ظهور عصر النهضة التقنية في القرن الميلادي العشرين، وبزوغ فجر جديد في تاريخ الحضارة البشرية باختراع الحاسوب، ثم تمكن المبرمجين من إحداث نقلة نوعية بالتعامل مع هذا الجهاز التقني عبر اللغات البشرية فضلاً عن اللغات البرمجية؛ يزيد الإلحاح على أبناء العربية كي يلحقوا بركب الحضارة، ويواكبوا ما استجد في مجال تطويع الحاسوب – الذي يعد ذروة التقنيات الحديثة – للتعامل مع اللغة – التي هي قمة علوم الإنسانيات – بإعادة توصيف قواعد العربية على نحو يتجاوز المعرفة التقليدية المتداولة في مجال الدرس وقاعات المحاضرة، ويكون من شأنه أن يمهد لمبرمجي الحاسوب تمكين الحاسوب من معالجة اللغة العربية معالجة آلية تكشف عن دخائل البنية الدفينة للغة العربية، وتحدد خصائصها ذات المغزى لأمور معالجتها آلياً. وكما أن متعلم اللغة العربية من الناطقين بغيرها يحتاج إلى توصيف للقواعد قد لا يحتاج إليه ابن اللغة، فإن من البديهي أن يحتاج الحاسوب إلى توصيف لقواعد العربية قد لا يحتاج إليها العقل البشري المجهز بالقدرة الفطرية. وتعد الدراسات اللغوية الحديثة التي تعنى بإعادة توصيف قواعد العربية لتكشف عما يعد فضلة للحدس البشري عمدة لبرمجة اللغة في الحاسوب؛ تعد تلك الدراسات جوهر تخصص (اللسانيات الحاسوبية) ومادته الأولية. للمزيد من بسط القول ينظر : نبيل علي، اللغة العربية والحاسوب، مؤسسة تعريب، الكويت، 1988م، التمهيد ص 1.

الحاسوب (وسيلة تعليم) للغة العربية أو (وعاء) تدون فيه اللغة العربية أو (مادة درس) وسيطه العربية؛ فإن ذلك يصب في اهتمام مجالات معرفية أخرى. إن الباحث المتخصص في اللسانيات الحاسوبية لا يتحتم عليه أن يكون مهندس حاسوب يجيد صيانة أجزائه، ولا مبرمجاً للحاسوب يتقن لغة البرمجة، ولا مستخدماً بارعاً لبرامج الحاسوب يحيط بتفاصيلها ودقائق إمكاناتها. إن المتخصص في اللسانيات الحاسوبية لغوي في الدرجة الأولى يعي النظام اللغوي للعربية، ويجيد التعامل مع أمهات الكتب إجادة يعين بها المبرمج على تمثيل اللغة تمثيلاً شكلياً دقيقاً من شأنه أن يزود الحاسوب بعدة الذكاء الاصطناعي، ويعوضه عما يفتقر إليه من الحدس البشري.



- اللغة وعصر المعلومات:

لقد أثرت النهضة تأثيراً كبيراً في مسار نشأة العلوم وتطورها، إذ أصبح الباحث لا يكاد يستغني عن ملامح هذه النهضة مهما كان تخصصه العلمي، لما أعطته التقنية من معلومات قيمة في أسرع وقت. فالباحث قديماً كان يبذل جهداً كبيراً ووقتاً طويلاً في سبيل البحث عن معلومة بسيطة، **لكن الباحث المعاصر يستطيع الحصول على المعلومة نفسها في أجزاء من الثانية دون أي عناء أو تعب بفضل دخول غمار الحوسبة**[1] .

ولا يخفى على أحد أن التقنية أصبحت شرياناً مهماً من شرايين البحث اللغوي، إذ تبوأت اللغات وعلومها مكانة مهمة في الأدوات التقنية، وعلى رأسها الحاسوب.

من هنا، فإن دخول اللغات على الحاسوب الآلي، وعلاقة ذلك باللغات كان تجاذب بين نقاط التماس المهمة في تطور كل منهما، فقد كان: (لدخول اللغات الحية وما يقابلها من علوم إنسانية وأدبية عالم الحاسبات أثر واضح في زيادة الإقبال على استخدام الحاسب في جميع الحقول والميادين)[2].

1- خليفة الميساوي، المصطلح اللساني وتأسيس المفهوم، دار الآمان، الرباط ، الطبعة الأولى، 1434 هـ ،2013 م ، ص 31.
2- العجيلي، عبد ذياب، الحاسوب واللغة العربية، ط1، منشورات جامعة اليرموك 1996م، ص 14.

فصار العقل الإلكتروني يسهل للعقل البشري[1] كثيراً من الأمور المتعلقة بالدرس اللغوي المتشعب المباحث[2]، وتعمقت الصلة بين هذين الطرفين حتى: (دانت اللغة للمعالجة الآلية من خلال التحليل الرياضي والمنطقي والإحصائي، وتهيأ الحاسوب للقائه مع اللغة بالسرعة الفائقة وضخامة الذاكرة، وضآلة الحجم، وأساليب الاصطناعي، ولغات البرمجة الراقية)[3].

وهكذا انعكس الالتقاء بين الحاسوب واللغة بوضوح على الأبحاث اللغوية، فأصبحت الحاجة ماسة إلى: (المراجعة الشاملة للمنظومة اللغوية ككل، لسبر أغوارها، واكتشاف ضوابطها، وإعادة تمحيص حالات شذوذها، وصياغة قوانينها في صورة دقيقة، ولا يختلف الحال بالنسبة للحاسوب. فقد استوجب الالتقاء تخلصه من معمارية تصميمه التقليدي)[4].

ونتيجة للتمازج بين العلوم اللغوية والحاسوب نشأ علم حديث يتبع لعلم اللسانيات، ويعنى بدراسة علاقتها بالحاسوب، يسمى علم اللسانيات الحاسوبية، ويعد هذا العلم حلقة الوصل بين التقنية الحديثة ودارس اللغة، إذ عمقت هذه الحلقة سبل التفاعل بين أبناء اللغة، والمعطيات اللغوية في هذه الأجهزة، فكان من نتاج هذا التفاعل ظهور جملة من التطبيقات اللغوية الحاسوبية المتطورة من أهمها:

- **أولا:** ظهور طرق الترجمة الآلية: (وتمثل الترجمة الآلية بين اللغات الإنسانية إحدى الغايات الأساسية للسانيات الحاسوبية، وقد تبين للعلماء إثر تجارب مريرة أنهم بعيدون

1- وتجدر الإشارة هنا إلى أن البحث في لطائف حقيقية عمل العقل البشري واستكناه العمليات الذهنية المعقدة التي يسلكها في تعامله مع اللغة يعد من صميم مباحث علم اللغة النفسي أحد فروع اللسانيات الحديثة، وأن هذا العلم، على ما أحرزه من تقدم واكتشافات، لا يزال يتصف بصعوبته وغموض مداخله وقصور المعرفة الإنسانية في مباحثه. الأمر الذي يعكس مدى صعوبة الخوض في مضمار اللسانيات الحاسوبية تبعاً لقصور معلومات المصدر الذي يستقي منه معطياته الأولية ومدخلاته الأساسية. بيد أن النظرة التكاملية للمعرفة الإنسانية تستشرف إضافات نوعية من شأن اللسانيات الحاسوبية أن تحرزها في فتح آفاق تخدم فروع اللسانيات جميعها.

ويؤكد ما نذهب إليه أن ثمة نواميس لغوية لا يعول فيها العقل البشري إلى معيار مصرح به منصوص عليه، ولكنه من ناحية أخرى يخدم تمثيل نظام اللغة للحاسوب بشكل رياضي حاسم. من ذلك أن الدراسات الإحصائيات لمواد المعاجم العربية – خلُصت مما خلُصت إليه – إلى أن العين لا تلي الهاء في أصل عربي مجرد. مما ينبؤ عن استحالة أن يكون في العربية الفعل (هعد) مثلاً. وهي حقيقة لغوية لا يسع الحدس البشري أن يبت فيها فيها بالوجود أو الانتفاء، لكنها في الوقت ذاته توصيف إضافي يعين الحاسوب على تصحيح خطأ في نحو (هعد أعضاء المجلس إلى الرئيس بالمهمة). ولا شك أن حدس ابن اللغة لا يستعين بقاعدة أو معيار في تعرفه على هذا الخطأ لاكتفائه بمعجمه الذهني.

2- كان لاستعمال الحاسوب أثر كبير في حفظ اللغات بشكل خاص، وحفظ التراث الفكري قديمه وحديثه بشكل عام، حيث قام العلماء بتخزين كم هائل من هذا التراث وحفظه في برامج مخصصة لذلك، بحيث يسهل تخزينها واسترجاعها في أي وقت، وتكون هذه البرامج متعددة التقنيات.

3- نبيل علي، اللغة والحاسوب، ط1، دراسة بحثية، تقديم: د. أسامة الخولي، دار تعريف 1988، ص 116.

4- المرجع نفسه، ص 117.

عن بلوغ هذه الغاية، وعلى الرغم من ذلك، وصفت اللسانيات الحاسوبية أنظمة برامج، يمكنها أن تيسر عمل المترجمين، وتحسن إنتاجياتهم تحسيناً آلياً)[1].

❖ **ثانياً:** ظهور ما يسمى **بالمعاجم الحاسوبية**[2] التي أصبح الهدف من وضعها عند العلماء يتمثل أساساً في (تقييس الدماغ البشري لسانياً، وذلك في محاولة لاستكشاف القدرة المعجمية عند الفرد العربي، وتقعيد إنتاجية هذا المستوى معلوماتياً بمساعدة الحاسوب)[3]، وأصبحت له وظائف محددة ومميزات خاصة، أضفت مزيداً من الإيجابيات على البحث اللغوي.

وقد أشار محمد الزركان إلى هذا بقوله: "يتجاوز المعجم الحاسوبي الحديث عن جميع ما يتجه على المعاجم التقليدية قديمها وحديثها من مآخذ في المادة والشرح والمنهج. ويشتمل على توصيف[4] معلوماتي لمواده، ويتضمن المعطيات اللازمة لتحديد المادة، وخصائصها النحوية

1- الزركان، محمد علي: اللسانيات وبرمجة اللغة العربية في الحاسوب، (الرياض، السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات التي نظمتها مكتبة الملك عبد العزيز، 1992، ص 54.

2- يشكل المعجم اللغوي ركيزة أساسية للتقدم المعرفي لدى الأمم فهو أحد الأدوات المهمة في التعليم والبحث العلمي واكتساب المعرفة والترجمة والتأليف . وإدراكا لأهمية المعجم في إثراء المحتوى العربي ومواكبة عصر المعلومات لما له من دور في تقدم الأمة وإثراء المعرفة الإنسانية فقد أنجزت بعض المعاجم الإلكترونية "كدليل مفهرس لمعرفة معاني الكلمات ومرادفاتها في اللغات الأخرى، تمت برمجتها إلكترونياً، ويمكن استخدامها من خلال جهاز إلكتروني خاص، أو عبر جهاز الحاسوب "، فالمعجم الإلكتروني يمكن تصنيفه حسب الخدمة التي يقوم بها، فهناك بعض المعاجم الإلكترونية التي تقوم باستخراج أصل الكلمة من المعاجم اللغوية العربية المعروفة مثل : لسان العرب، أو القاموس المحيط، أو مختار الصحاح، أو غيرها من المعاجم اللغوية. أما النوع الآخر فهو بعض المعاجم التخصصية التي تقوم بنفس عمل القاموس الذي ذكر آنفاً، إلا أنها تخص تخصصاً معيناً من حيث التعريف بمصطلحاته وشرحها، أو إيجاد المرادف لها من لغات أجنبية أخرى، مثل : المعاجم النفسية، والهندسية، والطبية، وهكذا...، يأتي التفصيل في التعريف بالمصطلح. في حينه من هذا البحث، لهذا ستكتفي في عرضه هنا بالقدر الذي يسمح لنا بوضعه كأحد التطبيقات الأساسية للسانيات الحاسوبية.

3- الموسى، نهاد، العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية. ط1 (بيروت – المؤسسة العربية للنشر والتوزيع 2000م، ص 54).

4- قبل أن نشرع في عرض المنهج المتبع لتوصيف قواعد العربية للحاسوب، تجدر بنا – في سياق عرض الإطار العام للسانيات الحاسوبية – الإشارة إلى ما يمكن أن يلتبس على بعض الباحثين من مباحث وتطبيقات أخرى تمت للحاسوب واللغة كليهما بصلة، لكنها لا تعد من صميم عمل اللسانيات الحاسوبية. فلا يعد من مجال اللسانيات الحاسوبية تعليم الحاسوب باللغة العربية على غرار ما تقوم به معاهد تعليم الحاسوب في الدول العربية، سواء كانت المادة المدروسة تعنى بطريقة استخدام برامج الحاسوب أو تصميم تلك البرامج وتعديلها أو صيانة الأجزاء الصلبة منه؛ حيث يكون الحاسوب عندئذ مادة الدرس وتكون العربية وسيط التعليم ولغة التخاطب، فلا يعد من تخصص اللسانيات الحاسوبية ما بات يعرف بمصطلح (تعليم اللغة بمساعدة الحاسوب) الذي يعنى بثلاثة مجالات هي: تعليم اللغة الثانية بمساعدة الحاسوب، واختبارات اللغة الثانية بمساعدة الحاسوب، وأبحاث تعلم اللغة الثانية بمساعدة الحاسوب حيث يعد الحاسوب في هذا التخصص وسيلة من وسائل تعليم اللغة الثانية شأن السبورة وجهاز العرض فوق الرأسي وبالبطاقات واللوحات وأشرطة التسجيل السمعية والبصرية. كذلك لا يعد من صميم اللسانيات الحاسوبية الأبحاث التقنية التي تعنى بآليات تخزين المعلومات في وسائط التخزين الحاسوبية المختلفة التي وثبت وثبات سريعة لتنتقل من القرص المرن إلى القرص الصلب فالمدمج فالذاكرة الضوئية المتميزة بصغر الحجم وضخامة السعة. إذ يعد الحاسوب في هذا المجال مجرد وعاء للغة شأن الورقة البيضاء أو سعف النخل وألواح الخشب بادئ أمر التدوين. أما آلية استدعاء تلك المعلومات المخزنة من خلال محرك بحث يلبي ترشيح مخرجات ذات خصائص لغوية موجهة؛ فتمثل ثمرة ما تقدمه اللسانيات الحاسوبية لتمكين الحاسوب من معالجة اللغات الطبيعيةللمزيد من التفصيل ينظر :

*/ عبد ذياب العجيلي، الحاسوب واللغة العربية، جامعة اليرموك، إربد، 1996م، ص 20.

*/ محمد الزركان، اللسانيات وبرمجة اللغة العربية، السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات، مكتبة الملك عبد العزيز، الرياض، 1993م، ص 125.

*/ كارول شابل، تطبيقات الحاسب الآلي في اكتساب اللغة الثانية، أسس للتعليم والقياس والبحث العلمي، ترجمة: سعد القحطاني، جامعة الملك سعود، الرياض، 2007م، ص 60.

والصرفية والدلالية والصوتية، ومعدلات استخدامها وغير ذلك. مما يوفر حاجة المستخدمين أو النظم اللغوية الآلية، مثل: المعالجات الصرفية والنحوية والدلالية، والفهم الآلي للغة، والترجمة بمساعدة الحاسوب، واسترجاع المعلومات وفهرستها، والمصححات الإملائية والنحوية، وتعليم اللغات، وتركيب الكلام وفهمه"[1].

- **ثالثاً:** المساعدة في دراسة النظرية اللغوية العربية انطلاقاً من منطقها الأصيل، وكما وضعها النحاة الأوائل: (فالنظرية اللغوية المتماسكة، أي التي لا تحتوي على غموض في تحديد مفاهيمها ولا تخلط بين هذه المفاهيم، ولا تقتصر على بعض أشكال هذا التحديد دون بعض، وهي الغاية المنشودة التي يجب أن يحققها اللغويون الذين يتعاونون مع الحاسوبيين في المعالجة الآلية للغة)[2]،

فاللسانيات الحاسوبية شكلت نقطة انطلاق لكثير من الدراسات اللغوية التي شاعت وانتشرت في هذا العصر. فساعدت العمليات الحاسوبية على الكشف من مرونة اللغة العربية في بنيتها الاشتقاقية **وطبيعة السوابق واللواحق للكلمات العربية**[3]، وعليه ما زال لازماً وميسوراً في الوقت نفسه تحديث اللغة العربية من خلال الحوسبة وهذا لمواكبة عصر المعلومات ودخول ميدان حوسبة المعجم العربي.

ومن خلال ما سبق، سأحاول في هذا الفصل توظيف أهم المفاهيم و النظريات الأساسية لمصطلح اللسانيات الحاسوبية في خدمة وتطوير المعجم الآلي للغة العربية في ضوء مقاربات منهجية وتصورية في معالجة اللغة العربية آليا .

1- الحناش محمد، نظرية حاسوبية لسانية لبناء المعاجم الآلية، الرياض، السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات، مكتبة الملك عبد العزيز، 1992، ص 365.

2- عبد الرحمن الحاج صالح، منطق النحو العربي والعلاج الحاسوبي للغات، (الرياض، السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات التي نظمتها مكتبة الملك عبد العزيز، 1992، ص 28.

3- محمد المغنم، مسألة السوابق واللواحق وطرق معالجتها، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 24، الرباط، المغرب، 1985/1984، ص 97.

2 ـ مفهوم اللسانيات الحاسوبية :

2 -1- إشكالية التعريف :

إن تعريف اللسانيات الحاسوبية، تعريفاً جامعاً ومانعاً وشاملاً، أمر صعب للغاية، إن لم نقل مستحيلاً، نظراً لاختلاف المنطلقات الفكرية والمرتكزات المعرفية والمنهجية التي ينطلق منها هذا الاتجاه أو الآخر في تعريف هذا العلم الجديد من جهة، ونظراً لتشعب وتباين المجالات الإمبريقية التي يهتم بها المجال نفسه من جهة أخرى[1].

وعلى هذا الأساس، يمكن القول إن مجمل التعريفات التي قدمت في هذا المضمار تتراوح بين التركيز على الجانب النظري، وبين التركيز على الجانب العمل التطبيقي، وبين جمع الجانبين معا في تعريف واحد. ومهما يكن من أمر هذا الاختلاف الإبستمي والنظري الذي يعتبر مفيداً في التاريخ والتوفيق لهذا العلم، خصوصاً إذا وضعنا في الاعتبار أنه في نقطة البداية، وكل بداية معرفية تتطلب الجمع والترتيب والتبويب والتراكم.

ولهذا السبب، سنقف عند بعض التعريفات، لأهميتها المذكورة، وفي المقابل سنقتصر على البعض منها، رغبة منا في تعميم الفائدة على القراء، وسعياً في إثارة النقاش حول إشكال التعريف.

إن أبسط تعريف للسانيات الحاسوبية هو أنها "علم يلم الشمل بين علمين قد تبدو المسافة بينهما متباعدة، هما علما(اللّسانيات) و(الحاسوب)؛ فهي "**علم حساب اللغات بنوعيها المنطوق والمكتوب**"[2]، بمعنى آخر إنها دراسة علمية للغة الطبيعية من منظور حاسوبي، وهذه الدراسة لا يمكن أن تتم إلا ببناء برامج حاسوبية لأنظمة اللغات البشرية، من خلال تقييس ومحاكاة نظام عمل الدماغ البشري لنظام عمل الدماغ البشري لنظام عمل الحاسب الآلي. ويرى البعض أنها **الدراسة العلمية للنظام اللغوي في سائر مستوياته بمنظار حاسوبي، يتجلى هدفها في تطبيق النماذج الحاسوبية على الملكة اللغوية**[3].

1– مهديوي عمر، تأهيل اللغة العربية وتهيئتها حاسوبياً، الندوة الدولية عن اللغة العربية ومتطلبات التهيء اللغوي – معهد التعريب، مرجع سابق.

2– مهديوي عمر، اللغة العربية والهندسة العربية، مدخل نظري، الندوة الدولية عن اللغة العربية والنظريات اللسانية، كلية الآداب والعلوم الإنسانية، فاس، سايس، 21 – 22 نوفمبر 2007.

3– نهاد الموسى، اللغة العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، 2000، ط1، ص 53.

في حين يعتبرها آخرون الميدان العلمي الذي يقع في نقطة الوصل بين اللغويات والحاسوبيات، وهي تنتسب إلى العلوم المعرفية، والتي تعتبر علوم الذكاء الاصطناعي شعبة من شعبها، موضوعها دراسة النماذج الحاسوبية للمعرفة البشرية[1].

2-2 - المصطلح بين النظري والتطبيق :

لقد استطاعت اللّسانيات الحديثة أن تحقق إنجازات عظيمة على المستويين النظري والتطبيقي، واندمجت مع العلوم الإنسانية، والطبيعية، والطبية والتقنية، فتوالد من هذا الاندماج فروع لسانية كثيرة، ومتباينة فكان من ذلك اللّسانيات الاجتماعية، والنفسية، والتربوية، والأدبية، والعصبية، والفيزيائية والحاسوبية....واللّسانيات البيولوجية... إضافة إلى فروع أخرى: كفن الترجمة، وفن صناعة المعاجم[2]...الخ.

ولعلّ اللّسانيات الحاسوبية تكون أحدث أفرع اللّسانيات الحديثة، ولعلّها تكون أهم هذه الفروع جميعا في عصر المعلومات[3].

إنّ هذا العلم الجديد ينتسب نصفه إلى اللّسانيات وموضوعها اللّغة، ونصفه الآخر حاسوبي، وموضوعه ترجمة اللّغة إلى رموز رياضية يفهمها الحاسوب ويعالجها[4].

فاللّغة هي القائم المشترك لعلوم اللّسانيات، لذا فإنّ دخول تطبيقات الحاسوب في مجال اللّسانيات لا بدّ أن يستند أساسًا إلى إخضاع اللّغة أولاً للسيطرة الآلية[5].

وقبل أن أسترسل في تعريف مصطلح اللّسانيات الحاسوبية يمكن أن نطرح تساؤلاً يكون بمثابة رسم للملامح الكبرى لطبيعة هذا العلم. **فمالمقصود باللّسانيات الحاسوبية؟ وما هي أهدافها ومجالاتها التطبيقية؟**

تعرف **اللّسانيات**[6] أو ما يسمى بعلم اللّغة بأنّه الدراسة العلمية للغة[7]. وقيل إنّه الدراسة العلمية للسان[8] وقد جاء في كتاب "بحوث ودراسات في علوم اللّسان" لعبد الرحمن الحاج صالح:

1- نهاد الموسى، المرجع نفسه، ص .53.

2- عبد السلام المسدي، قاموس اللّسانيات، الدار العربية للكتاب، تونس 1994، ص 155.

3- وليد العناتي، خالد الجبر، دليل الباحث إلى اللّسانيات الحاسوبية العربية، دار جرير، ط1، 1428- 2007، ص 13.

4- المرجع نفسه، الموضع نفسه.

5- نبيل علي، اللّغة العربية والحاسوب، ص 101.

6 - Linguistique

7 -The New Encyclopedia Britinnica, , encyclopedia Britinica Inc Translated, USA 1996 ,Vol 23,PP: 49 .

8 -John Lyons (Introduction to théorial linguistics),Combridge 1960, PP.1.

وينظر:

-André Martinet (Eléments de linguistique Générale) Edition, Paris. 1976. PP.06

"أنّ علم اللّسان يقوم على مفاهيم علمية، ذكر منها الإحاطة بمفردات اللّغة، ومعرفة دلالتها . فكان هذا المفهوم ناتج عن مقابلته لكلمة "نحو" مقابلة الشيء لقسيمه، وكذلك معرفة قواعد اللّغة المتعلقة بالمفردات من قبيل اشتقاقها وأبنيتها الصرفية[1].

فعلم اللّسانيات يهتم بتطبيق النظرية والاكتشافات اللّسانية على مجالات علمية متنوعة منها: مجال التعليم ووسائل الترجمة ولغة الحاسوب[2] فهي تقدم للباحث اللّساني الأسس العلمية التي يمكن الاعتماد عليها لبناء قواعد بيانية لصياغة نظرية لغوية ما[3].

إن **اللّسانيات الحاسوبية:** Linguistique Informatique: هو نظام بينيٌ بين اللّسانيات من جهة، وعلم الحاسوب من جهة أخرى[4] ويمكن التعبير عن ذلك بقولنا: أنّها دراسة اللّغة من جهة نظم حاسوبية[5].

ويُعرَفُ هذا العلم بعلم اللّغة الحاسوبي[6] ومعالجة اللّغة الطبيعية[7]، وهو دراسة الجوانب الحاسوبية للّغة، والمشاكل الشائعة التي تواجه المعالجة الحاسوبية للغة المكتوبة والمنظومة، فهي تُعنى بدراسة أنظمة الحاسوب[8] لغرض فهم وتوليد اللّغة اللّغة الطبيعية[9].

واللّسانيات الحاسوبية، أو الهندسة التكنولوجية[10] للغة الطبيعية تسعى إلى محاكاة[11] البشري لغويًا فهي تهدف إلى معالجة اللّغة الطبيعية حاسوبيا في مجالات البحث العلمي والتعليم وأهميته بالنسبة للغة العربية في وصفها للطريقة التي يجب إتباعها لإدماج العربية في الحاسوب أو ما يمكن بتسميته بحوسبة اللغة[12]، أو حوسبة النّص اللّغوي[13].

1- ينظر: عبد الرحمن الحاج صالح، بحوث ودراسات في علوم اللّسان، دار موفم للنشر والتوزيع، 2007، د ط، ص 36- 37.

2- ميشال زكريا، الألسنية – علم اللّغة الحديث- قراءة تمهيدية، المؤسسة الجامعية للدراسات، النشر والتوزيع بيروت، 1985، ص 206.

3- المسّدي عبد السلام، اللّسانيات من خلال النصوص- النشرة الأولى- الدار التونسية للنشر، تونس، 1984، د ط، ص 152- 153.

4- نهاد الموسى، العربية نحو توصيف جديد في ضوء اللّسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، الأردن، ط1، 2000، ص53.

5 − Colin Beardon and other (Natural Language and Computional Linguistics An Introduction, Ellis Horwood, 1991, PP: 10.

6 −Linguistique Informatique.

7 − Traitement de la langue naturel.

8 −Les Systèmes Informatiques

9- صلاح راشد جاسم محمد الناجم، دكتوراه في اللّغة الحاسوبي- عنوان الدكتوراه Computional Linguistics.

عنوان الموقع: Computional Approaches Arabic Morphology-/ http://www. Arabic linguistics. Net

http://www. Alnajm. Com

10 −Technologie de génie.

11 −Simulation.

12 −L'informatisation de la langue.

13- منتدى اللّسانيات، موقع الويب :http:// www. Lissaniat.net

ويعرفها "توفيق محمد شاهين" بأنّها: « ذلك الفرع الذي يعنى بدراسة اللغة لتطويع المادة اللّغوية، ومن أبرز مواضيع هذا العلم هو الترجمة الآلية[1] وتخزين المعلومات واستفادتها من ذاكرة الدماغ الإلكتروني Recherche Information »[2].

وهو بدوره فرع من علوم الذكاء الاصطناعي[3]، بشكل عام[4] وهذا الحقل الواسع من علم اللسانيات[5] أي اللّسانيات الحاسوبية، يتناول ويستكشف الآليات الأساسية التي عليها اللّغة والعقل وذلك بوصفها وصياغتها رياضيًا باستخدام اللّغات الصورية والاصطناعية لوضعها في نماذج ومن ثم محاولة محاكاتها في البرامج الحاسوبية[6].

أمّا "مازن الوعر" فيصفها بالإطار العلمي «الذي يبحث في "اللّغة البشرية كأداة طيّعة لمعالجتها في الآلة...»، وتتألف مبادئ هذا العلم من اللّسانيات العامة بجميع مستوياتها التحليلية، الصوتية والنحوية، والدلالية ومن علم الحواسيب الإلكترونية ومن الذكاء الاصطناعي[7] والمنطق[8] ثم الرياضيات مشكلة بذلك مبادئ اللّسانيات الحاسوبية[9].

ومن جانبها حاولت جمعية اللّسانيات ACL تعريفها قائلة:

" إنها الدراسة العلمية للغة من زاوية الإدراك الحاسوبي، وبالتالي فإنّها تهتم بحوسبة مختلف النماذج والظواهر اللّسانية، التي بإمكانها أن تكون قاعدة للمعارف[10] على سبيل المثال، أما العمل بها فيبرره الإدراك العلمي بخصوصيات الظواهر اللّغوية، أو ظواهر علم النفس اللّساني تارة والاحتياجات التكنولوجية لدراسة الخطاب أو أنظمة اللّغة الطبيعية تارة أخرى"[11].

1- عبد الله بن حمد الحميدان، مقدمة في الترجمة الآلية، (دراسات لغوية) مكتبة العبيكان، السعودية، الطبعة الأولى، 1421- 2001، ص 8.

2- توفيق محمد شاهين (علم اللّغة العام) مكتبة وهبة، الطبعة الأولى- 1400- 1980، ص 31- 32.

3 -L'intelligence artificielle.

4 - Carré, René Etal. « Langage Humain et Machine », Presses du CNRS, Paris, P: 300 , 1991.

5- ينظر: عبد السلام المسدي، مباحث تأسيسية في اللسانيات، دار الكتاب الجديدة المتحدة، ليبيا،الطبعة الأولى، 2010، ص12.

6 -Homiedan A, « Basis of the decision making process exercised by the translator, college of language, Ain shams university, philology P: 22 , 1994.

7 -Intelligence Artificielle

8 -Logique

9- مازن الوعر، قضايا أساسية في علم اللّسانيات الحديث، دار طلاس للدراسات والترجمة والنشر، دمشق،الطبعة الأولى 1988م، ص 406..وينظر: المؤلف نفسه، دراسات لسانية تطبيقية، ص317-320، د.رمزي بعلبكي، معجم المصطلحات اللغوية، ط1،دار العلم للملايين، بيروت – لبنان، 1990م، ص110، 301، د.محمد علي الخولي، معجم علم اللغة التطبيقي، ط1، مكتبة لبنان بيروت، 1986م، ص21 .

10 -Base de Connaissance.

11 -http://acl. Idc- Upenn. Edu/ P/ P: 79.

فاللّسانيات الحاسوبية تقوم على رؤى لسانية وأخرى حاسوبية بغية الوصول إلى غاية مُثلى مفادها أن تبلغ بالحاسوب مبلغ الذكاء الإنساني وأن نملكه كفاية لغوية[1] كالتي للإنسان حتى يصير قادرًا على فهم اللّغة وإنتاجها[2].

ملاحظة:

لقد أشرنا سابقًا إلى أن مصطلح اللّسانيات الحاسوبية عُرف بعدة نعوتات تعلم اللغة الحاسوبي، ومعالجة اللّغة الطبيعية، وهندسة التكنولوجية للغة الطبيعية، غير أنّه اصطلح عند بعض أهل الاختصاص بما يسمى اللّسانيات المعلوماتية[3] وهو ما يتضح عند Philip Miller و Thérèse Torris في كتابهما:

" Formalismes syntaxiques pour le traitement automatique de la langue naturelle "

حينما يقولان: « إن اللّسانيات المعلوماتية في مفهومها الواسع، هي تحاول أن تطور أنظمتها وبرامجها لمعالجة اللّغة الطبيعية في الحاسوب»[4].

انطلاقًا من هذه التعاريف نستنتج أن موضوع اللّسانيات الحاسوبية يبقى إطارًا رهين الفهم، وطبيعة التداول مع درجة الاستثمار الذي يمنحه له الباحث اللّغوي- الحاسوبي بالدرجة الأولى فيبقى بالتالي عملاً لغويًا آليًا بامتياز، يعالج مُثل كيان موّحد ومنظم في قلب الحواسيب الإلكترونية، وكأنّه إجراء دقيق لبعض العمليات الخوارزمية[5]؛ فعمل اللّساني يكمن في مراقبة وتوصيف اللّغة الطبيعية، ويقوم دور الحاسوبي عندئذٍ على نقل هذا التوصيف[6] إلى العالم الداخلي للحاسوب وذلك عن طريق إيجاد "بنى المعطيات والخوارزميات"[7] المناسبة والتي تحاكي كيفية استخدامنا للغة[8].

---

1 -Insuffisance linguistique.

2- وليد العناتي، خالد جبر، دليل الباحث إلى اللّسانيات الحاسوبية العربي، سابق، ص 13.

3 -Linguistique Informatique.

4 -Philip Miller et thèse torris (Formalismes Syntaxiques pour traitement automatique de la langue naturelle) Hermes. Paris. 1990. P: 15.

5- خوارزمية هي مجموعة من القواعد والعمليات الحسابية والرياضية المتسلسلة على شكل خطوات، يؤدي تنفيذها إلى الوصول إلى النتيجة المطلوبة.للمزيد من بسط القول ينظر المدخل ص..

6- أما التوصيف فإنه ينتظم الوصف اللغوي المجرد، مضافاً إليه العناصر التي يتعرفها الإنسان بالحدس والسليقة والقرائن المتعددة اللفظية والمعنوية والموقفية. ولما كان الحدس أظهر ما يتكئ عليه الإنسان في تعرفه اللغة وأدائها، ولما كان الحاسوب يفتقر إلى هذا العنصر البشري الخالص، وجب على المُوَصِّف أن يتدارك هذا النقص، ليبلغ بالحاسوب مَبْلَغَ المعرفة الإنسانية باللغة.

يعد مفهوم " التوصيف" عنصراً رئيساً في الأطروحة التي أسسها الدكتور نهاد الموسى في تهيئة العربية للتمثيل الحاسوبي. للمزيد من بسط القول في ضبط مصطلح التوصيف ينظر : - نهاد الموسى، العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية.

- نهاد الموسى، من الوصف إلى التوصيف، مقاربة في حوسبة العربية. فصل من كتاب الثنائيات في قضايا اللغة العربية من عصر النهضة إلى عصر العولمة.

7 - Structures de données et algorithmes.

8- عبد الله بن حمد الحميدان (مقدمة في الترجمة الآلية) سابق، ص 17.

إنّ اللّسانيين يراقبون اللّغة من أجل التعرف على الظواهر المميزة لها بكل امتداداتها وبالتالي من أجل وصفها باستخدام صياغات صورية أو لفات صناعية تقريبية منسجمة مع أسس البرمجة الحاسوبية ومع عالم المعلوميات يقصدون بذلك توضيح القواعد الصرفية والنحوية والدلالية للمعلوماتيين وإعطاءهم الوصف الموافق والمفيد لبنية المفردات والجمل في اللّغة أو لوظائفها وبعد هذا كلّه ينبغي على المبرمجين أن يسعوا دائمًا إلى تحسين أداء برامجهم بزيادة سرعتها وكفاءتها وتطوير طرق التعامل مع مستخدمي نظم الترجمة الآلية[1].

وكيفما كان الحال، إذا كانت اللّسانيات العامة تدرس اللّغة الطبيعية الدراسة العلمية الكفيلة بتحليل مستوياتها ووسائطها، تحليلاً يبرز مجالاتها، ويبرّر ظواهرها كنتاج اجتماعي من جهة، وكأداة للتواصل من جهة ثانية، فإن الحاسوب هو عبارة عن آلة إنجاز خام يتم توجيهها من خلال البرمجيات لتنفيذ وظائف محددة وفق قواعد معطيات أعدت خصيصًا لدراسة ظاهرة لغوية ما[2]. فقد أضحى هذا الجانب المعرفي ذاك الإطار التقني الّذي تنصهر داخله تجليات اللّغة الطبيعية ومظاهرها، في تفاعل يحاول ربط هذه الأخيرة (أي اللغة) بالحاسوب وأنظمته، أو أنّها الفرع الّذي يبحث في مختلف طرق وكيفيات الاستفادة من قدرات الحاسوب في تحليل اللّغة ومعالجتها وتعليمها وتعلمها... عن طريق الحوسبة.

فمعالجة اللّغة الطبيعية في الحواسيب هي القيام بدراسة نظم اللّغات الطبيعية مكتوبة ومنطوقة للتعرف على مكوّناتها وتحديد العلاقات بين هذه المكونات بغرض الوصول إلى قيام الحاسوب:

- بفهم والتعرف على الكلام المنطوق والمكتوب.
- بالاستجابة للكلام بتوليد الكلام مسموعًا ومكتوبًا[3].

وهنا جاء دور الحاسوب ليقوم بالتحليل ثم التوصيف ثم إعطاء برنامج حاسوبي وفق برامج يتعاون على وضعها عالم اللّغة وعالم الحاسوب في فرع معرفي يدعى اللّسانيات الحاسوبية[4].

1- المرجع السابق، ص: 17 و 18.

2- نبيل علي، البرمجيات ومهمة توظيف المعرفة، مجلة العربي- العدد 541- ديسمبر 2003- ص: 136.

3- بربارة سهيلة، الترجمة بمساعدة الحاسوب، من الانجليزية إلى العربية، مذكرة ماجستير من إشراف: أ. مختار محمد صباحي، جامعة الجزائر، قسم الترجمة، 2005- 2006، ص 20.

4- المرجع نفسه، ص 21.

وبتعبير آخر، إنّها محطة تقنية تدرس بنية اللّغة الطبيعية[1] في ارتباطها بجهاز الحاسوب على اعتباره أساس المنظومة المعرفية الأولى في تطويع[2] عناصر اللّغة كمادة تكنولوجية تُبرمجُ في الحواسيب الإلكترونية[3].

وعليه فخلاصة هذه المفاهيم، أنّ اللّسانيات الحاسوبية عبارة عن علم بيني يجمع بين اللّسانيات والحاسوب موضوعها تهيئة اللّغة الطبيعية لتكون لغة تخاطب وتحاور مع الحاسوب، مما يفضي إلى أن يؤدي الحاسوب كثيرًا من الأنشطة اللّغوية التي يؤديها الإنسان، مع إقامة الفرق في الوقت والكُلفة.

فهذا المجال من البحث اللغوي يقوم على تصور نظري يتخيّل الحاسوب عقلاً بشريًا، محاولة استكناه العمليات العقلية والنفسية التي يقوم بها العقل البشري حين ينتج اللّغة ويستقبلها، من ثمَّ يفهمها ويدركها، ولكنّها تستدرك على الحاسوب أنّه جهاز أصمٌّ لا يُستعمل إلاّ وفق البرنامج الّذي صممه الإنسان له، ولذلك ينبغي أن يوصّف للحاسوب المواد اللّغوية توصيفًا دقيقًا يستنفذ الإشكالات اللّغوية التي يدركها الإنسان بالحدس.

وكغيره من فروع اللّسانيات ينتظم اللّسانيات الحاسوبية مكوّنان: أحدهما نظري والآخر تطبيقي[4].

فأما الجانب الأول (النظري) فيبحث «في الإطار النظري العميق الذي به يمكننا أن نفترض كيف يعمل الدماغ الإلكتروني لحل المشكلات اللغوية»[5].

وأما الجانب الآخر (التطبيقي) فهو يُعنَى « بالناتج العملي لنمذجة الاستعمال الإنساني للغة... وإنتاج برامج ذات معرفة باللغة الإنسانية »[6].

ومن المعروف أن الجانب التطبيقي – هو الجانب الأهم في اللغويات الحاسوبية – يتمثل في تسخير العقل الإلكتروني لحل القضايا اللغوية، وهنا يبرز الدور الرئيس والأثر الفاعل لالتقاء اللغويين والحاسوبيين، والتعاون فيما بينهم، وما يثمر عنه من نتائج تسهم إلى حد كبير

---

1 −La structure du la langage naturel.

2 −Adaptation.

3 −Ordinateurs électroniques.

4− مازن الوعر، اللسانيات والعلم والتكنولوجيا " نحو تعريب موحد للسانيات التطبيقية العربية وبرمجتها في الحاسبات الإلكترونية"، دورية اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 22، الرباط، المغرب، 1983/1982، ص 20.

5− نهاد الموسى،العربية نحو توصيف جديد في ضوء اللّسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر،الأردن، ط1− 2000، ص 53− 54. وينظر: مازن الوعر، قضايا أساسية في علم اللسانيات الحديث، ص407 .

6− نهاد الموسى، العربية – نحو توصيف جديد ...، ص 53 .وينظر: د.مازن الوعر، قضايا أساسية في علم اللسانيات الحديث، ص407، اللسانيات والحاسوب واللغة العربية، صحيفة رؤى ثقافية (سورية)، العدد 4، 2003م، ص23 .

في تذليل العقبات وحل المشكلات التي تواجه التحليل الحاسوبي للغة العربية، هذه العقبات والمشكلات بعضها يتصل بطبيعة اللغة العربية، أصواتاً، وبنيةً، وتركيباً، ودلالة، ومعجماً، وبعضها يتعلق بنظام الكتابة العربية، وبعضها يتصل بالمصطلح العلمي التكنولوجي للسانيات العربية، كما أن هناك مشكلات أخرى تتعلق بالبرمجيات، إعداداً، واختياراً للمادة اللغوية العربية (أنموذج لساني عربي)، وتعريباً للبرمجة. وثالث هذه المشاكل يكمن في الجهاز الحاسوبي (الكمبيوتر)، وأنظمة تمثيل المعرفة على الحاسوب باللغة العربية.

وأمّا منتهي الغاية التي تجتهد اللّسانيات الحاسوبية أن تُحصّلها فهي أن نُهيّأ للحاسوب كفاية لغوية[1] تشبه ما يكون للإنسان حين يستقبل اللّغة ويدركها ويفهمها ثم يعيد إنتاجها على وفق المطلوب.

ورغم ما يكتنف التعريفات السابقة من لبس، فإننا نسجل في الإطار نفسه. أنها تتكامل وتتقاطع في نقطة واحدة هي الهدف المتوخى في اللسانيات الحاسوبية بجانبيها النظري والتطبيقي والمتمثل أساساً في إيجاد حل للقضايا والأسئلة التي تطرحها المعالجة الآلية للغات الطبيعية.

ثانيا: اللسانيات الحاسوبية والمعالجة الآلية للغات الطبيعية :

لقد فتح الدرس اللساني الحديث أفاق نحو تطبيقات حاسوبية متطورة خاضعة لقاعدة بيانات لسانية، وذلك بسبب نضج النظريات اللسانية اللآلية التي وظفت لمقاربة الموضوع اللساني في سائر مستوياته، وبالتالي تم الشروع في بناء مدققات إملائية ونحوية، ومحللات صرفية ونحوية، الشيء الذي غذى حقل اللسانيات الحاسوبية بالترسانة المعرفية الازمة والضرورية.

ومن هنا بدأ التمييز بين اللسانيات الحاسوبية[2]، باعتبارها دراسة علمية للغة الطبيعية باستخدام الحاسوب الآلي[3]، وبين تطوير معالج آلي للغة الطبيعية[4]، لأن الكثير من غير المتخصصين والفنيين في هذا المجال لا يفرقون بين المصطلحين. في حين هناك فرق واضح بين المجالين لم ينتبه إليه أهل الاختصاص في العصر الحديث، سنوضحه فيما يأتي من الكلام.

1- ينظر نهاد الموسى،الأساليب مناهج ونماذج في تعليم اللّغة العربية، دار الشرق- الأردن، 2003، ط1، ص 123- 124.

2 -Linguistique Informatique

3- عرفت اللسانيات الحاسوبية تعريفات متباينة، بحسب المجال المعرفي الذي وظف فيه المصطلح نفسه. لهذا أكدنا على هذا التعريف من خلال ما سبق للرفع اللبس من تداخل في التعريف بالمصطلح ومجالاته التطبيقة كما ورد في بعض التعريفات التي ذكرناها في ثنايا هذا البحث.

4 -Elaboration automatique du langage naturel

* إن "تطوير معالج آلي للغة الطبيعية(EALNA) هو المجال اللغوي الذي يستفيد من تقنية الحاسوب ولا سيما التقنيات المعلومياتية غير الذكية في إنجاز تطبيقات لغوية كالإحصاء المعلومياتي المطبق على المدونات المعجمية، ونستشهد هنا بالأعمال التي قام بها **علي حلمي موسى**[1] في إحصاء مدونات المعاجم العربية القديمة، وحصر نسبة تردد الحروف المتآلفة أو المتنافرة صوتياً في بنية الجذر. إن مثل هذه الأعمال لا علاقة لها بأي حال من الأحوال باللسانيات الحاسوبية، وإنما هي أقرب ما تكون إلى تطوير معالج إحصائي آلي للمعجم العربي، وبنفس الأسلوب تم التعامل مع ألفاظ القرآن الكريم مع **نبيل علي ومحمد زكي خضر**[2]. إن اللسانيات الحاسوبية بهذا المعنى تقبل التطبيق على مختلف نواحي البحث اللساني النظري مما يتصل بالتركيب، والصرف والتداولية، وتحليل الخطاب وغيرها.

أما اللسانيات الحاسوبية فهي فرع من فروع اللسانيات التطبيقية والصورية، شأنها في ذلك شأن، الاكتساب اللغوي، وتعليم اللغة وتعلمها، وتحليل الخطاب، والأسلوبية. وبهذا المعنى فهي اختبار لعلوم الحاسوب الآلي على أنظمة اللغات الطبيعية باستخدام التقنيات المعلوماتية المتطورة المستمدة أساساً من هندسة الحاسوب، وهندسة المعرفة، والذكاء الاصطناعي[3].

فقد حصل تأثير متبادل بين علم اللّغة (اللّسانيات) وعلم الحاسوب في ميدان النظريات اللّغوية، سواء كانت صرفية أو نحوية أو معجمية، من حيث مناهجها التحليلية مهدّ بظهور علم جديد أو مصطلح جديد عرف باسم اللّسانيات الحاسوبية.

ومن بين المجالات اللّسانية التي حصل فيها أيضًا تقدّم وكانت سببًا في الثورة المعرفية الحديثة التوصل إلى حوسبة اللّغة في عدد من جوانبها سواء كانت نحوية أو صرفية أو دلالية، وتوفيق عدد من التقنيات التي يصعب القيام بها دون الآلة، مثل الإحصاء والتخزين الذي يتطلب ذاكرة ضخمة....

وقد طرح هذا الوضع الجديد عددًا من الأسئلة وأفرز متطلبات جديدة، وكان لا بدّ أن تقتحم اللّغة العربية على غرار غيرها من اللّغات، ميدان الحاسوب وتدخله من الباب الواسع. إلاّ

1– علي حلمي موسى، دراسة إحصائية لجذور معجم الصحاح باستخدام الكومبيوتر، جامعة الكويت، رقم 33، 1973.
إحصائية جذور معجم لسان العرب باستخدام الكمبيوتر، جامعة الكويت، 1979، 19.
علي حلمي موسى، وعبد الصابور شاهين، دراسة إحصائيةلجذور معجم تاج العروس باستخدام الكمبيوتر، جامعة الكويت، رقم 1973، 32.
2– المعجم المفهرس لألفاظ القرآن الكريم.

3– Mario Montelone ;Lexicographie et Dictionnaires Electroniques des usages linguistiques aux bases de données lexicales ; Thèse de Doctorat, Université de Marne-la-Vallée, Déc. 2003 ; P.8.

أنّ هناك شروطًا لتهيأ الأدوات الوصفية والصورية الضرورية لجعلها قادرة على اقتحام الميدان، وهذا بصياغة نظريات لغوية ضمن قواعد بيانات تكون مضبوطة وهذا لحوسبتها وفق برامج حاسوبية والمشكل المطروح في مثل هذه الدراسات يتمثل في البحث في قواعد اللّغة (أو نحوها)، والبحث في مفردات اللّغة (أو معجمها)، فالبحث في هذين الميدانين مازال متعثرًا، ومازلنا بحاجة إلى صورة شاملة للقواعد والضوابط التي تجعل المعرفة النّحوية والمعرفة المعجمية قابلتين لأن تتخذا قاعدة للمعطيات في الحاسوب حيث يمكن توظيف هذه القواعد إلى جانب المجالات الصرفية والنحوية، أمّا بالنسبة لقاعدة المعطيات المعجمية فلا بدّ من تهيأ المداخل المعجمية بالشكل الكافي واللاّفت[1].

ثالثا : مكونات اللسانيات الحاسوبية ووظائفها:

إنّ أهم ما تصبو إليه اللّسانيات الحاسوبية، هو محاولة تلقين الحاسوب لغة بشرية تمكنه من التحاور مع مريديه شفويًا وكتابيًا حتى يصبح قادرًا على فهم الكلام وإنتاجه وتوليده، كما هو الحال بالنسبة للدماغ البشري، بمعنى آخر أن يصبح الحاسوب قادرًا على محاكاة أو على الأقل مجاراة التفكير والإنجاز البشري[2]. فعملية التخاطب عمومًا بين البشر والحاسوب تقتضي معرفة تامة بقواعد الحوار من جهة وبعض المعارف الثقافية والاجتماعية"[3].

ونشير هنا أن اللّسانيات الحاسوبية هي فرع من فروع اللّسانيات الصورية فهي تعالج التحوّلات للنماذج اللّغوية وفق قواعد مضبوطة، وهذا الانضباط ينبغي أن يكون أكثر دقة وصرامة، وفق ما يعرف باللّسانيات الإجرائية لأنّها تدرس نظريات تطبيقية لتفعيل النماذج الصورية، فقد بات من الواضح أن المعالجة الصورية للغات الطبيعية لم تعد ترفاً فكرياً، كما كان يتصور البعض، في وقت من الأوقات، بل أضحت اليوم ضرورة علمية ملحة أكثر من أي وقت مضى، وذلك نظراً للتحديات الكبرى التي تطرحها العولمة والمعلوماتية في الألفية الثالثة

1- منظمة الأمم المتحدة للتربية والعلوم والثقافة، تقدم اللّسانيات في الأقطار العربية، وقائع ندوة جهوية، أبريل 1987- الرباط، ط1، 1991، ص 25- 26.

2- إنّ مجهودات علماء الحاسوب في هذا الشأن، مازالت عبارة عن محاولات لبناء نظم قادرة على فهم اللّغة الإنسانية وإنتاجها تماما كما يفعل سائر البشر إلاّ أنّ الأمر لم يكن باليسير عليهم، رغم ما توقعه اليابانيون في العقدين الأخريين، حين اختراعهم لجهاز حاسوبي ينتظرون منه معرفة ما يريده مستخدموه عن طريق "فعل السؤال والجواب" /Acte: Question/ Repense، ثم جعله قادرًا على مناظرة الإنسان بلغته القومية وفهم الكلام والصورة. كما يؤكد على ذلك الخبير في مجال الحواسيب، الياباني. إدوارد فايجنبومE. A. Feignebaum . ينظر: ادوارد فايجنبوم وباميلاماك كوردك، الجيل الخامس: الذكاء الاصطناعي والتحدي الياباني للعالم، د ط،1987، ص 286.

3- شحادة فارع وآخرون، مقدمة في اللغويات المعاصرة، نشر وائل، الجامعة الأردنية، د.ط، د.ت، ص319.

على جميع الأصعدة عامة، وعلى مستوى اللغات الإنسانية خاصة. واليوم هذه الإجراءات في معظمها تنفذ بواسطة الحاسوب[1].

الغرض من هذا الفرع الجديد من فروع البحث اللّساني هو معالجة اللّغة الطبيعية بطريقة آلية، وتعتمد اللّسانيات الحاسوبية على شتى المعارف والتخصصات لكي تعني بغرض الوصف الحاسوبي للسانيات:

1- تستمد اللّسانيات الحاسوبية من اللّسانيات العامة مادتها الأولى، والمقصود بمادتها الأولى اللّغة الطبيعية وما يتصل بهذه اللّغة الطبيعية من أدوات الوصف والمصطلحات.

2- تستمد اللّسانيات الحاسوبية من المفهوم الفلسفي للظاهرة اللّغوية، ويعرف هذا الفرع المعرفي بفلسفة اللّغة والمنطق ودورهما في فهم علاقة اللّغة بالفكر[2].

3- تستمد اللّسانيات الحاسوبية من الذكاء الاصطناعي، الّذي له دور في توظيف هذه الهندسة في التحليل اللّساني الآلي (**تحليل صرفي، نحوي، صوتي، معجمي...**) ويحوّل اللّسانيات الحاسوبية اللّغة الطبيعية إلى تمثيلات صورية ونماذج حسابية لهذه اللّغة الطبيعية. تسعى اللّسانيات الحاسوبية لمحاكاة الدماغ البشري لغويًا"[3].

4- تستمد اللّسانيات الحاسوبية من الرّياضيات كعلم الإحصاء، ونظرية المخططات في فهمها الآلي للغة الطبيعية[4] ومجموعة من القواعد الحسابية متسلسلة تسمى الخوارزميات[5].

تتكون اللغويات المعلومياتية من مكونين رئيسيين وهما:[6]

*** المكون التطبيقي للسانيات الحاسوبية أو اللسانيات المعلومياتية التطبيقية:**

تهتم اللسانيات الحاسوبية التطبيقية بدراسة الجانب العملي الإنجازي في النموذج اللغوي، وتعتبر قواعد الإنتاج عنصراً أساسياً وضرورياً لتحقيق التفاعل والتحاور والتواصل بين المستخدم البشري وآلة الحاسوب. ويتمثل الهدف الرئيس لهذا المكون في صناعة أنظمة إنتاج المعرفة الخاصة باللغات البشرية، سيما وأن المشكلة الحقيقية بين الإنسان والحاسوب تكمن في الفهم أي أن الحاسب الآلي يجد صعوبة في فهم اللغة الطبيعية البشرية، كما أن الإنسان هو الآخر يجد صعوبة في فهم اللغة الاصطناعية، والتي تختلف جذرياً عن بنية اللغة المخزنة في دماغ الإنسان.

---

1 -Alain Polguère, (Structuration et Mise en jeu Procédurale d'un modèle linguistique) déclaratif dans un cadre de génération de texte- thèse doctorat (PH. D) en linguistique, 1990.PP:1.

2- عبد الرحمن الحاج صالح، بحوث ودراسات في علوم اللّسان، دار موفم للنشر، د. ط- 2007، ص 30- 31.

3- آلان بونيه، الذكاء الاصطناعي واقعه ومستقبله، ترجمة علي صبري فرغلي، عالم المعرفة، 1993 ، رقم 172، ص 45- 46.

4- ينظر: نبيل علي، اللّغة العربية والحاسوب، ص 85.

5- عبد الحسن الحسيني، معجم مصطلحات علوم الكمبيوتر، عربي انجليزي- دار القلم- بيروت، ط1، 1407/ 1987، ص 65.

6 - ينظر :

Mario Montelone ; Lexicographie et Dictionnaires Electroniques ; des usages linguistiques aux bases de données lexicales ; Thèse de Doctorat, Université de Marne-la-Vallée, Déc. 2003 ; P.8.

وينظر: مازن الوعر، قضايا أساسية في علم اللغة الحديث، دار طلاس للدراسات، والترجمة والنشر، ط1، 1988، ص. 407 – 408.

وأصبح من المؤكد، أن هذا المشكل قد تجوز في كثير من جوانبه وحيثياته، إن لم نقل، إن العناصر الأساسية فيه تم العثور لها عن حلول، وقد تم ذلك، بالفعل، بفضل التعاون المثمر بين المهندسين واللسانيين في غضون الأعوام الأخيرة، إذ استطاعت اللسانيات الحاسوبية أن تحقق نتائج عملية إمبريقية مهمة وأساسية في دراسة وتحليل خصائص الأنظمة اللغوية الطبيعية.

وهكذا، إذا كانت البرامج اللسانية الحاسوبية ترمي إلى توثيق المعلومات التي ترد على الذهن البشري، فإن هناك إمكانات عديدة لحوسبة هذه المعلومات/ العمليات. لذا فإن فهم الآلة للغة الطبيعية يعني توظيف اللغة وتقييس[1] استخدامها في الحاسوب أصبح أمراً بالغ الأهمية على مستوى تشغيل البرامج والاستعمال الإنتاجي.

إن مستقبل المكون اللساني الحاسوبي التطبيقي إذن رهين بمدى النجاح الذي يمكن أن تحققه أنظمة تشغيل البرامج، ذلك أن تقييس ومحاكاة[2] الملكة اللغوية عند البشر متوقف على التقدم الذي من المنتظر أن تحققه الأبحاث والدراسات في هذا المجال، والتي تتوخى صياغة أنظمة قادرة على تيسير عمل الإنسان، ولا سيما فيما يتصل ببرامج نحو الكلمة.

لكن في مقابل ذلك نجد أنّ Sylvain Surcin يتأسف على عدم تحقيق مثل هذا الرّجاء بقوله:«...ذلك أنّه كلّما تقدمت خطوات استقصاء المعلوماتيين لأنماط انتظام المكونات اللّغوية المختلفة ووقائعها المتنوعة، واستكشاف أبعاد اشتغالها المعقدة كلّما صادفتهم مشاكل عديدة تحطم رؤاهم الأولية، وتصوراتهم البسيطة التي كانوا يعتبرون اللّغة بموجبها تكديسًا من الأنظمة المتتالية والمستقلة عن بعضها البعض»[3] ونجد Augustin Lux حين يقول:

1- أن للتّقييس (المحاكاة) زمنا بدأ مع العلوم الدقيقة والتطبيقية خاصة في مجالي الهندسة والعلوم، فكان استعماله يقتصر على اختبار صلاحية فرضية ما عندما لا يُتاح اختبار النظام الحقيقي. في هذا الإطار يقول الدكتور عادل عوض عن المحاكاة الحاسوبية "بعد أن يجري تطوير نموذج رياضي ديناميكي لعملية ما، فإنّ المعادلات التي تشكل النموذج يجب حلّها كي تسمح بتقدير سلوك العملية بالعلاقة مع الزمن وعملية التطوير هذه تعرف بالمحاكاة ويمكن الوصول إليها باستعمال النموذج لاكتشاف تأثير الحالات المتغيرة في النظام الحقيقي". ينظر : عادل عوض، تطبيقات النمذجة والمحاكاة الحاسوبية في الأنظمة البيئية- ألمانية، عالم الفكر العدد، 3،المجلد 32- 2004 ص 126.

2- لقد طرح "مينسكي" (1965) السؤال: هل من الواجب على الحاسوب أن يحاكي سلوك البشر في التفكير - وقبله "Aranturing" (1950)- حتى نفاه مرتين وأثبته في واحدة قائلاً: «لا يستطيع الحاسوب التفكير، لو حدّدنا هذا المفهوم في ارتباطه بأي نشاط إنساني خاص...، ولا يفكر كذلك لو كان ما يستنتج من الفكر غير قابل للتحليل، كأن يكون شيئًا أسطوريًا أو خياليًا أو حتى سريًا».

وقد أثبته مرة واحدة بشرط لما قال: « نعم يستطيع الحاسوب أن يفكر لو وجد السؤال مساره نحو التجربة، وكانت الملاحظة إجراء حاسم في مقارنة السلوكات الآلية بالسلوكات البشرية، ليتبيّن لنا عندها من هو أهل للتفكير" وبناءًا على هذا الإثبات، فهل يعني ذلك أن نظام التعرف والفهم الإنساني» الّذي أتى به "Rossi Mario"(1996) في مقالة: المعارف والمعالجة الآلية للكلام"2 نظام بأساس معرفي- Système à base de Connaissance. ينظر : مصطفى بوعناني، الفونولوجيات الحاسوبية والمسارات المعرفية للإنجاز الكلامي، ملفات معرفية، الملف الأول مطبعة أبي، الطبعة الأولى، 2003، ص 5.

3- مصطفى بوعناني، الفونولوجيات الحاسوبية والمسارات المعرفية للإنجاز الكلامي، ملفات معرفية، الملف الأول مطبعة أبي، الطبعة الأولى، 2003، ص 5، 6.

« إنّ الواقع يتجاوز التوقع »[1] فكان يعني أنّ التحديات والأمل في بناء آلة تحاكي السلوك والتفكير البشريين أمرًا يبقى مرهونًا بمدى تقبلنا أصلاً للفكرة المطروحة. وهذا كله في حضرة استفهام تشوبه عدّة إشكالات في خضم هذا الصراع الفكري المعلوماتي؟

لقد ترك هذا التساؤل حيرة على كل المختصين في المجالين اللّساني والحاسوبي في محاولة فردية لإيجاد توافق ولو بسيط بين المجالين رغم أن البعض منهم لا يعترف بدور اللّسانيين المهتمين بالسلوك البشري لإنجاح ذاك النظام المعرفي، فتكون طبيعة هذا تأكيد عدم قدرة الحاسوب أو أي نظام آخر مهما بلغت قوته وسرعته على التكافؤ مع الإنسان إن لم يتأسس على تقييس (محاكاة) Simmulation تام ودقيق للسلوك الإنساني، لأن المشكل لم يكن ليطرح لو توفرت الآلة على مخ إنساني[2].

فقد أثارت هذه النقطة كثيرًا من الجدل بين باحثي الذكاء الاصطناعي تركزت معظم تساؤلاتهم حول إمكانية محاكاة برامج الذكاء الاصطناعي الطريقة التي يتبعها الإنسان في حل المسائل، أم أن الطريقة لا تهم طالما يتوصل البرنامج في النهاية إلى حل بشكل أو بآخر[3].

*** المكون النظري للسانيات الحاسوبية أو اللسانيات الحاسوبية النظرية:**

تعتبر اللسانيات الحاسوبية أهم مكون في التطبيقات المستقبلية المتعلقة بأنظمة المعلومات، إذ كلما زاد اعتماد الإنسان على الحاسب الآلي، كلما زادت الحاجة الملحة إلى تحويل بنيات اللغة الطبيعية مما يتصل بالصرف والنحو والدلالة وغيرها إلى صورة رقمية، وذلك من أجل توفير تطبيقات البحث السريعة عن المعلومات لا البيانات فقط، ولتوفير إمكانية التعامل مع الحاسوب بالتخاطب الطبيعي، مما سيؤثر جذرياً على مجالات واسعة من حياة البشر، وخاصة في ميداني التعليم والأعمال وغيرهما.

يُعنى مكون اللسانيات الحاسوبية النظرية بدراسة خصائص المعرفة اللغوية عند البشر، ويتوخى مد الإنسان بالأدوات والإجراءات التي تساعده على فهم اللغة وتوليدها.

إن نظريات المكون اللساني الحاسوبي تنصب على دراسة الأشكال اللغوية المستخدمة في الحاسوب. لهذا يسعى اللسانيون والحاسوبيون، جنباً إلى جنب، إلى صياغة نماذج صورية

---

1 –Augustin Lux, (sur l'évolution de l'intelligence Artificielle) la machine perceptive 11/12 janvier 1990, P: 12.

2– المرجع نفسه، ص 91.

3– آلان بونيه، الذكاء الاصطناعي واقعه ومستقبله– ترجمة علي صبري فرغلي– عالم المعرفة– العدد 172، 1993 ، ص 19.

تحاكي نظام اشتغال الملكة اللغوية عند البشر، والتي تقبل البرمجة الحاسوبية، وتعتبر هذه البرامج قاعدة صلبة لتطوير النظريات والمناهج اللسانية وبلورتها[1].

والغاية من ذلك تحقيق الكفاية[2] أو المقدرة اللّغوية للحاسوب، أو ما يصطلح عليه الآن بالكفاية الحاسوبية "Compétence informatique" والتي من خلالها يمكن للحاسوب أن يعرف اللّغة شكلاً ومضموناً وبوجهيها الكلامي والمفهومي، وبعبارة أدق هي النّظام الّذي يولّد وينتج ويفسِّر الجمل تفسيرًا دلاليًا صحيحًا ومفهومًا[3].

وأمام هذه التحديات والصعوبات، لجأ الباحثون في علم اللّسانيات الحاسوبية إلى العمل في نماذج تعنى بمجالات لغوية محدودة، حين انصب اهتمامهم على دراسة الجانب اللّغوي فجزؤوه إلى نحو ومعجم وصرف ودلالة... فعدلوا (مؤقتا) عن فكرة بناء نموذج (يحاكي) مقدرات الإنسان اللّغوية وغير اللّغوية، وركزوا مجهوداتهم على بناء نماذج حاسوبية معينة وفي حقول محددة.

1- مهديوي عمر، سابق، 1999، ص 59، 16.

2- والكفاية هنا هي المؤدي الضمني لمفهوم تشومسكي، وهي تتألف على المستوى النظري من:

أولاً: تداخل قواعد اللّغة العربية في نظامها الصوتي، وأنساقها الصرفية، وأنماط نَظْمها الجُمْلي، وأنحاء أعاريبها، ودلالات ألفاظها، ووجوه استعمالها وأساليبها في البيان وأحكام رسمها الإملائي.

وبيان ذلك أن لكلّ لغة نظمها الصوتية والصرفية والنحوية...الخاصة التي تتأسس على قواعد يكتسبها الناطق باللغة على نحوٍ غيرِ واعٍ، ويؤديها بتلقائية ومن أمثلة ذلك من العربية:

- ففي النّظام الصوتي يَصْدُرُ العربي في نطق لغته عن قوانين صوتية لا تتهيأ له بالمعرفة المباشرة، فهو يُماثل في (ال) الشمسية ويُقصِّر الصلة قبل المساكن، ويراعي قوانين التركيب المقطعي، فلا يبدأ بصامتين، ولا ينتهي بهما إلاّ في الوقف حصرًا...
- وفي النّظام الصرفي يعرّف الأبنية الصرفية للاسم والفعل.. وما هو مشترك بينهما ويعرف أبنية اسم الفاعل واسم المفعول واسم المكان، فإذا التبست الأبنية الصرفية إستعان بالسياق ليغض اللّبس.
- وفي النّظام النحوي يعرف قواعد إعراب الأسماء والأفعال والحروف، ويعرف قواعد تركيب الجملة الاسمية وتركيب الجملة الفعلية وما يعرض لكل واحدة منهما من امتدادات من اليسار أو اليمين.
- وفي الدلالة يعرف معاني المفردات المعجمية، ويستعين بالسّياق ليفضّ الالتباس بين الألفاظ المشتركة، وكذا القول في النظام الأسلوبي والهجائي. وهذه القواعد اللّغوية متناهية مهما بلغ عددها.

ثانيا: إنتاج ما لا يتناهى من الأداءات اللّغوية الصحيحة، فبالرغم من أن بناء اللّغة على قواعد محدودة، إلاّ أنّنا نستطيع أن نولّد من هذه القواعد عددًا غير محدود ولا معدود من الأداءات اللّغوية، وهذا ما عرفته العرب بـــ "القياس والتمثيل" وعرفه تشومسكي بـــ"لاّ تناهي" وهذان مثلان من العربية يوضحان المقصود "بلاّتناهي".

ثالثا: مرجعٌ في تمييز الخطأ من الصواب: ويتشكل هذا المرجع من قواعد النظام اللّغوي الذي استدخله النّاطق باللّغة اكتسابًا وصقله بالتدريب والتمرين فيصير هذا النّظام محتكمه ومرجعه في رد ما يرد عليه من أخطاء وزلات وذلك أننا نفسر الخطأ ثم نصوبه بالعودة إلى القاعدة التي نعرفها، فإن كتب أحدهم: ناقش الأستاذ القصةُ رددناه إلى الصواب بقولنا: إن القصة: منصوبة لوقوع الفعل عليها.

وإن قال أحدهم: سَبَرْتُ عليك كثيرًا، فالصحيح من أن (السَبْر) لا يكون في هذا السياق) وإنّما القصد (صيرتُ).....

رابعًا: ومن تمام هذه الكفاية اللّغوية كفاية تواصلية*، أن يستخدم اللّغة وفقًا لمقتضيات السياق (سياق المقام وسياق المقال)، كما قالت العرب قديمًا: "لكل مقام مقال" والنّاس الناطقون الأصلاء بلغة ما يميزون هذا باكتساب اللّغة عُرفيًا وذلك أنّهم يميزون ما ينبغي قوله من عبارات في مواقف على التعيين فلا يضل عنهم أن للتعزية عبارات خاصة وللمباركة عبارات خاصة ولمخاطبة الكبير آدابًا...

- الكفاية التواصلية: Compétence Communicative : مصطلح للساني الأمريكي (دل هايمز) وهو استدراك أقامه على مفهوم الكفاية اللّغوية عند تشومسكي، إذ يرى (هايمز) أن معرفة قواعد اللّغة وحدها لا تكفي للتواصل بل ينبغي أن يعرف النّاطق باللّغة كيفية استخدام هذه القواعد والسياقات التي ينبغي أن تقال فيها. (ينظر نهاد الموسى، الأساليب مناهج ونماذج في تعليم اللّغة العربية).

3- ينظر: نهاد الموسى، الأساليب مناهج ونماذج في تعليم اللغة العربية، دار الشروق، الأردن، 2003، ط1، ص123-124.

رابعا: الجوانب التطبيقية للسانيات الحاسوبية:

مع تقدم الحاسوب، وتطور أساليبه ظلّت تطبيقاته في مجال اللّغة تتفرع وتتعمق حتى غطت معظم أركان المنظومة اللّغوية، ونفذت إلى طبقاتها العميقة وجوانبها الخفية وتتراوح التطبيقات اللّغوية هذه ما بين تلك التي لا تُرى في الآلة الالكترونية إلاّ قوتها الحسابية الهائلة، إلى تلك المُقامة على أدق الأسس اللّغوية ناظرة إلى الحاسوب كآلة ذكية قادرة على أن تعلِّم وتتعلم. وعلى الرّغم من تعدد وتفاوت أغراض استخدامات الحاسوب كأداة للغة[1]، فقد حققت تطبيقات الذكاء الاصطناعي لجعل الحاسوب يفهم اللّغة الطبيعية إنجازات هامة مما قاد إلى تطور هذا الفرع (اللّسانيات الحاسوبية) وهو العامل المشترك لعلمي اللّسانيات والحاسوب، فاللّسانيات هي الأرضية النظرية لهذه التطبيقات بكل ما يتعلق باللّغة وأساليب وطرق صياغتها بشكل علاقات رياضية ومنطقية، أما علم الحاسوب فيتولى بلورة الجوانب اللّغوية وصياغتها بشكل برامج آلية تكون قادرة على جعل الحاسوب يفهم لغة البشر ويتعامل معها[2].

تحاول تطبيقات اللّسانيات الحاسوبية إلى التوصل إلى إمكانية تعامل الإنسان مع الحاسوب باستخدام لغته الطبيعية، وعلى الرغم من حداثة هذا المجال، إلاّ أنه يعدّ فردًا من فروع العلم الّذي يهتم بتطبيقات نظريات الحاسوب في الدراسات العلمية للغات البشر وتصميم منظومات تستطيع معالجة وتحليل وتوليد اللّغات المكتوبة والمنطوقة[3].

إنّ جوانبه العلمية اللّغوية تسعى للنهوض والتعمق في نظريات ودراسات المعجم، والنّحو، والدلالة، وتوليد وتحليل اللّغة، وكيفية اكتساب اللّغة، أمّا جوانبه العملية فهي تركّز على تطوير تطبيقات حاسوبية لاستخدام اللّغة الطبيعية في الجوانب العملية اليومية للبشر، وهي تهدف في النهاية إلى جعل الحاسوب يفهم لغة البشر[4]. وبناءًا على ما تقدم يمكن تصنيف اللّسانيات الحاسوب إلى عدّة مجالات رئيسية لعل أهمها[5]:

1- الإحصاء اللّغوي.

2- التحليل والتركيب النّحوي والصرفي.

3- الإعراب الآلي[6].

1- نبيل علي، اللغة العربية والحاسوب، سابق، ص130.

2- بربارة سهيلة، الترجمة بمساعدة الحاسوب من الإنجليزية إلى العربية، سابق، ص24.

3- المرجع نفسه، ص24.

4- المرجع نفسه، ص25.

5- نبيل علي، اللغة العربية والحاسوب، سابق، ص131، وينظر: د.وليد العناني وينظر: د.خالد الجبر، دليل الباحث إلى اللسانيات الحاسوبية العربية، ص14.

6- يعدّ هذا المنحى استكمالاً لمطلب تمكين الحاسوب من القراءة الصحيحة للنصوص العربية غير المشكولة، وقد ألمحنا إلى النصف الأول لاستيفاء هذا المطلب في توصيف النظام المورفولوجي المعني بحركات المبنى، وسنختصر الكلام هنا على سبيل الإشارة لا التفصيل . أما الإعراب المتداول في قاعات الدرس بما هو عملية ذهنية تنتظم في نسقها الخاص ونموذجها الهيكلي المنضبط بمقولات عيار مخصوص؛ فقد تصدّر له من الباحثين من عوّل في ذلك على أربع خطوات إجرائية تعتمد أربع قوائم تحوي فيها كل قائمة مجموعة من البدائل، ويقع الاختيار على إحدى تلكم البدائل بناء على ثبت مستدخل في ذاكرة الحاسوب مسبقاً ويستدل فيه بقرائن السياق. على أنه لا بأس أن نحاول نهجاً آخر في توصيف نظام الإعراب للحاسوب، وحسبنا في هذا المبحث لزوم القصد، وغاية القصد من توصيف نظام الإعراب تمكين الحاسوب من تعيين حركات أواخر الكلم .للمزيد من التفصيل ينظر: نهاد الموسى، العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، ص 158 – 168.

4- الترجمة الآلية.

5- تعلم اللّغة باستخدام الحاسوب.

6- الفهم الآلي للسياق.

7- استنطاق النصوص وتخزينها من الصورة المكتوبة إلى المنطوقة.

8- إدخال النصوص وتخزينها بالمسح الضوئي.

9- بنوك المصطلحات[1] و قاعدة المعلومات.

10- قواعد المعارف.

11- الفهرسة.

12- التحليل الإحصائي.

13- ميكنة المعجم أو ما يعرف بـــ المعاجم الالكترونية[2].

وترتيب مجالات التطبيق- كما أوردناه هنا- لا يعكس التسلسل الزمني لبداية ظهورها، فقد بدأت محاولات الترجمة مثلاً مع بداية ظهور الحاسوب، وإذا أردنا أن نراعي ترتيب هذه المجالات من حيث التدرج المنطقي فإنّه يكون كما يلي:

1- الإحصاء اللّغوي[3].

2- التحليل والتركيب اللّغويات.

3- الفهم الأوتوماتي للسياق.

4- تحليل النصوص (أو النتاج اللّغوي).

5- ميكنة المعاجم.

6- الترجمة الآلية.

7- تعلم اللّغة باستخدام الحاسوب.

1- للمزيد من التفصيل ينظر : علي القاسمي، نحو تطوير بنوك المصطلحات "أداة للبحث المصطلحي والعلمي"، مجلة اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 28، الرباط، المغرب، 1987، ص 79 وما بعدها.

2- يأتي تفصيل هذا العنصر في هذا البحث، لهذا سنكتفي في عرضه هنا بالقدر الذي يسمح لنا بوضع الإطار العام للسانيات الحاسوبية لمنظومة اللغة ككل.

3- أما في مجال الإحصاء اللغوي – وهو كما سبق الميدان الأول لتطبيق استخدام الحاسوب في البحث اللغوي العربي المعاصر – فلا يخفى أن استخدام الإحصاء الرياضي في اللغة يحقق تقييماً كمياً «لبعض الخواص النوعية للغة، كمعدلات استخدام الحروف، والكلمات، والصيغ الصرفية، والموازين الشعرية، وأنواع الأساليب النحوية، أو التوزيع النسبي للأفعال المعتلة والصحيحة، أو للإفراد والتثنية والجمع، أو لحالات الإعراب المختلفة». كما يحقق توصيفاً كمياً لبعض العلاقات اللغوية، كالعلاقة بين طول جذر الكلمة وعدد مرات تكراره، والعلاقة بين طول الكلمة ومعدل استخدامها داخل النصوص. ينظر: د. نبيل علي، اللغة العربية والحاسوب، ص131، وينظر: ص52،54 من هذا الكتاب .

حيث يمكن النظر إلى الثلاث مجالات الأولى بصفتها تطبيقات أساسية، في حين يسود الثلاث الأخيرة طابع التطبيق العملي[1].

1-الإحصاء اللّغوي[2]:

ويقوم الإحصاء بتفسير بعض الظواهر اللغوية وتحليلها[3]. ليس هذا فحسب، بل هناك «إحصاء جديد يستطيع أن يتعامل مع البنية المعقدة للسياق اللغوي، حتى يكشف لنا عن علاقات الترابط والتماسك بين فقراته وجمله وألفاظه، وتلك التي تربط بين ظاهر العبارات وما تبطنه من معان وإشارات»[4] .

لا يختلف دور الإحصاء في النطاق اللّغوي عن ذلك الّذي في المجالات العلمية الأخرى، إذ يوكل إليه ثلاث مهام رئيسية:

أ- **التقييم الكمي لبعض الخواص النوعية للغة:** كمعدلات استخدام الحروف، والكلمات، والصيغ الصرفية، والموازين الشعرية، وأنواع الأساليب النحوية، أو التوزيع النسبي للأفعال المعتلة والصحيحة، أو للإفراد والتثنية والجمع، أو لحالات الإعراب المختلفة، وسنكتفي هنا بمثالين؛ الأول عن المعدل النسبي لاستخدام علامات التشكيل النص "القرآني" والثاني عن أكثر الحروف الإنجليزية تواترا[5].

المثال الأول في اللغة العربية	
الحركة	**المعدل النسبي**
الفتحة (القصيرة)	%43.92
الكسرة (القصيرة)	%18.89
الضمة (القصيرة)	%13.90
ألف المد (الفتحة الطويلة)	%14.76
الياء الساكنة (الكسرة الطويلة)	%4.24
الواو الساكنة (الضمة الطويلة)	%5.39

المثال الثاني في اللغة الإنجليزية	
الرف	**المعدل النسبي**
E	%12.7
T	%9.8
A	%7,9
O	%7,8
I	%7,1

<u>الجدول رقم 01</u>:المعدل النسبي لاستخدام علامات التشكيل النص "القرآني"، وأكثر الحروف الإنجليزية تواترا.

1- نبيل علي، اللغة العربية والحاسوب، ص 131 .

2 - langue statistique.

3- اللغة العربية والحاسوب، سابق، ص 132، 135 .

4- نبيل علي، الثقافة العربية وعصر المعلومات، المجلس الوطني للثقافة والفنون والآداب، الكويت (سلسلة عالم المعرفة، رقم 265)، 2001م، ص 254.

5- اللغة العربية والحاسوب، سابق، ص131 .

وخلاصة، يتضح من هذه النتائج بشكل قاطع طغيان التحريك بالفتح في (العربية) وميل "الإنجليزية" إلى الكسر، وكذلك استذكار كلتيهما للحركة المضمومة، ومنه أيضا يستفاد انحياز الحركة نحو القصر في العربية"[1].

ب-**التوصيف الكمي لبعض العلاقات اللغوية:** ونعطي هنا مثالين؛ أحدهما من "العربية" والثاني من "الإنجليزية".

<u>المثال الأول</u>: العلاقة بين طول جذر الكلمة وعدد مرات تكراره في معجم "لسان العرب"[2]

طول الجذر	عدد مرات التكرار
3	6538 (%80,5)
4	2458 (%27,5)
5	187 (%2)

<u>الجدول رقم 02</u> : العلاقة بين طول جذر الكلمة وعدد مرات تكراره في معجم "لسان العرب.

<u>المثال الثاني</u>: العلاقة بين طول الكلمة في المعجم الإنجليزي ومعدل استخدامها داخل النصوص:

طول الكلمة (عدد حروفها)	عدد الكلمات ذات الطول نفسه
4	879
5	936
6	895
7	771
8	527
9	362
10	260
11	137
12	79
13	30
14	8

<u>الجدول رقم 03</u> : العلاقة بين طول الكلمة في المعجم الإنجليزي ومعدل استخدامها داخل النصوص.

تشير هذه الإحصائيات إلى التناقض الحاد في معدلات استخدام الجذور العربية والكلمات ولتأكيد فكرة الانتظام الإحصائي، بغض النظر عن عنصر البناء اللغوي الذي نتخذه مؤشرا لذلك، نعرض هنا العلاقة بين طول "الآية القرآنية" مقاسا بعدد ألفاظها، وعدد مرات تكرار الآيات بهذا الطول، وذلك في النص القرآني الشريف، وكذلك العلاقة بين طول الجملة

1- المرجع السابق ، ص 132.

2- المرجع نفسه ، ص 132.

الإنجليزية وتواترها في عينة من النصوص، ومرة أخرى يتضح ذلك الاستذكار في اللغة في استخدام التعبيرات الطويلة، ويؤكد هذا بشكل كمي الخاصية الاقتصادية للاستخدام اللغوي، والتي تعكس نفسها في كثير من أمور اللغة...[1]

ج- تفسير الظواهر اللغوية: يستخدم التحليل الإحصائي كذلك في تفسير الظواهر اللغوية، ومثال لذلك تفسير بعض ظواهر عدم الاطراد في قواعد الاشتقاق وتركيب الكلمة الإنجليزية[2].

طول الآية	معدل تكرارها	طول الآية	معدل تكرارها
من 1 إلى 5	1526	من 46 إلى 50	35
من 6 إلى 10	1517	من 51 إلى 55	18
من 11 إلى 15	1255	من 56 إلى 60	10
من 16 إلى 20	832	من 61 إلى 65	10
من 21 إلى 25	429	من 66 إلى 70	3
من 26 إلى 30	295	من 71 إلى 75	4
من 31 إلى 35	159	من 81 إلى 85	5
من 36 إلى 40	80	من 86 إلى 145	3
من 41 إلى 45	53		

الجدول رقم 04: معدلات تواتر الأطوار المختلفة للآيات القرآنية (مقاسة بعدد الكلمات)[3]

طول الجملة	معدل تكرارها
من 1 إلى 5	2
من 6 إلى 10	18
من 11 إلى 15	26
من 16 إلى 20	18
من 21 إلى 25	15
من 26 إلى 30	9
من 31 إلى 35	6
من 36 إلى 40	5
من 41 إلى 45	1

الجدول رقم 05: معدل تكرار الأطوار المختلفة للجمل الإنجليزية (مقاسا بعدد الكلمات). في عينة صغيرة من النصوص[4]

1- المرجع السابق، ص 133.
2- المرجع نفسه، الصفحة نفسها.
3- المرجع نفسه ، ص 134.
4- المرجع نفسه، الموضع نفسه.

ملاحظة: انطلاقا من الإحصاءات الواردة في الجداول السابقة، نلاحظ أنها تتناول كلها تقريبا ظواهر مشتركة بين "العربية" و"الإنجليزية"، وهو اختيار متعمد قد روعي فيه تأكيد "عمومية" هذا الانتظام الذي هو أحد مظاهر العموم الغوي بصفة عامة وظاهرة الفائض اللغوي بصفة خاصة[1].

هذا، ويعتمد الإحصاء اللغوي على ثلاثة نماذج إحصائية أساسية:

- نموذج "زيبف"[2]، ويعتمد على نظرية الاحتمالات.
- نموذج "شانون"[3] ويعتمد على نظرية المعلومات.
- نموذج ماركوف[4] ويعتمد على السلاسل الزمنية والتنبؤ الإحصائي.

يعيب النهج الإحصائي اعتماده على قوة الحاسوب، واقتصاره على التجليات السطحية القابلة للتقييم الكمي، ذلك لأنه يأتي مرحلة وسيطة إذا قصد الباحث التعامل مع عدم اليقين العلمي، وهو بذلك لا يعادل الرياضيات والمنطق في القدرة على كشف المكنونات والمتداخلات[5].

2- التحليل والتركيب اللغويان:

يستخدم الحاسوب في كثير من العمليات الأساسية للتحليل والتركيب اللغويين[6]، والتي تعد إحدى الخطوات التمهيدية للنظم المختلفة لمعالجة اللغة آليا، وفيما يأتي أمثلة بيانية لعمليات التحليل والتركيب، مع التركيز على مفهومها[7] كثنائية وفي الوقت نفسه كطورين أساسيين لعمل المنظومة اللغوية، علما أنهما شائعان في مجال اللسانيات الحاسوبية[8].

من خلال المجالات[9]:

أولا: (على مستوى الصوتيات) تمييز الكلام وتوليده آليا.

ثانيا: (على مستوى الكتابة) القراءة الآلية للنصوص المكتوبة وطباعتها.

ثالثا: (على مستوى الصرف) تحليل الكلمات وتركيبها آليا.

رابعا: (على مستوى النحو) تحليل الجمل (إعرابها) وتوليدها آليا.

غير أن الكلام هنا لن يشمل كافة هذه المجالات لأني سأركز على جانب واحد منا، يتمثل في المستوى الصرفي الذي يعنى بتحليل الكلمات وتركيبها آليا.

1- المرجع السابق، ص 135.

2 - ZEPF Modèle.

3 - Shannon modèle

4 - Modèle de Markov

5- نبيل علي،اللغة العربية والحاسوب دراسة بحثية، ص 135.

6 - Analyse et structure de la langue

7- المقصود: ثنائية التحليل والتركيب

8- نبيل علي، اللغة العربية والحاسوب، سابق، ص135-136.

9- المرجع نفسه، ص 138.

أ- المستوى الصرفي: تحليل الكلمات وتركيبها آليا:

يقصد بالتحليل الصرفي[1] النظام الآلي بتحليل ما يغذي إليه من كلمات إلى عناصرها الصرفية الأولية (الصرفيمات) وتحديد سماتها الصرف صوتية، والصرف نحوية، مثال ذلك:

الكلمة "المدخلة"للمحلل الصرفي الآلي[2]، هي: "والإستيطاني"

* مبدأ عمل المحلل الصرفي الآلي:

*الكلمة "المدخلة" للمحلل الصرفي الآلي[3] هي: "والاستيطاني"

*مبدأ عمل المحلل الصرفي الآلي[4]:

عناصر الكلمة: "و" + "الـــ" + "استيطان" + "ياء النسب"+ "علامة الضمة"

جذع الكلمة: "استطان"

جذر الكلمة: وطن

الصيغة الصرفية: إستفعال

الميزان الصرفي: إستيعال

قسم الكلم: اسم معنوي (مجرد)، مفرد، متصرف.

الحالة الإعرابية: مرفوع

العمليات الفونولوجية: إبدال الواو ياء، وكسر النون قبل ياء النسب وتعد عملية التحليل الصرفي إحدى العمليات الأساسية لمعالجة اللّغة آليًا وتدخل في العديد من التطبيقات مثل:

اكتشاف الأخطاء الإملائية وتصحيحها، ميكنة المعاجم، **(المعاجم الإلكترونية)**، تحليل النصوص، تمييز الكلام، استرجاع المعلومات...الخ.

يمثل التركيب الصرفي[5] العملية العكسية للتحليل الصرفي، إذ يغذي إلى المركب جذر الكلمة، (أو جذعها)، والحالة الاشتقاقية والتصريفية والإعرابية المراد إخراج الكلمة على صورتها[6].

مثال: من العربية[7]:

العناصر "المدخلة" المغذاة إلى المركب الصرفي الآلي هي: جذر الكلمة: ق- و- م

---

1 - Analyse morphologique.

2 - Analyse morphologique automatique.

3 -Analyse morphologique automatique.

4- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 138.

5 -Structure morphologique.

6- اللّغة العربية والحاسوب، سابق، ص 139.

7- المرجع نفسه، ص 139.

الصيغة الصرفية المطلوبة: المضارع المزيد من صيغة "إستفعل" بيانات التصريف: الجزم، لجمع المؤنث.

الكلمة المخرجة، المركب الصرفي: "يستقمن".

3- الفهم الآلي للسّياق اللّغوي[1]:

الهدف الرئيسي للنظم الآلية للتحليل اللّغوي هو الوصول إلى نظام آلي لفهم السّياق اللّغوي في صورته المنطوقة والمكتوبة[2]، ومعظم المحاولات التي تجري حاليًا للفهم الإنساني للسياق اللّغوي، وذلك من خلال تطوير نظم تتكامل فيها جميع جوانب المعرفة اللّغوية والسياقية لاستخلاص المعنى المقصود من السياق. لقد شبّه البعض عملية الفهم الآلي بعملية حل ألغاز تكوين الأشكال من قطع صغيرة ومتداخلة، فالفهم ليس في جوهره عملية تحليل مطرد[3].

وتعد نظم الفهم الآلي للسياق اللّغوي ذات أهمية بالغة للبرامج التعليمية، وذلك لاستخلاص شبكة المفاهيم الأساسية التي تضمنها نصوص المادة التعليمية[4].

4- استخدام اللّسانيات الحاسوبية في تحليل النّصوص:

ومن أهم المجالات التي استخدمت فيها أساليب الإحصاء والتحليل اللّغوي، تحليل الناتج اللّغوي الأدبي: تراثه، وحديثه، منثوره، ومنظومه، وذلك للأغراض التالية[5]:

1- **تحقيق التراث:** وذلك لتحديد تاريخ نشأة الوثائق ومصدرها، ومن أبرز الأمثلة على ذلك الدراسة التي استخدم فيها الحاسوب في تحليل ملحمة (الإلياذة: 15694 بيتًا شعريًا، 112000 كلمة) للتأكد من تفرد "هوميروس" بإنشائها[6]، فقد أكدت الدراسة الإحصائية من خلال تحليل الإلياذة، تجانس التوزيع الإحصائي للنغمة الشعرية[7].

2- التقييم الكمي لخصائص الأساليب الأدبية.

3-التحديد الموضوعي لمدى تأثر الأدباء والشعراء ممن سبقهم.

---

1 -Compréhension du contexte automatique de la langue.

2- نبيل علي،اللّغة العربية والحاسوب، سابق، ص 140.

3- المرجع نفسه، ص 141.

4- المرجع نفسه، ص 151.

5- المرجع نفسه، ص 142.

6- المرجع نفسه، الموضع نفسه.

7- المرجع نفسه، الصفحة نفسها.

ويعتمد تحليل الناتج اللّغوي الأدبي على انتقاء المؤثرات التي تدين بحكم طبيعتها للتقويم الإحصائي، مثل تواتر استخدام المفردات، والصيغ الصرفية، وأنواع الأساليب النحوية، وكلمات الربط كالمبينات في العربية، ويعدّ التحليل الإحصائي للأساليب الأدبية بوادر لدخول العلوم الأدبية مصاف العلوم المضبوطة[1].

5- ميكنة المعاجم[2]:

تستخدم المجامع اللّغوية إمكانية الحاسوب لميكنة المعاجم اللّغوية، وتشمل تطبيقات النّظم الآلية في المجال المعجمي الأغراض الرئيسية التالية:

أ- تخزين المعاجم إلكترونيًا[3]: على وسائط ممغنطة أو ضوئية، كالأقراص الممغنطة والأقراص الضوئية الرقمية (سي دي- روم)[4] والشرائح الإلكترونية، وذلك لاستخدامها مع المعالجات اللّغوية المختلفة وفي أغراض الترجمة الآلية، واكتشاف وتصحيح الأخطاء الإملائية[5]...الخ.

ب- تحليل العلاقات التي تربط بين مفردات المعجم: كالعلاقة بين جذور الكلمات والصيغ الصرفية (أو قواعد تكوين الكلمات) المنطبقة عليها[6]. أو العلاقات المتعلقة بمعاني الألفاظ، كعلاقات الترادف والاشتراك اللّفظي، والتضاد...الخ[7].

جـ- تحليل تعريفات المعجم: والمستخدمة في شرح معاني مفرداته، وتعدّ هذه الدراسات التحليلية ذات أهمية خاصة لبحوث الدلالة المعجمية ونظم الفهم الآلي، واسترجاع المعلومات، فهي تسعى إلى إبراز المفاهيم الأساسية، والأبعاد المحورية لمدلولات الألفاظ المترابطة والوصول إلى نواة هذه اللّغة أو قاعدة المفاهيم الأساسية التي يقام عليها المعجم[8].

1- المرجع السابق، ص 144.

2 -La mécanisation de dictionnaire.

3- جورج مصري، صناعة المعجم العلمي المختص من منظور اللسانيات الحديثة،اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد50، الرباط، المغرب، 1421هـ/2000م، ص 86.

4- CD.ROOM.

5- نبيل علي، اللغة العربية والحاسوب، سابق، ص 144.

6- المرجع نفسه، ص 144- 145.

7- المرجع نفسه، ص 145.

8- المرجع نفسه، الصفحة نفسها.

6- الترجمة الآلية[1]:

تعدّ الترجمة الآلية إحدى الغايات النهائية التي تصب فيها معظم روافد نظم التحليل والتركيب اللّغويين، وبجانب كونها تطبيقا قائما بذاته. ينظر البعض إليها كنموذج آلي شامل للمنظومة اللّغوية، وقاعدة فعالة للمعارف اللّغوية بالنسبة لبحوث اللّسانيات الحاسوبية[2].

تواجه الترجمة الآلية كثيرًا من العقبات والمشكلات اللّغوية من أهمها[3]: مشاكل الإبهام والازدواج الصرفي والمفرداتي الدلالين، التعقيد النحوي، والاختلاف المعجمي بين اللّغات، الصيغ المجازية المخالفة للقواعد، وما إلى ذلك من المشاكل اللّغوية التي تعترض تطوير نظم فعالة؛ بما فيها ضخامة حجم المعاجم[4].

7- في تعليم اللّغات:

اتسعت تطبيقات الحاسوب التعليمية وبخاصة في مجال تعلم اللّغات، بغرض تنمية حصيلة المفردات لدى المتعلم إلى تلك التي تتبع مناهج متكاملة لتعليم قواعد النّحو والصرف، ويواجه تعليم اللّغات[5] بالحاسوب عدّة مصاعب أهمها[6]:

- صعوبة محاكاة المواقف الطبيعية للاستخدامات اللّغوية لتأكيد مبدأ الوظيفة الاتصالية في اكتساب مهارات اللّغة.
- احتياج معظم البرامج الحاسوبية إلى حيز سعة تخزين هائلة لحفظ المادة التعليمية[7].

كما أن هناك فروع تطبيقية للسانيات الحاسوبية نذكر منها[8]:

- نظم استرجاع المعلومات.
- نظم قواعد المعارف (النظم الكبيرة).
- لغات البرمجة الراقية.
- نظم التعامل باللّغة الطبيعية.
- معمارية الجيل الخامس.
- تطبيقات الذكاء الاصطناعي.

---

1 -Machine Translation.

2- اللّغة العربية والحاسوب، سابق، ص 145.

3- عبد الله بن حمد الحميدان، مقدمة في الترجمة الآلية، سابق، ص 19.

4- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 146.

5- المرجع نفسه، ص 146.

6- المرجع نفسه، ص 146، 147.

7- المرجع نفسه، الموضع نفسه.

8- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 152.

ونلخصها في الجدول الآتي، الذي يمثل علاقات الترابط بين الفروع التطبيقية للسانيات الحاسوبية ومجالات استخداماتها[1]:

طـــرفـــا العـــلاقـــة		وصف العلاقة
مـــن	إلـــى	
استرجاع المعلومات نظم قواعد المعارف (النظم الكبيرة)	نظم قواعد المعارف (النظم الكبيرة) نظم استرجاع المعلومات	– تعد قواعد المعلومات مصدرًا أساسيًا من مصادر المعرفة. – استخدام النظم الكبيرة في عمليات البحث في قواعد المعلومات البيبلوغرافية والمصدرية.
نظم التعامل باللّغة الطبيعية	نظم قواعد المعارف (النظم الكبيرة)	تمثل نظم التعامل باللغة الطبيعية أحد عناصر المواجهة الرئيسية بين المستخدم والنظم الكبيرة، حيث يفضل التعامل معها باللّغة الطبيعية.
لغات البرمجة الراقية	لغات استرجاع المعلومات	تطوير لغات برمجة راقية لاسترجاع المعلومات
لغات البرمجة الرّاقية	نظم قواعد المعارف	استخدام لغات البرمجة الراقية "تصميم هياكل نمطية لنظم قواعد المعارف، والتي تستخدم لتصميم التطبيقات المختلفة للنظم الكبيرة".
لغات البرمجة الراقية	معمارية الجيل الخامس	استخدام لغة البرمجة الراقية (برولوج) في تطوير حواسيب الجيل الخامس.
معمارية الجيل الخامس	لغات البرمجة الرّاقية	تساعد على تصميم لغات برمجة تتصف بالمرونة، والقوة، يمهد ذلك إلى الترجمة بين لغات البرمجة الأخرى.
نظم قواعد المعارف	معمارية الجيل الخامس	تمثل قواعد المعارف إحدى الركائز الأساسية في تصميم وتطوير وتشغيل حواسيب الجيل الخامس.
نظم التعامل باللّغة الطبيعية	معمارية الجيل الخامس	تمثل نظم التعامل باللّغة الطبيعية وسيلة تواصل رئيسية بين المستخدم بنظم الحيل الخامس.
معمارية الجيل الخامس	تطبيقات الذكاء الاصطناعي	يمثل الجيل الخامس عنصر البنية التحتية لتطبيقات الذكاء الاصطناعي

الجدول رقم 06: علاقات الترابط بين الفروع التطبيقية للسانيات الحاسوبية ومجالات استخداماتها.

وهكذا فالرهان الحقيقي للسانيات الحاسوبية يتجلى أساساً في إمكانية التقريب والتنسيب بين المناهج والاستراتيجيات العلمية من جهة، وبين المتخصصين في هذا المجال في سائر دول العالم من جهة أخرى. ونعتقد أن البحث اللساني الحاسوبي الغربي قد راكم من الدراسات والبحوث النظرية والتطبيقية التي تهم مختلف خصائص اللغات الأوروبية والأمريكية والأسيوية،

1– المرجع السابق، ص 167، 168.

إلا أن نظيره العربي ما يزال في بدايته، ويحاول تلمس الطريق الصحيح، دون أن نلغي محاولات بعض الأفراد والجماعات المتفرقة في البلاد العربية أو الغربية، والتي حققت نتائج مهمة وأساسية، والتي لا يمكن التغاضي عنها بأي حال من الأحوال.

خامسا :اللسانيات الحاسوبية في خدمة الدراسات اللغة العربية: تجارب ورؤى مستقبلية:

1:التعاون والتنسيق بين اللغويين وعلماء الهندسة والحاسوب[1]:المطلب الذي لا بدّ منه:

لقد قطع البحث اللساني الحاسوبي العربي في السنوات الأخيرة مراحل مهمة، وحقق منجزات هي غاية في الأهمية سواء على مستويات المدقق الإملائي أو النحوي أو الصرفي أو ترميز الحرف أو الترجمة الآلية وبنوك المصطلحات[2] ... لكن إذا قارناه بنظيره الغربي سنجد ثمة فجوة بيننا وبين الآخرين فيما يتصل بالمعالجة الحاسوبية والهندسية للغة، فقد قطعت اللسانيات الحاسوبية أشواطا متقدمة في معالجة الألسن الغربية[3]. ولعل الأسباب كثيرة، قد سبق التلميح إلى بعضها.

ومن هنا نرى أن الحوار بين الإنسان العربي والآلة بواسطة اللغة العربية ما يزال في بداية الطريق، إن لم نقل، يبدو متعثراً. ومع ذلك فإننا لا ننقص من قيمة ما أنجزه العرب وغير العرب، أفراداً ومؤسسات، من أعمال مهمة[4].

وفي هذا المقام نلح على ضرورة تفعيل آلية الحوار بين اللغويين الخبراء المتمكنين من النظريات اللسانية الحديثة والقديمة ومن علوم اللغة العربية صرفها ونحوها وبلاغتها ومعجمها، والمهندسين ذوي الخبرة الرصينة في علوم الحاسوب وهندسة المعرفة والذكاء الاصطناعي[5].

ونعتقد أن هذا الاقتراح يتقاطع مع من يدعو إليه البعض باستحداث مشروعات بحثية[6] مشتركة بين تخصصات متعددة يشارك فيها علماء اللغة إلى جانب علماء الهندسة والمعلومياتية

1- مازن الوعر، اللسانيات والعلم والتكنولوجيا " نحو تعريب موحد للسانيات التطبيقية العربية وبرمجتها في الحاسبات الإلكترونية"، دورية اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 22، الرباط، المغرب، 1983/1982، ص 20،21.

2- علي القاسمي، نحو تطوير بنوك المصطلحات "أداة للبحث المصطلحي والعلمي"، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 28، الرباط، المغرب، 1987، ص 80،79.

3- خليفة الميساوي، المصطلح اللساني وتأسيس المفهوم، دار الآمان، الرباط ، الطبعة الأولى، 1434 هـ ،2013 م ، ص 31.

4- للمزيد من بسط القول ينظر: نبيل علي، اللغة والحاسوب، مرجع سابق، ص 537،536.

5- إن التعاون والتنسيق في هذا الميدان – إن تَمَّ – ستكون نتائجه غاية في الأهمية، وستدفع به خطوات واثقة إلى الأمام، وسوف تتغير قناعات ومسلمات كانت أشبه بالحقائق التي لا تقبل الجدل والنقاش حولها، كما ستتأكد قضايا في اللغة كانت أقرب ما تكون إلى الاحتمال والرجحان والظن، وهذا ما توفره اللسانيات الحاسوبية للعلوم الإنسانية بعامة .

والواقع أن هذا الأمر، أعني التقاء علماء اللغة وعلماء الهندسة والحاسوب، كان – ومازال – مطلباً ملحاً، وضرورة قصوى، لأي عمل ناجع في هذا الميدان على وجه الخصوص، وهو ما أكدت عليه جميع المؤتمرات التي عقدت لمناقشة قضية الحاسوب واللغة العربية.

6- اللغة والحاسوب، مرجع سابق، ص 539،538،...550.

والذكاء الاصطناعي والشبكات العصبية للقيام بأبحاث خلاقة حول معالجة اللغة آلياً مثل استحداث وسائل متقدمة للفهرسة، التكشيف والتخليص والبحث، واستخدام أدوات الذكاء الاصطناعي في تحليل النصوص العربية وفهمها آلياً[1].

لحسن الحظ أنّ عددًا من الأعمال اللّسانية العربية أصبحت توفر نماذج للدراسة النّحوية والصرفية الكافية والمصورنة التي يمكن اعتمادها لبناء قواعد لغوية تامة. فهي توفر فرصة للساني الحاسوبي العربي لتوظيفها في حوسبة اللّغة.

فلا يمكن تقدير أبعاد التحول الذي بدأ يطال اللغة العربية اليوم دون استقراء بعض الجهود الأولية لواضعي أسس هذا التحول من الباحثين المنظرين الذين ما فتئوا يدعون إلى ضرورة اعتماد التكنولوجيا ورعايتِها، وينبّهون على مدى خطورة الإبطاء في استيعاب "الثقافة المعلوماتية" وعدم تضمينها في المنظومتين التعليمية والإنتاجية للعالم العربي، وإذا أردنا الحديث عن جهود بعض العلماء العرب المعاصرين في مجال علم اللغة الحاسوبي يمكن القول بأن كتاب الدكتور نبيل علي(*) (اللغة العربية والحاسوب) يُعَدُّ أول مؤلَّف يتناول موضوع اللسانيات الحاسوبية مطبقةً على أنظمة اللغة العربية، صوتاً، وصرفاً، ونحواً، ومعجماً، مع المعالجة الآلية لهذه النظم اللغوية جميعها.

وكان تاريخ صدوره لأول مرة سنة 1988م[2]. وقد حالف التوفيق المؤلف في كثير من القضايا المتصلة بالحاسوب واللغة، وذلك حينما انطلق في عمله هذا من وضع دراسات تقابلية بين العربية والإنجليزية شاملة لكل النظم اللغوية، بالنظر إلى أن الإنجليزية هي اللغة الأم لتقنيات نظم الحاسوب والمعلومات، وهذا ما نتج عنه معرفة أوجه الاختلاف والاتفاق بين اللغتين، وكان هذا النهج بمثابة الأرض الصُّلْبة والقاعدة المتينة التي هيأت للمؤلِّف منهجية وموضوعية، ومكَّنته من الإسهام الإيجابي في جهود تعريب الحاسوب من جهة، والمعالجة الآلية للغة العربية من جهة أخرى.

1- المرجع السابق، ص 252.

(*) - الدكتور نبيل علي أحد أبرز المتخصصين في بحوث اللغويات الحاسوبية، ويعمل منذ سنوات طويلة في مجال الكمبيوتر ونظم المعلومات، برمجة، وتصميماً، وإدارةً، وبحثاً.وهو صاحب فكرة مشروع كمبيوتر صخر والعالمية للبرامج. له من المؤلفات: اللغة العربية والحاسوب، العرب وعصر المعلومات، الثقافة العربية وعصر المعلومات، صورة الثقافة العربية والحضارة العربية والإسلامية على الانترنت .

2- صدر عن مؤسسة تعريب،الكويت، سنة 1988م. وكان المؤلف قد مهد لصدور هذا الكتاب ببحث نشره في مجلة عالم الفكر، المجلد 18، العدد 3، 1987م، ص59-118 بعنوان (اللغة العربية والحاسوب)، ثم صدر هذا الكتاب في طبعته الثانية سنة 1998م عن دار غريب بالقاهرة .ولهذا الكتاب عرض تحليلي للدكتور علي فرغلي، نشره في مجلة عالم الفكر (الكويت)، المجلد 20، العدد 3، 1989م، ص255-278، كما أن له مراجعة أخرى للدكتور نهاد الموسى نشرها في المجلة العربية للعلوم الإنسانية (جامعة الكويت)، العدد 38، المجلد 10، 1990م، ص244-251 .

ولعلي لا أبالغ في القول بأن هذا هذا الكتاب يمثل حجر الزاوية في مسيرة البحث اللغوي العربي في اللسانيات الحاسوبية، بل إنه كما وصفه الدكتور نهاد الموسى – بحق – «خطوة واسعة واثقة، تنتظم مشروعاً مستوعباً لتأسيس اللسانيات الحاسوبية في العربية، على أساس نظري وتطبيقي في آن واحد معاً»[1] .

وبعد نشر هذا الكتاب بسنوات، صدر كتاب الدكتور عبد ذياب العجيلي (الحاسوب واللغة العربية)[2]، وهو – كما يقول الدكتور نهاد الموسى – : «خطوة جزئية إيجابية نحو معالجة مسائل متنوعة من العربية بلغة برولوج Prolog .وهو يمثل جهداً حميداً في هذا الاتجاه البيني (اللسانيات العربية الحاسوبية)»[3] .

وفي سنة 2000 م صدر كتاب الدكتور نهاد الموسى (العربية – نحو توصيف جديد في ضوء اللسانيات الحاسوبية)[4]. والكتاب – كما يذكر مؤلفه – « محاولة في الانتقال من وصف العربية إلى توصيفها، وذلك في ضوء الأطروحة العامة للسانيات الحاسوبية»[5] .

وقد اشتمل الكتاب على رؤى حاسوبية حاول المؤلف إسقاطها على أنظمة العربية، وخاصة النحو (الإعراب)، والصرف (البنية)، والمعجم، إضافة إلى التصويب اللغوي (الأخطاء النحوية، والصرفية، والإملائية)[6]. كما يعد كتاب "المعالجة الآلية للغة العربية : المشاكل والحلول " للدكتورة سلوى حمادة[7] والذي أشرف على إصداره معهد بحوث الاليكترونيات التابع للمركز القومي للبحوث في القاهرة، أول كتاب لباحثة عربية في مجال اللسانيات الحاسوبية. وقد

1– نهاد الموسى، كتاب اللغة العربية والحاسوب لنبيل علي (مراجعة) المجلة العربية للعلوم الإنسانية (جامعة الكويت)، العدد 38، المجلد 10، 1990م، ص251 .

2– صدر هذا الكتاب سنة 1996م، عن جامعة اليرموك – إربد (الأردن) .

3– نهاد الموسى، العربية ، نحو توصيف جديد، مرجع سابق، ص45 .

4– صدر عن المؤسسة العربية للدراسات والنشر، بيروت .

5– نهاد الموسى، العربية، مرجع سابق، ص 288 .

6– ومن الأعمال الرائدة في ميدان اللسانيات الحاسوبية، عمل نهاد الموسى بعنوان "العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية". يشبه هذا الكتاب أن يكون سبراً للدماغ البشري حين يستقبل اللغة العربية ويفهمها وينتجها، إذ أنه ينفذ إلى أعماق ما يدور في دماغ المتحدث العربي متتبعاً بدقة متناهية كيفية تشكل الأداء اللغوي على نحو مضبوط. وغايته أن ينقل المعرفة باللغة العربية من اللاوعي إلى الوعي، ثم تقديمها للحاسوب، قصد أن يهيئ له كفاية لغوية تشبه ما يكون للمتحدث العربي.

7– الدكتورة سلوى حمادة باحثة في معهد بحوث الإلكترونيات بمصر، وحصلت على الماجستير من كلية الهندسة جامعة عين الشمس بعنوان:" تطويع اللغة العربية للحاسب الآلي"، وكذلك الدكتوراة بعنوان : "تصميم نموذج لتفهم معاني اللغة العربية عن طريق الحاسبات" . وتعمل محاضرة – جزءا من الوقت – بجامعة القاهرة، كلية الحاسبات . وهي الباحثة الأولى التي استعملت اللغة العربية في أبحاثها المنشورة في مجال الهندسة اللغوية وتقنية المعلومات وقد تعدت أربعين بحثا. وقد ركزت اهتمامها بأن تزيد عدد الأبحاث المنشورة باللغة العربية، و أن تحصل على قاعدة واسعة من قرّاء العربية. كتبت مقالات عديدة عن المناهج الحاسوبية لمعالجة اللغات الطبيعية، وركزت اهتمامها على اللغة العربية.

وقد دعيت إلى العديد من المؤتمرات الدولية في تونس، الإمارات، ولبنان، والمملكة العربية السعودية، وصبت جل اهتمامها بشكل رئيسي بتحليل اللغة العربية، وتقييم برمجياتها، والمعجم العربي، والتعليم الإلكتروني، والبحث على الشبكات، والترجمة الآلية. وقد رشحت لعدة جوائز علمية نالت جائزة الكندي لأفضل باحث عربي معلوماتي من سوريا، تعتبر الباحثة عضوة في عدة جمعيات لغوية داخل وخارج مصر والوطن العربي. أشرفت د.سلوى على عدد من الرسائل العلمية في المجال، و دعيت لتحكيم مشاريع عربية وعالمية أهمها مشروع أفضل محلل صرفي عربي و الذي تولته المنظمة العربية للتربية والثقافة والعلوم بتونس.

أشارت الدكتورة سلوى حمادة أن الكتاب خارطة طريق إلى البحوث الرامية إلى إنجاز محلِّلٍ صرفي مفتوح المصدر للغة العربية، يمتاز بدقته العالية وسهولة تطويره وإمكان دمجه في مستوياتٍ أخرى من معالجة اللغة العربية وتطبيقاتها المختلفة التي منها تحليل النصوص والترجمة الآلية من اللغات الاجنبية إلي اللغة العربية والعكس حتى تأخذ هذه اللغة حقها وتصبح لغة عالمية عبر الانترنت. حيث تركز المؤلفة في الباب الثاني على مشكلات حوسبة الصرف والاشتقاق، وتبحث فى توليد العديد من الكلمات من خلال أوزان ومشتقات مختلفة.

أما البحوث والمقالات الخاصة باللسانيات الحاسوبية[1]، فمنها ما نشر في مجلات علمية، ومنها ما أُلقي أو قُدِّم في الندوات والمؤتمرات التي خُصِّصت أصلاً للغويات الحاسوبية، أو اللسانيات التطبيقية، أو لتكنولوجيا الحاسوب ومجالات استخدامه في العلوم الإنسانية، ثم نشرت هذه البحوث ضمن أعمال تلك المؤتمرات والندوات[2].

1- ولعل جلّ أصحاب تلك البحوث والمقالات هم من اللغويين الأكثر حضوراً وفاعليةً على الساحة اللغوية، "من أمثال الدكتور محمد الحناش (المغرب)، والدكتور محمود إسماعيل صيني (السعودية)، والأستاذ أحمد الأخضر غزال (المغرب)، والدكتور عبد القادر الفاسي الفهري (المغرب)، والدكتور مازن الوعر (سورية)، والدكتور محمود فهمي حجازي (مصر)، والدكتور عبد الرحمن الحاج صالح (الجزائر)، والدكتور سالم الغزالي (تونس)، والدكتور داود عبده (الأردن)، وبعضهم من المتخصصين في الحاسوب أو الهندسة الحاسوبية، كالدكتور يحيى هلال (المغرب)، والدكتور محمد مراياتي (سوريا)، والدكتور نبيل علي (مصر)، والدكتورة نادية حجازي (مصر). ويلاحظ على أغلب هذه البحوث أنها انتقلت باللسانيات الحاسوبية من مجالها النظري أو التنظيري إلى الجانب التطبيقي، وهو تطور إيجابي يُحسب لأصحاب هذا الاتجاه"، ينظر : توظيف اللسانيات الحاسوبية في خدمة الدراسات اللغوية العربية ــ جهودٌ ونتائج،بحثٌ قُدِّم لمؤتمر (حضارة الأمة وتحدي المعلوماتية)، المنعقد في كلية الآداب بجامعة الزرقاء الأهلية (الأردن)، في الفترة 18ــ 20 مايو 2004م، الموافق 3/29 – 4/1/ 1425هـ.،ونُشر في مجلة مجمع اللغة العربية الأردني،العدد73، 1428هـ – 2007م،ص47 – 96.

2- لعل من أهم المؤتمرات والندوات التي تناولت بحوثها قضية اللسانيات الحاسوبية العربية ما يلي :

أ – ندوة استخدام اللغة العربية في تقنية المعلومات، التي نظمتها مكتبة الملك عبد العزيز العامة بالرياض، سنة 1992م.

ب – ندوة استخدام اللغة العربية في الحاسب الآلي، التي عقدت بالكويت عام 1985م .

ج – مؤتمر اللسانيات التطبيقية العربية ومعالجة الإشارة والمعلومات، الذي عقده المركز القومي للتنسيق والتخطيط للبحث العلمي والتقني ومعهد الدراسات والأبحاث للتعريب بالمغرب (الرباط) عام 1983م (صدرت أعماله في كتاب بالإنجليزية والفرنسية) .

د – المؤتمر الأول والثاني للغويات الحسابية العربية، اللذان عقدا في الكويت، سنة 1989م

هـ– ندوة اللغويات الحسابية العربية، التي عقدت في القاهرة، سنة 1992م .

و – أعمال المناظرة المنظمة بمؤسسة الملك عبد العزيز آل سعود للدراسات الإسلامية والعلوم الإنسانية تحت عنوان : اللغة العربية والتقنيات المعلوماتية المتقدمة، الدار البيضاء، 1993م .

ز – مؤتمر استخدام اللغة العربية في المعلوماتية، الذي عقدته المنظمة العربية للتربية والثقافة والعلوم، تونس، 1996م .

ح – الملتقى الرابع للسانيات (اللسانيات العربية والإعلامية) الذي عقده مركز الدراسات والأبحاث الاقتصادية والاجتماعية بالجامعة التونسية سنة 1987م .

ط – المؤتمر الدولي الثاني في اللغة والترجمة (دور التكنولوجيا الحديثة في تعلم اللغات وتعليمها) الذي عقده مركز أطلس العالمي للدراسات والأبحاث بعمان، في 14 – 15 /2002/12م (صدرت أعماله في كتاب سنة 2005م) .

فتلك الجهود على كثرتها – بل يستحيل – حصرها في مقام كهذا، كون هذه البحوث والمقالات قد أضحت تمثل تياراً واضحاً في الجهود اللسانية الحاسوبية، وهذا ما جعل كلاً من الدكتور وليد العناتي و زميله الدكتور خالد الجبر، من جامعة البترا الأهلية (الأردن)، يقومان بوضع دليل بيبليوغرافي لها أسمياه (دليل الباحث إلى اللسانيات الحاسوبية العربية)[1]، حاولا فيه أن يستقصيا جميع ما وقفا عليه من أعمال علمية تنتظم في ميدان اللسانيات الحاسوبية[2].

الدليل..... من كتاب ورقي إلى موقع إلكتروني:

ومن الأعمال الرائدة في ميدان اللسانيات الحاسوبية، عمل د. وليد أحمد العناتي[3] بعنوان الدَّليلُ نحو بناء قاعدة بيانات[4] للِّسانيّاتِ الحاسوبيّة العَرَبِيّة يقصد من خلال هذا العمل إلى التعريف بكتاب "دليل الباحث إلى اللسانيات الحاسوبية العربية" الذي أَنْجَزَهُ وزميليه د. خالد الجبر فهو

1– لتفاصيل وافية عن ما أُنْجِزَ في اللسانيات الحاسوبية العربية؛ ينظر "دليل الباحث إلى اللسانيات الحاسوبية العربية" للدكتور وليد العناتي و الدكتور خالد جبر. وهو يمثِّل، أولاً: مفتاحاً للسانيات الحاسوبية العربية من حيث هي خطوة إضافية على طريق وصف العربية واستبطان نظامها في إنشائها وتحليلها، وهو، ثانياً: إسهام في الطَّمْح العربي لدخول عصر اقتصاد المعرفة واستثمار حوسبة العربية في تطبيقات عملية وعلمية وتعليمية وإعلامية وعالمية" وأحسب أنه سيكون أول مصدر جامع يتضمن جُلَّ ما أنجزه الباحثون العرب في حوسبة العربية و نَمْذَجَتِها وتطويعها للحوسبة. ولا شك في أنه سيمثل مقدمة مهمة لدخول عصر المعرفة، وتهيئة قاعدة بيانات، ولو كانت ورقية، تختصر على الباحثين وقتاً طويلاً وتحفظ لهم كثيراً من الجهد؛ إذ يحملهم على تجنب القول المعاد المكرور. و قد قدَّم د. نهاد الموسى لهذا الكتاب بالقول: "ويقدِّم الدكتور وليد والدكتور خالد بهذا الدليل المستفيض في سياق واحد جامع لكل باحث في اللسانيات الحاسوبية العربية جُلَّ ما أنجزه المشتغلون بهذا الحقل. وهما يَسُدّان بذلك حاجة مُؤَرِّقة كان الباحث العربي يواجهها وهو يلتمس مصادر موضوعه في المظانِّ المتنائية يتوسل بها بالمراسلات المؤجلة ومخاطبة المكتبات المتعددة وتحرِّي الدوريات المبعثرة". ينظر: "دليل الباحث إلى اللسانيات الحاسوبية العربية"،وليد العناتي،الدكتور خالد جبر،دار جرير للنشر والتوزيع،الطبعة الأولى، 1428هـ – 2007م، عمان، الأردن، ص 12.

2– سنشير نحن هنا إلى أهم تلك البحوث والمقالات:

– بحث بعنوان "توظيف قواعد النحو والصرف في بناء محلل صرفي للغة العربية " إعداد: مجدي صوالحه وإيرك أتول، جامعة ليدز، المملكة المتحدة. بحث مقدّم لاجتماع خبراء المحلّلات الحاسوبية الصرفيّة للغة العربية, دمشق, 2009.

– بحث بعنوان" سبل تطوير محلل الصرف الآلي المستخدم في حوسبة الموسوعات العربية" إعداد: حسن مظفر الرزّو.

– دراسة نقدية لغوية لعدد من المحللات الصرفية، مروان البواب، عضو مجمع اللغة العربية بدمشق.

– اللغة العربية والحاسوب : قراءة سريعة في الهندسة اللسانية للأستاذ الدكتور محمد الحناش،مجلة التواصل اللساني، مج 9،س 2003

– " نظام آلي لفهم اللغة العربية " إعداد: ميساء عبد الكريم ناصر ومرتضى عباس فالح وصبا عبد الواحد، مجلة بابل,المجلد(الرابع عشر) العدد الثالث

– " تطوير وتصميم وإنتاج محرك بحث اشتقاقي خاص بالنص القرآني" إعداد: أ.د محمد رفعت علي الحفني، أعمال ندوة القرآن الكريم والتقنيات المعاصرة، وزارة الشؤون الإسلامية والأوقاف والدعوة والإرشاد.

– " تصميم نظام استخلاص المعلومات من بعض فقرات النصوص العربية " بحث مستل من أطروحة دكتوراه، م.د علياء سلمان صابر، م.د.سلمى عبد الباقي محمود، أ.م.د علي فاضل مرهون.

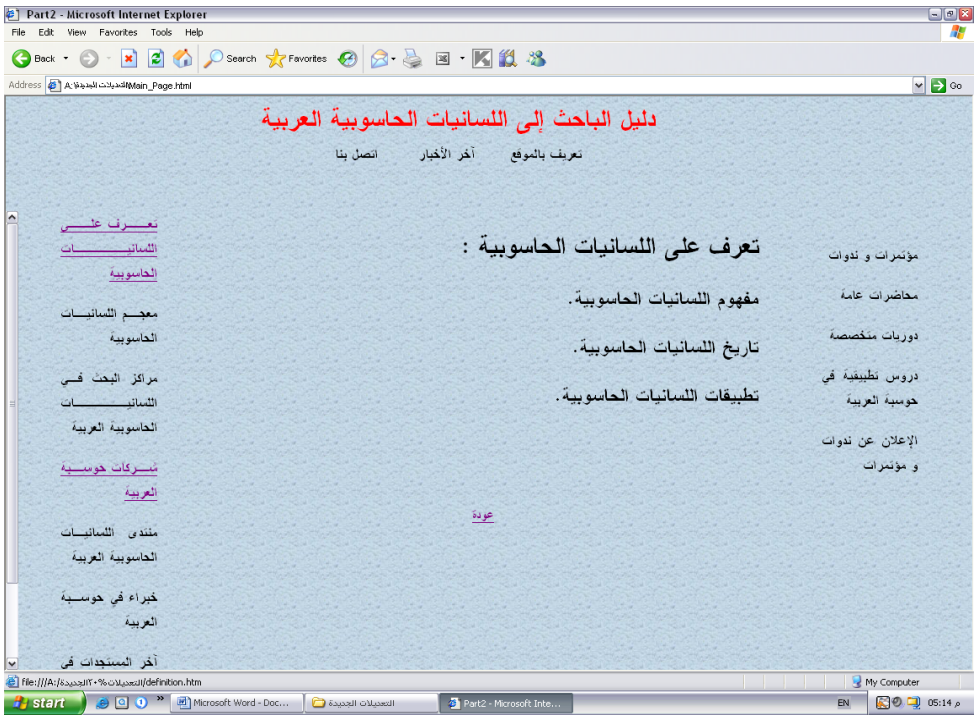
– المحللات الصرفية للغة العربية، إعداد: د. سلوى السيد حمادة، بحث مقدّم لاجتماع خبراء المحلّلات الحاسوبية الصرفيّة للغة العربية , دمشق, 2009.

3– أستاذ اللسانيات المشارك في قسم اللغة العربية بجامعة البترا الأردنية الخاصة.

4– و تُعَرَّفُ قاعدة البيانات بأنها " مجموعة من المعلومات أو البيانات المتصلة، ذات العلاقة المتبادلة فيما بينها المخزنة بطريقة نموذجية ودون تكرار". ينظر: منيب قطيشات، قواعد البيانات، ط2، دار وائل للنشر والتوزيع، عَمَّان،2005، ص23.

يطمح في منتهى غايته أنْ يدل على أهمية[1] تحويله من كتاب ورقيّ إلى قاعدةِ بياناتٍ وموقعٍ إلكترونيٍ على الإنترنت" نحو قاعدة بياناتٍ للسانيات الحاسوبية العربية"، ويتناول بتركيز كبير كيفية تحويل الدليل من كتاب ورقي إلى قاعدةِ بيانات أولاً، وكيفية تحويله إلى موقعٍ إلكترونيٍ على الإنترنت. مستعرضا في ثنايا العرض، أهمية نشر الدليل نشراً إلكترونياً مبرزا في الوقت نفسه ماحققته اللسانيات الحاسوبية من إنجازات هائلة في مجال حفظ المعلومات وتوثيقها وفهرستها وترتيبها واسترجاعها، وتواترت مراحل التطور حتى بلغت درجة عالية من الدقة والتنظيم ومدى الاستيعاب متاحاً للجميع[2]. وتأسيساً على ذلك سيكون **موقع "دليل الباحث إلى اللسانيات الحاسوبية العربية"** الموقع الرئيسي الذي يفيء إليه الأكاديميون وخبراء الحوسبة والشركات وجميع المهتمين؛ إذ سيوفر لهم جُلَّ ما أُنْجِزَ في حوسبة العربية من البحوث النظرية والتطبيقية وتجارب الشركات المختلفة، وتمثل اللوحات التالية نماذج من صفحات الموقع المقترح[3]:

الشاشة الأولى:

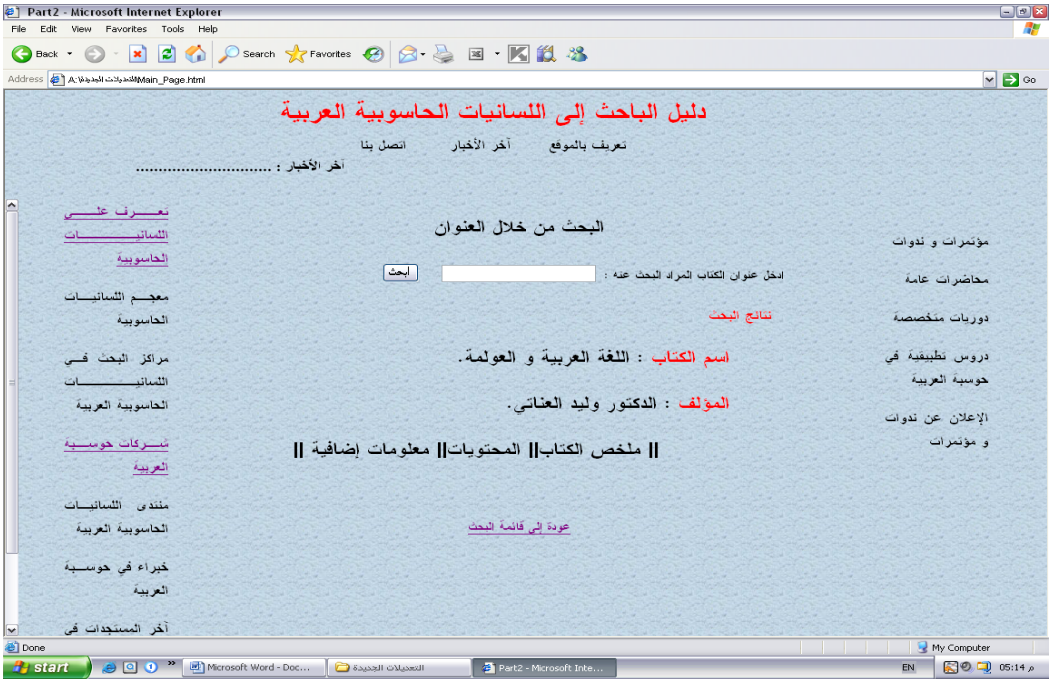


1- ينتهي تحويل "دليل الباحث إلى اللسانيات الحاسوبية العربية" إلى قاعدة بيانات إلى فائدة عظيمة تتجاوز الفائدة المحَصَّلة من الكتاب في صورته الورقية؛ إذ إنه يسهل على الباحث أن يسترجع معلوماتٍ محددة ودقيقة في وقت قصير جداً، في حين أن استرجاع هذه المعلومات من الكتاب الورقي قد يستغرق ساعات طويلة. ومن ذلك أنَّ قاعدة البيانات ستوفر تطبيقاتٍ جديدةً ليست مستعملة في التوثيق التقليدي، كترتيب البحوث حسب تسلسلها التاريخي، ورصد مُنْجَزِ دولة على التعيين، و تقديم ملخصات البحوث.

2- وليد أحمد العناتي، الدّليلُ نحو بناء قاعدة بيانات للِّسانيّاتِ الحاسوبيّة العَرَبِيّة،، ندوة تقنية المعلومات والعلوم الشرعية والعربية، ندوة (تقنية المعلومات والعلوم الشرعية و العربية) في كلية علوم الحاسب والمعلومات بجامعة الإمام في 16-17/صفر/1428هـــ ـ الموافق 6-7/مارس 2007م، ص 671.

3- الدّليلُ نحو بناء قاعدة بيانات للِّسانيّاتِ الحاسوبيّة العَرَبِيّة، المرجع نفسه، ص 709،710،711.

الشاشة الثانية:



وبلا شك فإن هذا الموقع الإلكتروني سَيَسُدُّ – بعد صدوره إن شاء الله – ثغرة واضحة في المكتبة اللغوية الإلكترونية بعامة، واللسانيات الحاسوبية بخاصة[1].

ويلاحظ على أغلب هذه البحوث[2] أنها انتقلت باللسانيات الحاسوبية من مجالها النظري أو التنظيري إلى الجانب التطبيقي، وهو تطور إيجابي يُحسب لأصحاب هذا الاتجاه . ومن

1– إنَّ قاعدة البيانات "دليل الباحث إلى اللسانيات الحاسوبية العربية" ستوفر للباحثين معلومات قيمة تتناول ما يتعلق بحوسبة العربية ومعالجتها وخبرائها وبرامجها التطبيقية والمشكلات التي تواجهها العربية في سياق الحوسبة. ولعل تحويل قاعدة بيانات "الدليل" إلى موقع إلكتروني متكامل سيكون أجدى وأنفع للباحثين والدارسين وطلبة الدراسات العليا والشركات، وسيكون أجدى وأنفع للغة العربية أيضاً، ولعل ذلك راجع إلى طبيعة الفرق بين قاعدة البيانات و الإنترنت من الناحية التقنية ومن ناحية الاستعمال؛ فقاعدة البيانات لا تكون متاحة أمام الجميع وفي الوقت الذي يختارونه، فهي مقيدة، غالباً، بمدير هذه القاعدة الذي يتحكم في إدارتها والنفاذ إليها وتنظيم أمورها حذفاً وزيادةً وتحديثاً، أما الموقع الإلكتروني على الشبكة فإنه متاحٌ للجميع، غالباً وإذا لم يكن هناك اشتراك، إذ يستطيع المُتَصَفِّحُ الدخولَ إلى الموقع متى شاء وتحصيل جميع المعلومات التي يريدها. كذلك فإن قاعدة البيانات غالباً ما توفر معلومات بيانية مختصرة وتمثل مفاتيح للمعرفة، أما الموقع الإلكتروني فإنه يتيح لك تحصيل المعلومات كاملة وبالتفصيل؛ إذ يمكنك قراءة نص البحث أو الكتاب كاملاً، وينضاف إلى ذلك تميز الإنترنت بالتفرع والانفتاح على خيارات كثيرة ومتعددة قد لا تكون في حسبان المُتَصَفِّحُ، وهي بذلك توفر له فرصاً عظيمة لتحصيل المعلومات من جميع مصادرها وبلغات متعددة. ويمكن في كثير من المواقع إضافة معلومات ومساهمات جديدة بالنظر في خيارات الإضافة المتاحة التي يسهل الوصول إليها بالنقر على الأيقونة ثم اتباع التعليمات المنصوص عليها .

2– ذكرت بعض الجهود في اللسانيات الحاسوبية وتوظيفها لخدمة العربية – وهناك الكثير من تلك الإسهامات لم أتمكن من الاطلاع عليها، أو عرضها في هذه الدراسة – فإن الطريق في هذا المضمار مازال شاقاً وطويلاً، والأمل معقود على جميع الباحثين الذين أوقفوا أنفسهم على هذا المجال النادر من مجالات العلم والمعرفة الإنسانية، أن تتكاتف جهودهم لتذليل العقبات وحل المشكلات التي تحيط بلغة القرآن الكريم إزاء الثورة المعلوماتية – الحاسوبية المعاصرة. والمحصلة النهائية لهذه الجهود تصب في خانة قدرة العربية على استيعاب لغة العصر، وتمثل تقنياته التكنولوجية بكل كفاية واقتدار، وهذه قضية من القضايا التي واجهتها – وما زالت تواجهها – كينونة الأمة العربية وحضارتها اللغوية، وهويتها الثقافية .

المعروف أن الجانب التطبيقي – وهو الجانب الأهم في اللغويات الحاسوبية – يتمثل في تسخير العقل الإلكتروني لحل القضايا اللغوية، وهنا يبرز الدور الرئيس والأثر الفاعل لالتقاء اللغويين والحاسوبيين، والتعاون فيما بينهم، وما يثمر عنه من نتائج تسهم إلى حد كبير في تذليل العقبات وحل المشكلات التي تواجه التحليل الحاسوبي للغة العربية.

إذن فاللّسانيات الحاسوبية كدراسة تطبيقية تعتمد على مجالين، المعرفة اللّسانية، والمعرفة الحاسوبية، أي أن هناك حوار بين اللّساني والحاسوبي[1].

فالتقدم وتطوير مجال اللّسانيات الحاسوبية متوقف على التعاون والتحاور المتين بين الطرفين، بعيدًا عن الأهداف التجارية المحضة التي غالبًا ما تقف عائقًا دون وجود حوار فعلي وصريح[2].

إن علم "اللّسانيات الحاسوبية"[3] وهندسة اللّغة، صاحبهما ثورة علمية حقيقية في معظم فروع اللّسانيات ما زال صداها يتردد إلى الآن، حتى أيقن الجميع أن دخول اللّغة مصاف العلوم المضبوطة شرط أساسي لكي تتبعها في ذلك العلوم الإنسانية بمختلف مجالاتها[4].

وهذا ما يؤكده د. نبيل علي من خلال استخدام اللغة كأداة للحاسوب و تعدد مجالات الاستخدام[5] من:

- "استخدام اللّغة في نظم استرجاع المعلومات".
- استخدام اللّغة في نظم قواعد المعارف (النظم الخبيرة).
- استخدام اللّغة في تصميم لغات البرمجة الراقية.
- استخدام اللّغات الطبيعية في التعامل المباشر مع الحاسوب.
- معمارية الجيل الخاص.
- تطبيقات الذكاء الاصطناعي.

لا شك أن اللغة العربية هي المستفيد الأول من استخدام تقنية الحاسوب، وأن الحاسوب يمكن تطويع آلياته وأنظمته لتتواءم مع خصوصية اللغة العربية، على جميع مستوياتها اللغوية، الصوتي منها، والصرفي، والنحوي، والمعجمي، والدلالي. وتم بجهود الباحثين العرب، اللغويين منهم والحاسوبيين، تمثيل الكلام المنطوق وتوليده آلياً، وتحليل الكلمات المفردة وتركيبها آلياً،

1– نهاد الموسى، العربية – نحو توصيف جديد...مرجع سابق، ص 26.

2– المرجع نفسه، الصفحة نفسها.

3- Linguistique informatique.

4– نبيل علي، (العرب وعصر المعلومات) عالم المعرفة، رقم 184، ص 331.

5– نبيل علي، اللغة العربية والحاسوب، دراسة بحثية، ص152.

وتوصيف الجمل وتوليدها وإعرابها آلياً، وقراءة النصوص المكتوبة وتصحيحها ومعالجتها آلياً، وصناعة المعاجم الآلية، وإنشاء البنوك المصطلحية، وتصويب الأخطاء النحوية والصرفية والإملائية آلياً، وتصميم البرامج الحاسوبية للترجمة الآلية، وتعليم اللغة العربية لأبنائها وغير أبنائها بوساطة الحاسوب .

إن هذه الجهود مؤشر حقيقي على نجاح الحاسوب في خدمة اللغة العربية، وتوظيفه في معالجة قضاياها المختلفة، تحليلاً، وتوليداً، وترجمة، وتعليماً، وصياغتها صياغة رياضية دقيقة وفق علاقة متبادلة بين المقاييس العلمية والمقاييس اللغوية .

والمحصلة النهائية لهذه الجهود تصب في خانة قدرة العربية على استيعاب لغة العصر، وتمثل تقنياته التكنولوجية بكل كفاية واقتدار، وهذه قضية من القضايا التي واجهتها – وما زالت تواجهها – كينونة الأمة العربية وحضارتها اللغوية، وهويتها الثقافية[1].

و من هنا يجب التركيز على بعض الجزيئات ، أحسب أنها تسهم في توطين هذا التوجه في الدراسات اللغوية المعاصرة، وتدفع به نحو الأمثل والأفضل، وهي على النحو الآتي[2] :

أولاً : أن تتضافر الأعمال في مجال اللسانيات الحاسوبية العربية وتتآزر بين اللغويين والحاسوبيين في أي مشروع علمي يهدف إلى برمجة الأنظمة اللغوية للعربية، وتحليلها، ومعالجتها آلياً[3]. وأرى في هذا المقام أن أي عمل منفرد، أو مستقل عن الطرف الآخر من المعادلة، أو غير منسق بين هذه الفئات العلمية، إنما هو بمثابة جهد ضائع لا طائل من ورائه.

ثانياً : ترجمة جميع الأعمال العلمية في مجال اللسانيات الحاسوبية العربية، التي كتبت باللغات الإنجليزية، والفرنسية، والألمانية، وغيرهن من اللغات، التي وضعها باحثون عرب وأجانب، ونقلها إلى العربية.

1– إن علم اللسانيات الحاسوبية لا يزال خصباً يعوزه العمل الدؤوب الجاد، وإن ما أسعدتنا به الجهود المبذولة في اللسانيات الحاسوبية وتوظيفها لخدمة العربية تبقى حالمة ترتقب المزيد من إسهامات اللسانيين والحاسوبيين على السواء، فرغم أن الطريق في هذا المضمار مازال شاقاً وطويلاً، والأمل معقود على جميع العلماء والباحثين الذين أوقفوا أنفسهم على هذا المجال النادر من مجالات العلم والمعرفة الإنسانية، أن تتكاتف جهودهم لتذليل العقبات وحل المشكلات التي تحيط بلغة القرآن الكريم إزاء الثورة المعلوماتية – الحاسوبية المعاصرة ، فاللغة العربية لا زالت تناشد أبناءها خدمتها من أجل تمكين الحاسوب احتواءها احتواءً يلبي أغراض أهلها، ويواكب مطالب التقنية المتسارع خطاها. وسيتخلف أهل العربية عن الركب ما لم يشمروا عن ساعد الجد، وما لم يجترؤوا خوض العالم الرقمي الحديث.

2– عبد الرحمن بن حسن العارف، توظيف اللسانيات الحاسوبية في خدمة الدراسات اللغوية العربية،" جهود ونتائج"،اللسانيات، مجلة في علوم اللسان وتكنولوجياته،مركز البحث العلمي والتقني لتطوير اللغة العربية، العددان 12،13،الجزائر،2007، ص 37.

3– إن التعاون والتنسيق في هذا الميدان – إن تَمَّ – ستكون نتائجه غاية في الأهمية، وستدفع به خطوات واثقة إلى الأمام، وسوف تتغير قناعات ومسلمات كانت أشبه بالحقائق التي لا تقبل الجدل والنقاش حولها، كما ستتأكد قضايا في اللغة كانت أقرب ما تكون إلى الاحتمال والرجحان والظن، وهذا ما توفره اللسانيات الحاسوبية للعلوم الإنسانية بعامة .

والواقع أن هذا الأمر، أعني التقاء علماء اللغة وعلماء الهندسة والحاسوب، كان – ومازال – مطلباً ملحاً، وضرورة قصوى، لأي عمل ناجع في هذا الميدان على وجه الخصوص، وهو ما أكدت عليه جميع المؤتمرات التي عقدت لمناقشة قضية الحاسوب واللغة العربية .

ثالثاً : عدم ترك أمر البرمجيات الحاسوبية العربية بيد الشركات ومراكز البحوث الغربية، بل ينبغي أن يصممها أبناؤها، فهم أقرب الناس رحماً بهذا المجال .

رابعاً : صناعة معجم موحد لمصطلحات اللسانيات الحاسوبية، بالعربية والإنجليزية، وفق المتعارف عليه علمياً في هذا النوع من المعاجم الاصطلاحية أو المصطلحية.

خامساً : أن يكون علم اللغة الحاسوبي مقرراً دراسياً معتمداً في أقسام اللغة العربية بكليات الآداب، كما هو الحال في كلية الآداب بالجامعات العربية والغربية .

سادساً : إنشاء قسم خاص للغويات الحاسوبية في الكليات والجامعات العربية يمنح درجة ليسانس– ماستر في هذا التخصص، على غرار ما هو موجود في بعض الجامعات العربية[1].

سابعاً : نشر تلك الرسائل العلمية التي كتبت أصلاً باللغة العربية عن قضايا استخدام اللغة العربية في الحاسوب، وهي تُشَكل فيما اطلعت عليه إسهاماً فاعلاً في ترسيخ هذا الاتجاه، ومعالجةً موضوعيةً وعلميةً لكثير من مشكلات تعامل العربية مع الحاسوب .

ثامناً : أن ما طرحه الدكتور نبيل علي في كتابه (اللغة العربية والحاسوب) من قائمة مقترحة في مجال بحوث اللسانيات الحاسوبية مطبقة على اللغة العربية، يحسن أن يكون قاعدة جيدة للانطلاق منها نحو تفعيل النشاط البحثي وتطويره في هذا المجال[2] .

صحيح إن بعض ما اقترحه كان بعد ذلك ميدان عناية الباحثين، ولكن كثيراً منها ما يزال في انتظار الأقلام المخلصة والعقول النيرة التي ستضيف بعداً جديداً للاتجاه الحاسوبي في دراسة اللغة، فالركام المعرفي يظل حقيقياً بالبحث والتنقيب، وإن النظرة التكاملية للعلوم الإنسانية تنبؤ عن مخرجات مرجوة للسانيات الحاسوبية تخدم اللغة العربية على الصعيدين التعليمي والتطبيقي، فالغاية المتوخاة من توصيف العربية لتلك الآلة المفتقرة إلى الفهم ستثمر عما تكتنفه العربية من أنظمة ضمنية غير مصرح بها من شأنها رسم صورة متكاملة تعين من ينشد تعرف العربية في رحلة اكتشافه لأبعادها الغائبة المستترة وخبايا نظامها الدقيق الفريد.

1– إن الصورة الحقيقية والفعلية لرصد الجهود في ميدان اللسانيات الحاسوبية العربية فتمثل– كما تقدم– في إنشاء قسم خاص لعلم "اللسانيات الحاسوبية" في جامعة الأمير سلطان الأهلية بالرياض (السعودية)، ويقول د عبد الرحمن بن حسن العارف – أنه أول قسم متخصص في هذا المجال بالجامعات العربية، وتم إنشاءه نتيجة متطلبات السوق الإقتصادية من جهة، وتتويجا – من جهة أخرى – للجهد الحاسوبية العربية التي أصبحت تشكل اتجاها عاما في الدراسات اللغوية المعاصرة. ينظر: عبد الرحمن بن حسن العارف، توظيف اللسانيات الحاسوبية في خدمة الدراسات اللغوية العربية،" جهود ونتائج"،اللسانيات، مجلة في علوم اللسان وتكنولوجياته،مركز البحث العلمي والتقني لتطوير اللغة العربية، العددان 12،13،الجزائر،2007، ص 36.

2– للمزيد من التفصيل ينظر : نبيل علي، اللغة والحاسوب، مرجع سابق، ص 538،...550.

سادسا : اللسانيات الحاسوبية في خدمة المعجم العربي:

لا شكّ أن العلوم المعجمية الحديثة تعتمد على النظريات اللسانية لتطوير نفسها وتحديد مواضيع بحثها وطرقها المنهجية[1]، إذ استفادت صناعة المعاجم الحديثة من لسانيات المدونة وطرق اشتغالها وكيفية دراستها لبنية الكلمة وتصنيفها للمقولات المعجمية النظرية والتطبيقية من اشتقاق وتوليد ونحت وتركيب ودلالة[2].

فأدرك الباحثون في اللسانيات الحاسوبية على وجه التخصيص – بلا تردد– أن مساحة استفادة المستوى المعجمي من الحاسوب واسعة جداً، لذلك ظهر ما يسمى بالمعاجم الحاسوبية أو المعاجم الآلية، حتى بدأ يأخذ بالبروز بوصفه علماً مستقلاً، أو فرعاً من فروع علم اللغة الحاسوبي الذي يصطلح عليه علم المعجم الحاسوبي؛ وتلك هي بوادر ظهور المعاجم الآلية أو الإلكترونية .

ويمكن توضيح علاقة المعجم العربي باللسانيات الحاسوبية، وهذا بالنظر في تداخل عدة علوم معرفية تحليلية للتعامل مع البحث اللساني الحاسوبي، وبالأخص المعالجة الآلية للغة العربية التي تزداد أهمية مع ارتفاع وتيرة تقدم التكنولوجيا والمعلوماتية.

إن الحاسوب مثله مثل المعجم، والمعجم وحده لا يكفي ليجعل من القارئ متضلعًا في لغته متمكنًا منها، وإن كان يساعده على ذلك، كما أن الحاسوب وحده لا يخول المتعامل معه لأن يكون متمكنًا خاصة من الرياضيات وإن كانت تساعده على ذلك[3].

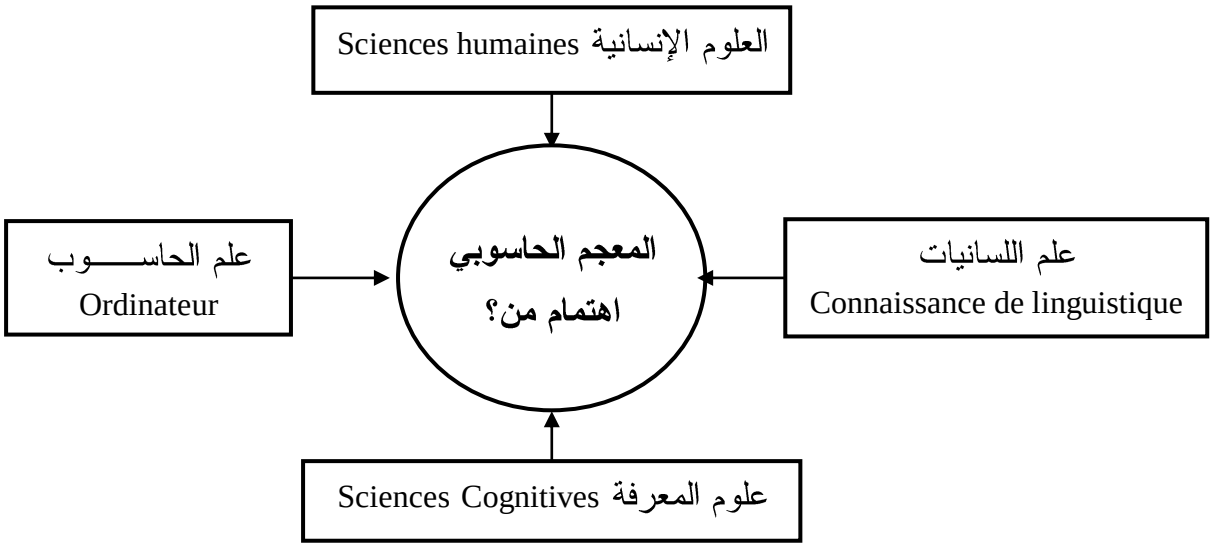
إن اللسانيات وسيلة، والحاسوب أداة، المعجم غاية...وفي هذا المجال المعرفي الحيوي يمكن استثمار هذا الثالوث في مجالات تطبيقية من منظور لساني – حاسوبي. وليس من الشك في أن ما يمتاز به هذا العمل المحوسب من دقة مع إشراك جميع العلوم المعرفية لإنجاح مثل هذه المشاريع، كالعلوم الإنسانية وعلوم الحاسوب وعلوم المعرفة وهذا بعد ضبط صارم للمصطلحات حتى يكون البرنامج الحاسوبي وعلم اللسانيات مضبوط، وفق قواعد بيانية[4] صارمة، وبإشراك مثل هذه العلوم المعرفية يكون له الأثر الفعال في إعداد معجم حاسوبي يتمتع بميزة الشمول؛ ويمثل الشكل (2) إجابة لإشكال كثر الكلام حوله: ما هي اهتمامات المعجم الحاسوبي؟:

1– جورج ماطوري، منهج المعجمية، ترجمة عبد العلي الودغيري، مطبعة المعرف الجديدة، الرباط، 1993، ص 170.

2– خليفة الميساوي، المصطلح اللساني وتأسيس المفهوم، سابق، ص 32.

3 – أحمد زياد محبك بن مصطفى، مقال:الحاسوب وتنمية المقدرة عند الطفل، مجلة المجمع الجزائري للغة العربية، المؤسسة الوطنية للفنون المطبعية، العدد السابع، السنة الثالثة، جمادى الثاني1429هـ/جون 2008م، ص44.

4- Les rrègles graphiques.



الشكل رقم 02: اهتمامات المعجم الحاسوبي[1].

وغني –عن الإشارة– في هذا المقام، ما سيتحقق للمعجم الحاسوبي من سهولة الوصول إلى مواده ومفرداته وذلك بتمكين القراء آليًا من العثور على ما يريدونه من مفردات المعجم، وسيتمكن المعجميون من تخزين هذه المواد (المدخلات) آليًا، والتعديل – متى شاءوا– في متونها، إضافةً أو حذفًا، أو تصحيحًا[2].

وبطبيعة الحال لا يمكن الحصول على معجم حاسوبي وفق المعايير المعمول بها، معجم حاسوبي شامل يواكب مسيرة اللغة العربية بمختلف تحدياتها، إلا في ظل توفر شروط علمية يوفرها الباحثون من اللسانيين والحاسوبيين وذلك باتباع نظرية قوامها جمع المادة اللغوية وصياغتها حاسوبيًا، بعد ضبط المصطلحات المراد صياغتها وتنميطها وتنسيقها، لوضع نظرية لسانية تمكن الباحث من وضع قواعد بيانية وفق ما يسمى شبكة العلاقات المعجمية، وبهذا العمل المتناسق يمثل المعجم الحاسوبي (الإلكتروني) مصدر بيانات أساسي لمعظم البرامج الحاسوبية[3] من محللات صرفية[4] ومركبات لغوية[5].

ولن يتأثر المعجم الحاسوبي من ضخامة مادته، فهو يتمتع بذاكرة قوية واسعة تمكنه من استيعاب المفردات والنصوص أيًا كان عددها[6].

1– جمانة كمال حنا، مؤسسة صخر لبرامج الحاسوب– القاهرة، موقع الويب، http//www.arabisation.org.ma.

2– صادق عبد الله أبو سليمان، مقال: نحو استثمار أفضل للحاسوب في مجالات خدمة اللغة العربية وعلومها، مجلة المجمع الجزائري للغة العربية، المؤسسة الوطنية للفنون المطبعية، العدد السادس، السنة الثالثة، ذو الحجة 1428هـ/ ديسمبر 2007م، ص 74.

3- Logiciels Informatique.

4- Analyseurs Morphologique.

5 – نبيل علي، اللغة العربية والحاسوب، سابق، ص 151.

6 – صادق عبد الله أبو سليمان، مقال: نحو استثمار أفضل للحاسوب في مجالات خدمة اللغة العربية وعلومها، سابق، ص 74.

وتعد الإفادة من الحاسوب في مقدمة هذه المشاريع لما تمتلكه هذه الأداة من القدرات الفائقة في إجراء العملات الإحصائية، وتنظيم قاعدة البيانات[1]، وتبويب المدخلات وقاعدة المعارف[2] والبيانات اللغوية.

ويتوقف المشروع في برنامج الصناعة المعجمية الحاسوبية على حجم المدونة التي تُؤخذ منها الوحدات اللغوية كمادة مدخلة، لأن المشكلة الأساسية تتصل بهذا الحجم وإمكانية إعداده، وتنضيجه لبرنامج حاسوبي[3].

ملاحظة: ولأهمية الموضوع خصصنا الفصل الثاني للحديث عن دور الحاسوب في حوسبة المعجم العربي من خلال التطرق إلى أهم المجالات الأساسية في الصناعة المعجمية الحديثة، مبرزين أهميتها في الإنجازات المعجمية الحاسوبية إذ يتمكن المعجميون من تخزين هذه المعجمات آليًا والتعديل متى أرادوا في متونها ومداخلها. إضافةً، أو حذفًا، أو تصحيحًا، أو تغيير مواقع أو صياغة الأمر الّذي سيمكّن من إيجاد معجم عربي حديث يستجيب لمتطلبات تكنولوجيا العصر، ويوفّر للعربية ما يمكن تسميته بالمعجم المتنامي، الشامل المواكب لمسيرتها والمعبّر عن خصائصها البنيوية[4].

1 - Base de données.

2 - Base de connaissances.

3 – عبد القادر عبد الجليل، علم اللسانيات الحديثة، نظم التحكم وقواعد البيانات، دار صفاء، عمان، ط1، 2002م/ 1422هـ، ص589.

4 – صادق عبد الله أبو سليمان: مجلة المجمع الجزائري للغة العربية، العدد السادس، السنة الثالثة، ذو الحجة 1428/ديسمبر 2007، ص74.

الفصل الثاني

حوسبة المعجم العربي

* نحو مقاربة لسانية حاسوبية لمعاجم آلية عربية *

أولاً- المعجم العربي من العين إلى الحاسوب

ثانياً- المعجم العربي والحوسبة العربية

ثالثا- من المعاجم الورقية إلى المعاجم المحوسبة: نظرة تاريخية

رابعاً- المعجم الإلكتروني للغة العربية : المشروع الذي لابد منه

خامساً- قواعد المعطيات النصية لتطوير المعجم الحاسوبي (الإلكتروني)

سادسا- الفرق بين المعجم الورقيّ والمعجم الإلكتروني

أولا: المعجم العربي من العين إلى الحاسوب:

إن المعجم هو الوعاء اللغوي لحضارات الإنسانية على تنوعها واختلافها، وهو فهرس معارفها وخبراتها، وهو خط التقاء اللغة قواعدها ومبادئها وهياكلها، وأنماطها وقيودها.

وتلتقي علوم اللسانيات في ساحة المعاجم مع علوم الحاسوب، حيث يبدو المعجم موضوعًا مثيرًا للمعالجة الآلية يشير إلى وجود حالات مماثلة له، ولعل السبب في ذلك أن المعجم يجمع في كيان واحد ثالوث "البيانات، المعلومات، المعارف"؛ أما البيانات فتتمثل في قوائم المفردات والمعلومات المنبعثة من شبكة العلاقات الكثيفة التي تربط بين هذه المفردات والعدد الهائل من التحليلات والمؤشرات والإحصائيات المترتبة عن هذه العلاقات والمعارف، سواء كانت لغوية أو حاسوبية[1].

فالحوسبة المعجمية تعتبر من أهم مجالات اللسانيات الحاسوبية وأكثرها تلبيةً للمتطلبات العلمية والثقافية والتقنية في الدول المتقدمة في العالم المعاصر. إذ يقدم الحاسوب خِدْمَاتٍ كبيرة للبحث اللغوي من خلال المعاونة في إعداد مدونات معجمية، باعتبارها أساسيات المعطيات المعجماتية من جهة وأساسيات المعطيات القاموساتية من جهة ثانية[2] . الهدف من هذه المدونات هو العمل على الإعداد المعجمي لمجموع الكلمات البسيطة والتعبيرات المركبة الواردة في نص معجمي محدد. وتتجلى أهمية الحاسوب في صناعة المعجم في مراحل متعددة أملتها العلاقة المتقدمة للهندسة المعلوماتية واللسانيات، فكلاهما يستفيدان من بعهما البعض، على الأقل في الجانب الخوارزمي، ذلك أن اللسانيات بدون هندسة معلوماتية تبقي نظريات وإن كانت غير قابلة فعليا للمعاينة، وأن المعلومات بدون لسانيات لن تبلغ أبدا هدفها في التواصل والاستمرارية، فهناك هناك متطلبات عامة للبناء لا يمكن أن يقوم أي عمل بالاستغناء عنه، فقد أصبح استخدام الحاسوب في جمع المادة اللغوية وترتيبها وسيلة هامة في صناعة المعاجم، حيث ظهر للعالم

1- نبيل علي، اللغة العربية والحاسوب: دراسة بحثية، تقديم: أسامة الخولي، دار تعريب، الرياض، المملكة السعودية، طبعة 1988م، ص457.

2- لهذان أساسيان أهداف كبرى تتمثل للإجابة عن التساؤلات المتعلقة للمفردات في سياقاتها وتواردتها الممكنة، فهي كما تقول بيرنالد كيمادا هي التي تجد أساسيات المعطيات النصية الخاضعة للمعالجة الحاسوبية بحيث يتم تدقيق اختبار المداخل المناسبة انطلاقا من سياقاتها بدقة، كما تتمثل في طبيعة عناصر مواد المعجم المقترحة للتساؤل، لذلك وجب تصنيف المعطيات حسب خصوصيات كل معجم فمثلا يتم تسجيل وتخزين المعاجم الورقية على الحاسوب وفق شفارات تحليلية خاصة تمكن من الوصول إلى مختلف المضامين والتعرف بسرعة وبدقة وبالتعديل لمعطياتها حذفا وإضافة، وهذا ما استخدم في المعاجم المحوسبة العربية حيث تم بها وضع معاجم في نسخ تامة ومضبوطة مهيأة للتصحيح والطبع والاختصار ولانجاز معاجم وقواميس مدرسية ومعاجم متخصصة أو كذلك مزدوجة وكل هذا حسب الأهداف المتوخاة من كل عمل معجمي. للمزيد ينظر: د. عبد الغاني أبو العزم، الحاسوب والصناعة المعجمية. اللسان العربي، ع 47.

فرع حديث يصطلح عليه: "المعجم الحاسوبي"[1]، وهو أحد فروع[2] اللسانيات الحاسوبية[3]، وهما كما أشرت مصطلحان جديدان في مجال البحث اللغوي[4].

وبما أن المعلوميات أصبحت في متناول الجميع فهي في الآن نفسه، استفادت من التطورات التكنولوجية والعلمية، كما أنها ألغت – كما يقول د. عبد الغاني أبو العزم– كل الاعتراضات المتعلقة بخاصية اللغة الاشتقاقية وأن أي معجم إلكتروني بإمكانه أن يمد مستعمله بكل ما يطلبه، سواء تعلق الأمر بمجمل الكلمات المشتقة من جذر واحد، أو حسب كتابتها الخطية[5] .

وللحاسوب إضافة إلى دوره في تدبير وتخزين أساسيات المعطيات اللغوية لاستغلالها في مستويات لغوية وإجرائية متعددة، فإن دوره كبير في تهيئة وتنظيم النصوص المعجمية. فهو يقوم بمتابعة نمو مفردات اللغة بسيطها ومركبيها في جميع مراحل الصناعة المعجمية، مما أدى إلى بلورة قواميس حاسوبية هي من البرمجيات الأساسية للعديد من الأنظمة المعلوماتية المتقدمة، ومما لاشك فيه أن بناء معجم محوسب يقتضي معرفة ملامح الباحث ومجالات الاستخدام الممكنة، أو المتوخاة من هذه الصناعة كالترجمة الآلية وفهم وترتيب الكلام واستخراج المعلومات ومعالجة النصوص، إلى غير ذلك[6].

ومع الانتشار الواسع للهندسة المعلوماتية في العالم العربي، ابتدأ في نقل مجموعة من القواميس والمعاجم العربية القديمة منها والحديثة إلى الحاسوب ومنها تجربة "جامع معاجم العربية" التي ضمت باقة قواميس عربية[7] (قيمة وحديثة) هي الأكثر استعمالا من لدن الطلاب والباحثين. مما ساعد على بناء معاجم محوسبة متقدمة كان لها الأثر الكبير في صياغة نظرية معجمية جديدة ستكون بحق قاطرة في انجاز معجم إلكترونية عربية وهذه التجربة تمثلت في

1– بدأ علم المعجم الحاسوبي يبرز كمجال مستقل من المعرفة حينما عقدت عدة ورشات عمل وأجريت بحوث مشتركة تتحدث عن أتمتة المعجم والتقييس المعجمي، والمعاجم الالكترونية، وعلم المعجم الحاسوبي، كما خصصت مجالات علم اللغة الحاسوبي (اللسانيات الحاسوبية)، إعدادًا منها لعم المعجم الحاسوبي ووجدت الحاجة إل أتمتة المعاجم نتيجة تضخم المواد اللغوية التي تهتم بها.

2 – وليد العناتي وخالد الجبر، دليل الباحث إلى اللسانيات الحاسوبية، دار جرير، الأردن، ط1، 1428هـ/2007م، ص 14.

3 – علم اللسانيات الحاسوبي هو أحد فروع اللسانيات التطبيقية، يهتم بالإفادة من معطيات الحاسوب في دراسة قضايا اللسانيات المتعددة، مثل رصد الظواهر اللغوية وفقًا لمستوياتها الصوتية والصرفية والنحوية والبلاغية والعروضية، يهدف هذا العلم إل معالجة عمليات تحليل ظواهر لغوية ونصوص تراثية وإجراء العمليات الإحصائية وصناعة المعاجم، والترجمة الآلية وتعليم اللغات وإجراء التحاليل على السياقات اللغوية المتعددة، كما تبحث اللسانيات الحاسوبية عن وضع اللغات البشرية في صيغ وأطر رياضية وذلك للاقتراب من هذا العلم الذي يبحث في اللغة على أنها ظاهرة حاسوبية معلوماتية يمكن معالجتها في الحاسوب، وذلك من أجل السرعة والدقة العلمية في البحوث اللغوية ومن أجل ترجمة النصوص اللغوية ترجمة آلية فورية أيضًا. للمزيد من التفصيل ينظر الصفحة .... من هذا البحث .

4 – جاء تفصيل في الكلام حول هذين الحقلين في الجانب المخصص لهما من البحث.

5– انظر، عبد الغني أبو العزم، مقدمة المعجم العني الشامل، مرقون(2007)، ص. 9.

6– عبد القادر الفاسي الفهري، المعجمة والتوسيط: نظرات جديدة في قضايا اللغة العربية، المركز الثقافي العربي، ط1، 1997م، ص61.

7– يضم برنامج "جامع معاجم العربية" باقة من القواميس.

"معجم الغني" لصاحبه عبد الغني أبو العزم والذي أصدرته شركة صخر ميديا بالقاهرة. فهذه المعاجم المحوسبة تعتبر في نفس الآن، حلقة وصل ما بين الصناعة المعجمية التقليدية والصناعة المعجمية الإلكترونية المعاصرة من جهة، وتهيئة لأساسيات المعطيات المعجمية للقواميس الإلكترونية من جهة أخرى. ؛ وبناء على ما تقدم يمكننا أن نطرح مجموعة من التساؤلات يمكن من خلالها وضع تصور منهجي لإعداد معجم حاسوبي[1]: ما علاقة المعجم العربي باللسانيات الحاسوبية، وبالأخص المعالجة الآلية للغة العربية؟، وما مدى تخصص هذا الفرع في إعداد المعاجم المحوسبة (الإلكترونية)؟، وهل بإمكاننا وضع نموذج لساني معلوماتي لصناعة معجم عربي محوسب؟

ثانيا- المعجم العربي والحوسبة العربية[2]:

منذ بداية الستينات كانت صناعة المعجم تتصف بالطابع التجريبي الذي لا يعتمد على منهجية علمية حقيقية وإنما على الخبرة، العلمية التي اكتسبها واضعوا المعاجم في العصور المتقدمة؛ فتطورت بذلك اللسانيات الحديثة عامة واللسانيات الحاسوبية على الخصوص، وقد ظهرت ملامح ذلك التطور في المعاجم اللغوية العامة وكانت حينئذ تسعى إلى إرساء أسس علمية موضوعية تستند إلى نتائج البحوث اللسانية والتطبيقات الحاسوبية التي أقيمت في مختلف ميادين الحوسبة[3].

فالهدف من هذا المبحث هو أولاً وضع أسس علمية لصناعة المعجم العربي الحديث، اعتمادًا على المكتسبات المنهجية التي نتجت عن اللسانيات الحديثة (اللسانيات الحاسوبية)، التي يعود لها فضل ظهور المعاجم الإلكترونية[4]، فظهرت إمكانيات الحاسوب الهائلة للباحثين خاصة في الحقول المعجمية وما يجري حاليًا من جهود لميكنة المعاجم اللغوية[5]، ويرى نبيل علي أن

---

1 – هذا الإشكال سياتي جوابه في فصل خاص بالمعجم الحاسوبي في حينه وموضعه.

2– إن هدف الحوسبة هو الوصول إلى تقنية الفهم الآلي للغة أو إلى درجة من درجات الفهم، ثم تطوير تطبيقات تبنى على هذه التقنية لتستخدم في التعليم والأعمال وجميع مجالات الحياة التي يدخل فيها الحاسوب، بحيث تحفظ للعربي لغته عند تعامله مع هذه التطبيقات فلا تضمحل صلته بموارد هويته بانتشار التقنية إلى التعليم والأعمال وبقية مناحي الحياة. وعليه فالواضح أن الحوسبة وسيلة لا غاية في ذاتها.

وعند الحديث عن العلاقة بين حوسبة اللغة العربيّة واللسانيّات الحديثة، واعتمادها عليها أو عدمه، ينبغي التّنبّه أوّلاً إلى قضيّتين: الأولى أنّ عمليّة الحوسبة –أيّة حوسبة– هي بطبيعتها تحتاج ابتداءً إلى وجود نموذج وصفيّ دقيق لما يُراد حوسبتُهُ، وأنّ غياب مثل هذا النّموذج يتنبّأ لنا بالحال الّتي ستكون عليها عمليّة الحوسبة من العجز والقصور، والقضيّة الثّانية هي أنّ كثيراً من العاملين في مجال حوسبة اللغة العربيّة هم أيضاً باحثون لسانيّون، وبينما تقودنا النّقطة الأولى إلى سؤال عن إمكان الاعتماد على نظريّة لسانيّة حديثة لم تكتمل بعدُ في إجراء حوسبة للغّة، فإنّ النّقطة الثّانية تقودنا إلى تعليل حدوث هذه الحوسبة واقعيّاً برغم عدم اكتمال النّظريّة، بسبب أنّ العاملين في الحقلين هم أنفسهم.

3 – Informatisation.

4 – Dictionnaire électronique.

5 – نبيل علي، اللغة العربية والحاسوب، سابق، ص330.

أهمية الحاسوب للمعجم اللغوي العربي، تفوق بكثير أهميته بالنسبة للمعجم الإنجليزي، وأساس هذه الرؤية هو الطبيعة الخاصة لهيكل المعجم العربي المبني على أساس الجذور[1]

من أبرز الإنجازات والجهود الحاسوبية في مجال المعاجم[2] ما تم إنجازه للقرآن الكريم، حيث توجد مجموعة من البرامج تؤدي مهام "المعجم المفهرس لألفاظ القرآن الكريم"[3]، فضلاً على ذلك فإنه يمكن عن طريق البرنامج الحاسوبي معرفة الآية أو الآيات التي ترد فيها الكلمة الهدف، مع سماع الترتيل، وغير ذلك من إمكانات قد لا تكون على صلة بالمعجم[4]، ومن الجهود العلمية الحاسوبية الرائدة في مجال المعاجم تلك الدراسة التي قام بها علي حلمي موسى، باستخدام الحاسوب في الإحصاء اللغوي لجذور اللغة العربية في معجم الصحاح" للجوهري[5].

هذا البحث يهدف إلى عرض جذور مفردات اللغة العربية المستخدمة والواردة في المعجم، على شكل جداول تبين تردد الحروف في المواقع المختلفة من الكلمة تتابع الحروف، واختلافها من حرف إلى آخر[6].

لقد اجتهد خبراء الحواسيب في توصيف حوسبة المعجم، إذ يقوم الحاسوب اعتمادًا على قواعد الاشتقاق بتوليد جميع القياسات من مشتقات ومزيدة ومصادر، وتخوله قواعد التعريف والإسناد من تحديد صيغة الفعل بصورها المتنوعة (الماضي. المضارع. الأمر)، مسندًا إلى جميع الضمائر، كما أن التوصيف المعجمي للكلمة العربية المدخلة يتم تحديدها من خلال عدة حقول: الحقل الصرفي، الحقل النحوي، الحقل الدلالي، الحقل الصوتي، والحقل الإحصائي[7].

إن الكلام في هذا المبحث سيكون عن مستقبل المعجم العربي الحديث في عصر المعلومات وذلك من حيث الأهداف التي يمكن لعلماء اللغة أن يبلغوها، لمحو ما كان في أذهان القدماء من المشاكل والصعوبات التي وجدوها في صناعة معاجمهم اللغوية.

فالحديث عن هذا المستقبل المنشود لا يخلو من الصعوبات وذلك لاختلاف الباحثين في الأهداف التي يمكن النظر إليها أو معالجتها في بناء معجم يتطلب كل مقتضيات العصر ويساير العولمة الحاصلة في ظل سيطرة اللغات الأجنبية على سائر العلوم التجريبية وحتى الإنسانية.

---

1 – المرجع السابق، ص 330.

2– سيأتي تفصيل في أبرز الجهود والإنجازات الحاسوبية في بناء المعاجم الإلكترونية في الباب الثالث من هذا البحث، لهذا سنكتفي في عرضه هنا بالقدر الذي يسمح لنا بوضع الإطار العام للسانيات الحاسوبية في سبيل حوسبة المعجم العربي لمنظومة اللغة ككل.

3 – فتح الله سليمان، دراسات في علم اللغة، سابق، ص 102.

4 – المرجع نفسه، ص 102.

5 – علي حلمي موسى، دراسة إحصائية لجذور معجم الصحاح باستخدام الكومبيوتر، الهيئة المصرية العامة للكتاب، (د.ب)، طبعة 1978–1979م.

6 – المرجع نفسه، ص 10.

7 – محمد علي الزركان، اللسانيات وبرمجة اللغة العربية في الحاسوب، ندوة استخدام اللغة العربية في تقنية المعلومات، الرياض 1992.

وإن كانت هذه الأهداف تبدو واضحة لبعض الدارسين، فإن آخرين يلمسون الغموض في معرفة آفاق تطور هذه المعاجم؛ ولكن السؤال المطروح هو: أي **مستقبل وأي هدف تقدمه المعاجم اللغوية العربية في ظل مواكبتها لمستجدات العصر، من حيث الرصيد المفرداتي وكمية المصطلحات العلمية الحديثة؟؛**

2-1- أهمية الحاسـوب ودوره في صناعة المعجم العربي:

إن نمو الدراسات اللغوية وتطورها مازالا في استمرار، وإن المعجم في خلال ذلك يتبوأ مكانة بارزة، والمعاجم اللغوية العربية تعد صناعة قائمة بذاتها، وعلم مؤسس على قواعد متينة، وله مناهج وتطبيقات، وما قدمه الخليل أو الزبيدي أو ابن منظور أو الأزهري او غيرهم من جهود في هذا الشأن خير دليل على أهمية تلك الصناعة.

وتقضي الضرورة تحديث أساليب التعامل مع المعجم وتطوير طرق البحث و سبل الاستفادة منه، خاصة إذا كان أحد المصادر الأساس التي لا غني للباحث المتخصص او غير المتخصص عنها، ومن هذه الضرورة بدأ البحث والتفكير في قضايا وتحديات حوسبة المعجم اللغوي العربي[1] للتحرر من صعوبة الوصول إليها من طرف الباحثين، وقد أصبح الحاسوب[2] انطلاقًا من ذلك أداة بحث لا مناص من الاعتماد عليها لإنجاح أي عمل معجمي، كما يشيع في أوساط البحث اللساني[3].

فما الذي يمكن للحاسوب أن يضيفه للمعجم العربي؟، وما هي الخصائص الآلية التي أصبحت تتوفر عليها الأداة الحاسوبية، في سبيل تنمية المعجم اللغوي وتطويره؟.

إن للحاسوب قدرة غير محدودة في إدماج آلاف الكتب بعضها مع بعض، وكذلك في إدماج نصوص مع أخرى، بفضل برامج معلوماتية، وهذه ميزة خاصة به دون وسائل البحث الأخرى[4]، فالحاسوب يحول البحث إلى قاعدة من المعطيات اللغوية التي يكون تصفحها آليًا، وتتم عملاً بالخطوات الآتية:

1- عبد الله أبو هيف، مستقبل اللغة العربية " حوسبة المعجم العربي ومشكلاته اللغوية والتقنية أنموذجا"، اللغة العربية، المجلس الأعلى للغة العربية، الجزائر، العدد العاشر، 2004، ص 127 ومابعدها.

2- إن الهدف في هذا البحث يتمثل في استخدام مصطلح "الحاسوب" مرادفًا لمصطلح "كومبيوتر" مع العلم أن هذا الأخير أكثر شيوعًا من الأول.

3 – عبد الغني أبو العزم، مقال: الحاسوب والصناعة المعجمية، مجلة اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق تعريب، 1998م، ص28 .

4 – عبد الرحمن الحاج صالح، بحوث ودراسات في اللسانيات العربية، المؤسسة الوطنية للفنون المطبعية، 2007م، ص167.

1-حصر كل السياقات لكل كلمة وردت فيها أو في جزء منها مع ذكر المرجع الكامل لها، وهذا العمل هو عبارة عن فهرستها. "فكل كلمة يذكر بكل ما اقترنت به من القرائن في نص واحد أو كل النّصوص، وهذا ما نسميه باللّغات الأجنبية بـــ Concordance أو Index.

وهذا بالنسبة للنص القرآني قد أنجز قبل اليوم ومثال على ذلك "المعجم المفهرس لألفاظ القرآن الكريم لمحمد فؤاد عبد الباقي قبل ظهور الحاسوب.

فالحاسوب يستطيع أن يفهرس أي كلمة من العناصر من مدونة- ولو جاءت في ملايين من النصوص في وقت قصير[1].

2-حصر جميع العبارات الخاصة، لا المفردات فقط، المتآلفة الألفاظ، بسياقات كل واحد منها وذكر المرجع الكامل لها. ونعني بالعبارة هنا أن تركيب تقترن فيه كلمتان أو ثلاث بشيء من التلازم.

3-تحليل آلي بالحاسوب لكل كلمة متمكنة تأتي في نص إلى مادتها الأصلية ووزنها.

4-إحصاء آلي في نص أو عدة نصوص لكل وحدة لغوية أيا كانت بما في ذلك العبارات والجذور والأوزان[2].

وجملة هذا العرض الوجيز يمكن أن نجمل أهم الخصائص الّتي يقدّمها الحاسوب في سبيل تطوير المعجم في:

1- تخزين النّصوص وبرمجتها.

2- ترتيب المفردات ورسم الألفاظ وجذوره، لتصبح معالجتها معالجة آلية بعد إدخالها في أنظمة دقيقة ممّا يسمح بالتطبيق العلمي المباشر الاستعمالات المتعددة[3].

إنّ هاتان الميزتان الأساسيتان في التطبيقات اللّسانية الحاسوبية في مجال صناعة المعاجم تخضع لضوابط وقواعد معطيات تختص بتنظيم بنوك المعطيات[4] لا يمكن الاستغناء عنها في مجال الصناعة المعجمية خاصة ونحن في عصر تحكمه الآلات الحاسوبية كأداة لها دور في النّهوض بمثل هذه الدراسات الحديثة لما لها من سرعة وإتقان في مجال ضبط القواعد اللّسانية. وهذا لبرمجتها وفق أسس ونظريات.

1- المرجع السابق، ص167.

2- المرجع نفسه، ص 168.

3- عبد الغني أبو العزم، الحاسوب والصناعة المعجماتية، عبد الغني أبو العزم (مجلة اللسان العربي)، ص28 .

4- للمزيد من بسط القول ينظر: ليلى المسعودي، علم المصطلحات وبنوك المعطيات، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 28، الرباط، المغرب، 1987، ص 86،87.

إنّ ما أصبح يوفره الحاسوب في مجال صناعة المعجم يمكن تحديده في النقاط التالية[1]:

1. إنجاز معاجم متنوعة ومتخصصة وبدقة وبأقصى ما يمكن من السرعة.
2. توفير الإرهاق. والتعب+ العمل المتقن وفي وقت وجيز.
3. ربح مصاريف باهظة.

فلقد تم التغلب على العديد من الصفات التي كانت ترهق المعجمي بفضل تقنية الحاسوب. فالجهود العلمية ما زالت تتواصل وذلك من أجل:

- حصر وضبط وتسجيل مختلف المعطيات.
- إنجاز برامج إلكترونية.
- طبع ونشر النتائج.

فالحاسوب في هذا الصدد يتحول إلى وعاء ضخم سواء في تعامله مع الجزء الموسع للألفاظ، أو في تعامله مع كل الاستعمالات في مجموع النصوص الموثقة بداخله. وهو في هذه الحالة يقدم للمستعمل مادة ضخمة منسقة ومضبوطة ومرتبة ترتيبًا ألفبائيًا[2].

إنّ ما يقدمه الحاسوب من ضبط للقواعد المعطيات واختصار الوقت في إنجاز البرامج الحاسوبية لابدّ من اللّساني والحاسوبي استغلال هذه الميزات التي يوفّرها الحاسوب. لتسخير كل الطاقات المعرفية للنهوض بهذا المشروع. وفي سبيل تطوير برامج حاسوبية وهذا من أجل خدمة الصناعة المعجمية التي تدير المجال الحيوي للغة العربية. لما يقتضيه هذا المجال من مجالات نحوية وصرفية ودلالية.

فالصناعة المعجمية تعتمد على مدى طاقة الحاسوب الاستيعابية بفضل اتساع ذاكرته والإمكانات التي يوفرها في تعامله مع النصوص وهذا وفق أربعة مجالات أساسية[3]:

1- أساسيات المعطيات Les bases de données
2- أساسيات المعطيات النّصية Les bases de données textuelles
3- أساسيات المعطيات المعجماتية Les bases de données lexicographiques
4- أساسيات المعطيات القاموسية Les bases de données dictionnairiques

1- الحاسوب والصناعة المعجماتية، مجلة اللّسان العربي، ص: 29، و ينظر: عبد الرحمن الحاج صالح، بحوث ودراسات في اللسانيات العربية، سابق، ص 167.

2 - الحاسوب والصناعة المعجماتية، سابق، ص 29.

3- المرجع نفسه، ص 30.

وليس من شك في أن ما يمتاز به العمل المحوسب من دقة سيكون له الأثر الفعّال في إعداد معجمات تتمتع بميزتي الشمول واطراد ترتيبها الخارجي والداخلي وهما من أهم ما يحتاجه العمل المعجمي النّاجح ويحرص المعجميون على تحقيقهما في معجماتهم[1].

فالمعجم المحوسب يحقق السهولة في الوصول إلى مواده ومفرداته وذلك بتمكين القرّاء آليا من العثور على ما يريد به من مفردات المعجم وتراكيبه، وتحديد سياقات استعمالاتها[2].

2.2- استخدام الحاسوب في حوسبة المعجم العربي:

لقد صنّف نبيل علي في كتابه اللّغة العربية والحاسوب المجالات الرئيسية لاستخدام الحاسوب في ميكنة المعجم بثلاثة حقول[3]:

1- تخزين المعاجم إلكترونيًا: على شرائح ممغنطة، أو ضوئية، أو إشارات إلكترونية وذلك لاستخدامها في المعالجات اللّغوية المختلفة، وفي أغراض الترجمة الآلية، والتعليم، واكتشاف وتصحيح الأخطاء الإملائية، وما إلى ذلك.

2- تحليل العلاقات التي تربط بين مفردات المعجم: كالعلاقة بين جذور الكلمات والصّيغ الصرفية (قواعد تكوين الكلمات)، والعلاقات المتعلقة بمعاني الألفاظ، كعلاقات الترادف، والاشتراك اللّفظي، والتضاد، أو العلاقات الموضوعية (مصطلحات طبية، مصطلحات زراعية...الخ).

3- تحليل مفردات المعجم: والمستخدمة في شرح معاني مفرداته، وهي دراسات ذات أهمية خاصة لبحوث الدلالة المعجمية، ونظم الفهم الآلي. واسترجاع المعلومات، إذ أنّها تسعى إلى إبراز المفاهيم الأساسية، والأبعاد المحورية لمدلولات الألفاظ المترابطة (كتلك التي تربط بين ألفاظ علاقة القرابة أب، أم، شقيق- شقيقة، خال- عم- صهر، جدة، حماة..) والوصول إلى نواة هذه اللّغة أو قاعدة المفاهيم الأساسية التي يقام عليها، معجمها، وتحديد شكل شجرة مفرداتها ومدى عمق مستويات تفريعها (مثال: عالم، قارة، إقليم، بلد، مدينة، قرية، شارع، حارة...)[4].

---

1 – صادق عبد الله أبو سليمان: مجلة المجمع الجزائري للغة العربية، العدد السادس، السنة الثالثة، ذو الحجة 1428/ديسمبر 2007، ص73.

2 – المرجع نفسه، ص 74.

3 – نبيل علي: اللغة والحاسوب (دراسة بحثية) تقديم أسامة الخولي، دار تعريب، د.ط، 1988، ص 144.

4 – المرجع نفسه، ص 144، 145.

وتشير الدراسات إلى أنّ حوسبة المعجم العربي[1] تمر حاليا بنقلة نوعية حادة ترجع إلى عدّة عوامل أهمها[2]:

1- ما وفّرته تكنولوجيا المعلومات من وسائل لتجميع المّادة المعجمية وتصنيفها وتحريرها، ومن أهمّ هذه الوسائل ما يعرف بدفاتر النصوص المحوسبة التي تخزن بها عينة من النصوص العقلية التي تمثل الاستخدام الواقعي للغة يتم على أساسها تحديد معاني الكلمات داخل سياقاتها الفعلية، وذلك على عكس ما يحدث في المعجمية التقليدية التي يتولى فيها المعجميون تحديد معاني الكلمات على أساس حصيلتهم اللّغوية التي لا يمكن- مهما اتسعت وتوسعت- أن تغطي الكم الهائل من معاني اللّغة.

2- التوسع في نظم معالجة اللّغة آليًا، خاصة نظم الترجمة الآلية، والقلة النوعية الحالية نحو النّظم الآلية لتحليل مضمون النّصوص وفهمها أتوماتيكيًا، وهو ما يتطلب تحليلاً عميقًا لمتن المادة المعجمية في صورة ما يعرف بالشيكات الدلالية، ومخططات المفاهيم.

3- الثورة التنظيمية التي نجم عنها تصميمات مبتكرة لقواعد البيانات وقواعد المعارف المعجمية.

4- الإقبال الشديد على تعلم الانجليزية لغة ثانية، ممّا أدّى إلى تنويع الخدمات المعجمية، وإنتاج معاجم (فردية/ ثنائية) لخدمة الأغراض والمراحل المختلفة وهو ما اضطرّ الناشر المعجمي إلى أن يلجأ إلى تكنولوجيا المعلومات لإخراج إصدارات متنوعة.

5- تسارع عملية إنتاج المعرفة الجديدة وما نجم عن ذلك من زيادة لطلب المصطلحي الّذي يتطلب بدوره تصميم نظم آلية لدعم عملية توليد المصطلح.

ومن خلال هذه الإطلالة الخفيفة: يمكن القول أن للحاسوب دور مهم في صناعة المعجم لما يتوفر عليه من طاقة استيعابية، وبرامج وأنظمة آلية ذات أبعاد متعددة الجوانب على مستوى الجذور والمشتقات والتعابير، أو على مستوى القضايا الصرفية والنّحوية، أو على مستوى الترادف والأضداد أو على مستوى الشواهد، وكذلك "قدرته الفائقة على الاسترجاع، وذلك لتجميع المصطلحات وتبويبها وحفظ عينات النصوص والوثائق الأصلية المستخدمة لهذه المصطلحات. واسترجاع المرادفات المكافئة لها في معاجم اللغات الأخرى[3].

1- عبد الله أبو هيف،مستقبل اللغة العربية " حوسبة المعجم العربي ومشكلاته اللغوية والتقنية أنموذجا"، اللغة العربية، المجلس الأعلى للغة العربية، الجزائر، العدد العاشر، 2004، ص 127.

2- نبيل علي، نادية حجازي، (الفجوة الرقمية: رؤية عربية لمجتمع المعرفة)، عالم المعرفة، الكويت، ع318، 2005، ص357.

3- نبيل علي، اللغة والحاسوب، دراسة بحثية، سابق، ص499.

وبذلك فإنّ المخزون في هذه الحالة بإمكانه أن يمدّ الباحث في مجاله بما يحتاجه وتلبية رغباته بالسّرعة المطلوبة.

يتم التحكم في ضبط الكلمات والتراكيب في ضوء أنظمة قائمة على قواعد منهجية مضبوطة، وفق منظومة متكاملة المعطيات التي تُعتبر أساس الصناعة المعجمية ممّا يسمح بالإنجاز السريع من جهد والدّقة التامة، لا فيما يخص المداخل المعجمية والاشتقاقات فقط، بل وفي مجال تنوع التراكيب مما يعطي لأي معجم ينجز في ظلّ هذه الشروط قيمة علمية دقيقة[1]. صفوة القول، إن المعجم منظومة معقدة، بل غاية في التعقيد، لدرجة يتعذر معها فهم أدائها ومحاصرة ظواهرها في ظل القيود التي تفرضها الوسائل اليدوية التقليدية لبناء المعاجم وتحديثها واستخدامها، ولم تكن ميكنة المعجم باستخدام الحاسوب أمرا من قبيل الرفاهية الفنية، بل مطلباً أساسياً، تفرضه طبيعة المعجم، طبيعة مضمونه، وتنظيمه، وتحديثه، وخدماته للبشر ولنظمهم الآلية، وذلك بجانب ما تفرضه بالطبع طبيعة العلم باعتباره معرفة، أو باعتباره منهجاً لتطوير المعرفة، وهو ما ينعكس في النهاية في صورة مفاهيم تحتاج إلى مصطلحات تدل عليها.

والآن في واقعنا، حيث تصدمنا أزمة المعجم العربي المزمنة، علة كبرى من علل دائنا اللغوي، وهي الأزمة التي تعددت مظاهرها وترسخت إلى الحد الذي أدى إلى عزوف معظم الناطقين بـــ "العربية" عن استخدام معجم لغتهم الأم، وقد زاد من تفاقم الأزمة تعقد بنية المعجم العربي نظراً لطغيان الخاصية الاشتقاقية للغة العربية[2] على ظاهرة تكوين الكلمات العربية، والتي تضيف طبقات أخرى لعمق بنيته، وأكاد أزعم أن فهم أسس وتفاصيل هذه البنية المعقدة كانت، ومازالت، وستظل إن بقى الحال على ما هو عليه، دون متناول معظم متخصصينا وباحثينا، ولا أمل في سبر أغوار المعجم العربي إلا باللجوء إلى الحاسوب، وطرح الكثير من المسلمات المعجمية واللغوية التي أدت إلى جمود النظرة إلى المعجم العربي في الماضي، فتطور العلوم المعجمية من المعاجم الورقية إلى المعاجم الإلكترونية بفضل ارتباط الرقمنة باللسانيات، هو مجال رغم حضوره في الدراسات العربية، لا يزال بعيدا عن حاجة اللسان العربي المعجمية إلى التطوير والإلمام بشتى قضايا المعجم التي لا تكون فاعلة الاّ بوجود مدونة لسانية تجمع شتات العربية وتعالجها معالجة آلية ولسانية[3] .

1- عبد الغني أبو العزم، الحاسوب والصناعة المعجمية، (اللسان العربي)، سابق، ص 44،45 .

2- تمام حسّان، اجتهادات لغوية، عالم الكتب،القاهرة، الطبعة الأولى، 2007، ص 142.

3- خليفة الميساوي، المصطلح اللساني وتأسيس المفهوم، مرجع سابق، ص 32.

وللإلمام بالجونب الأساسية لهذه الدراسة سنتطرق في المبحث التالي إلى تاريخ حوسبة المعجم الإلكتروني، يتبعه سرد موجز للبنية الأساسية للمعجم، ثم مناقشة الجوانب الرئيسية لحوسبة المعجم العربي، وتحديد منطلقات تحديثه وتطويره. على ضوء هذه الخلفية، يتم رسم الملامح العامة لمنظومة المعجم، نطرحها هنا في صورتها الموسعة بافتراض حوسبته، لنقوم بعد ذلك بتوصيف مكونات هذه المنظومة، والعلاقات البينية التي تربط بين هذه المكونات. لينتهي الفصل بتحديد المطالب المختلفة لحوسبة المعجم العربي.

ثالثا: من المعاجم الورقية إلى المعاجم المحوسبة: نظرة تاريخية:

يتطلّع جلّ الباحثيين في صناعة المعاجم الحاسوبية[1] إلى رؤية العمل المعجمي يفيد من المعلوماتية أقصى الفائدة لأنّه أطوع علوم اللسان إلى الحوسبة. وتكون وجهة المعجم نحو الحاسوب بطريقتين، أولاهما أن يتحوّل المعجم الورقي إلى حاسوبيّ ينسب إلى الحاسوب انتسابا مثلما كان ينسب إلى الورق، وليس هو من صلب نظامه، والثّانية أن يعدّ المعجم سلفا ويصاغ منذ بدايته بحسب متطلّبات النّظام المعلوماتي ليكون معجما محوسبا. فالفرق إذن حقيقي بين المعجم الحاسوبيّ والمعجم المحوسب.

فإن المتأمل في تراثنا المعجمي على ما فيه من عيوب ونقائص أثارها الباحثون لا يشك في أنّه تراث بالغ الأهميّة في الحفاظ على رصيد لغويّ للسان قديم يمتد على أكثر من أربعة عشر قرنا[2]، ويمثّل بمختلف مستوياته كلاّ لا يتجزّأ [3].

1- اقتحم العرب مجال الإحصاء اللغوي الحاسوبي منذ السبعينات، وعقد المؤتمرات والندوات العالمية والعربية، كان أهمّها: المؤتمر الثاني حول اللغويات الحسابية العربية الذي انعقد بالكويت عام: 1989، وسبقه الملتقى الرابع للسانيات العربية والإعلامية بتونس، كما يعدّ على رأس المشتغلين بحوسبة الدراسات اللغوية د.عبد الرحمان الحاج صالح الذي قدّم عدّة أعمال رائدة حول العلاج الآلي للنصوص العربية، وبحوث د. عبدالقادر الفاسي الفهري حول حوسبة المعجم العربي، و د.محمّد الحنّاش الذي قام بدراسات حول المعجم الالكتروني للغة العربية، واقترح مشروع نظرية حاسوب لسانية في سبيل بناء معاجم آلية للغة العربية، و د. محمّد حشيش الذي قدّم بحوثا حول معالجة اللغة العربية بالحاسوب، و د. نبيل علي الذي قدّم أعمالا رائدة منها: اللغة العربية والحاسوب، وميكنة المعجم العربي باستخدام المعالج الصرفي الآلي وبحثه: الجيل الخامس ومعالجة اللغة العربية آليا.

إنّ استخدام الحاسوب في صناعة المعاجم وتصنيفها يعدّ ظاهرة مثالية للمعالجة اللغوية، فقد استطاعت الدراسات اللسانية تحقيق تقدم ملموس ظهرت نتائجه على شكل معاجم آلية قابلة للاستعمال العادي من قبل الباحثين والمتعلمين، إذ تساعد حوسبة المعجم العربي على تسهيل معجمية الرصيد اللغوي العربي الثري في حافظات برمجية جاهزة للتسيير وفق الأغراض المعجمية المنشودة من حيث الإحصاء والوصف والتعدد الدلالي والتوزع اللغوي الصوتي أو الصرفي أو النحوي أو البلاغي أو الاصطلاحي والمجالات الابلاغية والاتصالية .

2- هذا اللسان كان في وقت من أوقات حمّال حضارات العرب والفرس والهنود واليونان ظلّ منذ تتالت عهود الانحطاط على الغرب يعمل بقدر ضئيل من طاقته لأنّ مستعمليه يجهلون قدرا كبيرا من إمكاناته المعجميّة ويبحثون في الوقت ذاته عمّا يمكن أن يسدّوا به ما يجدونه ثغرات في معجمهم، كالذي لا يدري أنّه ينام على كنز ويتضور جوعا. ومثال ذلك أنّ حديقة حيوانات فيها فيل وناقة وفرس وظبي ودبّ ونعامة ومعها صغارها فكم من مستعمل للّسان العربيّ سيقول: هذا فيل ودغفل، وهذه ناقة وحوار، وهذا فرس ومهر، وهذا ظبي وخشف، وهذا دب وديسم، وهذه نعامة ورأل؟ وكم متكلّم يذكر عند وصفه الألوان الأمهق والأزهر والأقهب والأبرش والأرمك والأصهب وعشرات أسماء الألوان الأخرى.

3- إبراهيم بن مراد: مسائل في المعجم، دار الغرب الإسلامي، 1997، ص 6 .

إنّ إحياء المعجم في حياة النّاس ليسمّوا الأشياء بأسمائها وبكلّ دقّة حاجة أكيدة إذن، وليس أفضل من قواميسنا المنجزة لتحقيق تلك الغاية. ولها عندما تصبح حاسوبيّة فضائل أخرى، أهمّها سرعة وصول الطّالب إلى طلبته وإمكانيّة البحث في عدد من القواميس في آن واحد. ولا ينتهي دور معاجمنا التّراثية عند هذا الحدّ، بل إنّ لها كذلك دورها في وضع المعاجم الحديثة المحوسبة.

1- المعاجم المحوسبة :

يتطلّب وضع مثل هذه القواميس تضافر جهود صنفين من الخبراء: خبراء اللغة من ناحية وخبراء تقنيات الحوسبة من ناحية ثانية. وسنتكلّم فيما يخصّنا في الصّف الأوّل[1].

فالمعجميّ يحتاج في سبيل وضع المعجم الحديث المحوسب إلى إنجاز أمور، أوّلها إعداد مدوّنة سياقيّة تضمّ الألفاظ مستعملة في سياقتها مستخرجة من المؤلّفات والمجلاّت والصحف و وسائل الإعلام وألسنة النّاس[2].

والأمر الثّاني هو إعداد قاعدة معجميّة فيها مجمل الألفاظ المستخرجة من المدونّة السّابقة ومن جميع قواميسنا العربيّة القديمة منها والحديثة التي تمّ تحويلها إلى قواميس حاسوبيّة.والأمر الثّالث استخراج الوحدات المعجمية[3] من الألفاظ ووضع نصوصها القاموسيّة، والأمر الرّابع هو إعداد قاعدة في التعالّق المعجميّ الذي يخصّ كل لفظ من ألفاظ القاعدة المعجميّة[4].

2- تطور تقنيات الحوامل و الصناعة المعجمية:

للإبحار في عالم المعجم والمعجمية، لابد أن نذكر، باقتضاب، بتاريخ تطور الصناعة المعجمية الوثائقية التي انتقلت عبر العصور من الشفاهي إلى المكتوب، والمدون فالمحفوظ بواسطة

1- هلال بن الحسين، "حاجة القاموسية العربية إلى حوسبة القواميس العربية الورقية ووضع القواميس الإلكترونية الجديدة"، الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، الرباط، المغرب، العدد السابع و الثامن، 1430هـ 2009م، ص 196.

2- يعمل النّاس على سدّ الثّغرات في المعجم دون انتظار آراء المختصين في ذلك،ولهم أحيانا طرق أنيقة في الاشتقاق حتّى ولو كان ذلك من الأعجميّ، لأنّه اشتقاق ينضوي تحت نظم اللسان العربيّ، ويمكن حينئذ إجازة مثل هذه المولّدات وإلحاقها بقاعدة الألفاظ، ومثالها " ميّل " أي أرسل رسالة إلكترونيّة وهي مشتقة من e-mail تأتي على وزن فعّل؛ و "مجور" التي اشتقّها الطّلبة من Etre major dans sa promotion وتروج كثيرا على ألسنتهم، يعنون بها الطّالب الذي يكون الأول في دفعته، وتأتي على وزن الرباعي فعلل.

3- يتطلب الوصف المعجمي الشّامل للدلالة والتّعالق المعجميّين اعتماد مفهوم الوحدات المعجميّة، وهي الوحدات التي ستدرج مداخل في القاموس، ذلك أنّ مدخل النص المعجمي في سبيل أن تستقصي دلالاته وطرق تعالقه المعجميّ لا يمكن إلاّ أن يختصّ بدلالة واحدة، وهي مستخرجة من خلال اللفظ بحسب معايير معجمية. ينظر: Introduction čVocable , Mel'u k (I),Clas (A) et Polguère (A) , à la l'lexicologie explicative et combinatoire, Louvin-la-Neuve (Belgique), Ducolot, 1995 pp. 156, 159, 160, 167, 184, 214 (ILEC).

4 – هلال بن الحسين، "حاجة القاموسية العربية إلى حوسبة القواميس العربية الورقية ووضع القواميس الإلكترونية الجديدة"، الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، سابق، ص 196.

الأحجار والرسومات ثم المخطوطات البردية في أول الأمر، والأوراق والأشرطة المغناطيسية والأقراص الإلكترونية المرنة والمدمجة والصلبة في عصرنا الحالي، وكل الحوامل[1] المستعملة في تدبير المعلومات والبيانات. كل هذه الوسائل ساهمت إلى حد بعيد، في صياغة مناهج التصنيف والترتيب والتبويب والتأليف[2]. ففي بداية الكتابة طرحت بحدة مسألة تنظيم الوثائق والبحث عن المعلومات وخاصة مع النصوص الدينية والطبية والعلمية والفلسفية والأدبية. وتعود أول خزانة في تاريخ البشرية إلى حوالي 688 سنة قبل الميلاد وكانت تضم الرسائل الرسمية والنصوص الدينية ومخطوطات الطب والهندسة وكل الوثائق المتعلقة بالحضارة البابلية والآشورية، حيث كانوا يصنفون المعلومات على قوائم مُجَدْوَلة[3]، كما أن الفراعنة في عهد رامسيس (القرن الثالث عشر قبل الميلاد) وضعوا الوثائق على مُدرجَات المرسومة ومصنوعة من الورق البري. ويعود مصطلح المكتبة إلى الحضارة اليونانية(في القرن السادس قبل الميلاد) حيث كانت الغاية تصنيف كل الأعمال الوثائقية الكونية (حوالي 500000 مُدْرَجاً بَرْدياً). وفي العهد الروماني صنفت تلك المدرجات في خزائن خشبية متحركة إذ في ذلك الزمن برزت عملية تصنيف الكتب ونشرها[4]. أما في العصور الوسطى فقد احتضنت الكاتدرائيات والكنائس الوثائق السياسية والدينية والعلمية، بدافع الاهتمام بالدراسات الدينية فقد صمم نظام الفهرسة لتسهيل عملية وضع المعلومات والبحث عنها. وقد ازدهرت حقبة الحضارة العربية بشكل مثير فضمت في خزاناتها ببغداد والقاهرة وقرطبة ومراكش بصفة خاصة (أزيد من 400000 كتاب). وقد اتسم القرن الخامس عشر بابتكار أول مطبعة سميت بمطبعة كوتنبورغ حيث أصبح منذ تلك الفترة الإنتاج غزيراً للمطبوعات والمنشورات المتنوعة الأشكال والأحجام ومنها المؤلفات المعجمية.

إن تطور الصناعة المعجمية عبر هذه الحوامل التي ذكرناها، من المخطوط البرديّ مروراً بالمخطوطات والمصنفات الورقية جعل الكتاب وبالتالي المعاجم الورقية تُنْشَر على أوسع نطاق، بعد ذلك بقليل، بعد الثورة التكنولوجية أصبحت المعاجم والقواميس محوسبة وإلكترونية، كما استقر عليه الحال الآن، وبعد تأثر هذه الصناعة بالثورة المعلوماتية الحديثة والمعاصرة استفادت الصناعة المعجمية والنظرية المعجمية، على حد سواء من هذه الثورة التكنولوجية

1 − Supports.

2 − Claude de Loupy, Evaluation de l' apport de connaissances Linguistique en Désambuisation sémantique et rechercher Documentaire, thèse de Doctorat présentée à l Université d Avignon et des Pays de Vaucluse. 2000.p.10.

3− المرجع نفسه، ص11.

4− ظهرت في تلك الحقبة التاريخية مصنفات قابلة للنشر والتوزيع عبارة عن مصنف "الكتاب" بالمعنى الحالي للكلمة تقريبا حيث يلف في طياته مجموعة من الأوراق مرتبطة فيما بينها بحسب الموضوعات مسفرٌ من الواجهتين بغلاف متين.

الهامة فكيف أمكن لها ذلك؟، وكيف أفاد العمل المعجمي بدوره هذا التقدم التكنولوجي، خاصة في مجالات معالجة اللغات الطبيعية وهندسة المعرفة والذكاء الاصطناعي[1]؟.

كل هذه الأسئلة أصبحت تشغل بال علماء المعجم (معجماتيين ومعجميين) وتؤرقهم هذه الثورة الإلكترونية السريعة، وكيف كانت النتيجة بالتفاعل الجدلي فاستلهم بعضهما البعض صورنةً وتوصيفاً. وفي ضوء هذا التطور بدأ هؤلاء العلماء إدراك مدى أهمية إعادة النظر في صياغة المفاهيم والتصورات العلمية[2] لتكون أكثر واقعية وقرباً من الظواهر الطبيعية وهذا ما أدى إلى إعادة ضبط المفاهيم من قبيل مفهوم المدخل الذي يشكل إحدى عوائق الصناعة المعجمية الإلكترونية.

3- الهندسة اللغوية والصناعة المعجمية:

3-1- الهندسة اللغوية :

لم يظهر مصطلح الهندسة اللغوية إلا في وقت قريب[3] مقارنة مع مصطلح اللسانية الحاسوبية أو المعلوماتية أو الآلية، ومعالجة اللغات الطبيعية آليا، أو تكنولوجيا اللغات .

لقد ترجم المصطلح[4] إلى العربية بمقابلات عديدة نختار منها على سبيل الذكر لا الحصر الهندسة اللسانية، الهندسة اللغوية، هندسة اللغات، علم اللغات الهندسي، اللغويات الهندسية، الهندسيات اللغويات، هندسة اللغويات وغيرها كثير[5]. وهذه المصطلحات وإن كانت متعددة ومختلفة من حيث الدلالة الصوتية، فإنها تكاد تشترك في الدلالة الاصطلاحية من حيث كونها تدل على العلم الجديد الذي ظهر بموجب الزواج بين الإنسان والحاسوب بواسطة اللغة الطبيعية قصد معالجة الإشكاليات التي تطرحها المعالجة الآلية للغة الطبيعية.

الهندسة اللغوية إذن محاولة لتجريب المعرفة اللغوية على أنظمة حاسوبية ذكية قادرة على التعرف على اللغة وفهمها وتأملها وإعادة إنتاجها في صور وأشكال متخصصة، وهي بهذا المعنى تمثل اتجاها هندسيا لسانيا يوظف المعارف والنظريات التي تساعد على فهم اللغات الطبيعية وإنتاجها[6].

---

1 – L’intelligence artificielle.

2– انظر، جان بريفو، المعاجم والتقنيات الجديدة، مراجعة وعرض الكتب، مجلة الدراسات المعجمية ع 3 و 4 يناير 2005. ص237. وللمزيد انظر Jean Pruvost, Dictionnaire et nouvelles technologies, coll Ecritures électronique, PUF. Paris. 2000.

3– علي فرغلي، الهندسة اللغوية والثورة المعلوماتية، المؤتمر الثاني عن التحيز: حوار الحضارات والمسارات المتنوعة للمعرفة، القاهرة، 2007.

4 - Engineering Language أو Ingenierie Linguistique أو Ingénierie des langues.

5– مهديوي عمر، اللغة العربية والهندسة اللغوية، سابق،أ 2007م وب 2008م.

6– مهديوي عمر، التوليد الصوري للأسماء البسيطة في اللغة العربية، المؤتمر الدولي الثاني حول هندسة اللغة وهندسة اللغة العربية، 27،28.

وهكذا فالمجالات التي تشتغل عليها المعالجة الآلية للغات أو الهندسة اللغوية نوعان رئيسيان: معالجة اللغة المكتوبة، ومعالجة اللغة المنطوقة أو معالجة الكلام، وهذا المجال الأخير يندرج ضمن ما يصطلح عليه بمعالجة الإشارة[1]. وتجب الإشارة إلى أن أغلب الدراسات والأبحاث الحاسوبية واللسانية قد تركزت حول القسم الأول المرتبط أساسا باللغة المكتوبة، بينما بقي القسم الثاني مهمشا لمدة ليست بالقصيرة، ولم يحظ بعناية محدودة نظرا لما تطرح معالجته من صعوبات وتعقيدات.

تعتمد الهندسة اللسانية على أدوات تقنية وأخرى لسانية:تعبر الأدوات التقنية عن البرامج الحاسوبية أو المعلوماتية[2] ، فيما تشكل الدوات اللسانية قواعد معرفية قابلة للتشغيل والتفعيل وفق برامج حاسوبية ذات طبيعة صورية (خوارزمية).

3-2- الصناعة المعجمية للمعجم الحاسوبي :

بعدما ذكرنا بتطور تقنيات الحوامل عبر التاريخ والتي غيرت من معطيات النظرية المعجمية بالتدريج فأصبح لزاماً أن تغير منهجها مع أنواع الأقراص الصلبة والمرنة والمدمجة والنشر الإلكتروني والافتراضي على الشابكة. أمام هذه الثورة التي عرفتها الصناعة المعجمية منهجا وإنتاجاً تحرر المعجمي، كما يقول جان بريفيو[3] من مجموعة من الاكراهات كانت تعترض سبيله وتحول بينه وبين انجاز المعجم الشامل الجامع الأمثل الذي يحلم به وبين تحديث معجمه

1- تهتم الهندسة اللغوية بدراسة تطبيقات لغوية آلية من قبيل:

- التصحيح الآلي للأخطاء كأخطاء الطباعة والإملاء والنحو والأسلوب.
- استخراج الكلمات من النص وفهرستها، وهذا يرتبط اشد الارتباط بالتكشيف الآلي والتوثيق الآلي ومحركات البحث Les moteurs de recherche.
- بناء قواعد البيانات أو المعارف الصرفية والمعجمية في اللغات الطبيعية.
- التوليد الآلي للنصوص، ويتوخى هذا الميدان توليد المفردات والجمل بواسطة قواعد صورية (خوارزمية) موجهة لخدمة أغراض في الفونولوجيا والمورفولوجيا والتركيب والمعجم.
- دراسة اللغة الطبيعية بواسطة الأتومات والربوت .
- المشكل الآلي للنصوص جزئيا أو كليا.
- الترجمة الآلية ثنائية الاتجاه من اللغة المصدر إلى الهدف أو العكس.
- التلخيص الآلي، وينقسم إلى قسمين: تلخيص مبني على فهم سطحي للنص وتلخيص مبني على فهم عميق للنص.
- الاستخراج الآلي للخصائص اللغوية أو ما يسمى بنظم استرجاع المعلومات.
- القاموس الإلكتروني وتطبيقاته في اللغات الطبيعية.

2 -Les logiciels informatiques.

3- انظر، جان بريفيو مرجع سابق، ص 236.

بإصدارات جديدة. فهذه هي المزية الكبرى التي ألغت تماماً زمن المعجم الثابت الذي يدوم سنوات بل قرون.

وإذا أردنا أن نقوم بإحصاء متون المعاجم المحوسبة في اللغات العالمية في السنوات الأخيرة سنجدها جلها مكونة ومعتمدة على نصوص معجمية لقواميس ورقية أو لمكانز وبنوك للمعطيات المعجمية[1]، فمثلا بالنسبة للغة الفرنسية هناك متن قاموسي كبير استخرج من قاعدة بيانات أعدت لإنجاز بنك اللغة الفرنسية (TLF) ونفس الشيء بالنسبة لأولى القواميس على القرص (روبير الكبير 1986) ومع بزوغ القرن الواحد والعشرين نشر أول بنك للغة الفرنسية في نوفمبر 2004 (بتعاون المركز الوطني للبحث العلمي الفرنسي) وهو عبارة عن قرص مدمج، وتذهب كريستين جاكيت بفو (Christine Jaquet-pfau)إلى اعتبار "أننا إزاء تصور جديد وأصيل للقاموسية، ذلك أنه عمل لا مثيل له من حيث جرأته في طرح القاموسية المعلوماتية للتداول لدى الناشرين، الشأن لا يتعلق ببناء القواميس فحسب، بل الهدف منه كذلك هو معالجة اللغات الطبيعية في سائر مستوياتها وتطبيقاتها. ولكن الأمر سرعان ما تحول فسُمح به إلى حد اعتبار شكلا آليا مُبَنْيناً[2]، ومن بات أكيداً استثمار هذا التوجه في العمل المعجمي، ولكن ثمة تساؤلات تطرح دوماً وهي كيف تجعل هذا العمل صميم عمل القواميس المحوسبة؟ فإذا حصل أن غيرت التكنولوجيا المعلوماتية طرق ولوجها للبيانات الموجودة بالملفات التي يراد بها صناعة القواميس، علاوة على مغالاة الناشرين في استعمال هذه المنهجية الجديدة، فإنه سيفسر بالعديد من أنواع القيود – تضيف كريستين جاكيت برفو – تقف عند تطوير التطبيقات الجديدة، وقد بلغ هذا حده بسرعة كبيرة تمثل في محطات كثيرة. ناهيك على تشجيع الناشرين المتخصصين في مثل هذه القواميس بدوافع تربوية واقتصادية وللاحترام التقليدي للقواميس الورقية ولاعتبارات أخرى مرتبطة بثمن ويسر الولوج. هذه الأسباب الجوهرية التي تعيق شيئاً ما الصناعة القاموسية الإلكترونية، فهل يمكن الذهاب مع الذين يعتبرون ذلك تحد[3] ؟.

1 − Christine Jaquet-pfau, pour un nouveau dictionnaire informatisé, in Ela, revue dedidactologie des langues-cultures et lexicultrurologie, numéro spécial, sur Dictionnaires et Innovation, n 137, janvier – Mars 2005.p.53.

2− المرجع نفسه، الصفحة نفسها.

3− المرجع نفسه، ص 54.

4- المعاجم الورقية وظهور مجتمع المعلومات:

كما في أوروبا جميعها وفي فرنسا وفي الولايات المتحدة الأمريكية، ما زالت المعاجم الورقية حتى الآن تسيطر بشكل كبير في دور النشر[1]. ولقد أدى النجاح الشامل لهذه التكنولوجيا القاموساتية إلى اندثار مغامرات النشر وجداول الوراقات والمكتبات والفضاءات الكبيرة. فتوفير قرص متجدد في السوق، رغم بعض القيود التي تطرحها الوسائط المتعددة، أصبح متداولا، مع أن القاموس الورقي يؤمّنُ على الأقل ديمومة التصفح المتجدد باستمرار، وبقصد التعرف على مستجدات الإملاء والتعريف والشرح أو حتى الجرد وغالبا ما يكون الحافز في الشكل الأنيق للبرمجي. ومع ذلك فقد تحولت الأمور بشكل سريع ودقيق، فإذا تمعنا قليلا سوف نجد غلاف نسخ لاروس الصغير (2005) مثلا. يحمل كعنوان صغير ملحق به "**ومعه القرص المدمج**" وكذلك إمكانية الطباعة الفوقية. لننظر بشكل شفاف صور لاكروى حيث الأقراص المدمجة المحفورة. يمكن فقط فتح الصفحة الأولى للقاموس وستجد مرفقاً على ظهر الغلاف، ظرف بلاستيكي شفاف يحمل القرص نفسه، مغلق في جيب. حيث تمثل الكتابة هنا غلاف النشرة الورقية للقاموس، وتبدو هذه الطريقة مفضلة من قبل مستهلكي القواميس الورقية- الحاسوبية [2]. إن الفرق بينهما واضح وجلي: فالورقي قد يصل وزنه إلى ثلاث كيلوغرامات ويكفي تصفحه بحركة بسيطة للأصابع ومنظره جميل، بينما الإلكتروني لا يتعدى مائة الغرامات ويستدعي حاسوباً لتشغيله، أو ربما كذلك يكون امتلاكه عام بين الناس. ولكن الاثنين معاً يتوفران على نفس المحتوى والمضمون "على الوجه من التطابق" تماما لأن ذلك هو ما الاختلاف بينهما، فعلى "وجه التشابه" هذا يفسر تنظيم المعلومات والبحث عنها واستخراجها في القاموس الآلي وحده وبالطريقة التي سيشهدها هو ذاته في غياب تام للنسخة الورقية عنه.

1- في الذكرى المائوية لصناعة إحدى القواميس ذات الشهرة الكبيرة مثل (le millésime 2005) والذي أنتجه معجم لاروس المصور الصغير وبهدف الاحتفال بالطبعة المائة، قررت هذه الدار إضافة الثقة وطابع الجدة على القاموس من خلال طبعة (la dent- de – lion et la Semeuse) ونفس الأمر بالنسبة للرسم والخطوط على واجهة القاموس.

2- لقد نفذت نسخ ورقية- إلكترونية بعدد مثير وكبير مثل ما حصل مع قاموس: -LE PETIT LAROUSSE 2000, Havas La rousse :HER, 1999, CD –ROM PC « Larousse multimédia » Interactive,

وكذلك قاموس " مساعد روبير الصغير على القرص (2001 بالفرنسية)، هذا ما حصل أيضا في العالم العربي حين ظهور نسخ حاسوبية مع النسخة الورقية لقاموس المعجم الوسيط، أو الأقراص المدمجة منفردة مثل معاجم "جامع معاجم العربية" وغيرها من القواميس الأحادية والثنائية اللغة.

رابعا : المعجم الإلكتروني للغة العربية : المشروع الذي لابد منه :

يحتلُّ المعجم مكانةً سامية عند جميع الأمم التي تحافظ على لغاتها وتراثها، فهو ديوان اللّغة، وقد أصبح علم المعجم علمًا واسعًا ذا جوانب عديدة، وأصبح له نظريات تتناول أسس صناعته، وأصبحت الدراسات المعجمية تحتلُ حيزًا كبيرًا من الدراسات اللّغوية الحديثة، إذ أخذت هذه الصناعة تخضع لقواعد وأسس دقيقة، وصارت توزن بمعايير ثابتة تدّل على نضج هذا العلم.

إنّ المعجم العربي وسيلة لغوية كانت حرفة سابقًا، وأصبحت صناعة اليوم- والمعجم حرفة وصناعة قبل كل شيء- تتعلق بجمع اللّغة ووضعها؛ وقد سعى إلى وضع أسس تتصل باللّغة وبالخصوص بمفرداتها ومفاهيمها التي ترتبط ارتباطًا متينًا بعلوم لسانية شتى منها علم الدلالة والنحو والصرف[1].

إنّ المعاجم العربية هي أداة من أدوات هذه المواجهة التي يمكن أن نبدأ منها وبها. وهنا نجد أن أبرز ما لفت نظر الباحثين المعاصرين في هذه المعاجم هو المآخذ والعيوب التي أخذت تتردد في كل دراسة حول المعاجم العربية[2]، فعلى عظم الخدمات التي أدّتها للعربية وطلابها، وما زالت تؤديها حتى الآن، ظلت في الحقيقة عاجزة عن مسايرة النّهضة العربية الحديثة، وقاصرة عن متابعة التطور الكبير في عصر المعلومات، ممّا زاد الباحثين في مختلف ديارهم شُعورًا بالحاجة إلى معجم حديث يضاهي المعاجم المعروفة في اللّغات الأجنبية، ويتسع لمصطلحات العلوم المختلفة، على أن هذا الشعور مشروط بوجوب إنشاء معجم عربي حديث يواكب تكنولوجيا المعلومات بطريق الإفادة التي تشمل عليها المعجمات القديمة، استنادًا إلى خصائص اللّغة العربية ومرونتها إلى حدٍّ يمكن معه أن تستوعب كل جديد تدعو إليه ضرورة أو مصلحة[3].

فكان لا بدّ من خلال هذا التحدّي أن نتعرض في هذه الدراسة إلى إمكانية توصلنا إلى معجم حديث الذي يحقق طموحات الباحثين في ضوء استخدام "اللّسانيات العربية الحديثة" لحوسبة المعجم العربي[4].

فكان هذا الموضوع ليناً طبعًا للحوسبة إلى درجة أنّنا تعززت رؤانا تلك حول ماهية عملية حوسبة المعجم وأنّها مسألة تحدٍّ حاسوبي بالدرجة الأولى.

1- محمد رشاد الحمزاوي، "من قضايا المعجم العربي"، قديمًا وحديثًا دار الغرب الإسلامي، تونس، الطبعة الأولى، 1986، ص 05.

2- حلمي خليل، "مقدمة لدراسة التراث المعجمي العربي"، دار النهضة العربية، بيروت، الطبعة الأولى، 1997، ص 08.

3- محمد علي عبد الكريم الرديني، المعجمات العربية دراسة منهجية، دار الهدى، الجزائر، الطبعة الثانية، 2006، ص 201.

4- عبد الله أبو هيف، مستقبل اللغة العربية " حوسبة المعجم العربي ومشكلاته اللغوية والتقنية أنموذجا"، اللغة العربية، المجلس الأعلى للغة العربية، الجزائر، العدد العاشر، 2004، ص 127.

إن ظهور المعاجم وإشراقها في إطار ما تشهده التكنولوجيا من تطور مذهل، يقود المعجميين بصفة حتمية، إلى مراجعة أعمالهم بصفة دورية. بيد أنه في المقابل، في واقع العمل المعجمي الحالي، هناك سيادة لتوجه في أوساط الناشرين القاموسيين يهدف إلى تشجيع مستعملي القواميس التقليدية[1] على اقتناء تلك الموضوعة على الأقراص المدمجة، هذه العملية تيسر ولوج المعلومات وتسهلها فتتركهم بين اختيارين، حسب إمكانياتهم، هذا ما له الدور الإيجابي من حيث النتائج.

إنه تطور أكثر منه ثورة، بحيث أن المرور من النسخة الورقية إلى النسخة المحوسبة لا ينفي بعضهما البعض. إذ هناك احترام يتمثل بالارتباط الوثيق بالنسخة الورقية من قبل الناشرين الذين يفضلون ما يسمى **بالانتقال الهادئ**. ذلك أن المحوسبة لازالت في لغات عديدة ومنها اللغة العربية على الأقل، مجرد تسجيل إلكتروني للنسخة الورقية ليس إلا[2].

أما في العالم العربي فقد سلك في هذا الإطار ثلة من علماء المعجمية، فأعَدوا معاجم حاسوبية تستجيب لمطلب التكنولوجيا المعاصرة. فكان أن أنجز أحد هؤلاء وهو المعجمي الكبير عبد الغني أبو العزم عملا معجميا[3] أصدرته شركة صخر بالقاهرة على قرص مدمج بدون أن يمر بالنسخة الورقية[4].

نحن حقيقة، بحاجة ماسة لتطوير المعاجم التراثية والحديثة الورقية منها والمحوسبة شكلا ومضمونا من خلال رصد دقيق لألفاظ الحضارة ولمصطلحات العلوم في كل فروع

1- إن السيرورة هنا تقتضي العمل على تكييف القاموس اليدوي انطلاقا من تحويل المحتوى الموجود إلى وثيقة محوسبة، بحيث تسمح أدوات البحث بإيجاد بعض أنواع المعلومات. فمنذ ذلك الحين وبعيداً عن كل ما يثير الجودة في محتوى هذه المعاجم، فإن الأمر يطرح على مستوى مدى جدوى الاستغلال المعلومياتي لهذه القواميس، وبالدخول مباشرة في بناء معاجم إلكترونية حقاً وليدة الأدوات العقلانية للتكنولوجيا والعلوم المنبثقة عنها. ألا يمكننا بهذا المنظور ترك النسخة الورقية؟ ونقم مباشرة بوضع تصور جديد يستجيب لمتطلبات النشر المحوسب؟ إن الحل الأمثل هنا لن يكون مع حالة الموازنة بين الورقي والإلكتروني، بل باستغلال إيجابيات بعضهما البعض، فالمعاجم العربية المحوسبة فهي على غرار المعاجم في اللغات العالمية الموضوعة على الحواسيب والأقراص المدمجة، إذ لم تضف سوى بعض التعديلات الطفيفة مثل التعديل الخاص بالإخراج المريح وكذا استعمال الألوان بمثابة حقائب حقول المعلومات أو ربما قد تخضع لبعض التعديلات حسب ما تتيحه المستجدات التكنولوجية1. العديد من موضوعات المعاجم في لغات كثيرة غير متوفرة على أساسيات المعطيات المعجمية مثل قاموس الأسماء الأعلام وقاموس المركبات وقاموس المتلازمات، وقد بدأ الاهتمام كما هو معلوم، نظراً للطلب المتزايد للتطبيقات التكنولوجية مثل الترجمة الآلية وأنظمة الاتصال وما تطرحه من تحديات أمام الأدوات اللسانية الطبيعية. وهذا العمل سيوفر علينا كثيرا من العناء وسيكون أكثر إفادة ومنفعة في كثير من التطبيقات الإلكترونية.

2- عز الدين غازي، بناء المعجم الآلي للغة العربية:معالجة التعبيرات المركبة والمتلازمة،أطروحة جامعية لنيل الدكتوراه، جامعة الحسن الثاني، عين الشق، الدار البيضاء،2008/2007، ص 134.

3- "معجم الغني" لصاحبه د عبد الغني أبو العزم اعتمد فيه مجموعة من الإجراءات أهمها: جمع المادة المدونة اللازمة لعمله المعجمي من نصوص ومعاجم تراثية ودينية ونصوص معاصرة وأدبية وإعلامية، فكان النص المعجمي بالصورة التي هي عليه الآن في المعجم المذكور. ثم اختار مداخل المعجم ومكوناته بحسب الرؤية التي تصورها في نظريته المعجمية، أما من حيث أسلوب تخزين المعجم فقد جاء بالشكل الحاسوبي الأنيق ميسراً الولوج بأدوات تربوية وتوصيلية جذابة غير مرهقة

4- عز الدين غازي، بناء المعجم الآلي للغة العربية:معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص 135.

المعرفة الإنسانية والعلوم والتكنولوجيا حتى تكتمل صورة اللغة العربية عبر القرون من جهة، ومن جهة أخرى الاستغلال العقلاني للنظريات المعجمية العربية وغير العربية القديمة والحديثة والمعاصرة[1].

أما فيما يتعلق بالاستفادة من التقنيات التكنولوجية المتاحة فقد طرحت مجموعة من المشاكل هي من صميم النظام اللغوي للعربية وخاصة الضبط بالحركات والتخزين الآلي للمداخل المعجمية. إن حوسبة المعجم العربي كما تصوره الدارس محمد الحناش في العديد من أعماله[2]. هدفه إعداد معجم يكون بمثابة أداة لنظام التعلم. وقد وضعت لهذه المنهجية متكاملة تمثلت في معالجة جميع الجذور العربية باعتبارها، صرفياً، أصل كل اشتقاق مفرداتي كتوليد الأسماء والأفعال منها. ثم معالجة المفردات البسيطة المعربة والمفردات المركبة المعربة وذلك على غرار المعاجم الآلية للغة الفرنسية[3].

4-1- المعاجم الحاسوبية والمحوسبة: نظرة عامة:

تعددت المعاجم المحوسبة لأغراض وأهداف وظيفية أملتها الثورة التكنولوجية والهندسة المعلوماتية والتطبيقات الحاسوبية. وسنعرض بإيجاز لأهم فئات هذا النوع من المعاجم[4]:

أ- المعاجم الحاسوبية أحادية اللغة :

إن المعاجم العربية أحادية اللغة تتبع ثلاث طرق رئيسية في ترتيب المداخل[5]، فهنا قد تكون أداة لفهم النصوص أو التعبير بلغة ما. ولكن يكفينا هنا التمييز بين معاجم الألفاظ والتعبيرات

1- نذكر على سبيل المثال من المعاصرين منهم، اللغوي الألماني هانز فير الذي كرس حياته في جميع بطاقات لمادة عربية من الصحف ومن الكتب العربية وقد توج بعمل كبير يضم المداخل العربية مع الشرح بالألمانية كما قد سبقه فرايتاج (1861-1788) الذي صنف المعجم العربي اللاتيني في أربعة أجزاء وعنه أخذ كازيمارسي في معجمه في معجمه العربي الفرنسي بتعاون مع فيستنفلد: معجم البلدان لياقوت إلى جانب كوزيجارتن (1850-1792) الذي وضع مقارنة معجمية بين العربية واللغات السامية (للمزيد انظر، الدكتور محمد لهلال، جهود اللغوين الألمان في تطوير اللغة العربية، المؤتمر العالمي ألمانيا والعالم العربي، الإمارات العربية المتحدة 2001). وقد عممت هذه التجربة على باقي اللغات حيث نجد على مستوى اللغة العربية أعمال كل ممن المصري الدكتور محمود فهمي حجازي والتونسيين الدكتور رشاد حمزاوي والدكتور إبراهيم بن مراد والمغربي الدكتور عبد الغني أبو العزم وغيرهم من جهابذة المعجم العربي ممن لم نذكرهم هنا.

2- محمد الحناش، مرجع سابق، (1992) و(1993) و(1996).

3- ماكس سلبر شتاين، مرجع سابق، (1988) و(1993).

4- عز الدين غازي، بناء المعجم الآلي للغة العربية:معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص 136.

5- اختلف العلماء والباحثون في تقسيم المعاجم العربية، فقد قسمها الدكتور حسين نصار إلى أربعة أقسام حسب ترتيب مداخلها: مدرسة الترتيب الصوتي والتقاليب ومدرسة الترتيب حسب الأبنية ومدرسة الترتيب الألفبائي حسب الأواخر ومدرسة الترتيب الألفبائي حسب الأوائل. وقسمها الدكتور علي القاسمي إلى سبعة أقسام: الترتيب العشوائي والترتيب المبوب والترتيب الموضوعي والترتيب الدلالي والترتيب النحوي والترتيب الجذري والترتيب الهجائي.

المتلازمة والأمثال من جهة، فكلها أدوات معينة في فهم النصوص المسموعة والمقروءة. ثم هناك معجم المعاني والمترادفات من جهة أخرى، وهي أدوات معينة في التعبير والترجمة، حيث تساعد الباحث في التعرف على الألفاظ الدقيقة المناسبة للتعبير عن فكرة ما في سياق ما. وهناك عدد من معاجم الألفاظ العربية والأجنبية المحوسبة التي يستطيع المترجم الاستفادة منها. بل إن بعض برامج تنسيق النصوص تتوفر بها مكانز عبارة عن معاجم للمترادفات وللمشترك اللفظي.

ب- المعاجم الحاسوبية الثنائية:

يميل البعض إلى تسمية هذا النوع من المعاجم بالقواميس الحاسوبية، أي قاموس ثنائي اللغة الذي يقدم للمترجم المقابلات اللفظية بين اللغات المخزونة في ذاكرة الحاسوب.

وتتوفر في الأسواق أعداد من هذه المعاجم في صورة أقراص للاستعمال على الحواسيب الشخصية أو في صورة حواسيب بحجم الهواتف المتنقلة (حواسيب الجيب). وهي في تطور مستمر من حيث النوع وحجم مخزوناتها. فبعضها لا يكتفي بتقديم الألفاظ ومقابلاتها بل تقدم كذلك بعض التعبيرات والجمل الشائعة ومقابلاتها في اللغة الأجنبية. كذلك نجد أن بعض حواسيب الجيب تتيح إمكانية إضافة في صورة شرائح > لألفاظ أو مصطلحات متخصصة في حقول المعرفة والعلوم المختلفة. وجدير بالذكر أن بعض برامج تنسيق النصوص متعددة اللغات أو ثنائية توفر معجما أليا ثنائي اللغة أو حتى متعدد اللغات لمستخدميها[1]، إضافة إلى مدققات إملائية وتركيبية آلية...

جـ- الموسوعات المحوسبة:

من الأدوات المساعدة التي يغفل عنها كثير من المترجمين، الموسوعات أو دوائر المعارف العامة والمتخصصة أو حتى الأطالس، والتي تعتبر معيناً مهمّاً للمعلومات المفيدة للمترجمين. ونجد في الأسواق عددا من هذه المووعات في صور محوسبة.

1- هناك عديد من الشركات والمراكز البحثية التي استثمرت في الترجمة الآلية المتعددة اللغات: نذكر على سبيل المثال شركة صخر ميديا (www.sakhr.com) وشركة (أي بي إم) (www.ibm.com) وشركة (إنفو أراب) وشركة (كوتلك) (www.coltec.com) وشركة سيموس التي أنجزت المترجم الآلي (Nakel) المتركز على الترجمة الدلالية التي أعطت الأهمية البالغة للتعبيرات المتلازمة للمزيد هنا ينظر محمد عز الدين مدير الشركة بالموقع التالي (www.cimos.com). وهناك كذلك خدمات محركات البحث الكثيرة(...google altavisa) والتي توفر بموقعها المترجم الآلي المتعدد للغات للمستخدم حيث يقوم بترجمة النص المرغوب فيه أثناء البحث والتصفح وإخضاع معالجة اللغة آليا حسب المحللات اللغوية الآلية المتاحة على البرامج.

ونشير إلى أن هذه المراجع العامة والمتخصصة تأتي في صور مختلفة، أهمها الأقراص المضغوطة – أو المدمجة كما يسميها البعض – أو عن طريق شبكة الإنترنت، وذلك بالإضافة إلى معاجم الجيب وبرامج وبيئات التنمية اللغوية لتحليل ومعالجة النصوص.

4-2- تسنين المعاجم الإلكترونية:

تهدف هذه العملية، وهي ضرورة تقنية حتمية، إلى بناء مجموعة من الاتفاقات على مستوى المداخل المعجمية على مستوى القيم العليا المتمثلة في صفحة العنوان والأدوات المدخلية تقسيم الكتل إلى مقولات نحوية مثل الاسم والفعل والحرف والظرف والصفة وأسماء الأعلام وكذا إلى لغات ثنائية مختلفة. فهذه القيم المتمثلة في الرموز والأرقام هي من نفس الطبيعة التي تتميز بها أنماط أخرى مؤلفة للنص.

طبقت هذه العملية على لغات طبيعية عالمية، وبصفة خاصة على القواميس المتوسطة الحجم (CE) (كلاروس الصغير وربير الصغير والقاموس الإنجليزي كولينس). أما القواميس التاريخية التقليدية كقاموس أو كسفورد للغة الإنجليزية أو أيضا مكنز اللغة بقيت جانبا ولم تعالج بهذه الطريقة وإنما وضعت في أقراص مدمجة. مثل قاموس (TLF) الفرنسي[1].

ومن حيث المكونات الأساسية لأي معجم محوسب، هناك مجموعة من المعلومات محددة تظهر كرموز في المداخل القاموسية: معلومات حول شكل المفردة من حيث الإملاء والنطق والكتابة ...ومعلومات أخرى نحوية تخص المقولة النحوية والأصناف الصغرى والقيم الصرفية ...علاوة على ذلك على التعاريف والاستشهاد والترجمة والنحت والإحالة وأصناف المداخل وبيانات الاستعمال والأمثلة وكل هذه المقولات والأجزاء ترفق بأرقام وقيم ورموز يراها اللساني الحاسوبي ضرورية في تسهيل عملية البرمجة من جهة وعملية البحث من جهة أخرى.

1- للمزيد من التفصيل ينظر : Christine Jaquet-pfau, pour un nouveau dictionnaire informatisé, in revue ELA , n 137, janvier – Mars 2005.p.47.

خامسا- قواعد المعطيات النصية لتطوير المعجم الحاسوبي (الإلكتروني) :

5-1- الملامح الأساسية لحوسبة المعجم العربي :

إن أهمية المعجم كخزان للغة وأدائها في مواكبة تطورها أدت بالدراسات اللّسانية إلى خوض غمار هذا التحدّي بتطويع تقنيات الحاسوب للمعجم العربي كموضوع مثير للمعالجة الآلية، فقد عدّت حوسبة المعجم من أهم مجالات اللّسانيات الحاسوبية وأكثرها تلبية للمتطلبات العلمية[1]. كالمعالجة الآلية للعناصر المعجمية وبرمجتها من حيث التصنيف والتخزين والمرجعية وتحويل النّص المعجمي[2] إلى نصّ إلكتروني ممنّهل.

إن المعجم منظومة معقدة بل غاية في التعقيد، لدرجة يتعذر معها فهم أدائها ومحاصرة ظواهرها في ظل القيود التي تفرضها الوسائل اليدوية لبناء المعاجم وتحديثها واستخدامها ولم تعد معالجة المعجم آليًا أمرًا من قبيل الرفاهية الفنية، بل مطلبًا أساسيًا تفرضه طبيعة المعجم من حيث مضمونه وتنظيمه وتحديثه وخدماته للبشر ولنظمهم الآلية[3]، فقيمته تتجلى أساسًا في قدرته على تخزين المادة وترتيبها طبقًا للنظام المطلوب، ولا أمل في سبر أغوار المعجم العربي إلاّ باللّجوء إلى الحاسوب من خلال استغلال طاقة تخزينه الهائلة وقدرته الفائقة على الاسترجاع[4].

ولعلَّ إبراز حتمية معالجة المعجم العربي آليًا وتحديد منطلقات تحديثه وتطويره يؤدي بنا إلى رسم الملامح العامة لمنظومة المعجم نطرحها في هذا المشروع بافتراض ميكنته لنقوم بتوصيف مكونات هذه المنظومة والعلاقات البينية التي تربط بين هذه المكونات، بتحديد المطالب المختلفة التي أدت إلى معالجة المعجم آليًا، والتي نجملها في النقاط التالية[5]:

- قصور حاد في المصطلحات العلمية الحديثة.
- جمود النظرة إلى تكوين الكلمات.
- أحادية تنظيم مواد المعجم على أساس الجذر.
- إهمال العلاقات المعجمية.

1- محمود فهمي حجازي، "المعجمات العربية وموقعها بين المعجمات العالمية"، المجلس الوطني للثقافة والفنون والآداب، بحوث (ندوة خاصة بمناسبة الانتهاء من تحقيق وطباعة معجم تاج العروس من جواهر القاموس، الكويت، 9- 10، فبراير 2002.

2- محمد رشاد الحمزاوي، النص المعجمي في المولدات والأعجميات، مجلة المعجمية، عدد11،1995، ص19/9،

وينظر: محمد رشاد الحمزاوي، مقترح لوضع نموذج للمعجم العربي الحديث، المعاجم العربية الواقع والآفاق، الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية،العدد السادس،2007، ص74.

3- نبيل علي، "اللّغة العربية والحاسوب" دراسة بحثية، دار تعريب، الطبعة الأولى- الكويت 1988- ص 459.

4- المرجع نفسه، ص 499.

5- المرجع نفسه، ص 471، 473.

– ضمور عنصر الدلالة في الدراسات المعجمية.

– إغفال الجهود الجارية للسانيات الحاسوبية في تحليل بنية المعجم واستخدام تكنولوجيا المعلومات في مكينة[1] المعجم ودعم جهود العمل المصطلحي.

إن معظم هذه المشاكل لا يمكن حلها دون اللجوء إلى الحاسوب واستخدام أساليب نظم المعلومات والمعارف، وهي النتيجة المنطقية التي ستتضح لنا بشكل واضح عند مناقشتنا للجانب التقني للدراسة.

فالوسائل التكنولوجية– وأخص بالذكر الحاسوب في أحدث صوره– هي التي تمكننا من تدوين العدد الهائل من المادة اللغوية والإمكانات التي يوفرها في التعامل مع النصوص وفق مجالات أساسية:

1– قواعد المعطيات ← قاعدة بيانية معجمية[2].

2– قاعدة المعطيات النصية[3] ← قاعدة ذخيرة نصية.

3– قاعدة معطيات معجمية.

إن فكرة المعجم كقاعدة معطيات[4] تلجأ إليها الآلة هو أمر جد مهم لا يمكن الاستغناء عنه، غير أن هناك فرضيات ينبغي أن تطرح[5] ؟:

أ. مكونات المعجم الالكتروني، أو قواعد المعطيات[6] كبنية تحتية رئيسية، لبناء قاعدة معطيات نصية تكون بمثابة قاعدة ذخيرة.

ب. ترتيب هذه القواعد يكون على أساس نظريات ومواد لغوية، اعتمادًا على قاعدة معطيات معجمية. أما إذا فرضنا أن قاعدة الذخيرة النصية التي تشكل وفق قواعد معطيات تبنى على استنباط المثل[7]، وإجراؤها بعد صياغتها إلى خوارزمات[8]. فإننا مضطرون حينئذٍ إلى مراعاة هذه القواعد الأساسية من قاعدة تخزين ذخيرة نصية أو قاعدة بيانية نصية معجمية.

---

1 –Automation.

2 - Base de données lexicographique.

3 - Base de données textuelles.

4 – قاعدة المعطيات: مجموعات الاستعلامات أو كتل المعطيات التي تحدد، تعرف، وتتعلق بمسألة أو موضوع معين، هذه المعطيات يجري إدخالها

وتخزينها بتركيبة معينة ترابطية سهلة الانتقاء والبلوغ.

5 – عبد الرحمن الحاج صالح، بحوث ودراسات في اللسانيات العربية، سابق، 93/1.

6 – عبد الحسن الحسيني، معجم مصطلحات علوم الكومبيوتر، إنجليزي– فرنسي–عربي، دار القلم، بيروت، ط1، 1407هـ/1987هـ، ص143.

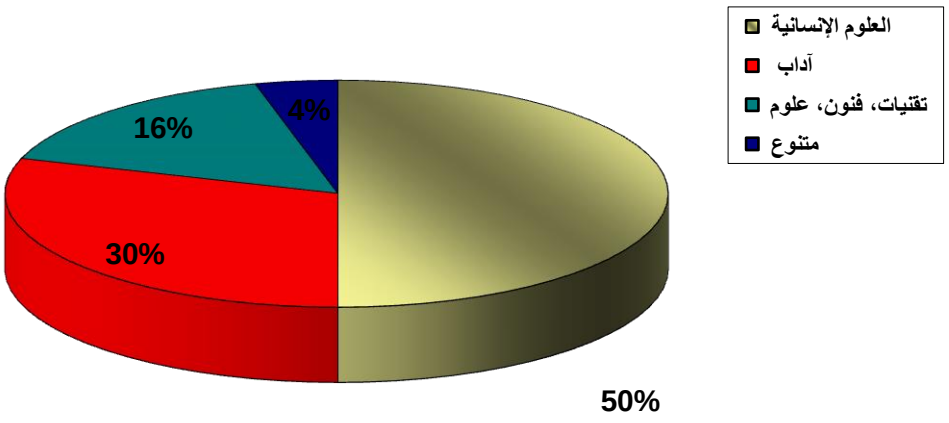
7 - Schèmes Générateurs.

8 – الخوارزمية: Algorithme مجموعة من القواعد والعمليات الحسابية والرياضية المتسلسلة، على شكل خطوات، يؤدي تنفيذها إلى الوصول إلى النتيجة المطلوبة، ينظر: عبد الحسن الحسيني، سابق، ص65.

والحديث عن الذخيرة معناه الحديث عن بنوك المعلومات الآلية، فهي كذلك ذخيرة للنصوص، تمثل الإنتاج الفكري العربي المعاصر، وتمثل الإنتاج الثقافي القديم إلى جانب الحديث، وللبنك المعلوماتي جانبان: أحدهما لغوي والآخر ثقافي، أما اللغوي:فلأنه يمثل الاستعمال الفعلي للغة العربية القديمة والحديثة من خلال المدونات العربية الموروثة عن السابقين، وما يقابل ذلك في عصرنا الحالي، وأما الثقافي: فهو يشمل التربوي والعلمي وطرق التعليم، وشتى مجالات البحث مناهجها (العلمية، التقنية، التاريخية، الاجتماعية، وغيرها)، ومن أجل ذلك أصبح الهدف صناعة موسوعات علمية في ظل التقنيات الحديثة[1].

العدد	الموضوعات الرئيسية
154	العلوم الإنسانية
90	آداب
48	تقنيات، فنون، علوم
11	متنـــوع

الجدول رقم 06: مجالات ذخيـــرة النصـــوص.



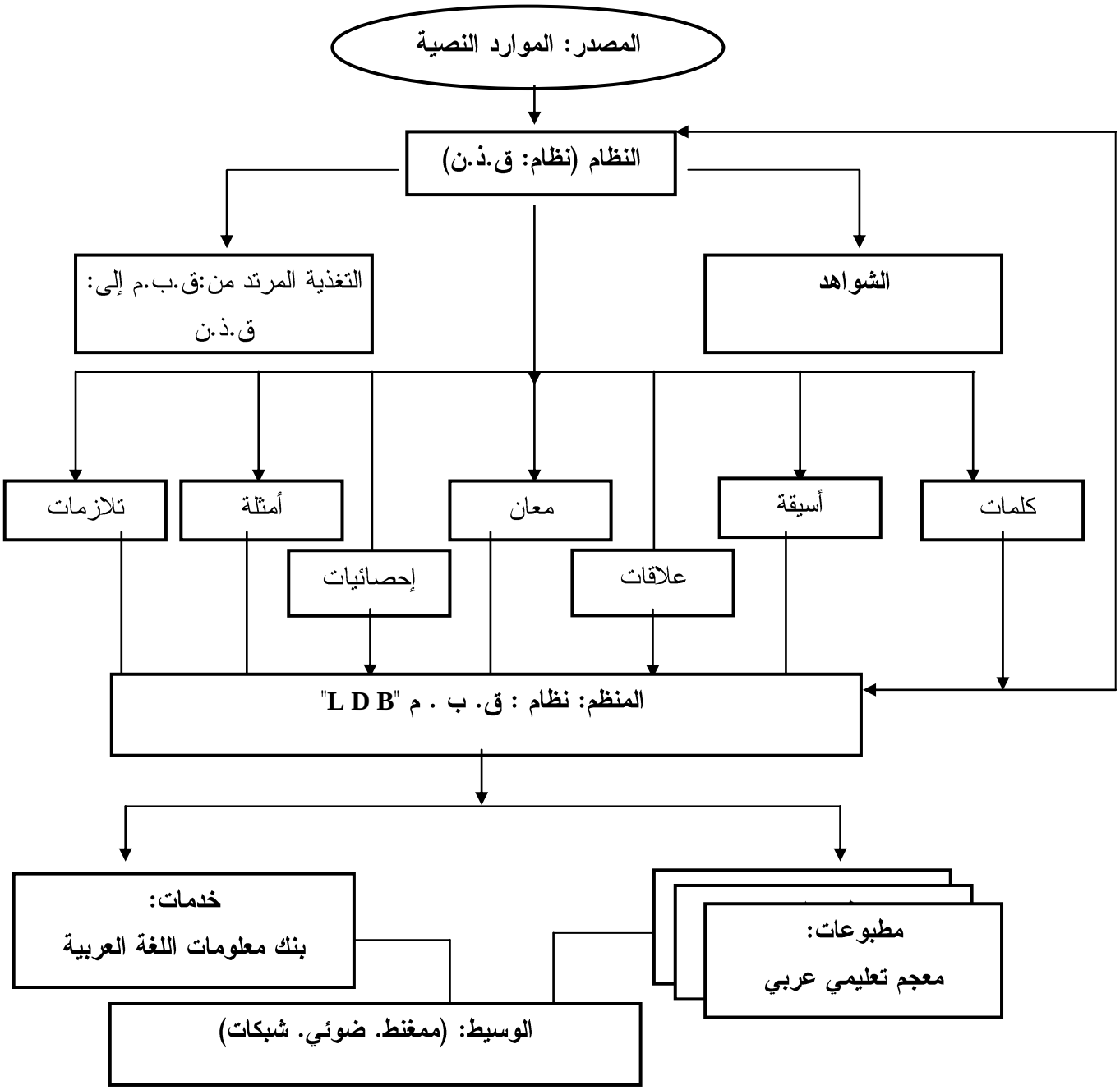
الشكل رقم 04: دائرة نسبية لذخيرة النصوص

1 – عبد الرحمن حاج صالح، بحوث ودراسات في اللسانيات العربية، موفم للنشر، الجزائر، طبعة 2006م، 2/ 153.

ويتضح بعد الاطلاع على الجدول والرسم البياني المرافق له (الدائرة النسبية) أن الحقول المعرفية أو مجالات البحث تتفاوت من حيث احتوائها للمواد المعرفية أو ما يصطلح عليه "ذخيرة النصوص"، وانطلاقًا منه فقد انفردت العلوم الإنسانية بحصة الأسد أي بما يعادل نصف الذخيرة (50%)، ذلك لأنها تشمل اللغات كافة، كما أنها مجال نسبي غير قار لا يقبل التجريب، يتغير بتغير الأحوال والأزمان، بينما الآداب جاءت في المرتبة التالية بثلاثين بالمائة (30%)، ثم العلوم التكنولوجية والعلوم الطبيعية بمقدار (16%)، وما بقي فهو متنوع بين مجالات أخرى.

وبناء على ذلك: فالمعجم على حد تعريف "أندرسون" هو «المحل الهندسي لكل ما لا يمكن رصده من خلال تطبيق القواعد والتنبؤ به على ضوء العلاقات المطردة والقرائن المنطقية التي تتضمنها التعبيرات اللغوية»[1]، ولإماطة اللثام عن التفاصيل الدقيقة لبنية المعجم الداخلية لذلك نضيف الشكل (5) لعرض الأنظمة القاعدية التي ترتكز عليها ذخيرة النصوص ومعطياتها في المعاجم الحاسوبية الإلكترونية وهو كما ياتي:

1 – نبيل علي، اللغة العربية والحاسوب، سابق، ص458.



<u>الشكل رقم 05:</u> مراحل تطوير المعجم العربي المقام على قاعدة ذخيرة النصوص (ق.ذ.ن)[1].

1– ق.ذ.ن: المقصود منها: قاعدة ذخيرة النصوص؛ أما نظام: ق.ب.م فمعناه: قاعدة البيانات المعجمية.

2-5- الأدوات الحاسوبية لتطوير المعجم العربي:

لقد ساعد التطور السريع الذي عرفته تكنولوجيا الحواسيب في الآونة الأخيرة عل معالجة المعلومات غير العددية بأقل تكلفة وبأسرع ما يمكن من الوقت، مما يتصل بالتخزين وإنشاء قواعد معطيات، نظريات لسانية انطلاقاً من قواعد لغوية مضبوطة وصارمة.

ومن أهم الأدوات الحاسوبية لدعم تطوير المعجم الإلكتروني نجد:

1) **القارئ العربي الآلي**[1] ← التعرف الضوئي على الحروف (OCR)[2]، يتميز بالدقة في تحويل الصورة إلى نص قابل للتعديل – مستندات مطبوعة ← تحويلها إلى نصوص إلكترونية قابلة للتعديل... كيف؟ القارئ الآلي.. الحل الوحيد اللغة العربية ذات طبيعة معقدة ← وتأتي بتحديات أساسية ← القارئ الآلي يواجه تلك التحديات بشكل فعال ← دقة عالية، سرعة متميزة، واجهة التطبيق ذات شكل نعين، وسهلة الاستخدام ← إنجاز العديد من المهام في آن واحد ← التكامل مع البرامج الأخرى[3].

2) **المحلل العربي**[4]، فهو عبارة عن نظام آلي لاستخلاص العناصر الأولية لبنية الكلمة وتحديد سماتها العرضية و"الصرف–نحوية"، والدلالية التي يمكن استنباطها من بنية الكلمة[5].

يقوم المحلل[6] بالتعرف على جميع أشكال جذر الكلمة، أي انه يقوم باستخلاص أصل الكلمة، بعد تجريدها من اللواصق[7]، فالمحلل الصرفي أساسي لأي نظام للمعالجة الآلية للغة الطبيعية فهو يسمح بتجميع الكلمات إلى طبقات أو فصول صالحة للاستعمال من طرف مستويات التحليل الصرفي التي تختلف باختلاف ميزة المعالجة[8].

3) **المدقق الإملائي**[9]، يقوم المدقق بتصحيح الأخطاء الإملائية التي تحدث أثناء كتابة النصوص العربية أو عن طريق إعطاء بدائل متعددة للكلمة الخطأ، ليقوم المستخدم باختيار إحداها، ولا يقتصر دور المدقق على تصحيح الأخطاء الإملائية بل يتعدى ذلك إلى تصحيح الخطاء النحوية والإعرابية في النصوص العربية؛ ومن مميزات المدقق الإملائي[10]:

1- Lecteur automatique.

2 – وهذا الرمز استحدثه الدارسون في البحث اللساني وأصله: Reconnaissance Optique De Caractères

لبسط الفهم في هذا أنظر: Anderson, « OCR Enters the Practical Stage »Dataation,Vol 17, 1971,p22-27 PL.

3 – موقع صخر: منتجات القارئ الآلي، موقع الويب: http://www.sakhr.com/product_a/OCR/.

4- Analyser Morphologique.

5 – نبيل علي، اللغة العربية والحاسوب، سابق، ص306.

6 – عبد الحسن الحسين، معجم مصطلحات علوم الكومبيوتر، إنجليزي–فرنسي– عربي، سابق، ص68.

7 – تقنيات التحليل الصرفي، موقع الويب: http://www.sakhr.com

8- Mohamed El Amine ABDERRAHIM, Reconnaissance des unités linguistiques signifiantes, Thèse pour obtenir le titre de doctorat, spécialité informatique encadré par Mr Fethi BEREKSI REGUIG et Mme Laurence BALICCO, Université de Tlemcen, 2007-2008, p.6.

9- Correcteur Orthographique.

10 – تقنيات المدقق الإملائي، موقع الويب: http://www.sakhr.com .

1- التصحيح التلقائي.

2- تقليل عدد الاقتراحات.

3- مراعاة السياق.

4- التعامل مع المصطلحات.

5- التعامل مع التشكيل.

وتجدر الإشارة إلى أن هناك أدوات حاسوبية يمكن توظيفها في تطوير المعاجم من ورقية إلى إلكترونية، وتتعلق هذه الأدوات ببناء قاعدة المعطيات الموجهة لصناعة معجم آلي باستخدام تقنيات معلوماتية حديثة لمعالجة اللغة العربية آليًا، في المستوى المعجمي مقروناً بالمستوى الصرفي[1].

إلى جانب ذلك نشير إلى أن هناك أدوات حاسوبية تساهم في تطوير المعجم الآلي المنشود، هي: قاعدة ذخيرة النصوص- نظام التشكيل التلقائي- نظام الإعراب - قاعدة بيانات المعجم.

وبيان ما تقدم نورده في الجدول الآتي:

OUTIL	الأداة	الغـــــرض
A- O C R	القارئ العربي الآلي	اقتناء النصوص آليا
MMMP	المحلل الصرفي الآلي	تحليل النصوص صرفيا – فك اللبس الصرفي
A- SPELLING CHECKER	المدقق الإملائي العربي	تصويب أخطاء النصوص المغذاة
CAC	قاعدة ذخيرة النصوص	إمداد الشواهد المعجمية نواة قاعدة ذخيرة النصوص المعجمية
MMSP	نظام الإعراب الآلي	فك اللبس الصرفي – والدلالي تحليل مادة تعريف المعجم لاستخراج نواة التعريف
ADG	نظام التشكيل التلقائي	تشكيل النصوص آليا
AFG	نحو العربية	توصيف البيانات المعجمية النحوية
AFTDB	قاعدة النصوص الكاملة العربية	البحث العميق والتنوع داخل قاعدة النصـــوص وداخـــل المعجم
LDB	قاعدة البيانات المعجمية	إدارة بيانات بناء المعجم – ضمان الاكتمال والاتساق
LEXICAL SEMANTICS	بحوث الدلالة المعجمية	السمات اللغوية- القيود الانتقائية- علاقات الأسماء

الجدول رقم 07: الأدوات الحاسوبية لدعم تطوير المعجم العربي.

1 – وليد العناتي، وخالد الجبر، دليل الباحث إلى اللسانيات الحاسوبية، سابق، ص109.

5-3- ماذا سيقدم المعجم الإلكتروني؟:

إن ما سميناه بمشروع المعجم الإلكتروني يرمي إلى ضغط بنك آلي (حاسوبي)، من قواعد معطيات نصية يتميز عن المعاجم الورقية المتوفرة[1] :

1-الشمولية الكاملة في المكان والزمان.

2-اندماج المعطيات النصية، كأنه نص واحد حاسوبيًا ومرتبة ومصنفة بحسب العصور والفنون المعرفية.

3-كونه آليًا وعالميًا إذ سيكون له موقع في شبكة الأنترنت.

4-استرجاع الكلمات بسهولة من قاعدة البيانات المعروضة[2].

وتقدم المعاجم الإلكترونية لمستخدم المعجم جملة من المميزات، أهمها[3]:

1-التدقيق الآلي[4].

2-الفهرسة الآلية[5].

3-الترجمة الفورية[6].

4-التلخيص الآلي[7].

---

1 – عبد الرحمن الحاج صالح، مقال: المعجم العربي والاستعمال الحقيقي للغة العربية، مجلة المجمع الجزائري للغة العربية، عدد1، السنة الأولى، ربيع1/1426هـ، ماي 2005م، ص19-20.

2 – أحمد عمر مختار، صناعة المعجم الحديث، سابق، ص183.

3- من بين التطبيقات التي يقدمها المعجم الإلكتروني للمستخدم:

- الفهم الآلي للغة (Automatic language understanding).
- التحليل والتوليد الصرفيان (Morphological analysis and generation)
- التحليل والتوليد التركيبيان (Syntactic analysis and generation)
- الترجمة الآلية (Automatic transtalation)
- التكشيف الآلي وأنظمة استخراج / استرجاع معلومات (Automatic, indexing and Information retrieval systems)
- المشكل الآلي (Automatic Vowelization).
- ضغط النصوص (Text Compression).
- التعرف الآلي على الخصائص اللغوية (Characters recognition)
- توليف الكلام وتوليده (Speech Synthesis and generation)

للمزيد من التفصيل ينظر: - Mahmoud Mostafa Mohamed, The developement of A Computer- based Arabic Lexicon, M.Sc. Theisis, Helwan University, Jan, 2004, p: 3

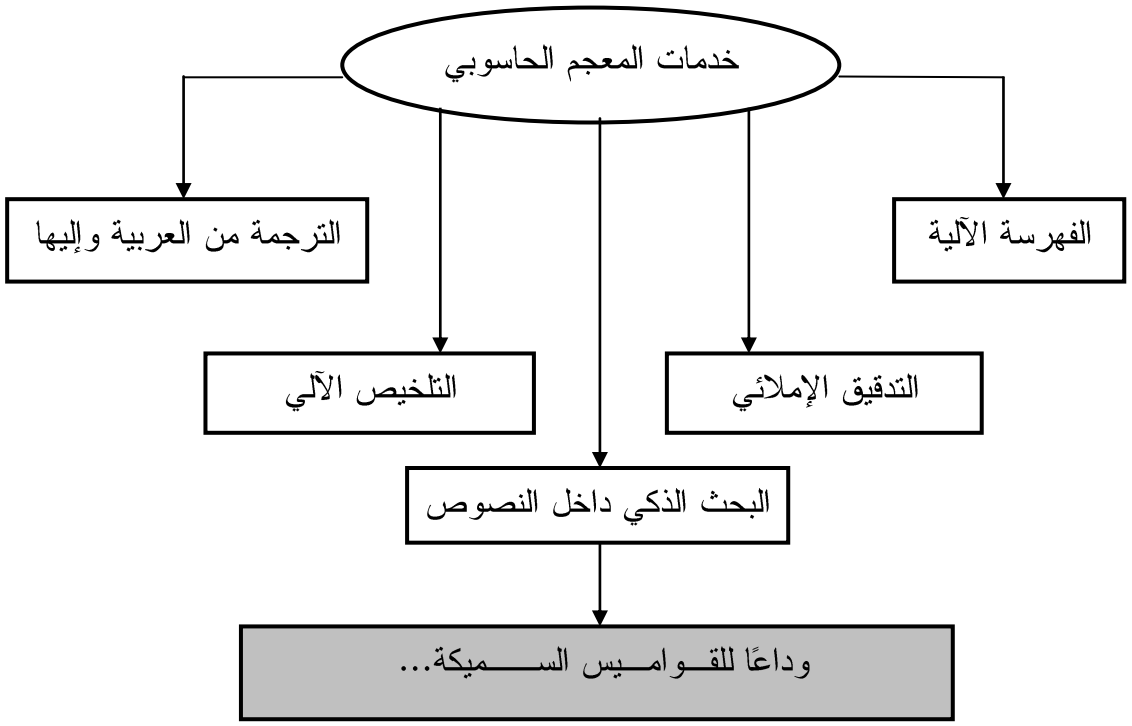
4 - Contrôles Automatisés.

5 - Indexation Automatique.

6 -Interprétation.

7 -Résumé automatique.

ذلك ما هو مبين في الشكل (1) الآتي:



الشكل رقم 06: المجالات التطبيقية للمعجم الحاسوبي (الإلكتروني).

سادسا- الفرق بين المعجم الورقي[1] والمعجم الإلكتروني:

يتجلى الفرق بين المعجم الإلكتروني والمعجم العادي، في أن الأول مبني وموجه أساسا لاستخدام البرامج الحاسوبية التي لا تربطها أية صلة بالمعرفة اللسانية. بينما المعاجم الورقية العادية مثل القواميس المتخصصة كقاموس التجارة الذي يبنى أساسا لخدمة أغراض تجارية، حيث يرمي إلى توسيع مدارك الأشخاص في ميدان المعرفة التجارية والاقتصادية والمبادلات. في حين أن القواميس اليدوية أو العادية يمكن أن تختزن على شكل مذكرات منسوخة أو في أقراص مدمجة، وتقبل الخضوع للبرمجة الحاسوبية أو الحوسبة، وهذا لا يمنع من أن توظف برامج التحليل اللساني، يسوغ ذلك الاعتبارات التالية[2]:

1- محمد الحناش، نظرية لسانية حاسوبية لبناء المعاجم الآلية، السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات 1992 منشورات مكتبة الملك عبد العزيز العامة بالرياض 1993، ص .363-394، وعلى القاسمي، "المعاجم العربية المتخصصة ومساهمتها في نقل التكنولوجيا"، في كتاب "المعجمية العربية بين النظرية والتطبيق 193-203.

2- عمر المهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، إشراف الدكتور عبد الغني أبو العزم، جامعة الحسن الثاني، الدار البيضاء، 2008/2007، ص ،91،90.

1/ المعجم الإلكتروني ويشترط فيه أن يكون شاملا وعاما، لأن البرنامج اللساني المعدّ للمعالجة الآلية لا ينبغي أن يفشل في العثور على أية معلومة كيف ما كان نوعها وكيفما اتفق، لأن أي خطأ في المعلومات المدخلة من شأنه أن يتسرب إلى باقي مفردات الجملة أو كل النص برمته، ومن ثمّ يعرقل عملية اشتغال البرنامج.

2/ يشترط في المعلومات أن تكون في المعجم الإلكتروني واضحة بينما في المعجم العادي غالبا ما يؤشر على المعلومات الأكثر شهرة أو تداولا بكلمة معروف أو بديهي، في حين لا شيء بديهي في الحاسوب، فكل المعلومات أساسية وضرورية، وينبغي أن تكون واضحة ومضبوطة.

3/ ترتبط المعاجم الإلكترونية ببرامج المعالجة الآلية، ذلك أن المعلومات المتضمنة فيها ينبغي أن تتلاءم والمداخل المعجمية.

ولا شك أن ثمة افتراقات متعددة بين المعجم الحاسوبي والمعجم الورقيّ، وهي افتراقات تفرضها طبيعة الوسيلة التي بني الحاسوب وفقا لها؛ وأهداف كلّ معجم منهما. ويمكن أن نحدد هذه الافتراقات في مجالات ثلاثة، وهي[1]:

- اختلافات في البنية.
- اختلافات في المادة.
- اختلافات في الاستخدام.

1- وليد أحمد العناتي، نحو معجم حاسوبي أحادي للناطقين بغير العربية،الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، مطبعة المعارف الجديدة، الرباط، 2009، ص 227.

أولاً: اختلافات في البنية:

ومرجع هذه الاختلافات إلى المعجم الورقي بني ليستعمل يدويا، ولذلك فإن غايته الرئيسية ألا يثقل على المتعلم بكثرة حجم الصفحات، وتشعب المعلومات وتنافرها، كمايحرص على أن يوفر للمتعلم المواد المطلوبة بأقصر الطرق وأيسرها لذلك تجده يحاول تقريب المفاهيم بجميع السبل اللغوية والتوضيحية المتيسرة، كالرسوم التوضيحية[1] والمتضادات والمترادفات مع تقانات الوسائط المتعددة[2].

أما المعجم الحاسوبي يتم بناؤه حاسوبيا ؛ ذلك أنه يعتمد على ما تتيحه تقنيات الحوسبة من وسائل عرض المعلومات اللازمة وتقديمها، كالصوت، والتحليل الصرفي، والتحليل النحوي، والخطوط الجملية، والألوان الجذابة، والصور الحقيقة غير المرسومة ...إلخ. وتعد الوسائط المتعددة اكتشافا ممتازا ييسر تقديم خدمات متعددة في الآن، إذ يستطيع المتعلم أن يشاهد، ويسمع ويكتب في وقت واحد[3].

ثانياً: اختلافات في المادة:

وبيان هذه الاختلافات ماثل في أن مادة المعجم الورقي محدودة[4]، مهما كبر حجم المعجم. وتتمثل هذه المحدودية في انتقاء عدد محدد من المداخل ومشتقاتها وما يرتبط بها من مواد أخرى، وغالبا ما يكون الحجم عاملا مهماً في تصميم المعجم الورقيّ وبنائه؛ لذلك فإنه يعتمد على المختصرات والإشارات المرجعية بكثر، اختصارا للمساحة الورقية. وبناء على هذا فإننا نتوقع أن يخلو المعجم الورقيّ من خصيصة الاتعاب النحوي والصرفي والاسلوبي؛ ذلك أنه لا يمكن أن تجد، مثلا في المعجم الورقي جميع المشتقات الصرفية، والأهاريب النحوية، والاستخدامات الاصطلاحية، والتنوعات الأسلوبية، لمدخل معين؛ فإذا نظرنا مثلا في المعجم

1- إن استخدام الرسوم التوضيحية في المعاجم العربية قديم نسبيا، إذ يعود إلى اوائل القرن الخامس عشر الميلادي، إلا أن الأخذ بهذه الوسيلة ظل محدود فقد ظهرت بعض الصور والرسوم التوضيحية في المعاجم العربية الحديثة في إطار ضيق جدا، كما استعانت بها المعاجم العربية المعاصرة، ولما تطور من تقنية اثباتها ضمن منهجية واضحة.للمزيد من بسط القول ينظر :

*/ حلام الجيلالي، تقنيات التعريف في المعاجم العربية المعاصرة، دراسة، اتحاد الكتاب العرب، دط،1999 ص 244.

*/ خليل ابراهيم الحماش، الرسوم التوضيحية ومكانتها في المعجم، اللسان العربي، مكتب تنسيق التعريب، العدد 22، الرباط، ص 129.

2- وليد أحمد العناتي، نحو معجم حاسوبي أحادي للناطقين بغير العربية،الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، مطبعة المعارف الجديدة، الرباط، 2009، ص 227.

3- المرجع نفسه، ص 228.

4- حلام الجيلالي، تقنيات التعريف في المعاجم العربية المعاصرة، دراسة، منشورات اتحاد الكتاب العرب، دمشق، دط ، 1999، ص81 وما بعدها.

العربي الأساسي فهل نجد،مثلا، في مادة (ضرب): اسم فاعل، واسم مفعول، وصيغة مبالغة ومعانيها الاصطلاحية...إلخ لا يمكن أن يحدث هذا.

أما في المعجم الحاسوبي الذي يمثله مشروعنا المقترح فإنه يكون مزودا ببرمجيات مساندة[1] تتألف من محلل صرفي، ومحلل نحوي، وينبغي أن أشير هنا إلى هذه المعلومات المطلوبة ليست مدخلة إدخالاً يدويا، وانما ينفذها الحاسوب ببرامج مُعِدّة سلفا.

ثم إنّ واضعي المعجم الورقي يحاولون، قدْر الوُسعِ، أن يجعلوه معجما شاملا نسبيا؛ فهو يتضمن دلالات المنفردة المشهورة، وأسماء الأعلام، ومعلومات ثقافية وحضارية، إضافة إلى معلومات لغوية تتناول تعريفا باللغة العربية ومستوياتها المتعددة[2].

أما المعجم الحاسوبي المفترض فإنه يمكن أن يتضمن عددا كبيرا من المعاجم التي ترتبط بطريقة فنية تمكن المتعلم من تحصيل كل ما له علاقة بالكلمة المرادة، فإذا أراد هذا المتعلم أن ينظر في معنى كلمة (صلح) فإنه يستطيع تحصيل معلومات من مثل[3]:

صلح: يصلح صلاحا (مصدر)

صلاح الدين الأيوبي (سيرة موجزة) من معجم أعلام الحضارة الإسلامية.

صلاح الدين/ صلاح: اسم علم مذكرة (ذكر). ويستعمل مضافا وقد يختص إلى (صلاح) من معجم أسماء العرب.

وكذا القول في كل ما يتعلق هذا المدخل.

ولا شك أن المعجم الورقي لن يكون متضمنا لكل هذه المعلومات.

ومن هذه الاختلافات أن المعجم الورقي، مثلا، كتاب مصمت، إذا لا يمكن تضمينه نطقا بشريا سليما، وإن كان يستعين على ذلك بضبط الكلمات كلها، ولكن هذا الضبط لن يدل على مواضع النبر أو التنغيم، وبيان الافترقات الأسلوبية، أما في المعجم الحاسوبي فإن تضمينه مادة مسجلة يعد من أبسط أمور التقنية إذ يمكن تسجيل مادة المداخل كاملة بصوت بشري خالص، فهو معجم ناطق بامتياز، ما يضاعف قيمته في تعليم نظام العربية الصوتي لغير الناطقين بالعربية، إذ يوفر مادة منطوقة باستمرار.

1- للمزيد من التفصيل ينظر : وليد أحمد العناتي، خالد الجبر، دليل الباحث إلى اللسانيات الحاسوبية العربية، ط1، دار جرير، عمان، 2007.

2- وليد أحمد العناتي، نحو معجم حاسوبي أحادي للناطقين بغير العربية،مرجع سابق، ص 228.

3- المرجع نفسه، ص 229.

ومن أهم هذه الاختلافات أنه يصعب جدا أن تخصص جزءا من المعجم الورقي للذخيرة النصية أو المتون، وهي نصوص مساندة تعد تمثيلا لواقع استخدام العربية في شؤون حياتنا المختلفة: كالصحافة والأعلام، والأدب، والثقافة، والتجارة والسياحة...إلخ، أما المعجم الحاسوبي فإنه من السهل تضمينه هذه المتون، ومن السهل استدعاؤها مكتوبة ومنطوقة ومحللة تحليلا صرفيا ونحويا وتداولياً، كل ذلك في سياق واحد جامع[1].

ثالثاً: في الاستخدام:

يعتمد استعمال المعجم الورقي اعتمادا أساسيا على فهم المتعلم لمنهج المعجم وطريقته في ترتيب المواد، ولعل هذا يصعب الاستخدام حين يكون المتعلم غير متثبت من المنهج (الترتيب)، كأن لا يحفظ ترتيب الحروف العربية، أو أن يكون غير قادر على تجريد الكلمة من الزوائد والوصول إلى جذرها[2].

ولا شك أن استعمال المعجم الحاسوبي أسهل من الورقي ذلك أنه يعتمد على معرفة مجمعة بسيطة من المفردات للتعامل معه والدخول في أعماقه، وهي معلومات تتعلق بمفاتيح الاستعمال ينبغي أن يكون المتعلم حصلها في بداية تعلمه العربية،أما تعليمات استعمال المعجم الورقي فقد تطول إلى حد ممل وصعب يجعل المتعلم ينفر منه من الأول[3].

1- المرجع السابق، ص 230.

2- المرجع نفسه، الموضع نفسه.

3- المرجع نفسه، الصفحة نفسها.

خلاصة

من المؤكّد أنّ النّظريّة القديمة – على وضعها الحاليّ- لا يمكن أن تنبنيَ عليها حوسبة متكاملة للعربيّة، لكنّ من الواضح أيضاً أنّ النّظريّة اللسانيّة الحديثة لم تضع نموذجاً متكاملاً لعلوم العربيّة يصلح قاعدة لحوسبة متكاملة. وإنّ هذه المشكلة تمثّل –في تصورنا- ذروة التّناقض عندما يراد حوسبة شيءٍ غير مؤطَّرٍ نظرياً.

فالمعجم الحاسوبي يخضع لما يخضع له المعجم العام، من أركان التأليف (الصناعة المعجمية)، إذ ينبغي أن ينطلق في تأليفه من تحديد المقتضيات العلمية والمنهجية الخاصة بثلاث مسائل: الأولى هي الوظيفة، فلا بد من تحديد ملامح المستعمل الذي يؤلف له المعجم، والثانية هي المادة المجموعة، ولا بد أن يراعى فيه التوفيق بين قواعد البيانات والمستويات اللغوية، والثالثة هي منهج الوضع، فيكون الترتيب بحسب النموذج المقترح[1].

انطلاقًا من هذا الباب التمهيدي أردت أن أضع تصورًا منهجيًا لحوسبة المعجم العربي من منظور لساني حاسوبي، وحاولت أن ألامس في هذه الأطروحة في شقها التمهيدي جل القضايا التي تخص مراحل الحوسبة من منطلق المعرفة اللسانية نحو انشاء نظرية لسانية في ضوء اللسانيات الحاسوبية كميدان أساسي لمثل هذه التطبيقات، مما دفعني لتوظيف الجهود لعقد الباب الثاني لمعالجة موضوع المعالجة الآلية للغة العربية[2] على وجه التخصيص، كميدان خصب لمثل هذه التطبيقات، فقد دانت الأساليب والدوافع لاستثمار تقنيات الحاسوب في مجال اللغويات عامة و المعجم خاصة.

1- لقد جاء في تراثنا اللغوي الكثير من الإشراقات اللغوية العلمية التي طرحتها النظريات اللسانية الحديثة. لكن هذا برأينا لا يغني عن استخدام مناهج البحث اللغوي الحديثة في وصف اللغة وفي البحث اللغوي. ولاشك أن التطور من قضايا العلوم الطبيعية، فاللغة الفصحى وإن كانت ثابتة وخالدة بحكم ارتباطها بالقرآن، إلا أن وصفها ووسائل تعليمها قابلة للتغير. واللغويون القدماء لم يضعوا قواعد اللغة، بل استخلصوها من الواقع اللغوي ووصفوها بالوسائل التي كانت متوفرة لهم في ذلك الوقت. والآن وقد ظهرت أدوات جديدة للوصف اللغوي فلا نرى بدّاً من الأخذ بها حتى تستمر اللغة في الوجود مواكبة للتطور العلمي.

2- سأفصل الكلام حول هذا الحقل في الجانب المخصص له من الأطروحة، فهو يعتبر من أهم مجالات البحث اللغوي في إعداد مثل هذه المشاريع.

الباب الثاني

المعالجة الآلية للغة العربية لبناء قاعدة البيانات المعجمية آليا

الفصل الأول

المعالجة الآلية للغة العربية

* النظام الصرفي العربي – أنموذجا *

أولاً- الإطار العام لمعالجة اللغة العربية آليًا

ثانيا- تحديد الإطار العام لمعالجة المعلومات و المعارف

ثالثاً- المعالجة الآلية للصرف في اللغة العربية

رابعاً- الإطار العام للمعالجة الآلية للصرف العربي

خامساً- تطبيقات المعالجة الآلية للغة العربية

سادساً- المعالجة الآلية للدلالة في اللغة العربية

أولا: الإطار العام لمعالجة اللغة العربية آليا:

1- اللغة والحاسوب:

إن العلاقة بين اللّغة والحاسوب هي بلا شك أهم القضايا في تهيئة هذه اللّغة للدخول في عصر المعلومات، إنّ تمحيصنا الدقيق لخصائص هذه العلاقة هو –إلى جانب كونه مدخلاً مهمًّا لمناقشة الجوانب المختلفة للقضية– مصدر إلهام أساسي لمنطلقات الحلول لكثير من المشكلات التي تشملها هذه الدراسة[1].

تبدأ قصة الاتصال العلمي بين الحاسوب والبحث اللغوي العربي في سبعينيات القرن الماضي حينما ذهب الدكتور ابراهيم أنيس إلى جامعة الكويت سنة 1971م للعمل بها أُستاذاً زائراً، وهناك التقى بالدكتور علي حلمي موسى، أستاذ الفيزياء النظرية في جامعة الكويت، وطرح عليه فكرة الاستعانة بالحاسوب في إحصاء الحروف الأصلية لمواد اللغة العربية[2]، بُغْية الوقوف على نسج الكلمة العربية. وقد رحب بهذه الفكرة واستحسنها، وكان من ثمرة ذلك صدور الدراسة الإحصائية للجذور الثلاثية وغير الثلاثية لمعجم الصحاح، للجوهري (324هـ)، فكانت بحق ابتكاراً جديداً لم يسبق إليه من قبل، بل هي المرة الأولى في العالم العربي التي تجري فيها هذه الإحصائيات على أسس علمية حديثة ودقيقة.

هذا التوجُّه في الفكر العربي المعاصر قد فتح الباب واسعاً للباحثين في الدراسات اللغوية العربية الولوج إلى عالم الحاسوب، وتسخيره لخدمة البحث اللغوي العربي.

لقد تهيأت اللّغة للمعالجة الآلية بدخولها مجالات التحليل الرياضي والمنطقي والإحصائي، وتهيأ الحاسوب من جانبه للقائه مع اللّغة بالسرية الفائقة وضخامة الذاكرة وضآلة الحجم وأســـاليـــب الذكاء الاصطناعي[3] ولغات البرمجة الرّاقية، أمّا التطبيق فقد شقَّ هو الآخر

1– نبيل علي، اللّغة العربية والحاسوب، سابق، ص 168.

2– سليمان فياض، أنظمة تصريف الأفعال العربية (الأفعال الثلاثية)، دار المريخ للنشر، الرياض،الطبعة1، 1989م، ص14.

3–Daniel Kayser, Bernard Levrat ,Traitement automatique du langage naturel, Ed, HERMES science Publication, Paris Volume 20- n°=03 ;2001- P .02

* الذكاء الاصطناعي: l'intelligence artificielle يهدف هذا العلم إلى فهم طبيعة الذكاء الإنساني عن طريق عمل برامج للحاسوب الآلي قادرة على محاكاة السلوك الإنساني المتسم بالذكاء، وتعني قدرة برنامج الحاسوب على حل مسألة ما، أو اتخاذ قرار في موقف ما– بناء على وصف لهذا الموقف– أن البرنامج نفسه يجد الطريقة التي يجب أن تتبع لحل المسألة أو التوصل إلى القرار بالرجوع إلى العديد من العمليات الاستدلالية المتنوعة التي غذي بها البرنامج. ويعتبر هذا نقطة تحول هامة تتعدى ما هو معروف باسم "تقنية المعلومات" التي تتسم فيها العملية الاستدلالية عن طريق الإنسان وتنحصر أهم أسباب استخدام الحاسوب في سرعته الفائقة.

فالذكاء الاصطناعي يدخل بوصفه أساسًا في مكونات جميع نظم التحكم الآلي والنظم الكبيرة. فهو يعنى بدراسة الملكات العقلية للإنسان باستخدام النماذج الحاسوبية لإكساب الحاسوب بعضًا منصا. للمزيد من التفصيل ينظر:

– آلان بونيه، الذكاء الاصطناعي واقعه ومستقبله، ترجمة علي صبري فرغلي، عالم المعرفة، الكويت العدد 172– أبريل 1993.

– عبد الله بن حمد الحميدان، مقدمة في الترجمة الآلية، مكتبة العبيكان– الرياض الطبعة الأولى 1421– 2001، ص 16.

– ادوارد فايجنبوم وباميلاماك كوردك ،الجيل الخامس: الذكاء الاصطناعي والتحدي الياباني للعالم، د ط،1987، ص 286.

طريقه النجاح إلى مجالات متعددة[1]: التعليم والعلوم الإنسانية والنظم الكبيرة[2].

هذا عن لقاء المعالجة الآلية مع اللّغة بصفة عامة؛ فماذا عن طبيعة هذا اللّقاء مع اللّغة العربية؟.

إنّ هذا الإشكال يؤدي بنا إلى صياغة عدّة أسئلة مهمة تُعنى بهذا الفصل كونه العمود الفقري في هذه الدراسة– في الجانب الخاص، بالمعالجة الآلية للغة العربية– الدراسة التطبيقية).

فالحاجة تتزايد يومًا بعد يوم للمعالجة الآلية للغة العربية يتزايد معها حجم الدراسات والأبحاث والمقاربات التي أنجزت وتنجز في هذا الإطار[3].

فسياق الحديث يؤدي بنا إلى طرح عدّة تساؤلات تخص الجوانب النظرية والتقنية لهذه الدراسة؟ كالشروط التي ينبغي أن يتوفر عليها البرنامج اللّغوي لنجاح هذه المعالجة؟، وهل هناك خصائص تعتمد عليها اللغة العربية حتى تجعلها لغة طبيعية للحاسوب؟.

تستند التطبيقات في المعالجة الآلية للغة مثل استخراج المعلومة أو استخراج جذور الكلمات العربية على التطبيقات اللّغوية وفق قواعد معارف صرفية ونحوية ومعجمية إذ تعتمد هذه التطبيقات على الخبرات اللسانية لتحديد تصور وأسس المعالجة الآلية.

فلهذا تتوخى مشاريع المعالجة الآلية للصرف العربي توفير كم هائل من الموارد اللسانية تشمل الجذور، والصّيغ الصرفية، والصور البسيطة أو المصرفة والمركبة "بالسوابق واللواحق"[4].

في هذا الإطار يعرض هذا الجانب من هذه الدراسة مرحلة مهمة تُعنى بالتحليل الصرفي في المعالجة الآلية للغة العربية فهذه المرحلة أساسية ومهمة لأنها أساس جميع تطبيقات المعالجة الآلية للغات الطبيعية[5]. ولهذا خصصنا في هذه الدراسة مبحثًا نتحدث فيه عن تطبيقات المعالجة الآلية للغة العربية في إطار المعالجة الآلية لمنظومة الصرف العربي.

1– نبيل علي، العرب وعصر المعلومات، عالم المعرفة الكويت، العدد 184– أبريل 1994، ص 346.

2 –Expert systeme.

3– فريق المعالجة الآلية للغة العربية– (التحليل الصرفي للأسماء العربية) تحت إشراف الأستاذ عبد الفتاح حمداني (المعالجة الآلية للغة العربية) ندوة دولية CITAL '07 18– 19– يونيو 2007.

معهد الدراسات والأبحاث للتعريب– جامعة محمد الخامس السويسي– الرباط– يونيو 2007، ص 51.

4– عبد الفتاح حمداني، خالد الأشهب، محمد الراضي (المعالجة الآلية للغة العربية) وقائع الندوة الدولية يونيو 2006– معهد الدراسات والأبحاث للتعريب، جامعة محمد الخامس– السويسي– الرباط، يناير 2007– ص 65.

5 –Bessou sadik, Analyse de données Textuelles pour la classification Automatique par les techniques de Text Mining à la Langue Arabe, Mémoire de Magister, option: Sciences et technologies de l'information et de la Communication, Université FERHAT ABBAS- SETIF (UFAS) 2007. P. 63.

وتجدر الإشارة إلى أن معظم النظم المتوفرة حاليًا لمعالجة اللغة آليًا قد أقيمت على أساس نظريات ومفاهيم سادها النّحو[1] وأغفل فيها الصرف بدرجة كبيرة فمن المؤكد أن هذه النظم تحتاج إلى تغيرات جذرية لتطويعها لمطالب المعالجة الآلية للغة العربية والذي للصرف فيها دور حاسم ومحوري[2]. وقد أوردنا هذه الدراسة في مجال المعالجة الآلية للغة العربية، إيماناً منا بأنّها تُعد أساسيات هذا البحث وكذلك أحد النقاط القوة في بناء المعجم الإلكتروني، بل إن بعضها يدخل في صميم عمل المعجم،إذن في ضوء هذه التحديات،

- هل يمكن استخدام اللغة العربية لتطوير برامج حاسوبية؟
- **فكيف يمكن توظيف منظومة الصرف العربي في معالجة اللّغة العربية آليًا ؟**
- **وكيف يمكن استثمار هذه المنظومة في بناء معجم الكتروني للمعالجة الآلية للغة العربية ؟.**

2- خصائص اللغة العربية من منظور لساني :

تتسم المنظومة اللّغوية بعدّة خصائص مميزة تحدد طبيعتها وأداءها، وتفرض قيودًا على أساليب تناولها، وأهم من هذا كلّه أنّها– أي الخصائص[3]– هي المصدر الرئيسي لمنطلقات حلول المسألة اللّغوية على اختلاف صورها ونطاقاتها.

1– محمد عبد العزيز عبد الدّايم،النظرية اللّغوية في التراث العربي، دار السّلام، القاهرة، الطبعة الأولى 2006،1427، ص 51.

2– نبيل علي،اللّغة العربية والحاسوب، سابق، ص 251.

3– الدارس للغة العربية لا يستطيع أن ينكر العديد من المزايا و الخصائص التي تميزها عن أي لغة أخرى. بدايةً فقوة أي لغة تكمن في قدرتها على التعايش مع العصر الذي تعيشه، وعلى التأقلم مع مفرداته، واستيعاب مستجداته، كل ذلك في أسلوب علمي يكفل الدقة في توصيل المفاهيم بأقصى كفاءة، واللغة أي لغة ليست مجرد أسلوب كتابة، بل هي نسق كامل من المفاهيم والمشارب التي تعبر عن نفسها حديثاً وكتابةً وقراءةً بل وتفكيراً، لنقل أنها وعاء كامل لتراث الأمة.

وللغة العربية بناء رياضي فريد لا يتوافر لأي لغة أخرى، فأسلوب الاشتقاق الفريد من الأفعال والأسماء بنسق رياضي دقيق يتيح لها استيعاب أي مصطلح جديد والتعبير عنه بطريقة تلقائية مبسطة تقترب من العامة قبل الخاصة، بل وتطويع مثل هذا المصطلح إلى موسيقاها الخاصة بما يتفق مع الأذن العربية السليمة، ويضاف إلى ذلك قوانين رياضية واضحة لإضافة السوابق واللواحق (Préfixes et suffixes) لأي كلمة بما يضيف إلى معناها ولا ينقص منه.

ولو تأملنا فكرة تغيير الموقع الإعرابي للكلمة في وسط الجملة تبعاً للمعنى المقصود، وما يعنيه ذلك من تغيير في تشكيل آخرها أو في حروفها، وتأثير ذلك على وضوح معناها، لتيقنا كم هي قوية هذه اللغة، وكم تستحق من مجهود وعمل.

ويمكن أن نجمل أهم الخصائص في بعض النقاط التي تميز اللغة العربية عن غيرها من اللغات...

أولاً: كإحصائيات مقارنة سريعة بين اللغتين العربية والإنجليزية...

فان إجمالي عدد الكلمات باللغة الإنجليزية يتراوح بين أربعمئة إلى خمسمائة ألف كلمة مستخدمة.

بينما تتعدى الكلمات المستخدمة في اللغة العربية عدة ملايين(أخذاً في الاعتبار الكلمات المشتقة والمزيدة).

ثانياً: في الوقت نفسه فإن جذور الكلمات العربية أقل بكثير من مثيلاتها الإنجليزية، ولكنها تخضع لجداول حاكمة يمكن من خلالها الاشتقاق المباشر لهذا العدد الكبير جداً.

ثالثاً: يمكن حصر حوالي مئة وأربعة وأربعين جدولاً رياضياً سليماً تتعامل تقريباً مع جميع الأفعال العربية من دون استثناء.

رابعاً: تمثل تركيبات السوابق واللواحق في اللغة العربية ثراء لا حد له، ويمكن من خلالها التعبير بالكلمة نفسها عن الإفراد والتثنية وضمائر الفاعل والمفعول والأزمنة والاستفهام ... الخ.

خامساً: يمثل البناء النحوي للغة العربية عمقاً دلالياً لا نظير له، حيث تقود حركات الإعراب المختلفة إلى حسم العديد من أوجه اللبس حسماً لا يدع مجالاً لسوء الفهم.

في هذا السبيل لابد أن نبدأ بالتعرف على الخصائص المميزة للغة العربية وما أكثرها فقد لا يكون من قبيل الزهو أو المبالغة أن نقول إن اللغة العربية هي أفضل اللغات[1]؛ فهي اللغة التي نزل بها القرآن الذي هو أفضل الكتب، ولغة الإسلام الذي هو خير الأديان، ولغة الأمة الإسلامية التي هي خير الأمم، ولغة الحضارة الإسلامية التي هي أعرق الحضارات وأنفعها للبشرية، وهي بعد ذلك لغة خالدة خلود التاريخ تولد لغات وتموت، وتبلى لغات وتنقرض أخرى، وهي باقية بقاء العصور إلى أن يرث الله الأرض ومن عليها ؛ وذلك لأنها محفوظة بحفظ الله تعالى القائل : ﴿ إنا نحن نزلنا الذكر وإنا له لحافظون ﴾[2].

ومن أبرز الاستنتاجات التي توصل اليها الباحثون[3] في اللغات عن خصائص اللغة العربية كالآتي:

أ- أنها اللغة (التامَّة الحروف، الكاملة الألفاظ، لم ينقص عنها شيء من الحروف فيشينها نقصانه، ولم يزد فيها شيءٌ فيعيبها زيادته، وإن كان لها فروع أخرى من الحروف فهي راجعة إلى الحروف الأصلية. وسائر اللغات فيها حروف مُولَّدة وينقصُ عنها حروف أصيلة)[4].

وقد تحدث العقاد عن هذه الخاصيَّة بقولـــه: (فإذا قيس اللسان العربي بمقاييس علم الألسنة، فليس في اللغات أوفى منه بشروط اللغة في ألفاظها وقواعدها)[5].

ثُمَّ يعلل ذلك بقوله: (ويحق لنا أن نعتبر أنها أوفى اللغات جميعًا بمقياس بسيط واضح لا خلاف عليه وهو مقياس جهاز النطق في الإنسان، فإنَّ اللغة العربية تستخدم هذا الجهاز الإنساني على أتمه وأحسنه، ولاتهمل وظيفة واحدة من وظائفه، كما يحدث ذلك في أكثر الأبجديات اللغوية...)[6].

1- القول بتفاضل اللغات لا يتنافى ــ مع المنهج العلمي ، وإن خالف في ذلك أكثر علماء اللغة المعاصرين ؛ ذلك أن التفاضل كائن بين كل شيء : بين البشر وبين الدواب، وفي الأطعمة، والأشربة والثمار والأمكنة وغيرها وحتى بين الرسل والملائكة والقرآن يقرر هذه الحقيقة في أكثر من موضع، إذ يقول تعالى :﴿تلك الرسل فضلنا بعضهم على بعض﴾ (سورة البقرة 253)، ويقول سبحانه عن الأطعمة :" ونفضل بعضها على بعض في الأكل" (سورة الرعد 4)، ونحن نرى البشر يتفاوتون في مواهبهم وملكاتهم العقلية والنفسية والجسمانية، ونرى الشعوب والمجتمعات تتفاوت في خصائصها ومواردها وظروفها، فكيف لا تتفاضل اللغات ؟

2- سورة الحجر، الآية (9)

3- ولو أن هذا الكلام عن أفضلية العربية وتميزها صدر عن العلماء العرب وحدهم لقيل إنه تعصب ومفاخرة،فكيف إذا قاله كبار علماء الغرب ومستشرقوه ؟، والحق ما شهدت به الأعداء

4- القلقشندي: صبح الأعشى 149/1، طبعة، 1383 هـــ-1963 م، عن المؤسسة المصرية العامة..، القاهرة، وانظر: نايف معروف: خصائص العربية وطرائق تدريسها: ص 38،39 الطبعة الثانية 1407 هـــ-1987 م، عن دار النفائس – بيروت.

5- أشتات مجتمعات في اللغة والأدب: ص 11، عن دار المعارف – مصر، الطبعة السادسة، (بدون تاريخ)، وانظر: محمد مصطفى بن الحاج عالميَّة اللغة العربية، مرجع سابق، ص 270.

6- أشتات مجتمعات في اللغة والأدب، سابق ص 11، 12،، وانظر: محمد مصطفى الحاج: عالمية اللغة العربيَّة، سابق ص 270.

واعترف بعض المستشرقين بهذه الخاصية، إذ يقول (رينان): (من أغرب المدهشات أن تثبت تلك اللغة... وتصل إلى درجة الكمال وسط الصحارى عند أُمَّة من الرُّحَّل، تلك اللغة التي فاقت أخواتها بكثرة مفرداتها ودقة معانيها وحسن نظام مبانيها)[1].

ويستمر في وصف خاصية عميقة تفردت بها لغة القرآن، فيقول: (ولم يعرف لها في كل أطوار حياتها طفولة ولا شيخوخة، ولانكاد نعلم من شأنها إلاَّ فتوحاتها وانتصاراتها التي لاتبارى، ولانعرف شبيهًا بهذه اللغة التي ظهرت للباحثين كاملة من غير تدرج وبقيت حافظة لكيانها من كل شائبة)[2].

ب – لاتقتصر اللغة العربية على كونها وسيلة تعبير كما تذهب إليه البحوث النظرية في الغرب التي لاتفرق بين لغة وأخرى؛ لأنها كلها وسائل تعبير وتواصل وتفاهم وإنَّما تتميَّز اللغة العربية بأنها ذات مضامين[3] علميَّة ومنهجيَّة وموضوعيَّة وحضارية، وتميَّزت في ذلك كله بالبيان والسهولة والوضوح على الرغم مِمَّا قد يبدو من صعوبة تعلمها في بادئ الأمر[4].

ويذكر الباحثون المختصون إنَّ واقعها (منحها نوعًا من التميُّز والتفرد إذ أكسبتها تجربتها الحضارية على مدى قرون ثروة هائلة من البنى. واحتبست تعابيرها في أرحامها قدرات خفيَّة على العطاء وعلى الإيحاء وعلى تنويع التعبير... كما أكسبها انتشارها الواسع في بقاع فسيحة من الأرض وتفاعلها مع جماعات لغويَّة كثيرة ألوانًا من الغنى، تأثرًا وتأثيرًا، فهي إذن ليست... اللغة الأوليَّة البدائية التي تحاول أن تصبو إلى مقاربة الحضارة أو ملاحقتها أو الاندماج فيها... وإنَّما هي اللغة ذات التجربة السابقة وما كان لظاهرة ما اجتماعية أو إنسانية أن تقوى على التخلي عن تجاربها السابقة، فهذه التجارب جزءٌ منها)[5].

1- الحاج، مرجع سابق، ص 274، وانظر: نائف معروف، خصائص العربيَّة وطرائق تدريسها: ص 40، .

2- ويعـــد هذا سرٌ اختصت به اللغة العربية مِمَّا جعل الباحثين يجتهدون في أصل نشأة اللغة العربية، وذهب بعضهم إلى أنها لغة توقيفية، واستدلوا بقوله تعالى: ﴿ وَعَلَّمَ ءَادَمَ ٱلۡأَسۡمَآءَ كُلَّهَا ثُمَّ عَرَضَهُمۡ عَلَى ٱلۡمَلَٰٓئِكَةِ ﴾ سورة البقرة: الآية (31)]، للمزيد من التفصيل ينظر: الحـــاج: المرجع نفسه: ص 274، وانظـــر: نائف معروف: المرجع نفســه: ص 40 وقال بعضهم الآخر: إنها اصطلاحية... لزيادة الاطلاع على هذه القضية؛ انظر: ابن فارس: الصاحبي: ص 6 – 31، (مرجع سابق)، وابن جني: الخصائص 44/1 – 47، تحقيق: محمد علي النجار، (بدون تاريخ ولم يذكر الناشر)، وانظر: السيوطي: المزهر في علوم اللغة وأنواعها: 8/1 – 55، تحقيق: محمد أحمد جاد المولى وآخرين، طبعة دار إحياء الكتب العربيَّة (عيسى البابي الحلبي وشركاه)، بدون تاريخ، وانظر: ابن حزم: الإحكام في أصول الأحكام، سابق 29/1 – 35، وانظر: محمد الخضر حسين: دراسات في العربية وتاريخها: ص 10، الطبعة الثانية 1380 هـــ–1960 م، عن المكتب الإسلامي، دمشق، وعددًا آخر من المراجع الحديثة، ونوقشت هذه القضية حتى عند اللغويين في الغرب وأكثروا القول فيها وطال الجدل حولها مِمَّا دعا الجمعية اللغوية الفرنسية إلى اتخاذ قرار بمنع البحث في هذه المسألة في قاعات الجمعية لعدم جدوى البحث فيه ولعقم النظريات المطروحة بصدده حسب زعمهم؛ انظر: نايف معروف،مرجع سابق، ص 23.

3- للمزيد من التفصيل ينظر: ابن فارس: الصاحبي في فقه اللغة.. ص 77 – 81، (مرجع سابق)، وانظر: شكري فيصل: قضايا اللغة العربية المعاصرة...، بحث مدرج في: من قضايا اللغة العربية المعاصرة، المرجع السابق نفسه: ص 32.

4- جميل عيسى الملائكة: اللغة العربيَّة ومكانتها في الثقافة العربيَّة الإسلاميَّة، سابق ص 129، وانظر: محمد سعيد بن رسلان: فضل العربية، سابق، ص 40 – 44.

5- شكري فيصل، قضايا اللغة العربيَّة المعاصرة، سابق، ص 32.

ومن أعظم تجاربها أنَّها خالطت لغات كثر (فلم تفسد في ألفاظها ولا في اشتقاقاتها، ولا في تراكيبها وأساليبها، أو في بيانها الدقيق المشرق، ولم يتعد تأثيرها بها عددًا محدودًا من الألفاظ التي تعربت استجابة لمتطلبات تطور أنماط الحياة وتنظيمها وازدهار الجوانب الاقتصادية والاجتماعية والعقلية والعسكرية والعمرانية في ربوع بلاد المسلمين)[1].

وإذا كان من سنن الله في خلقه أن تتأثر اللغة سلبًا وإيجابًا بواقع الأُمَّة فتعز بعزها وتذل بذلها فإنَّ اللغة العربية خضعت لهذه السنة ردحًا من الزمن ولكن جذوتها لم تنطفئ، والسر في ذلك ارتباطها بالقرآن الكريم فحفظت بحفظه[2]﴿ إِنَّا نَحْنُ نَزَّلْنَا ٱلذِّكْرَ وَإِنَّا لَهُۥ لَحَـٰفِظُونَ ﴾[3].

إن رمي اللغة العربية بالقصور وعدم الكفاية العلميَّة تهمة لاتتفق مع حقيقة اللغة العربية؛ لأنَّها لغة حيَّة عمليَّة لها طاقة هائلة على استيعاب المعاني الغزيرة في الكلمات القليلة[4]، يقول الإمام الشافعي رضي الله عنه عن هذا الجانب في اللغة العربية: (ولسان العرب أوسع الألسنة مذهبا، وأكثرها ألفاظًا، ولانعلمه يحيط بجميع علمه إنسان غيرُ نبيٍّ)[5].

وقد أقرت هيئة الأمم المتحدة عالمية اللغة العربية وأدرجتها في اللغات المعتمدة (كلغة سادسة لشعوب الأرض كافة، يتكلمها ما يزيد على (180) مليون من العرب، ويقدسها المسلمون؛ لأنَّهَا لغة القرآن الكريم ولغة الرسول ﷺ)[6].

ومِمَّا جاء في قرار الجمعيَّة العامَّة للأمم المتحدة الذي اتخذته بالإجماع في دورتها الثامنة والعشرين لسنة (1973 م): (إنَّ اللغة العربية أدت دورًا مهمًّا في الحفاظ على حضارة الإنسان وتراثه الثقافي، وفي العمل على نشرهما)[7].

فإلى يومنا هذا تتعرض اللغة العربية لهجوم شديد من الداخل والخارج، وذلك بغرض تدمير وسحق الدعامة الأساسية للحضارة العربية والتراث العربي والإسلامي.

1- محمد مصطفى الحاج، عالمية اللغة العربيَّة، سابق، ص 258 .

2- المرجع نفسه: ص 258.

3- سورة الحجر: الآية (9).

4- نايف معروف: خصائص العربية، سابق، ص 74.

5- الرسالة، تحقيق: أحمد محمد شاكر، مرجع سابق، ص 42.

6- أحمد محمد جمال: تأمر الأعداء على لغة القرآن، (مجلة رابطة العالم الإسلامي، محاضرات موسم 1394/93 هـ): ص 98، تصدر عن رابطة العالم الإسلامي، مكة المكرمة، وانظر: إسماعيل العرفي: اللغة العربية أم اللغات ولغة البشرية، دار الفكر، دمشق، الطبعة الأولى 1406 هـ - 1985 م، ، ص 21 .

7- أحمد محمد جمال، مرجع سابق، ص 98.

وآخر الادعاءات، أو الافتراءات، توجه ضدها هوقصورها عن استعاب علوم العصر[1] وعدم قابليتها للمعالجة الآلية باستخدام الحاسوب[2].

إن إثبات زيف وتضليل هذه الادعاءات، على المستويين اللغوي والتقني، ليس فقط ضربا من الحمية القومية اللغوية، ولكنه عامل مهم لتخليص أخصائي الحاسوب ومستخدميه أسر هذا المفهوم الخاطئ الذي يحد من طموحهم، ويضع قيودا مصطنعة على البشر وما يبدعونه من نظم آلية.[3]

ولئن خرجت اللغة العربية من صراعها أو كادت أن تخرج بهذه النتيجة فإنها في حقيقة الأمر اللغة الأولى ويكفيها شرفًا أنَّ الله اختارها لكلامه المجيد، وما أصدق ما قاله (حافظ إبراهيم) في رده على شبهات المستشرقين وتلاميذهم على لسان اللغة العربية[4]:

1- أما في مجال الدرس اللغوي للعربية، فإننا نسمع كثيراً من الأصوات التي تسم لغتنا بالصعوبة والتعقيد، وتصفها بالتخلف، وعدم القدرة على مجاراة الركب الحضاري المتجدد، وتطالب بالانحراف عنها، والابتعاد عن قوالبها وقواعدها، واستبدالها باللغات الأجنبية، أو العاميات العربية، ولاشك لدينا، في أن مرد ذلك كله يعود إلىإلصاق تهمة الصعوبة والتعقيد بلغتنا العربية، على وجه خاص، كان يصدر في الأعمِّ الأغلب عن جهة مشبوهة تبتغي من وراء دعوتها تشويه الوجه المشرق للغة العربية، التي كانت أيام قوة أبنائها وعزتهم عبر قرون ممتدة، لغة حضارة تمكّنت من تجسيد الفكر، والوجدان الإنساني في أسمى صورة، وأدق تعبير، وهم بدعوتهم التي يلبسونها أثواباً براقة خادعة، يهدفون إلى دفع أبناء هذه الأمة إلى العزوف عن لغتهم الغنية، التي وسموها بالصعوبة والتعقيد، واللجوء إلى اللهجات العامية الفقيرة، التي يسمونها بالسلاسة والسهولة، والتي تؤدي-كما يخططون- إلى تفتيت الأمة عقدياً، وفكرياً، وسياسياً. وهذه الجهة يعرفها كثيرون من أبناء هذه الأمة، ممن شرفوا بأمانة حراسة هذه اللغة، والدفاع عنها بأفكارهم، ودراساتهم، وممارساتهم.

ولكن كلامنا هذا لا يعني بحالٍ من الأحوال الاستسلام إلى الواقع اللغوي الذي تشوبه، أو يمكن أن تشوبه، وتشوهه، بعض الهنات، أو النقائص، أو التعقيدات، كما يسميها بعضهم، وقدنسي أو تناسى من يدعي جمود اللغة العربية عن مواكبة العصر، أن اللغة أي لغة لا تجمدبنفسها، ولا تتخلف بطبيعتها، كما أنها في المقابل لا تنمو وتزدهر منعزلة عن مجتمعهاوما يجري فيه من أحداث.

يقولالدكتور كمال بشر: "إن جمود اللغة وتخلفها، ونموها وازدهارها، كل أولئك يرجع أولاًوآخراً إلى وضع أهليها، وإلى نصيبهم من التعامل والتفاعل مع الحياة، وما يجري، في العالم من أفكار وثقافات ومعارف جديدة ومتنامية، فإن كان لهم من ذلك كله حظ موفورانعكس أثره على اللغة، وإن قل هذا النصيب أو انعدم، بقيت اللغة على حالها دون حراك أو تقدم، اللغة لا تحيا ولا تموت بنفسها، وإنما يلحقها هذا الوجه أو ذاك بحسب الظروف والملابسات التي تحيط بها، فإن كانت الظروف فاعلة غنية بالنشاط العلميوالثقافي والفكري، كان للغة استجابتها الفورية ورد فعلها القوي تعبيراً عن هذهالظروف وأمارة ما يموج به المجتمع من ألوان النشاط الإنساني، وإن حرمت اللغة من هذاالتفاعل ظلت على حالها وقدمت للجاهلين فرصة وصمها بالتخلف والجمود، في حين أن قومهاهم الجامدون المتخلفون"، ويردعلى هذه الشبهة أيضاً الدكتور عبدالرحمن رأفت الباشا فيقول: "وأما قضية جمود اللغةوعدم تطورها مع الزمن كما يرجف المرجفون فتلك قضية باطلة ودعوة على ظاهرها ملامحالرحمة، وتكمن في باطنها صنوف العذاب، فلقد أفض مضاجع الأعداء أن تكون اللغة العربية هي الوحيدة بين لغات الأرض التي اتصل تليد تراثها بطريفه خلال خمسة عشر قرناً امتدت منذالنابغة في الجاهلية إلى شوقي في العصر الحديث، والتي يستطيع الملايين من أبنائنافي العصر الحاضر تلاوة القرآن الكريم والحديث الشريف وأن يفقهوا معانيهما، وأنيدركوا هديهما، وأن يستشعروا عظمتها، وأن يتملوا مما حفلا به صلاح وإصلاح". وللمزيد حول الرد على هذه الشبهة ينظر: عبد التواب، رمضان، الفصحى وتحديات العصر، بحث ضمن كتاب بحوث ومقالات في اللغة، مكتبة الخانجي، القاهرة، ص167-172، وينظر: اللغة العربية بين الوهم وسوء الفهم، ط. دار غريب، القاهرة، 1999م، ص54.

2- نبيل علي، اللغة العربية والحاسوب، دراسة بحثية، تعريب، د.ط، 1988، ص 10

3- المرجع نفسه، ص 10

4- قصيدة قالها شاعر النيل " حافظ إبراهيم " مدافعا ومنا فحا عن اللغة العربية، اللغة التي يفتخر بها العرب والمسلمون ويعتزون بها، فهي تحفظ كتابهم وتشريعهم، وتعبر عن علومهم وآدابهم.. حين تعالى الهمس واللمز حولها في أوساط رسمية وأدبية، وعلى مسمع ومشهد من أبنائها واشتد الهمس وعلا، واستفحل الخلاف وطغى، فريق يؤهلها لاستيعاب الآداب والمعارف والعلوم الحديثة، وفريق جحود، يتهمها بالقصور والبلى وبالضيق عن استيعاب العلوم الحديثة.. ولكن حافظاً الأمين على لغته الودود لها يصرخ بوجوه أولئك المتهامسين والداعين لوأدها في ربيع حياتها بأن يعودوا إلى عقولهم ويدركوا خزائن لغتهم فنظم هذه القصيدة يخاطب بلسانها قومه ويستثير ولاءهم لها وإخلاصهم لعرائسها وأمجادها.

وَسِعْــتُ كِتَــابَ اللهِ لفظًــا وغَايَــةً وَمَــا ضِقْتُ عن آيٍ بــهِ وعِظَــاتِ

فَكَيْفَ أَضِيقُ اليَــــوْمَ عَنْ وَصْفِ آلَةٍ وَتَنْسِيــقِ أَسْــماءٍ لِمُخْــتَرَعَــاتِ

أَنَــا البَحْرُ فــي أَحْشَائِهِ الــدُرُّ كَامِنُ فَهَــــلْ سَأَلُوا الغَوَّاصَ عَن صَدَفَاتِي؟[1]

أمَّا الاعتراف للغة العربيَّة بأنها حافظت على تراث الإنسان وعملت على نشره فإنَّ ذلك جزءٌ من الحقيقة، وجزؤها الآخر هو ما أسهمت به اللغة العربيَّة من صنع الحضارة الحديثة في مختلف مجالاتها. ونستطيع القول هنا أن أحد الأهداف الرئيسية لهذا البحث هو تبديد تلك المفاهيم الخاطئة عن العلاقة بين اللغة العربية والحاسوب، وإبراز خصائص هذه اللغة التي تؤهلها كموضوع مثير للمعالجة الآلية[2]. وكل ذلك يدل على كمال هذه اللغة وشرفها وفضلها، وقدرتها على مواكبة تطورات العصر في العلوم والتقنية، وهو ما شهد به علماء العرب والغرب على السواء.

إنّ الأهمية الكبرى لهذه الخصائص هو مطلب أساسي لفهم طبيعة العلاقة بين اللّغة والحاسوب ومعالجتها آليًا[3] وسنحاول في هذه العُجالة أن نتجاوز حدود الأحكام القيمية لمنظومة اللّغة العربية مركزين في الوقت نفسه على أهم خصائص اللّغة العربية[4] باعتبارها إحدى الحالات الخاصة للمنظومة اللّغوية العامة، وتجدر الإشارة أننا نبرز أهم الخصائص ذات المغزى لمعالجة اللّغة العربية آليًا، فالهدف من هذا كلّه هو توسيع النظرة إلى هذا الجانب الهام في دراسة اللّغة العربية[5].

1- ديوان حافظ إبراهيم (ضبط وتصحيح وشرح وترتيب): أحمد أمين، أحمد الزين، إبراهيم الأبياري، 253/1، 254، عن دار الجيل – بيروت، (بدون تاريخ)، ص 254،253.

2- لم يقتصر الأمر في الدعوة إلى العناية بحوسبة اللغة العربية وعلومها على المبادرات الفردية للباحثين وبعض المهندسين العرب الذين اشتهروا بحبهم للغة العربية وغيرتهم عليها، إذ سرعان ما بدأت المجامع العلمية المهتمة باللغة العربية وعلومها تعقد الندوات العلمية لمقاربة موضوع اللغة العربية والمعلوميات، نذكر منها على سبيل المثال ندوة الجزائر الدولية في ديسمبر 2002حول "اللغة العربية وتكنولوجيا المعلومات"، وقد راهنت هذه الندوة على إبراز دور اللغة العربية وقدرتها على مسايرة التطورات التكنولوجية الحديثة.

3- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 43.

4- تمام حسان، اللّغة العربية معناها ومبناها، سابق، ص 29.

5- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 69.

2-1- خصائص اللغة العربية من منظور صرفي ـ حاسوبي :

إنّ اللّغة العربية أعقد اللّغات "الساميّة" وأغناها صوتًا[1]، وصرفًا[2]، ومفردات[3] ومن أهم خصائص منظومة "اللّغة العربية" نذكر[4]:

- ❖ التوسط اللّغوي: فهي تسهل تطويع النماذج البرمجية المصممة للغة العربية لتلبية مطالب اللّغات الأخرى وعلى رأسها الانجليزية[5].
- ❖ حدّة الخاصية الصرفية[6]: والتي تتميز بخاصية الاشتقاق المبني على أنماط الصّيغ[7].

1- تخضع الفونيمات العربية لقواعد فونولوجية تحكم تتابعها في سياق الكلمة أو الجملة، وقد يطرأ على الفونيم تغيير في صفته تأثراً بما يسبقه أو ما يليه من الفونيمات. ويعني ذلك أن التوصيف الفوناتيكي لأصوات العربية لا يكفي وحده ليمكّن الحاسوب من تمييز الفونيمات، وأنه يحتاج إلى توصيف آخر لما يعتري فونيمات العربية – في سياقها الوظيفي – من تغييرات تحكمها القواعد الفونولوجية على مستوى التشكيل الصوتي. فالحرف المضعف مثلا تزيد المدة الزمنية التي تستغرق نطقه1، والدال في (أردت) تنطق تاء بسبب التأثير الرجعي، والباء في (سبت) يصيبها من أثر التاء المهموسة ما يزيل عنها صفة الجهر، ولفظ الجلالة (الله) بعد غير الكسر تكون فيه اللام مفخمة من قبيل التنوع الألوفوني، وكلمتا (سوط – صوت) قد يتطابقان في سياق الأداء النطقي1، ولام (ال) لا تنطبق إذا تلتها إحدى الحروف الشمسية.

إن تصميم برنامج حاسوبي يستدخل الحدث الكلامي المنطوق ويحوله إلى نص مكتوب يتطلب الوقوف على تلك المواضع التي تتغير فيها صفات الفونيمات تأثراً بما يجاورها على مستوى الكلمة. وإذا ما انتقلنا إلى مستوى الجملة لزمنا التنبه إلى ما ينطق من الفونيمات ولا يكتب كالتنوين والمد في (هذا وأخواتها – الله – لكن)، وكذلك ينبغي التحرز من الفونيمات التي تكتب ولا تنطق كهمزة الوصل واللام الشمسية وألف (مائة) والتاء المربوطة عند الوقف والألف الفارقة بعد واو الجماعة، والألف المقصورة المتلوة بـــ (ال).

وجدير بالذكر هنا الإشارة إلى ما يواجه الحاسوبيين من إشكالات في استدخال النصوص المطبوعة عبر الماسح الضوئي scanner، وأن ثمة حروفاً يشكل على الحاسوب معرفتها. ويلجأ الحاسوبيين – أجل معالجة تلك المشكلات آلياً – إلى تزويد الحاسوب بثبت لما يتقارب من الحروف رسماً، وثبت آخر يرصد تتابع الحروف الأكثر شيوعاً مرتبة حسب النسب المئوية لدرجة الشيوع.

منصور الغامدي، الإدراك الآلي للتضعيف في اللغة العربية، السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات، مكتبة الملك عبد العزيز، الرياض، 1993م، ص 83.للمزيد من التفصيل ينظر : سالم الغزالي، المعالجة الآلية للكلام المنطوق: التعرف والآلية، من كتاب: استخدام اللغة العربية في المعلوماتية، المنظمة العربية للتربية والثقافة والعلوم، تونس، 1996م، ص 134.و محمد البكاء، الحاسوب لطلبة العلوم الإنسانية: اللغة العربية، مكتبة الفلاح، الكويت، ط1، 2005م، ص 139.

2- ينظر:

-Bessou Sadik (Analyse de donnée Textuelles pour la Classification Automatique par les techniques de text Mining, application à la langue Arabe) thèse Magister 2007- Université Ferhat ABBAS. P 63.

-Bessou Sadik ,Louail Mohamed, Refoufi Allaoua, Kadem Zehour, et Touahria Mohamed (un système de lemmatisation pour les applications de TALN) colloque International. CITALA'07- 18-19, juin/2007.

3- ينظر: نبيل علي،الثقافة العربية وعصر المعلومات، رؤية لمستقبل الخطاب الثقافي العربي. عالم المعرفة-يناير 2001- العدد 265- ص 242.

4- نبيل علي،اللّغة العربية والحاسوب، سابق، ص 60.

5- أثبتت بحوث الكاتب نبيل علي بما لا يدع مجالاً للشك إمكان استخدام نظم الإعراب والصرف الآلية المصممة للغة العربية في مجال اللّغة الانجليزية، ينظر: نبيل علي، الثقافة العربية وعصر المعلومات، رؤية لمستقبل الخطاب الثقافي العربي، سابق، ص 242.

6- ينبغي ابتداء أن نلمح إلى أن توصيف النظام المورفولوجي يهدف إلى تمكين الحاسوب من التعامل مع المورفيمات (الوحدات الصرفية) التي تشكل التراكيب اللغوية المتنوعة. وينبغي ونحن في هذا الصدد التمييز بين مطلبين متمايزين: المطلب الأول المتمثل في القدرة على (التوليد) أي صوغ الأبنية اللغوية السليمة بخطوات إجرائية محددة على نحو شامل مستقص، والمطلب الثاني الذي يمكن الحاسوب من (التحليل) بتعيين المعاني الصرفية وتفكيك التراكيب اللغوية إلى وحداتها الصرفية الصغرى حين ترد في سياق الجملة أو في سياق النص.

أما قواعد تتابع المورفيمات في الأنساق اللغوية المتنوعة، وحصر ما تجيزه العربية من أشكال رصف المورفيمات وما تمنعه على مستوى الجملة؛ فيستعان به برد أنماط الجمل إلى نسقين رئيسيين هما الجملة الاسمية المكونة من مبتدأ وخبر فحسب، والجملة الفعلية المكونة من فعل لازم وفاعل وحسب. ومن ثم النظر فيما تسمح به قواعد اللغة من دخول مورفيمات إضافية إلى ذينك النسقين، وما تتمتع به من مرونة في التقديم والتأخير والحذف. والحق أن هذا المطلب ضخم مجهد عسير لما تتميز به اللغة العربية من تعدد أنماطها الجملية كونها من اللغات ذات النسق الحر، وقد خلُص أحد الباحثين إلى أن الأنماط الجملية الممكن تأليفها على وفق قواعد النظم في العربية تصل إلى بضعة عشر ألف نمط.لبسط القول ينظر:*سعد مصلوح، في البلاغة العربية والأسلوبيات اللسانية، جامعة الكويت، الكويت، ط1، 2003م، ص 113.

7- محمّد المبارك، فقه اللّغة وخصائص العربية، دراسة تحليلية مقارنة للكلمة العربية وعرض لمنهج العربية الأصيل في التجديد والتوليد، سابق- ص 276 وما بعدها.

- ❖ المرونة النحوية[1]: من خلال ترتيب الكلمات داخل الجمل، تبرز هذه المرونة في التقديم و التأخير، وتعد هذه الخاصية تحديا لمعالجة النحو آليا, نظرا لكون النماذج المتاحة مصممة أصلا لتلائم اللغة الإنجليزية.
- ❖ الانتظام الصرفي في العربية التي يتميز بظاهرة الاطراد ويتميز نظامه بالبساطة.
- ❖ ظاهرة الإعراب.
- ❖ الحاسية السياقية: تعدّ أحد العوامل الهامة في تحديد نظام التقعيد اللّغوي.
- ❖ تعدد نظم الكتابة: كتابة تامة التشكيل، كتابة مشكولة جزئيًا، كتابة خالية من التشكيل.
- ❖ اعتماد المعجم على الجذور: إذ يرتبط تنظيم المعجم بصلة وثيقة مع قواعد تكوين كلمات اللّغة، لذا كان من الطبيعي أن يأتي المعجم العربي معتمدًا على الجذور وليس على لترتيب الأبجدي للكلمات كما في الانجليزية.
- ❖ شدة التماسك بين عناصر منظومة "العربية".
- ❖ وفرة الفائض اللّغوي.
- ❖ ثنائية الفصحى والعامية.
- ❖ تمثل تركيبات السوابق واللّواحق في اللّغة العربية ثراءً لا حدّ له، ويمكن من خلالها التعبير بالكلمة نفسها عن الإفراد والتثنية وضمائر الفاعل والمفعول والأزمنة والاستفهام[2]...الخ.
- ❖ يمثل البناء النّحوي للغة العربية عمقًا دلاليًا لا نظير له، حيث تقود حركات الإعراب المختلفة إلى حسم العديد من أوجه اللّبس حسمًا لا يدع مجالاً لسوء الفهم[3].

ولقد تجلت حكمة المولى عز وجل باختياره وتشريفه للغة العربية بترتيل القرآن الكريم بلسان عربي مبين، فالبناء الرياضي[4] غير المتكرر للغة العربية أتاح لها القدرة على استيعاب العلوم الحديثة والتطور غير المحدود تبعاً لمقتضيات العصر من دون أن تفقد طابعها المميز ولا سماتها الأساسية من خلال هذه الخصائص التي ميزتها عن سائر اللغات السامية.

1- يوضح الأستاذ أنور الجندي أن اللغة العربية تملك من المرونة مالا تملكه لغة حية أخرى فالألماني المعاصر مثلا لا يستطيع فهم كلمة واحدة من اللهجة التي كان يتحدث بها أجداده من ألف عام بينما العرب المحدثون يستطيعون فهم لغتهم التي كتبت في الجاهلية قبل الإسلام ".ينظر: بحوث لغوية، لأحمد مطلوب، عمان، دار الفكر، 1987، ص 26 .

2- علاء الدين صلاح العجماوي،المعالجة الآلية للغة العربية بين الواقع والتحديات، المحاضرة رقم 2، الموسم الثقافي – التاسع عشر 2001- مجمع اللّغة العربية الأردني، ص 07.

3- المرجع نفسه، ص 07.

4 -Structure mathématique

في نهاية حديثنا عن خصائص اللّغة العربية، وجبت الإشارة إلى أننا قد ركزنا على تلك الخصائص فالهدف هو توسيع النظرة إلى هذا الجانب الهام من هذه الدراسة، ومناقشته بشكل أكثر دقة وشمولية[1].

3- المعالجة الآلية للغة العربية:

3 -1- ضبط المصطلحات:

أ-المعالجة[2]: المعالجة من وجهة نظر علم اللّغة الحاسوبي هي التطبيق الآلي[3] على مجموعة من نصوص اللّغة وذلك بتغييرها وتحويلها، وإبداع شيء جديد اعتمادًا عليها، ويتم كل ذلك باستعمال تقنيات وأدوات من علوم اللسانيات والإعلام الآلي والنمذجة[4]، عملية المعالجة تتم باستعمال برنامج خاص يسمح بإدخال النصوص وتنقيحها ومن ثم إخراجها بأشكال متعددة[5].

ويجب التفرقة عند المعالجة بين وصف المعارف باعتبارها وظيفة من وظائف اللّسانيات، والتعبير عن هذه المعارف في نماذج باستخدام تقنيات واستراتيجيات فعّالة مستمدة من علوم الحاسوب وهي وظيفة علم اللّغة الحاسوبي[6].

ب- الآلية[7]: العمليات الآلية هي تجري عن طريق الآلة والتي تقابلها العمليات التي تجري بواسطة الإنسان[8].

والآلة التي تستعمل في المعالجة الآلية للغة هي الحاسوب الذي اخترع لإجراء العمليات الحاسوبية؛ لذا يجب تطويره لمعالجة المعلومات ذات الطبيعة اللّسانية، حيث أنّ المعالجة الآلية

1- للمزيد من التفصيل، ينظر: نبيل علي: اللّغة العربية والحاسوب، سابق، ص 60 وما بعدها.
وانظر أيضًا: نبيل علي،العرب وعصر المعلومات، عالم المعرفة- العدد 184- أبريل 1994- ص 331 وما بعدها.

2- Le traitement.

3- Application automatisée.

4- Cacaly Serg- Model. In: Dictionnaire encyclopédique de l 'information et de la documentation. 1998-p. 403.

- لقد بدأ الاهتمام بعمليات النمذجة Modélisation أواخر الأربعينيات وبداية الخمسينات حينما عمل الباحثون على تطوير آلة متناهية Finite automaton مماثلة لنموذج ماركوف Marcov ذي الحالات المتناهية، بهدف نمذجة بعض العمليات، ومنها العمليات اللّغوية، فالنمذجة تعمل على توفير الأدوات الصورية التي تمكن من بناء أوصاف تركيبية دقيقة للغات الطبيعية الأمر الذي يسمح للحاسوبي بتطوير خوارزميات متطورة تمكنه من تجاوز بعض المشاكل التقنية التي تعترض تطوير برامج حاسوبية ملائمة. للمزيد من التفصيل ينظر:

- الرّحالي محمد "بعض الخصائص الصورية للنمذجة اللّسانية" ضمن: قضايا في اللّسانيات العربية- ط1- الدار البيضاء. منشورات كلية الآداب والعلوم الإنسانية- ابن مسيك 1992، ص 21.

- بن عيسى زغبوش مصطفى بوعناني"المعالجة الآلية للغة" -مسارات تحليل الجمل العربية باعتماد نماذج شبكات الانتقال- مجلة كلية الآداب والعلوم الإنسانية- ظهر المهراز (المهراز)، فاس- العدد- 14- 2006، ص من 55، 84.

5- ينظر: عبد الحسن الحسيني، معجم مصطلحات علوم الكمبيوتر، انجليزي فرنسي- عربي انجليزي- سابق، ص 411.

6 - DELAFOSSE, RIONEL-Automatique- In: Glossaire de linguistique Computationnelle. - [Enline].
http :// page sperso-orange. Fr/idela fossaire/A-htm.

7 - Automatique.

8 - Delafosse, Rionel- Idem.

والآلة التي تستعمل في المعالجة الآلية للغة هي الحاسوب الذي اخترع لإجراء العمليات الحاسوبية؛ لذا يجب تطويره لمعالجة المعلومات ذات الطبيعة اللّسانية، حيث أنّ المعالجة الآلية هي تتابع حركات حسابية تقوم بها الآلة وفق تسلسل زمني أي أن برنامج المعالجة الآلية[1] يمكن أن يكون كلّي أو جزئي، حيث أنّ:

1-كلّي "total": الآلة تؤدي وظائف معينة بشكل آلي ودون تدخل الإنسان "Automate"[2].

2-جزئي:"Partiel": يتدخل الإنسان في بعض المراحل.

ومعالجة شيء لساني من وجهة الآلة يواجه بقيود "contraintes" كثيرة في وصف النصوص اللّغوية نفسها، لذا يجب نمذجة مكونات النصوص بطريقة واضحة ومتناسقة "coherence"[3].

جـــ- اللّغة[4]: يختص الإنسان باللّغة[5] التي هي وسيلة للتعبير عن النفس[6] والاتصال[7] فاللّغة نظام من الرموز، ومعنى هذا أن اللّغة تتكون من مجموعة من الرموز تكون نظامًا متكاملاً[8] فهي عبارة عن عدد في الأنظمة يتألف كل واحد منها من مجموعة من "المعاني" تشمل مجموعة من الوحدات التنظيمية أو "المباني"[9] المأخوذة من نظامين: نظام (نظام) صرفي، وحدات صرفية "**المورفيمات**"[10].

1- traitement automatique./ Programme automatique.

2- ينظر: عبد الحسن الحسيني، معجم مصطلحات علوم الكمبيوتر، إنجليزي- فرنسي- عربي- سابق، ص 77.

3- فارس شاشة،المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي إعرابي للفعل العربي،مذكرة ماجستير في علم المكتبات والتوثيق، كلية العلوم الاجتماعية والإنسانية، قسم علم المكتبات والتوثيق جامعة الجزائر، 2008، ص 13.

4-La Langue.

5- من أقدم هذه التعريفات وأشهرها في التراث العربي التعريف الّذي وضعه العالم العربي "أبو الفتح عثمان بن جنّي (ت 392هـ) يقول فيه: "أمّا حدّها فهي أصوات يعبر بها كل قوم عن أغراضهم"

-ينظر: ابن جني، أبو الفتح عثمان "الخصائص" ج 1 تحقيق محمد علي النّجار مطبعة دار الهدى بيروت، ط2، ص 33.

6- يحاول علم اللّغة أن يجد طرقًا لدراسة اللّغة باعتبارها ظاهرة إنسانية عامة، طرقًا تصلح لدراسة جميع الأشكال الكلامية التي تصطنعها الجماعات البشرية على اختلافها.

- ينظر: محمد رشاد الحمزاوي "المصطلحات اللّغوية الحديثة في اللّغة العربية" معجم عربي أعجمي وأعجمي عربي- الدار التونسية للنشر، تونس- المؤسسة الوطنية للكتاب- الجزائر- د ط- 1987- ص 135.

7- نايف خرما "أضواء على الدراسات اللّغوية المعاصرة"- سلسلة عالم المعرفة- العدد 9- سبتمبر 1978- ص 26.

8- محمود فهمي حجازي مدخل إلى علم اللّغة. دار قباء. القاهرة- د ط. د. ت، ص 10.

9- تمام حسان اللّغة العربية معناها ومبناها. دار الثقافة. الدار البيضاء- د ط. 1421- 2001- ص 34.

10- المورفيم: صرفيم:Morphème : جاء هذا المصطلح معرفًا ضمن إطار النظريات البنيوية؛ إذ يشير إلى تلك الوحدات الدنيا في اللّسان، المتضمنة لشقي الدال والمدلول. ينظر: ماري نوال غاري بريور، المصطلحات المفاتيح في اللّسانيات، ترجمة عبد القادر فهيم الشيباني الجزائر – 2007- ص 72.

إذا كان الاتصال مكتوبًا، ومن نظام صوتي، "حروف" "الفونيمات"[1]، إذا كان الاتصال شفهيًا، فاللّغة تدرس الظواهر اللّغوية في جميع أشكالها من حيث خصائصها وتطورها وقوانينها[2] . فاللّغة أعم من أن تقتصر على البحث في نشأة الكلمة ومراحل تطورها وأشكال ضبطها[3].

وقد تناول هذا التعريف عدد من الباحثين في علم اللّغة[4]. وذلك على ضوء ما انتهى إليه الفكر اللّغوي الحديث وانتهوا جميعًا إلى أنّ هذا التّعريف يتضمن عدّة حقائق تتصل بماهية اللّغة ووظيفتها وهي[5]:

1-الطبيعة الصوتية للغة.

2-الوظيفة الاجتماعية للغة من حيث كونها أداة للتعبير والاتصال.

3-اختلاف اللّغة باختلاف المجتمع.

3-2- ضبط المفاهيم:

3-2-1- تعريف المعالجة الآلية للغة: (TAL)[6]

تقوم بمعالجة معطيات لغوية (نصوص) المعبّر عنها في لغة طبيعية، وحتى يمكن معالجة هذه المعطيات يجب توضيح قواعد اللّغة والتعبير عنها بواسطة نماذج[7] تكون عملياتية وحسابية باستخدام برامج حاسوبية[8] قادرة على معالجة اللّغة الطبيعية حاسوبيًا[9].

1- الفونيم: (صوتيم): Phonème: يمثل ذلك الصوت الذي يؤدي داخل اللّسان، دور تمييز وحدتين مفردتين بمعنيين مختلفين، فهو معرف إذا بوظيفته التمييزية؛ تمثل الفونيمات الوحدات القاعدية للسان إذا نجدها محدودة العدد داخل كل لسان، الفونيم هو الذي يمثل وحدة دنيا على صعيد الشكل.

- للمزيد من بسط القول ينظر: ماري نوال غاري بريور (المصطلحات المفاتيح في اللّسانيات) ترجمة: عبد القادر فهيم الشيباني- الجزائر 2007- ص 77- 78.

2- محمّد المبارك، "فقه اللّغة وخصائص العربية"، دراسة تحليلية مقارنة للكلمة العربية وعرض لمنهج العربية الأصيل في التجديد والتوليد، دار الفكر، بيروت- 1425- 2005. د. ط- ص 20.

3- محمّد سمير نجيب اللّبدي "معجم المصطلحات النحوية والصرفية" سابق، ص 205.

4- انظر على سبيل المثال ما كتبه د. حلمي خليل حول ذلك في كتابه، دراسات في اللّغة والمعاجم، دار النهضة العربية- بيروت- الطبعة الأولى- 1998. وانظر أيضًا: د. عبده الراجحي فقه اللّغة في الكتب العربية.

5- ينظر: حلمي خليل،دراسات في اللّغة والمعاجم، سابق، ص 158. ود./رمضان عبد التوّاب "المدخل إلى علم اللغة ومناهج البحث اللّغوي"، مكتبة الخانجي، القاهرة، دار الرفاعي بالرياض- الطبعة الأولى- 1403/ 1983هـ - ص 137 وما بعدها.

6 - Traitement automatique de la langue. TAL.

7 -Modèles.

8- فارس شاشة،المعالجة الآلية للغة الطبيعية، إنشاء نموذج لساني صرفي إعرابي للفعل العربي- سابق، ص: 14.

9-PIERRETTE BOUILLON: avec la collaboration de: Françoise Vandooren- lyne Dasylva- lanrence jacqmin- Sabine lehmann- Graham- Russell et Evelyne Viegas (Traitement automatique des langues naturelles) Edition Duculot. AUPELEF- Uref. 1998. Paris. P:5.

فالمعالجة الآلية تتطلب الحكم على صحة التعبير اللّغوي من عدمه مع الأخذ بالاعتبار أن نظم المعالجة الآلية لابدّ أن تتعامل مع اللّغة على اتساعها أي على أساس من لا نهائيتها[1].

وإنجاز نظام معالجة اللّغة يتطلب البحث عن المعلومات وإجراء دراسات في ميداني فهم النصوص وتعميمها[2] « Génération de texte ".فالنظام يضم مجموعة من العناصر أو القواعد الفاعلة التي تتحكم في قدرات نظم المعلومات[3].

3-2-2- المعالجة الآلية للغة: وتشتمل على شقين أساسيين[4]:

أ-الشق الأوّل: ويشمل نظم البرمجة المستخدمة في المعالجة الآلية بواسطة الكمبيوتر للفروع اللّغوية المختلفة مثل:

<u>أولاً</u>-نظام الصرف الآلي: الّذي يقوم بتحليل الكلمات إلى عناصرها الاشتقاقية والتصريفية، أو يعيد تركيبها في هذه العناصر.

(مثال: تحليل كلمة "بإيجادها" إلى:

حرف الجر: "الباء"

والضمير المتصل: "ها"

وساق الكلمة: "إيجاد" الذي يُحلل إلى جذر "وجد" على صيغة "إفعال".

<u>ثانيًا</u>- نظام الإعراب الآلي: الّذي بإعراب الجمل آليًا.

1- نبيل علي، الثقافة العربية وعصر المعلومات، سلسلة عالم المعرفة- العدد 265- أبريل 2001- ص 270.

2 -Delafosse, Rionel- La génération de texte- In: GLOSSAIRE de linguistique computationnelle- [Enline]. http://pages perso- orange, Fr/ idelafosse/glossaire/AihTM.

3 -Alain Couillault « traitement automatique des langues dans les industries de L'information » Livre blanc. Janvier 2005. Edition. APIL- GFIL. P :04.

4- نبيل علي، الثقافة العربية وعصر المعلومات، "رؤية لمستقبل الخطاب الثقافي العربي"، عالم المعرفة- العدد 265- يناير 2001، ص 290- 291.

ثالثًا–نظام التحليل الدلالي الآلي: الذي يستخلص معاني الكلمات استنادًا إلى سياقها، ويحدد معاني الجمل استنادًا إلى ما يسبقها وما يلحقها من جمل.وذلك علاوة على قواعد البيانات المعجمية والقواميس الالكتروني[1] ومنهجيات هندسة اللّغة[2].

ب–الشق الثاني: ويتضمن التطبيقات التي تقوم على النظم اللّغوية الآلية السابقة الذكر والتي تشمل على سبيل المثال لا الحصر: الترجمة الآلية– التدقيق الهجائي واللغوي– الفهرسة والاستخلاص الآلي– البحث العميق داخل مضمون النصوص– فهم الكلام ونطقه آليًا[3].

3-2-3- تعريف المعالجة الآلية للغة الطبيعية: هي إحدى المجالات التطبيقية للسانيات الحاسوبية يدور محور هذا المجال حول الآلة (الحاسوب) والإنسان فهناك علاقة تفاعل وتواصل (اتصال)[4]. فمن أهداف هذه المعالجة جعل الحاسوب يفهم اللّغة الطبيعية[5].

تمكن هذه الدراسة من استعمال الحاسوب لمعالجة النصوص اللّغوية بكيفية آلية[6]. وميادين التطبيق بالنسبة لهذه المعالجة كثيرة ومشهورة، كالتوثيق الآلي، والفهرسة الآلية والتركيب الآلي للكلام....الخ[7]

1–Dictionnaires électroniques.

2– هندسة اللّغة Langage Engineering: تمثل اللّغة موضوعًا متميزًا ومثيرًا للتناول الهندسي، إذا ما نظرنا إليها كنظام معقد متشابك ونظرنا إلى الهندسة– بصفتها في السيطرة على النظم المعقدة. وهكذا ظهر إلى الوجود مصطلح هندسة اللّغة، كنوع متخصص من فروع هندسة المعرفة والذكاء الاصطناعي.

تتميز الهندسة بقدرتها على تناول الموضوعات التي تفتقد الأساس النظري المكتمل، وذلك بفضل أساليبها التقريبية وأغراضها العملية وفي ظل هذا المفهوم تصبح اللّغات عموما، اللّغة العربية بصفة خاصة في حاجة إلى الهندسة من أجل سد النقص النظري والعملي. فيمكن للهندسة بأساليبها التقريبية أن تسد فجوات التنظير اللّغوي.

وعلاقة اللّغة بهندسة الحاسوب هي علاقة "أخذ وعطاء" فكما أنّه يمكن استخدام الحاسوب حاليًا لإقامة نماذج لغوية، وتحليل الفروع اللّغوية المختلفة، فإنه يمكن تحديد مجالات التطبيقات الحاسوبية على المستوى اللغوي فيما يلي:

– الصرف الحاسوبي.– النحو الحاسوبي.– الدلالة الحاسوبية.

– المعجمية الحاسوبية.– علم النفس اللّغوي الحاسوبي.

لقد أصبحت معالجة اللّغات الطبيعية آليًا بواسطة الحاسوب أحد المقومات الأساسية في تصميم معمارية نظم المعلومات، يمكننا القول علم اللّغة الحديث، قد دخل إلى مصاف العلوم الدقيقة من المدخل السليم، فقد قام على النموذج الرياضي للنحو التوليدي، والّذي يتميز بقابلية عالية للمعالجة الآلية وبالتالي للتطبيق الهندسي العملي.

3– ينظر: نبيل علي، الثقافة العربية وعصر المعلومات، "رؤية لمستقبل الخطاب الثقافي العربي" سابق، ص 260 وما بعدها.

4 –Daniel Kayser, Bernard Levrat, (Traitement automatique du Langage naturel)Ed, HERMES Science Publication Paris, Volume 20 n°=03/2001 P:02.

5– سهاد مال الله كاظم، هالة بهجت عبد الوهاب، إخلاص خلف اكباش، تمثيل قواعد اللّغة العربية كحدود منطقية في أنظمة معالجة اللّغات الطّبيعية، مجلة الهندسة والتكنولوجيا، المجلد 26، العدد 11، 2008، العراق، ص 02.

6 –Dominique ARCHAMBAULT, Bassano, J-C (Directeur de thèse).« Proposition de réseau neuromimétique par le traitement du Langage naturel » thèse de doctorat 3ème cycle 1995. Université d'Orléans, Orléans, France.

7– عبد الرحمن الحاج صالح (بحوث ودراسات في اللّسانيات العربية) الجزء الأوّل، سابق ص 84– 85.

إن نظام المعالجة الآلية للغة الطبيعية يتكون من ثلاث أنظمة رئيسية[1]:

-نظام التحليل.

-نظام التفسير.

-نظام التوليد.

ويُعتبر نظام التحليل والتوليد كأساس لنظم معالجة المعلومات والمعارف[2].

ونظرًا لتعقيد المعرفة اللاّزمة لنظم المعالجة الكلية للغة الطبيعية شملت مستويات هذه المعالجة[3] العناصر المورفولوجية[4] والنحوية[5] والدلالية[6]، والمعنوية[7] معرفة معالجة المعلومات ومعالجة المعارف لأنّ المهام اللّغوية للنظام الآلي تتفاوت بينهما إلى حد كبير[8].

1- Mohamed. Tayeb. LASKRI et R. MAHDJOUBI « Traitement automatique de la langue arabe en vue d'une traduction automatique des textes vert la langue Française » Proc. 3eme journées Internationales de traitement de données textuelles. JADT'95 ROME ; Italie. 11- 13. DECEMBRE 1995.

-Karima MEFTOUH « Un réseau simplement récurrent pour la génération d'une représentation interne du sens en langue arabe basée sur les cas Sémantique » Mémoire de magister Université de Annaba .2000.

2- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 187.

3- للمزيد من بسط القول في هذا الموضوع ينظر:

- Nacera TAIBI Née MELLITE « Contribution à l'étude du traitement automatique des erreurs dans un texte écrit en arabe » Mémoire de Magister English 1997.

&

- Mohamed El Amine ABDRAHIME « Reconnaissance des unités linguistiques signifiantes. Thèse de doctorat spécialité informatique. Université de Tlemcen 2007- 2008.

4- المورفولوجية Morphologie: تعد المورفولوجيات فرعًا من النظرية اللّسانية، وتعنى بدراسة أشكال الكلمات وانتظامها، تهتم "المورفولوجيات الصرفية" بتحولات الكلمة طبقًا للجنس، والعدد، والزمن...للمزيد ينظر: ماري نوال غاري بريور (المصطلحات المفاتيح في اللّسانيات، ترجمة عبد القادر فهيم الشيباني، ص: 74 الجزائر، ط1- 2007، ص 74 .

5- العناصر النحوية Syntaxique: تعد التركيبات أو العناصر النحوية فرعًا من النظرية اللّسانية، وهي تعنى بوصف القواعد الانتظامية لوحدات اللّسان عند تأليف الجمل: يمكن الإشارة إلى المجالات الأساسية التي تختص العناصر النحوية بدراستها:

1 - الوظائف: المحددة لبنية الجملة البسيطة.

2- البنى المختلفة لأنماط المجموعات المؤلفة للجملة: مج إ، مج ف، مج ن وغيرها. Syntagme.

3- النماذج المختلفة لانتظام الجمل المركبة انطلاقًا من جمل بسيطة.....للمزيد من التفصيل

ينظر: ماري نوال غاري بريور،المصطلحات المفاتيح في اللّسانيات، سابق، ص: 104- 105.

6- العناصر الدلالية: Sémantique: تهتم الدلاليات؛ بوصفها فرعًا من النظرية اللّسانية، بوصف مختلف أشكال إقامة المعنى داخل اللّسان. تحاول العناصر الدلالية تبيان معنى الوحدات المفرداتية والجمل والملفوظات ضمن علاقاتها فيما بينها. للمزيد من التفصيل.

ينظر:" ماري نوال غاري بريور،المصطلحات المفاتيح في اللّسانيات، سابق، ص 92.

7- المعنوية Pragmatique: عمد السيميائي شارل موريس (1901) إلى إدراج التقابل بين كل من التركيبات (علاقة العلامات داخل الجملة) والدلالية وعلاقة الجمل بحالات الأشياء التي تدل عليها، و"المعنوية" وعلاقة الجمل بالمتلفظين والمؤولين لها. للمزيد من التفصيل

ينظر: ماري نوال غاري بريور، المصطلحات المفاتيح في اللّسانيات، سابق، ص 82.

8- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 189.

3-2- 4- المعالجة الآلية للغة العربية: المعالجة الآلية لفروع اللّغة العربية المختلفة: صوتًا وحرفًا وصرفًا ونحوًا ومعجمًا تمثل تلك المعالجات اللّغوية الأساسية البنية التحتية التي يمكن أن تقام عليها نظم أعمق لمعالجة اللّغة العربية مثل: **نظم الفهم الآلي العميق لمحتوى النصوص، ونظم التلخيص الآلي** التي تعتمد على فهم هذا المحتوى، لا مجرد الاعتماد على إحصائيات تواتر الألفاظ والجمل[1].

فالمعالجة الآلية للغة العربية هي عبارة علم معرفي أين نجد فيه المزيد من الأبحاث والتكنولوجيات المتعلقة بخصوصيات هذه اللّغة، فهي تفترض أدوات لسانية لتطويرها حاسوبيًا وفق ما يعرف بالمعالجة الآلية[2].

تحتاج المعالجة الآلية للغة العربية إلى أربعة مستويات مختلفة في التحليلات المتداخلة[3] وهي:

- التحليل الصرفي[4].
- التحليل النّحوي[5].
- التحليل الدلالي[6].
- التحليل المعجمي[7].

وهناك بعض الباحثين من يضيف مستوى آخر مثل المستوى الصوتي[8] الذي يعتبر مستوى متنقلاً في اللّغة الانجليزية مثلاً، بينما يندرج التحليل الصوتي[9] ضمن التحليل الصرفي في اللّغة العربية.

1- نبيل علي، الثقافة العربية وعصر المعلومات، رؤية لمستقبل الخطاب الثقافي العربي، سابق، ص 292.

2 - Motasem ALRAHABI, Ghassem MOURAD, Brahim DJIOUA, Filtrage sémantique de textes en arabe en vue d'un Prototype de résumé automatique. (le traitement automatique de la langue arabe-, JEP-TALN 2004, Fès 19-22 avril 2004

3 -Ghayda AL- Talib « An Arabic Natural Language Interface To Data Base System Using Prolog » Master thesis, College of Science, Department of computer Science, University of Mosul, Iraq, 1991, pp:8.

4-Analyse morphologique.

5 - Analyse Syntaxique.

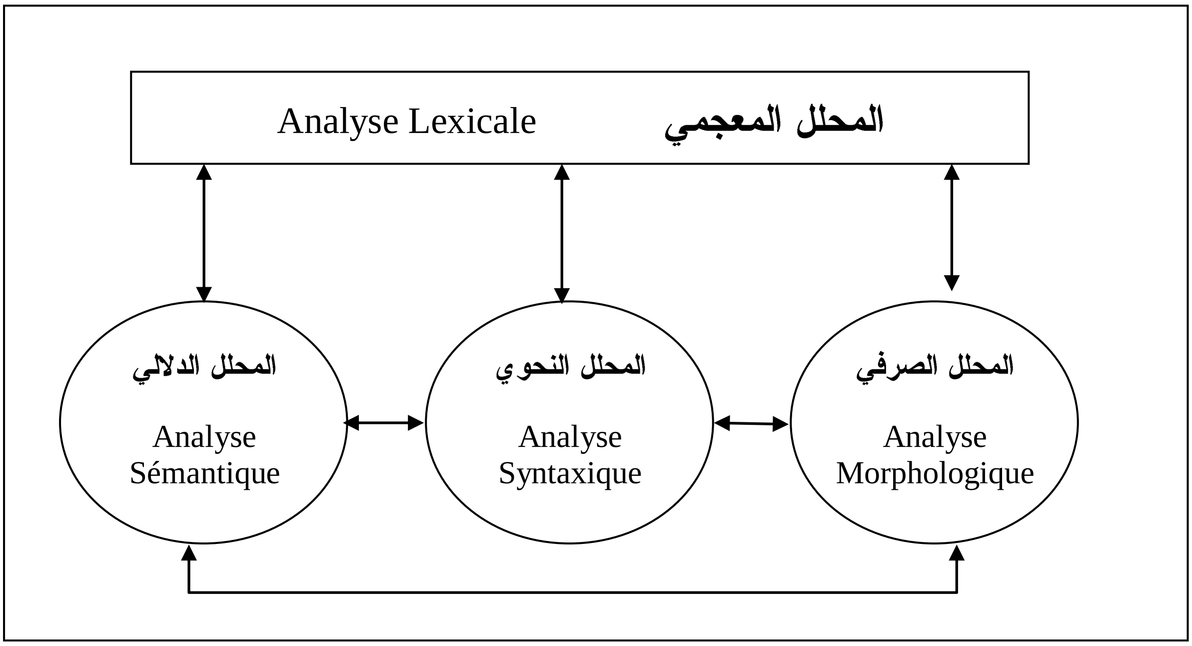
6 - Analyse Sémantique.

7 - Analyse lexical.

8- Analyse phonologique.

9- التحليل الصوتي: يستهدف التحليل الصوتي تجهيز المحتوى اللغوي للمعجم الحاسوبي للتطبيقات الصوتية التي يمكن أن تبنى عليه مثل تقنية تحويل النص المكتوب إلى كلام منطوق TTS، وتقنية التعرف الآلي على الصوت البشري SR ، ويأخذ التحليل الصوتي أكثر من صورة، وقد يشمل التحليل الصوتي مادة المعجم كلها، مداخل وشروحًا وأمثلة وشواهد، وقد يقتصر على بعضها حسب هدف المعجم.

ويبين الشكل رقم -1- العلاقة المترابطة بين هذه المستويات[1].



الشكل رقم09 : العلاقة بين مختلف المستويات اللغوية للمعالجة الآلية.

4- الوظائف الأساسية لنظم المعالجة الآلية للغة:

4-1- ثنائية التحليل والتركيب:

إن معظم التطبيقات في حقل المعالجة الآلية للغة تعتمد على التحليل والتركيب ومن ذلك الصرف، فاشتقاق الجذور هو تحليل أما تكوين الكلمات فهو تركيب، وفي بعض الأحيان يتداخل التحليل والتركيب لدرجة أنّنا نجد بعض النّظم تعتمد على أساليب التحليل بالتركيب والتركيب بالتحليل[2].

يمكن القول، بوجه عام، أن نظم المعالجة الآلية للغة على اختلاف مستوياتها تقوم بوظيفتين أساسيتين[3]:

- وظيفة تمييز العنصر اللّغوي (الشق التحليلي).
- وظيفة توليد العنصر اللّغوي (الشق التركيبي).

1- الزبيدي، فالح حسن عويد، "تطبيق النظرية التوليدية التحويلية على اللّغة العربية حاسوبيًا"، رسالة ماجستير، جامعة بغداد، بغداد، العراق، 1992، ص 14.

2- بربارة سهيلة،الترجمة بمساعدة الحاسوب من الانجليزية إلى العربية، سابق، ص 65، 66.

3- نبيل علي،اللّغة العربية والحاسوب، سابق، ص 187.

يوضح الجدول (1) الآتي تصنيفاً لبعض تطبيقات المعالجة الآلي للغة على ضوء الوظيفة الأساسية التي تقوم بها، (تحليلاً وترتيبًا) وعنصر اللّغة الأساسي الذي تتعامل معه (صوتًا، حرفًا- كلمة- جملة- سياقًا)[1].

إنّ المقصود من عملتي التحليل والتركيب هو الطبيعة العامة والسائدة للتطبيق اللّغوي ومن المناقشة[2]، فالتحليل والتركيب يكثر استخدامهما في حالات اللّبس اللّغوي المختلفة إذ على المحلل أن يفترض إحدى القراءات الممكنة للتعبير عن اللّبس كأساس لتحليله، ليقوم بعد إتمام التحليل بإعادة التركيب لمقارنة ما أعاد تركيبه مع الدخل الأصلي المغذي إليه، فإذا كان مطابقاً كان الافتراض صحيحًا، وإلاّ فعليه أن يعيد عملية التحليل مرّة أخرى بافتراض جديد[3].

التطبيق	الوظيفة الأساسية للنظام	
	تمييز (تحليل)	توليد (تركيب)
عنصر اللّغة	"**على مستوى عنصر التحليل النحوي**"	
الصوت	تمييز الكلام التعرف على المتكلم آليًا	- توليد الكلام آليًا - تغيير الكلام آليًا من صوت الذكور لصوت الإناث، أو العكس.
الحرف	تمييز الحروف آليًا (القراءة الآلية).	- توليد أنماط الحروف آليًا (طباعتها وإظهارها).
الكلمة	التحليل الصرفي الآلي	- التركيب الصرفي الآلي.
الجملة	التحليل النحوي الآلي	- التركيب النحوي الآلي.
السياق	الفهم الآلي للنصوص	- توليد النصوص آليًا.
عنصر اللّغة	"**على مستوى نظام التطبيق الآلي**"	
معالجة البيانات	- إدخال البيانات	- طباعة المستخرجات.
الترجمة الآلية	- تحليل تعبيرات لغة "المصدر"	- توليد التعبيرات المقابلة بلغة (الهدف).
معالجة النصوص	- اكتشاف الأخطاء الإملائية	- تصحيح الأخطاء الإملائية.
معالجة المعلومات	- تحليل طلب البحث	- الاستجابة لطلب البحث.
معالجة المعارف	- اكتساب المعارف من مصادرها	-استخدام آلة الاستدلال. المنطقي لتوليد الإجابات.

الجدول رقم 07: طرفا التحليل والتركيب لنظم المعالجة الآلية للغة[4].

1- المرجع السابق، ص 187.

2- المرجع نفسه، ص 187.

3- المرجع، نفسه، ص 180.

4- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 188.

بعد التحدث عن طبيعة الوظائف الأساسية التي يقوم بها نظم المعالجة الآلية، ننتقل الآن إلى تحديد الإطار العام للمعالجة "اللّغة العربيّة" آليًا، وذلك على ضوء مبدأين أساسيين[1]:

– المبدأ الأوّل: تحديد الإطار العام على مستوى معالجة المعلومات[2]، ومعالجة المعارف[3]. كون المهام اللّغوية للنظام الآلي تتفاوت بينهما إلى حدّ كبير.

– المبدأ الثاني: تحديد الإطار العام للغة الانجليزية أولاً، ثم بعد ذلك "اللّغة العربية" لتضح الفروق بينهما، وتبرز الإضافات المطلوبة لتضم معالجة اللّغة العربية آليًا بصفتها أعقد من نظم اللّغة الانجليزية.

ثانيا- تحديد الإطار العام لمعالجة المعلومات والمعارف:

2-1- معالجة المعلومات:

2-1-1- بالنسبة لـــلانجليزية:

إن لنظم استرجاع المعلومات مكانتها البارزة ضمن دراسات المعاني الآلية للغة الطبيعية وتعد بحق ذروة سنام التخصص، وتشير إلى الحقائق والمعارف التي تم التوصل إليها في دراسات المكتبات والمعلومات النظرية والتطبيقية على حد سواء، كالفهرسة الآلية، والتصنيف الآلي والتحليل الموضوعي والتكثيف والاستخلاص والببليوغرافيا[4] فالمجال مفتوح لتوليد دراسة معالجة المعلومات يكون ميدانها معالجة اللّغة العربية لمختلف الأغراض المتصلة بمفردات وصياغة اللّغات الطبيعية والاصطناعية المستخدمة في وصف وتحليل المعلومات واسترجاعها بدقة سواء أكانت النصوص حرة كما في العناوين والمستخلصات أم كانت اللّغة محكمة كما في المكانز ورؤوس الموضوعات[5].

إنّ الأساليب المختلفة لنظم استرجاع المعلومات النّصية منها: البتر وخوارزميات التجريد، والمحللات الصرفية تهدف إلى تحسين أداء النظام[6] فمعالجة المعلومات بالنسبة للغة الانجليزية يتكون من العناصر الرئيسية التالية[7]:

العنصر الأول: قاعدة المعلومات.

1– المرجع السابق، ص 189.

2 –Traitement de l'information.

-ينظر: معجم مصطلحات علوم الكمبيوتر (انجليزي، فرنسي، عربي) سابق، ص 222.

3 –Traitement de la Connaissance.

4– محمد سالم، غنيم نظم استرجاع المعلومات العربية، مظاهر الغموض وآفاق الحلول مكتبة الملك فهد الوطنية، ط1، الرياض، 2008، ص 42.

5– علي السليمان الصوينع، كشافات التباديل واسترجاع المعلومات في اللّغة العربية، مكتبة الملك فهد الوطنية، الرياض،السلسلة الثانية– 1– 1988، ص 11.

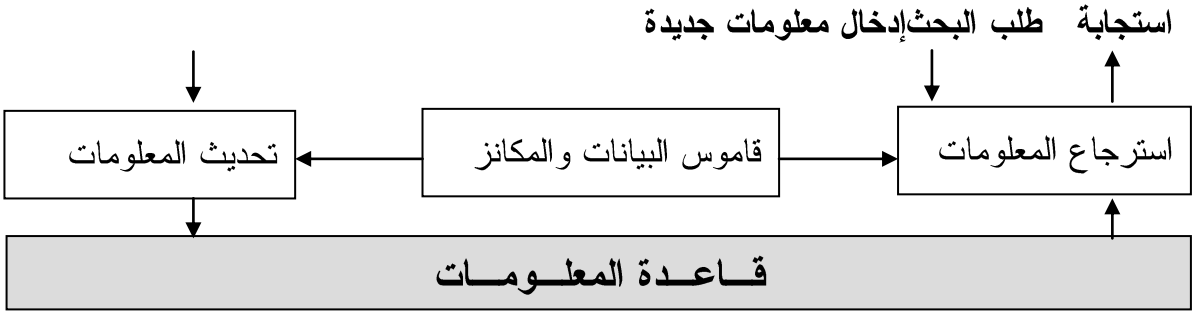
6– محمد سالم غنيم (نظم استرجاع المعلومات العربية) مظاهر الغموض وآفاق الحلول، سابق، ص 85.

7– نبيل علي، اللّغة العربية والحاسوب، سابق، ص 189– 190.

العنصر الثاني: وسيلة تحديث المعلومات.

العنصر الثالث: وسيلة استرجاع المعلومات.

العنصر الرابع: قاموس البيانات الّذي يتضمن تعريفات المصطلحات المستخدمة في قاعدة المعلومات والمكانز التي تتضمن قوائم بعناوين الموضوعات العامة والتفصيلية التي تم اختيار الكلمات المفتاحية منها.



الشكل رقم 10: معالجة المعلومات "نموذج اللّغة الانجليزية".

2-1-2-طريقة استخراج المعلومات:

يتم إدخال المعلومات الجديدة سواء باستخدام لوحات مفاتيح، أو بقراءة الوثائق آليًا من خلال الآلات لقارئة التي تقوم بتمييز النصوص المكتوبة تلقائيًا، وتشمل عملية إدخال المعلومات الجديدة على عمليات التحقق من صحة البيانات، واكتشاف وتصحيح الأخطاء الإملائية، وكذلك عملية تشفير[1] عناصر المعلومات بما يتفق والمواصفات المحددة في قاموس البيانات. تتلقى وسيلة استرجاع المعلومات على هيئة طلبات بحث عن معلومات معينة.

وتتم صياغة طلبات البحث عادة باستخدام الكلمات المفتاحية[2] التي يتم تحديدها بالرجوع إلى المكانز العامة والمتخصصة- بناء على طلب البحث تقوم وسيلة استرجاع المعلومات بمقارنة البيانات الواردة في طلب البحث مع ما تتضمنه المعلومات الإشارية للوثائق (كالعناوين، والكلمات المفتاحية، والمستخلصات) واستخراج قائمة بالوثائق المطابقة وفقًا للكلمات المفتاحية التي يتضمنها طلب البحث[3].

---

1 -Processus de cryptage: - مجموعة من الرموز والقواعد تستعمل لتمثيل المعلومات عند الإرسال أو التسجيل في الذاكرة.

2 -Mots clés.

3- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 190، للمزيد من التفاصيل انظر: ف. و. لانكستر و أ. ج. وورنر، أساسيات استرجاع المعلومات (نظم استرجاع المعلومات)، ترجمة حشمت قاسم، مطبوعات مكتبة الملك فهد الوطنية، السلسلة الثانية- 29- الرياض- 1418- 1997، ط3 مزيدة ومنقحة، ص341.

2-1-3- بالنسبة للّغة العربية:

يختلف الإطار العام لمعالجة المعلومات في اللّغة العربية (الشكل رقم 11) عن ذلك لمعالجتها في اللّغة الانجليزية[1] في ثلاث نقاط أساسية[2]:

– النقطة الأولى: من الطبيعي أن نفترض دخول المعلومات الجديدة وبيانات طلب البحث، بدون تشكيل أو بتشكيل ناقص، ولإزالة اللّبس النّاجم عن غياب التشكيل أو عدم اكتماله، يجب تزويد نظم معالجة المعلومات العربية بمعالج للتشكيل الآلي[3]. وذلك لاستكمال التشكيل في حالة "الدخل" وإضافة علامات التشكيل اللاّزمة في حالة "الحرج" وذلك لضمان القراءة الصحيحة للمستخرجات.

– النقطة الثانية: يلزم لعمل معالج التشكيل الآلي وجود معجم للجذور العربية والصيغ الصرفية المطبقة عليها، لأنّ المعالج يقوم بعمليات تحليل وتركيب لغوية مختلفة (صرفية ونحوية ودلالية يلزم لها كثير من المعطيات المعجمية، فلقد توصلت بعض الدراسات إلى أن الأسلوب الصرفي الدلالي قد حقق نسبة استدعاء **91%** لنظم استرجاع المعلومات، إذ قام د. الطيّار بالبحث عن مدخل مبتكر للبحث في نظم استرجاع المعلومات العربية ويسمى هذا المدخل الجديد **الأسلوب الصرفي الدلالي** الذي يعتمد على روابط دلالية للصيغ الصرفية محصورة في معجم للصيغ الصرفية[4].

1 –Kasem Hishmat M. A." Arabic In Specialist Information Systems": a study In Linguistic Aspects of information Transfer. Ph- D- thesis- Univ. Of London 1978. P 226.

– للمزيد من التفصيل ينظر: ف. و. لانكستر و أ. ج. وورنر. أساسيات استرجاع المعلومات (نظم استرجاع المعلومات ترجمة: حشمت قاسم.

– محمد سالم غنيم، نظم استرجاع المعلومات العربية، مظاهر الغموض وآفاق الحلول.

2– نبيل علي، اللّغة العربية والحاسوب، سابق، ص 190.

3– نظرًا لأنّ معظم الوثائق العربية لا تحتوي على تشكيل، قامت صخر بتطوير المشكل الآلي لمعالجة هذه المشكلة ووضع علامات التشكيل على النصوص غير المشكلة لمساعدة محرك النطق الآلي على نطق النّص العربي بصورة صحيحة.

يزيل التشكيل أي التباس في معاني الكلمات وهو يعتمد على أبحاث مكثفة شاملة في مجال معالجة اللّغة الطبيعية (TLN).

يعتمد التشكيل التلقائي على مستويات مختلفة في معالجة اللّغة وتحليلها بدءًا من "الصرف" وانتهاء بنك الالتباس في معاني الكلمات. ويتم ذلك بتوظيف أبحاث "معالجة اللّغة الطبيعية" (TLN) إضافة إلى قواعد البيانات اللّغوية الضخمة التي قامت شركة صخر بتطويرها على مدار أعوام عديدة يقدم محرك المشكل الآلي نصًا عربيًا مشكلاً حيث تصل دقة التشكيل إلى 98% ويسمح لك بتحديد مستوى التشكيل المطلوب: كامل أو لازم أو تشكيل نهايات الكلمات فقط. (الإعراب).

4– محمد سالم غنيم، نظم استرجاع المعلومات العربية، مظاهر الغموض وآفاق الحلول، سابق، ص 85– 86.

للمزيد من التفصيل:

–AL Tayyar, Musaid Salah. (Arabic Information Retrieval System Based on Morphological Analysis (AIRSMA): A comparative study of word, Stem, root and morpho- semantic methods- PH. D. thesis- De MONT FORT University, July. 2000.

2-1-4- اللّغة الطبيعية في استرجاع المعلومة:

هناك أربعة سبل محتملة لمعالجة اللّغات التي يمكن أن تستخدم في التعبير عن ناتج التعرف على المستوى الموضوعي لقاموس البيانات[1] والمكنز[2] وإجراء عمليات البحث في نظم الاسترجاع[3]:

1- التحكم في اللّغة في المدخلات والمخرجات ويستخدم هذا السبيل لغة مقيدة مسبقًا كالمكّنز التقليدي، أو قاموس البيانات.

2- عدم التحكم في أي نوع في المدخلات أو المخرجات: ويرتبط هذا السبيل بنظم الاسترجاع المعتمدة على اللّغة الطبيعية.

3- التحكم في اللّغة في المدخلات وعدم التحكم في مرحلة المخرجات حيث يمكن للمستعمل البحث استخدام كل ما يروق له من مصطلحات ويقوم الحاسوب "برسم خريطة" هذه المصطلحات المقيدة للنظام أو بأي وسيلة أخرى.

5- لا تحكم في مرحلة المدخلات مع تحكم نسبي في مرحلة المخرجات: ويستخدم هذا السبيل بالاعتماد على المكنز المخصص للبحث فقط ويكمن أن يسمى باللغة التي يتحكم فيها لاحقًا.

-النقطة الثالثة: يتم توسيع نطاق البحث في النصوص العربية بتوليد جميع المشتقات والتصريفات الممكنة ذات الصلة الدلالية بالكلمات المفتاحية المختارة، وذلك باستخدام معالج الصرف الآلي[4].

---

1 -Dictionnaire de données.

2- المكنز (Corpus): لقد عرف المكنز على صعيد اللّغات اللاّتينية تطورًا مدهشًا منذ بدايات الستينات وأصبح يتوفر على تقنيات هائلة في مجال تخزين المعلومات والنصوص والمفردات والمصطلحات واسترجاعها وفق معايير وأنساق منهجية، ولم يعد بالإمكان الاستغناء عما يقدمه من خدمات وحلول آنية في مجال اللّغة نحوها وصرفها ومعجمها، وتخزين مختلف نصوص المعرفة هذا بالإضافة إلى ركنه الأساسي في إعداد وتجهيز واختبار كل البرامج المبنية على تطبيقات معالجة اللّغات الطبيعية الإنسانية (TLN).

إنّ حجم المكنز اللّغوي له قدرة ضخمة وهائلة على تخزين المعلومات والنصوص والمفردات والمصطلحات ويمكن الإشارة هنا إلى أنّ المكنز بالحرف اللاّتيني في السبعينات تمكن من تخزين 470 مليارًا من الحرف وعلى هذا الأساس يمكن أن يستوعب مليون كتاب، أي بمعدل مائتي صفحة لكل كتاب، نشير في هذا المقام أن شركة صخر أنجزت المكنز العربي الذي بلغ حجمه في المرحلة الأولى 10 ملايين كلمة ثم وصل في المرحلة الثانية إلى 160 مليون كلمة، وقد تم تحديد البيانات الخاصة بها بواسطة برنامج المشكل الآلي تلقائيًا، حيث يقوم بإجراء عمليات فك اللّبس الصرفي والنحوي بالاعتماد على تقنيات متطورة ومن المفروض أن يصل حجمه إلى 500 مليون كلمة.

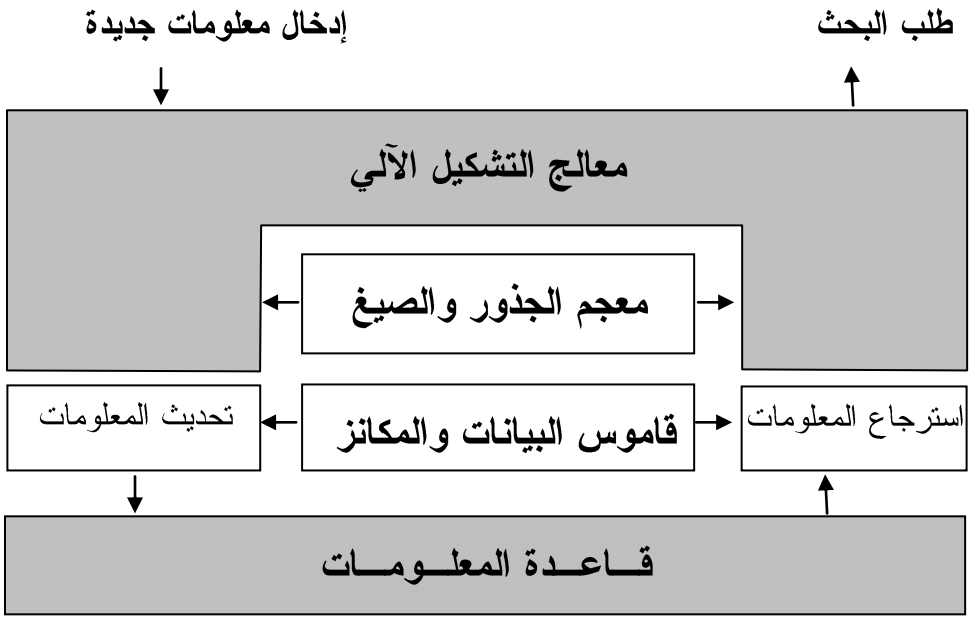
ومن أهداف وغايات المكنز نذكر أهم الإنجازات التي يقدمها للباحث من جهة وللبرامج والنظم الآلي التي يتم إعدادها من جهة أخرى: المحلل الصرفي الآلي- قاعدة البيانات المعجمية التي تستمد قوتها من المكنز- التشكيل والإعراب الآلي، المصحح الآلي- المصحح النحوي- المفهرس الآلي- الملخص الآلي- الرابط الآلي- البحث الموضوعي. القارئ الآلي OCR الإملاء العربي الآلي- فقد تركزت مهام المكنز العربي في تجهيز واختبار كل البرامج العربية الآلية السابقة الذكر. والمرتبطة بتطبيقات معالجة اللّغة الطبيعية.

3- ف. و لانكسر و أ. ج. وورنر (أساسيات استرجاع المعلومات) "نظم استرجاع المعلومات" سابق، ص 154.

4- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 191.

للمزيد من التفصيل حول هذا الموضوع؛ ينظر:

- Christian Jacquemin, "Traitement automatique de langues pour la recherche d'information" volume 49-n°2/2000. Hermes science Publication, Paris. P: 473.



الشكل رقم 11: معالجة المعلومات "نموذج اللّغة العربية".

2-2- معالجة المعارف:

يوضح الشكل -12- الإطار العام لمعالجة المعارف للغة الانجليزية؛ وهو يشمل العناصر الرئيسية[1]:

أ: قاعدة المعارف التي تشتمل على شبكات المفاهيم التي تمثل عناصر المعرفة (المفاهيم) والعلاقات التي تربط بينها.

ب: عناصر التحليل اللّغوي المختلفة (صرفي، نحوي، دلالي، سياقي منطقي)، والتي تقوم بتحليل المعلومات المدخلة للنظام لغويًا، كتمهيد لاستخلاص المفاهيم التي تتضمنها النصوص، ولا يعني التسلسل الموضح في الشكل بين عناصر التحليل اللّغوي المختلفة أنّها تعمل بصورة متتالية، إذ أن جميع عناصر التحليل هذه تعمل عادة بشكل متكامل يغلب عليه الطابع المتوازي أو شبه المتوازي.

جـــ: آلة الاستدلال[2] (أو الاستنتاج) المنطقي، وتقوم بمسح مضمون قاعدة المعارف بحثًا عن إجابات الأسئلة أو حل الأسئلة الذي تطرح على نظام معالجة المعارف على أساس البحث في قواعد المعارف وإيجاد الحل للمسألة المطروحة[3].

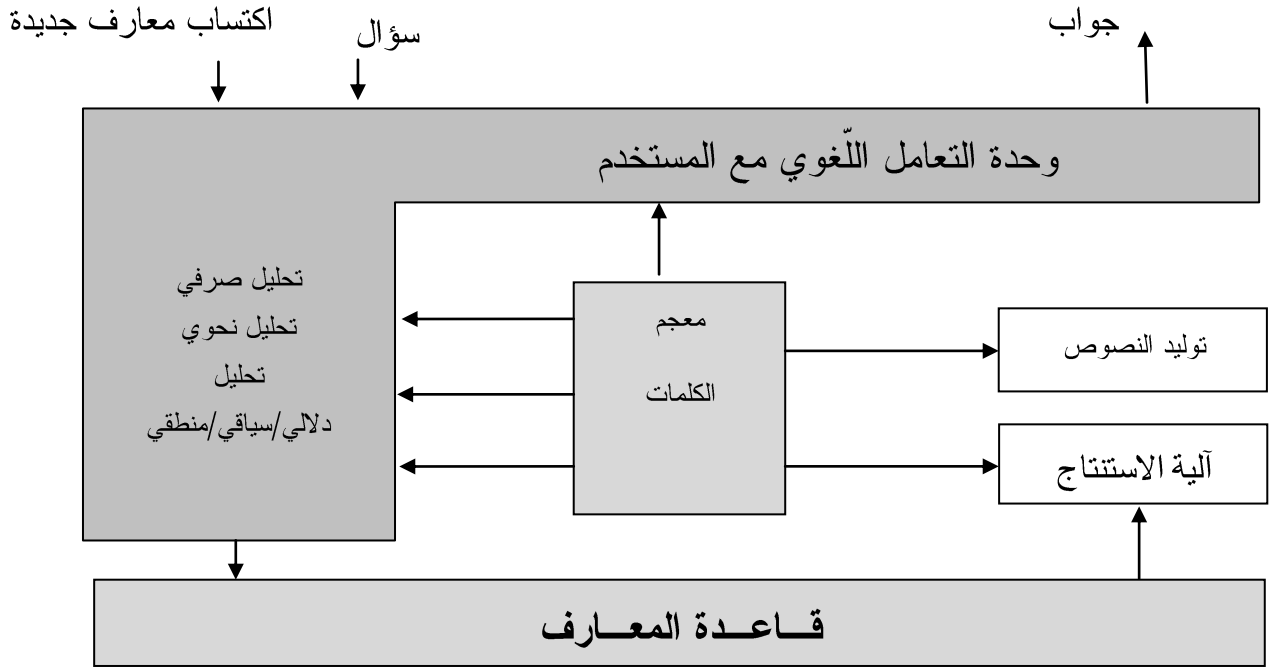
1- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 192.

2 -Inference machine.

3- عز الدين غازي، مفهوم قواعد المعارف، موقع الويب: http://www.Ahewar.org

د: وحدة توليد النصوص آليًا لتحويل سلسلة خطوات الإجابة أو الحل والتي خلصت إليها آلة الاستدلال في صياغة لغوية يفهمها المستخدم.

هـــ: معجم الكلمات الذي يمد جميع عناصر التحليل والتوليد اللّغويين بالمعطيات المعجمية[1] اللاّزمة لعملها.



الشكل رقم 12: معالجة المعارف "نموذج اللّغة الانجليزية[2].

يختلف الإطار العام لمعالجة المعارف باللّغة العربية (الشكل رقم 13) في إطار معالجتها في الانجليزية في أربع نقاط[3]:

1- النقطة الأولى: إضافة معالج التشكيل الآلي كخط مواجهة يفصل بين النظام ومصادر (الدخل) و(الخرج) حتى يتكمل التشكيل الناقص في حالة "الدخل" أو يضيف التشكيل اللاّزم في حالة "الخرج".

2- النقطة الثانية: استخدام عمليتي التحليل الصرفي والنحوي في إجراءات إزالة اللّبس الناجم عن غياب التشكيل أو نقصانه.

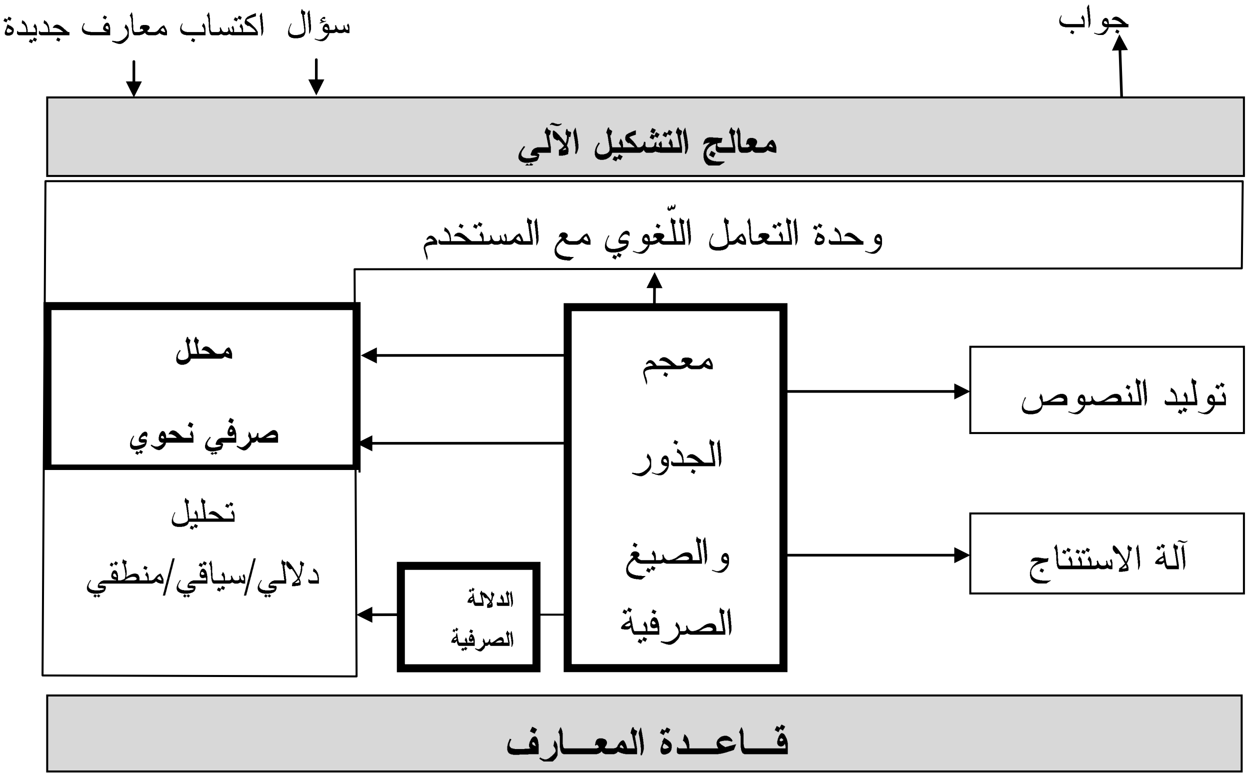
1- المعطيات المعجمية: أو قاعدة بيانات المعجمية وتتجلى أهميتها في كونها مهيأة و مجهزة لتحليل الأنماط المعجمية وقدرتها على الإجابة على كل التساؤلات المتعلقة بالمفردات داخل سياقاتها الدلالية والتركيبية وهو ما يجسد أهمية قاعدة المعطيات النصية الخاضعة لمعالجة آلية وفق منهجية دقيقة ومكثفة وذات مردودية عالية.

2- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 192.

3- المرجع نفسه، ص 193.

3- النقطة الثالثة: استخدام معجم الجذور[1] والصّيغ الصرفية بدلاً من معجم الكلمات لتسهيل عمليات التحليل اللّغوي المختلفة.

1- النقطة الرابعة: وجود عنصر الدلالة الصرفية كحلقة وصل بين المعجم وعنصر التحليل الدلالي، إذ تلعب المعاني الصرفية دورًا مهمًا في الفهم الآلي للتعبيرات العربية.



<u>الشكل رقم 13</u>: معالجة المعارف "نموذج اللّغة العربية".

3-2- المقارنة بين معالجة "العربية" و"الانجليزية" آليًا:

نلخص هنا الفروق الأساسية بين اللّغتين العربية والانجليزية على مستوى الإطار العام لمعالجة المنظومة اللّغوية الشاملة[2]:

1- تفترض النظم الانجليزية علاقة "واحد إلى واحد" في كثير من أمورها بينما تتسم "العربية" ببروز التعدد اللّغوي فالحرف له أكثر من شكل، والكلمة لها أكثر من تشكيل، والمعنى الواحد له أكثر من مرادف، والفعل له أكثر من باب، والباب النّحوي له أكثر من صيغة صرفية، وهكذا...ومن الجلي أن هذا يضيف عبئًا إضافيًا على جميع عناصر التحليل والتركيب اللّغويين.

1- فاطمة الخلوفي، تمكين اللغات، التكنولوجيا والإتصال، اشراف عبد القادر الفاسي الفهري، منشورات معهد الدراسات والأبحاث للتعريب، الرباط ، المغرب، المجلد2، أكتوبر 1997، ص 20 و ما بعدها.

2- اللّغة العربية والحاسوب، سابق، ص 193- 194.

2- يطغى على النظم الانجليزية الطابع التركيبي، إذ يهيمن النّحو على عملية معالجتها آليًا، في حين يطغى الصرف على "العربية" إذ يعدّ بحق محور النظم الآلية للتحليل والتركيب اللّغويين وحلقة وصل مهمة بين المعجم وعنصر التحليل الدلالي[1].

3- يجب أن تتميّز النظم الآلية لمعالجة اللّغة العربية بقدرات هائلة لإزالة اللّبس[2] الناجم عن غياب التشكيل أو نقصانه. لذا يجد معالج التشكيل الآلي عنصرًا أساسيًا يجب أن يسبق وظائف التحليل اللّغوي، في نفس الوقت الذي يحتاج هذا المعالج دالة إلى عمليات تحليل لغوي متعددة. ومن هنا تتضح العلاقة الحلقية في معالجة اللّغة العربية آليًا وهذه العلاقة التي لا ينفع معها أسلوب التحليل الهرمي متعدد المستويات، مما يوحي باقتراح أسلوب "شبكي" تتصل فيه معظم العناصر فيما بينها لتتضافر وتتآزر بعضها مع بعض لفض اللّبس الناجم عن غياب التشكيل كليًا أو جزئيًا، وهذا الّذي خلصنا إليه كنتيجة منطقية لشدة تماسك منظومة اللّغة العربية.

4-نظرًا لاختلاف الجوهري في تنظيم المعجم العربي (على أساس الجذر) والمعجم الانجليزي، فإنّ أسلوب معالجته آليًا (ميكنته) ودمجه في نظم المعالجة الآلية يختلف اختلافًا شاسعًا عمَّا يتم على مستوى اللّغة الانجليزية[3].

ثالثا: المعالجة الآلية للصرف في اللغة العربية:

إن توظيف المعالجة الآلية في معالجة الصرف العربي هو تذليل العقبات أمام المستخدم للوصول إلى ضالته المنشودة في ساحة معالجة اللّغة العربية آليًا مع زيادة سرعة الوصول، وكفاءة الأداء الّذي يتصف به التطبيق البرمجي، مع توظيف الموارد اللّسانية الصرفية كأداة فاعلة صوب إنشاء قاعدة معطيات صرفية لبناء برامج حاسوبية في ميدان حوسبة الصرف العربي، واعتمادها كإحدى الأدوات الفاعلة لتطوير برامج آلية للمعالجة الآلية للغة العربية.

1- ينظر:

Aida KHEMAKHEM (ARABIC LDB: Une base lexicale normalisée pour la langue- arabe Mémoire master en (système d'information et nouvelles Technologies) université de Sfax. Faculté des sciences Economiques et de gestion. 2005- 2006- P-21- 33.

2- اللّبس: Ambiguïté: لا يختلف معنى هذا المصطلح تأثيلياً في اللّسانيات عنه في الاستعمال اللّغوي- فالغموض "ما يحتمل فيه الكلام معنيين أو أكثر" ينطبق الغموض حصريًا على كل تأليف تعبيري (مركب أو جملة)؛ أما في حال الوحدة المفرداتية المتضمنة لمعان عدّة فإننا نتحدث عن "التعدد الدلالي". يؤدي الغموض دورًا مهمًا داخل حقل اللّسانيات، كونه يعد من معطيات الوصف اللّساني، يمكن للملفوظ أن يكون غامضًا على الصعيد "التركيبي" أو على الصعيد "الدلالي"، لعل الغموضات التركيبية هي الأكثر عددًا؛ ذلك راجع لإمكانية تقديم تحليلات تركيبية متعددة للجملة، وفي كل الأحوال يتضاعف الغموض التركيبي في شكل غموض دلالي. نشير إلى أنّه ثمة عدد من علماء اللّسان ممّن يعرفون مصطلح "الغموض" بالنمط التركيبي، الّذي كنّا قد أتينا على بيانه سلفًا دون غيره.

للمزيد من التفصيل ينظر:

-FUCHS Catherine ,Les ambiguïtés du Français. Paris Ophrys. 1996.

3- اللّغة العربية والحاسوب، سابق، ص 194.

وقد أظهرت عملية التحرّي أنّ هذا الأسلوب في المعالجة الآلية يمتلك خصائص إيجابية تتجلى في سرعة وكفاءة البرامج التي توظفه كأداة ناجعة للبحث تمكّن من الوصول إلى جملة من الإيجابيات والارتقاء بهذه التقنية "المعالجة الآلية للغة العربية" إلى مستويات أداء أفضل، وذلك عن طريق إنتاج برامج حاسوبية لإنشاء معاجم الكترونية للغة العربية.

ولكي تتجلى لنا أمامنا الآفاق المتاحة للمعالجة الآلية لنظام اللّغة العربية، سنحاول التعرف على نظام من أنظمة لغتنا العربية والإمكانيات المتاحة لإقامة علاقة متبادلة مع إمكانية الحاسوب، تجعل من بنية النظام الصرفي العربي نظامًا متوافقًا مع الآليات التي يعمل بها الحاسوب، وقابلاً للمعالجة التي تستلزمها التطبيقات الحاسوبية.

3-1- استخدام النظام الصرفي العربي[1] كأنموذج للمعالجة الآلية في الدراسة التطبيقية:

إنّ الصرف بلا شك هو مصدر التوسع اللّغوي بما يوفره من وسائل عديدة لتكوين كلمات جديدة، وإعادة تحليل تلك القائمة بالفعل يمثل الصرف والصرف العربي خاصة، وضعًا مثاليًا لإبراز ثنائية التحليل والتركيب وهي الثنائية التي يوصف من خلالها كثير من الظواهر اللّغوية، في نفس الوقت الّذي تعدّ فيه أحد المفاهيم الأساسية في تصميم المعالجة الآلية للغات[2].

فالصرف يعدّ مدخلاً أساسيًا لوصف النظام الشامل لهذه اللّغات وتفسير الكثير من ظواهره وتحديد أسلوب معالجته آليًا[3].

إنَ معالجة نظام الصرف العربي آليًا[4] تعدُّ الخطوة المنطقية لمعالجة اللّغة العربية آليًا حيث يرتقي هذا النظام الآلي من بدائية التعامل مع الحروف إلى معالجة الكلمات، وهي العملية التي

1- في المستوى الصرفي، فقد تمت المعالجة الآلية له في ضوء أهمية الصرف العربي بالنسبة لنظام اللغة ككل. وقد تناولت هذه المعالجة الآلية بعض جوانب الصرف العربي، كالخاصية الثلاثية للجذور العربية، وأصل الاشتقاق، والأنماط الصرفية، وثنائية الصيغة الصرفية والميزان الصرفي، والإنتاجية الصرفية، والفائض الصرفي، واللبس الصرفي. للمزيد من التفصيل ينظر: نبيل علي،اللغة العربية والحاسوب، ص247-296.

2- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 247.

3- المرجع نفسه، ص 248.

4- في هذا الصدد قدَّم الدكتور نبيل علي إطاراً عاماً لمعالجة الصرف العربي آلياً، وأورد عدة نماذج للتحليل الصرفي الآلي بوجه عام، مبيناً مدى ملائمتها لمطالب الصرف العربي. وتبعاً لهذا قام بعرض نموذج وضعه لمعالجة الكلمات العربية صرفياً في أطوار التشكيل المختلفة، وهو نموذج التحليل بالتركيب، وأطلق عليه اسم (المعالج الصرفي المتعدد الأطوار). ويشتمل هذا النموذج المبتكر على عناصر أربعة هي : المعالج الصرف – نحوي، والمعالج الاشتقاقي، والمعالج الإعرابي، ومعالج التشكيل. ينظر: ينظر: المصدر السابق، ص301-314، ص181-182 .

تعدّ طلبًا أساسيًا لميكنة المنظومات اللّغوية الأخرى، ونقصد بها منظومات النّحو والدلالة والمعجم[1] عن طريق القليل الصرفي للكلمات إلى عناصرها الأولية الاشتقاقية والصرفية.

فمهمة المعالجة الآلية للصرف العربي هو قيام النظام الحاسوبي باستخلاص العناصر الأولية لبنية الكلمة وتحديد سماتها العرضية والصرف–نحوية[2].

ولعل من أهم الخصائص والتي تدعو للاستعانة بالمعالجة الصرفية في بناء نظم معالجة اللغة العربية أن 80% من كلمات اللغة العربية تم إنتاجها من جذور ثلاثية أو رباعية أو خماسية . كما أن هذه الكلمات يتم تكوينها بإضافة سوابق أو لواحق على الجذر مع تطبيق عدد من القواعد الصرفية وباستخدام الأنماط الصرفية، وقد يتم في هذه العملية حذف أو تعديل بعض حروف الجذور .

تجدر الإشارة إلى أنّ جميع الأبحاث اللّسانية الحاسوبية اعتبرت أن الصرف فرع رئيسي لأن تبنى عليه جميع الدراسات والمستويات اللّسانية، إذ سنخص هذا الجانب من الدراسة الصرفية لمفهوم الصّيغ والأوزان الصرفية والجذر وما يتصل به من زوائد (سوابق ولواحق) ولكن بطريقة حديثة في التناول، على أساس مقاربة العديد من قضايا الصرف في اللّغة العربية[3]، وذلك لمعالجتها آليًا عن طريق بناء قواعد معطيات نصيّة وقواعد معطيات معجمية في إطار المشروع المراد تحقيقه لبناء معجم الكتروني للمعالجة الآلية لغة العربية[4].

فأهمية هذا المبحث من الدراسة تؤدي بنا إلى طرح الإشكال التالي: كيف يمكن استثمار منظومة الصرف العربي في معالجة المفردة العربية عن طريق استخلاص العناصر الأولية لبنية الكلمة العربية ؟ ومباشرة تحديد سماتها الصرفية والنحوية؟

ملاحظة: قد تأتي مصطلحات بدلالات متقاربة، ولكن لكل مصطلح مقام خاص به، تفرقه عن غيره استعمالات معينة.

ففي هذا المبحث ترد علينا مصطلحات بمعانٍ متداخلة يتحتم علينا تحديد فروقها المعجمية والاصطلاحية ليسهل علينا توظيفها بطريقة علمية خاصة، ونحن نعلم أنه سيتكرر استعمالها في

1– اللغة والحاسوب، سابق، ص 248.

2– نبيل علي، آفاق النشر الالكتروني، المنظور اللّغوي مقال منشور في مجلة العربي– الكويت.

عنوان الموقع: آفاق النشر الالكتروني، موقع الويب: http://www.alarabimag .com

3– ينظر: عز الدين غازي (الصرافة الحاسوبية العربية)، محاولة في التأصيل، – عنوان الموقع: http://www.ahewar.org

4– سيأتي الحديث عن ذلك مفصلاً في حينه من خلال الدراسة التطبيقية للمشروع.

حالات كثيرة، لذا رأينا توضيحها لازمًا ليتسنى لنا أن نُفرِّق بين مختلف المفاهيم التي تحملها هذه المصطلحات منها: الصرف، الصيغة والوزن، الأصل، والجذر، السوابق واللّواحق...الخ.

3-2- دراسة البنية الصرفية في ضوء اللسانيات الحديثة :

الصرف[1] هو فرع من اللّسانيات الّذي يتعامل مع البنية الداخلية للكلمات من حيث تكوين عناصرها الأولية؛ والتفاعل بين هذه العناصر، وللصرف دور رئيسي في تحديد العلاقات التي تربط بين مفردات المعجم، والتي تحكم بالتالي تنويبه وتنظيمه، إن مهمة الصرف داخل المنظومة اللّغوية لا تقتصر على الأمور المتعلقة "بشكل" الكلمات، بل تتعداها إلى دراسة الصلة بين مباني الكلمات ومعانيها وخصائصها النّحوية والفونولوجية[2].

3-3- الصرف والتصريف في الدرس اللساني :

3-3-1- ضبط المصطلح في اللغة والإصطلاح:

التقى مصطلحا الصرف والتصريف في الدَّال اللغوي المبني على مفهومي التغيير والتحويل، فنجد أن المعجم اللغوي[3] يكاد يتفق في بيان معنى التصريف، بأنه التغيير والتحويل والإنتقال[4]. فهو من صرّف بمعنى غيّر، فتقول: "صرّفت الشيء بمعنى غيرته". و"صرّف الشيءَ: أعملَه في غير وجه كأنه يصرفه عن وجه إلى وجه". فالتصريف إذن تحويل الأصل الواحد أي تغيره. ويشترك معه في هذا المعنى مصطلح الصرف الذي هو رد الشيء عن وجهه. يقال: "صَرَفَه يصرْفه صرفا، وصَرَفَ الشيءَ: أعمله في غير وجهه"، كأنه يصرفه من وجه إلى وجه و"الصَّرْفُ: ردُّ الشيء عن وجهه"، صرفه يصرفه صرفًا فانصرف، "**<u>فالصرف أن تصرف الفعل الثاني عن معنى الفعل الأول</u>**[5] "

وصارف الشيء: صرفها عنه[6]، وقوله تعالى:﴿ وإذا ما أُنْزِلَتْ سُورةٌ نَّظَر بَعْضُهُم إلى بعضٍ هَلْ يَراكُمْ من أحدٍ ثم انصرَفُوا صرف الله قلوبهم بأنّهم قوم لا يفقهون﴾[7] أي رجعوا عن

1- الصرف: La Morphologie: علم يبحث في صيغ الكلمات واشتقاقها؛ يبحث في اللّفظ المفرد من حيث بناؤه ووزنه وما طرأ على هيكله من نقصان أو زيادة. سيأتي الحديث بالتفصيل في حينه من خلال هذا البحث، للمزيد من بسط القول: ينظر: محمد رشاد الحمزاوي:"المصطلحات اللّغوية الحديثة في اللّغة العربية"، معجم عربي أعجمي وأعجمي عربي- الدار التونسية للنشر- المؤسسة الوطنية للكتاب- د ط. 1987- تونس- الجزائر- ص 131، وينظر : - محمد سمير نجيب اللّبدي "معجم المصطلحات النحوية والصرفية"، سابق، ص 125.

2- اللّغة العربية والحاسوب، سابق، ص 247.

3- ابن منظور، لسان العرب، ط3(دار إحياء التراث العربي، بيروت، 1999م) (مادة صرف) 189/9، الفيروز أبادي، القاموس المحيط، ط(عالم الكتب، بيروت، د.ت)مادة صرف، 513/3 .

4- ابن منظور، لسان العرب، نفسه،[ص ر ف] المجلد 9، ص 189 وما بعدها.

5- عصام نور الدين، المصطلح الصرفي، مميزات التذكير والتأنيث، دار الكتاب العالمي، بيروت، الطبعة الأولى، 1409هـ ،1988 مــ ، ص 47.

6- لسان العرب، [ص ر ف]المجلد 9، ص 189، سابق.

7- سورة التوبة، الآية: 127.

المكان الّذي استمعوا فيه، وقيل انصرفوا عن العمل بشيء مما سمعوا، صرف الله قلوبهم أي أضلهم مجازاةً على فعلهم، وصَرفتُ الرجل عنّي فانصرف. و"الصَّرْفُ: التَّقلُّب والحيلة"[1].

فعلم الصرف هو أصول وقواعد تعرف بها أحوال أبنية الكلمة صيغها الأصلية والعارضة وما يلابسها من تغير معنوي في مدلولها[2].ولما كان الصرف يهتم بدراسة بنية الكلمة في ضوء ما يطلق عليه المورفيم الذي هو أصغر وحدة حرفية ذات معنى[3] فهو بذلك يبحث في الكلمة المفردة من حيث البناء والوزن وما يطرأ عليها من تغيرات بالزيادة أو النقصان[4]، وفي اللّغة الفرنسية يعرف علم الصرف "بأنّه جزء من علم القواعد الّذي يدرس شكل الكلمات والتحولات التي تطرأ عليها"[5].ويعرّف جون ديبوا "Jean Dubois" علم الصرف من وجهة اللّسانيات الحديثة بأنّه يأخذ مفهومين[6]:

وصف القواعد البنيوية الداخلية للكلمات أي قواعد تركيب وجمل المورفيمات Morphème لتشكيل الكلمات ودراسة الإضافات التي تلحق الكلمات في أوّلها وآخرها وكذلك دراسة مختلف التحولات التي تطرأ على الجذر حسب صيغ الجمع والمفرد، النوع، الوقت وحسب الحالة (الإعراب الاسمي والفعلي).

وصف القواعد التي تشكل البنية الداخلية للكلمات وطرق التركيب النحوي للجمل وهي هنا تتداخل مع تركيب الكلمات وإعرابها ومع بناء المعجم والصوتيات وفي هذه الحالة يسمى علم الصرف النحوي[7].

1- عصام نور الدين، المصطلح الصرفي، مميزات التذكير والتأنيث، دار الكتاب العالمي، بيروت، الطبعة الأولى، 1409هـ ،1988 مـ ، ص 47.

2- فخر الدين قباوة،"تصريف الأسماء والأفعال"، جامعة حلب، كلية الآداب سوريا، الطبعة الثانية، 1401/1981، ص 13.

3- حازم علي كمال الدين، نظرية القوالب من نظريات علم اللغة الحديث، مكتبة الآداب، القاهرة، دط، ص 15، وينظر: محمود سليمان ياقوت، معاجم الموضوعات في ضوء علم اللّغة الحديث، سابق، ص 253.

4- ينظر: مباني التقسيم، تمام حسان (اللّغة العربية معناها ومبناها)، سابق، ص 83. وينظر:محمد سالم محيسن، تصريف الأفعال والأسماء في ضوء أساليب القرآن، دار الكتاب العربي، بيروت، الطبعة الأولى، 1407ه-1987م، ص17.

5 – Le petit Larousse-France: maison d'édition, 2006- P: 707.

6 –Jean Dubois- La morphologie- In: Dictionnaire de linguistique et des Sciences de langages ; maison d'édition Larousse, 2001, P: 311.

7 –Morphosyntaxe.

فمدلول مصطلح علم الصرف هو دراسة الكلمة المفردة[1] في حالة انعزالها عن التركيب، بمعنى أنّه يتعامل مع الكلمة وبنيتها عن طريق تحليلها إلى عناصر الصرفية المختلفة[2].

وورد مصطلح التصريف عند القدماء "ولكنهم لم يوضحوا معناه توضيحا كافيا، فذكر سيبويه التصريف في قوله:" هذا باب ما بنت العرب من الأسماء والصفات، والأفعال غير المعتلة، والمعتلة، وما قيس من المعتلّ الذي لا يتكلمون به، ولم يجئ في كلامهم إلا نظيره من غير بابه، وهو الذي يسميه النحويون: التصريف والفعل"[3]، فمصطلح التصريف عنده يقوم على الدّال اللغوي المبني على مفهوم التغيير والتحويل؛ وذلك بتغيير الكلمة من وزن إلى وزن آخر، سواء أكان ذلك من المعتل أم من غير المعتل، على نسق كلام العرب الذي تكلموا به في غير باب المعتل أو غير المعتل[4]، وما معرفة قوانين البدل والحذف والقلب إلاّ لِتُعين على مسائل التصريف، وإلاّ فهي ليست تصريفاً[5]، وهكذا ظهر عنده مصطلح التصريف على أنه جزء من أجزاء النحو بمعناه الواسع، لم يحدّه بتعريف جامع مانع[6]، وأقامه على مفهومي التغيير والتحويل.

أمّا ابن السّراج فينص على مفهوم التغيير[7] في إيراده مصطلح التصريف دون الصرف، يقول هو "ما عرض في أصول الكلام وذواتها من التغيير"[8]، وذكر أن هذا " سمّي تصريفًا لتصريف الكلمة الواحدة بأبنية مختلفة، وخصّوا به ما عرض في أصول الكلام، وذواتها من

1- عبده الراجحي، التطبيق الصرفي، سابق، دار المعرفة الجامعية، مصر، د.ط، د.ت، ص7.

2- مجدي إبراهيم محمد إبراهيم، بحوث ودراسات في علم اللّغة، "الصرف. المعاجم. الدلالة". مكتبة النهضة المصرية- القاهرة، د ط، 2004، ص 23.

3 – سيبويه، الكتاب، تحقيق، عبد السلام هارون، ط1(دار الجيل، بيروت، د.ت) 242/4. وسيبويه في هذا الباب يفْصل بين ما بنته العرب من الأسماء والصفات والأفعال، وما قيس عليه. أما ما بنته فلم يذكر تعلّق الصرف به، وذكر فيه ما أحصاه من أبنية هذه الأقسام. وأما المقيس عليه مما لم يجئ في كلامهم إلا نظيره من غير المعتل والمضعف فهو الذي سماه النحويون التصريف والفعل، كما قال سيبويه، وهذا ما كان الصرف يتعلق به آنذاك.ينظر: جهود ابن جني في الصرف وتقويمها في ضوء علم اللغة الحديث، غنيم الينبعاوي، ط1(المكتبة التجارية، مكة المكرمة، 1995م) ص7، 8.

4- يُقاس الصحيح على وزن للمعتل لم يأت الصحيح عليه، والعكس أيضًا، وهذا يكون في مسائل التمارين والتدريبات؛ لترويض قوانين البدل والقلب والحذف، ومعرفة الأبنية، والميزان الصرفي.

5- مهدي بن علي بن مهدي آل ملحان القرني،" الترتيب الصرفي في المؤلفات النحوية والصرفية إلى أواخر القرن العاشر الهجري"، مجلة جامعة أم القرى، ج13، ع21، ديسمبر 2000م، ص737.

6- وهو ما سار عليه سيبويه في الكتاب، فلم يحدّ أي مصطلح من المصطلحات التي وردت في الكتاب بتعريف جامع مانع، للمزيد من بسط القول ينظر : خديجة بن عبد الرزاق الحديثي، أبنية الصرف في كتاب سبويه، منشورات مكتبة النهضة، بغداد، الطبعة الأولى، 1385هـ/1965م، ص 23.

7- وينص على أن مدلول مصطلح التصريف هو تغيير يحدث في أصل الكلمة وهيئتها، وهذا الفهم مستقى من دلالة المصطلح اللغوية، فضلا عن الاصطلاحية، وكذلك لا يختلف كثيرًا عما فُهم من قول سيبويه والمبرد، غير أنه يجعل التصريف خمسة أقسام: زيادة، وإبدال، وحذفٌ، وتغيير بالحركة والسكون، وإدغام. ينظر: ابن السرّاج،الأصول في النحو،231/3.

8- أبو بكر محمد بن سهل بن السرّاج، الأصول في النحو، تحقيق، عبد الحسين الفتلي،ط3(مؤسسة الرسالة، بيروت، 1988م) 231/3.

التغيير، وهو ينقسم خمسة أقسام: زيادةٌ، وإبدال، وحذفٌ، وتغييرٌ بالحركة والسكون، وإدغامٌ؛ وله حدٌّ يعرف به[1]".

وكان السيرافي تابعا لسيبويه، فقد أورد مصطلح التصريف في شرحه، مبينا أن مراد سيبويه من التصريف[2] هو التغيير والتحويل وذلك بتغيير الكلمة بالحركات والزيادات والقلب للحروف التي رسمنا جوازها، حتى تصير على مثال كلمة أخرى.

نجد هذا التداخل في إيراد المصطلحات المبني على الدال اللغوي عند ابن يعيش فقد ورد مصطلح التصريف عنده في مواضع متفرقة[3]، منها قوله: إن "التصريف مصدر، وُضِع كالعَلَم على هذا العِلْم؛ للفَرْق، خصُّوا به ما عرض في أصول الكلم وذواتها من التغيير، كاختصاصهم عِلْم العربية بالنحو"، ويضيف أن التصريف:"كلام على ذوات الكلم وفعله: صرّفته أصرِّفه تصريفا. يقال: صرَّفتُه فتصرَّف، أي: طاوع وقبل التصريف"[4]، مطرد يشمل جميع أنواع التغيير[5] **الذي يطرأ على بنية الكلمة .** ، ويتداخل هذا المصطلح عنده مع الاشتقاق، وذلك عند يذكر أن التصريف "اشتقاقه من تصريف الحديث والكلام، وهو تغييره بحمله على غير الظاهر. ومنه تصريف الرياح، وهو تحويلها من حال إلى حال"[6].

1- ابن السرّاج،الأصول في النحو،231/3.

2 – السيرافي النحوي في ضوء شرحه لكتاب سيبويه، تحقيق، عبد المنعم فائز، ط1(دار الفكر، دمشق، 1983م) ص592.

3- أما مصطلح التصريف في شرح المفصل ط(دار الكتب، بيروت) فقد ورد في المواضع التالية: في قوله:" وقالوا في النسب إلى أمية وسنقف عليه في التصريف"10/6. وقوله:" فأما قلب الواو ياء فسيذكر في موضعه من التصريف" 35/5. ويقصد من هذين الموضعين علم التصريف المعروف. وقوله:" أقوم ونقوم وتقوم يستوي فيها ضمير المخاطب والمتكلم والغائب في الاستتار وعدم الظهور000 لأن تصريف الفعل وما في أوله من حروف المضارعة يدل على المعنى"109/3. ويقصد هنا الصيغ . وقوله:" والتصريف يوجب أن الواو إذا تحركت وانفتح ما قبلها قلبت ألفا" 146/5. وقوله:" وهذا القسم الرابع آخر أقسام الكتاب إذ كان مشتملا على نكت هذا العلم وتصريفه" 9/ 53. وهنا يقصد القواعد التصريفية. وقوله:"وكذلك الواو لا تزاد أولا في حكم التصريف"155/9 وهنا كذلك يقصد قواعد التصريف. وقوله:"فإن الأمومة حكاها ثعلب والتصريف الفاسد ما لا يدفع عنه" 5/10. وهنا يقصد القواعد التصريفية والقياس عليها. وفي عموم المصطلح لم يستخدم المصطلح فيما يدل على الرياضة العقلية عند غيره بل أطلقه على التغيرات التي تلحق كلم العربية المستخدم فقط. وكان يستخدم في غير موضع من شرح المفصل مصطلح الاشتقاق للدلالة على التغيرات الصرفية، نحو قوله:"وأما العدل فهو اشتقاق اسم من اسم على طريق التغيير له، نحو اشتقاق"عمر" من "عامر" والمشتق فرع على المشتق منه، والفرق بين العدل والاشتقاق الذي ليس بعدل، أن الاشتقاق يكون لمعنى آخر أخذ من الأول كضارب من الضرب والعدل هو أن تريد لفظا ثم تعدل عنه إلى لفظ آخر، فيكون المسموع لفظا والمراد غيره، ولا يكون العدل في المعنى وإنما في اللفظ" ابن يعيش، شرح المفصل، 61/1 ــ 62. ومن هذا النص وغيره ندرك أن الاشتقاق عند ابن يعيش هو أخذ لفظ من لفظ آخر لمعنى، ويقصد به الاشتقاق الصغير، أي اشتقاق الفعل من المصدر واشتقاق المشتقات من الفعل، وهو بهذا لم يخرج عن حدود الصيغ الصرفية المعروفة كاسم الفاعل والمفعول يضاف إليها الزيادة الصرفية، كما في قوله" وأما "تمسكن" و"تمدرع" فهو قليل كالمشتق من الاسم بالزيادة، نحو سبحل وحمدل" 154/9. و هذه الجوانب تنضوي في الأساس تحت علم التصريف، وبالتالي يكون الاشتقاق عنده جانبا من جوانب التصريف البارزة، واستخدمه للدلالة على بناء الألفاظ المختلفة حسب الصيغ الصرفية المختلفة.

4- ابن يعيش، شرح الملوكي في التصريف، تحقيق، فخر الدين قباوة، ط1(المكتبة العربية، حلب، "سوريا"1973م)ص18 ــ 19.

5- محمد بن مالك الطائي النحوي (ت 276هـ)،إيجاز التعريف في علم التصريف، تحقيق محمد عثمان، مكتبة الثقافة الدينية، الطبعة الأولى،2009،1430،القاهرة، ص 81.

6- المرجع نفسه، ص 19.

فأمّات المعاجم تبين أنّ دالّهما اللغوي هو التغيير وهو ما أخذ به علماؤنا، وتداولوه في مؤلفاتهم، فدار المعنى الاصطلاحي في فلك تغيير الأصل وتقلباته أو تحولاته[1].

من خلال هذه التعريفات يمكن القول أن هناك تداخل في المفهوم العلمي لمصطلح الصرف والتصريف[2] وهذا التداخل في المفهوم العلمي للمصطلحين مككننا من طرح التساؤلات التالية: بما أن الصرف والتصريف يدوران في فلك التغيير والتحويل، فهل هما مترادفان؟ ولم اختير التحويل على التغيير؟ وهل التغيير هو التحويل؟ اختار بعضهم التحويل على التغيير لما في التحويل من معنى النقل، فورد أن التحويل هو: نقل الشيء من موضع إلى موضع. فأنت مثلا تنقل حروف "الضرب" إلى: ضَرَبَ، ويَضْرِب، وغيرهما، فيكون التحويل أولى من التغيير[3]. والعلاقة بينهما علاقة خصوص وعموم فلا يفسر التصريف لغة بالتحويل؛ لأنه أخص من التغيير.

أما لم اختر التصريف على الصرف مع أنه بمعناه؟ قال العزي لأن في هذا العلم تصرفات كثيرة فاختير لفظ يدل على المبالغة والتكثير[4]. وعليه فلتصريف معنيان[5]: أحدهما: لغوي، وهو ما دونه المعجم اللغوي. وثانيهما: اصطلاحيّ، وهو ما وضَعَه له أهل هذه الصناعة، والمراد

1- أما لفظة التصريف في القرآن فقد وردت في قوله تعالى:"وتصريف الرياح والسحاب المسخر بين السماء والأرض"(سورة البقرة: الآية 164) وقوله" وتصريف الرياح آيات لقوم يعقلون"(سورة الجاثية: الآية 4). قال الإمام القرطبي:"تصريفها: إرسالها عقيما وملقحة، وقيل تصريفها إرسالها جنوبا وشمالا" الإمام القرطبي، الجامع لأحكام القرآن، ط(دار الشام، بيروت، د.ت)197/2. وهي بهذا المعنى تفيد التدبير والتوجيه. وورد في القاموس المحيط(وتصريف الآيات: تبيينها)، وتفيد(في الدراهم والبيعات إنفاقها) الفيروز أبادي، القاموس المحيط، مادة صرف، 513/3. ويرى الفاخري أن المعاني التي استعملت فيها لفظة صرف تدور حول التغيير والتحويل وهو يتفق مع التدبير والتوجيه في كثير من جوانبه إذ لا يخالفه إلا فيما يقتضيه التضعيف من كثرة ومبالغة، فإذا قلت(صرف) كان المعنى محدودا أما إذا قلت (صرّف) فإن الصيغة تقتضي أن يكون كثيرا ومبالغا فيه. صالح الفاخري، علم التصريف العربي،، ط(ELGA، مالطا، 1999م) ص20.

2- نجد هذا الربط والتداخل بين المصطلحات عند الجرجاني الذي أورد الصرف عنوانا لكتابه"المفتاح في الصرف" إلا أنه في أثناء التطبيق استخدم مصطلح التصريف، يقول:"اعلم أن التصريف تفعيل من الصرف وهو أن تصرف الكلمة المفردة فتتولد منها ألفاظ مختلفة ومعان متفاوتة". وهذا الاستخدام مطرد في هذا الكتاب، وإن ورد مصطلح الصرف فهو على سبيل الترادف بينهما. ينظر: عبد القاهر الجرجاني، المفتاح في الصرف، تحقيق، علي الحمد، ط(مؤسسة الرسالة، بيروت، 1987م) ص26. ومثله الميداني الذي جاء الصرف عنوانا لكتابه"نزهة الطرف في فن الصرف" فقد استخدم في التعريف المصطلحين معا إلا أنه أظهر التصريف على الصرف، وهما مترادفان عنده، يقول عندما عرّفَ الصرف:"التصريف تفعيل من الصرف وهو أن تصرف الكلمة الواحدة فتتولد منها ألفاظ مختلفة ومعان متفاوتة، مثل أن تقول من الضرب ضرب يضرب"وعنده أن التصريف لا يختص بالأفعال دون الأسماء بل يطلق عليهما جميعا، فالاسم له واحد وجمع وتعريف وتنكير ونسبة وتصغير كما للأفعال ماض ومستقبل وأمر ونهي وفاعل ومفعول ويطلق عليه حكم الصحة والاعتلال كما يطلق على الأسماء. وهذه مرتبطة بالتحول والتغيير والتنقل. ينظر: أحمد الميداني، نزهة الطرف في علم الصرف، تحقيق، لجنة إحياء التراث، ط1(دار الآفاق الجديدة، بيروت، 1981) ص4. وكما ونجد هذا التداخل في إيراد المصطلحات المبني على الدال اللغوي عند ابن يعيش فقد ورد مصطلح التصريف عنده في مواضع متفرقة، منها قوله: إن "التصريف مصدر، ووُضع كالعَلَم على هذا العِلْم؛ للفَرق، خصُّوا به ما عرض في أصول الكلم وذواتها من التغيير، كاختصاصهم عِلْم العربية بالنحو"، ويضيف أن التصريف:"كلام على ذوات الكلم وفعله: صرّفته أصرِّفه تصريفا. يقال: صرَّفْتُه فتصرَّف، أي: طاوع وقبل التصريف"، وهنا يبين أن التصريف عَلَم على هذا العلم الذي يبحث فيما عرض لأصول الكلم وذواتها من التغيير، وهذا حدّه عنده. فهو مطرد يشمل جميع أنواع التغيير التي تطرأ على بنية الكلمة، ويتداخل هذا المصطلح عنده مع الاشتقاق، وذلك عندما يذكر أن التصريف "اشتقاقه من تصريف الحديث والكلام، وهو تغييره بحمله على غير الظاهر. ومنه تصريف الرياح، وهو تحويلها من حال إلى حال". ينظر: ابن يعيش، شرح الملوكي في التصريف، تحقيق، فخر الدين قباوة، ط1(المكتبة العربية، حلب، "سوريا"1973م)ص18 ــ 19.

أما ابن الحاجب فعنده رؤية خاصة به تمثلت عندما صرّح أن التصريف علم في قوله:"التصريف علمٌ بأصول يُعرف بها أحوال أبنية الكلم التي ليست بإعراب " وهو ما أخذ به المتأخرون . فمدلول المصطلح شمل عنده أبنية الكلمة وأحوالها التي ليست بإعراب ولاً بناء. ينظر : رضي الدين الاستراباذي، شرح شافية ابن الحاجب، تحقيق، محمد نور الحسن وآخرين، ط(دار الكتب العلمية، بيروت، 1982م)1/1.

3- التفتازاني، شرح مختصر التصريف العزي في فن الصرف، تحقيق، عبد العال سالم مكرم،ط1(ذات السلاسل، الكويت، 1983م)، ص 26.

4- المرجع نفسه، ص 28.

5- المرجع نفسه، ص 24، 25.

هنا صناعة التصريف، وهو تحويل الأصل[1] الواحد أو تغييره[2]، والأصل ما يبنى عليه شيء. والمراد هنا المصدر (إلى أمثلة) أي أبنية وصيغ، وهي الكلم باعتبار الهيئات التي تعرض لها من الحركات، والسكنات، وتقديم بعض الحروف على بعض وتأخيره عنه،(مختلفة) باختلاف الهيئات كضرَب يَضْرب، ونحوهما من المشتقات(لمعان). فالتصريف = تحويل الأصل أي المصدر إلى أمثلة مختلفة لأجل حصول معان(مقصودة لا تحصل) تلك المعاني(إلا بها) أي بهذه الأمثلة. والتصريف ههنا غير علم التصريف الذي هو معرفة أحوال الأبنية. ومن جهة أخرى نرى صاحب العين ينتهج نهج المقاربة في مدلول التصريف عندما يعرفه بالاشتقاق، يقول:" التصريف: اشتقاق بعض من بعض"[3]. وهو بهذا يعطي التصريف بعدا جديدا يصدر عن نظرة بعدية، فهو بمعنى الاشتقاق من منطلق أخذ كلمة من كلمة بدلالة التنكير، بإيحاء أن الكلم مرتبط بعضه ببعض.

وبتتبع مؤلفات الصرف والتصريف في آثار الدارسين، ومن خلال القائمة التي قدمها أحمد هريدي[4] وعَرَضَ فيها لمؤلفات الصرف يظهر أن الغالب على تسمية العلم هو "التصريف"، أما "الصرف" فيسير على استحياء إلى جانب التصريف، بالرغم من أننا نجد المصطلحين ينهضان في الكتب جنبا إلى جنب، وإن كانت التسمية بداية معنونة بالتصريف، فقد أدمج القدماء لفظ (التصريف) بلفظ (الصرف) في دلالة واحدة بحيث يتوهم الدارس أنهما دالتان لمعنى واحد لا يختلف وهما مختلفان اشتقاقا ومختلفان اصطلاحا، فمن حيث اختلافهما اشتقاقا أن الصرف مصدر (صرف) والتصريف مصدر الرّباعي (صرّف)[5].

1- وهناك خلاف بين البصريين والكوفيين حول (الأصل)، فهو عند البصريين المصدر، وعند الكوفيين الفعل.

2- ينص ابن السّراج على مفهوم التغيير في إيراده مصطلح التصريف دون الصرف، يقول هو "ما عرض في أصول الكلام وذواتها من خمسة أقسام: زيادةٌ، وإبدال، وحذفٌ، وتغييرٌ بالحركة والسكون،وإدغامٌ؛ وله حدٌّ يعرف به،"وينص من هذا على أن مدلول مصطلح التصريف هو تغيير يحدث في أصل الكلمة وهيئتها، وهذا الفهم مستقى من دلالة المصطلح اللغوية، فضلا عن الاصطلاحية، وكذلك لا يختلف كثيرًا عما فُهممن قول سيبويه والمبرد، غير أنه يجعل التصريف خمسة أقسام: زيادة، وإبدال، وحذفٌ، وتغييربالحركة والسكون، وإدغام. وهذه محاولة لتوسيع دائرة مفهوممصطلح التصريف، مع العلم بأن هذه المباحث هي نفسها التي ذكرها سيبويه باستثناء أبنية الأسماء والأفعال. للمزيد من التفصيل ينظر: أبو بكر محمد بن سهل بن السرّاج، الأصول في النحو، تحقيق، عبد الحسين الفتلي، ط3(مؤسسة الرسالة، بيروت، 1988م) 231/3، التغيير"، وذكر أن هذا " سمّي تصريفاً التصريف الكلمة الواحدة بأبنية مختلفة، وخصّوا به ما عرض في أصول الكلام، وذواتها منالتغيير، وهو ينقسم وينظر : أبو محمد عبد اللهبن علي بن إسحاق الصيمري، التبصرة والتذكرة، تحقيق، فتحي أحمد مصطفى عليّ الدين، ط1(جامعة أم القرى، مكة المكرمة 1982م)،788/2.

3- الخليل بن أحمد الفراهيدي،العين، تحقيق، مهدي المخزومي، وإبراهيم السامرائي، ط (دار الهلال،د.م، د.ت) مادة صرف،109/7.

4- ينظر: ابن هشام، نزهة الطرف في علم الصرف، تحقيق، أحمد هريدي، ط (مكتبة الزهراء، القاهرة، 1990م)، ص55 وما بعدها.

5- هادي نهر، الصّرف الوافي، عالم الكتب الحديث، الأردن، 2010، ص9.

أما في العصر الحديث فظهرت بعض المصنفات تتخذ من الصرف عنوانا لها[1]، هذا من ناحية ومن ناحية أخرى نجد عددا من الدراسات الحديثة تميل إلى استخدام مصطلح التصريف في مقابل مصطلح(morphology)، ولذا يعرف هذا المصطلح بأنه"فرع من علم القواعد يبحث في تركيب الكلمات من حيث السوابق واللواحق والدواخل والجذور"[2] ويطلق على ما يهتم به علم التصريف مورفيم (morpheme) ويجمع على مورفيمات(morphemes)[3].

3-4- الصرف والتصريف ومباحثهما في الدراسات اللسانية الصورية الحديثة:

اقترن مصطلح الصرف والتصريف عند علماء اللغة المحدثين كذلك بالدال اللغوي للمصطلح وهو التغيير الذي جاء على ضربين[4]: تغيير صرفي محض، وتغيير شامل، ومن الضرب الأول: التغيير في صوت العلة من(foot) إلى(feet) ومن(man) إلى(men) ،أما التغيير الشامل فهو تغيير في شكل الأصل يتناول الصيغة كلها في تحويل الأصل الواحد إلى أمثلة مختلفة لمعان مقصودة لا تحصل إلاّ بها[5].، ولا يقتصر على بعض الفونيمات، كما في(went) ماضي(go)[6].

فالمستوى الصرفي من مستويات علم اللغة، يقول ماريوباي"مستوى الصرف(morphology) أو مستوى دراسة الصيغ اللغوية وبخاصة تلك التغييرات التي تعتري صيغ الكلمات، فتحدث معنى جديدا، مثل اللواحق التصريفية على سبيل المثال(s) التي

1- نحو:" شذا العرف في فن الصرف" للحملاوي، و"الصرف الواضح" لسعيد النائلة، و"عمدة الصرف" لكمال إبراهيم، و"التطبيق الصرفي" لعبده الراجحي، و"المغني الجديد في علم الصرف" لمحمد خير حلواني. غير أنه رغم الإطباق من بعض المحدثين على استخدام مصطلح الصرف، فإن عددا منهم فضل استخدام مصطلح التصريف على الصرف، نحو: فخر الدين قباوة في مصنفه"تصريف الأسماء والأفعال"، ومحمد محمود هلال في" الوافي الحديث في فن التصريف" غير أنهما لم يناقشا المصطلح، ويبينا سبب اختيارهما له. ويرى الفاخري أن سبب اختيارهما مصطلح التصريف "من قبيل التمسك بالتسمية التي درج عليها المتقدمون". ولا نستطيع أن نجزم بهذا فقد ورد عندهما مصطلح الصرف كما ورد مصطلح التصريف. وهناك أيضا كتاب" التصريف العربي من خلال علم الأصوات الحديث"للطيب البكوش، تناول فيه أبنية الفعل الثلاثي المجرد من خلال علم الأصوات، وهو من أحدث ما ألف في التصريف من منظور علم الأصوات الحديث، وكان يمكن أن يؤدي إلى تطور في الدرس اللغوي بعامة والدرس الصرفي بخاصة، لو قام مؤلفه بتعميم دراسته على مختلف المباحث الصرفية وهو ليس بصعب على من قدم تلك الدراسة المتميزة عن أبنية الفعل الثلاثي.ينظر : صالح الفاخري، علم التصريف العربي، ص21، ص27.

2- محمد الخولي، معجم علم اللغة النظري، ط (مكتبة الحياة، لبنان، 1983م) ص255.

3- أما عند الأوروبيين فإنهم يطلقون عليه(morphology)، ويعني بالنظر في المورفيمات(morphemes) وهي قسمان: 1) مورفيم حر وهو الكلمة التي يمكن استعمالها مستقلة عن سواها مثل: كتب، ونظر، وجلس، ورجل وبيت2) مورفيم متصل أو مقيد: وهو الكلمة التي لا يتحدد معناها إلا بانضمامها إلى غيرها، نحو أحرف المضارعة،(أنيت)ونحو: زوائد الصيغ(استفعل) وعلامة التثنية، والجمع، والتأنيث(اللواحق والسوابق). محمد الخولي، معجم علم اللغة النظري، ص174.

4- حسن هنداوي، مناهج الصرفيين ومذاهبهم، ص23.

5- عبد الصبور شاهين، المنهج الصوتي للبنية العربية، مؤسسة الرسالة، بيروت، 1980، ص23.

6- ماريوباي، أسس علم اللغة، ترجمة أحمد مختار عمر، ط (طرابلس، 1973م،) ص106.

تضاف إلى(cat) فتصيرها جمعا. والسوابق مثل(er) قبل(tell) لتعطيها معنى يخبر مرة ثانية. والتغييرات الداخلية مثل تغيير حرف العلة في (sing) إلى(sang) لإفادة الماضي"[1]. فربط ماريوباي مصطلح الصرف بدراسة الصيغ وتغيراتها[2].

وهذا عموما ما نحا إليه علم اللغة الحديث فيما يتعلق بمصطلح الصرف والتصريف نحو دراسة الصيغة والعلامة الصرفية التي تدل على المورفيمات، والوحدات الصرفية والصور الصرفية[3].

فأصبحت كل دراسة تتصل بالكلمة أو أحد أجزائها، وتؤدي إلى خدمة العبارة والجملة هي في محور مصطلح الصرف[4].ونجد ذلك في مفهوم هذا المصطلح عند حسان الذي بين أن الصرف هو" ذلك العلم الذي يتناول الناحية التشكيلية التركيبية للصيغ والموازين الصرفية[5]. وعندها نجد أن المصطلح قد اتخذ بعدا إضافيا تمثل في البحث"في الوحدات الصرفية وأهم أمثلتها الكلمات وأجزاؤها ذات المعاني الصرفية كالسوابق واللواحق"[6] سواء أكانت هذه اللواحق

1- ويطلق الغربيون عموما على التصريف لفظ (morpholgogy) ويرون أنه يختص بدراسة الصيغ، ويسمون النحو(syntax) ويعرفونه بأنه تنظيم الكلمات في شكل مجموعات أو جمل، ويجمعها عندهم مصطلح(علم القواعد ــ grammar) أو التركيب(structure)، أو التركيب القواعدي(grammatical structure) ماريوباي، أسس علم اللغة، ص43 ــ 44.

2- ماريوباي، أسس علم اللغة، ص53.

3- ويفرق علم اللغة الحديث بين الوحدات الصرفية والصور الصرفية، فالصور الصرفية لها وجود مباشر منطوق مسموع، وأما الوحدة الصرفية الجامعة للصور الصرفية فهي موجودة غير أنه وجود غير مباشر، وذلك نحو ضرب واضطرب، فالفرق بين الصيغتين من ناحية البنية الصرفية هو الفرق بين قرب واقترب، لكن التغيير ليس واحدا من الناحية الصوتية على الرغم من اتحاد الوظيفة في بنية اللغة، ومثل ذلك أمر التاء والدال في قرب واقترب من جانب، وزهر وازدهر من الجانب الآخر، فكل من الطاء والتاء والدال تأتي في جوار صوتي بعينه، وتدخل إحداها في هذه البنية، ويقودنا هذا إلى القول بوجود ثلاث صور صرفية لوحدة صرفية واحدة، ينظر: محمود فهمي حجازي، مدخل إلى علم اللغة، القاهرة، 1978م، ص58، و حسن هنداوي، مناهج الصرفيين ومذاهبهم، ص22، 23.

4- كمال بشر، دراسات في علم اللغة، ط (دار المعارف، مصر، 1973م)، ص221. أما عن رؤية بشر بضرورة إبعاد بعض الظواهر اللغوية التي عدها علماء اللغة من مباحث التصريف وإلحاقها بعلم الأصوات، مثل الأوزان وصيغ جمع التكسير، وبعض صور الإبدال كالذي يحدث لتاء الافتعال إذا جاءت بعد حرف من حروف الإطباق(ص، ض، ط، ظ) فهذه الظواهر ليس كما قرر فهي تفيد دراسة الجملة، فهناك أوزان تفيد لزوم الفعل أو تعديه أو دلالته على معنى من المعاني كالتعجب وغيره مما يساعد على فهم شكل الجملة. وجمع التكسير يفيد دراسة الجملة أيضا، فهناك أوزان خاصة تمنع من الصرف، فضلا عن أن الجمع نفسه بعامة، يقتضي علاقة خاصة بالفعل وهذا كله له أثره الواضح في الجملة.أما تاء الافتعال فإنها ككل زيادة تزاد على الفعل المجرد وحدة صرفية تؤدي إلى معان نحوية، كالتعدية وعليه فالعرب محقون في جعلها من موضوعات علم التصريف. وأما عن رأيه المتضمن إلحاق موضوعات بعلم التصريفهي الآن من مباحث فرع آخر من فروع الدراسة اللغوية فإننا ــ حسب رأي الفاخري ــ نظن أنه بحاجة إلى إعادة نظر، فتقسيم الكلمة والنظر إليها من حيث النوع والعدد، يحتاج إليها الباحثون في العلمين، فالكلمة هي موضوع علم التصريف كما أنها هي الركن الأساسي في الجملة التي هي موضوع علم النحو، ودراستها ينبغي أن تكون من خلالهما. ينظر: الفاخري، علم التصريف العربي، ص32. وهذا الرأي يلتقي مع مما قرره فندريس من أن "تصنيف الفصائل عمل من أعمال الصرف العام" وأن التسمية التي تطلق عليه هي الفصائل النحوية أو القواعدية Grammatical categories. فندريس، اللغة، ترجمة الدواخلي والقصاص، ص126. و الفاخري، علم التصريف العربي، ص32.

5- تمام حسان، مناهج البحث في اللغة، ص170.

6- كمال بشر، دراسات في علم اللغة، ص.11

صدورا أو أحشاء أو أعجازا[1]، فهو إذن "يدرس الصيغ والمفردات من حيث ترتيب أصواتها وأصالتها وزيادتها واشتقاقات الصيغ"[2] وعلى هذا فهو "أحد مستويات البحث الحديث في اللغة"[3].

ويرى عبد الصبور شاهين قد فرق بين مصطلحي الصرف والتصريف، فالمعنى العلمي4 هو مدلول(الصرف)، والمعنى العملي هو مدلول(التصريف)،وبهذا يتخصص كل من المصطلحين لدلالة واحدة، ويضيف أنه بذلك يقترب معنى(الصرف) من معنى (المورفولوجيا) في الدراسات اللغوية الحديثة[5].

ونلاحظ اقتران مصطلح الصرف أيضا بمصطلح التحويل من صيغة إلى صيغة، وإن كان هناك إشارات عند القدماء إلى التحويل في الصيغ الصرفية إلا أن الاهتمام بهذا المصطلح قد زاد في الدراسات الحديثة مع ظهور المنهج التحويلي عند تشومسكي[6].

ونجد أحيانا في علم اللغة الحديث[7] جمعا بين مصطلحي(النحو) و(الصرف) تحت اسم واحد هو(التركيب القواعدي) مما يدل على التداخل بين المصطلحين وعدم وضوح الحدود بينهما. فضلا عن أن هناك تبادلا مطردا بين الصرف والنحو، إذ تستغني بعض اللغات بأحد المصطلحين عن الآخر[8]. أما مسألة التعامل مع الأصل فليست موجودة في اللغات السامية، وليس" للنحو والقواعد صلة بالأصول ولكن له صلة بالكلمات الكاملة"[9]. أما موضوع التصريف

1- تمام حسان، مناهج البحث في اللغة، ص175.

2- علي أبو المكارم، الظواهر اللغوية والتركيبية، ص220.

3- كمال بشر، دراسات في علم اللغة، ص11.

4- والتعريف العلمي هو أن الصرف علم بأصول يعرف بها أحوال أبنية الكلمة التي ليست بإعراب. وسار الحملاوي في هذا الاتجاه من حيث التفريق بين الناحية العملية والناحية العلمية للتصريف، يقول الصرف ويقال له التصريف هو لغة التغيير ومنه تصريف الرياح أي تغييرها واصطلاحا بالمعنى العملي تحويل الأصل الواحد إلى أمثلة مختلفة لمعان مقصودة لا تحصل إلا بها كاسمي الفاعل والمفعول واسم التفضيل والتثنية والجمع إلى غير ذلك وبالمعنى العلمي علم بأصول يعرف بها أحوال أبنية الكلمة التي ليست بإعراب" ينظر: أحمد الحملاوي، شذا العرف في فن الصرف، ص3. وواضح أنه يفرق بين الناحية العلمية والناحية العملية للتصريف، ولكنه لم يحدد لكل ناحية مصطلحا كما فعل عبد الصبور شاهين.

5- عبد الصبور شاهين،المنهج الصوتي للبنية العربية، ط(مؤسسة الرسالة، بيروت، 1980م)، ص23

6- استخدم القدماء عبارات تشير إلى مصطلح التحويل في الصيغ الصرفية، نحو: صيغ المبالغة مثلا فرع عن اسم الفاعل، لأنها محولة عنها. ونحو: ينوب (فعيل) عن (مفعول)، مثل، (قتيل) بمعنى (مقتول) والتحويل في مسألة الأصل والفرع فـــ(قال) محولة عن(قول). ينظر: ظاهرة التحويل في الصيغ الصرفية، محمود سليمان ياقوت، ط (دار المعرفة الجامعية، الإسكندرية، 1985م)ص7. وينظر: عبده الراجحي،النحو العربي والدرس الحديث، ط (دار النهضة العربية، بيروت، 1979م)، ص24.

7- ماريوباي، أسس علم اللغة، 45، 53، 54 . وينظر: حسن هنداوي، مناهج الصرفيين ومذاهبهم، ص33.

8- كارل بروكلمان، فقه اللغات السامية، ترجمة، رمضان عبد التواب، ط(الرياض،1977)، ص83.

9- المرجع نفسه، ص83.

في اللغات السامية فيشمل: الاسم والضمير والعدد والأدوات والفعل، وهذا كما ذكره موسكاتي في كتابه "المدخل في النحو المقارن للغات السامية"[1] .

وإنما تدخل هذه الأقسام في علم الصرف في اللغات السامية ــ كما يذكر هنداوي[2] ــ لأنها في مجال بنية الكلمة، والخلاف كما يذكر هنداوي في تقسيم التصريف ينحصر في مسألتين، هما:الخلاف في الأقسام والخلاف في الأسماء مع أن المسميات التي وسموها بها تبقى واحدة[3].

4- الصرف والاشتقاق في الدرس اللساني الحديث:

ورد مصطلح الصرف مقترنا بالاشتقاق، فقد جاء في التسهيل" الصرف أعم من الاشتقاق لأن بناء مثل قردد من الضرب يسمى تصريفا ولا يسمى اشتقاقا لأنه خاص بما بنته العرب"[4]، وهنا نجد ذكرا للمصطلحات الثلاثة: الصرف والتصريف والاشتقاق. ففي البداية ذكر مصطلح الصرف، ويؤكد أن بناء كلمة من كلمة هو التصريف، فـ"التصريف في اللغة هو التغيير، والاشتقاق هو تغيير الصيغة إلى صيغ أخرى تخالفها في الوزن، فليس هناك ما يمنع أبدا إدراج الاشتقاق في التصريف"[5]. وكذلك مصطلح الصرف في المعاجم يرد أيضا بمعنى التغيير[6]؛ وذلك لأنهم تناولوا الكلمة بعيدة عن معناها الوظيفي فتناوله على أساس معناه المعجمي فقط وفرق بعيد بين المعنيين ذلك لأن المعنى المعجمي قاصر عن المعنى الدلالي فلا نتوقع منهم أن يذكروا لنا دور الصرف في إضفاء الدلالات على الكلمة ولا أن يوضحوا أهميته على مستوى العبارة فنراهم محصورين في دائرة معنى الصاد والراء والفاء معجميا. وبناء على هذا المعنى المعجمي جاء المعنى الاصطلاحي الذي اتفق عليه النحاة وإن كان ينطبق على حالات خاصة من المفردات اللغوية ومن هنا كانت بعض الكتب تمزج بين التعريفين اللغوي

1-An introduction to the comparative grammar of the Semitic languages,. Moscati,2nd edition.1969.wisbaden p. 75, 102 115,120, 122,

2- حسن هنداوي، مناهج الصرفيين ومذاهبهم، ص40.

3- المرجع نفسه، ص44ــ 45.

4- السيوطي، المزهر في علوم اللغة وأنواعها، تحقيق، محمد أحمد جاد المولى،علي البجاوي، محمد أبو الفضل، ط(دار إحياء الكتب العربي، عيسى البابي الحلبي، د.ت) 1/ 351.

5- حسن هنداوي، مناهج الصرفيين ومذاهبهم، ص20.

6- الدناع، دور الصرف في منهجي النحو والمعجم، ص30 ــ 31.

والاصطلاحي في تعريف واحد، وبعضهم وفق في الفصل بينهما كالصبان في شرحه على الأشموني[1].

فالاشتقاق هو عملية استنباط وتوليد صيغة من صيغة أو لفظ من لفظ[2]، بحيث تكون الصيغتان واللفظان متفقين في المعنى العام وفي الحروف الأصلية، أو بمعنى آخر هو أن تأخذ من اللفظ ما يناسبه في التركيب فتجعله دالا على ما يناسبه، فمجال الاشتقاق إذن البحث في مادة الكلمة من وجوهها المختلفة[3]. فهو إذن تغيير في بنية الكلمة لغرض لفظي أو معنوي"وهذا التعريف ينطبق على الصرف أيضا، وهذا ما اتضح من أن الاشتقاق من صميم الصرف يقول صاحب التصريح إن الصرف تغيير في بنية الكلمة لغرض معنوي أو لفظي"[4]، والتغيير اللفظي كتحويل(قول) من الأجوف و(غزو) من الناقص إلى (قال) و(غزا) مع أن هذا يمكن أن يفسر على أساس صوتي. والاشتقاق عند العرب المحدثين يقوم"على مجرد العلاقة بين الكلمات واشتراكها في شيء معين"[5]، وهذا يخرجنا من الخلافات حول الأصل والفرع الذي كان موضع خلاف بين مدرستي البصرة والكوفة أن الاسم مشتق من الفعل أو أن الفعل مشتق من الاسم[6].

1- يقول الصبان : إن الصرف يشمل تعريفه اصطلاحا ثلاثة أمور:

أولا: تحويل الكلمة إلى أبنية مختلفة لأداء المعنى الذي يقتضيه المقام كالتصغير واسم الفاعل والتكسير.

ثانيا: تغيير الكلمة لا لمعنى طارئ ولكن لغرض آخر كالزيادة والحذف والقلب والنقل والإدغام والإمالة.

ثالثا: العلم بأحكام بنية الكلمة بما لحروفها من أصالة وزيادة وصحة ويقصد ببنية الكلمة الصيغة التي توجد الكلمة عليها في حالة كونها مفردة فإذا ركبت مع جارتها أصبح البحث فيها خاصا بالنحو. السياق هو وسيلة نحوية غير صرفية ولكنه" يدخل في تحديد المعنى الصرفي المراد عند الحاجة" يُلحظ التداخل والامتزاج بين النحو الصرف، فالنحو في خدمة الصرف، كما أن الصرف ببحثه في ذات المفردات وطبيعة الكلمات سواء أكانت أسماء أو أفعالا أو أدوات أو ضمائر يكون دائما في خدمة اللغة على مستوى العبارة. ما تقدم تعريفات للصرف تتناول معناه العلمي بمفهومه التقليدي ولكن بمعناه العملي بذلك المفهوم أيضا فيعني تحويل الأصل الواحد إلى أمثلة مختلفة لمعان مقصودة لا تحصل إلا بها كتحويل مادة الفهم إلى يفهم إفهاما فهو فاهم. والصرف بمعناه التقليدي يشترك مع الاشتقاق في أحد أنواعه وهو الاشتقاق الكبير، ويرى الدناع أن الاشتقاق الكبير أبدال كالإبدال الصرفي إلا أنه إبدال لغوي يكون في جميع حروف الهجاء ما عدا الألف كما أن الإبدال الصرفي لأجل الإدغام يكون في جميع حروف الهجاء ما عدا الألف. وإذا كان الصرف يبحث في حروف الكلمات من حيث الزيادة والحذف والقلب والإبدال وتغيير الحركات والسكون والإدغام فإن بالاشتقاق تعرف أصول الكلمات وفروعها والعلاقات بينها وطرق صوغ بعضها من بعض، فتبين ما بين الصرف والاشتقاق من نسب قريب وأصل متصل كما يرى ابن جني، ولكن ينبغي أن يؤخذ كـ"خطوة أساسية في تصنيف الصيغ إلى أوزان ذات قيم صرفية خاصة وهي قيم ذات أهمية كبيرة في الجمل والتراكيب التي هي مجال البحث في النحو". ينظر : تمام حسان، اللغة بين المعيارية والوصفية، ص 179، الدناع، دور الصرف في منهجي النحو والمعجم، ص 31، كمال بشر، دراسات في علم اللغة، ص93.

2- إبراهيم أنيس، من أسرار اللغة، ص46.

3- محمد رشاد الحمزاوي،المعجمية:" مقاربة نظرية ومطبقة" مصطلحاتها ومفاهيمها "، مركز النشر الجامعي، تونس، د.ط،2004، ص 250.

4- الدناع، دور الصرف في منهجي النحو والمعجم، ص 39.

5- تمام حسان، مناهج البحث في اللغة، ص181.

6 - الأنباري، الإنصاف في مسائل الخلاف، المسألة رقم32، وينظر: خالد محمد عيال سلمان، أثر المحتسب في الدراسات الصرفية، دار ومكتبة الحامد للنشر والتوزيع،عمان، الطبعة الأولى،2011، ص 81،80.

كانت مسائل الصرف في البداية متناثرة ضمن المسائل النحوية، وقد فرق علماؤنا القدماء بين اصطلاح النحويين واصطلاح اللغويين[1]، يقول ابن عقيل:"وإنما قال المصنف"كلامنا" ليعلم أن التعريف إنما هو للكلام في اصطلاح النحويين؛ لا في اصطلاح اللغويين"[2] ومن هنا ــ كما يذكر الدناع ــ وضح الانفصال التام بينهما ولكن مع تطور وجهة نظر الباحثين التحمت هذه الدراسات واعتمدت على بعضها ويرى بشر أن الصرف "وسيلة وطريق من طرق دراسة التركيب والنص اللذين يقوم بالنظر فيهما علم النحو"[3]. ويبعد الصرف بمعناه التقليدي الأفعال الجامدة لأنها لا تقبل التصريف والاشتقاق، نحو إبعادهم فعلي التعجب "ما أفعله" و"أفعل به" مع أن أصلها أفعال ثلاثية متصرفة، لأنهم يرجعون نحو"ما أحسن" إلى "حسن" بالتشديد و"أحسن به" إلى "حسن" أي أن أصلهما فعلان متصرفان.وكذلك يبعد الحروف لأنها لا تتصرف ومجهولة الأصل. ويرى الدناع أن اعتماد العرب على الأصل جعلهم يجهدون أنفسهم في البحث عن الحرف الثالث حتى تستقيم الكلمة مع الميزان الصرفي. ونتيجة لتحكم هذا الميزان في مباحثهم فإنهم لم يولوا الحرف أية عناية فجعلوه في خدمة غيره من أسماء وأفعال وسلبوه المعنى بنفسه وأضافوه إلى غيره[4].

5- منزلة الصرف في المعجم العربي:

إن المعجم العربي في بداية الصناعة أسّس على تقنية هامة من تقنيات علم الصرف أو علم التصريف –كما يصطلح عليه آخرون– لذلك اعتمدت المعاجم القديمة في مجملها تحديد الجذور وإحصائها. وإرجاعُ المشتقات إلى جذورها كل قضية صرفية؛ ذلك ما نتبينه في مرحلة الكــلام عن أثر الصرف في منهجية ترتيب المعاجم ودوره في صناعة المعاجم[5].

1- وجدنا تداخلا بين مصطلحي التصريف والاشتقاق، فقد اتفقا من جهة واختلفا من جهة أخرى، وتداخلا من جهة ثالثة واندمجا معا من جهة رابعة، ولا مشكل في ذلك فمنطلق دلالتهما واحد. ويكمن التداخل بين المصطلحين بإطلاق مصطلح التصريف على الاشتقاق، والاشتقاق على التصريف، فظهر مصطلح الصرف على هذا العلم، ليشمل المسائل والقواعد التي وضعها العلماء، ولعلّ ظهوره يواكب استقلال هذا العلم عن النحو؛ ولهذا فإنّ بعضهم يَعُدّ التصريف هو المعنى العملي، والصرف هو المعنى العلمي؛ أي أن التصريف يرتبط بكثرة دوران الأبنية واشتقاقها والعمل فيها، والصرف يرتبط بالأصول الكلية التي ينبني عليها معرفة أحوال المفردات.

2- ابن عقيل، شرح ابن عقيل على ألفية ابن مالك، 14/1.

3- كمال بشر، دراسات في علم اللغة، ص84.

4- الدناع، دور الصرف في منهجي النحو والمعجم، ص 35 ــ 36.

5- يمثل الصرف أهم أسس تنظيم المعجم وتختلف شدة العلاقة بينهما من لغة إلى أخرى، إلا أن هذه العلاقة من التداخل والتماسك بحيث ينظر البعض إلى نظام الصرف بأكمله كتابع للمعجم. يمد المعجم الصرف بالمعطيات الفونولوجية والصرفية والنحوية والدلالية والإيتمولوجية (الأصل المعجمي) والتي يحتاج إليها في تطبيق قواعد الصرف المختلفة. وعادة ما يرتبط عدم انطباق هذه القواعد بتوفر مفردات معجمية بديلة للكلمات المراد اشتقاقها. ينظر نبيل علي، اللغة والحاسوب، ص 280

كما أن هناك أمر آخر يبين منزلة الصرف في المعاجم العربية هي تلك المواد الهائلة التي تضمها المعاجم في ثناياها من مسائل صرفية ولو جمعت ما في المعاجم من مسائل الصرف لخرج لك علم كامل قد يغنيك درسه عن مراجعة كتب الصرف المتخصصة، فأغلب تلك المعاجم تطرح في مقدماتها قضايا ومباحث صرفية موسعة.

فإذا نظرنا في علم الصرف وجدناه وثيق الصّلة هو أيضا بعلم المعجم[1]. على أن هذه الصلة لا تظهر إلا إذا فصلْنا بين مفهومين مختلفين للصّرف، لكنهما متداخلان تداخلا كبيراً في أذهان كثيرين من الناس، ونعني بهما[2]: الصّرف من حيث هو علم الوحدات الصرفيّة المعجميّة[3]، وهذا هو المسمّى علم الصّرف الاشتقاقي[4]أو علم الصرف المعجمي[5]أو علم الصرف من حيث هو علم تصريف الوحدات الصّرفية المعجمية، وهذا هو علم التصّريف[6].

6- الصرف وحوسبة اللغة:

يعد علم الصرف أساسا متيناً في التعامل مع اللغة العربية . أما حوسبة الصرف فتعد مطلباً أساسياً لكثير من التطبيقات اللغوية التي تعتمد على النظم الآلية[7] .

فقد استطاع الباحثون في ميدان معالجة اللغة العربية آليًا وضع برامج حاسوبية لسانية طُبقت فيها جميع الخوارزميات الصورية، فكان الصرف العربي أهمّ أسس هذه المقاربة، فهو يعدّ في اللّغة العربية مدخلاً أساسيًا في فهم الأداء الشامل للغة العربية هذا من الجانب اللّغوي أمّا من الجانب الحاسوبي فتعدّ معالجة الصرف العربي آليًا مطلبًا أساسيًا لحوسبة عمليات تحليل النصوص ومعالجتها آليًا، علاوةً على كونه أساسًا لا غنى عنه لحوسبة المعاجم واسترجاع المعلومات واستخراج الكلمات[8].

إنّ الصرف العربي بلا شك، من أرقى فروع علوم اللّغة العربية[9] ويمثل من الوجهة النظرية النموذج الأمثل لدراسة الظاهرة الصرفية بصفة عامة والمعجم بصفة خاصة[10]. ولهذا

1- نبيل علي، اللغة والحاسوب، مرجع سابق، ص 270.

2- ابراهيم بن مراد، مقدمة لنظرية المعجم، دار الغرب الإسلامي،بيروت، الطبعة الأولى،1997، ص 41،40.

3 -Morphèmes lexicaux

4- Morphologie dérivationnelle

5- Morphologie lexicale

6- Morphologie flexionnelle

7- نبيل علي، اللغة العربية والحاسوب، مرجع سابق، ص 248.

8- المرجع نفسه ، ص 271.

9- من المعلوم بداهة أن خصائص أي نظام طبيعي يراد بناؤه ستؤثر بشكل كبير على طريقة تصميمه وتنفيذه آليا، وتتجلى هذه الحقيقة في بناء تطبيقات معالجة اللغات الطبيعية وخصوصاً اللغات السامية منها وأهمها اللغة العربية. تعتمد اللغة العربية – كما هو حال جميع اللغات السامية – على الصرف اعتماد كبيراً بل إنها تعتبر من فئة اللغات الطبيعية الأكثر تعقيداً والأغنى صرفياً. ينظر: نبيل علي، اللغة والحاسوب، مرجع سابق، ص 271.

10- عبد الرزاق تورابي، حوسبة الصرف العربي: "الموارد والخبرات اللسانية"،وقائع الندوة الدولية، اشراف واعداد عبد الفتاح حمداني، خالد الأشهب، محمد الراضي، المعالجة الآلية للغة العربية، وقائع الندوة الدولية، معهد الدراسات والأبحاث للتعريب، الرباط، المغرب، يونيو 2006، ص 48 وما بعدها .

يعدّ المستوى الصرفي أكثر المستويات اللّسانية قابليةً للمعالجة الآلية، وتصبح اللّغة من وجهها الصرفي قابلة للاستخدام في المنظومة الحاسوبية[1]، فالصرف يمثل أهم أسس تنظيم المعجم وتختلف شدة العلاقة بينهما من لغة إلى أخرى، فالمعجم يمدّ الصرف بالمعطيات الفونولوجية والصرفية والنحوية والدلالية والإيتمولوجية[2] والتي يحتاج إليها في تطبيق قواعد الصرف المختلفة.

ولهذا فإن المعجم يشغل ضمن المعالجة الآلية للصرف العربي خاصة واللّغة العربية عامة موقعًا مهمًا، فلا يمكن أن نتصور إنجاز معالجات عميقة للنصوص العربية بدون معلومات صرفية كافية عن قاعدة للمعطيات المعجمية للكلمة المفردة فمعالجة المعجم آليًا يتطلب أخذ بعين الاعتبار أمرين وهما **أولاً:** ضرورة الاستناد إلى إطار لساني صوري كنموذج لقاعدة المعطيات قادر على توصيف مستويات اللّغة الهدف.

ثانيًا: الاعتماد على مقدار كبير من المعلومات التي ترتب وتبوب أساسًا انطلاقًا من قواعد بيانات آلية.

6-1- خصائص الصرف العربي للمعالجة الآلية : نظرة عامة:

سنركز من خلال الحديث عن خصائص الصرف العربي على تلك النّواحي ذات الصلة بمعالجته الآلية.

بداية كما أشرنا في مقدمة هذا المبحث، يعد الصرف "في العربية" مدخلاً أساسيًا في فهم الأداء الشامل للغة[3]، هذا من الجانب اللّغوي، أما على الصعيد الحاسوبي فتعد معالجة الصرف العربي آليًا مطلبًا أساسيًا لميكنة عمليات تحليل النصوص المكتوبة والمنطوقة وفهما وتوليدها

1- محمد خليفة الدناع، دور الصرف في منهجي النحو والمعجم، منشورات جامعة قار يونس،بنغازي، ليبيا، د. ط، 1991، ص 04.

وينظر:

Cohen David, Etudes de linguistique sémantique et arabe ; « Junua Linguarum, série practica, 81 », the Hague, – Mouton, Paris, 1970.

2- نبيل علي، اللغة العربية والحاسوب، سابق ص 270.

3- يعد علم الصرف أساسا متيناً في التعامل مع اللغة العربية . أما حوسبة الصرف فتعد مطلباً أساسياً لكثير من التطبيقات اللغوية التي تعتمد على النظم الآلية . فلقد اهتم العرب الأقدمون بالصرف كثيراً حيث كان للصرف دوراً رئيساً في بناء المعاجم وتبويبها وتنظيمها . ولم يقتصر اهتمامهم بالصرف على ذلك بل استثمروه في دراسة العلاقة بين مباني الكلمات ومعانيها وخصائصها النحوية والصوتية . و من المعلوم أن خصائص أي نظام طبيعي يراد بناؤه ستؤثر بشكل كبير على طريقة تصميمه وتنفيذه آليا، وتتجلى هذه الحقيقة في بناء تطبيقات معالجة اللغات الطبيعية وخصوصاً اللغات السامية منها وأهمها اللغة العربية. تعتمد اللغة العربية على الصرف اعتماد كبيراً بل إنها تعتبر من فئة اللغات الطبيعية الأكثر تعقيداً والأغنى صرفياً . ويتجلى غناها الصرفي في إمكانية بناء عدة مئات من الكلمات النهائية وبمعاني مختلفة من جذر واحد وذلك باستخدام الموازين والسوابق واللواحق الصرفية المختلفة . للمزيد من التفصيل ينظر: اللغة العربية والحاسوب، سابق، ص 248، ص271، و المعجم المفصل في علم الصرف، مراجعة إميل بديع يعقوب، دار الكتب العلمية، لبنان، ط 1، 1413هـ/1993م.

آليًا، علاوة على كونه أساسًا لا غنى عنه كحوسبة المعاجم واسترجاع المعلومات[1] و"استخراج جذور الكلمات العربية"[2].

والصرف العربي بلا شك من أرقى علوم اللّغة العربية، ويمثل من الوجهة النظرية النموذج الأمثل لدراسة الظاهرة العرفية بصفة عامة. فنموذج الصرف العربي يَجُبّ معظم النماذج الأبسط كاستخراج جذور الكلمات.

يتميز الصرف العربي بعدّة خصائص[3]:

(1)- حدّة الخاصية الاشتقاقية ووضوح مسار عملية الاشتقاق.

(2)- اطراد التصريف.

(3)- الصلة القوية بين مباني الكلمات بالإضافة، وكرهه لتكوين الكلمات من خلال المزج والاختصار.

(4)- انتظام بنية الكلمة العربية لثبوت رتبة عناصرها "الصرف- نحوية".

(5)- شدة التداخل بين الصرف والفونولوجي من حيث تعدد قواعد الإبدال والإعلال وعمليات التغير "الصرف- صوتية" الأخرى[4].

(6)- الصلة العضوية بين الصرف العربي والمعجم العربي.

(7)- ثراء المعجم واعتماده على الجذور[5].

وسنتعرض هنا لبعض جوانب الصرف العربي ذات الأهمية بالنسبة لمعالجته آليا، هي:[6]

- الخاصية الثلاثية.

1- ينظر: نبيل علي، اللّغة العربية والحاسوب، سابق، ص 271.

2- سيأتي الحديث عن هذه النقطة في مبحث خاص من خلال الدراسة التطبيقية.

3- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 273.

4- ينظر: د. أحمد شامية "في اللّغة" دراسة تمهيدية منهجية متخصصة في مستويات البنية اللّغوية. دار البلاغ للنشر والتوزيع، الجزائر، الطبعة الأولى 1423- 2002، ص 74.

5- يرتبط تنظيم المعجم في أي لغة بصلات وثيقة مع طبيعة عمليات تكوين الكلمات بها، لذا فقد كان من الطبيعي أن يأتي تنظيم معجمنا العربي على أساس الجذور التي تعد أصل تكوين الكلمات لا على الترتيب الأبجدي كما في الإنجليزية مثلا، حيث تتميز شجرة المفردات العربية شأنها في ذلك شأن جميع اللغات السامية بقلة الجذور وتعدد الأوراق تعددا مضاعفا لذا فهي توصف بأنها شجرة (ثقيلة القاع) فرغم صغر نواة المعجم العربي أقل من 10 آلاف جذر تتعدد المفردات بصورة هائلة وذلك بفعل خاصية الاشتقاق أو الإنتاجية الصرفية العالية، حيث يمكن الاشتقاق من جذر (فاعل) 15 صيغة من مزيدات الأفعال تتيح كل صيغة منها صيغا مطردة وغير مطردة للمصادر وأسماء الأفعال والمفعول وأسماء المكان والزمان وصيغ التفضيل والمبالغة وصيغ الجموع، علاوة على ذلك تتميز العربية بثراء مفرداتها وكثرة مترادفاتها. إذا فالمعجم العربي ليس مجرد قائمة مفردات كما يجوز القول بالنسبة للإنجليزية بل بنية معقدة من العلاقات التي تربط بين مشتقات الجذور وصيغ الأفراد والجمع والمترادفات وما شابه ناهيك عن علاقات الاشتقاق الأعظم كما أوردها بن جني في كتابه الخصائص ليبرز ظاهرة اتصال المعنى وإن اختلف ترتيب حروف الأصل، هذا التعقيد في بنية المعجم العربي يتطلب الكثير من الجهد و البحث من قبل المتخصصين، ولا سبيل إلى سبر أغوار بنية المعجم العربي إلا باستخدام نظم المعلومات والكمبيوتر، وميكنة المعجم ليست رفاهية فكرية أو ذهنية، بل مطلبا أساسيا نابعا من طبيعة المعجم وضرورة تحديثه وإعادة تنظيمه بما يلبى المطالب المتوقعة منه وخاصة تسهيل تعلم الأجانب للغة العربية. للمزيد من بسط القول ينظر: د. نبيل علي، العرب وعصر المعلومات، عالم المعرفة، 1994، الكويت، ص339،338.

6- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 274.

- الأنماط الصرفية.
- ثنائية الصيغة الصرفية والميزان الصرفي.
- الإنتاجية الصرفية.
- الفائض الصرفي.
- اللّبس الصرفي
- تركيب الكلمات
- تعزيز عمليات تكوين الكلمات في اللّغة العربية

ودون الدخول في الجوانب التفصيلية لجوانب الصرف العربي سنتعرض إلى أهم النقاط الأساسية التي تهمنا في الدراسة التطبيقية من خلال بناء قاعدة معطيات صرفية لبناء المعجم الالكتروني للمعالجة الآلية للغة العربية.

❖ **الخاصية الثلاثية:**

إنّ معظم الكلمات العربية مشتقة من أصل ثلاثي[1]، وتعدّ الخاصيّة الثلاثية إحدى القسمات البارزة للصرف العربي، والتي تتضح في كثير من الظواهر أهمّها[2]:

(أ)–**__على مستوى المعجم__**، يفوق عدد الجذور الثلاثية، وبكثير عدد الجذور الرباعية والخماسية، ويزداد الانحياز نحو الثلاثية في معدلات تواتر الجذور المختلفة داخل النّصوص العربية بغض النّظر عن موضوعها أو أسلوب كتابتها.

(ب)–تتميز الجذور الثلاثية باتساع نطاق اشتقاقها (مقاسًا بعدد الصّيغ الممكن انطباقها عليها).

توضح القائمة التالية مدى التفاوت في الخاصية الاشتقاقية للجذور وفقًا لطولها[3]:

طول الجذر	عدد الصّيغ المنطبقة
ثلاثي	320
رباعي	60
خماسي	أقل من عشرة

__الجدول رقم 08:__ التفاوت في الخاصية الاشتقاقية للجذور وفقًا لطولها.

1– تمام حسّان، اجتهادات لغوية، سابق، ص 142.

2– اللّغة العربية والحاسوب، سابق ، ص 274.

3– المرجع نفسه، الموضع نفسه.

(جـــ)-عدم اطراد المعنى المعجمي للكلمات ذات الأصول الثلاثية مع المعنى الصرفي الذي يمكن استنتاجه من الجذر والصيغة الصرفية، وذلك إذا ما قورنت بالكلمات ذات الأصول الرباعية والخماسية، ويعني ذلك أن المشتقات الثلاثية ذات مرونة معجمية أعلى بكثير من مرونة المشتقات ذات الأصول الأطول؛ ويقصد بالمرونة هنا قدرة معانيها الفعلية على الابتعاد، أو الانفصام أحياناً عن المعاني المستنتجة من مبانيها الصرفية، بقول آخر فإنّ "الأصول الثلاثية هي المصدر الرئيسي للتوسّع المعجمي"[1].

وممّا يؤكد شيوع الأصل الثلاثي في اللّغة العربيّة[2]، هو وروده بكثرة في القرآن الكريم[3] واعتبار الأصل الثلاثي أبدل الأصول ليس لقلة حروفه فقط، لأنّه لو كان الأمر كذلك، لكان الثنائي وما جاء على حرف واحد أكثر استعمالاً من الثلاثي[4]. حاول ابن جني أن يجيب عن ذلك بقوله:« فتمكن الثلاثي إنّما هو لقلة حروفه- لعمري- ولشيء آخر، وهو حجز الحشو الذي هو عينه بين فائه ولامه وذلك لتباينهما ولتعادي حاليهما»[5]؛ وهناك من اللّغويين من يعتبر أنّ الثنائي هو أقدم أصل في العربية، وأن الثلاثي هو الأوسع انتشارًا[6].

❖ الإنتاجية الصرفية:

لا يمكن تطبيق أي من الصيغ الصرفية على جميع أصول المعجم، فعملية الاشتقاق محكومة بعوامل صوتية، وصرفية، ونحوية، ودلالية، بل و"إيتمولوجية" (أصل نشأة الكمات) أحياناً. تختلف الصيغ الصرفية صورة حادة من حيث قابلية تطبيقها على الأصول المعجمية المختلفة، وهي الخاصية المعروفة بـــ "الإنتاجية الصرفية"[7]، ويقصد بذلك أن عملية الاشتقاق

1- المرجع السابق، ص 275.

2- ثريا عبد الله عثمان إدريس، "الصّيغ الفعلية في القرآن الكريم"، أصواتًا وأبنية ودلالةً. رسالة مقدمة لنيل درجة الدكتوراه في اللّغة.المجلّد 1، كلية اللّغة العربية، جامعة أم القرى، المملكة العربية السعودية، 1410هـــ/1989م، ص02.

3- نجاة عبد العظيم الكوفي،أبنية الأفعال "دراسة لغوية قرآنية"، دار الثقافة للنشر والتوزيع القاهرة، د ط، 1409- 1989، ص 14.

4- المرجع نفسه، ص 13.

5- ابن جني، أبو الفتح عثمان بن جني،الخصائص، تحقيق محمد علي النجار، دار الهدى للطباعة والنشر بيروت- ج1، الطبعة الثانية، ص56.

6 -C. Brockelmann, « Précis de linguistique sémitique, traduit de l'allemand par w. Marçais et M. cohen, Librairie Paul Geuthner, Paris 1910, P: 124.

7- مع كون ظاهرة الإنتاجية الصرفية هي صلب المعجم العربي، وبلا منازع، إلا أنها لم تحظ إلا بقدر قليل للغاية من التحليل والدراسة، ولم يتجاوز الأمر في كثير من الأحيان حدود ذكر الأمثلة، والإشارة إلى بعض القيود الخاصة بتطبيق قواعد الاشتقاق. والسبب في رأيي، أن عملية الإنتاجية الصرفية تتطلب "قلباً" لمواد المعجم العربي بأكملها، لتصبح الصيغة، لا الجذر، هي البعد الأساسي، والتي تقابلها بالتالي كل الجذور التي تنطبق عليها كل صيغة من الصيغ المختلفة، ولا شك أن العملية المذكورة لقلب مواد المعجم تتجاوز طاقة العمل اليدوي من حيث الجهد والدقة، في نفس الوقت الذي تمثل فيه مجالاً نموذجياً لاستغلال إمكانيات الحاسوب الذي يمكنه القيام بعملية القلب هذه بصورة آلية دون خطأ أو نقصان. للمزيد من التفصيل ينظر: اللغة والحاسوب، مرجع سابق،ص 282.

باستخدام صيغة صرفية معينة تعتبر عالية الإنتاجية أو منخفضتها وفقاً لاتساع أو ضيق نطاق الأصول المعجمية التي يمكن تطبيق الصيغة عليها[1].

7- الصيغة الصرفية والميزان الصرفي في اللغة العربية:

7-1- الصّيغ[2]:

ونتعرف عليها من جانبين:

أ-دلالتها المعجمية:

ورد الجذر (ص و غ) في المعاجم للدلالة على معانٍ متعددة:

ب-الدلالة على الهيئة والسبك :

جاء في تاج العروس: صاغ الشيء يصونه صوغًا: "هيّأه على مثال مستقيم وسبكه عليه **(فانصاغ)**[3].

ج-الدلالة على القدر:

ذكرها صاحب الصحاح، فقال: "صغتُ الشيء أصوغه صوغًا...وهذا صوغ هذا إذا كان على قدره[4].

د-الدلالة على التقدير والمثال والصورة:

ورد في المصباح:"الصيغة العمل والتقدير وصيغة القول كذا أي مثاله وصورته"[5].

و- الدلالة على الترتيب:

ويشير إليها صاحب اللّسان قائلاً: "وصاغ شِعرًا وكلاَمًا أي وضعه ورتبه"[6].

فالصيغة في ضوء الدلالة المعجمية جملة من المعاني المتقاربة تدل على نظام خاص.

7-1-1-دلالتها الاصطلاحية: هي عبارة عن أشكال وقوالب[7] للتعبير عن المعاني.

ويرى تمام حسان: أن« الصيغة تلخيص شكلي لجمهرة من العلامات لا حصر لها ترد على ألسنة المتكلمين باللّغة الفصحى كل يوم بل في كل ثانية من دقيقة من ساعة من يوم والنّاس ينطقون العلامات ولا ينطقون هذه التلخيصات الشكلية»[8] ثم نبه على أنّه هناك فرقًا بين "الصيغة" و"الميزان".فالصيغة **(مبنى صرفي)** والميزان **(مبنى صوتي)**"[9]

1- نبيل علي، اللغة والحاسوب، مرجع سابق، ص 282.

2 -Les Formules- L'aspect.

3- الزبيدي (مرتض)، ""تاج العروس من جواهر القاموس" تحقيق. عبد الستار أحمد فراج وغيره، مطبعة الكويت. الكويت، 1965- 1978. الجزء 23- ص 533.

4- الجوهري، "الصحاح"، "تاج اللّغة وصحاح العربية"، سابق- الجزء الرابع- 4- ص 1324.

5- الفيومي، المصباح المنير، في غريب الشرح الكبير،المكتبة العصرية، بيروت- ط 1، 1996، ج 1/ ص352.

6- ابن منظور، لسان العرب، مجلد 4؛ مادة (صوغ) ص 443.

7- يرى ادريس العلمي أن الأوزان أشبه ما تكون بقوالب المصانع التي تصبّ فيها مادّة الإنتاج، فتصوغ لك منتجات على أشكالها وحجمها .

8- تمام حسان، اللّغة العربية معناها ومبناها، سابق، ص 144.

9- المرجع نفسه، ص 145.

أما مصطفى النّحاس يذهب إلى أنّ الصّيغ في اللّغة العربية:«ما هي إلاّ قوالب فكرية تصب فيها المعاني العامة تحددها وتعطيها حجمها ومعناها"[1]؛ كما أنّها مرادفة لهذه المصطلحات: "الهيئة، البناء، القالب، الصورة، الشكل»[2]، وهناك من يرى أن «الصيغة هي الوزن وهناك من يفرق بينهما على أساس أن الوزن هو أصل المادة والصيغة هي هيئة المادة بعد الزيادة أو التعديل»[3].

ويخصّها بعض اللغويين بالأبنية المقيسة في الأكثر ولها أوزانها التي لا تختلف في عمومها وغالب أمرها[4].

غير أنّ بعضهم الآخر[5] اتجه إلى تحديد المصطلحات، فذكر أن المراد بأوزان العربية: أبواب الأفعال من ثلاثية ومزيد فيها، والمراد بصيغها: أوزان الأسماء من مشتقة، وذلك أمنا للبس. ونبّه على أنّه لا فرق بين معنى "**وزن**" و"**صيغة**" وهو بخلاف الرأي الذي ذهب إليه تمام حسان.

فالصيغة تمثل شكل المادة اللّغوية وبواسطتها نتمكن من التوصل إلى المعنى الّذي يميزه عن غيره من الأشكال اللّفظية المشتقة من مادة لغوية واحدة، فالصيغة هي التي تمكننا من التمييز بين (كاتب وكتاب ومكتوب...الخ)[6].

فالصّيغ في واقعها ما هي إلاّ تشكيلات صوتية متنوعة تعرض للجذر الثلاثي (ف.ع. ل) بإضافة الصوائت أو تكرار أحد أصواته، لتأدية دلالات اللّغة المتنوعة.

7-1-2- أهمية الصّيغة الصرفية في اللغة العربية:

في ضوء التعريفات السابقة يمكن تلخيص أهمية الصّيغ:

1- إنّ هذه الصّيغ تصلح لأن تُستخدم أداة من أدوات الكشف عن الحدود بين الكلمات في السّياق...ولذلك فإنّ اتخاذ الصّيغ الصرفية أداة من أدوات خلف الحدود بين الكلمات في السّياق ميزة اللّغة العربية من كبريات ميزاتها التي تفاخر بها[7].

1- عبد العزيز عتيق، "مدخل إلى دراسة الصرف العربي"، دار النهضة، مصر، الطبعة الثانية 1974، ص 13، وينظر : مصطفى النّحاس، من قضايا اللغة، مطبوعات جامعة الكويت، الكويت، الطبعة الأولى،1415هـ/1995 م، ص 141 وما بعدها .

2- محمد المبارك، فقه اللّغة وخصائص العربية، سابق، ص 113.

3- أحمد شامية، في اللّغة دراسة تمهيدية منهجية متخصصة في مستويات البنية اللّغوية، دار البلاغ، الجزائر، ط1، 1423/ 2002، ص 33.

4- محمد سمير نجيب اللّبدي، معجم المصطلحات النحوية والصرفية، سابق، ص 128- 129.

5- عبد الرحمن شاهين، في تصريف الأسماء، مكتبة الشباب، القاهرة، د. ط، 1977م، ص 118.

6- سيدي محمد غيثري، التركيب الفعلي العربي، دراسة لسانية حاسوبية، أطروحة جامعية للحصول على درجة دكتوراه الدولة في اللّسانيات الحاسوبية، معهد اللّغة العربية وآدابها، جامعة تلمسان، السنة الجامعية- 1419- 1998، ص 148.

7- تمام حسان، مناهج البحث في اللّغة، دار الثقافة، الدار البيضاء، المغرب، د. ط، 1979، ص 210.

2- وسيلة التوليد والارتجال في اللّغة[1]، ولذلك فإن العناصر القابلة للتحول والتطور في اللّغة هي المفردات ذات الصّيغ (أي العناصر ذات الصّيغ الاشتقاقية)[2].

3- أحد العناصر الأساسية في تكوين الكلمة.

4- تخصص المعنى وتحدده.

5- الصيغة تحدد المادة الأصلية من المزيدة في الكلمة[3].

أمّا المجال الذي يتناول الصّيغ بالدراسة والبحث فلا يقتصر- كما هو شائع- على علم الصرف أو ما يعرف حديثًا بالمورفولوجي[4] وهو علم دراسة أبنية الصيغ أو الأبنية الصرفية[5]، إذ يبدئ بعض المحدثين[6] أن صّيغ الأفعال الثلاثية لا تدخل في باب الصرف وإنّما هي من مباحث متن اللّغة والمعجمات لكونها من قضايا الثروة اللّفظية.

7-2- الـوزن[7]:

7-2-1- الوزن في اللّغة: "ثقل شيء بشيء مثله كأوزان الدراهم، وقد يطلق الوزن على السنجة التي يوزن بها ويطلق الميزان على الآلة التي يوزن بها فيقال: وزن فلان الدراهم وزنًا بالميزان"[8].

7-2-2- المعنى الاصطلاحي: الوزن هو مقابلة اللّفظ بحروف الميزان- وهي الفاء والعين واللاّم لمعرفة ما فيه من حروف أصلية أو زائدة ولضبط ما في مبناه من حركات أو سكون[9] والوزن كذلك اسم يستعمل في تعداد الأشكال والصّيغ المقررة للأسماء والأفعال كأن يقال: للفعل الثلاثي المجرد ستة أوزان أي: ستة أشكال وصّيغ[10] تكفيها حركة عين الفعل[11].

1- ابراهيم بن مراد، مقدمة لنظرية المعجم، دار الغرب الإسلامي، بيروت، الطبعة الأولى، 1997،ص 160،159.

2- اللّغة العربية معناها ومبناها، سابق، ص 151.

3- ينظر: "فقه وخصائص العربية"، سابق، ص 115، 118.

4- الصرف "البنية" La Morphologie: ومن طبيعة هذه الدراسة أن تتناول الناحية الشكلية التركيبية للصّيغ والموازين الصرفية وعلاقاتها التصريفية من ناحية، والاشتقاقية من ناحية أخرى.

-ينظر: د محمد رشاد الحمزاوي " المصطلحات اللّغوية الحديثة في اللّغة العربية" معجم عربي أعجمي وأعجمي عربي- سابق، ص 88.

5- كريم زكي حسام الدّين، أصول تراثية في علم اللّغة، مكتبة الأنجلو المصرية-القاهرة- الطبعة الثانية 1985م، ص 203.

6- كمال محمد بشر، دراسات في علم اللّغة، دار المعارف، القاهرة- الطبعة التاسعة. ط9. 1986م، ص 235.

7 -Le Schème.

8- لسان العرب، مادة وزن، ج 17، ص 337.

9- معجم المصطلحات النحوية والصرفية، سابق، ص 239.

10- المرجع نفسه، ص 239.

11- المصطلحات اللّغوية الحديثة في اللّغة العربية، معجم عربي أعجمي، وأعجمي عربي، ص 107.

والوزن عند علماء الصّرف هو معرفة حركات الكلمة وسكناتها، وكان ذلك بعد أن استقرأ علماء الصرف الكلمات العربية ووجدوها تتألف من ثلاثة أصول أو أربعة أو خمسة، إلاّ أنّهم وجدوا الثلاثي هو الغالب[1]؛ فاتخذوا لذلك ميزانًا ينوه على ثلاثة أحرف ليزنوا به وحدات اللّغة العربية ذات الأصل المعجمي، فكان الوزن هو "فعل" ومكنّهم هذا الميزان الصرفي من معرفة ما في هذه الوحدات من أصول وزوائد وتقديم وتأخير وظهور وحذف[2].

فقد أراد علماء الصرف وضع مقياس موحد، تخضع له جميع المفردات العربية ويكون وسيلة دقيقة، لتحديد صيغة الكلمة من بين أنواع الأسماء والأفعال ولذلك اتخذوا مادة ثلاثية الأصول، توزن بها جميع المفردات وهي: (ف.ع.ل)[3] كوحدة قياس صرفية (قالب) ، ويرى الأستاذ إدريس العلمي: " أن هذه الأوزان أشبه ما تكون بقوالب المصانع التي تصبّ فيها مادّة الإنتاج ، فتصوغ لك منتجات على أشكالها وحجمها"[4] تقاس بها الصّيغ والأبنية اللّغوية[5].

7-2-3- تنبيه : الفرق بين الوزن والصيغة الصرفية: من خلال هذه التعاريف يمكن القول أنّ: الصيغة الصرفية هي الهيئة اللّغوية الحاصلة من الوزن، بينما الوزن هو القالب الّذي يصاغ عليها الكلمات فهو "أصل المادة"[6]. يعني أن الصيغة في اللغة العربية دلالة، والوزن معنى فهنا تكمن أهمية و دور الصيغة والوزن للكلمة العربية في تحديد معناها[7].

وورد في شرح الشافية في الميزان الصرفي: قال: "وَيُعَبَّرُ عَنْهَا بِالْفَاءِ والعَيْنِ واللاَّمِ، وَمَا زَادَ بلاَمٍ ثَانِيَةٍ وثَالثَةٍ، ويُعَبَّرُ عَنِ الزَّائِدِ بِلَفْظِهِ، إلاَّ المُبْدَلَ مِنْ تاءِ الافْتِعَالِ فَإنّه بِالتَّاءِ، وإلاَّ المُكَرَّرَ لِلإلْحَاقِ أوْ لِغَيْرِهِ فإنَّهُ تَقَدَّمَهُ وَكَانَ مِنْ حُرُوفِ الزِّيَادَةِ إلاَّ بِثَبَتٍ، ومنْ ثَمَّ كانَ حِلتيتٌ فِعْلِيلاً لاَ فِعْلِيتًا وسُحنونٌ وعُثْنون فُعْلولاً لاَ فُعْلونًا لذلك ولِعَدَمِهِ، وَسَحْنُونٌ إنْ صَحَّ الفَتْحُ فَفَعْلُونٌ

1- رمضان عبد الله رمضان، الصيغ الصرفية في العربية في ضوء علم اللغة المعاصر، مكتبة بستان المعرفة، ا الإسكندرية ،الطبعة الأولى، 2005، ص 31.

2- محمد خير الحلواني، الواضح في النّحو والصرف، قسم الصرف، دار الحامون للتراث دمشق، سوريا، ط2- 1987، ص 11.

3- فخر الدين قباوة، تصريف الأسماء والأفعال، جامعة حلب، كلية الآداب، سوريا، الطبعة الثانية، 1401-1981، ص 16.

4- إدريس العلمي، في المصطلح، الدار البيضاء، المغرب، دط،2001 مـــ ، ص 17، وينظر : مهدي صالح سلطان الشمري، في المصطلح ولغة العلم، كلية الآداب، جامعة بغداد، دط،2012، ص 47.

5- أحمد شامية، في اللّغة، دراسة تمهيدية منهجية متخصصة في مستويات البنية اللّغوية، دار البلاغ للنشر والتوزيع، الجزائر، الطبعة الأولى، 1423- 2002، ص 33.

6- المرجع نفسه، ص 33.

7- بلقاسم بلعرج، لغة القرآن الكريم، دراسة لسانية للمشتقات في الربع الاول، دار العلوم، الجزائر، دط، 2005، ص 34.

لاَ فُعْلُولٌ كَحَمْدونٍ: وهو مُخْتَصٌّ بالعَلَمِ لِندورِ فُعْلُولٍ وهو صَعْفُوقٌ، وَخَرْنوبٌ ضَعيفٌ، وسَمْنَانُ فَعْلانٌ وَخَزْعَالٌ وبَطْنَان فُعْلان، وقُرطاسٌ، ضَعيفٌ مع أنَّهُ نقيضُ ظُهْرانٍ"[1].

وطريقة الوزن في اللّغة العربية تعتمد على استبدال الحروف الأصول بالفاء والعين واللاّم، وزن (ف ع ل)، أي نضع في الوزن مكان الحروف الأصلية، هذه الحروف الثلاثة في مثل: ضَرَبَ على وزن: فَعَلَ، واتفق الصرفيون على استعمال حروف فعل دون غيرها لتكون ميزنًا صرفيًا[2] للألفاظ لكون هذه الحروف تمثل بمجموعها فعل الفاعل وهو ما ينطبق على كل فعل أو ما قام به[3].

أمّا سبب اختيار لفظ "فعل" دون غيره من الألفاظ لأمرين: الأوّل لمعرفة أصول الكلمة، وتمييز الزوائد فيها، والثاني للاختصار[4].

7-2- 4- ملاحظة:

تمثل ثنائية الصيغة والميزان أحد مواضع الخلاف[5] الحاد في دراسة الصرف العربي وتدريسه بل ومعالجته الآلية أيضًا، فهناك من ينادي بطرح فكرة الصيغة والاكتفاء بالميزان الصرفي[6]، فالدكتور نبيل علي يرى أن في هذا تجن شديد على أحد المبادئ المحورية للصرف العربي وحرمانه من أحد آليات "اختزاليته" الفعالة لاختصار قواعده وتعميقها، والرأي المنادي بالاكتفاء بأحد العنصرين سواء الصيغة أو الميزان، ليس له ما يبرره وذلك على ضوء النظرية اللّغوية الحديثة التي بنيت على مبدأ ثنائية البنية العميقة (الصيغة الصرفية في حالتنا) والبنية السطحية (ويمثلها الميزان الصرفي هنا)، ومن جهة نظرا للمعالجة الآلية، لا يوجد ما يبرر اتجاه توحيد الصيغة والميزان أيضًا، فالنموذج ذو المستويين الذي طوره "كوسكينمي" والّذي شاع استخدامه في الحالات الصرفية الآلية قد تم بناؤه على أساس التعامل مع البنية العميقة والبنية السطحية في نفس الوقت[7].

1- الشيخ رضا الدّين محمد بن الحسن الاستراباذي النّحوي، شرح شافية ابن الحاجب مع شرح شواهده، تحقيق الأساتذة: محمد نور الحسن، محمد الزفزاف، محمد محي الدين عبد الحميد، القسم الأوّل- الجزء الأول- دار الكتب العلمية لبنان- 1982- ص 10-11.

2- الميزان الصرفي: هو ما وضعه الصرفيون من وسيلة لوزن الكلمات ومعرفة أصولها من زوائدها وحركاتها من سكناتها وقد اختاروه مؤلفًا من الفاء والعين واللاّم على أن يقابل الحرف الأوّل الأصلي في اللّفظ الموزن بالفاء من الميزان ويسمى فاء الكلمة ويقابل الثاني بالعين ويسمى عين الكلمة ويقابل الثالث باللاّم ويسمّى لام الكلمة وما يضعف من عين الكلمة أو لامها يضعف مقابلة في الميزان، ينظر: د. محمّد سمير نجيب اللّبدي "معجم المصطلحات النحوية والصرفية" ص 239- 240.

3- معجم المصطلحات النحوية والصرفية، مرجع سابق، ص 240.

4- عبد الحليم عبد الباسط محمد المرصفي، من صيغ وأوزان العربية "أفعل" دار مرجان للطباعة- القاهرة، ط1، 1978م، ص 14- 15.

5- ينظر: د أحمد شامية،"في اللّغة"، دراسة تمهيدية منهجية متخصصة في مستويات البنية اللّغوية، سابق، ص 33.

6- إنّ صاحب فكرة الميزان الصرفي في العربية هو الخليل بن أحمد الفراهيدي (175هــ) فالخليل قام بدراسة الكلمات المعروفة من المواد نفسها تحديد حروفها الأصول، و ينظر: د. محمود فهمي حجازي، "مدخل إلى اللّغة"، دار قباء للطباعة والنشر والتوزيع، القاهرة طبعة جديدة مزيدة ومنقحة، 1997، ص 98.

7- اللّغة العربية والحاسوب، ص 281- 282

8- أهمية الميزان الصرفي:

يعدّ الميزان الصرفي[1] دُرّةَ الصرف العربي، إذ تنصب الدراسة الصرفية للصيغ على تحديد الشكل الخارجي لها ووظيفتها ووبيان أوجه الإتفاق والإختلاف بينهما[2]، هذا النموذج القائم في أساسه على التمييز بين الحروف والحركات من جهة، وعلى الجذر (الحروف الأصلية) وحروف الزيادات الصرفية من جهة أخرى[3]؛ وهو يمثل عملاً عربيًّا خالصًا لا يشترك معهم فيه أحد بل مثّل عَملهم هذا تحديًّا علميًّا للدراسات الغربية التي لم تألف في الصرف هذا النمط من التجريد فهو يعدّ مفتاح فهم طبيعة بنية الكلمة لا في العربية فحسب بل في اللّغات السّامية الأخرى أيضًا[4].

يتمثل أهمّ ما ينتهي إليه هذا الجانب من هذه الدراسة في مناقشته لثنائية الصيغة الصرفية والميزان الصرفي فيما يلي[5]:

(1)- اقتضاء الوزن الذي ترد عليه الكلمات في العربية، ويمثل دالة صرفية لنموذج الميزان الصرفي، حيث لا يمكن أن تقدم أوزان الكلمات من خلال نموذج آخر.

(2)- كشف الميزان الصرفي عن دقّة اللّغويين العرب في التمييز بين التفسيرات الصرفية.

(3)- كفاءة الميزان الصرفي بانفراده دون غيره من نماذج التحليل الأخرى في التفريق بين العناصر الصرفية والعناصر النّحوية في التراكيب اللّغوية.

(4)- كفاءة نموذج الميزان الصرفي في تَكَفُّلِهِ باستغراق تغييرات التركيب الصرفي للكلمات التي لا يستغرقها منهجَا الجداول التصرفية والعلامة بواسطة الميزان الصرفي نستطيع معرفة أصول الكلمات وما يطرأ عليها من زيادة أو حذف أو إعلال (Défection)[6].

1- لم يكتف بعض الباحثين بالميزان الصرفي بل قاموا بتعريف مقاييس آخرى مثل : النمط الصرفي (أو الصيغة الصرفية) ويعرفونه بأنه القالب الذي يشمل الحركات وحروف الزيادة (حروف سألتمونيها) ومواضع حروف الجذر بتسلسل ورودها داخل القالب . أما الهيكل الصرفي فيعرفونه بالصيغة الصرفية دون تشكيل . وعلى خلاف الميزان الصرفي فإن الصيغة الصرفية تشمل دائما الحروف (ف،ع،ل) . ولمعرفة العلاقة بين الصيغ والموازين الصرفية فيمكن القول أن الصيغة الصرفية تمثل "البنية العميقة." أما الميزان الصرفي فيمثل "البنية السطحية" حيث يمكن استنتاج الميزان الصرفي من الصيغة الصرفية بعد إجراء عمليات الإبدال والإعلال عليها , وكمثال على ذلك فكلمة (سماء) ميزانها الصرفي (فعاء) وجذرها (س،م،و) أما صيغتها الصرفية فهي (فعال) حيث قبلت الواو المتطرفة همزة عن طريق عمليات الإبدال والإعلال. ومن جهة أخرى فإن الصيغ الصرفية تختلف فيما بينها من حيث إمكانية تطبيقها على مختلف الجذور .

تمثل هذه الصيغ مقاييس لغوية لنمذجة الكلمات من حيث حروفها الأصلية والزائدة وعلامات التشكيل كما تمثل دوالاً مستخدمة تطبق على الجذور لاشتقاق كلمات منها . تمكن هذه المقاييس الباحث من إيجاد قواعد منضبطة في التعامل مع الصرف العربي وبالتالي التحكم بجميع الحالات المؤثرة على تكوين واشتقاق الكلمات مع مراعاة جميع التغيرات الصرفية والنحوية والصوتية .

2- عبد المقصود محمد عبد المقصود، دراسة البنية الصرفية في ضوء اللسانيات الوصفية، الدار العربية للموسوعات، بيروت،الطبعة الأولى،1427هـــ، 2006مـــ ، ص 101.

3- محمد عبد العزيز عبد الدّايم،نظرية الصرف العربي، دراسة في المفهوم والمنهج، حوليات الآداب والعلوم الإجتماعية، دورية علمية محكمة، الرسالة 185، الحولية الحادي والعشرون، كلية دار العلوم،2001 مـــ، ص 47.

4- اللّغة العربية والحاسوب، سابق، ص 191.

5- محمود فهمي حجازي، "مدخل إلى علم اللّغة"، دار قباء للطباعة والنشر والتوزيع، القاهرة، طبعة جديدة مزيدة ومنقحة، 1997، ص96- 97.

6- أنطوان الدحداح، "معجم قواعد اللّغة العربية في جداول ولوحات" زائد مسرد بالمصطلحات النحوية، عربي- انجليزي- فرنسي"، مكتبة لبنان- ناشرون، الطبعة رقم 01- 1997- ص 31.

8-1- الأسرة اللفظية والأوزان الصرفية:

إن خاصية الاشتقاق، التي تمتاز بها العربية، تزود المعجمي بأداة فعالة في تسيير إمداد القارئ بالمعلومات الدلالية. فكل مفردة في الأسرة اللفظية[1] تصاغ وفق وزن صرفي له معنى معلوم شعورياً أو لا شعورياً لدى الناطقين باللغة العربية.

9- الأصول والجذور في اللغة العربية:

9-1- الأصل والفرع :

الأصل[2] في اللّغة العربية هو مجموع المادة التي يتكون منها، وهو يمثّل العناصر الأصلية التي تكون مصدر اشتقاق الصيغ المختلفة[3].

فقد اعتبر الصرفيون أن الأصل والفرع في دراسة الظواهر الصرفية كظاهرة الإشتقاق[4]، إذ عرَف الرماني (296 هـ/909م- 384هـ/994م) الإشتقاق بأنه اقتطاع فرع من أصل يدور في تصاريفه على الأصل[5]، وبين العبكري أن الأصل ههنا يُراد به الحروف الموضوعة على المعنى وضعا أوليا[6]، والفرع لفظ يوجد في تلك الحروف مع نوع تغيير ينضم إليه معنى زائد على الأصل[7]، يقول ابن يعيش: "الأصل الحروف اللازمة للكلمة كيف تصرفت"[8]، وهو الّذي تتولد منه الألفاظ أشبهه ما يكون بالرابطة النسبية بين النّاس[9] وهو أساس الشيء- فيما يبنى عليه غيره[10].

وللأصل في عرف المعجميين معنيان: **الأوّل** بمعنى الفصيح؛ و**الآخر** بمعنى ما أخذ منه اللّفظ أي مصدره، بمعنى أدق المعنى الرئيسي الذي يؤخذ منه بقية المعاني[11].

1- تجرد الإشارة إلى أن المعنى العام للأسرة اللفظية ليس مطابقاً تماماً للمعنى الأصلي الجذر، فالجذر يمكن أن يكون له أكثر من معنى أصلي واحد، للمزيد من التفصيل ينظر:علي القاسمي،المعجمية العربية بين النظرية والتطبيق، مكتبة لبنان ناشرون، بيروت-لبنان، الطبعة الأولى، 2003، ص 78، 79.

2- الأصل: La Racine: إنّ الأصل "يعطي المعنى العام للكلمة"، للمزيد من التفصيل ينظر: د/ محمد رشاد الحمزاوي (المصطلحات اللّغوية الحديثة في اللّغة العربية)، سابق، ص 24.

3- محمود فهمي حجازي، مدخل إلى علم اللغة، دار قباء للطباعة والنشر والتوزيع، القاهرة، دط، دت، ص 96.

4- حسن خميس الملخ، نظرية الأصل والفرع في النحو العربي، دار الشروق،عمان- الأردن، الطبعة الأولى، 2001، ص 20..

5- أبو الحسن علي بن عيسى بن علي بن عبد الله الرماني، الحدود، تحقيق إبراهيم السامرائي، دار الفكر – عمان،1984 ، ص 3.

6- العُكبري عبد الله بن الحسين بن عبد الله أبو البقاء، مسائل خلافية في النحو، تحقيق الدكتور عبد الفتاح سليم،مكتبة الآداب، القاهرة، الطبعة الثالثة،2007، ص 62.

7- المرجع نفسه، ص 62

8- ابن يعيش موفق الدين يعيش بن علي، شرح الملوكي في التصريف، تحقيق فخر الدين قباوة،المكتبة العربية، حلب، الطبعة الأولى، 1972، ص 108.

9- أبو بكر محمد بن الحسن بن دريد، جمهرة اللّغة، الجزء الأوّل، طبعة حيدر آباد 1344هـ- ص 1/ 13.

10- معجم المصطلحات النحوية والصرفية، سابق، ص 11.

11- حُسام قدوري عبده، تأصيل الجذور السامية وأثره في بناء معجم عربي حديث، دار الكتب العلمية، بيروت، الطبعة الأولى 1428- 2007، ص 26.

والخليل بن أحمد أوّل من أخضع هذه الأصول[1] إلى نظام التقاليب في معجم "العين"، حيث كان يبحث الكلمة ويعرض إلى الصّور الممكن تكوينها من أصول هذه الكلمات[2] والأصول كما وردت في كتب التصريف تنقسم إلى نوعين هما الأصل الثلاثي والرباعي وهو أقل استعمالاً من الأصل الثلاثي[3]. والمقارنة بين أوزان الثلاثي والرباعي تؤكد أن الأصل الثلاثي هو أعدل الأصول تركيبًا وأكثرها استعمالاً[4].

وربّما كان الأصل الثلاثي أعدل[5] الأصول لتوسطه بين أصلين ثنائي مهمل يشبه الجذور المطمئنة تحت الثرى، ورباعي مستعمل متفرع منه[6].

والملاحظ أن هناك اختلافات حول الأصول تجعلنا نميل إلى الأصل الثلاثي[7] لشيوع استعماله، ولتوسطه بين الأصلين الثنائي والرباعي، وهذا لا يعني أنّه الأصوب إلاَّ أنّنا نعتقد أن ما وصلت إليه الدراسات حول هذه المسألة هو استمرارية الاختلاف بين الباحثين، فكل فريق يبحث عن مبرّرات لتحرير رأيه، وتقديم ما يبرر به رفضه للرأي الآخر[8].

1- لقد لاحظ الخليل في دراسته لبعض الكلمات أنّ أسماء لفظها جاءت على حرفين وتمامها على ثلاثة أحرف، مثل: يد ودم وفم، فالخليل يتجاوز الثنائية التي تبدو في هذه الكلمات ليحاول دراسة الكلمات المعروفة من المواد نفسها تحديد حروفها الأصول، ومنهجه واضح في هذا فهو يقول:"فإذا أردت معرفتها فاطلبها في الجميع والتصغير). وهنا نلاحظ أن كلمة يد وردت إلينا في كلمات أخرى مثل: أيديهم ويديّه (بتشديد الياء الثانية وفتحها)، وهنا يلاحظ الخليل أن هذه الكلمة لابدّ أن أصولها ثلاثية، وأنّها تكونت في الأصل من الياء والدّال والياء فهي على هذا ثلاثية الأصول وما يبدو منها في كلمة "يد" ليس كل الأصول بل يظهر حرفان أصليان من ثلاثة، وشبيه بهذا ما فعله الخليل بن أحمد في دراسته لكلمة "دم"، فقد لاحظ وجود صيغ في هذه المادة وبها أصول ثلاثة بدت له هي الحروف الأصول. ودرس الخليل كذلك "فم" ولاحظ التثنية:"فموان" والجمع "أفواه"، فاستقرّ لديه أنّ الأصل هنا ثلاثي مكون من الحروف ف و م، ينظر: د. محمود فهمي حجازي" مدخل إلى علم اللّغة"،سابق، ص98.

2- د. سيدي محمد غيثري. التركيب الفعلي العربي. دراسة لسانية حاسوبية، سابق، ص 109.

3- المرجع نفسه، ص 109، ينظر: د. رمضان عبد التواب،"التطور النّحوي للغة العربية"، محاضرات ألقاها في الجامعة المصرية عام 1929 المشترق الألماني برجشتراسر G.Bergsträsser، مكتبة الخانجي، القاهرة، الطبعة الثانية، 1417هـ- 1997 م، ص 103.

4- نجاة عبد العظيم الكوفي، أبنية الأفعال، دراسة لغوية قرآنية، دار الثقافة للنشر والتوزيع، القاهرة، د ط. 1409- 1989، ص 13.

5- على مستوى المعجم، يفوق عدد الجذور الثلاثية، وبكثير، عدد الجذور الرباعية والخماسية، ويزداد الانحياز نحو الثلاثية في معدلات تواتر الجذور المختلفة داخل النصوص العربية بغض النظر عن موضوعها، أو أسلوب كتابتها، ويندر كذلك وجود كلمات مقترضة ذات أصل ثلاثي في "العربية"، للمزيد من التفصيل ينظر :نبيل علي،اللغة والحاسوب،ص 275،274.

6- نجاة عبد العظيم الكوفي، أبنية الأفعال، دراسة لغوية قرآنية، مرجع سابق، ص 14.

7- ترتكز بنيانية المعجم العربي على أصول ثلاثية ولئن كانت هناك جذور ثنائية فهي لم تشكل سوى مرحلة تاريخية من مراحل تطور اللغة العربية فقد تم تحول بنيانية المعجم العربي إلى أصول ثلاثية بفعل تحولات داخلية بحتة نذكر منها المد والتضعيف وزيادة حرف واحد على الحرفين الأصليين مما أدى إلى أصول بلغت حد الاكتناز ويكون المنطلق البنياني لمعظم كلمات المعجم العربي هو الأصل الثلاثي ويحدث ذلك نظرياً على الشكل التالي: 28 × 27 × 26 (بإهمال تنوع حركات عين الأصول الثلاثية) 19656 وقد يحد من فعالية هذا العدد الهائل من الأصول مبدأ التنافر الصوتي والاستعمال وقد تكتفي اللغة بعدد صغير من الجذور (3000 تقريباً) يتم بموجبها وضع معظم الكلمات العربية. تتولد الكلمات من الأصول الثلاثية في إحداثيتين (abscisse et ordonnée) هما إحداثية مزيدات الأصول الثلاثية وإحداثية مشتقات الأصل ومزيداته أي خط الصفات والأسماء وكلما تقاطع خطان منبثقان عن الإحداثيتين تولدت مفردة قد تدخل إلى حيز الاستعمال المباشر أو قد تبقى في حيز الكمون تنتظر الفرصة الملائمة لتشق لنفسها طريق الوجود. وخلاصة القول أن معظم الكلمات في اللغة العربية ينشأ عن أصول ثلاثية (ثلاث حروف صامتة وغير مصوتة) هي حجر الزاوية في إقامة صرح التنظيم الرياضي اللغوي المتكامل. للمزيد من التفصيل ينظر : ريمون طحان، الألسنية العربية- مقدمة- الأصوات- المعجم- الصرف، دار الكتاب اللبناني، بيروت،الطبعة الثانية،1981، ص 79،78.

8- توفيق محمد شاهين، أصول اللّغة العربية بين الثنائية والثلاثي، مكتبة وهبة، الطبعة 1، 1980، ص 92.

يمكن الإشارة إلى أن هناك جملة من الطرق التي استعان بها الصرفيون لمعرفة الأصل، يقول ابن جني:"الأصل عبارة عند أهل الصناعة عن الحروف التي تلزم الكلمة في كل موضع من تصرفها إلا أن يحذف شيء من الأصول تخفيفًا أو لعلة عارضة فإنّه لذلك في تقدير الثبات"[1] فما يثبت من حروف الكلمة حال تقليبها وتصريفها هو الأصل منها، وما يعرض عليه التغيير هو الزّائد[2].

9-2- الجذور:

إنّ اللّغة العربية هي لغة ذات جذور[3] كغيرها من الساميات، ويغلب على جذورها أنّها ثلاثية الصوامت وقد تكون ثنائية أو رباعية ويلعب الاشتقاق دورًا كبيرًا في توليد هذه الجذور وتوسيعها مع الاحتفاظ بالصوامت الأصلية (الجذر).

فالجذر[4] في اللّغة العربية يتكون من الحروف الأصلية للوحدة المعجمية التي تتكون منها الوحدات اللّغوية الدّالة في حالتها المجردة من كلّ زيادة/ ضرب/ في "مضرب"، ولابدّ من الإشارة إلى الحدود المأخوذة في الوحدات الدالة، أي: (ما لا يقصد بجزء منه الدلالة على جزء معناه)، والمقصود من/ضرب/أنّ أجزاءه هي ذوات حروفه الثلاثة التي تتمثل في /ض ر ب/ أي أن هذه الوحدات غير مقصود بها الدلالة ولا تدل على معنى[5].

فالجذور إذن تتكون من الحروف الأصلية، وقد رأينا أن الحروف الأصلية بدورها يدور حولها الخلاف في عددها وفي طبيعتها وذلك راجع لعدم الاستقرار[6]. والجذور هي نفسها لم تسلم من هذا الاختلاف فمنهم من يراها هي والأصل بمعنى واحد[7] .

فالجذر في اصطلاح المعجميين (الأحرف المشتركة بين عدد من الكلمات، يعتقد بأنّها تتصل ببعضها اتصالاً اشتقاقيًا، وقد يتعذر لفظ هذه الأحرف في كثير من الحالات لأسباب

1- أبي الفتح عثمان بن جنّي "التصريف الملوكي"، عُني بتصحيحه محمد سعيد النعسان، مطبعة شركة التمدّن الصناعية، مصر، الطبعة الأولى،د،ت، ص 05.

2- للمزيد من التفصيل ينظر: د حسام قدوري عبد، "تأصيل الجذور السّامية وأثره في بناء معجم عربي حديث"، سابق، ص 74 وما بعدها.

3- عبد الصبور شاهين، "في التطور اللّغوي"، مكتبة دار العلوم، القاهرة، الطبعة الأولى 1975 م، ص 31.

4 -Radical, Racine.

5- جمال الدين عبد الله ابن أحمد ابن علي ابن محمد الفاكهي، شرح الحدود النّحوية، حققه وقدمه الدكتور محمد الطيب الإبراهيم، دار النفائس، بيروت، الطبعة الأولى، 1996، ص 68.

6- أسعد أحمد علي "تهذيب المقدمة اللّغوية للعلايلي" – دار السؤال للطباعة والنشر. دمشق الطبعة الثانية، 1981- ص 114.

7 -Henri Fleich, traité de philologie Arabe ; Vol. 1, Imprimerie catholique, Beyrouth, 1961. P:247,248,249.

وينظر: - ينظر: أحمد شامية،"في اللّغة" دراسة تمهيدية منهجية متخصصة في مستويات البنية اللّغوية، سابق ص34.

صوتية فيُعمد إلى إضافة أصوات علة بينها لتسيير لفظها)[1] فالجذر هو الذي يحمل المعنى الأساسي للكلمة.

9-2-1 المعنى الأصلي للجذر:

لقد استند اللغويون والمعجميون العرب القدامى إلى نظرية دلالية مفادها أن كل جذر، ثنائياً كان أو ثلاثياً أو رباعياً، يحمل معنى أصلياً عاماً (وأحياناً أكثر من معنى أصلي عام واحد، واقترح أن تسمى هذه الظاهرة بالاشتراك الجذري تمييزاً لها عن ظاهرة الاشتراك اللفظي، إن لم تكن سميت بهذا الاسم من قبل)، ويتجلى المعنى الأصلي للجذر[2]

1- حسام قدوري عبد، "تأصيل الجذور السّامية" وأثره في بناء معجم عربي حديث، ص 26.

2- يمكن أن نعد نظرية المعنى الأصلي للجذر نظرية متينة لها تطبيقات واسعة في المعجم العربي. وإلى جانب هذه النظرية، شاعت نظرية أخرى يمكن نعتها بالركيكة لقلة تطبيقها، وصعوبة تسويغها، تلك هي نظرية الاشتقاق الأكبر التي تزعم أن تقليبات الجذر المختلفة تشترك في معنى أساسي عام. فالجذر الذي ذكرناه آنفاً (ع ب ر) له ستة تقليبات إذا أخذنا كل حرف من الحروف وأتبعناه بالحرفين الآخرين، مرة متبوعاً بالحرف الثاني ومرة متبوعاً بالحرف الثالث. وهذه التقليبات هي: ع ب ر، ع ر ب، ب ع ر، ب ر ع، ر ع ب، ر ب ع.

وكان الخليل بن أحمد الفراهيدي أول من استعمل تقليبات الجذر في صناعة المعجم العربي. وكانت غايته حصر جميع الألفاظ المستعملة والمهملة في اللغة العربية، والتأكد من شمولية معجمه (العين)، إذ يقول في مقدمة معجمه: "هذا ما ألفه الخليل بن أحمد البصري من حروف: أ ب ت ث، مع ما تكملت به فكان كلام العرب وألفاظهم، فلا يخرج منها عنه شيء.

ويرى الأستاذ هيوود أن الخليل بن أحمد لم يروج لنظرية المعنى الأساسي المشترك لتقليبات الجذر الواحد، ولكن بطريقة التقليبات التي اتبعها في ترتيب مادة معجمه (العين) ربما أعطت الانطباع للغويين والمعجميين في عصور تالية بأن تقليبات الجذر الواحد مرتبطة دلالياً كما هي مرتبطة صوتياً، أو حتى الاعتقاد "بوجود بعض السحر في تشكيلات الحروف".

وفي حقيقة الأمر، فإن صاحب هذه النظرية هو اللغوي الكبير أبو الفتح ابن جني (المتوفى عام 1002م)، الذي عزاها لنفسه وأطلق عليها اسم الاشتقاق الكبير في مقابل الاشتقاق الصغير الذي يقتصر على اشتقاق الألفاظ من جذر واحد من دون تقليباته. وكان المثل الذي ضربه ابن جني على نظرية الاشتقاق الكبير معنى الجذر (ق و ل) في تقليباته الستة: ق و ل، ق ل و، و ق ل، و ل ق، ل ق و، ل و ق.

وإذا كانت نظرية المعنى الأساسي المشترك لتقليبات الجذر الواحد (أو نظرية الاشتقاق الكبير، كما يسميها ابن جني) لم تلق صدى كبيراً في مصنفات المعجميين العرب، فإن نظرية المعنى الأصلي للجذر حققت نجاحاً باهراً، ووجدت لها تطبيقات عملية في عدد من المعاجم في طليعتها المعجم القيم الموسوم بـــ (المقاييس) الذي صنفه ابن فارس. فكل مدخل في هذا المعجم يبدأ بالمعنى الأصلي للجذر، قبل أن تسرد الكلمات الكلمات المشتقة منه ومعانيها. ولنضرب مثلاً على ذلك بالمدخلين (ب س ط) و (ب ر أ) في معجم (المقاييس):

"(ب س ط): الباء والسين والطاء أصل واحد، وهو امتداد الشيء في عرض أو غير عرض. فالبساط ما يبسط. والبساط الأرض وهي البسيطة. يقال مكان بسيط وبساط. قال:

ودون يد الحجاج من أن تنالني بساط لأيدي الناعجات عريض

ويد فلان بسط إن كان منافقاً، والبسطة في كل شيء السعة. وهو بسيط الجسم والباع والعلم. قال الله تعالى: ﴿وزادهم بسطة في العلم والجسم﴾.الآية 247 من سورة البقرة.

أما مدخل (ب ر أ) فنجد فيه:

"(ب ر أ): الباء والراء والهمزة فأصلان إليهما ترجع بعض فروع الباب: أحدهما، الخلق، يقال: برأ الله الخلق يبرؤهم برءاً. والبارئ الله جل ثناؤه ...

والأصل الآخر: التباعد من الشيء ومزايلته، من ذلك البرء، وهو السلامة من السقم، يقال برئت وبرأت ... من المرض. ومن ذلك برأت إليك من حقك ... أنا بريء منك ...".

وتجدر الملاحظة إن إعطاء المعنى الأصلي للجذر قبل التطرق إلى معاني مشتقاته وسيلة معينة على التذكر، في زمن كانت فيه المعاجم تحفظ عن ظهر قلب من قبل الدارسين والمتخصصين، ولم تُقتنَ لتوضع على الرف للرجوع إليها عند الحاجة فقط. للمزيد من بسط القول في هذه النقطة ينظر : الخليل بن أحمد الفراهيدي، كتاب العين، تحقيق مهدي المخزومي وإبراهيم السامرائي بغداد: دار الرشيد للنشر، 1980، ج1: ص 47، و وينظر :. ابن جني، الخصائص، تحقيق محمد علي النجار (بيروت: دار الكتاب العربي، ب ت) ج1، ص 5.

في جميع الألفاظ المشتقة من ذلك الجذر. وإضافة إلى ذلك، فإن كل لفظ من تلك الألفاظ له معنى خاص به[1]. وعلاوة على ذلك، فإن كل لفظ يصاغ على وزن معين من أوزان الصرف العربي، والوزن ذاته يحمل معنى خاصاً به. وهكذا يتكون معنى اللفظ من حاصل الجمع الدلالي، لا الحسابي، للمعاني الثلاثة. وبعبارة أخرى[2]:

معنى اللفظ = المعنى الأصلي للجذر + المعنى الصرفي + المعنى الخاص للفظ أو:

م ل = م أ + م ص + م خ.

ملاحظة:

تشير بعض الدراسات الحديثة إلى أنّ هناك عدّة مصطلحات متقاربة المعنى، وتؤدي نفس الغرض، لكلمة "الجذر".

- القاعدة[3]
- الجذع[4]

1-القاعدة: (Base) هي ما تبقى من الكلمة بعد حذف الزوائد (Affixes) الإعرابية.

2-الجذع: (Racine)[5]: وهي وحدة خيالية تحتوي على المعنى المشترك لكل الكلمات التي تتكون من هذا الجذع[6].

3-الجذر (Radical): وهو ما تبقى من الكلمة بعد حذف الزوائد[7].

1- علي القاسمي،المعجمية العربية بين النظرية والتطبيق،مكتبة لبنان ناشرون، بيروت-لبنان، الطبعة الأولى، 2003، ص 78، 79.

2- المعجمية العربية بين النظرية والتطبيق، المرجع نفسه، ص 79.

3 - Base.

4 -Racine.

5 -Delphine Bernard, « apprentissage de connaissances morphologiques pour l'acquisition automatique de ressources lexicales », thèse doctorat, sciences cognitives, université Grenoble I, Faculté de science de la langue 2006, P: 194.

6- تتخذ معظم المعاجم اللغوية من جذع الكلمة مدخلا أساسيا لها، ويقصد بالجذع، هنا، الكلمة خالية من الزوائد التصريفية، وأدوات التأنيث والتعريف، وعلامات الإعراب، وكل ما يتصل بها من ضمائر متصلة (أمثلة: "يستخرج"، "والد"، "انبهار").

نظرا للخاصية الصرفية العالية للغات السامية، مثل "العربية" و"العبرية"، استخدم "الجذر"، (أو ما يسمى بأصل الاشتقاق) بدلا من "الجذع" كأساس لتنظيم معاجمها، وتجدر الإشارة هنا إلى أن البعض قد نادى في تنظيم معاجم اللغات السامية إلى الجمع "الجذر" (أو ما يسمى بأصل الاشتقاق) بدلا من "الجذع" كأساس لتنظيم معاجمها، والصفات فعلى أساس الجذوع. للمزيد من بسط القول ينظر : اللغة العربية والحاسوب، دراسة بحثية، نبيل علي، ص 460.

7 -Oxford dictionary, Oxford ; university press ; 2006 , P: 116.

9-3- أهمية الأصول الثلاثية في المعجم العربي:

ترتكز بنيانية المعجم العربي على أصول ثلاثية[1] ولئن كانت هناك جذور ثنائية فهي لم تشكل سوى مرحلة تاريخية من مراحل تطور اللغة العربية فقد تم تحول بنيانية المعجم العربي إلى أصول ثلاثية بفعل تحولات داخلية بحتة نذكر منها المد والتضعيف وزيادة حرف واحد على الحرفين الأصليين مما أدى إلى أصول بلغت حد الاكتناز ويكون المنطلق البنياني لمعظم كلمات المعجم العربي هو الأصل الثلاثي[2] ويحدث ذلك نظرياً على الشكل التالي[3]:

28 × 27 × 26 (بإهمال تنوع حركات عين الأصول الثلاثية) = **19656**

وقد يحد من فعالية هذا العدد الهائل من الأصول مبدأ التنافر الصوتي والاستعمال وقد تكتفي اللغة بعدد صغير من الجذور (3000 تقريباً) يتم بموجبها وضع معظم الكلمات العربية.

تتولد الكلمات من الأصول الثلاثية في إحداثيتين(abscisse et ordonnée) هما إحداثية مزيدات الأصول الثلاثية وإحداثية مشتقات الأصل ومزيداته أي خط الصفات والأسماء وكلما تقاطع خطان منبثقان عن الإحداثيتين تولدت مفردة قد تدخل إلى حيز الاستعمال المباشر أو قد تبقى في حيز الكمون تنتظر الفرصة الملائمة لتشق لنفسها طريق الوجود. وخلاصة القول أن معظم الكلمات في اللغة العربية ينشأ عن أصول ثلاثية (ثلاث حروف صامتة وغير مصوتة) هي حجر الزاوية في إقامة صرح التنظيم الرياضي اللغوي المتكامل[4].

إن القول بوجود بعض الأصول الرباعية في اللغة العربية قضية هامة تحتاج إلى دراسة مستوعبة. لم تستثمر اللغة العربية حتى الآن الإمكانيات الكامنة والناجمة عن الأصول الثلاثية ولو أنها استثمرت كافة إمكانيات الأصول الرباعية النظرية لأفضى بها الأمر إلى لغة رمزية تفوق فيها وسائل التعبير المفاهيم التي قد يستوعبها الفكر البشري ونستطيع أن نقوم بعملية حسابية بسيطة تبين لنا عدد الأصول النظرية التي قد تنشأ عن استثمار الأصول الرباعية في اللغة العربية[5]:

1- أنطوان عبدو، مصطلح المعجمية العربية، الشركة العالمية للكتاب، بيروت، الطبعة الأولى، 1991، ص 122.

2- أنطوان عبدو، مصطلح المعجمية العربية، سابق، ص 128.

3- المرجع نفسه، ص 123.

4- وقد أحصى العلماء الجذور الحقيقية التي هي موجودة في الاستعمال وهو ما يسميه الخليل بالمستعمل (من التقاليب الحرفية). فقد ذكر للزبيدي في مختصر العين هذه الأرقام : تبلغ جميع الجذور المستعملة في اللغة العربية: 5626 جذرا وتنقسم إلى : ثنائي وتبلغ جذوره 489، وللثلاثي 4261 جذرا، وللرباعي 820 جذرا، وللخماسي أخيرا42 جذرا (ذكره المزهر،1/75-76). للمزيد من بسط القول ينظر: عبد الرحمن الحاج صالح، منطق العرب في علوم اللسان، منشورات المجمع الجزائري للغة العربية، ص 220.

5- أنطوان عبدو، مصطلح المعجمية العربية، الشركة العالمية للكتاب، بيروت، الطبعة الأولى، 1991، ص 123.

$28 \times 27 \times 26 \times 25 = ... 4914$.

تضاف إلى هذا العدد الهائل من الجذور مشتقات الرباعي!.. ويتبين لنا من العودة إلى واقع اللغة العربية أن الأصول الرباعية تتمتع بميزات تشبه تلك التي تتمتع بها الأصول الثلاثية أي حق التفريع والتوليد بموجب إحداثيتين ثابتتين هما إحداثية المزيدات الفعلية وإحداثية المشتقات الخاصة بالأسماء والصفات. لا تعيش لغة بمعزل عن غيرها فكان من الطبيعي أن تتعرض لغة العرب للتأثر، فتتسرب إليها كلمات أجنبية، تتأقلم وتتغير أبنيتها وتعرب وتفقد عجمتها وتحاكي الكلمات العربية الأصل. وفي هذا الأمر نفع محقق يؤدي إلى توسيع شبكة مفردات اللغة وإلى تنمية مواد حقولها المفهومية وهو يتيح الفرصة للباحثين لكي يعثروا على مصطلحات جديدة تدعو إليها الحاجة الأدبية وبعض متطلبات الحضارة الحديثة (هناك رأي خاص في طرق وضع المصطلحات العلمية ولا نتكلم هنا إلا عن المصطلحات الحضارية)[1].

10- اللواصق في العربية:

ونتعرف عليها من جانبين:

10-1- المعنى اللّغوي: إذا التمسنا الجذر اللّغوي لمادة [ل ص ق] في المعجمات اللّغوية نجد أنّها من (لَصِقَ يَلْصِقُ لُصوقًا)[2]، ويقال (لَصِقَ الشيءَ بالشيء يَلْصِقُ لُصوقًا)[3]. إذ دل على ملازمة الشيء للشيء[4]، واللّصوق دواءٌ يُلْصَقُ بالجرح[5].

10-2- المعنى الاصطلاحي: فهو عملية إضافة اللّواصق إلى الجذور التي تظهر في اللّغات الهندوأوربية[6] بواسطة لواصق تلصق بالجذر، وهذه اللّواصق تقع على شكل سوابق، أو مقحمات، أو لواحق، لتبين وظيفة قواعدية[7].

وجملة القول: أنّ اللّصق هو ملازمة الشيء للشيء، ثم تحولت اللّفظة من المعنى المذكور إلى اصطلاح خاص، يطلق على مجموعة معينة من اللّغات الإنسانية إذ تسمى باللّغات

1- انظر: Raymond TAHHAN, Formation des sciences et Langue Arabe,– Travaux et Jours – N° 31 Avril – Juin 1969 – Beyrouth.

2- الخليل بن أحمد الفراهيدي، كتاب العين: تحق، د. عبد الحميد هنداوي، منشورات محمد علي بيضون، دار الكتب العلمية، بيروت، ط1، 1424هـ/ 2003 م، 5/ 64.

3- ابن منظور، لسان العرب، 369/10، والمصباح المنير: 216/2.

4- ابن فارس، معجم مقاييس اللّغة، مصدر سابق، 5/ 248.

5- لسان العرب، 369/10.

6 –Sapir, Edward. Language: An Introduction to the study of speech, New York: Harcourt, Bace & world, 1921, P: 92.

7 –Sohnlyons « Introduction », In New Horizons in Linguistics, Penguin Books, UK ; P: 96.

اللّصقية، لأن تفسير معنى الكلمة الأصل (الجذر) يشار إليه بمكونات لغوية تلتصق به فتوضع هذه المكونات إمّا قبله، أو وسطه، أو بعده[1].

10-3- المكونات الأساسية للواصق في العربية:

تعتمد عملية اللّصق[2] على مكونات لصقية أساسية تعرف بـــالسّوابق[3] والمقحمات[4] واللّواحق[5] في تحديث وحدات لغوية جديدة[6]، وقبل تحديد هذه العناصر الثلاثة لابدّ من القول أنّ الباحثين العرب ترجموا (Suffixes, Infixes, Préfixes) إلى مصطلحات متعددة، إذ أطلقوا على السوابق مصطلحات كـــالصدور[7]، والبوادئ واللّواحق القبلية والإسباق واللّواحق الأولية[8]....

وأطلقوا على المقحمات تسمية الدواخل[9]، والأحشاء[10]، والأواسط[11]، وعلى اللّواحق الإعجاز[12]، والكسع[13]، واللّواصق البعدية[14]...إن كل مجموعة في هذه المصطلحات تدل على واحد فلا فرق بين مصطلح وآخر من حيث الاستخدام[15].

أمّا **الزوائد**[16] في الكلمات العربية فهي محدودة محصورة جمعها علماء الصرف في "سألتمونيها"[17] ويمكن تصنيف اللّواصق على الموقع الذي تشغله في الجذر إلى:

1- أشواق محمّد النّجار،"دلالة اللّواصق التصريفية في اللّغة العربية، دار دجلة،عمان، الطبعة الأولى، 2007، ص 62.

2- اللّواصق أو الزوائد: Affixes، يطلق مصطلح اللّواصق أو الزوائد في مجال الموفولوجيات على تلك المورفيمات التي تستخدم ضمن عملية الاشتقاق لتكوين كلمات مركبة انطلاقًا من مورفيم أو كلمة قاعدية بسيطة وتنقسم إلى قسمين: السوابق واللّواحق، للمزيد من بسط القول ينظر : ينظر: ماري نوال غاري بريور "المصطلحات المفاتيح في اللّسانيات"، سابق، ص 15.

3 –Préfixes.

4 –Infixes.

5 – Suffixes.

6- أشواق محمّد النّجار، "دلالة اللّواصق التصريفية في اللّغة العربية، دار دجلة، عمان، الطبعة 1، 2007، ص 67.

7- الأمير مصطفى الشهابي، "المصطلحات العلمية في اللّغة العربية في القديم والحديث، سابق، ص 14.

8- أشواق محمّد النّجار، "دلالة اللّواصق التصريفية في اللّغة العربية"، سابق، ص 67.

9- محمد رشاد الحمزاوي، "المصطلحات اللّغوية الحديثة في اللّغة العربية"، معجم عربي أعجمي وأعجمي عربي، سابق، ص 61 ود. محمود فهمي حجازي "مدخل إلى علم اللّغة"، سابق، ص 89..

10- المصطلحات العلمية في اللّغة العربية في القديم والحديث، ص 14، و"المصطلحات اللّغوية الحديثة في اللّغة العربية"، ص 52.

11- تمام حسان، "مناهج البحث في اللّغة"، سابق، ص 221.

12- المصطلحات اللّغوية الحديثة في اللّغة العربية، سابق، ص 127.

13- المصطلحات العلمية في اللّغة العربية في القديم والحديث، ص 14.

14- دلالة اللّواصق التصريفية في اللّغة العربية، ص 67.

15- يستعمل البحث مصطلحات السوابق واللّواحق والمقحمات.

16- الزوائد يقصد بها اللّواصق.

17- محمد المبارك، "فقه اللّغة وخصائص العربية، سابق، ص 74، 75، وينظر: اللّغة العربية، معناها ومبناها، سابق، ص 160 وما بعدها.

- **السوابق**[1]: Préfixes: اتفق اللّغويون على تعريفها بالعناصر (Eléments) التي تضاف إلى أول الكلمة، مثال ذلك في اللّغة العربية، لواصق المضارعة التي تدخل على أول الفعل المضارع نحو: أكْتُبُ، ونَكْتُبُ، وتَكْتُبُ، ويَكْتُبُ[2]، فالسوابق هي التي تسبق الاسم أو الفعل أو الحرف مثل (ال) التعريف، وحروف الزيادة في الأفعال (سألتمونيها) وغيرها، ويدخل ضمن السوابق ما يعرف بالموجه : وتمثله حروف مثل (السين) أو (ســـــوف) أو (قد)، فهذه الحروف دورها تكييف معنى الزمن للفعل فـــ (قد) مثلا تدل على تمام الحدث أو وقوعه بالفعل عندما يكون الفعل ماضيا وتدل على التشكك أو الإمكان مع الأفعال المضارعة، أما (سوف) فهي تدل على معنى المستقبل وكذلك (السين) ، فهذه الحروف تعتبر موجهة[3].

-المقحمات: Infixes: هي العناصر التي تتوسط الجذر[4]، مثال ذلك في اللّغة العربية: (التضعيف[5]).

أما **اللواحق**: Suffixes: فهي العناصر التي تضاف إلى نهاية الجذور، لتغيير وظائفها أو معانيها الأصلية، مثال ذلك في اللّغة العربية لاحقة الألف والنون (ا ن) والواو والنون، للدلالة على التنبيه وجمع مذكر السالم[6]، فاللواحق هي التي تلحق آخر الاسم أو الفعل أو الحرف كالضمائر المتصلة، وعلامات الجمع، ومـــن أمثلة ذلك (رجلان)، حيث يعتبر(رجل) وَ(ان) مورفيمان في علم اللغة التركيبي الحديث، أو وحدتان ذواتا معنى، تحمل إحداهما المعنى الأساسي للكلمة، وتحمل الثانية معلومات عن العدد، وعلى كل حال فالتفرقة بين اللفظين ربما تتم عن طريق تسمية الأولى باسم المورفيم الحر: أي الذي يمكـــــن أن يستعمل بمفرده والثاني باسم المورفيم المتصل : أي الذي لا يستعمل منفردا، وإنما متصلا بمورفيم آخر[7].

1- السوابق في العربية هي عنصر من عناصر الإشتقاق الصغير والنحت المشتق حسب ابن فارس في مقاييسه، وتظهر في الصيغ المزيدة على الثلاثي، وتضيف إليها معاني متعددة منها التعدية، والدخول في ...، والطلب، واستحالة الشيء، وتعتبر السابقة تحيلا سطحيا يطرأ على البنية العميقة التي كثيرا ما تكون ثلاثية في العربية، وينشأ منها توليدات دلالية هي على قدر وجوه مصدر السابقة وصيغها.للمزيد من بسط القول ينظر : محمد رشاد الحمزاوي، المعجمية،(مقربة نظرية ومطبقة، مصطلحاتها ومفاهيمها)، مركز النشر الجامعي، تونس، د.ط، 2004، ص 235.

2- دلالة اللّواصق التصريفية في اللّغة العربية، سابق، ص 68 والمصطلحات اللّغوية الحديثة في اللّغة العربية، ص 78.

3- عبد القادر الفاسي الفهري، البناء المتوازي نظرية في بناء الكلمة وبناء الجملة، مكتبة بستان المعرفة للطباعة والنشر والتوزيع، 1990، ص52.

4- ينظر: المرجع نفسه، ص 49 و 68.

5 -Réduplication.

6- ينظر: دلالة اللّواصق التصريفية في اللّغة العربية، ص 49 و 68 والمصطلحات اللّغوية الحديثة في اللّغة العربية، ص 127.

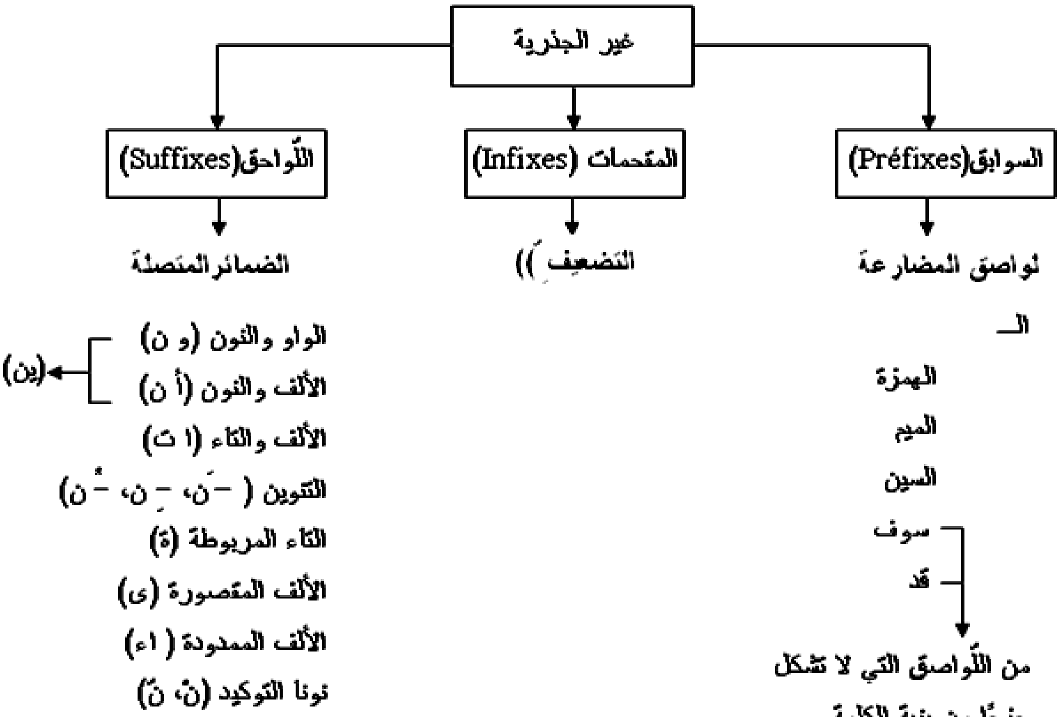
7- عبد القادر الفاسي الفهري، البناء المتوازي نظرية في بناء الكلمة وبناء الجملة، مكتبة بستان المعرفة للطباعة والنشر والتوزيع، 1990، ص53.

ومهما يكن من أمر، فإن هذه العناصر الثلاثية عناصر تجريدية، فالزيادة عند العلايلي هي توسيع في الجذور العربية[1]، فهي تساعد على خلق كلمات أو أصول موجودة بالفعل[2] فالزّيادة تتوقف عند السداسي فقط، لأن الزيادة هنا بلغت ضعف الأصل[3].

ويقول بلومفليد Bloomfield: (إن استخدام السوابق واللّواحق يقرر لنا تحديد بداية، أو نهاية، أو جذر كل كلمة)[4].

– **اللّواصق** (Affixes) مصطلح أعمّ من السوابق، والمقحمات واللّواحق، لأنّها تضم العناصر الثلاثة، فهي عبارة عن وحدات صرفية تضاف إلى الجذر[5]، ومع ذلك يجوز استخدام اللّواصق بدلاً من السوابق والمقحمات، واللّواحق، لأنّ هذه العناصر أجزاء من اللّواصق[6].

ويمكن تصنيف أشهر اللّواصق التصريفية وتفريعاتها في العربية على النّحو الآتي[7]:



<u>الشكل رقم 14</u>: أهم اللّواصق التصريفية وتفريعاتها في العربية.

1– أسعد أحمد علي، "تهذيب المقدمة اللّغوية للعلايلي"، سابق، ص 202.

2– دلالة اللّواصق التصريفية في اللّغة العربية، ص 68.

3– سيدي محمد غيثري، التّركيب الفعلي العربي، دراسة لسانية، حاسوبية، سابق، ص 140.

4– دلالة اللّواصق التصريفية في اللّغة العربية، ص 68.

5– وجيه حمد عبد الرحمن، منهجية وضع المصطلحات الجديدة في الميزان، دورية اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 24، الرباط، المغرب، 1985/1984، ص 58،57.

6– المرجع نفسه، الصفحة نفسها.

7– المرجع نفسه، ص 70.

رابعا- الإطار العام للمعالجة الآلية للصرف العربي:

تتبوأ المعالجة الآلية للصرف العربي مكانة متميزة في ميادين التعامل مع اللّغة العربية على وقف منظور معلوماتي، وتعدّ مدخلاً أساسًا، وقاسما مشتركًا لإدارة معظم نظمها الآلية التي تعنى بإنشاء نظم المعلومات والمعارف.

4-1- مفاهيم أساسية:

يقصد بالمعالجة الآلية للصرف العربي اعتماد نظم حوسبة متقدمة تستند إلى خوارزميات برمجية تستثمر المنطق لصرفي العربي في معالجة المفردة العربية عن طريق استخلاص العناصر الأولية لبنية الكلمة ومباشرة تحديد سماتها الصرفية[1]، والصرف نحوية، والصرف دلالية، القابلة للاستنباط في هذه البنية[2].

وقد أفرزت جهود العلماء والباحثين في هذا المضمار[3] ظهور مجموعة من النماذج المقترحة لمعالجة الصرف آليًا بيد أن الدراسة المتأنية لهذه النماذج تظهر بوضوح وجود

1- نبيل علي، اللغة العربية والحاسوب، سابق ،ص308.

2- حسن مظفر الرزّو "سبل تطوير محلل الصرف الآلي المستخدم في حوسبة الموسوعات العربية" مقال منشور، موقع الألوكة، دراسات، سبل تطوير محلل الصرف الآلي. موقع الويب:Alukah.net.http://www .

3- بعض الدراسات السابقة ذات الصلة في مجال المعالجة الآلية للصرف العربي:

* وأبتدئ بالرسائل العلمية وأخص بالذكر أطروحة دكتوراه بالجامعة الأردنية 2005م بعنوان "النظام الصرفي للعربية في ضوء اللسانيات الحاسوبية – مثل من جمع التكسير" للباحثة هدى سالم عبد الله آل طه، فقد تناولت الدراسة توصيف النظام الصرفي للعربية في ضوء اللسانيات الحاسوبية متخذة من جمع التكسير ظاهرة، في محاولة تجاوز وصف القدماء إلى توصيف يؤهلها للحوسبة. وقد اتكأ وصفها على حدس ابن اللغة وقدرته على الاستبطان والاستنباط لاسيما في الجانب التحليلي من المعالجة الصرفية. لكن الطابع العام لهذه الدراسة اقتصر على الجانب النظري، تاركة مجال التطبيق العملي لها مشروعاً في المستقبل القريب. وستفيد هذه الدراسة من تلك الأطروحة وذلك للتقارب من حيث المضمون في تناولهما المستوى الصرفي.

* ومن الأطروحات أيضاً بحث عمر مهديوي: توليد الأسماء من الجذور الثلاثية المعتلة: مقاربة لسانية حاسوبية، وهو بحث لنيل دكتوراه السلك الثالث (دبلوم الدراسات العليا) في اللسانيات الحاسوبية العربية، كلية الآداب والعلوم الإنسانية، ظهر المهراز، فاس، 1999.

* ومن الدراسات التي تناولت المستوى الصرفي "خوارزميات توليد الأسماء في اللغة العربية" لعمر مهديوي، مركز البحوث العلمية والتقنية لترقية اللغة العربية بتعاون مع أكاديمية الجزائر بالجزائر في 27 – 28 يونيو 2005م. إذ عالجت هذه الدراسة اللسانية الحاسوبية مستوى هاماً من النظام الصرفي أي توليد الأسماء البسيطة من الجذور المعتلة والخوارزميات والزوائد التصريفية التي تدخل على بنية الجذر، سواء كانت سوابق أو أواسط أو لواحق. وتقارب هذه الدراسة ما ننوي القيام به من خلال توصيف المشتقات الواردة في ديوان السياب ومعالجتها آلياً. بيد أن عمل الباحث كان أقرب للمنطق الرياضي في عمليات المعالجة.

* وفيما يتعلق بالتصريف والاشتقاق لابد من ذكر ما أنجزه الأساتذة (مروان الباب، ومحمد مراياتي، ومحمد حسن الطيان، وأسامة رجب، ويحي مير علم) للنظام الاشتقاقي التصريفي للكلمات العربية الذي أصدره المعهد العالي للعلوم التطبيقية والتكنولوجيا بدمشق الإصدار الأول (1998)، والفائدة ستكون كبيرة من هذا النظام وذلك من حيث الاستفادة من الجداول والمعطيات التي يحتويها ناهيك عن منهجه وطريقته في تصريف المشتقات.

* ومما قدم في ميدان الصرف دراسة بعنوان "تعليم الصرف العربي بدعم الحاسوب" لسالم الرامي، وقدمت ضمن وقائع ندوة "التوليد والنسقية والترجمة الآلية"، التي نظمها معهد الدراسات والأبحاث للتعريب في المغرب، تناولت هذه الورقة، حقل النظم التعليمية الآلية، من خلال النطاق الثلاث التالية:

أنسجة البرامج التعليمية المدعومة بالحاسوب.

دراسة البرمجيات التعليمية المتوفرة، لتمكين المدرسين من تبين معايير انتقائها، ودرجات فعاليتها، ثم حدودها.

= تحديد البرمجية التعليمية الذكية التي نقترحها لتعليم تصريف الأفعال في اللغة العربية، وهي نظام مبني على التوليد أساساً.=

اختلافات من حيث ملا ئمتها لمتطلبات الصرف العربي، ومدى كفاءتها في أداء المهام التي تسندُ إليها وسنقوم هنا باستعراض النموذج الذي قام به الدكتور نبيل علي لمعالجة اللّغة العربية صرفيًا بصورة وجيزة كونه يشبه النموذج الذي نصبو إلى صياغته لبناء المعجم الالكتروني للمعالجة الآلية للغة العربية من حيث نقطة ارتكازه على خصائص الصرف العربي "الصيغة والميزان الصرفي".

4-2- المعالج الصرفي متعدد الأطوار:

هو نموذج التحليل بالتركيب الّذي قام بوضعه د. نبيل علي لشركة صخر العالمية[1] يمتاز بآلية عمل متقدمة من الناحيتين البرمجية البحتة والمعالجة الصرفية واللّغوية، الأمر الذي جعله يتبوأ مكانة بارزة في ميدان البحث الحاسوبي بالموسوعات، ونظرًا لخلو الساحة العربية من محللات صرفية آلية ترقى إلى مستواه الرّصين في المنطق اللّغوي والبرمجي المحكم، سأحاول التركيز على الجانب المعالجة الاشتقاقية لهذا النموذج لما له أهمية في هذه الدراسة[2].

=* ومن البحوث التي اهتمت بالحاسب والبنى الصرفية للغة العربية، ثلاث بحوث في هذا الإطار خلال الندوة الدولية الأولى عن الحاسب واللغة العربية في مدينة الملك عبد العزيز للعلوم والتقنية في الرياض 2007م وكانت على النحو الآتي:

(1) الحوسبة التوليدية للصرف العربي، لبوشعيب راغين. وقد تعرض الباحث لأبرز المحاولات التي سعت لمعالجة الصرف العربي آلياً عربياً وعالمياً من خلال توزيعها إلى شقين يستغرق أولهما النموذج التحليلي للصرف، ويستغرق الثاني النموذج الاشتقاقي الصرفي.

وخصص النقطة الثالثة لتقديم أسس مقاربة حاسوبية جديدة مبنية على توليد الصرف مستدلاً عليها من خلال تدقيق ميكانزمات التوليد، وتحديد أهم المصافي الصوتية والصرفية والخطية بشكل مبسط للتمثيل فقط. واقترح في نهاية البحث المولد الصرفي للصور المعجمية العربية، مع تطبيق أولي على تصريف الأفعال العربية، كيفما كان نوعها.

(2) خصائص حاسوبية في بناء الدلالة اللغوية، نماذج من المعجم العربي لمحمد غاليم. فقد تناول الخصائص الحاسوبية التي تتصف بها قواعد بناء الدلالة اللغوية. وذلك من خلال بنيات دلالية معجمية عربية.

(3) المعجم العربي في ضوء اللسانيات الحاسوبية لعمر مهديوي، إذ تطرق إلى المعجم العربي في ضوء المنظور اللساني الحاسوبي الحديث، مع الوقوف عند معمارية المعجم الورقي التقليدي ومقوماته وثغراته. وتناول الوسائل والأدوات الحديثة التي يمكن اعتمادها في تحديث المعجم العربي بنوعيه القديم والحديث متخذاً من لسان العرب لابن منظور منطلقاً لاختبار تجربة حوسبة المعجم التقليدي.

* وهناك دراسات قدمت في هذا المجال نذكر منها على سبيل الذكر لا الحصر:

أ-التحليل الصرفي للغة العربية باستخدام الحاسوب، مأمون الحطاب وحسان عبد المنان فقد عرضنا تجربتهما في بناء محلل صرفي باستخدام الحاسوب، من خلال رد الكلمة إلى جذرها، وتبيين وزنها، وما دخلها في الزيادة.

ب-تعامل الأجهزة والمعدات مع الحرف العربي، محمد مراياتي. إذ تناول في سياق بحثه مسألة تعرف الكلام بتمثيله للحاسوب ليحوله إلى رسم مكتوب.

ج-المعالجة الآلية للكلام المنطوق: التعرف والتأليف، سالم الغزالي. وذهب في بحثه إلى اختيار الكلية التي ترمي إلى التعرف على الكلمة دون تجزئتها لتفادي المشاكل المنجزة عن التصاحب النطقي.

وهناك العديد من الرسائل العلمية في اللسانيات الحاسوبية نوقشت في الجامعات الأردنية، منها:

أ- النظام الصرفي للعربية في ضوء اللسانيات الحاسوبية " إعداد : هدى سالم آل طه، إشراف : الأستاذ الدكتور نهاد الموسى، الجامعة الأردنية، 2005.

ب- التركيب الإضافي في العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، إعداد: جنات علي أحمد، إشراف: الأستاذ الدكتور نهاد الموسى، الجامعة الأردنية، 2007.

1- سبل تطوير محلل الصرف الآلي المستخدم في حوسبة الموسوعات العربية، سابق، وينظر: اللّغة العربية والحاسوب، ص 308.

2- للمزيد من التفصيل ينظر: اللّغة العربية والحاسوب، ص 308 وما بعدها.

يتألف المحلل الصرفي الآلي متعدد الأطوار من أربع معالجات ثانوية تعمل بصورة متكاملة لتحقيق الغاية التي تهدف إليها عملية الحوسبة الصرفية للنصوص العربية، وهي:

1- المعالج الصرفي النّحوي؛ **2**- المعالج الاشتقاقي، **3**- المعالج الإعرابي، **4**- معالج التشكيل.

أ- المعالج الاشتقاقي:

يباشر المعالج الاشتقاقي عمله باستخلاص الجذور، والصيغة الصرفية من الجذع أو الجذوع، التي قام بتفكيكها المعالج الصرف النحوي، مع افتراض عدم وجود علامات التشكيل، تبدأ عملية الاستخلاص بمقارنة سلسلة حروف الجذع مع قائمة قوالب الهياكل الصرفية التي تتضمنها قاعدة بياناته، فعند تطابق سلسلة الحروف مع أحد الهياكل الصرفية، يقوم المستخلص بتسيير حروف الجذور المناظرة لحروف (الفعللة) الواردة في الهيكل المطابق[1].

بعد استخلاص الجذر يقارن بمعجم الجذور المسموح بها في اللّغة العربية للوصول إلى جذر وأحد المراد الحصول عليه. أمّا إذا لم يفلح المستخلص في الوصول إلى الجذر الصحيح "لعدم عثوره على طول جذع مكافئ في الصيغ الصرفية المتاحة"، فيعمد إلى افتراض وجود حالة أو جملة حالات من الإبدال والإعلال، ليقوم بناءًا على ذلك بتوظيف سلسلة من الافتراضات المحتملة لعكس عمليات الإبدال والإعلال، ثم يكرر بعدها الخطوات الإجرائية سالفة الذكر لحين وصوله إلى الجذر الصحيح.

ثم يقدم المستخلص الاشتقاقي مخرجاته إلى المعالج الإعرابي على هيئة سلسلة الجذور، والهياكل الصرفية (الأوزان أو الصيغ) المحتملة لجذور الكلمة.

ثم يأتي دور المعالج الإعرابي حين يعمد إلى تمييز الحالة الإعرابية بناءًا على الوسم الإعرابي الذي تتضمنه، وبعد ذلك يعمد معالج التشكيل إلى إعادة تركيب عناصر الكلمة لغرض مقارنتها بالكلمة المدخلة إلى ساحة المعالجة الآلية؛ وذلك للتأكد من التحليل الصرفي لبنيتها، ويتم ذلك عن طريق إعادة تركيب جذع الكلمة، عبر مطابقة الجذر مع جميع الصّيغ الصرفية المناظرة للهيكل الصرفي الذي حدّده المستخلص والتي يسمح المعجم اللّغوي بانطباقها عليها[2].

فعلى سبيل المثال، وفي حالة الوزن الصرفي، "مفعل" يقوم مُعالج التشكيل باسترجاع الصّيغ الصرفية المناظرة له من ملف الصّيغ الصرفية وهي: "مَفعل، مُفعل، مِفعل، منفعل..." بعد ذلك يقوم بالتأكد من صحة انطباق الصيغة الصرفية المحتملة على جذر الكلمة قيد الدراسة،

1- اللّغة العربية والحاسوب، سابق، ص 311.

2- المرجع نفسه، ص 311، 312، 313.

ليقوم بعد ذلك بمطابقة الجذر مع كل صيغة صرفية منطبقة عليه، ثم ينفذ جميع عمليات الإبدال والإعلال الواجبة تحت توجيه سلسلة القواعد الصرفية التي يتعامل معها، بعدها يدمج عناصر الكلمة الأخرى (السوابق واللّواحق) مع جذع الكلمة التي أعاد تركيبه[1] فنلاحظ أن من خلال تقنية عمل هذا النموذج أنّ الخاصية الصرفية تطرح نفسها كمدخل منطقي لعملية استخراج جذور الكلمات العربية[2].

من خلال هذا العرض الموجز لأهم نموذج من نماذج المعالجة الآلية للصرف العربي يتبين لنا كيف يمكن أن يستثمر الصرف العربي في معالجة الكلمة العربية من خلال استخلاص العناصر الأولية لبنية الكلمة، وهذا عن طريق تحليلها ثم تحديد سماتها الصرفية والنّحوية، هذا من جهة ومن جهة أخرى يمكن استثمار منظومة الصرف العربي في البحث في كيفية معالجة الصرف العربي آليًا وهذا من خلال[3]:

- الاهتمام بشقّ المعنى صرفيًا ومعجميًا.
- وضع العلاقة النّحوية الصرفية في صورة رسمية تتمشى مع النظريات اللّغوية الحديثة.
- التعمق في الإحصاء الصرفي باستغلال الإمكانيات التي تُتيحها المعالجة الصرفية الآلية (كالمعالج الصرفي المتعدد الأطوار).
- دفع الجهود في مجال ميكنة المعجم العربي، وتفريعها لتشمل الجوانب المعجمية المختلفة.
- استغلال الحاسوب لعلاج المشكلة المتفاقمة لنقص المصطلحات.
- وضع الأساس النظري للكلمات المركبة العربية لتوسيع نطاق تكوين الكلمات.

3-4- حوسبة المعاجم:

ننهي الحديث عن تطبيقات المعالجة الآلية للصرف العربي بتأكيد أهمية هذه المعالجة الآلية في حوسبة المعاجم العربية، وتوسيع نطاق مصطلحاتها، وتعزيز معطياتها اللّغوية: الفونولوجية، والصرفية، والنّحوية والدلالية.

إنّ أهمية الإمكانات الهائلة للحاسوب في الحقل المعجمي والمصطلحي وما يجري حاليًا من جهود لحوسبة المعاجم يعد دليلاً وللاتجاهات المعجمية الحديثة.

1- المرجع السابق، ص 301.

2- المرجع نفسه، ص 323.

3- المرجع نفسه، الموضع نفسه.

إنّ أهمية الحاسوب للمعجم العربي تفوق بكثير أهميته بالنسبة للمعجم الأجنبي (الانجليزي مثلاً)، ويرى نبيل علي أن: « أساس هذا الرأي هو الطبيعة الخاصة لهيكل المعجم العربي المبني على أساس الجذور»[1]، ولهذا خصصنا الفصل الثاني من هذا البحث للحديث عن الجوانب المختلفة لهندسة المعجم الالكتروني، والدور الذي يلعبه الصرف العربي في هذا الجانب من هذه الدراسة .

4-4- أسس مقترحة لمعالجة الصرف العربي آليا:

تعرضنا في الفقرات السابقة لهذا المبحث إلى تعريف الصرف وخصائصه وظواهره المميزة وانتهينا على التو من طرح بعض المصطلحات التي كان لابد للتعرض لها لما لها من أهمية كبرى في مشروعنا المراد تحقيقه وحان الوقت لاقتراح الأسس التي نرى ضرورة إقامة هذه المعالجة عليها، وهي الأسس التي لا بد وأن تستغل خصائص الصرف العربي، وتتعامل مع دخائله، وتتصدى لمشاكله وتتجاوز مع أوجه قصوره؛ وأريد التأكيد مرةً أخرى إلى أهمية الصرف العربي، فهو يمثل مجالاً نموذجيًا لتزاوج الحاسوب واللّغة، ومرجع ذلك هو نمطية الاشتقاق، واطراد التصريف، وانتظام قواعد الإبدال والإعلال، واتساق بنية الكلمة.

وتتمثل أهم أسس معالجة الصرف العربي آليًا في[2]:

– كمبدأ عام في تصميم النظم الآلية، يجب أن يكون نظام الصرف الآلي "تجزيئيًا"، أي مكونًا من عدة آليات متخصصة يربط بها علاقات ترابط واضحة ومحدودة ويجب أن تؤسس تجزئة النظام الآلي على أساس لغوي، لا فنّي، ونقصد بذلك تقسيم النظام إلى عناصر تعكس الوظائف الأساسية للمنظومة الصرفية، لا الخطوات الإجرائية للبرنامج.

1– المرجع السابق، ص 330.

2– اللّغة العربية والحاسوب، سابق ، ص 300.

– يجب أن تتعامل المعالجة الآلية للصرف العربي مع ثنائية الصيغة الصرفية والميزان الصرفي (البنية العميقة والبنية السطحية)[1].

– يجب أن تكون المعالجة الآلية للصرف العربي بشكل عام أي غير مقيدة بنظام معيّن ويجب أن تكون شاملة لجميع الجذور المعجمية الممكنة والصيّغ الصرفية المسموح بها.

– فصل القواعد الصرفية عن البرنامج الآلي، وذلك لإمكان تعديل هذه القواعد بأقل كمية من التغيرات في برنامج المعالجة، وربّما يمكن ذلك في مرحلة متقدمة في جعل المعالجة الصرفية لها القدرة على التعامل مع عدّة لغات بمجرد تغيير الشق الخاص بقواعد الصرف خاصة وأن هذا الأمر يبدو أكثر واقعية بالنسبة للصرف العربي.

– مراعاة ثنائية التحليل والتوليد، بأن تكون المعالجة الآلية قادرة على العمل في الاتجاهين ويقصد بذلك إمكانية تحليل الكلمات القائمة إلى عناصرها الأولية، أو تركيب هذه العناصر في هيئة كلمات نهائية، وفي هذا المقام، علينا أن نشير إلى أن قواعد الصرف العربي في كثير من جوانبها تنحاز نحو التوليد، وعلى مصمم المعالج الصرفي الآلي أن يسعى لإحداث التوازن الوظيفي، وذلك بجعل قواعده وبرامجه[2] أقل حساسية لغرض المعالجة واتجاهها.

– ضرورة توفر عنصري الكفاءة والسرعة حيث سيتم دمج المعالج الصرفي في كثير من النظم اللّغوية العامة، وهو ما يفرض قيودًا قاسية على مطالب هذا المعالج من العتاد والبرمجيات حتى لا يستهلك جزءًا كبيرًا من طاقة الحاسوب وموارده التي يجب أن تتوزع على عدد أكبر من النّظم اللّغوية الفرعية التي يحتويها النظام الآلي الأشمل[3].

1– عرّف تشومسكي البنية العميقة بأنها الأساس الذهني لمعنى معين يوجد في ذهن المتكلم ويرتبط بتركيب جملي أصولي، ويكون هذا التركيب رفداً وتجسيداً للمعنى، وهي النواة التي لا بد منها لفهم الجملة ولتحديد معناها الدلالي، ومن مثال ذلك: (يشرح المحاضر الدرس بقلم يكتب به على السبورة)، فهذه الجملة تتكون في الأصل من ثلاث جمل أصولية تمثل كل واحدة منها معنى عقلياً في ذهن المتكلم، وهي: (يشرح المحاضر الدرس، يكتب المحاضر بالقلم، يكتب المحاضر على السبورة)، ويربط بين الجمل الثلاث محور رئيس أو علاقة بين عناصر رئيسة (المحاضر، الدرس، السبورة، القلم). وهذه الكلمات تمثل البنية العميقة التي يعبّر عنها عَبرَ بناء جملي تحويلي يعبّر عن العلاقة بين الكلمات السابقة كالآتي: (يشرح المحاضر الدرس بقلم يكتب به على السبورة)، فقد يقدم جزء من الجمل النواة على الآخر، مثلاً: (يكتب المدرس بالقلم على السبورة وهو يشرح الدرس). وقد يقدم الجزء الثالث على الثاني أو على الأول ...إلخ، حيث إن هذا كله من تقديم وتأخير لا يؤثر في المعنى الذي في ذهن المتكلم أو في الكشف عنه، فالبنية السطحية هي الكلام المنطوق أو الألفاظ التي ارتبطت ارتباطاً نحوياً وفق قواعد اللغة، وتعبر عن معنى ذهني مجرد لكلمات محسوسة منطوقة. انظر البنية العميقة والبنية السطحية في: Chomsky, Noam.1965. Aspects of the theory of Syntax. P.18.؛ والراجحي، عبده، النحو العربي والدرس الحديث، ص140؛ وعمايرة، خليل، في نحو اللغة وتراكيبها، ص65؛ والسيد، صبري إبراهيم، تشومسكي: فكره اللغوي وآراء النقاد فيه، دار المعرفة الجامعية، الإسكندرية، 1989م، ص66؛ وحسام الدين، كريم زكي، أصول تراثية في علم اللغة، ص66؛ وخرما، نايف، أضواء على الدراسات اللغوية المعاصرة، سلسلة عالم المعرفة، الكويت، 1978م، ص308، وغيرها من المراجع الحديثة التي تناولت نظرية تشومسكي.

2– اللّغة العربية والحاسوب، سابق ، ص 300.

3– المرجع نفسه، الصفحة نفسها.

– ضرورة تعريض المعالج الصرفي لاختبارات قاسية للتحقق من سلامة الأسس اللّغوية المقام عليها، والتأكد من اكتمال تغطيته لجميع الحالات الصرفية والمعجمية للكلمة العربية[1].

– الاهتمام بشق المعنى، والذي يمثل الغاية القصوى للتنظير اللّغوي للصرف وبالتالي معالجته آليًا[2].

– الالتزام بما خلص إليه البحث الصرفي الحديث من حيث اعتبار الكلمة دون غيرها هي أساس تكوين الكلمات، واستخدام الأساليب المنهجية الحديثة في بناء المعاجم الآلية وصياغة القواعد الصرفية وتبويبها[3].

وإدراكاً منا لأهمية منظومة الصرف في اللغة العربية من خلال بناء قاعدة بيانات صرفية، يمكن توظيفها في بناء برنامج حاسوبي بالإعتماد على التقنيات اللغوية للمعالجة الآلية للغة العربية، وزيادةً في الإيضاح نستعرض في نهاية هذا المبحث مقتطفات من المعجم الإلكتروني وتبيان أهمية المداخل الصرفية للغة العربية في البرنامج، فمعظم هذه البرامج المقترحة تُبنى على فكرة استخدام نظام قواعد البيانات المعجمية والتي تتكون من مجموعة من المعاجم الصغيرة للجذور في اللغة العربية ومعجم للأوزان الصرفية، مع مجموعة من القواعد الصرفية اللازمة للتحليل والتركيب . ويبقى التفاوت فيما بين هذه الحلول في كيفية تنفيذ قواعد البيانات والقواعد الصرفية وعن مدى التزامها بأصول لغوية منضبطة، وكيفية استغلال واستثمار البيانات الصرفية في اللغة العربية لبناء المعاجم الآلية للغة العربية .

تم وضع هذا المعجم ببرنامج توليد قاعدة بيانات الصرف العربي المرفقة صورة من واجهته[4].

1– المرجع السابق، ص 301.

2– اللّغة العربية والحاسوب، سابق، ص: 301.

3– المرجع نفسه، الصفحة نفسها.

4– واجهة تخاطبية Interface صممت لتوليد أسماء الذات آليا، وهي إحدى الواجهات التخاطبية التي صممت لتوليد مكونات المعجم الآلي للغة العربية، صممت بلغة بيسك(basic). تنقسم إلى مجموعتين من النوافذ: الأولى تتضمن جميع الجذور العربية وعددها يفوق 12،000 جذر (ثلاثي ورباعي)، والثنية تحتوي على مجموع الأوزان الخاصة باسم الذات في اللغة العربية، يفوق عددها المائتي وزن، أما بقية الأحياز فهي عبارة عن خصائص الاسم المعرب في اللغة العربية: التعريف والتنكير، والتأنيث والتذكير، والتصغير ...إلخ.

الشاشة رقم 01: المعجم الإلكتروني بقاعدة بيانات صرفية

وبقليل من التأمل لشاشة عرض البرنامج يدرك الباحث مدى أهمية المعالجة الآلية للصرف العربي حينما يتم توظيفها واستثمارها وفق أسس لغوية دقيقة ومحكمة في توليد برامج حاسوبية هدفها خدمة الأنظمة اللغوية عامة، وتطوير وإنماء المحللات الصرفية والمعاجم الإلكترونية للغة العربية على الخصوص.

خامسا: تطبيقات المعالجة الآلية للغة العربية:

هناك مجموعة من الأبحاث والدراسات وأوراق العمل قد تناولت موضوع المعالجة الآلية للغة العربية، وكانت في مجملها تتناول جزئية من جزئيات هذا الموضوع، يناقش هذا المبحث تطبيقات المعالجة الآلية للغة العربية في الحاسوب، وهي التحليل الصرفي، والتحليل النحوي، والتحليل الدلالي، والتحليل المعجمي، فالدراسة تناقش أهمية هذه التطبيقات في إمكانية إقامة قاعدة بيانية لبناء برامج الكترونية تعليمية كبناء قاعدة بيانات صرفية، نحوية ودلالية معجمية لبناء معجم الكتروني، إن تصميم نظام آلي بهذا الحجم هو أحد مشارف البحث الهامة في تطويع اللّغة العربية لمطالب النظم الآلية باللّغات الطبيعية، فتتميز هذه النماذج أو النظم الآلية لمعالجة اللّغة بقدرات هائلة لإزالة أي مشكلة تنجم عن معالجة اللّغة العربية آليًا– فمثلاً كاللّبس النّاجم عن التّشكيل أو نقصانه– على مستوى معالجة قاعدة البيانات النحوية[1].

وتمثل هذه التطبيقات أحد المقومات الرئيسية لبناء أي نظام حاسوبي التي تتعاظم أهميتها مع التوسع في استخدام أساليب الذكاء الاصطناعي ونظم قواعد المعارف.

إنّ الهدف من هذه التطبيقات هو إكساب البرنامج الآلي القدرة على شرح معاني مفرداته وتحليل العلاقات التي تربط بين مفردات المعجم، كالعلاقة بين جذور الكلمات والصّيغ الصرفية (أو قواعد تكوين الكلمات) أو العلاقات المتعلقة بمعاني الألفاظ كعلاقات الترادف والاشتراك اللّفظي والتضاد....وتكمن أهمية هذه البرامج الآلية لبحوث الدلالة المعجمية ونظم الفهم الآلي واسترجاع المعلومة عن طريق استخراج الكلمات باستخدام قاعدة بيانات صرفية، إذ أنّها تسعى في الأخير إلى إبراز المفاهيم الأساسية لمعالجة اللّغة العربية آليًا والأبعاد المحورية لمدلولات كل برنامج آلي ومجالاته التطبيقية.

فهنا تكمن العلاقة بين هذه الأنظمة الآلية سواء كانت صرفية أو كونية أو دلالية أو معجمية هذه العلاقة التي لا ينفع معها أسلوب التحليل الهرمي متعدد المستويات في إنشاء قاعدة البيانات مما يوحي باقتراح أسلوب شبكي، تتصل فيه معظم مستويات هذه الأنظمة فيما بينها لتتضافر بعضها مع بعض لفض أي لبس في المعالجة الآلية لأي مستوى كان كليًا أو جزئيًا. وهكذا نخلص إلى نتيجة منطقية تفسر شدّة تماسك منظومة اللّغة العربية في جميع مستوياتها اللّغوية ممّا يسهّل معالجتها آليًا.

1– سلوى حماده، المعالجة الآلية للغة العربية، "المشاكل والحلول"، دار غريب للطباعة والنشر والتوزيع، القاهرة،2009، ص 150 وما بعدها.

وعلى ضوء ما سبق فإن هذه الأنظمة تمثل همزة وصل فيما بينها من خلال المستويات التحليلية لكل نظام على أساس مبدأ المعالجة الآلية للغة العربية، فأهمية هذه التطبيقات تؤدي بنا إلى طرح التساؤل التالي: هل يمكن استثمار هذه العلاقة بين الأنظمة لبناء معجم الكتروني للغة العربية يعالج جميع المستويات التحليلية لمنظومة اللّغة العربية؟

5-1- المحلل الصرفي[1]: Analyse morphologique:

5-1-1- المقصود بالمحلل الصرفي الآلي:

هو عبارة عن نظام آلي يقوم باستخلاص العناصر الأولية لبنية الكلمة وتحديد سماتها الصرفية و"الصرف- النحوية" والتي يمكن استنباطها من بنية الكلمة[2] فهو يهدف إلى ربط الكلمة بعناصرها الأولية التي تدخل في تكوينها، وربطها أيضًا بالقيم النحوية والدلالية المتعلقة بها[3]، يقوم هذا المحلل باستخراج جذع الكلمة المدخلة بعد حذف السّوابق واللّواحق، مثال: كلمة (الوالِدين) بعد حذف السّابقة واللّاحقة تصبح (والد) وهو جذع الكلمة[4].

تقوم تقنية المحلل الصرفي على شقين أساسيين هما التحليل والتركيب، ففي التحليل يقوم المحلل بتحليل الكلمة وإعطاء البيانات الصرفية الكاملة للكلمة هذا بالإضافة لوضع السمات الصرفية الخاصة بكل كلمة على حدّة فهو أشبه بمعجم صرفي متكامل للغة العربية، ثم يأتي دور المركب ويقوم بدور عكسي، حيث يعيد توليد الكلمة من بياناتها الصرفية من سوابق ولواحق والجذر والوزن الصرفي[5]؛ فالمحلل الآلي عبارة برنامج تحليل صرفي للكلمات، فالأمر يتعلق بمشكلة واضحة المعالم لتنفيذ برامج فعّالة[6].

1- لقد صممت شركة صخر المحلل الصرفي بطريقة تقوم على إعادة توليد الكلمة من جديد من عناصرها الأولى المحللة منها أو بعبارة أخرى، يمكننا اعتبار التحليل صحيحًا إذا كان المعالج الصرفي قادرًا على إعادة توليد التحليل الصرفي من جديد ومقارنة الناتج بالتحليل الأصلي للكلمة.

2- اللغة العربية والحاسوب، ص 306.

3- يحي هلال، " التحليل الصرفي للعربية" ندوة استخدام اللّغة العربية في الحاسب الآلي، معهد الكويت للأبحاث العلمية، دار الرّازي، د. ط، 1985، ص 265.

4- مجدي صوالحه و إيرك أتول، توظيف قواعد النحو والصرف في بناء محلل صرفي للغة العربية، بحث مقدم في جامعة ليدز – المملكة المتحدة، ص14.

5- تقنيات التحليل الصرفي، موقع شركة صخر:

موقع الويب:
.Comhttp://www.skhar

6- ينظر:

Jaques PITRAT, « textes, ordinateurs et compréhension », Eyrolles Paris, 1985, P: 77.

للمزيد من التفصيل ينظر: Mohamed EL AMINE ABDERRAHIM « Reconnaissance des unités linguistiques signifiantes « thèse pour obtenir le titre de doctorat spécialité informatique. Université Tlemcen – 2007- 2008. PP: 06, 07.

يقوم المحلل الصرفي الآلي بعدّة وظائف آليًا منها[1]:

– تحليل الكلمة إلى جذرها، سواء أكان ثلاثيًا أو رباعيًا أم خماسيًا، مع تبيين الأصل الذي عليه الكلمة وتبيين الزيادة التي طرأت على الكلمة.

– تبيين ما زيدَ على الكلمة خارجًا عنها، وهو ما يعرف بالسوابق "**Préfixes**" واللّواحق "**Suffixes**".

– تبيين الوزن الّذي صيغت عليه الكلمة.

– إرجاع الحروف المنقلبة إلى أصولها.

– صياغة أمثلة على وزن معين من الأصل بهدف تعليمي أو دلالي.

فالهدف الذي يقوم به المحلل هو البحث فيما إذا كانت الكلمة جزء من الكلمة المعالجة أم لا، ويكون تحليل الكلمة إلى مورفيمات[2] دون الأخذ بالاعتبار العلاقات النحوية فيما بينها[3]. فبرنامج المحلل العرضي يعدّ أساس البرامج الآلية وخاصة البرامج التي تُعنى بالبحث والفهرسة إذ يمكن اعتباره تطبيقًا مباشرًا لها.

***/ مثال تطبيقي:**

تأتي السوابق واللّواحق متصلة بالكلمة على اختلاف نوعها، فيسبق الاسم أداة التعريف (الـــ) ومجموعة حروف الجرّ، وأدوات الاستفهام الفردية؛ كقولنا:"أفبالحق تمترون" فكلمة "أفبالحق" تتكون من كلمة "حق" وأربعة سوابق وهي: الهمزة الاستفهامية، والفاء، وحرف الجر "بـــ"، وأداة التعريف "الـــ"، ويمكن أن يتصل بالاسم ثلاثة أنواع من اللّواحق، مثل: "ياء النسب"، و"ياء الجمع"، أو "واو الجماعة" وضمير الملكية المتصلة (مثال: مادياتهم) ويمكن توضيح ذلك من خلال الجدول الآتي الّذي يبين "البنية الصرف– نحوية" للكلمة العربية[4].

1– مأمون حطاب، "التحليل الصرفي للغة العربية باستخدام الحاسوب" الموسم الثقافي الرابع عشر لمجمع اللّغة العربية الأردني، من منشورات المجمع، مجمع اللّغة العربية، الأردن، 4 ماي، 22 جوان، 1996، ص 55.

ينظر:

2 –Attia Mohammed « Developing Robust Arabic Morphological transducer using Finite state technology in 8 th annual cluk Research colloquim.

3 – Bessou Sadik « Analyse de données Textuelles pour la classification Automatique pour les techniques de Text Mining application à la Langue Arabe » Mémoire présenté ä la faculté des sciences de l'ingénieur, Département d'informatique pour l'obtention du diplôme de Magister, option ; Sciences et technologie de l'information et de la Communication, Université, FERHAT ABBAS- SETIF (UFAS) ALGERIE. 2007. P. 64.

4– ينظر: اللّغة العربية والحاسوب، سابق، ص 266.

س: سابق	مجموعة السوابق **Préfixes**			أ- حروف "أنيتُ" للفعل ب- التعريف للاسم	جذع الكلمة	مجموعة اللّواحق Suffixes		
ل: لاحق	س1	س2	س3		جذر- صيغة صرفية	ل1	ل2	ل3
فستستفتونهم **(فعل)**		ف حرف	س حرف	ت حرف مضارعة	الجذر: "فتو"- الصيغة: "يستفعِلُ"	و ن	هم	ضمير منصوب
أفباستعماريّيهم **(اسم)**	أ استفهام	ف حرف	ب حرف		الجذر: عمر- الصيغة: "استفعال"	يّ ياء النسب	ي للجمع المذكر	هم ضمير الملكية

<u>الجدول رقم 09</u> : البنية "الصرف- نحوية" للكلمة العربية.

ويمكن أن يسبق الفعل:"الهمزة" و"الفاء" و"السّين"، كقولنا: "أفستمطِرُ"، وقد يلحق الفعل:"الواو والنون" و"هم"[1]. ويبين الجدول (10) مجموعة السوابق واللّواحق التي يمكن أن تتصل بالاسم، كما يبيّن الجدول (11) مجموعة السوابق واللّواحق التي يمكن أن تتصل بالفعل.

مجموعة السوابق **Préfixes**				جذر الكلمة **Racine**	مجموعة اللّواحق **Suffixes**		
سابق 1	**سابق 2**	**سابق 3**	**سابق 4**		**لاحق 1**	**لاحق 2**	**لاحق 3**
أ الاستفهام	ف و	ب، لـــ كـــ	الـــ		ـــةً ـــت	ـــان، و ن ـــيْن، ـــين ـــات	ـــه، ـــها، ـــهن ـــهم، ـــكم، ـــكَ ـــكِ، ـــهما، كما. ـــكن- نا، ي.

<u>الجدول رقم 10</u>: مجموعة السّوابق واللّواحق التي يمكن أن تتصل بالاسم.

1- المرجع السابق، ص 266.

مجموعة السوابق Préfixes				جذر الكلمة Racine	مجموعة اللّواحق Suffixes		
سابق 1	سابق 2	سابق 3	سابق 4		لاحق 1	لاحق 2	لاحق 3
أ الاستفهام	ف و	س	**أ، ن، ي، ت**		ا، وا ـتِ، ـتُ ـتَ، نَا ـتم، ـتنَّ	ـان، ون ـن، ـين	ـه، ـها، ـهن، ـهم ـهم كم، كن، ك، كِ، ني

<u>الجدول رقم 11</u>: مجموعة السّوابق واللّواحق التي يمكن أن تتصل بالفعل.

وعندما يقوم المحلل الصرفي بتجريد الكلمة من سوابقها ولواحقها وتحديد نوع هذه السوابق واللّواحق يربطها بالمعجم لمعرفة ماهيتها وصفاتها.

ومثال ذلك تحليل كلمة: "**وبإيصاله**"[1].

السوابق: Préfixes: "و" حرف عطف، "ب" حرف جر

جذع الكلمة: Racine: "إيصال".

الرتبة النحوية: Niveau de grammaire: مصدر ثلاثي مطرد.

قسم الكلم: اسم جامد، محسوس، قابل للعد، أو اسم مجرد غير قابل للعد.

جذر الكلمة: Radical: وصل.

الصيغة الصرفية: L'aspect: إفعال.

الميزان الصرفي: Le schème: إيعال.

الحالة التصريفية: مفرد، مذكر، مجرور.

العلامة الإعرابية: الكسرة الظاهرة.

اللّواحق: Suffixes: ضمير الملكية للمفرد الغائب (هـ).

1- اللّغة العربية والحاسوب، ص 306.

1- عمليات التعديل الفونولوجي:

"إبدال" "الفاء" الجذر (الواو "وصل") و"ياءً" ومماثلة حركة الضمير المتصل (هاء الغائب) مع علامة الإعراب التي تسبقه الكسرة.

تعد عملية التحليل (أو التعرف أو التمييز) الصرفي أكثر صعوبة من عمليات التوليد وذلك لطبيعتها العكسية لردّ الفرع إلى أصله أو استرداد البنية العميقة من البنية السطحية[1].

يمكن تقسيم المحلل الصرفي إلى مراحل متتالية هي[2]:

- تحضير النّص المدخل المراد تحليله: يتم في هذه المرحلة تقطيع النّص إلى كلمات بالاستعانة بعلامات الوقف والحروف البيضاء[3].

- معالجة كل كلمة على حدّة ليتم التعرف على صفاتها وهذا بالاستعانة بمجموعة من البيانات يوفرها المحلل الصرفي على شكل قواعد للبيانات مثلاً، قواعد بيانات الأدوات العاملة[4]، قواعد بيانات الأسماء والأفعال.

أنظر الشكل الموالي[5]:

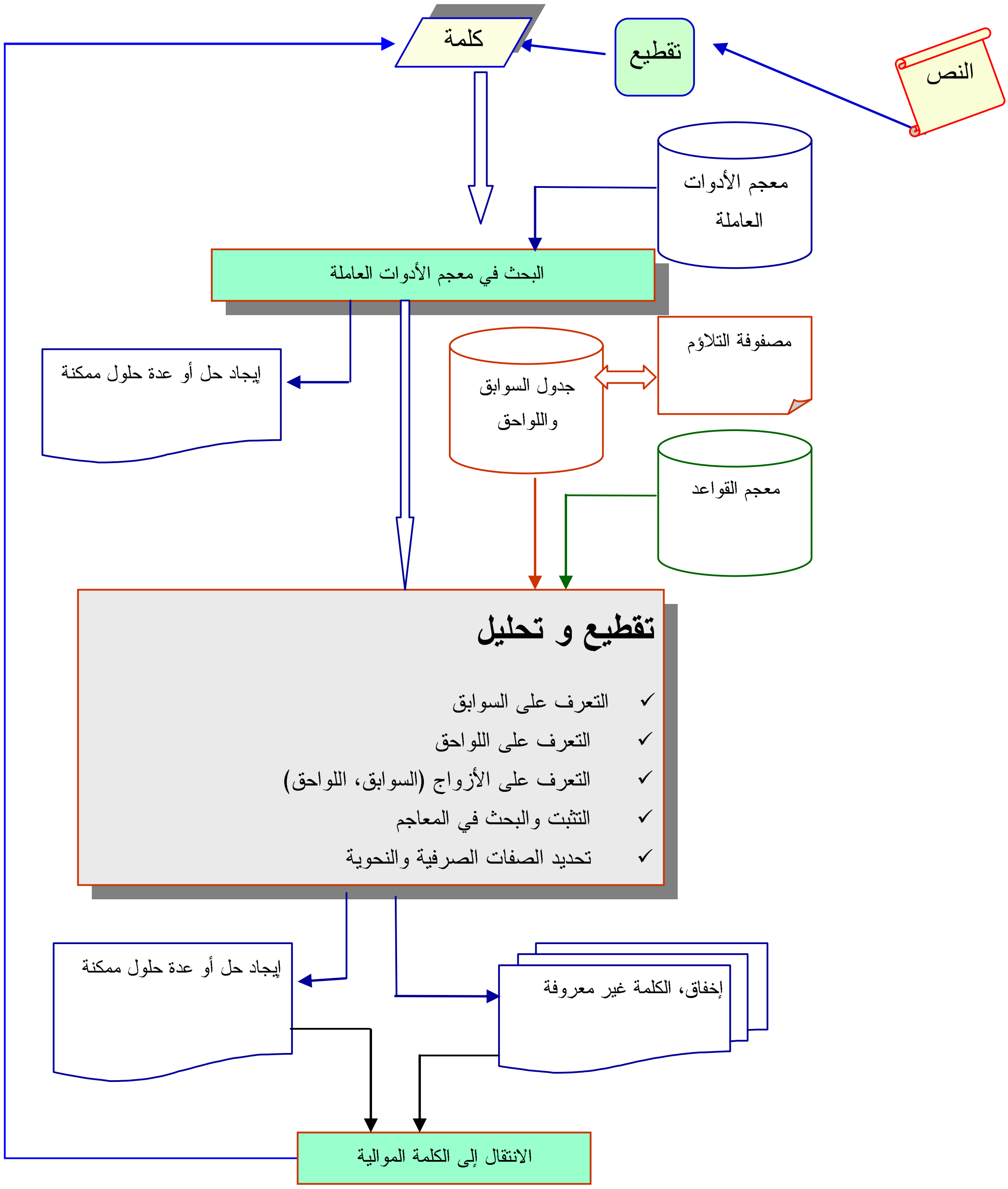
1- المرجع السابق، ص 306.

2 -ABDERRAHIM MOHAMED EL AMINE, Un analyseur morphologique pour l'arabe voyellé ou non: 1ère conférence Internationale « Système d'information et Intelligence Economique » SIIE' 2008 Hammamet-Tunisie ; 14- 16 Février 2008- Proceedings tome II The édition ; PP: 324- 339.

3 -caractère blanc.

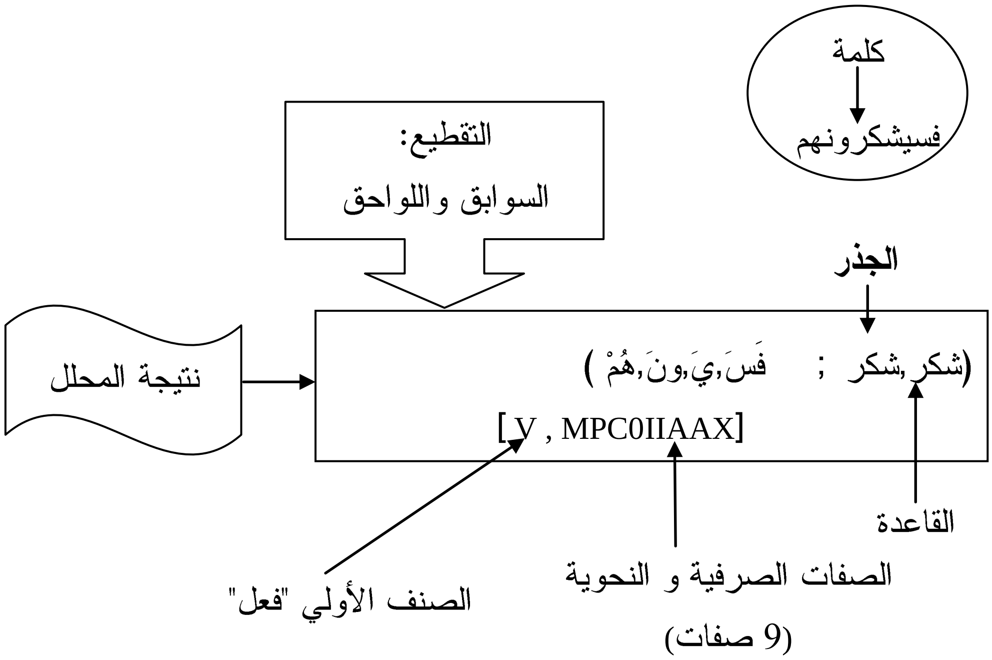
4 -Les mots outilles.

5 - Mohamed EL AMINE ABDERRAHIM « Reconnaissance des unités linguistiques signifiantes « thèse pour obtenir le titre de doctorat spécialité informatique. Université Tlemcen – 2007- 2008. PP: 117.



الشكل رقم 15: بنية المحلل الصرفي

يمثل الشكل رقم 16 الموالي[1] مثال لنتيجة التحليل الصرفي- لكلمة: فسيشكرونهم.



الشكل رقم 16: مثال لنتيجة التحليل الصرفي لكلمة فسيشكرونهم.

5-1-2- فوائد المحلل الصرفي وأهميّته:

أ- **في خدمة المحلل النّحوي:** تعدّ مستخرجات المحلل الصرفي مدخلات للمحلل النّحوي، فعلى سبيل المثال السوابق المتصلة ببداية الكلمة تبين للمحلل النّحوي حروف العطف، والجر، والاستفهام، وأمّا اللّواحق تبين نوعها: (اسم، فعل، حرف،...) وهذه أهم الأمور التي على أساسها يعطي المحلل النّحوي الكلمات مواقعها الإعرابية في الجملة.

ب- **خدمة المعاجم وكتب الصرف والنّحو**[2]: يمكّننُا المحلّل الصرفي من فهرسة الألفاظ وموازينها في النّصوص الكبيرة، فيزيد ذلك من غنى وقوّة المعاجم وفيما اختلف فيه النحاة المتقدمون، فيساعدنا المحلل الصرفي بالتزامه قاعدة واحدة دون تناقض أو تغيير.

1- المرجع السابق، ص117، وينظر:
ABDERRAHIM M.A ; Un analyseur morphologique pour l’arabe voyellé ou non 1ére conférence internationale –système d’information et intelligence economique-SIIE-2008 hammamet,tunisie,14-16 février 2008,procedings, tome 2 ;IHE,èdition,tome 2. PP: 324- 339.

2- مأمون حطاب، "التحليل الصرفي للغة العربية باستخدام الحاسوب" مرجع سابق، ص 58.

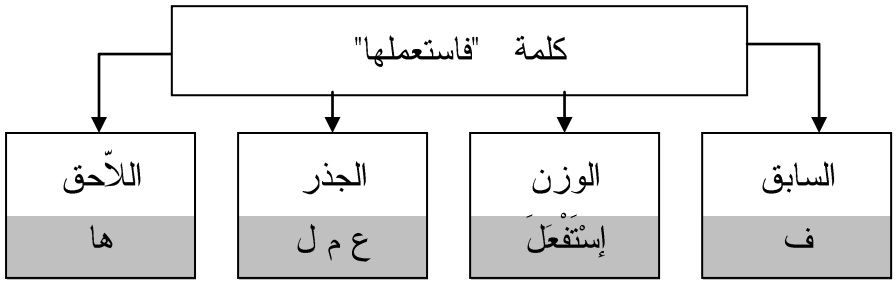
ج- في خدمة الموسوعات وغيرها من الكتب والدراسات: يستفاد من المحلل الصرفي في توفير الوقت على العاملين في الموسوعات[1].

د- مجال التعليم: قد يستخدم المحلل الصرفي في برامج تعليم مادة الصرف، فيزيد فهم الطالب للمادة، ويستخدم في التمرين على موضوعات في الصرف العربي التي يجد فيها صعوبة مثل: موضوع الإعلال والإبدال.

ولأنّ منظومة الصرف تعد المحور الأساسي في منظومة اللّغة العربية لارتباطها بباقي عناصر المنظومة نحوًا، وصوتًا ومعجمًا ودلالةً، وقد قامت الشركة الهندسية لتطوير نظم الحاسبات **RDI** [2] ببناء محلل صرفي آلي © Arab Morpho ليكون قلب تقنيات المعالجة الآلية للغة العربية. ولتبني عليه بعد ذلك تقنياتها الأخرى الهامة للتحليل الصرفي، والتشكيل الآلي، والتعامل مع اللّبس.

5-1-3- أهم الخصائص والإمكانيات التي يقدمها المحلل الصرفي الآلي:

يظهر الشكل التالي كيفية عمل المحلل الصرفي[3] التي تقوم بتحليل كلماتها العربية إلى عناصرها الصرفية وتحديد نوع كل كلمة وجذرها الاشتقاقي ووزنها الصرفي وسوابقها ولواحقها.



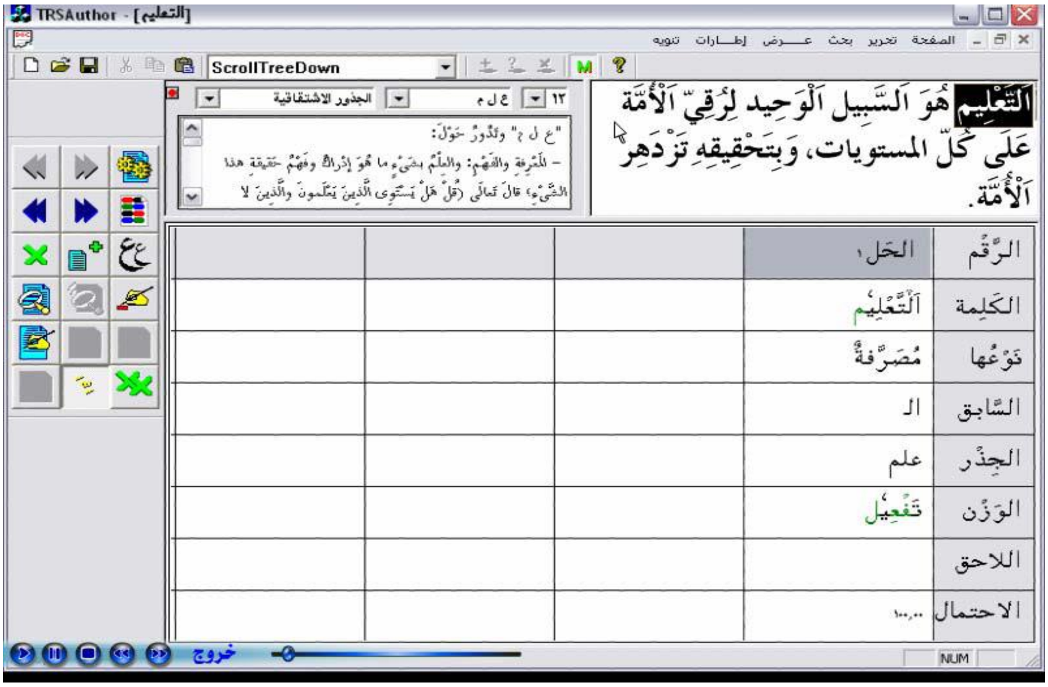
الشكل رقم 17: مثال على آلية التحليل الصرفي

1- ومثال ذلك: موسوعة "المعجم المفهرس لألفاظ الحديث النبوية" التي استمر العمل بها من سنة 1936 م حتى 1987م، حيث توفي صاحب الفكرة والمبادئ فيها سنة 1939م، وأكمل العمل فريقٌ عمل حتى صدر آخر مجلد سنة 1987، فكم من السنوات مضت حتى أنجز هذا العمل باستخدام الجهد اليدوي، بينما لو استخدم المحلل الصرفي لذلك الهدف لأنجز العمل هذا في وقت قصير جدًّا مقارنة معه، طبعًا إذا كانت مادة النصوص المراد تحليلها متوفرة، وهي الآن موجودة ومتوفرة.

2- الشركة الهندسية لتطوير نظم الحاسبات RDI مصر، موقع الويب: -eg. comhttp://www.rdi

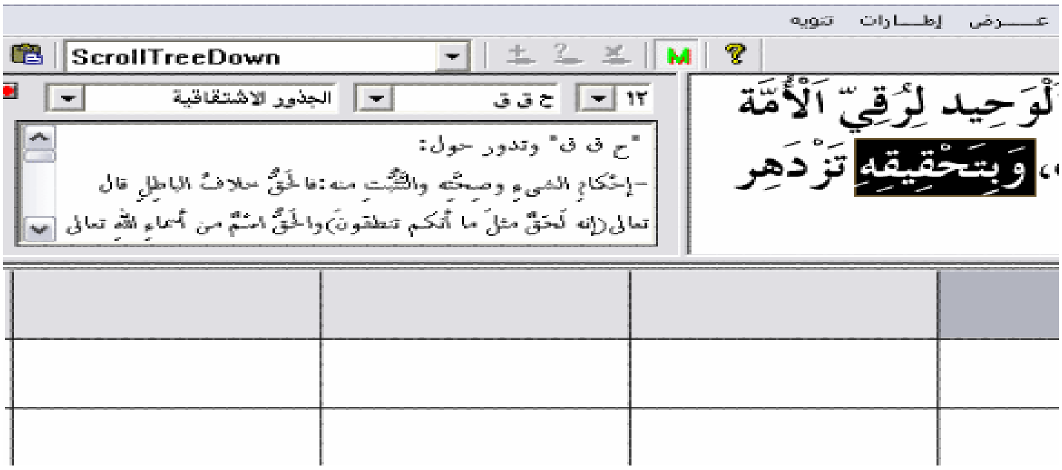
3 - Arab Morpho ©

كما يظهر الشكل أدناه واجهة المستخدم ويظهر فيها التحليل الصرفي لكلمة "التعليم"[1].



الشاشة رقم 02 : التحليل الصرفي لكلمة "التعليم"

كما توفر التقنية معلومات معجمية تفصيلية عن كل محتويات الكلمة – الشكل الموالي يوضح المعالجة التقنية لكلمة "**وبتحقيقه**"[2].



الشاشة رقم 03: معلومات معجمية، المعالجة الآلية لكلمة "وبتحقيقه".

1– المرجع السابق، ص 57.

2– المرجع نفسه، ص 58.

أما الكلمات التي تحتمل أكثر من حل صرفي فإنّ المعالجة الآلية تقوم بإعداد إحصائيات الحلول الأكثر رجحاناً من الناحية الإحصائية، وتقوم بترجيح أحد الحلول وذلك طبقًا لما تدرب عليه البرنامج على النحو الوارد في كلمة **المستويات**[1].

<u>الشاشة رقم 04</u>: المعالجة الآلية للكلمات التي تحتمل أكثر من حل صرفي[2]

5-1-4- مميزات المحلّل الصرفي العربي:

- التغطية الشاملة لمفردات اللّغة العربية بمستوياتها التاريخية والجغرافية المختلفة.
- إمكانية زيادة تدريب النظام (حسب الطلب) على نوع معين من النصوص لزيادة دقة التحليل.
- التعامل مع الأخطاء الشائعة وتصويبها.
- التعامل بمرونة مع الكلمات المدخلة سواء أكانت مضبوطة أم لا.
- سرعة المعالجة الصرفية للنصوص والتي تتجاوز 300 كلمة في الثانية.
- توليد التصنيفات الدلالية المعجمية الممكنة لكل كلمة تم تحليلها صرفيًا.

1- المرجع السابق، ص 59.

2- المرجع نفسه، الصفحة نفسها.

5-2- المحلل النحوي: Analyse syntaxique:

تنطوي معالجة النّحو آليًا على تحديات قاسية، سواء على الصعيد اللّغوي أو الحاسوبي لذا فقد كانت وما تزال منهلاً وفيرًا لأساليب البرمجة الحديثة؛ إنّ المعالجة الآلية للنّحو العربي هي التحدّي الحقيقي لعلماء اللّغة وعلماء الحاسوب وهي جديرة بلا شك بأن نوليها أقصى درجات اهتمامنا ليس فقط لكونها أحد المقومات الأساسية لمعالجة اللّغة العربية، بل أيضًا باعتبارها وسيلة لا غنى عنها لتحديث التنظير اللّغوي بصفة عامة.

على أساس مبدأ النّحو العام والدور الأساسي الّذي يلعبه كمستوى من مستويات التحليل للمعالجة الآلية للغة العربية، بدا منطقيًا أن نطرح السؤال التالي: لماذا نشغل أنفسنا بمعالجة نحوية آلية خاصة بالنّحو العربي، ما دام هذا النّحو ما هو إلاّ مجردّ خاصة من النّحو العام؟ أو بقول آخر؛ لماذا لا يكون هناك معالج نحوي عام يصلح للمعالجة الآلية؟

يرى نبيل علي أنّه لا يوجد أي اعتراض أساسي من النّاحية النّظرية البحتة على مثل هذا التصوّر، إلاّ أنّه من النّاحية العلمية، لا يبدو واقعيًا، فالنّحو العام بطبيعته موغل في التجريد لدرجة يصعب معها تطويعه للمعالجة الآلية، دون النزول إلى مستويات أدنى من التجريد، فلا بديل حاليًا إلاّ التّصدي لهذه المشكلة الأساسية على المستوى الخاص باللّغة العربية[1].

5-2-1- مفهوم المحلل النّحوي:

هو برنامج آلي يقوم بتحليل بنية الجملة، وذلك من حيث ترتيب عناصرها والعلاقات التركيبية والوظيفة التي تربط بينها، فالمحلل النحوي له صفتان: الإصدار والتدقيق إذا يقوم المحلل بإصدار عدد كبير من الجمل ذات تركيبات مختلفة كما أنّه يقوم بالتدقيق في الجمل التي يدخلها المتعلم ويصحح الأخطاء[2].

1- اللّغة العربية والحاسوب، ص 390/ 391.

2- سناء محمد حافظ غانم "مواكبة العربية للغة المعلومات وعصر التقنية" مقال منشور.موقع الويب: http:// www. Iseco. Org. Ma

ويمكن التعبير عن اللّغة باستخدام نموذج رياضي كالآتي[1]:

اللّغة=[م، ن، ق، ج] وترمز الأربعة أجزاء إلى الآتي:

- م: مجموعة من الرموز الطرفية[2]: وهي مفردات المعجم وما يشتق منها من كلمات.
- ن: مجموعة من الرموز الأطرفية[3]: والتي تعبّر عن تصنيف النّظام النّحوي المتبع.
- ق: مجموعة من قواعد الإنتاج[4]: والتي تصنف الرموز الأطرفية معبرًا عنها برموز طرفية ورموز لا طرفية.
- ج: مجموعة الجمل التامة.

5-2-3- العناصر الأساسية للمعالج النحوي الآلي للجمل العربية:

إن المكونات الأساسية للمعالج النّحوي للجمل العربية تمكن النظام الآلي على التعامل مع الأطوار المختلفة لتشكيل النّصوص العربية: تامة التشكيل، ناقصة التشكيل، والخالية من التشكيل، ويمكن توضيح ذلك من خلال الشكل 18، سنعرض في جولة سريعة الجوانب الأساسية للمعالج النحوي[5].

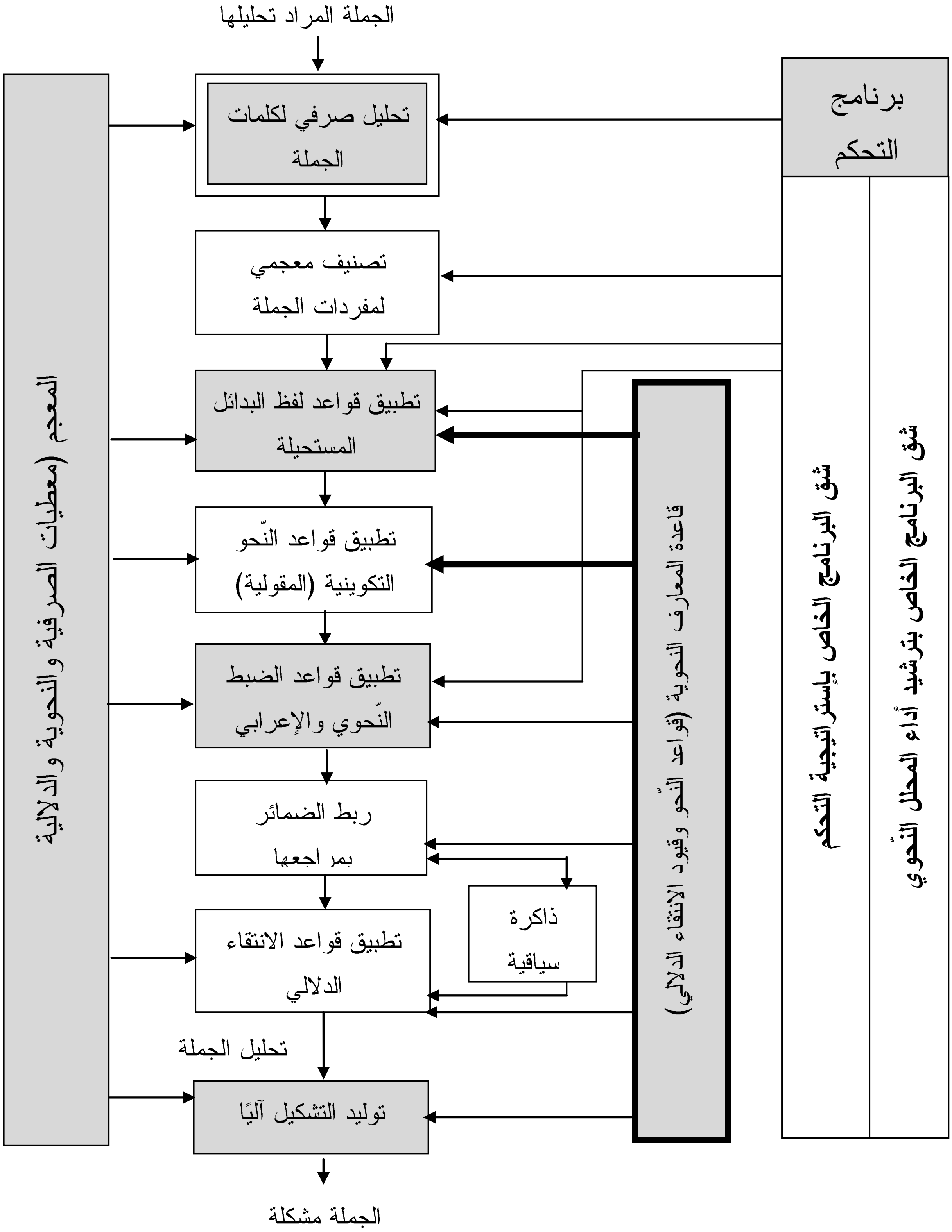
1 –Kenneth Slonneger, and others « Formal syntax and semantics of programming language, addison- wesely publishing company ; 1995 ,PP: 02.

2 – Symboles terminaux.

3 – Symboles Non terminaux.

4 –Règles de production.

5– اللّغة العربية والحاسوب، سابق، ص 406.



الشكل رقم 18: المكونات الرئيسية للمعالج النحوي للجملة العربية[1]

1- المرجع السابق، ص 407.

يتكون المعالج الآلي من عدّة مكونات رئيسية هي[1]:

(1)- المعجم: بعد ميكنته (حوسبته)، متضمناً المعطيات النحوية والدلالية للمفردات.

(2)- قاعدة المفردة النحوية، وتشمل قواعد النّحو، وقيود الانتقاء الدلالي التي تضمن توافق الأفعال مع عناصر إسنادها والأسماء مع مكملاتها وملحقاتها.

(3)- روتينات برنامج المعالجة وهي عبارة عن سلسلة من الإجراءات البرمجية التي تتعامل مع المعجم، وقاعدة المعارف النحوية.

(4)- برنامج التحكم وهو الّذي يحدّد التسلسل الذي يتم به تنفيذ الروتينات المختلفة بحيث يمكن تحليل الجملة آليًا في أقصر وقت ممكن، وبأقل موارد ممكنة.

تظل المهمة الرئيسية للتحليل النّحوي الآلي هي توفير المعطيات اللاّزمة للتحليل اللّغوي الأعمق، وهو الفهم الآلي للنصوص اللّغوية إلى جانب ذلك هناك تطبيقات عملية تستخدم المعالج النّحوي الآلي نذكر منها[2]:

(1)- المصحح الإملائي للأخطاء النّحوية.

(2)- التخاطب مع قواعد البيانات باللّغة الطبيعية.

(3)- الترجمة الآلية من وإلى اللّغة العربية.

(4)- تعليم النحو بواسطة الحاسوب.

- ملاحظة:

لقد تم اقتراح محللين نحويين من قبل باحثين عرب باستخدام الحاسوب الآلي يعملان بأسلوبين مختلفين الأول يعتمد على توصيف تراكيب الجمل العربية بأسلوب Clause de la grammaire définie وهو يسمح بالتحقق النّحوي من سلامة العلاقات الدلالية بين الكلمات في الجملة، كما يسمح بالتأكد من سلامة المعنى في بعض الحالات.

أما المحلل الثاني: فيحلل الجملة بفحص كل كلمتين متتاليتين فيها ومدى صلاحية تتابعهما، مع الأخذ بعين الاعتبار مطابقة الجملة كلها لبعض خصائص الجملة العربية النّحوية والخصائص الصرفية المتعلقة بكل كلمة، ويضبط هذا المحلل الحرف الأخير، ويعالج الجمل العربية البسيطة[3].

1- المرجع السابق، ص 406، 407.

2- المرجع نفسه، ص 416.

3- دليل الباحث إلى اللّسانيات الحاسوبية العربية،سابق، ص 148، 149.

وجملة القول أنّ معالجة النّحو آليًا على نطاق اللّغات الطبيعية قد قطعت أشواطًا في دمج وامتزاج الدراسات اللّسانية بالدراسات الحاسوبية، وأصبحت العلاقة بينهما علاقة تبادل واقتراض، وأي تنظير للغات الطبيعية أصبح مرهونًا بمعرفة دقيقة للأساليب المنهجية والتحليلية لعلوم الحاسوب، الأمر الّذي يؤهل اللّغوي لكي يستخدمها في صياغة جديدة للمنظومة النّحوية، وتطوير علم الدلالة وكل ما يتعلق بالصناعة المعجمية.

5-3- المحلل المعجمي: Analyse lexical:

إنّي لا أقصد بالمعجم هنا القاموس المتداول لمعرفة معاني المفردات، وإنّما نقصد به المعجم الآلي[1] الّذي يتكون من قاعدة بيانات تتركب من جزئيات تحتاجها المعالجة الآلية في مستوياتها الثلاثة الأخرى، هذه الجزئيات كالمعلومات الصرفية[2] والتركيبية للكلمات المستخدمة في اللّغة، فكل رحلة تحليلية سترجع للمعجم للتدقيق ومعرفة ما تحتاجه من صفات للمفردة. ويكاد يتفق خبراء الحواسب على وجود نوعين من المعاجم الآلية العربية:

5-4- المعجم الآلي للمفردة في مستواها الإفرادي:

حيث يتضمن هذا المستوى تقديم وصف صوري للمفردات المكونة للغة حول الكيفيات التي ترسم بها الحروف المؤلفة للوحدة المعجمية، ووصف دقيق حول كيفية نطق الكلمات والحركات التي تحدد رسم المتوالية اللّغوية ووضع رصيد صوري لجميع المعطيات المورفولوجية التي قد تتجلى من خلالها الوحدة المعجمية وتقديم كل المعلومات المرتبطة بالقيم الصرفية والنّحوية إضافة إلى المعلومات الدلالية التي تحمل من طرف الجذور والمباني الصرفية التي تتجلى من خلالها الكلمة[3].

5-5- المعجم الآلي للمفردة في مستواها التركيبي:

هو تكميلي للمعطيات التي تُسجل للمفردة في مستواها الإفرادي، يتم هذا المستوى تقديم كل المعطيات اللّغوية التي يمكن أن تكتشف المفردة من خلال ما يقتضيه المستوى التركيبي الّذي يعطي للمفردة وظيفتها اللّغوية المناسبة في سائل السياقات المختلفة[4].

---

1 –Dictionnaire automatique.

2– للمزيد من بسط القول في هذا المبحث ينظر: صافية زفنكي، التطورات المعجمية والمعجمات اللغوية العامة والعربية الحديثة، منشورات وزارة الثقافة، دمشق، د.ط، 2007، ص 210.

3– محمد الحنّاش "نظرية حاسوب– لسانية في بناء معاجم آلية للغة العربية– مجلة التواصل اللّساني المجلّد الثاني، العدد الثاني، 1990، مطبعة النجاح، الدار البيضاء المغرب، ص 46 وما بعدها.

4– المرجع نفسه، الموضع نفسه.

وقد جاءت بعض الدراسات والأبحاث لتبحث تركيب هذه المعاجم، منها دراسة لباحث مغربي[1] اقترح معجمًا شاملاً (جذور الكلمات العربية، المفردات البسيطة، المفردات البسيطة المعربة، المفردات المركبة، المفردات المركبة المعربة)، ومنها دراسة لفريق عمل في القاهرة تحت عنوان: "قاموس عربي آلي: مستلزمات بناء قاعدة معطيات للمفردات اللّغوية العربية)، ولكنها ذات صبغة تمهيدية[2] وما يلزمنا عند بناء المعاجم أن تكون متصفة بالشمول والاطراد والوضوح وقابلة للتوسع المستقبلي، وأن تخدمنا من الناحية التطبيقية.

يمثل بناء محلل معجمي آلي على أساس قاعدة معطيات صرفية نحوية ودلالية أحد أهم المقومات الأساسية لبناء نظم المعالجة الآلية للغة العربية، إذ تقوم فكرة المحلّل المعجمي أو المعجم الالكتروني على تحليل تعريفات معاني الكلمات منطقيًا ودلاليًا لصياغتها في هيئة شبكات دلالية[3] أو مخططات مفاهيم[4]، تولد هذه المعاجم نواتج فرعية عديدة لدعم البحوث اللّغوية والمعجمية وبناء نظم بنوك المصطلحات[5].

نكتفي بذكر أهمية المحلل المعجمي أو ما يعرف بالمعجم الآلي من منظور المعالجة الآلية للغة العربية، وعلاقته بالمستويات اللّغوية الأخرى خشية الإطالة في الموضوع، وتجدر الإشارة أنّه قد خصّصنا في الباب الثالث من هذا البحث فصلاً للحديث عن المعجم الالكتروني للغة العربية من منظور لساني حاسوبي.

سادسا- المعالجة الآلية للدلالة في اللغة العربية:

إنّ للدلالة أهمية كبيرة في المعالجة الآلية للغة العربية، فلا يمكن أن نتصور إنجاز قاعدة معطيات للنصوص العربية بدون معلومات كافية عن دلالة الألفاظ والعلاقات الدلالية بين الكلمات المكونة لتلك النصوص.

تتطلب المعالجة الآلية للدلالة في اللّغة العربية أمرين أساسيين: **أولاً:** ضرورة الاستناد إلى إطار لساني صوري قادر على توصيف مستويات اللّغة الهدف، **ثانيًا:** ضرورة الاعتماد على مقدار كبير من المعلومات التي ترتب وتبوب انطلاقًا من قواعد بيانات آلية.

1- محمد الحنّاش، "نظرية حاسوب- لسانية في بناء معاجم آلية للغة العربية: محاولة في التأصيل المسجل العالمي لندوة استخدام اللّغة العربية في تقنية المعلومات، مكتبة الملك عبد العزيز العامة، الرياض 1993.

2- محمود الصيني " نحو معجم عربي للتطبيقات الحاسوبية"، السّجل العلمي لندوة استخدام اللّغة العربية في تقنية المعلومات، مكتبة الملك عبد العزيز العامة، الرياض، 1993، ص 513.

3 -Sémantique des Filets.

4 - Graphes conceptuels.

5- نبيل علي "مسح للمحتوى العربي الرقمي وبرمجياته وتطبيقاته وتقييم احتياجاته، دراسة أعدت ضمن تطوير مشروع تعزيز تطوير صناعة المحتوى العربي في حاضنات تكنولوجيا المعلومات والاتصالات، ص 74.

إن وضع معجم آلي حديث يستوجب إعطاء أهمية لخصائص الكلمات ومدلولاتها وذلك بغية تجاوز ثغرات المعاجم التقليدية، ومن ثم إيجاد حل ناجع للكثير من القضايا اللّغوية التي تواجه المعالجة الآلية للغة العربية من قبيل اللّبس الدلالي للمفردات والعبارات والنصوص والترجمة الآلية،... وغيرها من التطبيقات الآلية.

فقضية الترادف والمشترك اللّفظي والتضاد وغيرها من العلاقات الدلالية لها قيمة كبيرة في بناء معجم آلي وفق معطيات بيانية محكمة ومستمدة من المفاهيم اللّسانية وبالتالي التمكن من حل مشاكل التحليل اللّغوي الآلي بصفة عامة.

1.6- أهمية المعالجة الآلية للدلالة في العربية:

تكمن أهمية معالجة الدلالة آليًا في إظهار أهمية قضايا الترادف والمشترك اللّفظي والتضاد باعتبارها مفتاحًا لحل الكثير من القضايا اللّغوية على مستوى المعالجة الآلية، ونذكر منها ما يلي[1]:

1- وضع الخصائص والسمات للكلمات والحقول الدلالية.

2- التحليل الآلي للنصوص.

3- فهم النصوص.

4- حل قضايا اللّبس اللّغوي.

5- الترجمة من اللغة المصدر إلى اللّغة الهدف.

2.6- الدلالة من منظور لساني- حاسوبي:

لقد اهتم علماء اللّغة بظواهر الاشتراك والتضاد والترادف والتعميم في إطار ما يسمى العلاقات الدلالية، أو ما يمكن التعبير عنه بنظرية الحقول الدلالية التي تعود في الأساس إلى نظرية الدلالة البنيوية.

والعمل في إطار معالجة الدلالة حاسوبيًا يتطلب مراعاة أمرين أساسيين:

1- ضرورة الاعتماد على إطار لساني نظري يتوفر على مفاهيم قادرة على صياغة الظواهر الآلية اللّغوية وفق قاعدة معطيات من الخوارزميات.

2- ضرورة بناء قاعدة بيانات للمعطيات اللّسانية.

1- عمر مهديوي، سلوى السّيد حمادة، "المعالجة الدلالية الآلية للغة العربية" نحو بناء قاعدة بيانات معجمية للعلاقات الدلالية بين الكلمات" مجلة النادي العربي للمعلومات، موقع الويب: مجلة العربية 3000 - web net : http://www.arabcin: -
وينظر: مجلة "فكر ونقد" العدد 82، المغرب، السنة 2006.

3.6- قواعد بيانات المعاني والدلالة للمعالجة الآلية:

إن قاعدة البيانات الدلالية على جانب كبير من الصعوبة بالنسبة للغات الحديثة فكيف باللّغة العربية الواسعة المعاني والدلالة[1] فالهدف من المعالجة الآلية بناء قواعد بيانية معجمية للمفردات العربية ترمي إلى صناعة معجم آلي عربي عصري، يتخطى النقص التي تعاني منها المعاجم العربية الورقية.

وتتجلى خطوات المعالجة الآلية للدلالة في العربية[2]:

1-جمع المادة المعجمية من الكلمات وسياستها- ويتم من خلالها التصنيف والتخزين.

2-تحليل النصوص عن طريق التعديل والإضافة.

3-تبويب المادة المعجمية تحت حقول دلالية وفي مستويات مختلفة.

4-مراعاة الكلمات المستخدمة في شرح المعاني لا تخرج عن نطاق كلمات المعجم.

ويكمن بناء قاعدة بيانات معجمية دلالية لتحديد السّمات الدلالية للمترادفات وذلك عن طريق[3]:

-تحديد المعنى الدقيق عن طريق السمات الدلالية للمترادف.

-إثراء المعجم عن طريق البحث في العلاقة بين معاني المترادفات والمشترك اللّفظي.

وجملة القول من خلال هذا العرض السريع للمعالجة الآلية للدلالة في العربية هو إبراز أهمية بناء قواعد بيانات معجمية موجّهة أساسًا لخدمة المعجم الآلي للغة العربية، وذلك بالتركيز أساسًا على مستوى الدلالة باعتبارها أصعب مستوى[4]، إن لم نقل أنّها تشكل تحديًا كبيرًا بالنسبة للمعالجة الآلية للغة الطبيعية، لأنّ الدلالة ترتبط أصلاً بالسياق، وهو ما يتعذر على الحاسوب إدراكه، وتنميطه، وفهمه على الرغم من تطور البرامج الحاسوبية والذكية، ولابد من توافر قدر كبير من المعلومات والمعطيات حول الكلمات والعلاقات الدلالية بينها إن أردنا صناعة معاجم الكترونية للغة العربية تشمل الدلالة والصرف والنّحو[5].

1- أ. د محمد زكي خضر "نحو معالجة الدلالة في اللّغة العربية عبر قواعد البيانات" "دراسة أولية لنص القرآن الكريم" المؤتمر الوطني السابع عشر للحاسب الآلي (المعلوماتية في خدمة ضيوف الرحمن) جامعة الملك عبد العزيز، المدينة المنورة، 2004، ص 12.

2- المعالجة الدلالية الآلية للغة العربية: نحو بناء قاعدة بيانات معجمية للعلاقات الدلالية بين الكلمات، سابق.

3- المرجع نفسه.

4- يمثل توصيف النظام الدلالي للغة العربية نواة المعالجة الآلية وعموده الفقري؛ ذلك أن جل عمليات المعالجة الآلية للتراكيب اللغوية تستند بشكل من الأشكال عليه وتحيل إليه. ويعد عند المعنيين باللسانيات الحاسوبية أعسر المباحث تناولاً لتعلق دلالات الألفاظ بداهة بـــ (الفهم) البعيد المنال عن الحاسوب. ذلك أن الدلالة قد تكون معجمية أو صرفية أو نحوية أو مجازية أو إيحائية، وهي تعني في الذاكرة البشرية كل ما يتداعى إلى ذهن ابن اللغة من علائق ومعان تتصل باللفظ.

وواقع الأمر أن توصيف النظام الدلالي لن يكون على سنة غيره من الأنظمة اللغوية السابقة، فلسنا نرسم قواعد شاملة تنتظم الجزئيات، وإنما نلتزم فيه رفد كل مفردة معجمية بتفاصيل ما يرد إلى الذهن من علائق دلالية، على أن النظام الدلالي سيبقى يفيد من قواعد بقية الأنظمة اللغوية بغية أن تكون معالجته للمفردات أكثر آلية.

ولعل الأولى ونحن في سبيل إنشاء هذا النظام الدلالي أن تستعين ببنك للنصوص العربية الفصيحة يتم اختيار نصوصه سلفاً وفق معايير علمية يرتضيها المتخصصون في العربية بحيث يكون توصيفنا للنظام الدلالي مبنياً على استقصاء لأحوال الكلم وسياقاتها المستعملة فعلياً في المتون العربية التراثية والمعاصرة؛ لا أن يُترك الحبل على الغارب واجتهاد المبرمج أو مرجعية الأفراد. إن المعجم الذهني الجمعي ينبغي أن يُتمثّل بتدبير منهجي موضوعي قابل للتحديث كلما طرأت الحاجة إلى توليد الألفاظ وكلما استجدت المعاني والدلالات.

5- للمزيد من التفصيل ينظر: د/ محمد زكي خضر "نحو معالجة الدلالة في اللّغة العربية عبر قواعد البيانات" دراسة أولية لنص القرآن، سابق.

4.6- ملاحظة:

لم تتضمن الدراسة الحالية دراسة مستفيضة خاصة بالمعالجة الآلية لعنصر الدلالة في اللّغة العربية أسوة بما اتبعته في المعالجة الآلية للصرف العربي، فمقارنة الدلالة بالصرف والنّحو والمعجم يعد عنصر الدلالة من أقل الفروع اللّغوية فيما يخص التباين اللّغوي[1]، فبالرغم من أن للدلالة لها دور كبير في صناعة المعاجم الالكترونية إلا أنّ دراستي التطبيقية كانت مقتصرةً على الجانب الصرفي في بناء قاعدة معطيات للمعجم الالكتروني، كون أن هناك قصور كبير في الدراسات الدلالية على مستوى اللّغة العربية بما لا يوفر الحد الأدنى من الحصاد النظري الذي يمكن أن تؤسس عليه تصورات محددة عن معالجة الدلالة العربية آليًا[2].

5.4- خلاصة:

وكخلاصة عامة لهذا المبحث نستنتج أن أهم تطبيقات المعالجة الآلية للغة العربية تكمن في الثالوث: المعالج الصرفي الآلي، والمعالج النحوي الآلي، والمحلل المعجمي "المعجم الآلي" الذي يعدّ بمثابة القاعدة الأساسية التي يمكن من خلالها استثمار العلاقة بين هذه المستويات في بناء برامج الكترونية، تعليمية، وبتوفير هذه النماذج الآلية يجب على الباحثين استغلال إمكانياتها العديدة في بناء برامج حاسوبية باللّغة العربية ولا بديل عن ذلك لتطوير قدراتنا الذاتية في هذا المجال في سبيل تحقيق قفزة نوعية للنهوض بمشروعات طموحة تخدم اللغة العربية على وجه الخصوص.

فإمكانية الوصول إلى الكلمة عن طريق المحلل النحوي والصرفي من خلال جذرها أو سابقتها أو لاحقتها يحقق لمستخدم المعجم الالكتروني جملة من المميزات أهمها[3]:

- "استرجاع الكلمات بسهولة من قاعدة البيانات المعروضة".
- استغناء الباحث عن اقتناء عدد من المعاجم بالرجوع إلى قاعدة البيانات التي يمكنه أن يأخذ منها ما يشاء.

وفقاً لِما سلف ذِكرهُ في مقدمة البحث، عالج هذا الفصل موضوع مذكرتنا من زاوية خاصة، فهو ينفردُ بدراسة المعالجة الآلية للغة العربية والتعريف بها وخصائصها، ومجالات تطبيقاتها، أما الجزء الثاني فأردناه للمعالجة الآلية للغة العربية،النظام الصرفي كأنموذج للدراسة

1- نبيل علي، اللّغة العربية والحاسوب، سابق، ص 58، 59.

2- المرجع نفسه، ص 532.

3- ينظر: صناعة المعجم الحديث، ص 183.

التطبيقية مع إبراز أهم تطبيقات المعالجة الآلية للغة العربية في إعداد معجم إلكتروني للغة العربية.

تمثل استقراؤُنا من خلال هذه الدراسة في اعتبارها منطلقاً واعداً لإعداد مشاريع لسانية في مجال المعلوماتية وفق منظومة اللغة العربية لبناء برامج حاسوبية العربية.

وفي ضوء ما سبق، نؤكد أن المعالجة الآلية للغة العربية و المعالجة الآلية للصرف العربي على وجه الخصوص قد قطعا أشواطاً هامة خلال العقدين باهتمام المختصين في علم الحاسوب وعلم اللغة لتطوير وسائل وأساليب المعالجة الحاسوبية للغة، وكذلك الإنجازات الإبداعية التي تحققت في هذه المعالجة، في جوانب الصرف، والتركيب، والدلالة، وبناء المعاجم، وغيرها.. مما سمح لنا بتكريس الفصل الثاني للحديث عن رؤية جديدة لمعجم آلي للمعالجة الآلية للغة العربية من خلال ما تمت دراسته في هذا الفصل النظري محل التطبيق في استخدام البرمجة الحاسوبية لبناء معجم إلكتروني للمعالجة الآلية للغة العربية كدراسة تطبيقية لهذا البحث.

الفصل الثاني

هندسة المعجم الإلكتروني

رؤية جديدة لمعجم آلي للمعالجة الآلية للغة العربية

أولاً- الإطار العام للمعجم العربي الإلكتروني

ثانياً- المعجم العربي الإلكتروني: نماذج لقاعدة البيانات المعجمية للمعاجم الإلكترونية

ثالثاً- بناء قاعدة بيانات معجمية لمعجم إلكتروني

رابعاً- سبل تطوير قاعدة بيانات معجمية لبناء معجم آلي عربي

أولا- الإطار العام للمعجم العربي الإلكتروني:

تطرح المعالجة الآلية للغة العربية بشكل عام جملة في الصعوبات وتزداد المشكلة تعقيدا عندما يتعلق الأمر ببناء معجم إلكتروني للمعالجة الآلية للغة العربية، وتجب الإشارة هنا إلى أن اللغات الأجنبية كالإنجليزية والفرنسية وغيرها قد قطعت أشواطا مهمة في مجال البحث اللساني الحاسوبي والهندسي فالوضع بالنسبة للعربية مختلف، إذ الأبحاث اللسانية والحاسوبية ما تزال في بداية الطريق.

في هذا الفصل، سأحاول تقديم المفاهيم والأدوات المناسبة لبناء المعجم الآلي للغة العربية من خلال منظومة جديدة للمعاجم الالكترونية التي استغلت التقنيات المعلوماتية الحديثة في معالجة اللغات الطبيعية في مستوياتها المختلفة.

ان بناء معجم آلي متكامل للغة العربية يرتكز على وصف المتواليات اللسانية بسيطها ومركبها مع ربط هذا الوصف بالمستويات الصرفية والتركيبية والدلالية .ومع تفاقم المشاكل التي تعترض العلاج الآلي للغة العربية مثل التدقيق الإملائي للنصوص المكتوبة والتشكيل الآلي وإشكالية غموض معالجة المتواليات المركبة، كانت الحاجة ملحة لبناء مثل هذا المعجم.

فلقد أصبح من المؤكد اليوم أن المعالجة الآلية للغة الطبيعية تعتبر ميدانا أساسيا في تطوير مختلف الأنظمة والبرامج الحاسوبية، وأصبحت من الأولويات الأساسية الّتي تعتمد عليها اللغات الطبيعية في تطبيقاتها ومعالجة أنظمتها[1]. فبدخول الحاسوب في مجال الدراسات العربية بشكل عام والمعجمات بشكل خاص، وعلى الرغم من قلة الاستفادة المعجمية من هذه الدراسات قياسا إلى فائدته في شتى المجالات الأخرى، إلا أن التطور التقني الحديث الذي طرأ على الصناعة المعجمية الحاسوبية مؤخرا يجعلنا نتفاءل بإمكانية التوسع في الانتفاع بما يقدمه الحاسوب في إمكانات وبرامج لخدمة المعجم والصناعة المعجمية.

فالمعالجة الآلية للمعجم العربي تحتاج إلى نظرة ثورية في تصميمه وتوسيع نطاق استخدامه، وتسهيل عملية اندماجه داخل النظم الآلية، وفي ظل هذه الشروط والتحديات **هل يمكن لنا التصريح بمشروع نظام آلي يكون بمثابة معجم الكتروني للمعالجة الآلية للغة العربية؟، وهل بإمكاننا بناء قاعدة معطيات ممنهجة لصناعة المعجم الالكتروني للغة العربية؟.**

1- للتذكير فإن كل عمل في المعالجة الآلية لكل اللغات الطبيعية يقتضي شرطين:

*/أولهما بناء قاعدة معطيات للمفردات العربية بسيطها ومركبها.

*/وثانيهما الاستناد إلى نظرية لسانية محددة تستطيع وصف الظواهر اللغوية بدقة عالية .

وهذان الشرطان يجنبان وضع القواعد اعتمادا على الظن مما يعرقل التطبيق الآلي، كما يجنبنا مزالق وضع قواعد جزئية في البرنامج المولد في بيانات اللغة . ومع ذلك فقد تعددت المقاربات المنهجية والتصورية في معالجة الظواهر المعجمية في الكثير من اللغات الطبيعية بتعدد الاساليب المنبثقة عن تصور وتشعب أساليب الهندسة الإلكترونية.

1- المعجم الإلكتروني: إشكالية التعريف:

1-1- تحديدات أولية لمصطلح المعجم الإلكتروني:

يبدو أن تعدد التصورات والمرجعيات العلمية والنظريات وكذا الأساليب والتقنيات، ساهم بشكل كبير في بلورة أنواع وأصناف من المعاجم الآلية. فهي مكون أساسيٌّ لكل نظام يروم معالجة اللغات الطبيعية. هنا سيصبح البناء مرحلة بعد التصور والإدراك، لأن هناك نقلة نوعية من الطبيعي إلى الاصطناعي، مما يؤسس لحوامل جديدة تُيسِّر الولوج والتحليل والبحث عن الدلالات والمعاني الممكنة للكلمات / المداخل، وأيضا الانطلاق من المفاهيم للحصول على الكلمات التي تعبر عنها. فحاجيات المستخدم لبيانات المعاجم تختلف من واحد لآخر، مما أفرز أنواعا عديدة تجيب عن مختلف هذه الاستعمالات[1] ،ومنها : المعاجم الأحادية والثنائية والمتعددة اللغات ومعاجم المترادفات والموسوعات وبنوك المصطلحات[2] وغيرها كثير. **فما المقصود بالمعجم الإلكتروني؟ وماهي أنواعه؟**

تلك أهم الأسئلة سنقترح عناصر للإجابة عنها إما حينها وإما خلال ثنايا فصول هذا البحث الذي سينصب على تصور عام للمعاجم الإلكترونية العربية كما سنرى. إلا أن ما يثير اهتمامنا، هو ما أنجز في المختبرات العالمية للبحث الآلي لأنها، بلا شك، متقدمة بشكل كبير عما نفكر فيه على مستوى اللغة العربية، ذلك فيما أعدته مجموعات البحث في تطبيق نظرية العجم – التركيبي، وسنكتفي هنا بالوقوف عند أعمال موريس كروس وزيليغ هاريس سابقا، وغيرهما ممن أسهموا في بناء معاجم آلية للغات العالمية ومنها اللغة الفرنسية في مختبر المعالجة الآلية والتوثيق اللساني.

سأحاول في هذا المبحث، إلقاء الضوء على جوانب مهمة من هذه الأطروحة، حيث سنبدأ بتقديم مفهوم المعجم الإلكتروني ومقارنته بالنسخة الورقية، ثم نبين أهمية المعجم الإلكتروني العربي في مجتمعنا الذي يتحول بسرعة إلى مجتمع المعرفة، ثم أضع تصوراً متكاملا للمعجم الآلي العربي من حيث الكفاية المعجمية أولا، ومن حيث استفادته من الأنظمة المعلوماتية المتقدمة ثانياً، وذلك بأن تعالج جانب من معجمنا العربي[3]، علاجاً آلياً هدفا نحو تحقيق تمثيل أعمق لمعلوماته في إطار تحضير نظام لغتنا العربية لتحديات التكنولوجيا وعصر المعارف. ثم نتعرض إلى أهمية المعجم الإلكتروني بالنسبة للمستخدم العربي ، ونمر بعد ذلك إلى تفصيل طرق هيكلة وبناء المعجم العربي الحديث من خلال التعرف على كيفية بناء قاعدة

---

1 – Michael Zock – John Carroll , les dictionnaires électroniques revue TAL n 44, 2003 p .7

2– علي القاسمي، نحو تطوير بنوك المصطلحات (آداة للبحث المصطلحي والعلمي)، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 28، الرباط، المغرب، 1987، ص 79،80.

3– اثبتت الدراسات والتجارب اللسانية قدرته الإنتاجية العالية في كفاية المتكلم العربي حيث تصل الى حوالي 40 %.

بيانات معجمية للمعجم الإلكتروني . ونختم بمناقشة موضوع تقييس المعاجم الإلكترونية والجهود التي تبذل في هذا المجال بالنسبة للغة العربية.

1-2- تعريف المعجم الإلكتروني:

إن المقصود بالمعجم الإلكتروني الذي يبدو غامضا إلى حد كبير، هو ما يحيل تارة إلى المكون المعجمي للنظام المعالج للتحليل والتوليد ، ما يروم ضبط الموارد المعجمية الأساسية التي تساعد الإنسان على إنتاج اللغة وانجازها (بما يُسمَّى بالقراءة المعاونة بالحاسوب) ومع ذلك فهما يؤديان نفس الوظيفة المنوطة بهما، كما يصلحان لكلا المستخدمين بشرا كان أم آلة[1].

ظهر مفهوم المعجم الإلكتروني[2] ليتيح الفرصة لبناء مفرداته على الحاسوب، وعلى الرقائق الإلكترونية، بحيث أصبح في الإمكان الحصول على أية مفردة بكل يسر، بل وتوظفها في تطبيقات أكثر تعقيدا كمعالجة النصوص والتدقيق الإملائي والترجمة الآلية وغيرها. ومن ثمة ظهرت قواميس عامة، وأخرى متخصصة في المجالات وقطاعات عديدة. ويمكن النظر إلى المعجم الإلكتروني نظريا، من زاويتين : زاوية اللسانيات الحاسوبية، زاوية المعجمية أو المعجماتية الحاسوبية التي تهتم بكيفية صناعة المعاجم المحوسبة والإلكترونية، وتحديد مبادئها وآلياتها.

ورغم المجهودات المبذولة في الصناعة المعجمية التقليدية فإنها محدودة من حيث الحجم (monotàches) وإمكانية الولوج إليها، إذ تنحصر في طرق الترتيب المألوفة

---

1 – للمزيد من التفصيل ينظر: Michael Zock – John Carroll , les dictionnaires électroniques revue TAL n 44, 2003 p

2– المعجم الإلكتروني هو نسخة حاسوبية معدلة من النسخة الورقية. فهو يتكون من عدد كبير من المداخل يحتوي كل واحد منها على المعلومات التي يمكن تجميعها حوله. تختلف هذه المعلومات من معجم إلى آخر حسب الأهداف التي بني من أجلها وأصناف المستخدمين المستهدفين. وقد بدأ الاهتمام بالمعاجم الإلكترونية منذ منتصف القرن الماضي؛ حيث اقتصر في البداية استعمال هذه المعاجم كموارد لغوية للتحليل الآلي للغات الطبيعية على المستوى الصرفي والنحوي والدلالي، فكانت المعاجم بمثابة قواعد بيانات تحتوي على معلومات مشفرة لا يفهمها إلا البرنامج الذي يستغلها. وقد تطورت وتنوعت هذه المعاجم من حيث المحتوى وكذلك من حيث الهيكلة لمواكبة تطور تطبيقات المعالجة الآلية للغات الطبيعية التي تستغلها مثل: تقطيع الجمل، تحليل النصوص واسترجاعها، البحث عن المعلومات، التدقيق الإملائي، التلخيص الآلي للوثائق والترجمة الآلية. فإذا أخذنا على سبيل المثال المعجم المصمم للتدقيق الإملائي نجده يختلف تماما عن المعجم المصمم للتحليل النحوي. فالأول يقتصر على قائمة كلمات اللغة، والثاني يستوجب تمثيل المعلومات الصرفية والنحوية (مثل: قسم الكلم، التعدية واللزوم، أدوات التعدية،...) بالنسبة لكل مدخل. وتجدر الإشارة هنا إلى أن هذه المعاجم التي صممت للآلة يصعب استغلالها من طرف الإنسان. ولبلوغ هذه الغاية وجب تطوير قواعد البيانات لتحويل المعلومات المشفرة التي تحتويها إلى نصوص وإضافة تفسيرات لها ليتمكن الإنسان من قراءتها وفهمها، وكذلك تصميم برامج بينية متطورة للبحث عن المعلومات المتوافرة بالقاعدة وعرضها على الشاشة بطريقة مستساغة واستعمال وسائط متعددة (نص، صورة، وصوت فيديو وصور متحركة). التطور الذي حصل على مستوى قواعد البيانات والبرمجيات المصاحبة، بالإضافة إلى تطور الأجهزة الحاسوبية، قد ساعدا في بداية الثمانينات على ظهور أول المعاجم الإلكترونية المتاحة إلى الجمهور العريض على سطح المكتب أو على الإنترنت (سواء عن طريق الاشتراك أو مفتوحة المصدر) أو على الأقراص المدمجة (CD - ROM) والأقراص الأخرى. ومن بين القواميس الإلكترونية الغربية العديدة والمتنوعة، نذكر هنا بعض الأمثلة للغة الإنكليزية والفرنسية: ذخيرة اللغة الفرنسية المحوسبة (TLFi)، وقاموس الأكاديمية الفرنسية (2 طبعات)، و"البيبليورم لاروس" الفرنسي ومعجم "أوكسفورد" الإنجليزي OED (الطبعة الثانية) والمعجم والمكنز الإلكتروني "كولينز". أما بالنسبة إلى اللغة العربية فثمة تقصير واضح في إعداد معاجم حاسوبية تلبي حاجيات المستخدم العربي بالرغم من إتاحة أمهات المعاجم العربية على أقراص المدمجة أو على الإنترنت. فهذه الخطوة ليست كافية؛ لأن هذه المعاجم تفتقر لأبسط مقومات المعاجم الحاسوبية بالمعنى الحديث للكلمة. إنها مجرد "نسخ مرقمنة" لتلك المعاجم الورقية (في صيغة "Doc" أو "HTML") لا يمكن الاستفادة منها بالشكل المطلوب (Ait Taleb, 2005). بالإضافة إلى ذلك، فالأدوات المصاحبة للبحث عن المعلومة بسيطة كما يتضح في المشاريع المعروفة مثل "عجيب" (لنظام صخر) "وكلمات" و"لسان العرب" / "القاموس المحيط". هذه السلبيات تعود أساسا إلى ضعف الهيكلة الحاسوبية التي بنيت عليها مداخل هذه المعاجم.

(monocritères)[1]. ولهذا السبب، وبفضل ما وفرته الهندسة الالكترونية، أصبحت المعاجم الحالية أكثر غنىً من ذي قبل، حيث المعلومات، وأكثر تشبعاً لإمكانية الانفتاح على النص المترابط (hypertexte) أثناء البحث الذي يعتمد معايير متعددة (multicritere) تستجيب لرغبات المستخدم. مما يمنح إمكانية الابحار الكبيرة والمستفيضة[2] (بنقرة واحدة فقط)، عن طريق روابط مشفّرة داخل نظام القاموس والموضوعة بشكل محسوب ومقرون دلاليا، لذلك فإن قاعدة بياناتها اللغوية تكون مشفرة أيضا، لتشمل جميع المستويات اللسانية كالأصوات والصرف والتركيب، إضافة إلى بناء معاجم إلكترونية للدلالة على الأقل في مستواها الصوري الذي يحدد العلاقات المنطقية[3] بين مختلف مكونات المتواليات اللسانية المقبولة في وجهيها الحقيقي والمجازي. كما تقوم تقنية التخزين الاستقرائي كما تقوم على بناء معاجم الكترونية للغات الطبيعية وفق الاستعمال العادي لكل لغة، مع ابراز خصائصها الصورية في جميع المستويات اللسانية التي صنفتها النظريات اللسانيات .

هذا وتعتبر المعاجم الالكترونية اللبنة الأولى والترجمة الآلية المسعفة بالحاسوب، التي تعتمد كثيراً في جميع المراحل والعمليات على تدخل بشكل كبير في ترجمة تلك التراكيب، كما أن هذا النوع من الترجمة يكون غير قادرٍ على ترجمة المعاني المعجمية للمفردات التي تكتسبها في السياق، فهي تقتصر في الغالب على ترجمة المعاني المعجمية للمفردات خارج سياقها التركيبي. كل هذه التطبيقات تعتمد المعاجم الالكترونية التي تحتاج الخبرة نفسها في التصور والبناء. فلا يمكن مثلا أن نتصور وجود مدقق إملائي لا يقوم على الخوارزميات اللسانية في المستويين الصوتي والصرفي. كما لا يمكن أن نتصور مترجما آليا لا يقوم على خوارزميات التعرّف على المتواليات اللسانية بسيطها ومركبها.

1- تبنى القواميس التقليدية لغايات اقتصادية بأحجام محدودة وهي عادة ما توضع مداخلها للتحليل، كما أن الولوج إليها يعتمد منهجية معينة في ترتيب المداخل مثل ما هو معهود : كالترتيب الألفبائي والصوتي والأبجدي، علاوة على محدودية الأجوبة التي توفرها أثناء البحث. لأن نصوصها المعجمية عادة ما تكون عبارة عن مسارد محصورة في زمن وضعها. لكن كل هذا كان عاديا جدا، لأن الحوامل الوحيدة للتخزين التي كانت متوفرة هي الورق فقط. أما في الوقت الحالي فقد وفرت المعوميات العناء في تخزين المعلومات أو حتى استرجاعها بأساليب خلفية وأمامية سهلة المنال على المستعملين بمختلف مستوياتهم ومشاربهم الثقافية والعلمية حيث أصبحت مداخل القواميس الالكترونية تتوفر على شروحات مصتفيضة تدخل في الحسبان جميع الأغراض الطبيعية مثل الترجمة والمعلومات الدلالية(تعريف المداخل والقواعد النحوية والمعارف العامة الموسوعية المرتبطة بتلك المداخل.

2- للمزيد من التفصيل ينظر: Michael Zock – John Carroll, les dictionnaires électroniques revue TAL n 44, 2003 p

3- لهذه العلاقات دور في تصنيف وترتيب المفاهيم والتصورات والدلالات والمعاني التي تجمع فئة دون غيرها في النص المعجمي لذلك تأتي المعاجم الالكترونية بأشكالها المختلفة في الترتيب المدخلي والموضوعاتي والعرض بالمجالات والتخصصات وتردد المواد المعجمية والوضع في عين الاعتبار المستوى اللغوي المطلوب.....(نفس المرجع والصفحة).

إن البحث في المعجم يقودنا إلى فهم القواعد التي تتحكم في تأليف المفردات، والحصول على معرفة واضحة بخصوص المظاهر التركيبية لها. ولقد ظهر مفهوم المعجم الإلكتروني في النصف الثاني من القرن العشرين[1]، مما أتاح الفرصة لتوفر مفرداته على الحاسب، وعلى الرقائق الإلكترونية، بحيث أصبح في الإمكان الحصول على أية مفردة بكل يسر، بل وتوظيفها في تطبيقات أكثر تعقيداً كمعالجة النصوص والتدقيق الإملائي، والترجمة الآلية، وغيرها، ومن ثمة ظهرت معاجم عامة، وأخرى متخصصة في الطب واللسانيات والهندسة والفلك.

نظرياً، يمكن النظر إلى المعجم الإلكتروني من زاويتين: زاوية اللغويات المعلوماتية، وزاوية المعجمية أو المعجماتية التي تهتم بكيفية صناعة المعاجم الورقية، وتحديد مبادئها وآلياتها.

يشترك المعجم الإلكتروني مع المعجم الورقي في مستوى المقولة، والوصف المورفوتركيبي للمقولة المعجمية. ومن وجهة نظر معلوماتية ولسانية، يشكل المعجم الإلكتروني مادة خام لتقديم وصف شامل للمفردات صرفياً ونحوياً. وباختصار، إن المعجم الإلكتروني عبارة عن قاعدة معطيات مشفرة.

يعتبرالمعجم نموذجاً للكفاية اللغوية لدى المتكلم (في دماغه البشري)[2]. وهو عبارة عن منظومة من الوظائف المسجلة في الكفاية المعجمية على شكل متواليات لغوية يصطلح عليها في علم المعجم بالمداخل المعجمية[3]، إنه لائحة طويلة من المفردات مقترنة بمعطيات، ومرتبة ترتيباً خاصاً، وتعدى هذه اللائحة من المفردات بالمدونة أو المتن[4].

1- محمد منصور الغامدي وآخرون، الترجمة عبر الشبكة العالمية: نظام مقترح للترجمة من وإلى العربية، مجلة جامعة الملك سعود، "اللغات والترجمة"، 2006، ص. 8.

2- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، إشراف الأستاذ الدكتور: عبد الغني أبو العزم، جامعة الحسن الثاني، الدار البيضاء، 2009/2008، ص 89.

3- للمزيد من بسط القول ينظر : جواد حسني سماعنه، المعجم العلمي المختص، (المنهج والمصطلح)، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 48، الرباط، المغرب، ديسمبر1420هـ/1999م، ص 42، 43.

4- المرجع نفسه، الموضع نفسه.

والمعجم عبارة عن زوج: المفتاحية[1] والمعطيات[2]، كما يمكن أن نقترح تعريفات وتحديدات[3] أخرى بناء على مناهج وطرق التصديق[4].

إن المعجم الإلكتروني كما يعرفه د. محمد الحناش "قاعدة البيانات اللغوية المشفرة، بل تشمل جميع المستويات اللسانية كالأصوات والصرف والتركيب، بالإضافة إلى بناء معاجم إلكترونية للدلالة على الأقل في مستواها الصوري الذي يحدد العلاقة المنطقية بين مختلف مكونات المتواليات اللسانية المقبولة في وجهيها الحقيقي والمجازي"[5]. فهو عبارة عن آلة إلكترونية متخصصة في الربط بين مستويات المعجم بما توفره تكنولوجيا الاسترجاع الخلفي والأمامي من إمكانية للاستعمال والاستغلال[6]. كما تقوم تقنية التخزين الاستقرائي على بناء معاجم إلكترونية للغات الطبيعية وفق الاستعمال العادي لكل لغة، مع إبراز خصائصها الصورية في جميع المستويات اللسانية التي صنفتها اللسانيات العامة.

وبهذا يكون المعجم الإلكتروني عبارة عن آلة إلكترونية فيها جميع المعلومات القاموسية لكل لغة اختيرت، فهي متعددة الوسائط توظف آخر المبتكرات التكنولوجيا للصورة والصوت، وقد ظهرت إحدى القواميس الإلكترونية في شكل حاسوب جيبي أو آلة رقمية محمولة صغيرة

1- يقصد بالمفتاحية: المفردة في المعجم.

2- يقصد بالمعطيات: المعلومات الصرفية والنحوية المقدمة عن كل مفردة، وقد تكون مسننة أي مشفرة، أو عبارة عن قاعدة، كما هو الأمر في المعجم الاستعمالي.للمزيد من التفصيل : ينظر: صافية زفنكي،التطورات المعجمية والمعجمات اللغوية العامة العربية الحديثة، منشورات وزارة الثقافة، دمشق، د.ط،2007، 197 و ما بعدها.

3- يمثل المتغير ت حجم القاموس أو المعجم على شكل نص مكون من حروف تسمى بـــ Les octets، أو مفردات الذاكرة (Mots-mémoire)، كما قد يمثل على شكل مفاتيح أو معلومات في المعجم أو مفاتيح فقط.

يمثل الرابط T رابط الضغط القائم بين حجم المعجم "ت" في صورة نص والحجم التمثيل 'T الممثل لها بالمعادلة التالية: [;T'T] F = R

ولكي يتحقق هذا الضغط، لابد أن يكون الرابط أقل من واحد. ويمثل المتغير n عدد المداخل المعجمية في القاموس أو المعجم، ويرمز إلى وظيفة الخوارزميات بـــ Log bc التي تستخدم باعتبارها وظيفة لضبط تعقيدات الخوارزميات، ومن أجل ذلك نوظف الترميز الآتي: o(f(n))

إن الدالة g(n) تنتمي إلى التعقيدo(f(n)) شريطة:

إذا وجدت c توجد No، بشرط أن تكون n >Nog(n)←c.f(n)

حيث إن: c وN ثابتين/ ساكنين.

إن المتغير (n) لابد أن يكون منسجماً مع الخوارزميات (المعطيات)، ذلك أن كتابة/ تحرير المعجم على شكل نص عملية خطية تتوقف على حجم المعجم الذي يتخذ التركيبة التالية: o(I).

وبمجرد تطبيق هذا الإجراء، إن العملية الأساس لن تكون خطية. وفيما يتعلق بوحدة التخزين نستعمل الميغا أوكتي المرموز لها بـــ: Mo.

4 – Les méthodes d'accès.

5 – انظر محمد الحناش، اللغة العربية والحاسوب، أكتوبر 2002. ص .12.

6 – H.C. Domine, Téchnique de l' intelligence artificielle , un guide structuré DUNODinformatique. 1988. P.51

الحجم يطلق عليها " المساعد الرقمي الشخص(PDA)[1]، نظرا لاستعمالاتها المتعددة في استرجاع المعلومات وتخزينها.

لقد تأثرت صناعة المعاجم الآلية كثيرا بنظرية الأوتومات التي أبانت عن جدارتها ونجاعتها في التعرف على جميع الأشكال والمتواليات اللسانية التي تثير الغموض واللبس أثناء معالجتها، كما أن الانتشار الواسع والاستخدامات المتعددة للمعاجم الآلية في الشبكة (الانترنت) في معالجة جميع مستويات المعلومات والمعارف الرمزية قد أدى إلى بناء معاجم آلية معتمدة على تخزين وفهرسة الوثائق واستخراج المعلومات مبنية بلغة التطوير (XML) أو(JAVA) مثلا أو غيرها، هذه التقنية المتطورة استطاعت أن تمنح الأدوات اللسانية القدرة على التمثيل المعياري للنصوص اللغوية.

فالمعجم الالكتروني : هو نتاج تطبيق علم الالكترونيات، وعلوم الحاسوب في مجال الصناعة المعجمية، ويعرفه أهل الاختصاص بأنه مخزون من المفردات اللغوية المرفوقة بمعلومات عنها، ككيفية النطق بها وأصلها واستعمالها ومعانيها وعلاقاتها بغيرها، محفوظ في نظام معين في ذاكرة ذات سعة تخزين كبيرة، ويقوم برنامج آلي بإدارة هذه المعطيات وتدبيرها وفق نموذج محدد سلفا[2]. "فهو عبارة عن قاعدة معطيات لغوية مشفرة تشمل جميع المستويات اللسانية: الأصوات والصرف والتركيب".

يحتوي هذا المعجم على قواعد معطيات وجداول ويتضمن قوانين تمكنه من عرض جميع المعارف المعجمية بسهولة ويسر، ومن إجراء عمليات بحيث متنوعة، فهو بذلك يبيّن حاجة المعلّمين والمتعلمين والمختصين وغير المختصين على حدّ سواء[3]. فالمعاجم الإلكترونية جلبت طريقة جديدة للاستخدام من حيث طريقة البحث عن الكلمات مع تقديم تحسينات جديدة من حيث

1– هذا اختصار لما بالمساعد الرقمي الشخصي، وقد شرع في تداول هذا المصطلح بتاريخ 7 يناير 1992 خلال معرض دولي في لاس فيكاس بالولايات المتحدة الأمريكية حينما قدم حون سكا لي قاموسه الإلكتروني الصغيرة الحجم سماه "نيوتون" كانت ابتكرته شركة شارب، تتميز هذه الآلات بمظهرها الجميل وخفتها وكذا سعة ذاكرتها القوية التي تستغل في العديد من العمليات منها استرجاع المعلومات والبحث عنها وعن التي استعصى فهمها، علاوة على تخزين الصورة والموسيقى باستخدام تقنية قواعد المعارف والإضافة عليها، وفي هذا الإطار تعتبر اللغة اليابانية رائدة في هذا المجال بحيث تم ابتكار آلة قاموسية ثنائية اللغة (ياباني– انجليزي) من وضع "كازيو" منذ سنة 1981، مجهزة بأحدث المكونات الرقمية والميكرو تقنية بذاكرة كبيرة الحجم إذ يحتوي على أزيد من 2020 كلمة وتعبير متلازم ومسكوك يتميز بتميز بمرونة عالية في اختيار اللغة للمستعمل ووفق أزرار ومفتاح لغوي متطور.

2- Ann Moseley, Interaction, Fourth edition and American Heritage Dictionary College, edition Houghton Mifflin company, 4ème édition, Dictionary of the English; 2003.

ينظر مقال : المعاجم العربية الإلكترونية وآفاق تطورها، للدكتور عز الدين البوشيخي، قدّم في المؤتمر الدولي الرابع في اللغة والترجمة، وموضوعه "الصناعة المعجمية: الواقع والتطلعات"، تنظيم مركز أطلس العالمي للدراسات والأبحاث وجامعة الشارقة، الشارقة، 20 و21 أبريل 2004.

3– مروان البواب، "نحو معجم حاسوبي للغة العربية". عنوان الموقع : في معجم حاسوب للغة العربية. موقع الويب:www.alesco.org.tn

الأداء وسهولة البحث في مختلف قاعدة المعطيات الموجودة في المعجم[1]، وجملة القول أنّ: المعجم الإلكتروني هو الخاضع للبرمجة الحاسوبية في أدقّ تفاصيلها[2]. فهو معجم مبني وفق ضوابط الصناعة المعجمية الحديثة، يوفر المادة اللغوية العربية من مفردات عامة ومصطلحات علمية[3]، فهو "معجم آلي غير ورقي" تسمح آليته بإمكانات التخزين الكبيرة، والإغناء المستمر.

يدمج المعجم في صورته الآلية مع كثير من النظم اللغوية الآلية، مثل المعالجات الصرفية والنحوية والدلالية، ونظم الفهم الأتوماتي ونظم التخاطب مع قواعد البيانات ونظم الخبيرة، والترجمة الآلية، واسترجاع المعلومات، والفهرسة الآلية، والمصطلحات الإملائية والنحوية، وبرامج تعليم اللغات[4].

ملاحظة: يمكن أن يكون المعجم الإلكتروني معجم متعدّد اللغات، يورد اللفظ أو المصطلح باللغات الإنجليزية والفرنسية والعربية، وقد يتسع لغيرها من اللغات، وترجمة اللفظ أو المصطلح إلى إعداد من اللغات الحيّة، كون إمكانات البحث المتنوعة بتنوع الغرض منها انطلاقا من اللفظ أو المصطلح أو من المجال العلمي أو الموضوعي وباللغة الّتي يختارها المستعمل[5].

2 - أهمية المعجم بالنسبة إلى اللغة العربية:

2-1 - أهمية المعجم الإلكتروني للمستخدم العربي:

المعجم الإلكتروني ضروري للفرد العربي في مجتمع المعلومات والاقتصاد القائم على المعرفة. فقد تطورت وضيفة المعجم السنوات الأخيرة ليصبح وسيلة عمل بالنسبة للطالب وللأستاذ وللباحث وللمترجم ولعدة فئات اجتماعية ومهنية أخرى.

وتتضح هذه الضرورة أكثر إذا تمعنا في واقع صناعةُ المعاجم في اللغات العالمية، وقارناه بواقع صناعة معاجمنا العربية، خاصة الإلكترونية منها.

1- Zaafrani Riadh, "un dictionnaire électronique pour apprenant de l'arabe (langue second) basé sur corpus, JEP-Taln 2004, Traitement automatique de l'arabe, fès, 20 Avril 2004.

وينظر : د عبد الغني أبو العزم وآخرون، "المعجم الإلكتروني العربي : التصور والمنهجية.

2– أ.د عبد الغني أبو العزم، "التصور المنهجي للمعجم العربي الحديث في أفق تحويله إلى معجم حاسوبي" موقع الويب : اجتماع خبراء المعجم الحاسوبي للغة العربية ALECSO، www.alesco.org.tn

3 – إدريس القاسمي، "المعجم الآلي الشامل"، الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية، المنظمة العربية للتربية والثقافة والعلوم، مدينة الملك عبد العزيز للعلوم والتقنية، 1429/5/2-4/29هـ (2008/5/7-5/5)، ص 1.

عنوان الموقع : الأوراق المقدمة في ورشة عمل المعجم الحاسوبي العربي، مقهى علوم اللغة العربية وآدابها.

موقع الويب : –http://www.almujam.org/DOC/Qasimi.pdf

4 – نبيل علي، اللغة العربية والحاسوب، سابق، ص 492.

5 – إدريس القاسمي، المعجم الآلي الشامل، سابق.

وتكمن أهميته بالنسبة إلى تعليم اللغة في إقبال التلاميذ أو الطلاب على النسخة الإلكترونية لما توفره من إغراء على مستوى البحث عن المعلومة والعرض باستعمال وسائل متعددة الوسائط.

التحديث السريع يجعل من المعجم الإلكتروني مواكبا لتطور اللغة العربية على جميع مستوياتها.

يمكن أن نضيف أخيرا أن المعجم الإلكتروني يسهل عملية بناء أنتولوجيا للغة العربية[1].

2-2 ـ أهمية المعجم للمعالجة الآلية للغة العربية:

يعتبر المعجم الإلكتروني أداة ضرورية وأساسية لتطبيقات المعالجة الآلية للغات الطبيعية[2]، فجودة النتائج التي تفرزها هذه التطبيقات مرتبطة كليا بجودة المعجم من حيث عدد مداخله وشمولية المعلومات التي يحتويها كل مدخل[3]. يتضح هذا الترابط إذا تمعنا مثلا في عمل المدقق الإملائي؛ فنجد أنه يعتبر الكلمات الصحيحة خاطئة إن لم يجدها ممثلة في المعجم الذي يعتمد عليه في عملية التدقيق. كذلك عند التصحيح لا يقترح كلمات غير موجودة بالمعجم من بين المرشحة للتصحيح.

ومن هنا نستخلص أن تطوير تطبيقات حاسوبية للغة العربية تلبي حاجة المستخدم العربي وتجعل اللغة العربية مواكبة لتطور المجتمع رهين وجود معاجم إلكترونية عربية مناسبة وذات جودة عالية وتخضع إلى مقاييس عالمية على مستوى المحتوى والهيكلة. وعدم توافر هذه المعاجم بالجودة المطلوبة له انعكاس سلبي على استعمال اللغة العربية في الوطن العربي وانتشارها عالميا، لأنه أصبح من المؤكد أن اللغة التي لا يقع الاهتمام بها من حيث الحوسبة يتناقص دورها تدريجيا كلغة ناقلة، وتصبح عرضة للزوال. والأمثلة تكاثرت في العشريات الأخيرة التي شاع فيها استعمال التكنولوجيات الحديثة بين الناس[4].

---

1 − Baccar F., Gargouri B., Ben Hamadou A. "Methodology for Generating Core Domain Ontologies from LMF Standardized Dictionaries ". In International Journal of Applied Ontology. Mark Musen and Nicola Guarino (Eds.), IOS Press, Amsterdam, 2011.

2− تعتبر المعاجم الالكترونية أدوات أساسية بالنسبة للتحليل والتوليد الآليين في اللغات الطبيعية، وهي أيضاً عنصر أساس لمعالجة اللغة البشرية، واللغة المستخدمة في الحاسوب. لذا ينبغي أخذهما بعين الحسبان، بالتركيز على اللغة البشرية على مستوى الاكتساب، وعلى اللغة الاصطناعية على مستوى البناء، فبناء معاجم آلية يتوخى خدمة أغراض المدقق الإملائي والمترجم الآلي، وغيرهما من مجالات وتطبيقات المعالجة الآلية للغات الطبيعية (NLP).

3− إن أهمية المعجم تكمن في كمية المعلومات التي يتوفر عليها، والأدوات والتقنيات التي يعتمد عليها في الكشف عن معلومات المداخل المعجمية. إن استراتيجيات تفوق المعجم إذن تحدد من ناحية في المعارف الضرورية التي من المسلم أن يقدمها عن كل مدخل (مفتاح معجمي)، ومن ناحية أخرى، في طريقة ومنهجية المعالجة والمدة الزمنية التي يستغرقها التفتيش عن المعلومة. فعلى سبيل المثال لا الحصر، عادة ما يبدأ البحث بتقديم معنى المفردات على مستوى التحليل، ثم يأتي بعد ذلك دور التأليف أو التركيب، حيث يتم ضبط مفاهيم وتصورات المفردات.

4− عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه،الموسم الثقافي التاسع والعشرين لمجمع اللغة العربية الأردني "اللغة العربية ووحدة الأمة"مقال منشور بتاريخ الأربعاء 27 ذو الحجة 1432ه/الموافق لـــ 23تشرين الثاني 211 م، ص292.

والمسؤولية هنا مشتركة بين اللغويين (المعجميين) والمعلوماتيين المتخصصين في اللغويات الحاسوبية، وكذلك صناع القرار لتوفير الدعم المادي والأدبي، لأن بناء وتطوير المعاجم الإلكترونية عمل متعدد الاختصاصات ويتطلب اعتمادات تفوق غالبا إمكانيات البلد الواحد[1]. وفي هذا الإطار يمكن ذكر بعض المشاريع المعروفة بالنسبة للغات الأجنبية مثل مشروع EUREKA الممول من طرف إيطاليا وإسبانيا لتطوير معجم مقيس، وكذلك المشاريع العديدة الممولة من طرف الإتحاد الأوروبي في مجال تطوير موارد معجمية متعددة اللغات والترجمة الآلية مثل مشروع EUROTRA[2].

3- مزايا المعجم الإلكتروني :

للمعجم الإلكتروني مزايا عديدة متأتية أساسا من التطور التكنولوجي الذي حصل السنوات الأخيرة على مستوى سرعة معالجة البيانات وسعة تخزين المعلومات وكذلك على مستوى البرمجة ومعالجة قواعد البيانات. هذه المزايا يمكن حصرها في النقاط التالية[3]:

أ/ تنوع طرق البحث عن المعلومة: يمكن لمستخدم المعجم الإلكتروني أن يصل إلى المعلومة عبر الجذر أو الجذع (البحث البسيط) أو عبر المعنى (البحث المتقدم). مثلا يمكن البحث عن كلمة "هضبة" باستعمال المعنى الآتي "أرض مرتفعة". كما يمكن البحث عبر الإبحار داخل المعجم باستعمال (الروابط النصية)[4].

ب/ طاقة التخزين الواسعة وتطور تقنيات قواعد المعطيات تتيح بناء معاجم كبيرة الحجم تجمع بين القديم والمعاصر ومتعددة اللغات والوسائط. هذه المعاجم تمتاز بالدقة والشمولية من حيث أنها توفر لكل كلمة معانيها الأساسية والفرعية وتعطي لذلك أمثلة وشواهد متنوعة[5].

ت/ إمكانية التوليد الآلي[6] لبعض الكلمات القياسات بدون الحاجة إلى تمثيلها بالمعجم وذلك بالاعتماد على قواعد الاشتقاق. هذه الميزة يصعب (أو يستحيل) توفيرها بالنسبة إلى المعجم الورقي لأن إيراد المشتقات القياسية لجميع الأفعال الممثلة سيضاعف حجم المعجم ويجعله غير قابل للاستعمال[7].

1- اللغة والحاسوب، نبيل علي، ص 568،569.

2- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص294.

3- المرجع نفسه، ص 291.

4 -les liens de texte.

5- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص291.

6- من مزايا المعجم الحاسوبي يمكن إعتماده على التويد الآلي من خلال التطبيقات الحاسوبية التي تبنى عليها قاعدة المعطيات، فيمكنه أن يعتمد التوليد أكثر من التخزين، فهو يمتاز بالسهولة في تدبير المعطيات اللغوية والقواعد. للمزيد من بسط القول : ينظر : التوليد والنسقية والترجمة الآلية، إشراف عبد القادر الفاسي الفهري، إعداد أحمد بريسول، خالد الأشهب، المجلد1،منشورات معهد الدراسات والأبحاث للتعريب، الرباط ،2001.

7- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص292.

ث/ احتواء المعجم الإلكتروني على عدة تطبيقات لغوية مهمة يمكن للمستخدم أن يستفيد منها مثل: تصريف الأفعال والأسماء، البحث عن المترادفات، المعالجة على المستوى الصوتي لتحويل المكتوب إلى منطوق، التدقيق الإملائي لتصويب الكلمات المدخلة،...هذه الخدمات غير متوفرة في المعجم الورقي.

جـــ/سهولة تعديل المعجم الإلكتروني بإضافة مداخل جديدة أو بتحيين مداخل موجودة. هذه الميزة يجب التعامل معها بكل حذر للحفاظ على مصداقية المعجم وجودته. تعديل محتوى المعجم يجب أن يتم من طرف معجميين وفق تراتيب مقننة تضمن تماسك المحتوى.

د/الاعتماد على الوسائلِ الحاسوبيّةِ الحديثة المتعددةِ الوسائط[1] من نصوص، وأصوات، وصور ثابتة ومتحركة، وأفلام الفيديو لعرض المعارف. هذه الخاصية لها تأثير إيجابي على استساغة وفهم المعارف المعروضة.

وتجدر الإشارة هنا إلى أن للمعجم الورقي والمعجم الإلكتروني مزايا مشتركة لكن بدرجات متفاوتة نذكر بالخصوص إمكانية الاستعمال في كل مكان. مع تطور الحواسيب المحمولة على مستوى الحجم والاستقلالية عن التزود بالطاقة الكهربائية أصبح استغلال المعجم الإلكتروني متاح في كل مكان لكن بدرجة أقل من المعجم الورقي[2].

4- طرق بناء المعجم الإلكتروني:

بغض النظر عن الطريقة المعتمدة، فإن بناء معجم إلكتروني ليس بالعمل الهين، فهو يتطلب مجهودا جبارا يقوم به فريق يتكون من معجميين ومعلوماتيين. يهتم المعجميون بتجميع المادة اللغوية من مدونات ومعاجم ورقية وانتقاء المداخل المعجمية[3] وتحديد المعلومات الملحقة بكل مدخل.

ويهتم المعلوماتيون أو الحاسوبيون بالتصميم والهيكلة لضبط (البنية الفوقية)[4] التي تعنى بترتيب مداخل المعجم و(البنية المصغرة)[5] التي تهتم بترتيب مكونات المدخل الواحد. كذلك يهتم

---

1 – Multimedia.

2– وبالرغم من ما ذكر من مزايا ما زال المعجم الورقي ينعم باستحسان القراء والمستعملين لكن المستقبل سيكون حتما لفائدة المعجم الإلكتروني.

3– إن تنظيم الذاكرة المعجمية ،المشكلة للمعجم الذهني الذي يكونه متكلم اللغة الطبيعية ،لابد وأن يراعي مجموعة من المداخل منها :

*الاهتمام بالخصائص النطقية للكلمات كالنبر والروم والتنغيم أو ما يسمى بالظواهر فوق مقطعية .

*الاهتمام بالسمات الصوتية للكلمات وهنا لا بد من معرفة السمات الصوتية للصوامت والصوامت ؛كالقلقلة ،والجهر والهمس والتوسط .

*الاهتمام بالمعنى ؛أي الحد أو مفهوم المفردة وعلاقتها بالمفردات الأخرى كالترادف أو التطابق ...

*الاهتمام بالمعلومات التركيبية :الصنف المقولي والإطار التفريعي، الخصائص الاعرابية،والمصفاة الاعرابية ،قيود التوارد ،الشبكة المحورية هذه إذن بعض خصائص المداخل المعجمية ،الضرورية لبناء قاموس عربي محوسب، ويمكن البحث في هذه المداخل المعجمية من معرفة خصائص المفردات العربية صوتيا وتركيبيا ،ودلاليا وصرفيا ،بالتالي تحاشي تكرار الكثير من المعلومات المتداخلة والمتحاقلة بين بحث المحلل النحوي والمحلل المعجمي، وقد أشار إليها الفاسي الفهري في مؤلفاته ، للمزيد من التفصيل ينظر: جواد حسني سماعنه، المعجم العلمي المختص، (المنهج والمصطلح)، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 48، الرباط، المغرب، ديسمبر 1420هـ/1999م، ص 42، 43.

4 – Macro-structure .

5 – Micro-structure.

الحاسوبيون بتصميم البرامج الضرورية لإدخال المعلومات المعجمية وتحيينها والبحث عنها وعرضها ولتوفير مختلف الخدمات الملحقة بالمعجم من تدقيق إملائي وتصريف الأسماء والأفعال وغيرها. ولبلوغ درجة عالية من الدقة والجودة يجب أن يعمل أعضاء هذا الفريق بطريقة متعاونة ومتكاملة وكذلك متوازية ربحا للوقت.

وفي ما يلي أهم طرق بناء المعاجم الإلكترونية[1]:

أ/ الطريقة اليدوية :

تعتبر الطريقة اليدوية من أسهل طرق بناء المعاجم لكنها مكلفة من حيث الجهد والوقت والمال.

يتم بناء المعجم باتباع المراحل التالية[2]:

- الاتفاق على هيكل المعجم (ترتيب المداخل ومحتوى كل مدخل) انطلاقا مما يسمى بكراس الشروط الذي يحدد الهدف من بناء المعجم، المستخدم المستهدف (مبتدئ، خبير،...) طريقة استغلال المعجم (على الشابكة، على سطح المكتب أو على قرص مضغوط،..).
- تحديد المصادر اللغوية (مدونات، معاجم ورقية،...) التي سيقع اعتمادها كمادة أولية.
- انتقاء وتجميع المادة المعجمية من المصادر المذكورة في جذاذات ورقية.
- تصميم وبناء قاعدة البيانات التي ستحتوي المعجم.
- بناء برمجية لإدخال المعلومات المعجمية يدويا من الجذاذات الورقية.
- إدخال المعلومات يدويا من الجذاذات الورقية.
- تجريب المعجم للتأكد من تحقيقه كافة الوظائف المطلوبة منه.
- تصويب الأخطاء اللغوية والبرمجية.

المرحلة الأولى يقوم بها المعجميون بتنسيق مع الحاسوبيين **المرحلة الثانية والثالثة** يقوم بها اللغويون مع إمكانية الاستعانة بالحاسوبيين لمدهم ببعض الأدوات التي تساعدهم على تحليل وجرد المدونات أو القيام ببعض العمليات الإحصائية.

المرحلة الرابعة والخامسة يقوم بها الحاسوبيون.

المرحلة السادسة يقوم بها المدونون.

المرحلة السابعة يقوم بها لغويون والمرحلة الأخيرة يقوم بها المعجميون و الحاسوبيون.

1- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص295.

2- المرجع نفسه ، الموضع نفسه.

ب/ انطلاقا من معجم ورقي محوسب:

هذه الطريقة ممتازة جدا لأنها تقتصر المسافات وتقتصد المال والجهد. الفكرة المركزية هي تحويل معجم ورقي مرقمن إلى معجم إلكتروني يخضع إلى مواصفات دقيقة وحديثة. هذا التحويل يتم عن طريق برنامج حاسوبي يقع تصميمه للغرض. النتيجة المتحصل عليها قاعدة بيانات معجمية. ومن مزايا هذه الطريقة إمكانية الاعتماد على أكثر من معجم ثم يقع دمج قواعد البيانات المتحصل عليها للحصول على قاعدة ثرية على مستوى عدد المداخل وكذلك على مستوى محتوى المداخل. وقد قام فريق من مخبر ميراكل بجامعة صفاقس باقتراح وتجربة هذه الطريقة[1] على معجم "الغني"[2]، وكذلك على معجم الوسيط بالتعاون مع مدينة الملك عبد العزيز للعلوم والتقنية.

❖ **مراحل بناء قاعدة البيانات المعجمية[3]:**

- اختيار المعجم (أو المعاجم) الذي سيعتمد عليه في صيغته المرقمنة.
- تصميم قاعدة البيانات المعجمية التي ستحتوي مداخل المعجم.
- تصميم برنامج تعبئة قاعدة البيانات المعجمية انطلاقا من النسخة أو النسخ المرقمنة.
- تجريب القاعدة.
- تصحيح الأخطاء اللغوية والحاسوبية.
- إثرائها يدويا إذا اقتضت الحاجة.

ج/انطلاقا من مدونة نصية:

هذه الطريقة تعتبر من أصعب الطرق وأقلها دقة. مزاياها تكمن في أنها غير مكلفة على المستوى البشرى والمادي وكذلك تمكن من الحصول على معجم مواكب.

---

1 – Khemakhem A., Elleuch I., Gargouri B., Ben Hamadou A., (2009). Towards an automatic conversion approach of editorial Arabic dictionaries into LMF- ISO 24613 standardized mode. MEDAR 2009, Cairo, Egypt.

2– "معجم الغني" لصاحبه عبد الغني أبو العزم والذي أصدرته شركة صخر ميديا بالقاهرة، يعتمد على منهجية تستجيب لآخر النظريات والابتكارات العلمية والتكنولوجية، والتي تتمثل في الأبعاد المعرفية اللاخطية للدماغ البشري في الإدراك والاكتساب. ولذلك فهي تهدف إلى تمكين المتلقي من اكتساب ألفاظ وتعابير اللغة العربية معرفياً. بهذه الصفة، هو الذي يدله على جذر الكلمة وليس جذر الكلمة الذي يفترض معرفةً مسبقةً للملتقى به ليتمكن من معرفة ما يرغب فيه ، وتلك هي الغايات الكبرى التي أصبحت تؤطر دوما النظرية المعجمية الحديثة، سنبسط القول في تجربة "معجم الغني" في الباب الثالث، الفصل الأول المخصص لتجارب تطبيقية للمعجم الإلكتروني : قراءة في بعض الدراسات الغربية و العربية.

3– عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص296.

بناء معجم بهذه الطريقة يتم باتباع المراحل التالية[1]:

- تحديد المدونة أو بناؤها.
- إنجاز برمجية لاستخراج المداخل والعلاقات بينها والمعلومات المعجمية الخاصة بكل مدخل بصفة آلية من المدونة.
- تدقيق المادة المعجمية المتحصل عليها من طرف المعجميين.
- إدخال المادة المعجمية المدققة في قاعدة البيانات التي ستحتوي المعجم.

د/ الطريقة التعاونية (ويكي):

المقاربة التعاونية (الويكي) تمكن من بناء معاجم إلكترونية بتكلفة زهيدة وذلك بتكاثف جهود المتطوعين المتخصصين في مجال المعجمية. هذه المقاربة برزت ونجحت مع ظهور الموسوعة "ويكيبيديا"[2]. فهي تسمح لكل متطوع يرى في نفسه القدرة على المساهمة البناءة بالعبور إلى المعجم وتغيير محتواه سواء بإضافة مدخل جديد أو بإثراء محتوى مدخل موجود[3].

ولتقليل الأخطاء يمكن لكل مستخدم أن يتعرف إلى التغييرات الحديثة التي أدخلت على المعجم والتثبت فيها مع إمكانية تصويب الخطأ.

***/ من أهم مزايا هذه الطريقة أنها[4]:**

– تضمن المواكبة السريعة للتطورات التي تحدث على المعجم من بروز مفردات جديدة وتوَسع مجال مفردات موجودة.

– تعتبر أن بناء المعرفة ليس مجرد مسألة شخصية. بل هو مقاربة جماعية من أجل الوصول إلى توافق في الآراء.

– لا تقصي أحداً من عملية بناء المعجم. فهي تضع على قدم المساواة المبتدئ (الطالب مثلا) والخبير (المعجمي).

1– المرجع السابق، ص296.

2– ويكيبيديا هي مشروع موسوعة رقمية، متعددة اللغات، حرة المحتوى. يستطيع أي شخص التحرير فيها بدون تسجيل، ويستطيع أي شخص الاستفادة من المحتوى، واستغلاله بهدف تجاري أو غيره وفقًا لترخيص الموسوعة .هناك 292 نسخة من ويكيبيديا بلغة مختلفة. وفي النسخة العربية، يساهم أكثر من 4,109 مساهمين نشطين في كتابة ما يزيد عن 426,133 مقالة، ويجري آلاف الزوار، من مختلف أنحاء العالم، الكثير من التعديلات، وينشئون الكثير من المقالات الجديدة يوميًا.

منذ أن أنشئت ويكيبيديا في عام 2001، نمت وتطورت بسرعة لتصبح واحدة من أكبر المواقع على الإنترنت، ولتجذب أكثر من 100 مليون زائر شهريًا، في أكتوبر 2014. بدأت النسخة العربية من الموسوعة الحرة في حزيران/يوليو 2003، وما زالت في مرحلة بناء المحتويات، لذا فإن أي تعديل أو إضافة، مهما بدت بسيطة، هي ذات قيمة كبيرة لهذه الموسوعة. عنوان الموقع : ويكيبيديا عن .

موقع الويب: ويكيبيدياعن https://ar.wikipedia.org/wiki/،آخر تعديل لهذه الصفحة كان يوم 21 أبريل 2016 الساعة 15:52.

3– عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص297.

4– المرجع، نفسه، ص297

بالنسبة للعربية يمكن ذكر مشروع "ويكاموس" الذي انطلق في شهر مايو 2004 والذي تقدم أشواطا مهمة لكنه لا يزال يشكو من ثغرات عديدة في المحتوى لأن بناءه لم يكن ممنهجا، بل موكول إلى إرادة المساهمين في عملية الإثراء[1].

5- تصنيف المعاجم الإلكترونية:

إن أبسط تصنيف للمعاجم الإلكترونية الذي يمكن أن يُستخلص مما ذكر، يُميز بين المعاجم الموجهة إلى الآلة ، والمعاجم الموجهة للإنسان، علما بأن الجهود قد اتجهت السنوات الأخيرة إلى دمج هذين الصنفين، وذلك باقتراح منهجية بناء معاجم مقيسة تراهن على تلبية حاجة الإنسان والآلة في نفس الوقت[2]. وفي ما تبقى سنهتم بالمعجم الإلكتروني الموجه للإنسان.

التصنيف الحديث للمعاجم الإلكترونية الموجه للإنسان يأخذ بعين الاعتبار الجوانب الأساسية التالية: ترتيب المداخل (حسب الجذور، حسب الجذوع،...)، وظيفة المعجم (للترجمة، للتعليم،...)، نوعية المحتوى (نصي/ متعدد الوسائط، لغوي/ متخصص،أحادي اللغة/ متعدد اللغات،....)، المحمل (سطح المكتب، صفحات الواب، أقراص مدمجة،..).

انطلاقا من هذه الجوانب يمكن تصنيف المعاجم الإلكترونية كما يلي[3]:

- **معاجم لغوية:** متكونة من عينة من المفردات. يحتوي كل مدخل على المعلومات اللغوية الأساسية التالية: تعريف الكلمة، خصائصها الصرفية والنحوية، طريقة الكتابة (الإملاء)، المعاني المختلفة مع أمثلة وشواهد لمختلف الاستعمالات.

- **معاجم متخصصة:** تحتوي على المفردات المستعملة لعلم ما أو فن: قاموس الرياضيات، قاموس الطب، قاموس الاقتصاد، قاموس الحاسوب، قاموس أسماء العلم.

- **معاجم متعددة اللغات:** تعطي ترجمة الكلمات إلى لغة أو لغات أجنبية.

- **معاجم بصرية:** تحتوي على مجموعة من الصور وأشرطة الفيديو مبوّبة حسب المواضيع التي تعالجها (بنايات، حيوانات، ألعاب، ملابس، وسائل نقل،....).

6- التمثيل الحاسوبي للبنية الأساسية للمعجم الإلكتروني:

يتكون المعجم من نواة وثلاث وحدات اختيارية (انظر الرسم البياني رقم1)[1]:

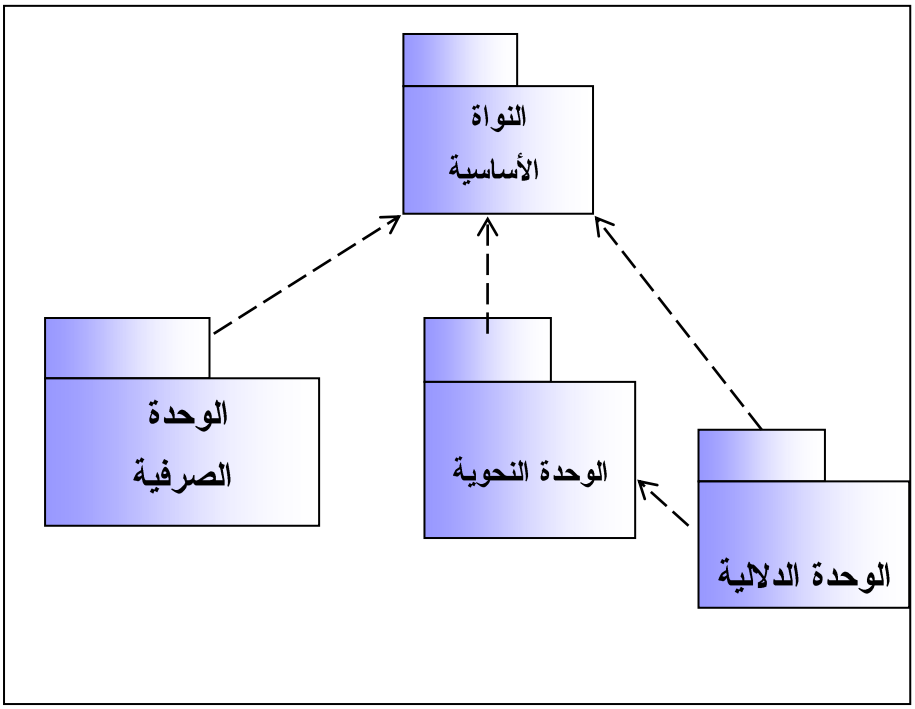
1- المرجع السابق، الموضع نفسه.

2 -Francopoulo G. & George M. (2008). ISO/TC 37/SC 4 N453 (N330 Rev.16). Language resource management- Lexical markup framework (LMF).

3- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه،الموسم الثقافي التاسع والعشرين لمجمع اللغة العربية الأردني "اللغة العربية ووحدة الأمة"مقال منشور بتاريخ الأربعاء 27 ذو الحجة 1432ه/الموافق لـ 23تشرين الثاني 211 م، ص292.

1- المرجع نفسه، ص 303.

- "**النواة الأساسية**" :تجمع المعلومات العامة للمدخل مثل أصل الكلمة، تعريفها، استعمالاتها (شواهد وأمثلة).
- وحدة صرفية تهتم بالمستوى الصرفي[1].
- وحدة نحوية تهتم بتمثيل السمات النحوية للكلمة[2]
- وحدة دلالية تهتم بالعلاقات الدلالية بين الكلمات.



الرسم البياني رقم1 : النواة الأساسية والوحدات المختصّة للمعجم الإلكتروني.

6-1- النواة الأساسية:

تتكوّن هذه النواة من ستة أصناف من المعلومات أهمها (انظر الرسم البياني رقم2)[3]:

❖ **المعلومات العامّة:** تحتوي على معلومات تخص نسخة المعجم مثل الجهة أو الجهات الممولة للمشروع، تاريخ الإصدار، رمز النسخة، المؤلف/ المؤلفون، العلامات الاصطلاحية، ملاحظات مختلفة،....

1- في هذا المستوى يتم توصيف المداخل المعجمية للكلمات المدخلة (الأفعال)، وتتضمن هذه المرحلة تحديد النوع والباب. ويتم في هذه العملية:(تحديد نوع الفعل وأبوابه التصريفية، تحديد مزيدات "مشتقات" الفعل المجرد بنوعيه، تحديد مصادر الفعل المجرد بأنواعه العادي، الميمي ، المدة والهيئة...،تحديد التغيرات التي تطرأ على الفعل كالإدغام والإبدال والحذف والقلب ..).ينظر: الدراسات المعجمية، مجلة تصدر عن الجمعية المغربية للدراسات المعجمية، العدد الثالث والرابع، مطبعة فضالة،المحمدية، المغرب،2005، ص 35.

2- في هذا المستوى يتم تحديد الخصائص النحوية للكلمات المدخلة، ينظر: الدراسات المعجمية، مجلة تصدر عن الجمعية المغربية للدراسات المعجمية، العدد الثالث والرابع، مطبعة فضالة،المحمدية، المغرب،2005، ص 35.

3- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص 304.

❖ **المدخل المعجمي:** يشمل معلومات خاصة بالمدخل الذي يمكن أن يكون (الجذع أو الجذر)[1] والتي تصنف كما يلي[2]:

- **النوع:** هذه المعلومة تمكن من تحديد نوعية المدخل إن كان جذرا[3] أم جذعا[4].
- **قسم الكلام:** اسم أو فعل أو حرف،... هذه المعلومة تخص الجذع.
- **الوزن الصرفي:** خاصّة بالأسماء المشتقّة والأفعال المشتقّة.
- **معدلات الاستخدام:** معلومة إحصائية حول مدى شيوع الكلمة (كثيرة الاستعمال/ متوسّطة الاستعمال/ قليلة الاستعمال)، و يمكن تحديد هذه المعلومة بالرجوع إلى المدوّنة.
- **التاريخ:** لتحديد الحقبة الزمنية التي ظهرت فيها الكلمة (قديمة أم حديثة،...).
- **أصل الكلمة:** لتحديد أصل الكلمة[5].

❖ **الجذع:** يمثل الكلمة في صيغتها المجردة (بدون تصريف)[6]. يحتوي هذا الصنف على المعلومات التالية[7]:

- طريقة كتابة الجذع.
- وطريقة نطقه (كتابة صوتية أو تسجيل صوتي).

❖ **المعنى:** يحتوي هذا الصنف على المعلومات التالية:

- المعرّف الوحيد .
- معدل استخدامه: كثير الاستعمال/ متوسّط الاستعمال/ قليل الاستعمال. تستغل هذه المعلومة الإحصائية لترتيب المعاني حسب شيوعها وشهرتها عند عرضها على الشاشة.

تتكرّر هذه المعلومات بحسب تعدد معاني المدخل. كل معنى مرتبط بتعريف أو بأكثر.

1- تشير بعض الدراسات الحديثة إلى أنّ هناك عدّة مصطلحات متقاربة المعنى، وتؤدي نفس الغرض، لكلمة "الجذر"هي كلمة الجذع، فالجذع: (Racine): وهي وحدة خيالية تحتوي على المعنى المشترك لكل الكلمات التي تتكون من هذا الجذع.

2- عبد المجيد بن حمادو،المعجم العربي الإلكتروني، مرجع سابق، ص 304.

3- نظرا للخاصية الصرفية العالية للغات السامية، مثل "العربية" و"العبرية"، استخدم "الجذر"، (أو ما يسمى بأصل الاشتقاق) بدلا من "الجذع" كأساس لتنظيم معاجمها، وتجدر الإشارة هنا إلى أن البعض قد نادى في تنظيم معاجم اللغات السامية إلى الجمع "الجذر" (أو ما يسمى بأصل الاشتقاق) بدلا من "الجذع" كأساس لتنظيم معاجمها، والصفات فعلى أساس الجذوع. للمزيد من بسط القول ينظر : اللغة العربية والحاسوب، دراسة بحثية، نبيل علي، ص 460.

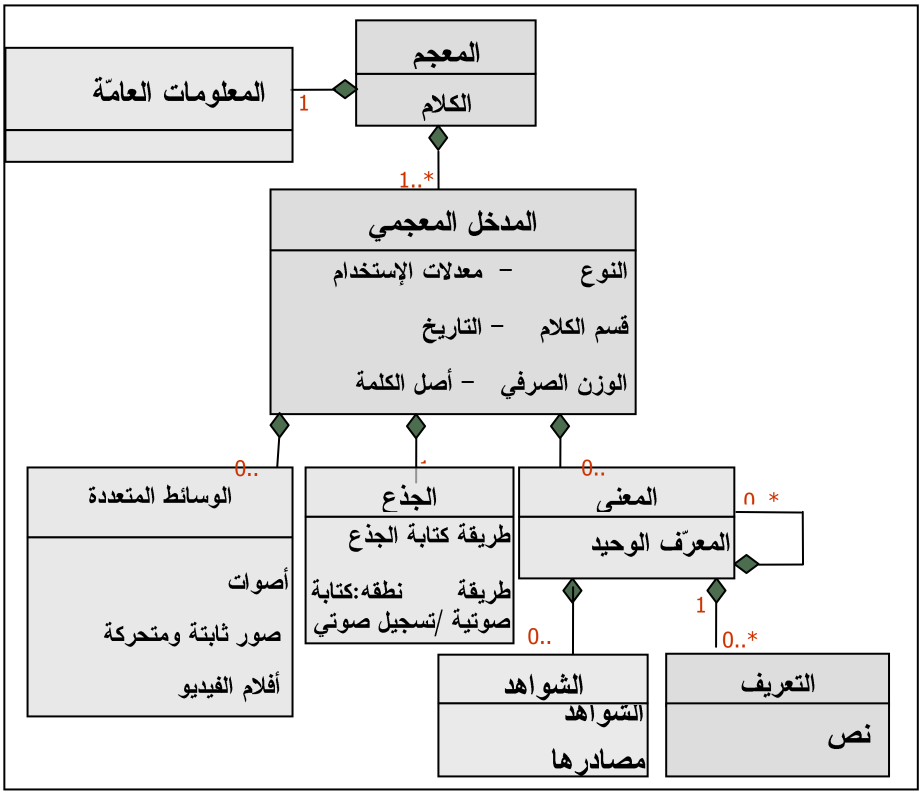
4- ويقصد بالجذع، هنا، الكلمة خالية من الزوائد التصريفية، وأدوات التأنيث والتعريف، وعلامات الإعراب، وكل ما يتصل بها من ضمائر متصلة (أمثلة: "يستخرج"، "والد"، "انبهار").

5- للمزيد من التفصيل ينظر: الباب الثاني من الأطروحة، الفصل الأول: المعالجة الآلية للغة العربية، النظام الصرفي العربية- أنموذجا، ص169.

6- للمزيد من التفصيل ينظر: الباب الثاني من الأطروحة، الفصل الأول: المعالجة الآلية للغة العربية، النظام الصرفي العربية- أنموذجا، ص 171.

7- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص 304.

- ❖ **التعريف:** يشتمل على نص لتقديم تفسير المعنى الواحد وتدقيقه.
- ❖ **الشواهد:** يشتمل هذا الصنف على[1]:
  - الشواهد (أو الأمثلة)
  - مصادرها: معاجمِ قديمة وحديثة، كتب الأدب، القرآن الكريم، الحديث النبوي.
- ❖ **الوسائط المتعددة:** يحتوي هذا الصنف على **أصوات، صور ثابتة ومتحركة**[2]، **أفلام الفيديو** وتفسيرات عامة حولها[3].



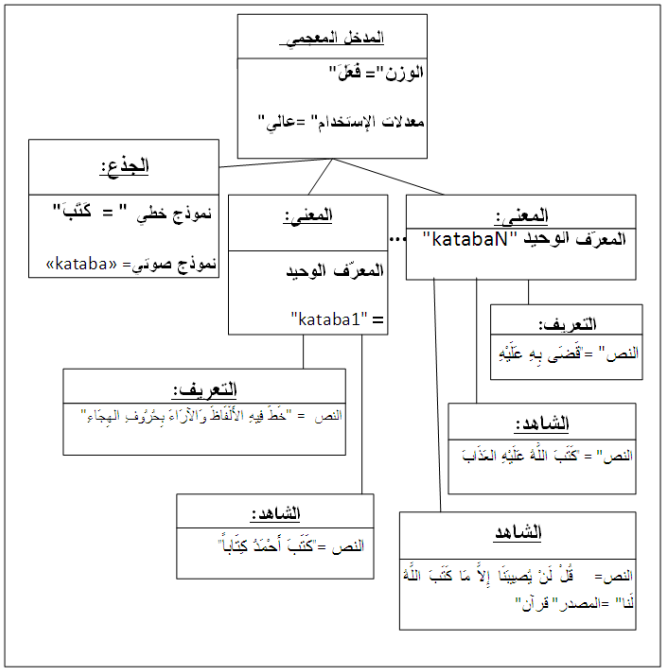
الرسم البياني رقم 2: مخطط النواة الأساسية للمعجم الإلكتروني.

وفيما يلي أمثلة لتوضيح التصميم المقترح للنواة الأساسيّة تخص الفعل "كتب" والاسم "سيناريو".

1- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص 305.

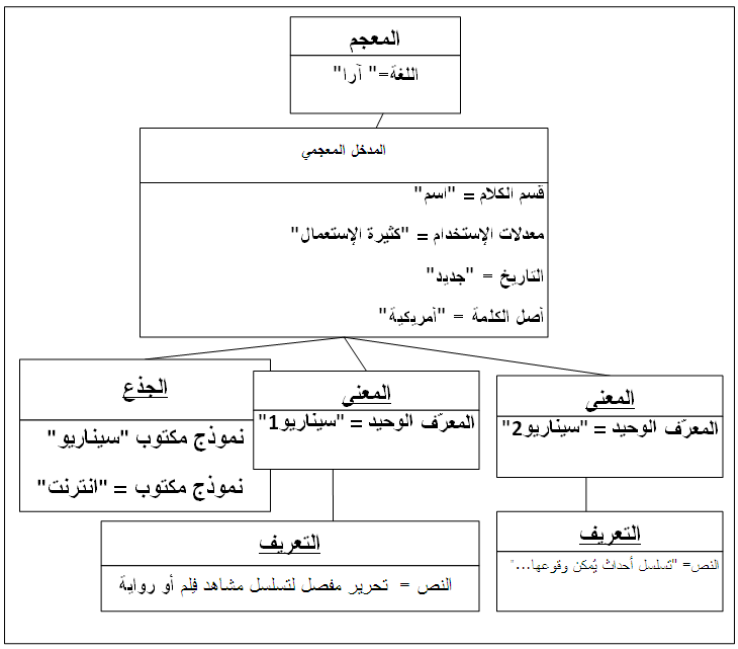
2- تلعب الرسوم التوضيحية دورا بارزا ف] إيضاح دلالات الكلمات بحيث تعد إحدى الضرورات الأساسية لأي معجم متكامل كان سواء أحادي اللغة أو ثنائي اللغة .للمزيد من بسط القول ينظر: دخليل ابراهيم الحماش، الرسوم التوضيحية ومكانتها في المعجم، اللسان العربي، دورية متخصصة، مكتب تنسيق التعريب،المنظمة العربية للتربية والثقافة والعلوم، العدد22، الرباط ، المغرب، ص 129.

3- المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص 305.



الرسم البياني رقم 3: تطبيق النواة الأساسية على فعل كتب

نلاحظ في هذا الرسم البياني أنّ فعل كتب ليس بحاجة لتحديد أصله لذلك وقع الاستغناء على هذه المعلومة بالنسبة لهذا المدخل[1].



الرسم البياني رقم 4: النواة الأساسية لكلمة حديثة ودخيلة "سيناريو"

1- المرجع السابق، ص 307.

نلاحظ أن بالنسبة لكلمة سيناريو وقع الاستغناء على تمثيل الوزن الذي كان ضروريا بالنسبة لكلمة كَتَبَ. بالمقابل يصبح أصل الكلمة معلومة أساسيّة[1].

6-2- الوحدة الصرفية:

المصطلح الأساسي في التحليل الصرفي الحديث هو مصطلح "المورفيم"[2] أي الوحدة الصرفية، التي تهتم بتمثيل الخصائص الصرفية للمدخل والعلاقات الصرفية بين المداخل[3]، وهي الوحدة الدلالية التي تعد أقل من كلمة، وتتمثل في المورفيم المتصل

ويشمل ذلك السوابق واللواحق، فالأولى مثل أحرف المضارعة، والسين الدالة على الإستقبال... والثانية مثل الضمائر المتصلة[4]. مثل الواو في الدلالة على الجمع في كتبوا[5].

تتكوّن الوحدة الصرفية من صنفين من المعلومات (انظر الرسم البياني رقم5)[6]:

- كلمة مُصرّفة : هذا الصنف يمكّنُ من تمثيل المعلومات التالية[7]:
  - العدد: مفرد، مثنى، جمع.
  - الجنس: مذكّر، مؤنّث.
  - الضمير: المتكلم، الغائب، المخاطب.
  - الصيغة: الماضي، المضارع، الأمر.
  - البناء : المعلوم، المجهول.
  - الإعراب: المرفوع، المجزوم، المنصوب.
  - قابلية الجمع: نعم، لا.
  - قابلية التعريف: نعم، لا....

الروابط بين المداخل: هذا الصنف يمكننا من ربط الكلمة بجذرها والعكس[8]:

1- المرجع السابق، ص 308.

2- حاتم صالح الضّامن، علم اللغة، بيت الحكمة،بغداد، دط، 1989، ص 58.، وينظر: محمود فهمي حجازي، مدخل إلى علم اللغة، دار قباء للطباعة والنشر والتوزيع، القاهرة، دط، دت، ص 90.

3- يرى الدكتور محمود فهمي حجازي أن هناك فرق بين الوحدة الصرفية والصورة الصرفية، فالصور الصرفية لها وجود مباشر منطوق مسموع، وأما الوحدة الصرفية الجامعة للصور الصرفية فهي موجودة غير أنه وجود غير مباشر، وذلك نحو ضرب واضطرب، فالفرق بين الصيغتين من ناحية البنية الصرفية هو الفرق بين قرب واقترب،لكن التغيير ليس واحدا من الناحية الصوتية على الرغم من اتحاد الوظيفة في بنية اللغة، ومثل ذلك أمر التاء والدال في قرب واقترب من جانب، وزهر وازدهر من الجانب الآخر، فكل من الطاء والتاء والدال تأتي في جوار صوتي بعينه، وتدخل إحداها في هذه البنية، ويقودنا هذا إلى القول بوجود ثلاث صور صرفية لوحدة صرفية واحدة، ينظر: محمود فهمي حجازي، مدخل إلى علم اللغة، القاهرة، 1978م، ص58، و حسن هنداوي، مناهج الصرفيين ومذاهبهم، ص22 ــ23.

4- مصطفى إبراهيم عبد الله، د ا رسات فى علم اللغة،دط، 1995 م، ص13

5- أحمد مختار عمر، علم الدلالة – عالم الكتب - الطبعة الرابعة،1993مـــ، ص 11.

6- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص308.

7- المرجع نفسه، ص 309.

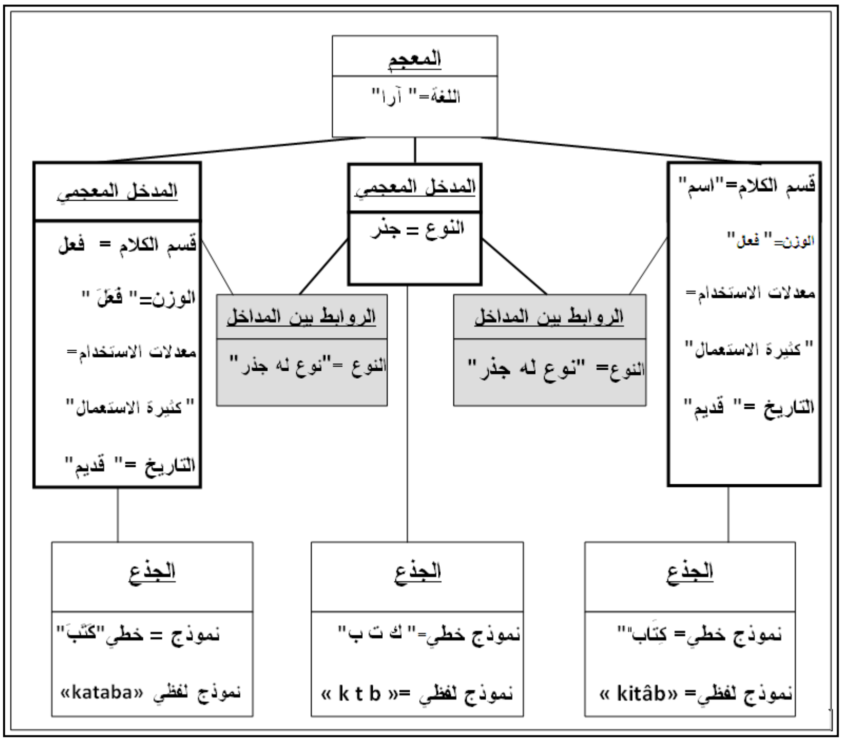
8- المرجع نفسه، الموضع نفسه.

❖ الرّابط: (يحتوي على الجذر) يمكن من الانتقال من الكلمة المشتقة إلى جذرها. مثلا من فعل كَتَبَ إلى جذره "ك، ت، ب".

❖ الرّابط يمكن من الانتقال من الجذع المشتق إلى الجذع الأصلي مثل زادَ و اِسْتَزَادَ.

❖ تمكّن هذه الرّوابط من تيسير التعرف (أو البحث) إلى جميع مشتقات المدخل الواحد (جذر أو جذع).

الرسم البياني رقم 5 يقدم توضيحا للمخطط المقترح للوحدة الصرفية بالنسبة إلى المدخل الخاص بالجذر "ك ت ب". كما يُبيّن العلاقة بين هذا المدخل والمدخلين المشتقين "كِتَاب" و"كَتَبَ[1]".

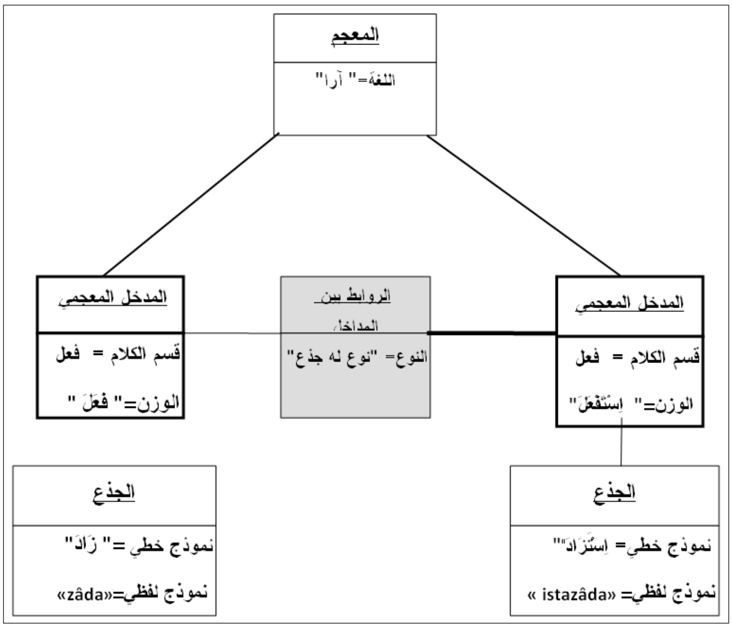


الرسم البياني رقم 5: مثال للرابط "يحتوي على الجذر "

الرسم البياني رقم 6 يقدم توضيحا لكيفيّة الرّبط بين فعل ثلاثي مجرّد مثل "زَادَ" و أحد المشتقات القياسية المزيدة مثل " اِسْتَزَادَ"[2].

1- المرجع السابق ، ص310.

2- المرجع نفسه ، ص 311.



الرسم البياني رقم 6: مثال للرابط "له جذع"

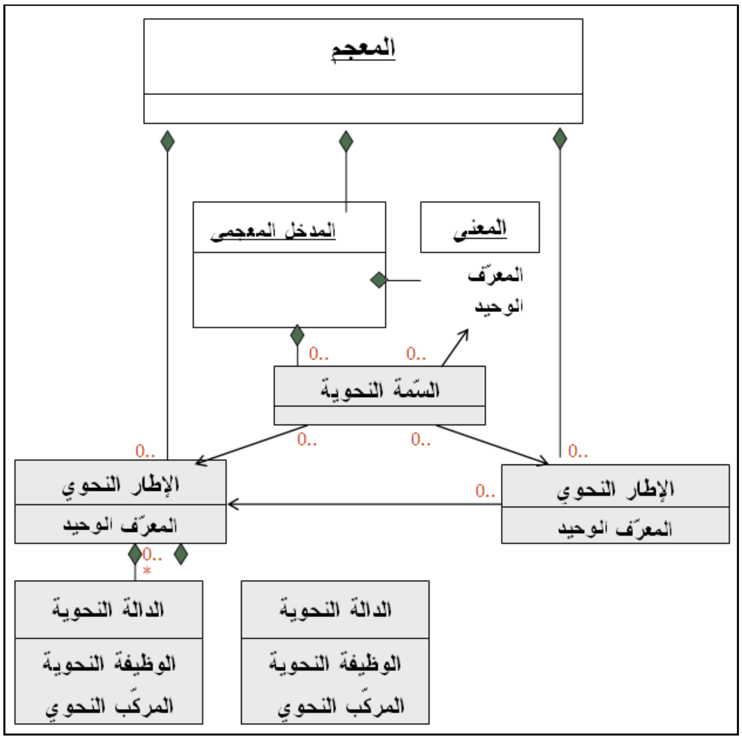
6-3- الوحدة النحوية:

تهتم الوحدة النحوية بِتمثيل السمات النحوية للمداخل المعجمية. فهي تمكّن من تعداد التراكيب الممكنة للكلمة (خاصة بالنسبة للفعل والحرف) وبالتالي يمكن من تحديد تعدي الفعل ولزومه، وكذلك تقسيم الأفعال المتعدية إلى ثلاثة أنواع: متعد لمفعول بنفسه، متعد لمفعول بأداة "حرف جر" أو "ظرف"، متعد لمفعولين[1].

تتكون هذه الوحدة من أربعة أصناف من المعلومات وهي (انظر الرسم البياني رقم7)[2]:

1- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص311.

2- المرجع نفسه، الموضع نفسه.



الرسم البياني رقم 7: الوحدة النحويّة

- **السّمة النحوية:** يجمّع هذا الصنف التراكيب النحوية المختلفة التي ترد فيها الكلمة. ويمكن تحديد المعنى الذي تؤديه الكلمة إذا ذكرت في هذا السياق.

يقع تدقيق مكونات كل تركيبة بوضعها في إطارها النحوي الخاص بها.

- **الإطار النحوي:** يقوم هذا الصنف بتفصيل السمة النحوية للكلمة[1]. فهو يهتم بتركيبة ما من جملة التراكيب الممكنة للكلمة، ويشار إلى ذلك باستعمال نمط تركيبي يضع هذه الكلمة في سياقها[2]. هذا النمط يمكن أن يكون من بين الأنماط التالية[3]:

- فعل/فاعل: إذا كان الفعل لازما.
- فعل/فاعل/مفعول: إذا كان الفعل متعد لمفعول واحد بنفسه،
- فعل/فاعل/أداة /مفعول: إذا كان الفعل متعديا لمفعول بأداة.
- فعل/فاعل/أداة /مفعول1/مفعول2: إذا كان الفعل متعديا لمفعولين.
- ...

وتجدر الملاحظة أن استعمال هذه الأطر وتدقيقها لا يقتصر على تحديد التعدي واللزوم فحسب، بل يمكّن من التعرف إلى الاستعمالات الممكنة للكلمة في الجملة.

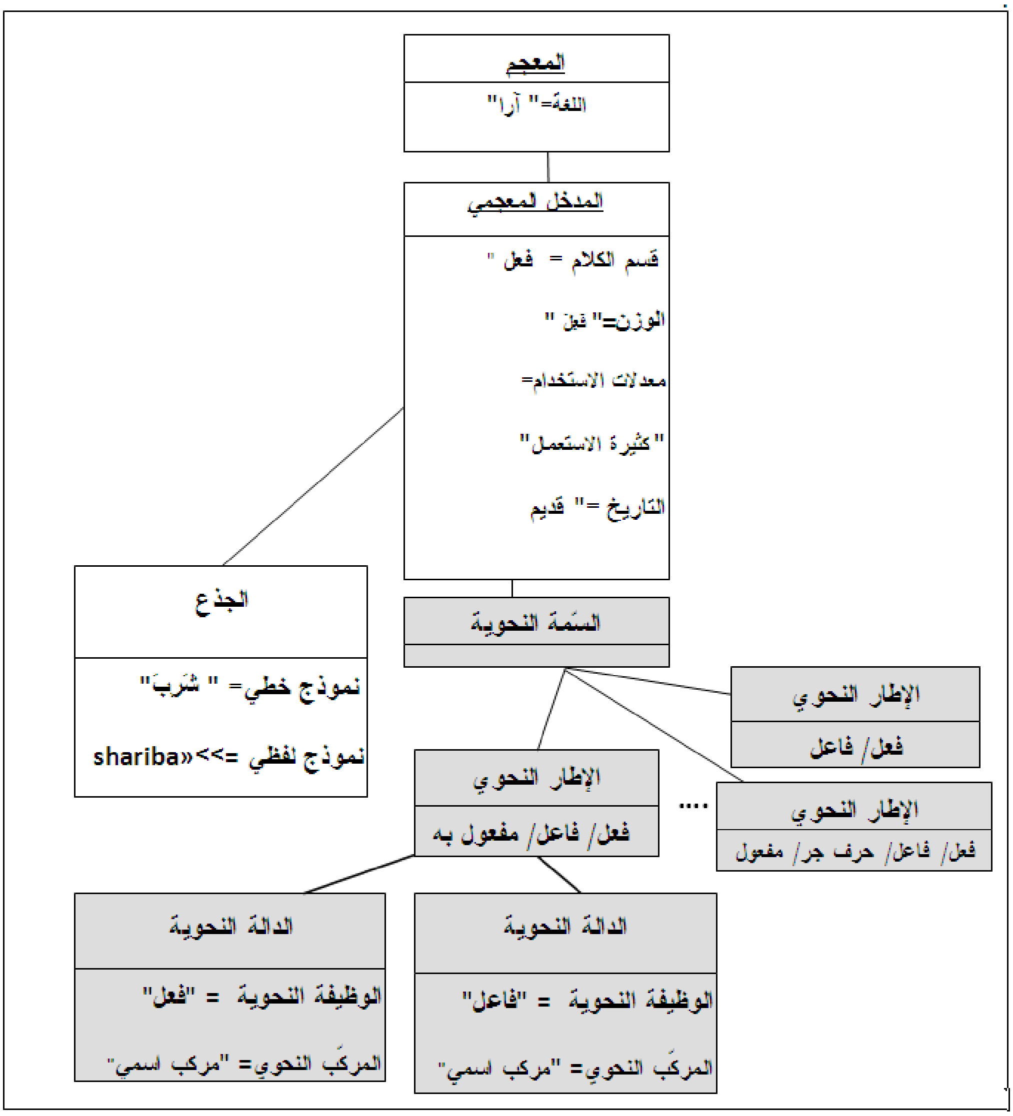
1- نبيل علي، اللغة والحاسوب، مرجع سابق ، ص 408،409.

2- المرجع نفسه، ص 410.

3- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص312.

❖ **الدالة النحوية:** وهي "الدلالة المحصلة من استخدام الألفاظ أو الصور الكلامية فى الجملة المكتوبة أو المنطوقة على المستوى التحليلي أو التركيبى"[1]، يستعمل هذا الصنف لتدقيق أحد خصائص الإطار النحوي[2]. وذلك بتحديد المعلومات التالية[3]:

- **الوظيفة النحوية:** فعل ، فاعل.
- **المركّب النحوي:** مركب اسمي...



الرسم البياني رقم 8: الوحدة النحوية لفعل "شَرِبَ"

1– فريد عوض حيدر، علم الدلالة، دراسة نظرية تطبيقية، مكتبة النهضة المصرية القاهرة – الطبعة الأولى،1998، ص12.

2– تقوم هذه الخطوة بتطبيق القيود الدلالية، والتي يتظمنها المعجم في صورة تصنيفات نحوية وأطر دلالية لمفرداته، لتستبعد من البنى النحوية التي تم استخراجها في تلك التي تتنافى مع منطق دلالات الألفاظ، بعد ترشيح البنى النحوية غير المتسقة دلاليا، تبقى لدينا بنية الجملة الصحيحة وتنتهي بذلك عملية التحليل النحوي الآلي.للمزيد من التفصيل ينظر: نبيل علي، اللغة والحاسوب، مرجع سابق ، ص 309،310.

3– عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص312.

يوضح هذا الرسم البياني كيف يقع ربط فعل "شَرِبَ" مع الإطارات النحوية الممكنة التي تخصه وهي[1]:

فعل/ فاعل (شَرِبَ الوَلَدُ)

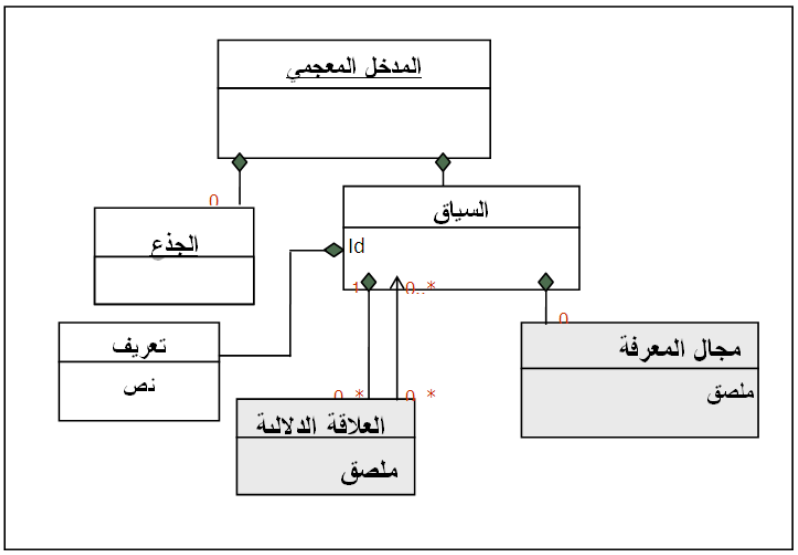
فعل/ فاعل/ مفعول به (شَرِبَ الوَلَدُ المَاءَ)

فعل/ فاعل/ حرف جر/ مفعول به (شَرِبَ الوَلَدُ بِسُرْعَةٍ)

6-4- الوحدة الدلالية:

تهتم الوحدة الدلالية بتدقيق معاني المداخل وذلك بتمثيل العلاقات الدلالية بينها (الترادف والتضاد) وكذلك بتحديد مجالاتها[2].

تحتوي الوحدة الدلالية على صنفين من المعلومات (انظر الرسم البياني عدد9)[3]:



الرسم البياني رقم 9 : مخطط الوحدة الدلالية الذي تعنى بالعلاقة الدلالية في المعجم الإلكتروني

❖ **العلاقة الدلالية (Sense Relation):** تربط هذه العلاقة بين معنيين مختلفين.

يمكن تصنيف هذه العلاقة الدلالية إلى عدة أنواع: الترادف، التضاد، التعميم (طالب – إنسان)، التخصيص (إنسان– طالب)، جزء من (اليد – جسم الإنسان)، التسلسل الزمني (الخميس – الجمعة)،... هذا الصنف من المعلومات يلعب دورا مهمّا في بناء شبكة دلالية "أنتولوجيا" انطلاقا من المعجم[1].

1- المرجع السابق، ص 314.

2- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص314.

3- المرجع نفسه، ص 315.

1- المرجع نفسه، الموضع نفسه.

مجال المعرفة (Subject Field): يعنى هذا الصنف بمجال معنى الكلمة[1]. (الكلمة الواحدة يمكن أن تتعدد مجالات استعمالها: رياضيات، علم النبات، لغة، طب،...)

7 - تقييس المعاجم الإلكترونية :

تقييس الموارد المعجمية هو عمل جماعي تقوم به مجموعة من الخبراء لوضع تصور موحد لهذه الموارد يشمل جميع اللغات. هذا التصور يصبح مقياسا إذا وافقت عليه جهة رسمية وطنية أو عالمية مثل منظمة إيزو "ISO"[2] أو اتحاد شبكة الويب العالمية [3](W3C). وكمثال نذكر مقياس "LMF" Lexical Markup Framework)) لمنظمة "إيزو" والمعرف بـ ISO 24613. لتقييس الموارد المعجمية مزايا عديدة نذكر أهمها[4]:

- المواصفات المقترحة صالحة لبناء معجم إلكتروني للاستعمال البشري وأيضا للمعالجة الآلية.
- تيسير عملية تبادل المعاجم بين الأشخاص والمؤسسات قصد الإثراء المشترك وذلك باستغلال البنية الموحدة.
- إمكانية دمج معاجم ثنائية اللغة بقصد توليد معجم متعدد اللغات للترجمة الآلية مثلا.

1- المرجع السابق، ص 315.

2- المنظمة الدولية للمعايير :أيزو بالفرنسية: (International Organization for Standardization) و بالإنجليزية : (Organisation internationale de normalisation):

هي منظمة تعمل على وضع المعايير، وتضم هذه المنظمة ممثلين من عدة منظمات قومية للمعايير. تأسست هذه المنظمة في 23 فبراير 1947 وهي تصرح عن معايير تجارية وصناعية عالمية. يكمن مقر هذه المنظمة في جنيف، سويسرا .بالرغم من أن الأيزو تعرف عن نفسها كمنظمة غير حكومية، ولكن قدرتها على وضع المعايير التي تتحول عادة إلى قوانين (إما عن طريق المعاهدات أو المعايير القومية) تجعلها أكثر قوة من معظم المنظمات غير الحكومية. تؤلف منظمة الأيزو عمليا حلف ذو صلات قوية مع الحكومات.

تأتي كلمة أيزو (ISO) من اليونانية) ἴσος إسوس (/ísos/ بمعنى المساواة؛ وبالتالي فهي ليست اختصاراً لعبارة International Organization for Standardization\الإنجليزية ولا Organisation Internationale de Normalisation الفرنسية وإنما إشارة إليهما. قد ترد الكلمة على صورة أخرى — إيزو أو أيزو، تشبيها باللفظ الفرنسي /izo/ أو الإنجليزي /ˈaɪsoʊ/ و /ˈaɪsəʊ/ و /ˈaɪzo/أيضا. بما إن الأعضاء المؤسسين للمنظمة يعلمون أن الأحرف ISO لا تشكل بالضرورة الأحرف الأولى من كلمات اسم المنظمات بجميع اللغات لذلك قد قرروا أن يكون ISO الاسم الدولي للمنظمة. يعكس هذا الأمر هدف المنظمة وهو المساواة بين الثقافات المختلفة.

وثائق الأيزو محمية بحقوق النشر وتأخذ المنظمة رسوم معينة مقابل الحصول على نسخ من معظم المعايير. ولكن الحصول على معظم المسودات بشكل إلكتروني مجاني. رغم أنها مفيدة، يجب الحرص على استخدام هذه المسودات فهناك إمكانية لتغيير كبير قبل أن يصبح المعيار نهائيا. بعض المعايير من قبل المنظمة وممثلها الرسمي للولايات المتحدة (واللجنة الكهروتقنية الدولية في الولايات المتحدة عن طريق اللجنة الوطنية وإتاحتها) متوفرة بشكل مجاني، لدى المنظمة 162 عضوا وطنيا، من مجموع 195 بلدا في العالم

3- معايير الويب: (W3C) : هي مجموعة من المعايير تم انشائها من قبل منظمة رابطة الشبكة العالمية أو بالعربية :اتحاد شبكة الويب العالمية بهدف تسهيل قابلية الوصول إلى المعلومات، و لتسهيل التعامل معها باستخدام المنتجات المتخصصة في شبكة الويب، كالمتصفحات و برامج تحرير المواقع و برامج الإدارة، و إزداد الاهتمام بمعايير الويب خاصة في ظل التطور الكبير في أدوات تصفح الإنترنت و منها: الهاتف النقال، و الحاسب، و لتلبية حاجات الناس جميعا في تصفح الإنترنت إن كان الإنسان صحيحا أو معقوفا.

و من فوائد معايير الويب : تصغير حجم الموقع المبني بمعايير الويب و التقليل من وزنه، فالمواقع التي يتم كتابتها بمعايير W3C القياسية تكون عادة أصغر حجما بمقدار 20% إلى 60% من المواقع التي لا يتم كتابتها بهذه المعايير، و استخدام المعايير بالشكل الصحيح يؤدي إلى تقليل تكلفة الصفحات كبيرة الحجم.

يوجد العديد من المنظمات التي تضع معايير الويب كمنظمة ECMA حيث وضعت هذه المنظمة بما يفوق 270 مواصفات لمعايير الويب.

4- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص299.

- بناء معاجم تستجيب لحاجيات المستخدم.
- تبادل الأدوات الحاسوبية مثل:
  - البحث عن المعلومات داخل قواعد البيانات والمعاجم.
  - التدقيق الإملائي.
  - التحليل الصرفي.
  - التشكيل الآلي.
  - التصنيف الآلي.
  - احتساب البيانات الإحصائية.

اعتمادا على مقياس LMF وقع بناء العديد من المعاجم الإلكترونية لأغلب اللغات (الإنجليزية، الفرنسية، الإسبانية، الإيطالية، البنغالية، وعدة لغات آسيوية،...).

بالنسبة إلى اللغة العربية يمكن القول إن عملية تقييس المعاجم لم تلقَ الاهتمام الكبير من الجهات المختصة. وفي ما يلي سنركز على الأبحاث التي أنجزت بمخبر "ميراكل" بجامعة صفاقس بالتعاون مع منظمة "إيزو" للتقييس ترتكز على مقياس LMF وأفضت إلى بناء قاعدة بيانات معجمية للغة العربية[1] . من هذه القاعدة يمكن توليد معاجم متعددة حسب الحاجة كما يمكن توليد شبكة دلالية عربية "أنتولوجيا".

يمتاز مقياس LMF بمرونته وبإمكانية تغطيته مختلف مستويات اللغة (الصرفي، النحوي، الدلالي،...). يعتمد LMF على تصميم وَحَداتي[2] يمكّن من[3]:

- اختزال حيز تمثيل المعلومات المعجمية بعدم تمثيل المعلومات الخاصة بالصيغ المصرفة القياسية والاقتصار على تمثيل القواعد التي تمكّن من توليدها.
- الفصل بين المستويات اللغوية لمداخل المعجم في شكل وحدات متخصصة تعنى كل وحدة بمستوى معيّن مع إمكانية الربط بين هذه الوحدات عند الحاجة إلى ذلك. هذا الفصل يمكّن من إضافة وحدات جديدة دون المساس بالوحدات الموجودة.

لقد تطرقت في هذا المبحث إلى جوانب عديدة تميز المعجم الإلكتروني عن المعجم الورقي من حيث الهيكلة والمحتوى وطرق الاستعمال. بينا بالخصوص أهمية المعجم الإلكتروني العربي بالنسبة إلى مجتمعنا الذي يتحول بسرعة إلى مجتمع المعرفة؛ حيث تلعب فيه التكنولوجيا دورا محوريا. لقد أصبح المعجم الإلكتروني أداة عمل لا يمكن الاستغناء عنه بالنسبة إلى الكثير من

1- ينظر : Baccar F., khemakem A., Gargouri B., haddar K., Ben Hamadou A. (2008). "LMF standardized model for the editorial electronic dictionaries of Arabic", NLPCS 2008, 12-13 june, Barcelone

2 – modulair

3- عبد المجيد بن حمادو،المعجم العربي الإلكتروني: أهميته وطرق بنائه، مرجع سابق، ص300.

الفئات الاجتماعية والمهنية، وهو كذلك حجر الزاوية للعديد، إن لم نقل لجل، تطبيقات المعالجة الآلية للغة. هذه التطبيقات تساهم بصفة جدية في إشعاع اللغة العربية واستمرارها كلغة ناقلة. كذلك تطرقنا إلى مختلف طرق بناء المعجم العربي وختمنا بموضوع تقييس المعاجم الإلكترونية لجعلها تخضع إلى مواصفات المعاجم الحديثة من خلال تجربة مخبر ميراكل. إتاحة مثل هذه المعاجم للغة العربية سيدعم حتما دورها في مجتمعنا الذي يعيش منافسة شرسة من لغات أجنبية أهمها الإنجليزية والفرنسية، وفي ضو ء هذه التحديات سأتعرض في المبحث التالي إلى نماذج لقاعدة البيانات المعجمية للمعاجم الإلكترونية تؤسس لبناء معاجم إلكترونية من منظور نظرية لسانية حاسوبية ،فمسؤوليتنا اليوم أعظم مما كانت عليه بالأمس لأن للغة أبعاداً ثقافية، واقتصادية واستراتيجية.

ثانيا- المعجم العربي الإلكتروني بين النظرية والتطبيق: نماذج لقاعدة البيانات المعجمية للمعاجم الإلكترونية:

إن المشتغل في الصناعة المعجمية، سواء في وضع المعاجم العامة الورقية أو بناء المعاجم الإلكترونية سيجد نفسه أمام مناهج متشعبة وبأهداف واستراتيجيات متشابهة ومشتركة في مراقبة وتحليل وتمثيل المعطيات اللسانية في أكبر عدد ممكن منها . وبما أن جهد الانجاز والالحاحية متطابقان بينهما تماما. وأن اللغة هي الموضوع الأساس بالنسبة لهما، فإن الحد الفاصل بينهما هو أن القاموس اليدوي يمكن تحويله إلى قاموس محوسب وفق إجراءات تكنولوجية لا يمكن تجاوزها، باستثناء ما يمكن استغلاله من مدونة معجمية قد تصبح مصدرا لأساسيات المعطيات النصية. ما عدا ذلك لن يصبح قاموسا مبينا على متون إلكترونية، لأن ذلك عمل من صميم المعجميات الحاسوبية المحضة. ولقد لخص فيلب أمييل (P.Amiel)[1] المعجم الالكتروني لكونه ليس معجما لنصوص واصفة للغة، بل العمل المعجمي في الحالة الالكترونية هو وضع دلالة الوحدات المعجمية وترتيبها وشرحها وتسنينها في لغة اصطناعية لمستعملي القاموس[2]. مما أفرز ثلاث أجيال متتالية وهي :

- قواميس الكترونية ذات قواعد بيانات نصية مفهرسة .
- قواميس الكترونية ذات قواعد بيانات معجماتية (تعريف، شواهد..) مُبْنِيَنَة.
- قواميس الكترونية ذات قواعد بيانات معجماتية ولسانية بحيث يكون ولوج أي مكون أو نص معجمي ممكنا '(حالة مشروعي(Hachette,Eurolexe).

1- يعمل فيليب أمييل مديرا بشركة انتلكسيس Antelellexis بفرنسا، للمزيد انظر مقالته المنشورة تحت عنوان: p.Amiel vers les dictionnaires electroniques 2004 وهي بالموقع الالكتروني للشركة المذكورة.

2- للمزيد من بسط القول ، ينظر: محمد الراضي ، نادية عمري، الخصائص التركيبية للوحدات المعجمية، المعجم العربي المولد،إشرف عبد القادر الفاسي الفهري، منشورات معهد الدراسات والأبحاث للتعريب، 2002، ص 156 ومابعدها.

وفي الحالة الأخيرة نلاحظ اهتمام متزايد للبنية الصغرى للقاموس[1] حيث تمثلها العلاقات الشجرية المتداخلة بين المواد الرئيسية للمعجم فللأفعال التالية"قاء" و"لقي وخرم" بنية شجرية تتضمن جميع الكلم المرتبطة بهما والتعبيرات التي تنتمي إليهما دلاليا، بحيث نجد في نصهما علاوة على التراكيب البسيطة التعبيرات المركبة والمتلازمة مثل "قاءَ"، "زيدٌ"، "نَفْسَهُ" أو "لقيت هند حتفها" أو "خرمته الخوارم" والتي تدل كلها على الموت. أما أسلوب تمثيل المعلومات المعجمية قد يكون بالطريقة الشجرية كما قد يكون بالطريقة الأقواسية.

وعليه، في هذا المبحث سنركز قدر الإمكان، على بعض المقاربات الرائدة عالمياً في بناء المعاجم الآلية وخاصة مقاربة المعجم التفسيري التأليفي المؤطرة بنظرية معنى – نص لصاحبها إيغور ملتشوك وزملاؤه[2] الذين سعوا إلى معالجة المعجم والمعجماتية والقاموس مرة واحدة[3]. كما سنبسط القول في تجربة المعاجم الإلكترونية ديلا (DELA) التي أرسى دعائمها الرواد موريس كروس وماكس شلبررشتاينوبلاندينكورتواز وغيرهم كثير، حيث نظرت هذه المقاربة إلى مختلف الإشكاليات المتعلقة بالبناء المعماري للمعاجم الإلكترونية بجميع أصنافها، ساعية إلى تطبيق لغة عقلانية وأدوات معلوماتية تعالج **المداخل المعجمية**[4]، باتباع منظومة من المعاجم المبنية والمترابطة فيما بينها، مرتكزة على المعجم الآلي للجذور والمعجم الآلي للمفردات البسيطة والمعربة والمعجم الآلي للمفردات المركبة والمتلازمة المعربة[5].

---

1 – microstructure.

2– يشرف على هذا مجموعة من رواد نظرية معنى – نص التي أسست لمقاربة قاموس المعجم التفسيري التأليفي الرجل الكبير إيغور ميلتشوك الواسع الاطلاع بخبايا ظواهر اللغات الطبيعية بحيث قام بجرد حوالي ستة وخمسين وظيفة معجمية معيار تغطي العلاقات الجدولية بين العجمات (lexies) كالترادف و الاشتقاق والتضاد وبين العلاقات المركبة كأفعال = العماد والتوسيم والتفعيل وذلك على مستوى العديد من لغات العالم، وقد أسس لهذا الغرض محترفا لسانيا لنظرية معنى – نص (OLST) « Observatoire de linguistique Sens – Texte » والتي أطرت المقاربة المعجمية التي نحن بصدد التعريف بها وبمفاهيمها وتطبيقاتها الآلية التي تهمنا هنا، للمزيد ينظر محاضرة متلشوك بالكوليج دو فرانس (vers une linguistique sens-texte, collège de France, chaire internationale) و ايغور ملتشوك واندري كلاس والان بولغير في (Introduction à la lexicologie explicative et combinatoire, 1995 ed. Duculot).

3– هناك نماذج عديدة من المعاجم الآلية المرتكزة على قواعد المعرفة الذي يلائم نظام اللغة العربية في كثير من جوانبه تطرق إليها الدكتور عز الدين غازي في أطروحته الموسومة: بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، و من باب الأمانة العلمية للبحث. ولبسط القول في هذا المبحث سأتطرق بنوع من التفصيل محاولا تسليط الضوء قدر الإمكان على مختلف النماذج كما ذكرها الدكتور عز الدين غازي في أطروحته.

4– ينظر : جواد حسني سماعنه، المعجم العلمي المختص، (المنهج والمصطلح)، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 48، الرباط، المغرب، ديسمبر1420هـ/1999م، ص 42، 43.

5– لقد عولجت هذه المركبات بصفة معمقة في لغات طبيعية عديدة مع مجموعة من الدارسين أهمهم: ماكس شلبرشتاين (M. SILBERZTEIN 93) موريس كروس (M. GROSS 1988) وايريك لابورت (LA PORTE. E. 1988) وجاك لابيل (LABELLE. J. 1988) وصاحب نظرية المقاربة المشهورة معنى – نص = ايغور ميلتشوكديكو (Dico) وفق أساليب الفحص والإتمام والاستنباط المنطقي المستندة إلى قيم معرفية ودلالية للمداخل المعجمية.

2-1- القاموس المفهومي الإلكتروني المرتكز على قواعد المعارف:

يعتبر القاموس الالكتروني المفهومي للغة اليابانية[1] (EDR) مشروعا متقدما يهدف إلى تطوير آلة إلكترونية[2] مصممة لغرض الإحاطة الشاملة بالمعارف البشرية. وهي تروم فهم اللغة الطبيعية وتوليدها والاضافة عليها اعتمادا على نظام القواعد المعرَّفة[3] ، بما في ذلك التعامل مع الخوارزميات اللسانية في سائر مستوياتها وأشكالها الصورية وتمثيل المعلومات تمثيلا دقيقا[4]. ولهذا السبب، اِعْتَمَدَتْ هذه المقاربة المفهومية كمًّا هائلاً من النصوص تُسمَّى بأساسيات المعطيات النصية المبنية – تحليلا وتوليدا- على المعرفة البشرية والكفاية الطبيعية. ومن خصائص هذا القاموس أنه[5]:

*- يشمل جميع معطيات اللغة، مفردات وتعبيرات وبنيات، كما يستند إلى حجم ضخم من النصوص تضم كل المعارف البشرية.

*- يكون مزودا بنظام لقواعد المعرفة اللازمة للتدقيق في البنيات الدلالية التي غالبا ما تكون غامضة.

ويحتوي هيكل القاموس المفهومي المذكور على خمس معاجم تعتبر حجر الزاوية للقاموس المفهومي وهي:

1. قاموس ثنائي (ياباني –انجليزي) و(إنجليزي–ياباني).
2. قاموس مفهومي ويضم القاموس المفهومي الرئيسي و القاموس المفهومي التصنيفي القاموس المفهومي الوصفي.
3. القاموس السياقي ويضم القاموس السياقي للغة الياباني و القاموس السياقي للغة الإنجليزية.
4. قاموس المدونة اللغوية ويحوي المدونة اللغوية الانجليزية.

1- القاموس الالكتروني المفهومي الياباني تجربة عالمية رائدة لبناء آلة Dractable قادرة على معالجة اللغات الطبيعية بشكل مرن إعتمادا على قواعد البيانات معرفية متطورة تشمل جميع المعارف والخبرات البشرية في انتاج البنيات اللسانية انطلاقا من مجموعة من النصوص الغاوية الأساسية :مفردات وتعبيرات متلازمة ومركبات لفظية ومصطلحية تقنية وعلمية ومفاهيم وحقول دلالية وعلاقات قضوية منطقية وتسلسل هرمي /شجري وابراز الوحدات المشاركة في الحدوث العلاقات الدلالية والسياقية والتواترية للمداخل والجمل.وقد أشرف على بناء هذا المشروع معهد البحوث الإلكترونية بطوكيو ،بشراكة مع (Japan Key Technology) وبالتعاون مع جهات وشركات عديدة نذكر منها شركة شارب وماتسوشيتا وأوكي وتوشيبا وميتسوبيشي وهيتاتشي للمزيد من المعلومات أنظر للموقع الالكتروني للمشروع http/ www.iignet.or.gp/edr/g_index.html

2 – Tractable.

3 – Knowledge base

4- عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، أطروحة جامعية لنيل الدكتوراه، إشراف الأستاذ الدكتور: عبد الغني أبو العزم، والأستاذ الدكتور: محمد لهلال، كلية الآداب والعلوم الإنسانية، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجميات،جامعة الحسن الثاني، عين الشق- الدار البيضاء، السنة الجامعية:2008/2007، ص171.

5- عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، المرجع نفسه، ص 171.

وينطلق هذا القاموس بتحديد الحقول المفهومية أولا، وبعده يتعرف على المكونات التي قد تشكل هذا الحقل أو ذاك مثلا: لغظة "تفاحه" هي الرابط المفهومي لأشكال لفظية مثل:المطبخ+السكين+العصير+كأس....إن العلاقة هنا بين مفهوم وآخر تقتضيها العلاقة بين الحقول بالتحاد والتقاطع والتسلسل الهرمي الذي أنشأها مما يحدد درجة القرب أو البعد والاتصال أو الانفصال، وهذا ما حذا. بمصممي هذا القاموس باعداد ما يسمى بالمفاهيم الرئيسة للمعجم ليسهل التزويد والإضافة والتعديل في استخدام تلك المفاهيم من قبل المستخدمين لهذا القاموس وفق الإجراءات التالية[1]:

أ/ القيام بمعالجة ذكية لكلمات اللغة: تعزيز التحويل بدقائق معرفية تجنبا للغموض الذي قد يكتنف بعض المتواليات الموجودة في النصوص المعجمية.

ب/ إعطاء البعد الدلالي حقه في تحديد مستوى المفاهيم.

ت/ استرجاع المعلومات بشكل ذكي، وذلك بالتعرف على المعلومات اللازمة عن طريق الاستدلال[2] ووفقا لقاعدة وقائع مضبوطة.

ث/ فهم محتويات وثيقة معينة وذلك بتلخيصها وفق أساليب ذكية.

ج/ استعراض وتصحيح الأخطاء النحوية والتدقيق فيها.

ح/ وصف اللغة في سائر مستوياتها بواسطة برمجيات متطورة.

خ/ اعتماد نظام خبير[3] يكتسب المعرفة ويتغذى من النص بشكل دائم .

و/ تفهم الاستعلام الفوري عند استفسار المستخدم للنظام بتفهم مضمون الحوار وحل الغموض المترتب عن الحوار ثم توليد خطاب الرَّدِّ بشكل دقيق وذكيٍ.

ويفيد هذا القاموس أساسا في المعالجة الآلية للغات الطبيعية والترجمة الآلية وتطوير التكنولوجيا ذاتها والعديد من التطبيقات الحاسوبية واللسانية. كما يحتوي على مجموعة من الأصناف والمحتويات ومنها الألفاظ والمفاهيم القاموسية والقاموس الثنائي اللغة[1]. والهدف

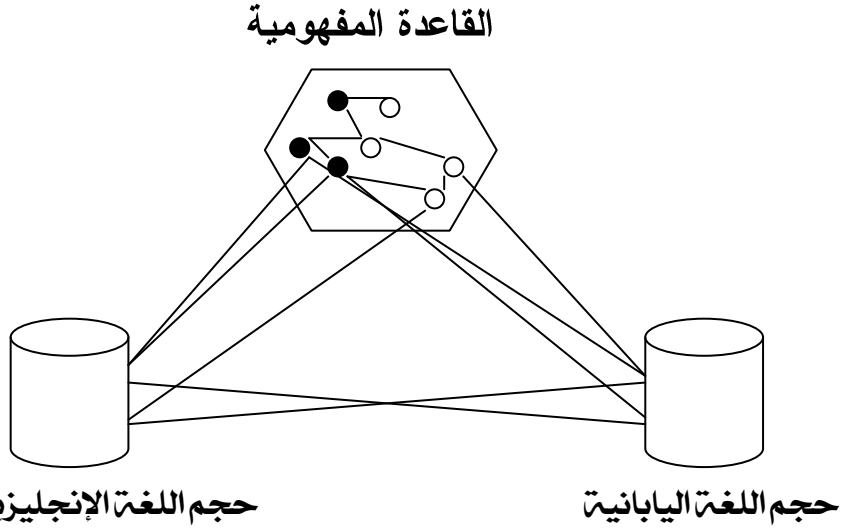
1– عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، أطروحة جامعية لنيل الدكتوراه، إشراف الأستاذ الدكتور: عبد الغني أبو العزم، والأستاذ الدكتور: محمد لهلال، كلية الآداب والعلوم الإنسانية، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجميات،جامعة الحسن الثاني، عين الشق– الدار البيضاء، السنة الجامعية:2008/2007، ص172.

2 – inférence.

3 – Système expert.

1– عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، أطروحة جامعية لنيل الدكتوراه، إشراف الأستاذ الدكتور: عبد الغني أبو العزم، والأستاذ الدكتور: محمد لهلال، كلية الآداب والعلوم الإنسانية، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجميات، جامعة الحسن الثاني، عين الشق– الدار البيضاء، السنة الجامعية:2008/2007، ص173.

المتوخى من هذا المشروع–الذي يعتبر ثمرة مجهود طال اثنتي عشرة سنة[1]، وهي ما بين (1982–1994)– هو ارساء بنية معرفية للمعالجة الآلية للمعلومات تساهم في بناء آلة قارئة تصنف المعارف المعجمية للغتين الانجليزية واليابانية بما فيها الألفاظ القاموسية والقاموس المزدوج للغة وقاموس المتلازمات (القاموسس الانسياقي) علاوة على بناء بنوك معطيات هدفها المساعدة على تصنيف المفاهيم(القاموس المفهومي) وقواعد معطيات المدونة متن . وحتى يتبين عمل هذا القاموس نورد الشكل التالي[2]:



الشكل رقم 01: يمثل القواعد المفهومية وعلاقتها بحجم المعجم الثنائي اللغة.

إن بناء مثل هذا القاموس الثنائي اللغة والمؤسس على قواعد المعرفة التي هي ليست كقواعد المعطيات التقليدية التي تعتمد على مبدأ التجميع ومراكمة المعلومات، بل على طبقية من المعلومات تتوسطها عمليات للتصفية[3] لاختيار المعرفة المناسبة التي تستثمر عبر خطوات ومراحل ووسائل وطرق ذات صلة متراصة ومبنينة والمعرفة بتعبير روبين كلارك (Clarck)[4] هي الآلة الوحيدة للانتاج وبحسب جيمس مادسون[1] هي الوسيلة التي سوف تحكم الجهل .هذا ما أكدته أدلة وصحة ثورة المعلومات، وهي أن "المعرفة " معيار النجاح والتفوق والقوة ووسيلة

1– عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة،المرجع نفسه ، ص173.

2 – Mathieu Mangeot-lexicologie et lexicographie multilangues,master ICA, Université de savoie France 2006 page 27.

3 – filtre

4– عز الدين غازي ، نظام قواعد معرفة صرافي –صواتي للغة العربية، مقاربة الفعل، رسالة لنيل دبلوم الدرسات العليا تحت اشراف د.محمد الحناش كلية الاداب ظهر المهراز فاس ،1999 ص 37.

1 – المرجع نفسه، ص37.

التعامل والتفاعل، مهي نوعان: معرفة ثابتة ومعرفة متغيرة تتغير نتيجة اكتساب الخبرة في مجال معين[1].

ولما أصبح الحاسوب أداة لحل المسائل، برز سؤال أساسي هو هل المعلومات هي المورد الأساسي لحل المسائل؟ من هذا المنطلق، ظهر للجميع أن قواعد البيانات[2] الضخمة لا تحل كثيرا من المشاكل، فبادر مهندسوا الحاسوب إلى بناء نظام أكثر تفوقا ونجاعة من الأنظمة المعتمدة على قاعدة البيانات، التي ما هي إلا قوائم معطيات عن مجال ما وليس معرفة بجوهر بنية المجال.

وعلاوة على ذلك، فإن أسلوب قواعد المعارف ذو فائدة كبيرة في مجال تحديث معالجات اللغات الطبيعية عموماً، بداية من معاجمها وسائر مستوياتها اللغوية. وهوما أراده الدكتور عز الدين غازي في أطروحته **بناء نظام لمعجم مفهومي تفاعلي باستطاعته أن يضيف شيئا ما على قاعدة المعارف لتحسين الأداء نتيجة الاستفادة من الهفوات والأخطاء السالفة، بمعنى أنه كلما توغل المرء في حل مسألة لغوية مثلا، إلا وبحث عن أساليب أنجع في القيام بهذه الحلول وبذلك نكون دائما مع هذه القواعد، بصدد الإضافة واختيار الأفضل والأمثل**[3].

أما من حيث أساليب تمثيل المعارف لبناء نظام معلومات ذكي، لابد من بناء ما يسمى ب "قواعد المعرفة" وليس " قاعدة البيانات" التي تحتوي على المعلومات غير المبنينة، عكس قاعدة المعرفة التي تشمل على كل من المعارف الثابتة والمتغيرة في مجال مخصوص. ولبناء هذه القاعدة لابد من معها وبالتالي معالجتها. وفي الواقع يوجد العديد من أساليب تمثيل المعرفة وهو فرع هام من مجالات علم الذكاء الاصطناعي. ونذكر بعض هذه الأساليب كما يلي :

أ- **أسلوب الشبكات الدلالية**[4] **:** وتساعد على استخلاص الحقائق الخمس المستخرجة من المثال التالي[5] :

"لون البرتقالة أصفر" وتمثل بقاعدة المعرفية التالية :(لون <البرتقالة أصفر>)

"مذاق البرتقالة حامض" وتمثل بالقاعدة المعرفية التالية (مذاق<البرتقالة حامض>)

1- عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، أطروحة جامعية لنيل الدكتوراه، إشراف الأستاذ الدكتور: عبد الغني أبو العزم، والأستاذ الدكتور: محمد لهلال، كلية الآداب والعلوم الإنسانية، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجميات،جامعة الحسن الثاني، عين الشق- الدار البيضاء، السنة الجامعية:2008/2007، ص174.

2 – Data Base

3- عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، أطروحة جامعية لنيل الدكتوراه، مرجع سابق، ص174.

4 – Semantic networks

5- عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، ص174.

ب- أسلوب القاعدة[1] : ويعتبر أسلوب القاعدة من أبسط وأهم الطرق المستخدمة في تمثيل المعرفة الثابتة. في هذا الأسلوب تكتب الحقائق على شكل القاعدة الشرطية [إذا كان (شرط) إذن (استنتاج)]. كما توضح الأمثلة التالية وهي في مختلف المجالات لنأخذ مثالا:

إذا كان الجو مستقرا وإذا كان الجو صحوا ، إذن سيكون الجو جميلا

ج- أسلوب الإطارات[2] : يستخدم هذا الأسلوب في تمثيل الحقائق والمعلومات الثابتة كما يستخدم لتجزئة بيانات شبكة دلالات الألفاظ ووضعها في مجموعة من الإطارات البسيطة.

د- أسلوب السيناريوهات[3] : ويستخدم لوصف المعرفة المتغيرة الناتجة من تعرض الإنسان لمواقف تحتوي على مجموعة من المواقف والأحداث غير متوقعة.

أما التجربة العربية المتواضعة في بناء المعاجم الإلكترونية المرتكزة على قواعد المعرفة على تجربتي نبيل علي[4] المتميزة، ثم عبد الرحمان الحاج صالح في الجزائر التي انصبت في بناء مشروع الذخيرة اللغوية[5] والذي كان هدفه انجاز بنك آلي للغة العربية، أو ما سمي ببنك النصوص ومعجم آلي للغة العربية (معجم المفردات) كما كان هدفه يروم بناء مشروع يهدف إلى توثيق موضوعي للمعاجم العربية والدراسات اللغوية،بالرصد الشامل والدقيق العربية، مفردات ومصطلحات وتصفح ومعاني الكلمات من خلال سياقاتها وتحديد تواترها فيه، ومعاجم ألفاظ الحضارة قديما وحديثا ومعاجم للطفل. لكن هذا المعجم لم يرى النور لحد الآن، نظرا للصعوبات التقنية والفنية واللسانية التي يلقاها. أما نموذج نبيل علي[6] الذي أغنى الخزانة العربية بكتب قيمة في اللسانيات الحاسوبية وثقافة المعلومات[7]، وأما ما يسمى لدى هذا الرجل بمكينة المعجم فلابد أن تخضع لثنائية التحليل والتركيب[1]، وشأنها شأن المكونات اللغوية الأخرى. والمعجم الآلي بالنسبة إليه يشمل المعطيات الصرفية والنحوية والدلالية. كما أنه يعتبر مصدراً

---

1 – Rules

2 – Frams

3 – Scripts

4- للمزيد من التفصيل في تجربة نبيل علي، ينظر الباب الأول والثالث من الأطروحة : ص 266،181،49، و للمزيد من بسط القول في هذا في هذا المبحث ، ينظر : اللغة العربية والحاسوب، دراسة بحثية لدكتور نبيل علي ، تقديم: د. أسامة الخولي، دار تعريب، الكويت،1988 ص569،457.

5 انظر، عبد الرحمان الحاج صالح (مدير مركز العلوم لترقية اللغة العربية، جامعة الجزائر)، ورقة حول مشروع الذخيرة اللغوية، مجلة اللسان العربي، ع. 45.

6- يعتبر أول مهندس لغوي عربي، قام بأبحاث رائدة عربيا فيما يتصل للمنظومة اللغوية العربية المؤسسة على قواعد المعرفة.

7- وفي مقدمتها كتابيه "اللغة العربية والحاسوب"، و "الثقافة العربية في عصر المعلومات" ، بالإضافة إلى العديد من المقالات والبحوث المهمة التي قدمها في مؤتمر وندوات علمية عربيا ودوليا.

1- نبيل علي، اللغة العربية والحاسوب، دار التعريب والنشر الكويت (1988).

أساسيا لا غنى عنه في دراسة الإنتاجية الصرفية فقد ارتأى أن ميكنة المعجم في أربع مستويات مترابطة ومتكاملة فيها بينها[1]:

1/ قائمة المفردات المتداولة في اللغة.

2/ مصفوفات وجداول بالعلاقات النحوية والصرفية والدلالية.

3/ قاعدة بيانات خاصة.

4/ قاعدة معرفة متكاملة.

2-2- المعجم الآلي التفسيري التأليفي (DiCo)[2]:

قبل الحديث عن هذه المقاربة ، تطرق الدكتور عز الدين غازي في أطروحته إلى التعريف ببعض المفاهيم التي تبدو لنا ضرورية في تصور بناء المعجم التفسيري التأليفي الآلي للغة العربية كما بين باقتضاء شديد[3] – لأن كل هذه المفاهيم تحتاج أبحاث مستقلة، مما لا تتسع القيام الآن – كل هذه الفرضيات التي تنطلق منها مليتشوك 1995 في مستندات العُجمة[4] الدلالية والمعيارية الصورية في تحديد مفاهيم أضحت شائكة في معالجة المكون الدلالي[5] والمعجمي. وقد نجحت نظرية المعجم التفسيري التأليفي إلى حد ما في صياغة هذه القواعد المؤطرة للغات الطبيعية بما فيها لغتنا العربية[6]. كما أسس الدكتور عز الدين غازي لمعيار رئيسي سمي "بمعيار التوارد بالمصاحبة المنسجم" والذي أفاد كثيرا في معالجة المركبات والمتلازمات في مجموعة

1– عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص176.

2– يعتبر قاموسDiCoالآلي قاعدة معجمية للغة الفرنسية طور خلال أعوام مضت في مختبر اللسانيات معنى – – نص التابع لجامعة مونريال تحت إشراف اكور ميلتشوك ألان بولكير، غاية هذه القاعة هي وصف العجمة الظاهرة في مدونة القاموس بمقتضى محورين: الاشتقاقات الدلالية التي تربط باقي العجمات بعضها بعض والمتلازمات (المسكوكة جزئيا) التي تراقبها. ويعتبر الوسم الدلالي السمة الأساسية التي بني من أجلها القاموس والتي هي في نفس الآن قيمة وصفية في المعجمية الصورية وخاصة ما يفيد تزويد معاجم أنظمة معالجة اللغات الطبيعية.(مالتشكوك 97.)، ينظر: عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، أطروحة جامعية لنيل الدكتوراه، ص 176 ، وينظر : الدراسات المعجمية: مجلة تصدر عن الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص81.

3– عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، أطروحة جامعية لنيل الدكتوراه، إشراف الأستاذ الدكتور: عبد الغني أبو العزم، والأستاذ الدكتور: محمد لهلال، كلية الآداب والعلوم الإنسانية، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجميات، جامعة الحسن الثاني، عين الشق– الدار البيضاء، السنة الجامعية:2008/2007، ص176.

4 – Lexie

5– جرهارد هلبش، تاريخ علم اللغة الحديث، ترجمه و علق عليه د/ سعيد حسن بحيري، مكتبة زهراء الشرق، القاهرة، الطبعة الأولى، 2003 مـــ ، ص 531 وما بعدها.

6– يقول الدكتور عز الدين غازي في هذه النقطة : كنت في اتصال شخصي مع العالم ايغور ملتشوك وأكد لي مرارا أهمية هذه المبادئ في مقاربة لغتنا العربية بهذا المنظور ورغم أنني لم أعتمد في بحثي استراتيجية هذه المقاربة فقد آثرت على تقديمها للإفادة في التفكير في بناء معجم عربي آلي مبنين وفق المبادئ الأساسية المسطرة في النظرية المذكورة، ينظر: عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، أطروحة جامعية لنيل الدكتوراه، ص 176.

من اللغات. كما سنعرف بالقاموس التأليفي الآلي (**DEC**) وقاموس التواردات (**Dico**)[1] والبرنامج المعلومياتي (**Edico**) و(**DEC**)[2] و(**Dico**) أساسيات المعطيات المعجمية لمعالة اللغات الطبيعية وسنحاول التركيز على بعض المبادئ لمقاربة اللغة العربية آليا وفق هذا المنظور[3].

وتعتبر المعجمية التفسيرية التأليفية نتاج نظرية معنى- نص والتي أرسى دعائمها الغوي الروسي الأصل أيغور ميلتشوك (Igor MELCUK)[4]، فهي واحدة من النظريات المعمية المعاصرة التي تسعى إلى معالجة المعجمية والقاموسية والصناعة المعجمية في آن واحد، إذ تهدف إلى ملامسة كل ما يتم بصلة إلى هذه العلوم الثلاثة بحكم ترابطها وتداخلها، فهي من المحاولات الرائدة التي تندرج في مال المعجمية الواصفة[5] والقاموسية الواصفة[6].

انطلاقا من الترابط النسقي، تهدف النظرية معجمية التفسيرية التأليفية إلى وصف وتحليل الظواهر المعجمية والدلالية بطريقة دقيقة ومنظمة معتمدة على تمثلات رياضية صورية. غير أن ما يجعل نظرية معنى-نص أكثر جرأة من سابقتها من النظريات اللسانية البنيوية والتوليدية هو قيامها على فرضية الانطلاق من التمثيلات الدلالية من أجل بناء قواعد تركيبية بمساعدة المعجم[7]، فالمعجم بحسب هذه النظرية سابق على التركيب والصَّرف والصِّواتة[8] . إن ميلتشوك لم يكن له بلوغ الهدف من وراء هذا التصور لولا اعتماده على المنطق والصورنة[1] والدقة في تحديد العُجمات وتعريفها. وانسجاما مع هذا التصور اقترح ملتشكوك وزملاءه إحدى عشر قاعدة تندرج القواعد الست الأولى ضمن ما يسمى **بميكروبنية** القاموس التفسيري التأليفي، أي

1- يتأسس هذا القاموس على وحدات معجمية تمزج بين الوحدات المألوفة في الصناعة المعجمية وتلك التي ذات طبيعة معقدة، بمعنى آخر، الأقل معجمية في البنيات التي تحسم في تركيبا وليس بحسب للقيم النحوية فقط. للمزيد ينظر، A. Polguere, typologie des entités lexical d' une base de données explicative et combinatoire, OLSTm Univ de Montreal.

2- يمكن الاطلاع أكثر على المعجم التفسيري التأليفي للغة الفرنسية لدي إيغور ميلتشوك وزملاءه بصفة خاصة ملتشوك (1984) (1988) (1992) (1999) .

3- عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص177، وينظر: الدراسات المعجمية: مجلة تصدر عن الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص82.

4- شرع هذا العالم في صياغة نظرية معنى- نص منذ أن كان عضوا باحثا في أكاديمية العلوم بموسكو حيث أصدر العديد من المؤلفات والمقالات منذ السبعينات، لينتقل سنة 1978 إلى جامعة مونريال بكندا حيث تابع مسيرته العلمية فقام اثر ذلك بتطوير مقاربة معنى- نص بمعية فريق من الباحثين المرموقين مثل : ألان بولغير واندي كلاس والكسندر زلكوفكي.وآخرون وقد ساعده على ذلك اتقانه للغات عديدة منها الروسية الفرنسية والانجليزية واطلاعه على أهم المصادر اللسانية والإلمام بخبايا الظواهر الغوية ورصد خصائص اللغات الطبيعية على بلورة هذه المقاربة. وهو الآن أستاذ زائر بالكويلج دو فرانس وعضو بارز في مجموعة البحث في الدلالة المعجمية والمصطلحية. وهذا الرجل له العديد من الأعمال العلمية أهمها 84، 88، 92،95،97، 99،01.

5 – méta-lexicologie

6 – méta-lexicographie

7- انظر محند الركيك ، قراءة جديدة في المعجمية العربية، نحو مقاربة معجمية تفسيرية تأليفية ، أطروحة لنيل الدكتوراه في اللسانيات العربية ، فاس 1 ،2000، ص، 172.

8- عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص177،

1- محمد الراضي، محلل تركيبي للغة العربية، التوليد والنسقية والترجم الآلية، عبد القادر الفاسي الفهري، المجلد 1، منشورات معهد الدراسات والأبحاث للتعريب، الرباط ، المغرب ، 2001، ص 45،44.

تنظيم مادة القاموس بطريقة نسقية[1] ، بمعنى التنظيم الداخلي لهذه المادة. في حين تندرج القواعد الخمس الأخرى ضمن **ماكروبنية** القاموس التفسيري التأليفي التي تعنى بتنظيم مجموع القاموس بكامله انطلاقا من الحقول الدلالية التي تمثلها الفوكابلات (vocables).[2] وبناء على هذا ميز ملتشوك بين أربعة جوانب هي[3]:

المعرف الخاص والمعرف به الخاص ومجموع المعرفات به داخل القاموس التفسيري التأليفي والعلاقة بين المعرف به والمعرف الخاصين. ولما كان لهذا القاموس مرجعية ومنطقية رياضية فقد وضع صياغة صورية تمثلت في القواعد التالية[4]:

2-2-1- قاعدة الأشكال الافتراضية : (formes propositionnelles):

للترابط المنطقي دور في تحديد العوامل الدلالية لذلك التجأ ملتشوك إلى إدراجها في المعرف به[5] وكذلك المعرف[6] الذي يضم أكثر من عجْمة واحدة، فهذه الأخير ة هي متغيرة للعوامل الدلالية، فهي وإن كانت مسندا ، فإنها تقوم بالتعريف نحو:

انتقل الموظف من وزارة 1 إلى وزارة 2، يتم تحديد كل ما يتعلق بالفعل "انتقل" كما في الصوغ التالي[7]: انتقل س من [أ إلى ب] / انتقل س إلى [جوار ربه].

إن العُجْمات التي يكون معناها مسندا دلاليا[1]هي التي تستجيب للأشكال الافتراضية والتي تسمى عادة بالمعرفات الممثلة بالتعبيرات[2] وبناء على هذا يصوغ ملتشوك الشكل الافتراضي التالي: " إذا تعلق الأمر بتعريف العجمة **(L)** المسند فينبغي على المعرف أن يتخذ شكلا افتراضيا حيث تمثل المتغيرات عوامل **(L)** الدلالية. إن الشكل الافتراضي في الحقيقة هو

---

1 – محند الركيك ، قراءة جديدة في المعجمية العربية، نحو مقاربة معجمية تفسيرية تأليفية، أطروحة لنيل الدكتوراه في اللسانيات العربية مرجع سابق، ص173.

2– حاولت الباحثة المغربية سامية سومية الزاهد إعداد خمس حقول دلالية للفوكابلات في اللغة العربية كما اقترحها إلكنتسي وهي في الحقيقة بنية المادة المعجمية للقاموس التفسيري التأليفي (منطقة المدخل والمنطقة الدلالية للفوكابل والمنطقة التأليفية التركيبية والمنطقة التأليفية المعجمية زمنطقة التوارد بالمصاحبة).

3– عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص178، وينظر : الدراسات المعجمية: مجلة تصدر عن الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص83.

4– أنظر :ايغور ملتشوك 1995 Melc'uk Igor,introduction a la lexicologie explicative et combinatoire,1995

5 – le définissant

6 – le défini

7– عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص178، وينظر : الدراسات المعجمية: مجلة تصدر عن الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص84.

1– نشير هنا إلى أن المسند الدلالي عند ملتشوك هو المفهوم الإجرائي الذي إعتمده موريس كروس وهو يختلف عن مفهوم الاسناد لدى النحاة العرب القدامى ، فهو هنا يدل على الحالات والاحداث والعلاقات والافعال والخصائص

2 – expressions

معرف معجمي تقسمه العُجمة الرئيسية[1] المدمجة في الشكل الافتراضي (المتغير) والمعرف به الذي يصف المعنى وهو جهاز في اللغة الواصفة للقاموس التفسيري التأليفي.

2-2-2- قاعدة التجزيء: (règle de composition):

تجاوزا للعطب الابستيمي الذي حصل مع المعجم التقليدية، اقترح ملتشوك قاعدة التجزيء التي تقوم بتحليل العجمة الرئيسية بالصوغ الصوري التالي[2]:

كل عُجْمة رئيسية(L1) تعرف بالعُجمات الثانويةL1+ L2 + L3 + L4 +...

Ln حيث العجمات الفرعية أسهل دلاليا من العجمة الرئيسية L ...

طبقا لذلك فتعريف السيارة مثلا هو: (عربة + محرك + عجلات...) فكل مكونات التعريف تندرج دلاليا، مما يجعلنا نتجنب التعريف الحلزوني الدائري[3] الذي لا يكترث بدور البساطة والتعقيد في وضع العُجمات. غير أن هذه المنهجية لا تتضمن في الترادف إذ بعضها يعرف بكلمة واحد فقط نحو: (والدة = أم) . نفس الشأن بالنسبة للعُجمات ذات المعاني المتعددة والمتنافرة دلاليا.

2-2-3- قاعدة الكتلة الكبرى:(règle de bloc maximal):

على عكس القاعدة السالفة يعتمد القاموس التفسيري التأليفي قاعدة الكتلة الكبرى، أو ما تسميه ويرزبيكا (**Anna Wiezbicka**) بالأولويات الدلالية[4]. فكل عُجمة رئيسية مُصاغة بعجمات فرعية **L1+L2+L3+.......+Ln**، إذ لا يمكن تعويضها دلاليا بعجمة واحدة، أي عنصر من العناصر المكونة للعجمة الرئيسية. فالمحرك لا يمكن أن يعرف السيارة لأنه قد يكون جزءا من آلات أخرى. وهذا التشكيل يمكن أن يحلّ إشكال وصف جملة من بيانات التي لايمكن تفكيكها وتحليلها إلى أجزاء صغرى من قبيل التعبيرات المتلازمة والمركبات العربية مثل (لقي زيد حتفه)لا يمكن فهمها انطلاقا من عناصرها بل اعتمادا على الاكتساب والفهم الكلي لها.

2-2-4- قاعدة التعبير: (règle de la standardisation):

لكل العُجْمات لغةٌ تصاغ بها وفق ضوابط مقيدة ومحدودة، فصياغة التعريف ينبغي أن يتفادى فيه كل الغموض الذي يطال المعرف به. هذا يعني أن القاموس التفسيري التأليفي استعان باللغة الواصفة التي عادة ما تقوم بإجراءات تقنية للترميز والتشفير، مثل ترقيم العُجْمات كل واحدة على

---

1 – lexie vedette

2- عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص179، وينظر : الدراسات المعجمية: مجلة تصدر عن الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص85.

3- أنظر :ايغور ملتشوك،1995، مرجع سابق Melc'uk Igor,introduction a la lexicologie explicative et combinatoire,1995

4- انظر محند الركيك ، قراءة جديدة في المعجمية العربية، نحو مقاربة معجمية تفسيرية تأليفية ، أطروحة لنيل الدكتوراه في اللسانيات العربية ، فاس 1 ،2000 ،ص 181

حدة حتى لا يختلط الأمر. نحو((ضرب 1 زيد عمراً) و(ضرب 2 في الارض)...)[1]. غير أن هذه العملية لا تصلح دائما، خاصة عند التعريف المترادف لأن كل معنى تقابله عُجْمة واحدة وهذا ما يتعذر فعله هنا عند البحث عن معاني مطابقة ودالة على أشكال مختلفة. فالتوارد المعجمي بالمصاحبة لا يكون بالضرورة قابلا للتطبيق على لغة التعريف.

2-2-5- قاعدة التعويض: (règle substituabilité):

وإذا كان الوصف التعبيري لا يساوي دائما بين المعرف والمعرف به، فإن التعويض يحل أمر هذه المسألة، يمكن لعُجْمة أن تعوض أخرى دلالياً[2] (عُجمة أ =عُجمة ب). فإذا عرفنا أستاذ مثلا لا يعني أنه يدرس في الجامعة أو المدرسة بل قد يكون محاميا أو قاضيا. فالتكافؤ الدلالي نسبي بحيث يصعب تطبيقه في السياقات الصعبة التي يلجأ عادة إلى فحص العناصر الأساسية فيها.

2-2-6-قاعدة الترتيب ذات الانسجام الدلالي:(règle de pertinence sémantique) :

تقوم عملية الترتيب بناء على الانسجام الدلالي، أي بين القرب والبعد الدلالي للعناصر فيما بينهما، ومن هنا فإن كل خاصية أكثر انسجاما على المستوى الدلالي، يشترط فيها أن تكون ذات أولوية في الترتيب المنطقي الدلالي للعجمة، مثل (اللبن مغذي وصحي) فالصفة مغذي أكثر انسجاما من الصفة صحي. فتبعا لأسبقية مكون على آخر يتم اعتماد هذه القاعدة التي لا تصلح إلا نسبيا لأنها تستدعي تعاريف مركزة ودالة مفهومة ومتلاحمة[3].

2-2-7- قاعدة التضمين الدلالي:(règle connotation sémantique) :

ارتباطا بالقاعدة السالفة، فإن قاعدة التضمين كما يشير معناها تعمل على تضمين عجمة(1) ذات التعريف المضمن في العجمة (2) عملا بمبدأ الأسبقية. فمثلا (عين (عضو الابصار) سابقة عن عين (ينبوع الماء) فمعنى الأول يجب يكون مضمنا في المعنى عين الثاني . ويمكن تطبيق هذه القاعدة بسائر مستويات اللغوية الصرفية والنحوية (اللزوم والتعدي مثلا) والدلالية والتداولية[1].

1- عز الدين غازي، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص179، وينظر : الدراسات المعجمية: مجلة تصدر عن الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص85.

2- محند الركيك، مرجع سابق، قراءة جديدة في المعجمية العربية، نحو مقاربة معجمية تفسيرية تأليفية ، أطروحة لنيل الدكتوراه في اللسانيات العربية ، فاس 1 ،2000 ، ص 194.

3- عز الدين غازي، بناء المعجم الآلي للغة العربية، مرجع سابق، ص180، وينظر : الدراسات المعجمية: مجلة تصدر عن الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص87،86.

1- المرجع السابق، ص 180.

2-2-8- قاعدة التقارب الدلالي: (règle de rapprochement sémantique) :

إن درجة القرب والبعد في دلالة العُجمات هي لا تحدد العلاقة بين العُجمة (1) والعجمة (2) والعجمة(3). فالمعنى في الفعل "بلغ" (بلغ زيد إلى الأعالي) ليس هو المعنى نفسه في (بلغ زيد غايته) وليس هو كذلك في (بلغ السيل الزّبى) فرتبة بلغ (1) التي تقابل العُجمة (1) سابقة عن رتبة بلغ (2) التي تقابل العُجمة (2) وهكذا، لذلك فالمعنى الوارد في الجملة (1) أقرب دلالياً للمعنى (2)، التي هي الأقرب بالتالي إلى المعنى (3)[1].. وما يمكن ملاحظته هو أن نوع المجاز المرسل قد يبدو هنا أكثر قربا من الاستعارة وهذا ما يعبر عنه إزوالد ديكرو (O.Ducrot) في أعماله بالمستوى التمجيزي للغة الطبيعية، وكذلك مايسميه ألان بولغير بالشبكة المعجمية للغة[2].

2-2-9- قاعدة أولوية المجاز المرسل:

خاصية التقارب الدلالي، تحدد المعاني المجازية للعجمات، فالمجاز المرسل يسبق رتبة الاستعارة. وقد أورد السكاكي في مفتاح العلوم[3] مثلا في هذا الصدد حينما تطرق لمفهوم الاستعارة التي تعني ذكر أحد طرفي التشبيه وتريد به الطرف الآخر مدعيا دخول المشبه في جنس المشبه به، كما تقول في الحمام الأسد وأنت تريد بها الشجاع مدعيا أنه من جنس أسود فتثبت للشجاع ما يخص المشبه به وهو إسم جنسه مع سد طريق التشبيه به بافراده في الذكر وكما تقول"أن المنية أَنْشَبَتْ أَظْفَارِهَا" وأنت تريد بالمنية السبع..." فعلاقة المشابهة بين الحقيقي والمجاز هي التي ترقى باللفظ بواسطة قرينة أو قرائن إما علاقة كلية أو جزئية أو استعارة ذلك ما يطلق عليه إيغور مالتشوك بالمجاورة الذهنية[4]. وإذا عمقنا النظر في المتداول اليومي، سنجد أن اللغة الطبيعية تمتد بنياتها من الحقيقي البسيط إلى المجاز المركب فالاستعاري[5]. ولما كانت اللغة مؤسسة إجتماعية تعبر على المستوى الحضاري والثقافي للشعوب كما تعبر عن سيكولوجية الفرد وعلاقاته بالواقع والطبيعة فقد كان لابد من أن نجد هذا الزخم الكبير من التعبيرات المتلازمة والمأثورة والمسكوكة في اللغة وما يلفت الانتباه هنا أن المعاجم اللغة العربية القديمة التي كانت تهتم بتعجيم الظواهر البلاغية مثل أساس البلاغة للزمخشري ولسان العرب لإبن منظور قد اهتمت بالمداخل المتلازمة والمركبة أكثر من غيرها من القواميس والمعاجم العربية في شتى مراحلها[6].

1- هذه المسائل بسّط فيها علماء البلاغة القول كثيرا ولا يتسع المقام للتطرق إلى ذلك بالتفصيل.

2 -A.Polguere , typologie des entités lexicales d' une base de données explicative et combinatoire, in OLST,Univ de Montreal.p.3 .2007.

3- انظر السكاكي ابي بكر محمد، مفتاح العلوم، مكتبة التعليمية الجديدة لبنان ص.174.

4- انظر محمد الركيك، قراءة الجديدة في المعجمية العربية، نحو مقاربة معجمية تفسيرية تأليفية، أطروحة لنيل الدكتوراه في اللسانيات فاس 1 س2000 ص.214.

5- عز الدين غازي، بناء المعجم الآلي للغة العربية، مرجع سابق، ص181، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص88،87.

6- عز الدين غازي ،مرجع سابق، ص181، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص88.

2-2-10- قاعدة المكونات الدلالية :(règle de composantes sémantiques) :

تتميز هذه القاعدة بكونها تربط بين المجازي والاستعاري من جهة بين المكونات كعجم الفرعية والمكونات الكبرى كعجمات رئيسة، فالمرور من عجمة1 إلى عجمة2 لابد من ولوج القنطرة الدلالية[1]. بواسطة مشترك ظاهر بين العجمتين. فإذا عرفنا مفردة (الفوكابل) "اليد" "اليد" سنجدها قنطرة دلالية بين "اليد1" كعضو طبيعي في الجسم و"اليد2" كآلة معينة (يد الفأس). وتمثل "اليد2" المكون التحتي بالنقل الاستعاري في حالة ورودها في تعبير مجازي أو تعبير متلازم، فترقى إلى تعبيرات نحو يد الله مع الجماعة (يد النعمة والاحسان). هذا التسلسل في تضمين[2] يسعى إلى إبراز الانسجام الدلالي وفق أولوية والأسبقية ويعتبر المستوى المعجمي في نظرية المعنى-النص مستوى مركزي ومحور باق المستويات اللسانية (الدلالة والتركيب والصرافة والصواتة بالترتيب)[3].

2-2-11- قاعدة السلوك التخصيصي للعُجمَات:

إن كل عجمة فرعية ذات تخصيص معين[4] ضمن الحقل المعجمي ككل، ويعرف مالتشوك لكون "عجمة1 تخصيصية بالنسبة لحقلها المعجمي وهي أسبق من العجمة2فمثلا الكلب تخصيصة للحقل المعجمي" الحيوانات الأليفة بينما قدم ليست تخصيصة للحقل المعجمي "أعضاء الجسم" لأنها تستعمل مجازيا في التعبير التالي"لزيد قدم في الكرم"

تعتبر القواعد من سبعة إلى إحدى عشر بمثابة ماكروبنية القاموس التفسيري التأليفي الذي يعني بترتيب العُجْمات ترتيبا منطقياً مبنياً على مبدأ الأسبقية والأولوية[5].

2-2-12- قاعدة المدخل الرئيسي (الفوكابل) : (Règle de l'entrée principale vocable)

لكل قاموس ومعجم ومادته الرئيسية (الفوكابل) التي تعرف بالنص المعجمي[6]، فهي عبارة عن مجموعة من العجمات(عجمة1+عجمة2+عجمة3+...+عجمة ن) التي تستجيب لشرطين هما[7]:

---

1 – Pont sémantique

2- انظر القاعدة السادسة من هذا المبحث : قاعدة الترتيب ذات الانسجام الدلالي، ص252 .

3- عز الدين غازي، مرجع سابق، ص182، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص89.

4 – être typique

5- عز الدين غازي، مرجع سابق، ص182، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص89.

6- النص المعجمي هو ركيزة المعجم، وهو مثله مثل النص الأدبي والعلمي و كان د. محمد رشاد الحمزاوي أول من نبه إلى قضاياه في العربية ومن اقترح نماذج منه مشيرا إإى الموسوعي منه والموسوعي الفوضوي والمتداخل، والضحل سواء في "فصيحه" ومعربه ودخيله. للمزيد من بسط القول، ينظر : محمد رشاد الحمزاوي، النص المعجمي في المولدات والأعجميات، مجلة المعجمية، عدد11،1995، ص19/9، وينظر : محمد رشاد الحمزاوي، مقترح لوضع نموذج للمعجم العربي الحديث، المعاجم العربية الواقع والآفاق، الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية،العدد السادس،1428هـ/2007م، ص74.

7- عز الدين غازي، مرجع سابق، ص182، ص 182.

*/الدوال المتطابقة ل(عجمة1،عجمة2،.....،عجمة ن).

*/المدلولات المكونة من عجمات (عجمة1،عجمة2،...،عجمة ن) المرتبطة فيما بينها.

تبعا لذلك فالمادة الرئيسية يقابلها المترادف حيث تجمع العجمات المتفرعة عنه[1]. وتفرز العجمة الأساس[2] من بين العجمات الفرعية للمدخل الرئيسي. نورد مثالا للتوضيح كما يلي (فالكتاب بمعنى المؤلف والعمل هي بمثابة عجمات أساسية ولذلك فالقول: كتاب الله وأهل الكتاب وكتاب القرآن كلها عجمات أساسية للكتاب كمدخل رئيسي في المعجم[3]. وكل هذا يعمل وفق تقارب الدلالي للمعانى المشتقة من نفس الجدر الواحد. وقد لاحظ الدكتور عز الدين غازي في تصوره للقاموس التفسيري التأليفي، هو اعتبار المداخل على النحو لا خطية[4]. لأنها تتحكمها العلاقات الدلالية المتشعبة تراتبيا، فمن خلال هذه الرؤية وجد الدكتور غازي أنه لا بد من مدخل الرئيسي الفوكابل وهو ما يسمى بالمسافة الدلالية[5]. وكل مستوى لهذه المسافة يرمز له برقم مميز في المعجم تحسبا للترادف الدلالي فالعلاقة التي تربط عجمتين هي دلالية ملازمة لخاصية تركيبية متمثلة في بنية التعبير، وهذا ما يطلق بالتوارد المعجمي بالمصاحبة[6] فكل مسافة يسند إليها نوع من أنواع الأرقام المعروفة لكي تتميز عن بعضها في المسافة الدلالية[7]. فمدخل "ضرب" تدلنا على المسافة الحقيقية بين[8] :

1. ضرب زيد عمرا.3. ضرب القلب.2. ضرب في الماء.

4. ضرب الليل بظلامه.5. ضرب في الأرض

ولهذا المدخل المعجمي حوالي 45 معنىً تفصلها مسافات دلالية تتصاعد من الحقيقي إلى المجازي بتراتبية وأولوية معينة حتى تصل إلى المراتب الإستعارية فنقول "ضرب زيد في الأرض" و" ضرب أخماسا في الأسداس" ولرصد هذا التدرج الدلالي يصوغ إيغو ملتشوك الشكل التالي[9]:

1- المرجع السابق، الصفحة نفسها.

2 – Lexie de base

3- عز الدين غازي ،مرجع سابق، ص182،

4- المرجع نفسه، الصفحة نفسها.

5- المرجع نفسه، ص183.

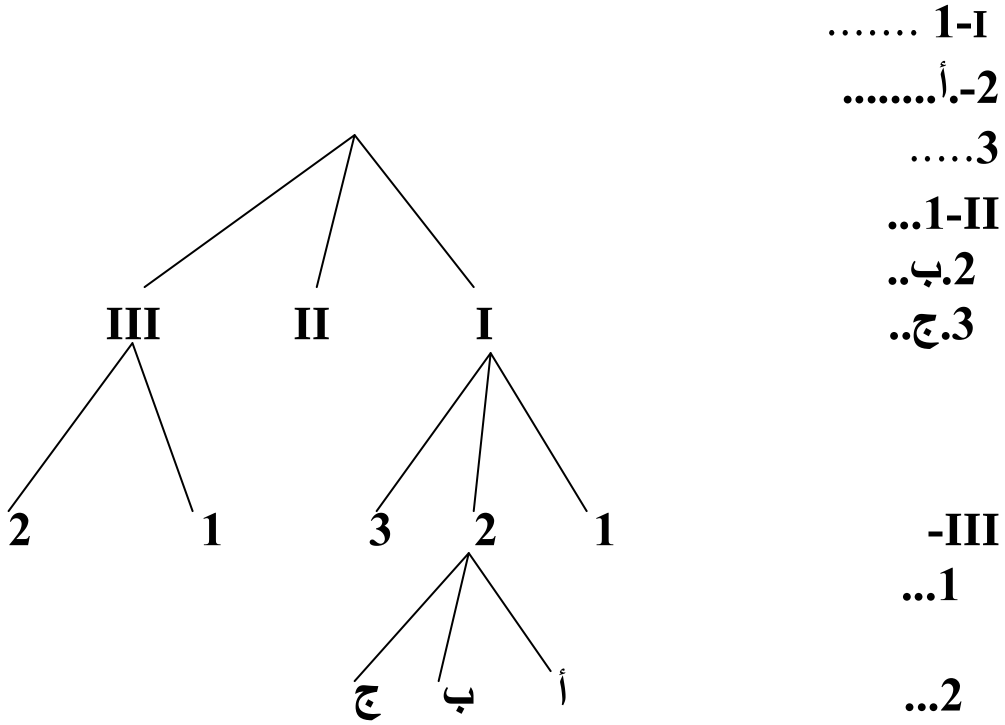
6 – cooccurrence lexicale

7- عز الدين غازي ، مرجع سابق، ص183.

8- المرجع نفسه، الصفحة نفسها.

9- المرجع السابق، الموضع نفسه.

المدخل الرئيسي (الفوكابل)



الشكل رقم :2 : يمثل ترميز المسافات الدلالية القائمة بين العجمات.

ويشير الدكتور عز الدين غازي في هذا الصدد، أن ملتشوك اعتمد معايير في المكون الدلالي إذا كانت ملتبسة بأن تحتمل تأويلا دلاليا متعددا ومناسبا كمعيار التوارد بالمصاحبة المنسجم، ذلك أن المكونات الدلالية تنفصل وتفترق داخل تعريف معجمي واحد، كما يقول الدارس الروسي أبريسان (Apresjan)[1] .وعليه فعندما طبق الدكتور عز الدين غازي نظرية معيار التوارد بالمصاحبة المنسجم على التعبيرات المتلازمة في اللغة العربية من خلال المثال التالي: فعل "ضرب" وجده ذو ترادفات عديدة منها[2]:

1- ضرب موعدا مع مريم .

2- ضرب السكة باسمه

3- ضرب أخماسا في أسداس

عندما يحصل الاختلاف الدلالي بين التعبيرات الثلاث المذكورة في 1و2و 3 لا يمكننا أن نؤلف جملة ذات توارد بالمصاحبة المنسجم نظرا للتنافر الدلالي وتشتته في المدخل الرئيسي"ضرب".

1- للمزيد من بسط القول ينظر: محند الركيك، سابق، 2000، ص.258.

2- عز الدين غازي ، مرجع سابق، ص184، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص91،92.

ويقول الدكتور في أطروحته[1] أن إيغور ملتشوك حاول بناء على هذه الفرضيات، حوسبة القاموس التفسيري التأليفي (DEC) وراء إثر ذلك أن الاشكالية تحيل إلى إعدادالأدوات اللسانية المعجمية التالية[2]:

*/حوسبة المصادر اللسانية واستغلال المدونة المحوسبة لإعداد قاموس تفسيري تأليفي إلكتروني والاعتماد على برامج إحصائية للمعالجة الأوتوماتيكية للغرض نفسه.

*/ترجمة المعلومة التي يتضمنها القاموس التفسيري التأليفي إلى لغة المعلوميات، واتخاذها شكل قواعد المعطيات المحوسبة ، على شكل جذاذيات حاسوبية مبنية بطريقة تسمح بتدبير أوتوماتيكي للمعلومة التي تتضمنها هذه الجذاذيات .

*/اعتماد الحوسبة في العمل اللغوي من شأنه أن يؤدي إلى إعداد القواميس لأهداف متعددة ومنها الترجمة الآلية وتطبيقات معالجة اللغات الطبيعية.

تعود أول محاولة لمعالجة بيانيا القاموس التفسيري التأليفي إلى سنة 1986، حينما وضع ميشال ديكاري[3] (Michel Decari) . أسس هذا المشروع ليطوره إيغور ملتشوك وزملاءه[4] بعد ذلك إذ اقترحوا صيغة مبسطة ومختزلة تمثلت في قاموس(DEC) الذي يصف الظواهر اللغوية التي ترد مصاحبة في القاموس المخصص للتواردات بالمصاحبة(Dico)[5] .

إن حوسبة قاموس التواردات (Dico) هو حوسبة أولية لمواد قاموس اللغة الطبيعية، وللقيام بذلك تم بناء قاموس (Edico)[6] اختزالا لناشر قاموس التواردات الذي يعتبر بمثابة نماذج تطبيقية لحوسبة المعجم. ولتجاوز ما قد يطرحه هذا القاموس ثم ابتكار برنامج معلوماتي منسجم كليا مع المبادئ العامة للقاموس التفسيري التأليفي (DEC)[7]. الذي اختزل إلى (Dico) الذي يعمل وفق تصور رياضي ومنطقي دقيق، لأنه يستطيع تفكيك قضاياه المنطقية فيجعلها مبسطة وسهلة وقد استفاد هذا القاموس بصفة عامة كما استفادت جميع المعاجم الإلكترونية في العالم من الأدوات العقلانية للمعطيات النظرية والهندسة الإلكترونية[8].

1- عز الدين غازي ،مرجع سابق، ص184.

2- Igor MELCUKE et al, introduction à la lexicologie explicative et combinatoire ,1995 .p.206

3- انظر محند الركيك/ قراءة جديدة في المعجمية الجديدة نحو مقاربة معجمية تفسيرية تأليفية، أطروحة لنيل الدكتوراه في اللسانيات العربية، فاس، س.2000.ص . 333.

4- انظر، Igor MELCUKE et al 1995.pp. 205-239

5- نظرا لما تطرحه الدلالة من صعوبات في الترجمة فقد فضل قاموس التوارد بالمصاحبة عن غيره لأنه لايتوفر على تعاريف معجمية مما يجعله أكثر قابلية للصورنة وأكثر مرونة من غيره لعمليات الحوسبة. وعلى الرغم من كونه يحاول تقديم الأوصاف الكاملة والشاملة للمداخل المعجمية فإنه يبقى في مرماه مبسطا ومختزلا. للتفصيل ينظر : عز الدين غازي ،مرجع سابق، ص،185،184، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص93،92.

6 - Editeur du dictionnaire de coovvurrences.

7- عز الدين غازي ،مرجع سابق، ص185.

8- المرجع نفسه، الصفحة نفسها.

وأهم ما يتعرض العمل المعجمي هو خصوصية النظام اللغوي وطبيعته من حيث الاطراد والحشو[1].

1- التحشية: annotation : إضافة معلومات في داخل النص ، مثل االوسم tagging. وهناك أنواع مختلفة للوسم ، منها:

(1) قسم الكلام: (اسم ، فعل ، حرف). ويجدر ملاحظة أن التقسيم العربي التقليدي يجتاج إلى مراجعة وتفصيل أكثر عند وسم الكلمات ، حتى يكون التحليل أكثر فائدة ودقة. (انظر: محمود إسماعيل الصيني "نحو معجم عربي للتطبيقات الحاسوبية" وانظر تطبيقا عمليا لنوع جديد من التقسيم في المعجم الذي أعده بكوولتر وباركنسون Buckwalter and Parkinson ، وهما من رواد العمل في مجال مدونات اللغة العربية في الولايات المتحدة الأمريكية .

يجدر بالذكر أن كتب النحو العربي تزخر بمصطلحات تتعلق بتصنيف الكلمات العربية يمكننا الاستفادة منها ، من مثل : ضمير، اسم موصول ، اسم جامد ، اسم مشتق(اسم فاعل ، اسم مفعول ، صيغة مبالغة ، صفة مشبهة ، اسم آلة ، اسم مكان) ، اسم جمع ، اسم نوع ، اسم علم ، مصدر،ظرف (مكان ، زمان) ، فعل (لازم ، متعد لمفعول ، متعد لمفعولين ، ناقص ، ماض ، مضارع ، أمر...) , حرف (جر، نصب، جزم ، عطف ، نفي ...) . ويجب وسم الكلمات بمثل هذه التصنيفات لإجراء كثير من الدراسات المعجمية والنحوية. (انظر الدراسة الممتازة في هذا المجال لمجدي صوالحة. انظر أيضا Kais Dukes and Nizar Habash "Morphological Annotation of Quranic Arabic"

(2) الوسم النحوي أو الإعرابي :(لبيان الوظيفة النحوية / الحالة الإعرابية (مبتدأ، فاعل، مسند إليه ...)

مثالان للوسم الإعرابي:

الجملة : الإيمان نور القلوب

الوسم الإعرابي : (مب (الإيمان) خب (مض (نور) مض إل (القلوب)))

الشرح:الإيمان = مبتدأنور القلوب = خبرنور= مضافالقلوب= مضاف إليه

المستوى الأول : الإيمان (مبتدأ) + نور القلوب (خبر)المستوى الثاني: نور (مضاف) + القلوب (مضاف إليه)

الجملة : الإيمان ينير القلوبالوسم الإعرابي: (مب (الإيمان) خب (فع (ينير) مف (القلوب)))

الإيمان = مبتدأينير القلوب = خبر

ينير = فعلالقلوب = مفعول به

المستوى الأول : الإيمان (مبتدأ) + ينير القلوب (خبر)المستوى الثاني: ينير (فعل) + القلوب (مفعول به)

للمزيد من التفصيل ، يمكن الرجوع إلى الدراسة التي تمت في جامعة ليدز البريطانية:

(انظر: Kais Dukes, Eric Atwell and Abdul-Baqee Sharif, "Syntactic Annotation Guidelines for the Qur'anic Arabic Dependency Treebank") ، كذلك الدراسة التي يقوم بها فريق من الباحثين في جامعة كولومبيا الأمريكية Habash, Nizar, Reem Faraj and Ryan Roth. "Syntactic Annotation in the Columbia Arabic Treebank").

(3) الوسم الدلالي : المحددات الدلالية (مثلا: لكلمة "عين" : البصر، الماء، توكيد).

أساليب الوسم المختلفة:

أما طريقة الوسم فيستحسن اتباع ما يعرف بلغة التعليم المعيارية العامة General Standard Markup Language (GSML). لأمثلة تطبيقية على اللغة العربية .

ومن الأمثلة على أساليب الوسم المختلفة استخدام القوس المثلث < > قبل وبعد الجزء المطلوب تعليمه. مثلا، لبيان أن النص من الشعر الأموي نكتب: < شعر أموي> أبيات من الشعر الأموي <شعر أموي/> ، حيث تشير المعلومات بين القوسين المثلثين إلى أن النص من الشعر الأموي، كما تبين المعلومات بداية النص ونهايته . ومن الأمثلة الأخرى استخدام الأقواس المربعة ، مثلا [S....S] للدلالة على بداية الجملة ونهايتها.

لا بد من التأكيد على ضرورة بيان حدود العبارات التي تزيد عن كلمة واحدة للتعرف على بداية العبارة ونهايتها. وهذا ينطبق على التعبيرات الاصطلاحية والأمثال والأفعال المتعدية بحروف وماشابها، إذا أردنا للحاسوب أن يعيننا في حصرها وإحصائها آليا.

ويمكننا وسم الكلمات بطريقة بسيطة ، مثل : كتب-س ، كتب-ف للتمييز بين الكلمة بوصفها اسما وبوصفها فعلا ، وذلك في غياب برنامج للتحليل الصرفي.

يذكر Bowker and Peason (2002) أن هناك نظاما يسمى معيار تشفير المدونة Corpus Encoding Standard (CES) اتفق عليه الباحثون في أوربا وأمريكا لتعليم النصوص وتحشيتها ، تنقسم إلى ثلاث فئات هي:

(أ)- التوثيق ، ويقصد به الإشارة إلى معلومات ببليوغرفية عن النص ولغته ، مثلا، (ب)- بيانات أساسية: وتشمل معلومات عن مكونات النص، مثل الفقرات والعناوين والهوامش، ألخ، إضافة إلى معلومات عن الرسومات graphics .

(جـ)- تحشية لغوية : وتشمل المعلومات اللغوية داخل النص، مثل سمات الكلمات الخاصة بقسم الكلام part of speech والمعلومات الخاصة بالخطاب والوسم النحوي والوسم الدلالي ...

ولذلك أول ما فكر فيه ميلتشوك وزملاءه هو الاستناد إلى نظرية لسانية واقعية ترتكز على المتكلم الفطري للغة وليس الاعتماد على معيارية اللغة[1].

وبفضل صورية هذا القاموس استطاع أن يلج بقوة معالجة اللغات الطبيعية (TALN) من خلال كونه النموذج القاموسي الصوري المحوسب الذي هيأ المناخ لمعالجة الظواهر التواردات بالمصاحبة[2].

وما يميز هذا القاموس عن غيره من القواميس الإلكترونية كونه مبنيا مومبنينا على التصور النظري لنظرية القاموس التفسيري التأليفي، فهو ذو نظام يقوم بحسابات منطقية صرفة ما يمكن للساني ان يقوم بها. ويرى ملتشوك أن وضع قانون القاموس في جذاذيات حاسوبية من شأنه أن يجعل من العمليات اللغوية عمليات مرنة ومبسطة، استنادا إلى فحصه الدقيق والشامل للجميع التواردات الواردة بقاعدة البيانات المعجمية المؤسس عليها العمل[3].

2- 3- خصائص المعجم التفسيري التأليفي الآلي:

ينطلق القاموس التفسيري التأليفي الآلي من البحث عن الأجوبة التي ترتبط بما يطرحه من اشكالات تحديد هوية المتواليات الهندسية اللغوية وذلك بشمل ذكي من قبيل[4]:

***/ ما هي العجمات التي تمثل المحمولات الدلالية والقيم المرتبطة بها؟** مثلا :

أ(+–انسان)+ ب(+ فعل).

***/ وما هي العجمات الأفعال التي يتحقق عاملها التركيبي العميق بشكل إجباري ويتخض صورة اسم مسبوق بحرف** (Prés position) ؟

***/ وما هي العجمات التي تمثل عنصر قيمة في الوظيفة المعجمية؟**

كما يعتمد هذا القاموس الآلي للعجمات وذلك بأسلوب التطابق غير المقبول دلاليا وتركيبا بين عدد العوامل الدلالية في تعريف ما بين الانجازات التركيبية للعوامل المعروضة في خانة جدول النظام[5]. هذا علاوة على الوظائف المعجمية التي تشكل العمود الفقري في اللغة كالترادف والتضاد..أما فيما يتعلق بالاستنباط المنطقي فإنه يقوم بالبحث عن طبيعة العلاقة القائمة بين قيمة

1- عز الدين غازي ، مرجع سابق، ص،185، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص93.

2- المرجع السابق، الصفحة نفسها.

3- المرجع نفسه، الصفحة نفسها، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص93.

4- عز الدين غازي ، مرجع سابق، ص،186.

5- المرجع نفسه، الصفحة نفسها.

الوظيفة المعجمية وتعزيزها برمز يقترحه ميلتشوك لهذا الغرض وهو Magn[1] يحدد القيمـــة الوظيفية والقيمـــة الدلاليـــة في تعريف العجمة[2].

كما أنه يقدم برنامجا للكشف عن أوجه التشابه الموجودة في نظام الأفعال التي يظم تعريفها ما يسمى بمكون المعاملة Stransaction ناهيك عن ملائمته لأنظمة الاستغلال في إصداراتها المختلفة Windows مثلا...[3]

2- 3-1- بنية وصورية معجم التواردات بالمصاحبة (Dico):

سيتم التركيز في هذا الجانب على الوصف الصوري للمعجم وعلى توضيح بنيته العامة ومادته الرئيسية، والتعريف بخصائص الفوكابل المعجمية والدلالية والاكتفاء ببعض النماذج التطبيقية على التعبيرات المتلازمة والمركبة في اللغة العربية ولتوضيح عمل البنية العامة لمدونة القاموس الواصفة للمادة الرئيسية، تخزن وتجهز (جذاذيات- نص)[4] بالمواد الرئيسية للقاموس فتسند إليه رموز مشفرة وفق الشكليات ومحارف (Caractere)[5] مناسبة لتسهل قراءة المعطيات محللة ومؤولة لبرنامج معلومات ليتبنى الأشكال التي تصف المادة الرئيسية كما في الشكل التالي[6]:

<عدد صحيح >←—— قانون
< مادة رئيسية (الفوكابل) >
< ملاحظات >

الشكل 3- بنية جذاذية القاموس ديكو "Dico"

يعتبر <........> رمز الدعامتان[7] مكون واصف .

إن المعلومة المذكورة في الجذاذية بمثابة قانون للمادة الرئيسية <عدد صحيح> التي تقع على يمين القانون، فهو بشير إلى التقدم الذي حصل على مستوى إعداد المادة القاموسية.[8]

1- انظر، Margarita Alonso RamosK Construction d'une base de données des collocations bilingues français espagnol,in langages N°134pp10.11.

2- وتجدر الاشارة ، أن القاموس يملك مواصفات تقنية هائلة تتمثل في عملية التشفير التي يتسم بها النص المترابط Hypertexte خاصة فيما يتعلق بمكونات التعريف وحالات الأسماء(علما أن (DEC) يفرض على القاموس ايراد جميع المكونات الأساسية في تعريف العجمة).

3- عز الدين غازي ، مرجع سابق، ص186.

4 – Fichier –texte

5- تحتوي جذاذيات القاموس أعلاه على نوع من المحارف Fonte واحد فقط ولم تتضمن أي أشكال أخرى للكتابة ثاخنة أو سفلية أو علوية.

6- عز الدين غازي ، مرجع سابق، ص187.

7 – chevrons

8- عز الدين غازي ، مرجع سابق ص187.

ولإعداد المادة القاموسية بسك ميلتشوك أربع قوانين[1] شكلية وأساسية للمراقبة والـــتأكد هدفها الاحتواء على المعلومات الضرورية لوصف التواردات بالمصلحة لعجمات هذه المواد بشكل شامل وعام دون إغفال أي معنى مرادف للمادة الرئيسة. ولجعل الوثيقة نهائية وصالحة يلتجأ ملتشوك إلى القانون 0 الذي يعطي للبنية الكبرى للقاموس دوره في وصف المواد الرئيسية في إطار الحقول الدلالية حتى تتأكد القيمة المعرفية والمنهجية المرجوة في (Dico). كما يمنح للبنية الصغرى الفرصة للتدقيق في توافق المادة الرئيسية، إن كانت مقولة اسمية أم فعلية، فيجب أن يكون هناك ارتباط اشتقاقي معلن في القيمة الممثلة للوظيفة المعجمية (V0) نحو: "كتب" (فعل) "كتاب" (اسم). ارتباط تحويلي للمادة الرئيسية نحو[2]:

"دخل" و"خرج" وتتكون المادة الرئيسية من المكونات التالية كما في الشكل الصوري التالي[3]:

< اسم المادة الرئيسية >
< مقولات نحوية >
<رمز الليكسيم الأول>
< وصف الليكسيم الأول >
.....
< رمز الليكسيم الأخير >
< وصف الليكسيم الأخير >

الشكل 4: يمثل مكونات المادة الرئيسية

يعمل الناشر المعلومياتي(Edico) بشكل آلي حيث يترك للمستخدم / المعجمي سوى ملء الفراغات، وهذه المرحلة تتسم بالتدخل البشري، وهذا مبدـــأ للمعالجة المماثلة – الآلية والبشرية – التي ينبني عليها القاموس التفسيري الآلي . فبناءً على طلب المستخدم يتم تلبية ذلك، فإن أراد مثلا مادة رئيسية ما نحو : ضرب (فوكابل فعلي) ، فسوف يبرز له بنية المعطيات التالية[4]:

1 –M.Decary an editeur specialise pour le dictionnaire explicatif et combinatoire du français contemporain ,unversité de Monterial 1986.pp.34-53.

2– عز الدين غازي ، مرجع سابق، ص187، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص96.

3– المرجع نفسه، الصفحة نفسها.

4– عز الدين غازي ، مرجع سابق، ص،189، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص97.

قانون
"ضرب
< مقولات نحوية>
_ <عُجْمة
خ ن < خصائص نحوية
خ د < خصائص دلالية
مل < ملاحظات أخرى
ج ن < جدول النظام
و م < وظائف معجمية

تتميز العلاقات بين مختلف المواد الرئيسية بالسلامة والدقة، وتبادل الفوكابلات وتلازم بعضها البعض . والقالب الذي أوردناه أعلاه يفسر مدى ولوج البنية انطلاقا من المفتاح " ". وهذه البنية تتصمن سبع خانات رئيسية،لكنها في نظر الدكتور محند الركيك بقيت ناقصة لأنها لم تضف ما يتعلق بخصوصية بعض اللغات الإعرابية كاللغة العربية من حيث جانب الصرف والإعراب والتعبيرات المجازية وغيرها[1].

وأما مهمة(Dico) كما أرادها صاحبه ايغور ملتشوك هي بناء المداخل المعجمية[2] محددة برموز لا تتغير أمنا للبس أو عطب في البرنامج كما يلي [3]:

< بداية المادة الرئيسية
> نهاية المادة الرئيسية
مدخل إلى بعض مقولات الفوكابل
بداية خانة العجمة.

2- 3-2- بنية قاعدة معطيات القاموس التفسيري التأليفي (Dico):

لقد كان المسعى الرئيسي والمبدئي من وراء بناء هذا القاموس هو بناء قواعد معطيات ذات تراتبية ومعمارية مُبْنَيْنَة[4] للمعجم الذري للغة الفرنسية[5] وقد أُعدّت لهذا الغرض قوالب لسانية وبرامج للتوليد النصي والمعجمي[6].

---

1 – محند الركيك، مشروع القاموس التفسيري التأليفي الإلكتروني للغة العربية، مجلة فكر ونقد، العدد82.اكتوبر 2006، ص 92،93 .

2– جواد حسني سماعنه، المعجم العلمي المختص، (المنهج والمصطلح)، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 48، الرباط، المغرب، ديسمبر1420هـ/1999م، ص 42، 43.

3– محند الركيك،مشروع القاموس التفسيري التأليفي الإلكتروني للغة العربية، سابق، ص.93.

4 – يعتبر هذا القاموس، كما يجري الآن تطويره نموذجا للقواميس المبنينة التي تعمل بقاعدة معجمية متعددة اللغات، وهذا ما ينكب عليه الآن مشروع (Papillon).

5 –Alin Polguére,Etiquetage des lexies dans la base de données DiCo.in TALN 44.N°2/2003.p.41.

6 – عز الدين غازي، مرجع سابق، ص،189، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص98.

لقد انفرد القاموس التفسيري الآلي بميزة خاصة تتمثل في اعتماده التعبيرات المتلازمة والمأثورات والأساليب والأمثال كمداخل معجمية مكونة إما من تركيبات نوعية (متلازمات) أو ليكسيمات[1] ، ولذلك يقع الخلط على مستوى هذا القاموس لأنه يتضمن العُجُمات[2] كلها. فهذا القاموس يساعد مستعمله على ايجاد التعبيرات المتلازمة بسهولة ويسر وذلك من خلال إدراجه لمختلف الاستعمالات الممكنة لتعبيرات معينة، فإن أردنا أن نبحث عن التعبيرات التالية[3] في قاعدة القاموس مثلا[4]:

ضرب له موعدا (حدده وعينه).

ضرب عصفورين بحجرة واحدة.

ضرب القاضي على يد فلان (حجر عبه ومنعه التصرف).

ضرب الدهر بين القوم (فرق وباعد)

ضرب به عرض الحائط (أهمله).

ضرب الدرهم (سكه وطبعه).

ينبغي تفحص المدخل المعجمي " ضرب" والقيام بالبحث في الجذاذة القاموسية عنها لايجاد التعبيرات المذكورة ضمن المشتقات الواردة في المدخل المذكور[5].

واعتبار لأهمية التعبيرات المتلازمة في الكفاية اللغوية الطبيعية وما تطرحه من مقومات في العديد من التطبيقات الحاسوبية والبرمجيات اللغوية والترجمة الآلية[6]، فقد كانت بداية للعديد من الابحاث العلمية التي استلهمت التطبيقات النظرية المعجم –التركيبي التي انصبت على اعطاء الأولوية للتعبيرات المتلازمة ضمن معمارية المعاجم الإلكترونية باللغات الفرنسية والانجليزية والاسبانية وغيرها من اللغات التي طبقت عليها هذه المقاربة .

2-3- المعاجم الآلية المبنية بنظام ديلا: (DELA):

نظام المعاجم الالكترونية(DELA)[7] ينتمي إلى أنظمة المعاجم الآلية التابعة لمختبر العلاج الآلي والتوثيق اللساني LADL فهو نظام لساني حاسوبي يجمع بين قوالب الوصف في نظام اللغة الفرنسية وبرامج معالجة المعطيات، الهدف الرئيسي منه هو معالجة النصوص وتحقيق

---

1 – lexémes

2 – lexies.

3- يعتبر على المعالجة الآلية والتوثيق اللساني (LADL) نموذجا من المراكز البحثية الرائدة في العالم على مستوى المعالجة الآلية للغات الطبيعية، والذي تمكن من بلوزة معاجم إلكترونية للغة الفرنسية استنادا إلى نظرية المعجم التركيبي (Lexique-Grammaire)

4- عز الدين غازي ، مرجع سابق، ص،189،

5- المرجع نفسه، ص 190.

6- عبد القادر الفاسي الفهري، التوليد والنسقية والترجم الآلية، المجلد 2، منشورات معهد الدراسات والأبحاث للتعريب، الرباط ، المغرب، 2001،ص 81،80.

7- يعتبر مخبرالمعالجة اللآلية والتوثيق اللساني LADL نموذجا من المراكز البحثية الرائدة في العالم على مستوى المعالجة الآلية للغات الطبيعية والذي تمكن من بلورة معاجم إلكترونية للغة الفرنسية استنادا إلى نظرية المعجم–التركيبي Lexique –Grammaire.

التواصل باللغة الطبيعية مع الآلة التي تستوعب النماذج المصورنة بدقة متناهية انسجاما وانتظاما، وللإجابة عن هذه المتطلبات ظهرت هذه المعاجم باعتبارها قواعد معطيات مبنينة حيث وحدات اللغة مصنفة تصنيفا صوريا علاوة على الخاصيات اللسانية للغة الموصوفة بدقة تكون مناسبة للبرمجة الآلية[1].

وبمأن الآلة في حاجة إلى قواميس لإغناء دورها في تقديم المعلومات بالسرعة الفائقة، مع أن القواميس ذاتها تبقى غير مكيفة وملائمة بالشكل مع التقنيات المعلوماتية المتطورة من جهة، وغير متوازية ومكتملة من حيث المحتوى من جهة أخرى. وإذا قارناها في العمل المعجمي الدؤوب والمستمر[2].

2-3-1- الأدوات المعلوماتية المطبقة في معالجة مداخل المعجم الإلكتروني :

أظهرت المحولات والأتمتة بأنواعها المختلفة، قدرتها على المعالجة الحاسوبية للبنيات والمتواليات اللسانية ذلك لأنها تمنح القدرة على تمثيل مدمج للمعاجم الآنية كما أنها أساس خوارزميات مراحل المعالجة الآلية للغات الطبيعية ومراحل التحليل الفومولوجي للتعرف على الكلام الطبيعي والتمثيل التركيبي للنصوص[3]. أما الأدوات المعلوماتية المستعملة[4] في معالجة التعبيرات المتلازمة والمركبة العربية هي اتومات الأوضاع النهائية والمحولات النهائية التي تقوم بدورها في عمليات الوسم المعجمي والتعرف على الوحدات في مرحلة التحليل. مع أن السمة الأساسية لهذا النوع من الأدوات العقلانية هيي المرونة الملازمة والسرعة ويسر مراحلها الإجرائية، ابتداءا للوصف الوحدات المدمجة والمقروءة إلى الشكل الإجرائي الوصفي الذي يتيح المعالجة الناجعة للمتواليات المعجمية. فغالبا ما يكون هما زمن المعالجة زمنا خطيا إذا ما قارناه مع حجم النص، كما يكون في نفس الآن مستقلا عن حجم الأتومات لما يحتويه من أنحاء محلية مناسبة. وأثناء هذه العمليات يمكن اختيار بعض الأدوات الأخرى المناسبة كأتومات خاصة بالتحليل المعجمي مع أن هذه الأدوات المعلوماتية مبدئيا، قابلة للاستغلال والتعديل خارج

1- عز الدين غازي ، مرجع سابق، ص،191، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص99،98.

2- يقول أحد أعمدة المعجمية الفرنسية صاحب روبير الكبير، بول روبير: في مقدمة روبير الكبير باللغة الفرنسية "إن إعداد قاموس عام للغة يتطلب مضنيا دؤوبا ومتواصلا لسنوات طوال ، ذلك لبلوغ الهدف بالحاح ومجهود قوي. فعلى الرغم من وجود القواميس التقليدية المحوسبة على الأقراص المدمجة، تبقى مضمونها محدود ومحددة سلفا من قبل الناشر. هذا ناهيك عما يفترضه صانعي القواميس في كون قارئيها يجب أن يكونوا بصفة ضمنية، ملمين بمستوى معين من المعاني كما يجب أن تكون لهم القدرة الكافية لتأويل الألفاظ والتعاريف وتناولها عن طريق قياسها بألفاظ أخرى وبمكنزمات الإبداع المعروفة لدى الإنسان .

3- عز الدين غازي ، مرجع سابق، ص،191، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص100،99.

4 في الادوات المعلوماتية سنطبق العبارات العقلانية، اتومات الأوضاع النهائية ذات الحالات النهائية ،خاصية كلين،التومات والغموض، الرسم الحلقي وغير الحلقي والمحولات التي دورها أمن اللبس وشبكة الانتقال المعززة .

الإجراءات التي يختارها أثناء المعالجة. فيما لا تستعمل أدوات التعليم الإحصائي دائما إلا من داخل شبكة العصابية والمعرفية وبالأحرى وفق نظام إحصائي للتعلم محدث لهذه الأدوات.

ويستند كل تحليل نصي على التعرف على المتواليات النصية المؤلفة من عدد نهائي من الوحدات الدنيا مثل[1] : أكل خالد التفاحة.

فمن ناحية الطباعة نجد أن الوحدات الدنيا هي عبارة عن حروف وعددها في المثال أعلاه أربعة عشر حرفا، ومن الجانب الإملائي يمكن اعتبار النص مؤلف من ثلاثة أشكال بسيطة ونهائية وهي: الفعل والفاعل والمفعول به.

أما بالنسبة للمستوى المورفولوجي فإن الأشكال البسيطة هي نفسها عبارة عن مورفيمات (الواصف) وهي نهائية أيضاً، فمثلا شكل (التفاحة) هو سلسل مكون من (أل) + (تفاحة) . أما من حيث المستوى التركيبي والدلالي فإن الوحدات الدنيا هي الكلمات، ففي المثال أعلاه نجدها ثلاثة وهي (أكل) + (خالد) + (التفاحة).

ومهما يكن المستوى التحليل المعجمي، فإن الوحدات هي رموز نهائية وأن مجموعة الرموز هي مفردات وأن كل سلسل للرموز هو متوالية وأن مجموع المتواليات هي لغة. وهذا ما يطلق عليه بلغات Kleene[2] التي تحدثنا عنها في المدخل العام للأطروحة. فيما يخص العبارات العقلانية، أما في ما يتعلق بالأوتومات النهائي فهو أيضا عبارة عقلانية تستخدم للتمثيل، وتميز فيها فيما بين عقدين واح أولية (دخل) وأخرى نهائية (خرج) ونرمز للعقدة التي تملأ بالعناصر الصورية ب [3]:

(< E >) ولكن العقدة النهائية تبقى غير مسومة لأنها تمثل الخرج النهائي للأوتومات.

نشير هنا إلى أن هذه الأنظمة التقنية المتطورة (التي تعالج بنجاعة الأشكال الصورية للغات الطبيعية بغية بناء أنظمة المعالجة والفهم) قد استخدمت في بيئة (INTEX)[4] بمختبر (LADL) بباريس7، وهذا ما شكل تحديا لنظام (INTEX) الذي استطاع أن يدققها ويحددها

---

1 – M.silberztein,dictionnaires electroniques et analyse automatique de système INTEX,masson 1993.pp.6

2– المرجع نفسه، الموضع نفسه.

3– المرجع نفسه، ص 192، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص100.

4– نظام (INTEX) هو برنامج معلوماتي أنشأه مختبر العلاج الآلي والتوثيق اللساني التابع لباريس 7 تحت إشراف العالم الكبير موريس كروس من أهدافه صورنة تطبيقات المعجم – التركيبي للغة الفرنسية ووصفها من أجل تحليلها ومعالجة بالحاسوب اعتمادا على قاعدة بيانات لسانية موسعة تستجيب لكل البرامج المعالجة وكما هو مسطر في أعمال المختبر والمتمثلة في أصحابها (J.P.Boon, A.Guillet,C.Leclère,M .Gross, 1976.1979.1982.1988.) تمثيلا لا حصرا، كانت بداية حول التمثيل التركيبي من خلال بناء جداول المعجم – التركيبي متضمنة جميع ألفاظ اللغة الفرنسية وتعبيرات ومركباتها، هذا علاوة على الوظائف الشكلية الأخرى المندمجة في مختلف مكونات النظام هو صميم أهداف المعاجم الإلكترونية(DELA) كما سنفصل القول في ذلك (للمزيد ينظر، Blandine, Courtois , M.Silberztein, les dictionnaires électroniques DELAS et DELAC .In (.LinguisticaCommunicatio V1. N1.1989.41-47.

ويوسمها ويفهرسها بنجاح كبير كما فعل الباحث ماكس سلبرشتاين في عمله عن المعاجم الإلكترونية التحليل الآلي للنصوص تحت إشراف العالم الكبير مويس مروس[1]، فأعطت نتائج جيدة على المستوى الإلكتروني للغة الفرنسية، وفي مشروع بحثنا هنا نريد نعرف بأهمية هذا الأسلوب المتقدم إلى جانب التقنيات الذكية لأنظمة الذكاء الاصطناعي (وبخاصة أنظمة قواعد المعرفة) في معالجة المعجم اللغة العربية وحوسبته. وسنحاول التطرق باختصارإلى تطبيقات هذه التقنيات على المفردات الكلمات المدخلة المركبة، وهي[2]:

2-3-1-1- **تقنية المحولات** (transducteurs):

في إطارنظرية المعجم التركيبي، أن الجملة هي الأساس، فتمثيلها يجب أن يخضع لدالة جبرية رياضية، تستطيع بقدرتها العالية وصف أي لغة طبيعية كانت، كما أن لها القدرة على تطوير تطبيقات المعالجة حسب الأنحاء المحلية المناسبة لكل متوالية. لقد أصبحت هناك امكانية عالية لربط الجمل بالتمثيلات الميتالسانية لبناء نظام للترجمة الآلية وذلك بفضل (الرسومات)[3] الذي يعتمده نظام التطوير اللغوية (INTEX) المرتكز على مبدأ المحولات[4]بحيث يستطيع الناشر أن يغير ويبدل في عناصر السلسلة الواصفة والمتمثلة في عقد هذه الرسومات التي تعمل بصفة في السلسلة المعنية. وقد اعتمد على نظام الترجمة الآلية في بناء معماريته مجموعة من الأوتوماتات تستخدم لتمثيل معلومات المقولات التركيبية والدلالية الكامنة في الجملة عن طريق تلك المحولات كما يظهر من خلال ثلاث قوالب خوارزمية على الشكل التالي[5]:

شكل 5- خوارزم التحويل

1- M.Silberztein, Dictionnaires électronique et analyse automatique se texte, le système INTEX, Masson 1993

2- الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص102.

3 – graphes.

4- تستخدم هذه التقنية لتمثيل خوارزميات الناشر الآلي الذي يمثل كذلك باسم (graphe) وكل عملية تمثيلية ميتا لسانية تمر عبر قوالب التعرف والتحويل (transduction) والتوليد . وتعمل أنظمة الترجمة الآلية بهذا الأسلوب بحيث أراد الدكتور عز الدين غازي الترجمة من اللغة 1 إلى اللغة 2 مثلا فلابد لك عبور خوارزميات التمثيل التي ذكرها في أطروحته، للمزيد من بسط القول ينظر: عز الدين غازي ، مرجع سابق.

5- حاول كل من بلانكو وروي (1999) بناء للترجمة الآلية يستغل المعاجم الإلكترونية للبلوغ إلى تمثيلات ميتا لغوية قبلية للجملة المعالجة للترجمة الآلية عبر المحولات التي توظف في التعرف على المتوالية اللسانية ذات البعد المعجمي، وقد ذهب بيير بوفي إلى حد اعتبار أن إطار نظرية المعجم – التركيبي ومنهجية نظرية معنى – نص متشابهان من حيث كونهما معجميتان في تناولهما لإجراءات التحويل التركيبي والمعجمي للوحدات الاسمية والفعلية.للمزيد انظر Pierre- Andre Buvet. Représentation métalinguistiques de phrases a partir de traducteurs, Centre Lucien Teniere, Universite de Franche- Comte. A paraitre c (2002)

فكل مرحلة من هذه المراحل مزودة بخوارزم خاص يقوم بإجراءات معينة، فإن أردنا أن نترجم بنية لغوية من لغة إلى أخرى، فما علينا إلا أن نمر بهذه القوالب كما يظهر من خلال الشكل التالي[1]:

الشكل 6- يبين خوارزم التحويل من لغة إلى أخرى

2-3-1-1- تقنية أوتومات الأوضاع النهائية: (Automates à état finis):

يعتبر اوتومات الأوضاع النهائية[2] عبارة عقلانية تستخدم للتمثيل، وتميز فيما بين عقديتن واحدة أولية (دخل) وأخرى نهائية (خرج) ونرمز للعقدة التي تملأ بالعناصر الصورية ب (< E >) بينما تبقى العقدة النهائية غير موسومة، لأنها تمثل الخرج النهائي للأوتومات. ولما كان هذا النوع من الأوتومات ينتمي لفصيلة العبارات العقلانية فإنه بشكل من الأشكال يقوم بوظيفة المحولات ذات الأوضاع النهائية في مختلف العمليات[3]، أي تلك التي لا تنتج أية معلومة غير تلك المعلومة الثنائية المتعلقة بـــ[4]:

أ. المتواليات المعروفة.

ب. المتواليات غير المعروفة.

1- الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص103.

2- يعتبر جهاز أوتومات الأوضاع النهائية الذي استغلته العديد من التحليلات الحاسوبية وطبقتها على اللغة العربية في مستوياتها الصرفية والصوتية والتركيبية، أداة عقلانية قد تقع في عدد من الأعداد النهائية للحالات، بمعنى آخر، يمكن تتحول من حالة إلى أخرى. وهذه تسمى بالمرحلة الانتقالية. فحينما ينطلق الأوتومات، فإنما ينطلق من حالة أولية إلى حالة نهائية حيث ينتهي عمله في تلك النقطة الأخيرة. ومن هنا يتلقى مدخلاته (input) وهي عبارة عن متواليات من الرموز آتية من مجموعة متناهية من الرموز تسمى بالأبجديات (alphabet)، إذ تسند حالة لكل رمز حيث المرور من حالة إلى بيد أن الفهم الآلي للغة يتطلب عمليات أعقد لأنه نظام من الأنشطة النفسية، تشكل الذاكرة أحد ثوابته الأساسية. وأثناء المعالجة يخلق النظام وحدات الذاكرة المختصة ويغيرها ويمسحها. وتستخلص المعلومات المخزنة في السجلات إما من النص بعد تحويلها بواسطة قوانين أو تبنى على شكل شروح جزئية. وهذه العملية تقابل الذاكرة البعيدة لدى الإنسان، وبهذا يستطيع النظام تكوين ميتا- معرفة (Méta connaissance) يحفظها أيضا في سجلات معينة. وأهم ما يميز النظام هو توظيفه للذاكرة كي يحتفظ معالجه بالنقطة التي خرج منها من الشبكة ليدخل في أخرى، وبالنقطة التي فيها باختيار معين من بين مجموعة من التعليمات الممكنة.

3- تعتبر العبارات العقلانية وسيلة سريعة لإدخال الأوتومات النهائي البسيط في النظام، من دون اعتماد الرسم، حيث تكون المتواليات الصورية المبحوثة عنها قصد المعالجة تستلزم وتقضي دائما متواليات أخرى هي دائما عبارة مفردات، بفضل السرعة المميزة لهذه العبارات أمكن إدخال هذه المفردات بشكل مباشر، وفي حالة تعقد الأمر نستعيض عن ذلك بالرسوم (graphes)

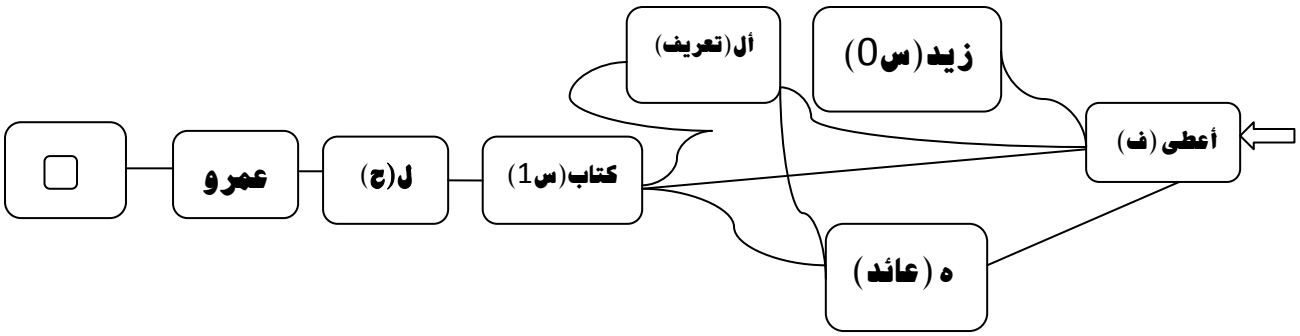
4- عز الدين غازي ، مرجع سابق، ص،195،

وإذا كانت الأوتومات كالعبارات العقلانية تصلح لتمثيل مجموعة من المتواليات اللغوية فإنها تتعرف على المتواليات اللسانية بسهولة. فإن أردنا أن نفحص مدى أوتومات معين على متوالية معينة، فإننا نقوم بتطبيق الخوارزم التالي[1]:

" ننطلق من العقدة الأساسية للأوتومات ونتبع الممرات التي تعبرها العقد حيث تتطابق قراءة الرموز بها مع البطائق. وإذا بلغنا العقدة النهائية بقراءة جميع المتواليات اللسانية، فإننا نقول أن المتوالية معروفة، أما إذا لم نستطع الوصول إلى العقدة النهائية، أو أننا نصل بدون أن نكون قد قرأنا المتوالية، فقل أنها غير معروفة".

فلنتأمل الجملة التالية: أعطى زيد كتابًا لعمرٍ

التي يمكن تمثيلها بواسطة الأتومات النهائي كما يلي[2]:



الشكل رقم 7: يمثل نموذج الأتومات " أعطى زيد كتابًا لعمرو"

عند قراءة المتوالية التالية " أعطاها الكتاب" فلن يكون هناك التطابق بين الممر الذي سنصل به إلى العقدة المتوالية الموسومة بالضمير "ها"، لأن القراءة تبلغ مداها ن=مباشرة عند تعرّفها على المتوالية "أعطاها كتابا" وبالتالي فإن الأتومات لن يتمكن بالتالي من أن يتعرف على المتوالية"أعطاها كتابا " ذلك أنها غير مطابقة تماما في الرسم. ومع ذلك يمكن منح نوع من الحرية للأوتومات من خلال إعطاء العقد وسما مركبا مؤلفا من متواليات الرموز غير متصلة فيما بينها، وهذا ما يفضي إلى بناء ممرات أخرى مما يجعله خفيفا . فيتم كتابة رسومات أخرى توازي الأوتومات النهائي المفترض[3] .

---

1 –Max. Silbertzein.(1993).p.9.

2– عز الدين غازي، معمارية المعجم العربي الإلكتروني : رؤية جديدة لمعجم آلي مُبنين، الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، يناير، مطبعة المعارف الجديدة، الرباط، 2009، ص 103.

3– المرجع نفسه، الصفحة نفسها.

2-3-2- طبقات المعاجم الإلكترونية من البسيط إلى المركب:

في البداية لا بد من الإشارة إلى أن إعداد معجم إلكتروني عربي مبنين يقتضي اعداد قاعدة بيانات ذات منظومة متناسقة تضم طبقة متراصة من قواعد معطيات خاصة لكل مستوى من المستويات اللغوية العربية، كالجذور والمفردات البسيطة والمفردات البسيطة المعرّبة والمفردات المركبة والمتلازمة كمثال تطبيقي في هذا المبحث .

2-3-2-1- المفردة في المعجم العربي الآلي:

تتألف قاعدة المعطيات من جميع مفردة اللغة العربية وفي مختلف أشكالها التصريفية والصرفية: الاسم في جميع أوجه ظهوره الممكنة في الاستعمال اللغوي العربي، والفعل المصرف في جميع الأزمنة، وكذلك الظرف والصفة، إلخ وبما أن هذا الوصف سوف يتعامل مع برامج تهدف إلى التحليل الآلي، فقد استدعى ذلك إعطاء معنى محدد للمفردة، التي تم تعريفها على أنها متوالية الحروف التي تتضمنها الأبجدية العربية محاطة ببياض بينها وبين ما يقابلها وما بعدها[1]. وهذا التعريف يشمل الأسماء والأفعال والظروف والصفات والأدوات وغيرها مع استثناء تلك المفردات التي خضعت لتركيبات الاضمارية مثل "أعطيتكه". كما يخرج من هذا التعريف الأسماء المركبة والمتلازمة منها والعادي مثل: أسداسا في أخماس وبصيص أمل فهذه المتواليات تشكل معطيات هذا معجم المفردات المركبة، كما تشكل أساس المعجم – التركيبي المخصص للتعبيرات المتلازمة. وبما أن هذا المعجم معجم مفردات فإن هذا الجزيئات سوف تدخل إلى المعجم مستقلة عن تراكيبها فمثلا (وبمدرستهم) سوف تضبط (و)في المعجم برقم معين كمرجع لها في البرنامج المعلوماتي (ح،. رقم) وكذلك الباء أم بالنسبة ل(هم) فسوف تظهر مرقمة بالشكل (ض،.رقم) (مدرسة،. رقم) وهكذا[2].

2-3-2-2- تقنية بناء المادة المعجمية في المعجم الآلي العربي:

تعتمد تقنية بناء المعجم الآلي للغة العربية –على غرار المعجم الآلي الفرنسي[3] على ثلاث مستويات متكاملة[4] وهي مستوى الجذور ومستوى المفردات البسيطة والتي تتفرع بدورها إلى المفردات البسيطة المعرّبة. أما المستوى الثالث فيتأسس على قاعدة بدورها إلى تجمع

1- عز الدين غازي ، بناء المعجم الآلي للغة العربية، معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص197، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص104.

2- عز الدين غازي ،مرجع سابق، ص197.

3- يقتضي إعداد قاعدة معطيات خاصة على مستويات لغوية وهي مستوى الجذور ومستوى المفردات البسيطة. أما المستوى الثالث يرتكز على قاعدة تجمع المفردات المركبة والمتلازمة.

4- انظر، Blandine courtois et max silberztein, les dictionnaires electroniques DELAS et DELAC in languistica comunicatio,vol1.1N°2.2.1989.pp.64-80.

المفردات المركبة والمتلازمة، والتي تتفرع بدورها على مفردات مركبة تشمل العادي منها والمتلازم مفردات مركبة معرّبة[1].

إجرائيا يتم تقسيم المعطيات إلى جميع أقسام الكلم في اللغة العربية كما يتم التعامل معها على شكل مداخل معجمية مستقلة تماما عن اطاراتها التركيبية، مخزنة في قاعدة معطيات المعجمية على أساس نحوي وصرفي[2] كما هو واضح في الجدول التالي[3]:

المرجع الصرفي	المرجع النحوي	نوع المعجم
الرقم الخاص بالصنف التصريفي	الرمز الخاص بكل مدخل (ف، س،ص، ظ، ح)	معجم المفردات البسيطة

إن بناء قاعدة معطيات الخاصة بالمفردات البسيطة سيكون بالفنية التالية، على أساس أن يثبت الرمز المحيل على الصنف النحوي أولا ثم يتبع مباشرة بفاصلة بعدها نقطة تسبق الرقم المحيل على الصنف التصريفي كما في المثال ف،.1 بهذا سيصبح من السهل على البرنامج للقيام في عملية البحث التي ستمكن من تكوين معاجم مختلفة حسب ما يطلب منها[4].

غير أن هذا الأسلوب سيصبح عاجزا في حالة بعض المفردات المركبة وبالتالي يقتضي الأمر النظر إليها من الجانب التركيبي المعجمي، بتوظيف قائمة الرموز المسطرة في الجدول الخاص بذلك، كما بينها الدكتور الحناش في الطريقة المبينة[5] في الجدول التالي:[6]

المدخل العجمي: المفردة المركبة	
متن الاسم المركب	البنية الصورية لكل صنف للإسم المركب(رمز1+ ...)
الجزء الأول مصرف دون الجزء الثاني	البنية الصورية (رمز1+ رمز2+رمز3)

1- تتأسس تقنية بناء هذا المعجم على القيم الصرفية والتركيبية والإعرابية التي تساعد البرنامج المعلومياتي على التعرف مباشرة على المفرد في السياق، واعتماد على بيانات إحصائية مصنفة وموصوفة يتم اسناد رقم معين للبنيات المصنفة مورفو-تركيبيا، علاوة على المعلومات الإضافية التي تساعد بنسبة لا تقل عن 10 %لاجتناب الهامش الخطأ لدى المعجم وهو ما يعرف ب Reconnaitre les données génératricesde bruits للمزيد انظر المرجع السابق ص 64.

2- لبناء المعجم الآلي المفردات المركبة العربية، اقترح الد.محمد الحناش تقنية تجمع بين الجانب التركيبي والمعجمي وذلك بتوظيف قائمة الرموز المسطرة في جدول خاص بذلك، انظر المعجم الآلي للغة العربية، بناء قواعد البيانات، مجلة التواصل اللساني م4.ع1.مارس 1992 ص85.

3- عز الدين غازي ، بناء المعجم الآلي للغة العربية، معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص198، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص105.

4- المرجع السابق،الموضع نفسه.

5-.محمد الحناش، نظرية حاسوب لسانية لبناء معاجم الآلية للغة العربية (محاولة في التأصيل)، مجلة التواصل اللساني ،المجلد الثاني،ع2.1990 ص75.

6- عز الدين غازي ، بناء المعجم الآلي للغة العربية، معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص199، وينظر : الدراسات المعجمية: مرجع سابق، العدد السابع والثامن، منشورات الجمعية الجمعية المغربية للدراسات المعجمية، مطبعة المعارف الجديدة، الرباط، 2009، ص106.

حيث أن الصنف التصريفي هو الطريقة التي يصرف بها الفعل والاسم في اللغة العربية.

حيث يظهر من المعالجة أن عملية التصريف قابلة للتأطيرفي أصناف متجانسة، يشمل كل صنف منها عددا من المداخل المصنفة كما يتبين من الأمثلة المدرجة في الجدول 3 التالي[1] :

أخماسا في أسداس	س ح س (اسم+ حرف+ اسم)
مرض عضال	س ص (اسم+ صفة)
حيص بيص	س س (اسم+ اسم)
بصيص أمل	ص س (صفة+ اسم)

يتضح أن قاعدة المعطيات التي ينبني عليها المعجم الآلي تتألف من نوعين من المعطيات وهما[2]:

أ- يتأسس على مداخل منعزلة خارج السياق التركيبي، وهو مصنف بناء على قواعد صرفية ومعجمية خاصة ، وتكون مادته المعاجم التالية: (DEMASF,DEMAS,DEAR).

ب- يتأسس من قواعد تركيبية – معجمية، ويتعلق الأمر بالمعاجم التالية: (DEAMC, DEAMSC).

أما المعجم –التركيبي الخاص ببنيات اللغة العربية فيتأسس من قاعدة معطيات تتألف من التراكيب العربية بنوعيها: العادي(البنيات الحرة)، والمتلازم(البنيات المركبة).

3-2-4- مشاكل وصعوبات في تمثيل المعلومات في المعجم الآلي العربي:

ومن جملة المشاكل التي تعترض المعالجة الآلية وبالتالي بناء هذه المعاجم هي إشكالية غموض بعض المتواليات المركبة (أو حتى البسيطة في بعض الأحيان) التي يصعب التعرف عليها في النص المعجمي المعالج[3]، ولحل هذا الأمر نُلْتَجَأُ إلى طريقة التسنين (codage) التي تسهل عملية ترتيب المداخل، وهي اسناد رمز لكل صنف عند بناء قاعدة المعطيات في سائر مستويات المتواليات معجمية كما يبين الجدول التالي[4]:

1- المرجع السابق،الموضع نفسه.

2- أنظر:

Blandine, Courtois, M.Silberztein, les dictionnaires électroniques DELAS et DELAC .In (Linguistica Communication V1. N2.1989.pp64-80.

3- إن المشاكل المطروحة التي يواجهها المحلل المعجمي عند تطبيقاته المعلوماتية تتجلى بصفة خاصة: في نظام تجميع المعلومات والتعرف الآلي على المتواليات، فالتعرف على المفردات المركبة أثار مشاكل كثيرة لأنه مدعى لما يظهر خلال عملية التحليل والمعالجة والتصنيف والبناء ومنها أولا : مشكلة التفرقة بين المداخل التي تتجانس في الصيغة الصرفية الإشتقاقية مع انتمائها إلى نوعين من أقسام الكلم.للمزيد من بسط القول ينظر : محمد الحناش، المعاجم الآلية للغة العربية ، بناء قاعدة معطيات، مجلة التواصل اللساني، المجلد 4.ع.1.1992.ص.86.

4- انظر، محمد الحناش، المعاجم الآلية للغة العربية ، بناء قاعدة معطيات، مجلة التواصل اللساني، المجلد 4.ع.1.1992.ص.86.

اسم	س:	فعل	ف:
ظرف	ظ:	صفة	ص:
ضمير	ض:	حرف	ح:
مخاطب	ك:	تعريف	تع:
مخاطبة	خة:	مخاطبة	خ:
غائبة	غة:	غائب	غ:
اسم مفعول	مف:	اسم فاعل	فا:
ماضي	م :	مضارع	ظا:
مجهول	جه:	أمر	أ:
جمع	ج:	مفرد	ر:
مصدر	مص:	مثنى	ث:
المعجم الآلي العربي للجذور	DEAR	اسم متلازم	مت:
المعجم الآلي العربي للمفردات البسيطة المعربة	DEAMSF	المعجم الآلي العربي للمفردات البسيطة المعربة	DEAMSF
المعجم الآلي العربي للمفردات البسيطة المركبة	DEAMCF	المعجم الآلي العربي للمفردات البسيطة المركبة	DEAMCF

2-3-2-4-معجم الجذور الآلي العربي للجذور[1]: (DEAR):

يعتبر معجم الجذورأساس بنية المعاجم الآلية العربية، فهو الذي يخضع لتعددية التوليد ولكل مستعمل أو مهمل في اللغة، فانطلاقا من جذر واحد[2]: يمكن القيام بمجموعة من الاحتمالات الاشتقاقية. ومهما يكن من أمر، فإن معجم الجذور هو معجم صوري يمثل رافدا معلوماتيا أكثر منه رافدًا علمياً[3]. فالتوليد يتم انطلاقا من الجذر[4] وليس العكس، وفي هذا الصدد يقترح محمد الحناش التعامل معه بالطريقة التالية مثلا[5]:

1- المرجع السابق، ص ،87.

2- يؤكد د.عبد الغني أبو العزم أن الجذر لا يعتبر المحدد للبنيات اللسانية وللمدخل المعجمي ، لأن المادة المعجمية هي التي تدل المتلقي على الجذر .

3- محمد الحناش، المعاجم الآلية للغة العربية ، بناء قاعدة معطيات، مجلة التواصل اللساني، المجلد 4.ع.1.1992.ص.86.

4- ينظرالدكتور عبد الغني أبو العزم إلى الجذر برؤية لا خطية، فعلى الرغم من أن العمليات الاشتقاقية تتم انطلاقا منه، فإذا الشأن لبناء بالنسبة لبناء المادة المعجمية للتعبيرات المتلازمة والمركبة سيختلف تماما لأنها هي التي تدله على الجذر، مما يعني أن هناك حلقة دائرية خوارزمية تتم على مستوى الاكتساب المعرفي في الكفاية المعجمية الطبيعية، وهذا ما ستتمكن في اعتقادي التقنيات الجديدة المتمثلة في الأدوات العقلانية التي نطبقها ونُنَمْذِجها في المعجم العربي من حل مجموعة من مشاكل التوليد الآلي للمتواليات اللسانية.

5- محمد الحناش، المعاجم الآلية للغة العربية ، بناء قاعدة معطيات، مجلة التواصل اللساني، المجلد 4.ع.1.1992.ص.88.

جذر (ض.ر.ب): < ف = 0 + > ، < مص = 0 + > ، < فا = 1 + > ، < فو = 2 + > ، < س = 0 + >، إلخ[1] .

1 = أو أي عدد يفيد عدد الحروف الزائدة عن الجذر لتوليد الصيغة المشار إليها بالحرف قبل علامة (=) ، وعندما يتعلق الأمر بنقص الجذر يشار إليها بالرمز (ن) التي معناها نقصان .

إن الصعوبة الوحيدة التي تواجه المعجمي أثناء بناءه لهذا المعجم هي في كيفية تسنين الأدوات والحروف التي لا تعرف جذورا مطردة أو عادية كما هو الحال مع بقية أقسام الكلم، لهذا يلتجأ إلى معاملتها بشكل الإستعمالي لا بصورتها الأصلية غير الموجودة أصلا[2].

2-3-2-5- معجم المفردات البسيطة(DEAMS):

لقد خصصت معظم المعاجم العربية الورقية والمحوسبة في تصنيف جميع مداخل اللغة العربية البسيطة[3] فنظرت إليها من وجهة نظر نحوية صرفة،حيث أنها تميز بين الأفعال اللازمة والأفعال المعتدية، أو التي تتعدى بحرف أو بدونه، كما صنفتها إلى أفعال عادية وأفعال قلوب وأفعال ناقصة[4]. أما التصنيف الذي اعتمد[5] في بناء هذا المعجم فإنه يقوم على أقسام الكلم والتي تراعي الخصوصية النحوية للمدخل المعجمي (المفردة)، وربطها بالمستوى المرفولوجي المناسب لها (الصيغة والاشتقاق) على النمط التالي[6]:

ف: الفعل[7] في الزمن الماضي للمفرد الغائب المذكر.

س : للمفرد المذكر المؤنث.

ص: الصفة وسائر المشتقات بجميع أنواعها.

ظ: الظروف بجميع أنواعها.

ح: للحروف والأدوات بجميع أنواعها.

1- حيث أن : (+ = الحروف الزائدة عن الجذر، ــ = الحروف الناقصة في الجذر، 0 = ليست هناك زيادة في حروف الجذر).

2- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، إشراف الأستاذ الدكتور عبد الغني أبو العزم، جامعة الحسن الثاني، الدار البيضاء، المغرب، 2009/2008، ص101.

3- إذا كانت اللغة الفرنسية أو لغة أخرى إلحاقية تصنف بعض أقسام الكلم ضمن بيانات الأسماء كالصيغة المنتهية للفعل، فإنه في اللغة العربية لا تدرج هذه الصيغة مباشرة مع الاسم بل مع أصناف أخرى، وهذا ما يتطلب الدقة في التصنيف.

4- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، مرجع سابق، ص 101.

5- قامت الباحثة بلاندين كوراتواز بمعية ماكس سلبر شتاين عند بنائهم للمعجم الآلي للمفردات البسيطة في اللغة الفرنسية على تصنيف المقولات النحوية، اعتمادا على تغيراتها التصريفية بغرض التعرف عليها وتوليدها آليا وهي: الاسم والفعل والظرف والصفات ...وهكذا يمكن أن يتولد عن اسم مفرد مثلا كثير من الأشكال المعربة حسب النوع والعدد والجنس، وذهب محمد الحناش على على نفس المنوال باعتماده تصنيفا يراعى كل خصوصية للمقولات النحوية في نظام اللغة العربية، سواء كانت فعلا أم اسما أم حرفا وربطها بالتغييرات التصريفية الممكنة في النص المعجمي.

6- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، مرجع سابق، ص 101.

7- للفعل أدوار توزيعية في العمليات التركيبية وقد أكدت هذه الفرضية نظرية النحو التأليفي / التحولي ، كما أنه له أبنية وصيغ صرفية تراوح مكانها بين الصحة والاعتلال، علما بأن الصيغة المنتهية (infinitive) غير موجودة في اللغة العربية على غرار اللغات الإلحاقية مثلا، مما يجعل مقولة نحوية صعبة التناول في الوصف والتفسير، وهذا ما انطلق منه النجاة سواء منهم المتقدمين أو المتأخرين، انظر على سبيل المثال لا الحصر، سيبويه في الكتاب، ج.1 .تحقيق عبد السلام هارون بيروت 1988، ص269.

ويعتبر المصدر مفردة بسيطة لذلك يستحسن إدراج رمزه ضمن معجم المفردات البسيطة المعربة، و من المفروض أن يهتم المعجم الآلي بالظروف[1] بشكل مستقل عن بقية أقسام الكلم في نظام اللغة العربية[2]، فهي تشكل مداخل مستقلة عن الصفات والأسماء والأفعال، وهذا بالرغم من قلة عددها مقارنة مع أقسام الكلم الأخرى، لتبقى بذلك مكونات خمس[3] يتأسس عليها نظام أقسام الكلم للغة العربية[4].

2-3-2-6- معجم المفردات البسيطة المعربة (DEAMSF):

إن أحد استعمالات المعاجم الآلية تتمثل في تقديم معلومات عن المفردات التي يتألف منها نص ما، وذلك بهدف علاجها آليا.

1- تدخل الظروف في باب المعجم الآلي للمفردات المركبة، مع أن بعضها يخضع للتحويلات التي تجري على الظروف المركبة وخاصة في عملية التوسيم، وقد أعد ماكس سلبر شتاين جداول المعجم التركيبي خاص بها بحيث يمكن أن نسمها بقيم الإيجاب والسلب حسب المتغيرات التي تطرأ عليها نحو: صباح مساء التي يمكن أن نعرفها: الصباح والمساء ، لغة أجنبية : لغات أجنبية. للمزيد انظر تأسيم الظروف والإحالة والحذف والتعوي انظر: M.silberztein,dictionnaires electroniques et analyse automatique de textes Ed.masson 1993.pp106-

2- يتألف نظام اللغة العربية من أكثر من مستوى، فهناك النظام الصوتي، والنظام الصرفي، والنظام المعجمي، والنظام الدلالي، والنظام التركيبي، ولكل نظام من هذه الأنظمة قوانينه ومبادئه التي تضبط العلاقات الداخلية بين مكوناته وعناصره، وتنظم علاقته ببقية أنظمة اللغة، وتتداخل أنظمة اللغة فيما بينها تداخلا لا يمكن فصله على مستوى الاستعمال الحقيقي للغة، وإن كان ذلك ضروريا على المستوى البحثي، والتمثيل الحاسوبي للغة العربية عليه أن يستوعب الخصائص اللغوية لنظام اللغة العربية، بما يتسم به من تداخل بين مستوياته، واقتصاد في وحداته اللغوية، ووضوح وإفادة في التعبير عن مراد المتكلم، وثراء في البدائل الممكنة للتعبير عن الشيء الواحد، و يتميز نظام اللغة العربية بالعديد من الخصائص اللغوية التي اختص بها عن سائر اللغات مما فرض تحديا كبيرا أمام معالجتها آليا، ولذلك فعلى علماء اللغة الحاسوبيين أن يستقروا الظواهر اللغوية المختلفة بهدف الوصول إلى قوانينها التي تحكم أصولها وفروعها، وتقديمها على نحو يوافق القوانين الرياضية التي انبنى عليها الحاسوب، كما أن علماء الحاسوب اللغويين عليهم أن يبنوا النماذج الرياضية التي يمكنها أن تحاكي سلوك العقل البشري في تعامله مع اللغة، وتستوعب النظام اللغوي على ضخامته وتناميه، كل ذلك جعل من المعالجة الآلية للغة مغامرة يبرر خوضها ومواصلة السير في ضروبها أهمية الأهداف المرجوة من تحقيقها.

إن الوحدة اللغوية التي يتعامل معها المعجم الحاسوبي هي الكلمة، فهو يقدم لها المعلومات المتعلقة بها سواء كانت صوتية أو صرفية أو معجمية أو دلالية أو تركيبية وقد يكتفي ببعضها حسب أهدافه،

3- يقوم التصنيف الذي اعتمد في بناء هذا المعجم على أقسام الكلم والتي تراعي النظام الآتي:

أ- تسنين الفعل : يعتبر الفعل من أصعب المقولات ضبطا في النص المعجمي نظرا لتشعب علاقته بسائر أقسام الكلم وخاصة منها المصدر وسائر أنواع الصفة. وبالتالي فقد تم تقسيم الأفعال إلى مجموعات تصريفية يتمثل فيها أقصى حد ممكن من الاطراد في الانتماء إلى الصنف المناسب. وهذه الأصناف تحدد من خلال الحركات التي يأخذها الفعل في المضارع، وحسب الصحة والاعتلال، وحسب الزيادة أو عدمها.

ب- الأسماء البسيطة: إذا كانت الأفعال تستجيب للترقيم المطرد اعتمادا على نظام الصيغ التصريفية، فإن ذلك سيتعذر مع الأسماء التي لا تعرف أي انتظام تصريفي يمكن اعتماده في إقامة تصنيف معجمي من نفس القبيل. ولهذا تتبع طرق أخرى لتصنيفها.

ج- الصفات والمصادر: مراعاة للتلاؤم التوزيعي الحاصل في تصنيف الأسماء البسيطة والصفات والمصادر وخاصة التطابق. تم التصنيف بالطريقة التالية وهي :

الصنف الأول: وهو مخصص للصفات الخاصة بالذكر، مثل عضال، في مرض عضال ، والتي ليس لها مؤنث من لفظها.

الصنف الثاني: فقد تم تخصيصه للصفات التي ليس لها مذكر من لفظها، مثل حامل، للمرأة الحامل

الصنف الثالث: فقد تم تخصيصه للصفات التي يشترك فيها المذكر والمؤنث، مثل جميل ← جميلة

الصنف الرابع: فقد تم تخصيصه للصفات المستعملة في بعض التخصيصات العلمية، مثل: أحمر قان

د- الظروف: من المفروض أن يهتم المعجم الآلي بالظروف بشكل مستقل عن بقية أقسام الكلم في نظام اللغة العربية فهي تشكل مداخل مستقلة عن الصفات والأسماء والأفعال، وهذا بالرغم من قلة عددها مقارنة مع أقسام الكلم الأخرى.

هـ- الحروف والأدوات: لقد تم تخصيص هذا الصنف من المداخل والأدوات في اللغة العربية ويعني بها حروف الجرّ والأدوات التي تقوم بدور الربط بين الكلمات دون أن تكون هي نفسها ذات دلالة مرجعية وهذه المداخل تدرجها كتب النحور تحت باب الحروف أو حروف المعاني، للمزيد من بسط القول ينظر : محمد الحناش، المعاجم الآلية للغة العربية ، بناء قاعدة معطيات، مجلة التواصل اللساني، المجلد 4.ع.1.1992. ص.89.

4- عز الدين غازي ، بناء المعجم الآلي للغة العربية، معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص202.

فبالإضافة إلى مهمة التعرف على أقسام الكلم يجب إعطاء معلومات أساسية مرفقة بكل مفردة معرّبة[1].

– **بالنسبة للأفعال** : تحدد طريقة تصريفها ونسبتها إلى الضمائر ومختلف المورفمات المتعلقة بالجنس والعدد والزمن.

– أما **بالنسبة للأسماء:** فتحدد طريقة تصغيرها ونسبتها وعلامتها الإعرابية الممكنة مع كل واحد منها. وبذلك سوف يتم رفق كل نوع من المداخل المعجمية التي يتضمنها معجم المفردات البسيطة بهذه المعلومات. وبعبارة أخرى، فإن كل شكل من أشكال هذا المعجم يجب أن تعطى عنه جميع الاحتمالات التصريفية والصرفية التي يوجد عليها في اللغة العربية، وهذا ما يشكل موضوع القسم الآخر من قاعدة المعطيات والتي يطلق عليها: معجم المفردات البسيطة المعربة على النمط التالي[2]:

الشكل أ =: جذر ب، صرف ج[3]

2-3-3- قاعدة بيانات المعجم العربي الآلي للتعبيرات المركبة غير المعربة والمعربة(DEAMC):

2-3-3-1- معجم المفردات المركبة:

لا يمكن للبرامج المعدة للتعامل مع المعجم الآلي أن تتعرف على جميع المكونات النص المكتوب إذا لم يتضمن هذا المعجم جميع أنواع الأسماء المركبة في اللغة العربية، نحو علامة المرور-أحد عشر – جواز السفر، وهي مداخل غير قابلة للتجزئة، لأنها تشكل في أغلبها مداخل متلازمة وبالتالي تتطلب وصفا صرفيا وآليا من نوع خاص[4]. فالأمر يدعو إلى الجمع بين نوعين من الوصف: والاهتمام أولا بالوضع المورفولوجي للعناصر التي يتألف منها الاسم المركب، ثم ثانيا وصف وضعه الجديد في التأليفية التي أنشأها منه نظام العربيةّ. خاصة وأن الاسم المركب يجمع في الغالب بين فصيلتين من الأسماء الموجودة في معجم المفردات

1 – محمد زايد، دراسة في المعاجم الفرنسية الحاسوبية، الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية، مدينة الملك عبد العزيز للعلوم والتقنية، أبريل 2008. ص ص.1-28 ، وينظر : عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، مرجع سابق، ص 102.

2- نريد من خلال هذا النمط إعطاء فكرة عن معجم المفردات البسيطة بأشكالها المعرّبة، لأننا نفترض تماما بأن بناء معجم المفردات المركبة يستند في جزء هام من بياناته عن متواليات البسيطة التي تكون عنصرا من مكوناته.

3- يقصد بالشكل: المدخل المعجمي(المفردة البسيطة)، وبالجذر: الصيغة التي يظهر بها في نفس المعجم، وبالصرف: مختلف الأوجه التصريفية التي تلحق المدخل البسيط، وهو ما يقصد به بالإعراب (Forme fléchie).

4- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، مرجع سابق، ص 102.

البسيطة: بين الاسم والصفة، أو بين الاسم والحروف، أو بين الاسم والظرف ، إلخ .هذا الجمع يعني أن الاسم المركب يتألف من فصيلة صرفية أيضا، وهو ما يتطلب معالجة خاصة[1] .

2-3-3-2- معجم التعبيرات المتلازمة:

عملا بمبدأ كون الوحدة المعجمية لا تكتفي بالعنصر المفرد والبسيط، بل تتجاوزها لتشمل الوحدات التركيبية التي تحتوي على أكثر من عنصر واحد، كما ترى نظرية المعجم-التركيبي، فقد أصبح لازما وضع معاجم تعتمد نظاما نتراصّا ومبنياً من التراكيب الثابتة والصمّاء التي تقف في وجه كل تأليف أو تحويل لأحدى عناصرها، مما يطرح مشكلة لا تقل أهمية وتتعلق بعملية الترتيب في المعجم. لأن المتواليات المراكبة هي متواليات لسانية تتألف من أكثر من عنصر مستمر كما يحتوي على أكثر من بياض واحد. فدماغ الانسان كما ذكرنا سابقا، مزود على المستوى الصوري بقدرات لسانية معقدة تأخذ على المستوى انجاز شكلين من التراكيب[2]:

يقوم الاول صوريا على مبدأ التوزيع التكاملي بين العناصر الاسمية والحرفية والفعلية ويتميز الثاني باستحالة التأليف بالتعتيم الدلالي وبعدم الاستجابة للقوانين للشكل الأول . لقد اعتمد الدكتور عز الدين غازي على منهجية المعجم – التركيبي من زاوية أساسية وهي إعداد النظر في مفهوم المفردة وقد تتضمن أكثر من متوالية لسانية متضمنة لأكثر من منطقتين فارغتين، لكن غير دالة على أكثر مفردة واحدة.

وهذا ما وضحه في المثال التالي[3]:

يضرب س.(أخماسا)1 في (أسداس) 2

تتأسس المفرد/ المتوالية: (أخماسا في أسداسا)على أكثر من بياضيين، ويمكن لأي برنامج توليد للمفردات اللغوية أن يدخل في الاعتبار أكثر من بياضين بينما تبقى البياضات الأخرى بدون أي دور توليدي في عملية التعرف عليها مقطع ذلك لأن هذا النوع من المركبات الاسمية لا يقبل التجزيئ إلى مكونات صغرى دالة تنفرد كل واحدة منها ببياضيها في الاستعمال العادي للغة العربية، بل ولا تسمح حتى باستبدالها بمتوالية لسانية بسيطة معادلة لها دلالياً ، فهذا النوع من المركبات الاسمية يبقى هكذا منطقة سوداء معتمة لا يمكن النفاذ إليها إلا بالإحصاء والتصنيف حتى تقدم إلى المعجم الآلي بشكل يسمح له بتوظيفها في العمليات التي ننتظر منه أن

1- عز الدين غازي ، بناء المعجم الآلي للغة العربية، معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص 204،205.

2- المرجع السابق، ص205.

3- المرجع نفسه، الموضع نفسه.

ينجزها[1]. غير أن هذا لا يجب أن يفهم منه أن كل التعبيرات المتلازمة قائمة على المتواليات الاسمية المركبة فهناك العديد منها تقوم على عناصر اسمية بسيطة، مثلا[2] .

ضرب س.(السمكة)

وقع س.(في الفخ)

شد س .(الرحال)1 (إلى مظ)2

فالمفردة المركبة،إذا ، هي متوالية لسانية اسمية مؤلفة من أكثر من مدخل معجمي[3] بسيط من مداخل معجم المفردات البسيطة. وهي تتألف تارة من عنصر وآخر جامد تارة أخرى من عنصرين جامدين. إن المفردة المركبة تارة لا تستجيب للتغيرات الصرفية أو حتى التوليد بظاهرة التوسيم مثلا[4]:

< س س >: = شذر مذر

< س س >: = حيص بيص

يصعب مع هذه الامثلة تحليل هذه المركبات تحليلا مورفو–تركيبيا، بل وحتى تلك المبنية على مصدر أو اسم الفاعل حيث يصعب كذلك العودة بها إلى البنية الأصلية. وذلك مثل:

مقود السيارة.

هناك فقدان للعلاقة الدلالية بين هذه المركبات وما قد يكون أصلا لها.

هذا ولقد عولجت هذه التراكيب بصفة معمقة في اللغة الفرنسية[5] وكذلك بعض محاولات في اللغة العربية[6] وقد خلص هؤلاء إلى أن هذه التراكيب تتصف بدرجة معينة من التلازم والتضام حيث منها المتلازمة جزئيا ومنها المتلازمة كليا[7].

1– المرجع نفسه، ص 206.

2 – محمد الحناش، المعاجم الآلية للغة العربية بناء قاعدة المعطيات، مجلة التواصل اللساني، م 4 ،ع.1 ،س.92 ،ص ص100–101.

3– مدخل معجمي: Entrée Lexical: و هو معروف عند القدماء بـــ "المادة "، ويعنى به الكلمة المدخل المرتبة ترتيبا معينا بالمعجم، لا سيما الترتيب الألفبائي، والمعروضة حسب ذلك الترتيب، مصحوبة بتفسيرها وبشرحها صوتيا وصرفيا ونحويا، ودلاليا وبلاغيا وأسلوبيا حسب الإمكان.للمزيد من بسط القول ينظر: محمد رشاد الحمزاوي، المعجمية،(مقربة نظرية ومطبقة، مصطلحاتها ومفاهيمها)، مركز النشر الجامعي، تونس، د.ط، 2004، ص 203.

4 –M.Elhannach, Lexique-Grammaire de l’arabe :classe des verbes qualitatifs,(suite)in linguistica communicatio,V.1,N.2.1989.pp.31-41.

5– مع مجموعة من الدارسين أهمهمJ1985 LA BELLE .M.SILBERZTEIN (1993). LA PORTE .E (1988).M.GROSS (1988)والباحثة الايطالية M CONENNA 1985

6– أنظر أحمد أبو سعد 1987 والحناش 1992 و1993 و2002.

7– إذا كانت المتلازمة كليا تتميز بعتمة منطقتها الثابتة، فغالبا ما يصعب ربطها بجملة وفي هذا الحالة ستصبح مدخلا مستقلا بنفسها رغم وجود بياضات أو عناصر مشتقة أو جامدة، منطقتها ترفض أي استبدال كان لأن عناصرها صامدة في وجه أي تغيير تركيبيي .وتتميز المفردة المركبة بوقوعها في الغالب في موقع المفعول وهي بذلك تبتعد عن منطقة تطابق مع الفعل.

لقد اعتمد الدكتورمحمد الحناش في هذا الإطار[1] إلى ادراج الأسماء المركبة الحملية المستقلة في هذا الباب والتي لا تعرف أصلا جمليا مثل :

أثار زيد حفيظة عمر.

إن التعبير المتلازم لا يترك عنصرا يندمج في مكوناته مثلا:

أخماسا في أسداسٍ

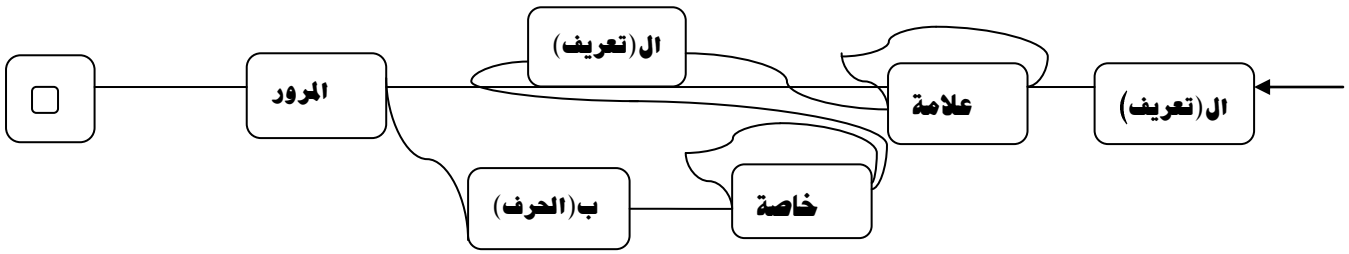
لكن *أخماسا X في أسداس

خلافا للمتلازمة جزئيا مثل علامة المرور .

علامة المرور

العلامة الخاصة بالمرور

وعليه يمكن تمثيل ذلك بالأتومات الأوضاع النهائية التالية[2]:



الشكل رقم: 8- يمثل تثنية وجمع بعض أصناف المفردات المركبة

ومنه استخلص الدكتور عز الدين غازي إلى هذه أن هذه المركبات غير متجانسة في تكوينها وفصيلتها التركيبية، لذلك وجب التعامل معها بتصنيفها ووصفها ثم معالجتها وفق طرق خاصة، فالمعجم يجب أيضا أن يشمل كل هذه الأنواع من المركبات بدون استثناء. فدراسة هذه التعبيرات وفق إطار المعجم –التركيبي وخاصة تقنيات التصنيف والتعرف على مناطقها المعتمة(بحيث لا تختلط بمعجم المأثورات والأمثال) تضمنت من الأسماء المتلازمة مع الفعل: أسماء بسيطة وأخرى مركبة منها (الاسم والضمير) (قضت هند نحبها)، الاسم المتجانس (وجها لوجه)، الاسم المتخالف(ذهابا وإيابا) المركب الإضافي (رد الاعتبار)، مركب حرفي(بمنأى عن كذا)، ومتلازم صفة (أصاب عمر مرض عضال)[3].

1 – M.Elhannach, Lexique-Grammaire de l'arabe :classe des verbes qualitatifs,(suite)in linguistica: communicatio,V.1,N.2.1989.pp.09-18.

2– عز الدين غازي ، بناء المعجم الآلي للغة العربية، معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص207.

3– محمد الحناش، نظرية حاسوب لسانية لبناء معاجم الآلية للغة العربية، محاولة في التأصيل (سلسة خاصة عن استخدام اللغة العربية في تقنية المعلومات، سلسة ندوات، المجلد1، مجلة التواصل اللساني .س.1992.ص71.

إن هذا يتطلب وصفا صرفيا آليا من نوع خاص لأنها بنيات تختلف عن باقي المفردات البسيطة، لذلك فإن استدعاء مفاهيم وأدوات إجرائية جديدة، تعالج البنية الصرفية للعناصر التي تتألف منها المركب وقد تكون من المفروض موجودة في معجم المفردات البسيطة المعربة[1].

غير أنها تختلف من حيث كونها تجمع بين المعجم والتركيب رغم أن الجانب الصرفي سيصبح لاحقا في وضع ثانوي بعد التركيب[2].

وصفيا يظهر إلى حدود الآن أن هذا النوع من البنيات اللسانية باستثناء المجهودات القليلة، لم يلق اهتماما رغم انتاجيته الكبيرة في قدرة المتكلم اللغوية والمعجمية بخاصة. وتحتاج هذه البنيات إلى بناء قواعد خاصة بها[3] .كما فعلت مجموعة(GETA4) وهذا ما يستدعي مزيدا من المعالجة بهدف تزويد تطبيقات الحاسوبية المتعددة للترجمة وتعليم اللغات والنشر والتدقيق الإملائي وذلك ببناء معجم آلي عربي يستجيب للتقنيات المتبعة في المعاجم الأجنبية من هذا الحجم،(انظر مثلا التعبيرات المتلازمة في المعجم الالكتروني في اللغة الفرنسية DELAC إذ نجد حوالي 20000 تعبير متلازم في قاعدة بيانات اللغة الفرنسية[5].

ولما كانت هذه التعبيرات خاضعة لجميع اللغات طبيعية لنفس عمليات التصنيف بحسب رؤية المعجم- التركيبي فإننا تبعا لذلك، سنحاول اختيار وتجريب التقنيات الذكية، إضافة إلى تقنيات الكتابة الأقواسية وطريقة الأجرميات التي اقترحها د.محمد الحناش[6].

باقتراح قواعد خاصة للمعالجة المعجمية النحوية المستفادة من المحللات التركيبية والصرافية والدلالية تراعي خصوصية ونحو كل صنف أولا وثانيا التعامل معها بطريقة الأتومات النهائي وأنواعه ولاسيما أتومات الإدماج التي تسمح بعض عقدها، حسب أنواع

1- إن غنى المتن المعجمي العربي و وجود الزخم الهائل من هذه التعبيرات المتلازمة والمركبة التي أغنتها بشكل واضح المداخل المولدة عبر الترجمة والتواصل بين اللغات، إذ نجد ألفاظ مركبة قد تكون انتقلت إلى العربية انطلاقا من لغات أجنبية ومنها المصطلحات العلمية بسبب المتطلبات التكنولوجية والعلمية والثقافية1، اضافة إلى المركبات الأصلية العربية المؤلفة من صفة واسم حرف واسم، اسم وصفة، حرف واسم وحرف، اسم وحرف واسم، اسم واسم، أشكال أخرى. كل ذلك يشكل قاعدة معجمية ضخمة يجب أن نستفيد منها في بناء معجم إلكتروني للمتلازمات العربية، حتى يمكننا تزويد مختلف التطبيقات الليسانية والحاسوبية والبرمجيات المناسبة لذلك، ينظر: انظر عز الدين البوشيخي، مجلة السان العربي ع49.ص14 .

2- عز الدين غازي ، بناء المعجم الآلي للغة العربية، معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص208.

3 – محمد الحناش، في قاعدة البيانات الخاصة بالتعابير المسكوكة والمركبة 1992 ونظام الترجمة المتعددة اللغات الذي انتجته شركة CIMOS للمزيد يرجى الاطلاع على الموقع الالكتروني:.www.cimos.com

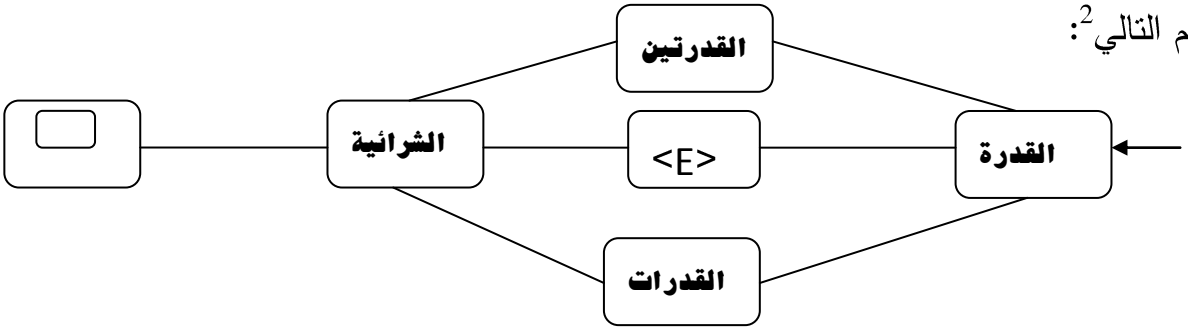
4 – د.محمود اسماعيل الصيني، التواصل اللساني مرجع سابق، 1993.

5 –Max Silberztein 1993

6- محمد الحناش ،بنامج اللساني الحاسوبي للتعرف الآلي على التعابير المسكوكة في اللغة العربية، مجلة التواصل اللساني .م.3سنة،1996، ص 86-88.

المركبات، إما بإدماج عنصر ما بين مكوناته أو تصريف أحد من عناصره كأن نثني ونجمع رأس التعبير التالي : القدرتين والقدرات على التوالي في[1] : على القدرة الشرائية

فأتومات الاوضاع النهائية يستطيع تمثيل هذه التراكيب آليا وفي مراحل إنتاجية معجمية (حالات إعرابية محلية).مختلفة لها تحتاج إلى خوارزمات متميزة كما في التومات ممثل في الرسم التالي[2]:



الشكل رقم 9- يمثل تثنية وجمع بعض الأصناف المفردات المركبة

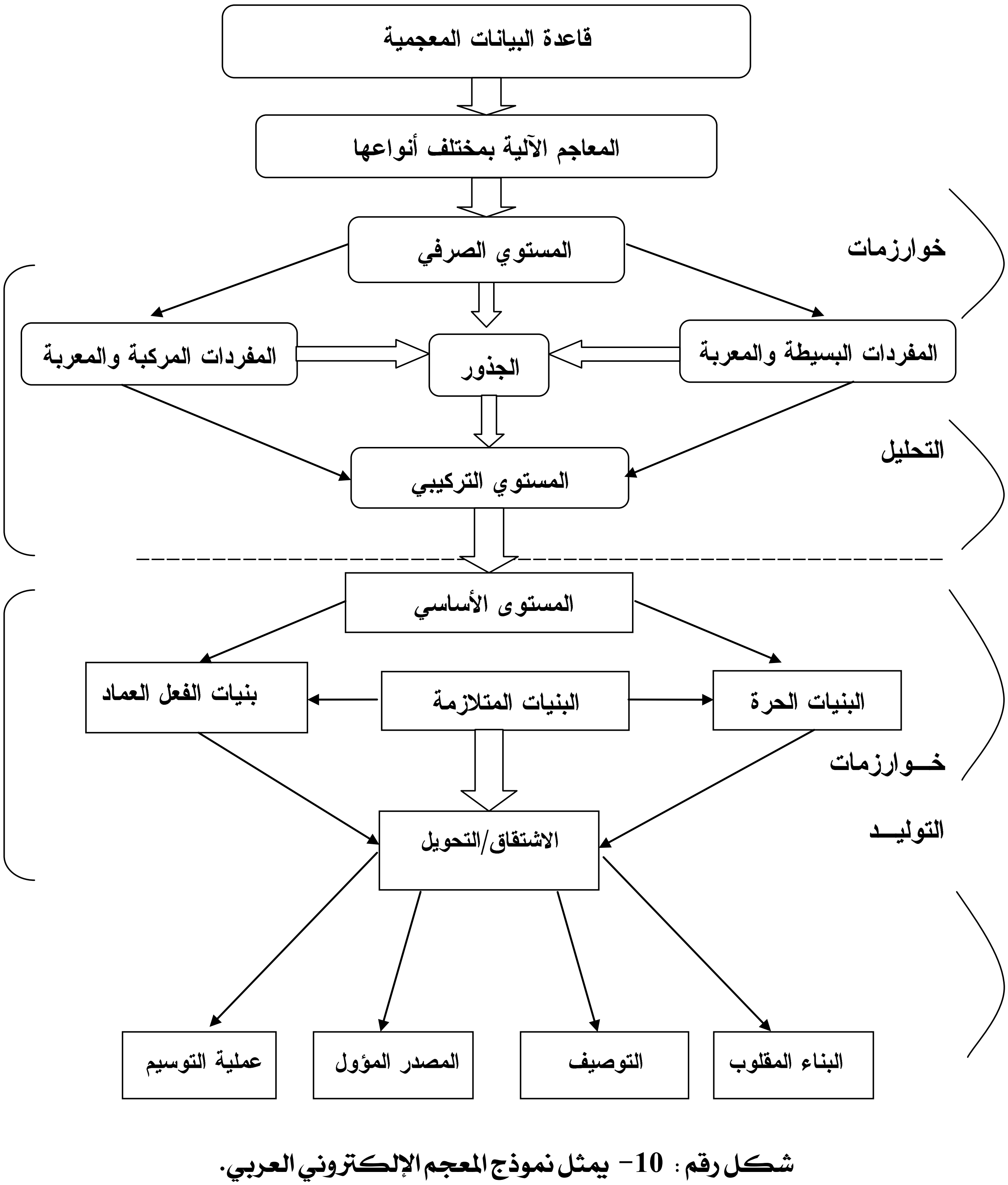
وإذا كان المعجم العربي الآلي الذي ننشده يأخذ في عين اعتبار المنهجية الفنية والتقنية التي اثرت بشكل كبير في تحديث المدخل المعجمي المتضمن البسيط والمركب وأن ذلك يستجيب للغة العقلانية المعلوماتية فإن بناء معجم آلي للتعبيرات المركبة والمتلازمة لابد أن يرتكزعلى منظومة من معاجم التي تعمل في تناسق تام ومكتمل ومُبنين يتمكن مُستعمله من التعامل معه بشكل تفاعلي ذكي، وقد صور الدكتور عز الدين غازي المعجم العربي على شكل التالي[3] كما في أنظمة اللغات الطبيعية المعدة لهذا الغرض، وبناء على هذه المعمارية اقترح الشكل التراتبي لمنظومة المعاجم الآلية كما يلي[4]:

1- عز الدين غازي ، بناء المعجم الآلي للغة العربية، معالجة التعبيرات المركبة والمتلازمة، مرجع سابق، ص209.

2- المرجع نفسه، الموضع نفسه.

3- المرجع نفسه، ص 210.

4- هذا الشكل التراتبي للمعجم الآلي العربي هو نسخة معدلة للشكل الذي اقترحه د محمد الحناش في بحث له حسب تطورات التي خضعت إليها منظومة معاجم DELA وخاصة مع ما جاءت به البيئة المجانية المفتوحة للمصدر (نوج). للمزيد انظر نظرية حاسوب لسانية لبناء معاجم آلية للغة العربية (محاولة في التأصيل، سلسلة خاصة عن استخدام نظرة عربية في تقنية المعلومات سلسلة ندوات المجلد 1 مجلة التواصل اللساني ص71.



شكل رقم : 10- يمثل نموذج المعجم الإلكتروني العربي.

ثالثا: بناء قاعدة بيانات معجمية لمعجم إلكتروني:

كيف نبني قاعدة بيانات معجمية للمفردات اللغوية موجهة لخدمة المعاجم الإلكترونية للغة العربية؟ ثم ما هي الإمكانات التي توفرها الآلة للمعجمي؟

3-1- قاعدة المعطيات المعجمية: Base de données lexicales:

يتطلب إعداد معجم إلكتروني للغة العربية مصادر تكون بمثابة قاعدة أساسية لثراء المعجم بمختلف المعلومات المراد البحث عنها ، فهذه المصادر تكون بمثابة المدونة اللغوية الأساسية للمعجم الإلكتروني، بعد إعداد المعطيات اللغوية أو ما يعرف بالمدونة اللغوية ، توضع في قاعدة معطيات Access[1]، وهذه القاعدة مستقلة عن برامج المعجم ، وبذلك يمكن تعديل محتواها (تغييرا وإضافة وحذفا) دون أن يؤثر ذلك في برامج المعجم[2].

فما المقصود بالمعطيات المعجمية؟ وكيف يتم تثبت الملفات فيها؟ وما علاقتها بقاعدة المعطيات النصية؟ وما دورها في بناء المعجم الإلكتروني؟

تحدد المعلوميات قاعدة المعطيات بأنها مجموعة من المعلومات الدقيقة وغير الحشوية والضرورية لسلسلة من التطبيقات الآلية، والتي تنبني على نظام رياضي منطقي يتولى تدبيرها[3].

إن بناء قاعدة معطيات[4] معينة عملية تنطلق من ملاحظة الحالة المدروسة للوصول إلى نتائج مضبوطة تشكل حجر الأساس لقاعدة المعطيات. وبهذا المعنى، إن قاعدة المعطيات هي كم من المعلومات المجمعة في ميدان معرفي معين بطريقة منظمة وصارمة في الآن نفسه[5].

1- قاعدة المعطيات Base de données : هي مجموعة برامج تقوم بتنفيذ وظائف استغلال المعطيات. هذه الوظائف تتمثل في إضافة حذف أو تغيير المعطيات ، الاستعلام أو البحث عن المعطيات وفق معايير معينة ، ترتيبها أو فرزها وعرضها في شكل معين.

ولإنشاء واستعلام قاعدة معطيات باستخدام البرامج ينبغي الإلمام بالعناصر التالية :

* الجدول Tableau
* الاستعلامات Requête
* الاستثمارات أو النماذج Formulaire
* التقارير Etat
* صفحات الويب Page web
* الماكرو Macros
* المقاييس Modules

إنّ هذه المعطيات يجري إدخالها وتخزينها بتركيبة معينة ترابطية سهلة الانتقاء والبلوغ.

ينظر : د الحسن الحسيني، (معجم مصطلحات علوم الكومبيوتر) ، سابق ، ص 143.

وكذلك على موقع الويب: عنوان الموقع : مفهوم قاعدة المعطيات : http://.www.aun.edu.eg

2- د.عمر المهديوي،المعجم الإلكتروني: قراءة في بعض التجارب الغربية والعربية،الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، مطبعة المعارف الجديدة ، الرباط، 2009، ص 287.

3- ينظر:عبد القادر الفاسي الفهري، المعجم العربي المولد، منشورات معهد الدراسات والأبحاث للتعريب،مطبعة معهد تعريب، 2002،الرباط، ص 126.

4- قاعدة المعطيات مرادف الكلمة معجم البيانات قاعدة بيانات تحتوي على بيانات حول جميع قواعد البيانات المؤلفة لنظام قاعدةالبيانات حيث يتألف المعجم من سرد لجميع أسماء الملفات والفهارس بالاضافة الى بنية قاعدة البيانات وأية معلومات أخرى متعلقة بصيانة وادارة قاعدة البيانات.

5- د.عمر المهديوي، المعجم الإلكتروني: قراءة في بعض التجارب الغربية والعربية،الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، مطبعة المعارف الجديدة ، الرباط، 2009، ص 288.

وحسب علم الجبر، فإن قاعدة المعطيات تعد بمثابة مجموعة نهائية تحتوي على عناصر مقرونة بخصائصها المشتركة، ووصفها يتم بطريقة محددة، وبناؤها يستند إلى معايير صارمة تشكل أساس التنظيم الصوري الذي يتحكم في محتوى المجالات غير الموحدة والملتبسة قصد تطبيقها على كل عناصر المجموعة المستخدمة في الوصف.

وعليه فإن هذا النوع عليه تقوم معالجة قواعد المعطيات الآلية بتطبيق وإنجاز أنظمة تدبير قاعدة المعطيات[1] المتخصصة، والتي تمكن من القيام بعمليات سريعة على مستوى المراقبة أو المعالجة، واسترجاع المعلومات إدخالاً وإخراجاً من وإلى قاعدة المعطيات.

تمثل قواعد المعطيات بالنسبة للمعجمية إحدى الركائز الأساسية التي ينبني عليها العمل المعجمي، وقواعد المعطيات تعد مهمة وأساسية بالنسبة لأي تخصص كيفما كان نوعه. بيد أن مفهوم (قاعدة المعطيات)[2] يتداخل مع مفهوم (بنك المعطيات)[3] ، إذ يقترح لهما برنار كيمادا تعريفين إجرائيين[4]، يرسم من خلالهما الحدود الفاصلة بين المفهومين، حيث يقول:

تحدد قاعدة المعطيات بأنها جهاز أو عتاد حاسوبي، نتاج تجميع مجموعة من المعطيات، أحداثاً (وقائعاً) كانت أم معلومات، تشترك في خاصية مشتركة ألا وهي أنها تكون مسجلة على الحاسب وفق قواعد التنظيم والترتيب والتنفيذ (الإدخال) المضبوطة وأن برنامج أو برامج المعالجة أو الاستفسار (المساءلة) تخضع بشكل خاص لهذه القواعد. بينما لبنك المعطيات بعد تكميلي أو إضافي، إنه يتناسب مع المؤسسة التوثيقية الآلية التي تدبر أو تشغل جهازاً أو أجهزة في بناء المعطيات وموجهة لخدمة جماعة من المستفيدين، ولها أهداف خاصة تتمثل في جمع وحفظ المعلومات المجمعة، والأكثر شمولية (استيعاباً) بهدف تأمين الوثائق المستعملة من قبل المستفيدين أو المنفذين. إنها تتوخى حل مشاكل التدبير والتسيير الإداري والقانوني والتي يبقى بعضها بدون حل، ولاسيما ما يتعلق بالجانب المالي، والأمن، والخاصية الافتتاحية والعلمية، ...الخ.

إن قاعدة المعطيات فكرة مبتذلة أو تافهة من حيث المبدأ، ولكن ما يهم المعجمي بشكل أساسي هو العتاد المزدوج: فمن جهة هناك النصوص (نصوص أو عناصر نصية عادية، نصوص طبيعية، مواد قاموسية، ...الخ)[5]. ومن جهة أخرى، هناك رهان خوارزميات المساءلة أو الاستفسار (برامج، قواميس آلة أو آلية، أنظمة توثيقية)، والتي انطلاقاً منها تصبح المعالجات

1- نظام تدبير قاعدة المعطيات – (DBMS) Data Base Management Systems

2 – Base de données.

3 – Banque de données.

4 –Bernard Quemada, Bases de données informatisées et dictionnaires, Lexique 2, Presses Universitaires de LILLe, p. 105.

5- إن الوضائف الأساسية التي تقوم بها نظم قواعد النصوص يمكن حصرها في :

(إدخال البيانات يدويا أو آليا، حفظ النصوص وضغطها، تحليل مضمون النصوص-إحصائيا-، الفهرسة الآلية للنصوص،المحافظة عي سرية النصوص، عمليات دمج ملفات النصوص..)، للمزيد من بسط القول ،ينظر : نبيل علي،اللغة العربية والحاسوب، دراسة بحثية، تقديم د.أسامة الخولي،دار تعريب، الكويت ،1988، ص 329.

أو التحليلات الآلية للنصوص سهلة ومتيسرة للغاية. على أن درجة تعقيدها تحدد نوعية ومستوى تطوير النتائج المحصل عليها.

وبالنسبة للأعمال المعجمية، إن التحليلات تنقسم إلى نوعين: التحليلات البسيطة من النوع التسلسلي/ التعاقبي، تمكن من كشف أو استكشاف المتتاليات أو تلازمية الكلمات حسب ترتيبها في الخطاب (توزيعات، وتواردات)، والنوع الآخر من التحليلات يمكن الحصول عليه من العناصر التي تنتمي إلى هذه الأصناف، والمقالات المعجمية، ...الخ.

وانطلاقاً مما سبق، يعتبر المعجم الالكتروني بمثابة قاعدة معطيات لاعتبارات عدة نحصرها فيما يلي[1]:

- أنه يحيل على مجال معرفي معين له صلة مباشرة باللغة، وبشكل محدد المعجم؛
- أن بنيته مبنية على شروح وتفسيرات خاصة بالمظاهر الاستعمالية للمعجم؛
- أن العناصر التي يتضمنها هي وحدات معجمية تشترك في خصائص مشاركة؛
- أنه قابل للاستخدام في الدعامات والركائز الاشعاعية، وفي أنظمة التدبير المنطقي، وبمعنى أوضح برامج التحليل اللساني.

يعتبر المعجم الإلكتروني إذن بمثابة قاعدة معطيات معجمية قابلة للاستخدام الوظيفي كمحرك بحث بالنسبة لبرامج اللسانيات على الشابكة.

كما يعد أيضاً أداة معلومياتية، إذ أنه يفيد في تحقيق العديد من التطبيقات[2] ، نذكر على سبيل الذكر أن حوسبة الكتابة تقدم خدمات إيجابية تتمثل في سرعة الأداء، والدقة في الإنجاز، وضعف في الكلفة[3].

1- عمر المهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، مرجع سابق، ص 94.وينظر: عمرالمهديوي، المعجم الإلكتروني: قراءة في بعض التجارب الغربية والعربية،الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، مطبعة المعارف الجديدة ، الرباط، 2009، ص 290.

2- إن استخدام الحاسوب بالنسبة للمعجمي لا ينحصر فقط في رصد الخصائص المميزة لهذه التقنية بقدر ما يتعداه إلى الإمكانية التي يتيحها الحاسوب على مستوى تطبيقات الآلية، إذ يساعده على تنظيم بنوك المعطيات الضخمة، وهو ما يتيسر له القيام به اعتمادا على ذاكرته المحدودة.

ومن بين الإمكانيات التي تتيحها الآلة للمعجمي نذكر ما يلي:

- تحرير النص على الآلة.
- استرجاعه بكل يسر وفي أسرع وقت ممكن.
- ضعف كلفة تخزين المعلومات وتكشيفها.
- استخدام المنهجية الصورية (الخوارزميات) في معالجة النصوص.
- الاستفادة من تقنية المدقق الإملائي والنحوي في مراجعة النص المعالج.
- توظيف تقنية القارئ الضوئي.
- التمكن من تجاوز المراحل التقليدية التي كان يعتقد خطأ أنها إجبارية في معالجة نص من النصوص.
- يمده بطيف من المعلومات التي تسهل عمله.
- يساعد على التسريع في إنجاز العمليات والقيام بالتعديلات الممكنة إما بالحذف أو الزيادة في محتوى الكتاب، مما يسمح بتطوير العمل المعجمي وبلورته في أية لحظة.

3- عمر المهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، مرجع سابق، ص 95.

3-2- خصائص قواعد المعطيات المعجمية الآلية:

تنقسم قواعد المعطيات الآلية إلى أصناف وأنماط، كل صنف له خصائصه الذاتية التي يتقاسمها مع الصنف الآخر، نذكر منها ما يلي[1]:

أ- بنك المعطيات المعجمية: تحدد بنوك المعطيات[2] بأنها مجموعة من الأدوات والتقنيات التي تقبل التطبيق بسهولة على الإنجاز المعجمي. وتتسم بالسمات التالية[3]:

- **الإجبارية:** يتضمن المعجم الأحادي أو الثنائي أو المتعدد اللغة، اللغوي أو الموسوعي، العام أو المتخصص (جذور، عبارات، حكم وأمثال ...)، معلومات نحوية صرفية (فعل، نوع، ...، طريقة النطق، إملاء، إيقاع)، ودلالية: (ترادف، اشتراك، تضاد، ...). وهي مجبرة للاستفادة من تقنية المعلومات وما تتيحه من إمكانيات وافرة فيما يتصل ببنوك المعطيات.

- **البنية:** تتسم بنوك المعطيات ببنيتها العلاقية أو الترابطية، وهذا يسمح بتمثيل المعلومات بشكل أوضح وأدق.

وتعد بنوك المعطيات المعجمية في الأساس، بنوكا نصية مختصة موجهة لخدمة التحليلات المعجمية، وهي كما يظهر من تسميتها تعنى بالعتاد الحاسوبي الذي يستجيب لمطالب الملفوظات بنوعيها المكتوب والمسجل ذات الخصائص خارج لسانية التي تؤشر وتحدد ضوابط الإنتاج، والغايات، والنقل، ومختلف التأليفات الممكن قيامها بين هذه العناصر برمتها، فمعجم النصوص الموجهة للمعالجة الآلية تتطلب تهيئة وإعداداً خاصاً، قبل مساءلتها بواسطة البرامج المعدة لهذا الغرض[4].

ب- بنك النصوص: les banques de textes:

تمثل تجميعاً لأكبر عدد من النصوص المتضمنة لجميع أصناف الأشكال والصيغ والوثائق المكتوبة أو المنسوخة والمضبوطة وفق المعايير التي تحقق مطالب الجماعة المستعملة[5].

تهدف "المكتبات الإلكترونية" على تغذية وإثراء قواعد المعطيات النصية بالمعلومات الأولية، حيث تضطلع الوثائق النصية بدور تجميع الوثائق الآلية وحفظها[6]. إن بنوك النصوص تختص بالنصوص العلمية والتقنية والمنهجية، ونأخذ مثالاً لذلك أرشيفات ووثائق **جامعة أكسفورد** المخزنة في بنوك نصية، وبنك النصوص الفرنسية **لمعهد اللغة الفرنسية** القديمة والحديثة والمعاصرة.

1- عمر المهديوي، المعجم الإلكتروني: قراءة في بعض التجارب الغربية والعربية،الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، مطبعة المعارف الجديدة ، الرباط، 2009، ص 291.

2- ليلى المسعودي، علم المصطلحات وبنوك المعطيات، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 28، الرباط، المغرب، 1987، ص 86،87.

3 -René Moreau et Isabelle Warnesson ; Ordinateur et lexicographie, Lexique 2, Presses Universitaires de LILLE ; 1983, p. 126.

4- عمر المهديوي، المعجم الإلكتروني: قراءة في بعض التجارب الغربية والعربية،الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، مطبعة المعارف الجديدة ، الرباط، 2009، ص 292.

5- Bernard QuemadaIbid, p. 108.

6- عمر المهديوي، المعجم الإلكتروني: قراءة في بعض التجارب الغربية والعربية، مرجع سابق، ص292.

إن وظائف بنوك النصوص لا تنحصر فقط في المهام والأغراض التقليدية للمكتبات، وإنما ترتبط بمهام أخرى لها صلة بالتشغيل الحاسوبي في حالة النصوص على الآلة. إن مشكلة بنك النصوص إذن تتجاوز مرحلة القدرات والاهتمامات الخاصة بالمعجميين والقاموسيين[1].

ج- قواعد المعطيات النصية les bases de données textuelles: تعتبر بنوك المعطيات[2] النصية مكوناً مهماً وأساسياً في المعالجة الآلية للغات الطبيعية، ولا سيما الحوسبة المعجمية، إنها تستجيب لمطلب العتاد الحاسوبي[3].

3- 3- تنظيم المعجم الإلكتروني للغة العربية في هيئة قاعدة بيانات معجمية:

يتم تنظيم المعجم الآلي في صورة قاعدة بيانات لتنويع خدمات البحث، ولإتاحة استخدامه لفئات المستخدمين على اختلاف مستوياتهم، وكذلك لتسهيل عملية دمجه مع النظم الآلية على اختلاف أغراضها وتصميماتها.

يجب أن يتوفر في بناء قاعدة بيانات المعجم العربي عدة شروط أساسية، من أهمها[4]:

- استغلال خاصية **الفائض المعجمي**[5] تحاشيا لـــ "حشو" قاعدة البيانات بالمعطيات المتكررة التي يمكن استنتاجها بتطبيق **القواعد المعجمية**[6] المطردة[7].

1- المرجع السابق، ص292.

2- ليلى المسعودي، علم المصطلحات وبنوك المعطيات، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 28، الرباط، المغرب، 1987، ص 86،87.

3- المعجم الإلكتروني: قراءة في بعض التجارب الغربية والعربية، مرجع سابق، ص292.

4- نبيل علي، اللغة والحاسوب، سابق، ص 513،514.

5- الفائض المعجمي لا تظهر على مستوى المفردات بل على مستوى خصائصها الدلالية التفصيلية، من أمثلة هذا الفائض أن كل الأفعال ذات السمة [+إجرائي] تتطلب فاعلاً ذا سمة [+حيوي] بغض النظر عن خصائصها الدلالية الأخرى، للمزيد من بسط القول ينظر: اللغة والحاسوب، نبيل علي، مرجع سابق، ص 463،469

6- يفرق الفاسي بين نوعين من القواميس :قاموس واسع ،ومقعد.فالواسع يضم لائحة كبيرة من المعلومات ولائحة طويلة من المداخل الحشوية وهي ذاكرة لا تتوفر في الحواسيب الشخصية فهو إذن ذاكرة معلوماتية ضخمة للمفردات .

أما المقعد فإنه يستغني عن كثير من المعلومات ،والمداخل الحشوية التي نصادفها في القاموس الواسع ؛فهو يستغني عن المعلومات المكررة فهو إذن بهذا المعنى ذاكرة مختصرة للمعلومات ، وحاذفة للمداخل الحشوية .

وينبغي ها هنا قبل بناء معجم عربي محوسب تحديد نوع القاموس المراد بناءه أهو واسع أم مقعد؟ ،ثم تحديد المادة اللغوية المراد حوسبتها ؛بمحددات زمنية ومكانية ، مع الانفتاح على ما استجد من مصطلحات فرضها التطور التكنولوجي والصناعي والتقني المعاصر .

وتهدف هذه البحوث المعجمية حسب الفاسي الفهري تحقيق الغايات الآتية :

•إعادة بناء أنساق اللغة العربية صوتيا وصرفيا وتركيبيا ودلاليا ،وإصلاح ما أفسده النحاة قدامى ومحدثين .

•إعادة المعلومات التصورية والانتقائية والمحورية للغة العربية من الناحية الدلالية

المراد من هاتين الغايتين فصل بين عمل المحلل النحوي والمحلل المعجمي ،وإن كان هناك صلات وصل على المستوى الدلالي كظواهر اللواصق كالمطاوعة والانعكاس والتفاعل وغيرها .

7- المقصود بقواعد الحشو المعجمي كمفهوم توليدي تحويلي تلك الاستراتيجيات التي تمكننا من التمييز بين عمل المحلل النحوي والمحلل المعجمي وهو ما دعي بقواعد الحشو المعجمي تشومسكي (1967)و ليز (1960)ودجاكندوف (1975)والفاسي الفهري (1985)

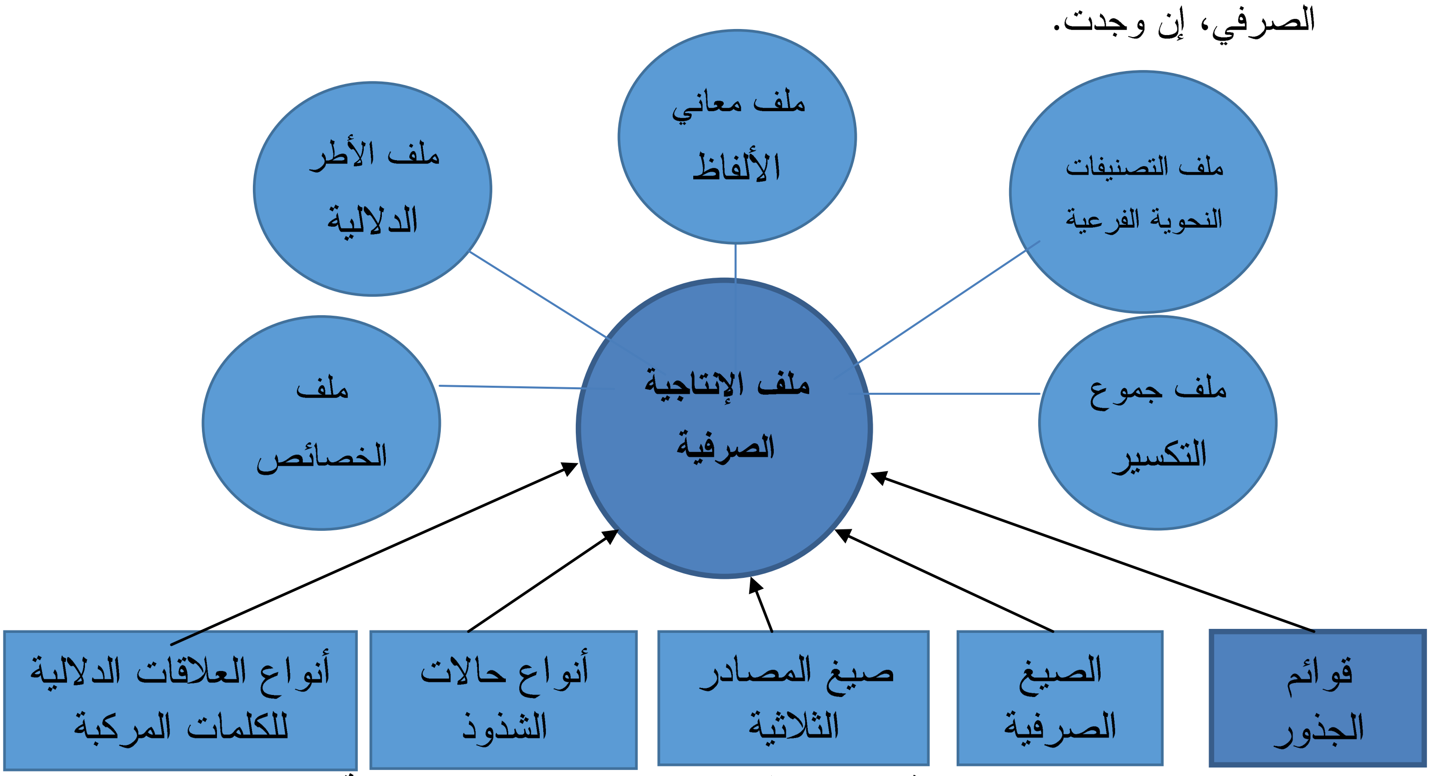
وتروم هذه القواعد الحشوية البحث عن مقاييس تسمح بحل مشكل التمايز بين المعجم والتركيب .ومعروف أن المفردات متعالقة اشتقاقا وصرفا ودلالة ومعجما، ويستدعي هذا البحث عن استراتيجيات في مقاربة العلاقة بين الخصائص التركيبية ودلالتها أي "القيام بتصنيف المفردات إلى طبقات باعتماد مقاييس تركيبية محضة على أمل أن هذا التصنيف ستدعمه الدلالة ،باعتبار أن هذه الطبقات لها عناصر دلالية مشتركة ".

إن كل مفردة بهذا المعنى ؛لها خصائص تركيبة اشتقاقية أو صرفية فكتب ،كتابة وكاتب ومكتوب ومكتب وفلان أكتب من فلان ،إن هذه التشكيلات المعجمية تحيل بالضرورة إلى دلالات مختلفة ؛مما يدل على دور القواعد المعجمية في تمكيننا من تفادي الكثير من التحشيات والتكرار المفرط ،لا سيما وأن الغاية هو بناء معجم عربي محوسب ؛ مما يمكن من سهولة الاكتساب الفهم والإنتاج،على اعتبار أن"تقدير كمية المعلومات يمكن أن يستخلص من عدد رموز المعجم وتتم المفاضلة بين المعاجم المتكافئة على أساس أن المعجم الأقل رموزا هو المعجم الأكثر قيمة "

- خدمة أغراض البحث المختلفة من بشر ونظم آلية.
- كفاءة عالية في تصميم الملفات لضغط بياناتها نظرا لضخامة حجمها، وتعقد العلاقات التي تربط بينها.
- التركيز على العلاقات المحورية (الخاصة بمصفوفات العلاقات).
- التخلص من طغيان الخاصية الصرفية بتضمين المعجم المعطيات النحوية والدلالية الخاصة بالمفردات.
- تصميم قاعدة البيانات المعجمية بحيث يمكن استخلاص مجموعة معاجم فرعية منها، كمعاجم الموضوعات المتخصصة، ومكانز الكلمات المفتاحية العامة والمتخصصة، ومعاجم للصغار، ومعاجم الترادف والتضاد.
- مراعاة سهولة التحديث تعديلا وحذفا وإضافة.

يمثل شكل ( ) **الإطار العام لقاعدة البيانات المعجمية**، والتي تشتمل على المجموعة التالية من الملفات والقوائم المترابطة مع بعضها[1]:

*/ **ملف رقم 1**: ملف الإنتاجية الصرفية[2]، وهو ملف يتضمن سجلا لكل جذر، ويتضمن هذا السجل كل الصيغ الصرفية المنطبقة على هذا الجذر، وبعض بيانات عن حالات الشذوذ الصرفي، إن وجدت.



شكل رقم 11: الإطار العام لقاعدة البيانات المعجمية

1- اللغة والحاسوب، سابق، ص 514.

2- مصفوفة الإنتاجية الصرفية هي مصفوفة ثلاثية الأبعاد، بعدها الأول هو الجذر، وبعدها الثاني: هو الصيغ الصرفية المنطبقة على هذا الجذر، ويقترح ترتيب هذه الصيغ في مجموعتين لصيغ الأفعال والأسماء، أما البعد الثالث فهو المعاني الصرفية لصيغ المزيدات في حالة الفعل، والتصنيفات الدلالية للأسماء والصفات في غير ذلك. ينظر، اللغة والحاسوب، مرجع سابق، ص512.

*/ **ملف رقم 2: ملف جموع التكسير**[1]، ويتضمن قائمة من المدخلات، يتكون كل مدخل من جمع تكسير مقرونا بمفرده، علاوة على جواز الجمع بين جمع التكسير وجمع المذكر، أو جمع المؤنث السالم.

*/ **ملف رقم 3: ملف الخصائص الدلالية للمفردات**، وهو ملف يخصص سجلا لكل مفرد، أو عدة سجلات له في حالة تعدد معانيه، ويتضمن السجل السمات الدلالية والنحوية

1- نقف هنا قليلا عند مشكلة جمع التكسير في "العربية"، وهي المشكلة التي لم تحظ عناية – رغم تفشيها – وكثرة الحديث عنها – بما تستحقه من دراسة مستفيضة ومتعمقة. وربما كان السبب في ذلك هو عدم توفر الإحصائيات، وغياب عنصر الدلالة، حيث اقتصرت معظم الدراسات على الجوانب الصرفية لهذه المسألة ذات الأبعاد المتعددة. نستثني من ذلك الدراسة التي قام بها "محمد عبده" من جامعة ستانفورد والذي سعى إلى وضع هيكلة عامة لصيغ جمع التكسير، وتصنيفها حسب الأسماء والصفات، وربط هذه الصيغ بخصائص مفرداتها الفونولوجية، والتصريفية كالتأنيث الحقيقي والمجازي واللفظي.

هناك عدة عوامل تدخل في اشتقاق جمع التكسير من مفرده، نلخصها فيما يلي:

(أ) ارتباط جمع التكسير بصيغة المفرد في حالتي المذكر والمؤنث كأن يجيء جمع "فعلان" الذي مؤنثه "فعلي" على صيغة "فعالى"، وينطبق في هذه الحالة المذكر والمؤنث (مثال: سكارى جمع "سكرى' أو "سكران").

(ب) ارتباط جمع التكسير بتاء التأنيث الحرفية الزائدة، حيث يقتصر جمع التكسير هنا أيضاً على المذكر دون المؤنث (من أمثلة ذلك: "دور" وجمعها "أدوار" – "دورة" وجمعها "دورات"، و"مكان" وجمعها "أمكنة" – "مكانة" وجمعها "أمكنة").

(ج) ارتباط جمع التكسير بتاء التأنيث الحرفية الزائدة دون اقتصاره على المذكر فقط، كما في (أ) و(ب) ٠لاه (مثال: "طريق" وتجمع "طرق"، "طريقة" وتجمع "طرائق" أو "طرق").

(د) ارتباط جمع التكسير بجنس مفرده البيولوجي، حيث يقتصر عادة على المذكر دون المؤنث الذي يجمع في هذه الحالة عادة على هيئة جمع مؤنث سالم، أو بصيغة جمع تكسير خاصة به (مثال1: "أرنب" وجمعها "أرانب"، "أرنبة" جمعها "أرنبات"، مثال2: "شاعر" وجمعها "شعراء"، "شاعرة" وجمعها "شواعر").

(هـ) ارتباط جمع التكسير بخاصية العاقل وغير العاقل في الأسماء أو الصفات، حيث يختلف جمع التكسير وفقا لهذه الخاصية (مثال: "طراق" جمع "طارق"، "طوارق" جمع "طارق" لغير العاقل).

(و) اختلاف جمع التكسير أحيانا مع تغير معنى اللفظ في حالة الاشتراك اللفظي (من أمثلتها: "أمر" وجمعها "أوامر" أو "أمور" وفقا لمعناها).

(ز) اعتماد صيغ الجمع على صيغة المفرد ودلالته، كأن يجمع الوصف المفرد من صيغة فعيل والدال على هلاك على صيغة فعلى (مثل: "جرحى" جمع "جريح").

جرت العادة في دراسة ظاهرة جمع التكسير في العربية على النظر إليها كظاهرة ثنائية الأبعاد حيث يرتكز الحديث عن تناظر صيغ المفرد مع صيغ جموعه. إلا أن الظاهرة أعقد من ذلك بكثير، حيث هي ظاهرة ذات خمسة أبعاد رئيسية:

البعد الأول: الجذر

البعد الثاني: الباب الصرفي (اسم فاعل، اسم مفعول، صيغة، مبالغة، ...) حيث ترتبط صيغة الجمع عادة بالباب الصرفي، فعلى سبيل المثال، يكثر جمع صيغة "فعيل" على صيغة "فعلى" عند استخدامها في اسم المفعول، وعلى صيغة "فعال" عند استخدامها في الصفة المشبهة .

البعد الثالث: صيغة المفرد.

البعد الرابع: معنى المفرد حيث يمكن ان يختلف جمع التكسير في حالة الاشتراك اللفظي للمفرد (مثال: "أمر" وتجمع على "أوامر"، "أمر" بمعنى شأن وتجمع على "أمور").

البعد الخامس: صيغة جمع التكسير

وقد أجريت دراسة إحصائية على جموع التكسير في المعجم "الوسيط"، نورد هنا عينة مما وصلت إليه من نتائج، والتي تمثل أكثر أزواج صيغ المفرد والجمع شيوعاً لكل باب صرفي.

وعلى ما يبدو، فالمعنى هو المدخل لفهم ظاهرة جمع التكسير في العربية، والحالات الواردة في (د)، (هـ)، (و)، (ز) أعلاه تؤكد هذه الحقيقة بدرجات متفاوتة، علاوة على ذلك، فإن ربط جمع التكسير بمعنى مفرده يجيء مع النظرية الصرفية الحديثة التي تعتبر الكلمة مرتبطة بمعناها هي الأصل في تكوين الكلمات.

المرتبطة بالمعنى المحدد للمفرد، والمعطيات الخاصة بالإزاحة الدلالية لمعناه عن المعنى الصرفي (الصوري) لصيغته الصرفية.

***/ ملف رقم 4: ملف التصنيفات النحوية الفرعية للأفعال والأسماء والصفات وحروف الجر[1]–**

***/ ملف رقم 5: ملف الأطر الدلالية للأفعال والصفات[2] –**

تدعم قاعدة البيانات المذكورة مجموعة من قوائم المعطيات الثابتة، وتشتمل على قوائم الجذور، والصيغ الصرفية، وصيغ المصادر الثلاثية، وبيان حالات الشذوذ، وأنواع العلاقات

1– التصنيفات النحوية الفرعية" ويقصد بها الأنماط التي تحدد السياق النحوي الذي ترد به المفردات داخل الجمل.

2– المعطيات الدلالية: ويقصد بها توصيف لموضوعات (أو عناصر الإسناد) التي تقترن بالمحمولات، أفعالاً كانت أم صفات، ولتوضيح ما نقصده بالمحمول والموضوع هنا نذكر أن الفعل بصفته "محمولا" (أو دالة الإسناد) يمكن أن يكون له موضوع واحد (أو عنصر إسناد) في حالة لزومه وهو الفاعل في هذه الحالة، أو موضوعان في حالة تعديته، وهما الفاعل والمفعول، أو ثلاثة موضوعات في حالة التعدية المضاعفة. ويتخذ تمثيل المعطيات الدلالية صورة ما يعرف بـــ "الأطر الدلالية"، والتي تحدد الأدوار الوظيفية والسمات الدلالية لموضوعات الفعل ومكملاته، أو موصوف الصفة وملحقاتها، ويتم تحديد الأدوار الوظيفية وفقا لقائمة موحدة تشمل الوظائف المنطقية (الحقيقية)، ويقصد بالمنطقية هنا تلك التي تتجاوز صورية التصنيفات النحوية مثل الفاعل، والمفعول به، والمضاف، والمضاف إليه، والمجرور وهلم جرا ، فعلى سبيل المثال، "زجاج" في "انكسر الزجاج" فاعل نحوي ولكنه ليس الفاعل المنطقي (أو الحقيقي). تشمل قائمة الأدوار الوظيفية وظائف مثل: منفذ، معان، متأثر، مفعول به، متلق، ضحية، محور، أداة، هدف، مستفيد وإليك بعض أمثلة الأطر الدلالية لبعض الأفعال العربية.

الفعل	إطاره الدلالي
عاتب	(المنفذ [+عاقل] – الملتقي [+عاقل] – المحور [+مجرد] مثال: "عاتب الرجل صديقه على تقصيره"
يفكر	(المعاني [+عاقل] – المحور [+مجرد] – الأداة [+عنصر بشري]) مثال: "يفكر الفيلسوف في الأمر بعقله"
دمر	(بمعناها الحرفي) (متأثر [+جماد]، الأداة [–مجرد] مثال: "دمرت المدينة بالقنابل".

يرى البعض أن هذه الأدوار الوظيفية ذات طبيعة عامة تشترك فيها جميع اللغات، وهناك عدة تحليلات لها تختلف من حيث عدد الأدوار ومسمياتها، وتستخدم هذه المعطيات الدلالية في تحديد عناصر البنية الوظيفية للجمل،أي تحديد الفاعل الحقيقي والمعاني والمتأثر، والمفعول والمتلقي والمستفيد، والأداة، والهدف، وخلافه، وتعد هذه العناصر مدخلات أساسية لنظم الفهم الأوتوماتي، فهي الوسيلة العملية للنفاذ إلى البنية الدلالية العميقة للجمل، والتي لا تعكسها عادة بنيتها النحوية السطحية.

تستخدم الأطر الدلالية أيضاً مع التصنيفات النحوية الفرعية في تطبيق قيود الانتقاء الدلالي في المعالجات النحوية الآلية، وذلك لضمان توفق الأفعال مع فواعلها ومفعولها، أو مفاعيلها، ومكملاتها، وذلك لمنع توليد جمل غير معقولة، والمثال التقليدي لها هو المثال الشهير لـــ "تشومسكي": "نامت الأفكار الخضراء عديمة اللون في غضب".

ومرة أخرى، وفيما يخص النصوص العربية غير المشكولة، نشير هنا إلى أهمية تطبيق قيود الانتقاء الدلالي والتصنيف النحوي معا في إجلاء اللبس الناجم عن غياب التشكيل أو نقصانه، وإليك مثلاً لتوضيح المقصود بذلك:

"يوشك السمك الكبير أن يأكل السمك الصغير، وفك فمه مفتوح ليلتهمه"

في هذه الجملة غير المشكولة هناك لبس في كلمتي "السمك"، و"فك"، فالكلمة الأولى يمكن أن تفسر بأنها "السَمَك"، الحيوان المائي المعروف، أو "السُمك" وهو خاصية فيزيائية للأشياء الجامدة. أما كلمة "فك" فيمكن أن تقرأ على أنها الفعل "فَكَّ" أو الإسم "فَك" الدال على جزء عضوي من رأس إنسان أو حيوان. على ضوء التصنيفات النحوية والدلالية للفعل "يأكل" يتم اختيار كلمة "السَمَك" في الحالتين حيث يتطلب هذا الفعل المذكور فاعلا حيا [+حي]، ومفعولا له السمة [+غذاء]، أما بالنسبة لكلمة "فك" فيرفض النظام احتمال أن تكون فعلاً، حيث ينتقي الفعل "فك"، عادة، مفعولا به جامدا، أي لا يمكن أن يكون "ضحيته" عنصرا لكائن حي (فم السمك في حالتنا هذه).

ربما يأتي البعض بأمثلة تناقض ذلك، كما في التعبير الإصطلاحي: "فك رقبة" على سبيل المثال، وتبرز بذلك قضية أخرى ذات أهمية بالغة لنظم الفهم الأوتوماتي ألا وهي: قيام الاستخدام المجازي بتقويض أثر قيود الانتقاء الدلالي في فض اللبس اللغوي بصفة عامة، وذلك الناجم عن غياب التشكيل بصفة خاصة.

الدلالية التي تربط بين الأسماء المركبة، وما شابه. تقلل هذه القوائم ذات البيانات الثابتة من حجم الملفات الرئيسية بدرجة كبيرة، فهي تمكن من استخدام الأكواد الرقمية في كثير من حقول هذه الملفات بدلا من استخدام البيانات الأصلية المطولة[1].

وكما يؤكد الدكتور نبيل علي أن، هناك نوعان أساسيان من أنواع نظم قواعد البيانات[2]:

- نظم قواعد البيانات العلاقية.
- نظم قواعد البيانات الهيكلية أو السلمية.

تتضمن نظم قواعد البيانات العلاقية بسهولة تمثيل العلاقات المختلفة التي تربط بين ملفات قاعدة البيانات، وسجلاتها، وبمرونة عالية في صياغة طلبات البحث عن المعلومات التي تضمنها قاعدة البيانات، يعيب نظم البيانات العلاقية عدم قدرة المصمم على التخلص نهائيا من تكرارية البيانات واستغلال خاصية الفائض المعجمي[3]، في المقابل تتميز قواعد البيانات الهيكلية بقدرتها العالية في التخلص من تكرارية البيانات، أي في استغلال خاصية الفائض المذكور، في حين يعيبها عدم المرونة، وصعوبة عمليات التحديث، وعدم القدرة على التشكل ديناميا مع مطالب المستخدم. يميل المؤلف إلى التوصية باستخدام نظم قواعد البيانات العلاقية في ميكنة المعجم العربي نظرا لشدة العلاقات التي تربط بين مفرداته وعناصره.

3 – 4 – : تنظيم المعجم العربي الآلي في هيئة قاعدة المعارف:

يرتكز البحث المعجمي الحديث على تحليل تعريفات مفردات المعجم سعيا إلى الكشف عن البنية الدلالية الدفينة له، وهناك قناعة حاليا بإمكانية دراسة تعريفات المعجم كما تدرس فصائل النباتات والحيوانات، وما إلى ذلك ، إن المعجم هو وعاء لمعارف العالم، وما دام هذا العالم، ومعارفه، يدين للتحليل البنيوي، فلابد أن تعريفات المعجم تخضع لهذا التحليل أيضاً، أو بقول آخر، إن وراء عفوية المعجم واعتباطيته الظاهرية تكمن هيكلية متعددة المستويات، وشبكة مفاهيم متشعبة تمثل "الكيان الحديدي" الذي يقيم صلب المعجم، إذن: **كيف يمكن الوصول إلى هذه البنية الداخلية المعقدة؟**

1– اللغة والحاسوب، سابق، ص 515.

2– المرجع نفسه، الصفحة نفسها.

3– نبيل علي، المعجم من منظور الترجمة الآلية، الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، الرباط، المغرب، العدد الثالث والرابع، 1425هـ ،2005م، ص 17،18.

يتم ذلك من خلال تحليل المفاهيم الداخلية التي تضمنها تعريفات المعجم، على أساس اعتبار أن كل تعريف لأحد معاني المفردات هو بمثابة مفهوم مركب تمثله شبكة دلالية تربط بين عدد من المفاهيم الأبسط، وهكذا نتحرك صعودا حتى نصل إلى أعقد المفاهيم التي يتضمنها المعجم.

إذ تعد علاقات المعارف المعجمية[1] مقوما رئيسيا في نظم الفهم الأتوماتي حيث تستخدمها آلة الاستدلال المنطقي لهذه النظم لاقتفاء أثر علاقات التطابق، والاشتمال، والتفريع، ومعرفة العلاقات الدلالية التي تربط بين الأفعال وموضوعاتها، والاسماء وملحقاتها، وما إلى غير ذلك[2].

يحتاج تطبيق تقنية قاعدة المعارف على المعجم العربي إلى مراجعة دقيقة للغة تعريفه بغرض تنميطها، ووضع قواعد عامة لكيفية إجلاء اللبس في حالة المفردات متعددة المعاني، ويقصد بذلك اختيار الألفاظ الحاكمة في تعريفات هذه المفردات، والتي يتوقف عليها فك اللبس المذكور، وإليك مثالا لتعريفات بعض معاني كلمة "عين" في العربية موضحا بها الكلمات الحاكمة في إجلاء اللبس.

العين: عين الحيوان التي يبصر بها.

العين: النقد من الدراهم.

العين: مطر أيام لا يقلع.

العين: عين البئر وهو مزج مائها.

العين: نفس الشيء.

العين: الجاسوس.

نظرا للجهد الكبير المطلوب لتنميط تعريفات المعجم العربي بالطريقة السالفة الذكر، نقترح هنا أن نبدأ بتحديد البنية الموضوعية[3] للمعجم العربي، يمكن الاستعانة بالمكانز لتحديد قائمة الموضوعات وتصنيفاتها الفرعية، ولإعطاء القارئ فكرة أكثر تحديداً، نشير هنا إلى أن المكنز الذي استخدم في تحديد البنية الموضوعية للمعجم العبري قد شمل 15 قسما تم تفريعها إلى 750 حقلا دلاليا، وإليك قائمة الأقسام المذكورة مرتبة بحيث تتجاور فيها الأقسام المتقاربة دلاليا فهي تتدرج من المجردات إلى الماديات ثم من الماديات إلى العواطف[4]:

1- تعد تعريفات المعجم المصدر الرئيسي لبناء شبكات المعارف المعجمية، والتي تعد بدورها أحد المقومات الرئيسية لنظم الفهم الأوتوماتي، فاستخدام معان مرادفة أو قريبة لمدلول الكلمة، أو إعطاء سلسلة من المرادفات اللفظية لها، أو وضعها في شبه جملة، أو جملة تامة، لإبانة المعنى المقصود من السياق، ونظرا لخاصية الاشتراك اللفظي يمكن أن يكون للفظة الواحدة أكثر من معنى، ولكل معنى تعريفه الخاص به، وما يرتبط بهذا المعنى من تصنيف نحوي فرعي، وإطار دلالي، حيث تتوقف هذه الخصائص المعجمية على المعنى المحدد للفظ.

2 – نبيل علي، اللغة والحاسوب ، سابق ، ص 517.

3- ويقصد بالبنية الموضوعية هنا تحديد الانحياز الموضوعي لكل جذر من جذور المعجم، وذلك بتحديد الموضوعات الأكثر استخداما لمشتقات هذا الجذر لهذا الغرض

4- اللغة والحاسوب، سابق، ص518.

1 مجرد	2 كمية	3 نظام	4 عدد	5 زمن
6 سبب	7 مساحة	8 شكل	9 حركة	10 مادة
11 حواس	12 فكرة	13 إرادة	14 حيازة	15 عواطف

لتحديد الانحياز الموضوعي لجذر معين، يتم حصر جميع مشتقاته، ثم يحدد عدد مشتقاته المستخدمة في كل من الأقسام الخمسة عشر المذكورة. يحدد الانحياز الموضوعي للجذر بدلالة الموضوعات التي تستخدم من مشتقاته نسبة تزيد عن 30% مثلا من إجمالي عدد المشتقات، وقد تم اختبار هذه الطريقة بصورة تقريبية على بعض جذور المعجم العربي، نورد هنا السمات الموضوعية لأربعة جذور منها، وهي: "دخل"، "ركب"، "قدم"، "شكل"، بغرض إبراز المقصود من الفكرة فقط[1]:

الجذر "دخل": [+ مجرد، + حركة، + مادة]

الجذر "ركب": [+ مادة، + حركة، + مجرد]

الجذر "قدم": [+ نظام، + حركة، + زمن]

الجذر "شكل": [+ شكل، + مجرد]

3-5- قاعدة عينات النصوص المختارة لقاعدة البيانات المعجمية للاستخدام المعجمي:

يقصد بقاعدة النصوص تخزين عينات مختارة من النصوص اللغوية يتم تجميعها من مصادر مختلفة، وتختار هذه العينات وفقا للمبادئ الإحصائية السليمة لتجيء ممثلة للكم الهائل من النصوص المستخدمة في لغة ما. تستخدم قاعدة النصوص لإعطاء مؤشرات كمية وأمثلة واقعية عن الاستخدام الفعلي للمفردات مع تنوع المجالات والأساليب وأشكال مصادر النصوص (وثائق – كتب – تقارير – كشوف – مسرحيات – قصص – ...). وتعد قاعدة عينات النصوص[2] حاليا أحد المقومات الرئيسية لمؤازرة العمل المعجمي وترشيده، وهي الوسيلة العملية لتحاشي انعزال المعجم عن جماعته اللغوية.

كانت جامعة "براون" في الولايات المتحدة الأمريكية سباقة في هذا الميدان، حيث جمعت قاعدة نصوص للغة الإنجليزية (الأمريكية) يقدر حجمها بمليون كلمة، تم تنميط جميع الكلمات بطرق آلية، وشبه آلية، تمهيدا لتحليل مضمونها إحصائيا ولغويا، تضمنت قاعدة النصوص 500 عينة موزعة على 15 نوعا من أنواع النصوص المنثورة.

نورد هنا قائمة بها مقرونة بعدد عينات النصوص لكل نوع منها[3]:

1- المرجع السابق، الصفحة نفسها.

2- قاعدة عينات النصوص المختارة تمثل إحصائيا الاستخدام الفعلي للمفردات من قبل الجماعة اللغوية. على ضوء التغذية المرتدة تكيف عمليات التوليد المعجمي عملها لتحتوي مظاهر الإزاحة الدلالية لمعاني الألفاظ، وتلتقط الرائج من التعابير الاصطلاحية المبتكرة من صلب النصوص الفعلية بغرض تعميمها وتقييسها، وتلفظ المهجور وغير المستساغ، وتعيد إحياء المقبول من القديم، وما إلى ذلك.

3- اللغة والحاسوب، ص 519.

أولا: <u>كتابة إخبارية:</u>

(أ) صحافة: تحقيقات صحفية	44
(ب) صحافة: مقالات ومراسلات	27
(ج) صحافة: عروض	17
(د) دين	17
(هـ) مهارات وهوايات	36
(و) قراءات عامة (كتب ومجلات)	48
(ز) مذكرات وسير ذاتية	75
(ح) كتابات متنوعة	30
(ي) مواد تعليمية	80

ثانيا: <u>كتابة روائية (خيالية):</u>

(ك) روايات عامة	29
(ل) روايات غامضة وبوليسية	24
(م) خيال علمي	6
(ن) روايات مغامرات	15
(ش) روايات رومانسية وغرامية	29
(ص) فكاهة	9
إجمالي	500

وقد قسم كل نوع من أنواع الكتابة المذكورة إلى تقسيمات فرعية، فعلى سبيل المثال، قسمت التحقيقات الصحفية إلى: سياسية، اجتماعية، رياضية، أخبار الساعة، مالية، ثقافية.

أصدرت الجامعة كتابا يشكل نتائج التحليلات الإحصائية ، وهو يعد حاليا مرجعا أساسيا في الدراسات والبحوث اللغوية والمعجمية، وتصميم نظم المعلومات ومعالجة اللغة آليا[1].

تقوم قاعدة النصوص بتقديم عدة خدمات أساسية، نذكر منها[2]:

(أ) إحصائيات عن معدل استخدام كل كلمة من الكلمات، المفردة والمركبة، داخل صلب الوثائق، وفي عناوينها، وعدد الأنواع الرئيسية للنصوص، وعدد عيناتها، التي وردت بها الكلمة، يوضح المثال الموضح بيانا بالمعطيات التي يوفرها دليل جامعة "براون" لكل كلمة من كلمات قاعدة النصوص.

(ب) إحصائيات عن طول الجمل، وأنواع أساليبها.

(ج) استخلاص فقرات من النصوص التي تتضمن لفظا معينا، أو عدة ألفاظ متتاخمة أو متباعدة، مرتبة أو غير مرتبة، وذلك للاستعانة بها في تحديد التصنيفات النحوية الفرعية، والأطر الدلالية للألفاظ، وكذلك لدراسة العلاقات الأخرى التي تربط بين الكلمات.

1- المرجع السابق، ص 520.

2- المرجع نفسه ، الموضع نفسه.

يجب أن يتوفر في قاعدة النصوص المقترحة لمواكبة أعمال التطوير والتحديث للمعجم العربي، عدة شروط إضافية، من أهمها[1]:

1/ نظرا لحدة الخاصية الصرفية في "العربية"، يجب بناء قاعدة النصوص على أساس صرفي، ويقصد بذلك عدم الاكتفاء بتصنيف المفردات وفقا لأقسام الكلم (كما هو الحال في قاعدة "براون" للغة الإنجليزية)، بل يجب النفاذ إلى البنية الصرفية للكلمات. لهذا الغرض، يمكن استخدام المعالج الصرفي الآلي متعدد الأطوار[2]، في القيام بعملية التحليل الصرفي المذكورة بصورة تلقائية، وهو ما سيتيح بدوره الحصول عل إحصائيات هامة عن معدل استخدام الجذور، والصيغ الصرفية، والدمج الصرف – نحو (دمج الأدوات، وزوائد التصريف، والضمائر المتصلة مع جذع الكلمة)، إن بناء قاعدة النصوص على أساس صرفي يجيء متسقا مع **تنظيم المعجم العربي على أساس الجذور**[3].

ننتقل الآن إلى سلسلة الخطوات اللازمة لإعداد قاعدة النصوص العربية[4]:

(أ) اختيار عينات النصوص على أساس إحصائي بحيث تلبي مطالب التغطية الموضوعية، وتنويع أساليب الكتاب، وتنوع أشكال المستندات والوثائق (تقارير، كتابة صحفية، أدبيات، جداول، فهارس، ...) ويمكن الاستهداء هنا بالتصنيفات الواردة في دليل جامعة "براون".

أما في يخص تنوع أساليب الكتاب، فيجب توخي الدقة العالية في اختيارها نظرا للتباين الشديد في أساليب الكتابة العربية الحديثة بفعل المرونة النحوية التي تتسم بها "العربية".

(ب) وضع مجموعة من القواعد والمبادئ لتنميط النصوص العربية تهيئة لإدخالها للحاسوب، ويقصد بذلك وضوع الرموز الخاصة بأقسام الكلم، (والتي تحتاج حتما إلى تصنيفات أدق من تلك المستخدمة حاليا)، وأنواع الأساليب الجملية، وتنميط استخدام علامات الترقيم، وعلامات التشكيل، وكتابة الياء والألف المقصورة والهمزات، وطرق كتابة الأرقام والأعداد، وعلى ضوء هذه القواعد يتم تنميط عينات النصوص بأساليب آلية.

نظرا للترابط السياقي الشديد الذي يربط بين الجمل داخل النصوص العربية، والتي من مظاهرها كثرة استخدام حروف العطف والاستدراك والاستئناف، يجب القيام بتحديد فواصل الجمل بصورة يدوية على ضوء عدة قواعد واضحة للحكم على تمام الجمل وصحتها.

(ج) استخدام المحلل الصرفي الآلي لتحليل النصوص المختارة لاستخلاص جذورها وصيغها الصرفية، وتحديث معدلات استخدام الجذور والصيغ الصرفية وقفا لنتائج التحليل الصرفي.

1– المرجع السابق، الموضع نفسه.

2– للمزيد من التفصيل ينظرالباب الثاني ،الفصل الأول، الصفحة181.

3– إن تنظيم مواد المعجم على أساس الجذر والاكتفاء بهذا التنظيم دون غيره، بالرغم من وجاهته المعجمية والصرفية، إلا أن تنظيم المعجم على أساس الجذور يضع حاجزاً عازلا بين المعجم ومستخدمه، والذي لا يهمه أصل الاشتقاق بقدر ما يهمه اللفظ المشتق ذاته، علاوة على ذلك، فقد افترض التنظيم المبني على الجذر معرفة المستخدم بالتفاصيل الدقيقة للتحليل الصرفي، وقواعد الإبدال والإعلال، والتي يتعذر بدونها استخلاص جذور الكلمات تهيأة لعملية الكشف عنها، وهي قدرات لغوية لا تتوفر للكثيرين، خلاصة القول، لقد طغى الصرف على المعجم العربي على حساب النحو والدلالة، بل يجوز أن نضيف هنا، وعلى حساب المستخدم أيضاً، للمزيد من بسط القول، ينظر:اللغة والحاسوب، ص 440،441.

4– اللغة والحاسوب، مرجع سابق، ص 522.

(د) طبع نتائج التحليلات الإحصائية وفقا لمجموعة من المؤشرات الصرفية والنحوية والمعجمية.

3- 6 - : الخدمات المعجمية لقاعدة البيانات المعجمية:

تتضمن وحدة الخدمات المعجمية في المنظومة المعجمية نطاقا عريضا من خدمات البحث والاسترجاع والاستدلال، ويتم تقديم هذه الخدمات لفئات عريضة من المستخدمين، أو يتم دمجها في إطار عدد كبير من النظم الآلية يتضمن جدول الآتي[1] بيانا بهذه الخدمات مقرونة بفئات المستعملين ونوعيات النظم الآلية المستفيدة منها، يتضح من الجدول مدى اتساع نطاق الخدمة المعجمية بصورة تعجز الوسائل اليدوية في تلبيتها مهما كانت درجة كفاءتها.

جدول يمثل قائمة الخدمات المعجمية:

الخدمـــات المـــعجمية	بشر				نظـــم آليـــة						
	أطفال	كبار	لغويين	باحثين	صرف آلي	نحوي آلي	ترجمة آلية	استرجاع معلومات	فهم أتوماتي	تحليل نصوص	بالحاسوب
معاني المفردات	•	•	•	•	•	•	•	•	•	•	•
إعطاء قائمة بسلاسل المفردات التي ترتبط بعلاقة معينة يختارها المستخدم، أو النظام الآلي كسلاسل المترادفات وسلاسل أفعال الحركة، وسلاسل أسماء النباتات، وذلك على سبيل المثال	•	•	•	•			•	•	•	•	•
تحديد مسار الاشتقاق، أي توضيح الخطوات المتدرجة التي سار فيها اشتقاق كلمة ما ابتداء من الجذر (مثال: أكابر – أكبر – كبير – كبر)			•	•	•	•	•	•	•	•	•
علاقات الإنتاجية الصرفية للجذور، أي استرجاع جميع الصيغ المنطبقة على جذر معين، أو العكس، جميع الجذور التي ترد على صيغة صرفية			•	•	•						•
عرض حالات الشذوذ المختلفة وفقا لما يحدده المستخدم أو النظام الآلي			•	•	•	•					•
تقديم قائمة السمات النحوية والدلالية للفظة معينة			•	•	•	•			•	•	
تقديم مخطط المفاهيم للفظ معين									•	•	•
تقديم سلميات أو شبكة المفاهيم المرتبطة بموضوع ما، أو بمفهوم شامل معين				•			•		•	•	•
البحث العكسي في المعجم، أي إعطاء لفظ، أو عدة ألفاظ معينة، لتطابق، أو تقترب من تعريف معين يحدده المستخدم أو النظام الآلي				•			•				•
إعطاء بيانات عن الأصل التاريخي لنشأة الألفاظ، ومسار تطور معناها				•							•
إعطاء قائمة الموضوعات التي يستخدم فيها جذر معين، أو مفرد معين				•			•	•		•	•
إعطاء جميع الصفات، أو الأفعال المتوافقة مع اسم معين	•	•	•	•		•	•	•	•	•	•
إعطاء قائمة بالكلمات المركبة التي ترتبط مفرداتها بعلاقة دلالية معينة	•	•	•	•	•	•	•	•	•	•	•

1- اللغة والحاسوب، سابق، ص 524 .

3-7- العلاقات الداخلية لقاعدة البيانات المعجمية لمنظومة المعجم:

ترتبط العناصر الفرعية لمنظومة المعجم مع بعضها بعلاقات وثيقة يمكن تصنيفها في مستويين[1]:

*/**المستوى الأول:** العلاقات بين عمليات تحديث المعجم، والتي تشمل صك المفردات، ولفظ المهجور. وجمع وتنميط التعابير الاصطلاحية، وعمليات الإزاحة الدلالية، وتمثل وحدة تحديث المعجم رابطة العقد بين هذه المعليات الفرعية.

*/**المستوى الثاني:** العلاقات التي تربط بين وحدة تحديث المعجم نفسها مع وحدات البحث والتنظيم، وقاعدة عينات النصوص، سنركز هنا على علاقات المستوى الثاني[2]:

أولاً- العلاقة بين تحديث المعجم وتنظيمه: يحدد شكل تنظيم المعجم نوعية المعطيات الواجب إمدادها بواسطة وحدة تحديث المعجم، وذلك لكل من المفردات الجديدة التي تضيفها هذه الوحدة، تتوقف العلاقة بين التحديث والتنظيم على عدة عوامل أهمها[3]:

- أسلوب ترتيب مفردات المعجم.
- طبيعة نظام قواعد البيانات المستخدمة في تنظيم مواد المعجم.
- طبيعة نظام قواعد المعارف المستخدمة في بناء شبكة المفاهيم المعجمية.

ثانياً- العلاقة بين تحديث المعجم وقاعدة عينات النصوص: تغذي قاعدة عينات النصوص وحدة تحديثالمعجم بفقرات النصوص اللازمة لتحديد التصنيفات النحوية الفرعية للمفردات وسماتها الدلالية، وكذلك قائمة التعبيرات الاصطلاحية المستحدثة التي تبتدعها الجماعة اللغوية، علاوة على معدلات استخدام المفردات والتعابير الاصطلاحية داخل نوعيات النصوص المختلفة.

ثالثاً- العلاقة بين تحديث المعجم ووحدة الخدمات المعجمية: بصورة مثالية، يفترض في وحدة الخدمات المعجمية أن تمد وحدة تحديث المعجم بالمطالب الجديدة لفئات مستخدميها، ونظمها الآلية، حتى تسعى وحدة تحديث المعجم إلى توفير المادة المعجمية[4] التي تفي بهذه

1- اللغة والحاسوب، سابق، ص526.

2- المرجع نفسه، الصفحة نفسها.

3- المرجع نفسه، ص 526.

4- قبل التطرق في أي بحث معجمي لا بد من تحديد للمادة المعجمية المراد الاشتغال عليها لبناء معجم أو قاموس ،والمادة اللغوية هي الملكة اللغوية التي تكون لدى متكلم لغة ما بحيث تجعله بين مفردات هذه اللغة بحيث يعرف المستعمل والمهمل ..الخ وقد بدأ البحث المعجمي العربي مع الخليل بن أحمد الفراهيدي ت(170هـ)،صاحب معجم "العين " الذي حاول حصر مستعمل اللغة العربية ،وترك مهملها ،غير أن المعجميين من بعد الخليل لم يلتزم بمنهج اللغة كالأزهري صاحب معجم" تهذيب اللغة"، وابن دريد صاحب "الجمهرة "وفيروز أبادي صاحب "القاموس المحيط" وابن منظور صاحب "لسان العرب" فقد اختاروا ما فاه به البدو دون الحضر ،وتجنبوا ما استجد من المصطلحات الجديدة كالرياضيات والطب والنبات والفلك .. يقول الفاسي الفهري: " وإذا كانت الإسهامات المعجمية العربية في القرون الهجرية الأولى ذات أهمية تاريخية كبرى بالنظر إلى تنوعها نهجا ومادة وتأليفا ، فإن المعاجم العربية رغم بعض الجالية رغم بعض الجهود تظل قاصرة لا تلبي حاجات مستهلكيها وهي تفتقد إلى الكفاية الوصفية ولا تغطي المادة المعجمية الجديدة ولا المعاني الجديدة للمفردات ولا تهتم بجوانب النطق والصرف والتركيب و الدلالة بصفة نسقية منتظمة أو جانب الأثالة ".

المطالب، مثال ذلك، أن يحدد اللسانيون الحاسوبيون مطالب نظم الفهم الأوتوماتي فيما يخص العلاقات الدلالية للكلمات المركبة والتعابير المجازية[1].

رابعاً- العلاقة بين تنظيم المعجم وقاعدة عينات النصوص: تمد وحدة تنظيم المعجم قاعدة عينات النصوص بالمعطيات اللازمة لتنميط نصوصها حتى تجيء متسقة مع مطالب تنظيم المعجم، وتمد قاعدة عينات النصوص وحدة تنظيم المعجم بإحصائيات متنوعة تفيد عملية تصميم هذا التنظيم على أسس كمية.

خامساً- العلاقة بين تنظيم المعجم ووحدة الخدمة المعجمية: تمد وحدة تنظيم المعجم وحدة الخدمات المعجمية بجميع المعطيات اللازمة لتلبية طلبات البحث من مفردات، ومعان، وعلاقات....، وتمد وحدة الخدمات المعجمية وحدة تنظيم المعجم بإحصائيات متنوعة عن طلبات البحث: موزعة وفقا للموضوع، وفئة المستخدم، وما إلى ذلك، وتفيد هذه الإحصائيات في ترشيد عملية تنظيم المعجم.

سادساً- العلاقة بين قاعدة عينات النصوص ووحدة الخدمات المعجمية: تمد قاعدة عينات النصوص وحدة الخدمات المعجمية بفقرات النصوص التي تتضمن كلمة، أو عدة كلمات معينة، يقوم المستخدم بتحديدها وذلك لأغراض تعليمية أو بحثية.

رابعا: سبل تطوير قاعدة بيانات معجمية لبناء معجم آلي عربي:

لتطوير وإعداد قاعدة بيانات معجمية يجب علينا[2] :

(أ) ربط عملية تطوير معجم اللغة بتحديث نظام التقعيد لها نظرا للعلاقة العضوية التي تربط بينهما.

(ب) الفصل بين اللغة العربية الحديثة، واللغة العربية القديمة، ومراعاة العامل الجغرافي في الاستخدام اللغوي على مستوى الوطن العربي.

(ج) تلبية المطالب المختلفة لفئات المستخدمين، كبارا وصغارا، مترجمين وكتاب، لغويين وتربويين.

(د) الانطلاق من منظور دلالي، وهو ما يتطلب بدوره اهتماما خاصة بلغة تعريف المعجم العربي، والتي تحتاج إلى مراجعة شاملة لتقييس أنماطها، وزيادة دقتها في التعبير عن مدلولات المعاني.

(هـ) وضع الأسس للتصنيف المعجمي الدقيق، والتي تشمل وضع قائمة قياسية للسمات الصرفية والنحوية والدلالية للألفاظ العربية كأساس لتصنيف المفردات، وتحديد العلاقات التي تربط بينها، يجب أن يراعى في اختيار هذه السمات مطالب النصوص العربية غير المشكولة، والتي تحتاج في إجلاء لبسها أكثر دقة وتفصيلا.

1- اللغة العربية والحاسوب، سابق ص 527.

2- المرجع نفسه، ص 528،529.

(و) أقصى استغلال لمظاهر الفائض المعجمي، والتي من أهمها تلك المرتبطة بالإنتاجية الاشتقاقية العالية للجذور العربية.

(ز) الاهتمام بالعلاقات التي تربط بين الكلمات، وإعادة تمحيص ظاهرة الترادف والاشتراك اللفظي في "العربية".

(ح) استغلال الرصيد المعجمي الهائل للغة العربية بإحياء كثير من المفردات المستخدمة في العربية القديمة.

(ط) إعادة النظر بصورة شاملة لآليات تكوين الكلمات في العربية على أساس النظرية الصرفية الحديثة، وتجاوز الحدود الضيقة لتكوين الكلمات بالاشتقاق الصرفي، يرتبط بذلك إعطاء اهتمام خاص لظاهرة تركيب الكلمات في "العربية"، وحصر وتبويب وتصنيف التعابير الاصطلاحية بها.

(ي) ربط المعجم العربي بالجماعة اللغوية المستخدمة له، وذلك من خلال دورة التغذية المرتدة، والتي يمكن توفيرها من خلال قواعد عينات النصوص.

(ك) استغلال الإمكانيات العديدة التي يتيحها المعالج الصرفي الآلي متعدد الأطوار في ميكنة المعجم العربي، وإعادة تنظيمه، وإعداد قاعدة نصوص ذات طابع صرفي لمؤازرة الأعمال المختلفة للتطوير المعجمي.

(ل) متابعة دقيقة لمشروع "المعجم" الجاري تنفيذه حالياً في معهد MIT والذي يركز على الجوانب الدلالية للمعجم في ظل نظرية الربط العاملي لـــ "تشومسكي".

(م) عزل البرجماتيات المرتبطة بالبيئة الصحراوية من معجم العربية الكلاسيكية، وفصل الاعتباطيات عن المظاهر المطردة والمنتظمة التي تربط بين مفردات المعجم، حيث يمثل الاطراد والانتظام أساس تنظيم المعجم، وزيادة فاعلية دمجه مع النظم الآلية.

4-1- بعض الجهود المبذولة لتطوير قاعدة بيانات معجمية دلالية للمعجم الإلكتروني في العالم العربي:

هناك عدة مؤسسات ومنظمات عربية تسعى حاليا لميكنة المعجم العربي، نذكر منها[1]:

1- مشروع العالمية للبرامج بالكويت لميكنة المعجم العربي في إطار مشروعها الخاص بالفهم الأوتوماتي للنصوص العربية المكتوبة..

2- مشروع المركز العلمي لشركة "آي. بي. إم" بالقاهرة لتطوير قاعدة بيانات معجمية.

3- تشكيل اللجنة الاقتصادية لمنطقة غرب آسيا (ECWA) لمجموعة عمل لبحث تطوير قاعدة بيانات للمعجم العربي.

4-تشكيل لجنة لشؤون الحاسوب في المجمع اللغوي بالقاهرة.

1- اللغة والحاسوب، سابق، ص529.

خلاصة

وفقاً لِما سلف ذِكرهُ في مقدمة البحث، عالج هذا الفصل موضوع مذكرتنا من زاوية خاصة، فهو ينفردُ بدراسة المعالجة الآلية للغة العربية، النظام الصرفي كأنموذج للهذه الدراسة البحثية، مع إبراز أهم تطبيقات المعالجة الآلية للغة العربية في إعداد معجم إلكتروني للغة العربية، أما الجزء الثنائي فأردناه للمعالجة الآلية للمعجم الإلكتروني للغة العربية، وكان الحديث فيه عن الخلفية اللسانية في بناء قاعدة معطيات معجمية للمعالجة الآلية لبناء وتطوير معاجم آلية عربية.

تمثل استقراؤُنا من خلال هذه الدراسة في اعتبارها منطلقاً واعداً لإعداد مشاريع لسانية في مجال المعلوماتية وفق منظومة اللغة العربية لبناء برامج حاسوبية العربية، وعليه تم التركيز في هذا الفصل على نماذج من المعاجم الإلكترونية، من خلال تسليط الضوء على نماذج لقاعدة بيانات معجمية من معاجم إلكترونية مثل القاموس المفهومي المبنية أساسا على قواعد المعرفة والمعجم التفسيري الآلي الذي يلائم نظام اللغة العربية في كثير من جوانبه مما جعلنا نركز في بحثنا على هذا المعجم ولما لا تطويره مستقبلا ليصبح برنامجا أصيلا للغة العربية سهل المنال والاستخدام، كما تطرقنا في تفصيل إلى المعاجم الآلية DELA التي أثبتت جدارتها ونجاعتها في مقاربة البنيات البسيطة والمركبة المتلازمة[1] للغة العربية التي يتم استخراجها من النصوص المعجمية ببناء قواعد بياناتها بفضل تطبيقات الآلية لبيئة التنمية اللغوية .

ومما تجدر ملاحظته في هذه الدراسة هو مدى انتظام ونسقية ومعمارية المعاجم المذكورة، سواءا تعلق الأمر ببناء قواعد بيانات ضخمة وموسعة أو بإيراد العلاقات المنطقية والصورية التي تجمع البنيات اللغوية الحرة والمقيدة. بحيث تظهر الكفاية مجسدة على مستوى الحاسوب وما علينا إلا التفكير في بناء مثل هذه الأساسيات للمعطيات المعجمية العربية تكون بحق شاملة ومشمولة لجميع الأشكال والبنيات في اللغة العربية، كما تشمل أيضا القدرة الإنتاجية العالية والمتمثلة في قدرة المتكلم العربي الطبيعية بمعالجتها آليا بأساليب وتقنيات المعلوماتية المتقدمة، والهدف من ذلك هو تحقيق تمثيل أعمق لمعلومات ومعارف لغتنا العربية بغية تحضير نظامها لملائمة الأدوات التكنولوجية العصرية للتطبيقات الجديدة.

إن المعالجة الآلية للغة العربية والصرف العربي على وجه الخصوص قد قطعا أشواطاً هامة خلال العقدين باهتمام المختصين في علم الحاسوب وعلم اللغة لتطوير وسائل وأساليب

1- مقاربة التعبيرات المركبة والمتلازمة إعتمد عليها الدكتور عز الدين غازي في أطروحته لبناء معجم آلي للغة العربية.

المعالجة الحاسوبية للمعجم، وكذلك الإنجازات الإبداعية التي تحققت في هذه المعالجة، في جوانب الصرف، والتركيب، والدلالة، وبناء المعاجم، وغيرها.. مما سمح لنا بتكريس الفصل الباب الثالث والأخير لوضع ما تمت دراسته في هذا الباب النظري محل التطبيق في استخدام البرمجة الحاسوبية من خلال إستعراض مجموعة من المحطات وتجارب سابقة اعتمدت التقنيات الحاسوبية وسيلة لها أدت إلى حوسبة المعاجم العربية، من خلال بعض الأعمال المعجمية التي نريد ملامسة كيفيّة الإفادة من الحوْسبة للارتقاء بالمعاجم العربيّة لإعداد معاجم آلية جديدة على غرار التّجارب الجاريّة على ألسن أخرى مثل الإنجليزية أو الفرنسية، ثم عرض أنموذج تطبيقي نحو بناء معجم إلكتروني للمعالجة الآلية للغة العربية كدراسة تطبيقية لهذه الدراسة.

الباب الثالث

توظيف تقنيات المعالجة الآلية للغة العربية لإعداد وتطوير معجم إلكتروني للغة العربية

الفصل الأول

المعجم الإلكتروني:
ـ نظرة في بعض الدراسات المعجمية ـ
* قراءة في الحصيلة واستشراف للمستقبل *

أولاً – الدراسات الغربية

ثانياً – الدراسات العربية: التجربة العربية: الواقع والمأمول

تمهيد:

إن المعجمَ الإلكتروني المنشودَ أصبح ضرورةً فرضها عصرُ المعلوميّات الذي نعيشُ فيه، كما فرضتْها طبيعةُ المعجمِ العربيِّ من حيثُ مضمونُه وتنظيمُه وتطويرُه وخدمتُه للمستثمرين ولِنُظُمِ المعالجةِ الآليّة. فهو يتميز بخاصيات تفتقدها المعاجم التقليدية، كالشمول، والانتظام، والاطراد، والدقة والوضوح، وقابليتها التطوير والعديل والتحليل[1]. وتتجلى هذه الضرورةُ لدى أدنى مقارنةٍ بين ما وصلت إليه صناعةُ المعاجم في اللغات العالمية، وبين واقع معاجمنا العربية الإلكترونية خاصة.

بالنسبة إلى اللغة العربية فثمة تقصير واضح في إعداد معاجم حاسوبية تلبي حاجيات المستخدم العربي بالرغم من إتاحة أمهات المعاجم العربية على أقراص المدمجة أو على الإنترنت. فهذه الخطوة ليست كافية؛ لأن هذه المعاجم تفتقر لأبسط مقومات المعاجم الحاسوبية بالمعنى الحديث للكلمة. إنها مجرد "نسخ مرقمنة" لتلك المعاجم الورقية (في صيغة " Doc" أو "HTML") لا يمكن الاستفادة منها بالشكل المطلوب[2]. بالإضافة إلى ذلك، فالأدوات المصاحبة للبحث عن المعلومة بسيطة كما يتضح في المشاريع المعروفة مثل "عجيب" (لنظام صخر) "وكلمات" و"لسان العرب" / "القاموس المحيط". هذه السلبيات تعود أساسا إلى ضعف الهيكلة الحاسوبية التي بنيت عليها مداخل هذه المعاجم.

فمن خلال بعض الأعمال المعجمية نريد ملامسة كيفيّة الإفادة من الحوْسبة للارتقاء بالمعاجم العربيّة لإعداد معاجم جديدة تكون مُحوْسبة على غرار التّجارب الجاريّة على ألسن أخرى مثل الإنجليزية أو الفرنسية، فالتطور الذي حصل على مستوى قواعد البيانات والبرمجيات المصاحبة، بالإضافة إلى تطور الأجهزة الحاسوبية، قد ساعدا في بداية الثمانينات على ظهور أول المعاجم الإلكترونية المتاحة إلى الجمهور العريض على سطح المكتب أو على الإنترنت (سواء عن طريق الاشتراك أو مفتوحة المصدر) أو على الأقراص المدمجة (CD - ROM) والأقراص الأخرى.

ويأتي موضوع هذا الفصل للحديث عن أهم الإنجازات في صناعة المعجم الإلكتروني ودور اللساني والحاسوبي في هذا الإطار، حيث نستعرض في هذا الفصل مجموعة من

---

1 – ينظر: د. أحمد مختار عمر، صناعة المعجم الحديث، ص179–188، وينظر: محمود فهمي حجازي، البحث اللغوي، ص71–78، د. محمد الحناش، مشروع نظرية حاسوب لسانية في بناء معاجم آلية للغة العربية، مجلة التواصل اللساني، المجلد2، العدد2، 1990م، ص43، د.عبد الغني أبو العزم، الحاسوب والصناعة المعجماتية، مجلة اللسان العربي، العدد 46، 1998م، ص28–39 .

2– ينظـــر: - Ait Taleb S., (2005). Dictionnaires électroniques arabes: le modèle des dictionnaires de Sakhr, revue de l'Association Marocaine des Etudes Lexicographiques, Numéro 3-4, 15-31.

المحطات وتجارب سابقة اعتمدت التقنيات الحاسوبية وسيلة لها أدت إلى ما آلت إليه حوسبة المعاجم العربية،. والتي تعتبر في الحقيقة حلقة وصل ما بين المعاجم الورقية التقليدية والمعاجم الالكترونية ذات المعمارية الجديدة المبنينة وفق قواعد وأنظمة الهندسة الإلكترونية واللسانية على السواء، لبناء معاجم إلكترونية للغة العربية من منظور نظرية لسانية حاسوبية، من خلال الاستفادة من التجارب العلمية للأمم والشعوب الأخرى.

إذن بناءا على ما تقدم من معطيات وفي ضوء التحديات الراهنة ، ما هي أهم الإنجازات الغربية والعربية التي عرفتها صناعة المعاجم الآلية والمحوسبة؟

أولا:ـ الدراسات الغربية: تجربة LADL[1]**:**

يتألف نظام المعاجم الإلكترونية للغة الفرنسية[2] من مستويات[3]:

1ـ **المعجم الإلكتروني للمفردات البسيطة:** DELAS : يتضمن هذا المعجم أكثر من 80000 مدخل معجمي، ويسند لكل مدخل رقم هو عبارة عن تشفير، يعتبر جزءاً من الخطاب، كما يسند إليه تشفير رقمي يحدد (الإعراب)[4]، مثل:

(جناح، س21)

حيث يشير (21) إلى أن الاسم مذكر، وتضاف إليه لاحقة الجمع (ات).

ويوظف DELAS حوالي 500 رقم، ويتراوح التشفير الرقمي ما بين (1 – 500)، كما يستعمل 2 Mo لتخزين ملفات النصوص[5].

تمكن الأرقام في DELAC من توليد حوالي 60000 شكل لغوي مأثور: صيغ تصريف الأفعال، الجمع والإفراد بالنسبة للأسماء، المذكر والمؤنث، المفرد والجمع بالنسبة للصفات، ويسمى المعجم الإلكتروني للصيغ المأثورة[6] DELAF

1– يعتبر مخبر المعالجة الآلية والتوثيق اللساني LADL نموذجا للمراكز البحثية الرائدة في العالم على مستوى المعالجة الآلية للغات الطبيعية، إذ تمكن من تطوير معاجم إلكترونية للغة الفرنسية استنادا إلى نظرية المعجم التركيبي (Lexique Grammaire).

2– من بين القواميس الإلكترونية الغربية العديدة والمتنوعة، نذكر هنا بعض الأمثلة للغة الإنكليزية والفرنسية: ذخيرة اللغة الفرنسية المحوسبة (TLFi)، وقاموس الأكاديمية الفرنسية (2 طبعات)، و"البيبليورم لاروس" الفرنسي ومعجم "أوكسفورد" الإنجليزي OED (الطبعة الثانية) والمعجم والمكنز الإلكتروني "كولينز".

3– عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مرجع سابق،99،98.

4 – flexion.

5– عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مرجع سابق،99.

6 – B. Courtois 1984.

تتألف مداخل: (DELAF) من مفردة وتشفير رقمي، ويوظف أقل من 8500 سنن رقمي، يتراوح ما بين (1 – 8500)، ويشغل مساحة 12 Mo مخصصة لملفات النصوص، يشغل منها DELAF فقط 6,7 Mo[1].

DELAC وDELAF مداخل الأول هي المفردات المركبة [Si–90] [Gr][2]. وتعرف المفردة المركبة بأنها متوالية من المفردات البسيطة، ذات خصائص تركيبية غير صالحة للمفردات البسيطة نفسها[3]. ومن الناحية الصورية فإن مداخل DELAC هي سلسلة من المفردات البسيطة المقترنة بمعلومات ترتبط بالخطاب والإعراب، وعلى العموم فإن الأرقام الموظفة في DELAC هي نفسها الموظفة في DELAF.

بالإضافة إلى عناصر مجموعة المعاجم الإلكترونية المبلورة من قبل (LADL) هناك معاجم أخرى لا تقل أهمية عما سبق ومنها: DELAP وDELAPF.

ـ **المعجم الإلكتروني الفونيمي** :DELAP : مداخله تصل إلى 80000، مقترنة بأوصاف صوتية. المفتاحية فيه هي المفردة، والمعلومة المقدمة هي الفونيمية[4] التابعة للمعلومات في DELAS، وما يميز هذا النوع من المعاجم هو كونه مركب أو معقد بمعنى آخر، كل مفتاح أي مدخل تسند إليه معلومات مغايرة[5].

– **المعجم التركيبي** :Grammaire Lexique وهو آخر عنصر من مجموعة المعاجم الإلكترونية، ويتعلق أساساً بقائمة الأرقام المسندة للجمل، والممثلة انطلاقاً من فعلها، وتشكل مجموع هذه الأرقام الصيغ النحوية للمفردة[6].

إن تشغيل معاجم LADL يطرح مشاكل كثيرة على مستوى قواعد المعطيات، لأن مفاتيح التفوق أو النجاح تتسم بالتغير مما يجعل من المساحة المخصصة للتمثيل المعلومياتي، والزمن المستغرق لنجاح المعطيات لا يسران على قدم المساواة، ومن ثمة ذلك سينعكس لا محالة على النتائج[7].

1– توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مرجع سابق،99.

2– المرجع نفسه، الصفحة نفسها.

3– المرجع نفسه، الموضع نفسه.

4 – Phonémisation

5– عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مرجع سابق،99.

6– المرجع نفسه، ص 100.

7– المرجع نفسه، الموضع نفسه.

ثانيا: الدراسات العربية: التجربة العربية: الواقع، والمأمول:

لقد حيكت جدالات وإشكالات حول القضايا المتعلقة ببناء معجم إلكتروني للغة العربية، من حيث مستوياته، وحقله المعجمي، ومحتوياته، فقد عقدت له ندوة خاصة نظمتها مدرسة الملك فهد العليا للترجمة بطنجة (المغرب) عام 1995م، وكان عنوانها "التقنيات الحاسوبية في خدمة المصطلح العلمي والمعجم "المختص"[1]. كما خُصص له جلسة نقاش سميت: "بناء المعجم حاسوبياً خلال: ندوة استخدام اللغة العربية في تقنية المعلومات[2].

ويُصنف الدكتور محمد الحناش في مقدمة اللغويين العرب المعاصرين الذين اهتموا بالمعجم الحاسوبي اهتماما خاصا، انطلاقا من دراساته المتعددة عن بناء المعاجم الآلية في اللغة العربية، والمعجم الإلكتروني، والمعجم التركيبي للغة العربية[3]...إلخ، وهو صاحب كتاب: "المعجم التركيبي للغة العربية – مقدمات في المعالجة الحاسوبية للغات الطبيعية"[4] ؛ ونحا نحوه آخرون أمثال نهاد الموسى، وأحمد مختار عمر، وعبد القادر الفاسي الفهري، ومحمد مراياتي وزملاؤه بوساطة نظام خبير للغة العربية[5].

ومن المشاريع العلمية ضمن هذا التوجه أيضاً ما قام به معهد اللغة العربية بجامعة الملك سعود بالرياض (السعودية) من إصدار المكنـــز Corpus الوجيز (معجم في المترادف والمتوارد)، ومعجم التعبيرات الاصطلاحية، بإشراف محمود إسماعيل صيني[6].

بالإضافة إلى جهود أخرى تبذل في بناء معجم إلكتروني للغة العربية من قبل المراكز العلمية والمؤسسات التجارية، كما يحصل في مشروع الشركة العالمية للبرامج (صخر) بالكويت، ومشروع الشركة العالمية لبرامج الحاسوب بالقاهرة عن المكنز الآلي أو قاعدة بيانات المادة المعجمية العربية، ومشروع المركز العلمي لشركة آي.بي.إم بالقاهرة لتطوير قاعدة بيانات

---

1 – للمزيد من بسط القول في أبحاث هذه الندوة ينظر: مجلة اللسان العربي، مرجع سابق، العدد 48، 1999، ص 169–229 .

2 – ينظر : السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات، ص 781–796.

3 – جاءت هذه البحوث على النحو التالي : المعجم الإلكتروني للغة العربية، مؤتمر الكويت الأول للحاسوب، 1989م، مشروع نظرية حاسوب – لسانية في بناء معاجم آلية للغة العربية، مجلة التواصل اللساني (المغرب) المجلد 2، العدد 2، 1990م، ص 40–55 (وقد نشر هذا البحث أيضاً ضمن السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات، ص 363–401)، المعجم التركيبي للغة العربية – معالجة المصادر والأسماء، مجلة التواصل اللساني، المجلد 2، العدد 1، 1990م، ص 42–49 .المعاجم الآلية للغة العربية – بناء قاعدة المعطيات، مجلة التواصل اللساني، المجلد4، العدد 1، 1992م، ص 81–108.

4 – صدر هذا الكتاب عن مجلة التواصل اللساني، فاس (المغرب)، 1992م .ويقع في جزءين، عدد صفحاتهما 350 صفحة .

5 – للمزيد من التفصيل ينظر : د. نهاد الموسى، العربية – نحو توصيف جديد...، ص 247–262، د. أحمد مختار عمر، المعجم العربي الحديث والخروج من الدائرة، مجلة كلية دار العلوم، العدد 21، 1997م، ص 16 فما بعدها، د. عبد القادر الفاسي الفهري، المعجمة والتوسيط، – نظرات جديدة في قضايا اللغة العربية. ط1، المركز الثقافي العربي، بيروت – لبنان، الدار البيضاء – المغرب، 1997م، ص 59–71، د.محمد مراياتي وزملاؤه، المعجم الحاسوبي نظام خبير للغة العربية، ضمن (بحوث المؤتمر العلمي الأول حول الكتابة العلمية باللغة العربية – واقع وتطلعات)، ص 339 فما بعدها.

6 – ينظر في هذا مداخلة الدكتور محمود إسماعيل صيني في الجلسة التي عقدت بهامش ندوة استخدام اللغة العربية في تقنية المعلومات، ضمن (السجل العلمي لهذه الندوة) ص 787–788 .

معجمية[1]؛ ومع ما ذكرته في عُجالة من جهود عربية في سبيل بناء وتطوير بناء معاجم إلكترونية للغة العربية، فالواقع أن جهود العلماء العرب المعاصرين والمؤسسات العلمية في هذا المجال يمكن حصرها خاصة بالبرامج والنظم التي وضعت لبناء معاجم إلكترونية للغة العربية، سواء ما كان منها فرديا، أو نتاجا مشتركا، أو عملا تجاريا عاما. ولعل من أهم الإنجازات في هذا المجال نذكر:

1. شركة صخر للمعلوماتيات والإلكترونيات: تجربة SAKHR:

تعتبر شركة صخر[2]، الشركة العربية الوحيدة التي لديها حالياً أطبر طيف من التقنيات والمنتجات، الموجهة لمعالجة اللغة العربية. وتتوفر حالياً على مدقق إملائي ونحوي، ومشكل آلي، تسميه "المصحح الآلي الثنائي اللغة" يمكن أن يدمج في برنامج مايكروسوفت وورد، أو يستخدم مع بعض البرامج الأخرى، ناهيك عن التقنيات الأساسية التي طورتها في مجال المعالجة الآلية للغة العربية، مما يتصل بـــ قاعدة بيانات المعلومات اللغوية"[3]، والمعالج الصرفي متعدد الأوجه[4] ، والمحلل النحوي[5] متعدد الأوجه والمشكل الآلي ، بينما تتضمن قاعدة بيانات المعلومات اللغوية: "قاعدةة بيانات المعاجم العربية"، و"قاعدة بيانات النحو العربي" و"الذخيرة اللغوية العربية".

1– للاطلاع على هذه الجهود المبذولة لإنتاج برامج معجمية بواسطة الحاسوب ينظر:د.نبيل علي، اللغة العربية والحاسوب، ص529، د.أحمد مختار عمر، صناعة المعجم الحديث، ص 169، 189 فما بعدها .

2- www. Sakhr.com

3– لابد لنا من الإشارة إلى مدونة شركة صخر (انظر عبد الغني أبو العزم "اللغة العربية والمعالجة الآلية: برامج صخر نموذجا") الذي يذكر فيه أنه يتوقع أن يصل حجم ما أسماه "مكنز صخر" إلى 500 مليون كلمة، كما يعطي أبو العزم تفاصيل مهمة عن هذه المدونة وطرق استفادة الشركة منها). فيما يلي نورد وصفا لهذه المدونة:

يذكر أبوالعزم في دراسته عن مدونة صخر اللغوية أن أهم معايير اختيار نصوص المدونة المذكورة كانت:

أ – التركيز على النصوص الأدبية والدينية ذات الطبيعية السردية. ب – انتقاء نصوص معبرة تاريخيا بما فيها نصوص ذات أساليب مجازية.

ج – تجنب النصوص الحوارية. د – إبعاد النصوص الشعرية نسبيا.

لقد أفرزت هذه المعايير في البداية مصادر أساسية نذكر من بينها:

أ – القرآن الكريم. ب – كتب السنة (صحيحا البخاري ومسلم، ومسند أحمد وموطأ مالك، وسنن الترمذي وأبي داود والنسائي والدارمي).

ج – كتب معرفية وأدبية سردية تتضمن موضوعات مختلفة. د – مقالات متنوعة مأخوذة من صحف ومجلات عربية ذات انتشار واسع، ولقد روعي في اختيار النصوص الأدبية الصفة التمثيلية لمؤلفيها على نطاق أقطار العالم العربي.

4– يعتمد المعالج الصرفي لمركز التراث على معالجة مفردة واحدة حسب اختيار المستخدم، ويباشر بإظهار أهم الجذور المتاحة لتلك المفردة في الموسوعة قيد الدراسة، مع إتاحة الفرصة لانتقاء الجذر المناسب لعملية البحث، بالمقابل فإن المحلل الصرفي الآلي للمنظمة العربية للتربية والثقافة والعلوم فيعمل بمفرده بعيدًا عن دائرة الموسوعات، لتوفير معلومات مفيدة عن خصائص الصرف العربي، بيد أن بيئة عمله تنحصر ضمن بيئة MS-DOS التي لا توفر الإمكانيات الرسومية الفريدة التي تتيحها بيئتي Windows 98/ 2000 فلا يرقى في إمكانياته إلى سابقيه، ويقتصر في استخدامه على المستخدم الذي يعنى بالصرف العربي ضمن المستويات المبتدئة فحسب.

5– يقوم هذا المعالج بتفكيك الكلمة قيد التحليل عن طريق إزالة ما يتصل بها من سوابق (كالأدوات والحروف)، ولواحق (كالضمائر المتصلة، وزوائد التصريف، وعلامات الإعراب)، كذلك يعمد إلى رد التعديلات الفونولوجية التي تتم على حدود عناصر البنية الصرفية للكلمة إلى أصلها، كي تصبح صالحة لعملية المعالجة.

فعلى سبيل المثال، عند مباشرة هذا المعالج بعملية تفكيك الكلمة (والدين)، يظهر أمامه أكثر من احتمال لعملية التفكيك الصرف نحوي، وكما يلي:

الاحتمال الأول: "والد" + "ين" (مثنى مذكّر). الاحتمال الثاني: "والد" + "ين" (جمع مذكّر سالم).

الاحتمال الثالث: "و" + "الدَّين" (كلمة الدَّين من الفعل الثلاثي "دان"). الاحتمال الرابع: " و " + " الدِّين " (كلمة الدِّين من الفعل الثلاثي "دان").

ولتلافي الضياع في متاهات تحليل المبنيات؛ كالضمائر، وحروف الجر، وظرفي الزمان والمكان، تم توفير آلية ذكية يمكن أن توظف من خلال هذا المعالج لمقارنة جذع الكلمة المفككة مع قائمة من الكلم المبني باللغة العربية

2 ـ تجربة الدكتور الحناش[1] في المعاجم الآلية للغة العربية:

يتألف المعجم الإلكتروني للغة العربية من ثلاثة مستويات متشابكة ومتكاملة فيما بينها، كل مستوى منها يجسد مستوى لسانياً بدءاً بالمستوى الصوتي، ومروراً بالمستوى الصرفي، وانتهاءاً بالمستوى التركيبي[2].

2-1- المعجم الإلكتروني العربي للجذور: DEAR :

وهو عبارة عن قاعدة معطيات من المعلومات الصرفية[3] والنحوية[4] التي تتحكم في استدراج الجذر من الكلمة في اللغة العربية. وعلى الرغم من تنوع مدونة الجذور العربية وتعددها، فإنه قد تم بناء عذا المعجم، وذلك بإرفاق كل جذر لغوي بجميع الاحتمالات الاشتقاقية الممكنة، مما طبعه بالشمولية[5] والسهولة. وقد بني هذا المعجم على الشكل التالي[6]:

ض.ر.ب: (ف = 0+)، (مص = 1+)، (فو = 2+)، (س = 0+) ...الخ.

حيث أن: + = الحروف الزائدة عن الجذر.

− = الحروف الناقصة عن الجذر.

0 = ليست هناك زيادة في بنية الجذر.

1 = أو أي عدد آخر يفيد عدد الحروف الزائدة عن الجذر لتوليد الصيغة المشار إليها بالحرف قبل علامة =، وعندما يتعلق الأمر بنقص في عدد حروف الجذر يشار إليها بالرمز (ن) التي تعني نقصان.

إن صعوبة هذا المعجم تكمن في كيفية تشفير الحروف والأدوات التي لا تعرف جذوراً مطردة وعادية، كما هو الحال مع أقسام الكلم، فقدم تم النظر إليها من خلال صورتها الاستعمالية، لا صورتها الأصلية غير الموجودة أصلاً[7].

1- أ .د / محمد محمد أحمد الحناش أستاذ اللسانيات الحاسوبية والعامة في جامعة سيدي محمد بن عبد الله ، مدير مسؤول عن مجلة (التواصل اللساني) المجلة الدولية المحكمة في اللسانيات العامة – باحث في طرق تعليم اللغة العربية بها وبغيرها بأحدث الطرق البيداغوجية – وخبير في مهارات الاتصال في اللغة العربية، فاس – المغرب .

2- للمزيد من بسط القول ينظر الجانب النظري من هذه الأطروحة، الباب الثاني، الفصل الثاني، ص283،...،287.

3- : صافية زفنكي،التطورات المعجمية والمعجمات اللغوية العامة العربية الحديثة، منشورات وزارة الثقافة، دمشق، د.ط،2007، 210.

4- المرجع نفسه، ص 222.

5- الشمولية، ويقصد بها استيعاب كل عناصر الظاهرة اللغوية التي تقع في إطار هدف المعجم، سواء كان ذلك على مستوى المداخل أو الشروح والتفسيرات:

أ- شمولية مداخل المعجم: تعني احتواء المعجم على جميع المداخل القياسية والسماعية للغة العربية.

ب- شمولية الشروح والتفسيرات: تعني اشتمالها على جميع المعلومات الدلالية التي تكشف عن معنى المدخل المشروح وتبين سلوكه الدلالي والتركيبي.

6- عمر مهديوي، توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مرجع سابق،100.

7- المرجع نفسه ، ص 101.

2-2- المعجم الإلكتروني للمفردات البسيطة: DEAMS:

وفيه تثبت كل مفردة في قاعدة المعطيات بناء على المعلومات النحوية والصرفية المرتبطة بكل واحدة منها. كما يسند لكل مدخل معجمي رقماً هو عبارة عن سنن أو شفرة، يعتبر جزءاً في الخطاب اللساني، وسنن رقمي يمثل الإعراب[1].

لقد صنفت معظم المعاجم العربية المفردات اللغوية العربية بناء على خصائصها النحوية، وذلك بتمييزها بين الأفعال اللازمة والأفعال المتعدية بحرف واحد أو أكثر، كما صنفتها إلى أفعال عادية وأفعال قلوب وافعال ناقصة. في حين انبنى التصنيف في هذا المعجم على أساس أقسام الكلم مع الأخذ بعين الاعتبار الخاصية النحوية للمدخل المعجمي (المفردة)، وعلاقتها بالمستوى المورفولوجي المناسب لها (الصيغة والاشتقاق)[2]:

ف: الفعل في الزمن الماضي للمفرد الغائب المذكر.

س: للمفرد المذكر والمؤنث.

ص: الصفة وسائر المشتقات بجميع أنواعها.

ظ: الظرف بجميع أنواعه.

أما المصدر فلم يخصص برمز معين، إذ تم إدماجه في معجم المفردات البسيطة المعربة.

2-3- المعجم الإلكتروني للمفردات المعربة: DEAMSF:

وفيه تستند المعلومات الأساسية لكل مفردة معربة (تصريف الفعل ونسبته إلى الضمائر وعلاماته الإعرابية مع كل واحد منها الخ) بمعنى آخر، أن كل شكل من أشكال هذا المعجم يجب أن تعطى عنه جميع الاحتمالات التصريفية والصرفية التي يوجد عليها في اللغة العربية، وهذا ما يشكل القسم المتعلق بقاعدة معطيات المفردات البسيطة المعربة[3]:

الشكل أ = جذر ب، صرف ج

حيث يراد بالشكل المدخل المعجمي (المفردة البسيطة) وبالجذر الصيغة التي يظهر بها في المعجم نفسه، وبالصرف مختلف الأوجه التصريفية والصرفية التي تلحق المدخل المعجمي البسيط وهو ما يسمى بالإعراب الصرفي في المعجم الآلي[4].

1- المرجع السابق،الموضع نفسه.

2- المرجع نفسه ، ص 101.

3- المرجع السابق، ص 102.

4- المرجع نفسه، الصفحة نفسها.

2-4- المعجم الإلكتروني للمفردات المركبة: DEAMC:

إن أي برنامج معد للمعالجة الآلية للمفردات في نص معين لابد وأن يتعرف على جميع مفردات النص/ الهدف لكي تكون المعالجة شاملة وخالية من أي نقص وذلك بأخذ بعين الحسبان المفردات المركبة التي تشغل حيزاً مهما في الذاكرة اللغوية مثل[1]:

علامة المرور، وأحد عشر، ...

وهي مداخل معجمية لا تقبل التجزيء ولا التقطيع، لأنها تشكل مداخل مسكوكة، ومن ثمة تستدعي وصفاً صرفياً وآلياً من نوع خاص، أي الجمع بين نوعين من الوصف: العناية بوضعها المورفولوجي للعناصر التي يتألف من الاسم المركب من جهة، ومن جهة أخرى، وصف وضعها في التأليفية التي أنشأها منها الاسم المركب[2]. هذا الصنف من المعاجم يصنف على أساس قاعدة معطيات المفردات المركبة، والتي تفرع بدورها إلى مفردات مركبة، تضم العادي منها والمسكوك، ومفردات مركبة معربة[3].

3- تجربة الدكتور أبو العزم[4]: من الغني إلى الغني/ الشامل[5]:

إن أي معجمي يود صناعة معجم عربي حديث (ورقي أو إلكتروني)، لا غنى له من الاعتماد على معجم الغني باعتباره ذخيرة لغوية مفتوحة المصدر، لها من المتطلبات المعجمية المعجماتية ما يسهل عمله، ويجعله يتخطى ثغرات المعاجم العربية القديمة والحديثة.

أولا- خصائص مميزة لمعجم الغني:

يعتبر معجم الغني لصاحبه عبد الغني أبو العزم من المعاجم الأساسية التي تخدم مباشرة طلاب المراحل الأساسية والثانوية ومستخدمي المعاجم الحديثة[6]. فهو معجم وظيفي

1- المرجع السابق، ص 102.

2- المرجع نفسه ، الصفحة نفسها.

3- المرجع نفسه، الموضع نفسه.

4- يعتبر الأستاذ عبد الغني أبو العزم عميد الدراسات المعجمية المعجماتية بالمغرب بامتياز، لما أنتجه من معاجم ودراسات قيمة أغنت الخزانة المغربية والعربية، وسدت ثغرة في باب البحث المعجمي تأليفاً ونشراً. ويعد معجم الغني أشهر معجم عربي محوسب، حظي بالسمعة والتداول ما لم يحظ به غيره من معاجم اللغة العربية المؤلفة حديثاً، فهو يتميز بخصائص معجمية ومعجماتية قلما توجد في معجم آخر، فقد عب من معين التقنيات اللسانية عامة، وتقنيات الصناعة المعجماتية والآلية خاصة، حيث تجاوز الكثير من المعاجم الحديثة سواء على مستوى الشكل أو المضمون، مما يجعله بحق مصدراً لا غنى عنه في صناعة معاجم إلكترونية للغة العربية.

كانت مداخل الغني في البداية لا تتعدى 20000 مدخل، لكن سرعان ما أصبحت 30000 مدخل بفعل الإضافات والتعديلات التي أدخلها عليه المعجمي أبو العزم، مما أدى إلى استخلاص رصيد لغوي يضم طائفة موسعة من المفردات غير الموجودة أصلاً في المعجم، ومعان جديدة، ومصطلحات أدبية وعلمية، ناهيك عن عدد مهم من الاستشهادات القرآنية والنبوية، والأقوال المأثورة، والأمثال والحكم، والمتلازمات.

5- أي النسخة المعدلة والمزيدة من الغني.

6- إن مهمة المعجم" معجم الغني" هي لوصف اللغة العربية شرحا وتفسيرا، فترى أنه يزود المداخل المعجمية ويكثف المعلومات حولها سواء بدافع التوصيف الصرفي والتركيبي أو بدوافع التنويع في العلاقات الدلالية للمفردات المعجمية وجعلها تنفتح على كل مشترك لفظي أو ترادف من شأنه أن يغني هذه المداخل في مدارك المتلقن، وهذا ما سيصبح في اعتقادنا قاعدة أساسية لبناء معجم إلكتروني قائم بنصوصه الفائقة والمترابطة دلاليا يضم مئات الشواهد الأدبية.

عملي، مرتب ترتيبا ألفبائياً، يعزز الشواهد بتعريفات بسيطة وشواهد أدبية. يشرح الكلمات والتعبيرات المتداولة وقديمها وحديثها. كما يضم عددا ضخما من المصطلحات الأدبية والعلمية والتقنية وألفاظ الحضارة ويحتوي على رسوم إيضاحية لمختلف الحيوانات والنباتات.كما أن هذا المعجم يساعد على معرفة دلالة الكلمة ومضامينها الحقيقية والمجازية ومعرفة صيغها وتراكيبها بالإضافة إلى تثبيت اللغة والعلم لدى المستعملين طلابا وباحثين. وما يجعله أكثر تميزا هو كونه يعتمد الترتيب الألفبائي مصحوبا بالنطق الإملائي للمداخل المعجمية .

ثانيا-اختيار مدونة[1] المداخل اللغوية:

يتميز هذا المعجم عن سائر المعاجم المحوسبة الأخرى بما يقدمه من شواهد من القرآن والحديث والأمثال والنصوص الأدبية قديما وحديثا، وبما هو متداول ومستعمل بمعانيه المختلفة في لغة الأدب والصحافة المكتوبة والسمعية والبصرية والكلمات المولدة المعربة والدخيلة التي أقرتها المعاجم اللغوية، وكل هذا يعبر عن رغبة في اعتماد مدونة شاملة وغنية[2]. وهذا من شأنه أن يغني رصيد المستعملين اللغوي ويقويه، كما ينمي الملكة الأدبية التعبيرية لديهم ببساطة ويسر وسهولة،بما أن المعاني في مدونته تتعدد فإنه أرفق لكل معنى رقما[3]. حتى يسبح المتلقي في مراتب الدلالات بنوع من التنظيم والانسياب والمرونة، إن اعتماد صاحب المعجم على

1- إن التصور الذي وضعه صاحب المعجم، الدكتور عبد الغني أبو العزم، قاده إلى مدونة مرنة قابلة للتعديل والتطور والديمومة. الذي اعتمد فيه المصادر المعجمية قديمها وحديثها فإنه ينظر إلى اللغة في سيرورتها التاريخية منذ العصر الجاهلي إلى اليوم دون فصل أو تمييز، فلا وجود، إذا للغة عربية قديمة وأخرى معاصرة 1 ويذهب الدكتور مصطفى غلفان، تعقيبا على هذا الرأي، إلى أن اعتماد مصادر معجمية هي بهدف تكريس معجم عربي موحد تحت مبررات تربوية، فنحن أمام لغات عربية وأمام تنوع لغوي يضم لغات متقاربة جدا1 وليست عربية واحدة بمعجم واحد لكن وتجنبا لهذه الاشكالية في الاحتجاج والوضع اهتدى صاحب "معجم الغني" إلى استقصاء مدونته إنما يكون من جانب تقني قد سقط سهوا منه. فهذا بحثنا عن مادة "أدنى " في "لسان" مثلا فإننا لن نعثر على ذلك (وربما بباقي المعاجم القديمة أيضا) بينما " معجم الغني" يورد مثل هذا المدخل، وهذا يدلل على أن أدنى بمعنى "أخفض" هو معنى معاصر وليس معنى قديم. مما يؤكد ارتباط نصوصه المعجمية بلغة العصر ونورد للتمثيل كذلك، مقارنة بين مدخل "بتّ" في " معجم الغني " و"لسان العرب"1 كما في الجدول التالي:

مادة"بت"	معجم الغني	لسان العرب
	بَتْ-[ب ت ت] (ف: ثلا. لازمتع. م. بحرف). بَتَتْتُ، أَبَتّ، بُتّ، مص.بَتّ. 1."كَلَفَهُ بأَنْ يَبُتّ في الأَمْرِ":أَنْ يَحْسِمَ فِيهِ. 2."بَتّ الأَمْرَ":أَنْفَذَهُ، أَمْضَاهُ. 3."بَتّ الصِّيَامَ من اللَّيْلِ"نَواهُ وقَطَعَ به. 4."بَتّض طَلاَقَ إمرأَتَه" :جَعَلَهُ بَاتّاً لا رِجْعَةً فِيهِ.5."بَتّ الرَّئيسُ في بُنُود القَانُون" قرَّرَ. 6."بَتّ عَلَيْهِ الشهَادَةَ" :أَلْزَمَهُ بِهَا	البَتُّ القَطْعُ المستَأصل .يقال:بَتَتُّ الحبل فانبتّ .ابن سيده: بتّ الشيء يبُتّهُ وبِتّهُ بتًّا وأَبَتَّه قطعه قطعاً مستأصلاً.

للمزيد من التفصيل ينظر: د.مصطفى غلفان ، مراجعات وعرض للكتب، مجلة الدراسات المعجمية ع 3و4 يناير 2005 ص.232/231.
وينظر مادة "بت" في معجم الغني (القرص) والمدخل المعجمي "بت" في لسان العرب(برنامج جامع معاجم اللغة).

2- يعتزم المؤلف في اصداره الثاني "معجم الغني الشامل"، على أن يعتمد مدونة موسعة شاملة لجميع ما هو مستعمل ومتداول في اللغة العربية بالمعاني والدلالات المجازية والأسلوبية المختلفة،معززة بالشواهد والعبارات والأمثلة المسكوكة. والهدف من هذا هو أن تخضع المدونة للإطار العام لطبيعة اللغة وأن تكون النسبة في التوارد أعلى بكثير من تلك التي مع معجم الغني حاليا وهي ما بين 3 بالمائة إلى30 بالمائة.للمزيد ينظر، مقدمة المعجم (مرقون) 2007

3- يعتبر "معجم الغني " مرجعا لحفظ معاني المفردات والجمل والمصطلحات وغيرها، ولم يقتصر على المفردات في شرح المداخل بل أراد به صانعه أن يصبح مجالا للبحث لتعليم الأساليب اللغوية السليمة لتمكين المستعملين من بناء نصوص عربية سليمة يتوفر فيها الحد الأقصى من الفصاحة اللغوية. وهذا لم يتأتاه إلا باتباع تصورٍ شاملا حول المعجم العربي مكنه من وضع صورة حية لفهم البرنامج اللغوي الكامن في الكفاية اللغوية للمتكلمين العرب، لذلك يركز " معجم الغنيّ" على تقديم أكثر من تعبير واحد يكون شرحا لدلالة المدخل (حوالي 45 شرحا).

مجموعة من المعاجم العربية القديمة والحديثة دليل على استقصائه لمختلف مراحل اللغة وتطور دلالات تعبيراتها وأساليبها ومجالاتها المعرفية والحضارية والعلمية والتكنولوجية مما سيؤدي إلى اكتساب المعرفة بأساليب العربية وبلاغتها مجازاتها والتمكن من الدقة في التعبير والكتابة[1]. وهذا يقتضي الاهتمام بإيراد أكبر عدد من مترادفات للمفردة الواحدة لتوسيع آفاق معانيها المتعددة وتنمية المعرفة اللغوية[2]، التي ستتشكل من خلالها قواعد معطيات نصية معجمية للمعجم الغني الشامل وهذا ما جعله يفكر في وضع مداخل معجمية متناسقة ومترابطة فيما بينها من حيث معالجتها في جزئياتها وفي شموليتها[3]. (المدونة في "معجم الغني" موسعة)[4]، إن لم نقل قابلة للتوسع، تقبل التعديل والإضافة شريطة الجودة في توضيح المعاني والدلالات.

ثالثا- المنهجية المعجمية المتبعة في "معجم الغني":

يجمع معجم الغني بين المناهج المعجمية القديمة والمناهج المعجمية الحديثة والمعاصرة ولا أريد هنا الغوص في قضايا المعجم العربي الكثيرة والمتشعبة، بل أود الإشارة فقط إلى الاتجاهات التي رسمتها النظرية المعجمية الحديثة والمتمثلة في التحديث والتيسير وذلك باستغلال التكنولوجيا المتطورة وخاصة الوسائط المتعددة من صوت وصورة. مما أدى إلى نقل بعض المعاجم الورقية[5] ووضعها على أقراص مدمجة ومكثفة مثل " جامع معاجم اللغة " الذي يحتوي مجموعة من القواميس العربية القديمة المحوسبة، وإما إلى وضع معاجم حاسوبية جديدة تتماشى مع تكنولوجيا البرمجيات اللغوية المتقدمة مثل ما أصدرته شركة صخر ميديا والمتمثل في "معجم الغني" في اصداره الأول الذي يحمل الرقم: 1.01 الذي تميز عن غيره في الترتيب والتعريف والشرح والشاهد[6]، لقد تحكمت المعجم غايات وظيفية كبرى مما جعله يعتمد منهجية مغايرة في الترتيب على الرغم من كونها على الطريقة الألفبائية، لقد كان هذا الرجل واعيًا تمام الوعي بالبعد اللاخطي الذي تحكم النظرية المعجمية المعاصرة بأبعادها المعرفية والصورية، لذلك منح الابعاد المعرفية حقها شكلا وصوتا وكل ما من شأنه أن يعيد للمعجم العربي رونقه وريادته، فتراه حاول مزج تصور نظرية الخليل المعجمية بشكل ذكي حينما وظف التقنيات الصوتية في وضع المداخل مدغَمة، إذ لم يرد تمثيلها تمثيلا خطيا كما عهد في العديد من

1- د.عبد الغني أبو العزم، مقدمة المعجم الشامل (مرقون) 2007 ص6.

2- المرجع نفسه، الموضع نفسه.

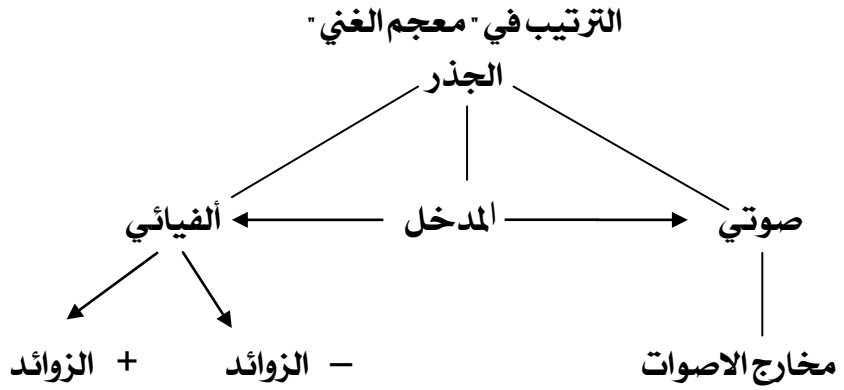
3- المرجع نفسه، ص 12.

4 – corpus extensionnelle

5- اضافة إلى المعاجم الورقية المحوسبة التي جمعت في باقة "جامع المعاجم العربية" هناك أيضا معاجم أخرى كقاموس المحيط ومحيط المحيط والوسيط ونجعة الرائد صادرة عن شركة صخر ميديا للبرمجيات العربية ومؤلفات معجمية موضوعة رقميا في مواقع إلكترونية عربية مثل موقع مكتبة الوراق أو المدونة.

6- ما يميز هذا المعجم عن سائر المعاجم الأخرى أنه يقدم شواهد من القرآن الكريم والحديث والأمثال والنصوص الأدبية قديما وحديثا مما يغني الرصيد اللغوي للمستخدمين بشتى أصنافهم كما ينمي الملكة الأدبية التعبيرية لديهم بيسر وسهولة.

القواميس العربية القديم منها والحديث، بل ذهب إلى اعتبار البعد اللاخطي في وضع المداخل المعجمية[1] أمرا ذي أهمية بالغة، مما يستجيب لصياغة صورية متقدمة سيتم استغلالها فيما بعد في بناء معجم اللغة العربية الصوتي على غرار اللغات الأجنبية مثل الفرنسية والانجليزية[2]. وللتوضيح فيما يلي شكل يبين عملية الترتيب في "معجم الغني".



الشكل رقم 04 : يبين الترتيب في معجم الغني.

رابعا ـ معطيات إحصائية عن المعجم:

معجم الغني هو قاموس عربي– عربي يحتوي على تعريف ل 30,000 مادة ومصطلح بوجود أكثر من 195,000 كلمة مشتقة فيه[3]. من صفات القاموس أنه مرتب الكلمات ألفبائياً ، بالإضافة إنه يحدد العربية الأصل من الكلمات الدخيلة والمعربة. وتدلنا الأرقام عن المعجم على أن هناك حوالي 41 ألفاً[4] من التعبيرات المتلازمة والمسكوكة أدرجت في ثنايا شرح مواد المعجم بما يمثل نسبة حوالي 31% من مجموع كلمات المعجم كما هو مبين في الجدول أسفله:

1– للمزيد من بسط القول ينظر : جواد حسني سماعنه، المعجم العلمي المختص، (المنهج والمصطلح)، سابق، ص 42، 43.

2– يمكن استغلال هذه المعطيات في بناء برمجيات لغوية متقدمة مثل توليف الكلام synthèse de la parole والتعرف الآلي على الكلام (reconnaissance automatique de la parole)والنقل الآلي للنص المكتوب إلى الكلام المنطوق والعكس =(text-to-speech) وغير ذلك من البرمجيات التطبيقية العربية والتي أنجزت منها شركة صخر مما قد يحتاجه الباحثون والمترجمون والمعالجون اللغويون.

3– صرح د. عبد الغني أبو العزم في مقدمة المعجم على أن الإصدار الأول منه كان منصبا لبناء معجم للأطفال، وهذا ما هيأ لرصيد معجمي يستهدف المراحل التعليمية المتوسطة. مما جعله يعتمد على مدونة معجمية جلها مدرسية رائد الطلاب والقاموس المدرسي والقاموس الجديد والرصيد اللغوي لمدارس المغرب العرب والرصيد اللغوي المغربي ومنجد الطلاب والمنجد الإعدادي، ينظر مصطفى غلفان مراجعات وعرض الكتب مجلة الدراسات المعجمية ع 3 و4 يناير 2005) . وهذا مما حصر مدخله في 30.000 مدخل هي قاعدة معطيات معجم الغني الذي أصدرته شركة صخر، وهذا التركيز جعل التواتر يصل إلى نسب مابين 30 % و 90 % وأحيانا مابين 5 % إلى 30 % على عكس المدونة المعجمية في معجم "الغني الشامل" التي تشكل توارده مابين 3 % و30% لأن المدونة خضعت للإطار العام لطبيعة اللغة وهذا ما سيجعله معجما شاملا لجميع الشرائح ذووا المستويات العلمية والثقافية المختلفة. (انظر مقدمة معجم الغني الشامل مرقون 2007) .

4– لقد تم اختيار عينة المتلازمات للقيام بالعملية الإحصائية للتعبيرات المتلازمة في "معجم الغني" وعلى هذا الأساس تم تجريد كل مدخل وكانت نتيجة ذلك الحصول على المداخل التي توجد به أعلى نسبة (أكثر من عشرة أسلوب وتعبير متلازمة ومركب) وعينة لا توجد بها أي نسبة تذكر (أقل من أسلوب أو تعبير واحد) فتم جمع العدد الإجمالي للمداخل واحتسبنا النسبة الإجمالية انطلاقا من المجموع العام وكانت النتيجة الحصول على حوالي 41 ألف من التعبيرات المركبة والمتلازمة والمسكوكة في النصوص المعجمية لهذا المعجم .

معجم الغني	متوسط المواد للحرف	عدد المواد	عدد المشتقات	عدد الكلمات	عدد التعبيرات المتلازمة والمتواليات المركبة	النسبة المائوية لمجموع التعبيرات المتلازمة والمركبة في المعجم
	1.071	30.000	195.000	2.000.000	41000	%31

جدول يوضح نسبة التعبيرات المتلازمة والمركبة في "معجم الغني"

وجملة القول :

إن متصفح هذا المعجم الذي اعتمد صاحبه على قاعدة معطيات معجمية شاملة[1] يخال أنه أمام معجم يسير في توفير الوقت للمستخدم، وذلك بإجراء السرعة في البحث والجمع والنقل المباشر من القرص الحامل للمعجم إلى البرامج المعالجة للنصوص الحاسوبية. كل هذا أملته الأدوات العقلانية للمعلوميات النظرية وما توصلت إليه اللسانيات الصورية والمعرفية، مما يستدعي إعادة تحديد التمثيل الهرمي المفاهيمي للظواهر المعجمية بصفة عامة، لتوضيح العلاقات القائمة بينها والطريقة التي تندرج فيها. والهدف هو إعادة تشكيل البنيات اللسانية والمعجمية لتحديد بدقة متناهية الحقول والتصورات والمفاهيم والمصطلحات والألفاظ المركبة والبسيطة الدخيلة والمولدة في أفق بناء قاعدة معارف معجمية للصور العربية المولدة[2] مما يستدعي ضبط اللغة في مختلف المسارات تجديدا وإضافة[3].

1- اعتمد صاحب " معجم الغني" مدونات وقواعد بيانات معجمية موسعة مما جعل هذا المعجم قابل للإستفادة بشكل كبير رمن أساسيات المعطيات النصية الشاملة للغة العربية في قديمها وحديثها والاستزادة منها، فحضور الجانب التعليمي بقوة في المعجم، هو ما يجعل منه بالفعل أداة عملية ووسيلة للتربية والتلقين واكتساب المعارف والمهارات. وهذا ما أدى به إلى إعتماد أساسيات موسعة للمعطيات النصية والمعجماتية ستمكنه من وضع التعبيرات المتلازمة في منزلة الشرح التعريف والارتقاء بالمتعلم إلى مراتب عليا في اللغة، رغم وضعها الثانوي في خدمة المداخل المعجمية، ضمن سياقاتها وأوضاعها البلاغية والاستعارية في معجم اللغة العربية. وعلاوة على مهمته اللغوية المتمثلة في توضيح دلالة المفردات،فإنه يسعى إلى توفير كافة المعلومات الأساسية التي يحتاجها المتعلم لغويا ومعرفيا فأما اللغوية فنقصد بها الإمكان الذي يتيحه المعجم المذكور من تقديم بحث في مترادفات المفردات وأضدادها والوقوف على المشترك اللفظي وتقديم التعبيرات المتلازمة والأمثال اللغوية والمتداولة .فإن "معجم الغني"يقدم الشروح اللغوية لكل المفردات ،ولكن أيضا لبعض العبارات المستعملة والمفردات المركبة والمقترضة من لغات أجنبية مثل (اليوبيل الذهبي/اليوبيل الفضي/اليوبيل الماسي(انظر مادة يوبيل)/الأمم المتحدة/يونيسكو/ايسيسكو/يونيسف) وما شابهها من ألفاظ الحضارة وعباراتها المستحدثة في الثقافة العربية الحديثة والمعاصرة.

2- عبد القادر الفاسي الفهري انشاء قاعدة معجمية عربية مولدة ،المعجم العربي المولد ،منشورات معهد الدراسات الأبحاث للتعريف بالرباط، ص19.

3 – عبد الغني أبو العزم ، مقدمة المعجم ، مرقون 2007 ص1.

4- تجربة مروان البواب[1] : إعداد المعجم العربي الحاسوبي للغة العربية:

يحتوي هذا المعجم على قواعدِ معطياتٍ وجداولَ، ويتضمَّن قوانينَ تُمكّنه من عرضِ جميع المعارف المعجمية بسهولةٍ ويسر، ومن إجراء عملياتِ بحثٍ متنوعة[2]. ومن مزايا أذكرُ منها ما يلي:

1- قدرته على إيراد جميع المفرداتِ الأصليةِ والفرعية والقياسيّةِ؛ أما المفردات الأصلية فيعتمد في إيرادها على قواعد المعطيات المخزنة فيه، شأنُه في ذلك شأنُ المعاجم التقليدية. وأما المفردات الفرعية، كالمؤنث والمثنى والنسبة والتصغير، فيعتمد في إيرادها على وجود مقابلاتها الأصلية. وأما المفردات القياسيّة، كاسمِ الفاعل واسمِ المفعول ومصادرِ الأفعال فوق الثلاثية... فيعتمد في إيرادِها على قواعدِ الاشتقاق وعلى وجود الأصل المشتق منه في جداول المعطيات.

2- قدرتُه على تصريفِ الأفعال والأسماء في جميع حالاتِها الصرفيةِ والنحوية، فهو يصرّفُ الأفعالَ في صيغة الماضي، وفي صيغة المضارعِ مرفوعةً ومنصوبةً ومجزومةً ومؤكّدةً، وكذلك يصرّفُ فعلَ الأمرِ مؤكّداً وغير مؤكّد. أما الأسماءُ فيصرِّفُها في جميع صورِها: المفرد والمثنى والجمع، والمذكرِ والمؤنث، في حالة الرفع أو النصب أو الجر، سواء أكانت مُعَرَّفَةً أم مُنَكّرةً أم مضافةً. ولا شك في أن هذه التصاريف الكثيرة للأفعال والأسماء يستحيل وضعها في المعاجم الورقية، لأن حجمها سيصير أضعافاً مضاعفة، [قرابة 50,000 صفحة لتصريف الأفعال وحدها]. ولا يتحقق هذا المطلب (أي القدرة على تصريف الأفعال والأسماء) إلا إذا تضمن المعجم الحاسوبي نظاماً حاسوبياً* للاشتقاق والتصريف.

3- اعتمادُه في عرضِهِ للمعارفِ اللغويةِ على الوسائلِ الحاسوبيّةِ الحديثةِ المتعددةِ الوسائط (Multimedia)، [النصوص، والأصوات، والصور الثابتة والمتحركة، وأفلام الفيديو]، إضافةً إلى إمكانية التحكمِ في أحجامِ الخطوطِ وأنواعِها وألوانِها.

4- سهولَةُ التعاملِ معه وسرعةُ أدائِهِ، إضافةً إلى إمكانيّة عملِهِ على حواسيبَ محمولةٍ، صغيرةِ الحجمِ خفيفةِ الوزن.

5- سهولةُ تحديثه، لمواكبة ما يستجد في المفردات والمعاني والأمثلة وغيرها.

1- أ. مروان البواب، (منهج إعداد المعجم العربي الحاسوبي للغة العربية)، الاجتماع الثاني لخبراء المعجم الحاسوبي التفاعلي للغة العربية، 1429/5/2-4/29هـ (2008/5/7-5/5) .

2- ذكر مروان البواب في بحثه المقدم في المؤتمر السابع لمجمع اللغة العربية بدمشق والمعنون بـ: "أثر التقانات الحديثة في تجديد المعجم العربي" أهم التقنيات المستعملة في تجديد المعجم العربي كالمكتبات الإلكترونية، والكتب الإلكترونية، ومواقع الصحف والمجلات، والماسح الضوئي، ومواقع البحث في الإنترنت، وبرامج المعالجة المختلفة وغيرها، كما بين أوجه الاستفادة من هذه التقنيات الحديثة في تجديد المعجم العربي، وأعطى الكثير من الأمثلة على وجوه تلك الاستفادات، كما عدد بعض مزايا استخدام المعجم الحاسوبي في خدمة اللغة العربية ودارسيها من حيث إحكام تبويبه، ووضوح أسلوبه، ودقة تعاريفه، وإمكانية إجراء عمليات بحث متقدمة فيه، وسهولة تحديثه لمواكبة ما يستجد في المفردات و المعاني، وغيرها من المزايا.

6- القدرةُ على التعاملِ مع أنظمةِ معالجةٍ حاسوبيّةٍ للغة العربيّة، كالتحليلِ الصرفِيِّ[1] والنحويِّ والدلاليِّ والصوتيِّ وغيرِها. فجميعُ هذه الأنظمةِ تحتاجُ إلى معجمٍ حاسوبيّ يزوّدُها بالمعارفِ اللغويّةِ المطلوبةِ في أثناء المعالجة. ولا يَخفى ما لهذه المعالجةِ من أهميةٍ تتجلى في التطبيقاتِ العديدةِ المتوخاةِ منها، كالترجمةِ الآليّةِ، وتعلّمِ العربيّة وتعليمِها، واكتشافِ أخطاءِ النصوصِ وتصحيحِها، وتعرُّفِ الكلامِ وتركيبِه، والقراءةِ الآليّةِ للنصوصِ المكتوبة، والكتابةِ الآليّةِ للنصوصِ المحكيّة، وغيرِها وغيرِها...

7- قدرته على احتواء عدد من معجماتِ اللغة العربيّة القديمة *والحديثة، كالقاموسِ واللسانِ والتاجِ والمعجمِ الوسيطِ وغيرها، وبذلك يستطيع الحاسوبُ عرضَ ما يطلبه المستثمر من المعجم الذي يختاره.

وفي هذا اختصار للجُهْدِ والزمنِ اللازمَيْن للبحثِ في عددٍ غيرِ قليلٍ من المجلّدات، وغُنْيَةٌ عن شراء عشراتِ المجلّداتِ من المعجماتِ المطبوعة، بل وعن مكتبةٍ لوضع هذه المجلدات!! ويقدّر هذا الأمر حَقَّ قدْرِهِ مَنْ يُضطَرُّ إلى مراجعةِ المعجماتِ مرّاتٍ ومرّات.

- **ممّا تتكون قاعدة المعطيات ؟**

إنّ قاعدة المعطيات[2] عبارة عن أربعة جداول أساسية ، يرتبط بها واحد وثلاثون جدولا فرعيا ، أمّا الجداول الأساسية فهي : "**جدول الأفعال** ، **جدول الأسماء** ، **جدول الأدوات** ، **جدول الإحصاء**" .

ويمكن توضيح ذلك من خلال الشكل رقم 1[3]:

1- يتمثل التوصيف الحاسوبي للجانب الصرفي في المعجم، فيما يلي:

أ- تحليل مداخل المعجم إلى عناصره الصرفية الأساسية، وهي: الجذر والوزن والسابقة واللاحقة.

ب- تحديد الصفات الصرفية الأساسية لكل مدخل، كتحديد نوعه من الخصائص التالية: التعيين، النوع، العدد، الجمود والاشتقاق ... إلخ.

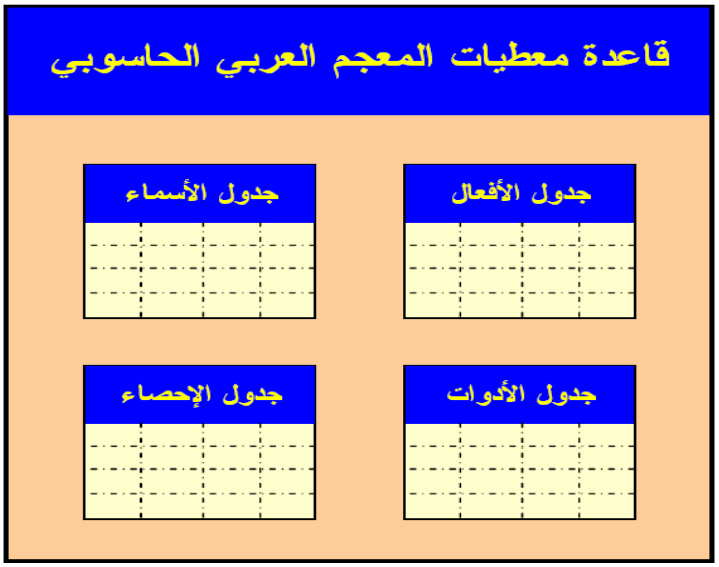
ت- الاشتمال على جميع الجذور اللغوية التي تقع في دائرة اهتمام مستخدم المعجم.

ث- الاشتمال على جميع التنوعات الصرفية المنبثقة عن الجذر اللغوي الواحد، سواء كانت قياسية أو سماعية، ما دام أنها تقع في إطار الظاهرة اللغوية التي يعالجها المعجم. وهذا الإجراء قد يترتب عليه تضخم حجم المعجم، ويتم التغلب على ذلك بالاستعانة بمحلل صرفي يمكنه الانتقال من الجذر اللغوي إلى أي تنوع صرفي منبثق عنه، كما يمكنه العكس، وبذلك تشتمل قاعدة بيانات المعجم الحاسوبي على نوعين من المداخل، النوع الأول يمكن تسميته بمداخل حيادية وهي الصيغ المجردة من السوابق واللواحق، والملتزمة حالة الإفراد والتذكير إذا كانت اسما، وحالة المضي والنسبة لضمير الغيبة المفرد المذكر إذا كانت فعلا، وهي المداخل الموجودة بذواتها في المعجم. والنوع الثاني يمكن تسميته بالمداخل الافتراضية وهي غير موجودة بذواتها، ولكن يمكن للمحل الصرفي توليدها إذا احتاج إليها مستخدم المعجم، أو الربط بينها وبين صورتها الحيادية الموجودة في المعجم.

ج- تحديد طبيعة العلاقات الصرفية المختلفة بين الصيغ الحيادية وكل تنوع صرفي مرتبط بها.

2 - Base des données.

3- للمزيد من التفصيل ينظر : أ. مروان البواب، (منهج إعداد المعجم العربي الحاسوبي للغة العربية)، الاجتماع الثاني لخبراء المعجم الحاسوبي التفاعلي للغة العربية، 29/4-2/5/1429هـ (5/5-7/5/2008) ص 8.



الجدول رقم 01: مكونات قاعدة المعطيات في المعجم الإلكتروني.

ويمكن توضيح الهدف من إنجاز هذه الجداول فيما يلي :

ب- **جدول الأفعال** :

يشتمل هذا الجدول الأفعال الثلاثية (المجردة والمزيدة) والأفعال الرباعية (المجردة والمزيدة)، أمّا الأفعال الجامدة وأسماء الأفعال والحروف المشبهة بالفعل فموضعها في جدول الأدوات[1]. يتألف هذا الجدول من سبعة أعمدة توضع فيها المعلومات الآتية :

1) **العمود الأوّل** : رقم مفتاح الفعل.

يستعمل هذا الرقم لربط الجداول بجداوله الفرعية ، وهذا ما سنوضحه بعد قليل.

2) **العمود الثاني** : جذر الفعل.

تراعى في جذر الفعل النقاط الآتية[2] :

- تكتب حروف الجذر على الشكل الآتي :

مثال الجذر (كَتَبَ) ← (ك ت ب).

- إذا كان الفعل مضعّفا يفكّ تضعيف جذره.

مثال الجذر (عدَّ) ← (ع د د)

1- المرجع السابق ، ص 9.

2- للمزيد من التفصيل ينظر : أ. مروان البواب ، (منهج إعداد المعجم الحاسوبي للغة العربية) ، ص 9.

- تردّ الألف إلى أصلها الواوي أو اليائي

مثال الجذر (قال يقول) ← (ق و ل)

- تكتب همزة الجذر على السطر :

مثال فجذر الفعل (أَمَلَ يَأمُلُ) ← (ء م ل)

- لا توضع حركات على حروف الجذر.

- تدرج الأفعال الرباعية المضاعفة في جذرها الرباعي.

فجذر (زَلْزَلَ يُزَلْزِلُ) ← (ز ل ز ل)

- تدرج الأفعال الملحقة بالرباعي في جذرها المستقل بها.

مثال : جذر الفعل (سَيْطَرَ يُسَيْطِرُ) ← (س ي ط ر) لا (س ط ر)

3) **العمود الثالث** : الفعل (الثلاثي المجرد والمزيد ، والرباعي المجرد والمزيد)،يذكر الفعل في صيغتي الماضي والمضارع ، تام الضبط.

4) **العمود الرابع** : وزن الفعل.

تراعى في وزن الفعل النقاط الآتية[1] :

- إذا كان الفعل ثلاثيا مجردا ، يذكر رقم بابه التصريفي كما في الجدول الآتي[2] :

الباب	وزن الفعل	رمز الوزن
الأوّل	فعَل – يَفعُل	1
الثاني	فعَل – يَفعِل	2
الثالث	فعَل – يَفعَل	3
الرابع	فعِل – يَفعَل	4
الخامس	فعُل – يَفعُل	5
السادس	فعِل – يَفعِل	6

1- المرجع السابق ، ص 9.

2- المرجع نفسه ، ص 10.

– وإذا كان الفعل ثلاثيا مزيدا أو رباعيا ، ذكرنا الوزن بصيغته ، فأوزان الثلاثي المزيد : أَفْعَلَ، فَعَّلَ ، فَاعَلَ ، انْفَعَلَ ، افْتَعَلَ ، أَفْعَلَّ ، تَفَعَّلَ ، تَفَاعَلَ ، اسْتَفْعَلَ ، افْعَوْعَلَ ، افعوَّلَ، افْعَالَّ.

– ووزن الرباعي المجرد : فَعْلَلَ ، نحو : سَيْطَرَ ، جَوْرَبَ...الخ.

وأوزان الرباعي المزيد : تفعلل ، افعنلل ، افعللَّ ، نحو تمسكن ، تشيطن ، تمنطق...الخ.

– وإذا كان الفعل مبنيا للمجهول يذكر في حقل وزن الفعل : مبني للمجهول.

5) **العمود الخامس** : نوع الفعل من جهة اللزوم والتعدية[1] :

– فإذا كان الفعل لازما ، يكتب في حقل نوع الفعل : (لازم).

– وإذا كان متعديا بنفسه ، يكتب (متعد بحرف).

– وإذا كان متعديا إلى مفعولين أحدهما بنفسه والآخر بحرف يكتب (متعدّ بنفسه وبحرف).

– وإذا كان متعديا بنفسه إلى مفعولين يكتب (متعدّ إلى مفعولين).

– وإذا كان متعديا إلى ثلاثة مفاعيل يكتب (متعد إلى ثلاثة مفاعيل).

6) **العمود السادس** : حرف التعدية.

يكتب في هذا الحقل حرف التعدية إن كان الفعل متعديا بحرف.

7) **العمود السابع** : معنى الفعل.

ينفرد كل معنى من معاني الفعل بسطر مستقل. وهذا ما يوضحه الجدول الآتي :

1– المرجع السابق، ص11.

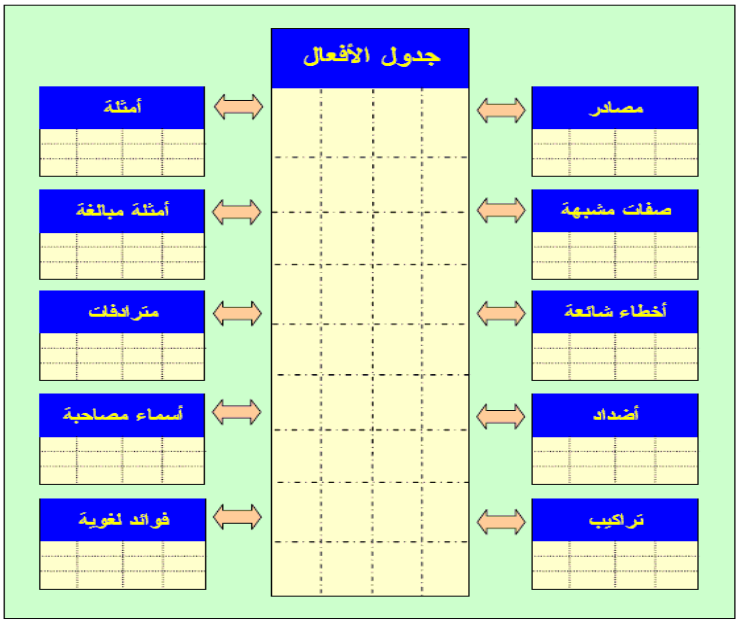
رقم الفعل	الجذر	الفعل	وزن الفعل	اللّزوم والتعدية	حرف التعدية	المعنى
24842	ص د ر	صَدَرَ يَصْدُرُ	1	لازم		صدر الحكم أو الأمر : وقع وتقدّر
24843	ص د ر	صَدَرَ يَصْدُرُ	1	لازم		صدر الكتاب ونحوه : ظهر
24844	ص د ر	صَدَرَ يَصْدُرُ	1	لازم		صدر الناس: خرجوا من قبورهم إلى الحشر ﴿يَوْمَئِذٍ يَصْدُرُ النَّاسُ أَشْتَاتًا لِّيُرَوْا أَعْمَالَهُمْ﴾[1]
24845	ص د ر	صَدَرَ يَصْدُرُ	1	متعد بحرف	عن	صدر الشيء عن غيره : نشا ونتج.
24846	ص د ر	صَدَرَ يَصْدُرُ	1	متعد بحرف	عن	صدر فلان عن المكان والورد: رجع وانصرف.
24847	ص د ر	صَدَرَ يَصْدُرُ	1	متعد بحرف	إلى	صدر فلان إلى المكان : انتهى إليه.
24848	ص د ر	صَدَرَ يَصْدُرُ	1	متعد		صدر فلاناً : اصاب صدره.
24849	ص د ر	صُدِرَ يُصْدَرُ	مبني للمجهول	لازم		صُدر فلانٌ : شكا صدره.
24850	ص د ر	أَصْدَرَ يُصْدِرُ	أَفْعَلَ	متعد		أصدر الأمر : أذاعه وأمر به.
24850	ص د ر	أَصْدَرَ يُصْدِرُ	أَفْعَلَ	متعد		أصدر الناشر كتابا : نشره ووزّعه.
24851	ص د ر	أَصْدَرَ يُصْدِرُ	أَفْعَلَ	متعد		أصدر الارعي ماشيته : صرفها عن الماء بعد ارتوائها. ﴿ قَالَتَا لَا نَسْقِي حَتَّى يُصْدِرَ الرِّعَاءُ﴾[2]
24852	ص د ر	أَصْدَرَ يُصْدِرُ	أَفْعَلَ	متعد بنفسه وبحرف		أصدر فلانا عن الشيء : صرفه عنه.
24853	ص د ر	صَادَرَ يُصَادِرُ	فَاعَلَ	متعد بنفسه		صادرت الدولة الأموال أو الممتلكات: استولت عليها.
24854	ص د ر	صَادَرَ يُصَادِرُ	فَاعَلَ	متعد بنفسه وبحرف		صادره على كذا: طالبه فيه بإلحاح.
24855	ص د ر	صَدَّرَ يُصَدِّرُ	فَعَّلَ	متعد		صدّر فلانا: قدّمه وأجلسه في صدر المجلس.
24856	ص د ر	صَدَّرَ يُصَدِّرُ	فَعَّلَ	متعد		صدّر المؤلف الكتاب: افتتحه بمقدمة.
24857	ص د ر	صَدَّرَ يُصَدِّرُ	فَعَّلَ	متعد		صدّر البضاعة والسلعة: أرسلها من بلد إلى آخر.
24858	ص د ر	تَصَدَّرَ يَتَصَدَّرُ	تَفَعَّلَ	لازم		تصدَّر فلانٌ: جلس في صدر المجلس.
24860	ص د ر	اسْتَصْدَرَ يَسْتَصْدِرُ	اسْتَفْعَلَ	متعد		استصدر الأمر: طلب إصداره.

1- سورة الزلزلة ، الآية 6.

2- سورة القصص ، الآية 23.

يبيّن هذا المثال أنّ للجذر (ص د ر) الأفعال التالية[1] :

- الفعل الثلاثي المجرد (صَدَرَ يَصْدُرُ) : ورد لازما في ثلاث معان ومتعديا بـــ (عن) في مثالين ، وبـــ (إلى) في معنى واحد ، ومتعديا بنفسه في معنى واحد ، ومبنيا للمجهول في معنى واحد.
- الفعل الثلاثي المزيد (أَصْدَرَ يُصْدِرُ) ورد متعديا بنفسه في ثلاث معان ومتعديا بنفسه وبحرف في معنى واحد.
- الفعل الثلاثي المزيد (صَادَرَ يُصَادِرُ) ، ورد متعديا بنفسه في مثالين ومتعديا بنفسه وبحرف في معنى واحد.
- الفعل الثلاثي المزيد (صَدَّرَ يُصَدِّرُ) ، ورد متعديا بنفسه في ثلاث معان.
- الفعل الثلاثي المزيد (تَصَدَّرَ يَتَصَدَّرُ) ، ورد لازما في معنيين.
- الفعل الثلاثي المزيد (اسْتَصْدَرَ يَسْتَصْدِرُ) ، ورد متعديا بنفسه في معنى واحد.
- يرتبط بجدول الأفعال عشرة جداول فرعية هي : جدول المصادر، وجدول أمثلة الأفعال ، وجدول الصفات المشبهة ، وجدول مبالغة اسم الفاعل، وجدول الأضداد، وجدول الأسماء المصاحبة للأفعال ، وجدول التراكيب اللغوية والعبارات الاصطلاحية، وجدول الفوائد اللغوية والنحوية والصرفية كما هو واضح في الشكل الآتي[2]:



الجدول رقم 02: جدول الأفعال.

1- أ. مروان البواب ، منهج إعداد المعجم الحاسوبي للغة العربية ، ص 13.

2- المرجع السابق ، ص 13.

وفيما يلي سنقدّم شرحا مختصرا لهذه الجداول وتبيان أهميتها في بناء قاعدة المعطيات.

1- **جدول المصادر** :

يتضمن هذه الجدول مصادر الأفعال الثلاثية المجردة الّتي وردت في جدول الأفعال، ويرتبط كل مصدر بفعله اعتمادا على رقم مفتاح الفعل.

يتألف جدول المصادر من ثلاثة أعمدة [1]:

- العمود الأوّل : رقم مفتاح الفعل ، وهو نفسه الذي ورد في "جدول الأفعال".
- العمود الثاني : مصدر الفعل (يجب اختيار من المصادر أشيعها وأكثرها دورانا).
- العمود الثالث : وزن الفعل.

ويمكن توضيح ذلك من خلال الجدول الآتي [2] :

رقم مفتاح الفعل	المصدر	وزن الفعل
24842 24843 24844 24845	صدْر	فَعْل
24842 24843 24844 24845	صُدُور	فُعُول
24846 24847 24848	صدْر	فَعْل
24846 24847 24848	صَدَر	فَعَل

الجدول رقم 03: بعض مصادر الفعل الثلاثي المجرد (صَدَرَ يَصْدُرُ).

يتبيّن من هذا المثال أنّ مصادر الفعل الثلاثي المجرد (صَدَرَ يَصْدُرُ) هي :

- (صَدْرٌ) و(صُدُورٌ) للمعاني الأربعة الأولى.
- (صَدْرٌ) و(صَدَرَ) للمعاني الثلاثة الثانية.

أمّا مصادر الأفعال الثلاثية المزيدة ، والأفعال الرباعية (المجردة والمزيدة) فهي مصادر قياسية لا توردها معظم المعجمات[1].

1- المرجع السابق ، ص 14.

2- المرجع نفسه ، ص 15.

1- باعتبار نظام الاشتقاق والتصريف المدمج في هذا المعجم يولّد هذه المصادر اعتمادا على قوانين توليدها ، ويعرضها في مكانها المخصص.

2- **جدول أمثلة الأفعال** :

يتضمن هذا الجدول أمثلة الأفعال الّتي وردت في "جدول الأفعال" ، ويرتبط كل مثال بفعله اعتمادا على رقم مفتاح الفعل[1].

يتألف "جدول أمثلة الأفعال" من عمودين :

- **العمود الأوّل** : "رقم مفتاح الفعل" (وهو نفسه الذي ورد في جدول الأفعال).
- **العمود الثاني** : المثال : (يُكتب الفعل الذي ه موضع المثال بخط واضح).

ينفرد كل مثال من أمثلة الفعل بسطر مستقل.

ويمكن توضيح ذلك بالجدول الآتي[2] :

رقم مفتاح الفعل	المثـــــال
4495	﴿ كُلُواْ مِن ثَمَرِهِ إِذَا أَثْمَرَ وَآتُواْ حَقَّهُ يَوْمَ حَصَادِهِ ﴾[3]
4495	فمرّة يتشبه بالبدر إذا أقمر ، وتارة يتمثل بالغصن إذا أثمر.
4495	قال أنس : دفنت من صلبي مئة ولد ، وإن نخلي يُثمر في السنة مرّتين.

الجدول رقم 04: أمثلةعن الفعل الثلاثي المزيد أَثْمَرَ.

3- **جدول الصفات المشبهة** :

يتضمن الجدول الصفات المشبهة للأفعال الّتي وردت في جدول الأفعال وترتبط كل صفة مشبهة بفعلها اعتمادا على رقم مفتاح الفعل.

يتألف جدول الصفات المشبهة من ستّة أعمدة[4] :

- **العمود الأوّل** : رقم مفتاح الفعل (وهو نفسه الذي ورد في جدول الأفعال).
- **العمود الثاني** : الصفة المشبهة للمذكر المفرد.
- **العمود الثالث** : وزن الصفة المشبهة.
- **العمود الرابع** : الصفة المشبهة للمؤنثة المفردة.
- **العمود الخامس** : الصفة المشبهة لجماعة الذكور.
- **العمود السادس** : الصفة المشبهة لجماعة الإناث.

1- يقصد هنا "برقم مفتاح" بالكود "un code".

2- للمزيد من التفصيل ينظر : "منهج إعداد المعجم الحاسوبي للغة العربية" ، ص 16.

3- سورة الأنعام ، الآية 141.

4- ينظر : منهج إعداد المعجم الحاسوبي للغة العربية ، ص 17.

تنفرد كل صفة من الصفات الشبهة بسطر مستقل.

وسنوضح ذلك بالجدول التالي :

رقم مفتاح الفعل	الصفة المشبهة	وزنها	مؤنثها	جمع المذكر	جمع المؤنث
7954	حَزِنٌ	فعِل	حَزِنَةٌ	حزنون وحزان وحزناء	حزنات وحزان وحزناء
7954	حَزِينٌ	فعِيْل	حَزِينَةٌ	حزينون وحزّان وحزناء	حزينات وحزان وحزناء
7954	حَزْنَانٌ	فَعْلان	حَزْنَانَةٌ	حزنانون وحزانى	حزنانات وحزانى

الجدول رقم 05: أمثلة الصفات المشبهة.

ملاحظة :

وطلبا للإيجاز[1] قدمت في هذا العرض لهذا العمل التطبيقي أهمّ الجزئيات الأساسية الّتي اعتمدت عليها قاعدة المعطيات الّتي أنجزت من أجل إعداد معجم إلكتروني ، وللمزيد من التفصيل في هذا العمل يمكن الرجوع إلى العمل الذي قدّمه الأستاذ مروان البواب من خلال تقديم تصور منهجي للمعجم الحاسوبي للغة العربية.

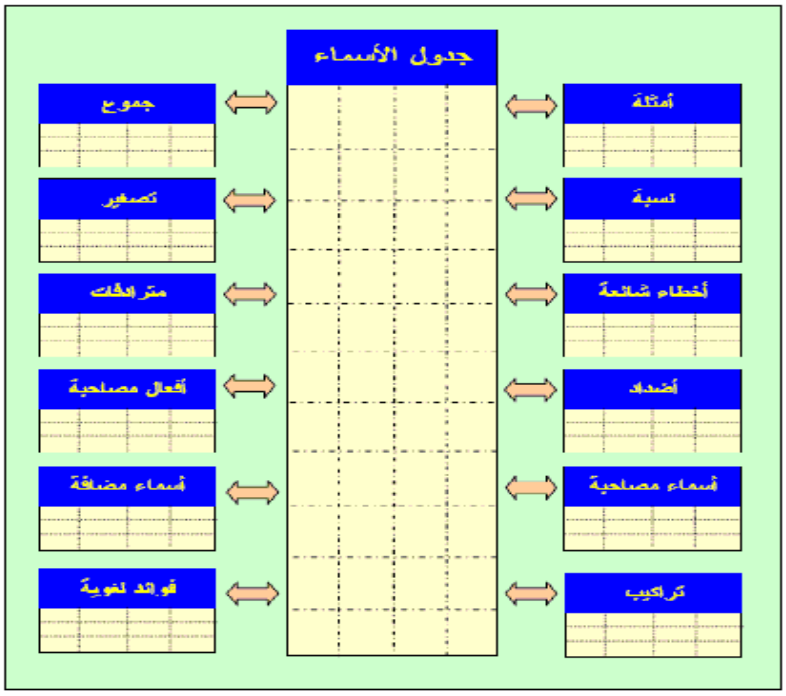
4- **جدول الأسماء** :

بخلاف جدول الأفعال فإنّ جدول الأسماء يرتبط باثني عشر جدول فرعيا هي كالآتي[2]: جدول أمثلة الأسماء ، جدول جموع الأسماء ، جدول النسبة، جدول التصغير، جدول الأخطاء الشائعة في الأسماء ، جدول المترادفات والمجالات الدلالية المتعلقة بالأسماء، جدول الأضداد المتعلقة بالأسماء، جدول الأفعال المصاحبة للأسماء، جدول الأسماء المصاحبة للصفات، جدول الأسماء المضافة إلى الأسماء، جدول التراكيب اللغوية والعبارات الاصطلاحية المتعلقة بالأسماء، جدول الفوائد اللغوية والنحوية والصرفية.

ويمكن توضيح ذلك في الشكل الآتي :

1- نفس العمل في إعداد قاعدة البيانات بالنسبة لكل من الجداول الآتية : جدول مبالغة اسم الفاعل ، جدول الأخطاء الشائعة في الأفعال، جدول المترادفات الدلالية المتعلقة بالأفعال ، جدول الأضداد المتعلقة بالأفعال ، جدول الأسماء المصاحبة للأفعال ، جدول التراكيب اللغوية والعبارات الاصطلاحية المتعلقة بالأفعال ، جدول الفوائد اللغوية والنحوية والصرفية المتعلقة بالأفعال.

2- ينظر : منهج إعداد المعجم الحاسوبي للغة العربية ، ص 25.



الجدول رقم 06: جدول الأسماء.

ومن باب الإشارة والتوضيح فقط سنختصر على عرض مثال واحد من الجداول المذكورة في جدول الأسماء كمثال تطبيقي يمكن تطبيقه في قاعدة المعطيات للجداول المتبقية[1].

أ- **جدول أمثلة الأسماء** :

يتضمن هذا الجدول أمثلة الأسماء الّتي وردت في "جدول الأسماء" ويرتبط كل مثال بالاسم المتعلق به اعتمادا على "رقم مفتاح الاسم" ، وعليه : فجدول الأسماء يتكون من عمودين :

- **العمود الأوّل** : رقم مفتاح الاسم : (وهو نفسه الذي ورد في جدول الأسماء).
- **العمود الثاني** : المثال (يكتب الاسم الذي هو موضع المثال بخط غامق).

ينفرد كل مثال من أمثلة الاسم بسطر مستقل ، كما هو موضح في الجدول الآتي[2]:

رقم مفتاح الاسم	المثـــــال
10536	فلقد رأيته يدخل **الخيط** في الإبرة وهو ابن ثمانين.
10536	قد ذاب حتّى رأيت **الخيط** يشبهه...من الهزال وعين الضرّ تبكيه
10536	الكتب عقل شوارد الكلم والخط **خيط** فرائد الحكم

الجدول رقم 07: أمثلة الأسماء

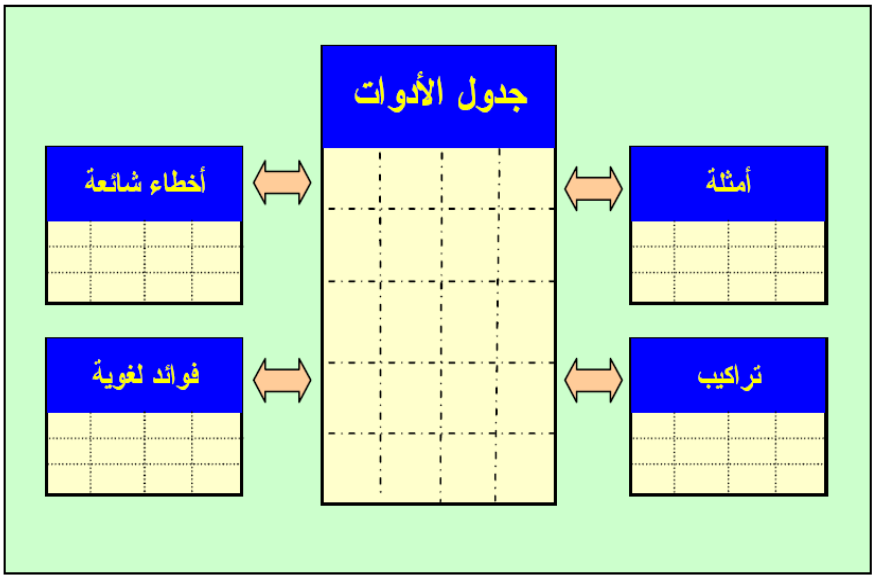
1- للمزيد من التفصيل ينظر : أ. مروان البواب ، منهج إعداد المعجم الحاسوبي للغة العربية، ص 26 وما بعدها.

2- المرجع السابق، ص 26.

ب- جدول الأدوات :

وأمّا جدول الأدوات يرتبط بأربعة جداول فرعية هي : جدول أمثلة الأدوات ، جدول الأخطاء الشائعة في الأدوات وجداول التراكيب اللغوية والعبارات الاصطلاحية ، وجدول الفوائد اللغوية والنحوية والصرفية.

كما هو موضح في الشكل الآتي [1]:



الجدول رقم 08: جدول الأدوات.

وسنقتصر في عرضنا هذا على جدول واحد كمثال توضيحي فقط.

1- جدول أمثلة الأدوات :

يتألف "جدول أمثلة الأدوات" من عمودين [2] :

- العمود الأوّل : رقم مفتاح الأداة : (وهو نفسه الذي ورد في جدول الأدوات).
- العمود الثاني : المثال : (تكتب الأداة الّتي هي موضع المثال بخط غامق).

وبنفس الطريقة الّتي أعهدت بها الجداول السابقة يفصل أو ينفرد كل مثال بسطر مستقل ، ويمكن توضيح ذلك بالجدول الآتي بالأمثلة على الأداة (مُنْذُ).

1- طلبا للإيجاز وتمكين القارئ أخذ ولو صورة عن كيفية إعداد قاعدة المعطيات للمعجم الإلكتروني قمت بعرض مثال تطبيقي يمكن الاستعانة به لإعداد بقية الجداول المتبقية.

وللمزيد من التفصيل في هذه الحالة ينظر : أ. مروان البواب ، (منهج إعداد المعجم الحاسوبي للغة العربية) ، ص 35 وما بعدها.

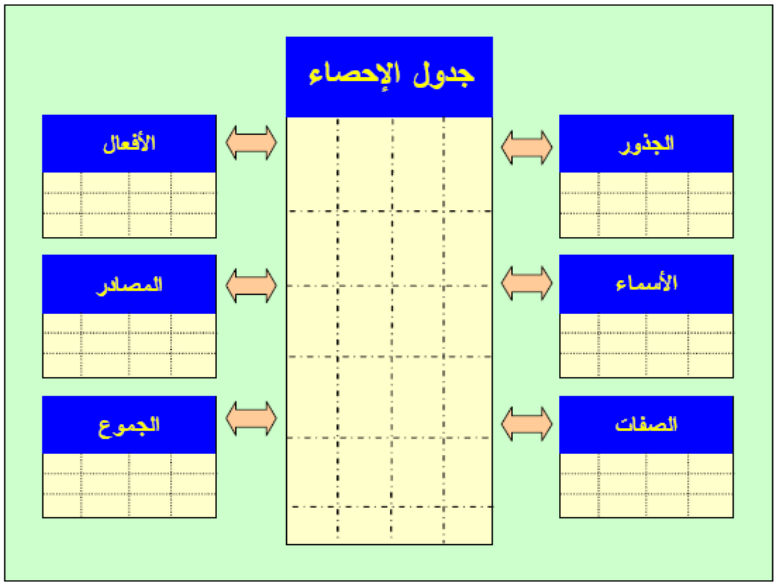
2- أ. مروان البواب ، منهج إعداد المعجم الحاسوبي للغة العربية، ص 35.

رقم مفتاح الاسم	المثـــــال
219	وأمّا النية والأمنية فأنا والله أتمنى هذا مُنْذُ أنا صبي (دخلت "مُنْذُ" على جملة اسمية.
219	باشرت الحروب مُنْذُ أنا غلام إلى الآن (دخلت " مُنْذُ" على جملة اسمية)
219	كم لك مُنْذُ نزلت بالعراق؟ (دخلت " مُنْذُ" على جملة فعلية).

الجدول رقم 09: أمثلة الأدوات.

2- جدول الإحصاء :

إنّ جدول الإحصاء يتضمن إحصائيات عن محتويات المعجم باستعمال برامج خاصة ، وهو يرتبط بجداول فرعية كجداول الأسماء والأفعال والأدوات ، إنّ البرامج المتخصصة في إحصاء المحتويات الخاصة بالجداول تقوم بعملية إحصائية لمكونات المعجم من جذور وأفعال ، وأسماء ، ومصادر ، وصفات ، يمكن وضعا بعد ذلك في جداول إحصائية خاصة ، يمكن من خلالها إجراء عملية البحث. وتكمن أهمية هذه الجداول الإحصائية في الاستفادة من خلال تسهيل عملية البحث بالنسبة للعاملين في حقل اللغة العربية ولاسيما العاملين في ميدان المعالجة الآلية للغة العربية[1]. يرتبط بجدول الإحصاء ستّة جداول فرعية إحصائية هي : جدول إحصاء الجذور، جدول إحصاء الأفعال، جدول إحصاء الأسماء، جدول إحصاء المصادر ، جدول إحصاء الصفات، جدول إحصاء الجموع ، كما هو موضح في الشكل الآتي[2] :



الجدول رقم 10: جدول الإحصاء.

1- المرجع السابق، ص 38.

2- المرجع نفسه، ص 38.

سنقتصر في هذه العينة من الدراسة على شرح مختصر لجدول إحصاء الجذور[1].

3- جدول إحصاء الجذور :

وتشتمل على عدد من الإحصائيات من ذلك مثلا : إحصائيات الجذور[2].

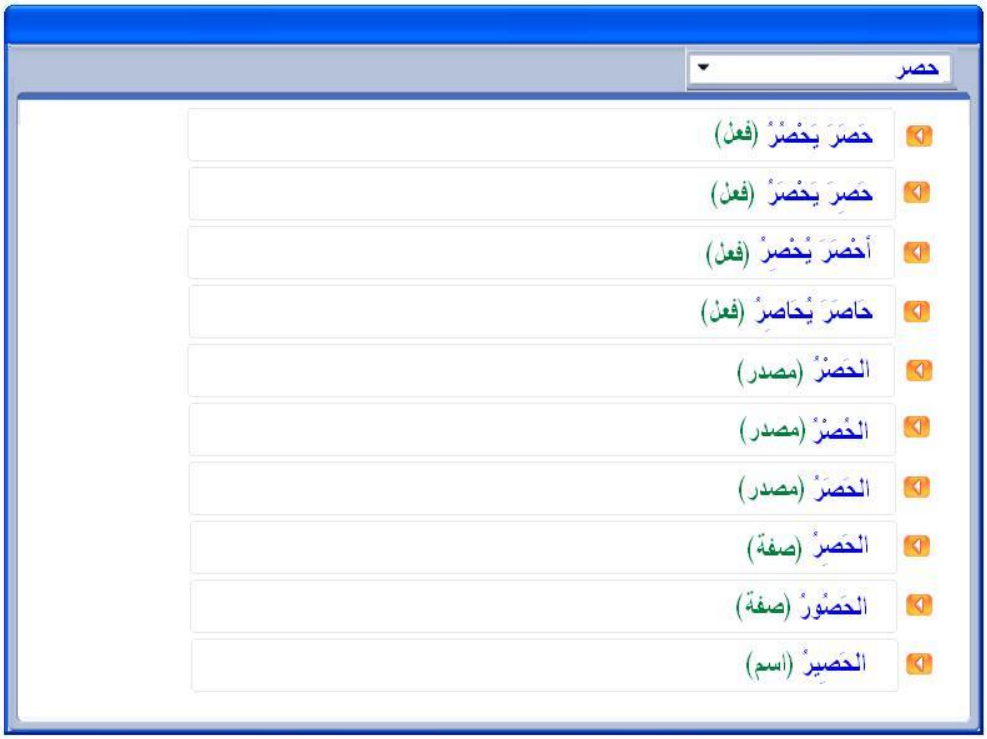
1. بحسب عدد الحروف : الجذور الثنائية والثلاثية والرباعية ، والخماسية.

2. بحسب مواقع الحروف : الجذور الّتي تبدأ بحرف معين ، والتي تنتهي بحرف معين ، والتي تتضمن حرفا معين في موقع محدّد.

3. بحسب تتابع الحروف : الجذور الّتي تتضمن حرفين معينين في موقعين مختلفين (الأوّل والثاني من الجذر ، والثاني والثالث.. والأوّل والثالث...والأوّل والأخير...).

طريقة الكشف في المعجم الحاسوبي:

أ/ عن طريق الجذر :



الشاشة رقم: 01 – توضح عمليات البحث في المعجم الحاسوبي :

1- للمزيد من التفصيل في هذه الحالة التطبيقية ينظر : منهج إعداد المعجم الحاسوبي للغة العربية لمروان البواب ، ص 39 ، 40.

2- المرجع نفسه، ص 39.

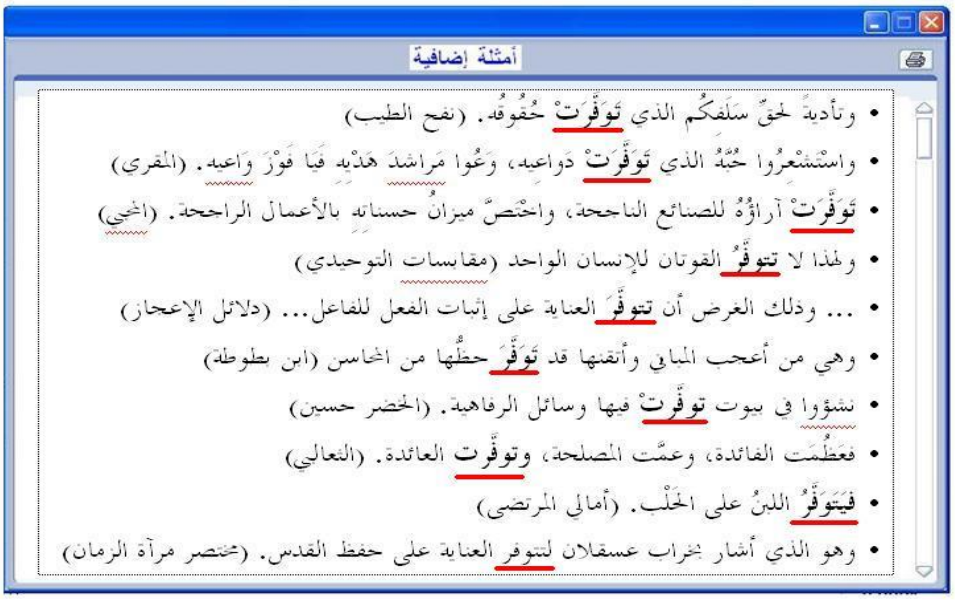
الشاشة رقم : 02 - نظام الإشتقاق والتصريف في المعجم الحاسوبي

الشاشة رقم 03 و 04 توضح عملية التصريف في المعجم الحاسوبي :

الشاشة رقم : 03 توضح أمثلة عن تصريف الفعل

الشاشة رقم : 04 ـ توضح أمثلة عن تصريف الفعل

أمثلة إضافية:



الشاشة رقم: 05

نستخلص من هذا العرض سريع للمعجم الحاسوبي للغة العربية المكلف بإعداده الأستاذ البواب يجب ان يكون شاملاً ووافياً كما ينبغي ان تكون الجهود المبذولة لاعداده موجهة نحو الجانب اللغوي فيه, فالتقدم الهائل في التقانات كفيل بتسهيل العمل وتيسيره لتحقيق معجم حاسوبي متطور ، إذا ما توفرت المادة اللغوية الدقيقة والمتكاملة ثم إن حسم اوجه الخلاف بين المعاجم من أهم النقاط التي يجب ان تراعى عند إعداد أي معجم حاسوبي, لأن طبيعته تفرض على الباحث اللساني أن يدخل فيه معطيات كاملة التوصيف من جهة ودقيقة التحديد من جهة اخرى وأن يسع المعجم جميع مواد العربية ويضمن كل ما دخلها واكتسب خصائصها ووزن بأوزانها.

5- تجربة الدكتور نبيل علي:

يعتبر نبيل علي أول مهندس لغوي عربي، قام بأبحاث رائدة في العالم العربي فيما يتصل بالمعالجة الحاسوبية للمنظومة اللغوية العربية، وقد أغنى الخزانة العربية بكتب قيمة في اللسانيات الحاسوبية وثقافة المعلومات، وفي مقدمتها كتابيه "اللغة العربية والحاسوب"، و"الثقافة العربية في عصر المعلومات"، بالإضافة إلى العديد من المقالات والبحوث المهمة التي قدمها في مؤتمرات وندوات علمية عربياً ودولياً.

إن ميكنة المعجم في نظر نبيل علي لا بد أن تخضع لثنائية التحليل والتركيب[1]، شأنها شأن المكونات اللغوية الأخرى. والمعجم الآلي بالنسبة إليه يشمل المعطيات الصرفية والنحوية والدلالية. كما أنه يعتبر مصدراً أساساً لا غنى عنه في دراسة الإنتاجية الصرفية، وعلاقة الترابط للعناصر المعجمية المختلفة.

وهكذا فقد ارتأى أن تنظم ميكنة المعجم في أربع مستويات مترابطة ومتكاملة فيما بينها[2]:

أولا- قائمة مفردات اللغة. **ثانيا**- مصفوفات علاقات نحوية وصرفية ودلالية.

ثالثا- قاعدة بيانات. **رابعا**- قاعدة معرفة متكاملة.

6- تجربة الحاج صالح: مشروع الذخيرة اللغوية العربية [3]:

يعتبرعبد الرحمن الحاج صالح[4] من الباحثين الرواد الذين أسهموا ببحوث ودراسات هامة في مجال اللسانيات بمجاليها النظري والتطبيقي، حيث لها الأثر البالغ الأهمية في تأصيل المعالجة الآلية للغة العربية من خلال إعداد تصور حول "أنموذج لساني للعلاج الآلي للغة العربية"[5].

1- نبيل علي، اللغة العربية والحاسوب، مرجع سابق، ص 17.

2- للمزيد من بسط القول في هذا المبحث ينظر : ميكنة المعجم العربي، نبيل علي، اللغة العربية والحاسوب، مرجع سابق، ص 457 وما بعدها.

3- للمزيد من التفصيل ينظر : عبد الرحمن الحاج صالح، الذخيرة اللغوية العربية، اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد 27، الرباط، المغرب، 1987/1986، ص 45 وما بعدها.

4- ولد الدّكتور عبد الرّحمن حاج صالح في مدينة وهران بالجزائر سنة 1927م/ ت 2017 م، وبعد أن أتمّ تعليمه المدرسيّ ، بدأ في دارسة الطب وفي سنة 1954 توجه إلى مصر ليكمل دراسة التخصص في جراحة الأعصاب ، ولما كان يتردد على جامع الأزهر وكان يحضر إلى بعض دروس اللغة العربية ، وإذا به يجد نفسه يعيد اكتشاف ذاته من جديد ويتعرف على تراث اللغة العربية بوعي جديد؛ فحوّل اهتمامه من حقل الطب إلى الدراسات اللغوية المعاصرة ، وبعد الاستقلال أكمل دراسته الجامعية وفي سنة 1968 كان أستاذا زائرا بجامعة فلوريدا حيث التقى بالعالم اللساني آنذاك نعوم تشومسكي، حاصل على جائزة الملك فيصل للّغة العربيّة والأدب سنة 2010 تقديرا لجهوده العلمية المتميزة في تحليله النظرية الخليلية وعلاقتها بالدّراسات اللسانية المعاصرة، ودفاعه عن أصالة النحو العربي، وإجرائه مقارنات علمية بين التراث ومختلف النظريات في هذا الموضوع، بالإضافة إلى مشاركاته في الدراسات اللسانية بحثا وتقويما وتعليما، وجهوده البارزة في حركة التعريب.

وكان الحاج صالح رئيس المجمع الجزائري للغة العربية وصاحب مشروع (غوغل العربي)، أو البنك الآلي العربي قد انتخب على رأس مؤسسة الذخيرة العربية، واختيرت الجزائر أن تكون المقر الرسمي لهذه المؤسسة.

5- من المعروف أن الجانب التطبيقي – وهو الجانب الأهم في اللغويات الحاسوبية – يتمثل في تسخير العقل الإلكتروني لحل القضايا اللغوية، وهنا يبرز الدور الرئيس والأثر الفاعل لالتقاء اللغويين والحاسوبيين، والتعاون فيما بينهم، وما يثمر عنه من نتائج تسهم إلى حد كبير في تذليل العقبات وحل المشكلات التي تواجه التحليل الحاسوبي للغة العربية، هذه العقبات والمشكلات بعضها يتصل بطبيعة اللغة العربية، أصواتاً، وبنيةً، وتركيباً، ودلالة، ومعجماً، وبعضها يتعلق بنظام الكتابة العربية، وبعضها يتصل بالمصطلح العلمي التكنولوجي للسانيات العربية، كما أن هناك مشكلات أخرى تتعلق بالبرمجيات، إعداداً، واختياراً للمادة اللغوية العربية (أنموذج لساني عربي)، وتعريباً للبرمجة. وثالث هذه المشاكل يكمن في الجهاز الحاسوبي (الكمبيوتر)، وأنظمة تمثيل المعرفة على الحاسوب باللغة العربية.

الذخيرة اللغوية العربية[1] عبارة عن بنك آلي أو محوسب من النصوص القديمة والحديثة، يشتمل على التراث العربي والعالمي، وعلى الإنتاج العلمي العربي والعالمي[2].

وتهدف الذخيرة إلى تحقيق الأهداف التالية:

- ✓ أن تكون قاعدة معطيات نصية (أي مدونة لغوية) يحرر على أساسه عدد من المعاجم وخاصة "المعجم الجامع للغة العربية".
- ✓ أن تكون مصدراً موثوقاً من المعلومات العلمية والتكنولوجية والأدبية واللغوية والاجتماعية والثقافية عامة، موجهة للباحثين والمهتمين على حد سواء.
- ✓ أن تستخدم كأداة تعليمية وتربوية في تعليم اللغة العربية خاصة، واللغات الأجنبية عامة.

أما فوائدها فتتحدد في النواحي الآتية:

– الناحية اللغوية والمعجمية، فإذا كانت أكثر المعاجم العربية المؤلفة اليوم تركز أساساً على المعاجم القديمة، فإن الذخيرة باعتبارها مجموعة نصوص، تهتم باللغة المستعملة من أقدم العصور إلى اليوم.

– الناحية التربوية والثقافية: تعتبر الذخيرة مصدراً موثوقاً للمعلومات التي تستجيب لحاجيات وطلبات التلميذ والطالب الجامعي والباحث بصفة عامة.

إنها تمكن من تعلم مهارات في اللغة كتعلم العربية أو الفرنسية أو الإنجليزية.

1– الذخيرة العربية مشروع أمضى فيه عالم اللسانيات الجزائري، عبد الرحمان الحاج صالح، جهدا كبيرا من وقته وفكره من أجل أن يراه حيا يزود هذه الأمة بجميع أنواع المعارف والعلوم والفتوحات العلمية التي تتحقق على المستوى العالمي، إضافة إلى الانجازات العلمية والفكرية التي أنجزتها الحضارة العربية الإسلامية على مدار 14قرنا من الزمن .

2– يقول الدكتور الحاج عبد الرحمان صالح في تعريفه بمشروع الذخيرة أوالإنترنيت العربي: "نحن نريد أن نبحث عن المعلومات العلمية والتكنولوجية والدينية، المعلومات التي لها قيمة، والتي يصعب على الناس الوصول إليها في الحين، الذخيرة موجهة من الصبي إلى العالم، إلى السياسي إلى كل طبقات الأمة"، ويسجل سبب طرحه لفكرة هذا المشروع قائلا" :لاحظت أن كتابا واحدا ترجم وصار هو المرجع الوحيد للسانيين العرب مدة 40 سنة"، ويرى العلامة عبد الرحمان الحاج صالح أن تأخر العرب هو عدم وصول المعلومات لكافة الشعب لأن في الشعب أشخاص لهم مواهب عظيمة وليس لهم الإمكانيات لتنمية هذا الذكاء فلشمولية تعميم المعرفة يرتقي الشعب. أما الفائدة التي نحصل عليها من الذخيرة –حسبه– فهي متعددة منها: أن المعاجم الحالية التي تخص اللغة العربية والمحررة في البلدان العربية تعتمد أساسا على المعاجم القديمة، أما الذخيرة، كمجموعة نصوص، فميزتها الأساسية أنها تخص اللغة المستعملة بالفعل من أقدم العصور إلى زمننا هذا، ومن جهة أخرى سيكون تحت تصرف أي باحث كل المعطيات التي تهمه ويمكن أن يتحصل في الحين على كل الألفاظ التي يطلبها في قائمة شاملة تغطي كل المجال المفهومي الذي يختاره، أما أنواع الأعمال فقد تم تحديدها في حيازة نصوص في ذاكرة الحاسوب وفي موقع الإنترنيت، وترجمة نصوص علمية إلى العربية، وكذا ترجمة بعض الكتب العلمية أوعدة موسوعات متخصصة صغيرة الحجم، النصوص التي ستدخل في الذخيرة، نصوص من التراث العربي ومن الإنتاج العلمي والأدبي العربي الحديث ومن الحياة العامة، بالإضافة إلى نصوص للتعريف بأحدث التطورات العلمية، كما تشمل الذخيرة العصر القديم والعصر الحديث ونصوص دراسية وموسوعية كالمعاجم وهذه تنضوي تحت لجان، كلجان التنسيق والمتابعة لجنة التراث، لجنة الترجمة للبحوث العلمية، لجنة الحاسوبيات، لجنة النصوص الإعلامية. مشروع الذخيرة العربية هاهو ينتقل من التفكير إلى التجسيد والانجاز بعد أن أصبح مؤسسة قائمة بذاتها من مؤسسات الجامعة العربية.

7- تجربة سلوى حمادة : نحو إنشاء قاعدة بيانات معجمية آلية في شكل معجم كوني هرمي (مكة): معجم عصر المعلوماتية" :

خصصت الدكتورة سلوى حمادة[1] فصلا كاملا من كتابها[2] للحديث عن المعجم الآلي تحت عنوان "نحو إنشاء قاعدة بيانات معجمية آلية في شكل معجم كوني هرمي (مكة): معجم عصر المعلوماتية"، حيث تحدثت المؤلفة عن أهمية الوعي المعجمي عند أبناء العربية، وشرحت أسباب ضرورة وجود المعاجم الآلية مبينة صعوبة البحث في أوراق المعجم مع طول الوقت المطلوب، كما تطرقت إلى أهم مشكلات المعاجم الحالية، وإلى متطلبات معجم عصر المعلوماتية و محتوياته، وخطوات إنشاء قاعدة البيانات المعجمية الآلية، كما بينت الجديد في معجم عصر المعلوماتية، وما يمكن أن يقدمه لحل القضايا اللغوية، ويعد العمل في حد ذاته خطوة ناجحة في سبيل حوسبة المعجم العربي من جهة، وفيما أوردته بخصوص المعاجم الآلية من جهة أخرى بالرغم من اختلاف المعجم المزمع بنائه من حيث تفاعليته.

8- تجربة عبد المجيد بن حمادو :

تطرق المجيد بن حمادو[3] في تجربته حول **المعجم العربي الإلكتروني من أهميته وطرق بنائه** إلى جوانب عديدة تميز المعجم الإلكتروني عن المعجم الورقي من حيث الهيكلة[4] والمحتوى وطرق الاستعمال، وبين بالخصوص أهمية المعجم الإلكتروني العربي بالنسبة إلى مجتمعنا الذي يتحول بسرعة إلى مجتمع المعرفة؛ حيث تلعب فيه التكنولوجيا دوراً محورياً، كما تحدث عن كون المعجم الإلكتروني أداة عمل لا يمكن الاستغناء عنه بالنسبة إلى الكثير من الفئات الاجتماعية والمهنية، وهو كذلك حجر الزاوية للعديد، إن لم نقل لجل، تطبيقات المعالجة الآلية للغة. هذه التطبيقات تساهم بصفة جدية في إشعاع اللغة العربية واستمرارها كلغة ناقلة. كذلك تطرق إلى مختلف طرق بناء المعجم العربي[5]، وختم بموضوع تقييس المعاجم الإلكترونية لجعلها تخضع إلى مواصفات المعاجم الحديثة من خلال تجربة مخبر ميراكل. كما أكد على أن إتاحة مثل هذه المعاجم للغة العربية سيدعم حتما دورها في مجتمعنا الذي يعيش منافسة شرسة مع لغات أجنبية أهمها الإنجليزية والفرنسية. فمسؤوليتنا اليوم أعظم مما كانت عليه بالأمس لأن للغة أبعاداً ثقافية، واقتصادية واستراتيجية.

1- الدكتورة سلوى حمادة، خبيرة المعلوماتية واللغويات حائزة على جائزة الكندي للمعلوماتية (المرتبة الأولى) للعام 2008 في مجال أفضل باحث عربي معلوماتي، عن مجمل أعمالها العلمية والأكاديمية.

2- للمزيد من بسط القول ينظر: سلوى حماده، المعالجة الآلية للغة العربية، المشاكل والحلول، دار غريبللطباعة والنشر والتوزيع، الطبعة الأولى،القاهرة،2009، ص 57،17.

3- الأستاذ الدكتور عبد المجيد بن حمادو، جامعة صفاقس،تونس.

4- للمزيد من التفصيل في هذه الدراسة البحثية ينظر الباب الثاني من الأطروحة، الفصل الثاني، ص ص226،...ص241.

5- أشارالدكتور عبد المجيد بن حمادو إلى واقع العمل في هذا المعجم حيث انتهوا – في جامعة صفاقس التونسية- من إعداد قاعدة البيانات المعجمية والنظام الذي يمكّن المعجمي من إدخال كافة متطلبات المعجم، وقد شدد على أن هذا العملٌ مضنٍ يحتاج لإدخال بشريّ ضخم، والدعوة مفتوحة لكل من لديه مادة معجمية صحيحة متكاملة شاملة مرتبة مدروسة وواضحة المعالم لإدخالها وفقً قاعدة البيانات المعدَّة سابقاً.

9 - تجربة مراد لوكام لإنجاز مشروع المعجم الحاسوبي التفاعلي للغة العربية:

ومن المقالات العلمية في هذا المجال ،ما قدمه مراد لوكام[1] مقالا قيما بعنوان **مشروع المعجم الحاسوبي التفاعلي للغة العربية:** مقترحات حول إعداد المشروع عرض في هذه الوثيقة بعض أعماله السابقة و الحالية في مجال حوسبة اللغة العربية التي يمكن الإفادة منها لإنجاز مشروع معجم اللغة العربية. نذكر منها على الخصوص: المجال الصرفي ونظام التجذير / الاشتقاق، هذا وقد بيَّن المؤلف الأهداف المرجوة من وضع المعجم الحاسوبي، وهو استعماله في تطبيقات مختلفة للغة العربية يجعل من تقييس المادة المعجمية شيئا بالغ الأهمية، واقترح أن يقتبس من النظرية المعرفية اللسانية HPSG طريقتها في تمثيل المعارف. كما لفت الانتباه إلى بعض الجوانب التي قد تكون لها آثار سلبية على خطة إعداد المشروع، والتي يمكن حصرها في نقطتين: وجود عوائق لتشكيل مدونة متعددة الوسائط بالمعجم، وعدم مناسبة المدة المقترحة للإنجاز اللغوي. حيث ستركز الباحثة على التقليل من وجود تلك العوائق.

10- تجربة معجم حاسوبي أحادي للناطقين بغير العربية للدكتور وليد أحمد العناتي[2]:

1.10- توصيف المعجم المقترح:

أ. أهداف المعجم:

إن الغاية الرئيسية لهذا المعجم المقترح غاية تعليمية خالصة[3]، وتتبع هذه الغاية من بنية هذا المعجم وطبيعة الفئة المستهدفة[4] منه. وبيان ذلك أنه ليس معجم مفردات حسب، إنه معجم لغوي موسوعي يقصد في منتهى غايته إلى أن يهيئ لمتعلمي العربية من الناطقين بغيرها كفاية

1- الأستاذ مراد لوكام أستاذ بجامعة حسيبة بن بولعيد الشلف ، الجزائر.

2-.أستاذ اللسانيات المشارك في قسم اللغة العربية بجامعة البترا الأردنية الخاصة.

3- معلوم أن الغاية الرئيسية من بناء المعاجم هي الغاية التعليمية، وإن تفاوتت هذه المعاجم في مجال اهتمامها وأهدافها الرئيسية.فالمعاجم الثنائية تقصد إلى إعانة متعلمي اللغات الأجنبية على فهم مفردات اللغة التي يتعلمونها وتبين سلوكها النحوي والصرفي واستعمالاتها السياقية والإصلاحية المختلفة بما ينتهي إلى تمكن المتعلم من فهم قدر من المفردات والمعلومات والقواعد يعينه على التواصل باللغة الأجنبية وقد تجاوزت المعاجم الثنائية هذا الهدف إلى غايات الترجمة التطبيقية ثم تطبيقات الترجمة الآلية. وأما المعاجم أحادية اللغة فإنها تقصد إلى تنمية ثروة المتعلم اللفظية وتعريفه بمفردات لغته ولا سيما كان قديما منها ليتمكن من التعامل مع نصوص لغته كلها في امتدادها الزمني المتعاقب.

4- صُمِّم هذا المعجم للناطقين بغير العربية من الجنسيات كلها، وهو موجه إلى جميع مستويات المتعلمين المبتدئين إلى المتقدمين وقد يتساءل بعض الناس: كيف يمكن أن يساوي المعجم بين مستويات المتعلمين جميعها؟ ونقول إن المادة إلتي يتضمنها المعجم في مستوياتها المتعددة تتدرج من الجذر الأصلي إلى كل ما يتعلق هذا الجذر من بناء صوتيومشتقات صرفية، وأعاريب نحوية، وتعبيرات اصطلاحية، ومدونات نصية متكاملة تتفاوت في أطوالها ...إلخ. وهذه المعلومات لا يستغني عنها أي متعلم أجنبي مهما علا قدر تحصيله من العربية، على أن تصميم المعجم انتهى إلى رسم طريقين للبحث في محتوياته:

الطريق الأول: وهو البحث المباشر في الصفحة الأولى وذلك بالنقر مباشرة على الحرف الذي يبدأ به الجذر الأصلي، ثم النقر على الجذر. وغالبا ما يخدم هذا الطريق القادرين على معرفة الجذر الأصلي.

الطريق الثاني: ويتمثل في إتاحة خيار البحث من البداية وذلك بكتابة الكلمة المراد معرفة معانيها أو أي معلومات أخرى، فينتهي البحث بتجريد الكلمة من الزوائد وتحديد الجذر الأصلي. ومع ذلك فإن من يمتلكون معارف تأسيسية في العربية أقدرعلى استعمال هذا المعجم والإبحار فيه والاستفادة منه، ثم إن هذا المعجم يقدم خدمة عظيمة لمعلمي العربية للناطقين بغيرها؛ ذلك أنه يجمع في مكان واحد كما هائلا من المعلومات والمعارف والمهارات التي تغنيه عن كثير من مصادر التعلم، ولا سيما مع الاستفادة من تقنيات الوسائط المتعددة. وتتعاظم فائدته عندما يُستَعمَل في بلد غير ناطق بالعربية.

لغوية تتجلى في أداء لغوي سليم يبين عن تمكن المتعلم من العربية في نظمها المتعددة: النظام الصوتي، والنظام الكتابي، والنظام الصرفي، والنظام النحوي إعرابا وتركيبا، والنظام الدلالي بعلاقته المتعددة، والنظام الأسلوبي، والنظام النصي واستراتيجيات الخطاب[1]، فالمعجم يسعى إلى تزويد المتعلم بكفاية ثقافية ومعرفية بالعربية يمثلها قدر مناسب من الثقافة العربية الإسلامية ومنجزاتها الحضارية المادية والمعنوية بما ينبئ عن ثقافة اللغة العربية ومحمولات مفرداتها الثقافية والحضارية على مر عصور الحضارة العربية الإسلامية[2].

ب. ترتيب المعجم:

تفاوتت الاجتهادات حول المنهج الأمثل لمعجم مخصص لغير الناطقين بالعربية؛ وقد عرض علي القاسمي لهذه المسألة بالتفصيل؛ إذ تناول المناهج المتعارفة في الترتيب المعجمي دالا على فوائد كل منهج ونقائصه فقد عرض الترتيب العشوائي والترتيب المعنوي(الموضوعي) والترتيب الألفبائي، فالدكتور العناتي يرى أن الترتيب في المعجم الحاسوبي ليس ضروريا جدا، لاعتماده على تقنية البحث الآلي فإن الترتيب الهجائي (الألفبائي) الجذري هو أفضل مناهج الترتيب لمعجم عربي للأجانب، ولعل أهم فوائد الترتيب تكون[3]:

1- إبراز العلاقة بين المفردات المنتمية إلى أسرة لفظية واحدة، فييسر عملية التعليم.

2- إن هذا الترتيب أكثر ملائمة لمعاجم اللغات السامية التي تميل إلى الاشتقاق وتتخذ منه وسيلة صرفية لزيادة المفردات أكثر من استخدامها للنحت كما هي الحال في اللغات الجرمانية.وتشترك مفردات الأسرة اللفظية الواحدة في المعنى العام للأصل ثم تتخذ كل لفظة منها معنى خاصا بها يضيفه عليها الوزن الصرفي الذي صيغت عليه. يؤدي الأخذ بالترتيب الجذري إلى الاقتصاد في حجم المعجم؛ ولذلك لعدم اضطرار المعجمي إلى إعادة تعريف كل لفظة مشتقة، لأن المشتقات جميعها تشترك في معنى عام.

3- يعمل هذا الترتيب على تفهيم المتعلمين الآلية الرئيسة في توليد الصيغ الصرفية ومعانيها الجديدة في اللغة العربية؛ إذ يدرك المتعلم مع كثرة استعمال المعجم كيفية الاشتقاق و طرائقه.

ويرى القاسمي أن لهذا الترتيب عيبا واحدا عندما نتبعه في تصنيف معجم للناطقين بغير العربية، ويتمثل هذا العيب في افتقار المتعلم غالبا إلى القدرة على تخمين جذر الكلمة التي

1- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، مجلة الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، يناير 2009، ص 231.

2- المرجع نفسه،ص 232.

3- علي القاسمي، المعجمية العربية بين النظرية والتطبيق، ص 119 ، 120، وأيضاً علي القاسمي: ترتيب المداخل في المعجم العربي، وقائع تعليم اللغة العربية لغير الناطقين بها، ج1، مكتب التربية العربي لدول الخليج، ص. 42 – 43. وقد نشر البحث نفسه في وقائع الدورة التدريبية لصناعة المعجم العربي لغير الناطقين بالعربية، الرباط، ص. 20 – 51. والنقاط الثلاث الأولى للقاسمي أما الرابعة فهي للباحث.

يروم البحث عنها في المعجم.فأنى له أن يعرف بأن "مضيفة" من "ضاف" و "صفة " من "وصف" ...وهكذا[1].

ج. لغة المعجم:

يراد لهذا المعجم الموسوعي أن يكون معجما أحادي اللغة بصورة كاملة[2]؛ إذ ستكون لغة المتن (المداخل/ الجذور) هي العربية،وكذلك ستكون لغة الشرح (التوضيح/ التفسير/ التعريف) هي العربية أيضاً[3]. وهكذا فان هذا المعجم ينطلق من الطريقة المباشرة في تعليم اللغات الأجنبية؛ أي استعمال اللغة الهدف، العربية هنا، دون استعمال اللغة الوسيطة: لغة المتعلم الأم أو أي لغة أجنبية أخرى تكون وسيط التفاهم بين المعلم والطالب كالإنجليزية مثلاً[4].

لذلك فانه "ينبغي أن يكون الأسلوب الذي تصاغ به المعلومات بسيطاً خالياً من التعقيد الذي يسبب عسر الفهم، فليس من المعقول أن يصاغ التعريف بأسلوب معقد أو أن تستخدم فيه مفردات أصعب من المدخل المعرَّف، وليس من المعقول أن يكون تركيب الشواهد والأمثلة التوضيحية ومفرداتها من الصعوبة بمكان بحيث تحتاج هي بدورها إلى شرح وإيضاح"[5].

ويرى الدكتور العناتي أن مما يندرج تحت سهولة أسلوب العرض والشرح[6] الطرائق التالية[7]:

- استخدام طريقة الشرح والتعريف المناسبة للكلمة للمفردة[8].
- ولعله يحسن أن تشرح المفردات بألفاظ وردت في المعجم، ولا ينبغي أن نشرحها بألفاظ جديدة غير موجودة في مداخل المعجم. ولا شك أن هذا الإجراء سيفضي إلى بناء ثروة لفظية

1- علي القاسمي، المعجمية العربية بين النظرية والتطبيق، سابق، ص، 120.

2- إن المعجم أحادي اللغة لغير الناطقين بالعربية يمثل تحديا كبيرا للمشتغلين في إعداده، فهو أصعب كثيراً من المعجم ثنائي اللغة. ومن هنا فانه ينبغي التعامل بدقة متناهية مع معاني المفردات وطرائق تقديمها.

3- لعله يكون تدبيراً سديداً أن نترجم المقدمات العامة إلى لغات عدة تمثل أكثر لغات الجنسيات إقبالاً على تعلم العربية، كالإنجليزية، والمالايوية، والبهاسا، والأردو، والفرنسية، ...الخ.

4- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، مجلة الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، يناير 2009، ص 236.

5- علي القاسمي، المعجمية العربية بين النظرية والتطبيق، سابق، ص 137.

6- لعل أفضل طريقة لتحقيق هذه السهولة والكفاية الدلالية والتفسيرية هي التركيز على شرح الجذر شرحا دقيقا دالاً؛ ذلك لأنه هو المدخل الرئيسي لعائلة الفعل كلها، ثم إن وضوح دلالة الفعل ومعناه الدقيق سيجعل من السهل تبين دلالات مشتقات الفعل وما يتصل به من معان أخرى.

7- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، مجلة الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، يناير 2009، ص 237.

8- إذا كانت الكلمة المراد شرحها تدل على معنى مشخص ظاهر استعملنا الصور الدقيقة، كأن نستعمل صورة كتاب أو سيارة أو قميص ...الخ. ومعلوم أن المعجم الحاسوبي المقترح، بما يتاح له من برامج معالجة مساندة، سيتجاوز المعجم الورقي كثيراً في الشرح والإبانة، كالألوان والحركة والصوت والوسائط المتعددة. وأحسب أنه يمكن لنا في كثير من الأفعال (الجذور) أن نستعمل الصور المتحركة أو المشاهد الحقيقية المصورة، فإذا أردنا الدلالة على معنى (يكتب) أظهرنا صورة يد وهي تكتب، وإذا أردنا أن نشرح معنى الفعل (يسبح) قدمنا مشهدا حيا يمثل فعل السباحة. و هكذا.

نامية ومتطورة، ولعله يسهم في تحقيق غاية بعيدة: أن يفكر المتعلم الأجنبي بالعربية لفهم العربية.

2.10- البرامـــج المسانـــدة:

إن المعجم المقترح سيعتمد على **البرامج المساندة**[1] التالية[2]:

1- البرامج الصوتية. وينبغي أن أشير إلى أن البرامج الصوتية التي ستستعمل هنا إنما هي برامج معالجة الصوت وتحسين خصائصه، وليست برامج القراءة الآلية أو استنطاق النصوص، ذلك أن البرامج الصوتية في العربية لا يمكن أن تخدم الهدف التعليمي الذي بُني المعجم لإنجازه.

2- المدقق الإملائي.

3- المحللات الصرفية، بشقيها التحليلي و التركيبي.

4- المحلل النحوي والتركيبي، ويشمل: التدقيق النحوي، والتشكيل الآلي، والإعراب الآلي.

5- البرامج الفنية، كالألوان والخطوط والإضاءة والمعينات البصرية والصوتية المتعددة.

6- محرك بحث متطور ودقيق.

1- من أهم وجوه افتراق هذا المعجم عن المعجم الورقي أنه سيستثمر كثيرا من منجزات اللسانيات الحاسوبية العربية وما طورته من برامج تطبيقية في أنظمة العربية كلها. ولا شك أن هذا الفرق يترتب عليه فرق آخر ينتهي إلى ضخامة سعة تخزين هذا المعجم مقارنة بالمعجم الورقي، ذلك أننا هنا لن نخزن المعلومات الصرفية والنحوية والأسلوبية؛ وإنما نُحَصِّلُها بتفعيل هذه البرامج لنحصل على ما نريد بسرعة كبيرة. للمزيد من بسط القول ينظر: وليد أحمد العناتي وخالد الجبر، دليل الباحث إلى اللسانيات الحاسوبية العربية، ط1، دار جرير، عمان، 2007. ويمكن تحصيل معلومات تفصيلية أيضاً على مواقع شركات حوسبة النص العربي ومعالجته، مثل: شركة صخر، ودار حوسبة النص العربي، ودار التراث ...الخ

2- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، مجلة الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، يناير 2009، ص 238.

3.10- المدونـــة اللغويـــة:

ويقصد بالمدونة اللغوية جميع المصادر اللغوية[1] والثقافية التي يستعين بها المعجمي لبناء هذا المعجم، وهي تشمل المعاجم والموسوعات الثقافية والصحف اليومية الراقية، ومجموع المقالات المتميزة المنشورة في مجلات ثقافية أو متخصصة، أو تلك المجموعة في كتاب، ومصادر الأدب العربي على تفاوت عصورها[2].

1- ويمكن أن تضم هذه المدونة المصادر التالية (مثلاً):

أولا: المعاجم: - لسان العرب، لابن منظور- تاج العروس، للزبيدي - أساس البلاغة، للزمخشري- المنجد في اللغة والأعلام، للويس معلوف- المعجم الوسيط، لإبراهيم مذكور وآخرين (مجمع اللغة العربية بالقاهرة) - المعجم العربي الأساسي، مجموعة باحثين في المنظمة العربية للتربية والثقافة والعلوم- المعجم السياقي للتعبيرات الاصطلاحية، لمحمود صيني وآخرين- معجم التراكيب والعبارات الاصطلاحية العربية القديم منها والمولد، لأحمد "أبو" سعد- كتاب "مبادئ اللغة" للخطيب الاسكافي. - معجم أسماء الأدوات واللوازم في التراث العربي، لنايف النوايسة- معجم ألفاظ الحياة العامة في الأردن، مجمع اللغة العربية الأردني- موسوعة السلطان قابوس لأسماء العرب- الأعلام للزركلي.

ثانيا: المصادر الأدبية (ومنها): - المستطرف، للإبشيهي- الأذكياء، لابن الجوزي- الأمالي، لأبي علي القالي.

ثالثا: المراجع الأدبية الحديثة (ومنها): - فيض الخاطر، لأحمد أمين- إلى ولدي، لأحمد أمين- مع الله في السماء مع الله في الأرض، لأحمد كمال زكي- سلسلة العبقريات، لطه حسين- الأعمال القصصية الكاملة، غسان كنفاني- العيد من النافذة الأخرى، لسميرة عزام - الصرة الصغيرة، لأمين فارس ملحس- بعض قصص نجيب محفوظ. - بعض قصائد نزار قباني- بعض قصائد محمود درويش الأولى- قصائد مختارة من الشعر العربي في عصوره الممتدة.

رابعا: المقالات الثقافية والعلمية المنشورة في مجلات رصينة، (ومنها): - مجلة العربي (الكويت) - كتاب العربي (الكويت) - مجلة أفكار (الأردن) - المجلة الثقافية (الجامعة الأردنية - الأردن) - مجلة الثقافة الأجنبية (الكويت) - مجلة المعرفة (السعودية) - مجلة الهلال (مصر).

خامسا: الصحف العربية ومنها: - الرأي والغد (الأردن) - القدس (فلسطين) - الأهرام (مصر) - القدس العربي (لندن) - القبس (الكويت) - الرياض (السعودية) - البيان (الإمارات) - السفير (لبنان).

سادسا: الثقافة والحضارة الإسلامية: - كتب في مبادئ الشريعة الإسلامية- كتاب تقديمي في السيرة النبوية وسير الخلفاء الراشدين.

ولعل معترضا يسأل عن سبب إدخال المعاجم القديمة ضمن مدونة هذا المعجم! ولعل إجابة الحصيف الواثق تتمثل في أن معظم الألفاظ الواردة في المعاجم القديمة كاللسان والتاج ما تزال قائمة فينا، ومستعملة في زمن الناس هذا، وأما الألفاظ التي هجرها الاستعمال فإنما تمثل نسبة ضئيلة من مجموع المعجم العربي.

أما الدراسات العلمية الإحصائية التي سيستفيد منها بناء المعجم فهي:

دراسات مروان البواب:

- دراسة إحصائية لحروف اللغة العربية.

- دراسة إحصائية لدوران الحروف العربية المشكولة.

أ- دراسة محمد جواد النوري: أبواب الفعل الثلاثي ... دراسة لغوية تحليلية إحصائية باستخدام الحاسوب.

ب- دراسة محمد حسن الطيان ويحي مير علم: قاعدة معطيات للجذور العربية.

ج- دراسة محمد حسن الطيان: تنافر الحروف ودورانها في نسج الكلمة العربية بين التراث والدراسات الإحصائية الحديثة.

د- قائمة المفردات الشائعة في اللغة العربية، لداود عبده.

و- قائمة مكة للمفردات الشائعة، أعدتها لجنة من معهد اللغة العربية بجامعة أم القرى.

ز- مشروع الرصيد العربي الأساسي، الرباط.

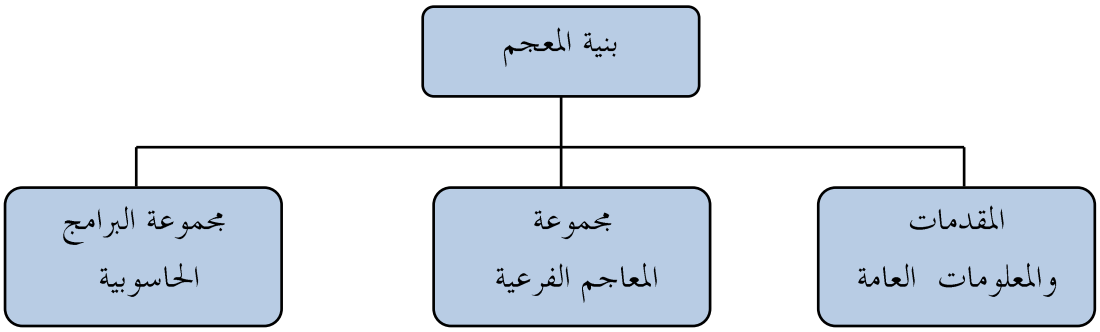
ي- مشروع الذخيرة اللغوية العربية.

2- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، مجلة الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، يناير 2009، ص 239.

4.10- بنيـــة المعجم وطريقة عمله:

بنية المعجم:

تقوم بنية هذا المعجم الحاسوبي متعدد الأغراض على ثلاثة عناصر رئيسية يمثلها الشكل التالي[1]:



- المقدمات والمعلومات العامة.
- مجموعة من المعاجم الفرعية.
- مجموعة البرامج الحاسوبية المساندة الناجزة؛ وقد عرضنا لها في سياق توصيف المعجم.

أولا: المقدمات والمعلومات العامة[2]:

وتتضمن هذه المقدمات الكلية[3]:

1- التعريف بالمعجم. ويتضمن هذا التعريف معلومات تفصيلية يحتاجها المستخدم، ومنها: أهداف المعجم وأغراضه، وطريقة بنائه، والفئة المستهدفة منه، وطريقة استخدامه ولا سيما كيفية البحث عن معلومات متعددة في الوقت نفسه، كبيان كيفية ربط معجم المفردات بمعجم الأعلام أو معجم التعبيرات الاصطلاحية ...الخ.

1- المرجع السابق، ص 244.

2- يمكن وضع هذه المقدمات بناء على الدراسات التي وضعت في ذلك، ومنها علي القاسمي: ماذا نتوخى في المعجم العربي للناطقي باللغات الأخرى، في كتاب "المعجمية العربية بين النظرية والتطبيق، ص.131 – 139. وأيضاً: توصيات الدورة التدريبية لصناعة المعجم العربي لغير الناطقين بالعربية، المبادئ الأساسية في تصنيف المعجم العربي، ص. 5 – 12. وقد نشرت هذه التوصيات أيضاً في: وقائع تعليم اللغة العربية لغير الناطقين بها، ج1، مكتب التربية العربي لدول الخليج، 46 – 53.

3- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، مجلة الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، يناير 2009، ص 245.

2- معلومات موجزة عن اللغة العربية. ويمكن أن تتضمن معلومات تعريفية بنظم العربية المتعددة، بحيث يركز على الخصائص العامة لكل نظام مما يحتاجه المتعلم[1].

3- المعاجم الفرعية المتخصصة:

ويمثل المخطط التالي بنية المعجمات الفرعية المتخصصة التي يتألف منها المعجم المقترح من طرف الدكتور العناتي، وهي تمثل العنصر الرئيسي من المعجم الحاسوبي[2].

1- يعتمد نظام المعجم على :

النظام الصوتي: وينبغي أن يتضمن معلومات مهمة ومنها: الفونيمات الرئيسية في العربية (الصوامت والصوائت) وكيفية نطق كل واحد منها منفردا ثم في سياقات مختلفة، والتمييز بين العلل القصيرة والطويلة وأثر ذلك في المعنى. وينضاف إلى ذلك القضايا التعاملية وقيود التتابع في العربية، كمنع البدء بصامتين، والوقوف على صامت لا متحرك، وتحويل التاء المربوطة إلى هاء عند الوقف عليها، وتقصير العلة الطويلة عندما تتبع بكلمة مبدوءة بأل التعريف. ومن ذلك التفريق بين الأصوات المتشابهة (/ س/و/ص/، و/د/و/ض/...الخ). كل ذلك في سياق يجمع بين النطق والكتابة.

- النظام الصرفي: ويركز فيه على خصيصة الاشتقاق في العربية في صورتها البسيطة (الاشتقاق الصغير)، وينبغي أن يشار في هذا الموضوع إلى قضيتين مهمتين: الأولى: احتفاظ المشتقات بالمعنى الأصلي الذي يحمله الجذر مضافا إليه المعنى المكتسب من الوزن الصرفي، والثانية: احتفاظ الجذر الأصلي بترتيب حروفه في المشتق دون أي تغيير.

- النظام النحوي والتركيبي: ويفضل أن يكون في قسمين: الإعراب والتركيب. ويمكن في الإعراب ربط الحركة الإعرابية بالوظيفة النحوية، وإنما يكون ذلك بمجرد الأسماء المرفوعة والأسماء المنصوبة والأسماء المجرورة، وكذا القول في المبنيات من الأسماء والأفعال، على أن تكون المعالجة هنا مقتضبة ولا تتجاوز أن تكون إشارات عامة يأتي تفصيلها في المعجم النحوي. وأما التركيب فيركز فيه على القضايا التركيبية الرئيسية في نظام العربية، ومنها:

أ- صدارة أسماء الاستفهام وأسماء الشرط للجملة العربية.

ب- تأخر الصفة عن الموصوف.

ت- تتابع المضاف والمضاف إليه.

- النظام الدلالي والمعجمي: وهما يقترنان اقترانا كبيرا بالمستوى الصرفي وخصيصة الاشتقاق،ولكنه ينبغي الإشارة إلى أهم الظواهر الدلالية في العربية، كالترادف والاشتراك والتضاد، والحقيقة والمجاز، والمعنى اللغوي والمعنى الاصطلاحي ...الخ.

- النظام الكتابي. وسيكتفى هنا بإشارات سريعة لنظام الكتابة العربية، لأنه سيكون هناك رابط مستقل للنظام الكتابي.

- الكتابة العربية: وقد رأى أحمد العناتي أن يكون لها رابط مستقل يتناول النظام الكتابي للعربية بالتفصيل مسنودا بنظام النطق وتقنية الكتابة الحاسوبية المتقدمة التي يمكنها أن تظهر كيفية كتابة الحروف واتجاهها. وينبغي أن يتضمن هذا الرابط معلومات منها:

أ. أشكال الحروف العربية على اختلاف مواضعها من الكلمة، حين تكون في أول الكلمة، أو في وسطها، أو في آخرها، وحين تكون منفصلة أو متصلة ...الخ.

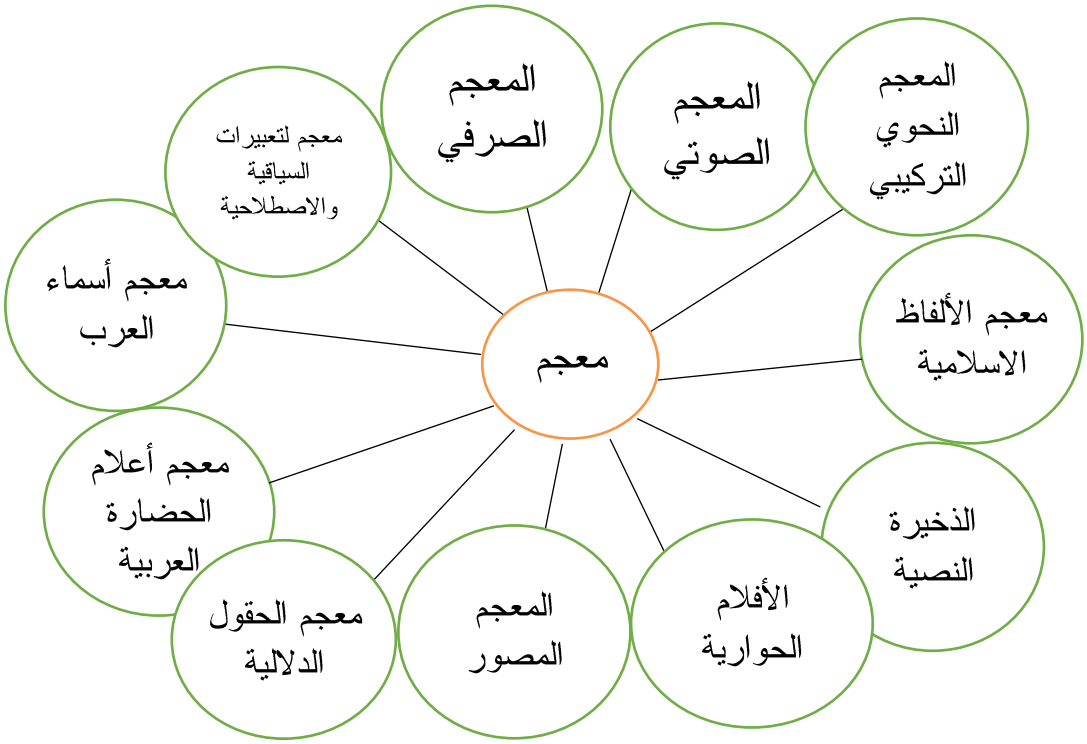
ب- الحروف التي تكتب ولا تنطق.

ت- الحروف التي تنطق ولا تكتب.

ث- اختلاف رسم الهمزة حسب موقعها.

ويمكن أن يتضمن هذا الرابط مسحة فنية جمالية، وذلك بتضمينه لوحات مكتوبة بالخطوط العربية المختلفة، وليس القصد من ذلك تعليمهم أنواع الخطوط، وإنما نقصد من ذلك بعدا ثقافيا وحضاريا يعرف بمنجز الحضارة الإسلامية، ولعله يكون عامل جذب وتحبيب للطلبة باللغة العربية وخطوطها، ولا سيما إن اقترنت بلوحات فنية تبرز خصوصية الخط العربي وخصوصية الحضارة العربية الإسلامية فنياً.

2- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 248.



1- المعجم الصوتي:

وتتمثل أهمية المعجم الصوتي في أنه يمثل صورة العربية المنطوقة في جميع تحققاتها الفعلية، بدءا بالصوت المفرد وانتهاء بالجمل والنصوص المكتملة، ويختلف هذا المعجم عن المعاجم الأخرى المتضمنة هنا في أنه لن يعتمد على برنامج جاهز وإنما سيخزن تخزينا، ذلك أن برامج استنطاق النص العربي لم تبلغ مرتبة راقية في إخراج الصوت بصورة متميزة[1]. وسيكون دور المعجم الصوتي دوراً مسانداً، إذإنه يمكن النفاذ إليه من المعاجم الأخرى. فمن معجم المفردات يمكن تحصيل نطق المفردة ومعانيها المتعددة، كما يمكن النفاذ من المعجم النحوي لقراءة التراكيب الجملية المحفوظة هناك. كذلك يعد المعجم الصوتي مركزياً في الاستماع إلى الذخيرة النصية المحفوظة؛ حيث يمكن للمتعلم الاستماع إليها غير مرة، والقول نفسه ينطبق على معجم الأفلام الحوارية المسجلة والمضمنة في المعجم، وصولاً إلى تمكين المتعلم من نظام العربية الصوتي ولا سيما الفونيمات فوق القطعية التي تحتاج إلى استماع مستمر لنصوص العربية في واقعها التداولي الفعلي، كالنبر والتنغيم والوقف ... الخ[2].

1- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 248.

2- المرجع نفسه، الموضع نفسه.

2- معجم المفردات:

يمثل هذا المعجم قلب المعجم المقترح؛ فهو نقطة البدء التي ينطلق منها المتعلم لمعرفة معاني المفردات وتشبعاتها الدلالية والنحوية والصرفية. وغالباً ما يدخل المتعلم إلى هذا المعجم من الواجهة الرئيسية مباشرة ؛ إذ يمكنه النفاذ من لوحة الحروف الهجائية العربية إلى ما يريد.

الحروف الهجائية العربية	
أ	ب
ت	ث
ج	ح
خ	د
ذ	ر
ز	س
ش	ص
ض	ط
ظ	ع
غ	ف
ق	ك
ل	م
ن	ه
و	ي

وظاهر أن هذه اللوحة[1] تتيح خيار البحث المباشر في الأحرف العربية، وأكثر المستفيدين منه من يعرفون آلية تخليص الكلمات من الزوائد ومعرفة الجذور. ويكون تحصيل المادة المرادة بالضغط على الحرف المراد، ولتأخذ حرف (س) نموذجاً. إن النقر على (س) سيفضي بالمتعلم (مثلاً) إلى الجذور التالية[2]:

1- المرجع السابق، ص 250.

2- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 250.

سأب	سأد	سأر	سأسأ	سأف
سأل	سئم	سبأ	سبَّ	سَبَتَ
سبح	سبخ	سبر	ستر	سدر
سعد	سغب	سفه	سكر	سلم

وهكذا يمضي الباحث حتى يحصل الجذر الذي يريده، فإذا كان يبحث عن (سأل) فعليه أن ينقر عليها ليحصل على المعلومات التالية مثلاً: (من الوسيط).

سَأَلَ ... يَسْأَلُ: استخبره. وسأله حاجة: طلبَ ... الخ.

وإذا أراد المتعلم تحصيل معلومات إضافية مختلفة فاء إلى المعاجم الأخرى المساندة.

وأما الذين لا يعرفون الجذر الأصلي، وهم غالباً من المبتدئين فإنه بإمكانهم استخدام محرك البحث مباشرة، على النحو التالي[1]:

ابحث	

وهنا يكتب المتعلم الكلمة التي يريد معرفة معناها ومزيداً من المعلومات عنها، ولنأخذ مثلاً كلمة (السائل):

السائل

وبعد تشغيل محرك البحث يحصل المتعلم على المواد التي يتضمنها مدخل (سأل)، حيث يظهر الجذر بلون مختلف عن بقية الأحرف، ويحصل المتعلم على معلومات متعددة عن هذه الكلمة (سائل)، كما يمكن العودة إلى الجذر (سأل) والنفاذ منه مباشرة إلى المعجمات الأخرى التي تقدم كل ما مخزن أو معالج[2].

3- المعجم الصرفي:

ويتضمن هذا المعجم الصرفي أبرز البنى الصرفية في العربية انطلاقاً من قواعد الاشتقاق الصرفي في العربية. ولن يكون هذا المعجم الصرفي معجماً مخزناً وإنما سيعتمد على

1- المرجع السابق، ص 250.

2- المرجع نفسه، ص 251.

برامج المعالجة الصرفية المساندة، ولعله يحسن أن توضح القواعد الصرفية ببساطة وبوضوح قبل المعالجة الصرفية المحوسبة سلفاً.

وتجري آلية العمل في المعجم الصرفي على النحو التالي[1]:

– ينقر المتعلم على أيقونة المعجم الصرفي، فيحصل على المخطط التالي (مثلاً):

المشتقات	الفعل المجرد والمزيد	المصادر	إسناد الأفعال إلى الضمائر
الجنس	العدد	الصحيح والمعتل	الاسم المنسوب

- يمثل هذا الشكل (مثلاً) المعالجات الصرفية التي يتضمنها المعجم الصرفي، فإذا أردنا أن نسند الفعل الماضي، مجرداً كان أم مزيداً، إلى الضمائر نقرنا على "إسناد الأفعال إلى الضمائر، لنحصل على المخطط التالي[2]:

أنا	نحن
أنت	أنت
أنتما	أنتما
أنتم	أنتن
هو	هي
هما	هما
هم	هن

وهذا المخطط يتضمن ضمائر اللغة العربية، وهو مدعوم بمحلل صرفي يسند الفعل المراد إسناده إلى الضمير مباشرة، مراعياً التغييرات الصوتية والصرفية والكتابية التي تطرأ على الفعل بعد الإسناد، ويبين الشكل التالي كيفية عمل المحلل الصرفي، حيث يكتب المتعلم في المستطيل الأول الفعل المراد تصريفه، ثم ينقر على الضمير المراد تصريفه في الفعل معه، فيظهر الفعل مصرفاً إلى جوار الضمير في المستطيلين السفليين، كما هو موضح في الشكلين التاليين[3]:

1– المرجع السابق، ص 251.

2– وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 252.

3– المرجع نفسه، ص 252،253.

الفعل المراد تصريفه

الضمير	الفعل بعد التصريف

دَرَسَ

أنا	دَرَسْتُ

– إن الخطوتين السابقتين في البحث في المعجم الصرفي تتمان مباشرة دون العودة إلى معجم المفردات، ولكنه يمكن الانطلاق من معجم المفردات إلى المعجم الصرفي لاستكمال هاتين الخطوتين وغيرهما من العمليات الصرفية التي يريدها المتعلم[1].

– يمكن للمتعلم مثلاً أن يستدل على البناء الصرفي لكلمة معينة بالنظر مباشرة في المعجم الصرفي لتحصيل بنيتها الصرفية وما يتعلق بها. فمثلاً إذا وقف المتعلم على كلمة "منزِل" في الذخيرة النصية، وأراد أن يتعرف بنيتها الصرفية فما عليه إلا أن ينقر على رابطة الكلمة في المعجم الصرفي، فيحصل عندها على معلومات تفيده بأن الكلمة اسم مكان، وهكذا[2].

4.المعجم النحوي التركيبي:

ويمكن أن يكون المعجم النحوي التركيبي على إحدى الصور التالية[3]:

– **الصورة الأولى:** وفيها يُضَمَّنُ المعجم معظم القواعد النحوية والتركيبية العربية في صورتها الميسرة التي تتجاوز التعقيد والخلافات النحوية؛ إذ المقصود منها تعريف الطالب بأنماط الجملة العربية، وقوانين التركيب العربي، والوظائف النحوية الأساسية في العربية، والعلاقات الأفقية والخطية التي تحكم عناصر الجملة العربية، ومنزلة السياق في فض اللبس. وتكون هذه القواعد مشفوعة بالأمثلة الدالة والشافية.

– **الصورة الثانية:** ربط المعجم النحوي بالمعاجم الأخرى ربطاً مباشراً، بحيث يتضمن هذا المعجم بياناً تفصيلياً وإعراباً ميسراً لجميع الجمل والشواهد والمتون الواردة في المعجم الكلي. فإذا ورد نص قصير في الذخيرة النصية كان لزاماً أن يكون هذا النص قد أعرِب وشرحت عناصره النحوية الرئيسية بطريقة ميسرة وموضحة.

1– المرجع السابق، ص 253.

2– المرجع نفسه، الموضع نفسه.

3– وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 253.

– **الصورة الثالثة:** وهي الاعتماد على محلل نحوي موثوق ومجرب قد ثبت نجاحه في التشكيل والإعراب الآلي؛ ليكون برنامجاً يستعين به المتعلم لفهم ما غمض عليه من قواعد العربية وتراكيبها.

– **الصورة الرابعة:** وهي الاستعانة بعدد كبير من الجمل والنصوص التي تمثل العربية الواقعية وأنماطها التداولية المختلفة، ثم قرنها بأعاريب مفرداتها بالتفصيل. (وليس المقصود هنا بالتفصيل كما هو للعربي وإنما الدلالة على الوظائف حسب:

3- معجم التعبيرات السياقية والاصطلاحية[1]:

ويمثل هذا المعجم ركيزة مهمة لتحقيق المتعلم كفاية تواصلية مناسبة باللغة العربية.

ويفترض في هذا المعجم أن يتضمن عدداً كبيراً من التعبيرات الاصطلاحية المتداولة بالعربية في المجتمع العربي. وتتحمل هذه التعبيرات بقيم ثقافية تدل على سياقات تداول هذه التعبيرات، وغالياً ما تكون معانيها واستعمالاتها مخالفة لمعاني كلماتها مفردة. ومن هنا كان فإن هذا المعجم يمثل الواقع التداولي للعربية، ولعل الطريقة المناسبة لتقديم هذه التعبيرات تقديمها مقرونة بالسياقات والمواقف الواقعية الحقيقية التي تستعمل فيها، موضحة قيود التداول التي تحكمها.

4- معجم أسماء العرب[2]:

ويضمّ هذا المعجم عددا كبيرا من الأسماء التي يسمي بها العرب أبناءهم وبناتهم، وتأصيل هذه الأسماء إن كانت عربية الأصل أو مقترضة من لغات أخرى. ثم إنه من الضروري إضافة معلومات تفصيلية حول وجوه استخدام الإسم. ويمكن البحث في هذا المعجم مباشرة من رابط فرعي داخله، حيث توجد لوحة تظهر الأحرف الهجائية العربية مرتبة، فما على الطالب إلا أن ينقر على الحرف المطلوب فيحصل على جميع الأعلام العربية المستخدمة3، أما ترتيب هذا المعجم فإنه سيترتب هجائياً حسب الحرف الأول منه دون حاجة إلى استرجاع الجذر.

1– اختلفت تعريفات "التعبير الاصطلاحي" بين المعجميين العرب بين: التعبير الاصطلاحي، والتعبير السياقي، والتعبير المسكوك. لمزيد من التفاصيل، ينظر: – على القاسمي، معالجة التعابير الاصطلاحية والسياقية في المعجم العربي، في كتاب المعجمية العربية بين النظرية والتطبيق، ص 87،111.

– محمد الحناش، برنامج لساني حاسوبي للتعرف الآلي على التعابير المسكوكة في اللغة العربية، وقائع مناظرة اللغة العربية والتقنيات المعلوماتية المتقدمة، مؤسسة الملك عبد العزيز آل سعود، الدار البيضاء، 1993، ص. 72 – 90.

– محمد الحناش، قاعدة بيانات التعابير المسكوكة في اللغة العربية، المؤتمر الأول حول اللغويات الحسابية العربية، القاهرة، 1992.

– محمد صيني ومختار الطاهر حسين وسيد الدوش، المعجم السياقي للتعبيرات الاصطلاحية؛ عربي – عربي، ط.1، مكتبة لبنان ناشرون، بيروت، 1996.

2– يمكن الاستفادة من الدراسية البحثية الجيدة: موسوعة السلطان قابوس لأسماء العرب، ط. 2، مكتبة لبنان، المطابع العالمية بمسقط، 1991.

3– للمزيد من بسط القول ينظر : وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 256.

5- معجم أعلام الحضارة العربية الاسلامية[1]:

ويقصد هذا المعجم إلى التعريف الموجز بأهم أعلام الحضارة العربية الإسلامية في جميع وجوهها، وقد يبدو مناسباً تقسيمه إلى حقول معرفية أو علمية أو ثقافية متعددة، كأعلام الأدب العربي، وأعلام الفلسفة، وأعلام الفقه، وأعلام القادة والعسكريين والسياسيين، وأعلام الأطباء ...الخ. ويمكن اتباع الترتيب التاريخي، وذلك كتقسيم الأدباء العرب إلى حقول ومن ثم ترتيب هذه الحقول تاريخياً؛ فإذا كان المقصود أعلام الشعر بدأنا بالعصر الجاهلي ثم صدر الإسلام ...الخ[2].

ولعل هذا المعجم يكون سهل التنفيذ؛ ذلك أن معظم التآليف التي توثق لتاريخ الحضارة العربية الإسلامية وأعلامها تعتمد الترتيب التاريخي والموضوعي، فلا يحتاج الأمر إلا إلى تحويل المادة من ورقية إلى حاسوبية[3].

ويمكن النفاذ إلى ها المعجم مباشرة بالنظر في اسم العلم مباشرة، أو بالنظر في الحقل العلمي الذي برز فيه، إذا كان المتعلم عارفاً ذلك، وإلا فإنه أيسر له البحث عن العَلَمِ مباشرة.

6- معجم الألفاظ الإسلامية:

ويفترض في هذا المعجم أن يتضمن أهم ألفاظ الديانة الإسلامية في جميع المجالات، كالصلاة، والزكاة والصيام، وسائر العبادات، والمعاملات ...الخ. وينبغي أن تقدم هذه الألفاظ على شكل تعريفات بصورة ميسرة ومبسطة.

7- معجم الحقول الدلالية:

ويضم هذا المعجم عدداً مناسباً من مجاميع الألفاظ قد صنفت وفقاً لدلالتها العامة، أو ما يعرف بالحقول الدلالية، بحيث تغطي هذه الحقول جميع مجالات الحياة الإنسانية العامة، وما تختص به البيئة العربية والثقافية العربية الإسلامية. ومن الحقول التي سيتضمنها المعجم الحقول التالية[4]:

1- يمكن الاعتماد على موسوعة أعلام العلماء والأدباء العرب والمسلمين، المنظمة العربية للتربية والثقافة والعلوم، والمشروع في طور الإنجاز.

2- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 257.

3- المرجع نفسه، الموضع نفسه.

4- المرجع نفسه، ص 258.

الحقل	نماذج من مفردات الحقل
وسائل النقل	الطائرة، السيارة، الدراجة الهوائية، السفينة، القطار، الحيوان.
وسائل الاتصال	البريد، الهاتف بأنواعه، الإنترنيت.
الفواكه	التفاح، البرتقال، الموز، الخوخ، الجوافة، التين، الدُّرّاق.
الألعاب	كرة القدم، كرة الطائرة، كرة السلة، كرة المضرب، السباحة، ...
التضاريس	جبل، هضبة، وادي، تلة، نهر، بحر، ... الخ.
الألوان	الأبيض، الأسود، الأخضر، الأصفر، البرتقالي، البني، ...الخ.
مكونات البيت	غرفة الطعام، غرفة النوم، غرفة المعيشة، المطبخ، الحمام، ...
الحيوانات	حصان، حمار، أسد، زرافة، جمل، كلب، نمر، ...

ولعله يحسن أن نشير إلى الملاحظات التالية[1]:

– ترتب هذه الحقول ترتيباً هجائياً من الحرف الأول ودون حاجة إلى إرجاع الكلمة إلى جذرها إن كانت كلمة مشتقة.

– ينبغي تقسيم كل حقل دلالي[2] عام إلى حقول فرعية أخص من الحقل العام؛ فالحيوانات تقسم إلى حيوانات مفترسة، وطيور، وزواحف، وثدييات ...

– ويمكن أن يعمل هذا المعجم متصلاً بالمعجم المصور، ولاسيما في المعاني المادية المحسوسة؛ إذ يمكن الوصول إليها برابط معين.

– لعله من المفيد هنا استثمار تقنيات الألوان والصوت والصورة بما يخدم المعنى المراد توضيحه، فإذا كان القصد توضيح الألوان أظهرنا كتابة كل لون باللون الذي يمثله، فالأحمر

1– المرجع السابق، الموضع نفسه.

2– حقل دلالي : champ semantique : يعني به مجموع استعمالات كلمة مدخل من مداخل المعجم، وما توفره من معان باعتبار السياقات التي ترد فيها، من ذلك أن مدخل" بان" لازما ومتعديا، يتوفر على ثمانية معان في المعجم الوسيط لمجمع اللغة العربية بالقاهرة منها :

1– بان الشيء بيانا : ظهر واتضح

2– بان فلان : رحل وابتعد

3– بانت الفتاة : تزوجت

4– بانت الشجرة : طالت طولا ظاهرا.

ولقد سبق لأحمد بن فارس أن أطلق على هذا الحقل مصطلح الفرع ج فروع. ويعول على الحقل الدلالي النحوي على وجه الخصوص، لتوليد المعاني والدلالات من المداخل الأصول ج الأصل. ينظر : محمد رشاد الحمزاوي، المعجمية،(مقربة نظرية ومطبقة، مصطلحاتها ومفاهيمها)، مركز النشر الجامعي، تونس، د.ط، 2004، ص 186.

يكتب بالأحمر، والأزرق يكتب بالأزرق، وهكذا. أما إذا كان المعنى المراد حيواناً فلعل صورة متحركة ممزوجة بالصوت تدل على المقصود.

8- المعجم المصور:

ويتضمن هذا المعجم المصور لوحات أكثر تفصيلاً مما هو موجود في معجم الحقول الدلالية. ويعتمد هذا المعجم على الرسوم والأشكال التوضيحية المعتمدة على تقنيات التوضيح الحاسوبية المتطورة، كالألوان، والأصوات، وتقطيع الصورة، وبيان الأجزاء المقصودة بالإضاءة، وسوى ذلك[1].

9- الذخيرة النصية[2]:

ونقصد بالذخيرة النصية مجموعة كبيرة من النصوص العربية على اختلاف أعصرها وموضوعاتها وفنونها، فهي تجمع بين الجاهلي والإسلامي والمعاصر، وتجمع القصيدة والقصة والمقالة والبحث العلمي، وتضم مقالات ونصوصاً من القرآن والحديث والتاريخ الإسلامي، والشعر العربي، والقصة العربية والمقالة ...الخ[3].

ولعله يحسن التنبيه هنا إلى أن القصد هنا تقديم صورة للعربية في مختلف تحققاتها الفكرية والإبداعية، بحيث تمثل هذه النصوص المختارة وجوهاً من تداول العربية في مستواها الفني المكتوب، فليس القصد إذن حشد النصوص وجمعها كيفما اتفق، وإنما ينبغي أن تستند إلى معايير محددة كالشيوع، والتمثيلية؛ بحيث تصلح هذه النصوص أن تكون نموذجاً[4] ممثلاً للفن والموضوع الذي تنتمي إليه[5]. وأن تكون قصيرة ومتوسطة الطول ...الخ[6].

1- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 259.

2- يمكن الاعتماد هنا على ما أنجزه مشروع الذخيرة اللغوية العربية الذي تعمل عليه المنظمة العربية للتربية والثقافة والعلوم.

3- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 261.

4- يمكن أن نستفيد من البرامج المساندة في فهم النصوص وتحليلها، وبيان ذلك فيما يلي:

- يمكن معرفة معلومات مختصرة عن موضوع النص بالنقر على الاسم؛ فتستدعى مثلا المعلومات من معجم أعلام الأدب العربي المتفرع من معجم أعلام الحضارة العربية الإسلامية.

- يمكن الاستماع إلى النص قراءة بصوت واضح ودال. ويمكن أن نستعين بتسجيل حقيقي لصاحب النص إن كان على قيد الحياة؛ إذ يمكن أن نسجل نصاً لمحمود درويش، مثلاً.

- يستدعي المتعلم معجم المعاني لمعرفة معنى كلمة "أغبطه" وهنا يعمل المعجم الدلالي على المعجم الصرفي الذي يفصل الضمير من الكلمة ويعيدها إلى أصلها الثلاثي.

- يمكن للمتعلم استدعاء التركيب الاصطلاحي "من عرق جبينه" من المعجم السياقي والاصطلاحي؛ لإتمام فهم هذا التركيب داخل السياق.

- يمكن أن يبحث المتعلم عن إعراب مختصر لكلمة "استجداء" باستدعاء المعرب أو المحلل النحوي.

وهكذا تعمل البرامج المساندة في خدمة الذخيرة النصية، بحيث تنتهي هذه الخدمات إلى الإجابة عن كل ما يمكن أن يخطر في نفس المتعلم.

5- تمثل هذه النصوص ميداناً رحباً لتفعيل البرامج المساندة التي يتضمنها المعجم، كالقراءة (البشرية المسجلة)، والتحليل الدلالي، والتحليل الصرفي، والتحليل النحوي، والأسلوبي ...الخ.

6- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 261.

10- أفلام حوارية وتعليمية:

ولعله يبدو مفيداً جدا أن يتضمن هذا المعجم مجموعة منتقاة من المشاهد التمثيلية والحوارية القصيرة التي تمثل واقع استعمال اللغة العربية في ميادينها المختلفة. ويمكن أن تعد هذه المشاهد إعداداً خاصاً للمعجم، ويمكن أن تنتقي بعناية من البرامج جيدة المستوى مما تعرضه الفضائيات العربية. ولعله يحسن أن يراعى في هذه البرامج المنتقاة ما يلي[1]:

- أن تكون باللغة العربية الفصحى المعاصرة التي تمثل قدراً مشتركاً من صورة العربية في الوطن العربي كله؛ وهذا يعني الابتعاد عن السمات اللهجية المحلية قدر الإمكان.

- أن تكون ممثلة لجميع الموضوعات السياسية، والاجتماعية، والاقتصادية ....الخ.

- أن تتفاوت في أطوالها بين برامج قصيرة وسريعة وبرامج متوسطة.

- أن تكون هذه البرامج وظيفية؛ أي أن تقدم مادة وظيفية يحتاجها المتعلم في حياته اليومية، للتعبير بالعربية.

ولا شك أن هذا المعجم التمثيلي سيمثل منتهى تفاعل المعجمات الفرعية الأخرى؛ إذ سيقدم العربية في صورتها الواقعية المتداولة منطوقة في مواقف حقيقية يومية مما يلابسه المتعلم الأجنبي في المجتمع العربي[2]. ولعل هذا الجزء من المعجم يمثل الاختبار الحقيقي لمدى تحقق أهداف المعاجم الفرعية الأخرى: بلوغ الكفاية اللغوية[3]، التي تنتهي بكفاية تواصلية تهيئ للمتعلم الأجنبي فرصة التفكير والتواصل بالعربية.

1- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 263.

2- المرجع نفسه، الصفحة نفسها.

3- وقد عرّفها تشومسكي بأنها القدرة على إنتاج عدد هائل من الجمل من عدد محدود من الفونيمات الصوتية والقدرة على ربط الأصوات المنتجة وتجمعها في مورفيمات تنتظم في جمل مناسبة. للمزيد من التفصيل ينظر في:

- Chomsky, Noam. 1965. Aspect of the theory of Syntax. Longman Group. First Publisher. London. P. 1.
Kempson, Ruth. M.1997. Semantic Theory. Cambridge University Press. P. 93

ولعل الشكل التالي يصلح أن يكون واجهة للبرنامج؛ الصفحة الرئيسية[1]:

نـون والقلم								
معجم حاسوبي لغير الناطقين بالعربية								
تعريف بالمعجم		محتويات المعجم		تعريف باللغة العربية			الكتابة العربية	
الحروف الهجائية العربية								
أ	ب	ت	ث	ج	ح	خ		المعجم الصوتي
د	ذ	ر	ز	س	ش	ص		معجم المفردات
ض	ط	ظ	ع	غ	ف	ق		المعجم الصرفي
ك	ل	م	ن	هـ	و	ي		المعجم النحوي التركيبي
								معجم التعبيرات السياقية والاصطلاحية
								معجم أسماء العرب
بحث			موضع كلمة البحث					معجم أعلام الحضارة الإسلامية
								معجم الألفاظ الإسلامية
								معجم الحقول الدلالية
								المعجم المصور
								الذخيرة النصية
								(نصوص عربية)
								أفلام حوارية وتعليمية

صفوة القول:

من خلال تجربة الدكتور وليد العناتي فإن تنفيذ معجم موسوعي حاسوبي أحادي لغير الناطقين بالعربية يمثل ضرورة ملحة لرفع كفاية متعلمي العربية الأجانب، ولاستثمار منجزات التقنية الحديثة في نشر العربية وتعميمها في الخارج. وبما أن كثيراً من الشعوب الإسلامية ما

1- وليد أحمد العناتي، " نحو معجم حاسوبي أحادي للناطقين بغير العربية "، سابق، ص 264.

تزال تستعمل الحرف العربي، وهي تعمل جاهدة على تطويع الحاسوب لاستعمالاته،فهذا يعتبر عاملاً مهماً يساعد في تعزيز نشر العربية حاسوبياً..

11ـ تجربة المعجم الاسلامي للدكتورة فاتن الخولي[1] :

اعتمدت هذه المقاربة على:

أولاً:الآليات اللسانية:

يعتمد التأليف المعجمي مراحل عديدة تشمل:جمع المعلومات، اختيار المداخل، وترتيبها طبقا لنظام معين، تحرير المواد، ثم نشر النتاج النهائي[2]، يسبق هذه المراحل إعداد تصور مبدئي للمنهج وتحديد آليات التنفيذ وفق جدول زمني محدد[3].

- **جمع المعلومات:**

جمع المعلومات أولى مراحل إعداد المعجم وفيها يتحدد المستوى اللغوي[4] الذي يعتمده المعجم والمصادر التي يستقي منها مفرداته.

تضم المصادر المستعملة[5] من مفردات المعاجم السابقة والمدونات الأدبية والعلمية المختلفة، فتتنوع ما بين الأساس الثانوي والرافد، على النحو التالي[6]:

- المصادر الأساسية، وتحوي جميع المادة الحية المستقاة من نصوص واقعية.
- المصادر الثانوية، وتشمل المعاجم السابقة.

1- الأستاذة فاتن الخولي ،شركة حرف لتقنية المعلومات ،قسم النشر الإلكتروني،مصر والباحثة في علم اللغة.

شاركت د.فاتن الخولى ببحث عن المعجم الإلكترونى وجهود شركة حرف فى بناء المعاجم والقواميس الإلكترونية، وقد حضرتها – ممثلة لشركة حرف فى ندوة دولية يومى 21 و 22 مارس عام 2008 نظمها فريق البحث فى اللغة والمجتمع واللغات السامية بالكلية، والجمعية المغربية للدراسات المعجمية، وجمعية المعجمية العربية بتونس. واستضافتها كلية الآداب والعلوم الإنسانية بجامعة ابن زهر بمدينة أكادير فى دولة المغرب العربى الشقيقة. وكان موضوع الندوة :المعجم العربى الحديث وتحديات عصر العولمة – المعجم والحاسوب وقد تحدثت الأستاذة فاتن فى كلمتها عن تجربة حرف الناجحة والرائدة فى بناء المعاجم والقواميس الإلكترونية ممثلة فى برنامج القاموس الإسلامى الذي يعد أول معجم إلكترونى متخصص يضم مجموعة كبيرة من المصطلحات التى تغطى جميع فروع المعارف والعلوم الإسلامية مع ترجمتها إلى اللغات الإنجليزية والفرنسية والإندونيسية والماليزية فجمع بذلك بين ميزة المعجم المتخصص وميزة القاموس متعدد اللغات. وشركة حرف مستمرة فى مسيرتها الناجحة فى خدمة اللغة العربية وتطويعها لتقنيات الحاسوب من خلال تطوير برنامج المعجم الشامل، ونظام التحليل الصرفى الآلى للغة العربية، ومحرك البحث العربى، والمصنف الآلى المبنى على المكانز الموسوعية كالمكنز العربي الشامل والمكنز الإسلامي الشامل، والملخص الآلي المبني على تحليلات لغوية دقيقة للكلمة والجملة العربية... إلخ

2- علي القاسمي، علم اللغة وصناعة المعجم، مرجع سابق.

3- أحمد مختار عمر، صناعة المعجم الحديث، عالم الكتب، ط1، 1998 ص. 66، 75.

4- أما المستوى اللغوي فيتوزع ما بين الفصحى والعامية القديم منها والحديث، يشتمل المستوى الفصيح على:ما دُوِّنَ في المعاجم القديمة، والمولد الدي وضعه المولدون في مختلف عصور العربية بعد القرن الأول الهجري خاصة،والكلمات الأعجمية المعربة أو الدخيلة

5- تنتسب مصادر المعجم العربي المعاصر إلى جميع عصور اللغة (من القرن الخامس إلى القرن العشرين الميلاديين) ومن مختلف المستويات اللغوية:من القرآن الكريم حتى الصحافة اليومية.

6- فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، مجلة الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، يناير 2009، ص 144.

– المصادر الرافدة، وتضم مجموعة مراجع للتوثيق وتحديد العبارات المسكوكة...[1].

وتتطلب هذه الشمولية من المعجمي بحثاً معمقاً واستقصاء كاملاً ليحقق التوازن بين مصادره مما يجعل عمله يكلل بالصعوبة[2]، ومن هنا يتبدى الدور الذي تلعبه تقانة المعلومات الحديثة في هذا الشأن، لاسيما فيما تغافل عنه المعجميون المعاصرون من وسيلة الجمع الميداني لمفردات اللغة الحية، تلك الوسيلة التي انتهجها القدماء في جمع مادتهم المعجمية.

- **اختيار المداخل:**

تفتقت عبقرية الخليل بن احمد الفراهيدي عن طريقة مبتكرة، استطاع من خلالها إحصاء جميع الجذور الممكنة في اللغة العربية، اعتمدت طريقته هذه نهجا رياضيا يسهل حوسبته، يقوم على تصنيف جذور العربية إلى ثنائي فثلاثي فرباعي فخماسي، ثم تقليب حروف تلك الجذور مع بعضها البعض على النحو التالي[3]:

– الثنائي (2×1)=2

– الثلاثي (3×2)= 6

– الرباعي (4×6)= 24

– الخماسي (5×24) = 120

ينتج عن هذه الطريقة جميع احتمالات الجذور في العربية، المستعملة في الواقع اللغوي أو المهملة، ومع تركيبها على أوزان العربية تمكن من تحديد مداخل معجمه، مع الإشارة إلى المستعمل منها والمهمل.

على الرغم من احتواء تراثنا المعجمي على هذه الآلية الرياضية الفريدة إلا أن علماء اللغة التطبيقيين لم يتوصلوا "إلى قواعد علمية أو مبادئ معلومة تحكم اختيار المداخل"[4].

1– أحمد مختار عمر، صناعة المعجم الحديث، مرجع سابق، ص 77.

2– علي القاسمي: الخصائص المميزة الرئيسية للمعجمية العربية، مجلة اللسان العربي، ع 49، 199.

3– فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، مجلة الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية، العدد السابع والثامن، يناير 2009، ص 145.

4– نجدفي هذه النقطة محاولة د.أحمد مختار لوضع إجراءات تحكم اختيار الوحدات المعجمية قبل الشروع في إعداد المعجم، تتمثل في:

–إعداد بيان تقديري بعدد المداخل أو المواد في الحرف الواحد؛ مما يسهم في تحديد حجم المعجم، وحجم كل من المادة الموسوعية والمصطلحية فيه.

– وضع قاعدة للتعامل مع الكلمات متعددة المعنى، مع مراعاة تغليب المعيار الدلالي.

– اتخاذ قرار بشأن الكلمات الوظيفية، حيث يصعب تعريفها، هل تعرف بدلالتها أم بوظائفها.

– اختيار منهج للتعامل مع الكلمات المركبة والمنحوتة.

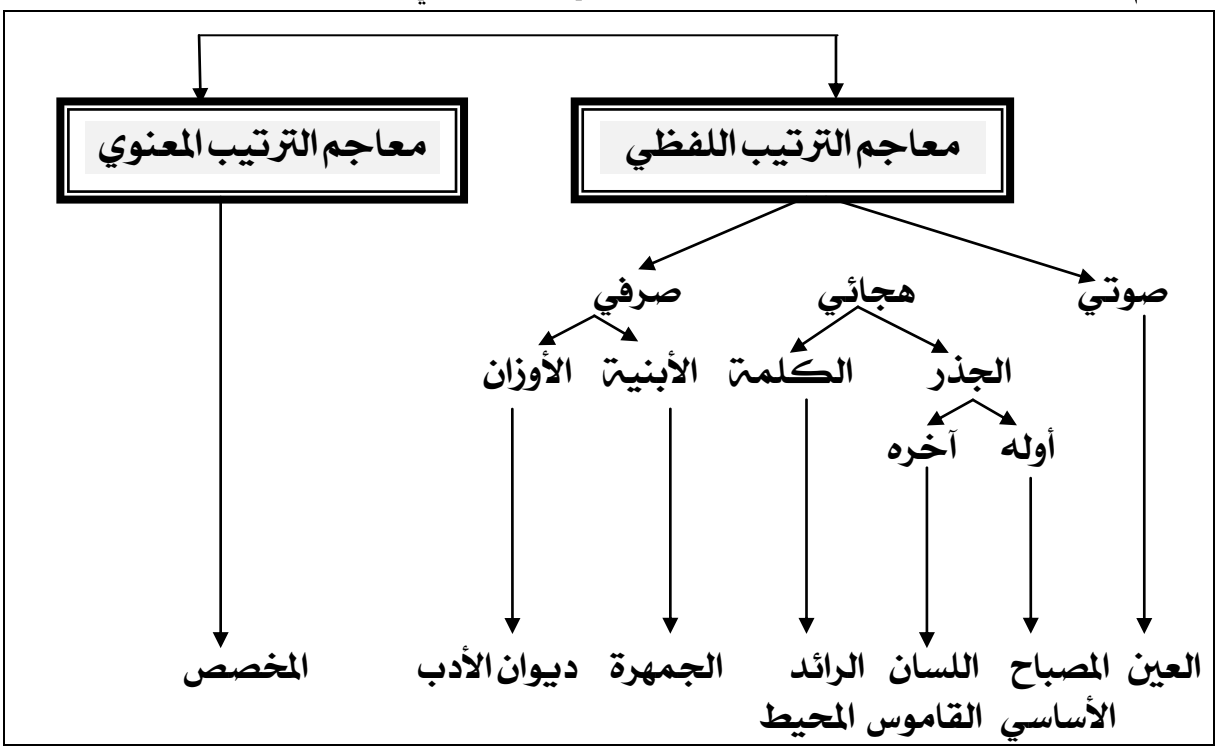
يمكن إضافة إجراءين آخرين هما:

– تحديد نوع المدخل ومواده من حيث الاستعمال والإهمال.

– إعداد قائمة بالكلمات المستحدثة التي لم يرد لها معان في المعاجم السابقة.

- **ترتيب المداخل:**

شهد تاريخ الصناعة المعجمية العربية على مدار مئات السنين تعدد منهجيات الترتيب المدخلي للمعجم، غير أن معظمها قد انطلق من الجذور على النحو التالي:



تتمثل أهم فوائد الترتيب الجذري في جمع العائلة اللفظية في مدخل واحد مما يجعل التعريفات أقصر وأيسر على الفهم، غير أن له مساوئ عديدة، منها[1]: عدم استطاعة غير المتخصص الوقوف على جذر الكلمة التي يبحث عن معناها في المعجم، فضلا عن تعدد الجذور لبعض مفردات اللغة على نحو ما نجد في: ميناء ومدينة.

وعلى الرغم من اختفاء كثير من هذه المناهج، السابقة، في المعاجم المعاصرة إلا أنها لم تزل متنازعة ما بين الترتيب الألفبائي حسب الجذر وحسب الكلمة، هذا وقد تعددت طرق الترتيب الداخلي للمدخل كذلك، ما بين الترتيب: الألفبائي، والصرفي من حيث الفعلية والاسمية والتجريد والزيادة والدلالي من حيث العموم والشيوع والحسية والعقلية[2]...

1- فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، سابق ص 147.

2- المرجع نفسه، الموضع نفسه

- **تحرير المواد**[1]

تتمثل المعلومات التي وضعتها الدكتورة فاتن الخولي لبنلء قاعدة البيانات المعجمية كالآتي[2]:

– معلومات صوتية، تعتمد تحديد هجاء الكلمة وضبطها بالشكل كذلك معالجة كل من التنوع النطقي والهجائي لها.

– معلومات صرفية، تتضمن: تصريفات الفعل ومشتقاته، الإشارة إلى طريقة الإسناد، تحليل المركبات والنحوتات...

– معلومات نحوية أساسية تشمل: التعدي واللزوم، المطابقة،الأفعال الناسخة، أفعال المدح والذم، الممنوع من الصرف، التمييز والحال والاستثناء،إعراب الأدوات، تعيين الشواهد، الإشارة الى المسائل النحوية...

– بينما تشمل المعلومات الدلالية: (المعاني، التعبيرات السياقية، تحديد المستوى اللغوي: فصحي، عامي، تحديد تاريخ الكلمة: قديم، حديث، معاصر...، تأصيل الكلمات الدخيلة أو المعربة.)

– وتتطلب حوسبة المعجم تحديد المنطلقات التأسيسية في اللسانيات العربية، وهي مجموعة المعايير والمقاييس المعتمدة لدى اللسانيين القدامى والمحدثين، والتخفيف من تعددية الرؤى بالتواضع على هذه القواعد الكلية المستمدة من استقراء اللغة في مصادرها الطبيعية: القرآن والحديث النبوي والشعر والنثر[3]...

ثانياً: الآليات الحوسبية:

اعتمدت الدكتورة فاتن الخولي عدة إجراءات تنظيمية تتمثل في[4]:

– **إعداد قاعدة عينات نصية الكترونية:** مختارة للاستخدام اللغوي، وفق المستوى اللغوي والمصادر التي سبق الحديث عنها في آلية جمع اللسانية، يستفاد منها في تحديد المداخل ومعانيها والإمداد بالشواهد والتعبيرات السياقية، وكذا المصطلحات على اختلاف مجالاتها. يستخدم المحلل الصرفي عادة في تحليل تلك العينات النصية لاستخلاص المفردات وجذورها

1– يعرف التحرير المعجمي بالتصميم الداخلي للوحدة المعجمية، الذي يقدم معلومات شكلية ودلالية مفصلة عن رأس المدخل. لمزيد التفصيل ينظر: اللغة العربية والحاسوب، صناعة المعجم الحديث، عبد العاطي إبراهيم: المحيط معجم اللغة العربية، في ضوء تقنيات المعجمية الحديثة، رسالة ماجستير، آداب القاهرة، 2004.

2– فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، سابق ص 148.

3– عبد الله أبو هيف، حوسبة المعجم العربي ومشكلاته اللغوية والتقنية نموذجاً، مجلة التراث العربي، اتحاد الكتاب العرب، دمشق،العدد94/93، السنة الرابعة والعشرون،2004.

4– فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، سابق ص 149.

وصيغها الصرفية للشروع في إعداد قوائم المفردات والمعاني والتعبيرات السياقية والمتلازمات اللفظية، وإحصاء الأكثر شيوعا في كل من تلك المعطيات، كما يفضل استخدام المصنف الآلي لتصنيف تلك النصوص إلى حقول دلالية تساعد على استخلاص المعاني بيسر وسهولة وفقاً لأسبقية الكلمات الواردة في كل حقل على حدا[1].

– **تنظيم قاعدة معارف**[2]:التي يتم اكتسابها من مصادرها الأولية كالمدونات اللسانية والقرارات المجمعية أو من خلال لقاءات مهندسي المعرفة مع المعجميين لاستخلاص حصيلة خبراتهم بصورة منهجية منطقية يمكن تخزينها ومداومة تحديثها، ويتم تمثيل هذه المعرفة بواسطة طرق ونظريات تمثيل المعرفة، مثل: الشبكات الدلالية[3]، القواعد الشرطية[4] ...الخ، تكمن فائدة هذه القاعدة في مرحلة تحديد المعجم لاسيما في الامداد بقواعد صك الكلمات[5]: الاشتقاق والتركيب والنحت والتعريب، وإجراءات تنميط المصطلحات للتعرف الآلي عليها، كذلك آليا تحديد المتلازمات اللفظية والتعبيرات السياقية المستحدثة[6].

– **تنظيم قاعدة البيانات** المعجمية طبقا لقواعد معينة في وسائط تخزينية، باستخدام إحدى نظم قواعد البيانات المعروفة:ORACLE، SQL، للاستفادة من هذه البيانات في عدد من التطبيقات المختلفة، وفي مرحلة التحديث كذلك.

وتتهيكل قادة البيانات المعجمية في عدد من الجداول، يتكون كل جدول من مجموعة من التسجيلات (السجلات)، وكل تسجيلة تحتوي على عدد من الحقول، يمكن توضيح ذلك على النحو التالي[7]:

– **رقم التعريف** (ID): يمثل عادة حرفا هجائيا إن كان حجم البيانات محدود يمكن حصره في عدد الحروف الهجائية، أو عددا رقميا إن كان حجم البيانات كبير، ويكون التسلسل هنا تلقائيا ضمانا لعدم حدوث تكرار.

1– المرجع السابق، ص. 519، وللمزيد من بسط القول ينظر : د. رأفت الكمار: الحاسوب وميكنة اللغة العربية، دار الكتب العلمية، القاهرة، ط1، 2006، 158.

2– عز الدين غازي، مفهوم قواعد المعارف، ، الحوار المتمدن، ع 1649، أغسطس – 2006، ص158، وينظر:رأفت الكمار، الحاسوب وميكنة اللغة العربية، دار الكتب العلمية،القاهرة، ط1، 2006، ص. 164.

3 – les réseaux sémantiques.

4 – les regles conditionelles.

5– تعد قاعدة المعارف هذه لب النظام الحوسبي للمعجم، يمكن أن تتعد لتشمل قواعداً للمعارف الدلالية وأخرى للسياقية و ثالثة للصرفية ورابعة لنصوص لغوية... وهكذا، مما يتطلب جهدا متضافرا يبذل من قبل كل من اللسانيين الحوسبيين ومهندسي المعرفة.

6– فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، سابق ص 150.

7– المرجع نفسه، الصفحة نفسها.

- **التسجيل** (Record): يصاغ من مجموعة من الحقول، تنشأ على أساسها بيانات تعطي في مجملها المعلومات المراد تخزينها كاملة: الجذور والمداخل والمعاني...

جذر	كلمة	معنى	شاهد

- **الحقل** (Field): عبارة عن وحدة بناء السجل، يتنوع ما بين العددية والحرفية، ويضم معلومات محددة حسب الغرض من التخزين؛ حقل: الجذر/الكلمة/المعاني/الشواهد/ التعبيرات...

جذر
كلمة
معنى

أنموذج إيضاحي (افتراضي):

ID	كلمة
1	كرب
1	كربة

الجدول 1

ID	جذر
1	ك ر ب
2	ك ر م

الجدول 2

ID	كلمة	معنى	شاهد	تعبير سياقي
1	كرب	الحزن والغم	فرج الله كربه	
1	كربة	كرب	من فرج عن مؤمن كربه من كرب الدنيا فرج الله عنه كربة من كرب يوم القيامة	

الجدول 3

لقد استخلصت الدكتورة فاتن الخولي أن الجدول الأول تعددت فيه الكلمات وتوحد الجذر، لذا وجب وجود رقم التعريف في الجدول الثاني مع الجذر حتى يتسنى تحويل الجذور في الجدول الثاني إلى أرقام في مرحلة تنظيم شبكة العلاقات التي تربط بين الجداول الثلاثة، حيث تكون العلاقة بين ج 2 وج1 علاقة واحدة إلى متعدد، في حين تكون العلاقة بين ج1 وج3 علاقة واحد إلى واحد[1].

- **تنظيم شبكة العلاقات:** من خلال تحديد العلاقات التي تربط بين مفردات المعجم: كالعلاقة بين الجذور والكلمات، والصيغ الصرفية خاصتها، أو العلاقات الدلالية كعلاقة الترادف والاشتراك اللفظي والتضاد أو الحقول الدلالية (مصطلحات طبية/زراعية...)، توفر تلك الشبكة من

1- المرجع السابق، ص 152.

العلاقات: تحسين عملية البحث، ضغط البيانات وتشفيرها، كذلك تسهم في استخلاص معاجم فرعية من المعجم الأساس كمعاجم: التعبيرات السياقية، والمتلازمات اللفظية، والترادف، والتضاد، وألفاظ الحضارة...[1]

- **تنظيم قائمة المفردات** الخاصة بالمداخل والمواد المعجمية المختلفة المختارة من قاعدة عينات النصوص، مرتبة ترتيبا ألفبائياً،وتحل كل مفردة رقما محددا ليتم التعرف عليه آليا، ويستفاد منها في استخلاص الإحصائيات المعجمية المختلفة سواء للأفعال أم الأسماء والصفات والجموع، كذلك إعداد فهارس متنوعة لكل معطى معجمي...[2]

- **توفير آلية تحديث** للمعجم تمكن من حذف بعض المعلومات الواردة فيه وأُهمل، هذا ويتم التحديث[3] من خلال[4]:

تحديث قاعدة العينات النصية.

- تحديث قائمة المفردات المعجمية بإضافة المستحدث وحذف المهجور.

- تحديث المعاني والشواهد والتعبيرات السياقية والمتلازمات اللفظية وغيرها من المعطيات المعجمية.

- تحديث شبكة العلاقات التي تربط بين المفردات وصيغها ودلالاتها...

- تحديث وحدة الخدمات المعجمية، بالإضافة أو الحذف حسب المستجدات المعجمية والتقنية.

- **إعداد وحدة خدمات معجمية**، تيسر للمستخدم عملية استرجاع المعلومات بطرق البحث المتعارف عليها (بالجذر والمشتق والمطابق والمترادف...)، وتوفر كذلك للمستخدم المتخصص خيارات عرض معلومات عن: أصل المفردة، وتاريخها ومجالها الدلالي، والتعبيرات السياقية خاصتها ومتلازماتها اللفظية، والشواهد الواردة لها، كما يمكن تزويده بإحصائيات الشيوع والتكرار وعدد المشتقات... وغيرها من الخدمات التي تجذب المستخدم وتوفر له احتياجاته المعجمية المختلفة في يسر وبسرعة مرضية[5].

1- فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، سابق ص 152.

2- المرجع نفسه، الموضع نفسه.

3- عد الدكتور نبيل علي هذه الآلية لب المنظومة المعجمية حيث قال: "عملية تحديث المعجم هي قلب منظومته، وهي التي تضفي على هذه المنظومة عنصري الدينامية والتماسك، إنها العملية التي تربط بين العمليات المعجمية الأخرى لتصل بين الأطراف المختلفة للنشا المعجمي، بين مصادر الحاجة لنشأة المصطلحات الجديدة وآليات تكوين المفردات والتعابير الاصطلاحية، بين صك الجديد ولفظ المهجور، وبين ما يتضمنه المعجم وما يستخدم فعلا من قبل الجماعة اللغوية.للمزيد من التفصيل ينظر : نبيل علي، اللغة العربية والحاسوب، سابق، ص 498.

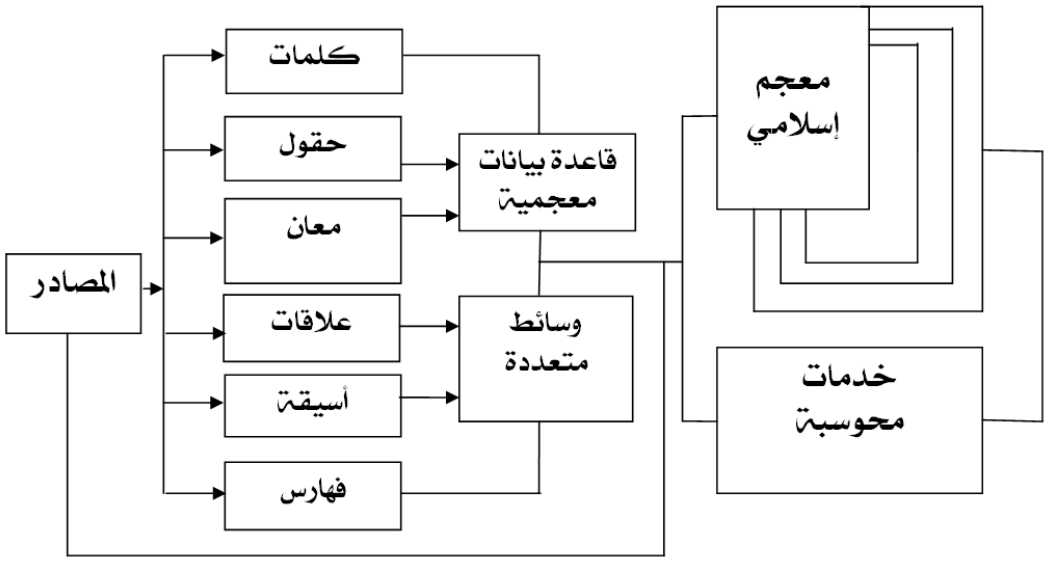
4- فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، سابق ص 153.

5- المرجع السابق، ص 153.

12ـ جهود حرف في مجال المعاجم الإلكترونية:

سعت حرف[1] للإسهام في تطوير الجهود الرامية إلى تدعيم المسلمين غير الناطقين بالعربية علميا وكذا مساعدة العاملين في مجال الترجمة الإسلامية من خلال بناء معجم/ قاموس إلكتروني متخصص يحوي المصطلحات الإسلامية في حقولها الدلالية المختلفة استجابة لمتطلبات المحيطين الثقافي والعملي،وأطلقت عليه اسما تجاريا هو القاموس الإسلامي[2].

القاموس الإسلامي: Dictionnaire Islamique:



شكل يوضح دورة بناء القاموس الإسلامي

1- تأسست شركة حرف في عام 1997 كامتداد لمركز التراث الإسلامي الذي مارس نشاطه منذ عام1985 ، وظلت أنشطة "حرف" وخدماتها تتسع وتتنوّع فاتخذت مقرًا رئيسيًا لها في الرياض وفرعاً للتطوير في القاهرة، وامتدت شبكة وكلائها وممثليها لتغطي منتجاتها وخدماتها العالم العربي والإسلامي والاتحاد الأوربي وجنوب شرق آسيا وأفريقيا وأستراليا وأمريكا.

قامت حرف بتطوير المصادر الإسلامية الكبرى على وسائل إلكترونية مثل: القرآن الكريم مع أشهر تفاسيره. وأكثر من 62000 حديث نبوي شريف بشروحها وكافة الدراسات المتعلقة بها. كما تراث الفقه الإسلامي أصبح متاحا على قاعدة بيانات إلكترونية تعدل حوالي 750000 صفحة مطبوعة بكافة الدراسات والتحليلات ذات الصلة. بالإضافة إلى العديد من التطبيقات الأخرى مثل: السيرة النبوية، والتاريخ الإسلامي، والقاموس الإسلامي، وفقه الصلاة، وفقه المواريث، وفقه المعاملات، والفتاوى الاقتصادية، والحج والعمرة، وغير ذلك.

كما قامت حرف باستخدام تقنيات المالتيميديا في كثير من برامجها، ومن أهمها البرامج الموجهة للطفل المسلم، حيث أنتجت سلسلة قصص الأنبياء التي نالت إعجاب كثير من الجهات المتخصصة ونالت الجوائز، بالإضافة إلى مجموعة قصص رشيد: رشيد في عالم القرآن، رشيد والكنز، رشيد والساعة، رشيد والمكتبة، وغير ذلك. وقد دخلت حرف مجال التعليم الإلكتروني عن بعد بقوة عندما تراءى لها تنامي الاهتمام به عالميا، وأنه سيصبح رافدا أساسيا من روافد التعليم في المستقبل، فأرادت أن توفر تقنيات التعليم الإلكتروني للمؤسسات التدريبية والتعليمية العربية، لتكون مواكبة للتطورات العالمية في هذا المجال، وذلك من خلال نظام تدراس للتعليم الالكتروني وما أنتج من خلاله من برمجيات مثل فصول وسين وخبير. ومن منطلق حرصها الدائب على مد خدماتها لتشمل المستخدمين غير العرب، فقد ترجمت كثير من برامج حرف وتطبيقاتها للغات الإنجليزية والفرنسية والألمانية والتركية والإندونيسية والماليزية بالإضافة إلى لغات أخرى ما زالت قيد التوسع كالإسبانية والأوردية والهوسا. وتقوم حرف بتطوير برامجها لمختلف البيئات الحاسوبية، مما يجعلها متوافقة مع كافة وسائل النشر. كما تنشر العديد من منتجاتها على الإنترنت بلغات عديدة، وبإمكانات بحث فائقة، والذي يتيح لزائري الإنترنت الاستخدام المجاني لكثير من برامج الشركة، كما أنه يعطي معلومات عن منتجات وتقنيات الشركة واتجاهاتها المستقبلية. . للمزيد من التفصيل ينظر الموقع الرسمي لشركة حرف : شركة حرف لتقنية المعلومات،

موقع الويب: http://www.harf.com/cms.aspx?ContentID=109

2- فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، سابق ص 154.

يعد القاموس الإسلامي لشركة حرف أول معجم إلكتروني إسلامي متخصص يضم مجموعة كبيرة من المصطلحات التي تغطي جميع فروع المعارف والعلوم الإسلامية وشروحها مع مقابلاتها في لغات أربع: الإنجليزية والإندونيسية والماليزية، ثم الفرنسية في الإصدار الثاني فجمع بذلك بين خصائص المعجم المختص والقاموس متعدد اللغات وقد أطلق عليه تجاريا القاموس الإسلامي[1].

انتهجت حرف منهجا علميا[2] في بناء معجمها وفقا للآليات اللسانية وحوسبية[3]، واقتصر **تحرير المداخل** على تقديم معلومات صوتية تتعلق بالهجاء والضبط، وأخرى دلالية تخص شرح المصطلح فقط دون الحاجة إلى غيرها من المعلومات اللسانية الأخرى كالمعلومات الصرفية والنحوية[4].

هكذا تم بناء المعجم الإسلامي لسانيا، ثم شرع في بنائه حوسبياً على النحو التالي مع مراعاة تداخل العمل اللساني مع الحوسبي في بعض مراحل البناء الكلي لاسيما فيما يتعلق بإعداد النصوص والتشكيل والتدقيق الآليين، ويمكن توضحيحه من خلال الشكل الآتي[5]:

1- المرجع السابق، الصفحة نفسها.

2- قام فريق العمل اللغوي بتحديد المصادر المتخصصة لجمع المعطيات المعجمية المبتغاة، فتنوعت تلك المصادر ما بين تراثية ومعاصرة (أسد الغابة في تمييز الصحابة، الإصابة في تمييز الصحابة، الأعلام للزركلي، برنامج التاريخ الإسلامي، تاريخ الإسلام، معجم المنجد في اللغة والأعلام ، معجم المصطلحات الإسلامية للخضري...) ثم اختيار المداخل وفقا لطبيعة المعجم الاصطلاحية فكانت المداخل اسمية إلا ما ندر وصنفت حقولا دلالية وفقا للمجالات الإسلامية المتعددة، ثم رتبت ترتيباً ألفبائياً تمهيدا لتحريرها.

3- وتتميز الشركة في بناء وتطوير برامجها بما يلي:

- النظم الآلية في كل مرحلة من مراحل عمليات التطوير بدءاً من القراءة الآلية وانتهاءً باختبارات ما قبل الإصدار وتوكيد الجودة.
- يستخدم المطورون أحدث لغات البرمجة، كما يستخدمون أحدث التقنيات لتطبيقات الإنترنت. (Visual C++\ Java\ ASP.Net\ C#\ Visual Basic\ SQLServer\ Oracle\ Macromedia Flash Director)
- اتصال فريق المطورين والمتخصصين الفنيين بالشبكة الدولية عن طريق الإنترنت والبريد الإلكتروني للاطلاع على كل جديد في المجال التقني لتطوير أدواتهم.
- امتلاكها لمكتبة إسلامية إلكترونية ضخمة تكونت من خلال خبرتها الطويلة في إنتاج البرامج الإسلامية الموسوعية.
- اعتمادها على تقنية التحليل الصرفي في معالجة وفهرسة البيانات باللغة العربية لضمان دقة وسرعة البحث والاسترجاع.
- تعتمد أقسام الترجمة والنشر على قاعدة بيانات ضخمة وبرامج الترجمة القياسية. بالإضافة إلى تطوير الأدوات المساعدة لإضفاء سمة الأتوماتيكية على عمليات الترجمة والنشر.
- تتصل الشركة بشبكة محلية وبريد داخلي بغرض تحسين التعاون والتفاعل بين الأقسام المختلفة.

وقد حصلت الشركة على العديد من الجوائز العربية والعالمية. ثم توجت هذا الالتزام بحصولها على شهادة المطابقة لمواصفات الجودة العالمية ISO 9001/2000 في عام 2001 و CMMI المستوى الثالث في عام 2006. للمزيد من التفصيل ينظر الموقع الرسمي لشركة حرف : شركة حرف لتقنية المعلومات، موقع الويب: http://www.harf.com/cms.aspx?ContentID=109

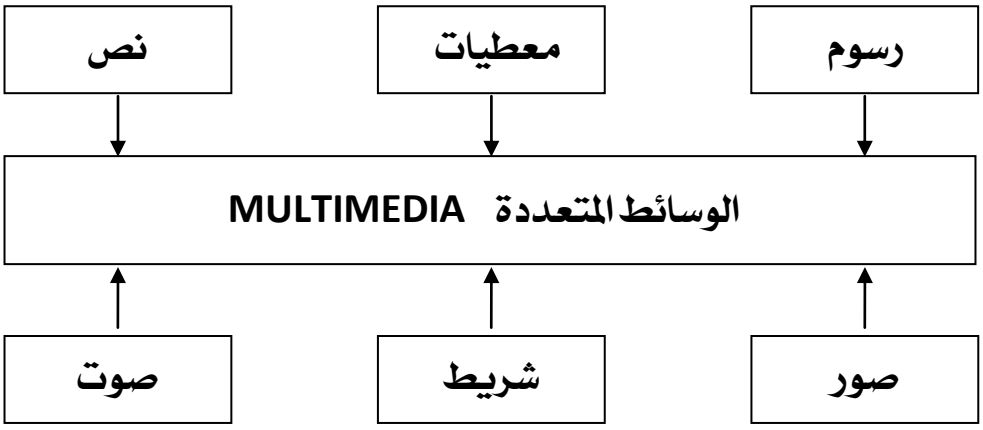
4- كانت مداخل معجم الأساسية باللغة العربية ثم وضعت مقابلاتها في اللغات الأربع مع شروحها المترجمة، وقام فريق عمل متخصص بإعداد الكتابة الصوتية الخاصة بكل مصطلح على حدا تيسيرا للمستخدم غير الناطق بالعربية حتى يتسنى له التعرف على طريقة النطق الصحيحة.

5- فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، سابق ص 156.

رسم توضيحي لبناء القاموس الإسلامي حوسبياً

لبناء هذا المعجم حوسبياً تم إعداد **قاعدة عينات نصية** من مصادر إسلامية متنوعة المجالات لاستخلاص المعطيات الإسلامية الخاصة به، وقد تنوعت تلك المعلومات ما بين مصطلحات، أحداث تاريخية، شخصيات إسلامية، بعد ذلك تم **إعداد قوائم بالمصطلحات الإسلامية** التي شملت أكثر من 55000 مصطلح ومقابلاتها في الإنجليزية والإندونيسية والماليزية، ثم الفرنسية في الإصدار الثاني، **وعرض للكتابة الصوتية**، مع توفير شروح للمصطلحات الأساسية منها، وتصنيفها إلى حقول دلالية تشمل مختلف المعارف الإسلامية، وذلك بعد مرورها بمرحلتي التدقيق الإملائي والتشكيل الآليين[1].

من ناحية أخرى توفير المصاحبات التصويرية للمعالم الإسلامية التي تضمنها القاموس كذلك لقطات فيديو تصاحب الألفاظ وتبين مدلولاتها في القرآن الكريم، وإعداد تدريبات لغوية وألعاب ترفيهية، كل هذا باستخدام الوسائط المتعددة؛ بهدف تنمية حصيلة المستخدم اللغوية واثراء معلوماته في إطار ترفيهي مشوق[2].



مكونات الوسائط المتعددة للقاموس الإسلامي

1- للمزيد من بسط القول حول التعريف بعمل المدقق الإملائي والتشكيل الآليين، ينظر الباب الثاني من الأطروحة، المعالجة الآلية للغة العربية لبناء قاعدة البيانات المعجمية آليا، الفصل الأول، ص 128 وما بعدها.

2- فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، سابق ص 157.

في هذا الإطار تم **إعداد قاعدة بيانات للقاموس الإسلامي** باستخدام SQL Server ولغة VC البرمجية، وبناء شبكة علاقات تربط بين المعطيات المعجمية بعضها البعض وبينها وبين معطيات الوسائط المتعددة المتعلقة بها من جانب آخر، مما وفر إمكانية إعداد فهرسة كاملة لألفاظ القرآن الكريم (150.000 لفظ ومعناه) مع ربطها بأسيقتها القرآنية، وعرض لأسماء الله الحسنى وأسماء الأنبياء والقيامة والجنة والنار التي وردت بالقرآن الكريم، كذلك بناء شجرة للموضوعات الاقتصادية التي ترتبط موضوعاتها بآيات في القرآن الكريم، وتم تتويج هذا العمل بإعداد وحدة خدمات معجمية تعتمد طرق استرجاع معلوماتي متداول قائم في الأساس على التحليل الصرفي[1].

اعتمدت حرف عدة أدوات حوسبية في بناء المعجم الإسلامي[2]، يكشف عنها الجدول التالي[3]:

الأداة	الهدف
المحلل الصرفي الآلي MMMP	تحليل المداخل والمواد المعجمية صرفيا
المدقق الإملائي العربي SpellingChecker	تصويب أخطاء النصوص المنتقاة
نظام التشكيل التلقائي	تشكيل المداخل وشروحها آليا،مع التدقيق يدويا
قاعدة البيانات المعجمية LDB	إدارة بيانات بناء المعجم وشبكة العلاقات
Photoshop	تنزيل الصور
VideoPremisere	تنزيل مقاطع الفيديو

أدوات بناء المعجم الإسلامي

جاء تطوير حرف لتقنية البحث واسترجاع المعلومات فيه من خلال تحسين ربط الكلمات المتلبسة شكليا بجذورها الصحيحة، فلفظة "قائلون" في سورة الأعراف/ 4 ترجع إلى الجذر (ق ي ل)، في حين لفظة "قائل" في سورة يوسف/ 56 ترجع إلى الجذر (ق ول) كما في الشكلين التاليين[4]:

قائل- ق و ل

هذا ولم تتوقف جهود حرف المعجمية عند هذا الحد بل تعدت ذلك إلى الدخول في مجال تحويل المعاجم الورقية التراثية إلى معاجم إلكترونية تيسيرا على المستخدم دون عناء التنقيب بين الأوراق.

1- المرجع السابق، ص 158،157.

2- يعد القاموس الإسلامي من إنتاج شركة حرف في مجال المعجم المحسوب، ثم قام بتطوير آلياتها وأدواتها على الصعيد اللساني والحواسبي وضمنت ذلك في معجمها لألفاظ القرآن الكريم، وطورت فيه من تقانتها في مجال البحث في القرآن الكريم.

3- فاتن الخولي، المعجم بين اللغة والحاسوب "بناء المعجم الإسلامي من منظور لساني حاسوبي "، سابق ص 158.

4- المرجع السابق، ص 159.

خلاصة

كانت لتقانة المعلومات دورها الفاعل في تطوير وبناء معاجم إلكترونية أدى إلى النهوض بالعلوم التطبيقية، والعلوم اللسانية خاصة في مثل هذه المشاريع ، حيث تلتقي الأخيرة مع الحاسوبية على ساحة المعجم خاصة، فيتبدى موضوعا مثيرا للمعالجة الآلية، من خلال احتوائه على"ثالوث (البيانات والمعلومات والمعارف)، البيانات ممثلة في قوائم مفرداته المطولة، والمعلومات التي تنبث من شبكة العلاقة الكثيفة التي تربط بين هذه المفردات والعدد الهائل من التحليلات والمؤشرات والاحصائيات المترتبة على هذه العلاقات، والمعارف ويقصد بها حصيلة المفاهيم البسيطة والمركبة التي تتضمنها تعريفات المعجم والتي تحدد به مدلولات مفرداته، وفي ضوء هذه التحديات عالجت في هذا الفصل أساسيات بناء المعجم الإلكتروني في اللغات الطبيعية، وعلاقته بقاعدة المعطيات الآلية، والتطبيقات الحاسوبية المنجزة في بعض الأقطار، مركزة على بعض التجارب الغربية والعربية، وأخص بالذكر أعمال الأستاذ الدكتور عبد الغني أبو العزم من خلال معجم الغني، وأعمال الدكتور الحناش محمد، وأعمال الدكتور نبيل علي، وأعمال الدكتور الحاج صالح، كما اقتصرت على تجربة غربية فرنسية في مختبر المعالجة الآلية والتوثيق اللساني (LADL) باعتباره تجربة رائدة في ميدان المعالجة الحاسوبية للغات الطبيعية، تمكنت من إنجاز معاجم إلكترونية للغة الفرنسية.

الفصل الثاني

نحو بناء معجم إلكتروني للمعالجة الآلية للغة العربية باستخدام نظام آلي لاستخراج جذور الكلمات العربية

مفردات ألفاظ القرآن الكريم "أنموذجاً"

أولاً - إعداد تصور منهجي للمعجم الإلكتروني اللغة العربية

ثانياً- عرض معجم إلكتروني للمعالجة الآلية للغة العربية

ثالثاً- المعالجة الآلية للمدونة: *مفردات ألفاظ القرآن الكريم*

رابعاً- الدراسة التطبيقية للمشروع

خامساً- الدراسات السابقة

سادساً- آلية التنفيذ : التطبيق

سابعاً- نتائج الدراسة التطبيقية

أولا: إعداد تصور منهجي للمعجم الإلكتروني اللغة العربية:

لاشكّ في أنّ المعجم الإلكتروني المنشود يجب أن يكون شاملا ووافيا ودقيقا، وأن يتجنب ما يؤخذ على المعجمات الحالية من مآخذ وعيوب، فالجهود المبذولة للنهوض بهذا المعجم الإلكتروني ينبغي أن تكون في معظمها موجهة نحو الجانب اللغوي منه، ذلك أنّ التقدم الهائل في مجال المعلوماتية كفيل بتسهيل العمل وتيسيره لتحقيق معجم إلكتروني متطور، إذا ما توفرت المادة اللغوية الدقيقة والمتكاملة.

فبعد تطرقنا في فصول هذا البحث إلى ماهية المعجم الإلكتروني و أهميته وفوائده، سننتقل إلى الحديث في هذا الفصل إلى الخطوات المنهجية نحو إعداد معجم إلكتروني في خطة عملية تمكننا من أخذ صورة شاملة حول كيفية إنشاء قاعدة لغوية انطلاقا من قواعد مضبوطة وصارمة. فليس من السهل على الباحث أو أي إنسان عادي أن يتحدث عن المعجم الإلكتروني للغة العربية، ومردّ ذلك أنّنا لسنا بصدد الحديث عن معجم عادي، فنحن أمام معجم من طبيعة أخرى وتداول آخر لمفهوم المعجم في حدّ ذاته.

في ضوء هذه الحقيقة البسيطة، يمكن القول بأنّه لا يمكن إدراك قيمة المعجم الإلكتروني للغة العربية المنشود إخراجه، إلاّ باستحضار المدونة اللغوية للمعجم، وهو مدونة خصبة جدا فعالة لما تحتوي عليه من قاعدة بيانات شاملة تجعل من المعجم الإلكتروني للغة العربية بوابة التطور في مجال الصناعة المعجمية الحديثة.

ثانيا : عرض معجم إلكتروني للمعالجة الآلية للغة العربية باستخدام نظام آلي لاستخراج جذور الكلمات العربية : مفردات ألفاظ القرآن الكريم "أنموذجا ":

يهدف هذا المبحث إلى تقديم عرض معجم إلكتروني للمعالجة الآلية للغة العربية "مفردات ألفاظ القرآن الكريم أنموذجا"، نقدّم من خلال هذه الدراسة وجهة نظر تتعلّق ببناء قاعدة المعطيات الموجهة أساسا لبناء المعجم الإلكتروني وذلك باستخدام تقنيات معلوماتية حديثة لمعالجة اللغة العربية آليا في المستوى المعجمي مقرونا بالمستوى الصرفي الذي كان أساس هذا العمل التطبيقي.

تبرز التحديات في هذا العمل التطبيقي في كيفية بناء نظام آلي لاستخراج جذور الكلمات العربية بمثابة قاعدة أساسية للمعجم المنشود.

وعليه نحاول من خلال هذه الدراسة العملية استغلال و استثمار كل الموارد اللسانية والحاسوبية، وإلى كل ما تعرضنا إليه من خلال الدراسة النظرية لعرض طريقة فعّالة لبناء واستغلال معجم إلكتروني مميّز للغة العربية.

إنّ هذا العمل المنجز يعتمد أساسا على خصائص اللغة العربية الّتي تتميز تميزا واضحا عن اللغات الأجنبية[1].

فالمعاجم الأجنبية إن لم نقل كلّها تعتمد في تنظيم مصطلحاتها على الترتيب الأبجدي لكل الكلمات دون استثناء سواء كانت أفعالا أو أسماء أو حروفا، بينما اللغة العربية فإنّ الأمر مختلف، إذ أنّ جل معاجمنا تعتمد على مفهوم "جذر الكلمة" وكل جذر يميز مجموعة من الكلمات تشترك عادة كلّها في الحروف الثلاثة لهذا الجذر وفق تراكيب خاصة تسمى "التفعيلات".

انطلاقا من هذه الملاحظة، عمدنا إلى استغلال هذه الخصائص بإنشاء معجم مكوّن من مجموعات مختلفة من الكلمات ذات الجذر المشترك يتمّ بلوغها أثناء عملية البحث بإيجاد الجذر أوّلا، ومن ثمّ إيجاد الكلمة داخل هذه المجموعة، وبالتالي فإنّ كل الجذور تكون مرتبة ترتيبا أبجديا عن المعجم بينما تكون الكلمات مرتبة ترتيبا أبجديا داخل المجموعة.

ملاحظة:

سنعتمد في هذه الدراسة التطبيقية على بعض المصطلحات التي يمكن استخدامها في هذا العمل من خلال استثمار واستغلال كل ما ذكرناه من مفاهيم وتعريفات لهذه المصطلحات في الدراسة النظرية وتوظيفها كل حسب مجاله، من باب الأهمية العلمية للمصطلح وكيفية استغلاله في الدراسة التطبيقية، هذه من جهة، ومن جهة أخرى، حتى يكون هناك تكامل بين الدراسة النظرية والتطبيقية.

تركزت الدراسة في مجالها اللساني والحاسوبي على بعض المصطلحات العلمية باعتبارها كلمات مفاتيح[2] لهذا العمل التطبيقي:

المعالجة الآلية للغة العربية[3]، المعجم الإلكتروني[4]، استخراج الجذر[5]، تحليل صرفي[6]، الوزن[7]، السوابق واللواحق[8]، مفردات القرآن الكريم.

1- للمزيد من التفصيل في هذا الموضوع، ينظر : محمد المبارك، فقه اللغة العربية وخصائص العربية، مرجع سابق، وينظر : أنيس فريحة، نظريات في اللغة، دار الكتاب اللبنانب، بيروت، لبنان، الطبعة الثانية، 1981، ص 68 وما بعدها.

2- Mots Clés.
3- Traitement automatique de la langue arabe (TALN Arabe).
4- Dictionnaire automatique.
5- Extrait de racine (Lemmatisation).
6- Analyse morphologique.
7- Le Schème.
8- Préfixes et Suffixes.

1- موضوع الدراسة التطبيقية للمشروع:

بدأنا نشهد منذ بداية عقد الثمانينات أبحاث لتقنيات حاسوبيه تهدف إلى معالجة اللغة العربية ذاتها, حيث تمتاز اللغة العربية لخصائص عديدة, منها طبيعة اللغة الثرية المرتفعة المستوى بل إنها تعتبر من فئة اللغات الطبيعية الأكثر تعقيدا والأغنى صرفيا. ويتجلى غناها الصرفي في إمكانية بناء عدة مئات من الكلمات وبمعاني مختلفة من جذر واحد وذلك باستخدام الموازين واللواحق الصرفية المختلفة. كما أن خصائص أي نظام طبيعي يراد بناؤه ستؤثر بشكل كبير على طريقة تصميمه و تنفيذه آليا.

إن الكلمة في اللغة العربية تشترك كلها في عدد معين من الحروف يكون عادة ثلاثة يسمى الجذر، فالجذر كما ذكرنا سابقا يتكون من الحروف الأصلية للوحدة المعجمية[1]، مكونة من ثلاثة حروف أساسية وهي "الفاء" التي ترمز إلى الحرف الأول منها، "العين" وترمز إلى الحرف الثاني، وأخيرا "الآم"، فترمز إلى الحرف الأخير وبذلك تكون الكلمة الرمز هي "فعل".

فعل ← النموذج الذي يجب أن تطابقه جميع الجذور (سجد، عهد، بعث ...).

إن المعجم العربي يعتمد على ترتيب هذه الجذور التي يشير كل واحد منها إلى مجموعته ويتغير بالتالي أسلوب البحث عن كلمة ما باستخراج جذرها ثم البحث عنه في المعجم، وبعد ذلك إيجاد هذا الجذر، يتحول العمل إلى عملية البحث عن الكلمة في مجال أضيق وهو مجال مجموعة الكلمات التي يمكن تكوينها من خلال هذا الجذر[2].

2- آليات الدراسة:

يقترح المشروع بناء نظام آلي بمثابة معجم إلكتروني للمعالجة الآلية للغة العربية، يعتمد هذا النظام طريقة لاستخراج جذور الكلمات العربية بالاعتماد على التقنيات اللغوية للمعالجة الآلية للغة العربية[3]. ترتكز هذه الطريقة على مفهوم الوزن أو الميزان العرفي "أحد نقاط القوة في الصرف العربي"[4].

1- ينظر الباب الثاني، الفصل الأول (المعالجة الآلية للصرف العربي) من البحث، ص 170 .

2- عند تحليل أو توليد الكلمات في اللغة العربية يمكن الاستفادة من عدة معطيات أهمها: المعجم حيث يتم عن طريقة معرفة جذور وجذوع الكلمات العربية. كما يمكن الاستفادة من المعطيات النحوية للكلمة داخل الجملة كحالتها الإعرابية و غير ذلك. وأخيرا يستفاد من السياق أو الدلالة الذي بدوره يؤثر على اختيار المعنى الصرفي أو الوظيفي المراد صياغة الكلمة في قالبه.

3 - TALN.

4- يعد علم الصرف أساسا متينا في التعامل مع اللغة العربيةُ أما حوسبة الصرف فتعد مطلبا أساسيا لكثير من التطبيقات اللغوية التي تعتمد على النظم الاليه. يبحث علم الصرف في التغييرات التي تطرأ على بناء الكلمات وصورها المختلفة من الداخل و أسس ترتيب عناصرها و ظواهر التفاعل بين هذه العناصر. ويتيح الصرف معرفة مكان الحرف الأصلي, ومكان الحرف الزائد في الكلمة, بالإضافه إلى الاهتداء إلى أصلها الاشتقاقي أو أصلها المعجمي. و للصرف ثمراته التي من أهمها صون اللسان عن الخطأ في المفردات و مراعاة قانون اللغة في الكتابة.

الجانب الإيجابي في هذه الطريقة أنها لا تستعمل معاجم الكلمات بل تقوم بالتعرف على مختلف الكلمات المدخلة بطريقة ذكية عن طريق استخراج الجذر والوزن المناسبين لكل كلمة محللة.

[الكلمة (ن)] ← [خوارزم (ن)] ← جذر (ن)

انطلاقا من هذا التصور النظري، هل يمكننا بناء نظام آلي لاستخراج جذور الكلمات العربية يكون بمثابة معجم إلكتروني للمعالجة الآلية للغة العربية يمكن تطويره فيما بعد؟

2.1 ـ مفردات ألفاظ القرآن الكريم كعينة في الدراسة التطبيقية:

لابد للمعجم الإلكتروني للمعالجة الآلية للغة العربية أن يحتوي مفردات جامعة ونصوصا رصينة[1] وليس هناك ما يحقق هاتين الخاصيتين أفضل من نصوص القرآن الكريم (الكلمات المدخلة)، هذا بالإضافة إلى أن الكلمات والتعابير الأكثر استعمالا في اللغة العربية واردة في القرآن الكريم في غالبيتها العظمى[2].

إن عدد الكلمات التي أوردناها ضمن قادرة البيانات للمعجم في حدود ثلاث آلاف كلمة (3000 كلمة)[3] تغطي جميع المواضيع التي تطرق لها القرآن الكريم[4]، وهو في الحقيقة لم يهمل أي موضوع تصديقا لقوله تعالى: ﴿ **ما فرطنا في الكتاب من شيء**﴾[5].

يحتوي المعجم على:

1- جميع مفردات القرآن الكريم وبالتالي فإن هذا المعجم يتميز بالشمول، دون الاجتهاد بأن هذه الكلمة معروفة أو غير معروفة، سهلة أو صعبة.

2- يحتوي المعجم على قاموس الأوزان الصرفية كقاعدة أساسية لقاعدة المعطيات، وهنا أخص بالذكر أنني استعنتُ في تحديد وإحصاء الأوزان الصرفية **بمعجم الأوزان الصرفية**[6].

3- عدد الجذور الثلاثية في اللغة العربية هو 3377 جذر بنسبة 92%.

ماذا يمكن أن تقدم قواعد بيانات القرآن الكريم للغة العربية؟

إن صغر حجم بيانات القرآن الكريم بالمقارنة مع اللغة العربية جعل بالإمكان إكمال بعض قواعد البيانات بجهد معقول . أما البيانات الكاملة للغة العربية فتحتاج إلى جهود كبيرة . لكن

1- النص القرآني هو المدونة المثالية للغة العربية فهو بمثابة المرجعية الصوتية لنطق هذه الكلمات. وينظر:د شرف الدين الراجحي، د سامي عياد حنا، مبادئ علم اللسانيات الحديث، تقديم أ.د عبده الراجحي،دار المعرفة الجامعية،مصر،دط،2003،ص80.

2- أ.د محمد زكي خضر، قواعد بيانات القرآن الكريم كأساس للمعجم الآلي الموسع للغة العربية

3- عدد الجذور الثلاثية في اللغة العربية هو 3377 جذر بنسبة 92 %

4- السيوطي (المزهر في علوم اللغة العربية وأنواعها) الجزء الأول، سابق، ص 165، وللمزيد من التفصيل ينظر : قواعد بيانات القرآن الكريم كأساس للمعجم الآلي الموسع للغة العربية، الموقع الإلكتروني: http:// www.al-Mishkat.com

5- سورة الأنعام، الآية 38.

6- إميل بديع يعقوب، معجم الأوزان الصرفية، عالم الكتب، الطبعة الأولى، 1413هـ/1993م.

الأمثلة التي عرضت في هذا البحث تعطي إشارة واضحة إلى سهولة الكثير من هذه الجهود وإمكان برمجتها . ويمكن تلخيص ما يمكن أن تقدمه بيانات القرآن الكريم للغة العربية كالآتي :

1- تعطي قواعد بيانات القرآن الكريم أمثلة واضحة للكثير من المعالجات التي يمكن أن تتم على اللغة العربية . وما تحتاجه اللغة العربية سوى توسيع بعض هذه القواعد لكي تشمل كل الكلمات والمعاني المستعملة في اللغة .

2- إن البرامج التي كتبت وفق قواعد معينة لمعالجة مفردات كلمات القرآن الكريم يمكن استخدامها نفسها على اللغة العربية بشكل عام.

3- إن الكثير من الأبحاث على قواعد بيانات مفردات القرآن الكريم لم تؤت ثمارها بعد لحاجتها إلى الكثير من المعلومات اللغوية، لكن تصميم هذه القواعد يعطي فوائد للغة وفي الوقت نفسه يعود بالفائدة على أبحاث القرآن الكريم في بناء برامج حاسوبية تخدم القرآن الكريم خاصة، واللغة العربية بشكل عام.

3.1 ـ المعالجة الآلية للمدونة: (مفردات ألفاظ القرآن الكريم):

إن المعالجة الآلية في تطبيقاتها المختلفة تنظر إلى تحقيق هدفين متزامنين؛ هما: دقةُ النتائج، وتعميمُها على الظاهرة اللغوية كلها، وتتفاوت نسبة تحقق هذين الهدفين باختلاف المستوى اللغوي المستهدف من المعالجة الآلية، ففي المستوى الصوتي والصرفي بلغت نسبة عالية تجاوزت 95%، أما معالجة المستوى الدلالي فلا تزال في بداية الطريق، وذلك لصعوبة الظاهرة اللغوية على المستوى الدلالي[1]، وتعقد النماذج الرياضية المستخدمة، ولذلك فإن الخطأ المتوقع في المعالجة الآلية للغة في هذه المرحلة قد يكون كبيراً في بعض التطبيقات، وذلك من أجل تعميم النتائج على أوسع رقعة من الظاهرة اللغوية.

هذا وإن كانت نسبة الخطأ مقبولة في هذه المرحلة ولاسيما على المستوى الدلالي في المعالجة الآلية لعموم نصوص اللغة العربية ؛ فإنها غير مقبولة في التعامل مع نصوص القرآن

1- التحليل الدلالي من أصعب أنواع التحاليل اللغوية الخاصة باللغات الطبيعية، فهو يتعامل مع دلالة الجمل والتعرف على معانيها وتحليلها تحليلا ذا معنى. وهو يجمع ما بين التحليل الصرفي والنحوي علاوة على تراكيب الجمل من حيث المعنى والدلالة. ولا تكاد أغلب النظم الآلية تصل إلى هذا المستوى من التحليل حتى في اللغات التي لها باع طويل في اللغويات الحاسوبية مثل الإنجليزية والفرنسية وغيرهما. لأن معنى الكلام يختلف من سياق لآخر.ويمكن القول إن مستويات التحليل للكلام (المكتوب والمنطوق) يبدأ بالتعرف على مكونات الكلمة وهو التحليل الصرفي، ثم يرتقي مستوى التحليل ليتعامل مع تراكيب الكلمات مع بعضها البعض. وذلك عن طريق التحليل النحوي إلى أن يصل التحليل إلى التعرف على معنى الكلام ودلالته وذلك عن طريق التحليل الدلال, وكل مستوى من مستويات التحليل له تطبيقاته وفوائده، وهناك أنواع أخرى من التحليل اللغوي تختلف بحسب الحاجة مثل التحليل المقامي، وتحليل العبارات المركبة. وتركز برمجيات الحاسب الحالية والمستقبلية على محاولة تبسيط طريقة التخاطب مع الحاسب وتقريب ذلك، بحيث يستطيع الإنسان أن يتعامل مع الحاسب بلغته الطبيعية سواء على مستوى لغات البرمجة أو التطبيقات الأخرى.

الكريم، وذلك لأمرين: أحدهما أن القرآن الكريم كلام الله تعالى ولا يجوز في حقه ما يجوز في كلامنا، والآخر هو أنه نص محدود في كلماته وتراكيبه، فهي ثابتة محفوظة لا تتغير زيادة أو نقصاً أو تبدلاً، وإن كانت معانيها لا تخلق على كثرة الرد، فعلى الرغم من أن القرآن الكريم نص لغوي يتصف بما تتصف به اللغة العربية من خصائص صوتية وصرفية وتركيبية ودلالية، فإنه يزيد عليها ما يجعله معجزاً خارجًا عن قدرات البشر في الإبداع اللغوي، وإن كان بإمكانهم فهمه وتبينه، وهذا يعني أن المعالجة الآلية للنص القرآني تختلف في مفهومها وأهدافها عن المعالجة الآلية لعموم نصوص اللغة العربية، إذ يتوجه اهتمامها الأول إلى تحري الدقة ما استطاعت، أما غاية تعميم النتائج على عموم النص القرآني فيمكن بلوغها عن طريق بناء قواعد بيانات شاملة تستوعب جميع مفردات النص القرآني، وهذا ما لا يمكن إدراكه في معالجة عموم نصوص اللغة العربية، وتبقى هذه المحاولات محفوفة بمحاذير، وتحتاج إلى وَضْعِ ضوابطَ دقيقة تُّحَدّدُ سابقاً.

3- فرضية الدراسة:

لقد مكنتنا الدراسة النظرية للمشروع في مقاربته اللسانية الحاسوبية من اتباع خطوات إجرائية لبناء المعجم الإلكتروني للمعالجة الآلية للغة العربية، والرسم التالي يبين الخطوات المتبعة في بناء المعجم الإلكتروني مع التركيز الشديد على المستوى الرئيسي التي اعتمدت عليه هذه المقاربة في البحث اللساني الحاسوبي: "الصرف".

يعتبرّ المحلل الصرفي الآلي أحد أهم تطبيقات المعالجة الآلية للغة العربية وتبرز أهميته في كونه مرجعا مهما للتطبيقات الأخرى، كالمحللات النحوية الآلية، المعاجم الآلية، برامج ضبطِ النصوص وتشكيلها، برامج ترجمة النصوص، برامج التصحيح الاملائي...إلخ.

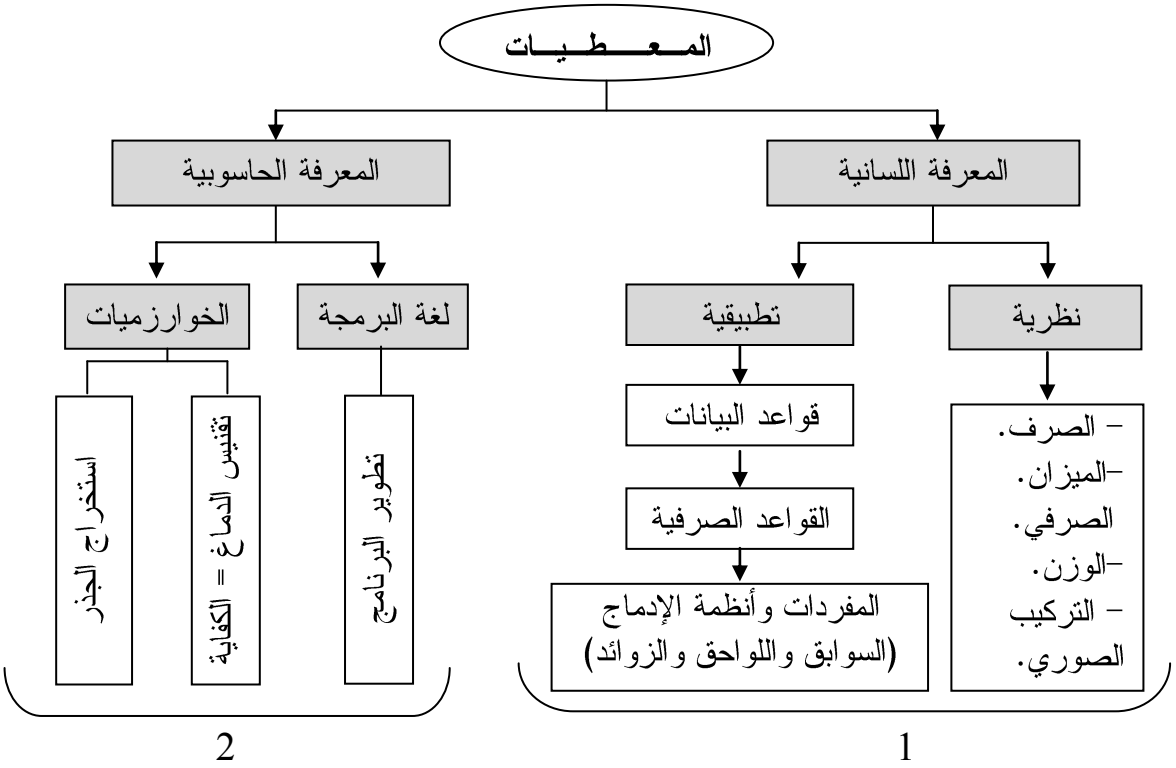
حيث تعد مستخرجاتُ المحلل الصرفي[1] مدخلاتٍ للمحلل النحوي، فعلى سبيل المثال السوابق المتصلة بالكلمة تبين للمحلل النحوي العطف والجر والاستفهام، أما اللواحق المتصلة بنهاية الكلمة فتبين الإضافة والمفعولية، وهذه أهم الأمور التي يعتمدها المحلل النحوي لإبراز مواقع الكلمات داخل التركيب ومحلها الإعرابي[2].

أما فيما يخص المعاجم والكتب الإلكترونية والموسوعات الضخمة، فهو يساعد على فهرستها وتوفير الجهد والوقت على الباحثين والطلاب للوصول إلى المعلومة في يسر وسهولة.

1- هو التحليل الخاص بمستوى الكلمة فقط من حيث تركيبها. والتحليل الصرفي يحلل ويفكك الكلمة إلى أجزائها الصرفية الصغرى (المورفيم) مثل السوابق واللواحق التي تلحق أصل الكلمة وغير ذلك. وهناك الكثير من التطبيقات الخاصة بالتحليل الصرفي مثل تعليم اللغة (خصوصا تركيب الكلمات وتصريفها) , واستخدامه في الترجمة الآلية, وكذلك في محركات البحث واسترجاع المعلومات .

2- بعد أن يتم تحليل الكلمة إلى أجزائها عن طريق المحلل الصرفي. يأتي دور التحليل النحوي (الإعرابي) الذي يتعامل مع الكلمات وفق تراكيبها النحوية. والتحليل النحوي يقوم بتحديد دور كل الكلمة في السياق والتعرف على موقع الكلمة في الجملة (فعل، فاعل، مفعول..). ويفيدنا هذا التحليل في تعليم النحو والإعراب عن طريق الحاسوب.

يستخدم المحلل الصرفي أيضا في مجال تعليم اللغة العربية للناطقين وغير الناطقين بها، فيزيد من فهم الطالب لمادة الصرف، وعن طريق التمارين الموجهة يمكنه تجاوز الصعوبات والإلمام بالمواضيع الصرفية، كاستخراج جذر الكلمة ووزنها وجذعها والسوابق واللواحق المتصلة بها، فهم قواعد الإبدال والإعلال، تصريف الأفعال، إيجاد نوع الكلمات (اسم، فعل، حرف)،...ألخ.



الشكل رقم 06: الخطوات الإجرائية المتبعة لبناء برنامج حاسوبي (المعجم الإلكتروني للمعالجة الآلية للغة العربية)

رابعا- الدراسة التطبيقية للمشروع:

تمت هذه الدراسة وفق المراحل الآتية:

أ- العملية الإحصائية: وبناء قاعدة المعطيات[1] للمعجم الإلكتروني من خلال تحديد الجذور التي تكون ضمن المعجم.

ب- الموارد اللسانية:

- حصر التفعيلات الممكنة.

1 – قاعدة المعطيات: une base de données: هي عبارة عن ملف (كبير) يدوي، تتكون من جزاءات، بطاقات، تقنيات تبديل معطيات منظمة في شكل أبواب / أركان rubrique، تقنيات: حقول: chomps، كل باب / ركن، الهدف منه هو الإحاطة بكل المعلومات التي حدد من أجلها.

- وضع لائحة محصورة وشاملة تتضمن مختلف السوابق وأخرى تحتوي على جميع اللواحق التي يجوز لها أن تلتحق بالكلمات في اللغة العربية.

ج- إنشاء المعجم.

د- صياغة إجراءات الاستعمال: كيفية استخراج الجذر.

هـ- لغة البرمجة المستخدمة.

1.4- العملية الإحصائية وبناء قاعدة المعطيات للمعجم الإلكتروني من خلال تحديد الجذور التي تكون ضمن المعجم:

وقد تمت معالجة النقاط التالية: - شرح كيفية بناء المعجم المنشود مع تقديمه كعينة لهذه الدراسة يشمل مفردات ألفاظ القرآن الكريم، حيث سنحاول إلقاء الضوء على هذا المعجم والمراحل التي مر بها انطلاقا من بناء قاعدة المعطيات التي أنجزناها ومكوناتها والنتائج المتوخاة منها.

- إحصاء وتحديد الجذور، ويتم من خلال ضبط القائمة التسلسلية والمرتبة ترتيبا ألفبائيا لكل الكلمات المدخلة (الثلاثية) أي الجذور التي تستنبط منها الكلمات.

2.4- الموارد اللسانية:

أ- قاعدة بيانات الصرف:

هناك من الخطوات التي ينبغي إجراؤها قبل البدء بتكوين قاعدة بيانات الصرف للمدونة. أولى هذه الخطوات هي تجزئة مفردات القرآن الكريم المدخلة إلى لواصقها الأولى ولواصقها النهائية وأدوات التشكيل ويمكن كحد أعلى اعتبار القطع 4 لواصق أولية كحد أعلى و2 من اللواصق الآخرية ونواة الكلمة وتشكيل للنواة وتشكيل آخر.

وبذلك فإن الكلمة يحجز لها 9 حقول . خذ مثلا كلمة " **أفبالباطل** " تحتوي 4 لواصق أولية هي " أ - فَـــ - بِـــ - الْـــ " ونواتها هو باطل يعقبها كسرة في الأخير . أما كلمة " **فسيكفيكهم** " فتحتوي على لواصق أولية عددها 2 هما " فَـــ - ـــسَـــ " والنواة يكفي واللواصق الآخرية هي " كَـــ " و " هم " والتشكيل على النواة هو السكون (غير الظاهر على الياء) وتشكيل آخر اللواصق هو السكون أيضا (الذي يظهر أو يختفي حسب الكلمة التالية للميم) ومثل ذلك في كلمة "أنلزمكموها". وتجدر الإشارة إلى أن ألف لام التعريف عند حذفها يكون الحرف الأول من الكلمة مشددًا إن كان من الحروف الشمسية لذلك يجب حذف هذه الشدة من نواة الكلمة، أما إن كان من الحروف القمرية فإن اللام (من ألف لام) تحمل سكوناً يجب حذفه. أما نواة الكلمة فيمكن أن تكون مشتقة من فعل ثلاثي على وزن ما . ويمكن أن يتم ذلك بواسطة برنامج خاص يقوم باستنباط الوزن الذي اشتقت منه هذه الكلمة . وهذا الاشتقاق يمكن أن يكون بسيطًا إذا لم يكن بين أحرف الجذر حرف علّة . أما عندما يكون أحدها (أو أكثر من واحد) حرف علة فإن قواعد تحويل الواو إلى ياء أو ألف أو بالعكس تجعل القواعد أكثر تعقيدا كما هو معروف.

أما إذا لم يكن النواة مشتقًا فيعني ذلك عدم وجود جذر للكلمة . وعلى هذا فإن عدد حقول قاعدة بيانات الصرف تصبح 12 حقلا هي: الكلمة، 4 حقول للواصق الأولية، نواة الكلمة، تشكيلها، لواصق آخرية، تشكيل الآخر، جذر النواة، وزن النواة. هذا وقد استعنتُ في تحديد جذور الكلمات القرآنية بالمعجم المفهرس لألفاظ القرآن الكريم[1] ولسان العرب[2]. انظر الجدولين (1) و(2).

كلمة(جذر)	عدد	كلمة(جذر)	عدد	كلمة(جذر)	عدد	كلمة(جذر)	عدد	كلمة(جذر)	عدد
عزز	121	بصر	149	ضلل	192	حقق	288	هو	482*
ءمم	122	رود	149	حسن	195	ذكر	293	رسل	514
يدي	122	وعد	152	خير	197	نزل	294	بـــ	518*
ءهل	124	وحد	154	نور	197	قبل	295*	شيء	520
بشر	124	غير	155*	جنن	202	نفس	299	بين	524*
رزق	124	صدق	156	عند	202*	ظلم	316	كفر	526
خوف	125	بعض	159	ملك	207	سمو	318	ءتي	550
دخل	127	نصر	160	حكم	211	كتب	320	لم	575*
بل	128*	شهد	161	دعو	212	ءو	326	قوم	661
خلف	128	نبء	161	ءنت	218*	هدي	327	ءذ	694*
عظم	129	كبر	162	ءم	221*	رءي	329	ءلى	743*
جمع	130	بنو	164	نوس	231	ءلي	332	له	806*
سءل	130	مع	165*	غفر	235	رحم	340	من	829*
طوع	130	موت	166	بعد	236	ثم	343	علم	855
لعل	130	سوء	168	ءمر	249	جعل	347	ءمن	880
نظر	130	كثر	168	ءخر	251	عمل	361	ربب	982
لكن	131	شرك	169	ءيي	253	عذب	374	كون	1359
نذر	131	قلب	169	وقى	259	كل	377	على	1443*
دنو	134	مثل	170	ولي	260	ءيه	383	ءلل	1471
قدر	134	قتل	171	خلق	262	قد	407	في	1702*
موسى	137*	تبع	175	ءخذ	274	لك	421	قول	1723
سلم	141	سبل	177	عبد	276	هم	445	لا	1731*
حتى	143	صلح	181	لو	276*	ءرض	462	ما	2181*
دون	145	خرج	183	هذا	276*	عن	465	ءله	2852
نعم	145	سمع	186	جيء	279	يوم	476	مـــن	3100*
لقي	147	حيي	190	كذب	283	ذلك	481	ءن	4038*

الجدول 01: قائمة الجذور والكلمات الأكثر ترددا في القرآن الكريم المستخدمة كقاعدة بيانات في البرنامج

1- محمد فؤاد عبد الباقي، المعجم المفهرس لألفاظ القرآن الكريم، مكتبة نويد إسلام، قم المقدسة، ط4، 1383هـ.

2- ابن منظور(أبو الفضل جمال الدين محمد بن مكرم الإفريقي)، لسان العرب، دار صادر، بيروت، ط6. 1417هـ /1997م.

الجذر2	الفعل	الجذر	الكلمة المدخلة	تشكيلها	اللواصق	تشكيل	تشكيل آخر	الوزن
رحم	رَحْمَن	رحم	الرَّحْمَنِ	الرَّحْمَن	رَحْمَن	الرَّحْمَنِ	ا?لرَّحْمَ!نِ	فعلان
رحم	رَحِيم	رحم	الرَّحِيمِ	الرَّحِيم	رَحِيم	الرَّحِيمِ	ا?لرَّحِيمِ	فعيل
عالم	عَالَم	علم	الْعَالَمِينَ	الْعَالَمِين	عَالَمِين	الْعَالَمِينَ	ا?لْعَ!لَمِينَ	فاعلين
ملك	مَالِك	ملك	مَالِكِ	مَالِك	مَالِك	مَالِكِ	مَ!لِكِ	فاعل

الجدول 02: بعث الأمثلة عن قاعدة بيانات الصرف

أ- **حصر التفعيلات الممكنة:** التفعيلة هي النموذج أو الوزن التي يمكن مصاقتها في اللغة العربية، إذ يتم إحصاؤها لمطابقتها بالكلمة المراد البحث عنها بغرض الحصول على الجذر.

أمثلة:

الكلمة	الوزن
عابد	فاعل
جهاد	فعال
سبيل	فعيل

ب- الأوزان الصرفية:يحتوي المعجم على قاموس الأوزان الصرفية كقاعدة أساسية لقاعدة المعطيات[1]، وتم تتحدد طبقات الأفعال (الصيغ) في الأبنية المقدمة في الجدول الموالي حيث تشير النجمات إلى أن الصيغة نادرة الورود في اللغة العربية[2].

الثلاثي المجرد	الثلاثي المزيد
فَعَلَ يَفْعِل	أفْعَلَ يُفْعِلُ
فَعَلَ يَفْعُل	فعَّل يُفَعِّل
فعِل يَفْعَل	فاعَل يُفَاعِل
فعُل يَفْعُل	تفعَل يتفعَّل
* فعِل يفْعِل	تفَاعَلَ يتفَاعَلُ
	إنفَعَلَ يَنْفَعِل
	افْتَعَلَ يَفْتَعِل
	اسْتَفْعَلَ يسْتَفْعِل
	** افعَوْعَل يفْعَوْعِل
	افعوَّل يَفْعَوِّل
	افْعلّ يَفْعلّ

1- إميل بديع يعقوب، معجم الأوزان الصرفية، عالم الكتب، الطبعة الأولى، 1413هـ/1993م

2- لقد تم حصر قاعدة البيانات في الكلمات المدخلة في الثلاثي فقط.

افعالّ يَفْعَالّ

ج- قائمة السوابق واللواحق:

- **السوابق** les Préfixes: إن أشهر السوابق التصريفية في اللغة العربية:[1]
- **لواحق المضارعة:** (أ، ن، ت، ي).

تضم (الهمزة، والنون، والتاء، والياء) وهي خاصة بالأفعال.[2]

- **اللواحق** les suffixes: نذكر منها: (ي، ا، وا، أن، ين، ن، تن، تم، تما، ت، نا، يات، ة، ية، ات).

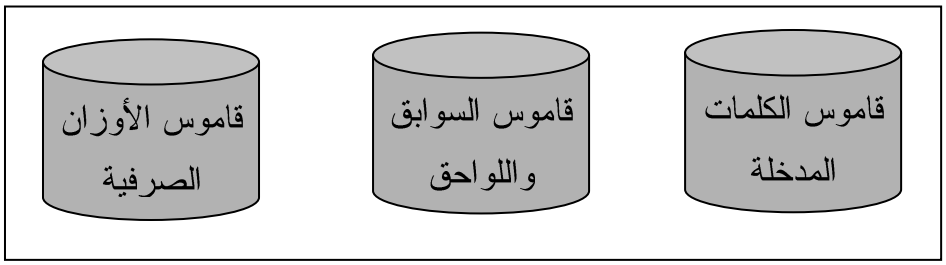
ويمكن تلخيص ما سبق في الجدول الآتي:

السوابق	أ				ن				ت			ي			
اللواحق	ي	ا	وا	ان	ين	ن	تن	تم	تما	تم	نا	يات	ة	ية	ات

الجدول 03: قائمة السوابق واللواحق.

3.4- إنشاء المعجم: قاعدة المعطيات:

تتطلب هذه المرحلة ضبط قاعدة المعطيات للمعجم وذلك من خلال تحديد هيكل المعطيات[3] لتمثيل المعجم على مستوى الحاسوب، ويمكن توضيح أهم مكونات المعجم في الشكل الآتي:



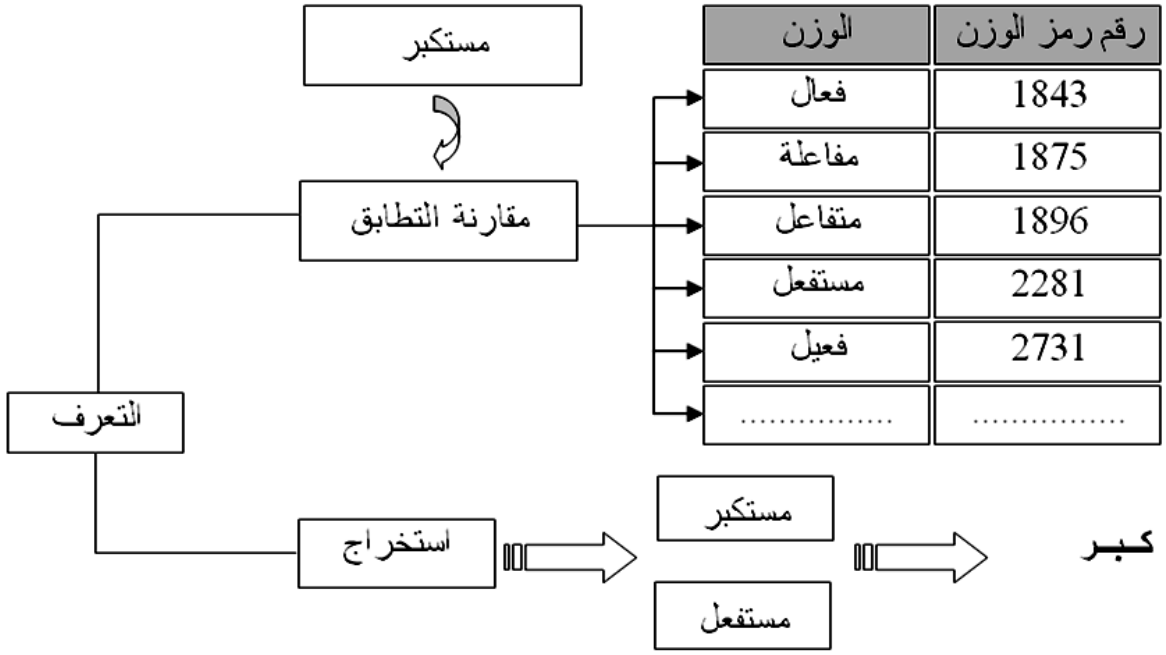
الشكل رقم 07: مكونات المعجم الإلكتروني TALA

يتم إنشاء في هذه المرحلة جدول (Table) يمكنه احتواء جميع التفعيلات الممكنة في اللغة العربية (ينظر الجدول 12)، والذي نستعمله لاحقا لمطابقة الكلمات المدخلة بغرض استخراج جذورها ويقابل كل واحدة من الأوزان الصرفية رمز رقمي (Code numérique) يستعمل لتسهيل عملية المعالجة، ويمكن توضيح آلية التعرف من خلال الشكل الآتي:

1 – أشواق محمد النجار (دلالة اللواحق التصريفية في اللغة العربية)، سابق، ص 72.

2 – للمزيد من التفصيل في هذا الموضوع ينظر، الفصل الأول، المعالجة الآلية للصرف العربي.

3 - La structure des données.



الشكل رقم 08: آلية التعرف على الكلمة المدخلة في البرنامج

بعد إنجاز جدول الأوزان الصرفية وجدول الكلمات المدخلة المحددة مسبقا بـــ (3000) كلمة مدخلة، نقوم بإضافة جدول آخر لاحتواء الجذور المحددة مرتبة ترتيبا ألفبائيا. وبناءا على هذا الجدول يتم إنشاء جدول آخر أكبر حجما يحتوي على جميع الكلمات وفق الطريقة التالية:

- نبدأ بأول جذر في القائمة ثم نقوم بتحديد جميع الكلمات التي يمكن استخراجها منه، ونقوم بترتيبها ترتيبا ألفبائيا داخل الجدول الثاني، ثم نمر إلى الجدول الثاني ونقوم بنفس العمل لإدراج الكلمات المستنبطة منه داخل الجدول تبعا للكلمات السابقة، وهكذا ينطبق العمل على باقي الجذور، وللربط بين الجدولين، نستعمل داخل جدول الجذور حقلا آخر (Champ) إضافة إلى الحقل الذي يحتوي على الجذر، يوضع فيه عنوان بداية مجموعة الكلمات.

ويمكن توضيح قاعدة المعطيات بثلاث جداول أساسية هي:

- جدول الأوزان الصرفية: التفعيلات.
- جدول الكلمات المدخلة (مفردات القرآن الكريم).
- جدول نحدد من خلاله جميع الكلمات التي يمكن استخراجها من الجدول الثاني، ونقوم بترتيبها ألفبائيا.

كما هو موضح في الشكل الآتي:

حـــــقـــــل Champ (للربط بين الجدولين)								
جدول -1-			جدول -2-			جدول -3-		
جدول الأوزان الصرفية			جدول الكلمات المدخلة[1]			تحديد جميع الكلمات المدخلة وتقوم بترتيبها ترتيبا ألفبائيا		
رقم رمز الوزن	الوزن		رقم رمز الكلمة المدخلة	الكلمة المدخلة		رقم رمز الجذر	الجذر	خ د ع
105	فَعَلَ	⇔	154	بحر	⇔	673		خدع
106	فَعْلٌ		162	برع		674		خديعة
107	فعول		364	جحر		675		خدعة
108	فعِيلٌ		442	جهد		677		مخدوع
109	فعلانٌ		477	حجب				....
110	مُستفعِلٌ		597	حلم			الجذر	ع ه د
111	فعُولٌ	⇔	673	خدع	⇔	1842		عهد
112	فِعَالٌ		692	خشع		1843		عِهاد
113	فعَّالٌ		1100	سبح		1844		عهدة
114	مَفعولٌ		1118	سجل		1845		تعهّد
115	مِفْعَلٌ		1352	شفع		1846		عهود
116	مَفْعَلٌ		1392	شهد				....
117	فِعَالَةٌ		1842	عهد			الجذر	ك ت ب
118	أفْعَالٌ	⇔	2279	كتب	⇔	2279		كَتَب
119	أفْعِلَةٌ		2465	مجد		2280		كَتَب
	....			.....		2281		كِتاب
						2282		كتّاب
						2283		كتابة
						2284		اكتتاب
						2285		كاتِبٌ
						2286		مكتوب
						2287		كتائب
								....

الجدول 04: قاعدة البيانات لإنشاء المعجم الإلكتروني(TALA)

1- الكلمات المدخلة في قاعدة المعطيات هي عينة من مفردات القرآن الكريم.

4.4- صياغة إجراءات الاستعمال:

نقوم خلال هذه المرحلة بتحديد الإجراءات المختلفة لاستغلال المعجم بكيفية استخراج الجذر وكيفية البحث عن الكلمة.

- **كيفية استخراج جذور الكلمات العربية وطريقة البحث عن الكلمة:**

خامسا- الدراسات السابقة:

إن الدراسات السابقة التي عالجت موضوع استخراج جذور الكلمات العربية، تنوعت ما بين دراسات أكاديمية أو بعض البرمجيات التجارية التي قدمت برامج تقوم بتحليل الكلمة إلى جذور وزوائد.

أ- برنامج التقسيم Glor arab [1]

يعمل البرنامج على تقسيم الكلمة إلى مكوناتها الجزئية: **السوابق، اللّواحق والجذر** وذلك وفق مرحلتين:

- **المعالجة الآلية:** ويتم ذلك وفق الخطوات التالية:

- تقسيم الجملة إلى وحداتها الجزئية مع إضافة معلومات عنها، مثل ترتيب الكلمة داخل الجملة وترتيب الجملة داخل النّص.
- إعادة بناء النّص بعد حذف الوحدات المكررة.
- تشكيل كشاف كلمات النّص.

- **المعالجة اللّسانية:** ويتم ذلك وفق ثلاث خطوات:

- البحث عن المداخل المعجمية لكل كلمة.
- البحث عن جذر كل مدخل معجمي متحصل عليه.
- التأكد من سلامة التحليلي اعتمادًا على مصادر خارجية (المعاجم اللّغوية).

تكمن أهمية هذا البرنامج في دوره، بحيث يقدم تحليل مورفولوجي للكلمات العربية ويحدد جذر كل كلمة.

ب- معجم للمعالجة الآلية للغة العربية (معجم اللّغة العربية المؤتمت) (DIINAR)[2]

مصدر هام للمعالجة الآلية للغة العربية، يقدم المداخل المعجمية مع معلومات مورفولوجية ونحوية لكل مدخل، يتكون معجم DIINAR من مجموعة قواعد البيانات (الفعل، الاسم، الحروف أسماء الآلة)[3].

---

1 -Laurence , Teirlinckx, La lemmatisation de l'arabe non classique- centre d'étude sur Grégoire. France. 2003- P: 08.

2 -DIINAR: Dictionnaire informatisé de l'arabe.

3 -Abbes, Ramzi, la conception et la Réalisation d'un concordancier électronique pour l'arabe. Institut national des sciences appliques, Science de l'informatique thèse doctorat 2004, P: 65.

	الجذر	المشتقات
الفعل	19457	5970488
الاسم	39099	1781316
أسماء الآلة	445	11731
أسماء الأعلام	1384	11403
المجموع	60385	7774938

الجدول 05: يبين مكونات قاموس DIINAR

ويُمكننا هذا المعجم من:

- اشتقاق الأسماء الممكنة من الجذور المختلفة.
- تصريف الأفعال.
- لكل اسم أو فعل يمكن الحصول على معلومات صرفية نحوية خاصة بالمدخل.

ج- أعمال شركة صخر: SAKHR

- **التحليل الصرفي:** يتيح المعالج الصرفي المتعدد الأطوار من شركة صخر المعالجة العميقة للكلمة العربية المفردة، ويعطي هذا المحلل نطاق الكلمات العربية بالكامل الحديث منها والقديم. يقوم المحلل بالتعرف على جميع أشكال جذر الكلمة أي أنه يقوم باستخلاص أصل الكلمة بعد تجريدها من اللّواحق ولا يتوقف عند هذا الحدّ بل يتخطى ذلك لاستخلاص البيانات الصرفية للكلمة مثل: الجذر، والميزان الصرفي لها[1].

سادسا- آلية التنفيذ (التطبيق):

6-1- متطلبات إنشاء البرنامج:

لغة البرمجة المستخدمة[2]: Langage de Programmation:

تمثل هذه المرحلة الهيكلية الحاسوبية للمعجم المنشود ويتطلب تصميم هيكل المعجم استخدام لغة برمجية وقادرة على استيعاب أوامر المبرمج أولا والتكيف مع محتويات قاعدة المعطيات

1- للمزيد من التفصيل ينظر: موقع شركة صخر- http://www.sakhr.com.

2- البرمجة : La programmation : تعني مجموعة من الأوامر والتعليمات الّتي تُعطى للحاسوب في صورة برنامج مكتوب بلغة برمجة معينة بواسطة معالج نصوص ، ويتكون مصدر البرنامج من عدّة سطور وكل سطر يعتبر جملة ، ويتعامل الحاسوب مع كل جملة بترتيب معين لإنجاز الأمر الذي صمم البرنامج لتحقيقه. وهناك أمثلة للغات البرمجة إلاّ أنّه يوجد فيها نوع من التخصص حيث يعتمد اختيار اللغة على المهام الّتي نريد من الحاسوب إنجازها حيث أنّ لكل لغة ما يناسبها وهناك مهام يصعب تنفيذها بلغة معينة ولكنها تكون ميسورة بلغة أخرى ، ومن اللغات الّتي ظهرت في فترة الستينات لغة "البسيك" Basic وهي لغة تدرس للمبتدئين ، وظهرت بعدها لغة فيجوال بيسك الّتي وضعها ميكروسوفت وأيضا فيجوال بيسك سكريبت الّتي تقدّم برمجيات تعمل ضمن صفحات الإنترنت، وأيضا لغة سي++ ائعة الصيت ، وهي لغة برمجة تمّ ابتكارها كتطوير للغة سي الّتي هي الأخرى تحسين للغة بي وكذلك هناك لغات أخرى أدهشت العالم كلغة الجافا Java.

للمعجم ونقترح بخصوص لغة البرمجة المستخدمة في هيكلة المعجم لغة دلفي لما تتمتع به من ليونة وسهولة الاستخدام في صناعة وتطوير البرامج الحاسوبية

يقترح البرنامج استخدام لغة برمجة تناقلية[1] وتعمل على شريحة واسعة من الأجهزة الحاسوبية بما فيها الأجهزة الكفية[2] والجوالات ومختلف النظم التشغيلية ، وباستعراض سريع للغات البرمجة يمكن الأخذ بأحد اللغتين التاليتين :

- لغة جافا[3] "Java".
- لغة سي++[4] "C++".

إن إنشاء محلل صرفي آلي[5] يتطلب خبرة ومعرفة جيدة بأحد لغات البرمجة الكثيرة، لقد اعتمدت في انشاء هذا المعجم على برنامج الجافا، فهو يتميز بالسهولة والسرعة والدقة.

---

1 - Portable.

2- الأجهزة الكفية : الهدف من هذه الأجهزة المحمولة هو تشجيع وتيسير إنشاء برمجيات مفتوحة المصدر للاستخدام على أجهزة الكمبيوتر. للمزيد من التفصيل ينظر موقع الويب: http://Translate.google.ae

3- لغة جافا "Java" ابتكر لغة الجافا المهندسان James Gosling و Patrik Naughton "جيمس جوسلينغ وباتريك نوتون" إثناء عملهما في مختبرات "San Microsystems".

وقد كانت لغة الجافا تطويرا للغة سي++ ، تتميز لغة الجافا بمميزات خاصة مما يجعلها أكثر لغات البرمجة إثارة حيث تمكننا من:

- إضافة الحركة والصوت إلى صفحات الويب.
- كتابة الألعاب والبرامج المساعدة.
- إنشاء برامجيات واجهة مستخدم رسومية.
- تصميم برمجيات تستفيد من كل مميزات الأنترنت.

إنّ لغة الجافا هي برمجة كائنية التوجه ومصممة للعمل على آلة افتراضية بحيث لا تحتاج إلى الترجمة من جديد عند استخدام برامجها على منصة التشغيل أو نظام تشغيل جديد. لقد تمّ فتح المصدر لهذه اللغة بين نوفمبر 2006 وما 2007 وأصبحت Open Source Code (مفتوحة المصدر).

توفر لغة الجافا بيئة تفاعلية عبر الشبكة العنكبوتية ، وبالتالي تستعمل لكتابة برامج تعليمية للإنترنت عبر برمجيات المحاكاة الحاسوبية للتجارب العلمية وبرمجيات الفصول الافتراضية للتعليم الإلكتروني والتعليم عن بعد، للمزيد من التفصيل ينظر : جافا (لغة البرمجة) ويكيبديا ، الموسوعة الحرة، موقع الويب : http://ar.wikipedia.org

إنّ المميزات الّتي تتميز بها لغة الجافا على لغات البرمجة الأخرى جعلتنا نختارها كلغة البرمجة في العمل التطبيقي لبناء المعجم الإلكتروني للمعالجة الآلية للغة العربية.

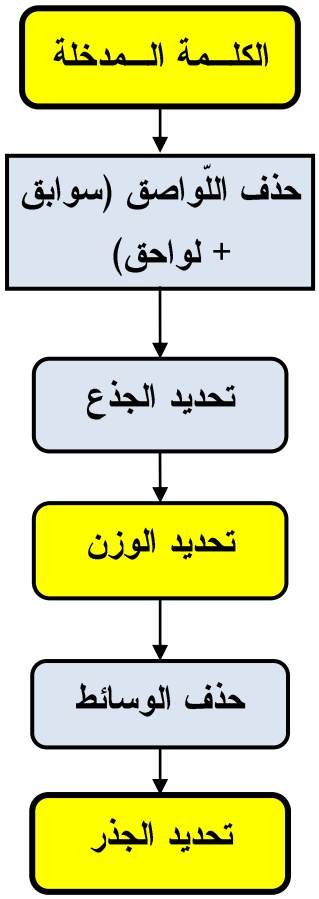
تنبيه : يجب عدم الخلط بين "الجافا" و"الجافاسكريبت" فرغم تشابه اللغتين في السياق والتراكيب إلاّ أنّهما لغتان مختلفتان ولهما أهداف مختلفة.

4- لغة سي++ "C++" هي لغة برمجة للاستخدامات العامة، تعتبر لغة السي++ لغة برمجة كائنية "البرمجة كائنية المنحنى "La programmation orientée objet" إذ يتمّ بناء البرنامج بواسطة استخدام الكائنات وربطها مع بعض وواجهة البرنامج الخارجية باستخدام هيكل البرنامج وواجهات الاستخدام الخاصة بكل كائن ، و لغة C++ من لغات البرمجة العالية والمستوى ، وفي نفس الوقت قريبة من لغة التجميع ذات المستوى المحدود كما أنّها تعدّ لغة برمجة إجرائية (يمكن كتابة برنامج يحتوي على إجراءات وتوابع فقط... وهي لغة برمجة ناشئة من لغة C الذي قام بتطويرها Bayarne Stroustrup". للمزيد من التفصيل ينظر موقع الويب : سي++ (لغة البرمجة) ويكيبديا ، الموسوعة الحرة : http://ar.wikipedia.org

5- المحلل الصرفي : يمكنه التعامل مع الكلمات العربية تحليلا وتوليدا، حيث يمكنه تحليل الكلمة إلى عناصرها الصرفية الأساسية، وهي: الجذر والوزن والسابقة واللاحقة، مثل كلمة: [المسلمين] يحللها إلى: الجذر [سلم] والوزن [مُفْعِل] والسابقة [ال] واللاحقة [ين]، ويمكنه أيضًا توليد الصيغ الصرفية المختلفة التي يمكن أن تشتق من أحد الجذور اللغوية، مثل الجذر [سلم] تولد منه الصيغ [يسلم، أسلم، مسلم، مسلمين، إسلام، مسلمة..إلخ] .

وللمحلل الصرفي دور مهم في بناء قاعدة بيانات المعجم الحاسوبي، فعن طريقه يمكن اختزال كلمات قاعدة البيانات إلى عدد محدود من الكلمات الحيادية وتبقى مهمة التعامل مع الكلمات الفرعية التي تشتق عنها من خلاله.

يعمل برنامج المحلل الصرفي للمعجم الإلكتروني الذي قمنا بإنشائه على تحليل الكلمة المدخلة والكشف عن جميع أجزائها الصرفية المكونة لها: سوابق، لواحق، وسائط (Infix)، جذر، جذع، كما يقوم أيضا بتحديد وزنها، والقالب الصرفي الذي صيغت عليه، وهو يعمل على ست خطوات، مبينة كما يلي:



شكل يوضح عمل برنامج استخراج الجذر في المعجم الإلكتروني

6-2- الطريقة التي يعمل بها برنامج المحلل الصرفي:

يعتمد هذا المحلل الصرفي في عمله على طريقة خوارزمية التحليل الصرفي التبادلي، حيث يتم مقارنة الكلمة المدخلة (Input Word) مع قاوائم معدة مسبقاً تحوي الكلمات الجامدة والسوابق واللواحق والأوزان والجذور، هذه المقارنة تتم وفقاً لخوارزمية تبادلية تقوم باختبار جميع تباديل حروف الكلمة المدخلة مع قائمة الكلمات الجامدة (Invariable Word) إذا لم تتطابق الكلمة مع هذه القائمة يعني أن الكلمة ليست بجامدة، فتنتقل الخوارزمية إلى الخطوة الثانية وهي مقارنة حروف الكلمة بقائمة السوابق ثم اللّواحق أين يتم حذف الحروف الزائدة، وبالتالي يتم تحديد جذع الكلمة (Stem) ومنها استخراج الوزن انطلاقا من قائمة الأوزان، في الأخير تحذف الفونيمات المتبقية ويتم استخراج الجذر وهو المطلوب.

هذه الطريقة تتطلب الوقت والجهد وقوائم اختبار كبيرة، وقد عمل بهذا الأسلوب بعض الباحثين مثل الفداغي والسعدون، والشلبي، والأفندي[1].

ا/التحليل الصرفي:

في هذه المرحلة سنعرض طريقتين لاستخراج جذور الكلمات العربية نبين من خلالها أهمية الجذر في مراحل التحليل ، وتحديد جذر الكلمة لتحقيق الربط بين مختلف الاشتقاقات الصرفية لها، كما هو موضح في المخطط الآتي.

تشترك هاتان الطريقتان في الخطوات التالية :

- استخراج جذع الكلمة (أساسها) وذلك باستبعاد اللواصق.
- يتم مقارنة الصيغ الصرفية المستنبطة لجذع الكلمة بقائمة من الصيغ الصرفية الممكنة.
- عند تحديد صيغ صرفية يتم استخراج الجذر ومقارنته بقائمة الجذور لموجودة في اللغة العربية.
- إذا تم تحديد الجذر فإن العملية تنتهي ويكون لدينا الجذر واللواصق والصيغ الصرفية.
- إذا لم يتم تحديد الجذر فانه يتم اختيار صيغا صرفية أخرى ويتم تجربتها.
- اذا لم يتم تحديد الجذر باستخدام كل الصيغ فإن عملية إزالة اللواصق يتم إعادتها مع إزالة لواصق أصغر.

6-3- مبدأ عمل برنامج المحلل الصرفي :

تتمثل مهمة المحلل الصرفي، في معالجة الكلمات المدخلة إليه وإعطاء تحليل صرفي كامل لها، و تتمثل خطوات عمل المحلل فيما يلي:

1- البحث عن الكلمة في قائمة الكلمات الجامدة وحسب نتيجة البحث تنجز المراحل التالية:

- إن وُجدت الكلمة انتهت العملية.
- إن لم يعثر على الكلمة في قائمة الكلمات الجامدة، تنتهج الخوارزمية الخطوة التالية:

2- البحث في قائمتي السوابق واللواحق لحذف جميع السوابق واللواحق المتصلة بها وبالتالي يحدد الجذع.

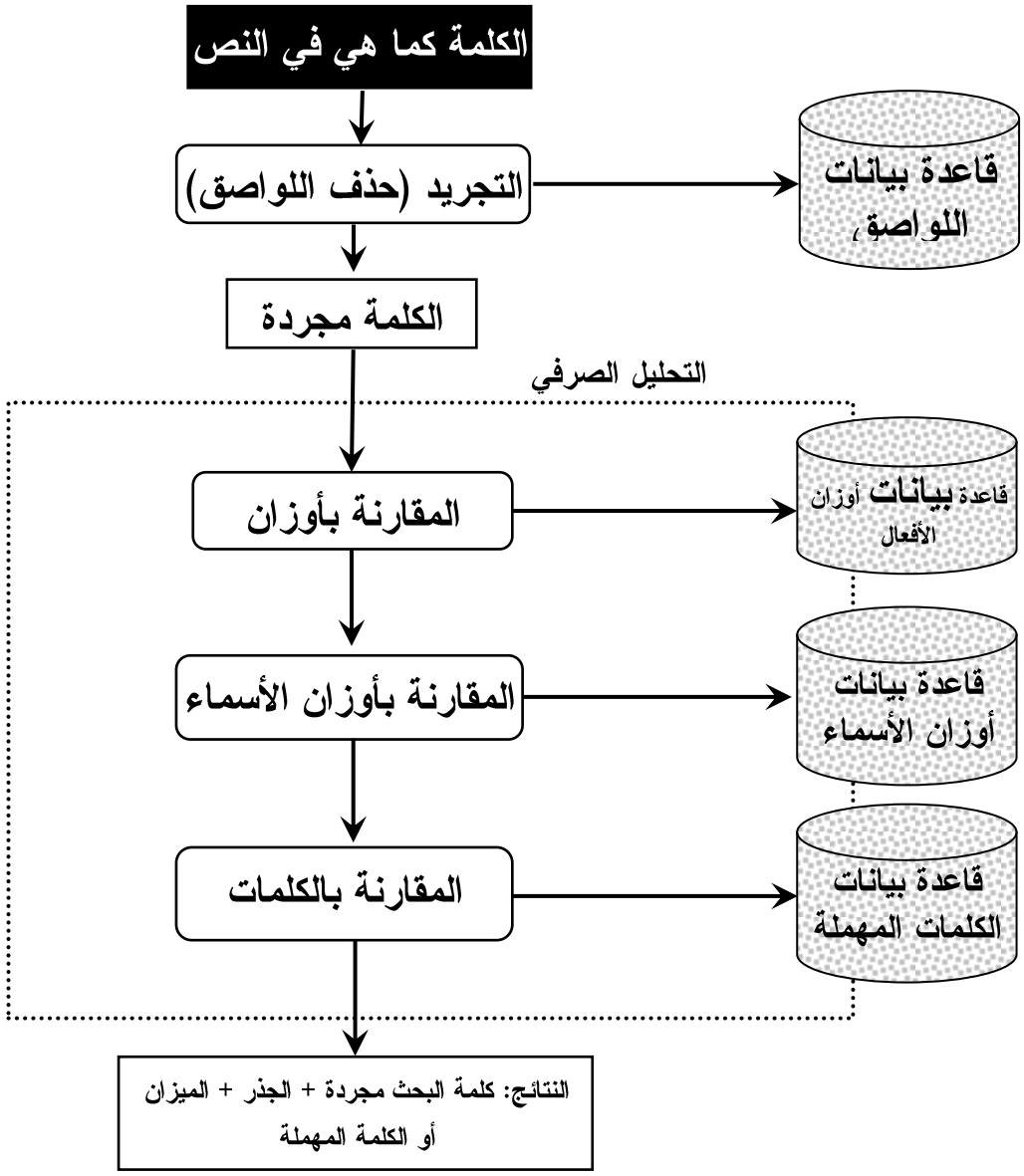
3- البحث في قائمة الأوزان وبالتالي يحدد الوزن.

4- حذف الفونيمات المتبقية.

5- استخراج الجذر ومقارنته بقائمة الجذور،

- إذا عثر عليه في قائمة الجذور، انتهت العملية.
- إذا لم يعثر عليه، هناك خطأ ما.

1- د. إبراهيم الـخراشي، Survey of Arabic Morphological Analysis Techniques، ص19، الكتاب متوفر على الرّابط التالي: http://ccisdb.ksu.edu.sa/files/rep1120000.doc



الشكل (1)

رسم توضيحي يبين مبدأ عمل برنامج المحلل الصرفي في المعجم الإلكتروني

من خلال هذا المخطط سنقدم طريقتين لاستخراج جذور الكلمات العربية نبين من خلالها أهمية التحليل الصرفي في عملية استخراج الجذر :

<u>الطريقة الأولى (المقاربة الأولى)</u>: المقارنة التقريبية (النوافذ المنزلقة):[1]

الوزن + الجذر = القاعدة [2] Stem

1 - Les fenêtres coulissantes.

2- القاعدة (la base) هي ما تبقى من الكلمة بعد حذف الزوائد الإعرابية (Affixes)، وينظر:

-Delphine Bernard, « Appentissage de connaissances morphologique pour l'acquisition automatique de ressources lexicalesù p.194. IDM.

وفكرة هذه الطريقة هي أن نأخذ الكلمة المدخلة ثم نحاول إيجاد الوزن، وبعد أن نجده نبدأ المقارنة بين الوزن والمدخل حتى يتكافأن ومن ثم ينتج الجذر.

كيفية استخراج الجذور والبحث عن الكلمات:

بعد الحصول على الكلمة المدخلة نقوم بمطابقتها بالأوزان الموجودة في قاعدة المعطيات، وعند الحصول على الوزن المناسب نأخذ كل حرف يكون في نفس المرتبة أي الحرف الأصلي من الحروف الثلاثة ⇐ **المطابقة بين الوزن (التفعيلة) والكلمة المدخلة حتى ينتج الجذر.**

ويمكن توضيح ذلك عمليا:

مثال: الكلمة المدخلة: "المستقبلون".

الكلمة المدخلة	ا	ل	م	س	ت	ق	ب	ل	و	ن
الوزن										
الجذر										

بعد تجريب أوزان كثيرة منها الخاطئ ومنها الوزن الصحيح "مستفعل" فعملية البحث تكون بعد مطابقة جميع الأوزان الموجودة في قاعدة المعطيات المنجزة، ستكون نتيجة البحث الحصول على الوزن "التفعيلة" – مستفعل.

<u>الحالة الأولى</u>: سنقوم في مثالنا هذا بتجريب هذا الوزن الصحيح وذلك لتسهيل العملية الإجرائية، وسيمثل الوزن في النافذة كالتالي:

الكلمة المدخلة	ا	ل	م	س	ت	ق	ب	ل	و	ن
الوزن	م	س	ت	ف	ع	ل				
الجذر										

من خلال النظر جديدا في الوزن نلاحظ أن الحروف الزائدة عن الوزن الأساسي "ف ع ل" هي حرف الميم وحرف السين، وحرف التاء. إذن سوف نقارن هذه الحروف مع كل حرف يقابلها في حروف الكلمة المدخلة حتى يحصل تطابق ...

الآن سنزيح الوزن بمقدار واحد إلى اليسار كما في الرسم أدناه.

الكلمة المدخلة	ا	ل	م	س	ت	ق	ب	ل	و	ن
الوزن		م	س	ت	ف	ع	ل			
الجذر										

الحالة الثانية: نلاحظ أن حرف الميم لم ينطبق مع اللام وحرف السين لم ينطبق مع الميم وحرف التاء لم ينطبق مع السين.

سنزيح الوزن لليسار أيضا مقدار واحد كما في الرسم التالي:

الكلمة المدخلة	ا	ل	م	س	ت	ق	ب	ل	و	ن
الوزن			م	س	ت	ف	ع	ل		
الجذر										

عند المقارنة نلاحظ أن حرف الميم انطبق مع الميم في الكلمة المدخلة.

الكلمة المدخلة	ا	ل	م	س	ت	ق	ب	ل	و	ن
الوزن			م	س	ت	ف	ع	ل		
الجذر										

وحرف السين أيضا تطابق مع السين في الكلمة المدخلة كما هو واضح في الجدول أعلاه: وكذلك نفس الشيء لحرف التاء مع التاء في الكلمة المدخلة.

الآن نأخذ الحروف في الكلمة المدخلة المقابلة لحروف الوزن الأساسي – ف ع ل – وهي حرف "القاف" الذي يعلو الفاء وحرف الباء الذي يعلو حرف العين، وحرف اللام في الكلمة المدخلة الذي يعلو حرف اللام، من الوزن. وننزلها كجزء في الخانة السفلى كما يوضحه الجدول التالي:

الكلمة المدخلة	ا	ل	م	س	ت	ق	ب	ل	و	ن
الوزن			م	س	ت	ف	ع	ل		
الجذر						ق	ب	ل		

فالجذر الذي تحصلنا عليه هو (ق ب ل).

ملاحظة: لقد صادفنا عند المعالجة الآلية لقاعدة المعطيات للمعجم بعض المشاكل فهذه الطريقة لا تخلو من بعض المشاكل فمثلا الكلمة المدخلة (قيل) لا يوجد لها وزن لأنه يجب إرجاع الكلمة إلى أصلها، وهي (قول) وهناك أمثلة أخرى كالفعل (ع) في الجملة (ع درسك) "أي إنتبه للدرس" أو "إحفظ الدرس" ففعل الأمر (ع) لا تستطيع هذه الخوارزمية إيجاد الجذر.

البحث عن الكلمة:

يتم البحث بفضل نتائج المرحلة السابقة، إذ يؤخذ الجذر ليطابق بكلمات جدول الجذور (الكلمات المدخلة) وحال العثور عليه يستعمل العنوان الذي يقابله لبلوغ بداية مجموعة الكلمات المكونة منه، كما يفيدنا العنوان المصاحب للجذر التالي في تحديد نهاية المجموعة حتى لا تخرج عملية البحث عن المجال المطلوب كما هو موضح في الجدول الآتي:

جدول الجذور (الكلمات المدخلة)		جدول الجذور (الكلمات المدخلة)	
ك ت ب		ع ه د	
	كَتْب		عَهْد
	كَتَبٌ		عِهاد
	كِتابٌ		عُهْدَةٌ
	كَتّابٌ		تَعَهُّد
	كِتابةٌ		تعَاهَد
	كُتَيِّب		عُهودٌ
	مَكْتَبٌ		مُعاهَدَةٌ
.....		.....	مُتعهِّد
	.....		.....
خ د ع		خ ل ع	
	خدْع		خَلْع
	خِدع		خُلْع
	خديعة		خَالِعٌ
	خُدعَةٌ		خَلِيعٌ
	مَخدُوعٌ		مَخْلُوعٌ
	مَخْدَعٌ		مُخالعة
	مُخَادِعٌ		خِلعَة
.....	.....	.....	.....
	.....		.....

<u>الجدول رقم14</u> : كيفية البحث عن الجذر باستعمال العنوان الذي يقابل كل مجموعة من الكلمات المدخلة.

***/ مرحلة البحث وعرض النتائج:**

1- في حالة البحث المطابق يكون البحث في قاعدة بيانات ، هذا إذا افترضنا أن المدونة عبارة عن نص قرآني يمكن من خلالها تحديد كلمات النص وتحديد الآيات التي تحتوي على كلمة البحث.

2- في حالة البحث باللواصق أو البحث بالمشتقات تُحَلّل الكلمة صرفيّاً باستخدام المحلل الصرفي الخاص بالقرآن الكريم المقتصر فقط على الجذور الموجودة بالقرآن.

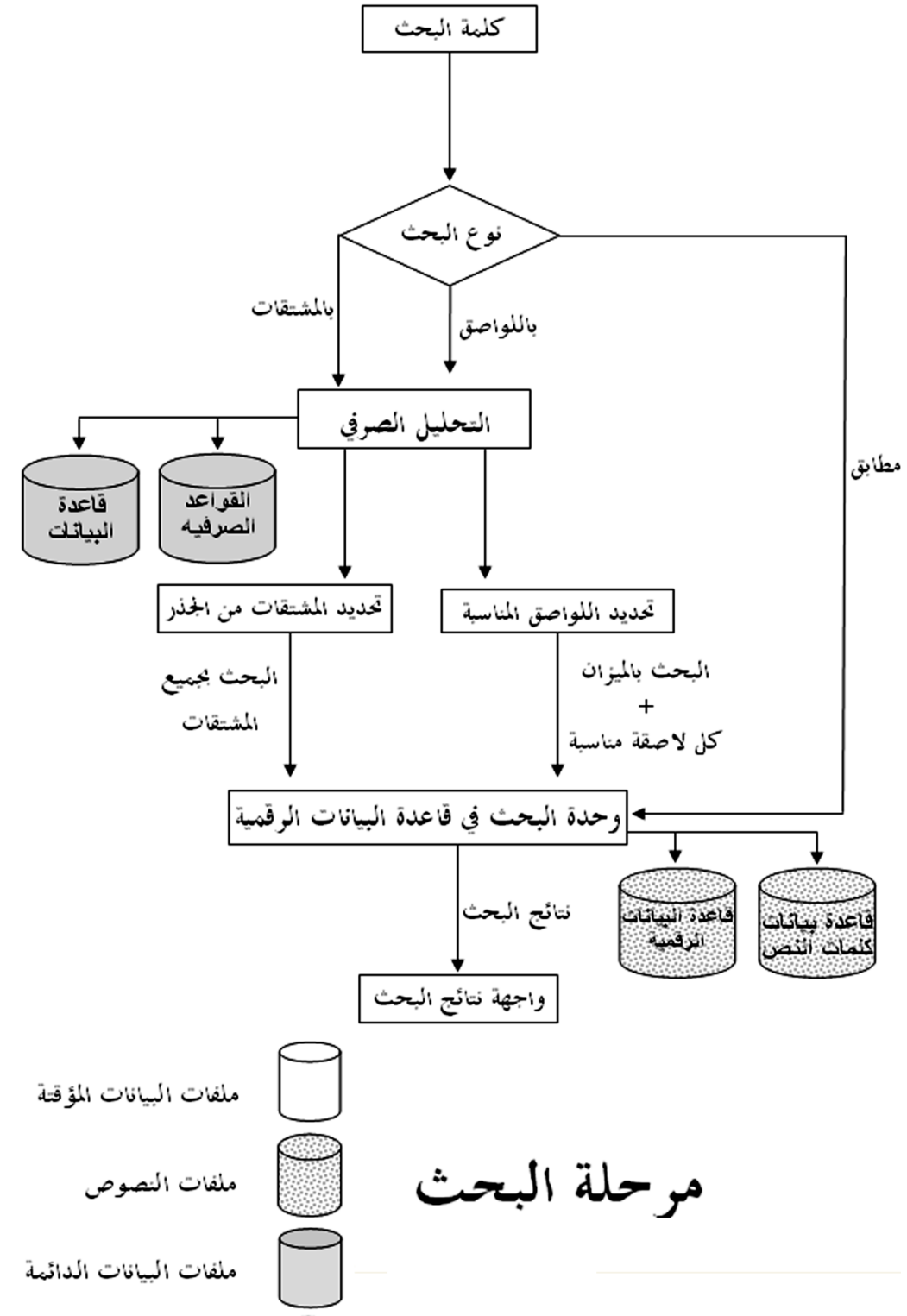
3- إتاحة الإمكانية للمستخدم لاختيار الجذر.

4- تحديد الآيات التي تحتوي على هذا الجذر باستخدام قاعدة البيانات الرقمية الناتجة عن عملية البناء.

5- في حالة البحث باللواصق تحدد الآيات فقط التي تحتوي على الجذر نفسه، وبها كلمة البحث نفسها كما هي أو باللواصق.

6- عرض النتائج وتلوين الآيات والكلمات في نتائج البحث.

الشكل الآتي يوضح خطوات تلك المرحلة:



2-الطريقة الثانية: تعرف بـــ التشذيب[1] "Stemming"

تقترح الطريقة الثانية طريقة عملية لاستخراج جذور الكلمات العربية باستخدام تقنيات المعالجة الآلية للغة العربية[2].

ترتكز قاعدة معطيات هذه الطريقة على مفهوم الوزن (التفعيلة) والجانب الإيجابي في هذه الطريقة أنها تقوم بالتعرف على مختلف الكلمات بطريقة ذكية عن طريق استخراج الجذر والوزن المناسبين لكل كلمة مدخلة دون استعمال معجم الكلمات.[3]

تتطلب هذه المقاربة بناء قاعدة معطيات أكثر صرامة، فالإشكال المطروح، هو كيف يمكننا أن نتمكن من استخدام هذه المعلومات واستثمارها في بناء نظام آلي لاستخراج جذور الكلمات العربية؟ وعليه لابد من برنامج لجمع ومعالجة، ونشر واستغلال قاعدة المعطيات لمساعدة المستعملين لتجنب هذه الصعاب والعقبات المعلوماتية.[4]

في هذا الإطار يعرض هذا العمل التطبيقي طريقة لاستخراج جذور من الكلمات العربية أثناء مرحلة التحليل الصرفي في المعالجة الآلية للغة الطبيعية، تعتبر هذه المرحلة أساسية ومهمة لأنها أساس جميع تطبيقات المعالجة الآلية للغات الطبيعية:[5]

الترجمة الآلية، الفهرسة، الاسترجاع، الإجابة عن الأسئلة، وحتى التنقيب والتصنيف الآلي.

إن تطبيق تقنيات المعالجة الآلية للغات الطبيعية على اللغة العربية يطرح عدة مشاكل في كيفية استثمار قاعدة المعطيات لاستخدامها في إعداد برامج حاسوبية منها:[6]

- مشكل الغموض الناتج عن غياب الحركات في الكلمات[7] مما يتطلب قواعد صرفية معقدة.[8]

---

1 – التشذيب: Lemmatisation (Stemming): هو إحدى الطرق المستخدمة لزيادة كفاءة البحث في نظام استرجاع البيانات كإزالة السوابق واللّواحق (الأحرف الزائدة) وإرجاع الكلمات إلى أصولها ليتم استخراج الجذر.

ينظر: خطوات عملية باتجاه تطبيق مفهوم البحث بالمفاهيم لمحركات البحث والاسترجاع العربية، حيدر العميد، شيخة الكتبي، آمنة الكعبي، خديجة الشبلي، نيلا الشامسي، نورة النعيمي، شيخة المهيري، جامعة الإمارات العربية المتحدة، كلية تقنية المعلومات، قسم هندسة البرمجيات.

2– عبد السلام نويصري، التهاكوك، سكولور، معالجة النصوص العربية لأجل الفهرسة والاسترجاع، المعالجة الآلية للغة العربية، وقائع الندوة الدولية CITALA'07، معهد الدراسات والأبحاث للتعريب، الرباط ، المغرب، يونيو 2007، ص 32.

3 - Bessou Sadik, Louail Mohamed, Refoufi Allaoua, Kadem Zehoure et Touahria Mohemmed. (un système de lemmatisation pour application de TALN) Traitement Automatique de langue Arabe, Colloque International CITALA 07.18-19 Juin 2007, Institut d'étude et de Recherche pour L'arbisation 2007. p.35.

4 -LEHMAM A, rt Bouvet, un Résumé de textes application a la langue arabe, TALAN (2004).

5 - Bessou S, Saadi A, Touahria M (un système d'indexation et de recherche de textes en arabe (SIRTA), LANIA 2007, chlef, Algerie, (2007).

6 - ATTIA Mohamed A, « Devloping Robust Arabic Morphological Tranducer Using Finite State Technology In 8 th annual CLUK research colloquium (2005).

-ATTIA Mohamed A, (Alarge – Scale Computational processor of the Arabic Morphology : A Master's thesis, Cairo University, (Egypt) 2000.

7- ALI, Nabil (the Second wave of Arabic Natural Language processing (2003).

8- Bouzidia, Fouad Soufiane, Guy LAPALME, « un système de resumé de textes en arabe, 2ème Congres International sur l'ingénierie de l'arabe et l'ingénierie de la langue : Alger, (2005).

- مشكل التعرف على الوحدة الخطية للكلمة الواحدة.[1]

1- **الهدف:** إن إرجاع الكلمة العربية (الكلمة المدخلة) إلى أصلها هو إجراء يرتكز على إزالة العلامة الإعرابية والاشتقاقية للكلمة ليحولها إلى شكلها الموحد وهو "الفعل الثلاثي"، عن طريق هذه العملية يمكن التأكد من أن الكلمة تنتمي إلى مجموع الكلمات العربية أم لا ؟

2- **مبدأ عمل النظام (المحلل):** ترتكز الطريقة على تحليل قاعدة المعطيات صرفيا، فالهدف إذن هو التحقق ما إذا كانت كل كلمة ما تنتمي إلى داخل النظام يعتمد هذا التحليل على تقطيع الكلمات إلى وحدات[2] بغض النظر عن العلاقات النحوية بينها.[3]

3- **آلية عمل النظام:**

أ- في البداية يقوم المحلل بتسوية الكلمات[4] عن طريق توحيد رسم بعض الحروف مثل الهمزة والألف المقصورة، هذه العملية ضرورية بسبب التغيرات المختلفة عند كتابة نفي الكلمة، ثم تأتي مرحلة التقطيع حيث تقسم الكلمة إلى وحدات صرفية بإزالة السوابق واللواحق.[5] الجزء المتبقي يتمثل في الجذر.

ب- لا يمكن لهذا النظام أن يعمل دون قواميس تحتوي على الوحدات الصرفية هذه المرحلة تسمح بالتحقق من توافق الكلمة المنحلة والجذر المراد استخراجه، وهذا بعد عملية استخدام قاموس الأوزان الصرفية وقاموس السوابق واللواحق عن طريق المقارنة والتطابق.[6]

جـ- بعد هاتين المرحلتين تأتي المرحلة الثالثة التي تعتمد على إعطاء تفسير لكل وحدات الكلمة بالاعتماد على مفهوم الوزن (الميزان الصافي) وكنتيجة عملية تعطي القيم الصرفية لمختلف الوحدات الخطية للكلمة المفردة.

1- ATTIA Mohamed A, « Report on the Intreduction of arabic to pargram, the pargram Fall Metting, National Centre for language technology ; Ireland (2004).

2- ATTIA Mohamed A, « Devloping Robust Arabic Morphological Tranducer Using Finite State Technology In 8th annual CLUK research colloquium.

3- تجدر الإشارة إلى أن هذا التحليل يختلف عن التحليل التقليدي الكلاسيكي الخاص بلغات البرمجة لأنه يرتكز على مفاهيم تتعلق بالصرف وقواعد اللغة، وذلك باستغلال واستثمار كل المفاهيم التي تطرقنا إليها في الفصل الثاني، الخاص بالمعالجة الآلية للصرف العربي، واستخدامها في إعداد قاعدة معطيات هذا النظام.

4- LARKEYL. S, BALL ESTEROSL AND CONNELLM, Improving Stemming For Arabic Information Retrieval : Light Stemming and co – Occurrence analysis, C Sigir 2002à Tampere, Finland August 2002.

- ينظر، محمود الديكي (توصيف مركب العدد في اللغة العربية للحاسب الآلي) المعالجة الآلية للغة العربية، وقائع الندوة الدولية، جوان 2006. إشراف عبد الفتاح حمداني، خالد الأشهب، محمد الراضي، جامعة محمد الخامس، السويسي، معهد الدراسات والأبحاث للتعريب الرباط، جانفي 2007، ص 28.

5- Mohammed El Amine Abderahim, « Reconnaissance des Unités linguistique Signifiantes, thèse Doctorat Spécialité : Informatique (2007 – 2008).

6- Sadik Bessou, « Anayse de données textuelles pour la classification Automayique pour les Technique de text Mining application a la langue Arabe.

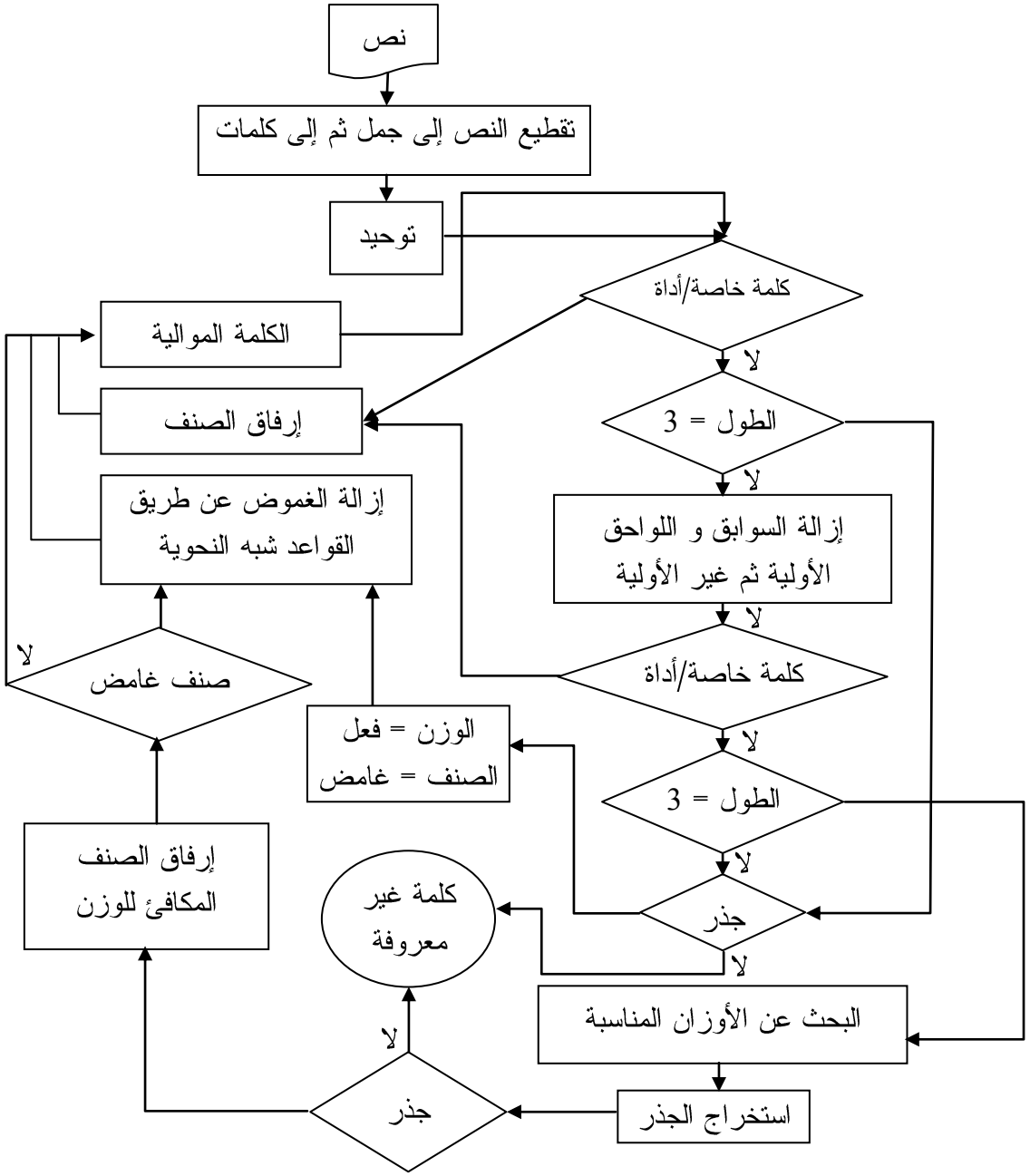
Mémoire pour l'obtention du diplôme de magister Option : Sciences et technologies de l'information et de la Communication.

Université Frhat ABBAS – SETIF 2007, p.64.

للتحقق من أن الكلمة المدخلة تنتمي إلى المفردات العربية المكونة لهذا النظام يكفي أن نجد لها جذرا ووزنا مناسبين.[1]

يرتكز هذا النظام في عملية التحليل الصرفي على ثلاثة مراحل أساسية هي:[2]

– التقطيع – البحث عن الجذر والوزن المناسبين – التفسير، كما هو موضح في الشكل الآتي:[3]



الشكل رقم 09: مخطط يوضح أهمية استخراج الجذر أثناء عملية التحليل الصرفي في البرنامج.

1– المرجع السابق، ص 64 ، 65.

2– المرجع نفسه، ص 65.

3– المرجع نفسه، ص 66.

7- القواميس المستخدمة في هذا النظام وبنيتها:

7-1- قوائم البيانات المدخلة في برنامج المحلل الصرفي:

قمنا بإنجاز قوائم البيانات باستخدام برنامج الأكسس (Access) وهو ينتمي إلى عائلة المايكروسوفت أوفيس (Microsoft Office)، يعود سبب اختيار الأكسس لكونه سهل ويأخذ بعين الاعتبار الحروف العربية.

أ- قاموس الأوزان:

نجذ في هذا القاموس كل المعلومات الضرورية أثناء تشفير الكلمات مع ذكر الصنف المكافئ لكل ميزان صرفي، هل هو لفعل أو لاسم كما نجد في هذا القاموس تشفيرا رقميا خاصا يسهل التعرف على وزن الكلمة المحللة واستخراج الجذر المرافق.[1]

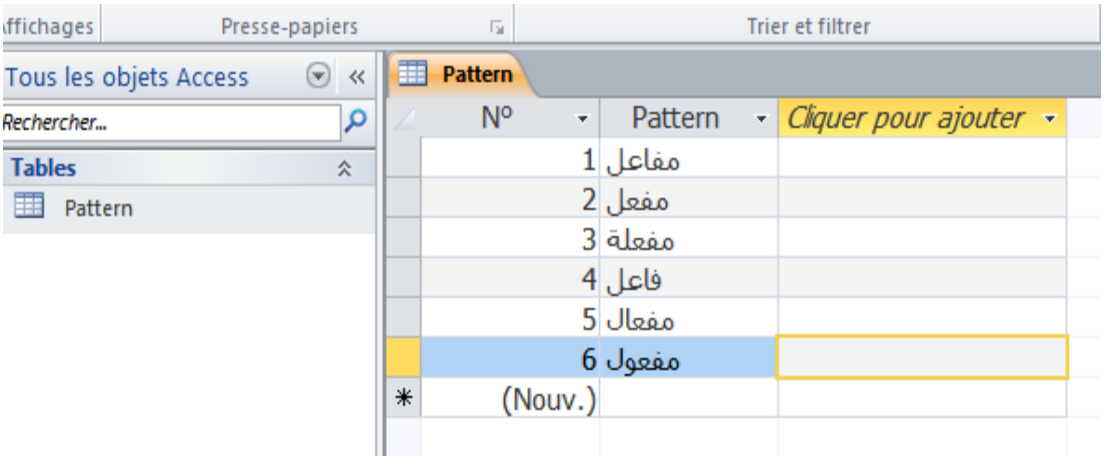
مثال: الكلمة المدخلة: مفاتيح، الوزن، مفاعيل.

الوزن	قائمة الزوائد	الصنف
مفاعيل	135	اسم

<u>الشكل رقم 01</u>: بنية قاموس الأوزان.

يحتوي الحقل (وزن) كل حروف الوزن، يعطي حقل (قائمة الزوائد وضعية الحروف الزائدة ماعدا حروف الجذر (فعل)، أما حقل (الصنف) فهو يعطي الصنف النحوي المكافئ للوزن.

إن عملية التشفير (وضع مفتاح رقمي)[2] الخاص بالأوزان يسهل عملية البحث عن الأوزان كما سنوضحه لاحقا في كيفية البحث عن الأوزان.



الشاشة رقم (1): قائمة الأوزان المستخدمة في البرنامج

1- المرجع السابق، ص 67.

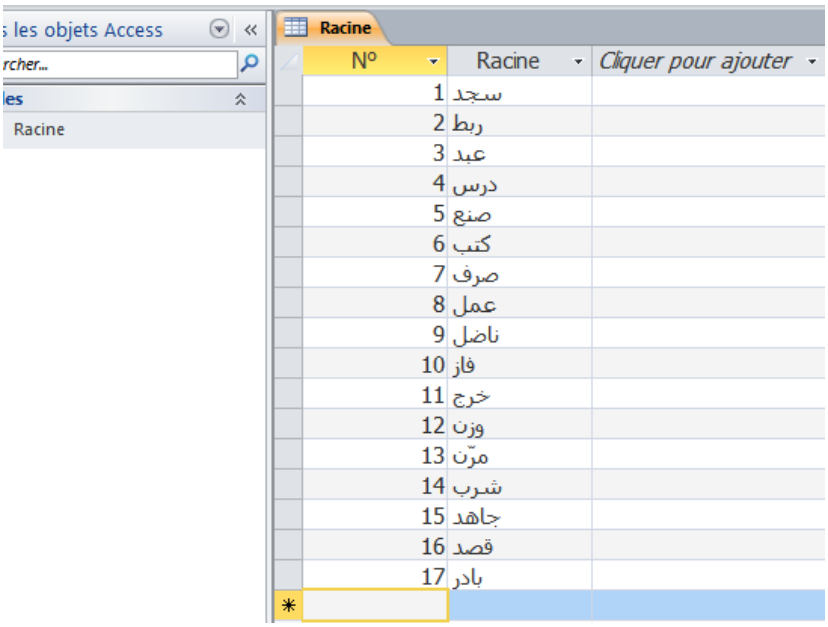
2 - Une Clé numérique (Code).

ب-قاموس الجذور:

في هذا القاموس تخزن جذور الكلمات العربية، مما يسمح بالتأكد من توافق الجذر والوزن لتفادي أي تقطيع خاطئ للكلمة المحللة.[1]

أصل
.....
سجد
.....
فتح

الشكل رقم 11: بنية قاموس الجذور.



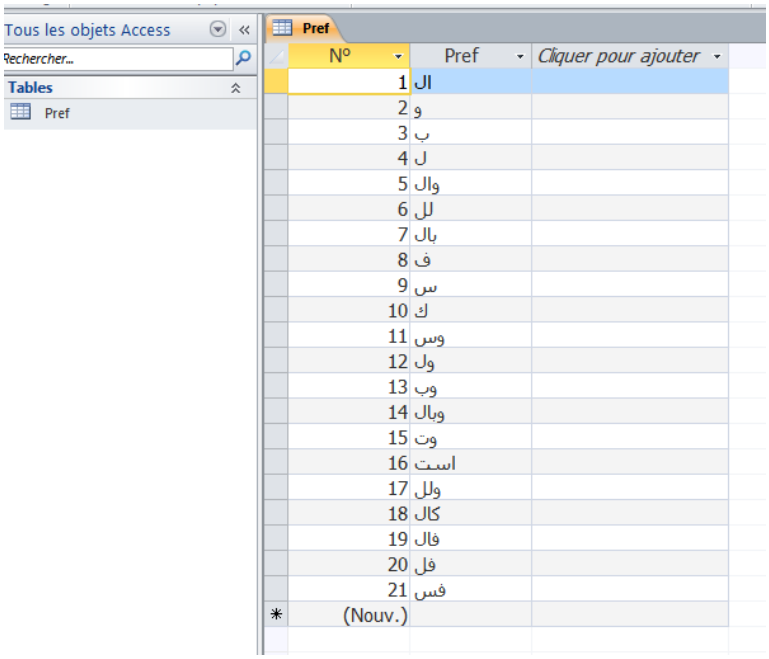
الشاشة رقم (2): قائمة الجذور المستخدمة في البرنامج

ج- قاموس الأدوات:

يمكن تحديد قاموس الأدوات بقائمة السوابق وقائمة اللواحق المدخلة فب البرنامج.

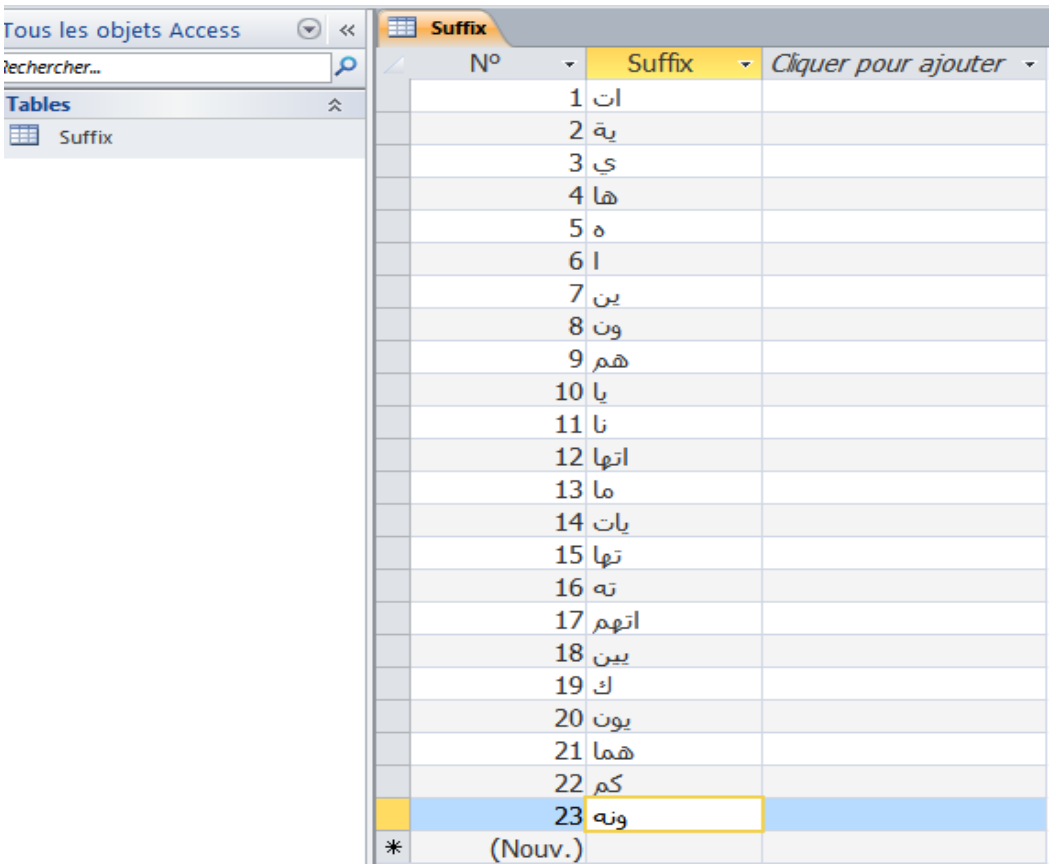
1 – Sadik Bessou, « Anayse de données textuelles pour la classification Automayique pour les Technique de text Mining application a la langue Arabe, p.67

أ/ قاموس السوابق:



الشاشة رقم (3): قائمة السوابق المستخدمة في البرنامج.

ب/ قاموس اللواحق:



الشاشة رقم (4): قائمة اللواحق المستخدمة في البرنامج

نعتبر الأداة كل كلمة تبقى ثابتة مهما تغير موضعها في الجملة ماعدا الكلمات الخاصة وأسماء العلم، مثل الضمائر، والحروف، كما أن هذا القاموس لا يحتوي إلا على الأدوات المنعزلة أما تلك التي تلتحق بالكلمات فلها معالجة خاصة أثناء عملية المعالجة في مرحلة إزالة اللواحق.[1]

الحرف	صنف اللاحق
في	إسم
قد	فعل
و	مشترك

الشكل رقم 12: بنية قاموس الأدوات.

الحقل (حرف) يحتوي على حروف الأداة، والحقل (صنف اللاحق) يحدد صنف الكلمة الموالية للأداة هل هو إسم أو فعل أو مشترك يحتمل الاختيارين.[2]

د- **قاموس الكلمات الخاصة:**

ليست كل الكلمات في اللغة العربية قابلة للتحليل إلى جذر ووزن إضافة إلى الأدوات يوجد نوع آخر يدعى الأسماء الخاصة والتي تتكون من الكلمات التي ليس لها أصل عربي وأسماء العلم، وبالتالي فإن إنشاء قاموس خاص بهذه الكلمات ضروري.[3]

جامد
أحمد
تلمسان
كمبيوتر
تكنولوجيا

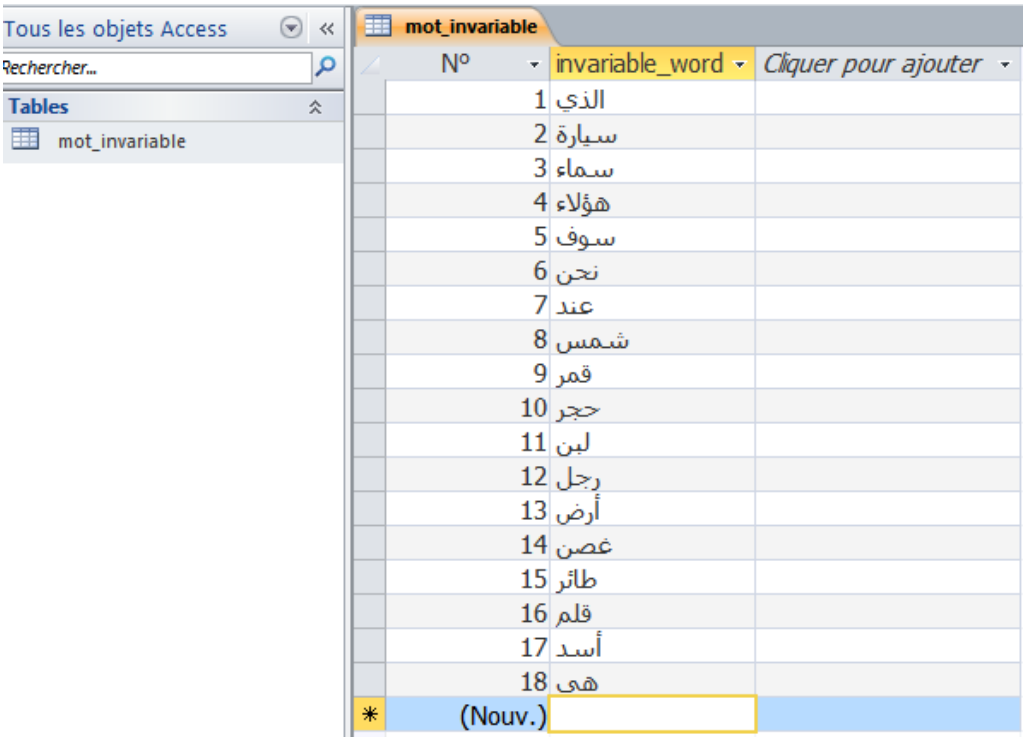
الشكل رقم 13: بنية قاموس الكلمات الخاصة.

1- المرجع السابق، ص 69.

2- المرجع نفسه، 69.

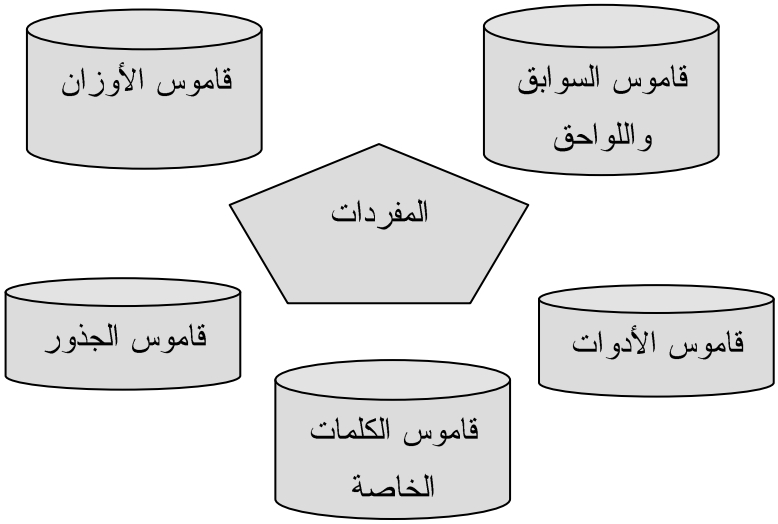
3- ينظر:

- Sadik Bessou, « Analyse de données textuelles pour la classification Automatique pour les Technique de text Mining application à la langue Arabe, p.69



الشاشة رقم (5): قائمة الكلمات الجامدة المستخدمة في البرنامج

الحقل الوحيد في هذا القاموس يحتوي على حروف الكلمات الخاصة ويمكن توضيح القواميس المستعملة في هذا النظام في الشكل الآتي:



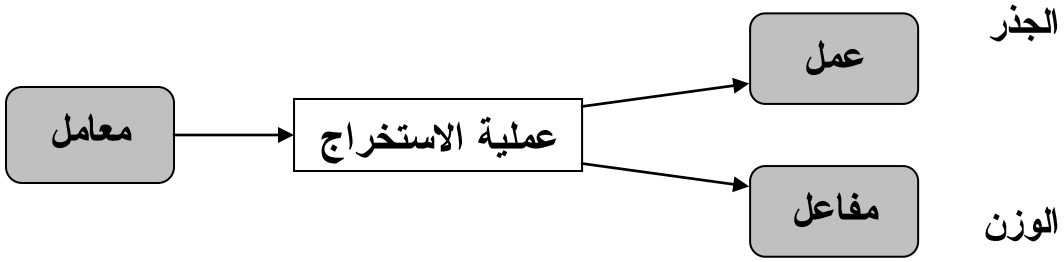
<u>الشكل رقم 14</u>: القواميس المستعملة من طرف النظام لعملية التحليل الصرفي

5- منهجية عملية النظام:

من المعلوم أن كل كلمة عربية ماعدا أسماء العلم، الأسماء الخاصة والأدوات، هي نتيجة ارتباط الجذر مع الميزان الصرفي خاص.

الكلمة = جذر + الوزن

يمكن أن نمثل استخراج الوزن والجذر من كلمة **(معامل)** بالمخطط الآتي:



الشكل رقم 15: استخراج الوزن والجذر.

للتحقق من أن الكلمة المدخلة تنتمي إلى مفردات النظام يكفي أن نجد لها وزنا وجذرا منا سبين وعليه فالنظام يرتكز على ثلاث قواعد أساسية، في مراحل التحليل الصرفي: التقطيع البحث بالمطابقة عن الجذر والوزن المناسبين، التفسير.[1]

أ- **التقطيع:**

يتمثل التقطيع أو ما يعرف بإزالة الزوائد في استخراج الأجزاء المختلفة للكلمة المدخلة (السابق + الجذر + اللاحق). سيعتمد النظام على الطريقة التي تتمثل في إلغاء السوابق واللواحق الزائدة في الكلمة[2] ثم نستمر في التقطيع حتى استخراج الجذر والوزن المناسبين.

تناولت العديد من البحوث موضوع التقطيع نذكر منها:[3] أعمال لعسكري[4] (1994)، زميرلي (1996) [5] خوجة (1999) [6] لاركي (2002) [7]، درويش (2003) [1] عبد الرحيم (2008).[2]

1- المرجع السابق، ص 70.

2 - DARWISH, K. And D.OARD. CLIR Experiments at Maryland For TERC 2002 : Evidence Combination for arabic English Retrievnl In TREC, Gaithersburg.

3- Abdefattah Hamdani, « Traitement automatique de la langue arabe », Actes du colloque juin 2006 ; Université Mohamed V Souissi Institut d'atudes et de recherches pour l'aribisation, janvier 2007, p.30, 91.

4 - LASAKRI Mohamed Taib, Sémantique du language naturel à travers unsystème Support de thésaurus, thèse d »état, 1994.

5 - ZEMIRLI Zouhir, un analyseur destiné à l'aide à la construction d'une base de données lexicales de la langue arabe, colloque International « langue situées, technologie et communication » IERA 1996.

6 - Khoja S, Orad, « Stemming Arabic text, technical report, coputing de partement, lancaster University, (1999).

7 - LARKEYL. S, BALLESTEROSL, And connel M, Improving Stemming for Arabic Information Retrieval : Light Steming and Co-occurrence Analysis, (SIGIR 2002). Tampere, Finland August (2002).

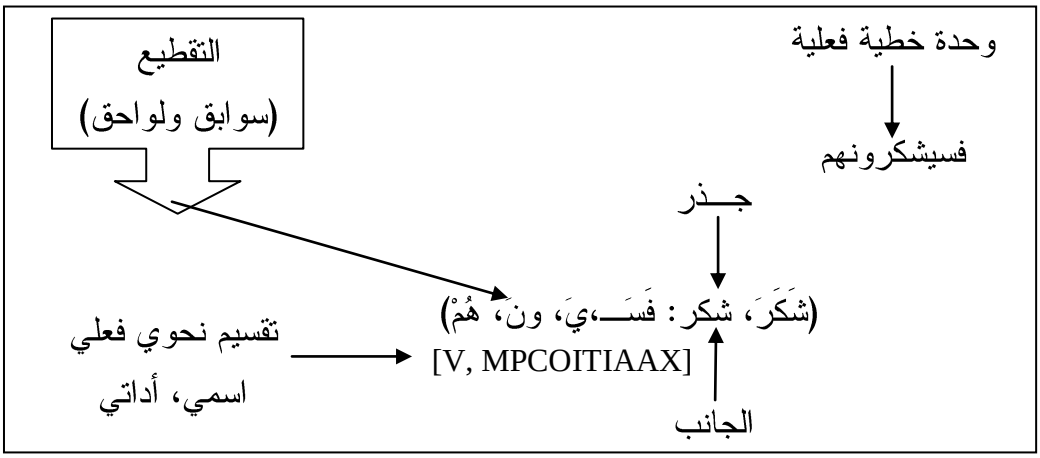
1 - DARWISH, K. probabilisitic Methods For Scarching OCR- Degraded Arabic Text, Doctoral disseration, university of maryland (2003).

2 - Mohamed El Amine ABDRRAHIM, Reconnaissance des Unité liguistiques Signifiantes thèse Doctoral, Spécialité Informatique Université Abou Bakr Belkaid Tlemcen 2007 – 2008, p.

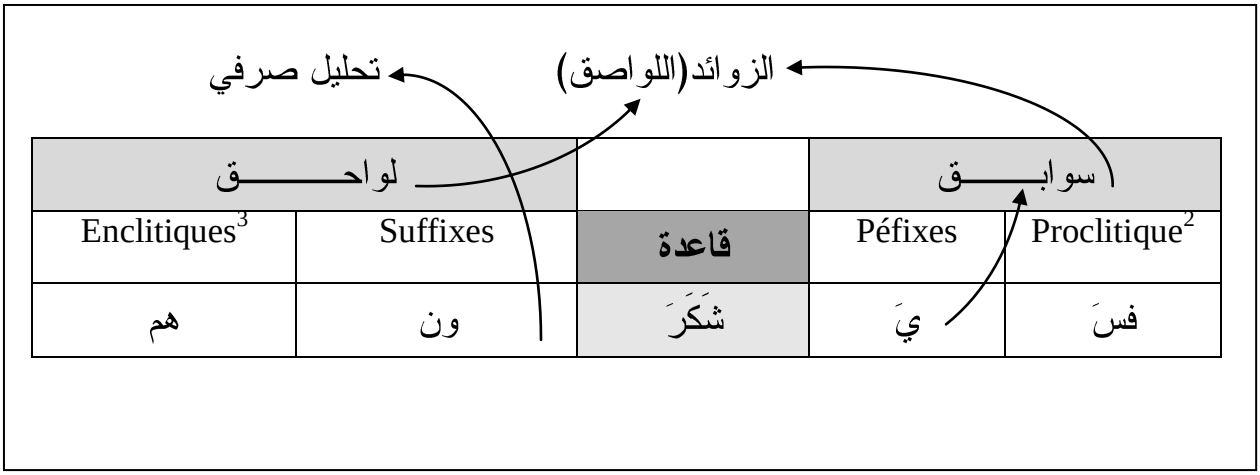
يتمثل مبدأ التقطيع في:

- تقطيع الكلمة إلى: سابق أولي + قاعدة أولى + لاحق أولي.
- تقطيع القاعدة الأولى إلى: سابق + قاعدة + لاحق.
- استخراج الجذر والوزن من القاعدة إذا كان ممكنا وإلا تعتبر الكلمة مفردة خاصة وبالتالي التقطيع ينحصر في المرحلتين الأوليتين.

ويمكن توضيح ذلك بالمثال الآتي:[1]



الشكل رقم 15: نتيجة التحليل الصرفي لكلمة: فَسَيَشْكُرُونَهُمْ



الشكل رقم 16: التعرف على الوحدة الخطية لكلمة: فَسَيَشْكُرُونَهُمْ.

1- د محمد الأمين عبد الرحيم، د غيتري سيدي محمد، "المعالجة الآلية للنحو العربي" الملتقى الوطني الثالث لقضايا النمو العربي الواقع والآفاق تيارت 21/20 أفريل 2009.

2- سابق أولي : Proclitique.

3- لاحق أولي: Enclitique.

2- **تطبيق الطريقة:**

يتم التحليل عبر ثلاث مراحل:

- إزالة اللواحق الأولية.
- إزالة اللواحق.
- استخراج الوزن والجذر المناسبين للكلمة إذا لم تكن كلمة خاصة أو أداة.

3- **التعرف على السوابق واللواحق الأولية:**

يقوم النظام بعد التعرف بإزالة الأحرف الزائدة من الكلمات العربية ورغم معرفة الزوائد في اللغة العربية من ضمائر وأحرف، إلا أنه توجد هناك أنظمة، تقتصر على إزالة بعض الأحرف وتترك البعض الآخر، ومنها من يقوم بإزالة معظم اللواحق.[1]

إن قائمة السوابق واللواحق الأولية في اللغة العربية محصورة ومعدودة فيمكننا أن نستعمل القائمة المقترحة من طرف الكثير من عناصر هذه القائمة لأجل استخراج الجذر.[2]

إن تقطيع الكلمة إلى سابق + قاعدة + لاحق لا يتمثل في البحث عن السابق و اللاحق في القائمة (الجدول رقم 15) من البداية إلى النهاية بل أيضا يقوم بالتحقق من التوافق بين الأطراف الموجودة في الكلمة المحللة (الجدول رقم 16).

السوابق الأولية[3]	ب	ك	ل	ف	س	أ	ال	كال	لل	فب	فس	فال	فك	فل	فلل	أف	أس	فبال	فكال	بال
اللواحق الأولية[4]	ه	ي	ك	هم	هن	هما	ها	كم	كن	كما	ني	نا								

الجدول رقم 15: قائمة السوابق واللواحق الأولية.

الجدول التالي: يوضح الحالات المختلفة للتوافق وعدم التوافق بين السوابق واللواحق الأولية.

تمثل الخانات التي تحوي x حالة عدم توافق بين السابق واللاحق إذا كان سابق (س) متوافق مع لاحق (ع) فهذا يعني التقطيع صحيح.

1- Nwersi A, Tahaghghis, Falk Scholer, « traitement des textes arabes pour l'indexation et la recherche d'information, CITALA 2007, Rabat Maroc, (2007).

2- CHEN A. And Gey F ; « Building an Arabic Stemmer For Information Retrieval », Proccedings of the Eleventh text Retrieval Conference (TREC 2002) National Institute of Strandard and technology, Nov 18/22 (2002).

اللواحق الأولية / السوابق الأولي	ه	ي	ك	هم	هن	هما	ها	كم	كن	كما	ني	نا
ب											x	
ك											x	
ل												
ف												
س		x										
أ		x										
ال	x	x	x	x	x	x	x	x	x	x	x	x
بال	x	x	x	x	x	x	x	x	x	x	x	x
كال	x	x	x	x	x	x	x	x	x	x	x	x
للـــ	x	x	x	x	x	x	x	x	x	x	x	x
فبـــ											x	
فســـ		x										
فال	x	X	x	x	x	x	x	x	x	x	x	x
فك											x	
فل											x	
فلل	x	x	x	x	x	x	x	x	x	x	x	x
أف												
أس		x										
فبال	x	x	x	x	x	x	x	x	x	x	x	x
فكال	x	x	x	x	x	x	x	x	x	x	x	x

<u>الجدول رقم 16</u>: التوافق بين السوابق واللواحق الأولية.

4- **طريقة التحقق من التوافق بين أجزاء الكلمة:**

بعد استخراج السابق واللاحق من الكلمة المحللة هتان الوحدتان تلحقان ببعضهما للتحقق من توافقهما عن طريق جدول يحتوي الوحدات غير المتوافقة متلصقة ببعضها (شكل17) إذا كانت الوحدتان لا تنتميان إلى الجدول فهذا يعني أن السابق يتوافق مع اللاحق[1] ويعمل النظام على تطابق السوابق اللواحق المستخرجة بتطابقها مع جدول السوابق واللواحق الأولية المعد كقاعدة معطيات.

فبـــ + هم **فبهم** ⇐ هذه السلسلة من الحروف متوافقة حسب الجدول ⇐ إذن فالتقطيع صحيح.

السلسلة
بني
.....
كالهما
.....
فكالنا

الشكل رقم 17: قائمة السلاسل غير متوافقة.

ملاحظة:

يحتوي الجدول السلاسل غير المتوافقة لأن العدد أقل.

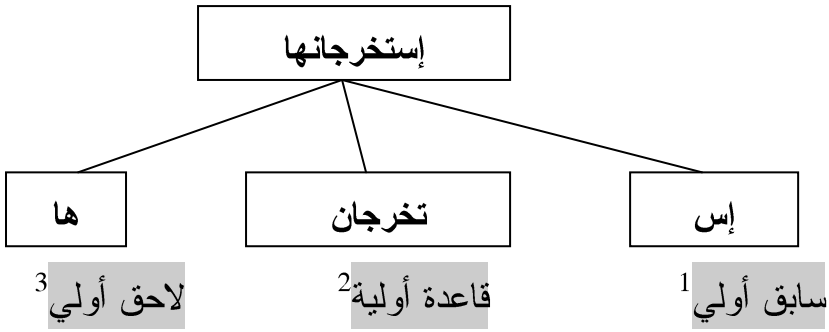
5- **مبدأ التحليل:**

أثناء تقطيع الكلمة إلى سابق + قاعدة + لاحق: يقوم البرنامج بتحديد أصول سابق وأطول لاحق في الكلمة ثم يقارنها مع جدول السوابق أو اللواحق لمعرفة توافقها من عدمه، فإذا كان التوافق يتطابق، فهذا التقطيع يعتبر مقبولا ويحزن في جدول نتائج هذه المرحلة، وإذا كان التقطيع مرفوض، نمر إلى تقطيع آخر محتمل لمعالجة جميع الحالات الممكنة.[2]

مثال: تقطيع كلمة استخرجانها تقطيع الكلمة حسب الطريقة التالية:

1 - Bessous, Louali M, Refoufi, A, KADEM. Z, Touahria M, « Unsystème de lemmatisation pour les applications de TALAN.-Traitement automatique de la langue arabe colloque Internationale CITALA 07, 18-19 juin 2007.
Université Mohammed V, Souissi, Institut d'études et de Recherches pour l'arabisation, juin 2007, p.45.

2- المرجع السابق، ص 45.



الشكل رقم 18: تقطيع الكلمة إلى سابق أولي وقاعدة أولية ولاحق أولي

نلاحظ أن التقطيع السابق صحيح.

6- **التعرف على السوابق واللواحق:**

مبدأ هذه المرحلة عمليا مماثل للمرحلة السابقة في السوابق واللواحق الأولية ماعدا في هذه المرحلة نستعمل قائمة السوابق واللواحق (جدول 17) ونستعمل جدولا آخر لتوافق السوابق واللواحق وهذا ما يوضحه الجدول الآتي[4]:

السوابق	إ	ن	ي	ت	ا											
اللواحق	ات	ية	ة	يات	نا	ت	تما	تم	تن	ن	ين	ان	ون	وا	ا	ي

الجدول رقم 17: السوابق واللواحق.

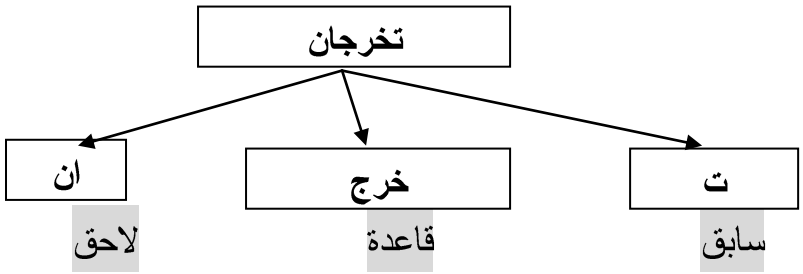
الجدول التالي هو جدول التوافق بين السوابق واللواحق.

اللواحق / السوابق	ات	ية	ة	يات	نا	ت	تما	تم	تن	ن	ين	ان	ون	وا	ا	ي
ا	x	x	x	x	x	x	x	x	x	x	X	x	x	x	x	x
ت	x	x	x	x	x	x	x	x	x							
ي	x	x	x	x	x	x	x	x	x		X					x
ن	x	x	x	x	x	x	x	x	x	x	X	x	x	x	x	x
إ	x	x	x	x	x	x	x	x	x	x	X	x	x			

الجدول رقم 18: جدول التوافق بين السوابق واللواحق.

1 - Proclitique.
2 - Base 1.
3 - Enclitique.
4 - Traitement automatique de la langue arabe colloque Internationale CITALA 07. p.47.

مثال: تقطيع القاعدة السابقة (تخرجان) يعطي:



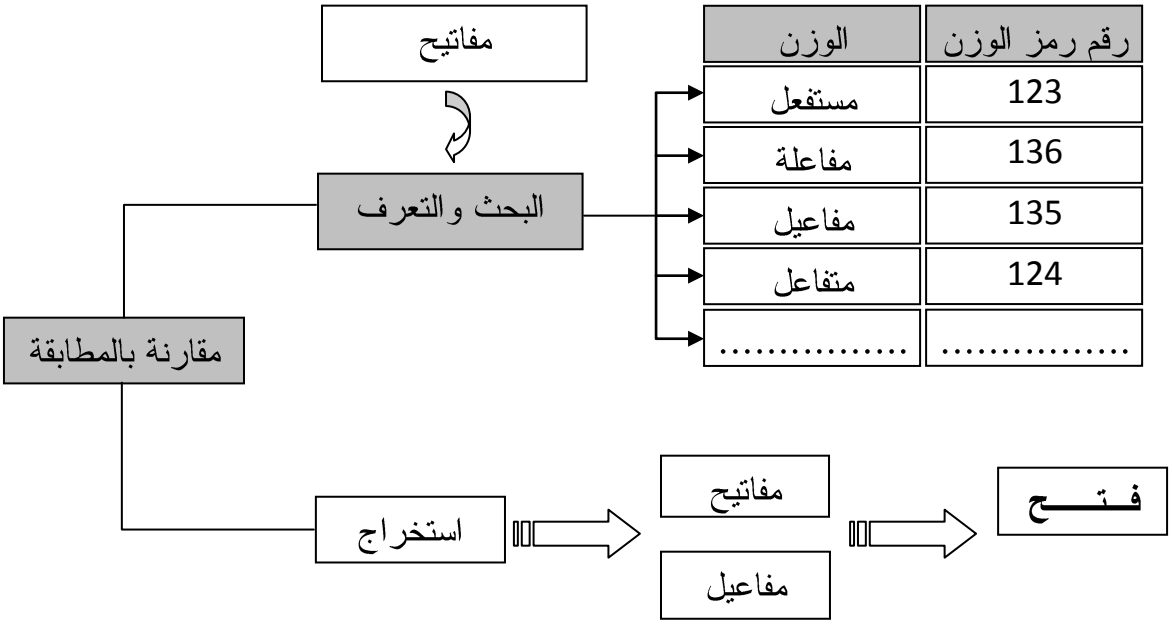
الشكل رقم 19: تقطيع القاعدة الأولى إلى سابق ولاحق وقاعدة.

7 – البحث عن الوزن والجذر:

أ- البحث عن الوزن:

مبدأ طريقة البحث بسيط من أجل كلمة (س) والوزن (ع) في قاموس الأوزان إذا كان عدد حروف الوزن (ع) مساويا له في الكلمة (س) وكل الحروف الموافقة لحقل قائمة الزوائر في جدول الأوزان توجد في نفس الوضعيات في الكلمة (س).[1]

وتتم آلية التعرف واستخراج الجذرمن خلال الشكل الآتي:



الشكل رقم 20: البحث عن الجذر والوزن المناسبين والتعرف عليهما

1- Analyse de données textuelles pour la classification Automatique pour les Techniques de text Mining application a la langue Arabe, p.80.

ب- **آلية التعرف:**

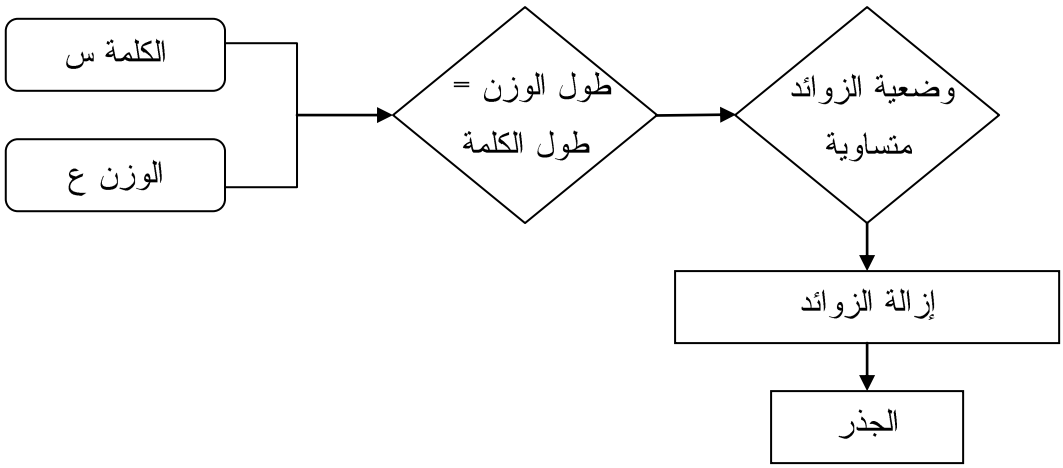
مثال: الكلمة المدخلة: مساجدٌ.

يقوم البرنامج بتفقد كل خانات الأوزان التي طولها يساوي طول القاعدة حتى الوصول إلى الوزن (فاعل)، يحتوي الحقل قائمة الزوائد (السوابق واللواحق) بقاموس الأوزان المكافئ للوزن (مفاعل) رقم 2، والحرف (أ) يوجد في نفس الوضعية في الكلمة (مساجد) إذن فالوزن (مفاعل) هو الوزن المناسب.

سـ|ـاجـ|ـد ⇔ فـ|ـاعـ|ـل[1]

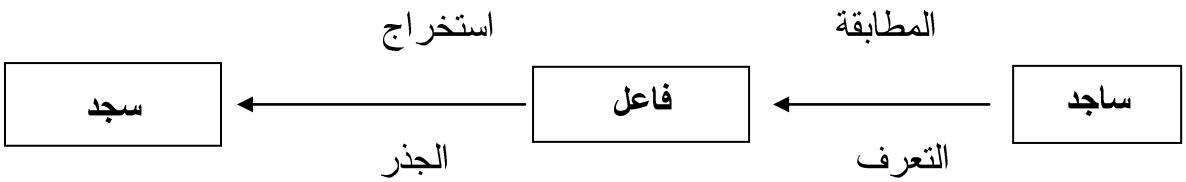
8- البحث عن الجذر:

بعد تحديد الوزن يقتصر استخراج الجذر على حذف كل الحروف الزائدة المحددة في حقل الحروف الزائدة للوزن المناسب، وهذا ما يوضحه شكل الآتي[2]:



الشكل رقم 21 : استخراج الجذر و الوزن

مثال: إزالة الحرف رقم (2) من الكلمة **(ساجد)** يعطي الجذر **سجد.**



الشكل رقم 22 : التعرف على الجذر بواسطة الوزن.

1- المرجع السابق، ص 80.

2- المرجع نفسه، ص 79.

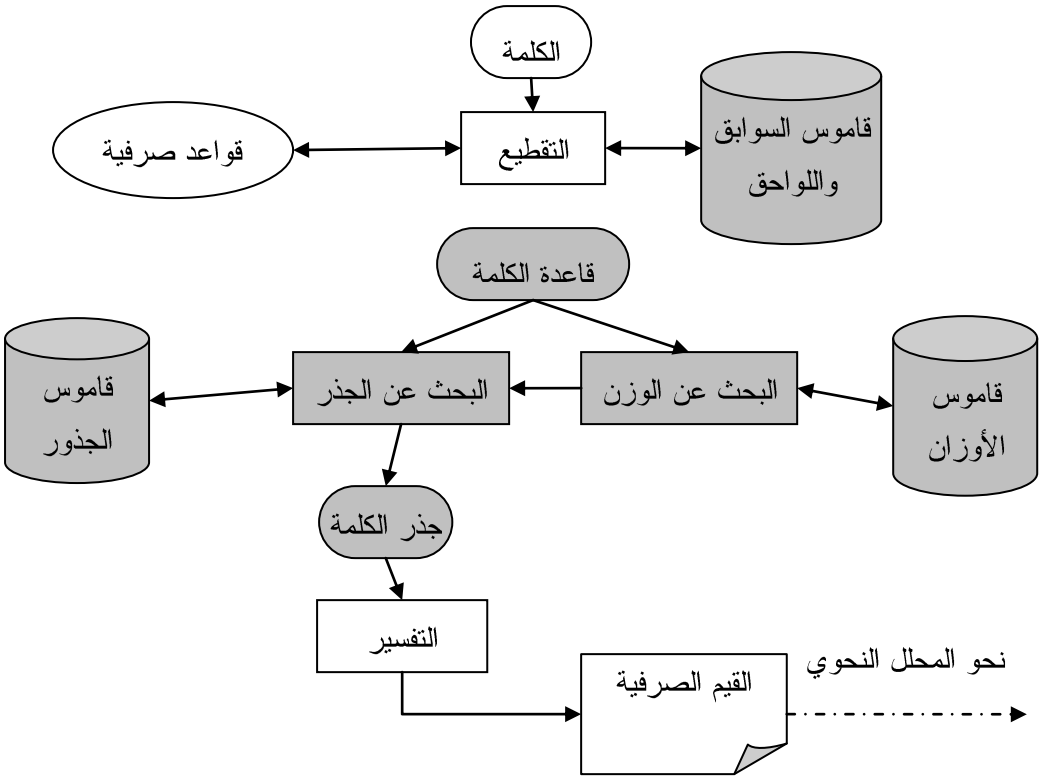
9- تفسير عمل برنامج المعجم الإلكتروني للمعالجة الآلية للغة العربية : "الأمين":

بعد استخراج الجذر والتأكد من وجوده في قائمة الجذور العربية، نمر إلى إنتاج القيم الصرفية للكلمة المحللة هذه القيم تستخرج من خلال جدول الأوزان هذا الأخير يقوم بإرفاق قيم نحوية لكل ميزان صرفي، هذه القيم تستعمل من طرف المحلل النحوي.

مثال: كلمة (مكتوب) يكافئها الوزن (مفعول) في الحقل (الصنف) في قاموس الأوزان نجد الصنف (إسم) إذن نرفق للكلمة (مكتوب) الصنف (إسم).

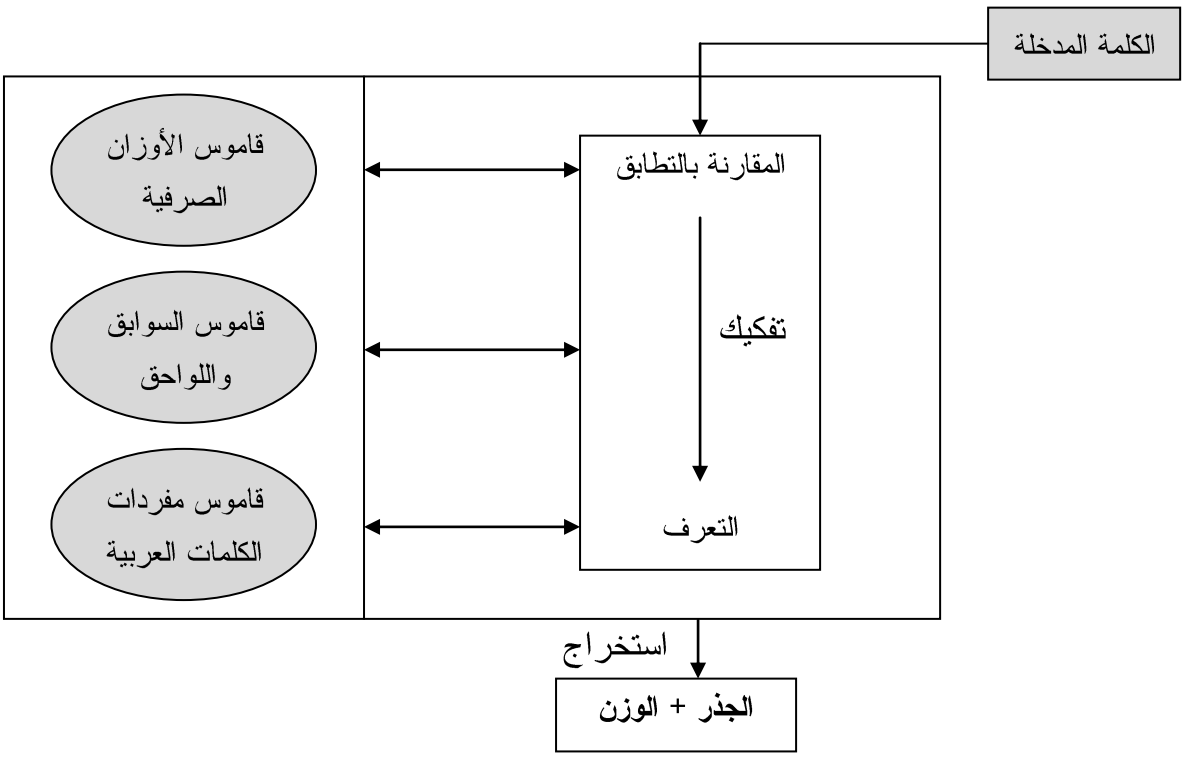
مكنتنا هذه الطريقة من تحليل الكلمات العربية تحليلا صرفيا من خلال إنجاز برنامج آلي يقوم بتقطيع الكلمة إلى قاعدة وزوائد ثم استخراج الجذر والوزن المناسبين للكلمة المحللة.

فالبرنامج يقوم بالتعرف على الكلمات بطريقة ذكية عن طريق قواعد لغوية ودون سابق معرفة بهذه الكلمات ثم إرفاق الكلمة بقيمها الصرفية.كما هو موضح في الشكل الآتي:



<u>الشكل رقم 23</u>: مرا حل استخراج الجذر والوزن.

ويمكن أن نلخص آلية عمل برنامج المحلل الصرفي للمعجم الإلكتروني للمعالجة الآلية للغة العربية في المخطط الآتي:



<u>الشكل رقم 24</u>: مخطط يوضح آلية عمل المعجم الإلكتروني للمعالجة الآلية للغة العربية باستخدام قواميس: الأوزان الصرفية، والسوابق واللواحق، و جذور الكلمات العربية.

10- عرض بعض الشاشات للبرنامج كأمثلة توضيحية من المعجم الإلكتروني :

برنامج المعجم الإلكتروني للمعالجة الآلية للغة العربية: "معجم الأمين[1]" مفردات القرآن الكريم*أنموذجا*

لتوضيح عملية استخراج الجذر والوزن المناسبين كما وضحناه سابقا نستعرض بعض الشاشات كأمثلة حية من البرنامج تبين مراحل آلية التحليل للكلمة المدخلة : **<u>المساجد</u>**

1- تمت تسميت برنامج المعجم الإلكتروني للمعالجة الآلية للغة العربية "بمعجم الأمين"لانفراده بخاصية فريدة من نوعها، فالبرنامج يقترح طريقة فعالة لإنشاء واستغلال معجم إلكتروني للمعالجة الآلية للغة العربية. تستخدم هذه الطريقة نظاما آليا لاستخراج جذور الكلمات العربية اعتماداً على التقنيات اللغوية للمعالجة الآلية للغة العربية، كما نبين من خلالها أهمية التعرف على الجذر في مراحل التحليل، ترتكز هذه المقاربة على مفهوم الوزن أو الميزان الصرفي (أحد نقاط القوة في الصرف العربي) وفق تراكيب خاصة تسمى "التفعيلات"، الجانب الإيجابي في هذه الطريقة أنها لا تستعمل معاجم الكلمات، بل تقوم بالتعرف على مختلف الكلمات بطريقة ذكية عن طريق استخراج الجذر والوزن المناسبين لكل كلمة محللة.

شاشات عرض برنامج المعجم الإلكتروني للمعالجة الآلية للغة العربية: "معجم الأمين":

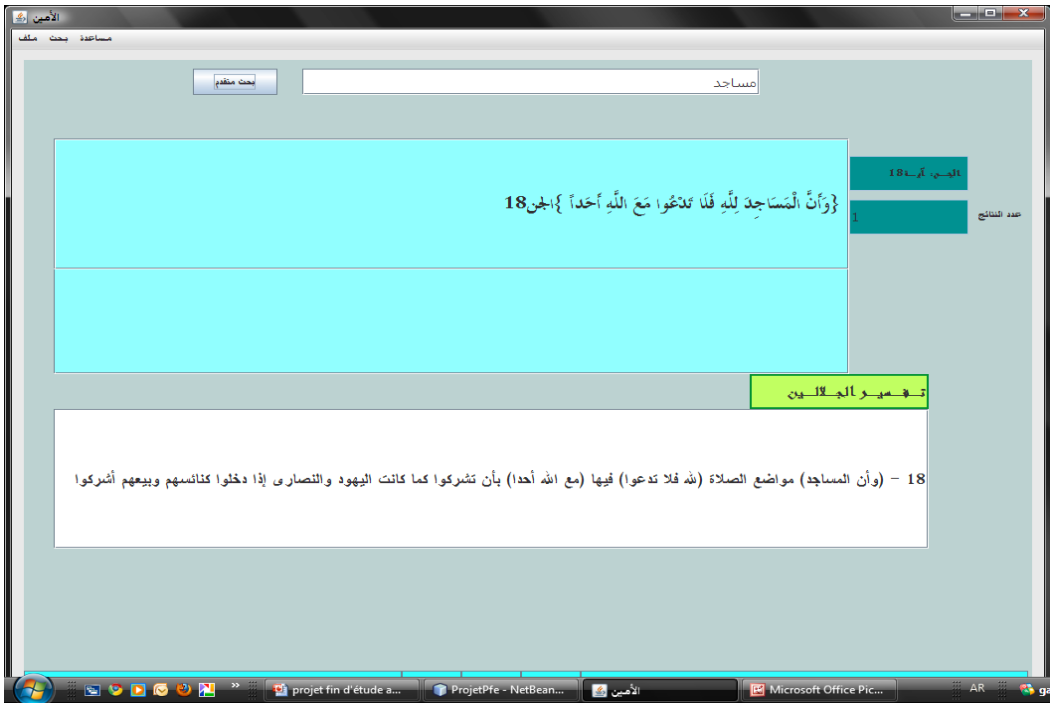
الشاشة الأولى: نافذة برنامج معجم الأمين

الشاشة الثانية:

توضح شاشة البرنامج عملية ادخال كلمة مساجد وعند النقر على موضع البحث بالأوزان وبعد عملية البحث تم التعرف على الوزن المناسب للكلمة المدخلة

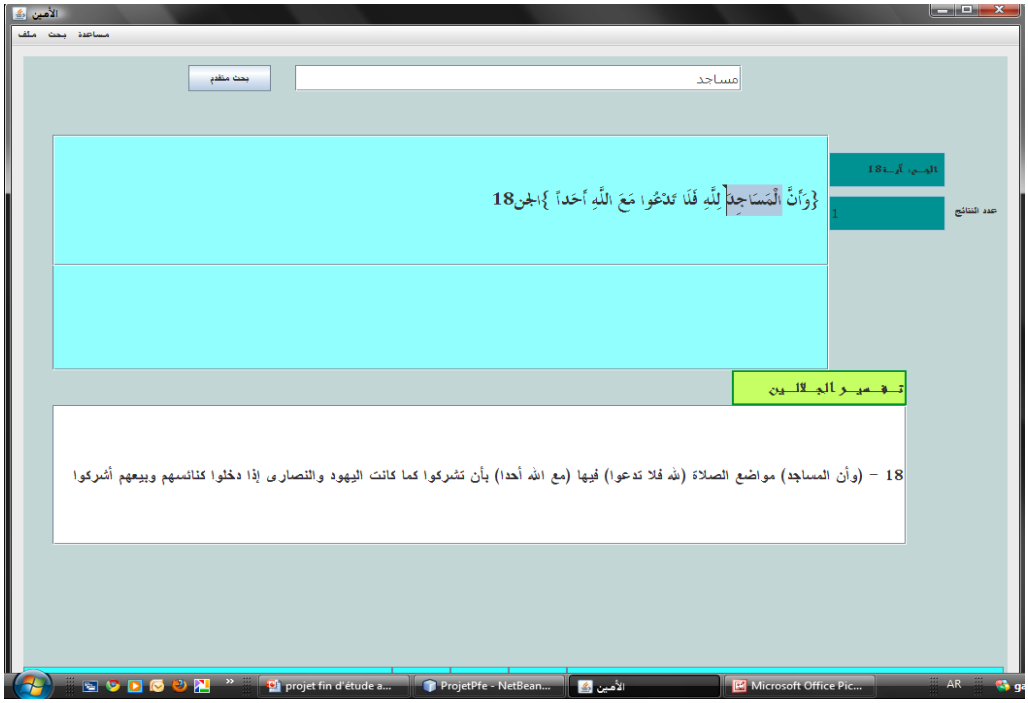
الكلمة المدخلة : **مساجد** – الوزن المناسب : **مفاعل**

الشاشة الثالثة: شرح الكلمة المدخلة

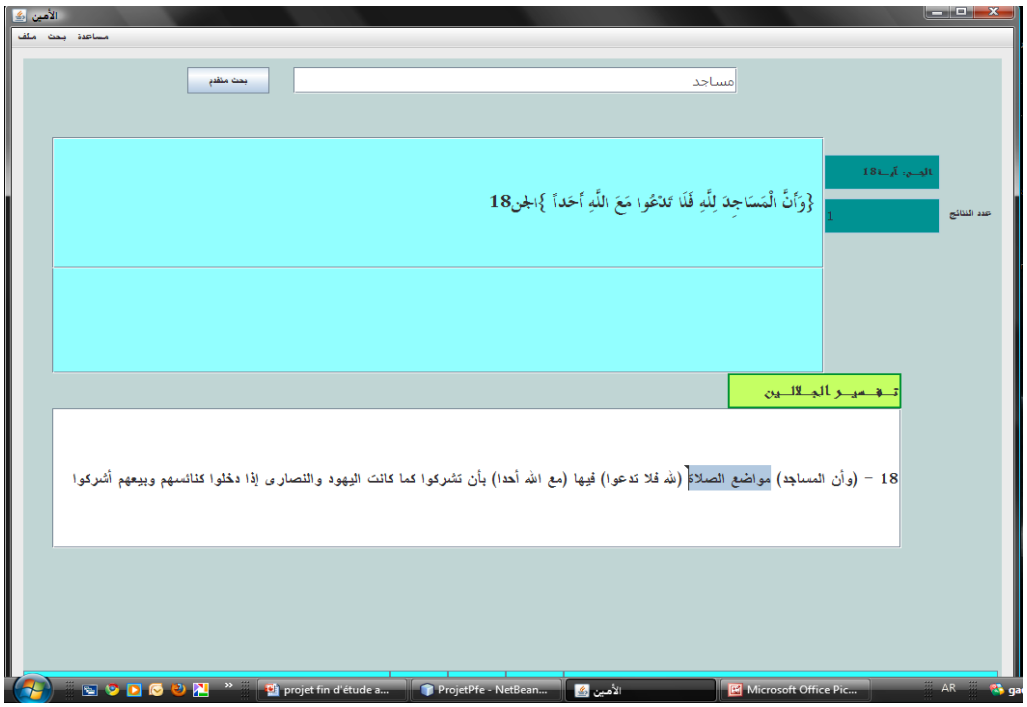


الشاشة الرابعة : عملية الشرح

بعد استخراج الجذر والوزن المناسبين للكلمة المدخلة، يأتي الدور والغاية الأسمى للمعجم ألا وهي شرح المعنى ، وهو ما توضحه الشاشتين الرابعة والخامسة من خلال المثال الموضح على نافذة البرنامج.



الشاشة الخامسة: عملية الشرح



سابعا- نتائج الدراسة التطبيقية:

1- قدمت الدراسة التطبيقية عرض وافي لأهم خطوات بناء معجم إلكتروني للمعالجة الآلية للغة العربية.

2- وأهم النّقاط ونتائج هذه الدراسة التي يمكن الخروج بها من هذا العرض هي الطريقة الآلية التي يمكن الحصول بها على الجذر.

3- إن العرض المقدّم اعتمدت فيه على عرض طريقتين في كيفية استخراج جذور الكلمات العربية كأهم خطوة لبناء معجم إلكتروني للمعالجة الآلية للغة العربية؛ من خلال عرض أنموذجين لقياس مقدار الكفاءة التشغيلية لكل منهما وللوصول إلى التوصية باستخدام أفضلهما في المستقبل لتطوير المعاجم الإلكترونية للمعالجة الآلية للغة العربية أو محركات البحث العربية التي تعاني من قلة المرونة والمحدودية في التعامل مع الكلمات المفهرسة من حيث الاسترجاع.

4- قدم العرض شرح وافي عن كل ما يتعلق بإجراءات بناء معجم إلكتروني للمعالجة الآلية للغة العربية باستخدام نظام آلي لاستخراج جذور الكلمات العربية بالاعتماد على التقنيات اللّغوية للمعالجة الآلية للغة العربية. يرتكز العرض على طريقة مفهوم الوزن (الميزان الصرفي) (أحد نقاط القوة في الصرف العربي).

5- الجانب الإيجابي في هذا البرنامج أنّه يقوم بالتعرف على مختلف الكلمات المدخلة بطريقة ذكية عن طريق استخراج الجذر والوزنين المناسبين لكل كلمة محللة دون استعمال معاجم الكلمات (بالنسبة للطريقة الثانية).

6- لقد مكنتنا هذه المقاربة من إنشاء معجم إلكتروني للمعالجة الآلية للغة العربية باستخدام طريقتين مختلفتين لاستخراج جذور الكلمات العربية لكن كلتا الطريقتين اعتمدت فيهما على مفهوم الوزن (الميزان الصرفي) للحصول على الجذر والوزن المناسبين:

أ-الطريقة الأولى أو المقاربة تعرف بـــ **المقارنة التقريبية (النوافذ المنزلقة).**

ب- الطريقة الثانية: تعرف بـــ التشذيب "Stemming":

الجذر+ الوزن= القاعدة Stem

- التشذيب (Stemming)[1]: هو عملية إزالة كل سوابق الكلمة ولواحقها للحصول على الجذر، أو الأصل. فهو طريقة رياضية تجمع كل الكلمات التي تتقاسم نفس الأصل وتملك بعض العلاقات الدلالية.

- تصنف خوارزميات التشذيب العربية طبقًا لمستوى التحليل المطلوب: على أساس الأصل أو على أساس الجذر.

- لماذا المعجم الإلكتروني للمعالجة الآلية للغة العربية؟.

من خلال ما تقدم من نتائج ومميزات هذا البرنامج يمكن أن نضع تعريفًا للمعجم الإلكتروني للمعالجة الآلية للغة العربية.

1- **فالمعجم الإلكتروني للمعالجة الآلية للغة العربية:** يشبه المعجم الورقي إلا أن طريقة استعماله تتم بواسطة الآلة (الحاسوب) فالمعجم يستعمله الحاسب فقط، فهو غير موجه للاستعمال البشري، والهدف منه هو الاستعمال لغاية البحث أو الترجمة مثلاً.

2- يمكن استعمال هذا المعجم لتطوير محركات البحث والمحللات والترجمة الآلية وأنظمة استرجاع النصوص.....الخ

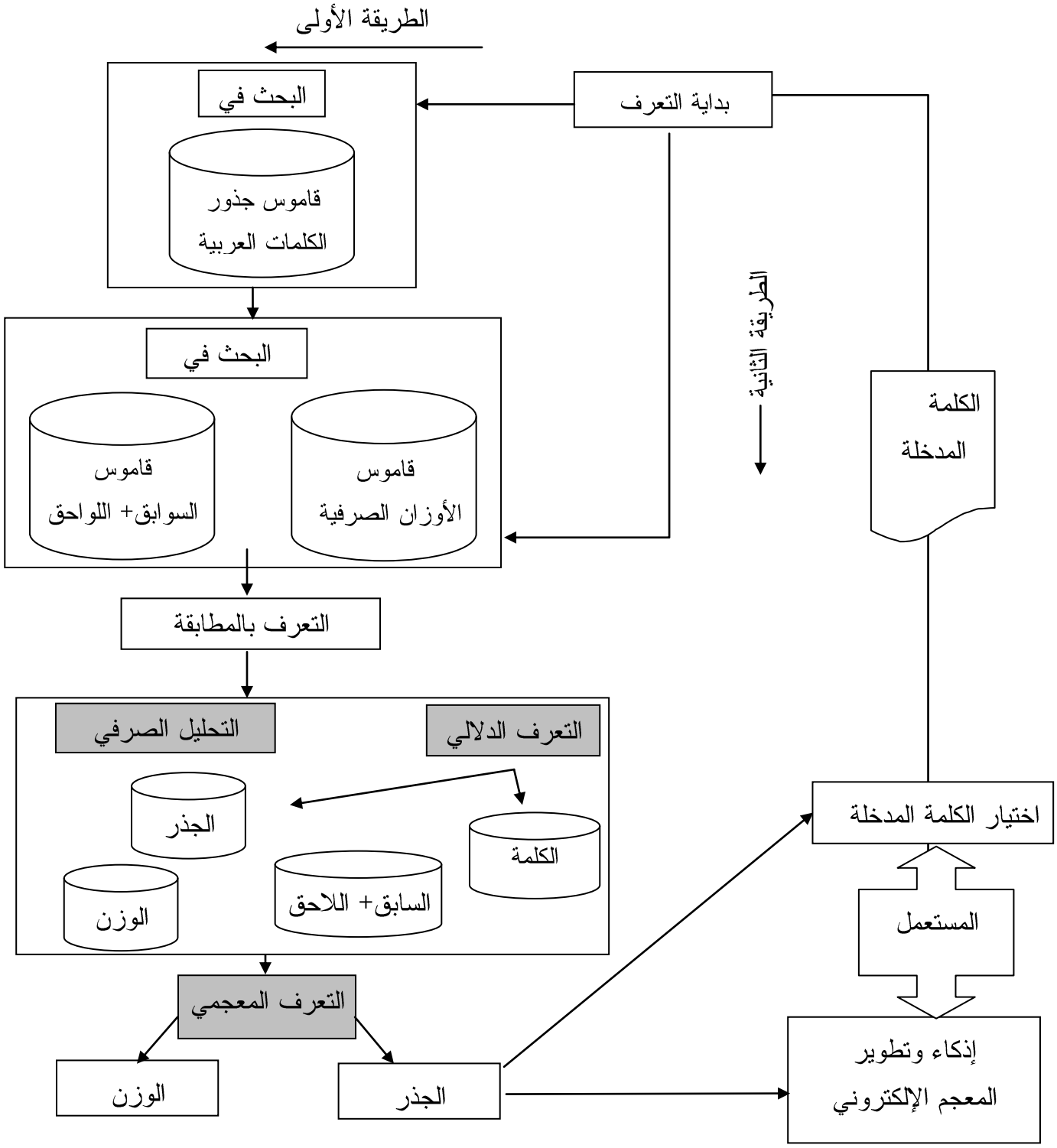
وهنا يكمن الاختلاف بين الميزة التي ينفرد بها هذا المعجم الإلكتروني للمعالجة الآلية للغة العربية ومختلف المعاجم الإلكترونية الموجودة.

3- تمت ميكنة المعجم على هيئة قائمة من الجذور والصيغ الصرفية+ قائمة السوابق واللّواحق الممكن تطبيقها على المعجم الآلي وفق معطيات صرفية اللاّزمة للتطبيقات المعالجة الآلية للغة العربية، ويعدّ هذا المعجم الإلكتروني للمعالجة الآلية للغة العربية عُنصرًا أساسيًا لجميع المعالجات اللّغوية الآلية كونهُ مصدرًا لا غنى عنهُ لدراسة الإنتاجية الصرفية للغة العربية.

وصفوة القول: إنّ لهذه الطريقتان في إنشاء واستغلال المعجم قد تفيدنا أيما فائدة في معالجة اللّغة العربية آليًا لأنّ حصولنا على الجذر والتفعيلة يساعدان كثيرًا في الحصول على معنى الكلمة إضافة إلى وظيفتها النحوية في كثير من الحالات، فمن خلال الوزن (التفعيلة) يمكن معرفة جزء كبير من المعنى المصاحب للكلمة، فمثلاً (مَفْعَلة) يدُل أحيانًا على مكان يكثر

1-Un système de lemmatisation pour les application de TALN.

فيه الشيء كقولك (مكتبةٌ) وهو مكان يدلُ فيه الكثير من الكتب أو (مسمكة) وهو مكان يكثر فيه السّمك أو الوزن (فاعل) يدل على من يقوم الفعل كقولك (ساجدٌ) والشخص الّذي سيجدُ وهذا ما يمكن استغلاله في المعالجة الدلالية للغة (الشكل25):



الشكل رقم 25: آلية التعرف على الكلمات العربية عن طريق استخراج الجذر والوزن المناسبين في المعجم الإلكتروني للمعالجة الآلية للغة العربية

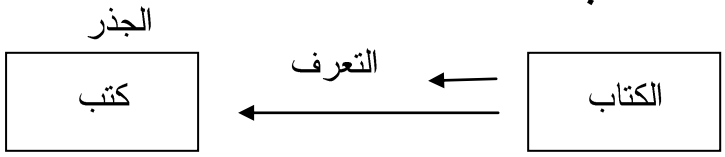
وفي الأخير؛ يمكننا القول أن عملي مكنّني من إثبات أهمية بعض الخصوصيات التي تتميز بها اللّغة العربية عن غيرها؛ وضرورة التنبيه إلى معالجة هذه اللّغة انطلاقًا من هذه الخصوصيات بعيدًا عن المحاولات المختلفة التي تهدف إلى استعمال طرق أخرى تخصُ اللّغات الأجنبية بصفة خاصة.

ملاحظة هامة:

من أبرز النتائج العلمية والتقنية لهذه الدراسة أنّنا قُمنا بتجريب بعض البرامج واستخدام آلية الطريقة التي اعتمدنها في إعداد برنامج المعجم الإلكتروني للمعالجة الآلية للغة العربية وتطبيق مبدأ البرنامج في آلية التعرف و استخراج جذور الكلمات.

- إدخال الكلمة للحصول على الجذر

مثال: الكلمة المدخلة: **الكتاب**



فقمنا بنفس العملية باستخدام برنامج المصحف الإلكتروني للقرآن الكريم.

من باب البحث عن المعنى الدلالي للكلمة المدخلة وترددها في المصحف الإلكتروني (البرنامج) لكن بعد تجريب بعض الكلمات لاحظنا وجود أخطاء في البرنامج في طريقة التعرف على الكلمة المدخلة. ويمكن توضيح ذلك من خلال تقديم شروحات أكثر بعرض بعض الكلمات التي أدخلناها في البرنامج كالآتي:

مثالاً: بعد إدخال كلمة: طَعَنَ في البرنامج حصلنا على النتائج التالية:

﴿ **وَقَالُوا سَمِعْنَا وَأَطَعْنَا غُفْرَانَكَ رَبَّنَا وَإِلَيْكَ المَصِيرُ**﴾[1].

﴿ **فَإِنْ أَطَعْنَكُمْ فَلاَ تَبْغُوا عَلَيْهِنَّ سَبِيلاً إِنَّ اللهَ كَانَ عَلِيًا كَبِيرًا**﴾[2].

﴿ **وقَطَعْنَا دَابِرَ الّذِينَ كَذَّبُوا بِآيَاتِنَا ومَا كَانُوا مُؤْمِنِين**﴾[3].

﴿ **لَوْ اسْتَطَعْنَا لَخَرَجْنَا مَعَكُمْ يَهْلِكُونَ أَنْفُسَهُمْ واللهُ يَعْلَمُ إِنَّهُمْ لَكَاذِبُونَ**﴾[4].

ويمكن توضيح ذلك من خلال شاشتي المصحف الإلكتروني:

1- البقرة، الآية (285).

2- النساء، الآية (34).

3- الأعراف، الآية (72).

4- التوبة، الآية (42).

الشاشة الأولى:

الشاشة الثانية:

وهناك أمثلة أخرى قمنا بتجريبها في برنامج المصحف الإلكتروني وقد لاحظنا نفس الخطأ يتكرر في طريقة التعرف على الكلمة المدخلة.

مثال: قمنا بتجريب الكلمات التالية: حجّ، عام، تاب؛ وقد اكتفينا بهذا القدر من الأمثلة من باب الاختصار.

إنّ أهمية هذه الملاحظة تكمن في أهمية الدراسة التطبيقية التي قمنا بها من خلال إعداد برنامج المعجم الإلكتروني للمعالجة الآلية للغة العربية وأهمية التعرف وكيفية استخراج الجذر (جذور كلمات اللّغة العربية).

والنتيجة التي يمكننا الخروج بها هي أن بواسطة النظام الآلي للمعجم الإلكتروني للمعالجة الآلية للغة العربية يمكن استخدامه لتطوير برامج مختلفة ولما استخدامه في تطوير المصحف الإلكتروني للقرآن الكريم من خلال ضبط آلية التعرف واستخراج الكلمات المدخلة لمفردات القرآن الكريم؛ زيادة على استعمالات أخرى كتطوير برامج المحللات الصرفية، المدققات الإملائية، محركات البحث العربية....الخ.

تنبيه:

لقد بينت هذه الدراسة التطبيقية أن أهمية قاعدة المعطيات تتجلى في بقائها مفتوحة وقابلة للإستغلال والتطبيق في الحث العلمي في أي وقت، وهو الهدف من بنائها للمساهمة في وضع لبنات من الأبحاث اللغوية في طريق مستقبلي موجه نحو الوصف الشامل للغة العربية لدمجها في الدراسات الحاسوبية .

خـــاتـــمـــــة

خـــــاتمة

نتائج وتوصيات:

قصارى القول: إن هذه الرحلة العلمية التي ألزمتني حصاد قراءات عديدة لأمَّات الكتب في حقول البحث اللّغوي عامة و والمعالجة الآلية للغة العربية كموضوع للدراسة النظرية وحوسبة المعجم العربي خاصة كتطبيق للمجال النظري ، فقد حاولتُ في هذا التطواف الشاق الجمع بين مجالين من المعرفة كمتطلبات لهذه الدراسة؛ فقد اتخذت من اللّغة منطلقًا أساسيًا لها لتأتي المعالجة اللّغوية الآلية كركيزة تقام عليها تطبيقات الدراسة النظرية. وعلى هذا النهج كان دأب الباحثين في هذا الميدان المعرفي.

لقد كان اجتهادي في هذه الدراسة اختزال جملة من النتائج التي ألخص فيها فصول بحثي المتواضع بل القاصر في النقاط التالية:

- إنّ هذه الدراسة قامت عن قناعة راسخة بأنّ العمل في حقل اللّسانيات الحاسوبية يتطلب إلمامًا متوازنًا بالنّواحي اللّغوية والنّواحي التقنية المرتبطة بمعالجة اللّغة العربية آليًا منظورًا إليها نظرة لسانية كونها تتناول اللّغة في محيطها العام.

- أنْ أضع بين يدي قارئ هذا البحث صورة عن أهمية حوسبة المعجم العربي في ظلّ التحديات الكبرى التي تعرفها اللّغة العربية بكلّ ما فيها من جهد وآراء الباحثين في مجال بناء المعاجم الإلكترونية للغة العربية.

- مدى أهمية المعالجة الآلية للغة العربية وما لذلك من انعكاسات إيجابية على البحوث في هذا المجال وعلى التطبيقات العامة في كل ما يتعلق باستخدام الحاسوب في التعامل مع اللّغة العربية.

- أهميّة استخدام اللّغة العربية في المعالجة الآلية وطرق معالجتها آليًا واستغلالها في بناء معجم إلكتروني للمعالجة الآلية للغة العربية.

- تمكين الحوار بين اللّساني والحاسوبي، وتضافر الجهود في مجال تطوير اللّسانيات الحاسوبية يفضي إلى تعاون ملح وضروري لمعالجة اللّغة العربية آليًا.

- تقديم صورة مطورة لدراسة بحثية في مجال اللّسانيات الحاسوبية يصف نظامًا حاسوبيًا لفهم اللغة العربية، يستخدم هذا النّظام مزيجًا من أصول علم اللّسانيات الضرورية للفهم وبعض المعارف البرمجية المرتبطة بمجال المعالجة الآلية لإنتاج برنامج حاسوبي يفهم اللّغة العربية.

وقد انتهت الدراسة بتقديم نموذج لساني حاسوبي في كيفية بناء معجم إلكتروني للمعالجة الآلية للغة العربية، يمكن استخدامه كأداة لتطوير برامج حاسوبية أخرى.

– لا يمكن الحديث عن معجم إلكتروني للمعالجة الآلية للغة العربية إلاّ بعد استكمال بناء قاعدة معطيات صرفية إذ بدونه يستحيل الدخول في أي مشروع علمي حاسوبي يعود بالنفع على اللّغة العربية.

– أن غياب الخبرة اللّسانية في هذا المجال هو سبب الفشل الّذي مُنيت به معظم البرامج الحاسوبية العربية، فقد اعتمدت التطبيقات السابقة على الخبرة اللّسانية مع إغفال يكاد يكون تامًا للخبرة اللّسانية الصورية، فقد اعتمد الحاسوبي على بعض اللّغويين ذوي التكوين التقليدي، إلاّ أن هؤلاء غالبًا ما يكونون غير قادرين على فهم متطلبات الحاسوب فيزودوا المبرمج بمعلومات أكاديمية صحيحة إلاّ أنّها غير قادرة على الاستجابة لمتطلبات الحاسوبي، ممّا ينعكس سلبًا على تطوير البرنامج الموضوع.

– أنّ اللّساني الّذي يحتاج إليه في وضع برامج الهندسة اللّسانية هو المتمكن من مختلف التطورات النّظرية التي تعرفها اللّسانيات الحاسوبية اليوم القادر على وضع الخوارزميات اللّسانية لمختلف مستويات نظام اللّغة العربية، وهذا النّوع من الباحثين هم الّذين يتمكنون من الدفع باللّغة العربية إلى مصاف اللّغات العالمية في الحوار مع الآلة.

– لقد اجتهد البحث في رصد كيفية بناء معجم إلكتروني للمعالجة الآلية للغة العربية وذلك بناء قاعدة معطيات صرفية فالغاية تكمن في الأهمية التي يمثلها المعجم الإلكتروني كمصدر بيانات أساسي بالنسبة لمعظم المحللات والمركبات اللّغوية. ومن أمثلة ذلك استخدام المحلل الصرفي الآلي للغة العربية لمعجم الجذور ومعجم الأوزان الصرفية؛ وهناك تكمن أهمية المعجم الإلكتروني في تطوير مثل هذه البرامج الحاسوبية.

– ظهر جليًا في ثنايا هذه الدراسة أن للمعالجة الآلية للصرف العربي أهمية كبيرة يمكن أن نجملها كما يلي:

1– يجب بذل كافة الجهود الممكنة لخدمة اللغة العربية.

2– أصبح من شبه المتفق عليه أن التعامل مع الحاسوب يجب أن يتم باستخدام لغات طبيعية لا اصطناعية.

3- تضخم حجم المعلومات وبالتالي ضرورة وجود طرق فعالة وسريعة في متابعة هذه المعلومات جمعًا ومعالجة.

4- تعتمد اللغة العربية على الصرف كثيرًا حيث يمثل أساسًا متينًا لها.

5- تمثل المعالجة الآلية للصرف العربي مطلبًا مهمًا في بناء مختلف التطبيقات اللغوية الحاسوبية.

- هناك عدد من الأفكار والاقتراحات المتعلقة بمعالجة الصرف العربي آليًا ومنها:

1- تحديد التطبيقات المختلفة التي يمكن ان تستفيد من علم الصرف وتحديد احتياجات هذه التطبيقات، وخصوصًا ما يتعلق بدرجة تعقد ودقة نظام حوسبة الصرف العربي.

2- دراسة نظم معالجة الصرف العربي آليًا في اللغات الأجنبية وبيان ما يمكن الاستفادة منها.

3- ضرورة حوسبة الصرف العربي على أسس لغوية سليمة وتبني آراء ما اتفق عليه اللغويون العرب.

-توصي هذه الدراسة بعدد من الأمور التي ينبغي مراعاتها في مثل هذه المشاريع، أهمها:

1- تكوين فرق عمل لدراسة مجال معالجة اللغة العربية الطبيعية آليًا.

2- القيام بمشروع لتوحيد المصطلحات اللغوية وكذلك مرادفاتها في اللغة الإنجليزية والفرنسية.

3- ضرورة تبني عمل إطار كامل لمنظومة معالجة اللغة العربية آليًا ومنظومة الصرف في اللغة العربية، ومن ثم بناء مختلف خوارزمياته بطريقة علمية.

- إنّ الصرف في اللّغة العربية يعدُّ مدخلاً أساسيًا في فهم الأداء الشامل للغة هذا من الجانب اللّغوي، أما على الصعيد الحاسوبي فتعدُ معالجة الصرف العربي آليًا أساسا لا غنى عنه للمعالجة الآلية للمعجم العربي واسترجاع المعلومات....

- إن أهمية تطبيقات الصرف العربي في حوسبة المعجم العربي تكمن في توسيع نطاق مصطلحات المعجم، وتعزيز معطياته اللّغوية: الفونولوجية والصرفية والنحوية والدلالية.

- تطوير آليات العمل اللّساني الحاسوبي في مجالاته المختلفة في مثل هذه المشاريع ممّا يستدعي تشكيل فرق عمل من اللّغويين والحاسوبين من أجل بناء معجم إلكتروني للمعالجة الآلية للغة العربية كنموذج من أجل بناء برامج حاسوبية أخرى.

- إنّ أهمية المعجم الإلكتروني للمعالجة الآلية للغة العربية تكمن في استعماله من خلال تقديم معلومات عن الكلمة المدخلة وذلك بهدف معالجتها آليًا، فبالإضافة إلى مهمة التعرف على أقسام الكلم فهو يقدم معلومات أساسية مرفقة بكل كلمة مدخلة آليًا.

- أنّ هذه الدراسة خضعت للبرمجة الحاسوبية وفي هذا دعم للمعالجة الآلية للغة العربية وبرهان على قابلية اللّغة العربية لأخذ مكانتها من التطور العلمي جنبًا إلى جنب مع لغات العالم الأخرى.

أعودُ- نهايةً- وأحمدُ الله ربّي على جُودِهِ وكرَمه، ذلك أنْ سخَّرَ لي يدًا ترْعاني في سَيْري خلال دراستي وفي بحثي المتواضع، وتُرشدُني إلى غايتي، وتتعهّدني بالنّظر، ممثلةً في شخص أستاذي ومعلِّمي الدكتور سيدي محمد غيثري حفظه الله ورعاهُ،

ومن نعمته علي كذلك- له الحمدُ والشكرُ وحدهُ- أنْ حباني بثلّةٍ ممنْ يُشْهَدُ لهم بالعلم في موضوع بحثي هذا؛ ليُنوِّروا- بتنقيحهم وتمحيصهم ونقدهم إلى نقضهم- طريق البحث، فيقوّموا اعوجاجه ويصوّبوا زلاّته، أساتذةً دكاترةً في اللّجنة العلمية الموقرة خاصة البحث المتواضع، باركهم الله سبحانه.

وفي الختام ، نزعم لأنفسنا أن هذا البحث إضافة جديدة فيما تناولته أقلام الباحثين في مجالات اللّسانيات الحديثة، و نسأل الله العلي القدير أن نكون قد وفقنا في إشباع هذا الموضوع بحثاً ودراسة، فإذا كان البحث موفقًّا في مجاله، فهذا فضلٌ من الله كبير وإن كان غير ذلك فلا يكلّف الله نفسًا إلا وسعها والله من وراء القصد وهو الهادي إلى سواء السبيل.

تمّت...

الطالب: أمين قدراوي

قائمة المصادر والمراجع

قائمة المصادر والمراجع

القرآن الكريم، برواية ورش عن نافع

أ- العربية:

1- إبراهيم أنيس:

- من أسرار اللغة، مكتبة الأنجلو المصرية، القاهرة، الطبعة الرابعة،1978م.

2- إبراهيم بن مراد:

- دراسات في المعجم العربي، دار الغرب الإسلامي، بيروت، الطبعة الأولى،1987م.
- مقدمة لنظرية المعجم، دار الغرب الإسلامي، بيروت، الطبعة الأولى،1997م.

3- إبراهيم محمد نجا:

- المعاجم اللغوية، دار الحديث، القاهرة، طبعة 1429هـ/2008م.

4- إبراهيم السامرائي:

- في الصناعة المعجمية، دار الفكر، عمان، ط1، 1419هـ/1992م.

5- أحمد أمين، أحمد الزين، إبراهيم الأبياري:

- ديوان حافظ إبراهيم ، الجزء الأول، دار الجيل، بيروت، دط، دت.

6- أحمد بريسول و كنزة بن عمر:

- المعجم العربي العصري وإشكالاته، منشورات معهد الدراسات والأبحاث للتعريب، جامعة محمد الخامس، السويسي- الرباط، المغرب، طبعة 2007م.

7- أحمد حسن الزيات وآخرون:

- المعجم الوسيط، المكتبة الإسلامية، استانبول/تركيا، دار المعارف، مصر، طبعة 1997م.

8- أحمد سليمان ياقوت:

- في علم اللغة التقابلي، دراسة تطبيقية، دار المعرفة الجامعية، القاهرة، د.ط، 1985م.

9- أحمد مختار عمر:

- صناعة المعجم الحديث، عالم الكتب، القاهرة، ط1418،1هـ/ 1998م.

10- أحمد مطلوب:

- بحوث لغوية، عمان، دار الفكر، دط،1987م.

11- أحمد عبد الغفور عطار:

- مقدمة الصحاح، دار الملايين، بيروت، ط3، 1404هـ/ 1984م.

12- أحمد شامية:

- في اللغة، دراسة تمهيدية منهجية متخصصة في مستويات البنية اللّغوية. دار البلاغ للنشر والتوزيع، الجزائر، الطبعة الأولى 1423هـ/ 2002م.

13- أحمد محمد قدور:

- مبادئ اللّسانيات، الطبعة الثانية، دمشق، سوريا، 1999م.

14- أحمد محمد عبد الدّايم:

- معجم الأبنية العربية، الأسماء والأفعال والمصادر، مكتبة لبنان ناشرون، بيروت، ط1، 2002 م.

15- أحمد عمايرة:

- المستشرقون ونظرياتهم في نشأة اللغة، وائل للنشر والتوزيع، عمان الأردن، ط3، 2002 م.

16- أحمد الميداني:

نزهة الطرف في علم الصرف، تحقيق، لجنة إحياء التراث، دار الآفاق الجديدة، بيروت، الطبعة الأولى،1981م.

17- إسماعيل العرفي:

اللغة العربية أم اللغات ولغة البشرية، دار الفكر، دمشق، الطبعة الأولى 1406 هـ - 1985م.

18- إميل بديع يعقوب:

- معجم الأوزان الصرفية، عالم الكتب، الطبعة الأولى، 1413هـ/1993م.

19- إدريس العلمي:

- في المصطلح، الدار البيضاء، المغرب، دط،2001 م.

20- الأمير مصطفى الشهابي:

- المصطلحات العلمية في اللغة العربية في القديم والحديث، دار صادر، بيروت، ط3.

21- أنيس فريحة:

- نظريات في اللغة، دار الكتاب اللبناني، المكتبة الجامعية، بيروت، الطبعة الأولى، 1983م.
- في اللغة العربية و بعض مشكلاتها، دار النهار للنشر، بيروت، الطبعة الأولى، 1980م.

22- أشواق محمد النجار:

- دلالة اللواصق التصريفية في اللغة العربية، دار دجلة، عمان، ط1، 2007م.

23- أنطوان الدحداح:

- معجم قواعد اللّغة العربية في جداول ولوحات زائد مسرد بالمصطلحات النحوية، عربي- انجليزي- فرنسي، مكتبة لبنان- ناشرون، الطبعة رقم 01- 1997م.

24- أنطوان عبدو:

- مصطلح المعجمية العربية، الشركة العالمية للكتاب، بيروت، الطبعة الأولى، 1991.

25- الأنباري أبو البركات:

- نزهة الألباء، تحق: محمد أبو الفضل إبراهيم، المكتبة العصرية، بيروت، ط1، 1424هــ/2003 م.

26- أسعد أحمد علي:

- تهذيب المقدمة اللّغوية للعلايلي، دار السؤال للطباعة والنشر. دمشق الطبعة الثانية، 1981م..

27- أبو بكر محمد بن سهل بن السرّاج:

- الأصول في النحو، تحقيق، عبد الحسين الفتلي، مؤسسة الرسالة، الجزء الثالث، بيروت، الطبعة الثالثة، 1988م.

28- أبو الحسن علي بن عيسى بن علي بن عبد الله الرماني:

- الحدود، تحقيق إبراهيم السامرائي، دار الفكر، عمان، دط، 1984م.

29- أبو العباس شهاب الدين أحمد بن علي بن أحمد القلقشندي:

- صبح الأعشى، الجزء الأول، المؤسسة المصرية العامة، القاهرة، دط، 1383 هــ/1963م،

30- أبو الطيب اللغوي:

- مراتب النحويّين، تحق: محمد أبو الفضل إبراهيم، المكتبة العصرية، صيدا بيروت، ط1، 1423هــ/ 2002 م.

31- ابن جني (أبو الفتح):

- سر صناعة الإعراب، تحق: محمد حسن محمد حسن إسماعيل وأحمد حسن رشدي شحاته عامر،دار الكتب العلمية، بيروت،الطبعة الثانية،2007 م.
- الخصائص، ت: محمد علي النّجار، عالم الكتب، بيروت، ط1، 1427هــ/2006 م.
- التصريف الملوكي، ت: محمد سعيد النعسان، مطبعة شركة التمدن الصناعية، مصر، ط1، د.ت.

32- ابن حزم:

- الإحكام في أصول الأحكام، تحقيق: إحسان عباس، الجزء الأول، دار الآفاق الجديدة، بيروت، 1983م.

33- ابن دريد(أبو بكر):

- جمهرة اللغة، ج1، طبعة حيدر أباد، 1344هــ.

34- ابن منظور(أبو الفضل جمال الدين محمد بن مكرم الإفريقي):

- لسان العرب، دار صادر، بيروت، ط6. 1417هــ /1997م.

35- ابن فارس (أبو الحسين أحمد بن زكريا):

- معجم مقاييس اللغة، تحق: عبد السلام هارون، دار الجيل بيروت، ط1، 1411هـ/1991م.

36- ابن سلام الجمحي:

- طبقات الشعراء، إعداد اللجنة العلمية لنشر التراث العربي، دار النهضة العربية، بيروت (د.ت).

37- ابن يعيش موفق الدين يعيش بن علي:

- شرح الملوكي في التصريف، تحقيق فخر الدين قباوة، المكتبة العربية، حلب، الطبعة الأولى، 1972م.

38- ابن القطاع الصقلي:

- أبنية الأسماء والأفعال والمصادر، ت: أحمد محمد عبد الدايم، دار الكتب والوثائق القومية، مركز تحقيق التراث، القاهرة، طبعة 1999م.

39- ابن خلدون (عبد الرحمن):

- المقدمة (محققة بلونين)، دار الفكر، طبعة 1424هـ / 2004م.

40- بلقاسم بلعرج:

- لغة القرآن الكريم، دراسة لسانية للمشتقات في الربع الاول، دار العلوم، الجزائر، دط، 2005م.

41- البخاري:

- صحيح البخاري، لأبي عبد الله البخاري، ضبطه ورقمه وذكر تكرار مواضعه وشرح ألفاظه وجمعه وخرّج أحاديثه في صحيح مسلم ووضع فهارسه: د مصطفى ديب البغا، مج4/ كتاب المغازي رقم 67، باب 10. 1477. موفم للنشر/الجزائر- دار الهدى للطباعة والنشر/ عين مليلة، 1992م،

42- جمال الدين عبد الله ابن أحمد ابن علي ابن محمد الفاكهي:

- شرح الحدود النّحوية، حققه وقدمه الدكتور محمد الطيب الإبراهيم، دار النفائس، بيروت، الطبعة الأولى، 1996م.

43- الجوهري (أبو نصر اسماعيل بن حماد):

- تاج اللغة وصحاح العربية، تحق: د إميل بديع يعقوب و د محمد نبيل طريفي، دار الكتب العلمية، بيروت، ط1، 1420هـ/1999م.

44- دزيرة سقال:

- نشأة المعاجم العربية وتطورها: معاجم المعاني-معاجم الألفاظ، دار الفكر العربي، بيروت، ط1، 1997م.

45- وليد العناتي وخالد الجبر:

- دليل الباحث إلى اللسانيات الحاسوبية، دار جرير، الأردن، ط1، 1428هـ/2007م.

46- الزبيدي (مرتض):

- تاج العروس من جواهر القاموس" تحقيق. عبد الستار أحمد فراج وغيره، الجزء23، مطبعة الكويت. الكويت، 1965م/ 1978م

47- زكي المبارك:

- النثر الفني في القرن الرابع الهجري، دار الجيل، دون تاريخ أو بلد.

48- حاتم صالح الضّامن:

- علم اللغة، بيت الحكمة، بغداد، دط، 1989م.

49- حازم علي كمال الدين:

- نظرية القوالب من نظريات علم اللغة الحديث، مكتبة الآداب، القاهرة، دط، دت.

50- حسين نصار:

- المعجم العربي، نشأته وتطوره، دار مصر للطباعة، القاهرة، ط2، 1988م.

51- حلمي خليل:

- مقدمة لدراسة التراث المعجمي العربي، دار النهضة العربية، بيروت، ط1، 1997م.
- دراسات في علم اللغة والمعاجم، دار النهضة العربية، ط1، 1998م.

52- حلام الجيلالي:

- تقنيات التعريف في المعاجم العربية المعاصرة، دراسة، منشورات اتحاد الكتاب العرب، دمشق، دط ، 1999م.

53- حسن خميس الملخ:

- نظرية الأصل والفرع في النحو العربي، دار الشروق، عمان، الأردن، الطبعة الأولى، 2001.

54- حُسام قدوري عبده:

- تأصيل الجذور السامية وأثره في بناء معجم عربي حديث. دار الكتب العلمية، بيروت، الطبعة الأولى 1428هـ/ 2007م.

55- خديجة بن عبد الرزاق الحديثي:

- أبنية الصرف في كتاب سبويه، منشورات مكتبة النهضة، بغداد، الطبعة الأولى، 1385هـ/1965م.

56- خليفة الميساوي:

- المصطلح اللساني وتأسيس المفهوم، دار الآمان، الرباط ، الطبعة الأولى، 1434 هـ ،2013 مـ .

57- يسرى عبد الغني عبد الله:

- معجم المعاجم العربية، دار الجيل بيروت، ط1، 1411هـ/1991م.

58- كمال محمد بشر:

- دراسات في علم اللّغة، دار المعارف، القاهرة- الطبعة التاسعة. ط9. 1986م.

59- كريم زكي حسام الدّين:

- أصول تراثية في علم اللّغة، مكتبة الأنجلو المصرية-القاهرة- الطبعة الثانية 1985م.

60- لطيفة إبراهيم النّجار:

- دور البنية الصرفية في وصف الظاهرة النحوية وتقعيدها" دار البشير، عمّان الأردن، الطبعة الأولى، 1414هـ/ 1994م.

61- مازن الوعر:

- قضايا أساسية في علم اللّسانيات الحديث، دار طلاس للدراسات والترجمة والنشر،ط1، 1988م.

62- مجدي إبراهيم محمد إبراهيم:

- بحوث ودراسات في علم اللّغة، "الصرف. المعاجم. الدلالة". مكتبة النهضة المصرية- القاهرة، د ط، 2004م.

63- ميشال زكريا:

- الألسنية(علم اللّغة الحديث) قراءة تمهيدية، المؤسسة الجامعية للدراسات، بيروت، 1985م.
- الألسنية التوليدية والتحويلية، المؤسسة الجامعية للدراسات والنشر والتوزيع، ط1، 1982م.

64- محمد بن مالك الطائي النحوي:

- إيجاز التعريف في علم التصريف، تحقيق محمد عثمان، مكتبة الثقافة الدينية، القاهرة، الطبعة الأولى، 1430هـ/2009م.

65- محمد البكاء:

- الحاسوب لطلبة العلوم الإنسانية: اللغة العربية، مكتبة الفلاح، الكويت، ط1، 2005م

66- محمد بن مالك الطائي النحوي:

- إيجاز التعريف في علم التصريف، تحقيق محمد عثمان، مكتبة الثقافة الدينية، القاهرة، الطبعة الأولى،1430،2009.

67- محمد رشاد الحمزاوي:

- من قضايا المعجم العربي، قديمًا وحديثًا دار الغرب الإسلامي، تونس، الطبعة الأولى، 1986م
- المصطلحات اللّغوية الحديثة في اللّغة العربية" معجم عربي أعجمي وأعجمي عربي، الدار التونسية للنشر، تونس/ المؤسسة الوطنية للكتاب- الجزائر طبعة 1987م.
- المعجمية، "مقاربة نظرية و مطبقة" "مصطلحاتها ومفاهيمها"، مركز النشر الجامعي ، تونس ، دط، 2004م.

68- محمد خير الحلواني:

- الواضح في النّحو والصرف، قسم الصرف، دار الحامون للتراث دمشق، سوريا، ط2- 1987م.

69- محمد خالد عيال سلمان:

- أثر المحتسب في الدراسات الصرفية، دار ومكتبة الحامد للنشر والتوزيع، عمان، الطبعة الأولى،2011 .

70- محمد خليفة الدناع:

- دور الصرف في منهجي النحو والمعجم، منشورات جامعة قار يونس، بنغازي، ليبيا، د.ط،1991م.

71- محمد علي عبد الكريم الرديني:

- المعجمات العربية (دراسة منهجية)، دار الهدى عين مليلة، الجزائر، ط2، 2006م.

72- محمد حسين آل ياسين:

- الدراسات اللغوية عند العرب، دار مكتبة الحياة، بيروت، 1980م.

73- محمد سويسي:

- لغة الرياضيات في العربية، دار التونسية للنشر والمؤسسة الوطنية للترجمة والتحقيق والدراسات، بيت الحكمة، د.ط، 1989م.

74- محمد سالم غنيم:

- نظم استرجاع المعلومات العربية، مظاهر الغموض وآفاق الحلول مكتبة الملك فهد الوطنية، ط1، الرياض، 2008م.

75- محمد عبد العزيز عبد الدّايم:

- النظرية اللّغوية في التراث العربي، دار السّلام، القاهرة، الطبعة الأولى 1427هـ/2006م.

76- محمد فؤاد عبد الباقي:

- المعجم المفهرس لألفاظ القرآن الكريم، مكتبة نويد إسلام، قم المقدسة، ط4، 1383هـ.

77- محمد الديداوي:

- الترجمة والتعريب بين اللغة البيانية والحاسوبية، المركز الثقافي العربي، الدار البيضاء، 2002م.

78- محمد محمد داود:

- العربية وعلم اللغة الحديث، دار غريب للطباعة والنشر والتوزيع، القاهرة، د ط، 2001م.

79- محمد الخضر حسين:

دراسات في العربية وتاريخها، المكتب الإسلامي، دمشق، الطبعة الثانية 1380 هـ-1960 م.

80- محمد الخولي:

- معجم علم اللغة النظري، مكتبة الحياة، لبنان، بيروت، الطبعة الأولى، 1983م.

81- محمّد المبارك:

- فقه اللّغة وخصائص العربية"، دراسة تحليلية مقارنة للكلمة العربية وعرض لمنهج العربية الأصيل في التجديد والتوليد، دار الفكر، بيروت، د. ط، 1425هـ/ 2005م.

82- محمد سالم محيسن:

- تصريف الأفعال والأسماء في ضوء أساليب القرآن، دار الكتاب العربي، بيروت، الطبعة الأولى، 1407هـ/1987م.

83- محمد سمير نجيب اللبدي:

- معجم المصطلحات النحوية والصرفية، مؤسسة الرسالة.بيروت/ قصر الكتاب.البليدة/ دار الثقافة - الجزائر، د.ط، د.ت.

84- محمود سليمان ياقوت:

- معاجم الموضوعات: في ضوء علم اللغة الحديث، دار المعرفة الجامعية، طبعة 2002م.

85- محمود فهمي حجازي:

- مدخل إلى اللّغة، دار قباء للطباعة والنشر والتوزيع، القاهرة طبعة جديدة مزيدة ومنقحة، 1997م.
- البحث اللغوي. دار غريب للطباعة والنشر. القاهرة، الطبعة الأولى،1994م.
- اللغة العربية في العصر الحديث (قضايا ومشكلات)، دار قباء للطباعة والنشر والتوزيع، القاهرة، طبعة1992م.

86- مصطفى بوعناني:

- الفونولوجيات الحاسوبية والمسارات المعرفية للإنجاز الكلامي، ملفات معرفية، الملف الأول مطبعة أبي، الطبعة الأولى، 2003م.

87- مصطفى النّحاس:

- من قضايا اللغة، مطبوعات جامعة الكويت، الكويت، الطبعة الأولى،1415هـ/1995 م

88- مهدي صالح سلطان الشمري:

- في المصطلح ولغة العلم، كلية الآداب، جامعة بغداد، دط،2012.

89- منيب قطيشات:

- قواعد البيانات، ، دار وائل للنشر والتوزيع، عَمَّان، الطبعة الثانية ،2005.

90- المنجد في اللغة العربية المعاصرة، دار المشرق، بيروت، ط2، 2001م.

91- نايف معروف:

- خصائص العربية وطرائق تدريسها،، دار النفائس، بيروت، الطبعة الثانية 1407 هـ-1987 م

92- نايف خرما:

- أضواء على الدراسات اللّغوية المعاصرة، سلسلة عالم المعرفة، العدد 9- سبتمبر 1978م.

93- نبيل علي:

- اللّغة العربية والحاسوب، دراسة بحثية، تقديم: أسامة الخولي، دار تعريب، الطبعة الأولى، الكويت 1988م.
- الثقافة العربية وعصر المعلومات، "رؤية لمستقبل الخطاب الثقافي العربي"،عالم المعرفة- العدد 265- يناير 2001م.
- الثقافة العربية وعصر المعلومات، عالم المعرفة، المجلس الوطني للثقافة والفنون والآداب، الكويت، العدد 265، يناير 2001م.
- العرب وعصر المعلومات، عالم المعرفة، الكويت، العدد 184- أبريل 1994م.
- البرمجيات ومهمة توظيف المعرفة، - العدد 541- ديسمبر 2003

94- نبيل علي/ نادية حجازي:

- الفجوة الرقمية، رؤية عربية لمجتمع المعرفة، سلسلة عالم المعرفة، العدد 318، أوت 2005م.

95- نجاة عبد العظيم الكوفي:

- أبنية الأفعال "دراسة لغوية قرآنية"، دار الثقافة للنشر والتوزيع القاهرة، دط، 1409هـ/ 1989م.

96- نور الهدى لوشن:

- مباحث في علم اللغة ومناهج البحث اللغوي، المكتبة الجامعية، الإسكندرية، طبعة 2000م.

97- نهاد الموسى :

- العربية نحو توصيف جديد في ضوء اللّسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، الأردن، ط1، 2000م.
- الأساليب مناهج ونماذج في تعليم اللغة العربية، دار الشروق، الأردن، ط1، 2003م.

98- سالم الغزالي:

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

99- سلوى حماده:

- المعالجة الآلية للغة العربية، "المشاكل والحلول"، دار غريب للطباعة والنشر والتوزيع، القاهرة، دط، 2009م.

100- سهاد مال الله كاظم، هالة بهجت عبد الوهاب، إخلاص خلف اكباش :

- تمثيل قواعد اللّغة العربية كحدود منطقية في أنظمة معالجة اللّغات الطّبيعية، مجلة الهندسة والتكنولوجيا، المجلد 26، العدد 11،العراق، 2008م.

101- السيوطي(جلال الدين):

- المزهر في علوم اللغة وأنواعها، تحق: د محمد أبو الفضل إبراهيم و محمد جاد المولى وعلي محمد البجاوي، المكتبة العصرية، بيروت، ط1، 1425هـ/ 2004م.
- الإقتراح في علم أصول النحو، تحق: أحمد محمد قاسم، مطبعة السعادة، القاهرة، ط1. 1396هـ/1976م.

102- سعيد الشرتوني:

- اقرب الموارد في فصح العربية والشوارد، مكتبة لبنان، لبنان، ط2، 1992م.

103- عادل عوض:

- تطبيقات النمذجة والمحاكاة الحاسوبية في الأنظمة البيئية- المائية، عالم الفكر العدد3،المجلد 32/2004م.

104- عادل فاخوري:

- اللّسانيات التوليدية والتحويلية، دار الطليعة- بيروت، ط2، 1988م.

105- عصام نور الدين:

- المصطلح الصرفي، مميزات التذكير والتأنيث، دار الكتاب العالمي، بيروت، الطبعة الأولى، 1409هـ ،1988 مـ ،

106- عمّار بوحوش، محمد محمود الذنيبات :

- مناهج البحث العلمي و طرق إعداد البحوث، ديوان المطبوعات الجامعية، الجزائر، 1999م .

107- عبد الحسن الحسيني :

- معجم مصطلحات علوم الكمبيوتر، عربي انجليزي- دار القلم- بيروت، ط1، 1407هـ/ 1987م.

108- عبد الصبور شاهين:

- في التطور اللّغوي"، مكتبة دار العلوم، القاهرة، الطبعة الأولى 1975 م.
- المنهج الصوتي للبنية العربية، مؤسسة الرسالة، بيروت، دط، 1980م.

109- عبد الكريم خليفة:

- اللّغة العربية على مدارج القرن الواحد والعشرين، دار الغرب الإسلامي، ط1، 2003م.

110- عبد الناصر جندلي:

- تقنيات ومناهج البحث في العلوم السياسية والاجتماعية، ديوان المطبوعات الجامعية، الجزائر، ط 2005م.

111- عبده الراجحي:

- التطبيق الصرفي،دار المعرفة الجامعية،اسكندرية، طبعة 1988م.

112- عبد الله العزازي:

- المعاجم العربية، القاهرة، ط3، 1403هـ/ 1981م.

113- عبد السميع محمد :

- المعاجم العربية (دراسة تحليلية)،دار الفكر العربي، طبعة 1393هـ/ 1974م.

114- عبد المقصود محمد عبد المقصود:

- دراسة البنية الصرفية في ضوء اللسانيات الوصفية، الدار العربية للموسوعات، بيروت، الطبعة الأولى،1427هـ، 2006 مـ ،

115- عبد المجيد الحر:

- المعجمات والمجامع العربية: نشأتها أنواعها نهجها تطورها، دار الفكر العربي، بيروت، ط1،1994م.

116- عبد القادر الفاسي الفهري:

- البناء المتوازي نظرية في بناء الكلمة وبناء الجملة، مكتبة بستان المعرفة للطباعة والنشر والتوزيع، 1990،.
- المعجم العربي، نماذج تحليلية جديدة، المعرفة اللسانية أبحاث ونماذج، دار توبقال للنشر، دار البيضاء، المغرب، الطبعة الثانية، 1999م.
- المقارنة والتخطيط في البحث اللساني العربي، دار توبقال للنشر، الدار البيضاء، المغرب، الطبعة الأولى، 1998م.
- المعجمة والتوسيط: نظرات جديدة في قضايا اللغة العربية، المركز الثقافي العربي، الدار البيضاء، المغرب، الطبعة الأولى، 1997م.
- المعجم العربي المولد، منشورات معهد الدراسات والأبحاث للتعريب، 2002
- التوليد والنسقية والترجم الآلية، ، المجلد 1، منشورات معهد الدراسات والأبحاث للتعريب، الرباط ، المغرب ، يناير 2001،
- تمكين اللغات، التكنولوجيا والإتصال، ، منشورات معهد الدراسات والأبحاث للتعريب، الرباط، المغرب، المجلد2، أكتوبر 1997.
- أزمة اللغة العربية في المغرب بين اختلالات التعددية وتعثرات "الترجمة". منشورات زاوية الطبعة الأولى 2005.

117- عدنان الخطيب:

- المعجم العربي بين الماضي والحاضر، منشورات معهد الدراسات العربية لجامعة الدول العربية، القاهرة، طبعة 1967م.

118- عبد الله بن حمد الحميدان :

- مقدمة في الترجمة الآلية، (دراسات لغوية) مكتبة العبيكان، السعودية، الطبعة الأولى، 1421هـ/ 2001م.

119- عبد العزيز قلقيلة:

- لغويات، مكتبة الأنجلو المصرية، المطبعة الفنية الحديثة، د.ط، القاهرة،1977م.

120- عبد العلي الودغيري:

- دراسات معجمية، تحت قاموس عربي تاريخي وقضايا أخرى، مطبعة النجاح الجديدة الطبعة الأولى 2001م.

121- عبد القادر عبد الجليل:

- المدارس المعجمية دراسة في البنية التركيبية، دار الصفاء، عمان، الأردن، ط1، 1420ه-1999/م.
- علم اللسانيات الحديثة، نظم التحكم وقواعد البيانات، دار صفاء، عمان، ط1، 1422هـ/2002م .

122- عبد القاهر الجرجاني:

- المفتاح في الصرف، تحقيق، علي الحمد، مؤسسة الرسالة، بيروت، دط،1987م.

123- علي عبد الواحد وافي:

- علم اللغة، دار النهضة، مصر، ط6، 1387هـ/1967م.
- فقه اللغة، نهضة مصر، الطبعة الثالثة، 2004م.

124- عبد الرحمن الحاج صالح:

- بحوث ودراسات في اللسانيات العربية، المؤسسة الوطنية للفنون المطبعية موفم للنشر،ج2/1، 2007م.

125- عبد السلام المسدي:

- قاموس اللّسانيات، الدار العربية للكتاب، تونس 1994م.
- اللّسانيات من خلال النصوص- النشرة الأولى- الدار التونسية للنشر، تونس، د ط، 1984.
- اللسانيات وأسسها المعرفية، الدار التونسية للنشر، تونس ، د ط ، د ت.
- مباحث تأسيسية في اللسانيات، دار الكتاب الجديدة المتحدة، ليبيا،الطبعة الأولى، 2010م

126- عبد العزيز عتيق:

- مدخل إلى دراسة الصرف العربي"، دار النهضة، مصر، الطبعة الثانية 1974م.

127- عبد الرحمن شاهين:

- في تصريف الأسماء، مكتبة الشباب، القاهرة، د. ط، 1977م.

128- عبد الحليم عبد الباسط محمد المرصفي:

- من صيغ وأوزان العربية "أفعل" دار مرجان للطباعة- القاهرة، ط1، 1978م.

129- علي حلمي موسى:

- دراسة إحصائية لجذور معجم الصحاح باستخدام الكومبيوتر، الهيئة المصرية العامة للكتاب، (د.ب)، طبعة 1978هـ/1979م.

130- علي القاسمي:

- المعجمية العربية بين النظرية والتطبيق، مكتبة لبنان ناشرون، بيروت، الطبعة الأولى، 2003م.
- علم اللغة وصناعة المعجم، مكتبة لبنان ناشرون، بيروت، الطبعة الأولى، 2004م.

131- علي السليمان الصوينع:

- كشافات التباديل واسترجاع المعلومات في اللّغة العربية، مكتبة الملك فهد الوطنية، الرياض،س1/1988م.

132- غنيم الينبعاوي:

- المصباح جهود ابن جني في الصرف وتقويمها في ضوء علم اللغة الحديث، المكتبة التجارية، مكة المكرمة، الطبعة الأولى،1995م.

133- الفيومي:

- المصباح المنير، في غريب الشرح الكبير،المكتبة العصرية، بيروت- ط 1، ج1، 1996م.

134- الفيروز ابادي:

- القاموس المحيط، دار الكتب العلمية، بيروت، ط1، 1415هـ/1995م.

135- فتح الله سليمان:

- دراسات في علم اللغة، دار الآفاق العربية، القاهرة، ط1، 1429هـ/ 2008م.

136- فخر الدين قباوة:

- تصريف الأسماء والأفعال"، جامعة حلب، كلية الآداب سوريا، الطبعة الثانية، 1981/1401.

137- صافية زفنكي:

- التطورات المعجمية والمعجمات اللغوية العامة والعربية الحديثة، منشورات وزارة الثقافة، دمشق، د.ط، 2007م.

138- صالح بلعيد:

- في إعداد المناهج اللّغوية وإعداد الأبحاث، دار هومه، الجزائر، د. ط، 2005م.

139- صالح الفاخري:

- علم التصريف العربي،ELGA، مالطا دط، 1999م.

140- رياض زكي قاسم:

- معجم علم اللغة النظري، دار المعرفة، بيروت، ط1، (د.ت).
- المعجم العربي: بحوث في المادة والمنهج والتطبيق، المعرفة، بيروت، ط1، 1407هـم/ 1987م.

141- رأفت الكمار:

- الحاسوب وميكنة اللغة العربية، دار الكتب العلمية، القاهرة، ط1، 2006م.

142- رمضان عبد الله رمضان:

- الصيغ الصرفية في العربية في ضوء علم اللغة المعاصر، مكتبة بستان المعرفة، الإسكندرية ،الطبعة الأولى، ، 2005م .

143- رمضان عبد التوّاب:

- المدخل إلى علم اللغة ومناهج البحث اللّغوي، مكتبة الخانجي، القاهرة، دار الرفاعي، الرياض، الطبعة الأولى، 1403هـ/ 1983م .
- التطور النّحوي للغة العربية، محاضرات ألقاها في الجامعة المصرية عام 1929 م، المشترق الألماني برجشتراسر G.Bergsträsser، مكتبة الخانجي، القاهرة، الطبعة الثانية، 1417هـ- 1997 م.
- الفصحى وتحديات العصر، بحث ضمن كتاب بحوث ومقالات في اللغة، مكتبة الخانجي، القاهرة، دط، دت.

144- رمزي البعلبكي:

- معجم المصطلحات اللغوية، دار العلم للملايين، بيروت، 1990م.

145- كمال بشر:

- دراسات في علم اللغة ، دار المعارف، مص، د.ط، 1973م .

146- شحادة فارع (وآخرون):

- مقدمة في اللغويات المعاصرة، نشر وائل، الجامعة الأردنية، د.ط، د.ت.

147- الشيخ رضا الدّين محمد بن الحسن الاستراباذي النّحوي:

- شرح شافية ابن الحاجب مع شرح شواهده، تحقيق الأساتذة: محمد محي الدين عبد الحميد، محمد نور الحسن، محمد الزفزاف، دار الكتب العلمية، بيروت، طبعة 1402هـ/1982م.،ق1/ج1.

148- شرف الدين الراجحي وسامي عياد حنا:

- مبادئ علم اللسانيات الحديث، تقديم: عبده الراجحي،دار المعرفة الجامعية، مصر، دط، 2003م.

149- توفيق محمد شاهين:

- أصول اللّغة العربية بين الثنائية والثلاثي، مكتبة وهبة، الطبعة الأولى،1400هـ،1980م.
- علم اللغة العام (دراسات لغوية)، مكتبة وهبة، د.ب، ط1، 1400هـ/1980م.

150- تمام حسان:

- مناهج البحث في اللّغة، دار الثقافة، الدار البيضاء، المغرب، د. ط، 1979م.
- اللغة بين المعيارية والوصفية، عالم الكتب، القاهرة، الطبعة الرابعة، 2000م.
- اللّغة العربية معناها ومبناها. دار الثقافة. الدار البيضاء، د ط. 1421هـ، 2001م.
- اجتهادات لغوية، عالم الكتب،القاهرة، الطبعة الأولى، 2007.

151- التفتازاني:

- شرح مختصر التصريف العزي في فن الصرف، تحقيق، عبد العال سالم مكرم، ذات السلاسل، الطبعة الأولى، الكويت، 1983م

152- الخليل بن أحمد الفراهيدي:

- كتاب العين، تحق: د عبد الحميد هنداوي، منشورات محمد علي بيضون، دار الكتب العلمية، بيروت، ط1،1424هـ/ 2003م.

153- العُكبري عبد الله بن الحسين بن عبد الله أبو البقاء:

- مسائل خلافية في النحو، تحقيق الدكتور عبد الفتاح سليم،مكتبة الآداب، القاهرة، الطبعة الثالثة،2007.

154- هادي نهر:

- الصّرف الوافي، عالم الكتب الحديث، الأردن، دط، 2010م.

ب- المترجمة:

1- ادوارد فايجنبوم وباميلاماك كوردك:

- الجيل الخامس: الذكاء الاصطناعي والتحدي الياباني للعالم، د ط،1987م.

2- آلان بونيه:

- الذكاء الاصطناعي واقعه ومستقبله، ترجمة علي صبري فرغلي، عالم المعرفة، رقم 172،1993م.

3- جيرهارد هلبش:

- تاريخ علم اللغة الحديث، ترجمه و علق عليه د/ سعيد حسن بحيري، مكتبة زهراء الشرق، القاهرة، الطبعة الأولى، 2003 مـ .

4- جون ليونز:

- مدخل إلى اللغة واللسانيات، جامعة كامبريدج، لندن، نيويورك، د ط، 1981م.

5- جورج ماطوري:

- منهج المعجمية، ترجمة عبد العلي الودغيري، مطبعة المعرف الجديدة،الرباط، 1993

6- شارل بوتون:

- اللسانيات التطبيقية، ترجمة د/ قاسم المقداد، محمد رياض المصري، دمشق، سوريا، د ط، د ت.

7- ماريوباي:

- أسس علم اللغة، ترجمة أحمد مختارعمر، طرابلس، دط، 1973م.

8- ماري نوال غاري بريور:

- المصطلحات المفاتيح في اللسانيات، ترجمة عبد القادر فهيم الشيباني، الجزائر، ط1، 2007م.

9- ف. و. لانكستر و أ. ج. وورنر:

- أساسيات استرجاع المعلومات (نظم استرجاع المعلومات)، ترجمة حشمت قاسم، مطبوعات مكتبة الملك فهد الوطنية، السلسلة الثانية- 29، الرياض، ط3 مزيدة ومنقحة، 1418هـ/ 1997م.

10- كارول شابل:

- تطبيقات الحاسب الآلي في اكتساب اللغة الثانية، أسس للتعليم والقياس والبحث العلمي، ترجمة: سعد القحطاني، جامعة الملك سعود، الرياض، 2007م.

11- نعوم تشو مسكي:

- البنى النحوية، ترجمة يؤيل يوسف عزيز ومراجعة مجيد الشماطة، منشورات عيون، الدار البيضاء، 1987م.

12- يوهان فك:

- العربية، دراسات في اللغة اللهجات والأساليب، ترجمة عبد الحليم النجار، المركز القومي للترجمة، الطبعة الأولى، 2014م.

ج- المراجع الأجنبية :

1- اللغة الفرنسية:

1. **Abbes, Ramzi**, la conception et la Réalisation d'un concordancie électronique pour l'arabe. Institut national des sciences appliques, Science de l'informatique thèse doctorat 2004.
2. **Aida Khemakhem** , (ARABIC LDB: Une base lexicale normalisée pour la langue arabe ; Mémoire master en(système d'information et nouvelles Technologies) université de Sfax. Faculté des sciences Economiques et de gestion. 2005- 2006.
3. **Alain Couillault** , « traitement automatique des langues dans les industries de L'information » Livre blanc. Janvier 2005. Edition. APIL- GFIL..
4. **Alain Polguère** (Structuration et Mise en jeu Procédurale d'un modèle linguistique) déclaratif dans un cadre de génération de texte- thèse doctorat (PH. D) en linguistique, 1990.
5. **André Martinet** (Eléments de linguistique Générale) Edition, Paris. 1976
6. **Augustin Lux** (sur l'évolution de l'intelligence Artificielle) la machine perceptive 11/12 janvier 1990.
7. **Bernard Quemada**, Bases de données informatisées et dictionnaires, Lexique 2, Presses Universitaires de LILLe.

8. **Blandine courtois et Max Silberztein**, les dictionnaires electroniques DELAS et DELAC in languistica comunicatio,vol1.1N°2.2.1989.
9. **Bessou sadik**, Analyse de données Textuelles pour la classification Automatique par les techniques de Text Mining à la Langue Arabe, Mémoire de Magister, option: Sciences et technologies de l'information et de la Communication, Université FERHAT ABBAS-SETIF (UFAS) 20.
10. **C. Brockelmann,** « Précis de linguistique sémitique, traduit de l'allemand par w. Marçais et M. cohen, Librairie Paul Geuthner, Paris 1910.
11. **Carré, René Etal**. « Langage Humain et Machine », Presses du CNRS, Paris, , 1991.
12. **Christian Jacquemin**; Traitement automatique de langues pour la recherche d'information volume 49- n°=2/2000. Hermes science Publication, Paris.
13. **Christine Jaquet-pfau**, pour un nouveau dictionnaire informatisé, in Ela, revue dedidactologie des langues-cultures et lexicultrurologie, numéro spécial, sur Dictionnaires et Innovation, n 137, janvier – Mars 2005
14. **Claude de Loupy**, Evaluation de l' apport de connaissances Linguistique en Désambuisation sémantique et rechercher Documentaire, thèse de Doctorat présentée à l Université d Avignon et des Pays de Vaucluse. 2000
15. **Daniel Kayser, Bernard Levrat**,Traitement automatique du langage naturel, Ed, HERMES science Publication, Paris Volume 20- n°=03 ;2001-
16. **Delphine Bernard**, « apprentissage de connaissances morphologiques pour l'acquisition automatique de ressources lexicales », thèse doctorat, sciences cognitives, université Grenoble I, Faculté de science de la langue 2006.
17. **Dominique ARCHAMBAULT, Bassano, J-C** (Directeur de thèse).« Proposition de réseau neuromimétique par le traitement du Langage naturel » thèse de doctorat 3ème cycle 1995. Université d'Orléans, Orléans, France.
18. **Eric Laporte**, mots et niveau lexicale ,in TAL ,vol n° 2 « Etat de l'art 1997.
19. **FUCHS Catherine**, Les ambiguïtés du Français. Paris Ophrys. 1996.
20. **Henri Fleich,** traité de philologie Arabe ; Vol. 1, Imprimerie catholique, Beyrouth, 1961.
21. **Hassoun M**, Conception d'un dictionnaire pour le traitement automatique de l'arabe dans différents contextes d'application, Thèse de doctorat d'état, Univ. Lyon-1,1987.
22. **J. Roy. Debore**, in language. Edition CEPL Paris. 1973..
23. **Jaques Pitrat**, « textes, ordinateurs et compréhension », Eyrolles Paris, 1985.
24. **Jean Dubois**, La morphologie- In: Dictionnaire de linguistique et des Sciences de langages ; maison d'édition Larousse, 2001.
25. **Jean Senellart**, Outils de reconnaissance d`expressions linguistiques complexes dans des grands corpus. Thèse de doctorat, , Janvier 1999, LADL, Paris.
26. **Jean Pruvost**, Dictionnaire et nouvelles technologies, coll Ecritures électronique, PUF. Paris. 2000.
27. **John Lyons**, (Introduction to théorial linguistics) .Combridge 1960
28. **Karima Meftouh**, « Un réseau simplement récurrent pour la génération d'une représentation interne du sens en langue arabe basée sur les cas Sémantique » Mémoire de magister Université de Annaba .2000.

29. **Laurence, Teirlinckx**, la lemmatisation de l'arabe non classique- centre d'étude sur Grégoire. France. 2003.
30. **Le petit Larousse** - France: maison d'édition, 2006.
31. **Lerat.. Coll** ; Les langues Spécialisées, linguistique nouvelle presses universitaires de France, Paris.
32. **Les bases de données du LADL**: Analyse automatique des langues naturelles, Aspects technologiques, Paris, 1989.
33. **M.Elhannach,** Lexique-Grammaire de l'arabe :classe des verbes qualitatifs,(suite)in linguistica communicatio,V.1,N.2.1989
34. **M.Jardino et M.EL-Béze**, Modélisation probabiliste du language naturel,(in T A L ,volume 44-n°1 /2003..
35. **Mel'učk (I), Clas (A) et Polguère (A)** , :Introduction à la l'lexicologie explicative et combinatoire, Louvin-la-Neuve (Belgique), Ducolot, 1995
36. **Max Silberztein,** Dictionnaires électroniques et analyse automatique de textes, , masson, Paris, 1993.
37. **Maurice Gross**, Méthodes en syntaxe, , Hermann, Paris 1975
38. **Maurice Gross, André Lentin,** Notions sur les grammaires formelles, collection « programmation » . Gauthier Villers, Editeurs, Paris 6/1970.
39. **Marie Nöelle, Gary Prieur**, les termes clés de la linguistique,Paris,édition, Seuil,1996.
40. **Margarita Alonso Ramos K**, Construction d'une base de données des collocations bilingues français espagnol,in langages N°134.
41. **Mohamed El Amine Abderrahim**, Reconnaissance des unités linguistiques signifiantes, Thèse pour obtenir le titre de doctorat, spécialité informatique encadré par Mr Fethi BEREKSI REGUIG et Mme Laurence BALICCO, Université de Tlemcen, 2007-2008,
42. **Michael Zock – John Carroll**, les dictionnaires électroniques revue TAL n 44, 2003.
43. **Nacera TAIBI Née Mellite**, « Contribution à l'étude du traitement automatique des erreurs dans un texte écrit en arabe » Mémoire de Magister English 1997.
44. **N.Chomesky , G.A.Miller**, l' Analyse Fromelle des naturelles,traduit par Ph. Richard et N.Ruwet Paris1971.
45. **Philip Miller et Thèse torris (**Formalismes Syntaxiques pour traitement automatique de la langue naturelle) Hermes. Paris. 1990.
46. **Pierrette Bouillon,** avec la collaboration de: Françoise Vandooren- lyne Dasylva- lanrence jacqmin- Sabine lehmann- Graham- Russell et Evelyne Viegas (Traitement automatique des langues naturelles) Edition Duculot. AUPELEF- Uref. 1998. Paris..
47. **Pierre-Andre Buvet**. Représentation métalinguistiques de phrases a partir de traducteurs, Centre Lucien Teniere, Universite de Franche- Comte. A paraitre c (2002)
48. **Jacqueline LEON,** Traitement automatique des langues, histoire Epistémologie langage, Paris 2001
49. **Eric Laporte,** Traitement automatique du mot: Etat de l'art, IGM, 2001
50. **Unitex**, Programme de recherche Unicode, IGM, Université Marne la Vallée, Paris 2002
51. **Zaafrani Riadh**, "un dictionnaire électronique pour apprenant de l'arabe (langue second) basé sur corpus, JEP-Taln 2004, Traitement automatique de l'arabe, fès, 20 Avril 2004.

2- اللغة الإنجليزية :

1. **Abdehadi** Soudi, A computational Lexem – based Treatment of arabic Etat thesis in Linguistics, faculty of letters and human ‘morphology, Doctora d sciences, med v yniversity 2002
2. **AL Tayyar, Musaid Salah.** (Arabic Information Retrieval System Based on Morphological Analysis (AIRSMA): A comparative study of word, Stem, root and morpho- semantic methods- PH. D. thesis- De MONT FORT University, July. 2000.
3. **ALI, Nabil** (the Second wave of Arabic Natural Language processing (2003)
4. **Anderson, PL,** « OCR Enters the Practical Stage »Dataation,Vol 17, 1971,
5. **Ann Moseley,** Interaction, Fourth edition and American Heritage Dictionary College, edition Houghton Mifflin company, 4ème édition, Dictionary of the English; 2003.
6. **Moscati,** An introduction to the comparative grammar of the Semitic languages,.,2nd edition.1969.wisbaden.
7. **ATTIA Mohamed A,** (Alarge – Scale Computational processor of the Arabic Morphology : A Master’s thesis, Cairo University, (Egypt) 2000.
8. **ATTIA Mohamed A,** « Devloping Robust Arabic Morphological Tranducer Using Finite State Technology In 8 th annual CLUK research colloquium (2005).
9. **ATTIA Mohamed A,** « Report on the Intreduction of arabic to pargram, the pargram Fall Metting, National Centre for language technology ; Ireland (2004).
10. **Attia Mohammed** « Developing Robust Arabic Morphological transducer using Finite state technology in 8 th annual cluk Research colloquim.
11. **Britannica**, Deluxe edition ,CD-ROM(Dictionary),2001
12. **Cacaly Serg- Model**. In: Dictionnaire encyclopédique de l 'information et de la documentation. 1998-
13. **Carr Micheal**, "Internet dictionaries and lexicography", international journal of lexicography, 1-8 October 1996.
14. **Colin Beardon and other** (Natural Language and Computional Linguistics An Introduction, Ellis Horwood, 1991,
15. **Chomsky, Noam.** Aspect of the theory of Syntax. Longman Group. First Publisher. London, 1965.
16. **DARWISH, K. And D.OARD.** CLIR Experiments at Maryland For TERC 2002 : Evidence Combination for arabic English Retrievnl In TREC, Gaithersburg.
17. **DARWISH, K.** probabilisitic Methods For Scarching OCR- Degraded Arabic Text, Doctoral disseration, university of maryland (2003).
18. **David Joff,** De Schryver Gilles-Maurices, "On how electonic dictionarie are Really used" Euralex 2004, proceeding.
19. **De Schryver Gilles-Maurices,** online dictionaries on the internet : an overview for the Africaine language, lexikos 13.1-20,2003.
20. **Farid Meziane,** "Development and Implementation of a dervational for the arbic language", A software engineering perspective.
21. **Geisler C,** "Computer-dased dictionary use : creation of a new Anglo-Swedish dictionary for the internet", ASLA information, 25.(2) – 1999.

22. **Gérard SABAH,** Intelligence Artificielle et langage, Processus de compréhension T 2 hermes. 1990.
23. **Ghayda AL- Talib** « An Arabic Natural Language Interface To Data Base System Using Prolog » Master thesis, College of Science, Department of computer Science, University of Mosul, Iraq, 1991,
24. **Homiedan A,** « Basis of the decision making process exercised by the translator, college of language, Ain shams university, philology, 1994.
25. **Kasem Hishmat M. A.**; Arabic In Specialist Information Systems" a study In Linguistic Aspects of information Transfer. Ph- D- thesis- Univ. Of London 1978.
26. **Kenneth Slonneger,** and others « Formal syntax and semantics of programming language, addison- wesely publishing company ; 1995.
27. **Khoja S, Orad,** « Stemming Arabic text, technical report, coputing de partement, lancaster University, (1999
28. **Nesi Hilary (ND),** Dictionaries on computer, How different markets have created different product, center for english language teacher education, University of warwick, CV4 7AL, United Kingdom.1987.
29. **James Martin,** Speech and language processing , An introduction to naturel language processing, compititional languistics and Speech recognition ,second edition ,Spring.
30. **Nesi Hilary,** "Do dictionaries help students write? ERIC document reproductionservice-NO.ED, 1987.
31. **Optical Character Recognition**, Reconnaissance Optique De Caractères
32. **Oxford dictionary,** Oxford ; university press ; 2006.
33. **Select paper on Lexicon** - Grammar, Vol. : 1-3, LADL, Paris (1973-1999)
34. **Sapir, Edward**. Language, An Introduction to the study of speech, New York: Harcourt, Bace & world, 1921.
35. **Sohnlyons** « Introduction », In New Horizons in Linguistics, Penguin Books, UK .
36. **Storre A/ Freesek** "Wörterbücher in internet", in deutsche sprache N° 24/2-1996. -
37. The New Encyclopedia Britinnica, Vol 23, , encyclopedia Britinica Inc Translated, USA 1996.
38. **Z.S. Harris,** Dunod, Structure mathématique du langage, Paris, 1971

د- المخطوطات (الرسائل الجامعية):

1- بربارة سهيلة :

- الترجمة بمساعدة الحاسوب، من الانجليزية إلى العربية، مذكرة ماجستير من إشراف: أ. مختار محمد صباحي، جامعة الجزائر، قسم الترجمة، 2005- 2006.

2- الزبيدي فالح حسن عويد:

- تطبيق النظرية التوليدية التحويلية على اللّغة العربية حاسوبيًا"، رسالة ماجستير، جامعة بغداد، بغداد، العراق، 1992.

3- سيدي محمد غيثري:

- التركيب الفعلي العربي، دراسة لسانية- حاسوبية، أطروحة جامعية للحصول على درجة دكتوراه الدولة في اللّسانيات الحاسوبية، معهد اللّغة العربية وآدابها، جامعة تلمسان، السنة الجامعية- 1419- 1998.

4- فارس شاشة:

- المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي إعرابي للفعل العربي، مذكرة ماجستير في علم المكتبات والتوثيق، كلية العلوم الاجتماعية والإنسانية، قسم علم المكتبات والتوثيق جامعة الجزائر، 2008.

5- ثريا عبد الله عثمان إدريس:

- الصّيغ الفعلية في القرآن الكريم، أصواتًا وأبنية ودلالةً. رسالة مقدمة لنيل درجة الدكتوراه في اللّغة.المجلّد 1، كلية اللّغة العربية، جامعة أم القرى، المملكة العربية السعودية، 1410هـ/1989م.

6- عمر مهديوي:

- توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية، مقاربة لسانية حاسوبية، أطروحة جامعية لنيل الدكتوراه، الجزء الأول، إشراف الدكتور عبد الغني أبو العزم، كلية الآداب والعلوم الإنسانية، عين الشق، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجمات، جامعة الحسن الثاني، الدار البيضاء،2008م/2009م.

7- عز الدين غازي:

- نظام قواعد صرفي – صواتي للغة العربية، مقاربة الفعل، رسالة لنيل دبلوم الدراسات العليا في الآداب، تحت إشراف الدكتور.محمد الحناش، كلية الآداب والعلوم الإنسانية ظهر المهراز، فاس 1999م.
- بناء المعجم الآلي للغة العربية: معالجة التعبيرات المركبة والمتلازمة، أطروحة جامعية لنيل الدكتوراه، إشراف الأستاذ الدكتور: عبد الغني أبو العزم، والأستاذ الدكتور: محمد لهلال، كلية الآداب والعلوم الإنسانية، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجميات، جامعة الحسن الثاني، عين الشق- الدار البيضاء، السنة الجامعية:2007/2008م.

8- محند الركيك:

- قراءة جديدة في المعجمية العربية، نحو مقاربة معجمية تفسيرية تأليفية ، أطروحة لنيل الدكتوراه في اللسانيات العربية ، فاس، المغرب، 2000م.

هـــ- الدوريات والمجلات والحوليات:

- **شكري فيصل**، قضايا اللغة العربيَّة المعاصرة، المجلة العربية للدراسات اللغوية، الجزء 2، عدد 1، 1983م.
- **محمد رشاد الحمزاوي**، النص المعجمي في المولدات والأعجميات، مجلة المعجمية، عدد11، 1995م.
- **عبد الإله ديوه جي**، مفاهيم أساسية حول تقنية المعلومات، مجلة عالم الفكر، م18. ع.3، 1987م.
- **أبحاث لسانية**، منشورات معهد الدراسات والأبحاث للتعريب:
  - المجلد 12، العدد 2/1، 2007م .
  - المجلد 13، العدد 2/1، 2008م.
- **مجلة الدراسات المعجمية**، منشورات الجمعية المغربية للدراسات المعجمية، الرباط، المغرب:
  - العدد الثالث والرابع، 1425هـــ ،2005م
  - العدد السادس، 1428هـــ/2007م
  - العدد السابع و الثامن، 1430هـــ 2009م
- **مجلة اللسانيات:** مجلة في علوم اللسان وتكننولوجياته ، مركز البحث العلمي والتقني لتطوير اللغة العربية، العددان 12،13،الجزائر،2007،
- **محمد عبد العزيز عبد الدّايم**، نظرية الصرف العربي، دراسة في المفهوم والمنهج، حوليات الآداب والعلوم الاجتماعية، دورية علمية محكمة، الرسالة 185، الحولية الحادي والعشرون، كلية دار العلوم،2001 مـــ
- **مجلة الهندسة والتكنولوجيا:** تمثيل قواعد اللّغة العربية كحدود منطقية في أنظمة معالجة اللّغات الطّبيعية، المجلد 26، العدد 11،العراق، 2008م.
- **مجلة تعريب:** المركز العربي للتعريب والترجمة والتأليف والنشر، دمشق، العدد 36، جوان 2009.
- **مجلة العربي:**
  - العدد 541- ديسمبر 2003
  - العدد 543، فبراير 2004.
- **مجلة التراث العربي:** إتحاد الكتاب العرب، دمشق، عدد77، السنة التاسعة عشر، 1420هـــ/1999م.

❖ **مجلة المعجمية :** العدد الخامس والسادس، تونس، 1989 .1990 .

❖ **مجلة اللغة العربية:** المجلس الأعلى للغة العربية، الجزائر، العدد العاشر، 2004.

❖ **مجلة اللسان العربي:** المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق تعريب الرباط.

- العدد 22، 1982م/1983م.
- العدد 24، 1984م/1985م.
- العدد 27، 1986م/1987م
- العدد 28، 1987م.
- العدد 35، 1991م.
- العدد 42، 1998م.
- العدد 48، 1999م.
- العدد 49، 2000م.
- العدد 50، 2000م.

❖ **مجلة المشعل:** مخبر المعالجة الآلية للغة العربية، جامعة أبي بكر بلقايد، تلمسان، الجزائر، عدد3 يناير 2008.

❖ **مجلة المجلس الأعلى للغة العربية:** المجلس الأعلى للغة العربية، الأبيار، الجزائر، عدد2، 1999م.

❖ **مجلة المجمع الجزائري للغة العربية:** المؤسسة الوطنية للفنون المطبعية

- العدد6، السنة الثالثة، ذوالحجة 1428هـ/ ديسمبر 2007م.
- العدد 7، السنة الثالثة، جمادى الثاني 1429هـ/ جوان 2008م.
- العدد1، السنة الأولى، ربيع1/1426هـ، ماي 2005م.
- العدد 2، السنة الأولى، ذوالقعدة 1426هـ/ديسمبر2005م.

❖ **مجلة كلية الآداب والعلوم الإنسانية:** ظهر المهراز (المهراز)، فاس- العدد- 14- 2006.

❖ **مجلة التواصل اللّساني:** المجلّد الثاني، العدد الثاني، 1990، مطبعة النجاح، الدار البيضاء المغرب.

❖ **منشورات كلية الآداب والعلوم الإنسانية:** (ابن مسيك)، ط1، الدار البيضاء، 1992م.

❖ **المجلة العربية للعلوم الإنسانية**، (جامعة الكويت)، العدد 38، المجلد 10، 1990م.

❖ **مجلة عالم الفكر** (الكويت)، المجلد 20، العدد 3، 1989م

و- المؤتمرات والندوات:

1- العربية:

- **إدريس القاسمي:** المعجم الآلي الشامل، الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية، المنظمة العربية للتربية والثقافة والعلوم، مدينة الملك عبد العزيز للعلوم والتقنية، (-1429هـ 2008/5/7-5/5 م)
- **حسن السيد وزكريا الكردي:** المعجم الحاسوبي للغة العربية، الجانب الحاسوبي الجوانب الحاسوبية للمعجم الإلكتروني للغة العربية، الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية، المنظمة العربية للتربية والثقافة والعلوم، إدارة العلوم والبحث العلمي، مدينة الملك عبد العزيز للعلوم والتقنية، دمشق 15 أبريل2008م.
- **مأمون حطاب:** التحليل الصرفي للغة العربية باستخدام الحاسوب" الموسم الثقافي الرابع عشر لمجمع اللّغة العربية الأردني، من منشورات المجمع، مجمع اللّغة العربية، الأردن، 4 ماي، 22 جوان، 1996م.
- **محمد الزركان:** اللّسانيات وبرمجة اللّغة العربية في الحاسب، ندوة استخدام اللّغة العربية في تقنية المعلومات، الرياض،1992م.
- **محمد الحنّاش:** نظرية حاسوب- لسانية في بناء معاجم آلية للغة العربية: محاولة في التأصيل المسجل العالمي لندوة استخدام اللّغة العربية في تقنية المعلومات، مكتبة الملك عبد العزيز العامة، الرياض 1993م.
- **محمد صيني:**
  - ❖ نحو معجم عربي للتطبيقات الحاسوبية، ندوة استخدام اللّغة العربية في تقنية المعلومات، الرياض 1992م.
  - ❖ نحو معجم عربي للتطبيقات الحاسوبية"، السّجل العلمي لندوة استخدام اللّغة العربية في تقنية المعلومات، مكتبة الملك عبد العزيز العامة، الرياض، 1993م.
- **محمد زكي خضر:** نحو معالجة الدلالة في اللّغة العربية عبر قواعد البيانات" "دراسة أولية لنص القرآن الكريم" المؤتمر الوطني السابع عشر للحاسب الآلي (المعلوماتية في خدمة ضيوف الرحمن) جامعة الملك عبد العزيز، المدينة المنورة، 2004م.
- **مراد لوكام:** مشروع المعجم الحاسوبي التفاعلي للغة العربية، مقترحات حول إعداد المشروع، الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية، أبريل 2008م ، المنظمة العربية للتربية والثقافة والعلوم، مدينة الملك عبد العزيز للعلوم والتقنية، الرياض.

- **محمد حسن عبد العزيز**، محمد يونس الحملاوي، المعتز بالله السعيد طه: المعجم الحاسوبي للغة العربية، الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية، أبريل 2008.
- **محمد زايد:** دراسة في المعاجم الفرنسية الحاسوبية، الاجتماع الثاني لخبراء المعجم الحاسوبي التفاعلي للغة العربية، 1429/2004/29هـ (2008/5/7-5/5).
- **محمود فهمي حجازي:** المعجمات العربية وموقعها بين المعجمات العالمية، المجلس الوطني للثقافة والفنون والآداب، بحوث (ندوة خاصة بمناسبة الانتهاء من تحقيق وطباعة معجم تاج العروس من جواهر القاموس، الكويت، 9- 10، فبراير 2002.
- **محمود الديكي:** توصيف مركب العدد في اللغة العربية للحاسب الآلي، المعالجة الآلية للغة العربية، وقائع الندوة الدولية، جوان 2006. إشراف عبد الفتاح حمداني، خالد الأشهب، محمد الراضي، جامعة محمد الخامس، السويسي، معهد الدراسات والأبحاث للتعريب الرباط، جانفي 2007 م.
- **مروان البواب:** منهج إعداد المعجم العربي الحاسوبي للغة العربية، الاجتماع الثاني لخبراء المعجم الحاسوبي التفاعلي للغة العربية، 1429/5/2-4/29هـ (2008/5/7-5/5).
- **منظمة الأمم المتحدة للتربية والعلوم والثقافة:** تقدم اللّسانيات في الأقطار العربية، وقائع ندوة جهوية، أبريل 1987 م- الرباط، ط1، 1991 م.
- **فريق المعالجة الآلية للغة العربية:** التحليل الصرفي للأسماء العربية، تحت إشراف الأستاذ عبد الفتاح حمداني (المعالجة الآلية للغة العربية) ندوة دولية CITAL '07
18- 19- جويلية، معهد الدراسات والأبحاث للتعريب- جامعة محمد الخامس السويسي- الرباط- يونيو 2007 م.
- **عبد الله أبو هيف**، مستقبل اللغة العربية " حوسبة المعجم العربي ومشكلاته اللغوية والتقنية أنموذجا"، اللغة العربية، المجلس الأعلى للغة العربية، الجزائر، العدد العاشر، 2004م.
- **عبد الواحد خيري**، معجمية الإدراكات والتصورات في الأبنية اللغوية، وقائع الندوة الدولية حول المعجم العربي واقع وآفاق أيام 27/26/25 سبتمبر 2003م بالدار البيضاء.
- **عبد الفتاح حمداني، خالد الأشهب، محمد الراضي:** المعالجة الآلية للغة العربية، وقائع الندوة الدولية يونيو 2006، معهد الدراسات والأبحاث للتعريب، جامعة محمد الخامس،السويسي، الرباط، يناير 2007م.
- **عبد الغني أبو العزم وآخرون:**
  - وقائع الندوة الدولية ، المعاجم العربية الواقع والآفاق، أيام 27/26/25 سبتمبر 2003م بالدار البيضاء.

❖ المعجم الحاسوبي العربي: التصور والمنهجية، الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية، المنظمة العربية للتربية والثقافة والعلوم، مدينة الملك عبد العزيز للعلوم والتقنية، الرياض، السعودية، أبريل 2008م.

- **علاء الدين صلاح العجماوي:** المعالجة الآلية للغة العربية بين الواقع والتحديات، المحاضرة رقم 2، الموسم الثقافي – التاسع عشر 2001– مجمع اللّغة العربية الأردني.
- **علي فرغلي**، الهندسة اللغوية والثورة المعلوماتية، المؤتمر الثاني عن التحيز: حوار الحضارات والمسارات المتنوعة للمعرفة، القاهرة، 2007.
- **عمر مهديوي:**

❖ التوليد الصوري للأسماء البسيطة في اللغة العربية، المؤتمر الدولي الثاني حول هندسة اللغة وهندسة اللغة العربية، 27–28، يونيو2005م.

❖ تأهيل اللغة العربية وتهيئتها حاسوبياً، الندوة الدولية عن اللغة العربية ومتطلبات التهيء اللغوي ، معهد التعريب.

❖ اللغة العربية والهندسة العربية، مدخل نظري، الندوة الدولية عن اللغة العربية والنظريات اللسانية، كلية الآداب والعلوم الإنسانية، فاس، سايس، 21 – 22 نوفمبر 2007م.

- **غيتري سيدي محمد، محمد عبد الرحيم الأمين:** المعالجة الآلية للنحو العربي" الملتقى الوطني الثالث لقضايا النمو العربي الواقع والآفاق تيارت 21/20 أفريل 2009.
- المعالجة الآلية للغة العربية، وقائع الندوة الدولية CITALA'07، معهد الدراسات والأبحاث للتعريب، الرباط ، المغرب، يونيو 2007م.
- **نهاد الموسى**، مقدمة في تمثيل الكفاية اللغوية للحاسوب، ورقة مقدمة إلى ندوة "الهوية اللغوية والعولمة" جامعة البترا الأردنية الخاصة، 2003.
- **وليد أحمد العناتي**، الدَّليلُ نحو بناء قاعدة بيانات لِلِّسانيّاتِ الحاسوبيّة العَرَبِيّة،، ندوة تقنية المعلومات والعلوم الشرعية والعربية، ندوة (تقنية المعلومات والعلوم الشرعية و العربية) في كلية علوم الحاسب والمعلومات بجامعة الإمام في 16–17/صفر/1428هـــــ ــ الموافق 6–7/مارس 2007م.
- **محمد رشاد الحمزاوي**، مقترح لوضع نموذج للمعجم العربي الحديث، المعاجم العربية الواقع والآفاق، الدراسات المعجمية، الجمعية المغربية للدراسات المعجمية،العدد السادس،2007م.

2- الأجنبية:

1. **Bessou Sadik ,Louail Mohamed, Refoufi Allaoua, Kadem Zehour, et Touahria Mohamed** (un système de lemmatisation pour les applications de TALN) colloque International. CITALA'07- 18-19, juin/2007 Université Mohammed V, Souissi, Institut d'études et de Recherches pour l'arabisation, juin 2007.
2. **Abdefattah Hamdani,** « Traitement automatique de la langue arabe », Actes du colloque juin 2006 ; Université Mohamed V Souissi Institut d'atudes et de recherches pour l'aribisation, janvier 2007.
3. **ABDERRAHIM MOHAMED EL AMINE,** Un analyseur morphologique pour l'arabe voyellé ou non: 1ère conférence Internationale « Système d'information et Intelligence Economique » SIIE' 2008 Hammamet- Tunisie ; 14- 16 Février 2008- Proceedings tome II The édition.
4. **Ahmed Haddad, Henda Benghezala,** "structuration automatique du lexique au sein du dictionnaire arabe" 4 international conférence : science of electonic technologies of information and telecommunication SETIT 2007, March 25-27 2007, Tunisia.
5. **ATTIA Mohamed A,** « Devloping Robust Arabic Morphological Tranducer Using Finite State Technology In 8 th annual CLUK research colloquium.
6. **Bessou S, Saadi A, Touahria M** (un système d'indexation et de recherche de textes en arabe (SIRTA), LANIA 2007, chlef, Algerie, (2007)
7. **Bouzidia, Fouad Soufiane, Guy LAPALME,** « un système de resumé de textes en arabe, 2ème Congres International sur l'ingénierie de l'arabe et l'ingénierie de la langue : Alger, (2005).
8. **CHEN A. And Gey F ;** « Building an Arabic Stemmer For Information Retrieval », Proccedings of the Eleventh text Retrieval Conference (TREC 2002) National Institute of Strandard and technology, Nov 18/22 (2002).
9. **LARKEYL. S, BALL ESTEROSL AND CONNELLM**, Improving Stemming For Arabic Information Retrieval : Light Stemming and co – Occurrence analysis, C Sigir 2002à Tampere, Finland August 2002.
10. **LARKEYL. S, BALLESTEROSL, And CONNEL M,** Improving Stemming for Arabic Information Retrieval : Light Steming and Co-occurrence Analysis, (SIGIR 2002). Tampere, Finland August (2002).
11. **Mohamed. Tayeb. LASKRI et R. MAHDJOUBI** « Traitement automatique de la langue arabe en vue d'une traduction automatique des textes vert la langue Française » Proc. 3eme journées Internationales de traitement de données textuelles. JADT'95 ROME ; Italie. 11-13. DECEMBRE 1995.
12. **Motasem ALRAHABI, Ghassem MOURAD, Brahim DJIOUA ;** Filtrage sémantique de textes en arabe en vue d'un Prototype de résumé automatique. (le traitement automatique de la langue arabe-, JEP-TALN 2004, Fès 19-22 avril 2004.
13. **Melc'uk Igor,** Introduction a la lexicologie explicative et combinatoire,1995
14. **Nwersi A, Tahaghghis, Falk Scholer,** « traitement des textes arabes pour l'indexation et la recherche d'information, CITALA 2007, Rabat Maroc, (2007).

15. Zaafrani Riadh, "un dictionnaire électronique pour apprenant de l'arabe (langue second) basé sur corpus, JEP-Taln 2004, Traitement automatique de l'arabe, fès, 20 Avril 2004.

16. ZEMIRLI Zouhir, un analyseur destiné à l'aide à la construction d'une base de données lexicales de la langue arabe, colloque International « langue situées, technologie et communication » IERA 1996.

ح – البحوث والمقالات:

1– مجدي صوالحه وإيرك أتول، "توظيف قواعد النحو والصرف في بناء محلل صرفي للغة العربية"، جامعة ليدز، المملكة المتحدة. بحث مقدّم لاجتماع خبراء المحلّلات الحاسوبية الصرفيّة للغة العربية، دمشق, 2009.

2– حسن مظفر الرزّو،" سبل تطوير محلل الصرف الآلي المستخدم في حوسبة الموسوعات العربية"، دراسة نقدية لغوية لعدد من المحللات الصرفية، مروان البواب، عضو مجمع اللغة العربية بدمشق.

3– ميساء عبد الكريم ناصر ومرتضى عباس فالح وصبا عبد الواحد، " نظام آلي لفهم اللغة العربية " ، مجلة بابل,المجلد(الرابع عشر) العدد الثالث.

4 – محمد رفعت علي الحفني، " تطوير وتصميم وإنتاج محرك بحث اشتقاقي خاص بالنص القرآني"، أعمال ندوة القرآن الكريم والتقنيات المعاصرة، وزارة الشؤون الإسلامية والأوقاف والدعوة والإرشاد.

5 – علياء سلمان صابر، سلمى عبد الباقي محمود، علي فاضل مرهون.

" تصميم نظام استخلاص المعلومات من بعض فقرات النصوص العربية " بحث مستل من أطروحة دكتوراه، .

6– سلوى السيد حمادة، المحللات الصرفية للغة العربية ، بحث مقدّم لاجتماع خبراء المحلّلات الحاسوبية الصرفيّة للغة العربية ، دمشق،2009م.

7– هلال بن الحسين، "حاجة القاموسية العربية إلى حوسبة القواميس العربية الورقية ووضع القواميس الإلكترونية الجديدة"، الدراسات المعجمية، منشورات الجمعية المغربية للدراسات المعجمية، الرباط، المغرب، العدد السابع و الثامن، 1430هـ 2009م،

خ– مواقع الويب:

- موقع شبكة الفصيح لعلوم اللغة العربية:

http://www.alfaseeh.net

- المعجم والقاموس، المنظمة الإسلامية لتربية والعلوم والثقافة:

http://www.arabization.org.ma

- موقع صخر:

http://www.sakhr.com

- منتجات القارئ الآلي:

http://www.sakhr.com/product_a/OCR

- تقنيات التحليل الصرفي:

http://www.sakhr.com

- تقنيات المدقق الإملائي:

http://www.sakhr.com

- عنوان الموقع:

http://www. Arabic linguistics. Net Computional Approaches Arabic Morphology

- عنوان الموقع:

fossaire/A-htm.Glossaire de linguistique sperso-orange. Fr/idelahttp :// page Computationnelle

- عنوان الموقع: مفهوم قواعد المعارف:

.ahewar.orghttp://www

- عنوان الموقع: آفاق النشر الالكتروني:

.comhttp://www.alarabimag

- الصرافة الحاسوبية العربية محاولة في التأصيل:

.http://www.ahewar.org

- موقع الألوكة، دراسات، سبل تطوير محلل الصرف الآلي.موقع الويب:

. alukah.nethttp://www

- الشركة الهندسية لتطوير نظم الحاسبات RDI مصر، موقع الويب:

eg.Comhttp://www.rdi-

- عنوان الموقع: مواكبة العربية للغة المعلومات وعصر التقنية،

http:// www. Iseco. Org. Ma

- موقع الويب: مجلة العربية 3000

.nethttp://www.arabcin

- في معجم حاسوب للغة العربية:

www.alesco.org.tn

- مقهى علوم اللغة العربية وآدابها:

http://www.almujam.org/DOC/Qasim

- عنوان الموقع: سي++ (لغة البرمجة) ويكيبديا، الموسوعة الحرة :

http://ar.wikipedia.org

- عنوان الموقع: منتديات الحاسب الآلي ونظم المعلومات. موقع الويب:

http://www.uqucs.com

- عنوان الموقع: شركة حرف لتقنية المعلومات،موقع الويب:

http://www.harf.com

ثبت الموضوعات

ثبت الموضوعات

الموضوع	الصفحة
مقدمة	أ – ح
مــدخــل: المعجم العربي بين اللغة والحاسوب -آفاق الالتقاء-	**03 – 30**
أولا: وضع نموذج الكفاية المعجمية لبناء معجم إلكتروني للغة العربية	06 – 09
ثانيا: الهندسة المعجمية للغة العربية	09 – 10
ثالثا: قراءة في موضوع البحث : فرضية الدراسة	10 – 17
الطبيعة الرياضية والانصهارية للغة العربية	10 – 15
الاطراد الخوارزمي ونظام اللغة العربية	15 – 17
رابعا: خوارزميات توليد و تحليل المفردات العربية	17 – 21
الخوارزميات الحاسوبية	18 – 19
الخوارزميات لسانيا	19 – 20
خوارزميات التوليد والتحليل	20
خوارزميات التحليل الصرفي	20 – 21
خامسا: مقاربة التحليل الصرفي	21 – 24
خوارزميات المحلل الصرفي	22 – 23
خوارزميات التحليل الصرفي التبادلي	23
خوارزميات التحليل الصرفي المصدري	23 – 24
سادسا: خوارزميات التوليد الصرفي في بناء معجم إلكتروني للغة العربية	24 – 26
التوليد الصرفي المعجمي	24
التوليد الصرفي النصي	24
التحليل المعجمي – التركيبي وبناء المعاجم الآلية للغة العربية	25– 26
سابعا: أهداف الدراسة	27 – 28
ثامنا: العوائق والحلول المقترحة للنهوض بمشروع المعجم الإلكتروني للغة العربية	28– 30

الباب الأول : نحو معجم إلكتروني للغة العربية في ضوء اللسانيات الحاسوبية : - قضايا نظرية ومنهجية -	**31- 113**
الفصل الأول :المعجم العربي: نحوتوصيف جديد في اللسانيات الحاسوبية	32- 76
أولاً: الإطار العام لمفهوم اللّسانيات الحاسوبية	33- 46
تحديدات أولية	33 – 38
مفهوم اللّسانيات الحاسوبية	39- 40
المصطلح بين النظري والتطبيق	40- 46
ثانيا: اللسانيات الحاسوبية والمعالجة الآلية للغات الطبيعية	46- 48
ثالثا: مكونات اللسانيات الحاسوبية ووظائفها	48- 52
رابعا: الجوانب التطبيقية للسانيات الحاسوبية	53 -63
خامسا: اللسانيات الحاسوبية في خدمة الدراسات اللغة العربية: تجارب ورؤى مستقبلية	64- 73
التعاون والتنسيق بين اللغويين وعلماء الهندسة والحاسوب : المطلب الذي لا بدّ منه	64- 73
سادسا: اللسانيات الحاسوبية في خدمة المعجم العربي	74- 76
الفصل الثاني :حوسبة المعجم العربي: ـ نحو مقاربة لسانية حاسوبية لمعاجم آلية عربية	**77- 113**
أولاً: المعجم العربي من العين إلى الحاسوب	78- 80
ثانيا: المعجم العربي والحوسبة العربية	80- 88
أهمية الحاسوب ودوره في صناعة المعجم العربي	82- 85
استخدام الحاسوب في حوسبة المعجم العربي	85- 88
ثالثا: من المعاجم الورقية إلى المعاجم المحوسبة: نظرة تاريخية	88- 94
المعاجم المحوسبة	89- 91
الهندسة اللغوية والصناعة المعجمية	91- 92
الصناعة المعجمية للمعجم الحاسوبي	92- 93
رابعا: المعجم الإلكتروني للغة العربية : المشروع الذي لابد منه	95
المعاجم الحاسوبية والمحوسبة: نظرة عامة	97- 98
تسنين المعاجم الإلكترونية:	99
خامسا: قواعد المعطيات النصية لتطوير المعجم الحاسوبي (الإلكتروني)	100- 108
الملامح الأساسية لحوسبة المعجم العربي	100- 104
الأدوات الحاسوبية لتطوير المعجم العربي	105- 106
سادسا: الفرق بين المعجم الورقيّ والمعجم الآلي	108- 112
خلاصة	113

الباب الثاني :المعالجة الآلية للغة العربية لبناء قاعدة البيانات المعجمية آليا	**114 - 302**
الفصل الأول :المعالجة الآلية للغة العربية - النظام الصرفي العربي – أنموذجا -	115- 214
أولاً: الإطار العام لمعالجة اللغة العربية آليًا	116- 143
اللغة وعصر المعلومات –	116- 118
خصائص اللغة العربية من منظور لساني	118- 131
خصائص اللغة العربية من منظور صرفي – حاسوبي	131- 134
المعالجة الآلية للغة العربية:	134- 141
الوظائف الأساسية لنظم المعالجة الآلية للغة:	141- 143
ثانيا: تحديد الإطار العام لمعالجة المعلومات و المعارف	143- 150
معالجة المعلومات	143- 147
معـــالجة المعـــارف	147- 149
المقارنة بين معالجة "العربية" و"الانجليزية" آليًا	149- 150
ثالثا: المعـــالجة الآلية للصرف في اللغة العربية	150- 185
استخدام النظام الصرفي العربي كأنموذج للمعالجة الآلية في الدراسة التطبيقية	151- 153
الصرف وحوسبة اللغة	165- 166
خصائص الصرف العربي للمعالجة الآلية : نظرة عامة	166- 169
الصيغة الصرفية والميزان الصرفي في اللغة العربية	169- 174
أهمية الميزان الصرفي	175
الأصـــول والجـــذور في اللغة العربية	176- 180
أهمية الأصول الثلاثية في المعجم العربي	181- 182
رابعا: الإطار العام للمعالجة الآلية للصرف العربي	186- 193
المعالج الصرفي متعدد الأطوار	187- 188
أسس مقترحة لمعالجة الصرف العربي آليا	190- 193
خامسا: تطبيقات المعالجة الآلية للغة العربية	194- 210
المحلل الصرفي	195- 204
المحلل النّحوي	205- 209
المحلل المعجمي	209- 210
سادسا: المعالجة الآلية للدلالة في اللّغة العربية	210- 214
أهمية المعالجة الآلية للدلالة في العربية	211

الدلالة من منظور لساني– حاسوبي	211
قواعد بيانات المعاني والدلالة للمعالجة الآلية	212
الفصل الثاني: هندسة المعجم الإلكتروني-رؤية جديدة لمعجم آلي للمعالجة الآلية للغة العربية-	215– 300
أولاً: الإطار العام للمعجم العربي الإلكتروني	216– 243
المعجم الإلكتروني:إشكالية التعريف	217– 222
أهمية المعجم للمعالجة الآلية للغة العربية	224– 225
مزايا المعجم الإلكتروني	225– 226
طرق بناء المعجم الإلكتروني	226– 230
التمثيل الحاسوبي للبنية الأساسية للمعجم الإلكتروني	230– 243
ثانيا: المعجم العربي الإلكتروني بين النظرية والتطبيق: نماذج لقاعدة البيانات المعجمية للمعاجم الإلكترونية	243– 283
القاموس المفهومي الإلكتروني المرتكز على قواعد المعارف	245– 250
المعجم الآلي التفسيري التأليفي (DiCo)	250– 261
بنية قاعدة معطيات القاموس التفسيري التأليفي (Dico)	264– 265
المعاجم الآلية المبنية بنظام ديلا (DELA)	265- 266
الأدوات المعلوماتية المطبقة في معالجة مداخل المعجم الإلكتروني	266– 270
طبقات المعاجم الإلكترونية من البسيط إلى المركب	271–273
مشاكل وصعوبات في تمثيل المعلومات في المعجم الآلي العربي	273– 274
معجم الجذور الآلي العربي للجذور: (DEAR)	274– 275
معجم المفردات البسيطة(DEAMS)	275– 276
معجم المفردات البسيطة المعربة :(DEAMSF)	276– 277
قاعدة بيانات المعجم العربي الآلي للتعبيرات المركبة غير المعربة والمعربة :(DEAMC)	277- 283
ثالثا: بناء قاعدة بيانات معجمية لمعجم إلكتروني	284–299
قاعدة المعطيات المعجمية	284– 286
تنظيم المعجم الإلكتروني للغة العربية في هيئة قاعدة بيانات معجمية	288– 292
الخدمات المعجمية لقاعدة البيانات المعجمية	297– 298
العلاقات الداخلية لقاعدة البيانات المعجمية لمنظومة المعجم	298– 299
رابعا: سبل تطوير قاعدة بيانات معجمية لبناء معجم آلي عربي	299– 300

الباب الثالث توظيف تقنيات المعالجة الآلية للغة العربية لإعداد وتطوير معجم للغة العربية	**303**
الفصل الأول :المعجم الإلكتروني بين النظرية والتطبيق: نظرة في بعض الدراسات المعجمية، " قراءة في الحصيلة واستشراف للمستقبل"	304- 365
أولا: الدراسات الغربية: تجربة :LADL	306- 307
المعجم الإلكتروني للمفردات البسيطة: DELAS	306- 307
المعجم الإلكتروني الفونيمي : DELAP	307
المعجم التركيبي: Grammaire Lexique	307
ثانيا: الدراسات العربية: التجربة العربية: الواقع، والمأمول	308 – 364
شركة صخر للمعلوماتيات والإلكترونيات: تجربة SAKHR	309
تجربة الدكتور الحناش في المعاجم الآلية للغة العربية	309- 312
تجربة الدكتور أبو العزم من الغني إلى الغني/ الشامل	312- 316
تجربة مروان البواب إعداد المعجم العربي الحاسوبي للغة العربية	317- 332
تجربة الدكتور نبيل علي	333
تجربة الحاج صالح : مشروع الذخيرة اللغوية العربية	333 – 334
تجربة سلوى حمادة : نحو إنشاء قاعدة بيانات معجمية آلية في شكل معجم كوني هرمي (مكة)	335
تجربة عبد المجيد بن حمادو	335 – 336
تجربة مراد لوكام لإنجاز مشروع المعجم الحاسوبي التفاعلي للغة العربية	336
تجربة معجم حاسوبي أحادي للناطقين بغير العربية للدكتور وليد أحمد العناتي	336- 353
تجربة المعجم الاسلامي للدكتورة فاتن الخولي	354- 360
جهود حرف في مجال المعاجم الإلكترونية	361- 364
خلاصة	365
الفصل الثاني : نحو بناء معجم إلكتروني للمعالجة الآلية للغة العربية باستخدام نظام آلي لاستخراج جذور الكلمات العربية مفردات ألفاظ القرآن الكريم "أنموذجا.	366
أولا: إعداد تصور منهجي للمعجم الإلكتروني اللغة العربية	367
ثانيا: عرض معجم إلكتروني للمعالجة الآلية للغة العربية باستخدام نظام آلي لاستخراج جذور الكلمات العربية	367- 370
مفردات ألفاظ القرآن الكريم كعينة في الدراسة التطبيقية	370- 371
ثالثا: المعالجة الآلية للمدونة: (مفردات ألفاظ الفرآن الكريم)	371- 373
رابعا: الدراسة التطبيقية للمشروع	373- 374

الموارد اللسانية: قاعدة بيانات الصرف	374 -380
خامسا: الدراسات السابقة	380 -381
سادسا: آلية التنفيذ (التطبيق)	381
لغة البرمجة المستخدمة	381 -383
تفسير عمل برنامج المعجم الإلكتروني للمعالجة الآلية للغة العربية	406 -407
عرض بعض الشاشات للبرنامج كأمثلة توضيحية من المعجم الإلكتروني	407 -410
سابعا: نتائج الدراسة التطبيقية	410 -415
خاتمة	417 -420
قائمة المصادر والمراجع	422 -451
ثبت الموضوعات	

RESUME :
Cette recherche aborde la question relative aux principales applications linguistiques informatiques sur la langue, à savoir le traitement automatique de la langue arabe vers l'élaboration d'un dictionnaire électronique qui n'est autre que le but de cette recherche qui est en fait une contribution à relever le grand défi à justifier le pouvoir de la langue arabe à adopter et transférer diverses technologies modernes, ce projet propose une méthode efficace pour créer et utiliser le dictionnaire électronique pour le traitement automatique de la langue arabe en se basant sur les techniques linguistiques pour le traitement automatique de la langue arabe.

L'approche repose sur les notions de schème (un des points forts de la morphologie de la langue arabe) selon des structures spécifiques appelées schématisation. L'avantage de cette méthode est qu'elle n'utilise pas le dictionnaire des mots. En effet l'identification des mots se fait d'une façon intelligente et ce par l'extraction de la racine et du schème approprié à chaque mot analysé, Cette recherche est un maillon d'une longue chaîne relative à un gros projet dans le cadre d'une théorie linguistique informatique qui repose sur les caractéristiques de la langue arabe.
Mots clés : Linguistique informatique ,Traitement automatique de la langue arabe (TALA), dictionnaire automatique, le schème, extraire la racine, analyse morphologique.

ABSTRACT :
This paper tries to tackle on important issue that concerns computational linguistics.It tries to apply the field of computational linguistics on the Arabic language in building an electronic dictionary so as to challenge other languages in terms of its potential vocabulary and ability to embrace and transmit different modern techniques and coping with them. This project offers an effective method to create and use the electronic dictionary for automatic processing of the Arabic language, based on linguistic techniques for automatic processing of Arabic.

This approach builds on the concepts of scheme (one of the strengths of the morphology of the Arabic language). The advantage of this approach is that it does not use dictionary words, but identification of individual words in an intelligent manner by the root extraction and appropriate pattern of words each word is analyzed, This research is a part of a big project that deals with a computational approach whose aim is to rely upon the characteristics of Arabic language.
Key words: computational linguistics ,Automatic processing of Arabic (TALA), automatic dictionary, Arabic morphology, the scheme, the root extract, morphological analysis.

ملخـــــــص:

يعالج البحث قضية من أبرز قضايا تطبيقات اللسانيات الحاسوبية على اللغة ألا وهي تطبيقات المعالجة الآلية للغة العربية لبناء معجم إلكتروني كهدف لهذه الدراسة، وإسهاما كذلك في رفع التحدي الكبير الذي يبرر قدرة لغة الضاد على احتضان ونقل مختلف التقنيات الحديثة والتجاوب معها، ، فالمشروع يقترح طريقة فعالة لإنشاء واستغلال معجم إلكتروني للمعالجة الآلية للغة العربية اعتماداً على التقنيات اللغوية للمعالجة الآلية للغة العربية.

ترتكز هذه المقاربة على مفهوم الوزن أو الميزان الصرفي (أحد نقاط القوة في الصرف العربي) وفق تراكيب خاصة تسمى "التفعيلات"، الجانب الإيجابي في هذه الطريقة أنها لا تستعمل معاجم الكلمات، بل تقوم بالتعرف على مختلف الكلمات بطريقة ذكية عن طريق استخراج الجذر والوزن المناسبين لكل كلمة محللة، فالبحث هو حلقة من حلقات مشروع ضخم وذلك في إطار نظرية لسانية حاسوبية، هدفها الاعتماد على خصائص اللغة العربية.
الكلمات المفتاحية: اللسانيات الحاسوبية، المعالجة الآلية للغة العربية، المعجم الالكتروني،الصرف العربي ،الميزان الصرفي، استخراج الجذر، التحليل الصرفي.