حوسبة النظام اللغوي العربي
ـالمعجم الآلي عند البروفسور عبد الرحمن الحاج صالح
-أنموذجا-

د/ راضية بن عريبة

جامعة حسيبة بن بوعلي -الشلف-

إنّ الثورة التكنولوجية الحاسوبية الحديثة ألقت بظلالها على اللّغاتِ الطّبيعية محدثةً الانقلاب التاريخيّ في المجالين المعرفي واللّغوي، وغدتِ الحاجةُ ماسّةً لاستجابةِ اللّغة الطّبيعية لذلك التّأثيرِ لمواكبةِ ظاهرةِ التّسْريعِ الّتي وُسِمَ بها هذا العنصر.[1]

لذلك نبتت اللّسانيات الحاسوبيّةُ العربيّةُ على غرارِ اللّسانيات الحاسوبيّةِ العامّةِ استجابةً لدواعي حضارية واستراتيجية يُنشِدها مستقبل اللّغة العربيّةِ.[2]

1/ **تعريف اللسانيات الحاسوبيّة**: اللّسانيات الحاسوبيّة (computational linguistics)، وتسمّى أيضاً علوم اللّغة الحاسوبيّة، وهي علومٌ حديثةٌ تستخدمُ الحواسيب في تحويل النّصوصِ والمعلوماتِ اللّغويّة إلى لغاتِ الحاسبِ الرّقمية لتحليلها، وترجمتها للغاتٍ أخرى[3]

تعدُّ دراسةُ اللّغة العربيّةِ باستخدامِ اللّسانيات الحاسوبيّةِ من أحدثِ الاتّجاهاتِ اللّغوية في اللّسانيات العربيّة المعاصرةِ.

2. نشأة اللسانيات الحاسوبية: تمّ اختراعُ جهازِ الحاسوب - كما تذكر المصادر- في أواخرِ النّصفِ الأوّلِ من القرنِ العشرين، وتحديداً عام 1948م وأصبح منذ ذلك التّاريخ متاحاً للإفادةِ منه في جميعِ مجالاتِ الحياةِ، ومختلفِ العلومِ والمعارفِ الإنسانية، وتطوّرت تقنية هذا الجهاز عبر السّنواتِ تطوّراً مذهلاً منذ ظهورِ الجيلِ الأوّلِ من الحواسيبِ الآليّة سنة 1951م وحتّى ظهورِ الجيلِ الخامسِ منه سنة 1991م[4].فعلى المستوى الأمريكيّ يذكر الدّكتور (مايكل زار تشناك) أستاذ علمِ الدّلالةِ ومنظّم البرمجة اللّسانية الآليّة بجامعة (جورج تاون) أنّ العمل في اللّسانيات الآلية بدأ في قسم اللّسانيات بجامعة (جورج تاون) سنة 1954م، وذلك في حقل التّرجمة الآلية من اللّغاتِ الأخرى إلى اللّغة الإنجليزية وهذا يعني أنّ بداية الخمسينيّات من القرن الماضي شهدت ولادة المعالجة الآليّة للغاتِ البشريّة[5]

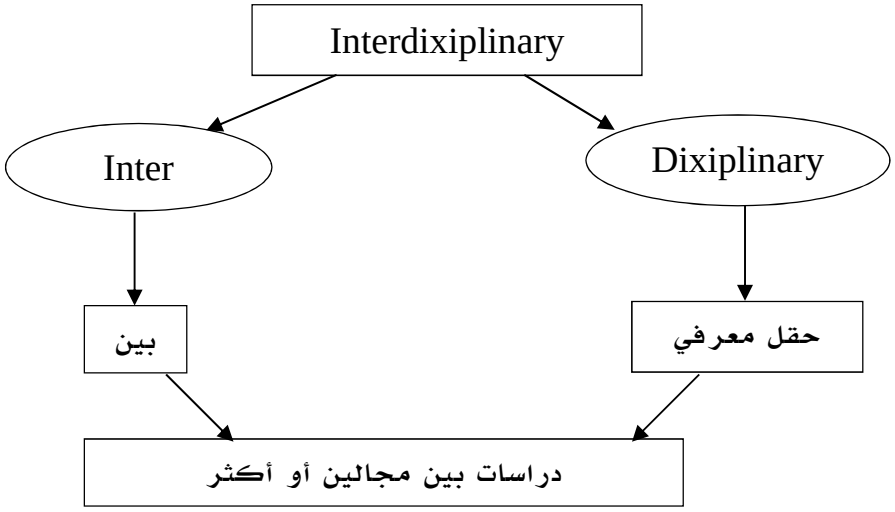
أمّا على المستوى الأوربيّ فتذكر المصادر أنّ أقدم محاولةٍ لدراسة اللّغة بواسطةِ الحاسوبِ كانت سنة 1961م، بجامعةِ قوتبرغ (Göteborg) السّويدية.

والبداية العلميّة لهذا الاتّجاهِ كانت في مركز التّحليلِ الآليّ للّغة بمدينة قالارات (Gallarat) بإيطاليا سنة 1962م[6] أمّا قضيّة

الاتّصال العلميّ بين الحاسوب والبحثِ اللّغويّ العربيّ، كما يذكر إبراهيم إنّيس، فكانت في الكويت عام 1971م، ويعدّ كتاب (اللّغة العربيّة والحاسوب للدّكتور (نبيل علي) سنة 1988م أوّل كتابٍ يتناول موضوع اللّسانيات الحاسوبية.

3-اللسانيات الحاسوبية والدّراسات البينيّة:

الدّراسة البينيّة: لقد أدّى تواصل المعارف وتشابكها إلى إيجاد ما يعرف بالدّراسات البينيّة (Interdixiplinary) التي تتداخل فيها التخصّصات، حيث تمّ التأكيد على طبيعة العلاقات بين المعارف الإنسانية المستمرّة في تقدّمها[7] تتكوّن كلمة البينيّة Interdixiplinary من مقطعين أساسيّين مقطع (Inter) وتعني بنية وكلمة نظام(dixiplinary) وتعني مجالاً دراسيًا معيّنًا، كما يوضّحه الشّكل الآتي[8]:

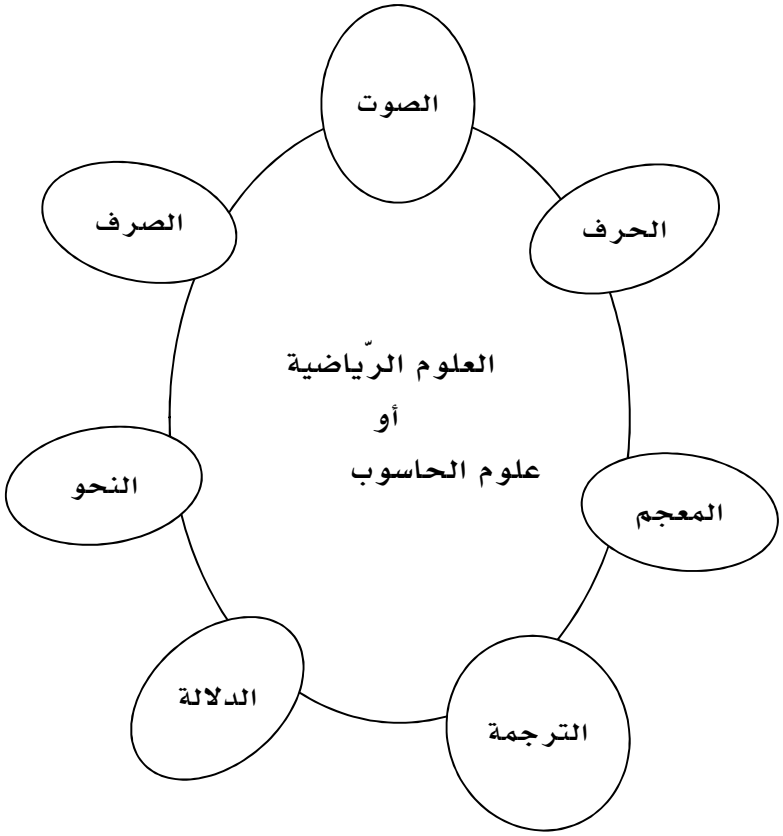


ومن هذا المنطلق فقد تمّ تعريف الدّراسات البينيّة من قبل (كْلايْن وويليم) على أنّها دراسات تعتمد على حقلين أو أكثر من حقول المعرفة الرّائدة، أو العمليّة التي يتمّ ُبموجبها الإجابة عن بعض الأسئلة أو حلّ بعض المشاكل أو معالجة موضوع واسع جداً يصعب التعامل معه بشكلٍ كافٍ عن طريق نظامٍ أو تخصّصٍ واحدٍ.[9]

فالنّظام فرعٌ من فروعِ المعرفة أو حقلٌ دراسيٌّ يقع ضمن أيٍّ من الثّلاث مجموعات أو التخصّصات التالية:[10]

- ✓ العلوم: الفيزياء، الكيمياء، البيولوجيا، الجيولوجيا، الزّراعة الهندسة.
- ✓ العلوم الاجتماعية: علم النّفس، القانون، علم الإنسان، الاقتصاد والعلوم السياسية، علم الاجتماع.
- ✓ العلوم الإنسانية: الفنون، الأدب، التاريخ، الفلسفة، الدّين، المسرح والموسيقى.

ويمكن إسقاط الدراسات البينيّة ضمن المعالجة الآلية للّغة العربيّة في المخطّط الآتي:



رسم توضيحي يبيّن إمكانية الدّمج بين العلوم الرّياضيّة أو علوم الحاسوب بعلوم اللّغة العربيّة

4ـ الهدف من الدراسات البينية: تهدف الدّراسات البينيّة في مجال المعالجة الآلية للّغة العربيّة بشكلٍ عامٍ إلى:

1. دمج المعرفة اللّسانية بالحاسوب.
2. الإبداع في طرق التفكير اللّساني الحاسوبي.
3. تحقيق التكامل بين مختلف العلوم اللّغويّة والعلوم الرّياضيّة والعلوم الحاسوبية.
4. إنّتاج المعرفة الكلّية لتوحيد الرّؤى الجديدة.
5. الاطلاع على التّجارب المحلّية والإقليميّة والعالميّة الخاصّة بهذا المجال.

5ـ اللسانيات الحاسوبيّة والمعالجة الآلية للغة:

1/المعالجة الآلية للّغة: تهتمّ المعالجة الآلية للّغة بدراسة الجوانب الحاسوبية للّغة والمشاكل اللّسانية والحاسوبية التي تواجه هذه المعالجة سواء أكانت هذه اللّغة منطوقة أم مكتوبة. وبناء نظام معالجة اللّغة العربيّة مهمّة معقّدة وصعبة، وذلك لصعوبة إدماج المعارف الصّوتيّة والصّرفية والنّحويّة والدّلالية في هذا النّظام.

تعريف بالمعالجة الآلية:

المعالجة: هي التّطبيق الآلي على مجموعة من نصوص اللّغة، وذلك بتغييرها وتحويلها، وإبداع شيء جديد اعتماداً عليها، ويتمّ ذلك باستعمال تقنيات وأدوات من علوم اللّسانيات والإعلام

الآلي والنّمذجة (Modélisation)، ويجب التفرقة بين وصف المعارف التي هي من وظيفة اللّسانيات والتعبير عن هذه المعارف في نماذج باستخدام تقنيات واستراتيجيات فعّالة مستمدّة من علم الحاسوب، وهي وظيفة علم اللّغة الحاسوبي[11]

الآلية: العمليّات الآلية هي التي تجري عن طريق الآلة، والتي تقابلها العمليات التي تجري بواسطة الإنسان والحاسوب هو الآلة التي تستعمل في معالجة اللّغة- الذي اخترع إجراء العمليات الحسابية، ووجب تطوره لمعالجة المعلومات ذات الصّلة بالطّبيعة اللّسانية، حيث أنّ المعالجة الآلية هي تتابعه حركات حسابية تقوم بها الآلة وفق تسلسل زمني أي إنّ برنامج المعالجة الآلية (Automatique Programme) يمكن أن يكون كلّياً(Total) أو جزئياً (Contraintes) حيث أنّ: [12]

أ- كلّي: يقوم الحاسوب بكلّ شيء.

ب- جزئي: يتدخّل الإنسان في بعض المراحل.

ومعالجة شيء لساني في الآلة كثيرة في وصف النّصوص اللّغويّة نفسها لذا يجب نمذجة مكوّنات النّصوص بطريقة واضحة ومتناسقة(Cohérence)[13]

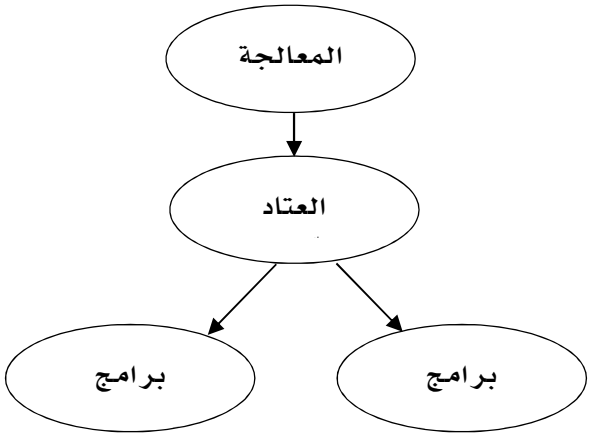
حيث يقرّ البحث اللّساني المعرفي بواقع علميّ مفاده أنّ النّماذج قد أصبحت تفرض نفسها كتقنية ومنهج. وهدف هذه التيّارات المعرفيّة بسبب طابعها العلمي والتكنولوجي [14]

2-مجالات البحث في المعالجة الآلية للغة: تنقسم مجالات البحث في المعالجة الآلية للّغة إلى ثلاثة مجالات أساسية:

الأوّل: مجال مشترك ونقصد به العتاد اللّساني.

الثاني: برامج بحثية.

الثالث: برامج تعليميّة.



العتاد اللساني:

- المعاجم الالكترونية المتكاملة (أصوات، صرف، تركيب)؛
- مولّد ومحلّل صرفيّات؛
- مولّد ومحلّل نحويّات؛
- المدقّق النّحوي والإملائي؛
- المشكّل الآلي.

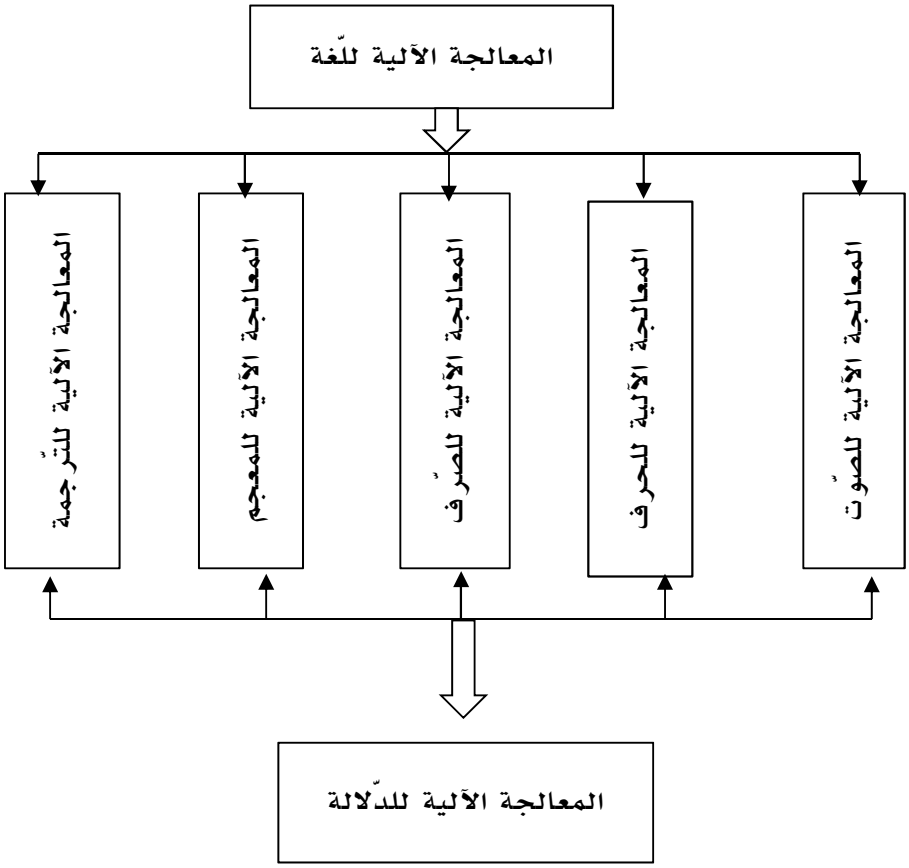
1- البحث العلمي:

- التعرّف البصري على الحروف سواء منها المطبوع أم اليدوي؛
- الترجمة الآلية والترجمة المسعفة بالحاسوب؛
- التوليف الصّوتي؛
- التوثيق الآلي؛
- الفهم الآلي للنّصوص.

2- التعليم:

- تصميم برامج تعليمية على عتاد إلكتروني؛
- تصميم برامج على الإنّترنت برامج عالمية؛
- وإذا خصّصنا أيّ برنامج تعليمي مصمّم لأيّ مرحلة من مراحل التعليم، سنجده يحصر خبرته في الجانب التربوي والجانب الحاسوبي، أمّا الخبرة اللّسانية التي تشكل عصبه المركزي فقلما تؤخذ في الحسبان، ممّا جعل منها برامج غير

قادرة على القيام بمهمّتها في عالمنا المتطوّر الّذي أصبحت فيه تقنية التّعليم تحتلّ الصّدارة في صناعة البرامج التعليميّة والبحثيّة[15] .ومن خلال هذه المحاضرات المحصورة في هذه المداخلة تنحصر المعالجة الآلية في المستويات الآتية:



المعالجة الآلية للمعجم العربي – عند البروفسور (عبد الرحمن الحاج صالح) أنموذجا

تتبوأ المعالجة الآلية للمعجم العربي مكانة متميزة في ميادين التعامل مع اللّغة العربية على وفق منظور معلوماتي، وتعدّ مدخلا أساسا وقاسما مشتركا لإدارة معظم نظمها الآلية التي تعنى بإنّشاء نظم المعلومات والمعارف.

يقصد بالمعالجة الآلية للمعجم العربي اعتماد نظم الحوسبة المتقدمة تستند إلى خوارزميات برمجية تستثمر المنطق المعجمي العربي في معالجة المفردة العربية عن طريق استخلاص العناصر الأوليّة لبنية الكلمة، ومباشرة تحديد سماتها المعجمية.

يعتمد المعجم الآلي المعجمي على معالجة مفردة واحدة حسب اختيار المستخدم، ويباشر بإظهار أهمّ الجذور المتاحة لتلك المفردة في الموسوعة قيد الدراسة، مع إتاحة الفرصة لانتقاء الجذر المناسب لعملية البحث.

ويلاحظ أنّ المعالج المعجمي الآلي متعدد الأطوار يمتاز بآلية عمل متقدمة من الناحيتين البرمجية البحتة والمعالجة المعجمية واللّغوية، الأمر الذي جعله يتبوأ مكانة بارزة في ميدان البحث الحاسوبي بالموسوعات، ونظرا لخلو الساحة العربية من المحللات المعجمية الآلية ترقى إلى مستواه

الرصين في المنطق اللّغوي والبرمجي المحكم، سنحاول دراسة آلية عمله لفهم المزيد عن هيكلته المعجمية اللغوية البرمجية.

أولا :المحلل المعجمي الآلي متعدد الأطوار:

المعالج الصرف معجمي: 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 ردّ التعديلات الفونولوجية التي تتمّ على حدود عناصر البنية الصرفية للكلمة إلى أصلها، كي تصبح صالحة لعملية المعالجة.

فعلى سبيل المثال عند هذا المعالج بعملية تفكيك الكلمة (والدين) يظهر أمامه أكثر من احتمال لعملية التفكيك الصرف معجمي، وكما يلي[16]:

* الاحتمال الأول: والد + ين ⟸ مثنى مذكر.

* الاحتمال الثإنّي: والد+ ين ⟸ جمع مذكر سالم.

* الاحتمال الثالث: و+الدين ⟸ كلمة الدين من الفعل الثلاثي دان.

*الاحتمال الرابع: و+ الدين ⟸ كلمة الدين من الفعل الثلاثي دان.

ولتفادي الضياع في متاهات تحليل المبنيات كالضمائر وحروف الجر وظرفي الزمان والمكان، ثم توفير آلية ذكية يمكن أن توظف من خلال هذا المعالج لمقارنة جذع الكلمة المعجمية المفكّكة مع قائمة من الكلم المبني باللّغة العربية .

وقد حاول الدكتور عبد الرحمن الحاج صالح ذلك من خلال المثال المولّد للاسم (أو حده الإجرائي) في ذخيرته.

ثانيا المعالج الإشتقاقي: يباشر المعالج الاشتقاقي عمله باستخلاص الجذور والصيغة الصرفية، الجذع أو الجذوع التي قام بتفكيكها المعالج الصرف معجمي، مع افتراض عدّ وجود علامات التشكيل تبدأ عملية الاستخلاص بمقارنة سلسلة حروف الجذع مع قائمة قوالب الهياكل الصرفية التي تتضمنها قاعدة بيانات، فعند تطابق سلسلة الحروف مع أخذ الهياكل، يقوم المستخلص بتمييز حروف الجر المناظرة لحروف العلّة الواردة في الهيكل المطابق[17].

بعد استخلاص الجذر، يقارن بمعجم الجذور المسموح بها للوصول إلى جذر واحد أو بضعة جذور لهذه الكلمة.

وبعدها يقدم المستخلص الاشتقاقي مخرجاته إلى المعالج الإعرابي على هيئة سلسلة الجذور والهياكل الصرفية المحتملة لجذور الكلمة وعلى الوجه التالي:

الجذر	الوزن	الاحتمال الصرفي
1-دين	فعل	والدين
2-		والدين
3- ولد	فاعل	والدين
4-	ولد	والدين

ثالثا: المعالج الإعرابي: تصبح الحالة الإعرابية للكلمة هدفا لهذا النظام، حيث يعمد إلى تمييزها بناء على الوسم الإعرابي الذي تتضمنه ممثلا في الإعراب بالحروف أو علامات حركات التشكيل الظاهرة، أو الحذف أحيانا[18]

رابعا: معالج التشكيل: يعمد معالج التشكيل إلى إعادة تركيب عناصر الكلمة لغرض مقارنتها بالكلمة المدخلة إلى ساحة المعالجة الآلية، وذلك للتأكد من التحليل الصرف المعجمي لبنيتها ويتم ذلك عن طريق إعادة تركيب ج\ع الكلمة، عبر صهر الجذر مع جميع الصيغ الصرفية المناظرة للهيكل الصرفي الذي حدده المستخلص والتي يسمح المعجم اللغوي بانطباقها عليها 19 فعلى سبيل المثال وفي حالة الهيكل الصرفي (مفعل) يقوم معالج التشكيل باسترجاع الصيغ الصرفية المناظرة له من ملف الصيغ الصرفية وهي: مفعل، منفعل ... بعد ذلك يقوم بالتأكد من صحة انطباق الصيغة الصرفية المحتملة على جذع الكلمة قيد الدراسة، ليقوم بعد ذلك بصهر الجذر مع كلّ صيغة صرفية منطبقة عليهن ثم ينفذ جميع

عمليات الإبدال والإعلال الواجبة تحت توجيه سلسلة قواعد الصرف الصوتية التي يتعامل معها، بعدها يدمج عناصر الكلمة الأخرى (السوابق واللواحق). مع جذع الكلمة الذي أعاد تركيبه.

وتستمر فعالية المعالج التشكيلي تركيب الكلمة المعجمية تامّة التشكيل لحين المطابقة التامة.

3- سبل تطوير أداء المحلل المعجمي في ضوء الذخيرة اللغوية:

يبدو جليا أنّ المحلل المعجمي الآلي متعدد الأطوار ذو فائدة كبيرة في ميدان البحث عن النصوص في الموسوعات العربية (على مستوى الكلمة مجموعة كلمات، أو جذر الكلمة) ويتصف بكفاءة جيدة وسرعة كبيرة، كذلك فإنّ التعديلات التي أجرتها العالمية على إصداراتها الجديدة قد غطت كثيرا من الجوانب التي كانت تشكل ثغرات ملحوظة في أداء المحلل المعجمي في الإصدار الأوّل.

بيد أنّ هذا الأمر لا يعني اكتمال دورة العمل المعجمي البرمجي في هذا المجال، بل يجب أن يشمل على جملة من الإضافات الجديدة إلى الخوارزميات لتطوير النموذج المعجمي الآلي ومن أهمّها: [20]ضرورة توفير بيانات معجمية تربط بين جذر الكلمة والصيغ الصرفية المنطبقة عليه من جانب وبين الكلمات المشتقة ومعانيها الصرفية معجمية من جانب آخر.

1- اعتماد مبدأ الإحصاء المعجمي لسلاسل الرموز اللغوية لترشيد حالات اللبس الزائفة وشبه الزائفة عن طريق توفير عدد معقول من الإحصائيات الصرفية لمعدلات استخدام الجذور والصيغ الصرفية والحالات التصريفية والإعرابية.

2- الاهتمام بزيادة كفاءة أداء المعالج المعجمي الآلي بحيث يمتلك القدرة على التعامل مع النصّ الخالي من التشكيل، لتجاوز عقبة الحاجة للنصّوص التامّة التشكيل، وجزئية التشكيل، واللّتان سوف تنصهران في ظاهرة غياب تشكيل المفردة العربية.

3- توظيف مدقّق إملائي لتمييز الأخطاء الإملائية للكلمة المدخلة إلى المعالج الآلي تلقائيا، في قطاعي الكلمة المشكلة وغير المشكلة للتقليل من حجم الضياعات بالتفتيش عن الكلمات.

4- استخدام آليات جديدة لتحسين أداء المعالج المعجمي الآلي في استرجاع المعلومات العربية – وذلك ما تسعى إليه الذخيرة اللّغوية – وتجاوز عقبة الاختيار اليدوي للمستخدم من قائمة احتمالات الكلمات والجذور.

خاتمة:

من خلال ما تمّ ذكره ولو بشكل مبسط، يمكن التوصل إلى النتائج التالية:

1-إنّ التحليل العربي إلى الكلمة المعجمية إلى مادة أصلية وصيغ يساعد المهندسين اللّسانين في وضع برمجيات لعلاج المفردات في الحاسوب وذلك بحصر كل الصيغ بمدلولاتها من الاستعمال الحقيقي للغة.

2-إنّ عملية الزيادة في المثال المولد للاسم عند الدكتور عبد الرحمن الحاج صالح تساعد على التحويل (تحويل الأصل) وتحديد المواضيع والمواقع الخاصّة أثناء المعالجة الآلية.

3-من خلال المحلّل المعجمي الآلي يمكن لنا دراسة النصوص آليا.

4-إمكانية فهرسته بكيفية آلية لكلّ الألفاظ المعجمية وذلك لوضع معجم إلكتروني شامل للّغة العربية.

5-الاعتماد على البنك المعجمي الآلي في البحث عن التطور المعجمي للألفاظ العربية، ومن ثمّ إمكانية وضع معجم تاريخي آلي دقيق للّغة العربية وذلك ما يسعى إليه مشروع الذخيرة اللغوية العربية.

6-استثمار مفاهيم الدكتور (عبد الرحمن الحاج صالح) في ظلّ أعمال الذخيرة اللّغوية وتطبيقها على أرض الواقع وذلك لخدمة اللّغة العربية.

الهوامش:

[1]- ينظر: فعاليّة اللّسانيات الحاسوبيّة العربيّة، عمود ديدوح، مجلّة ألأثر، الآداب واللّغات، جامعة قاصدي مرباح، ورقلة، العدد 08، 2009، ص: 90.

[2]- ينظر: المرجع نفسه، ص: 91.

[3]- ينظر: اللّغة العربيّة والحاسوب، نبيل علي، مؤسّسة تعريب، الكويت، 1988م، ص: 01.

[4]- ينظر: توظيف اللّسانيات الحاسوبية في خدمة الدّراساتِ اللّغوي العربيّة (جهود ونتائج)، عبد الرحمن بن حسن العارف، جامعة أمّ القرى، مكّة، ص: 01.

[5]- ينظر: المرجع نفسه، ص: 02.

[6]- ينظر: توظيف اللّسانيات الحاسوبيّة في خدمة الدّراسات اللّغويّة العربيّة (جهود ونتائج(، عبد الرحمن بن حسن العارف، ص: 03.

[7]- ينظر: برامج الدّراسات البينيّة واحتياجات السّوق، مركز البحوث والدّراسات الغرفة الصّناعية التجارية، الإدارة العامّة للبحوث والمعلومات، الرياض، د ط، 1432هـ، ص: 07.

[8]- ينظر: مقال: الدّراسات البينيّة، رؤية لتطوير التعليم الجامعي، إعداد أ. د عمار بن عبد المنعم أمين، عميد كلّية علوم الأرض، جامعة الملك بن عبد العزيز، ص: 02.

[9]- ينظر: الدراسات البينيّة، عمّار بن عبد المنعم أمين، ص: 02.

[10]- ينظر: المرجع نفسه، ص: 02.

[11]-ينظر: المعالجة الآلية للّغة العربيّة، إنّشاء نموذج لساني صرفي إعرابي للفعل العربي، فارس شاشة، مذكرة ماجستير في علم المكتبات والتوثيق، جامعة الجزائر، كلية العلوم الاجتماعية والإنسانية، الجزائر، 2008، ص: 13.

[12]- ينظر: المعالجة الآلية للّغة العربيّة، فارس شاشة، ص: 13.

[13]- ينظر: المعالجة الآلية للّغة، مسارات تحليل الجمل العربيّة باعتماد نماذج شبكات الانتقال، بن عيسى زغبوش، مصطفى بوعنّاني، مجلّة كلّية الآداب والعلوم الإنسانية، فاس، المغرب، العدد 14، 2006، ص: 80.

[14]- ينظر: المرجع نفسه والصّفحة. 3

[15]- ينظر: اللّغة العربيّة والحاسوب، محمّد الحنّاش، ص: 24.

[16]-ينظر: اللّسانيات-المجال والوظيفة و المنهج –سمير شريف اسيتية-الأردن- عمان-عالم الكتب الحديث ،ط1،2005،ص548.

[17]- ينظر : المرجع نفسه والصفحة.

[18]- ينظر: سبل تطوير محلل الصرف الآلي – حسن مظفر الرزو – ص 12

[19]- ينظر: سبل تطوير محلل الصرف الآلي- حسن مظفر الرزو- ص: 12

[20]- ينظر المرجع نفسه- ص 10