حوسبة المعجم اللساني عند اللغويين المحدثين العرب

د/ راضية بن عريبة
جامعة حسيبة بن بوعلي الشلف
الجزائر
radiabenariba@yahoo.fr

الهاتف:06.70.35.51.16

تاريخ الاستلام:2018/02/26 تاريخ لقبول :2018/05/30

ملخص

حظيت اللّغة العربية بجهودٍ كبيرةٍ من لدن الباحثين منذ ظهورِ الإسلامِ إلى يومنا هذا، وعملوا على جمع مواردها ووصفِ ظواهرها، ووضعِ النّظرياتِ حول بنائها واشتغالها، وتأصيلِ العلوم التي تعالج مواضيعها، وصنّفت في ذلك التصانيف المتعدّدة الأحجام والأغراض، وفي العصرِ الحاضرِ ساهم الباحثون العرب، والأجانب في إثراءِ البحوثِ في اللّغة العربيّة، وساعدهم في ذلك تقدّم الدّراسات اللّسانية، وانتشار الابتكارات العلميّة والتقنيّة.

كلُّ ذلك أدّى إلى ظهور مجالاتٍ علميّةٍ تتداخل فيها الدّراساتُ اللّغويّةُ بالدّراساتِ العلميّةٍ والتقنيّة من أجلِ تطويرِ اللّغة العربية ومعالجتها آلياً؛ أو كما أطلق عليها بالمصطلح العلميّ "اللّسانيات الحاسوبيّة".ومن بين فروعها المعجم اللساني الحاسوبي الإلكتروني .

1/المقصود بحوسبة للمعجم اللساني :

يقول العلامة عبد الرحمن حاج صالح :" إن الدراسات والبحوث العلمية في اللسانيات الرّتابية (الحاسوبية) ازدهرت في الوطن العربي في هذه الآونة ،وتكاثر إلى حدّ ما الباحثون في هذا الميدان الذي تتلاقى فيه علوم الحاسوب و علوم اللسان ،وهو

ميدان علمي وتطبيقي واسع جدا كما هو معروف ،إذ يشمل التطبيقات الكثيرة كالترجمة الآلية والاصطلاح الآلي و الأخطاء المطبعية وتعليم اللغات بالحاسوب ..."[1]

تقوم اللّسانيات الحاسوبيّة من هذا المنطلق على جانبين مهمّين:[(2)]

أ- **الجانب النّظري:** يبحث في الإطار النّظريّ العميقِ الّذي به يمكننا أن نفترض كيف يعمل الدّماغ الإلكتروني لحلّ المشكلات اللّغويّة.

ب- **الجانب التطبيقيُّ:** يعنى بالنّاتج العلميّ لنمذجة الاستعمال الإنسانيّ للّغةِ، وإنتاجِ برامجَ ذاتِ معرفةٍ باللّغة الإنسانيّةِ.

ومن المعروف أنّ الجانب التطبيقيَّ هو الجنب المهمُّ في اللّغويّاتِ الحاسوبيّةِ، والّذي يتمثّل في تسخيرِ العقلِ البشريّ لحلّ القضايا اللّغويّةِ؛ وهنا يبرز الدّور الرّئيس، والأثر الفاعل لالتقاء اللّغويّين والحاسوبيين.

وتعد حوسبة المعجم من أهم مجلات علوم اللغة الحاسوبي .

حيث 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 الأولية لبنية الكلمة، ومباشرة تحديد سماتها المعجمية[3].

2/**أهمية الحاسوب في صناعة المعجم :**

تقوم حوسبة المعجم بخدمات كبيرة للبحث اللغوي ،وتتجلى أهمية الحاسوب في صناعة المعجم كما يقترح اللغوي "عبد الله أبو الهيف" فيما يلي [4]:

-التعرّف على الحروف والكلمات آليا .

-تخزين المادّة

-ترتيب المادة طبقا لنظام الحاسوب

-استرجاع المادة أو بعضها .

-استكمال أجزاء من المادّة أو من الشرح .

- تعديل بعض المعطيات .

- حذف بعض المعطيات .

-النقل المباشر إلى المطبعة .

-تجديد المعجمات بسهولة

-الحصول على أجزاء محدّدة داخل المادّة المخزونة لبحثها .

فهو بذلك يثمّن مشروع حوسبة المعجم إذ يقرّ بأنها تساعد على تسهيل معجمية الرصيد اللغوي العربي الثري في حافظات برمجية جاهزة للتسيير وفق الأغراض المعجمية المنشودة من حيث الإحصاء والوصف والتعدّد الدلالي والتوزّع اللغوي الصّوتي أو الصّرفي أو النحوي والمجالات الإبلاغية و الاتصالية[5] .

3/**المحلل المعجمي الآلي متعدد الأطوار:**

يقوم هذا المعالج حسب اللساني سمير شريف استيتية- بتفكيك الكلمة قيد التحليل عن طريق إزالته ما يتصل بها من سوابق (كالأدوات و الحروف)، ولواحق(كالضمائر المتصلة،وزوائد التصريف، وعلامات الإعراب)،كذلك يعمد إلى رد التعديلات الفونولوجية التي تم 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 :[(6)]

* الاحتمال الأول: والد + ين ⟸ مثنى مذكر.

* الاحتمال الثاني: 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.

و لتفادي الضياع في متاهات تحليل المبنيات كالضمائر و حروف الجر، وظرفي الزمان و المكان ، ثم توفير آلية ذكية يمكن أن توظف من خلال هذا المعالج لمقارنة جذع الكلمة المعجمية المفككة مع قائمة من الكلم المبني باللغة العربية .

ويلاحظ أن المعالج المعجمي الآلي متعدد الأطوار يمتاز بآلية عمل متقدمة من الناحيتين البرمجية البحتة والمعالجة المعجمية واللغوية، الأمر الذي جعله يتبوأ مكانة بارزة في ميدان البحث .

وقد حاول الدكتور عبد الرحمن حاج صالح ذلك من خلال المثال المولد للاسم (أو حده الإجرائي) توضيح ذلك في مخططه الإجرائي :

نقصد بالمعالجة الآلية للصرف العربي اعتمادا على الحاسوب بمختلف أنماطه البرمجية، وذلك بوضع نماذج Modèles لمعالجة الصرف آليا و قد ظهرت جملة من أنظمة هذا التحليل من أهمها : المحلل الصرفي متعدد الأطوار لشركة صخر العالمية [7] والمحلل الصرفي الآلي الذي هو نحت قيد المعالجة من طرف مخبر المعالجة الآلية للغات الطبيعية، بولاية تلمسان بالجزائر.

حيث يقوم الحاسوب بتوليد الأسماء، والأفعال المشتقة انطلاقا من الجذر المدخل عليه.

ويعتمد نظام اشتقاق الكلمة على معجم حاسوبي ضمن قاعدة معطيات وعلى القوانين الصرفية والنحوية لقواعد الاشتقاق.

وقد أنجز خبراء المعلوماتية في سوريا المعجم الحاسوبي ضمن قاعدة بيانات وعلى القوانين الصّرفية و النحوية لقواعد الاشتقاق ويحتوي على جميع الجذور المعجمية الثنائيـة والثلاثية والرباعية والخماسية وقد بلغ عددها في أحد الإحصاءات 11347جذرا توزعت كما يلي [8]:

115 : جذر ثنائي => جذور تراكيب لا اشتقاق منها (مثل أفعال) الأمر: قل، دع...)

7198 : جذر ثلاثي => هي أكثر الجذور خصوبة.

3739 : جذر رباعي => هي دون الثلاثية في الخصوبة.

295 : جذر خماسي=> أقل الجذور خصوبة.

واعتمدت هذه الإحصائيات على خمسة معاجم أصول هي :

جمهرة اللغة " بن دريد "، تهذيب اللغة " الأزهري "، المحكم " ابن سيده "، لسان العرب " ابن منظور "، القاموس المحيط " فيروز أبادي ".

كما يحتوي هذا المعجم الحاسوبي على جميع الأفعال الثلاثية والرباعية المجردة والمزيدة والتي بلغ عددها في الإحصائيات 32490 فعلا .

كما اشتمل تصريف الأفعال وحروف التعدية والصفات المشبهة ...[2]

و قد حاول الأعضاء وضع مصفوفة ثلاثية البعد يمثل بعدها

الأول: فاء الفعل

الثاني: عين الفعل

الثالث: لام الفعل

وذلك عن طريق وضع الأرقام لكل حرف لإسقاط الأسماء أو الأفعال في الميزان الصرفي الآلي بطريقة سريعة و دقيقة .

أما أعضاء شركة صخر المصرية فهم يركزون على:

1- المعالج صرف نحوي.
2- المعالج الاشتقاقي.

دون إغفال جانبي الإعراب والتشكيل.

أ- **<u>المعالج صرف نحوي:</u>**

يقوم هذا المعالج بتفكيك الكلمة وتجريدها من اللواصق واللواحق فيذهب بذلك إلى أصل الكلمة لتصبح صالحة للمعالجة مثل كلمة : والدين .[9]

فهو في هذه الكلمة يقترح وجود عدة احتمالات من بينها :

1- والد + يْن => مثنى المذكر.
2- والد + ين => ج مذكر السالم.
3- و + الدين=> من الفعل "دان".
4- وَ + الدِّين=> من الفعل "دان".

ب- **المعالج الاشتقاقي:**

حيث يقوم باستخلاص جذر الكلمة و الصيغة الصرفية التي قام بتفكيكها المعالج الصرف نحوي.

الكلمة	الجذر	الوزن	الاحتمال الصرفي
دين	دَيْن	فعْل	والدّين
دَيْن	دَان	فعَل	وَ الدَّين
وَالِد	وَلد	فاعِل	وَالدِين
وَالِد	وَلد	فاعِل	وَالِدَيْن

أما من ناحية الإعراب فالهدف من وراء هذا النظام تمييز الكلمة بناء على الحركة الإعرابية (العلامات : حركة التشكيل الظاهرة أو المحذوفة ...)[10]

وأما عن التشكيل فهو يقوم بإعادة تركيب عناصر الكلمة المفككة سابقا للتأكد من التحليل الصرفي لبنيتها ومطابقتها بالصيغة الصرفية المحتملة.

حيث يتم التحليل عن طريق الانتقال من الكلمة العربية إلى جذرها الأصلي أي أن الحاسوب يعالج الكلمات العربية المشكولة جزئيا أو كليا ، أو غير المشكولة ويحدد نوعها ، وميزانها الصرفي ، وسابقتها و لاحقتها ، وحالتها الإعرابية ودلالتها ... فإذا احتوت الكلمة المراد معالجتها على حروف غير مشكولة وضع الحاسوب الحركات الممكنة لها اعتمادا على إحصائيات تلائم الحركات مع الحروف تمهيدا لتحليلها، ومن المعلوم أن خلو الكلمة من الشكل يجعلها متعددة الأشكال ومن ثم المعاني، ما دامت مستقلة عن النص.

فكلمة "ورد" مثلا يمكن أن تكون لها الإمكانيات التالية:

ورد = وَرَدَ - وَرَّدَ - وُرِدَ... => أفعال.

= وَرْدٌ - وِرْدٌ ... => أسماء.

= و+ ردَّ - و+ رُدَّ ... => حرف عطف+أفعال.

= وَ+ رَدٌّ - و+ رَدٍّ ... => حرف عطف+أسماء.

والكلمة المشكولة إذا عولجت مستقلة عن سياق النص فلا يمكن شكلها (أي ضبطها بالحركات) لإمكانية اشتراكها في الاسمية والفعلية والحرفية، فمن أمثلة

الحالة الأولى كلمة "أحمد" فهي إسم في نحو ذهب أحمد إلى المدرسة، وفعل في نحو :أحمد لك حسن صنيعك.

ومن أمثلة الحالة الثانية كلمة أنَّ، فهي حرف في نحو: تعلمون أن السماء لا تمطر ذهبا، وفعل في نحو: أنَّ المريض من الألم.

وهذا يعني أن الحاسوب عليه أن يعالج الكلمة عن تحليلها على أنها اسم أو فعل أو حرف وأن يعطي جميع الإمكانيات المحتملة لها مع مراعاة الحالات التي تحدد نوعها (

فالكلمة المنونة لا تكون إلا اسما) :و بعد ذلك يقوم الحاسوب باختيار الإمكانية المناسبة التي تتوافق مع اختيار النص.

فيفترض الحاسوب أن الكلمة المدخلة إليه هي فعل، فيقوم بتحديد كل من سابقته ولاحقته وصيغته وبنائه للمعلوم أو المجهول ، وتد تجرده أو زيادته و وزنه و أصله المشتق منه، وحالته الإعرابية والضمير المسند إليه و دلالته أي أن يعطي وصفا كاملا عن حالة الفعل الصرفية مستقلة عن النص[11].

كما إذا كانت الكلمة المدخلة إليه إسما، فيحدد كلا من سابقته ولاحقته ووزنه وأصله المشتق منه وحالته الإعرابية، ونوعه من حيث الاشتقاق والتذكير والتأنيث والإفراد والتثنية والجمع والنسبة والتصغير ... .الخ[12]

نموذج حاسوبي

صممت لتوليد أسماء الذات آليا، وهي إحدى الواجهات Interfaceواجهة تخاطبية التخاطبية التي صممت لتوليد مكونات المعجم الآلي للغة العربية، صممت بلغة بيسك. 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 وهلم جرا[13].

نموذج عن صناعة المعجم اللساني الحاسوبي:

وتعتمد في اللسانيات الحاسوبية كإجراء رياضياتي لمختلف المستويات، فمثلاً: إذا قمنا بتطبيق قوانين "المزهر"[13] على الكلمات التي تحوي ثلاث أحرف(كنموذج) في معرفة المعرّب من الفصيح، فوجدناه يحتوي على 2412 كلمة فصيحة بالتّقريب سواء كانت الجيم (الفاصلة بين المعرّب والفصيح في العربيّة) في أوّل الكلمة أو وسطها أو آخرها.

قمنا بالخطوات التالية:

1- معرفة رتبة الجيم تطبيق نظام التقاليب

ج ق-	ق ج –	– ق ج	6 تقاليب
ج – ق	ق – ج	- ج ق	

2- السّحب يتمّ على الكلمات التي تحوي 3 أحرف:

أ ن= 6^3= 6×6×6= 216 كلمة.

216 كلمة معرّبة تحوي على (ج، ق)

وهكذا مع (ج، ص)- (ج، ت)، (ج، ط)

فنحصل –مقاربة رياضية- على 216×4= 864 كلمة، وبذلك نجد:

3276- 864= 2412 كلمة.

2412 كلمة –بالتّقريب-فصيحة تحوي حرف الجيم كاملة لكلمات ثلاثية الأحرف، سواء كانت الجيم في أوّل الكلمة أو وسطها أو آخرها[14]

في هذه الحالة جمعنا بين الإحصاء وعملية الاحتمالات للحصول على الرقم الدّقيق للمعجم.

و لا ينحصر هذا فقط على الدّراسة المعجميّة بل حتّى الصّوتية والصّرفية والنّحوية والدّلالية.

مشاكل المعالجة الآلية للكلمة المعجمية :

يعترض المعجم اللساني الحاسوبي مجموعة من المشاكل يمكن حصرها في :

1- غياب الشكل (الحركات) أحيانا.
2- غياب كتابة الألف أحيانا.
3- عدم التفريق بين الياء والألف المقصورة.
4- تعدد المعاني => تعدد الاحتمالات مثل: كلمة عين، فهي تحمل معاني كثيرة

منها عين الجاسوس، ينبوع الماء – طليعة الجيش – ذات الشيء-معجم العين.

5- اللبس الصرفي: فمن الكلمات ما يشترك في الاسمية والفعلية.

مثل:كاتبوه ، ومنها ما يشترك في اسم الفاعل والمفعول مثل المختار ومنها ما يشترك في صيغة الأمر و الماضي ، مثل: اصطادوا... .الخ.

و الحاسوب في هذا كله قد لا يكتشف اللبس إلا في مرحلة متأخرة ، مثل كلمة " ليمون" أو " زيتون " فقد يصنفها ضمن الأسماء ، في جمع المذكر السالم و ليس المفردة.

خاتمة :

من خلال هذا البحث توصّلنا إلى :

1/إن المعجم اللغوي العربي يتصف بحجم محدود من الجذور والصيغ الصرفية مع تعدد مفرداته الخصبة نتيجة الإنتاجية الصرفية والمعجمية العالية، الأمر الذي يجعله موردا طيبا للمعالجة الآلية للمعجم.

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

3/خلو الساحة اللغوية من البرمجة الحاسوبية للمعجم نظرا لقلة التخصصات في ميادين المعالجة الآلية للغة والحوسبة اللغوية .

4/المعجم اللساني الحاسوبي يحتاج إلى عمل مؤسساتي وتضافر الجهود في ذلك وتوفير البرامج

- [1]-بحوث ودراسات في اللسانيات العربية –عبد الرحمن حاج صالح –موفم للنشر –الجزائر –دط –دت –ج:01- ص :230

[2] ينظر: توظيف اللّسانيات الحاسوبية في خدمة الدّراسات اللّغوية العربية"جهود ونتائج"، عبد الرحمن بن حسن العارف، ص: 03

3-ينظر: ينظر اللغة العربية و الحاسوب – علي نبيل – بيروت – دار التعريب -ط : 01 – 1988 – ص : 308

2-ينظر: حوسبة المعجم العربي والقضايا المعاصرة-أحمد هاشم السامرائي –مجلة سرّ من رأى – جامعو سامراء –كلية التربية-المجلد:09-العدد34-2013-ص:07 وما بعدها .

- ينظر: حوسبة المعجم العربي –الواقع والآفاق –جيلالي بن يشو –مقال محمّل من الانترنيت على الرابط : [5]

//www.voceofrabic.net/ar/articles/1901.:http

[6].ينظر: اللسانيات-المجال والوظيفة و المنهج –سمير شريف استيتية-الأردن-عمان-عالم الكتب الحديث ،ط2005،1،ص548.

[7]ينظر اللغة العربية و الحاسوب – علي نبيل – بيروت – دار التعريب -ط : 01 – 1988 – ص : 308 .

8 -ينظر: مستقبل اللغة العربية –حوسبة المعجم العربي ومشكلاته اللغوية والتقنية أنموذجا –عبد الله أبو هيف –مجلة التراث العربي –اتحاد الكتاب العرب –دمشق –سوريا –العدد 93/94-2004- ص :114 وما بعدها .

[9]سبل تطوير محلل الصرف الآلي – حسنى مظفر الرزو مقال صدر في الإنترنت : 2008/05/12 على موقع Google – ص : 04.

[10]ينظر: اللغة العربية و الحاسوب – علي نبيل – ص : 312.

11- لقاء مع البروفسور سيدي محمد غيثري حول حوسبة المعجم وجهود المخبر في ذلك –مدير مخبر المعالجة الآلية للغة العربي –جامعة أبي بكر بلقايد –تلمسان –الجزائر –بتاريخ 2017/12/12م-

-المرجع نفسه . [12]

[13] -ينظر: محاضرة في موضوع: اللغة العربية والحاسوب -(قراءة سريعة في الهندسة اللسانية العربية)أومقاربة في محاكاة الدماغ العربي لغويا- أ.د/ محمد محمد الحناش- جامعة الإمارات العربية المتحدة- كلية العلوم الإنسانية والاجتماعية -قسم اللغة العربية وآدابها-أكتوبر 2002-ص :23

[13] ينظر: المزهر في علوم اللغة وأنواعها، السيوطي، شرح وتعليق محمد جاد المولى بك، محمد أبو الفضل إبراهيم، علي يد محمد البجاوي، المكتبة العصرية، صيدا بيروت، 1986م، دط، ج:01، من ص: 271 إلى 273.

[14] ينظر: صوت الجيم في المنطوق التلمساني، دراسة صوتية مخبرية، راضية بن عريبة رسالة ليسانس- جامعة أبو بكر بلقايد، تلمسان، الجزائر، 2004/2003، ص: 52- 53.