

إسهامات اللسانيات الحاسوبية في تألية قواعد اللغة العربية

Efforts Of Computational Linguistics in the machinism of Arabic Grammar

الأستاذة: نعيمة عيوش
n.ayouche@univ-dbkm.dz

الدكتور: العربي بو عمران بوعلام
boualemlarbi16@hotmail.fr

جامعة الجيلالي بونعامة - خميس مليانة- الجزائر

تاريخ النشر: 2019/12/31

تاريخ القبول: 2019/11/11

تاريخ الإرسال: 2018/09/24

الملخص:

إن عصرنا الحالي يشهد تطور هائل في العلوم نتيجة ظهور التقانات والتكنولوجيات الحديثة، التي استطاعت أن تجتاح مختلف المجالات المعرفية من بينها الدراسات اللغوية، إذ صار لزاما معالجة هذه اللغة آليا وإدخالها الحاسوب فتمخض عن ذلك ظهور علم جديد هو اللسانيات الحاسوبية التي أخذت على عاتقها استحداث برامج آلية تخدم اللغة، وقد نهضت الدراسات اللغوية على هذا الأساس فحوضا يضاهي التطور الحديث في سائر أنماط النشاط البشري و أثبتت الأعمال المقدمة في هذا الجانب أنها يمكن أن تنبعث بمزيد من الضبط والدقة والتصحيح والتقدم إلى ما هو أفضل وأكثر فائدة لدارسها، باستعمال تقنية الحاسوب،

ظهرت العديد من المحاولات التي ما زالت تتوالى من قبل الدارسين ومن بين هذه البرامج نذكر تلك التي تتعلق بتعليم قواعد اللغة نحوها وصرفها، إذ كان لابد من استخدام البرامج الحاسوبية في العملية التعليمية وإدراجها ضمن المناهج التربوية مما يسمح بتطوير طرق التدريس والانتقال من الطريقة التقليدية التي أصبح ينفر منها المتعلم إلى طريقة مثلى تتلاءم مع التطورات الحاصلة في هذا العصر.

المهدف من هذه الدراسة هو اقتراح بعض الحلول من أجل بناء برامج حاسوبية تستوعب النظام الكلي للغة العربية. دور اللسانيات الحاسوبية في تألية النظام اللغوي، لذا حاولنا تسليط الضوء حول بعض البرامج الحاسوبية التي تدخل ضمن مجال اللسانيات الحاسوبية التعليمية مثل: "برنامج الخليل الصربي - برنامج اعراب _ برنامج التشكيل الآلي".

الكلمات المفتاحية: اللسانيات، الحاسوبية، التعليمية، البرامج، القواعد، اللغة العربية.

Summary: The current era is witnessing a tremendous development in science as a result of the emergence of modern technologies and technologies, which was able to sweep the various fields of knowledge, including language studies, as it became necessary to process this language automatically and the introduction of computer, resulting in the emergence of a new science is computer linguistics, which took on the development of programs Linguistic studies on this basis have risen in the same way as modern development in other types of human activity and the works presented in this aspect have proved that they can be emitted with more control and accuracy, correction and progress to the best and most useful for its students. Computer technology money,

Keywords: Linguistics, Computational, Educational, Software, Grammar, Arabic.

مقدمة

لم تبقى اللغة العربية حبيسة الدراسات اللسانية، وإنما أحدثت قفزة نوعية في مجال اللسانيات الحاسوبية، التي حاولت ربط كل ما هو لغوي بالحاسوب من أجل مواكبة حركية التطور العلمي والتكنولوجي، التي شملت مختلف مجالات الحياة وظهور إمكانيات جديدة لتناقل المعارف بين الأمم والشعوب وظهور برامج وآليات وطرائق مستحدثة في مجال معالجة البيانات.

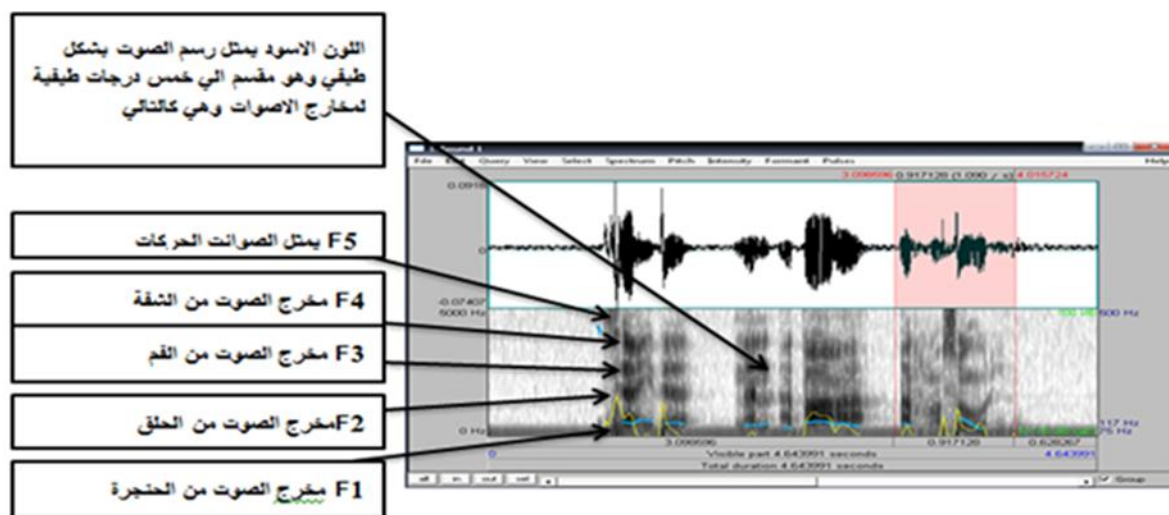
إن وعي الدارسين والباحثين بضرورة حوسبة اللغة العربية ومواكبة الرقمنة، أثمر بظهور العديد من الدراسات الحاسوبية العربية، رغم العديد من المعوقات التي في مجملها تعود الى طبيعة اللغة العربية، التي تقتضي منها دقياً في التعامل مع نظامها اللغوي.

لذا يتعين على المختصين في اللسانيات الحاسوبية، إتباع المنهج الوصفي من أجل تألية قواعد اللغة العربية واستقراء النظام الكلي للغة العربية، الذي نص عليه النحاة والدارسين اللغويين قديماً وحديثاً، إذ أن منتهى الغاية من هذا المجال هو تهيئة الحاسوب لتحصيله كفاية لغوية، تشبه ما يكون بالإنسان حين يستقبل اللغة ويدركها ويفهمها ثم يعيد إنتاجها، ولا يتأتى ذلك إلا باستدخال قواعد اللغة العربية بنظامها الصوتي، والصرفي، والتركيب، والدلالي، لذا سنحاول تسليط الضوء حول بعض القضايا اللغوية المهمة، التي لا بد من مراعاتها أثناء حوسبة هذه القواعد بدءاً بالنظام الصوتي الفونيتيكي، والفونولوجي، ثم النظام الصرفي (المورفولوجي)، يليه النظام النحوي التركيبي، وأخيراً النظام الدلالي.

1- تألية النظام الصوتي اللغوي:

تعتمد اللغة العربية في نظامها الصوتي على أربعة وثلاثين فونيماً، ثمانية وعشرون صامتاً تتمثل في حروف الأبجدية العربية، وستة صوائت تمثلها الحركات الثلاث القصيرة الفتحة، والكسرة، والضممة، والحركات الثلاث الطويلة المد، لذا تعنى اللسانيات الحاسوبية بتمثيل هذه الأصوات طيفياً، مع مراعاة العديد من القضايا الصوتية كالزمن، والشدة، والتردد، كل ذلك يمثل رياضياتياً وبأرقام دقيقة جداً باستطاعتها تمييز الأصوات عن بعضها البعض، إذ أن الإشارة الصوتية يمكن أن تتمثل بالاعتماد على معالم متنوعة مثل التمثيل باستخدام معلمات المجال الزمني للإشارة، أو المعلمات الطيفية ليتم استخراج خصائص الأصوات المطلوبة من قبل النظام المصمم، إن اختيار نوع التمثيل للأصوات يؤدي دوراً كبيراً ومهماً في إظهار خصائص الأصوات¹.

هناك العديد من البرامج الحاسوبية التي تعمل على معالجة الأصوات اللغوية بدقة متناهية نذكر على سبيل المثال برنامج (PRAAT).



الرسم البياني frequency يمثل التردد F1 من 0 إلى 1000 هرتز و F2 من 1000 إلى 2000 هرتز و F3 من 2000 إلى 3000 هرتز و F4 من 3000 إلى 4000 هرتز و F5 من 4000 إلى 5000 هرتز.

الشكل (1): التقطيع الصوتي بواسطة برنامج (PRAAT) لتحديد مخارج الاصوات

ما يزيد الأمر تعقيدا في محاولة تألية النظام الصوتي هي القواعد الفونولوجية، حيث أن الأصوات تكون بشكل إصاقي متابعي، وبالتالي تتأثر خصائصها الصوتية بما يسبقها وبما يليها، لذا لا يحتاج الحاسوب إلى إدراك الفونيمات لوحدها، وإنما لا بد من أن يدرك جملة من التغيرات التي تطرأ على الفونيمات، ولنا أن نعطي بعض الأمثلة على ذلك: الحروف المضعفة تزيد مدتها الزمنية النطقية، تغير نطق بعض الفونيمات نظرا لتداخل مخرجها أو صفتها مع فونيمات أخرى مجاورة لها مثلا كلمة (أردت) و(سبت)، عدم نطق اللام في (ال) التعريفية إذا تلتها إحدى الحروف الشمسية، بالإضافة إلى أن هناك فونيمات تنطق ولا تكتب كالنتين والمد، وبعض الفونيمات التي تكتب ولا تنطق كهزمة الوصل، والتاء المربوطة عند الوقف، والألف الفارقة بعد واو الجماعة.

وبالتالي إذا أردنا إيجاد برنامج حاسوبي يتعرف على المنطوق ويحوله إلى مكتوب، أو العكس يتعرف على المكتوب ويقرؤه، يتطلب الإمام بمختلف التغيرات التي تطرأ على الفونيمات مع ضبطها بالتشكيل التام، نذكر من بين هذه البرامج .Balabolka- Read Speaker - DSpeech- Easy Speech2 Text

2- تألية النظام الصرفي (المورفولوجي) اللغوي:

نقصد بتألية القواعد الصرفية للغة العربية تمكين الحاسوب من التعامل مع الوحدات الصرفية، من خلال القدرة على صياغة الأبنية اللغوية السليمة بخطوات اجرائية منتظمة، مثال على ذلك مقدراته على توليد الأبنية الزمنية، كالفعل الماضي بنوعيه المجرد والمزيد، والفعل المضارع، والأمر، لأي مادة معجمية مدخلة، وتقديم توضيحا شاملا لما يطرأ على الجذر من تغيير لدى البناء للمعلوم ولل مجهول، ثم ما يعتريها من زيادات وتغيير في حركات المبنى لدى صوغ الأفعال المزيدة المبنية للمعلوم ولل مجهول، وذكر ما يعتريها من تغيير في الحركات، أو الاعتلال، والقلب، أو الابدال، في بعض الوحدات المعجمية، بسبب تتابع الأصوات مع مراعاة التغيير الذي يطرأ على كل نوع من الأفعال، من حيث الصحة

والاعتلال، إذ يتباين الفعل السالم بنوعيه المضعف والمهموز، والمعتل بأنواعه المثال، فالأجوف، فالناقص، فاللغيف، وبالتالي سيكون من المفيد جدا أن يدعم كل برنامج بهذه القواعد الصرفية، حتى يغدو قادرا على توليد كل الصيغ الصحيحة للوحدات المعجمية.

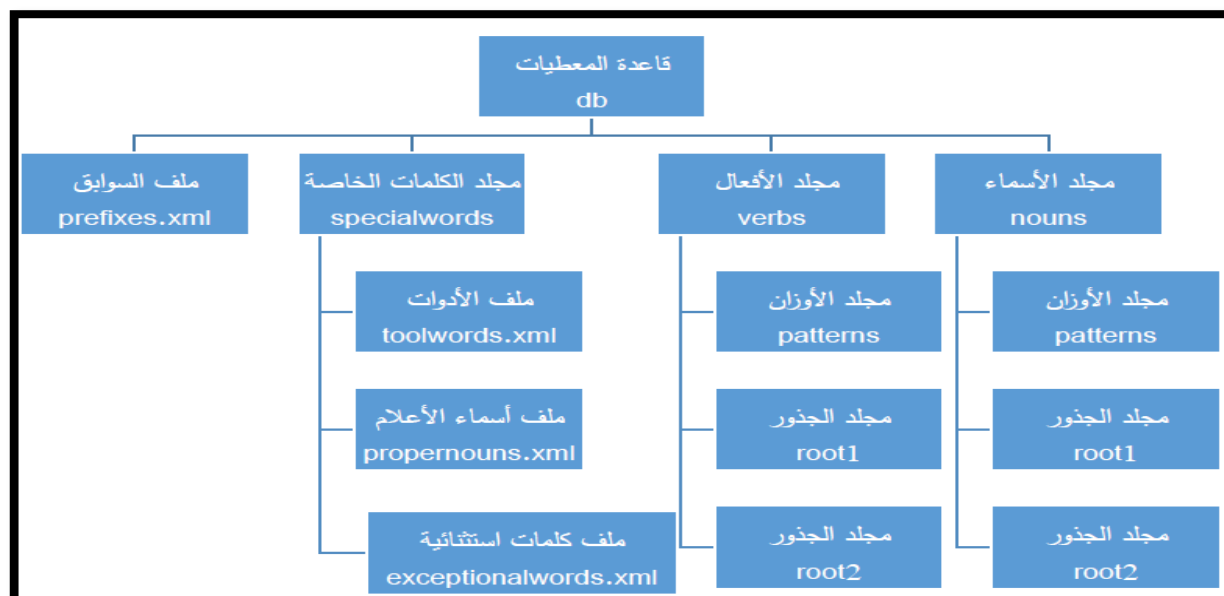
كما يجب أن يتمكن الحاسوب من تعيين المعاني الصرفية، وتفكيك التراكيب اللغوية إلى وحداتها الصرفية الصغرى حين ترد في سياق الجملة أو في سياق النص، بالاعتماد على محددات وأدلة نحوية وتركيبية، يميز بها عما قد يلتبس به من المباني الشبيهة خاصة إذا كانت غير مشكولة، فقد يلتبس عليه التفرقة بين الفعل الماضي والأمر (استخرج)، وبالتالي يلجأ الحاسوب إلى الأدلة النحوية التي تسبق الفعل مثل (قد - حتى - ما) أو اتصاله ببعض الضمائر.

وبالتالي يقصد بالتحليل الصرفي الآلي للكلمة في اللغة العربية، ربط كلمات النص بالعناصر الصرفية الأولية التي تدخل في تكوينها، وكذلك بالقيم النحوية دون اعتبار موقعها، فيتم في التحليل الانتقال من الكلمة إلى جذرها الاصلي، أي أن الحاسوب يعالج الكلمات العربية المشكولة جزئيا، أو كليا، أو غير المشكولة، فيصف ما يطرأ عليها من تغيير، زيادة، أو نقصان، أو إعلاالا أو إبدالا، أو إدغاما، أو قلبا، حيث يحدد نوعها، وميزاتها الصرفي، وسابقتها ولاحققتها وحالتها الاعرابية، ودلالاتها²، هناك العديد من البرامج الحاسوبية التي تعنى بالقواعد الصرفية من بينها برنامج التحليل الصرفي.

تعريف برنامج الخليل الصرفي:

أنجز نظام الاشتقاق والتصريف، برنامج الخليل الصرفي، تحت مظلة المنظمة العربية للتربية والثقافة والعلوم خلال العام 2010، بالتعاون مع جامعة محمد الأول بالمملكة المغربية، ومعهد بحوث الحاسب في مدينة الملك عبد العزيز للعلوم والتقنية بالمملكة العربية السعودية، يرمي هذا النظام إلى تمكين الحاسوب من اشتقاق (توليد) الأفعال والأسماء المشتقة والمصادر، وتصريفها، انطلاقا من جذورها الثلاثية أو الرباعية، بالاعتماد على قوانين النحو والصرف وعلى المعجم الحاسوبي، ومن أهم مزاياه أنه مفتوح المصدر، وبرنامج، وخوارزمياته وقواعد معطياته، وملفاته، جميعها متاحة للإضافة أو التغيير وبذلك يستطيع العاملون في هذا المجال المعالجة الآلية اللغة العربية، أو في مجال المعلوماتية تطوير البرنامج وتعديل قواعد المعطيات إضافة وحذف وتحديثا³.

ويعد هذا البرنامج من أبرز الاعمال شركة صخر إذ أنه يقوم بالمعالجة العميقة للكلمة العربية، حيث نستطيع من خلاله التعرف على أشكال جذر الكلمة، وتجريدها من حروف الزيادة مع توضيح حالتها الاعرابية. يعتمد برنامج الخليل الصرفي على قاعدة معطيات تم ترتيبها على شكل مجلدات وملفات، بحيث أنه مزود بكم هائل من الملفات التي تحوي كل القواعد الصرفية، والخطاطة التالية تلخص أهم مراحل إعداد برنامج الخليل الصرفي:



الشكل (2): قاعدة معطيات برنامج الخليل الصرفي

يوضح الشكل (2) قاعدة المعطيات التي ينبنى عليها برنامج الخليل الصرفي إذ تتكون من مجلدات وملفات يحتوي كل مجلد أو ملف على مبحث صرفي كالجذور أو الأوزان أو الأسماء والأفعال، الغرض من البرنامج هو تحليل الكلية لتحديد المعارف الصرفية المتعلقة بها كجذورها وسابقتها ولاحقتها ونوعها ووزنها كما يوضح الشكل التالي:

الكلمة المشكولة	الساكن	نوع الكلمة	وزن الكلمة	جزء الكلمة	الحالة الإعرابية	اللاحق
استقبل	#	فعل ماض مبني للمعلوم	استقبل	قبل	ثلاثي مزيد مسند إلى الغائب (هو) لازم	#
استقبل	#	فعل أمر	استقبل	قبل	ثلاثي مزيد مسند إلى المخاطب أنت لازم	#
استقبل	#	فعل ماض مبني للمجهول	استقبل	قبل	ثلاثي مزيد مسند إلى الغائب (هو) لازم	#

الشكل (3): تحليل الفعل استقبل من قبل برنامج الخليل الصرفي

يبين الشكل تعرف البرنامج على الكلمة المدخلة (استقبل) من دون تشكيل، إذ حدد نوع الكلمة على أنها فعل إما أن تكون فعل ماضي مبني للمعلوم، أو للمجهول، أو فعل أمر، كما أعطى مجموعة من المعلومات كوزن الكلمة وجذورها وحالتها الإعرابية.

يعمل برنامج الخليل الصرفي على تحليل الكلمة، وإظهار جميع الوجوه الممكنة للكلمة المعالجة، إذ يستوعب جميع الأفعال الثلاثية والرابعة، المجردة والمزيدة، وأما الاسم فيشتمل على جميع الأسماء والمصادر والصفات والظروف، وأما الحرف فيتضمن جميع حروف الجر والنصب والجزم والعطف،

وبسبب غياب الشكل يمكن أن تشترك وجوه التحليل الصرفي للكلمة في:

الاسمية والفعلية؛ نحو: ورد = وَرَدَ (اسم)، وَرَدَ (فعل).

أو الاسمية والحرفية؛ نحو: رب = رَبُّ (اسم)، رَبٌّ (حرف).
 أو الفعلية والحرفية؛ نحو: عدا = عَدَا (من الفعل: عَدَا يَعْدُو)، عَدَا (حرف جرّ).
 أو الفعلية والاسمية والحرفية؛ نحو: بل = بَلَّ (فعل)، بَلَّ (مصدر)، بَلَّ (حرف).
 لذا فإن نظام التحليل الصرفي يفترض الكلمة التي يعالجها إما فعلٌ واسمٌ وحرف، على هذا فإن خرج نظام التحليل الصرفي يتضمن جميع الوجوه الممكنة للكلمة المعالجة.
 العمل الجديد في هذا البرنامج أنه يبين الحالة الإعرابية للكلمة المدخلة فإذا كانت فعل يبين البرنامج إذا كان لازماً أو متعدداً، أما إذا كانت الكلمة اسماً، أعطى المحلل النتائج التالية:

- 1- سابقته (و، أل، ب).
 - 2- لاحقته (ضمير الجر المتصل).
 - 3- وزنه (إفعال، فَعِيل، تَفَعَّل).
 - 4- نوع الاسم (مصدر، صفة، ظرف، اسم علم).
 - 5- جذره.
 - 6- اسم جامد أو اسم مشتق.
 - 7- نوعه من جهة التصرف (ممنوع من الصرف، مصروف).
 - 8- نوعه من جهة التذكير والتأنيث (مذكر، مؤنث).
 - 9- مثنى أو جمع مذكر سالم أو جمع تكسير.
 - 10- نوعه من جهة النسبة.
 - 11- نوعه من جهة التصغير.
- تظهر النتائج من خلال إدخالنا لكلمة مكتبة من خلال الشكل التالي:

اللاحق	الحالة الإعرابية	جذر الكلمة	وزن الكلمة	نوع الكلمة	السابق	الكلمة المستكولة
#	مفرد مؤنث منصوب نكرة	كتب	مَفْعَلَةٌ	اسم زمان أو مكان	#	مَكْتَبَةٌ
#	مفرد مؤنث مرفوع نكرة	كتب	مَفْعَلَةٌ	اسم زمان أو مكان	#	مَكْتَبَةٌ
#	مفرد مؤنث مجرور نكرة	كتب	مَفْعَلَةٌ	اسم زمان أو مكان	#	مَكْتَبَةٌ
#	مفرد مؤنث منصوب في حالة الإضافة	كتب	مَفْعَلَةٌ	اسم زمان أو مكان	#	مَكْتَبَةٌ
#	مفرد مؤنث مرفوع في حالة الإضافة	كتب	مَفْعَلَةٌ	اسم زمان أو مكان	#	مَكْتَبَةٌ
#	مفرد مؤنث مجرور في حالة الإضافة	كتب	مَفْعَلَةٌ	اسم زمان أو مكان	#	مَكْتَبَةٌ
#	مفرد مؤنث منصوب نكرة	كتب	مَفْعَلَةٌ	اسم مفعول	#	مَكْتَبَةٌ
#	مفرد مؤنث مرفوع نكرة	كتب	مَفْعَلَةٌ	اسم مفعول	#	مَكْتَبَةٌ
#	مفرد مؤنث مجرور نكرة	كتب	مَفْعَلَةٌ	اسم مفعول	#	مَكْتَبَةٌ
#	مفرد مؤنث منصوب في حالة الإضافة	كتب	مَفْعَلَةٌ	اسم مفعول	#	مَكْتَبَةٌ

الشكل (4): التحليل الصرفي للاسم (مكتبة)

3- تألية القواعد النحوية (التركيبية) اللغوية:

كلنا نعلم أن القواعد النحوية هي دقيقة جدا كونها عملية ذهنية، وقصد إدخالها الحاسوب لابد من تمكين الحاسوب من تعيين حركات اواخر الكلم، إذ لابد من تصنيف الكلمات العربية إلى مبنية ومعربة، لأن المبنى لا تتغير هيئته وبالتالي يحتمل طريقة واحدة لقراءتها، أما المعرب فنركز على الاسم المعرب، والفعل المعرب، نستثني الحروف والفعل الماضي وفعل الأمر والأسماء المبنية، كونها تلزم بنية وحيدة في جميع أحوالها ومواقعها السياقية، كما نستثني ما يجيء في باب المعرب مما لا تبدل حركة آخره كالمثنى، وجمع المذكر السالم، والأسماء الخمسة، والفعل المضارع المجزوم، معتل الآخر، نركز في هذه المعربات على التشكيل.

لذا يتعين التركيز على الفعل المضارع وضبط حركات أواخره بالضم والفتح والسكون، إذ يكون الفتح بعد أدوات النصب، ويكون السكون بعد أدوات الجزم، مع الأخذ بعين الاعتبار الأفعال معتلة الآخر لأن الحركة لا تظهر في أواخرها.

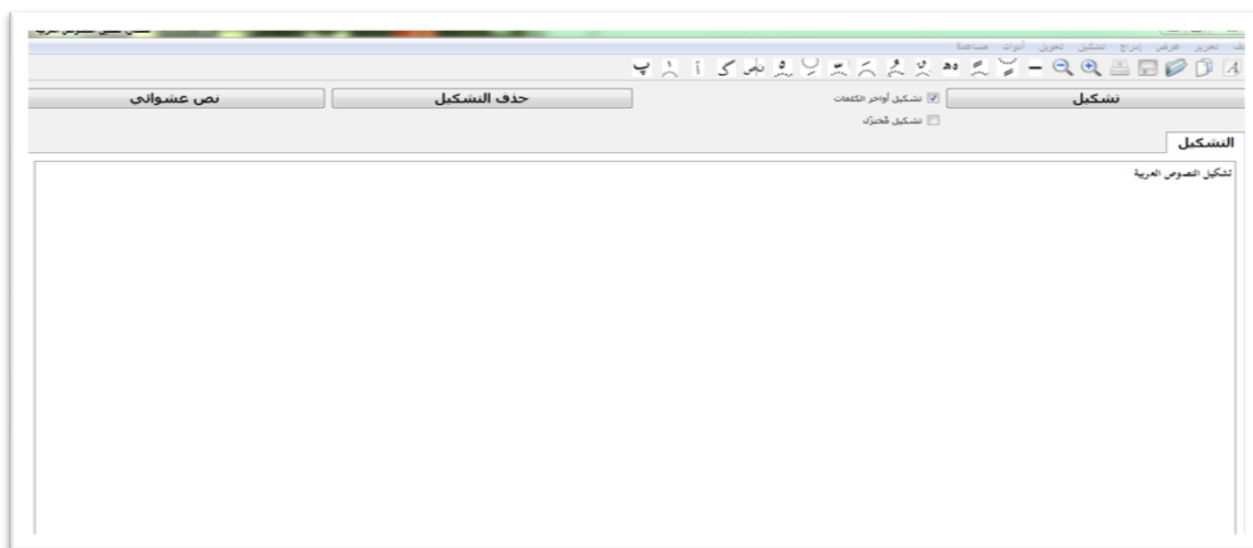
كما نركز أيضا على الأسماء التي يتغير آخرها كالاسم المفرد وجمع المؤنث السالم والممنوع من الصرف، لأن جميعها تحتمل النصب، كما يمكن الاعتماد على بعض المحددات في التعرف على بعض الحالات كالجر والرفع، إذ أن الجر يكون بعد حروف الجر وبعد المضاف، أما الرفع فيكون في المبتدأ والخبر والفاعل ونائبه واسم كان وأخواتها وخبر إن وأخواتها. وبالتالي لابد من تزويد الحاسوب بقاعدة معطيات تحوي كل القواعد النحوية بشكل منظم ومصنف بدقة، تكون على شكل أبواب (باب الأفعال، باب الأسماء المنصوبة والمرفوعة والمجرورة، وباب التوابع كالعطف والبدل والتوكيد والنعت)، وبالتالي تألية القواعد النحوية يساعد في اختصار الوقت وذلك لسرعته الفائقة في المعالجة، وكذا قدرته العالية من حيث سعة التخزين، إذ بإمكان الحاسوب تشكيل نص في ثوان معدودة وبدقة متناهية.

دور البرامج الحاسوبية في تألية القواعد النحوية:

أ- برنامج المشكل الآلي:

تحتاج أغلبية البرامج الحاسوبية التي تعالج اللغة العربية إلى ضرورة تشكيل المدخلات سواء كانت حرفا، أو كلمة، أو جملة، أو نصا، كون أن الحركات في اللغة العربية تلعب دور جد هام في تحديد الدلالة أو الوظيفة اللغوية المنوطة بها، وبالتالي فإن التشكيل اليدوي يكون مرهقا جدا ويستهلك الكثير من الوقت، خاصة إذا كنا نتعامل مع نصوص كبيرة الحجم، لذلك سيكون من المفيد ونحن في عصر التقنية الاستعانة بمنتجاتها لتسهيل العديد من أمور حياتنا اليومية، والتوصل إلى برامج معلوماتية وخوارزميات تنوب عنا في تشكيل النصوص المكتوبة باللغة العربية، قد لا تبلغ منتجاتها الكمال، لكنها في المقابل ستقتصد الكثير من الوقت والجهد، خصوصا إذا استحضرننا عدم كتابة الحركات على معظم لوحات المفاتيح، وعدم تمرن الكثير على استخدامها بسلاسة.

وهناك برامج كثيرة لتشكيل النصوص العربية، ومنها برنامج "مشكال" الذي يعتبر من أفضل المواقع في مجال تشكيل النصوص العربية، حيث يتيح لمستخدميه إمكانية تصحيح التشكيل المقترح للنص، وذلك عبر النقر على الكلمة المشكّلة، وتحديد التشكيل الصحيح من الاختيارات المتاحة.



الشكل (5): واجهة برنامج مشكال النصوص العربية

2- برنامج اعراب الجمل (بالعربي):

يعد هذا البرنامج من أفضل التطبيقات لإعراب الجمل، حيث يقوم بتقسيم الجملة إلى كلمات منفردة، وتقديم إعراب مفصل لها، إلا أن هذا البرنامج ما يزال ناقصا بعض الشيء من حيث استيعابه لكل القواعد النحوية.



الشكل (6): واجهة برنامج اعراب الجمل العربية (بالعربي)

3- تألية النظام الدلالي اللغوي:

يعد هذا الجانب من أعسر المباحث على الإطلاق، وذلك لتعلق دلالات الألفاظ بالفهم، أي تدخل العقل البشري، وهو أمر شبه بعيد المنال عن الحاسوب، ذلك أن الدلالة قد تكون معجمية أو صرفية أو نحوية، أو مجازية

أو إيجابية، وإدراكها يعود إلى فهم مختلف العلائق الدلالية من خلال السياقات التي ترد فيها، أو مختلف المعاني التي قد تحيل إليها وبالتالي لا يمكن ضبط هذا النظام بقواعد محددة.

لكن لا يمكن أن نقف عند هذا الحد، وننظر إليه كعائق أمام تألية هذا النظام الذي يرتبط بكل الأنظمة الأخرى (الصوتي، الصرفي، النحوي)، وإنما لا بد من السعي من أجل حوسبة ورقمنة هذا النظام المتشعب، وأفضل السبل لذلك هو بناء قاعدة بيانات ضخمة أو بنك للنصوص العربية، نركز فيها على مختلف الاستعمالات اللغوية في سياقاتها متعددة، ثم نستعين ببرامج للمعالجة الآلية ومحاولة تصنيفها وفق الدلالات المتواضع عليها، حالياً تهتم المعاجم الإلكترونية برصد كلي للألفاظ العربية مع تحديد مختلف استعمالاتها الدلالية، وذلك لأنها تأخذ بعين الاعتبار السياقات اللغوية المختلفة التي قد ترد فيها الوحدات المعجمية كما يمكن الافادة من المعاجم المتخصصة كمعاجم فقه اللغة، ومعاجم المعاني ومعاجم البلاغة، ومعاجم المصطلحات.⁴

صعوبات حوسبة قواعد اللغة العربية:

إن محاولة قبولية اللغة العربية في الحاسوب من أهم المشاكل التي تعترض طريق اللسانيات الحاسوبية، بدء بمحاولة إدخالها البرامج الحاسوبية التي تعتمد على لغة الجافا وحصر الأوزان العربية حصراً دقيقاً وتحليلها، وهو ما يمكن من وضع رموز رياضية لها في الحاسوب ثم يتم بعد ذلك تطوير برامج آلية يمكنها استيعاب القواعد النحوية و الصرفية العربية.⁵

لعل المعوق الأساسي هو أن معظم البرامج الآلية تبنى على خوارزميات هي في الأصل غربية المنشأ، بالإضافة إلى ما تتميز به اللغة العربية عن بقية اللغات بأنها تكتب وتقرأ من اليمين إلى اليسار، كما أن حروفها تكتب بأشكال مختلفة تبعاً لموقعها والحروف المجاورة لها، وتختلف طريقة نطق الحرف وبالتالي معنى الكلمة وموقعها الإعرابي بناءً على حركة التشكيل الموجودة عليه، بالإضافة إلى أن العربية لغة اشتقاقية وليست إصاقية، حيث يعد نظامها الصرفي من أكثر النظم الصرفية تقدماً، فهو مبني على تصريف الجذور وفقاً لمجموعة محددة من الأوزان للحصول على كلمات ذات دلالات مختلفة من نفس الجذر بالإضافة إلى غنى المعجم العربي مقارنة باللغات الأجنبية فالمعجم الإلكتروني الفرنسي DELAF يحتوي في إصداره الحالي 700000 مدخل معجمي أما المعجم الإلكتروني العربي DIINAR يحتوي على 7774938 مدخل معجمي، هذه الخصائص جعلت الباحثين في علوم اللسانيات الحاسوبية يعتبرونها لغة متميزة تنفرد بخصائص لا توجد في لغات أخرى ومن ثم يصعب وصفها اعتماداً على النماذج الغربية للغات الأخرى، وكل ما سبق ذكره يمثل تحديات لمقننة التحليل الصوتي والصرفي والإعرابي والدلالي للغة العربية.

إضافة إلى أن الحاسوب يفتقر إلى الفهم والحس اللغوي فيأخذ بعين الاعتبار التشابه الحاصل بين بعض الظواهر اللغوية لذا فإنه يتنبه إليها فقد يظن الأفعال (التفت، التحق) اسماً لا بدائها ب(ال) وكذا الأمر بالنسبة إلى (ياسر، يابس، ياسمين) لالتباسها بالنداء وتنبه أيضاً إلى المباني التي يتقاطع فيها الاسم والفعل نحو (أحمد، يزيد) مما قد يرد في النص فعلاً ويتوقع أن يلتبس على الحاسوب كونه اسماً.

خاتمة

ما يمكن قوله أن اللسانيات الحاسوبية أصبحت اليوم مجالا رائجا بين أوساط الدارسين، نظرا لأهميته لذا لا بد من مضاعفة الجهودات وتوجيه أنظارنا نحو إيجاد تقانات جديدة، والعمل على تطوير البرامج الحاسوبية من أجل أن تناسب لغتنا العربية واحتوائها في الحاسوب لمواكبة عصر الرقمنة.

الهوامش:

¹ غانم سعيد، تقطيع الكلمة العربية إلى أحرف وتمييزها، مجلة الراصد للعلوم الحاسبات، المجلد 2، العدد 2، 2005، ص 124.

² منصور بن محمد الغامدي زاحرون، مدخل الى اللسانيات الحاسوبية، ط 1، دار وجوه للنشر والتوزيع، الرياض: 2017م، ص 50.

³ أمين القلق ومنصور الغامدي، برنامج التحليل الصرفي في اللغة العربية — برنامج الخليل الصرفي، جامعة الدول العربية المنظمة، المنظمة العربية — التربية والثقافة والعلوم، كلية دار العلوم، جامعة القاهرة، 2011م، ص 02.

⁴ أحمد السامرائي، حوسبة المعجم العربي والقضايا المعاصرة، مجلة سر من رأى، جامعة سامراء، المجلد 9/ العدد 34، تموز 2013، ص 13.

⁵ سعد بن هادي قحطاني، تحليل اللغة العربية بواسطة الحاسوب، مركز اللغة الانجليزية، معهد الإدارة، الرياض، 2011، ص 6.