المعالجة الآلية للنحو العربي

Automated treatment of Arabic grammar

أ.أحلام سعيدي*

جامعة الجزائر-2-

البريد الإلكتروني: ahla30699@gmail.com

تاريخ الوصول.....،تاريخ القبول ..... تاريخ النشر.....

ملخص:

يعتبر النحو أحد أهم فروع علوم اللغة، ورغم أن الدراسات النحوية القديمة منها أو الحديثة قد حققت شوطا كبيرا حصرت من خلاله قواعد النحو وضبطت آليات اشتغال النظام اللغوي ككل، إلا أن معالجة النحو آليا لا تزال تنطوي على تحديات كبيرة سواء على الصعيد اللغوي أو الحاسوبي، وهذه التحديات في مجملها تطرحها طبيعة اللغة العربية وخصائصها والتي لا يلائمها كثير من الأساليب التي تم تصميمها لنحو اللغة الإنجليزية الذي يختلف عنها اختلافا جوهريا، مما يستلزم عدم الاكتفاء بالنماذج المستوردة الجاهزة والعمل على تصميم برامج وتطبيقات خاصة باللغة العربية.

الكلمات المتاحية: النحو، العربي، المعالجة ، الآلية.

Abstract:

The grammar is one of the most important branches of language science. Although ancient or modern grammatical studies have made great strides in limiting grammar rules and controlling the functioning mechanisms of the language system as a whole, the automated processing of grammar continues to present significant challenges, both linguistic and computer.

Keywords: Grammar, Arabic, Processing, Mechanism.

* أ.أحلام سعيدي ahla30699@gmail.com

تمهيد:

يعتبر النحو أحد أهم فروع علوم اللغة، ورغم أن الدراسات النحوية القديمة منها أو الحديثة قد حققت شوطا كبيرا حصرت من خلاله قواعد النحو وضبطت آليات اشتغال النظام اللغوي ككل، إلا أن معالجة النحو آليا لا تزال تنطوي على تحديات كبيرة سواء على الصعيد اللغوي أو الحاسوبي، وهذه التحديات في مجملها تطرحها طبيعة اللغة العربية وخصائصها والتي لا يلائمها كثير من الأساليب التي تم تصميمها لنحو اللغة الإنجليزية الذي يختلف عنها اختلافا جوهريا، مما يستلزم عدم الاكتفاء بالنماذج المستوردة الجاهزة والعمل على تصميم برامج وتطبيقات خاصة باللغة العربية، ويظل التحليل النحوي يحظى باهتمام المعلوماتين في سبيل إيجاد خوارزميات ملائمة وتصميم برامج فعالة وناجحة يمكن الاستفادة منها في خدمة حوسبة اللغة العربية، وأمام هذا الطرح تتطرق هذه الدراسة إلى الحديث عن المعالجة الآلية للنحو متطلباتها وكذلك التحديات الكبرى التي تواجهها.

1- المعالجة الآلية للنحو:

إن المعالجة الآلية للنحو نتاج لاقتران الدراسات النحوية مع الحاسوب لخدمة الأهداف والغايات المتعلقة به، والتي تتحدد في حل المشكلات المعقدة التي تتصل بحوسبة النحو.

وتكمن الغاية من المعالجة الآلية للنحو في" تقديم توصيف شامل ودقيق للنظام اللغوي للحاسوب تجعله يضاهي الإنسان في كفاءته النحوية"[1]، وهنا يصبح الحاسوب قادرا على تركيب القواعد وتحليلها تماما كما يفعل الإنسان، وذلك بالاعتماد على ذاكرته الرقمية والخوارزميات الدقيقة المخزنة فيه، فبمجرد إدخال جملة ما إلى الحاسوب يتعرف عليها ويعطيها هذا الأخير تمثيلا رقميا خاصا ويقوم بإجراء التحليل المطلوب عليها.

"ويعتبر التحليل الآلي جزءا هاما من تطبيقات التدقيق اللغوي الموجودة في برامج معالجة الكلمات حيث إن معرفة صحة أو خطأ جملة معينة يعتمد على معرفة بنيتها، كما يتدخل التحليل النحوي الآلي بشكل كبير في تطبيقات التحليل الدلالي والترجمة الآلية وأنظمة استرجاع المعلومات"[2]، فالتحليل النحوي أحد ركائز التحليل اللغوي وكل تطبيق آلي متعلق باللغة سواء كان الترجمة الآلية، نظام استرجاع المعلومات، التدقيق الإملائي أو غيرها من التطبيقات فإنه يستلزم الاستفادة من مخرجات

التحليل النحوي الآلي، وهو الأمر الذي يحتم على الحاسوب وكل تقاناته الاستيعاب الكامل للمنظمة النحوية والاستناد إليها.

2 -الإطار العام للتحليل النحوي الآلي:

إن من أهم أجزاء التحليل اللغوي الآلي ما يعرف بالتحليل النحوي فهو محور مهم فيه من خلال تكامله مع باقي الأنظمة اللغوية الصوتية والصرفية والدلالية والتركيبية وفيما يلي تفصيل موجز لأهم خطوات التحليل النحوي الآلي:

أ- الخطوة التمهيدية من قبل مصمم الحاسب:

حيث يقوم الحاسوبي بتصميم نماذج أولية تساعده على القيام بهذا التحليل، ويجب أن يكون تصميمه مجردا ومبسطا بداية الأمر ثم يطوره تدريجيا من خلال الإضافة والتعديل عليه ليصل في نهاية الأمر إلى تشكيل نظام آلي متكامل وفق متطلبات التحليل المرجو القيام به.

وهذه العملية لا يمكن أن يقوم بها إلا مبرمج لديه خبرة ومعرفة كاملة بالميكانيزمات الحاسوبية من جهة وكذلك اللسانية باعتبار أنه سيقوم بتحليل لساني قبل كل شيء.

ب- الخطوات اللغوية في التحليل النحوي الآلي:

ومدار هذه الخطوة بداهة هو تحليل الجملة لفهم معناها وتحديد أجزائها لمعرفة بنيتها وتكوينها وينفذ هذا حاسوبيا بترجمة لغة الجملة المدخلة إلى أجزاء مجردة دلالية غير مركبة عن طريق الرمز أو تحويلها إلى لغة مصطنعة من اللغات التي صمم الحاسوب عليها كاللغة التي صمم الحاسوب عليها كاللغة الخاصة بأنظمة قواعد البيانات[3].

ونشير هنا إلى أن أغلب الدراسات والتطبيقات المنجزة والمتعلقة بتحليل اللغة ومستوياتها أخذت في مجملها منحنيين أو اتجاهين يقوم أحدهما على عرض النص المراد تحليله على قواعد وقوانين لغوية قد أعدت وبرمجت في الجهاز الآلي سابقا ثم طبقت تلك القواعد والقوانين على النص المراد تحليله"[4]

أي أن هذا الاتجاه يعتمد على النص وينطلق من النص المدخل فيقوم بإدخاله في جهاز الحاسوب ويحلله وفق القوانين الواردة في البرامج المصممة في جهاز الآلي سابقا.

وهذا على عكس الاتجاه الثاني الذي يقوم فيه الباحثون "بتخزين عدد كبير من النصوص المحللة يدويا في الجهاز الآلي، لتكون نموذجا لغويا يقاس عليه فيعرض الجهاز النص المدخل على تلك النماذج للمقارنة واستنتاج الاحتمالات الصحيحة بحسب أوجه التوافق المتوافرة بين النصين"[5]

والمقصود هنا هو أن التحليل يعتمد على المقارنة بين النص المدخل المراد تحليله ونموذج لنص محلل ومخزن سابقا في الجهاز الحاسوب الذي يعتبر مقياسا يقاس عليه وبالتالي تحدد التوافقات والاختلافات وتحدد طبيعة النص أو الجملة المدخلة حسبه.

وبالعودة إلى ضوابط التحليل وآلياته فإنه من البديهي أن يقوم المحلل بتحليل نصه بتخزينه وتقسيم الجملة إلى مفرداتها الأصلية والتي أساسها في اللغة العربية اسم وفعل وحرف، كما يتفق على ذلك النحو العربي وقد ورد عن ألفية ابن مالك:[6]

كلامنا لفظ مفيد كاستقم اسم وفعل وحرف الكلم

فهذه العناصر الثلاثة هي بؤرة التحليل النحوي وهي التي تقوم عليها التراكيب والجمل وكل وحدة لغوية في اللغة ويكون مردها إما اسما أو فعلا أو حرفا.

وبعد القيام بتصنيف عناصر الجملة يجب أن يتم تجميع العناصر الرئيسية لكل فئة ثم تصنيف كل عنصر إلى جميع أجزائه إلى أن يبقى شيء يقبل التجزئة فكل عنصر كان(اسما أو فعلا أو حرفا) ينطوي في حد ذاته على أنواع وتصنيفات وتقسيمات خاصة به وهذه التقسيمات من عدة نواحي وبإمكانها أن تتفرع بدورها إلى تقسيمات أخرى:

-المجموعة الاسمية:

تصنف فيه عناصر الاسم في اللغة العربية: المعرفة والنكرة، المفرد والمثنى والجمع، المذكر والمؤنث، العاقل وغير عاقل، المجرد والمزيد، الجامد والمشتق، ويصنف المشتق إلى اسم الفاعل واسم المفعول والصفة المشبهة و اسم الآلة، اسم المكان واسم الزمان، المصدر ويصنف إلى المصدر الصريح و المؤول والمصدر الميمي واسم الهيئة، اسم المرة، الاسم المنقوص والمقصور والممدود والصحيح وأيضا المبني المعرب، البسيط والمركب وهكذا.[7]

-المجموعة الفعلية:

تصنيف عناصر الفعل في هذه المجموعة إلى الفعل الماضي، المضارع والأمر، الصحيح و المعتل والذي ينقسم بدوره إلى مثال أجوف ناقص لفيف، وأيضا من ناحية التجريد والزيادة يصنف الفعل إلى مجرد ومزيد، المبني والمعرب، والتام والناقص كما يصنف الفعل تحت كل أوزانه التي لا يخرج عنها[8].

-المجموعة الحرفية:

وتصنف إلى: أدوات الشرط والاستفهام والنفي والنهي والجر(الإضافة) والنداء والجواب والتفسير والعطف والتنبيه والاستثناء والمصدرية والتعليل والعلة والتفصيل والقسم والخطاب والتثنية والجمع والتثنية والتعريف والتأنيث والإضراب والتنفيس والاستدراك ثم تقسم كل مجموعة إلى أجزاء ذرية أعني أجزاء مجردة غير قابلة للتقسيم والتجزئة نحو (مذكر ومؤنث غيبة خطاب تكلم إفراد جمع تثنية وهكذا.....

فهذه هي الفئات التي تتكون منها الدراسة النحوية والصرفية وفي داخل كل فئة سمات يبدأ بتعرفها وتخص كل سمة برمز يميزها ثم تدخل تلك الرموز في معادلات رياضية في الحاسوب مثال ذلك تحليل الفعل **(كتبت)**:[9]

الفئة: فعل

الزمن: ماض

الوزن: فعل

السمات المرتبطة: مفرد، مؤنث، غائب

الدلالة المعجمية: الكتابة

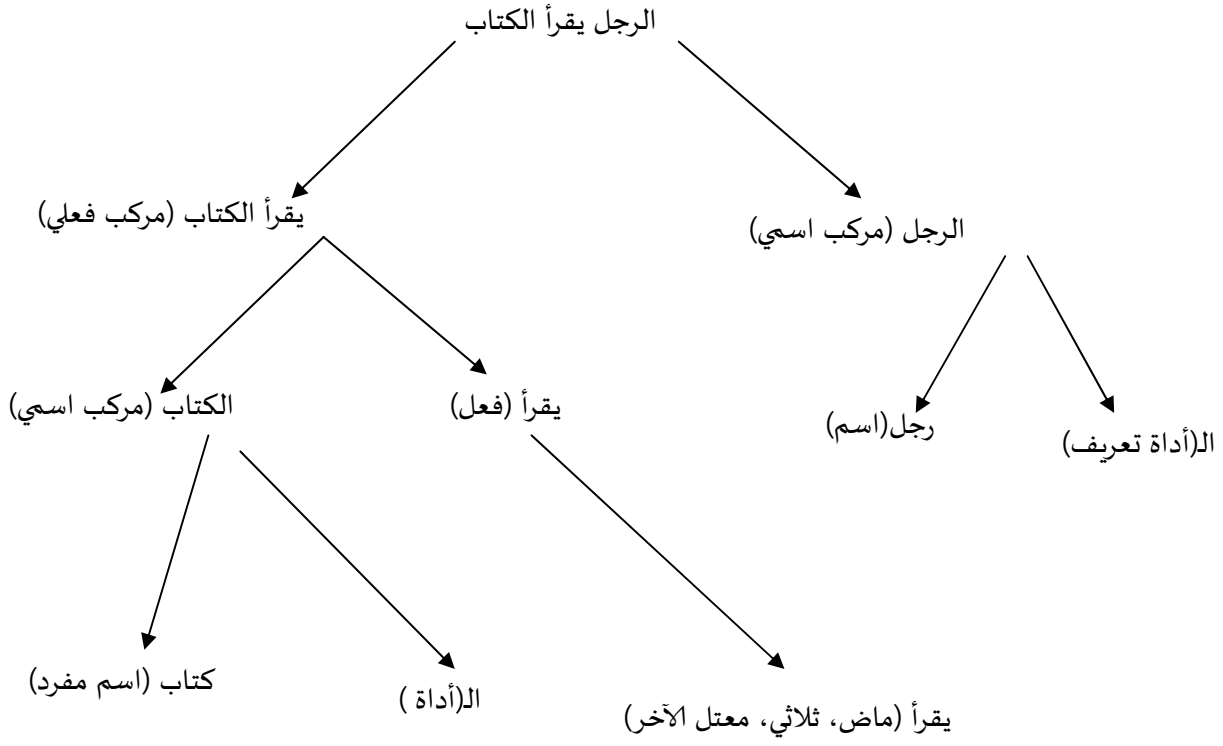
المواد الصوتية : ب ت ت ك: كتبت

ترتيبها: ك ت ب ت

فكل المعلومات والبيانات النحوية السابقة بتصنيفاتها وتقسيماتها يجب أن تبرمج في جهاز الحاسوب حتى يتمكن من فهم النظام النحوي للغة العربية.

ج- بناء نماذج تراعي قوانين التركيب النحوي في الجملة العربية يتم الاحتكام إليها والتصميم وفقها:

يعتمد نجاح المعلومات النحوية التي يظهرها الحاسب على صحة المعادلات المدخلة فيه وفي لغة الحاسب الآلي أنظمة تعرف بأبنية السمات، وظيفتها تقديم وسيلة سليمة للتعبير عن الضوابط النحوية التي قد يصعب التعبير عنها باستخدام أنظمة وقوانين التركيب النحوي في الجملة العربية، ومن أمثلتها قوانين الجملة الفعلية وجملة النداء وجملة الشرط والجملة الاسمية، فإذا أخذنا كمثال الجملة الاسمية وقانونها المعروف بالمركب الاسمي ومركب الفعلي، نحو: **محمد يكتب**، فلابد من تحقيق التوافق المعجمي بين الاسم المسند إليه والكتابة لأنه لا يصح إسناد الكتابة لمن لا تعقل منه، كما أنه لابد من اتحاد الضمير ومرجعه فلا يصح نحو **عائشة يكتب** أو **محمد تكتب** ونحو ذلك، فلكل مفردة أو وحدة سماتها الخاصة في نظام اللغة وتتكون الجمل عن طريق قوانين بنية العبارة باختيار الوحدات المعجمية والصرفية لتكون الجملة مقبولة وصحيحة[10]، كما يظهر في المثال الآتي لتحليل جملة" **الرجل يقرأ الكتاب**":



د-تخزين نتائج التحليل النحوي حاسوبيا وتوظيفها في بناء برامج تحليل نحوية آلية:

بعد إنهاء خطوات التحليل النحوي للكلمات والجمل لابد من تخزينها في ذاكرة الجهاز لاسترجاعها عند الطلب، ومن هنا وضع المبرمجون الحاسوبيون سجلات التخزين رقمية دقيقة جدا، تحمل في طياتها جميع الفئات النحوية التي سبق تحليلها كاملة ليستعان بها في تحليل أي نص جديد يراد تحليله.

كما يمكن الاستفادة من هذه المعطيات في بناء وتصميم برامج وتطبيقات حاسوبية خاصة بالمستوى النحوي وهذه التطبيقات يمكن الاستعانة بها في عدة برامج أخرى كالترجمة الآلية واسترجاع المعلومات وتحليل النصوص وعمليات الإحصاء الآلي، وكذلك التصحيح الآلي للنصوص وغيرها.

3-التحديات التي تواجه التحليل النحوي الآلي:

تواجه المعالجة الآلية للنحو العربي تحديات كثيرة ومتداخلة ولعل أهمها:

أ-غياب صياغة رسمية للنحو العربي:

يلزم لمعالجة النحو العربي آليا وضع قواعد في صياغة رسمية مكتملة، ووفقا للنموذج النحوي المتبع فهناك عدة نماذج نحوية وهي مختلفة من حيث منطلقاتها الأساسية وقابليتها للمعالجة والتعامل مع مشاكل النحو للتحليل اللغوي الأعمق.[11]

ب-إسقاط علامات التشكيل في معظم النصوص العربية:

فإسقاط علامات التشكيل عن الكلمات سيؤدي حتما إلى حدوث لبس وكل هذا له تأثيره على المعالجة الآلية للنحو.

ج-تعدد حالات اللبس النحوي وتداخلها:

فاللبس وبلا منازع هو أكبر مشكلة تواجه التحليل النحوي الآلي ولا يقتصر اللبس في العربية على ذلك اللبس الناجم عن غياب التشكيل بل هناك أنواع أخرى من اللبس تظل موجودة حتى مع وجود التشكيل[12]، ككلمة (ساق).... يمكن أن تدل على الساق كعضو في جسم الإنسان ويمكن أن تدل على الفعل ساق بمعنى أوصل، وأيضا لفظة يزيدُ يمكن أن تلتبس بين كونها فعلا مضارعا أو اسما.

د- المصاعب الناتجة عن المرونة النحوية:

تمثل المرونة النحوية للعربية إحدى المشاكل الأساسية للتحليل الآلي، فعلى المحلل أن يقتفي أثر مواضع التقديم والتأخير وأن يتعرف على جميع الحالات النحوية الممكنة من تغييرات حاصلة مقارنة بالجمل الأصلية[13]، إضافة إلى إشكالية طول الجمل العربية وتعقدها مما يصعب تحليلها آليا وهذا نظرا للتماسك بين الجمل وفقرات النصوص العربية والتي تميل لاستخدام حروف العطف والوصل والاستئناف والاستدراك وما إلى ذلك.

و-ظاهرة الحذف النحوي:

إذ يصادف أن يتم حذف لأحد العناصر في الجملة العربية وعلى المحلل الآلي أن يخمن مواضعه ويستنبط ما تم حذفه،[14] ويشمل ذلك خاصة حذف الضمير (الاستتار) وحذف الفعل أو الفاعل أو المفعول به وغير ذلك.

ه-قصور المعجم العربي نحويا ودلاليا:

فالمعجم لابد له أن يتضمن جميع المعطيات اللغوية الصوتية والصرفية والنحوية والدلالية والمقامية والتي لا يمكن استنتاجها باللجوء إلى القواعد أو المبادئ العامة،[15] و هذه المعطيات ضرورية ولا غنى عنها في المعالجة الآلية النحوية.

ر-عدم توفر الإحصائيات النحوية:

تحتاج المعالجة الآلية للنحو العربي إلى كثير من الإحصائيات اللازمة لترشيد أدائها ففي مواجهة قائمة من البدائل المحتملة لبنية الجملة، وخاصة عندما يتعلق الأمر بغياب التشكيل، على النظام الآلي أن يحدد الأولوية التي يتناولها هذه البدائل على أساس معدلات حدوثها داخل النصوص العربية[16].

4-بعض المقترحات للمساهمة في تفعيل المعالجة الآلية للنحو:

وفي ختام هذه الدراسة نحاول إيراد مجموعة من المقترحات التي من شأنها المساهمة في تفعيل المعالجة الآلية للنحو، ونذكر منها:

-يجب الاستفادة القصوى من المظاهر النحوية المختلفة(المطابقة، الإعراب، التضام، الرتبة...) واستغلالها في بناء أنظمة المعالجة الآلية، وذلك لإمداد النظام الآلي بالقرائن التي تعينه على فك اللبس الناجم عن غياب التشكيل.

-تصميم نظام نحوي آلي بحيث يكون قادرا على التعامل مع حالات اللبس المتنوعة، وأن تتوفر لديه وسائل فنية ومادية لاحتواء مجموعة الاحتمالات الناشئة عن التفاعل ما بين حالات اللبس المذكورة[17].

-فصل النحو والمعجم عن برنامج المعالجة ذاته، أو بقول آخر فصل المعارف اللغوية عن شق البرمجة الإجرائية الذي يتعامل معها حتى لا يؤدي التغيير في قواعد النحو أو معطيات المعجم إلى تغييرات البرنامج[18].

-ربط التحليل النحوي بالتحليل الصرفي في المعالجة الآلية العربية، فالانطلاق في التحليل النحوي يبدأ من خلال الاستفادة من مخرجات التحليل الصرفي والمعلومات المقدمة منه وكل هذا من شأنه أن يساعد على فك اللبس خاصة عند غياب علامات التشكيل.

-استغلال الإحصائيات النحوية والصرفية في ترشيد أداء المحلل النحوي الآلي لتجديد الأولويات التي يتناول بها قائمة البدائل المحتملة لبنية الجملة رهن التحليل[19].

-التركيز على اللغة العربية الفصحى الحديثة وبالتالي تخليص المعجم من الألفاظ والصيغ المهجورة، حيث يؤدي ذلك إلى تضييق نطاق حالات اللبس الناجمة عن غياب التشكيل خاصة وأن العربية القديمة أكثر حساسية لغياب التشكيل من العربية الحديثة[20].

-توفير الوسائل التشخيصية العملية للتعامل مع حالات الخطأ وإظهار أسباب فشل المحلل الآلي في تحليل جملة ما، وتحديد موضع حدوث هذا الفشل، وذلك باقتفاء أثر المحلل الإعرابي بتسجيل خطوات عمل النظام بصورة تمكن من اكتشاف مصدر الخطأ والتحاور مع المستخدم لتصحيحه[21].

الهوامش:

1-أحمد على لقم، حاسوبية النحو العربي بين المعوقات والمقترح، سلسلة دراسات وبحوث محكمة، المجلة العلمية للجمعية المصرية للكومبيوتر التعليمي، العدد، 2018، ص180

2-صلاح نجم، التحليل النحوي الآلي (مقرر علم اللغة الحاسوبي)، د-ت، ص1(www.alnajm.com)

3-عبد الله محمد بن مهدي الأنصاري، الدرس النحوي في ضوء الحاسب الآلي، (http://www.albagadi.net)ص271.

4- ينظر: المرجع نفسه، ص271

5-ينظر: يحي الحاج، عبد الأنصاري، تحديد الخصائص الصرفية للمفردات العربية آليا اعتمادا على بنية الجملة، ندوة المصادر اللغوية وأدواتها الحاسوبية،د-ت، ص241

6- عبد اللطيف بن محمد الخطيب، متن ألفية ابن مالك، دار العروبة للنشر، الكويت، 2010، ص1

7- عبد الله محمد بن مهدي الأنصاري، الدرس النحوي في ظل الحاسب الآلي، ص272

8- المرجع نفسه،ص272

9-المرجع نفسه، ص272

10-المرجع نفسه، ص275

11-نبيل علي، اللغة العربية والحاسوب، دار تعريب للنشر، د-ب، 1988، ص391

12-ينظر: المرجع نفسه، ص 392

13- المرجع نفسه، ص394

14-المرجع نفسه، ص398

15- ينظر: المرجع نفسه، ص399

16- المرجع نفسه، ص401

17-المرجع نفسه،ص402

18-المرجع نفسه،ص402

19-المرجع نفسه،ص403

20-المرجع نفسه، ص404

21-المرجع نفسه،ص404

قائمة المراجع:

الكتب:

- نبيل علي، اللغة العربية والحاسوب، دار تعريب للنشر، د-ب، 1988

-عبد اللطيف بن محمد الخطيب، متن ألفية ابن مالك، دار العروبة للنشر، الكويت، 2010

المقالات والدوريات:

-أحمد على لقم، حاسوبية النحو العربي بين المعوقات والمقترح، سلسلة دراسات وبحوث محكمة، المجلة العلمية للجمعية المصرية للكومبيوتر التعليمي، العدد، 2018.

-صلاح نجم، التحليل النحوي الآلي (مقرر علم اللغة الحاسوبي)، د-ت، ص1(www.alnajm.com)

-عبد الله محمد بن مهدي الأنصاري، الدرس النحوي في ضوء الحاسب الآلي. (http://www.albagadi.net)

-يحي الحاج، عبد الأنصاري، تحديد الخصائص الصرفية للمفردات العربية آليا اعتمادا على بنية الجملة، ندوة المصادر اللغوية وأدواتها الحاسوبية،د-ت.