



المعجم العربي ودعم الدراسات القرآنية الحاسوبية*

عمر مهديوي¹، سلوى حمادة السيد²

¹جامعة مولاي إسماعيل، المملكة المغربية

²جامعة طيبة، جامعة عين شمس، جمهورية مصر العربية

الخلاصة: تتوخى هذه الورقة التعريف باللغويات الحاسوبية (Computational Linguistics) وتطبيقاتها العملية في معالجة النصوص اللغوية الطبيعية عامة، والنصوص الدينية على وجه الخصوص (النص القرآني نموذجاً)، خصوصاً وأن تقنية المعلومات تتيح جملة من المفاهيم الصورية من قبيل الأتمتة والخوارزميات وقاعدة المعطيات، وقاعدة المعارف والتحليل والتوليد الآليين، وغيرها. كل هذه المفاهيم تساعد على تحليل المعطيات بدقة، والتوصل إلى نتائج موضوعية، ومن ثم التوصل إلى إدخال المعلومات القرآنية إلى عقل الآلة واسترجاعها وفق نظم معلوماتية متطورة وذكية، وبالتالي بناء مدونات لغوية حاسوبية للنصوص.

كما تتوخى أيضاً تبيان أهمية المعجم العربي بنوعيه الورقي والإلكتروني في دعم الدراسات القرآنية، على اعتبار أن ما أنجز لحد الساعة في هذا المجال ما يزال قليلاً، وحتى إن وجد، فإنه لا يستوعب مستويات النص القرآني المعجز لفظاً ومعنى.

لهذا سنحاول في هذه الدراسة أن نوضح المجالات الأساسية في النص القرآني التي يمكن أن تستفيد من تطبيقات الحاسب الآلي، والتي من شأنها أن تساعد على بلورة وتنفيذ برمجيات حاسوبية إسلامية خاصة بالقرآن الكريم، ومن ثم السعي الحثيث نحو بناء معجم حاسوبي للقرآن الكريم، موجه للمسلمين وغير المسلمين، يساعد جميعاً على فهم النص الديني واستكشاف خصائصه اللغوية والأسلوبية والصرفية والنحوية.

الكلمات الجوهرية: اللغويات الحاسوبية، النصوص الدينية، القرآن الكريم، قاعدة المعطيات، الخوارزميات، تطبيقات الحاسب اللغوية، المعجم العربي، الدراسات القرآنية الحاسوبية، مدونة القرآن الكريم.

1. المقدمة

أ. أهمية الدراسة وأهدافها:

تكمن أهمية هذه الدراسة في كونها تسعى إلى بناء وتطوير برمجيات حاسوبية خاصة بالقرآن الكريم من بين غاياتها تقديم النص القرآني إلى الناس كافة، في صورة تمكنهم من تلاوته تلاوة مجودة، وفهم معانيه، وتيسير البحث عن مفرداته أو تراكيبه أو معانيه أو بناء مدونات قرآنية. ففي إطار تعليم تلاوة القرآن الكريم باستخدام الحاسوب، بذلت جهود مهمة من قبل علماء اللغة والتجويد والحاسوب، إذ توصل هؤلاء إلى توصيف الأصوات اللغوية توصيفاً حاسوبياً من خلال ضبط مخارجها وصفاتها، وتقديم المعايير الصوتية الدقيقة التي يختص بها الأداء القرآني. وقد مكّنهم ذلك من تحقيق نتائج غاية

* - البحث في الأصل نسخة منقحة عن ورقة مقدمة إلى الورشة الدولية الأولى عن المعالجة الآلية للغة العربية، كلية الحاسبات وتقنية المعلومات 23-24 تشرين أول، 2015 جامعة تبوك، المملكة العربية السعودية.

في الدقة والموضوعية بسبب استخدام تقنيات معلوماتية متطورة، ومن ثم تحقيق درجات عالية من التفاعلية بين الإنسان المُتعلِّم والحاسوب المُعلِّم.

أما فهم النص القرآني وتفسيره، فقد سعت بعض البرامج المحوسبة إلى تقديم تفاسير القرآن الكريم في صور إلكترونية، وربط كل آية بتفاسيرها المختلفة، وما قيل فيها من إعجاز قرآني بمختلف أنواعه وأشكاله، بحيث يستطيع المتصفح أن يربط بين الآية وما جاء من أقوال وآراء، وهذا جهد طيب حقق بعض ما استهدفه من غايات ومرام.

وأما مجال الدراسة الحالية فهي تصبو إلى استغلال تقنيات الخطاب المعجمي في دعم الدراسات القرآنية الحاسوبية أخذاً بعين الاعتبار التداخل المشترك بين المعاجم والمدونات القرآنية المحوسبة.

ب. منهجية الدراسة:

يندرج هذا العمل في إطار اللغويات الحاسوبية التي تعنى أساساً بدراسة النصوص اللغوية بنوعيتها العادي وغير العادي أي تطويع الحاسوب لخدمة اللغة العادية أو الدينية من خلال تطوير برمجيات معلوماتية متطورة تهم الأصوات والصرف والنحو والدلالة والمعجم. وبهذا فقد وزعنا محتويات هذه الدراسة إلى العناصر الرئيسة التالية:

1- ملخص الدراسة، ويضم أهمية الدراسة وأهدافها؛ ومنهجية البحث.

2- تمهيد

3- نظرة عن المعجم العربي المعاصر ودوره في دعم الدراسات القرآنية

4- القرآن الكريم والتطبيقات الحاسوبية

5- خاتمة البحث: النتائج والتحديات.

2. تمهيد الدراسة:

تعتبر المعالجة الآلية للنصوص اللغوية الطبيعية عامة، والنصوص الدينية خاصة فرعاً تطبيقياً من تطبيقات اللغويات الآلية التي تعنى أساساً بمعالجة اللغات الطبيعية حاسوبياً، من خلال إقامة حوار بين الإنسان والآلة عبر الوسيط اللغوي، بهدف تطوير وبلورة برمجيات حاسوبية تدعم الدراسات القرآنية وغير القرآنية.

أ. نظرة عن المعجم العربي المعاصر ودوره في دعم الدراسات القرآنية

يشكل المعجم اللغوي ركيزة أساساً للتقدم المعرفي لدى الأمم، فهو أحد الأدوات المهمة في التعليم والبحث العلمي واكتساب المعرفة والترجمة والتأليف. وقد أدرك علماء الأمة في صدر الحضارة الإسلامية العربية أهمية المعجم، فكان من أول ما ألفوه معجم اللغة العربية المسمى بـ "العين" الذي كتبه الخليل بن أحمد الفراهيدي في أوائل القرن الثاني الهجري، ثم توالى الكتب العربية الأخرى تبعاً في شتى العلوم والمعارف كالطب والصيدلة والبصريات والرياضيات وغيرها.

إن الغنى في المعارف والحاجيات ومتطلبات الصناعة المعجمية بنوعيتها الورقية والآلية فرض التمييز بين أنماط المعاجم:

1. المعاجم أحادية اللغة: وينحصر دورها في وصف اللغة في المعجم، إنها تقوم بدور الإحالة الوصفية بالنسبة

لمتكلم اللغة الخاصة، ومنها ما يتخذ طابعاً موسوعياً كلسان العرب لابن منظور؛

2. المعاجم ثنائية اللغة: يسمى هذا النوع بمعاجم الترجمة التي تنقل من لغة إلى أخرى، لأنها تقوم بدور الوصف من

اللغة المصدر إلى اللغة الهدف؛

3. المعاجم متعددة اللغة: هي التي توظف لغات متعددة، وتنقل المعارف والمعلومات بين متكلمين مختلفين؛

4. معجم المترادفات: وهو معجم يبحث في العلاقات الدلالية بين المفردات² كالترادف والتضاد والمشارك اللفظي أو المعنوي.

5. الذخيرة اللغوية؛ هي عبارة عن بنك آلي من النصوص اللغوية الطبيعية.

6. المعجم الحاسوبي يطلق عليه في الأدبيات الرقمية بالمعجم الرقمي أو الآلي وهو عبارة عن قاعدة معطيات مشفرة، تشمل جميع المستويات اللسانية كالأصوات والصرف والتركيب، بالإضافة إلى بناء معاجم إلكترونية للدلالة على الأقل في مستواها الصوري الذي يحدد العلاقات المنطقية بين مختلف مكونات المتواليات اللسانية المقبولة في وجهيها الحقيقي والمجازي. وتعتبر قواعد المعطيات اللسانية في مختلف مستوياتها والهندسة المعلوماتية عماد هذا النوع من المعاجم.

7. المعجم المحوسب (Lexique informatisé): هو عبارة عن مدونات لغوية صغيرة أو كبيرة الحجم، تخزن في أقراص، باستخدام تقنية المعلومات، يظهر الجانب الحاسوبي فيها على مستوي الإدخال والإخراج، حيث تنجز كل عملية من تينك بسرعة، وبأقل كلفة. وتزخر الشبكة Internet بنسخ إلكترونية من هذا النوع من المعاجم في مجالات شتى، وهي ليست بمعاجم آلية البنية، والجانب الحاسوبي يكمن في بنيتها التركيبية لا أقل ولا أكثر. لهذا لا يمكن أن تقوم بدور قاعدة المعطيات الآلية التي تشكل حجر الأساس في بناء معاجم حاسوبية للغات الطبيعية من جهة، وبالنسبة للتحليل الآلي للنصوص اللغوية من جهة أخرى.

وعليه، يمكن تقسيم المعاجم العربية، بوجه عام، إلى صنفين كبيرين هما: المعجم الورقي/ العادي، والمعجم الإلكتروني الحاسوبي. وينقسم الورقي أي المطبوع إلى قديم وحديث كما بينا ذلك في الفقرة أعلاه، أما الآلي فينقسم بدوره إلى حاسوبي ومحوسب.

يعتبر المعجم الإلكتروني إذن أداة تطبيقية أساسية من تطبيقات المعالجة الآلية للغات الطبيعية، إلى جانب المحلل النحوي، والمحلل الصرفي، والترجمة الآلية، وتعليم اللغة باستخدام الحاسب، وغيرها. ولهذا فقيمة هذا النوع من المعاجم تكمن أساسا فيما يقدمه من معلومات صوتية ونحوية وصرفية ودلالية بطريقة صارمة ودقيقة، وأيضا في الأدوات والتقنيات التي يوظفها لتحقيق هذه الخدمة³.

وانطلاقا مما تقدم يمكن القول إن المعجم العربي بنوعيه الورقي والحاسوبي يشكل آلية لدعم التطبيقات الحاسوبية الخاصة بالقرآن الكريم اعتبارا للوظائف الرئيسة التي يقوم بها من جهة، واعتبارا أيضا من أن الدراسات المعجمية القرآنية ما تزال محدودة للغاية من جهة أخرى، وحتى إن وجدت، فهي معدودة على رؤوس الأصابع، لا تستوعب خصائص النص القرآني المعجز لفظا ومعنى، وقليلة هي المحاولات العربية التي تسعى إلى تغطية هذه الثغرة في البحث اللساني الحاسوبي الموجه أساسا لخدمة النصوص الشرعية وفي مقدمتها النص الديني القرآني. وعلى هذا الأساس، سعت هذه المحاولة إلى توضيح أهمية نقانة المعلومات في تطوير برمجيات حاسوبية للقرآن الكريم بهدف تحقيق خدمات جلى للمسلم وغير المسلم كما ورد في ملخص البحث.

3. القرآن الكريم والحاسب الآلي: نظرة تاريخية

تعود بداية الربط بين الحاسب الآلي و القرآن الكريم إلى ثمانينيات وتسعينات القرن الماضي وانتشار أنظمة ويندوز على الحواسيب الشخصية، مما فتح الباب لانطلاق العديد من البرامج الحاسوبية الخاصة بالقرآن الكريم وعلومه، ومن أشهرها

2- مهديوي عمر وسلوى حمادة السيد، المعالجة الدلالية الآلية للغة العربية، مجلة العربية، ع3، 2006.

3 - عمر مهديوي، المعجم الإلكتروني: قراءة في بعض التجارب الغربية والعربية، مجلة الدراسات المعجمية، ع7-8 يناير 2009، ص. 281.

(أبجد) و(سلسيل) وبرامج التفسير. إلى جانب هذا فقد ساعدت ثورة الشبكة المعلوماتية على التوسيع من دائرة الاستفادة من التقنية الحديثة، مما أسهم بالكاد من تطوير المحتوى الرقمي للقرآن الكريم على الويب من خلال تطبيقات حاسوبية متنوعة كما سنبين ذلك في ثنايا الفقرات الآتية.

وبعد ظهور شبكة المعلومات ازدادت خدمات تقنية المعلومات للقرآن الكريم، ويتجلى ذلك في ارتفاع عدد المواقع الإلكترونية التي تعنى بالقرآن وعلومه سواء تلك التي تهتم بتفسير القرآن أو معاني المفردات القرآنية أو ترجمة القرآن، ومشاكل الترجمة أو مخطوطات القرآن الكريم ونسخه أو تعليم القرآن الكريم وحفظه، أو تعليم التجويد وأحكام التلاوة، أو تعدد أشكال العرض بالصورة والصوت، أو تقديم كتب الأولين بسهولة ويسر. بالإضافة إلى بناء محرك للبحث في كلمات القرآن وجذوره ومواضيعه وغيرها من القضايا التي استأثرت باهتمام الباحثين والدارسين في هذا المجال. هذا وقد لاحظنا خلال السنوات الأخيرة، اهتماماً متزايداً من قبل عدد من المؤسسات العلمية بتنظيم عدد مهم من المؤتمرات والندوات العالمية تهم القرآن والتقانة الحديثة سواء في المملكة العربية السعودية أو المملكة الهاشمية الأردنية أو ماليزيا. ناهيك عن تأسيس جمعيات ومراكز بحثية موجهة لخدمة العلاقة بين تقنية المعلومات والقرآن الكريم⁴.

وتتخصص مجالات استفادة القرآن الكريم من تقنية المعلومات في المجالات التطبيقية الآتية:

-حوسبة الحرف القرآني، وتتم من خلال تجميع قواعد معطيات خاصة بالحروف القرآنية من جوانبها كافة وصولاً إلى كتابة المصحف بالرسم العثماني، وانتهاء بكتابة الخط العربي الحديث والضرورات الطباعة الحديثة. ويمكن الحديث في هذا المضمار عن العمل الجاد الذي قدمه الدكتور محمد زكي محمد خضر، والذي يرمي إلى بناء قاعدة معطيات خاصة بكلمات القرآن الكريم وحروفه، ناهيك عن أعمال جامعات أجنبية كجامعة ليدز ببريطانيا.

-أنطولوجيا القرآن الكريم: ويقصد بالأنطولوجيا اللغوية وضع الحجر الأساس للمعطيات اللغوية، حيث تحتوي على تعريف دقيق للمعنى الدلالي للبيانات المراد تبادلها، وتكون هذه التعريفات مكتوبة بلغة المنطق، بحيث يستطيع أي نظام فهمها وحسابها، بل والاستنتاج منها. فلقد ظهر في السنوات العشر الأخيرة الكثير من التطبيقات التي تُعتبر فيها الأنطولوجيا بالغة الأهمية مثل الحكومات الإلكترونية، التجارة الإلكترونية، محركات البحث، المكتبات الإلكترونية والمعاجم الإلكترونية والمحسوبة والمدونات الرقمية، وغيرها من التطبيقات. ويتضمن مشروع الأنطولوجيا القرآنية التطبيقات التالية⁵:

- بناء محرك بحث دلالي: تتوخى محركات البحث على اختلافها، تمكين المستخدم من الوصول إلى مبتغاه بسهولة ودقة في الانجاز. وتتعدد محركات البحث، وتتعدد مناهجها وفقاً لحاجات مستخدميها ومنها الأنواع التالية:

النوع الأول: محرك البحث البنيوي، وحدته الأساس الكلمة، وله أنواع نذكر منها:

أ. البحث المعتمد على المطابقة في جميع حروف الكلمة. إذ يستدعي محرك البحث الكلمات المتطابقة في جميع الحروف مع كلمة البحث مثل: عليم لا يستدعي لها إلا كلمة عليم؛

4 - مركز أبحاث تقنية المعلومات لخدمة القرآن الكريم وعلومه(نور)

<http://nooritc.org/?q=ar/project/details/112>

5 - ينظر حسين البسومي، المعالجة الحاسوبية للنص القرآني، مفهومها وأهدافها

<dspace.medi.u.edu.my:8181/xmlui/.../012.doc?...1>

ب. البحث المعتمد على المطابقة في بعض حروف الكلمة، سواء أكانت المطابقة في بنية الكلمة الأصلية أو في السوابق واللاحق مثل: كلمة (زب) يمكن أن تستدعي لها محرك البحث الكلمات التالية: رب- ري- ربه- قرب-قربان..)؛

ج. البحث المعتمد على المطابقة في الجذر، وإن اختلفت صيغ الكلمات مثل كلمة (جنة) لها محرك البحث الكلمات: جنة- جنتان- جنتين-جنات)؛

د. البحث المعتمد على المطابقة في الجذر، وإن اختلفت صيغ الكلمات مثال: كلمة عليم تستدعي لها محرك البحث للكلمات التالية: عالم- علم- علماء- يعلم....).

النوع الثاني: محرك البحث الدلالي، وحدته الأساس هي المعنى، وليس بنية الكلمة سواء أكان معنى المفردة أو معنى الآية القرآنية. وله صورتان:

الصورة الأولى هي المسماة بالبحث الموضوعي، وحدته الأساس هي معنى الآية القرآنية، وهو يعتمد على توزيع سابق للآيات القرآنية إلى الموضوعات المتصلة بها.

الصورة الثانية: البحث المفهومي، وحدته الأساس هي معنى المفردة القرآنية، ويقصد به البحث عن مفهوم أو أكثر متصل بالمفهوم المبوَّح به بعلاقة أو أكثر من العلاقات الدلالية التي تترابط بها مفردات الثروة اللفظية القرآنية، وهذا النوع من البحث ما تزال المكتبة القرآنية تفتقده مع أهميته الكبرى.

النص القرآني نص ثري وغني بالمعارف على جميع أنواعها. وعلى الرغم من وجود آليات عديدة للبحث القرآني، إلا أنها مازالت لا تعطي الدراسات القرآنية ما تصبو إليه من مساعدات فعالة. وقد حظيت اللغة الإنجليزية ببرمجيات شبيهة، وقد خطت خطوات لا بأس بها في ميدان البحث في النصوص. وقد خاضت البرمجيات المطبقة على اللغة الإنجليزية - في مجال البحث الذكي للنصوص أيضا، وخرجت بنتائج مقبولة لم يمكننا - إلى الآن - الاستفادة منها بالدرجة نفسها في التطبيق العملي على اللغة العربية. ومن هنا جاءت فكرة برنامج شامل وقوي لكل ما يمكن أن يفيد في فهم النص أو يجلي ما غمض منه. نقدم هنا آلية صقر لفحص النصوص واستخراج المعلومات والبحث الذكي. وهذا بالطبع يحتاج للدعم بكافة صوره. فإن هذا البحث نواة لمشروع كبير جدا لفحص النصوص اللغوية عامة، والقرآن الكريم خاصة واكتشافهما بشكل لا مثيل له في أية لغة حتى الآن، وقد قطعنا شوطا لا بأس به في مرحلة البحث الصرفي والإحصائي والتنقيب المعلوماتي. ويمكن لصقر أن يزودنا بمعلومات هائلة عن الكلمات والتراكيب والجمل المترابطة كما سيتضح في سياق البحث. ومن أهم ما استخدم فيه البحث الإحصائي هو بناء قائمة الكلمات الأكثر شيوعا في القرآن الكريم. ويمكن ألا تنقيد فكرة البرنامج بلغة ما حيث يمكن تطبيقه على اللغة الإنجليزية أيضا.

4. التطبيقات الحاسوبية للقرآن الكريم: المدونة القرآنية نموذجا

1.4 بناء المدونات الحاسوبية المذيلة:

1.1.4 - مفهوم التذييل

التذييل هو إضافة معلومات لغوية تفسيرية عن المدونة، ومنها على سبيل المثال: رموز أقسام الكلام؛ الاسم والفعل والحرف، يمكن أن تكون مفيدة في تمييز الكلمات التي لها الهجاء نفسه، لكن المعنى أو اللفظ مختلف، فعلى سبيل المثال الكلمة "كتب" قد تكون فعلا ماضيا ثلاثيا (كَتَبَ) أو اسما لمجموعة (كُتِبَ) أو فعلا بمعنى أرغمه على الكتابة. هذه الكلمات

لها الشكل نفسه، بينما علامات التشكيل تعطي أكثر من 12 معنى⁶، حيث إن علامات التشكيل في اللغة العربية تمثل جزءاً من الكلمة يتم إهماله في الكتابة غالباً. إن استعمال الطريقة البسيطة لرموز أجزاء الكلام تربط الكلمات برموز تشير لسمات ما. والتذييل متعدد الوظيفية، ففي حالة أجزاء الكلام POS يمكن للمعلومات نفسها عن أنواع الكلمات أن تستخدم في تأليف القواميس والمعاجم، و في الإعراب، وفي حصر قوائم التردد، وفي تأليف الخطاب، وفي العديد من التطبيقات الأخرى.

والاستعمالات المستقبلية للمدونات المذيلة تتغير دائماً عن الاستعمالات الأصلية بطريقة لا يمكن توقعها⁷. وتلعب المدونات اللغوية المذيلة دوراً مهماً في الإحصاء اللغوي الحاسوبي، أو ما يعرف بعلم اللغة الإحصائي.

2.1.4-أنواع التذييل المختلفة

وبالإضافة لأقسام الكلام، توجد أنواع أخرى من تذييل المدونات تبعا لمستويات التحليل المختلفة، ومن أمثلته:

أ. التذييل الصوتي:

هو إضافة معلومات مثل كيفية نطق الكلمة في المدونة المنطوقة ومن أنواع التذييل الصوتي: إضافة معلومات عن السمات مثل التشديد والترنيم، والوقوف، والإجهاد، والترنيم، والإيقاع وغير ذلك.

ب. التذييل المعجمي والصرفي:

هو معلومات عن أوزان الكلمات والاشتقاق والسوابق واللاحق ونوع الكلمة، وغير ذلك من المعلومات الصرفية عن الكلمات. وينقسم إلى:

- التقسيم: تعرف الوحدات المعجمية من بعضها بعضاً عن طريق المسافات بينها. وهذا الأمر قد يكون صحيحاً في حالة (الكلمة)، ولكن هناك وحدات معجمية متعددة الكلمات لا تنطبق عليها وجهة النظر هذه مثل التعبيرات الاصطلاحية والمسكوكة والمتلازمات اللفظية والتراكيب وغير ذلك مما يعامل معاملة الوحدة الواحدة أي أن مجموع معنى كلماته لا يساوي معناه مثل: "رجعت بخفي حنين".
- التجذيع: وتعني تبسيط الكلمات بالمدونة لأصغر عنصر للكلمة. أي الشكل الأساس للكلمة، مثلما يوجد مع الكلمة كالكلمة الرئيسة في القاموس (ومثال ذلك: - كتب، سيكاتيون).

ج. التذييل النحوي:

هو معلومات مضافة عن كيفية تحليل الجملة، ويساهم في إزالة غموض حد الجملة. فنهايات الجمل عادة ما تكون (نقطة، علامات تعجب- استفهام-..). يمكن أن تكون النقطة نهاية الجملة أو جزء من مختصر (ج.ع.م) أو تكون تكملة (إلخ...) وقد تقوم بدور مزدوج كما في الحالة الأخيرة. ويضيف من ناحية التحليل النحوي معلومات عن الوحدات من العبارات وأشباه الجمل.

د. التذييل الدلالي:

هو معلومات مضافة حول التصنيف الدلالي للكلمات مثال الإسمي: "لعبة الفوت بول" والمصطلحي: "الرياضة" ويرجع إلى الأصناف الدلالية المختلفة، على الرغم من أنه ليس هناك اختلاف في التهجئة أو التلفظ. كما يفيد ترميز المعنى الكلمة في (حل لبس المعنى)⁸.

6 - مهديوي وسلوى حمادة، المعالجة الدلالية الآلية للغة العربية، مجلة فكر ونقد، عدد 82، أكتوبر 2006.

7- Martin Wynne, "Developing Linguistic Corpora: a Guide to Good Practice;2004.

8 - Martin Wynne, "Developing Linguistic Corpora: a Guide to Good Practice;2004.

هـ. التذييل التداولي:

هو معلومات حول أنواع النص الخطابى ودور الحوار، وكلمات الحوار -وهذا يحدث في لغة الحوار- وهو يرتبط بالمناسبات المختلفة؛ فقد يكون اعترافاً أو طلباً للتعليقات، أو مرحلة جديدة من المناقشة.

و. تذييل الخطاب:

هو معلومات مضافة حول مرجعية الضمائر في النص⁹، فعلى سبيل المثال مرجعية الضمير "هم" على سابق في الجملة "يعلم المدرس التلاميذ ويفهمهم الدرس" وهو "التلاميذ"، وتنفذ هذه المهمة غالباً من قبل المحللين من البشر للحصول على دقة عالية.

ز. التذييل الأسلوبى:

هو معلومات مضافة حول عرض الفكرة والخطاب والحديث المباشر، أو غير المباشر. ومعلومات مثل أنواع الأساليب من تعجب أو استفهام أو طلب أو تأكيد إلخ..

ح. أنواع أخرى من التذييل كما يقدمها [مارتن، 2004]:

من المحتمل ابتداء أنواع هائلة من التذييل قد تكون مفيدة لأنواع معينة من البحث، ومثال ذلك تذييل الفصاحة. فمن يعمل على البيانات المنطوقة، قد يرغب في تذييل مدونة الخطاب التلقائي بأنواع مختلفة من الرموز، ومن مجالات بحث المدونات التي ازدهرت في السنوات العشرة الأخيرة؛ المدونات التعليمية، فمثل هذه المدونات تشمل الكتابات التي أنتجها متعلمو اللغة الثانية حيث تذييل المدونات برموز للخطأ تشير إلى موضع خطأ المتعلم، و أنواع هذه الأخطاء. وتستخدم هذه المدونات في دعم تعليم اللغة لتلافي هذه الأخطاء أو اختبار المتعلمين الجدد¹⁰.

5. من أهداف الإحصاء اللغوي في المدونات اللغوية:

يعدنا الإحصاء اللغوي باستخدام صقر؛ بالكثير من المعلومات عن النص في جميع مستويات التحليل اللغوي، وهذا ينطبق أيضاً على مدونة القرآن الكريم. ومن مستويات الإحصاء اللغوي:

1.5 الحصر الإحصائي الشكلي:

الجدول 1: بعض مخرجات البحث الإحصائي الشكلي في المدونات غير المذيلة

حصر تكرار الكلمات	حصر المفردات الأكثر شيوع
حصر تكرار الكلمات التي لها التكرار نفسه في النص.	حصر تكرار الكلمات التي لها التكرار نفسه في الجملة نفسها.
حصر تكرار المقاطع.	حصر تكرار المركبات.
حصر تكرار كلمة معينة.	حصر تكرار ببداية الكلمات.
حصر تكرار بنهاية الكلمات.	حصر تكرار عبارة.
حصر تكرار حرف.	

أنظر فهرس (1) لبعض نتائج بعض أمثلة الحصر السابقة.

9 - نفسه، مرجع سابق.

10 - ولمزيد من التفاصيل راجع [سلوى حمادة، 2006]، و[نبيل علي، 2000].

2.5 الحصر من خلال قوائم أو معاجم:

ويمكن أن تستخدم أيضا في التحليل الإحصائي المبني على قواعد بيانات تبني من خلال قوائم أو معاجم أو غير ذلك. نعني بالبحث الإحصائي بالقوائم أن ندخل للبرنامج قوائم يبحث عن عناصرها داخل الذخيرة النصية. ومن أمثلة هذه القوائم:

- قوائم بالضمائر؛
- قوائم بالأسماء العربية؛
- قوائم بالأسماء الأجنبية؛
- التعبيرات الاصطلاحية والمتلازمة؛
- المصطلحات العلمية والتقنية.
- الكلمات المشهورة أو أي قوائم يبحث عنها المستخدم.

وعلى ذلك يمكن تقديم:

الجدول 2: بعض مخرجات البحث الإحصائي من خلال قوائم في المدونات غير المذيلة

حصر تكرار الضمائر	حصر تكرار الأدوات
حصر تكرار الأفعال.	حصر الأسماء:- أسماء الأعلام والأسماء المركبة.
حصر تكرار قائمة التعبيرات الاصطلاحية.	حصر تكرار تعبيرات مشهورة.
حصر تكرار قائمة بالأسماء العربية.	حصر تكرار معاجم خاصة.
إحصاء تردد كل وحدة لغوية (كلمة أو مركب أو فقرة أو جملة).	حصر تكرار المترادفات الأكثر شيوعا، وكذلك الأضداد والمتضادات، وجميع الظواهر اللغوية الأخرى.
حصر تكرار مصاحبات لغوية.	

3.5 الحصر اللغوي:

وتقوم فكرته الأساس على إيجاد علاقة بين المفردات المدرجة اعتمادا على المتصاحبات اللغوية للمفردات الرئيسة، وذلك من خلال علم اللغة الإحصائي اللغوي وتطبيقاته على المدونات المذيلة.

والبرنامج يتصل بقاعدة بيانات تنقسم إلى ثلاثة أقسام:

- القاعدة البنيوية: وتضم معلومات البنية والاشتقاق الخاصة بالمفردات؛
- القاعدة التركيبية: وفيها حصر للأنماط التركيبية الواردة في القاعدة الأساس؛
- القاعدة الدلالية: ويرتكز هذا القسم على الربط بين المفردات الرئيسة ومتصاحباتها اللغوية.

وقد عملنا في بحث سابق¹¹ على بناء قاعدة البيانات الأساس انطلاقا من مدونة لغوية مستمدة من لغة الصحافة المعاصرة، بالإضافة إلى مجموعة من الكتب والمؤلفات العربية المعاصرة في شتى العلوم والمجالات. حيث يتم حصر

¹¹ - سلوى حمادة، تحليل للجملة العربية مبني على مدونة نصية من النصوص المعاصرة، المؤتمر العاشر لجمعية هندسة اللغة، 15-16 ديسمبر، 2010، القاهرة، 2010.

تكرار التصاحبات (تركيب أكثر من كلمة، والبحث في مدى التصاحب بينهما) ⁽¹²⁾. عن طريق معادلة (PMI) والتي تختص بقياس الترابط بين كلمتين أو أكثر - اعتمادا على مدى التصاحب- من خلال حدوث كل منهما على حدة وحدثهما معا. ويعرض فهرس (2) الكلمات المتصاحبة وعدد الكلمات المتصاحبة ومدى التصاحب =2.

4.5 الحصر الصرفي:

الجدول 3: بعض مخرجات البحث الإحصائي الشكلي في المدونات المذيلة

حصر تكرار الجذر	حصر تكرار كلمات مشتقة محددة
حصر تكرار مشتقات كلمة محددة.	حصر الأسماء المشتقة Derived المستخدمة.

انظر فهرس (3) يعرض البحث لاستخدام الجذر.

5.5 الحصر التركيبي:

ويطبق على مدونات مرمزة صرفيا وتركيبيا ونحويا، ومن أمثلتها بنوك الأشجار العربية المشار إليها سابقا، والمدونة التي أشرفنا عليها وقمنا بتحليلها تدعى "مساك"¹³. انظر فهرس⁽⁴⁾ بعض ما أمكن حصره من خلال المدونة:

- صور الجمل العربية بتركيبيها وأشكالها وأصنافها المختلفة؛
- صور الأساليب العربية بتركيبيها وأصنافها المختلفة.
- مفردات اللغة المدروسة وتراكيبها؛ كما يمكن التمييز بين المهمل منها والمستعمل؛
- الصفة والموصوف والجار والمجرور والمضاف والمضاف إليه؛
- حصر الأفعال وعلاقاتها بغيرها من أجزاء الكلام؛
- علاقة الفعل بحرف الجر وتأثير ذلك على اختلاف المعاني والدلالات؛
- علاقة المشتقات بالفعل أو بجذره وتأثير ذلك على اختلاف المعاني والدلالات.

6.

7. التطبيق

1.6 تجهيز النصوص للبحث:

نقوم بتجهيز ذخيرة النصوص التي نريد البحث فيها عن طريق:

- تحديد النصوص التي نريد البحث فيها؛
- حذف الصور وتحويل الجداول التي تحتوي وحدات لغوية إلى نصوص؛
- تحويل الملفات لصورة txt إن أمكن، ويمكن التعامل معها بأي صورة، ولكن تكون موحدة لجميع الملفات؛

¹² - يمكن الاستفادة من هذه الوظيفة في استخراج أكثر المتلازمات والتعابير المسكوكة شيوعا ومحاولة معالجتها، ومن ثم توظيفها في عملية فك اللبس في البنية والتركيب. ويمكن كذلك دراسة أنماط تركيبية مثل (الجار والمجرور - النعت والمنعوت -... إلخ) والاستفادة من ذلك عند فك اللبس في البنية والتركيب.

¹³ - ينظر سلوى حمادة، تحليل للجملة العربية مبني على مدونة نصية من النصوص المعاصرة"، المؤتمر العاشر لجمعية هندسة اللغة، 15-16 ديسمبر، 2010، القاهرة.

- تجميع النصوص على ملف واحد للبحث فيه¹⁴. وسيطور البرنامج فيما بعد للبحث في النصوص الموجودة على الشبكات والتي نختارها بكلمات مفتاحية محددة.

2.6 متطلبات البحث:

- إجراء التحديدات الأساس التي قد يتطلبها البحث، كأن نقوم بتحديد الكلمات المهمة أو الكلمات المطلوب البحث عنها فقط، ويمكن أيضاً أن نحدد نهايات الجمل إذا تطلب البحث عن عدد الجمل؛
- بناء قاعدة بيانات المترادفات، ذلك أن المفهوم الواحد قد يمثل بكلمات مختلفة كثيرة (أي مترادفات)، وفي هذه الحالة نخزن قاعدة بيانات للمترادفات. وعند طلب البحث عن أي وحدة لغوية يسأل البرنامج (بحث مطابق أم عرض المترادفات أيضاً)؛
- بناء قاعدة بيانات المصاحبات اللغوية: إن الكلمة الواحدة قد تصف العديد من المفاهيم المختلفة (تعدد المعاني)، وفي هذه الحالة نخزن بعض المصاحبات لكل وحدة لغوية تحت مجال محدد، وعند طلب البحث عن الوحدة اللغوية يسأل البرنامج (حدد المجال، ثم يقدم خيارات للمستخدم).

3.6 صقر والفهرسة الآلية في اللغة العربية بناء على تكرار تردد المفردات:

يعمل المفهرس صقر¹⁵ على إيجاد مدى تكرار تردد الكلمات في نصوص اللغة الطبيعية لمجال ما دلالة على أهمية هذه الكلمات لأغراض التعبير عن المضمون، ومن ثم تكوين معجم هذا المجال، وإحصاء عدد مرات تردد الكلمات التي لها دلالة موضوعية، وترتيبها تنازلياً وفقاً لعدد مرات تردها في النص. وتعد الكلمات التي ترد على القمة، بعيداً عن الألفاظ الشائعة، والألفاظ غير ذات دلالة، هي الكلمات التي اختيرت لتكون مصطلحات أو كلمات مهمة في مجال الوثيقة، ويمكن تحديد عدد المصطلحات لكل وثيقة بناء على معايير محتملة، كتحديد عدد يتناسب وطول النص، أو الاكتفاء بالكلمات التي يزيد عدد مرات تكرارها عن حد معين، وكذلك يمكن إحصاء تكرار العبارات، حيث يمكن فهرسة الوثائق بالجمع بين الكلمات والعبارات.

وهناك مشاكل تتجم عن تعدد الأشكال النحوية للكلمات، وللتغلب عليها يمكن الاعتماد على برامج تتعامل مع جذور الكلمات. أما في اللغة العربية، فإن نظم التحليل الصرفي أو التجريد لا تزال في مراحلها المبكرة، وقد لاقت مؤيدين ومعارضين في خطواتها لإدخال الحركات في معالجة الكلمات العربية وانعكس هذا الانقسام على تطوير خوارزميات التحليل الصرفي في اللغة العربية التي تدعم الفهرسة¹⁶. وقد كان استرجاع المعلومات سبباً من أسباب تطوير هذه الخوارزميات، ولكنه كان الهدف الأساس لتطوير أساليب العربية. وقد أجرينا بعض التجارب في هذا السياق التي تهدف إلى دراسة الفهرسة واسترجاع البيانات باللغة العربية وخلصنا إلى أن استعمال جذور الكلمات وجذوع الكلمات يسفر عن نتائج أفضل.

4.6 أهمية آلية صقر: ما هو برنامج صقر؟

لكي نجيب عن هذا السؤال، لا بد أن نحدد في البداية من هو مستخدم فاحص النص صقر؟

¹⁴ - راجع سلوى حمادة، 2010.

- برنامج قامت بتصميمه وتنفيذه أ.د. سلوى حمادة. ¹⁵

¹⁶ - ينظر مهديوي، خوارزميات صرفية للغة من التحليل إلى التوليد، المؤتمر الخامس لمجمع دمشق، نوفمبر، 2006.

إن برنامج صقر المقترح سيفيد كل من يعمل في كتابة النصوص، حيث يحل محتوى النص، ويزود المستعمل بقائمة أو قاموس من الكلمات، العبارات، وترددها في الاستعمال ضمن محتويات النص. وهذه المعلومات التي ينتجها البرنامج تمكن من التدقيق في الاستعمال المفرط للكلمات أو تكرار الكلمات أو العبارات ضمن أية وثيقة، وهو ما يفيد المعجميين، ذلك أن المساهمات المقترحة للبحث الإحصائي بصقر في المدونات لدعم المعجم الحاسوبي المعاصر بما فيه من الألفاظ القرآنية:

- إنشاء معجم آلي يضم السمات اللغوية على جميع المستويات¹⁷. وهذا بدوره يؤدي إلى سهولة فك اللبس في اللغات الطبيعية¹⁸؛
 - استخلاص سمات مهمة للمعنى، وقواعد الكلمة لم تلاحظ من قبل اللغويين الذين لم يكن عندهم بيانات للعمل واضحة وميسرة من قبل؛
 - استعمال تردد الكلمة في معلومات المداخل المعجمية؛
 - تحديد المصاحبات اللغوية وفك اللبس عن طريق توظيفها في فكها؛
 - تحديد التعبيرات الاصطلاحية والمتلازمة واستخدامهما ومعانيهما من خلال السياق. وقد اكتشفنا تنوع المعاني واكتسابها معاني جديدة من خلال السياق - بحث -؛
 - دعم التصنيف الدلالي للحقول الدلالية؛
 - تدعيم المداخل المعجمية في المعجم العربي بسمات وخصائص ودلالات جديدة من المعلومات الإضافية التي استخلصت عن طريق تحليل المدونة (ومثال على ذلك: - الأشكال الأكثر شيوعاً، التضمين، الخ¹⁹. أنظر إثراء المعاجم العربية بالعلاقات والسمات الدلالية²⁰؛
 - مد المعجم بأمثلة من اللغة الحية متنوعة وفي جميع المجالات؛
 - دعم المعاني وشرحها - شواهد لغوية - يمكن استخدامها في تعليم اللغة للناطقين بها أو بغيرها؛
 - يوفر أداة للمقارنات بين اللغات البشرية.
- وبطبيعة الحال يمكن اختبار كل هذه التطبيقات على النص القرآني.

1.4.6 الدراسات الأسلوبية:

ويمكن أن يفيد الحاسوب في مجال النصوص الأدبية وأساليب الكتابة، وكذلك في تحليل أساليب النص القرآني مثلاً: من حيث أطوال الجمل والمفردات والتراكيب النحوية واستخدامها للضمائر والعطف والظروف والجمل الاعتراضية والظواهر اللغوية وغيرها. ويمكن عندها حصر وشرح هذه الأساليب. ويمكن أن توفر أداة برمجية لغوية لتقييم الأسلوب من حيث مستوى الصعوبة، وموقعه آلياً وشرح النص بصياغة أبسط. يتطلب تحديد مستوى الصعوبة ومواقعها آلياً بحثاً مستفيضة في مجال الأسلوبية، ووضع معايير دقيقة لتحديد درجة الصعوبة معجمياً ونحوياً ومنطقياً وهو أمر غير متوافر للغة العربية حتى على المستوى البشري.

2.4.6 مستوى التعليم:

يفيد في التعليم، والعديد من الدراسات مع جميع أنواع النصوص مثل: النص القرآني للبحث عن ظواهر الإعجاز اللغوي والعديدي وغير ذلك من الأبحاث في هذا المجال:

¹⁷ - ينظر مهديوي وسلوى حمادة، المعالجة الدلالية الآلية للغة العربية، مجلة فكر ونقد، عدد 82، أكتوبر 2006.

¹⁸ - نفسه، مرجع سابق.

¹⁹ - بخصوص إثراء المعاجم العربية بالعلاقات والسمات الدلالية، ينظر مهديوي وسلوى حمادة، 2006.

²⁰ - نفسه، مرجع نفسه.

3.4.6 النص اللغوي والقرآني:

- إن البحث عن المفردات القرآنية في النصوص العربية الحديثة يعود لعدة أسباب منها:
- دحض فكرة أن لغة القرآن لا تتفق مع الكتابات المعاصرة، وإنما هي لغة كتاب الله والدين فقط؛
- استغلال فصاحة القرآن في تفصيح لغة الكتابة المعاصرة؛
- شرح القرآن من خلال هذه الكتابات المعاصرة وتقريب المعنى لقرآني لأذهان متعلمي اللغة؛
- الكشف عن جانب لغوي هام في الكتابات والمنطوقات العربية الحديثة والقديمة وهو التأثير بالنص القرآني؛
- الكشف عن نقاط لغوية يمكن الاعتماد عليها واستغلالها في كثير من مسائل النقد الأدبي، وتاريخ الأدب، وأهم هذه النقاط فكرة تقسيم الجملة القرآنية إلى أنماط، وفكرة تغليب نمط على آخر، وتفضيل بعض النماذج، والإكثار منها كملح لغوي للكاتب سواء كان ملمحاً عاماً مطرداً في جميع أعماله أم كان خاصاً برواية أو مرحلة معينة²¹.
- ويمكن أن يطبق ذلك على أعمال أكثر من كاتب من خلال المدونة. كما يمكن أن تطبق على أكثر من كاتب خلال مراحل زمنية مختلفة أو أنواع نصية مختلفة أو مناطق جغرافية مختلفة وهكذا؛
- الكشف عن الإمكانات التركيبية الواسعة جداً للجملة العربية، وذلك من خلال التحويرات التي يجريها الكاتب على الجملة القرآنية سواء من الناحية الشكلية أو التركيبية²². ويمكن أن يطبق ذلك على أعمال أكثر من كاتب من خلال المدونة؛
- استغلال هذه الأنماط والنماذج القرآنية في تعليم وشرح اللغة العربية لأهلها أو متعلميها من الأجانب.
- استغلال هذه الإمكانات في التعرف على شخصية الكاتب وأسلوبه بصفة عامة أو في حالة التباس بعض الأعمال ما بين كاتبين أو أكثر؛
- تقديم نموذج للتدريب اللغوي يبدأ من الجملة القرآنية النصية، ثم يتدرج في التدريب على الجمل التي حور فيها الكاتب بعض التحويرات -ويمكن اتخاذ هذا النموذج معياراً أو منطلقاً لتدريبات أخرى²³- والعكس. والعودة للغة الفصحى في الكتابات العربية من خلال النمط القرآني والمحافظة عليه.

4.4.6 النصوص اللغوية:

- دراسة مواد القراءة الإضافية كالقراءة القرآنية، وقصص الأطفال، وكتب المتابعة لتحديد صلاحيتها من جهة، والمرحلة التي تصلح لها من جهة أخرى²⁴.
- وضع الاختبارات لتحديد المستوى اللغوي للممتحنين بشكل عام ولثروتهما اللفظية بشكل خاص، مما يساعد على وضع الدارسين في المستوى الذي يناسبهم²⁵.
- تحديد المترادفات الأكثر شيوعاً، وكذلك الأضداد والمتضادات، وجميع الظواهر اللغوية الأخرى.

21 - محمد حماد، صور الجملة القرآني في أدب نجيب محفوظ"، كلية دار العلوم، 1992..

-نفسه²² -

23 - محمد حماد، صور الجملة القرآني في أدب نجيب محفوظ"، كلية دار العلوم، 1992.

24 -ريما سعد الجرف، آلية اللغات والترجمة، جامعة الملك سعود، الجمعية الدولية للمترجمين العرب، 2008.

25 - نفسه، مرجع نفسه.

5.4.6 صقر والقرآن الكريم:

أ. ترميز مدونة القرآن الكريم

ولأن المدونات المذيلة للقرآن الكريم تتم يدويا وهي غالية الثمن، وللتذليل القرآني سمات خاصة حيث يمكن للكلمة والمركب والسياق أن يذيل بأكثر من طريقة وهذه القضية ليست مطروحة للمناقشة هنا فقد تم الاعتماد على مدونة قرآنية محللة آليا من قبل شركة صخر لإظهار أهمية هذا النوع من المدونات وكيف يستخدم في تحليل النص القرآني وشرحه للدارسين.

ب. فهرسة القرآن الكريم

يقوم البرنامج بإظهار عدد ورود كل كلمة في القرآن الكريم بالترتيبات المختلفة وفي فهرس (4) تم اختيار الترتيب من حيث الأكثر ورودا فجاءت الحروف في المراتب الأولى بالطبع. ويمكن عن طريق الضغط على كلمة معينة أن يظهر لنا البرنامج جميع سياقات هذه الكلمة كما هو واضح في السياق التالي.

ويمكن حصر التكرارات على مستوى السور والأرباع والأحزاب والأجزاء. حصر تكرارات السور المدنية والسور المكية. حصر السور المتساوية في عدد معين من التكرارات أو الكلمات أو الآيات. حصر الآيات التي لها نفس البدايات أو نفس النهايات.

ج. البحث في المدونات المذيلة/ المرمزة بأجزاء الكلام

تم عمل مدونة تمثل أغلب النصوص العربية وتحتوي أكثر من نصف مليون كلمة 26. وتم عمل تذليل لهذه المدونة على المستوى الصرفي والنحوي والأسلوبي من قبل لغويين بشريين مما كان مكلفا جدا كبحت فردي. وقد تم حصر التذييلات وعمل رسوم بيانية للتوضيح كما في فهرس (4).

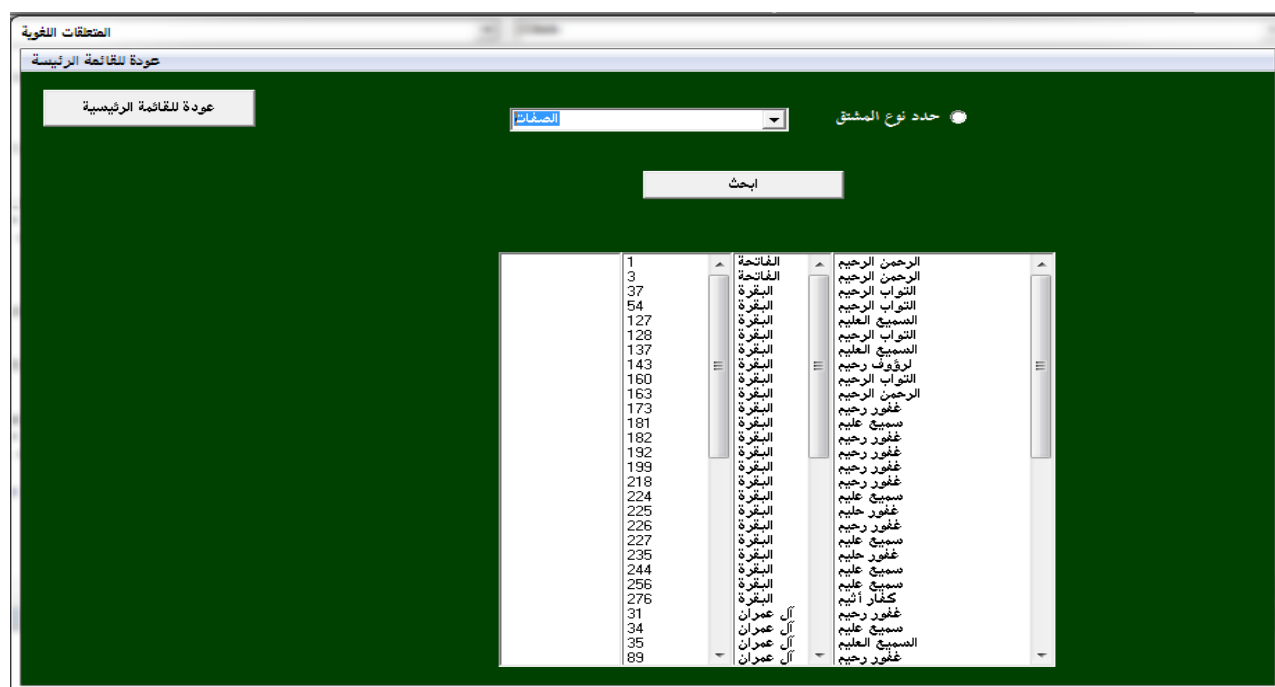
وتوضح معلومات التذليل بأجزاء الكلام على سبيل المثال ما يلي:

الجدول 4: حصر التذييلات

احتمال أن تكون الكلمة فعلا مشتقا.	احتمال أن تكون الكلمة فعلا جامدا.
احتمال أن يكون قبل الفعل سابق.	احتمال أن يكون بعد الفعل لاحق.
احتمال أن تكون الكلمة اسما مشتقا.	احتمال أن تكون الكلمة اسما جامدا.
احتمال أن تكون الكلمة اسما معرفا.	احتمال أن تكون الكلمة اسما نكرة.
احتمال أن تكون الكلمة حرفا.. وهكذا.	

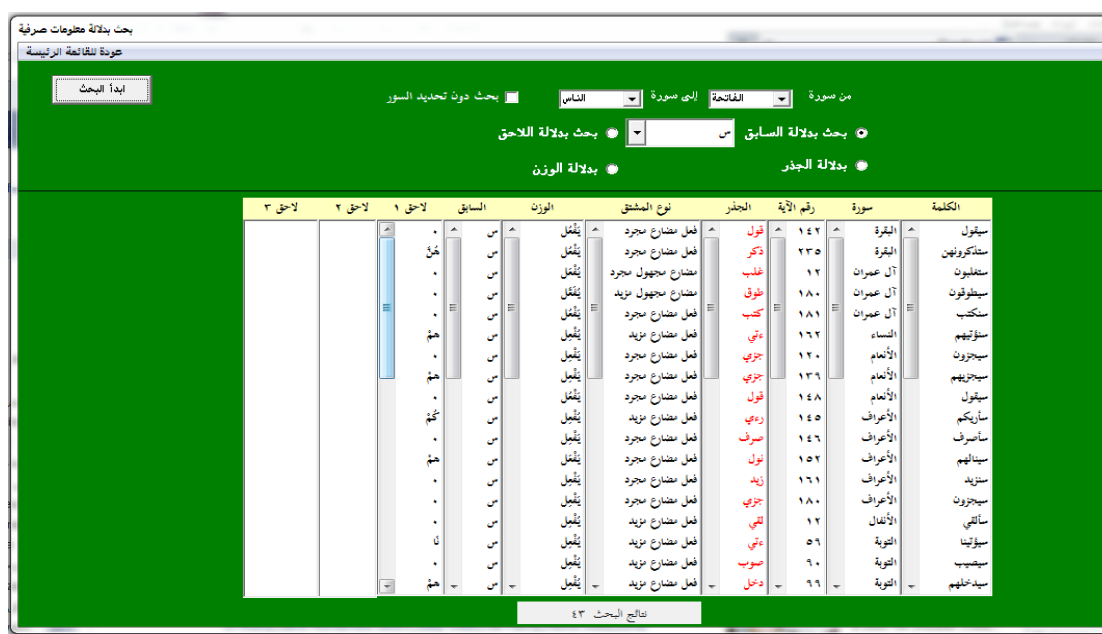
مثال: البحث عن الصفات في القرآن الكريم في مدونة قرآنية مذيلة:

²⁶ - سلوى حمادة، "تحليل للجملة العربية مبني على مدونة نصية من النصوص المعاصرة"، المؤتمر العاشر لجمعية هندسة اللغة، القاهرة، 15-16 ديسمبر، 2010.



شكل 1: نتائج الحصر الصرفي لسمة الصفات في مدونة قرآنية مذيلة صرفيا

مثال: البحث عن السابقة (س) في القرآن الكريم في مدونة قرآنية مذيلة:



شكل 2: نتائج الحصر الصرفي عن السابقة (س) في مدونة قرآنية مذيلة صرفيا.

البحث باستخدام الجذر:

مثال: البحث على مستوى الجذر، فمثلا نطلب من البرنامج إيجاد كل الكلمات التي تنتمي للجذر "درس" في القرآن الكريم. أنظر فهرس (3) إظهار مواضع مشتقات كتب في القرآن الكريم في سياقاتها.

البحث النكي:

ويقصد به قراءة معجمية صرفية للنص واستخلاص العديد من المعلومات مثل:

- إجراء إحصائيات عن الكلمات والتراكيب والجمل. أنظر فهرس (4)؛
- بحث عن كلمات سوابق ولواحق كلمة: هل "ماذا" يليها اسم أم فعل أم أيهما؛
- إظهار كل ورود للكلمة (المؤمن) في المدونة؛
- استخدام المفرد والجمع لكلمة "الصحيفة" و"الصحف" أيهما أكثر استعمالاً؛
- كتابة الأعداد في صورتها الرقمية والحرفية والمقارنة بين استخداماتها؛
- استخدام كلمة البحث كجزء من اسم مركب؛
- كلمات شائعة تلي كلمة البحث؛
- كلمات شائعة تسبق كلمة البحث؛

أما بالنسبة للمدونات المذيلة صرفياً ونحوياً، فإن كلمة البحث تبحث في:

- العبارات الثابتة؛
- بناء العبارة الفعلية؛
- بناء الفعل؛
- إظهار كل محذوفات المفعول به المباشر؛
- إظهار كل الجمل التي بها حذف؛
- إجابة الأسئلة مثل:
- هل يمكن أن تستعمل الكلمة (ذهب) كفعل؟.
- كلمة البحث في الحس المجازي؛
- كيف يظهر حرف (ب: الباء) كحرف جر، وكيف يظهر كرابط؟
- صور العبارات التي يمكن أن تحدث بين الفقرة اسمية والضمير العائد عليها؟
- هل هناك جمل بلا مكمل ؟

أجزاء تحت التنفيذ:

- تطبيق عمل البحث على اللغة الإنجليزية ولغات أخرى.
- بالنسبة للمدونات المذيلة دلاليًا:

- استخدام كلمة البحث كاسم - بمعانيه المختلفة (مشترك لفظي)؛
- البحث عن المضادات لكلمة البحث.

البحث النكي واستخلاص المعلومات:

يحتاج هذا البحث لدراسات أسلوبية وعمل علاقات بين الكلمات والجمل ستساعد فيها الآلية في مراحلها السابقة:

- نوعية الضمائر المستخدمة قد توضح شخصية كاتب النص مثل كثرة استخدامه لضمير (أنا)؛
- زمن الأفعال السائد في نصه مثل زمن الفعل الماضي في كتابات طه حسين؛
- تحديد ملامح كتابة كاتب ما من خلال تحليل أكثر من نص له.

سمات أخرى:

- تحديد الكلمات التي تتساوى أو تتقارب في التكرار مثل الإسلام والإرهاب في خطب الرئيس الأمريكي السابق جورج بوش؛
- تحديد صور الجمل التي تحتوي مفهوماً أو مصطلحاً معيناً، وتحديد ما يسبقه وما يليه من كلمات، حتى يمكن التنبؤ بأسماء المصطلحات دون معرفة مسبقة بها: فمثلاً في النشرات الطبية نبحث عن كل الجمل التي تتحدث عن مرض يصيب الإنسان، وتقسم الجملة إلى (سابقة ونرمز لها بـ (سا) و(اسم المرض) ولاحقة ونرمز لها بـ (لا). ونخزن كل صور الجمل الممكنة، وللبحث عن مرض جديد غير موجود في البيانات نبحث عن: (سا) ، أو (لا) ، أو (سا) ، لا. حتى نتأكد من الصورة التي يظهر بها اسم المرض في الجملة.

المقارنة بين النصوص:

عند وجود أكثر من نص فإنه يساعد على استخلاص المعلومات عنها بطريقة سريعة. ونطمح مستقبلاً أن نزوده بأداة للتقييم والمقارنة بين نصين أو أكثر.

كيف يقدم صقر هذه الخدمات:

باستخدام آلية صقر يمكن التنقيب في النص لاستخلاص المعلومات والتنبؤ بما يحتمل حدوثه فهو:

يوفر العديد من أنواع الترددات:

بالإضافة للكلمات الفردية، فإن صقر قادر على حساب ترددات العبارات وشكل الكلمات. ويقيم برنامجنا العبارات بصرف النظر عن علامات الترقيم أو المسافات. فترددات الكلمة المماثلة إذن تقاس كما تقاس أيضاً ترددات العبارات المماثلة، بغض النظر عن سطور الدخول أو المواقع المختلفة للفاصلة. إضافة إلى ذلك فإنه يمكنك أن تحدد حروف النهاية للجمل (وافترضنا النهايات هي: النقطة وعلامات الاستفهام وعلامات التعجب). ونعني بشكل الكلمة هنا أي تتابع للحروف في الكلمة. فهذا النمط في البحث المتقدم يسمح باكتشاف الكلمة المستعملة كثيراً حتى عندما تتخذ أشكالاً مختلفة في النص، ومثال ذلك: - اختلاف اللاحقة و/ أو البادئة (جميع المشتقات).

يوفر صوراً مختلفة لتقديم النتيجة:

- بعد مسح النص، يعطيك صقر نتائج في شكل قائمة مصنفة. هذه القائمة يمكن أن تظهر بصور إحصائية مختلفة وتظهر:
- بترتيب الورد في النص؛
 - بالترتيب الأبجدي؛
 - بترتيب تصاعدي أو تنازلي اعتماداً على قيمة تكرار الوحدة اللغوية؛
 - اعتماداً على نوع النص ، اسم المؤلف، سنة الكتابة،....

بالإضافة إلى ذلك، يمكن أن تبحث في النتائج عن نمط معين، يميز حدوث الكلمة أو العبارة التي نبحث عنها في النص الأصلي.

يوفر خيارات متعددة للبحث:

يضم صقر تشكيلة واسعة من الخيارات الشكلية التي تساعد في الحصول على نتائج معينة، فيمكن أن تحدد قائمة من كلمات لإهمالها وطول عبارة السياق المطلوب البحث فيها. وفي الوقت نفسه تسمح الخيارات المبدئية بعمل المهام الأكثر شيوعاً.

8. الخاتمة:

حاول هذا البحث توضيح أهمية تقنية المعلومات في خدمة النص القرآني من خلال التعريف بمجال علم اللغويات الحاسوبية ومجالاته التطبيقية، مع تبيان الجوانب الأساسية في القرآن الكريم التي يمكن أن تستفيد من المعالجة الحاسوبية. كما بين أن الاشتغال على التطبيقات الحاسوبية للقرآن الكريم مايزال مجالاً بكرًا، يحتاج إلى بحوث ودراسات معمقة، لا بد أن تتم بتضافر بين الخبراء في المعلوماتية والخبراء في اللغويات. ولعل وضع تصور لآلية تساعد على فحص النصوص بجميع صورها الإلكترونية وأوعيتها التخزينية يمكن أن يمثل تطبيقاً في هذا المضمار. إذ يشرح كيف يمكن أن تسهم هذه الآلية في تحليل النصوص العربية عامة، والنص القرآني خاصة، ولقد تم التطبيق على النص القرآني أولاً، وعلى مدونة لغوية مذيبة صنعت خصيصاً لغرض إيضاح الحصر النحوي والأسلوبي ثانياً. كما تم التطبيق أيضاً على مدونة قرآنية مذيبة لإثبات صحة الفكرة وقيمتها. وقد أرفقت فهارس لتوضيح أنواع مختلفة من البحث على مدونة القرآن الكريم في نهاية البحث.

9. نتائج الدراسة:

- إن الأبحاث اللسانية والحاسوبية الموجهة لخدمة القرآن الكريم يجب أن تتم بالتعاون بين الخبراء في الدراسات الإسلامية والخبراء في الدراسات اللغوية الرقمية و الخبراء في علوم الحاسوب والهندسة؛
- من الضروري للمختصين في الدراسات الإسلامية استخدام التقانة الحديثة خدمة للقرآن، إن أرادوا بناء مجتمع معرفة إسلامي؛
- إن تطوير برمجيات إسلامية لا ينبغي أن يكون حكراً على الفنيين، بل لابد من أن يوكل الأمر إلى ذوي الخبرة والعلم والمعرفة بعلوم الدين وعلوم اللسان وعلوم الحاسوب؛
- إن دور المعجم العربي بنوعيه الورقي والرقمي في تعزيز الدراسات القرآنية ودعمها لا يحتاج إلى دليل، على اعتبار أن المعجم أداة معرفية مشتركة بين العلوم والمعارف.

10. المصادر والمراجع

[1]

جميل أحمد إطميزي، "تقنيات التعليم الإلكتروني وأدواته في خدمة القرآن الكريم بين النظرية والتطبيق"، جمعية الاتصالات العربية للحاسب، المجلد الأول، 2011.

[2]

حسين البسومي، المعالجة الحاسوبية للنص القرآني مفهومها وأهدافها،
dspace.medi.u.edu.my:8181/xmlui/.../012.doc?...1

[3]

محمد سالم سعد الله، علم اللغة الحاسوبي وحوسبة القرآن الكريم، جامعة مالايا / ماليزيا، المؤتمر القرآني الدولي السنوي الثاني، 22-23 فبراير 2012.

[4]

حسين محمد علي البسومي، قاعدة بيانات معجمية لألفاظ القرآن الكريم وتطبيقاتها، ندوة القرآن الكريم والتقنيات المعاصرة (تقنية المعلومات)، صص 4-51.

[5]

محمد حماد، "صور الجملة القرآني في أدب نجيب محفوظ"، كلية دار العلوم ، 1992.

[6]

ريما سعد الجرف، "آلية اللغات والترجمة، جامعة الملك سعود"، الجمعية الدولية للمترجمين العرب، 2008.

[7]

سلوى حمادة، "المعالجة الآلية للغة العربية: المشاكل والحلول"، دار غريب للطباعة والنشر، القاهرة، 2009.

[8]

سلوى حمادة " المدونات اللغوية والمفهرسات الآلية :عرض مفهرس صرفي لمدونة للغة العربية المعاصرة "،

[9]

The 4th International Conference on Computer Science Practice in Arabic (CSPA '08), Doha, Qatar, 2008, <http://www.phillips-publishing.com/cspa08/>

[10]

سلوى حمادة، محمود السابيس، " دعم تعلم وترجمة القرآن بالإبحار في قاعدة بيانات القرآن الكريم بآلية مجانية على شبكات الإنترنت "، ندوة القرآن الكريم والتقنيات المعاصرة، المدينة المنورة، 2009.

[11]

سلوى حمادة، "تحليل للجملة العربية مبني على مدونة نصية من النصوص المعاصرة"، المؤتمر العاشر لجمعية هندسة اللغة، 15-16 ديسمبر، 2010، القاهرة.

[12]

Martin Wynne , "Developing Linguistic Corpora: a Guide to Good Practice 2004", [مارتن، 2004]: <http://www.ahds.ac.uk/creating/guides/linguistic-corpora/index.htm>

[13]

نبيل علي، "اللغة العربية والحاسوب"، مجلة عالم الفكر، 2000.

[14]

حسين محمد علي البسومي، قاعدة بيانات معجمية لألفاظ القرآن الكريم وتطبيقاتها، ندوة القرآن الكريم والتقنيات المعاصرة (تقنية المعلومات)، صص 4-51.

[15]

مصطفى جرار، نحو تأصيل منهجي لبناء أنطولوجيا اللغة العربية/ArabicOntology، <http://sites.birzeit.edu/comp/>

[16]

محمد سالم سعد الله، علم اللغة الحاسوبي وحوسبة القرآن الكريم، جامعة مالايا / ماليزيا، المؤتمر القرآني الدولي السنوي الثاني، 22-23 فبراير 2012.

[17]

مهديوي عمر، خوارزميات صرفية للغة العربية من التحليل إلى التوليد، المؤتمر السنوي الخامس لمجمع اللغة العربية بدمشق نوفمبر 2006.

[18]

مهديوي وسلوى حمادة، المعالجة الدلالية الآلية للغة العربية، مجلة فكر ونقد، عدد 82، أكتوبر 2006.

[19]

مهديوي عمر، المعجم العربي في ضوء اللسانيات الحاسوبية، الندوة الدولية الأولى عن الحاسب واللغة العربية، الرياض، 2007.

[20]

مهديوي، رقمنة المعاجم العربية: قاعدة معطيات المعجم التاريخي نموذجاً، المؤتمر الدولي الرابع للمجلس الدولي للغة العربية، دبي الإمارات العربية المتحدة، 6-10 ماي، 2015.

[21]

محمد زكي خضر، مجموعة مقالات ودراسات منشورة على الرابط:

http://www.al-mishkat.com/khedher/?page_id=17:

[22]

مركز أبحاث تقنية المعلومات لخدمة القرآن الكريم وعلومه (نور) <http://nooritc.org/?q=ar/project/details/112>

11. جدول الألفاظ الأجنبية ومقابلاتها العربية:

English	عربي
Corpus Linguistics	علم لغة المدونة أو لغويات المدونة
Computational Linguistics	اللغويات الحاسوبية
Forms of Arabic	صور اللغة العربية
Traditional Arabic Language	اللغة العربية التقليدية
Modern Standard Arabic	اللغة العربية الفصحى
Hybrid Modern Standard Arabic	اللغة العربية المختلطة
Statistical Linguistics	اللغويات الاحصائية
Linguistic corpus	مدونة لغوية
Point-wise Mutual Information	معامل الارتباط
POS (part-of-speech tagging)	أقسام الكلام
Multi-functionality	متعدد الوظيفية
Lexicography	صناعة قاموسية أو صناعة معجمية
frequency lists	قوائم التردد
speech synthesis	تأليف الخطاب

Different kinds of annotation	أنواع التذييل المختلفة
Phonetic annotation	التذييل الصوتي
spoken corpus	المدونة المنطوقة
Tokenization	التقسيم
Lemmatisation	التجذيع
Lemma	(لما) أصغر عنصر للكلمة
Phrases	العبارات أو الجمل
Clauses	شبه الجمل
Disambiguation /sense resolution	حل لبس المعنى
pragmatic annotation	التذييل الواقعي أو التداولي
Discourse annotation	تذييل الخطاب
Anaphoric	مرجعية الضمير
Antecedent	سابق
Stylistic annotation	تذييل الأسلوب
Dysfluency	تذييل الفصاحة
excessive use	الاستعمال المفرط
most frequent words co-occurring	الكلمات الأكثر تكرارا
immediate context	السياق الآني
hard measurable evidence	شواهد لغوية
Variation and consistency	الاختلاف والاتساق

12. الخلاصة باللغة الانجليزية

Lexicon Arab support and Quranic studies computer

Abstract: This paper envisages definition of Computational Linguistics and its applications in the language texts General natural treatment, and religious texts in particular (the Quranic text mode), since the information technology allows a number of image concepts such as automation and algorithms and database, and knowledge base, analysis and generation robots, and other. All of these concepts help to analyze the data accurately, and to reach substantive results, and then enter the access information to the Quranic mind and retrieval machine according to a sophisticated and intelligent.

It also envisages to demonstrate the importance of the Arab lexicon of both types paper and electronic support in Quranic studies, on the grounds that what has been achieved. The time limit in this area is still a little bit, and even if any, it does not accommodate the Quranic text miraculous levels and rude to us.

For this we will try in this study to clarify the key areas in the Quranic text that could benefit from computer applications, and that will help to develop and implement computer software, especially Islamic Qur'an, and there is the relentless pursuit towards building a glossary of computer of the Koran, addressed to Muslims and non-Muslims, helping them all to understand the religious text linguistic and stylistic characteristics and morphological and grammatical.

Keywords: computational linguistics, religious texts, the Koran, the database, algorithms, computer application language. Quran corpus.