



منقب ومفهرس اللغة العربية ودوره في تطوير الدراسات اللغوية والقرآنية

سلوى حمادة

معهد بحوث الإلكترونيات - القاهرة - الدقي

hesalwa@hotmail.com

الخلاصة: يقدم البحث آلية تساعد الباحثين على إيجاد ضالتهم من الوحدات اللغوية داخل النص اللغوي والقرآني وتوفر لهم مستويات متنوعة من البحث أيضا، وتتكون من أدوات تستخلص معلومات من النص العربي بصفة عامة والنص القرآني بصفة خاصة. ونقترح - في الوقت ذاته - تطبيقات لاستخدام هذه الأدوات لتوضيح أهميتها وأسباب الحاجة إليها، وذلك حتى نتمكن من أن ننقل من طور استهلاك التقنية إلى طور إنتاج التقنية، بل وتصديرها. وينتهي البحث بالتقييم والتقييم هنا لن يكون لقيمة آلية البحث من هذه النوعية وفائدتها - والتي لا جدال عليها - لكنه متعلق بالناحية الفنية في البرمجة من حيث السرعة واستيعاب أكبر حجم من المعلومات وسهولة الاستخدام وخلافه.

الكلمات الجوهرية: المدونات - مستويات التحليل اللغوي - البحث في القرآن - الإحصاء - الفهرسة - قوائم الكلمات.

1 المقدمة

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 وقطعت شوطا لا بأس به في إطار 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 جلبي 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 قطعت شوطا لا بأس به في مرحلة البحث الصرفي والإحصائي والتقييم 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.

2 الحاجة لبناء المدونات

من خلال تفحص الكثير من الأبحاث في جميع المجالات العلمية يمكن الجزم بأن الأبحاث التي لا تنطلق من خلال مدونة تمثل المجال الذي تطبق فيه تصبح أبحاثاً ضعيفة وغير موثقة. والمدونة بصورتها الخام لها أهمية كبيرة أيضاً ويمكن استخدامها في عمليات البحث عن المعلومات وإحصاء العديد من الكلمات والمركبات والتعبيرات والجمل. أما ظهور علم لغة المدونة -الذي يشمل ترميز / تذييل المدونات وفحصاً بعدة طرق- فقد زودنا بأدوات مدهشة لتحري أحجام كبيرة من النصوص لكي نجد الأجوبة على الأسئلة المختلفة عن استعمال اللغة. كان من الصعب الإجابة عن هذه الأسئلة من قبل؛ حيث كان لا بد للغويين حينئذ أن يعتمدوا على سعة ذاكرتهم البشرية وإمكاناتهم المحدودة لإنهاء أبحاثهم لأن تقنية الحاسبات كانت محدودة. وقد اهتمت الدول المتقدمة - في المجال اللغوي والتقنيات اللغوية الحديثة كالتحليل الآلي للنصوص والترجمة الآلية والاختصار الآلي وغيرها من التطبيقات - بالمدونات والتدوين اللغوي منذ أكثر من ستين عاماً.

3 المدونة القرآنية كمدونة للغة العربية

هناك ما يزيد على 550 مليون شخص يتحدثون العربية في العالم. وتعتبر اللغة العربية من اللغات الرسمية الأولى في أكثر من ثلاثين دولة. وعلى الرغم من ذلك فهناك عدد محدود من متحدثي اللغة الفصحى سواء الكلاسيكية القديمة أو الحديثة، فهي تدرس فقط بالمدارس أو الكليات المتخصصة. ومتحدثو العربية يستخدمون اللهجات العربية جنباً إلى جنب مع اللغة المعاصرة. ويطلق بعض اللغويين على اللغة العربية المعاصرة "لغة العقل"، وعلى اللهجات العربية المحلية "لغة القلب". وتتأثر اللغة بعوامل شتى منها؛ أسلوب الحياة، والبيئة، والمستوى الاجتماعي، والتعليم، وربما المستوى الديني أيضاً؛ كذلك تتأثر بنوع الإنسان سواء أكان ذكراً أم أنثى.

ومن صور اللغة العربية:

- اللغة التقليدية الأصلية: وهي لغة الجزيرة العربية قبل الإسلام.
 - اللغة العربية الفصحى: وهي اللغة المستخدمة في كتب اللغة والآداب.
 - اللغة العامية: وهي لغة الكلام، وهي تختلف من مدينة إلى أخرى على مستوى البلد، ومن بلد عربي لآخر.
 - اللغة العربية المختلطة: وهي خليط من اللغة العربية الفصحى واللغة العامية، وهي لغة الإعلام والجراند اليومية.
- أما لغة القرآن الكريم فهي لغة عربية فصيحة، معجزتها في البناء اللغوي على كل مستوياته.

4 البحث في المدونات المرمزة/ المذيلة

أ. البحث في اللغة العربية ومشاكله

تم شرح هذا الجزء بالتفصيل في [سلوى حمادة، محمود الساييس، 2009].

ب. نظم البحث في القرآن الكريم المتاحة على شبكات الاتصال

تم التعرض إلى عديد من هذه النظم في تم شرح هذا الجزء بالتفصيل في [سلوى حمادة، محمود الساييس، 2009] ، [جميل، 2011].

ج. بناء المدونات: راجع [سلوى حمادة، 2008].

د. المدونات المذيلة

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 كتب أو اسما لمجموعة كتب أو فعلا بمعنى أرغمه على الكتابة. هذه الكلمات لها الشكل نفسه بينما علامات التشكيل تعطي أكثر من 12 معنى [سلوى حمادة، 2006] حيث إن علامات التشكيل في اللغة العربية تمثل جزءا من الكلمة يتم إهماله في الكتابة غالبا. 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، و في الإعراب، وفي حصر قوائم التردد، وفي تأليف الخطاب، وفي العديد من التطبيقات الأخرى.

والاستعمالات المستقبلية للمدونات المذيلة تتغير دائما عن الاستعمالات الأصلية بطريقة لا يمكن توقعها [مارتن، 2004]. وتلعب المدونات اللغوية المذيلة دورا مهما في الإحصاء اللغوي الحاسوبي، أو ما يعرف بعلم اللغة الإحصائي.

أنواع التذييل المختلفة

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

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.

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 بعضا عن طريق المسافات بينها. وهذا الأمر قد يكون صحيحا في حالة (الكلمة) ولكن هناك وحدات معجمية متعددة الكلمات لا تنطبق عليها وجهة النظر هذه مثل التعبيرات الاصطلاحية 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 الأساسي للكلمة، مثلما يوجد مع الكلمة كالكلمة الرئيسية في القاموس (ومثال ذلك: - كتب، سيكاتيون).

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 وشبه الجمل.

التذييل الدلالي: هو معلومات مضافة حول التصنيف الدلالي للكلمات مثال الاسمي: "لعبة الفوت بول" والمصطلحي: "الرياضة" ويرجع إلى الأصناف الدلالية المختلفة، على الرغم من أنه ليس هناك اختلاف في التهجى أو التلفظ. كما يفيد ترميز المعنى الكلمة في (حل لبس المعنى) [مارتن، 2004].

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

تذييل الخطاب: هو معلومات مضافة حول مرجعية الضمائر في النص [مارتن، 2004]. فعلى سبيل المثال مرجعية الضمير "هم" على سابق في الجملة "يعلم المدرس التلاميذ ويفهمهم الدرس" وهو " التلاميذ"، وتنفذ هذه المهمة غالباً من قبل المحللين البشر للحصول على دقة عالية. وترى سلوى بعض الحلول الآلية لهذه المشكلة والتي تصل بالدقة لدرجة معقولة [سلوى حمادة، 2006].

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

أنواع أخرى [مارتن، 2004]: من المحتمل ابتداء أنواع هائلة من التذييل التي قد تكون مفيدة لأنواع معينة من البحث. ومثال ذلك تذييل الفصاحة. فمن يعمل على البيانات المنطوقة قد يرغب في تذييل مدونة الخطاب التلقائي بأنواع من 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 ؟ وما أنواع هذه الأخطاء؟. وتستخدم هذه المدونات في دعم تعليم اللغة لتلافي هذه الأخطاء أو اختبار المتعلمين الجدد. ولمزيد من التفاصيل راجع [سلوى حمادة، 2006]، [نبيل علي، 2000].

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:

أ. الحصر الإحصائي الشكلي

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

حصر تكرار الكلمات.	حصر المفردات الأكثر شيوعاً.
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.	

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

ب. الحصر من خلال قوائم أو معاجم

ويمكن أن تستخدم أيضاً 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 الإصطلاحية.
- المصطلحات العلمية في مجالات متعددة.
- الكلمات المشهورة أو أي قوائم يبحث عنها المستخدم.

وعلى ذلك يمكن:

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 شيوعاً، وكذلك الأضداد والمتضادات، وجميع الظواهر اللغوية الأخرى.
حصر تكرار مصاحبات لغوية.	

الحصر اللغوي

وتقوم فكرته الأساسية على إيجاد علاقة بين المفردات المدرجة اعتماداً على المتصاحبات اللغوية للمفردات الرئيسية، وذلك من خلال علم اللغة الإحصائي اللغوي وتطبيقه على المدونات المذيلة.

والبرنامج يتصل بقاعدة بيانات تنقسم إلى ثلاثة أقسام:

- القاعدة البنوية: وتضم معلومات البنية والاشتقاق الخاصة بالمفردات.
 - القاعدة التركيبية: وفيها حصر للأنماط التركيبية الواردة في القاعدة الأساسية.
 - القاعدة الدلالية: ويرتكز هذا القسم على الربط بين المفردات الرئيسة ومتصاحباتها اللغوية.
- وقد اعتمد بحث سابق [سلوى حمادة، 2010] في بناء قاعدة البيانات الأساسية على مدونة لغوية مستمدة من لغة الصحافة المعاصرة، بالإضافة إلى مجموعة من الكتب والمؤلفات العربية المعاصرة في شتى العلوم والمجالات. حيث يتم حصر تكرار المتصاحبات (تركيب أكثر من كلمة والبحث في مدى التصاحب بينهما)⁽¹⁾. عن طريق معادلة (PMI) والتي تختص بقياس الترابط بين كلمتين أو أكثر - اعتماداً على مدى التصاحب- من خلال حدوث كل منهما على حدة وحدوثهما معاً. ويعرض فهرس (2) الكلمات المتصاحبة وعدد الكلمات المتصاحبة ومدى التصاحب = 2.

(1) يمكن الاستفادة من هذه الوظيفة في استخراج أكثر المتلازمات والتعابير المسكوكة شيوعاً 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.

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.

الحصر التركيبي

ويطبق على مدونات مرمزة صرفيا وتركيبيا ونحويا ومن أمثلتها بنوك الأشجار العربية السابق الإشارة إليها في الجزء السابقة، والمدونة التي قامت الباحثة بالإشراف وتحليلها وتدعى "مساك" [سلوى حمادة، 2010] 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.
- ولا سبيل لحصر العمل الإحصائي أبداً.

6 التطبيق

أ. 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.
 - تجميع النصوص على ملف واحد للبحث فيه.
- راجع [سلوى حمادة، 2010] وسيطور البرنامج فيما بعد للبحث في النصوص الموجودة على الشبكات والتي نختارها بكلمات مفتاحية محددة.

ب. متطلبات البحث

- عمل التحديدات لو تطلب البحث ذلك كأن نقوم بتحديد الكلمات المهمة أو الكلمات المطلوب البحث عنها فقط ويمكن أيضاً أن نحدد نهايات الجمل إذا تطلب البحث عن عدد الجمل.

- عمل قاعدة المترادفات: أن المفهوم الواحد قد يمثل بكلمات مختلفة كثيرة (مترادفات)، وفي هذه الحالة نخزن قاعدة بيانات للمترادفات وعند طلب البحث عن أي وحدة لغوية يسأل البرنامج (بحث مطابق أم عرض المترادفات أيضاً).
- عمل قاعدة المصاحبات: أن 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).

ت. صقر والفهرسة الآلية في اللغة العربية بناء على تكرار تردد المفردات:-

يعمل المفهرس صقر 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 تنجم 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 مؤيديها ومعارضين في خطوتها لإدخال الحركات في معالجة الكلمات العربية وانعكس هذا الانقسام على تطوير خوارزميات التحليل الصرفي في اللغة العربية التي تدعم الفهرسة. وقد كان استرجاع المعلومات سبباً من أسباب تطوير هذه الخوارزميات ولكنه كان الهدف الأساسي لتطوير أساليب العربية. وأجرينا بعض التجارب في هذا السياق التي تهدف إلى دراسة الفهرسة واسترجاع البيانات باللغة العربية وخلصت هذه الدراسة إلى أن استعمال جذور الكلمات وجذوع الكلمات يسفر عن نتائج أفضل.

7 أهمية آلية صقر

لكي نجيب عن هذا السؤال يجب أن نحدد من يستخدم فاحص النص صقر؟ إن برنامج صقر المقترح سيفيد كل من يعمل في كتابة النصوص حيث يحلل هذا البرنامج محتوى النص ويزودهم بقائمة أو قاموس من الكلمات، العبارات، وتردها في الاستعمال ضمن محتويات النص. وهذه المعلومات التي ينتجها البرنامج تمكن من التدقيق من الاستعمال المفرط للكلمات أو تكرار الكلمات أو العبارات ضمن أية وثيقة، وهو ما يفيد:

أ. المعجميين

المساهمات المقترحة للبحث الإحصائي بصقر في المدونات لدعم المعجم الحاسوبي المعاصر بما فيه من الألفاظ القرآنية:

- إنشاء معجم آلي يضم السمات اللغوية على جميع المستويات انظر [سلوى حمادة، 2006]. وهذا بدوره يؤدي إلى سهولة فك اللبس انظر [سلوى حمادة، 2006].
- 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 فكه.
- تحديد التعبيرات الاصطلاحية واستخدامها ومعانيها من خلال السياق. وقد اكتشفنا تنوع المعاني واكتسابها معاني جديدة من خلال السياق - بحث -.
- دعم التصنيف الدلالي للمجالات.
- تدعيم المداخل معجمية الموجودة للمعجم العربي بسمات وخصائص ودلالات جديدة من المعلومات الإضافية التي استخلصت عن طريق تحليل المدونة (ومثال على ذلك: - أشكال الأكثر شيوعاً، التضمين ، الخ. انظر إثراء المعاجم العربية بالعلاقات والسمات [سلوى حمادة، 2006].
- مد المعجم بأمثلة من اللغة الحية متنوعة وفي جميع المجالات.
- دعم المعاني وشرحها - شواهد لغوية - يمكن استخدامها في تعليم اللغة لابن اللغة والأجنبي عنها.
- يوفر أداة للمقارنات بين اللغات.
- وبطبيعة الحال يمكن عمل كل هذه التطبيقات على النص القرآني.

ب. 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.

ت. التعليم

يفيد في التعليم والعديد من الدراسات مع جميع أنواع النصوص مثل:

النص القرآني

للبحث عن ظواهر الإعجاز اللغوي والعدي وغير ذلك من الأبحاث في هذا المجال.

النص اللغوي والقرآني

- البحث عن المفردات القرآنية في النصوص العربية الحديثة لعدة أسباب:
- 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 ملهما عاما مطردا في جميع أعماله أم كان خاصا برواية أو مرحلة معينة[حماد،1992]. 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 جدا للجملة العربية وذلك من خلال التحويرات التي يجريها الكاتب على الجملة القرآنية سواء من الناحية الشكلية أو التركيبية [حماد،1992]. 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 معيارا أو منطلقا لتدريبات أخرى- [حماد،1992] والعكس.
- والعودة للغة الفصحى في الكتابات العربية من خلال النمط القرآني والمحافظة عليه.

النصوص اللغوية

- دراسة مواد القراءة الإضافية كالقراءة القرآنية وقصص الأطفال وكتب المتابعة لتحديد صلاحيتها من جهة والمرحلة التي تصلح لها من جهة أخرى [ريما،2008].
- وضع الاختبارات لتحديد المستوى اللغوي للممتحنين بشكل عام ولثروتهم من المفردات بشكل خاص، مما يساعد على وضع الدارسين في المستوى الذي يناسبهم [ريما،2008].
- تحديد المترادفات الأكثر شيوعا، وكذلك الأضداد والمتضادات، وجميع الظواهر اللغوية الأخرى.

8 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 لكلمة 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.

ت. البحث في المدونات المذيلة/ المرمزة بأجزاء الكلام

تم عمل مدونة تمثل أغلب النصوص العربية وتحتوي أكثر من نصف مليون كلمة [سلوى حمادة، 2010]. وتم عمل تذييل 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 التذييل بأجزاء الكلام على سبيل المثال ما يلي:

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

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

المتعلقات اللغوية		
عودة للقائمة الرئيسية		
عودة للقائمة الرئيسية	الصفات	حدد نوع المشتق
البحث		
1	الفتاححة	الرحمن الرحيم
3	الفتاححة	الرحمن الرحيم
37	البقرة	التواب الرحيم
54	البقرة	التواب الرحيم
127	البقرة	المصمغ العليم
128	البقرة	التواب الرحيم
137	البقرة	المصمغ العليم
143	البقرة	لوقوف رحيم
160	البقرة	التواب الرحيم
163	البقرة	الرحمن الرحيم
173	البقرة	غفور رحيم
181	البقرة	مهيمن عليم
182	البقرة	غفور رحيم
192	البقرة	غفور رحيم
199	البقرة	غفور رحيم
218	البقرة	غفور رحيم
224	البقرة	مهيمن عليم
225	البقرة	غفور رحيم
226	البقرة	غفور رحيم
227	البقرة	مهيمن عليم
235	البقرة	غفور رحيم
244	البقرة	مهيمن عليم
256	البقرة	مهيمن عليم
276	البقرة	كفار أثيم
31	آل عمران	غفور رحيم
34	آل عمران	مهيمن عليم
35	آل عمران	المصمغ العليم
89	آل عمران	غفور رحيم

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

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

بحث بدلالة معلومات صرفية

عودة للقائمة الرئيسية

ابداً البحث

من سورة: الفاتحة إلى سورة: الناس

● بحث بدلالة السابق من ● بحث بدلالة اللاحق

● بدلالة الجذر ● بدلالة الوزن

الكلمة	سورة	رقم الآية	الجذر	نوع المشتق	الوزن	السابق	لاحق ١	لاحق ٢	لاحق ٣
سبحون	البقرة	١٤٢	قول	فعل مضارع مجرد	يُتَعَلَّل	س	حُ		
مستكبرون	البقرة	٢٣٥	ذكر	فعل مضارع مجرد	يُتَعَلَّل	س	حُ		
ستغفلون	آل عمران	١٢	غلب	مضارع مجزول مجرد	يُتَعَلَّل	س	حُ		
سيطوفون	آل عمران	١٨٠	طوف	مضارع مجزول مزيد	يُتَعَلَّل	س	حُ		
سكتك	آل عمران	١٨١	كتب	فعل مضارع مجرد	يُتَعَلَّل	س	حُ		
سؤيتهم	النساء	١٦٢	أي	فعل مضارع مزيد	يُتَعَلَّل	س	حُ		
سبحون	الأأنام	١٢٠	جزي	فعل مضارع مجرد	يُتَعَلَّل	س	حُ		
سبحون	الأأنام	١٣٩	جزي	فعل مضارع مجرد	يُتَعَلَّل	س	حُ		
سبحون	الأأنام	١٤٨	قول	فعل مضارع مجرد	يُتَعَلَّل	س	حُ		
سأريكم	الأعراف	١٤٥	رأي	فعل مضارع مزيد	يُتَعَلَّل	س	حُ		
سأصرف	الأعراف	١٤٦	صرف	فعل مضارع مجرد	يُتَعَلَّل	س	حُ		
سينالهم	الأعراف	١٥٢	نول	فعل مضارع مجرد	يُتَعَلَّل	س	حُ		
سنزيد	الأعراف	١٦١	زيد	فعل مضارع مجرد	يُتَعَلَّل	س	حُ		
سبحون	الأعراف	١٨٠	جزي	فعل مضارع مجرد	يُتَعَلَّل	س	حُ		
سألني	الأأنام	١٢	لني	فعل مضارع مزيد	يُتَعَلَّل	س	حُ		
سؤيتنا	التوبة	٥٩	أي	فعل مضارع مزيد	يُتَعَلَّل	س	حُ		
سيعيب	التوبة	٩٠	صوب	فعل مضارع مزيد	يُتَعَلَّل	س	حُ		
سيدخلهم	التوبة	٩٦	دخل	فعل مضارع مزيد	يُتَعَلَّل	س	حُ		

نتائج البحث ٤٣

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

البحث باستخدام الجذر

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

ث. البحث الذكي

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

- عمل إحصائيات عن الكلمات والتراكيب والجمل انظر 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 الاستخدام.
 - استخدام كلمة البحث كجزء من اسم مركب.
 - كلمات شائعة تلي كلمة البحث .
 - كلمات شائعة تسبق كلمة البحث .
- أما بالنسبة للمدونات المذيلة صرفياً ونحوها
- كلمة البحث في العبارات الثابتة.
 - كلمة البحث في بناء العبارة الفعلية.
 - كلمة البحث في 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 (مشارك لفظي).
- البحث عن المضادات لكلمة البحث.

ج. البحث الذكي 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 خطاب الرئيس الأمريكي السابق جورج بوش.
- تحديد صور الجمل التي تحتوي مفهوم أو مصطلح معين وتحديد ما يسبقه وما يليه من كلمات حتى يمكن التنبؤ بأسماء المصطلحات دون معرفة مسبقة لها: فمثلا في النشرات الطبية نبحث عن كل الجمل التي تتحدث عن مرض يصيب الإنسان وتقسم الجملة إلى (سابقة ونرمز لها ب (سا) و(اسم المرض) ولاحقة ونرمز لها ب (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. ونطمع مستقبلا أن نزوده بأداة للتقييم والمقارنة بين نصين أو أكثر.

9 كيف يقدم صقر هذه الخدمات:

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

أ. 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 أيضا ترددات العبارات المماثلة، بغض النظر عن سطور الدخل أو المواقع المختلفة للفاصلة. إضافة إلى ذلك فإنه يمكنك أن تحدد حروف النهاية للجمل (وافترضيا 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 كثيرا حتى عندما تتخذ أشكالا مختلفة في النص، ومثال ذلك: - اختلاف اللاحقة و/ أو البادئة (جميع المشتقات).

ب. يوفر صورا 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 اعتمادا 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.

ت. 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 شيوعا.

10 الخاتمة

حاول هذا البحث مد القارئ الراغب في العمل باستكشاف وفحص النصوص والمدونات بكل ما يلزمه من معلومات. وهناك ندرة هذه النوعية من الأبحاث على اللغة العربية فحاول هذا البحث وضع تصور لآلية تساعد على فحص النصوص بجميع صورها الإلكترونية وأوعيتها التخزينية. وشرح كيف يمكن أن تسهم هذه الآلية في تحليل النصوص العربية والنص القرآني خاصة وتم التطبيق على النص القرآني في البحث. والتطبيق على مدونة لغوية مذيبة صنعت خصيصا لغرض إيضاح الحصر النحوي والأسلوبي. كما تم التطبيق على مدونة قرآنية مذيبة لإثبات صحة الفكرة وقيمتها. وقد أرفقت فهارس لتوضيح أنواعا مختلفة من البحث على مدونة القرآن الكريم في نهاية البحث.

11 المراجع

[جميل،2011]

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

[حماد،1992]

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

[ريما،2008]:

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

[سلوى حمادة،2006]:

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

[سلوى حمادة،2008]:

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

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

[سلوى حمادة، محمود الساييس، 2009]

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

[سلوى حمادة،2010]:

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

[إمارتن،2004]:

Martin Wynne , "Developing Linguistic Corpora: a Guide to Good Practice", 2004

<http://www.ahds.ac.uk/creating/guides/linguistic-corpora/index.htm>

[نبيل علي،2000]:

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

12 جدول الألفاظ

عربي	English
علم لغة المدونة	corpus linguistic
صور اللغة العربية	Forms of Arabic
اللغة التقليدية الأصلية	Traditional Arabic Language
اللغة العربية الفصيحة	Modern Standard Arabic

Modern Standard Arabic Hybrid	اللغة العربية المختلطة
Statistical Linguistics	علم اللغة الإحصائي اللغوي
Linguistic corpus	مدونة لغوية
Point-wise Mutual Information	معامل الارتباط
part-of-speech tagging (POS)	أقسام الكلام
Multi-functionality	متعدد الوظيفة
lexicography	تأليف القواميس
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
Lemmatization	التجذيع
lemma	(لما) أصغر عنصر للكلمة
phrases	العبارات
clauses	شبه الجمل
sense resolution /Disambiguation	حل لبس المعنى
pragmatic annotation	التذييل الواقعي
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 الآتي
hard measurable evidence	شواهد لغوية
Variation and consistency	الاختلاف والاتساق

13 الخلاصة باللغة الانجليزية

Fast Arabic Language Concordance (FALCON)

Salwa Hamada

hesalwa@hotmail.com

Abstract:

This paper introduces a tool that helps the researchers to find their needs of language units in the texts. It provides them with various levels of search also. It helps to extract knowledge from the text and it was name (FALCON). At the same time it had been proposed some applications to use these tools to clarify their importance and needs. The goal of producing these tools is to move from the phase of consumption of technology to production of technology. The evaluation here of course do not related to the importance of the tool but to its technical quality.

Keywords: Corpora, Search, Statistic, Concordance, Word lists.

الفهارس

فهرس (1) تكوين (معجم النص) والبحث عن الكلمات

بحث عن الكلمة كاملة	بحث ببداية الكلمة
بحث بنهاية الكلمة	بحث بالجذر

الكلمات المفردة (المعجم) وتكرار كل كلمة في النص

[illegible]

عدد سطور الصفحة وعدد الكلمات

شكل 3: نتائج الحصر الشكلي لكلمة (ولا) في مدونة قرآنية مذيلة صرفيا

البحث عن كلمة

1) ولأخذ كلمة "اللغة" مثلاً: باختيار كلمة اللغة من جدول التكرارات قام البرنامج بعرض جميع السياقات التي وردت بها في شاشة العرض موضحاً ما يسبقها وما يليها من كلمات.

اللغة
كل من عن الكلمة ...

بحث

الكلمة	التكرار
في	77
من	72
اللغة	48
أن	34
العربية	30
على	28
بين	21
التراث	18
لا	18
إلى	18
عن	18
هذا	16
اللغوي	16
الفصحى	16
ما	15
التي	15
العامة	14
العربي	13
يمكن	12
و	10
العرب	10
لغة	10
الدراسات	10
اللسانية	10
هذه	9
الحاسب	9
حيث	9

السابق	الكلمة	التالي
استخدام إمكانيات	اللغة	العربية الفصحى ..
منذ فجر التاريخ ...	اللغة	للهدف الأساسي ..
في الكلمة تغير من اللغة	اللغة	العربية الفصحى ..
من أصل معناها ...	اللغة	العربية الفصحى ..
العربية الفصحى ...	اللغة	التي تستخدم في
إلا القلة الأجنبية ..	اللغة	العربية نفسها لقد
حاولت هذه ...	اللغة	العربية ولكن هذه
في تطوير الفكر ...	اللغة	العربية وحدهم هم
أنه لا توجد ...	اللغة	العربية الحديثة ...
القديم أطرف الآراء	اللغة	وطبيعة الألسنية ..
علماء الغرب إلا في	اللغة	الألمان في النصف
العرب هم مستنبط اللغة	اللغة	الحديث ولم يكن ..
هناك بد من تنمية	اللغة	و الحاسوب حيث ..
بين اللغة و ...	اللغة	هي تجسيد جوهري
مجال الترجمة ...	اللغة	العربية ويرمجات
أن يضيف الحاسب	اللغة	علاقة تبادلية حيث
اللغة علاقة تبادلية	اللغة	بمنظور منطقي ...
تكشف القناع عن	اللغة	والقدرة اللغوية ..
وكيفية قيام الذهن	اللغة	واكتساب الخبرات
واسترجاع ...	اللغة	في تطوير أنظمة ..
عطاء تبادلية بين ...	اللغة	وأدى هذا التبادل
مجال التنظير ...	اللغة	الياً ووضعها في ...
وتضيف إليها الكثير	اللغة	والحاسب إلى إقام
نماذج حاسوبية لفهم اللغة	اللغة	العربية الاهتمام ..
الترجمة الآلية ...	اللغة	العربية وعلاقتها ..
تطبيق أساليب ...	اللغة	العربية غير ...
تحليل دقيق ...	اللغة	العربية وتقنيات ...
والظواهر اللغوية ..	اللغة	لتحقيق المرجو من
تعاليمها اللغة لتحقيق اللغة	اللغة	العربية من ندرة ...
من ندرة الكفاءات ..	اللغة	بمنظور علمي ...
التخلف الحضاري ..	اللغة	هي شمس ...
شمس الحضارات ..	اللغة	أيضاً من مشكلة ..
مثلاً ورغم ذلك ...	اللغة	العربية العامة أما
تمثل صعوبة اللغة	اللغة	العامة فهي لغة ..

تمثل صعوبة اللغة	اللغة	العامة فهي لغة ..
يحتلها عن طبيعة ...	اللغة	الفصحى من جهة ..
تصحيحها لكي ...	اللغة	الفصحى وليست ..
قصور العامة في ..	اللغة	العامة قادرة فعلاً ..
حيث شئت جملة ..	اللغة	العربية الفصحى ..
القومية بين أقطار ..	اللغة	الفصحى العامل ..
واضطراب ثقافي ..	اللغة	العامة في يكتفها ..
العامة في يكتفها ..	اللغة	العامة أي مشكلة
العامة أي مشكلة	اللغة	الفصحى لأن الناس
الناس لحيه رغبة	اللغة	لكن سياق الحياة ..
يخص من المعلم	اللغة	العربية الفصحى ..
تعلم اللغة العربية	اللغة	العامة في الأصل
لغة أجنبية وقد ...	اللغة	الفصحى قروياً ...
باللغة العامة ...	اللغة	الفصحى عن طريق
بين الناس عن ...	اللغة	العربية لدارستها ..
الحواسيب استخدام ...	اللغة	في وسائل الإعلام

شكل 4: عرض سياقات كلمة اللغة في شاشة العرض

البحث ببداية الكلمات

ولنأخذ مثالا الكلمات التي تبدأ ب "است" مثالا

المفهرس
الملف خصائص أنماط

تنسيق البحث أحمايات

المُفهرس/Indexer

است

بحث ببداية الكلمة

▼

بحث

التكرار	الكلمة
77	في
72	من
48	اللغة
34	أن
30	العربية
28	على
21	بين
18	التراث
18	لا
18	إلى
18	عن
16	هذا
16	اللغوي
16	الفصحي
15	ما
15	التي
14	العامة
13	العربي
12	يمكن
10	و
10	العرب
10	لغة
10	الدراسات
10	اللسانية
9	هذه
9	الحاسب
9	حيث

السابق	الكلمة	التالي
بين اللغات فكل لغات	استخدام	إمكانات العصر ...
اللغة العربية وحدهم	استثناء	مجموعة من ...
التحليل القاعدي	استعداد	مكتهم العلمية ...
استثماراً يخالف به ..	استثماراً	يخالف به هدف ...
وترائنا اللغوي هو ...	استقى	منها لغوي العرب ...
ضغط النصوص ...	استرجاع	النصوص باستخدام ...
إلى العامة وأسبابها	استخدام	العامة بدلاً من ...
الشخصية القومية ...	العرب جميعاً	يعد أن ...
الثقافي المدون ...	استبدالها	أيضاً بحروف لاتينية .
من الأذان والمفاهيم	استخدامها	في برامج الأطفال ..
المدرسة أيضاً ث ...	استخدام	العامة بالمدراس ...
استخدام العامة ...	الفصحي ج العودة ..	استخدام
أن نطبق ما سبق ..	استغلال	إمكانات الحاسب في ...
وعمل ملخصات ...	استعادة	المعلومات عن ...

استخدام إمكانيات العصر الحديث للرقى باستخدام اللغة العربية الفصحى

الملخص:-

مقدمة:-

الاتصال اللغوي: منذ فجر التاريخ والإنسان يستخدم اللغة للهدف الأساسي الذي نشأت من أجله، ألا وهو تبادل الأفكار والآراء والإحساس بين الناس، ولا شك أن هناك جوانب مشتركة بين اللغات فكل لغات العلم جلا استثناء- مجموعة من الأصوات تنظم في أسلوب معين يستطيع المتحدث و السامع أن يفهما بينهما دون جهد أو عناء، اللغات جميعاً مجموعة من الكلمات تعبر عن الأشياء والأفعال والمفاهيم والأفكار وكل لغة أشكال نحوية وتغييرات في الكلمة تغير من أصل معناها[1].

اللغة العربية الفصحى[1]: اللغة العربية الفصحى هي اللغة التي تستخدم في تدوين الشعر والنثر والإنتاج الفكري والعلمي عامة، وهي تخضع لقوانين ضبطها، وتحكم عبارتها، وهي تتوخى الإيضاح والأمانة، والإعراب إحدى وسائلها إلى تحقيق هذه الغاية، غاية الإيضاح والإفصاح عن صلات الكلمات بعضها بعض وعن نظم تكوين الجمل بالحوالات المختلفة لها.

ما مكانة التراث اللغوي العربي من الدراسات اللسانية الأجنبية الحديثة؟ إن الحضارة الإسلامية من أوائل الحضارات التي اهتمت بالدراسات اللغوية، وذلك مرجعه أن معجزة الإسلام في القرآن الكريم وما يظهر فيه من إعجاز لغوي إلهي لا حد له بمعنى أنه كلما تقدم العلم تغيرت الدلالات وضحت مفاهيم كانت مبهمه، مثل مفهوم الينابيع مثلاً كان يفهم على أنه ينباق البرعم عن الزهور في النبات ثم فسر على أنه بعد التقدم العلمي التفت الذي.

إذا حاولنا أن نحدد مكانة التراث اللغوي العربي من الدراسات اللسانية يجب أن نقوم على هذا التراث بكل موضوعية، وأن ننصحه في الإطار العالمي للدراسات اللغوية. على الرغم من ذلك يجعل العلماء اللسانيون الأجانب التراث العربي جهلاً تاماً، إلا القلة الأجنبية العاملة في مجال اللغة العربية نفسها لقد حاولت هذه الدراسات الأجنبية الحديثة إبراز اللغة العربية، ولكن هذه الدراسات ليست كافية لإجلال قيمة التراث اللغوي العربي وإسهامه في تطوير الفكر الإنساني لأن أبناء اللغة العربية وحدهم هم القادرون على استعادة مكانتهم العلمية لموروثهم اللغوي إذا تزودوا بالمعرفة اللسانية الحديثة، بمعنى أنه لا توجد نظريات منبثقة من اللغة العربية الحديثة يمكن للعلماء الأجانب الرجوع إليها.

والرغم مما توصل إليه النخبة العرب القدامى في ميدان وصف الظاهرة الصوتية، والذي يعد إسهاماً مهماً لا يمكن الاستغناء منه وهو معزول عن مراحل المعرفة اللسانية التي تندرج منها اللسانيات، في تراثنا اللغوي القيم أطراف الآراء وأدق التعريفات لمميزات اللغة وطبيعة الأسس الصوتية ووظيفتها الاجتماعية مثل مؤلفات ابن جني خصوصاً كما

السابق 1

التالي

Number of Lines : 106
Number of Words : 4953

شكل 5: عرض الكلمات التي تبدأ ب "است"

البحث بنهاية الكلمات

ولنأخذ مثالا الكلمات التي تنتهي ب "ون" مثالا

X

الملف خصائص أيقونات
 تنسيق البحث أخصائيات

استخدام إمكانيات العصر الحديث للرقى باستخدام اللغة العربية الفصحى

الملخص:-

مقدمة:-

الانتماء اللغوي: منذ فجر التاريخ والإنسان يستخدم اللغة للهدف الأساسي الذي نشأت من أجله، ألا وهو تبادل الأفكار والآراء والإحساس بين الناس، ولا شك أن هناك جوانب مشتركة بين اللغات فكل لغات العلم جلا استثناء- مجموعة من الأصوات تنظم في أسلوب معين يستطيع المتحدث و السامع أن يفرقا بينهما دون جهد أو عناء ، للغات جميعاً مجموعة من الكلمات تعبر عن الأشياء والأفعال والمفاهيم والأفكار و لكل لغة أشكال نحوية وتغييرات في الكلمة تغير من أصل معناها[1].

اللغة العربية الفصحى[1]: اللغة العربية الفصحى هي اللغة التي تستخدم في تدوين الشعر والنثر والإنتاج الفكري والعلمي عامة، وهي تخضع لقوانين تضبطها، وتحكم عباراتها، وهي تتوخى الإيضاح والأصالة، والإعراب إحدى وسائلها إلى تحقيق هذه الغاية ، غاية الإيضاح والإفصاح عن صلات الكلمات بعضها بعض وعن نظم تكوين الجمل بالحوالات المختلفة لها.

ما مكانة التراث اللغوي العربي من الدراسات اللسانية الأجنبية الحديثة؟ إن الحضارة الإسلامية من أوائل الحضارات التي اهتمت بالدراسات اللغوية، وذلك مرجعه أن معجزة الإسلام في القرآن الكريم وما يظهر فيه من إعجاز لغوي إلهي لا حد له بمعنى أنه كلما تقدم العلم تغيرت الدلالات وضحت مفاهيم كانت مبهمه، مثل مفهوم البتائق مثلاً كان يفهم على أنه ابتئاق البرعم عن الزهور في النبات ثم فسر على أنه بعد التقدم العلمي التفتت الذي.

إذا حاولنا أن نحدد مكانة التراث اللغوي العربي من الدراسات اللسانية يجب أن نقوم على هذا التراث بكل موضوعية، وأن نضعه في الإطار العلمي للدراسات اللغوية. على الرغم من ذلك يجهل العلماء اللسانيون الأجانب التراث العربي جهلاً تاماً رلاً القلة الأجنبية العاملة في مجال اللغة العربية نفسها لقد حاولت هذه الدراسات الأجنبية الحديثة إبراز اللغة العربية، ولكن هذه الدراسات ليست كافية لجلاء قيمة التراث اللغوي العربي وإسهامه في تطوير الفكر الإنساني لأن أبناء اللغة العربية وحدهم هم القادرون على استعادة مكانتهم العلمية لموروثهم اللغوي إذا تزودوا بالمعرفة اللسانية الحديثة، بمعنى أنه لا توجد نظريات منبثقة من اللغة العربية الحديثة يمكن للعلماء الأجانب الرجوع إليها.

والرغم مما توصل إليه النخبة العرب القدامى في ميدان وصف الظاهرة الصوتية، والذي يعد إسهاماً مهماً لا يمكن الاستغناء منه وهو معزول عن مراحل المعرفة الإنسانية التي تتدرج منها اللسانيات. في تراثنا اللغوي القديم أطراف الآراء وأدق التعريفات لمميزات اللغة وطبيعة الألسنية الصوتية ووظيفتها الاجتماعية مثل مؤلفات ابن جني خموماً كنا

المُفهرس/Indexer

بحث بنهاية الكلمة

بحث

السابق	الكلمة	التالي
المتحدث و السامع أ دون	جهد أو عناء لغات ..	...
على الرغم من ذلك	لللسانيون	الأجانب التراث ...
لأن أبناء اللغة ...	القادرون	على استعادة مكانته
إلى نشوز واضح في	وكون	هوة سخيفة بين ...
والقضايا اللغوية ...	والمترجمون	على إيجاد مقليل ...
الآلية للخلفية اللا	مضمون	الخبرة البشرية بكل
على حساب الوقت	تكون	في هذه الحالة منابع
وطائف الذهن اللغوية	تكون	علاقة عطاء تبادلية .
بالنسبة لتعلم الأطفال	تعليمون	بالفصحى في ...
المدارس حيث ...	ويتحدثون	العامة بالمنزل ...
فهو لغة الحديث	التأشؤون	العامة ويجري بها ..
ولذلك فهي لا تصلح	تكون	أكثر من تخاطب في
أن تكون أكثر من ...	التأشؤون	العامة ومن ثم لا ...
من أعراض لأنها لا	تكون	لغة عامة لاختلاف ..
الأجيال العربية ...	المدون	بالحروف العربية ...
أقرانه من أنشطة	يلعبون	منصتاً لكل ما ...
وهو يراقبهم يلعبون	يقولون	متنهرزاً الفرسه ...
اللغة في وسائل ...	تكون	قريبة من الأذان ...

1 السابق

التالي

Number of Lines : 106
Number of Words : 4953

شكل 6: الكلمات التي تنتهي ب"ون"

29

فهرس (2) الإحصاء

الكلمات المتصاحبة: عدد الكلمات المتصاحبة = 2 مدى التصاحب

المفهرس

الملف خصائص أخطاء

تنسيق البحث احتماليات

Dependency

N: 2

Collocation بحث عن

	Score PMI	Frequen	Gram	Num
1	12.584962500721156	1	أذهب بسمعهم	1
2	12.584962500721156	1	وحيرل وميكال	2
3	12.584962500721156	1	التصف تعرفهم	3
4	12.584962500721156	1	عباده فيأورا	4
5	12.584962500721156	1	ويستحيون نساءكم	5
6	12.584962500721156	1	وأوفوا بعهدي	6
7	12.584962500721156	1	أموالا فأحياكم	7
8	12.584962500721156	1	فادوا والنصارى	8
9	12.584962500721156	1	تفتنهمهم وأخرجهم	9
10	12.584962500721156	1	تبتهم فلهم	10
11	12.584962500721156	1	تتمتع بالعمرة	11
12	12.584962500721156	1	والكتاب والنبين	12
13	12.584962500721156	1	اضرب بعصاك	13
14	12.584962500721156	1	فبشرون ومنذرين	14
15	12.584962500721156	1	وإله إيتك	15
16	12.584962500721156	1	والعذاب بالمعقرة	16
17	12.584962500721156	1	البرق يخطف	17
18	12.584962500721156	1	كفارا حسدا	18
19	12.584962500721156	1	بالعروة الوثقى	19
20	12.584962500721156	1	الوسطى وقوموا	20
21	12.584962500721156	1	أخنته العزة	21
22	12.584962500721156	1	ادعوه بأنبياء	22
23	12.584962500721156	1	أكلها مضعين	23
24	12.584962500721156	1	للطافين والعاكفين	24
25	12.584962500721156	1	فأروا وماروت	25
26	12.584962500721156	1	وعمل صالحا	26
27	12.584962500721156	1	فحسبه جهنم	27
28	12.584962500721156	1	دعاء ودهاء	28
29	12.584962500721156	1	ليطفن قلبى	29
30	12.584962500721156	1	تخشوهم واخشوني	30
31	12.584962500721156	1	المغرب فبهت	31
32	12.584962500721156	1	بملائكة	32

بسم الله الرحمن الرحيم

الم

1

ذلك الكتاب لا ريب فيه هدى للمتقين

الذين يؤمنون بالغيب ويقيمون الصلاة وما رزقناهم ينفقون

والذين يؤمنون بما آتينا اليك وما أنزل من قبلك ولاخروهم يؤفون

أولئك على هدى من ربهم وأولئك هم المفلحون

إنا الذين كفروا سوء عليهم المذنبهم إن لم تنذروهم لا يؤمنون

ختم الله على قلوبهم وعلى سمعهم وعلى أبصارهم غشاوة ولهم عذاب عظيم

ومن الناس من هوى له مال الله وألوانه الأخر وما هم بمؤمنين

يغارون الله والذين آمنوا وما يخذعون إلا أنفسهم وما يشعرون

في قلوبهم مرض فزادهم الله مرماة ولهم عذاب أليم بما كانوا يكفون

وإذا قيل لهم لا تفسدوا في الأرض قالوا إنما نحن مصلحون

ألا أنهم هم المفسدون ولكن لا يشعرون

وإذا قيل لهم أين ما آمن الناس قالوا أنؤمن كما آمن السفهاء ألا إنهم هم السفهاء

ولكن لا يعلمون

وإذا قيل لهم أين ما آمنوا قالوا آمنوا أنما نكذب في شفاطينهم قالوا إنا معكم إذا نحن

مستهزؤون

الله يستعزدهم ويهدهم في صراطهم يعلمون

أولئك الذين اشتروا الضلالة بالهدى فما ربحت تجارتهم وما كانوا مهتدين

فلهم كمثل الذي استوفد نارا فلما أمانت ما حوله ذهب الله بنورهم وتركهم في

ظلمات لا يبصرون

ممن بكم عني فإني لا يرجعون

أو نصيب من السيمة في ظلمات وعد وقرى يجعلون أصابعهم في آذانهم من

المواعظ حذر الموت والله محيط بالكافرين

بئس البرق يخطف أبصارهم كلما أضاء لهم مشوا فيه وإذا أظلم عليهم أقبلوا وقوله

الله أنذهب بسمعهم وأبصارهم إنا الله على كل شيء قدير

يا أيها الناس اعبدوا ربكم الذي خلقكم والذين من قبلك

السابق 1 التالي

Number of Lines : 574

Number of Words : 6431

شكل 7: المتصاحبات اللغوية

فهرس(3) البحث بالجذر

ولنأخذ الجذر "درس" مثالا

المفهرس
الملف خملص أخطاء
تنسيق البحث أخصيات

Indexer/المفهرس

درس
بحث بالجذر

77 في
72 من
48 اللغة
34 أن
30 العربية
28 على
21 بين
19 التراث
18 لا
18 لاني
18 عن
16 هذا
16 اللغوي
16 الفصحى
15 ما
15 التي
13 العامة
12 العربي
10 يمكن
10 و
10 العرب
10 لغة
10 الدراسات
10 اللسانية
9 هذه
9 الحاسب
9 حيث

الكلمة
الكرار

ما مكانة التراث ...
الإسلامية من أوائل ...
نحدد مكانة التراث ..
وأن نضعه في ...
اللغة العربية نفسها ...
الحديثة ليراز اللغة ..
القائمة الآن إلا أنها ..
نشأة العلم المذكور ...
التراث اللغوي ...
من لغوي العرب ...
نحو أدى إلى نشوز ...
المجال عن ذكرها ...
حقيقة هذا التراث ..
الملحوظة بين الباحث ...
المضروبة عليه ...
لها أسسها ...
الدراسات اللسانية ..
في الوصف ...
بين اللغة علاقة ...
والدالية المعجمية ..
التي تنشأ بالنسبة ..
في المدارس حيث ..
لم يكن اهتمام ...
عن طريق أ تبسيط ..
مشوقة حتى بألفها ...
قبل دخول المدرسة ...
منع المدرسين من ..
دار المويخ للنشر ...
التربية الإسلامية ...

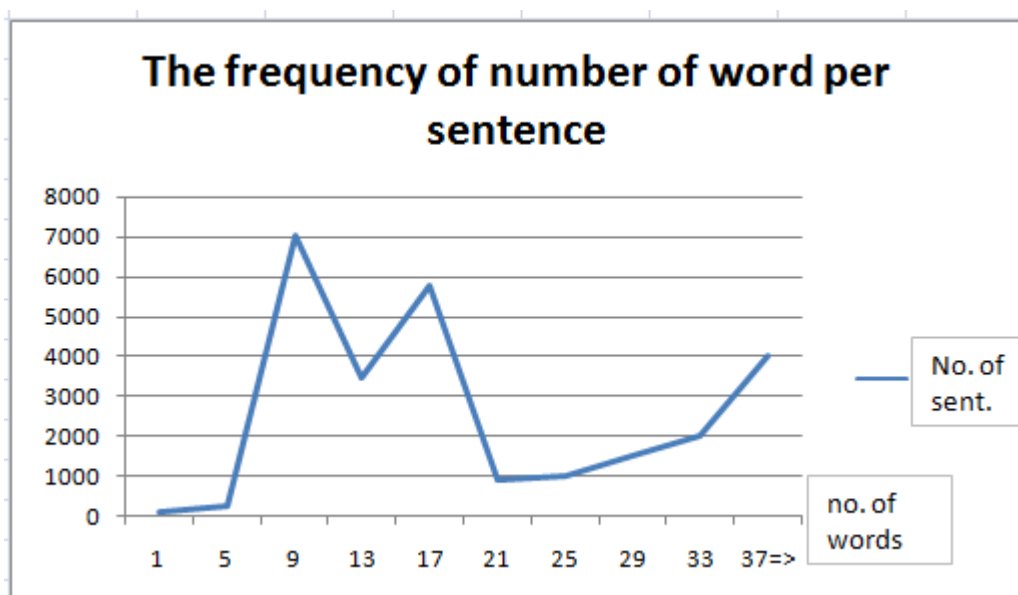
السابق
التالي

1
Number of Lines : 106
Number of Words : 4953

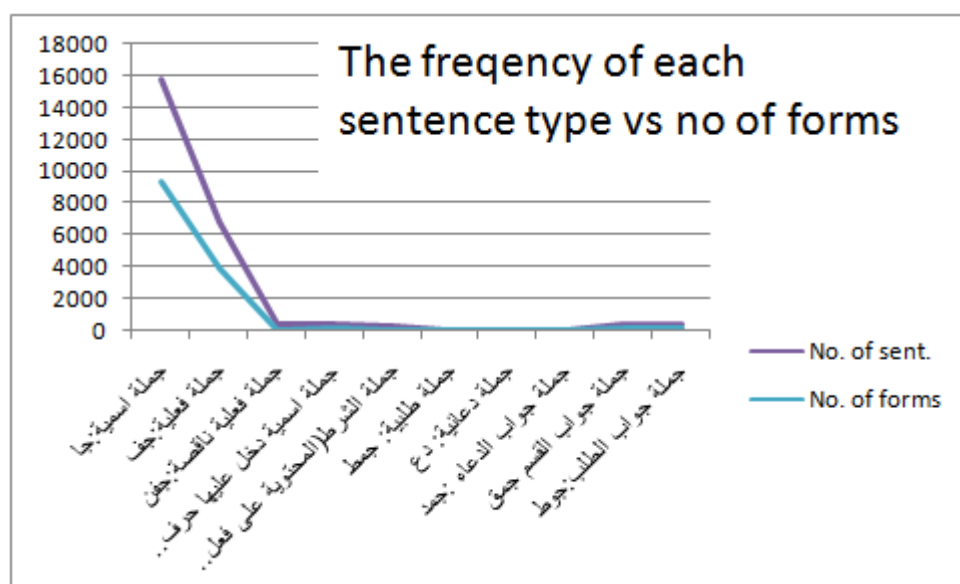
السابق	الكلمة	التالي
ما مكانة التراث ...	الدراسات	اللسانية الأجنبية ...
الإسلامية من أوائل	بالدراسات	اللغوية وذلك مرجعه
نحدد مكانة التراث ..	الدراسات	اللسانية يجب أن تقو
وأن نضعه في ...	للدراسات	اللغوية على الرغم .
اللغة العربية نفسها	الدراسات	الأجنبية الحديثة ...
الحديثة ليراز اللغة ..	الدراسات	ليست كافية لجلاء ..
القائمة الآن إلا أنها .	للدراسات	اللغوية التقليدية ...
نشأة العلم المذكور	الدراسات	الحديثة للسانيات مع
التراث اللغوي ...	الدراسات	اللسانية الأجنبية ...
من لغوي العرب ...	لدراسة	العربية بمناهج ...
نحو أدى إلى نشوز .	الدراسات	التطبيقية وكون هوة
المجال عن ذكرها ...	بالدراسات	اللسانية من العرب أ
حقيقة هذا التراث ..	الدراسات	اللسانية الحديثة إذ .
الملحوظة بين الباحث	بالدراسات	اللسانية الغربية ...
المضروبة عليه ...	مدرسة	لسانية عربية لها ...
لها أسسها ...	الدراسات	اللسانية الحديثة هنا
الدراسات اللسانية .	الدراسة	اللسانية والباحث ...
في الوصف ...	الدراسة	اللسانية اليوم العلاء
بين اللغة علاقة ...	دراسة	اللغة بمنظور منطقي
والدالية المعجمية .	الدراسات	نصل إلى قواعد تمك
التي تنشأ بالنسبة .	المدارس	حيث يتعلمون ...
في المدارس حيث .	المدرسة	ويتحدثون العامية ...
لم يكن اهتمام ...	بدراسة	العامية وعلاقتها ...
عن طريق أ تبسيط ..	لدارسيها	وتيسير النحو ...
مشوقة حتى بألفها	المدرسة	أيضاً ت منع ...
قبل دخول المدرسة	المدرسين	من استخدام العامية
منع المدرسين من .	بالمدراس	وتشجيع الطلبة على
دار المويخ للنشر ...	دراسات	في التربية الإسلامية
التربية الإسلامية ...	للدراسات	والنشر والتوزيع

شكل 8: إظهار مواضع مشتقات درس في القرآن الكريم في سياقاتها.

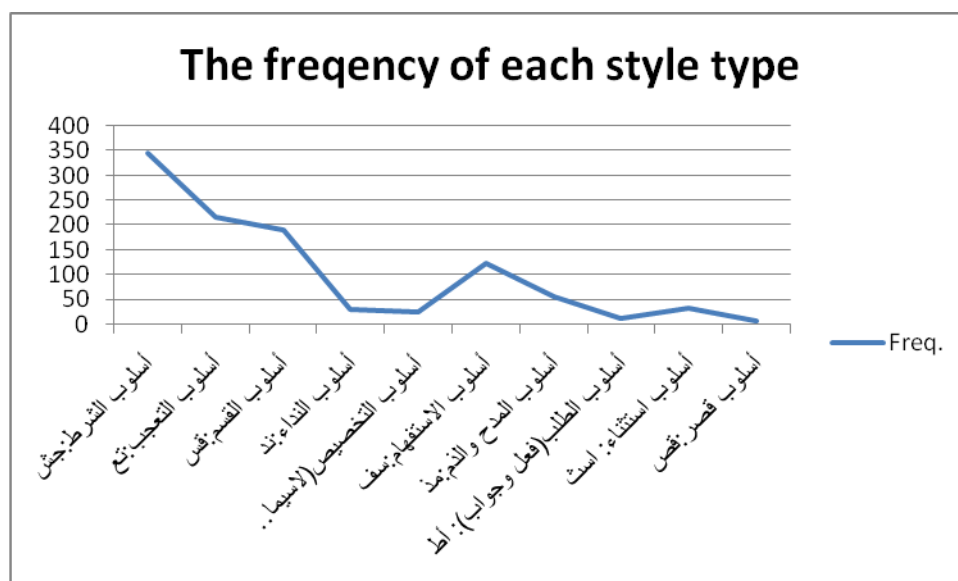
فهرس(4) البحث في المدونات المذيلة



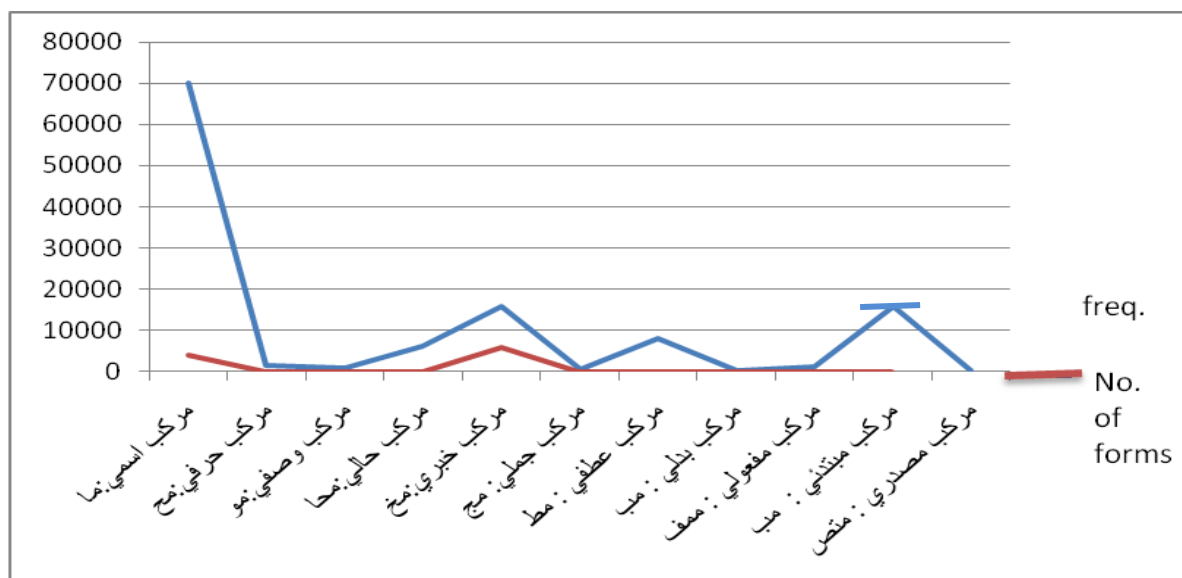
شكل 9. عدد الجمل بالنسبة لعدد الكلمات في مساك [سلوى حمادة، 2010]



شكل 10. تردد الجمل بالنسبة لأنواعها. types of sentences number versus their forms.



شكل 11. تردد جميع الأساليب. The frequency of each style type.



شكل 12. تردد جميع أنواع المركبات النحوية

The frequency of each syntatic structure type.