

# الذخيرة النصية الفصحى حجر الأساس لدراسة معاني كلمات القرآن الكريم

في ضوء نماذج الدلالة المتوزعة

عبدالمك سلمان السلطان

قسم علوم الحاسب

جامعة الملك سعود

الرياض، المملكة العربية السعودية

salman@ksu.edu.sa

مها سليمان الربيعية

قسم علوم الحاسب

جامعة الملك سعود

الرياض، المملكة العربية السعودية

msrabiah@gmail.com

إيريك أتويل

كلية الحوسبة

جامعة ليدز

ليدز، المملكة المتحدة

e.s.atwell@leeds.ac.uk

بعض التجارب الأولية التي أجريتها على الذخيرة النصية الفصحى  
والنتائج التي توصلنا إليها.

الكلمات المفتاحية- القرآن الكريم؛ الذخيرة النصية؛ العربية  
الفصحى؛ نماذج الدلالة المتفرقة؛ اللغويات الحاسوبية.

## 1. مقدمة

يتميز النص القرآني الشريف عن غيره من نصوص اللغة العربية  
الفصحى بالمتانة وبلاغة العبارة وانتقاء المفردة بالإضافة إلى احتوائه على  
العديد من التراكيب الصرفية والنحوية الفريدة والمختلفة. كما يتميز النص  
القرآني الشريف بتوظيفه لنفس المفردة كي تخدم معانٍ متعددة بناءً على  
السياق الذي وردت فيه. فعلى سبيل المثال، نجد أن كلمة "فتنة" قد  
وردت في القرآن بأحد عشر معنىً مختلفاً<sup>1</sup>، كما أن بعض المفردات  
القرآنية تحتل عدة معانٍ مختلفة في نفس الآية والسياق الذي وردت فيه

الملخص- تعد نماذج الدلالة المتوزعة إحدى الطرق التجريبية  
لدراسة تركيب اللغة وتصميمها، وهي مبنية على بناء نماذج لمعاني  
الكلمات عن طريق الدراسات الإحصائية لمدى توزع تلك  
الكلمات وانتشارها في نصوص ضخمة جداً.

وفي هذه الورقة، سنعرض الذخيرة النصية الفصحى لجامعة الملك  
سعود والتي تعد حجر الأساس لدراسة تعنى ببناء نموذج للدلالات  
المتوزعة لمعاني كلمات القرآن الكريم. وهي ذخيرة نصية ضخمة  
جداً ومتاحة مجاناً، قوامها ما يربو على خمسين مليون كلمة من  
مؤلفات تنتمي إلى الحقبة الزمنية ما بين العصر الجاهلي وحتى  
القرن الرابع الهجري. وستتطرق للأسس التي بنيت عليها الذخيرة  
النصية الفصحى بما في ذلك الهدف من بنائها ومدى توازنها  
 وتمثيلها للغة الفصحى، وكذلك حجم المقاطع النصية المأخوذة  
من المؤلفات وحقوق النشر الخاصة بها. بالإضافة إلى ترميز  
حروف المؤلفات النصية وكيفية ترتيبها في مجلدات. كما سنعرض

<sup>1</sup> [www.islamqa.com](http://www.islamqa.com)

[1]، كقوله تعالى: ﴿وَمَا قَلَّوهُ يَقِينًا﴾ النساء: ١٥٧. أضيف إلى ذلك أن النص القرآني يزخر بالعبارات المسبوكة والتي توصل للقارئ المعنى المراد بل ومعانٍ متعددة باستخدام أقل المفردات، كقوله تعالى: ﴿وَلَكُمْ فِي الْقِصَاصِ حَيَوةٌ يَا أُولِي الْأَلْبَابِ لَعَلَّكُمْ

تَتَّقُونَ﴾ البقرة: ١٧٩. ولعل مما يزيد من بلاغة النص القرآني الشريف وفصاحته كثرة استخدام الاستعارة والتشبيه مما يضيف أبعاداً إضافية للمعنى الذي يراد إيصاله للقارئ، كقوله تعالى: ﴿قَالَ رَبِّ إِنِّي

وَهَنَ الْعَظْمُ مِنِّي وَاسْتَعَلَ الرَّأْسُ شَيْبًا وَلَمْ أَكُنْ بِدَعَائِكَ رَبِّ شَقِيًّا﴾ مريم: ٤. كل هذه الخصائص الفريدة جعلت من النص القرآني الشريف المرشح الأمثل لتمثيل اللغة العربية الفصحى، والنص العربي الفصيح الذي يشكل التحدي الأكبر أمام الفهم الآلي للغة العربية.

## 2. نماذج الدلالة المتوزعة

معظم المحاولات السابقة لمعالجة القرآن الكريم واللغة العربية الفصحى آلياً كانت مبنية على النهج العقلاني لاكتساب اللغة، والذي يتبنى مبدأ أن العقل البشري مهياً بوحدة عامة للغة تحتوي على العديد من القواعد والأنماط والتراكيب التي يعتمد عليها الأطفال في عملية اكتساب وتوليد اللغة. بل إن معظم البحوث العلمية في مجالي المعالجة الآلية للغات الطبيعية واللغويات الحاسوبية في الفترة ما بين 1960م إلى 1985م كانت مبنية على النهج العقلاني، حيث تم بناء العديد من الأنظمة الذكية والمزودة بالعديد من القواعد اللغوية المبنية في داخلها بالإضافة إلى خوارزميات تعلم متعددة. إلا أنه قد تم نقد مثل هذه الأنظمة والتي تتبنى النهج العقلاني لاعتمادها المباشر على القواعد التي يغذيها بها الإنسان بالإضافة إلى عدم قدرتها على التعامل بفعالية مع النصوص الضخمة جداً [2]، الأمر الذي حدى بالعلماء إلى العودة إلى النهج التجريبي لتعلم اللغة الذي كان سائداً في الفترة ما بين 1920م إلى 1960م والذي لا يؤيد وجود تلك الوحدة العامة للغة في العقل البشري بل يفترض أن الله قد زوده بالبيات ذكاء مثل المقدرة على التعميم والتعرف على الأنماط المتشابهة والتي عندما يطبقها الأطفال على ما يتلقونه عن طريق حواسهم يتمكنون من خلاله تعلم التركيب المعقد للغة [2].

وبناءً على ذلك، تعتنى الطرق التي تتبع النهج التجريبي بدراسة الاستخدام الواقعي للغة والذي من خلاله تستطيع أن تستشف تركيب اللغة والخصائص المميزة لها. ولعل الطريقة الأمثل لتوفير ذلك الاستخدام الواقعي للغة لتلك الطرق هو عن طريق الذخائر اللغوية والتي يمكن تعريفها على أنها مجموعة من النصوص المكتوبة أو (و) المنطوقة والمصممة بشكل منهجي بحيث يمكن الاستفادة منها في دراسة اللغويات [3].

ومن هذا المنطلق، كانت إحدى الخطوات نحو فهم القرآن الكريم عن طريق الذكاء الصناعي والمعالجة الآلية للغات الطبيعية هي محاولة استخدام الطرق التي تتبع المنهج التجريبي لدراسة التركيب الفريد للغة القرآن ومعاني مفرداته. وربما تعد نماذج الدلالة المتوزعة (Distributational Semantic Models) أحد أهم تلك الطرق. حيث تعتنى بشكل رئيسي ببناء نماذج لمعاني الكلمات مستخدمة الدراسات الإحصائية لمدى توزع تلك الكلمات وانتشارها في ذخائر نصية ضخمة جداً [4]. وهي مبنية على فرضية التوزع (Distributational Hypothesis) والتي تعود للعالم هاريس (1970م) وتنص على أن "التغير في المعنى له علاقة بالتغير في التوزع" [5]، أو بمعنى آخر "أن الكلمات التي لها معاني متقاربة تظهر في سياقات متشابهة" [6]، وعلى هذا فإن "الكلمات التي تظهر في نفس السياقات، تميل إلى أن يكون لها معانٍ متشابهة" [7].

وبالرغم من أن هاريس لم يوضح في فرضيته ما هية هذا الاختلاف في توزع الكلمات، إلا أن اللغوي السابق دي سوسور قد أشار إليه مسبقاً وأنه نابع من نوعين من العلاقات، وهما السينتاجماتية (syntagmatic) والبرادجماتية (paradigmatic) [8]. فالعلاقات السينتاجماتية هي علاقات خطية اندماجية تنشأ بين الكلمات التي تتوارد مع بعضها في نص متتابع، مثل كلمتي "الغيوم" و"المطر". بينما العلاقات البرادجماتية هي علاقات استبدالية تنشأ بين الكلمات التي تقع في نفس السياق وتشارك في نفس الكلمات المجاورة ولكنها لا تتوارد مع بعضها في نفس النص، مثل كلمتي "المحبة" و"المودة".

ومن التعريف السابق لفرضية التوزع، يمكن تحديد عنصرين أساسيين لبناء نماذج الدلالة المتوزعة؛ وهما توفر الذخائر اللغوية الضخمة جداً

بالإضافة إلى النماذج الحسابية أو الرياضية التي سوف تطبق على تلك الذخائر [9].

### أ. الهدف من الذخيرة النصية الفصحى

الهدف الرئيسي من بناء الذخيرة الفصحى هو استخدامها في بناء نماذج دلالة متوزعة لمفردات وتراكيب القرآن الكريم، إلا أنه وكما ذكرنا قد صممت بطريقة عامة تتيح استخدامها في أبحاث لغوية أخرى، منها على سبيل المثال لا الحصر:

- بناء المعاجم.
- بناء المعاجم الدلالية.
- البحث عن الكنوز اللغوية من مرادفات وتراكيب لغوية في حقل من الحقول الدلالية في الذخيرة النصية الفصحى وإحيائها مرة أخرى عن طريق تضمينها في المناهج التعليمية وكذلك الكتابات الأدبية.
- دراسة التطور الدلالي للمفردات والتراكيب اللغوية منذ العصر الجاهلي وحتى عصرنا هذا، وكذلك استعارة المفردات للدلالة على معاني أخرى في مجالات أخرى.
- دراسة المصطلحات والتعابير المسكوكة وظاهرة التصاحب اللغوي.
- دراسة ارتباط بعض المفردات بتراكيب نحوية محددة.
- استنباط الفروق الدقيقة بين المترادفات.
- دراسة العلاقة بين التراكيب النحوية ونوع النص.
- إيجاد العلاقات الدلالية (الترادف، التضاد، جزء من، يتضمن...).
- بناء أنتولوجيا للقرآن الكريم، واللغة العربية الفصحى.
- تعليم اللغة العربية الفصحى للعرب ولغيرهم.

### ب. التمثيل والتوازن

يعد التمثيل والتوازن من المقومات التي يفضل توفرهما في أي ذخيرة لغوية وذلك لأهميتهما في تحديد مدى دقة النتائج التي سيحصل عليها الباحث من الذخيرة النصية. ويمكن تعريف الذخيرة اللغوية المثلثة بأنها الذخيرة التي تمثل استخدام اللغة في الواقع وتغطي جميع مجالاتها ويمكن

### 3. الذخيرة النصية الفصحى

تعد الذخيرة النصية الفصحى لجامعة الملك سعود<sup>1</sup> حجر الأساس لبناء مشروع يهدف إلى دراسة نماذج الدلالة المتوزعة لمفردات القرآن الكريم. فكما هو معلوم بأن النص القرآني نزل بلسان عربي مبين وبلغة هي غاية في الفصاحة، لذا فقد كان لزاماً أن تتم دراسة مفرداته وتراكيبه في ضوء نصوص مقارنة له في فصاحة اللغة، وهي النصوص العربية التي ألفت في قرون الاحتجاج، من العصر الجاهلي وحتى القرن الرابع الهجري.

وعلى حسب علم الباحثين فإنه لا توجد سوى ذخيرتين نصيتين متخصصتان في اللغة العربية الفصحى؛ فالأولى هي جزء من المدونة العربية التابعة لمدينة الملك عبدالعزيز للعلوم والتقنية<sup>2</sup>، والثانية هي مدونة اللغة العربية الفصحى [3]. إلا أن كليهما لم يكن مناسباً للبحث، وذلك لأن الأولى لم تغط العديد من المجالات التي كان التأليف فيها سائداً في ذلك الوقت، كما أن قوامها حوالي 17 مليون كلمة فهي لا تعد ضخمة جداً، أما الثانية فإنها صغيرة نسبياً حيث تتكون من خمسة ملايين كلمة فقط.

لذا كان لابد من العمل أولاً على مشروع جمع وتصميم ذخيرة نصية خاصة باللغة العربية الفصحى بحيث تكون كبيرة الحجم وعلى مستوى مقبول من التمثيل والتوازن مما يناسب تطبيق نماذج الدلالة المتوزعة عليها. وعلى هذا فقد تم جمع وبناء الذخيرة النصية الفصحى لجامعة الملك سعود وقوامها أكثر من خمسين مليون كلمة وحاولنا أن يكون تصميمها عاماً وذلك لإتاحتها لأبحاث علمية أخرى.

<sup>1</sup> الذخيرة النصية متوفرة مجاناً على الإنترنت

<http://ksucorpus.ksu.edu.sa/ar>

<sup>2</sup> موقع المدونة العربية على الإنترنت

<http://www.kacstac.org.sa/Pages/search.aspx>

المجال	المجال الثانوي
علم الاجتماع	الجغرافيا والرحلات
	الطب
	الفيزياء
	الفلك
	الفلسفة
	السياسة
	علوم متفرقة
السير والتراجم	الآداب والرقائق
	علم الأنساب
	سيرة النبي محمد صلى الله عليه وسلم
	التراجم والطبقات

ثم قمنا بعد ذلك بتصنيف النصوص وفقاً للمجال الرئيسي/الفرعي الذي تنتمي إليه، وأسفر التصنيف عن نسب غير متساوية من النصوص في كل مجال رئيسي/فرعي، كما في الجدول رقم 2. إلا أن هذا يتوافق مع ما نعرفه عن توجه التأليف في تلك الفترة الزمنية من تاريخ العرب، الأمر الذي يشير إلى توازن ذخيرتنا النصية.

جدول 2: وصف عام لمحتوى الذخيرة النصية الفصحى

النسبة المئوية	عدد الكلمات	عدد النصوص	المجال
46.73 %	23645087	150	الدين
14.02 %	7093966	56	اللغة
14.28 %	7224504	104	الأدب
12.71 %	6429133	42	العلوم
5.36 %	2709774	32	علم الاجتماع
6.92 %	3499948	26	السير والتراجم
100 %	50602412	410	المجموع

### ت. حجم المقاطع النصية

تم اعتماد كامل النص للمؤلفات في الذخيرة النصية وذلك، وفقاً لسنكلير [10]، يعد الاختيار الأنسب لاكتشاف الخصائص اللغوية والمعاني التي ما تكون غالباً متوزعة بشكل متساوي في ثنايا النص.

### ث. إذن حقوق الطبع والنشر

المؤلفات المشمولة في الذخيرة النصية الفصحى تعود لقرون طويلة ماضية ولا تخضع لحقوق الملكية الفكرية.

تعميم النتائج التي يتم الحصول عليها منها على اللغة ككل. أما الذخيرة اللغوية المتوازنة فهي التي يكون فيها عدد المقاطع المأخوذة من كل مجال من المجالات متناسباً مع حجم النصوص المصنفة ضمن ذلك المجال في الواقع.

ونظراً لقلّة النصوص الرقمية الموثقة والمؤلفة في الفترة من العصر الجاهلي وحتى القرن الرابع الهجري فقد حاولنا أن نجمع كل ما توفر لدينا، ولم يكن هنالك حاجة لعمل اختيار عشوائي من عينات مجتمع البحث المتوفر لدينا وذلك لقلّة عدد النصوص كما ذكرنا، وكانت المكتبة الشاملة على الإنترنت أحد أهم المصادر التي اعتمدنا عليها<sup>1</sup> في توفير مجتمع البحث.

وبعد دراسة المجالات التي تنتمي إليها النصوص المتوفرة ولكي نصل إلى ذخيرة نصية ممثلة، قسمنا مجتمع البحث إلى ستة مجالات رئيسية تغطي معظم المجالات التي كان التأليف فيها سائداً في ذلك الوقت، وهي: الدين واللغة والأدب والعلوم وعلم الاجتماع والسير والتراجم. كما قمنا بتقسيم كل من هذه المجالات الرئيسية إلى مجالات فرعية بلغ مجموعها 27 مجالاً، كما يظهر في الجدول رقم 1.

جدول 1: توزيع مجتمع البحث إلى مجالات رئيسية وثانوية

المجال	المجال الثانوي
الدين	القرآن الكريم
	الحديث الشريف
	تفسير القرآن الكريم
	علوم القرآن
	علوم الحديث
	العقيدة
	الفقه
اللغة	أصول الفقه
	النحو والصرف
	اللغة
	المعاجم
الأدب	الأمثال
	الشعر
	الرواية
العلوم	الأدب والبلاغة
	التاريخ

<sup>1</sup> موقع المكتبة على الإنترنت <http://shamela.ws>

ج. ترميز المحارف  
تستخدم الذخيرة النصية الفصحى ترميز المحارف المعروف (UTF-8) والذي يتوافق مع الترميز السابق الـ (ASCII) ومدعوم من قبل العديد من الأنظمة [11].

جدول 4: معاني كلمة "ماء" في القرآن الكريم

عدد مرات التوارد	الكلمات المتواردة	المعنى
22	أنزل، السماء	ماء المطر
10	شرب	ماء الشرب
3	الخلق، الإنسان، دافق، مهين	المني

عند محاولة إيجاد متصاحبات كلمة "ماء" في الذخيرة النصية الفصحى باستخدام المعيار الإحصائي (MI.log<sub>f</sub>) في برنامج (Sketch Engine) [12] كان التقييم الأعلى هو للمتصاحبات التي في الجدول رقم 5.

جدول 5: معاني كلمة "ماء" في الذخيرة النصية الفصحى

MI.log <sub>f</sub>	المتصاحبات	المعنى
63.325	الشعير	الأعشاب الممزوجة بالماء
54.434	العسل	الشراب المخلوط بالماء
54.210	السماء	ماء المطر
51.729	تجدوا	ماء الشرب
50.467	زمزم	ماء زمزم
46.671	دافق	المني
46.055	الرمان	عصير

ونستنتج من الجدول السابق أن كلمة "ماء" استخدمت في العربية الفصحى بنفس المعاني التي وردت في القرآن الكريم بالإضافة إلى معاني أخرى نتجت بسبب تغير الكلمات المصاحبة لها. ويرجع ارتفاع تقييم أول معنيين لوجود العديد من كتب الطب في الذخيرة النصية الفصحى والتي استخدمت كلمة "ماء" كثيراً.

ب. الفهرس السياقي لمفردات القرآن في الذخيرة النصية

#### الفصحى

يعد الفهرس السياقي من أهم الأدوات التي تعين الباحث على التعرف على الكيفية التي تستخدم بها الكلمة في النص بحيث يستطيع من خلاله استنباط المعاني المختلفة للكلمة والقرائن التي أدت إلى نشوء

ح. ترتيب الملفات

تتكون الذخيرة النصية الفصحى من عدة ملفات نصية بحيث يمثل كل ملف مؤلف واحد (كتاب أو قصيدة). وتم ترتيب هذه الملفات في ستة مجلدات أساسية تمثل المجالات الرئيسية للذخيرة النصية، والتي تحتوي بدورها على عدة مجلدات تمثل المجالات الثانوية. كل من المجالات الأساسية والثانوية تم الرمز إليها بحروف الأبجدية الإنجليزية بشكل يضمن عدم تطابق أي مجلدين في التسمية. وكل ملف يتكون اسمه من حرف المجال الرئيسي ملحفاً بحرف المجال الثانوي ملحفاً برقم الملف التسلسلي في المجال الثانوي. الجدول 3 يوضح طريق تسمية ملفات مجال (الدين).

جدول 3: طريقة تسمية ملفات مجال (الدين)

المجال (الرمز)	المجال الثانوي (الرمز)	اسم الملفات
(A) الدين	(A) القرآن الكريم	AA1
	(B) الحديث الشريف	AB1-44
	(C) تفسير القرآن الكريم	AC1-13
	(D) علوم القرآن	AD1-29
	(E) علوم الحديث	AE1-10
	(F) العقيدة	AF1-23
	(G) الفقه	AG1-26
	(H) أصول الفقه	AH1-4

#### 4. أمثلة على تحليل بعض المفردات القرآنية باستخدام الذخيرة النصية الفصحى

أ. مقارنة استخدام كلمة "ماء" في القرآن والذخيرة النصية الفصحى

بإمكاننا استخدام الذخيرة النصية الفصحى لمعرفة الكيفية التي كانت تستخدم بها مفردات القرآن الكريم في لغة العرب. ولنأخذ على

هذا المعنى دون غيره من المعاني الأخرى. الشكل 1 يعرض بعض السطور من الفهرس السياقي لكلمة "الفتنة" مستخرجة من الذخيرة النصية الفصحى وباستخدام برنامج (Sketch Engine) [12].

وقال خليفة وذكر <O> </O> . سمعنا النذر **الفتنة** فقال : اتكلم الدعوات ترسي بالفتن ثم التي : والفتح : الجفاف . والفتح : الحركة . والفتح **الفتنة** والفتح : هيجان الدم أو الجراح أو التوق : قال أبو سعيد : البهجة <O> </O> . وبعثت **الفتنة** والفتح والإخلاص . يدل منه : قد هربتم النوم أنس في البيئات فرد . حتى به الفتن **يقول** في **الفتنة** **يقول** أبو زيد <O> </O> . لا يدرى من قلة الله جل وجن : [ ولا أرضوا خلتكم بيموتكم **الفتنة** التوبة : 47 ] أوحشت في السور : أسرعت [ ] اليتم : أراد والأرضوا مراكبكم خلتكم بيموتكم **الفتنة** **يقول** <O> </O> . وجعل خلتكم بعلمى وبسلكم . وبنى هند <O> </O> . ( ) يدى الساعة جل **الفتنة** والخلل : الجن <O> </O> . والهرج والقتل وأتانا فتن زين الطيخة - أي : زين <O> </O> **الفتنة** **يقول** اللطيفي : طاح فتن <O> </O> . والحرب **يقول** : هو يصرى في صرة الظهر ويشك في صرة **الفتنة** **يقول** <O> </O> . وصرة الموت : شدة هجمه . **يقول** جل وجن : [ بيموتكم <O> </O> . عليه عنة **الفتنة** ويكسب ساعون لهم ] ( التوبة : 47 ) . يعزرون لهم ] ( التوبة : 47 ) . يعزرون : يعزرون لكم **الفتنة** **يقول** كتب بن زهير : إذا ما نلتنا <O> </O> . أبو سعيد : ومن هذا حديث أبي موسى حين أهدت عثمان . قال : ( إن هذه **الفتنة** **الفتنة** بعد مقتل عثمان . قال : ( إن هذه **الفتنة** ( بقرة كداء البطن لا يدرى إلى يولى له **الفتنة** بعد مقتل عثمان . قال : ( إن هذه **الفتنة** إذا عشيتم . وكفى في سلاحه إذا تحطى به المرح <O> </O> . الرحمن : 19 ) أي أحرأها **الفتنة** **يقول** <O> </O> . المشكلة . والمرج الفساد جماعات مختلفين وطوائف متباينين . سائلين إلى **الفتنة** <O> </O> . مثلين إلى الدنيا وزخرفها . لولا إذا قام . ومعه قول خليفة حين ذكر **الفتنة** قال : فإذا كان ذلك . وأما أبو الرازي فتن : جماع ملى <O> </O> . كان غير وطى **الفتنة** في كتم العرب الإنلاء والامتحان وأصلها والفتنة أشد من الفتن ] ( الشعراء : 191 ) فسمى **الفتنة** **سقطوا** ( التوبة : 49 ) أي الذين لم في التحلف أي استسلموا في <O> </O> . ( الحديد : 14 **الفتنة** **يقول** : أمتومها . قال : والفتنة الإسجال . : المذري عن طلب عن ابن الأعرابي أنه قال **الفتنة** الاختيار . والفتنة الصحة . والفتنة المال : بالأراء . والفتنة الإجراء بالذلل . **يقول** **الفتنة** المظن في التورق المظلم . **يقول** فتن ممتون ممتون يطلب الدنيا إلى قد عا في طلبها . وجماع **الفتنة** <O> </O> . في كتم العرب : الإنلاء والامتحان أبو زيد : فتن الرجل يعفن فتردا إذا وقع في **الفتنة** أو تحول من حال حسنة إلى حال سيئة . وفتن . مسرعاً بالجماعة الذين يرون رأيه ولم يجر على **الفتنة** والذلل في كتم العرب على وجوده . من ذلك .

شكل 1: الفهرس السياقي لكلمة "الفتنة" في الذخيرة النصية الفصحى.

الروضة بالرياض على الجهد الذي بذلوه في جمع وإدخال هذا الكم الكبير من كتب التراث العربي والإسلامي.

المراجع

[1] ابن عاشور، "التحرير والتنوير"، دار سحنون، تونس، 1997.

[2] C.D. Manning and H. Schuetze, Foundations of Statistical Natural Language Processing, 1st ed., The MIT Press, 1999.

[3] A. Elewa, "Did they translate the Qur'an or its exegesis?", 3rd Languages and Translation Conference and Exhibition on Translation and Arbization in Saudi Arabia, Riyadh, Saudi Arabia, 2009.

[4] M. Sahlgren, "The word-space model: using distributional analysis to represent syntagmatic and paradigmatic relations between words in high-dimensional vector spaces", Ph.D. dissertation, Department of Linguistics, Stockholm University, 2006.

[5] Z. Harris and H. Hiz, "Papers on syntax", Springer, 3-22, 1981.

[6] H. Rubenstein and J. Goodenough, "Contextual correlates of synonymy", Communications of the ACM, 8 (10), 627-633, 1965.

[7] P. Pantel, "Inducing ontological co-occurrence vectors", In Proceedings of the 43rd Conference of the Association for Computational Linguistics, ACL'05, 125-132, 2005.

[8] M. Sahlgren, "The distributional hypothesis", Rivista di Linguistica (Italian Journal of Linguistics), 20 (1), 33-53, 2008.

[9] A. Lenci., "Distributional semantics in linguistic and cognitive research", Italian Journal of Linguistics, 20(1), 1-31, 2008.

[10] J. Sinclair, Corpus and text - basic principles. in developing linguistic corpora: a guide to good practice, ed. M. Wynne. Oxford: Oxbow Books: 1-16, 2005, Available online at <http://ahds.ac.uk/linguistic-corpora/>

[11] A. McEnery and R. Xiao, Character encoding in corpus construction in developing linguistic corpora: a guide to good practice, ed. M. Wynne. Oxford: Oxbow Books: 47-58, 2005, Available online at <http://ahds.ac.uk/linguistic-corpora/>.

[12] متوفر على الموقع <http://www.sketchengine.co.uk>

5. الخاتمة

الذخيرة النصية الفصحى هي ذخيرة نصية رائدة قوامها أكثر من خمسين مليون كلمة تحمل ثقافة أمة، وهي مشروع له أبعاد علمية ولغوية وثقافية واجتماعية ودينية. تم جمعها وتصميمها بهدف خدمة الأبحاث العلمية في مجال اللغويات واللغويات الحاسوبية، إلا أنه يمكن استخدامها من قبل باحثين في مجالات أخرى كالتاريخ والأدب.

ستتيح الذخيرة النصية الفصحى للباحثين معرفة كيف استخدمت المفردات القرآنية في كلام العرب، كما ستتيح عمل فهرس سياقي لكلمات القرآن الكريم مما سيمكنهم من استنباط الفروق الدقيقة بين المترادفات واستنتاج أسرارها اللغوية، كما أنها ممكن أن تستخدم كنواة لاستخراج متصاحبات الألفاظ القرآنية ومترادفاتها وغيرها من الخصائص اللغوية آلياً عن طريق النماذج الدلالية المتوزعة، الأمر الذي سوف يفتح المزيد من الآفاق لتدبير كلام الله واستخراج المزيد من عجائبه.

شكر وتقدير

نود أن نتوجه بالشكر وننوه بجهود القائمين على موقع المكتبة الشاملة على الإنترنت وكذلك المكتب التعاوني لدعوة الجاليات فرع