



الاستفادة من لسانيات المدونة الحاسوبية في إعداد برامج لتعليم اللغة العربية للناطقين بغيرها.

ظافر بن علي المشهوري الشهري*

* أستاذ اللغويات التطبيقية، معهد تعليم اللغة للناطقين بغيرها - جامعة الملك عبدالعزيز

dalmshhori@kau.edu.sa

المخلص:

استفادت الدراسات الحديثة كثيراً من التقدم الحاصل في مجال علوم الحاسب ومن الإمكانيات الهائلة في تخزين المعلومات، وبناء قواعد بيانات للنصوص، وذلك لإنجاز مرحلة تجميع المعلومات، فظهر بذلك ميدان علمي تعددت مسمياته من نحو (لغويات المدونات) أو (لسانيات المدونات)، وظهرت المعاجم الإلكترونية للقواميس اللغوية، وللنصوص المكتوبة والشفوية، فكانت العلوم اللغوية والنفسية والاجتماعية تتسابق على هذه المدونات لدراساتها وتحليلها باعتبارها عينة تمثل المادة التحليلية في علومهم التي يسعون من خلالها لإثبات نظرية أو تفسيرها وفق أحكام وقوانين تلك العلوم، فكانت تتناول مصادر التحليل من القرآن الكريم والحديث الشريف، وكتب الأدباء ودواوين الشعراء، وبعض الروايات والقصص المشهورة، ولعل هذه الدراسة قد تفسح المجال للباحثين في علم اللغة النصي للاستفادة من هذا المجال في الدراسات الحديثة.

الكلمات المفتاحية: علم اللغة الحاسوبي، المدونة اللغوية، التحليل النصي، اللغة العربية للناطقين بغيرها.

Research Article

Utilizing Blog Computer Linguistics in Preparing Arabic Language Programs for Non-Native Speakers

Dhafer Ali Alshehri

King Abdulaziz University.

dalmshhori@kau.edu.sa

Abstract:

Recent studies have greatly benefited from the progress in the field of computer science and from its tremendous potentials in storing information and building databases for texts, in order to complete the stage of collecting information, thus a scientific field appeared with many names such as (linguistics Corpus) or (linguistics blog) and electronic dictionaries for written and oral texts. Linguistic, psychological and social sciences were competing for studying these corpora and analyzing them as a sample representing the analytical material in their sciences through which they sought to prove a theory or interpreting it according to the provisions and laws of those sciences. The researcher have surveyed researches and scientific studies and how it use linguistics corpus in textual analysis, but the researcher does not find - within his knowledge - whoever studies or analyzes the texts of corpus from the perspective of linguistics of text, since all previous textual studies dealt with the sources of analysis from the Holy Qur'an and Prophet Tradition, writers' books and poets' collections, and some famous novels and stories, perhaps this study may give a way to researchers in textual linguistics to benefit from this field in recent studies .

Keywords: Computational Linguistics, Corpus linguistic, Linguistic data analysis, Arabic for non-native speakers.

1. المقدمة:

استفادت الدراسات الحديثة كثيراً من التقدم الحاصل في مجال علوم الحاسب ومن الإمكانيات الهائلة في تخزين المعلومات، وبناء قواعد بيانات للنصوص، وذلك لإنجاز مرحلة تجميع المعلومات، فظهر بذلك ميدان علمي تعددت مسمياته من نحو (لغويات المدونات) أو (لسانيات المدونات)، وظهرت المعاجم الإلكترونية للقواميس اللغوية، وللنصوص المكتوبة والشفوية، فكانت العلوم اللغوية والنفسية والاجتماعية تتسابق على هذه المدونات لدراساتها وتحليلها باعتبارها عينة تمثل المادة التحليلية في علومهم التي يسعون من خلالها لإثبات نظرية أو تفسيرها وفق أحكام وقوانين تلك العلوم، وقد قام الباحث باستقصاء موسع للبحوث والرسائل العلمية ومدى الإفادة من المدونات اللغوية في التحليل النصي؛ إذ لم يجد الباحث - في حدود اطلاعه - من قام بدراسة وتحليل نصوص المدونات من منظور علم لغة النص، إذ إن جميع الدراسات النصية السابقة كانت تتناول مصادر التحليل من القرآن الكريم والحديث الشريف، وكتب الأدباء ودواوين الشعراء، وبعض الروايات والقصص المشهورة، ولعل هذه الدراسة قد تفسح المجال للباحثين في علم اللغة النصي للاستفادة من هذا المجال في الدراسات الحديثة.

وسوف أقوم في هذا المبحث باستعراض لمفهوم المدونة اللغوية والمدونة المحوسبة، ومجالات استخدامها في العلوم الحديثة، ثم أقوم بشرح موجز عن مشروع المدونة اللغوية لمتعلمي اللغة العربية، وأهم ما احتوته، وطريقة جمع النصوص إلكترونياً، ثم أختتم بالحديث عن البرنامج الذي سوف أستخدمه في التحليل النصي، وطريقة عمل البرنامج.

2. مفهوم المدونة اللغوية.

ورد في المعجم العربي الأساسي في مادة (دان): "دَوْنٌ يدون تدويناً الكتب جمعها ورتبها، ومُدَوَّنَةٌ جمعها مدَوَّنَات: مجموعة أحكام قانونية أو فقهية (مدونة القانون المدني)، ومدونة مالك: أشهر كتب الإمام مالك الفقهية". المنظمة العربية للتربية والثقافة والعلوم (1989) مادة دون.

أما في المعجم الكبير فقد شرح كلمة (الديوان) شرحاً طويلاً وأضاف: "والمدونة الكبرى مجموعة فقهية جمعها قاضي القيروان وأصبحت من أهم أدوات التشريع في الغرب الإسلامي كله، والمدونة المدنية مصنف يضم قواعد القانون المدني" مجمع اللغة العربية (2006) مادة دون.

وعرض هارتمان (Hartman) وستورك (Stork) معنى المدونة اللغوية وأهميتها ووظيفتها في معجمهما (Dictionary of language and linguistics) (معجم اللغة واللغويات)، فَوَرَدَ في مادة (Corpus)، أي المدونة ما نصّه: "هي بيانات لغوية غير منتظمة جُمعت خلال عمل ميداني، أو نصوص مكتوبة، ويقوم اللغوي بتحليل هذه البيانات ليكتب تقارير عن الملامح الوظيفية، أو الكتابية، أو النحوية، أو المعجمية للغة ما" (R.R.K.Hartman and F.C.Stork 1972).

وورد في معجم (A dictionary of linguistics and phonetics)، أي (معجم للغويات والصوتيات) ما يلي: "مجموعة بيانات لغوية سواء أكانت نصوصاً مكتوبة أم نسخة طبق الأصل عن كلام مسجل، ويمكن أن تعد هذه البيانات نقطة البداية للقيام بوصف لغوي، أو وسيلة لإثبات صحة افتراضات معينة عن اللغة" (David Crystal 1985. P77).

وفي معجم المصطلحات اللغوية، للمؤلف (رمزي البعلبكي) أورد تعريفاً للمدونة اللغوية (Corpus) فيقول: "هي مجموعة المواد اللغوية المدونة إما بالكتابة أو الصوت بغرض الدراسة والتحليل، التي تشكل بنية لغة بعينها" (البعلبكي، منير 1990 ص128).

ومن التعريفات السابقة يمكن تحديد دلالة المدونة بأنها تمثل النصوص المجموعة في عصرها أو عصور سابقة، وتتنوع مجالاتها ومستوياتها، إذ تُبنى المدونات اللغوية في الأصل لدراسة اللغة، ولايزال اختلاف المهتمين باللغة حول ما إذا كانت المدونة تمثل اللغة كلها أو تمثل جانباً محدداً منها، يقول (عودة خليل): "ولا أجد سبباً مقنعاً لهذا الاختلاف، إذ إن المدونة تُبنى وفق هدف أو

أهداف محددة للبحث، وربما يمكن لعدد من المدونات في الموضوعات المتقاربة وفق تخطيط دقيق أن تكون من تمثيل لغة ما تمثيلاً قريباً من الدقة والصواب" (بركة، بسام وآخرون 2014، ص251). وقد أجاب (علي القاسمي) عن هذه الملاحظة إجابة علمية دقيقة في مقالة له بعنوان: (الخطوة العلمية للمعجم التاريخي للغة العربية) جاء فيها: "نقترح أن تتألف مدونة المعجم التاريخي للغة العربية من بليون كلمة [مليار=ألف مليون كلمة]، وأن تقسم إلى عشرين مجاًلاً موضوعياً، ونصوص كل مجال من هذه المجالات تسمى مدونة فرعية (sub corpus)، وكل مدونة فرعية مقسمة هي الأخرى إلى مجموعة من المجالات الموضوعية المتخصصة، فإذا أخذنا المدونة الفرعية المتعلقة بمجال الرياضة مثلاً، نقسمها إلى عدة مجالات متخصصة، مثل: كرة القدم، كرة السلة ... بحيث يمكن القول إن تركيب هذه المدونة متوازنة موضوعياً" (عبدالعزیز، حسن 2008، ص2017).

ولحساسية موضوع قدرة المدونة على تمثيل اللغة، وعدم الاتفاق في شأنه يرى الباحث أن هذا الاختلاف عائد إلى عدم التفريق بين أنواع المدونات إذ إن تمثيل اللغة يختلف تحقيقه من مدونة إلى أخرى، والحقيقة أن محط اهتمام صاحب المدونة هي الأنماط الشائعة التي تظهر في استخدامات اللغة كما تُظهرها المدونات اللغوية على المستويات كلها: اللفظية والنحوية والصرفية والقدرة على دراسة اللغة بشكل جديد بفضل الإمكانيات التقنية المتاحة لمعالجة هذه النصوص.

ونظراً لهذا الاختلاف سوف يقوم الباحث بعرض موجز لأنواع المدونات، ذكرها (عبد المحسن الثبيتي) نقلاً عن العالم (سنكلير)، ثم أختتم هذا المبحث بالحديث عن مفهوم المدونات المحوسبة التي يعمل الباحث على تحليل نصوصها من وجهة نظر علم لغة النص.

ذكر (الثبيتي) سبعة أنواع للمدونة اللغوية، (بركة، بسام وآخرون 2014، ص297). وسوف نلقي الضوء عليها بصورة موجزة:

- المدونات المرجعية: هي المدونات التي تُصمم بحيث تعطينا معلومات مفصلة بقدر الإمكان عن استخدامات اللغة.
- المدونات الراصدة: وهي مدونات مشابهة للمدونات المرجعية، ولكنها تُنقَّح وتُحدَّث بنصوص جديدة بصفة مستمرة، لتعطي صورة عن تغيرات اللغة وتطور دلالات الألفاظ عبر الزمن من مثل: (الربيع العربي).
- المدونات المقارنة: وهي مجموعات متشابهة من النصوص من لغات عدة مثل: التعليقات السياسية الصحفية في الصحافة العربية والإنكليزية والفرنسية وقد تكون من لغة واحدة.
- المدونات المتوازية: مجموعة من النصوص وترجماتها إلى لغات أخرى، وتستخدم عادة في أعمال الترجمة.
- المدونات المتخصصة: وهي مجموعة من النصوص ذات طابع محدد، مثل المقالات العلمية في مجال الفيزياء، أو النصوص القانونية وغيرها.
- المدونات التاريخية: وهي مجموعة من النصوص لفترات زمنية معينة والهدف منها دراسة تطور اللغة.
- مدونات المتعلمين: وهي مجموعة من نصوص متعلمي اللغة مثل: كتابات الطلاب في التعليم العام، أو في معاهد تعليم اللغة العربية للناطقين بغيرها، وتشمل المدونة اللغوية المنطوقة والمكتوبة.

تلك المدونات السابقة أصبح لزاماً حوسبتها ثم تحليلها من خلال برامج خاصة في التعامل مع النصوص للاستفادة منها، والوصول للنص أو الكلمة أو الحرف بطريقة سريعة حتى يسهل على الباحث في أي مجال علمي من البحث والتحليل لتلك النصوص، من أجل إثبات نظرية أو تحليل نحوي أو نفسي أو غير ذلك، لذا وجب علينا أن نبين مفهوم المدونة المحوسبة ومجالاتها، وماذا قدمت للباحثين في مختلف العلوم.

3. مجالات استخدام المدونة اللغوية المحوسبة

1.3 مفهوم المدونة المحوسبة:

ذكر (الدكتور بن حمادو) تعريفاً للمدونة المحوسبة فيقول: " مجموعة ضخمة من النصوص اللغوية (منطوقة أو مكتوبة) يقع اختيارها من بين ما هو متوافر بحسب معايير محددة قصد استغلالها في إعداد دراسات لغوية مختلفة" (بركة، بسام وآخرون 2014، ص268). إذن عندما تتحول هذه النصوص إلى صيغتها الإلكترونية وتخزن في قواعد بيانات لتسهيل الوصول لها تسمى مدونة محوسبة، وذلك عن طريق برامج عديدة منها مفتوح المصدر، مثل برنامج (NOOJ) أو غيرها من البرامج الخاصة ببعض الشركات التجارية، وتتوحد هذه المدونات من حيث القيمة العلمية وجودة النتائج، وذكر الدكتور (حمادو) بعض المدونات المحوسبة المهمة في الوطن العربي، والتي يرى الباحث أنه يجدر بالباحث اللساني استغلالها في مجال البحوث التطبيقية في العصر الحديث، وهي (بركة، بسام وآخرون 2014، ص276):

- 1- المدونة اللغوية العربية لمدينة الملك عبد العزيز للعلوم والتقنية. (الرياض)
- 2- المدونة العربية العالمية. (الإسكندرية)
- 3- المدونة العربية مفتوحة المصدر. (معتر سعد)
- 4- مدونة العربية الحديثة. (جامعة ليدز)
- 5- مدونة المعجم التاريخي للغة العربية. (القاهرة)
- 6- مدونة الوطن والخليج. (مراد عباس)

ولا تزال البحوث مستمرة في هذا المجال رغم قلتها على مستوى اللغة العربية التي لم تنل حظها من الحوسبة الإلكترونية كبقية اللغات، لذا نجد الأستاذ الدكتور (محمود إسماعيل صالح) يؤكد على أهمية هذه المدونات التي تمثل اللغة المطلوب تحليلها، (صالح، محمود، 2015). فيقول: وتكمن أهمية المدونات المحوسبة فيما يلي:

- 1- الواقعية والتمثيل الحقيقي للغة.
- 2- الشمول من حيث المصادر والتنوعات والاستعمالات اللغوية والأساليب، والأجناس الأدبية، والتخصصات العلمية، والتقنية.
- 3- إمكانية إخضاعها للتحليل الإحصائي من جوانب مختلفة ولأغراض مختلفة، مثل التعرف على شيوع الكلمات، ومصاحباتها اللفظية وسياقات استعمالها (من خلال الكشافات السياقية).
- 4- التعرف على شيوع الكلمة وشيوع معانيها المختلفة ونسبة شيوع الكلمة مقارنة بمجموع الكلمات في المدونة، إضافة إلى شيوعها من عدمه في أنواع النصوص المختلفة، وهو ما يفيد في استخلاص المصطلحات الشائعة في كل تخصص من التخصصات العلمية والتقنية.
- 5- إمكانية التعرف على شيوع الأوزان والصيغ الصرفية المختلفة.
- 6- إمكانية إجراء أنواع من التحليل النحوي.
- 7- إمكانية إجراء التحليل الصوتي (بوصف الحروف تمثيلاً للأصوات العربية من حيث شيوعها ومواقعها في الألفاظ إلى غير ذلك).

كما لا يفوتنا في هذا المبحث ذكر العالم الكبير الدكتور (نبيل علي) الذي فتح آفاقاً واسعة للمدونات المحوسبة في اللغة العربية، وذلك بإقامة نماذج حاسوبية لفهم الأداء الشامل لمنظومة اللغة العربية، بالإضافة لتصميم نظم المعالجة الآلية للغة العربية في مؤلفاته التي جاء من أبرزها كتاب (اللغة العربية والحاسوب) والذي يعد من المصادر المهمة في إعداد المدونات في اللغة العربية، وخاصة مجال استخدام الحاسوب في الإحصاء اللغوي والتحليل الذي وضعه الدكتور (نبيل علي) قواعده في حديثه عن المعالجة الآلية للغة العربية، (علي، نبيل 1988).

2.3. مجالات استخدام المدونة المحوسبة:

اللسانيات الحاسوبية هي أحد فروع اللسانيات التي تهتم بالإفادة من معطيات الحاسوب في دراسة قضايا اللسانيات المتعددة، ويمثل الدكتور (مازن الوعر) لهذه الظواهر اللغوية فيقول: "وتهدف إلى رصد الظواهر اللغوية وفقاً لمستوياتها الصوتية والصرفية والنحوية والبلاغية والعروضية، وتهدف كذلك إلى معالجة عمليات تحليل الظواهر اللغوية والنصوص التراثية وإجراء العمليات الإحصائية وصناعة المعاجم والترجمة الآلية وتعليم اللغات وإجراء التحليل على السياقات اللغوية المتعددة، إلى غير ذلك من العمليات الحاسوبية المتعلقة باللغات الطبيعية، واستعمال الحاسوب يكفل الدقة العلمية والسرعة وتوفير الجهد" (الوعر، مازن 1989، ص 25).

ذكر الباحث (الفيفي) في مقدمة مشروعه عن المدونة اللغوية لمتعلمي اللغة العربية بعض المجالات لاستخدام المدونة الطلابية المحوسبة والتي تركز على تعلم واكتساب اللغة وصناعة معاجم الطلاب، فذكر سبع مجالات مهمة، نذكر بعضاً منها بشكل موجز، وهي: (العصيمي، صالح وآخرون 1436هـ)

- 1- الأبحاث في مجال اكتساب اللغة: وقد أجريت دراسات حول مدى الاستفادة من مدونات المتعلمين ذكر منها الباحث (الفيفي) دراستين أجرتهما (قرانجر 1988م) لفهم آلية اكتساب اللغة الثانية، فهي تعطي حسب قول (ييب 1995م) صورة أشمل لبحث جوانب من مثل الأشكال اللغوية التي يصعب على الطلاب تعلمها واستخدامها بشكل طبيعي.
- 2- تأليف معاجم الطلاب: أصبحت صناعة المعاجم المتخصصة لتعلم اللغة - وخصوصاً الإنجليزية - تعتمد على نتائج تحليل المدونات الطلابية في تصحيح الأخطاء الشائعة لدى متعلمي اللغة، ويعد معجم (Longman Essential Activator) (1997م)، أول معجم يعتمد على مدونة طلابية، حيث استُخدمت مدونة طلاب (لونجمان) في تأليف هذا المعجم من خلال البحث عن طريقة استخدام الطلاب للكلمات التي اشتمل عليها المعجم.
- 3- تدريس اللغة وتصميم المواد التعليمية: حيث ترى (قرانجر 2003م) أنه قد بدأ العمل فعلياً على إنتاج بعض كتب تعليم اللغة المبنية على نتائج أبحاث المدونات الطلابية، ورغم بطء هذا المجال لحدائته إلا أن هناك جهوداً لمدونات طلابية تساعد مصممي المواد التعليمية، فيمكن للمدونات أن تساعد على اختيار وتصنيف مواد تعليم اللغة حسب مستويات الطلاب اللغوية وذلك من خلال تعرفهم على الشائع من المفردات والتراكيب اللغوية.

4. مشروع المدونة اللغوية لمتعلمي اللغة العربية.

أطلق هذا المشروع باحث الدكتوراه في (جامعة ليدز) الدكتور (عبدالله الفيفي) تحت إشراف الدكتور (إرك أتول) أستاذ الحاسب الآلي بالجامعة (والمدونة مشروع مفتوح المصدر ومرخصة ضمن Creative Commons Attribution، ومتوفرة على الشابكة: <http://www.arabiclearnercorpus.com>)، وتضمن المشروع التعريف بالمدونات اللغوية لمتعلمي اللغة العربية بعد الاطلاع على (159) مدونة تعليمية حول العالم للوقوف على أفضل النماذج، وكان الهدف من هذا المشروع توفير مدونة لغوية لنصوص حررها متعلمو اللغة العربية في المملكة العربية السعودية، وتتألف من مجموعة من المواد المكتوبة والمنطوقة، والمدونة مشروع مفتوح المصدر، ومرخص ضمن (Creative Commons Attribution)، (يمكن الاطلاع على الترخيص والاتفاقية على الرابط: <http://creativecommons.org/licenses/by-nc/4.0/legalcode>) وسوف نتحدث في هذا المبحث عن أبرز ما تضمنته هذه المدونة وطريقة بنائها، وتنسيق النصوص بشكل موجز، وكيفية الاستفادة الباحث من هذه النصوص في مجال الترابط النحوي والتماسك النصي.

1,4. المدونة اللغوية لمتعلمي اللغة العربية: (ALC: Arabic Learner Corpus)

أُنشئت المدونة اللغوية لمتعلمي اللغة العربية لتكون مصدرًا لغويًا لأبحاث تعليم اللغة العربية، وكذلك للبحث في مجال معالجة اللغة الطبيعية، ويمكن تحميل نصوص المدونة من شبكة الإنترنت مجانًا، وهي تتألف من مجموعة من المواد المكتوبة والمنطوقة التي حررها متعلمو اللغة العربية في المملكة العربية السعودية خلال العامين (٢٠١٢م) و (٢٠١٣م) وتضم (١٥٨٥) نصًا (٢٨٢،٧٣٢) كلمة، شارك في تحريرها (٩٤٢) طالبًا من (٦٧) جنسية، و (٦٦) لغة أم مختلفة، يدرسون في (٢٥) مؤسسة تعليمية ويبلغ متوسط طول النص الواحد (١٧٨) كلمة. وتضم نصوص المدونة نوعين أدبيين، الأول: سردي يحكي فيه الطالب رحلته في إحدى الإجازات، والثاني: يناقش فيه الطالب اهتماماته الدراسية ومستقبله العلمي. ويستطيع الباحثون من خلال المعلومات الوصفية للطلاب (metadata) التعرف على خصائص النص اللغوي، كذلك مؤلف النص مما يضيف عمقًا أكثر لتحليل بيانات المدونة، كما تفيد أيضًا في مجالي التعرف الآلي على النص المكتوب بخط اليد، والتعرف على الكلام المنطوق. كما أن أصول الأوراق - المكتوبة يدويًا من قبل الطلاب - متاحة للتنزيل بعد أن أُدخلت عن طريق الماسح الضوئي، ثم حُفظت في ملفات بصيغة (PDF). وكذلك الحال بالنسبة للتسجيلات الصوتية لأولئك الطلاب الذين أذنوا بنشرها على الإنترنت فهي متاحة للتحميل في صيغة صوتية رقمية (MP3)، ما مجموعه ثلاث ساعات ونصف تقريبًا. وقد تم جمع المعلومات والأرقام السابقة من مقال لصاحب المدونة، انظر: (العريان، يوسف 1436 هـ، ص32).

سُمِّيَتْ ملفات المدونة بطريقة تسمح بالتعرف على الخصائص الأساسية للنص وكذلك المؤلف، نحو: S038_T2_M_Pre_NNAS_W_C) وهي بالترتيب من اليسار: رقم الطالب، رقم النص، جنس الطالب، المرحلة العامة، ناطق بالعربية باعتبارها لغته الأم أو ناطق بغيرها، نوع النص (مكتوب أم منطوق) مكان تحرير أو تسجيل النص (في الصف أو في المنزل). ويمكن تنزيل نصوص المدونة من موقعها الرئيس: (<http://www.arabiclearnercorpus.com>) كما يمكن البحث فيها عن طريق الإنترنت من خلال الموقع التالي: (<http://www.alcsearch.com>) أو موقع (Sketch Engine). وتم إنشاء موقع خاص للبحث من قبل الباحث (الفيفي) على العنوان التالي: (<http://arabicorpus.byu.edu/>).

استُخدمت المدونة اللغوية في تقنية تهدف للتعرف على النصوص التي حررها الناطقون باللغة العربية، وتمييزها عن تلك التي حررها الناطقون بغير العربية، وكذلك تصحيح الأخطاء آليًا. حيث تم إنشاء خمسة أنواع من الملفات، (هذه المعلومات التقنية تم اختصارها بشكل موجز، وهي متاحة للاطلاع في موقع المدونة على الشبكة: <http://www.arabiclearnercorpus.com>) كما يلي:

- (1) ملفات نصية دون المعلومات الوصفية.
- (2) ملفات نصية مع المعلومات الوصفية باللغة العربية.
- (3) ملفات نصية مع المعلومات الوصفية باللغة الإنجليزية.
- (4) ملفات بلغة الترميز الممتدة (XML) مع المعلومات الوصفية باللغة العربية.
- (5) ملفات بلغة الترميز الممتدة (XML) مع المعلومات الوصفية باللغة الإنجليزية.

وأنشأ الباحث جدولاً لتصنيف الأخطاء لاستخدامه في رسم الأخطاء اللغوية في مدونات المتعلمين العربية، وقد استفاد هذا التصنيف من مجموعة من تصنيفات الأخطاء المصممة لمدونات المتعلمين باللغة الإنجليزية، ليكون مناسباً للأخطاء اللغوية التي يقع فيها متعلمو اللغة العربية، كما استُخدم هذا التصنيف لتطوير أداة هي عبارة عن محرر لوسم أخطاء الطلاب في المدونة اللغوية لمتعلمي اللغة العربية، والذي يمكن الاستفادة منه في رسم أي مدونة من مدونات المتعلمين.

5. التحليل النحوي والنصي لكلمات المدونة.

1.5. (الكشاف السياقي / concordance).

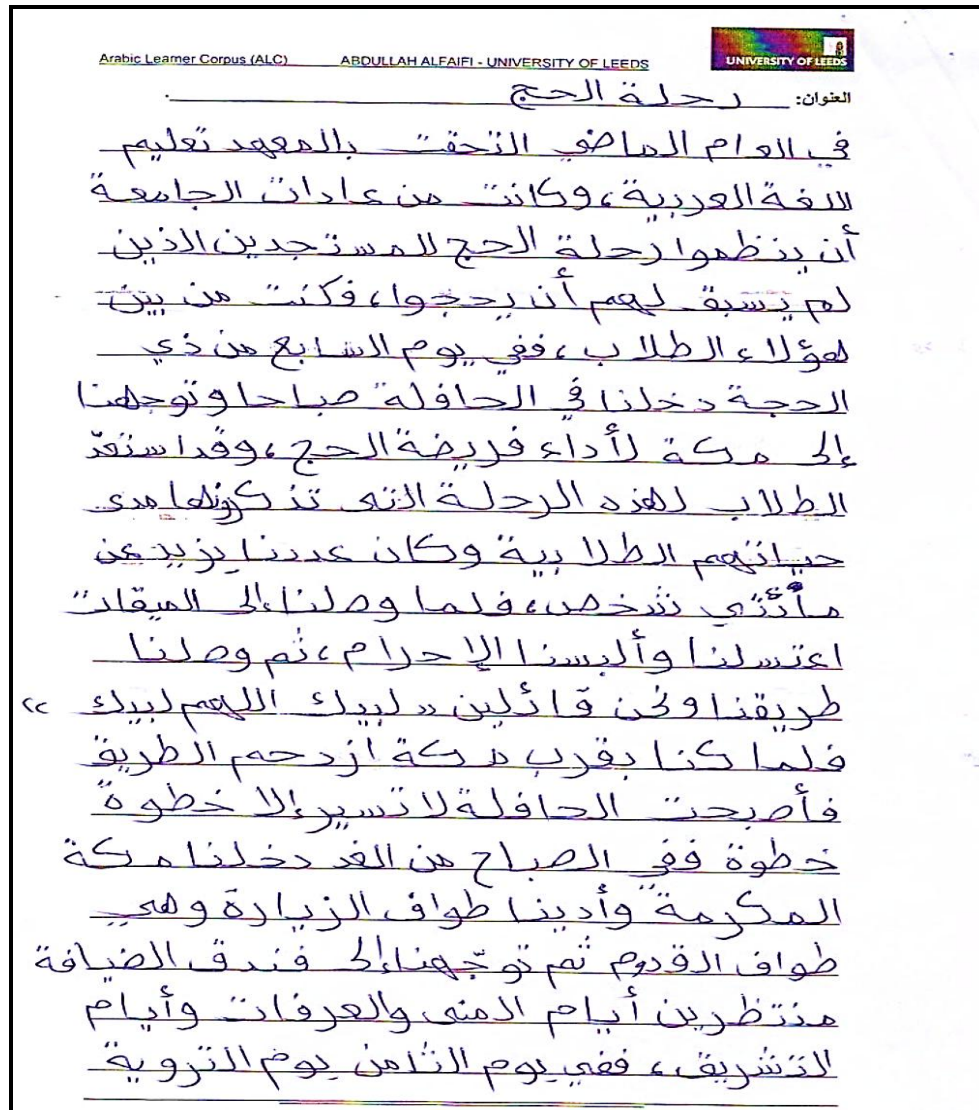
كما أشرنا سابقاً في بداية البحث بأن المدونات المحوسبة كانت محط اهتمام اللسانيات اللغوية، وخصوصاً في علم اللغة التقابلي وتحليل الأخطاء، كونها تساعد في عملية تعلم اللغة واكتسابها، وتصميم المواد التعليمية لمتعلمي اللغة، وقد رأى الباحث الاستفادة من هذه المدونة في مجال علم لغة النص من خلال ترابط النصوص وتماسكها لمتعلمي اللغة العربية الناطقين بلغة ثانية، وذلك من خلال الكشف عن الدور المتوقع حول قياس مدى استعمال متعلمي العربية لغة ثانية لأدوات الترابط النحوي والتماسك النصي، وما يرتبط بذلك من الحكم على النص بالجودة أو عدم الانتظام والتفكك ومدى مساهمة ذلك في تحسين العملية التعليمية لدى متعلمي العربية لغة ثانية في المستويات الدراسية.

وقد اعتمد الباحث في اختيار عينة الدراسة من المدونة اللغوية لمتعلمي اللغة العربية على الطريقة الطبقيّة المنتظمة (Systematic Sampling)، إذ قام الباحث بجمع نصوص كل جنسية في مستند نصي وترميزها باللغتين العربية والإنجليزية تسهيلاً لإدخالها في برنامج الكشف السياقي (Concordance)، على النحو التالي:

(ط001-ج1) - (S001-N1)، فحرف الطاء يعني نص الطالب رقم واحد، وحرف الجيم يعني الجنسية، ثم قمتُ باختيار عينة ممثلة من المدونة ومنتظمة من الأعداد الفردية من نصوص كل جنسية بما يعادل (25%) ليكون مجموع نصوص الدراسة التطبيقية لجميع الجنسيات (200) نص.

وفيما يلي نموذج لنص أحد الطلاب في المدونة اللغوية لمتعلمي اللغة العربية (وهي متاحة للتحميل والاطلاع من موقع المدونة على الشبكة www.arabiclearnercorpus.com)، بالإضافة لجميع المعلومات الخاصة بكل نص:

[العمر: 21-الجنس: ذكر]-[الجنسية: مالي]-[اللغة الأم: السونينكية دولة مالي]-[أصالة اللغة: ناطق بغير العربية]-[عدد اللغات التي يتحدثها: 3]-[عدد سنوات تعلم اللغة العربية: 13]-[عدد السنوات في بلدان عربية: 1]-[المرحلة العامة: ما قبل الجامعة]-[المرحلة الدراسية: برنامج الدبلوم] - [السنة/المستوى: المستوى الأول]-[المؤسسة التعليمية: معهد اللغة بجامعة الإمام]-[معلومات النص: النوع: سردي]-[مكان التأليف: في الصف]-[سنة التأليف: 1435هـ]-[دولة التأليف: السعودية]-[مدينة التأليف: الرياض]-[محدد بوقت: نعم]-[استخدام مراجع: لا]-[استخدام كتاب قواعد: لا]-[استخدام معجم أحادي: لا]-[استخدام معجم ثنائي: لا] [استخدام مراجع أخرى: لا]-[الشكل: مكتوب]-[الوسيط: مكتوب يدوياً]-[الطول: 218 كلمة].



شكل رقم 1-1 نموذج لنص أحد الطلاب مكتوبة بخط اليد.

2.5. الكشف السياقي: (concordance)

تحتاج أي مدونة لغوية إلى برامج خاصة للتعامل مع النصوص واستخراج ما يتطلبه مجال الدراسة من معلومات ومعارف، ويعد برنامج (a Concorde) موقع البرنامج على الشبكة: <http://www.andy-roberts.net/coding/aconcorde>. وهو جهود مثمرة لبرامج تحليل نصوص المدونات العربية، ويمكن الاطلاع على تطوير البرنامج في بحث (Roberts, A., Al-) (Sulaiti, L., & Atwell, E. 2006) ومن أشهر البرامج التي تهتم بالمسح اللغوي لكل الكلمات في المدونة والسياق الذي وردت فيه، ويمكن الرجوع للدراسة التي أجراها (الفيفي وأتويل) لمقارنة ست أدوات لتحليل المدونات اللغوية العربية - ومنها a Concorde (Alfaifi, Abdullah and Eric Atwell (2015)). ويقول الصوينع: "وكان ظهور مجال لغويات الحاسب الآلي (computational linguistics) - الذي يُعنى بتطبيق تحليلات الحاسب على التركيبات والمعاني اللغوية - من أحد الركائز التي جعلت التحليل الآلي للغة موضوعاً للبحوث الجادة في الخمسينيات عند الغرب، وقد أصبحت الكشافات السياقية جزءاً من ظاهرة عالمية تنمو سريعاً تسمى بالحاسب والإنسانيات (Humanities and Computer) التي تؤثر في كل ما يتعلق بعلم اللغويات" (الصوينع، علي، ص17).

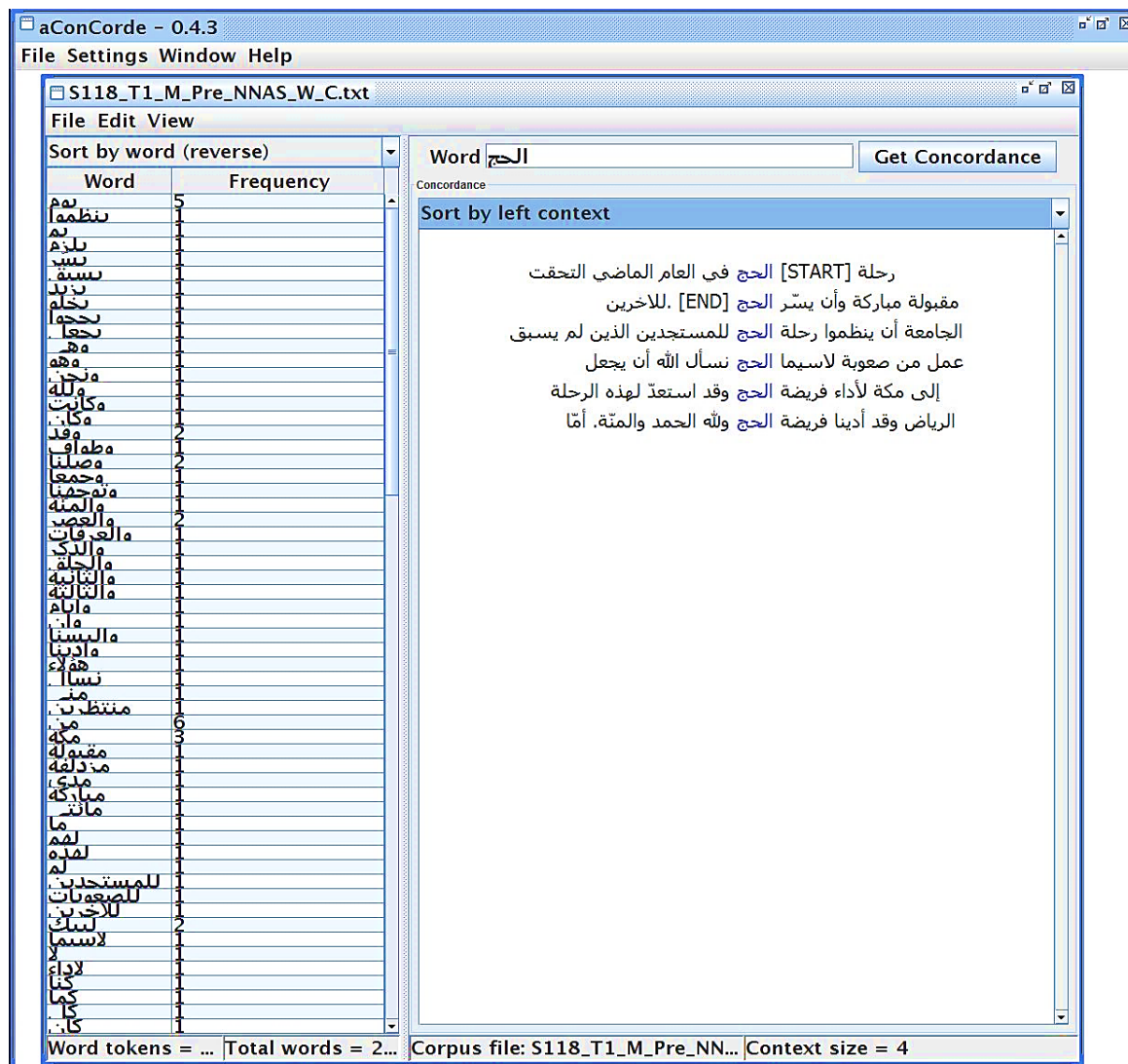
وقد صمم الباحثون في جامعة (ليدز) مجموعة من الأدوات المخصصة لمعالجة اللغة العربية، يقول (الفيفي): " ولهذا فقد بُنيَ كشاف السياقات (a Concorde) في قسم الحاسب الآلي بجامعة (ليدز) وهو أداة مجانية ومفتوحة المصدر، صممت في المقام الأول لمساعدة الباحثين في لسانيات المدونات العربية (Roberts, et al. 2005, 2006). وقد أشادت دراسة علمية أجريت حول كشافات السياقات في ذلك الوقت (Wiechmann and Fuhs, 2006) ببرنامج (a Concorde) الذي يوفر دعمًا كاملاً للنصوص العربية، ويتمثل هذا الدعم في عدة جوانب منها واجهة المستخدم التي يمكن تحويلها بين اللغتين العربية والإنجليزية، وتبنيته لترميز الحرف المعياري (Unicode)، إضافة إلى دعمه لمجموعة من الخطوط العربية وكذلك اتجاه النص العربي من اليمين إلى اليسار " (العريان، يوسف 1436هـ، ص32).

ونذكر هنا بعض الفوائد المهمة في استخدام الكشاف السياقي في التحليل النصي ذكرها ((الصوينع، علي، ص17)). ومنها:

- 1- انجاز كشافات أو مسارد سياقية (concordance) تأخذ المستفيد إلى مواضع كلمات بعينها في النص.
- 2- معرفة الطرق المختلفة لبداية الجمل ونهايتها في نص ما، وذلك من خلال البحث عن علامة (الترقيم- النقطة) داخل النص.
- 3- تجمع كل السياقات التي وردت الكلمة فيها بالنص في مكان واحد، وهذا يجعل مقارنة فقرات النص ببعضها أمراً سهلاً.
- 4- تجميع التعبيرات والمصطلحات المترادفة أو المتضادة في النص.

وسوف يعتمد الباحث الإصدار الحالي من برنامج (Aconcorde-0.4.3) تصميم: (Roberts Andrew) ، والذي

بالإمكان تحميله من موقع: <http://www.andy-roberts.net/coding/acorde>



شكل رقم 1-2 صورة لتحليل نص أحد الطلاب عن طريق برنامج (aConcorde).

6. الخاتمة

تعد المدونة اللغوية لتعليمي اللغة العربية- التي هي محل التطبيق- مجالاً خصباً للدراسة اللغوية لنصوص متعلمي اللغة العربية لغة ثانية، غير أن كثرة الأخطاء النحوية والإملائية كانت عائقاً لدى الباحث وخاصة في تحليل نصوص التماسك النصي. ومن أبرز نتائج هذا البحث ما يلي:

6. تكمن أهمية الربط الإحالي وتعامله مع النصوص في وجود بعض العناصر اللغوية التي لا تكفي بذاتها، مما يجعل من الضروري العودة إلى ما تشير إليه أو تحيل عليه من أجل تأويلها، ونجد ذلك واضحاً في نصوص الإحالة إذ بلغت نسبة عدم تحقق السلامة التركيبية في نصوص العينة (41,5%) من إجمالي نصوص الإحالة المجموعة، مما يعني ضعفاً في سبك النص واستمرارية المفاهيم، إذ إن الإحالة عملية مرجعية تلعب العناصر اللغوية من مثل الضمائر وأسماء الإشارة وغيرها دوراً في تحديد الذات التي تعود عليها، وهو ما شكل صعوبة لدى متعلمي اللغة العربية لغة ثانية، وكان واضحاً في كتاباتهم.

7. تأتي أهمية الوصل باعتباره بنية للنص يربط مجموعة من الجمل أو المتواليات المتعاقبة، وعليه حلّ الوصل في المرتبة الثانية في نسبة عدم تحقق السلامة التركيبية إذ بلغت (5،12%) من إجمالي نسبة ورودها في نصوص العينة مما يمثل ضعفاً في توظيف هذه الأداة بالشكل الصحيح، وبالتالي كان أثره واضحاً في ضعف السبك داخل نصوص العينة.
 8. أظهرت الدراسة ضعفاً في استخدام أدوات الترابط النحوي إذ بلغت نسبة عدم تحقق السلامة التركيبية (5،61%) من إجمالي نصوص العينة، بينما بلغت نسبة تحقق السلامة التركيبية (5،38%) من إجمالي نصوص العينة، وهذا يشير إلى ضعف في سبك نصوص العينة.
 9. حظيت وسيلة التقابل المضموني العكسي باهتمام جل النحويين والبلاغيين العرب، وفي نصوص العينة احتلت هذه الوسيلة المرتبة الأولى إذ بلغت نسبة ورودها (5،25%) من نصوص العينة، وجاءت نسبة تحقق السلامة المضمونية فيها (23%) من نصوص العينة، فساعدت هذه الوسيلة في تماسك نصوص العينة.
 10. جاءت وسيلة الارتباط الافتراضي في المرتبة الثانية في نصوص العينة إذ بلغت نسبة ورودها (5،20%) من نصوص العينة، وجاءت نسبة تحقق السلامة المضمونية فيها (18%) من نصوص العينة، فساعدت هذه الوسيلة في تماسك نصوص العينة.
 11. أظهرت الدراسة توازناً في استخدام وسائل التماسك النصي إذ بلغت نسبة عدم تحقق السلامة التركيبية (22%) من إجمالي نصوص العينة، بينما بلغت نسبة تحقق السلامة التركيبية (78%) من إجمالي نصوص العينة، مما يشكل توازناً في حيك نصوص العينة.
 12. يلاحظ في هذه الدراسة ضعف السبك النصي مع توازن الحيك النصي في نصوص العينة، مما يفقد الوحدة والترابط والانسجام للبنية اللغوية لنصوص العينة، فلا يمكن عزل السبك عن الحيك لارتباطهما بالأداء اللغوي والصحة اللغوية في مسرح النص.
 13. ومن أهم المقترحات والتوصيات التي أرجو أن تلقى القبول:
- يقترح الباحث على أقسام تعليم اللغة العربية لغة ثانية اهتمامها بمثل هذه الدراسات النصية لقياس مدى تمكن الطلاب من تحقيق السلامة التركيبية والمضمونية في النصوص المكتوبة أو المنطوقة.
 - يقترح الباحث عمل معجم لغوي لجميع المناهج المقررة في معاهد اللغة وقياس استخدام الطلاب للتراكيب اللغوية، ومدى استفادتهم وتوظيفهم لمفردات المعجم بطريقة صحيحة في النصوص.
 - التراجع عن الوقوف عند الصور التقليدية التي تعالج النصوص على شكل الجمل والأبنية والوحدات الكبرى، فالنص مكون بلاغي تتحقق كينونته في أبنيتة اللغوية الخاصة ودلالاتها المجازية لأن كل ما يصير نصاً هو بلاغي.
 - إذا أدرجنا نتائج البحوث القائمة على مبادئ ومفاهيم لسانيات النص في المناهج الدراسية لتعليم اللغة العربية لغة ثانية، فإننا سنتوصل إلى نتيجة مهمة هي الرفع من قدرة متعلم اللغة الثانية إلى الفهم الصحيح للنصوص العربية الأصلية، وإنشاء نصوص تقترب منها، وهذا من شأنه أن يؤدي إلى تحسين فهمه لقواعد العربية وتحسين مستوى التواصل بينه وبين متحدث اللغة الأم.
 - يقترح الباحث لحل مشكلة ضعف السبك والحيك في النصوص إدراج تطبيقات لسانية في المناهج الدراسية مما يساعد متعلم اللغة الثانية على الفهم النظري والتطبيقي للغة الهدف، ويمكن ذلك بإدراج مقرر تعليمي بعنوان (لسانيات النص في العربية المعاصرة)، فمن خلال نتائج الدراسة يرى الباحث ضرورة تضمين موضوعات قد تساعد في حل مشكلة عدم تحقق السلامة التركيبية والمضمونية في النصوص، ومن تلك الموضوعات ما يلي:

- 2- الصياغة اللغوية للنص.
- 3- الربط بين دلالات جمل النص.
- 4- البنية الفعلية للنص (نظرية الفعل الكلامي).
- 5- المضمون النصي (السرد، الوصف، الحجاج...).
- 6- أنواع النصوص (علمية، أدبية، صحفية...).

ويكون الفهم التطبيقي لهذا المقرر من خلال تمارين تطبيقية بترتيب نصّ عربي أصيل من خلال سبك النص وحبكه في صورته الأصلية، أو تلخيص النص أو ترجمة نصّ من لغته الأم إلى العربية لقياس الفهم التطبيقي لدى متعلم اللغة وغيرها من التمارين التي تجعل متعلم اللغة الثانية يستكشف تراكيب النص العربي ومضمونه، ليعرف أن النص العربي يختلف عن لغته الأم وخصوصاً في الترابط بين الجمل، إذ وجد الباحث أن أغلب نصوص العينة يعتمد فيها الطالب على فهمه بلغته الأم فتظهر كأنها ترجمة حرفية لما يريد التعبير عنه مما أوقع الكثير من تلك النصوص في ضعف الترابط النحوي والتماسك النصي.

وأخيراً أود أن أقول:

هذا البحث المتواضع والمحاولة في دراسة المدونات التعليمية دراسة نصية أردت أن أشارك بها في ميدان البحث العلمي خدمة للدراسات اللغوية النصية وما هو إلا جهد المقل، ولا أدعي فيه كمالاً، فالنقص سمة البشر والكمال المطلق لله تعالى وحده. ربنا لا تؤاخذنا إن نسينا أو أخطأنا، وآخر دعوانا أن الحمد لله رب العالمين.

المراجع:

- [1] البعلبكي، رمزي (1990). *معجم المصطلحات اللغوية: إنكليزي-عربي*. بيروت: دار العلم للملايين.
- [2] بركة، بسام، وآخرون. (2014). *نحو معجم تاريخي للغة العربية* (ط. 1). قطر: المركز العربي للأبحاث ودراسة السياسات.
- [3] الصوينع، علي. (1987). *كشافات النصوص وتطبيقاتها في نصوص كشافات القرآن والحديث*. مجلة المكتبات والمعلومات العربية، 7. (4)
- [4] عبدالعزيز، محمد (2008). *المعجم التاريخي للغة العربية: وثائق ونماذج*. القاهرة: دار السلام.
- [5] العريان، يوسف (1436هـ). *الحرف العربي والتقنية: أبحاث في حوسبة العربية*. الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية.
- [6] العصيمي، صالح، وآخرون. (1436هـ). *المدونات اللغوية: بناؤها وطرائق الاستفادة منها* (ط. 1، ص ص. 114-125). الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية.
- [7] علي، نبيل (1988). *اللغة العربية والحاسوب: دراسة تحليلية* (ص ص. 113-244). الرياض: مطبعة تعريب.
- [8] المنظمة العربية للتربية والثقافة والعلوم. (1989). *المعجم العربي الأساسي للناطقين بالعربية*. لاروس.
- [9] مجمع اللغة العربية. (2006). *المعجم الكبير*. القاهرة: مجمع اللغة العربية، مادة [دون].
- [10] الوعر، مازن (1989). *دراسات لسانية تطبيقية* (ط. 1). دمشق، سوريا: دار طلاس.
- [11] برنامج (a Concorde) على الشبكة: <http://www.andy-roberts.net/coding/aconcorde> . تاريخ الاستطلاع: 2016/12/6م، الساعة: 11 صباحاً GMT.
- [12] ترخيص المدونة والاتفاقية: <http://creativecommons.org/licenses/by-nc/4.0/legalcode> (تاريخ الاستطلاع: 2015/12/4م، الساعة: 10 صباحاً GMT).
- [13] صالح، محمود إسماعيل. الجانب اللغوي للمعجم الحاسوبي للغة العربية، ضمن البحوث المقدمة في اجتماع خبراء المعجم الحاسوبي التفاعلي، انظر مدونة الدكتور على الشبكة: <http://dr-mahmoud-ismail-saleh.blogspot.com/> (تاريخ الاستطلاع: 2015-12-4م، الساعة 3 GMT).
- [14] المدونة اللغوية لمتعلمي اللغة العربية: <http://www.arabiclearnercorpus.com>، (تاريخ الاستطلاع: 2015/12/4م، الساعة: 10 صباحاً GMT).

المراجع الأجنبية:

- [1] Alfaifi, A., & Atwell, E. (2016). Comparative evaluation of tools for Arabic corpora search and analysis. *International Journal of Speech Technology*, 19(2), 347–357. <https://doi.org/10.1007/s10772-016-9365-6>
- [2] Crystal, D. (1985). *A dictionary of linguistics and phonetics* (2nd ed.). Oxford: Basil Blackwell & Andre Deutsch.
- [3] Hartmann, R. R. K., & Stork, F. C. (1972). *Dictionary of language and linguistics*. London: Applied Science Publishers.
- [4] Roberts, A., Al-Sulaiti, L., & Atwell, E. (2006). AConCorde: Towards an open-source, extendable concordancer for Arabic. *Corpora*, 1(1), 39–60. <https://doi.org/10.3366/cor.2006.1.1.39>