استثمار المدوّنات النصّية المحوسبة للذّخيرة العربيّة في الصّناعة المعجميّة

Investing Computational Text Corpus in Lexicography

سارة لعقد
جامعة الجزائر 2-الجزائر
sara.laked@univ-alger2.dz

تاريخ الإرسال: 22/ 06/ 2022-تاريخ القبول:2022/10/25- تاريخ النشر: 2023/06/30

الملخص:
لقد بدأت أفكار الأستاذ عبد الرحمان الحاج صالح في الذخيرة العربية تصبح واقعا علميا ملحوظا، فمع التطور التكنولوجي الذي يعرفه العالم بأسره بات بإمكاننا اليوم القيام بجمع أكبر عدد من المدوّنات النصية العربية في وقت قصير جدا باستخدام الحاسوب، وعليه فإّن العمل على إنجاز الذخيرة العربية لابدّ أن ينطلق من وضع المدوّنة النصية المحوسبة ممّا هو مستعمل من اللغة العربية وذلك بمحاولة جمع أكبر عدد ممكن من المدوّنات النصية بطريقة آلية؛ منها: مدوّنة الجرائد العربية ومدوّنة قصص الأطفال، ومدوّنة الرواية العربية، ومدوّنة الكتب المدرسية، والغرض من هذا هو جمع مدوّنة نصيّة محوسبة تكون أساسا في الصناعة المعجمية. ومن هنا نطرح الإشكال التالي: كيف يمكن استثمار المدوّنات النصّية المحوسبة للذخيرة في الصناعة المعجمية؟

الكلمات المفتاحية: الحاسوب؛ المدوّنة النصّية؛ مشروع الذخيرة العربية؛ المدوّنة النصّية المحوسبة؛ الصناعة المعجمية.

Abstract :

The ideas of Professor Abderrahman Hadj Salah concerning Dhakhira al-Arabiyya project have begun to be realized. The technological advancement, witnessed by the whole world knows, enables us to constitute large textual Arabic language corpora in a very short time using the computer and various programs and applications. In fact, a number of corpora have already been constituted: the corpus of Arabic newspapers, the corpus of children stories corpus, the Arabic novel corpus, and the corpus of textbooks. The purpose being to help lexicographers in their work.

Keywords : Computer; Text Corpus ; al-Dhakhira al-Arabiyya ; Computational Text Corpus ; Lexicography

.

مقدمة:

إنّ الانفجار التكنولوجي الذي نعيشه في السنوات الأخيرة جعل الكثير من الباحثين في مجال العلوم اللغوية يفكر في الكيفية التي يمكن من خلالها حفظ التراث اللغوي والثقافي العربي، خصوصا مع ظهور كمٍّ هائل من الأجهزة التي تمكّن من حفظ جملة من المعطيات اللغوية التي يمكن الاستفادة منها وتخزينها وحتى معالجتها. ونظرا لأهمية اللغة العربية بالنسبة لأهلها فإنّهم عملوا جاهدين على ترقيتها، ولعل أبرز ما سعوا في سبيله هو محاولاتهم الجادة من أجل رقمنتها، ذلك لأنّ العالم أصبح اليوم يسير في اتجاه رقمي في شتّى المجالات إن لم نقل كلّها. وعندما نتحدث عن رقمنة اللغة العربية لابدّ أن نتحدث عن الأستاذ عبد الرحمان الحاج صالح اعتبارا من كونه من أبرز علماء اللغة العربية الذين حاولوا استغلال الحاسوب في سبيل رقمنتها، ما دفعه إلى التفكير في محاولة استثمار مخرجاته - الحاسوب- والاستعانة بمختلف البرامج والتطبيقات المتاحة من خلاله في إنجاز مشروع الذخيرة العربية؛ وفي هذا السياق قال في ورقة بحثية قدّمها في أحد المؤتمرات التي انعقدت بالمملكة المغربية: «إنّ هذا المشروع نشأ من فكرة الاستعانة بالكومبيوتر -الحاسوب- واستغلال سرعته الهائلة في علاج المعطيات وقدرته العجيبة في تخزين الملايين من هذه المعطيات في ذاكرته؛ لإنشاء بنك آلي[1] من المعطيات يحتوي أهم ما حرّر بالعربية مما سينتجه على مرّ السنين، وسيكون البنك الآلي تحت تصرف أي باحث في أي مكان من العالم[2]». وعندما نتحدث عن البنك الآلي للغة العربية فنحن في خضم الحديث عن علاقة اللغة العربية بلسانيات المدوّنة Linguistique de corpus وباللسانيات الحاسوبية Linguistique computationnelle. ثم إنّ أهم النتائج التي ستفضي إليها الذخيرة العربية متمثلة في العدد الكبير من البنوك المصطلحية الآلية في شتّى التخصصات المعتمدة في إنجاز المدوّنات النصّية بدايةً.

تعدّ اللسانيات الحاسوبية نقطة انطلاق في محاولة استغلال كل ما يتيحه الحاسوب من برامج وتطبيقات تمكّن من برمجة اللغة و رقمنتها والاستفادة منها في إنجاز المحتوى الرقمي العربي؛ وللإشارة فإنّ اللسانيات الحاسوبية (أو علم اللغة الحاسوبي كما ذهب إلى ذلك علي القاسمي) هي : «وضع نظام آلي لمعالجة اللغة

الطبيعية ويحتاج إلى مدوّنة لتطبيق أو اختبار النظام الذي يضعه على نماذج ممثّلة لجميع أنواع النصوص التي من المتوقع أنّ يعالجها ذلك النظام وكذلك بناء معجم مختص في مجال معين» [3]. ظهر هذا التخصص في منتصف الستينيات ويرجح أنّ أول من استخدم المصطلح للدلالة على هذا المجال هو دافيد هايرس "David Hayrs"عندما كان عضوا في اللجنة الاستشارية للمعالجة الآلية للغة في الأكاديمية الوطنية للعلوم في الولايات المتحدة، وجدير بالذكر أنّ تشومسكي "Chomsky" كان من الذين أسهموا بشكل كبير في تطور هذا التخصص الذي أصبح له دور مهم خصوصا مع تطور مجالي الذكاء الاصطناعي Artificial Intelligence[4]، ومعالجة اللغة الطبيعية Natural Language Processing [5]، إضافة إلى تطور الصناعة المعجمية بانتقالها من الطابع التقليدي إلى طابع إلكتروني يعتمد أيضا المدوّنات اللغوية لحل بعض الإشكالات أبرزها معنى المفردات والمصطلحات والسياقات التي وردت فيها.

ولعل الحاجة إلى مدوّنة عربية تكون القاعدة الأساس للمعطيات اللغوية المراد العمل بها وعليها هو الذي جعل الأستاذ عبد الرحمان الحاج صالح يفكر في مشروع الذخيرة العربية المحوسبة، وهي بمثابة المحرك العربي على الشابكة العالمية، ومن جهة أخرى فإنّ هذه المدوّنة العربية هي الأساس في إنشاء عدد من المعاجم العربية الإلكترونية العامة والخاصة، ولعل الإشكالية التي تتبادر إلى أذهاننا في هذا السياق: كيف يمكن استثمار المدوّنات النصّية المحوسبة للذخيرة في الصناعة المعجمية ؟

1. المدوّنة النصية المحوسبة والذخيرة العربية:

قبل الخوض في كل من المدوّنة النصية المحوسبة والذخيرة العربية لابدّ من الوقوف على معنى كل من مصطلح المدوّنة والعلم الذي تأسس في مجمله للبحث فيها -المدوّنة-؛ فالمدوّنة وهي: اسم مفعول مشتقّ من الفعل "دوّن يدوّن تدوينا" وهي من الكلمة الفارسية المعرّبة "ديوان" وهو عبارة عن دفتر يجمع فيه كل أسماء العمال والجند، كما تطلق كلمة ديوان على المكان الذي تحفظ فيه هذه الدفاتر[6].

استخدم العرب مصطلح "مدوّنة" مقابلا للمصطلح الأجنبي الفرنسي والإنجليزي Corpus[7] ، وهو من اللاتينية Corps والتي تعني الجسد، و Corpus

جمع لكلمة Corpora وهي مجموعة من البيانات اللّغوية؛ سواء أكانت نصوصا مكتوبة أم تسجيلات صوتية -نصوصا منطوقة-[8]، وبعبارة أخرى هي ما يشكل مجموع المعطيات اللّغوية التي يخضعها الباحث للتحليل والدرس[9].

واعتبارا من التعريفات التي جاءت في كل من المعاجم الأجنبية والعربية فإنّ المدوّنة هي: عبارة عن عدد كبير من النصوص المتنوعة المكتوبة منها والمنطوقة والتي تنتمي إلى مختلف التخصصات؛ وعليه يمكننا أن نتحدث عن نوعين من المدوّنات: المدوّنة النصية المكتوبة، والمدوّنة النصية المنطوقة التي يمكن أنّ تدوّن باستخدام برامج تقنية مخصصة في نقل النصوص من طابعها المنطوق إلى طابعها المكتوبة مثل ما أشار إلى ذلك "David Crystal" في معجمه: Dictionary of Linguistics Phonetics.

وتكمن أهمية المدوّنة في إمكانية استخدامها كنقطة انطلاق للوصف اللغوي أو كوسيلة للتحقق من الفرضيات حول لغة معينة[10]، فالمدوّنة في أبسط ما يمكن أن يقال عنها إنّها تجميع لعدد من النصوص المتعلقة بجملة من التخصصات فهي تفيد في استرجاع السياقات التي وردت فيها بعض الكلمات أو المصطلحات ولكنها ليست قادرة على الإجابة عن الظواهر اللغوية أو مختلف العمليات التي تحتاج نوعا من التحليل والتفسير[11]؛ كالإجابة عن بعض الظواهر اللّغوية المعقّدة، ومن الأفضل أن يتدخل الباحث للإجابة عنها.

وقد أتاحت التكنولوجيا الحديثة للمهتمين إمكانية إنشاء المدوّنات وتصميم الأنظمة التي تتبنى فكرة المدوّنة وتهتم باللغات الطبيعية في محاولة للإجابة عن الكثير من الأسئلة وكذا المساهمة بشكل فعّال في سير عملية البحث عبر العالم؛ فيمكننا استخدام مختلف الوسائل التكنولوجية وأبرزها الحاسوب لتشفير اللغة في أي سياق. يمكننا على سبيل المثال الحديث عن مجموعة اللغة المكتوبة ومجموعة اللغة المنطوقة، إضافة إلى ذلك وجود مجموعة الإشارات التي تمّ تسجيلها من خلال مجموعة من الفيديوهات .[12]

1.1. لسانيات المدوّنة:

بدأ علماء اللغة في مطلع النصف الثاني من القرن العشرين الاهتمام بما يعرف بلسانيات المدوّنة أو المدوّنات اللسانية أو المدوّنات اللغوية وغيرها من

المصطلحات المقابلة للمصطلح الفرنسي linguistique de Corpus والمصطلح الإنجليزي Corpus Linguistics والواضح أنّ هناك من اعتمد الترجمة عن الإنجليزية وهناك من اعتمد الترجمة عن الفرنسية ليخلق ذلك تعددا اصطلاحيا للدلالة على مفهوم واحد هو كما جاء في كتاب "Corpus Linguistics"؛ فهي مجموعة من النصوص الشفوية أو المكتوبة الموثّقة[13]، أو دراسة بيانات اللغة على نطاق واسع، حيث يقوم الحاسوب بتحليل لمجموعة كبيرة من الكلام المكتوب أو النصوص المكتوبة[14] لتتشكل لدينا مجموعة كبيرة من النصوص التي تصبح بمثابة المرجع الذي تحيل إليه مراكز الإنترنت the centrality of the internet فتحيل إلى مواقع الويب Websites ومواقع الخدمات[15] Webservices اعتبارا من كونها مدونات لغوية في الأساس.

تستعمل المدونة عموما للدلالة على المعنى الورقي القديم لها؛ فهي في نظر أغلب المستعملين للمصطلح بهذا المفهوم تدل على كتاب يضم مجموعة من النصوص التي تصب في مجال واحد، أو مجموعة من المجالات، أو الدال على مجموعة من الكتب لمؤلف واحد، أو لمجموعة من المؤلفين في بيئة واحدة، أو في حقبة زمنية معينة أو غير ذلك، وفي المقابل نجد من يستعمل المصطلح بمعناه الحديث المتعلق باللسانيات والمرتبط بالحاسوب فبرز إضافة إلى مصطلح لسانيات المدوّنة مصطلح المدوّنة الحاسوبية والتي عرّفت على أنّها مجموعة من النصوص المقروءة آليا -المكتوبة أو المنطوقة-بشكل متزايد حيث تم جمع مدوّنة كبيرة خاصة باللغة الإنجليزية وحيث تعتمد المدوّنة الحاسوبية في تطوير برامج معالجة اللّغة الطبيعية، وفي الصناعة المعجمية أو إنشاء المعاجم الإلكترونية والتعرف على الكلام والترجمة الآلية[16]، وربما تعود المحاولة الأولى لإنشاء مدونة آلية إلى العالمين كيوسيرا Kucera وفرانسيز Francis حيث قاما بإنشاء مدونة لغوية لأكثر من مليون كلمة للغة الإنجليزية عام 1960 ثم توالى إنشاء الكثير من المدوّنات نذكر على سبيل التمثيل: المدوّنة البريطانية BNC "British National Corpus"، والمدوّنة الأمريكية ANC "American National Corpus".

2.1. المدوّنة النصية المحوسبة للذخيرة العربية:

لقد حاول الأستاذ عبد الرحمان الحاج صالح إقناع المؤسسات والهيئات

العلمية في مختلف الدول العربية بمشروع الذخيرة العربية التي عرّفها على أنّها: «بنك آلي من النصوص؛ وليست مجرد مدونة أدخلت في ذاكرة الحاسوب، وهي ليست CDROM كما يقولون، بل مجموعة من النصوص أدمجت على الطريقة الحاسوبية، حتى يتمكن الحاسوب من مسحها كاملة، أو جزئيا، ولها عدد من البرامج الحاسوبية وضعت خصيصا لإلقاء أنواع خاصة وكثيرة من الأسئلة على الذخيرة» [17]. وقد عُرِض المشروع أول مرة في ديسمبر عام (1988م) على المنظمة العربية للتربية والثقافة والعلوم فوافق أعضاؤها على تبنيه، وفي عام (1991م) نظّمت جامعة الجزائر 2 بمشاركة المنظمة العربية للتربية والثقافة والعلوم ندوة لمناقشة كيفية إنجاز المشروع والتدابير اللازمة، ثمّ أعدّ الأستاذ عبد الرحمن الحاج صالح المشروع وقدمته الحكومة الجزائرية إلى جامعة الدول العربية فصادق عليها مجلسها سنة 2008، وأنشأت له منظمة متخصصة -على غرار المنظمات التابعة لجامعة الدول العربية- هي "الهيئة العليا للذخيرة العربية"، واختيرت الجزائر لتكون مقرا لهذه الهيئة[18]. وبناء على ما سبق فالذخيرة العربية عبارة عن مدوّنة نصية محوسبة فهي مجموعة كبيرة من النصوص التي تجمع آليا لتشكّل بنكا آليا عربيا؛ وهذه العملية لا تكون إلا وفق جملة من المراحل نوجزها فيما يلي:

- تحديد الهدف من الذخيرة العربية وهو الشيء الذي وضّحه الأستاذ عبد الرحمان الحاج صالح في الوثيقة التي قدّمها للمنظمة العربية منذ البداية.
- تحديد المصادر التي تُعتمد في إنشاء الذخيرة العربية والتي حددها كذلك الأستاذ عبد الرحمان الحاج صالح سابقا والمتمثلة في: الكتب والمعاجم العربية القديمة والحديثة إضافة إلى الجرائد والكتب المدرسية والقصص والروايات والمحاضرات لأنّها تمثل الاستعمال الحقيقي للّغة.
- الحصول على الموافقة القانونية من المؤسسات المعنية بتلك المؤلفات.

1.2.1. أهداف مشروع الذخيرة العربية:

تتمثل أهداف إنشاء الذخيرة العربية حسب ما ذكره الأستاذ عبد الرحمان الحاج صالح في مختلف المؤتمرات والمحافل الدولية فيما يلي:

- **الذخيرة بنك معلومات آلي:**

وقد حدثنا عن ذلك الأستاذ عبد الرحمان الحاج صالح في ندوة له مع طلبة الدكتوراه عمّا أطلق عليه ب:"Google عربي" وهو بنك نصي آلي يعدّ نموذجا حيا عن الاستعمال الحقيقي للغة وتكون هذه النصوص في جميع التخصصات على مرّ العصور، ويكون هذا البنك الآلي قاعدة معطيات دائمة حيث تقبل الزيادة والتصحيح على الدوام بسبب تطور المعلومات من خلال الاستعمال الحقيقي للغة العربية، ومنه إنجاز المعجم العربي الجامع.

- **الذخيرة العربية المحوسبة والمعاجم الآلية:**

إنّ أهم الأهداف التي تحدث عنها الأستاذ عبد الرحمان الحاج صالح إمكانية استثمار الذخيرة العربية في إنجاز عدد من المعاجم الآلية نذكر على سبيل التمثيل المعجم التاريخي للغة العربية.

2.2.1. منهجية جمع النصوص المحوسبة للذخيرة العربية:

تم انتقاء مجموعة من الصحف من كل بلد عربي عضو في جامعة الدول العربية بالإضافة إلى مجموعة من الصحف المشهورة التي تصدر في الخارج مثل الشرق الأوسط وجريدة الحياة. وكان المعيار الأساس لاختيار الجرائد هو الشهرة والشيوع والمقروئية الواسعة وكونها جريدة حكومية ما أمكن.

- المحتوى الرقمي للصحف والجرائد العربية: نذكر على سبيل التمثيل: جريدة "الوطن"، وجريدة "عُمان" لسلطنة عُمان، وجريدة "الصباح" للمغرب، جريدة "المدى"، وجريدة "الصباح" للعراق اعتمادا على مجموعة من البرامج التقنية والتي تسهم بشكل كبير وفعّال في تنزيل عدد ضخم من المقالات في وقت وجيز.

1. جمع المادة من عدد كبير من الصحف والمجلات في مختلف دول الوطن العربي أو الناطقة باللغة العربية.
2. تحديد المجالات التي يُعمل على جمع مقالاتها كل مجال على حدة: في المجالات التالية كل على حدة فنجد: مجال السياسة، ومجال الصحة، ومجال الرياضة، ومجال الثقافة... إلخ.

- حيازة مجموعة من قصص الأطفال والكتب المدرسية.

2. استثمار المدوّنة النصية للذخيرة العربية في الصناعة المعجمية:

لقد أولت العرب اهتماما كبيرا بالصناعة المعجمية منذ القدم فتنوعت المعاجم التقليدية بتنوع المنهج المتّبع في تأليفها مثل معاجم التقاليب والمعاجم الهجائية والمعاجم الموضوعاتية...إلخ، وكذا من حيث المادة التي يقدّمها المعجم، وهو ما يجعله يندرج إما في خانة المعجم العام أو المعجم الخاص. ومع تطور العلوم والتقنيات واعتبارا مما قدّمه الأستاذ عبد الرحمان الحاج صالح فيما يخص المدوّنة النصية للذخيرة العربية فإنّه أكّد على أنّها تمكن من إنجاز عدد كبير من المعاجم الإلكترونية العامة والخاصة.

1.2. الصناعة المعجمية العامة:

إنّ عملية البحث التقليدي عن معنى كلمة في كثير من الأحيان تستغرق وقتا طويلا؛ غير أنّ عملية البحث هذه ستصبح ممكنة ومتاحة لأي باحث خصوصا مع إنجاز مشروع الذخيرة العربية والذي سيفضي إلى مجموعة من المعاجم الآلية العامة؛ ومعنى قولنا عامة أنّها تجمع كل ألفاظ اللغة العربية في مختلف المجالات وعلى مرّ العصور إلى يومنا هذا؛ وحسب ما جاء في وثائق مشروع الذخيرة التي قدّمها الأستاذ عبد الرحمان الحاج صالح فيمكن التمثيل لهذا النوع بـ:

- المعجم الآلي الجامع لألفاظ اللغة العربية:

يعد المعجم الآلي الجامع لألفاظ اللغة العربية النموذج الصادق عن استثمار الجانب الحاسوبي في مجال المعجميات حيث يحتوي «على جميع المفردات العربية التي وردت في النصوص قديمة وحديثة، وتحدد فيه معاني كل مفردة باستخراج هذه المعاني من السياقات التي ظهرت فيها، ثم يضاف إلى ذلك تحديدات العلماء»[19]. إنّ أهم ما يمكن أن نستفيده من هذا المعجم أن تصبح للغة العربية أكبر مدوّنة نصيّة محوسبة تتضمن جميع مفرداتها القديمة والحديثة مشروحة، أضف إلى ذلك السرعة في البحث، والدقة حيث تقوم الآلة بتقديم الكلمة ومعناها والسياقات التي وردت فيها هذه الكلمة كما يقدّم لنا المقابلات الفرنسية والإنجليزية.

2.2. الصناعة المعجمية المتخصصة:

يعدّ الاهتمام بالمعجم الخاص نابعا من اهتمام الباحثين باللغات العامة واللغات الخاصة ومحاولة الفصل بين كليهما بوضع الحدود الإبستيمية التي تفصل هذه عن تلك أو تتداخل فيها معها ما أدى إلى القول بالمصطلح والمفهوم في اللغة الخاصة مقابلين للكلمة والمعنى في اللغة العامة، وعليه برزت العناية بالمعجم الخاص الذي يضم عددا كبيرا أو قليلا من المصطلحات العلمية أو التقنية، وهذا الأمر بالتحديد جعل الأستاذ عبد الرحمان الحاج صالح يشير إلى إمكانية استثمار الحاسوب في وضع معاجم متخصصة في اللغة العربية وإنجازها.

1.2.2. المعجم التاريخي للغة العربية:

ينطلق المعجم التاريخي من محاولة : «رصد التطور الدلالي للغة العربية عبر العصور التاريخية، وفهم التراث المعرفي والعلمي فهما صحيحا بفهم دلالة الألفاظ ومفاهيم المصطلحات في سياقاتها التاريخية وسدّ الفجوة المعجمية الناتجة من قصور المعاجم العربية عن مواكبة التطور اللغوي»[20] فالمعجم التاريخي من أكبر المدوّنات اللغوية المحوسبة ذلك لأنّها قابلة للزيادة والنقصان أضف إلى ذلك أنّها ستقدم معطيات وبيانات كان البحث عنها في ثنايا الكتب والمعاجم شاقًّا ويستغرق وقتًا طويلا.

2.2.2. معاجم المصطلحات العلمية والتقنية:

لا يفصل البعض بين المصطلح العلمي والمصطلح التقني ويعتقدون أنّهما شيء واحد والصواب أن الأول هو عبارة عن ألفاظ اتفق العلماء والباحثون عليها للدلالة على شيء معين كعنصر أو عضو أو ظاهرة ما، نقول مثلا: المبيض، والتركيب الضوئي، والأكسدة...إلخ، أما الثاني فقد اتفق على أن يدل على جهاز نقول مثلا: الحاسوب، البيشر، الميكروسكوب...إلخ. فإنجاز مشروع كالذخيرة العربية للأستاذ عبد الرحمان الحاج صالح يفضي إلى وضع معاجم للمصطلحات العلمية والتقنية وسيكون لهذه المعاجم الدور الفعّال في توحيدها في الوطن العربي، ويحكم هذا التوحيد شيوع الاستعمال والتداول فبدلا من استخدام السيتوبلازم، والهيولى، والسائل الهلامي نستعمل مصطلح السيتوبلازم اعتبارا من كونه الأكثر شيوعا

واستعمالا في الوطن العربي أو نتفق على غيره بناءً على ما ستقدّمه الآلة من تكرارات للمصطلح الواحد حيث يعبر التكرار عن كثرة الاستعمال والشيوع.

3.2.2. معجم الطفل:

ذهب الأستاذ الحاج صالح إلى أنّ الذخيرة العربية يمكنها أنّ تكون كدعم وتطوير للرصيد اللغوي الوظيفي[21]؛ حيث يهتم مشروع الذخيرة العربية المحوسبة بالطفل وما يجب أنّ يحيط به من مفردات وعبارات لغوية ضرورية يحتاجها في حياته.

خاتمة:

توصلنا من خلال بحثنا في الموضوع الموسوم:" استثمار المدوّنات النصّية المحوسبة للذّخيرة العربية في الصّناعة المعجمية" إلى جملة من النتائج نوجزها فيما يلي:

1. يندرج مشروع الذخيرة العربية الذي اقترحه الأستاذ عبد الرحمان الحاج صالح في إطار لسانيات المدوّنة، لأنه عبارة عن مدوّنة كبيرة من النصوص المحوسبة.
2. يعد مشروع الذخيرة العربية أداة ووسيلة نافعة للصناعة المعجمية وإنشاء بنوك المصطلحات.
3. لا يمكن في الوقت الراهن بناء معجم ما مهما كان نوعه بالطريق التقليدية أي بالاعتماد على ما ورد في المعاجم التراثية والاتكاء على محتوياتها في انتقاء المادة المعجمية وترتيبها وشرحها، بل يجب الانطلاق من مدونات محوسبة معدّة سلفا بطريقة علمية وجرد ما فيها من مادة معجمية ثم معالجتها بطريقة آلية عن طريق برمجية خاصة تحدد جذور الألفاظ وتأتي بها في سياقاتها التي وردت فيها داخل نصوص المدوّنة.
4. مشروع الذخيرة العربية رغم تعثره في بداياته الأولى استطاع أن يؤسس رؤية علمية مستقبلية للصناعة المعجمية ولإنشاء بنوك المصطلحات وذلك من خلال الأداة التقنية التي طورها فريق المهندسين التابعين له، ومن خلال المدوّنات التي جمعها خبراء مختصون، لتوضع بين أيدي الباحثين المهتمين بهذا المجال.

5. استفادة بعض الهيئات من أفكار الأستاذ عبد الرحمن الحاج صالح خاصة في مجال المعاجم؛ حيث ظل يدعو إلى ضرورة الانطلاق من النصوص المحوسبة في ذلك، واعتماد برمجيات إلكترونية لتحقيق الغاية. وتتجلى الإفادة من هذه الفكرة في المعجم التاريخي للغة العربية الذي لم يعرف طريقه إلى الظهور إلا بعد أن اعتمد هذين المبدأين: مبدأ الانطلاق من المدوّنات النصية المحوسبة، ومبدأ اعتماد أدوات تقنية تسهل العمل وتسرع عملية الإنجاز.

الإحالات والهوامش:

[1] شاع استعمال مصطلح البنك الآلي في الدراسات اللسانية للدلالة على جملة من البنوك التي تنتمي إلى قواعد البيانات مثل: بنك البيانات، وبنك الكلمات، وبنك المصطلحات.

[2] عبد الرحمان الحاج صالح، ورقة حول مشروع الذخيرة العربية، مجلة اللسان العربي، مكتب تنسيق التعريب، ع47، المغرب، 1998، ص107.

[3] علي القاسمي، علم المصطلح أسسه النظرية وتطبيقاته العلميّة، ط2، مكتبة لبنان ناشرون، بيروت، لبنان، 2019، ص711.

[4] الذكاء الاصطناعي يرمز له عموما في المؤلفات التي تناولته AI اختصارا للمصطلح المركب "Artificial Intelligence" يتشكل المصطلح من شقين الأول اصطناعي وهو كل شيء مصنوع أو غير طبيعي، أما الثاني فهو الذكاء وهو القدرة على التفكير والفهم. ينظر: أحمد حبيب بلال، عبد الله موسى، الذكاء الاصطناعي ثورة في تقنيات العصر، ط1، دار الكتب المصرية، القاهرة، مصر،2019م، ص18.

[5] ينظر: أحمد روبي محمد أحمد الرفاعي وآخرون، مدخل إلى اللسانيات الحاسوبية، ط1، مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الرياض، المملكة العربية السعودية،2017م، ص6.

[6] ينظر: علي القاسمي، علم المصطلح أسسه النظرية وتطبيقاته العلميّة،2019م، ص700.

[7] Sandrine Zufferey, Introduction à la Linguistique de Corpus, IST Edition LTD, British, 2020, p.3.

[8] David Crystal, Dictionary of Linguistics and Phonetics, 6Th Edition, Blackwell Publishing, USA, 2008, p.117.

[9] ينظر: مكتب تنسيق التعريب، المعجم الموحد لمصطلحات اللسانيات-إنجليزي-عربي-فرنسي، ط2، مطبعة النجاح، الدار البيضاء، المغرب، 2002، ص38.

[10] David Crystal, Dictionary of Linguistics and Phonetics, 117.

[11] أيمن الدكروري، المدوّنات اللغوية ودورها في معالجة النصوص العربية، دار وجوه للنشر والتوزيع، الرياض، المملكة العربية السعودية، ط1، 2018، ص38-39.

[12] Andrew Hardie, and Tony Mcenery, Corpus Linguistics Method-Theory and Practice, Cambridge University press, New York, USA, 2012, p.3.

[13] سلطان ناصر المجيول وآخرون، المدوّنات اللغوية العربية بناؤها وطرائق الإفادة منها، ط1، مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الرياض، المملكة العربية السعودية، 2015، ص19.

[14] Andrew Hardie, and Tony Mcenery, Corpus Linguistics, p. 1.

[15] Ibid., p.12.

[16] David Crystal, Dictionary of Linguistics and Phonetics, p.117.

[17] عبد الرحمان الحاج صالح، مشروع الذخيرة العربية، مجلة المجمع الجزائري للغة العربية، ع2، 2005، ص287.

[18] ينظر: عبد الرحمان الحاج صالح، ورقة حول مشروع الذخيرة العربية، مجلة اللسان العربي، ص107.

[19] عبد الرحمان الحاج صالح، مشروع الذخيرة اللغوية العربية وأبعادها العلمية والتطبيقية، مجلة الآداب، 1996م، ص7.

[20] عز الدين البوشيخي وآخرون، نحو معجم تاريخي للغة العربية، المركز العربي للأبحاث ودراسة السياسات، الدوحة، قطر، الطبعة1، 2014م، ص20-21.

[21] عبد الرحمان الحاج صالح، بحوث ودراسات في اللسانيات الحديثة، منشورات المجمع الجزائري للغة العربية، الجزائر، 2007، ج2، 144-145.

قائمة المصادر والمراجع:

أ- باللّغة العربيّة:

1- البوشيخي عز الدين وآخرون، نحو معجم تاريخي للغة العربية، المركز العربي للأبحاث ودراسة السياسات، الدوحة، قطر، الطبعة 1، 2014م.

2- الحاج صالح عبد الرحمان، بحوث ودراسات في اللسانيات الحديثة، منشورات المجمع الجزائري للغة العربية، الجزائر، 2007م، الجزء الثاني.

3- الحاج صالح عبد الرحمان، مشروع الذخيرة العربية، مجلة المجمع الجزائري للغة العربية، العدد2، 2005م.

4- الحاج صالح عبد الرحمان، مشروع الذخيرة اللغوية العربية وأبعادها العلمية والتطبيقية، مجلة الآداب، 1996م.

5- الحاج صالح عبد الرحمان، ورقة حول مشروع الذخيرة العربية، مجلة اللسان العربي، مكتب تنسيق التعريب، العدد47، المغرب، 1998.

6- حبيب بلال أحمد وموسى عبد الله، الذكاء الاصطناعي ثورة في تقنيات العصر، الطبعة1، دار الكتب المصرية، القاهرة، مصر،2019م.

7- الدكروري أيمن، المدوّنات اللغوية ودورها في معالجة النصوص العربية، دار وجوه للنشر والتوزيع، الرياض، المملكة العربية السعودية، الطبعة 1، 2018م.

8- الرفاعي أحمد روبي محمد أحمد وآخرون، مدخل إلى اللسانيات الحاسوبية، الطبعة 1، مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الرياض، المملكة العربية السعودية،2017م.

9- القاسمي علي، علم المصطلح أسسه النظرية وتطبيقاته العلميّة، الطبعة 2، مكتبة لبنان ناشرون، بيروت، لبنان، 2019م.

10- المجيول ناصر سلطان وآخرون، المدوّنات اللغوية العربية بناؤها وطرائق الإفادة منها، الطبعة1، مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الرياض، المملكة العربية السعودية، 2015.

11- مكتب تنسيق التعريب، المعجم الموحد لمصطلحات اللسانيات-إنجليزي-عربي-فرنسي، الطبعة 2، مطبعة النجاح، الدار البيضاء، المغرب، 2002م.

ب- باللغات الأجنبية:

1- Crystal David, Dictionary of Linguistics and Phonetics, 6th Edition, Blackwell Publishing, USA, 2008.

2- Hardie Andrew, and Tony Mcenery, Corpus Linguistics Method-Theory and practice, Cambridge University press, New York, USA, 2012.

3- Zufferey Sandrine, Introduction à la Linguistique de Corpus, first published, ISTE Edition LTD, British, 2020.