

واقع رقميّة المعجم العربي وآفاقه - معجم الرائد أنموذجاً -

د. بن يشو الجيلالي

جامعة عبد الحميد بن باديس -
مستغانم (الجزائر)

ط.د. اللوط بدرة

جامعة عبد الحميد بن باديس -
مستغانم (الجزائر)

تاريخ القبول: 2023/02/09

تاريخ الاستلام: 2022/11/26

ملخص:

يُعتبر المعجم من أهم روافد اللّغة والمعرفة وذاكرة أمة، ودُستورها النفيس، وبما أنّ العالم اليوم يشهد انفتاحاً واسعاً على البيّنة الرقمية والتكنولوجيا المعرفية، يوصفهما عاملين أساسيين لتطوير اللّغة والتّهوّض بها.

ومع انتشار الحاسب الآلي وتطوّر استعمالاته، أضحت الجامع اللّغويّة في أمسّ الحاجة لاستخدامه، واستثمار إمكانياته، لما يُوفّره من طاقات لتخزين المعلومات المعجمية، وتخيّناتها، واسترجاعها، والبحث عنها بسرعة، ولما يُتيحها من قدرات لمعالجة البيانات المتعدّدة الوسائط للإحاطة بالمعاجم اللّغوية؛ والتي أضحت ونظراً لضخامتها، وتنوّع مجالاتها وحقوقها استحالة تنظيم وتحليل وتفسير الكمّ الهائل لمعلوماتها يدوياً، فبات من الضروريّ بما كان الاستعانة بالرقمنة واستثمارها في المعاجم.

انطلاقاً من ذلك تروم ورقه بحثي حول واقع رقميّة المعجم العربي، والتطوّرات التي عرفها في ظلّ عالم الحوسبة والتحوّلات الحديثة التي أفرزت لنا مجموعة من المصطلحات، والمفاهيم في بناء المعجم الحاسوبي، وما آلت إليه من تطوّرات، وآفاق لتطوير المعاجم الالكترونية بسرعة فائقة والتي أصبحت تُنافس بصفة جدّية مثيلاتها من المعاجم الورقية، وقد اعتمدت في دراستي النموذجية على "معجم الرائد" في نسخته الالكترونية، والذي يُعتبر من المعاجم الإلكترونية الرائدة في هذا المجال.

الكلمات المفتاحية: المعجم العربي، الرقمنة، حوسبة المعجم، معجم الرائد.

تمهيد:

إنّ العالم اليوم يعيش في تراكم معرفيٍّ من شتى المجالات خاصة بما يتعلّق بالتضخم التكنولوجي، فصار إلزامياً مواكبة هذا الزخم الحضاري والدخول لعالم الرقمنة، ولا يتأتى ذلك إلا باللغة، بكونها الوسيلة الأساسية لتناقل العلوم وتبادل المعارف، فكان لا بد من تطويرها وجعلها تتلاءم مع مختلف الإبداعات الرقمية.

* ط.د. اللوط بدرة.

من المعروف أن الأمة العربية تمتلك أكبر رصيد معجمي لغوي، إذ إن حركة التأليف المعجمي بدأت منذ القرن الأول للهجرة في مرحلة جمع اللغة، على يد ثلة من البعث اللغويين؛ نتج لنا تراث لغوي والرصيد المعجمي امتد ليومنا هذا، وما زالت الأمة العربية تبتهل في صناعة المعاجم نظراً لما يطرأ على اللغة من أحوال تندثر فيها ألفاظ تارة وتتطور فيها تارة أخرى، وهنا نقف أمام قضية العصر والتحديات الرقمية الحاصلة في هذا الزمن، لتساءل عن موقع المعجم العربي في الترجمة اللسانية الحاسوبية الرقمية؟.

1- مفهوم الرقمنة:

تعددت مفاهيم الرقمنة ولعل من أشملها أنها: "عملية تحويل المعلومات على اختلاف أشكالها من الكتب والدوريات والتسجيلات والصور المتحركة إلى شكل مقروء بواسطة تقنية الحاسبات الآلية، عبر النظام الثنائي (bits)، وتشير شارلوت بيرسي (chargette bursi) إلى الرقمنة على أنها منهج يسمح بتحويل البيانات والمعلومات من النظام التناظري إلى النظام الرقمي"¹؛ أي التحويل من الصيغة الورقية الملموسة إلى الصيغة الرقمية الافتراضية المحسوسة.

ويرى البعض الآخر "بأنها التحول في الأساليب التقليدية المعهود بها إلى الحفظ الإلكتروني هذا التحول يستدعي التعرف على كل الطرق والأساليب القائمة واختيار ما يتناسب مع البيئة الطالبة لهذا التحول، وتعتمد الرقمنة على المسح، وتحويل الصفحة إلى صور أو المسح مع استخدام نظام التعرف على الأحرف (OCR)، أو إدخال الوثيقة من جديد عن طريق لوحة المفاتيح، أو إدخال الوثيقة وفق معيار النصوص الفاتقة lHtm²، باختصار؛ هي المعالجة الحاسوبية للغة العربية أو هي طريقة جديدة تحفظ المعلومات بشكل حديث، مختصر وسريع.

وقد واجهت الرقمنة بعض من عوائق في العالم العربي نذكر أهمها:

1. انتشار الأمية المعلوماتية، حيث كانت نسبة بعض الدراسات حولها في العالم العربي بنحو 40%.
2. ضعف البنى التحتية لشبكات الإنترنت في عدة دول عربية بسبب الحالة المادية، مما منعها من الاستفادة والحضور في العالم الرقمي ومواكبة الالركب التكنولوجي، والذي يُنظرُ إليه في المستويات الاقتصادية العالمية على أنه ترفٌ فكري وعصري.

3. غياب دور مراكز البحث والجامعات في رقمنة الدراسات والبحوث الأكاديمية التي تعدّها هذه المؤسسات، وإذا وجدت هذا الأمر في مركز غاب في عشر مراكز، وهذا أيضاً مرتبط بقضية الموارد وشُحّها.
4. انتشار ما يُعرف باللغة الفرانكو عربية (franco-arabe)، بمعنى كتابة اللغة العربية بحروف أجنبية، وهذا الشائع في مواقع التواصل الاجتماعي، وهو ما جنى على اللغة العربية جناية عظيمة.
5. مشكلة التدقيق الآلي وشيوع الأخطاء، فمعظم محركات البحث عندها خوارزميات ثابتة تتعلق بشيوع الاستخدام، فكلما كان اللفظ مستخدماً أكثر تعاملت معه على أنه هو الصحيح، بغض النظر عن السلامة اللغوية.
6. غياب الموسوعات العربية الموثوقة والمصادر المفتوحة أو ما يُسمى "MOOCs"، وهو اختصار للعبارة الإنجليزية "Massive Open Online Course"، وهي مواقع متخصصة في كل فرع من العلوم، وفيها معلومات موثقة، فغياب هذه الموسوعات يزيد الفجوة المعرفية ويُتيح المجال لانتشار المعلومات الخاطئة.

2- تاريخ رقمنة المعجم:

يُعرّف المعجم الرقّمي على أنه "معجم تقليدي محوسب وممكن، يُخزّن على شكل شرائح إلكترونية أو وسائط ممغنطة كالأقراص الممغنطة أو الضوئية، وذلك لاستخدامها لأغراض الترجمة الآلية والتعليم واكتشاف الأخطاء الإملائية علاوة على أغراض المعالجة الآلية الأخرى"³، وقد نال المعجم الرقّمي مساحة واسعة من الاهتمام كونه "قاعدة بيانات آلية تقنية للوحدات اللغوية، وما تعلق بها من قبيل كفاءات النطق بها، وأصولها الصرفية ومحملها الدلالية، وكيفية استخدامها ومفاهيمها المخصصة التي تُحفظ بنظام معين في ذاكرة تخزين سعة كبيرة، ويقوم جهاز آلي بإدارة المعطيات الفنية والمضمونية التي يتضمنها المعجم الإلكتروني (الرقّمي) وفق برنامج محدد سلفاً"⁴.

فمنذ منتصف القرن الماضي؛ بدأت المعاجم الإلكترونية تلقى الاهتمام لما تتطلبه تكنولوجيا العصر، بدايةً انحصر اهتمامهم في استخدام هذه المعاجم كمصادر لغوية للتحليل الآلي للغات الطبيعية على المستويات اللغة من نحو وصرف وصوت...، فكانت المعاجم في أول الأمر عبارة عن قواعد بيانات تحتوي على معلومات مشفرة على شكل رموز لا يفقهها إلا البرنامج الذي يحملها.

وقد تنوعت وتعددت هذه المعاجم وتطورت من حيث المحتوى ومن حيث الهيكل ليواكب تطور البرامج وتطبيقات المعالجة الآلية للغة التي يشتغل عليها؛ "كنقطة الجمل، تحليل النصوص واسترجاعها، البحث عن المعلومات، التدقيق الإملائي، التلخيص الآلي للوثائق والترجمة الآلية. فإذا أخذنا على سبيل المثال المعجم المصمم

للتدقيق الإملائي نجدّه يختلف تماما عن المعجم المصمم للتحليل النحوي. فالأول يقتصر على قائمة كلمات اللغة، والثاني يستوجب تمثيل المعلومات الصرفية والنحوية (مثل: قسم الكلم، التعدية واللزوم، أدوات التعدية،...) بالنسبة لكل مدخل⁵.

فأضحت المعاجم الرقمية من أبرز المعالجات الآلية للغات الطبيعية، واعتبرت من أهم المصادر الحديثة لحفظ الذاكرة اللغوية لأي أمة ما، فوجب الاهتمام بها وتحيينها لمعايشة الركب المعلوماتي الهائل والتدفق المصطلحي المتسارع الذي تشهده الساحة اللغوية عامة والعربية بشكل خاص.

3- آليات بناء المعجم الرقمي:

رغم تعدد الطرق المعتمدة صناعة المعجم الرقمي إلا أن هذا الفعل ليس بالعمل الهين، فالأمر يتطلب جهداً ووقت و تركيز وعمل جماعي يجمع بين فريق مكون من مجموعتين؛ معجميين لغويين ومهندسين في الإعلام الآلي أو مختصين في المعلوماتية وعلم الحاسوب، ونتاج هذا العمل يندرج تحت ما يسمى علم اللسانيات الحاسوبية، فيهتم المعجميون بجمع المادة اللغوية من المعاجم ورقية والمصادر اللغوية ثم يتم اختيار المداخل وترتيبها بإلحاق تعاريف لكل مدخل، والفريق الثاني يختص بكل ما تعلق بتحويل وحوسبة المادة اللغوية وتصميم البرامج الضرورية لإدخال المعلومات المعجمية، مع الأخذ بعين الاعتبار "الجوانب الأساسية ترتيب المداخل (حسب الجذور، حسب الجذوع)، وظيفة المعجم (للترجمة، للتعليم...)، نوعية المحتوى (نصي/ متعدد الوسائط، لغوي/ متخصص، أحادي اللغة/متعدد اللغات...)، الحمل (سطح المكتب، صفحات الواب، أقراص مدمجة...)"⁶.

كما سبق وذكر أن الحاسب الآلي أداة من أدوات العمل المهمة في بناء المعجم الرقمي، "وتتجلى أهمية الحاسوب في الصناعة المعجمية الإلكترونية فيما يلي:

أ- تظهر قيمته في تخزين المادة وترتيبها طبقا للنظام المطلوب وتعديل وحذف بعض المعطيات، وتحديد المعجمات بسهولة، إضافة إلى التقليل المباشر إلى المطبعة والحصول على أجزاء محدودة من داخل المادة المخزونة"⁷.

ب- يساهم استعمال الحاسب الآلي الصناعة المعجمية في "تطوير الترجمة الآلية وتحسينها يرتد ذلك في شمولية النظام اللغوي الذي يزود به الحاسوب، إذ يستطيع أن يضع أمام المترجم المعنى الدقيق للمصطلح في كل فرع، فمن المعروف أن معنى المصطلح الذي يستعمل فيه في حين يعجز معجم واحد عن سرد معاني المصطلحات المختلفة في مختلف فروع العلوم والتكنولوجيا"⁸.

ج- وكذلك يقوم بتحليل "العلاقة بين مفردات المعجم وعناصره كالعلاقة بين جذور الكلمات والصيغ الصرفية أو قواعد تكوين الكلمات المطبقة عليها العلاقات الموضوعية التي تجمع بين عائلات المفردات (مصطلحات رياضية، مصطلحات طب...)"⁹.

د- يعمل على تحليل لغة تعريف المعجم "وهي التي تستخدم لتوضيح دلالات مفرداته وتعدّ هذه الدراسة ذات أهمية خاصة لبحوث الدلالة المعجمية، حيث تسعى للوصول إلى نواة المعجم ذاتها وكذلك المبادئ الأساسية التي تحكم صياغة مفرداتها"¹⁰.

ومن هنا جاءت الدعوة إلى "التعاون بين اللسانيين والمعلوماتيين لبناء قواعد البيانات الصحيحة للغة العربية والتي ستكون أساس أي برنامج للعلاج الآلي لهذه اللغة"¹¹.

فالمعجم الرقمي يبنى مبدئياً وفق "عملية التجزئة إلى وحدات كرافية على شكل رسومات، ولذلك فإن المعجم يبنى أساساً لتحديد كل واحدة من هذه الوحدات من وجهات نظر مختلفة: صورية، دلالية، صرفية، صوتية"¹²، وتتم هذه المراحل في طرق بُني المعجم الإلكتروني عليها نحصها فيما يلي:

1- الطريقة البدوية:

بالرغم من استهلاكها للوقت والجهد والمال، إلا أنها تُعد من أبسط الطرق وأسهلها وأكثرها شيوعاً، تقوم هذه الطريقة لبناء المعجم الرقمي على مراحل هي:

أ- "الاتفاق على هيكل المعجم (ترتيب المداخل ومحتوى كل مدخل) انطلاقاً مما يسمى بكراس الشروط الذي يحدد الهدف من بناء المعجم، المستخدم المستهدف (مبتدئ، خبير،...) طريقة استغلال المعجم (على الشابكة، على سطح المكتب أو على قرص مضغوط،...)".

ب- تحديد المصادر اللغوية (مدونات، معاجم ورقية،...) التي سيقع اعتمادها كمادة أولية.

ج- انتقاء وتجميع المادة المعجمية من المصادر المذكورة في جذاذات ورقية.

د- تصميم وبناء قاعدة البيانات التي ستحتوي المعجم.

هـ- بناء برمجية لإدخال المعلومات المعجمية يدوياً من الجذاذات الورقية.

و- إدخال المعلومات يدوياً من الجذاذات الورقية.

ز- تجريب المعجم للتأكد من تحقيقه كافة الوظائف المطلوبة منه.

ح- تصويب الأخطاء اللغوية والبرمجية¹³.

2- طريقة الانطلاق من معجم ورقي مرقمن:

أي تحويل هذا الأخير إلى معجم الكتروني مرقمن يمر عبر برنامج حاسوبي بغية تصميمه، لنستخلص في آخر الأمر قاعدة البيانات المعجمية يمكن دمجها بالارتكاز على عدة معاجم أخرى، وتعد هذه الطريقة الأكثر اقتصادا للمال و الجهد والأسرع وقتا، وتتم هذه الطريقة وفق المراحل التالية:

أ- اختيار المعجم (أو المعاجم) الذي سيعتمد عليه في صيغته المرقمنة.

ب- تصميم قاعدة البيانات المعجمية التي ستحتوي مداخل المعجم.

ج- تصميم برنامج تعبئة قاعدة البيانات المعجمية انطلاقا من النسخة أو النسخ المرقمنة.

ح- تجريب القاعدة.

خ- تصحيح الأخطاء اللغوية والحاسوبية.

د- إثرائها يدويا إذا اقتضت الحاجة¹⁴.

2- طريقة الإسناد إلى مدونة نصية:

وهي الانطلاق من مدونة أو عدد كبير من النصوص تعد هذه الطريقة من الطرق الصعبة وأقلها دقة بالمقارنة مع الطرق التي سلفنا ذكرها، يتبع فيها المراحل الآتية:

أ- "تحديد المدونة أو بناؤها.

ب- إنجاز برمجية لاستخراج المداخل والعلاقات بينها والمعلومات المعجمية الخاصة بكل مدخل بصفة آلية من المدونة.

ج- تدقيق المادة المعجمية المتحصل عليها من طرف المعجمين.

د- إدخال المادة المعجمية المدققة في قاعدة البيانات التي ستحتوي المعجم¹⁵.

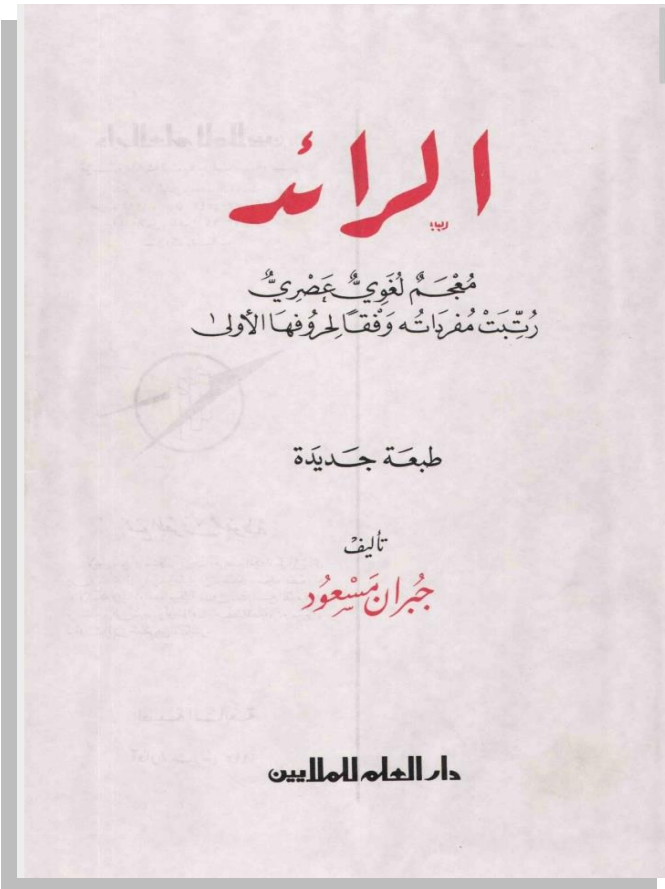
3- الطريقة التعاونية (ويكي):

ظهرت هذه الطريقة مع ظهور الموسوعة (ويكيبيديا) "تُمكن من بناء معاجم إلكترونية بتكلفة زهيدة وذلك بتكاتف جهود المتطوعين المتخصصين في مجال المعجمية ... نتاج هذه الطريقة بالنسبة للغة العربية مشروع (ويكاموس) الذي انطلق في شهر مايو 2004 والذي تقدم أشواطاً مهمة لكنه لا يزال يشكو من ثغرات عديدة في المحتوى لأن بناءه لم يكن ممنهجاً، بل موكول إلى إرادة المساهمين في عملية الإثراء"¹⁶.

وسنفصل أكثر عن هذا في تطبيق هذه الورقة البحثية، على معجم من المعاجم الالكترونية الرائدة في هذا العصر الذي تم بناءه انطلاقاً من معجم ورقي مرقم.

مُعْجَمُ الرَّائِدِ:

يعد من المعاجم الالكترونية العربية المحوسبة الموجود على الانترنت بصيغة PDF كنسخة رقمية لمعجم الرائد الورقي، ذو الطبعة السابعة، في مارس سنة 1992م، عن دار العلم للملايين للنشر ببيروت، لمؤلفه جبران مسعود، تعتبر هذه الطبعة الالكترونية؛ مجانية التحميل والأولى من نوعها في طرق حوسبة المعاجم وتنقيحها احتوت على 912 صفحة.



تمت عملية رقمنة "معجم الرائد" وفق طريقة؛ الانطلاق من معجم ورقي التي ذكرناها من قبل ، في كيفية بناء معجم إلكتروني كنوع من أنواع الصناعة المعجمية الرقمية، كما كان لابد من الأخذ بالاعتبار الجوانب الأساسية التي قام عليها هذا المعجم من:

أ- اعتمادهم في اختيار المداخل على ترتيب مفرداته وفقاً لحروفها الألفبائي.

ب- تصنيفه كمعجم مدرسي كون وظيفة هذا المعجم تختص المتعلم في الدرجة الأولى.

ت- معجم عربي/عربي (أحادي اللغة).

ث- تواجده على الانترنت ويُحْمَلُ بطريقة مجانية

على صيغة PDF.

وقد تمت حوسبة معجم "الرائد" باتباع الخطوات معينة أخرجته من لباسه الورقي إلى حلة رقمية عصرية، التي عُدت من أبسط الطرق وأكثرها اقتصادا للوقت والمال ففي البداية كان لا بد من القائمين على إنجازها في البداية يتم:

أ- اختيار المعجم الورقي (الرائد)؛ الذي سيعتمد عليه في إنجاز النسخة الرقمية، هذا الفعل يختص به المعجميون .

ب- ثم يتشارك المعجمي والمعلوماتي في إنشاء قاعدة البيانات المعجمية التي تحتوي مداخل المعجم الأول من جانب اللغة والترتيب المداخل أما الثاني لتصميم القاعدة البيانات.

ت- اختيار برنامج لتعبئة قاعدة البيانات المعجمية بالرجوع للنسخة الورقية .

ث- تجريب القاعدة وصحيح الأخطاء اللغوية والحاسوبية

ج- ترك المجال مفتوحا للتعديل والإضافة يدويا إذا اقتضت الحاجة، من قبل القائمين على صناعة هذا المنجز الإلكتروني، فيقول صاحب الرائد: " فقد راجعنا إلى كل لفظة مشروحة من ألفاظه، وحيث وجدنا حاجة و نفعا في زيادة الشروح زدنا، كما أننا زدنا الشروح بالمزيد من الأمثلة المناسبة، هذا فضلا عن فصل ضم معلومات وقواعد لغوية و إعرابية تفيدك في رحلتك الأدبية واللغوية"¹⁷.

ح- الجميل في النسخة الإلكترونية أنه مزود في آخره بفصل جُمعت فيه مجموعة من المصطلحات الحضارية أُرِفقت بالشواهد الصورية لِيَسْهُلَ تلقيها.

- أصبحت حوسبة اللغات الطبيعية ومعالجتها آليا أمرا ضروريا ومشروعا حيويا، تنافست الأمم في إنجازها ضمن ما يسمى المجتمع الشبكي.
- المعجم الإلكتروني هو نتاج تطبيق علم الالكترونيات وعلوم الحاسوب في مجال الصناعة المعجمية واللسانيات الحاسوبية.
- يحتوي المعجم الإلكتروني على عدة تطبيقات لغوية مهمة يمكن للمستخدم أن يستفيد منها كتصريف الأفعال والأسماء، البحث عن المترادفات، المعالجة على المستوى الصوتي لتحويل المكتوب إلى منطوق، التدقيق الإملائي لتصويب الكلمات المتداخلة.
- إن الحاجة لبناء معجم رقمي عربي أضحت مطلبا ضروري في ظل ما يشهده العالم من عصنة وتطور تكنولوجي، بغية صياغة تراث الدراسات اللغوية صياغة جديدة في ظل ما تحققة اللسانيات الحاسوبية من إنجازات في اللغات الأخرى.
- معجم الرائد من المعاجم التي لاقت رواجاً وإقبالاً رهيباً بالأخص فئة المتمدرسين كون الحصول عليه سهل ومجاني، وكذا احتوى على مصطلحات تواكب العصنة والتطور الحضاري، وتطور آداب اللغة العربية وعلومها.
- ويبقى معجم الرائد منجزاً؛ يعمل في حدود جمع المادة اللغوية الموجودة بالفعل بشكل ورقي، ثم صياغتها رقمياً وهذا في حد ذاته نقلة حديثة في مجال حوسبة المعجم العربي.

الشواهد:

- ¹ - المحتوى الرقمي باللغة العربية والبرمجيات، مخبر الممارسات اللغوية في الجزائر، منشورات مخبر الممارسات اللغوية في الجزائر، 2014، عدد خاص، ص512.
- ² - ارشادات مشاريع رقمنة مجموعات الحق العام، هبة ملحم، قراءة ومراجعة، عبد اللطيف صوني، 2013، ص 15، 16.
- ³ - اللسانيات الحاسوبية والترجمة الآلية، بعض الثوابت النظرية و الإجرائية، سناء منعم، تقديم مصطفى بوعناني، عالم الكتب الحديث، ط1، إربد، الأردن، 2015م ص102.
- ⁴ - المعاجم الإلكترونية، العربية وآفاق تطويرها، فعاليات المؤتمر الدولي الرابع في اللغة والترجمة، الصناعة المعجمية، الواقع والتطلعات، غز الدين البوشيخي، مركز أطلس العالمي للدراسات والأبحاث، جامعة الشارقة، الإمارات، يومي 20، 21 أبريل 2004، ص102.

- ⁵ - المعجم العربي الإلكتروني أهميته وطرق بنائه، عبد المجيد حمادو، الموسم الثقافي التاسع والعشرون لمجمع اللغة العربية الأردني، الأردن، 2011م، ص290.
- ⁶ - المرجع نفسه، ص292.
- ⁷ - المعجمات العربية وموقعها في المعجمات العالمية، محمود فهمي حجازي، (ندوة خاصة بمناسبة الانتهاء من معجم تاج العروس)، □ مجلس الوطني للثقافة والفنون، الكويت، 10 أكتوبر 2002، ص183.
- ⁸ - مقدمة في علم المصطلح، علي القاسمي، مكتبة النهضة، القاهرة، ط2، 1988م، ص 174.
- ⁹ - اللغة العربية والحاسوب، نبيل علي، دار تعريب، 1988م، ص93.
- ¹⁰ - المرجع نفسه، ص93.
- ¹¹ - مستقبل اللغة العربية، حوسبة المعجم العربي ومشكلاته اللغوية والتقنية، عبد الله أبو الهيف، مجلة التراث العربي، اتحاد الكتاب العربي، دمشق، 2004، ص144.
- ¹² - ، المعاجم الإلكترونية للغة العربية، محمد محمد الحناش ، 22 سبتمبر 2019، ص6.
- ¹³ - المعجم العربي الإلكتروني أهميته وطرق بنائه، عبد المجيد حمادو، ص297.
- ¹⁴ - المرجع نفسه، ص 299.
- ¹⁵ - المرجع نفسه، ص 299-300.
- ¹⁶ - بتصرف، المرجع نفسه، ص300.
- ¹⁷ - رائد الطلاب، جبران مسعود، دار العزة والكرامة للكتاب، بيروت لبنان، في خلفية المعجم.