

الجهود العربية في صناعة المحتوى اللغوي الإفرادي؛ من القوائم اللغوية والرصيد الوظيفي إلى المدونات الرقمية

Arab Efforts in Lexical Content Development: From Word Lists and

Functional Inventories to Digital Corpora

* جيلالي فاسي

Djilali FACI

جامعة البليدة 2، لونيسلي علي (الجزائر)

BLIDA 2 University, LOUNICI Ali (Algeria)

d.faci@univ-blida2.dz

تاريخ النشر: 2025/12/02

تاريخ القبول: 2025/11/16

تاريخ الإرسال: 2025/09/15

مَلِكُ حَصْرِ الْبَحْثِ

يتوجه البحث إلى تتبع الجهود العربية الفردية منها والجماعية في صناعة المحتوى اللغوي الإفرادي لأغراض تعليمية تتمثل في تقديم خدمة للمجالس المكلفة بإعداد المناهج التربوية والكتب المدرسية، وللمعلمين في تحضير الأنشطة التعليمية، وللمتعلمين في إثراء رصيدهم اللغوي. وكان المنطلق من ثلاثينات القرن العشرين بظهور القوائم اللغوية، ثم الرصيد اللغوي الوظيفي الخاص بثلاث دول من المغرب العربي، إلى أن توسع العمل بعد ذلك من أجل وضع قاعدة بيانات للمفردات تُمهّد للشروع في حوسبة اللغة العربية. ولقد أثمرت تلك الجهود المكثفة في توفير رؤية واضحة لهذه العملية، فظهرت المدونات الرقمية، وسمحت بإنجاز المعجم التاريخي للغة العربية، في انتظار تحقيق مشروع الذخيرة العربية. الكلمات المفتاحية: محتوى إفرادي؛ رصيد لغوي؛ قاعدة بيانات؛ معجم تاريخي؛ ذخيرة عربية.

Abstract:

This research traces the individual and collective Arab efforts to produce lexical content for educational purposes, aiming to serve councils responsible for preparing educational programs and textbooks, assist teachers in designing learning activities, and help learners enrich their lexical repertoire. The journey began in the 1930s with the emergence of linguistic lists, followed by the development of functional lexical inventories specific to three Maghreb countries. The work later expanded toward creating a lexical database to pave the way for the computerization of the Arabic language. These intensive efforts resulted in a clear vision of the process, leading to the

* جيلالي فاسي d.faci@univ-blida2.dz

creation of digital corpora and enabling the compilation of the historical dictionary of the Arabic language, while awaiting the completion of The Arabic Thesaurus project.

Keywords: lexical content, lexical repertoire, database, historical dictionary, arabic thesaurus.



مقدمة:

تحقيق التواصل بين الناس هو أهم وظيفة للغة، ومع التطور الصناعي والتقني الذي لا يتوقف اليوم في مجال الاتصالات المتعددة، وانتشار نقاط البث الإذاعي، وتعدد القنوات التلفزية الفضائية، واستعمال الهواتف الذكية بتطبيقاتها الكثيرة، وسهولة الولوج إلى مواقع الأنترنت، كل ذلك أسهم في تفعيل حركة التواصل الذي يبنى أساسا على اللغة. غير أن تلك الحركة النشيطة التي يشهدها التواصل الإنساني فرضت على اللغات الطبيعية تحديات جديدة من خلال الممارسات اللغوية الحديثة التي أثبتت وجودها بفعل سلطة الانتشار الواسع والسريع في آن واحد، وصار لها أسلوب خاص ومفردات جديدة واجتهاد جريء في نظام المنطوق والمكتوب. وهناك ما يفوق هذه التحديات خطورة إذا علمنا أن نصف لغات العالم مهددة بالانقراض، وأن من بين أكثر من ستة آلاف لغة مستعملة اليوم، هناك خمسمائة (500) لغة فقط -ومنها العربية- استطاعت أن تجد لها مكانا ولو ضعيفا على شبكة الأنترنت. إن هذا الأمر يستدعي النظر والتفكير من اللسانيين وأصحاب القرار في الحلول المناسبة لترقية اللغة العربية، وحمايتها من الاضطرابات التي تتعرض لها وسط هذا الصراع في عالم الاتصال، فاللغة العربية التي تحافظ على النظام اللغوي السليم وفق قواعد الفصح والأفصح لم يعد لها وجود إلا في مساحة ضيقة: في قاعات التدريس ومدركات أقسام اللغة العربية وبعض المنابر الخطابية والإعلامية. وتبدأ الحلول من توجيه الجهود الجماعية إلى إعداد محتوى إفرادي لتعزيز الرصيد اللغوي للمتعلمين، وفي هذا الإطار تبرز الإشكالية التي يسعى البحث إلى الإجابة عنها، وهي: كيف توجهت الجهود العربية من الأفراد والهيئات الرسمية إلى صناعة المحتوى اللغوي الإفرادي، وما المراحل التي مرت بها العملية؟

1. القوائم اللغوية:

ظهرت منذ الثلث الثاني من القرن العشرين مجموعة كبيرة من القوائم اللغوية التي أحصت المفردات الأكثر شيوعا في العربية، إذ أنجز "بريل" سنة 1940م قائمة للمفردات الجارية على الجرائد العربية اليومية، فأحصى 5981 مفردة لمساعدة المعلمين ومؤلفي الكتب المدرسية في انتخاب الكلمات السليمة¹. وفتح هذا الإنجاز المجال لظهور قوائم لغوية متتالية؛ فكان منها على سبيل التمثيل²:

- قائمة فاخر عاقل التي ارتبطت بالمفردات الشائعة في كتب القراءة للسنوات الثلاث الأولى في فلسطين ولبنان وسوريا ومصر والسعودية والعراق (1953م)؛

- قائمة مفردات النثر العربي الحديث ليعقوب لاندوا لمساعدة بني جنسه من اليهود على تعلم العربية (1959م)؛

- قائمة رشدي خاطر، اعتمدت على قائمة فاخر، وبلغت 1069 كلمة نوعية (1959م)؛
 - قائمة محمد قدرى لطفي، وهي جزء من رسالته للدكتوراه التي قدمها لجامعة شيكاغو (1964م)؛
 - القائمة المدرجة لمتعلمي العربية، أشرف عليها فريق مشترك من جامعة متشجن والجامعة الأمريكية بالقاهرة (1967م)؛
 - قائمة فتحي يونس بعنوان: الكلمات الشائعة في كلام تلاميذ الصفوف الأولى من المرحلة الابتدائية وتقويم بعض مجالات تدريس اللغة في ضوءها، وهي في الأصل رسالته للدكتوراه عام 1974م؛
 - قائمة الرياض التي جمعها داود عبده (1979م)؛
 - قائمة معهد الخرطوم الدولي التابع للمنظمة العربية للتربية والثقافة والعلوم (1981م)؛
 - قائمة مكة للمفردات الشائعة الصادرة عن معهد اللغة العربية بجامعة أم القرى (1981م)، واعتمدت اللجنة أثناء إعدادها على قوائم الرياض والخرطوم والجامعة الأمريكية بالقاهرة، وعلى الرصيد اللغوي الوظيفي للمغرب العربي الذي سيأتي الحديث عنه بعد قليل؛
 - قائمة رشدي أحمد طعيمة للمفردات شائعة الاستخدام في البلاد العربية (1982م)؛
 - قائمة المفردات الشائعة في اللغة العربية من العصر الجاهلي إلى العصر الحديث، من إعداد خالد السادات إبراهيم والسيد العربي يوسف والبراء صفوان عبد الغني وإشراف فتحي علي يونس، وهي في طبعها الأولى من ثلاثة أجزاء (2018م).
- الملاحظ على تلك القوائم أنها جهود فردية إلا قائمة الخرطوم التي صدرت عن هيئة رسمية، وقوائم الرياض ومكة والقاهرة التي تنتمي كل واحدة منها إلى مؤسسة جامعية، وأن أصل بعضها رسالة دكتوراه، وبعضها يعتمد على بعض آخر سابق مما يؤدي في النتيجة إلى تشابه في مفردات القوائم، ولكن هذا التشابه لا يمنع ورود مفردات متباينة تميز كل قائمة عن الأخرى؛ لأن لاختلاف الأقاليم العربية أثرًا في اختلاف المفردات. وليس ذلك عيبًا، بل هو دليل على ثراء اللغة العربية وتنوع أدائها، وقد حصل مثل ذلك بين القبائل العربية في زمن السماع اللغوي مثلما هو معلوم. كما يلاحظ أن بعض القوائم لا تنتمي إلى العملية التعليمية بشكل مباشر، مثل قوائم بريل ولانداو وطعيمة، والقائمة التي أحصت المفردات من العصر الجاهلي إلى العصر الحديث، ولذلك جاء عدد مفرداتها كبيرًا لكنها لا تؤدي مفاهيم كثيرة بسبب التداخل بين المفردات مما لا يسمح باستثمارها في ميدان تعليم الأطفال، وهي المشكلة ذاتها التي يعاني منها الكتاب المدرسي، إذ بينت العمليات الإحصائية لعدد من الكتب المقررة على المرحلة الابتدائية في بلدان الوطن العربي أنها تعاني من فقر في المفاهيم والمعاني مع حشو كبير للمفردات³.
2. الرصيد اللغوي الوظيفي:

الرصيد اللغوي العربي هو مجموعة المفردات العربية التي تؤدي مفاهيم الطفل في سن معينة، وتستجيب لحاجاته، والتي يمتلكها الطفل عند دخوله المدرسة للانطلاق منها في المناهج الدراسية، والبناء عليها لتنميتها وتطويرها⁴. وظهر في هذا الباب مشروع "الرصيد اللغوي الوظيفي للمرحلة الأولى من التعليم الابتدائي" الذي أعده فريق اللجنة الدائمة للرصيد الوظيفي المكونة من:

- عبد الرحمن الحاج صالح: من معهد العلوم اللسانية والصوتية بالجزائر؛
 - أحمد العابد: من قسم اللسانيات بمركز الدراسات والأبحاث الاقتصادية والاجتماعية بتونس؛
 - أحمد الأخضر غزال: من معهد الدراسات والأبحاث للتعريب بالرباط.
- كان ذلك تبعا لقرار ندوة وزراء التربية والتعليم بالمغرب العربي التي عُقدت بتونس من 14 إلى 20 فبراير 1967، وصدر المشروع بتونس سنة 1976م، وقد أنجز في دول المغرب العربي الجزائر وتونس والمغرب من خلال منهجية العمل التالية⁵:
- جرد الكتب المستعملة في الدول الثلاث في المرحلة الأولى من التعليم الابتدائي وإحصاء مفرداتها وتواترها؛
 - إجراء تحريات لغوية في مناطق مختلفة من الدول المعنية وتسجيل محاورات تلقائية لأطفال من سن الخامسة إلى التاسعة؛
 - معالجة التسجيلات إلكترونيا لتحديد درجة الشيع والتواتر؛
 - بعد ذلك، بدأت اللجنة العمل النهائي لترتيب المفردات وفق المقاييس التالية:
 - التواتر: فلا بد للمفردة من ورودها أكثر من مرة؛
 - التوزع: أي تكون المفردة مشتركة بين الأقطار الثلاثة أو على الأقل بين قطرين منها؛
 - الابتعاد عن الترادف قدر الإمكان: فالغرض من هذا الرصيد هو الحصول على أدنى عدد من المفردات خلافا للقوائم اللغوية المذكورة سابقا؛
 - الكمون: ومعناه إضافة مفردات لم ترد في التسجيلات لإمكان توقع ورودها على لسان الطفل، والسبب في ذلك محدودية التسجيلات؛
 - ضرورة التدخل: ويكون التدخل من قبل اللجنة باختيار أصلح الألفاظ وضبط مصطلحات تقنية وعلمية وسد الفراغات؛
 - اعتبار البعد العربي: وذلك لمنع القطعية في الزمان والمكان، فواجب الرصيد هو توثيق الصلة بالماضي وبالبلدان العربية الأخرى.

ووضعت اللجنة آليات لضبط الاستعمال الحسن للرصيد الوظيفي بالشكل التالي⁶:

- ترتيب المفردات في الرصيد هو ترتيب ألفبائي حسب تتابع الحروف لا حسب الجذور؛
 - ترتيب الهمزة حسب وضعية كتابتها: الألف ثم الواو ثم الياء، والألف الممدودة بعد الألف المهموزة؛
 - الحركات مرتبة كما يلي: الفتحة ثم الضمة ثم الكسرة فالكسرة والتشديد؛
 - يُبحث عن الفعل الأجوف في الألف لا في الواو أو الياء (قال وليس قول)؛
 - يبحث عن الفعل المضعّف في لفظه الثاني لا الثلاثي (شدّ وليس شدد)؛
 - ترك الإشارة إلى جذر الفعل الثلاثي والمزيد على وزن فعّل لوضوح ذلك.
- يلاحظ على هذا المشروع اعتماده على اللغتين العربية والفرنسية كما هو واضح في الشكل التالي من باب الهمزة:

!	!
eh ! ô, oh !	أ (أداة نداء)
est-ce que ?	أ...؟ (أداة استنفاع)
n'est-ce pas ?	أليس كذلك؟
brique	أجرّ و يَجُرّ (ة)، [أجر]
autre	آخر، ج: ون، م: أخرّ، ج: أخرّيات
	[أخر]

الشكل رقم 01: طريقة التأليف في الرصيد اللغوي الوظيفي (ص01)

والسبب في ذلك أن المدرسة المغاربية في ذلك الزمان كانت تتبنى نظاما تربويا مزدوج اللغة؛ لقرب عهدها بالاستعمار الفرنسي ولقلة المعلمين ذوي الكفاءة التعريبية اللازمة. هذا، ووعدت اللجنة الدائمة بطباعة نسخة أخرى إنجليزية عربية وعربية إنجليزية، مع إضافة شروح للكلمات في الطبعة الثانية وتدعيمها بالصور.

كما عُقد بالقاهرة بين 29 و31 يناير 1989م اجتماع الخبراء العرب في اللسانيات الحاسوبية الذي نظّمته اللجنة الاقتصادية والاجتماعية لغربي آسيا، تحت رعاية مكتب الأمم المتحدة، بالتعاون مع البنك الإسلامي للتنمية. ولم يقتصر العمل هذه المرة على بلدان المغرب العربي، بل شارك فيه خبراء من دول عربية أخرى، وتشكل فريق عمل مختلط بين باحثين لسانيين ومهندسين في الحوسبة، وهم كالتالي⁷:

- سالم غزالي: معهد بورقبة للغات الحية، تونس؛
- شريف الدي: مركز الكويت العلمي IBM؛
- عبد الرحمن الحاج صالح: معهد علوم اللسان، جامعة الجزائر؛
- عبد القادر الفاسي الفهري: كلية الآداب بجامعة محمد الخامس، المغرب؛
- كمال بشر: كلية دار العلوم، مصر؛
- محمد فرحات، قسم الهندسة بمعهد الكويت للأبحاث العلمية؛
- محمد حشيش، رئيس مركز القاهرة العلمي IBM؛
- محمد مندورة: عميد كلية الحاسبات والمعلومات بجامعة الملك سعود، السعودية؛
- مظفر الجراح: رئيس قسم الأبحاث بالمركز القومي للحاسبات الإلكترونية، العراق؛
- ناديا حجازي: مركز الإلكترونيات بمجلس البحث العلمي، مصر؛
- نبيل علي: مدير قسم الدراسات والأبحاث والتطوير، العالمية للإلكترونيات، مصر؛

- يحيى هلال: قسم المعلوماتية بجامعة محمد الخامس، المغرب.

قدّم في الاجتماع ممثل الجزائر عبد الرحمن الحاج صالح تقريراً حرّره في الموضوع يتعلق بمستلزمات بناء قاعدة آلية للمفردات العربية⁸، وسجّل التقرير اختلافاً في وجهات النظر والفهم لدى المشاركين حول كيفية تحديد قاعدة معطيات (Une base de données) للمفردات العربية التي ستبنى عليها الأعمال اللغوية الحاسوبية، وعلى رأسها بناء المعجم، وما يتعلق بذلك من تحديد نوعية اللغة التي يجب اعتمادها، والنظر في إشكالية المداخل والقواعد الخاصة بالمعطيات الإفرادية.

3. مستلزمات بناء قاعدة معطيات للمحتوى الإفرادي:

لاحظ الحاج صالح وجود تباين في وجهات النظر لدى الخبراء المشاركين في الاجتماع المذكور، فمن خلال استعراض عناوين البحوث المقدمة تجلّى الاختلاف في الفهم بين هؤلاء، ولكن جمعتهم الفكرة الأساسية حول بناء قاعدة المعطيات، واشتركت رؤيتهم للموضوع في جملة من الاقتراحات ذكر التقرير منها⁹:

1.3. اعتماد اللغة الحديثة:

رأى الخبراء في المواصفات العامة لقاعدة المعطيات ضرورة اعتماد اللغة الحديثة بمعانيها المستحدثة، ودعوا إلى الانطلاق من مدونة تشتمل على مجموعة نصوص حديثة. ويمكن اليوم أن يستفيد الباحثون العرب من المشروع الذي أنجزه جون سنكلير (John Sinclair) وزملاؤه في جامعة برمنغهام والمعروف بمشروع كويبلد (Cobuild)، وهو بمنزلة

بنك لغوي يهدف إلى وضع معجم تعليمي للغة الإنجليزية المعاصرة، يساعد متعلميها من أهلها ومن غير أهلها على فهم أعمق لتراكيبها واستعمالاتها اللغوية بعد حصر المفردات والتراكيب الشائعة، وبيان نسبة شيوع كل منها. كما اهتم صانعو المعجم الإنجليزي بوضع أكبر عدد ممكن من الاستعمالات اللغوية لأقل عدد من المفردات، وهو ما يجعل تعلم اللغة والتكلم منها أمراً ميسوراً¹⁰. ولقد كانت النتائج محمودة جداً، فاستطاعوا أن ينجزوا معجماً في غاية الجودة والإنقان، مع سهولة الاستخدام، بعد أن عالجوا الصعوبات التي تعترض المشروع، ومنها:

— حصر المشكلات الأساسية في تعلم اللغة؛

— أهمية الاستعمال اللغوي والوصف البراغماتي لمفردات اللغة؛

— التركيز على الإحصاء ووضع قوانين لذلك.

2.3. الاكتفاء بالمفردات الشائعة:

من الاقتراحات المهمة والمتفق عليها بين جميع الخبراء المشاركين في اجتماع القاهرة ضرورة الاكتفاء بالشائع من المفردات واعتماد المتواتر منها، ويتطلب ذلك — بالمقابل — اجتناب الغريب غير المعروف عند الجميع. وأشار أحد المتدخلين إلى أن المعاجم العربية التي تحمل صفة المعاصرة كالمعجم الوسيط مثلاً — لم تواكب التغيرات المعجمية والدلالية التي هي صفة لازمة لجميع اللغات ومنها العربية. وسبب هذه المشكلة راجع إلى البناء غير الصحيح لمدونة النصوص التي تشكل القاعدة الأساسية للمعطيات الإفرادية، فإلى جانب شرط الحدثة الواجب تحقُّقه في النصوص المختارة، لا بد من تحقق شرط آخر: هو أن تكون النصوص منبثقة من واقع التبادل اللغوي الفعلي، لا من أمثلة يخترعها مؤلفو القاموس¹¹.

لكن في تحديد المقصود بالغريب غموض شديد، فالتقرير ينص على اتفاق الباحثين على ضرورة الاكتفاء بالشائع والمتواتر من المفردات، واجتناب الغريب غير المعروف عند الجميع. فما معيار الشيوع الذي ينبغي الالتزام به؟ أهو الشيوع بالنسبة للباحثين أم للمتعلمين؟ إن سبب هذا التساؤل هو الحشية من أن تُعدَّ مفرداتٌ من قبيل: ما فتى، وما انفك، واخلوق، ودري... غريبة لأنها لم تُعدَّ مستعملةً ولا شائعة لدى الغالب الأعم من الطلاب، واجتناب مثل هذه المفردات من شأنه أن يؤدي إلى انحسار مساحة الدلالة في استعمالات قليلة بناءً على اعتماد مبدأ الشيوع.

3.3. التصنيف الدقيق للمفردات:

حسب هؤلاء الخبراء، فإن الاكتفاء بالمفردات الشائعة واجتناب الغريب منها يجعل التصنيف أكثر دقة، ويكون البحث الآلي أكثر سهولة، لأن حجم المعطيات المخزونة يصبح أقل ضخامة. ويرتكز التصنيف الدقيق للمفردات على جانبين متلازمين هما¹²:

— التصنيف من حيث اللفظ والفصائل النحوية؛

— التصنيف من حيث الدلالة وأجناس المعاني.

4.3. إشكالية المداخل:

تناول المشاركون إشكالية المداخل التي تُعتمد في التصنيف: هل يكون الترتيب جذريا كسائر المعاجم العربية أم أبجديا؟ وقد بين بعضهم الصعوبات التي يلاقها الطلاب أثناء عملية البحث في المعاجم العربية بسبب الترتيب الجذري الذي يقتضي الكثير من المعلومات الصرفية، فأشار أحد الباحثين إلى أن الاعتماد على الجذر ضروري؛ لأن الجذر هو أصل الأصول، ولكن ذلك لا يمنع من أن ندخل في الترتيب الجذري وفي مكانه الأبجدي كل الألفاظ التي فيها مشكلة ونحيلها إلى جذرها. وهناك مجموعة من المعايير اللغوية التي يجب أن تراعى عند تعامل محرك البحث مع المفردات، أهمها ما يلي¹³:

1.4.3. التدقيق اللغوي للمداخل:

عندما يكتب المستخدم في محرك البحث المفردة المراد البحث عنها، يقوم المحرك بالتدقيق اللغوي ويتدارك الخطأ إن حصل، فينبه المستخدم ويقترح المفردة الصحيحة ثم يعرض النتائج الخاصة بها مثلما هو مستعمل في محرك البحث جوجل Google. فعلى سبيل المثال: إذا كتب المستخدم كلمة "الإسلام" فمن المحتمل أن يكتب في مكان الحرف الصحيح "إ" واحدا من الحروف التالية "أ" - "آ"، فيراعي محرك البحث تلقائيا الصورة الصحيحة لكتابة المفردة المبحوث عنها.

2.4.3. أشكال الكلمات المعرّبة:

تختلف أشكال الكلمات المعربة من بلد عربي إلى آخر، فبعض الكلمات المعربة عن الفرنسية أو الإنجليزية يقع اختلاف في رسم شكلها العربي، ولنأخذ الأمثلة التالية:

— Google: جوجل/غوجل،

— Kilogramme: كيلوغرام/كيلوجرام،

— Romantique: رومانسي/رومانتيكي/رومانطيتي.

3.4.3. أشكال الكلمات المركبة:

هناك كلمات مركبة من جزأين بينهما مسافة مثل: عبد الله، أبو ظبي، ومن المستخدمين من يترك المسافة بين الجزأين، ومنهم من يستغني عنها، ففي المغرب يكتبون الأسماء التي أولها "بن" ملتصقة، مثل: بنعبد الله وبنشقرون، وعلى محرك البحث مراعاة كل الاستخدامات الواردة.

4.4.3. الكلمات التي تحتوي على حروف تنطق ولا تكتب:

قد يقع الخطأ في كتابة بعض الكلمات مثل:

— الرحمن: الرحان،

— لكن: لائن،

— داود: داوود.

فعلى محرك البحث مراعاة هذه الحالة التي لا يحترم فيها المستخدم قواعد الإملاء الصحيحة.

5.4.3. زوائد الكلمة:

يلحق الكلمة بعضُ الزوائد من السوابق واللاحق، مثل: كتاب- الكتاب- كتابي- بالكتاب- بكتابه... ففي هذه الحالة يقدم محرك البحث لمستخدميه جميع الحالات الواردة، كما أنه يراعي الحالات الإعرابية من حيث الأفراد والتنشئة والجمع، مثل:

- كتاب- كتابان- كتابين- كتب- كتباً؛
- تلميذ- تلميذة- تلميذان- تلميذتان- تلاميذ- تلميذات؛
- مسلم- مسلمة- مسلمان- مسلمتين- مسلمون- مسلمين- مسلمات.

6.4.3. الترادف:

كل المعايير السابقة تتعلق بالمستوى الصرفي، وهناك معايير أخرى دلالية أهمها ضرورة مراعاة الترادف. وبما أن اقتراحات الخبراء نصّت على اعتماد اللغة العربية الحديثة بمعانيها المستحدثة فإنه من الضروري أن يشير محرك البحث إلى جميع المفردات المرادفة للكلمة المبحوث عنها من الفصحى واللهجات وحتى الكلمات المعربة، ومثال ذلك:

- بنك/مصرف/مؤسسة مالية...
- حاسوب/حاسب/كمبيوتر...
- هاتف نقال/جوال/موبايل/خلوي...
- أمين عام/كاتب/سكرتير...

7.4.3. المشترك اللفظي:

إذا كان للمفردة الواحدة أكثر من دلالة، فإن محرك البحث يكتشف ذلك فيعرض على المستخدم جميع الخيارات الممكنة مع توجيه سؤال: أي المعاني تقصد؟ وما على المستخدم سوى تعيين معنى المفردة التي يريد البحث عنها. والمراد بالمشارك اللفظي هنا هو المعنى الأعم لا المعنى الاصطلاحي؛ فإذا كان البحث مثلاً عن كلمة "بجر" فإن محرك البحث يعرض اقتراحات عديدة منها:

- البحر الأبيض المتوسط؛
- البحر الميت؛
- البحر في عروض الشعر؛
- البحر المحيط لأبي حيان،
- البحر المحيط للزركشي...

وعندما يختار المستخدم المعنى المراد تأتي نتائج البحث محصورة في ذلك المعنى دون غيره.

8.4.3. اللبس الصرفي:

قد يؤدي اللبس الصرفي في بعض المفردات إلى عشوائية البحث بسبب الاختلاف الدلالي الذي يؤديه، مثل: اللّين/الَّذين، العِلْم/العَلَم، الكلام/الكَلَام... وقد يكون اللبس حاصلًا في بعض المفردات بين الاسمى والفعلية مثل: التَّهم/الْتَّهم، رُكِب/رُكِب، سَمِع/سَمِع...

ففي مثل هذه الحالات يمكن ضبط التشكيل من قبل المستخدم إذا كان على دراية بموضوع البحث، وإلا كان على المحرك اقتراح جميع الحالات الواردة في هذا اللبس الصرفي، ثم يقوم المستخدم بتعيين المعنى المراد.

5.3. تحديد النظرية اللغوية المناسبة:

تطرق بعض المشاركين في اجتماع القاهرة إلى كيفية التحليل اللغوي وتحديد النظرية اللغوية التي يجب اعتمادها في ذلك، فالتحليل الآلي الذي يحقق بناء قاعدة للمفردات العربية لا يمكن أن يتأسس إلا على الأنماط اللغوية والرياضية التي يتابعها الخبير اللساني بالاشتراك مع الخبير الرياضي في إطار العمل العلمي متعدد الاختصاصات. ونلمح هنا الحاجة إلى أن يشتغل اللساني مع الرياضي في حقل واحد، إذ يهيئ الأول الأنماط اللغوية حتى يتمكن الثاني من إدخالها في ذاكرة الحاسوب¹⁴.

وتحديد النظرية اللغوية المناسبة لهذا العمل ضروري جدا بعدما اتضح من المشاركين اختلاف مذاهبهم اللسانية من البنوية الأوربية أو الأمريكية إلى التوليدية التحويلية، بل هو أهم عنصر في المواصفات العامة للمشروع. وكان مثل الجزائر يشير إلى النظرية الحليلية التي يراها أنسب النظريات وأقربها إلى موضوع الاجتماع، ومن شأنها أن توفر على المهندس واللساني جهودا كبيرة في مجال البحث عن الآليات الحاسوبية. إن مفاهيم النحو الأصيل للخليل وسيبويه ومن نحا نحوهما - كفهوم المثال وحد اللفظة - كثيلة بأن تزيل التعقيد في عمليات التحليل التي تظهر أثناء الانطلاق من الأفكار البنوية الغربية وتطبيقها بكثير من التّعسف على الكلام العربي.

4. المدونات الرقمية:

بعد تلك الأشواط الطويلة، انتقلت الجهود العربية إلى إصدار مدونات لغوية رقمية، وفي تعريف المدونة اللغوية يقول ماكزري وزميله McEnery et al إنها مجموعة من نصوص اللغة الطبيعية¹⁵. ومن التسميات الشائعة للمدونات:

– الذخيرة اللغوية: عبد الرحمن الحاج صالح؛

– الذخيرة النصية: مها الربيع في مشروعها المسمى: الذخيرة النصية الفصحى لجامعة الملك سعود؛

– المكنز: عبد الغني أبو العزم في مقاله: اللغة العربية والمعالجة الآلية برامج صخر نموذجاً؛

– لسانيات المتون: صالح فهد العصبي في دراسته: لسانيات المتون قضايا أساسية في التأصيل والتطبيق والمنهج. إن أهم ما يميز المدونات اللغوية هو حجمها الكبير، فعلى سبيل المثال، تضم مدونة صخر خمسمائة (500) مليون كلمة، وفي المدونة العربية لمجمع الملك سلمان أكثر من مليار كلمة بالتكرار، وأكثر من تسعة ملايين كلمة من دونه¹⁶. وأما مدونة ArTen Ten -وهي مدونة إلكترونية مركبة في موقع Sketch Engine- فقد حوت حسب إحصائيات سنة 2024م

مليارا ونصف مليار كلمة، وثلاثمائة مليون جملة، وهي مدونة تتطلب التسجيل من خلال اسم المستخدم وكلمة المرور¹⁷. غير أننا نرغب في الوقوف على مشروعين اثنين: المعجم التاريخي للغة العربية، ومشروع الذخيرة العربية.

1.4. المعجم التاريخي للغة العربية:

تعود فكرة المعجم التاريخي إلى تأسيس مجمع فؤاد الأول للغة العربية بالقاهرة عام 1932م، وكان من أعضائه المستشرق الألماني أوغست فيشر August Fischer الذي شرع عام 1936م في تأليف معجم عربي على منهج علمي حديث، ورحب الأعضاء بذلك إذ كان من أهداف المجمع -مثلا تنص عليه المادة الثانية- إنشاء معجم تاريخي للغة العربية، ونشر أبحاث دقيقة في تاريخ الكلمات وتغير دلالاتها. ولكن لم يستمر العمل بعد اندلاع الحرب العالمية الثانية ثم مرض فيشر وموته في 1949م، فتابعت جهود المجمع إلى إصدار المعجم الوسيط ثم المعجم الكبير في جزئه الأول عام 1970م. وفي عام 2013م، أطلق المركز العربي للأبحاث ودراسة السياسات في قطر مشروع معجم الدوحة التاريخي، وهو مشروع يتكون من ثلاث مراحل¹⁸:

– المرحلة الأولى: تمتد من القرن الخامس قبل الهجرة إلى نهاية سنة 200هـ؛

– المرحلة الثانية: من 201هـ إلى 500هـ؛

– المرحلة الثالثة: من 501هـ إلى العصر الحاضر، فهي مرحلة مفتوحة.

أنجزت إلى اليوم المرحلتان الأولى والثانية، وبلغ عدد المداخل المعجمية مائة وثمانين (180) ألف مدخل، ولا يزال العمل جاريا لاستكمال المداخل المعجمية التي تمتد إلى العصر الحديث. وشارك في بناء المعجم أساتذة من بلدان عربية مختلفة منها الجزائر، ومن خصائصه التي تميزه عن غيره من المعاجم أنه:

– معجم مفتوح وقابل للتحديث المستمر؛

– مبني من مدونة نصية واسعة؛

– يرجع باللفظ إلى استعماله الأولى في النقوش إن وجدت؛

– يرصد ألفاظ الحياة العامة وفق ضوابط علمية ومنهجية محددة؛

– يرصد المصطلحات في مجالاتها العلمية والمعرفية والفنية؛

– يربط الألفاظ بنظائرها في اللغات السامية متى أمكن ذلك؛

– يرجع باللفظ المقترض إلى أصله الفارسي أو اليوناني أو غير ذلك؛

– لا يستطرد في التعريفات اللغوية والاصطلاحية.

هذا عن معجم الدوحة، وأما معجم الشارقة التاريخي فكان من إعداد اتحاد المجمع اللغوية العلمية العربية الذي أقامه حاكم الشارقة سلطان بن محمد القاسمي، وللمجلس الأعلى للغة العربية الجزائري مساهمة حسنة فيه، وكان المنهج أن تقسم العصور إلى خمسة:

- ما قبل الإسلام: (... قبل الهجرة - 1 قبل الهجرة)؛
 - العصر الإسلامي: (1هـ - 132هـ)؛
 - العصر العباسي: (133هـ - 656هـ)؛
 - عصر الدول والإمارات: (657هـ - 1213هـ)؛
 - العصر الحديث: (1214هـ/1798م - 1441هـ/2020م)
- صدر المعجم شهر نوفمبر 2024م في مائة وسبعة وعشرين (127) مجلدا، وله اليوم موقع إلكتروني قابل للتحديث، إذ بلغت معطياته إلى غاية اليوم ما يمكن تسجيله في الجدول التالي¹⁹:

28471	الأفعال	78079	المفردات
49448	الأسماء		
160	الأدوات		
11987			الجنود
29000	قبل الإسلام	408956	الشواهد
80000	الإسلامي		
194000	العباسي		
53000	الدول والإمارات		
54000	العصر الحديث		

الجدول رقم 01: إحصاء المفردات والجنود والشواهد في معجم الشارقة

2.4. مشروع الذخيرة العربية:

الذخيرة اللغوية في تصورها هي بنك رقمي ذو استعمال واسع في مجال البحث العلمي، فهو يوفر الوقت والجهد، ويعرض للباحثين مجموعة من المراجع لم يكونوا بالغيها لولاه إلا بشق الأنفس. وتخضع النصوص اللغوية التي يحتويها هذا البنك إلى قواعد محددة من حيث أساليب الجمع، ونسب التمثيل، وطرق المعالجة قبل عملية الجمع وبعدها، ومنهجيات الترميز حسب متطلبات البحث اللغوي. وقد جرى الغرف بين علماء اللغة أن المقصود من مصطلح "ذخيرة لغوية" ذلك العدد الهائل من الشواهد اللغوية الطبيعية الذي يُستخدم أساسا لإجراء البحوث اللغوية²⁰. ولا تقتصر الذخيرة اللغوية على اللغة التحريرية فقط، ولكنها تحتوي أيضا على عيّينات صوتية من اللغة الشفهية، ومنذ أربعين عاما أصبح هذا المصطلح يقترن بالمادة اللغوية الطبيعية التي تُخزن عن طريق الحاسوب²¹.

أما ما يتعلق بالذخيرة العربية، فقدّم عبد الرحمن الحاج صالح هذا المشروع الذي بدأت ملامحه الأولى في المؤتمر الواحد والستين (61) لمجمع اللغة العربية بالقاهرة عام 1995م، وكان نصّ المشروع قد نُشر من قبل ذلك في مجلة المجمع

الأردني للغة العربية عام 1986م، ثم في مجلة المجمع العلمي العراقي عام 1988م، وعُرض في نفس السنة على المنظمة العربية للتربية والثقافة والعلوم فتنبأها مجلسها التنفيذي، وعُقدت ندوة دولية في الجزائر في 1991م للبحث في كيفية إنجازه، ثم رُحِّب به اتحاد المجمع العربية في 1998م. وعلى إثر ذلك، قرَّر المجلس الوزاري بجامعة الدول العربية تبني المشروع بقرار رقم 6457.د.ع (121) بتاريخ 2004/9/14 باعتباره مشروعاً قومياً. وبعد أن كان يقتصر على الجانب اللغوي؛ فإنه الآن يتجاوزه إلى الجوانب الثقافية والعلمية، فسمي بالذخيرة العربية بدلا من الذخيرة اللغوية.²²

يرتكز التصور العام للذخيرة العربية على البرمجة الحاسوبية وعلى النظرية الخليلية الحديثة التي تزوِّم إعادة إحياء التراث اللغوي العربي وتطوِّره ليتوافق مع مستجدات الدرس اللساني الحديث في مجال الحوسبة اللغوية، ومتطلبات تعليم اللغة العربية بمراعاة التقنيات الحديثة، وعن الهدف الأول للمشروع يقول الحاج صالح: "إن الهدف الرئيسي لمشروع الذخيرة هو أن يمكِّن الباحث العربي أيا كان وأينما كان من العثور على معلومات شتى من واقع استعمال العربية بكيفية آلية وفي وقت وجيز. وهذا سيتحقق بإنجاز بنك آلي للغة العربية المستعملة بالفعل، يتضمن أمحات الكتب التراثية الأدبية والعلمية والتقنية وغيرها، وعلى الإنتاج الفكري العربي المعاصر في أهم صوره، بالإضافة إلى العدد الكبير من الخطابات والمحاورات العفوية بالفصحى في شتى الميادين"²³.

وعليه، يمكن بيان فوائد المشروع في النقاط التالية:²⁴

– الاعتماد في وضع المصطلحات والبحث عنها على كل المعطيات اللغوية في ميدان معين من واقع الاستعمال للغة العربية قديماً كان أم حديثاً؛

– الاعتماد في اختيار اللفظ على مقياس الشيع والدفقة في دلالة المعنى المراد؛

– الاعتماد على هذا البنك النصي الآلي في البحث عن التطور الدلالي للألفاظ العربية، ومن ثم إمكانية وضع معجم تاريخي دقيق للغة العربية؛

– إمكانية فهرسة آلية لكل النصوص العربية ذات القيمة العلمية والأدبية مما طبع وما سيطبع وينشر على مستوى الوطن العربي؛

– إقامة الدراسات العلمية المقارنة في مختلف الميادين حول مجموعة معينة من المفاهيم العلمية؛

– البحث المنتظم عن تطور الفكر العلمي العربي بالاعتماد على تطور دلالات الألفاظ العلمية في داخل حقول دلالية عبر الزمان؛

– إمكانية وضع معجم شامل للغة العربية المستعملة بالفعل تخصص لكل مدخل دراسة لغوية دقيقة.

إن الذخيرة العربية مشروع لا يستغني عنه أي باحث، ومنبِّغ للكثير من الدراسات في العلوم الإنسانية والتجريبية، لأنها تمكِّن كل من يلج ميدان البحث من إجراء العمليات التالية:

- رصد دقيق وشامل لاستعمال العربية في إقليم خاص في عصر من العصور؛
- رصد منتظم للاستعمال الحقيقي لمصطلحات ميدان فتي معين؛
- تصفح معاني الكلمات من خلال سياقاتها عبر الزمان، وتحديد تاريخ ظهور بعض الكلمات الفصيحة المولدة أو اختفائها؛

- تحليل لغة كاتب أو شاعر أو خطيب وإحصاء مفرداته بكيفية آلية. وعلى هذا الأساس، تُتاح الفرصة لكي تُولف أنواع كثيرة من المعاجم مثل²⁵:

- المعجم التاريخي للغة العربية؛
 - معاجم خاصة بأساء الأعلام والأماكن وغيرها؛
 - معاجم فنية في كل الميادين؛
 - معاجم أساسية ووظيفية لتعليم العربية؛
 - معاجم لألفاظ الحضارة قديما وحديثا؛
 - معاجم للغة الطفل العربي.
- يتضح من هذا العرض أن فكرة الذخيرة العربية أوسع بكثير من المعجم التاريخي، بل إن المعجم التاريخي لا يمثل سوى جزء يسير منها، ولهذا السبب شهد المشروع تأخرا في إنجازه، فلقد مات عبد الرحمن الحاج صالح عام 2017م وفي نفسه شيء من الذخيرة. وإنه لمن الوفاء له وللغة والدين والوطن أن يسعى الباحثون من بعده إلى مواصلة العمل من أجل إكمال هذا المشروع، فيكون علما يُنتفع به. ولا مجال لليأس إذا نظرنا إلى مشروع المعجم التاريخي الذي بدأت فكرته في ثلاثينات القرن الماضي، وكتب له أن يرى النور في عام 2018م.

5. خاتمة:

بعد مرور عقود على اجتماع القاهرة في 1989م، تبقى جملة التوصيات التي أقرها الخبراء المجتمعون صالحة للاستعانة بها في ميدان اللسانيات الحاسوبية من أجل وضع قاعدة معطيات آلية للمحتوى الإفرادي، والتي يمكن للقائمين على إعداد المناهج التربوية الخاصة بالمدرسة الابتدائية الرجوع إليها والاطلاع على بنك المفردات التي توفرها لهم، وقد اعترف الحاج صالح في تقريره بأن الطريق المؤدي إلى العلاج الآلي للنصوص العربية والذي يمر عبر بناء قاعدة البيانات الإفرادية طريق طويل وشاق، والمهم في كل ذلك هو وجود الإرادة القوية لتحقيق هذا الإنجاز خدمة للغة العربية وكل مستعمل لها. وفي النهاية، يتجه هذا البحث إلى تقرير النتائج التالية:

- الانتباه المبكر للباحثين المحدثين إلى أهمية صناعة المحتوى اللغوي الإفرادي لخدمة المتعلمين والمعلمين منذ ظهور القوائم اللغوية الأولى.
- مشروع الرصيد اللغوي الوظيفي قدّم خدمة معجمية لتوحيد المحتوى الإفرادي للمتعلمين والمعلمين في الدول التي شاركت في إنجازه، ومثل مرحلة أولى لظهور المدونات الرقمية.
- إصدار المعجم التاريخي للغة العربية بُسّختيه في الدوحة والشارقة إنجازٌ علمي كبير، ولا يزال بحاجة إلى التعريف به وحث الطلاب والباحثين على استخدامه فضلا عن الجمهور العام.
- تأخر إنجاز مشروع الذخيرة العربية راجعٌ إلى ضخامة العمل وقلة التنسيق بين الهيئات المعنية، ولا بد من تضافر الجهود واتحادها لتحقيق ذلك.
- أهمية اقتراب أصحاب الاختصاص بعضهم من بعض، فالباحث اللساني بحاجة إلى جهود المهندس في مجال الحوسبة، والمهندس يفتقر إلى معلومات اللساني، فالتعاون في المجالات البينية أمر ضروري.
- حوسبة اللغة العربية عملية مهمة لربط الأجيال بترائهم من جهة، ولانفتاح اللغة العربية على العالم المعاصر من جهة ثانية، وفي ذلك خدمة جليلة لها، ومعظم جهود المجلس الأعلى للغة العربية تتصل بهذا الميدان.

هوامش:

- 1 - يونس، فتحي علي وآخرون، قائمة المفردات الشائعة في اللغة العربية، مكتبة وهبة للطبع والنشر والتوزيع، القاهرة، ط1، 2018، ج1 ص20.
- 2 - ينظر:
- قائمة المفردات الشائعة في اللغة العربية (مرجع سابق)، ج1 ص20.
- الغالي، ناصر عبد الله، و: عبد الله، عبد الحميد، أسس إعداد الكتب التعليمية لغير الناطقين بالعربية، دار الاعتصام، دار النصر للطباعة الإسلامية، 1991، ص81، 82.
- قائمة مكة للمفردات الشائعة، سلسلة دراسات في تعليم اللغة العربية 1، وحدة البحوث والمناهج بمعهد اللغة العربية، جامعة أم القرى، مطابع الصفا، مكة، ص18.
- 3 - تازروتي، حفظة، الرصيد اللغوي العربي والتأليف المدرسي، مجلة اللغة العربية، المجلس الأعلى للغة العربية، المجلد 5، العدد 1، جوان 2003، ص245. [277-243]
- 4- ينظر: بن نافلة، يوسف، المعالجة الآلية وهندستها للمدونات اللغوية في ضوء اللسانيات الحاسوبية، أعمال الندوة الوطنية للغة العربية والتقانات الجديدة، منشورات المجلس الأعلى للغة العربية، الجزائر، ص17.
- 5 - الحاج صالح، عبد الرحمن وآخرون، الرصيد اللغوي الوظيفي، المعهد التربوي الوطني، الجزائر، ط1، 1975، ص ج.
- 6 - الرصيد اللغوي الوظيفي، ص ك.

- 7 - اللجنة الاقتصادية والاجتماعية لغربي آسيا والبنك الإسلامي للتنمية، مستلزمات بناء قاعدة معطيات للمفردات اللغوية العربية، 1989م، ص 03.
- 8 - لتحليل التقرير والأعمال المقدمة في اجتماع الخبراء العرب، يرجى الاطلاع على موقع المكتبة الرقمية للأمم المتحدة من خلال الضغط على الرابط التالي: <https://digitallibrary.un.org/record/1292382?ln=ar&v=pdf>
- كما أن التقرير مذكور في: بحوث ودراسات في اللسانيات العربية، ج 1 ص 97.
- 9- ينظر: الحاج صالح، عبد الرحمن، بحوث ودراسات في اللسانيات العربية، موف للنشر، الجزائر، ط 1، 2012، ج 1 ص 98.
- 10 - جمعة، عمرو، تقنيات اللغة العربية الحاسوبية -معايير التقييم ورؤى التطوير- مجمع الملك سلمان العالمي للغة العربية، الرياض، 1445هـ، ص 113.
- 11- ينظر: بحوث ودراسات في اللسانيات العربية، ج 1 ص 99.
- 12- نفسه، ج 1 ص 98.
- 13- ينظر: تقنيات اللغة العربية الحاسوبية، ص 17.
- 14- ينظر: بحوث ودراسات في اللسانيات العربية، ج 1 ص 101.
- 15- العصيمي، صالح بن فهد وآخرون، المدونات اللغوية العربية بناؤها وطرائق الإفادة منها، مجمع الملك سلمان العالمي للغة العربية، الرياض، 1445هـ، ص 20.
- 16 - ينظر موقع المدونة على الرابط التالي: <https://falak.ksaa.gov.sa/corpora/arabiccorpus>
- 17 - ينظر موقع المدونة على الرابط التالي: <https://www.sketchengine.eu/artenten-arabic-corpus>
- 18 - الكيسي، علي بن أحمد، معجم الدوحة التاريخي للغة العربية: نبذة تعريفية، مقال ضمن السجل العلمي للأوراق المقدمة في ملتقى الصناعة المعجمية العربية، مجمع الملك سلمان العالمي للغة العربية، ط 1، 2025م، ص 381 [379-388] ولمزيد من المعلومات يرجى الاطلاع على موقع المعجم في الرابط التالي: <https://www.dohadictionary.org>
- 19 - جرى الاطلاع على الموقع بتاريخ 2025/08/04م من خلال الرابط التالي: <https://almojam.org/statistics>
- 20 - خوانغ تشانغ نينغ - لي جوان تري، علم الذخائر اللغوية، تر هشام موسى المالكي، المركز القومي للترجمة، ط 1 سنة 2016، ص 25.
- 21 - نفس المرجع.
- 22 - بحوث ودراسات في اللسانيات العربية، ج 1 ص 395.
- 23 - نفسه، ج 1 ص 396.
- 24 - نفسه، ج 1 ص 410-411.
- 25 - نفسه، ج 1 ص 398.

قائمة المراجع:

1. بن نافلة، يوسف، المعالجة الآلية وهندستها للمدونات اللغوية في ضوء اللسانيات الحاسوبية، أعمال الندوة الوطنية للغة العربية والتقانات الجديدة، منشورات المجلس الأعلى للغة العربية، الجزائر.
2. تازروتي، حفيظة، الرصيد اللغوي العربي والتأليف المدرسي، مجلة اللغة العربية، المجلس الأعلى للغة العربية، المجلد 5، العدد 1، جوان 2003.
3. جمعة، عمرو، تقنيات اللغة العربية الحاسوبية -معايير التقييم ورؤى التطوير- مجمع الملك سلمان العالمي للغة العربية، الرياض، 1445هـ.

4. الحاج صالح، عبد الرحمن وآخرون، الرصيد اللغوي الوظيفي، المعهد التربوي الوطني، الجزائر، ط1، 1975.
 5. الحاج صالح، عبد الرحمن، بحوث ودراسات في اللسانيات العربية، موف للنشر، الجزائر، ط1، 2012.
 6. خواف تشانغ نينغ - لي جوان تزي، علم الذخائر اللغوية، تر هشام موسى المالك، المركز القومي للترجمة، ط1 سنة 2016.
 7. سلسلة دراسات في تعليم اللغة العربية 1، قائمة مكة للمفردات الشائعة، وحدة البحوث والمناهج بمعهد اللغة العربية، جامعة أم القرى، مطابع الصفا، مكة.
 8. العصبي، صالح بن فهد وآخرون، المدونات اللغوية العربية بناؤها وطرائق الإفادة منها، مجمع الملك سلمان العالمي للغة العربية، الرياض، 1445هـ.
 9. الغالي، ناصر عبد الله، وعبد الله، عبد الحميد، أسس إعداد الكتب التعليمية لغير الناطقين بالعربية، دار الاعتصام، دار النصر للطباعة الإسلامية، 1991.
 10. الكيسبي، علي بن أحمد، معجم الدوحة التاريخي للغة العربية: نبذة تعريفية، مقال ضمن السجل العلمي للأوراق المقدمة في ملتقى الصناعة المعجمية العربية، مجمع الملك سلمان العالمي للغة العربية، ط1، 2025م.
 11. اللجنة الاقتصادية والاجتماعية لغربي آسيا والبنك الإسلامي للتنمية، مستلزمات بناء قاعدة معطيات للمفردات اللغوية العربية، 1989م.
 12. يونس، فتحي علي وآخرون، قائمة المفردات الشائعة في اللغة العربية، مكتبة وهبة للطبع والنشر والتوزيع، القاهرة، ط1، 2018.
- المواقع الإلكترونية:
13. مدونة ArTen Ten: <https://www.sketchengine.eu/artenten-arabic-corpus>
 14. المدونة اللغوية لمجمع الملك سلمان العالمي للغة العربية: <https://falak.ksaa.gov.sa/corpora/arabiccorpus>
 15. المعجم التاريخي للغة العربية بالشارقة: <https://almojam.org/statistics>
 16. معجم الدوحة التاريخي للغة العربية: <https://www.dohadictionary.org>
 17. المكتبة الرقمية للأمم المتحدة: <https://digitallibrary.un.org/record/1292382?ln=ar&v=pdf>