موارد المُعجم التَّاريخيّ للُّغة العربيَّة

د. المُعتزّ بالله السَّعيد

جامعة القاهرة

moataz@cu.edu.eg

• **مُلَخَّص**

الكلِمات الدَّالَّة:

المُعجم التَّاريخيّ، صناعة المُعجم، موارد اللُّغة، اللِّسانيَّات الحاسوبيَّة.

تَسعى الدِّراسةُ إلى تقديم منهجيَّةٍ لإعداد مواردِ المعجم التَّاريخيّ للُّغة العربيَّة، وتهدفُ إلى إيجاد رؤيةٍ واضحةٍ لجوانبِ الإفادة من الموارِدِ المُتاحة (اللُّغويَّة والمُعجميَّة والحاسُوبيَّة) في صناعة المُعجم التَّاريخيّ المنشُود، كما تقترحُ الدِّراسةُ منهجيَّةً لتوظيف هذهِ الموارِدِ في صناعة المُعجم. وتأتي الدِّراسةُ في ستَّة محاورَ أساسِيَّة تتضَمَّن مُقدِّمةً ثُمَّ عرضًا لإشكالات الدِّراسة. ويلي ذلك تقديمُ المنهجيَّة المُقترَحة لإعداد موارد المُعجم التَّاريخيّ للُّغة العربيَّة، ثُمَّ عرضٌ لمنهجيَّة توظيفِ الموارِدِ ومراحل الصِّناعة المُعجميَّة في ضوء هذه الموارد. وأخيرًا يستعرضُ الباحثُ نتائجَ الدِّراسة، ويعرضُ خُلاصةَ بحثه.

The Historical Arabic Dictionary Resources

Almoataz B. Al-Said

Abstract

Keywords:
Historical Dictionary, Lexicography, Language Resources, Computational Linguistics.

The study seeks to provide a methodology for preparation resources of the desired historical Arabic dictionary. The study aims to set a clear vision to take advantage of the available resources in building the historical Arabic dictionary. The study also suggest a deployment methodology of the resources in building the desired dictionary. The study comes in six main sections including the introduction then a presentation of the main problems, followed by an introduction of the proposed methodology for preparation of the historical Arabic dictionary resources; then a display of a deployment methodology of the resources and the stages of building the dictionary. Finally, a review of the results is introduced then the conclusions are discussed.

١. مُقدِّمة.

١,١. في ماهيَّة المُعجم التَّاريخيّ.

ظهرَت فِكرةُ المُعجَمات اللُّغَوِيَّة التَّارِيخِيَّة *Historical Dictionaries* في النِّصف الأوَّل من القرن التَّاسِع عشر نتيجةً لظهور منهج التَّحليل التَّاريخيّ في الدَّرس اللُّغَوِيّ. وكانَت الغايةُ من هذه المُعجمات: حفظ التُّراث الإنسانيّ، والتَّعرُّف على مراحل تطوُّر الألفاظ والمُفرَدات في الشَّكل والمعنى منذُ أوَّل ظُهور مُسَجَّلٍ لها، ودِراسة أوجُه التَّبايُن اللُّغَويّ بينَ الحِقَب الزَّمنيَّة المُتعاقبة؛ بالإضافة إلى مُتابَعة التَّطوُّرات الحادِثة في ظواهر اللُّغات الإنسانيَّة وَفقًا لمُستويات التَّحليل اللُّغَوِيّ الَّتي تهتَض بها عُلُومُ اللُّغَة (الأصوات *Phonology*، والبِنية *Morphology*، والتَّركِيب *Syntax*، والدّلالة *Semantics*). وساهمَت المُعجماتُ التَّاريخيَّةُ المُنجَزة للُّغات الجرمانيَّة (الألمانيَّة والهولنديَّة والإنجليزيَّة والسّويديَّة) في إرساء قواعد الصِّناعة المُعجميَّة الحديثة وتطوير مناهج البحث في الدَّرس اللُّغويّ بمختلف ميادينه وفُرُوعه (١)، كما ساعَدَت المادَّةُ التَّأثيليَّةُ الَّتي تحويها على إنجاز العديد من الدِّراسات والبُحوث حولَ إعادة بناء اللُّغة الأمِّ في الفصيلة اللُّغَويَّة الواحدة.

١,٢. موارد الصِّناعة المُعجميَّة.

تقومُ الصِّناعةُ المُعجميَّةُ على مجموعةٍ من الإجراءات الَّتي تتحدَّدُ معالمُها في ضوء الغاية من المُعجم المنشُود. وقد كانَ للطَّفرة العلميَّة والمعلوماتيَّة الَّتي شَهِدَها العصرُ الحديثُ أثرٌ بالغٌ في تقويم صناعة المُعجمات اللُّغَويَّة؛ حيثُ اتَّجَهَ صُنَّاعُ المُعجمات إلى الإفادة من مواردَ وأدواتٍ لم تكُن موجودةً من قبلُ في بناء مُعجماتهم وتحريرها، بعدما كانوا يعتمدونَ على المُعجمات القديمة ونُصُوص التُّراث وأبحاث اللُّغَوِيِّين حولَ أصول المُفردات ومعانيها. ولم نلحظ تغيُّرًا ملموسًا في مناهج الصِّناعة المُعجميَّة إلَّا في منتصف القرن الثَّامن عشر. فمع ظُهُور المُدَوَّنات اللُّغَويَّة *Linguistic Corpora* سعى المُعجميُّونَ إلى توظيفها في صناعة المُعجم (٢). وكانت البدايةُ في عام ١٧٤٦ عندما صَنَعَ الأديبُ والمُعجميُّ الإنجليزيُّ صموئيل جونسون *Samuel Johnson* (١٧٠٩-١٧٨٤) مُدَوَّنةً لُغويَّة – بمُساعدة ستَّة من تلاميذِه – للإفادة منها في صناعة مُعجمه الموسوم بـ "مُعجم اللُّغة الإنجليزيَّة *A Dictionary of the English Language*" الَّذي نَشَرَهُ في عام ١٧٥٥ (٣). ومنذ هذا الوقت أصبَحَت المُدَوَّناتُ اللُّغَويَّةُ موردًا رئيسًا في الصِّناعة المُعجَميَّة.

وفي النِّصف الآخَر من القرن التَّاسع عشر ازدهَرَت صناعةُ المُعجمات التَّأثيليَّة *Etymological Dictionaries* الَّتي تُعنى بالبحث في أصول المُفرَدات، بعدَ عدَّة قرون من البحث في تاريخ اللُّغات. وكانَ من أبرز ما ظَهَرَ في هذه الحِقبة: المُعجمُ التَّأثيليّ للُّغة الإنجليزيَّة *A Dictionary of English Etymology* الَّذي صَنَعَهُ الإنجليزيّ هنسلي ويدجود *Hensleigh Wedgwood* (١٨٠٣-١٨٩١) ونشَرَهُ في عام ١٨٥٩ (٤)، ومُعجَم اللُّغات الرُّومانسيَّة اللاَّتينيَّة *Lateinisch-romanisches Wörterbuch* الَّذي صَنَعَهُ اللُّغَويّ الألمانيّ جوستاف كيرتنج *Gustav Körting* (١٨٤٥-١٩١٣) ونشَرَهُ في عام ١٨٩١ (٥). وأصبَحَ هذا النَّوعُ من المُعجمات أحدَ أهمِّ الموارد المُستخدَمة في صناعة المُعجمات التَّاريخيَّة؛ لاسيَّما في التَّجارب الفرنسيَّة الَّتي تعتمدُ عليها بصورةٍ كُلِّيَّة في أحيانٍ كثيرة.

وكانَ التَّأثيرُ الحقيقيُّ مع ظُهُور اللِّسانيَّات الحاسُوبيَّة في النِّصف الآخَر من القرن العشرين؛ حيثُ أمكَنَ – من خلالها – تطويرُ أدواتٍ لمُعالجة نُصُوص المُدَوَّنات اللُّغَويَّة الكبيرة نسبيًّا بجميع ما تحويه من المُفرَدات والمعاني، كما ساعَدَ الحاسُوبُ في تعيين البيانات الإحصائيَّة للنُّصُوص، واستخلاص المادَّة المُعجميَّة الَّتي يتضَمَّنُها المُعجمُ، ومُعالجة هذه المادَّة في مراحل الصِّناعة المُعجميَّة – جمعًا وتحريرًا ونشرًا؛ وساعَدَ الحاسُوبُ – كذلكَ – في توفير العديد من الموارد الأخرى، مثل: قواعد البيانات المُعجميَّة *Lexical Databases* وشبكات الكلمات *WordNets* وأنظمة التَّحرير المعجَميّ *Dictionary Writing Systems (DWS).*

١,٣. خلفيَّة تاريخيَّة حولَ مُحاولات إنجاز مُعجمٍ تاريخيّ للعربيَّة.

طَرَحَ المُستشرقُ الألماني أَوْجُست فيشر *August Fischer* (١٨٦٥-١٩٤٩) فكرةَ بناء مُعجمٍ عَصريّ للُّغة العربيَّة الفُصحى لعهدها القديم؛ وذلكَ في مؤتمر المُستشرقينَ الألمان الَّذي عُقِدَ في مدينة بازل بسويسرا في عام ١٩٠٧. وشرعَ فيشر في بناء مُعجمِه بعدَ تأسيس معاهد الأبحاث السَّكسونيَّة في مدينة لَيبْتسك *Leipzig* عامَ ١٩١٤؛ حيثُ توافَرَ له الدَّعمُ الَّذي كانَ يطمحُ إليه. ولم يستمرَّ العملُ طويلًا، إذ توقَّفَ نتيجةَ الانهيار الماليّ لألمانيا في أعقاب الحرب العالميَّة الأولى. وفي عام ١٩٣٢ أُنشِئَ مجمعُ اللُّغة العربيَّة بالقاهِرة. وجاء في المادَّة الثَّانِية من المرسوم الصَّادِر بشأن إنشاء المجمَع أنَّ من أغراضِه "أن يقومَ بِوَضْع مُعجَمٍ تاريخِيٍّ للُّغة العربيَّة، وأن ينشُرَ أبحاثًا دقِيقةً في تاريخ بعضِ الكلمات وتَغَيُّر مدلولاتِها" (٦). ولأنَّ الموارِدَ والأدوات المُستخدَمة في إنجاز هذا النَّوع من المُعجمات لم تكن مُهيَّأةً في ذلكَ الوقت، فقد انصَرَفَ المجمعُ المصريُّ عن صناعة المُعجم التَّاريخيّ للعربيَّة حتَّى حين، ثُمَّ رأى المجمعُ أن ينتفعَ بما أنجَزَهُ فيشر في مشروعه، نزولًا على اقتراح المُستَشرِق الإيطالي كارلو نلينو *Carlos Nallino* (١٨٧٢-١٩٣٨). وبالفعل، تبنَّى مجمعُ اللُّغة العربيَّة مشروعَ فيشر لبناء مُعجمه الموسوم بـ "المُعجم التَّاريخيّ للُّغة الآداب العربيَّة" في عام ١٩٣٦ بدعمٍ من الحُكومة المصريَّة. ولم تستمرَّ مُحاولة فيشر أكثرَ من ثلاث سنوات؛ حيثُ غادَرَ مصرَ إلى ألمانيا في عام ١٩٣٩ في أعقاب الحرب العالميَّة الثَّانية، ثُمَّ أقعَدَه المرضُ فلم يتَمَكَّن مِن العَودة واستئناف العمل حتَّى وفاتِه في عام ١٩٤٩ (٧). وتناثَرَت أوراقُ فيشر بينَ مصرَ وألمانيا، فلم يَجِد المجمَعُ المصريُّ ما يَصلُحُ للنَّشر مِنها إلاَّ مُقَدِّمةً – أعَدَّها فيشر – ونموذجًا يبدأ مِن حرف الهمزة إلى مادَّة "أبد"؛ فَنَشَرَهُما المَجمَعُ عامَ ١٩٥٠ بِعُنوان (مُعجَم فيشر "مُقَدِّمة ونموذَج مِنه")، ثُمَّ أعادَ نَشرَهُما في ١٩٦٧ بِعُنوان (المُعجَم اللُّغَوِيّ التَّارِيخِيّ).

وأعقَبَت محاولةَ فيشر عدَّةُ مُحاولاتٍ أخرى، فكانت مُحاولةُ الجمعيَّة الألمانيَّة للاستشراق في عام ١٩٥٧ لصناعة "مُعجم اللُّغة العربيَّة الفُصحى *Wörterbuch der Klassischen Arabischen Sprache*" بقيادة المُستشرق الألمانيّ يورج كريمر *Jörg Kraemer* (١٩١٧-١٩٦١). وتحوَّلَ مسارُ هذا المُعجم ليُصبحَ مرحليًّا لثلاثة القرون الأولى مِن الهِجرة وقَرَنَينِ قبلها، ولا يزالُ العملُ فيه قائمًا بقيادة مانفرِد أولمان *Manfred Ullmann* بعدَما تعاقَبَت عليه عدَّةُ أجيال. وبُعِثَت فكرةُ بناء مُعجمٍ تاريخيٍّ للعربيَّة مرَّةً ثالثة في عام ١٩٩٠ عندما أعلَنَت جمعِيَّة المُعجَمِيَّة العَرَبِيَّة بتُونُس عن مشروع جديد للمُعجم. ولم يتجاوز المشروعُ الفكرةَ حتَّى عام ١٩٩٦؛ إذ توافرَ للجمعيَّة دعمٌ مادِّيٌّ من وَزارةِ التَّعلِيم العالِي والبَحث العِلمِيّ بالجُمهورِيَّة التُّونُسِيَّة؛ وتوقَّفَ المشروعُ عندَ بناء مُدَوَّنةٍ لُغَوِيَّةٍ مُؤَرِّخةٍ للعصر الجاهِلِيّ، تَضُمُّ نُصُوصًا تنتمي إلى الحقبة من عام ٢٠٠م إلى عام ٦٠٤م انطلاقًا مِن أقدم ما اطَّلَعَ علَيه فَرِيقُ البَحثِ مِن نُصُوصٍ مُوَثَّقة؛ ولم يُنشَر من أعمال هذا المشروع إلَّا مجموعةٌ من المقالات والأوراق البحثيَّة لبعض أعضاء الجمعيَّة. وفي عام ٢٠٠٤ أعلَنَ اتِّحادُ المجامع العَرَبِيَّة عن تأسيس "هيئة المُعجم التَّاريخيّ للُّغة العربيَّة" بالقاهرة؛ وتكوَّنَ لها مجلسٌ علميٌّ يرأسُهُ الأمينُ العامُّ لاتِّحاد المجامع العربيَّة – حينئذٍ – اللُّغَويّ المصريّ كمال بشر، ويضُمُّ المجلسُ عددًا من الخُبراء في عُلُوم اللُّغة والمُعجم. ولم تتجاوز أعمالُ هذه الهيئة مجموعةً من الاجتماعات والمُؤتمرات حتَّى ديسمبر من عام ٢٠٠٦؛ حيثُ تلقَّت الهيئةُ دعمًا من حكومة الشَّارقة لإنجاز المشروع. وكانَ آخرُ ما أُنجِزَ من الأعمال العلميَّة لهذه الهيئة تقريرٌ من رئيس المجلس العلميّ في أبريل من عام ٢٠٠٧، استعرَضَ فيه أعمالَ المجلس وما أعَدَّتهُ لجانُهُ من الخُطط المُستقبليَّة (٨). وأخيرًا، أعلَنَ المركزُ العربيُّ للأبحاث ودراسة السِّياسات في الدَّوحة عامَ ٢٠١٣ عن مشروع "مُعجم الدَّوحة التَّاريخيّ للُّغة العربيَّة" بعدَ ثلاثة اجتماعاتٍ لأساتذةٍ وخُبراء مُتخصِّصين؛ وتكوَّنَ لهذا المشروع مجلسٌ علميٌّ برئاسة اللُّغَويّ اللُّبنانيّ رمزي بعلبكِّي، وهيئةٌ تنفيذيَّة برئاسة اللُّغَويّ المغربيّ عزّ الدِّين البوشيخي، ثُمَّ أُعلِنَ عن استقلال المشروع عن المركز العربيّ للأبحاث في ٢٠١٤ ليكونَ مؤسَّسةً مُستقلَّةً تتبعُ إداريًّا معهدَ الدَّوحة للدِّراسات العُليا. وقد انتهَت الهيئةُ التَّنفيذيَّةُ للمشروع من إعداد مُدوَّنةٍ لُغويَّةٍ تمتدُّ من بداية التَّأريخ للعربيَّة حتَّى نهاية القرن الثَّاني الهجريّ، كما انتهَت من إعداد ببليوغرافيا مرجعيَّة بمصادر مادَّة المُدَوَّنة اللُّغَويَّة (٩).

٢. إشكالات إعداد موارد المُعجم التَّاريخيّ للعربيَّة.

ثمَّةَ بعضُ الإشكالات الَّتي تُواجهُ إعدادَ موارد المُعجم التَّاريخيّ للعربيَّة، لعلَّ أبرزَها:

١. نُدرة موارد التَّأثيل المُعجميّ للعربيَّة.

من العَقَبات الأساسيَّة الَّتي تُواجهُ صناعةَ مُعجمٍ تاريخيٍّ للعربيَّة أنَّها – إلى الآن – لا تملكُ مُعجمًا تأثيليًّا. وبالنَّظر إلى التَّجارب النَّاجحة في اللُّغات الأخرى، نجدُها جميعًا قد اعتمَدَت على مُعجماتٍ تأثيليَّة أفِيدَ منها في إثراء المادَّة المُعجميَّة. لقد عَرَفَت الفرنسيَّةُ – مثلًا – طريقَها إلى المُعجمات التَّأثيليَّة في القرن السَّابع عشر، واعتمَدَت عليها موردًا رئيسًا في أوَّل مُحاولةٍ تعرفُها البشريَّةُ لبناء مُعجمٍ تاريخيٍّ عامَ ١٧٧٨. وأفادَت الفرنسيَّةُ أيضًا من المُعجمات التَّأثيليَّة في تجربتها النَّاجحة عامَ ١٩٩٢ لبناء "المُعجَم التَّارِيخِيّ لِلُّغَة الفَرَنسِيَّة *Dictionnaire Historique De La Langue Française*". ولا يختلفُ الحالُ كثيرًا في تجارب اللُّغات الألمانيَّة (بينَ عامَي ١٨٣٨ و ١٩٦١) والهولنديَّة (بينَ عامَي ١٨٤٩ و ١٩٩٨) والإنجليزيَّة (بينَ عامَي ١٨٥٩ و ١٩٢٨) والسِّويديَّة (مُنذ عام ١٨٨٤ إلى عام ٢٠١٧، حسبما أعلَنَت الأَكاديميَّة السِّويديَّة). فقد اعتمَدَت هذه اللُّغاتُ جميعًا على مُعجماتٍ تأثيليَّةٍ مُنجَزةٍ سلفًا. أمَّا اللُّغةُ العربيَّة، فلم تعرف إلَّا مُحاولةً حديثةَ العهد، أعلَنَت عنها كُلِّيَّةُ العُلُوم الإنسانيَّة في جامعة أوسلو بالنِّرويج عامَ ٢٠١٣ بإشراف الباحث ستيفان جوث *Stephan Guth* لبناء "المُعجم التَّأثيليّ للُّغة والثَّقافة العربيَّة *Etymological Dictionary of Arabic Language and Culture*". ومع ما يبدو في هذه المُحاولة من جُهُود في الإعداد المنهجيّ وما يتوافرُ لها من دعم المؤسَّسات العلميَّة، إلَّا أنَّها لاتزالُ في مراحلها الأولى؛ حيثُ لم يبدأ العملُ فيها – فعليًّا – إلَّا قريبًا من مُنتصف ٢٠١٤ (١٠).

٢. حداثة عهد العربيَّة بالموارد اللُّغَويَّة.

يغلُبُ على صناعة المُعجم العربيّ الحديث أن تقتصِرَ في مواردها على المُعجمات اللُّغَويَّة المُنجَزة سلفًا، دونَ الإفادة من الموارد اللُّغَويَّة الَّتي تعكسُ واقعَ اللُّغة وتُثري المادَّةَ المُعجميَّة. ومع ما نراهُ من عنايةٍ بالموارد اللُّغَويَّة في صناعة مُعجمات اللُّغات الأخرى، تظلُّ الصِّناعةُ المُعجميَّةُ العربيَّةُ قاصرةً عن توظيف هذه الموارد؛ بل نراها قاصِرةً عن مُجرَّد بنائها وتطويرها. نُلاحظُ مثلًا أنَّ المُدَوَّنات اللُّغَويَّة لم تُستخدَم في صناعة المُعجم العربيّ إلَّا في مشروعاتٍ محدودة لمُعجماتٍ ثُنائيَّة اللُّغة، على نحو ما نجدُ في "مُعجم نايمَيخِن العربيّ – الهولنديّ *Nijmegen Arabic/Dutch Dictionary*" الَّذي أشرَفَ عليه المعجميّ الهولندي يان هوخلاند *Jan Hoogland* بينَ عامَي ١٩٩٧ و ٢٠٠٣ (١١)، أو من خلال بعض النَّماذج المُعجميَّة المُتضَمَّنة في أطروحاتٍ علميَّةٍ أو دراساتٍ حولَ بناء المُدَوَّنات اللُّغَويَّة واستخدامها لأغراض الصِّناعة المُعجميَّة. ولعلَّ المُدَوَّناتِ اللُّغَويَّةَ أوفرُ حظًّا من مواردَ أخرى، مثل: شبكة الكلمات أو الأنطولوجيا المعرفيَّة *Ontology* الَّتي لم يكتمل لها في اللُّغة العربيَّة مشروعٌ واحدٌ حتَّى الآن. من ذلكَ مثلًا مشروعُ "شبكة الكلمات العربيَّة *Arabic WordNet (AWN)*" الَّذي سعى إلى إنجازه فريقٌ بحثيٌّ مُشترَكٌ بينَ جامعة برينستون في الولايات المُتَّحدة الأمريكيَّة وجامعة برشلونة في أسبانيا وجامعة مانشستر في المملكة المُتَّحدة بإشراف الباحثة الألمانيَّة كريستيان فيلباوم *Christiane Fellbaum* بينَ عامَي ٢٠٠٥ و ٢٠٠٧ (١٢)؛ حيثُ توقَّفَ المشروعُ مُبكِّرًا نتيجةَ ضعف التَّمويل؛ ولم تتجاوز مُخرَجاتُهُ نموذجًا مُترجمًا لبعض أجزاء شبكة الكلمات الإنجليزيَّة الموسُومة بـ "شبكة كلمات برينستون *Princeton WordNet*" (١٣). ومن ناحيةٍ أخرى، لا يزالُ العملُ جاريًا في مشروعاتٍ عربيَّةٍ حديثةٍ، مثل: مشروع "أنطولوجيا اللُّغة العربيَّة *Arabic Ontology*" الَّذي ترعاهُ جامعة بيرزيت في فلسطين منذ عام ٢٠١١، ويُشرفُ عليه الأَكاديميّ الفلسطينيّ مُصطفى جرَّار؛ وكذلكَ مشروع "الشَّبكة الدِّلاليَّة للُّغة العربيَّة *SEmantic Wordnet of ArabiC (SEWAC)*" الَّذي ترعاهُ المُنظَّمة العربيَّة للتَّربية والثَّقافة والعُلُوم *ALECSO* منذُ عام ٢٠١٣، ويُشرفُ عليه الباحثُ المغربيّ كريم بُوزُوبع.

٣. حداثة عهد العربيَّة باللِّسانيَّات الحاسُوبيَّة.

ظَهَرَت اللِّسانيَّاتُ الحاسُوبيَّةُ *Computational Linguistics* في الخمسينيَّات من القرن العشرين عبرَ ميادين التَّرجمة بينَ اللُّغات. وأمكنَ الإفادةُ من الأبحاث المُنجزة في هذا الوقت في تطويع الآلة لفهم اللُّغات الطَّبيعيَّة ومُحاكاة الذَّكاء الإنسانيّ. وفي مُدَّة قصيرة انتَقَلَت تقنيات اللِّسانيَّات الحاسُوبيَّة من الولايات المُتَّحدة الأمريكيَّة – مهدها الأوَّل – إلى الأمريكتَين وأوروبا، ثُمَّ إلى مناطقَ أخرى. ومع التَّطوُّر السَّريع لهذه التِّقنيات، ظَهَرَ ميدانٌ جديدٌ يُعنى بتوظيف أدوات مُعالجة اللُّغات الطَّبيعيَّة *Natural Language Processing* في صناعة المُعجم؛ واستقلَّ هذا الميدانُ لاحقًا فيما يُعرَفُ بـ "الصِّناعة المُعجميَّة الحاسُوبيَّة *Computational Lexicography*"، فصارَ بالإمكان أن يُستفادَ من تقنيات الحاسُوب في الصِّناعة المُعجميَّة – جمعًا وتحريرًا ونشرًا – من خلال مواردَ وأدواتٍ حاسُوبيَّة تُساعدُ في مُعالجة المادَّة المُعجميَّة وتعيين مُخرجاتها، الأمرُ الَّذي ساهمَ في توفير الوقت والجهد. ويكفي أن نعرفَ أنَّ سلاسلَ مُعجمات: لونجمان *Longman* الإنجليزيَّة، وكولينز كوبيلد *Collins COBUILD* الأنجلوأمريكيَّة، وروبير *Dictionnaires Le Robert* الفرنسيَّة، وفيدمانشي *Weidmannsche* الألمانيَّة تعتمدُ جميعًا على تقنيات الحاسُوب – في جميع مراحل الصِّناعة – مُنذُ مطلع الألفيَّة الثَّالثة. وعلى صعيد المُعجمات التَّاريخيَّة، فقد أُنجِزَ أكثرُها قبلَ أن تتَّضِحَ أساليبُ الصِّناعة المُعجميَّة الحاسُوبيَّة؛ إلَّا أنَّ هذا لم يمنع من استثمار أدوات المُعالجة الآليَّة في رَقمنة هذه المُعجمات أو تطويرها. فقد قامَ المَركَز الألمانيّ للتَّنميَة الإلكترونيَّة ونَشر العُلوم الإنسانيَّة بجامِعة تريير *University of Trier* برقمَنة المُعجم الألمانيّ بينَ عامَي ١٩٩٨ و ٢٠٠٤، وقامَ المعهد الهولنديّ للمُعجميَّة *Institute for Dutch Lexicology* برقمَنة المُعجم الهولنديّ بينَ عامَي ١٩٩٩ و ٢٠٠٧. ومنذُ عام ١٩٩٣ تتولَّى مطبعة جامعة أكسفورد *Oxford University Press* إعادةَ إنتاج المُعجم الإنجليزيّ في نُسخته الثَّالثة؛ ويُستخدَمُ الحاسُوبُ في جميع مراحل الصِّناعة المُعجميَّة لهذه النُّسخة الَّتي من المُزمَع أن تكتملَ عامَ ٢٠٣٧، وفقَ تقدير فريق العمل الَّذي يقودُهُ المُعجميّ الإنجليزيّ مايكل بروفيت *Michael Proffitt* منذ عام ٢٠١٣.

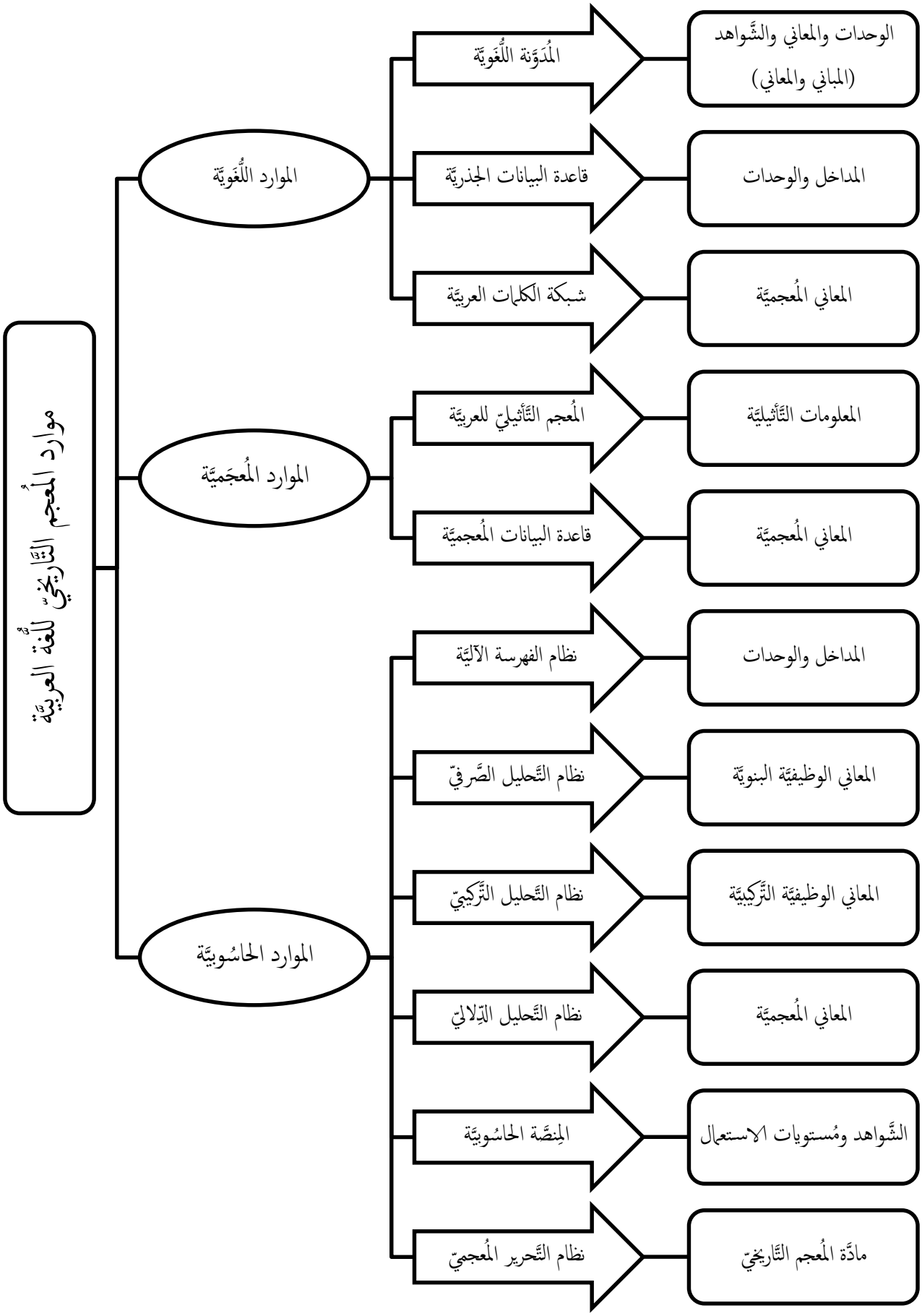
أمَّا اللُّغةُ العربيَّةُ فقد تأخَّرَت عن مُلاحقة هذا التَّطوُّر المعلوماتيّ بنحو أربعينَ سنة؛ إذ لم تعرف الطَّريقَ إلى اللِّسانيَّات الحاسُوبيَّة إلَّا مع مطلع التِّسعينيَّات من القرن العشرين، باستثناء بعض التَّجارب الأمريكيَّة والأوروبيَّة في ميدان التَّرجمة الآليَّة *Machine Translation*، الأمرُ الَّذي انعكسَ – سلبًا – على أدوات مُعالجة اللُّغة العربيَّة، سواءٌ على مُستوى المباني أم المعاني؛ وانعكَسَ – كذلكَ – على حوسبة المُعجم العربيّ. وإذا تجاوَزنا هذا إلى الحديث عن الواقع، فسنجدُ أنَّ العربيَّةَ لا تُعاني مُشكلةً في حوسبة مُعجمها في مرحلة النَّشر المُعجميّ الَّتي تُمثِّلُ المرحلةَ الثَّالثةَ من مراحل الصِّناعة. وهي في ذلكَ مثلُ كُلِّ اللُّغات الإنسانيَّة الَّتي تخضعُ في نشر مُعجماتها لضوابط ومعاييرَ واضحة ومُوَحَّدة. لكنَّ المُشكلة الحقيقيَّة ستبدو في مرحَلَتي الجمع والتَّحرير اللَّتَين تُمثِّلان مرحلَتَي الصِّناعة الأولى والثَّانية. ففي المرحلة الأولى، لاتزالُ أدواتُ المُعالجة قاصرةً عن التَّعامُل مع النُّصُوص العربيَّة بصورةٍ كاملة. وبالنَّظر إلى الأدوات الأساسيَّة الَّتي يُمكنُ – من خلالها – مُعالجةُ النُّصُوص، نجدُ قُصُورًا في المُفهرسات الآليَّة نتيجةً لطبيعة اللُّغة العربيَّة الاشتقاقيَّة ونظامها الكتابيّ، ونجدُ قصورًا في أدوات التَّحليل الصَّرفيّ نتيجةَ الالتباس الَّذي يُحدثُهُ خُلُوُّ الكلمات من علامات الضَّبط، ونجدُ قُصُورًا في أدوات التَّحليل التَّركيبيّ نتيجةَ الالتباس الَّذي يقعُ بينَ أقسام الجُملة العربيَّة، ونجدُ قُصُورًا في أدوات التَّحليل الدِّلاليّ نتيجةً لتعدُّد معاني الكلمة الواحدة. وفي المرحلة الثَّانية من مراحل الصِّناعة، نجدُ أكثرَ أنظمة التَّحرير المُعجميّ تتعاملُ مع العربيَّة باعتبارها لُغةً التصاقيَّة، الأمرُ الَّذي يتطلَّبُ التَّدَخُّلَ اليدويَّ بصورةٍ مُستمرَّةٍ لاستدراك أخطاء الآلة (١٤). ومع هذا، فقد نجحَت بعضُ التَّجارب الحديثة في الإفادة من تقنيات الحاسُوب وتوظيفها في الصِّناعة المُعجميَّة – جمعًا وتحريرًا ونشرًا – على نحو ما نجدُ مثلًا في "المُعجم التَّكراريّ للُّغة العربيَّة *A Frequency Dictionary of Arabic*" الَّذي صَنَعَهُ الباحثان الأمريكيَّان تيم بكوالتر *Tim Buckwalter* وديلوورث باركنسون *Dilworth Parkinson* ونشراهُ عام ٢٠١١ (١٥).

٣. منهجيَّة إعداد موارد المُعجم التَّاريخيّ للُّغة العربيَّة.

إنَّ إنجازَ المعجَم التَّاريخيّ للُّغة العربيَّةِ ليسَ عملاً إبداعيًّا يعكُفُ عليه المُؤلِّفونَ والمُفكِّرون؛ وإنَّما هو صناعةٌ، ينبغي أن تُوكَلَ إلى أهلِها العارفِينَ بدقائِقِها وضوابطِها. والواقعُ أنَّ العَقَبَةَ الرَّئيسةَ أمامَ إنجاز المعجم المنشُود – من وُجهةِ نظر الباحث – تتمثَّلُ في أنَّنا لا نأخُذُ بعينِ الاعتبار ضرورةَ اعتماد هذا المعجم على مواردَ رئيسَـةٍ تتكوَّنُ عنها مادَّتُه على نحوٍ يتوافقُ مع مناهجِ الصِّناعة المُعجميَّة الحديثة. لقد توقَّفَت المحاولاتُ السَّابقةُ لصناعة مُعجمٍ تاريخيٍّ للعربيَّة من حيثُ بدأت. فقد انتهى بعضُها عندَ مرحلة إيجاد مادَّةٍ نَصِّيَّةٍ تستوعبُ اللُّغةَ العربيَّةَ عبرَ عُصُورها، أو عندَ مرحلة التَّأثيل لها؛ وتوقَّفَ البعضُ الآخرُ عندَ مرحلة التَّخطيط لصناعة المعجم. وكانت النَّتيجةُ أنَّنا الآنَ – بعدَ ما يربو على مئة عامٍ من المحاولة الأولى – لم نخطُ خُطواتٍ تُذكَر في هذا الشَّأن. وبالنَّظَر إلى التَّجارب النَّاجحة في المعجمات التَّاريخيَّة للُّغات الجرمانيَّة – على سبيل المثال – نجدُ اعتمادَها بصورةٍ رئيسةٍ على مُدَوَّناتٍ لُغَويَّةٍ – وإن اختلَفَت مناهجُ صناعتها – مَثَّلَت قاعدةً لمادَّة هذه المعجمات؛ في حين اعتمَدَ المعجم التَّاريخيّ للفرنسيَّة على المُعجمات التَّأثيليَّة الَّتي اهتمَّ بها المعجميُّونَ الفرنسيُّونَ في وقتٍ مُبكِّر. وعلى النَّهج ذاته تسيرُ المشروعاتُ الَّتي لا تزالُ قائمةً لإنجاز مُعجماتٍ لُغَويَّةٍ تاريخيَّة، على نحو ما نجدُ – مثلًا – في مشروع "المُعجم التَّاريخيّ الحديث للُّغة الأسبانيَّة *Nuevo diccionario histórico del español*" الَّذي ترعاهُ الأكاديميَّة الملكيَّةُ الأسبانيَّة في مدريد (١٦)؛ حيثُ يعتمدُ أساسًا على مجموعةٍ من المُدَوَّنات اللُّغَويَّة الكبيرة نسبيًّا وقواعد البيانات اللُّغَويَّة والمُعجميَّة؛ وكذلكَ الحالُ في مشروع "المُعجم التَّاريخيّ للُّغة العبريَّة *The Historical Dictionary Project of the Hebrew Language*" الَّذي يرعاهُ مجمعُ اللُّغة العبريَّة في القُدس (١٧)؛ حيثُ يعتمدُ على مجموعةٍ كبيرةٍ من النُّقُوش والمُدَوَّنات اللُّغَويَّة وقواعد البيانات المُعجميَّة. أضف إلى ذلكَ أنَّ جميعَ هذه المعجمات وغيرَها تعتمدُ على الموارد الحاسُوبيَّة في إصداراتِها المرقمنة.

وعلى مُستوى اللُّغة العربيَّة، فقد بُذِلَت جُهُودٌ محمودةٌ في صناعة المُعجم العربيّ – قديمًا وحديثًا، كما ساعدَت المناهجُ اللُّغَويَّةُ الحديثةُ وتقنياتُ الحاسُوب في بناء العديد من الموارد الَّتي أُنجِزَت في صُورة مُدَوَّناتٍ أو مُعجماتٍ لُغَويَّة أو مُصَنَّفاتٍ أو أدواتٍ حاسُوبيَّة. وسيكونُ من المُفيد أن يُستفادَ ممَّا أنجِزَ من جُملة ما تحويه المكتبةُ العربيَّةُ والمُجتَمَعُ المَعرفيُّ من مواردَ لصناعة المُعجم التَّاريخيّ المنشود، توفيرًا للوقت والجُهد من ناحيةٍ، وإثراءً لمادَّة المُعجم من ناحيةٍ ثانية، ومُحاكاةً للتَّجارب النَّاجحة في بناء المُعجمات التَّاريخيَّة للُّغات الأخرى من ناحيةٍ ثالثة. وتحقيقًا للفائدة المنشودة، تسعى الدِّراسةُ – فيما يلي – إلى الوُقُوف على الموارد الَّتي يُمكنُ الإفادةُ منها في بناء المُعجم التَّاريخيّ للعربيَّة، بالنَّصِّ على المُتاح منها – حالَ وُجُوده، واقتراح البدائل الَّتي قد تكونُ مُناسبةً ما لم تتوافر الموارِدُ الَّتي ننشُدها، بعدَ استقراء هذه البدائل والمُفاضَلَة بينها.

وانطلاقًا من تجارب اللُّغات الأخرى والمُحاولات السَّابقة لصناعة مُعجمٍ تاريخيٍّ للُّغة العربيَّة، تقومُ منهجيَّةُ هذه الدِّراسة على تقسيم موارد المُعجم المنشود إلى ثلاثة أقسام، هي: الموارد اللُّغويَّة *Linguistic Resources*؛ وتشتملُ على: المُدَوَّنة اللُّغَويَّة، وقاعدة البيانات الجذريَّة، وشبكة الكلمات العربيَّة؛ والموارد المعجميَّة *Lexical Resources*؛ وتشتملُ على: المُعجم التَّأثيليّ للعربيَّة، وقاعدة البيانات المُعجميَّة؛ والموارد الحاسوبيَّة *Computational Resources*؛ ونعني بها أدوات المُعالجة والتَّحرير؛ وتشتملُ على: نظام الفهرسة الآليَّة، ونظام التَّحليل الصَّرفيّ، ونظام التَّحليل التَّركيبيّ، ونظام التَّحليل الدِّلاليّ، والمِنَصَّة الحاسُوبيَّة، ونظام التَّحرير المُعجميّ. وتُشكِّلُ هذه الموارِدُ – مُجتمِعةً – منظومةً مُتكاملةً، يُمكنُ من خلالها استخلاصُ المادَّة المُعجميَّة، على النَّحو المُوَضَّح في (الشَّكل ١)؛ حيثُ تُستخلَصُ المداخلُ والوحداتُ من المُدَوَّنة اللُّغَويَّة وقاعدة البيانات الجذريَّة باستخدام نظام الفهرسة الآليَّة، وتُستخلَصُ المعاني المُعجميَّةُ من المُدَوَّنة وشبكة الكلمات وقاعدة البيانات المُعجميَّة بمُساعدة نظام التَّحليل الدِّلاليّ، وتُستخلَصُ المعلوماتُ التَّأثيليَّةُ من المُعجم التَّأثيليّ، وتُستخلَصُ الشَّواهدُ من المُدَوَّنة اللُّغَويَّة عبرَ المِنَصَّة الحاسُوبيَّة، وتُستخلَصُ المعاني الوظيفيَّة من المُدَوَّنة بمُساعَدَة نظامَي التَّحليل الصَّرفيّ والتَّركيبيّ، وتُحَدَّدُ مُستوياتُ الاستعمال عن طريق المنصَّة الحاسُوبيَّة؛ وتَصُبُّ جميعُ المُخرَجات في نظام التَّحرير المُعجميّ.



الشَّكل ١: موارد المُعجم التَّاريخيّ للُّغة العربيَّة ومُخرجاتها المُعجميَّة

٣,١. الموارد اللُّغويّة.

٣,١,١. المُدَوَّنة اللُّغويّة.

المُدَوَّنةُ اللُّغَويَّةُ *Linguistic Corpus* هي المورِدُ الرَّئيسُ للمُعجم التَّاريخيّ. فهي مصدرُ المداخل والوحدات المُعجميَّة، وهي مصدرُ المعاني المُعجميَّة لهذه الوحدات، وهي مصدرُ الشَّواهد المُعجميَّة للوحدات ومعانيها. ومن خلال المُدَوَّنة نستطيعُ أن نتعقَّبَ تاريخَ المُفرَدات ومراحلَ تطوُّرها على مُستوى المباني والمعاني؛ ونستطيعُ – أيضًا – أن نقفَ على المُستعمَل والمُهمَل والمُمات من المُفرَدات والألفاظ. ولمَّا كانَ المأمولُ مِن المُعجَم التَّاريخيّ المنشود أن يكونَ دِيوانًا للعَرَبيَّة، يُؤَرِّخُ لألفاظِها ومعانيها، ويُبيِّن ما طرأ على تلك الألفاظ من تَحوُّلٍ وتغيُّرٍ، كان لابُدَّ لمُدَوَّنة هذا المُعجَم أن تَضُمَّ قدرًا وافِيًا مِن النُّصُوص الَّتي تعكِسُ واقِعَ اللُّغة العَرَبيَّة عبرَ تاريخها المديد، وفي بيئاتِها ومراكِزِها الثَّقافِيَّةِ والعِلميَّة والحضارِيَّة الَّتي شَهِدَت مراحِلَ نُمُوِّها وتَطَوُّرِ ألفاظِها وتراكِيبِها، لتكونَ بمثابة الذَّخيرة الجامِعة لآثارِها الأدَبِيَّة وتُراثِها العِلميّ المُدَوَّن (١٨).

ونظرًا لحداثة مُحاولات بناء مُدَوَّناتٍ لُغَويَّةٍ للمُعجم التَّاريخيّ للعربيَّة وعدم اكتمال أكثرها، تقترحُ الدِّراسةُ أن يُستفادَ من (مُدَوَّنة المُعجم التَّاريخيّ للُّغة العربيَّة) الَّتي أنجَزَها الباحثُ المُعتزّ بالله السَّعيد، ضِمنَ أطروحته للدُّكتوراه في جامعة القاهرة، والمَوسُومة بـ "مُدَوَّنة مُعجمٍ تاريخيٍّ للُّغة العربيَّة، مُعالَجة لُغَويَّة حاسُوبيَّة". وقد اشتمَلَت هذه المُدَوَّنةُ على نُصُوصٍ من اللُّغة العربيَّة المكتوبة منذُ أقدم ما عُثِرَ عليه من الوثائق في القرن الخامس قبلَ الهجرة إلى وقت إنجازها؛ وجُمعت نصوصُها من التُّراث العربيّ المُدَوَّن، في صورة نُصُوصٍ مُنتقاةٍ من المصنَّفات العربيَّة في مختلف المعارف بحيثُ تغطي الحِقَبَ الزَّمنيَّة للعربيَّة، وبحيثُ تغطِّي الحُدُودَ الجُغرافيَّةَ الَّتي مثَّلت المراكزَ الثَّقافيَّةَ والحضاريَّةَ للعربيَّة. وقُسِّمَت نُصُوصُ المدوَّنة زمنيًّا إلى مرحلتين: تمتدُّ الأولى مُنذ أقدم وثائقها عامَ (٤٠٣ ق.هـ) حتَّى نهاية القرن الثَّاني الهجريّ، ويغلُبُ على نُصُوصها الانتقالُ بالرِّواية، واشتملت المدوَّنةُ على جميع ما أتيحَ لصانعها من نُصُوص هذه المرحلة، وتمتدُّ المرحلةُ الأخرى منذ بداية القرن الثالث الهجري إلى أحدث وثائق المُدوَّنة في القرن الخامس عشر الهجري (وَقت صناعتها)، ويغلب على نصوصها الانتقالُ بالتدوين، واقتصرت المُدَوَّنةُ في هذه المرحلة على أمَّهات الكُتُب العربيَّة في مختلف العُلُوم والفُنُون والآداب بما لا يقلُّ عن مليون كلمة لكلِّ قرن زمنيّ (وهو الحدُّ المعياريُّ الأدنى لاعتبار نُصُوص القرن الزَّمَنيّ مُدَوَّنةً لُغويَّةً مُستقلَّة)، مع مُراعاة التَّنوُّع الجُغرافيّ والموضوعيّ للنُّصوص. أمَّا عن تصنيف مادَّة المُدَوَّنة؛ فقد صُنِّفَت تاريخيًّا إلى خمسة حُقُول، هي العُصُورُ الأدبيَّة للعربيَّة (الجاهليّ، والإسلاميّ، والعبَّاسيّ، والوَسيط، والحديث)، وجُغرافيًّا إلى سبعة حُقُول، هي: (شِبه الجزيرة العربيَّة، والشَّام، وبلاد الرَّافدَين، ووادي النِّيل، والأندلُس، ومنطقة المغرب العربيّ وصقلِّية، ومنطقة فارس وآسيا الوُسطى)، وموضوعيًا إلى خمسةَ عشرَ حقلًا، هي: (الشِّعر، والقرآن الكريم، والحديث النَّبويّ، والكتاب المُقدَّس، والنَّثر الأدبيّ، وعُلُوم العربيَّة وآدابها، والنَّثر العلميّ واصطلاحات العُلُوم، والتَّاريخ والأنساب، والطَّبَقات والتَّراجم، والجُغرافيا والرِّحلات والبُلدان، والقوانين والأحكام، والمِلل والعقائد، والموسُوعات والمُعجمات، والصِّحافة العربيَّة، والمعارف الأخرى) (١٩).

وتجدُرُ الإشارةُ إلى أنَّ هذه المُدَوَّنة قد أنجِزَت في إصدارَين؛ حيثُ امتدَّ العملُ في الإصدار الأوَّل من عام ٢٠٠٨ إلى عام ٢٠١١؛ واشتَمَلَت المُدَوَّنةُ في هذا الإصدار على أكثر من مئة مليون كلمة، تَضُمُّها ثمانمئة وتسعةٌ وستُّونَ وثيقة؛ واستُخدِمَ هذا الإصدارُ فعليًّا في بناء نموذجٍ كاملٍ للمُعجم التَّاريخيّ المنشود للعربيَّة، يحوي سبعةً وخمسينَ وحدةً مُنسدلةً عن المداخلَ المُعجميَّةَ الخمسة (ء د م، ج م هـ ر، س ق ف، غ س ق، و ث ب). أمَّا الإصدارُ الآخَرُ فقد امتدَّ العملُ فيه من عام ٢٠١١ إلى عام ٢٠١٤؛ وسعى الباحثُ فيه إلى تنمية مادَّة المُدَوَّنة في الإصدار الأوَّل ومُعالَجة جوانب القُصُور فيها. واشتَمَلَت المُدَوَّنةُ في هذا الإصدار على أكثرَ من مليار كلمة (عشرة أمثال حجمها في الإصدار الأوَّل)، تحويها أكثر من عشرة آلاف وثيقة. ولا يزالُ العملُ قائمًا في مُراجعة وثائق المُدَوَّنة وتدقيق مادَّتها وتطويرها وإعداد قاعدة بياناتها الببليوغرافيَّة، سعيًا إلى إخراجها على النَّحو الَّذي يتحقَّقُ معه الهدفُ من المُعجم التَّاريخيّ للُّغة العربيَّة.

٣,١,٢. قاعدة البيانات الجذريَّة.

قاعدة البيانات الجذريَّة *Roots Database* هي المورِدُ الوسيطُ بينَ المُدَوَّنة اللُّغَويَّة للمُعجم والمادَّة المُعجَميَّة المُستخلَصَة؛ حيثُ تتمُّ تعبئتُها بكلمات المُدَوَّنة *Corpus Words* ومعلومات التَّكرار *Frequency Information* وجُذُور الكلمات *Roots* والوحدات المُعجميَّة *Lexemes*. وهي – بذلكَ – موردٌ فرعيٌّ، ترتبطُ مادَّتُه بما تحويه مُدَوَّنةُ المُعجَم من النُّصُوص. وبعبارةٍ أخرى، هي هيكلٌ مُعجَميٌّ مُفرَّغٌ للمُدَوَّنة اللُّغَويَّة بعدَ مُراجعة مادَّتها واستخلاص مبانيها من المداخل والوحدات المُعجَميَّة *Entries & Headwords (Lexemes)* الَّتي سيحويها المُعجَم.

وتُساعدُ قاعدةُ البيانات الجذريَّة في توفير جُهُود المُحرِّرينَ والمُعالجينَ – إلى درجةٍ كبيرةٍ – حيثُ تَمُدُّهُم بقوائم الكلمات والجُذُور والوحدات؛ ويكونُ تركيزُهُم في استخلاص المعاني والشَّواهد والمعلومات المُعجَميَّة الأخرى. وتُساعدُ قاعدةُ البيانات الجذريَّة – أيضًا – في استكشاف أخطاء اللُّغة والإملاء الَّتي يُمكنُ أن تقعَ في المُدَوَّنة اللُّغَويَّة للمُعجم؛ الأمرُ الَّذي يُمكنُ معه استدراكُ هذه الأخطاء قبلَ البدء في مرحلة التَّحرير المُعجميّ. وسعيًا إلى إثراء هذه القاعدة وتحقيقِها للهدف المنشُود، تقترحُ الدِّراسةُ تزويدَها بأقسام الكلام الأساسيَّة (الاسم *Noun* والفعل *Verb* والأداة *Particle*)؛ إذ يَسهُلُ معها استخلاصُ المعاني الوظيفيَّة أثناءَ المُعالَجة. ويُوَضِّحُ (الجدول ١) نموذجًا لقاعدة بياناتٍ جذريَّةٍ، استمدَّ الباحثُ مادَّتها من مُدَوَّنة المُعجم التَّاريخيّ – المذكورة آنفًا – حتَّى نهاية القرن الثَّاني الهجريّ.

قسم الكلام	الوحدة	الجذر	التَّكرار	الكلمة
v	ابتَدَرَ	بدر	٦	ابتدار
v	ابتَدَرَ	بدر	١	ابتدارا
v	ابتَدَرَ	بدر	١	ابتدارها
v	ابتَدَرَ	بدر	١٤	ابتدر
v	ابتَدَرَ	بدر	٣	ابتدرا
v	ابتَدَرَ	بدر	١	ابتدراه
v	ابتَدَرَ	بدر	٤	ابتدرت
v	ابتَدَرَ	بدر	٦	ابتدرته
v	ابتَدَرَ	بدر	٧	ابتدرن
v	ابتَدَرَ	بدر	٢	ابتدرنا
v	ابتَدَرَ	بدر	١	ابتدره
v	ابتَدَرَ	بدر	٩	ابتدرها
v	ابتَدَرَ	بدر	١٦	ابتدروا
v	ابتَدَرَ	بدر	٢	ابتدروه
v	ابتَدَعَ	بدع	٣	ابتداعا
v	ابتَدَعَ	بدع	١	ابتداعها
v	ابتَدَعَ	بدع	١٩	ابتدع
v	ابتَدَعَ	بدع	١	ابتدعا
v	ابتَدَعَ	بدع	٦	ابتدعت

الجدول ١: نموذج توضيحيّ لقاعدة البيانات الجذريَّة للمُعجم التَّاريخيّ للعربيَّة

٣,١,٣. شَبكة الكلمات العَرَبيَّة.

تعملُ شبكةُ الكلمات بمثابة قاعدة بياناتٍ مُعجميَّة دلاليَّة *Lexical-Semantic Database*، تُصَنَّفُ المُفرَداتُ فيها إلى أقسام الكلام *Parts of Speech (PoS)*، ثُمَّ تُصَنَّفُ في سلاسلَ ومجمُوعاتٍ من المُتَرادفات *Synonyms* أو المفاهيم *Synsets* الَّتي تتَّفقُ في معانيها وتترابطُ فيما بينها بشبكةٍ من العلاقات الدِّلاليَّة *Semantic Relations*. وتُستَخدَمُ شبكاتُ الكلمات باعتبارها موردًا لُغَويًّا يُستفادُ منهُ في تطوير صناعة المُعجَم *Lexicography* وتطبيقات الذَّكاء الاصطناعيّ *Artificial Intelligence* ومُعالَجَة اللُّغات الطَّبِيعِيَّة *Natural Language Processing* (٢٠).

وتبدو الحاجةُ ماسَّةً إلى شبكةٍ للكلمات العربيَّة في صناعة المُعجم المنشود، نظرًا لما ينبغي أن يكونَ عليه مُحتوى المُعجم من الشُّمُوليَّة واتِّساع المادَّة. ذلكَ أنَّ مُدَوَّنة المُعجم ستحوي قدرًا كبيرًا من النُّصُوص العربيَّة غير المشكولة، الأمرُ الَّذي يُحدثُ التباسًا دلاليًّا في مجموعة الكلمات الَّتي تتَّفقُ في شكلها الكتابيّ. وعلى سبيل المثال، يُفتَرَضُ أن تَرِدَ في مُدَوَّنة المُعجم مجموعة الأشكال الكتابيَّة (من، بل، لكن)؛ ويُمكنُ أن تُصَنَّفَ هذه الأشكالُ باعتبارها اسمًا أو فعلًا أو أداة. فإذا جاءَت على الصُّورة (مَنٌّ، بَلٌّ، لَكْنٌ) فهي أسماء، وإذا جاءَت على الصُّورة (مَنَّ، بَلَّ، لَكَنَ) فهي أفعال. أمَّا إذا جاءَت على الصُّورة (مِنْ، بَلْ، لَكِنَّ) فهي أدوات. وتختلفُ دلالاتُ هذه المُفرَدات في كُلِّ قسمٍ من أقسام الكلام إلى عدَّةِ معانٍ مُعجميَّة. ووفقًا لطبيعة المُعجم المنشُود، يُمكنُ الإفادةُ من شبكة الكلمات في تعيين أقسام الكلام ومعانيها المُعجميَّة عبرَ علاقة (التَّرادُف) الَّتي تُمَثِّلُ إحدى العلاقات الدِّلاليَّة الرَّئيسة في شبكات الكلمات. ومع ما نَشهدُهُ من عدم اكتمال أيٍّ من مشروعات بناء شبكات الكلمات العربيَّة، فإنَّ الدِّراسةَ تقترحُ أن يُستفادَ بالمُحتوى المُتاح من شبكة الكلمات العربيَّة *AWN* في مراحل الصِّناعة المُعجميَّة الأوَّليَّة (٢١)، بعدَ إغناء قاعدة بياناتها بالمُترادفات العربيَّة، وفقًا لما تقتضيه طبيعةُ المُعجم المنشود. وتشتملُ الشَّبكةُ – في صُورتها الحاليَّة – في ١٧٥٦١ مُفردة، تَضُمُّها ٧٨٢٢ مجموعة ترادُفيَّة؛ تنتظمُ كما في (الشَّكل ٢).

الشَّكل ٢: شبكة الكلمات العربيَّة *AWN - v.01*

٣,٢. الموارد المُعجميَّة.

٣,٢,١. المُعجم التَّأثيليّ للعربيَّة.

تعودُ أقدمُ وثائق اللُّغة العربيَّة الَّتي يُمكنُ الاعتمادُ عليها في التَّأريخ لها إلى القرن الخامس قبلَ الهجرة؛ وتحديدًا إلى عام وفاة زعيم قبائل الأزد "مالك بن فَهْم الدُّوسيّ" (٤٠٣ ق.هـ). لكِنَّ هذا التَّاريخَ لا يُمَثِّلُ البِدايَةَ الحَقِيقِيَّةَ لوُجُود اللُّغَةِ العَرَبِيَّة، وإن كانَ يُعَبِّرُ عن حِقبةٍ متقَدِّمَةٍ من حِقَب التَّأريخ لنُصُوصِها. لقَد بَدَت العَرَبِيَّةُ مُكتمِلَةَ الأركان وواضِحَةَ المَعالِم في الوَثائِق والمَخطُوطات المُؤَرَّخَة في تلكَ الآوِنَة والَّتي تَلِيها، الأمرُ الَّذي يُؤَكِّدُ نُشُوءَ اللُّغَة العَرَبِيَّةِ واكتِمالَ هَيئَتِها قبل ذلك بِزَمَنٍ طَوِيل، وإن تَعَدَّدَت لهجاتُها وطُرُقُ النُّطْقِ بها؛ ومَعلُومٌ أنَّ هذا الشَّكْلَ الَّذي بَدَت عليه العَرَبِيَّةُ قَد سُبِقَ بِمراحِلَ أُخرَى، كَغَيرِها مِن اللُّغات المُشتَرِكَةِ معها في فَصِيلَتِها اللُّغَوِيَّة (السَّامِيَّة الحامِيَّة). ومِن ناحِيَةٍ أخرى، فقد تأثَّرَت العَرَبِيَّةُ بثَقافاتٍ وحَضاراتٍ عبرَ تاريخِها المَدِيد؛ وأدَّى اتِّصالُ العَرَبِيَّةِ بغَيرِها من اللُّغات إلى تَطَوُّرٍ في مبانِيها ومَعانِيها وأساليبها واستِعمالاتِها، فَدَخَلَتْها مُفرَدَاتٌ واندَثَرَت عنها مُفرَدَاتٌ أخرى، وتَبَدَّلَت دلالاتُ العَدِيد مِن الألفاظ والتَّراكيب فيها. ولأنَّ اللُّغَةَ العربيَّةَ لا تملكُ مُعجمًا تأثيليًّا يُؤصِّلُ لمُفرَداتها قبلَ أن تجريَ على ألسنة النَّاطقينَ بها وفي كتاباتهم، فإنَّ البديلَ المُناسبَ هو الاستعاضةُ عن المُعجم التَّأثيليّ بالموارد الَّتي يُعتَمَدُ عليها في التَّأثيلُ المُعجميُّ *Lexical Etymology* للمُفرَدات العربيَّة. ولا تخرُجُ هذه الموارِدُ – من وُجهة نظر الباحث – عن أربعةٍ، هي: مُعجمات النُّقُوش القديمة، وكُتُب اللَّهجات العَرَبيَّة، ومُعجَمات اللُّغات السَّاميَّة الحاميَّة، ومصادر المُعَرَّب والدَّخيل. ومن هذا المُنطلَق، تقترحُ الدِّراسةُ الاعتمادَ على الموارد الآتية للتَّأثيل:

٣,٢,١,١. مُعجمات النُّقُوش القديمة.

ويُمكنُ الاعتمادُ فيها على:

- كتاب "مجموعة نُصُوص النُّقُوش السَّاميَّة الشَّماليَّة *A text-book of North-Semitic inscriptions Moabite, Hebrew, Phoenician, Aramaic, Nabataean, Palmyrene, Jewish*"، الَّذي وَضَعَهُ الكَنَسِيّ الإنجليزيّ جورج كوك *George Albert Cooke* (١٨٦٥-١٩٣٩)، ونَشَرَهُ في أكسفورد عامَ ١٩٠٣ (٢٢).
- "مُعجم النُّقُوش السَّاميَّة الشَّماليَّة الغربيَّة *Dictionary of the North-West Semitic Inscriptions*"، الَّذي صَنَعَهُ فريقٌ من الباحثينَ في النُّقُوش واللُّغات السَّاميَّة، على رأسهم الأكاديميّ الأمريكيّ ريتشارد ستاينر *Richard Steiner*. وقد نُشِرَ في لَيدن بهولندا عامَ ٢٠٠٣ (٢٣).

٣,٢,١,٢. كُتُب اللَّهجات العَرَبيَّة.

ويُمكنُ الاعتمادُ فيها على الكُتُب الآتية:

- "العربيَّة: دراسات في اللُّغة واللَّهجات والأساليب *Arabiya,: Untersuchungen zur arabischen Sprach- und Stilgeschichte*"، الَّذي وَضَعَهُ المُستشرق الألمانيّ يوهان فُك *Johann Fück* (١٨٩٤-١٩٧٤)، ونَشَرَهُ في برلين عامَ ١٩٥٠ (٢٤).
- "اللهجات العربية القديمة في غرب الجزيرة العربية *Ancient West-Arabian*"، الَّذي وَضَعَهُ اللُّغَويّ الألمانيّ الأصل حاييم رابين *Chaim Rabin* (١٩١٥-١٩٩٦)، ونَشَرَهُ في لندن عامَ ١٩٥١ (٢٥).
- "في اللَّهجات العَرَبيَّة" للُّغَوِيّ المِصرِيّ إبراهِيم أنيس (١٩٠٦-١٩٧٧). نَشَرَهُ في القاهرة عامَ ١٩٦٥ (٢٦).
- "في اللَّهجات العَرَبيَّة القَدِيمة" للُّغَوِيّ العِراقِيّ إبراهِيم السَّامرَّائِيّ (١٩٢٣-٢٠٠١). يُلخِّصُ العديدَ من الجُهُود السَّابقة عليه؛ وقد نَشَرَهُ في بيروت عامَ ١٩٩٤ (٢٧).

٣,٢,١,٣. مُعجَمات اللُّغات السَّاميَّة الحاميَّة.

ويُمكنُ الاعتمادُ فيها على:

- "المُعجم الألمانيّ العبريّ لمُفردات العهد القديم، مع العناية بأسماء الأماكن والمُفرَدات الكلدانيَّة *Hebräisch-deutsches Handwörterbuch über die Schriften des Alten Testaments : mit Einschluss der geographischen Nahmen und der chaldäischen Wörter beym Daniel und Esra*"، الَّذي وَضَعَهُ المُستشرق الألمانيّ فيلهلم جزينيوس *Wilhelm Gesenius* (١٧٨٦-١٨٤٢)، ونَشَرَهُ عامَ ١٨١٠ (٢٨). ويُعنى المُعجمُ بتأثيل مُفرَدات العهد القديم في اللُّغات السَّاميَّة.
- "المُعجم العبريّ الآراميّ للعهد القديم *Hebräisches und Aramäisches Handwörterbuch über das Alte Testament*". وهو مُعجمٌ آخَرٌ لفيلهلم جزينيوس، نشَرَهُ في عام ١٨١٥ (٢٩). ويُعنى هذا المُعجمُ بتأثيل مُفرَدات العهد القديم في اللُّغات السَّاميَّة والنُّقُوش العبريَّة والآراميَّة القديمة.
- "قامُوس الكِتاب المُقَدَّس: في الآثار والتَّراجم والجُغرافيا والتَّاريخ الطَّبيعيّ *A Dictionary of the Bible: Antiquities, Biography, Geography and Natural History*"، الَّذي وَضَعَهُ المُعجَميّ الإنجِليزِيّ ويليام سمِيث *Sir William Smith* (١٨١٣-١٨٩٣)، ونَشَرَهُ في لندن عامَ ١٨٦٣ (٣٠). ويُعنى هذا المُعجمُ بتأثيل مُفرَدات الكِتاب المُقَدَّسِ في اللُّغات السَّامِيَّةِ والحامِيَّة.
- "المُعجم المُقارن للُّغات الهندوأوروبيَّة والسَّاميَّة *Vergleichendes indogermanisch-semitisches Wörterbuch*"، الَّذي وَضَعَهُ اللُّغَويّ الدَّنماركيّ هرمان مُولر *Hermann Möller* (١٨٥٠-١٩٢٣)، ونَشَرَهُ في كوبنهاجن عامَ ١٩١١ (٣١).
- "الجُذُور المُشتَرَكة في مُفردات اللُّغات السَّاميَّة والهندوأوروبيَّة: مُحاولة تأثيليَّة *Die Gemeinsamen Wurzeln Des Semitischen und Indogermanischen Wortschatzes: Versuch Einer Etymologie*"، الَّذي وَضَعَهُ السَّويسريّ لينُوس بُرونر *Linus Brunner* (١٩٠٩-١٩٨٧)، ونَشَرَهُ في ميونخ عامَ ١٩٦٩ (٣٢).
- "مُعجَم الجُذُور السَّامِيَّة كما تبدو في اللُّغات السَّاميَّة *Dictionnaire des racines sémitiques ou attestées dans les langues sémitiques*"، الَّذي أنجَزَهُ فريقٌ من الباحثين، وأشرَفَ عليه اللُّغَوِيّ الفَرَنسِيّ – من أصل تونُسيّ – دِيفِيد كوهِين *David Cohen* وفقَ المنهج الَّذي وَضَعَهُ أستاذُهُ المُستَشرِق الفَرَنسِيّ جان كانتينو *Jean Cantineau* (١٨٩٩-١٩٥٦)، ونُشِرَ في سلسلةٍ من عشرة أجزاء، آخرها في عام ٢٠١٢ (٣٣). ويتميَّزُ هذا المُعجمُ بأنَّهُ يُلخِّص جُهودَ السَّابقين في مُقارنة اللُّغات السَّاميَّة.

٣,٢,١,٤. مصادر المُعَرَّب والدَّخيل.

ويُمكنُ الاعتمادُ فيها على الكُتُب الآتية:

- "المُعَرَّب من الكلام الأعجميّ على حُرُوف المُعجَم" لأبي منصُور الجواليقيّ البغداديّ (ت ٥٤٠ هـ) (٣٣).
- "تَكْمِلَة المُعجَمات العَرَبيَّة *Supplement Aux Dictionnaires Arabes*" للمُستَشرِق الهولنديّ رينهارت دوزي *Reinhart Dozy* (١٨٢٠-١٨٨٣). وقد نَشَرَهُ في لَيدِن عامَ ١٨٨١ (٣٥).
- "الألفاظ الفارِسيَّة المُعَرَّبة" لآدي شير الكِلْدانيّ (١٨٦٧-١٩١٥)؛ وقد نَشَرَهُ في بيروت عامَ ١٩٠٨ (٣٦).
- "غرائب اللُّغة العَرَبيَّة" للأب روفائيل نخلة (١٨٩٠-١٩٧٣)؛ وقد نَشَرَهُ في بيروت عامَ ١٩٦٠ (٣٧).
- "مُعجم الدَّخيل في اللُّغة العربيَّة ولهجاتها"؛ وهُو مُعجمٌ حديثٌ، وضَعَهُ الباحثُ الهنديّ فانيامبادي عبد الرحيم *Vaniyambadi Abdur Rahim*، ونشَرَهُ في دمشق عامَ ٢٠١١ (٣٨).

٣,٢,٢. قاعدة البيانات المُعجَميَّة.

قاعدةُ البيانات المُعجميَّة *Lexical Database* هي بيئةٌ برمَجيَّةٌ تحوي المعلوماتِ المُعجميَّة الَّتي يُمكنُ الإفادةُ منها في تحرير المُعجمات وبناء الموارد اللُّغويَّة. ووفقًا لطبيعة المُعجم التَّاريخيّ المنشود، سنكونُ بحاجةٍ إلى قاعدة بياناتٍ مُعجميَّةٍ تشتملُ على المعاني المُعجميَّة لمُفرَدات العربيَّة عبرَ تاريخِها المديد. وعليه، تقترحُ الدِّراسةُ – لتحرير المعاني المُعجميَّة – الاعتمادَ على قاعدة بياناتٍ تضُمُّ مُعجماتٍ للُّغة العربيَّة القديمة والوسيطة والحديثة (٣٩) على النَّحو الآتي.

٣,٢,٢,١. مُعجمات اللُّغة العربيَّة القديمة.

ويُمكنُ الاعتمادُ فيها على:

- "مُعجم (العَيْن)" للخليل بن أحمد (ت ١٧٣ هـ) (٤٠).
- "مُعجَم (تَهذِيب اللُّغَة)" لأبي مَنصُور الأزهَرِيّ (ت ٣٧٠ هـ) (٤١).
- "مُعجَم اللُّغة العَرَبيَّة الفُصحى *Wörterbuch der Klassischen Arabischen Sprache*"، الَّذي يَعمَلُ على إنجازه فريقٌ من المُستشرقين في الجمعيَّة الألمانيَّة للاستشراق. وقد صَدَرَ منهُ – حتَّى الآن – جُزءان، نُشِرا بينَ عامَي ١٩٧٠ و ٢٠٠٩ (٤٢)؛ ويَضُمَّان بابَي المُعجم العربيّ (الكاف) و (اللَّام).

٣,٢,٢,٢. مُعجمات اللُّغة العربيَّة الوسيطة.

ويُمكنُ الاعتمادُ فيها على:

- "القامُوس المُحِيط" لمُحَمَّد بن يعقوب الفَيْرُوزَابَاديّ (ت ٨١٧ هـ) (٤٣).
- "تاج العَرُوس مِن جَواهِر القامُوس" لمُرتَضى الزَّبِيدِيّ (ت ١٢٠٥ هـ)؛ وهو شرحٌ للقاموس المُحيط (٤٤).
- "مَدّ القاموس *An Arabic-English Lexicon, derived from the best and the most copious Eastern sources, comprising a very large collection of words and significations omitted in the Kámoos, with supplements to its abridged and defective explanations, ample grammatical and critical comments, and examples in prose and verse*". وهُوَ مُعجمٌ كبيرٌ، كانَ قد بدأَ العملَ فيه المُستَشرِق الإنجليزِيّ إدوارد لين *Edward Lane* (١٨٠١-١٨٧٦) في عام ١٨٤٢م، وانتَهى فِيه إلى حَرف القاف، ثُمَّ وافَتهُ المَنِيَّة قَبلَ أن يُتِمَّه (٤٥). وقد طُبِعَ المُعجَم كامِلًا بعد وفاة إدوارد لين، حيث أَكمَلَهُ – بِناءً على رَغبَتِهِ – ابن أخيه المُستَشرِق وعالِم الآثار ستانلي لين بول *Stanley Lane-Poole* (١٨٥٤-١٩٣١)، وإن كانَ بول قد استَقى أَكثرَ الموادّ المُتَبَقِّيَة مِن مُعجَم "تاج العروس" للزَّبيديّ.

٣,٢,٢,٣. مُعجمات اللُّغة العربيَّة الحديثة والمُعاصرة.

ويُمكنُ الاعتمادُ فيها على:

- "مُعجَم اللُّغة العَرَبيَّة الحديثة المكتُوبة *Arabisches Wörterbuch für die Schriftsprache der Gegenwart*" للمُستَشرِق الألمانِيّ هانز فِير *Hans Wehr* (١٩٠٩-١٩٨١). وقد نَشَرَهُ في لَيبتسك بألمانيا عامَ ١٩٥٢ (٤٦).
- "مُعجم اللُّغة العربيَّة المُعاصرة"، الَّذي حرَّرَه المُعجميُّ المصريّ أحمد مختار عُمَر (١٩٣٣-٢٠٠٣) بمُساعدة فريق عمل. وقد طُبِعَ في القاهرة عامَ ٢٠٠٨ (٤٧).
- "المُعجم الكبير" الَّذي يعملُ على إنجازه مَجمَعُ اللُّغة العربيَّة في القاهرة. وقد صَدَرَت منه – حتَّى الآن – سبعةُ أجزاء، نُشِرَت بينَ عامَي ١٩٥٦ و ٢٠١٢؛ وتضمُّ أبوابَ المُعجم العربيّ من (الهمزة) إلى (الرَّاء) (٤٨).

٣,٣. الموارد الحاسُوبيَّة.

٣,٣,١. نظام الفهرسة الآليَّة.

تشتملُ أنظمةُ الفهرسة الآليَّة على أداتَين رئيسَتَين، هما: المُفهرس الآليّ للنُّصُوص *Text Indexer*؛ ويُستَخدَمُ في حصر وإحصاء وترتيب كلمات المُدَوَّنة اللُّغَويَّة؛ والكشَّافُ السِّياقيُّ *Concordancer*؛ ويُستَخدَمُ في استخلاص السِّياقات الَّتي ترِدُ فيها كُلُّ كلمةٍ على حِدة. وتبدو الحاجةُ ماسَّةً إلى نظام الفهرسة الآليَّة مع ضخامة حجم المُدَوَّنة اللُّغَويَّة للمُعجَم المنشُود؛ إذ المُفتَرَضُ أن تَضُمَّ مئات الملايين من الكلمات الَّتي تَصعُبُ مُعالَجتُها يدويًّا. ويعملُ النِّظامُ على استخلاص كلمات المُدَوَّنة، بما يُساعدُ في تعيين المداخل والوحدات المُعجميَّة لاحقًا، كما يؤدِّي دورًا مُهمًّا في تعيين المعاني المُعجميَّة من خلال السِّياقات الَّتي يعرضُها (٤٩). وتُوجدُ أنظمةُ الفهرسة الآليَّة الَّتي تدعمُ اللُّغةَ العربيَّة في ثلاثة أنواع، هي:

- أنظمة الفهرسة الآليَّة الجذريَّة: تُرَتَّبُ فيها المُفرداتُ بحسب الجُذُور *Roots*. ويَصعُبُ استخدامُ هذه الأنظمة في مُعالَجة المُدَوَّنة اللُّغَويَّة للمُعجم، نظرًا لما تُعانيه من التباسٍ صرفيٍّ ناتجٍ عن الطَّبيعة الاشتقاقيَّة للعربيَّة.
- أنظمة الفهرسة الآليَّة الجذعيَّة: تُرَتَّبُ فيها المُفرداتُ بحسب الجُذُوع *Stems*. ويَصعُبُ استخدامُها – كذلكَ – نظرًا لما تُعانيه من قُصُورٍ في استخلاص الجُذوع أو استيعاب مجموعات النُّصُوص الكبيرة نسبيًّا.
- أنظمة الفهرسة الآليَّة الألفبائيَّة: تُرَتَّبُ فيها المُفرداتُ ألفبائيًّا؛ وهي الَّتي يُمكنُ الاعتمادُ عليها؛ حيثُ يُلتَزَمُ فيها بقواعدَ معياريَّة للإحصاء والتَّرتيب، كما تسمحُ باستيعاب مئات الملايين من الكلمات. وثمَّةَ العديدُ من أنظمة الفهرسة الآليَّة الألفبائيَّة الَّتي تتفاوتُ في كفاءتها ودقَّة مُخرَجاتها. وفي ضوء طبيعة المُعجم التَّاريخيّ المنشُود، تقترحُ الدِّراسةُ أن يُستفادَ من نظام الفهرسة المُلحَق ببرمجيَّة (*tlCorpus*) مُغلَقة المصدر، الَّتي أنجزَتها شركة تقنيات اللُّغات الإنسانيَّة *TshwaneDJe* (٥٠). ويُوَضِّحُ (الشَّكل ٤) مُخرَجات هذا النِّظام.

الشَّكل ٣: مُخرَجات نظام الفهرسة الآليَّة المُلحَق ببرمجيَّة *tlCorpus – v.7.1*

٣,٣,٢. نظام التَّحليل الصَّرفيّ.

يقومُ نظامُ التَّحليل الصَّرفيّ *Morphological Analysis System* بتعيين المداخل والوحدات المُعجَميَّة، مُتَمِّمًا – بذلكَ – عملَ نظام الفهرسة الآليَّة، كما يقومُ بدوره الأساسيّ في تعيين المعاني الوظيفيَّة البنويَّة السَّماعيَّة، كأبواب الفعل الثُّلاثيّ المُجرَّد وصيغ الجُمُوع ونحوها. وقد ساعَدَ نظامُ التَّصريف العربيّ القياسيّ على تطوير خوارزميَّة التَّحليل الصَّرفيّ العربيّ وتحسين أداء المُحلِّلات الصَّرفيَّة لها، كي تصلَ إلى درجاتٍ عاليةٍ من الدِّقَّة والكفاءة؛ وإن كانت تُعاني بعضَ جوانب القُصُور (٥١). وثمَّةَ العديدُ من أنظمة التَّحليل الصَّرفيّ الَّتي يُمكنُ الإفادةُ منها في مُعالجة مادَّة المُعجم التَّاريخيّ المنشود؛ منها:

- المُحلِّل الصَّرفيّ *Aramorph* مفتوح المصدر. وقد صَنَعَهُ الباحثُ الأمريكيّ تِيم بكوالتر *Tim Buckwalter* ضمنَ عمله في مؤسَّسة البيانات اللُّغويَّة *Linguistic Data Consortium (LDC)* بجامعة بنسلفانيا الأمريكيَّة (٥٢)؛ وإن كانَ يَعيبُهُ ضعفُ نتائجه في مُعالجة النُّصُوص العربيَّة المُستمَدَّة من فُصحى التُّراث، نظرًا لاعتمادهِ على قواعد بياناتٍ لنُصُوص اللُّغة العربيَّة المُعاصرة. أضف إلى ذلكَ أنَّهُ يعرضُ نتائجَهُ بحُرُوفٍ لاتينيَّة، الأمرُ الَّذي يستدعي القيامَ بإجراءٍ حاسُوبيّ لرَوْمَنة الحُرُوف العربيَّة *Romanization of Arabic*.
- نظام "الخليل" الصَّرفيّ مفتوح المصدر. وهو نتاجُ عملٍ مُشتَرَكٍ بينَ الباحث (المُعتزّ بالله السَّعيد) وفريق مُعالجة اللُّغات الطَّبيعيَّة بجامعة مُحمَّد الأوَّل المغربيَّة، برعاية المُنظَّمة العربيَّة للتَّربية والثَّقافة والعُلُوم *ALECSO* ومدينة الملك عبد العزيز للعُلُوم والتِّقنية بالرِّياض (٥٣). ويَعيبُ هذا النِّظامَ كثرةُ احتمالات التَّحليل الصَّرفيّ الَّتي يُوردُها لكلّ كلمة على حدة، نظرًا لكبر حجم قاعدة بياناته وخُلُوِّه من آليَّةٍ لتنقية المُخرَجات القياسيَّة المُهملة في العربيَّة، الأمرُ الَّذي يستدعي التَّدخُّلَ اليدويَّ لمُعالجة المُخرَجات.
- نظام *SARF* مُغلَق المصدر. وهُوَ نتاجُ مشروعٍ بحثيٍّ أنجَزتهُ شركة *Microsoft* الأمريكيَّة الرَّائدة، لبناء آليَّةٍ للتَّحليل الصَّرفيّ للنُّصُوص العربيَّة (٥٤). وتراهُ الدِّراسةُ مُقَدَّمًا على نِظامَي "*Aramorph*" و "الخليل"؛ إذ يستدركُ جوانبَ القُصُور فيهما، كما يُعنى بترتيب النَّتائج والمُخرَجات بحسب أكثرها دورانًا في اللُّغة، اعتمادًا على مُدَوَّناتٍ لُغَويَّةٍ مُمثِّلةٍ لواقع اللُّغة. ويُوضِّحُ (الجدول ٢) نموذجًا لمُخرَجات هذا النِّظام (٥٥).

الحكم	التَّحليل ١	التَّحليل ٢	التَّحليل ٣
الكلمة مشكولة ***Diacritized Word***	اَلْحُكْم	اَلْحَكَم	اَلْحِكَم
السَّوابق	ال	ال	ال
الجذع	حُكْم	حَكَم	حِكَم
اللَّواحق	--	--	--
قسم الكلام ***Part of Speech***	مصدر مُجرَّد	اسم ذات	اسم ذات "تكسير"
التَّوصيف النَّحويّ [البنويّ - التَّركيبيّ] ***Morpho-Syntactic***	*Masculine, Singular, GendMinus, DefinedThe*	*Masculine, Singular, DefinedThe, NameUsedAdjective*	*Feminine, NumberPluralBroken, DefinedThe*
الاحتمالات	٠,٧٣٦٨٠٧٩	٠,٢٦٢٧٢٥١	٠,٠٠٠٤٦٦٩٥٥٤

الجدول ٢: مُخرجات نظام *SARF* للتَّحليل الصَّرفيّ *(Microsoft SARF – 2014)*

٣,٣,٣. نظام التَّحليل التَّركيبيّ.

يُستفادُ من نظامُ التَّحليل التَّركيبيّ *Syntactic Analysis System* في تعيين المعاني الوظيفيَّة التَّركيبيَّة السَّماعيَّة؛ مثل: أحوال الفعل بين اللُّزوم والتَّعدّي، ومعاني الأدوات، ونحو ذلك. وتَبدُو خوارزميَّةُ التَّحليلِ التَّركِيبِيّ للعَرَبِيَّةِ مُضطَرِبَةً ومُعَقَّدَةً إلى حَدٍّ بعيد، نتيجةً لطبيعة اللُّغة العربيَّة الاشتقاقيَّة ونظامها الكتابيّ من ناحية، وطول الجُملة العربيَّة وتعدُّد أنماطها مُقارنةً لها مع بقيَّة اللُّغات الطَّبيعيَّة الأخرى من ناحيةٍ ثانية، وتداخُل خوارزميَّة التَّحليل التَّركيبيّ مع خوارزميَّة التَّحليل الصَّرفيّ من ناحيةٍ ثالثة. ويجعلُ هذا الاضطرابُ والتَّعقيدُ بِناءَ مُحلِّلٍ تَركِيبِيٍّ كاملٍ للُّغةِ العَرَبِيَّةِ أمرًا بالِغَ الصُّعُوبة عالِيَ التَّكلِفَة، سواءٌ على الباحِثِين الأفرادِ أم على المُؤَسَّسَات المَعنِيَّة بِحَوسَبةِ اللُّغة (٥٦). ومع هذا، توجدُ بعضُ الأدوات الَّتي يُمكنُ توظيفُها في التَّحليل التَّركيبيّ لمادَّة المُعجم، والَّتي تُعرَفُ بـ "مُعَنونات أقسام الكلام *Parts of Speech Taggers*"؛ حيثُ تُساعدُ هذه الأدواتُ في التَّمييز بينَ أقسام الكلام (الاسم والفعل والأداة)، كما تُساعدُ في توصيف كُلِّ قسمٍ منها على حدة. ومن الأدوات الَّتي يُمكنُ الإفادةُ منها في الصِّناعة المُعجميَّة:

- مُعنون نُصُوص ستانفورد *Stanford Log-linear Part-Of-Speech Tagger* مفتوح المصدر. أنجَزَهُ الفريقُ البَحثِيّ بمجموعة مُعالَجة اللُّغات الطَّبِيعِيَّة في جامِعة ستانفُورد الأمريكِيَّة. يدعمُ اللُّغات الإنجليزيَّة والألمانيَّة والصِّينيَّة والعربيَّة (٥٧)؛ وإن كانَ يَعيبُهُ تعدُّد أقسام الكلام فيه بما يتناسبُ مع اللُّغات الجرمانيَّة، على خلاف ما تتطلَّبُهُ الصِّناعةُ المُعجميَّة العربيَّة، الأمرُ الَّذي يتطلَّبُ التَّدخُّلَ اليدويَّ لاختزالها.
- المُعنون النَّحويّ للمُفرَدات العربيَّة *RDI's Arabic Part-of-Speech Tagger (ArabTagger)* مُغلَق المصدر. وهو نظامٌ للمُعالجة التَّركيبيَّة للنُّصُوص العربيَّة، أنتَجَتْهُ الشَّركة المصريَّة لتطوير النُّظُم الرَّقميَّة *RDI* ضمنَ مجموعة تقنياتٍ لمُعالَجة اللُّغة العربيَّة (٥٨). ومُخرَجاتُهُ مُوَضَّحةٌ في (الجدول ٣).

الحل	الكلمة	*PoS Tags vector*
١٥/١	بَلْ	لا سابِق - عَطْف - لا لاحِق
١٥/٢	بِلّ	لا سابِق - اسْم - لا لاحِق
١٥/٣	بَلّ	لا سابِق - اسْم - لا لاحِق
١٥/٤	بَلَّ	لا سابِق - مَبْنِيّ لِلمعلوم - فِعْل - ماضٍ - لا لاحِق
١٥/٥	بِلّ	لا سابِق - مَبْنِيّ لِلمعلوم - فِعْل - مُضارِع أو أَمْر - لا لاحِق
١٥/٦	بُلْ	لا سابِق - مَبْنِيّ لِلمعلوم - فِعْل - مُضارِع أو أَمْر - لا لاحِق
١٥/٧	بِلْ	لا سابِق - مَبْنِيّ لِلمعلوم - فِعْل - مُضارِع أو أَمْر - لا لاحِق
١٥/٨	بُلّ	لا سابِق - اسْم - لا لاحِق
١٥/٩	بُلّ	لا سابِق - مَبْنِيّ لِلمعلوم - فِعْل - مُضارِع أو أَمْر - لا لاحِق
١٥/١٠	بَلّ	لا سابِق - مَبْنِيّ لِلمعلوم - فِعْل - مُضارِع أو أَمْر - لا لاحِق
١٥/١١	بُلَّ	لا سابِق - مَبْنِيّ لِلمجهول - فِعْل - ماضٍ - لا لاحِق
١٥/١٢	بَلِّ	لا سابِق - مَبْنِيّ لِلمعلوم - فِعْل - مُضارِع أو أَمْر - لا لاحِق
١٥/١٣	بَلِّ	لا سابِق - مَبْنِيّ لِلمعلوم - فِعْل - مُضارِع أو أَمْر - لا لاحِق
١٥/١٤	بَلٍ	لا سابِق - اسْم - صِيغَة مُبالَغة - لا لاحِق
١٥/١٥	بَلٍ	لا سابِق - اسْم - صِيغَة مُبالَغة - لا لاحِق

الجدول ٣: مُخرَجات المُعنون النَّحويّ للمُفرَدات العربيَّة *RDI's Arabic PoS Tagger*

٣,٣,٤. نظام التَّحليل الدِّلاليّ.

مع اتِّساع حجم المُدَوَّنة اللُّغويَّة للمُعجم التَّاريخيّ، تبدو الحاجةُ ماسَّةً إلى وسيلةٍ لإزالة الالتباس الدِّلاليّ الواقع في المباني مُتعدِّدة المعاني. ويتمُّ هذا من خلال نظام التَّحليل الدِّلاليّ *Semantic Analysis System*؛ حيثُ يُساعدُ على تمييز المعاني المُعجميَّة لكلِّ وَحدةٍ تحملُ أكثرَ من معنى. والواقعُ أنَّ المُجتمَعَ المعلوماتيَّ العربيَّ لم يعرف طريقَهُ بعدُ إلى أنظمةٍ فعَّالةٍ للتَّحليل الدِّلاليّ، إلَّا من خلال مجموعةٍ من الدِّراسات والبُحُوث الَّتي لا تزالُ في طور التَّطوير. وتدورُ هذه الدِّراساتُ في فلك ما يُعرفُ بآليَّة فَكِّ الالتباس الدِّلاليّ للكَلِماتِ *Word Sense Disambiguation (WSD)*؛ وهي آليَّةٌ تستمدُّ فكرتَها من المُتصاحبات اللَّفظيَّة *Collocations* الَّتي تتَكَوَّنُ عن سِلسِلَةٍ مِن كلِمَتَين أو أكثَر، تتَلازَمُ مُفرَداتُها في عَلاقَاتٍ تَركِيبِيَّة؛ وإن كانَت الآليَّةُ تأخُذُ بُعدًا آخَرَ؛ حيثُ يَكُونُ التَّلازُمُ على مُستوى السِّياقِ بأكمَلِه، بالنَّظر إلى مجموعة الكلمات الَّتي ترتبطُ بكلمةٍ مُعيَّنةٍ في حَيِّز الجوار، سواءٌ أكانَت سابِقَةً لها أم لاحِقَةً بِها، وسواءٌ أكانَت مُصاحِبَةً لها أم مُنفَصِلةً عنها (٥٩).

وسعيًا إلى توفير نظامٍ للتَّحليل الدِّلاليّ لمادَّة المُعجم المنشُود يُعتَمَدُ عليه في تمييز المعاني المُعجميَّة، تقترحُ الدِّراسةُ أن يُستَفادَ من "مُصَنِّف بايز المُبَسَّط *Naïve Bayes Classifier*" المُنبثق عن خوارزميَّة بايز للفيلسوف والإحصائيّ الإنجليزيّ توماس بايز *Thomas Bayes* (١٧٠٢-١٧٦١). ويشيعُ استخدامُ مُصَنِّف بايز المُبَسَّط لدقَّة نتائجه مُقارنةً بغيره من وسائل فكّ الالتباس الدِّلاليّ للكلمات؛ وإن كانَ يتطلَّبُ وقتًا وجهدًا في عمليَّتَي الإعداد والتَّدريب. ويفترضُ هذا المُصَنِّفُ استِقلالَ دِلالات الكَلِمة في السِّياق عن بعضِها، كما يَفتَرِضُ أن تَتساوى احتِمالاتُ المعلومات الدِّلاليَّة المُصاحِبة لهذه الكَلِمة. وعلى سبيل المِثال: يُمكِنُ التَّعبِيرُ عن كَلِمة (أسَد) بتوصِيفِها دِلاليًّا باستِخدام مُصَنِّف بايز المُبَسَّط، باعتبار المُتَصاحِبات الثَّلاثة: (حيوان / مِن الثَّدييَّات / مِن آكِلي اللُّحُوم). ووفقًا لقاعدة المُصَنِّف، تُؤَدِّي كُلُّ مُتَصاحِبَة دَورَها في التَّعبير عن الكلمة (أسد). ويُوَضِّحُ (الجدول ٤) نتائجَ تجربةٍ بحثيَّة مُشتركة قامَ بها الباحثُ مع فريق عملٍ بكلِّيَّة الهندسة في جامعة القاهرة ومعهد بُحُوث الإلكترونيَّات بالمركز القوميّ للبُحُوث في القاهرة والشَّركة المصريَّة لتطوير النُّظُم الرَّقميَّة *RDI*، لتمييز المعاني المُعجميَّة في مُدَوَّنةٍ لُغويَّةٍ تحوي ٥٤٠٢ نموذج باستخدام مُصَنِّف بايز المُبَسَّط. وقد تأكَّدَت لفريق العمل فاعليَّة هذا المُصَنِّف في التَّحليل الدِّلاليّ للعربيَّة من خلال نتائجه (٦٠).

الكلمة	المعاني المُعجميَّة	عدد النَّماذج	نسبة تميز الكلمات	نسبة تميز الجذوع
سَرطان	حيوانٌ بحريٌّ من القشريَّات أحد بروج السماء، بين الجوزاء والأسد ورم خبيث يتولد في الخلايا الظاهرية الغُدّية	٧٥٧	٨٦,٥٦ ٪	٨٧,٣٥ ٪
جَدول	مَجرًى صَغير للمِياه صَفحَة يُخَطّ فيها خُطُوط مُتَوازِيَة ومُتَقاطِعَة	٩١٤	٨٢,٧١ ٪	٨٣,٥٦ ٪
مَشروع	ما سَوَّغَهُ الشَّرْع الأمر يُهَيَّأ لِيُدرَس ويُقَرَّر	١١٨٤	٨٧,٠٠ ٪	٨٧,٠٠ ٪
مَرسُوم	المُعَيَّنُ مِن الأشياء رَسْمًا وخَطًّا ووَصْفًا قانُون ذو صبغَة تَشرِيعِيَّة	١٣٥٨	٩٢,٥٠ ٪	٩٢,٧٠ ٪
حاجِب	الشَّعْر النَّابِت فَوقَ العَيْن خازن الباب وحارِسه عَلَم (من أسماء الذُّكور)	١١٨٩	٨٢,٢٦ ٪	٨١,٠٠ ٪

الجدول ٤: نتائج التَّحليل الدِّلاليّ في العربيَّة باستخدام "مُصَنِّف بايز المُبَسَّط *Naïve Bayes Classifier*"

٣،٣،٥. المِنصَّة الحاسُوبيَّة.

المِنَصَّةُ الحاسُوبيَّة *Computational Framework* هي بيئةٌ معلوماتيَّةٌ واسعةُ النِّطاق، تضُمُّ مجموعةَ الموارد والأدوات المُستخدمة في مُعالَجاتٍ حاسُوبيَّةٍ للمُدَوَّنات والمُعجمات اللُّغويَّة، كما يُمكنُ توظيفُها في العديد من تطبيقات مُعالجة اللُّغات الطَّبيعيَّة، كالتَّنقيب في النُّصُوص *Text Mining* والبحث في مُستودعات البيانات *Data-Warehouse* وغيرها (٦١). وسيكونُ من المُفيد أن يُستَفادَ من المنصَّة الحاسُوبيَّة في استخلاص الشَّواهد المُعجميَّة وتعيين مُستويات الاستعمال في المُعجم التَّاريخيّ المنشود؛ حيثُ تُزَوَّدُ المِنَصَّاتُ الحاسُوبيَّةُ – عُمُومًا – بمُحرِّكاتٍ بحثيَّة *Search Engines* قادرة على التَّعامُل مع المجموعات الضَّخمة من النُّصُوص. وتُشبهُ هذه المُحرِّكاتُ في عملها طريقةَ الكشَّافات السِّياقيَّة في إدراج السِّياقات الَّتي ترِدُ فيها الكلمة الواحدة، وإن كانَت أكثرَ قُدرةً وكفاءةً في المِنصَّات الحاسُوبيَّة، نظرًا لضخامة قواعد بياناتها نسبيًّا وكثرة الموارد والأدوات الَّتي يُمكنُ توظيفُها في المُعالَجة الآليَّة للمُدخَلاتِ من النُّصُوص أو المُعجمات.

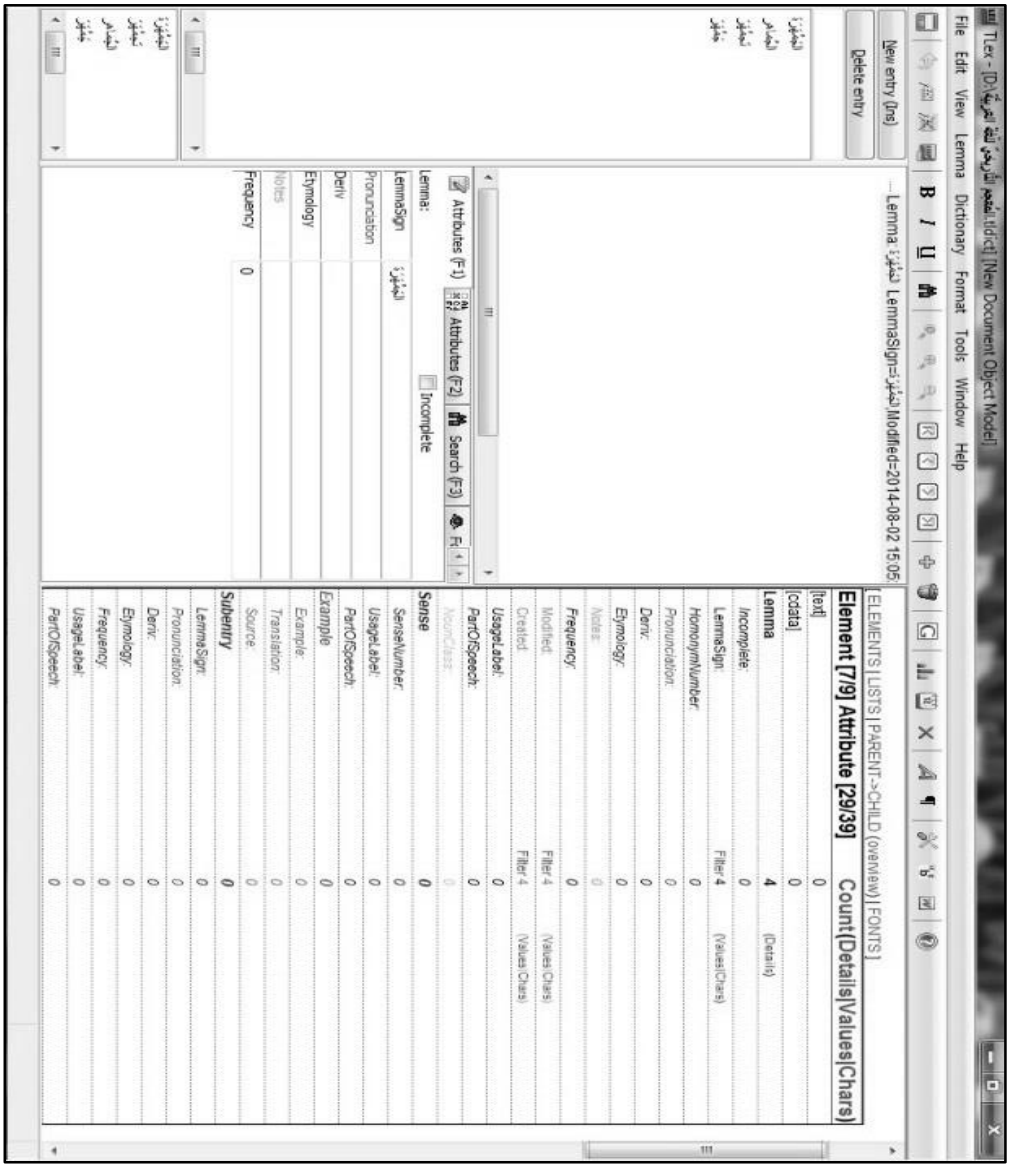
وتحقيقًا للهدف المنشُود، تقترحُ الدِّراسةُ أن يُستَفادَ من مِنصَّة *Nooj* مفتوحة المصدر. وهي بيئةُ تطويرٍ لُغَويَّة، صَمَّمَها الباحثُ الفرنسيُّ ماكس سِلبرزتن *Max Silberztein* في عامَ ٢٠٠٢، ثُمَّ قامَ بتطويرها مع العديد من الباحثينَ حولَ العالَم (٦٢). وتدعمُ هذه المِنصَّةُ ثلاثًا وعِشرينَ لُغة طبيعيَّة، منها العربيَّة، كما تحتوي على مجموعةٍ من الموارد اللُّغَويَّة والمُعجَميَّة والأدوات المُستخدَمة في المُعالجة الآليَّة للُّغات، بما فيها الكشَّاف السِّياقيّ الَّذي يعملُ كمُحرِّك بحثٍ عن النُّصُوص؛ وتمتازُ بإمكانيَّة التَّحكُّم في مواردها وأدواتها بالإضافة أو التَّعديل (٦٣). ويُوضِّحُ (الشَّكل ٤) نموذجًا لهذه المِنَصَّة.

الشَّكل ٤: نموذج توضيحيّ لِمنصَّة *Nooj* (٢٠١٤)

٣,٣,٦. نظام التَّحرير المُعجميّ.

يُمثِّلُ نظامُ التَّحرير المُعجَميّ بيئةً حاسُوبيَّةً تفاعُليَّة، يُمكنُ من خلالها التَّحكُّمُ في المعلومات المُعجميَّة المُستخلَصَة من مُخرَجات الموارد اللُّغَويَّة والمُعجميَّة والحاسُوبيَّة. وتُساعدُ هذه البيئةُ في تنظيم المعلُومات المُعجميَّة وترتيب الموادّ المُدرَجة في هيكل المُعجم وفقًا لما يقتضيه النِّظامُ اللُّغَويُّ وما تَفرضُهُ طبيعةُ الصِّناعة المُعجميَّة (٦٤). ويُعوَّلُ على أنظمة التَّحرير المُعجميّ في إخراج مادَّة المُعجَم في صيغةٍ قياسيَّةٍ ومُنتظمة تتوحَّدُ فيها مناهج المُحرِّرين، ليبدو المُعجمُ عملًا مُتوازنًا ومُتجانسًا، لا يتأثَّرُ بأهواء صُنَّاعه وأفكارهم الشَّخصيَّة الَّتي قد تكونُ مُغايرةً لمفهوم الصِّناعة، وإن تعاقبت عليه أجيالُ المُحرِّرين.

وتقترحُ الدِّراسةُ أن يُستفادَ في تحرير مادَّة المُعجم التَّاريخيّ المنشود بنظام التَّحرير *TshwaneLex* مُغلَق المصدر، الَّذي أنتَجَتْهُ شركة تقنيات اللُّغات الإنسانيَّة *TshwaneDJe* ضمنَ حزمة برمجيَّاتٍ لمُعالجة النُّصُوص (٦٥). ومع ما لهذا النِّظام من كفاءة وقُدرة على التَّعامُل مع النُّصُوص العربيَّة، إلَّا أنَّهُ يتطلَّبُ التَّدخُّلَ اليدويَّ لاستدراك أخطاء الآلة ومُراجعة المادَّة المُخرَجة. وهو في ذلك مثلُ غيره من أنظمة التَّحرير المُعجميّ الَّتي لا تدعمُ العربيَّةَ بصورةٍ كاملةٍ، نظرًا لعدم شُيُوع استخدامها في الصِّناعة المُعجميَّة العربيَّة – حتَّى الآن. ويُوضِّحُ (الشَّكل ٥) نموذجًا لهذا النِّظام.



الشَّكل ٥: نموذج توضيحيّ لنظام التَّحرير المُعجميّ – مُغلَق المصدر *TshwaneLex – v.7.1*

٤. توظيف الموارد في صناعة المُعجم التَّاريخيّ للُّغة العربيَّة.

أكثرُ الموارد الَّتي عَرَضَت لها الدِّراسةُ – آنفًا – لم يتوافر مثلُها لأصحاب التَّجارب السَّابقة في صناعة المُعجمات التَّاريخيَّة، نظرًا لارتباط وُجُودها بتقنيات الحاسُوب والمُجتَمَع المعلوماتيّ. وقد كانَ للصِّناعة المُعجميَّة اليدويَّة التَّقليديَّة أثرٌ بالغٌ على هذه التَّجارب؛ إذ استغرقَت كُلٌّ منها عشرات السِّنين، وتعاقبَت عليها أجيالٌ عديدة (٦٦). نُلاحظُ – مثلًا – أنَّ التَّجربةَ الألمانيَّةَ الرَّائدة قد استغرَقَت أكثرَ من مئةٍ وعشرينَ عامًا؛ ونُلاحظُ – أيضًا – أنَّ مادَّةَ هذا المُعجم تمتدُّ لقريب من خمسمائة عامٍ؛ حيثُ تبدأ بمجموعة المقالات الَّتي وضَعَها المُصلِح الكنَسِيّ مارتن لوثر *Martin Luther* (١٤٨٣-١٥٤٦). وقياسًا على هذه التَّجربة، فإنَّ صناعةَ مُعجمٍ تاريخيٍّ للُّغة العربيَّة – على نفس النَّهج – تستغرقُ أربعمئة عامٍ – على أقلّ تقدير؛ حيثُ يمتدُّ تاريخُ اللُّغة العربيَّة المُدَوَّن إلى حُدُود ألفَي عام. ورُبَّما كانَ من حظِّ العربيَّة أن يتأخَّرَ إنجازُ مُعجَمها التَّاريخيّ إلى ذلكَ الوقت الَّذي أمكنَ فيه توجيهُ الآلة إلى فهم اللُّغات الطَّبيعيَّة ومُحاكاة الذَّكاء البشريّ.

ولأنَّ إنجازَ المُعجم التَّاريخيّ المنشُود للعربيَّة يتطلَّبُ قدرًا كبيرًا من المُرُونة، فعلينا أن نسعى إلى الإفادة من الموارد المُتاحة فعليًّا، بعدَ تجربتها والتَّأكُّد من مدى فاعليَّتها. وبالنَّظر إلى الموارد اللُّغَويَّة والمُعجميَّة والحاسوبيَّة الَّتي يُمكنُ الإفادةُ منها في استخلاص مادَّة المُعجم وتحريرها، نجدُ أنَّ توظيفَ بعضِها أمرٌ حَتميٌّ لا يُمكنُ الاستغناءُ عنه أو استبداله بغيره، كونه عمادَ الصِّناعة المُعجميَّة؛ ونجدُ أنَّ بإمكاننا البحثَ عن بدائلَ عن بعضها الآخر الَّذي يُستفادُ منه في توفير الوقت والجُهد أو زيادة جودة المادَّة المُعجميَّة. وعليه، يُمكنُ تقسيمُ موارد المُعجم إلى قسمَين رئيسَين، على النَّحو الوارد في (الجدول ٥).

		المورد	التَّصنيف	التَّوصيف
الموارد الأساسيَّة	١	المُدَوَّنة اللُّغَويَّة	مَورد لُغَويّ	المورد الرَّئيس للمادَّة المُعجميَّة
	٢	المُعجم التَّأثيليّ للعربيَّة	مَورد مُعجميّ	المورد الرَّئيس للمعلومات التَّأثيليَّة
	٣	قاعدة البيانات المُعجميَّة	مَورد مُعجميّ	المورد الرَّئيس للمعاني المُعجميَّة عبرَ العُصُور
	٤	نظام الفهرسة الآليَّة	أداة / مَورد حاسُوبيّ	وسيلة إحصاء المُفردات واستخلاص السِّياقات
الموارد الثَّانويَّة	٥	قاعدة البيانات الجذريَّة	مَورد لُغَويّ	موردٌ مُساعدٌ في تعيين المداخل والوحدات
	٦	شبكة الكلمات العربيَّة	مَورد لُغَويّ	موردٌ مُساعدٌ في تعيين المعاني المُعجميَّة
	٧	نظام التَّحليل الصَّرفيّ	أداة / مَورد حاسُوبيّ	موردٌ مُساعدٌ في تعيين المعاني الوظيفيَّة البنويَّة
	٨	نظام التَّحليل التَّركيبيّ	أداة / مَورد حاسُوبيّ	موردٌ مُساعدٌ في تعيين المعاني الوظيفيَّة التَّركيبيَّة
	٩	نظام التَّحليل الدِّلاليّ	أداة / مَورد حاسُوبيّ	موردٌ مُساعدٌ في تعيين المعاني المُعجميَّة
	١٠	المِنصَّة الحاسُوبيَّة	أداة / مَورد حاسُوبيّ	أداةٌ لاستخلاص الشَّواهد ومُستويات الاستعمال
	١١	نظام التَّحرير المُعجميّ	أداة / مَورد حاسُوبيّ	وسيلة مُساعدة في التَّحرير المُعجميّ

الجدول ٥: الموارد الأساسيَّة والثَّانويَّة للمُعجم التَّاريخيّ

وفي ضوء هذه الموارد يُمكنُ إعدادُ منهجيَّةٍ واضحة المعالم لصناعة المُعجم التَّاريخيّ المنشود للعربيَّة. وبعبارةٍ أخرى، نستطيعُ توظيفَ هذه الموارد جميعًا في مراحل الصِّناعة المُعجميَّة – جمعًا وتحريرًا ونشرًا – على التَّوالي والتَّوازي. ووفقًا لمنهجيَّة الصِّناعة، ستكونُ مُحصِّلةُ المرحلة الأولى [مرحلة الجمع] أن نستخلصَ المعلومات المُعجميَّة على مُستوى المباني والمعاني؛ وستكونُ مُحصِّلةُ المرحلة الثَّانية [مرحلة التَّحرير] أن نستخلصَ المادَّةَ المُعجميَّةَ المُحرَّرة. وأخيرًا، ستكونُ مُحصِّلةُ المرحلة الثَّالثة [مرحلة النَّشر] أن نصلَ إلى الهدف المنشود بإنجاز المُعجم التَّاريخيّ. ويُوضِّحُ (الجدول ٦) هيكلًا مُقترَحًا لمراحل صناعة المُعجم التَّاريخيّ للُّغة العربيَّة في ضوء الموارد الَّتي عَرَضَت لها الدِّراسة.

مراحل الصِّناعة المُعجميَّة المُتوالية			
مراحل الصِّناعة المُعجميَّة المُتوازية	المرحلة ١: الجمع		
	بناء المُدوَّنة اللُّغويَّة	رَقمنة النُّصُوص	بناء المُعجم التَّأثيليّ
		تنقية النُّصُوص	
		تحرير النُّصُوص	
		التَّدقيق الإملائيّ	
		التَّشكيل الآليّ	
		التَّرميز	
	مُعالجة المُدوَّنة اللُّغويَّة	الفهرسة الآليَّة	استخلاص المعلومات التَّأثيليَّة
		التَّحليل الصَّرفيّ / التَّحليل التَّركيبيّ	
		بناء قاعدة البيانات الجذريَّة	
		استخلاص معلومات المباني	
		إعداد شبكة الكلمات / التَّحليل الدِّلاليّ	
		استخلاص معلومات المعاني	
	المرحلة ٢: التَّحرير		
	تحرير المبنى	تحرير المداخل والوحدات المُعجميَّة *"Entries & Headwords"*	
		تحرير معلومات الهجاء والنُّطق *"Alphabet & Pronunciation"*	
		تحرير المعاني الوظيفيَّة *"Grammatical Functions"*	
	تحرير المعنى	تحرير المعلومات التَّأثيليَّة *"Etymological Information"*	
		تحرير المعاني المُعجميَّة *"Lexical Meanings"*	
		تحرير الشَّواهد المُعجميَّة ومُستويات الاستعمال *"Lexical Evidence & Usage Levels"*	
	المرحلة ٣: النَّشر		
	النَّشر الورقيّ	تنسيق المادَّة المُعجميَّة	
		تمييز المادَّة المُعجميَّة بالخُطُوط الكتابيَّة	
		إعداد الرُّسُوم التَّوضيحيَّة أحاديَّة وثنائيَّة الأبعاد (من وسائل شرح المعنى)	
		تجربة الطِّباعة الورقيَّة	
	النَّشر الرَّقميّ	إعداد الهيكل الحاسُوبيّ للمُعجم	
		إعداد الرُّسُوم التَّوضيحيَّة أحاديَّة وثنائيَّة الأبعاد / إعداد الرُّسُوم التَّوضيحيَّة ثُلاثيَّة الأبعاد	
		إعداد الصَّوتيَّات	
		إعداد المرئيَّات	
المُعجم التَّاريخيّ للُّغة العربيَّة			

الجدول ٦: هيكل مُقترَح لمراحل صناعة المُعجم التَّاريخيّ للُّغة العربيَّة

٥. نتائج الدِّراسة.

١. إنَّ إنجازَ المعجَم التَّاريخيّ للُّغة العربيَّة ليسَ عملاً إبداعيًّا يعكُفُ عليه المُؤلِّفُونَ والمُفكِّرون؛ وإنَّما هو صناعةٌ، ينبغي أن تُوكَلَ إلى أهلِها العارفِينَ بدقائِقِها وضوابطِها.

٢. يتطلَّبُ إنجازُ المُعجم التَّاريخيّ للُّغة العربيَّة إعدادَ الموارد الَّتي يَستمدُّ منها المُعجمُ مادَّتَهُ أو الَّتي يُمكنُ من خلالها تيسيرُ العمل فيه على الوجه المنشود. وثَمَّةَ بعضُ الإشكالات الَّتي تُواجهُ إعدادَ موارد المُعجم التَّاريخيّ للعربيَّة؛ منها: نُدرة موارد التَّأثيل المُعجميّ للعربيَّة وافتقار المكتبة العربيَّة إلى مُعجمٍ تأثيليٍّ لمُفرداتها، وحداثة عهد العربيَّة بالموارد اللُّغَويَّة – بصورةٍ عامَّة، وحداثة عهد العربيَّة باللِّسانيَّات الحاسُوبيَّة.

٣. قامَت منهجيَّةُ هذه الدِّراسة على تقسيم موارد المُعجم التَّاريخيّ المنشود إلى ثلاثة أقسام، هي: الموارد اللُّغويَّة؛ وتشتملُ على: المُدَوَّنة اللُّغَويَّة، وقاعدة البيانات الجذريَّة، وشبكة الكلمات العربيَّة؛ والموارد المعجميَّة؛ وتشتملُ على: المُعجم التَّأثيليّ للعربيَّة، وقاعدة البيانات المُعجميَّة؛ والموارد الحاسوبيَّة؛ ونعني بها أدوات المُعالجة والتَّحرير؛ وتشتملُ على: نظام الفهرسة الآليَّة، ونظام التَّحليل الصَّرفيّ، ونظام التَّحليل التَّركيبيّ، ونظام التَّحليل الدِّلاليّ، والمِنَصَّة الحاسُوبيَّة، ونظام التَّحرير المُعجميّ.

٤. عَرَضَت الدِّراسةُ توصيفًا لكلِّ موردٍ على حِدةٍ، كما أبانَت عن وُجود تجاربَ مُنجَزةٍ لبعض الموارد المُستَهدَفة (مثل: المُدَوَّنة اللُّغَويَّة، والمُفهرس الآليّ، والمُحلِّل الصَّرفيّ)؛ وأبانت الدِّراسةُ عن قُصُورٍ في بعض الموارد الأخرى (مثل: شبكة الكلمات العربيَّة، ونظام التَّحليل التَّركيبيّ، ونظام التَّحليل الدِّلاليّ)؛ وأبانت – أيضًا – عن افتقار مُجتمَع المَعرفة العربيّ إلى بعض الموارد (مثل: المُعجم التَّأثيليّ للعربيَّة).

٥. سَعى الباحثُ إلى استقراء الموارد المُتاحة فعليًّا، وأبانَ عن كيفيَّة الاستفادة منها، كما اقترحَ حُلُولًا لمُعالجة القُصُور الواقع في الموارد غير المُكتملة، واقترَحَ البدائلَ الَّتي رآها مُناسبةً للموارد الَّتي يفتقرُ إليها المُجتمَعُ المعرفيُّ العربيّ، بعد استقرائها والمُفاضَلة بينها.

٦. اقترَحت الدِّراسةُ توظيفَ موارد المُعجم المنشود بتصنيفها إلى قسمين رئيسَين؛ بحيثُ يضُمُّ القسمُ الأوَّلُ الموارِدَ الأساسيَّةَ الَّتي تُستمدُّ منها مادَّةُ المُعجم بصورةٍ مُباشرة، ويضُمُّ القسمُ الآخَرُ الموارِدَ الثَّانويَّة الَّتي يُستفادُ منها في استخلاص مادَّة المُعجم أو تيسير إجراءات الصِّناعة المُعجميَّة. واقترَحت الدِّراسةُ – أخيرًا – هيكلًا لمراحل صناعة المُعجم التَّاريخيّ المنشود، في ضوء الموارد الَّتي عَرَضَت لها.

٦. الخُلاصة.

بدأ التَّفكيرُ في بناء مُعجمٍ تاريخيّ للعربيَّة مُنذُ أكثرَ من مئةِ عامٍ؛ ولم يُكتَب لأيٍّ من مُحاولات إنجازه النَّجاحُ إلى الآن. ولعلَّ السَّببَ الرَّئيسَ في ذلك أنَّ المُحاولاتِ السَّابقةَ لم تَسْعَ إلى توفير الموارد الَّتي يُستفادُ منها في بناء المُعجم المنشود قبلَ البدء في إعداد منهجيَّةِ صناعته. والواقعُ أنَّ ثَمَّةَ إشكالاتٍ تُواجهُ إعدادَ موارد المُعجم المنشود؛ منها: افتقار المكتبة العربيَّة إلى مُعجمٍ تأثيليٍّ لمُفرداتها، وحداثةُ عهد العربيَّة بالموارد اللُّغَويَّة، وحداثة عهد العربيَّة باللِّسانيَّات الحاسُوبيَّة. وانطلاقًا من واقع اللُّغة العربيَّة والتَّجارب النَّاجحة في صناعةِ مُعجماتٍ تاريخيَّةٍ للُّغات الأخرى، تقترحُ الدِّراسةُ منهجيَّةً لإعداد موارد المُعجم التَّاريخيّ المنشود (اللُّغَويَّة، والمُعجميَّة، والحاسُوبيَّة). وقد أبانَ الباحثُ – من خلال منهجيَّة الدِّراسة – عن الموارد المُتاحة والبدائل الَّتي يُمكنُ الإفادةُ منها للموارد غير المُتاحة، كما سَعَى إلى وضعِ آليَّةٍ لتوظيف موارد المُعجم (الأساسيَّة والثَّانويَّة)، ثُمَّ قَدَّمَ هيكلًا مُقترَحًا لمراحل صناعة المُعجم التَّاريخيّ المنشود في ضوء الموارد الَّتي تقترحُها الدِّراسة.

هوامش الدِّراسة

(١) للمزيد حولَ مفهوم المُعجم التَّاريخيّ ونشأته، راجع: السَّعيد (المُعتزّ بالله): مُدَوَّنة مُعجَم تاريخيّ للُّغة العربيَّة "مُعالَجة لُغَوِيَّة حاسوبيَّة"، أطروحة دكتوراه، جامعة القاهرة، ٢٠١١م، ص ٢ : ٤.

(٢) للمزيد حولَ المُدوَّنات اللُّغويَّة وتوظيفها في صناعة المُعجم، راجع: السَّعيد (المُعتزّ بالله): المدوَّنات اللُّغَوِيَّة، ضِمنَ كتاب (مُقَدِّمة في حوسَبة اللُّغة العربيَّة)، مجموعة من المؤلِّفين، تحرير: مُحسن رَشوان، والمعتزّ بالله السَّعيد، قَيد النَّشر بمدينة الملك عبد العزيز للعُلُوم والتِّقنية، الرِّياض، الجزء الأوَّل، ٢٠١٤م.

(٣) طُبِعَ المُعجمُ في طبعاتٍ مُتتالية، آخرُها في عام ٢٠٠٧. راجع:

Johnson, S. (2007). A Dictionary of the English Language. Reprint. PLB.

(٤) للمُعجم طبعةٌ حديثةٌ في عام ٢٠١٢. راجع:

Wedgwood, H. (2012). A Dictionary of English Etymology. Reprint. Nabu Press.

(٥) للمُعجم طبعةٌ حديثةٌ في عام ٢٠١٣. راجع:

Körting, G. (2013). Lateinisch-Romanisches Worterbuch (Etymologisches Worterbuch Der Romanischen Hauptsprachen). Reprint. BiblioBazaar.

(٦) مَرسُوم بإنشاء مَجمَع مَلَكِيّ للُّغَة العَرَبيَّة، مجلَّة مجمع اللُّغة العربيَّة الملكِيّ، الجزء الأوَّل، القاهرة، ١٩٣٥م، ص ٧.

(٧) حول مُعجم فيشر ومراحِل العَمل فيه، راجِع: فيشر (أوجُست): المُعجَم اللُّغَوِيّ التَّارِيخِيّ، القِسم الأوَّل، مجمع اللُّغة العَرَبيَّة، القاهِرة، ١٩٦٧م، ص ٢٩ : ٣٤؛ وراجع: بدوي (عبد الرَّحمن): موسوعة المُستَشرِقين، دار العِلم للملايين، بيروت، ط٣، ١٩٩٣م، ص ٤٠٥-٤٠٦؛ وراجِع أيضًا: المُنجد (صلاح الدِّين): المستشرقون الألمان: تراجمهم وما أسهموا به في الدِّراسات العَرَبيَّة، ج١، دار الكِتاب الجديد، بيروت، ١٩٧٨م، ص ١٣٥. وراجع حولَ قصَّة المُعجم ومنهجه: السَّعيد (المُعتزّ بالله): مُدَوَّنة مُعجَم تاريخيّ للُّغة العربيَّة، ص ٢٢ : ٢٥.

(٨) حولَ تفصيل هذه المُحاولات، راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص ٢٦ : ٣١.

(٩) شاركَ الباحثُ في الاجتماعات الثَّلاثة المُمَهِّدة للمشروع، ثُمَّ عملَ عُضوًا بالهيئة التَّنفيذيَّة له. وقامَ – من خلال عمله – ببناء وتطوير ومُراجعة المُدَوَّنة اللُّغَويَّة للمرحلة الأولى، منذ القرن الخامس قبلَ الهجرة إلى نهاية القرن الثَّاني الهجريّ، كما قامَ بإعداد ببليوغرافيا مصادر هذه المرحلة بمُشاركة الدُّكتور عبد العلي الودغيري.

(١٠) راجع حولَ مشروع بناء المُعجم التَّأثيليّ للُّغة والثَّقافة العربيَّة:

Scientific Report from the ESF Exploratory Workshop Breaking the Grounds for an Arabic Etymological Dictionary (Oslo, 22-23 June, 2013).

وراجع أيضًا:

Mikaili, P. (2013). An Etymological Dictionary of Arabic. ESF Exploratory Workshop Breaking the Grounds for an Arabic Etymological Dictionary (Oslo, 22-23 June, 2013).

(١١) الموقع الإلكترونيّ للمُعجم:

http://wba.ruhosting.nl/index.htm.

(١٢) الموقع الإلكترونيّ لشبكة الكلمات العربيَّة:

http://globalwordnet.org/arabic-wordnet/.

(١٣) حولَ مشروع شبكة الكلمات العربيَّة *"AWN"* ومنهجيَّة العمل، راجع:

Black, W., Elkateb, S., Rodriguez, H, Alkhalifa, M., Vossen, P., Pease, A. and Fellbaum, C., (2006). Introducing the Arabic WordNet Project, in Proceedings of the 3rd International WordNet Conference, Sojka, Choi, Fellbaum and Vossen eds.

(١٤) حولَ إشكالات حوسبة المُعجم التَّاريخيّ للعربيَّة، راجع: السَّعيد (المُعتزّ بالله): حَوسَبة المُعجم التَّاريخيّ للُّغة العربيَّة، مجلَّة اللِّسان العربيّ، قَيد النَّشر بالعدد ٧٤، مكتب تنسيق التَّعريب بالرِّباط، ٢٠١٤.

(١٥) حولَ منهجيَّة المُعجم التَّكراريّ للُّغة العربيَّة، راجع:

Buckwalter, T., Parkinson, D. (2011). A Frequency Dictionary of Arabic: Core Vocabulary for Learners. Routledge.

(١٦) صفحة المُعجم على الموقع الإلكترونيّ للأَكاديميَّة المَلَكيَّة الإسبانيَّة:

http://www.rae.es/obras-academicas/diccionarios/nuevo-diccionario-historico-del-espanol.

(١٧) صفحة المُعجم على الموقع الإلكترونيّ لمجمع اللُّغة العبريَّة:

http://hebrew-academy.huji.ac.il/English/HistoricalDictionaryProject/.

(١٨) راجع: السَّعيد (المُعتزّ بالله): مُدَوَّنة مُعجَم تاريخيّ للُّغة العربيَّة، ص ٣٤، بتصرُّف.

(١٩) حولَ منهج بناء مُدَوَّنة المُعجم التَّاريخيّ للُّغة العربيَّة في إصدارها الأَوَّل، راجع: السَّعيد (المُعتزّ بالله): مُدَوَّنة مُعجَم تاريخيّ للُّغة العربيَّة، ص ٣٣ : ٥٢.

(٢٠) راجع: السَّعيد (المُعتزّ بالله): شبكة الكلمات العَرَبيَّة "المنهجِيَّة والتَّطبيق"، مجلَّة مجمَع اللُّغة العَرَبيَّة بدمشق، المُجلَّد ٨٧، الجُزء الثَّاني، ٢٠١٢، بتصرُّف.

(٢١) الشَّبكة مُتاحةٌ على موقع البرمجيَّات الحرَّة *sourceforge*؛ ورابطها:

http://sourceforge.net/projects/awnbrowser/.

(٢٢) للكتاب طبعةٌ حديثةٌ في عام ٢٠١٣. راجع:

Cooke, G. A. (2013). A text-book of North-Semitic inscriptions: Moabite, Hebrew, Phoenician, Aramaic, Nabataean, Palmyrene, Jewish. ATLA monograph preservation program. Reprint. HardPress Publishing.

(٢٣) راجع:

Hoftijzer, J., Jongeling, K., Steiner, R.C., Porten, B., Moshavi, A. M., (2003). Dictionary of the North-West Semitic Inscriptions. Handbook of Oriental Studies. Part I. Brill Academic Pub.

(٢٤) راجع:

Fück, J. W. (1950). Arabiya,: Untersuchungen zur arabischen Sprach- und Stilgeschichte. Abhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig, Philologisch-Historische Klasse. Akademie-Verlag.

(٢٥) راجع:

Rabin, C. (1951). Ancient West-Arabian. Taylor's Foreign Press.

(٢٦) راجع: أنيس (إبراهيم): في اللَّهجات العربيَّة، مكتبة الأنجلو المصريَّة، ٢٠١٠.

(٢٧) راجع: السَّامرَّائي (إبراهيم): في اللَّهجات العربيَّة القديمة، دار الحداثة، ٢٠٠٧.

(٢٨) للمُعجم طبعةٌ حديثةٌ في عام ٢٠١٠. راجع:

Gesenius, W. (2010). Hebräisch-deutsches Handwörterbuch über die Schriften des Alten Testaments: mit Einschluss der geographischen Nahmen und der chaldäischen Wörter beym Daniel und Esra. Reprint. Nabu Press.

(٢٩) للمُعجم طبعةٌ حديثةٌ في عام ٢٠١٣. راجع:

Gesenius, W., Meyer, R. D., (2013). Hebräisches und Aramäisches Handwörterbuch über das Alte Testament. Reprint. Springer.

(٣٠) للمُعجم طبعةٌ حديثةٌ في عام ٢٠١٢. راجع:

Smith, W. (2012). Smith's Bible Dictionary. Reprint. Barbour Publishing, Inc.

(٣١) راجع:

Möller, H. (1970). Vergleichendes indogermanisch-semitisches Wörterbuch. Raabe,samtliche Werke. Vandenhoeck & Ruprecht.

(٣٢) راجع:

Brunner, L. (1969). Die gemeinsamen Wurzeln des semitischen und indogermanischen Wortschatzes. Francke.

(٣٣) راجع:

Cohen, D., Bron, F., Lonnet, A., Cantineau, J., (2012). Dictionnaire des racines sémitiques ou attestées dans les langues sémitiques. III. Fascicule 10, H, T-T', Ŧ, Y, K. Paris; Louvain: Peeters, DL 2012.

(٣٤) راجع: الجواليقي (أبو منصور): المُعَرَّب من الكلام الأعجميّ على حُرُوف المُعجَم، تحقيق: أحمد شاكر، مطبعة دار الكُتُب، ط٢، ١٩٦٩.

(٣٥) للمُعجم طبعةٌ حديثةٌ في عام ٢٠٠٦. راجع:

Dozy, R. (2006). Supplement Aux Dictionnaires Arabes. Gorgias Historical Dictionaries. Reprint. Gorgias PressLlc.

(٣٦) راجع: الكلداني (آدي شير): الألفاظ الفارِسِيَّة المُعَرَّبة، المطبعة الكاثوليكيَّة، بيروت، ١٩٠٨.

(٣٧) راجع: اليَسُوعيّ (روفائيل نخلة): غرائب اللُّغة العَرَبيَّة، دار المشرق، ١٩٩٦.

(٣٨) راجع: عبد الرَّحيم (فانيامبادي): مُعجم الدَّخيل في اللُّغة العربيَّة ولهجاتها، دار القلم، ٢٠١١.

(٣٩) من وُجهة نظر الباحث، تمتدُّ اللُّغةُ العربيَّة القديمة من أقدم نصٍّ مُدَوَّنٍ لها إلى نهاية عصر الرِّواية في حُدُود عام ٢٠٠ هـ؛ وتضُمُّ عربيَّة العصور الثَّلاثة (الجاهليّ والإسلاميّ والأمويّ) وشطرًا من العصر العبَّاسيّ. وتمتدُّ العربيَّةُ الوسيطة منذُ بداية عصر التَّدوين إلى ما قبلَ عصر النَّهضة (في حُدُود عام ١٨٠٤)؛ وتضُمُّ القدرَ الأَكبرَ من عربيَّة العصر العبَّاسيّ وما يليه من عصر الدُّوَل والدُّويلات. وتمتدُّ اللُّغةُ العربيَّةُ الحديثةُ منذ بداية عصر النَّهضة إبَّانَ حُكم مُحمَّد علي باشا (١٧٦٩-١٨٤٩) إلى الآن؛ وتضُمُّ عربيَّةَ العصر الحديث، بما في ذلكَ اللُّغة العربيَّة المُعاصرة.

(٤٠) راجع: الفَراهِيدِيّ (الخَليل بن أحمد): كتاب العَين، تحقيق: داود سلوم، داود العنبكي، إنعام داود سلوم، لبنان ناشرون، ٢٠٠٤.

(٤١) راجع: الأزهريّ (أبو منصور): تهذيب اللُّغة، حَقَّقَهُ: عبد السَّلام هارون وآخرون، دار المعارف، القاهرة، ١٩٧٦.

(٤٢) راجع:

Ullmann, M. [Edetor]. (2009). Wörterbuch der klassischen arabischen Sprache. 2,4. Lawā bis līwān. Deutsche Morgenländische Gesellschaft. Otto Harrassowitz Verlag.

(٤٣) راجع: الفيروزابادي (مُحمَّد بن يعقُوب): القاموس المُحيط، تحقيق: مُحمَّد نعيم العرقسوسي، مُؤسَّسة الرِّسالة، ٢٠٠٥.

(٤٤) راجع: الزَّبِيدِيّ (السَّيِّد مُرتضى): تاج العَرُوس مِن جَواهِر القامُوس، تحقيق: مجموعة من المُحقِّقين، مطبعة حُكومة الكويت والمجلس الوطنيّ للثَّقافة والفُنُون والآداب، الكويت، من ١٩٦٥ إلى ٢٠٠٢.

(٤٥) للمُعجم طبعةٌ حديثةٌ في عام ٢٠١٠. راجع:

Lane, E. W. (2010). An Arabic-English Lexicon: Derived from the Best and the Most Copious Eastern Sources. Reprint. Nabu Press.

(٤٦) للمُعجم طبعةٌ حديثةٌ في عام ٢٠١١. راجع:

Wehr, H. (2011). Arabisches Wörterbuch für die Schriftsprache der Gegenwart: Arabisch-Deutsch. Edited by: Lorenz Kropfitsch. Otto Harrassowitz Verlag.

(٤٧) راجع: عُمَر (أحمد مُختار): مُعجم اللُّغة العربيَّة المُعاصرة، عالَم الكُتُب، القاهرة، ٢٠٠٨.

(٤٨) راجع: مَجمع اللُّغة العَرَبِيَّة: المُعجَم الكبير، القسم الأوَّل من حرف الرَّاء، مطبوعات مجمع اللُّغة العربيَّة، القاهرة، ط١، ٢٠١٢. (وهُو آخرُ ما طُبِعَ من المُعجَم حتَّى الآن).

(٤٩) للمزيد حولَ نظام الفهرسة الآليَّة ودوره في صناعة المُعجم، راجع: السَّعيد (المُعتزّ بالله): حَوسَبة المُعجم التَّاريخيّ للُّغة العربيَّة، مرجع سابق.

(٥٠) صفحة البرمجيَّة على الموقع الإلكترونيّ لشركة *TshwaneDJe*:

http://tshwanedje.com/corpus/.

(٥١) للمزيد حولَ آليَّة التَّحليل الصَّرفيّ، راجع: السَّعيد (المُعتزّ بالله): مُدَوَّنة مُعجَم تاريخيّ للُّغة العربيَّة "مُعالَجة لُغَوِيَّة حاسوبيَّة"، ص ٨٤ : ٩٤.

(٥٢) البرمجيَّة مُتاحةٌ على موقع البرمجيَّات الحرَّة *Sourceforge*؛ ورابطها:

http://sourceforge.net/projects/aramorph.

(٥٣) البرمجيَّة مُتاحةٌ على موقع البرمجيَّات الحرَّة *Sourceforge*؛ ورابطها:

http://sourceforge.net/projects/alkhalil.

(٥٤) صفحة المشروع على الموقع الإلكترونيّ لشركة *Microsoft*:

http://research.microsoft.com/en-us/projects/sarf.

(٥٥) (الجدول ٢) مُستمدٌّ من صفحة المشروع على موقع شركة *Microsoft*.

(٥٦) راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص ٧٨، بتصرُّف.

(٥٧) البرمجيَّة مُتاحةٌ على موقع جامعة *Stanford*؛ ورابطها:

http://nlp.stanford.edu/software/tagger.shtml.

(٥٨) صفحة البرمجيَّة على الموقع الإلكترونيّ لشركة *RDI*:

http://www.rdi-eg.com/technologies/POS.aspx.

(٥٩) راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص ١٠٢، بتصرُّف.

(٦٠) حولَ تطبيق مُصَنِّف بايز المُبَسَّط في التَّحليل الدِّلاليّ للنُّصُوص العربيَّة، راجع:

Eid, S. M. & Al-Said, A. & Wanas, N. M. & Rashwan, M. A. & Hegazy, N. H. (2010). A Comparative Study of Rocchio Classifier Applied to supervised WSD Using Arabic Lexical Samples. 10th Conference on Language Engineering. Ain Shams University, Cairo, Egypt.

وراجع أيضًا:

El-Gamml, M. & Fakhr, M. & Rashwan. M. & Al-Said, A. (2011). A Comparative Study for Arabic Word Sense Disambiguation using document preprocessing and machine learning techniques. Arabic Language Technology International Conference (ALTIC), Bibliotheca Alexandrina, Alex, Egypt.

(٦١) للمزيد حولَ المِنصَّة الحاسُوبيَّة ودورها في صناعة المُعجم، راجع: السَّعيد (المُعتزّ بالله): حَوسَبة المُعجم التَّاريخيّ للُّغة العربيَّة، مرجع سابق.

(٦٢) حَولَ مِنَصَّة *Nooj* ومنهجيَّة عملها، راجع:

Silberztein, M. (2004). NooJ : an Object-Oriented Approach. In INTEX pour la Linguistique et le Traitement Automatique des Langues, C. Muller, J. Royauté M. Silberztein Eds, Cahiers de la MSH Ledoux. Presses Universitaires de Franche-Comté, pp. 359-369.

(٦٣) الموقع الإلكترونيّ لِمنصَّة *Nooj*:

http://www.nooj4nlp.net.

(٦٤) حولَ أنظمة التَّحرير المُعجميّ وآليَّات عملها، راجع:

Abel, A. (2012). Dictionary writing systems and beyond. In: Granger, S. & Paquot, M. (Ed.). Electronic Lexicography. Oxford University Press, 2012. Pp.83-106.

(٦٥) صفحة البرمجيَّة على الموقع الإلكترونيّ لشركة *TshwaneDJe*:

http://tshwanedje.com/tshwanelex/.

(٦٦) حولَ مُستقبل صناعة المُعجمات اللُّغَويَّة التَّاريخيَّة، راجع:

Brewer, C. (2013). The future of historical dictionaries. In: Jackson, H. (Ed.). The Bloomsbury companion to lexicography. A&C Black, 2013. Pp.341-354.

مراجع الدِّراسة

أوَّلًا: المراجع العربيَّة

الكُتُب

- أنيس (إبراهيم): في اللَّهجات العربيَّة، مكتبة الأنجلو المصريَّة، ٢٠١٠.
- بدوي (عبد الرَّحمن): موسوعة المُستَشرِقين، دار العِلم للملايين، بيروت، ط٣، ١٩٩٣م.
- السَّامرَّائي (إبراهيم): في اللَّهجات العربيَّة القديمة، دار الحداثة، ٢٠٠٧.
- السَّعيد (المُعتزّ بالله): المدوَّنات اللُّغَوِيَّة، ضِمنَ كتاب (مُقَدِّمة في حوسَبة اللُّغة العربيَّة)، مجموعة من المؤلِّفين، تحرير: مُحسن رَشوان والمعتزّ بالله السَّعيد، قَيد النَّشر بمدينة الملك عبد العزيز للعُلُوم والتِّقنية، الرِّياض، ٢٠١٤.
- فيشر (أوجُست): المُعجَم اللُّغَوِيّ التَّارِيخِيّ، القِسم الأوَّل، مجمع اللُّغة العَرَبِيَّة، القاهِرة، ١٩٦٧م.
- المُنجد (صلاح الدِّين): المستشرقون الألمان: تراجمهم وما أسهموا به في الدِّراسات العَرَبِيَّة، ج١، دار الكِتاب الجديد، بيروت، ١٩٧٨م.
- اليَسُوعيّ (روفائيل نخلة): غرائب اللُّغة العَرَبيَّة، دار المشرق، ١٩٩٦.

المعاجم

- الأزهريّ (أبو منصور): تهذيب اللُّغة، حَقَّقَهُ: عبد السَّلام هارون وآخرون، دار المعارف، القاهرة، ١٩٧٦.
- الجواليقي (أبو منصور): المُعَرَّب من الكلام الأعجميّ على حُرُوف المُعجَم، تحقيق: أحمد شاكر، مطبعة دار الكُتُب، ط٢، ١٩٦٩.
- الزَّبيدِيّ (السَّيِّد مُرتضى): تاج العَرُوس مِن جَواهِر القامُوس، تحقيق: مجموعة من المُحقِّقين، مطبعة حُكومة الكويت والمجلس الوطنيّ للثَّقافة والفُنُون والآداب، الكويت، من ١٩٦٥ إلى ٢٠٠٢.
- عبد الرَّحيم (فانيامبادي): مُعجم الدَّخيل في اللُّغة العربيَّة ولهجاتها، دار القلم، ٢٠١١.
- عُمَر (أحمد مُختار): مُعجم اللُّغة العربيَّة المُعاصرة، عالَم الكُتُب، القاهرة، ٢٠٠٨.
- الفَراهِيدِيّ (الخَلِيل بن أحمد): العَين، حَقَّقَهُ: داود سلوم وآخَرُون، لبنان ناشرون، ٢٠٠٤.
- الفيروزابادي (مُحمَّد بن يعقُوب): القاموس المُحيط، تحقيق: مُحمَّد نعيم العرقسوسي، مُؤسَّسة الرِّسالة، ٢٠٠٥.
- الكلداني (آدي شير): الألفاظ الفارِسِيَّة المُعَرَّبة، المطبعة الكاثوليكيَّة، بيروت، ١٩٠٨.
- مَجمع اللُّغة العَرَبيَّة: المُعجَم الكبير، القسم الأوَّل من حرف الرَّاء، مطبوعات مجمع اللُّغة العربيَّة، القاهرة، ٢٠١٢.

الأطروحات العلميَّة

- السَّعيد (المُعتزّ بالله): مُدَوَّنة مُعجَم تاريخيّ للُّغة العربيَّة "مُعالَجة لُغَوِيَّة حاسوبيَّة"، أطروحة دكتوراه، جامعة القاهرة، ٢٠١١م.

الأوراق البحثيَّة

- السَّعيد (المُعتزّ بالله): حَوسَبة المُعجم التَّاريخيّ للُّغة العربيَّة، مجلَّة اللِّسان العربيّ، قَيد النَّشر بالعدد ٧٤، مكتب تنسيق التَّعريب بالرِّباط، ٢٠١٤.
- السَّعيد (المُعتزّ بالله): شبكة الكلمات العَرَبيَّة "المنهجِيَّة والتَّطبيق"، مجلَّة مجمَع اللُّغة العَرَبيَّة بدمشق، المُجلَّد ٨٧، الجُزء الثَّاني، ٢٠١٢.

ثانيًا: المراجع الأجنبيَّة

الكُتُب

- *Abel, A. (2012). Dictionary writing systems and beyond. In: Granger, S. & Paquot, M. (Ed.). Electronic Lexicography. OUP, 2012.*
- *Brewer, C. (2013). The future of historical dictionaries. In: Jackson, H. (Ed.). The Bloomsbury companion to lexicography. A&C Black, 2013.*
- *Fück, J. W. (1950). Arabiya,: Untersuchungen zur arabischen Sprach- und Stilgeschichte. Abhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig. Akademie-Verlag.*
- *Rabin, C. (1951). Ancient West-Arabian. Taylor's Foreign Press.*
- *Silberztein, M. (2004). NooJ : an Object-Oriented Approach. In INTEX pour la Linguistique et le Traitement Automatique des Langues, C. Muller, J. Royauté M. Silberztein Eds, Cahiers de la MSH Ledoux. Presses Universitaires de Franche-Comté.*

المعاجم

- *Brunner, L. (1969). Die gemeinsamen Wurzeln des semitischen und indogermanischen Wortschatzes. Francke.*
- *Buckwalter, T., Parkinson, D. (2011). A Frequency Dictionary of Arabic: Core Vocabulary for Learners. Routledge.*
- *Cohen, D., Bron, F., Lonnet, A., Cantineau, J., (2012). Dictionnaire des racines sémitiques ou attestées dans les langues sémitiques. III. Fascicule 10, H, T-T', Ŧ, Y, K. Paris; Louvain: Peeters, DL 2012.*
- *Cooke, G. A. (2013). A text-book of North-Semitic inscriptions: Moabite, Hebrew, Phoenician, Aramaic, Nabataean, Palmyrene, Jewish. ATLA monograph preservation program. Reprint. HardPress Publishing.*
- *Dozy, R. (2006). Supplement Aux Dictionnaires Arabes. Gorgias Historical Dictionaries. Reprint. Gorgias PressLlc.*
- *Gesenius, W. (2010). Hebräisch-deutsches Handwörterbuch über die Schriften des Alten Testaments: mit Einschluss der geographischen Nahmen und der chaldäischen Wörter beym Daniel und Esra. Nabu Press.*
- *Gesenius, W., Meyer, R. D., (2013). Hebräisches und Aramäisches Handwörterbuch über das Alte Testament. Springer.*
- *Hoftijzer, J., Jongeling, K., Steiner, R.C., Porten, B., Moshavi, A. M., (2003). Dictionary of the North-West Semitic Inscriptions. Handbook of Oriental Studies. Part I. Brill Academic Pub.*
- *Johnson, S. (2007). A Dictionary of the English Language. PLB.*
- *Körting, G. (2013). Lateinisch-Romanisches Worterbuch (Etymologisches Worterbuch Der Romanischen Hauptsprachen). Reprint. BiblioBazaar.*
- *Lane, E. W. (2010). An Arabic-English Lexicon: Derived from the Best and the Most Copious Eastern Sources. Reprint. Nabu Press.*
- *Möller, H. (1970). Vergleichendes indogermanisch-semitisches Wörterbuch. Raabe,samtliche Werke. Vandenhoeck & Ruprecht.*

- *Smith, W. (2012). Smith's Bible Dictionary. Barbour Publishing, Inc.*
- *Ullmann, M. [Edetor]. (2009). Wörterbuch der klassischen arabischen Sprache. 2,4. Lawā bis līwān. Deutsche Morgenländische Gesellschaft. Otto Harrassowitz Verlag.*
- *Wedgwood, H. (2012). A Dictionary of English Etymology. Nabu Press.*
- *Wehr, H. (2011). Arabisches Wörterbuch für die Schriftsprache der Gegenwart: Arabisch-Deutsch. Edited by: Lorenz Kropfitsch. Reprint. Otto Harrassowitz Verlag.*

الأوراق البحثيّة

- *Black, W., Elkateb, S., Rodriguez, H, Alkhalifa, M., Vossen, P., Pease, A. and Fellbaum, C., (2006). Introducing the Arabic WordNet Project, in Proceedings of the 3rd International WordNet Conference, Sojka, Choi, Fellbaum and Vossen eds.*
- *Eid, S. M. & Al-Said, A. & Wanas, N. M. & Rashwan, M. A. & Hegazy, N. H. (2010). A Comparative Study of Rocchio Classifier Applied to supervised WSD Using Arabic Lexical Samples. 10th Conference on Language Engineering. Ain Shams University, Cairo, Egypt.*
- *El-Gamml, M. & Fakhr, M. & Rashwan. M. & Al-Said, A. (2011). A Comparative Study for Arabic Word Sense Disambiguation using document preprocessing and machine learning techniques. Arabic Language Technology International Conference (ALTIC), Bibliotheca Alexandrina, Alex, Egypt.*
- *Mikaili, P. (2013). An Etymological Dictionary of Arabic. ESF Exploratory Workshop Breaking the Grounds for an Arabic Etymological Dictionary (Oslo, 22-23 June, 2013).*

ثالثًا: المواقع الإلكترونيّة

- *http://globalwordnet.org/arabic-wordnet/.*
- *http://hebrew-academy.huji.ac.il/English/HistoricalDictionaryProject/.*
- *http://nlp.stanford.edu/software/tagger.shtml.*
- *http://research.microsoft.com/en-us/projects/sarf.*
- *http://sourceforge.net/projects/alkhalil.*
- *http://sourceforge.net/projects/aramorph.*
- *http://sourceforge.net/projects/awnbrowser/.*
- *http://tshwanedje.com/corpus/.*
- *http://tshwanedje.com/tshwanelex/.*
- *http://wba.ruhosting.nl/index.htm.*
- *http://www.nooj4nlp.net.*
- *http://www.rae.es/obras-academicas/diccionarios/nuevo-diccionario-historico-del-espanol.*
- *http://www.rdi-eg.com/technologies/POS.aspx.*

مُصطلحات الدِّراسة

Alphabet & Pronunciation	معلومات الهجاء والنُّطق
Artificial Intelligence	الذَّكاء الاصطناعيّ
Collocations	المُتَصاحِبات اللَّفظِيَّة
Computational Framework	المِنَصَّة الحاسُوبيَّة
Computational Lexicography	الصِّناعة المُعجميَّة الحاسُوبيَّة
Computational Linguistics	اللِّسانيَّات الحاسُوبيَّة
Computational Resources	الموارد الحاسوبيَّة
Concordancer	الكشَّاف السِّياقيّ
Corpus Words	كلمات المُدَوَّنة
Data-Warehouse	مُستَودعات البيانات
Dictionary Writing Systems (DWS)	أنظمة التَّحرير المعجَميّ
Etymological Dictionaries	المُعجمات التَّأثيليَّة
Etymological Information	المعلومات التَّأثيليَّة
Frequency Information	معلومات التَّكرار
Grammatical Functions	المعاني الوظيفيَّة
Historical Dictionary	المُعجم التَّاريخيّ
Language Resources	موارد اللُّغة
Language Resources	الموارد اللُّغويَّة
Lexemes (Headwords)	الوحدات المُعجميَّة
Lexical Database	قاعدةُ البيانات المُعجميَّة
Lexical Databases	قواعد البيانات المُعجميَّة
Lexical Entries	المداخل المُعجميَّة
Lexical Etymology	التَّأثيل المُعجميّ
Lexical Evidence	الشَّواهد المُعجميَّة
Lexical Meanings	المعاني المُعجميَّة
Lexical Resources	الموارد المعجميَّة
Lexical-Semantic Database	قاعدة بياناتٍ مُعجميَّة دلاليَّة
Lexicography	صناعة المُعجم
Linguistic Corpora	المُدوَّنات اللُّغويَّة
Linguistic Corpus	المُدَوَّنة اللُّغَويَّة
Machine Translation	التَّرجمة الآليَّة

Morphological Analysis System	نظامُ التَّحليل الصَّرفيّ
Morphology	علم البِنية
Naïve Bayes Classifier	مُصَنِّف بايز المُبَسَّط
Natural Language Processing (NLP)	مُعالجة اللُّغات الطَّبيعيَّة
Noun	الاسم
Ontology	الأنطولوجيا المعرفيَّة
Particle	الأداة
Parts of Speech (PoS)	أقسام الكلام
Parts of Speech Taggers	مُعَنونات أقسام الكلام
Phonology	علم الأصوات
Roots	جُذُور الكلمات
Roots Database	قاعدة البيانات الجذريَّة
Search Engines	مُحرِّكاتٍ بحثيَّة
Semantic Analysis System	نظام التَّحليل الدِّلاليّ
Semantic Relations	العلاقات الدِّلاليَّة
Semantics	علم الدِّلالة
Stems	الجُذُوع
Synonyms	المُتَرادفات
Synsets	المفاهيم
Syntactic Analysis System	نظام التَّحليل التَّركيبيّ
Syntax	علم التَّركِيب
Text Indexer	المُفهرس الآليّ للنُّصُوص
Text Mining	التَّنقيب في النُّصُوص
Usage Levels	مُستويات الاستعمال
Verb	الفعل
Word Sense Disambiguation (WSD)	فَكِّ الالتباس الدِّلالِيّ للكَلِماتِ
WordNet	الشَّبكة الدِّلاليَّة [شبكة الكلمات]