حَوسَبة المُعجَم التَّاريخيّ للُّغة العَرَبيَّة

د. المُعتزّ بالله السَّعيد
جامعة القاهرة

1. مُقَدِّمة

1.1. مفهُوم المُعجَم التَّاريخيّ

المعجَم التَّاريخِيُّ *HISTORICAL DICTIONARY* مُعجَمٌ لُغَوِيٌّ عامٌّ، يَستَمِدُّ مادَّتَه مِن التُّراث الإنسانِيّ المكتوبِ والمدَوَّنِ عبرَ العُصُور والأمكِنة في مُختَلِف العُلُوم والفُنُون والآداب، ويَضُمُّ مُفرَداتِ اللُّغة، ويُبيِّن معانِيَها ومُشتَقَّاتِها وأساليبَها؛ ويُعنى بتأثيلِها وتاريخِ استعمالِها أو إهمالِها، ودِراسَة تَطَوُّر مبانِيها ومعانِيها عبرَ المراحِل الزَّمَنِيَّة المتَعاقِبة للُّغة[1]. وتُرَتَّبُ المداخلُ والوحداتُ المُعجَمِيَّةُ في هذا النَّوع من المُعجَمات ترتيبًا ألفبائيًّا أو جذريًّا وفقَ ما يتناسبُ مع طبيعة اللُّغة المَعنِيَّة.

والمُعجمُ التَّاريخيُّ – بذلكَ – لا يقتصرُ على احتواء المادَّة المُعجميَّة بالمفهوم التَّقليديّ لدى أهل الصِّناعة المُعجميَّة؛ وإنَّما يتجاوزُ ذلكَ ليكونَ ديوانًا كبيرًا للُّغة؛ إذ يُساعدُ على فهم التُّراث الإنسانِيّ بِرُمَّتِه، ويُبيِّنُ كيفَ تُؤَثِّر اللُّغةُ في الشُّعوب وكيفَ تتأثَّرُ بها، ويُبرزُ جوانِبَ التَّبايُنِ بين لُغة الفَردِ ولُغة المُجتَمَع، كما يفتحُ آفاقًا جديدةً وواسعة للبحث في اللِّسانيَّات *LINGUISTICS*، وعلم الأعراق (دِراسَة خصائِص الشُّعوب) *ETHNOLOGY*، وعِلم الإنسان *ANTHROPOLOGY*،

1 – السَّعيد (المُعتزّ بالله): مُدَوَّنة مُعجَم تاريخيّ للُّغة العربيَّة "مُعالجَة لُغَوِيَّة حاسوبِيَّة"، أطروحة دكتوراه، جامعة القاهرة، 2011م، ص 2، بتصرُّف.

وعِلم الآثار *ARCHAEOLOGY*؛ وهو – كذلِك – أداةٌ للبَحث في تارِيخ العُلُوم *HISTORY OF SCIENCE*[2].

1.2. مُحاولات إنجاز مُعجَمٍ تاريخيٍّ للُّغة العَرَبِيَّة

كانت هناك العديد من المحاولات السَّابقة لبناء مُعجَمٍ تاريخيٍّ للعربيَّة، لعلَّ أبرزَها تلكَ المُحاولة الَّتي قامَ بها المستشرق الألماني أوجُست فيشر *AUGUST FISCHER* (1865-1949)، حيثُ أعلنَ عنها في مؤتمر اللُّغَوِيِّين الألمان في بازل عام 1907[3]، وتبنَّاها مجمعُ اللُّغة العربيَّة بالقاهرة عامَ 1935 بناءً على مُقترَح عُضو المجمع في ذلكَ الوقت – المُستَشرِق الإيطالي كارلو نلينو *CARLOS ALFONSO NALLINO* (1872-1938). وقد أعقبَت محاولة فيشر عدَّةُ مُحاولاتٍ أخرى، منها: مُحاولة الجمعيَّة الألمانيَّة للاستشراق *DEUTSCHE MORGENLÄNDISCHE GESELLSCHAFT* الَّتي سَعَت في عام 1957 إلى إحياء مشروع فيشر بعدَ وفاتِهِ وتناثُر جُزازاتِه بينَ مِصر وألمانيا، من خلال المُعجم الموسُوم بـ "مُعجم اللُّغة العربيَّة الفُصحى *WÖRTERBUCH DER KLASSISCHEN ARABISCHEN SPRACHE*" بريادة المُستشرقين الألمان: يورج كريمر *JÖRG KRAEMER* (1917-1961)، وهيلموت جيتييه *HELMUT GÄTJE* (1927-1986)، وأنطون شبيتالر *ANTON SPITALER* (1910-2003)؛ ثُمَّ المُحاولة الأولى لجمعيَّة المعجَمِيَّة العربيَّة بتونس في عام 1990، والمُحاولة الثَّانية لها في عام 1996، ومحاولة هيئة المعجم التَّاريخيّ للُّغة العربيَّة بالقاهرة في عام 2004 بدعم من حُكومة الشَّارقة[4]، وأخيرًا المحاولة القائمة في المركز العربيّ للأبحاث في الدَّوحة منذ 2013.

والواقعُ أنَّ أيًّا مِن هذه المُحاولات لم يُكتَب لها النَّجاحُ إلى الآن. فقد توقَّفَت مُحاولةُ فيشر بعدما اضطرَّتهُ الظُّرُوفُ للعودة إلى ألمانيا في أعقاب الحرب العالميَّة

2 – السَّعيد (المُعتزّ بالله): السَّابق، ص 2، بتصرُّف.

3 – فيشر (أوجُست): المُعجَم اللُّغَوِيّ التَّارِيخِيّ، القِسم الأوَّل، مجمع اللُّغة العَرَبِيَّة، القاهِرة، 1967م، ص 25.

4 – للمزيد حولَ هذه المُحاولات، راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص 23 : 31.

الثَّانيّة؛ ونتجَ عن مُحاولتِه مجموعةٌ من الجزازات، لم يَجِد مجمعُ اللُّغة العربيَّة بالقاهرة ما يَصلُحُ للنَّشر مِنها إلَّا مُقَدِّمةً – أعَدَّها فيشر بِنَفسِه – ونموذجًا يبدأ مِن حرف الهمزة إلى مادَّة "أبد"؛ فَنَشَرَهُما المَجمَعُ في عام 1950 بِعُنوان (مُعجَم فيشر "مُقَدِّمة ونموذَج مِنه")، ثُمَّ أعادَ نَشرَهُما في عام 1967 بِعُنوان (المُعجَم اللُّغَوِيّ التَّارِيخِيّ). أمَّا مُحاولة الجمعيَّة الألمانيَّة للاستشراق فقد غيَّرت مسارَها ليُصبحَ مُعجَمُها مرحليًّا، يُغَطِّي حقبة اللُّغة العربيَّة للقُرُون الثَّلاثة الأولى مِن الهِجرة وقَرنَينِ قبلها[5]. أمَّا بخُصُوص مُحاولتَيْ جمعيَّة المعجَمِيَّة العربيَّة فقد توقَّفَت عندَ بناء "مُدَوَّنة مُعجَمِيَّة مُؤَرِّخة للعَصر الجاهِلِيّ"، تَضُمُّ نُصُوصًا تنتمِي إلى الفترة من عام 200م إلى عام 604م[6]. وتوقَّفَ العملُ في هيئة المعجم التَّاريخيّ للُّغة العربيَّة بالقاهرة عندَ إعداد "منهجٍ لإنجاز مُعجَم اللُّغَة العَرَبِيَّة التَّارِيخِيّ". وأمَّا المحاولة القائمة في المركز العربيّ للأبحاث في الدَّوحة منذ 2013، فقد خطا القائمونَ عليها خُطواتٍ أوَّليَّة، ترتَّبَ عليها إعدادُ ببليوغرافيا مرجعيَّة لمصادر المُدَوَّنة اللُّغَويَّة للمُعجَم حتَّى نهاية القرن الثَّاني الهجريّ، وبناءُ هذه المُدَوَّنة فعليًّا وإعدادُ منهجيَّةٍ لصناعة المُعجم التَّاريخيّ المنشُود.

1.3. في حَوسَبة المُعجَم اللُّغَوِيّ

أدَّى ظُهورُ اللِّسانيَّات الحاسُوبيَّة *COMPUTATIONAL LINGUISTICS* في مطلع النِّصف الأخير من القرن العشرين إلى طفرةٍ معلُوماتيَّةٍ هائلة، أمكَنَ معها تطويعُ الآلة لمعالجَة المجمُوعات الضَّخمة من النُّصُوص – أو ما يُعرَفُ بالمدَوَّنات اللُّغَويَّة

---

5 – راجع في ذلك:

KRAEMER, J. & GÄTJE, H. & SPITALER, A. & ULLMANN, M. (1970). WÖRTERBUCH DER KLASSISCHEN ARABISCHEN SPRACHE. OTTO HARRASSOWITZ. WIESBADEN. BD.I. S.I. P.XIII.

وراجع أيضًا:

FISCHER, W. & GÄTJE, H. EDITORS. (1982). GRUNDRISS DER ARABISCHEN PHILOLOGIE: SPRACHWISSENSCHAFT. REICHERT. WIESBADEN. S.40.

6 – ابن مُراد (إبراهِيم): في مفهوم المُعجَم التَّاريخِيّ وتَطبيقاته على العَرَبِيَّة، بحث مُقَدَّم إلى ندوة اتِّحاد المعاجم العربية حول المعجم التاريخيّ للُّغة العربِيَّة، الشَّارِقة، 2006م، ص 15 : 17، بتصرُّف

NATURAL – *LINGUISTIC CORPORA* بالإضافة إلى مُعالجَة اللُّغات الطَّبيعيَّة *LANGUAGE PROCESSING (NLP)* عبرَ مُستويات التَّحليل اللُّغَويّ، بما في ذلكَ مُستوى التَّحليل المعجَمِيّ. وصارَ بالإمكان الاستغناءُ عن الجُهود البشريَّة الجبَّارة في مراحل الصِّناعة المُعجَمِيَّة – جمعًا وتحريرًا ونشرًا[7]. وأصبحَ دورُ الحاسُوب يتجاوزُ توفيرَ الجهُود والطَّاقات إلى توحيد معايير الصِّناعة المعجَمِيَّة وتيسير عمليَّات التَّحرير والمراجعة؛ فأخَذَت الصِّناعةُ المعجَمِيَّة – بذلكَ – منحى جديدًا، أمكنَ معهُ إخضاعُ الآلة لمتَطلَّبات الصِّناعة عبرَ جميع المراحل، بدايةً من مرحلة الجمع، بمراحلها الفرعيَّة المتَمَثِّلة في إعداد المُدَوَّنة اللُّغَويَّة وتهيئة النُّصُوص المُمثِّلة لواقعَ اللُّغة المعنِيَّة، ثُمَّ مُعالجَة هذه النُّصُوص على مُستوى المباني – لُغَويًّا وآليًّا وإحصائيًّا – بما يتناسبُ مع طبيعة اللُّغة ويتوافقُ مع أنظمة ترميزِها سعيًا إلى تعيين المداخل والوحدات المعجَمِيَّة والمعلُومات النَّحويَّة السَّماعيَّة [الصَّرفيَّة والتَّركيبيَّة] الَّتي يُفتَرَضُ أن يتضَمَّنَها المعجَمُ في سياق مبانيه؛ ثُمَّ مُعالجَة النُّصُوص دلاليًّا – على مُستوى المعاني – بهدف تعيين الدِّلالات المُستخدَمة للوحدات المُعجَمِيَّة وبيان مواضع الاستشهاد المناسبة. ومُرُورًا بمرحلة التَّحرير، بمراحلها الفرعيَّة المتَمَثِّلة في بناء قواعد البيانات المعجميَّة، وتعيين المعلومات المعجمِيَّة المتَضَمَّنة، وتطويع المدوَّنة اللُّغَوِيَّة لعمليَّات التَّحرير المعجَمِيّ، والرَّبط الآليّ بينَ مادَّة المدوَّنة اللُّغَوِيَّة المجموعة وقواعد البيانات المعجَمِيَّة، والتَّوسيم الآليّ للمداخل والوحدات، والتَّرتيب الآليّ للمعلومات المعجَمِيَّة، وتطويع أنظمة التَّحرير لتتوافقَ مع طبيعة اللُّغة المعنِيَّة من ناحيةٍ والهدف من المعجم المنشُود من ناحيةٍ أخرى. وانتهاءً بمرحلة النَّشر، بمراحلها الفرعيَّة المتَمَثِّلة في بناء الهيكل الحاسُوبيّ

7 – للمزيد حولَ دور الحاسُوب في حوسبة المعاجم التَّاريخيَّة، راجع:

PAJZS, J. (2000). MAKING HISTORICAL DICTIONARIES BY COMPUTER. PROCEEDINGS OF EURALEX 2000. ULRICH HEID ED. UNIVERSITY OF STUTTGART, STUTTGART, 2000. PP. 249-259.

للمعجَم، وتصميم الواجهة/البوَّابة الإلكترونيَّة، وتطوير آليَّات البحث والتَّحرير، والرَّبط بينَ المادَّة المعجَمِيَّة المحَرَّرة ووسائل شرح هذه المادَّة – من الرُّسُومات والصَّوتيَّات والمرئيَّات – وإضافة آليَّات إدارة الهيكل المعجمِيّ المخرَج، وتعيين الصِّيَغ الإلكترونيَّة الَّتي تُؤدِّي إلى ظهور المعجَم على الوجه المنشُود، وإتاحة المعجم عبرَ صفحات الشَّبكة العنكبُوتيَّة.

وقد ظهرت بعضُ التَّجارب النَّاجحة لحوسبة المُعجمات اللُّغَويَّة في الحقبة الأخيرة من القَرن العشرين ومع مطلع الألفيَّة الثَّالثة. منها – على سبيل المثال – تجربة مُعجَم كزلينز–كوبيلد الإنجليزيّ التَّعليميّ[8]، الَّذي أنجزته مؤسَّسَة كولينـز–كوبيلد *COLLINS COBUILD*. فقد خضَعَ المُعجمُ لتقنيّات الحاسُوب في جميع مراحلِه؛ وتمثَّلَت الغايةُ الَّتي التَمَسَها صُنَّاعُ المعجَم في توفير مادَّةٍ مُعجَمِيَّة تعليميَّة للغة الإنجليزيَّة المعاصِرة، لمساعدة مُتَعلِّميها – من النَّاطِقِينَ بالإنجليزِيَّة ومِن غَيرِهِم – على فهمٍ أعمقَ لتراكيبِها واستعمالاتها اللُّغَوِيَّة بعد حصر المفردات والتَّراكيب الشَّائعة وبيان نسبَة شيوع كلٍّ منها؛ كما اهتمَّ صُنَّاعُ المعجَم بوضع أكبر عدد ممكنٍ من الاستعمالات اللُّغَوِيَّة لأقلِّ عدد من المفردات، وهو ما يجعل تَعَلُّمَ اللُّغة والتَّمكُّنَ منها أمرًا ميسورًا. وقد ظهرَ المُعجَمُ الَّذي أشرَفَ عليه المعجميُّ الإنجليزيّ الاسكتلنديّ جون سينكلير *JON SINCLAIR* (1933-2007) في صُورَتِهِ الورقيَّة للمرَّة الأولى بعُنوان *ENGLISH DICTIONARY HELPING LEARNERS WITH REAL ENGLISH* في عام 1987؛ ثُمَّ أعيدَ نشرُهُ في العديد من الصُّوَر الورقيَّة والإلكترونيَّة. وقد أتاحَت مُؤسَّسة كولينز–كوبيلد – مُؤخَّرًا – العديدَ من أشكال النَّشر الإلكترونيّ لهذا المعجَم التَّعليميّ على الأقراص المدمجَة *CDs* وعبرَ صفحات الويب *ON-LINE* وعلى الجوَّالات والحواسيب اللَّوحِيَّة *TABLETS*؛ كما حَرَصَ صُنَّاعُ

8 – للمزيد حولَ تجربة كولينز–كوبيلد، راجع:

SINCLAIR, J. (2009). COLLINS COBUILD ADVANCED DICTIONARY OF ENGLISH. HEINLE CENGAGE LEARNING.

المعجَم على تنمية مُحتوى المدَوَّنة اللُّغَوِيَّة المُستخدمة في المُعجم، فتجاوزَ عددُ كلماتها 650 مليون كلمة في عام 2014، حوتها قاعدةُ البيانات النَّصِّيَّة الَّتي وَسَمَتها مُؤسَّسةُ كولينز-كوبيلد بـ "بنك اللُّغة الإنجليزيَّة *BANK OF ENGLISH*"[9]. وتجدُرُ الإشارةُ إلى عناية فريق العمل في المُعجَم بتوظيف قاعدةِ بياناتِه باعتبارِها موردًا للمداخل والوحدات والمعاني المُعجَمِيَّة؛ كما قاموا بربط قاعدة البيانات المعجَمِيَّة الرَّئيسة ومُعطياتِها بأدوات التَّحرير المعجَمِيّ بصُورةٍ آليَّة، أدَّت إلى ظُهور مادَّة المُعجَم في صُورةٍ مُنتظِمةٍ ومُتجانسة.

2. إشكالات حوسبة المُعجم التَّاريخيّ للعربيَّة

ثمَّةَ بعضُ الإشكالات الَّتي ينبغي الوُقُوفُ عليها قبلَ إعداد منهجيَّة حوسبة المُعجَم التَّاريخيّ للُّغة العربيَّة، سعيًا إلى معرفة العقبات الَّتي يُمكنُ أن تقفَ حائلًا أمامَ حوسبة المُعجم والعمل على إزالتها. وتتمثَّلُ أولى هذه الإشكالات في أنَّنا لا نملكُ تجربةً واحدةً لحوسبة المُعجمات التَّاريخيَّة عبرَ مختلف مراحل الصِّناعة المُعجَمِيَّة. فقد سارَت التَّجارب السَّابقة بصُورةٍ تقليديَّة، أدَّت إلى إنجاز العمل المُعجَميّ في عشرات السَّنوات، وعلى أيدي أجيالٍ مُتعاقبة من المُعجَمِيِّينَ واللُّغَوِيِّينَ والمُحرِّرين. وبالنَّظر إلى هذه التَّجارب نجدُ أنَّ إعدادَ المُعجَم الألمانيّ *DEUTSCHES WÖRTERBUCH (DW)* قد استغرقَ ما يربو على مئةٍ وعشرينَ عامًا، حيثُ بدأ العملُ فيه عامَ 1838 وفقَ منهج الأخوين يعقوب جريم *JACOB GRIMM* (1785-1863) وفيلهِلم جريم *WILHELM GRIMM* (1786-1859)، واكتَمَلَ في عام 1961، بإشرافِ الأكادِيمِيَّة البروسِيَّة للعُلُوم والدِّراسات الإنسانِيَّة في برلين *PREUSSISCHE AKADEMIE DER WISSENSCHAFTEN ZU BERLIN* وأكادِيمِيَّة العُلُوم في جوتنجِن *AKADEMIE DER WISSENSCHAFTEN ZU GÖTTINGEN*. وعلى نحوٍ قريبٍ سارَ العملُ في مُعجَم اللُّغة الهولنديَّة *WOORDENBOEK DER*

9 – الموقع الإلكترونيّ لسلسلة معاجم كولينز:

HTTP://WWW.COLLINSDICTIONARY.COM.

NEDERLANDSCHE TAAL (WNT) إذ استغرَقَ بناؤُه ما يقرُبُ من مئةٍ وخمسينَ عامًا، فبدأ العملُ فيه عامَ 1849، واكتملَ في عام 1998. وكانَ الأوفرَ حظًّا من هاتين التَّجربتين مُعجمُ أكسفورد للإنجليزِيَّة *THE OXFORD ENGLISH DICTIONARY (OED)* الَّذي استغرَقَ بناؤُه أقلَّ من ستِّينَ عامًا إذ بدأ العملُ فيه عامَ 1859، واكتملَ في عام 1928. أمَّا مُعجم الأكاديميَّة السِّويديَّة *SVENSKA AKADEMIENS ORDBOK (SAOB)* فقد بدأ العملُ فيه عامَ 1884، ومن المُزمَع أن يكتملَ في عام 2017 حسبما أعلَنَت الأكاديميَّة السِّويديَّة[10]. ولم تتجاوز حوسبةُ أيٍّ من هذه المُعجماتِ مرحلةَ النَّشر، باستثناء مُعجم أكسفورد للإنجليزِيَّة الَّذي يُعادُ تحريرُه، تمهيدًا لإصداره في نُسخةٍ ثالثة *OED3* [بعدَ نُسختيه الأولى *OED1* في عام 1928، والثَّانية *OED2* في عام 1989]، حيثُ يقودُ المُعجميّ الإنجليزيّ مايكل بروفيت *MICHAEL PROFFITT* [11] فريقًا من أكثر من ثلاثمئة باحثٍ ومُحرِّرٍ بهدف إخراج المُعجم في صورةٍ جديدةٍ؛ ويستعينونَ بتقنيّات الحاسُوب في جميع مراحل الصِّناعة. وعليه، فإنَّنا سنكونُ أمامَ عملٍ ضخم، نفتقرُ فيه إلى تجربةٍ ناجحةٍ يُمكنُ الاقتداءُ بها لنُوفِّرَ الوقتَ والجُهدَ والطَّاقةَ البشريَّة. وتزيدُ صُعُوبةُ الأمر إذا عرَفنا أنَّ اللُّغةَ العربيَّةَ لا تكادُ تعرفُ طريقَها إلى حوسَبة المُعجم – بالمفهوم الدَّقيق – حتَّى الآن. ففي الوقت الَّذي يقتصرُ فيه المعجَمِيُّونَ العربُ – في مواردهم – على المعاجم القديمة وما تناثرَ من معلوماتٍ مُعجَمِيَّة في بُطُون ما تحويه المكتبةُ العربيَّة، نجدُ العديدَ من المؤسَّسات المَعنيَّة بصناعة المُعجمات الجرمانيَّة واللَّاتينيَّة تضُمُّ نُخبًا من المعجَمِيِّينَ العارفينَ بأصُول الصِّناعة؛ ونجدُ هؤلاء يعتمدونَ في مواردهم على مُدَوَّناتٍ لُغَويَّةٍ مُحوسَبةٍ تعكسُ واقعَ اللُّغة المعنِيَّة، وعلى قواعد بياناتٍ لُغَويَّة مُنَقَّحة، وعلى أدواتٍ حاسُوبيَّةٍ وإحصائيَّة لمعالجَة المادَّة المعجَمِيَّة على

10 – للمزيد حولَ هذه التَّجارب، راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص 8 : 22.

11 – "مايكل بروفيت" هُو رئيسُ التَّحرير الثَّامن لمُعجم أكسفورد للإنجليزيَّة؛ ويتولَّى مهامَّ عملِهِ منذ نوفمبر 2013، بعدَ تقاعُد سلفِهِ جون سيمبسون JOHN SIMPSON الَّذي تولَّى رئاسة تحرير المُعجم الإنجليزيّ بينَ عامَي 1993 و2013.

مُستوى المباني والمعاني؛ بالإضافة إلى اعتمادهم على معاجمهم القديمة وما تحويه المصَنَّفاتُ في مكتباتهم. وتظلُّ الصِّناعةُ العربيَّةُ التَّقليديَّة هي الغالبة على المدارس العربيَّة، باستثناء بعض التَّجارب الَّتي حاولَ أصحابُها أن يُفيدوا من التِّقنيات الحاسُوبيَّة في بعض مراحل الصِّناعة، كتجربة (المكنز الكبير) الَّذي أنجزَهُ فريقُ عملٍ بإشراف الدُّكتور أحمد مُختار عُمَر؛ حيثُ جُمِعَت مادَّةُ المكنز – وهو مُعجَمٌ دلاليٌّ – من أمَّهات الكُتب العربيَّة المعنِيَّة بعلاقَتَي التَّرادُف والتَّضادّ، وصُنِّفَت في ضوء نظريَّة الحُقُول الدِّلاليَّة في قواعدَ بياناتٍ – استخدمَها الفريقُ باعتبارها قوالب مُعجَمِيَّة مُحوسَبة؛ واتُّخِذَت هذه القوالبُ معيارًا للمخرَج المعجَمِيّ[12]. وكذلكَ تجربة (المعجَم الحاسُوبيّ التَّفاعُليّ) الَّذي أشرَفَت على صناعتِهِ مدينةُ الملك عبد العزيز للعُلُوم والتِّقنية بالرِّياض، حيثُ جُمِعَت مادَّتهُ من المعجَم الوسيط، وأعيدَ بناءُ هيكلِهِ في قاعدة بياناتٍ مُعجَمِيَّة؛ ثُمَّ ذُيِّلَ بوسائل شرح المعنى في صُورةٍ رقميَّة، عُنِيَ فيها بالرُّسُومات والصَّوتيَّات والمرئيَّات[13]. ولَمَّا كُنَّا نفتقرُ إلى تجربةٍ لحوسبة مُعجمٍ تاريخيّ، ونفتقرُ كذلكَ إلى تجربةٍ لحوسبة المُعجم العربيّ، فإنَّ علينا أن نصنعَ منهجيَّةً لحوسبة المُعجم التَّاريخيّ للُّغة العَرَبيَّة، تجمعُ بينَ الأصالة والحداثة، وتُساعدُ على إنجاز المُعجم المنشُود وفقَ معايير الصِّناعة المُعجَمِيَّة.

أمَّا الإشكال الثَّاني، فيتمثَّلُ في ضبابيَّة الرُّؤية المنهجيَّة وحاجة الصِّناعة المُعجَمِيَّة العربيَّة إلى طفرةٍ حقيقيَّة، تُعيدُ للعرب ريادتَهم لهذا الميدان وتضعُ قواعدَ راسخةً لصناعة المُعجم العربيّ. فالعربُ قد عرفوا طريقَهُم إلى الصِّناعة المعجَمِيَّة في وقتٍ مُبكِّرٍ، مُقارنةً بالأمم الأخرى. ونتجَ عن هذه المعرفة ظهورُ العديد من المدارس المعجَمِيَّة العربيَّة الَّتي تُعنى بحصر مُفردات العربيَّة وتفسير معانيها، انطلاقًا من المباني – كما في معاجم الألفاظ – أو المعاني – كما في معاجم

12 – راجع: عُمَر (أحمد مُختار): المكنز الكبير، "معجم شامل للمجالات والمُترادفات والمتضادَّات"، مؤسَّسة سُطور، 2000

13 – الموقع الإلكترونيّ للمُعجم:

HTTP://WWW.ALMUAJAM.ORG.

الموضوعات. وكانَ لظهور عدد من المعاجم العربية مُبكِّرًا أثرٌ بالغٌ في اكتمال أركان الصِّناعة المعجَمِيَّة العربيَّة؛ وهو أمرٌ طبيعيٌّ يفرضُهُ تنوُّعُ المدارس المعجَمِيَّة، بدءًا بمدرسة التَّحليل الصَّوتيّ الَّتي استمَدَّت قواعدها من مُعجَم (العَين) للخليل بن أحمد (ت 173 ه)، وانطلاقًا بمدرسة الأبنية الَّتي أرساها أبو بكر مُحَمَّد بن الحسَن بن دُريد (ت 321 ه) من خلال مُعجَمِهِ (جمهرة اللُّغة)؛ ثُمَّ مدرسة التَّقفية [القافية] الَّتي أسَّسَها إسماعيل بن حمَّاد الجوهريّ (ت 393 ه) من خلال مُعجَمِهِ (تاج اللُّغة وصحاح العربيَّة)، وانتهاءً بمدرسة التَّرتيب الألفبائيّ الَّتي أسَسَ لها أبو المعالِي مُحَمَّد بن تَمِيم البَرْمَكِيّ (ت 411 ه) في مُصَنَّفِه (المنتَهَى في اللُّغَة) الَّذي أعادَ فِيه تَرتِيبَ مُعجَم الصِّحاح (تاج اللُّغَة وصِحاح العَرَبِيَّة)، وسارَ عليه أبو القاسم الزَّمَخشَرِيّ (ت 538 ه)، ثُمَّ سارَت عليها جُلُّ المعجَمات العَرَبِيَّة. وأدَّت الجُهُودُ المبذولة في الارتقاء بصناعة المُعجم إلى تنوُّع المادَّة المعجَمِيَّة الَّتي تحويها المعاجم العربيَّة – بمُختلف أنواعها. إلَّا أنَّ قُصُورًا في مناهج المعجَمِيِّينَ أدَّى إلى تأخُّر العرب في الصِّناعة المعجَمِيَّة – مُقارنة بالعديد من الأمم الأخرى. فمع تقدُّم المعجَمِيِّينَ العرب قديمًا في هذه الصِّناعة، إلَّا أنَّ انشغالَ تابعيهم بتنقيح ما صَنَّفَهُ القُدماءُ وتهذيبه وشرحه دونَ العناية بسدّ الفجوات المتروكة أو مُراعاة التَّطَوُّر اللُّغَوِيّ الحادث في اللُّغة العربيَّة، كُلُّها أمورٌ أدَّت إلى توسيع الفجوة الَّتي تُعانيها المكتبةُ العربيَّة. ولنا أن نلحَظَ هذا في المعاجم المعاصرة الَّتي تشرحُ الكلمات – في كثيرٍ من الأحيان – بعباراتٍ صعبةٍ على مُستخدم المعجَم، وتزخرُ في مادَّتها بالحُوشيّ والغريب من المفرَدات الَّتي يقلُّ استخدامُها في عموم العربيَّة، وكثيرٌ منها قد اندثرَ في العربيَّة المعاصِرة. وفي الوقت ذاتِهِ، تخلو هذه المعاجمُ من كثيرٍ ممَّا استُحدِثَ من المفردات والألفاظ. ولا عجبَ حينئذٍ أن نجدَ المعجَم الوسيط، الَّذي وَضَعَهُ نخبةٌ من أعضاء مجمع اللُّغة العربيَّة بالقاهرة – على سبيل المثال – يزخرُ بمُفردات مهجورةٍ على شاكلة (بجبجَ، وبحشَلَ ..) ويُهمِلُ [في طبيعتِه الثَّالثة المزيدة والمنَقَّحة – 2004] كلماتٍ صارت أساسيَّةً في العربيَّة المعاصِرة، مثل:

حاسُوب وجوَّال[14]. والأمثلةُ على ذلك كثيرة، في الوجيز والوسيط وغيرهما من المعاجم العربيَّة المعاصِرة. إنَّ تتبُّعَ مادَّةِ المعاجم العربيَّة المعاصرة تُؤكِّدُ أنَّ جُهودَ المعجَمِيِّينَ العرب المعاصرين تكاد تنحصرُ في اختصار المعاجم العربيَّة القديمة وإعادة إنتاجها لتُلائمَ روحَ العصر الحديث، وتهذيب منهج المدرسة الألفبائيَّة في ترتيب المداخل والوحدات المعجَمِيَّة؛ بالإضافة إلى جهودٍ ميمونة في تعريب المصطلحات العلميَّة والانتقال من محدوديَّة المعاجم العامَّة إلى سعة المعاجم المختَصَّة؛ ولا تزالُ الصِّناعةُ المعجَمِيَّة التَّقليديَّة مُسيطرة على المجامع اللُّغَوِيَّة والمؤَسَّسات المعنِيَّة بصناعة المعجَم العربيّ، حيثُ تعتمدُ بصورةٍ رئيسةٍ على استخلاص المادَّة المعجَمِيَّة من المعاجم العربيَّة القديمة – بما فيها من حوشيٍّ أو غريب – ثُمَّ إعادة تحرير هذه المادَّة يدويًّا بعدَ تهذيبِها. ومع ما في جُهود المجامع اللُّغَوِيَّة والمؤسَّسات من عنايةٍ بالصِّناعة المعجَمِيَّة العربيَّة، إلَّا أنَّها لا تُؤَدِّي إلى نتائجَ فعليَّةٍ ملموسة، ولا تُحَقِّقُ الهدفَ المنشُود.

إنَّ حوسبةَ المُعجَم التَّاريخيّ للُّغة العَرَبِيَّة أسلُوبٌ مُتقدِّمٌ من أساليب الصِّناعة المُعجَمِيَّة، يتطلَّبُ انتهاجُهُ الانطلاقَ من رؤيةٍ واضحةٍ ومنهجٍ مُكتمل الأركان، والوُقُوفَ على قاعدةٍ صلبة تُساعدُ القائمينَ على صناعة المُعجم المنشُود في رسم خارطة طريقٍ واضحة المعالِم، بحيثُ يعرفُونَ جميعَ مراحل العمل المُعجَميّ وإطارَهُ المنهجيّ وإطارَهُ الزَّمَنيّ وتكلفتَهُ ومواردَهُ البشريَّة ومواردَهُ اللُّغَويَّة والحاسُوبيَّة والمُعجَمِيَّة. والواقعُ أنَّ تحقيقَ هذه الأمور مُجتمعةً يستدعي إخلاصًا للهدف المنشود وإنكارًا للذَّات من أجل تحقيقِه وفهمًا دقيقًا لمفهوم الصِّناعة المُعجَمِيَّة ودِراسةَ التَّجارِب النَّاجحة للمُعجَمات اللُّغَوِيَّة التَّارِيخِيَّة، ومُحاولة الإفادة منها في بِناءِ منهجٍ مُتكامِلٍ للمُعجَم العَرَبِيّ المَنشود؛ ثُمَّ تقييم

14 – راجع: مَجمع اللُّغة العَرَبِيَّة بالقاهِرة: المُعجَم الوسيط، مكتَبَة الشّروق الدَّولِيَّة، القاهِرَة، ط3، 2004.

المُحاوَلاتِ السَّابِقَةِ لبناء مُعجَمٍ تاريخِيٍّ للعَرَبِيَّة للوُقُوف على جوانِب القُصُور ومُحاوَلَة مُعالجتها بِطريقَةٍ منهجِيَّة تَتَّفِقُ مع أساليب صِناعَة المُعجَم الحَدِيث.

ويتمثَّلُ الإشكالُ الثَّالثُ في طبيعة اللُّغة العربيَّة الاشتقاقيَّة ونظامها الكتابيّ. فهي – على مُستوى البنية – لُغةٌ اشتقاقيَّة، يُمكنُ تحليلُ الكلمات فيها إلى: الجُذور *ROOTS* والجُذُوع *STEMS* والفُروع *LEMMATA* والزَّوائد *AFFIXES* الَّتي تأتي في هيئة سوابق *PREFIXES* ولواحق *SUFFIXES*. وهي – على مُستوى التَّركيب – لُغةٌ مُعرَبة، تَتغيَّرُ هيئةُ كلماتها تبعًا لموقعها الإعرابيّ، إذ تظهرُ علاماتُ الإعراب فيها وفقًا لقوانين الإعراب والبناء الَّتي تُطَبَّقُ على الكلمة العربيَّة في أقسامِها الثَّلاثة (الاسم، والفِعل، والحَرف). وعلى مُستوى الدِّلالة، تتميَّزُ المفردات العربيَّة بأنَّها مُتَطَوِّرة في دلالاتِها ومعانيها، حيثُ تتغيَّرُ معاني الكلمات من زمنٍ إلى زمن ومن مكانٍ إلى مكان ومن سياقٍ إلى سياق، وفقًا لتأثير العوامل الاجتماعيَّة والثَّقافيَّة والحضاريَّة والتَّاريخيَّة[15]؛ كما يبدو أثرُ البيئة العربيَّة واضِحًا في تعدُّد دَلالة المُفرَدة الواحدة من ناحية، وتعدُّد المُفرَدات الَّتي تُشِيرُ إلى دَلالةٍ واحدةٍ من ناحيةٍ أخرى. وتتميَّزُ اللُّغةُ العربيَّةُ – كذلكَ – بنظامٍ كتابيٍّ خاصٍّ، حيثُ تُكتَبُ فيها الوحدات الكتابيَّةُ [الجرافيمات *GRAPHEMES*] بصُورةٍ مُتشابكة، الأمرُ الَّذي يُؤدِّي إلى تعدُّد أشكال الجرافيم الواحد – وفقًا لموضِعِهِ في الجُملة وظهورِه في مجموعة من الأشكال الكتابيَّة [الألوجرافات *ALLOGRAPHS*]. وسعيًا إلى حوسبة المُعجم التَّاريخيّ للُّغة العربيَّة، ينبغي أن تُراعَى طبيعةُ العربيَّة عندَ بناء مواردها الحاسُوبيَّة وأدواتِ مُعالجتِها آليًّا. فطبيعةُ اللُّغة العربيَّة ونظامُها الكتابيّ يستدعيان مُعالجةً آليَّةً خاصَّةً على جميع مُستوياتها، بما في ذلكَ مُستوى التَّحليل المُعجَمِيّ؛ وحالَ الاعتماد على أدواتٍ حاسُوبيَّةٍ مُنجزةٍ سلفًا، ينبغي التَّأكُّدُ من دعم هذه الأدوات للعربيَّة

15 – حولَ أسباب تغيُّر المعنى ومظاهره، راجع: السَّعيد (المُعتزّ بالله): علم الدِّلالة ونظريَّة المعنى، دار الهاني، القاهرة، ط1، 2013، ص 130 : 139.

أوَّلًا، واستيعابِ قواعد بياناتِها للمجموعات الضَّخمة من النُّصُوص ثانيًا، والقُدرة على التَّحكُّم فيها بما يُحَقِّقُ الهدفَ المنشودَ ثالثًا. ومع هذا، يفرضُ الواقعُ تطويعَ بعض الأدوات الَّتي تدعمُ العربيَّةَ بصورةٍ جُزئيَّة [كالمُفهرسات الآليَّة وأدوات التَّحرير المُعجَمِيّ]، إذ يستدعي توفيرُها بدعمٍ كاملٍ للعربيَّة مزيدًا من الأبحاث والدِّراسات والمُحاولات الَّتي تتطلَّبُ وقتًا وجهدًا كبيرَين، ويصعُبُ التَّنبُّؤ بنتائجها.

وثمَّةَ إشكالٌ رابعٌ ينبغي النَّظَرُ إليه بعين الاعتبار، مفادُه أنَّ كثيرًا من المعجَمِيِّينَ والقائمينَ على المؤَسَّسات المعنِيَّة بالصِّناعة المعجَمِيَّة العربيَّة لا يُمَيِّزونَ فعليًّا – على المستوى التَّطبيقيّ – بينَ مُصطَلَحَي المعجَمِيَّة *LEXICOLOGY* الَّذي يُشارُ به إلى العِلم الَّذي يُعنى بِدِراسة كلمات اللُّغة الإنسانيَّة المُعَيَّنة ويبحَث في بِنيَتِها الدِّلالِيَّة وتطوُّر هذه البِنية عبرَ الزَّمان والمكان؛ وصِناعة المعجَم *LEXICOGRAPHY*، وهي الصِّناعة المتَمَثِّلة في مجموعة الإجراءات المتعاقبة لإنتاج عملٍ مُعجميٍّ مُعَيَّنٍ – جمعًا وتحريرًا ونشرًا – وفقَ ضوابط ومعاييرَ ومناهج تتحدَّدُ في ضوء طبيعة اللُّغة والهدف من المعجَم المنشُود، وإن كانَ التَّمييزُ قائمًا بينَ المصطلَحَين من النَّاحية التَّنظيريَّة[16]. إنَّ الخلطَ القائمَ بينَ هذين المصطلَحَين يُؤدِّي إلى تفاوُتٍ كبيرٍ بينَ المادَّة المعجَمِيَّة المتَضَمَّنة في مُعجَمٍ واحدٍ، تبعًا لتغيُّر ظروف التَّأليف أو الصِّناعة وتعدُّد المعجَمِيِّينَ والمحَرِّرينَ القائمينَ على إعداد المعجَم. ويظهرُ هذا الأمر في المعاجم الكبرى الَّتي يتطلَّبُ تأليفها سنواتٍ طويلة، ويتعاقبُ عليها أجيالٌ من المحَرِّرين. ونستطيعُ التَّمثيلَ على ذلكَ بـ "المعجَم الكبير" الَّذي بدأ العملُ فيه مع إنشاء مجمع اللُّغة العربيَّة بالقاهرة في عام 1934، ولا يزالُ العملُ قائمًا فيه. فالمطَّلعُ المتَخَصِّص

16 – للتَّمييز بينَ المُعجميَّة وصناعة المُعجم، راجع: حجازي (مَحْمُود فهمي): البحث اللُّغَوِيّ، دار غريب، القاهرة، ط1، (د.ت)، ص 47، بتصرُّف. وراجِع أيضًا:

Vermon, L. (2012): Lexicology and Lexicography: Words and Ways. Webster's Digital Services

على مادَّة المعجم يستطيعُ اكتشافَ أنَّ مُحرِّرَ المادَّة المعجَمِيَّة لباب (الهمزة) يختلفُ عن مُحرِّرَ المادَّة المعجَمِيَّة لباب (الجيم)؛ وكذلكَ يُمكنُه اكتشافُ أنَّ مُحرِّرَ المادَّة المعجَمِيَّة لباب (الدَّال) يختلف عن مُحرِّرَي (الهمزة والجيم). ذلكَ أنَّ كُلًّا من هؤلاء يجتهدُ في تحرير المادَّة المعجَمِيَّة استنادًا على معرفَتِهِ وفهمه لطبيعة المعجم، دونَ الاعتماد على ضوابط مُحَدَّدة سلفًا، الأمر الَّذي يُؤدِّي إلى تعدُّد مناهج التَّحرير المعجَمِيّ بتعدُّد المحَرِّرين أنفسِهم. أضف إلى ذلكَ أنَّ هذا التَّبايُنَ يجعلُنا غيرَ قادرينَ على معرفة موعد اكتمال العمل في تحرير المعجَم – ولو على وجه التَّقريب. ولا نلمحُ هذا الأمرَ – مثلًا – في مُعجَم الأكاديميَّة السِّويديَّة *SOAD* الَّذي بدأ العملُ فيه عامَ 1884 – كما ذكرنا آنفًا – وسيستمرُ إلى عام 2017؛ وكذلك الحالُ في النُّسخة الثَّالثة من مُعجم أكسفورد للإنجليزِيَّة *OED3*، حيثُ بدأ العملُ فيها مع مطلع التِّسعينيَّات من القرن العشرين، وسيستمرُّ إلى عام 2037 – وفقَ تقدير مؤسَّسة أكسفورد. فقد اعتمدَ صُنَّاعُ المُعجَمَين على معاييرَ واضحة، يستشعرُ معها المطَّلعُ أنَّ مُحرِّرَهُ شخصٌ [أو فريقٌ] واحدٌ، وليسَ مئات المحَرِّرين، على نحو ما هُو واقعٌ فعلًا.

3. منهجيَّة حوسبة المُعجَم التَّاريخيّ للُّغة العَرَبِيَّة

تَمُرُّ الصِّناعةُ المعجَمِيَّةُ بثلاث مراحل أساسيَّة؛ هي: الجمع، والتَّحرير، والنَّشر. وتشهد صناعة المعجَم المعاصِرة عنايةً كبيرةً بالحوسبة المعجَمِيَّة عبرَ المراحل الثَّلاثة، حيثُ يُعنى المعجَمِيُّونَ بالإفادة من تقنيات الحاسُوب في جمع الموارد المعجَمِيَّة، سواءٌ أكانت مُدَوَّناتٍ لُغَوِيَّةً مصنوعةً لأغراض الصِّناعة، أم قواعد بيانات مُعجَمِيَّة، أم مواردَ أخرى. كذلكَ يُعنى المعجَمِيُّونَ – لاسيَّما في أمريكا وأوروبا – بالإفادة من أنظمة التَّحرير المعجَمِيّ *DICTIONARY WRITING SYSTEMS (DWS)* في تحرير معاجمهم وتعيين مُخرجاتها آليًّا. أمَّا الإفادة من تقنيات الحواسيب في المرحلة الثَّالثة المعنِيَّة بالنَّشر المعجَمِيّ فقد أصبَحَت أمرًا حتميًّا في

معاجم اللُّغات الإنسانيَّة بصفةٍ عامَّة، حيثُ تفرضُ الثَّورةُ المعلوماتِيَّةُ الهائلة الَّتي يشهدُها القرنُ الحادي والعشرون عنايةً فائقة بالنَّشر الحاسُوبيّ عبرَ الحواسيب الشَّخصِيَّة واللَّوحِيَّة والهواتف النَّقَّالة.. وغيرها. وسعيًا إلى إعداد منهجيَّةٍ لحوسبة المُعجَم التَّاريخيّ للُّغة العربيَّة، ينبغي تعيينُ المُدخلات أو الموارد الرَّئيسة للمُعجم المنشُود وما سينتُجُ عنها من المُخرجات المُعجَمِيَّة. ووفقًا لتوصيف المُعجم التَّاريخيّ، فإنَّ الموردَ الرَّئيسَ له هو مُدَوَّنةٌ لُغَويَّة، تعكسُ واقعَ اللُّغة العربيَّة خلالَ الحِقبة الزَّمنيَّة الَّتي يُغَطِّيها المُعجم؛ وهي بالنِّسبة للعربيَّة قريبةٌ من ألفي عامٍ تُمثِّلُ حِقبةَ التَّأريخ للُّغة العربيَّة في صُورتها المكتوبة. أمَّا المُخرجاتُ المُعجَمِيَّة، فهي: المداخل والوحدات المُعجَمِيَّة *ENTRIES & LEXEMES*، والتَّوصيفُ النَّحويّ [البنويّ والتَّركيبيّ] للوحدات، والتَّهجئة، والمعلومات الصَّوتيَّة، والمعلومات التَّأثيليَّة *ETYMOLOGICAL INFORMATION*، والمعاني المُعجَمِيَّة، والشَّواهد المُعجَمِيَّة، ومُستويات الاستعمال بينَ الشُّيُوع والإهمال أو الإباحة والحَظر.

لسنا بصدد الحديث عن موارد المُعجم التَّاريخيّ المنشود؛ ولكن يعنينا أن نعرفَ أنَّ المُدَوَّنةَ اللُّغَويَّةَ للمُعجم هي المورِدُ الوحيدُ للمداخل والوحدات المُعجَمِيَّة؛ وهي كذلكَ المورِدُ الوحيدُ الَّذي يُعتمَدُ عليه في تعيين مُستويات الاستعمال واستخراج الشَّواهد المُعجَمِيَّة. وهي موردٌ أساسيٌّ للمعاني الوظيفيَّة والمعاني المُعجَمِيَّة للوحدات، مع مواردَ فرعيَّةٍ تَضُمُّ مُصَنَّفات النَّحو العربيّ وقواعد البيانات المُعجَمِيَّة للمعاني والعلاقات الدِّلاليَّة بينَ المُفردات؛ بالإضافة إلى المُعجَمات العربيَّة المُمَثِّلة للُّغة عبرَ عُصُورها [العربيَّة القديمة، والعربيَّة الوسيطة، والعربيَّة الحديثة]. ويُستَمَدُّ ما تبقَّى من المُخرجات المُعجَمِيَّة من مواردَ أخرى – غير المُدَوَّنة اللُّغَويَّة المصنوعة لغرض صناعة المُعجَم. فمعلُوماتُ التَّهجئة مصدرُها معاجمُ الألفاظِ والمُصَنَّفاتُ في قواعد الإملاء والكتابة العربيَّة [السَّماعيَّة والقياسيَّة]؛ والمعلوماتُ الصَّوتيَّةُ مصدرُها القواعد والقوانين المُعتَمَدَة

في الأَلِفبائِيَّة الصَّوتِيَّة الدَّولِيَّة *INTERNATIONAL PHONETIC ALPHABET (IPA)* الَّتي أَقَرَّتها الجَمعِيَّة الصَّوتِيَّة الدَّولِيَّة بباريس *ASSOCIATION PHONÉTIQUE INTERNATIONALE* في عام 1888م. أمَّا المعلوماتُ التَّأثيليَّة فمصدرُها المعاجم اللُّغَويَّة التَّأثيليَّة للعربيَّة وقريناتها من اللُّغات السَّاميَّة؛ بالإضافة إلى المُصَنَّفات في الألفاظ المُعَرَّبة والدَّخيلة واللَّهجات العربيَّة.

وحتَّى تتَّضِحَ منهجيَّةُ حوسبة المُعجم التَّاريخيّ المنشُود، ينبغي أن نعرفَ أنَّ دورَ الحاسُوب هو توفيرُ الوقت والجهد والطَّاقة البشريَّة، وليسَ إنجازَ المُعجم بجميع مراحِلِه. إنَّهُ وسيلةٌ مُساعدةٌ فحسب، يُخضِعُها الإنسانُ لحاجتِه، ويتحكَّمُ فيها بالكيفيَّة الَّتي تُحقِّقُ أهدافَه. وعودٌ على ما سبقَ بشأن مراحلِ الصِّناعة المُعجَمِيَّة وموارد المُعجَم التَّاريخيّ ومُخرجاتِه، تقومُ منهجيَّةُ حوسبة المُعجم التَّاريخيّ للُّغة العربيَّة على ثلاث مراحلَ أساسيَّة، حيثُ تُعنى المرحلةُ الأولى بحوسبة المُدَوَّنة اللُّغَوِيَّة للمُعجَم عبرَ مرحلَتَي البناء والمُعالجَة؛ وينتُجُ عن هذه المرحلة المُخرجات المُعجَمِيَّة المُستمدَّة أساسًا من المُدَوَّنة دونَ غيرِها من الموارد. وتُعنى المرحلةُ الثَّانية بالتَّحرير المُعجَميّ؛ وينتُجُ عنها المُخرَج المُعجَميّ الكامل وقواعد بيانات المُعجَم بعدَ توحيد منهجها وإخضاعِها لما يُعرَفُ بـ "النَّمذجة المُعجَمِيَّة *LEXICAL MODELING*". وتُعنى المرحلةُ الثَّالثة بالنَّشر المُعجَمِيّ؛ وينتُجُ عنها الهيكل المُعجميّ النِّهائيّ [في صُورَتَيه: الورقيَّة والحاسُوبيَّة/الإلكترونيَّة] والبوَّابة الإلكترونيَّة للمُعجم. وسنعرضُ - فيما يلي - لهذه المراحل الثَّلاث تفصيلًا، مع الأخذ في الاعتبار أنَّ مراحلَ حوسبة المُعجم التَّاريخيّ تتوالى فيما بينها، وتتوازى مع مراحلَ أخرى من مراحل الصِّناعة المُعجَمِيَّة الَّتي لا يتَّسِعُ المقامُ للحديث عنها في هذه الدِّراسة.

3.1. حَوسبة المُدَوَّنة اللُّغَويَّة

3.1.1. مرحلة بنـــاء المُدَوَّنة اللُّغَوِيَّة للمُعجَم

مادَّةُ المُدَوَّنة اللُّغَوِيَّة ليست نُصوصًا تقييديَّةً أو عشوائية؛ لكنَّها كتلةٌ غيرُ منتظمةٍ من النُّصوص الَّتي تخضعُ لمجموعةٍ من الأسُس والمعايير، يُحَدِّدها الهدفُ المنشودُ منها[17]. ولَمَّا كانَ الهدفُ إنجازَ مُعجمٍ تاريخيٍّ للُّغة العربيَّة، لَزِمَ أن تكونَ مُدَوَّنَتُه اللُّغَوِيَّة مُرَقمَنة، بحيثُ يُمكنُ التَّحكُّمُ فيها آليًّا – بالإضافة أو الحذف أو التَّعديل. ولَزِمَ – كذلكَ – أن تكونَ كبيرةً نسبيًّا، لتعكسَ واقِعَ اللُّغة العربيَّة عبرَ تاريخِها المديد؛ وأن تكونَ مُدَقَّقةً إملائيًّا، بحيثُ لا تتجاوزُ نسبةُ الأخطاء فيها النِّسبةَ المسموح بها [وهي – من النَّاحية المعياريَّة – خمسة بالمئة من مجموع كلمات المُدَوَّنة]، وذلكَ التماسًا لدقَّة النَّتائج؛ وأن تكونَ مشكولةً في بعض أجزائِها، بصورةٍ كُلِّيَّة أو جُزئيَّة، ليُستفادَ منها في ضبط المداخل والوحدات والشَّواهد المُعجَمِيَّة؛ ولَزِمَ – كذلكَ – أن تكونَ المُدَوَّنةُ مُرَمَّزةً، من خلال ما يُعرَفُ بـ "ترميز المحارف *CHARACTER ENCODING*"، لتكونَ النُّصُوصُ قابلةً للتَّفاعُل مع قواعد البيانات ومُهَيَّئةً للمُعالجَةِ الآليَّة باستخدام الأدوات الدَّاعمة للُّغة العربيَّة[18].

يستدعي إعدادُ المُدَوَّنة اللُّغَويَّة للمُعجم المنشُود في صُورةٍ مُرَقمَنة أن يتمَّ تحويلُها من شكلِها الورقيّ الَّذي تبدو عليه في الكُتُب والمُصَنَّفات والوثائق إلى شكلٍ رَقْمِيّ مُنَقًّى من الحواشي والشُّرُوح والتَّذييلات والكشائد والفراغات الزَّائدة والحُرُوف اللاتينيَّة وغير ذلكَ ممَّا سوى المُتون النَّصِّيَّة الَّتي وَضَعَها مُصَنِّفو الكُتُب ومُؤلِّفُوها ممَّا يعكسُ طبيعةَ اللُّغة العربيَّة عبرَ تاريخِها. وبخلاف الوسيلة

---

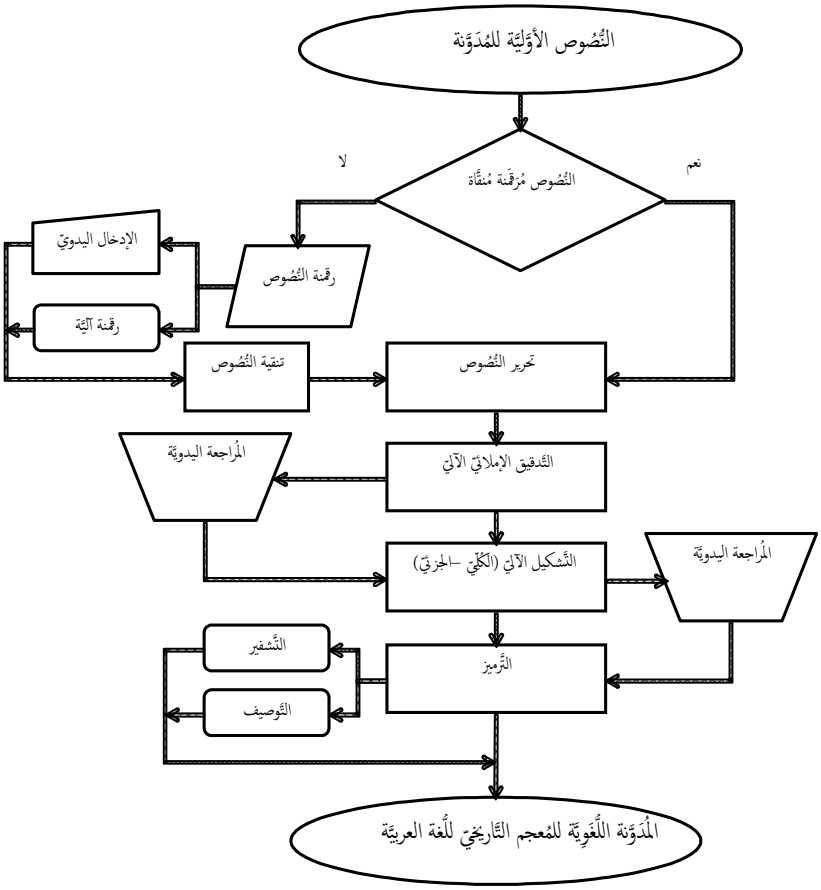
17 – للمزيد حولَ المُدوَّنات اللُّغويَّة ومناهجها وأنواعها، راجع: السَّعيد (الـمُعتزّ بالله): المدوَّنات اللُّغَوِيَّة، ضِمنَ كتاب (مُقَدِّمة في حوسَبة اللُّغة العربيَّة)، مجموعة من المؤلِّفين، تحرير: مُحسن رَشوان، والمعتزّ بالله السَّعيد، قَيد النَّشر بمدينة الملك عبد العزيز للعُلُوم والتِّقنية، الرِّياض، الجزء الأوَّل، 2014م.

18 – حولَ منهج بناء مُدوَّنة المُعجم التَّاريخيّ للُّغة العربيَّة والأدوات المُستخدمة، راجع: السَّعيد (المُعتزّ بالله): مُدَوَّنة مُعجَم تاريخيّ للُّغة العربيَّة، ص 33 : 52.

التَّقليديَّة لإدخال النُّصُوص إلى الحاسُوب، تُستخدمُ النُّصُوصُ المُتاحةُ إلكترونيًّا عبرَ صفحات الويب أو الأقراص المُدمجَة *CDs*؛ وتُتاحُ هذه النُّصُوصُ في العديد من الصِّيَغ أو الأنماط الشَّكليَّة، مثل: صِيغة الوثيقة المُتَنَقِّلة *(PDF)*، وصِيغة صَفحات الوِيب *(HTML)*، وصِيَغ الوثائق النَّصِّيَّة *(DOC& DOCX)*، والصِّيغة *(RTF)*، والصِّيغة *(TXT)*. كما يُمكنُ إدخالُ النُّصُوص آليًّا باستخدام الماسحات الضَّوئيَّة *SCANNERS* الَّتي تقومُ بإرسال الوَثائِق النَّصِّيَّة إلى الحواسيب بعدَ استِكشافِها ضَوئِيًّا، ثُمَّ التَّعَرُّف على الجرافيمات أو المحارف وتَحوِيلها من صِيغتها السَّاكنة إلى صِيغةٍ نَصِّيَّةٍ يُمكنُ التَّحكُّمُ فيها باستخدام تقنية القارئ الآليّ *OPTICAL CHARACTER RECOGNITION (OCR)*. أمَّا عن إدارة المُدَوَّنة اللُّغَويَّة الَّتي يُفتَرَضُ أن تحويَ قدرًا هائلًا من النُّصُوص، فهذا يستدعي استخدامَ المُحرِّرات النَّصِّيَّة المصنوعة خُصُوصًا للتَّعامُل مع مجموعات النُّصُوص الكبيرة، بالإضافة إلى قواعد البيانات النَّصِّيَّة الَّتي تُراعي طبيعةَ المُدَوَّنة اللُّغَويَّة والحجمَ التَّقريبيَّ الَّذي يُتَوقَّعُ أن تصِلَ إليه. ويستدعي تدقيقُ النُّصُوص استخدامَ مُدَقِّقٍ إملائيٍّ آليٍّ، تمَّ تدريبه/تجريبه – أثناءَ بنائِهِ – على وثائقَ نصِّيَّةٍ مُمثِّلةٍ للمُدَوَّنة اللُّغَويَّة الَّتي ننشُدُ تدقيقَها. وأخيرًا، يستدعي تشكيلُ/ضبطُ النُّصُوص [كُلِّيًّا أو جُزئيًّا] استخدامَ مُشكِّلٍ آليٍّ يستوعبُ المجموعات الكبيرة من النُّصُوص، بعدَ إمدادِهِ بالموارد اللُّغَويَّة من معاجمَ ومُدَوَّناتٍ لُغَويَّةٍ مشكولة.

تُمَثِّلُ هذه الإجراءات [رَقمنة المُدَوَّنة، وتحرير النُّصُوص، والتَّدقيق الإملائيّ، والتَّشكيل] مجموعةَ الإجراءات اللُّغَويَّة، حيثُ ينبغي أن تتوافرَ المعرفةُ اللُّغويَّة في القائمينَ عليها، وإن لم يكونوا لُغَوِيِّين أو مُعجَمِيِّين. أمَّا الإجراءُ الأخير [التَّرميز] فهو إجراءٌ حاسُوبيّ، يتمُّ – من خلالِه – تحويلُ نُصُوصِ المُدَوَّنة اللُّغويَّة مِن صُورَتِها النَّصِّيَّة الأوَّلِيَّة إلى صُورةٍ مَشرُوحَةٍ مُفَصَّلَة، أو بمعنى آخر: تحويلُها مِن نُصُوصٍ خامٍ غَيرِ مُعنونة *RAW TEXT (UNANNOTATED TEXT)* – وهُو الشَّكل الَّذي تظهَر فيه عِندَ تحريرِها أوَّلِيًّا – إلى نُصُوصٍ مُعنونة *ANNOTATED TEXT* مُذَيَّلةٍ بالخَصائِص الشَّكلِيَّة لنُصُوص المُدَوَّنة اللُّغَوِيَّة (وتَشتَمِلُ على: لُغة التَّرمِيز

المُستَخدَمة، وصِيغة تحوِيل المحارف، وأنواعُ الخُطوط وأحجامها وألوانها) والمعلومات الببليوغرافِيَّة (وتَشتَمِلُ على: عُنوان الوثِيقة، واسم المُصَنِّف، واسمُ المُحرِّر، وتاريخ التَّحرير). ويستدعي ترميزُ النُّصُوصِ في المُدَوَّنة اللُّغَوِيَّة تشفيرُها أوَّلًا بتحويلِها إلى إحدى صيغ نِظام الحُرُوف الدَّوْلِيّ المُوَحَّد *UNICODE*، مثل: (*UTF-8* و*UTF-16*)، أو صيغة تشفير المحارف العربيَّة *CP-1256*؛ ثُمَّ توصيفِها باستخدام إحدى لُغات التَّوصيف القياسيَّة، مثل: لُغة التَّوصِيف القابِلة للامتِداد *EXTENSIBLE MARKUP LANGUAGE (XML)*.



الشَّكل 1: مُخَطَّط انسيابي لحوسبة المُدَوَّنة اللُّغَويَّة في مرحلة البناء

3.1.2. مرحلة مُعالجَة المُدَوَّنة اللُّغَوِيَّة للمُعجَم

تُعالَجُ نُصُوصُ المُدَوَّنات اللُّغَوِيَّة المصنوعة لأغراض الصِّناعة المُعجَمِيَّة – من النَّاحِية المنهجيَّة – على مُستويين رئيسَين، هما: مُستوى المباني، ومُستوى المعاني. ولمَّا كانت المُدَوَّنةُ اللُّغَويَّة للمُعجم التَّاريخيّ موردًا وحيدًا للمداخل والوحدات ومُستويات الاستعمال والشَّواهد المُعجَمِيَّة، وموردًا أساسيًّا – مع مواردَ أخرى – للمعاني الوظيفيَّة [البِنَوِيَّة والتَّركيبيَّة] والمعاني المُعجَمِيَّة، فإنَّ هذه المُخرَجاتِ جميعًا تُمَثِّلُ الهدفَ المنشُودَ من مُعالجة المُدَوَّنة اللُّغَويَّة. وبعبارةٍ أخرى، تُستَخدَمُ في مُعالجة المُدَوَّنة اللُّغَوِيَّة للمُعجم التَّاريخيّ مجموعةٌ من الأدوات الحاسُوبيَّة والإحصائيَّة الَّتي تُساعدُ في تعيين مجموعة المُخرجات: (المداخل والوحدات المُعجَمِيَّة، ومُستويات الاستعمال والشَّواهد المُعجَمِيَّة، والمعاني الوظيفيَّة والمعاني المُعجَمِيَّة). ومع أنَّ بعضَ هذه المُخرَجات يتطلَّبُ الجمعَ بينَ مستوى المباني ومُستوى المعاني عندَ المُعالجَة [مثل: تعيين المعاني الوظيفيَّة] التماسًا للدَّرجة القُصوى من الدِّقَّة، إلَّا أنَّ واقعَ المُعالجة الآليَّة للُّغة العربيَّة لم يصل بعد إلى إتاحة ذلك. وعليه، تُصَنَّفُ هذه المُخرجاتُ إلى مجموعتين، حيثُ تشتملُ مُخرجاتُ المجموعة الأولى على: المداخل والوحدات المُعجَمِيَّة والمعاني الوظيفيَّة؛ وتُعالَجُ المُدَوَّنةُ آليًّا لتعيينها على مُستوى المباني. وتشتملُ مُخرجاتُ المجموعة الأخرى على: المعاني المُعجَمِيَّة ومُستويات الاستعمال والشَّواهد المُعجَمِيَّة والمعاني المُعجَمِيَّة؛ وتُعالَجُ المُدَوَّنةُ آليًّا لتعيينها على مُستوى المعاني. وتحقيقًا لهذا، تتطلَّبُ مُعالجَةُ المُدَوَّنة اللُّغَوِيَّة للمُعجَم تطويعَ آليةٍ لفهرسة النُّصُوص بهدف تعيين المداخل والوحدات المُعجَمِيَّة، وآليَّةٍ للتَّحليل الصَّرفيّ [البنويّ] وأخرى للتَّحليل التَّركيبيّ بهدف تعيين المعاني الوظيفيَّة، وآليَّةٍ للتَّحليل الدِّلاليّ لتعيين المعاني المُعجَمِيَّة؛ وأخيرًا سنكونُ بحاجةٍ إلى مِنَصَّةٍ حاسُوبيَّةٍ لاستخلاص الشَّواهد المُعجَمِيَّة وتعيين مُستويات الاستعمال. ونعرضُ فيما يلي لكلِّ آليَّةٍ على حِدة.

3.1.2.1. آليَّة فهرسة النُّصُوص.

الوظيفةُ الأساسيَّة للمُفهرس الآليّ [أو ما يُعرَفُ بالكشَّاف السِّياقيّ *CONCORDANCER*] حصرُ مُفردات النُّصُوص في المُدَوَّنة اللُّغَويَّة وبيان نسبة وُرُودِ كُلِّ مُفردةٍ على حِدة، مع إدراج جميع السِّياقات الَّتي تردُ فيها هذه المُفرَدات. وتوجدُ هذه الآليَّةُ في اللُّغة العربيَّة على ثلاثة أنواع: الفهرسة الألفبائيَّة – الَّتي تُرَتَّبُ فيها المُفرداتُ ألفبائيًّا، والفهرسة الجذعِيَّة – الَّتي تُرَتَّبُ فيها المُفرداتُ بحسب الجُذُوع *STEMS*، والفهرسة الجذريَّة – الَّتي تُرَتَّبُ فيها المُفرداتُ بحسب الجُذُور *ROOTS*. ومع ما يبدو من مُناسبة الفهرسة الجذريَّة لطبيعة الصِّناعة المُعجَمِيَّة العربيَّة – الَّتي تميلُ في مدارسِها المُعاصِرة إلى ترتيب المداخل المُعجَميَّة لمعاجم الألفاظ جذريًّا – إلَّا أنَّ الواقعَ خلافُ ذلك، لأنَّ بناءَ مُفهرسٍ جذريٍّ مُناسبٍ لطبيعة المُعجم العربيّ يتطلَّبُ تزويدَهُ بآليَّةٍ أخرى – عالية الدِّقَّة – لإزالة الالتباس الدِّلاليّ بينَ المُفردات؛ وهو أمرٌ لم تصل إليه اللُّغةُ العربيَّةُ بعد. فالمُفهرساتُ الجذريَّة المُتاحةُ للعربيَّة تواجهُ إشكالَ اشتراك المُفردة الواحدة في مجموعةٍ من الجُذور، الأمرُ الَّذي تلتزمُ معه بإيراد جميع هذه الجُذور في هيئة مُقترحات؛ ويكونُ على المُعالج أن يتدخَّلَ يدويًّا لاختيار المُقترَح الصَّحيح لكلِّ مُفردةٍ تحتملُ أكثرَ من جذر واحد. فالمُفردة (فَقَعَ) – على سبيل المثال – تحتملُ الجذرين (ف ق ع) و(و ق ع)؛ ويزيدُ الأمرُ صُعوبةً إذا خلَت المُفردةُ من علامات الضَّبط، حيثُ تزيدُ احتمالاتُ الجُذور، على نحو ما نجدُ – مثلًا – في المُفردة (وعد) الَّتي تحتملُ مجموعةَ الجُذور (و ع د) و(ع د د) و(ع ود) و(ع ي د). وعليه، تميلُ الدِّراسةُ إلى تعيين المداخل والوحدات المُعجَمِيَّة باستخدام آليَّة فهرسة النُّصُوص ألفبائيًّا أو جذعيًّا، ثُمَّ إعادة تعبئة المُفردات يدويًّا في قاعدة بياناتٍ جذريَّة، يُراعى فيها التَّرتيبُ المِعياريُّ للوحدات المُعجَمِيَّة؛ ويُمكنُ الاستعانةُ حينئذٍ بآليَّة التَّحليل الصَّرفيّ، الَّتي سيردُ الحديثُ عنها لاحقًا، حيثُ تُساعدُ على تجميع المُفردات ذات الجذر اللُّغَويّ الواحد. ويُوضِّحُ [الشَّكل 2] و[الشَّكل 3] نموذجين لمادَّتَين مُفهرَسَتَين ألفبائيًّا وجذعيًّا، كما يُوضِّح [الشَّكل 4] و[الشَّكل 5]

نموذجين لمُفهرسين [كشَّافَين سياقِيَّين] ألفبائيّ وجذعيّ على التَّرتيب، هما: المُفهرس الألفبائيّ *NOOJ CONCORDANCE* المُلحق بمِنصَّة *NOOJ* [19] والمُفهرس الجذعيّ *CONCAPP* [20]. والمُفهرسان يخضعان لنِظام الحُرُوف الدَّوْلِيّ المُوَحَّد *UNICODE* ويدعمان اللُّغةَ العربيَّة مع مجموعةٍ أخرى من اللُّغات الطَّبيعيَّة.

الشَّرَائِعَ، غَيَّرُوا الْفَرِيضَةَ، نَكَثُوا	**الْعَهْدَ**	الأَبَدِيَّ. لِذلِكَ لَعْنَةٌ أَكَلَتِ
وَيَصْغَى إِلَى الَّذِينَ تَرَكُوا	**الْعَهْدَ**	الْمُقَدَّسَ. وَتَقُومُ مِنْهُ أَذْرُعٌ
حَقِيرَةً وَلَا تَرْتَفِعَ، لِتَحْفَظَ	**الْعَهْدَ**	فَتَثْبُتَ. فَتَمَرَّدَ عَلَيْهِ بِإِرْسَالِهِ
الإِلهُ الْعَظِيمُ الْمَخُوفُ، الْحَافِظُ	**الْعَهْدَ**	وَالرَّحْمَةَ لِمُحِبِّيهِ وَحَافِظِي وَصَايَاهُ،
حَسَنًا . أَنَا أَقْطَعُ مَعَكَ	**عَهْدًا**	إِلَّا إِنِّي أَطْلُبُ مِنْكَ
فَقَطَعَ الْمَلِكُ دَاوُدُ مَعَهُمْ	**عَهْدًا**	فِي حَبْرُونَ أَمَامَ الرَّبِّ
وَأَخِي عَانِرَ . وَكَانُوا أَصْحَابَ	**عَهْدٍ**	مَعَ أَبْرَامَ . فَلَمَّا سَمِعَ
اِحْتَرِزُوا مِنْ أَنْ تَنْسَوْا	**عَهْدَ**	الرَّبِّ إِلهِكُمُ الَّذِي قَطَعَهُ
مِنْ أَجْلِ أَنَّهُمْ تَرَكُوا	**عَهْدَ**	الرَّبِّ إِلهِهِمْ وَسَجَدُوا لآلِهَةٍ
هِيَ الأَرْضُ؟ يَقُولُونَ :اقْطَعْ	**عَهْدَكَ**	مَعِي، وَهُوَذَا يَدِي مَعَكَ
بَنِي إِسْرَائِيلَ قَدْ تَرَكُوا	**عَهْدَكَ**	وَنَقَضُوا مَذَابِحَكَ، وَقَتَلُوا أَنْبِيَاءَكَ
الَّذِي ازْدَرَى قَسَمَهُ وَنَقَضَ	**عَهْدَهُ**	فَعِنْدَهُ فِي وَسْطِ بَابِلَ

19 – الموقع الإلكترونيّ للبرمجيَّة:

HTTP://WWW.NOOJ4NLP.NET.

20 – الموقع الإلكترونيّ للبرمجيَّة:

HTTP://CONCAPP.SOFTWARE.INFORMER.COM.

لِشَعْبِهِ . أَقَامَ إِلَى الأَبَدِ	**عَهْدَهُ**	قُدُّوسٌ وَمَهُوبٌ اسْمُهُ .رَأْسُ
أَحْكَامُهُ . ذَكَرَ إِلَى الدَّهْرِ	**عَهْدَهُ**	كَلَامًا أَوْصَى بِهِ إِلَى
الرَّبِّ إِلٰهِهِمْ، بَلْ تَجَاوَزُوا	**عَهْدَهُ**	وَكُلَّ مَا أَمَرَ بِهِ

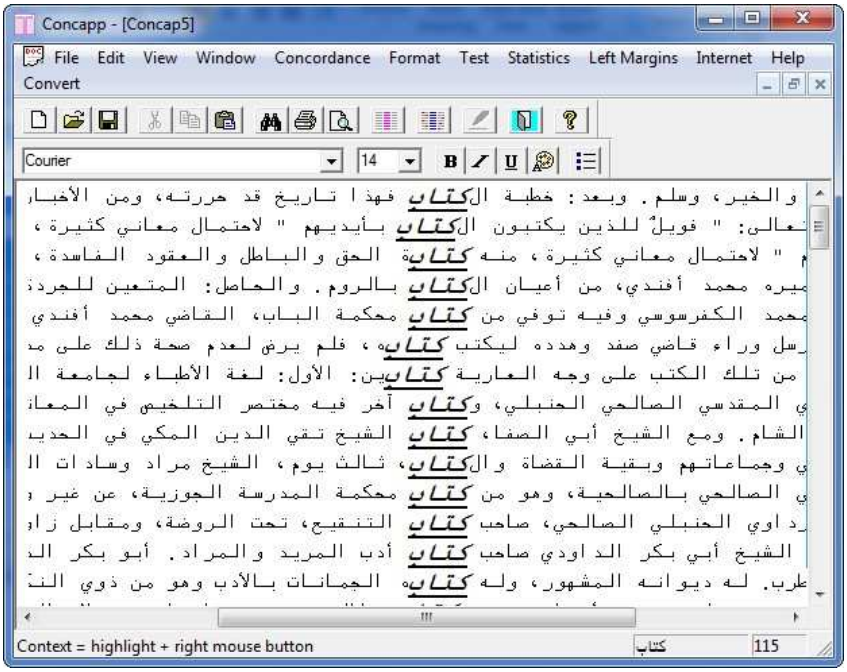
الشَّكل 2: نموذج لمادة مفهرَسة ألفبائيًا – من: العهد القديم

1	وبرز في علم القراءات وعمر م**درس**ة للقراء وسماها دار القرآن
2	أنه لم يراجع شيئًا حتى أورد **درس**ه ومثله لايستكثر عليه ذلك
3	وهمذان ودمشق وحران من خلائق و**درس** بالقدس الشريف في الصلاحية
4	ثلاث عشرة سنة انتهى. ثم **درس** بالرواحية وهو أول من **درس** بها
5	بعد أبيه مدةً قليلة ثم عزل و**درس** بالغزالية مدة كما سيأتي وباشر
6	أنه لم يراجع شيئًا حتى أورد **درس**ه ومثله لايستكثر عليه ذلك
7	سنة تسع وأربعين فسكنت الم**درس**ة الرواحية وبقيت سنتين لم أضع
8	الحسن الفارقي خطيب دمشق وم**درس** الشامية والناصرية الجوانية
9	أكبر والفضلاء أكثر كان ال**درس** أنظر وأنضر وأحلى وأنصح وأفصح
10	الشيخ كمال الدين بالناصرية ي**درس** بها عشرين سنة ثم انتزعها من
11	جمال الدين أبن القلانسي و**درس** في الناصرية كمال الدين بن
12	يومًا ثم عاد إلى القضاء وقد **درس** بمصر والشام بمدارس كبار

13	الأرض وكان قوتي بها جراية المدرسة لا غير وحفظت التنبيه في نحو
14	سمع منه البرزالي وغيره ودرس بمدرسة جدة وبدار الحديث
15	وتخرج به الفقهاء انتهى ثم درس بها ولده بعده عز الدين. قال
16	الدرس ويتكلم الحاضرون فيه ودرس بالجوزية وكان بيده نصف

الشَّكل 3: نموذج لمادة مفهرسة جذعيًا – من كتاب: الدَّارس في تاريخ المدارس، للنُّعيميّ

الشَّكل 4: المُفهرس الآليّ الألفبائيّ *NOOJ CONCORDANCE* – المادَّة من: القُرآن الكريم

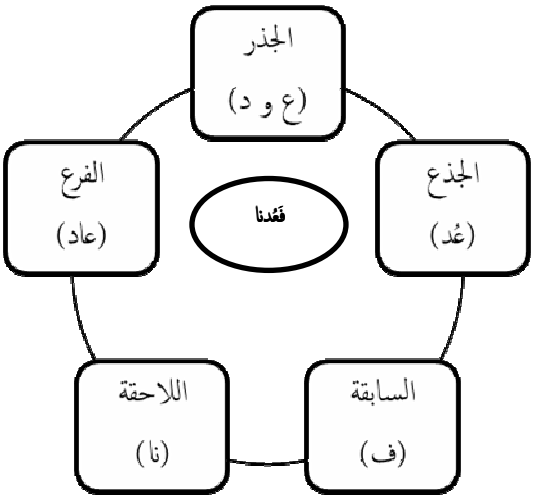


الشَّكل 5: المُفهرس الآليّ الجذعيّ *CONCAPP* – المادَّة من: يوميَّات شاميَّة، لابن كنان

3.1.2.2. آليَّة التَّحليل الصَّرفيّ.

يُستخدم المحلِّل الصَّرفيّ/ البِنويّ *MORPHOLOGICAL ANALYZER* في تحليل الكلمة إلى عناصِرِها الأوَّليَّة الخمسة (الجذر، والجذع، والفرع، والسَّابقة، واللَّاحِقة)، وهي العناصرُ الَّتي تتكوَّنُ عنها مجموعةَ الوحدات الصَّرفيَّة المُجَرَّدة (المورفيمات *MORPHEMES*)؛ كما يُعنى بالتَّوصيف الصَّرفيّ/ البِنويّ لكلِّ مُفردةٍ على حِدة[21].

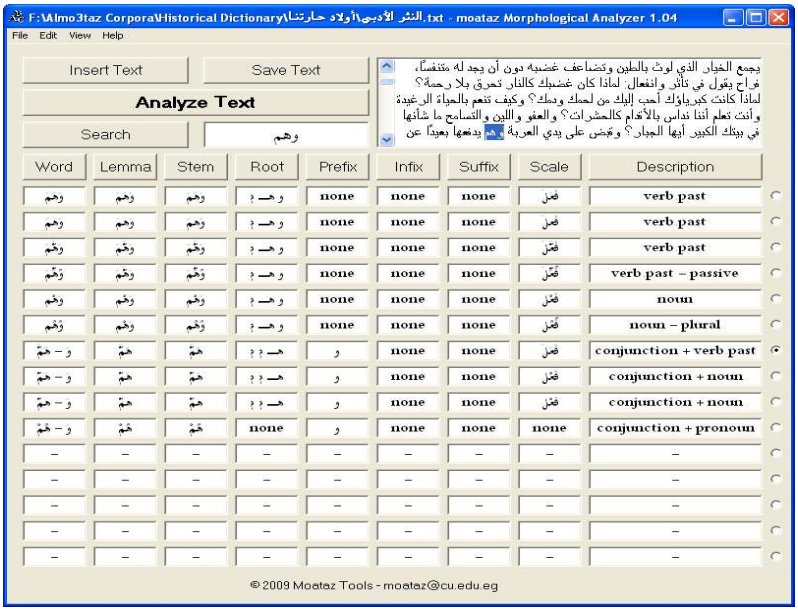
21 – للمزيد حولَ آليَّة التَّحليل الصَّرفيّ، راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص 84 : 94.



الشَّكل 6: العناصر الأوَّليَّة المكوِّنة للكلمة العربيَّة – نموذج: [فَعُدنا]

والمُحلِّلُ الصَّرفيّ – بذلكَ – يُتمِّمُ عملَ المُفهرس الآليّ من ناحية، حيثُ يُساعدُ في تعيين المداخل والوحدات المُعجَمِيَّة بإحالة المُفردات إلى جُذورها [باعتبارها المداخل المُعجميَّة] وإلى فُروعها [باعتبارها الوحدات المُعجَمِيَّة]. ومن ناحيةٍ أخرى يؤدِّي دورَهُ الأساسيّ في تعيين المعاني الوظيفيَّة البنويَّة [السَّماعيَّة، كأبواب الفعل الثُّلاثيّ المُجرَّد]. وتجدُرُ الإشارةُ إلى وُجود آليَّتين تتفرَّعان عن آليَّة التَّحليل الصَّرفيّ، هما: آليَّة التَّجذيع *STEMMER*، الَّتي تُشبهُ المُفهرسَ الآليَّ الجذعيَّ حيثُ تُعنى بتعيين جُذُوع المُفردات، وإن لم تهتمَّ بترتيبها وحصر السِّياقات الَّتي تردُ فيها؛ وآليَّة التَّفريع *LEMMATIZER*، ووظيفتُها تعيين فُرُوع الكلمة *LEMMATA* الَّتي تنتُجُ عنها الوحدات المُعجَمِيَّة *LEXEMES*. ويُفترضُ في أدوات التَّحليل الصَّرفيّ للعربيَّة – عُمومًا – أن تُراعيَ النِّظامَ الكتابيَّ للعربيَّة من حيث ضبط وتشكيل الحروف، لتكونَ قادرةً على التَّحليل الصَّرفيّ للمُفردات في جميع حالاتها الَّتي تُوجدُ عليها، سواءٌ أكانت مشكولةً كُلِّيًّا أم جُزئيًّا أم غيرَ مشكولة، كما في [الشَّكل 7][22].

22 – الآليَّة خاصَّة بالباحث. للمزيد حولها، راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص 89.



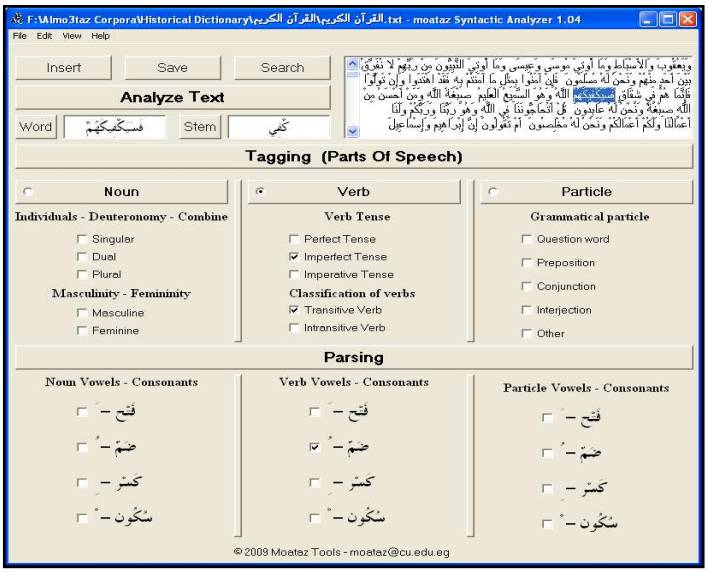
الشَّكل 7: آلية التحليل الصرفي *(MORPHOLOGICAL ANALYZER 1.04)* – المادَّة من: أولاد حارتنا، لنجيب محفوظ

3.1.2.3. آليَّة التَّحليل التَّركيبيّ.

يُعنى المُحلِّلُ التَّركيبيُّ *SYNTACTIC ANALYZER* في العربيَّة بتعيين أقسام الكلام *PARTS OF SPEECH (POS)* وبيان وظائِفِها في الجُملة [تحليل أقسام الكلام *POS TAGGING*]، ثُمَّ توصيف هذه الأقسام – تركيبيًّا – بتعيين حركاتِها في حالتَي الإعراب والبناء [إعراب أقسام الكلام *POS PARSING*][23]. والواقعُ أنَّ اللُّغةَ العربيَّةَ تُعاني قُصُورًا كبيرًا في آليَّة التَّحليل التَّركيبيّ بالمفهوم الدَّقيق لهذه الآليَّة. ذلكَ أنَّ بناءَ مُحلِّلٍ تركيبيٍّ قادرٍ على تعيين وظائف أقسام الكلام وتوصيفها تركيبيًّا يتطلَّبُ تزويدَهُ بآليَّةٍ أخرى – عالية الدِّقَّة – لإزالة الالتباس الدِّلاليّ بينَ المُفردات؛ وهو الإشكالُ ذاته الَّذي عرضنا له سلفًا عندَ الحديث عن المُفهرس الآليّ الجذريّ. إلَّا أنَّ ذلكَ الأمرَ لا يُمثِّلُ عقبةً كبيرةً بالنَّظر إلى طبيعة الصِّناعة المُعجَمِيَّة

23 – للمزيد حولَ آليَّة التَّحليل التَّركيبيّ، راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص 77 : 83.

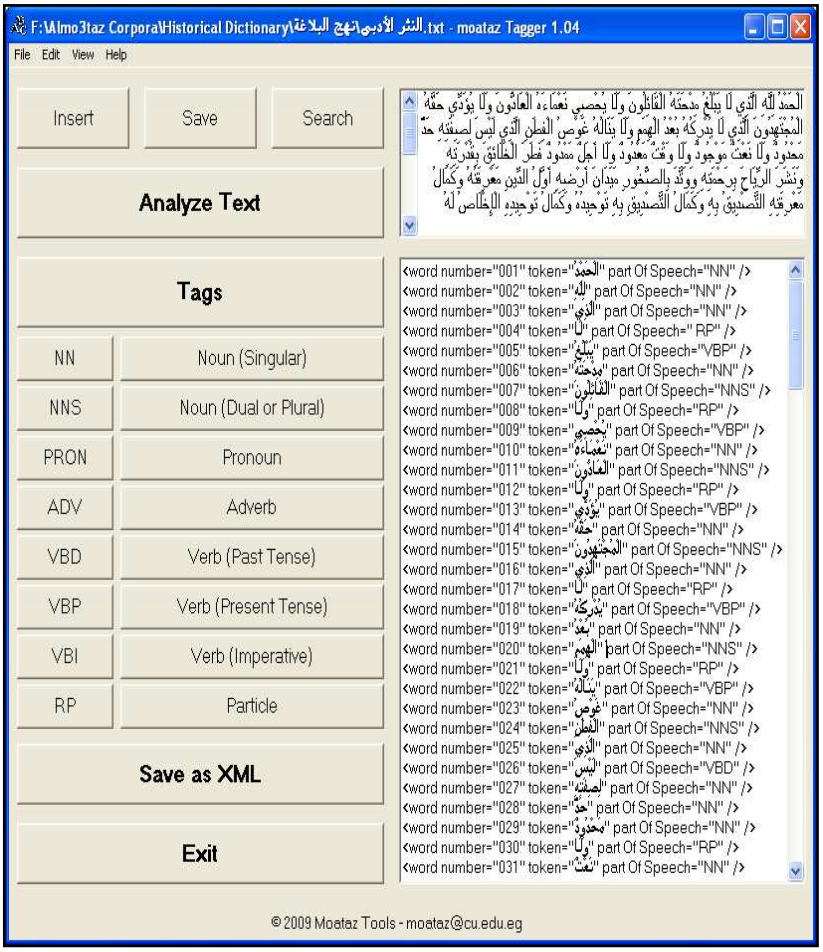
من ناحية وإلى المعاني الوظيفيَّة المُبتغاة من ناحيةٍ أخرى. فما نحنُ بحاجةٍ إليه فعليًّا عندَ التَّحليل التَّركيبيّ لنُصُوص المُدَوَّنة اللُّغَويَّة هو تعيينُ أقسام الكلام، وهو أمرٌ تُتيحُه أدواتُ التَّحليل التَّركيبيّ – مع التَّدخُّل اليدويّ لضبط المُخرَجات. أمَّا المعاني الوظيفيَّة التَّركيبيَّة، فتنقسمُ إلى قسمَين: المعاني القياسيَّة، وهذه مصدرُها مُصنَّفاتُ النَّحو العربيّ ولا حاجةَ للمُعجم إليها، والمعاني السَّماعيَّة [كاللَّازم والمُتعدِّي من الأفعال]، وهذه الأخرى تغلُبُ عليها المُعالجة اليدويَّة كونها تُستنبَطُ من الشَّواهد المُعجميَّة ولا تقومُ على قواعدَ مُطَّردة. ويُوضِّحُ [الشَّكل 8] نموذجًا للمُحلِّل التَّركيبيّ[24]، كما يُوضِّحُ [الشَّكل 9] نموذجًا لآليَّة تحليل أقسام الكلام[25]، وهي الَّتي تقترحُ الدِّراسةُ الاقتصارَ عليها عندَ إخضاع مُدوَّنة المُعجم التَّاريخيّ للتَّحليل التَّركيبيّ.



الشَّكل 8: آلية التحليل التركيبي *SYNTACTIC ANALYZER 1.04* – المادَّة من: القُرآن الكريم

24 – الآليَّة خاصَّة بالباحث. للمزيد حولها، راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص 83.

25 – الآليَّة خاصَّة بالباحث. للمزيد حولها، راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص 82.



الشَّكل 9: آلية تحليل أقسام الكلام *PoS Tagger 1.04* – المادَّة من: نهج البلاغة، المنسُوب للإمام عليّ

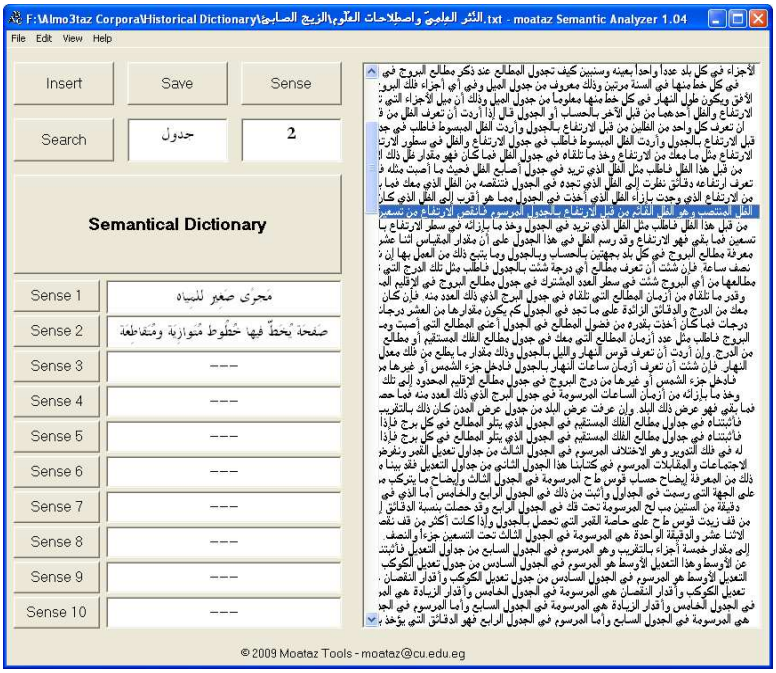
3.1.2.4. آليَّة التَّحليل الدِّلاليّ.

تُعتَبَرُ آلِيَّةُ التَّحلِيل الدِّلاليّ *Semantic Analyzer* إحدى تطبيقات الذَّكاء الاصطِناعِيّ المُتَقَدِّمة، حيثُ ترمي إلى إخضاع الآلة لفهمِ مُفرَدات اللُّغَةِ في مجموعةٍ من السِّياقات والأنماط التَّركِيبِيَّة بالاعتِمادِ على مُعطَياتٍ سابِقَةٍ يُحَدِّدُها الهَدَفُ الَّذي يُبنَى لأجلِهِ المُحَلِّلُ الدِّلاليّ. ولَمَّا كانَت المُدَوَّنةُ اللُّغويَّةُ موردًا أساسيًّا

للمعاني المُعجميَّة لوحدات المُعجم المنشود، فإنَّنا سنكونُ بحاجةٍ إلى أداةٍ حاسُوبيَّةٍ [تعملُ عملَ المُحلِّل الدِّلاليّ] قادرةٍ على تَعيين دِلالات المُفرَدات ومعانيها المُعجميَّة بتحلِيل سِياقاتِها في المُدَوَّنة اللُّغويَّة بما تحويه من النُّصُوص. ومن النَّاحية العمليَّة في حوسبة اللُّغة، تتبلورُ مثلُ هذه الأداة فيما يُعرفُ بآليَّة فَكِّ الالتباس الدِّلاليّ للكَلِماتِ *(WORD SENSE DISAMBIGUATION (WSD*. وَتَستَمِدُّ هذه الآليَّةُ فِكرَتَها الأساسيَّةَ مِن المُتَصاحِبات اللَّفظِيَّة *COLLOCATIONS* الَّتي تَتَكَوَّنُ عن سِلسِلَةٍ مِن كلِمَتَين أو أكثَر، تَتَلازَمُ مُفرَداتُها في عَلاقَةٍ تَركِيبِيَّة، كَعَلاقَة الفِعل والفاعِل (مِثل: صاحَ الدِّيك، سادَ الصَّمْت) وعَلاقَة الفِعْل والمَفعُول (مِثل: قَدَّمَ استِقالَتَهُ، أحرَزَ هَدَفًا) وعَلاقَة الإضافة (مِثل: عابر سَبِيل، قاطِع طَرِيق). والواقعُ أنَّ الآلِيَّةَ تأخُذُ بُعدًا آخَرَ إذ يَكُونُ التَّلازُمُ فيها على مُستوى السِّياقِ بأكمَلِه؛ فالكَلِمَةُ العَرَبِيَّةُ تَرتَبِطُ في دِلالاتِها بِكلِماتٍ أخرى في حَيِّز الجوار، سواءٌ أكانَت سابِقَةً لها أم لاحِقَةً بِها، وسواءٌ أكانَت مُصاحِبَةً لها أم مُنفَصِلةً عنها[26] ؛ ويَتَكَرَّرُ التَّصاحُبُ بينَ الكَلِمَة ومُتجاوراتها بصُورةٍ كبيرَةٍ نسبِيًّا، الأمر الَّذي يجعلُنا قادِرين على الرَّبط بين الكلمات الوارِدةِ في مجموعةٍ من السِّياقات، ويَقُودُنا ذلك إلى تَعيين المعانِي المُعجَمِيَّة لمُفرَدات النُّصُوص. وعليه تَقُومُ الفِكرَةُ الأساسِيَّةُ لآلِيَّة فَكِّ اللَّبس الدِّلالِيّ *WSB* على استِدعاءِ مَعنى الكَلِمَة مُتَعَدِّدَة الدِّلالات من خِلال الكَلِمات المُصاحِبَة لها في السِّياق *CONTEXT*، إذ يَغلُبُ على هذه الكلِمات أن تَشتَرِكَ مَعَها في حَقلٍ دِلالِيٍّ واحِد.

26 – السَّعيد (المُعتزّ بالله): السَّابق، ص 102، بتصرُّف. وللباحث تجربتان في بناء هذه الآليَّة، إحداهما مُنفردًا، يُمكن الاطِّلاعُ عليها في: السَّعيد (المُعتزّ بالله): السَّابق، ص 107 : 109، وأخرى مع فريق عمل، يُمكنُ الاطِّلاعُ عليها في:

Eid, S. M. & Al-Said, A. B. & 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



الشَّكل 10: نموذج لعمل آليَّة التَّحليل الدِّلاليّ *SEMANTIC ANALYZER 1.04*[27] – المادَّة من: الزَّيج الصَّابئ، للبتَّاني

3.1.2.5. المِنَصَّة الحاسُوبيَّة للمُدَوَّنة.

يُقصَدُ بالمِنَصَّة الحاسُوبيَّة *COMPUTATIONAL FRAMEWORK* ذلكَ المُجتَمَع الَّذي يَضُمُّ عددًا من الموارد اللُّغَوِيَّة والمُعجَمِيَّة والحاسُوبيَّة الَّتي تتلاحمُ مع بعضِها لتقُومَ بمُعالجَاتٍ تتحَقَّقُ من خِلالِها الغايةُ من المُدَوَّنة اللُّغَوِيَّة. وتُعَدُّ – بذلكَ – موردًا حاسُوبيًّا بالغَ الأهمِّيَّة للمُعجم التَّاريخيّ المنشود. ذلكَ أنَّ المِنَصَّات الحاسُوبيَّة تُزَوَّدُ بمُحرِّكاتٍ بحثيَّة *SEARCH ENGINES* تصلُحُ وسيلةً لاستخلاص الشَّواهد المُعجَمِيَّة وتعيين مُستويات الاستعمال اعتمادًا على إحصاءاتِها. وبالإضافة إلى إمكانيَّة توظيف المنصَّة لتكونَ قاعدةَ بياناتٍ بحثيَّة لنُصُوص

27 – للمزيد حولَ آليَّة التَّحليل الدِّلاليّ، راجع: السَّعيد (المُعتزّ بالله): السَّابق، ص 111.

المُدَوَّنة اللُّغَويَّة، يُمكنُ توظيفُها – كذلكَ – في تجميع أدوات المُعالجة الآليَّة للنُّصُوص والانتقال من مرحلةٍ إلى أخرى بصورةٍ مُنتظمة. وثمَّةَ العديدُ من المِنَصَّات الحاسُوبيَّة الَّتي تدعمُ اللُّغةَ العربيَّةَ ويُمكنُ الإفادةُ منها في المُعجم المنشود؛ لعلَّ أبرزَها مِنَصَّات *NOOJ* [28]، و*GATE* [29]، و*SAFAR* [30]. ويُوضِّحُ [الشَّكل 11] نموذجًا لمُدَوَّنةٍ لُغويَّةٍ تاريخيَّة مُدرجة في مِنصَّة *NOOJ* تمهيدًا لمُعالجتها واستخلاص شواهدها.

الشَّكل 11: نموذج توضيحيّ لمُدَوَّنةٍ لُغويَّةٍ تاريخيَّة مُدرجة في مِنصَّة *NOOJ*

28 – الموقع الإلكترونيّ للمنصَّة:

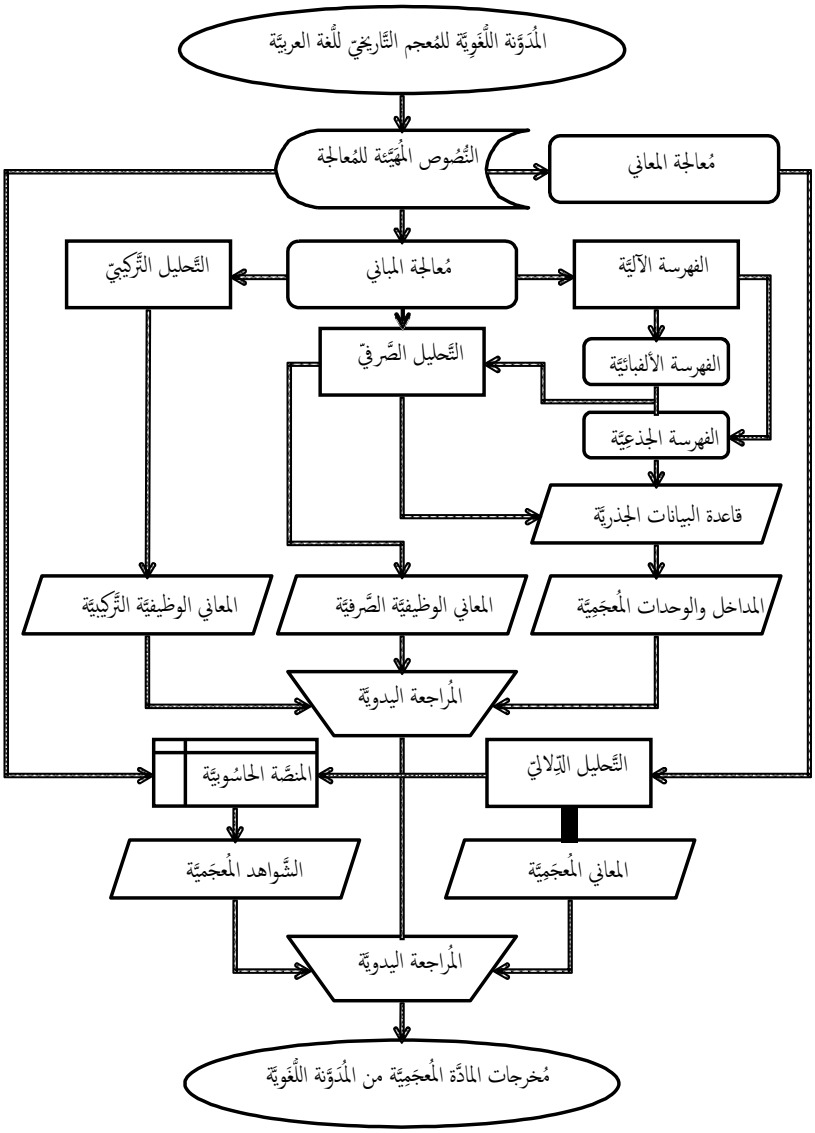
HTTP://WWW.NOOJ4NLP.NET.

29 – الموقع الإلكترونيّ للمنصَّة:

HTTP://GATE.AC.UK.

30 – الموقع الإلكترونيّ للمنصَّة:

HTTP://SIBAWAYH.EMI.AC.MA/SAFAR.



الشَّكل 12: مُخَطَّط انسيابي لحوسبة المُدَوَّنة اللُّغَويَّة في مرحلة المُعالجة

3.2. التَّحرير المُعجَميّ

يقومُ تحريرُ المعاجم الحديثة على نظام التَّحرير المُعجَميّ *DICTIONARY WRITING SYSTEM (DWS)* – أو ما يُعرَفُ بنظام صناعة المُعجم *DICTIONARY PRODUCTION/PUBLISHING SYSTEM (DPS)*. وهو بيئةٌ حاسُوبيَّةٌ تفاعُليَّة، تشتملُ على قاعدة بياناتٍ للمعلومات المُعجميَّة وأداةٍ للتَّحرير المُعجميّ وأدواتٍ لإدارة الموارد اللُّغَوِيَّة والمُعجَمِيَّة والحاسُوبيَّة المُستخدمة في صناعة المُعجم، بالإضافة إلى الواجهة التَّفاعُليَّة لنظام التَّحرير. ويُساعدُ نظامُ التَّحرير المُعجَميّ على إخراج هيكل المُعجم في صيغةٍ قياسيَّةٍ ومُنتظمة تتوحَّدُ فيها مناهج المُحرِّرين المُشاركينَ في صناعة المُعجم[31]. ومع أنَّ صناعةَ المُعجم العربيّ لم تعرف الطَّريقَ إلى مثل هذه الأنظمة الحاسُوبيَّة، إذ تميلُ إلى التَّأليف المُعجميّ المُغاير لمفهوم الصِّناعة، إلَّا أنَّ طبيعةَ المُعجم المنشود للعربيَّة تقتضي السَّعيَ إلى الإفادة من نظامٍ تحريريٍّ يُساعدُ على توفير الوقت والجهد من ناحية، وتوحيد منهج صناعة المُعجم، وإن تعاقبت عليه أجيالُ المُحرِّرينَ، من ناحيةٍ أخرى. وثمَّةَ العديدُ من أنظمة التَّحرير الَّتي يُمكنُ الإفادةُ منها في المُعجم التَّاريخيّ، مثل: نظام التَّحرير المفتوح *MĀTĀPUNA* [32]، ونظام *LÉACSLANN* [33]، والنِّظام الأشهر *TSHWANELEX* [34]، وغيرها من الأنظمة التَّفاعُليَّة. ويُوضِّح [الشَّكل 13] هيكلًا توضيحيًّا لأحد هذه الأنظمة.

---

31 – 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.

32 – البرمجيَّة مُتاحةٌ على موقع البرمجيَّات الحرَّة. *SOURCEFORGE*؛ ورابطها:

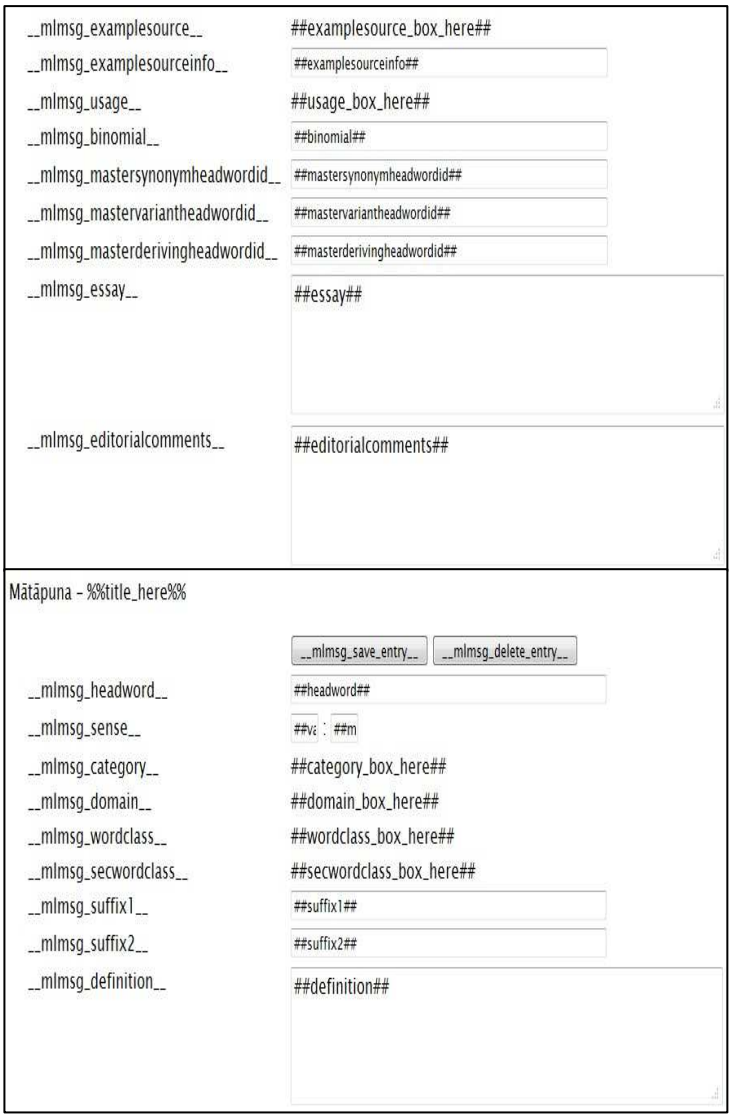
HTTP://SOURCEFORGE.NET/PROJECTS/MATAPUNA/.

33 – الموقع الإلكترونيّ للبرمجيَّة:

HTTP://LXLN.PRETTYDATA.EU/.

34 – الموقع الإلكترونيّ للبرمجيَّة:

HTTP://TSHWANEDJE.COM/TSHWANELEX/.



الشَّكل 13: هيكل توضيحيّ لنظام التَّحرير المُعجميّ – مفتوح المصدر *MĀTĀPUNA-0.47*

وسنُحاولُ – فيما يلي – أن نعرضَ مدى الإفادة من أنظمة التَّحرير المُعجميّ في تعيين المعلومات المُعجميَّة الأساسيَّة في المُعجم التَّاريخيّ المنشُود للعربيَّة، كُلٌّ منها على حِدة.

3.2.1. المداخل والوحدات المُعجَميَّة *(Entries & Headwords (Lexemes*

تُستمدُّ المداخلُ والوحدات المُعجميَّة من قاعدة البيانات الجذريَّة للمُدَوَّنة اللُّغويَّة بعدَ مُراجعتها يدويًّا. وبالتَّالي لا يُستفادُ من نظام التَّحرير المُعجميّ فيها بصورةٍ مُباشرة. ويقتصرُ دورُ النِّظام على أن يعملَ وعاءً للمداخل والوحدات، فيُساعد على ترتيبها وتنظيمها وعرضِها أوَّلًا، ويُساعدُ على الرَّبط بينَها وبينَ المعلومات المُعجميَّة الأخرى ثانيًا.

3.2.2. معلومات الهجاء والنُّطق *Alphabet & Pronunciation*

لا تميلُ المُعجماتُ العربيَّةُ إلى العناية بمعلومات الهجاء والنُّطق، إذ تفترضُ أن يُقَدَّمَ المُعجمُ للقارئ العربيّ دونَ غيره، كما تفترضُ المعرفةَ الكاملة بقواعد الهجاء والنُّطق العربيَّة لدى هذا القارئ. وهما أمران ينفيهما الواقعُ الَّذي يؤكِّدُ عنايةَ المُستشرقينَ والباحثينَ في معارف الشَّرق بالمعاجم العربيَّة، ويؤكِّدُ – كذلكَ – صعوبةَ البحث والقراءة في المعاجم العربيَّة على القارئ العربيّ. ومع أنَّ معلومات الهجاء والنُّطق تخضعُ لقواعد الألفبائيَّة الصَّوتيَّة الدَّوليَّة، إلَّا أنَّ أكثرَ هذه المعلومات قياسيَّة، الأمرُ الَّذي ساعدَ على استخدام ما يُعرَفُ بآليَّة رومنة الحُرُوف العربيَّة *Romanization of Arabic* في تعيين هذه المعلومات. وتوجدُ هذه الآليَّة في أدواتٍ مُستقلَّة للرَّومنة *Romanizer*، أو مُلحقة بأدواتٍ وأنظمةٍ أخرى على النَّحو الوارد في [الشَّكل 14]. ويُستفادُ من نظام التَّحرير المُعجميّ في تعيين هذه المعلومات وإدراجها في قاعدة البيانات المُعجميَّة، إمَّا من خلال آليَّةٍ مُلحقةٍ به لرومنة الحُرُوف العربيَّة أو بتوظيف النِّظام ليعملَ أداةَ ربطٍ بينَ أداةٍ مُستقلَّة للرَّومنة وقاعدة البيانات المُعجميَّة في نظام التَّحرير، مع ضرورة مُراجعة المعلومات يدويًّا قبلَ إدراجها في مُخرجات المادَّة المُعجَميَّة.

```
INPUT STRING: النافذة
LOOK-UP WORD: AlnAf*p
  SOLUTION 1: (AlnAfi*ap) [nAfi*_1]
Al/DET+nAfi*/ADJ+ap/NSUFF_FEM_SG
    (GLOSS): the + effective/operative/valid + [fem.sg.]
  SOLUTION 2: (AlnAfi*ap) [nAfi*ap_1]
Al/DET+nAfi*/NOUN+ap/NSUFF_FEM_SG
    (GLOSS): the + window + [fem.sg.]
```

الشَّكل 14: نموذج توضيحيّ لرومنة الحروف العربيَّة في نظام *AraMorph 1.2.1* الصَّرفيّ

3.2.3. المعلومات التَّأثيليَّة *Etymological Information*

لا يتَّسعُ المقامُ – كما لا تسمحُ طبيعةُ هذه الدِّراسة – للحديث عن إشكالات المعلومات التَّأثيليَّة في المُعجم التَّاريخيّ المنشود للعربيَّة. وخُلاصة ما يعنينا أن المعلومات التَّأثيليَّة هي تلكَ الَّتي تؤصِّلُ للمُفردات في مرحلة ما قبلَ التَّأريخ، أو بالأحرى قبلَ أن تجريَ المُفرداتُ في النُّصُوص المكتوبة وعلى ألسنة النَّاطقينَ بها. وبما أنَّ المكتبةَ العربيَّةَ تكادُ تخلو من المعاجم التَّأثيليَّة الَّتي يُمكنُ الإفادةُ منها في تأثيل مُفردات المُعجم، فإنَّنا سنكونُ بحاجةٍ إلى موارد أخرى – لُغويَّة ومُعجميَّة – كما سنكونُ بحاجةٍ إلى منهجيَّةٍ واضحةٍ لتأثيل المُفردات. ولأنَّ المعلومات التَّأثيليَّة تخرُجُ عن إطار الحوسبة، فمن المُفترض أن يقتصرَ دورُ نظام التَّحرير المُعجميّ على إدراجها في قاعدة البيانات المُعجميَّة وربطها بالوحدات المُتضَمِّنة لها.

3.2.4. المعاني الوظيفيَّة [النَّحويَّة] *Grammatical Functions*

عرضنا آنفًا لطبيعة المعاني الوظيفيَّة من حيثُ إنَّها تُعنى بجانِبَي المنظومة النَّحوِيَّة: البِنيَة والتَّركِيب. ولمَّا كانت المعلومات النَّحويَّة [أو المعاني الوظيفيَّة]

المُتَضَمَّنة في العمل المُعجميّ – بشكلٍ عامّ – قاصرةً في أكثرها على الجانب السَّماعيّ دونَ القياسيّ، فقد حَرَصَ مُطَوِّرو أدوات التَّحليل الصَّرفيّ والتَّركيبيّ للنُّصُوص العربيَّة على تزويد هذه الأدوات بقواعد بياناتٍ تتضمَّنُ كثيرًا من المعاني الوظيفيَّة السَّماعيَّة، لتعملَ هذه القواعدُ – جنبًا إلى جنب – مع خوارزمات التَّحليل النَّحويّ للمعاني الوظيفيَّة القياسيَّة. وبسبب عدم تضمين الجانب السَّماعيّ في هذه الأدوات بصورةٍ كُلِّيَّة، لزمَ أن يُستفادَ من مواردَ أخرى في تعيين المعاني الوظيفيَّة للمُعجم. وعليه فإنَّ المعاني الوظيفيَّة للوحدات المُعجميَّة تُستمدُّ من المُدَوَّنة اللُّغويَّة للمُعجم بعدَ مُعالجتها صرفيًّا وتركيبيًّا، بالإضافة إلى موارد أخرى، سواءٌ أكانت المُصَنَّفاتُ في قواعد النَّحو العربيّ أم المعاجم اللُّغويَّة المُتقدِّمة. ويسمحُ نظامُ التَّحرير المُعجميّ ببناء هيكلٍ معلوماتيٍّ للمعاني الوظيفيَّة النَّاتجة عن التَّحليل الصَّرفيّ والتَّركيبيّ لمُفردات المُدَوَّنة اللُّغويَّة؛ كما يؤدِّي دورَهُ في استقبال المعاني الوظيفيَّة السَّماعيَّة المُستمدَّة من مواردَ أخرى وتعبئتها في قاعدة البيانات المُعجميَّة.

3.2.5. المعاني المُعجميَّة *LEXICAL MEANINGS*

المَعنَى المُعجَمِيُّ هو ذلكَ المعنى الَّذي يَخضَعُ للعُرف العامّ بينَ الجماعة اللُّغَوِيَّة. ولأنَّ المُدَوَّنةَ اللُّغويَّةَ للمُعجم المنشود هي التَّمثيل الحقيقيُّ لذلك العُرف، فهي المورِدُ الأساسيُّ الَّذي ينبغي أن نستمدَّ منهُ المعاني المُعجميَّة. أمَّا عن تحرير هذه المعاني وصياغتها، فيُمكن الإفادةُ من المُعجمات اللُّغويَّة للعربيَّة عبرَ عُصُورِها. وتؤدِّي أنظمة التَّحرير المُعجميّ دورًا بالغَ الأهمِّيَّة في تحرير المعاني المُعجميَّة؛ إذ تُدرَجُ المعاني فيها بأسلوبٍ منهجيٍّ يتَّفقُ مع ما ينبغي أن تكونَ عليه الصِّناعةُ، لا على الواقع الَّذي يطغى فيه التَّأليفُ بصورةٍ عشوائيَّة. وبعبارةٍ أخرى، يُمكنُ القولُ إنَّ أنظمة التَّحرير المُعجميّ تُساعدُ على تحرير المعاني وفقَ منهجيَّةٍ لنمذجة المُعجم على النَّحو الَّذي تتوحَّدُ فيه أنماطُ المعاني. ولتوضيح ذلك، نلاحظُ – مثلًا – طريقة المُعجم الوسيط في تحرير معانيه. فمعنى (الأسد) جنس من

الفصيلة السِّنَّوريَّة ..، ومعنى (الفهد) سبع من الفصيلة السِّنَّوريَّة بَين الْكَلْب والنمر ..، ومعنى (النَّمر) حَيَوَان مفترس أرقط من الفصيلة السِّنَّوريَّة ورتبة اللَّواحم ..، ومعنى (الظَّربان) حَيَوَان من رُتْبَة اللَّواحم والفصيلة السِّنَّوريَّة ..، ومعنى (القطّ) الهرُّ وَهُوَ جنس من الفصيلة السِّنَّوريَّة ورتبة اللواحم .. وهكذا[35]. إنَّ الكلمات الخمسة (الأسد، الفهد، النَّمر، الظَّربان، القطّ) تُشيرُ جميعًا إلى (جنس حيوانٍ من الفصيلة السِّنَّوريَّة). وقد أبانَ المُعجمُ عن هذا المعنى، لكنَّهُ لم يستخدم نمطًا واحدًا لها جميعًا، وإنَّما استخدمَ خمسةَ أنماطٍ مُختلفة، الأمرُ الَّذي يتنافى مع معايير الصِّناعة المُعجميَّة ويؤدِّي إلى صُعوبة التَّعامُل مع المادَّة المُعجميَّةٍ حاسوبيًّا حالَ الحاجة إلى ذلكَ مُستقبلًا.

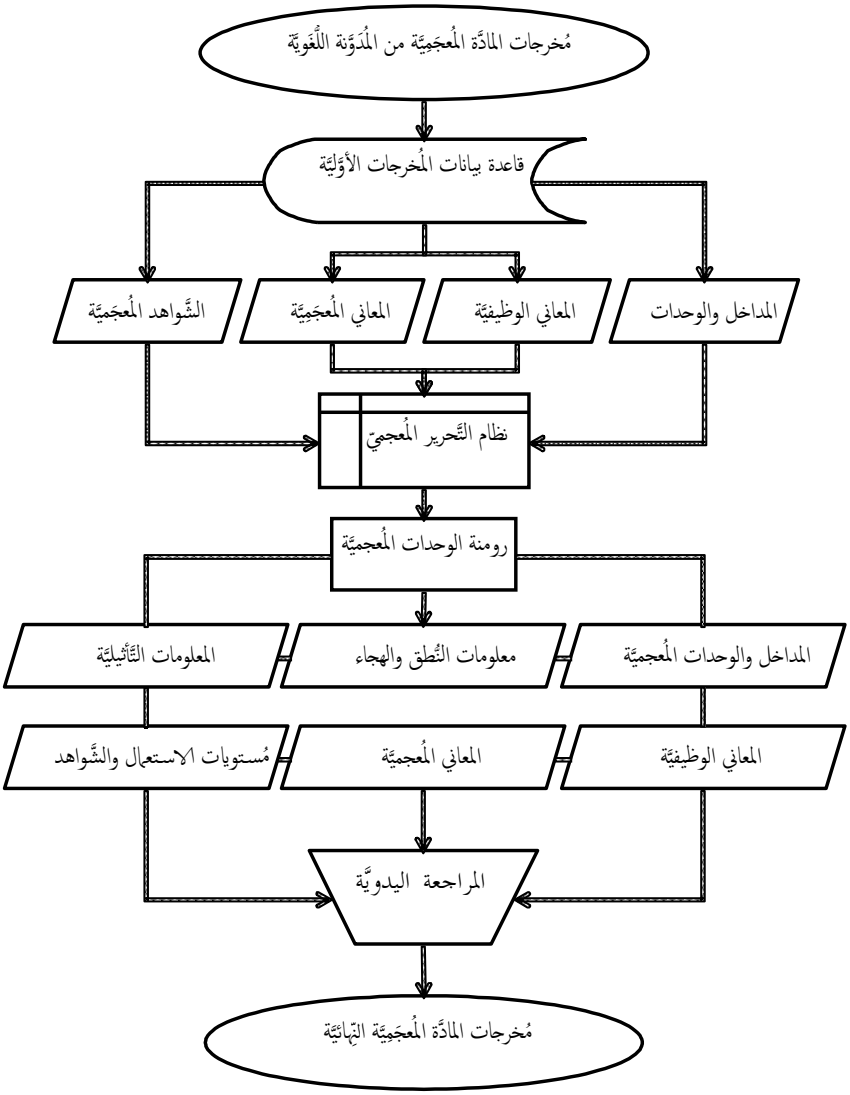
3.2.6. مُستويات الاستعمال *USAGE LEVELS*

ثَمَّةَ اعتِباراتٌ لتَعيين مُستَوَيات استِعمال الوحدات المُعجَمِيَّة، كاعتِبار الشُّيُوع والإهمال (شائِع، مُطَّرِد، نادِر، شاذّ)، واعتِبار الإباحة والحَظْر (مَحظُور، مُبتَذَل، سُوقِيّ، سَوادِيّ)، وغيرها. ويتطلَّبُ تعيينُ مُستويات الاستعمال في المُعجم المنشود الوُقُوفَ على إحصاءاتٍ دقيقةٍ لوحدات المُعجم المُستمدَّة من المُدَوَّنة اللُّغويَّة. ويُستفادُ من أنظمة التَّحرير المُعجميّ في استخلاص هذه الإحصاءات على نحوٍ دقيقٍ من خلال الأدوات الإحصائيَّة المُرفقة بها.

3.2.7. الشَّواهد المُعجميَّة *LEXICAL EVIDENCE*

الشَّاهدُ المُعجميُّ هو المثالُ الحيُّ على استخدام الجماعة اللُّغويَّة لوحدةٍ مُعجميَّةٍ في معنى مُعَيَّن. وثَمَّةَ مجمُوعَةٌ مِن السِّمات الَّتي يَنبَغي تَوافُرُها فِي شَواهِدِ المُعجَم التَّاريخيّ المنشود، منها: الوُضُوح والإيجاز، والتَّعبير عن المَعنى، والتَّرتيب الزَّمنيّ، ودِقَّةُ التَّوثِيق – نِسبَةً وتأرِيخًا. ويُستفادُ من أنظمة التَّحرير المُعجميّ في استخلاص الشَّواهد المُعجميَّة وترتيبها وربطها بالوحدات من خلال أدوات البحث والتَّرتيب المُرفقة بنظام التَّحرير المُعجميّ المُستخدَم.

35 – راجع هذه الموادّ ومعانيها في: مَجمع اللُّغة العَرَبِيَّة بالقاهِرة: السَّابق.



الشَّكل 13: مُخَطَّط انسيابي لمرحلة التَّحرير المُعجميّ

3.3. النَّشر المُعجَميّ

تُساعدُ أنظمةُ التَّحرير المُعجميّ إلى درجةٍ كبيرةٍ في إعداد الهيكل المُعجميّ النِّهائيّ [في صُورَتَيه: الورقيَّة والإلكترونيَّة]. والتماسًا لتحقيق الإفادة الكاملة من المُعجم التَّاريخيّ المنشود، ينبغي التَّخطيطُ للهيكل المُعجميّ في مراحل الصِّناعة الأولى. والصِّيغةُ المثاليَّة الَّتي تلتمسُها الدِّراسة هي صيغة المُعجم المقروء آليًّا *(MACHINE-READABLE DICTIONARY (MRD*؛ وهي إحدى صيغ المُعجم الإلكترونيّ، وإن كانت تزيدُ في أنَّها تمثيلٌ حاسوبيٌّ كاملٌ للمُعجم وموارِدِهِ وقواعد البيانات المُعجميَّة والأدوات المُستخدمة في المُعالجات على مادَّة المُعجم[36]. والواقع أنَّ الفائدةَ من هذه الصِّيغة قد لا تبدو على المدى القريب، ورُبَّما لا تخدم المعجمَ بصورةٍ مباشرة؛ لكنَّها تحقق النفع مستقبلًا. ذلكَ أنَّ إنجازَ المُعجم التَّاريخيّ للُّغة العربيَّة يتطلَّبُ العديدَ من الموارد اللُّغَوِيَّة والمُعجَمِيَّة والحاسُوبيَّة الَّتي يستغرقُ إعدادُها وقتًا وجهدًا كبيرين. وستُساعدُ إتاحةُ المُعجم المنشُود في هذه الصِّيغة على بناء العديد من الموارد الأخرى اعتمادًا على الموارد المُنجزة لهدف صناعة المُعجم، الأمرُ الَّذي يُساعدُ على إنجاز العديد من المشروعات الكبرى الَّتي تعودُ بالنَّفع على اللُّغة العربيَّة في مختلف ميادينها. فعلى مُستوى صناعة المُعجم، يُستفادُ من المُدَوَّنة اللُّغويَّة والمادَّة المُعجميَّة في صناعة مُعجماتٍ مرحليَّة للعربيَّة [القديمة والوسيطة والمُعاصرة] ومُعجمات مُتخصِّصة ومُعجمٍ تأثيليّ ومُعجمٍ للغة الطِّفل ومُعجمٍ للغة المرأة؛ ويُستفادُ من المادَّة المُعجميَّة وقواعد البيانات في صناعة مُعجمٍ موضوعيّ ومكنزٍ لُغويّ *TRESAURUS* وشبكة دلاليَّة *WORDNET* وأنطولوجيا معرفيَّة *ONTOLOGY*. وعلى مُستوى لسانيَّات المُدَوَّنة، يُستفادُ من المُدَوَّنة اللُّغويَّة للمُعجم في بناء مُدوَّنةٍ لُغويَّة للعربيَّة القديمة وثانيةٍ للعربيَّة

36 – حول مفهوم المُعجم المقروء آليًّا، راجع:

MITKOV, R. (2003). THE OXFORD HANDBOOK OF COMPUTATIONAL LINGUISTICS. OXFORD UNIVERSITY PRESS, 2003. P.747

الوسيطة وثالثةٍ للعربيَّة المُعاصرة ومُدَوَّناتٍ لأغراض تعليم العربيَّة وأخرى لدراسة ظواهرها البنويَّة والتَّركيبيَّة والدِّلاليَّة. وعلى مُستوى حوسبة اللُّغة، يُستفادُ من موارد المُعجم في بناء أدوات مُعالجة المباني وتدريب واختبار أدوات مُعالجة المعاني وتطوير قواعد بيانات أدوات مُعالجة العربيَّة عمومًا؛ والأمرُ يتجاوزُ ذلكَ إلى العديد من جوانب الإفادة الَّتي يصعُبُ حصرُها.

وفي إطار النَّشر المُعجميّ ينبغي بناء البوَّابة الإلكترونيَّة *ELECTRONIC PORTAL* للمُعجم المنشود مع مراحل الصِّناعة الأولى – كذلك. وتحريًا للدِّقَّة، تجدُرُ الإشارةُ إلى أنَّ البوَّابة الإلكترونيَّة لا تُمثِّلُ المُخرجات المُعجميَّة بالمفهوم الدَّقيق؛ وإنَّما هي أشبه بواجهةٍ إعلاميَّة، تشتملُ على تعريفٍ بالمُعجم والقائمينَ عليه والجهات المُتعاونة على إنجازه والأدوات والموارد المُتاحة له ومراحل العمل في صناعة المُعجم والمادَّة المُعجميَّة أو نماذج منها. وينبغي – عندَ بناء البوَّابة الإلكترونيَّة المُحاكية للبوَّابات الإلكترونيَّة لمُعجمات اللُّغات الأخرى – مُراعاة طبيعة اللُّغة العربيَّة، لاسيَّما فيما يتعلَّقُ بآليَّات البحث والتَّحرير. فإذا كانت البوَّاباتُ الإلكترونيَّة للمعاجم التَّاريخيَّة للألمانيَّة أو الإنجليزيَّة أو السِّويديَّة – مثلًا – تلتزمُ وسيلةَ البحث بالمدخل المُعجميّ؛ فهذا لأنَّ هذه اللُّغات التصاقيَّة؛ أمَّا العربيَّة فهي لُغةٌ اشتقاقيَّة يقتضي نظامُها المُعجميُّ أن تتاحُ آليَّاتُ البحث الإلكترونيّ فيها بالمدخل المُعجميّ المُمَثَّل في الجُذور، وبالوحدة المُعجميَّة المُمثَّلة في الفُرُوع أو اللِّكسيمات؛ بالإضافة إلى البحث بالمُطابقة الكاملة على النَّحو الَّذي تسيرُ عليه مُحرِّكاتُ البحث – عمومًا – بصرف النَّظر عن طبيعة اللُّغة. وكما تُعنى البوَّابةُ الإلكترونيَّة بتقديم المعلومات العامَّة عن المُعجم، فإنَّها تُعنى – كذلكَ – بتقديم خدماتٍ عامَّة، مثل: توفير نشرة إخباريَّة عن المُعجم، أو تكوين مجموعاتٍ ومنتدياتٍ إلكترونيَّة للنِّقاش بينَ أهل الصِّناعة وذوي العناية بالمُعجم، أو تقديم خدمات الاستفادة من الموارد والموادِّ المُعجميَّة الَّتي تُتيحُها البوَّابة، أو فتح المجال للمُتطوِّعينَ الرَّاغبينَ في تقديم خدماتهم للمُساعدة في إنجاز المُعجم المنشود؛ مع الأخذ في الاعتبار أنَّ إدارةَ البوَّابة تُتيحُ الخدمات في صورةٍ مُتباينة بينَ مُستخدمي

المُعجم. فما يُتاحُ لهيئة التَّحرير قد لا يُتاح للمُتطوِّعين، وما يُتاحُ للمُتطوِّعينَ قد لا يُتاحُ للمُستخدم العاديّ، وهكذا. والبوَّابةُ الإلكترونيَّةُ بهذا التَّوصيف ليسَت شكلًا ثابتًا أو خاملًا؛ وإنَّما هي شكلٌ مُتطوِّرٌ يُعنى بمراحل الصِّناعة المُعجميَّة بأكملها ويُؤرِّخُ لها، وهي كذلك مصدرٌ يُمكنُ من خلالِهِ إتاحةُ المُعجم عبرَ الشَّبكة العنكبوتيَّة مُستقبلًا. أمَّا عن الهيكل المُعجميّ النِّهائيّ الَّذي يُتاحُ للمُستخدم، فيوجدُ في صورتَين: ورقيَّة وإلكترونيَّة؛ ونعرضُ للصُّورتين فيما يلي.

3.3.1. النَّشر الورقيّ للمُعجم

يستمدُّ المُعجم التَّاريخيُّ المنشودُ – في صُورتِهِ الورقيَّة – مُعطياتِهِ من نظام التَّحرير المُعجميّ – على نحو ما أسلفنا. وبما أنَّ النَّشرَ الورقيَّ يُواجهُ العديدَ من التَّحدِّيات بعدَ ظهور النَّشر الرَّقميّ [الإلكترونيّ] في أشكالٍ مُتعدِّدة، فسيكونُ من المُفيد أن تتوافرَ عواملُ الجذب والدِّقَّة والوُضُوح في الصُّورة الورقيَّة للمُعجم التَّاريخيّ المنشُود. ويُمكنُ الإفادةُ من الحاسُوب في توفير هذه العوامل من خلال التَّحكُّم في أنواع الخُطوط المُستخدمة في مادَّة المُعجم وأحجامها وألوانها – بما يُلبِّي حاجةَ المُستخدم؛ وكذلك باستخدام الرُّموز والإشارات والعلامات المرجعيَّة الَّتي تُساعدُه على الوُصُول إلى المعلومة الَّتي ينشُدُها؛ وأخيرًا بتوفير الصُّورة التَّوضيحيَّة المُصاحبة باعتبارها وسيلةً مُساعدةً لشرح المعنى، سواءٌ أكانت صورة فوتوغرافيَّة *PHOTO* أم صورةً مرسومة *PICTURE* أم صورةً افتراضيَّة *IMAGE*. وثمَّةَ أربعة مصادرَ أساسيَّة للصُّور الَّتي يُمكنُ استخدامُها في النُّسخة الورقيَّة للمُعجم المنشود. أول هذه المصادر الصُّوَر الفوتوغرافيَّة الَّتي يُمكنُ التقاطُها خُصُوصًا لهدف الصِّناعة المُعجميَّة؛ والمصدرُ الثَّاني – وهو المصدرُ التَّقليديّ – بالاستعانة برسَّام قادرٍ على إخراج الصُّورة ببساطةٍ ووُضُوح؛ والمصدرُ الثَّالث بالاعتماد على الصُّورِ التَّوضيحيَّة الحُرَّة المُتاحة إلكترونيًّا على الويب، كتلكَ الَّتي تخضعُ لرخصة جنو للوثائق الحُرَّة *GNU FREE DOCUMENTATION LICENSE* [37]. أمَّا المصدرُ الرَّابع

37 – يُمكن الاطِّلاعُ على صيغة الرُّخصة في إصدارها الأخير (نوفمبر 2008) عبرَ الموقع الإلكترونيّ: HTTP://WWW.GNU.ORG.

فباستخدام أدوات تحرير الصُّورة، سواءٌ أكانت مفتوحة المصدر، مثل: برمجيَّة *GIMPHOTO* [38] وبرمجيَّة *GNU IMAGE MANIPULATION PROGRAM* [39]، أم تجاريَّة ومُغلقة المصدر، مثل: البرنامج الأشهر *ADOBE PHOTOSHOP* [40] وبرمجيَّة *COREL PHOTO–PAINT* [41]. وتحسُّبًا للتَّبايُن في توجُّهات مُستخدمي المُعجم يُمكن إتاحة النُّسخة الورقيَّة في صورتَين تحوي الأولى كاملَ المادَّة المُعجميَّة وتحوي الأخرى المادَّةَ مُختصرة، على النَّحو الَّذي جاءَ عليه مُعجم أكسفورد للإنجليزيَّة *OED* الَّذي أتيحَ في صورته الكاملة سالفة الذِّكر، وأتيحَ مُختصرًا بعُنوان: *SHORTER OXFORD ENGLISH DICTIONARY (SOED)*.

3.3.2. النَّشر الحاسُوبيّ للمُعجم

سبقت الإشارةُ إلى أنَّ الصِّيغةَ المثاليَّة الَّتي تلتمسُها الدِّراسة للمُعجم هي صيغة المُعجم المقروء آليًّا *MRD*. ومع أنَّ هذه الصِّيغة تُعَدُّ إحدى صُور النَّشر الإلكترونيّ، إلَّا أنَّها قد لا تكونُ مُناسبةً للمُستخدم العاديّ، نظرًا لما يبدو فيها من ضوضاء الوسائط وزحمة المعارف والمعلومات. وبالتَّالي فالمُفترضُ أن تُتاحَ هذه الصِّيغةُ لفئاتٍ مُعيَّنةٍ قادرةٍ على الإفادة الفعليَّة منها. أمَّا عن النُّسخة الإلكترونيَّة الرَّسميَّة للمُعجم المنشُود، فينبغي أن يُراعى في إعدادها التَّبايُنُ في توجُّهات المُستخدمين واختلافُ مُستوياتهم الفكريَّة والثَّقافيَّة، كما ينبغي أن تُتاحَ النُّسخةُ الإلكترونيَّة في صورةٍ تسمحُ لذوي الاحتياجات الخاصَّة بالإفادة الكاملة منها. وتحقيقًا لهذا، يُفترَضُ أن تُزوَّدَ النُّسخةُ الإلكترونيَّةُ من المُعجم المنشود بآليَّات

38 – الموقع الإلكترونيّ للبرمجيَّة:
HTTP://WWW.GIMPHOTO.COM.

39 – الموقع الإلكترونيّ للبرمجيَّة:
HTTP://WWW.GIMP.ORG.

40 – الموقع الإلكترونيّ للبرمجيَّة:
HTTP://WWW.PHOTOSHOP.COM.

41 – الموقع الإلكترونيّ للبرمجيَّة:
HTTP://WWW.COREL.COM.

البحث الألفبائيّ والجذريّ والمطابق، كما يُفترَضُ أن تُزوَّدَ بالوسائط المُتعدِّدة *MULTIMEDIA* من الصُّور والرُّسوم المُتحرِّكة والصَّوتيَّات والمرئيَّات الَّتي تُعَدُّ وسيلةً لشرح المعاني. أمَّا عن الصُّورة، فيُمكنُ الإفادةُ في تحريرها من المصادر الأربعة المُستخدمة في تحرير صُور النُّسخة الورقيَّة، مع إمكانيَّة الإفادة من أدواتٍ أخرى لتوفير الصُّورة ثلاثيَّة الأبعاد وإعداد الرُّسُوم المُتحرِّكة، سواءٌ أكانت مفتوحة المصدر، مثل: برمجيَّة *WINGS 3D* [42]، أم تجاريَّة مُغلقة، مثل: برمجيَّة *AUTODESK 3DS MAX* [43]. وأمَّا الصَّوتيَّاتُ فينبغي أن يُراعى فيها سلامة اللِّسان من العُيُوب ووُضُوحُ المخارج الصَّوتيَّة والإحاطة بقواعد اللُّغة العربيَّة. ويُمكنُ الإفادة من برمجيَّاتٍ لتحرير الصَّوتيَّات، مثل: برمجيَّة *AUDACITY* [44] مفتوحة المصدر، وبرمجيَّة *WAVELAB* [45] المُغلقة. أمَّا المرئيَّات – حالَ توافُرها في صيغةٍ قانونيَّةٍ تسمحُ باستخدامها – فينبغي أن تخضعَ لعمليَّات الضَّبط والتَّركيب والمونتاج التماسًا للإطار الأخلاقيّ فيها، بالبُعد عن المحظورات. ويُمكنُ الإفادةُ من برمجيَّاتٍ لتحرير المرئيَّات، مثل: برمجيَّة *VIVIA* [46] مفتوحة المصدر، وبرمجيَّة *ADOBE PREMIERE* [47] المُغلقة. وفي كلِّ الأحوال، ينبغي أن تُراعى البساطة والدِّقة والوُضُوح والخُلُوّ من الضَّوضاء في الوسائط المُتعدِّدة للمُعجم التَّاريخيّ المنشُود.

---

42 – الموقع الإلكترونيّ للبرمجيَّة:

HTTP://WWW.WINGS3D.COM.

43 – الموقع الإلكترونيّ للبرمجيَّة:

HTTP://WWW.AUTODESK.COM.

44 – البرمجيَّة مُتاحةٌ على موقع البرمجيَّات الحرَّة. *SOURCEFORGE*؛ ورابطها:

HTTP://AUDACITY.SOURCEFORGE.NET/.

45 – الموقع الإلكترونيّ للبرمجيَّة:

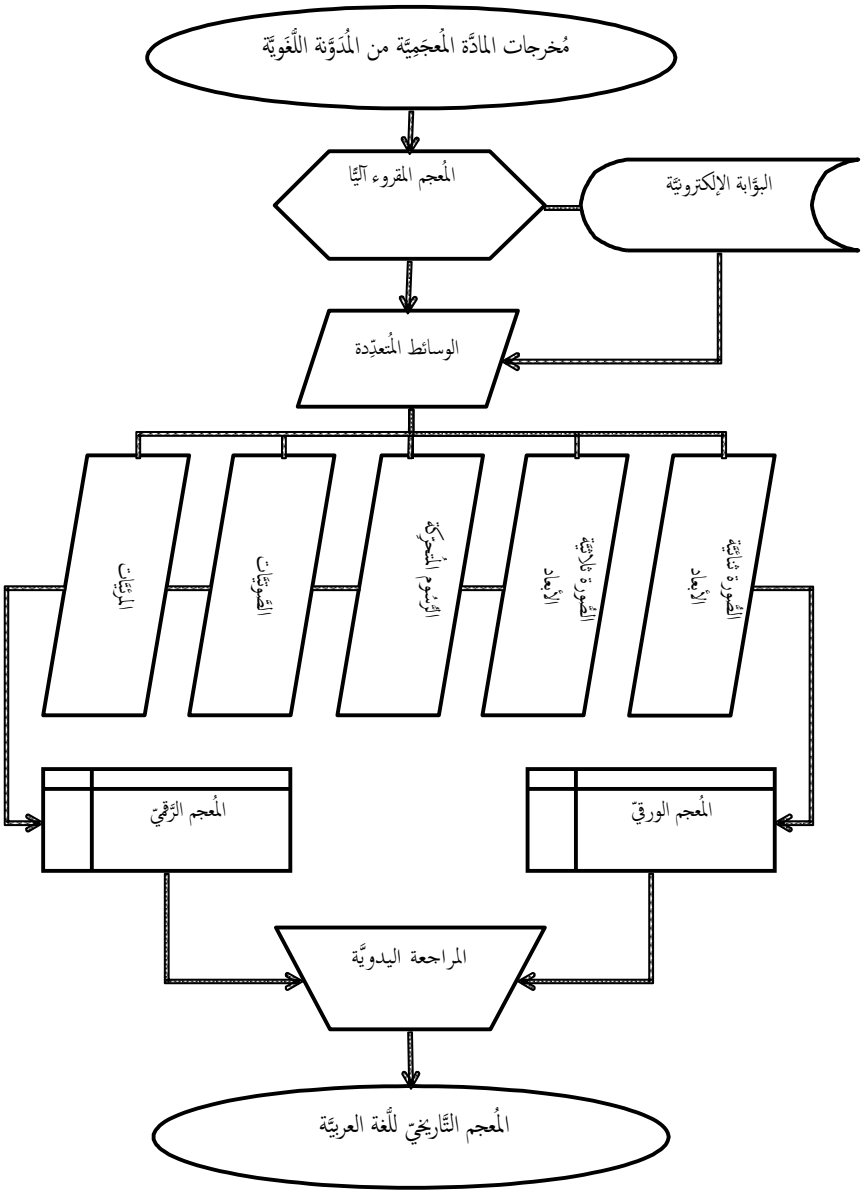
HTTP://WWW.STEINBERG.NET.

46 – البرمجيَّة مُتاحةٌ على موقع البرمجيَّات الحرَّة. *SOURCEFORGE*؛ ورابطها:

HTTP://SOURCEFORGE.NET/PROJECTS/VIVIA.

47 – الموقع الإلكترونيّ للبرمجيَّة:

HTTP://WWW.ADOBE.COM.



الشَّكل 14: مُخَطَّط انسيابي لمرحلة النَّشر المُعجميّ

4. نتائج الدِّراسة

1. تستدعي حوسبةُ المُعجم التَّاريخيّ للُّغة العربيَّة الوُقُوفَ على الإشكالات النَّاتجة عن طبيعة اللُّغة العربيَّة وواقع اللِّسانيَّات الحاسُوبيَّة، والانطلاقَ من رؤيةٍ واضحة المعالم ومنهجٍ مُكتمل الأركان وفهمٍ دقيق لمفهوم الصِّناعة المُعجميَّة، والقُدرة على التَّنبُّؤ بالعقبات الَّتي يُمكنُ أن تواجهها الصِّناعة المُعجميَّة.

2. تقومُ منهجيَّةُ حوسبة المُعجم التَّاريخيّ للُّغة العربيَّة على ثلاث مراحلَ أساسيَّة؛ حيثُ تُعنى المرحلةُ الأولى بحوسبة المُدَوَّنة اللُّغَوِيَّة للمُعجَم عبرَ مرحلَتَي البناء والمُعالجَة، وينتُجُ عنها المُخرجات المُعجَمِيَّة المُستمدَّة من المُدَوَّنة دونَ غيرِها من الموارد. وتُعنى المرحلةُ الثَّانية بالتَّحرير المُعجَميّ، وينتُجُ عنها المُخرَج المُعجَميّ الكامل. وتُعنى المرحلةُ الثَّالثة بالنَّشر المُعجَمِيّ، وينتُجُ عنها الهيكل المُعجميّ النِّهائيّ [في صُورَتَيه: الورقيَّة والرَّقميَّة].

3. يستدعي بناءُ المُدَوَّنة اللُّغَويَّة للمُعجم أن يتمَّ تحويلُها من شكلِها الورقيّ الَّذي تبدو عليه في الكُتُب والمُصَنَّفات والوثائق إلى شكلٍ رَقْمِيّ مُنَقَّى، يُمكنُ التَّحكُّمُ فيه بالإضافة أو الحذف أو التَّعديل، ويصلُحُ للمُعالجة الآليَّة.

4. تتمُّ مُعالجة المُدَوَّنة اللُّغويَّة للمُعجم على مُستوى المباني والمعاني لاستخلاص المعلومات المُعجميَّة الرَّئيسة من المداخل والوحدات والمعاني والشَّواهد. ويُعتمَدُ – في ذلكَ – على آليَّاتٍ خمسة، هي: (آليَّة فهرسة النُّصُوص، وآليَّة التَّحليل الصَّرفيّ، وآليَّة التَّحليل التَّركيبيّ، وآليَّة التَّحليل الدِّلاليّ، والمنصَّة الحاسُوبيَّة).

5. يستدعي تحريرُ المُعجم في صورةٍ مُمنهجة ومُنتظمة الاعتمادَ على نظامٍ للتَّحرير المُعجميّ *DWS* يتوافقُ مع طبيعة اللُّغة العربيَّة، ويكونُ قادرًا على التَّعامُل

مع موارد المُعجم ومُزوَّدًا بالأدوات المُساعدة على تحرير المُعجم على الوجه المنشُود، وتتحقَّقُ من خلاله الإفادةُ الكاملة بتعيين المعلومات المُعجميَّة الأساسيَّة.

6. اقترَحَت الدِّراسةُ أن يُتاحَ المُعجمُ في صيغة المُعجم المقروء آليًّا *(MRD)*؛ حيثُ تُساعدُ هذه الصِّيغة على بناء العديد من الموارد الأخرى، الأمرُ الَّذي يُساعدُ على إنجاز العديد من المشروعات الكبرى الَّتي تعودُ بالنَّفع على اللُّغة العربيَّة – مُستقبلًا – في مختلف ميادينها.

7. اقتَرَحت الدِّراسة أن يُستفادَ من تقنيات الحاسُوب في النَّشر المُعجميّ بصورتيه الورقيَّة والإلكترونيَّة، عبرَ توفير وسائل شرح المعنى المُناسبة. كما اقترحَت الدِّراسةُ توفيرَ النُّسخة الورقيَّة من المُعجم في صورتَين، تحوي الأولى كاملَ المادَّة المُعجميَّة وتحوي الأخرى المادَّةَ مُختصرة، على النَّحو الَّذي جاءَ عليه مُعجم أكسفورد للإنجليزيَّة.

5. الخُلاصة

تواجهُ حوسبةُ المُعجم التَّاريخيّ للُّغة العربيَّة أربعة إشكالاتٍ رئيسة، خُلاصتُها أنَّنا لا نملكُ تجربة واحدة لحوسبة المُعجمات التَّاريخيَّة عبرَ مختلف مراحل الصِّناعة المُعجَمِيَّة، وأنَّنا نُعاني ضبابيَّة الرُّؤية المنهجيَّة وحاجة الصِّناعة المُعجَمِيَّة العربيَّة إلى طفرةٍ حقيقيَّة، وأنَّ للُّغة العربيَّة طبيعةً اشتقاقيَّة ونظامًا كتابيًّا خاصًّا بها، وأنَّ كثيرًا من المعجَمِيِّينَ والقائمينَ على المؤَسَّسات المعنِيَّة بالصِّناعة المعجَمِيَّة العربيَّة لا يُمَيِّزونَ بينَ مُصطَلَحَي المعجَمِيَّة وصِناعة المعجَم. وتُحاولُ الدِّراسةُ الوُقُوفَ على هذه الإشكالات سعيًا إلى حلِّها. ولأنَّ الصِّناعةَ المُعجميَّة تمرُّ بثلاث مراحل، هي الجمع والتَّحرير والنَّشر، فقد اقترَحَت الدِّراسةُ منهجيَّةً لحوسبة المُعجم التَّاريخيّ المنشود للعربيَّة، انطلاقًا من مفهوم المُعجم التَّاريخيّ

وطبيعة اللُّغة العربيَّة على مُستوى المباني والمعاني، وفي ضوء ما تُتيحُهُ اللِّسانيَّاتُ الحاسُوبيَّةُ من أدواتٍ للمُعالجة الآليَّة. وتتدرَّجُ الدِّراسةُ في تقديم المنهجيَّة، فتبدأ بحوسبة المُدَوَّنة اللُّغويَّة للمُعجم المنشود عبرَ مرحلَتي البناء والمُعالجَة بما يُساعدُ على استخلاص المادَّة المُعجميَّة، ثُمَّ تعرضُ للتَّحرير المُعجميّ بما يُساعدُ على تعيين مُخرجات المُعجم، وأخيرًا تعرضُ للنَّشر المُعجميّ بصُورتيه الورقيَّة والإلكترونيَّة.

مراجع الدِّراسة

أوَّلًا: المراجع العربيَّة

الكُتُب

- حجازي (مَحْمُود فهمي): البحث اللُّغَوِيّ، دار غريب، القاهرة، ط1، (د.ت).
- السَّعيد (الـمُعتزّ 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، 2014م.
- السَّعيد (المُعتزّ بالله): علم الدِّلالة ونظريَّة المعنى، دار الهاني، القاهرة، ط1، 2013.
- فيشر (أوجُست): المُعجَم اللُّغَوِيّ التَّارِيخِيّ، القِسم الأوَّل، مجمع اللُّغة العَرَبِيَّة، القاهِرة، 1967م.

المعاجم

- عُمَر (أحمد مُختار): المكنز الكبير، "معجم شامل للمجالات والمُترادفات والمتضادَّات"، مؤسَّسة سُطور، 2000.
- مَجمع اللُّغة العَرَبِيَّة بالقاهِرة: المُعجَم الوسيط، مكتَبَة الشّروق الدَّولِيَّة، القاهِرَة، ط3، 2004.

الدَّوريَّات

- ابن مُراد (إبراهِيم): في مفهوم المُعجَم التَّاريخِيّ وتَطبيقاته على العَرَبيَّة، بحث مُقَدَّم إلى ندوة اتِّحاد المعاجم العربية حول المعجم التاريخيّ للُّغة العربيَّة، الشَّارِقة، 2006م.

الأطروحات العلميَّة

- السَّعيد (المُعتزّ بالله): مُدَوَّنة مُعجَم تاريخيّ للُّغة العربيَّة "مُعالجَة لُغَوِيَّة حاسوبِيَّة"، أطروحة دكتوراه، جامعة القاهرة، 2011م.

ثانيًا: المراجع الأجنبيَّة

الكُتُب

- *Abel, A. (2012). Dictionary writing systems and beyond. In: Granger, S. & Paquot, M. (Ed.). Electronic Lexicography. Oxford University Press, 2012.*
- *Fischer, W. & Gätje, H. editors. (1982). Grundriss der arabischen Philologie: Sprachwissenschaft. Reichert. Wiesbaden.*
- *Mitkov, R. (2003). The Oxford Handbook of Computational Linguistics. Oxford University Press, 2003.*
- *Vermon, L. (2012): Lexicology and Lexicography: Words and Ways. Webster's Digital Services.*

المعاجم

- *Kraemer, J. & Gätje, H. & Spitaler, A. & Ullmann, M. (1970). Wörterbuch der Klassischen Arabischen Sprache. Otto Harrassowitz. Wiesbaden. Bd.I.*
- *Sinclair, J. (2009). Collins COBUILD Advanced Dictionary of English. Heinle Cengage Learning.*

الأوراق البحثيَّة

- *EID, S. M. & AL-SAID, A. B. & 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.*
- *PAJZS, J. (2000). MAKING HISTORICAL DICTIONARIES BY COMPUTER. PROCEEDINGS OF EURALEX 2000. ULRICH HEID ED. UNIVERSITY OF STUTTGART, STUTTGART, 2000.*

الوثائق

- *"GNU FREE DOCUMENTATION LICENSE". VERSION 1.3, 3 NOVEMBER 2008, ON THE WEB PAGE: HTTP://WWW.GNU.ORG/COPYLEFT/FDL.HTML.*

ثالثًا: المواقع الإلكترونيَّة

- *HTTP://AUDACITY.SOURCEFORGE.NET/.*
- *HTTP://CONCAPP.SOFTWARE.INFORMER.COM.*
- *HTTP://GATE.AC.UK.*
- *HTTP://LXLN.PRETTYDATA.EU/.*
- *HTTP://SIBAWAYH.EMI.AC.MA/SAFAR.*
- *HTTP://SOURCEFORGE.NET/PROJECTS/MATAPUNA/.*
- *HTTP://SOURCEFORGE.NET/PROJECTS/VIVIA.*
- *HTTP://TSHWANEDJE.COM/TSHWANELEX/.*
- *HTTP://WWW.ADOBE.COM.*
- *HTTP://WWW.ALMUAJAM.ORG.*
- *HTTP://WWW.AUTODESK.COM.*

- *HTTP://WWW.COLLINSDICTIONARY.COM.*
- *HTTP://WWW.COREL.COM.*
- *HTTP://WWW.GIMP.ORG.*
- *HTTP://WWW.GIMPHOTO.COM.*
- *HTTP://WWW.GNU.ORG.*
- *HTTP://WWW.NOOJ4NLP.NET.*
- *HTTP://WWW.NOOJ4NLP.NET.*
- *HTTP://WWW.PHOTOSHOP.COM.*
- *HTTP://WWW.STEINBERG.NET.*
- *HTTP://WWW.WINGS3D.COM.*

مُصطلحات الدِّراسة

AFFIX	**الزَّائدة**
ALLOGRAPH	**الشَّكل الكتابيّ للجرافيم [الألوجراف]**
ALPHABET & PRONUNCIATION	**معلومات الهجاء والنُّطق**
ANNOTATED TEXT	**نُصُوص مُرَمَّزة [مُعنونة]**
ANTHROPOLOGY	**علم الإنسان**
ARCHAEOLOGY	**علم الآثار**
CHARACTER ENCODING	**ترميز المحارف**
COLLOCATIONS	**المُتَصاحِبات اللَّفظِيَّة**
COMPUTATIONAL FRAMEWORK	**المِنَصَّة الحاسُوبيَّة**

COMPUTATIONAL LINGUISTICS	اللِّسانيَّات الحاسُوبيَّة
CONCORDANCER	المُفهرس الآليّ [الكشَّاف السِّياقيّ]
CONTEXT	السِّياق
DICTIONARY PRODUCTION SYSTEM (DPS)	نظام صناعة المُعجم
DICTIONARY WRITING SYSTEMS (DWS)	أنظمة التَّحرير المعجَمِيّ
ELECTRONIC PORTAL	البوَّابة الإلكترونيَّة
ETHNOLOGY	علم الأعراق
ETYMOLOGICAL INFORMATION	المعلومات التَّأثيليَّة
EXTENSIBLE MARKUP LANGUAGE (XML)	لُغة التَّوصِيف القابِلة للامتِداد
GNU FREE DOCUMENTATION LICENSE	رخصة جنو للوثائق الحُرَّة
GRAMMATICAL FUNCTIONS	المعاني الوظيفيَّة [النَّحويَّة]
GRAPHEME	الوحدة الكتابيَّة [الجرافيم]
HISTORICAL DICTIONARY	المُعجم التَّاريخيّ
HISTORY OF SCIENCE	تاريخ العُلُوم

HYPER TEXT MARKUP LANGUAGE (HTML)	**لُغَة تَوصِيف النُّصُوص التَّشَعُّبِيَّة [صيغة صفحات الويب]**
INTERNATIONAL PHONETIC ALPHABET (IPA)	**الأَلِفبائِيَّةُ الصَّوتِيَّة الدَّولِيَّة**
LEMMA - LEMMATA	**الفرع اللُّغويّ – فُرُوع الكلمة**
LEMMATIZER	**آليَّة التَّفريع**
LEXEME – HEADWORD	**الوحدة المُعجميَّة**
LEXICAL ENTRY	**المدخل المُعجميّ**
LEXICAL EVIDENCE	**الشَّواهد المُعجميَّة**
LEXICAL MEANINGS	**المعاني المُعجميَّة**
LEXICAL MODELING	**النَّمذجة المُعجَمِيَّة**
LEXICOGRAPHY	**صناعة المُعجم**
LEXICOLOGY	**علم المُعجم [المُعجميَّة]**
LINGUISTIC CORPORA	**المُدوَّنات اللُّغويَّة**
LINGUISTICS	**اللِّسانيَّات**
MACHINE-READABLE DICTIONARY (MRD)	**المُعجم المقروء آليًّا**
MORPHEME	**الوحدة الصَّرفيَّة المُجَرَّدة [المورفيم]**
MORPHOLOGICAL ANALYZER	**المحلِّل الصَّرفيّ/ البِنويّ**
MULTIMEDIA	**الوسائط المُتعدِّدة**

Natural Language Processing (NLP)	مُعالجة اللُّغات الطَّبيعيَّة
Ontology	الأنطولوجيا المعرفيَّة
Optical Character Recognition (OCR)	تقنية القارئ الآليّ [التَّعرُّف الضَّوئيّ على المحارف]
Parts of Speech (PoS)	أقسام الكلام
Portable Document Format (PDF)	صِيغَة الوَثِيقَة المُتَنَقِّلَة
PoS Parsing	إعراب أقسام الكلام
PoS Tagging	تحليل أقسام الكلام
Prefix	السَّابقة
Raw Text – Unannotated Text	نُصُوص خام [غير مُعنونة]
Romanization of Arabic	رومنة الحُرُوف العربيَّة
Romanizer	آليَّة الرَّومنة
Root	الجذر اللُّغويّ
Search Engine	مُحرِّك البحث
Semantic Analyzer	آلِيَّةُ التَّحلِيل الدِّلالِيّ
Stem	الجذع اللُّغويّ
Stemmer	آليَّة التَّجذيع
Suffix	اللَّاحقة

SYNTACTIC ANALYZER	**المُحلِّلُ التَّركيبيُّ**
TRESAURUS	**المكنز اللُّغويّ**
UNICODE	**نِظام الحُرُوف الدَّوْلِيّ المُوَحَّد**
USAGE LEVELS	**مُستويات الاستعمال**
WORD SENSE DISAMBIGUATION (WSD)	**فَكُّ الالتباس الدِّلالِيّ للكَلِماتِ**
WORDNET	**الشَّبكة الدِّلاليَّة [شبكة الكلمات]**