حوسبة اللغة العربية: واقع وتحديات

الاستـلام: 21/أكتوبر/2023
التحكيم: 18/ نوفمبر /2023
القبـــول: 25 / نوفمبر /2023

د. نجاة حسين [1،*]

© 2023 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2023 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] قسم اللغة العربية -كلية الآداب والفنون-جامعة شلف-الجزائر.
* عنوان المراسلة: nadjet.hocine@univ-chlef.dz

https://doi.org/10.20428/jss.v29i4.2243

حوسبة اللغة العربية: واقع وتحديات

الملخص:

تعد اللسانيات الحاسوبية من أهم العلوم الحديثة التي تواكب التقنية، وترتقي باللغة البشرية عبر الاستفادة من مزايا الحاسوب ومعطياته وقدراته وأدواته المتاحة، لتوظيف قواعد بيانات اللغة والموارد اللسانية المتوفرة والمتاحة، والاستفادة كذلك من علوم عديدة بحتة، كالرياضيات، والمنطق.

ترتبط اللغة بالحاسوب ارتباطًا وثيقًا وقويًا ، فبإدخال اللغة في الحاسوب يعني إقامة حوار بين الإنسان والآلة، وكلما تطورت تقنيات الحاسوب كانت أقرب إلى محاكاة قدرات الإنسان في طريقة عمله وتفكيره، واتخاذ قراراته أيضًا. وتنفرد اللغة العربية بخصائص لسانية صورية جعلتها من وجهة نظر حاسوبية قابلة للتوصيف والنّمذجة بعد كثير من جهود اللسانيين والحاسوبيين في مختلف مستوياتها التحليليّة، الصوتيّة، والصرفيّة، والنّحويّة، والمعجميّة، والدّلالية، ولكن السؤال الذي يطرح نفسه هل هذه الخصائص والسمات التي تمتاز بها لغتنا كافية لحوسبتها أم هناك مشاكل تستدعي منّا النّظر للنّهوض بها؟

الكلمات المفتاحية: اللسانيات الحاسوبيّة، حوسبة اللغة العربيّة، التوصيف، النّمذجة.

Computerization of the Arabic Language: Reality and Challenges

Dr. Najat Hocine [1,*]

Abstract

Computer linguistics is one of the most important modern sciences that keep pace with technology and advance the human language by taking advantage of the advantages of computers, its data, capabilities and available tools, to employ language databases and available and available linguistic resources, as well as benefit from several pure sciences such as mathematics and logic.

Language and computers are closely and strongly linked. Inserting language into a computer means establishing a dialogue between humans and machines. The more computer technologies develop, the closer they are to simulating human capabilities in the way they work, think, and make decisions as well. It is worth noting that the Arabic language is unique in its formal linguistic characteristics that make it, from a perspective A computational view that can be described and modelled after many efforts by linguists and computers at its various levels of analysis: phonetic, morphological, grammatical, lexical, and semantic. However, the question that arises is whether these characteristics and features that characterize our language are sufficient to computerize it? Or are there problems that require us to look into improving them?

Keywords: computational linguistics, computing the Arabic language, characterization, modelling.

1 Department of Arabic Language - Faculty of Arts - University of Chlef - Algeria.
* Corresponding Email: nadjet.hocine@univ-chlef.dz

https://doi.org/10.20428/jss.v29i4.2243

المقدمة:

1- موضوع الدراسة:

لقد شهد النّصف الثاني من القرن العشرين تطورات مذهلة في الاتصال ووسائله، ولعلّ أبرز هذه التطورات ظهور ما يسمى بالحاسوب، هذه الآلة الصغيرة التي أحدثت نقلةً نوعيّةً في التاريخ؛ لما لها من قدراتٍ فائقةٍ على إنجاز عملياتٍ متعددة، كتخزين كمٍّ هائلٍ من المعلومات واسترجاعها، ومعالجتها بطرق مختلفة تعجز عن مثلها القدرات البشريّة.

إنّ ما يتصف به الحاسوب من قدرة على تمثيل الحقائق والأفكار جعله يبثّ منطقه وآلياته في مختلف الميادين الإنسانيّة، ومن هنا أخذ علماء اللسانيات يدرسون مبادئ هذا العلم وآلياته؛ حتى يتمكنوا من الإفادة منه في معالجته اللغات البشريّة معالجةً آليةً، ليس بمعنى أن تعمل الحاسبات على اللغة البشريّة، بل أن تعمل اللغات البشريّة في الحاسبات الإلكترونيّة (الوعر، 1989 ص316-320).

ومضى الحاسوبيون جادين في اتخاذ لغة الإنسان الطبيعيّة ومعالجتها آليًا، لتصبح أداةً لحوار بين الإنسان والحاسوب، لاسيما في حواسيب الجيل الخامس والسادس لتنحية اللغات الاصطناعيّة من جهة، وليصبح التخاطب بين الإنسان والحاسوب كما التخاطب بين إنسان وإنسان آخر (عيسى ، 2007، ص 39)، وقد نجم عن معالجة الحاسوبيين للغات الطبيعيّة ومحاولة إفهامها له إلى ظهور ما يسمى "باللسانيات الحاسوبية" التي صارت فرعًا من فروع علم الحاسوب.

2- أهمية الدّراسة:

تأتي أهمية هذه الدّراسة استجابةً لما نعيشه اليوم من ثورةٍ علميّةٍ وتقنيّةٍ ومعلوماتيّةٍ في مختلف المجالات، وأمام هذا الوضع تجد اللغة العربيّة نفسها مضطرةً إلى مواكبة هذا التطور الحاصل، ولن يتحقق هذا إلاّ باستغلال التقنيات التي يقدمها الحاسوب لمعالجتها آليًا.

3- أهداف الدّراسة:

تهدف هذه الدّراسة إلى:

- إبراز أهمية اللسانيات الحاسوبية، وخاصة في اللغة العربية.
- تبيان الجهود العربية المقدمة في هذا الميدان، من بحوث، ومؤتمرات، ودراسات، والصعوبات التي مازالت تشكّل عائقًا في معالجة اللغة العربية آليًا.

4- منهجية الدّراسة:

لمعالجة هذه الدّراسة اتبعنا المنهج الوصفي التحليلي؛ باعتباره المنهج الأنسب لطبيعة الموضوع.

أولًا: مفهوم اللسانيات الحاسوبيّة (Computational Linguistics)

لعلّ اللّسانيات الحاسوبيّة تكون أحدث فروع اللّسانيات، ولا نبالغ إذا قلنا إنّها أهم هذه الفروع جميعًا في عصرٍ تتعاظم فيه أهمية الآلة، والتقنية، والمعرفة.

والظاهر جليًا أنّ هذا العلم فرع بيني، نصفه منتسب إلى اللّسانيات التطبيقية وموضوعها اللغة، والنّصف الآخر حاسوبي وموضوعه توليد اللغة إلى رموز رياضيّة يفهمها الحاسوب، حتى يتأتى له القيام بكثيرٍ من الأنشطة اللغوية

التي يؤديها العقل البشري (عيسى، 2007، ص 39)، وبالتالي فإنّ هذا العلم يعمل من أجل فهم آلية الدّماغ البشري، ومحاكاته في استقبال اللغة وانتاجها.

وقد تعدّدت تعريفات وتسميات اللسانيات الحاسوبية*، مما يصعب علينا إعطاء تعريفٍ جامع وشاملٍ لها، ومن بين هذه التعريفات أنها: أحد فروع اللسانيّات التطبيقيّة تهتم بالإفادة من معطيات الحاسوب في دراسة قضايا اللسانيات المتعددة، مثل رصد الظواهر اللغوية وفقًا لمستوياتها الصوتيّة، الصرفيّة، النحويّة، البلاغيّة، والعروضيّة، واجراء العمليّات الإيحائيّة، وصناعة المعاجم، والترجمة الآليّة، وتعليم اللغات (عبد القادر، 2002، ص 181)، فهذا العلم يهتم باللغة، فيعمل على استخدام اللغة كأداة طبيعية لمعالجتها في الحاسبات الإلكترونية، ويتألف من اللسانيات العامة، ومن علم الحاسبات الإلكترونية، وعلم الذكاء الاصطناعي، ومن علم المنطق، وكذا علم الرياضيات، حتى تتناسق هذه الفروع، وتتآلف لتشكّل مبادئ علم اللسانيات الآلي (حافظ وآخرون، 2009 ص 111)، أي تحاول أن تعالج اللغة الطبيعيّة آليًا بوضع دماغٍ آليٍ قادرٍ على استعمال اللغة مثلما يستعملها الإنسان، وبما أنّه علم يجمع بين اللسانيات والحاسوب فإنّ موضوعه اللغة والحاسوب فهو: "ترجمة اللغة إلى رموز رياضيّة يفهمها الحاسوب، أو تهيئة اللغة الطبيعية لتكون لغة تخاطب وتحاور مع الحاسوب بما يفضي إلى أن يؤدي الحاسوب كثيرًا من الأنشطة اللغوية التي يؤديها الإنسان، مع إقامة الفرق في الوقت والكلفة" (العناتي، 2005، ص 62).

ثانيًا: نشأة اللسانيات الحاسوبيّة

إنّ اللسانيات الحاسوبيّة جاءت نتيجة جهودٍ متفرقةٍ من قبل المختصين والباحثين البارزين في هذا الميدان، وهذا ما جعل صعوبة في وضع تاريخٍ زمنيٍّ محددٍ لهذا العلم، لكن رغم ذلك مرّت بحقبٍ زمنيةٍ مختلفةٍ ودولٍ عديدة، وعلى غرار هذا القول، سوف نحاول التطرق إليها، عبر عرض مراحل نشأتها عند الغرب والعرب.

1.2 بالنسبة للغرب:

كان ظهور هذا العلم في أمريكا من خلال حقل الترجمة الآلية من اللغات الأخرى إلى اللغة الإنجليزية بجامعة (جورج تاون) سنة (1954)، وقد ساهمت أيضًا بحوث الألسني الأمريكي نعوم تشومسكي (Noam Chomsky) في تطور اللسانيات الحاسوبية، والذي سعى دومًا إلى التأقلم مع متطلبات المعالجة الآلية للغة، وعمل بالتنسيق مع الحاسوبين (طانيوس، 2012، ص 17).

أما في أوروبا فكان ظهورها سنة (1961) بجامعة قوتبرك (Goteborg) السويدية، لكنها كانت محاولة ذات طابعٍ محلي، أما البداية الفعلية فكانت في مركز التحليل الآلي للغة بمدينة قالارات (Gallarat) بإيطاليا، حيث وضع روبيرتوا بوزا (Roberto busa) سنة (1962) الدعائم الأولى لاستخدام الحاسوب في دراسة اللغة، ثم توالت بعدها فتح مراكز حاسوبية للغة (العارف، 2007، ص 15).

* ذكرت مجموعة من المسميات للسانيات الحاسوبية، وذلك نتيجة اختلاف المنطلقات والمرجعيات والخلفيات، ومن تلك التسميات على سبيل المثال: علم اللغة الحاسوبي، اللغويات المعلوماتية.

https://doi.org/10.20428/jss.v29i4.2243

2.2 بالنسبة للعرب:

أما بالنسبة للعالم العربي، فقد كانت بداية الاستفادة من هذا العلم في السبعينيات، وذلك في مجال العلوم الشرعية، إذ اقتصرت في بادئ الأمر على إدخال أجزاء معينة من القرآن الكريم في الحاسوب، ثم أتت أول محاولة لتعريبه، وذلك باستبدال الحروف اللاتينية بالحروف العربية.

أما في مجال الدّرس اللغوي، فقد كانت البدايات إحصائيات بالدرجة الأولى، وبتوجيه من الدكتور إبراهيم أنيس الذي انتهز فرصة تدريسه بجامعة الكويت سنة (1971)، ليلتقي بأستاذ الفيزياء المصري "علي موسى" ويشرح له فكرة الإحصاءات اللغوية وأهميتها في البحث اللغوي، ثم اتفقا على البدء بدراسة إحصائية لجذور اللغة، كما جاءت في معجم الصحاح للجوهري (أنيس، 1973، ص211).

كانت هذه البداية الأولى لحوسبة التراث العربي عبر هذه الدّراسة الإحصائية، ثم تبعتها دراسات إحصائية أخرى كإحصاء جذور معجم لسان العرب لابن منظور، وإحصاء جذور معجم تاج العروس للزبيدي، وإحصاء ألفاظ القرآن الكريم (موسى، 2001)، ومن ثم توالت جهود عربية لأفرادٍ ومؤسساتٍ علميةٍ وشركاتٍ للولوج في هذا المكان عبر نشر مؤلفات، ومقالات علمية، وإقامة الندوات، وإعداد برامج، ونظم من أجل حوسبة العربية.

ويمكن إدراج تاريخ لهذه المساهمات التي سعت لخدمة اللغة العربية وتطويرها في مجال اللسانيات الحاسوبية، وأول ما يمكن أن نبدأ به هو أهم الندوات والمؤتمرات.

ويعدّ المؤتمر الذي عُقد بالرباط بين (26) سبتمبر و(5) نوفمبر (1983) من بواكير الجهود في مجال اللسانيات الحاسوبية، وقد ساهم في عقده "المركز القومي للتوثيق والتخطيط للبحث العلمي والتقني في المغرب"، و"معهد الدراسات والأبحاث والتعريب في المغرب"، وقد نشرَت أعماله في كتاب "اللسانيات العربية التطبيقية والمعالجة الإشارية والمعلوماتية".

وقد قسِّم الكتاب إلى شطرين، يتناول الشطر الأول منه إطلالة على اللسانيات والصوتيات للمهندسين، وخُصِّص الشطر الثاني لوصف خصائص العربية، وجرى التركيز على قواعد التصريف والاشتقاق التي منها ينتج عدد كبير من الكلم والصيغ المعجمية ذات الأنماط والدّلالات المحددة، وذلك كله من جذرٍ واحد (الموسى، 2000، ص 35).

أما ندوة استخدام اللغة العربية في الحاسب الآلي التي عقدت في الكويت من (14) إلى (16) أبريل (1985م)، فتناولت مباحث عربية حاسوبية في اتجاه تمثيل النّظام الصوتي، وأدلة قراءة النّظام الكتابي، ومعالجة النّظام الصرفي، وفهم النّص وتحليله بمدخلين معجم وصرفي (الموسى، 2000، ص 36).

وفي سنة (1996م) جعل مجمع اللغة العربية الأردني موضوع محور موسومه الثقافي الرابع عشر حول: "الحاسوب في خدمة اللغة العربية"، وكانت محاضراته متنوعة، حيث عرض نبيل علي بحثه حول "الحاسوب والنّحو العربي"، أما مأمون الحطاب فتناول "التحليل الصرفي للغة العربية باستخدام الحاسوب"، أما "المعجمات العلمية العربية المتخصصة ودور الحاسوب، فألقاها إبراهيم بن مراد" (حليمة، 2015 ،ص 31).

كما عرض محمد زكي خضر بحثه الموسوم بـ "الحرف العربي والحوسبة" في الموسم الثقافي لمجمع اللغة العربية الأردني في سنة 2001 (العناتي، 2006، ص 114).

بينما عقد "اتحاد ومجامع اللغة العربية" ندوة في عمان من (16 إلى 19) سبتمبر (2002م)، عرض فيها عبد الرحمن حاج صالح "دور النظرية الخليلية في النهوض بالبحوث الحاسوبية الخاصة باللغة العربية".

أما المؤلفات فهي قليلة جداً في هذا المجال، إذ يعدّ كتاب نبيل علي"اللغة العربية والحاسوب" الذي صدر سنة (1988م) أول مؤلف تناول موضوع اللسانيات الحاسوبية مطبقة على أنظمة اللغة العربية صوتًا، وصرفاً، ونحواً، ومعجماً مع المعالجة الآلية لهذه النظم اللغوية (العارف ،2007،ص 19).

والكتاب يقع في (591) صفحة، شاملًا تسعة فصول، بما في ذلك المقدمة، والخاتمة، وأربعة ملاحق، وقد حدّد في المقدمة الدّوافع التي تقف وراء البحث في موضوع اللغة والحاسوب، فكانت منها دوافع تقنية، ولغوية، وعامة، وكانت أغلب الفصول تتحدّث عن المعالجة الآلية لمنظومة اللغة العربية، والكتابة العربية، والصرف العربي، والنّحو العربي، والكلام العربي (الخناش، الخناش،2003،ص 256-296).

إنّ هذا الكتاب يمثل حجر الأساس في مسيرة البحث اللغوي العربي في اللسانيات الحاسوبية، بل حتى أنّ نهاد الموسى وصفه بأنه: "خطوة واسعة واثقة تنتظم لتأسيس اللسانيات الحاسوبية في العربية على أساس نظري وتطبيقي في آنٍ واحدٍ معًا" (الموسى، 2000، ص25).

وبعده بسنوات، نشر عبد ذياب العجيلي كتاب "الحاسوب واللغة العربية"، وهو كما قال عنه نهاد الموسى بأنه: "خطوة جزئية إيجابية نحو معالجة مسائل متنوعة من العربية بلغة برولوج (prolog)، وهو يمثل جهداً حميداً في هذا الاتجاه البيني (اللسانيات العربية الحاسوبية" (الموسى، 2000، ص27).

وفي سنة (2000) نشر كتاب "العربية نحو توصيفٍ جديدٍ في ضوء اللسانيات الحاسوبية" لنهاد الموسى"ويعدّ هذا الكتاب أول مؤلف في هذا العلم اللغوي الحديث يصدر عن متخصص في العربية وعلومها" (العارف، 2007 ص20).

ولذا فهو يمثل نقلةً نوعيةً في توظيف اللسانيات الحاسوبية لخدمة علوم اللسانيات العربية، والكتاب كما يذكر مؤلفه أنه محاولة في الانتقال من وصف العربية إلى توصيفها، وذلك في ضوء الأطروحة العامة للسانيات الحاسوبية. وقد اشتمل الكتاب على رؤى حاسوبية، حاول منها المؤلف إسقاطها على أنظمة العربية خاصة النحو، والصرف، والمعجم، إضافةً إلى تصويب الأخطاء النحوية والصرفية والإملائية (العارف، 2007 ص20).

وفي نفس السنة قام محمد جواد النوري بدراساتٍ إحصائيةٍ لغويةٍ حاسوبيةٍ للغة العربية، معتمداً في معطياته على المادة اللغوية لجذور الأفعال الثلاثية الواردة في المعجم الوسيط، الذي يعدّ نتاج هيئةٍ علميةٍ مرموقةٍ في عالمنا العربي، وقد أخرج الدكتور النوري هذا العمل اللغوي الحاسوبي في ثلاثة كتب، صدرت عن دار الجندي للنشر والتوزيع في (2016) نذكرهم كالآتي:

- لغويات حاسوبية، دراسة صوتية صرفية في أبواب الفعل الثلاثي في المعجم الوسيط باستخدام الحاسوب.
-لغويات حاسوبية، دراسة صوتية صرفية في جذور الأفعال الثلاثية.
- لغويات حاسوبية، دراسة صوتية صرفية في الأفعال الثلاثية المزيدة (باهي،2019 ، ص387).

كما أصدرت سلوى السيد حمادة كتاباً موسوماً بـ "المعالجة الآلية للغة العربية - النظرية والتطبيق-" الكتاب يقع في نحو (374) صفحة، جاء مقسماً إلى سبعة أبواب، تعلق الباب الأول بتقديم التحليل الحاسوبي، بينما اختص الباب الثاني بالتحليل الصوتي، وعالج الباب الثالث التحليل الصرفي، أما الباب الرابع فكان عن التحليل النحوي، والباب الخامس عن التحليل الدّلالي، والباب السادس يحتوي نموذجاً تطبيقياً لبرامج عديدة تعمل على صحة فروض الجزء النظري، والباب السابع تناول فيه العمل المستقبلي للكتابة (باهي، 2019، ص387-388).

ولا ننسى جهود محمد مراياتي بالتعاون مع زملائه العاملين في مركز الدّراسات والبحوث العلمية في سوريا، تلك الدّراسة التي تدور حول إحصائية الجذور العربية.

https://doi.org/10.20428/jss.v29i4.2243

فقد درس مراياتي الجذور العربية المنتشرة في المعاجم والقواميس العربية القديمة دراسة حديثة، معتمداً في ذلك على الحاسبات الإلكترونيّة التي تساعد كثيراً في ضبط العملية الإحصائية والسرعة العلمية فيها، وهو ما دفع "مراياتي" لأن يحصي النسب المئوية للجذور الثنائية الثلاثية والرباعيّة والخماسية في العربية، وقد دفعه أيضًا لأن يحصي الدّرجات المئوية التي يمكن فيها للأصوات العربية أن تندمج مع بعضها بعضًا، أو تنفصل عن بعضها بعضًا، ثم القوانين التي تحكم هذا الدّمج والانفصال، والواقع أنّ هذه الدّراسات الإحصائيّة لجذور الكلمات العربية مهمة، بحيث يمكن استخدام نتائجها في الترجمة الآلية من اللغة العربية إلى اللغة الأجنبية الأخرى أو بالعكس، ولاسيما من حيث مقابلة المركبات الصوتية الأجنبية، ومن حيث التحليل والتركيب، وقد دعا "مراياتي" هذا الإجراء تنافر الأصوات العربية وانسجامها، وإمكانية اكتشاف مثل هذا التنافر والانسجام مبرمجًا في الحاسبات الإلكترونية (ديدوح، 2009، ص 89).

وأيضًا جهود عبد القادر الفاسي الفهري حول حوسبة المعجم العربي، وجهود محمد الحناش الذي قام بدراسة حول المعجم الإلكتروني، كما اقترح مشروع نظرية حاسوب لسانيات في سبيل بناء معاجم آلية للغة العربية، ونجد أيضًا محمد حشيش الذي قدم لنا بدوره بحوثًا حول معالجة اللغة العربية بالحاسوب (بن يشو ، 2014).

رغم هذه الجهود التي بذلت من طرف الباحثين العرب، إلا أنّها في واقع الحال جهود يعوزها التنظيم والتنسيق على المستوى العربي.

ثالثًا: مجالات اللسانيات الحاسوبية:

للسانيات الحاسوبية مجالات كثيرة، ولعلّ أهمها:

1.3الترجمة الآليّة:

إنّ مصطلح (ترجمة) يشير إلى الترجمة بمعناها القديم والشائع، أي إلى ما ينبغي أن ندعوه الآن (الترجمة البشرية) بعد أن ظهر أسلوب جديد في الترجمة، هو الترجمة الآلية، وتعدّ الترجمة من أهم مجالات اللسانيات الحاسوبية، وإحدى غاياتها الأساسية.

جاءت مصطلحات تعليم الترجمة تحت مادة الترجمة الآلية، وهي الترجمة بواسطة برنامج معلوماتي معدّ لتحليل النص المصدر، ولإنتاج النص الهدف من غير أي تدخل بشريٍّ في الترجمة الآلية، ويكون المترجم في خدمة الآلة، فيما تكون الآلة في خدمة المترجم في حال اللجوء إلى الترجمة بمساعدة الحاسوب (جوزف ، 2012 ، ص 33).

وتعدّ الترجمة أول تطبيقات بحوث (الذكاء الاصطناعي)، ويطلق الذّكاء الاصطناعي "على محاكاة ذكاء الإنسان، وكيفية استخدام خبرته المكتسبة في مجالٍ معينٍ بواسطة الآلة، وخاصة باستخدام أنظمة الحاسبات، وتتضمن عملية المحاكاة التعلم، أو الحصول على المعلومات، وقواعد استخدام تلك المعلومات للوصول إلى استنتاجات محددة، أو تقريبية، وتصحيح الأخطاء ذاتيًا، ويستخدم الذكاء الاصطناعي فيما يسمى بالأنظمة الخبيرة" (الربيعي وآخرون، 2001، ص13).

تحاول تطبيقات الذكاء الاصطناعي أن تجعل الآلات تفعل أشياء تتطلب الذكاء إذا ما فعلها البشر، إذ يتطلب من الحاسوب تحويل المعنى من اللغة المترجم منها (اللغة المصدر) إلى اللغة المترجم إليها (اللغة الهدف)، فلإنجاز هذا التحويل يحتاج الحاسوب إلى بعض الصياغات اللغوية المناسبة ليتمكن من استيعاب المفاهيم، وبالتالي ليتمكن من نقل المعنى من لغة المصدر إلى لغة الهدف بالكفاءة نفسها، والقدرة التي يتمتع بها الإنسان في هذه العملية.

-الوسائل الداعمة للترجمة الآلية:

من أهم الوسائل التي تعتمدها الترجمة الآلية برامج التعرف الآلي على الكلمة، والمعاجم الأحادية اللغة، أو الثنائية، أو متعددة اللغات، وبرامج التحليل الصرفي للغة، والتحليل الدّلالي، والتحليل التركيبي النحوي (مدكور، 2011، ص 902).

أ. التعرف الآلي للكلمة:

فعند إدخال النّص المراد ترجمته -سواءً أكان بالعربية أو بلغةٍ أخرى كالإنجليزية – يحتاج أولًا التعرف على الكلمات التي أدخِلت، فإن كان الإدخال بالعربية، كان التعرف هو الخطوة الأولى في عملية الترجمة، وإن كان الإدخال بلغةٍ أخرى، والمخرج باللغة العربيّة، كان التعرف للغة الأخرى ضرورةً أوليّةً لإتمام الترجمة الآلية، ثم يأتي التعرف الآلي على الكلمات العربية المكافئة للمدخل (مدكور، 2011، ص 902-903).

ب. التحليل الصرفي

ويُقصد به استخلاص العناصر الأولية لبُنية الكلمة، وتحديد سماتها الصرفيّة والدّلالية التي يمكن استنباطها من البُنية (علي، 1987، ص 227)، ويتطلب المحلل الصرفي قائمةً بالسوابق، وثانيةً باللواحق، وثالثةً بالأوزان الصرفية، ورابعةً بالجذور، وخاصةً بالكلمات الجامدة (مدكور، 2011، ص 904)، ويستخدم المعالج الصرفي في أعمال حاسوبية متعددة مثل: اختصار النّصوص العربية، واسترجاع النّصوص، وتصحيح الأخطاء الإملائية المتعلقة بالصرف، ودعم التشكيل العربي (علي، 1987، ص 315).

ج. التحليل التركيبي النحوي:

ويُقصد به استخلاص العناصر الأولية لبُنية الكلمة، وتحديد سماتها الصرفيّة والدّلالية التي يمكن استنباطها من البنية، وللمعالج النّحوي مستويان رئيسان هما: مستوى التمييز النحوي، وفيه يكون الحكم على التركيب بالصحة النّحوية ومستوى الإعراب، أي التحليل التركيبي لبُنية الجملة، ووظائف عناصرها، والتقديم والتأخير، واستنباط الحذف، وبذلك يقوم المحلل برد البُنية السطحية إلى بُنية عميقة (علي، 1987، ص 388).

ولكن هناك مشكلات تواجه التحليل التركيبي النحوي للغة العربية، أهمها:

- مرونة العربية الكبيرة، كما في الذكر والحذف، والتقديم والتأخير، مثل التناوب بين التركيب المبدوء بالفعل (فعل+فاعل+مفعول به) والتركيب المبدوء بالاسم (اسم *مبتدأ*+فعل-فاعل مستتر (ضمير مستتر وقد يكون متصلاً يعود على المبتدأ) +مفعول به.
- تعدد العلامات الإعرابية، كالاسم الواقع بعد لاسيما والمستثنى والمنادى (مدكور، 2011، ص 905-906).

د. التحليل الدلالي:

يقع التحليل الدّلالي في بؤرة الأهمية للترجمة الآلية، فمن خلاله يفصل في الكلمات المتعددة المعنى والأضداد، والمترادفات، وتحديد أي المكافئات في اللغة الهدف أقرب لنقل الدّلالة نقلًا دقيقًا، وكانت البرمجيات القديمة تعتمد على تحليل الكلمة المفتاح، "وفيها يقوم برنامج معالجة اللغات الحيّة (nlpp) بالبحث في الجملة عن الكلمة المفتاح... ويمكن أن تستخدم هذه الكلمة في فهم الحاسوب للمدخلات" (راغب، 2000، ص 49). ولكن هذا الأسلوب محدود الفائدة، لذلك لجأ الباحثون إلى التحليل الدّلالي الذي يعتمد على نظريات علم الدّلالة مثل: المجالات الدّلالية التي تساعد برنامج الترجمة الآلية على تحديد مجال النّص، فورود الكلمة (الجذر) يحتمل مجالاتٍ علمية متعددة، لغوية، وزراعية، ورياضية، وتحديد المجال الدلالي للكلمة أو المصطلح هنا يعين على وضع

https://doi.org/10.20428/jss.v29i4.2243

المكافئ الصحيح باللغة الأخرى، خاصة حال وجود كلمة في لغة ما تكافأ بكلمتين أو أكثر في اللغة المقابلة، ومثال ذلك كلمة (uncle) التي تكافأ في اللغة العربية بـ (عم وخال) وربما يحتوي هذا المثال على صعوبة، فالكلمات (عم وخال) تقعان ضمن مجال دلالي واحد هو مجال الأقارب (مذكور، 2011، ص 907).

وتأتي الترجمة الآلية تتويجًا لكل الجهود المبذولة للمحللات الصرفية، والتركيبية، والدّلالية، والمعجم، والتعرف الآلي على كلمات النّص المدخل، فهي "مجال الذكاء الاصطناعي الذي يستخدم برامج معالجة اللغات الحية ...لفهم النص المكتوب باللغة الأولى، وترجمته إلى المعنى المرادف له في اللغة الأخرى عن طريق برامج إنتاج اللغة الحيّة" (مذكور، 2011، ص 908).

2.3- تحليل النصوص آليًا:

إن الكمّ الهائل من النصوص اللغوية يستدعي استخدام برمجياتٍ متخصصةٍ في تحليل النصوص، حيث تعمل هذه البرمجيات على إنجاز المهام الآتية:

-الإحصاء العددي ((count، وتتضمن استخراج النسبة المئوية لتكرار كلمة معينة في النص.

-تصنيف الكلمات، أي ترتيبها وتبويبها في فئات نحوية، كالأسماء، والأفعال، والصفات.

-توافق كلمة مع أخرى، أي معرفة الموافقة لكلمة معينة.

-السياق الذي ترد فيه الكلمة في النص. ويشمل الإحصاء العددي لكلمات النص اللغوي غير المعالج الخطوات الآتية: (بلعيد، 2016، ص 20).

-إحصاء العدد الإجمالي للكلمات في النص.

-إحصاء مقدار تكرار الكلمة الواحدة في النص.

-إحصاء العدد الإجمالي للجُمل في النص.

-عدد الأسماء والأفعال، والصفات، والحروف في النص.

ولكي نقوم بإعداد النّصوص العربية للمعالجة الآلية، نحتاج إلى عمليات عديدة:

-عملية التشكيل الإلقائي للنّصوص، وذلك تحاشيًا عن غياب التشكيل.

-عملية إعراب النّصوص العربية آليًا.

-ترقيم النّصوص آليًا (عبر وضع علامات الوقف) وذلك من أجل كشف البُنية السردية للجُمل والمقولات النحوية.

-استخدام معاجم الألفاظ المستخدمة (فهارس متخصصة للمصطلحات).

ويجب أن أنبه هنا إلى بعض الأمور التي يظن بعضهم أنها من صميم اللسانيات الحاسوبية وهي ليست كذلك "فلا يُعدّ من مجال اللّسانيات الحاسوبية تعليم الحاسوب باللغة العربية على غرار ما تقوم به معاهد تعليم الحاسوب في الدّول العربية، سواء أكانت المادة المدروسة تُعنى بطريقة استخدام برامج الحاسوب، أو تصميم تلك البرامج وتعديلها، أو صيانة الأجزاء الصلبة من حيث يكون الحاسوب عندئذٍ مادة الدرس، وتكون العربية وسيط التعليم ولغة التخاطب. ولا يُعدّ من تخصص اللّسانيات الحاسوبية ما بات يُعرف بمصطلح تعليم اللغة بمساعدة الحاسوب الذي يُعنى بثلاث مجالات، هي تعليم اللغة الثانية بمساعدة الحاسوب، واختبارات اللغة الثانية بمساعدة الحاسوب، وأبحاث تعلم اللغة الثانية بمساعدة الحاسوب، حيث يعدّ الحاسوب في هذا التخصص وسيلة من وسائل تعليم اللغة الثانية شأن السبورة،

وجهاز العرض فوق الرأسي، والبطاقات، واللوحات، وأشرطة التسجيل السمعية والبصرية، ويتميز الحاسوب بقدرته على توفير إمكانات كل تلك الوسائل مجتمعة، إلى جانب إمكانية التعلم الذاتي الذي لا يستعين فيه الطالب بمعلم.

كذلك لا يُعدّ من صميم اللسانيات الحاسوبية الأبحاث التقنية التي تعنى بآليات تخزين المعلومات في وسائط التخزين الحاسوبية المختلفة التي وثبت وثباتٍ سريعةٍ لتنتقل من القرص المرن إلى القرص الصلب فالمدمج فالذاكرة الضوئية المتميزة بصغر الحجم وضخامة السعة، إذ يُعدّ الحاسوب في هذا المجال مجرد وعاءٍ للغة، شأن الورقة البيضاء، أو سعف النخل، وألواح الخشب بادئ أمر التدوين" (كنالي، 1434 هـ، ص 08).

رابعًا: مشكلة حوسبة اللغة العربية

إنّ الدّافع لتوجه الحاسوبيين واللغويين العرب اليوم لحوسبة اللغة العربية أسباب كثيرة، نذكر منها ما يأتي: (علي، 1987، ص 78-80).

1 -إعادة الاعتزاز باللغة العربية وتراثها، وتكثيف تدريسها في مواد التعليم العام.

2-توجيه مستخدمي المعاجم العربية إلى أهمية المعاجم الإلكترونيّة.

3-استخدام نظام قواعد النصوص الكاملة في حفظ النص القرآني، واسترجاعه بتأثيرات إلكترونيّة متنوعة.

4- استخدام نظم المعلومات لتحليل مادة الحديث الشريف، وعدم الاكتفاء بعرض سلاسل الرواة، أو بتخريج الأحاديث، أو بالاهتمام بالمضمون وحسب، ومحاولة تنسيق ارتباط الحديث الواحد دلاليًا بغيره من الأحاديث المنتمية للموضوع الواحد.

5- إقامة قواعد بيانات تنظيمية للفهرسة للوثائق وللدّراسات، فضلًا عن الأحكام الخاصة بالعقيدة، والأخلاق، ونحو ذلك.

6- إقامة قاعدة بيانات لذخيرة النّصوص العربية في المجالات الإنسانية والعلمية من الفكر، والفنون، والأدب والشعر، وغير ذلك.

7- استخدام نظام المعلومات في تحقيق التراث باستخدام الحاسوب في تسهيل قراءته، عبر أساليب التكبير الرقمي، وتقديم الدعم المعجمي للباحثين عبر دمج المعاجم التاريخية بهذه النظم.

8- ضرورة تكثيف الجهود في مجال التعريب والترجمة، وتشجيع البحوث باللغة العربية في مجال العلوم الحديثة من أجل تعريب المصطلحات في هذا المجال.

9-تكثيف الصفحات العربية في الإنترنت، ووضع المعاجم اللغوية في الشبكة العنكبوتية، ووضع دروسٍ مناسبةٍ لتعليم اللغة العربية للناطقين بها وبغيرها.

ومن المعلوم أنّ اللغة العربية تتوفر على الكثير من الخصائص التي تجعلها قابلة للحوسبة، إذ هي تمتلك الخصائص التي تتوافق مع متطلبات الحوسبة من اعتماد معجمها على الجذور، وخاصية الاشتقاق الصرفي والمدونة النّحوية – أي الحرية النسبية في ترتيب الكلمات داخل الجمل – والصلة الوثيقة بين المبنى والمعنى، واطراد القياس في كثير من الحالات الصرفية، والإعرابية، والصوتية، والذي أدى ببعضهم إلى وصف العربية بأنها لغة جبرية (خضر ، 2007)، وقد وصفت اللغة العربية أيضًا "بأنها لغة ذات نظام دقيق وأنيق تركيبيًا ودلاليًا ومعجميًا" (الوعر، 1989 ، ص 377)، ولكن بالرّغم من هذه السّمات والخصائص تظل اللسانيات الحاسوبية العربية بعيدة كل البُعد عن نظيرتها عند الغربيين، الذين قطعوا أشواطًا كبيرةً في هذا المضمار، ومرد ذلك إلى:

https://doi.org/10.20428/jss.v29i4.2243

1-أنّ الحاسوب الذي ابتكره الغربيون قد صُمّم طبقًا للغاتهم الأجنبية، ومتوائمًا مع مواصفاتها، لذلك تواجه العلاقة بينه وبين اللغات الأخرى ومنها العربية صعوبات ومشكلات، مردها عدم ملاءمته لطبيعتها وخصائصها (علوي، العناتي، 2009، ص 137).

ومن بين هذه الصعوبات التي واجهتها اللغة العربية في الحاسوب صعوبة في إدخال حروف العربية، حيث إنّ الصعوبة لا تكمن من حيث عددها البالغ ثمانية وعشرين حرفًا، وإنما المشكلة في الحروف كانت تكمن في كونها حروفًا متصلة وليست منفصلة كالحروف الإنجليزية، إضافةً إلى أنها تبدأ من اليمين إلى اليسار عكس اللغات اللاتينية التي تبدأ من اليسار إلى اليمين، كما أنّ مشكلتها الأخرى تتعلق بالشكل، حيث إنّ الحركات القصيرة (الفتحة، والضمّة، والكسرة) هي حروف في اللغات اللاتينية: كالفرنسية والإنجليزية، تأخذ حيّزًا ومكانًا كالحرف، بينما لا يوجد مكان للحركة في العربية، ثم إنّ الحرف الواحد يتعدّد شكله عند إخراجه بالكتابة على شاشة الحاسوب في أول الكلمة له شكل، وفي وسطها له شكل، وفي نهايتها أيضًا له شكل، مثل: العين، ففي أول الكلمة يتخذ شكل(عـ)، وفي وسطها(ـعـ)، وفي آخرها (ـع، ع).

ولحل مشكلة الشكل في الحروف، كان لابد من تبنّي بعض المبادئ التصميمية، أوّلها مشروع الباحث المغربي أحمد الأخضر غزال، رئيس معهد الدراسات والأبحاث للتعريب سابقًا في الرباط، الذي وضع من تصميم أنجزه، الحركة بشكل منفصل عن الحرف، فكلمة "كتب" مثلًا تكتب هكذا (ك/ت/ب) (Kataba)، كما صمم طريقةً معينةً لكتابة الحروف بتوحيد شكل الحروف في بداية الكلمة ووسطها ونهايتها (علوي، والعناتي، 2009، ص 137).

ومن بين الصعوبات أيضًا نجد قضية التحت، فلا يمكن لجهاز الحاسوب أن يكون قادرًا على فهم ظاهرة التحت ما لم يزوّد بطريقةٍ قادرةٍ على اختيار الأحرف المطلوبة، وتجاهل البقية (يوسف، 2014، ص 31)، ومن أمثلته:

-البسملة، منحوتة من "بسم الله الرحمن الرحيم".

-الحمدلة، منحوتة من" الحمد لله".

-كبّر، منحوتة من "الله أكبر".

2- غياب الخبرة اللسانية التي تعتمدها جلّ التطبيقات الحاسوبية، حيث اعتمدت على بعض اللغويين ذوي التكوين التقليدي، إلا أنّ هؤلاء غالبًا ما يكونون غير قادرين على فهم متطلبات الحاسوب، فهم يقدمون معلومات ومعطيات أكاديمية صحيحة، إلا أنها غير قادرة على الاستجابة لمتطلبات المبرمجين مما ينعكس سلبًا على تطوير البرنامج الموضوع (نبيل، 1987، ص 20).

3-عدم تقديم الدّعم اللازم للقطاع الخاص المبادر في هذا المجال.

4-ندرة مراكز البحوث الأكاديمية النّظرية والتطبيقية في مجال اللسانيات الحاسوبية.

5-بعثرة الجهود الفردية، سواءً على المستوى النظري أو التطبيقي، فكل باحث وكل منظمة تعمل بمعزلٍ عن غيرها، إضافة إلى محدودية الدراسات في هذا المجال كمًّا ومستوى، فإنها تعاني من ضعف الانتشار وانعدام التكامل والتعامل بينها، ولا تكاد تتجاوز الملتقيات والندوات (الخطيب، 1998، ص 83).

6-الانفصال بين النّظري والتطبيقي في مجال اللسانيات الحاسوبية، وذلك نتاج الانفصال بين نظر اللغوي وتطبيق الحاسوبي، إذ لا يعقل أن ينهض المرء لمعالجة العربية بالحاسوب، وهو يفتقر إلى الحد الأدنى من المعرفة اللغوية (آل طه، 2005، ص 19)، لأن المعالجة الآلية لا يمكنها أن تتعامل إلا مع الدّقيق والمضبوط، والمكتمل، لذا فهي تتطلب

الكشف عن دخائل البنية الدفينة للغة العربية، وتقتحم الكثير من المجالات التي لم يتطرق إليها البحث من قبل، واتخاذ مواقف محددة تجاه الكثير من النقاط المتخلف منها (الموسى، 2000، ص360)

7-إنّ اللساني الذي نحتاج إليه في وضع برامج الهندسة اللسانية هو الذي يتمكن، ويساير مختلف التطورات النظرية التي تعرفها اللسانيات الصورية اليوم، والقادر على وضع الخوارزميات للسانية لمختلف مستويات نظام اللغة العربية، وهذا الصنف من الباحثين هم الذين يتمكنون من الدّفع باللغة العربية إلى مصاف اللغات العالمية في الحوار مع الآلة، ولم يعد العمل في حوسبة اللغة مجرد موضة وترف يمارسه بعض الهواة، بل هو ذلك الجسر الذي يمكن منه العبور نحو قرننا الحالي الذي نعيش فيه (فاهم، 2015، ص 133).

5- توصيات الدراسة:

إذا كانت اللغة العربية تقوم على مقومات مؤهلة لحوسبة اللغة فلابد:

1. أن نوفر للغة العربيّة محلّلات صرفية ونحوية، فعلى الرغم من أهمية هذه المحلّلات الموجودة إلا أنها تعاني نقصًا في جانب من جوانبها.
2. لابد أن نولي مكانة المعاجم اهتمامًا بارزا، فنحن ما زلنا نعاني من أنّ المعجم العربي يفتقر إلى معجم تاريخي.
3. لابدّ من إيلاء المحتوى العربي اللغوي الإلكتروني اهتمامًا بارزًا، فمازالت الموضوعات التي يتناولها الناشرون أقل من الطموح ولا نقلل من شأنها، ولكن أمام هذا الانفتاح يجب أن تكون هناك مواكبة.
4. الاهتمام بالمصطلحات العلمية، فلغة من غير مصطلحات علمية لن تكون قادرةً على المنافسة.
5. إدخال (مقرر حوسبة اللغة) ضمن برنامج اللغة العربية.
6. تصميم معاجم إلكترونية خاصة بالمصطلحات النقدية، والبلاغية، والأسلوبية، والنحوية، والصرفية، فضلاً عن اللغوية.
7. عقد لقاءات دورية بين برنامجي اللغة العربية وعلوم الحاسب الآلي في الجامعات، من شأنها أن تثري الدّرس اللغوي الحاسوبي.

https://doi.org/10.20428/jss.v29i4.2243

قائمة المراجع:

آل طه، هدى عبد الله (2005). *النظام الصرفي للعربية في ضوء اللسانيات الحاسوبية مثل جمع التكسير*، رسالة دكتوراه، الجامعة الأردنية عمان، ص(19).

أنيس، إبراهيم (1973). النظامة الإلكترونيّة تحصي جذور مفردات اللغة العربية، *مجلة اللسان العربي*، *10*(1)، ص.(211) .

باهي، فتحي (2019). اللسانيات الحاسوبية العربية واقع وتحديات، *دراسات وأبحاث*، *11* (1)، ص (387 – 388).

برهومة، عيسى (2007). *اللغة العربية وأسئلة العصر*، ط1، دار الشروق، عمان، ص (39).

بلعيد، عادل (2016). *أثر اللسانيات الحاسوبية في تعليمية اللغة العربية*، مذكرة مقدمة لنيل شهادة الماستر ، ص (20).

حافظ، إسماعيل علوي، والعناتي، وليد (2009). *أسئلة اللغة، أسئلة اللسانيات*، دار العربية للعلوم، ناشرون، الرباط، ص (111 – 127).

حوسبة المعجم العربي الواقع والآفاق د.جيلالي بن يشو ، الرابط :https://aladdin.7olm.org/t3654-topic 28 ي ناير 2014.

الخطيب، حسام (1998). العربية في عصر المعلوماتية – تحديات عاصفة ومواجهة متواضعة، *المركز العربي للتعريب والترجمة والنشر*، *8* (15)، ص.(83)

الخناش، محمد (2003). اللغة العربية والحاسوب، *مجلة التواصل اللساني*، *9* ، ص.(296-256) .

ديدوح، عمر (2009): فعالية اللسانيات الحاسوبية العربية، مجلة الآداب واللغات، العدد 8، الجزائر، ص.(89)

راغب، أحمد (2009). *الانحراف الدلالي في النص العربي المترجم آليًا عن الإنجليزية: مظاهره وآثاره*، ط1، العالم العربي للنشر والتوزيع، دبي، ص.(49) .

الربيعي، محمود ، وآخرون (2001). *المعجم الشامل للمصطلحات الحاسب الآلي والإنترنت*، مكتبة العبيكان، الرياض، ص.(13)

طانيوس، جوزف (2012). *المعلوماتية واللغة والأدب والحضارة (الرقم والحرف)*، ط1، المؤسسة الحديثة للكتاب، ص.(17) .

العارف، عبد الرحمن بن حسن (2007). توظيف اللسانيات الحاسوبية في خدمة الدراسات اللغوية العربية (جهود ونتائج)، *مجلة مجمع اللغة العربية الأردني*، *12* (2)، الأردن، ص (15-20)

عبد القادر، عبد الجليل (2002). *علم اللسانيات الحديثة*، ط1، دار الصفاء، الأردن، ص.(181) .

العناتي، وليد (2006). *دليل الباحث إلى اللسانيات الحاسوبية العربية*، ط1، دار جرير للنشر والتوزيع، ص.(114)

العناتي، وليد أحمد ، (2005). اللسانيات الحاسوبية (المفهوم، التطبيقات، الجدوى)، *مجلة الزرقاء للبحوث والدراسات*، *7*(2)، ص (62).

فاهم، سعيد (2015)، قراءة في الإسهامات اللسانية الحاسوبية العربية، *مجلة دراسات لجامعة الأغواط، 36*، ص(133).

كنالي، وجدان محمد صالح (1434). *اللسانيات الحاسوبية العربية المنهج والإطار*، بحث قدم في المؤتمر الدولي الثاني للغة العربية في دبي، 30 جمادى الآخرة، 1434 هـ، ص.(08)

اللغة العربية والحاسوب قراءة سريعة في الهندسة اللسانية العربية، ص (20)

مازن، الوعر (1989). *دراسات لسانية تطبيقية*، دار طلاس، دمشق، ص (320 – 377).

مدكور، عمرو محمد فرج، (2011). الترجمة الآلية مفهومها نماذج تطبيقية في اللغة العربية، *مجلة كلية دار العلوم*، (6)، ص (902 – 908).

موسى، علي حلمي، (2001). *حوسبة التراث العربي*، محاضرة ألقيت في مجمع اللغة العربية الأردني في (17) نيسان (أبريل) 2001.

الموسى، نهاد (2000). *العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية*، ط 1، المؤسسة العربية للدراسات والنشر، ص (25-35).

خضر ،محمد زكي (2007) . الحرف العربي والحوسبة، الجمعية *الدولية لمترجمي العربية*. تم الاسترجاع من الرابط https://www.atinternational.org/forums/showthread.php?t=2172 في 23 سبتمبر 2023.

نبيل، على (1987). اللغة العربية والحاسوب، *مجلة عالم أفكار، 18* (3)، ص (78 – 388).

يوسف، حليمة (2015). *اللسانيات الحاسوبية وآليات حوسبة الحرف القرآني*، ص(31)

https://doi.org/10.20428/jss.v29i4.2243