في حوسبة اللغة العربية: الواقع والتحديات

In computing the Arabic Language: reality and challenges

النذير ضبعي

جامعة العربي بن مهيدي –أم البواقي- الجزائر

email:Nadirdobai80@gmail.com

تاريخ الإرسال:29- 09-2022 تاريخ القبول: 09-12-2022

ملخص:

يروم هذا البحث تقديم رؤية واضحة عن واقع المعالجة الآلية للغة العربية، وسعيا لتحقيق هذه الغاية فقد لزم الوقوف على جملة من القضايا المتعلقة بحوسبة اللغة العربية وهندستها، انطلاقا مما تقدمه التكنولوجيا من تقنيات وأدوات وبرامج تجعل الآلة طيعة ومرنة أمام الاستخدام اللغوي المتشعب.

وللإحاطة بهذا الموضوع حاولنا التطرق لخصوصيات اللغة العربية التي تجعلها أكثر قابلية للمعالجة الحاسوبية في جميع مستوياتها الصوتية والصرفية والتركيبية والدلالية. مع ذكر التحديات التي تواجه حوسبة اللغة العربية، وآفاق حوسبتها التي توفرها الآلة ويمليها الذكاء الاصطناعي.

الكلمات المفتاحية:لسانيات حاسوبية؛ معالجة؛ تحديات.

Abstract:

This research aims to provide a clear vision of the reality of the automatic processing of the Arabic language, and in order to achieve this end, it was necessary to identify a number of issues related to the computerization and engineering of the Arabic language, based on the technologies, tools and programs offered by technology that make the machine pliable and flexible in the face of complex linguistic use

In order to understand this topic, we have tried to address the specifics of the Arabic language, which makes it more amenable to computer processing at all its phonemic, morphological, syntactic and semantic levels. With mentioning the challenges facing the computerization of the Arabic language, and the prospects for its computing provided by the machine and dictated by artificial intelligence.

Keywords:computational linguistics; processing; Challenges.

مقدمة:

لقد أحدثت التكنولوجيا الرقمية ثورة علمية كبرى مست جميع الميادين، ولم تكن اللغة في منأى عن ذلك، حيث حاول العلماء الاستفادة من تطبيقات الحاسوب لخدمتها، وهو ما ولد علما بينيا جديدا يطلق عليه: اللسانيات الحاسوبية، هذا الأخير يجمع بين علمين مستقلين هما علم الحاسوب واللسانيات.

وقد شهدت اللسانيات الحاسوبية تطورا كبيرا خاصة في السنوات الأخيرة، حيث ظهرت برامج وتطبيقات حاسوبية مكنت الباحثين من حوسبة اللغة وهندستها، الأمر الذي حقق نتائج عظيمة للدراسات اللغوية في جميع مجالاتها كالتعلم والتعليم والترجمة والإحصاء اللغوي...

ولا شك أن للغة العربية تحديات حالية ومستقبلية متعددة، مما يتطلب النهوض بواقعها ومجاوزة حالة التهميش والركود في معالجة مشكلاتها، ولن يتأتى ذلك إلا بإسهام عربي يواكب التطور الحاصل في مجال حوسبة اللغة، وذلك بتسخير علم الحاسوب لخدمة الدرس اللغوي العربي.

وفي كنف هذا التصور يندرج موضوع هذا البحث الذي نحن بصدده، من حيث إنه يعنى بواقع اللسانيات الحاسوبية العربية، ويسعى إلى بيان دور اللسانيات الحاسوبية في نهضة المنجز اللساني العربي، والتحديات التي تواجهها حوسبة اللغة العربية، مستندا في ذلك على المنهج الوصفي والمنهج التحليلي.

لأجل ذلك تخيرت للمقال العنوان الآتي: في حوسبة اللغة العربية: الواقع والتحديات.

ويقوم هذا البحث في جوهره على فكرة رئيسة تلخصها الإجابة عن الإشكال الآتي:
ما واقع اللسانيات الحاسوبية العربية، وما هي التحديات التي تواجهها؟

1-نشأة اللسانيات الحاسوبية:

كانت بدايات اللسانيات الحاسوبية حسب "مارتن كاي" عام 1949 عندما كتب "وارن ويفر" مذكرته الشهيرة التي يشير فيها إلى إمكانية بناء نظام للترجمة الآلية، وقد تلا ذلك

عقد أول مؤتمر للترجمة الآلية في معهد "ماساتشوستس" للتكنولوجيا سنة 1952، ثم صدرت مجلة علمية بعنوان الترجمة الآلية في 1954"[1].

أما مصطلح اللسانيات الحاسوبية فقد بدأ استعماله في منتصف الستينيات، ويرجح أن "ديفيد هيز" هو أول من "أطلق هذا المسمى على هذا المجال، ثم أتى عدد من المتخصصين الذين كان لهم دور في تطور هذا المجال مثل "تشومسكي"، و"جون كوك"[2].

وقد ظلت علوم اللغة العربية في منأى عن الانتفاع بها ردحا من الزمن، "حتى التقى أستاذ الفيزياء النظرية علي حلمي موسى بالدكتور إبراهيم أنيس في سبعينيات القرن الماضي، هذا الأخير الذي اقترح عليه إمكانية الاستفادة من الحاسوب في البحوث اللغوية، وكانت من ثمرة ذلك صدور الدراسة الإحصائية للجذور الثلاثية وغير الثلاثية لمعجم الصحاح للجوهري"[3].

وقد "جاءت نتائج الدراسة في صورة جداول إحصائية لجذور اللغة العربية وحروفه، وتتابع أصواتها وخصائص حروفها، مع دراسة تحليلية موجزة في التفسير اللغوي لما ورد في تلك الجداول. وقد فتح هذا التوجه في الفكر اللغوي العربي المعاصر بابا واسعا للباحثين في الدراسات اللغوية للولوج من خلاله إلى علم الحاسوب، وتسخيره في خدمة الدرس اللغوي"[4]. ثم "تبع ذلك دراسات أخرى لم تخرج عن هذا الإطار الإحصائي، منها دراسة إحصائية لمعجم لسان العرب عام 1973"[5].

وفي عام 1990 قام كل من "تيم باك والتر، وكين بيسلي"بتطوير محلل صرفي لكلمات اللغة العربية، يعتمد على أسلوب التحليل الصرفي ثنائي المستوى، ثم قام "كين بيسلي" باستخدام تقنية جديدة للتحليل والتوليد وكان ذلك في سنة 1996، حيث أصبح المحلل يقبل الكلمات العربية المشكولة كليا أو جزئيا، ويقدم عددا من الحلول الممكنة

---

[1] - الفيقي عبد الله، مدخل إلى اللسانيات الحاسوبية، (2017)، مركز الملك عبد الله عبد العزيز الدولي، ط1، داروجوه للنشر والتوزيع المملكة العربية السعودية، ص6.

[2] - المرجع نفسه، ص6.

[3] - رباحي محمد، (2019)، آفاق حوسبة المعاجم العربية، مجلة اللغة العربية، العدد 43، ص 133.

[4] - فاهم سعيد، 2015، قراءة في الإسهامات اللسانية الحاسوبية العربية، مجلة دراسات لجامعة الأغواط، العدد 36، ص130.

[5] - رباحي محمد، (2019)، آفاق حوسبة المعاجم العربية، ص 133.

للكلمة المحللة، ويقلل عدد تلك الحلول مع استخدام علامات التشكيل. وفي السنة نفسها قامت شركة حوسبة النص العربية في عمان ببناء محلل صرفي عربي يهدف إلى الوصول إلى محلل قواعدي يستخلص جذر الكلمة، ويقدم معلومات الضمائر[1].

وفي عام 2005 بدأ العمل في شبكة الكلمات العربية كأحد مشروعات شبكة الكلمات العالمية، وقد جرى في البداية إنشاء المحتوى المعجمي لشبكة الكلمات العربية نقلا -عبر الترجمة- عن محتوى الإصدار الثاني لشبكة "برينستون" للكلمات الإنجليزية. ويضم محتوى شبكات اللغة العربية حوالي تسعة آلاف فئة ترادف، وحوالي عشرين ألف مفردة، وحوالي مئة ألف من الأزواج المرتبة (مفردة، فئة ترادف)، كما أجريت بعض التحسينات على هذا المحتوى المعجمي لتوسعة تغطيته عبر توظيف أدوات معالجة لغوية لتمكين التعامل بمرونة مع البنية الصرفية للكلمات العربية. وقد كانت أهم وجوه النقد التي توجه لشبكة الكلمات العربية هي أن محتواها لم يُفَصّل من الأساس على مقاس اللغة العربية، بل جرت ترجمته من اللغة الإنجليزية، وكذلك ما واجه تطويرها من تعثر التمويل نهاية عام 2007 وبداية عام 2008. ورغم أن فريق العمل القائم على تطوير شبكة الكلمات العربية مازال يعدّها في طريق طويل من التطوير والتحسين، فإن هذا العمل يبقى ذا قيمة معتبرة خاصة وأنه متاح للجميع[2].

2- مفهوم اللسانيات الحاسوبية:

لقد أنتج هذا التطور التكنولوجي الهائل مباحث لسانية جديدة، تجمع بين اللغة والذكاء الاصطناعي، يطلق عليها اللسانيات الحاسوبية؛ وهي علم يجمع بين اللسانيات وعلم الحاسوب لمعالجة اللغة الإنسانية معالجة آلية، وتسخير التطبيقات الرقمية لهندسة اللغة وحوسبتها، ويعرفها "نيوقس" بأنها: "فرع من علمي اللغة والحاسب، يهدف إلى تصميم برامج رياضية للتراكيب اللغوية، للتمكن من معالجة اللغة آليا عن طريق الحاسوب. كما يعرفها من وجهة نظر لغوية على أنها تشكيل للنظريات والنماذج اللغوية

[1] - ينظر: المهوبي عبد العزيز، (2017)، مدخل إلى اللسانيات الحاسوبية، مركز الملك عبد الله عبد العزيز الدولي،ط1، دار وجوه للنشر والتوزيع المملكة العربية السعودية، ص57.

[2] - ينظر: عطيةمحمد، مقدمة في حوسبة اللغة العربية، ط1، مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الرياض، ص195.

أو تنفيذها على الآلة، ويرى أنه بإمكاننا النظر إليها على أنها وسيلة لتطوير نظريات لغوية جديدة بمساعدة الحاسوب"[1].

وتشمل اللسانيات الحاسوبية: حوسبة اللغة، والترجمة الآلية، والتدقيق اللغوي، وتعليمية اللغات، والذخائر اللغوية... وتتضمن مكونين أساسيين: مكونا نظريا وآخر تطبيقيا؛ يختص المكون النظري "بمعرفة كيفية عمل الدماغ الإلكتروني. أما المكون التطبيقي فهو يختص بتسخير ذلك العقل لحل القضايا والمشكلات اللغوية، مما يعني ضرورة الالتقاء بين اللغويين والحاسوبيين والتعاون فيما بينهم، للخروج بنتائج تسهم في تذليل الصعاب، وحل المشكلات التي تواجه التحليل الحاسوبي للغات البشرية"[2].

وتقوم اللسانيات الحاسوبية بمكونها النظري والتطبيقي "على تصور نظري يتخيل الحاسوب عقلا بشريا، لاستكناه العمليات العقلية التي يقوم بها العقل البشري لإنتاج اللغة وفهمها وإدراكها، إلا أنها تستدرك على الحاسوب أنه جهاز أصم لا يعمل إلا حسب البرنامج الذي صمّمه له الإنسان، لذا ينبغي توصيف المواد اللغوية له توصيفا دقيقا بحيث تستنفذ كافة الإشكالات التي يستطيع الإنسان إدراكها"[3].

إن هذه المعالجة الآلية للغة عمقت البحث اللغوي، إذ لم يعد الاهتمام منصبا على دراسة الظاهرة اللغوية في بعدها السطحي، بل تحول البحث اللساني مع اللغات الصورية الدقيقة إلى ما هو أعمق من ذلك بالبحث في سر تكوين الآلية اللغوية، وتخزينها في دماغ البشر، على شكل خوارزميات لغوية[4].

ومن أهم المجالات التي يعنى بها البحث اللغوي الحاسوبي[5]:

- المدونات الحاسوبية والتنسيق المؤسسي لاستعمالات اللغة، خاصة اللغات الخاصة.

---

[1] - الفيقي عبد الله، (2017)، مدخل إلى اللسانيات الحاسوبية، ص6.

[2] - العناتي وليد، (2003)، اللسانيات التطبيقية وتعليم اللغة العربية لغير الناطقين بها، دار الجوهرة للنشر، عمان، ط2، ، ص55.

[3] - المرجع نفسه، ص55.

[4] - ينظر: رباحي محمد، (2019)، آفاق حوسبة المعاجم العربية، ص 129.

[5] - ينظر: المجبول سلطان بن ناصر، (2018)، البحث اللغوي الآلي في المدونة الحاسوبية واللغة العربية، مجلة التواصل اللساني، جامعة سيدي محمد، فاس المغرب، المجلد19، العددان 1 و2، ص73.

- المدونات الحاسوبية والترجمة.
- المدونات الحاسوبية واللغويات الجنائية.
- المدونات الحاسوبية ووسائل التواصل والإعلام والاجتماع.
- المدونات الحاسوبية الخاصة بمتعلمي اللغة واكتسابها.
- المدونات الحاسوبية والتعليم الصفي.
- المدونات الحاسوبية وتصميم مواد تعليم اللغة.

3- المعالجة الآلية للغة العربية:

إن الاستفادة من التقنية الحاسوبية في الدراسات اللغوية أصبح ضرورة في عصرنا، هذا لأن استخدام الحاسوب في الأعمال اللغوية قد انتشر انتشارا واسعا، وقد نهضت الدراسات اللغوية على هذا الأساس نهوضا يضاهي التطور الحديث في سائر أنماط الحياة البشرية، وبخاصة النشاط الفكري الذي تعد اللغة أهم وسائله. ولقد أثبتت الأبحاث في هذا المجال أن الدراسات اللغوية يمكنها أن تنبعث بمزيد من الضبط والدقة والتقدم إلى ما هو أفضل باستعمال تقنية الحاسوب[1].

وتمتلك اللغة العربية نظاما خاصا يجعلها أكثر قابلية لأن تمثل حاسوبيا، فهي توفر مجالات كثيرة للمعالجة الآلية على اختلاف مستوياتها: الصوتية والمعجمية والصرفية والنحوية؛ لأنها لغة ذات نظام دقيق. "فلقد أثبتت اللغة العربية أنها من أكثر اللغات قابلية للمعالجة الآلية؛ لأنها تجمع بين كثير من الخصائص اللغوية المشتركة مع اللغات الأخرى، فأبجديتها مثلا ليست فونيمية صرفية كالإسبانية والفنلندية، حيث يقابل كل حرف صوتا واحدا، كما أنها ليست مقطعية كاليابانية، حيث رموز الأبجدية عبارة عن مقاطع يتكون كل منها من صامت يتبعه صائت مثل "ما، كي، فو" فالأبجدية العربية رغم كونها فونيمية أساسا، فإنها تتضمن حروفا ذات طبيعة مقطعية"[2].

[1] - ينظر: الأنصاري عبد الله، (1435هـ)، علاقة النحو العربي بالبرمجة الآلية للغة، جامعة الإمام محمد بن سعود الإسلامية، ص90.

[2] - بيومي سعيد أحمد، (2002)، أم اللغات، دراسة في خصائص اللغة العربية، ط1، ص105.

كما يتميز الصرف العربي بعدة خصائص تمكن المحلل الآلي من دراسة بنية الكلمة في اللغة العربية وتحديد سماتها، ومن أهمها[1]:

- وضوح مسار عملية الاشتقاق (الانتقال من الجذر إلى المشتقات الفعلية)
- اطّراد التصريف في العربية، باستثناء حالات نادرة.
- ميل الصرف العربي لتركيب الكلمات بالإضافة، وبعده عن تكوين الكلمات من خلال المزج والاختصار.
- انتظام بنية الكلمة العربية لثبوت رتبة عناصرها.
- شدة التداخل بين الصرف والفونولوجيا من حيث تعدد قواعد الإبدال والإعلال، وعمليات التغيير (الصرف- صوتية) الأخرى.
- الاشتقاق في العربية مبني على الأنماط الصرفية، حيث تتعدد هذه الأنماط مستخدمة عددا قليلا من حروف الزيادة.
- محورية مفهوم الجذر في العربية كعنصر ربط معجمي ودلالي.

إن كل هذه الخصائص جعلت اللغة العربية أكثر قابلية للمعالجة الآلية، وساعدت الشركات المهتمة ببرامج الحاسوب في تطوير البرمجيات الحاسوبية التي تربط اللغة العربية بشتى مجالات الحوسبة الصوتية والصرفية والنحوية والدلالية ...، حيث أنشأت هذه الشركات عدة برامج خاصة بحوسبة اللغة العربية من أهمها:[2]

- برنامج المكنز العربي الذي يُعنى بتخزين قواعد المعلومات والنصوص، وحصر جميع الفروع المعرفية المتاحة وفق قواعد علمية دقيقة ومعايير منهجية دولية، والتمكن من استرجاعها قصد تصنيفها وتحليلها حاسوبيا.
- المعاجم الإلكترونية: تُعنى بضبط الكلمات وتعريفها حاسوبيا.
- برنامج المحلل الصرفي الآلي: يعنى بتحديد بنية الكلمة العربية وأشكالها الصرفية وسماتها البنيوية.
- المحلل النحوي الآلي: يصحح الأخطاء النحوية ويضبط البنى التركيبية.

---

[1] - ينظر: المهوبي عبد العزيز، (2017)، مدخل إلى اللسانيات الحاسوبية، ص49.

[2] - ينظر: الميساوي خليفة ، (2012)، الذكاء الاصطناعي وحوسبة اللغة العربية: الواقع والآفاق، مجلة مدارات في اللغة والأدب، مركز مدارات للدراسات والأبحاث، تبسة، المجلد01، العدد 05، ص25-28.

- التشكيل والإعراب الآلي: يقوم بضبط قواعد التشكيل في الكلمات العربية وتعاملها حاسوبيا.
- برنامج المصحح الآلي: يصحح الأخطاء الإملائية والصرفية والنحوية.
- المصنف الآلي: يُعنى بتصنيف النصوص وتوثيقها آليا، وتحليل مصادرها المخزنة في قاعدة بيانات ضخمة، وتقديم بياناتها الإحصائية.
- المفهرس الآلي: يقوم بفهرسة الكتب آليا.
- الملخص الآلي: يقوم بتلخيص الوثائق، ويقدم الأفكار الأساسية لمحتوى الوثيقة.
- الرابط الآلي: يهتم بتنظيم المعلومات والحصول عليها داخل الوثيقة الواحدة، كتتبع كلمة في بحث.
- القارئ الآلي: يقوم بالتعرف على النصوص العربية وتخزينها وفق صيغ الكتابة.
- برنامج إدارة الوثائق العربية: يسهل التعامل مع الوثائق الإلكترونية.
- برنامج الأرشفة والمكتبات الرقمية: يقوم بتخزين الوثائق والكتب الرقمية ، ويسهل عملية الاطلاع على محتوياتها.

3-1- حوسبة المستويات اللغوية:

تستدعي المعالجة الحاسوبية للغة برامج ومحللات آلية، يختص كل محلل بمستوى لغوي معين، وله مهام متعددة؛ فالمحلل الصرفي الآلي يقوم بفضّ اللّبس الناجم عن نقص التشكيل أو غيابه، وذلك باستنباطه لجميع الاحتمالات الممكنة لتحليل الكلمة، ويتميز بقدرته على التعامل مع الأنماط اللغوية، وعناصر التشكيل وتداخل النحو والصرف، مع اهتمامه بما يلحق الكلمة من تغيرات صوتية ممثلة في عمليات الإعلال والإبدال والإدغام.

ويستطيع المحلل الصرفي الآلي استخلاص العناصر الأولية لبنية الكلمة في اللغة العربية وتحديد "سماتها الصرفية، والصرف صوتية والصرف نحوية، فيقوم بالكشف عن جذور الكلمة ووزنها الصرفي، ويبين ما يطرأ عليها من تغيير بالزيادة أو النقصان، والإعلال والإبدال، والإدغام والقلب، ويوضح ما يلحق عليها من سوابق، ولواحق وزوائد، إضافة إلى تقسيم الكلمة إلى اسم أو فعل أو حرف، وتقسيم الاسم إلى جامد أو مشتق، ومذكر أو مؤنث، ومفرد أو مثنى أو جمع... ويضم المحلل الصرفي مجموعة من قواعد المعطيات؛ هي قواعد معطيات معجمية لأوزان الكلمات العربية المستعملة،

وقواعد معطيات لأسماء الأعلام، وقواعد معطيات للأخطاء الإملائية والنحوية الشائعة في النصوص العربية[1].

وتأتي حوسبة المعاجم من ضمن أهم المباحث اللسانية الحاسوبية لما تحققه من أهداف تتعلق بالتداول والاستخدام، فحوسبة المعاجم فتحت مباحث جديدة، ووسعت مجالات الاستخدام المعجمي متجاوزة بذلك تلك المفاهيم التقليدية لدور المعاجم.

إن المعالجة الحاسوبية لا تقتصر على بعض مراحل الصناعة المعجمية دون غيرها، فهي وسيلة لإنجاز المراحل الثلاث في صناعة المعاجم، وهي: الجمع والتحرير والنشر، ففي مرحلة جمع المادة الأولية نفيد من المعالجة الحاسوبية في استخلاص المعلومات المعجمية، بداية بالمداخل والوحدات، ومرورا بالمعاني، وانتهاء بالشواهد، وتفيد في المعالجة الإحصائية في التأكد من صلاحية المدونة لتمثيل مجتمع اللغة، واستخلاص قائمة الكلمات الشائعة التي يتحدد من خلالها حجم المعجم وجانب من المعلومات المعجمية. وفي مرحلة التحرير نفيد من المعالجة الحاسوبية في استخلاص المادة المعجمية الكاملة عبر أنظمة التحرير المعجمي. ونفيد من المعالجة الإحصائية في توصيف مادة المعجم وإحصاء المداخل والوحدات والشواهد. وفي مرحلة النشر تفيدنا المعالجة الحاسوبية في بناء الهيكل الحاسوبي للمعجم، وتهيئة الوسائط المناسبة له كالرسومات والصوتيات وغيرها بما يلبي حاجة الفئة المستهدفة[2].

وهناك نموذج يطلق عليه "قواعد بيانات الدلالات المعجمية العربية عبر الحقول الدلالية" الذي نفذته الشركة الهندسية لتطوير نظم الحاسبات "آر- دي- آي" حيث عكفت على بناء وتطوير هذه الشبكة الدلالية المعجمية العربية بين منتصف عام 2005 حتى بداية 2011. ويشمل هذا المحتوى حوالي مئة ألف مفردة عربية صريحة، تندرج تحت مستويين من الحقول الدلالية؛ مستوى أساسي به نحو ألفين من المعاني

[1] - ينظر: المهوبي عبد العزيز، (2017)، مدخل إلى اللسانيات الحاسوبية، ص51.

[2] - ينظر: المعتز بالله السعيد، (2018)، نحو معجم للغة العربية للناطقين بغيرها-معالجة حاسوبية إحصائية- مجلة التواصل اللساني، جامعة سيدي محمد، فاس المغرب، المجلد19، العددان 1 و2، ص27.

والمفاهيم الكلية، ومستوى ثانوي يتشعب إلى حوالي ثمانية عشر ألفا وأربعمائة حقل دلالي أكثر تفصيلا لتصنيف المعاني الكلية وفق الصيغ الصرفية، ووفق بعض التصنيفات الدلالية، وترتبط هذه الحقول الدلالية بعضها ببعض، وفي محصلة ذلك كله، فإن لدينا في نسيج الشبكة أكثر من مئتين وثلاث وتسعين ألف وصلة دلالية تربط الحقول الدلالية بعضها ببعض[1].

أما المحلل النحوي "فيمكننا من التحقق من فرضيات النظريات اللغوية، والتأكد من فرضيات نحوية قائمة حول لغة معينة، فضلا عن رسم معالم واضحة لنظام الجملة في اللغة المدروسة، وتحديد خصائص علاقاتها التركيبية، مبينة وجوه الائتلاف والاختلاف في بنية الجملة، وهذا يسهل النظر في الغايات التربوية في تعليم اللغة ومعرفة خصائصها، كما يستطيع المعلم والمتعلم على حد سواء معرفة السمات اللغوية لكلمة معينة وسياقها في الجملة واختلاف معانيها اعتمادا على السياق والموضوع"[2].

كما تحاول المعالجة الآلية للغات فك اللبس الدلالي، وقد استخدمت فرضية مفادها أن لكل كلمة عددا معينا من المعاني المختلفة التي يمكن تخزينها في قاموس يضم الكلمات ومعانيها، بعد ذلك يستخدم برنامج حاسوبي للبحث عن المعاني المختلفة لأي كلمة في مخازن الذخيرة لاستعادتها، ثم يقوم بعملية اتخاذ القرار المعنى الأقرب من الصواب في السياق الكلامي[3].

وهناك عدد من الأبحاث التي تطرقت لفك اللّبس الدلالي في اللسانيات الحاسوبية العربية، وقد توصلت إلى أن إعادة الكلمات إلى جذورها يسهم بشكل كبير في رفع دقة البرنامج الآلي، ويسهم بشكل فعّال في تخفيف اللّبس الناجم عن غياب التشكيل في معظم النصوص العربية[4].

---

[1] - ينظر: عطية، محمد، (2019)، مقدمة في حوسبة اللغة العربية، ط1، مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الرياض، ص195-196.

[2] - روبي أحمد ، (2017)، مدخل إلى اللسانيات الحاسوبية، مركز الملك عبد الله عبد العزيز الدولي،ط1، دار وجوه للنشر والتوزيع المملكة العربية السعودية، ص87.

[3] - ينظر: إشراق علي الرفاعي، (2017)، مدخل إلى اللسانيات الحاسوبية، مركز الملك عبد الله عبد العزيز الدولي،ط1، دار وجوه للنشر والتوزيع المملكة العربية السعودية، ص113.

[4] - ينظر: المرجع نفسه، ص115.

وحري بالبيان أن المعالجة الآلية للغة لم تتوقف عند حدود الجملة، بل تعدت إلى معالجة النصوص، عن طريق عمليات تعدّ العمود الفقري لتطبيقات المعالجة الآلية للغات الطبيعية، حيث تفكك الوحدة النصية إلى مكوناتها الجزئية لمعرفة بنياتها الداخلية.

وتمر عملية تجزئة النصوص في اللغة العربية على ثلاثة مستويات[1]:

أ- التجزئة على مستوى الجملة: وذلك بالتعرف على معيار تحديد أبعاد الجملة، ويتم التقسيم حسب الإسناد والتركيب التام المفيد، وما بين الجمل من علاقات الربط بواسطة أدوات الاستئناف والعطف. معتمدة على علامات الترقيم.

ب- التجزئة على مستوى الوحدات/ العناصر الرئيسة: إن الكلمة في عرف الحاسوب هي حيز من الحروف المتشابكة، أو الحروف المفردة أو العلامات أو الرموز، يحيطها من جانبيها مساحات بيضاء.

ج- التجزئة على مستوى الوحدات/ العناصر الفرعية: قد تكون الكلمة مكونة من مورفيم واحد فقط، وقد تتعدى ذلك الى أكثر من مورفيم. فثمة أنواع من المورفيمات اللصيقة في اللغة العربية مثل الجذع واللواصق والزوائد.

- الجذع: هو الجزء الأساسي من الكلمة، يأتي جامدا أو مشتقا.

- اللواصق: هي مورفيمات تتعلق بجذع الكلمة، وهي نوعان: السوابق كنون المضارع، واللواحق كالواو والنون في جمع المذكر السالم.

- الزوائد: هي مورفيمات مقيدة بكلمات أخرى، مثل: حروف العطف والضمائر المتصلة.

ويشكل المستوى التداولي عقبة أمام حوسبة اللغات؛ ذلك لأن التوصيف الحاسوبي يمحو الفاعليات الخطابية المساهمة في بناء الخطاب، وبذلك فإن مقاصد المتكلمين غير معتبرة في التوصيف الحاسوبي، وإذا صرفت الفاعلية الخطابية ومحيت قصدية المتكلمين توجب اعتبار صورة العبارة النحوية مجردة من سياقاتها، وتوجب ضبط معانيها وإحكامها بطريقة تجعل منها غير قابلة للتأويل، ولأجل إتمام هذا المشروع لابد أن تُجرى في النحو

[1] - ينظر: روبي أحمد، (2017)، مدخل إلى اللسانيات الحاسوبية، مركز الملك عبد الله عبد العزيز الدولي،ط1، دار وجوه للنشر والتوزيع المملكة العربية السعودية، ص89- 92.

العربي مجموعة من التحويلات على المعطيات النحوية قبل الشروع في بناء المحتوى، ومن هذه التحويلات:

- إظهار المحذوفات النحوية.
- محو الفعاليات الخطابية.
- استبعاد اللّبس في الكلام[1]

2-3- التحديات التي تواجه حوسبة اللغة العربية:

تستدعي المعالجة الحاسوبية للغة العربية أساليب وأدوات قادرة على التعامل مع أنظمة اللغة العربية البنيوية والتركيبية والدلالية لاستخلاص المباني والمعاني التي تمثل هيكل النظام المعجمي وفقا لما تفرضه طبيعة اللغة. وتزداد صعوبة الأمر حين تلتقي الطبيعة الاشتقاقية للعربية مع نظامها الكتابي الذي قد يختلف شكل الحرف الواحد فيه تبعا لموضعه في الكلمة كحرف العين الذي يختلف شكله في بداية الكلمة عن شكله في وسطها، ويختلف شكله في بداية الكلمة ووسطها عن شكله في نهايتها، ويعبر فيه عن علامات الإعراب برموز خاصة تمثل جميع أحوالها. ويؤدي هذا الالتقاء إلى صعوبة المعالجة الحاسوبية للكلمة الواحدة، إذ يتم رسمها في مجموع كتابي واحد يمكن تقسيمه باستخلاص زوائده من السوابق واللواحق، ويعالج معجميا بتعيين الجذر والجذع والفرع الذي يعبر عن الوحدة المعجمية"[2]

كما تواجه الحوسبة العربية مجموعة من التحديات منها ما يتعلق بالمجتمع العربي، ومنها ما يتعلق بالنظام اللغوي نفسه:

أ- المعيقات الاجتماعية: هناك الكثير من المعيقات الاجتماعية التي حالت دون تحقيق حوسبة لغوية عربية تضاهي ما وصلت إليه حوسبة اللغات الأخرى منها[3]:

---

[1] - ينظر: المالكي طارق، (2015)، أنطلوجيا حاسوبية للنحو العربي، دار النابغة للنشر والتوزيع، جامعة طنطا، ط1، ص66-67.

[2] - المعتز بالله السعيد، نحو معجم للغة العربية للناطقين بغيرها-معالجة حاسوبية إحصائية- ص12.

[3] - ينظر: السعيطي عبد الخالق عبد الرحمان، (2021)، استخدام تكنولوجيا الحوسبة وأثرها في تعليم اللغة العربية في ليبيا، مجلة ابن منظور، العدد الرابع، ص 198.

- ضعف البنية التحتية لشبكات الأنترنيت في الوطن العربي، وضعف المستوى المادي لغالبية المجتمعات العربية، وهذا يمنعهم من الاستفادة من العالم الرقمي.
- غياب دور الجامعات ومراكز البحث والمراكز الإعلامية في حوسبة البحوث.
- ضعف حركة النشر وغياب شبكات التوزيع.
- كتابة اللغة العربية بحروف أجنبية.
- غياب تنمية المحتوى الرقمي العربي، وعدم توفر أدوات معلوماتية أساسية تعتمد على حوسبة اللغة العربية، وتحليلها بشكل علمي دقيق، مثل محركات البحث والمعاجم.
- غياب الموسوعات العربية الموثوقة والمصادر المفتوحة، وهي مواقع متخصصة في كل فرع من العلوم، وفيها معلومات موثقة، وغيابها يزيد الفجوة المعرفية، ويتيح المجال للمعلومات المغلوطة.
- ضعف الإنتاج الفكري العربي، وقلة ترجمته إلى اللغات الأخرى، وقد يرجع ذلك إلى كون النظم والأجهزة ومعظم التطبيقات البرمجية القوية صممت في دول أجنبية، وتقوم بخدمة لغات هذه الدول.
- عدم وجود محرك بحث عربي يتعامل بشكل علمي مع اللغة العربية، باستثناء بعض المحاولات.

ب- المعيقات اللغوية: تتميز اللغة العربية بسمات تميزها عن غيرها من اللغات، مع أن هذه السمات قد تمثل عناصر قوة لها، إلا أنها تمثل في الوقت ذاته عناصر تحد إذا أخضعت للحوسبة، ومن ذلك:

- الكتابة العربية: لا تفهم الكتابة العربية إلا إذا كانت حروفها متصلة، وهناك بعض الحروف التي لا تقبل الاتصال، إذ بينها وبين متجاوراتها مسافة قصيرة نسبيا. والواقع أن فصل الكلمات بعضها عن بعض مسألة ليست بسيطة، ففي أنظمة تجارية لبرامج التعرف على الحروف العربية نجد أن نسبة تقارب 10 بالمئة من الأخطاء على مستوى الكلمات يكون أصل الخطأ فيها خطأ في فصل الكلمات. كما أن كتابة الحروف متصلة تجعل مسألة التعرف على حروف هذه الكلمات مسألة أكثر صعوبة من التعرف على

النصوص المكتوبة باللغات اللاتينية، حيث تكتب جميع حروفها منفصلة وهو الغالب في الطباعة[1].

- غياب التشكيل الذي يمثل أكبر عقبة تواجه تحليل الكلمات العربية حاسوبيا، فهناك صيغ صرفية مختلفة لها هيكل حرفي واحد، مثل "كتب" يمكن أن تدل على الفعل أو الجمع أو غيرها. فالتشكيل يمثل جزءا هاما في الكلمة العربية، فهو بمثابة الحركات القصيرة التي تكتب في اللغات اللاتينية وغيرها. فإذا كان العقل البشري قادرا على استنباط ما لم يكتب من الحركات، فإن الأمر بالنسبة للحاسوب ليس بالسهولة نفسها.

إن هذا الأمر يسهل الكتابة العربية ، فمتوسط عدد حروف الكلمة العربية المكتوبة قد يزيد قليلا عن أربعة حروف، بينما يكون في اللغة الإنجليزية أكثر من ذلك، نظرا لكتابة كل الأصوات المنطوقة باستثناء الكلمات الوظيفية، ولكنه يزيد دائرة الالتباس عند الحاسوب، ويجعل المعالجة الحاسوبية ذات أبعاد أكثر من مثيلاتها في اللغات اللاتينية[2].

- يشكل الرسم الإملائي أيضا حاجزا لهذا البرنامج، فاختلاف كتابة بعض الكلمات مثل: "مسؤول- مسئول، أوروبا-أوربا" يؤثر على دقة عمل الحاسوب الصرفي.

- وجود أكثر من معنى للبنية الصرفية الواحدة.

- تطابق بعض أبنية الكلمات التي تنتهي بالألف والنون أو بالياء والنون مع المثنى والجمع المذكر السالم، إذ من السهل على الحاسوب تمييز أبنية المثنى، لكنه يعجز عن التفريق بينها وبين الكلمات التي تنتهي بالألف والنون" أو الياء والنون"، وهي ليست مثنى، وهذا يتطلب تزويد الحاسوب بقاعدة معطيات تمكنه من معرفة الكلمة بعد تجريدها من الألف والنون.

- دمج الضمائر المتصلة مع الكلمات في اللغة العربية، حيث تتغير صورة الكلمة عند اتصالها بالضمير، مثل: كتب- كتبنا"[3].

---

[1] - ينظر: رشوان محسن، (2019)، مقدمة في حوسبة اللغة العربية، ط1، مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الرياض، ص24.

[2] - ينظر: المرجع السابق، ص25.

[3] - المهوبي عبد العزيز، (2017)، مدخل إلى اللسانيات الحاسوبية، ص 64.

كما خلق ثراء اللغة العربية في جميع مستوياتها مشكلات حاسوبية عدة، "وكان من أثر ذلك أن استعار مطور المحللات الحاسوبية الآلية العربية حلولا من النظريات الخاصة للغة الإنجليزية، حيث لم تسهم تلك الحلول في استيعاب خصائص اللغة العربية وطاقاتها حاسوبيا؛ لأنها في شتى قواعدها أشمل، وأثرى من النموذج الإنجليزي، حيث يقف المحلل الصرفي للغة الإنجليزية –نظرا لخلو اللغة الانجليزية من خاصية الاشتقاق- عند حدود ساق الكلمة، فإذا أردنا أن نستخدمه في اللغة العربية فلن نجد أي ارتباط بين الجذر ومشتقاته"[1].

إن كل هذه المعيقات حالت دون تحقيق الغاية المنشودة من هذا العلم البيني الوليد في دراسة اللغة العربية، ولتجاوز هذه العقبات لابد من فرق متخصصة لغويا وحاسوبيا، وبناء محللات حاسوبية لغوية عربية تتماشى مع التطور التكنولوجي؛ لأن الدراسات اللغوية العربية القديمة غير كافية لبناء محلل لغوي حاسوبي للغة العربية، فالحاسوب لا يتعامل إلا مع الخوارزميات التي تضبط عمليات توليد الكلمات وتحليلها، مما يتطلب رصد دقائق بنية صرف الكلمة العربية، والإحاطة الكاملة بكلماتها.

3-3-آفاق حوسبة اللغة العربية:

ترتبط حوسبة اللغة بما تقدمه التكنولوجيا من تقنيات وأدوات وبرامج تجعل الآلة طيعة ومرنة أمام الاستخدام اللغوي المتشعب، فآفاق حوسبة اللغات كبيرة وواسعة بقدر ما توفره الآلة ويمليه الذكاء الاصطناعي، وتتمثل هذه الآفاق في[2]:

1- تخزين المفردات والتراكيب اللغوية بكيفية تكفل سرعة استرجاعها ومعالجتها وتوظيفها في مستهدفات أخرى كالترجمة الآلية والفورية وغيرها، وبذلك يمكن استخدام تلك المداخل المعجمية في التعريف بالألفاظ وإزالة الغموض عنها، وفي الترجمة الآلية من وإلى العربية، وفي نقل المحتويات اللغوية الأجنبية إلى العربية أو العكس، وهو ما يجعل المعاجم المحوسبة تقترب من الاشتغال الذهني.

---

[1] - ينظر: بيومي سعيد أحمد، (2002)، أم اللغات، دراسة في خصائص اللغة العربية والنهوض بها، ط1، 2002، ص108.

[2] - ينظر: رباحي محمد، (2019)، آفاق حوسبة المعاجم العربية، مجلة اللغة العربية، العدد43، ص134-136.

2- تحسين الترجمة الآلية والترجمة الفورية من اللغة العربية وإليها: وهذا من المستهدفات الكبرى للحوسبة اللغوية التي يمكن القول عنها إنها متأخرة مقارنة باللغات اللاتينية التي تكون الترجمة فيما بينها بنسبة صواب مرتفعة، في حين نجد أن ترجمة النصوص إلى العربية أو العكس، يعطينا مقابلات دون ترابط، وأحيانا معنى غير سليم.

3- الاستفادة من المعاجم المحوسبة واستخدامها في التدقيق النحوي والإملائي والتركيبي، وهذا متاح الآن، لكنه لم يصل إلى الجودة التي تمكن من التخلي ولو نسبيا عن المراجعة اللغوية البشرية.

4- تحويل الكلام المنطوق إلى نص مكتوب أو العكس.

كما نأمل أيضا في توسيع دوائر معالجة اللغة العربية بأدوات حاسوبية يمكن لها أن توفر للباحث اللغوي العربي مزيدا من التحليل، وذلك عن طريق تطوير معالجات التوسيم النحوي، أو التوسيم القواعدي، واستحداث بدايات جادة للتوسيم الدلالي، والإحالي والنبري التي ستفيد الباحث اللغوي بمزيد من وقائع التحليل للغة العربية[1].

خاتمة:

توصل البحث إلى جملة من النتائج، نذكر منها:

- فتحت اللسانيات الحاسوبية بابا واسعا للباحثين في الدراسات اللغوية العربية للولوج من خلالها إلى علم الحاسوب، وتسخيره في خدمة الدرس اللغوي، الأمر الذي أسهم في النهوض بالدراسات اللغوية العربية.

- إن دقة نظام اللغة العربية جعلها من أكثر اللغات قابلية للحوسبة الآلية، فهي توفر مجالات كثيرة لتناولها حاسوبيا على اختلاف مستوياتها: الصوتية والمعجمية والصرفية والنحوية.

- أنشأت الشركات المهتمة ببرامج الحاسوب عدة برامج خاصة بحوسبة اللغة العربية، وقد استطاعت تطوير البرمجيات الحاسوبية التي تربط اللغة العربية بشتى مجالات الحوسبة.

[1] - ينظر: سلطان بن ناصر المجيول، البحث اللغوي في المدونات العربية الحاسوبية بين الممكن والمحتمل والمأمول، مجلة التواصل اللساني، المجلد19، العددان 1و2، ط1، مركز الملك عبد الله بن عبد العزيز الدولي، 2015، ص263.

- تتطلب حوسبة اللغة العربية أدوات تستطيع التعامل مع أنظمتها الصوتية والصرفية والتركيبية والدلالية.

- تواجه المعالجة الآلية للغة العربية مجموعة من التحديات كضعف البنية التحتية لشبكات الأنترنيت في الوطن العربي، وضعف المستوى المادي لغالبية المجتمعات العربية، وغياب شبكات التوزيع، وعدم توفر أدوات معلوماتية أساسية تعتمد على حوسبة اللغة العربية...

قائمة المراجع:

الكتب:

1- إشراق علي الرفاعي، (2017)، مدخل إلى اللسانيات الحاسوبية،،ط1، دار وجوه للنشر والتوزيع، المملكة العربية السعودية.

2 - بيومي سعيد أحمد، (2002)، أم اللغات، دراسة في خصائص اللغة العربية، ط1، 2002.

3- المالكي، طارق، (2015)، أنطلوجيا حاسوبية للنحو العربي، ط1، دار النابغة للنشر والتوزيع، جامعة طنطا.

4- المهوبي عبد العزيز، (2017)، مدخل إلى اللسانيات الحاسوبية، ط1، مركز الملك عبد الله عبد العزيز الدولي ، دار وجوه للنشر والتوزيع المملكة العربية السعودية.

5- الأنصاري عبد الله، (1435هـ)، علاقة النحو العربي بالبرمجة الآلية للغة، جامعة الإمام محمد بن سعود الإسلامية.

6- عطية محمد، مقدمة في حوسبة اللغة العربية، ط1، مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الرياض.

7- العناتي وليد، (2003)، اللسانيات التطبيقية وتعليم اللغة العربية لغير الناطقين بها، ط2، دار الجوهرة للنشر، عمان.

8- الفيقي عبد الله، مدخل إلى اللسانيات الحاسوبية، (2017)، مركز الملك عبد الله عبد العزيز الدولي، ط1، دار وجوه للنشر والتوزيع المملكة العربية السعودية.

9- روبي أحمد ، (2017)، مدخل إلى اللسانيات الحاسوبية، ط1، دار وجوه للنشر والتوزيع المملكة العربية السعودية.

المقالات:

1- المجبول سلطان بن ناصر، (2018)، البحث اللغوي الآلي في المدونة الحاسوبية واللغة العربية، مجلة التواصل اللساني، جامعة سيدي محمد، فاس المغرب، المجلد19، العددان 1 و2.

2- الميساوي خليفة ، (2012)، الذكاء الاصطناعي وحوسبة اللغة العربية: الواقع والآفاق، مجلة مدارات في اللغة والأدب، مركز مدارات للدراسات والأبحاث، تبسة، المجلد01، العدد 05.

3- المعتز بالله السعيد، (2018)، نحو معجم للغة العربية للناطقين بغيرها-معالجة حاسوبية إحصائية- مجلة التواصل اللساني، جامعة سيدي محمد، فاس المغرب، المجلد19، العددان 1 و2.

4- السعيطي عبد الخالق عبد الرحمان، (2021)، استخدام تكنولوجيا الحوسبة وأثرها في تعليم اللغة العربية في ليبيا، مجلة ابن منظور، العدد الرابع.

5- فاهم سعيد، 2015، قراءة في الإسهامات اللسانية الحاسوبية العربية، مجلة دراسات لجامعة الأغواط، العدد 36.

6- رباحي محمد، (2019)، آفاق حوسبة المعاجم العربية، مجلة اللغة العربية، جامعة الدكتور مولاي الطاهر، المجلد 21، العدد 43.