قواعد خوارزمية عربية لرفع الالتباس الدلالي: دراسة لسانية حاسوبية

Algorithmic Rules for Semantic Disambiguation in Arabic: A Computational Linguistic Study

نبيلة بنت محمد الأسمري: تخصص: لغويات، كلية اللغة العربية وآدابها، جامعة أم القرى، المملكة العربية السعودية.

Nabila bent Mohammed Al-Asmari: Linguistics, Faculty of Arabic Language and Literature, Umm Al-Qura University, Saudi Arabia.

Email: nabelh105@gmail.com

DOI: https://doi.org/10.56989/benkj.v5i4.1446

تاريخ الاستلام: 2025-01-27 تاريخ القبول: 2025-02-10 تاريخ النشر 2025-04-01

الملخص:

تطوَّرَ الذكاء البشري، واستغلَّه الإنسان لصُنع الذكاء الاصطناعي، والذكاء الاصطناعي: هو قدرة الآلة على إنتاج بعض الكلمات، أو قدرة الإنسان الآلي أو (الإنسال) –كما يسميه بعض اللُّغويين– على تصحيح النصوص المكتوبة. ويهدف هذا البحث إلى عمل قواعد خوارزمية عربية لرفع الالتباس الدلالي من النصوص أو الجمل المراد ترجمتها من خلال الحاسوب. ومن أجل ذلك اعتمدت الباحثة المنهجي الوصفي والتحليلي ليتم من خلاله وسم الكلمات والقدرة على إنتاج عدد من الوسوم الحاسوبية للإبانة والإفصاح عن المعنى الصحيح بعيدا عن تداخل الكلمات وصعوبة تفسيرها. ولذلك توصي الباحثة بالنظر في موضوع القواعد الخوارزمية الحاسوبية لكونها نقلة كبيرة في تاريخ اللغة العربية في هذا الجانب. توصلت الدراسة إلى مجموعة من النتائج، أبرزها: ما زالت الجهود النظرية هي المرجع الأساسي للعمل الحاسوبي؛ وهناك صعوبة في إمكانية التعرُّف على طرق بناء المدوَّنات، كما أنَّ الأعمال النظرية موجودة، لكنها تفتقر إلى من يترجمها آليًّا.

الكلمات المفتاحية: القواعد الخوارزمية، الذكاء الاصطناعي، الوسم، الفكر اللغوي الجديد.

Abstract:

Human intelligence has evolved, and humans have exploited it to create artificial intelligence. Artificial intelligence is the ability of a machine to produce certain words, or the ability of a robot (or "robot")—as some linguists call it—to correct written texts. This research aims to develop Arabic algorithmic rules to remove semantic ambiguity from texts or sentences to be translated by computer. To this end, the researcher adopted a descriptive and analytical approach to tag words and generate a number of computer tags to clarify and reveal the correct meaning, avoiding the overlap of words and the difficulty of interpreting them. Therefore, the researcher recommends exploring the topic of computer algorithmic rules, as they represent a major shift in the history of the Arabic language in this regard. The study reached a number of conclusions, most notably: theoretical efforts remain the primary reference for computer work; there is difficulty in identifying methods for constructing corpora; and theoretical work exists, but it lacks the ability to translate them automatically.

Keywords: algorithmic rules, artificial intelligence, tagging, new linguistic thought.

المقدمة:

أشار تمام حسان في كتابه {الفكر اللغوي الجديد} إلى أن اللغة عبارة عن اللفظ والمعنى، ولكل منهما نموه وتطوره، فحينا يكون على نحو التراكمي، وحينا يكون على نحو التجدد عبر توالي الأزمنة، وقد كان الانسان مصدرا لها، وكان لابد للغة أن تمر بمراحل التغير إيجابيا أو سلبيا (تمام، ص11).

فقد بثَّ الله الناس في الأرض وجعل بينهم روابط؛ للتَّواصل والتَّعارف فيما بينهم، ونقل الثقافات من مكان إلى آخر، ويعود ذلك الفضل للغة؛ فقد اتَّحدتْ جهودُ العلماء واللغويين من أجل نشر الثقافة اللُّغوية التي تطورت وأصبحت تُدعَى بـ: (اللِّسانيات)، وبعدها استدعى الأمر أن تكون اللِّسانيات حاسُوبِيَّةً؛ لهدف الحفاظ على اللغة، ولكون (الحَوْسَبَة) أصبحت ذات رواج واسع؛ لتكون مسايرةً لمتطلَّبات العصر اللغوي الحديث. ولعل هذا دليل على أن اللغة قابلة للتطور وفق ما تقضيه حاجة اللغوي.

مشكلة البحث:

اختير هذا البحث لبيان حاجة اللغة واللسانيات إلى وضعها في نطاق حاسوبي له القدرة على التصحيح النحوي والصرفي والدلالي حاسوبياً، ولعل تجربة الدكتور محمد الشارخ {رحمه الله} خير برهان على نجاح تجربته.

1- ما أهمية التطور الحاسوبي في صياغة اللغة العربية؟

2- كيف من الممكن تجاوز اللبس الدلالي حاسوبياً؟

3- هل لتجربة الدكتور محمد الشارخ {رحمه الله} أهمية في وضع الأساس النحوي والصرفي الآلي؟

أهمية الدراسة:

تنبع أهمية الدراسة كونها تعزز القدرة على عمل خوارزمية عربية قادرة لتكون مرجع لساني حاسوبي لتفادي تداخل المعنى وتجنب اللبس الدلالي الذي ينتج من خلاله. لذلك وضع العلماء عددًا من العناصر والصيغ المختلفة لتكون عونا في وضع الجملة الصحيحة في السياقات المناسب ورد الخطأ النحوي والصرفي والدلالي فيها؛ لتكون معجمًا آليا في تصنيف اللغة العربية.

منهج الدراسة:

تعتمد الباحثة المنهجين "الوصفي والتحليلي" من خلال وصف الكلمات والنصوص وترميزها ووضع لها وسوم معينة للتمكن من بيان المعنى الصحيح.

أهداف الدراسة:

تهدف هذه الدراسة إلى عمل خوارزمية عربية حاسوبية؛ لها القدرة على تصحيح النصوص من الناحية النحوية والصرفية والتطور الدلالي. وكذلك بيان طرق الاستفادة من تجارب العلماء اللسانيين في حوسبة اللغة العربية.

هيكل الدراسة:

- أولاً: أهمية الخوارزمية العربية لتصحيح النصوص اللغوية.
- ثانياً: الذكاء الاصطناعي وقدرته على إنتاج نظم برمجية لغوية.
- ثالثاً: مفهوم التوسيم بوجه عام.
- رابعاً: خطوات عمل خوارزمية نحوية.
- خامساً: مفهوم اللسانيات الحاسوبية.

أولاً: أهمية الخوارزمية العربية لتصحيح النصوص اللغوية.

«أصبح من المؤكَّد اليوم أن موضوع اللِّسانيات الحديث لم يَعُدْ منحصرًا في الوصف الظاهر للظواهر اللُّغوية، بل أضحت دراسة الأولوِيَّات المعرفية والإدراكية التي تتحكم في إنتاج الفعل اللُّغوي واستعماله ساقيا، مع إمكانية نَمْذَجَة العمليات وحوسبتها في عقل الآلة؛ بمعنى آخر: البحث في القواعد الصُّورية (أي: الخَوارِزْميَّة) التي تقوم بتوليد اللغة، وتحديد فهمها، وتحليل عناصرها» (مهدويوي، ص11).

انتهج اللغويون الشرح والتفسير والتفصيل اللُّغوي، من الجذور حتى استقامة النص نَحويًّا وصرفيًّا ودَلاليًّا ومُعجميًّا، من خلال القواعد النظرية التي قامت عليها صحَّة الدرس اللُّغوي؛ فأصبحت مُنطلَقًا لعمل (القواعد الخَوارِزْميَّة)، ومعالجة اللغات الطبيعية.

«إن التطور المعرفيَّ والتِّقنيَّ الذي شهده العصر الحديث، ساعد على معالجة أنظمة اللُّغات الطبيعية على المستويات الصرفية والنَّحوية والدَّلالية والمُعجمية؛ باستخدام علوم الحاسب وتِقنياته، ولم تقف هذه المعالجة عند حدود الوصف الظاهري للموضوع اللُّغوي الطبيعي، بل تعدَّته إلى التفسير» (مهدويوي، ص20).

نَسْتذكر الحديث عن اللُّغويات الحاسُوبِيَّة: «إن أبسط تعريف للُّغويات الحاسُوبِيَّة: أنها علم حَوْسَبة اللغات الطبيعية بنوعيها المنطوق والمكتوب» (اللغة العربية والتنمية البشرية، ج2، ص922).

ولا يمكن أن تتم هذه الدراسة إلا ببناء برامج حاسُوبِيَّة لأنظمة اللُّغات البشرية، من خلال تقييس ومحاكاة نظام عمل الدماغ البشري لنظام عمل الحاسب الآلي(مهدويوي، ص28).

لابد من وضع قواعد بيانية آليَّة، تعتمد على القواعد النظرية المُعدَّة سابقًا، والتي تعتمد على بنك المعلومات.

ومن الضروري أن تُوسَم هذه القواعد بوُسُومٍ حاسُوبِيَّةٍ، وتُرَمَّز كذلك برموز يسهل على (الإنسال) الآلة فهمها وترجمتها.

عند بناء خَوارِزْميَّة، من المهم استكمال استخدام الذخيرة اللُّغوية التي تعتمد على بنك المعلومات، وبناء مُدوَّنَة خَوارِزْميَّة ذات قواعد للمُشْتَرَك اللَّفظي، وأيضًا جمع جميع النصوص والكلمات ذات المُشْتَرَك اللَّفظي، وحصر معاني اللفظة الواحدة.

ثانياً: الذكاء الاصطناعي وقدرته على إنتاج نظم برمجية لغوية

يحتاج هذا العمل إلى اتحاد جهود العلماء واللُّغويين والمبرمجين؛ من أجل القدرة على إنتاج عمل حاسوبيٍّ قائمٍ على محاولة حلِّ مشكلة اللغة نحويا وصرفيا ودلالياً.

مفهوم الذكاء الاصطناعي:

الذكاء: عملية معقَّدة، قادرة على تمكين الكائن الحي من التكيُّف المتجدِّد الذي يناظر فيه، والعمل على الوسائل والغايات (إسماعيل، ص45).

اصطناعي: كلمة ترتبط بالفعل: (يَصْطَنِع)؛ وبالتالي تُطلق الكلمة على الأشياء التي تنشأ نتيجة النَّشاط، أو الفعل الذي يتم من خلاله اصطناعُ وتشكيلُ الأشياء؛ تمييزًا عن الموجودة بالفعل، والموجودة بصورة طبيعية دون تدخل الإنسان (إسماعيل، ص45).

قدرة الحاسوب على الكلام والتحدث:

تزويد الحاسب على فهم الكلام البشري عن طريق تلقي الأصوات من الخارج وإعادة تجميعها والتعرف عليها ومن ثم الرد عليها، ويبحث هذا المجال في إيجاد طرق تسمح للحاسب بفهم التعليمات المعطاة إليه بصورة طبيعية، بمعنى يستطيع فهم لغة الإنسان بسهولة (إسماعيل، ص54)، لأن لها القدرة على تلاقي الأوامر من الإنسان لتطبيقه حاسوبيا.

كذلك يُعَدُّ الذكاء الاصطناعي علمًا من علوم الحاسب، وقد صنعه الإنسان من أجل حلِّ المشكلات والصعوبات التي تواجه اللغة.

ويشير مارتن ويك Martin Weik: أن الذكاء الاصطناعي هو قدرة الآلة على القيام بالمهام التي تحتاج للذكاء البشري عند أدائها مثل الاستنتاج المنطقي والتعلم والقدرة على التحليل والابتكار (إسماعيل، ص47).

من خلال التعريفات السابقة يتَّضح أن الذكاء الاصطناعي يعتمد على الذهن الإنساني في القدرة على ابتكار خوارزمية لُغوية، يستطيع استنطاق العقل الآلي، من خلال قدرته على التفكير؛ بوضع برامج حاسُوبِيَّة، وأَوْسِمَة، ورموز لها القدرة على إنتاج عمل حاسوبيٍّ خاصٍّ بالآلة نفسها ليتمكن من اشتقاق الكلمات وتصويبها نحويا والحرص على سلامة الكلمة تركيباً.

ثالثاً: مفهوم التوسيم بوجه عام

هو نوع الكلمة الذي يُسنَد إليها يدويًّا أو آليًّا، ويكون ذلك عن طريق لوحة المفاتيح الأَلِفْبائية والرقمية (التميمي، ص96).

بحسب الغرض يتم وَسْم المُدَوَّنة، وهناك أنظمة تُبنَى فقط لتعيين: (الاسم، والفعل، والحرف)، وهناك مُدَوَّنة تُبنَى لحاجات أدق وأعمق، تستلزم تحديد خصائص لُغوية لمعرفة الزمن الصرفي: (مضارع، أمر، ماض)؛ من أجل ذلك تختلف التقطيعات وَفْق كل مَطلَب؛ لأن الوَسْم الصرفي له نوع معيَّن من التقطيع، والوَسْم النَّحْوي له نوع آخر كذلك، وتكون عمليات التقطيع فيها مبنيَّة على ما يحتاجه العمل (التميمي، ص159).

يشكِّل التَّوسِيم الدَّلالي مهمةً أكثر تحدِّيًا من التَّوسِيم النَّحْوي والتَّوسِيم التَّركيبي؛ لأنه يعتمد على المعرفة الإدراكية، إذ يتطلب موارد معجمية وأنطولوجية؛ كالمعاجم، والمكانز (McEnery & Wilson, 2011, p.16).

وفي دراسة حديثة تتعلق بالعربية، قُدِّم منهجٌ معتمدٌ على القواعد، يقف على أنواع متعددة من أفعال الكلام: (استفهام، تعجب، نفي) وتم اختباره على مُدَوَّنة من (1500) جملة، وحقق دقة بلغت (98، 92٪) (التميمي، ص48).

رابعاً: خطوات عمل خوارزمية نحوية

إنَّ عمل خوارزمية نحوية يعتمد على الخطوات التالية:

- تحديد الغرض الذي من أجله سيُبنى النظام الآلي للتَّوسيم النَّحْوي.

- تصميم بناء المُدَوَّنة وَفْق الأُسُس المعيارية، وحسب مناهجها، وبما يخدُم نظام التَّوسِيم الدَّلالي.
- تحديد معايير المعالجة القبلية التي تُعرف بـ(مرحلة التفريق)؛ حيث يتم فيها تفريق الكلمات عن الأرقام والرموز، وعلامات الترقيم والمسافات التي تفصل بين الكلمات في بعض اللغات.
- اختيار وتطبيق المنهج الذي سيعتمد عليه بناء نظام التَّوسِيم الآلي، والمنهج المعتمد على تعليم الآلة، وفيها يتم تدرُّب الخَوارِزْميَّة لتعليم الآلة جزءًا كبيرًا من المُدَوَّنة الموسومة يدويًّا.
- قياس أداء نظام التَّوسِيم؛ سواء أكان نحويًّا، أم دلاليًّا... إلخ المستويات (التميمي، ص69-70).

إن لتجربة الدكتور محمد الشارخ أهمية في وضع الأساس الأول في ترجمة المصحف الكريم والحديث الشريف، وسلامة اللغة العربية من الناحية النحوية والصرفية، وفتح المجال إلى البحث عن التطور الدلالي، ولقد اهتم العلماء واللغويون العرب منهم والأعاجم بحوسبة اللغة.

خامساً: مفهوم اللسانيات الحاسوبية

اللسانيات:

تعرف اللسانيات linguistics وتسمى الألسنية، وعلم اللغة بأنها الدراسة العلمية للغة (يونس، ص9).

الحاسوبية: يقصد بها توظيف الحاسوب في خدمة اللغة العربية من حيث إصدار معاجم وكتيبات حاسوبية لها القدرة على ترجمة الكلمات من اللغة الأم إلى اللغة العربية.

اللغة العربية وحوسبتها أمر في غاية الصعوبة كما أنه يحتاج إلى متخصصين في لغة البرمجة وهذا المجال أصبح معقدا جدا لقلة مرتادي الحاسوب.

إن الحاجة إلى الحوسبة أدت إلى ظهور اللسانيات الحاسوبية التي تحاول محاكاة العقل البشري في فهم الظاهرة اللغوية تنظيرًا وإنجازًا، لذلك جمع هذا الحقل من المعرفة بين اللسانيات والذكاء الاصطناعي والإعلامية والرياضيات والمنطق؛ بهدف تحليل النظام اللغوي تحليلا آليا متعدد المستويات وبأسرع وقت ممكن (الميساوي، ص28-31).

التفكير الحاسوبي (CT): هو أسلوبٌ لحل المسائل التي تستخدم تقنيات علوم الحاسوب، وتم استخدام مصطلح التفكير الحاسوبيّ لأول مرة من قِبَل (سيمور بابيرت عام 1996م)، ويمكن استخدام التفكير الحاسوبي لإيجاد حلٍّ خَوارِزْمِيٍّ للمسائل المعقدة، وغالبًا ما يُستخدم لتحقيق تحسينات كبيرة في الكفاءة (إسماعيل، ص53).

أصبح من المؤكد اليوم أن موضوع اللسانيات الحديث لم يعد منحصرا في الوصف الظاهري للظاهرة اللغوية، بل أصبح له القدرة على نمذجة العمليات العقلية وحوسبتها في عقل الآلة بمعنى البحث في القواعد الصورية {الخوارزمية} التي تقوم بتوليد اللغة وتحديد فهمها وتحليل عناصرها (مهديوي ص11).

اتجهت المحاولات العربية والأجنبية التي تندرج في إطار تعريب الحاسوب في ثلاث اتجاهات:

1- إنجاز الوظائف التي تتعلق بإدخال المعطيات والمعلومات باللغة العربية وتشفيرها وتخزينها وإظهارها على الشاشة والطابعة.

2- تعريب التطبيقات الحاسوبية عن طريق برمجة تعتمد اللغة العربية أساسا لتشغيلها بالحرف العربي.

3- يهدف إلى تعريب نظم تشغيل الحواسب حسب أحجامها من الصغير إلى المتوسط والكبير (مهديوي، ص33-34).

التعبيرات الاصطلاحية ومشكلة فهمها حاسوبيا:

تعد التعبيرات الاصطلاحية ظاهرة لغوية عالمية، فما من لغة تخلو منها لكنها تتأثر وتختلف غالبا باختلاف البيئة واختلاف حضارات الأمم. وتمثل هذه التعبيرات مشكلة كبيرة عند محاولة فهمها وخاصة في حالة حوسبة النص والترجمة الآلية (مهديوي، ص89).

تشير المصطلحات في الغالب إلى عدم تمييزها ومعرفتها في كثير من الأحيان، يعني تكون مصطلحات دخيلة ليس لها معنى ثابت في اللغة العربية أو تكون لهجة أو كلمات أعجمية معربة يصعب على اللغوي فهمها وترجمتها آليا.

ولعل ذلك يفرض على الباحثين عمل معاجم متخصصة لهذه التعبيرات؛ وذلك ليتضح دورها في اللغة وشرحها وشرح طرق معالجتها اللغوية (مهديوي، ص90).

الخاتمة:

يتبين من خلال هذه البحث أن عمل خوارزمية عربية له أهمية كبيرة من خلال إدراج اللغة العربية في مجال الحاسوب وفق السياقات والتراكيب النحوية والصرفية والدلالية والاستعانة بالوسم الحاسوبية لبناء دالة أو عدد من الدوال قادرة على تصحيح النصوص اللغوية.

من أهم النتائج ما يلي:

1) الجهود النظرية ما زالت هي المرجع الأساسي للعمل الحاسوبي.
2) صعوبة حوسبة اللغة أو استعمال الحاسوب وفق معطياته: تُعدُّ صعوبةً للمتدرب، وتحتاج إلى الصبر والمهارة، والبحث والتَّعلُّم.
3) بناء مدوَّنات حاسُوبيَّة يحتاج إلى خزينة لُغوية، أو ما يُسمَّى: (بنك المعلومات الحاسوبي).
4) صعوبة إمكانية التعرُّف على طرق بناء المدوَّنات.

من اهم التوصيات:

1) توجيه الجهود إلى اختيار مجموعات للعمل الحاسوبي، يرتكز عملُهم على التدريس.
2) الأعمال النظرية موجودة، لكنها تفتقر إلى من يترجمها آليًّا.

قائمة المصادر والمراجع:

أولاً: باللغة العربية:

1- إسماعيل، عبد الرؤوف محمد (2017): تكنولوجيا الذكاء الاصطناعي وتطبيقاته في التَّعليم، ط1، عالم الكتاب.

2- التميمي، أفراح عبد العزيز، نظام آلي للتقطيع والتَّوسِيم النحوي العربي، ط1، كنوز المعرفة.

3- حسان، تمام، الفكر اللغوي الجديد، ط1، الكويت: عالم الكتب.

4- علي، محمد يونس، مدخل إلى اللسانيات الحديث، دار الكتاب الجديد المتحدة.

5- اللغة العربية والهندسة اللغوية (2011): مدخل ضمن كتاب اللغة العربية والتنمية البشرية الواقع والرهانات، الجزء2، منشور سرهو.

6- مهديوي، عمر (2019): اللسانيات الحاسوبية واللغة العربية-إشكال وحلول، ط1، مؤسسة كنوز المعرفة.

7- آل محرق، شمعة حسن ظيف الله؛ آل محرق؛ حمودة حسن ظيف الله؛ الدربي، صالحة حسن محمد؛ عسيري، سهام ضيف الله علي (2025): الاعتبارات الأخلاقية في دمج الذكاء الاصطناعي في رؤية المملكة العربية السعودية 2030 للتعليم، مجلة ابن خلدون للدراسات والأبحاث، مجلد 5، عدد 3، ص. 64-83.

8- الهذلي، بشائر نوار؛ السّعيديّ، الحسن العَّياشي (2022): المصادر الرقمية ودورها في تعليم اللغة العربية غير الناطقين بها أثناء الحجر الصحي نموذج معهد تعليم اللغة العربية لغير الناطقين بها جامعة أم القرى، مجلة ابن خلدون للدراسات والأبحاث، مجلد 2، عدد 10.

ثانيًا: باللغة الإنجليزية:

McEnery, T. Wilson, A. (2011): **Corpus Linguistics (An Introduction)**. University press, Edinburgh, UK, P16.