

الأخطاء اللغوية والإملائية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ:

دراسة لغوية وصفية تحليلية

د. إبراهيم محمد عبده سعد طلحة

أستاذ الدراسات اللغوية المساعد بكلية اللغات والترجمة
جامعة تعز – الجمهورية اليمنية
الإيميل: e738144443@gmail.com

الملخص:

هذا بحث عن: (الأخطاء اللغوية والإملائية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ)، هدف فيه الباحث إلى التعرف إلى أهم الأخطاء اللغوية والإملائية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ، واقفا على عدد من النماذج. وقد ارتضى الباحث المنهج الوصفي التحليلي إلى جوار المنهجين التاريخي والمقارن وغيرهما حال اللزوم، وتضمن البحث تمهيدا ومبحثين، أحدهما بعنوان: الأخطاء اللغوية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ، والآخر بعنوان: الأخطاء الإملائية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ، إضافة إلى خاتمة البحث ونتائجه، وقائمة بأهم مصادره ومراجعته.

الكلمات المفتاحية: اللغة – الذكاء الاصطناعيّ – المحتوى العربيّ – الأخطاء اللغوية – الأخطاء الإملائية – اللغة والنحو والقواعد.



المقدمة:

هذه دراسة بحثية عن: (الأخطاء اللغوية والإملائية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ)، نابعة من مشكلة مفادها تساؤل علمي يقول: ما أهم الأخطاء اللغوية والإملائية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ؟

وتتلخص أهم أهداف هذا البحث في الآتي:

- 1- التعرف إلى أهم المشكلات اللغوية والنحوية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ.
 - 2- التعرف إلى أهم المشكلات الكتابية والإملائية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ.
 - 3- إيجاد حلول فكرية ولغوية للمشكلات الفنية والتقنية المتعلقة باللغة والإملاء العربيين في وسائل التواصل الاجتماعي ووسائط التقنية الحديثة، وتزويد التقنيين العرب بالمادة النظرية المعينة لهم على تطوير المضمون التقني وتحسين جودة المحتوى العربي في الشبّكة.
 - 4- رصد المكتبة العربية الحديثة وإثرائها بدراسة نظرية تطبيقية في آن.
 - 5- الإسهام في معالجة اللغات الطبيعية من وجهة نظر عربية تفاعليّة.
- واعتمد الباحث في دراسته المنهج اللغويّ الوصفيّ التحليليّ، مع تداخله أحيانا مع المنهجين التاريخي والمقارن، وذلك في ضوء علم اللغة الاجتماعيّ.

وقد اقتضت طبيعة هذا البحث أن تتوزع مادّته على مقدمة هي هذه، وتمهيد يتحدث عن اللغات والذكاء الاصطناعي، ومبحثين أحدهما بعنوان: الأخطاء اللغوية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ، والآخر بعنوان: الأخطاء الإملائية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ، إضافة إلى خاتمة البحث ونتائجه، وقائمة بأهم مصادره ومراجعته، وهي مصادر ومراجع عربية وأجنبية.



تمهيد:

اللغة الإنسانية والذكاء الاصطناعي - مدخل تاريخي عام -

يعدُّ تاريخ العلاقة الإنسانية - الآلية تاريخاً عتيداً وليس وليد اليوم، فقد شهد التاريخ أنواعاً من التآلف (أو لنقل التناؤس) الإنساني - الآلي، ومحاولات شتى للاختراع والابتكار على مرِّ الزمن، وما هذا الانتقال النوعي من عصر الثورة الصناعيّة إلى عصر الثورة الرقمية إلا نموذج لزخم العلاقة التاريخية بين الإنسان والآلة.

"ويحفل تاريخ مختلف الأمم عتيقه وقريبه مثلما يحفل التراث العريض لأدب الخيال العلمي بالطرائف والنفائس من أمثلة تلك الآلات التي حاولت محاكاة البشر في ملمح أو أكثر، وتشير كلها دُونَ لبس إلى قَدَمٍ وتواصل الوَلَعِ بحلم صناعة "أشباه البشر"، وكثيراً ما يستخدم العامّة اصطلاح "الخدم الآليّ robot" بدلاً من "شبيه البشر humanoid" وخصوصاً في سياق التحكُّم والتسخير"^(١).

وقد مرَّ الروبوت (المصنوع) في إزاء الإنسان (المخلوق) بأطوار ومراحل قبل إيجاد هذه الصورة النهائية - إلى الآن - عن المنافس الآلي للإنسان، ولعلَّ أهمَّ مرحلة هي مرحلة الحوسبة التي بدأت في أربعينيات القرن الماضي، ثم تطورت إلى شكل الذكاء الاصطناعي الموضوع اليوم في منافسة مع الذكاء البشري، "بعد أن أظهرت أحدثُ مكتشفاتٍ تشريح المنظومة العصبية في الحيوانات وفي البشر أن المخ يتكوّن من شبكة كثيفة الترابط من الخلايا العصبية، دفعت هذه المكتشفات بعض الباحثين مثل (مارفين مينسكي) - أحد الآباء المؤسسين للذكاء الاصطناعي - إلى محاولة بناء شبكة عصبية اصطناعية كهربائية عسى أن تظهر بعض قدرات المخ الطبيعي، واستطاعت هذه الشبكة العصبية بعد عمل مطوّل في تطويرها إظهار بعض التصرفات التي تكافئ العمليات المنطقية الثنائية البسيطة"^(٢).

والملاحظ أنّ درجة العلاقة بين الإنسان والآلة في زمننا أصبحت كبيرة، وأسرع وتيرة من أيّ فترة.

وسنجد أنّ الحاسوب هو الذي ساعد الذكاء الاصطناعي على المنافسة فطالما ساعد هذا الجهاز الإنسان على تحديد مدخلاته لتكون مخرجاه في ضوئها، فيمكنه من استكشاف بنية اللغة ومعرفة تشكيلاتها، وتمييز مواطن الجمال فيها،

١) العربية والذكاء الاصطناعي، المعز بالله السعيد ومحمد عطية وآخرون، إصدارات مجمع الملك سلمان للغة العربية، ومركز الملك عبدالله للتخطيط والسياسات اللغوية، سلسلة مباحث

لغوية (٥٩)، ١٤٤٥ هـ. ص: ٢١.

٢) المرجع السابق نفسه، ص: ٣١.

من خلال الوقوف على مجموعة متنوعة من النصوص بمستوياتها البلاغية المتعددة والمتفاوتة في الوقت نفسه. وكما أصبح الحاسوب كتابًا معرفيًا جامعًا، أصبح اليوم يدعم عقلا اصطناعيا قارئًا^(١).

ولا شك في أنّ اللُّغة هي موضع التفرد في الإنسان، ويأتي دورها امتدادا لدور العقل البشري الذي تميز به، ومن هنا كانت محل اختبار قدرات الآلة على محاكاتها، ويمثل عصر الذكاء الاصطناعي نزوة التحدي بين الإنسان والروبوت وبين العقل الطبيعي والعقل الاصطناعي، وسنجد أنّ معظم اللغات الطبيعية الآن تسعى إلى تفعيل منظوماتها المعجمية لمجاراة المعالجات الآلية الحديثة.

والعربية واحدة من أهمّ لغات العالم وأقدمها تاريخا وحضارة، وتعدّ من اللُّغات الست الرئيسة في الأمم المتحدة، وكان لزاما أن تجاري تطوّرات التقانة، وتواكب جديد المجال المعرفي والمعلوماتي. وقد بدا واضحا أنّ مواكبة هذا المجال تستند إلى ضرورة تعاون اللغويين مع التقنيين لردم الفجوة الكائنة بين المستخدم العربي والفني العربي، أو غير العربي حتّى. فاليوم نحن في عالم يزداد فيه الاتصال بين الناس، وتنمو وسائله، حتى لم يعد في طوق أمة من الأمم، ولا فرد من الأفراد، أن يلزم بيته ويعبد ربه. فمصير كل أمة مرتبط بمصير غيرها من الأمم التي تصادقها أو تعادىها، ويرتبط كذلك مصير كل فرد في المجتمع بمصير بقية أفراد المجتمع الذي يعيش فيه، ويتزايد ارتباط مصيره هذا بمصير الإنسانية في عمومها في الوقت الحاضر، والذي ساعد على ذلك هو التطور في وسائل الاتصال المادي والفكري بين أطراف المعمورة^(٢). وإنّ أدوات اللُّغة التعبيرية والفنية ينبغي أن تستجيب لمقتضيات التنوع، وكذلك لخصوصيات الوسيلة الإعلامية وأدوات التكنولوجيا، ويجب أن تواكب التطورات والمستجدات قدر الإمكان^(٣).

وإنّ التكنولوجيا الرقمية في الوطن العربي في حاجة ماسّة إلى مراجعة شاملة ورؤية فكرية ذات مرجعية أكاديمية وعلمية وتعليمية وثقافية تأسيسية؛ ولا سيما مع التطورات التي طالت مناحي الحياة، ومنها ما نعائشه ونزامنه اليوم من أشكال الإعلام الجديد ومناهج الاتصال، في هذا العصر المتفق على تسميته بعصر التدفق الحرّ للمعلومات.

(١) ينظر: Using computers in Linguistics A partical guide: John M. Lawler and Helen Aristar Dry, Routledge, London

and New York, 1998 : 174

(٢) اللغة بين المعيارية والوصفية، تمام حسان، عالم الكتب. القاهرة، ط ٤، ٢٠٠٠ م، ص ٦١.

(٣) ينظر: الارتقاء بالعربية في وسائل الإعلام، نور الدين بلبليل، سلسلة كتاب الأمة الصادر عن وزارة الأوقاف والشؤون الإسلامية بقطر، العدد ٨٤، رجب ١٤٢٢هـ، ص ٥٩ وما بعدها.

المبحث الأول:

الأخطاء اللغوية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ: توطئة:

عُني العرب بسلامة لغتهم، وأنشؤوا لها المجامع اللغوية التي بدأت بتأسيس (المجمع اللغوي للوضع والتعريب) في القاهرة، عام 1892م، ثم (المجمع العلمي العربي) بدمشق عام 1919م، وتلتها مجامع عدة في الوطن العربي، في بغداد وعمّان وطرابلس والدوحة وغيرها، وقد وكل إلى هذه المجامع النظر في اللغة العربية وأوضاعها العصرية، ونشر آدابها وإحياء مخطوطاتها، وتعريب ما ينقصها من كتب العلوم والصناعات والفنون من اللغات الأوربية، وتأليف ما تحتاج إليه من الكتب المختلفة المواضيع على نمط جديد. وعني أيضاً بجمع الآثار القديمة، وجمع الكتب المخطوطة والمطبوعة⁽¹⁾. ومن ثم أوكل إليها ضمناً النظر في معايير اللحن اللغوي، ومعالجة الأخطاء التي تتفاوت درجاتها من حيث الحدّة، وتختلف مستوياتها من حيث الفداحة، كما كان للمجامع مواقف أخرى من الاستعمالات المختلفة، بتسجيل ما شاع من ألفاظ الحضارة وجعلها من مظانّها، ثم تهذيبها وإقرارها؛ وما لا سبيل إلى إقراره يدعه للزمن والاستعمال؛ كي يصلح من شأنه، ويقوم من عوجه⁽²⁾.

وأهم الأخطاء اللغوية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ هي:

١- أخطاء المخارج الصوتية:

تتنوع الأخطاء الصوتية في المحتوى الإبداعيّ العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ، وأبرزها أخطاء المخارج الصوتية، ولعل على رأسها: أخطاء نطق الظاء والذال زايًا، وأخطاء نطق الثاء سينا، وأخطاء نطق الضاد ظاء، وهذه الأخطاء وإن كانت نتيجة لكُنات المتكلمين ومما له علاقة بلهجاتهم المحلية إلا أنّها تعكس اختلالاً في منظومة معالجة اللغة العربية بالذكاء الاصطناعيّ، مما يعني أنه يمكن اليوم ملاحظة أنّ اللغة المنطوقة في الوسائط الرقمية وبرامج الذكاء الاصطناعيّ هي مقابل اللغة المكتوبة في مناهج التعليم والتعلم، ومن ثمّ فإنّ "كلام الشبكة تطوّر ذو أهمية؛ إذ إنه وسيط جديد للتواصل اللغوي لا يتكرر حدوثه كثيراً في تاريخ الجنس البشري"⁽³⁾. فينبغي معالجته،

(1) ينظر: تاريخ المجمع العلمي العربي، أحمد الفتيح، (مطبوعات المجمع العلمي العربي بدمشق) - مطبعة الترقى - 1956م، ص 11.

(2) ينظر: مجمع اللغة العربية في ثلاثين عاماً (ماضيه وحاضر)، إبراهيم مذكور، الهيئة العامة لشؤون المطابع الأميرية - القاهرة - 1964م، ص 59 - 60.

(3) اللغة والإنترنت، ديفيد كريستال، ترجمة/ أحمد شفيق الخطيب، المشروع القومي للترجمة، المجلس الأعلى للثقافة. القاهرة، ع 982، ط 1، 2005م، ص 294 وما بعدها.

ولا سيما وأن لكل قوم لغتهم ولكل لغة أصواتها الخاصة، بل إن اللغة أساساً: "أصواتٌ يُعبرُ بها كلُّ قومٍ عن أغراضهم"^(١).

ولاحظ الباحث أن مخارج الحروف وصفاتها في المحتوى العربي المعالج ببرامج الذكاء الاصطناعي وتطبيقاته غيرها في وصف علماء الأصوات، فعلى سبيل المثال لدينا: (الذال) من الصَّوامتِ الأسنانية - كما تسمى - إلى جوار الثاء والظاء، ومخرجه عند علماء الأصوات من طرف اللسان وأطراف الثنايا العليا^(٢)، ولكنه في موضوعات داخل المحتوى العربي المعالج ببرامج الذكاء الاصطناعي ينطق كالزاي، أي من طرف اللسان وفوق الثنايا السفلى. ومثل هذا الخطأ يوجد خطأ آخر يتمثل في إطالة المدى الزمني للصوائت، فالكسرة القصيرة، قد تتحول إلى كسرة طويلة عندهم أي إلى (ي)، في نحو قولهم: أنتي = أنتِ، ولكي = لكِ.

كما أن ثمة أغلاطا تتصل بالحركات والإعراب من نحو تسكين الآية الكريمة من قوله تعالى: ﴿بِلِسَانٍ عَرَبِيٍّ مُبِينٍ﴾^(٣)، فينطقها عدد من المتكلمين في برامج معالجة الذكاء الاصطناعي لدى الوصل بالتسكين، وهذا خطأ، فيقال: بلسانٍ عربيٍّ مبينٌ، مثلاً.

٢- أخطاء البنية الصرفية:

وقد تكون هذه الأخطاء نتيجة أخطاء المستوى الأول، أي الأخطاء الصوتية، أو العكس، فمثلاً في نطق كلمة القدر: القدر، في عبارة: الإيمان بالقضاء والقدر، هذا الخطأ في نطق الكلمة هو ناتج عن الخطأ في فهم البنية الصرفية لها وسياقها العام، فكان أن أدى إلى خطأ في نطق الكلمة.

وفي برامج ترجمة المحتوى على تطبيقات التواصل الاجتماعي ما يشبه ذلك، فقد نستمع إلى مقطع صوتي لأحد المتحدثين يقول فيه مثلاً: الصديق هو من تجده وقت الضيق، فتترجمها برامج الآلة كتابياً: الصديق، ولو لم ينطقها المتحدث سينا، فكان التطبيقات تنقل الأخطاء المنطوقة إلى عوالم الكتابة، والعكس أحياناً أخرى. وأما على صعيد أخطاء الخلط بين أشكال حروف المعاني فالأمر أكثر تعقيداً، كالخلط بين بنية كل من: أنا وإنا، أو: نحنُ ونحنو، أو إذن، وإذا، ونحوها. ونجد أحياناً خلطاً ما بين أوزان الصيغ الصرفية كالأفعال والمصادر، من نحو: أكلَ وأكُل، أو: شربَ وشُرِب، أو: أخذَ وأُخذ، أو: تباينَ وتباين، أو: تناقلَ وتناقل، فيقال مثلاً: إنَّ أكلَ التفاح يومياً مفيد، وإنَّ شربَ الماء يومياً

(١) الخصائص، لأبي الفتح عثمان بن جني (ت ٣٩٢هـ)، تحقيق: محمد علي النجار، الهيئة المصرية العامة للكتاب، ط٤، ١٩٩٩م، ج ١، ص: ٣٣.

(٢) ينظر: مناهج البحث في اللغة، تمام حسان، مكتبة الأنجلو المصرية، ١٩٩٠م، ص: ١٢٤ وما بعدها.

(٣) سورة الشعراء، الآية: ١٩٥.

ضروري للجسم، وإنَّ أخذَ خاطرك من أخيك أمر محزن، وإنَّ تبايَنَ الألوان أمر ملفت، وإنَّ تناقَلَ الحديث من غير تثبت عادة سيئة!

٣- أخطاء التراكيب النحوية:

وقد يكون من نافلة القول إنَّ هذه الأخطاء هي أخطاء سياقية ناتجة عن اختلال المستويين الصوتي والصرفي، فمن الطبيعي - إذن - أن نصادف أخطاء من قبيل تكثير المؤنث وتأنيث المذكر، كما في محتوى مرئي (ريلز) على (الفيسبوك) نُشر وتم الإطِّلاع عليه في ٢٨ / ١١ / ٢٠٢٥، تحت عنوان موضوع العلاقات الزوجية السليمة والمتوازنة - صفحة (لمحة معرفة) - تقول فيه المتحدثة: "المرأة لا تستطيع إخفاء مشاعرها فيظهر عليها تلميحات وإشارات كالنظرات الطويلة (إليك) وأن تمدح (رجولتك) واحمرار الوجه والنظرات الخجولة!" كذلك من هذه الأخطاء ما يتداخل فيه المستوى النحوي مع المستوى الكتابي، كما في تحدي (My IQ)، إذ يعطي الفرد فرصة لاختبار متوسط ذكائه ونقاطا مقابل ذلك، ولكن واجهته تحمل خطأ واضحا في قول: "درب عقلك يوميا وشاهد (ذكائك) ينمو"، والصواب أن يقال هنا: وشاهد (ذكاءك) ينمو، بفتح الهمزة، لا بكسرها.

٤- أخطاء الدلالات المعنوية:

من المحتمل والمتوقع بالتأكيد أن تؤدي الاستعمالات اللغوية الخاطئة على المستويات الصوتية والصرفية والنحوية إلى دلالات ومعانٍ لغوية فاسدة، فالدلالة والمعنى يقومان على النحو والتركيب، مضافين إلى المستويين الصوتي والصرفي بطبيعة الحال، وهذه التراتبية هي التي تقوم فكرة المعيرة في الصواب اللغوي عليها، إضافة إلى الاعتماد على المرجعية التراثية المتمثلة في كل من: القرآن الكريم، والشعر العربي واللغات العربية القديمة، والحديث النبوي، للاستشهاد بلفظٍ ما، أو الاستشهاد على أسلوب ما. ولم يخرج المحدثون عن القدامى في هذا الشأن، وخاصة لاعتبارهم أن اللغة العربية لغة حية، "تحتاج دائما إلى قليل من التهذيب؛ لمسيرة العصر الذي تعيش فيه"^(١).

(١) معجم الأخطاء الشائعة، محمد العدناني، مكتبة لبنان - بيروت، ط ٢، ١٩٨٣ م، ص: ٧.

وقد تقدم التنويه إلى عدد من الأخطاء والأغلاط الصوتية والصرفية والنحوية التي ستؤدي بدورها إلى أخطاء وأغلاط دلالية، ويمكن أن يضاف إليها أخطاء التأويل المعنوي غير المناسبة لسياقاتها، ففي بعض البرامج يقول المتكلم: إنَّ (فينيسوس) لاعب (ريال مدريد) سيحوّل مركزه، وكان مقام المقال يقتضي أن يقول: سيحوّل مركزه! ومثل هذا نجده لدى التعامل مع (Chatgpt) أحيانا في نحو إدخال جملة: (ما عندك) مثلا، من دُون علامة الاستفهام، سنجده يوزع العبارات بدلالاته هو، فيمكن أن يعدّه سؤالاً أحيانا، من غير التعامل على وفق منظومة التقييم المعروفة، وهكذا يقال في كثير من برامج الذكاء الاصطناعي وتطبيقاته التي تضع الإجابات على عواهنها. وما يهمّ في هذا السياق هو الاستعمال اللغوي على وفق المستوى الدلالي المراعي لِ (المعنى الأسلوبي)، الذي تقع في نطاقه مجموعة من الكلمات والتعبيرات، حيث إن كثيرا من الكلمات قد تحمل من المعنى ما يدلُّ على مستوى الكلام وأسلوبه وواسطته. ثمَّ لنضع في الاعتبار: هل تستخدم الكلمة في الكتابة أم المحادثة أم الخطابة؟ فمثلاً: (أيها الناس) ليست للمحادثة بل للخطابة. وهل الكلمة ذات مستوى رسمي أم عامي أم مبتذل؟، فهناك كلمات لا يجوز استخدامها في موقف رسمي، وهل الكلمة ذات استخدام نثري أم شعري؟ فمثلاً: (يا صاح)، تستخدم في الشّعر وليس في النثر. وإنّ تظهر مستويات الاستعمال هذه وظلال المعنى الأسلوبي في الكلمات مثل: أمي / ماما، والدي / دادي، أبي / بابا، كريمته / ابنته / بنته (في العامية)، عقيلته / حرمه / زوجته / امرأته / مرته (في العامية)، ونحوها، ونعرف أنّ لكل مقام استعمالٍ مقاله^(١).

المبحث الثاني:

الأخطاء الإملائية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ: توطئة:

شغل مجمع القاهرة منذ إنشائه بموضوع تيسير الكتابة العربية، وتقدم عدد من أعضائه بمقترحات حول هذا الموضوع، في يناير عام ١٩٤٤م (مشروع استبدال الحروف اللاتينية بالحروف العربية)، وفي فبراير عام ١٩٤٤م (مشروع تيسير الكتابة)، ومقترحات من ضمنها: إبقاء صور الحروف العربية على حالها، ومحاولة تصوير الكلمة كما تنطق، بإضافة زوائد تكتب في بنية الكلمة للدلالة على الحركات. وفي مؤتمر يناير عام ١٩٥١م، قُدِّم مقترح لتسهيل شكل الكتابة المطبعية، بالاقصصار على صورة واحدة للحروف، هي التي تقبل الاتصال من أول الكلمة، وذلك في الحروف التي تتصل بغيرها، وأقرت لجنة تيسير الكتابة عام ١٩٥٩م عدداً من القواعد لضبط الكلمات بالشكل في الكتب المدرسية. كما اتفق على أن تقوم طريقة تيسير الكتابة على أساسين:

(١) ينظر: علم الدلالة (علم المعنى)، محمد علي الخولي، دار الفلاح للنشر. عمان، ط ١، ٢٠٠٢م، ص ٧٧.

١- اختصار صور الحروف، بتمثيل الحرف بصورة واحدة، أيًا كان موقعه من الكلمة.

٢- الاحتفاظ بطبيعة الخط العربي، وتجنب المبادعة بين القديم والجديد.

كما شكل مجمع القاهرة لجنة للإملاء، حاولت معالجة الصعوبات التي يواجهها المبتدئ في تعلم الإملاء، في تقرير عرض في عام ١٩٤٨م على مؤتمر المجمع، وافق فيه المجتمعون على قواعد عدة منها رسم الهمزة، وفصل الأعداد من (ثلاث) إلى (تسع) عند إضافتها إلى مئة، وغير ذلك. وكانت هذه مقدمة لمعالجات الإملاء العربي الحديث؛ كونه موضوعاً مُشكلاً، إلى جانب موضوعات اللغة والنحو، فالكتابة من المهارات اللغوية الأربع^(١).

وأهم الأخطاء الإملائية الشائعة لدى صنّاع المحتوى الإبداعي العربي المعالج باستخدام تقنيات الذكاء الاصطناعي:

١- الهمزة:

ومن هذه الأخطاء الخط بين همزتي الوصل والقطع أو عدم الاعتداد بقاعدتهما، ومن ذلك مثلاً: ما لو أدخلت اسم (إبراهيم) بدون همز (إبراهيم) أو بهمزة فوق الألف (أبراهيم)، سنجد أن الذكاء الاصطناعي قد يمرره ولا ينبه إلى الخطأ في كتابته، ومثل ذلك ما سبق وضرربنا له مثلاً عند حديثنا عن تداخل الأخطاء النحوية مع الأخطاء الكتابية في جملة: شاهد (نكائك) ينمو، والصواب: شاهد نكائك؛ لأن (نكاء مفعول به) من حيث القاعدة النحوية، ولأن همزته متوسطة تخضع لقاعدة أقوى الحركات، من حيث قاعدة الإملاء والكتابة.

٢ - الألف اللينة:

ومن أمثلة الخطأ فيها: كتابة (Chatgpt) لخواطر أدبية فيها كلمات من قبيل: دعى والحلى، والصواب: دعا، والحلا؛ لأن أصلها الواو، من: يدعو، ويحلو، ولك قد يمرر عليه هذا التقييد، حال الانهماك معه في خاطرة أدبية أو مساجلة شعرية.

(١) ينظر: مجمع اللغة العربية في ثلاثين عامًا (مجموعة القرارات العلمية) - مطبعة الكيلاني - ط ٢ - القاهرة ١٩٧١م، قرارات محاضر جلسات - الدورة ١٠ - ص ٢٧٥ - ٣١٢،

ومحاضر جلسات - الدورة ١٩ - ص ٨٠، ٣٦٠، ومحاضر الجلسات - الدورة ١٩ - ص ٨٠، ٣٦٠.

ومن المفارقات العجيبة أن هذا الشكل من الخطأ قد ورد في دليل اللكء الاصطناعي التوليدي، مخصص لمعالجات لغوية، وذلك في كلمة (تلى) - والصواب: (تلا) من (يتلو) - في سياق جملة تعريفية تقول: (تلى ذلك مجموعة من الملتقيات)^(١).

وكذلك الاستفهام بإلامٍ وعلامٍ وحتامٍ ونحوها، قد يخلط بين جمل من نحو:

إلامٍ الخلف بينكم إلاما؟ وهذي الضجة الكبرى علاما؟!^(٢)

وأخرى تقول: أتدعوهم إلى ما كنت تنهاهم عنه؟!

فيكتب: إلامٍ وعلامٍ في القافية؛ لعدم تفرقه بين القافية المطلقة والأخرى المقيدة، أو يكتب جملة: إلى ما كنت تنهاهم عنه، بطريقة تشبه كتابتها استفهاما، أي: إلامٍ كنت تنهاهم عنه!

٣- التاء المربوطة والهاء:

ومن أمثلة هذا ما لو ورد سياق فيه: ردت عليه عليّة قائلة: من أنت؟ أو: ردت عليه عليه قائلة: من أنت؟ فربما سيُشكل عليه سياق التقديم والتأخير، هل الكلمة الأولى عليّة والثانية عليه أم الأولى عليه والثانية عليّة، وخاصة مع عدم تضمين الأساليب البلاغية فيه، فكثيرا ما يتعامل اللكء الاصطناعي مع المدخلات التي تنتظم وسائل التواصل وشبكات الإنترنت كخوارزميات، لكنه يغمغم في إجاباته التي تتعلق بالأسلوب لا بالشكل، فما يزال موضوع الأسلوب والبلاغة موضع تفاوت وفرق بين قدرات العقل البشري وقدرات اللكء الاصطناعي.

ولأن ضمير الغائب المتصل كثيرا ما يتداخل من حيث شكله مع التاء المربوطة، فإن هذه التاء المعروفة عند بعض اللغويين والنحاة العرب القدامى بالهاء؛ تجوزا واتساعا في الاستعمال، ما تزال موطن تيه اللكء الاصطناعي يصعب عليه معالجتها دون محمولات السياق بل وأحيانا مع وجودها. ومن أمثلة ما يشكل عليه هنا أيضا: (وفيه) و(وفية)، (أنه) و(أنه)، و(وجهه) و(وجهة)، ونحوها.

٤- الضاد والظاء:

وقد يرجع الخطأ فيهما إلى السياق أيضا، فضلا عن الخطأ في معرفة المخرجين الصوتيين للحرفين، فلو قيل: حضر أصدقاء صفحته، أو: حضر أصدقاء صفحته، سيُشكل على البرنامج معرفة الشكل السليم للجملة ما لم يوجد محمول سياقي لها؛ إذ قد تكون مثلا: حضر أصدقاء صفحته، أو: حضر أصدقاء صفحته، فالأولى من الحضور، والحاضرون هم الفاعل أي أصدقاؤه، والثانية من الحظر، والمحظورون هم أصدقاؤه، فهم المفعول به وهو الفاعل.

وإلى ذلك قد يُقدّم محتوى يتضمن حرفي الضاد والظاء بشكله المعياري المعروف من القواعد الإملائية ولكن غلطا، فيستند إلى تععيد عام دون أعمال النظر وإمعانه في المحمول السياقي أيضا، أو بتسرع في الحكم، كما في مثال:

(١) ينظر: اللكء الاصطناعي التوليدي، الهيئة السعودية للبيانات والذكاء الاصطناعي، سلسلة اللكء الاصطناعي التوليدي (١) نوفمبر، ٢٠٢٣م، ص: ٢٦.

(٢) الشوقيات، أحمد شوقي، دار المعارف المصرية، ط١، ١٩٢٤، ج ٢ / ص ٧١.

أنت محظوظ، فأول ما يتبادر أنه تعبير من الحظ، أي إنك صاحب حظ، لكن لن يكون هذا الحكم صحيحاً فيما لو نُظِرَ إلى سياق قبليّ يقال فيه مثلاً: قد حُضَّ الجميع على فعل الخير في هذا الموسم المبارك، حتى أنا محضوض وأنت محضوض على فعله، فإنني محظوظ وأنت محظوظ!!

٥- الحروف الزائدة:

وأمثلتها كثيرة، فمن الوارد أن يُقرأ في برامج الذكاء الاصطناعي اسم (عمرو) بنطق الواو، ولا سيما مع نطقه هكذا في لهجات محلية، ولكن هذا خطأ لدى النظر إليه من زاوية معيار الصواب اللغوي؛ إذ إنَّ (واو عمرو) وضعت للتفريق بين اسم (عمرو) واسم (عُمر). فهذه الواو من الحروف التي تُكتب ولا تُنطق، وهي تشبه في ذلك مثلاً حرف الـ K في اللغة الإنجليزية في كلمة نحو: Know بمعنى أعرف أو يعرف، فإنه يُكتب ولا يُنطق.

كذلك قد يُشكل الأمر حين ورود فعلٍ وُجِدَ فيه ألف بعد واو فلا يُدرى أي ألف فارقة لحقت واو جماعة، أم ألف اثنتين في مثني؟! نحو جملة: دعوا الله، فهي إمّا أن تكون: دَعُوا الله، أو: دَعَا الله، وكلتاها صحيحة ولكن في سياقها.

٦- الحروف المحذوفة:

ويستلزم من معرفة الحروف المحذوفة في اللغة العربية معرفة قواعد الكتابة القياسية والتفريق بينها وبين الكتابات غير القياسية، فرسم المصحف الشريف أو ما يعرف بالخط العثماني هو خط توقيفي عليه، والكتابة العروضية مثلاً هي كتابة غير قياسية، وعليه فإننا لو كتبنا بيت الفرزدق الذي يقول فيه:

هذا الذي تعرف البطحاء وطأته والبيت يعرفه والحل والحرم^(١)

سنلاحظ أننا نكتب (هذا) - وهو اسم إشارة - بدون ألف مدّ، أي إننا ننطق ألف المد ولا نكتبها في نحو: هذا وهذه وهؤلاء، ولكننا في الكتابة العروضية سنكتب:

هاذا الذي تعرفُ بطحا أوطأتهو

والبيتيُّ رفهُو ولجل لوالحرمو

أي يُحذف ما لا ينطق، ويُثبت ما يُنطق صوتياً.

(١) شرح ديوان الحماسة، أبو علي أحمد بن محمد بن الحسن المرزوقي الأصفهاني (ت ٤٢١ هـ)، تحقيق: غريد الشيخ، وضع فهرسه العامة: إبراهيم شمس الدين، دار الكتب العلمية،

بيروت - لبنان، ط ١، ١٤٢٤ هـ، ٢٠٠٣ م، ص: ١١٣٤.

٧-علامات الترقيم:

ومثال ذلك أن تضع لبرنامج الذكاء الاصطناعي مقترحا لكتابة موضوع إنشاء وتعبير وليكن عن الصداقة مثلا، فيعطي لك جُملا مع علامات ترقيمها، منها الجملة الآتية: "يجب علينا أن نحافظ على أصدقائنا، وأن نخترهم بعناية، لأن الصديق الصالح نعمة ثمينة".

ونلاحظ أنه وضع فاصلة قبل جملة: لأنَّ الصديق الصالح نعمة ثمينة، التي تحمل معنى التعليل، وهو ترقيم خاطئ؛ إذ كان حقها أن توضع قبلها فاصلة منقوطة.

وعلى الرغم من أن ثمة خلافات في علامات الترقيم، إلا أن من المفترض أن تكون برامج الذكاء الاصطناعي معدة على وفق ما لا خلاف فيه، أو ما لا يسعه الخلاف، وهذه النقطة هي نقطة اختلاف إلى الآن بينها وبين الإنسان والقدرات التي يمتلكها في مجال تقنين اللغة وتسخير أدوات التكنولوجيا لصالحه.

وهناك ما نسويه افتقادا للمرجعية المشتركة بين المرسل والمتلقي هنا، أو بين الطرفين (الإنسان والآلة)، فقد تتحول علامات الترقيم إلى علامات مائية مجردة من المعنى، أو على وفق الخوارزمية الافتراضية الأولية فقط، وهذا مخل بأركان التواصل المعهودة في العمليات اللغوية سواء التي تتّم من طرف الآلة أو التي تتّم من طرف الإنسان؛ فإنّ توافر أركان أساسية في عملية التواصل بين الطرفين يعدّ أمراً ضرورياً، وأهم تلك الأركان، هي:

١. التعبيرية أو الانفعالية.
٢. الإفهامية أو الذاتية.
٣. المقاصدية أو الاتصالية.
٤. المرجعية أو التعالق.
٥. الميّا لسانية أو ما بعد اللسانية.
٦. الشعرية أو الجمال^(١).

ولكنها هنا غير موجودة في حالة خوارزميات المدخلات والمخرجات اللغوية المفترض تجانسها وانسجامها بين الإنسان والآلة.

(١) ينظر: مدخل إلى الألسنية، بول فابر، المركز الثقافي العربي، الدار البيضاء، (د.ط)، ١٩٩٢ م : ص ٥٣ وما بعدها.

الخاتمة والنتائج والتوصيات:

درس هذا البحثُ الأخطاء اللغوية والإملائية الشائعة لدى صنّاع المحتوى العربيّ المعالج باستخدام تقنيات الذكاء الاصطناعيّ، وتناول حسن اختيار اللفظ والعبارة، ومراعاة الكلمة الصحيحة التي يستطيع المتلقي فهم مقاصدها، ودرجات الالتزام بقواعد اللّغة وحدودها في إطار الذكاء الاصطناعيّ، ودرجات التمكن لدى مستخدميه من قواعد اللّغة، والقدرة على النطق السليم للغة العربية، والتعامل الصحيح مع ألفاظ العربية وحروفها بطريقة صحيحة، في عصر المعرفة والمعلوماتية والمنافسة الإبداعية، لتتطور العربية سواء على مستوى اللغة المنطوقة، أو اللّغة المكتوبة من خلال تضيق المسافة بين لغة الخطاب ولغة الكتابة، ليكون للغتنا الحضور اللائق بها بين لغات العالم.

وقد توصل الباحث إلى نتائج عدّة وخرج بتوصياتٍ أخرى عدّة في دراسته، ويمكن إجمال ما توصل إليه على نحو ما هو آت:

أولاً: أهم النتائج:

- 1- أن المحتوى العربي المعالج بالذكاء الاصطناعي ما زال يعاني من المشكلات نفسها التي عانى منها المحتوى المستقى من مواقع الإنترنت.
- 2- أن عددا من النصوص المعالجة بالذكاء الاصطناعي وخاصة المعالجة صوتياً تحتاج إلى عمل معايير ومقاييس إضافية لضبط الجودة.
- 3- أن طابع المحتوى اللغوي العربي المعالج بالذكاء الاصطناعي شكلي، فإنه مركّز على الشكليات الخارجية للنص، كعلامات الترقيم، والعناوين الفرعية، دون المضمون، الذي ما زال مضطرباً.
- 4- أن اللّغة العربية في المحتوى المعالج بالذكاء الاصطناعي تراوح محلها بين مدخلات الإنترنت التي لا تخلو من الأخطاء الشائعة ومخرجات تطبيقات الذكاء الاصطناعي التي تضع التشكيل والعناوين الفرعية وباب قل ولا تقل، في رأس معالجاتها، في حين أن المحتوى موضوع المعالجة هو من حيث الفكرة فلا تجديد ولا إضافة إلى الفكرة الإنسانية، إلّا في حدود ما يدخله الإنسان.

٥- أن مراعاة السياق دائما هي ضرورة ملحة في التعامل مع المحتوى العربي؛ فإن لكل مقام مقالا، والجمل العربية - وحتى المفردات العربية - لا تُفهم إلا من سياقاتها.

٦- أن كتابة المذكرات ونحوها بالذكاء الاصطناعي تفتقد إلى القدرة على تختيار رابطة يبدأ بها جملة جديدة، أو رابطة يُستطرد بها للتعبير عن فكرة موافقة أو مخالفة، أو رابطة يُنقل منها إلى استلها مترح أو استنتاج رأي، أو تعليل وجهة نظر، إلى غير ذلك.

ثانيا: أهم التوصيات:

- ١- عقد مؤتمرات علمية لمعالجة المحتوى العربي الآلي، وليس لمجرد وصفه.
- ٢- تنفيذ مشروع قومي لضبط جودة المحتوى اللغوي المعالج آليا ولا سيما من الناحية الصوتية.
- ٣- اختراع تطبيقات لغوية عربية غير مستوردة من الخارج؛ لأن المحتوى العربي الحالي معالج في ضوء لغة النموذج المستورد.
- ٤- تأصيل أعمال المدونين العرب الفكرية ذات الطابع الذاتي وفصلها عن تلك المولدة من خلال تقنيات الذكاء الاصطناعي؛ حفظا للحقوق، ولدرء أشكال السرقات الفكرية الآلية.

مصادر البحث ومراجعته:

أولاً: القرآن الكريم.

ثانياً: مصادر البحث ومراجعته بالعربية:

- الآلة قوة وسلطة: التكنولوجيا والإنسان منذ القرن ١٧ حتى الحاضر، آر. إيه بوكانان، ترجمة/ شوقي جلال، عالم المعرفة . الكويت، ع ٢٥٩، ٢٠٠٠م.
- الارتقاء بالعربية في وسائل الإعلام، نور الدين بليبل، سلسلة كتاب الأمة الصادر عن وزارة الأوقاف والشؤون الإسلامية بقطر، العدد ٨٤، رجب ١٤٢٢هـ.
- تاريخ المجمع العلمي العربي، أحمد الفتيح، (مطبوعات المجمع العلمي العربي بدمشق) - مطبعة الترقى - ١٩٥٦م.
- الخصائص، أبو الفتح عثمان بن جني (ت ٣٩٢ هـ)، تحقيق: محمد علي النجار، الهيئة المصرية العامة للكتاب، ط ٤، ١٩٩٩م.
- "ذكاء الآلة"، صلاح الفضلي، عالم الفكر. الكويت، ع ١، مج ٣٥، ٢٠٠٦م.
- الذكاء الاصطناعي التوليدي، الهيئة السعودية للبيانات والذكاء الاصطناعي، سلسلة الذكاء الاصطناعي التوليدي (١) نوفمبر، ٢٠٢٣م.
- الذكاء الاصطناعي، واقعه ومستقبله، الآن بونيه، ترجمة/ علي صبري فرغلي، عالم المعرفة. الكويت، ع ١٧٢، ١٩٩٣م.
- شرح ديوان الحماسة، أبو علي أحمد بن محمد بن الحسن المرزوقي الأصفهاني (ت ٤٢١ هـ)، تحقيق: غريد الشيخ، وضع فهرسه العامة: إبراهيم شمس الدين، دار الكتب العلمية، بيروت - لبنان، ط ١، ١٤٢٤ هـ، ٢٠٠٣م.
- الشوقيات، أحمد شوقي، دار المعارف المصرية، ط ١، ١٩٢٤م.
- العربية والذكاء الاصطناعي، المعتر بالله السعيد ومحمد عطية وآخرون، إصدارات مجمع الملك سلمان للغات العربية، ومركز الملك عبد الله للتخطيط والسياسات اللغوية، سلسلة مباحث لغوية (٥٩)، ١٤٤٥هـ.
- علم الدلالة (علم المعنى)، محمد علي الخولي، دار الفلاح للنشر. عمان، ط ١، ٢٠٠٢م.
- قراءات في علم اللغة، أحمد شفيق الخطيب، دار النشر للجامعات. القاهرة، ط ١، ٢٠٠٦م.
- اللغة بين المعيارية والوصفية، تمام حسان، عالم الكتب. القاهرة، ط ٤، ٢٠٠٠م.
- اللغة العربية: أسئلة التطوير الذاتي والمستقبل، مجموعة من الباحثين، سلسلة كتب المستقبل العربي (٤٦)، مركز دراسات الوحدة العربية. بيروت، ط ١، ٢٠٠٥م.
- اللغة العربية معناها ومبناها، تمام حسان، عالم الكتب. القاهرة، ط ٤، ٢٠٠٤م.

- اللغة والإنترنت، ديفيد كريستال، ترجمة/ أحمد شفيق الخطيب، المشروع القومي للترجمة، المجلس الأعلى للثقافة . القاهرة، ٩٨٢ع، ط١، ٢٠٠٥ م.
- المجامع العربية وقضايا اللغة من النشأة إلى أواخر القرن العشرين: وفاء كامل فايد، عالم الكتب . القاهرة، ٢٠٠٤م.
- مجمع اللغة العربية في ثلاثين عامًا (ماضيه وحاضره)، إبراهيم مذكور، الهيئة العامة لشؤون المطابع الأميرية - القاهرة - ١٩٦٤م.
- مجمع اللغة العربية في ثلاثين عامًا (مجموعة القرارات العلمية) - مطبعة الكيلاني - ط ٢ - القاهرة ١٩٧١م.
- مدخل إلى الألسنية، بول فابر، المركز الثقافي العربي، الدار البيضاء، (د. ط)، ١٩٩٢م.
- معجم الأخطاء الشائعة، محمد العدناني، مكتبة لبنان . بيروت، ط ٢، ١٩٨٣ م.
- مناهج البحث في اللغة، تمام حسان، مكتبة الأنجلو المصرية، ١٩٩٠م.

ثالثًا: مصادر البحث ومراجعته بالإنجليزية:

- Artificial Intelligence: A Modern Approach (AIMA), Stuart Russel, Peter Norvig, press, 2021.
- The Art of editing the news. Robert Mc Giffert-Ed chilton. London 1972
- Linguistics : An Introduction, Andrew Radford and others, Cambridge, University press, - 1999
- The structure and function of communication. In society. Hard D. Lasswell-
- Translation in a post-colonial context; Early Irish Literature in English. Tymoczko,M. - Manchester 1999
- Using computers in Linguistics: A partical guide: John M. Lawler and Helen Aristar Dry, - .Routledge, London and New York, 1998

رابعًا: مصادر البحث ومراجعته على الإنترنت:

- . موقع التواصل الاجتماعي (إكس).
- . موقع التواصل الاجتماعي (فيس بوك).
- . موقع التواصل الاجتماعي (يوتيوب).
- . ويكيبيديا، الموسوعة الحرة على الإنترنت.