

Treating the Arabic text in the Word program and its problems

Author^{1*}: Habbachi Latifa & Boulebda Zineb

¹ Badji Mokhtar-Annaba University- Algeria (Department of Arabic Language and Literature)

* email: habbachilat@yahoo.fr

ABSTRACT:

Abstract: This article aims to monitor the developments of the Arabic text when processing it with the Word program in the 2007 and 2019 versions, and to count and classify the linguistic problems that each version faced separately, and then compare them.

There has been a discrepancy in the ability of the two versions to process the Arabic language system. The 2019 version has been able to successfully address most of the inputs that the 2007 version has had a problem dealing with, and developments that have reached all aspects of the program's conjugational, grammatical and semantic treatment have been evident

Keywords: Automated Processing ; text processing ; Word software.

معالجة النص العربي في برنامج وورد ومشكلاتها

^{1*} د/ لطيفة هباشي، ط/د. زينب بولبدة

⁻ جامعة باجي مختار - عنابة (قسم اللغة العربية وآدابها)

ملخص: يهدف هذا المقال إلى رصد التطورات الحاصلة على النص العربي عند معالجته ببرنامج الورد (word) نسختي 2007 و 2019، وإحصاء المشكلات اللغوية التي واجهتها كل نسخة على حدة وتصنيفها، ثم المقارنة بينهما.

وبعد إجراء الدراسة الإحصائية تبين وجود تفاوت في مدى قدرتهما على معالجة نظام اللغة العربية، وتمكن نسخة 2019 من معالجة جلّ المدخلات التي واجهت نسخة 2007 مشكلة في التعامل معها، وبرزت بجلاء التطورات التي مست كل جوانب المعالجة الصرفية والنحوية والدلالية للبرنامج مع تسجيل عدد من النقائص التي تتطلب تطويرا.

كلمات مفتاحية: المعالجة الآلية، المعالجة النصية، برنامج word .

1. مقدمة:

يمثل السلوك اللغوي أهم تحليلات الذكاء البشري المستهدفة بالبحث من لدن علماء الذكاء الاصطناعي، الذي اشتغل علماءه على الاستفادة من التخصصات والعلوم التي تسعى لكشف أسرارهِ وتفسيرهِ، وذلك من أجل وضع نمذجة محاكيه له في المجال التقني.

وقد تجسد هدفه عندما أقيمت صروح التضافر بين اللسانيات وعلم الحاسوب وهنا بزغ فجر علم جديد وهو اللسانيات الحاسوبية، والتي تعتبر المعالجة الآلية للغات الطبيعية موضوعها الأساس.

وباعتبار النص نسيجاً متكاملًا تتجسد فيه كل المستويات اللغوية للنظام اللغوي، فقد وُضعت برامج إلكترونية تختص بمعالجة النصوص، تهدف إلى تحليل النصوص المدخلة على صعيد كل المستويات اللغوية وفهم خصائصها، وذلك بما زودت به من معالجات ومدققات، حيث يتولى كل معالج معالجة النظام اللغوي الفرعي الذي خصص له.

* جامعة باجي مختار - عنابة / الجزائر (قسم اللغة العربية وآدابها)، البريد الإلكتروني: habbachilat@yahoo.fr

ولأنّ كل اللغات واللغة العربية -بصفة خاصة- تتغير مقاصدها بتغير السياقات فإنّ هذا يعكس طبيعتها المعقدة، التي تتطلب قدرات عقلية عالية للتعامل معها، ما رفع من درجة تعقّد عمل برامج المعالجة النصية، ذلك أنّها برامج كامنة في آلة صماء، يصعب ارتقاؤها إلى مستوى تدرك به عبقرية اللغة ومستعملها.

وهذا ما يجعلها تواجه العديد من المشكلات اللغوية، التي سنحاول كشفها في هذه الدراسة البحثية المعنونة بـ: (معالجة النصّ العربي في برنامج وورد)، وذلك بإدخالنا نصا باللغة العربية إلى برنامج **Word** نسختي 2007 و 2019، باعتباره أشهر برامج المعالجة النصية وأكثرها استخداما، ساعين وراء ذلك كله إلى الإجابة عن إشكالتين رئيسيتين هما:

- ماهي أنواع المشكلات اللغوية التي يواجهها برنامج المعالجة النصية **Word** بنسخته عند معالجته النص العربي؟ وفيما تتمثل أسبابها؟

- ما مدى قدرة نسختي (2007) و (2019) على معالجة النصوص العربية؟ وأيهما الأنجع في ذلك؟

وللإجابة عن هاتين الإشكالتين، سنقوم باختيار أنموذج نصي من القرآن الكريم تمثّل في سورة البقرة، تتم كتابته على نسختي (**Word** 2007 و 2019)، والوقوف على المشكلات اللغوية التي تواجهها كل منهما وإحصائها وتصنيفها، ومن ثمّ المقارنة بينهما لنحدد النسخة الأنجع في التعامل مع النظام العربي، أي مدى قدرة برنامج **Word** على معالجة النصّ العربي قبل التطوير وبعده.

وسنعرض في هذا المقال العناصر التالية:

-أهداف الدراسة وإشكالياتها.

-أنموذج الدراسة ومتطلباتها.

منهجية الدراسة.

مصطلحات الدراسة.

-نتائج الدراسة.

-المقارنة.

2. مكونات المقال:

1.2 أهداف الدراسة وإشكالياتها:

تهدف هذه الدراسة إلى:

الكشف عن المشكلات اللغوية التي يواجهها برنامج المعالجة النصية **word** عند معالجته للنماذج اللغوية العربية قبل التطوير وبعده، وتحديد أنواعها والوقوف على أسبابها الحقيقية.

الوقوف على مدى جدية المساعي العاملة على تطوير برنامج **word** للتعامل مع اللغة العربية.

حل مشاكل واقعية يواجهها المستخدم العربي في حياته اليومية عند إدخاله نصوص لغته إلى برنامج **word**.

ربط اللغة العربية بالجال التكنولوجي لتكون لغة من اللغات العصرية التي تستفيد من خدمات الحاسوب بنجاح كاللغات العالمية الأخرى.

2.2 أنموذج الدراسة ومتطلباتها:

-برنامج المعالجة النصية Word (2019/2007)

هو برنامج مخصص لكتابة النصوص، حيث يتيح إدخالها بصيغة إلكترونية، ومن ثم معالجتها وإخراجها بالشكل المطلوب، مع إمكانية طباعتها أو إبقائها على صيغتها الإلكترونية، وهو من أهم البرامج التي طرحتها مايكروسوفت في حزمة برامج، والتي تعمل على نظام التشغيل Windows XP، وإضافة إلى عملية كتابة وطباعة مختلف الوثائق والنصوص، يسمح البرنامج بالعديد من العمليات الأخرى التي تُجرى عليه، كـ "تغيير وتحريك وتخزين تلك الوثائق، وليس هذا فحسب بل تستخدم حزمة معالجة النصوص على مستوى نظام النصوص المدخلة ذاتها، إذ تقوم هذه الحزمة بتدقيق الإملاء والقواعد" (الصباغ، 2002، ص: 100)

ولا يوجد قدر محدد لكمية المدخلات التي يعالجها البرنامج ويحفظها، فهو يتمتع بنطاق واسع جداً في التعامل معها، وتخضع لتعديلات على مستواه بحسب تعليمات المستخدم، ففي المرحلة الأولى يتم إدخال النص من قبل المستخدم (الزبون) ليظهر على شاشة البرنامج، فيحفظها في ذاكرته الخارجية إلى حين إخضاعها للمعالجة، ليتم بعدها حفظها في ذاكرته بصورة دائمة.

وبصورة موجزة يمكننا القول إنّه برنامج من برامج المعالجة النصية، يسمح بكتابة النصوص وتخزينها ومعالجتها، اعتماداً على عدد من البرامج الأخرى للتحليل الصوتي والصرفي والنحوي والتدقيق اللغوي، يقوم بتوظيفها في معالجة النصوص اللغوية المدخلة والتدقيق في حروفها وكلماتها وتراكيبها، ومحاولة تقديم اقتراحات لتصحيح أكبر قدر ممكن من الأخطاء اللغوية الواردة فيها، وهو برنامج متعدد الخدمات إذ يفسح المجال للمستخدم بإدخال النصوص وحفظها واسترجاعها وتعديلها.

3.2 مدونة الدراسة:

يتمثل النص المدخل في سورة البقرة، برسمها العثماني المثبت في المصحف الشريف، وهي السورة الثانية من حيث الترتيب، "مدنية بلا خلاف، يبلغ عدد آياتها مائتان وثمانون وسبع آيات، تشتمل على ألف خبر وألف أمر وألف نهي، قال فيها النبي عليه الصلاة والسلام: لكل شيء سنام، وسنام القرآن البقرة" (ابن كثير، 2000، ص: 122)

أرقى أنموذج لغوي، فهي كلام الله الذي لا خطأ فيه ولا زلل، لنكون بهذا على يقين من صحة النص المدخل، وأن سبب اعتبار البرنامج لأي عنصر لغوي بأنه خاطئ يرجع في الحقيقة إلى خلل وقصور فيه.

من أطول النصوص العربية وهذا ما يجعل من برنامج Word بنسخته يتعامل مع أكبر قدر من المفردات والأبنية والأوزان الصرفية والقوالب النحوية المكونة للنظام اللغوي العربي.

4.2 منهجية الدراسة:

ستكون هذه الدراسة تطبيقية تعتمد فيها على المنهجين الوصفي والمقارن المدعّمين بآليات الإحصاء والتحليل والاستنباط، حيث سنقوم بكتابة سورة البقرة في البرنامج في كل من النسختين 2007 و 2019، ثم نستخرج جميع العناصر

اللغوية التي أشار البرنامج إلى أنها خاطئة عن طريق التسطير تحتها باللونين الأحمر أو الأخضر واقترح تصحيحات لها، لنحللها ونصنفها حسب المستويات اللغوية، ثم سنعتمد إلى إحصائها للوقوف على النسبة الأكبر من الأخطاء التي وقع فيها، بالإضافة إلى أننا سنناقش نماذج منها لمعرفة الأسباب التي أدت به إلى اعتبارها خاطئة بوضع فرضيات تبرر حدوث الخطأ، ومن ثمة البرهنة على صحتها أو خطئها، عن طريق ممارسة بعض التقنيات البحثية ككتابة مختلف اشتقاقات الكلمة الخاطئة من منظور البرنامج، من أجل تحديد طبيعة الخطأ (إملائي - صرفي - نحوي - دلالي) و استنتاج السبب المحتمل له، كما سنقيم مقارنة بين النسختين القديمة 2007 و الحديثة 2019 لتحديد النسخة التي كانت الأكثر قدرة على معالجة النظام اللغوي العربي بنجاح، لنخلص في النهاية إلى نتائج سنقوم بتحليلها ومحاولة تفسيرها.

3. مصطلحات الدراسة:

1.3 المعالجة الآلية للغات الطبيعية (مفهومها ومراحلها) :

هي : " عملية وضع أنظمة وبرامج معلوماتية تحلل وتولد النصوص المكتوبة آلياً، بما زودت به من قواعد بيانات تحوي معلومات مهيكلية بطريقة حاسوبية عن اللغات، وذلك لمساعدة مستعمل الحاسوب على حل بعض المسائل المتعلقة باللغة، مثل: برامج الترجمة الآلية والتصحيح الإملائي والنحوي، وتحليل الخطاب آلياً..." (تاويريت وشاشه، 2022، ص: 534). والنظر في هذا المفهوم يجعلنا نقف على الأبعاد الأساسية للمعالجة، من حيث الوسيلة والمادة والمراحل والهدف؛ إذ تمر المعالجة الآلية للغات الطبيعية بمرحلتين أساسيتين هما: التحليل والتوليد الآلي للغات، ونعني بالتحليل الآلي: "عملية تفكيك وتجزئ آلي للنصوص" (رامي وتورابي، 2001، ص: 81)، التي هي أكبر وحدة لغوية، إلى عناصر جزئية اتحدت ليتشكل النص، ويقابل قدرة الفهم عند الإنسان.

أما التوليد الآلي فهو: "عملية اشتقاقية لإيجاد الكلمات والجمل التنسيقية، سواء كانت مثبتة في معجمه الإلكتروني أو لم تكن مثبتة، ويكون التوليد بتحديد الإمكانيات التأليفية للأبجدية الصوتية والصرفية والنحوية" (رامي وتورابي، 2001، ص: 81)، وهو ما يقابل قدرة الإنسان الفطرية على إنتاج جمل جديدة إنتاجاً غير محدود. ويطبق التحليل والتوليد على كل المستويات (الصوتي، الصرفي، النحوي...) وفق تقسيم افتراضي، تنطلق فيه الآلة من أصغر وحدة (الصوت) إلى أكبرها (الدلالة) لكونها لا تستوعب النظام اللغوي دفعة واحدة.

2.3 المعالجة النصية (Traitement automatique textuelle)

تعرف على أنها: "استخدام الحاسوب في كتابة النصوص اللغوية، وتحويلها إلى لغات الحاسب الرقمية لتحليلها، وتسهم بصورة كبيرة في تسهيل دراستها، والوقوف على إحصائها بصورة سريعة ودقيقة، وهي تساعد في فهم خصائص النصوص اللغوية محل الدراسة" (محمود، 2012، ص: 39). حيث يكتب النص باللغة الطبيعية في برنامج المعالجة النصية، المزود ببرمجيات تحول ذلك النص إلى لغة الآلة، وبذلك يتحول من الطابع الإنساني إلى التقني، ليتمكن البرنامج من تحليله والتحقق من صحته والإشارة إلى الكلمات التي يعتبرها خاطئة، وتقديم اقتراحات لتصحيحها. ومن أبرز متطلباتها التي يقوم عليها اشتغالها نذكر:

1.2.3 أنظمة المعالجة اللسانية:

تتم معالجة النصوص وفق تقسيم افتراضي مرتبط بالمستويات اللغوية (صوتي، صرفي، نحوي، دلالي) وعلى أساسه خُصص لكل مستوى لغوي معالج خاص به، يتكون كل معالج من (محلل ومولد)، حيث يقوم المحلل بتفكيك البنية اللسانية إلى وحداتها الرئيسية والفرعية، من أجل تحديد خصائصها ووظائفها، في حين يقوم المولد بـ: العملية العكسية، إذ يعمل على تركيب تلك الوحدات من أجل الحصول على بنى لسانية مكتملة.

2.2.3 المدققات اللغوية:

تعتبر المدققات العمود الفقري لكل عملية معالجة، ذلك أنها هي التي تقوم باكتشاف الأخطاء وتصويبها، وتمثل في: المدقق الإملائي والمدقق النحوي.

المدقق الإملائي: هو "أحد برمجيات المعالجة الآلية للغة، التي يعتمد عليها لتصويب المدخلات واقتراح بدائل لبنى لغوية أصوب وأكثر دقة، حيث يقوم بالتحقق من مدى مطابقة المدخلات التي تكتب في خانة البحث بما زود به الحاسوب من بيانات وفي حال تعذر الحصول على نتيجة يقوم باقتراح بدائل أخرى للمستخدم" (تاويريت وشاشة، 2022، ص: 536)، وبناء على ذلك يعد المدقق الإملائي وسيلة لاكتشاف الأخطاء الإملائية وتصحيحها.

المدقق النحوي: هو: "آلية برمجية تتعامل مع النصوص، وتتلخص وظيفته في مراجعة صحة الكلمة إعرابياً، وصحة الجملة تركيباً، ومن ثم اقتراح بدائل نحوية صحيحة لهما، لأجل الوصول إلى الصواب النحوي أو التركيبي" (جمعة، 2012، ص: 191) ويمكن الفرق بين المدقق الإملائي والمدقق النحوي في كون وظيفة الأول هي مراجعة واكتشاف الأخطاء على مستوى نظام الكتابة وتصحيحها، في حين يختص الثاني بتحديد مدى صحة القواعد الضابطة للجملة، ويصوبها استناداً إلى ما بنيت عليه من قواعد صرفية ونحوية.

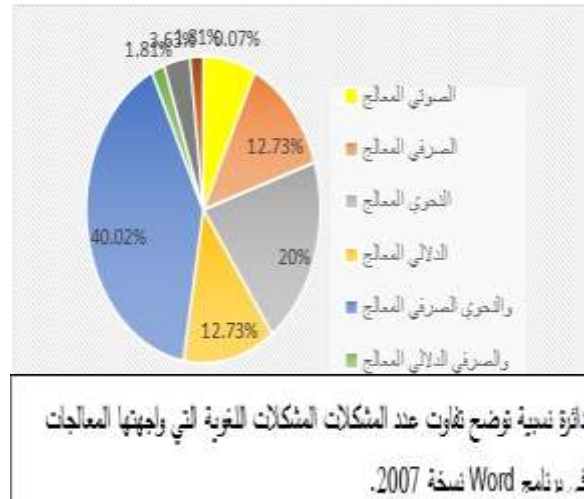
4. نتائج الدراسة:

حاولت ورقتنا البحثية تقييم قدرة برنامج المعالجة النصية **word** بنسخته القديمة 2007 والحديثة 2019 على معالجة النص العربي، وذلك بالوقوف على المشكلات اللغوية التي يواجهها عند إدخال نص خال من الأخطاء متمثل في سورة البقرة، وبعد رصد وتحليل كل ما أشار إليه على أنه خاطئ توصلنا إلى جملة من النتائج الخاصة بكل نسخة على حدة، نذكرها فيما يلي:

نسخة 2007:

فور إدخال سورة البقرة إلى البرنامج قام بوضع سطر تحت خمس وخمسين (55) من عناصرها اللغوية توزعت بين مفرداتها وتراكيبها، كإشارة على أنها خاطئة.

لقد شكلت هذه المفردات والتراكيب المدروسة جزءاً من المدخلات التي واجه برنامج وورد نسخة 2007 مشكلة في التعامل معها، عند معالجة سورة البقرة، وبعد قيامنا بتحليلها ومحاولة تحديد لبّ الإشكال، حرّينا أن نقوم بعملية إحصائية نحصى بها عدد المشكلات المنسوبة إلى كل معالج، لنصل إلى نتائج نقيم بها أداء كل منهما، وسنعرض نتائج هذه العملية الإحصائية في الدائرة النسبية التالية:



نلاحظ على هذه الدائرة النسبية مايلي:

— واجه المعالج الصوتي 7.27% من المشكلات اللغوية، بمعدل أربع مشكلات من أصل خمس وخمسين، وبهذا يعتبر المعالج الذي يواجه أقل عدد من المشكلات.

— اعترض المعالج الصرفي والمعالج الدلالي ما نسبته 12.73% من المشكلات اللغوية أي بمعدل سبع مشكلات.

— واجه المعالج النحوي 20% من المشكلات ما معدله إحدى عشرة مشكلة لغوية، ما يجعله المعالج الذي يعاني الوقوع في أكبر نسبة من المشكلات.

— تعرض المعالج الصوتي إلى 7.27% من المشكلات اللغوية، بمعدل أربع مشكلات من أصل خمس وخمسين، وبهذا يعتبر المعالج الذي يواجه أقل عدد من المشكلات.

— اعترض المعالج الصرفي والمعالج الدلالي ما نسبته 12.73% من المشكلات اللغوية أي بمعدل سبع مشكلات.

— واجه المعالج النحوي 20% من المشكلات؛ أي إحدى عشرة مشكلة لغوية، ما يجعله المعالج الذي يعاني الوقوع في أكبر نسبة من المشكلات.

— بعض المشكلات اللغوية شكلت عائقاً لأكثر من معالج واحد، حيث اشترك كلٌّ من المعالج الصرفي والنحوي والدلالي بنسب متفاوتة في الوقوع أمام مدخلات صُعُب التعرف عليها ومعالجتها.

— لم يتعرف برنامج وورد على العناصر اللغوية التي برزت فيها الخصائص الكتابية للخط العثماني الذي رسم به القرآن الكريم، مثل: حذف الألف بعد واو الجماعة في أربعة مواضع (فاء وعتو...).

— صعب على البرنامج معالجة الأسماء والأفعال المتصلة بسوابق ولواحق متعددة، مثل: فسَيَكْفِيْكُمْ.

— عجز البرنامج عن التعرف على بعض الأوزان الصرفية التي وردت وفقها المدخلات، مثل: فَعَالِي التي جاءت على وزنها لفظة (أسارى).

— لم يستطع البرنامج معالجة العديد من القوالب النحوية، نذكر منها: {أداة استفهام (الهمزة) + حرف عطف (فاء)} + فعل مضارع + واو الجماعة، مثل: أَقْتَطِعُونَ {

{لام التوكيد + فعل مضارع + نون التوكيد الثقيلة + ضمير متصل في محل نصب مفعول به، مثل: لَنَبْلُوَنَّكُمْ}.

- {حرف جر + اسم مجرور على هيئة ضمير متصل، مثل: به}.
- {حرف جر + اسم مجرور + مضاف اليه على هيئة ضمير متصل، مثل: ببعضها}.
- لم يتعرف البرنامج على الفعل: إدّراً.
- اعتبر البرنامج بعض الأفعال الواردة وفق قواعد نحوية معينة لأغراض بلاغية خاطئة، مثل: اتقون، والذي حذفت منه ياء المتكلم الزائدة المتصلة بأفعال الأمر.
- واجه البرنامج مشكلة في التعامل مع التراكيب المدخلة، التي تتضمن أفعالاً متصلة بأكثر من ضمير، مثل: فسَيَكْفِيكَهْم.
- لم يطبق البرنامج كل القواعد الصرفية والقوالب النحوية التي زُود بها لمعالجة المدخلات على كل الأفعال المدخلة ولم يعمّمها، فمثلاً: تعرّف البرنامج على التركيب التالي: (يحييكم: فعل مضارع + الضمير المتصل - كم - الذي هو في محل نصب مفعول به) واعتبره صحيحاً، في حين اعتبر التركيب التالي خاطئاً (يميتكم) رغم أنه مزود بالفعل (يميت).
- لم يتعرف البرنامج على المفردات التراثية القديمة مثل: شية، جنفا..
- لم يتعرف على الأسماء والأفعال غير الشائعة، مثل: ميكال، ننشزها..
- صعب على البرنامج التعامل مع العبارات المجازية.

ولضيق المقام سنعرض ثلاثة نماذج من المشكلات اللغوية التي واجهها على اختلاف أنواعها:

فأءو: قال تعالى: "لِلَّذِينَ يُؤْثِرُونَ مِنْ نِسَائِهِمْ تَرِيصُ أَشْهَرٍ فَإِنْ فَأءُو فَإِنَّ اللَّهَ غَفُورٌ رَحِيمٌ". (سورة البقرة: الآية 226)

اعتبر برنامج وورد هذه الكلمة خاطئة واقترح تصحيحها ب: فائو، فاء، فأءوا، فراءو، فأءا.

يُظهر الاقتراح الثالث (فأءوا) أن المعالج الصرفي يشتمل على جذر هذا الفعل، كما أن المعالج النحوي مزود به في الزمن نفسه الذي ورد وفقه (الماضي)، كما أننا لو قارنا بين (فأءو) المدخلة و(فأءوا) الاقتراح الثالث ذاته، سنلاحظ أن الفرق هو عدم رسم الألف الفارقة في الأولى ورسمها في الثانية.

ولذا نقول أنّ البرنامج غير مزود بمواضع رسم الألف الفارقة ومواضع حذفها في الحالات الاستثنائية كما هو الحال في الفعل (فأءو) إذا ما تمت كتابته بالخط العثماني، حيث قال أبو عمرو عثمان بن سعيد الداني: "واتفقت المصاحف على حذف الألف بعد واو الجمع في أصلين مطّردين وأربعة أحرف، فأما الاصلان فهما (جاءو و باءو) حيث وقعا، وأما الأربعة أحرف فأولها في البقرة (فإن فأءو)..."

وهذا مثال عن الأخطاء الناتجة عن الحالات الاستثنائية في اللغة العربية، التي نأمل أن تستدركها البرامج اللسانية الحاسوبية مستقبلاً.

فأنجيناكم:

قال تعالى: "وَإِذْ فَرَقْنَا بِكُمْ الْبَحْرَ فَأَنْجَيْنَاكُمْ وَأَغْرَقْنَا آلَ فِرْعَوْنَ وَأَنْتُمْ تَنْظُرُونَ" (سورة البقرة: الآية 50)

اقترح البرنامج التصحيحات التالية: فأنجيناكم، فأنديناكم، فأنجي، فأنجي، فأنجيك، ونرجح في هذا المقام أن تكون المشكلة مزدوجة، أي صرفية نحوية، فبالنسبة للمعالج الصرفي فمن المحتمل أن يكون قد واجه عجزاً عند تفكيكه لهذه المدخلة وهي

متصلة بسابقة ولاحتقين فأُنجِنَاكُمْ، والدليل على ذلك أننا عندما فصلناها عن السابقة (الفاء) واللاحقة الثانية (كم) وأصبحت (أُنْجِنَا) اعتبرها صحيحة.

وعليه فهو يعاني فقرا في قاعدة بياناته إزاء الحالات المختلفة التي ترد وفقها المدخلات.

كما تكمن المشكلة في المعالج النحوي، وذلك لعدم تعرفه على هذا القالب الذي وردت وفقه: (أُنْجِي_الفعل_+نون الجماعة_الفاعل_+كم_المفعول به_)، فحتى عندما غيّرنا ضمير الفاعل وحولنا المدخلة الى (أُنْجَاكُمْ) اعتبرها خاطئة كذلك، في حين أننا عندما أدخلناها مع الضمير ذاته وفصلناها عن ضمير المفعول به (كم) وأصبحت (فَأُنْجِنَا) اعتبرها البرنامج مدخلة صحيحة وتعرف عليها، وهذا لأن المعالج النحوي غير مزود بقالب نحوي يرد وفقه الفعل (أُنْجِي) متصلا بضمير المفعول به، فلو ورد هذا الأخير اسما لا ضميرا متصلا، لما واجه المعالج النحوي أي مشكلة.

وعليه فإنّ المشكلة تتمثل في: عدم احتواء قاعدة معطيات المعالج الصرفي للنماذج المختلفة التي ترد وفقها المدخلات، خاصة التي تكون فيها متصلة بالعديد من السوابق واللاحق، وكذلك فقر المعالج النحوي من القوالب النحوية المختلفة التي ترد وفقها التراكيب، كورود بعض الأفعال متصلة بالمفعول به على هيئة ضمير.

بِسْمِ اللَّهِ:

قال تعالى: "الْفُقَرَاءُ الَّذِينَ أُحْصِرُوا فِي سَبِيلِ اللَّهِ لَا يَسْتَطِيعُونَ ضَرْبًا فِي الْأَرْضِ يَحْسَبُهُمُ الْجَاهِلُ أَغْنِيَاءَ مِنَ التَّعَفُّفِ تَعْرِفُهُمْ بِسْمِ اللَّهِ لَا يَسْأَلُونَ النَّاسَ إِلْحَافًا وَمَا تُنْفِقُوا مِنْ خَيْرٍ فَإِنَّ اللَّهَ بِهِ عَلِيمٌ". (سورة البقرة، الآية: 273) عرض البرنامج الكلمات التالية بديلا: بستمهم، قسيمهم.

عندما نقوم بإدخال لفظة (سيمة) مفصولة عن الباء التي سبقتها، و(هم) التي لحقتها، يثبت البرنامج على حكمه في أنها لفظة خاطئة، إذ ليست المشكلة في السوابق واللاحق ولا في التركيب الذي وردت وفقه، بل هي مشكلة دلالية، تتمثل في أن المعالج الدلالي لم يزود بهذه المفردة، لذا لم يتعرف عليها ولا على معناها، وهو: "العلامة". وعليه نسجل قصورا في المعالج الدلالي ومخزون قاموسه الإلكتروني.

نسخة 2019:

نذكر من النتائج المتوصل إليها ما يلي:

— فور إدخال سورة البقرة الى البرنامج قام بالإشارة إلى ثلاثة عشر عنصرا لغويا عن طريق وضع سطر تحتها لكونه اعتبرها خاطئة.

— لم يتمكن البرنامج من التعرف على المفردات التي تنتهي برسم الألف الخنجرية في نهايتها كخاصية من الخصائص الكتابية للخط العثماني واعتبرها خاطئة، مثل: اصطفى، الهدى...

— واجه البرنامج مشكلة في التعامل مع المدخلات التي تتجلى فيها مواضع حذف الألف الفارقة بعد واو الجمع في المصحف الشريف مثل: باءو.

— وجد البرنامج صعوبة في معالجة بعض المدخلات التي وردت متصلة بلاحقة مهما كان نوعها، مثل: ترضاه...

— صعب على البرنامج معالجة بعض المدخلات التي ترد متصلة بضمير جمع مذكر المخاطب، مثل: مبتليكم.

— بعض المدخلات شكّلت مشكلة لغوية لكلتا النسختين تمثلت في: يتسنه، اذارثتم، باءو، أوكلما.
— تعرّف البرنامج على كل الأسماء المتعلقة بالحضارة الإسلامية مثل: ميكال، عيسى، موسى...
— يقدّم البرنامج مرادفات للمفردة التي اعتبرها خاطئة في البداية وتمت إضافتها إلى معجمه من طرف المستخدم.
— يعتمد البرنامج على السياق في تحديد معنى اللفظة التي تضاف الى معجمه لتقديم قائمة بمرادفاتھا، حيث يبحث في قواعد بياناته المزود بها عن السياقات المشابهة للسياق الذي وردت فيه المدخلة ويقابل بين معاني الالفاظ.
وسنعرض ثلاثة نماذج من المشكلات اللغوية التي واجهها على اختلاف أنواعها:

اصطفي:

قال تعالى: "وَوَصَّىٰ بِهَا إِبْرَاهِيمُ بَنِيهِ وَيَعْقُوبُ يَا بَنِيَّ إِنَّ اللَّهَ اصْطَفَىٰ لَكُمُ الدِّينَ فَلَا تَمُوتُنَّ إِلَّا وَأَنتُمْ مُسْلِمُونَ". (سورة البقرة، الآية: 132)

اعتبر البرنامج هذه المدخلة خاطئة، واقترح لها التصحيحات البديلة التالية: (اصطفي، اصطفي).
عند النظر في المقترح الأول نجد أنّه مطابق للمدخلة ذاتها، أي أنه مزود بنوع هذا الفعل وزمانه وجذره ووزنه، مما يجعلنا نستبعد ان تكون المشكلة في المعالج الصرفي او النحوي.
وبعد إضافة هذه المفردة إلى قاموسه والاطلاع على المرادفات التي يقدمها والتي كانت (اختار، اجتبي، استقى)، ظهر جليا من خلال المرادف الأول والثاني أنّه يعرف على معناها، مما يجعلنا كذلك نستبعد أن تكون العلة في المعالج الدلالي.
وعليه قمنا بالتدقيق في الجانب الإملائي المقابل للجانب الصوتي للمفردة، وأعدنا كتابتها بالخط العادي لا العثماني (دون ألف خنجرية في نهاية الكلمة)، فاعتبرها البرنامج صحيحة.
نستنتج أن المشكلة تتعلق بنظام كتابة النص العربي؛ ذلك أنّه غير مزود بقاعدة بيانات شاملة لكل خصائص الخط العثماني.

ترضاها:

قال تعالى: "قَدْ نَرَىٰ تَقَلُّبَ وَجْهِكَ فِي السَّمَاءِ ۚ فَلَنُوَلِّيَنَّكَ قِبْلَةً تَرْضَاهَا .." (سورة البقرة، الآية: 144)

اقترح البرنامج التصحيحات التالية: (ترداها، تعرضها، تراها).

للقوف على السبب الذي جعل البرنامج يعتبر المدخلة خاطئة قمنا بتغيير الضمير المتصل إلى ضمير متصل آخر (ترضاها، تعرضها..). فاعتبرها خاطئة كذلك، ثم قمنا بإدخال الفعل دون إلحاقه بأي ضمير (ترضى) فاعتبره صحيحا، فتبين أنّ هذا البرنامج مزود بهذا الفعل وجذره إلا أنّه غير مزود به متصلا بلاحقة، كما أنّه غير مزود به ضمن قالب نحوي يرد فيه على هذا النحو: (فعل + فاعل مستتر + مفعول به— على هيئة ضمير متصل—)، لتكون بذلك المشكلة كامنة في المعالج الصرفي والنحوي.

مبتليكم:

قال تعالى: "فَلَمَّا فَصَلَ طَالُوتُ بِالْجُنُودِ قَالَ إِنَّ اللَّهَ مُبْتَلِيكُمْ بِنَهَرٍ فَمَنْ شَرِبَ مِنْهُ فَلَيْسَ مِنِّي..".

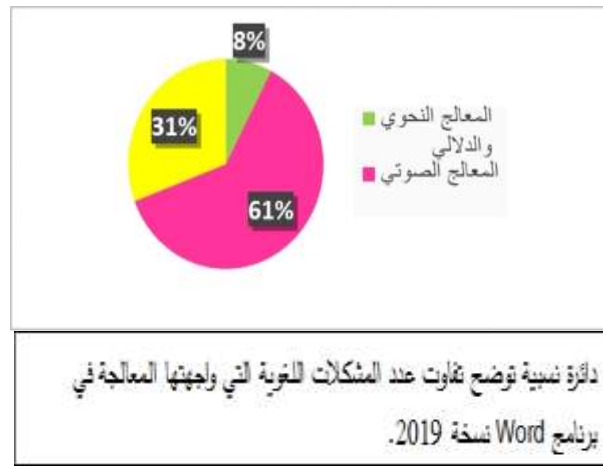
سورة البقرة الآية (249)

قدم البرنامج الاقتراحات التالية: (مبتليكم، مبتدليكم، متبذليكم).

لتحديد موطن القصور قمنا بتغيير الضمير المتصل (كم) إلى ضمائر أخرى ووصل المدخلة بها (مبتليهم، مبتليهما...) فأشار إلى أنها خاطئة كذلك، ثم قمنا بفصل هذه المدخلة عن الضمير المتصل بها (مبتلي) أي فصل الخبر عن المضاف اليه، فتعرف عليها البرنامج واعتبرها صحيحة.

إن المعالج الصرفي يعاني فقرا في قاعدة بياناته إزاء الحالات التي يرد وفقها اسم الفاعل متبوعا بلاحقة، كما أنّ هذه المشكلة اللغوية دالة على أنّ المعالج النحوي يفتقر إلى القالب النحوي الذي يرد وفقه (الخبر _ مبتلي _ + مضاف اليه _ الضمير المتصل _).

إنّ تحليل المشكلات اللغوية التي واجهها البرنامج نسخة 2019 أبرزَ بوضوح أنّ السبب كثيرا ما يرجع إلى فقر قاعدة بيانات المعالجات، والتي بدورها تفاوتت في عدد المشكلات التي واجهتها وشكّلت عائقا أمام قيامها بوظيفة معالجة النص المدخل، وسنوضح ذلك بدقة في الدائرة النسبية التالية:



حيث نلاحظ من المعطيات الواردة فيها ما يلي:

- واجهت المعالجة الصوتية والكتابية أكبر عدد من المشكلات اللغوية؛ أي ما نسبته 61%.
- جاءت المشكلات اللغوية الأخرى مشتركة بين بقية المعالجات.

المقارنة بين نسختي وورد 2007 و 2019:

بعد الوقوف على نتائج معالجة النسخة القديمة 2007 والنسخة الحديثة 2019 لبرنامج **Word** نستخلص ما يلي:

- عدد المشكلات اللغوية التي واجهتها نسخة 2007 بلغ خمسا وخمسين (55)، في حين كان عدد المشكلات التي واجهتها نسخة 2019 ثلاث عشرة (13) مشكلة فقط.

بعض المدخلات شكّلت مشكلة لغوية لكلتا النسختين تمثّلت في: يتسنه، اذارئتم، باءو، أوكلما.

- تمكّن برنامج **Word** بنسخته الحديثة 2019 من:

معالجة أغلب الأسماء والأفعال على اختلاف أوزانها وجذورها وعدد سوابقها ولواحقها، وتجاوز جل العقبات الصرفية التي كانت تواجهها نسخة 2007.

ومعالجة الحمل على اختلاف القوالب النحوية التي ترد وفقها وتجاوز أغلب المشكلات النحوية التي كانت تقف عائقاً أمام المعالج النحوي في النسخة القديمة.

التعرف على كل المفردات التراثية القديمة والأسماء المتعلقة بالحضارة الإسلامية. في حين واجهت نسخة 2007 مشكلة في معالجة الكثير منها، وبهذا تكون نسخة 2019 هي الأنجح في التعامل مع نظام اللغة العربية ومعالجتها.

إن نسخة 2019 الأكثر اعتماداً على السياق لتحديد معنى المدخلات، فحتى عند معالجتها للعناصر اللغوية التي لا يتعرف عليها في البداية فإنها فور إدراجها في قاموسها من طرف المستخدم يقوم بالعودة الى ما زُود به من سياقات مشابهة للسياق الذي وردت فيه ويستنتج معناها.

كما أظهر برنامج (Word) تفاوتاً في مدى قدرته على معالجة نظام اللغة العربية بين نسخته القديمة 2007 والحديثة 2019، حيث واجهت النسخة القديمة مشكلات لغوية عديدة ومتنوعة مسّت كل المعالجات وذلك راجع إلى:

- يعاني برنامج (Word) نسخة 2007 فقراً نسبياً متفاوتاً في قواعد بياناته الصوتية والصرفية والنحوية والدلالية.
- عدم تزويد البرنامج ببعض القواعد النحوية الاستثنائية في نظام اللغة العربية.
- تزويد البرنامج بما هو شائع من المفردات والمعاني والتراكيب والاوزان وما هو متداول فقط.
- التطور اللغوي وظهور مفردات جديدة، أو إحياء مفردات قديمة بمعان جديدة.
- عدم اثناء قاعدة بيانات البرنامج بالمفردات والسياقات التي تعكس الثقافة العربية والحضارة الإسلامية، لذا لم يتعرف على أسماء الملائكة.

في حين نجد أنّ نسخة 2019 قد سجلت نجاحاً نسبياً في التعامل مع النص اللغوي العربي رغم بعض المشكلات التي واجهتها، ونفسر ذلك ب:

- بذل مجهودات مقبولة للعمل على تزويد هذه النسخة بجلّ ما تفتقر إليه النسخة القديمة من قواعد بيانات على صعيد كل المستويات اللغوية.
- تم تزويدها بقاعدة بيانات ثرية الى حد مقبول عن الثقافة الإسلامية والحضارة العربية بما تحمله من أسماء وألفاظ وسياقات لغوية خاصة باللغة العربية.
- تم اثناء جلّ المعالجات بما يخص النظام اللغوي العربي بنجاح.

5. خاتمة:

لقد حاولنا في هذه الورقة البحثية تقييم قدرة برنامج المعالجة النصية **Word** بنسخته القديمة 2007 والحديثة 2019 على معالجة النص العربي المتمثل في النص القرآني "سورة البقرة"، والوقوف على المشكلات التي يواجهها، ولأنّ ما نصبو إليه من وراء ذلك هو الرقي بأدائه أكثر في معالجة النصوص العربية، فإننا نقترح:

- إثراء قاعدة بيانات المعالج الصوتي بأنواع الخطوط العربية كالخط العثماني والخط الكوفي، وبالخصائص الكتابية لكل منها.
- استخراج كل المفردات والأسماء والأوزان والأفعال والتراكيب والسياقات، التي جاءت في القرآن الكريم والأحاديث النبوية الشريفة صحيحة السند، باعتبارها مرجعاً أساساً للحضارة الإسلامية.

- تزويد البرنامج بنصوص من الشعر العربي التي نُظِّمَتْ في العصور المختلفة الجاهلي والإسلامي والحديث...، لكونها تعبّر عن التطور اللغوي الحاصل على مستوى المفردات والسياقات، وتجسد مراحل حياة اللغة العربية، فلا يواجه البرنامج مشكلة في التعامل مع ما هو قديم أو حديث.

- عدم الاقتصار على تزويد البرنامج بما هو شائع ومتداول.

- تكثيف الجهود اللسانية لبناء المكانز والذخائر والمعاجم الإلكترونية، والاستناد إليها في تزويد برنامج **Word** بقاعدة بيانات تضم كل: المفردات والأوزان، والجمل والقوالب النحوية، المتعلقة بما أنتج وما يتم إنتاجه باللغة العربية.
- إزالة الفجوة الحاصلة بين الجهود اللسانية والحاسوبية، وتظافر الجهود حتى تكون النتائج مرضية.

6. قائمة المراجع:

- المصحف الشريف، برواية حفص عن عاصم.

• المؤلفات:

- 1- بن كثير أبي الفداء إسماعيل بن عمر القرشي الدمشقي، (2000)، تفسير القرآن العظيم، بيروت، دار ابن حزم.
- 2- محمود، عصام، (2012)، اللسانيات الحاسوبية العربية، الإسكندرية، دار الوفاء لدنيا الطباعة والنشر.
- 3- جمعة، عمرو، (2012)، معالجة اللغة العربية آلياً-دراسة حاسوبية لظاهرة اللبس اللغوي في العربية المعاصرة، الإمارات العربية المتحدة، دار الكتب الوطنية.
- 4- الصباغ، عبد الوهاب، 2002، علم المعلومات، الأردن، دار الثقافة للنشر والتوزيع.

• المقالات:

- 1- حسام الدين تاوريرت وإيمان شاشه، (2022)، دور برمجيات المعالجة الآلية للغة في بناء المعاجم الحاسوبية -المدقق الإملائي والمحلل الصرفي أنموذجاً-، مجلة إشكالات في اللغة والادب، المجلد 11، العدد 04.
- 2- توراي، سالم رامي وعبد الرزاق، (2001)، المولد الصرفي للكلمات المعجمية العربية التوليدية والنسقية والترجمة الآلية، منشورات معهد الدراسات والأبحاث للتعريب، المجلد الأول.