

المدقق الإملائي للعربية في الحاسوب

Correcteur Orthographiques Arabe Dans Ordinateur

أبو بكر أحمد إبراهيم بوبكر*

باحث في اللسانيات والمعمية العربية) انواكشوط – موريتانيا

diaabou017@gmail.com

نشر: 2026-03-31

مقبول: 2026-05-06

استلم: 2022-06-09

ملخص: تعنى هذه الورقة البحثية بدراسة بعضا من جوانب الإشكالات الإملائية للغة العربية في الحاسوب، وتحاول أن تعالج هذه القضية من خلال النظر في المدققات الإملائية، مستعينة في ذلك بمنهج الوصف والتحليل، وذلك من خلال محاولة الإجابة عن الأسئلة التالية: ما هي الأسس التي تقوم عليها المدققات الإملائية؟ وكيف يتعامل الحاسوب مع التدقيق الإملائي للغة العربية؟ ما دور المدقق الإملائي في معالجة النصوص العربية؟ وكيف تطورت المدققات الإملائية للغة العربية في الحاسوب إلى هذا المستوى من التقدم؟ وما هي التحديات التي يواجهها المدقق الإملائي في الحاسوب؟ ثم كيف وصل الحاسوب إلى طباعة الحروف العربية؟

الكلمات المفتاحية: الحاسوب، المدققات الإملائية، المدونة

Abstract: This research paper explores several aspects of Arabic orthographic challenges within computing environments. It addresses these issues by examining spell-checking systems using a descriptive and analytical methodology. The study aims to answer the following core questions: What are the fundamental principles underlying spell-checkers? How does a computer process Arabic spell-checking? What role do spell-checkers play in processing Arabic texts? How have Arabic spell-checkers evolved to their current level of sophistication? What challenges do these digital spell-checkers face? Finally, how did computing systems achieve the capability of rendering and printing Arabic characters?

Keywords: Computer, Spell-checkers, Linguistic Corpus.

*أبو بكر أحمد إبراهيم بوبكر

المقدمة

الأصل في اللغة أن تكون منطوقة، ولكن أغلب اللغات العالمية تنطق وتكتب. والكتابة هي الإملاء، ويقصد بها تحويل الأصوات المسموعة رموزا مرئية، عن طريق المواضعة والاصطلاح في بيئة اجتماعية معينة. ويعد الحاسوب أهلا للتعامل مع اللغة، غير أن "الاستعمال الرئيس للحاسوب بالنسبة لمختلف اللغات، لا يزال مجال الكتابة، أما مجال النطق والأصوات، فقد حدثت فيه تطورات مهمة خلال العقود الأخيرة، إلا أنه لا يزال محدودا بالمقارنة مع مجال القراءة" (زكي، 1996، ص2).

وقد يعتري هذا الإملاء إشكالات على مستوى الكتابة، تتمثل في رسم بعض الحروف بأشكال مختلفة، حسب موقعها الجغرافي من الكلمة، وبعضها عبارة عن حذف أو زيادة بعض الحروف المنطوقة أو غير المنطوقة أثناء الكتابة، وتتمثل بعضها الآخر في الفصل والوصل بين الحرفين المتجاورين، سواء في كلمة واحدة أو كلمتين متجاورتين.

تعتمد الورقة البحثية على الخطة التالية لدراسة هذه الإشكالية

خطة البحث

تقسّم البحث من حيث هيكله، إلى:

تمهيد

- 1- نبذة تاريخية عن حوسبة اللغة العربية (النشأة - الطابعات - المحارف)
- 2- المدقق الإملائي وتدقيق النصوص (التعريف - المدونة - مدقق وورد - برامج التدقيق - مواقع للتدقيق الإملائي)
- 3- التحديات التي يواجهها المدقق الإملائي (أمثلة)

الخاتمة

تمهيد

يحتوي الحاسوب على العديد من الأدوات التي تتعامل بها مع اللغة العربية، من حروف ومعاجم ومدونات وبرامج تحليل بمستوياتها اللغوية. فهو قادر على معالجة النصوص العربية وفق الخوارزميات المبرمجة عليه، فيفهم النص ويقرأه، ويتعرف على الحروف والأصوات والنبر والتنغيم....، وعنده إمكانية كبيرة في تصحيح الأخطاء الإملائية وفق قاعدة بيانات معجمية مبرمجة فيه، والقاعدة " هي المورد المباشر الذي يعتمد عليه في بناء وتطوير آلية التدقيق الإملائي" (السعيد، د.ت. ص7).

وقد كان المتخصصون في اللغة قديما، من يقومون بالتدقيق الإملائي بعد تحرير النصوص، لأن النصوص في الغالب، تنتج عنها أخطاء إملائية، بسبب جهل كاتبه أو سرعة كتابته وطباعته، أو بسبب نسيانه تنقيط بعض الحروف أو تخطيه، أو بسبب وضعه علامات الترقيم في غير محلها، أو بسبب آخر غير هذا وذلك. غير أن الحاسوب اليوم يعد عند البعض المدقق الأدق في تصحيح النصوص، ومعالجة الأخطاء في مختلف اللغات، سواء تعلقت الأخطاء بتوليد ألفاظ ليست من أصل اللغة العربية، كتنوين ما لا ينون، والتشكيل والإضافة، والحذف والقلب والابدال....، أو تعلقت الأخطاء بمستوى الطباعة، كالفصل والوصل بين الحروف، على غير المعتاد، أو تشديد ما لا يشدد، أو غيرها من الأخطاء اللغوية، وذلك بتفعيل خاصية التصحيح التلقائي من الحاسوب. غير أن هذه الخاصية قد تكون سببا في توليد الأخطاء، كما سيتبين ذلك لاحقا.

1- نبذة تاريخية عن حوسبة اللغة العربية

1-1 النشأة

يذكر الباحث اللغوي بسام أغبر، أن بدء استعمال الحاسوب للبحث اللغوي، كان في بداية الخمسينيات من القرن العشرين، في أمريكا ثم في أوروبا، وكان بداية استعماله عند العرب مع إبراهيم أنيس، وعلي حلمي، في محاولة لإحصاء الحروف الأصلية لمواد اللغة العربية من معجم الصحاح، عام 1971م (2017). غير أنه قد تأكد في الحقيقة، أن عبد الصبور شاهين هو فعلا من قام بهذه الدراسة، حسب المعتز بالله السعيد (2019).

واللسانيات الحاسوبية للغة العربية، بمفهومها الحديث، بدأت بإدخال أجزاء من القرآن لحوسبتها، وذلك في السبعينات من القرن الماضي (علي، 1994، ص4).

وفي بداية ظهور اللسانيات الحاسوبية العربية، ظهر قلق واضح على اللغة العربية من الضياع، وقد تناول نبيل علي هذه القضية، بالتفصيل والإجمال، مشيراً إلى أن البيئة الحاسوبية غير متوافقة مع اللغة العربية في كثير من أحيائها، من حيث الهيكل الكتابي والتشكيل، ما قد يجعل اللغة العربية في حداد الموت (1994، ص338).

2-1 الطابعات

تجدر الإشارة إلى أنه لم يكن الحاسوب مزوداً بطابع للحروف العربية قبل الستينات من هذا القرن، وقد كانت أحزمة بيلتس (Belts) أول حروف للطباعة العربية، مع وجود صعوبة لقراءة منتجاتها الطباعية، ولتحسين شكل الكتابة العربية على الحاسوب، قامت الشركات بالاتصال بالجامعات العربية ومراكز الحواسيب الضخمة في العالم العربي، فتمكنت من إنتاج الطابعات النقطية (Matrix Printers) بحروف عربية رديئة، ولكنها سرعان ما تطورت بعد دراسات في الندوات والمؤتمرات حول الحاسوب (زكي، 1996، ص3).

3-1 المحارف

استطاعت الشركات أن تنتج أطقماً عربية تحمل الحروف العربية بأشكالها وكتابتها المختلفة، من التشكيل والهمزات والكشيدة والتاء المربوطة والألف المقصورة، واللامات الثمانية، وغيرها من الرموز التي تعني بخدمة الخط العربي (زكي، 1996، ص5).

ب	ب	ا	ئ	ا	ا	ا	ا	ا	ا	ا
ج	ج	ج	ث	ث	ث	ث	ث	ث	ث	ث
س	ز	ر	ذ	ذ	ذ	ذ	ذ	ذ	ذ	ذ
ض	ض	ض	ص	ص	ص	ص	ص	ص	ص	ص
غ	غ	ع	ظ	ظ	ظ	ظ	ظ	ظ	ظ	ظ
ل	ل	ك	ق	ق	ق	ق	ق	ق	ق	ق
ه	ه	ن	م	م	م	م	م	م	م	م
لا	لا	لا	لا	لا	لا	لا	لا	لا	لا	لا

0621	0622	0623	0624	0625	0626	0627	0628	0629	062A	062B
ث	ت	ة	ب	ا	ئ	ا	ا	ا	ا	ا
062C	062D	062E	062F	0630	0631	0632	0633	0634	0635	0636
ض	ص	ش	س	ز	ر	ذ	د	خ	ح	ج
0637	0638	0639	063A	063B	063C	063D	063E	063F	0640	0641
ف	-									
0642	0643	0644	0645	0646	0647	0648	0649	064A	064B	064C
ق	ك	ل	م	ن	ه	و	ي	ي	=	ق
064D	064E	064F	0650	0651	0652	0653	0654	0655	0656	0657
=	-	و	-	س	ه	-	ع	ع	ف	ف
0658	0659	065A	065B	065C	065D	065E	065F	0660	0661	0662
0663	0664	0665	0666	0667	0668	0669	066A	066B	066C	066D
٣	٤	٥	٦	٧	٨	٩	%	,	,	*
066E	066F	0670	0671	0672	0673	0674	0675	0676	0677	0678

(الشكل 2/1): ب

بتاريخ: 2019/10/16 <https://bit.ly/3ztbvZ1>

وقد اجتهدت منظمات في تشفير المحارف العربية بمقاييس معيارية، أشهرها المنظمة العربية للمواصفات والمقاييس، ومنظمة الآيزو، ومجمع اليونيكود، والجمعية الأوروبية لمصنعي الحواسيب. وهذا نموذج لتشفيرة منظمة الآيزو ومجمع اليونيكود:

Codepage 1089 - Arabic - ISO 8859-6

	-0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-A	-B	-C	-D	-E	-F
0-		0001	0002	0003	0004	0005	0006	0007	0008	0009	000A	000B	000C	000D	000E	000F
1-	0010	0011	0012	0013	0014	0015	0016	0017	0018	0019	001A	001B	001C	001D	001E	001F
2-	0020	0021	0022	0023	0024	0025	0026	0027	0028	0029	002A	002B	002C	002D	002E	002F
3-	0030	0031	0032	0033	0034	0035	0036	0037	0038	0039	003A	003B	003C	003D	003E	003F
4-	0040	0041	0042	0043	0044	0045	0046	0047	0048	0049	004A	004B	004C	004D	004E	004F
5-	0050	0051	0052	0053	0054	0055	0056	0057	0058	0059	005A	005B	005C	005D	005E	005F
6-	0060	0061	0062	0063	0064	0065	0066	0067	0068	0069	006A	006B	006C	006D	006E	006F
7-	0070	0071	0072	0073	0074	0075	0076	0077	0078	0079	007A	007B	007C	007D	007E	007F
8-	0080	0081	0082	0083	0084	0085	0086	0087	0088	0089	008A	008B	008C	008D	008E	008F
9-	0090	0091	0092	0093	0094	0095	0096	0097	0098	0099	009A	009B	009C	009D	009E	009F
A-	00A0															
B-																
C-		00C1	00C2	00C3	00C4	00C5	00C6	00C7	00C8	00C9	00CA	00CB	00CC	00CD	00CE	00CF
D-	00D0	00D1	00D2	00D3	00D4	00D5	00D6	00D7	00D8	00D9	00DA	00DB	00DC	00DD	00DE	00DF
E-	00E0	00E1	00E2	00E3	00E4	00E5	00E6	00E7	00E8	00E9	00EA	00EB	00EC	00ED	00EE	00EF
F-	00F0	00F1	00F2	00F3	00F4	00F5	00F6	00F7	00F8	00F9	00FA	00FB	00FC	00FD	00FE	00FF

(الشكل 3: <https://bit.ly/3ztIglH> بتاريخ: 2019/10/16)

Codepage 1256 - Arabic Windows

	-0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-A	-B	-C	-D	-E	-F
0-		0001	0002	0003	0004	0005	0006	0007	0008	0009	000A	000B	000C	000D	000E	000F
1-	0010	0011	0012	0013	0014	0015	0016	0017	0018	0019	001A	001B	001C	001D	001E	001F
2-	0020	0021	0022	0023	0024	0025	0026	0027	0028	0029	002A	002B	002C	002D	002E	002F
3-	0030	0031	0032	0033	0034	0035	0036	0037	0038	0039	003A	003B	003C	003D	003E	003F
4-	0040	0041	0042	0043	0044	0045	0046	0047	0048	0049	004A	004B	004C	004D	004E	004F
5-	0050	0051	0052	0053	0054	0055	0056	0057	0058	0059	005A	005B	005C	005D	005E	005F
6-	0060	0061	0062	0063	0064	0065	0066	0067	0068	0069	006A	006B	006C	006D	006E	006F
7-	0070	0071	0072	0073	0074	0075	0076	0077	0078	0079	007A	007B	007C	007D	007E	007F
8-	0080	0081	0082	0083	0084	0085	0086	0087	0088	0089	008A	008B	008C	008D	008E	008F
9-	0090	0091	0092	0093	0094	0095	0096	0097	0098	0099	009A	009B	009C	009D	009E	009F
A-	00A0	00A1	00A2	00A3	00A4	00A5	00A6	00A7	00A8	00A9	00AA	00AB	00AC	00AD	00AE	00AF
B-	00B0	00B1	00B2	00B3	00B4	00B5	00B6	00B7	00B8	00B9	00BA	00BB	00BC	00BD	00BE	00BF
C-	00C0	00C1	00C2	00C3	00C4	00C5	00C6	00C7	00C8	00C9	00CA	00CB	00CC	00CD	00CE	00CF
D-	00D0	00D1	00D2	00D3	00D4	00D5	00D6	00D7	00D8	00D9	00DA	00DB	00DC	00DD	00DE	00DF
E-	00E0	00E1	00E2	00E3	00E4	00E5	00E6	00E7	00E8	00E9	00EA	00EB	00EC	00ED	00EE	00EF
F-	00F0	00F1	00F2	00F3	00F4	00F5	00F6	00F7	00F8	00F9	00FA	00FB	00FC	00FD	00FE	00FF

(الشكل 4: <https://bit.ly/3ND3ZPP> بتاريخ: 2019/10/16)

2- المدقق الإملائي وتدقيق النصوص

1-2 التعريف

التدقيق الإملائي أو التدقيق اللغوي، أو المراجعة اللغوية، أو التصحيح اللغوي.... مصطلحات تطلق على عملية دراسة الحاسوب للنصوص اللغوية بقصد معالجة الأخطاء التي وقعت فيها، وذلك حسب المدقق الإملائي الذي هو " تطبيق برمجي، يستكشف الأخطاء الإملائية الواردة في النصوص المحررة والمكتوبة رقمياً، ويساعد على تصويب هذه الأخطاء عبر اقتراح مجموعة من احتمالات الصواب " (السعيد، د.ت، ص2) ولكن "يقتصر عمل المدققات الإملائية المتاحة للعربية على الشكل الخارجي للكلمة" (السابق، ص3).

ويعتبر التدقيق الإملائي مهماً للترجمة الحاسوبية، وللمعاملات الصرفية والتركيبية، إذ تقوم الخوارزميات اللسانية عليهما. ف"الهدف المتوخى من التدقيق الإملائي، إنما هو التعرف الآلي على بنية الكلمة من خلال القاعدة اللسانية، لا من خلال المعجم الصوري المخزن سلفاً في الكفاية الحاسوبية" (مهديوي، ص227).

وقد ذكر زروق وكبداني، أن جهد محمد زبير في برنامجه "الدؤلي" (Duali) وسمير في برنامجه "بغداد" (Baghdad) أول محاولة لوضع تدقيق إملائي للحاسوب، بعد سنة 2006. ثم تلتها برامج هانسبل (Hunspell) وأسبل (Aspell) اللذين اعتمدا على يونيكود للتعامل مع النصوص العربية. وكان القاموس الوحيد الذي زود به الحاسوب حينذاك هو تيم بوكولتر (Tim Buckwalter)، الذي تميز بنسبة كبيرة من الأخطاء الإملائية (2008، ص3). غير أنه قد تأكد عند المعتز بالله السعيد، أنه قد تضمن ميكروسوفت أوفيس من سنة 1990 برنامجاً للتدقيق الإملائي للغة العربية (2019).

2-2 المدونة

في كل برنامج مدقق لغوي مدونة، وتحتوي مدونة اللغة العربية في الحاسوب على مجموعتين من النصوص، تضم الأولى عدداً كبيراً من الكلمات من مختلف المصادر الخالية من الأخطاء، وتضم الثانية كلمات من نصوص مكتوبة يدوياً، سُدت فيها الأخطاء الإملائية في الكلمة الواحدة، أو إضافة فراغ وسط الكلمة، أو حذف فراغ بين كلمتين (السويلم وآخرون، 2018، ص36).

وتعتمد المصححات الإملائية على قاموس كبير من الكلمات تقارنها بالنصوص التي تتعامل معها، على أن "تعتمد على التحليل الصرفي، بشكل أساسي، لأنه يهتم بشكل المفردة لا التركيب، وذلك بالمقارنة مع المعجم المصحح" (اللتيني، 2019، ص268).

وقد زُود وورد Word بقاموس للتدقيق الإملائي، يقوم بتصحيح الأخطاء الإملائية تلقائياً أحيانا، أو يخط تحت الكلمة بلون أحمر، ويعطي خيارات بديلة للكلمة التي لم يتعرف عليها، فيتم تغييرها أو حفظها في القاموس للمرات القادمة. ولأن المدقق الإملائي والنحوي المبرمج على أوفيس، ليس ممن لَم بالعربية القديمة والحديثة، فإنهما مبنيان على الظن والوهم، لا يجنب الوقوع في الخطأ (مهديوي، 2018، ص227).

ولكن مهما يكن من أمر، فإن "هذا المشروع، يهدف إلى بناء خوارزميات ذكاء اصطناعي، قادرة على تصحيح الأخطاء الإملائية في النصوص العربية، وكذلك محاولة تصحيح الأخطاء في المعاني (السويلم وآخرون، 2018، ص33).

وقد ظهرت حواسيب لا تتضمن في محتواها المدقق الإملائي للغة العربية، فكان لا بد من تنصيب مدقق إملائي على بعض برامج التشغيل، كنظام OS X، ويتم ذلك عن طريق تنزيل القاموس الخاص باللغة العربية وتنصيبه على الجهاز، وإلا تم الاستعانة ببعض البرامج والمواقع الالكترونية، المعنية بالتدقيق الإملائي للغة العربية.

3-2 مواقع وبرامج للتدقيق الإملائي:

مواقع عربية	مواقع أجنبية	برامج حاسوب
أكتب صح مدقق دقق لي	reverso / Onlinecorrection Grammarly / Grammarcheck PolishWriting Spellcheckplus	Ginger softwar 2018/ Phrasexpress/ برنامج الكتابة بدون أخطاء إملائية

وقد ظهر في الآونة الأخيرة برنامج تحت اسم "دال"، و "هو مدقق إملائي يغطي أكثر من 98% من النصوص العربية الحديثة، مبني على موارد لغوية عالية الدقة. يتميز مدقق "دال" بقدرته على التدقيق في الكلمات المحركة وغير المحركة، وتحليل وتفكيك الكلمات عن الزوائد، كأدوات النصب والجرّ، والضمائر المتصلة، واعتماده على قدرات التوليد للأفعال والأسماء، وجموع التكسير في نظامه التشغيلي، يعزّز قدرة مدقق "دال" وفعاليتها وسرعته." (بكري، 2018، د.ص)

وللقرآن الكريم في الحاسوب مدونة خاصة، فالقرآن الكريم، تختلف كتابته الإملائية عن الكتابة الإملائية للغة العربية. وتنقسم مدونة القرآن الكريم إلى قسم بكلمات غير مشكولة، وقسم بكلمات مشكولة، مع حذف المكررات والعناية بالنمط الكتابي له، ويعتبر مخالفة كتابة الكلمة على هيئتها القرآنية في المدقق الإملائي تحريفاً لا خطأً (السعيد، د.ت. ص 5). فأصبحت كتابة المصحف بالرسم العثماني ممكناً وذلك بتنصيب برنامج فونت في مجلد خطوط الحاسوب مع تحميل المصحف على صيغة وورد وسيتم إضافة المصحف إلى القاموس الإملائي والتعرف على الآيات وتصحيحها إملائياً وبخطها العثماني المعهود

ب	ب	ا	ء	د	ئ	إ	ؤ	أ	آ	أ
ج	ج	ج	ث	ث	ث	ت	ت	ت	ة	هـ
س	ز	ر	ذ	د	خ	خ	خ	ح	ح	ح
ض	ض	ض	ص	ص	ص	ش	ش	ش	س	س
غ	غ	ع	ع	ع	ظ	ظ	ظ	ط	ط	ط
ل	ل	ك	ك	ق	ق	ق	ف	ف	ف	غ
هـ	هـ	هـ	ن	ن	ن	م	م	م	ل	ل
لا	لا	لا	لا	لا	ي	ي	ي	ي	و	و

(الشكل 5: <https://bit.ly/3QbQW9B> بتاريخ: 2019/10/16)

06D1	06D2	06D3	06D4	06D5	06D6	06D7	06D8	06D9	06DA	06DB
٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢
06DC	06DD	06DE	06DF	06E0	06E1	06E2	06E3	06E4	06E5	06E6
١٣	١٤	١٥	١٦	١٧	١٨	١٩	٢٠	٢١	٢٢	٢٣
06E7	06E8	06E9	06EA	06EB	06EC	06ED	06EE	06EF	06F0	06F1
٢٤	٢٥	٢٦	٢٧	٢٨	٢٩	٣٠	٣١	٣٢	٣٣	٣٤
06F2	06F3	06F4	06F5	06F6	06F7	06F8	06F9	06FA	06FB	06FC
٣٥	٣٦	٣٧	٣٨	٣٩	٤٠	٤١	٤٢	٤٣	٤٤	٤٥
06FD	06FE	06FF	---	---	---	---	---	---	---	---
٤٦	٤٧	٤٨	---	---	---	---	---	---	---	---
---	---	---	---	---	---	---	---	---	---	---
---	---	---	---	---	---	---	---	---	---	---
---	---	---	---	---	---	---	---	---	---	---
٤٩	٥٠	٥١	٥٢	٥٣	٥٤	٥٥	٥٦	٥٧	٥٨	٥٩
٦٠	٦١	٦٢	٦٣	٦٤	٦٥	٦٦	٦٧	٦٨	٦٩	٧٠

(الشكل 6: <https://bit.ly/398NADz> بتاريخ: 2019/10/16)

ولأن نصوص القرآن محدودة، فإن المدقق الإملائي لها، لن يخرج عن دائرة القواعد القياسية لها. مما يعني القدرة على الإحاطة بالمادة كلياً، وعدم الوقوع في الخطأ إلا نادراً (السعيد، د.ت، ص4).

3. تحديثات المدقق الاملائي

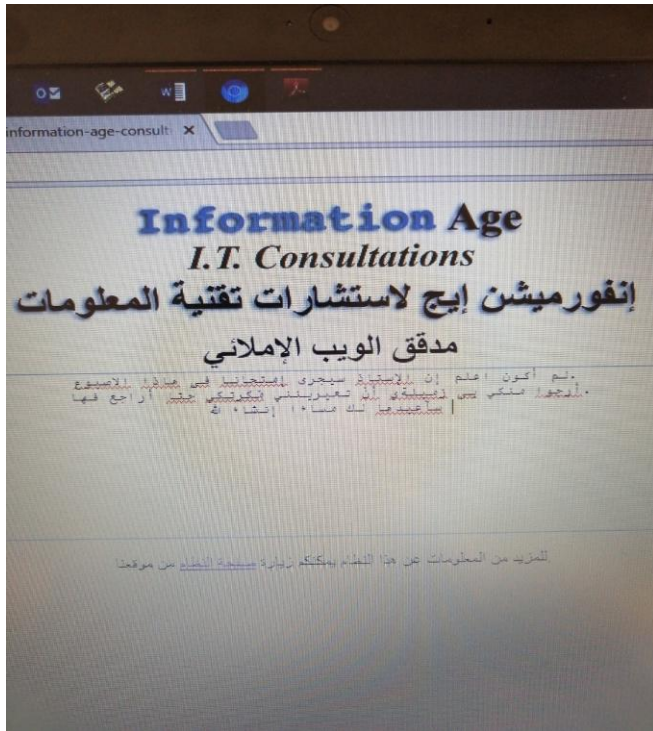
رغم التقدم العلمي الذي حظيه الحاسوب في مجال التدقيق الإملائي للغة العربية، إلا أنه لا يزال يحتاج إلى مزيد عناية. فمن التحديات التي يواجهها الحاسوب في التدقيق الإملائي، عدم القدرة على تعرّف بعض القواعد اللغوية، التي تعتبر أساساً في الفصحاة. فَمَنْ كتب صلى (المسلمين) مثلاً، لا يخطئه الحاسوب، لأنه قد تضمن في قاموسه جميع المفردات المتعلقة بلفظ مسلم، بلواحقه وسوابقه، في غير تركيب مستقيم، ولم يتمّ تحديد القاعدة النحوية المصحّحة للخطأ في القاموس الإملائي.

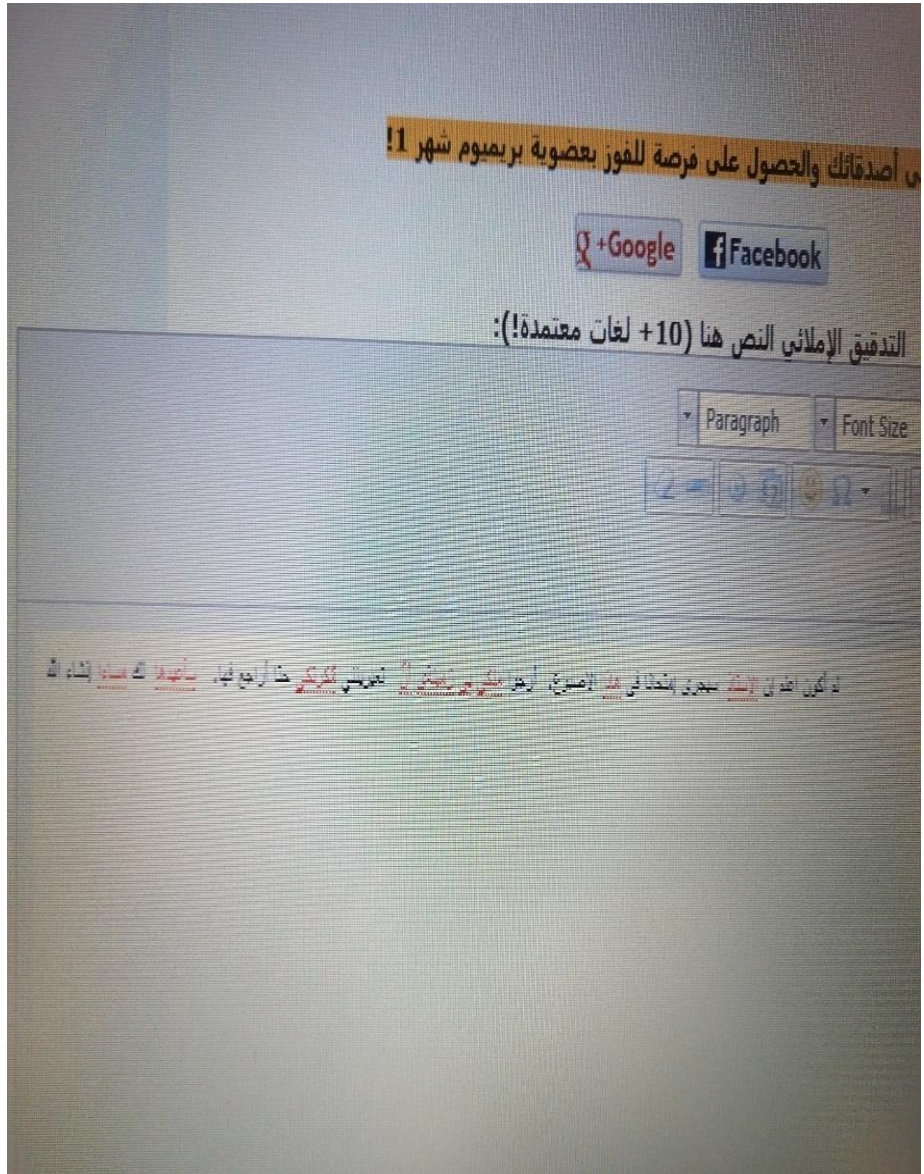
وكذلك عند إدخال بعض أسماء الأعلام المشهورة، لا يمكن للحاسوب أن يتعرف عليها بسبب عدم وجودها في القاموس الإملائي، فيخطئها، وإن كانت في الأصل صحيحة البنية، مثل كلمة (موتى).

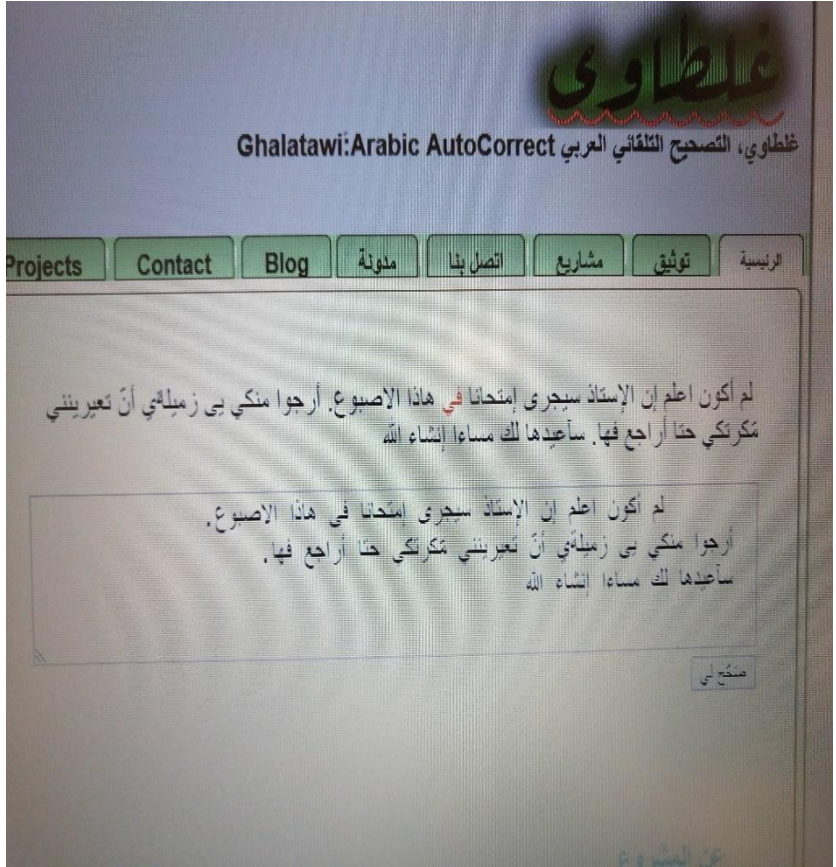
وقد يدخل على الكلمات التباس من حيث الشكل، فيعجز الحاسوب عن التفريق بينها، فكلمة (لغا/لغى) مختلفتان شكلا ومعنى، ولكن مع ذا، يحتفظ بهما الحاسوب على حالهما، في نحو لغى الرجل، ولغا الرجل.

وأحيانا عند تفعيل المصحح التلقائي، يقوم بتغيير بعض الكلمات الصحيحة البنية إلى كلمات أخرى ليست هي المقصد منها. والأمثلة التي لا يراعي فيها المدقق اللغوي التصحيح، كثيرة منها: رأيت يكتب المفيد/ حجت في الحكة / صمنا رمضان / رأيت أحمدًا/ فل هـ (فله) أن يقولي .../ يّام موسيا/ دنى مني... لذا فإنه لا بد من إعادة قراءة النص، بعد الانتهاء من التحرير، للتأكد من خلوه من الخطأ.

ولتوضيح الفكرة، تم إدخال نص بأخطاء إملائية كثيرة على بعض برامج التدقيق الإملائي، للتأكد من مدى قدرتها على كشف الأخطاء، فكانت النتائج كالتالي:

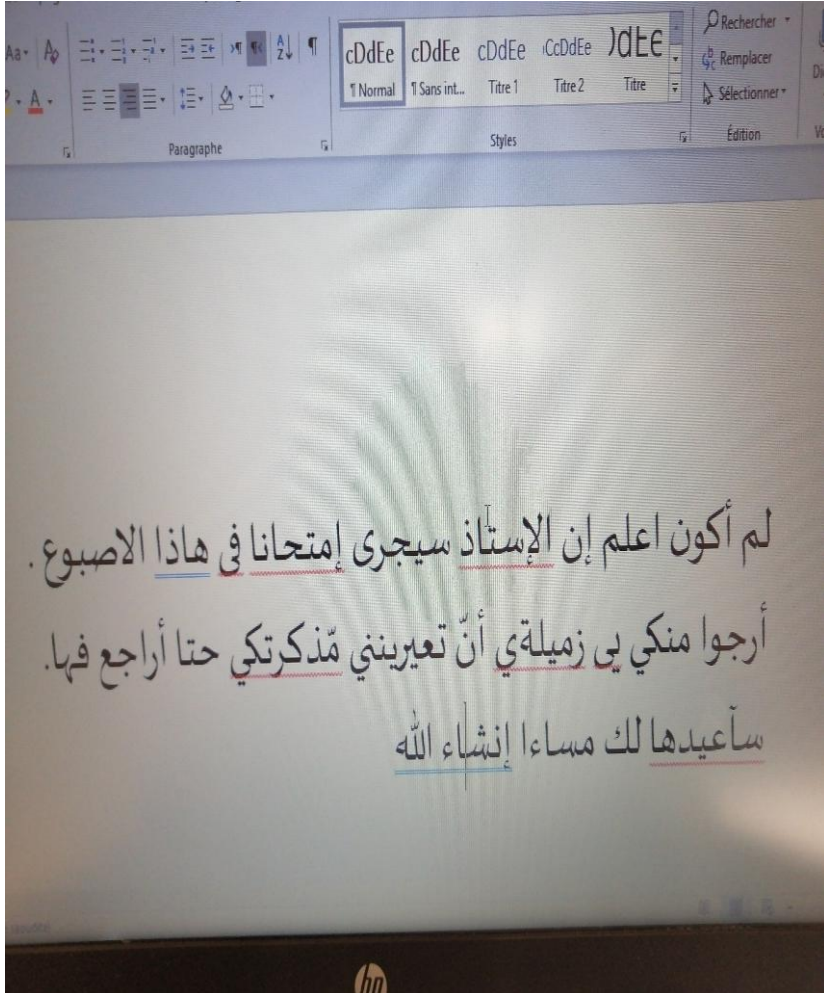




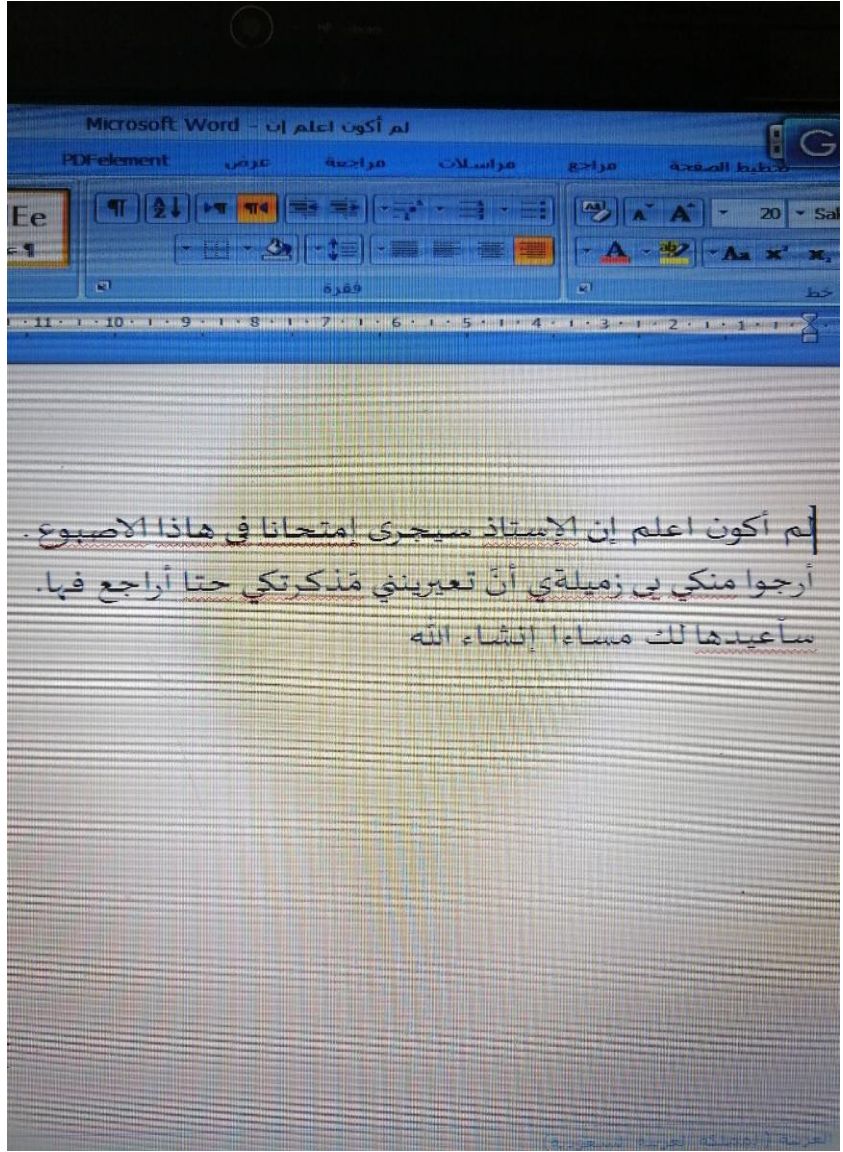


يلاحظ جليا أن برنامج "مدقق الويب" قد تعرف على الأخطاء الإملائية في النص، بنسبة 10 كلمات من أصل 25، بينما برنامج "التدقيق الإملائي" تعرف على 9 كلمات، وبرنامج "غلطاوي" خطأ كلمة واحدة فقط من أصل 25، ما يعني أن المدقق الإملائي تتفاوت في عملها، وأن اللغة العربية تحتاج إلى عناية أكبر من قبل اللسانين الحاسوبيين. ولكن مع ذا، يظهر المدقق الإملائي في برنامج وورد، بشكل أقدر على التعامل مع النصوص العربية من البرامج، ولكن ذلك لا يعني بالضرورة أنه يصحح جميع الأخطاء. فبعد وضع نفس النص السابق في وورد، ظهرت نتائجه على النحو التالي:

وورد 365



وورد 2007



هناك اختلاف كبير بين البرنامجين، حيث إن المدقق الإملائي فيهما على تفاوت في النتائج، فمن حيث التحديد، يلاحظ أن وورد 365 يعتبر بعض الأخطاء نحوية فيخط تحتها خطا أزرق، في حين أن وورد 2007 يعتبر جميع الأخطاء الواردة في النص أخطاء إملائية. والجدول التالي يفصل الفرق بينهما في الدقة:

العرف على الخطأ	مقترح وورد			الكلمة
	2007	العرف على الخطأ	365	
				الأستاذ
				امتحاننا
				في
				هاذا
	لم يفتح			الاصبوع
	يا		اقتراحات غير صائبة	ي
	اقتراحات بعيدة عن الصواب			منكي
	زميلتي		اقتراح قريب من الصواب (زميلي)	زميلة
				مذكرتي
	حتى			حنا
				سأعيدها
				إنشاء الله

ومع أن البرنامجين قد تعرفا على بعض الأخطاء الإملائية في النص السابق، إلا أن الكلمات التالية مع كونها خطأ فاضحا، لم يتعرفا عليها (أكون – اعلم- سيجري- أرجوا- أن تعيريني- فيها- مساء) وقد يكون السبب في ذلك، عدم وجود قاعدة بيانات تخطئها، أو أن القاموس الإملائي فيهما، قد تضمن الكلمات صحيحة.

الخاتمة

على الرغم من أن المدقق الإملائي للعربية قد حاز مكانة في الحاسوب، إلا أنه لا يزال يحتاج إلى مدونات تكميلية لما لها من أهمية في سد نقص موارد اللغة العربية التي يعتمد عليها لتدقيق وتصحيح الأخطاء اللغوية، وإن كان لا ينكر أن نسبة معطيات المدقق الإملائي العربي للحاسوب كبيرة، إلا أن الأمر يحتاج إلى إعادة مراجعة النصوص مرة تلو مرة، فمن شأنها أن تعين على تنقيحها من الشوائب التي قد تطرأ عليها بتقصير الحاسوب في تدقيقه لها. لأن الحاسوب مهما زود بذكاء اصطناعي، فإنه لن يضاهي الذكاء البشري في تعامله مع اللغة.

قائمة المصادر والمراجع

1. أغبر، بسام (2017). علم اللسانيات الحاسوبية.. تعريفه ومجالاته، متاح على الرابط التالي: <https://bit.ly/31tN219>
2. بكري، محمد (2018). مدقق "دال" الإملائي – الموارد الرقمية للغة العربية، متاح على الرابط التالي: <https://bit.ly/33UDC0z>
3. زروق، طه وكبداني، محمد (د.ت). المدقق الإملائي العربي الحر لمشروع آيسبل، الجزائر: جامعة أوبكر بلقايد
4. زكي، محمد (1996). الحروف العربية والحاسوب، (د.م): مجمع اللغة العربية الأردني
5. السعيد، المعتز بالله (2019). جلسة في مكتبه بتاريخ، 2019/10/16م، الدوحة – قطر: معهد الدوحة للدراسات العليا
6. السعيد، المعتز بالله (د.ت). نحو بناء مدقق إملائي للنص القرآني المطبوع رقمياً "معالجة لغوية حاسوبية" متاح على الرابط التالي: <https://scholar.cu.edu.eg/?q=moataz/files/moataz-madinah-02.pdf>
7. السويلم، عبد العزيز، وآخرون (2018). جهود مدينة الملك عبد العزيز للعلوم والتقنية في مجال تعريب العلوم، ورقة بحثية نشرت في: مؤتمر التعريب الثالث عشر. (التعريب وتوطين العلوم والتقنية)، الجزء الثاني، الرياض-المملكة العربية السعودية: معهد الملك عبد الله للترجمة والتعريب، (ص.ص. 8-40)،
8. علي، نبيل (1994). العرب وعصر المعلومات، الكويت: عالم المعرفة
9. اللتيني، يحيى (2019). أسس صناعة المعجم المحوسب، عمان: دار كنوز المعرفة للنشر والتوزيع
10. مهديوي، عمر (2018). مصطلحات الهندسة اللغوية بين الترجمة والتعريب: نحو بناء معجم موحد، ورقة بحثية نشرت في: مؤتمر التعريب الثالث عشر. (التعريب وتوطين العلوم والتقنية)، الجزء الثاني، الرياض-المملكة العربية السعودية: معهد الملك عبد الله للترجمة والتعريب، (ص.ص. 219-252)