المدقّق والمعرِب والمشكِّل الآلي

لقد جمعنا هذه المكونات تحت عنوان واحد، وذلك لما بينها من تكامل وترابط برمجي ولساني يحدد الهدف العام منها جميعا. **فهي كلها تعتمد المعاجم الإلكترونية والمكانز** التي تمدها بالمادة اللغوية المعيارية المستخدمة في كل تدقيق أو تصحيح أو إعراب أو ضبط بالشكل. كما لا يمكن أن نتصور وجود مدقق إملائي—مثلا- لا **يقوم على المحللات الصوتية والصرفية والنحوية**، لأن التدقيق المقصود هو تعرف الحاسوب على بنية الكلمة العربية من خلال القواعد التي يضعها اللسانيون، لا من خلال معجم الكلمات التي يكون مزودا بها سلفا.

1-المدقق الآلي:

يعد المدقق الآلي -أو المصحّح الاملائي والنحوي- من الأدوات الإجرائية التي تضع النصوص تحت مراقبة دقيقة مما يساعد على تلافي الأخطاء الشائعة، وبخاصة الأخطاء الاملائية .

وعلى هذا الأساس فإن المدقق الآلي يحدد ويبرز:

أ - صور وأنماط الأخطاء الإملائية والنحوية الشائعة في النصوص، وعلى سبيل المثال الأخطاء المتمثلة في الخلط بين همزتي القطع والوصل، والخطأ في رسم الهمزات المتوسطة والمتطرفة، وكذلك الخلط بين الهمزة والمدة، والتاء المربوطة والهاء، والألف المقصورة والممدودة، كما يقوم المصحح الآلي بتصويب الأخطاء الناتجة عن وجود حرف ناقص أو حرف خطأ، أو مكرر، أو زائد، وغير ذلك من الأخطاء الإملائية في كتابة اللغة العربية.

ب — إنجاز إحصائيات دقيقة وشاملة لأفضل اقتراحات التصويب.

ج — إضافة المتلازمات السياقية مما يعزز دقة التصويب من خلال الموارد التالية:

— القواعد اللغوية الموضوعة لتصويب الأخطاء الإملائية والنحوية في المكنز .

— قاعدة بيانات معجمية تحتوي على جميع التصنيفات النحوية والدلالية لعناصر الجملة والتعبيرات الاصطلاحية.

- قاعدة بيانات تحتوي على تعبيرات المسكوكة .

يمكن أن نجد المدققات الالية في برامج معالجة النصوص أو أي برامج أخرى مثل برنامج ووردword ، ويمكن أن تستخدم مباشرة وتكون مستقلة مثل برنامج غلطاوي.

ينقسم المصحح الآلي الى برنامجين هما:

1-1- المدقق الإملائيّ:

يحلّل هذا البرنامج النّصوص كلمة بعد أخرى ثم يقارنها بقاعدة بيانات محفوظة في ذاكرته لكلّ الكلمات الممكنة. وأما الكلمات التي لا يعرفها فإنه يعتبرها خاطئة أو محلّ شكّ فيقوم باقتراح البديل أو يترك الخيار للمستعمل في استبدالها أو المحافظة عليها وإدخالها في قاموس المستعمل.

يمكن أن يحلّل المدقّق الإملائيّ النّصوص المشكولة أو غير المشكولة، إلاّ أنّ مدقّقات النّصوص المشكولة لم تر النّور لأنّها تفترض محلّلات نحويّة جيّدة لا تعرفها الصّناعة اللغّويّة إلى اليوم. بل إنّ طريقة العرب في الكتابة (عدم رسم الهمزة وإهمال الشدّة، مثلا) تجبر برامج التّدقيق الإملائيّ على إتاحة الاختيار بين إمكان التّدقيق الإملائيّ الصّارم والتّدقيق الإملائيّ المتساهل.

2-1- المدقق النحوي:

وفي سياق التصحيح الآلي للأخطاء يأتي المدقق النحوي، الذي أصبح بإمكانه القيام بتصحيح الأخطاء النحوية التي نجدها شائعة في العديد من الكتابات الصحفية والإنشائية وغيرها؛ مثل تصحيح إعراب الكلمات التي تلزم حالة الرفع والجر، والأفعال التي تلزم حالة النصب والجزم، وكذلك أخطاء الممنوع من الصرف والمفعول المطلق، وعدم المطابقة النحوية بين الصفة والموصوف والمعطوف والمعطوف عليه، وأخطاء كتابة الأعداد البسيطة والمركبة والمعطوفة وتمييزها، وأخطاء الجملة الفعلية والاسمية، وغيرها من الأخطاء النحوية الأخرى في اللغة العربية.

يعرف المدقّق النحوي بأنه آلية برمجية تتعامل مع النص على مستويي الكلمة والجملة، وتتلخص وظيفته في مراجعة صحة الكلمة إعرابيا، وصحة الجملة تركيبا، ومن ثم اقتراح بدائل نحوية صحيحة للكلمة الخاطئة، أو اقتراح تعديل بديل لأجل الوصول إلى الصواب النحوي.

يعتمد المدقق النحوي على المحلل الصرفي، الذي يميز جذع الكلمة عن سوابقها ولواحقها، وذلك ما يسمح للمدقق النحويبالتعرف على الأخطاء واكتشافها؛ فمثلا حروف الجر المتصلة بالكلمة (ك-ب- ل) تجر الأسماء التي تدخل عليها.

وتختلف علامات الجر باختلاف الأسماء، ما بين الكسرة في حال المفرد، وما يجمع بالألف والتاء والفتحة في حال الممنوع من الصرف.

مثلا: عند معالجة المدقق النحوي لكلمة (بالوالدون) سيقوم المحلل الصرفي بتحليلها على النحو التالي: بال (سابقة) = (حرب جر) + آل (التعريفية)

والد = اسم

ـان (لاحقة) = علامة رفع المثنى.

لكن السابقة (بال – حرف جر – أل التعريفية) لا تتوافق نحويا مع اللاحقة (أن= علامة رفع المثنى)؛ إذ كيف يجتمع حرف الجر (بـ) مع علامة الرفع (ون) في (بالوالدون) ومن ثم يبدأ المدقق النحوي باقتراح احتمالات الصواب للكلمة، وعلى رأسها تصويب إعراب الكلمة إلى (بالوالدين).

2–المعرِب:

هو برنامج يقوم بإعراب الجمل أو النصوص المدخلة، عن طريق تمييز كل كلمة بالوظيفة النحوية التي تقوم بها داخل الجملة.

وقد يتعدى المعرب وظيفته الرئيسية هذه ليقوم بتشكيل الجملة (ضبطها بالشكل) فيصير مشكّلا، وذلك لأن معرفة الوظيفة النحوية للكلمة يؤدي مباشرة إلى ضبط آخر بالشكل.

إن المعرب عبارة عن مجموعة من قواعد اللغة العربية، يقوم بإعراب الجملة العربية سواء كانت فعلية أو اسمية بعد مرورها على مجموعة من القواعد الإعرابية الأساسية، حيث يقوم باختيار القالب (النسق) المناسب للجملة، وذلك لتشكيلها بالعلامات المناسبة، يتم إعراب الجملة المدخلة، بعد مرورها على المحلل النحوي، الذي يحلل الجملة إلى كلمات، ليختار النسق المناسب لها، ومن ثم إعرابها.

والخطوة الموالية هي إخضاع الجملة للمحلل الصرفي الذي يحلل كلمات الجملة إلى حروف، وذلك لتحديد ظهور أو تقدير الحركات على أواخرها.

1– المشكّل الآلي

هو برنامج يقوم بوضع علامات الضبط بالشكل على الكلمة حسب ما يناسبها لفظاً .

يفترض أن تيسّر برامج التّشكيل الآليضبط النصوص بالشكل، وهي عمليّة مُكْلفة في الوقت والجهد. لكنّ برامج التّشكيل المتوفّرة حاليّا تنقصها الدقّة لأنّها تشترط وجود محلّلات نحويّة جيّدة تعتمدها

للتعرّف على وظائف كلّ كلمة في سياق النّص، وهو ما لم يتوفّر بعد؛ إذ يعتبر المشكل الآلي مهمة صعبة عند تنفيذه؛ لأنه يحتاج إلى أبحاث عميقة في اللغة العربية. كما يعتمد على وجود المحللين الصرفي والنحوي وفي بعض الحالات المحلل الدلالي أيضا للوصول إلى اختيار التشكيل المناسب من بين عدّة حالات ممكنة.

ومع ذلك تظل برامج التّشكيل مهمّة لاعتماد النّصوص المشكولة بكثرة في تدريس الصّغار وكذلك في برامج القراءة الآلية للنّصوص ، إذ انّ النّصّ غير المشكول يثير في حدّ ذاته في بعض الأحيان إشكالات حقيقية.

وتوجد في الوطن العربيّ ثلاثة برامج معروفة تسمح بالتّشكيل هي على التّوالي: برنامج «الخليل للتّشكيل الآليّ» – وقد أنتجته جامعة محمّد الأوّل بالمغرب وهو مجانيّ – وبرنامج «المشكّل الآليّ» لشركة صخر، وبرنامج «مشكال» للتّشكيل الآليّ .وهو برنامج مفتوح المصدر قابل للتّطوير من انتاج جزائري لصاحبه طه زروقي، هذا الأخير سنحاول التعرف عليه فيما يلي:

مشكال برنامج لتشكيل النصوص العربية آليا لاستعمالها في القراءة والتعليم وفك الغموض عنها وهو برنامج مفتوح المصدر، مما يعطي للمستخدم والمطوّر حريّة استعماله، والهدف منه توفير اللبنات الأساسية لبناء نظم لمعالجة اللغة العربية، وتطبيقات مختلفة لها.

أهم ميزة لمشكال أنه يقترح تشكيلا آليا للنص، ومن ثمّ يمكن للمستخدم أن يصحح التشكيل، إذا وجد به خطأ، أو لم يكن المقصود كذلك.

عينة من مشكال

إِنَّ خَوْفَ الْمُسْتَبِدِّ مِنْ نقمة رَعِيَّته أَكْثَرُ مِنْ خَوْفِهِمْ مِنْ بَأسه ؛ لأَنَّ خَوْفَهُ يَنْشَأُ عَنْ عِلْمه بِمَا يَسْتَحِقُّهُ مِنْهُمْ ، وَخَوْفُهُمْ نَاشِئٌ عَنْ جَهْلٍ ؛ وَخَوَّفَهُ عَنْ عَجْزٍ حَقِيقِيٍّ فِيه ، وَخَوَّفَهُمْ عَنْ تَوَهُّمْ التَّخَاذُلِ فَقَطُّ ؛ وَخَوَّفَهُ عَلَى فَقْدِ حَيَاته وَسُلْطَانَهُ ، وَخَوَّفَهُمْ عَلَى لَقِيِّمَاتٍ مِنَ النَّبَاتِ وَعَلَى وَطَنٍ يَأْلَفُونَ غَيْرَه فِي أَيَّامٍ ؛ وَخَوَّفَهُ عَلَى كُلِّ شَيْءٍ تَحْتَ سَمَاءِ مَلِكِه ، وَخَوَّفَهُمْ عَلَى حَيَاةٍ تَعِيسَةٍ فَقَطُّ (. الْكَوَاكِبِيُّ)

مشكال النصوص العربية

مراجع المحاضرة:

1- عبد الغني أبو العزم: المعالجة الآلية للغة العربية برامج صخر أنموذجا (ديسمبر 2017) http://www.aljabriabed.net

2- عمرو جمعة: معالجة اللغة العربية آليا دراسة حاسوبية لظاهرة اللبس اللغوي في العربية المعاصرة الناشر هيئة البو ظبي للسياحة. الطبعة الاولى2015.