مدخل إلى اللسانيات الحاسوبية

أ. د. منصور بن محمد الغامدي
د. عبدالعزيز بن عبدالله المهيوبي
أ. أحمد روبي محمد عبدالرحمن
د. إشراق علي أحمد الرفاعي
د. صلاح راشد الناجم
د. وليد بن عبدالله الصانع

تحرير: د. عبدالله بن يحيى الفيفي

مباحث لغوية ٣٠

مركز الملك عبدالله بن عبدالعزيز الدولي
لخدمة اللغة العربية
King Abdullah Bin Abdulaziz Int'l Center for
The Arabic Language

www.alukah.net إهداء من شبكة الألوكة

مدخل إلى اللسانيات الحاسوبية

تأليف:

أ. د. منصور بن محمد الغامدي

د. عبدالعزيز بن عبدالله المهيوبي

أ. أحمد روبي محمد عبدالرحمن

د. إشراق علي أحمد الرفاعي

د. صلاح راشد الناجم

د. وليد بن عبدالله الصانع

تحرير :

د. عبدالله بن يحيى الفيفي

مركز الملك عبدالله بن عبدالعزيز الدولي
لخدمة اللغة العربية
King Abdullah Bin Abdulaziz Int'l Center for
The Arabic Language

الجديد و الطريف NEW & EXCLUSIVE

www.alukah.net إهداء من شبكة الألوكة

مدخل إلى اللسانيات الحاسوبية

الطبعة الأولى
١٤٣٨ هـ - ٢٠١٧م

جميع الحقوق محفوظة

المملكة العربية السعودية - الرياض
ص.ب ١٢٥٠٠ الرياض ١١٤٧٣
هاتف: ٠٠٩٦٦١١٢٥٨٧٢٦٨ - ٠٠٩٦٦١١٢٥٨١٠٨٢
البريد الإليكتروني: nashr@kaica.org.sa

ح/ مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية، ١٤٣٨هـ.
فهرسة مكتبة الملك فهد الوطنية أثناء النشر
الفيفي ، عبدالله بن يحيى
مدخل إلى اللسانيات الحاسوبية./ عبدالله بن يحيى الفيفي - الرياض، ١٤٣٨هـ
..ص؛ .. سم
ردمك: ٠-٢-٩٠٨٩٥-٦٠٣-٩٧٨
١- اللغة العربية-معالجة البيانات أ. العنوان
ديوي ٤١٨,٢٨٥ ٤١٨٣/ ١٤٣٨
رقم الإيداع: ٤١٨٣/ ١٤٣٨
ردمك: ٠-٢-٩٠٨٩٥-٦٠٣-٩٧٨

التصميم والإخراج

دار وجوه للنشر والتوزيع
Wojooh Publishing & Distribution House
www.wojoooh.com
المملكة العربية السعودية - الرياض
الهاتف: 4562410 الفاكس: 4561675
للتواصل والنشر:
info@wojoooh.com

لايسمح بإعادة إصدار هذا الكتاب، أو نقله في أي شكل أو وسيلة، سواء أكانت إلكترونية أم يدوية أم ميكانيكية، بما في ذلك جميع أنواع تصوير المستندات بالنسخ، أو التسجيل أو التخزين، أو أنظمة الاسترجاع، دون إذن خطي من المركز بذلك.

www.alukah.net إهداء من شبكة الألوكة

www.alukah.net إهداء من شبكة الألوكة

مقدمة المحرر [1]

الحمد لله رب العالمين، والصلاة والسلام على أشرف المرسلين، نبينا محمد وعلى آله وصحابته أجمعين، وبعد:

يعد مجال اللسانيات الحاسوبية (Computational Linguistics) أحد العلوم البينية (Interdisciplinary) التي تقع بين علمين مستقلين، وذلك لاتصاله بعلم اللغويات أو اللسانيات من جهة، وبعلم الحاسب الآلي من جهة أخرى. ويرى مارتن كي (Kay, ٢٠٠٣)[2] أن اللسانيات الحاسوبية قد برزت إلى حيز الوجود خلسة وبخجل، وأن بدايتها كانت في عام ١٩٤٩م عندما كتب وارن ويفر مذكرته الشهيرة التي يشير فيها إلى إمكانية بناء نظام للترجمة الآلية. ثم تلا ذلك عقد أول مؤتمر للترجمة الآلية في معهد

١- عبدالله بن يحيى الفيفي: أستاذ اللغويات الحاسوبية المساعد في جامعة الإمام محمد بن سعود الإسلامية في الرياض. درس البكالوريوس في اللغة العربية في جامعة الملك خالد في أبها، والماجستير في تعليم اللغة بمساعدة الحاسب في قسم اللغويات في جامعة Essex، والدكتوراه في اللغويات الحاسوبية في قسم الحاسب الآلي في جامعة Leeds، وكلاهما في بريطانيا. له عدة أبحاث منشورة حول تقنيات معالجة اللغة العربية آلياً، والمدونات اللغوية وبرامجها الحاسوبية، وكذلك مدونات المتعلمين، والمعاجم الحاسوبية لمتعلمي اللغة العربية، إضافة إلى مشاركته في تأليف بعض الكتب المتخصصة في اللسانيات الحاسوبية، والمدونات اللغوية وتطبيقاتها. عمل محكماً لدى عدد من الدوريات العلمية والمؤتمرات الدولية. أنشأ المدونة اللغوية لمتعلمي اللغة العربية، وأسس فريق معجم المفردات الشائعة لمتعلمي اللغة العربية "شائع".

2- Kay, Martin (2003) Introduction. In: Mitkov, Ruslan (Ed.), The Oxford Handbook of Computational Lnguistics. New York: Oxford University Press.

www.alukah.net إهداء من شبكة الألوكة

ماساتشوستس للتكنولوجيا (MIT) في ١٩٥٢م، ثم صدرت مجلة علمية بعنوان الترجمة الآلية في ١٩٥٤م. أما مصطلح اللسانيات الحاسوبية نفسه فقد بدأ استعماله في منتصف الستينات (1960s)، ويُرجح أن ديفيد هيز (David Hays) هو أول من أطلق هذا المسمى على هذا المجال عندما كان عضواً في اللجنة الاستشارية لمعالجة اللغة آلياً في الأكاديمية الوطنية للعلوم في الولايات المتحدة الأمريكية. ثم أتى بعد ذلك عدد من المتخصصين الذين كان لهم دور في ظهور هذا المجال مثل نعوم تشومسكي (Noam Chomsky) وجون كوك (John Cocke)، وغيرهم. واليوم يعد هذا التخصص من التخصصات ذات الأهمية المتزايدة لما له من دور كبير في التطور الحاصل في مجال معالجة اللغة الطبيعية (Natural Language Processing) والذكاء الاصطناعي (Artificial Intelligence)، إذ بات أحد الأسس المهمة في صناعة التطبيقات الحاسوبية المستعملة في كثير من مظاهر الحياة اليومية.

ويُعرف نيوقس اللسانيات الحاسوبية بأنها فرع عن علمي اللغة والحاسب، يهدف إلى تصميم نماذج رياضية للتراكيب اللغوية؛ للتمكن من معالجة اللغة آلياً عن طريق الحاسب، كما يعرفه من وجهة نظر لغوية على أنه تشكيل للنظريات والنماذج اللغوية أو تنفيذها على الآلة، ويرى أنه بإمكاننا النظر إليه على أنه وسيلة لتطوير نظريات لغوية جديدة بمساعدة الحاسب (Nugues, 2006) [1].

ولقد شهد البحث في مجال اللسانيات الحاسوبية تقدماً متسارعاً في السنوات القليلة الماضية، مما ساعد على بروز تطبيقات عملية استفادت من نتائج تلك الأبحاث بشكل مباشر وفي مجالات شتى، لعل من أبرزها تطبيقات التخاطب مع الآلة المسماة بتطبيقات المساعد الشخصي الذكي (Intelligent personal assistant) والتي نرى انتشارها بين أيدينا مثل: سيري (Siri) من شركة أبل (Apple)، وجوجل ناو (Google Now) من شركة جوجل (Google)، وكورتانا (Cortana) من شركة مايكروسوفت (Microsoft)، وأمازون إيكو (Amazon Echo) من شركة أمازون (Amazon)، وعشرات الأنظمة المشابهة التي تجمع عدداً من مستويات المعالجة اللغوية في تطبيق واحد. ومن هنا تبرز أهمية وجود مدخل إلى اللسانيات الحاسوبية باللغة العربية،

1- Nugues, Pierre M. (2006) An Introduction to Language Processing with Perl and Prolog. Berlin: Springer-Verlag.

www.alukah.net إهداء من شبكة الألوكة

لتعريف القارئ العربي بهذا المجال وببعض فروعه وتطبيقاته، وليكون تمهيداً لما يكتب بعده من مراجع متخصصة تتناول فروعه بتوسع أكثر. وهذا الكتاب موجه بالدرجة الأولى لطلاب الدراسات العليا في الجامعات، أو الراغبين في الاطلاع على هذا المجال من غير المتخصصين، إذ يقدم تعريفاً لعدد من مجالات اللسانيات الحاسوبية وهي:

- الصوتيات الحاسوبية Computational Phonetics.
- التحليل الصرفي Morphological Analysis.
- التحليل النحوي Syntactic Parsing.
- التحليل الدلالي Semantic Analysis.
- تحليل النصوص Texts Analytics.
- التدقيق الإملائي Spelling Checker.

ولقد حرص المشاركون في تأليف هذا الكتاب على أن يكون الطرح تعليمياً متدرجاً مع شرح المصطلحات قدر الإمكان، وتقريب المعلومات للقارئ بأمثلة واضحة تساعد على الفهم والتطبيق. وفيما يلي عرض موجز لمحتويات الكتاب اعتماداً على الملخصات التي سترد لاحقاً في بداية كل فصل من فصوله.

ففي **الفصل الأول** يتحدث منصور الغامدي عن **الصوتيات الحاسوبية**، مبتدئاً بمقدمة عامة لعلم الصوتيات، ثم يتطرق إلى الفروع الثلاثة لهذا العلم: الصوتيات النطقية، الصوتيات الأكوستية، الصوتيات السمعية. وتحت كل فرع يُورد مقدمة ثم يذكر التقنيات المتعلقة به من حيث الدراسة والبحث والتحليل. ولأن هذا العلم أساس لعدد من العلوم، فالفصل يذكر التطبيقات التقنية لعلم الصوتيات وخاصة في مجال التعرف الآلي على الكلام وتوليد الكلام آليا، والتعرف على المتحدث، مع الإشارة إلى المتطلبات التي تقوم عليها هذه التقنيات.

في **الفصل الثاني** يتحدث عبدالعزيز المهيوبي عن **التحليل الصرفي**، مبتدئاً بعرض موجز لخصائص الصرف العربي، ثم مفهوم التحليل الصرفي الآلي، وقواعد المعطيات المصاحبة للمحلل الصرفي. ينتقل بعد ذلك إلى الحديث عن مجموعة من الأسس المهمة لبناء محلل صرفي دقيق للغة العربية، ويقدّم نظرة تاريخية للتحليل الصرفي الآلي، مع استعراض مجموعة من أهم المحللات الصرفية العربية، مشيراً لأهمية التطبيقات الحاسوبية للتحليل الصرفي. ينتقل بعد ذلك إلى الحديث عن مجموعة من الضوابط

www.alukah.net إهداء من شبكة الألوكة

والمحددات التي تساعد في بناء المحللات الصرفية، مقسّماً إيّاها إلى ضوابط ومحددات شكليَّة ودلالية. ثم يستعرض المشكلات التي تواجه بناء محلل صرفي دقيق لكلمات اللغة العربية ونصوصها، وطرق عرض نتائجها، وكيفية توصيف القواعد الصرفية لبناء المحلل الصرفي الآلي. ثم يشير في عجالة إلى أسباب قصور المحللات الإنجليزية عن استيعاب خصائص اللغة العربية، متحدثاً بالتفصيل عن خطوات بناء المحلل الصرفي الآلي، ومتطلبات بنائه.

وفي **الفصل الثالث** يتحدث أحمد روبي عن **التحليل النحوي**، فيقدم رؤية شاملة عن التحليل النحوي الحاسوبي في إطار تطبيقي، محاولاً الوقوف على منطلقات التحليل النحوي (التمثيل النحوي - النظرية النحوية - المحتوى النحوي) وأدواته في صورة مبسطة، بحيث تكون مدخلا مبسطًا للقارئ العربي، يمكن من خلالها فهم الصورة العامة لإطار التحليل النحوي الحاسوبي. وسعيًا لتحقيق هذه الغاية، فإنه يقف على قوام العملية النحوية/ التركيبية ودورها في بناء التطبيقات الحاسوبية المختلفة التي تناظر الأداء الإنساني؛ فيأتي الفصل في خمسة محاور رئيسية: تتضمن مقدمة يعرض من خلالها تأصيل طرق التوصيف النحوي، ثم عرضًا لإرهاصات التحليل النحوي الحاسوبي، ثم أهمية التحليل النحوي الحاسوبي للدراسات اللغوية بصورة عامة وللغويات الحاسوبية أو معالجة اللغة الطبيعية بصورة خاصة، ويلي ذلك الخطوات الإجرائية اللازمة لبناء أية عملية تحليل نحوي حاسوبي، والتي يمكن تلخيصها في العناصر التالية على الترتيب: (النص الخام/ المدونة اللغوية - تجزئة النصوص - العنونة بالأجزاء الكلامية - الترميز بالعلاقات التركيبية)، وأخيرا يعرض الفصل بعض موارد التحليل النحوي المتاحة للغة العربية وكذلك تطبيقاته.

في **الفصل الرابع** تتحدث إشراق الرفاعي عن **التحليل الدلالي**، إذ يقدم الفصل نبذة تعريفية عن التحليل الدلالي، تشمل استعراضا لأهم المصطلحات المرتبطة بهذا المجال مثل المتضادات والمترادفات، إضافة للفرق بين المعنى الحرفي والمعنى العملي للنصوص، كما يشير الفصل إلى المنهج البحثي المستخدم في دراسة التحليل الدلالي، إضافة إلى أبرز الموارد اللغوية المتاحة مثل شبكة الكلمات العربية (Arabic WordNet). يتطرق الفصل فيما يلي ذلك إلى عدد من أبرز تطبيقات هذا المجال وهي: تحليل العواطف، وفك اللبس الدلالي، مع تعريف كل منهما، واستعراض أبرز ما أنجز فيهما من أبحاث

www.alukah.net إهداء من شبكة الألوكة

وتطبيقات. كما يتطرق الفصل إلى الحديث عن الكينونات كمفهوم مهم عند دراسة التحليل الدلالي، ويُعنى بدراسة الكلمات المجردة والعلاقات فيما بينها من حيث المعنى. يشير الجزء الأخير من الفصل إلى أبرز الجهود البحثية في مجال التحليل الدلالي فيما يخص اللغة العربية، والتي قُدمت من قبل مجموعات بحثية شهيرة حول العالم؛ حتى يتسنى للقارئ الاطلاع على المخرجات البحثية الأحدث في هذا المجال.

في **الفصل الخامس** يتحدث صلاح الناجم عن **تحليل النصوص**، فيتناول أهمية تحليل النصوص كتطبيق أساسي من تطبيقات المعالجة الحاسوبية للغة الطبيعية، وهو يساعد على اكتشاف وانتزاع معرفة هامة من نصوص حرة لا تسير وفق بنية منظمة (Unstructured Data). يشير الفصل كذلك إلى التطور الكبير في مجال البيانات الضخمة (Big Data) الذي أفرز كميات هائلة من البيانات النصية، ومنها على سبيل المثال لا الحصر مشاركات وحوارات وسائل التواصل الاجتماعي، إذ يتطلب تحليل هذه البيانات إيجاد تطبيقات ومنصات تحليلية ولغات برمجة وأدوات وخوارزميات (Algorithms) متخصصة للتعامل مع هذا الكم الهائل من البيانات النصية. وهنا تبرز أهمية تحليل النصوص كمجال بيني (Interdisciplinary) يدمج أكثر من مجال أكاديمي أهمها علم الحاسوب، اللسانيات الحاسوبية، استرجاع المعلومات (Information Retrieval)، تحليل البيانات (Data Mining)، تعلُّم الآلة (Machine Learning)، والإحصاء (Statistics). يتحدث الفصل أيضاً عن أهمية البيانات الضخمة، ومستويات ومراحل تحليل النصوص، ثم ينتقل إلى الحديث عن المعالجة الحاسوبية للنصوص وخطواتها، ثم يتناول بعض التطبيقات مثل تصنيف النصوص، وانتزاع المعلومات، وتحليل المزاج العام.

وفي **الفصل السادس** يتحدث وليد الصانع عن **التدقيق الإملائي**، فيستعرض أبرز التحديات التي تواجه مطوري المدققات الإملائية للغة العربية، إذ تعتبر اللغة العربية من اللغات المدعومة في كثير من أنظمة التشغيل وأجهزة الحاسب الآلي والبرمجيات، وقد قامت كبريات الشركات العالمية بتطوير مدققات إملائية للغة العربية. ويعد تطوير مدققات إملائية عربية تحديا يواجه مطوري هذا النوع من التطبيقات بسبب اختلاف صيغ الإملاء زماناً ومكاناً.

يُعرّج الفصل بعد ذلك على آليات اكتشاف الأخطاء الإملائية وإشكالياتها، ومن

www.alukah.net إهداء من شبكة الألوكة

ثم أبرز الطرق لتصحيح هذه الأخطاء الإملائية، كما يعطي نبذة سريعة عن بعض النظريات المتقدمة التي تستخدم في أبحاث تطوير المدققات الإملائية، وبعض المراجع الأساسية التي قد تفيد القارئ.

ختاماً، أتقدم بالشكر الوافر - بعد شكر الله عز وجل - إلى القائمين على مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية، على ما قدموا لفريق التأليف من دعم متصل وتذليل للعقبات، في سبيل خروج هذا الكتاب في أتم صورة، ليكون مرجعاً للمهتمين بهذا الميدان، وبداية للسالكين فيه من طلاب الجامعات وخصوصاً طلاب الدراسات العليا. كما أتقدم بالشكر الجزيل لجميع المشاركين في تأليف فصول هذا الكتاب الذين قدموا خلاصة فكرهم، مع ما أبدوا من التزام ودقة في العمل، كل ذلك في سبيل تحقيق غاية الكتاب، وانتظام عقد موضوعاته.

المحرر/ عبدالله بن يحيى الفيفي

الرياض

١٧ ربيع الثاني ١٤٣٨هـ

ayjfaifi@gmail.com

◇◇◇

www.alukah.net إهداء من شبكة الألوكة

الفصل الثاني: التحليل الصرفي

منذ ظهور الجيل الأول للحواسيب في عام ١٩٤٤م، وصلة الحواسيب باللغات الطبيعية تزيد وتتطور "وذلك لسبب أساسي وبسيط، وهو كون اللغة تجسيد لما هو جوهري في الإنسان، أي نشاطه الذهني بكل تجلياته، في الوقت نفسه الذي تتجه فيه الحواسيب نحو محاكاة بعض وظائف الإنسان، وقدراته الذهنية"[١]. وقد أولت الأمم المتقدمة أهمية كبيرة في عصر الرقمنة لمعالجة لغاتها الطبيعية بواسطة الحاسوب، ويُحتِّم هذا الأمر علينا دراسة لغتنا العربية محاولين توصيف قواعدها، وميكنتها بالحاسوب، مستفيدين من خصائصها في تطوير برمجيات، وبناء قواعد معطيات تساعد في معالجتها آلياً، «فمنذ الأربعينيات والمحاولات مستمرة لتحوير قواعد اللغات الطبيعية من الشكل الوصفي الأدبي إلى الشكل العلمي الدقيق، والذي يمكن برمجته حسب الإمكانيات التي يوفرها الحاسوب...، وتتوفر الآن لغات برمجة عالية المستوى تتسم بما يُطلق عليه الذكاء الاصطناعي، الذي يجعل الحاسوب يستقبل، ويحلل، ويُنفذ ما يُعطى إليه»[٢].

١- نبيل علي، اللغة العربية والحاسوب (دراسة بحثية)، تعريب، د.ط، ١٩٨٨م، ص١١٤.

٢- يُنظر: عبده ذياب العجيلي، الحاسوب واللغة العربية، منشورات جامعة اليرموك- عمادة البحث العلمي والدراسات العليا، الأردن، د ط، ١٩٩٦م، ص١٤.

www.alukah.net إهداء من شبكة الألوكة

وتمرُّ اللغات الطبيعية بمراحل لفهم الكلام، أهمها: التحليل الصرفي - وهو موضوع اهتمامنا هنا - والتحليل المعجمي، والتحليل الدلالي[1]، والتحليل النحوي[2]. والعربية لغة قابلة للمعالجة الحاسوبية» إذ تملك نظاماً خاصاً يجعلها أكثر قابلية لأن تُمثّل حاسوبياً، وتوفر اللغة العربية مجالات كثيرة لتناولها حاسوبياً على اختلاف مستوياتها، ولاسيما الصرفية؛ لأنها لغة ذات نظام دقيق تركيبياً، ودلالياً، ومعجمياً»[3]. لقد أثبتت اللغة العربية أنها من أكثر اللغات قابلية لاستخدام الحاسوب في معالجتها آلياً؛ «لأنها تجمع بين كثير من الخصائص اللغوية المشتركة مع اللغات الأخرى، فأبجديتها -مثلاً- ليست فونيمية صرفة كالإسبانية والفنلندية، حيث يقابل كلُّ حرف صوتاً واحداً، كما أنها ليست مقطعية كاليابانية، حيث رموز الأبجدية عبارة عن مقاطع يتكون كل منها من صامت (حرف) يتبعه صائت (حركة) مثل: «ما، كي، فو» فالأبجدية العربية رغم كونها فونيمية أساساً، فإنها تتضمن حروفاً ذات طبيعة مقطعية، مثل: «لا، لأ، ؤ»[4].

وتتعرض لغتنا العربية لهجوم كبير، ورَمْي بالقصور والعجز عن مواكبة التطور، وذلك بهدف تدمير اللغة التي يتعبد بها أكثر من مليار مسلم، ناهيك عن كونها الدعامة الأساسية، والمعبر الحقيقي للقوة التي تستند إليها الحضارة العربية، والتراث الإسلامي. ولعل آخر هجوم عليها هو عدم قابليتها للمعالجة الآلية باستخدام الحاسوب.

«وعلم الصرف من أهم العلوم العربية قديماً وحديثاً؛ فلا يمكن لنحوي، أو لغوي، أو معلم، أو طالب الاستغناء عنه؛ لأنه أساس العربية، وميزانها، به تتولد الكلمات[5]، وبه يتم الاشتقاق، كاسم الفاعل، واسم المفعول، والصفة المشبهة، وصيغة المبالغة

١- يستخلص المحلل الدلالي معاني الكلمات استناداً إلى سياقها، ويحدد معاني الجمل استناداً إلى ما يسبقها، وما يلحقها من جمل.

٢- يحدد المحلل النحوي صيغة الفعل في الماضي والمضارع والأمر، وكونه مؤكداً، أو غير مؤكد، ومعلوماً أو مجهولاً، ومبنياً أو معرباً، وعلامة كل منهما، ظاهرة أو مقدرة، حرفاً كانت أو حركة، ويبيّن ضمائر الرفع أو النصب التي أسند إليها الفعل، كما يبيّن المحلل النحوي حالة الاسم من حيث الإعراب والبناء، وعلامة كلّ منهما، أمّا الحرف فيبيّن المحلل سابقته، ولاحقته، ووظيفته الإعرابية، وعلامة بنائه.

٣- مازن الوعر، دراسات لسانية تطبيقية، دار طلاس، دمشق، ط١، ١٩٨٩م، ص٣٧٧.

٤- سعيد أحمد بيومي، أم اللغات -دراسة في خصائص اللغة العربية، والنهوض بها-، ط١، ٢٠٠٢م، ص١٠٥.

٥- أنواع الكلمات في اللغة العربية: جذور بدون إضافات، وجذور تضم سوابق فقط، وجذور تضم لواحق فقط، وجذور تضم سوابق ولواحق، وجذور تضم أواسط فقط، وجذور تضم أواسط وسوابق فقط، وجذور تضم أواسط ولواحق فقط، وجذور تضم أواسط وسوابق ولواحق.

www.alukah.net إهداء من شبكة الألوكة

وغيرها، وبه يُعرف الصحيح من المعتل، والمجرد من المزيد، والأوزان المختلفة، وبه تُعرف الأسماء تعريفاً وتنكيراً، وجنساً وعدّاً، ناهيك عن أنّه يقي اللسان من الوقوع في الخطأ، ويرشده إلى الصواب، ويصحّح القلم من الزّلات، وتتكئ عليه الحقول اللغوية المختلفة، الصوتية، والنحوية، والتركيبية، والمعجمية، والدلاليّة[1]».

«والصرف هو رابطة العقد لعناصر المنظومة اللغوية، فهو ركيزة الفونولوجي، ومدخل النحو، وأساس تنظيم المعجم، وفوق هذا كله فهو خط المواجهة الساخن لالتقاء مباني اللغة ومعانيها»[2]. ويُعد الصرف في اللغة العربية مصدر التوسع اللغوي بما يوفره من وسائل عديدة لتكوين كلمات جديدة، وإعادة تحليل تلك الكلمات، «كما يُعد الصرف العربي وضعاً مثاليّاً لإبراز ثنائية التحليل والتركيب[3]، وذلك نظراً لكون كلٍّ منهما صورة منعكسة من الآخر بصورة لا تتوافر في الفروع اللغوية الأخرى»[4].

وتتسم اللغة العربية باطّراد نظام صرفها، وظهور أثره داخل منظومتها اللغوية «لذا فهو مدخل أساسي لوصف النظام الشامل للغة العربية، وتفسير الكثير من ظواهرها، وتحديد أسلوب معالجتها آليّاً»[5]. ويُعد التحليل الصرفي إحدى مراحل معالجة اللغات الطبيعية آليّاً، ويدخل في الكثير من التطبيقات اللغوية، مثل: ميكنة المعاجم، وضغط النصوص، وتشكيلها، وتشفيرها، وتحليلها، وتمييز الكلام، وتوليده، وتصحيح الأخطاء الإملائية، والترجمة الآلية.

وأكاد أجزم أنَّ لمعالجة الصرف آليّاً دور حيوي في جميع الأمور المتعلقة بتناول اللغة العربية حاسوبيّاً ومعلوماتياً؛ حيث تُعد ميكنة العمليات الصرفية بالنسبة للغة العربية مدخلاً أساسياً، وقاسماً مشتركا لمعظم نظمها الآليّة، حيث يزعم الدكتور «نبيل علي» «أنَّ مدى نجاحنا في تعريب نظم المعلومات والمعارف، يتوقف بالدرجة الأولى على

١- محمود مصطفى عيسى خليل، إسناد الأفعال إلى الضمائر في ضوء اللسانيات الحاسوبي- ماجستير، كلية الدراسات العليا، الأردن، ٢٠١١م، ص٤٥.

٢- نبيل علي، اللغة العربية والحاسوب (دراسة بحثية)، تعريب، د.ط، ١٩٨٨م، ص٢٤٧.

٣- وهي الثنائية التي يوصف من خلالها كثير من الظواهر اللغوية في الوقت نفسه الذي تُعدُّ فيه أحد المفاهيم الأساسية في تصميم نظم المعالجة الآلية للغات.

٤- يُنظر: نبيل علي، اللغة العربية والحاسوب (دراسة بحثية)، تعريب، د.ط، ١٩٨٨م، ص ٢٤٧-٢٤٨.

٥- المرجع السابق، ص٢٤٨.

www.alukah.net إهداء من شبكة الألوكة

ما نستطيع أن نحققه على جبهة الصرف»[1]، ويقصد الدكتور «نبيل علي» هنا الصرف بمعناه الواسع: مبناه ومعناه، تصريفه وتركيبه، تحليله وتوليده، اطّراده وشذوذه.

وحتى نتمكن من تحليل اللغة العربية نحتاج إلى معرفة مفرداتها، وطريقة تركيب تلك المفردات في سياقات للحصول على جمل مفيدة. كذلك نحتاج إلى معرفة معاني تلك المفردات، وطرق استعمالها في الكلام.

كما أنَّ الصرف هو المسؤول عن بنية مفردات اللغة، تحليلاً وتوليداً. ويتلقى الصرف مدخلاته في اللغة العربية من ثلاثة مصادر، حيث تتسم المنظومة اللغوية بالتماسك الشديد بين عناصرها، وهذه المصادر هي: المعجم: حيث يُغذي الصرف بجذور المفردات، أو جذوعها، والدلالة: حيث تحدد المعنى الصرفي المراد صياغة الكلمة في قالبه، والنحو: حيث يعيّن الوظيفة النحوية للمفردة داخل الجملة، وحالتها الإعرابية»[2].

١. خصائص الصرف العربي[3]

تتسم اللغة العربية بخاصية الاشتقاق الصرفي المبني على أنماط الصيغ، إذ إنها تتميز بالاطّراد الصرفي المنتظم الذي أدى بالبعض إلى وصفها بالجبرية (نسبة إلى علم الجبر) بدرجة تقترب من حد الاصطناع. كما تتميز اللغة العربية بالتعالق الشديد بين مستوياتها، حيث يتعالق المستوى الصرفي مع المستوى الصوتي، فيعتمد الصرف اعتماداً كبيراً على نتائج علم الأصوات عند الحديث -مثلاً- عن الإعلال والإبدال.» كما أن النحو لا يتخذ لمعانيه مباني من أي نوع إلا ما يقدمه له الصرف من المباني، وهذا هو السبب الذي جعل النحاة يجدون في أغلب الأحيان أنه من الصعب أن يفصلوا بين الصرف والنحو، فيعالجون كلاً منهما علاجاً منفصلاً، ومن هنا جاءت متون القواعد مشتملة على مزيج من هذا وذاك، يصعب معه إعطاء ما للنحو للنحو، وما للصرف للصرف»[4]. وقد ارتبطت عملية الكشف على المعاجم بعملية التحليل الصرفي، علاوة على ذلك فالتماسك المعجمي ممثلاً في الاشتقاق، وكذلك في العلاقات الدلالية بين

١- المرجع السابق، ص٢٩٧.

٢- المرجع السابق ١٩٨٨م، ص٢٩.

٣- عبدالعزيز بن عبدالله المهيوبي، بناء خوارزمية حاسوبية لتوليد الأفعال في اللغة العربية وتصريفها - دكتوراه، معهد تعليم اللغة العربية - جامعة الإمام محمد بن سعود الإسلامية، ١٤٣٦هـ، ص١١٨.

٤- تمام حسّان، اللغة العربية معناها ومبناها، دار الثقافة، المغرب، ١٩٩٤م، ص١٧٨.

www.alukah.net إهداء من شبكة الألوكة

المفردات المشتركة في الصيغة الصرفية الواحدة، هو نتيجة طبيعية لشدة التماسك بين الصرف والمعجم.

وسنركِّز عند حديثنا عن خصائص الصرف العربي على تلك النواحي ذات الصلة بمعالجته آلياً، حيث تعد معالجة الصرف العربي آلياً مطلباً أساسياً لميكنة عمليات تحليل النصوص المكتوبة والمنطوقة، وفهمها وتوليدها، علاوة على أنّه أساس لا غنى عنه لميكنة المعاجم واسترجاع المعلومات وتحليل مضمون النصوص. حيث يتميز الصرف العربي بعدة خصائص من أهمها:

١- وضوح مسار عملية الاشتقاق (الانتقال من الجذور إلى المشتقات الفعلية).

٢- اطِّراد التصريف في العربية، باستثناء حالات نادرة.

٣- ميل الصرف العربي لتركيب الكلمات بالإضافة، وكرهه لتكوين الكلمات من خلال المزج والاختصار.

٤- انتظام بنية الكلمة العربية لثبوت رتبة عناصرها (الصرف-نحوية).

٥- شدة التداخل بين الصرف، والفونولوجي من حيث تعدد قواعد الإبدال والإعلال، وعمليات التغيير (الصرف-صوتية) الأخرى.

٦- قلة عدد جذور الأفعال وكثرة عدد فروعها.

٧- أن الاشتقاق في العربية مبني على الأنماط الصرفية [1]، حيث تتعدد هذه الأنماط مستخدمة عدداً قليلاً من حروف الزيادة.

٨- محورية مفهوم الجذر في العربية كعنصر ربط معجمي ودلالي.

٢. الحاسوب ومحاكاة تفكير الإنسان

سعى علماء اللسانيات الحاسوبية إلى بناء تطبيقات وأدوات للتحليل الصرفي الحاسوبي؛ بهدف محاكاة التفكير الإنساني في تحليل كلمات ونصوص اللغات الطبيعية من النواحي الإدراكية والنفسية. ولكن هل تمكَّنوا من ذلك؟ الجواب: لا، لأنَّ علماء اللسانيات الحاسوبية لم يتمكنوا من بناء محلل صرفي متكامل يحاكي تفكير الإنسان،

١- النمط الصرفي: عبارة عن قالب يشمل الحركات وحروف الزيادة ومواضع حروف الجذر بتسلسل ورودها داخل القالب.

www.alukah.net إهداء من شبكة الألوكة

على الرغم من كل المحاولات الجادة التي تُبذل لتحقيق هذا الهدف، محاولين «استكناه العمليات اللاإرادية التي تحدث في العقل البشري التي يمكن من خلالها إعطاء الحاسوب القدرة على فهم اللغة الطبيعية، وتحليلها، وإعادة إنتاجها، وكيفية تشكيلها في العقل البشري»[1].

٣. التحليل الصرفي

«يُقصد بالتحليل الصرفي الآلي للكلمة في اللغة العربية «ربط كلمات النص بالعناصر الصرفية الأولية التي تدخل في تكوينها، وكذلك بالقيم النحوية دون اعتبار موقعها»[2]. فيتم في التحليل الانتقال من الكلمة إلى جذرها الأصلي؛ أي أنَّ الحاسوب يعالج الكلمات العربية المشكولة جزئياً، أو كلياً، أو غير المشكولة، فيصف ما يطرأ عليها من تغيير؛ زيادة، أو نقصاناً، أو إعلالاً، أو إبدالاً، أو إدغاماً، أو قلباً، حيث «يحدد نوعها، وميزانها الصرفي، وسابقتها (أو سوابقها)[3]، ولاحقتها (أو لواحقها)[4]، وحالتها الإعرابية، ودلالتها، ... فإذا احتوت الكلمة المراد تحليلها على حروف غير مشكولة، وضع الحاسوب الحركات الممكنة لها اعتماداً على بيانات مخزنة. ومن المعلوم أن خلوَّ الكلمة من الشكل يجعلها متعددة الأشكال، ومن ثمَّ المعاني، مادامت مستقلة عن سياق النص.»[5]

«فكلمة (وجد) مثلاً يمكن أن تكون لها الإمكانيات التالية:

وَجَدَ، وَجَّدَ، وُجِدَ، وُجِّدَ ... = أفعال / وَجْدٌ = اسم / وَ+جَدَّ، وَ+جُدَّ ... = حرف عطف+أفعال / وَ+جَدٌّ = حرف عطف+اسم.

ومع ذلك فالكلمة المشكولة إذا عولجت مستقلة عن سياق النص، فلا يمنع شَكْلُها

١- نهاد الموسى، العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، بيروت، ط١، ٢٠٠١م، ص٥٧.

٢- يحيى هلال، التحليل الصرفي للعربية، وقائع مختارة من ندوة استخدام اللغة العربية في الحاسب الآلي في الكويت، عمّان، دار الرازي، ص٢٦٦.

٣- السوابق مجموعة من الحروف، والأدوات التي تسبق الكلمة، وتؤدي إلى تغيير معناها، أو وظيفتها النحوية.

٤- اللواحق مجموعة من الحروف التي تُضاف إلى آخر الكلمة، فتُغير معناها، أو وظيفتها النحوية.

٥- يُنظر: مروان البواب، ومحمد الطيّان، أسلوب معالجة اللغة العربية في المعلوماتية (الكلمة - الجملة)، استخدام اللغة العربية في المعلوماتية.

www.alukah.net إهداء من شبكة الألوكة

من إمكانية اشتراكها في الاسمية والفعلية، أو الفعلية والحرفية، فمن أمثلة الحالة الأولى كلمة (يَزِيدُ) فهي اسم في نحو قولنا: خَرَجَ يزيدُ من الغرفة. وفعلٌ في نحو قولنا: يزيدُ الله في خلقه ما يشاء. ومن أمثلة الحالة الثانية كلمة (أنَّ).

وهذا يعني أن على الحاسوب أن يعالج الكلمة عند تحليلها على أنها فعلٌ واسمٌ وحرف، وأن يعطي جميع الإمكانيات المحتملة لها، مع مراعاة الحالات التي تحدد نوعها، فالكلمة المنونة -مثلاً- لا تكون إلا اسماً. وبعد ذلك يقوم الحاسوب باختيار الإمكانية المناسبة التي تتوافق مع سياق النص»[1].

٤. المحلل الصرفي الآلي

هو تطبيق حاسوبي يقوم باستخلاص العناصر الأولية لبنية الكلمة في اللغة العربية، ويُحدد سماتها الصرفية، والصرف صوتية، والصرف نحوية، فيقوم المحلل الآلي بالكشف عن جذر الكلمة، ووزنها الصرفي، ويبيِّن ما يطرأ عليها من تغيير بالزيادة أو النقصان، والإعلال، والإبدال، والإدغام، والقلب، ويوضح ما يلحقها من سوابق، ولواحق، وزوائد، بالإضافة إلى تقسيم الكلمة إلى اسم، أو فعل، أو حرف، وتقسيم الاسم إلى جامد، أو مشتق، ومذكر، أو مؤنث، ومفرد أو مثنى أو جمع ...إلخ.

ويضم المحلل الصرفي مجموعة من قواعد المعطيات: هي قواعد معطيات معجمية لأوزان الكلمات العربية المستعملة، وقواعد معطيات لأسماء الأعلام، وقواعد معطيات للأخطاء الإملائية، والنحوية الشائعة في نصوص اللغة العربية.

٥. توأمة النحو والصرف في المعالجة الآلية

التداخل الكبير بين المستويين اللغويين الصرفي والنحوي في اللغة العربية «أوجب معالجتهما آلياً بأسلوب متداخل، فهناك تداخل مستمر بين الصرف والنحو ينشأ في بداية معالجتهما آلياً بأسلوب متداخل، ويستمر حتى نهايتها»[2]. فالمستويات عملية التحليل الصرفي الآلي للكلمات والجمل، اللغوية في اللغة العربية رغم تباينها فهي «متداخلة متكاملة دون تفاضل أو تمايز، فهي

١- عبدالعزيز بن عبدالله المهيوبي، بناء خوارزمية حاسوبية لتوليد الأفعال في اللغة العربية وتصريفها - دكتوراه، معهد تعليم اللغة العربية - جامعة الإمام محمد بن سعود الإسلامية، ١٤٣٦هـ، ص١٢٦ - ١٢٧.

٢- هدى آل طه، النظام الصرفي للعربية في ضوء اللسانيات الحاسوبية "مثل من جمع التكسير"، رسالة دكتوراه، الجامعة الأردنية، ٢٠٠٥م، ص٢١.

www.alukah.net إهداء من شبكة الألوكة

كلٌّ واحد تتآزر في بناء اللغة؛ فاللغة كالجسد الواحد، تربطه وحدة عضوية، وتصل بين أجزائه شرايين، وأعصاب قد تكون متناهية الدقة، إلا أنها تجعل سلامة عضو تعتمد على سلامة عضو آخر، بل حياته أيضاً»[1]، فالصرف يعتمد على الأصوات من جهة، وعلى النحو من جهة أخرى، فالعلاقة بين المستويين الصرفي والنحوي وطيدة، فهما توأمان يصعب الفصل بينهما.

٦. أهمية التحليل الصرفي

تتجلى في المحلل الصرفي أهم خصائص اللغة العربية في مجال المعالجة الحاسوبية، فيُتاح عن طريق المحلل الصرفي توليد جميع الكلمات التي يمكننا اشتقاقها من جذر معين، كما نستطيع من خلاله ردَّ أيَّ كلمة مشتقة إلى جذرها، أو أصلها الذي تعود إليه. كما يستطيع المحلل بعد الكشف عن جذر الكلمة توليد الأسماء المشتقة من الفعل المجرد، أو المزيد، ويولِّد مزيدات الفعل الثلاثي بحرف وبحرفين وبثلاثة أحرف، وكذلك مزيدات الفعل الرباعي بحرف وبحرفين، ويكتشف ما يصيب الكلمة من حالات الإعلال، أو الإبدال، أو الهمز، أو التضعيف.

تُعدُّ تطبيقات التحليل الصرفي لكلمات اللغة العربية بمثابة الأساس والقاعدة للتطبيقات الحاسوبية اللغوية الأخرى، إذ تستفيد منها بشكل، أو آخر، ولكنَّها تصبح أساسية بالنسبة لتطبيقات البحث والفهرسة، فهي تطبيق مباشر لها، حيث يقف المحلل الصرفي في مكان الصدارة بوصفه التطبيق الفاعل والسريع للمساعدة في الوصول إلى الكلمات المطلوبة عن طريق إعادة الكلمة المشتقة إلى جذرها، والتعرّف على الصور الصرفية لها. كما يُستخدم المحلل الصرفي في الترجمة الآلية، واسترجاع البيانات، "فيتولى المحلل ربط المفردات المختلفة للصيغ، مثل (كتب، الكتب، يكتبون، كاتبون، كتبتُ ...) التي يمكن استرجاعها تحت الجذر (ك ت ب) بالإضافة إلى إمكانية استرجاع الكلمات المختلفة حسب صياغاتها المتفاوتة، مع ما يتصل بها من سوابق أو لواحق"[2].

كما تبرز أهمية المحلل الصرفي عند التعامل مع النصوص العربية الكبيرة، مثل القرآن الكريم، وموسوعات الحديث النبوي على الحاسوب، فيكفي على سبيل المثال

١- نبيل علي، اللغة العربية والحاسوب (دراسة بحثية)، تعريب، د.ط، ١٩٨٨م، ص٤٠٣.

٢- علي السليمان الصوينع، استرجاع المعلومات في اللغة العربية، مطبوعات مكتبة الملك فهد الوطنية، السلسلة الثانية، الرياض، ١٩٩٤م، ص١٤٠.

www.alukah.net إهداء من شبكة الألوكة

أن تستخدم للبحث في القرآن الكريم جذراً، مثل «س ل م» فيستدعي المحلل جميع الآيات القرآنية التي وردت بها مشتقات هذا لجذر، مثل: (أسلم، سلام، سالمون، سليم، مسلمون، الإسلام ..إلخ)

ويُعد المحلل الصرفي إحدى الدعائم الأساسية التي يقوم عليها مشروع خدمة السنة النبوية، فحاجتنا للبحث - على مستوى الجذر- في الأحاديث التي تضم عدداً كبيراً من الألفاظ تتضاعف مع كثرة البحث، وتعدد أهدافه»[1]، كما يمكن للمحلل الصرفي دعم التشكيل الآلي للكلمات الخالية من التشكيل، حيث يعطي مجموعة من الخيارات لتشكيل الكلمة داخل النص. ويُساعد في التدقيق الإملائي للنصوص العربية؛ حيث يكتشف الخطأ الإملائي، ويقترح البدائل الصحيحة المحتملة، فالمحلل الصرفي عندما لا يكتشف الساق السليم للكلمة، والصيغة الصرفية التي يجوز انطباقها عليه، فإنه يَعدُّ الكلمة خاطئة، ثمَّ يقدم عدة احتمالات لتصويب الكلمة الخاطئة، عن طريق توليد احتمالات صحيحة، بحيث تكون أقرب ما تكون للصيغ الصرفية السليمة.

كما تدعم تطبيقات التحليل الصرفي محركات البحث في الإنترنت؛ حيث يمكننا البحث عن كلمة أو جملة أو مجموعة كلمات بحث مطابق، أو باللواصق، أو على مستوى الجذر، فإذا أردنا البحث عن كلمة (رأى) وكل ما يرتبط بها من كلمات داخل صفحات الشبكة العالمية، فسنحصل - بمساعدة تطبيقات التحليل الصرفي - على قائمة طويلة من الكلمات التي لا تشترك في بداياتها، أو نهاياتها، ولكنها تشترك في الجذر، مثل (نرى، يريكم، أرنا، يرون، تَرَ، أرني، رأيتموه، أراكم، رأيت، أراك، ليريه، فترى، يروا، أرأيتكم، ليريها، ترونهم، تراني، سأريكم، رأوا..إلخ.

٧. الهدف من بناء محللات صرفية آلية للغة العربية

إنَّ الهدف من بناء المحللات الصرفية الآلية هو بناء أداة لغوية تُمَكِّنُ الحاسوبَ من مشابهة الإنسان في كفايته، وأدائه اللغويين؛ «ليكون قادراً على تحليل نصوص اللغة العربية، وكلماتها، فيكتشف الأخطاء الإملائية عن طريق معرفة النظام الكتابي للغة العربية، ويحلل الصيغ الصرفية، ويتعرفها في سياق الكلام»[2].

١- محمود عوض المراكبي، تطويع تقنية المعلومات لخدمة العلوم الشرعية، السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات، الرياض، ١٩٩٢م، ص١٣٠.

٢- مسفر محماس الدوسري، برمجة الاسم المنسوب بياء النسب في العربية حاسوبياً - رسالة دكتوراه، جامعة اليرموك - كلية الآداب، ٢٠١٠م، ص٦.

www.alukah.net إهداء من شبكة الألوكة

٨. عرض نتائج التحليل

تتفاوت المحللات الصرفية في طريقة عرض نتائج تحليل الكلمات والنصوص، وترتيبها، وذلك تبعاً لقوة المحلل، والمنهج المتَّبع في الترتيب، حيث نجد أن بعض المحللات الآلية تكتفي بذكر نوع الكلمة وزوائدها، في مقابل ذلك اهتمَّ بعض مطوري المحللات بذكر سابقة الفعل، ولاحقته، وصيغته الصرفية، وتجرده، أو زيادته، ووزنه، وجذره، وإسناده، وبنائه للمعلوم، أو المجهول، وعلامات البناء، والإعراب، وضبطه بالشكل التام، وسابقة الحرف، ولاحقته، وعلامات بنائه، وسابقة الاسم، ولاحقته، ووزنه، وجذره، ونوعه من جهة التصرف، وعدمه، ومن جهة التذكير، والتأنيث، والنسبة، والتصغير، وعلامات بنائه، وإعرابه، وضبطه بالشكل التام.

٩. خطوات عمل المحلل الصرفي الآلي

عند تحليل الأفعال -مثلاً- فإن الحاسوب سيفترض أنَّ الكلمة المطلوب تحليلها هي فِعْلٌ، «فيقوم بتحديد سوابقها ولواحقها، ويحدد كذلك صيغة الفعل، وبناءه للمعلوم أو المجهول، وهل هو مجرَّد أم مزيد؟ ويحدد وزنه، وأصله المشتق منه، ويُبيِّن حالته الإعرابية، والضمير المسند إليه. أي أنه يعطي وصفاً كاملاً عن حالة الفعل الصرفية، والنحوية، والدلالية مستقلة عن سياق النص. وتمرُّ عملية تحليل الفعل بمراحل كثيرة»[١]، فعند تحليل كلمة «وحزن» يبدأ المحلل بالخطوات التالية:

الخطوة الأولى: وهي خطوة تشذيب[٢] الكلمة المدخلة، حيث ينزع المحلل الزوائد التي لحقت بساق الكلمة، ليكتشف أنها تحتمل أكثر من تحليل، فينزع في الاحتمال الأول (الواو) كحرف عطف ليكون ساق الكلمة هو (حزن)، أمّا في الاحتمال الثاني فيقتطع الحاسوب أحد حروف الفعل الأصلية (النون) ظناً منه أنها لاحقة زائدة لجمع المؤنث، ويُبقي على السابقة (الواو)، ولكنَّه يقوم بإعادتها بعد أن يفشل في تحليل ما تبقى من الكلمة (وحز) فيُعيد السابقة المقتطعة إليها، ثم يحللها مرة أخرى، والاحتمال الثالث هو

١- عبدالعزيز بن عبدالله المهيوبي، بناء خوارزمية حاسوبية لتوليد الأفعال في اللغة العربية وتصريفها - دكتوراه، معهد تعليم اللغة العربية - جامعة الإمام محمد بن سعود الإسلامية، ١٤٣٦هـ، ص١٢٧.

٢- التشذيب هو عملية إزالة كلِّ من بوادئ الكلمة ولواحقها لإنتاج الجذر او الجذع. وهي طريقة رياضية تجمع كل الكلمات التي تتقاسم الأصل نفسه، وتملك بعض العلاقات الدلالية، حيث تعمل طريقة التشذيب المبنية على الجذع على إزالة السوابق واللواحق المتصلة بالكلمة، في حين تحوِّل الطريقة المبنية على أساس الجذر الأصول إلى جذور.

www.alukah.net إهداء من شبكة الألوكة

نزع السابقة (الواو) واللاحقة (النون)، ليكون ساق الكلمة (حز)، وهي ساق سليمة عند تضعيف الحرف الثاني، أمّا الاحتمال الرابع فهو نزع السابقة (الواو) واللاحقة (النون) ليكون ساق الكلمة (حز) وهي ساق سليمة عند حذف حرف العلة، وأصلها (حوز).

الخطوة الثانية: يكتشف المحلل العمليات الصرف صوتية التي تمت على ساق الكلمة، وهي في الاحتمال الأول (حَزَنَّ) بنون النسوة، وفي الاحتمال الثالث (حُزَنَّ، حُزِّنَّ، حُزُّنَّ) بنون التوكيد الثقيلة، وفي الاحتمال الرابع (حُزْنَ) بنون النسوة، أمّا في الاحتمال الثاني فلا يجد المحلل أية أدلة على حدوث تغيرات صرف صوتية حدثتْ على ساق الكلمة.

الخطوة الثالثة: يقوم الحاسوب بعد ذلك بمقابلة ما تبقى من الفعل مع جداول الأفعال النموذجية المخزنة في الذاكرة. فيبحث عن الأفعال المساوية له في عدد الحروف، ليحصل على جميع الأفعال المفترضة الموافقة لهذا الفعل. وتحتوي جداول الأفعال النموذجية على جميع المعلومات المتعلقة بجذر الفعل، الباب الذي يتصرف منه، ووزنه، ومساره الاشتقاقي، حيث يحدد المحلل جذر الساق، وصيغته الصرفية للاحتمال الأول بأنها (ح ز ن) و (فعِل) أو (فعُل) أو (فَعَّلَ) أمّا الاحتمال الثالث فجذر الساق هو (ح ز ز) وصيغته الصرفية (فعل)، بالنسبة للاحتمال الرابع فيكون الجذر (ح و ز) وصيغته (فعل)، أما الاحتمال الثاني فلا تسفر عملية البحث عن جذر سليم، وصيغة صرفية يجوز انطباقها عليه.

تستمر بعد ذلك عمليات الفحص، وتحديد مكونات الفعل، حتى يعرض الحاسوب نتيجة التحليل الصرفي، والتي تسفر عن ثلاثة احتمالات سليمة، الاحتمال الأول: يفترض أن الفعل في الزمن الماضي، مسند إلى المفرد الغائب (هو) أو مسند إلى ضمير الغائبات (هُنَّ) ومسبوق بحرف العطف (الواو). أما الاحتمال الثاني، فيفترض أنَّ الفعل للأمر، مسند إلى المفرد المخاطب (أنتَ) أو المفردة المخاطبة (أنتِ) أو الجمع المذكر المخاطب (أنتم) في حالة التأكيد بالنون الثقيلة، والاحتمال الثالث، يفترض أن الفعل (حُزْنَ) للأمر مسند إلى الجمع المؤنث (أنتنَّ) أو أن الفعل في الزمن الماضي، مسند إلى الجمع المؤنث (هُنَّ).

www.alukah.net إهداء من شبكة الألوكة

١٠. نظرة تاريخية للتحليل الصرفي الآلي للغة العربية

حظيت اللغة الإنجليزية بنصيب الأسد في ميدان معالجة اللغات الطبيعية، أمّا بالنسبة للغة العربية فالأبحاث في هذا المجال انطلقت بصفة فردية في بداية السبعينيات في معامل بلدان أجنبية. فاعتمدت المحللات الصرفية العربية في بداياتها على الخبرة الفنية الحاسوبية، مع إغفال تام للخبرة اللسانية اللغوية، وربما اعتمد الحاسوبيون على بعض اللغويين التقليديين غير القادرين على فهم حاجات الحاسوب، حيث منيتْ تلك المحللات ذات الطابع الفني بفشل ذريع.

قام بعد ذلك مجموعة من الحاسوبيين واللغويين ببناء محللات صرفية آلية، تعتمد على قواعد صرفية تقوم باستخلاص عناصر بنية الكلمة، ويتميز هذا النوع من المحللات بمتانة أساسه اللغوي. فطوّر الدكتور «نبيل علي»، وأخصائية اللسانيات الحاسوبية «أمل الشامي» في عام ١٩٨٥م أول محلل صرفي آلي متعدد الأطوار[١]، وهو محلل صرفي قادر على التعامل مع أطوار التشكيل المختلفة للكلمة العربية، حيث يتكون المحلل من العناصر التالية:

١- المعالج الصرف نحوي:

يقوم هذا المعالج بدور المفكك، ليفصل جذع الكلمة عمّا يتصل به من السوابق واللواحق، كذلك يقوم برد التغيّرات الصوتية التي حدثت على عناصر بنية الكلمة الصرفية إلى أصلها؛ فعند تحليل كلمة «تمكَّنّا» يدرك المعالج وجود التضعيف على النون كأحد الحالات الممكنة لتشكيل الكلمة، فيقوم بتحليلها إلى «تمكَّن + نا».

٢- المعالج الاشتقاقي:

يستخلص المعالج الاشتقاقي في طور التحليل الجذر والصيغة الصرفية من الجذع[٢] الذي فكّكه المعالج الصرف نحوي، ويتم ذلك من خلال مقارنة سلسلة حروف الجذع مع قائمة قوالب الصيغ الصرفية دون علامات تشكيلها، وبعد استخلاص الجذر يتم مقارنته بمعجم الجذور الممكنة في اللغة العربية، فإذا فشل المعالج في الوصول إلى جذر مقبول يبدأ في افتراض وجود حالة أو أكثر من الإعلال أو الإبدال حتى يصل إلى جذر مقبول.

١- يُنظر: نبيل علي، اللغة العربية والحاسوب (دراسة بحثية)، تعريب، د.ط، ١٩٨٨م، ص٣٠٨.

٢- الجذع هو الكلمة التي يمكن أن تدخلها الزائدة الصرفية، وتُشكِّل مدخلاً معجمياً في العادة.

www.alukah.net إهداء من شبكة الألوكة

٣- المعالج الإعرابي:

يقدِّم المعالج الاشتقاقي خرجه إلى المعالج الإعرابي، حيث يقوم بتمييز الحالة الإعرابية بناء على الوسم الإعرابي الذي تتضمنه الكلمة.

٤- معالج التشكيل:

يقوم معالج التشكيل بتحديد عناصر التشكيل الغائبة أو الناقصة، ويقوم بالرجوع إلى المعجم ليتأكد من صحة انطباق الصيغة الصرفية على الجذر رهن المعالجة.

ويتميّز المحلل الصرفي الآلي متعدد الأطوار بقدرته على فضِّ اللبس الناجم عن غياب التشكيل، أو نقصانه، وذلك باستنباطه لجميع الاحتمالات الممكنة لتحليل الكلمة، كما يتميز بقدرته على التعامل مع الأنماط اللغوية، وعناصر التشكيل، وتداخل النحو والصرف، مع اهتمامه بما يلحق الكلمة من تغيّرات صوتية ممثلة في عمليات الإعلال، والإبدال، والإدغام.

بعد ذلك طوَّر كلّ من (تيم باك والتر، وكين بيسلي) من عام ١٩٨٨م إلى عام ١٩٩٠م محللاً صرفياً لكلمات اللغة العربية، يعتمد على أسلوب التحليل الصرفي ثنائي المستوى، وفي عام ١٩٩٦م طوَّر (كين بيسلي) المحلل، واستخدم تقنية جديدة للتحليل والتوليد، حيث يقبل المحلل الكلمات العربية المشكولة كليّاً أو جزئيّاً، ويقدم عدداً من الحلول الممكنة للكلمة المحللة، ويقل عدد تلك الحلول مع استخدام علامات التشكيل.

وفي عام ١٩٩٦م قامت شركة حوسبة النص العربي في عمّان ببناء محلل صرفي عربي بهدف الوصول إلى محلل قواعدي يستخلص جذر الكلمة، ويقدِّم معلومات الضمائر والحروف المضافة إليها. كما قامت الشركة الهندسية لتطوير نظم الحاسبات (Rdi) - وهي شركة مصرية - بتطوير المحلل الصرفي Arabmorph الذي يحلل الكلمات إلى جذورها وأوزانها، ويحدد توابعها، ويضم المحلل قواعد معطيات معجمية تفصيلية لكل كلمة، ويعتمد المحلل على السياق عند تحليل الكلمات التي تحتمل أكثر من معنى.

عُقد في دمشق خلال الفترة من ٢٦ إلى ٢٨ أبريل ٢٠٠٩م اجتماع خبراء المحللات الحاسوبية الصرفية للغة العربية، وذلك بدعوة من المنظمة العربية للتربية والثقافة والعلوم، وبالتعاون مع مجمع اللغة العربية بدمشق، ومدينة الملك عبد العزيز للعلوم والتقنية بالرياض، حيث شارك في الاجتماع باحثون جامعيون عرب، وأجانب من المملكة العربية السعودية، والمغرب، والجزائر، وتونس، ومصر، وسوريا، وبريطانيا،

www.alukah.net إهداء من شبكة الألوكة

وفرنسا، والولايات المتحدة الأمريكية. وتمحور جدول أعمال الاجتماع حول التعريف بالمحللات الصرفية الآلية للغة العربية المقدمة من قِبَل المشاركين في الاجتماع، مع عرض نتائج تطبيق معايير تقييم المحللات الصرفية على المحللات الصرفية التي قدَّمها المشاركون. وسنعرض في عجالة بعض تلك المحللات:

١- المحلل الصرفي للغة العربية لمخبر «ميراكل» صفاقس - تونس:

يقوم محلل مخبر «ميراكل» بالتعرّف على السوابق واللواحق مع اكتشاف زوائد الكلمات، واستخراج الخصائص النحوية، والصرفية الممكنة لها. وعندما حلّلنا كلمة «وزوجناكها» باستخدام المحلل حصلنا على النتيجة التالية:

الواو	حرف عطف
زوَّجنا	فعل غير ناسخ، مسند إلى ضمير متكلم جمع مذكر / مؤنث. الصيغة: الماضي، البناء: للمعلوم، الجذر: زوج، اللاحقة: ١- (ك) المخاطب مفرد مذكر / مؤنث. ٢- (ها) الغائب مفرد مؤنث

الجدول ١: تحليل كلمة «وزوجناكها»

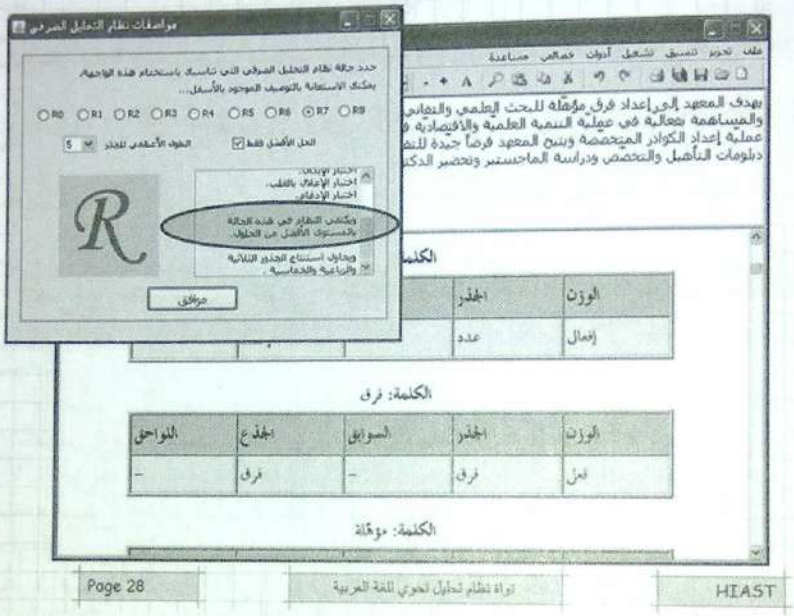
٢- محلل صرفي موجَّه بالتطبيقات - المعهد العالي للعلوم التطبيقية والتكنولوجيا - سوريا:

يهدف المشروع إلى تطوير محلل صرفي آلي يتمتع بالمرونة الكافية ليكون مناسباً للاستخدام في جميع المجالات. وتتكون الخوارزمية المتَّبعة في نظام المحلل من المراحل التالية:

- مرحلة تحديد نوع الكلمة: حيث يختبر المحلل كون الكلمة المدخلة أداة، أو كلمة معربة، وذلك باستخدام قائمة من الأدوات والكلمات الجامدة والمعربة.
- مرحلة الكشف عن الحروف الأصلية: وقد طوَّر المعهد خوارزمية خاصة للكشف عن الأحرف الأصلية، يصل المحلل بعد هذه المرحلة على حلول لـ ٦٠٪ من الكلمات.
- مرحلة عرض الاحتمالات: يعرض المحلل مجموعة من الاحتمالات والحلول، يتضمن كل حلّ الحروف الأصلية جميعها، ولا يتضمن أي حرف زائد، مع الإشارة إلى السوابق، واللواحق الصحيحة، وإلى حالات الإدغام، والإعلال، والإبدال.
- مرحلة تصحيح الحلول: حيث يصحِّح المحلل الحلول بتطبيق مجموعة من

www.alukah.net إهداء من شبكة الألوكة

الخطوات الاختبارية؛ للتأكد من وجود الوزن، والجذر، وتطبيق قواعد الإبدال، والإعلال. وعند تحليل كلمة «فرق» باستخدام المحلل حصلنا على النتيجة التالية: الكلمة «فرق» الوزن: فعل، الجذر: فرق، السوابق: لا يوجد، الجذع: فرق، اللواحق: لا يوجد. وقد أغفل المحلل ذكر الكثير من المعلومات الصرفية والنحوية المهمة للفعل «فرق» كنوع الفعل، وصيغته الصرفية، وحالته الإعرابية (انظر الشكل التالي)، كما أغفل المحلل ذكر العديد من الحالات القابلة للتحليل.



الشكل ١: تحليل كلمة «فرق»

٣- محلل صرفي للغة العربية باستخدام تقنيات الذكاء الاصطناعي - فاضل سكر، وسمر معطي - سوريا:

يجرِّد المحلل الصرفي الكلمة المدخلة [١] من السوابق واللواحق، ويبحث عن الوزن الصحيح، ثم الصيغة الصرفية الصحيحة، مع إيجاد العلاقة الصرف نحوية. وقد استخدم مطورو المحلل لغة visual prolog لأنها تختلف عن اللغات التقليدية بكونها الأكثر قرباً من لغة الإنسان، وتتضمن قواعد معطيات المحلل: أوزان الأسماء، والأفعال في اللغة العربية، بحيث يسمح محرك بحث المحلل بتوليد الافتراضات بناءً على الأوزان المخزنة في قواعد المعطيات.

١- دخل المحلل يكون كلمة مشكولة كليّاً أو جزئيّاً أو غير مشكولة.

www.alukah.net إهداء من شبكة الألوكة

٤- محلل صرفي لكلمات اللغة العربية خارج السياق وداخله - جامعة محمد الأول وجامعة قطر:

يعمل هذه المحلل معالجة على كلمات اللغة العربية بطريقتين:

الطريقة الأولى: معالجة الكلمات خارج السياق، حيث يعمد النظام إلى تفكيك الكلمات إلى لبناتها الصرفية من سوابق، وجذوع، ولواحق؛ مما يسمح بتحديد الحلول الصرفية المحتملة لها باللجوء إلى قواعد معطيات المحلل.

الطريقة الثانية: معالجة الكلمات داخل السياق؛ ويعتمد المحلل على نموذج إحصائي يسمح بتحديد الحل الأكثر رجحاناً لكل كلمة بالنظر إلى الحلول المقترحة في الكلمات السابقة من الجملة.

٥- محلل صرفي مصدري عربي للتطبيقات العامة - مدينة الملك عبد العزيز للعلوم والتقنية:

هو محلل مصدري، وليس صرفي لكلمات اللغة العربية، يهدف إلى الوصول إلى مصدر الكلمة وليس جذرها، ويحدد لواصقها، ويعتمد المحلل التشابه في الشكل الخارجي للمفردات المولدة على الميزان الصرفي نفسه، وباللواصق نفسها. يتميز هذا المحلل بالسرعة الكبيرة، ولا يحتاج لجداول كثيرة، وهو جيد لبناء التطبيقات الحاسوبية العامة لمحركات البحث، كما يمكننا استخدامه كخطوة أولى لتفكيك المفردة العربية قبل تحليلها صرفياً.

٦- برنامج مداد للتحليل الصرفي للكلمات العربية - شركة مداد لتقنية المعلومات:

يهدف محلل مداد إلى تفكيك النصوص العربية إلى كلمات، وتحليلها لتحديد أنواع كلماتها، ثم تحديد الزوائد من سوابق ولواحق، وحروف مزيدة، بهدف الوصول إلى الجذر، مع عرض التشكيلات الممكنة للكلمة. وتضم قواعد معطيات المحلل الجذور، والكلمات الشاذة، أمّا باقي الكلمات التي تأتي حسب القواعد فيستطيع المحلل التعرف عليها دون الرجوع إلى قواعد المعطيات.

www.alukah.net إهداء من شبكة الألوكة

١١. طرق التحليل الصرفي الآلي[١]

هناك العديد من الدراسات التي تناولت التحليل الصرفي، وقد اتَّبعت هذه الدراسات طرقاً مختلفة لمعالجة الكلمات صرفياً، حيث يمكننا إيجازها فيما يلي:

الطريقة الأولى: طريقة قوائم الكلمات المخزنة، وتعتمد هذه الطريقة على تخزين كلمات اللغة العربية جميعها في قوائم مع مكوناتها الصرفية على شكل جداول كبيرة في قواعد معطيات ضخمة، تضم الانزياحات الصرفية بأشكالها المختلفة، ويحلِّل البرنامج الكلمة المدخلة عن طريق البحث عنها في هذه الجداول، ومن ثمَّ معرفة جذرها ببساطة، ويمكن تطبيق هذه الطريقة على نصوص معينة، مثل القرآن الكريم، أو مجموعة كتب محددة. ويعيب هذه الطريقة ضخامة حجم المواد اللغوية التي نقوم بإدخالها إلى الحاسوب، وتضييقها للتحليل اللغوي، باعتمادها مواد المعجم مرجعاً وحيداً للتحليل.

الطريقة الثانية: الطريقة اللغوية، ويكون ذلك عن طريق توصيف قواعد اللغة العربية الصرفية، وتحويلها إلى خوارزميات حاسوبية، فيحاكي المحلل عمل اللغوي عند تصريف الكلمات، وتحليلها. وتعالج هذه الطريقة عيوب الطريقة الأولى.

الطريقة الثالثة: الطريقة الرياضية، وذلك عن طريق تحليل الكلمات بشكل آلي بطريقة التجربة والخطأ والتصحيح، فالكلمة هنا مجموعة من الحروف يأخذ المحلل ثلاثة أحرف منها، ويقارنها بقائمة الجذور المخزنة في قواعد المعطيات، فإذا لم يجد المحلل الجذر في القائمة، أخذ ثلاثة أحرف أخرى، ويستمر إلى أن يجد الجذر الأقرب إلى الصواب.

الطريقة الرابعة: طريقة الأوزان، وذلك بتوليد مجموعة من القواعد النصية الآلية عن طريق المقارنة بين قائمة كبيرة من الكلمات مع ما يقابلها من مصادرها، وتُستخدم هذه القواعد لمعرفة مصدر الكلمة. فإذا انطبقت على الكلمة أكثر من قاعدة واحدة، رجَّحَ المحلل أكثر القواعد تكراراً.

١- يُنظر: عبدالله بن عبدالرحمن الزامل، العلاقة الصرفية بين الجذور والأوزان، (الأوراق البحثية للندوة الدولية الأولى عن الحاسب واللغة العربية) الرياض- السعودية، مدينة الملك عبد العزيز للعلوم والتقنية، ٢٠٠٧م، ص٢٩٩-٢٣٠.

www.alukah.net إهداء من شبكة الألوكة

١٢. ضوابط ومحددات للمساعدة في بناء المحللات الصرفية [١]

يعتمد النظام الصرفي والنحوي للغة العربية على مجموعة من الضوابط، والمحددات التي تُشكّل بناء كلمات اللغة العربية وجملها، وتتوزع هذه الضوابط والمحددات بين ضوابط شكليَّة، وأخرى دلالية، وهي في مجموعها محددات يمكن للعقل البشري فهمها، واستنباطها بخلاف الحاسوب الذي لا يُدرك إلا الشكلي منها، وهذه الضوابط والمحددات يمكننا ترجمتها إلى مسائل منطقية، يسهل برمجتها حاسوبياً.

إنَّ توصيف قواعد تحليل كلمات اللغة العربية مع رصد ضوابطها، يقصد إلى تمكين الحاسوب من كشف صيغ المشتقات في النصوص المشكولة، وغير المشكولة. وتنقسم ضوابط الكلمات إلى نوعين: صرفية وهي العلامات أو الخصائص التي تميِّز الكلمة من حيث بنيتها الصرفية، وما يمكن أن تقبله من تغيّرات، وما يمكن أن يطرأ عليها من أحوال. و»ضوابط نحوية تتميز بها الكلمة من خلال وجودها في التركيب، ومن خلال ارتباطها مع غيرها من الأبنية بعلاقات تحددها طبيعة التركيب نفسه، فهي ضوابط لا يتحقق وجودها إلا في التركيب»[٢].

هناك محددات، وضوابط عامة خاصة بالأسماء دون الأفعال والحروف، مثل أل التعريف، وحروف الجر، والإضافة، والتاء المربوطة، والتنوين، واتصال (ون) بجمع المذكر السالم و(ات) بجمع المؤنث السالم ...إلخ. بينما هناك محددات تميز المشتقات عن الأسماء؛ فلكل مشتق صيغة خاصة به دون غيره من المشتقات؛ فصيغة (فعّال) -مثلاً- تختص بالمبالغة، غير أن بعض صيغ المبالغة قد توافق بعض صيغ الصفة المشبهة؛ كصيغة (فعيل) ولا يفصل بينهما إلا المعنى السياقي.

١٣. مشكلات تواجه بناء محلل صرفي دقيق لكلمات اللغة العربية ونصوصها

إنَّ بناء محلل صرفي دقيق لكلمات اللغة العربية ونصوصها ليس بالأمر الهين، « بل يتطلب الكثير من الجهد، كما يحتاج إلى فرق بحث متخصصة ذات تصور كامل حاسوبياً ولغوياً، فمعظم النظم والبرامج المجرَّبة على اللغات الإنسانية لم تسلم حتى الآن من

١- يُنظر: عزت جهاد عزت العجوري، توصيف لغوي صرفي لشعر بدر شاكر السيّاب في ضوء اللسانيات الحاسوبية، رسالة ماجستير، الجامعة الهاشمية، ٢٠٠٩ م، ص ٨٧-٩٠.

٢- لطيفة النجار، دور البنية الصرفية في وصف الظاهرة النحوية وتقعيدها، دار البشير، عمّان، ط١، ١٩٩٤م، ص٤٣.

www.alukah.net إهداء من شبكة الألوكة

الكثير من المشكلات والصعوبات سواء على المستوى المنهجي، أو الصوري للغة»[1].
ويمكننا تقسيم هذه المشكلات إلى:

١,١٣ مشكلات لغوية:

إنَّ الدراسات الصرفية القديمة غير كافية لبناء محلل صرفي حاسوبي للغة العربية، فمع ما أحرزه القدماء من تقدم في دراسة صرف اللغة العربية، ومحاولاتهم الجادة في ضبط نظامها الصرفي والصوتي، «لكنها بحد ذاتها غير كافية للتعامل العلمي مع اللغة، ذلك التعامل الذي يأخذ اللغة بوصفها ظاهرة»[2]. فالحاسوب لا يتعامل إلا مع خوارزميات تضبط عمليات توليد الكلمات، وتحليلها، مما يتطلب رصد دقائق بنية صرف اللغة العربية، والإحاطة الكاملة بكلماتها. ونوجز فيما يلي أبرز تلك المشكلات اللغوية:

١- غياب التشكيل، والذي يمثل -بلا منازع- أكبر عقبة تواجه تحليل الكلمات العربية حاسوبياً؛ فنتيجة لغياب التشكيل يمكن لعدة صيغ صرفية مختلفة أن تستخدم هيكلاً واحداً للحروف، وعلى الحاسوب أن يُخَمِّنَ الصيغة الصرفية المقصودة، مثل: «كتب» يمكن أن تكون: «كَتَبَ، كُتِبَ، كُتُبٌ، كَتَّبَ». فيأخذ في الاعتبار كل حالات اللبس[3] الممكنة، وبالتالي على القائمين على بناء تطبيقات التحليل الصرفي الآلي تغطية جميع حالات اللبس التي يسببها غياب التشكيل من خلال بناء مجموعة من الخوارزميات[4] لتغطية جميع الحالات الممكنة للكلمة.

٢- الرسم الإملائي: تختلف طُرق كتابة بعض الكلمات المعرَّبة في اللغة العربية، مما يؤثر على دقة عمل المحلل الصرفي الآلي، وذلك مثل: (مسؤول - مسئول، كمبيوتر - كومبيوتر، أوروبا - أوربا)

١- عزت جهاد عزت العجوري، توصيف لغوي صرفي لشعر بدر شاكر السياب في ضوء اللسانيات الحاسوبية، رسالة ماجستير، الجامعة الهاشمية، ٢٠٠٩ م، ص ١٨-١٩.

٢- حسام الخطيب، العربية في عصر المعلوماتية - تحديات عاصفة ومواجهات متواضعة، مجلة التعريب، المركز العربي للتعريب والترجمة والنشر، العدد الثاني، ١٩٩٨م، ص٧٧.

٣- اللبس نوعان: لبس حقيقي، يكون فيه للكلمات التشكيل نفسه كما في كلمة "كَمال" فهي تحتمل: "كمال = اسم علم" و "كَمال = كـ + مال". ولبس غير حقيقي، يكون ذلك عند غياب التشكيل، كما في "كتب" حيث تحتمل "كَتَبَ و كُتُبٌ و كُتِبَ".

٤- الخوارزميات: مجموعة قواعد وقوانين مكتوبة، تستعمل لوصف الخطوات المنطقية المتبعة لمعالجة البيانات الداخلة للحصول على المعلومات والنتائج المطلوبة. وقد سميت الخوارزميات بهذا الاسم نسبة إلى العالم العربي المسلم "أبو جعفر محمد بن موسى الخوارزمي" (٨٢٥م)، والذي اشتهر في مجال الرياضيات، وقد ألف كتابه المشهور "الجبر والمقابلة".

www.alukah.net إهداء من شبكة الألوكة

٣- الأسماء المترجمة: هناك اختلاف في كتابة الأسماء الأجنبية باللغة العربية، مثل: (كوفي أنان - كوفي عنان، وفرانسوا أولاند - فرانسوا أولند، ومحاضير بن محمد - مهاتير بن محمد).

٤- كلمات الوقف: وهي كلمات كثيرة الورود في النصوص، ولا تحمل معاني إذا فُصلتْ عن السياق، ولا تُكوِّنُ جملة مفيدة عند استخدامها وحدها، وهي حروف وأدوات لازمة لتركيب الكلام العربي، مثل: حروف الجر، والعطف، والاستفهام، والنفي، والتعجب، والنداء، والظروف، والضمائر ...إلخ.

٥- الأخطاء الإملائية: تكثر الأخطاء الإملائية في الكتابات المعاصرة، حيث يمكننا كتابة بعض الكلمات في صور إملائية مختلفة من بينها الصورة الصحيحة، ونوجز الأخطاء الإملائية في النقاط التالية:

- الخطأ في كتابة همزتي الوصل والقطع، والهمزة المتوسطة، واختلاف كتابة الهمزة باختلاف حالة الكلمة الإعرابية، فقد لوحظ أن الهمزة تلعب دورًا كبيرًا في التمهيد للمحلل الصرفي لتحليل الكلمة المطلوبة بسهولة، في حين يشكل عليه تحليلها بدون وجود الهمزة.

- اختلاف كتابة الياء المنقوصة، مثل «قاضي» فهي منقوطة في الكتابة الشامية، وغير منقوطة في الكتابة المصرية.

إنَّ لمشكلة تفاوت رسم الكلمات جوانب سلبية على عملية تحليل الكلمات، «وترجع ظاهرة التفاوت إلى سببين، أحدهما الأخطاء البشرية، وثانيهما اختلاف الممارسات، أو القواعد المتبعة لإملاء الكلمات المعربة، والأسماء الأجنبية -كما أشرنا إلى ذلك سابقاً- والتي يختلف رسمها بين المؤلفين العرب»[1]، مما يؤثر على دقة تحليل الكلمات.

٦- من السهل على الحاسوب أن يميِّز أبنية المثنى، ولكنَّه يعجز عن تمييز كلمات أخرى، إذا عُرضتْ عليه وكانت تنتهي بألف ونون (ان) أو ياء ونون (ين)، وهي ليست مثنى، إلا إذا كان المحلل الصرفي مُزوَّداً بقاعدة معطيات تكون دليلاً إلى معرفة الكلمة بعد تجريدها من الألف والنون أو الياء والنون، «فإنْ دلَّتْ بعد التجرّد -غالباً- على

١- علي السليمان الصوينع، استرجاع المعلومات في اللغة العربية، مطبوعات مكتبة الملك فهد الوطنية، السلسلة الثانية، الرياض، ١٩٩٤م، ص٧٥.

www.alukah.net إهداء من شبكة الألوكة

مفردة مفيدة كانت مثنى، وإن لم تدل كانت كلمة أخرى»[1]، فكلمة (رجلان) بعد تجريدها من الألف والنون (ان) تصبح (رجل) إذن الكلمة مثنى، وذلك بخلاف كلمة (كان) فبعد تجريدها من الألف والنون تصبح (كَ) إذن هي كلمة أخرى وليست مثنى.

٧- وجود أكثر من معنى للبنية الصرفية الواحدة، ونعني بذلك أنَّ «بنية الكلمة الواحدة تحتمل أكثر من معنى واحد، فكلمة (ظهور) تكون مصدراً للفعل (ظهر) أو جمعاً للمفرد (ظَهْر). وإزالة اللبس هنا يحتاج إلى العديد من الأدلة الإضافية التي تساعد على التمييز بين الكلمتين السابقتين»[2] كأن نضع الكلمتين في سياق لغوي. كما أنه يمكن أن يكون للصيغة الصرفية الواحدة في اللغة العربية أكثر من وظيفة نحوية؛ فصيغة «فُعول» يمكن أن تكون مصدراً، نحو: «جُلوس»، وجمع كثرة، نحو: «سُيوف».

٨- التغيّرات الصوتية، وهي تغيّرات تحدث في بنية الكلمة، وتطرأ على بعض أصوات اللغة العربية في سياقات صوتية معينة، حيث تتغيّر بعض أصول الكلمة بحذفها، أو إبدالها، أو قلبها، أو إدغامها مع صوت آخر، أو إعلالها، أو إعادتها إلى أصلها؛ فالواو -مثلاً- حُذفت في «يقف» ثم عادت للظهور في «وقف»، في حين عادت ألف «جرى» إلى أصلها في «يجري». والحذف يكون في الصوامت أيضا؛ كحذف نون المثنى وياء المخاطبة إذا وقعت نون التوكيد بعدهما، كما في «يكتبانِّ» و «تكتبِنَّ»، وكذلك حذف نوني المثنى، وجمع المذكر السالم عند الإضافة، كما في «كاتبا الصحيفة، وكاتبو الصحيفة»، وكذلك حذف نون الفعل المضارع إذا كان من الأفعال الخمسة في حالة النصب، أو الجزم، أو مع نون الوقاية.

ويكون التغيير أيضاً بإبدال حرف صحيح بحرف آخر، كإبدال تاء «اِفْتَعَلَ» طاءً إذا كانت فاؤها صاداً أو ضاداً أو طاءً أو ظاءً، نحو: «اصْطَبَرَ» وأصلها «اِصْتَبَرَ»، وكذلك إبدال تاء «اِفْتَعَلَ» دالاً، إذا كانت الفاء دالاً، أو ذالاً، أو زاياً، نحو: «اِدَّهَنَ» وأصلها «اِدْتَهَنَ»، ومن التغيّر بالإبدال أيضاً، إبدال تاء «اِفْتَعَلَ» ثاءً، إذا كانت الفاء ثاءً، نحو: «اِثَّأَرَ» وأصلها «اِثْتَأَرَ».

١- محمود مصطفى عيسى خليل، إسناد الأفعال إلى الضمائر في ضوء اللسانيات الحاسوبي- ماجستير، كلية الدراسات العليا، الأردن، ٢٠١١م، ص٤٩.

٢- نهاد الموسى، العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، بيروت، ط١، ٢٠٠١م، ص٢٠٢.

www.alukah.net إهداء من شبكة الألوكة

ويكون التغيّر بالإعلال بالقلب، كقلب الواو والياء ألفاً، كما في «جالَ» من «جَوَلَ»، وقلب الواو ياءً، نحو: «رضي» من «رضو»، وقلب الياء واواً، نحو: «موقِن» من «مُيْقِن». كما يكون الإعلال بالحذف، نحو: «طُفْ» وأصلها «طُوفْ»، ويكون الإعلال بالتسكين، نحو: «يَسْمُوْ» وأصلها «يَسْمُوُ». أمّا التغيّر بالإدغام، فكما في «حَدَّ» وأصلها «حَدَدَ» ثمَّ حُذفت الفتحة التي بين الحرفين الثاني والثالث، مما أدى إلى الإدغام، بسبب تجاور صوتين متشابهين. وتمثّل هذه التغيّرات تحدياً يواجه الباحثين في مجال اللسانيات الحاسوبية عند تصميم المحللات الصرفية الآليّة، نظراً لكثرة هذه التغيّرات وتنوعها. مما يؤدي إلى أعباء إضافية في ردِّ الفرع إلى الأصل عند تحليل الكلمة

٩- صعوبة تعرف المحللات الصرفية الآلية على المصدر الصناعي، وذلك نحو: اشتراكيّة، انتهازيّة، شموليّة...إلخ.

١٠- اللغة العربية ذات عمليات صرفية معقدة تعتمد على العدد (مفرد، ومثنى، وجمع) والضمائر المتصلة والمنفصلة.

١١- دمج الأدوات، والضمائر المتصلة مع الكلمات في اللغة العربية، حيث تتغير صورة الكلمة في اللغة العربية عند اتصالها بالضمير، مثل: كتبت، كتبنا، كتبوا...إلخ.

١٢- عدم توفر توصيف دقيق ومتكامل لقواعد الصرف العربي، حيث تكتفي معظم كتب الصرف بشرح عام للقواعد الصرفية، مقرونة ببعض الأمثلة عن حالات الشذوذ والاطّراد.

١٣- تمثل الكلمات المركَّبة في اللغة العربية صعوبة بالغة عند تحليلها آليّاً؛ وذلك بسبب وجود الفراغ الذي يفصل بين عناصر الكلمة المركبة، حيث تختلط تلك الكلمات مع الكلمات غير المركبة.

١٣, ٢ مشكلات حاسوبية:

هناك هوَّة كبيرة تفصل بين اللغويين والحاسوبيين، ولعل من أبرز مسبباتها «ذلك التسارع في التطور الحاسوبي من جهة، والتباطؤ في الدراسات اللغوية من جهة أخرى، إلى جانب المرجعية الغربية لتطبيقات الحاسوب، واللسانيات الحاسوبية»[1]، يضاف إلى

١- عزت جهاد عزت العجوري، توصيف لغوي صرفي لشعر بدر شاكر السيّاب في ضوء اللسانيات الحاسوبية، رسالة ماجستير، الجامعة الهاشمية، ٢٠٠٩ م، ص ٢٢.

www.alukah.net إهداء من شبكة الألوكة

ذلك أن معظم برامج التحليل الصرفي الآلي لكلمات اللغة العربية ونصوصها هي من تطوير الحاسوبيين، حيث انشغل الحاسوبيون بالمطّرد من قواعد الصرف العربي دون النظر إلى الظواهر الشاذة، مع الاهتمام بالجانب التوليدي للكلمات دون تحليلها.

١٤. كيفية توصف القواعد الصرفية لبناء المحلل الصرفي الآلي

الحاسوب آلة صماء، لا تملك عقلاً مدركاً، ولا يمكن أن يكون الحاسوب قادراً على تقدير الأمور إلا بمقتضى حدود البرمجة؛ فهو غير قادر على تمييز كلمة (انتقل) إذا جاءت خارج سياقها، أهي فعل أمر، أم فعل ماضٍ؟ لذا ينبغي أن يوصِّف اللغوي بمساعدة الحاسوبي قواعد اللغة للحاسوب.

«وتبدأ عملية التوصيف بإيداع الحاسوب القواعد، والأساسيات الابتدائية التي يختزنها العقل الإنساني، بهدف الوصول إلى الكفاية اللغوية، ويكون ذلك عن طريق عرض منهجي قادر على استقراء القواعد، وتفصيلها وفقاً لمستويات اللغة المتفاوتة (الصوتي والصرفي والنحوي)»[١] فعند توصيف الفعل ينبغي أن نبيِّن نوعه من حيث البناء للمعلوم أو المجهول، وعلامة بنائه، وتوصيفه من ناحية صرفيّةٍ ثلاثياً أو رباعياً، مجرداً أو مزيداً، صحيحاً أو معتلاً، مع الإشارة إلى ما أصاب الفعل من إعلال أو إبدال أو إدغام، وهكذا يتم توصيف الجانب الصوتي والصرفي والنحوي.

١٥. متطلبات بناء المحلل الصرفي الآلي[٢]

أولاً- متطلبات لغوية:

١- تحديد جذور الكلمات العربية، لمعرفة أصول الكلمات التي تتشابه فيها البنية والضبط مع اختلاف الجذر.

٢- تحديد الأعلام دون تحليلها إلى مستوى الجذر.

٣- تحديد الكلمات الثابتة[٣] التي لا تُشتق منها كلمات أخرى، وهي الكلمات التي

١- نهاد الموسى، العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، بيروت، ط١، ٢٠٠١م، ص٦١.

٢- يُنظر: عبدالعزيز بن عبدالله المهيوبي، إشكاليّات تطوير محلل صرفي حاسوبي دقيق للغة العربية (محلل الخليل نموذجا)، مجلة اللغة العربية وتعليمها للناطقين بغيرها، جامعة أفريقيا العالمية، العدد ٢١، ٢٠١٦م.

٣- وتُسمّى كلمات التوقف أو الوقف.

www.alukah.net إهداء من شبكة الألوكة

تثبت كما هي دون حاجة للاشتقاق منها، مثل (هؤلاء، ذلك ....).

٤- تحديد الفروق الدقيقة بين الكلمات الملبسة.

٥- بناء قاعدة معطيات للأوزان القياسية للأسماء، والأفعال المشتقة من كل جذر.

٦- بناء قاعدة معطيات للسوابق، واللواحق، والزوائد، التي يمكن أن تأتي في بداية كل كلمة أو نهايتها.

٧- بناء قاعدة بيانات لتخزين نتائج التحليل الصرفي للكلمات.

ثانياً- متطلبات تقنية (برمجية):

١- بناء قواعد المعطيات، وبرامج إدخال المواد اللغوية وبرامج تعديلها بعد الإدخال.

٢- بناء برنامج التحليل الصرفي الآلي باستخدام إحدى لغات البرمجة.

٣- بناء برنامج لربط الجذور بمشتقاتها المختلفة الموجودة في قواعد المعطيات.

٤- بناء برنامج للتشكيل الآلي للكلمات.

٥- بناء برنامج للتصحيح الإملائي[1].

١٦. قصور المحللات الإنجليزية عن استيعاب خصائص اللغة العربية

يواجه مطورو المحللات الصرفية العربية صعوبات تتعلق بثرائها الصوتي، والصرفي، والمعجمي الواسع، نظراً لقلة الأبحاث الأكاديمية، والتقنية المرتبطة بها، وتناثرها، وغياب التنسيق فيما بينها، سواء من الناحية النظرية أو العملية، مع قلة الإمكانيات المتاحة. «وكان من أثر ذلك أن استعار مطورو المحللات الصرفية الآلية العربية حلولاً من النظريات الخاصة باللغة الإنجليزية، حيث لم تسهم تلك الحلول في استيعاب خصائص اللغة العربية وطاقاتها حاسوبياً؛ لأنها في شتى قواعدها أشمل، وأثرى من النموذج الإنجليزي، حيث يقف المحلل الصرفي للغة الإنجليزية -نظراً لخلو اللغة الإنجليزية من خاصية الاشتقاق- عند حدود ساق الكلمة، فإذا أردنا أن نستخدمه في نطاق اللغة العربية، فلن نجد أي ارتباط بين الجذر ومشتقاته، فإذا بحثنا عن الفعل «اعلم» باستخدام محلل صرفي صُمِّمَ للغة الإنجليزية وجدناه في حرف الألف، بينما

١- يقوم المدقق الإملائي باكتشاف الأخطاء الإملائية، واقتراح التصحيحات المناسبة البديلة لها. ويُعدُّ مدقق صخر واحداً من أوائل المدققات الإملائية التجارية العربية.

www.alukah.net إهداء من شبكة الألوكة

نجد «تعلم» في حرف التاء، وهذا يجرِّد اللغة العربية من خاصية استدعاء الجذر لمشتقاته الذي شُيِّدتْ على أساسها ثروتها اللفظية في المعاجم، وكتب التراث المتداولة»[1].

كما يتّسم التصريف في اللغة العربية بالاطّراد التام عدا حالات نادرة، في حين يزخر تصريف الإنجليزية بحالات شذوذ متعددة. كما أن لظاهرة الإعراب أهمية كبيرة في اللغة العربية، بخلاف اللغة الإنجليزية التي تغيب عنها هذه الظاهرة بشكل شبه تام. وتتصف اللغة العربية بإمكانية دمج الضمائر والأدوات مع كلماتها، وغيابها في كلمات اللغة الإنجليزية.

ويتضح من هذا التباين أنه لا بديل من بناء نموذج لغوي لمحلل صرفي آلي مبتكر قادر على التعامل مع طبيعة اللغة العربية، يتم فيه توصيف القواعد الصرفية والنحوية بطريقة تناسب أساليب المعالجة الآلية، دون اللجوء إلى الحلول المستوردة من اللغات الأجنبية.

١٧. لماذا تفوقت المحللات الصرفية العالمية على العربية؟

لقد أصبحت مشاريع معالجة اللغة العربية حاسوبياً سلعة تجارية تصدَّتْ لها شركات تجارية، بسبب تأخر اللغويين والحاسوبيين العرب عن الخوض في هذا المجال، عدا بعض الأعمال الفردية التي افتقدت إلى الدعم المالي. وعلى العكس من ذلك تسعى الدول المتقدمة لدعم البحوث العلمية في مجال اللسانيات الحاسوبية، مع تقديم الدعم اللازم للقطاع الخاص، ومراكز البحوث. كما أنَّ لبعثرة جهود اللغويين والحاسوبيين العرب دور كبير في هذه الفجوة الرقمية بين النظم اللغوية العالمية، والنظم العربية، فكل باحث أو شركة عربية تعمل بمعزل عن الأخرى.

١٨. أسس مقترحة لبناء محلل صرفي دقيق للغة العربية

سنطرح هنا مجموعة من الأسس التي تهدف إلى استغلال خصائص تصريف كلمات اللغة العربية (كاطِّراد قواعد الإعلال والإبدال والإدغام)، وتتعامل مع دخائله، وتتصدى لمشاكله، وتستغل وضوحه، وتتحاور مع أوجه قصوره. ونعيد هنا لنؤكد أن الصرف العربي يمثل مجالاً نموذجياً لتزاوج الحاسوب واللغة، ونوجز هنا أهم هذه الأسس:

١- يُنظر: سعيد أحمد بيومي، أم اللغات -دراسة في خصائص اللغة العربية، والنهوض بها-، ط١، ٢٠٠٢م، ص١٠٨-١٠٩.

www.alukah.net إهداء من شبكة الألوكة

١- ضرورة تعامل المحلل الصرفي الآلي لكلمات اللغة العربية ونصوصها مع «أطوار التشكيل المختلفة للنصوص العربية (تامة التشكيل، والخالية من التشكيل، والمشكولة جزئياً) لذا ينبغي أن يتوافر في المحلل الصرفي الآلي الذكاء الاصطناعي الكافي؛ لتخمين النقص في عناصر التشكيل، وتغطية جميع الاحتمالات الممكنة صرفيّاً ومعجميّاً»[١].

٢- ينبغي أن يشير المحلل الصرفي إلى التغيرات الصوتية التي حدثت في الكلمة المراد تحليلها، فعند تحليل الفعل «رَدَّ» يذكر المحلل أن أصل الكلمة هو «رَدَدَ» فحُذفت حركة عين الكلمة، وأُدغمت عينها في لامها؛ بسبب تجاور صوتين متشابهين. وعند تحليل الفعل المضارع «يَرُدُّ» يذكر المحلل الصرفي أنَّ أصله «يَرْدُدُ» فحدث إعلال بنقل حركة عين الكلمة إلى فائها الساكنة، وإدغام العين باللام بسبب تجاور متشابهين. وعند تحليل الفعل «جالَ» يذكر المحلل أنَّ أصل الفعل «جَوَلَ» فقُلبت الواو ألفاً وحُذفت حركتها. أما الفعل المضارع «يَجُوْلُ»، فأصله «يَجْوُلُ»، حدث فيه إعلال بنقل حركة عين الفعل إلى فائه.

٣- أن يفرِّق المحلل الصرفي بين الصيغة الصرفية، والميزان الصرفي.

٤- أهمية التكامل بين المحللين الآليين الصرفي والنحوي، نظراً لتداخل المستويين الصرفي والنحوي، حيث يقدِّم المحلل النحوي توقعات نحوية لنوعية الكلمات، وخصائصها حسب موقعها في الجملة، ولهذه التوقعات أهمية بالغة في تسهيل عمل المحلل الصرفي عند تحليل نصوص غير مشكولة، حيث ينحصر نطاق اللبس الصرفي في حدود الاحتمالات الصرفية المقبولة نحوياً حسب مقتضيات الجملة رهن التحليل.

٥- أن يتعامل المحلل الصرفي مع جذور اللغة العربية جميعها (الثلاثية والرباعية والخماسية).

٦- فصل قواعد المعطيات[٢] المرافقة للمحلل، والقواعد الصرفية عن برنامج التحليل؛ ليكون تعديل القواعد، وتحديث قواعد المعطيات أيسر وأسهل، حيث عانت

١- نبيل علي، اللغة العربية والحاسوب (دراسة بحثية)، تعريب، د.ط، ١٩٨٨م، ص ٢٩٩.

٢- يقصد بقاعدة المعطيات (البيانات) مجموعة من الملفات ذات الصلة ببعضها، ففي قاعدة معطيات صرفية -على سبيل المثال- يمكن أن تكون هناك عدة ملفات مترابطة مع بعضها. مثل: ملف الجذور، وملف الأفعال الثلاثية المجردة والمزيدة، وملف الأفعال الرباعية المجردة والمزيدة، والأفعال الملحقة بالرباعي...إلخ. ولعل من أهم خصائص قواعد المعطيات: الشمول، والوضوح والدقة، وقابلية التوسع والتعديل.

www.alukah.net إهداء من شبكة الألوكة

المحاولات الأولى لبناء محلل صرفي آلي للغة العربية من الخلط بين الجانب اللغوي، والجانب البرمجي.

٧- ضرورة تعريض المحلل الصرفي لتجارب مختلفة للتأكد من سلامة بناء خوارزميات التحليل، وكفاية قواعد المعطيات.

٨- ضرورة توفر عنصر الكفاءة والسرعة في المحلل الآلي.

٩- الالتزام بما خلص إليه البحث الصرفي الحديث من حيث اعتبار الكلمة (كتب، استخرج) - دون غيرها، هي أساس تصريف الأفعال، وجعل الجذر (ك ت ب) أساساً لعملية الاشتقاق، واستخدام الأساليب المنهجية الحديثة في صياغة القواعد الصرفية وتبويبها[1].

وأخيراً فإنَّه لبناء محلل صرفي دقيق لكلمات اللغة العربية ينبغي تطوير محلل صرفي آلي يفسر جميع مكونات كلمات اللغة انطلاقاً من المورفيمات التي تتكون منها الكلمة، مع الأخذ في الاعتبار كل الظواهر الصوتية التي تظهر عند كتابة الكلمة.

١٩. منتهى غايتنا عند بناء محلل صرفي حاسوبي

إنَّ منتهى غايتنا عند بناء محلل صرفي حاسوبي لتحليل كلمات اللغة العربية ونصوصها هو أنْ نهيئ للحاسوب كفاية لغوية تشبه ما يكون للإنسان حين يستقبل كلمات اللغة، ويدركها، ويفهمها، ثمَّ يحللها، ويعيد توليدها. والكفاية اللغوية الحاسوبية مرتهنة بثلاثة ضوابط، هي: ١- الضابط الإملائي. ٢- الضابط الصرفي. ٣- الضابط النحوي.

١- يُنظر: علي، نبيل، اللغة العربية والحاسوب، تعريب، ١٩٨٨ م، ص٢٩٩-٢٣١.

www.alukah.net إهداء من شبكة الألوكة

٢٠. خاتمة

وفي ختام هذا المبحث يمكننا أن تستنتج أنَّ تطوير محللات صرفية آلية للغة العربية قد أحرز تقدماً كبيراً في السنوات الأخيرة، خاصة بعد توافر مجموعة من الأدوات التي مكَّنت اللغويين والحاسوبيين من اختبار مدى كفاية المحللات الصرفية الآلية التي طوَّرتها مراكز الأبحاث والشركات.

فكان هناك عدة محاولات لتطوير نظم آلية لتحليل كلمات اللغة العربية، بعضها يفترض وجود التشكيل الكامل، والبعض الآخر يفترض غيابه بالكامل، ومعظمها يتعامل مع الميزان الصرفي، أي الشكل النهائي للكلمة، لا مع الصيغة الصرفية؛ وذلك هروباً من مشكلات الإعلال والإبدال. غير أن الهدف بناء محلل صرفي آلي متعدد الأطوار، قادر على تحليل الكلمات المشكولة كليّاً أو جزئياً، أو غير المشكولة عن طريق تطبيق أساليب الذكاء الاصطناعي، ويعتمد على الصيغة الصرفية، ويكشف عن الانحرافات الصوتية التي حدثت في الكلمة؛ كحالات الإعلال والإبدال والتضعيف.

ويُعدُّ محلل الخليل الصرفي -الذي طوَّره مخبر البحث في الإعلاميات بجامعة محمد الأول بوجده في المملكة المغربية بالتعاون مع المنظمة العربية للتربية والثقافة والعلوم (ألكسو)، ومدينة الملك عبد العزيز للعلوم والتقنية بالمملكة العربية- بحق نقلة نوعية كبيرة جداً في مجال معالجة كلمات اللغة العربية حاسوبياً؛ حيث تمكن القائمون على هذا المحلل من جعل الحاسوب يتعامل مع اللغة بشكل يحاكي الطريقة التي يستخدمها الإنسان عند تحليل كلمات اللغة العربية.

وهنا لابدَّ من التنبيه إلى أنَّ نجاح برامج معالجة اللغات الطبيعية صرفياً يتوقف على مدى وضوح النظرية اللغوية التي يتبناها مطورو المحللات الآلية، ومدى قدرتها على تحليل الكلمات والنصوص تحليلاً صحيحاً.

www.alukah.net إهداء من شبكة الألوكة

المراجع

◇ **تمام حسّان**، اللغة العربية معناها ومبناها، دار الثقافة، المغرب، ١٩٩٤م.

◇ **حسام الخطيب**، العربية في عصر المعلوماتية - تحديات عاصفة ومواجهات متواضعة، مجلة التعريب، المركز العربي للتعريب والترجمة والنشر، العدد الثاني، ١٩٩٨م.

◇ **سعيد أحمد بيومي**، أم اللغات -دراسة في خصائص اللغة العربية، والنهوض بها-، ط١، ٢٠٠٢م.

◇ **عبدالعزيز بن عبدالله المهيوبي**، إشكاليّات تطوير محلل صرفي حاسوبي دقيق للغة العربية (محلل الخليل نموذجا)، مجلة اللغة العربية وتعليمها للناطقين بغيرها، جامعة أفريقيا العالمية، العدد ٢١، ٢٠١٦م.

◇ **عبدالعزيز بن عبدالله المهيوبي**، بناء خوارزمية حاسوبية لتوليد الأفعال في اللغة العربية وتصريفها - دكتوراه' معهد تعليم اللغة العربية - جامعة الإمام محمد بن سعود الإسلامية، ١٤٣٦هـ.

◇ **عبدالله بن عبدالرحمن الزامل**، العلاقة الصرفية بين الجذور والأوزان، (الأوراق البحثية للندوة الدولية الأولى عن الحاسب واللغة العربية) الرياض- السعودية، مدينة الملك عبد العزيز للعلوم والتقنية، ٢٠٠٧م .

◇ **عبده ذياب العجيلي**، الحاسوب واللغة العربية، منشورات جامعة اليرموك- عمادة البحث العلمي والدراسات العليا، الأردن، د ط، ١٩٩٦م.

◇ **عزت جهاد عزت العجوري**، توصيف لغوي صرفي لشعر بدر شاكر السيّاب في ضوء اللسانيات الحاسوبية، رسالة ماجستير، الجامعة الهاشمية، ٢٠٠٩م.

◇ **علي السليمان الصوينع**، استرجاع المعلومات في اللغة العربية، مطبوعات مكتبة الملك فهد الوطنية، السلسلة الثانية، الرياض، ١٩٩٤م.

◇ **لطيفة النجار**، دور البنية الصرفية في وصف الظاهرة النحوية وتقعيدها، دار البشير، عمّان، ط١، ١٩٩٤م.

◇ **مازن الوعر**، دراسات لسانية تطبيقية، دار طلاس، دمشق، ط١، ١٩٨٩م.

◇ **محمود عوض المراكبي**، تطويع تقنية المعلومات لخدمة العلوم الشرعية، السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات، الرياض، ١٩٩٢م.

www.alukah.net إهداء من شبكة الألوكة

◇ **محمود مصطفى عيسى خليل**، إسناد الأفعال إلى الضمائر في ضوء اللسانيات الحاسوبي- ماجستير، كلية الدراسات العليا، الأردن، ٢٠١١م.

◇ **مروان البواب، ومحمد الطيّان**، أسلوب معالجة اللغة العربية في المعلوماتية (الكلمة - الجملة)، استخدام اللغة العربية في المعلوماتية.

◇ **مسفر محماس الدوسري**، برمجة الاسم المنسوب بياء النسب في العربية حاسوبياً- رسالة دكتوراه، جامعة اليرموك - كلية الآداب، ٢٠١٠م.

◇ **نبيل علي**، اللغة العربية والحاسوب (دراسة بحثية)، تعريب، د.ط، ١٩٨٨م.

◇ **نهاد الموسى**، العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، بيروت، ط١، ٢٠٠١م.

◇ **هدى آل طه**، النظام الصرفي للعربية في ضوء اللسانيات الحاسوبية «مثل من جمع التكسير»، رسالة دكتوراه، الجامعة الأردنية، ٢٠٠٥م.

◇ **يحيى هلال**، التحليل الصرفي للعربية، وقائع مختارة من ندوة استخدام اللغة العربية في الحاسب الآلي في الكويت، عمّان، دار الرازي.

◇◇◇

هذا الكتاب منشور في