دار المنظومة
DAR ALMANDUMAH
الرواد في قواعد المعلومات العربية

العنوان:	تحليل اللغة العربية بواسطة الحاسب الآلي
المصدر:	علوم اللغة
الناشر:	دار غريب للطباعة والنشر والتوزيع
المؤلف الرئيسي:	القحطاني، سعيد بن هادي
المجلد/العدد:	مج 5, ع 3
محكمة:	نعم
التاريخ الميلادي:	2002
الصفحات:	227 - 232
رقم MD:	134543
نوع المحتوى:	بحوث ومقالات
قواعد المعلومات:	AraBase
مواضيع:	الأفعال ، اللغة العربية ، التحليل اللغوي ، النحو ، الحاسبات الإلكترونية ، الصرف ، الدلالات اللغوية ، الحروف العربية ، المعاجم اللغوية ، المصطلحات
رابط:	http://search.mandumah.com/Record/134543

© 2021 دار المنظومة. جميع الحقوق محفوظة.
هذه المادة متاحة بناء على الإتفاق الموقع مع أصحاب حقوق النشر، علما أن جميع حقوق النشر محفوظة. يمكنك تحميل أو طباعة هذه المادة للاستخدام الشخصي فقط، ويمنع النسخ أو التحويل أو النشر عبر أي وسيلة (مثل مواقع الانترنت أو البريد الالكتروني) دون تصريح خطي من أصحاب حقوق النشر أو دار المنظومة.

تحليل اللغة العربية بواسطة الحاسب الآلي

د. سعيد بن هادى القحطانى

الرياض

مما لاشك فيه أن محاولة إخضاع اللغة (أي لغة) للتحليل بواسطة الحاسب الآلي لابد وأن يعترضها العديد من الإشكاليات والعقبات وعندما تتشابه العقبات في لغات عديدة فإنه بلاشك تتشابه طرق حلها . غير أن تحليل اللغة العربية بواسطة الحاسب الآلي يكتنفه عقبات كثيرة ، أكثر من أي لغة أخرى . ومعظم هذه المشكلات يتعلق بالجوانب التي تختلف فيها العربية عن اللغات الأوربية ، تلك اللغات التي صممت معظم البرامج الحاسوبية أصلاً لتحليلها (سلوكم و آريستر (Slocum & Aristar) ١٩٩٢م) .

ولا شك أن محاولة قولبة اللغة العربية في الحاسب الآلي من أهم المشكلات التي تعترض طريق وضع المصطلحات العربية . ولذلك يقترح (Slocum & Aristar) وجوب حصر الأوزان العربية حصراً دقيقاً وتحليلها وفق نظام تصنيفي معين ، وهو ما سيمكن من وضع رموز رياضية لها في الحاسب الآلي . ويجب أن يكون هذا التصنيف من الشمولية بحيث يستوعب الفروق بين الكلمات الناتجة عن اختلاف التشكيل . ثم يتم بعد ذلك تطوير برامج آلية يمكنها استيعاب القواعد النحوية العربية . بحيث يتمكن الحاسب الآلي من تصويب الجمل الخطأ عند قراءتها . ومن الجانب النظري ، فإن تطوير مثل هذه البرامج ليس مستحيلاً .

وقد ظهرت محاولات كثيرة لوضع برامج آلية لتحليل اللغة العربية بحيث يمكنها التعامل مع المصطلحات وفق مخزون كبير من اللغة . ولكن هذه المحاولات يعترضها العديد من العوائق التي تعود في المقام الأول إلى التركيب الصرفي للغة العربية ، بالإضافة إلى بعض هذه العوائق الناشئة عن عملية التعريب نفسها ، أي تعدد الطرق المستخدمة في التعريب وتباينها فيما بينها . وفيما يلي سأستعرض بعض القضايا المتعلقة بمحاولة تحليل اللغة العربية بواسطة الحاسب الآلي .

طبيعة النحو العربي

من المعروف أن النحو العربي ينطلق من منطلقات رياضية توازنية صرفة . ويشير المولودي (١٩٨٦م) إلى أن دراسة القوانين النحوية العربية "القواعد" مبنية بشكل أساسي على الملاحظة ، والمقارنة ، والإلحاق . والنحو العربى في مجمله مبني على المقارنة والقياس. وقد حقق اللغوي العربي الخليل بن أحمد ومن تبعه عدداً من الخطوات المهمة في تحليل النحو العربي ، خصوصاً فيما يعرف بالمجموعات البنائية ، والأصول والفروع والقياس .

وفي الواقع فإن الخليل بن أحمد ومن تبعه وضعوا مراتب مختلفة لتحليل النحو العربي ، وتحدثوا عن تلك المراتب التي لم يسبقهم أحد إلى تحليلها . ومن هذه المراتب المرتبة الوسطى للدلالات ، والعلاقة بين البناء والوصل ، وهذه الأخيرة تشبه إلى حد كبير نظرية القواعد التي تحدث عنها الأمريكي الشهير تشومسكي (١٩٧٥م) .

ولكي يتسني لنا تحليل اللغة العربية وترميزها بواسطة الحاسب الآلي فإنه لابد من الفهم العميق لهذه العلاقات في اللغة العربية بنفس القدر الذي تم بواسطته تحليل اللغة الإنجليزية وترميزها . ومما لاشك فيه أن هذه البرمجيات المستخدمة حالياً بنيت على تصنيف دقيق لقواعد اللغة الإنجليزية .

ومن أهم العقبات التي تعترض استخدام تقنيات الحاسب في تحليل اللغة العربية اختلاف بناء الجملة العربية عن اللغات الأخرى . كذلك فإن معظم التقنيات الحاسوبية المستخدمة صممت أصلاً لمعالجة الجملة الإنجليزية ؛ لذا تظل مقصورة عن التعامل مع الصيغ الصرفية العربية بكفاءة .

ويقترح " بشاى ١٩٨٢م " وضع برنامج آلي يتعامل مع الصيغ الصرفية العربية، والدلالات الرئيسية للأسماء والأفعال والحروف بدلاً من تحوير عمل البرامج المستوردة التي صممت أصلاً للغة الإنجليزية ، بحيث يقوم مثل هذا البرنامج بفرز الصيغ التي تتفق مع قواعد اللغة العربية وحسب ، ولا يسمح بإفراز صيغ خارجة عن قواعد اللغة . وهناك بعض المواصفات التي لابد من توفرها في هذه البرامج لكي تتعامل مع الخصائص النحوية التي تتميز بها العربية :

● الفعل في اللغة العربية يسبق الفاعل وبالتـالي فإن ترتيب الجملة الفعلية العربية هو فعل ، فاعل ، مفعول به .

● عندما يكون الفاعل ضـميراً فإنه يتصل بالفعل ويكون في هـذه الحالة لاحقة، وليس كلمة منفصلة كما هو الحال في اللغة الإنجليزية .

● هناك كلمات استـدراكية في اللغة العربيـة مثل "لكن" و"لإن" وهي تتطلب وجود الفاعل قبل الـفعل ، وبالتالي فإن ترتيب الجمل التي تبدأ بهـذه الكلمات يختلف عن القاعـدة التي أشير إليها آنـفاً . كما أن الأفعـال العربية تتفق مع الفـاعل من حيث العدد والجنس بعكس اللغة الإنجليزية إذ أن الفعل يبقى مجردا .

● يدل الضمير في اللغة العربيـة على الفاعل ، الذي سبق ذكره في بداية الجملة في نحو "جاء الولد الذي رأيته " فالهاء تعود على "الولد" ، بينما في اللغة الإنجليزية لا يوجد مثل هذه الخاصية .

● تتوافق الصـفات مع الأسـماء في العربيـة من حيث الجنس والعـدد ، أما في الإنجليزية فالصفات مجردة .

ومن المقترحات حـول وضع برامج آلية خاصة بتحليل اللغة العربيـة استخدام ما يطلق عليـه "بالشـبكات الآلية الـعصـبيـة" "Neural Netwerk Systems" (أحمـد وآخرون ١٩٩٢م) . وهذه الشبكات هي أحد أنظمة الذكاء الاصطناعي الذي يقوم على وضع أنماط مستمدة من عدد هائل من المدخـلات ؛ حيث يقوم الحاسب الآلي بالتعرف على أنماط مـعينة بعـد تحليل عدد كبـير من المدخـلات . وقد تنجح هذه البـرامج ولو بشكل نسبي في التعرف على الأنماط النحوية العربية ، ومن ثم التعامل معها ، ويمكنها أيضا التعرف على الأشكال المتعددة للحرف العربي وتنميطها .

الحاسب الآلي والمعجمية العربية

يلعب الحاسب الآلي دوراً رئيسيًّـا في المعجمية الحديثة ، وخصوصـاً فيما يتعلق بحفظ مراحل التطور التاريخي لمعاني المفردات . فـباستخدام الحاسب الآلي يمكننا حفظ معلومات عن كل كلمـة، بل مقاطع من النصوص التي كـانت تستخدم فيـها في حقب زمنية مختلفـة . وإذا ما تمكنا من استخدام هذه التقنية في اللغـة العربية ؛ فإنه بإمكاننا

تتبع مراحل تغيير معاني المفردات العربية عبر مراحل تاريخية مختلفة ، ومعرفة مدى تردد كل مفردة ، وتغيير الدلالات في النطق . وهذا سيوفر لنا معلومات دقيقة عن التغيرات التي تحدث لكل جذر واشتقاقاته ، وما إلى ذلك. وبعبارة أخرى يمكننا حفظ سجل شامل للغة .

وفيما يتعلق باستخدام الحاسب الآلي في المعجمية العربية على وجه الخصوص ، هناك عقبات تعترض استخدام هذه البرامج بشكل فاعل ؛ ذلك أن معظم هذه البرامج أنتجت أصلاً للتعامل مع اللغات اللاتينية والإنجليزية على وجه الخصوص . كما أن هناك عقبات أخرى تتعلق بتعامل هذه البرامج مع الحرف العربي ، ومن هذه العقبات ما يلي :

- تعدد صيغ كتابة الحرف العربي واختلاف تلك الصيغ بحسب موقع الحرف في الكلمة (أول الكلمة ، وسطها ، آخرها) .
- علامات الترقيم ، والنقاط التي توضع على الحروف (تقارب صيغها ، وتعددها) .
- بعض النقط يكون استخدامها اختيارياً .
- وجهة القراءة من اليمين إلى اليسار (عكس الحرف اللاتيني) .
- بعض النقاط لها وظائف مزدوجة .

ثم إن معظم المعاجم الأجنبية المصنفة بالحاسب الآلي مصنفة حسب الصيغة المبنية (أي صيغة الفعل ، أو الاسم المجرد) وليس بالجذر كما في المعاجم العربية . وتصنف السوابق واللواحق كمدخلات أساسية في المعاجم الحديثة بينما لا تصنف كذلك في المعاجم العربية . وبما أن اللغة العربية تحتوي على صيغ صرفية داخلية (تحدث داخل الكلمة نفسها) وليست سوابق أو لواحق ؛ فإنه يتحتم التعامل مع الجذر وليس مع كل صيغة على حدة . وهذه الخصائص التي تختص بها العربية تجعل من الصعب استقطاب البرامج الآلية الحديثة التي صممت أصلاً للتعامل مع الإنجليزية (Slocum & Aristar ١٩٩٢م) .

الحاسب الآلي وبنوك المصطلحات

لقد أصبح الحاسب الآلي ضرورة ملحة في عملية تعريب المصطلحات خصوصاً فيما يتعلق ببنوك المصطلحات العالمية وإمكانية الإفادة منها . ولن تستطيع اللغة العربية مواكبة المصطلحات العلمية والتقنية مالم توظف تقنية الحاسب في هذا المجال ؛ ففي كل

يوم يرد المئات من المصطلحات العلمية والتقنية والحرفية ، وتبث بنوك المعلومات الدولية عدداً هائلاً من أسماء المنتجات الحديثة كل يوم (حجازي ١٩٩٢م) .

ويقترح الحمزاوي (١٩٩٢م) خمس خطوات رئيسية للإفادة من تقنية الحاسب في هذا المجال :

- وضع منهجية موحدة لعملية وضع المصطلح.
- وضع منهجية موحدة للترجمة .
- وضع نظام محدد للتقييس .
- وضع نظام آلي لمعالجة المصطلحات . (١٩٩٢م ، ص.٦)

ونظام التقييس هذا من شأنه القضاء على الازدواجية في وضع المصطلحات التي تختلف من قطر إلى آخر في الدول العربية ، فعندما يكون هناك نظام موحد مبني على منهجية موحدة ؛ فإنه بلا شك سيدفع عملية تعريب المصطلحات إلى وضع أفضل بكثير مما هو عليه الآن .

ويقترح بعض المهتمين بعملية تعريب المصطلح أيضاً وضع نظام للتقييس وفق الأسس التالية :

- درجة شيوع المصطلح : أي أن المصطلح المستخدم في أكثر من عشرة مصادر هو الأولى بالاستخدام ؛ لأنه الأكثر شيوعاً ، بينما المصطلح المستخدم في مصدر أو مصدرين هو بلا شك أقل شيوعاً .
- مبدأ الاقتصاد : بحيث يكون المصطلح قصيراً ومعبراً .
- مبدأ المواءمة : أي تواءم المصطلح مع المعنى المقصود بشكل مباشر .
- مبدأ الإنتاجية : أي أنه يمكن اشتقاق صيغ أخرى من المصطلح .

والخلاصة أن استخدام الحاسب الآلي في تحليل اللغة العربية أضحى ضرورة ملحة لابد من حشد كل الطاقات لإنتاج برامج حاسوبية قادرة على تحليل الجملة العربية وترميزها ، لكي تتم الإفادة من خدمات بنوك المعلومات الآلية المنتشرة في العالم ، ومن ثم نقل المصطلحات الأجنبية إلى العربية آلياً .

المراجع

أحمـد وآخرون (١٩٩٢م). "Arabic Script Recognition" . . ندوة استـخدام العربيـة في تقنية المعلومات. مكتبة الملك عبد العزيز ، الرياض .

بشاي (١٩٨٢م). 'Linguistic Problem in Computer Aided Translation from English to Arbic' دراسات في الترجمة الآلية. محمود صالح. الرياض .

حجازي (١٩٩٢م). "الحاسب الآلي والمعجمية العربية". ندوة اسـتخدام العربية في تقنية المعلـومات مكتبة الملك عبد العزيز ، الرياض .

الحمزاوي (١٩٩٢م) . "نحو نظرية مصطلحية عربية" . . ندوة اسـتخدام العربية في تقنية المعلومات. مكتبة الملك عبد العزيز، الرياض .

المولودي، (١٩٨٦). 'Arabic Language Planing : the case of Lexical Modernization' . رسالة دكتوراه ، جامعة جورج تاون ، الولايات المتحدة الأمريكية .

رقم الإيـداع ٦٨١٥