

تحليل اللغة العربية بواسطة الحاسوب

د. سعد بن هادي القحطاني

مركز اللغة الإنجليزية - معهد الإدارة/ الرياض

مما لا شك فيه أن محاولة إخضاع اللغة (أي لغة) للتحليل بواسطة الحاسوب لا بد وأن يعترضها العديد من الإشكاليات والعقبات. وعندما تنتشابه العقبات في لغات عديدة فإنه بلا شك تنتشابه طرق حلها. غير أن تحليل اللغة العربية بواسطة الحاسوب يكتنفه عقبات كثيرة، أكثر من أي لغة أخرى. ومعظم هذه المشاكل متعلقة بالجوانب التي تختلف فيها العربية عن اللغات الأوروبية، تلك اللغات التي صممت معظم البرامج الحاسوبية أصلاً لتحليلها (سلوكم وأريستر (Slocum & Aristar) ١٩٩٢م).

ولا شك أن محاولة قولبة اللغة العربية في الحاسوب من أهم المشاكل 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 الحاسوب. 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 الحاسوب من تصويب الجمل الخاطئة 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 يأتي سأستعرض بعض القضايا المتعلقة بمحاولة تحليل اللغة العربية بواسطة الحاسوب.

طبيعة النحو العربي

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

وفي الواقع أن 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 التحويلية التي تحدث عنها اللغوي الأمريكي الشهير تشومسكي (١٩٧٥).

ولكي يتسنى لنا تحليل اللغة العربية وترميزها بواسطة الحاسوب فإنه لا بد من الفهم العميق لهذه العلاقات في اللغة العربية بنفس القدر الذي تم بوساطته تحليل اللغة الإنجليزية وترميزها. ومما لا شك فيه أن هذه البرمجيات المستخدمة حالياً بنيت على تصنيف دقيق لقواعد اللغة الإنجليزية.

ومن أهم العقبات التي تعترض استخدام تقنيات الحاسوب في تحليل اللغة العربية هو اختلاف بناء الجملة العربية عن اللغات الأخرى. وحيث إن معظم التقنيات الحاسوبية المستخدمة صممت أصلاً لمعالجة الجملة الإنجليزية؛ فإنها تظل قاصرة عن التعامل مع الصيغ الصرفية العربية بكفاءة.

ويقترح "بشاي ١٩٨٢م" وضع برنامج آلي يتعامل مع الصيغ الصرفية العربية، والدلالات الرئيسية للأسماء والأفعال والحروف بدلاً من تحويل عمل البرامج المستوردة التي صممت أصلاً للغة الإنجليزية، بحيث يقوم مثل هذا البرنامج بفرز الصيغ التي تنفق مع قواعد اللغة العربية وحسب، ولا يسمح بإفراز صيغ خارجة عن قواعد اللغة.

وهناك بعض المواصفات التي لا بد من توافرها في هذه البرامج لكي تتعامل مع الخصائص النحوية الآتية التي تتميز بها العربية:

- الفعل في اللغة العربية يسبق الفاعل وبالتالي فإن ترتيب الجملة الفعلية العربية هو فعل، فاعل، مفعول به.
 - عندما يكون الفاعل ضميراً فإنه يتصل بالفعل ويكون في هذه الحالة لاحقة، وليس كلمة منفصلة كما هو الحال في اللغة الإنجليزية.
 - هناك كلمات استدرابية في اللغة العربية مثل "كن" و"إن" وهي تتطلب وجود الفاعل قبل الفعل، وبالتالي فإن ترتيب الجمل التي تبدأ بهذه الكلمات يختلف عن القاعدة التي أشير إليها آنفاً. كما أن الأفعال العربية تتفق مع الفاعل من حيث العدد والجنس بعكس اللغة الإنجليزية إذ إن الفعل يبقى مجرداً.
 - يدل الضمير في اللغة العربية على الفاعل، الذي سبق ذكره في بداية الجملة في نحو "جاء الولد الذي رأيتَه" فالهاء تعود على "الولد"، بينما في اللغة الإنجليزية لا يوجد مثل هذه الخاصية.
 - تتوافق الصفات مع الأسماء في العربية من حيث الجنس والعدد، أما في الإنجليزية فالصفات مجردة.
- ومن المقترحات حول وضع برامج آلية خاصة بتحليل اللغة العربية استخدام ما يطلق عليه بالشبكات الآلية العصبية "Neural Network Systems" (أحمد وآخرون ١٩٩٢م). وهذه الشبكات هي أحد أنظمة الذكاء الاصطناعي الذي يقوم على وضع أنماط مستمدة من عدد هائل من المدخلات؛ حيث يقوم الحاسوب بالتعرف على أنماط معينة بعد تحليل عدد كبير من المدخلات. وقد تتجح هذه البرامج ولو بشكل نسبي في التعرف على الأنماط النحوية العربية، ومن ثم التعامل معها. ويمكنها أيضاً التعرف على الأشكال المتعددة للحرف العربي وتمييزها.

الحاسوب والمعجمية العربية

يلعب الحاسوب دوراً رئيساً في المعجمية الحديثة، وخصوصاً فيما يتعلق بحفظ مراحل التطور التاريخي لمعاني المفردات. فباستخدام الحاسوب يمكننا حفظ معلومات عن كل كلمة، بل مقاطع من النصوص التي كانت تستخدم فيها في حقبة زمنية مختلفة. وإذا ما تمكنا من استخدام هذه التقنية في اللغة 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 على كل جذر واشتقاقاته، وما إلى ذلك. وبعبارة أخرى يمكننا حفظ سجل شامل للغة.

وفيما يتعلق باستخدام الحاسوب 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 يأتي:

- 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، وتعددتها).
- 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 بالحاسوب مصنفة حسب الصيغة المبنية (أي صيغة الفعل، أو الاسم المجرد) وليس بالجذر كما في المعاجم العربية. وتصنّف 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).

الحاسوب وبنوك المصطلحات

لقد أصبح الحاسوب ضرورة ملحة في عملية تعريب المصطلحات، خصوصاً فيما يتعلق ببنوك المصطلحات العالمية وإمكانية الإفادة منها. ولن تستطيع اللغة العربية مواكبة المصطلحات العلمية والتقنية ما لم توظف تقنية الحاسوب في هذا المجال؛ ففي كل يوم ترد المئات من المصطلحات العلمية والتقنية والحرفية، وتبث بنوك المعلومات الدولية عدداً هائلاً من أسماء المنتجات الحديثة كل يوم (حجازي ١٩٩٢م).

ويقترح الحمزواي (١٩٩٢م) خطوات رئيسة للإفادة من تقنية الحاسوب 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 الحاسوب في تحليل اللغة العربية أضحي ضرورة ملحة لا بد من حشد كل الطاقات لإنتاج برامج حاسوبية قادرة على تحليل الجملة العربية وترميزها، لكي تتم الاستفادة من خدمات بنوك المعلومات الآلية المنتشرة في العالم، ومن ثم نقل المصطلحات الأجنبية إلى العربية آلياً.

المراجع

- أحمد وآخرون (١٩٩٢م). "Arabic Script Recognition". ندوة استخدام العربية في تقنية المعلومات. مكتبة الملك عبد العزيز، الرياض.
- بشاي (١٩٨٢م). Linguistic Problem in Computer Aided Translation From English to Arabic. دراسات في الترجمة الآلية. محمود صالح. الرياض.
- حجازي (١٩٩٢). "الحاسب الآلي والمعجمية العربية". ندوة استخدام العربية في تقنية المعلومات. مكتبة الملك عبد العزيز، الرياض.
- الحمزوي (١٩٩٢م). "نحو نظرية مصطلحية عربية". ندوة استخدام العربية في تقنية المعلومات. مكتبة الملك عبد العزيز، الرياض.
- المولودي، (١٩٨٦). "Arabic Language Planning: the case of Lexical Modernization". رسالة دكتوراه، جامعة جورج تاون، الولايات المتحدة الأمريكية.