

المحاضرة الأولى

مشاريع حوسبة علوم اللغة

تسمى البرامج التي تعمل على إدخال البيانات الصوتية والصرفية والنحوية والمعجمية والدلالية إلى الحاسب بمشاريع حوسبة علوم اللغة ، وقبل الحديث عن هذه المشاريع وما تقوم به في مجال الحوسبة ينبغي الإشارة إلى عدة عناصر ، وذلك من خلال طرح بعض الأسئلة والإجابة عنها :

يتعلق السؤال الأول بما تم إنجازه في العالم من حوسبة للغات ، بمعنى : هل استطاعت لغة ما حوسبة علومها اللغوية بشكل كامل ؟

ج : نعم ، استطاعت كثير من اللغات إنجاز مشاريع مهمة في مجال حوسبة علومها ، وبخاصة اللغة الإنجليزية التي نجحت في إكساب الحاسوب القدرة على أن يتعامل معها بطريقة مماثلة للعقل البشري ، وذلك في مجالات الترجمة ، والتدقيق الإملائي والنحوي ، والاشتقاق الصرفي وتحويل النصوص المكتوبة إلى منطوقة والعكس ، وبهذا الإنجاز تمكن المتحدثون بالإنجليزية من التواصل بنجاح مع الحاسوب دون أية معوقات .

س : برأيك ، ما أسباب نجاح حوسبة علوم اللغة الإنجليزية ؟

ج : يعد السبب الأساسي لهذا النجاح ، هو أن برامج الحاسوب صممت في الأصل لخدمة اللغة الإنجليزية .

س : هل حققت اللغة العربية في حوسبة علومها ما حققته اللغة الإنجليزية ؟

ج : بالطبع لا .

س : علل : تأخر اللغة العربية في حوسبة علومها على الرغم مما تتمتع به من سمات لغوية رياضية تتناسب مع الحاسوب .

ج : يعود السبب إلى أمرين : الأول : انصراف كثير من اللغويين العرب عن هذا المجال الذي يتطلب العمل بروح الفريق ، ولن يتم إنجازه بالجهود الفردية ، والثاني : تنوع الخبرات التي يحتاجها إنجاز مثل هذا العمل .

س : ما أنواع الخبرات المطلوبة لحوسبة علوم اللغة العربية ؟

ج : يتطلب الأمر نوعين من الخبرات :

أحدهما : الخبرة اللغوية الجيدة بنظام اللغة العربية في جميع مستوياتها الصوتية والصرفية والنحوية والمعجمية والدلالية ، والقدرة على رصد كل وجوه الاستعمال اللغوي الممكنة للظاهرة اللغوية الواحدة ثم تحليلها ، وتجريدها ، وجعلها ذات طبيعة رياضية لا تقبل الظن .

والآخر : الخبرة النظرية والعملية بنظام عمل الحاسوب ، فالحاسوب يعمل بلغة رياضية صارمة لا تقبل الظن أو النسبة (التغير) ، لذا يجب صياغة القاعدة اللغوية في صورة حاسمة لا تقبل إلا بتأويل واحد لكل قضية أو ظاهرة لغوية .

وعلى أية حال قد يبدو أمر حوسبة علوم العربية صعبًا ، ولكنه ليس مستحيلًا .

نخلص مما سبق إلى أن مشاريع حوسبة علوم اللغة هي برامج تبني للحاسوب قاعدة بيانات لغوية كاملة تمكنه من التعامل الصحيح مع اللغة في مستوياتها التحليلية المعروفة (الصوتية والصرفية والنحوية والمعجمية والدلالية) .

ولكل مشروع من هذه المشاريع شقان ، أحدهما تحليلي يتم فيه إدخال البيانات للحاسب ليقوم بالتعرف عليها وتخزينها للتعامل معها ، والآخر توليدي يتمثل في كيفية إنتاج الحاسب للغة وفق ما يُطلب منه في ضوء المعطيات التي أدخلت إليه .

وعليه فإن مشاريع حوسبة علوم اللغة تتمثل إجمالًا في أربعة مشاريع ، هي :

- * مشروع المحلل والمولد الصوتي الحاسوبي .
- * مشروع المحلل والمولد الصرفي الحاسوبي .
- * مشروع المحلل والمولد النحوي الحاسوبي .
- * مشروع المحلل والمولد المعجمي والدلالي الحاسوبي .

أولاً : مشروع المحلل والمولد الصوتي الحاسوبي :

(أ) : المحلل الصوتي الحاسوبي :

تحليل الأصوات العربية بصفة عامة يعني فحصها وتصنيفها إلى أقسامها المتعددة باعتبارات مختلفة ، كاعتبار المخرج ، أو اعتبار الصفات من همس وجهر ، وانفجار واحتكاك ، وترقيق وتفتيح ، ونحو ذلك ...

وعليه فمشروع المحلل الصوتي الحاسوبي يعني : إكساب الحاسوب القدرة على التعامل مع الأصوات العربية وفحصها وتصنيفها وتحليلها ، كما تتعامل معها الأجهزة الصوتية في معامل الأصوات ؛ بحيث يستطيع التمييز بين مخارج الأصوات وصفاتها (في الإملاء الآلي عليه) .

ومن أبرز مهام المحلل الصوتي الحاسوبي التعرف على شخصية المتكلم الذي سبق تخزين أنماطه الصوتية (بصمة الصوت) .

(ب) : المولد الصوتي الحاسوبي :

التوليد الصوتي يعني إنتاج الأصوات بعد استيعاب الحاسوب لأصوات اللغة وقدرته على التمييز بينها ، وبمعنى آخر ، المولد الصوتي الحاسوبي هو : تمكين الحاسوب من أن ينطق أصوات اللغة العربية كما ينطقها العربي (في التحدث الآلي – الناطق الآلي) .

وتنقسم عملية توليد الحاسب للأصوات إلى قسمين :

الأول : نظم عدد محدد من المفردات ، وتخزينها منطوقة على الجهاز في شكل جمل قصيرة وربطها بمواقف معينة يتم استدعاؤها فيها (من ذلك الرسائل الصوتية المسجلة آلياً كتلك التي تنبه الراكب إلى دعاء ركوب السيارة عند تشغيلها ، أو تنبيه السائق برسالة صوتية إلى قرب نفاذ الوقود ، أو الرسائل الصوتية المسجلة في بعض لعب الأطفال ، ...) ، وهذا القسم من التوليد الصوتي يسهل تنفيذه .

الثاني : نظم عدد غير محدد من المفردات وتخزين طريقة نطق كل مفردة على حدة ، لإكساب الحاسوب القدرة على توليد الصوت اللغوي ، وتمكينه من القراءة الآلية لأي جملة أو نص يتم إدخاله إليه ، وهذا القسم يصعب تنفيذه لأنه يتطلب إدخال البيانات الصوتية لكل مفردات اللغة فضلاً عن لا نهائية الجمل التي يمكن تشكيلها من هذه المفردات ، بخلاف القسم الأول الذي يعتمد على مفردات وجمل ثابتة .

ويمكن إجمال ما يستفاد من مشروع المولد الصوتي الحاسوبي في نقاط ستة (للتفصيل انظر الكتاب صفحة ٣٣)

*** متطلبات إنجاز مشروع المولد الصوتي الحاسوبي :**

لبناء نظام المولد الصوتي الحاسوبي لا بد أن يمتلك الحاسوب قاعدة بيانات كاملة عن أصوات اللغة المراد نطقها ، وتتكون قاعدة البيانات من

- ملفات صوتية سبق أن سُجّلت لكلام متحدثين باللغة العربية .

- كتابة صوتية للكلام المسجل .

ويتوقف نجاح هذا المشروع على شمول قاعدة البيانات ودقتها وجودتها .

* ما تم إنجازه في مشروع المحلل والمولد الصوتي الحاسوبي : (ينظر تفصيل ما قدمه الباحثان د / فانت حجازي ، م / شادي دلية ، في الكتاب من ص ٣٤ : ص ٤٠)

ثانياً : مشروع المحلل والمولد الصرفي الحاسوبي :

(أ) : المحلل الصرفي الحاسوبي :

تعريفه : يقصد به تمكين الحاسوب من أن يقوم بتحليل الكلمات المغذاة إليه إلى عناصرها الأولية : ببيان الجذر اللغوي والسوابق واللواحق ، وما طرأ على الكلمة من تغيرات صرفية ومن ثم بيان ميزانها الصرفي .

مهمته : تتضح مهمة المحلل الصرفي الحاسوبي في تحديد نوع الكلمة (اسم – فعل – حرف) وتحديد وزنها (ثلاثي – رباعي ...) وبالتالي تحديد نوع الجملة (اسمية – فعلية – شبه جملة)

س : علل : للمحلل الصرفي الحاسوبي أهمية في إنجاز مشروع المحلل والمولد النحوي الحاسوبي .

ج : لأن المحلل الصرفي الحاسوبي يهتم بتحديد نوع الكلمة ، هذا التحديد يترتب عليه كثير من الأمور المتعلقة بالتحليل النحوي ، وبيان الوظيفة النحوية للكلمة وعلامتها الإعرابية .

* مثال للتحليل الصرفي الحاسوبي (انظر في الكتاب طريقة تحليل الفعل : يضربون)

(ب) : المولد الصرفي الحاسوبي :

تعريفه : هو عملية عكسية لعملية التحليل ، ويقصد به تمكين الحاسوب من توليد الاشتقاقات الكثيرة من الجذر اللغوي الأصلي المدخل إليه ، كتوليد بناء فاعل أو مفعول من مادة (ك ت ب)

مثال : * الجذر اللغوي : ك ت ب

* الصيغة الصرفية المطلوبة : اسم المفعول

* الكلمة : مكتوب

* حالات الكلمة : اسم ، مفرد ، نكرة ، مثبت .

س : علل : ثبت نجاح مشروع المحلل والمولد الصرفي بدرجة عالية مقارنة بالمشاريع الأخرى

ج : لأن اللغة العربية في باب الصرف (المستوى الصرفي) تتميز بكثير من السمات الرياضية الصورية التي تمكن من معالجتها معالجة حاسوبية صرفية ، فهي تقوم على مكونين رياضيين

هما : **الجزر** (المادة الأصلية نحو : ك ت ب ، ج ل س ، ن ص ر) وهو كما تراه متعددًا ويتولى وضع البنية الأساسية للكلمة .

الوزن (الميزان الصرفي المتمثل في فعل) ، وهو ثابت (فكل من كتب وجلس ونصر مثلاً على وزن واحد هو : فَعَلَ) ، ويقوم الوزن بتوزيع الحركات على مختلف حروف الكلمة ، كما يقوم بتوزيع السوابق واللواحق التي تضاف إلى الجذر بهدف توليد كلمات جديدة .

وهذا التشكيل الرياضي للغة العربية من الجذر والوزن جعل منها لغة انصهارية (اشتقاقية) والقول بالانصهارية يؤدي حتمًا إلى القول بالطبيعة الرياضية الجبرية للغة العربية .

* متطلبات إنجاز مشروع المحلل الصرفي الحاسوبي : (ينظر التفصيل ص ٤٥ و ص ٤٦)

- ١- تقديم وصف علمي دقيق لكل الظواهر الصرفية ، كل على حدة .
- ٢- لا بد أن تكون مناهج التوصيف السابقة واحدة وصورية قابلة للحوسبة ز
- ٣- يجب أن تكون الظواهر المدروسة مزودة بقوائم الكلمات الموجودة على هذا النحو في المعاجم القديمة والحديثة .
- ٤- كما يجب أن تكون الأوزان مزودة بالسلمات الفارقة (أي الحركات التي تفرق بين أوزان قد تشترك بين أكثر من نوع نحو : أفعل للتفضيل وأفعل للتعجب ...)
- ٥- وضع قوائم للكلمات الموجودة على هذا النحو في المعاجم العربية القديمة والحديثة .
- ٦- رصد كثير من السياقات التي وردت فيها هذه الأبنية .
- ٧ - تزويد الحاسوب بكل هذه المعلومات في شكل برامج تستفيد من كل المعطيات السابقة لبناء محلل صرفي حاسوبي للعربية .

* ما تم إنجازه في المشروع الصرفي الحاسوبي : هو أكثر المشروعات التي تم فيها محاولات جادة لحوسبتها لاتفاقها مع الطبيعة الرياضية للحاسوب ، وقد أنجزت فيه بالفعل عدة مشاريع وصل بعضها إلى درجة عالية من الدقة ، ولكن أكثرها يحتاج إلى مزيد من المراجعة والتدقيق .

(انظر تفصيل ما أنجزته شركة صخر في هذا المشروع من ص ٤٣ : ص ٤٧)

تمت المحاضرة ، والحمد لله أولاً وآخراً