المحاضرة الأولى: اللغة العربية والمعالجة الآلية

هل حقا تسبب طبيعة اللغة العربية صعوبة في معالجتها آليا وتحول دون تفاعل الحاسوب معها؟

يقول "نبيل علي" أن اللغة العربية أشد حاجة من غيرها إلى هندسة اللغة والمعالجة الآلية بسبب كثرة الفجوات في تنظيرها اللغوي الراهن، حيث تستطيع المعالجة الآلية بأساليبها العملية التجريبية سد جزء من هذا الفراغ، وإذا لم يحدث ذلك، سيطول الوقت انتظارا لاكتمال الأسس النظرية لمعالجة اللغة العربية الآلية، فهندسة اللغة العربية وتطوير التنظير لها لابد أن يسيرا جنبا إلى جنب؛ لأن كل منهما يتغذى على نتاج الآخر، وحسب رأيه فإن المعالجة الآلية للغة العربية بواسطة الحاسوب تنقسم إلى جانبين أساسيين:

- ✓ الأول يشمل الأنظمة البرمجة المستخدمة في المعالجة الآلية بواسطة الحاسوب للمستويات اللسانية المختلفة(الصوتي- الصرفي-المعجمي-النحوي-الدلالي)؛ أي مختلف الأنظمة التي تكون أساسا لأي معالجة لسانية آلية؛ مثل قواعد البيانات المعجمية والذخائر اللغوية والقواميس الإلكترونية، ونظام المعالجة الصوتية،والمعالج الصرفي الآلي، ونظام الإعراب الآلي الذي يقوم بإعراب الجمل آليا ونظام التحليل الدلالي الآلي، الذى يستخلص معانى الكلمات استنادا إلى سياقها، ويحدد معانى الجمل استنادا إلى ما يسبقها وما يتبعها من جمل.
- ✓ والجانب الثاني يتضمن التطبيقات (Les applications)التي تقوم على الأنظمة اللغوية الآلية السابقة الذكر، والتي تشمل على سبيل المثال: الترجمة الآلية، والتدقيق الاملائي والنحوي، والفهرسة، والتلخيص الآلي، وتحليل النصوص، وفهم الكلام ونطقه آليا.

1-خصائص اللغة العربية وعلاقتها بالمعالجة العربية:

يقول نبيل على في كتابه (العرب وعصر المعلومات) أن منظومة اللغة العربية تتصف بمجموعة من الخصائص التي إما تسهم في تسهيل وإما في تعقيد معالجتها آليا، بشكل يتفاوت من خاصية لأخرى، على النحو التالي:

1-1-خاصية الوسطية والتوازن: تتسم اللغة العربية بالتوازن في معظم خصائصها ، وهذا ما يجعلها تتصف بالوسطية وعدم المغالاة في اعتماد الظواهر والقواعد اللغوية، فلو تأملنا فكرة تغيير الموقع الإعرابي للكلمة في وسط الجملة (التقديم والتأخير) تبعاً للمعنى المقصود وما يلحقه من تغير في الحالة الاعرابية وحتى في حروفها، لأدركنا أن العربية يمكن أن تخرج عن القواعد العامة إلى قواعد خاصةبسياقات معينة، وهذا يعنى أن أمور معالجتها آليا ستتعرض للعديد من القضايا والصعوبات، وهذا يستوجب بذل جهود توصيفية كبيرة من أجل تذليل تلك الصعوبات.

1-2-خاصية الاشتقاق الصرفي:فاللغة العربية بناء رياضي فريد لا يتوافر لأي لغة أخرى، فأسلوب الاشتقاق من الأفعال والأسماء بنسق رياضي دقيق يتيح لها التوالد واستيعاب أي مصطلح جديد والتعبير عنه بطريقة دقيقة،.

1-3- خاصية الإلصاق وإضافة إلى **الاشتقاق الصرفي** تتمتع اللغة العربية بخاصية الإلصاق التي تشتهر بها اللغات الإلصاقية كالانجليزية والفرنسية، فالكلمات العربية أيضا تقبل إضافة السوابق واللواحق (Préfixes- Suffixes) لأي كلمة بما يضيف إلى معناها ولا ينتقص منه ، وذلك وفق قوانين رياضية واضحة.

تمثل تركيبات السوابق واللواحق في اللغة العربية ثراء مفرداتيا كبيرا، فيمكن من خلالها التعبير بالكلمة نفسها عن الإفراد والتثنية وضمائر الفاعل والمفعول والأزمنة والاستفهام ... الخ. ولهذه الخاصية أثر كبير في تسهيل معالجتها آليا.

1-4-خاصية الانتظام في القواعد الصوتية العربية، حيث يتسم نظام مقاطعها الصوتية ونبرها بالبساطة، فلابد أن تجمع المقاطع بحرف صامت ولا تتضمن أكثر من صامتين، وهذه الخاصية ذات أهمية كبرى في توليف الكلام العربي وتمييزه ، والتعرف عليه آليا.

1-5- الترابط بين العناصر اللغوية: ويمس هذا الارتباط أغلب المستويات اللغوية (الصوتي – المعجمي-النحوي الدلالي)، إذ تتأثر العناصر اللغوية سواء كانت أصواتا أو كلمات بما سبقها أو ما يليها؛ فعلى المستوى الصوتي قد تختلف الصورة النطقية لبعض الأصوات بحسب ما قبلها وما بعدها؛ ومثال ذلك "مسطرة- -مسيطر" حيث تنطق (السين) (كأنها (صاد) . وحتى شكل الحروف يعتمد على رتبته في الكلمة(بدايتها -وسطها-آخرها)؛ فحرف الثاء يكتب(ثـ - ـثـ - ـث)

أما على مستوى النحو يتمثل هذاالارتباط في علاقات المطابقة؛ مثل تطابق الصفة مع الموصوف والعدد مع المعدود، مما يجعل معالجة اللغة العربية آليا أكثر تعقيدا وصعوبة .

1-6- التداخل بين المستويات اللغوية: ومن أمثلة ذلك علاقة النحو بالصرف وشدة الصلة بين أشكال الصيغ الصرفية ومعانيها. وهذا التماسك سلاح ذو حدين على حد قول نبيل علي، فهو من جانب يزيد من صعوبة النظم الآلية حيث يصعب فصل المعالجات الآلية للنحو الآلي عن المعالجات الآلية للصرف والمعجم ، ومن جانب آخر يعمل هذا التماسك على إجلاء اللبس الناجم عن غياب التشكيل حيث توفر قرائن لغوية تكشف اللبس.

1-7- اعتماد المعجم العربي على الجذور وليس على الترتيب الأبجدي للكلمات، فرغم صغر نواة المعجم العربي (اقل من 10 آلاف جذر) تتعدد المفردات بصورة هائلة بفضل خاصية الاشتقاق الصرفي. فمثلايقدر إجمالي عدد الكلمات باللغة الإنجليزية بين أربعمئة إلى خمسمئة ألف كلمة مستخدمة. بينما تتعدى الكلمات المستخدمة في اللغة العربية عدة ملايين (أخذاً في الاعتبار الكلمات المشتقة والمزيدة). في الوقت نفسه فإن جذور الكلمات العربية أقل بكثير من مثيلاتها الإنجليزية، ولكنها تخضع لجداول منظمة يمكن من خلالها الاشتقاق المباشر لهذا العدد الكبير جداً.

1-8-غياب تشكيل الكلمات: يقول "علاء الدين صلاح العجماوي""فلا يخلو الأمر من بعض نواحي القصور في أسلوب الكتابة العربية الحديثة، قصوراً قد يؤدي إلى الكثير من السلبيات عند محاولة تطبيق التقنيات الحديثة في أساليب المعالجة الآلية للغة العربية، فقلّة حروف الحركة

(Vowels) العربية، والاعتماد الكبير على حركات التشكيل من فتح وضم وكسر التي قلما تستخدم في النصوص العربية، يؤثر سلبا على تقنيات المعالجة الآلية.

فلو قارنا اللغة العربية بلغة مثل الإنجليزية مثلاً، لوجدنا أن الإنجليزية تحتوى على سبعة حروف متحركة أساسية(a, e, i, o, u, y, w) مقارنة بثلاثة حروف عربية (أ، و، ي)، كما أن تركيبات حروف الحركة الإنجليزية كثيرة جداً (مثل ea, ee, ei, ie, oa, ui) مقارنة بندرة تركيب حروف الحركة العربية، وهذا الأمر مرده أساساً إلى اعتماد العربية على حركات التشكيل، وبكل أسف فإن الكتابات العربية الحديثة لا تستخدم حركات التشكيل إلا فيما ندر."

ينتج عن ذلك تشابه مفردات عربية كثيرة في حروفها الأساسية مع اختلاف بَين في معناها (مثل كتب تُقرأ كَتَبَ أو كُتُب)،

وكمثال عملي لنفترض مدققاً نحوياً يفحص عبارة "كتب المدرسين" هل هي خطأ أم صواب، بالتأكيد سوف يأخذ في الاعتبار حالة الفعل والفاعل وهي خطأ(كَتَبَ المدرسين)، وحالة المضاف والمضاف إليه وهي صواب(كُتُبُ المدرسين)، وللمترجم الآلي أهي"Teachers' books" أو " Teachers wrote".

يمثل غياب التشكيل عقبة أساسية أمام معالجة النصوص العربية آليا، ولهذا تحتاج معالجة النصوص العربية آليا إلى وسيلة برمجية لتشكيل النصوص تلقائيًا.

مراجع المحاضرة :

- نبيل علي: الثقافة العربية وعصر المعلومات، سلسلة عالم المعرفة ،عدد184 ، الكويت، أفريل 1994.

- علاء الدين صلاح العجماوي :المعالجة الآلية للغة العربية بين الواقع والتحديات https://arabic.jo/res/seasons/19/19-2.doc(فيفري2021)

-عمر مهديوي: الهندسة اللغوية والترجمة الآلية: المفهوم والوظيفة https://allissan.org/node/1316(فيفري2021)