المحاضرة الرابعة: الهندسة اللغوية والمعالجة الآلية للغات الطبيعية

1- مفهوم الهندسة اللسانية (Ingenierie linguistique):

لم ترتبط اللغة في علاقتها بالحاسوب مع الذكاء الاصطناعي فقط بل أيضا بالهندسة، إذ «تمثل اللغة موضوعا متميزا للتساؤل الهندسي بصفتها فن السيطرة على النظم المعقدة، وهكذا ظهر إلى الوجود مصطلح هندسة اللغة كفرع من فروع هندسة المعرفة والذكاء الاصطناعي»، نتيجة عمليات المصاهرة التي حصلت بين العلوم العرفانية والمعلوماتية، وفي مقدمتها هندسة الحواسيب وهندسة المعرفة بصفة عامة.

لم يظهر هذا المصطلح إلا في وقت قريب مقارنة بمصطلحات اللسانيات الحاسوبية والمعالجة الآلية والذكاء الاصطناعي، وهذا ما يؤكد حاجة هذه التخصصات إلى الهندسة اللغوية وارتباطها به.

وتجدر الإشارة إلى أن المصطلح الأجنبي (Ingenierie linguistique)، قد ترجم إلى العربية بمقابلات عديدة نذكر منها الهندسة اللغوية –الهندسة اللسانية– هندسة اللغويات، ولكن نختار تبني مصطلح الهندسة اللسانية للدلالة على العلم وللدلالة أيضا على عملية بناء الهياكل المنطقية وتمثيل العلاقات التي تربط الوحدات اللسانية، وهنا نشير إلى أن هذا المصطلح يميل الكثير من الباحثين إلى تعريفه بأنه فرع معرفي أو علم، في حين يستخدم أيضا للدلالة على عملية تمثيل اللغة آليا أو صوريا (أي على شكل رمز رياضية منظمة في معادلات وهياكل).

عرف"مازن الوعر" الهندسة اللسانية بقوله: «إنما العلم الذي يبحث في اللغة البشرية كأداة طيعة لمعالجتها في الآلة، وتتألف مبادئ هذا العلم من اللسانيات العامة بجميع مستوياتها التحليلية: الصوتية والنحوية، والدلالية، ومن علم الحاسبات الإلكترونية (الكمبيوتر)، ومن علم الذكاء الاصطناعي، وعلم المنطق، ثم علم الرياضيات».

وبغض النظر عن كون الهندسة اللسانية علما أو عملية، فإنها تتطلب نوعين من المعرفة هما:

–المعرفة اللسانية العميقة وصفا وتصنيفا، بمختلف جزئيات النظام اللغوي، على ضوء أحدث النظريات اللسانية المعاصرة، وبخاصة اللسانيات الصورية(*).

---

*– اللسانيات الصورية (linguistique formelle): هي الدراسة التي تقوم بإنشاء نماذج لسانية تستعمل فيها لغات صورية (رموز ورسوم) من أجل تمثيل العبارات، وضبط اللغة عن طريق وصف الأشكال اللسانية وتجريدها ثم تمثيلها في نماذج يمكن القياس عليها، تعتبر نتائج اللسانيات الصورية هي المادة الخام التي تبدأ منها الهندسة اللسانية الآلية، هذه الأخيرة التي تقوم بصورنة تلك النماذج اللسانية.

-**المعرفة الحاسوبية** ذات الصلة بمعالجة اللغات الطبيعية، وخاصة في جانبها البرمجي؛ أي التعامل المنظم والمنطقي مع الآلة أو الحاسوب.

والحاسوب هو عبارة عن منظومة خوارزمية، ذات قواعد صورية لا تقبل التأويلات والاحتمالات، والجمع بين هذين النوعين من المعرفة (اللسانية والحاسوبية) هو المحك المطلوب لتحقيق التقدم في مجال البحث في الهندسة اللسانية.

2- مفهوم المعالجة الآلية للغات (Traitement automatique des langages):

تعرّف المعالجة في أبسط صورها على أنها حلول لمشكلات عن طريق الاستعانة بوسائل وتقنيات ومواد. وبتقريب هذا المفهوم من المعالجة الآلية للغات، نستنتج أن هذه الأخيرة ما هي إلا طريقة في التعامل مع البيانات والمعلومات اللغوية بالاستعانة بالآلة، من أجل تقديم برنامج أو تطبيق من شأنه حل مشكلة ذات طبيعة لسانية أو تقديم خدمة للطرفين (اللغة والآلة).

وتجدر الإشارة إلى أن الإنسان أثناء تأدية نشاطاته يقوم بمعالجة عقلية، تحدث في ذهنه انطلاقا مما يخزنه من معلومات وخبرات، وما يدركه من حواس، فيقوم بعمليات تحليل وتركيب واستنتاج ونقد وتخزين، والعمليات نفسها يهدف الحاسوب إلى الوصول إليها حتى يصير ذكيا، فقد اتجهت الجهود إلى التقريب بين الحاسوب والإنسان عن طريق تسهيل حركة التواصل بينهما، وهي الحركة التي كانت بطيئة وغير مباشرة لا يقوم بها إلا المبرمجون (Programmateurs)، لأن لغة الحاسوب لغة رقمية ثنائية (0 و 1) فاتخذوا إلى ذلك سبيل المعالجة الآلية للغة الطبيعية، التي صارت في يومنا هذا أُسّ البرامج والتطبيقات.

إن المعالجة الآلية للغات تعتمد بشكل أساسي على الظواهر والقواعد اللغوية، فهي تتطلب دراسة لسانية معمقة، وهذا ما يحتم عليها الاستفادة من النظريات اللسانية، التي تمدها بوصف علمي دقيق المختلف الظواهر والقواعد اللسانية، وهذا الوصف الذي يتم توصيفه، وتهيئته للهندسة اللغوية، ومن ثمة الاستفادة منه مباشرة في عملية المعالجة.

تبنى عمليات المعالجة الآلية للغات الطبيعية أساسا على عمليات البرمجة الآلية (programmation)، التي تعني في عرف المتخصصين في الحاسوبيات "فن استغلال قوة الحاسوب

الصورنة: هي العملية المعرفية التي يتحول بمقتضاها الأنموذج إلى نسق مجرد ممثل برموز أو برسوم بيانية.

لأجل ستر حدة بلادته وآثار نقائصه، بل هي فهم وتفسير العلاقة الجامعية بين اللغة الطبيعية واللغة الاصطناعية (لغة الآلة) عن طريق وسيط بينهما يسمى لغة البرمجة".

2-1- تعريف لغة البرمجة (langage de programmation):

تحدثنا سابقا عن لغة الحاسوب الرقمية ثنائية الرموز (0-1)، وهي اللغة الوحيدة التي يفهمها الحاسوب مباشرة من دون وسيط، إذ تحوَّل جميع اللغات والرموز إلى لغة (0 و1)، فهي لغة سهلة من حيث الرموز ولكن صعبة من حيث التعامل.

وأما لغة البرمجة الوسيطة، فهي اللغة التي يتم بها إدخال البيانات والمعلومات إلى الحاسوب، إنها «بكل بساطة اللغة الصورية المترجمة لواقع اللغة الطبيعية، يمكن للآلة فهمها وقراءتها والتعامل معها، وتنبني أساسا على وحدة الخوارزم وحدة أساسية لكل تعامل حاسوبي مع معطيات اللغة الطبيعية، بل الوحدة الجوهرية لكل تعامل مع الحاسوب، فهو أساس كل برامج المعلوميات، والخوارزم هو لغة افتراضية، قابلة للتشفير، وضعت لكل مشاكل التطبيقات المرتبطة بلغة البرمجة دون قيود».

من أشهر لغات البرمجة نذكر: الفورتران (Fortran)، الباسكال (Paskal) و(C++) ودلفي (Delphi)، والبازيك (Basic).

تعتمد لغات البرمجة على الخوارزم وهو عبارة عن أمر يتضمن معلومات كتبت برموز لغوية ورياضية، فعملية البرمجة هي في أبسط تعريفاتها سلسلة من الخوارزميات التي تحمل أوامر وبيانات.

مهم: تندرج لغة الحاسوب الثنائية (0، 1) ولغة البرمجة ضمن اللغات الاصطناعية.

2-2- مستويات المعالجة الآلية للغة:

تعنى المعالجة الآلية للغة باستخدام أجهزة الحاسوب في معالجة اللغة المكتوبة والمنطوقة من أجل أغراض عملية مفيدة مثل الترجمة الآلية، والمعاجم الالكترونية، وبنوك المعلومات، والتدقيق الإملائي... ولأن المعالجة الطبيعية للغة تتم على مستويين هما: **الفهم** الذي يصل إليه الإنسان بعد تحليله للرسائل اللسانية في دماغه، و**الإنتاج** الذي يمارسه بتركيب عدد من الوحدات اللسانية المعبرة عن مقصوده، فإن المعالجة الآلية أيضا تتم على مستويين يُعبَّر عنهما بمصطلحي: **التوليد والتحليل**.

2-3- أهداف المعالجة الآلية للغة:

- ✓ تواصل أفضل مع الحاسوب: ويمكن التعامل المباشر معه بمختلف الأشكال اللغوية المنطوقة والمكتوبة.

- ✓ تواصل أفضل بين البشر: عن طريق توفير برامج تحل بعض المشكلات التواصلية كعائق اللغات الأجنبية، إذ توفر اليوم الترجمة الآلية خدمات مهمة للأفراد.
- ✓ التعامل اليسير مع المعلومات، وسهولة الحصول عليها، وتنظيمها، وفهرستها، وبخاصة في ظل خدمات الشابكة.

2-4- كيف تتم المعالجة الآلية للغات؟

رغم الاختلافات الكثيرة التي طالت تعريف المعالجة الآلية للغة إلا أنها تتفق جميعا في كونها تعني دراسة مشكلات التوليد والتحليل الآلي للغات الطبيعية، أي تحويل عينات ونماذج لغوية إلى تمثيل شكلي صوري يسهل على البرامج الآلية تطويعه والتعامل معه، ولكن بشكل عام يمكن تقسيم إجراءات المعالجة الآلية، إلى ثلاث خطوات رئيسة هي:

- ✓ وضع نمذجة منطقية للغة الطبيعية: أي بناء نماذج صورية تمثل مختلف القواعد والمعطيات اللغوية، ويتدخل اللساني في هذه الخطوة بقوة لتوفير وصف وتوصيف مناسب للغة، إضافة إلى تدخل الهندسة اللغوية التي تسهم في تكييف هذه النماذج وتهيئتها وتصميم مخططات العمل التي سيبني عليها البرنامج أو التطبيق المراد وضعه.
- ✓ صياغة تلك النمذجة على شكل خوارزميات: إذ يتم تجسيد تلك النماذج وكتابتها بلغة من لغات الحاسوب على شكل خوارزميات.
- ✓ إدخال هذه الخوارزميات إلى الحاسوب: وهذا ما يسمى البرمجة.

مراجع المحاضرة:

- نبيل علي: الثقافة العربية وعصر المعلومات، سلسلة عالم المعرفة ،عدد 184 ، الكويت، أفريل 1994.
- مازن الوعر: قضايا أساسية في علم اللسانيات الحديثة، دار طلاس، سوريا، 1988.
- عبد الواحد الدكيكي: منظور أوليات المعجم تركيب في المستوى التركيبي، اللسانيات وإعادة البناء، وقائع الندوة العلمية الثالثة للسانيات، جامعة منوبة، تونس، 2014.