

الأعمال الموجهة: تطبيق رقم 14. اللسانيات الحاسوبية. مفهوم اللسانيات الحاسوبية:

هي "دراسة علمية للغة الطبيعية من منظور حاسوبي، وهذه الدراسة لا يمكن أن تتم إلا ببناء برامج حاسوبية لأنظمة اللغات البشرية من خلال تقييس ومحاكاة نظام عمل الدماغ البشري لنظم عمل الحاسب الآلي" أو هي "الدراسة العلمية للنظام اللغوي في سائر مستوياته بمنظار حاسوبي، ويتجلى هدفها في تطبيق النماذج الحاسوبية على الملكة اللغوية".
وفي هذا الإطار، كانت اللسانيات الحاسوبية ذلك العلم الذي يحاول ربط علاقة بين علمي اللسانيات والمعلوماتية، قصد معالجة اللغات الطبيعية معالجة آلية، فجعلت من موضوعها الرئيس تعليم وتعلم اللغات كمجال أساس للدراسة والتجريب.

وقد هدفت اللسانيات الحاسوبية إلى تفسير كيفية اشتغال الذهن البشري في تعامله مع اللغة، معرفة واكتساباً واستعمالاً، بمعنى أنه أصبح في إمكان الحاسوب محاكاة نمط اشتغال العقل الإنساني وتقليده من خلال لغة صورية خوارزمية أشبه ما تكون باللغة الصناعية، كما أصبح أيضاً مجالاً تطبيقياً لاختبار الفرضيات حول الطريقة التي يشتغل بموجبها العقل الإنساني .
فبعد أن كان الحاسوب آلة ذات قدرة عظيمة في التعامل، وبسرعة فائقة، مع أعقد العمليات الحسابية وأطولها، أصبح في تطبيقات تكنولوجيا المعلومات المتقدمة آلة ذكية قادرة على تداول مختلف المعلومات وتحليلها وتداولها.

قدم الحاسوب خدمات جليلة للإنسان في مجالات عديدة، ويُعدُّ المجال اللغوي من أبرز تلك المجالات، لذلك كان الحاسوب الركيزة الأساسية في هذا الفرع اللساني التطبيقي الحديث المتصل بالذكاء الاصطناعي (Artificial Intelligence)، بوصف هذا الأخير يُركز على قواعد المعارف الأعمق والأشمل من قواعد البيانات، واللغة العربية استفادت منه كثيراً في إنجاز تطبيقات لغوية حققت بها قفزة نوعية؛ مما يفسر أنه لا حل لمعضلة اللغة دون اللجوء إلى أساليب الذكاء الاصطناعي وهندسة المعرفة، سعياً إلى عالميتها ووَحدتها.

توصلت بحوث الذكاء الاصطناعي إلى أن الوظيفة الأساسية للعقل البشري التي تميزه عن العقل الحيواني، هي مقدرته على إنتاج الأنظمة الرمزية واستعمالها، وعلى رأسها النظام الرمزي اللغوي المستعمل في: التواصل، وتمثيل المعلومة، وتخزين المعرفة، ونقلها، فقامت برامج الحاسوب على هذا الأساس.

تختلف اللسانيات الحاسوبية عن نظيرتها العامة في نقاط عديدة، منها:

أولاً: اللسانيات الحاسوبية فرع تطبيقي صوري اهتم بالتقنيات المعلوماتية والاتصالية،

ثانياً: تقوم على التخطيط والتنظيم والبرمجة،

ثالثاً: إن المهتم باللسانيات الحاسوبية يفتح على العلوم الحديثة والتطورات التكنولوجية والعلمية.

النشأة والتطور:

ترجع البداية الأولى للسانيات الحاسوبية إلى فترة ظهور الحاسوب عام 1948م؛ إذ شكّل أداة مُسخرة لكل المعارف والمعالجات، لتكون اللغة العربية من اللغات المُوجهة إلى المعالجة الآلية، وفي هذه الفترة، تم تحقيق ترجمة آلية باعتماد الحاسوب من لغة مصدر إلى أخرى هدف، وقد مثّلت اللغة الإنجليزية المحطة الأولى للمعالجات الحاسوبية، لكن هذه الترجمة لم تحقق الأهداف المتوخاة منها؛ نظراً لغياب العتاد اللساني القادر على استيعاب خصائص النقل من اللغة المصدر إلى اللغة الهدف.

ومما تجدر الإشارة إليه: أن هذه البادرة الأولى قام بها الغرب، أما عن نصيب العرب من العملية، فلم يتأت إلا عام 1971م، حين تم اعتماد الحاسوب قصد الدراسة الإحصائية للألفاظ، وهي عملية لا يمكن اعتبارها معالجة آلية، وإنما إحصاء لغوي باعتماد الحاسب.

3- المنجزات ومجالات الاشتغال: أهمها ما يأتي:

أولاً: التوثيق:

ويُعرف بأنه شكل من أشكال العمل البليوغرافي الذي يستخدم وسائل متعددة؛ كالكشافات والمستخلصات والمقالات البليوغرافية، إضافة إلى الوسائل والطرق التقليدية الأخرى؛ كالتصنيف والفهرسة؛ وذلك لجعل المعلومات سهلة المنال، والوصول إليها سهلاً أيضاً.

ويُتوخى من التوثيق معالجة الوثائق والمعلومات الواردة بها بشكل يسهل على الباحث ولوجها، واستدعاءها عن طريق التجميع والاستخلاص والنشر، وذلك كالتصنيف الأتوماتيكي للملفات، والمؤلفات تبعاً لفهارسها أو مؤلفيها أو مواضيعها، وعليه فإن البرامج اللسانية الحاسوبية ترمي إلى توثيق المعلومات التي ترد على ذهن البشري.

ثانياً: صناعة المعجم الإلكتروني وتطوير العمل المصطلحي:

يتطلب بناء المعجم الإلكتروني صياغة المصطلحات وتعميم استعمالها ونشرها وتداولها، إلا أن هذا الصنف من المعاجم يلم إماماً كبيراً بجميع مستويات اللغة؛ أي إنه يعتمد على أدوات مُشرفة، وقاعدة معطيات مُرمزة تخص مستويات اللغة؛ نحواً ودلالة، بحيث تخضع المادة المعجمية فيه لبناء يلم إماماً دقيقاً بفروع اللغة.

ويشترط في المعجم الإلكتروني "أن يكون شاملاً وعماماً؛ لأن البرنامج اللساني المُعدّ للمعالجة الآلية لا ينبغي أن يفشل في العثور على أية معلومة كيفما كان نوعها وكيفما اتفق؛ لأن أي خطأ في المعلومات المدخلة من شأنه أن يتسرب إلى باقي مفردات الجملة، أو النص برمته، ومن ثمة يعرقل عملية اشتغال البرنامج، كما يشترط في المعلومات أن تكون في المعجم الإلكتروني واضحة موائمة للمداخل المعجمية المراد معالجتها آلياً، ومن هنا يظهر توظيف واستغلال تقنيات المعلومات في الأعمال التطبيقية لصناعة معاجم مصطلحية باعتماد برامج معينة، وللمغاربة جهود كبيرة فيها.

كما تسعى اللسانيات الحاسوبية إلى إنشاء بنوك للمصطلحات، انطلاقاً من تخزين المصطلحات مُرفقة بمعلومات عن كل مصطلح مفرد، ستكون مساعدة للمترجمين والمحريين والمتعلمين، مع دعم الترجمة الآلية وبناء المعاجم المختصة، وتُحوّل البنوك المصطلحية تخزين معطيات دقيقة عن كل مصطلح في ضوء نصوص موثقة، مع ذكر مقابلاته بلغات متعددة، وتوضيح مجالات استخدامه، وأساليب توظيفه، وكذا الإشارة إلى مرجعه، سواء كان معجماً أو معهداً علمياً، أو نصّاً، أو وثيقة.

وبهذا، فإن المعجم الإلكتروني له علاقة بالبنوك المصطلحية، وذلك من خلال تنظيمه لها، وهو ما لا تيسره ذاكرة الإنسان المحدودة.

المعاجم الإلكترونية:

تعكس المعاجم الإلكترونية مدى ذكاء الإنسان وقدرته على الإبداع والاختراع من جهة، وعلى الدور الهام الذي يؤديه الحاسوب في حياتنا. بحيث تظهر قيمة الحاسوب في تخزين المادة وترتيبها

طبقا للنظام المطلوب وتعديل بعض المعطيات وحذفها، وتجديد المعجمات بسهولة، إضافة إلى النقل المباشر إلى المطبعة والحصول على أجزاء محدودة من داخل المادة المخزونة¹.

يقوم الحاسوب اعتمادا على قواعد الاشتقاق بتوليد جميع القياسات من مشتقات ومصادر، وتمكنه قواعد التعريف والإسناد من تحديد صيغة الفعل بمختلف صورها (الماضي، المضارع، الأمر) مسندا إلى جميع الضمائر، ومن تحديد نوع الفعل من حيث الصحة والاعتلال والهمز والتضعيف وما يتفرع عنها من تحديد صيغة التثنية والجمع السالم بنوعيه، كما أن التوصيف المعجمي للكلمة العربية التي ستكون المدخل للمعجم سواء أكانت فعلا أم اسما أم صفة أم ظرفا أم حرفا يتم تحديده من خلال عدّة حقول: الحقل الصرفي، الحقل النحوي، الحقل الدلالي، الحقل الصوتي والحقل الإحصائي².

ما هو المعجم الإلكتروني؟:

إنه نسخة حاسوبية معدلة من النسخة الورقية. يحمل كما هائلا من المداخل المقترنة بتعاريفها التي تختلف بدورها من معجم لآخر حسب المجال الذي يرتبط به.

اقتصرت استعمال المعجم في بداياته الأولى عند الظهور بوصفه مادة لغوية لتحليل الآلي للغات الطبيعية على المستوى الصرفي والنحوي والدلالي، فكانت المعاجم بمثابة قواعد بيانات تحتوي على معلومات مشفرة لا يفهمها إلا البرنامج الذي يستغلها. وقد تطورت وتنوعت هذه المعاجم من حيث المحتوى وكذلك من حيث الهيكلية لمواكبة تطور تطبيقات المعالجة الآلية للغات الطبيعية التي تستغلها مثل: تقطيع الجمل، تحليل النصوص واسترجاعها، البحث عن المعلومات، التدقيق الإملائي، التلخيص الآلي للوثائق والترجمة الآلية.

وسعى لتقريب العملية من الفهم وجب تطوير قواعد البيانات لتحويل المعلومات المشفرة التي تحتويها إلى نصوص، وإضافة تفسيرات لها ليتمكن الإنسان من قراءتها وفهمها، وكذلك تصميم برامج بينية متطورة للبحث عن المعلومات المتوافرة بالقاعدة وعرضها على الشاشة بطريقة مستساغة واستعمال وسائط متعددة (نص، صورة، وصوت فيديو وصور متحركة).

من القواميس الإلكترونية الغربية للغة الإنكليزية والفرنسية: ذخيرة اللغة الفرنسية المحوسبة (TLFi)، وقاموس الأكاديمية الفرنسية (2 طبعات)، و"البيبيورم لاروس" الفرنسي ومعجم "أوكسفورد" الإنجليزي OED (الطبعة الثانية) والمعجم والمكنز الإلكتروني "كولينز".

¹ محمود فهمي حجازي، المعجمات العربية وموقعها من المعجمات العالمية، بحوث ندوة خاصة بمناسبة الانتهاء من تحقيق وطباعة معجم تاج العروس، المجلس الوطني للثقافة والفنون والآداب، الكويت، 9-10 فبراير 2002 ص: 183.

² محمّد علي الزركان، اللسانيات وبرمجة اللغة العربية في الحاسوب، ص: 56.

أما على المستوى العربي فلا زلنا ننتظر بزوغ فجر جديد، يمجّد تراثنا العربي ويلحقه بعالم الحوسبة والرقمنة، ونحن بحاجة ماسة لإكمال المعاجم التراثية قلباً وقالباً من خلال رصد دقيق لألفاظ الحضارة للمصطلحات العلمية في كل فروع المعرفة الإنسانية والعلوم والتقنيات حتى تكتمل صورة اللغة العربية عبر القرون.

تصنيف المعاجم الإلكترونية:

يمكن التمييز بين نوعين من المعاجم:

1. المعاجم الموجهة إلى الآلة (Machine Readable Dictionary)

2. المعاجم الموجهة للإنسان (Human Oriented Electronic Dictionaries)،

علماً بأن الجهود قد اتجهت السنوات الأخيرة إلى دمج هذين الصنفين، وذلك باقتراح منهجية بناء معاجم مقيسة تراهن على تلبية حاجة الإنسان والآلة في نفس الوقت.

المعجم الإلكتروني الموجه للإنسان: إن التصنيف الحديث للمعاجم الإلكترونية الموجه للإنسان يأخذ بعين الاعتبار الجوانب الأساسية التالية:

- ترتيب المداخل: (حسب الجذور، حسب الجذوع،...)،

- وظيفة المعجم: (للترجمة، للتعليم،...)،

- نوعية المحتوى: (نصي/ متعدد الوسائط، لغوي/ متخصص، أحادي اللغة/ متعدد اللغات،...)،

- المحمل (سطح المكتب، صفحات الواب، أقراص مدمجة،..).

انطلاقاً من هذه الجوانب يمكن تصنيف المعاجم الإلكترونية كما يلي:

- معاجم لغوية: متكونة من عينة من المفردات. يحتوي كل مدخل على المعلومات اللغوية الأساسية التالية: تعريف الكلمة، خصائصها الصرفية والنحوية، طريقة الكتابة (الإملاء)، المعاني المختلفة مع أمثلة وشواهد لمختلف الاستعمالات.

- معاجم متخصصة تحتوي على المفردات المستعملة لعلم ما أو فن: قاموس الرياضيات، قاموس الطب، قاموس الاقتصاد، قاموس الحاسوب، قاموس أسماء العلم.

- معاجم متعددة اللغات تعطي ترجمة الكلمات إلى لغة أو لغات أجنبية.

- معاجم بصرية تحتوي على مجموعة من الصور وأشرطة الفيديو مبنية حسب المواضيع التي تعالجها (بنايات، حيوانات، ألعاب، ملابس، وسائل نقل،...).

مزايا المعجم الإلكتروني:

- تنوع طرق البحث عن المعلومة:
- طاقة التخزين الواسعة وتطور تقنيات قواعد المعطيات
- إمكانية التوليد الآلي لبعض الكلمات القياسات بدون الحاجة إلى تمثيلها بالمعجم وذلك بالاعتماد على قواعد الاشتقاق.
- سهولة تعديل المعجم الإلكتروني بإضافة مداخل جديدة أو بتحيين مداخل موجودة.
- الاعتماد على الوسائل الحاسوبية الحديثة المتعددة الوسائط (Multimedia) من نصوص، وأصوات، وصور ثابتة ومتحركة، وأفلام الفيديو لعرض المعارف. هذه الخاصية لها تأثير إيجابي على استساغة وفهم المعارف المعروضة.
- يمكن القول: إن تطوير تطبيقات حاسوبية للغة العربية تلبية حاجة المستخدم العربي وتجعل اللغة العربية مواكبة لتطور المجتمع رهين وجود معاجم إلكترونية عربية مناسبة وذات جودة عالية وتخضع إلى مقاييس عالمية على مستوى المحتوى والهيكلية.

مراحل بناء المعجم:

1. الاتفاق على هيكل المعجم (ترتيب المداخل ومحتوى كل مدخل) انطلاقاً مما يسمى بكراس الشروط الذي يحدد الهدف من بناء المعجم، المستخدم المستهدف (مبتدئ، خبير،...) طريقة استغلال المعجم (على الشاشة، على سطح المكتب أو على قرص مضغوط،...).
 2. تحديد المصادر اللغوية (مدونات، معاجم ورقية،...) التي سيقع اعتمادها كمادة أولية.
 3. انتقاء وتجميع المادة المعجمية من المصادر المذكورة في جذاذات ورقية.
 4. تصميم وبناء قاعدة البيانات التي ستحتوي المعجم.
 5. بناء برمجية لإدخال المعلومات المعجمية يدوياً من الجذاذات الورقية.
 6. إدخال المعلومات يدوياً من الجذاذات الورقية.
 7. تجريب المعجم للتأكد من تحقيقه كافة الوظائف المطلوبة منه.
 8. تصويب الأخطاء اللغوية والبرمجية.
- مراحل بناء قاعدة البيانات المعجمية:

1. اختيار المعجم (أو المعاجم) الذي سيعتمد عليه في صيغته المرقمنة.

2. تصميم قاعدة البيانات المعجمية التي ستحتوي مداخل المعجم.
 3. تصميم برنامج تعبئة قاعدة البيانات المعجمية انطلاقاً من النسخة أو النسخ المرقمنة.
 4. تجريب القاعدة.
 5. تصحيح الأخطاء اللغوية والحاسوبية.
 6. إثرائها يدوياً إذا اقتضت الحاجة.
- 3.5 انطلاقاً من مدونة نصية
- تعد هذه الطريقة من أصعب الطرق وأقلها دقة. تكمن مزاياها في أنها غير مكلفة على المستوى البشري والمادي وكذلك تمكن من الحصول على معجم مواكب.
- بناء معجم بهذه الطريقة يتم باتباع المراحل التالية:
1. تحديد المدونة أو بناؤها.
 2. إنجاز برمجية لاستخراج المداخل والعلاقات بينها والمعلومات المعجمية الخاصة بكل مدخل بصفة آلية من المدونة.
 3. تدقيق المادة المعجمية المتحصل عليها من طرف المعجميين.
 4. إدخال المادة المعجمية المدققة في قاعدة البيانات التي ستحتوي المعجم.