الأستاذ: أنور طراد

الدّرس الثامن: كيفية صناعة المعاجم الإلكترونية:

1. المعجم الإلكتروني: تاريخه وأهم مكوناته

المعجم الإلكتروني **هو نسخة حاسوبية معدّلة من النسخة الورقية** . بحيث يتكون من عدد كبير من المداخل (الوحدات المعجمية..) يحتوي كل واحد منها على المعلومات التي يمكن تجميعها حوله. وتختلف هذه المعلومات من معجم إلى آخر حسب المجال المستهدف، والأهداف التي بني من أجلها، وأصناف المستخدمين المستهدفين.

وقد بدأ الاهتمام بالمعاجم الإلكترونية منذ منتصف القرن الماضي ؛ حيث اقتصر في البداية على استعمال هذه المعاجم كموارد لغوية للتحليل الآلي للغات الطبيعية على المستوى الصّرفي ، والنحوي، والدلالي، **فكانت المعاجم بمثابة قواعد بيانات تحتوي على معلومات مشفرة لا يفهمها إلا البرنامج الذي يستغلها** . وقد تطوّرت وتنوعت هذه المعاجم من حيث المحتوى ، وكذلك من حيث الهيكلة لمواكبة تطور تطبيقات المعالجة الآلية للغات الطبيعية التي تستغلها مثل: تقطيع الجمل، وتحليل النصوص واسترجاعها، والبحث عن المعلومات، والتدقيق الإملائي، والتلخيص الآلي للوثائق ، والترجمة الآلية. فإذا أخذنا على سبيل المثال المعجم المصمم للتدقيق الإملائي نجده يختلف تماما عن المعجم المصمم للتحليل النحوي. فالأول يقتصر على قائمة كلمات اللغة، والثاني يستوجب تمثيل المعلومات الصرفية والنحوية (مثل: قسم الكلم، التّعدية واللّزوم، أدوات التّعدية...) بالنسبة لكل مدخل.

وتجدر الإشارة هنا إلى أن هذه المعاجم التي صُممت للآلة يصعب استغلالها من طرف الإنسان. ولبلوغ هذه الغاية 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).

التطور الذي حصل على مستوى قواعد البيانات والبرمجيات المصاحبة، بالإضافة إلى تطور الأجهزة الحاسوبية، قد ساعدا في بداية الثمانينات على ظهور أول المعاجم الإلكترونية المتاحة إلى الجمهور العريض على سطح المكتب أو على الإنترنت (سواء عن طريق الاشتراك أو مفتوحة المصدر) أو على الأقراص المدمجة (CD – ROM) والأقراص الأخرى.

ومن بين القواميس الإلكترونية الغربية: ذخيرة اللغة الفرنسية المحوسبة (TLFi)، وقاموس الأكاديمية الفرنسية، و"البيبليورم لاروس" الفرنسي ومعجم "أوكسفورد" الإنجليزي OED (الطبعة الثانية) والمعجم والمكنز الإلكتروني "كولينز".

أما بالنسبة إلى اللغة العربية فثمة تقصير واضح في إعداد معاجم حاسوبية تلبي حاجيات المستخدم العربي بالرغم من إتاحة أمّات المعاجم العربية على أقراص المدمجة أو على الإنترنت. فهذه الخطوة ليست كافية؛ لأن هذه المعاجم تفتقر لأبسط مقومات المعاجم الحاسوبية بالمعنى الحديث للكلمة. إنها مجرد "نسخ مرقمنة" لتلك المعاجم الورقية (في صيغة " Doc" أو " HTML") لا يمكن الاستفادة منها بالشكل المطلوب (Ait Taleb, 2005). بالإضافة إلى ذلك، فالأدوات المصاحبة للبحث عن المعلومة بسيطة كما يتضح في المشاريع المعروفة مثل "عجيب" (لنظام صخر) "وكلمات" و"لسان العرب" / "القاموس المحيط". هذه السلبيات تعود أساسا إلى ضعف الهيكلة الحاسوبية التي بنيت عليها مداخل هذه المعاجم.

وينبغي أن نفرق بين المعجم المرقمن والمعجم الإلكتروني، فالأول عبارة عن نسخة مصورة للنسخة الورقية لا تزيد عنها ولا يمكن تعديلها أو إضافة معلومات لها، أما المعجم الإلكتروني فهو عبارة عن قاعدة بيانات تسمح بالبحث فيها عبر محرك البحث، وقد يشتمل على عدة نسخ مرقمنة مدمجة، تختلف موضوعاتها، كما أنه يحتوي على معلومات تفوق المعجم الورقي والمرقمن. فعندما تبحث عن كلمة ما يعطيك معانيها ومرادفاتها، بالإضافة إلى سياقاتها، ومعلوماتها الصرفية والصوتية..وهذا لا نجده كثير في المعاجم الورقية.

التصنيف الحديث للمعاجم الإلكترونية الموجه للإنسان يأخذ بعين الاعتبار الجوانب الأساسية الآتية:

- **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،..).**

انطلاقا 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 الصّور وأشرطة الفيديو مبوّبة حسب المواضيع التي تعالجها (بنايات، وحيوانات، وألعاب، ملابس، وسائل نقل،....).

2 – مزايا المعجم الإلكتروني:

للمعجم الإلكتروني مزايا عديدة متأتية أساسا من التّطور التكنولوجي الذي حصل السنوات الأخيرة على مستوى سرعة معالجة البيانات وسعة تخزين المعلومات ، وكذلك على مستوى البرمجة ومعالجة قواعد البيانات. هذه المزايا يمكن حصرها في فيما يلي:

- **تنوع طرق البحث عن المعلومة** : يمكن لمستخدم المعجم الإلكتروني أن يصل إلى المعلومة عبر الجذر أو الجذع (البحث البسيط) أو عبر المعنى (البحث المتقدم) . مثلا يمكن البحث عن كلمة "هضبة" باستعمال المعنى الآتي "أرض مرتفعة". كما يمكن البحث عبر الإبحار داخل المعجم باستعمال الروابط النّصية (Hyper-text links).

- **طاقة التخزين الواسعة** وتطور تقنيات قواعد المعطيات تتيح بناء معاجم كبيرة الحجم تجمع بين القديم والمعاصر ومتعددة اللّغات والوسائط. هذه المعاجم تمتاز بالدّقة والشمولية من حيث إنها توفر لكل كلمة معانيها الأساسية والفرعية، وتعطي لذلك أمثلة وشواهد متنوعة.

- **إمكانية التوليد الآلي لبعض الكلمات** القياسات بدون الحاجة إلى تمثيلها بالمعجم ، وذلك بالاعتماد على قواعد الاشتقاق. هذه الميزة يصعب (أو يستحيل) توفيرها بالنسبة إلى المعجم الورقي ، لأن إيراد المشتقات القياسية لجميع الأفعال الممثلة سيضاعف حجم المعجم ويجعله غير قابل للاستعمال.

- **احتواء المعجم الإلكتروني على عدّة تطبيقات لغوية مهمّة** يمكن للمستخدم أن يستفيد منها ، مثل: تصريف الأفعال والأسماء، البحث عن المترادفات، المعالجة على المستوى الصوتي لتحويل المكتوب إلى منطوق، التدقيق الإملائي لتصويب الكلمات المدخلة،...هذه الخدمات غير متوفرة في المعجم الورقي.

- **سهولة تعديل المعجم الإلكتروني** بإضافة مداخل جديدة ، أو بتحيين مداخل موجودة. هذه الميزة يجب التعامل معها بكل حذر للحفاظ على مصداقية المعجم وجودته. تعديل محتوى المعجم يجب أن يتم من طرف معجميين وفق تراتيب مقننة تضمن تماسك المحتوى.

- **الاعتماد على الوسائلِ الحاسوبيّةِ الحديثة المتعددةِ الوسائط** (Multimedia) من نصوص، وأصوات، وصور ثابتة ومتحركة، وأفلام الفيديو لعرض المعارف. هذه الخاصية لها تأثير إيجابي على استساغة وفهم المعارف المعروضة.

- ولا تخفى مزايا المعجم الإلكتروني على سائر مجالات اللّسانيات الحاسوبية الأخرى، فتطور الترجمة الآلية ووصولها للمستويات عالية من الدقة مرهون بتوفر المعاجم الإلكترونية بأنواعها، بالإضافة إلى فائدته على تعليم وتعلم اللغات عبر البرامج سواء بالنسبة للتلاميذ أو غيرهم.

وتجدر الإشارة هنا إلى أن للمعجم الورقي والمعجم الإلكتروني مزايا مشتركة لكن بدرجات متفاوتة نذكر بالخصوص إمكانية الاستعمال في كل مكان. مع تطور الحواسيب المحمولة على مستوى الحجم والاستقلالية عن التزود بالطاقة الكهربائية أصبح استغلال المعجم الإلكتروني متاح في كل مكان لكن بدرجة أقل من المعجم الورقي.

وبالرغم من ما ذكر من مزايا ما زال المعجم الورقي ينعم باستحسان القرّاء والمستعملين لكن المستقبل سيكون حتما لفائدة المعجم الإلكتروني.

3 - طرق بناء المعجم الإلكتروني

بغض النظر عن الطريقة المعتمدة، فإن بناء معجم إلكتروني ليس بالعمل الهين، فهو يتطلب مجهودا جبارا يقوم به فريق يتكون من معجميين ومعلوماتيين. يهتم المعجميون بتجميع المادة اللغوية من مدونات ومعاجم ورقية وانتقاء المداخل، وتحديد المعلومات الملحقة بكل مدخل.

ويهتم المعلوماتيون أو الحاسوبيون بالتصميم والهيكلة لضبط البنية الفوقية (**Macro-structure**) التي تُعنى بترتيب مداخل المعجم والبنية المصغرة (**Micro-structure**) التي تهتم بترتيب مكونات المدخل الواحد . كذلك يهتم الحاسوبيون بتصميم البرامج الضرورية لإدخال المعلومات المعجمية وتحيينها والبحث عنها وعرضها ، ولتوفير مختلف الخدمات الملحقة بالمعجم من تدقيق إملائي وتصريف الأسماء والأفعال وغيرها. ولبلوغ درجة عالية من الدقة والجودة يجب أن يعمل أعضاء هذا الفريق بطريقة متعاونة ومتكاملة، وكذلك متوازية ربحا للوقت.

وفي ما يلي أهم طرق بناء المعاجم الإلكترونية.

أ – الطريقة اليدوية

تعد الطريقة اليدوية من أسهل طرق بناء المعاجم لكنها مكلفة من حيث الجهد والوقت والمال. ويتم بناء المعجم باتباع المراحل الآتية:

1. الاتفاق على هيكل المعجم (ترتيب المداخل ومحتوى كل مدخل) انطلاقا مما يسمى بكراس الشروط الذي يحدد الهدف من بناء المعجم، المستخدم المستهدف (مبتدئ، خبير،...) طريقة استغلال المعجم (على الشابكة، 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 يدويا من الجذاذات الورقية.

6. إدخال المعلومات يدويا من الجذاذات الورقية.
7. تجريب المعجم للتأكد من تحقيقه كافة الوظائف المطلوبة منه.
8. تصويب الأخطاء اللغوية والبرمجية.

المرحلة الأولى يقوم بها المعجميون بتنسيق مع الحاسوبيين، المرحلة الثانية والثالثة يقوم بها اللغويون مع إمكانية الاستعانة بالحاسوبيين لمدهم ببعض الأدوات التي تساعدهم على تحليل وجرد المدونات أو القيام ببعض العمليات الإحصائية. المرحلة الرابعة والخامسة يقوم بها الحاسوبيون. المرحلة السادسة يقوم بها كتبة. المرحلة السابعة يقوم بها لغويون والمرحلة الأخيرة يقوم بها المعجميون والحاسوبيون.

ب - انطلاقا من معجم ورقي مرقمن:

هذه الطريقة ممتازة جدا لأنها تقتصر المسافات، وتقتصد المال والجهد. الفكرة المركزية فيها هي تحويل معجم ورقي مرقمن إلى معجم إلكتروني يخضع إلى مواصفات دقيقة وحديثة. هذا التحويل يتم عن طريق برنامج حاسوبي يقع تصميمه للغرض. النتيجة المتحصل عليها قاعدة بيانات معجمية. ومن مزايا هذه الطريقة إمكانية الاعتماد على أكثر من معجم ثم يقع دمج قواعد البيانات المتحصل عليها للحصول على قاعدة ثرية على مستوى عدد المداخل، وكذلك على مستوى محتوى المداخل. وقد قام فريق من مخبر ميراكل بجامعة صفاقس باقتراح وتجربة هذه الطريقة على معجم "الغني" (A Khemakhem 2009)، وكذلك على معجم الوسيط بالتعاون مع مدينة الملك عبد العزيز للعلوم والتقنية.

مراحل بناء قاعدة البيانات المعجمية وفق هذه الطريقة تكون باتباع مايلي:

- اختيار المعجم (أو المعاجم) الذي سيعتمد عليه في صيغته المرقمنة.
- تصميم قاعدة البيانات المعجمية التي ستحتوي مداخل المعجم.
- تصميم برنامج تعبئة قاعدة البيانات المعجمية انطلاقا من النسخة أو النسخ المرقمنة.
- تجريب القاعدة.
- تصحيح الأخطاء اللغوية والحاسوبية.
- إثرائها يدويا إذا اقتضت الحاجة.

د – الطريقة التعاونية (ويكي)

المقاربة التعاونية (الويكي) تمكن من بناء معاجم إلكترونية بتكلفة زهيدة، وذلك بتكاثف جهود المتطوعين المتخصصين في مجال المعجمية. هذه المقاربة برزت ونجحت مع ظهور الموسوعة "ويكيبيديا". فهي تسمح لكل متطوع يرى في نفسه القدرة على المساهمة البناءة بالعبور إلى المعجم وتغيير محتواه سواء بإضافة مدخل جديد أو بإثراء محتوى مدخل موجود.

ولتقليل الأخطاء يمكن لكل مستخدم أن يتعرف إلى التغييرات الحديثة التي أدخلت على المعجم والتثبت فيها مع إمكانية تصويب الخطأ.

من أهم مزايا هذه الطريقة أنها:

- تضمن المواكبة السريعة للتطورات التي تحدث على المعجم من بروز مفردات جديدة وتوَسع مجال مفردات موجودة.
- تعتبر أن بناء المعرفة ليس مجرد مسألة شخصية. بل هو مقاربة جماعية من أجل الوصول إلى توافق في الآراء.
- لا تقصي أحداً من عملية بناء المعجم. فهي تضع على قدم المساواة المبتدئ (الطالب مثلا) والخبير (المعجمي).

بالنسبة للعربية يمكن ذكر مشروع "ويكاموس" الذي انطلق في شهر مايو 2004 والذي تقدم أشواطا مهمة لكنه لا يزال يشكو من ثغرات عديدة في المحتوى لأن بناءه لم يكن ممنهجا، بل موكولا إلى إرادة المساهمين في عملية الإثراء.

المراجع:

_ المعجم العربي الإلكتروني أهميته وطرق بنائه، عبد المجيد بن حمادو، جامعة صفاقس، تونس

_فاعلية اللسانيات الحاسوبية في حوسبة المعجم العربي، مراد مسعودة وبن يحي محمد، مقال منشور بمجلة علوم اللغة العربية وأدابها، العدد1، المجلد13.