المحاضرة العاشرة: المكانز اللغوية الآلية

1-تعريفها:

لم يكن ممكنا للمعالجة الآلية للغات أن تتقدم وتفتح لنفسها مجالات للتطبيق ثم تستشرف المستقبل دون أن يكون لها قلب نابض يستوعب مليارات الكلمات ، وتكون لديه القدرة على الوصول إلى أي منها، وفهم وتحليل ما بينها من علاقات، ليصبح مع الوقت بيتا للذاكرة ونبع متدفق للمعرفة، وعمليا تجسد هذا القلب النابض فيما يعرف بالمكانز اللغوية(الذخائر)، التي تعتمد في عملها على قوة الحاسبات في المعالجة وقدرات البرمجيات المختلفة في التحليل والفهرسة والرصد والاسترجاع، وسعات وحدات التخزين في استيعاب ما لا حصر له من الكلمات والأصوات. فما هي المكانز اللغوية؟

إن هذا المصطلح يشير إلى المجموعات الكبيرة من النصوص أو الملفات الصوتية التي تمثل عينة أو تخصص أو شريحة معينة من اللغة، وهذه النصوص تكون غالبا في صيغة إلكترونية يستطيع الحاسب قراءتها والبحث فيها، ومن الممكن أن يتكون المكنز اللغوي من نصوص خام فقط دون أي معلومات، أو يحتوى على معلومات لغوية خاصة تسمى بالحواشي أو الوصف.

وهى إما مكانز نصية، تحتوى على مجموعة هائلة من النصوص المكتوبة والمطبوعة في مختلف المجالات، أو مكانز صوتية منطوقة تحتوى على تسجيلات آلاف المحادثات والحوارات والخطب ونشرات الأخبار الإذاعية والبرامج الحوارية والمسلسلات والأفلام وغيرها.

2- مجالات استخدامها:

✓ صناعة المعاجم:

تساعد المكانز اللغوية على معرفة معلومات غاية في الأهمية لبناء المعجم، وهي متابعة الكلمات الجديدة التي تدخل اللغة وتحديد وقت دخولها، ومعرفة الكلمات الموجودة بالفعل التي

اكتسبت معنى جديدا، وبمساعدة المكانز الالكترونية يستطيع خبراء صناعة المعاجم البحث في ملايين الجمل والسياقات المختلفة، واستدعاء جميع الأمثلة لكلمة معينة لمعرفة استخداماتها والألفاظ التي ترد عادة قبلها أو بعدها، من أجل تحديد التعابير الاصطلاحية والمتتاليات اللغوية، الأمر الذى يسهّل تعلم اللغة على الأجانب.

✓ **فهم قواعد النحو:**

يمكن الاستفادة من المكانز في دراسة الملامح الصرفية، والبحث عن السوابق واللواحق المعينة التي تدخل على الكلمة؛ فكلمة (علم) تتعدد معانيها بإضافة سوابق أو لواحق مختلفة لتصبح (علمية، علمتنا، علماء، تعليم، علوم)، بجانب تحديد توزيع الكلمة وموقعها في الجملة، وهل تأتى قبل الاسم أم بعد الاسم، وقبل الصفة أم بعد الصفة.

✓ **تحديد الدلالة:**

في السنوات الأخيرة، ظهر اتجاه جديد يعتمد على استخلاص معنى الكلمة من المكانز اللغوية، وتتنوع المعلومات الدلالية بين الترادف والتضاد إلى علاقات أكثر تعقيدا، ويمكن استخلاص هذه المعلومات بسهولة من المكانز اللغوية، لتوظيفها في بناء المعاجم التاريخية التي تتبع تطور دلالات الكلمات.

تحسين تعليم اللغات الأجنبية:

من خلال المكانز من الممكن تحليل مدى تكرار وشيوع الكلمات، ومعرفة تأثير السياق أو الموقف على أسلوب اللغة، وهي معلومات مفيدة للغاية في وضع مناهج تعليم اللغات. أيضا باستخدام المكانز اللغوية، يستطيع واضع المراجع والمواد إنشاء تمرينات تعتمد على أمثلة حقيقية تقدم للمتعلمين فرصة اكتشاف خصائص استخدام اللغة. وبدلا من الاعتماد على البحث في معاجم تقليدية قديمة، يستطيع المتعلمين البحث بأنفسهم في برامج المكانز اللغوية، من خلال برنامج بحث وإحصاء لغوى متخصص، ويكتشفوا بأنفسهم استخدامات اللغة وقواعدها وخصائصها.

3-أهمية المكانز اللغوية لبناء معاجم اللغة العربية:

- حصر جميع المعاني: من مميزات البحث الإلكتروني في المكانز اللغوية إمكانية عرض جميع السياقات التي يمكن أن تظهر فيها الكلمة، وبالتالي إمكانية حصر كل المعاني المختلفة للكلمة نفسها حسب السياقات المختلفة، ومن أمثلة ذلك كلمة (قلب) التي تعنى جوهر وقلب يضخ الدماء ووسط أو منتصف.
- دراسة مدى شيوع الكلمات: تسهم المكانز اللغوية في معرفة أكثر الكلمات شيوعا في اللغة العربية على المستويين المنطوق والمكتوب، وذلك من أجل تعليم هذه الكلمات للأجانب وحتى يتمكنوا من فهم المعاجم العربية، ومما لا شك فيه أن معرفة مستوى شيوع كلمة ما يدل على أهميتها في كل اللغة المكتوبة أو المنطوقة، وبالتالي ترشد المتعلم إلى معرفة مدى ضرورة تعلمها أو إغفالها من عدمه.
- دراسة التنويعات المعجمية: دراسة تكرار الكلمة كمادة أو كفئة معجمية (اسم، فعل، صفة، غير ذلك)، ومن الممكن أن يكون للكلمة الواحدة أكثر من فئة معجمية حسب السياق، مثال ذلك كلمة (عين) التي ترد كاسم وكفعل مع التشديد بمعنى (وظف)
- دراسة استخدام المترادفات: تحتوى اللغة على كلمات عديدة تعتبر مترادفات لبعضها البعض، ومن خلال المكنز اللغوي يستطيع الباحث بسهولة معرفة مترادفات الكلمة ومعدل شيوعها.
- شكل الكلمة وفقا لحالتها الإعرابية: يتغير شكل الكلمة وفقا لحالتها الإعرابية (الرفع والنصب والجر)، ويمكن للباحث من خلال المكنز اللغوي معرفة التنويعات في شكل الكلمة.
- دراسة الكلمة وفقا لاشتقاقها الصرفي: فهناك كلمات ذات أكثر من اشتقاق صرفي مثل كتب (التي يمكن اشتقاق كاتب وكتاب ومكتبة ومكتوب وغيرها)، فالمكنز اللغوي يفيد كثيرا في هذه الحالات.
- معرفة مدى السلامة والصحة اللغوية: أحيانا يكون للكلمة أكثر من جمع أو هي نفسها أكثر من شكل، ويمكن اللجوء إلى ذخيرة المكنز اللغوية الهائلة والبحث فيها لمعرفة أكثر

هذه الأشكال شيوعا واستخداما في اللغة، مثل (شهور وأشهر) و(عيون وأعين)، وهكذا.

- معرفة المتصاحبات اللغوية والتعبيرات الاصطلاحية: هناك أوصاف أو كلمات دائما تقترن بأوصاف أو كلمات أخرى دون سبب ظاهر أو منطقي، مثل تعبيرات (فتح الباب على مصراعيه) و(رأب الصدع) و(العروة الوثقى) .ودراسة المكانز اللغوية تفيد للغاية في البحث عن أمثلة هذه المتلازمات، من أجل تسجيلها في موادها بالمعجم العربي الحديث. فمثلا في مادة عروة ينبغي أن يذكر القاموس أنه دائما تليها كلمة (وثقى)، وهكذا.
- الاستفادة من المكانز اللغوية في استخلاص القواعد اللغوية: على مستوى الكلمة والجملة والخطاب، والحصول على معلومات عن تركيب واستخدام العديد من التعبيرات اللغوية.

4-آلية عمل المكانز اللغوية الآلية:

4-1- جمع المادة النصية: عن طريق جمع المادة الالكترونية المتوفرة على الأقراص الضوئية أو الشبكات، أو من خلال الإدخال اليدوي عن طريق لوحة المفاتيح، وعن طريق تحويل المادة المنطوقة إلى مادة مكتوبة إن وجدت، أو عن طريق جهاز المسح الضوئي، الذى يحول النصوص من المستندات الورقية إلى الصيغة الإلكترونية.

4-2-تجهيز المادة اللغوية وهي مرحلة ما قبل المعالجة: يقصد بذلك تقسيم الجمل وعزل الصور وتوحيد الخطوط وحجم الخط، وعزل علامات الترقيم، وتمييز اللبس والاختصارات والقوائم بأنواعها؛ بمعنى تصفية المادة اللغوية .

4-3- إدخال معلومات الوصف والترميز والتسمية المناسبة : يقصد بها تحديد المعلومات اللغوية، وبيانات المؤلف، وتاريخ الإنشاء، والعنوان واللغة والمجال، وإدخال المعلومات اللغوية؛ مثل فئات أو أقسام الكلام والمشتقات والمعلومات النحوية، وبناء الجملة، والمعلومات الدلالية، ومعلومات عن الأسلوب، والصوت بصيغة يفهمها برنامج البحث في المكنز، وسائر أدوات

معالجة اللغة، وتتضمن معلومات الوصف تمييز حدود الجمل والعبارات والفقرات، وفك لبس حدود الجمل من خلال النقطة والمسافة وتمييز جذور الكلمات ومشتقاتها.

مراجع المحاضرة:

-محمد علي مناصر و عادل بن مرتاح و عمار عرباوي:كتب علوم القرآن الكريم كمن نسخة خطية إلى مكنز رقمي ثري و مهيكل للاستخدام الفعال في تقنية المعلومات، مركز الأبحاث الواعدة في أبحاث تقنية المعلومات لخدمة *القرآن الكريم* وعلومه (نور https://www.ketabonline.com/ar/books/105016/read?page=2)

-أميرة بهجت: المعالجة الألية (جانفي2021).https://amiraabahget.wordpress.com